Hamiltonian Reduction by Stages

Jerrold E. Marsden

Control and Dynamical Systems 107–81

California Institute of Technology

Pasadena, California 91125, USA

Gerard Misiołek

Department of Mathematics

University of Notre Dame

Notre Dame, IN 46556, USA

Juan-Pablo Ortega

CNRS and Département de Mathématiques

Université de Franche-Comté

F-25030 Besançon cedex, France

Matthew Perlmutter

Department of Mathematics

Massey University

Palmerston North 5301, New Zealand

Tudor S. Ratiu

Section de Mathématiques and Bernoulli Center

École Polytechnique Fédérale de Lausanne

CH-1015 Lausanne, Switzerland

Preface

This book is about *Symplectic Reduction by Stages* for Hamiltonian systems with symmetry. Reduction by stages means, roughly speaking, that we have two symmetry groups and we want to carry out symplectic reduction by both of these groups, either sequentially or all at once. More precisely, we shall start with a "large group" M that acts on a phase space P and assume that M has a normal subgroup N. The goal is to carry out reduction of the phase space P by the action of M in two stages; first by N and then by the quotient group M/N. For example, M might be the Euclidean group of $\mathbb{R}^3$, with N the translation subgroup so that M/N is the rotation group. In the Poisson context such a reduction by stages is easily carried out and we shall show exactly how this goes in the text. However, in the context of symplectic reduction, things are not nearly as simple because one must also introduce momentum maps and keep track of the level set of the momentum map at which one is reducing. But this gives an initial flavor of the type of problem with which the book is concerned.

As we shall see in this book, carrying out reduction by stages, first by N and then by M/N, rather than all in "one-shot" by the "large group" M is often not only a much simpler procedure, but it also can give nontrivial additional information about the reduced space. Thus, reduction by stages can provide an essential and useful tool for computing reduced spaces, including coadjoint orbits, which is useful to researchers in symplectic geometry and geometric mechanics.

Reduction theory is an old and time-honored subject, going back to the early roots of mechanics through the works of Euler, Lagrange, Poisson,

Liouville, Jacobi, Hamilton, Riemann, Routh, Noether, Poincaré, and others. These founding masters regarded reduction theory as a useful tool for simplifying and studying concrete mechanical systems, such as the use of Jacobi's *elimination of the node* in the study of the n-body problem to deal with the overall rotational symmetry of the problem. Likewise, Liouville and Routh used the elimination of cyclic variables (what we would call today an Abelian symmetry group) to simplify problems and it was in this setting that the *Routh stability method* was developed.

The modern form of symplectic reduction theory begins with the works of Arnold [1966a], Smale [1970], Meyer [1973], and Marsden and Weinstein [1974]. A more detailed survey of the history of reduction theory can be found in the first Chapter of the present book. As was the case with Routh, this theory has close connections with the stability theory of *relative equilibria*, as in Arnold [1969] and Simo, Lewis and Marsden [1991]. The symplectic reduction method is, in fact, by now so well known that it is used as a standard tool, often without much mention. It has also entered many textbooks on geometric mechanics and symplectic geometry, such as Abraham and Marsden [1978], Arnold [1989], Guillemin and Sternberg [1984], Libermann and Marle [1987], and McDuff and Salamon [1995]. Despite its relatively old age, research in reduction theory continues vigorously today and this book is a contribution to that theory.

Already in the original papers (such as Marsden and Weinstein [1974]), the issue of performing reduction by stages comes up. That is, one wants a framework in which repeated reduction by two successive symmetry groups can be performed and the result is the same as that of a single larger group. However, even this elementary question has some surprises.

For example, one of the nicest examples of reduction by stages is the theory of *semidirect product reduction* that is due to Guillemin and Sternberg [1980] and Marsden, Ratiu and Weinstein [1984a,b] and which is presented in Chapter 4 of this book. This theory is more than just a verification that reduction for a semidirect product can be done in two stages or, equivalently, all at once. In fact, information and procedures that are useful and powerful in a variety of examples, emerged from that effort. Application areas abound: the heavy top, compressible fluids, magnetohydrodynamics, the dynamics of underwater vehicles, etc. Mathematical techniques, such as the determination of coadjoint orbits in semidirect products, in conjunction with cotangent bundle reduction theory were also developed. Motivated by this success, it was only natural that generalizations would be sought.

In fact, work on a setting for a generalization of semidirect product theory was begun by two of us (JEM and TSR) during a visit to the Schrödinger Institute in Vienna in 1994. After a month or so of thinking about the question, it was realized that while the corresponding question for Poisson reduction was quite simple, the symplectic question was not so easy. It

was decided that the framework of starting with a "big group" M with a normal subgroup N and trying to first reduce by N and then by a variant of quotient group M/N was the right framework. Of course the semidirect product case, which was understood at the time, and which is a nontrivial special case, was an important and incentive that provided guidance.

With that modest start, the project slowly evolved and grew in various ways, with a progress report published as Marsden, Misiołek, Perlmutter and Ratiu [1998], and then ending up as this monograph. One of the ways in which it evolved was to ask corresponding questions in the context of Lagrangian reduction. That parallel effort resulted in several important works with Hernan Cendra, the most relevant one for this book being Cendra, Marsden, and Ratiu [2001a] on *Lagrangian Reduction by Stages*. Keeping contact with some of the key applications makes clear the importance of the magnetic terms that appear in the symplectic Hamiltonian and Lagrangian reduction of cotangent and tangent bundles, respectively. Joint work with Darryl Holm on a version of semidirect product reduction theory in the Lagrangian context was an important ingredient (see Holm, Marsden and Ratiu [1998]) in the general Lagrangian reduction by stages program. It was also a key component in the development of the Lagrangian averaged Euler (or Euler-α or LAE) equations as well as the Lagrangian averaged Navier-Stokes equations, also called the LANS-α equations.

Another ingredient that was a driver of some of the initial work was the attempt to understand the relation between cocycles for central extensions (such as the Bott–Virasoro cocycle) and curvatures of connections (such as the mechanical connection) that one uses in the theory of cotangent bundle reduction. These cocycles arise in the study of, for example, the KdV equation and the Camassa-Holm equation (see Ovsienko and Khesin [1987], Misiołek [1997, 1998] and Marsden and Ratiu [1999]) as well as in examples such as spin glasses, as in Holm and Kupershmidt [1988]. We believe that in this book, we have succeeded to a large extent in the interesting task of relating cocycles and magnetic terms. See Cendra, Marsden, and Ratiu [2003] for related ideas.

The work got several important boosts as it proceeded. One was from the PhD thesis of Matt Perlmutter (Perlmutter [1999]) and the second was from a productive visit to Caltech of Gerard Misiołek during 1997–1998. A preliminary version of our results were published in Marsden, Misiołek, Perlmutter and Ratiu [1998], appropriately enough in a celebratory volume for Victor Guillemin. Another boost came in discussions with Juan–Pablo Ortega about how the newly developing theory of optimal reduction based on a distribution-theoretic approach to Hamiltonian conservation laws (Ortega and Ratiu [2002], Ortega [2002]) might fit into the picture. His perspective led to an improvement on and an identification of situations where the hypotheses necessary for reduction by stages (the so called "stages hypothesis") are satisfied. The "optimal" oriented techniques also

enabled Juan–Pablo, among other things, to extend the reduction by stages method to the singular case, which appears as Part III of this book.

The singular case is important for many examples; for instance, using the result of Smale [1970] that the momentum map is not regular at points with nontrivial infinitesimal symmetry, Arms, Marsden and Moncrief [1981] showed under rather general circumstances (including in the infinite dimensional case) that the level sets of the momentum map have quadratic singularities. This sort of situation happens in interesting examples, such as Yang-Mills theory and general relativity. There are many other examples of singular reduction, such as those occurring in resonant phenomena (see, for example, Kummer [1981]; Cushman and Rod [1982] and Alber, Luther, Marsden, and Robbins [1998]. It was Sjamaar [1990] and Sjamaar and Lerman [1991] who began the systematic development of the corresponding singular reduction theory. These initial steps, while important, were also limited (they assumed the groups were compact, only dealt with reduction at zero, etc), but the theory rapidly developed in the 1990s and the early 2000s. We summarize some of the key results in this area in §1.4 and will survey additional literature in the historical survey in §1.3. A brief account of singular cotangent bundle reduction is given in §2.4.

Hopefully the above explains how, through this long saga, the three parts of the book came into existence. But it has a happy ending: the theory is not only very attractive, but is now also fairly comprehensive. Of course this does not mean that interesting questions are not left—there are many and we try to point out some of them as we proceed.

Structure of the Book.   The book has three parts. The first part gives a fairly complete treatment of regular symplectic reduction, cotangent bundle reduction and also gives an outline of the singular case. We do this for the convenience of the reader as this material is somewhat scattered in the literature. The second part develops the theory of Hamiltonian reduction by stages in the regular case, including a complete treatment of semidirect product reduction theory from the stages point of view. The third part develops this theory in the singular case, that is, the case when the reduced manifolds can have singularities, typically because the symmetry group action is not free, as was mentioned above. While Parts II and III use rather different techniques, the two together make the subject whole. Both theoretically and from the point of view of examples, our view is that it is not helpful to regard the regular case as a special case of the singular case. Thus, we have kept them separate in the two parts.

Prerequisites.   It will be assumed that the reader is familiar with the basics of geometric mechanics. While some of this will be recalled as we proceed, this will be mainly for purposes of establishing notation and conventions. In short, we assume everything that is in Marsden and Ratiu [1999], including the construction of momentum maps and their proper-

ties; we shall recall, for the reader's convenience, some of the basic theory of symplectic reduction that can be found in, for example, Abraham and Marsden [1978], Marsden [1992] or in one of the many other books on geometric mechanics. As we proceed, we shall review some additional material needed later as well, such as cotangent bundle reduction theory and the theory of principal connections. Again, this is primarily for the reader's convenience.

Part III on singular reduction by stages (by Juan-Pablo Ortega), will require material on singular reduction techniques, so at that point, the reader will need to consult other sources; the main terminology, tools, and techniques are outlined in §1.4 and details can be found in the book Ortega and Ratiu [2004a]. While Part III does not deal with cotangent bundles, an outline of what is known to the authors on singular cotangent bundle reduction is provided in §2.4 for the readers information.

Much of our work on Hamiltonian reduction by stages appears in this book for the first time. The work is just too large to publish in the journal literature without fragmentation and it seemed best to keep it together as a coherent whole.

What is not Covered in this Book. There is a lot that we *do not cover* in this book. As should be clear from the above remarks, we *do focus* on symplectic reduction by stages motivated by both applications and the intrinsic mathematical structure. There are many other aspects of reduction as well, such as Poisson reduction, Lagrangian reduction and Routh reduction. *We do not cover these topics in this work, but a discussion and references are given in the introductory chapter.* For example, Lagrangian reduction itself already deserves a separate monograph, although fairly comprehensive accounts already exist, such as Marsden, Ratiu and Scheurle [2000] and Cendra, Marsden, and Ratiu [2001a].

Another thing we do not cover in this book in a *systematic way* is the analytical (function space) theory in the infinite dimensional case, despite the fact that many of the most interesting examples are, in fact, infinite dimensional. Again this topic deserves a monograph of its own—the *general theory* of infinite dimensional Hamiltonian systems has some way to go, although there has been some general progress, as, for instance, Chernoff and Marsden [1974] and Mielke [1991] and references therein. There are also a number of research papers in this area and we give some specific references in the main text. We give a number of additional comments in §3.2 and, based on Gay-Balmaz and Ratiu [2006] and Gay-Balmaz [2007], outline one example of reduction by stages with all the functional analytic details taken care of in some detail, namely in §9.5 we discuss the case of a fluid in a symmetric container.

We also do not cover the interesting links that reduction theory has with representation theory and quantization and we do not touch all the other

interesting developments in symplectic topology. Other than a brief mention in connection with Teichmüller theory in §9.4 and its link to coadjoint orbits of the Bott-Virasoro group, we do not discuss the interesting applications to moduli spaces of connections (see Atiyah and Bott [1982], Goldman and Millson [1990] and Takhtajan and Teo [2004, 2006]).

Apology.   As usual in an advanced book with a relatively broad scope, we must apologize in advance to all the researchers in the area whose favorite topic or reference is not found here. Of course it is not possible to be complete with either task as the subject is now too developed and far reaching. However, we would be very happy to receive constructive suggestions for future printings.

Abbreviations.   We shall be referring to a few references often, so it will be convenient to have abbreviations for them;

We refer to Abraham and Marsden [1978] as [FofM].

We refer to Abraham, Marsden and Ratiu [1988] as [MTA].

We refer to Marsden and Ratiu [1999] as [MandS].

We refer to Marsden [1992] as [LonM].

We refer to Marsden, Misiołek, Perlmutter and Ratiu [1998] as [MMPR].

We refer to Ortega and Ratiu [2004a] as [HRed].

Notation.   To keep things reasonably systematic in the book, we have adopted the following universal conventions for some common maps:

Cotangent bundle projection: $\pi_Q : T^*Q \to Q$

Tangent bundle projection: $\tau_Q : TQ \to Q$

Quotient projection: $\pi_{P,G} : P \to P/G$

Tangent map: $T\varphi : TM \to TN$ for the tangent of a map $\varphi : M \to N$

Thus, for example, the symbol $\pi_{T^*Q,G}$ would denote the quotient projection from T^*Q to $(T^*Q)/G$.

Acknowledgments.   We are very grateful to many colleagues for their collaboration and for their input, directly or indirectly. We are especially grateful to Alan Weinstein, Victor Guillemin and Shlomo Sternberg for their incredible insights and work over the last few decades that was directly or indirectly inspirational for this volume. We also thank Hernan Cendra and Darryl Holm, our collaborators on the complementary efforts in the Lagrangian context. We would also like to thank Richard Cushman, especially for his helpful comments in the singular case, and Karl-Hermann Neeb and

May, 2007

Jerrold E. Marsden
Gerard Misiołek
Juan-Pablo Ortega
Matt Perlmutter
Tudor Ratiu

Contents

Part I. Background and the Problem Setting

The purpose of this first part is to provide background from those parts of geometric mechanics that are needed in the remainder of the book. The first chapter contains background on regular symplectic reduction and includes all the proofs of the main theorems, such as point reduction, coadjoint orbits and orbit reduction. It also gives an overview of related research topics in geometric mechanics.

The second chapter starts with a review, again including proofs, of connections on principal bundles, including curvature. This is needed background for one of the important constructions for the book, namely cotangent bundle reduction, whose reduced spaces involves non-canonical symplectic structures, namely magnetic, or curvature terms. The chapter also gives the problem setting and explains why reduction by stages is relatively routine in the Poisson setting, while being quite nontrivial in the symplectic setting. The chapter ends with a survey and discussion of various applications, related areas and future directions, such as swimming fish, loop groups, the Bott-Virasoro group, and multisymplectic geometry.

As explained in the preface, the book assumes that the reader is familiar with the essentials of geometric mechanics, up to, but not including symplectic reduction theory; this background is given in, for example, [MandS] (see the abbreviation code for the major references in the preface). Of course people with a strong background in geometric mechanics can proceed directly to Part II.

2   Contents

1
Symplectic Reduction

The purpose of this introductory Chapter is to both establish basic notation and also to give a reasonably complete account of symplectic reduction theory. The first section is a basic introduction, the second provides proofs and the third gives a history of the subject. This chapter focuses on reduction theory in the general setting of symplectic manifolds. The next chapter deals with, amongst other things, the important case of cotangent bundle reduction. Both of these cases are fundamental ingredients in the reduction by stages program.

1.1   Introduction to Symplectic Reduction

Roughly speaking, here is how symplectic reduction goes: given the symplectic action of a Lie group on a symplectic manifold that has a momentum map, one divides a level set of the momentum map by the action of a suitable subgroup to form a new symplectic manifold. Before the division step, one has a manifold (that can be singular if the points in the level set have symmetries) carrying a degenerate closed 2-form. Removing such a degeneracy by passing to a quotient space is a differential-geometric operation that was promoted by Cartan [1922].

The "suitable subgroup" related to a momentum mapping was identified by Smale [1970] in the special context of cotangent bundles. It was Smale's work that inspired the general symplectic construction by Meyer [1973] and the version we shall use, which makes explicit use of the properties of

momentum maps, by Marsden and Weinstein [1974].

Momentum Maps.   Let G be a Lie group, $\mathfrak{g}$ its Lie algebra, and $\mathfrak{g}^*$ be its dual. Suppose that G acts symplectically on a symplectic manifold P with symplectic form denoted by Ω. We shall denote the infinitesimal generator associated with the Lie algebra element ξ by ξ_P and we shall let the Hamiltonian vector field associated to a function $f : P \to \mathbb{R}$ be denoted X_f.

Recall that a **momentum map** is a map $\mathbf{J} : P \to \mathfrak{g}^*$, which is defined by the condition

$$\xi_P = X_{\langle \mathbf{J}, \xi \rangle} \tag{1.1.1}$$

for all $\xi \in \mathfrak{g}$, and where $\langle \mathbf{J}, \xi \rangle : P \to \mathbb{R}$ is defined by the natural pointwise pairing. We call such a momentum map **equivariant** when it is equivariant with respect to the given action on P and the coadjoint action of G on $\mathfrak{g}^*$. That is,

$$\mathbf{J}(g \cdot z) = \mathrm{Ad}^*_{g^{-1}} \mathbf{J}(z) \tag{1.1.2}$$

for every $g \in G$, $z \in P$, where $g \cdot z$ denotes the action of g on the point z, Ad denotes the adjoint action, and Ad^* the coadjoint action.[1] A quadruple $(P, \Omega, G, \mathbf{J})$, where (P, Ω) is a given symplectic manifold and $\mathbf{J} : P \to \mathfrak{g}^*$ is an equivariant momentum map for the symplectic action of a Lie group G, is sometimes called a **Hamiltonian G-space**.

Taking the derivative of the equivariance identity (1.1.2) with respect to g at the identity yields the condition of **infinitesimal equivariance**:

$$T_z \mathbf{J}(\xi_P(z)) = - \, \mathrm{ad}^*_\xi \mathbf{J}(z) \tag{1.1.3}$$

for any $\xi \in \mathfrak{g}$ and $z \in P$. Here, $\mathrm{ad}_\xi : \mathfrak{g} \to \mathfrak{g}; \eta \mapsto [\xi, \eta]$ is the adjoint map and $\mathrm{ad}^*_\xi : \mathfrak{g}^* \to \mathfrak{g}^*$ is its dual. A computation shows that (1.1.3) is equivalent to

$$\langle \mathbf{J}, [\xi, \eta] \rangle = \{ \langle \mathbf{J}, \xi \rangle, \langle \mathbf{J}, \eta \rangle \} \tag{1.1.4}$$

for any $\xi, \eta \in \mathfrak{g}$, that is, $\langle \mathbf{J}, \cdot \rangle : \mathfrak{g} \to \mathcal{F}(P)$ is a Lie algebra homomorphism, where $\mathcal{F}(P)$ denotes the Poisson algebra of smooth functions on P. The converse is also true if the Lie group is connected, that is, if G is connected then an infinitesimally equivariant action is equivariant (see [MandS], §12.3).

The idea that an action of a Lie group G with Lie algebra $\mathfrak{g}$ on a symplectic manifold P should be accompanied by such an equivariant momentum map $\mathbf{J} : P \to \mathfrak{g}^*$ and the fact that the orbits of this action are themselves

---

[1]Note that when we write $\mathrm{Ad}^*_{g^{-1}}$, we *literally* mean the adjoint of the linear map $\mathrm{Ad}_{g^{-1}} : \mathfrak{g} \to \mathfrak{g}$. The inverse of g is necessary for this to be a *left* action on $\mathfrak{g}^*$. Some authors let that inverse be understood in the notation. However, in this book, such a convention would be a notational disaster since we need to deal with both *left* and *right* actions, a distinction that is essential in mechanics.

symplectic manifolds both occur already in Lie [1890]; the links with mechanics also rely on the work of Lagrange, Poisson, Jacobi and Noether. In modern form, the momentum map and its equivariance were rediscovered by Kostant [1966] and Souriau [1966, 1970] in the general symplectic case and by Smale [1970] for the case of the lifted action from a manifold Q to its cotangent bundle $P = T^*Q$. Recall that the equivariant momentum map in this case is given explicitly by

$$\langle \mathbf{J}(\alpha_q), \xi \rangle = \langle \alpha_q, \xi_Q(q) \rangle, \tag{1.1.5}$$

where $\alpha_q \in T_q^*Q$, $\xi \in \mathfrak{g}$, and where the angular brackets denote the natural pairing on the appropriate spaces.

Smale referred to $\mathbf{J}$ as the "angular momentum" by generalization from the special case $G = SO(3)$, while Souriau used the French word "moment". Marsden and Weinstein [1974], followed usage emerging at that time and used the word "moment" for $\mathbf{J}$, but they were soon corrected by Richard Cushman and Hans Duistermaat, who suggested that the proper English translation of Souriau's French word was "momentum," which fit better with Smale's designation as well as standard usage in mechanics. Since 1976 or so, most people who have contact with mechanics use the term momentum map (or mapping). On the other hand, Guillemin and Sternberg popularized the continuing use of "moment" in English, and both words coexist today. It is a curious twist, as comes out in work on collective nuclear motion (Guillemin and Sternberg [1980]) and plasma physics (Marsden and Weinstein [1982] and Marsden, Weinstein, Ratiu, Schmid, and Spencer [1982]), that moments of inertia and moments of probability distributions can actually be the values of momentum maps! Mikami and Weinstein [1988] attempted a linguistic reconciliation between the usage of "moment" and "momentum" in the context of groupoids. See [MandS] for more information on the history of the momentum map and §1.3 for a more systematic review of general reduction theory.

Cocycles and Nonequivariant Momentum Maps.   Consider a momentum map $\mathbf{J} : P \to \mathfrak{g}^*$ that *need not be equivariant*, where P is a symplectic manifold on which a Lie group G acts symplectically. The map $\sigma : G \to \mathfrak{g}^*$ that is defined by

$$\sigma(g) := \mathbf{J}(g \cdot z) - \mathrm{Ad}^*_{g^{-1}} \mathbf{J}(z), \tag{1.1.6}$$

where $g \in G$ and $z \in P$ is called a **nonequivariance one-cocycle**. Clearly, *σ is a measure of the lack of equivariance of the momentum map*.

We shall now prove a number of facts about σ. The first claim is that *σ does not depend on the point $z \in P$ provided that the symplectic manifold P is connected* (otherwise it is constant on connected components). To prove this, we first recall the following *equivariance identity for infinitesimal generators*:

$$T_q \Phi_g \left(\xi_P(q) \right) = (\mathrm{Ad}_g \, \xi)_P(g \cdot q); \quad i.e., \quad \Phi_g^* \xi_P = \left(\mathrm{Ad}_{g^{-1}} \, \xi \right)_P. \tag{1.1.7}$$

This is an easy Lie group identity that is proved, for example, in [MandS], Lemma 9.3.7.

We shall now show that $\sigma(g)$ is constant by showing that its Hamiltonian vector field vanishes. This is done as follows. For each $\xi \in \mathfrak{g}$, we have

$$
\begin{aligned}
X_{\langle \sigma(g), \xi \rangle} &= X_{\Phi_g^* \langle \mathbf{J}, \xi \rangle} - X_{\langle \mathbf{J}, \mathrm{Ad}_{g^{-1}} \xi \rangle} \\
&= \Phi_g^* X_{\langle \mathbf{J}, \xi \rangle} - X_{\langle \mathbf{J}, \mathrm{Ad}_{g^{-1}} \xi \rangle} \\
&= \Phi_g^* \xi_P - \left(\mathrm{Ad}_{g^{-1}} \xi \right)_P,
\end{aligned}
$$

which vanishes in view of equation (1.1.7). In going from the first line to the second we used the fact that the action Φ_g is symplectic and in going to the last line, we used the definition of the momentum map.

Next, using the fact that $\sigma(g)$ is independent of z along with the basic identity $\mathrm{Ad}_{gh} = \mathrm{Ad}_g \, \mathrm{Ad}_h$ and its consequence $\mathrm{Ad}^*_{(gh)^{-1}} = \mathrm{Ad}^*_{g^{-1}} \mathrm{Ad}^*_{h^{-1}}$, we have the following calculation

$$
\begin{aligned}
\sigma(gh) &= \mathbf{J}((gh) \cdot z) - \mathrm{Ad}^*_{(gh)^{-1}} \mathbf{J}(z) \\
&= \left[\mathbf{J}(g \cdot (h \cdot z)) - \mathrm{Ad}^*_{g^{-1}} \mathbf{J}(h \cdot z) \right] \\
&\quad + \left[\mathrm{Ad}^*_{g^{-1}} \mathbf{J}(h \cdot z) - \mathrm{Ad}^*_{g^{-1}} \mathrm{Ad}^*_{h^{-1}} \mathbf{J}(z) \right] \\
&= \sigma(g) + \mathrm{Ad}^*_{g^{-1}} \sigma(h),
\end{aligned}
$$

which shows that σ satisfies the **cocycle identity**

$$
\sigma(gh) = \sigma(g) + \mathrm{Ad}^*_{g^{-1}} \sigma(h) \tag{1.1.8}
$$

for any $g, h \in G$.

The cocycle identity shows that σ produces a new action $\Theta : G \times \mathfrak{g}^* \to \mathfrak{g}^*$ defined by

$$
\Theta(g, \mu) := \mathrm{Ad}^*_{g^{-1}} \mu + \sigma(g) \tag{1.1.9}
$$

with respect to which the momentum map $\mathbf{J}$ is obviously equivariant. This action Θ is not linear anymore—it is an **affine action**.

For $\eta \in \mathfrak{g}$, let $\sigma_\eta(g) = \langle \sigma(g), \eta \rangle$. Differentiating the definition of σ, namely

$$
\sigma_\eta(g) = \langle \mathbf{J}(g \cdot z), \eta \rangle - \langle \mathbf{J}(z), \mathrm{Ad}_{g^{-1}} \eta \rangle
$$

with respect to g at the identity in the direction $\xi \in \mathfrak{g}$ gives

$$
\begin{aligned}
T_e \sigma_\eta(\xi) &= \langle T_z \mathbf{J} \cdot \xi_P(z), \eta \rangle + \langle \mathbf{J}(z), [\xi, \eta] \rangle \\
&= X_{\langle \mathbf{J}(z), \xi \rangle} [\langle \mathbf{J}(z), \eta \rangle] + \langle \mathbf{J}(z), [\xi, \eta] \rangle \\
&= - \left\{ \langle \mathbf{J}, \xi \rangle, \langle \mathbf{J}, \eta \rangle \right\} + \langle \mathbf{J}(z), [\xi, \eta] \rangle.
\end{aligned}
$$

This shows that

$$
T_e \sigma_\eta(\xi) = \Sigma(\xi, \eta), \tag{1.1.10}
$$

where $\Sigma(\xi, \eta)$, which is called the **infinitesimal nonequivariance two-cocycle** is defined by

$$\Sigma(\xi, \eta) = \langle \mathbf{J}, [\xi, \eta] \rangle - \{\langle \mathbf{J}, \xi \rangle, \langle \mathbf{J}, \eta \rangle\}. \tag{1.1.11}$$

Note that since σ does not depend on the point $z \in P$, neither does Σ. Also, it is clear from this definition that Σ *measures the lack of infinitesimal equivariance of* $\mathbf{J}$. Another way to look at this is to notice that from the derivation of equation (1.1.10), for $z \in P$ and $\xi \in \mathfrak{g}$, we have

$$T_z \mathbf{J}(\xi_P(z)) = -\operatorname{ad}_\xi^* \mathbf{J}(z) + \Sigma(\xi, \cdot). \tag{1.1.12}$$

Comparison of this relation with equation (1.1.3) also shows the relation between Σ and the infinitesimal equivariance of $\mathbf{J}$.

The map $\Sigma : \mathfrak{g} \times \mathfrak{g} \to \mathbb{R}$ is bilinear, skew-symmetric, and, as can be readily verified, satisfies the **two-cocycle identity**

$$\Sigma([\xi, \eta], \zeta) + \Sigma([\eta, \zeta], \xi) + \Sigma([\zeta, \xi], \eta) = 0, \tag{1.1.13}$$

for all $\xi, \eta, \zeta \in \mathfrak{g}$.

The Symplectic Reduction Theorem.   There are many precursors of symplectic reduction theory. When G is Abelian, the components of the momentum map form a system of functions in involution (i.e. the Poisson bracket of any two is zero). The use of k such functions to reduce a phase space to one having $2k$ fewer dimensions may be found already in the work of Lagrange, Poisson, Jacobi, and Routh; it is well described in, for example, Whittaker [1937].

In the nonabelian case, Smale [1970] noted that Jacobi's elimination of the node in $SO(3)$ symmetric problems can be understood as division of a nonzero angular momentum level by the $SO(2)$ subgroup which fixes the momentum value. In his setting of cotangent bundles, Smale clearly stated that the coadjoint isotropy group G_μ of $\mu \in \mathfrak{g}^*$ (defined to be the group of those $g \in G$ such that $g \cdot \mu = \mu$, where the dot indicates the coadjoint action), leaves the level set $\mathbf{J}^{-1}(\mu)$ invariant (Smale [1970], Corollary 4.5). However, he only divided by G_μ after fixing the total energy as well, in order to obtain the "minimal" manifold on which to analyze the reduced dynamics. The goal of his "topology and mechanics" program was to use topology, and specifically Morse theory, to study relative equilibria, which he did with great effectiveness.

Marsden and Weinstein [1974] combined Souriau's momentum map for general symplectic actions, Smale's idea of dividing the momentum level by the coadjoint isotropy group, and Cartan's idea of removing the degeneracy of a 2-form by passing to the leaf space of the form's null foliation. The key observation was that the leaves of the null foliation are precisely the (connected components of the) orbits of the coadjoint isotropy group (a fact we shall prove in the next section as the *reduction lemma*). An analogous

observation was made in Meyer [1973], except that Meyer worked in terms of a basis for the Lie algebra $\mathfrak{g}$ and identified the subgroup G_μ as the group which left the momentum level set $\mathbf{J}^{-1}(\mu)$ invariant. In this way, he did not need to deal with the equivariance properties of the coadjoint representation.

In the more general setting of symplectic manifolds with an equivariant momentum map for a symplectic group action, the fact that G_μ acts on $\mathbf{J}^{-1}(\mu)$ follows directly from equivariance of $\mathbf{J}$. Thus, it makes sense to form the **symplectic reduced space** which is defined to be the quotient space

$$P_\mu = \mathbf{J}^{-1}(\mu)/G_\mu. \tag{1.1.14}$$

Roughly speaking, the symplectic reduction theorem states that, under suitable hypotheses, P_μ is itself a symplectic manifold. To state this precisely, we need a short excursion on level sets of the momentum map and some facts about quotients.

Free and Proper Actions The action of a Lie group G on a manifold M is called a **free action** if $g \cdot m = m$ for some $g \in G$ and $m \in M$ implies that $g = e$, the identity element.

An action of G on M is called **proper** when the map $G \times M \to M \times M; (g,m) \mapsto (g \cdot m, m)$ is a proper map—that is, inverse images of compact sets are compact. This is equivalent to the statement that if m_k is a convergent sequence in M and if $g_k \cdot m_k$ converges in M, then g_k has a convergent subsequence in G.

As is shown in, for example, [MTA] and Duistermaat and Kolk [1999], freeness, together with properness implies that the quotient space M/G is a smooth manifold and that the projection map $\pi : M \to M/G$ is a smooth surjective submersion.

Locally Free Actions. An action of G on M is called **infinitesimally free** at a point $m \in M$ if $\xi_M(m) = 0$ implies that $\xi = 0$. An action of G on M is called **locally free** at a point $m \in M$ if there is a neighborhood U of the identity in G such that $g \in U$ and $g \cdot m = m$ implies $g = e$.

1.1.1 Proposition. *An action of a Lie group G on a manifold M is locally free at $m \in M$ if and only if it is infinitesimally free at m.*

Proof. If an action is locally free, it is clearly infinitesimally free because if it were not, there would be a nonzero $\xi \in \mathfrak{g}$ such that $\xi_M(m) = 0$; but then $\exp(t\xi)$ would be a family of group elements arbitrarily close to the identity that leave m fixed.

Conversely[2], suppose the action is infinitesimally free. Then it is easy to see that the map $G \times M \to M \times M$ given by $(g,x) \mapsto (g \cdot x, x)$ has an injective derivative at $x = m, g = e$. Thus, by the immersion theorem (see,

---

[2]We thank Alan Weinstein for pointing out this proof.

for example, [MTA]), the map itself restricted to a neighborhood of the identity in G and of m in M is injective. Thus it is locally free. ∎

A free action is obviously locally free. The converse is not true because the action of any discrete group is locally free, but need not be globally free. When one has an action that is locally free but not globally free, one is lead to the theory of orbifolds, as in Satake [1956]. In fact, quotients of manifolds by locally free and proper group actions are orbifolds, which follows by the use of the Palais slice theorem (see Palais [1957]). Orbifolds come up in a variety of interesting examples involving, for example, resonances; see, for instance, Cushman and Bates [1997] and Alber, Luther, Marsden, and Robbins [1998] for some specific examples.

Symmetry and Singularities.   If μ is a regular value of $\mathbf{J}$ then we claim that the action is automatically locally free at the elements of the corresponding level set $\mathbf{J}^{-1}(\mu)$. In this context it is convenient to introduce the notion of the **symmetry algebra** at $z \in P$ defined by

$$\mathfrak{g}_z = \{\xi \in \mathfrak{g} \mid \xi_P(z) = 0\}.$$

The symmetry algebra $\mathfrak{g}_z$ is the Lie algebra of the **isotropy subgroup** G_z of $z \in P$ defined by

$$G_z = \{g \in G \mid g \cdot z = z\}.$$

The following result (due to Smale [1970] in the special case of cotangent bundles and in general to Arms, Marsden and Moncrief [1981]), is important for the recognition of regular as well as singular points in the reduction process.

1.1.2 Proposition.   *An element $\mu \in \mathfrak{g}^*$ is a regular value of $\mathbf{J}$ if and only if $\mathfrak{g}_z = 0$ for all $z \in \mathbf{J}^{-1}(\mu)$.*

In other words, *points are regular points precisely when they have trivial symmetry algebra*. In examples, this gives an easy way to recognize regular points. For example, for the double spherical pendulum (see, for example, Marsden and Scheurle [1993a] or [LonM], one can say right away that the only singular points are those with *both* pendula pointing vertically (either straight down or straight up).

Proof.   By definition, z is a regular point if and only if $T_z\mathbf{J}$ is surjective. This is equivalent to saying that its annihilator is zero:

$$\{0\} = \{\xi \in \mathfrak{g} \mid \langle \xi, T_z\mathbf{J} \cdot v \rangle = 0, \text{ for all } v \in T_zP\}$$

Writing Ω_z (or sometimes $\Omega(z)$) for the value of Ω at z, we have

$$\langle \xi, T_z\mathbf{J} \cdot v \rangle = \Omega_z(\xi_P(z), v)$$

by definition of the momentum map. Therefore, z is a regular point if and only if

$$\{0\} = \{\xi \in \mathfrak{g} \mid \Omega_z(\xi_P(z), v) = 0 \text{ for all } v \in T_z P\}$$

which, as Ω_z is nondegenerate, is equivalent to $\mathfrak{g}_z = \{0\}$.   ∎

Notice that this result holds whether or not **J** is equivariant.

This result, connecting the symmetry of z with the regularity of μ, suggests that *points with symmetry are bifurcation points of* **J**. This observation turns out to have many important consequences, including some related key *convexity theorems*.

Now we are ready to state the symplectic reduction theorem. There is a variety of hypotheses that are important to understand, especially when we come later to the singular case. At the moment, we shall make one of the following two sets of hypotheses; other variants are discussed in the next section.

> **SR.** (P, Ω) is a symplectic manifold, G is a Lie group that acts symplectically on P and has an equivariant momentum map $\mathbf{J} : P \to \mathfrak{g}^*$.

> **SRFree.** G acts freely and properly on P.

> **SRRegular.** Assume that $\mu \in \mathfrak{g}^*$ is a regular value of **J** and that the action of G_μ on $\mathbf{J}^{-1}(\mu)$ is free and proper

From the previous discussion, note that condition **SRFree** implies condition **SRRegular**. The real difference is that **SRRegular** assumes local freeness of the action of G (which is equivalent to μ being a regular value, as we have seen), while **SRFree** assumes global freeness (on all of P).

1.1.3 Theorem (*Symplectic Reduction Theorem*).   *Assume that condition* **SR** *and that either the condition* **SRFree** *or the condition* **SRRegular** *holds. Then P_μ is a symplectic manifold, and is equipped with the **reduced symplectic form** Ω_μ that is uniquely characterized by the condition*

$$\pi_\mu^* \Omega_\mu = i_\mu^* \Omega, \tag{1.1.15}$$

where $\pi_\mu : \mathbf{J}^{-1}(\mu) \to P_\mu$ is the projection to the quotient space and where $i_\mu : \mathbf{J}^{-1}(\mu) \to P$ is the inclusion.

The proof of the symplectic reduction theorem is given, for the reader's convenience, in the next section along with variants of the hypotheses.

The above procedure is often called *point reduction* because one is fixing the value of the momentum map at a point $\mu \in \mathfrak{g}^*$. An equivalent reduction method called *orbit reduction* will be discussed shortly; see also §5.3.

Coadjoint Orbits. A standard example (due to Marsden and Weinstein [1974]) that we shall derive in detail in the next section, is the construction of the coadjoint orbits in $\mathfrak{g}^*$ of a group G by reduction of the cotangent bundle T^*G with its canonical symplectic structure and with G acting on T^*G by the cotangent lift of left (resp. right) group multiplication. In this case, one finds that $(T^*G)_\mu = \mathcal{O}_\mu$, the coadjoint orbit through $\mu \in \mathfrak{g}^*$. The reduced symplectic form is given by the ***Kostant, Kirillov, Souriau coadjoint form***, also referred to as the ***KKS form***:

$$\omega_{\mathcal{O}_\mu}^{\mp}(\nu)(\mathrm{ad}_\xi^* \nu, \, \mathrm{ad}_\eta^* \nu) = \mp \langle \nu, \, [\xi, \, \eta] \rangle, \qquad (1.1.16)$$

where $\xi, \eta \in \mathfrak{g}$, $\nu \in \mathcal{O}_\mu$, $\mathrm{ad}_\xi : \mathfrak{g} \rightarrow \mathfrak{g}$ is the adjoint operator defined by $\mathrm{ad}_\xi \eta := [\xi, \eta]$ and $\mathrm{ad}_\xi^* : \mathfrak{g}^* \rightarrow \mathfrak{g}^*$ is its dual. In this formula, one uses the minus sign for the left action and the plus sign for the right action. We recall that coadjoint orbits, like any group orbit is always an immersed manifold. Thus, one arrives at the following result (the details for which are given in the next section—see Theorem 1.2.3):

1.1.4 Corollary. *Given a Lie group G with Lie algebra $\mathfrak{g}$ and any point $\mu \in \mathfrak{g}^*$, the reduced space $(T^*G)_\mu$ is the coadjoint orbit $\mathcal{O}_\mu$ through the point μ; it is a symplectic manifold with symplectic form given by (1.1.16).*

This example, which "explains" Kostant, Kirillov and Souriau's formula for this structure, is typical of many of the ensuing applications, in which the reduction procedure is applied to a "trivial" symplectic manifold to produce something interesting.

Orbit Reduction. An important variant of the symplectic reduction theorem is called ***orbit reduction*** and, roughly speaking, it constructs $\mathbf{J}^{-1}(\mathcal{O})/G$, where $\mathcal{O}$ is a coadjoint orbit in $\mathfrak{g}^*$. In the next section—see Theorem 1.2.4—we show that orbit reduction is equivalent to the point reduction considered above.

Cotangent Bundle Reduction. The theory of cotangent bundle reduction is a very important special case of general reduction theory. Notice that the reduction of T^*G above to give a coadjoint orbit is a special case of the more general procedure in which G is replaced by a configuration manifold Q. The theory of cotangent bundle reduction will be outlined in the historical overview in this chapter, and then treated in some detail in the following chapter.

Mathematical Physics Links. Another example in Marsden and Weinstein [1974] came from general relativity, namely the reduction of the cotangent bundle of the space of Riemannian metrics on a manifold M by the action of the group of diffeomorphisms of M. In this case, restriction to the zero momentum level is the divergence constraint of general relativity, and so one is led to a construction of a symplectic structure on a space

closely related to the space of solutions of the Einstein equations, a question revisited in Fischer, Marsden, and Moncrief [1980] and Arms, Marsden and Moncrief [1982]. Here one sees a precursor of an idea of Atiyah and Bott [1982], which has led to some of the most spectacular applications of reduction in mathematical physics and related areas of pure mathematics, especially low-dimensional topology.

Singular Reduction.   In the preceding discussion, we have been making hypotheses that ensure the momentum levels and their quotients are smooth manifolds. Of course, this is not always the case, as was already noted in Smale [1970] and analyzed (even in the infinite-dimensional case) in Arms, Marsden and Moncrief [1981]. We give a review of some of the current literature on this singular case in §1.4 and in §2.4 as well as give a discussion of some of the history in §1.3; again, a thorough treatment of the singular case is given in [HRed] and we will rely on this theory in Part III.

Reduction of Dynamics.   Along with the geometry of reduction, there is also a theory of *reduction of dynamics*. The main idea is that a G-invariant Hamiltonian H on P 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].

Reduction theory has many interesting connections with the theory of integrable systems; we just mention some selected references Kazhdan, Kostant and Sternberg [1978]; Ratiu [1980a,b,c]; Bobenko, Reyman and Semenov-Tian-Shansky [1989]; Pedroni [1995]; Marsden and Ratiu [1986]; Vanhaecke [1996]; Bloch, Crouch, Marsden, and Ratiu [2002], which the reader can consult for further information.

1.2   Symplectic Reduction—Proofs and Further Details

To prove the symplectic reduction theorem, we prepare a few lemmas. The first refers to the reflexivity of the operation of taking the symplectic orthogonal complement.

1.2.1 Lemma.   *Let (V, Ω) be a finite dimensional symplectic vector space and $W \subset V$ be a subspace. Define the **symplectic orthogonal** to W by*

$$W^\Omega = \{v \in V \mid \Omega(v, w) = 0 \text{ for all } w \in W\}.$$

Then

$$(W^\Omega)^\Omega = W. \tag{1.2.1}$$

Proof.   That there is a natural inclusion $W \subset (W^\Omega)^\Omega$ follows directly from the definitions. To prove the converse inclusion we first show that

$$\dim V = \dim W + \dim W^\Omega, \tag{1.2.2}$$

even though $V \neq W \oplus W^\Omega$ in general. To prove (1.2.2), let $r : V^* \to W^*$ denote the restriction map, defined by $r(\alpha) = \alpha|W$, for $\alpha \in V^*$ and note that r *is onto* (since it is the dual of the inclusion map). Since Ω is nondegenerate, $\Omega^\flat : V \to V^*$, defined by $\Omega^\flat(v) = \Omega(v, \cdot)$, is also onto and thus $r \circ \Omega^\flat : V \to W^*$ is onto. Since $\ker(r \circ \Omega^\flat) = W^\Omega$, we conclude that V/W^Ω is isomorphic to W^*, whence we get from linear algebra that $\dim V - \dim W^\Omega = \dim W$, so (1.2.2) holds.

Applying (1.2.2) to W and then to W^Ω, we get

$$\dim V = \dim W + \dim W^\Omega = \dim W^\Omega + \dim(W^\Omega)^\Omega,$$

i.e.,

$$\dim W = \dim(W^\Omega)^\Omega.$$

This and the inclusion $W \subset (W^\Omega)^\Omega$ proves that $W = (W^\Omega)^\Omega$.   ∎

In what follows, we denote by $G \cdot z$ and $G_\mu \cdot z$ the G and G_μ-orbits through the point $z \in P$; note that if $z \in \mathbf{J}^{-1}(\mu)$ then $G_\mu \cdot z \subset \mathbf{J}^{-1}(\mu)$. Next we prove a key lemma that is central to the proof and is also useful in a number of situations.

1.2.2 Lemma (Reduction Lemma).   *Let P be a Poisson manifold and let $\mathbf{J} : P \to \mathfrak{g}^*$ be an equivariant momentum map of a Lie group action by Poisson maps of G on P. Let $G \cdot \mu$ denote the coadjoint orbit through a regular value $\mu \in \mathfrak{g}^*$ of $\mathbf{J}$. Then*

(i) $\mathbf{J}^{-1}(G \cdot \mu) = G \cdot \mathbf{J}^{-1}(\mu) = \{g \cdot z \mid g \in G \ and \ \mathbf{J}(z) = \mu\};$

(ii) $G_\mu \cdot z = (G \cdot z) \cap \mathbf{J}^{-1}(\mu);$

(iii) $\mathbf{J}^{-1}(\mu)$ *and* $G \cdot z$ ***intersect cleanly***, *i.e.*,

$$T_z(G_\mu \cdot z) = T_z(G \cdot z) \cap T_z(\mathbf{J}^{-1}(\mu));$$

(iv) *if (P, Ω) is symplectic, then $T_z(\mathbf{J}^{-1}(\mu)) = (T_z(G \cdot z))^\Omega$; i.e., the sets*

$$T_z(\mathbf{J}^{-1}(\mu)) \quad and \quad T_z(G \cdot z)$$

are Ω-orthogonal complements of each other.

Refer to Figure 1.2.1 for one way of visualizing the geometry associated with the reduction lemma. As it suggests, the two manifolds $\mathbf{J}^{-1}(\mu)$ and $G \cdot z$ intersect in the orbit of the isotropy group $G_\mu \cdot z$ and their tangent spaces $T_z\mathbf{J}^{-1}(\mu)$ and $T_z(G \cdot z)$ are symplectically orthogonal and intersect in the space $T_z(G_\mu \cdot z)$.
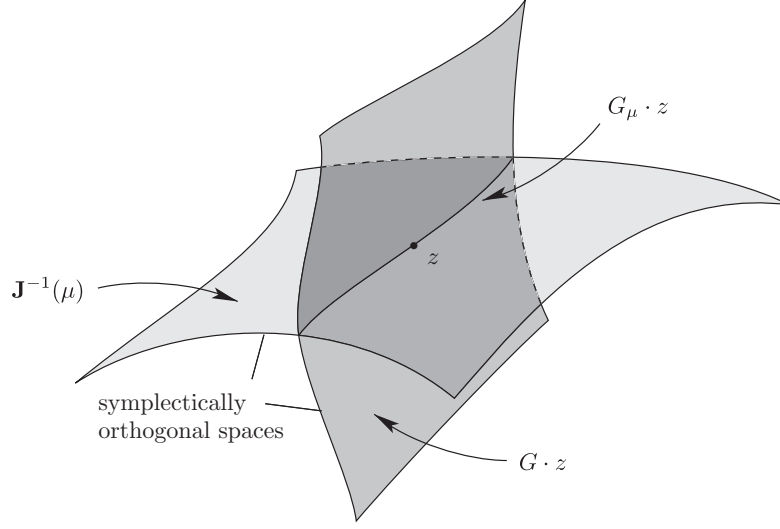
FIGURE 1.2.1. The geometry of the reduction lemma.

Proof of the Reduction Lemma.

(i) $z \in \mathbf{J}^{-1}(G \cdot \mu)$ if and only if $\mathbf{J}(z) = \mathrm{Ad}^*_{g^{-1}} \mu$ for some $g \in G$, which is equivalent to $\mu = \mathrm{Ad}^*_g \mathbf{J}(z) = \mathbf{J}(g^{-1} \cdot z)$, *i.e.*, $g^{-1} \cdot z \in \mathbf{J}^{-1}(\mu)$ and thus
$$z = g \cdot (g^{-1} \cdot z) \in G \cdot \mathbf{J}^{-1}(\mu).$$

(ii) $g \cdot z \in \mathbf{J}^{-1}(\mu)$ if and only if $\mu = \mathbf{J}(g \cdot z) = \mathrm{Ad}^*_{g^{-1}} \mathbf{J}(z) = \mathrm{Ad}^*_{g^{-1}} \mu$ if and only if $g \in G_\mu$.

(iii) First suppose that $v_z \in T_z(G \cdot z) \cap T_z(\mathbf{J}^{-1}(\mu))$. Then $v_z = \xi_P(z)$ for some $\xi \in \mathfrak{g}$ and $T_z \mathbf{J}(v_z) = 0$ which, by infinitesimal equivariance, gives $\mathrm{ad}^*_\xi \mu = 0$; *i.e.*, $\xi \in \mathfrak{g}_\mu$. If $v_z \in \xi_P(z)$ for $\xi \in \mathfrak{g}_\mu$ then $v_z \in T_z(G_\mu \cdot z)$. The reverse inclusion is immediate since by (ii) $G_\mu \cdot z$ is included in both $G \cdot z$ and $\mathbf{J}^{-1}(\mu)$.

(iv) The condition $v_z \in (T_z(G \cdot z))^\Omega$ means that $\Omega_z(\xi_P(z), v_z) = 0$ for all $\xi \in \mathfrak{g}$. This is equivalent to $\langle \mathbf{dJ}(z) \cdot v_z, \xi \rangle = 0$ for all $\xi \in \mathfrak{g}$ by definition of the momentum map. Thus, $v_z \in (T_z(G \cdot z))^\Omega$ if and only if $v_z \in \ker T_z \mathbf{J} = T_z(\mathbf{J}^{-1}(\mu))$. ∎

Notice from the statement (iv) that $T_z(\mathbf{J}^{-1}(\mu))^\Omega \subset T_z(\mathbf{J}^{-1}(\mu))$ provided that $G_\mu \cdot z = G \cdot z$. Thus, $\mathbf{J}^{-1}(\mu)$ *is coisotropic if* $G_\mu = G$; for example, this happens if $\mu = 0$ or if G is Abelian.

Proof of the Symplectic Reduction Theorem.   Since π_μ is a surjective submersion, if Ω_μ exists, it is uniquely determined by the condition $\pi_\mu^*\Omega_\mu = i_\mu^*\Omega$. This relation also defines Ω_μ in the following way.

For $v \in T_z\mathbf{J}^{-1}(\mu)$, let $[v] = T_z\pi_\mu(v)$ denote its equivalence class in $T_z\mathbf{J}^{-1}(\mu)/T_z(G_\mu \cdot z)$. We can use π_μ to identify the tangent space to the quotient at the point $[z] = \pi_\mu(z)$, namely, $T_{[z]}(\mathbf{J}^{-1}(\mu)/G_\mu)$ with the quotient of tangent spaces, namely $T_z(\mathbf{J}^{-1}(\mu))/T_z(G_\mu \cdot z)$. Then $\pi_\mu^*\Omega_\mu = i_\mu^*\Omega$ is equivalent to saying

$$\Omega_\mu([z])([v],[w]) = \Omega(z)(v,w)$$

for all $v, w \in T_z\mathbf{J}^{-1}(\mu)$. To see that this relation defines Ω_μ, let

$$y = \Phi_g(z), \quad v' = T_z\Phi_g \cdot v \quad \text{and} \quad w' = T_z\Phi_g \cdot w,$$

where $g \in G_\mu$. If, in addition $[v''] = [v']$ and $[w''] = [w']$, then

$$
\begin{aligned}
\Omega(y)(v'',w'') &= \Omega(y)(v',w') &&\text{(by the reduction lemma 1.2.2 (iv))}\\
&= \Omega(\Phi_g(z))(T_z\Phi_g \cdot v, T_z\Phi_g \cdot w)\\
&= (\Phi_g^*\Omega)(z)(v,w)\\
&= \Omega(z)(v,w) &&\text{(since the action is symplectic).}
\end{aligned}
$$

Thus Ω_μ is well-defined. It is smooth since $\pi_\mu^*\Omega_\mu$ is smooth. Since $\mathbf{d}\Omega = 0$,

$$\pi_\mu^*\mathbf{d}\Omega_\mu = \mathbf{d}\pi_\mu^*\Omega_\mu = \mathbf{d}i_\mu^*\Omega = i_\mu^*\mathbf{d}\Omega = 0.$$

Since π_μ is a surjective submersion, we conclude that $\mathbf{d}\Omega_\mu = 0$.

To prove nondegeneracy of Ω_μ, suppose that

$$\Omega_\mu([z])([v],[w]) = 0 \quad \text{for all} \quad w \in T_z(\mathbf{J}^{-1}(\mu)).$$

This means that

$$\Omega(z)(v,w) = 0 \quad \text{for all} \quad w \in T_z(\mathbf{J}^{-1}(\mu)),$$

i.e., that

$$v \in (T_z(\mathbf{J}^{-1}(\mu)))^\Omega = T_z(G \cdot z)$$

by Lemma 1.2.1 and the reduction Lemma 1.2.2 (iv). Hence

$$v \in T_z(\mathbf{J}^{-1}(\mu)) \cap T_z(G \cdot z) = T_z(G_\mu \cdot z)$$

by the reduction Lemma 1.2.2 (iii) so that $[v] = 0$, thus proving nondegeneracy of Ω_μ.   ■

Remarks on the Reduction Theorem.

1. Even if Ω is exact; say $\Omega = -\mathbf{d}\Theta$ and the action of G leaves Θ invariant, Ω_μ *need not be exact.* Perhaps the simplest example is a nontrivial coadjoint orbit of $\mathrm{SO}(3)$, which is a sphere with symplectic form given by the area form (by Stokes' theorem, it cannot be exact). That this is a symplectic reduced space of $T^*\mathrm{SO}(3)$ (with the canonical symplectic structure, so is exact) is shown in Theorem 1.2.3 below.

2. Continuing with the previous remark, assume that $\Omega = -\mathbf{d}\Theta$ and that the G_μ principal bundle $\mathbf{J}^{-1}(\mu) \to P_\mu := \mathbf{J}^{-1}(\mu)/G_\mu$ is trivializable; that is, it admits a *global* section $s : P_\mu \to \mathbf{J}^{-1}(\mu)$. Let $\Theta_\mu := s^* i_\mu^* \Theta \in \Omega^1(P_\mu)$. Then the reduced symplectic form $\Omega_\mu = -\mathbf{d}\Theta_\mu$ is exact.

   Indeed, making use of the fact that $\pi_\mu \circ s$ is the identity on P_μ and the defining relation of the reduced symplectic form Ω_μ, it follows that

$$\begin{aligned}
\Omega_\mu &= (\pi_\mu \circ s)^* \Omega_\mu = s^* \pi_\mu^* \Omega_\mu \\
&= s^* i_\mu^* \Omega = -s^* i_\mu^* \mathbf{d}\Theta \\
&= -\mathbf{d}s^* i_\mu^* \Theta = -\mathbf{d}\Theta_\mu.
\end{aligned}$$

   Note that this statement *does not imply that the one-form Θ descends to the reduced space*, only that the reduced symplectic form is exact and one if its primitives is Θ_μ. In fact, if one changes the global section, another primitive of Ω_μ is found which differs from Θ_μ by a closed one-form on P_μ.

3. The assumption that μ is a regular value of $\mathbf{J}$ is never really used in the proof. *The only hypothesis needed is that μ be a clean value of $\mathbf{J}$, i.e., $\mathbf{J}^{-1}(\mu)$ is a manifold and $T_z(\mathbf{J}^{-1}(\mu)) = \ker T_z\mathbf{J}$. This generalization applies, for instance, for zero angular momentum in the three dimensional two body problem, as was noted by Marsden and Weinstein [1974] and Kazhdan, Kostant and Sternberg [1978]; see also Guillemin and Sternberg [1984].

   Here are the general definitions: If $f : M \to N$ is a smooth map, a point $y \in N$ is called a ***clean value*** if $f^{-1}(y)$ is a submanifold and for each $x \in f^{-1}(y)$, $T_x f^{-1}(y) = \ker T_x f$. We say that f intersects a submanifold $L \subset N$ ***cleanly*** if $f^{-1}(L)$ is a submanifold of M and $T_x(f^{-1}(L)) = (T_x f)^{-1}(T_{f(x)}L)$. Note that *regular values of f are clean values and that if f intersects the submanifold L transversally, then it intersects it cleanly.* Also note that the definition of clean intersection of two manifolds is equivalent to the statement that the inclusion map of either one of them intersects the other cleanly. The reduction lemma is an example of this situation.

4. The freeness and properness of the G_μ action on $\mathbf{J}^{-1}(\mu)$ are used only to guarantee that P_μ is a manifold; these hypotheses can thus be replaced by the requirement that P_μ is a manifold and $\pi_\mu : \mathbf{J}^{-1}(\mu) \to P_\mu$ a submersion; the proof of the symplectic reduction theorem remains unchanged.

5. Even if μ is a regular value (in the sense of a regular value of the mapping $\mathbf{J}$), it need not be a ***regular point*** (also called a ***generic point***) in $\mathfrak{g}^*$; that is, a point whose coadjoint orbit is of maximal dimension. The reduction theorem *does not require that μ be a regular point.* For example, if G acts on itself on the left by group multiplication and if we lift this to an action on T^*G by the cotangent lift, then the action is free and so all μ are regular values, but such values (for instance, the zero element in $\mathfrak{so}(3)$) need not be regular. On the other hand, in many important stability considerations, a regularity assumption on the point μ is required; see, for instance, Patrick [1992], Ortega and Ratiu [1999b] and Patrick, Roberts, and Wulff [2004].

Nonequivariant Reduction.   We now describe how one can carry out reduction for a *nonequivariant momentum map. This result is needed in an essential way in the general reduction by stages construction.*

If $\mathbf{J} : P \to \mathfrak{g}^*$ is a nonequivariant momentum map on the connected symplectic manifold P with nonequivariance group one-cocycle σ consider the affine action (1.1.9) and let $\widetilde{G}_\mu$ be the isotropy subgroup of $\mu \in \mathfrak{g}^*$ relative to this action. Then, under the same regularity assumptions (for example, assume that G acts freely and properly on P, or that μ is a regular value of $\mathbf{J}$ and that $\widetilde{G}_\mu$ acts freely and properly on $\mathbf{J}^{-1}(\mu)$), the quotient manifold $P_\mu := \mathbf{J}^{-1}(\mu)/\widetilde{G}_\mu$ is a symplectic manifold whose symplectic form is uniquely determined by the relation $i_\mu^*\Omega = \pi_\mu^*\Omega_\mu$. The proof of this statement is identical to the one given above with the obvious changes in the meaning of the symbols.

When using nonequivariant reduction, one has to remember that G acts on $\mathfrak{g}^*$ in an *affine* and not a linear manner. For example, while the coadjoint isotropy subgroup at the origin is equal to G; that is, $G_0 = G$, this is no longer the case for the affine action, where $\widetilde{G}_0$ in general does not equal G.

Coadjoint Orbits as Symplectic Reduced Spaces.   We now prove Corollary 1.1.4; that is, we show that coadjoint orbits may be realized as reduced spaces. This provides a proof that they are symplectic manifolds[3] Historically, a direct argument was found first, by Kirillov, Kostant and Souriau in the early 1960's and the (minus) coadjoint symplectic structure was found to be

$$\omega_\nu^-\left(\mathrm{ad}_\xi^* \nu, \mathrm{ad}_\eta^* \nu\right) = -\left\langle \nu, [\xi, \eta] \right\rangle \tag{1.2.3}$$

---

[3]See [MandS], Chapter 14 for a "direct" or "bare hands" argument.

Interestingly, this is the symplectic structure on the symplectic leaves of the Lie-Poisson bracket, as is shown in, for example, [MandS]. (See the historical overview in §1.3 below and specifically, see equation (1.3.1) for a quick review of the Lie-Poisson bracket.)

The strategy of the reduction proof, as mentioned in the discussion in the last section, is to show that the coadjoint symplectic form on a coadjoint orbit $\mathcal{O}_\mu$ of the point μ, at a point $\nu \in \mathcal{O}$, may be obtained by symplectically reducing T^*G at the value μ. The following theorem (due to Marsden and Weinstein [1974]), and which is an elaboration on the result in Corollary 1.1.4, formulates the result for left actions; of course there is a similar one for right actions, with the minus sign replaced by a plus sign.

1.2.3 Theorem (Reduction to Coadjoint Orbits). *Let G be a Lie group and let G act on G (and hence on T^*G by cotangent lift) by left multiplication. Let $\mu \in \mathfrak{g}^*$ and let $\mathbf{J}_L : T^*G \to \mathfrak{g}^*$ be the momentum map for the left action. Then μ is a regular value of $\mathbf{J}_L$, the action of G is free and proper, the symplectic reduced space $\mathbf{J}_L^{-1}(\mu)/G_\mu$ is identified via left translation with $\mathcal{O}_\mu$, the coadjoint orbit through μ, and the reduced symplectic form coincides with ω^- given in equation (1.2.3).*

Proof.   Recall from, for example, [MandS], that $\mathbf{J}_L$ is given by *right* translation to the identity:

$$\mathbf{J}_L(\alpha_g) = T_e^* R_g \alpha_g \qquad (1.2.4)$$

Thus, $\mathbf{J}_L^{-1}(\mu)$ consists of those $\alpha_g \in P = T^*G$ such that $\alpha_g = T_g^* R_{g^{-1}}\mu$. In other words, if we extend μ to a *right* invariant 1-form α_μ on G, then its graph is $\mathbf{J}_L^{-1}(\mu)$. It is clear that the G action is free and proper on G and hence on T^*G. From this or directly, we see that each μ is a regular value (to see this directly, note that the derivative of $\mathbf{J}_L$ restricted to $\mathfrak{g}^*$ is already surjective).

Recall from the Lie-Poisson reduction theorem ([MandS], Chapter 13) that the reduction of T^*G by the *left* action of G is implemented by the *right* momentum map. Consistent with this, we claim that the map $\varphi : \mathbf{J}_L^{-1}(\mu) \to \mathcal{O}_\mu$ defined by $\alpha_g \mapsto \mathrm{Ad}_g^* \mu = T_e^* L_g \alpha_g$, *i.e.*, $\varphi = \mathbf{J}_R|\mathbf{J}_L^{-1}(\mu)$, induces a diffeomorphism $\overline{\varphi}$ between $\mathbf{J}_L^{-1}(\mu)/G_\mu$ and $\mathcal{O}_\mu$. There is indeed a map $\overline{\varphi}$ defined because if $\alpha_{hg} = T_g L_h \alpha_g$ for $h \in G_\mu$, then

$$\varphi(\alpha_{hg}) = \mathrm{Ad}_{hg}^* \mu = \mathrm{Ad}_g^* \mu = \varphi(\alpha_g).$$

Thus, $\overline{\varphi}$ is well-defined and it is readily checked to be a bijection. It is smooth since it is induced on the quotient by a smooth map. The derivative of φ induces an isomorphism at each point, as is readily checked (see the calculations below). Thus, $\overline{\varphi}$ is a diffeomorphism.

The reduced symplectic form on $\mathcal{O}_\mu \approx \mathbf{J}_L^{-1}(\mu)/G_\mu$ is induced by the canonical symplectic form Ω on T^*G pulled-back to $\mathbf{J}_L^{-1}(\mu)$. Let $\mathrm{ad}_\xi^* \nu$ and $\mathrm{ad}_\eta^* \nu$ be two tangent vectors to $\mathcal{O}_\mu$ at a point $\nu = \mathrm{Ad}_g^* \mu$. They are tangent

to the curves

$$c_\xi(t) = \mathrm{Ad}^*_{\exp(t\xi)} \mathrm{Ad}^*_g \mu = \mathrm{Ad}^*_{g \exp(t\xi)} \mu$$

and

$$c_\eta(t) = \mathrm{Ad}^*_{g \exp(t\eta)} \mu$$

respectively. Now notice that if we define

$$d_\xi(t) = \alpha_{g \exp(t\xi)} = T^*_{g \exp(t\xi)} R_{\exp(-t\xi)g^{-1}} \mu,$$

then $\varphi(d_\xi(t)) = c_\xi(t)$ and similarly for η. Now

$$d_\xi(t) = T^* L_{g^{-1}} T^* R_{\exp(-t\xi)} \nu = \Psi_g \Phi_{\exp(t\xi)} \nu$$

where Ψ denotes the *left* action and Φ denotes the right action. By the chain rule,

$$d'_\xi(0) = T\Psi_g \cdot \xi_{T^*G}(\nu).$$

Thus,

$$\Omega(\nu)(d'_\xi(0), d'_\eta(0)) = \Omega(\nu)(\xi_{T^*G}(\nu), \eta_{T^*G}(\nu)),$$

since Ψ_g is a symplectic map. Next, recall that $\Omega = -\mathbf{d}\Theta$ where Θ is the canonical 1-form, so by a standard formula for the exterior derivative of a 1-form γ on a manifold R (see [MTA]), namely,

$$\mathbf{d}\gamma(X,Y) = X[\gamma(Y)] - Y[\gamma(X)] - \gamma([X,Y]) \tag{1.2.5}$$

where X and Y are vector fields on R, we get

$$\Omega(X,Y) = -X[\Theta(Y)] + Y[\Theta(X)] + \Theta([X,Y])$$

for vector fields X and Y on T^*G. Thus,

$$\Omega(\nu)(\xi_{T^*G}(\nu), \eta_{T^*G}(\nu))$$
$$= -\xi_{T^*G}[\Theta(\eta_{T^*G})](\nu) + \eta_{T^*G}[\Theta(\xi_{T^*G})](\nu) + \Theta([\xi_{T^*G}, \eta_{T^*G}])(\nu).$$

Since we have a *right* action, $[\xi_{T^*G}, \eta_{T^*G}] = [\xi, \eta]_{T^*G}$. The definition of the canonical 1-form shows that on infinitesimal generators, it is given by $\Theta(\eta_{T^*G}) = \langle \mathbf{J}_R, \eta \rangle$, so the preceding displayed expression equals

$$-\{\langle \mathbf{J}_R, \eta \rangle, \langle \mathbf{J}_R, \xi \rangle\}(\nu) + \{\langle \mathbf{J}_R, \xi \rangle, \langle \mathbf{J}_R, \eta \rangle\}(\nu) + \Theta([\xi, \eta]_{T^*G})(\nu).$$

By equivariance of $\mathbf{J}_R$, the first and last (or second and last) cancel, leaving

$$-\langle \mathbf{J}_R(\nu), [\xi, \eta] \rangle = -\langle \nu, [\xi, \eta] \rangle,$$

which is the coadjoint orbit symplectic structure ω^-.   ∎

Remarks.

1. Notice that, as in the general Symplectic Reduction Theorem 1.1.3, this result does *not* require μ to be a regular (or generic) *point* in $\mathfrak{g}^*$; that is, arbitrarily nearby coadjoint orbits may have a different dimension.

2. As remarked earlier, the form ω^- on the orbit need not be exact even though Ω is. An example that shows this is $SO(3)$, whose coadjoint orbits are spheres and whose symplectic structure is, as shown in [MandS], a multiple of the area element, which is not exact by Stokes' Theorem.

3. The essence of the preceding proof can be stated as the identity

$$\Omega(\nu)\left(\xi_{T^*G}(\nu), \eta_{T^*G}(\nu)\right) = -\left\langle \nu, [\xi, \eta] \right\rangle. \qquad (1.2.6)$$

Orbit Reduction.   So far, we have presented what is usually called *point reduction.* There is another point of view that is called *orbit reduction*, which we now summarize. We assume the same set up as in the symplectic reduction theorem, with P connected, G acting symplectically, freely, and properly on P with an equivariant momentum map $\mathbf{J} : P \to \mathfrak{g}^*$.

The connected components of the point reduced spaces P_μ can be regarded as the symplectic leaves of the Poisson manifold $\left(P/G, \{\cdot, \cdot\}_{P/G}\right)$ in the following way. Form a map $[i_\mu] : P_\mu \to P/G$ defined by selecting an equivalence class $[z]_{G_\mu}$ for $z \in \mathbf{J}^{-1}(\mu)$ and sending it to the class $[z]_G$. This map is checked to be well-defined and smooth. We then have the commutative diagram

$$
\begin{array}{ccc}
\mathbf{J}^{-1}(\mu) & \xrightarrow{\ \ i_\mu\ \ } & P \\[0.5em]
\Big\downarrow{\scriptstyle \pi_\mu} & & \Big\downarrow{\scriptstyle \pi} \\[0.5em]
P_\mu & \xrightarrow{\ \ [i_\mu]\ \ } & P/G
\end{array}
$$

One then checks that $[i_\mu]$ is a Poisson injective immersion. Moreover, the $[i_\mu]$-images in P/G of the connected components of the symplectic manifolds (P_μ, Ω_μ) are its symplectic leaves (see [HRed] and references therein for details). As sets,

$$[i_\mu](P_\mu) = \mathbf{J}^{-1}(\mathcal{O}_\mu)/G,$$

where $\mathcal{O}_\mu \subset \mathfrak{g}^*$ is the coadjoint orbit through $\mu \in \mathfrak{g}^*$. The set

$$P_{\mathcal{O}_\mu} := \mathbf{J}^{-1}(\mathcal{O}_\mu)/G$$

is called the **orbit reduced space** associated to the orbit $\mathcal{O}_\mu$. The smooth manifold structure (and hence the topology) on $P_{\mathcal{O}_\mu}$ is the one that makes

$$[i_\mu] : P_\mu \to P_{\mathcal{O}_\mu}$$

into a diffeomorphism.

For the next theorem, which characterizes the symplectic form and the Hamiltonian dynamics on $P_{\mathcal{O}_\mu}$, recall the coadjoint orbit symplectic structure of Kirillov, Kostant and Souriau that was established in the preceding Theorem 1.2.3:

$$\omega_{\mathcal{O}_\mu}^-(\nu)(\xi_{\mathfrak{g}^*}(\nu),\, \eta_{\mathfrak{g}^*}(\nu)) = -\langle \nu,\, [\xi,\, \eta] \rangle, \tag{1.2.7}$$

for $\xi, \eta \in \mathfrak{g}$ and $\nu \in \mathcal{O}_\mu$.

We also recall that an injectively immersed submanifold of S of Q is called an **initial submanifold** of Q when for any smooth manifold P, a map $g : P \to S$ is smooth if and only if $\iota \circ g : P \to Q$ is smooth, where $\iota : S \hookrightarrow Q$ is the inclusion.

1.2.4 Theorem (Symplectic Orbit Reduction Theorem). *In the setup explained above, we have:*

(i) *The momentum map* $\mathbf{J}$ *is transverse to the coadjoint orbit* $\mathcal{O}_\mu$ *and hence* $\mathbf{J}^{-1}(\mathcal{O}_\mu)$ *is an initial submanifold of* P. *Moreover, the projection* $\pi_{\mathcal{O}_\mu} : \mathbf{J}^{-1}(\mathcal{O}_\mu) \to P_{\mathcal{O}_\mu}$ *is a surjective submersion.*

(ii) $P_{\mathcal{O}_\mu}$ *is a symplectic manifold with the symplectic form* $\Omega_{\mathcal{O}\mu}$ *uniquely characterized by the relation*

$$\pi_{\mathcal{O}_\mu}^* \Omega_{\mathcal{O}\mu} = \mathbf{J}_{\mathcal{O}_\mu}^* \omega_{\mathcal{O}_\mu}^- + i_{\mathcal{O}_\mu}^* \Omega, \tag{1.2.8}$$

where $\mathbf{J}_{\mathcal{O}_\mu}$ *is the restriction of* $\mathbf{J}$ *to* $\mathbf{J}^{-1}(\mathcal{O}_\mu)$ *and* $i_{\mathcal{O}_\mu} : \mathbf{J}^{-1}(\mathcal{O}_\mu) \hookrightarrow P$ *is the inclusion.*

(iii) *The map* $[i_\mu] : P_\mu \to P_{\mathcal{O}_\mu}$ *is a symplectic diffeomorphism.*

(iv) (Dynamics.) *Let* H *be a* G-*invariant function on* P *and define* $\widetilde{H} : P/G \to \mathbb{R}$ *by* $H = \widetilde{H} \circ \pi$. *Then the Hamiltonian vector field* X_H *is also* G-*invariant and hence induces a vector field on* P/G, *which coincides with the Hamiltonian vector field* $X_{\widetilde{H}}$. *Moreover, the flow of* $X_{\widetilde{H}}$ *leaves the symplectic leaves* $P_{\mathcal{O}_\mu}$ *of* P/G *invariant. This flow restricted to the symplectic leaves is again Hamiltonian relative to the symplectic form* $\Omega_{\mathcal{O}\mu}$ *and the Hamiltonian function* $\widetilde{H}_{\mathcal{O}_\mu}$ *given by*

$$\widetilde{H}_{\mathcal{O}_\mu} \circ \pi_{\mathcal{O}_\mu} = H \circ i_{\mathcal{O}_\mu}.$$

Note that if $\mathcal{O}_\mu$ is an embedded submanifold of $\mathfrak{g}^*$ then $\mathbf{J}$ is transverse to $\mathcal{O}_\mu$ and hence $\mathbf{J}^{-1}(\mathcal{O}_\mu)$ is automatically an embedded submanifold of P.

The proof of this theorem when $\mathcal{O}_\mu$ is an embedded submanifold of $\mathfrak{g}^*$ can be found in Marle [1976], Kazhdan, Kostant and Sternberg [1978], with useful additions given in Marsden [1981] and Blaom [2001]. For nonfree actions and when $\mathcal{O}_\mu$ is not an embedded submanifold of $\mathfrak{g}^*$ see [HRed]. Further comments on the historical context of this result are given in the next section.

Sketch of the Proof.   We focus on how the formula (1.2.8) is derived and refer to [HRed] for some of the technical details. The last part of the theorem is proved as in reduction of dynamics in the symplectic reduction theorem.

Part (i) of the previous theorem follows from the fact that the orbit of any smooth group action is an initial submanifold of the manifold on which the group acts. In particular, $\mathcal{O}_\mu$ is an initial submanifold of $\mathfrak{g}^*$. Additionally, the freeness of the action implies that the momentum map $\mathbf{J}$ is a submersion and hence transverse to $\mathcal{O}_\mu$. The claim on the initial character of $\mathbf{J}^{-1}(\mathcal{O}_\mu)$ follows from the transversality theorem for initial submanifolds (see [HRed] for a discussion). Since the G–action on $\mathbf{J}^{-1}(\mathcal{O}_\mu)$ is free and proper the quotient space $\mathbf{J}^{-1}(\mathcal{O}_\mu)/G$ is naturally endowed with a regular quotient manifold whose smooth structure coincides with the one that makes the map $[i_\mu]$ into a diffeomorphism.

Let us just write $\mathcal{O}$ for $\mathcal{O}_\mu$ for simplicity. For the proof of (ii) we will use the following.

1.2.5 Lemma.   *Under the hypotheses of the preceding Theorem, we have*

(i) $T_z(\mathbf{J}^{-1}(\mathcal{O})) = T_z(G \cdot z) + \ker(T_z\mathbf{J})$;

(ii) $\mathbf{J}^*_{\mathcal{O}}\omega^+_{\mathcal{O}}$ *restricted to* $T_z(G \cdot z) \times T_z(G \cdot z)$ *coincides with* Ω *restricted to the same space.*

Proof.   First we prove (i). The tangent space to $\mathbf{J}^{-1}(\mathcal{O})$ at z equals $T_z(\mathbf{J}^{-1}(\mathcal{O})) = (T_z\mathbf{J})^{-1}(T_{\mathrm{J}(z)}\mathcal{O})$. Thus by infinitesimal equivariance of $\mathbf{J}$,

$$T_{\mathrm{J}(z)}\mathcal{O} = \{-(\mathrm{ad}_\xi)^*\mathbf{J}(z) \mid \xi \in \mathfrak{g}\} = \{T_z\mathbf{J}(\xi_P(z)) \mid \xi \in \mathfrak{g}\}$$
$$= \{T_z\mathbf{J}(v) \mid v \in T_z(G \cdot z)\} = (T_z\mathbf{J})(T_z(G \cdot z)).$$

Applying $(T_z\mathbf{J})^{-1}$ gives the desired result.

To prove (ii), we use (i) and let $v = \xi_P(z) + v'$ and $w = \eta_P(z) + w'$, where $\xi, \eta \in \mathfrak{g}$, and $v', w' \in \ker(T_z\mathbf{J})$, be two arbitrary vectors in $T_z(\mathbf{J}^{-1}(\mathcal{O}))$. We have

$$(\mathbf{J}^*_{\mathcal{O}}\omega^+_{\mathcal{O}}(z)(v,w) = \omega^+_{\mathcal{O}}(\mathbf{J}(z))(T_z\mathbf{J}(v), T_z\mathbf{J}(w))$$
$$= \omega^+_{\mathcal{O}}(\mathbf{J}(z))(T_z\mathbf{J}(\xi_P(z)), T_z\mathbf{J}(\eta_P(z)))$$

Now using infinitesimal equivariance (equation $(1.1.3)$) and then the definition of $\omega_{\mathcal{O}}^+$, the preceding becomes

$$
\begin{aligned}
(\mathbf{J}_{\mathcal{O}}^* \omega_{\mathcal{O}}^+(z)(v,w) &= \omega_{\mathcal{O}}^+(\mathbf{J}(z))((\mathrm{ad}_\xi^*)\mathbf{J}(z), (\mathrm{ad}_\eta)^*\mathbf{J}(z)) \\
&= \langle \mathbf{J}(z), [\xi, \eta]\rangle = J([\xi,\eta])(z) = \{J(\xi), J(\eta)\}\,(z) \\
&= \Omega(z)(X_{J(\xi)}(z), X_{J(\eta)}(z)) = \Omega(z)(\xi_P(z), \eta_P(z))
\end{aligned}
$$

which proves (ii).   ▼

Now we are ready to prove equation $(1.2.8)$. Using part (i) of the Lemma, consider $\xi_P(z) + v,\ \eta_P(z) + w \in T_z(\mathbf{J}^{-1}(\mathcal{O}))$, where $v, w \in \ker(T_z\mathbf{J})$, which are two arbitrary vectors tangent to $\mathbf{J}^{-1}(\mathcal{O})$ at z. By part (ii) of the Lemma, the defining relation for $\Omega_{\mathcal{O}}$ is

$$
\Omega(z)(\xi_P(z) + v, \eta_P(z) + w) = \Omega_{\mathcal{O}}([z])([v],[w]) + \Omega(z)(\xi_P(z), \eta_P(z))
$$

or, by Ω-orthogonality of $T_z(G \cdot z)$ and $\ker(T_z\mathbf{J})$ (this is the Reduction Lemma $1.2.2$ (iv)),

$$
\Omega(z)(v,w) = \Omega_{\mathcal{O}}([z])([v],[w])
$$

for all $v, w \in \ker(T_z\mathbf{J})$, where

$$
[v] = T_z\pi_{\mathcal{O}}(v) \in T_z(\mathbf{J}^{-1}(\mathcal{O}))/T_z(G \cdot z) \cong T_{[z]}(\mathbf{J}^{-1}(\mathcal{O})/G),
$$

$[z] = \pi_{\mathcal{O}}(z)$, and similarly for w.

It is shown, just as in the Symplectic Reduction Theorem, that this relation defines $\Omega_{\mathcal{O}}$. Since Ω and $\omega_{\mathcal{O}}^-$ are closed and $\pi_{\mathcal{O}}$ is a surjective submersion, it follows that $\Omega_{\mathcal{O}}$ is also closed. It can be shown directly that $\Omega_{\mathcal{O}}$ is nondegenerate as in the Symplectic Reduction Theorem, or one can use (iii).

To prove (iii), notice that the relation $[i_\mu]^*\Omega_{\mathcal{O}} = \Omega_\mu$ is equivalent to $\pi_\mu^*[i_\mu]^*\Omega_{\mathcal{O}} = i_\mu^*\Omega$. Since $[i_\mu] \circ \pi_\mu = \pi_{\mathcal{O}} \circ i_\mu$, this says that $i_\mu^*\pi_{\mathcal{O}}^*\Omega_{\mathcal{O}} = i_\mu^*\Omega$. By (ii), we have

$$
i_\mu^*\pi_{\mathcal{O}}^*\Omega_{\mathcal{O}} = i_\mu^*(i_{\mathcal{O}}^*\Omega + \mathbf{J}_{\mathcal{O}}^*\omega_{\mathcal{O}}^-) = (i_{\mathcal{O}} \circ i_\mu)^*\Omega + (\mathbf{J}_{\mathcal{O}} \circ i_\mu)^*\omega_{\mathcal{O}}^- = i_\mu^*\Omega
$$

since $i_{\mathcal{O}} \circ i_\mu = i_\mu$ and $\mathbf{J}_{\mathcal{O}} \circ i_\mu = \mu$ on $\mathbf{J}^{-1}(\mu)$.   ∎

Remarks.

1. A similar result holds for right actions.

2. In this proof the freeness and properness of the G_μ-action on $\mathbf{J}^{-1}(\mu)$ were only used indirectly. In fact these conditions are sufficient but not necessary for P_μ to be a manifold. All that is needed is for P_μ to be a manifold and π_μ to be a submersion and the above proof remains unchanged.

3. Note that the description of the symplectic structure on $\mathbf{J}^{-1}(\mathcal{O})/G$ is not as simple as it was for $\mathbf{J}^{-1}(\mu)/G$, while the Poisson bracket description is simpler on $\mathbf{J}^{-1}(\mathcal{O})/G$. Of course, the symplectic structure depends only on the orbit $\mathcal{O}$ and not on the choice of a point μ on it.

Cotangent Bundle Reduction.   As mentioned in the last section, perhaps the most important and basic reduction theorem in addition to those already presented is the cotangent bundle reduction theorem. We shall give an exposition of the key aspects of this theory in §2.2 and §2.3, and give a historical account of its development, along with references in the next section. We shall also present, in Chapter 7, some generalizations that are needed in the reduction by stages context.

At this point, to orient the reader, we note that one of the special cases is cotangent bundle reduction at zero (see Theorem 2.2.2). This result says that if one has, again for simplicity, a free and proper action of G on Q (which is then lifted to T^*Q by the cotangent lift), then the reduced space at zero of T^*Q is given by $T^*(Q/G)$, with its canonical symplectic structure. On the other hand, reduction at a nonzero value is a bit more complicated and gives rise to modifications of the standard symplectic structure; namely, one adds to the canonical structure, the pull-back of a closed two form on Q to T^*Q. Because of their physical interpretation (discussed, for example, in [MandS]), such extra terms are called magnetic terms. In §2.1 and §2.2 and §2.3, we prove the basic cotangent bundle reduction theorems along with providing some of the other important notions, such as the mechanical connection and the locked inertia tensor. Other notions that are important in mechanics, such as the amended potential, can be found in [LonM].

1.3   Reduction Theory: Historical Overview

We have already given bits an pieces of the history of symplectic reduction and momentum maps. In this section we take a broader view of the subject to put things in historical and topical context.

History before 1960.   In the preceding sections, reduction theory has been presented as a mathematical construction. Of course, these ideas are rooted in classical work on mechanical systems with symmetry by such masters as Euler, Lagrange, Hamilton, Jacobi, Routh, Riemann, Liouville, Lie, and Poincaré. The aim of their work was, to a large extent, to eliminate variables associated with symmetries in order to simplify calculations in concrete examples. Much of this work was done using coordinates, although the deep connection between mechanics and geometry was already evident. Whittaker [1937] gives a good picture of the theory as it existed up to about 1910.

A highlight of this early theory was the work of Routh [1860, 1884] who studied reduction of systems with cyclic variables and introduced the amended potential for the reduced system for the purpose of studying, for instance, the *stability of a uniformly rotating state*—what we would call today a *relative equilibrium*, terminology introduced later by Poincaré.[4] Routh's work was closely related to the reduction of systems with integrals in involution studied by Jacobi and Liouville around 1870; the Routh method corresponds to the modern theory of Lagrangian reduction for the action of Abelian groups.

The rigid body, whose equations were discovered by Euler around 1740, was a key example of reduction–what we would call today either reduction to coadjoint orbits or Lie-Poisson reduction on the Hamiltonian side, or Euler-Poincaré reduction on the Lagrangian side, depending on one's point of view. Lagrange [1788] already understood reduction of the rigid body equations by a method not so far from what one would do today with the symmetry group SO(3).

Many later authors, unfortunately, relied so much on coordinates (especially Euler angles) that there is little mention of SO(3) in classical mechanics books written before 1990, which by today's standards, seems rather surprising! In addition, there seemed to be little appreciation until recently for the role of topological notions; for example, the fact that one cannot globally split off cyclic variables for the S^1 action on the configuration space of the heavy top. The Hopf fibration was patiently waiting to be discovered in the reduction theory for the classical rigid body, but it was only explicitly found later on by H. Hopf [1931]. Hopf was, apparently, unaware that this example is of great mechanical interest—the gap between workers in mechanics and geometers seems to have been particularly wide at that time.

Another noteworthy instance of reduction is Jacobi's elimination of the node for reducing the gravitational (or electrostatic) n-body problem by means of the group SE(3) of Euclidean motions, around 1860 or so. This example has, of course, been a mainstay of celestial mechanics. It is related to the work done by Riemann, Jacobi, Poincaré and others on rotating fluid masses held together by gravitational forces, such as stars. Hidden in these examples is much of the beauty of modern reduction, stability and bifurcation theory for mechanical systems with symmetry.

While both symplectic and Poisson geometry have their roots in the work of Lagrange and Jacobi, it matured considerably with the work of Lie [1890], who discovered many remarkably modern concepts such as the Lie-Poisson bracket on the dual of a Lie algebra.[5] How Lie could have viewed his wonderful discoveries so divorced from their roots in mechanics remains

---

[4]Smale [1970] eventually put the amended potential into a nice geometric setting.
[5]See Weinstein [1983b] and Marsden and Ratiu [1999] for more details on the history.

a mystery. We can only guess that he was inspired by Jacobi, Lagrange and Riemann and then, as mathematicians often do, he quickly abstracted the ideas, losing valuable scientific and historical connections along the way.

As we have already hinted, it was the famous paper Poincaré [1901a] where we find what we call today the Euler-Poincaré equations–a generalization of the Euler equations for both fluids and the rigid body to general Lie algebras. (The Euler-Poincaré equations are treated in detail in [MandS]). It is curious that Poincaré did not stress either the symplectic ideas of Lie, nor the variational principles of mechanics of Lagrange and Hamilton—in fact, it is not clear to what extent he understood what we would call today Euler-Poincaré reduction. It was only with the development and physical application of the notion of a manifold, pioneered by Lie, Poincaré, Weyl, Cartan, Reeb, Synge and many others, that a more general and intrinsic view of mechanics was possible. By the late 1950's, the stage was set for an explosion in the field.

1960-1972.   Beginning in the 1960's, the subject of geometric mechanics indeed did explode with the basic contributions of people such as (alphabetically and nonexhaustively) Abraham, Arnold, Kirillov, Kostant, Mackey, MacLane, Segal, Sternberg, Smale, and Souriau. Kirillov and Kostant found deep connections between mechanics and pure mathematics in their work on the orbit method in group representations, while Arnold, Smale, and Souriau were in closer touch with mechanics.

The modern vision of geometric mechanics combines strong links to important questions in mathematics with the traditional classical mechanics of particles, rigid bodies, fields, fluids, plasmas, and elastic solids, as well as quantum and relativistic theories. Symmetries in these theories vary from obvious translational and rotational symmetries to less obvious particle relabeling symmetries in fluids and plasmas, to the "hidden" symmetries underlying integrable systems. As we have already mentioned, reduction theory concerns the removal of variables using symmetries and their associated conservation laws. Variational principles, in addition to symplectic and Poisson geometry, provide fundamental tools for this endeavor. In fact, conservation of the momentum map associated with a symmetry group action is a geometric expression of the classical Noether theorem (discovered by variational, not symplectic methods).

Arnold and Smale.   The modern era of reduction theory began with the fundamental papers of Arnold [1966a] and Smale [1970]. Arnold focused on systems whose configuration manifold is a Lie group, while Smale focused on bifurcations of relative equilibria. Both Arnold and Smale linked their theory strongly with examples. For Arnold, they were the same examples as for Poincaré, namely the rigid body and fluids, for which he went on to develop powerful stability methods, as in Arnold [1969].

With hindsight, we can say that Arnold [1966a] was picking up on the basic work of Poincaré for both rigid body motion and fluids. In the case of

fluids, G is the group of (volume preserving) diffeomorphisms of a compact manifold (possibly with boundary). In this setting, one obtains the Euler equations for (incompressible) fluids by reduction from the Lagrangian formulation of the equations of motion, an idea exploited by Arnold [1966a] and Ebin and Marsden [1970]. This sort of description of a fluid goes back to Poincaré (using the Euler-Poincaré equations) and to the thesis of Ehrenfest (as geodesics on the diffeomorphism group), written under the direction of Boltzmann.

For Smale, the motivating example was celestial mechanics, especially the study of the number and stability of relative equilibria by a topological study of the energy-momentum mapping. He gave an intrinsic geometric account of the amended potential and in doing so, discovered what later became known as the mechanical connection. (Smale appears to not to have recognized that the interesting object he called α is, in fact, a *principal connection*; this was first observed by Kummer [1981]). One of Smale's key ideas in studying relative equilibria was to link mechanics with topology via the fact that relative equilibria are critical points of the amended potential.

Besides giving a beautiful exposition of the momentum map, Smale also emphasized the connection between singularities and symmetry, observing that the symmetry group of a phase space point has positive dimension if and only if that point is not a regular point of the momentum map restricted to a fiber of the cotangent bundle (Smale [1970], Proposition 6.2)—a result we have proved in Proposition 1.1.2. He went on from here to develop his topology and mechanics program and to apply it to the planar n-body problem. The topology and mechanics program definitely involved reduction ideas, as in Smale's construction of the quotients of integral manifolds, as in $I_{c,p}/S^1$ (Smale [1970], page 320). He also understood Jacobi's elimination of the node in this context, although he did not attempt to give any general theory of reduction along these lines.

Smale thereby set the stage for symplectic reduction: he realized the importance of the momentum map and of quotient constructions, and he worked out explicit examples like the planar n-body problem with its S^1 symmetry group. (Interestingly, he pointed out that one should really use the nonabelian group SE(2); his feeling of unease with fixing the center of mass of an n-body system is remarkably perceptive.)

Synthesis.   The problem of synthesizing the Lie algebra reduction methods of Arnold [1966a] with the techniques of Smale [1970] on the reduction of cotangent bundles by Abelian groups, led to the development of reduction theory in the general context of symplectic manifolds and equivariant momentum maps in Marsden and Weinstein [1974] and Meyer [1973], as we described in the last section. Both of these papers were completed by 1972.

Poisson Manifolds.   Meanwhile, things were also gestating from the viewpoint of Poisson brackets and the idea of a Poisson manifold was being initiated and developed, with much duplication and rediscovery (see [MandS] Section 10.1 for additional information).

A basic example of a noncanonical Poisson bracket is the ***Lie-Poisson bracket*** on $\mathfrak{g}^*$, the dual of a Lie algebra $\mathfrak{g}$. This bracket (which comes with a plus or minus sign) is given on two smooth functions on $\mathfrak{g}^*$ by

$$\{f,g\}_{\pm}(\mu) = \pm \left\langle \mu, \left[\frac{\delta f}{\delta \mu}, \frac{\delta g}{\delta \mu} \right] \right\rangle, \qquad (1.3.1)$$

where $\delta f / \delta \mu$ is the derivative of f, but thought of as an element of $\mathfrak{g}$. These Poisson structures, including the coadjoint orbits as their symplectic leaves, were known to Lie [1890], although, as we mentioned previously, Lie does not seem to have recognized their importance in mechanics. It is also not clear whether or not Lie realized that the Lie Poisson bracket is the Poisson reduction of the canonical Poisson bracket on T^*G by the action of G. (See [MandS], Chapter 13 for an account of this theory). The first place we know of that has this clearly stated (but with no references, and no discussion of the context) is Bourbaki [1971], Chapter III, Section 4, Exercise 6. Remarkably, this exercise also contains an interesting proof of the Duflo-Vergne theorem (with no reference to the original paper, which appeared in 1969). Again, any hint of links with mechanics is missing.

This takes us up to about 1972.

Post 1972.   An important contribution was made by Marle [1976], who divides the inverse image of an orbit by its characteristic foliation to obtain the product of an orbit and a reduced manifold. In particular, as we saw in Theorem 1.2.4, P_μ is symplectically diffeomorphic to an "orbit-reduced" space $P_\mu \cong J^{-1}(\mathcal{O}_\mu)/G$, where $\mathcal{O}_\mu$ is a coadjoint orbit of G. From this it follows that the P_μ are symplectic leaves in the Poisson space P/G. The related paper of Kazhdan, Kostant and Sternberg [1978] was one of the first to notice deep links between reduction and integrable systems. In particular, they found that the Calogero-Moser systems could be obtained by reducing a system that was trivially integrable; in this way, reduction provided a method of producing an interesting integrable system from a simple one. This point of view was used again by, for example, Bobenko, Reyman and Semenov-Tian-Shansky [1989] in their spectacular group theoretic explanation of the integrability of the Kowalewski top.

Noncanonical Poisson Brackets.   The Hamiltonian description of many physical systems, such as rigid bodies and fluids in Eulerian variables, requires noncanonical Poisson brackets and constrained variational principles of the sort studied by Lie and Poincaré. As discussed above, a basic example of a noncanonical Poisson bracket is the ***Lie-Poisson bracket*** on the dual of a Lie algebra. From the mechanics perspective, the remarkably

modern book (but which was, unfortunately, rather out of touch with the corresponding mathematical developments) by Sudarshan and Mukunda [1974] showed via explicit examples how systems such as the rigid body could be written in terms of noncanonical brackets, an idea going back to Pauli [1953], Martin [1959] and Nambu [1973]. Others in the physics community, such as Morrison and Greene [1980] also discovered noncanonical bracket formalisms for fluid and magnetohydrodynamic systems. In the 1980's, many fluid and plasma systems were shown to have a noncanonical Poisson formulation. It was Marsden and Weinstein [1982, 1983] who first applied reduction techniques to these systems.

The *reduction philosophy* concerning noncanonical brackets can be summarized by saying

> *Any mechanical system has its roots somewhere as a cotangent bundle and one can recover noncanonical brackets by the simple process of Poisson reduction. For example, in fluid mechanics, this reduction is implemented by the Lagrange-to-Euler map.*

This view ran contrary to the point of view, taken by some researchers, that one should guess at what a Poisson structure might be and then to try to limit the guesses by the constraint of Jacobi's identity.

In the simplest version of the Poisson reduction process, one starts with a Poisson manifold P on which a group G acts by Poisson maps and then forms the quotient space P/G, which, if not singular, inherits a natural Poisson structure itself. Of course, the Lie-Poisson structure on $\mathfrak{g}^*$ is inherited in exactly this way from the canonical symplectic structure on T^*G. One of the attractions of this Poisson bracket formalism was its use in stability theory. This literature is now very large, but Holm, Marsden, Ratiu and Weinstein [1985] is representative.

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

Reduction theory for mechanical systems with symmetry has proven to be a powerful tool that has enabled key advances in stability theory (from the Arnold method to the energy-momentum method for relative equilibria) as well as in bifurcation theory of mechanical systems, geometric phases via reconstruction—the inverse of reduction—as well as uses in control theory from stabilization results to a deeper understanding of locomotion. For a general introduction to some of these ideas and for further references, see Marsden, Montgomery and Ratiu [1990]; Simo, Lewis and Marsden [1991]; Marsden and Ostrowski [1996]; Marsden and Ratiu [1999]; Montgomery [1988, 1990, 1991a,b, 1993]; Blaom [2000, 2001], and Kanso, Marsden, Rowley, and Melli-Huber [2005].

Tangent and Cotangent Bundle Reduction. The simplest case of cotangent bundle reduction is the case of reduction of $P = T^*Q$ at $\mu = 0$; the answer is simply $P_0 = T^*(Q/G)$ with the canonical symplectic form. Another basic case is when G is Abelian. Here, $(T^*Q)_\mu \cong T^*(Q/G)$, but the latter has a symplectic structure modified by magnetic terms, that is, by the curvature of the mechanical connection.

An Abelian version of cotangent bundle reduction was developed by Smale [1970]. Then Satzer [1977] studied the relatively simple, but important case of cotangent bundle reduction at the zero value of the momentum map. The full generalization of cotangent bundle reduction for nonabelian groups at arbitrary values of the momentum map appears for the first time in Abraham and Marsden [1978]. It was Kummer [1981] who first interpreted this result in terms of a connection, now called the **mechanical connection**. The geometry of this situation was used to great effect in, for example, Guichardet [1984], Iwai [1987, 1990], and Montgomery [1984, 1990, 1991a]. We give a detailed account of cotangent bundle reduction theory in the following chapter, with some generalizations important for reduction by stages in Chapter 7.

The Gauge Theory Viewpoint. Tangent and cotangent bundle reduction evolved into what we now term as the "bundle picture" or the "gauge theory of mechanics". This picture was first developed by Montgomery, Marsden and Ratiu [1984] and Montgomery [1984, 1986]. That work was motivated and influenced by the work of Sternberg [1977] and Weinstein [1978a] on a "Yang-Mills construction" which is, in turn, motivated by Wong's equations, i.e., the equations for a particle moving in a Yang-Mills field. The main result of the bundle picture gives a structure to the quotient spaces $(T^*Q)/G$ and $(TQ)/G$ when G acts by the cotangent and tangent lifted actions. The symplectic leaves in this picture were analyzed by Zaalani [1999], Cushman and Śniatycki [1999], and Marsden and Perlmutter [2000]. The work of Perlmutter and Ratiu [2005] gives a unified study of the Poisson bracket on $(T^*Q)/G$ in both the Sternberg and Weinstein realizations of the quotient.

As mentioned earlier, we shall review some of the basics of cotangent bundle reduction theory in §2.2 and §2.3 along with some of the needed material from the theory of principal connections in §2.1. Further information on this theory may be found in [LonM] and [FofM] as well as a number of the other references mentioned above.

Lagrangian Reduction. A key ingredient in Lagrangian reduction is the classical work of Poincaré [1901a] in which the Euler–Poincaré equations were introduced. Poincaré realized that the equations of fluids, free rigid bodies, and heavy tops could all be described in Lie algebraic terms in a beautiful way. The importance of these equations was realized by Hamel [1904, 1949] and Chetayev [1941], but to a large extent, the work of Poincaré lay dormant until it was revived in the Russian literature in the 1980's.

The more recent developments of Lagrangian reduction were motivated by attempts to understand the relation between reduction, variational principles and Clebsch variables in Cendra and Marsden [1987] and Cendra, Ibort, and Marsden [1987]. In Marsden and Scheurle [1993b] it was shown that, for matrix groups, one could view the Euler-Poincaré equations via the reduction of Hamilton's variational principle from TG to $\mathfrak{g}$. The work of Bloch, Krishnaprasad, Marsden and Ratiu [1996] established the Euler–Poincaré variational structure for general Lie groups.

The paper of Marsden and Scheurle [1993b] also considered the case of more general configuration spaces Q on which a group G acts, which was motivated by both the Euler-Poincaré case as well as the work of Cendra and Marsden [1987] and Cendra, Ibort, and Marsden [1987]. The Euler-Poincaré equations correspond to the case $Q = G$. Related ideas stressing the groupoid point of view were given in Weinstein [1996]. The resulting reduced equations were called the *reduced Euler–Lagrange equations*. This work is the Lagrangian analogue of Poisson reduction, in the sense that no momentum map constraint is imposed.

Lagrangian reduction proceeds in a way that is very much in the spirit of the gauge theoretic point of view of mechanical systems with symmetry. It starts with Hamilton's variational principle for a Lagrangian system on a configuration manifold Q and with a symmetry group G acting on Q. The idea is to drop this variational principle to the quotient Q/G to derive a reduced variational principle. This theory has its origins in specific examples such as fluid mechanics (see, for example, Arnold [1966b] and Bretherton [1970]), while the systematic theory of Lagrangian reduction was begun in Marsden and Scheurle [1993b] and further developed in Cendra, Marsden, and Ratiu [2001a]. The latter reference also introduced a connection to realize the space $(TQ)/G$ as the fiber product $T(Q/G) \times \tilde{\mathfrak{g}}$ of $T(Q/G)$ with the associated bundle formed using the adjoint action of G on $\mathfrak{g}$. The reduced equations associated to this construction are called the *Lagrange–Poincaré equations* and their geometry has been fairly well developed. Note that a G-invariant Lagrangian L on TQ induces a Lagrangian l on $(TQ)/G$.

Until recently, the Lagrangian side of the reduction story had lacked a general category that is the Lagrangian analogue of Poisson manifolds in which reduction can be repeated. One candidate is the category of Lie algebroids, as explained in Weinstein [1996]. Another is that of *Lagrange-Poincaré bundles*, developed in Cendra, Marsden, and Ratiu [2001a]. Both have tangent bundles and Lie algebras as basic examples. The latter work also develops the Lagrangian analogue of reduction for central extensions and, as in the case of symplectic reduction by stages, cocycles and curvatures enter in a natural way.

This bundle picture and Lagrangian reduction has proven very useful in control and optimal control problems. For example, it was used in Chang, Bloch, Leonard, Marsden, and Woolsey [2002] to develop a Lagrangian and Hamiltonian reduction theory for controlled mechanical systems and in

Koon and Marsden [1997] to extend the falling cat theorem of Montgomery [1990] to the case of nonholonomic systems as well as to nonzero values of the momentum map.

Finally we mention that the paper Cendra, Marsden, Pekarsky, and Ratiu [2003] develops the reduction theory for Hamilton's *phase space principle* and the equations on the reduced space, along with a reduced variational principle, are developed and called *the Hamilton-Poincaré equations.* Even in the case $Q = G$, this collapses to an interesting variational principle for the Lie-Poisson equations on $\mathfrak{g}^*$.

Legendre Transformation.   Of course the Lagrangian and Hamiltonian sides of the reduction story are linked by the Legendre transformation. This mapping descends at the appropriate points to give relations between the Lagrangian and the Hamiltonian sides of the theory. However, even in standard cases such as the heavy top, one must be careful with this approach, as is already explained in, for example, Holm, Marsden and Ratiu [1998]. For field theories, such as the Maxwell-Vlasov equations, this issues is also important, as explained in Cendra, Holm, Hoyle and Marsden [1998] (see also Tulczyjew and Urbański [1999]).

Nonabelian Routh Reduction.   Routh reduction for Lagrangian systems, which goes back Routh [1860, 1877, 1884] is classically associated with systems having cyclic variables (this is almost synonymous with having an Abelian symmetry group). Modern expositions of this classical theory can be found in Arnold, Kozlov and Neishtadt [1988] and in [MandS], §8.9. Routh Reduction may be thought of as the Lagrangian analog of symplectic reduction in that a momentum map is set equal to a constant. A key feature of Routh reduction is that when one drops the Euler–Lagrange equations to the quotient space associated with the symmetry, and when the momentum map is constrained to a specified value (i.e., when the cyclic variables and their velocities are eliminated using the given value of the momentum), then the resulting equations are in Euler–Lagrange form not with respect to the Lagrangian itself, but with respect to a modified function called the *Routhian*.

Routh [1877] applied his method to stability theory; this was a precursor to the energy-momentum method for stability that synthesizes Arnold's and Routh's methods (see Simo, Lewis and Marsden [1991]). Routh's stability method is still widely used in mechanics.

The initial work on generalizing Routh reduction to the nonabelian case was that of Marsden and Scheurle [1993a]. This subject was further developed in Jalnapurkar and Marsden [2000] and Marsden, Ratiu and Scheurle [2000]. The latter reference used this theory to give some nice formulas for geometric phases from the Lagrangian point of view.

Semidirect Product Reduction.   We shall study the case of reduction by a semidirect product Lie group in some detail in Chapter 4. In the

simplest case of a semidirect product, one has a Lie group G that acts on a vector space V (and hence on its dual V^*) and then one forms the semidirect product $S = G \,\circledS\, V$, generalizing the semidirect product structure of the Euclidean group $\mathrm{SE}(3) = \mathrm{SO}(3) \,\circledS\, \mathbb{R}^3$.

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

The Lagrangian reduction analog of semidirect product theory was developed by Holm, Marsden and Ratiu [1998, 2002]. This construction is used in applications where one has advected quantities (such as the direction of gravity in the heavy top, density in compressible fluids and the magnetic field in MHD) as well as to geophysical flows. Cendra, Holm, Hoyle and Marsden [1998] applied this idea to the Maxwell–Vlasov equations of plasma physics. Cendra, Holm, Marsden and Ratiu [1998] showed how Lagrangian semidirect product theory fits into the general framework of Lagrangian reduction.

The semidirect product reduction theorem has been proved in Landsman [1995], Landsman [1998, Chapter 4] as an application of a stages theorem for his *special symplectic reduction method*. Even though special symplectic reduction generalizes Marsden-Weinstein reduction, the special reduction by stages theorem in Landsman [1995] studies a setup that, in general, is different to the ones in the reduction by stages theorems of this book.

Singular reduction. Singular reduction starts with the observation of Smale [1970] that we have already mentioned: $z \in P$ is a regular point of a momentum map **J** if and only if z has no continuous isotropy. Motivated by this, Arms, Marsden and Moncrief [1981, 1982] showed that (under hypotheses which include the ellipticity of certain operators and which can be interpreted more or less, as playing the role of a properness assumption on the group action in the finite dimensional case) the level sets $\mathbf{J}^{-1}(0)$ of an equivariant momentum map **J** have quadratic singularities at points with continuous symmetry. While such a result is easy to prove for compact group actions on finite dimensional manifolds (using the equivariant Darboux theorem), the main examples of Arms, Marsden and Moncrief [1981] were, in fact, infinite dimensional—both the phase space and the group. Singular points in the level sets of the momentum map are related to convexity properties of the momentum map in that the singular points in phase space map to corresponding singular points in the the image polytope.

The paper of Otto [1987] showed that if G is a Lie group acting properly on an almost Kähler manifold then the orbit space $\mathbf{J}^{-1}(\mu)/G_\mu$ decomposes into symplectic smooth manifolds constructed out of the orbit types of the

G-action on P. In some related work, Huebschmann [1998] has made a careful study of the singularities of moduli spaces of flat connections.

The detailed structure of $\mathbf{J}^{-1}(0)/G$ for compact Lie groups acting on finite dimensional manifolds was determined by Sjamaar and Lerman [1991]; their work was extended to proper Lie group actions and to $\mathbf{J}^{-1}(\mathcal{O}_\mu)/G$ by Bates and Lerman [1997], with the assumption that $\mathcal{O}_\mu$ be locally closed in $\mathfrak{g}^*$. Ortega [1998] and [HRed] redid the entire singular reduction theory for proper Lie group actions starting with the point reduced spaces $\mathbf{J}^{-1}(\mu)/G_\mu$ and also connected it to the more algebraic approach of Arms, Cushman, and Gotay [1991]. Specific examples of singular reduction, with further references, may be found in Lerman, Montgomery, and Sjamaar [1993] and Cushman and Bates [1997]. One of these, the "canoe" will be given in §2.4. In fact, this is an example of singular reduction in the case of cotangent bundles, and much more can be said in this case, as we shall see in §2.4. Another approach to singular reduction based on the technique of blowing up singularities, and which was also designed for the case of singular cotangent bundle reduction, was started in Hernandez and Marsden [2005] and Birtea, Puta, Ratiu, and Tudoran [2005], a technique which requires further development.

Singular reduction has been extensively used in the study of the persistence, bifurcation, and stability of relative dynamical elements; see Chossat, Lewis, Ortega, and Ratiu [2003]; Chossat, Ortega, and Ratiu [2002]; Grabsi, Montaldi, and Ortega [2004]; Lerman and Singer [1998]; Lerman and Tokieda [1999]; Ortega [2003b]; Ortega and Ratiu [1997, 1999a], Ortega and Ratiu [1999b, 2004b]; Patrick, Roberts, and Wulff [2004]; Roberts and de Sousa Dias [1997]; Roberts, Wulff, and Lamb [2002]; Wulff and Roberts [2002], and Wulff [2003].

Symplectic Reduction Without Momentum Maps.   The reduction theory presented so far needs the existence of a momentum map. However, more primitive versions of this procedure based on foliation theory (see Cartan [1922] and Meyer [1973]) do not require the existence of this object. Working in this direction, but with a mathematical program that goes beyond the reduction problem, Condevaux, Dazord, and P. Molino [1988] introduced a concept that generalizes the momentum map. This object is defined via a connection that associates an additive holonomy group to each canonical action on a symplectic manifold. The existence of the momentum map is equivalent to the vanishing of this group. Symplectic reduction has been carried out using this generalized momentum map in Ortega and Ratiu [2006a] and Ortega and Ratiu [2006b].

Another approach to symplectic reduction that is able to avoid the possible non-existence of the momentum map is based on the optimal momentum map introduced and studied in Ortega and Ratiu [2002] and Ortega [2002]. This distribution theoretical approach can also deal with reduction of Poisson manifolds, where the standard momentum map does not exist

generically. This point of view is presented in detail in Part III of this book and fully exploited in handling singular reduction by stages.

Reduction of Other Geometric Structures.   Besides symplectic reduction, there are many other geometric structures on which one can perform similar constructions. For example, one can reduce Kähler, hyper-Kähler, Poisson, contact, Jacobi, etc. manifolds and this can be done either in the regular or singular cases. We refer to [HRed] for a survey of the literature for these topics.

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

Nonholonomic Systems.   Nonholonomic mechanical systems (such as systems with rolling constraints) provide a very interesting class of systems where the reduction procedure has to be modified. In fact this provides a class of systems that gives rise to an almost Poisson structure, i.e. a bracket which does not necessarily satisfy the Jacobi identity. Reduction theory for nonholonomic systems has made a lot of progress, but many interesting questions still remain. In these types of systems, there is a natural notion of a momentum map, but in general it is not conserved, but rather obeys a *momentum equation* as was discovered by Bloch, Krishnaprasad, Marsden, and Murray [1996]. This means, in particular, that point reduction in such a situation may not be appropriate. Nevertheless, Poisson reduction in the almost Poisson and almost symplectic setting is interesting and from the mathematical point of view, point reduction is also interesting, although, as remarked, one has to be cautious with how it is applied to, for example, nonholonomic systems. A few references are Koiller [1992], Bates and Śniatycki [1993], Bloch, Krishnaprasad, Marsden, and Murray [1996], Koon and Marsden [1998], Blankenstein and van der Schaft [2001], Cushman and Śniatycki [2002], Planas-Bielsa [2004], and Ortega and Planas [2004]. We refer to Cendra, Marsden, and Ratiu [2001b] and Bloch [2003] for a more detailed historical review.

Multisymplectic Reduction.   Reduction theory is by no means completed. For example, for PDE's, the multisymplectic (as opposed to symplectic) framework seems appropriate, both for relativistic and nonrelativistic systems. In fact, this approach has experienced somewhat of a revival since it has been realized that it is rather useful for numerical computation (see Marsden, Patrick, and Shkoller [1998]). Only a few instances and

examples of multisymplectic and multi-Poisson reduction are really well understood (see Marsden, Montgomery, Morrison and Thompson [1986]; Castrillón López, Ratiu, and Shkoller [2000]; Castrillón López, Garcia Pérez, and Ratiu [2001]; Castrillón López and Ratiu [2003]; Castrillón López and Marsden [2003]), so one can expect to see more activity in this area as well.

Discrete Mechanical Systems.   Another emerging area, also motivated by numerical analysis, is that of discrete mechanics. Here the idea is to replace the velocity phase space TQ by $Q \times Q$, with the role of a velocity vector played by a pair of nearby points. This has been a powerful tool for numerical analysis, reproducing standard symplectic integration algorithms and much more. See, for example, Kane, Marsden, Ortiz, and West [2000]; Marsden and West [2001]; Lew, Marsden, Ortiz, and West [2004] for recent articles. This subject, too, has its own reduction theory. See Marsden, Pekarsky, and Shkoller [1999], Bobenko and Suris [1999] and Jalnapurkar, Leok, Marsden and West [2006]. Discrete mechanics also has some intriguing links with quantization, since Feynman himself first defined path integrals through a limiting process using the sort of discretization used in the discrete action principle (see Feynman and Hibbs [1965]).

1.4   Overview of Singular Symplectic Reduction

The preceding section gave a brief history of singular reduction and further historical information and references can be found in [HRed].

  This section gives an overview, without proofs, of some of the theory of singular symplectic reduction in the context of symplectic manifolds and from the viewpoint of [HRed]. It reviews only the minimal material needed for an understanding of the developments in Part III. The detailed theory of singular reduction can be found in [HRed] or in the original papers cited therein. *The material of this section is not needed for any developments in Parts I and II.* It is, however, necessary for understanding the results in Part III, where singular reduction by stages is carried out. An overview of singular reduction for the case of cotangent bundles is given in §2.4 and in that section an exposition of a basic example, namely the "canoe" (singular reduction for the phase space of the spherical pendulum with its S^1 action) is given.

The Setting for Singular Symplectic Reduction.   Up to this point, symplectic reduction was carried out for the *regular case*; that is, for free and proper actions. Recall that freeness of the action guarantees that the associated momentum map $\mathbf{J}$ is a submersion and hence that the level sets $\mathbf{J}^{-1}(\mu)$ are smooth manifolds. Freeness and properness also ensure that the point symplectic reduced spaces, namely the orbit spaces

$M_\mu := \mathbf{J}^{-1}(\mu)/G_\mu$, are regular quotient manifolds. Singular reduction describes the quotient topological spaces M_μ when the Lie group action is not free. As will be seen below, the main result reviewed in this section states that these quotients are symplectic Whitney stratified spaces in the sense that the strata are symplectic manifolds in a natural way. In addition, the local properties of this Whitney stratification make it into what is called a cone space. In essence, this is the rough statement of the Symplectic Stratification Theorem to be explained below. It adapts to the symplectic symmetric context, the classical stratification theorem of the orbit space of a proper Lie group action by using its orbit type manifolds.

The section begins with a quick review of the necessary definitions and results concerning stratified spaces before moving on to singular symplectic reduction itself.

Stratified Spaces.   This paragraph gives a quick overview of stratified spaces; the basic reference is Pflaum [2001].

A subset S of a topological space P is said to be **locally closed**, if S is open in $\bar{S}$, where $\bar{S}$ denotes the usual topological closure of S in P. Recall that any submanifold is locally closed in the ambient manifold. Conversely, if Q is a manifold and $f : Q \to P$ is an injective immersion such that $f(Q)$ is locally closed in P, then $f(Q)$ is an embedded submanifold of P.

Let $\mathcal{Z}$ be a locally finite partition of a topological space P into smooth manifolds $S_i \subset P$, $i \in I$. We assume that the manifolds $S_i \subset P$, $i \in I$, with their manifold topology are locally closed topological subspaces of P. The pair $(P, \mathcal{Z})$ is referred to as a **decomposition** of P with **pieces** in $\mathcal{Z}$ when the following condition is satisfied:

(DS) If $R, S \in \mathcal{Z}$ are such that $R \cap \bar{S} \neq \varnothing$, then $R \subset \bar{S}$. In this case we write $R \preceq S$. If, in addition, $R \neq S$ we say that R is **incident** to S or that it is a **boundary piece** of S and write $R \prec S$.

The **boundary** of a piece S is defined as $\partial S := \bar{S} \setminus S$. Note that this is *not* the usual topological closure of S which equals $\bar{S} \cap \overline{(P \setminus S)}$. Thus, $R \prec S$ if and only if $R \subset \partial S$. Condition **(DS)** is called the **frontier condition** and the pair $(P, \mathcal{Z})$ is called a **decomposed space**. The **dimension** of P is defined as $\dim P = \sup\{\dim S_i \mid S_i \in \mathcal{Z}\}$. If $k \in \mathbb{N}$, the k-**skeleton** P^k of P is the union of all the pieces of dimension smaller than or equal to k; its topology is the relative topology induced by P. The **depth** $\mathrm{dp}(z)$ of any $z \in (P, \mathcal{Z})$ is defined as

$$\mathrm{dp}(z) := \sup\{k \in \mathbb{N} \mid \exists\, S_0, S_1, \ldots, S_k \in \mathcal{Z} \text{ with } z \in S_0 \prec S_1 \prec \ldots \prec S_k\}.$$

Since for any two elements $x, y \in \mathcal{S}$ in the same piece $\mathcal{S} \in P$ we have $\mathrm{dp}(x) = \mathrm{dp}(y)$, the depth $\mathrm{dp}(S)$ of the piece S is well defined by $\mathrm{dp}(S) := \mathrm{dp}(x)$, $x \in \mathcal{S}$. Finally, the depth $\mathrm{dp}(P)$ of $(P, \mathcal{Z})$ is defined by $\mathrm{dp}(P) := \sup\{\mathrm{dp}(S) \mid S \in \mathcal{Z}\}$.

A continuous mapping $f : P \to Q$ between the decomposed spaces $(P, \mathcal{Z})$ and $(Q, \mathcal{Y})$ is a ***morphism of decomposed spaces*** if for every piece $S \in \mathcal{Z}$, there is a piece $T \in \mathcal{Y}$ such that $f(S) \subset T$ and the restriction $f|_S : S \to T$ is smooth. If $(P, \mathcal{Z})$ and $(P, \mathcal{T})$ are two decompositions of the same topological space we say that $\mathcal{Z}$ is ***coarser*** than $\mathcal{T}$ or that $\mathcal{T}$ is ***finer*** than $\mathcal{Z}$ if the identity mapping $(P, \mathcal{T}) \to (P, \mathcal{Z})$ is a morphism of decomposed spaces. A topological subspace $Q \subset P$ is a ***decomposed subspace*** of $(P, \mathcal{Z})$ if for all pieces $S \in \mathcal{Z}$, the intersection $S \cap Q$ is a submanifold of S and the corresponding partition $\mathcal{Z} \cap Q$ forms a decomposition of Q.

Let P be a topological space and $z \in P$. Two subsets A and B of P are said to be ***equivalent*** at z if there is an open neighborhood U of z such that $A \cap U = B \cap U$. This relation constitutes an equivalence relation on the power set of P. The class of all sets equivalent to a given subset A at z will be denoted by $[A]_z$ and called the ***set germ*** of A at z. If $A \subset B \subset P$ we say that $[A]_z$ is a ***subgerm*** of $[B]_z$, and denote $[A]_z \subset [B]_z$.

A ***stratification*** of the topological space P is a map $\mathcal{S}$ that associates to any $z \in P$ the set germ $\mathcal{S}(z)$ of a closed subset of P such that the following condition is satisfied:

(ST) For every $z \in P$ there is a neighborhood U of z and a decomposition $\mathcal{Z}$ of U such that for all $y \in U$ the germ $\mathcal{S}(y)$ coincides with the set germ of the piece of $\mathcal{Z}$ that contains y.

The pair $(P, \mathcal{S})$ is called a ***stratified space***. Any decomposition of P defines a stratification of P by associating to each of its points the set germ of the piece in which it is contained. The converse is, by definition, locally true.

The Strata.   Two decompositions $\mathcal{Z}_1$ and $\mathcal{Z}_2$ of P are said to be ***equivalent*** if they induce the same stratification of P. If $\mathcal{Z}_1$ and $\mathcal{Z}_2$ are equivalent decompositions of P then, for all $z \in P$, we have that $\mathrm{dp}_{\mathcal{Z}_1}(z) = \mathrm{dp}_{\mathcal{Z}_2}(z)$. Any stratified space $(P, \mathcal{S})$ has a unique decomposition $\mathcal{Z}_{\mathcal{S}}$ associated with the following maximality property: for any open subset $U \subset P$ and any decomposition $\mathcal{Z}$ of P inducing $\mathcal{S}$ over U, the restriction of $\mathcal{Z}_{\mathcal{S}}$ to U is coarser than the restriction of $\mathcal{Z}$ to U. The decomposition $\mathcal{Z}_{\mathcal{S}}$ is called the ***canonical decomposition*** associated to the stratification $(P, \mathcal{S})$. It is often denoted by $\mathcal{S}$ and its pieces are called the ***strata*** of P. The local finiteness of the decomposition $\mathcal{Z}_{\mathcal{S}}$ implies that for any stratum S of $(P, \mathcal{S})$ there are only finitely many strata R with $S \prec R$. In the sequel, the symbol $\mathcal{S}$ in the stratification $(P, \mathcal{S})$ will denote both the map that associates to each point a set germ and the set of pieces associated to the canonical decomposition induced by the stratification of P.

A ***stratified map*** $f : (P, \mathcal{S}_P) \to (Q, \mathcal{S}_Q)$ is a morphism of decomposed spaces relative to the canonical decompositions $\mathcal{S}_P$ on P and $\mathcal{S}_Q$ on Q.

Stratified Spaces with Smooth Structure.   Let $(P, \mathcal{S})$ be a stratified space. A ***singular*** or ***stratified chart*** of P is a homeomorphism $\phi : U \to \phi(U) \subset \mathbb{R}^n$ from an open set $U \subset P$ to a subset of $\mathbb{R}^n$ such that for

every stratum $S \in \mathcal{S}$ the image $\phi(U \cap S)$ is a submanifold of $\mathbb{R}^n$ and the restriction $\phi|_{U \cap S} : U \cap S \to \phi(U \cap S)$ is a diffeomorphism. Two singular charts $\phi : U \to \phi(U) \subset \mathbb{R}^n$ and $\varphi : V \to \varphi(V) \subset \mathbb{R}^m$ are **compatible** if for any $z \in U \cap V$ there exist an open neighborhood $W \subset U \cap V$ of z, a natural number $N \geq \max\{n, m\}$, open neighborhoods $O, O' \subset \mathbb{R}^N$ of $\phi(U) \times \{0\}$ and $\varphi(V) \times \{0\}$, respectively, and a diffeomorphism $\psi : O \to O'$ such that $i_m \circ \varphi|_W = \psi \circ i_n \circ \phi|_W$, where i_n and i_m denote the natural embeddings of $\mathbb{R}^n$ and $\mathbb{R}^m$ into $\mathbb{R}^N$ by using the first n and m coordinates, respectively. The notion of **singular** or **stratified atlas** is the natural generalization for stratifications of the concept of atlas existing for smooth manifolds. Analogously, one introduces compatible and maximal stratified atlases. If the stratified space $(P, \mathcal{S})$ has a well defined maximal atlas, then we say that this atlas determines a **smooth** or **differentiable structure** on P. The pair $(P, \mathcal{S})$ is called a **smooth stratified space**.

The Whitney Condition (B). Let M be a manifold and $R, S \subset M$ two submanifolds. Let $\phi : U \to \mathbb{R}^n$ be a smooth chart of M around the point z. The **Whitney condition (B)** at the point $z \in R$ with respect to the chart (U, ϕ) is given by the following statement:

(B) Let $\{x_n\}_{n \in \mathbb{N}} \subset R \cap U$ and $\{y_n\}_{n \in \mathbb{N}} \subset S \cap U$ be two sequences with the same limit $z = \lim\limits_{n \to \infty} x_n = \lim\limits_{n \to \infty} y_n$ and such that $x_n \neq y_n$, for all $n \in \mathbb{N}$. Suppose that the set of connecting lines $\overline{\phi(x_n)\phi(y_n)} \subset \mathbb{R}^n$ converges in projective space to a line L and that the sequence of tangent spaces $\{T_{y_n} S\}_{n \in \mathbb{N}}$ converges in the Grassmann bundle of dim S-dimensional subspaces of TM to $\tau \subset T_z M$. Then, $(T_z \phi)^{-1}(L) \subset \tau$.

If the condition (B) is verified for every point $z \in R$, the pair (R, S) is said to satisfy the Whitney condition **(B)**. It can be verified that Whitney's condition (B) does not depend on the chart used to formulate it. A stratified space with smooth structure such that for every pair of strata Whitney's condition (B) is satisfied is called a **Whitney space** or a **(B)- stratified space**.

Cone Spaces and Local Triviality. Let P be a topological space. Consider the equivalence relation $\sim$ in the product $P \times [0, \infty)$ given by $(z, a) \sim (z', a')$ if and only if $a = a' = 0$. The **cone** CP on P is defined as the quotient topological space $P \times [0, \infty)/\sim$. If P is a smooth manifold then the cone CP is a decomposed space with two pieces, namely, $P \times (0, \infty)$ and the **vertex** which is the class corresponding to any element of the form $(z, 0)$, $z \in P$, that is, $P \times \{0\}$. Analogously, if $(P, \mathcal{Z})$ is a decomposed (stratified) space then the associated cone CP is also a decomposed (stratified) space whose pieces (strata) are the vertex and the sets of the form $S \times (0, \infty)$, with $S \in \mathcal{Z}$. This implies, in particular, that $\dim CP = \dim P + 1$ and $\mathrm{dp}(CP) = \mathrm{dp}(P) + 1$.

   A stratified space $(P, \mathcal{S})$ is said to be **locally trivial** if for any $z \in P$ there exist a neighborhood U of z, a stratified space $(F, \mathcal{S}^F)$, a distinguished

point $\mathbf{0} \in F$, and an isomorphism of stratified spaces

$$\psi : U \to (S \cap U) \times F,$$

where S is the stratum that contains z and ψ satisfies $\psi^{-1}(y, \mathbf{0}) = y$, for all $y \in S \cap U$. Whenever F is given by a cone CL over a compact stratified space L then L is called the **link** of z.

An important corollary of **Thom's First Isotopy Lemma** guarantees that every Whitney stratified space is locally trivial. A converse to this implication needs the introduction of cone spaces. Their definition is given by recursion on the depth of the space.

1.4.1 Definition.   *Let $m \in \mathbb{N} \cup \{\infty, \omega\}$. A **cone space** of class C^m and depth 0 is the union of countably many C^m manifolds together with the stratification whose strata are the unions of the connected components of equal dimension. A **cone space** of class C^m and depth $d + 1$, $d \in \mathbb{N}$, is a stratified space $(P, \mathcal{S})$ with a C^m differentiable structure such that for any $z \in P$ there exists a connected neighborhood U of z, a compact cone space L of class C^m and depth d called the **link** and a stratified isomorphism*

$$\psi : U \to (S \cap U) \times CL,$$

where S is the stratum that contains the point z, the map ψ satisfies $\psi^{-1}(y, \mathbf{0}) = y$, for all $y \in S \cap U$, and $\mathbf{0}$ is the vertex of the cone CL.

*If $m \neq 0$ then L is required to be embedded into a sphere S^l via a fixed smooth global singular chart $\varphi : L \to S^l$ that determines the smooth structure of CL. More specifically, the smooth structure of CL is generated by the global chart $\tau : [z, t] \in CL \longmapsto t\varphi(z) \in \mathbb{R}^{l+1}$. The maps $\psi : U \to (S \cap U) \times CL$ and $\varphi : L \to S^l$ are called a **cone chart** and a **link chart**, respectively. Moreover, if $m \neq 0$ then ψ and ψ^{-1} are required to be differentiable of class C^m as maps between stratified spaces with a smooth structure.*

Orbit Type Stratification.   The most commonly encountered cone space appears as the quotient of a proper smooth Lie group action on a manifold. Since this stratification is also the backbone of the one encountered in symplectic singular reduction we quickly review it here. For more information, see Duistermaat and Kolk [1999] or [HRed] and references therein.

If H is a closed subgroup of a Lie group G define the conjugacy class

$$(H) := \{L \subset G \mid L = gHg^{-1}, \, g \in G\}.$$

The set of conjugacy classes of Lie subgroups of a Lie group G admits a partial order by defining $(K) \preceq (H)$ if and only if H is conjugate to a subgroup of K. We write $(K) \prec (H)$ if $(K) \preceq (H)$ and $(K) \neq (H)$.

Assume that the Lie group G acts smoothly and properly on a manifold M. Then the partially ordered subset $I(G, M)$ of conjugacy classes represented by an isotropy subgroup of the action is called the **isotropy lattice** of the G-action.

If H is a closed subgroup of G introduce the following subsets of M:

$$M_{(H)} := \{z \in M \mid G_z \in (H)\},$$
$$M^H := \{z \in M \mid H \subset G_z\},$$
$$M_H := \{z \in M \mid H = G_z\}.$$

The set $M_{(H)}$ is called the (H)-**orbit type submanifold**, M_H the H-**isotropy type submanifold**, and M^H is the H-**fixed point submanifold**. Collectively, these subsets are called **type submanifolds**. If the G-action is proper then the connected components of the type manifolds are indeed embedded submanifolds of M, but $M_{(H)}$ can have connected components of different dimensions. Moreover, under the properness assumption it follows that the connected components of the orbit type manifolds $M_{(H)}$ and their projections $M_{(H)}/G$ onto orbit space constitute a Whitney (B) stratification of M and M/G, respectively. Both stratifications are cone spaces. The stratification of M/G is minimal among all Whitney stratifications of M/G. This statement is the so called **Stratification Theorem**.

Later we shall need the following result for a proper Lie group action. If M/G is connected, there is a unique maximal conjugacy class $(H) \in I(G, M)$, that is, $(G_x) \preceq (H)$ for all $x \in M$. Its associated orbit type manifold $M_{(H)}$ is open and dense in M and the orbit space $M_{(H)}$ is connected.

The Symplectic Stratification Theorem.   With these preparations in mind we can state now the precise statement of the **Symplectic Stratification Theorem**.

1.4.2 Theorem.   *Let (M, ω) be a connected symplectic manifold acted smoothly, canonically, and properly upon by a Lie group G. Suppose that this action has an associated momentum map $\mathbf{J} : M \to \mathfrak{g}^*$ with non-equivariance one-cocycle $\sigma : G \to \mathfrak{g}^*$. Let $\mu \in \mathfrak{g}^*$ lie in the range of $\mathbf{J}$, G_μ the isotropy subgroup of μ with respect to the affine action defined by $(g, \mu) \in G \times \mathfrak{g}^* \mapsto \mathrm{Ad}^*_{g^{-1}} \mu + \sigma(g) \in \mathfrak{g}^*$, and let $H \subset G$ be an isotropy subgroup of the G-action on M. Let M_H^z be the connected component of the H-isotropy type manifold that contains a given element $z \in M$ such that $\mathbf{J}(z) = \mu$ and let $G_\mu \cdot M_H^z$ be its G_μ-saturation. Then the following hold:*

(i) *The set $\mathbf{J}^{-1}(\mu) \cap (G_\mu \cdot M_H^z)$ is a submanifold of M.*

(ii) *The set $M_\mu^{(H)} := \left[\mathbf{J}^{-1}(\mu) \cap (G_\mu \cdot M_H^z) \right] / G_\mu$ has a unique quotient differentiable structure such that the canonical projection*

$$\pi_\mu^{(H)} : \mathbf{J}^{-1}(\mu) \cap (G_\mu \cdot M_H^z) \longrightarrow M_\mu^{(H)}$$

is a surjective submersion.

(iii) *There is a unique symplectic structure $\omega_\mu^{(H)}$ on $M_\mu^{(H)}$ characterized by the identity*

$$i_\mu^{(H)\,*}\omega = \pi_\mu^{(H)\,*}\omega_\mu^{(H)},$$

*where $i_\mu^{(H)} : \mathbf{J}^{-1}(\mu) \cap (G_\mu \cdot M_H^z) \hookrightarrow M$ is the natural inclusion. The pairs $(M_\mu^{(H)}, \omega_\mu^{(H)})$ are called **singular symplectic point strata**.*

(iv) *Let $h \in C^\infty(M)^G$ be a G-invariant Hamiltonian. Then the flow F_t of X_h leaves the connected components of $\mathbf{J}^{-1}(\mu) \cap (G_\mu \cdot M_H^z)$ invariant and commutes with the G_μ-action, so it induces a flow F_t^μ on $M_\mu^{(H)}$ that is characterized by $\pi_\mu^{(H)} \circ F_t \circ i_\mu^{(H)} = F_t^\mu \circ \pi_\mu^{(H)}$.*

(v) *The flow F_t^μ is Hamiltonian on $M_\mu^{(H)}$, with **reduced Hamiltonian function** $h_\mu^{(H)} : M_\mu^{(H)} \to \mathbb{R}$ defined by $h_\mu^{(H)} \circ \pi_\mu^{(H)} = h \circ i_\mu^{(H)}$. The vector fields X_h and $X_{h_\mu^{(H)}}$ are $\pi_\mu^{(H)}$–related.*

(vi) *Let $k : M \to \mathbb{R}$ be another G-invariant function. Then $\{h, k\}$ is also G-invariant and $\{h, k\}_\mu^{(H)} = \{h_\mu^{(H)}, k_\mu^{(H)}\}_{M_\mu^{(H)}}$, where $\{\ ,\ \}_{M_\mu^{(H)}}$ denotes the Poisson bracket induced by the symplectic structure on $M_\mu^{(H)}$.*

(vii) *The quotient topological space $M_\mu := \mathbf{J}^{-1}(\mu)/G_\mu$ is a cone space when considered as a stratified space with strata $M_\mu^{(H)}$.*

As was the case for regular reduction, this theorem can be also formulated from the orbit reduction point of view. Using that approach one can conclude that the orbit reduced spaces $M_{\mathcal{O}_\mu}$ are cone spaces, symplectically stratified by the manifolds $M_{\mathcal{O}_\mu}^{(H)} := G \cdot (\mathbf{J}^{-1}(\mu) \cap M_H^z)/G$, that have symplectic structure uniquely determined by the expression

$$i_{\mathcal{O}_\mu}^{(H)\,*}\omega = \pi_{\mathcal{O}_\mu}^{(H)\,*}\omega_{\mathcal{O}_\mu}^{(H)} + \mathbf{J}_{\mathcal{O}_\mu}^{(H)\,*}\omega_{\mathcal{O}_\mu}^+,$$

where $i_{\mathcal{O}_\mu}^{(H)} : G \cdot (\mathbf{J}^{-1}(\mu) \cap M_H^z) \hookrightarrow M$ is the inclusion, $\mathbf{J}_{\mathcal{O}_\mu}^{(H)} : G \cdot (\mathbf{J}^{-1}(\mu) \cap M_H^z) \to \mathcal{O}_\mu$ is obtained by restriction of the momentum map $\mathbf{J}$, and $\omega_{\mathcal{O}_\mu}^+$ is the +− symplectic form on $\mathcal{O}_\mu$. Analogous statements to **(i) - (vii)** above are valid with obvious modifications.

As mentioned earlier, the case of singular reduction for cotangent bundles is discussed in §2.4.

2

Cotangent Bundle Reduction

This chapter gives some additional background on symplectic reduction theory, the main topic being one of the most important cases, namely the symplectic reduction of cotangent bundles. The main results concerning cotangent bundle reduction make use of the theory of principal connections and so we provide the necessary background on this theory in the first section. The chapter closes with a description of the setting for the major topic of the book: reduction by stages.

2.1   Principal Bundles and Connections

In preparation for the next section which gives a brief exposition of the cotangent bundle reduction theorem, we now give a review and summary of facts that we shall need about principal connections. An important thing to keep in mind is that the magnetic terms in the cotangent bundle reduction theorem will appear as the curvature of a connection.

Principal Connections Defined.   We consider the following basic set up. Let Q be a manifold and let G be a Lie group acting freely and properly on the left on Q. Let

$$\pi_{Q,G} : Q \to Q/G$$

denote the bundle projection from the configuration manifold Q to **shape space** $S = Q/G$. We refer to $\pi_{Q,G} : Q \to Q/G$ as a **principal bundle**.

One can alternatively use right actions, which is common in the principal bundle literature, but we shall stick with the case of left actions for the main exposition.

Vectors that are infinitesimal generators, namely those of the form $\xi_Q(q)$ are called **vertical** since they are sent to zero by the tangent of the projection map $\pi_{Q,G}$.

2.1.1 Definition. *A **connection**, also called a **principal connection** on the bundle $\pi_{Q,G} : Q \to Q/G$ is a Lie algebra valued 1-form*

$$\mathcal{A} : TQ \to \mathfrak{g}$$

where $\mathfrak{g}$ denotes the Lie algebra of G, with the following properties:

(i) *the identity $\mathcal{A}(\xi_Q(q)) = \xi$ holds for all $\xi \in \mathfrak{g}$; that is, $\mathcal{A}$ takes infinitesimal generators of a given Lie algebra element to that same element, and*

(ii) *we have **equivariance**: $\mathcal{A}(T_q\Phi_g(v)) = \mathrm{Ad}_g(\mathcal{A}(v))$*

for all $v \in T_qQ$, where $\Phi_g : Q \to Q$ denotes the given action for $g \in G$ and where Ad_g denotes the adjoint action of G on $\mathfrak{g}$.

A remark is noteworthy at this point. The equivariance identity for infinitesimal generators noted previously (see (1.1.7)), namely,

$$T_q\Phi_g\left(\xi_Q(q)\right) = (\mathrm{Ad}_g \xi)_Q(g \cdot q),$$

shows that *if the first condition for a connection holds, then the second condition holds automatically on vertical vectors.*

If the G-action on Q is a *right action*, the equivariance condition (ii) in Definition 2.1.1 needs to be changed to $\mathcal{A}(T_q\Phi_g(v)) = \mathrm{Ad}_{g^{-1}}(\mathcal{A}(v))$ for all $g \in G$ and $v \in T_qQ$.

Associated One-Forms. Since $\mathcal{A}$ is a Lie algebra valued 1-form, for each $q \in Q$, we get a linear map $\mathcal{A}(q) : T_qQ \to \mathfrak{g}$ and so we can form its dual $\mathcal{A}(q)^* : \mathfrak{g}^* \to T_q^*Q$. Evaluating this on μ produces an ordinary 1-form:

$$\alpha_\mu(q) = \mathcal{A}(q)^*(\mu). \tag{2.1.1}$$

This 1-form satisfies two important properties given in the next Proposition.

2.1.2 Proposition. *For any connection $\mathcal{A}$ and $\mu \in \mathfrak{g}^*$, the corresponding 1-form α_μ defined by (2.1.1) takes values in $\mathbf{J}^{-1}(\mu)$ and satisfies the following G-equivariance property:*

$$\Phi_g^*\alpha_\mu = \alpha_{\mathrm{Ad}_g^*\mu}.$$

Proof.   First of all, notice that from the definition of α_μ and then using first property of a connection,

$$
\begin{aligned}
\langle \mathbf{J}(\alpha_\mu(q)), \xi \rangle &= \langle \alpha_\mu(q), \xi_Q(q) \rangle \\
&= \langle \mathcal{A}(q)^*(\mu), \xi_Q(q) \rangle \\
&= \langle \mu, \mathcal{A}(q)(\xi_Q(q)) \rangle \\
&= \langle \mu, \xi \rangle .
\end{aligned}
$$

Since $\xi \in \mathfrak{g}$ is arbitrary, we conclude that $\mathbf{J}(\alpha_\mu(q)) = \mu$ and therefore, indeed, α_μ takes values in $\mathbf{J}^{-1}(\mu)$.

To establish invariance of the form α_μ, we compute in the following way. Let $v \in T_qQ$ and $g \in G$, and first use the definition of α_μ and the definition of the adjoint to get

$$
\begin{aligned}
(\Phi_g^*\alpha_\mu)(v) &= \alpha_\mu(g \cdot q)(T_q\Phi_g(v)) \\
&= \langle \mathcal{A}(g \cdot q)^*(\mu), T_q\Phi_g(v) \rangle \\
&= \langle \mu, \mathcal{A}(g \cdot q)(T_q\Phi_g(v)) \rangle .
\end{aligned}
$$

Next, make use of equivariance of $\mathcal{A}$ and convert the preceding expression back to one involving α_μ to get:

$$
\begin{aligned}
(\Phi_g^*\alpha_\mu)(v) &= \langle \mu, \mathrm{Ad}_g(\mathcal{A}(q)(v)) \rangle \\
&= \langle \mathrm{Ad}_g^*\mu, \mathcal{A}(q)(v) \rangle \\
&= \langle \mathcal{A}(q)^*\mathrm{Ad}_g^*\mu, v \rangle \\
&= \alpha_{\mathrm{Ad}_g^*\mu}(q)(v)
\end{aligned}
$$

so that we get the required equivariance property.   ∎

Notice in particular, if the group is Abelian or if μ is G-invariant, (for example, if $\mu = 0$), then α_μ is an *invariant 1-form*.

Horizontal and Vertical Spaces.   Associated with any connection are vertical and horizontal spaces defined as follows.

2.1.3 Definition.   *Given the connection $\mathcal{A}$, its **horizontal space** at $q \in Q$ is defined by*

$$
H_q = \{ v_q \in T_qQ \mid \mathcal{A}(v_q) = 0 \}
$$

*and the **vertical space** at $q \in Q$ is, as above,*

$$
V_q = \{ \xi_Q(q) \mid \xi \in \mathfrak{g} \}.
$$

The map

$$
v_q \mapsto \mathrm{ver}_q(v_q) := [\mathcal{A}(q)(v_q)]_Q(q)
$$

*is called the **vertical projection**, while the map*

$$
v_q \mapsto \mathrm{hor}_q(v_q) := v_q - \mathrm{ver}_q(v_q)
$$

*is called the **horizontal projection**.*

Because connections map infinitesimal generators of a Lie algebra elements to that same Lie algebra element, the vertical projection is indeed a projection for each fixed q onto the vertical space and likewise with the horizontal projection.

By construction, we have

$$v_q = \mathrm{ver}_q(v_q) + \mathrm{hor}_q(v_q)$$

and so

$$T_q Q = H_q \oplus V_q$$

and the maps hor_q and ver_q are projections onto these subspaces.

It is sometimes convenient to *define* a connection by the specification of a space H_q declared to be the horizontal space that is complementary to V_q at each point, varies smoothly with q and respects the group action in the sense that $H_{g\cdot q} = T_q\Phi_g(H_q)$. Clearly this alternative definition of a principal connection is equivalent to the definition given above.

Given a point $q \in Q$, the tangent of the projection map $\pi_{Q,G}$ restricted to the horizontal space H_q gives an isomorphism between H_q and $T_{[q]}(Q/G)$. Its inverse $\left[T_q\pi_{Q,G}|_{H_q}\right]^{-1} : T_{\pi_{Q,G}(q)}(Q/G) \to H_q$ is called the ***horizontal lift*** to $q \in Q$.

The Mechanical Connection.   As an example of defining a connection by the specification of a horizontal space, suppose that the configuration manifold Q is a Riemannian manifold. Of course, the Riemannian structure will often be that defined by the kinetic energy of a given mechanical system.

Thus, assume that Q is a Riemannian manifold, with metric denoted $\langle\!\langle\,,\rangle\!\rangle$ and that G acts freely and properly on Q by isometries, so $\pi_{Q,G} : Q \to Q/G$ is a principal G-bundle.

In this context we may *define the horizontal space at a point simply to be the metric orthogonal to the vertical space.* This therefore defines a connection called the ***mechanical connection***.

Recall from the historical survey in the introduction that this connection was first introduced by Kummer [1981] following motivation from Smale [1970] and [FofM]. See also Guichardet [1984], who applied these ideas in an interesting way to molecular dynamics. The number of references since then making use of the mechanical connection is too large to survey here.

In Proposition 2.1.5 we develop an explicit formula for the associated Lie algebra valued 1-form in terms of an inertia tensor and the momentum map. As a prelude to this formula, we show the following basic link with mechanics. In this context we write the momentum map on TQ simply as $\mathbf{J} : TQ \to \mathfrak{g}^*$.

2.1.4 Proposition.   *The horizontal space of the mechanical connection at a point $q \in Q$ consists of the set of vectors $v_q \in T_q Q$ such that $\mathbf{J}(v_q) = 0$.*

Proof. This follows directly from the formula for the momentum map for a Lagrangian that is given by the kinetic energy of a given Riemannian metric, namely,

$$\langle \mathbf{J}(v_q), \xi \rangle = \langle\!\langle v_q, \xi_Q(q) \rangle\!\rangle$$

and the fact that the vertical space at $q \in Q$ is spanned by the set of infinitesimal generators $\xi_Q(q)$. ∎

For each $q \in Q$, define the **locked inertia tensor** $\mathbb{I}(q)$ to be the linear map $\mathbb{I}(q) : \mathfrak{g} \to \mathfrak{g}^*$ defined by

$$\langle \mathbb{I}(q)\eta, \zeta \rangle = \langle\!\langle \eta_Q(q), \zeta_Q(q) \rangle\!\rangle \tag{2.1.2}$$

for any $\eta, \zeta \in \mathfrak{g}$. Since the action is free, $\mathbb{I}(q)$ is nondegenerate, so (2.1.2) defines an inner product. The terminology "locked inertia tensor" comes from the fact that for coupled rigid or elastic systems, $\mathbb{I}(q)$ is the classical moment of inertia tensor of the rigid body obtained by locking all the joints of the system. In coordinates,

$$I_{ab} = g_{ij} K_a^i K_b^j, \tag{2.1.3}$$

where $[\xi_Q(q)]^i = K_a^i(q)\xi^a$ define the **action functions** K_a^i.

Define the map $\mathcal{A} : TQ \to \mathfrak{g}$ which assigns to each $v_q \in T_q Q$ the corresponding *angular velocity of the locked system*:

$$\mathcal{A}(q)(v_q) = \mathbb{I}(q)^{-1}(\mathbf{J}(v_q)), \tag{2.1.4}$$

where L is the kinetic energy Lagrangian. In coordinates,

$$\mathcal{A}^a = I^{ab} g_{ij} K_b^i v^j \tag{2.1.5}$$

since $J_a(q, p) = p_i K_a^i(q)$.

We defined the mechanical connection by declaring its horizontal space to be the metric orthogonal to the vertical space. The next proposition shows that $\mathcal{A}$ is the associated connection one-form.

2.1.5 Proposition. *The $\mathfrak{g}$-valued one-form defined by* (2.1.4) *is the mechanical connection on the principal G-bundle $\pi_{Q,G} : Q \to Q/G$.*

Proof. First notice that $\mathcal{A}$ is G-equivariant and satisfies $\mathcal{A}(\xi_Q(q)) = \xi$, both of which are readily verified. In checking equivariance, one uses invariance of the metric and hence equivariance of $\mathbb{F}L : TQ \to T^*Q$, where L is the kinetic energy of the metric, equivariance of $\mathbf{J} : T^*Q \to \mathfrak{g}^*$, and equivariance of $\mathbb{I}$ in the sense of a map $\mathbb{I} : Q \times \mathfrak{g} \to \mathfrak{g}^*$; that is,

$$\mathbb{I}(g \cdot q)(\mathrm{Ad}_g \xi) = \mathrm{Ad}_{g^{-1}}^* \mathbb{I}(q)(\xi).$$

Thus, $\mathcal{A}$ is a connection.

The horizontal space of $\mathcal{A}$ is given by

$$H_q = \{v_q \in T_qQ \mid \mathbf{J}(\mathbb{F}L(v_q)) = 0\}. \tag{2.1.6}$$

Thus, by Proposition 2.1.4 and the fact that any two connections with the same horizontal spaces are equal, we get the result.   ∎

Given a general connection $\mathcal{A}$ and an element $\mu \in \mathfrak{g}^*$, we can define the μ-component of $\mathcal{A}$ to be the ordinary one-form α_μ given by

$$\alpha_\mu(q) = \mathcal{A}(q)^*\mu \in T_q^*Q; \quad \text{i.e.,} \quad \langle \alpha_\mu(q), v_q \rangle = \langle \mu, \mathcal{A}(q)(v_q) \rangle$$

for all $v_q \in T_qQ$. Note that α_μ is a G_μ-invariant one-form. It takes values in $\mathbf{J}^{-1}(\mu)$ since for any $\xi \in \mathfrak{g}$, we have

$$\langle \mathbf{J}(\alpha_\mu(q)), \xi \rangle = \langle \alpha_\mu(q), \xi_Q \rangle = \langle \mu, \mathcal{A}(q)(\xi_Q(q)) \rangle = \langle \mu, \xi \rangle.$$

In the Riemannian context, Smale [1970] constructed α_μ by a minimization process. Let $\alpha_q^\sharp \in T_qQ$ be the tangent vector that corresponds to $\alpha_q \in T_q^*Q$ via the metric $\langle\!\langle \, , \rangle\!\rangle$ on Q.

2.1.6 Proposition.   *The 1-form $\alpha_\mu(q) = \mathcal{A}(q)^*\mu \in T_q^*Q$ associated with the mechanical connection $\mathcal{A}$ given by (2.1.4) is characterized by*

$$K(\alpha_\mu(q)) = \inf\{K(\beta_q) \mid \beta_q \in \mathbf{J}^{-1}(\mu) \cap T_q^*Q\}, \tag{2.1.7}$$

where $K(\beta_q) = \frac{1}{2}\|\beta_q^\sharp\|^2$ is the kinetic energy function on T^*Q. See Figure 2.1.1.
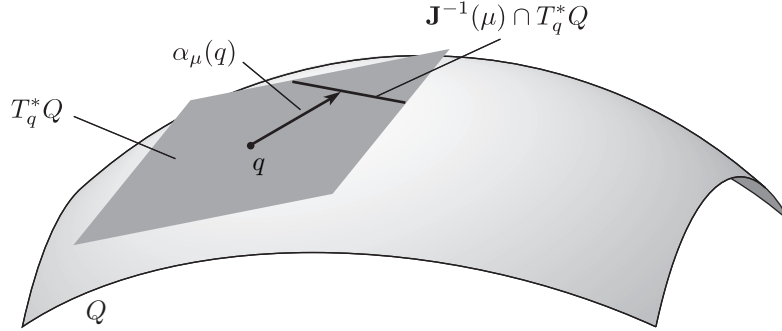
FIGURE 2.1.1. The extremal characterization of the mechanical connection.

The proof is a direct verification. We do not give here it since this proposition will not be used later in this book. The original approach of Smale [1970] was to take (2.1.7) as the definition of α_μ. To prove from here that α_μ is a smooth one-form is a nontrivial fact; see the proof in Smale [1970]

or of Proposition 4.4.5 in [FofM]. Thus, one of the merits of the previous proposition is to show easily that this variational definition of α_μ does indeed yield a smooth one-form on Q with the desired properties. Note also that $\alpha_\mu(q)$ lies in the orthogonal space to $T_q^*Q \cap \mathbf{J}^{-1}(\mu)$ in the fiber T_q^*Q relative to the bundle metric on T^*Q defined by the Riemannian metric on Q. It also follows that $\alpha_\mu(q)$ is the unique critical point of the kinetic energy of the bundle metric on T^*Q restricted to the fiber $T_q^*Q \cap \mathbf{J}^{-1}(\mu)$.

Curvature.   The curvature $\mathcal{B}$ of a connection $\mathcal{A}$ is defined as follows.

2.1.7 Definition.   *The **curvature** of a connection $\mathcal{A}$ is the Lie algebra valued two-form on Q defined by*

$$\mathcal{B}(q)(u_q, v_q) = \mathbf{d}\mathcal{A}(\mathrm{hor}_q(u_q), \mathrm{hor}_q(v_q)), \tag{2.1.8}$$

where $\mathbf{d}$ is the exterior derivative.

When one replaces vectors in the exterior derivative with their horizontal projections, then the result is called the ***exterior covariant derivative*** and one writes the preceding formula for $\mathcal{B}$ as

$$\mathcal{B} = \mathbf{d}^{\mathcal{A}}\mathcal{A}.$$

For a general Lie algebra valued k-form α on Q, the ***exterior covariant derivative*** is the $k+1$-form $\mathbf{d}^{\mathcal{A}}\alpha$ defined on tangent vectors $v_0, v_1, \ldots, v_k \in T_qQ$ by

$$\mathbf{d}^{\mathcal{A}}\alpha(v_0, v_1, \ldots, v_k) = \mathbf{d}\alpha\Big(\mathrm{hor}_q(v_0), \mathrm{hor}_q(v_1), \ldots, \mathrm{hor}_q(v_k)\Big). \tag{2.1.9}$$

Here, the symbol $\mathbf{d}^{\mathcal{A}}$ reminds us that it is like the exterior derivative but that it depends on the connection $\mathcal{A}$.

Curvature *measures the lack of integrability of the horizontal distribution* in the following sense.

2.1.8 Proposition.   *On two vector fields u, v on Q one has*

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

Proof.   We use the formula (1.2.5) for the exterior derivative of a 1-form to get

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

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

Given a general distribution $\mathcal{D} \subset TQ$ on a manifold Q one can also define its curvature in an analogous way directly in terms of its lack of integrability. Define ***vertical vectors*** at $q \in Q$ to be the quotient space $T_qQ/\mathcal{D}_q$ and define the curvature acting on two ***horizontal vector fields*** u, v (that is, two vector fields that take their values in the distribution) to be the projection onto the quotient of their Jacobi–Lie bracket. One can check that this operation depends only on the point values of the vector fields, so indeed defines a two-form on horizontal vectors.

Cartan Structure Equations.   We now derive an important formula for the curvature of a principal connection.

2.1.9 Theorem (Cartan Structure Equations).   *For any vector fields u, v on Q we have*

$$\mathcal{B}(u, v) = \mathbf{d}\mathcal{A}(u, v) - [\mathcal{A}(u), \mathcal{A}(v)], \qquad (2.1.10)$$

where the bracket on the right hand side is the Lie bracket in $\mathfrak{g}$. We write this equation for short as

$$\mathcal{B} = \mathbf{d}\mathcal{A} - [\mathcal{A}, \mathcal{A}].$$

If the G-action on Q is a *right action*, then the Cartan Structure Equations read $\mathcal{B} = \mathbf{d}\mathcal{A} + [\mathcal{A}, \mathcal{A}]$.

Proof.   To prove this theorem we prepare a lemma.

2.1.10 Lemma.   *The identity $\mathbf{d}\mathcal{A}(\mathrm{hor}(u), \mathrm{ver}(v)) = 0$ holds for any two vector fields u, v on Q .*

Proof.   Since this identity depends only on the point values of u and v, we can assume that $\mathrm{ver}(v) = \xi_Q$ identically. Then, as in the preceding proposition, we have

$$\begin{aligned}
\mathbf{d}\mathcal{A}(\mathrm{hor}(u), \mathrm{ver}(v)) &= (\mathrm{hor}(u))[\mathcal{A}(\xi_Q)] - \xi_Q[\mathcal{A}(\mathrm{hor}(u))] - \mathcal{A}([\mathrm{hor}(u), \xi_Q]) \\
&= \mathrm{hor}(u)[\xi] - \xi_Q[0] + \mathcal{A}[\xi_Q, \mathrm{hor}(u)] \\
&= \mathcal{A}[\xi_Q, \mathrm{hor}(u)]
\end{aligned}$$

since ξ is constant. However, the flow of ξ_Q is $\Phi_{\exp(t\xi)}$ and the map hor is equivariant and so

$$\begin{aligned}
[\xi_Q, \mathrm{hor}(u)] &= \left.\frac{d}{dt}\right|_{t=0} \Phi_{\exp(t\xi)}^* \, \mathrm{hor}(u) \\
&= \mathrm{hor}\left.\frac{d}{dt}\right|_{t=0} \Phi_{\exp(t\xi)}^*(u) \\
&= \mathrm{hor}[\xi_Q, u]
\end{aligned}$$

Thus, $[\xi_Q, \mathrm{hor}(u)]$ is horizontal and so it is annihilated by $\mathcal{A}$ and the lemma follows.   ▼

With this lemma in hand we can complete the proof of the Cartan Structure Equations. Making use of the lemma and writing $u = \mathrm{hor}(u) + \mathrm{ver}(u)$ and similarly for v, shows that

$$\mathbf{d}\mathcal{A}(u, \mathrm{ver}(v)) = \mathbf{d}\mathcal{A}(\mathrm{ver}(u), \mathrm{ver}(v))$$

and so we get

$$\mathcal{B}(u, v) = \mathbf{d}\mathcal{A}(u, v) - \mathbf{d}\mathcal{A}(\mathrm{ver}(u), \mathrm{ver}(v)).$$

Again, the second term on the right hand side of this equation depends only on the point values of u and v and so we can assume that $\text{ver}(u) = \xi_Q$ and that $\text{ver}(v) = \eta_Q$ for $\xi \in \mathfrak{g}$ and $\eta \in \mathfrak{g}$. Then

$$
\begin{aligned}
\mathbf{d}\mathcal{A}(\xi_Q, \eta_Q) &= \xi_Q[\mathcal{A}(\eta_Q)] - \eta_Q[\mathcal{A}(\xi_Q)] - \mathcal{A}([\xi_Q, \eta_Q]) \\
&= \mathcal{A}([\xi, \eta]_Q) = [\xi, \eta] \\
&= [\mathcal{A}(u), \mathcal{A}(v)],
\end{aligned}
$$

which, combined with the preceding formula for $\mathcal{B}$, proves the result. $\blacksquare$

The following Corollary shows how the Cartan Structure Equations yield a fundamental equivariance property of the curvature.

2.1.11 Corollary.   *For all $g \in G$ we have $\Phi_g^* \mathcal{B} = \text{Ad}_g \circ \mathcal{B}$. If the G-action on Q is on the right, equivariance means $\Phi_g^* \mathcal{B} = \text{Ad}_{g^{-1}} \circ \mathcal{B}$.*

Proof.   From the definition of $[\mathcal{A}, \mathcal{A}]$, we get for any $u_q, v_q \in T_q Q$ and $g \in G$

$$
\begin{aligned}
\left(\Phi_g^*[\mathcal{A}, \mathcal{A}]\right)(q)(u_q, v_q) &= [\mathcal{A}, \mathcal{A}](g \cdot q)\left(T_q\Phi_g(u_q), T_q\Phi_g(v_q)\right) \\
&= [\mathcal{A}(q)\left(T_q\Phi_g(u_q)\right), \mathcal{A}(q)\left(T_q\Phi_g(v_q)\right)] \\
&= [\left(\Phi_g^*\mathcal{A}\right)(q)(u_q), \left(\Phi_g^*\mathcal{A}\right)(q)(v_q)] \\
&= [\Phi_g^*\mathcal{A}, \Phi_g^*\mathcal{A}](q)(u_q, v_q),
\end{aligned}
$$

that is, $\Phi_g^*[\mathcal{A}, \mathcal{A}] = [\Phi_g^*\mathcal{A}, \Phi_g^*\mathcal{A}]$. Thus Definition 2.1.1(ii) gives

$$
\Phi_g^*[\mathcal{A}, \mathcal{A}] = [\Phi_g^*\mathcal{A}, \Phi_g^*\mathcal{A}] = [\text{Ad}_g \circ \mathcal{A}, \text{Ad}_g \circ \mathcal{A}] = \text{Ad}_g \circ [\mathcal{A}, \mathcal{A}]
$$

and hence

$$
\begin{aligned}
\Phi_g^*\mathcal{B} &= \Phi_g^*\left(\mathbf{d}\mathcal{A} - [\mathcal{A}, \mathcal{A}]\right) = \mathbf{d}\left(\Phi_g^*\mathcal{A}\right) - \text{Ad}_g \circ [\mathcal{A}, \mathcal{A}] \\
&= \mathbf{d}\left(\text{Ad}_g \circ \mathcal{A}\right) - \text{Ad}_g \circ [\mathcal{A}, \mathcal{A}] = \text{Ad}_g \circ \left(\mathbf{d}\mathcal{A} - [\mathcal{A}, \mathcal{A}]\right) = \text{Ad}_g \circ \mathcal{B}
\end{aligned}
$$

as required. The case of right actions is proved in a similar way. $\blacksquare$

Bianchi Identity.   The famous ***Bianchi Identity***, which states that the exterior covariant derivative of the curvature is zero, is another important consequence of the Cartan Structure Equations.

2.1.12 Corollary.   *If $\mathcal{B} = \mathbf{d}^\mathcal{A}\mathcal{A} \in \Omega^2(Q; \mathfrak{g})$ is the curvature two-form of the connection $\mathcal{A}$, then the **Bianchi Identity** holds:*

$$
\mathbf{d}^\mathcal{A}\mathcal{B} = 0.
$$

Proof.   Consider three vector fields $u, v, w \in \mathfrak{X}(Q)$. We want to show that $\mathbf{d}^\mathcal{A}\mathcal{B}(u, v, w)$ vanishes. Start by using the definition of the exterior

covariant derivative (2.1.9), the Cartan Structure Equations (2.1.10) and the identity $\mathbf{d}^2 = 0$ for the usual exterior derivative:

$$
\begin{aligned}
\mathbf{d}^{\mathcal{A}}\mathcal{B}(u,v,w) &= \mathbf{d}\mathcal{B}(\mathrm{hor}(u),\mathrm{hor}(v),\mathrm{hor}(w)) \\
&= \mathbf{d}\left(\mathbf{d}\mathcal{A} - [\mathcal{A},\mathcal{A}]\right)(\mathrm{hor}(u),\mathrm{hor}(v),\mathrm{hor}(w)) \\
&= -\mathbf{d}[\mathcal{A},\mathcal{A}](\mathrm{hor}(u),\mathrm{hor}(v),\mathrm{hor}(w)).
\end{aligned}
$$

Next, we make use of the following identity for the exterior derivative of a two form (see, for example, [MTA]): for ω a two-form and X, Y, Z vector fields, one has,

$$
\begin{aligned}
\mathbf{d}\omega(X,Y,Z) &= X[\omega(Y,Z)] + Y[\omega(Z,X)] + Z[\omega(X,Y)] \\
&\quad - \omega([X,Y],Z) - \omega([Z,X],Y) - \omega([Y,Z],X)
\end{aligned}
$$

Using this identity with $\omega = [\mathcal{A},\mathcal{A}]$, we continue with the preceding computation of $\mathbf{d}^{\mathcal{A}}\mathcal{B}(u,v,w)$ to get

$$
\begin{aligned}
\mathbf{d}^{\mathcal{A}}\mathcal{B}(u,v,w) = &- \mathrm{hor}(u)\big[[\mathcal{A}(\mathrm{hor}(v)),\mathcal{A}(\mathrm{hor}(w))]\big] \\
&- \mathrm{hor}(v)\big[[\mathcal{A}(\mathrm{hor}(w)),\mathcal{A}(\mathrm{hor}(u))]\big] \\
&- \mathrm{hor}(w)\big[[\mathcal{A}(\mathrm{hor}(u)),\mathcal{A}(\mathrm{hor}(v))]\big] \\
&+ [\mathcal{A},\mathcal{A}]([\mathrm{hor}(u),\mathrm{hor}(v)],\mathrm{hor}(w)) \\
&+ [\mathcal{A},\mathcal{A}]([\mathrm{hor}(v),\mathrm{hor}(w)],\mathrm{hor}(u)) \\
&+ [\mathcal{A},\mathcal{A}]([\mathrm{hor}(w),\mathrm{hor}(u)],\mathrm{hor}(v)).
\end{aligned}
$$

Each term in this expression is zero since $\mathcal{A}$ vanishes on horizontal vectors. Therefore, $\mathbf{d}^{\mathcal{A}}\mathcal{B}(u,v,w) = 0$.   ∎

This form of the Bianchi identity is implied by another version, namely

$$
\mathbf{d}\mathcal{B} = [\mathcal{B},\mathcal{A}]^{\wedge},
$$

where the bracket on the right hand side is that of Lie algebra valued differential forms, a notion that we do not develop here; see the brief discussion at the end of §9.1 in [MandS]. The proof of the above form of the Bianchi identity can be found in, for example, Kobayashi and Nomizu [1963] or in Bleecker [1981], Proposition 2.2.8, page 39.

Aside: Link to Riemannian Curvature.   Since connections and curvature are commonly developed in the setting of Riemannian geometry, it may be useful to point out at least the starting point of the link with the preceding approach of principal connections.

Given an oriented Riemannian manifold M with metric tensor g and associated Levi-Civita connection, first construct the oriented orthonormal frame bundle B over M. This is a principal bundle with the group $\mathrm{SO}(n)$ acting fiberwise in a free and proper way (one can make it either a left or

right action as one prefers). The quotient is clearly just M. Now define an $\mathfrak{so}(n)$-valued one form $\mathcal{A}_g$ on B in the following way.

Start with a point in B, namely an orthonormal frame $(e_1, \ldots e_n)$ based at a point x in M. Next, take a tangent vector w to this frame in B, which we can identify with an equivalence class of moving frames. These moving frames define a tangent vector v to M, namely that of the moving base point, say $x(t)$ of these frames. Parallel transporting (using the Levi-Civita connection) these frames back along the curve $x(t)$ to the point x produces a curve in $\mathrm{SO}(n)$ starting at the identity (since any two oriented orthonormal frames are uniquely related by an element of $\mathrm{SO}(n)$). Thus, a tangent vector to B produces, in a natural way, an element v of TM together with an element of the Lie algebra $\mathfrak{so}(n)$. The principal connection $\mathcal{A}_g$ associated to the metric tensor g is just the map taking w to this element of $\mathfrak{so}(n)$; one can routinely verify the needed properties of a connection.

Detailed development of this idea is closely related to differential form methods for computing curvature in concrete examples and to mathematical ideas related to the Gauss-Bonnet theorem, Chern classes, and the like. Unfortunately developing the full details of how this goes is a rather long story, and the literature is not very accessible, but the reader can consult, for example, Kobayashi and Nomizu [1963], Misner, Thorne, and Wheeler [1973] and Spivak [1979] for further information.

Curvature as a Two-Form on the Base.   We now show how the curvature two-form drops to a two-form on the base with values in the adjoint bundle.

The associated bundle to the given left principal bundle $\pi_{Q,G} : Q \to Q/G$ via the adjoint action is called the ***adjoint bundle***. It is defined in the following way. Consider the free proper action $(g, (q, \xi)) \in G \times (Q \times \mathfrak{g}) \mapsto (g \cdot q, \mathrm{Ad}_g \xi) \in Q \times \mathfrak{g}$ and form the quotient $\tilde{\mathfrak{g}} := Q \times_G \mathfrak{g} := (Q \times \mathfrak{g})/G$ which is easily verified to be a vector bundle $\pi_{\tilde{\mathfrak{g}}} : \tilde{\mathfrak{g}} \to Q/G$, where $\pi_{\tilde{\mathfrak{g}}}(g, \xi) := \pi_{Q,G}(q)$. This vector bundle has an additional structure: it is a ***Lie algebra bundle***; that is, a vector bundle whose fibers are Lie algebras. In this case the bracket is defined pointwise:

$$[\pi_{\tilde{\mathfrak{g}}}(g, \xi), \pi_{\tilde{\mathfrak{g}}}(g, \eta)] := \pi_{\tilde{\mathfrak{g}}}(g, [\xi, \eta])$$

for all $g \in G$ and $\xi, \eta \in \mathfrak{g}$. It is easy to check that this defines a Lie bracket on every fiber and that this operation is smooth as a function of $\pi_{Q,G}(q)$.

The curvature two-form $\mathcal{B} \in \Omega^2(Q; \mathfrak{g})$ (the vector space of $\mathfrak{g}$-valued two-forms on Q) naturally induces a two-form $\overline{\mathcal{B}}$ on the base Q/G with values in $\tilde{\mathfrak{g}}$ by

$$\overline{\mathcal{B}}(\pi_{Q,G}(q)) \left(T_q \pi_{Q,G}(u), T_q \pi_{Q,G}(v) \right) := \pi_{\tilde{\mathfrak{g}}}\left(q, \mathcal{B}(u, v) \right) \qquad (2.1.11)$$

for all $q \in Q$ and $u, v \in T_q Q$.

Let us prove that $\overline{\mathcal{B}}$ is well defined. If $q' = g \cdot q$ and $u', v' \in T_{q'}Q$ are such that

$$T_{q'}\pi_{Q,G}(u') = T_{q'}\pi_{Q,G}\left(T_q\Phi_g(u)\right) = T_q\pi_{Q,G}(u)$$

and

$$T_{q'}\pi_{Q,G}(v') = T_{q'}\pi_{Q,G}\left(T_q\Phi_g(v)\right) = T_q\pi_{Q,G}(v),$$

then $u' - T_q\Phi_g(u), v' - T_q\Phi_g(v) \in V_{q'}$ and hence by Corollary 2.1.11 we get

$$
\begin{aligned}
\pi_{\tilde{\mathfrak{g}}} &\left(q', \mathcal{B}(u', v')\right) \\
&= \pi_{\tilde{\mathfrak{g}}}\left(g \cdot q, \mathcal{B}(T_q\Phi_g(u) + (u' - T_q\Phi_g(u)), T_q\Phi_g(v) + (v' - T_q\Phi_g(v)))\right) \\
&= \pi_{\tilde{\mathfrak{g}}}\left(g \cdot q, \mathcal{B}(T_q\Phi_g(u), T_q\Phi_g(v))\right) \\
&= \pi_{\tilde{\mathfrak{g}}}\left(g \cdot q, (\Phi_g^*\mathcal{B})(u, v)\right) \\
&= \pi_{\tilde{\mathfrak{g}}}\left(g \cdot q, \mathrm{Ad}_g\left(\mathcal{B}(u, v)\right)\right) \\
&= \pi_{\tilde{\mathfrak{g}}}\left(q, \mathcal{B}(u, v)\right)
\end{aligned}
$$

which shows that the right hand side of (2.1.11) is independent of the choices made to define $\overline{\mathcal{B}}$.

Since (2.1.11) can be equivalently written as $\pi_{Q,G}^*\overline{\mathcal{B}} = \pi_{\tilde{\mathfrak{g}}} \circ (\mathrm{id}_Q \times \mathcal{B})$ and $\pi_{Q,G}$ is a surjective submersion, it follows that $\overline{\mathcal{B}}$ is indeed a smooth two-form on Q/G with values in $\tilde{\mathfrak{g}}$.

Associated Two-Forms. Since $\mathcal{B}$ is a $\mathfrak{g}$-valued two-form, in analogy with (2.1.1), for every $\mu \in \mathfrak{g}^*$ we can define the μ-***component*** of $\mathcal{B}$, an ordinary two-form $\mathcal{B}_\mu \in \Omega^2(Q)$ on Q, by

$$\mathcal{B}_\mu(q)(u_q, v_q) := \langle \mu, \mathcal{B}(q)(u_q, v_q) \rangle \tag{2.1.12}$$

for all $q \in Q$ and $u_q, v_q \in T_qQ$.

The adjoint bundle valued curvature two-form $\overline{\mathcal{B}}$ induces an ordinary two-form on the base Q/G. To obtain it, we consider the dual $\tilde{\mathfrak{g}}^*$ of the adjoint bundle. This is a vector bundle over Q/G which is the associated bundle relative to the coadjoint action of the structure group G of the principal (left) bundle $\pi_{Q,G} : Q \to Q/G$ on $\mathfrak{g}^*$. This vector bundle has additional structure: each of its fibers is a Lie-Poisson space and the associated Poisson tensors on each fiber depend smoothly on the base, that is, $\pi_{\tilde{\mathfrak{g}}^*} : \tilde{\mathfrak{g}}^* \to Q/G$ is a ***Lie-Poisson bundle*** over Q/G.

Given $\mu \in \mathfrak{g}^*$, define the *ordinary two-form* $\overline{\mathcal{B}}_\mu$ *on* Q/G by

$$
\begin{aligned}
\overline{\mathcal{B}}_\mu &\left(\pi_{Q,G}(q)\right)\left(T_q\pi_{Q,G}(u_q), T_q\pi_{Q,G}(v_q)\right) \\
&:= \left\langle \pi_{\tilde{\mathfrak{g}}^*}(q, \mu), \overline{\mathcal{B}}(\pi_{Q,G}(q))\left(T_q\pi_{Q,G}(u_q), T_q\pi_{Q,G}(v_q)\right)\right\rangle \\
&= \langle \mu, \mathcal{B}(q)(u_q, v_q) \rangle = \mathcal{B}_\mu(q)(u_q, v_q), \tag{2.1.13}
\end{aligned}
$$

where $q \in Q$, $u_q, v_q \in T_qQ$, and in the second equality $\langle\,,\,\rangle : \tilde{\mathfrak{g}}^* \times \tilde{\mathfrak{g}} \to \mathbb{R}$ is the duality pairing between the coadjoint and adjoint bundles. Since $\overline{\mathcal{B}}$ is well defined and smooth, so is $\overline{\mathcal{B}}_\mu$.

2.1.13 Proposition.   *Let $\mathcal{A} \in \Omega^1(Q; \mathfrak{g})$ be a connection one-form on the (left) principal bundle $\pi_{Q,G} : Q \to Q/G$ and $\mathcal{B} \in \Omega^2(Q; \mathfrak{g})$ its curvature two-form on Q. If $\mu \in \mathfrak{g}^*$, the corresponding two-forms $\mathcal{B}_\mu \in \Omega^2(Q)$ and $\overline{\mathcal{B}}_\mu \in \Omega^2(Q/G)$ defined by (2.1.12) and (2.1.13), respectively, are related by $\pi_{Q,G}^* \overline{\mathcal{B}}_\mu = \mathcal{B}_\mu$. In addition, $\mathcal{B}_\mu$ satisfies the following G-equivariance property:*

$$\Phi_g^* \mathcal{B}_\mu = \mathcal{B}_{\mathrm{Ad}_g^* \mu}.$$

Thus, if $G = G_\mu$ then $\mathbf{d}\alpha_\mu = \mathcal{B}_\mu = \pi_{Q,G}^* \overline{\mathcal{B}}_\mu$, where $\alpha_\mu(q) = \mathcal{A}(q)^*(\mu)$.

Proof.   The identity $\pi_{Q,G}^* \overline{\mathcal{B}}_\mu = \mathcal{B}_\mu$ is a restatement of (2.1.13). To prove the equivariance of $\mathcal{B}_\mu$, note that for $g \in G$, Corollary 2.1.11 yields

$$\Phi_g^* \mathcal{B}_\mu = \left\langle \mu, \Phi_g^* \mathcal{B} \right\rangle = \left\langle \mu, \mathrm{Ad}_g \circ \mathcal{B} \right\rangle = \left\langle \mathrm{Ad}_g^* \mu, \mathcal{B} \right\rangle = \mathcal{B}_{\mathrm{Ad}_g^* \mu}$$

as required.

The last relation is a consequence of Proposition 2.1.2. Indeed, if $G = G_\mu$ then $\Phi_g^* \alpha_\mu = \alpha_\mu$ for any $g \in G$ so taking the derivative of this relation relative to g at the identity yields $\boldsymbol{\mathcal{L}}_{\xi_Q} \alpha_\mu = 0$. However, we also know that α_μ takes values in $\mathbf{J}^{-1}(\mu)$ and hence

$$(\mathbf{i}_{\xi_Q} \alpha_\mu)(q) = \langle \alpha_\mu(q), \xi_Q(q) \rangle = \langle \mathbf{J}(\alpha_\mu(q)), \xi \rangle = \langle \mu, \xi \rangle,$$

that is, $\mathbf{i}_{\xi_Q} \alpha_\mu$ is a constant function on Q. Therefore,

$$\mathbf{i}_{\xi_Q} \mathbf{d}\alpha_\mu = \boldsymbol{\mathcal{L}}_{\xi_Q} \alpha_\mu - \mathbf{d}\mathbf{i}_{\xi_Q} \alpha_\mu = 0.$$

Now let $q \in Q$, $u_q, v_q \in T_q Q$. Then $u_q - [\langle \mathcal{A}(q), u_q \rangle]_Q(q)$ is the horizontal component of u_q and similarly for v_q. Therefore,

$$\begin{aligned} \mathcal{B}_\mu(q)(u_q, v_q) &= \langle \mu, \, \mathbf{d}\mathcal{A}(q)\,(u_q - [\langle \mathcal{A}(q), u_q \rangle]_Q(q), \, v_q - [\langle \mathcal{A}(q), u_q \rangle]_Q(q)) \rangle \\ &= \mathbf{d}\mathcal{A}(q)\,(u_q, v_q)\,, \end{aligned}$$

since each of the remaining three terms in the expansion is of the form $\mathbf{i}_{\xi_Q} \mathbf{d}\alpha_\mu = 0$, as was shown above. ∎

Further relations between α_μ and the μ-component of the curvature will be studied in the next section when discussing the magnetic terms appearing in cotangent bundle reduction.

The Maurer–Cartan Equations.   A consequence of the structure equations relates curvature to the process of left and right trivialization and hence to momentum maps.

2.1.14 Theorem (Maurer–Cartan Equations).   *Let G be a Lie group and let $\theta^R : TG \to \mathfrak{g}$ be the map (called the right Maurer–Cartan form) that right translates vectors to the identity:*

$$\theta^R(v_g) = T_g R_{g^{-1}}(v_g).$$

Then

$$\mathbf{d}\theta^R - [\theta^R, \theta^R] = 0.$$

Proof.   Note that θ^R is literally a connection on G for the left action. In considering this, keep in mind that for the action by left multiplication we have $\xi_Q(q) = T_e R_g(\xi)$. On the other hand, the curvature of this connection must be zero since the shape space G/G is a point. Thus, the result follows from the structure equations (2.1.10). ∎

Another proof of this result is given in [MandS], §9.1. Of course there is a similar result for the left trivialization θ^L, namely the identity

$$\mathbf{d}\theta^L + [\theta^L, \theta^L] = 0.$$

The results in this section are the main ones that we shall need for subsequent developments, but of course there is much more to this subject, such as the Bianchi identities, the link with classical connection theory, etc. We refer to Kobayashi and Nomizu [1963] for further basic information, to Bloch [2003] for applications to nonholonomic systems and to Cendra, Marsden, and Ratiu [2001a] for applications to Lagrangian reduction.

Computations for Homogeneous Spaces.   We now carry out the calculation of the mechanical connection and its curvature in an important special case that will be needed at several places later in the book. We will choose Q to be a Lie group G and will choose the group that acts on G to be a closed normal subgroup N acting on G by right translations. Thus, we are considering the principal bundle $G \to G/N$. Right translation of a given positive definite inner product on $\mathfrak{g}$ yields a right invariant Riemannian metric $\langle\!\langle\,,\rangle\!\rangle$ on G.

The goal is to now compute the mechanical connection on $G \to G/N$ and its curvature. To do this, we consider the splitting $\mathfrak{g} = \mathfrak{n} \oplus \mathfrak{n}^\perp$ and the corresponding orthogonal projection $\mathcal{P}_\mathfrak{n} : \mathfrak{g} \to \mathfrak{n}$. Since N is a normal subgroup, the adjoint action of G on $\mathfrak{g}$ induces a G-representation

$$\mathrm{Ad}_g^N = \mathrm{Ad}_g\,|_\mathfrak{n} : \mathfrak{n} \to \mathfrak{n}$$

obtained by taking the derivative of the conjugation of G on N. The splitting $\mathfrak{g} = \mathfrak{n} \oplus \mathfrak{n}^\perp$ is in general *not* equivariant with respect to these G representations on $\mathfrak{g}$ and $\mathfrak{n}$. Correspondingly, the adjoint action need not commute with the projection $\mathcal{P}_\mathfrak{n}$.

Since the metric $\langle\!\langle\,,\rangle\!\rangle$ is right G- and hence N-invariant, there is a corresponding mechanical connection defined by the horizontal distribution whose fibers are $H_g := \mathfrak{n}_G(g)^\perp := \{T_e L_g \eta \mid \eta \in \mathfrak{n}\}^\perp$. The horizontal spaces

are in fact the right translates of $\mathfrak{n}^\perp$ since

$$
\begin{aligned}
H_g &= \{w_g \in T_gG \mid \langle\!\langle w_g, T_eL_g\eta \rangle\!\rangle = 0, \text{ for all } \eta \in \mathfrak{n}\} \\
&= \{w_g \in T_gG \mid \langle\!\langle T_gR_{g^{-1}}w_g, \mathrm{Ad}_g\,\eta \rangle\!\rangle = 0 \text{ for all } \eta \in \mathfrak{n}\} \\
&= \{w_g \in T_gG \mid \langle\!\langle T_gR_{g^{-1}}w_g, \zeta \rangle\!\rangle = 0 \text{ for all } \zeta \in \mathfrak{n}\} \\
&= \{w_g \in T_gG \mid T_gR_{g^{-1}}w_g \in \mathfrak{n}^\perp\} \\
&= T_eR_g\left(\mathfrak{n}^\perp\right).
\end{aligned}
\tag{2.1.14}
$$

We now determine the formula for the mechanical connection $\mathcal{A} : TG \to \mathfrak{n}$ having this horizontal subbundle.

2.1.15 Theorem.   *Under the setup above, the following hold.*

(i) *The connection one-form $\mathcal{A} \in \Omega^1(G; \mathfrak{n})$ associated to the G-right invariant Riemannian metric $\langle\!\langle\,,\rangle\!\rangle$ whose horizontal subbundle is given by $H_gG = T_eR_g\left(\mathfrak{n}^\perp\right)$ has the expression*

$$
\mathcal{A}(g)(X_g) = \left(\mathrm{Ad}_{g^{-1}} \circ \mathcal{P}_\mathfrak{n} \circ \theta^R\right)(X_g),
\tag{2.1.15}
$$

where $X_g \in T_gG$ and θ^R is the right Maurer-Cartan one-form (see Theorem 2.1.14).

(ii) *$\mathcal{A}$ has the following equivariance property relative to right translation by $h \in G$:*

$$
R_h^*\mathcal{A} = \mathrm{Ad}_{h^{-1}} \circ \mathcal{A}
\tag{2.1.16}
$$

Proof.   To prove (i), we first compute the locked inertia tensor. Given $\xi_1, \xi_2 \in \mathfrak{n}$ and $g \in G$, we have

$$
\begin{aligned}
\langle \mathbb{I}(g)\xi_1, \xi_2 \rangle &= \langle\!\langle T_eL_g\xi_1, T_eL_g\xi_2 \rangle\!\rangle = \left\langle\!\!\left\langle \mathrm{Ad}_g^N\,\xi_1, \mathrm{Ad}_g^N\,\xi_2 \right\rangle\!\!\right\rangle \\
&= \left\langle\!\!\left\langle \left((\mathrm{Ad}_g^N)^\dagger \circ \mathrm{Ad}_g^N\right)\xi_1, \xi_2 \right\rangle\!\!\right\rangle
\end{aligned}
$$

where $(\mathrm{Ad}_g^N)^\dagger$ is the adjoint of the linear isomorphism Ad_g^N relative to the inner product $\langle\!\langle\,,\rangle\!\rangle$. This gives $\mathbb{I}(g)(\xi) = (\mathrm{Ad}_g^N)^\dagger \circ \mathrm{Ad}_g^N$ and hence $\mathbb{I}(g)^{-1} = \mathrm{Ad}_{g^{-1}}^N \circ (\mathrm{Ad}_{g^{-1}}^N)^\dagger$. Therefore,

$$
\begin{aligned}
\langle \mathbf{J}(\mathbb{F}L(X_g)), \xi \rangle &= \langle\!\langle X_g, T_eL_g\xi \rangle\!\rangle = \left\langle\!\!\left\langle T_gR_{g^{-1}}X_g, \mathrm{Ad}_g^N\,\xi \right\rangle\!\!\right\rangle \\
&= \left\langle\!\!\left\langle \mathcal{P}_\mathfrak{n}\left(T_gR_{g^{-1}}X_g\right), \mathrm{Ad}_g^N\,\xi \right\rangle\!\!\right\rangle \\
&= \left\langle\!\!\left\langle \left((\mathrm{Ad}_g^N)^\dagger \circ \mathcal{P}_\mathfrak{n} \circ \theta^R\right)(X_g), \xi \right\rangle\!\!\right\rangle,
\end{aligned}
$$

where we have used, crucially, the normality of N in G in the second equality by writing $(T_gR_{g^{-1}} \circ T_eL_g)(\xi) = \mathrm{Ad}_g^N\,\xi \in \mathfrak{n}$. This shows that

$\mathbf{J}(\mathbb{F}L(X_g)) = \left((\mathrm{Ad}_g^N)^\dagger \circ \mathcal{P}_{\mathfrak{n}} \circ \theta^R \right)(X_g)$ and hence

$$\begin{aligned}
\mathcal{A}(X_g) &= \mathbb{I}(g)^{-1}\left(\mathbf{J}(\mathbb{F}L(X_g)) \right) \\
&= \left(\mathrm{Ad}_{g^{-1}}^N \circ (\mathrm{Ad}_{g^{-1}}^N)^\dagger \circ (\mathrm{Ad}_g^N)^\dagger \circ \mathcal{P}_{\mathfrak{n}} \circ \theta^R \right)(X_g) \\
&= \left(\mathrm{Ad}_{g^{-1}}^N \circ \mathcal{P}_{\mathfrak{n}} \circ \theta^R \right)(X_g),
\end{aligned}$$

which proves (2.1.15).

To prove (ii), let $g \in G$ and $X_g \in T_g G$. We have

$$\begin{aligned}
R_h^* \mathcal{A}(g)(X_g) &= \mathcal{A}(gh)(T_g R_h X_g) \\
&= \mathrm{Ad}_{h^{-1}g^{-1}} \, \mathcal{P}_{\mathfrak{n}} \left(T_{gh} R_{(gh)^{-1}} T_g R_h X_g \right) \\
&= \mathrm{Ad}_{h^{-1}} \, \mathrm{Ad}_{g^{-1}} \, \mathcal{P}_{\mathfrak{n}} \left(T_g R_{g^{-1}} X_g \right) \\
&= \mathrm{Ad}_{h^{-1}} \, \mathcal{A}(g)(X_g),
\end{aligned}$$

which proves (ii) and thus the Theorem. ∎

Remark.   Notice that if $\mathcal{P}_{\mathfrak{n}}$ were G-equivariant with respect to the Ad^N and Ad-actions, (2.1.15) would become $\mathcal{A} = \mathcal{P}_{\mathfrak{n}} \circ \theta^L$, where θ^L is the left Maurer-Cartan form. However, this is only valid in the trivial case when we have a homomorphism from G to N, which is not the case for group extensions.

The next theorem gives an explicit formula for the curvature of $\mathcal{A}$.

2.1.16 Theorem.   *The curvature two form $\mathcal{B} \in \Omega^2(G; \mathfrak{n})$ of the connection $\mathcal{A} \in \Omega^1(G; \mathfrak{n})$ given in (2.1.15) is*

$$\begin{aligned}
&\mathcal{B}(X_g, Y_g) \\
&= \mathrm{Ad}_{g^{-1}} \left(- \mathrm{ad}_\xi \, \mathcal{P}_{\mathfrak{n}}(\eta) + \mathrm{ad}_\eta \, \mathcal{P}_{\mathfrak{n}}(\xi) + \mathcal{P}_{\mathfrak{n}}[\xi, \eta] + [\mathcal{P}_{\mathfrak{n}}(\xi), \mathcal{P}_{\mathfrak{n}}(\eta)] \right) \quad (2.1.17)
\end{aligned}$$

where $\xi := \theta^R(X_g)$, $\eta := \theta^R(Y_g) \in \mathfrak{g}$.

Proof.   Note first that the four terms in the parenthesis on the right hand side in (2.1.17) are all elements of $\mathfrak{n}$ since $\mathfrak{n}$ is an ideal in $\mathfrak{g}$. Next, recall that by the structure equation (2.1.10) we have

$$\mathcal{B} = \mathbf{d}\mathcal{A} + [\mathcal{A}, \mathcal{A}],$$

where the bracket $[\,,]$ is in $\mathfrak{n}$. There is a plus sign in front of the bracket since we are working here with *right* bundles and connections whereas formula (2.1.10) is for left bundles and connections.

We first compute $\mathbf{d}\mathcal{A}$ as follows

$$\mathbf{d}\mathcal{A}(X_g, Y_g) = X\left[\mathcal{A}(Y) \right](g) - Y\left[\mathcal{A}(X) \right](g) - \mathcal{A}([X, Y])(g), \qquad (2.1.18)$$

where X and Y are the right invariant vector fields on G such that $X(g) = X_g$ and $Y(g) = Y_g$.

Inside this formula, let us first compute the term $X\left[\mathcal{A}(Y)\right](g)$. Letting $\eta := T_g R_{g^{-1}} Y_g = \theta^R(Y_g)$, formula (2.1.15) implies that

$$\mathcal{A}(Y)(h) = \mathcal{A}(h)(Y(h))$$
$$= \left(\mathrm{Ad}_{h^{-1}}^N \circ \mathcal{P}_{\mathfrak{n}}\right)(T_h R_{h^{-1}} Y(h))$$
$$= \left(\mathrm{Ad}_{h^{-1}}^N \circ \mathcal{P}_{\mathfrak{n}}\right)(\eta)$$

for any $h \in G$. Let $\xi := \theta^R(X_g)$, that is, $X_g = \left.\frac{d}{dt}\right|_{t=0}(\exp t\xi)g$, so that

$$X\left[\mathcal{A}(Y)\right](g) = \left.\frac{d}{dt}\right|_{t=0} \mathrm{Ad}_{g^{-1}\exp(-t\xi)}^N \mathcal{P}_{\mathfrak{n}}(\eta)$$
$$= \left.\frac{d}{dt}\right|_{t=0} \mathrm{Ad}_{g^{-1}}^N \mathrm{Ad}_{\exp(-t\xi)}^N \mathcal{P}_{\mathfrak{n}}(\eta)$$
$$= -\mathrm{Ad}_{g^{-1}}^N \mathrm{ad}_\xi \mathcal{P}_{\mathfrak{n}}(\eta).$$

The second term in (2.1.18) is computed the same way. Now we turn to the third term. Since X, Y are right invariant vector fields, we have $[X, Y](g) = -T_e R_g[\xi, \eta]$ and hence

$$\mathcal{A}([X, Y])(g) = -\mathcal{A}(T_e R_g[\xi, \eta]) = -\mathrm{Ad}_{g^{-1}}^N \mathcal{P}_{\mathfrak{n}}([\xi, \eta]).$$

Substituting these results into equation (2.1.18), we get

$$\mathbf{d}\mathcal{A}(X_g, Y_g) = \mathrm{Ad}_{g^{-1}}^N \left(\mathrm{ad}_\xi \mathcal{P}_{\mathfrak{n}}(\eta) + \mathrm{ad}_\eta \mathcal{P}_{\mathfrak{n}}(\xi) + \mathcal{P}_{\mathfrak{n}}([\xi, \eta])\right). \qquad (2.1.19)$$

The last term in the structure equation for the curvature is computed in the following way:

$$[\mathcal{A}(X_g), \mathcal{A}(Y_g)] = [\mathrm{Ad}_{g^{-1}}^N \mathcal{P}_{\mathfrak{n}}(\xi), \mathrm{Ad}_{g^{-1}}^N \mathcal{P}_{\mathfrak{n}}(\eta)]$$
$$= \mathrm{Ad}_{g^{-1}}^N [\mathcal{P}_{\mathfrak{n}}(\xi), \mathcal{P}_{\mathfrak{n}}(\eta)].$$

Combining this with equation (2.1.19) gives the result. ∎

2.2 Cotangent Bundle Reduction: Embedding Version

As mentioned in the introduction, where we reviewed some of the main ideas and the history, the cotangent bundle reduction theorems are amongst the most basic and useful of the symplectic reduction theorems. In this and in the next section, we give the regular versions of the theorems. For

singular versions of general symplectic reduction theorems, see the outline in §1.4, as well as Part III and [HRed] and for the specific case of singular versions of cotangent bundle reduction theorems, see the outline in §2.4.

Cotangent bundle reduction theorems come in two forms—the *embedding cotangent bundle reduction* theorem and the *bundle cotangent bundle reduction* theorem. These are treated respectively in the following two sections. In either case, we start with a smooth, free, and proper, left action

$$\Phi : G \times Q \to Q$$

of the Lie group G on the configuration manifold Q and lift it to an action on T^*Q. This lifted action is symplectic with respect to the canonical symplectic form on T^*Q, which we denote Ω_{can}, and has an *equivariant* momentum map $\mathbf{J} : T^*Q \to \mathfrak{g}^*$ given by

$$\langle \mathbf{J}(\alpha_q), \xi \rangle = \langle \alpha_q, \xi_Q(q) \rangle \,,$$

where $\xi \in \mathfrak{g}$. Letting $\mu \in \mathfrak{g}^*$, the aim of this section is to determine the structure of the symplectic reduced space $((T^*Q)_\mu, \Omega_\mu)$, which, by Theorem 1.1.3, is a symplectic manifold. We are interested in particular in the question of to what extent $((T^*Q)_\mu, \Omega_\mu)$ is a synthesis of a cotangent bundles and a coadjoint orbit.

Cotangent Bundle Reduction: Embedding Version.   In this version of the theorem, we first form the quotient manifold

$$Q_\mu := Q/G_\mu,$$

which we call the μ-**shape space.** Since the action of G on Q is smooth, free, and proper, so is the action of the isotropy subgroup G_μ and therefore, Q_μ is a smooth manifold and the canonical projection

$$\pi_{Q,G_\mu} : Q \to Q_\mu$$

is a surjective submersion.

Consider the G_μ-action on Q and its lift to T^*Q. This lifted action is of course also symplectic with respect to the canonical symplectic form Ω_{can} and has an equivariant momentum map $\mathbf{J}^\mu : T^*Q \to \mathfrak{g}_\mu^*$ obtained by restricting $\mathbf{J}$; that is, for $\alpha_q \in T_q^*Q$,

$$\mathbf{J}^\mu(\alpha_q) = \mathbf{J}(\alpha_q)|_{\mathfrak{g}_\mu}.$$

Let $\mu' := \mu|_{\mathfrak{g}_\mu} \in \mathfrak{g}_\mu^*$ be the restriction of μ to $\mathfrak{g}_\mu$. Notice that there is a natural inclusion of submanifolds

$$\mathbf{J}^{-1}(\mu) \subset (\mathbf{J}^\mu)^{-1}(\mu'). \tag{2.2.1}$$

Since the actions are free and proper, μ and μ' are regular values, so these sets are indeed smooth manifolds. Note that, by construction, μ' is G_μ-invariant.

There will be two key assumptions relevant to the embedding version of cotangent bundle reduction. Namely,

CBR1. *In the above setting, assume there is a G_μ-invariant one-form α_μ on Q with values in $(\mathbf{J}^\mu)^{-1}(\mu')$.*

and the condition (which by (2.2.1), is a stronger condition)

CBR2. *Assume that α_μ in* **CBR1** *takes values in $\mathbf{J}^{-1}(\mu)$.*

For $\xi \in \mathfrak{g}_\mu$ and $q \in Q$, notice that, under the condition **CBR1**,

$$(\mathbf{i}_{\xi_Q}\alpha_\mu)(q) = \langle \mathbf{J}(\alpha_\mu(q)), \xi \rangle = \langle \mu', \xi \rangle,$$

and so $\mathbf{i}_{\xi_Q}\alpha_\mu$ is a constant function on Q . Therefore, for $\xi \in \mathfrak{g}_\mu$,

$$\mathbf{i}_{\xi_Q}\mathbf{d}\alpha_\mu = \boldsymbol{\mathcal{L}}_{\xi_Q}\alpha_\mu - \mathbf{d}\mathbf{i}_{\xi_Q}\alpha_\mu = 0, \tag{2.2.2}$$

since the Lie derivative is zero by G_μ-invariance of α_μ. It follows that

There is a unique two-form β_μ on Q_μ such that

$$\pi^*_{Q,G_\mu}\beta_\mu = \mathbf{d}\alpha_\mu.$$

Since π_{Q,G_μ} is a submersion, β_μ is closed (it need not be exact). Let

$$B_\mu = \pi^*_{Q_\mu}\beta_\mu,$$

where $\pi_{Q_\mu} : T^*Q_\mu \to Q_\mu$ is (following our general conventions for maps) the cotangent bundle projection. Also, to avoid confusion with the canonical symplectic form Ω_{can} on T^*Q, we shall denote the canonical symplectic form on T^*Q_μ, the cotangent bundle of μ-shape space, by ω_{can}.

2.2.1 Theorem (Cotangent Bundle Reduction—Embedding Version)**.**

(i) *If condition* **CBR1** *holds, then there is a symplectic embedding*

$$\varphi_\mu : ((T^*Q)_\mu, \Omega_\mu) \to (T^*Q_\mu, \omega_{\mathrm{can}} - B_\mu), \tag{2.2.3}$$

onto a submanifold of T^*Q_μ covering the base Q/G_μ.

(ii) *The map φ_μ in* (i) *gives a symplectic diffeomorphism of $((T^*Q)_\mu, \Omega_\mu)$ onto $(T^*Q_\mu, \omega_{\mathrm{can}} - B_\mu)$ if and only if $\mathfrak{g} = \mathfrak{g}_\mu$.*

(iii) *If* **CBR2** *holds, then the image of φ_μ equals the vector subbundle $[T\pi_{Q,G_\mu}(V)]^\circ$ of T^*Q_μ, where $V \subset TQ$ is the vector subbundle consisting of vectors tangent to the G-orbits, that is, its fiber at $q \in Q$ equals $V_q = \{\xi_Q(q) \mid \xi \in \mathfrak{g}\}$, and $^\circ$ denotes the annihilator relative to the natural duality pairing between TQ_μ and T^*Q_μ.*

Remarks.

1. As was mentioned, a history of this result can be found in §1.3.

2. As we saw in Proposition 2.1.2 and will elaborate shortly, α_μ may be constructed satisfying condition **CBR1** from a connection on the μ–shape space bundle $\pi_{Q,G_\mu} : Q \to Q/G_\mu$ and an α_μ satisfying **CBR2** can be constructed using a connection on the shape space bundle $\pi_{Q,G} : Q \to Q/G$.

3. Note that in the case of Abelian reduction, or, more generally, the case in which $G = G_\mu$, the reduced space is symplectically diffeomorphic to $T^*(Q/G)$ with the symplectic structure given by $\Omega_{\mathrm{can}} - B_\mu$. In particular, if $\mu = 0$, then the symplectic form on $T^*(Q/G)$ is the canonical one, since in this case one can choose $\alpha_\mu = 0$ which yields $B_\mu = 0$.

4. The term B_μ on T^*Q is usually called a **magnetic term**, a **gyroscopic term**, or a **Coriolis term**. The terminology "magnetic" comes from the Hamiltonian description of a particle of charge e moving according to the Lorentz force law in $\mathbb{R}^3$ under the influence of a magnetic field B. This motion takes place in $T^*\mathbb{R}^3$ but with the nonstandard symplectic structure $\mathbf{d}q^i \wedge \mathbf{d}p_i - \frac{e}{c}B$, $i = 1, 2, 3$, where c is the speed of light and B is regarded as a closed two-form: $B = B_x \mathbf{d}y \wedge \mathbf{d}z - B_y \mathbf{d}x \wedge \mathbf{d}z + B_z \mathbf{d}x \wedge \mathbf{d}y$ (see [MandS], §6.7 for details).

The strategy for proving this theorem is to first deal with the case of reduction at zero and then to treat the general case using a momentum shift.

Reduction at Zero.   The reduced space at $\mu = 0$ is, as a set,

$$(T^*Q)_0 = \mathbf{J}^{-1}(0)/G$$

since, for $\mu = 0$, $G_\mu = G$. Notice that in this case, there is no distinction between orbit reduction and symplectic reduction.

2.2.2 Theorem (Reduction at Zero).   *Assume that the action of G on Q is free and proper, so that the quotient Q/G is a smooth manifold. Then 0 is a regular value of $\mathbf{J}$ and there is a symplectic diffeomorphism between $(T^*Q)_0$ and $T^*(Q/G)$ with its canonical symplectic structure.*

Proof.   Since the action $\Phi : G \times Q \to Q$ of G on Q is free, so is the (left) lifted action of G on T^*Q. Thus, since all phase space points have no symmetry, they are all regular. The action of G being proper on Q implies that it is proper on T^*Q (and hence on $\mathbf{J}^{-1}(0)$) as well. Thus, $(T^*Q)_0$ and $T^*(Q/G)$ are smooth symplectic manifolds. Following our general notation, let $\pi_{Q,G} : Q \to Q/G$ denote the canonical projection.

Define the smooth map

$$\overline{\varphi}_0 : \mathbf{J}^{-1}(0) = \{\alpha_q \in T^*Q \mid \langle \alpha_q, \xi_Q(q) \rangle = 0 \ \ \text{for all} \ \ \xi \in \mathfrak{g}\} \to T^*(Q/G)$$

by

$$\langle \overline{\varphi}_0(\alpha_q), T_q\pi_{Q,G}(v_q) \rangle := \langle \alpha_q, v_q \rangle \qquad (2.2.4)$$

for any $v_q \in T_qQ$. This map is well–defined for if $v_q, v_q' \in T_qQ$ are such that $T_q\pi_{Q,G}(v_q) = T_q\pi_{Q,G}(v_q')$, there is some $\xi \in \mathfrak{g}$ such that $v_q' - v_q = \xi_Q(q)$ and hence $\langle \alpha_q, v_q' \rangle = \langle \alpha_q, v_q + \xi_Q(q) \rangle = \langle \alpha_q, v_q \rangle$, since $\langle \alpha_q, \xi_Q(q) \rangle = 0$ because $\alpha_q \in \mathbf{J}^{-1}(0)$.

We claim that the map $\overline{\varphi}_0$ is G-invariant. Indeed, for any $g \in G$ and $v_q \in T_qQ$ we have $T_{g \cdot q}\pi_{Q,G}(g \cdot v_q) = T_q\pi_{Q,G}(v_q)$ and hence

$$\langle \overline{\varphi}_0(g \cdot \alpha_q), T_q\pi_{Q,G}(v_q) \rangle = \langle \overline{\varphi}_0(g \cdot \alpha_q), T_{g \cdot q}\pi_{Q,G}(g \cdot v_q) \rangle = \langle g \cdot \alpha_q, g \cdot v_q \rangle$$
$$= \langle \alpha_q, v_q \rangle = \langle \overline{\varphi}_0(\alpha_q), T_q\pi_{Q,G}(v_q) \rangle,$$

by the definitions of the (left) lifted G-actions to TQ and T^*Q, respectively. This shows that $\overline{\varphi}_0(g \cdot \alpha_q) = \overline{\varphi}_0(\alpha_q)$ for any $\alpha_q \in \mathbf{J}^{-1}(0)$ and $g \in G$.

Next, we claim that $\overline{\varphi}_0$ is surjective. To see this, observe that if $\Gamma_{[q]} \in T^*_{[q]}(Q/G)$, where $[q] := \pi_{Q,G}(q)$, and if we define $\alpha_q \in \mathbf{J}^{-1}(0)$ by $\langle \alpha_q, v_q \rangle := \langle \Gamma_{[q]}, T_q\pi_{Q,G}(v_q) \rangle$, then $\overline{\varphi}_0(\alpha_q) = \Gamma_{[q]}$.

Thus, $\overline{\varphi}_0$ induces a smooth surjective map $\varphi_0 : (T^*Q)_0 \to T^*(Q/G)$ uniquely characterized by the relation $\varphi_0 \circ \pi_0 = \overline{\varphi}_0$, where $\pi_0 : \mathbf{J}^{-1}(0) \to \mathbf{J}^{-1}(0)/G$ is the canonical projection.

We now assert that φ_0 is injective. Let $\alpha_q, \alpha_{q'} \in \mathbf{J}^{-1}(0)$ be such that $\overline{\varphi}_0(\alpha_q) = \varphi_0(\pi_0(\alpha_q)) = \varphi_0(\pi_0(\alpha_{q'})) = \overline{\varphi}_0(\alpha_{q'})$. In particular, this implies that there is some $g \in G$ such that $q' = g \cdot q$. Since $\mathbf{J}(g \cdot \alpha_q) = \mathrm{Ad}^*_{g^{-1}} \mathbf{J}(\alpha_q) = 0$ by equivariance of $\mathbf{J}$ and the fact that $\alpha_q \in \mathbf{J}^{-1}(0)$, we conclude that $g \cdot \alpha_q, \alpha_{q'} \in \mathbf{J}^{-1}(0) \cap T^*_{q'}Q$. Thus, by G-invariance of $\overline{\varphi}_0$ it follows that the relation $\overline{\varphi}_0(\alpha_q) = \overline{\varphi}_0(\alpha_{q'})$ is equivalent to $\overline{\varphi}_0(g \cdot \alpha_q) = \overline{\varphi}_0(\alpha_{q'})$, which states that $\langle g \cdot \alpha_q, v_{q'} \rangle = \langle \alpha_{q'}, v_{q'} \rangle$ for all $v_{q'} \in T_{q'}Q$. Therefore $\alpha_{q'} = g \cdot \alpha_q$ and hence $\pi_0(\alpha_{q'}) = \pi_0(\alpha_q)$ thus proving that φ_0 is injective.

Since we have shown that $\varphi_0 : (T^*Q)_0 \to T^*(Q/G)$ is both injective and surjective, it is a smooth bijective map.

To show that φ_0 is canonical, by the unique characterization of the reduced symplectic form, it suffices to show that

$$\pi_0^* \varphi_0^* \theta_{\mathrm{can}} = i_0^* \Theta_{\mathrm{can}}, \qquad (2.2.5)$$

where

(a) θ_{can} is the canonical one-form on $T^*(Q/G)$,

(b) Θ_{can} is the canonical one-form on T^*Q,

(c) $i_0 : \mathbf{J}^{-1}(0) \hookrightarrow T^*Q$ is the inclusion map, and

(d) $\pi_0 : \mathbf{J}^{-1}(0) \to \mathbf{J}^{-1}(0)/G$ is the projection.

Indeed, taking the exterior derivative of (2.2.5) gives

$$\pi_0^* \varphi_0^* \omega_{\mathrm{can}} = i_0^* \Omega_{\mathrm{can}}, \tag{2.2.6}$$

where ω_{can} is the canonical symplectic form on $T^*(Q/G)$ and Ω_{can} the one on T^*Q. Equation (2.2.6) implies, by the Symplectic Reduction Theorem 1.1.3, that

$$\varphi_0^* \omega_{\mathrm{can}} = \Omega_0, \tag{2.2.7}$$

where Ω_0 is the reduced symplectic form on $(T^*Q)_0$.

It remains to prove (2.2.5). Denote by $\pi_{Q/G} : T^*(Q/G) \to Q/G$ the cotangent bundle projection. Let $\alpha_q \in \mathbf{J}^{-1}(0)$ and $v \in T_{\alpha_q} \mathbf{J}^{-1}(0)$. Then $\varphi_0 \circ \pi_0 = \overline{\varphi}_0$, the definition of the canonical one-form, the relation $\pi_{Q/G} \circ \overline{\varphi}_0 = \pi_{Q,G} \circ \pi_Q \circ i_0$, and (2.2.4), show that

$$
\begin{aligned}
\langle (\pi_0^* \varphi_0^* \theta_{\mathrm{can}})(\alpha_q), v \rangle &= \langle (\overline{\varphi}_0^* \theta_{\mathrm{can}})(\alpha_q), v \rangle = \langle \theta_{\mathrm{can}}(\overline{\varphi}_0(\alpha_q)), T_{\alpha_q} \overline{\varphi}_0(v) \rangle \\
&= \langle \overline{\varphi}_0(\alpha_q), T_{\overline{\varphi}_0(\alpha_q)} \pi_{Q/G} \left(T_{\alpha_q} \overline{\varphi}_0(v) \right) \rangle = \langle \overline{\varphi}_0(\alpha_q), T_{\alpha_q}(\pi_{Q/G} \circ \overline{\varphi}_0)(v) \rangle \\
&= \langle \overline{\varphi}_0(\alpha_q), T_{\alpha_q}(\pi_{Q,G} \circ \pi_Q \circ i_0)(v) \rangle = \langle \overline{\varphi}_0(\alpha_q), T_q \pi_{Q,G} \left(T_{\alpha_q} \pi_Q(v) \right) \rangle \\
&= \langle \alpha_q, T_{\alpha_q} \pi_Q(v) \rangle = \langle i_0^* \Theta(\alpha_q), v \rangle,
\end{aligned}
$$

which proves (2.2.5).

At this point we know that

$$\varphi_0 : (T^*Q)_0 \to T^*(Q/G)$$

is a smooth symplectic bijective map. Since every symplectic map is immersive, we conclude that φ_0 is an immersion. Finally, note that

$$\dim(T^*Q)_0 = 2\dim Q - 2\dim G = \dim T^*(Q/G),$$

which implies that φ_0 has bijective derivative and therefore is a local diffeomorphism. Hence φ_0 is a bijective local diffeomorphism and thus it is diffeomorphism. ∎

Next, we prove an additional important special case of the general cotangent bundle reduction theorem that, for example, includes the case of Abelian reduction. The key assumption here is that $G = G_\mu$, which indeed is always the case if G is Abelian.

2.2.3 Theorem.   *Assume that the action of G on Q is free and proper, so that the quotient Q/G is a smooth manifold. Let $\mu \in \mathfrak{g}^*$, assume that $G = G_\mu$, and assume that* **CBR2** *holds. Then μ is a regular value of $\mathbf{J}$ and there is a symplectic diffeomorphism between $(T^*Q)_\mu$ and $T^*(Q/G)$, the latter with the symplectic form $\omega_{\mathrm{can}} - B_\mu$; here, ω_{can} is the canonical symplectic form on $T^*(Q/G)$ and $B_\mu = \pi_{Q/G}^* \beta_\mu$, where the two form β_μ on Q/G is defined by*

$$\pi_{Q,G}^* \beta_\mu = \mathbf{d}\alpha_\mu.$$

Proof.   Note that in this case, the hypothesis **CBR2** on α_μ is that it is G-invariant and that $\alpha_\mu(Q) \subset \mathbf{J}^{-1}(\mu)$.

First, introduce the shift map

$$\mathrm{Shift}_\mu : T^*Q \to T^*Q,$$

which is the diffeomorphism defined by

$$\mathrm{Shift}_\mu(\alpha_q) := \alpha_q - \alpha_\mu(q) \tag{2.2.8}$$

and note that Shift_μ gives a G-equivariant diffeomorphism

$$\mathrm{shift}_\mu : \mathbf{J}^{-1}(\mu) \to \mathbf{J}^{-1}(0);$$

this is simply the statement that if α_q has momentum μ, then $\mathrm{shift}_\mu(\alpha_q)$ has momentum value zero. By construction we have the following commutative diagram; that is, the identity $i_0 \circ \mathrm{shift}_\mu = \mathrm{Shift}_\mu \circ i_\mu$.

$$
\begin{array}{ccc}
\mathbf{J}^{-1}(\mu) & \xrightarrow{\;\;i_\mu\;\;} & T^*Q \\[2pt]
\Big\downarrow{\scriptstyle \mathrm{shift}_\mu} & & \Big\downarrow{\scriptstyle \mathrm{Shift}_\mu} \\[2pt]
\mathbf{J}^{-1}(0) & \xrightarrow{\;\;i_0\;\;} & T^*Q
\end{array}
$$

Second, by the momentum shifting lemma (see [MandS] §6.6), which is, in fact, a simple calculation, we have

$$(\mathrm{Shift}_\mu)^*\Omega_{\mathrm{can}} = \Omega_{\mathrm{can}} + \pi_Q^* \mathbf{d}\alpha_\mu.$$

Letting i_μ^* act on each side and using the diagram above, we get

$$\mathrm{shift}_\mu^* \, i_0^* \Omega_{\mathrm{can}} = i_\mu^*(\Omega_{\mathrm{can}} + \pi_Q^* \mathbf{d}\alpha_\mu). \tag{2.2.9}$$

Third, by G-equivariance of shift_μ, we get an induced diffeomorphism

$$\widetilde{\mathrm{shift}}_\mu : \mathbf{J}^{-1}(\mu)/G \to \mathbf{J}^{-1}(0)/G.$$

Notice that it is at this point that we are using the hypothesis that $G = G_\mu$.

Now define the diffeomorphism

$$\varphi_\mu := \varphi_0 \circ \widetilde{\mathrm{shift}}_\mu : \mathbf{J}^{-1}(\mu)/G \to T^*(Q/G),$$

where $\varphi_0 : \mathbf{J}^{-1}(0)/G \to T^*(Q/G)$ is the diffeomorphism defined in the proof of Theorem 2.2.2. To show that $\varphi_\mu^*(\omega_{\mathrm{can}} - B_\mu) = \Omega_\mu$, we shall

check that $\varphi_\mu^*(\omega_{\mathrm{can}} - B_\mu)$ satisfies the property *uniquely* characterizing Ω_μ, namely

$$\pi_\mu^* \varphi_\mu^* (\omega_{\mathrm{can}} - B_\mu) = i_\mu^* \Omega_{\mathrm{can}}. \tag{2.2.10}$$

We shall use the definition of φ_μ and the identity $\widetilde{\mathrm{shift}}_\mu \circ \pi_\mu = \pi_0 \circ \mathrm{shift}_\mu$ to rewrite the left hand side of (2.2.10) as follows:

$$\begin{aligned}
(\varphi_\mu \circ \pi_\mu)^*(\omega_{\mathrm{can}} - B_\mu) &= (\varphi_0 \circ \widetilde{\mathrm{shift}}_\mu \circ \pi_\mu)^*(\omega_{\mathrm{can}} - B_\mu) \\
&= (\varphi_0 \circ \pi_0 \circ \mathrm{shift}_\mu)^*(\omega_{\mathrm{can}} - B_\mu) \\
&= \mathrm{shift}_\mu^* \pi_0^* \varphi_0^*(\omega_{\mathrm{can}} - B_\mu).
\end{aligned} \tag{2.2.11}$$

However, $B_\mu = \pi_{Q/G}^* \beta_\mu$ and $\varphi_0^* \omega_{\mathrm{can}} = \Omega_0$ (see (2.2.7)), so (2.2.11) becomes

$$\mathrm{shift}_\mu^* \pi_0^* \Omega_0 - \mathrm{shift}_\mu^* \pi_0^* \varphi_0^* \pi_{Q/G}^* \beta_\mu. \tag{2.2.12}$$

But

$$\mathrm{shift}_\mu^* \pi_0^* \Omega_0 = \mathrm{shift}_\mu^* i_0^* \Omega_{\mathrm{can}} = i_\mu^*(\Omega_{\mathrm{can}} + \pi_Q^* \mathbf{d}\alpha_\mu)$$

by (2.2.9). Also, from the results on reduction at zero, we have the identity $\pi_{Q/G} \circ \varphi_0 \circ \pi_0 = \pi_{Q,G} \circ \pi_Q \circ i_0$ and so the last term in (2.2.12) becomes

$$\mathrm{shift}_\mu^* \pi_0^* \varphi_0^* \pi_{Q/G}^* \beta_\mu = \mathrm{shift}_\mu^* i_0^* \pi_Q^* \pi_{Q,G}^* \beta_\mu = \mathrm{shift}_\mu^* i_0^* \pi_Q^* \mathbf{d}\alpha_\mu = i_\mu^* \pi_Q^* \mathbf{d}\alpha_\mu$$

since $\pi_Q \circ i_\mu = \pi_Q \circ i_0 \circ \mathrm{shift}_\mu$. Thus (2.2.12) becomes

$$i_\mu^*(\Omega_{\mathrm{can}} + \pi_Q^* \mathbf{d}\alpha_\mu) - i_\mu^* \pi_Q^* \mathbf{d}\alpha_\mu = i_\mu^* \Omega_{\mathrm{can}},$$

as required.   ∎

Proof of Theorem 2.2.1.   We shall prove (i) and (ii) together. As already mentioned, the strategy is to reduce the problem to reduction in the case in which $G = G_\mu$, which was dealt with in the preceding theorem.

From the inclusion of submanifolds (2.2.1), we get a symplectic inclusion of reduced symplectic manifolds

$$\mathbf{J}^{-1}(\mu)/G_\mu \subset (\mathbf{J}^\mu)^{-1}(\mu')/G_\mu; \tag{2.2.13}$$

the manifold $(\mathbf{J}^\mu)^{-1}(\mu')/G_\mu$ is the reduction at μ' of T^*Q for the G_μ-action since μ' is G_μ-invariant; that is,

$$(G_\mu)_{\mu'} = G_\mu. \tag{2.2.14}$$

We assert that the inclusion (2.2.13) is an equality precisely when $\mathfrak{g} = \mathfrak{g}_\mu$. Indeed, if $\mathfrak{g} = \mathfrak{g}_\mu$ then clearly $\mathbf{J}^{-1}(\mu) = (\mathbf{J}^\mu)^{-1}(\mu')$ and hence, by G_μ-invariance of both $\mathbf{J}^{-1}(\mu)$ and $(\mathbf{J}^\mu)^{-1}(\mu')$, we get

$$\mathbf{J}^{-1}(\mu)/G_\mu = (\mathbf{J}^\mu)^{-1}(\mu')/G_\mu. \tag{2.2.15}$$

Now assume, conversely, that (2.2.15) holds. It is a general and elementary fact that if a group K acts on two sets A and B, with $A \subset B$ and if $A/K = B/K$, then $A = B$. Thus, from the inclusion (2.2.1), their G_μ-invariance, and the equality (2.2.15), we have $\mathbf{J}^{-1}(\mu) = (\mathbf{J}^\mu)^{-1}(\mu')$. Now let $\alpha_q \in \mathbf{J}^{-1}(\mu) = (\mathbf{J}^\mu)^{-1}(\mu')$. The Reduction Lemma 1.2.2 (iv) applied to the G-momentum map $\mathbf{J}$ yields $\ker T_{\alpha_q}\mathbf{J} = (T_{\alpha_q}(G \cdot \alpha_q))^\Omega$, whereas applied to the G_μ-momentum map $\mathbf{J}^\mu$ gives $\ker T_{\alpha_q}\mathbf{J}^\mu = (T_{\alpha_q}(G_\mu \cdot \alpha_q))^\Omega$. Since the manifolds $\mathbf{J}^{-1}(\mu)$ and $(\mathbf{J}^\mu)^{-1}(\mu')$ are equal, so are their tangent spaces at α_q and therefore, taking Ω-orthogonal complements yields

$$\{\xi_Q(q) \mid \xi \in \mathfrak{g}\} = T_{\alpha_q}(G \cdot \alpha_q) = T_{\alpha_q}(G_\mu \cdot \alpha_q) = \{\eta_Q(q) \mid \eta \in \mathfrak{g}_\mu\}.$$

Freeness of the G-action (actually, local freeness suffices here) implies then that $\mathfrak{g} = \mathfrak{g}_\mu$.

Since the inclusion (2.2.13) is symplectic, it is necessarily an immersion and since (2.2.1) is an inclusion of submanifolds, so is (2.2.13). Therefore (2.2.13) is an embedding.

Applying Theorem 2.2.3, with the group being G_μ, with its momentum map $\mathbf{J}^\mu$, at the momentum value μ', and with $\alpha_{\mu'}$ chosen to be α_μ (satisfying **CBR1**), we see that $\left((\mathbf{J}^\mu)^{-1}(\mu')/G_\mu, \Omega_{\mu'}\right)$ is symplectically diffeomorphic to $(T^*(Q/G_\mu), \omega_{\mathrm{can}} - B_\mu)$. The hypotheses of this theorem hold by virtue of (2.2.14).

Composing the above symplectic diffeomorphism with the symplectic embedding $\left(\mathbf{J}^{-1}(\mu)/G_\mu, \Omega_\mu\right) \subset \left((\mathbf{J}^\mu)^{-1}(\mu')/G_\mu, \Omega_{\mu'}\right)$ gives a symplectic embedding

$$\varphi_\mu : ((T^*Q)_\mu, \Omega_\mu) \to (T^*(Q/G_\mu), \omega_{\mathrm{can}} - B_\mu).$$

The map φ_μ is explicitly given as follows. Denote by $[\alpha_q]_\mu$ the elements of the reduced space $(T^*Q)_\mu = \mathbf{J}^{-1}(\mu)/G_\mu$ and by $[\alpha_q]_0$ the elements of the reduced space $(\mathbf{J}^\mu)^{-1}(0)/G_\mu$. Then φ_μ is given by

$$\varphi_\mu([\alpha_q]_\mu) = \varphi_0([\alpha_q - \alpha_\mu(q)]_0) = \overline{\varphi}_0(\alpha_q - \alpha_\mu(q)). \qquad (2.2.16)$$

or, explicitly

$$\left\langle \varphi_\mu([\alpha_q]_\mu), T_q\pi_{Q,G_\mu}(v_q) \right\rangle = \langle \alpha_q - \alpha_\mu(q), v_q \rangle \qquad (2.2.17)$$

for any $v_q \in T_qQ$.

The image of the embedded submanifold $\varphi_\mu((T^*Q)_\mu)$ in $T^*(Q/G_\mu)$ obviously projects by π_{Q,G_μ} onto the μ–shape space $Q/G_\mu = Q_\mu$ (see (2.2.16)). Thus, we have proved (i) and (ii).

Now we turn to part (iii), which assumes **CBR2**; that is, $\alpha_\mu(Q) \subset \mathbf{J}^{-1}(\mu)$. First we give some needed preliminaries.

Let $V \subset TQ$ be the vertical bundle of the G-principal bundle $\pi_{Q,G} : Q \to Q/G$, that is, V is the vector subbundle of TQ whose fiber at each

$q \in Q$ equals $V_q = \{\xi_Q(q) \mid \xi \in \mathfrak{g}\}$. Since $T_q\pi_{Q,G_\mu} : V_q \to T_{\pi_{Q,G_\mu}(q)}(Q/G_\mu)$ has kernel $\{\eta_Q(q) \mid \eta \in \mathfrak{g}_\mu\}$ whose dimension does not depend on $q \in Q$ (because the G-action is free on Q), it follows that the rank of $T_q\pi_{Q,G_\mu}|_V$ is constant in q. Therefore, both the kernel and the range of $T\pi_{Q,G_\mu}|_V$ are subbundles of V and $T(Q/G_\mu)$ respectively. Thus the annihilator $[T\pi_{Q,G_\mu}(V)]^\circ$ is a vector subbundle of $T^*(Q/G_\mu)$.

To prove (iii) we shall show that

$$\varphi_\mu((T^*Q)_\mu) = [T\pi_{Q,G_\mu}(V)]^\circ. \tag{2.2.18}$$

Indeed, if $\alpha_q \in \mathbf{J}^{-1}(\mu)$ we have by (2.2.16) and (2.2.4) for any $\xi \in \mathfrak{g}$

$$
\begin{aligned}
\langle \varphi_\mu([\alpha_q]_\mu), T_q\pi_{Q,G_\mu}(\xi_Q(q)) \rangle &= \langle \overline{\varphi}_0(\alpha_q - \alpha_\mu(q)), T_q\pi_{Q,G_\mu}(\xi_Q(q)) \rangle \\
&= \langle \alpha_q, \xi_Q(q) \rangle - \langle \alpha_\mu(q), \xi_Q(q) \rangle = \langle \mathbf{J}(\alpha_q), \xi \rangle - \langle \mathbf{J}(\alpha_\mu(q)), \xi_Q(q) \rangle \\
&= \langle \mu, \xi \rangle - \langle \mu, \xi \rangle = 0.
\end{aligned}
$$

Note that the hypothesis **CBR2**, namely $\alpha_\mu(Q) \subset \mathbf{J}^{-1}(\mu)$, was essential in this argument and that the previous reasoning would not have worked had we assumed only **CBR1**, namely α_μ has values in $(\mathbf{J}^\mu)^{-1}(\mu')$. This shows that the image of φ_μ lies in $[T\pi_{Q,G_\mu}(V)]^\circ$. Conversely, if $\Gamma_{[q]} \in [T\pi_{Q,G_\mu}(V)]^\circ$ define $\alpha_q \in T_q^*Q$ by

$$\langle \alpha_q, v_q \rangle := \langle \Gamma_{[q]}, T_q\pi_{Q,G_\mu}(v_q) \rangle + \langle \alpha_\mu(q), v_q \rangle$$

for any $v_q \in T_qQ$. We claim that $\alpha_q \in \mathbf{J}^{-1}(\mu)$. Indeed, for any $\xi \in \mathfrak{g}$ we have

$$
\begin{aligned}
\langle \mathbf{J}(\alpha_q), \xi \rangle &= \langle \alpha_q, \xi_Q(q) \rangle \\
&= \langle \Gamma_{[q]}, T_q\pi_{Q,G_\mu}(\xi_Q(q)) \rangle + \langle \alpha_\mu(q), \xi_Q(q) \rangle \\
&= 0 + \langle \mathbf{J}(\alpha_\mu(q)), \xi \rangle \\
&= \langle \mu, \xi \rangle.
\end{aligned}
$$

From the definition of α_q it immediately follows that $\varphi_\mu([\alpha_q]_\mu) = \Gamma_{[q]}$, which shows that $[T\pi_{Q,G_\mu}(V)]^\circ$ is contained in the range of φ_μ. This proves (2.2.18). ∎

Example.   Consider the reduction of a general cotangent bundle T^*Q by $G = \mathrm{SO}(3)$. Here $G_\mu \cong S^1$, if $\mu \neq 0$, and so the reduced space is embedded into the cotangent bundle $T^*(Q/S^1)$. A specific example is the case of $Q = \mathrm{SO}(3)$. Then the reduced space $(T^*\mathrm{SO}(3))_\mu$ is $S^2_{\|\mu\|}$, the sphere of radius $\|\mu\|$ which is a coadjoint orbit in $\mathfrak{so}(3)^*$. In this case, $Q/G_\mu = \mathrm{SO}(3)/S^1 \cong S^2_{\|\mu\|}$ and the embedding of $S^2_{\|\mu\|}$ into $T^*S^2_{\|\mu\|}$ is the zero section.   ♦

Magnetic Terms and Curvature.   Using the results of the preceding section, we will now show how one can interpret the magnetic term B_μ as the curvature of a connection on a principal bundle.

We saw in the preamble to the Cotangent Bundle Reduction Theorem 2.2.1 that $\mathbf{i}_{\xi_Q} \mathbf{d}\alpha_\mu = 0$ for any $\xi \in \mathfrak{g}_\mu$, which was used to drop $\mathbf{d}\alpha_\mu$ to the quotient. In the language of principal bundles, this may be rephrased by saying that $\mathbf{d}\alpha_\mu$ is *horizontal* and thus, once a connection is introduced, *the covariant exterior derivative of α_μ coincides with $\mathbf{d}\alpha_\mu$.*

There are two methods to construct a form α_μ with the properties in Theorem 2.2.1. We continue to work under the general assumption that G acts on Q freely and properly.

First Method.   *Construction of α_μ from a connection $\mathcal{A}^\mu \in \Omega^1(Q; \mathfrak{g}_\mu)$ on the principal bundle $\pi_{Q,G_\mu} : Q \to Q/G_\mu$.*

To carry this out, we show that the choice

$$\alpha_\mu := \langle \mu', \mathcal{A}^\mu \rangle \in \Omega^1(Q)$$

satisfies the condition **CBR1** in Theorem 2.2.1, where, as above, $\mu' = \mu|_{\mathfrak{g}_\mu}$. We show first that α_μ is G_μ-invariant. Indeed, for any $g \in G_\mu$ and $v_q \in T_q Q$ we have

$$(\Phi_g^* \alpha_\mu)(q)(v_q) = \alpha_\mu(g \cdot q)(g \cdot v_q) = \langle \mu', \mathcal{A}^\mu(g \cdot q)(g \cdot v_q) \rangle$$
$$= \langle \mu', (\Phi_g^* \mathcal{A}^\mu)(q)(v_q) \rangle = \langle \mu', \mathrm{Ad}_g \, \mathcal{A}^\mu(q)(v_q) \rangle.$$

However, since $g \in G_\mu$ and $\mathcal{A}^\mu(q)(v_q) \in \mathfrak{g}_\mu$ it follows that $\mathrm{Ad}_g \, \mathcal{A}^\mu(q)(v_q) \in \mathfrak{g}_\mu$ and since $\mu' = \mu|_{\mathfrak{g}_\mu}$, we conclude that

$$\langle \mu', \mathrm{Ad}_g \, \mathcal{A}^\mu(q)(v_q) \rangle = \langle \mu, \mathrm{Ad}_g \, \mathcal{A}^\mu(q)(v_q) \rangle = \langle \mathrm{Ad}_g^* \mu, \mathcal{A}^\mu(q)(v_q) \rangle$$
$$= \langle \mu, \mathcal{A}^\mu(q)(v_q) \rangle = \alpha_\mu(q)(v_q)$$

because $g \in G_\mu$. This shows that $\Phi_g^* \alpha_\mu = \alpha_\mu$ for any $g \in G_\mu$. Second, we show that α_μ takes values in $(\mathbf{J}^\mu)^{-1}(\mu')$. Indeed, for any $\xi \in \mathfrak{g}_\mu$, we get

$$\langle \mathbf{J}^\mu(\alpha_\mu(q)), \xi \rangle = \langle \alpha_\mu(q), \xi_Q(q) \rangle = \langle \mu', \mathcal{A}^\mu(q)(\xi_Q(q)) \rangle = \langle \mu', \xi \rangle.$$

Thus this one-form α_μ satisfies the condition **CBR1** in Theorem 2.2.1.

Next, we shall give an interpretation of the two-form $\mathbf{d}\alpha_\mu$ in terms of curvature. Recall from (2.1.8) that the curvature two-form of $\mathcal{A}^\mu$ is given by

$$\mathcal{B}^\mu(q)(u_q, v_q) = \mathbf{d}\mathcal{A}^\mu(q)(\mathrm{hor}_q(u_q), \mathrm{hor}_q(v_q))$$

so that the μ'-component of $\mathcal{B}^\mu$ equals

$$\langle \mu', \mathcal{B}^\mu \rangle(q)(u_q, v_q) = \mathbf{d}\langle \mu', \mathcal{A}^\mu \rangle(q)(\mathrm{hor}_q(u_q), \mathrm{hor}_q(v_q))$$
$$= \mathbf{d}\langle \mu', \mathcal{A}^\mu \rangle(q)\big(u_q - [\mathcal{A}^\mu(q)(u_q)]_Q(q), v_q - [\mathcal{A}^\mu(q)(v_q)]_Q(q)\big)$$
$$= \mathbf{d}\alpha_\mu(q)\big(u_q - [\mathcal{A}^\mu(q)(u_q)]_Q(q), v_q - [\mathcal{A}^\mu(q)(v_q)]_Q(q)\big)$$
$$= \mathbf{d}\alpha_\mu(q)\,(u_q, v_q)$$

since each of the other three terms is of the form

$$\mathbf{d}\alpha_\mu(q)(\zeta_Q(q), w_q) = (\mathbf{i}_{\zeta_Q}\mathbf{d}\alpha_\mu)(q)(w_q) = 0$$

by (2.2.2), where $\zeta \in \mathfrak{g}_\mu$. Thus, $\mathbf{d}\alpha_\mu$ is the μ'-component of the curvature two-form. We summarize these results in the following statement.

2.2.4 Proposition.   *If the principal bundle* $\pi_{Q,G_\mu} : Q \to Q/G_\mu$ *with structure group* G_μ *has a connection* $\mathcal{A}^\mu$, *then* $\alpha_\mu(q)$ *can be taken to equal* $\mathcal{A}^\mu(q)^*\mu'$ *and* B_μ *is induced on* T^*Q_μ *by* $\mathbf{d}\alpha_\mu$ *(a two-form on Q), which equals the* μ'-*component of the curvature* $\mathcal{B}^\mu$ *of* $\mathcal{A}^\mu$.

Second Method.   *Construction of* α_μ *from a connection* $\mathcal{A} \in \Omega^1(Q; \mathfrak{g})$ *on the principal bundle* $\pi_{Q,G} : Q \to Q/G$.

Next, we show that the choice (2.1.1), that is,

$$\alpha_\mu := \langle \mu, \mathcal{A} \rangle \in \Omega^1(Q)$$

satisfies the condition **CBR2** in Theorem 2.2.1. By Proposition 2.1.2 we have for any $g \in G$ the identity $\Phi_g^*\alpha_\mu = \alpha_{\mathrm{Ad}_g^*\mu}$. Therefore, if $g \in G_\mu$ we get $\Phi_g^*\alpha_\mu = \alpha_\mu$. In addition, Proposition 2.1.2 states that this one-form α_μ takes values in $\mathbf{J}^{-1}(\mu)$. Thus this one-form α_μ satisfies the condition **CBR2** of Theorem 2.2.1.

As with the first method, we shall now give an interpretation of the two-form $\mathbf{d}\alpha_\mu$ in terms of curvature.

Unlike the previous case, $\mathbf{d}\alpha_\mu$ is *not* the μ-component of the curvature of $\mathcal{A}$ because the reasoning above used formula (2.2.2) that is valid only for elements in $\mathfrak{g}_\mu$. On the other hand, $\mathbf{d}\alpha_\mu$ does have an expression involving the curvature that is a direct consequence of the Cartan Structure Equations (2.1.10). In fact, pairing $\mathcal{B} = \mathbf{d}\mathcal{A} - [\mathcal{A}, \mathcal{A}]$ with $\mu \in \mathfrak{g}^*$ and recalling that $\alpha_\mu = \langle \mu, \mathcal{A} \rangle$ we get

$$\mathbf{d}\alpha_\mu = \langle \mu, \mathcal{B} \rangle + \langle \mu, [\mathcal{A}, \mathcal{A}] \rangle .$$

We summarize these results in the following statement.

2.2.5 Proposition.   *If the principal bundle* $\pi_{Q,G} : Q \to Q/G$ *with structure group* G *has a connection* $\mathcal{A}$, *then* $\alpha_\mu(q)$ *can be taken to equal* $\mathcal{A}(q)^*\mu$ *and* B_μ *is the pull back to* T^*Q_μ *of* $\mathbf{d}\alpha_\mu \in \Omega^2(Q)$, *which equals the* μ-*component of the two form* $\mathcal{B} + [\mathcal{A}, \mathcal{A}] \in \Omega^2(Q; \mathfrak{g})$, *where* $\mathcal{B}$ *is the curvature of* $\mathcal{A}$.

Coadjoint Orbits.   We now apply the Cotangent Bundle Reduction Theorem 2.2.1 to the case $Q = G$ and with the G-action given by *left* translation. The *right* Maurer-Cartan form θ^R is a flat connection associated to this action (see Theorem 2.1.14) and hence

$$\mathbf{d}\alpha_\mu(g)(u_g, v_g) = \left\langle \mu, [\theta^R, \theta^R](g)(u_g, v_g) \right\rangle = \left\langle \mu, [T_g R_{g^{-1}} u_g, T_g R_{g^{-1}} v_g] \right\rangle .$$

Recall from Theorem 1.2.3 that the reduced space $(T^*G)_\mu$ is the coadjoint orbit $\mathcal{O}_\mu$ endowed with the negative orbit symplectic form ω_μ^- and, according to the Cotangent Bundle Reduction Theorem, it symplectically embeds as the zero section into $(T^*\mathcal{O}_\mu, \omega_{\text{can}} - B_\mu)$, where $B_\mu = \pi_{\mathcal{O}_\mu}^* \beta_\mu$, $\pi_{\mathcal{O}_\mu} : T^*\mathcal{O}_\mu \to \mathcal{O}_\mu$ is the cotangent bundle projection, $\pi_{G,G_\mu}^* \beta_\mu = \mathbf{d}\alpha_\mu$, and $\pi_{G,G_\mu} : G \to \mathcal{O}_\mu$ is given by $\pi_{G,G_\mu}(g) = \text{Ad}_g^* \mu$. The derivative of π_{G,G_μ} is given by

$$T_g \pi_{G,G_\mu}(T_e L_g \xi) = \left.\frac{d}{dt}\right|_{t=0} \text{Ad}_{g\exp(t\xi)}^* \mu = \text{ad}_\xi^* \text{Ad}_g^* \mu$$

for any $\xi \in \mathfrak{g}$.

Therefore, if we let $\nu = \text{Ad}_g^* \mu \in \mathcal{O}_\mu$, $\text{ad}_\xi^* \nu, \text{ad}_\eta^* \nu \in T_\nu \mathcal{O}_\mu$, we get

$$
\begin{aligned}
\beta_\mu(\nu)\left(\text{ad}_\xi^* \nu, \text{ad}_\eta^* \nu\right) &= \left(\pi_{G,G_\mu}^* \beta_\mu\right)(g)\left(T_e L_g \xi, T_e L_g \eta\right) \\
&= \mathbf{d}\alpha_\mu(g)\left(T_e L_g \xi, T_e L_g \eta\right) = \left\langle \mu, [T_g R_{g^{-1}} T_e L_g \xi, T_g R_{g^{-1}} T_e L_g \eta]\right\rangle \\
&= \left\langle \mu, [\text{Ad}_g \xi, \text{Ad}_g \eta]\right\rangle = \left\langle \mu, \text{Ad}_g[\xi, \eta]\right\rangle = \left\langle \nu, [\xi, \eta]\right\rangle \\
&= -\omega_\mu^-(\nu)\left(\text{ad}_\xi^* \nu, \text{ad}_\eta^* \nu\right)
\end{aligned}
$$

which shows that $\beta_\mu = -\omega_\mu^-$. Thus, the embedding version of the cotangent bundle reduction theorem produces the following statement which, of course, can be easily checked directly.

2.2.6 Corollary.   *The coadjoint orbit $(\mathcal{O}_\mu, \omega_\mu^-)$ symplectically embeds as the zero section into the symplectic manifold $\left(T^*\mathcal{O}_\mu, \omega_{\text{can}} + \pi_{\mathcal{O}_\mu}^* \omega_\mu^-\right)$.*

2.3 Cotangent Bundle Reduction: Bundle Version

The embedding version of the cotangent bundle reduction theorem presented in the preceding section states that $(T^*Q)_\mu$ *embeds* as a vector subbundle of $T^*(Q/G_\mu)$. The bundle version of this theorem says, roughly speaking, that $(T^*Q)_\mu$ is a *coadjoint orbit bundle* over $T^*(Q/G)$ with fiber the coadjoint orbit $\mathcal{O}$ through μ.

Statement.   We state this version as follows. Again we will utilize a choice of connection $\mathcal{A}$ on the shape space bundle $\pi_{Q,G} : Q \to Q/G$. A key step in the argument is to utilize orbit reduction and the identification $(T^*Q)_\mu \cong (T^*Q)_\mathcal{O}$.

2.3.1 Theorem (Cotangent Bundle Reduction—Bundle Version).   *The reduced space $(T^*Q)_\mu$ is a locally trivial fiber bundle over $T^*(Q/G)$ with typical fiber $\mathcal{O}$.*

Sketch of Proof.   The proof proceeds in a manner similar to the previous arguments in a number of steps. We sketch these below.

Step 1: **Reduction at Zero.** In Theorem 2.2.2, we have shown that reduction at zero is given by

$$(T^*Q)_0 \cong T^*(Q/G). \tag{2.3.1}$$

Here $(T^*Q)_0 = \mathbf{J}^{-1}(0)/G$ and the symplectic form on $T^*(Q/G)$ is the canonical one.

Step 2: **Orbit Reduction.**   Theorem 1.2.4 shows that the reduced space $(T^*Q)_\mu = \mathbf{J}^{-1}(\mu)/G_\mu$ can be identified with the quotient $\mathbf{J}^{-1}(\mathcal{O})/G$, where $\mathcal{O}$ is the coadjoint orbit through μ.

Step 3: **Shifting.**   Use the ***shift map*** shift : $T^*Q \to T^*Q$ defined by

$$\text{shift}(\alpha_q) = \alpha_q - \mathcal{A}(q)^* \mathbf{J}(\alpha_q) \tag{2.3.2}$$

which restricts to define a map $\text{shift}_{\mathcal{O}} : \mathbf{J}^{-1}(\mathcal{O}) \to \mathbf{J}^{-1}(0)$.[1]

We assert that shift : $T^*Q \to T^*Q$ is equivariant with respect to the G-action. To see this, let $h \in G$, $v_q \in T_qQ$, and let $\Phi_h : Q \to Q$ denote the action by h. Now use the definitions to compute the h-action on the second term in the shift map:

$$\begin{aligned}
\langle T_q^* \Phi_h(\mathcal{A}(h \cdot q)^* \mathbf{J}(h \cdot \alpha_q)), v_q \rangle &= \langle \mathcal{A}(h \cdot q)^* \mathbf{J}(h \cdot \alpha_q)), T_q \Phi_h(v_q) \rangle \\
&= \langle \mathbf{J}(h \cdot \alpha_q), \mathcal{A}(h \cdot q)(T_q \Phi_h(v_q)) \rangle \\
&= \langle \mathrm{Ad}_{h^{-1}}^* \mathbf{J}(\alpha_q), \mathrm{Ad}_h \, \mathcal{A}(v_q) \rangle \\
&= \langle \mathcal{A}(q)^* \mathbf{J}(\alpha_q), v_q \rangle.
\end{aligned}$$

Thus, $T_q^* \Phi_h(\mathcal{A}(h \cdot q)^* \mathbf{J}(h \cdot \alpha_q)) = \mathcal{A}(q)^* \mathbf{J}(\alpha_q)$, which in other notation reads

$$h \cdot (\mathcal{A}(q)^* \mathbf{J}(\alpha_q)) = \mathcal{A}(h \cdot q)^* \mathbf{J}(h \cdot \alpha_q))$$

This computation proves our assertion that shift is equivariant.

Thus, $\text{shift}_{\mathcal{O}} : \mathbf{J}^{-1}(\mathcal{O}) \to \mathbf{J}^{-1}(0)$ drops to the quotient, producing the desired map

$$\widetilde{\text{shift}}_{\mathcal{O}} : \mathbf{J}^{-1}(\mathcal{O})/G \to \mathbf{J}^{-1}(0)/G = T^*(Q/G). \tag{2.3.3}$$

As we saw in orbit reduction (see Theorem 1.2.4), $(T^*Q)_\mu$ is symplectically diffeomorphic to $\mathbf{J}^{-1}(\mathcal{O})/G$, where $\mathcal{O}$ is the coadjoint orbit through μ, so one gets a map of $(T^*Q)_\mu$ to $T^*(Q/G)$.

---

[1] If we have a Riemannian manifold (invariant under the group action), and use it to identify TQ and T^*Q, the shift map is nothing more than the horizontal projection map for the connection.

We claim that the map $\widetilde{\mathrm{shift}}_{\mathcal{O}}$ has fiber that is diffeomorphic to $\mathcal{O}$, which then would prove the theorem.

To prove the claim, start with two points $[\alpha_q]_G, [\beta_r]_G \in \mathbf{J}^{-1}(\mathcal{O})/G$, that map to the same point in $\mathbf{J}^{-1}(0)/G$ by $\widetilde{\mathrm{shift}}_{\mathcal{O}}$. Thus, there is a $g \in G$ such that $g \cdot \mathrm{shift}_{\mathcal{O}}(\alpha_q) = \mathrm{shift}_{\mathcal{O}}(\beta_r)$. That is,

$$g \cdot (\alpha_q - \mathcal{A}(q)^* \mathbf{J}(\alpha_q)) = \beta_r - \mathcal{A}(q)^* \mathbf{J}(\beta_r)$$

which is the same as

$$g \cdot \alpha_q - \mathcal{A}(g \cdot q)^* \mathbf{J}(g \cdot \alpha_q) = \beta_r - \mathcal{A}(r)^* \mathbf{J}(\beta_r).$$

Since the base points are equal, we have $g \cdot q = r$, so the preceding equation becomes

$$g \cdot \alpha_q - \mathcal{A}(r)^* \mathbf{J}(g \cdot \alpha_q) = \beta_r - \mathcal{A}(r)^* \mathbf{J}(\beta_r).$$

Since $[\alpha_q]_G, [\beta_r]_G \in \mathbf{J}^{-1}(\mathcal{O})/G$, and $\mathbf{J}$ is equivariant, there is a group element $h \in G$ such that

$$\mathbf{J}(\alpha_q) = \mathrm{Ad}_{h^{-1}}^* \mathbf{J}(\beta_r) =: h \cdot \mathbf{J}(\beta_r).$$

Letting $\mu = \mathbf{J}(\beta_r)$, we get hence

$$g \cdot (\alpha_q - \mathcal{A}(q)^*(h \cdot \mu)) = \beta_r - \mathcal{A}(r)^* \mu.$$

Thus, we have shown that the classes $[\alpha_q]_G, [\beta_r]_G$ have representatives that have the same base point r and that these representatives differ by quantities parametrized by elements of $\mathcal{O}$, namely the element $h \cdot \mu \in \mathcal{O}$. Regarding β_r as fixed in a local trivialization, we see that the classes $[\alpha_q]_G$ that map to the same point are parametrized by elements of $\mathcal{O}$. ∎

This point of view is explored further and the exact nature of the coadjoint orbit bundle is identified and its symplectic structure will be elucidated shortly.

Poisson Version. This same type of argument as above shows the following, which we state slightly informally.

2.3.2 Theorem. *The Poisson reduced space $(T^*Q)/G$ is diffeomorphic to the coadjoint bundle of $\pi_{Q,G} : Q \to Q/G$. This diffeomorphism is implemented by a connection $\mathcal{A} \in \Omega^1(Q; \mathfrak{g})$. Thus the fiber of $(T^*Q)/G \to T^*(Q/G)$ is isomorphic to the Lie-Poisson space $\mathfrak{g}^*$.*

There is an interesting formula for the Poisson structure on $(T^*Q)/G$ that was originally computed in Montgomery, Marsden and Ratiu [1984], Montgomery [1986]. Further developments in Cendra, Marsden, Pekarsky, and Ratiu [2003] and Perlmutter and Ratiu [2005] gives a unified study of the Poisson bracket on $(T^*Q)/G$ in both the Sternberg and Weinstein realizations of the quotient. Finally, we refer to, for instance, Lewis, Marsden, Montgomery, and Ratiu [1986] for an application of this result; in this case, the dynamics of fluid systems with free boundaries is studied.

Coadjoint Orbit Bundles.   The purpose of the remainder of this section is to give the details of the nature of the bundle and its associated symplectic structure that was sketched in Theorem 2.3.1. We follow the exposition in Marsden and Perlmutter [2000]; see also Zaalani [1999], Cushman and Śniatycki [1999], and Perlmutter and Ratiu [2005]. The rest of this section can be omitted on a first reading and is included mainly for completeness.

As above, we let a Lie group G act freely and properly on a manifold Q so that the natural quotient map $\pi : Q \rightarrow Q/G$ defines a principal bundle. Let $\mathcal{A}$ be a principal connection this bundle and let $\tilde{\mathfrak{g}}$ denote the associated bundle to the Lie algebra $\mathfrak{g}$, namely $\tilde{\mathfrak{g}} = (Q \times \mathfrak{g})/G$, which we regard as a vector bundle over Q/G. We recall the following natural bundle isomorphisms (see Cendra, Holm, Marsden and Ratiu [1998]):

2.3.3 Lemma.   *There are bundle isomorphisms*

$$\alpha_{\mathcal{A}} : TQ/G \rightarrow T(Q/G) \oplus \tilde{\mathfrak{g}} \quad and \quad (\alpha_{\mathcal{A}}^{-1})^* : T^*Q/G \rightarrow T^*(Q/G) \oplus \tilde{\mathfrak{g}}^*$$

$$(2.3.4)$$

Proof.   Given $v_q \in T_qQ$, denote its equivalence class in $T(Q/G)$ by $[v_q]$. We claim that the mapping $\alpha_{\mathcal{A}} : [v_q] \mapsto T_q\pi(v_q) \oplus [q, \mathcal{A}(q)(v_q)]$ is well defined and induces the desired isomorphism of TQ/G with $T(Q/G) \oplus \tilde{\mathfrak{g}}$. To see this, consider another representative of the orbit $[v_q]$, given by $g \cdot v_q$ where we use concatenated notation for the tangent lifted action of G on TQ. We have $T_{g\cdot q}\pi(g \cdot v_q) = T_q\pi(v_q)$ and

$$[g \cdot q, \mathcal{A}(g \cdot q)(g \cdot v_q)] = [g \cdot q, (\phi_g^*\mathcal{A})] = [g \cdot q, \mathrm{Ad}_g \mathcal{A}(q)(v_q)]$$
$$= [q, \mathcal{A}(q)(v_q)].$$

The inverse of this map is given by $v_{[q]} \oplus [q, \xi] \mapsto [\mathrm{hor}_q v_{[q]} + \xi_Q(q)]$ as is readily verified. We therefore have a vector bundle isomorphism.

We next compute $(\alpha_{\mathcal{A}}^{-1})^* : T^*Q/G \rightarrow T^*(Q/G) \oplus \tilde{\mathfrak{g}}^*$, the dual of the inverse map:

$$\left\langle (\alpha_{\mathcal{A}}^{-1})^*([\alpha_q]), (u_{[q]}, [q, \xi])) \right\rangle = \left\langle [\alpha_q], [\mathrm{hor}_q \cdot u_{[q]} + \xi_Q(q)] \right\rangle$$
$$= \left\langle \alpha_q, \mathrm{hor}_q \cdot u_{[q]} \right\rangle + \left\langle \alpha_q, \xi_Q(q) \right\rangle$$
$$= \left\langle \mathrm{hor}_q^* \alpha_q, u_{[q]} \right\rangle + \left\langle \mathbf{J}(\alpha_q), \xi \right\rangle,$$

where $\mathrm{hor}_q^* : T_q^*Q \rightarrow T_{[q]}^*(Q/G)$ is dual to the horizontal lift map $\mathrm{hor}_q : T_{[q]}(Q/G) \rightarrow T_qQ$ so that we conclude $(\alpha_{\mathcal{A}}^{-1})^*([\alpha_q]) = (\mathrm{hor}_q^* \alpha_q, [q, \mathbf{J}(\alpha_q)])$. ∎

This bundle isomorphism can be recast as follows (see Cushman and Śniatycki [1999]). Consider the maps

$$\Delta : T^*Q \rightarrow T^*Q/G \rightarrow \tilde{\mathfrak{g}}^*; \quad \alpha_q \mapsto [\alpha_q] \mapsto [q, \mathbf{J}(\alpha_q)]$$

and
$$\Gamma : T^*Q \to T^*Q/G \to T^*(Q/G); \quad \alpha_q \mapsto [\alpha_q] \mapsto \mathrm{hor}_q^* \alpha_q.$$

Notice that the map $\alpha_q \mapsto \mathrm{hor}_q^* \alpha_q$ drops to T^*Q/G, since we have, for all $V_{[q]} \in T_{[q]}(Q/G)$,

$$\begin{aligned}
\left\langle \mathrm{hor}_{g \cdot q}^*(g \cdot \alpha_q), V_{[q]} \right\rangle &= \left\langle g \cdot \alpha_q, \mathrm{hor}_{g \cdot q} V_{[q]} \right\rangle = \left\langle \alpha_q, g^{-1} \cdot \mathrm{hor}_{g \cdot q} V_{[q]} \right\rangle \\
&= \left\langle \alpha_q, g^{-1} \cdot (g \cdot \mathrm{hor}_q V_{[q]} \right\rangle = \left\langle \alpha_q, \mathrm{hor}_q V_{[q]} \right\rangle
\end{aligned}$$

where we use the fact that $g \cdot \mathrm{hor}_q = \mathrm{hor}_{g \cdot q}$.

We can write $(\alpha_{\mathcal{A}}^{-1})^* = \Gamma \oplus \Delta$. A partial inverse to the projection Δ is given in the next proposition,

2.3.4 Proposition.   *Consider the map,*

$$\sigma : Q \times \mathfrak{g}^* \to T^*Q/G \tag{2.3.5}$$

given by $(q, \nu) \mapsto \mathcal{A}(q)^* \nu$. This map is equivariant with respect to the diagonal action of G on $Q \times \mathfrak{g}^*$ and the cotangent lifted action of G on T^*Q, and so uniquely defines a map on the quotient,

$$\widetilde{\sigma} : \widetilde{\mathfrak{g}}^* \to T^*Q/G \tag{2.3.6}$$

This is a fiber preserving bundle map which is injective on each fiber and satisfies $\Delta \circ \widetilde{\sigma} = id|_{\widetilde{\mathfrak{g}}^*}$

Proof.   Under the map σ, $g \cdot (q, \nu) = (g \cdot q, \mathrm{Ad}_{g^{-1}}^* \nu) \mapsto \mathcal{A}(g \cdot q)^*(\mathrm{Ad}_{g^{-1}}^* \nu)$. However, for all $v \in T_{g \cdot q} Q$,

$$\begin{aligned}
\left\langle \mathcal{A}(g \cdot q)^*(\mathrm{Ad}_{g^{-1}}^* \nu, v \right\rangle &= \left\langle \mathrm{Ad}_{g^{-1}}^* \nu, \mathcal{A}(g \cdot q) v \right\rangle = \left\langle \nu, \mathrm{Ad}_{g^{-1}} \mathcal{A}(g \cdot q) v \right\rangle \\
&= \left\langle \nu, (\psi_{g^{-1}})^* \mathcal{A}(g \cdot q) v \right\rangle = \left\langle \nu, \mathcal{A}(q)(T\psi_{g^{-1}} v) \right\rangle \\
&= \left\langle g \cdot \mathcal{A}(q)^* \nu, v \right\rangle,
\end{aligned}$$

from which we conclude equivariance of σ. Also, for each $[q, \nu] \in \widetilde{\mathfrak{g}}^*$,

$$\Delta(\widetilde{\sigma}([q, \nu])) = \Delta([\mathcal{A}(q)^* \nu]) = [q, \mathbf{J}(\mathcal{A}(q)^* \nu)] = [q, \nu],$$

since, for all $\xi \in \mathfrak{g}$, $\langle \mathbf{J}(\mathcal{A}(q)^* \nu), \xi \rangle = \langle \nu, \mathcal{A}(q)(\xi_Q(q)) \rangle = \langle \nu, \xi \rangle$.   ∎

We next determine the image, under $(\alpha_{\mathcal{A}}^{-1})^*$ of the symplectic leaves of the Poisson manifold T^*Q/G, which we know from the symplectic correspondence theorem (see Weinstein [1983a], Blaom [2000]) are given by $\mathbf{J}^{-1}(\mathcal{O})/G$ for each coadjoint orbit $\mathcal{O}$ in $\mathfrak{g}^*$.

2.3.5 Theorem.   *We have*

$$(\alpha_{\mathcal{A}}^{-1})^*(\mathbf{J}^{-1}(\mathcal{O})/G) = T^*(Q/G) \times_{Q/G} \widetilde{\mathcal{O}},$$

where $\widetilde{\mathcal{O}} = (Q \times \mathcal{O})/G$ is the associated bundle using the coadjoint action of G on $\mathcal{O}$.

Proof.   From the definition of the bundle isomorphism $\alpha_{\mathcal{A}}$,

$$(\alpha_{\mathcal{A}}^{-1})^*(\mathbf{J}^{-1}(\mathcal{O})/G) = \{(\Gamma(\alpha_q), \Delta(\alpha_q)) \mid \mathbf{J}(\alpha_q) \in \mathcal{O}\}$$
$$= \{(\mathrm{hor}_q^* \alpha_q, [q, \mathbf{J}(\alpha_q)]) \mid \mathbf{J}(\alpha_q) \in \mathcal{O}\}$$

First we characterize the sets $T_q^* Q \cap \mathbf{J}^{-1}(\mathcal{O})$, using the connection $\mathcal{A}$. Denote by $\mathbf{J}_q$, the restriction of $\mathbf{J}$ to $T_q^* Q$, and let $\sigma_q : \mathfrak{g} \mapsto T_q Q$, be the injective infinitesimal generator map. Using the fact that $\mathbf{J}_q = \sigma_q^*$, we have for all $\xi \in \mathfrak{g}$

$$\langle \mathbf{J}(\alpha_q + \mathcal{A}_\mu(q)), \xi \rangle = \langle \alpha_q, \xi_Q(q) \rangle + \langle \mu, \xi \rangle = \langle \mu, \xi \rangle \qquad (2.3.7)$$

where the second equality holds since $\alpha_q \in V^0$, the annihilator of the vertical sub-bundle of TQ. We conclude that

$$T_q^* Q \cap \mathbf{J}^{-1}(\mathcal{O}) = \{V_q^0 + \mathcal{A}_\mu(q) \mid \mu \in \mathcal{O}\}.$$

Recall that since σ_q^* is surjective, $\tau \circ \mathbf{J}^{-1}(\mathcal{O}) = Q$, where $\tau_{T^*Q} : T^*Q \to Q$ is the cotangent bundle projection. Now apply hor_q^* to each fiber over Q in $\mathbf{J}^{-1}(\mathcal{O})$. That is, for each $q \in Q$, we consider

$$\mathrm{hor}_q^*(\mathbf{J}^{-1}(\mathcal{O}) \cap T_q^* Q) \qquad (2.3.8)$$

First, note that for all $X_{[q]} \in T_{[q]}(Q/G)$,

$$\left\langle \mathrm{hor}_q^*(\mathcal{A}_\mu(q)), X_{[q]} \right\rangle = \left\langle \mu, \mathcal{A}(q)(\mathrm{hor}_q(X_{[q]})) \right\rangle = 0 \qquad (2.3.9)$$

so that $\mathrm{hor}_q^*(\mathcal{A}_\mu(q)) = 0$. Furthermore, since hor_q is injective, $\mathrm{hor}_q^* : T_q^* Q \to T_{[q]}^*(Q/G)$ is surjective with $\ker \mathrm{hor}_q^* = H^0$, where H^0 denotes the annihilator of the horizontal subbundle of TQ. Thus, as a linear map, $\mathrm{hor}_q^* : V^0 \to T_{[q]}^*(Q/G)$ is an isomorphism. Consider the set of pairs, $\{(\Gamma(\alpha_q), \Delta(\alpha_q)) \mid \mathbf{J}(\alpha_q) \in \mathcal{O}\}$. Each α_q can be uniquely expressed as $\beta_q + \mathcal{A}_\mu(q)$ for some $\beta_q \in V^0$ and $\mu \in \mathcal{O}$. For a fixed μ, let β_q range over V_q^0. This generates the set $T_{[q]}^*(Q/G) \times [q, \mu]$ since $\mathbf{J}$ vanishes on V^0. The result now follows by varying $\mu \in \mathcal{O}$.   ∎

Orbit Reduction.   Recall from Theorem 1.2.4 the reduction construction in terms of coadjoint orbits. Assume μ is a regular value of an equivariant momentum map $\mathbf{J} : P \to \mathfrak{g}^*$ of a left symplectic action of G on the symplectic manifold (P, Ω) and assume that the symplectic reduced space P_μ is a manifold with π_μ a submersion. Let $\mathcal{O}$ be the coadjoint orbit through μ in $\mathfrak{g}_+^*$. Then the inclusion map induces a symplectic diffeomorphism from P_μ to the orbit reduced symplectic manifold $(\mathbf{J}^{-1}(\mathcal{O})/G, \Omega_\mathcal{O})$, where the symplectic form $\Omega_\mathcal{O}$ is determined by

$$\iota_\mathcal{O}^* \Omega = \pi_\mathcal{O}^* \Omega_\mathcal{O} + \mathbf{J}_\mathcal{O}^* \omega_\mathcal{O}^+, \qquad (2.3.10)$$

where $\iota_{\mathcal{O}} : \mathbf{J}^{-1}(\mathcal{O}) \to P$ is the inclusion, $\mathbf{J}_{\mathcal{O}} = \mathbf{J}|\mathbf{J}^{-1}(\mathcal{O})$, and $\omega_{\mathcal{O}}^+$ is the "+" orbit symplectic structure on $\mathcal{O}$. Recall that Theorem 1.2.4 explains the sense in which the orbit reduced space $\mathbf{J}^{-1}(\mathcal{O})/G$ is symplectically diffeomorphic to the point reduced space P_μ. Note that if one assumes that $\mathbf{J}^{-1}(\mathcal{O})$ is closed, then it is a submanifold in the usual sense; otherwise it is an initial submanifold.

By considering a momentum shift we can realize a bundle isomorphism between $\mathbf{J}^{-1}(\mathcal{O})$ and the space $V^0 \times \mathcal{O}$. Since it will be shown that this isomorphism is G equivariant, it determines a unique diffeomorphism between $V^0 \times \mathcal{O}/G$ and $\mathbf{J}^{-1}(\mathcal{O})/G$. We will characterize the symplectic form on the former by pulling back the "characterizing" symplectic form on $\mathbf{J}^{-1}(\mathcal{O})$. Furthermore, it will be shown that $V^0 \times \mathcal{O}/G$ is diffeomorphic to $T^*(Q/G) \oplus \widetilde{\mathcal{O}}$, so that the reduced symplectic form can be expressed on this space as well. Since $\mathbf{J}^{-1}(\mathcal{O}) \subset T^*Q$, the reduced symplectic form is determined by the restriction of the canonical symplectic form in T^*Q to $\mathbf{J}^{-1}(\mathcal{O})$, which in turn is determined by the restriction of the canonical one-form to $\mathbf{J}^{-1}(\mathcal{O})$.

2.3.6 Lemma.   *There is a G-equivariant bundle isomorphism,*

$$\chi : V^0 \times \mathcal{O} \to \mathbf{J}^{-1}(\mathcal{O}) \tag{2.3.11}$$

that uniquely determines a diffeomorphism, $\bar\chi$ on the quotient spaces so that the following diagram commutes

Proof.   The map χ is given by

$$\chi(\alpha_q, \nu) = \alpha_q + \mathcal{A}(q)^*\nu \tag{2.3.12}$$

This map takes values in $\mathbf{J}^{-1}(\mathcal{O})$, as is seen from the proof of the previous theorem. From the characterization of the fibers of the bundle $\mathbf{J}^{-1}(\mathcal{O}) \to Q$, it follows that this map is onto. Since it is a momentum shift, it is clearly invertible with inverse

$$\alpha_q \mapsto \alpha_q - \mathcal{A}(q)^*\mathbf{J}(\alpha_q) \tag{2.3.13}$$

We check G-equivariance as follows:

$$\chi(g \cdot (\alpha_q, \nu)) = \chi(g \cdot \alpha_q, g \cdot \nu) = g \cdot \alpha_q + \mathcal{A}(g \cdot q)^*(g \cdot \nu)$$
$$= g \cdot \alpha_q + g \cdot (\mathcal{A}(q)^* \nu) = g \cdot (\chi(\alpha_q, \nu)$$

where the third equality uses the invariance properties of the connection.
∎

Because χ is G-equivariant, the pull back by χ of the G-invariant form on $\mathbf{J}^{-1}(\mathcal{O})$, $\pi_{\mathcal{O}}^* \Omega_{\mathcal{O}}$, given by $\chi^*(\pi_{\mathcal{O}}^* \Omega_{\mathcal{O}})$, is a G-invariant form on $V^0 \times \mathcal{O}$. In fact, the form drops to the quotient by the diagonal G action since

$$\chi^*(\pi_{\mathcal{O}}^* \Omega_{\mathcal{O}}) = \widetilde{\pi}_G^* \bar{\chi}^* \Omega_{\mathcal{O}}, \tag{2.3.14}$$

where $\widetilde{\pi}_G : V^0 \times \mathcal{O} \to V^0 \times \mathcal{O}/G$ denotes the projection. This follows since the diagram in the preceding theorem is commutative.

The Two-form on $V^0 \times \mathcal{O}$.   We proceed to characterize the structure of the two-form $\chi^*(\pi_{\mathcal{O}}^* \Omega_{\mathcal{O}})$ on $V^0 \times \mathcal{O}$. By construction,

$$\chi^*(\pi_{\mathcal{O}}^* \Omega_{\mathcal{O}}) = \chi^* \iota_{\mathcal{O}}^* \Omega - \chi^* \mathbf{J}_{\mathcal{O}}^* \omega_{\mathcal{O}}^+ \tag{2.3.15}$$

We claim that the *second term* is

$$\chi^* \mathbf{J}_{\mathcal{O}}^* \omega_{\mathcal{O}}^+ = \pi_2^* \omega_{\mathcal{O}}^+, \tag{2.3.16}$$

where $\pi_2 : V^0 \times \mathcal{O} \to \mathcal{O}$ is projection on the second factor. This follows since for all $\xi \in \mathfrak{g}$,

$$\langle \mathbf{J}_{\mathcal{O}} \circ \chi(\alpha_q, \nu), \xi \rangle = \langle \alpha_q + \mathcal{A}(q)^* \nu, \xi_Q(q) \rangle = \langle \alpha_q, \xi_Q(q) \rangle + \langle \mathcal{A}(q)^* \nu, \xi_Q(q) \rangle$$
$$= 0 + \langle \nu, \xi \rangle,$$

so that $\mathbf{J}_{\mathcal{O}} \circ \chi = \pi_2$.

The *first term* in (2.3.15) is a little more complicated and itself splits into a sum of terms. We begin by considering the pull back by χ of the restriction to $\mathbf{J}^{-1}(\mathcal{O})$ of the canonical one-form, and then we compute the exterior derivative of this one-form.

2.3.7 Lemma.   *We have*

$$\chi^* \iota_{\mathcal{O}}^* \Theta = \pi_1^* \iota_{V^0}^* \Theta + \varpi, \tag{2.3.17}$$

where $\iota_{V^0} : V^0 \to T^*Q$ is inclusion, Θ is the canonical one-form on T^*Q, $\pi_1 : V^0 \times \mathcal{O} \to V^0$ is projection on the first factor, and $\varpi \in \Omega^1(V^0 \times \mathcal{O})$ is given by

$$\varpi(\alpha_q, \nu)(X_{\alpha_q}, X_\nu^{\xi'}) = \langle \nu, \mathcal{A}(q)(T_{\alpha_q} \tau_{T^*Q}(X_{\alpha_q})) \rangle \tag{2.3.18}$$

for $(X_{\alpha_q}, X_\nu^{\xi'}) \in T_{(\alpha_q, \nu)}(V^0 \times \mathcal{O})$, where $X_\nu^{\xi'} \in T_\nu \mathcal{O}$ denotes the infinitesimal generator for the left action of G on $\mathcal{O}$, $X_\nu^{\xi'} = -\operatorname{ad}_{\xi'}^* \nu$.

Proof. Let $t \mapsto (\alpha_q(t), \mathrm{Ad}^*_{\exp -t\xi'}\nu)$ be a curve in $V^0 \times \mathcal{O}$, through the point (α_q, ν) such that $\frac{d}{dt}\big|_{t=0}\alpha_q(t) = X_{\alpha_q} \in T_{\alpha_q}V^0$. Since

$$\frac{d}{dt}\bigg|_{t=0} \mathrm{Ad}^*_{\exp -t\xi'}\nu = X_\nu^{\xi'},$$

we get

$$\chi^*\iota_{\mathcal{O}}^*\Theta(\alpha_q,\nu)(X_{\alpha_q}, X_\nu^{\xi'}) = \iota_{\mathcal{O}}^*\Theta(\alpha_q + \mathcal{A}(q)^*\nu)(T_{(\alpha_q,\nu)}\chi(X_{\alpha_q}, X_\nu^{\xi'}) \quad (2.3.19)$$

Now,

$$\begin{aligned}
T_{(\alpha_q,\nu)}\chi(X_{\alpha_q}, X_\nu^{\xi'}) &= \frac{d}{dt}\bigg|_{t=0}(\chi(\alpha_q(t), \mathrm{Ad}^*_{\exp -t\xi'}\nu)) \\
&= \frac{d}{dt}\bigg|_{t=0}(\alpha_q(t) + \mathcal{A}(q)^*(\mathrm{Ad}^*_{\exp -t\xi'}\nu)) \\
&= X_{\alpha_q} - \mathcal{A}(q)^*(\mathrm{ad}^*_{\xi'}\nu)
\end{aligned}$$

where we have used the fact that the curve $t \mapsto \mathcal{A}(q)^*(\mathrm{Ad}^*_{\exp -t\xi'}\nu)$ lies in the single fiber, T_q^*Q for all t. Thus, the right hand side in (2.3.19) becomes

$$\begin{aligned}
\iota_{\mathcal{O}}^*\Theta(\alpha_q &+ \mathcal{A}(q)^*\nu)(T_{(\alpha_q,\nu)}\chi(X_{\alpha_q}, X_\nu^{\xi'})) \\
&= \Theta(\alpha_q + \mathcal{A}(q)^*\nu)(T_{(\alpha_q,\nu)}\chi(X_{\alpha_q}, X_\nu^{\xi'})) \\
&= \Theta(\alpha_q + \mathcal{A}(q)^*\nu)(X_{\alpha_q} - \mathcal{A}(q)^*(\mathrm{ad}^*_{\xi'}\nu)) \\
&= \langle \alpha_q + \mathcal{A}(q)^*\nu, T\tau_{T^*Q}\cdot X_{\alpha_q}\rangle \\
&= \langle \alpha_q, T\tau_{T^*Q}\cdot X_{\alpha_q}\rangle + \langle \nu, \mathcal{A}(q)(T\tau_{T^*Q}\cdot X_{\alpha_q})\rangle \\
&= \pi_1^*\iota_{V^0}^*\Theta + \varpi
\end{aligned}$$

The third equality holds because, for all t,

$$\tau_{T^*Q}(\alpha_q(t) + \mathcal{A}(q)^*(\mathrm{Ad}^*_{\exp -t\xi'}\nu)) = \tau_{T^*Q}(\alpha_q(t)).$$

$\blacksquare$

Computing the exterior derivative,

$$\chi^*\iota_{\mathcal{O}}^*(-\mathbf{d}\Theta) = -\mathbf{d}(\pi_1^*\iota_{V^0}^*\Theta + \varpi) = \pi_1^*\iota_{V^0}^*\Omega - \mathbf{d}\varpi,$$

so that

$$\chi^*(\pi_{\mathcal{O}}^*\Omega_{\mathcal{O}}) = \chi^*\iota_{\mathcal{O}}^*\Omega - \chi^*\mathbf{J}_{\mathcal{O}}^*\omega_{\mathcal{O}}^+ = \pi_1^*\iota_{V^0}^*\Omega - \mathbf{d}\varpi - \pi_2^*\omega_{\mathcal{O}}^+.$$

Notice that the one-form ϖ defined in Lemma 2.3.7 is given by $\varpi = (\tau_{T^*Q} \times \mathrm{id})^*\alpha$, the pull back to $V^0 \times \mathcal{O}$ of the one-form α on $Q \times \mathcal{O}$ defined by

$$\alpha(q,\nu)(X_q, X_\nu) = \langle \nu, \mathcal{A}(q)(X_q)\rangle. \quad (2.3.20)$$

We are implicitly restricting the domain of τ_{T^*Q} to the sub-bundle V^0.

Computation of dα. The philosophy of the computation will be to make use of the connection to decompose tangent vectors to Q in terms of their horizontal and vertical parts. Of course we expect the curvature of the connection to appear in the resulting formula. However the presence of the pairing with ν, which varies over the coadjoint orbit $\mathcal{O}$ must be dealt with carefully.

We begin with an elementary but useful fact concerning the Jacobi-Lie bracket of vector fields on the cartesian product of two manifolds.

2.3.8 Lemma. *Let M and N be two smooth manifolds of dimension m and n respectively and consider their Cartesian product $M \times N$. Suppose we have two vector fields (X^M, X^N) and (Y^M, Y^N) on $M \times N$, each with the property that the tangent vector to M is independent of N, and that the tangent vector to N is independent of M. Then, the Jacobi Lie bracket of these two vector fields is also of this type. In fact, we have,*

$$[(X^M, X^N), (Y^M, Y^N)] = ([X^M, Y^M]_M, [X^N, Y^N]_N) \qquad (2.3.21)$$

This is readily proved using the local coordinate expression of the bracket.

To determine $\mathbf{d}\alpha \in \Omega^2(Q \times \mathcal{O})$, it suffices, by bilinearity and skew symmetry, to compute its value on pairs of tangent vectors to Q of the type

- $\mathrm{hor}_q, \mathrm{hor}_q$

- $\mathrm{hor}_q, \mathrm{ver}_q$

- $\mathrm{ver}_q, \mathrm{ver}_q$

To carry this out, we will extend each tangent vector to be horizontal or vertical in an entire neighborhood of the point in question and use the fact that $\mathbf{d}\alpha$ is a tensor.

Case 1. $X_q, Y_q \in \mathrm{Hor}_q Q.$ We consider

$$(X_q, X_\nu^{\xi'}), ((Y_q, Y_\nu^{\eta'}) \in T_{(q,\nu)}(Q \times \mathcal{O}).$$

Extend X_q to the horizontal vector field $\widetilde{X}_Q \in \mathrm{Hor}\, Q$ and similarly extend Y_q to $\widetilde{Y}_Q$. We extend the second components of each tangent vector in the obvious way to be infinitesimal generators of the given Lie algebra element. That is we extend $X_\nu^{\xi'}$ to $\xi'_\mathcal{O}$ and similarly for $Y_\nu^{\eta'}$. Denote by $\widetilde{X}$, the extended vector field on a neighborhood of $Q \times \mathcal{O}$ given by $(\widetilde{X}_Q, \xi'_\mathcal{O})$, and similarly for $\widetilde{Y}$. We then have,

$$\begin{aligned}
\mathbf{d}\alpha(q,\nu)&((X_q, X_\nu^{\xi'}), (Y_q, Y_\nu^{\eta'})) \\
&= (X_q, X_\nu^{\xi'}) \cdot \alpha(\widetilde{Y}_Q, \eta'_\mathcal{O}) - (Y_q, Y_\nu^{\eta'}) \cdot \alpha(\widetilde{X}_Q, \xi'_\mathcal{O}) - \alpha([\widetilde{X}, \widetilde{Y}])(q, \nu)
\end{aligned}$$
$$(2.3.22)$$

Notice that the first term vanishes since, if we take a curve $t \mapsto (q(t), \nu(t))$ through the point (q, ν) such that $(\dot{q}(0), \dot{\nu}(0)) = (X_q, X_\nu^{\xi'})$, we have

$$\frac{d}{dt}\bigg|_{t=0} \alpha(q(t), \nu(t))(\widetilde{Y}_Q(q(t)), \eta'_{\mathcal{O}}(\nu(t))) = \frac{d}{dt}\bigg|_{t=0} \left\langle \nu(t), \mathcal{A}(q(t)) \cdot \widetilde{Y}_Q(q(t)) \right\rangle$$
$$= 0 \qquad\qquad (2.3.23)$$

since for all t, $\widetilde{Y}_Q(q(t)) \in \mathrm{Hor}_{q(t)} Q$. Similarly, the second term vanishes. By Lemma 2.3.8, we have $[\widetilde{X}, \widetilde{Y}]_{Q \times \mathcal{O}} = ([\widetilde{X}_Q, \widetilde{Y}_Q]_Q, [\xi'_{\mathcal{O}}, \eta'_{\mathcal{O}}]_{\mathcal{O}})$ leaving

$$\mathbf{d}\alpha(q, \nu)((X_q, X_\nu^{\xi'}), (Y_q, Y_\nu^{\eta'})) = -\left\langle \nu, \mathcal{A}(q)([\widetilde{X}_Q, \widetilde{Y}_Q](q)) \right\rangle \qquad (2.3.24)$$

However, since $\widetilde{X}_Q$ and $\widetilde{Y}_Q$ are horizontal vector fields, it follows that

$$\mathcal{A}([\widetilde{X}_Q, \widetilde{Y}_Q]) = -\mathrm{Curv}_{\mathcal{A}}(\widetilde{X}_Q, \widetilde{Y}_Q) \qquad (2.3.25)$$

so that

$$\mathbf{d}\alpha(q, \nu)((X_q, X_\nu^{\xi'}), (Y_q, Y_\nu^{\eta'})) = \langle \nu, \mathrm{Curv}_{\mathcal{A}}(X_q, Y_q) \rangle. \qquad (2.3.26)$$

Case 2. $X_q \in \mathrm{Hor}_q Q, Y_q \in \mathrm{Ver}_q Q.$ Using the same notation for vector fields as in the previous case, we let $\widetilde{X}_Q$ denote the horizontal vector field extending X_q. Let $\eta = \mathcal{A}(q)(Y_q)$. Since Y_q is vertical we have $\eta_Q(q) = Y_q$. Then η_Q is a vertical extension of Y_q. With these extensions, we have

$$\mathbf{d}\alpha(q, \nu)(X_q, X_\nu^{\xi'}), (Y_q, Y_\nu^{\eta'})$$
$$= (X_q, X_\nu^{\xi'}) \cdot \alpha(\eta_Q, \eta'_{\mathcal{O}}) - (Y_q, Y_\nu^{\eta'}) \cdot \alpha((\widetilde{X}_Q, \xi'_{\mathcal{O}})) - \alpha(q, \nu)([\widetilde{X}, \widetilde{Y}])$$
$$(2.3.27)$$

Consider the first term. Let $t \mapsto (q(t), \nu(t))$ be a curve through (q, ν) with $(\dot{q}(0), \dot{\nu}(0)) = (X_q, -\mathrm{ad}_{\xi'}^* \nu)$. Then

$$(X_q, X_\nu^{\xi'}) \cdot \alpha(\eta_Q, \eta'_{\mathcal{O}}) = \frac{d}{dt}\bigg|_{t=0} \alpha(q(t), \nu(t))(\eta_Q(q(t)), \eta'_{\mathcal{O}}(\nu(t)))$$
$$= \frac{d}{dt}\bigg|_{t=0} \langle \nu(t), \mathcal{A}(q(t))(\eta_Q(q(t))) \rangle$$
$$= \frac{d}{dt}\bigg|_{t=0} \langle \nu(t), \eta \rangle = \left\langle -\mathrm{ad}_{\xi'}^* \nu, \eta \right\rangle$$

The second term vanishes since $\alpha(\widetilde{X}_Q, \xi'_{\mathcal{O}}) = 0$ for $\widetilde{X}_Q \in \mathrm{Hor}\, Q$. Recall that for $\widetilde{X}_Q$ a horizontal vector field, we have, for all $\eta \in \mathfrak{g}$,

$$[\widetilde{X}_Q, \eta_Q] \in \mathrm{Hor}\, Q \qquad (2.3.28)$$

This fact, together with Lemma 2.3.8, gives

$$\alpha(q,\nu)([\widetilde{X},\widetilde{Y}]) = \left\langle \nu, \mathcal{A}(q)([\widetilde{X}_Q, \eta_Q])\right\rangle = 0, \qquad (2.3.29)$$

so that

$$\mathbf{d}\alpha(q,\nu)(X_q, X_\nu^{\xi'}), (Y_q, Y_\nu^{\eta'}) = \left\langle -\operatorname{ad}_{\xi'}^* \nu, \eta\right\rangle. \qquad (2.3.30)$$

Case 3. $X_q, Y_q \in \operatorname{Ver}_q Q.$ Let $\xi = \mathcal{A}(q)(X_q)$ and $\eta = \mathcal{A}(q)(Y_q)$ We choose extensions that are globally vertical. Thus, $\widetilde{X}_Q = \xi_Q$ and $\widetilde{Y}_Q = \eta_Q$. Then we compute each term in the expression for $\mathbf{d}\alpha$.

The first term will again be

$$\left\langle -\operatorname{ad}_{\xi'}^* \nu, \eta\right\rangle \qquad (2.3.31)$$

since $\alpha(q,\nu)(\eta_Q, \eta'_{\mathcal{O}}) = \langle \nu, \eta\rangle$ i.e. $\iota_{(\eta_Q, \eta'_{\mathcal{O}})}\alpha : Q \times \mathcal{O} \to \mathbb{R}$ is independent of Q.

The second term is computed similarly to be $\left\langle \operatorname{ad}_{\eta'}^* \nu, \xi\right\rangle$. For the last term, recall that for left actions, $G \times Q \to Q$, we have

$$[\xi_Q, \eta_Q] = -[\xi, \eta]_Q \qquad (2.3.32)$$

so that

$$\alpha(q,\nu)([\widetilde{X},\widetilde{Y}]) = \langle \nu, \mathcal{A}(q)([\xi_Q, \eta_Q])\rangle = -\langle \nu, \mathcal{A}(q)([\xi,\eta]_Q)(q)\rangle = -\langle \nu, [\xi,\eta]\rangle.$$

Therefore,

$$\mathbf{d}\alpha(q,\nu)(X_q, X_\nu^{\xi'}), (Y_q, Y_\nu^{\eta'}) = \langle \nu, [\eta, \xi']\rangle + \langle \nu, [\eta', \xi]\rangle + \langle \nu, [\xi, \eta]\rangle. \quad (2.3.33)$$

We now collect these results to obtain a formula for the two form relative to a decomposition of the tangent vectors to Q into their horizontal and vertical projections.

2.3.9 Proposition.   *Let* $(X_q, X_\nu^{\xi'}), (Y_q, Y_\nu^{\eta'}) \in T_{(q,\nu)}(Q \times \mathcal{O})$*. Let*

$$\xi = \mathcal{A}(q)(X_q) \quad and \quad \eta = \mathcal{A}(q)(Y_q)$$

so that

$$X_q = \xi_Q(q) + \operatorname{Hor}_q X_q, \quad Y_q = \eta_Q(q) + \operatorname{Hor}_q Y_q \qquad (2.3.34)$$

where Hor_q *denotes the horizontal projection onto the horizontal distribution. We then have*

$$\mathbf{d}\alpha(q,\nu)((X_q, X_\nu^{\xi'}), (Y_q, Y_\nu^{\eta'}))$$
$$= \langle \nu, [\eta', \xi]\rangle + \langle \nu, [\eta, \xi']\rangle + \langle \nu, [\xi, \eta]\rangle + \langle \nu, \operatorname{Curv}_{\mathcal{A}}(q)(X_q, Y_q)\rangle \quad (2.3.35)$$

Proof.   The proof is a straightforward computation:

$$
\mathbf{d}\alpha(q,\nu)((X_q, X_\nu^{\xi'}), (Y_q, Y_\nu^{\eta'}))
$$
$$
= \mathbf{d}\alpha(q,\nu)((\xi_Q(q) + \mathrm{Hor}_q X_q, X_\nu^{\xi'}), (\eta_Q(q) + \mathrm{Hor}_q Y_q, Y_\nu^{\eta'}))
$$
$$
= \mathbf{d}\alpha(q,\nu)((\xi_Q(q), \tfrac{1}{2}X_\nu^{\xi'}) + (\mathrm{Hor}_q X_q, \tfrac{1}{2}X_\nu^{\xi'}), (\eta_Q(q), \tfrac{1}{2}Y_\nu^{\eta'})
$$
$$
\quad + (\mathrm{Hor}_q Y_q, \tfrac{1}{2}Y_\nu^{\eta'}))
$$
$$
= \mathbf{d}\alpha(q,\nu)((\xi_Q(q), \tfrac{1}{2}X_\nu^{\xi'}), (\eta_Q(q), \tfrac{1}{2}Y_\nu^{\eta'})) + \mathbf{d}\alpha(q,\nu)((\xi_Q(q), \tfrac{1}{2}X_\nu^{\xi'}),
$$
$$
\quad (\mathrm{Hor}_q Y_q, \tfrac{1}{2}Y_\nu^{\eta'})) + \mathbf{d}\alpha(q,\nu)((\mathrm{Hor}_q X_q, \tfrac{1}{2}X_\nu^{\xi'}), (\eta_Q(q), \tfrac{1}{2}Y_\nu^{\eta'}))
$$
$$
\quad + \mathbf{d}\alpha(q,\nu)((\mathrm{Hor}_q X_q, \tfrac{1}{2}X_\nu^{\xi'}), (\mathrm{Hor}_q Y_q, \tfrac{1}{2}Y_\nu^{\eta'}))
$$
$$
= \left\langle \nu, [\tfrac{1}{2}\eta', \xi] \right\rangle + \langle \nu, [\xi, \eta] \rangle + \langle \nu, [\xi, \eta] \rangle + \left\langle \nu, [\eta, \tfrac{1}{2}\xi'] \right\rangle
$$
$$
\quad - \left\langle \mathrm{ad}^*_{\frac{1}{2}\eta'} \nu, \xi \right\rangle + \left\langle - \mathrm{ad}^*_{\frac{1}{2}\xi'} \nu, \eta \right\rangle + \langle \nu, \mathrm{Curv}_{\mathcal{A}}(X_q, Y_q) \rangle
$$
$$
= \langle \nu, [\eta', \xi] \rangle + \langle \nu, [\eta, \xi'] \rangle + \langle \nu, [\xi, \eta] \rangle
$$
$$
\quad + \langle \nu, \mathrm{Curv}_{\mathcal{A}}(q)(X_q, Y_q) \rangle
$$

where we have used the relations determined in the previous discussion.   ∎

Now we are ready to start discussing the reduced form. Recall that

$$
\chi^*(\pi_{\mathcal{O}}^* \Omega_{\mathcal{O}}) = \chi^* \iota_{\mathcal{O}}^* \Omega - \chi^* \mathbf{J}_{\mathcal{O}}^* \omega_{\mathcal{O}}^+
$$
$$
= \pi_1^* \iota_{V^0}^* \Omega - \mathbf{d}\varpi - \pi_2^* \omega_{\mathcal{O}}^+
$$
$$
= \pi_1^* \iota_{V^0}^* \Omega - (\tau_{T^*Q} \times \mathrm{id})^* \mathbf{d}\alpha - \pi_2^* \omega_{\mathcal{O}}^+ \qquad (2.3.36)
$$

We have already established the G-invariance of this form. Notice that the first term is independently G-invariant since, if we denote the action of G on $V^0 \times \mathcal{O}$ by $\psi^{V^0 \times \mathcal{O}}$ and the action of G on T^*Q by ψ, we have

$$
(\psi_g^{V^0 \times \mathcal{O}})^* \pi_1^* \iota_{V^0}^* \Omega = \pi_1^* \psi_g^* \iota_{V^0}^* \Omega = \pi_1^* \iota_{V^0}^* \psi_g^* \Omega
$$
$$
= \pi_1^* \iota_{V^0}^* \Omega
$$

since $\pi_1 \circ \psi_g^{V^0 \times \mathcal{O}}(\alpha_q, \nu) = g \cdot \alpha_q = \psi_g \circ \pi_1(\alpha_q, \nu)$. Thus, the sum of the last two terms is G invariant. Furthermore, the G invariance of the last two terms as forms on $V^0 \times \mathcal{O}$, is really G invariance of a form on $Q \times \mathcal{O}$ since

$$
(\tau_{T^*Q} \times \mathrm{id}) \circ \psi_g^{V^0 \times \mathcal{O}}(\alpha_q, \nu) = (\tau_{T^*Q} \times \mathrm{id})(g \cdot \alpha_q, g \cdot \nu)
$$
$$
= (g \cdot q, g \cdot \nu) = \psi_g^{Q \times \mathcal{O}}(q, \nu)
$$
$$
= \psi_g^{Q \times \mathcal{O}} \circ (\tau_{T^*Q} \times \mathrm{id})(\alpha_q, \nu)
$$

The Part that Drops to $\widetilde{\mathcal{O}}$.   We begin with the proof that the two-form $\mathbf{d}\alpha + \pi_2^*\omega_{\mathcal{O}}^+$ vanishes on vertical vectors.

2.3.10 Proposition.   *The two form, $\mathbf{d}\alpha + \pi_2^*\omega_{\mathcal{O}}^+$ on $Q \times \mathcal{O}$ vanishes on vertical vectors of the bundle $Q \times \mathcal{O} \to \widetilde{\mathcal{O}}$. It therefore uniquely determines a two form on $\widetilde{\mathcal{O}}$.*

Proof.   For the two form $\mathbf{d}\alpha + \pi_2^*\omega_{\mathcal{O}}^+$ on $Q \times \mathcal{O}$ to drop to the quotient, $\widetilde{\mathcal{O}}$, we must have both G invariance of the form and also the property that it vanish on the vertical fibers. To see this, fix $\xi \in \mathfrak{g}$ and let $(Y_q, Y_\nu^{\eta'}) \in T_{(q,\nu)}(Q \times \mathcal{O})$. As usual, let $\eta = \mathcal{A}(q)(Y_q)$. Since the action of G on $Q \times \mathcal{O}$ is the diagonal action, we have

$$\xi_{Q \times \mathcal{O}}(q, \nu) = \left.\frac{d}{dt}\right|_{t=0} (\exp t\xi \cdot q, \mathrm{Ad}^*_{\exp -t\xi}\nu) = (\xi_Q(q), X_\nu^\xi) \qquad (2.3.37)$$

We then have,

$$(\mathbf{d}\alpha + \pi_2^*\omega_{\mathcal{O}}^+)(q, \nu)((\xi_Q(q), X_\nu^\xi), (Y_q, Y_\nu^{\eta'}))$$
$$= \mathbf{d}\alpha(q, \nu)((\xi_Q(q), X_\nu^\xi), (Y_q, Y_\nu^{\eta'}))$$
$$+ \omega_{\mathcal{O}}^+(-\mathrm{ad}_\xi^* \nu, -\mathrm{ad}_{\eta'}^* \nu)$$
$$= \langle \nu, [\eta', \xi] \rangle + \langle \nu, [\eta, \xi] \rangle + \langle \nu, [\xi, \eta] \rangle$$
$$+ \omega_{\mathcal{O}}^+(\mathrm{ad}_\xi^* \nu, \mathrm{ad}_{\eta'}^* \nu)$$
$$= \langle \nu, [\eta', \xi] \rangle + \langle \nu, [\xi, \eta'] \rangle = 0$$

Notice that the curvature term in the formula for $\mathbf{d}\alpha$ vanishes since it is evaluated on a vertical vector $\xi_Q(q)$. ∎

The Part that Drops to $T^*(Q/G)$.   We next characterize the first term of

$$\pi_1^*\iota_{V^0}^*\Omega - (\tau_{T^*Q} \times \mathrm{id})^*\mathbf{d}\alpha - \pi_2^*\omega_{\mathcal{O}}^+ \qquad (2.3.38)$$

as the pull back relative to Γ of the canonical form on $T^*(Q/G)$.

2.3.11 Proposition.   *Denote by*

$$\pi_{\mathcal{A}} : V^0 \to T^*(Q/G) \qquad (2.3.39)$$

the map given by $\alpha_q \mapsto \mathrm{hor}_q^* \alpha_q \in T_{[q]}^*(Q/G)$. Note that this is simply the map Γ restricted to V^0. Let Θ denote the canonical one-form on T^*Q and $\Theta_{Q/G}$ the canonical one-form on $T^*(Q/G)$. We then have

$$\pi_{\mathcal{A}}^*\Theta_{Q/G} = \iota_{V^0}^*\Theta \qquad (2.3.40)$$

from which it follows that

$$\pi_{\mathcal{A}}^*\Omega_{Q/G} = \iota_{V^0}^*\Omega \qquad (2.3.41)$$

Proof.   Let $X_{\alpha_q} \in T_{\alpha_q}V^0$. We have

$$
\pi_{\mathcal{A}}^* \Theta_{Q/G}(X_{\alpha_q}) = \Theta_{Q/G}(\text{hor}_q^* \alpha_q)
$$
$$
= \left\langle \text{hor}_q^* \alpha_q, T\tau_{Q/G} \circ T\pi_{\mathcal{A}} \cdot X_{\alpha_q} \right\rangle
$$

We need to compute the derivative of the composition,

$$
\tau_{Q/G} \circ \pi_{\mathcal{A}} : V^0 \to Q/G \tag{2.3.42}
$$

Let $t \mapsto \alpha_q(t) \in V^0$ be a smooth curve through α_q such that $\dot{\alpha}_q(0) = X_{\alpha_q}$. Let $q(t) = \tau_{T^*Q}(\alpha_q(t))$. Then

$$
\left.\frac{d}{dt}\right|_{t=0} \tau_{Q/G} \circ \pi_{\mathcal{A}}(\alpha_q(t)) = \left.\frac{d}{dt}\right|_{t=0} \tau_{Q/G}(\text{hor}_{q(t)}^*(\alpha_q(t)))
$$
$$
= \left.\frac{d}{dt}\right|_{t=0} [q(t)] = T\pi \circ T\tau_{T^*Q} \cdot X_{\alpha_q}
$$

Thus,

$$
\left\langle \text{hor}_q^* \alpha_q, T\tau_{Q/G} \circ T\pi_{\mathcal{A}} \cdot X_{\alpha_q} \right\rangle = \left\langle \alpha_q, \text{hor}_q \circ T\pi \circ T\tau_{T^*Q} \cdot X_{\alpha_q} \right\rangle \tag{2.3.43}
$$

On the other hand, we have

$$
\iota_{V^0}^* \Theta(\alpha_q)(X_{\alpha_q}) = \left\langle \alpha_q, T\tau_{T^*Q} \cdot X_{\alpha_q} \right\rangle
$$
$$
= \left\langle \alpha_q, \text{Hor}_q T\tau_{T^*Q} \cdot X_{\alpha_q} + \text{Ver}_q T\tau_{T^*Q} \cdot X_{\alpha_q} \right\rangle
$$
$$
= \left\langle \alpha_q, \text{Hor}_q T\tau_{T^*Q} \cdot X_{\alpha_q} \right\rangle
$$
$$
= \left\langle \alpha_q, \text{hor}_q \circ T\pi \circ T\tau_{T^*Q} \cdot X_{\alpha_q} \right\rangle
$$

where the third equality follows from the fact that α_q annihilates vertical vectors. ∎

A Final Piece of Diagram Chasing.   Recall that we have the following maps:

$$
V^0 \times \mathcal{O} \xrightarrow{\tau_{T^*Q} \times \text{id}} Q \times \mathcal{O} \xrightarrow{\pi_G} \widetilde{\mathcal{O}}.
$$

Define the map ϕ as follows:

$$
\phi(\alpha_q, \nu) = (\text{hor}_q^* \circ \pi_1, \pi_G \circ (\tau_{T^*Q} \times \text{id})) \tag{2.3.44}
$$

It is easy to see that ϕ is G-invariant, so that we have the following commutative diagram.

$$
\begin{array}{ccc}
V^0 \times \mathcal{O} & \xrightarrow{\quad \phi \quad} & T^*(Q/G) \times_{Q/G} \widetilde{\mathcal{O}} \\[2mm]
& \widetilde{\pi}_G \searrow \quad \quad \nearrow \widetilde{\phi} & \\[2mm]
& (V^0 \times \mathcal{O})/G &
\end{array}
$$

It is straightforward to check that the map $\widetilde{\phi}$ is invertible and therefore determines a bundle isomorphism.

Final Formula for the Symplectic Form.   Now we are ready to gather these pieces to give the main result.

2.3.12 Theorem.   *For $P = T^*Q$ a cotangent bundle, with the Lie group G acting freely and properly by cotangent lift, the symplectic reduced space P_μ is symplectically diffeomorphic to $T^*(Q/G) \times_{Q/G} \widetilde{\mathcal{O}}$ with the symplectic form ω_{red} that is given by the formula*

$$
\omega_{\mathrm{red}} = \Omega_{Q/G} - \beta
$$

where $\Omega_{Q/G}$ is the canonical form on $T^*(Q/G)$ and where β is the unique two form on $\widetilde{\mathcal{O}}$ determined by

$$
\pi_G^* \beta = \mathbf{d}\alpha + \pi_2^* \omega_{\mathcal{O}}^+
$$

and, as in Theorem 2.3.9,

$$
\begin{aligned}
&\mathbf{d}\alpha(q,\nu)((X_q, X_\nu^{\xi'}), (Y_q, Y_\nu^{\eta'})) \\
&\quad = \langle \nu, [\eta', \xi] \rangle + \langle \nu, [\eta, \xi'] \rangle + \langle \nu, [\xi, \eta] \rangle + \langle \nu, \mathrm{Curv}_\mathcal{A}(q)(X_q, Y_q) \rangle \quad (2.3.45)
\end{aligned}
$$

Proof.   The two-form ω_{red} is the unique two-form on $T^*(Q/G) \times_{Q/G} \widetilde{\mathcal{O}}$ such that $\widetilde{\phi}^* \omega_{\mathrm{red}} = \bar{\chi}^* \Omega_{\mathcal{O}}$, where $\bar{\chi}^* \Omega_{\mathcal{O}}$ (see equation 2.3.14) is the unique two-form on $(V^0 \times \mathcal{O})/G$ such that

$$
\widetilde{\pi}_G^* \bar{\chi}^* \Omega_{\mathcal{O}} = \pi_1^* \iota_{V^0}^* \Omega - (\tau_{T^*Q} \times \mathrm{id})^* \mathbf{d}\alpha - \pi_2^* \omega_{\mathcal{O}}^+.
$$

We then have

$$
\widetilde{\pi}_G^* \widetilde{\phi}^* \omega_{\mathrm{red}} = \pi_1^* \iota_{V^0}^* \Omega - (\tau_{T^*Q} \times \mathrm{id})^* \mathbf{d}\alpha - \pi_2^* \omega_{\mathcal{O}}^+
$$

However, since $\widetilde{\phi} \circ \widetilde{\pi}_G = \phi$, we have

$$
\phi^* \omega_{\mathrm{red}} = \pi_1^* \iota_{V^0}^* \Omega - (\tau_{T^*Q} \times \mathrm{id})^* \mathbf{d}\alpha - \pi_2^* \omega_{\mathcal{O}}^+ \qquad (2.3.46)
$$

from which we can read off ω_{red}:

$$
\begin{aligned}
\phi^* \omega_{\mathrm{red}}(\alpha_q, \nu)((X_{\alpha_q}, X_\nu), (Y_{\alpha_q}, Y_\nu)) = \\
\omega_{\mathrm{red}}(\mathrm{hor}^*_q \alpha_q, [q, \nu])\big((T(\mathrm{hor}^* \circ \pi_1)(X_{\alpha_q}, X_\nu), \\
T(\pi_G \circ \tau_{T^*Q} \times \mathrm{id})(X_{\alpha_q}, X_\nu)), (T(\mathrm{hor}^* \circ \pi_1)(Y_{\alpha_q}, Y_\nu), \\
T(\pi_G \circ \tau_{T^*Q} \times \mathrm{id})(Y_{\alpha_q}, Y_\nu)))
\end{aligned}
$$

Note that $T(\mathrm{hor}^* \circ \pi_1)(X_{\alpha_q}, X_\nu) = T\pi_{\mathcal{A}} X_{\alpha_q}$ and

$$
T(\pi_G \circ (\tau_{T^*Q} \times \mathrm{id}))(X_{\alpha_q}, X_\nu) = T_{(q, \nu)}\pi_G \cdot (T\tau_{T^*Q} X_{\alpha_q}, X_\nu)
$$

The right hand side of equation (2.3.46) becomes

$$
\begin{aligned}
\Omega(\alpha_q)(X_{\alpha_q}, Y_{\alpha_q}) - (\tau_{T^*Q} \times \mathrm{id})^* \pi_G^* \beta(X_{\alpha_q}, Y_{\alpha_q}) \\
= \Omega_{Q/G}(\pi_{\mathcal{A}}(\alpha_q))(T\pi_{\mathcal{A}} X_{\alpha_q}, T\pi_{\mathcal{A}} Y_{\alpha_q}) \\
- \beta(([q, \nu])(T_{(q, \nu)}\pi_G \cdot (T\tau_{T^*Q} X_{\alpha_q}, X_\nu), \\
T_{(q, \nu)}\pi_G \cdot (T\tau_{T^*Q} Y_{\alpha_q}, Y_\nu)),
\end{aligned}
$$

from which the the claim follows. ∎

The Extreme Cases. The obvious extreme cases are $Q = G$ and G Abelian. We first consider the case $Q = G$. Then, Q/G reduces to a point, and the associated bundle is simply the coadjoint orbit through ν_0. $\widetilde{\mathcal{O}} = Q \times \mathcal{O}/G = G \times \mathcal{O}/G \simeq \mathcal{O}$. Consider a tangent vector to the coadjoint orbit through a point ν, given by $-\mathrm{ad}^*_{\xi'} \nu$. Represent this tangent vector with the curve through ν, $t \mapsto \mathrm{Ad}^*_{\exp -t\xi'} \nu$. We must find a lift to $G \times \mathcal{O}$ of such a tangent vector. The projection, $\pi_G : G \times \mathcal{O} \to \mathcal{O}$ is given by $(g, \nu) \mapsto g^{-1}\nu$ since $[g, \nu] = [e, g^{-1}\nu]$. More generally, consider a curve through $(e, \nu) \in G \times \mathcal{O}$ denoted by $t \mapsto (g(t), \nu(t))$. Let $\xi = \dot{g}(0)$. Since $\mathcal{A}(e)(\dot{g}(0)) = \dot{g}(0)$, this is consistent notation. The projection of this curve to $\mathcal{O}$ is given by

$$
\pi_G(g(t), \nu(t)) = g(t)^{-1}\nu(t) = \mathrm{Ad}^*_{g(t)} \nu(t) \tag{2.3.47}
$$

and therefore we require

$$
\frac{d}{dt}\bigg|_{t=0} \mathrm{Ad}^*_{g(t)} \nu(t) = -\mathrm{ad}^*_{\xi'} \nu,
$$

which implies

$$
\mathrm{Ad}^*_g \dot{\nu}(o) + \mathrm{ad}^*_{\dot{g}(0)} \nu = -\mathrm{ad}^*_{\xi'} \nu, \tag{2.3.48}
$$

from which it follows that

$$
\dot{\nu}(0) = -\mathrm{ad}^*_{\xi'} \nu - \mathrm{ad}^*_{\xi} \nu
$$

Equations ($2.3.36$) and ($2.3.35$) give

$$\begin{aligned}
\omega_{\mathrm{red}}(\nu)(-\operatorname{ad}^*_{\xi'}\nu, \operatorname{ad}^*_{\eta'}\nu) &= \Omega_{Q/G} - \beta(e,\nu)((\xi, X_\nu^{\xi'+\xi}), (\eta, Y_\nu^{\eta'+\eta})) \\
&= -\big(\mathbf{d}\alpha + \pi_2^*\omega_{\widetilde{\mathcal{O}}}^+\big)(e,\nu)((\xi, X_\nu^{\xi'+\xi}), (\eta, Y_\nu^{\eta'+\eta})) \\
&= -\big(\,\langle\nu,[\eta'+\eta]\rangle + \langle\nu,[\eta,\xi'+\xi]\rangle + \langle\nu,[\xi,\eta]\rangle \\
&\qquad + \operatorname{Curv}_{\mathcal{A}}(e)(\xi,\eta) + \langle\nu,[\xi'+\xi,\eta'+\eta]\rangle\,\big) \\
&= -\langle\nu,[\xi',\eta']\rangle\,,
\end{aligned}$$

where the last equality follows from the fact that the curvature term vanishes on vertical vectors and an expansion of the Lie algebra brackets. Thus, we recover the coadjoint orbit symplectic structure, as we should.

For G Abelian, the fibers of the $\widetilde{\mathcal{O}}$ bundle collapse and we are left with just $T^*(Q/G)$. The reduced symplectic form, from equations $2.3.36$ and $2.3.35$ is then

$$\omega_{\mathrm{red}} = \Omega_{Q/G} - \langle\nu, \operatorname{Curv}_{\mathcal{A}}\rangle \tag{2.3.49}$$

since all brackets vanish.

2.4   Singular Cotangent Bundle Reduction

This section gives an outline of parts of cotangent bundle reduction in the singular case. This material is not strictly needed for later sections, and is included for completeness and the convenience of the reader. No proofs will be given here and the inclined reader can safely skip this section if their main goal is the regular case. On the other hand, this material would be needed background for singular cotangent bundle reduction by stages—a theory that is not treated in this book—or anywhere else as far as we know. Only the general theory of singular reduction by stages in the general context of *symplectic reduction* is presented in Part III.

History, Setting, Literature.   In §1.3 we gave a brief history of singular reduction in the symplectic case. As we have already seen in the regular case, it would be naive to think that one can dismiss the case of cotangent bundles as "merely a special case". Likewise, the case of singular cotangent bundle reduction should not be viewed as simply following easily from the case of singular symplectic reduction. In fact, the additional structure that one has is quite remarkable.

Historically, Montgomery [1983] made some interesting progress in the case of non-free actions and Emmrich and Römer [1990] and Schmah [2002] took some first steps in the case of singular cotangent bundle reduction in the case of momentum value zero and momentum values with trivial coadjoint orbits. As was mentioned in §1.3, Hernandez and Marsden [2005] and Birtea, Puta, Ratiu, and Tudoran [2005] began an approach to singular

cotangent bundle reduction using the blowing up technique, which *regularizes* singular points, a method that remains to be solidly linked with that outlined in this section. One of the interesting things about that approach is that it also deals with singular relative equilibria; it shows how new nonsingular relative equilibria *bifurcate* from singular ones, as is known to happen in the double spherical pendulum, for example (see Marsden and Scheurle [1993a]). The method also constructs a regularized (or blown up) amended potential, whose second variation gives an approach to stability of singular relative equilibria. Another very interesting approach to stability in the singular case is a singular version of the block diagonalization method of Simo, Lewis and Marsden [1991]; this is given in Rodríguez-Olmos [2006b].

As we have already seen in the regular case in the preceding sections of this chapter, the symplectic reduced spaces $(T^*Q)_\mu$ of a cotangent bundle T^*Q carry additional geometric structure. In the singular case, the general theory of singular reduction reviewed in §1.4 applies and we conclude from Theorem 1.4.2 that the reduced topological spaces $(T^*Q)_\mu$ determined by the cotangent lifted action of a Lie group G, are cone spaces whose strata $(T^*Q)_\mu^{(H)}$ are symplectic manifolds indexed by the orbit types (H) of this G-action on T^*Q. The symplectic form and the Poisson bracket on the strata $(T^*Q)_\mu^{(H)}$ are naturally induced from that of T^*Q as given in Theorem 1.4.2. All these strata have a rich additional structure that is presented below. At this point, the most general case has not been studied yet, so we shall review below what is known as of this writing. All these results can be found in Perlmutter, Rodríguez-Olmos, and Sousa Dias [2006], Perlmutter, Rodríguez-Olmos, and Sousa Dias [2007], Perlmutter and Rodríguez-Olmos [2006], and Rodríguez-Olmos [2006a].

Examples.   At the end of this section, we provide a simple, but key example, that illustrates singular cotangent bundle reduction, and hence singular symplectic reduction as well, namely that of the spherical pendulum (Lerman, Montgomery, and Sjamaar [1993], Cushman and Bates [1997], and Cushman and Śniatycki [2001]). Another, more sophisticated mathematical example that illustrates more of the features that one sees in the theory, can be found in Perlmutter, Rodríguez-Olmos, and Sousa Dias [2006].

Reduction at Zero Momentum Value.   In all that follows we shall use the notations and conventions established in §1.4. It is shown in Perlmutter, Rodríguez-Olmos, and Sousa Dias [2006] that the symplectic stratification in Sjamaar and Lerman [1991], reviewed in §1.4, refines in two ways:

(i) Each symplectic stratum of the reduced space at zero momentum admits an additional stratification called the secondary stratification with two distinct types of pieces, one of which is open and dense and symplectomorphic to a cotangent bundle.

(ii) The reduced space at zero momentum admits a finer stratification than the symplectic one; its pieces are coisotropic in their respective symplectic strata that include them. This finer stratification is called the coisotropic stratification.

Let us elaborate in more detail on these statements. We consider the cotangent lift of a proper G-action on the manifold Q. For simplicity of exposition, let us assume that the orbit types of this G-action on Q are all connected.[2] If $(T^*Q)_0 := \mathbf{J}^{-1}(0)/G$, we know from §1.4 that $(T^*Q)_0$ is a stratified symplectic space. However much more is true owing to the cotangent bundle structure. First, we consider the orbit types for the G-action on $\mathbf{J}^{-1}(0)$ which determine the symplectic stratification of $(T^*Q)_0$. Let $I(G, \mathbf{J}^{-1}(0))$ be the isotropy lattice for the G-action on $\mathbf{J}^{-1}(0)$ and $I(G, Q)$ the isotropy lattice for the G-action on the base manifold Q. For $(L) \in I(G, \mathbf{J}^{-1}(0))$, define $(T^*Q)_0^{(L)} := \mathbf{J}^{-1}(0)_{(L)}/G$.

Since $\alpha_q \in \mathbf{J}^{-1}(0) \cap T_q^*Q$ if and only if α_q annihilates the tangent space to the group orbit through q, namely $\mathfrak{g} \cdot q := \{\xi_Q(q) \mid \xi \in \mathfrak{g}\}$, it follows that the restriction of $\mathbf{J}^{-1}(0) \subset T^*Q$ to an orbit type manifold $Q_{(H)}$ of the base Q is a subbundle of T^*Q restricted to $Q_{(H)}$. This subbundle carries a G-action and its careful study reveals that it admits orbit type manifolds $\left(\mathbf{J}^{-1}(0)|_{Q_{(H)}}\right)_{(K)} \neq \varnothing$ if and only if $(K) \in I(G, Q)$ and $(H) \preceq (K)$. Notice that $\left(\mathbf{J}^{-1}(0)|_{Q_{(H)}}\right)_{(K)}$ is a fiber bundle over $Q_{(H)}$. Its quotient

$$S_{(H) \preceq (K)} := \left(\mathbf{J}^{-1}(0)|_{Q_{(H)}}\right)_{(K)}/G$$

by the G-action is called a **seam** from H to K; it is a fiber bundle over $Q_{(H)}/G$. Decomposing $\mathbf{J}^{-1}(0)$ over the orbit types of Q yields the following result.

2.4.1 Theorem.   *The isotropy lattice of $\mathbf{J}^{-1}(0)$ is identical to the isotropy lattice for Q, that is, $(L) \in I(G, \mathbf{J}^{-1}(0))$ if and only if $(L) \in I(G, Q)$. Therefore, $(T^*Q)_0^{(L)} \neq \varnothing$ if and only if $(L) \in I(G, Q)$. Furthermore, this symplectic stratum $(T^*Q)_0^{(L)}$ admits the following partition:*

$$(T^*Q)_0^{(L)} = \bigsqcup_{(H) \preceq (L)} S_{(H) \preceq (L)} = \left(\mathbf{J}_{(L)}^{-1}(0)/G\right) \bigsqcup \bigsqcup_{(H) \prec (L)} S_{(H) \preceq (L)}. \quad (2.4.1)$$

where $S_{(H) \preceq (L)}$ is the seam from H to L and $\mathbf{J}_{(L)} : T^*(Q_{(L)}) \to \mathfrak{g}^*$ is the momentum map for the lifted action of G on the orbit type manifold $Q_{(L)}$.

The restriction of the cotangent bundle projection to $\left(\mathbf{J}^{-1}(0)\right)_{(L)}$ is a continuous G-equivariant surjection $\tau_L : \left(\mathbf{J}^{-1}(0)\right)_{(L)} \to \overline{Q_{(L)}}$ and the quotient $(T^*Q)_0^{(L)}$ fibers over the *closure* of $Q_{(L)}/G$. This partition of $(T^*Q)_0^{(L)}$

---

[2]Otherwise one would need to state the following results in the category of Σ-manifolds which we shall not define and not use in this quick review.

admits frontier conditions determined by the partial order in the isotropy lattice $I(G, Q)$. Consider the first piece of this partition. It is not hard to show that if Q has only one orbit type, that is, $Q = Q_{(L)}$ for some $(L) \in I(G, Q)$, then the reduced space $\mathbf{J}_{(L)}^{-1}(0)/G$ at the zero momentum value consists of just one stratum symplectomorphic to $T^*(Q_{(L)}/G)$ endowed with its canonical symplectic form ω_L; see Emmrich and Römer [1990] and Schmah [2002]. We shall use below the notation $Q^{(L)} := Q_{(L)}/G$ and similarly for other upper indexed sets. An upper indexed map is obtained from quotienting.

2.4.2 Theorem.   *The partition* (2.4.1) *is a stratification of* $(T^*Q)_0^{(L)}$ *which we call the* **secondary stratification** *of* $(T^*Q)_0^{(L)}$. *The stratum* $C_L := \mathbf{J}_{(L)}^{-1}(0)/G$ *is open and dense in* $(T^*Q)_0^{(L)}$. *The frontier conditions are:*

(i) $S_{(H) \preceq (L)} \subset \partial C_L$ *for all* $(H) \prec (L)$.

(ii) $S_{(H') \preceq (L)} \subset \partial S_{(H) \preceq (L)}$ *if and only if* $(H') \prec (H) \prec (L)$.

The map $\tau^L : (T^*Q)_0^{(L)} \to \overline{Q^{(L)}}$ *is a stratified surjective submersion.*

The next theorem describes the symplectic properties of the secondary stratification of $(T^*Q)_0^{(L)}$. On an open dense set of the symplectic stratum $(T^*Q)_0^{(L)}$ the symplectic structure is the canonical symplectic structure of the reduced cotangent bundle $T^*(Q_{(L)}/G)$. Furthermore, the seams $S_{(H) \preceq (L)}$ are coisotropic submanifolds of $(T^*Q)_0^{(L)}$.

2.4.3 Theorem.   *The restriction to the open dense stratum* C_L *of the reduced symplectic form* $\omega_0^{(L)}$ *on* $(T^*Q)_0^{(L)}$ *coincides with the symplectic form on* $C_L := \mathbf{J}_{(L)}^{-1}(0)/G$ *obtained by reduction of* $T^*(Q_{(L)})$. *Moreover, this reduced symplectic form on* C_L *has a unique smooth extension to* $(T^*Q)_0^{(L)}$ *which is identical to* $\omega_0^{(L)}$.

Let ω_L *denote the canonical symplectic form on* $T^*Q^{(L)}$. *The stratification* (2.4.1) *has the following properties:*

(i) *The open dense stratum* C_L *is a maximal symplectic submanifold of the symplectic stratum* $\left((T^*Q)_0^{(L)}, \omega_0^{(L)}\right)$ *and is symplectomorphic to the canonical cotangent bundle* $\left(T^*Q^{(L)}, \omega_L\right)$.

(ii) *The seams* $S_{(H) \preceq (L)}$ *are coisotropic submanifolds of* $\left((T^*Q)_0^{(L)}, \omega_0^{(L)}\right)$.

Finally we consider the *finer* partition of $(T^*Q)_0$ whose elements are all the seams and all cotangent bundles found in the secondary stratifications of all the symplectic strata, that is,

$$(T^*Q)_0 = \bigsqcup_{(L)} C_L \bigsqcup_{(K') \prec (K)} S_{(K') \preceq (K)} \tag{2.4.2}$$

for all $(L), (K), (K') \in I(G, Q)$. We can also write this partition as

$$(T^*Q)_0 = \bigsqcup_{(H) \preceq (L)} S_{(H) \preceq (L)}, \qquad (2.4.3)$$

where the disjoint union is over all pairs in $I(G, Q)$ that are comparable in the partial order and we use the identity $S_{(L) \preceq (L)} = C_L$ for each $(L) \in I(G, Q)$. The next theorem describes the topological properties of this partition as well as the bundle structure of $(T^*Q)_0$.

2.4.4 Theorem.   *The partition* (2.4.2) *of* $(T^*Q)_0$ *is a stratification, called the **coisotropic stratification**, and has the following properties:*

(i) *If* $(H_0) \succeq (L)$ *for all* $(L) \in I(G, Q)$ *then* C_{H_0} *is open and dense in* $(T^*Q)_0$.

(ii) *The frontier conditions are:* $S_{(H) \preceq (L)} \subset \partial S_{(H') \preceq (L')}$ *if and only if* $(H') \succeq (H)$ *and* $(L') \succeq (L)$ *except for the simultaneous equalities* $(H') = (H)$ *and* $(L') = (L)$.

(iii) *The continuous projection* $\tau^0 : (T^*Q)_0 \to Q/G$ *is a stratified surjective submersion with respect to the coisotropic stratification* (2.4.2) *of* $(T^*Q)_0$ *and the usual orbit type stratification of* Q/G.

(iv) *If* $I(G, Q)$ *has more than one class then the coisotropic stratification is strictly finer than the symplectic stratification. If* $I(G, Q)$ *has only one class then the symplectic and coisotropic stratifications of* $(T^*Q)_0$ *are identical.*

Notice that by **(i)**, $(T^*Q)_0$ is almost a cotangent bundle, namely the cotangent bundle of $Q^{(H_0)}$.

Towards Reduction at Nonzero Momentum Values.   The previous considerations completely describe the additional stratified and symplectic structures that appear in the singular reduced space at the zero value of the momentum map. The situation for the reduction at nonzero values is considerably more involved as can be seen even in the simple case when the base manifold consists of just one single orbit type. As will be explained below, unlike the case of the reduced space at zero momentum value of a cotangent bundle whose base has only one orbit type, the reduced spaces at nonzero momentum values having orbits not equal to one point are nontrivial stratified spaces whose strata are determined purely by an action induced by the coadjoint action.

Recall that in the free case, reduction at nonzero momentum values involves a connection in the description of the reduced symplectic form. Even for the single orbit type case the generalization of a connection leads to an interesting theory in its own right (see Perlmutter and Rodríguez-Olmos

[2006]). Using this concept of singular connection that shall not be reviewed here, it is possible to carry out symplectic and Poisson reduction for the single orbit type base manifolds and the conclusion is that the reduced spaces $\mathbf{J}^{-1}(\mathcal{O}_\mu)/G$ are stratified. To describe the strata, assume that $Q = Q_{(H)}$ and consider the action $H \times \mathfrak{h}^\circ \to \mathfrak{h}^\circ$ induced by the coadjoint action, where $\mathfrak{h}^\circ \subset \mathfrak{g}^*$ is the annihilator of $\mathfrak{h}$. This action has its own isotropy lattice $I(H, \mathfrak{h}^\circ)$. The following statement is proved in Perlmutter and Rodríguez-Olmos [2006] (see also Hochgerner [2007]).

2.4.5 Theorem.   *For the case in which Q satisfies $Q = Q_{(H)}$, the symplectic reduced spaces $\mathbf{J}^{-1}(\mathcal{O}_\mu)/G$ are stratified by symplectic manifolds with strata given by $\{S_{(K)} \mid (K) \in I(H, \mathfrak{h}^\circ)\}$. Each $S_{(K)}$ is a fiber bundle over $T^*(Q/G)$ with typical fiber $\left(\mathcal{O}_\mu \cap \mathfrak{h}^\circ_{(K)} \right) / H$. The symplectic structure on $S_{(K)}$ depends on the choice of singular connection and, like in free cotangent bundle reduction, splits into three terms: the pull back of the canonical symplectic form on $T^*(Q/G)$, the pull back to each fiber of a natural symplectic form on $\left(\mathcal{O}_\mu \cap \mathfrak{h}^\circ_{(K)} \right) / H$, and a term involving the curvature of the singular connection. These strata satisfy frontier conditions determined by the lattice $I(H, \mathfrak{h}^\circ)$.*

Remark 1. The last two terms constitute an extension of the Sternberg minimal coupling procedure to the quotient $Q \to Q/G$ and a symplectic fibration over Q with symplectic fibers that look like $\left(\mathcal{O}_\mu \cap \mathfrak{h}^\circ_{(K)} \right) / H$.

Remark 2. If the orbit type $(H) = e$, i.e. the action of G on the base manifold Q is free, then we recover the orbit form of the cotangent bundle reduction theorem (see the preceding section and Marsden and Perlmutter [2000], Perlmutter and Ratiu [2005]).

Nonzero Momentum.   Now consider the general case when the base manifold admits multiple orbit types. A preliminary problem is to determine the structure of the full isotropy lattice $I(G, T^*Q)$ since its elements index the smooth strata of the Poisson reduced space $(T^*Q)/G$. There are two types of isotropy lattices that we have already encountered in the previous special cases: the base lattice $I(G, Q)$ (which appeared for the points in T^*Q with zero momentum) and, for each $(H) \in I(G, Q)$, the restricted coadjoint action lattice $I(H, \mathfrak{h}^\circ)$ (which determined the strata for the single orbit type case with nonzero momentum). It turns out that the entire orbit type lattice $I(G, T^*Q)$ can be constructed from these two lattices as shown in Rodríguez-Olmos [2006a].

2.4.6 Theorem.   *Let G act on Q and by cotangent lift on T^*Q. Let L be a closed subgroup of G. Then $(L) \in I(G, T^*Q)$ if and only if there exist $(H_1), (H_2) \in I(G, Q)$ and $(K) \in I(H_2, \mathfrak{h}_2^\circ)$ such that $(H_1) \succeq (H_2)$ and $L = H_1 \cap K$.*

Remark. One can obtain all the classes in $I(G, T^*Q)$ by the following algorithm. First, choose representatives of every class in $I(G, Q)$ such that if $(H') \succeq (H)$ the corresponding representatives satisfy $H' \subset H$. It is always possible to choose a complete set of representatives for all the classes of $I(G, Q)$ in this way and we will call them *normal representatives*. Let H_0 be the normal representative of the maximal class in $I(G, Q)$. We will say that H_0 has depth zero. Any other normal representative H will have depth $n+1$ if there is a normal representative H' with depth n such that $H' \subset H$ and there is no other normal representative H'' with $H' \subset H'' \subset H$. To compute all the elements of $I(G, T^*Q)$ we start by computing those of $I(H_0, \mathfrak{h}_0^\circ)$. Then, for any n and every normal representative H of depth n, intersect the classes of $I(H, \mathfrak{h}^0)$ with the H-class of every normal representative of depth $0, \ldots, n-1$ included in H. All the classes obtained in this way for any n can be made G-classes by conjugating in G. After removing the repeated ones, if any, we obtain all the elements of $I(G, T^*Q)$.

The Canoe.   In this paragraph[3] we illustrate some of these ideas through the study of the phase space reduction of the simple spherical pendulum, namely T^*S^2 with the natural cotangent lifted S^1 action given by rotations around the vertical axis. This example was studied in Lerman, Montgomery, and Sjamaar [1993], Cushman and Bates [1997], and Cushman and Śniatycki [2001] but in this paragraph we shall investigate it from the point of view of singular cotangent bundle reduction.

Consider the unit sphere S^2 in $\mathbb{R}^3$ and let S^1 act on S^2 by rotation around the z-axis and hence by cotangent lifts on T^*S^2. Following Cushman and Śniatycki [2001], page 749, the algebra of invariant polynomials for this lifted action is generated by

$$
\begin{array}{llll}
\sigma_1 &=& x_3 & \qquad \sigma_4 &=& y_1^2 + y_2^2 + y_3^2 \\
\sigma_2 &=& x_1^2 + x_2^2 & \qquad \sigma_5 &=& x_1 y_2 - x_2 y_1 \\
\sigma_3 &=& y_3 & \qquad \sigma_6 &=& x_1 y_1 + x_2 y_2
\end{array}
$$

where a point in T^*S^2 is identified with a pair $(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^3 \times \mathbb{R}^3$ satisfying $\|\mathbf{x}\| = 1$ and $\mathbf{x} \cdot \mathbf{y} = 0$. The relations among the generators are

$$
\begin{array}{llll}
1 &=& \sigma_1^2 + \sigma_2 & \qquad \sigma_5^2 + \sigma_6^2 &=& \sigma_2(\sigma_4 - \sigma_3^2) & \qquad \sigma_4 - \sigma_3^2 &\geq& 0 \\
0 &=& \sigma_6 + \sigma_1 \sigma_3 & \qquad \sigma_2 &\geq& 0.
\end{array}
$$

Note that $\mathbf{J}(\mathbf{x}, \mathbf{y}) = \sigma_5$.

The reduced space at zero, namely $(T^*S^2)_0$, is homeomorphic to the semi-algebraic set $\mathcal{C}$ obtained by projecting onto the hyperplane $\sigma_5 = 0$, the subset of $\mathbb{R}^6$ (thought of as having coordinates σ_i, $i = 1, \ldots, 6$) given by these five relations among generators; for details see Cushman and Bates

---

[3]The exposition in this paragraph was kindly provided by Miguel Rodríguez-Olmos. We thank him for this helpful contribution.

[1997], Cushman and Śniatycki [2001], or Lerman, Montgomery, and Sjamaar [1993]. Eliminating σ_2 and σ_6, we get the "canoe"

$$\mathcal{C} = \{\boldsymbol{\sigma} := (\sigma_1, \sigma_3, \sigma_4) \in \mathbb{R}^3 \mid \sigma_3^2 = (1 - \sigma_1^2)\sigma_4, \ |\sigma_1| \leq 1, \ \sigma_4 \geq 0\},$$

which is illustrated in Figure 2.4.1.
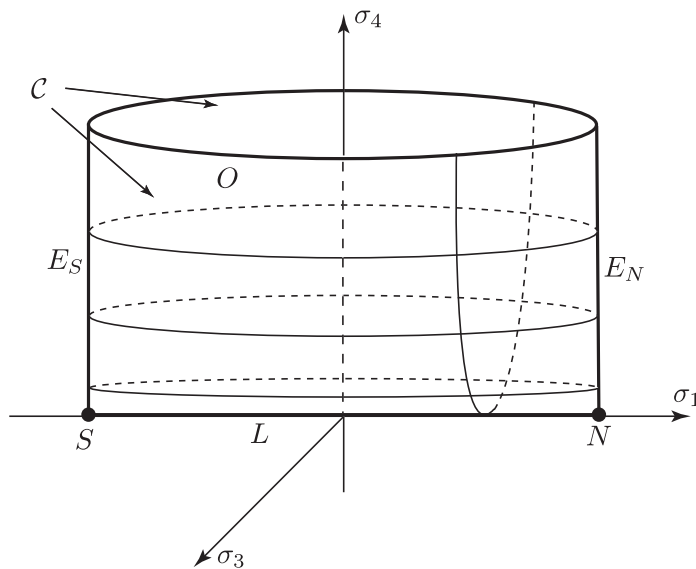
FIGURE 2.4.1. The "canoe" $\mathcal{C} = (T^*S^2)_0$ obtained by reducing T^*S^2 by the cotangent lifted S^1-action on S^2 around the z-axis.

In $\mathcal{C}$ we distinguish the subsets

$$E_S = \{(-1, 0, t) \mid t > 0\}$$
$$E_N = \{(1, 0, t) \mid t > 0\}$$
$$L = \{(t, 0, 0) \mid -1 < t < 1\}$$
$$S = \{(-1, 0, 0)\}, N = \{(1, 0, 0)\}$$
$$O = \{\boldsymbol{\sigma} \in \mathcal{C} \mid \sigma_1 \neq -1, 1\}.$$

Note that $L \subset O$.

The isotropy lattice of the S^1-action on S^2 has two elements: $I(S^1, S^2) = \{1, S^1\}$. The corresponding isotropy type manifolds are $S^2_{(S^1)}$, which is S^2 except the north and south poles, and $S^2_{(1)}$ which is the union of the north and south poles. It is easy to check that also $I(S^1, \mathbf{J}^{-1}(0)) = \{1, S^1\}$ in

agreement with Theorem 2.4.1. Since

$$\mathbf{J}^{-1}(0) = \{(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^3 \times \mathbb{R}^3 \mid x_1 y_2 - y_1 x_2 = 0, \; x_1^2 + x_2^2 + x_3^2 = 1,$$
$$x_1 y_1 + x_2 y_2 + x_3 y_3 = 0\},$$

the two orbit type manifolds are

$$\mathbf{J}^{-1}(0)_{(S^1)} = \{(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^3 \times \mathbb{R}^3 \mid x_1 = x_2 = y_1 = y_2 = 0\} \cap \mathbf{J}^{-1}(0)$$
$$\mathbf{J}^{-1}(0)_{(1)} = A \cup B,$$

where

$$A = \{(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^3 \times \mathbb{R}^3 \mid x_1 = x_2 = 0, \; y_1^2 + y_2^2 \neq 0\} \cap \mathbf{J}^{-1}(0)$$
$$B = \{(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^3 \times \mathbb{R}^3 \mid x_1^2 + x_2^2 \neq 0\} \cap \mathbf{J}^{-1}(0).$$

Note that we have $\mathbf{J}^{-1}(0) = A \cup B \cup \mathbf{J}^{-1}(0)_{(S^1)}$, and

$$
\begin{aligned}
A &= \mathbf{J}^{-1}(0)_{(1)} \cap T^* S^2|_{S^2_{(S^1)}} \\
B &= \mathbf{J}^{-1}(0)_{(1)} \cap T^* S^2|_{S^2_{(1)}} \\
\mathbf{J}^{-1}(0)_{(S^1)} &= \mathbf{J}^{-1}(0)_{(S^1)} \cap T^* S^2|_{S^2_{(S^1)}}.
\end{aligned}
$$

We shall describe below the different stratifications of $(T^* S^2)_0 \cong \mathcal{C}$.

The Symplectic Stratification. This is the stratification given by singular reduction and does not take advantage of the cotangent bundle structure. In view of the description of the two orbit type manifolds of $\mathbf{J}^{-1}(0)$, it is easy to see that

$$
\begin{aligned}
\mathbf{J}^{-1}(0)_{(S^1)}/S^1 &=: (T^* S^2)_0^{(S^1)} &= \{S, N\} \\
\mathbf{J}^{-1}(0)_{(1)}/S^1 &=: (T^* S^2)_0^{(1)} &= \mathcal{C}\backslash\{S, N\} = O \cup E_S \cup E_N.
\end{aligned}
$$

Regarding S^2 as the zero section of $T^* S^2$, it is clear that S^2/S^1 corresponds to the subset $S \cup L \cup N$. It is straightforward to obtain the orbit type stratification of S^2/S^1 in the canoe as:

$$
\begin{aligned}
S^2_{(S^1)}/S^1 &= \{S, N\} \\
S^2_{(1)}/S^1 &= L.
\end{aligned}
$$

The Secondary Stratifications. According to Theorems 2.4.1 and 2.4.2, since (S^1) is the minimal element of $I(S^1, S^2)$, it follows that $(T^* S^2)_0^{(S^1)}$ consists of a single secondary stratum, namely itself, which is symplectomorphic to $T^* \left(S^2_{(S^1)}/S^1 \right)$. Therefore,

$$T^* \{S, N\} = \{S, N\} = (T^* S^2)_0^{(S^1)}. \tag{2.4.4}$$

Note that O is diffeomorphic to T^*L, where the cotangent fibers are mapped onto parabolas in planes perpendicular to L. Then, since by the general theory we can identify $T^*\left(S^2_{(1)}/S^1\right)$ and $S_{(S^1)\preceq(1)}$ with the images in $\mathcal{C}$ of B and A respectively, we have

$$T^*\left(S^2_{(1)}/S^1\right) = \{\boldsymbol{\sigma} \in \mathcal{C} \mid \sigma_1 \neq -1, 1\} = O \tag{2.4.5}$$

$$S_{(S^1)\preceq(1)} = \{\boldsymbol{\sigma} \in \mathcal{C} \mid \sigma_1 = \pm 1,\ \sigma_4 > 0\} = E_S \cup E_N. \tag{2.4.6}$$

From (2.4.4), (2.4.5), (2.4.6) and Theorem 2.4.2 it is straightforward to obtain the secondary stratification of each symplectic stratum

$$(T^*S^2)_0^{(S^1)} = T^*\left(S^2_{(S^1)}/S^1\right), \quad \text{and}$$

$$(T^*S^2)_0^{(1)} = T^*\left(S^2_{(1)}/S^1\right) \bigcup S_{(S^1)\preceq(1)},$$

with secondary frontier conditions $S_{(S^1)\preceq(1)} \subset \partial T^*\left(S^2_{(1)}/S^1\right)$.

The Coisotropic Stratification. Applying Theorem 2.4.4 and (2.4.2) we conclude

$$(T^*S^2)_0 = T^*\left(S^2_{(1)}/S^1\right) \bigcup S_{(S^1)\preceq(1)} \bigcup T^*\left(S^2_{(S^1)}/S^1\right),$$

with frontier conditions

$$\begin{aligned} T^*\left(S^2_{(S^1)}/S^1\right) &\subset \partial S_{(S^1)\preceq(1)} \\ S_{(S^1)\preceq(1)} &\subset \partial T^*\left(S^2_{(1)}/S^1\right). \end{aligned}$$

The Symplectic Structure on the Reduced Spaces. To determine the symplectic structure on the reduced spaces, that is, on the strata of the symplectic stratification, one proceeds as follows. First, one computes the Poisson algebra on $\mathbb{R}^6$ in terms of the invariants $\sigma_1, \dots, \sigma_6$. This Poisson algebra has two Casimir functions, namely, $C_1(\sigma_1, \dots, \sigma_6) := \sigma_5^2 + \sigma_6^2 - \sigma_2(\sigma_4 - \sigma_3^2)$ and $C_2(\sigma_1, \dots, \sigma_6) := \sigma_5$. Since C_1 is a Casimir function on $\mathbb{R}^6$ and $(T^*\mathbb{R}^3)/S^1$ is described as a subset of $\mathbb{R}^6$ by the conditions $C_1 = 0$, $\sigma_2 \geq 0$, and $\sigma_4 - \sigma_3^2 \geq 0$, it follows that the restriction of this Poisson bracket to this C_1-level set remains unchanged. Second, note that $\mathcal{C} \subset (T^*\mathbb{R}^3)/S^1$ is obtained by imposing the three additional constraints $C_2 = 0$, $\sigma_1^2 + \sigma_2 = 1$, and $\sigma_6 + \sigma_1\sigma_3 = 0$. Using the Dirac formula for the induced Poisson bracket on a cosymplectic submanifold (see [HRed] Proposition 10.5.8, [MandS] Proposition 8.5.1, Cushman and Bates [1997], Appendix A4, page 296) for these constraints, one easily computes the Poisson algebra on $\mathcal{C} = (T^*S^2)_0$ to be given by

$$\{\sigma_1, \sigma_3\} = 1 - \sigma_1^2, \quad \{\sigma_1, \sigma_4\} = 2\sigma_3, \quad \{\sigma_3, \sigma_4\} = -2\sigma_1\sigma_4.$$

See e.g. Cushman and Bates [1997], page 155 for this explicit computation. The restrictions of this Poisson algebra on $\mathcal{C}$ to $(T^*S^2)_0^{(S^1)}$ and $(T^*S^2)_0^{(1)}$ induce the smooth reduced symplectic structures on these symplectic manifolds.

Finally, the bundle structure of the reduced space $(T^*S^2)_0 \to S^2/S^1$ corresponds to the projection $\mathcal{C} \to S \cup L \cup N$ induced by the map $(\sigma_1, \sigma_3, \sigma_4) \mapsto (\sigma_1, 0, 0)$.

The Spherical Pendulum.   In the previous picture, we can regard the coordinate σ_1 as the height value of any S^1-orbit in S^2 projecting onto the point $(\sigma_1, 0, 0)$.

The Hamiltonian of the spherical pendulum is the function on T^*S^2 given by $H(\mathbf{x}, \mathbf{y}) := \frac{1}{2}m\|\mathbf{y}\|^2 + mgx_3$. The reduced Hamiltonian is therefore the restriction of the function

$$h(\sigma_1, \sigma_3, \sigma_4) := \frac{1}{2}m\sigma_4 + mg\sigma_1$$

to $\mathcal{C}$. The level sets of h are planes P_h intersecting the (σ_1, σ_3)-plane in the line $\sigma_1 = \frac{h}{mg}$ with the acute angle

$$\theta = \arccos\left(\frac{1}{\sqrt{1 + 4g^2}}\right).$$

The nonempty intersection of a plane P_h with $\mathcal{C}$ is zero or one dimensional and closed, hence it corresponds to a periodic orbit of energy h. There are three different qualitative cases, which are illustrated in Figure 2.4.2:

1. $h = -mg$: Then $P_h \cap \mathcal{C} = \{S\}$ and the reduced system is at a fixed equilibrium. This corresponds to the downward equilibrium in the unreduced pendulum.

2. $h \in (-mg, mg)$: In this case the reduced system evolves on a closed orbit whose projection onto the shape space S^2/S^1 is the closed interval $-1 \leq \sigma_1 \leq h/mg$. This interval is the range of heights for the corresponding S^1-orbit of dynamic orbits that lie in a vertical plane of the unreduced system with total energy h and momentum 0. Notice that in this case $P_h \cap \mathcal{C}$ is completely contained in $(T^*S^2)_0^{(1)}$ and that every orbit whose energy h is in $(-mg, mg)$ intersects the two secondary strata $T^*\left(S^2/S^1\right) = O$ and $S_{(S^1)\preceq(1)} = E_S \cup E_N$. Therefore the projection of the dynamics onto the shape space crosses several orbit types.

3. $h = mg$: In this case $P_h \cap \mathcal{C}$ consists of two different dynamical orbits, due to the fact that P_h intersects the two symplectic strata, which are invariant under the Hamiltonian evolution. Indeed, $P_h \cap (T^*S^2)_0^{(S^1)} =$
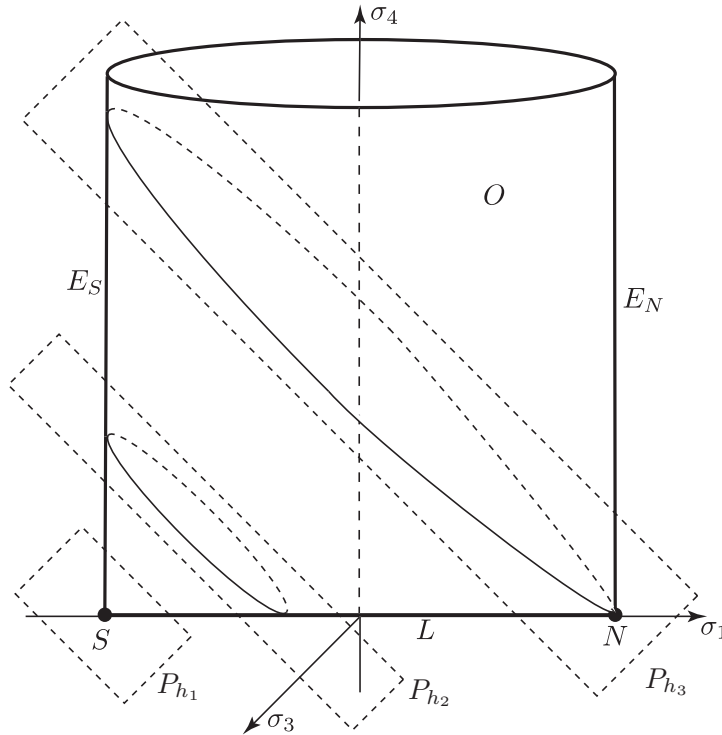
FIGURE 2.4.2. $h_1 = -mg$, $h_2 \in (-mg, mg)$, $h_3 = mg$.

$\{N\}$, which is a fixed equilibrium and $P_h \cap (T^*S^2)_0^{(e)}$ is a homoclinic orbit that projects onto $S^2/S^1 \backslash \{N\}$.

This analysis shows that the equilibrium $\{S\}$ is Lyapunov stable and that $\{N\}$ is unstable, both within $(T^*S^2)_0$, that is, relative to perturbations on $\mathcal{C}$ only. To see this, note that every open neighborhood of $\{S\}$ in $\mathcal{C}$ is filled with periodic orbits for $h \in [-mg, -mg+\epsilon]$. On the other hand, every orbit close to $\{N\}$ reaches the connected component E_S of the seam $S_{(S^1)\preceq(1)}$, and therefore $\{N\}$ is unstable.

3
The Problem Setting

3.1   The Setting for Reduction by Stages

The general problem and setting for reduction by stages is as follows. Consider a Lie group M that acts smoothly and properly on a manifold P; if we denote the M–action on P by $\Phi : M \times P \to P$, recall from the discussion of the symplectic reduction theorem in §1.2 that Φ is called a ***proper action*** if the following condition is satisfied: given two convergent sequences, $\{p_n\}$ and $\{g_n \cdot p_n\}$ in P, there exists a convergent subsequence $\{g_{n_k}\}$ in M.

Next, assume that P is a symplectic manifold and that the action of M has an equivariant momentum map $\mathbf{J}_M : P \to \mathfrak{m}^*$, where $\mathfrak{m}$ is the Lie algebra of M.

In Part I of the book, we assume that the M-action on P is not only proper, but also *free*. The definition of a *free action* (already discussed in §1.2) may be rephrased as the condition that the isotropy subgroup M_p of any element $p \in P$ is just the identity element of M. The main consequence of having a free and proper action is that the quotient space P/M is a smooth manifold with the property that the canonical projection of P to P/M is a smooth submersion. In Part II of the book, the freeness hypothesis will be dropped and in this case, one is in the world of singular reduction.

Now, let N be a closed 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 followed by a reduction by a group related to the quotient group M/N.

Reduction by Stages and Group Extensions.   There are many precursors to the general theory of reduction by stages, the main topic of this book. A simple version for the product of two groups was given in Marsden and Weinstein [1974]. Other versions are due to Sjamaar and Lerman [1991] and Landsman [1995, 1998].

The semidirect product reduction theorem (proved in Chapter 4) can be very nicely viewed using reduction by stages: one reduces T^*S by the action of the semidirect product group $S = G \, \circledS \, V$ in two stages, first by the action of V at a point a_0 and then by the action of G_{a_0}. Semidirect product reduction by stages for actions of semidirect products on general symplectic manifolds was developed and applied to underwater vehicle dynamics in Leonard and Marsden [1997]. Motivated partly by semidirect product reduction, [MMPR] gave the first indications of a comprehensive general theory of reduction by stages for group extensions. An interesting extension which is not a semidirect product is the Bott–Virasoro group, where the Gelfand-Fuchs cocycle may be interpreted as the curvature of a mechanical connection. We shall examine this example in detail later in the book.

Poisson Reduction by Stages.   Perhaps the simplest general context in which one can understand a reduction by stages result 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 [MandS] for an exposition of this standard theory). A relatively easy result that we prove in Proposition 5.3.1 is the following:

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

Symplectic reduction by stages is a more ambitious task, which keeps track of the symplectic leaves in this process.

Because of the close relation between point reduction and orbit reduction as well as the link between orbit reduction and Poisson reduction (that we have discussed in Theorem 1.2.4), one might think that orbit reduction would simplify the question of reduction by stages. In §5.3 we show that this is not the case; namely it leads to basically the same situation as in point reduction.

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

Commuting Reduction and Semidirect Products.   The simplest case of reduction by stages in the symplectic context is that of *commuting reduction*, due to Marsden and Weinstein [1974], which will be treated in §4.1. In this case, the result simply says that if one has the *direct product* of two groups, then one can reduce by the subgroups one at a time and in either order and get the same result as reducing by the product group all at once.

The next simplest case, but already a good deal more sophisticated, is the case of semidirect product reduction. Because of the work on semidirect products of Ratiu [1981, 1982], Guillemin and Sternberg [1980, 1984], Marsden, Ratiu and Weinstein [1984a,b], a nice bit of insight emerged. Namely, reduction by a semidirect product $S = G \,\circledS\, V$ of a Lie group G and a vector space V is most conveniently viewed as symplectic reduction by stages: first by the vector space V and second by an appropriate subgroup of G. This point of view naturally suggests a general setting for the problem: a big group M with a normal subgroup N. As we have mentioned earlier, this theory has many interesting applications, such as that of underwater vehicle dynamics, which was developed in Leonard and Marsden [1997], just to mention one. These applications were of great benefit in refining the setting and providing further motivation. For example, if one has the Euclidean group SE(3), then one must first reduce by translations and then by rotations to get the same effect as reducing by SE(3) all at once. We treat this case in detail in §4.2.

Lagrangian and Routh Reduction by Stages.   As was mentioned in the last section on the history of reduction theory, the work of Cendra, Marsden, and Ratiu [2001a] carries out the analogue of *Poisson reduction by stages* for the case of Lagrangian reduction. As in that work, we will need to keep track of cocycles, curvatures, as well as the closely associated notion of magnetic terms. As mentioned in the last section, the Lagrangian analog of *symplectic reduction*, is *Routh reduction*. Carrying out Routh reduction by stages, which has not yet been done, would be the Lagrangian analog of the main topic of this book.

For Lie groupoid and algebroid approaches to Lagrangian reduction by stages and related topics, see Mestdag [2005] and Cortés et al. [2006] and references therein.

Symplectic Induction and Induction by Stages.   Analogous to the reduction procedure is the method of *symplectic induction* that owes its name to the induction method in representation theory. Symplectic induction is a construction that does involve reduction methods, but is a development that is modeled after its representation counterpart.

We shall state a result of Mackey, called *induction by stages*. We mention this only for the information of the reader, who may have thought that it is closely related to reduction by stages. However, as we shall see, it is, in reality, a different construction.

First of all, we recall the notion of a unitary representation. Consider a separable locally compact topological group G and a Hilbert space V. A **unitary representation** of G on V is an action $\rho : G \times V \to V$ such that for each $g \in G$, the map $\rho_g : V \to V$ is a linear unitary operator for every $g \in G$, and, as usual for actions, $\rho_{g_1 g_2} = \rho_{g_1} \circ \rho_{g_2}$ for all $g_1, g_2 \in G$, and in addition, for each fixed $v \in V$ the map $g \mapsto \rho_g(v) \in V$ is continuous.

Classically, induced representations are obtained in the following manner (see, for example, Kirillov [1976a] or Mackey [1968]). Assume that H is a closed subgroup of G and that $\sigma : H \times W \to W$ is a unitary representation of H on a Hilbert space W, whose inner product is denoted $\langle \, , \, \rangle$. Let $G/H := \{Hg \mid g \in G\}$ be the quotient topological space of right cosets of H in G and assume that μ is an invariant Borel measure on G/H, that is, $\mu(Sg) = \mu(S)$ for any Borel set $S \subset G/H$ and any $g \in G$. The unitary H-representation σ naturally induces a unitary G-representation in the following manner. Define the *induced Hilbert space* $\mathrm{Ind}_H^G(W)$ to be the set of all Borel measurable functions $f : G \to W$ such that

(a) $f(hg) = \sigma_h f(g)$, for all $g \in G$ and $h \in H$,

(b) $g \in G \mapsto \langle f(g), f(g) \rangle \in \mathbb{C}$ is Borel measurable, and

(c) $\displaystyle \int_{G/H} \langle f(g), f(g) \rangle \mu(Hg) < \infty.$

Note that the expression in the last condition makes sense since

$$\langle f(hg), f(hg) \rangle = \langle \sigma_h f(g), \sigma_h f(g) \rangle = \langle f(g), f(g) \rangle$$

for any $h \in H$, because σ_h is a unitary operator for each $h \in H$. The vector space $\mathrm{Ind}_H^G(W)$ is a Hilbert space relative to the inner product

$$\langle f_1, f_2 \rangle := \int_{G/H} \langle f_1(g), f_2(g) \rangle \mu(Hg)$$

for $f_1, f_2 \in \mathrm{Ind}_H^G(W)$. Finally, define the unitary representation $\mathrm{Ind}_H^G(\sigma)$ of G on $\mathrm{Ind}_H^G(W)$ by

$$\left[\mathrm{Ind}_H^G(\sigma)_g(f) \right](g') := f(g'g),$$

where $g, g' \in G$. This unitary G-representation is said to be *induced* by the given unitary H-representation.

Mackey's *Induction by Stages Theorem* (see Mackey [1968], page 34) states that if $H \subset K \subset G$ is a chain of closed subgroups, then

$$\mathrm{Ind}_K^G \circ \mathrm{Ind}_H^K = \mathrm{Ind}_H^G.$$

There is an analogous symplectic induction construction that is given as follows. (See (Duval, Elhadad, and Tuynman [1992]; Guillemin and Sternberg [1982b, 1983]; Kazhdan, Kostant and Sternberg [1978]; Weinstein [1978a]; Zakrzewski [1982] for various versions of this construction). Let (P, Ω) be a symplectic manifold acted upon in a Hamiltonian fashion by a closed subgroup H of a Lie group G with associated equivariant momentum map $\mathbf{J}_P^H : P \to \mathfrak{h}^*$. The goal of the induction construction is to find

a Hamiltonian G-space $(\mathrm{Ind}_H^G(P), \Omega_{\mathrm{Ind}_H^G(P)}, G, \mathbf{J}_{\mathrm{Ind}_H^G(P)}^G)$ naturally induced by the Hamiltonian H-space $(P, \Omega, H, \mathbf{J}_P^H)$. To achieve this, one begins by considering the left action of H on G given by $h \cdot g := gh^{-1}$, for any $h \in H$ and $g \in G$. The standard momentum map $\mathbf{J}_{T^*G}^H : T^*G \to \mathfrak{h}^*$ of the cotangent lift of this H-action to $(T^*G, -\mathbf{d}\theta)$, where θ is the canonical one-form of the cotangent bundle T^*G, is given by $\mathbf{J}_{T^*G}^H(\alpha_g) = -T_e^* L_g(\alpha_g)|_{\mathfrak{h}}$. If one trivializes T^*G via left translations as $G \times \mathfrak{g}^*$, this momentum map becomes $\mathbf{J}_{G \times \mathfrak{g}^*}^H(g, \mu) = -\mu|_{\mathfrak{h}}$. On the symplectic manifold $(P \times T^*G, \Omega \oplus (-\mathbf{d}\theta))$ the H-action that is given in left trivialization by

$$h \cdot (z, g, \mu) := (h \cdot z, gh^{-1}, \mathrm{Ad}_{h^{-1}}^* \mu)$$

is free, proper (since H is closed), symplectic, and admits an equivariant momentum map $\mathbf{J}_{P \times G \times \mathfrak{g}^*}^H := \mathbf{J}_P^H + \mathbf{J}_{G \times \mathfrak{g}^*}^H : P \times G \times \mathfrak{g}^* \to \mathfrak{h}^*$; in other words, $\mathbf{J}_{P \times G \times \mathfrak{g}^*}^H(z, g, \mu) = \mathbf{J}_P^H(z) - \mu|_{\mathfrak{h}}$. Since the H-action is free, $0 \in \mathfrak{h}^*$ is a regular value and one can form the reduced symplectic manifold at zero, namely,

$$(\mathrm{Ind}_H^G(P), \Omega_{\mathrm{Ind}_H^G(P)}) := ((\mathbf{J}_{P \times G \times \mathfrak{g}^*}^H)^{-1}(0)/H, \Omega_0).$$

Now note that the lift to T^*G of left translations of G on itself, in left trivialization, is the G-action on $G \times \mathfrak{g}^*$ given by $g' \cdot (g, \mu) := (g'g, \mu)$ which has an equivariant momentum map given by $\mathbf{J}_{G \times \mathfrak{g}^*}^G(g, \mu) = \mathrm{Ad}_{g^{-1}}^* \mu$. Extend this action to $P \times G \times \mathfrak{g}^*$ by letting G act trivially on P to obtain a Hamiltonian G-action on $P \times G \times \mathfrak{g}^*$ with equivariant momentum map $\mathbf{J}_{P \times G \times \mathfrak{g}^*}^G(z, g, \mu) = \mathrm{Ad}_{g^{-1}}^* \mu$. Since this G-Hamiltonian action commutes with the H-action on $P \times G \times \mathfrak{g}^*$ and leaves $\mathbf{J}_{P \times G \times \mathfrak{g}^*}^H$ invariant, it induces the G-Hamiltonian action $g' \cdot [z, g, \mu] := [z, g'g, \mu]$ on $\mathrm{Ind}_H^G(P)$ whose G-equivariant momentum map $\mathbf{J}_{\mathrm{Ind}_H^G(P)}^G : \mathrm{Ind}_H^G(P) \to \mathfrak{g}^*$ is given by $\mathbf{J}_{\mathrm{Ind}_H^G(P)}^G([z, g, \mu]) = \mathrm{Ad}_{g^{-1}}^* \mu$. In this way one obtains the Hamiltonian G-space

$$\left(\mathrm{Ind}_H^G(P), \Omega_{\mathrm{Ind}_H^G(P)}, G, \mathbf{J}_{\mathrm{Ind}_H^G(P)}^G \right).$$

As in the case of representations, there is a *symplectic induction by stages theorem* stating that if $H \subset K \subset G$ is a chain of closed subgroups then

$$\mathrm{Ind}_K^G \circ \mathrm{Ind}_H^K = \mathrm{Ind}_H^G.$$

The proof of this theorem is the subject of a future paper.

The notion of symplectic induction that we just reviewed should not be confused with the symplectic induction procedure introduced in Landsman [1995] based on previous work of Xu [1991]. Due to this terminological conflict, Landsman's construction was renamed "special symplectic induction" in Landsman [1998].

3.2 Applications and Infinite Dimensional Problems

This book deals largely with mathematical applications of the theory, such as to semidirect product theory with cocycles, cotangent bundle reduction theory with cocycles, central extensions, frozen Poisson structures in the KdV equation and the Camassa-Holm equation, and the coadjoint orbit structure of the Bott–Virasoro group. Some of these applications are quite interesting; for instance, it is shown that the curvature of an appropriate mechanical connection gives, in a natural way, the Gelfand-Fuchs cocycle. In addition, just as the semidirect product reduction theory leads to a nice computational tool for computing coadjoint orbits for semidirect products, so reduction by stages gives considerable insight into the classification of coadjoint orbits for group extensions.

Remarks on Physical Applications. Of course, in general, one of the great successes of reduction theory has been the advancement of stability theory, starting with Poincaré [1901a,b, 1910] and Arnold [1966a], and continuing with the energy-momentum method developed in Simo, Lewis and Marsden [1991]; see also Bloch, Krishnaprasad, Marsden and Ratiu [1996]. Apart from the many applications of standard semidirect product theory, we have already mentioned the significant application to underwater vehicle dynamics and stability by Leonard and Marsden [1997]. Applications to compressible fluids and MHD by Holm and Kupershmidt [1983a] and Marsden, Ratiu and Weinstein [1984a] are typical of applications in continuum mechanics. Another interesting application is to the Landau-Lifschitz equations by Patrick [1999].

An especially exciting area where these ideas are useful is that of locomotion and articulated underwater bodies (think of fish with flexible bodies). These ideas involve the notion of geometric phases; see Marsden and Ostrowski [1996] and Kanso, Marsden, Rowley, and Melli-Huber [2005]. The problem of an articulated body in a fluid has a very interesting infinite dimensional symmetry group containing the diffeomorphism group of the fluid particle relabeling symmetry as a normal subgroup and the quotient group is the Euclidean group, corresponding to translations and rotations of the articulated body as a whole.

Infinite Dimensional Problems. As has been mentioned, in the bulk of this book, all configuration spaces and groups will be assumed to be finite dimensional for simplicity. We will make a few formal excursions into the infinite dimensional case, such as with the example of loop groups, the Bott–Virasoro group, and fluids in a symmetric container. Certainly most of the results formally work in the infinite dimensional context, but there are the usual functional analytic caveats. These caveats concern the choice of proper function spaces for the problem at hand. For instance, in fluids, to

deal with the PDE issues, one uses the appropriate Sobolev spaces; on the other hand, with these spaces, the underlying group, namely the volume preserving diffeomorphism group, is not literally a Lie group and so one must modify the theory somewhat. On the other hand, some other groups, such as the gauge group of electromagnetism, is a literal Lie group when Sobolev topologies are used. When infinite domains are used, conditions at infinity must be imposed.

To overcome these difficulties even without the complications of reduction by stages, various pieces of technology have been developed. For instance, for fluids, Ebin and Marsden [1970] developed an appropriate setting as did Arms, Marsden and Moncrief [1982] and Śniatycki [1999] for Yang-Mills theories.

Of course the number of infinite dimensional examples that are interesting is large as has been discussed in the preceding paragraphs. In the text that follows, we discuss a few examples in which the functional analysis can be carried out in the context of reduction by stages to give a specific indication of what is involved in the infinite dimensional context. The reader can consult with these references to get an indication of how the rigorous infinite dimensional version of the theory goes.

1. **Loop Groups.** In §9.2 we treat a central extension of a special case of loop groups, namely the group of maps of S^1 to S^1 under pointwise multiplication. In this case the functional analytic difficulties are relatively minor and we sketch in that section how they may be dealt with.

2. **The Bott–Virasoro Group.** The Bott–Virasoro group is the central extension of the orientation preserving diffeomorphism group of the circle and it appears in many interesting applications, such as the KdV and the Camassa–Holm equations. In this case, the functional analytic difficulties are overcome in Gay-Balmaz [2007], as is discussed in §9.4.

3. **Fluids with a Spatial Symmetry.** In §9.5 we discuss the motion of an ideal incompressible fluid moving in a container with spatial symmetry. There are two commuting groups that act on this problem, namely the particle relabeling symmetry (the volume preserving diffeomorphism group of the container acting on itself the right) and the symmetry group of the container itself. This is another case in which the functional analytical technicalities can be worked out in detail. We sketch how this goes, referring to Gay-Balmaz and Ratiu [2006] for details.

Some other examples that might also be of interest to carry out the functional analytic details for are as follows.

1. **Inhomogeneous Incompressible Fluids.** Models of stratified flow in geophysics often start with incompressible inhomogeneous fluids or various approximations built out of such fluids. The "big group" M in this case consists of the semidirect product of the volume preserving diffeomorphisms of a region, say Ω, and the space of real valued functions $\mathcal{F}$ on Ω. The normal subgroup is $\mathcal{F}$. This example comes under the heading of semidirect product theory and is, in the finite dimensional case, treated in detail in this book. For some of the needed function spaces to make the general theory work, we refer to Marsden [1976] and for applications to the stability theory of stratified flow, we refer to Abarbanel, Holm, Marsden, and Ratiu [1986]. Of course some of the techniques used are similar to those employed in the case of just one group, as occurs in the theory of a homogeneous fluid, studied by Arnold [1966a]. Many of the function space issues needed to make the infinite dimensional theory rigorous were already settled in Ebin and Marsden [1970].

2. **Charged Fluids and Plasmas.** The Hamiltonian structure of the Maxwell-Vlasov equation was derived by reduction techniques in Marsden and Weinstein [1982]. In this case, the "big group" M is the semidirect product of the symplectic diffeomorphism group and the gauge group of electromagnetism, the latter being the normal subgroup N. Related examples, also relevant to reduction by stages are the equations for a charged fluid, studied by Marsden, Weinstein, Ratiu, Schmid, and Spencer [1982] and the Einstein–Yang-Mills equations studied by Arms [1979]. Some of the rigorous function space issues needed to do these examples in a precise way were studied by Arms, Marsden and Moncrief [1982], Eardley and Moncrief [1982] and by Śniatycki [1999], for example.

   The paper Cendra, Holm, Hoyle and Marsden [1998] studies the variational structure of the Maxwell-Vlasov equations, overcoming an interesting degeneracy of the problem noted early on by Marsden and Morrison that is not so different, in principle, from a degeneracy that one already sees in the heavy top–namely that the kinetic energy term in the Hamiltonian is linear in one of the basic field variables, the plasma density. We mention this (amongst the multitude of papers in this area) mainly because it makes use of a reduction by stages philosophy.

3. **The Automorphism Group of a Principal Bundle.** We now describe another group that is somewhat related to the above example and that shows up in a number of interesting field theories, including Yang-Mills theory and liquid crystals (see, for example, Holm [2002]) and Yang-Mills fluids (see, Gay-Balmaz and Ratiu [2007]). Consider a principal bundle P over a manifold B with a structure group a Lie

group G. Then the group M of automorphisms of P that cover diffeomorphisms of B (possibly with a condition imposed, such as volume preserving) is the "big group". The normal subgroup N is the group of automorphisms that cover the identity. This group is essentially the gauge group of field theories such as Yang-Mills fields; as above, this group in the infinite dimensional case has no difficulties (namely derivative losses) with its Lie group structure; see, for instance, Arms [1979] and Nill [1983] and references therein. However, when coupled with the diffeomorphism group, the Lie group structure clearly will get more complicated. This example is closely related to those considered later in §10.6.

For Yang-Mills fluids one uses the Kaluza-Klein point of view to formulate the equations of motion in both Lagrangian and Hamiltonian forms. Correspondingly, the fluid part of these equations has an Euler-Poincaré and Lie-Poisson formulation. The associated Poisson bracket couples a Lie-Poisson bracket with cocycle on a semidirect product with the canonical bracket on the cotangent bundle of the space of connections. As a consequence of the Lagrangian approach one gets a Kelvin-Noether theorem (see, Gay-Balmaz and Ratiu [2007]).

The Multisymplectic View. Many infinite dimensional systems may be regarded from the *multisymplectic* point of view, as in (to mention just a few references), Vinogradov and Kupershmidt [1977], Gotay, Isenberg, and Marsden [1997], Marsden and Shkoller [1999]; Marsden, Patrick, and Shkoller [1998] and the subject of multisymplectic reduction is still in its infancy (see Castrillón López, Ratiu, and Shkoller [2000] and Castrillón López and Marsden [2003] and references therein).

Part II. Regular Symplectic Reduction by Stages

In this part, we prove a general symplectic reduction by stages theorem in the case of regular (nonsingular) symplectic reduction and give several applications. Recall from the Introduction that the general setting is the following: we consider the action of a Lie group M acting freely and properly on a symplectic manifold P and form the (Meyer-Marsden-Weinstein) symplectic reduced manifold P_σ at a momentum value $\sigma \in \mathfrak{g}^*$. Assume that M has a closed normal subgroup N. A goal is to realize this reduced space P_σ in a two step procedure: first reducing by N and then by an appropriate group that is a modification of M/N. This construction generalizes and unifies several well known techniques, such as semidirect product reduction, as we shall explain in the main text. In fact, we shall give a reasonably complete discussion of the case of semidirect product reduction from the point of view of reduction by stages.

The bulk of the present Part II will make use of a special *stages hypothesis*, which is satisfied by all the examples that we are aware of. Moreover, in Chapter 11 we show that under appropriate conditions, this hypothesis is always satisfied in a large set of situations. For instance, it holds when M and N are compact groups.

Besides giving a general result in the context of symplectic manifolds, we also study the case of cotangent bundles in some detail and make use of the curvature of the mechanical connection, which provides both a magnetic term as well as the cocycle associated with the group extension.

We apply this theory to several examples, including the Heisenberg group, loop groups, the Bott–Virasoro group, and the classification of coadjoint orbits for group extensions.

4
Commuting Reduction and Semidirect Product Theory

In this chapter we develop two of the basic results on reduction by stages, namely the case of commuting reduction and semidirect product reduction. While one could view these as special cases of more general theorems to follow in the next chapter, it is worthwhile to see them on their own as more structured preludes to more general cases. In addition, these cases are important in applications as well as for the historical development of the subject.

4.1 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 when one is considering say, particle mechanics. 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 in some detail, namely the case of a *direct product*. For instance, for a symmetric heavy top, the symmetry group is $S^1 \times S^1$, with the first S^1 being the symmetry of rotations about the vertical axis of gravity and the second S^1 being rotations about the symmetry axis of the body. These two group actions commute.

The version of commuting reduction given in 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, which we will go through rather carefully in a way that facilitates its generalization, is given in the following development.

The set up is as follows: 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 $\mathbf{J}_K : P \to \mathfrak{k}^*$. Let G be another group (with Lie algebra $\mathfrak{g}$) acting on P with an equivariant momentum map $\mathbf{J}_G : P \to \mathfrak{g}^*$. The first main assumption is

C1. The actions of G and K on P commute.

It follows that there is a well-defined action of $G \times K$ on P given by $(g, k) \cdot z = g \cdot (k \cdot z) = k \cdot (g \cdot z)$. Next, we claim that

$$\mathbf{J}_{G \times K} := \mathbf{J}_G \times \mathbf{J}_K : P \to (\mathfrak{g} \times \mathfrak{k})^* = \mathfrak{g}^* \times \mathfrak{k}^*$$

is a momentum map for the action of $G \times K$ on P. Indeed, for $\xi \in \mathfrak{g}$ and $\eta \in \mathfrak{k}$, we have

$$(\xi, \eta)_P(z) = \xi_P(z) + \eta_P(z),$$

as follows by noting that $\exp(t(\xi, \eta)) = (\exp(t\xi), \exp(t\eta))$. Note that

$$\mathbf{J}_{G \times K}^{(\xi, \eta)} = \mathbf{J}_G^{\xi} + \mathbf{J}_K^{\eta}.$$

Therefore,

$$\mathbf{i}_{(\xi, \eta)_P}\Omega = \mathbf{i}_{\xi_P}\Omega + \mathbf{i}_{\eta_P}\Omega = \mathbf{d}\mathbf{J}_G^{\xi} + \mathbf{d}\mathbf{J}_K^{\eta} = \mathbf{d}\mathbf{J}_{G \times K}^{(\xi, \eta)},$$

This proves the claim.

To ensure that $\mathbf{J}_{G \times K}$ is an equivariant momentum map, we make an additional hypothesis.

C2. $\mathbf{J}_G$ is K-invariant and $\mathbf{J}_K$ is G-invariant.

There are some remarks to be made about this condition. First of all, if $P = T^*Q$ and the actions are lifted from commuting actions on Q, then we assert that the condition **C2** automatically holds. This is because, in

the cotangent case, we can use the explicit formula for the equivariant momentum maps $\mathbf{J}_G$ and $\mathbf{J}_K$. Let $k \in K$, $\alpha_q \in T_q^*Q$, and $\xi \in \mathfrak{g}$. Then

$$\begin{aligned}
\langle \mathbf{J}_G(k \cdot \alpha_q), \xi \rangle &= \langle k \cdot \alpha_q, \xi_Q(k \cdot q) \rangle \\
&= \langle k \cdot \alpha_q, k \cdot \xi_Q(q) \rangle \\
&= \langle \alpha_q, \xi_Q(q) \rangle \\
&= \langle \mathbf{J}_G(\alpha_q), \xi \rangle .
\end{aligned}$$

There is a similar argument for $\mathbf{J}_K$. This proves our assertion.

The second remark we wish to make is that in a sense, one needs to only assume that "half" of **C2** holds. Namely, we claim that if $\mathbf{J}_K$ is G-invariant and K is connected, then $\mathbf{J}_G$ is K-invariant. Indeed, $\mathbf{d} \langle \mathbf{J}_K, \eta \rangle \cdot \xi_P = 0$ for all $\xi \in \mathfrak{g}$ and $\eta \in \mathfrak{k}$ and hence

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

from which we conclude K-invariance of $\mathbf{J}_G$ by connectedness of K, which proves the claim.

Now we have the ingredients needed to get an equivariant momentum map.

4.1.1 Proposition.   *Under hypotheses* **C1** *and* **C2**, $\mathbf{J}_{G \times K}$ *is an equivariant momentum map for the action of $G \times K$ on P.*

Proof.   For all $z \in P$ and $(g, k) \in G \times K$ we have

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

where we have used equivariance of each of $\mathbf{J}_G$ and $\mathbf{J}_K$, the fact that the actions commute (condition **C1**), and condition **C2**, the invariance of $\mathbf{J}_G$ and $\mathbf{J}_K$. ∎

We need one more assumption.

C3. The action of $G \times K$ on P is free and proper.

Let $(\mu, \nu) \in \mathfrak{g}^* \times \mathfrak{k}^*$ be given. Since we have a simple product, the isotropy group is $(G \times K)_{(\mu, \nu)} = G_\mu \times K_\nu$. Our goal is to show that the "one-shot" reduced space

$$P_{(\mu, \nu)} = (\mathbf{J}_G \times \mathbf{J}_K)^{-1}(\mu, \nu)/(G_\mu \times K_\nu)$$

is symplectically diffeomorphic to the space obtained by first reducing by K at ν to form the first reduced space $P_\nu = \mathbf{J}_K^{-1}(\nu)/K_\nu$ and then reducing this space by the G action. Note that the actions of K and G on P are free and proper as a consequence of **C3**.

Warning.   If each of the actions of G and K are free, this need not imply, conversely, that the action of $G \times K$ is free. For example, Let $G = K = S^1$ act on $\mathbb{R}^2$ minus the origin by standard rotations. The actions obviously commute, each one is free, but the product action is not free since $(e^{i\theta}, e^{-i\theta})z = z$ for any θ and any nonzero $z \in \mathbb{R}^2$.

Another example where this occurs is in the Lagrange top, that is, a rigid body with an axis of symmetry, rotating about a fixed point on that axis, and moving in a gravitational field. There are two commuting S^1 symmetry groups acting on the phase space $T^*\mathrm{SO}(3)$. These two actions are given by (the cotangent lift of) left translation corresponding to rotations about the axis of gravity and the other by right translation about the axis of symmetry; these two actions clearly commute. 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. In this problem, one should omit the "vertical" state of rotation of the body in order for the action of $S^1 \times S^1$ to be free, even though each action separately is free; see, for instance, Lewis, Ratiu, Simo, and Marsden [1992].

To carry out the second stage reduction, we need the following.

4.1.2 Lemma.   *The group G induces a free and proper symplectic action on P_ν, and the map $\mathbf{J}_\nu : P_\nu \to \mathfrak{g}^*$ naturally induced by $\mathbf{J}_G$ is an equivariant momentum map for this action.*

Proof.   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 **C2**, 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 smooth 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 $\Psi_g \circ i_\nu = i_\nu \circ \Psi_g^\nu$.

Recall from Theorem 1.1.3 that the symplectic form Ω_ν on the reduced space P_ν 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 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. 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 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.

We next prove that the action of G on P_ν is free and proper. First note that the action of G on $\mathbf{J}_K^{-1}(\nu)$ is free and proper. For $z \in \mathbf{J}_K^{-1}(\nu)$, let its class be denoted $[z]_\nu := \pi_\nu(z)$. The G action in this notation is simply $g[z]_\nu = [gz]_\nu$. To check freeness, assume that $[gz]_\nu = [z]_\nu$. Thus, there is a $k \in K_\nu$ such that $kgz = z$. But $kgz = (g,k)z$ and hence freeness of the action of $G \times K$ (condition **C3**) implies that $g = e, k = e$. Thus, the action of G on P_ν is free.

To prove properness, let $[z_n]_\nu \to [z]_\nu$ and $[g_n z_n]_\nu \to [z']_\nu$. Since the action of K_ν on $\mathbf{J}_K^{-1}(\nu)$ is free and proper, by the definition of the quotient topology, and the fact that proper actions have slices (see the discussions in, for example, [MTA] and Duistermaat and Kolk [1999]), there are sequences $k_n, k_n' \in K_\nu$ such that $k_n z_n \to z$ and $k_n g_n z_n = g_n k_n z_n \to z'$ (since the actions commute). By properness of the original action, this implies that a subsequence of g_n converges. $\blacksquare$

With the above ingredients, we can now form the second reduced space, namely $(P_\nu)_\mu = \mathbf{J}_\nu^{-1}(\mu)/G_\mu$.

4.1.3 Theorem (Commuting Reduction Theorem). *Under the hypotheses* **C1, C2, C3**, $P_{(\mu,\nu)}$ *and* $(P_\nu)_\mu$ *are symplectically diffeomorphic.*

Proof.   Composing the inclusion map

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

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 $\mathbf{J}_\nu \circ \pi_\nu = \mathbf{J}_G \circ i_\nu$. Thus, we get a map

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

such that $(i_\nu)_\mu \circ \kappa_\nu = \pi_\nu \circ j$, where we use the notation $(i_\nu)_\mu$ for the inclusion $\mathbf{J}_\nu^{-1}(\mu) \hookrightarrow P_\nu$. The map κ_ν 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

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

To show that this map is symplectic, it is enough to show that

$$\pi_{(\mu,\nu)}^* \left([\kappa_\nu]^*(\Omega_\nu)_\mu\right) = \pi_{(\mu,\nu)}^* \Omega_{(\mu,\nu)}, \qquad (4.1.1)$$

where we use self-explanatory notation; $\Omega_{(\mu,\nu)}$ is the symplectic form on $P_{(\mu,\nu)}$, $\pi_{(\mu,\nu)} : (\mathbf{J}_G \times \mathbf{J}_K)^{-1}(\mu,\nu) \rightarrow P_{(\mu,\nu)}$ is the projection, $(\pi_\nu)_\mu : \mathbf{J}_\nu^{-1}(\mu) \rightarrow (P_\nu)_\mu$ is the projection, and $(\Omega_\nu)_\mu$ is the reduced symplectic form on $(P_\nu)_\mu$. It is enough to establish equation (4.1.1) since $\pi_{(\mu,\nu)}$ is a surjective submersion. The right hand side of (4.1.1) is given by

$$i_{(\mu,\nu)}^* \Omega$$

by the unique characterization of the reduced symplectic form $\Omega_{(\mu,\nu)}$. The left side is

$$\pi_{(\mu,\nu)}^* \left([\kappa_\nu]^*(\Omega_\nu)_\mu\right) = \kappa_\nu^*(\pi_\nu)_\mu^*(\Omega_\nu)_\mu = \kappa_\nu^*(i_\nu)_\mu^* \Omega_\nu$$

because of the relation $[\kappa_\nu] \circ \pi_{(\mu,\nu)} = (\pi_\nu)_\mu \circ \kappa_\nu$ and the unique characterization of the reduced symplectic form $(\Omega_\nu)_\mu$. However, since $(i_\nu)_\mu \circ \kappa_\nu = \pi_\nu \circ j$, we get

$$\kappa_\nu^*(i_\nu)_\mu^* \Omega_\nu = j^* \pi_\nu^* \Omega_\nu = j^* i_\nu^* \Omega,$$

by the unique characterization of the reduced symplectic form Ω_ν. Since $i_\nu \circ j = i_{(\mu,\nu)}$ we get the desired equality. Thus, $[\kappa_\nu] : P_{(\mu,\nu)} \rightarrow (P_\nu)_\mu$ is a symplectic map.

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

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

as follows. Choose an equivalence class $[z]_\nu \in \mathbf{J}_\nu^{-1}(\mu) \subset P_\nu$ for $z \in \mathbf{J}_K^{-1}(\nu)$. 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 $z \in (\mathbf{J}_G \times \mathbf{J}_K)^{-1}(\mu,\nu)$ since by construction $z \in \mathbf{J}_K^{-1}(\nu)$ and also

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

Hence, it makes sense to consider the class $[z]_{(\mu,\nu)} \in P_{(\mu,\nu)}$. The result is independent of the representative, since any other representative of the same class has the form $k \cdot z$ 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

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

Chasing the definitions shows that this map is the inverse of the map $[\kappa_\nu]$. Thus, both are bijections. Since $[\kappa_\nu]$ is smooth and symplectic, it is an immersion. A dimension count shows that $(P_\nu)_\mu$ and $P_{(\mu,\nu)}$ have the same dimension. Thus, $[\kappa_\nu]$ is a bijective local diffeomorphism, so it is a diffeomorphism. ∎

The above theorem on commuting reduction may be viewed in the general context discussed in §3.1 by taking $M = G \times K$ with the normal subgroup N being chosen to be either G or K, so that the quotient group of M is the other group.

It is instructive to build up to the general reduction by stages theorem by giving direct proofs of some simpler special cases, such as the one at hand and the case of semidirect products treated in §4.2; these special cases not only point the way to the general case, but contain interesting constructions that are relevant to these more specific cases. The general case has some subtleties not shared by these simpler cases which will be spelled out as we proceed.

4.2   Semidirect Products

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 [1980b], Guillemin and Sternberg [1980], Ratiu [1981, 1982], Marsden [1982], Marsden, Weinstein, Ratiu, Schmid, and Spencer [1982], Holm and Kupershmidt [1983a] 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 cited and references therein. Another, and very useful, concrete application of this theory is to underwater vehicle dynamics; see Leonard and Marsden [1997].

Generalities on Semidirect Products.   We begin by recalling some definitions and properties of semidirect products. Let V be a vector space and assume that the Lie group G (with Lie algebra $\mathfrak{g}$) acts (on the left) by linear maps on V, and hence G also acts (also on the left) on its dual space V^*, the action by an element g on V^* being the transpose of the action of g^{-1} on V. As sets, the semidirect product $S = G \circledS V$ is the Cartesian product $S = G \times V$ and 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 $\xi_1 v_2$.

Perhaps the most basic example of a semidirect product is the Euclidean group $SE(3)$ of $\mathbb{R}^3$, which is studied in, for example, [MandS] and which will be treated in some detail in §4.4.

We will need the formulas for the adjoint action and the coadjoint action for semidirect products. Denoting these and other actions by simple concatenation, they are given by

$$(g, v)(\xi, u) = (g\xi, gu - \rho_v(g\xi)) \tag{4.2.1}$$

and

$$(g, v)(\mu, a) = (g\mu + \rho_v^*(ga), ga), \tag{4.2.2}$$

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.   Recall from [MandS] that the Lie-Poisson bracket on the dual of a Lie algebra $\mathfrak{g}^*$ comes with two signs and is given on two functions F, K of $\mu \in \mathfrak{g}^*$ by

$$\{F, K\}_\pm(\mu) = \pm \left\langle \mu, \left[\frac{\delta F}{\delta \mu}, \frac{\delta K}{\delta \mu} \right] \right\rangle \tag{4.2.3}$$

Recall also that this bracket is obtained naturally from the canonical bracket on T^*G by taking quotients—this is the Lie-Poisson reduction theorem that is found in [MandS], Chapter 13. The minus sign corresponds to reduction by the left action and the plus sign to reduction by the right action.

Next, we give the formula for the $\pm$ Lie-Poisson bracket on a semidirect product; namely, for $F, K : s^* \to \mathbb{R}$, their **_semidirect product bracket_** is given by:

$$\{F, K\}_\pm(\mu, a) = \pm \left\langle \mu, \left[\frac{\delta F}{\delta \mu}, \frac{\delta K}{\delta \mu} \right] \right\rangle \pm \left\langle a, \frac{\delta F}{\delta \mu} \cdot \frac{\delta K}{\delta a} - \frac{\delta K}{\delta \mu} \cdot \frac{\delta F}{\delta a} \right\rangle, \tag{4.2.4}$$

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 a smooth function $H : s^* \to \mathbb{R}$ is given by

$$X_H(\mu, a) = \mp \left(\mathrm{ad}_{\delta H/\delta \mu}^* \mu - \rho_{\delta H/\delta a}^* a, \frac{\delta H}{\delta \mu} \cdot a \right). \tag{4.2.5}$$

Semidirect Product Reduction Theorem—Statement. We next state the "classical" semidirect product reduction theorem and following this, we give a more general theorem concerning actions by semidirect products on symplectic manifolds. The strategy will be to obtain the classical result as a special case of the more general result, which we shall prove in detail.

The semidirect product reduction theorem states, roughly speaking, that for a semidirect product $S = G \circledS V$, one can first reduce T^*S by V and then by G and one gets the same result as reducing by S. The precise statement of the classical semidirect product reduction theorem is as follows.

4.2.1 Theorem (Semidirect Product Reduction Theorem). *As above, let $S = G \circledS V$, choose $\sigma = (\mu, a) \in \mathfrak{g}^* \times V^*$ and reduce T^*S by the action of S at σ, which, by Corollary 1.1.4 gives the coadjoint orbit $\mathcal{O}_\sigma$ through $\sigma \in \mathfrak{s}^*$. Then, 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.*

Remark. Note that in the semidirect product reduction theorem, only $a \in V$ and $\mu|\mathfrak{g}_a \in \mathfrak{g}_a^*$ are used in the equivalent description of the coadjoint orbit. 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.

The preceding result will next be shown to be a special case of a theorem we shall prove on reduction by stages for semidirect products acting on a symplectic manifold

Semidirect Product Actions. We now set the stage for the statement of the more general reduction by stages result for semidirect product actions. Start with a free and proper symplectic action of a semidirect product $S = G \circledS V$ 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 (closed, 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 $i_V : V \to \mathfrak{s}$, given by $v \mapsto (0, v)$, is the inclusion where and $i_V^* : \mathfrak{s}^* \to V^*$ is its dual.

We carry out the reduction of P by S at a value $\sigma = (\mu, a)$ of the momentum map $\mathbf{J}_S$ for S (it is a regular value because of the freeness assumption) in two stages using the following procedure.

- First, reduce P by V at the value a (it follows from the freeness assumption that this too is a regular value) to get the first reduced space $P_a = \mathbf{J}_V^{-1}(a)/V$.

- Form the isotropy subgroup 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 G_a acts freely and properly 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.

- Second, reduce P_a at the point $\mu_a := \mu|\mathfrak{g}_a$ to get the second reduced space $(P_a)_{\mu_a} = \mathbf{J}_a^{-1}(\mu_a)/(G_a)_{\mu_a}$.

4.2.2 Theorem (Reduction by Stages for Semidirect Product Actions)**.**
The two-stage reduced space $(P_a)_{\mu_a}$ is symplectically diffeomorphic to the "all-at-once" reduced space P_σ obtained by reducing P by the whole group S at the point $\sigma = (\mu, a)$.

We have made the free and proper assumption on the action of S in this case that is the analog of the hypothesis **SRFree** in the symplectic reduction Theorem 1.1.3. One can also make hypotheses analogous to **SR-Regular**, but these assumptions would need to be imposed at each of the stages. We have used the free and proper assumption since, as we shall see, it is automatically inherited in each of the two stages.

Special Cases. We recover the classical semidirect product reduction Theorem 4.2.1 by choosing $P = T^*S$ and using the fact that the first reduced space, namely reduction by V, is just T^*G with its canonical symplectic structure. We shall go through this in detail in §4.3.

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. The fact that the full commuting reduction theorem is not literally as special case 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, in fact, more general results in this direction 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 4.2.1 along lines somewhat different than we shall present here. The proofs we give in this book have the advantage that they work for more general reduction by stages theorems.

Classifying Orbits. Combined with the cotangent bundle reduction theorem (as mentioned in the introductory chapter, the reader may consult either [FofM] or [LonM] for an exposition), 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 problem in detail starting with the Euclidean group in §4.4.

Reducing Dynamics. There is a method for reducing dynamics that is associated with the geometry of the semidirect product reduction theorem. One can start with a Hamiltonian on either of the phase spaces $(P_a)_{\mu_a}$ or P_σ and induce one (and hence its associated dynamics) on the other space in a natural way.

Another view of reducing dynamics that is useful in many applications is as follows: 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, if one prefers, on the coadjoint orbits of $\mathfrak{s}^*$. For example, this observation is extremely useful in underwater vehicle dynamics (again, see Leonard and Marsden [1997]).

The Momentum Map for the V-action. We now work towards a proof of reduction by stages for semidirect product actions, Theorem 4.2.2. We first elaborate on the constructions in the statement of the theorem.

Thus, we start by considering a given 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 also 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.

Equivariance of $\mathbf{J}_S$ under G implies the following relation for $\mathbf{J}_V$:

$$\mathbf{J}_V(gz) = g\mathbf{J}_V(z), \tag{4.2.6}$$

where $z \in P$ and we denote the appropriate action of $g \in G$ on an element by concatenation, as before. To prove equation (4.2.6), one uses the fact that for the coadjoint action of S on $\mathfrak{s}^*$ the second component of that action is just the dual of the given action of G on V, which is evident from equation (4.2.2).

The Reduction by Stages Construction. We now elaborate on the reduction by stages construction given in Theorem 4.2.2. An important step will be to show that the construction is, in fact, well-defined.

The "one-shot" reduction step is, in principle, straightforward: one carries out reduction of P by S at a regular value $\sigma = (\mu, a)$ of the momentum map $\mathbf{J}_S$ for S.

On the other hand, in reduction by stages, one carries out the reduction in the following two stages (see Figure 4.2.1).

- First, reduce P by V at the value $a \in V^*$. Since the action of S was assumed to be free and proper, so is the action by V and hence a is a regular value. Thus, we get the reduced manifold $P_a = \mathbf{J}_V^{-1}(a)/V$. Since the reduction is by an Abelian group, the quotient is taken using the whole of V. We will denote the projection to the reduced space 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^*. We will need to show 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, using this action of G_a, reduce P_a at the point $\mu_a := \mu|\mathfrak{g}_a$ to get the reduced manifold $(P_a)_{\mu_a} = \mathbf{J}_a^{-1}(\mu_a)/(G_a)_{\mu_a}$.

To prove the result, we will systematically check these claims and after doing this, we will set up the symplectic isomorphism.

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

4.2.3 Lemma.   *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 (4.2.6) noted above, we have

$$\mathbf{J}_V(gz) = g\mathbf{J}_V(z) = ga = a.$$

Thus, G_a acts on the set $\mathbf{J}_V^{-1}(a)$.   ▼

4.2.4 Lemma.   *The action of G_a on $\mathbf{J}_V^{-1}(a)$ constructed in the preceding lemma, induces a free and proper 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 (relative to the action of G_a), then we claim that $g[z]_a = [gz]_a$ defines the action. We first show that it is well-defined.
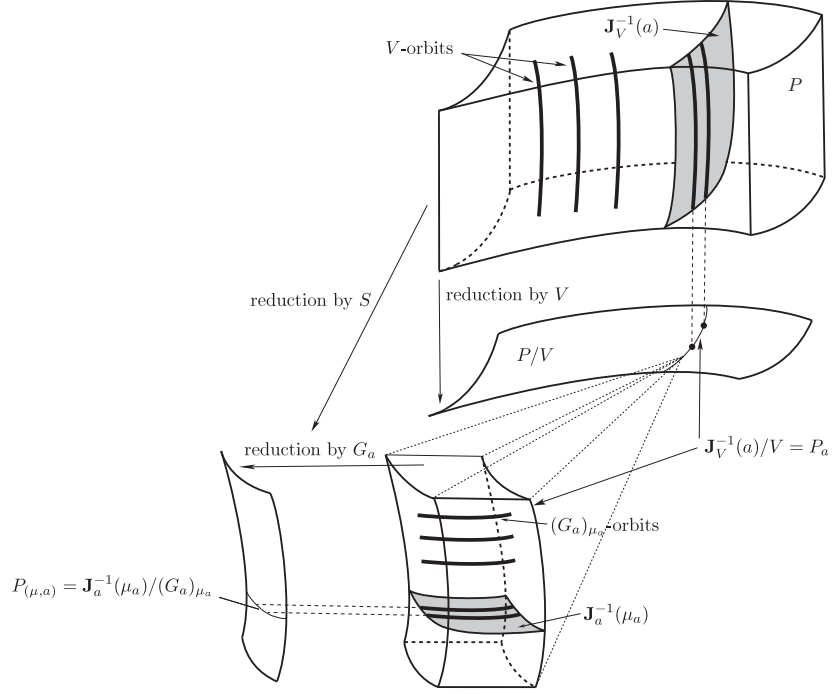
FIGURE 4.2.1. A schematic of reduction by stages for semidirect products.

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

$$
\begin{aligned}
[gvz]_a &= [(g, 0)(e, v)z]_a = [(g, gv)z]_a \\
&= [(e, gv)(g, 0)z]_a = [(gv)(gz)]_a \\
&= [gz]_a.
\end{aligned}
$$

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.

This action is free because if $[gz]_a = [z]_a$, then there is a $v \in V$ such that $vgz = z$. Since $vg = (g, v)$, freeness of the S-action implies that $g = e$ and $v = 0$.

To show properness, assume $[z_n]_a \to [z]_a$ and that $[g_n z_n]_a \to [z']_a$. We must find a convergent subsequence $g_{n_p} \in G_a$. There are sequences $v_n \in V$ and $v'_n \in V$ such that $v_n z_n = (e, v_n)z_n \to z$ and $v'_n g_n z_n = (g_n, v'_n)z_n \to z'$. Write

$$
\begin{aligned}
(g_n, v'_n)z_n &= (g_n, v'_n)(e, v_n)^{-1}(e, v_n)z_n \\
&= (g_n, v'_n - v_n)(e, v_n)z_n
\end{aligned}
$$

Thus, $(g_n, v'_n - v_n)$ has a convergent subsequence, by properness of the S action on P and hence the first components also form a convergent subsequence. Since G_a is closed and $g_n \in G_a$, we get a convergent subsequence in G_a.   ▼

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

4.2.5 Lemma.   *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 let the inclusion be denoted $i_a : \mathbf{J}_V^{-1}(a) \to P$. Denote by $\Psi_g : P \to P$ the action of $g \in G$ on P. The preceding lemma 4.2.4 shows 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 unique characterization $i_a^* \Omega = \pi_a^* \Omega_a$ of the reduced symplectic form Ω_a on P_a yields

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

Since π_a is a surjective submersion, we conclude that

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

Thus, the action of G_a on P_a is symplectic.   ▼

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

4.2.6 Lemma.   *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

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

In this equation, 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

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

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

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

Again, since π_a is a surjective submersion, it follows 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 4.2.2. Having established the preliminary facts in the preceding lemmas, 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

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

We assert that projection onto the first factor defines a smooth Lie group homomorphism $\psi : S_\sigma \to (G_a)_{\mu_a}$. In fact, 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. This proves the assertion.

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.

We will show that this map $[\pi_a \circ j]$ is a diffeomorphism by constructing an inverse. We will begin by showing how to define a map

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

Refer to Figure 4.2.2 for the spaces involved.

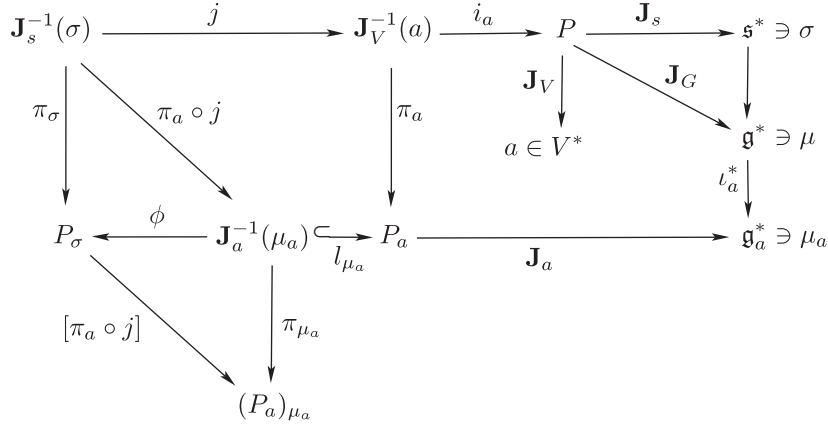FIGURE 4.2.2. Maps that are used in the proof of the semidirect product reduction theorem.

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

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

Thus, we require that

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

That this is possible follows from the next lemma.

4.2.7 Lemma (Annihilator Lemma).   *If* $\mathfrak{g}_a^\circ = \{\nu \in \mathfrak{g}^* \mid \nu|\mathfrak{g}_a = 0\}$ *denotes the annihilator of* $\mathfrak{g}_a$ *in* $\mathfrak{g}^*$, *then*

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

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 $F_0 \subset F$ 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 the element $\nu = \mu - \mathbf{J}_G(z)$, which is an element 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 $v \in V$ 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 then claim that we can 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.

---

[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.

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. Following the above procedure, choose 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.

Next, we 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 $g_0 \in (G_a)_{\mu_a}$. Choose $v \in V$ so that $vz \in \mathbf{J}_S^{-1}(\sigma)$ and let $u \in V$ be chosen so that $ug_0 z \in \mathbf{J}_S^{-1}(\sigma)$. We must show that $[vz]_\sigma = [ug_0 z]_\sigma$. 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 $\sigma = \mathbf{J}_S(vz) = \mathbf{J}_S(ug_0 z)$, 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 diffeomorphic 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 4.2.1.

2. The original proof of Theorem 4.2.1 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 concerning reduction by stages for SE(3), the special Euclidean group of $\mathbb{R}^3$. This illustrates some important aspects and applications of the classical semidirect product reduction Theorem 4.2.1.

4. We briefly describe two examples that require the more general result of Theorem 4.2.2.

    (a) First, consider a pseudo-rigid body in a fluid; that is, a body which can undergo linear deformations and moving in potential flow, as was the case for rigid bodies in potential flow 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.

    (b) 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.

    As we have mentioned before, the reduction by stages philosophy is quite helpful in understanding the dynamics and stability of underwater vehicle dynamics, as in Leonard and Marsden [1997].

4.3   Cotangent Bundle Reduction and Semidirect Products

The purpose of this section is to couple the semidirect product reduction theorem with cotangent bundle reduction to obtain a more detailed structure of the reduced spaces for the right cotangent lifted action of $G\,\textcircled{S}\,V$ on $T^*\left(G\,\textcircled{S}\,V\right)$. Of course, by Theorem 1.2.3 on reduction to coadjoint orbits, these reduced spaces are the coadjoint orbits of the group $G\,\textcircled{S}\,V$.

   To carry out this program, we first construct a mechanical connection on the bundle $G\,\textcircled{S}\,V \to G$ 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 cotangent bundle reduction again to complete the orbit classification.

Notation.   As in the preceding section, let $S = G\,\textcircled{S}\,V$ be the semidirect product of a Lie group G and a vector space V on which G acts, with multiplication

$$(g, v)(h, w) = (gh, v + gw), \tag{4.3.1}$$

where $g, h \in G$ and $v, w \in V$. The identity element is $(e, 0)$ and inversion is given by $(g, v)^{-1} = (g^{-1}, -g^{-1}v)$. Recall that the Lie algebra of S is the semidirect product $\mathfrak{s} = \mathfrak{g}\,\textcircled{S}\,V$ with the commutator

$$[(\xi, v), (\eta, w)] = ([\xi, \eta],\ \xi w - \eta v), \tag{4.3.2}$$

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 and to recall that we identify V with $\{e\} \times V$, a closed normal Lie subgroup of $G\,\textcircled{S}\,V$.

   The adjoint representation of S on $\mathfrak{s}$ given in equation (4.2.1) restricts to the S-representation on V given by $\mathrm{Ad}_{(g,v)}\, u = gu$ for any $g \in G$ and $u, v \in V$. Its derivative with respect to the group variable (g, v) in the direction $(\xi, w) \in \mathfrak{s}$ is $\mathrm{ad}_{(\xi,w)}\, u = \xi u$.

The Mechanical Connection.   Let $\langle\!\langle \cdot, \cdot \rangle\!\rangle_{\mathfrak{g}}$ and $\langle\!\langle \cdot, \cdot \rangle\!\rangle_V$ be two positive definite inner products on the Lie algebra $\mathfrak{g}$ and on the vector space V, respectively. Then

$$\langle\!\langle (\xi, v), (\eta, w) \rangle\!\rangle_{\mathfrak{s}} = \langle\!\langle \xi, \eta \rangle\!\rangle_{\mathfrak{g}} + \langle\!\langle v, w \rangle\!\rangle_V, \tag{4.3.3}$$

for any $(\xi, v), (\eta, w) \in \mathfrak{s}$, defines a positive definite inner product on $\mathfrak{s}$. Since the spaces $\mathfrak{g} \times \{0\}$ and $\{0\} \times V$ are orthogonal, the orthogonal $\langle\!\langle\, ,\, \rangle\!\rangle_{\mathfrak{s}}$-projection $\mathcal{P}_V : \mathfrak{s} = \mathfrak{g}\,\textcircled{S}\,V \to V$ is simply the projection on the second factor.

Extend the inner product (4.3.3) on $\mathfrak{s}$ to a *right*-invariant Riemannian metric on S by setting

$$\langle\!\langle (X_g, u), (Y_g, w) \rangle\!\rangle_{(g,v)}$$
$$= \langle\!\langle T_{(g,v)} R_{(g,v)^{-1}}(X_g, u), T_{(g,v)} R_{(g,v)^{-1}}(Y_g, w) \rangle\!\rangle_{\mathfrak{s}}, \qquad (4.3.4)$$

where $(g, v) \in S$, $(X_g, u), (Y_g, w) \in T_{(g,v)} S$, and $R_{(g,v)}$ is right translation[2] on S. The derivative of $R_{(h,w)}$ is readily computed from (4.3.1) to be

$$T_{(g,v)} R_{(h,w)}(X_g, u) = (X_g \cdot h, \ u + T_g \phi^w(X_g)), \qquad (4.3.5)$$

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

$$T_{(g,v)} R_{(g,v)^{-1}}(X_g, u) = \left(X_g \cdot g^{-1}, \ u - (X_g \cdot g^{-1})v \right), \qquad (4.3.6)$$

a formula that is useful in the subsequent computations.

The hypotheses of Theorem 2.1.15 hold for the bundle $G \circledS V \to V$ and hence the mechanical connection $\mathcal{A}^V \in \Omega(G \circledS V; V)$ associated to the Riemannian metric $\langle\!\langle \, , \rangle\!\rangle_{\mathfrak{s}}$ is given by formula (2.1.15) which in this case becomes

$$\mathcal{A}^V(g, v)(X_g, u) = \mathrm{Ad}_{(g,v)^{-1}} \left(\mathcal{P}_V T_{(g,v)} R_{(g,v)^{-1}}(X_g, u) \right)$$
$$= \mathrm{Ad}_{(g^{-1}, -g^{-1}v)} \left(\mathcal{P}_V \left(X_g \cdot g^{-1}, \ u - (X \cdot g^{-1})v \right) \right)$$
$$= g^{-1} \left(u - (X_g \cdot g^{-1})v \right), \qquad (4.3.7)$$

where $(g, v) \in S$ and $(X_g, u) \in T_{(g,v)} S$.

Notice that 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 later, invariance in this case will follow from the equivariance equation (2.1.16).

The Flatness Calculation.   The "reason" why the first reduced space is so simple is that the mechanical connection $\mathcal{A}^V$ is flat—that is, its curvature is zero. This is a direct consequence of Theorem 2.1.16 as will be shown below. Let $(X_g, \bar{u}), (Y_g, \bar{w}) \in T_{(g,v)} S$ and let

$$(\xi, u) = T_{(g,v)} R_{(g,v)^{-1}}(X_g, \bar{u})$$
$$(\eta, w) = T_{(g,v)} R_{(g,v)^{-1}}(Y_g, \bar{v}),$$

each of which is an element of $\mathfrak{s}$. We compute the curvature of the mechanical connection $\mathcal{A}^V$ with the assistance of the equation $\mathrm{ad}_{(\xi, w)} u = \xi u$,

---

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

using the formula $(2.1.17)$, which in this case becomes

$$\mathrm{curv}_{\mathcal{A}^V}\left((X_g, \bar{u}), (Y_g, \bar{w})\right)$$
$$= \mathrm{Ad}_{(g,v)^{-1}}\left(-\mathrm{ad}_{(\xi, u)}\,\mathcal{P}_V(\eta, w) + \mathrm{ad}_{(\eta, w)}\,\mathcal{P}_V(\xi, u) + \mathcal{P}_V\left[(\xi, u), (\eta, w)\right] + 0\right)$$
$$= \mathrm{Ad}_{(g,v)^{-1}}\left(-\xi w + \eta u + \xi w - \eta u\right) = 0.$$

We summarize this discussion in the following theorem.

4.3.1 Theorem.   *The mechanical connection $\mathcal{A}^V$ defined on the right principal V-bundle $S \to G$ by formula $(4.3.7)$ 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 $S = G \times V$, so that the connection 1-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 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.

Cotangent Bundle Structure of the Orbits.   We are now ready to establish the extent to which coadjoint orbits of $G \circledS V$ are cotangent bundles (possibly with magnetic terms). We will illustrate the methods with SE(3) in §4.4. As we have mentioned, the strategy is to combine the reduction by stages theorem with the cotangent bundle reduction theorem. In the course of doing this, we recover a result of Ratiu [1980a, 1981, 1982] regarding the embedding of the semidirect product coadjoint orbits into cotangent bundles with magnetic terms, but will provide a different proof here based on connections. We consider here the cotangent lift of right translation of S on T^*S (see Theorem $4.2.1$) and all connections are the mechanical connections associated to the right invariant metrics induced on S and G by the inner products $\langle\!\langle\,,\rangle\!\rangle_{\mathfrak{s}}$ and $\langle\!\langle\,,\rangle\!\rangle_{\mathfrak{g}}$, respectively..

4.3.2 Theorem.   *Let $S = G \circledS V$ and $\mathbf{J}_V : T^*S \to V^*$, $\mathbf{J}_V(\alpha_g, a) = g^{-1}a$, be the momentum map of the cotangent lift of right translation of V on S, where $\langle g^{-1}a, u\rangle := \langle a, gu\rangle$ for any $u \in V$, $a \in V^*$, and $g \in G$. Let $a \in \mathbf{J}_V(T^*S) \subset V^*$ and reduce T^*S at a. There is a right G_a-equivariant symplectic diffeomorphism*

$$(T^*S)_a := \mathbf{J}_V^{-1}(a)/V \simeq (T^*G, \Omega_a)\,, \tag{4.3.8}$$

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 cotangent lift of right translation of

S on itself at σ obtaining the coadjoint orbit $\mathcal{O}_\sigma$ through σ endowed with the plus orbit symplectic form. Let $\mathbf{J}_a : T^*G \to \mathfrak{g}_a^*$ be the momentum map of the cotangent lift of right translation of the isotropy subgroup $G_a = \{g \in G \mid ga = g\}$ on G, $\mathfrak{g}_a = \{\xi \in \mathfrak{g} \mid \xi a = 0\}$ the Lie algebra of G_a, and $\mu_a = \mu|_{\mathfrak{g}_a}$. Then there is a symplectic diffeomorphism

$$\mathcal{O}_\sigma \simeq (T^*G)_{\mu_a} := \mathbf{J}_a^{-1}(\mu_a)/(G_a)_{\mu_a} \tag{4.3.9}$$

and a symplectic embedding

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

where $\Omega_{\mu_a} = \Omega_{\mathrm{can}} - \pi^*\mathcal{B}_{\mu_a}$ with $\mathcal{B}_{\mu_a}$ a closed two-form on $G/(G_a)_{\mu_a}$ calculated in Theorem 4.3.3. The image of this embedding is a vector subbundle of $T^*\left(G/(G_a)_{\mu_a}\right)$. If G_a is Abelian, in which case $(G_a)_{\mu_a} = G_a$, this embedding is a diffeomorphism onto $T^*(G/G_a)$.

Proof. The fact that the spaces in (4.3.8) are symplectomorphic is a consequence of the standard cotangent bundle reduction theorem for Abelian symmetry groups in §2.2 combined with Theorem 4.3.1. As we have seen in §2.2, the symplectomorphism is induced by the shift map (which, recall, is also the projection to the horizontal part):

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

To show the equivariance it only suffices to check that

$$\mathrm{shift}_a\left(p_{(g,v)} \cdot (h,0)\right) = \left(\mathrm{shift}_a(p_{(g,v)})\right) \cdot (h,0), \tag{4.3.10}$$

for any $p_{(g,v)} \in T^*_{(g,v)}S$ and $h \in G_a$. However, if $(X,u) \in T_{(gh,v)}S$, formulas (4.3.7), (4.3.5), and $ha = a$ imply

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

which proves (4.3.10).

The fact that the map in (4.3.9) is a symplectomorphism follows from Theorem 4.2.1 and the G_a-equivariance in (4.3.8).

The rest of the statement is a direct consequence of the Cotangent Bundle Reduction Theorem 2.2.1: the magnetic term of the cotangent bundle $T^*\left(G/(G_a)_{\mu_a}\right)$ is the μ_a-component $\mathcal{B}_{\mu_a} := \langle\mu_a, \mathcal{B}\rangle$ of the curvature $\mathcal{B}$ of the mechanical connection $\mathcal{A}^{G_a}$ on the right principal bundle $G \to G/G_a$ associated to the inner product $\langle\!\langle,\rangle\!\rangle_\mathfrak{g}$ (see Proposition 2.2.5). ∎

Calculation of $\mathcal{A}^{G_a}$ and $\mathbf{d}\mathcal{A}^{G_a}$.   As promised in the preceding theorem, we now derive formulas for the mechanical connection and its curvature on the right principal G_a-bundle $G \to G/G_a$.

4.3.3 Theorem.   *The mechanical connection on the right principal bundle $G \to G/G_a$ associated to the inner product $\langle\!\langle\,,\rangle\!\rangle_{\mathfrak{g}}$ is given by*

$$\mathcal{A}^{G_a}(g)(X_g) = \mathbb{P}_a(T_g L_{g^{-1}} X_g)$$
$$+ \left(\left(\mathbb{P}_a \circ \operatorname{Ad}_g^T \circ \operatorname{Ad}_g \right)\Big|_{\mathfrak{g}_a} \right)^{-1} \left(\mathbb{P}_a \circ \operatorname{Ad}_g^T \circ \operatorname{Ad}_g \right) \left(\mathbb{P}_a^{\perp}(T_g L_{g^{-1}} X_g) \right), \quad (4.3.11)$$

where $\mathbb{P}_a : \mathfrak{g} \to \mathfrak{g}_a$ and $\mathbb{P}_a^{\perp} : \mathfrak{g} \to \mathfrak{g}_a^{\perp}$ are the orthogonal projections relative to the inner product $\langle\!\langle\,,\rangle\!\rangle_{\mathfrak{g}}$. Let $\mathcal{A}_{\mu_a}^{G_a} := \langle \mu_a, \mathcal{A}^{G_a} \rangle \in \Omega^1(G)$ be the μ_a-component of $\mathcal{A}^{G_a}$. The two-form $\mathcal{B}_{\mu_a} \in \Omega^2(G)$ is obtained by dropping $\mathbf{d}\mathcal{A}_{\mu_a}^{G_a}$ to the quotient $G/(G_a)_{\mu_a}$.

If $\operatorname{Ad}_g^T \circ \operatorname{Ad}_g$ leaves $\mathfrak{g}_a$ invariant, where $\operatorname{Ad}_g^T : \mathfrak{g} \to \mathfrak{g}$ is the transpose (adjoint) of Ad_g relative to $\langle\!\langle\,,\rangle\!\rangle_{\mathfrak{g}}$ (this holds, in particular, when $\langle\!\langle\,,\rangle\!\rangle_{\mathfrak{g}}$ is Ad-invariant, which can always be achieved if G is compact), the formulas for the connection and its differential simplify to

$$\mathcal{A}^{G_a} = \mathbb{P}_a \circ \theta^L \tag{4.3.12}$$

and

$$\mathbf{d}\mathcal{A}^{G_a}(g)(X_g, Y_g) = -\mathbb{P}_a \left([T_g L_{g^{-1}} X_g, T_g L_{g^{-1}} Y_g] \right), \tag{4.3.13}$$

where θ^L, defined by $\theta^L(X_g) = T_g L_{g^{-1}} X_g$, is the left-invariant Maurer-Cartan form on G (see Theorem 2.1.14).

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

$$\langle \mathbb{I}(g)(\xi), \eta \rangle = \langle\!\langle \xi_G(g), \eta_G(g) \rangle\!\rangle_g = \langle\!\langle T_e L_g \xi, T_e L_g \eta \rangle\!\rangle_g = \langle\!\langle \operatorname{Ad}_g \xi, \operatorname{Ad}_g \eta \rangle\!\rangle_e$$
$$= \left\langle\!\!\left\langle \operatorname{Ad}_g^T \circ \operatorname{Ad}_g \xi, \eta \right\rangle\!\!\right\rangle_e = \left\langle\!\!\left\langle (\mathbb{P}_a \circ \operatorname{Ad}_g^T \circ \operatorname{Ad}_g)(\xi), \eta \right\rangle\!\!\right\rangle_e.$$

Thus,

$$\mathbb{I}(g)(\xi) = \left\langle\!\!\left\langle (\mathbb{P}_a \circ \operatorname{Ad}_g^T \circ \operatorname{Ad}_g)(\xi), \cdot \right\rangle\!\!\right\rangle_e \in \mathfrak{g}_a^*. \tag{4.3.14}$$

Since the action is free, $\mathbb{I}(g)$ is invertible for every $g \in G$ and hence we conclude that $\left(\mathbb{P}_a \circ \operatorname{Ad}_g^T \circ \operatorname{Ad}_g \right)\Big|_{\mathfrak{g}_a} : \mathfrak{g}_a \to \mathfrak{g}_a$ is an isomorphism.

Next, we compute the value $\mathbf{J}\left(\langle\!\langle X_g, \cdot \rangle\!\rangle_g\right) \in \mathfrak{g}_a^*$ of the G_a-momentum map $\mathbf{J} : T^*G \to \mathfrak{g}_a^*$. For $\xi \in \mathfrak{g}_a$ we have

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

We conclude that

$$
\begin{aligned}
\mathbf{J}\left(\langle\!\langle X_g, \cdot \rangle\!\rangle_g\right) &= \langle\!\langle (\mathbb{P}_a \circ \mathrm{Ad}_g^T \circ \mathrm{Ad}_g)(T_g L_{g^{-1}} X_g), \cdot \rangle\!\rangle_e \\
&= \langle\!\langle (\mathbb{P}_a \circ \mathrm{Ad}_g^T \circ \mathrm{Ad}_g)(\mathbb{P}_a T_g L_{g^{-1}} X_g), \cdot \rangle\!\rangle_e \\
&\quad + \langle\!\langle (\mathbb{P}_a \circ \mathrm{Ad}_g^T \circ \mathrm{Ad}_g)(\mathbb{P}_a^\perp T_g L_{g^{-1}} X_g), \cdot \rangle\!\rangle_e . \qquad (4.3.15)
\end{aligned}
$$

Using (4.3.14) and (4.3.15) in the definition (2.1.4) of he mechanical connection yields (4.3.11).

Now assume that $\mathrm{Ad}_g^T \circ \mathrm{Ad}_g$ leaves $\mathfrak{g}_a$ invariant; since this linear operator is symmetric, it also leaves its orthogonal complement invariant. This implies that the second summand in (4.3.15) vanishes and hence

$$
\mathbf{J}\left(\langle\!\langle X_g, \cdot \rangle\!\rangle_g\right) = \langle\!\langle \left(\mathbb{P}_a \circ \mathrm{Ad}_g^T \circ \mathrm{Ad}_g\right) \left(\mathbb{P}_a(g^{-1} \cdot X_g)\right), \cdot \rangle\!\rangle_e . \qquad (4.3.16)
$$

Combining (4.3.14) and (4.3.16), we get

$$
\mathcal{A}^{G_a}(g)(X_g) = \left(\mathbb{I}(g)^{-1} \circ \mathbf{J}\right)\left(\langle\!\langle X_g, \cdot \rangle\!\rangle_g\right) = (\mathbb{P}_a \circ \theta^L)(X_g).
$$

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,

$$
\begin{aligned}
\mathbf{d}\mathcal{A}^{G_a}(g)(X_g, Y_g) &= \bar{X}[\mathcal{A}^{G_a}(\bar{Y})](g) - \bar{Y}[\mathcal{A}^{G_a}(\bar{X})](g) - \mathcal{A}^{G_a}(g)([\bar{X}, \bar{Y}](g)) \\
&= -\mathbb{P}_a([T_g L_g^{-1} \cdot X_g, T_g L_g^{-1} \cdot Y_g]), \qquad (4.3.17)
\end{aligned}
$$

where we have used the fact that $\mathcal{A}^{G_a}(\bar{Y})$, for example, is constant from the preceding expression for $\mathcal{A}^{G_a}$ and left invariance, and so the first two terms vanish. ∎

4.4   Example: The Euclidean Group

This section uses the results of the preceding section to classify the coadjoint orbits of the Euclidean group SE(3). We will also make use of mechanical connections and their curvatures to compute the the coadjoint orbit symplectic forms.

A Right Invariant Metric on $\mathrm{SE}(3)$**.**   Identify

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

and define the natural inner product at the identity (see (4.3.3))

$$\langle\!\langle (X, a), (Y, b) \rangle\!\rangle_{(I,0)} = -\frac{1}{2}\mathrm{tr}(XY) + \langle\!\langle a, b \rangle\!\rangle,$$

where, on the right hand side, $\langle\!\langle \cdot, \cdot \rangle\!\rangle$ denotes the Euclidean inner product. Requiring right invariance of the metric and use of the equation (4.3.6) gives

$$
\begin{aligned}
&\langle\!\langle (X_A, a_A), (Y_A, b_A) \rangle\!\rangle_{(A,\alpha)} \\
&= \langle\!\langle (X_A \cdot A^{-1}, a_A - (X_A \cdot A^{-1})\alpha), (Y_A \cdot A^{-1}, b_A - (Y_A \cdot A^{-1})\alpha) \rangle\!\rangle_{(I,0)} \\
&= -\frac{1}{2}\mathrm{tr}(X_A \cdot A^{-1} \cdot Y_A \cdot A^{-1}) + \langle\!\langle (X_A \cdot A^{-1})\alpha, (Y_A \cdot A^{-1})\alpha \rangle\!\rangle \\
&\quad - \langle\!\langle (X_A \cdot A^{-1})\alpha, b_A \rangle\!\rangle - \langle\!\langle (Y_A \cdot A^{-1})\alpha, a_A \rangle\!\rangle + \langle\!\langle a_A, b_A \rangle\!\rangle. \qquad (4.4.1)
\end{aligned}
$$

The mechanical connection for the principal $\mathbb{R}^3$-bundle $\mathrm{SE}(3) \to \mathrm{SO}(3)$, is given by (4.3.7):

$$\mathcal{A}^{\mathbb{R}^3}(A, \alpha)(X_A, a_A) = A^{-1}\left(a_A - (X_A \cdot A^{-1})\alpha\right)$$

and from Theorem 4.3.1, we see that $\mathrm{curv}\,\mathcal{A}^{\mathbb{R}^3} = \mathbf{d}\mathcal{A}^{\mathbb{R}^3} = 0$.

First Reduction.   We first reduce by the $\mathbb{R}^3$-cotangent lifted action. Let $a \in \mathbb{R}^{3*} = \mathbb{R}^3$. By Theorem 4.3.2, the cotangent bundle reduction theorem for semidirect products, we know that the first reduced space $(T^*\,\mathrm{SE}(3))_a = \mathbf{J}_{\mathbb{R}^3}^{-1}(a)/\mathbb{R}^3$ is symplectically diffeomorphic to the cotangent bundle $(T^*\,\mathrm{SO}(3), \Omega_{\mathrm{can}})$.

Second Reduction.   We first take the easy case in which $a = 0$. Then $G_a = \mathrm{SO}(3)$. Reduction by the $\mathrm{SO}(3)$-action therefore gives coadjoint orbits of $\mathrm{SO}(3)$. Thus $\mathcal{O}_{(a=0,\mu)} = S_\mu^2$, the two-sphere passing through $\mu \in \mathbb{R}^3$.

Next, assume $a \neq 0$. Then the group $\mathrm{SE}(3)_a/\mathbb{R}_a^3 \simeq \mathrm{SO}(3)_a \simeq S^1$ acts (on the *right*) on the first reduced space, $(T^*\,\mathrm{SO}(3), \Omega_{\mathrm{can}})$. Note that the map $[A] \in \mathrm{SO}(3)/\mathrm{SO}(3)_a \mapsto Aa \in S_a^2$, the two-sphere passing through $a \in \mathbb{R}^3$, is a diffeomorphism. Depending on whether $\mu = 0$ or $\mu \neq 0$, we have to consider two further subcases.

Suppose that $\mu = 0$. Reducing by the $\mathrm{SO}(3)_a = S^1$-action at $\mu_a = 0$ gives, by another application of the cotangent bundle reduction theorem for Abelian groups, the symplectic manifold $(T^*S_a^2, \Omega_{\mathrm{can}})$ (see also Theorem 4.3.2).

Finally, consider the subcase $\mu \neq 0$. The group $G_a = \mathrm{SO}(3)_a = S^1$ acts by cotangent lift of right translation on $T^*\,\mathrm{SO}(3)$. The S^1-principal

bundle $\mathrm{SO}(3) \to \mathrm{SO}(3)/\mathrm{SO}(3)_a \simeq S_a^2$ naturally inherits a metric from the principal $\mathbb{R}^3$-bundle $\mathrm{SE}(3) \to \mathrm{SO}(3)$, which is $\mathrm{SO}(3)$-invariant.

Let us compute the curvature of the mechanical connection on the bundle $\mathrm{SO}(3) \to S_a^2$. It is convenient to use the Lie algebra isomorphism $x \in \mathbb{R}^3 \mapsto \widehat{x} \in \mathfrak{so}(3)$ defined by the cross product, namely, $\widehat{x}u = x \times u$. The inner product on $\mathfrak{so}(3)$

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

where $x, y \in \mathbb{R}^3$, induces the right invariant Riemannian metric $\langle\!\langle \cdot, \cdot \rangle\!\rangle$ on $\mathrm{SO}(3)$ given on the tangent space at A by

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

where $X_A, Y_A \in T_A\,\mathrm{SO}(3)$. The Lie algebra of $\mathrm{SO}(3)_a \simeq S^1$ is $\mathrm{span}\{a\} \cong \mathbb{R}$, so the infinitesimal generator of $u \in \mathbb{R}$ is

$$u_{\mathrm{SO}(3)}(A) = \left.\frac{d}{dt}\right|_{t=0} A\exp(tu\widehat{a}) = uA\widehat{a}. \tag{4.4.2}$$

By right invariance of the metric and the identity $A\widehat{a}A^{-1} = \widehat{Aa}$, we get for any $u, v \in \mathbb{R}$,

$$\langle\!\langle \mathbb{I}(A)u\widehat{a}, v\widehat{a} \rangle\!\rangle = uv\,\langle\!\langle A\widehat{a}, A\widehat{a} \rangle\!\rangle_A = uv\,\langle\!\langle A\widehat{a}A^{-1}, A\widehat{a}A^{-1} \rangle\!\rangle_I$$

$$= uv\,\langle\!\langle \widehat{Aa}, \widehat{Aa} \rangle\!\rangle_I = uv\,\langle\!\langle Aa, Aa \rangle\!\rangle = \langle\!\langle ua, va \rangle\!\rangle. \tag{4.4.3}$$

To identify from this formula the locked inertia tensor $\mathbb{I}(A) : \mathrm{span}\{\widehat{a}\} \to \mathrm{span}\{a\}^*$ as a linear map from $\mathrm{span}\{\widehat{a}\}$ to $\mathbb{R}^3$ and to determine its one-dimensional range, we will make use of the isomorphism $\bar{\mu} \in \mathfrak{so}(3)^* \mapsto \mu \in \mathbb{R}^3$ given by $\langle \bar{\mu}, \widehat{x} \rangle = \langle\!\langle \mu, x \rangle\!\rangle$ for any $x \in \mathbb{R}^3$.

The projection $\mathbb{R}^3 \to \mathrm{span}\{a\}$ is given by $x \mapsto \frac{\langle\!\langle x, a \rangle\!\rangle}{\|a\|^2}a$ and composing it with the isomorphism $\widehat{x} \in \mathfrak{so}(3) \mapsto x \in \mathbb{R}^3$ gives the projection $\widehat{x} \in \mathfrak{so}(3) \mapsto \frac{\langle\!\langle x, a \rangle\!\rangle}{\|a\|^2}a \in \mathrm{span}\{a\}$. The dual $\mathrm{span}\{a\}^* \to \mathfrak{so}(3)^*$ of this map composed with the isomorphism $\bar{\mu} \in \mathfrak{so}(3)^* \mapsto \mu \in \mathbb{R}^3$ gives the embedding $\kappa : \varphi \in \mathrm{span}\{a\}^* \mapsto \frac{\langle \varphi, \widehat{a} \rangle}{\|a\|^2}a \in \mathrm{span}\{a\} \subset \mathbb{R}^3$. This isomorphism κ which identifies $\mathrm{span}\{a\}^*$ with $\mathrm{span}\{a\}$ is thus characterized by

$$\langle\!\langle \kappa(\varphi), a \rangle\!\rangle = \langle \varphi, \widehat{a} \rangle. \tag{4.4.4}$$

Thus, by (4.4.3), we get

$$\langle\!\langle \kappa\left(\mathbb{I}(A)u\widehat{a}\right), a \rangle\!\rangle = \langle\!\langle \mathbb{I}(A)u\widehat{a}, \widehat{a} \rangle\!\rangle = \langle\!\langle ua, a \rangle\!\rangle.$$

Therefore, identifying via κ the spaces $\mathrm{span}\{a\}^*$ and $\mathrm{span}\{a\}$, formula (4.4.3) shows that $\mathbb{I}(A) : \mathrm{span}\{\widehat{a}\} \to \mathrm{span}\{a\}$ is given by

$$\mathbb{I}(A)\widehat{a} = a. \tag{4.4.5}$$

Taking $u \in \mathbb{R}$, the $\mathrm{SO}(3)_a$-momentum map $\mathbf{J} : T^* \mathrm{SO}(3) \to \mathrm{span}\{a\}^*$ is given by

$$
\begin{aligned}
\langle \mathbf{J} \left(\langle\!\langle X_A, \cdot \rangle\!\rangle_A \right), u\widehat{a} \rangle &= \langle\!\langle X_A, u_{\mathrm{SO}(3)}(A) \rangle\!\rangle_A = u \langle\!\langle X_A, A\widehat{a} \rangle\!\rangle_A \\
&= u \langle\!\langle X_A \cdot A^{-1}, A\widehat{a}A^{-1} \rangle\!\rangle_I \\
&= u \langle\!\langle \mathrm{Ad}_{A^{-1}}(X_A \cdot A^{-1}), \widehat{a} \rangle\!\rangle_I \\
&= u \langle\!\langle A^{-1} \cdot X_A, \widehat{a} \rangle\!\rangle_I \\
&= \left(\frac{1}{\|a\|^2} \langle\!\langle A^{-1} \cdot X_A, \widehat{a} \rangle\!\rangle_I \, a, ua \right),
\end{aligned}
$$

so that, identifying $\mathrm{span}\{a\}^*$ with $\mathrm{span}\{a\}$ via the map κ, equation (4.4.4) gives

$$
\mathbf{J} \left(\langle\!\langle X_A, \cdot \rangle\!\rangle_A \right) = \frac{1}{\|a\|^2} \langle\!\langle A^{-1} \cdot X_A, \widehat{a} \rangle\!\rangle_I \, a \in \mathrm{span}\{a\}. \tag{4.4.6}
$$

Therefore, by (4.4.5) and (4.4.6), the mechanical $\mathrm{span}\{\widehat{a}\}$-valued connection one-form has the expression

$$
\begin{aligned}
\mathcal{A}^{\mathrm{SO}(3)_a}(A)(X_A) :&= \left(\mathbb{I}(A)^{-1} \circ \mathbf{J} \right) \left(\langle\!\langle X_A, \cdot \rangle\!\rangle_A \right) \\
&= \frac{1}{\|a\|^2} \langle\!\langle A^{-1} \cdot X_A, \widehat{a} \rangle\!\rangle_I \, \widehat{a}. \tag{4.4.7}
\end{aligned}
$$

If $\mu \in \mathbb{R}^3$, then $\bar{\mu}|_{\mathrm{span}\{a\}} \in \mathrm{span}\{a\}^*$ and hence $\kappa \left(\bar{\mu}|_{\mathrm{span}\{a\}} \right) = \frac{\langle\!\langle \mu, a \rangle\!\rangle}{\|a\|^2} a \in \mathrm{span}\{a\}$, which says that if we identify $\mathrm{span}\{a\}^*$ with $\mathrm{span}\{a\}$ via κ then $\mu_a = \frac{\langle\!\langle \mu, a \rangle\!\rangle}{\|a\|^2} a \in \mathrm{span}\{a\}$.

From (4.4.7) we see that if $\mu \in \mathbb{R}^3$, the μ_a-component $\mathcal{A}_{\mu_a}^{\mathrm{SO}(3)_a}$ of $\mathcal{A}^{\mathrm{SO}(3)_a}$ is given by

$$
\begin{aligned}
\mathcal{A}_{\mu_a}^{\mathrm{SO}(3)_a}(A)(X_A) &= \left\langle \bar{\mu}|_{\mathrm{span}\{\widehat{a}\}}, \mathcal{A}^{\mathrm{SO}(3)_a}(A)(X_A) \right\rangle \\
&= \left\langle \bar{\mu}, \frac{1}{\|a\|^2} \langle\!\langle A^{-1} \cdot X_A, \widehat{a} \rangle\!\rangle_I \, \widehat{a} \right\rangle \\
&= \frac{\langle\!\langle \mu, a \rangle\!\rangle}{\|a\|^2} \langle\!\langle A^{-1} \cdot X_A, \widehat{a} \rangle\!\rangle_I. \tag{4.4.8}
\end{aligned}
$$

To find the magnetic term we need to compute $\mathbf{d}\mathcal{A}_{\mu_a}^{\mathrm{SO}(3)_a}(A)(X_A, Y_A)$ for $A \in \mathrm{SO}(3)$ and $X_A, Y_A \in T_A \mathrm{SO}(3)$. 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 (4.4.7) and the identities $A^{-1}\widehat{x}A = \widehat{A^{-1}x}$, $\langle\!\langle \widehat{x}, \widehat{y} \rangle\!\rangle_I =$

$\langle\!\langle x, y \rangle\!\rangle$, imply

$$
\begin{aligned}
Y_A \left[\mathcal{A}^{\mathrm{SO}(3)_a}_{\mu_a}(\bar{X}) \right] &= \left.\frac{d}{dt}\right|_{t=0} \mathcal{A}^{\mathrm{SO}(3)_a}_{\mu_a}(\bar{X})((\exp t\widehat{y})A) \\
&= \left.\frac{d}{dt}\right|_{t=0} \mathcal{A}^{\mathrm{SO}(3)_a}_{\mu_a}((\exp t\widehat{y})A)(\widehat{x} \cdot (\exp t\widehat{y})A) \\
&= \left.\frac{d}{dt}\right|_{t=0} \frac{\langle\!\langle \mu, a \rangle\!\rangle}{\|a\|^2} \left\langle\!\left\langle A^{-1}\exp(-t\widehat{y})\widehat{x}(\exp t\widehat{y})A,\, \widehat{a} \right\rangle\!\right\rangle_I \\
&= \left.\frac{d}{dt}\right|_{t=0} \frac{\langle\!\langle \mu, a \rangle\!\rangle}{\|a\|^2} \left\langle\!\left\langle A^{-1}\exp(-t\widehat{y})x, a \right\rangle\!\right\rangle \\
&= \frac{\langle\!\langle \mu, a \rangle\!\rangle}{\|a\|^2} \left\langle\!\left\langle A^{-1}(x \times y), a \right\rangle\!\right\rangle.
\end{aligned} \tag{4.4.9}
$$

Similarly

$$
X_A \left[\mathcal{A}^{\mathrm{SO}(3)_a}_{\mu_a}(\bar{Y}) \right] = -\frac{\langle\!\langle \mu, a \rangle\!\rangle}{\|a\|^2} \left\langle\!\left\langle A^{-1}(x \times y), a \right\rangle\!\right\rangle. \tag{4.4.10}
$$

Finally, since $[\bar{X}, \bar{Y}])(A) = -[\widehat{x}, \widehat{y}] \cdot A$ (because $\bar{X}, \bar{Y}$ are *right* invariant vector fields), formula (4.4.7) yields

$$
\begin{aligned}
\mathcal{A}^{\mathrm{SO}(3)_a}_{\mu_a}\left([\bar{X}, \bar{Y}]\right)(A) &= \mathcal{A}^{\mathrm{SO}(3)_a}_{\mu_a}(A)(-[\widehat{x}, \widehat{y}] \cdot A) \\
&= -\frac{\langle\!\langle \mu, a \rangle\!\rangle}{\|a\|^2} \left\langle\!\left\langle A^{-1}[\widehat{x}, \widehat{y}]A, \widehat{a} \right\rangle\!\right\rangle_I \\
&= -\frac{\langle\!\langle \mu, a \rangle\!\rangle}{\|a\|^2} \left\langle\!\left\langle A^{-1}(x \times y), a \right\rangle\!\right\rangle.
\end{aligned} \tag{4.4.11}
$$

Formulas (4.4.9), (4.4.10), and (4.4.11) therefore give

$$
\begin{aligned}
\mathbf{d}&\mathcal{A}^{\mathrm{SO}(3)_a}_{\mu_a}(A)(X_A, Y_A) \\
&= X_A \left[\mathcal{A}^{\mathrm{SO}(3)_a}_{\mu_a}(\bar{Y}) \right] - Y_A \left[\mathcal{A}^{\mathrm{SO}(3)_a}_{\mu_a}(\bar{X}) \right] - \mathcal{A}^{\mathrm{SO}(3)_a}_{\mu_a}\left([\bar{X}, \bar{Y}]\right)(A) \\
&= -\frac{\langle\!\langle \mu, a \rangle\!\rangle}{\|a\|^2} \left\langle\!\left\langle A^{-1}(x \times y), a \right\rangle\!\right\rangle,
\end{aligned} \tag{4.4.12}
$$

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 4.3.3.

This two-form on $\mathrm{SO}(3)$ clearly induces a two-form $\mathcal{B}_{\mu_a}$, the magnetic term, on the sphere S^2_a by

$$
\mathcal{B}_{\mu_a}(Aa)(x \times Aa, y \times Aa) = -\frac{\langle\!\langle \mu, a \rangle\!\rangle}{\|a\|^2} \langle\!\langle x \times y, Aa \rangle\!\rangle. \tag{4.4.13}
$$

Invoking the cotangent bundle reduction theorem we classify the orbits of SE(3) as follows.

4.4.1 Theorem.   *The coadjoint orbits of* $\mathrm{SE}(3)$ *are of the following types.*

- $\mathcal{O}_{(a=0,\mu)} \simeq (S^2_\mu, \omega_\mu)$

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

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

where ω_μ is the orbit symplectic form on the sphere S^2_μ of radius $\|\mu\|$, $\mu_a = \frac{\langle\!\langle\mu,a\rangle\!\rangle}{\|a\|^2}a \in \mathrm{span}\{a\}$ 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 (4.4.13).

5
Regular Reduction by Stages

In this chapter we formulate the first of several reduction by stages theorems in the regular, that is, free actions, case. We state a sufficient condition, called the *stages hypothesis* under which the two step reduced space is symplectically diffeomorphic to the space obtained by reducing all at once by the original group. In Chapters 11 and 12 we will come up with alternative conditions for reduction by stages based on the use of distribution theory.

Some results in the general direction of reduction by stages for group extensions are those of Landsman [1998], Sjamaar and Lerman [1991], and Ziegler [1996]. The results in Landsman [1998] also make 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 even cover the important 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   Motivating Example: The Heisenberg Group

To motivate some of the issues beyond those encountered in the previous chapter that will come up in the general theory of reduction by stages in the following sections (as well as the theory for group extensions in the following Chapter), we first consider one of the basic examples, namely the Heisenberg group[1] denoted H. We will be interested in reducing T^*H by the action of H (thereby producing the coadjoint orbits of H) in two stages, first by the center of H (which is therefore a normal subgroup) and then by the quotient of H by its center. 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.   We start with the commutative group $\mathbb{R}^2$ with its standard symplectic form ω, the usual area form on the plane, that is,

$$\omega(u, v) = u_x v_y - u_y v_x,$$

where $u = (u_x, u_y), v = (v_x, v_y) \in \mathbb{R}^2$. Form the set $H = \mathbb{R}^2 \oplus \mathbb{R}$ with multiplication

$$(u, \alpha)(v, \beta) = \left(u + v, \alpha + \beta + \frac{1}{2}\omega(u, v) \right). \tag{5.1.1}$$

It is readily verified that this operation defines a Lie group, the ***Heisenberg group***. Note that the identity element is $(0, 0)$ and the inverse of (u, α) is given by $(u, \alpha)^{-1} = (-u, -\alpha)$.

In Chapter 6 we will recall the general theory of group extensions and in the terminology introduced there, one says that $\mathbb{R}^2$ has been ***centrally extended*** by $\mathbb{R}$ using $\omega/2$ as a ***group two-cocycle***. The subgroup $\mathbb{R}$, consisting of pairs $(0, \alpha)$ is the ***center*** of H in the sense that each $(0, \alpha)$ commutes with every other element of H. Also, every element of H that commutes with all other elements of H is of the form $(0, \alpha)$, as is seen using nondegeneracy of ω.

The Heisenberg group is isomorphic to the upper triangular 3×3 matrices with ones on the diagonal. The isomorphism is given by

$$(u, \alpha) \mapsto \begin{bmatrix} 1 & u_x & \alpha + \frac{1}{2}u_x u_y \\ 0 & 1 & u_y \\ 0 & 0 & 1 \end{bmatrix}$$

Conjugation in H is given by

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

---

[1] Another exposition of the Heisenberg group may be found in, for example, Guillemin and Sternberg [1984]

Differentiating with respect to (v, β), we see that the Ad operator is

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

and, differentiating once more, we get the ad operator; that is, the Lie bracket:

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

where $(X, a), (Y, b) \in \mathfrak{h} = \mathbb{R}^2 \oplus \mathbb{R}$.

Consistent with the general terminology to be reviewed in Chapter 6, we can read off from this bracket relation the **_Lie algebra cocycle_** C as the bilinear form

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

The exponential map $\exp : \mathfrak{h} \to H$ is the identity map of $\mathbb{R}^3$. Note that this shows that the exponential map of the Heisenberg group coincides with that of the vector Lie group $(\mathbb{R}^3, +)$.

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

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

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

- $\mathcal{O}_{(\mu,0)} = \{(\mu, 0)\}$

- $\mathcal{O}_{(\mu,\nu\neq0)} = \mathbb{R}^2 \times \{\nu\}$.

The Mechanical Connection.   Next, consider H as a 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

$$\langle (X, a), (Y, b) \rangle = (X, Y) + ab, \tag{5.1.2}$$

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 and multiplication of real numbers in the second. If $(X_{(u,\alpha)}, a_{(u,\alpha)}) \in T_{(u,\alpha)}H$, then the tangent of right multiplication is given by

$$T_{(u,\alpha)}R_{(v,\beta)}(X_{(u,\alpha)}, a_{(u,\alpha)}) = \left(X_{(u,\alpha)}, a_{(u,\alpha)} + \frac{1}{2}\omega(X_{(u,\alpha)}, v) \right)$$

$$\in T_{(u,\alpha)(v,\beta)}H$$

and, in particular,

$$T_{(u,\alpha)}R_{(u,\alpha)^{-1}}(X_{(u,\alpha)}, a_{(u,\alpha)}) = \left(X_{(u,\alpha)}, a_{(u,\alpha)} - \frac{1}{2}\omega(X_{(u,\alpha)}, u) \right) \in \mathfrak{h}.$$

Thus, the associated right invariant metric on H is given by

$$\left\langle\!\!\left\langle (X_{(u,\alpha)}, a_{(u,\alpha)}), (Y_{(u,\alpha)}, b_{(u,\alpha)}) \right\rangle\!\!\right\rangle_{(u,\alpha)} = (X_{(u,\alpha)}, Y_{(u,\alpha)}) + a_{(u,\alpha)} b_{(u,\alpha)}$$

$$- \frac{1}{2} a_{(u,\alpha)} \omega(Y_{(u,\alpha)}, u) - \frac{1}{2} b_{(u,\alpha)} \omega(X_{(u,\alpha)}, u) + \frac{1}{4} \omega(X_{(u,\alpha)}, u) \omega(Y_{(u,\alpha)}, u).$$

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

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

Combining these formulas and using the general formula (2.1.2) for the locked inertia tensor yields the expression of the associated locked inertia tensor in this case:

$$\mathbb{I}_{(v,\alpha)}(a)(b) = \left\langle\!\!\left\langle a_H(v, \alpha), b_H(v, \alpha) \right\rangle\!\!\right\rangle_{(v,\alpha)} = ab \qquad (5.1.3)$$

For any $(X_{(u,\alpha)}, a_{(u,\alpha)}) \in T_{(u,\alpha)} H$ and $b \in \mathbb{R}$, we have the momentum map for the $\mathbb{R}$ action given as follows:

$$\left\langle \mathbf{J}_{\mathbb{R}} \left(\left\langle\!\!\left\langle (X_{(u,\alpha)}, a_{(u,\alpha)}), \cdot \right\rangle\!\!\right\rangle_{(u,\alpha)} \right), b \right\rangle = \left\langle\!\!\left\langle (X_{(u,\alpha)}, a_{(u,\alpha)}), (0, b) \right\rangle\!\!\right\rangle_{(u,\alpha)}$$

$$= \left(a_{(u,\alpha)} - \frac{1}{2} \omega(X_{(u,\alpha)}, u) \right) b.$$

$$(5.1.4)$$

Thus, according to formula (2.1.4), the mechanical connection has the expression

$$\mathcal{A}(u, \alpha)(X_{(u,\alpha)}, a_{(u,\alpha)}) = a_{(u,\alpha)} - \frac{1}{2} \omega(X_{(u,\alpha)}, u). \qquad (5.1.5)$$

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

$$\mathbf{d}\mathcal{A}(u, \alpha)((X_{(u,\alpha)}, a_{(u,\alpha)}), (Y_{(u,\alpha)}, b_{(u,\alpha)})) = \omega(X_{(u,\alpha)}, Y_{(u,\alpha)}). \qquad (5.1.6)$$

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

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

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, according to cotangent bundle reduction by an Abelian group (see Theorem 2.2.1), the space

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

where $\simeq$ is a symplectic diffeomorphism that is readily seen to be equivariant with respect to the remaining $\mathbb{R}^2$-action, $\pi : T^*\mathbb{R}^2 \to \mathbb{R}^2$ denotes the cotangent bundle projection, and $\mathcal{B}_\nu = \nu\mathcal{B}$, that is,

$$\mathcal{B}_\nu(X,Y) = \nu\omega(X,Y), \tag{5.1.7}$$

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

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 nonequivariant 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 $\left(T^*Q, \Omega_{\mathrm{can}} - \pi_Q^*\mathcal{B}\right)$, where $\mathcal{B}$ is a closed two-form on Q, suppose that there is a linear map $\xi \mapsto \phi^\xi = \langle\phi, \xi\rangle$ from $\mathfrak{g}$ to functions on Q (so that $\phi : Q \to \mathfrak{g}^*$) such that for all $\xi \in \mathfrak{g}$, we have $\mathbf{i}_{\xi_Q}\mathcal{B} = \mathbf{d}\phi^\xi$ (where $\mathbf{i}_\xi$ denotes the interior product). Later on, in §7.1 these maps will be generalized and will be called $\mathcal{B}\mathfrak{g}$-*potentials* (see Theorem 7.1.1).

In these circumstances, a direct verification shows that the momentum map is given by $\mathbf{J} = \mathbf{J}_{\mathrm{can}} - \phi \circ \pi_Q$, where $\pi_Q : T^*Q \to Q$ is cotangent bundle projection.

In the case of the Heisenberg group, where $\mathcal{B}$ is $\mathcal{B}_\nu$ given by (5.1.7), which is a symplectic form, the ϕ is in fact a momentum map. Namely, in this case, it is the momentum map for the translation action of $\mathbb{R}^2$ on $(\mathbb{R}^2, \nu\omega)$, which, in view of the ν dependence of $\mathcal{B}_\nu$, we shall denote as ϕ^ν; it is given by

$$\phi^\nu(x,y) = \nu(y,-x) \tag{5.1.8}$$

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

$$\mathbf{J}_{\mathrm{can}}(x, y, p_x, p_y) = (p_x, p_y) \tag{5.1.9}$$

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

$$\mathbf{J}_\nu(x, y, p_x, p_y) = (p_x - \nu y, p_y + \nu x). \tag{5.1.10}$$

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

$$\mathbf{J}_\nu^{-1}(\rho) = \{(x, y, p_x, p_y) \mid p_x = \nu y + \rho_1, p_y = -\nu x + \rho_2\}. \tag{5.1.11}$$

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 [1970] 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., [MandS]). This affine G action on $\mathfrak{g}^*$ with respect to which the momentum map becomes equivariant is given by the general formula

$$g \cdot \mu = \mathrm{Ad}^*_{g^{-1}} \mu + \sigma(g) \tag{5.1.12}$$

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

5.1.1 Lemma.   *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 equation (1.1.6) or [MandS], §12.4) and using the fact that the coadjoint action of $\mathbb{R}^2$ is trivial (since $\mathbb{R}^2$ is Abelian), we get

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

that is,

$$\sigma^\nu(a,b) = \nu(-b, a). \tag{5.1.13}$$

Therefore, the affine action is

$$(a,b) \cdot (\rho_1, \rho_2) = (\rho_1, \rho_2) + \nu(-b, a) \tag{5.1.14}$$

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

The Second Reduced Space.   First consider the case $\nu = 0$. In this case, the first reduced space is $\big(T^*\mathbb{R}^2, \Omega_{\mathrm{can}}\big)$. 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 $\mathbf{J}_\nu$.

5.1.2 Proposition.   *The coadjoint orbit symplectic form for the orbit through the point $((\rho_1, \rho_2), \nu)$ is given by*

$$\omega_{\mathcal{O}_{(\rho,\nu)}}(x,y)(X,Y) = \nu\omega(X,Y). \tag{5.1.15}$$

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

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

We have, according to the definition of the reduced symplectic form,

$$
\begin{aligned}
&\omega_{\mathrm{red}}(x,y)(X,Y) \\
&\quad = \Omega_{\mathrm{can}}(\psi(x,y))\left(X_1 \frac{\partial}{\partial x} + X_2 \frac{\partial}{\partial y} + \nu X_2 \frac{\partial}{\partial p_x} - \nu X_1 \frac{\partial}{\partial p_y}, \right. \\
&\qquad \left. Y_1 \frac{\partial}{\partial x} + Y_2 \frac{\partial}{\partial y} + \nu Y_2 \frac{\partial}{\partial p_x} - \nu Y_1 \frac{\partial}{\partial p_y} \right) - \psi^*\pi^*\mathcal{B}_\nu(x,y)(X,Y) \\
&\quad = X_1\nu Y_2 - Y_1\nu X_2 + X_2(-\nu Y_1) - Y_2(-\nu X_1) \\
&\qquad - \psi^*\pi^*\mathcal{B}_\nu(x,y)(X,Y) \\
&\quad = 2\nu\omega(x,y)(X,Y) - \nu\omega(x,y)(X,Y) \\
&\quad = \nu\omega(x,y)(X,Y), \tag{5.1.16}
\end{aligned}
$$

where we have used the fact that $\pi \circ \psi = \mathrm{id}$. Thus, equation (5.1.15) holds. ∎

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 two sections.

5.2   Point Reduction by Stages

In this section, we introduce the *stages hypothesis* as a sufficient condition for a general reduction by stages theorem using *point reduction*. We will relate this result to the Poisson reduction and orbit reduction viewpoints in §5.3.

Interestingly, the stages hypothesis depends only on properties of the Lie group itself and not on the data contained in the action. We prove in §6.3 that the most general class of group extensions, roughly given by semidirect products with a cocycle, satisfies this sufficient condition. In particular, this will show that both central extensions and semidirect products with a vector space fit into this class.

There will be two other reduction by stages theorems, given in Chapter 11, where variants of the stages hypothesis are studied in detail. In particular, it is shown that if one keeps careful track of connected components,

there is a sense in which the stages hypothesis is not needed at all, so it is, perhaps, not surprising that it holds in the examples we give. In addition, the singular case is given in Part III.

The Setup.   As explained in the introduction, we start with a symplectic manifold (P, Ω) and a Lie group M that acts on P freely and properly 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 $\Phi_m : P \to P$.

The Full Reduced Space.   Assume that $\sigma \in \mathfrak{m}^*$ is a given value of $\mathbf{J}_M$, which is a regular value by the freeness assumption on the action. Then we can form the **full reduced space**

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

where $M_\sigma = \{m \in M \mid m\sigma = \sigma\}$ is the isotropy subgroup of σ for the action of M on $\mathfrak{m}^*$. It is this full reduction that we will break into two stages.

The Normal Subgroup N.   Assume that N is a closed 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. Since N is a normal subgroup, M acts on N by conjugation and therefore also on $\mathfrak{n}$ by the derivative of the conjugation at the identity element. This representation will be denoted by $(m, \xi) \in M \times \mathfrak{n} \mapsto m \cdot \xi$. Dualizing it, one obtains a representation of M on $\mathfrak{n}^*$ that satisfies

$$m \cdot \nu := m \cdot i^*(\sigma) = i^*(\mathrm{Ad}^*_{m^{-1}} \sigma), \tag{5.2.1}$$

where $\sigma \in \mathfrak{m}^*$ satisfies $i^*(\sigma) = \nu$.

The momentum map for the action of the group N on P is given by

$$\mathbf{J}_N(z) = i^*(\mathbf{J}_M(z)). \tag{5.2.2}$$

It is not only N equivariant, as one would expect, but also M equivariant with respect to the actions of M on P and on $\mathfrak{n}^*$. To see this, notice that

$$\mathbf{J}_N(m \cdot z) = i^*(\mathbf{J}_M(m \cdot z)) = i^*(\mathrm{Ad}^*_{m^{-1}}(\mathbf{J}_M(z)))$$
$$= m \cdot i^*(\mathbf{J}_M(z)) = m \cdot \mathbf{J}_N(z).$$

The First Reduced Space.   Let $\nu \in \mathfrak{n}^*$ be a value of $\mathbf{J}_N$ (which by the freeness of the action will be a regular value) and let N_ν be the isotropy subgroup of ν for the coadjoint action of N on the dual of its Lie algebra. We form the **first symplectic reduced space**

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

In what follows we let $\pi_\nu : \mathbf{J}_N^{-1}(\nu) \to P_\nu$ denote the natural projection and $i_\nu : \mathbf{J}_N^{-1}(\nu) \hookrightarrow P$ be the inclusion. The *reduced symplectic structure* on P_ν denoted by Ω_ν is uniquely determined by the condition $i_\nu^*\Omega = \pi_\nu^*\Omega_\nu$.

Induced Coadjoint Actions of Quotient Groups.   To prepare for the second reduced space, we start with some general remarks. 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, one sees that *the restriction of the action of M on $\mathfrak{n}^*$ to the subgroup N coincides with the coadjoint action of N on $\mathfrak{n}^*$.*

It may be useful to express the preceding result in a formula. Let, temporarily, $\mathrm{Ad}^*_{m^{-1},M}$ denote the coadjoint action of M on $\mathfrak{m}^*$ and similarly, $\mathrm{Ad}^*_{n^{-1},N}$ denote the coadjoint action of N on $\mathfrak{n}^*$. Then in this notation, the action of M on $\mathfrak{n}^*$ given in (5.2.1) reads

$$ m \cdot i^*(\mu) = i^* \mathrm{Ad}^*_{m^{-1},M} \mu. $$

The result mentioned in the preceding paragraph then can be written as

$$ i^* \mathrm{Ad}^*_{n^{-1},M} \mu = \mathrm{Ad}^*_{n^{-1},N} i^* \mu \tag{5.2.3} $$

for all $n \in N$ and $\mu \in \mathfrak{m}^*$.

Let M_ν denote the isotropy subgroup of $\nu \in \mathfrak{n}^*$ for the action of M on $\mathfrak{n}^*$. We now show, using the preceding remarks and normality of N in M, that

$$ N_\nu = M_\nu \cap N. \tag{5.2.4} $$

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 remarks, this means that n fixes ν using the coadjoint action. The other inclusion is obvious.[2]

It is an elementary fact that the intersection of a normal subgroup N with another subgroup is normal in that subgroup, and in particular, the subgroup $N_\nu \subset M$ is normal in M_ν. Thus, we can form the quotient group M_ν / N_ν. Since N_ν is normal, we can think of this quotient space as either the set of right cosets or left cosets. However, in what follows, it is more

---

[2]**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.

convenient to think of the class of $m \in M_\nu$ as $[m] = mN_\nu$. This quotient group will play an important role in what follows.

In the context of semidirect products, where $M = G \circledS V$ is the semidirect product of a group G with a vector space V on which G is represented, we write $\nu = a$ and M_ν/N_ν reduces to the isotropy group G_a for the contragredient action of G on V^*. As we saw in §4.2, this is a key ingredient in semidirect product theory.

Quotient Group Actions.   The following general lemma will be useful when we consider the action of quotient groups, and in particular, the group M_ν/N_ν.

5.2.1 Lemma.   *Let a Lie group K act in a free and proper way on a manifold R, and let $L \subset K$ be a (closed) normal subgroup. Then the quotient group $K/L := \{kL \mid k \in K\}$ acts in a free and proper way on the quotient space R/L.*

Proof.   First of all, the action of L on R is also free and proper, since L is a closed subgroup of K which, by assumption, acts freely and properly. Thus, R/L is a smooth manifold and the natural projection $\pi_{R,L} : R \to R/L$ is a surjective submersion.

Denoting the equivalence class of a point $x \in R$ in R/L by $[x]_L$, and classes in K/L by $[k]_L := kL$, the quotient action is

$$[k]_L \cdot [x]_L = [k \cdot x]_L$$

which is well–defined since for any $l, l' \in L$, we have

$$
\begin{aligned}
[kl]_L \cdot [l' \cdot x]_L &= [kll' \cdot x]_L \\
&= [(kll'k^{-1}) \cdot k \cdot x]_L \\
&= [k \cdot x]_L
\end{aligned}
$$

since $kll'k^{-1} \in L$ as L is a normal subgroup. Thus, the quotient action is well–defined. To show it is free, note that if

$$[k]_L \cdot [x]_L = [k \cdot x]_L = [x]_L$$

then there is an $l \in L$ such that $lk \cdot x = x$. Freeness of the K action implies that $lk = e$ and so $k \in L$; thus, $[k]_L$ is the identity element in K/L.

To check properness, assume that $[x_n]_L \to [x]_L$ and $[k_n \cdot x_n]_L \to [x']_L$. We need to show that there is a convergent subsequence $[k_{n_p}]_L$. Since $[x_n]_L \to [x]_L$, by the definition of the quotient topology, and the fact that proper actions have slices (see the discussions in, for example, [MTA] and Duistermaat and Kolk [1999]), there is a sequence $l_n \in L$ such that $l_n \cdot x_n \to x$. Similarly, there is a sequence $l'_n \in L$ such that $l'_n k_n \cdot x_n \to x'$. Now write

$$l'_n k_n \cdot x_n = k_n \bar{l}_n \cdot (l_n \cdot x_n) \to x',$$

where $\bar{l}_n = k_n^{-1} l n' k_n \in L$. Thus, by properness of the K action on R, there is a convergent subsequence $k_{n_p} \bar{l}_{n_p}$. Therefore, $[k_{n_p}]_L$ converges in K/L. ∎

This induced action can be rephrased in the following way. If $\Psi_k : R \to R$ denotes the given action of a group element $k \in K$, and $\Psi_{[k]_L}^L : R/L \to R/L$ denotes the quotient action of an element $[k]_L \in K/L$, then

$$\Psi_{[k]_L}^L \circ \pi_{R,L} = \pi_{R,L} \circ \Psi_k,$$

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

The Action of M_ν / N_ν on P_ν.   We are now ready to study the action of M_ν / N_ν on the first reduced space.

5.2.2 Lemma.   *There is a well-defined induced free and proper 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 $\Phi_m^\nu : \mathbf{J}_N^{-1}(\nu) \to \mathbf{J}_N^{-1}(\nu)$.

The general considerations given in Lemma 5.2.1 show that the group M_ν / N_ν has a well-defined free and proper 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,

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

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

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

for each $m \in M_\nu$.

Recall from the regular Symplectic Reduction Theorem that $i_\nu^* \Omega = \pi_\nu^* \Omega_\nu$. Therefore,

$$\pi_\nu^* \Psi_{[m],\nu}^* \Omega_\nu = (\Phi_m^\nu)^* \pi_\nu^* \Omega_\nu = (\Phi_m^\nu)^* i_\nu^* \Omega = i_\nu^* \Phi_m^* \Omega = i_\nu^* \Omega = \pi_\nu^* \Omega_\nu.$$

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.   The following lemmas show that there is a momentum map for the action of M_ν/N_ν on P_ν. This momentum map, $\mathbf{J}_\nu : P_\nu \to (\mathfrak{m}_\nu/\mathfrak{n}_\nu)^*$, is induced from the original momentum map on the total space, $\mathbf{J}_M : P \to \mathfrak{m}^*$. The example of the Heisenberg group, discussed in §5.1 (as well as in the next Chapter), shows that this momentum map, $\mathbf{J}_\nu$ *need not be equivariant.*

We prepare the following elementary but useful lemma.

5.2.3 Lemma.   *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

$$\mathrm{Ad}_n \xi - \xi = \frac{d}{dt}\bigg|_{t=0} [I_n(\exp(t\xi))]\exp(-t\xi)$$

$$= \frac{d}{dt}\bigg|_{t=0} (n[\exp(t\xi)]n^{-1})\exp(-t\xi)$$

$$= \frac{d}{dt}\bigg|_{t=0} n[\exp(t\xi))n^{-1}\exp(-t\xi)]. \qquad (5.2.5)$$

Since N is a normal subgroup, $\exp(t\xi))n^{-1}\exp(-t\xi)$ is a curve in N (that passes through the point n^{-1} at $t = 0$), so the resulting tangent vector to this curve is an element of $\mathfrak{n}$.   ▼

5.2.4 Lemma.   *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*

$$(r'_\nu)^* \circ \mathbf{J}_\nu \circ \pi_\nu = k_\nu^* \circ \mathbf{J}_M \circ i_\nu - \bar{\nu} \qquad (5.2.6)$$

where the notation used is as follows (see Figure 5.2.1; other notation in this diagram will be explained below). First of all,

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

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 $\mathfrak{m}_\nu$. Formula (5.2.6) is equivalently written as

$$\langle \mathbf{J}_\nu([z]), [\xi] \rangle = \langle \mathbf{J}_M(z), \xi \rangle - \langle \bar{\nu}, \xi \rangle \qquad (5.2.7)$$

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 $\mathfrak{m}_\nu/\mathfrak{n}_\nu$.

Remark. If N is a compact connected Lie group, all coadjoint isotropy subgroups N_ν are connected (see Theorem 3.3.1 in Duistermaat and Kolk [1999]).
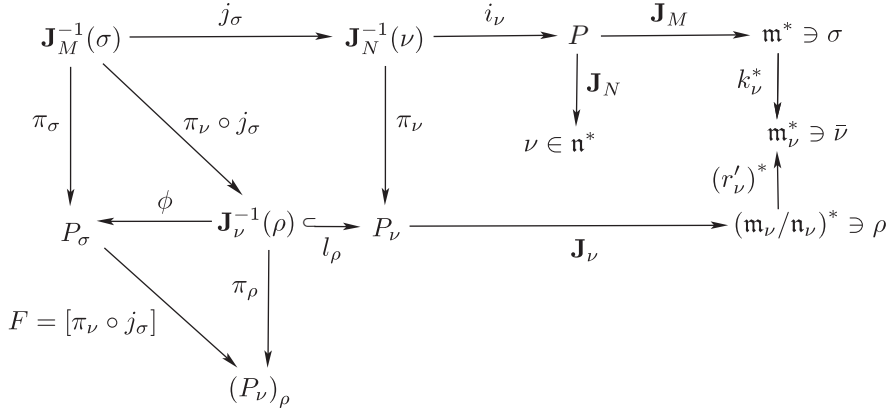

FIGURE 5.2.1. Some of the maps involved in symplectic reduction by stages.

The main point of this lemma is to show that the equation (5.2.6), or equivalently (5.2.7), gives a well-defined map $\mathbf{J}_\nu$, which takes some checking. We turn to this task in the following proof.

Proof.   First of all, we show that the definition of the map $\mathbf{J}_\nu$ in equation (5.2.7) is independent of the representative of $[\xi]$. To do this, it suffices to show that the right hand side of (5.2.7) 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

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

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

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

by Lemma 5.2.3, 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

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

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

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

However,

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

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.   ▼

5.2.5 Lemma.   *The map $\mathbf{J}_\nu$ 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 (5.2.6) 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 associated with the Lie algebra element $[\xi] = r'_\nu(\xi) \in \mathfrak{m}_\nu/\mathfrak{n}_\nu$. Since

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

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,

$$\begin{aligned}
[\xi]_{P_\nu}([z]) &= \left.\frac{d}{dt}\right|_{t=0} \exp_\nu tr'_\nu(\xi) \cdot \pi_\nu(z) \\
&= \left.\frac{d}{dt}\right|_{t=0} r_\nu(\exp t\xi) \cdot \pi_\nu(z) \\
&= \left.\frac{d}{dt}\right|_{t=0} \pi_\nu(\exp t\xi \cdot z) \\
&= T_z\pi_\nu(\xi_P(z)),
\end{aligned} \tag{5.2.8}$$

that is,

$$[\xi]_{P_\nu}([z]) = T_z\pi_\nu(\xi_P(z)). \tag{5.2.9}$$

Thirdly, denote by $J_M^\xi : P \to \mathbb{R}$, the map $J_M^\xi(z) = \langle \mathbf{J}_M(z), \xi \rangle$ and similarly for $J_\nu^{[\xi]} : P \to \mathbb{R}$ and note that equation (5.2.6) can be written as

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

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

$$\mathbf{d}J_\nu^{[\xi]}(\pi_\nu(z)) \cdot T_z\pi_\nu(v) = \mathbf{d}J_M^\xi(z) \cdot v \tag{5.2.10}$$

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 (5.2.9) and (5.2.10)

$$\begin{aligned}
\Omega_\nu([z])\left([\xi]_{P_\nu}([z]), T_z\pi_\nu(v)\right) &= \Omega_\nu(\pi_\nu(z))\left(T_z\pi_\nu(\xi_P(z)), T_z\pi_\nu(v)\right) \\
&= (\pi_\nu^*\Omega_\nu)(z)(\xi_P(z), v) \\
&= (i_\nu^*\Omega)(z)(\xi_P(z), v) = \mathbf{d}J_M^\xi(z) \cdot v \\
&= \mathbf{d}J_\nu^{[\xi]}(\pi_\nu(z)) \cdot T_z\pi_\nu(v),
\end{aligned} \tag{5.2.11}$$

which proves that $\mathbf{J}_\nu$ given by (5.2.6) is a momentum map for the M_ν/N_ν–action on P_ν.   ▼

The Cocycle for the Induced Momentum Map $\mathbf{J}_\nu$. Recall (see [MandS], Section 12.4) that for a cocycle associated to a *nonequivariant*

momentum map. to be well-defined, the underlying space should be connected. Thus, we shall assume tacitly that *the first reduced space is connected* (or that one restricts one's attention to one component of it). We now state the extent to which we have a lack of equivariance of $\mathbf{J}_\nu$ by exhibiting a defining equation for the associated cocycle.

5.2.6 Proposition.   *The $(\mathfrak{m}_\nu/\mathfrak{n}_\nu)^*$–valued one–cocycle 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. That is, we have the identity

$$\mathbf{J}_\nu([m][z]) - \mathrm{Ad}^*_{[m]^{-1}}\,\mathbf{J}_\nu([z]) = \varpi([m]).$$

Proof.   Since

$$\mathrm{Ad}_{r_\nu(m)}\,r'_\nu(\xi) = r'_\nu(\mathrm{Ad}_m\,\xi) \qquad\qquad (5.2.12)$$

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

$$
\begin{aligned}
\langle \mathbf{J}_\nu([m][z]) &- \mathrm{Ad}^*_{[m]^{-1}}\,\mathbf{J}_\nu([z]), [\xi]\rangle \\
&= \langle \mathbf{J}_\nu([mz]), [\xi]\rangle - \langle \mathbf{J}_\nu([z]), \mathrm{Ad}_{[m]^{-1}}[\xi]\rangle \\
&= \langle \mathbf{J}_M(mz), \xi\rangle - \langle \bar\nu, \xi\rangle - \langle \mathbf{J}_\nu([z]), [\mathrm{Ad}_{m^{-1}}\,\xi]\rangle \\
&= \langle \mathrm{Ad}^*_{m^{-1}}\,\mathbf{J}_M(z), \xi\rangle - \langle \bar\nu, \xi\rangle - \langle \mathbf{J}_M(z), \mathrm{Ad}_{m^{-1}}\,\xi\rangle + \langle \bar\nu, \mathrm{Ad}_{m^{-1}}\,\xi\rangle \\
&= \langle \bar\nu, \mathrm{Ad}_{m^{-1}}\,\xi - \xi\rangle \\
&= \langle \mathrm{Ad}^*_{m^{-1}}\,\bar\nu - \bar\nu, \xi\rangle. \qquad\qquad (5.2.13)
\end{aligned}
$$

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

$$
\begin{aligned}
\langle \mathrm{Ad}^*_{m^{-1}}\,\bar\nu, \xi\rangle &= \langle \bar\nu, \mathrm{Ad}_{m^{-1}}\,\xi\rangle \\
&= \langle \nu, \mathrm{Ad}_{m^{-1}}\,\xi\rangle \\
&= \langle m\nu, \xi\rangle = \langle \nu, \xi\rangle, \qquad\qquad (5.2.14)
\end{aligned}
$$

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

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

where $\mathfrak{n}_\nu^0 = \{\lambda \in \mathfrak{m}_\nu^* \mid \lambda|\mathfrak{n}_\nu = 0\}$ 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

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

Since

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

it follows that[3]

$$(r'_\nu)^* \left((\mathfrak{m}_\nu/\mathfrak{n}_\nu)^* \right) = \mathfrak{n}_\nu^0. \tag{5.2.15}$$

Because of (5.2.15) 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 (5.2.12) we get

$$
\begin{aligned}
(r'_\nu)^*(\varpi(m_1 m_2)) &= \mathrm{Ad}^*_{(m_1 m_2)^{-1}} \bar\nu - \bar\nu \\
&= \mathrm{Ad}^*_{m_1^{-1}} \mathrm{Ad}^*_{m_2^{-1}} \bar\nu - \mathrm{Ad}^*_{m_1^{-1}} \bar\nu + \mathrm{Ad}^*_{m_1^{-1}} \bar\nu - \bar\nu \\
&= \mathrm{Ad}^*_{m_1^{-1}} \left(\mathrm{Ad}^*_{m_2^{-1}} \bar\nu - \bar\nu \right) + \mathrm{Ad}^*_{m_1^{-1}} \bar\nu - \bar\nu \\
&= \mathrm{Ad}^*_{m_1^{-1}} (r'_\nu)^*(\varpi(m_2)) + (r'_\nu)^*(\varpi(m_1)) \\
&= (r'_\nu)^*(\varpi(m_1)) + (r'_\nu)^* \left(\mathrm{Ad}^*_{[m_1]^{-1}} \varpi(m_2) \right) \\
&= (r'_\nu)^* \left(\varpi(m_1) + \mathrm{Ad}^*_{[m_1]^{-1}} \varpi(m_2) \right). \tag{5.2.16}
\end{aligned}
$$

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

$$\varpi(m_1 m_2) = \varpi(m_1) + \mathrm{Ad}^*_{[m_1]^{-1}} \varpi(m_2). \tag{5.2.17}$$

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 $\xi \in \mathfrak{m}_\nu$,

$$
\begin{aligned}
\langle (r'_\nu)^*(\varpi(n)), \xi \rangle &= \langle \mathrm{Ad}^*_{n^{-1}} \bar\nu - \bar\nu, \xi \rangle \\
&= \langle \bar\nu, \mathrm{Ad}_{n^{-1}} \xi - \xi \rangle \\
&= \langle \nu, \mathrm{Ad}_{n^{-1}} \xi - \xi \rangle, \tag{5.2.18}
\end{aligned}
$$

since by Lemma 5.2.3, $\mathrm{Ad}_{n^{-1}} \xi - \xi \in \mathfrak{n}_\nu$. 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 (5.2.17) shows that it is a one-cocycle on M_ν/N_ν. ∎

---

[3]We take the expedient view that in infinite dimensions, this needs to be proven on a case by case basis.

Although the momentum map $\mathbf{J}_\nu$ depends on the extension of ν to $\mathfrak{m}_\nu^*$, different choices of extension lead to momentum maps that are equivalent with respect to the cohomology class of their cocycles. That is, if $\bar{\nu}$ and $\bar{\nu}'$ are two choices extending $\nu|_{\eta_\nu}$, then it follows from the defining property of the cocycle that for all $[m] \in M_\nu/N_\nu$,

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

for the unique $\lambda \in (\mathfrak{m}_\nu/\mathfrak{n}_\nu)^*$ that satisfies

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

This demonstrates that the two group one-cocycles on M_ν/N_ν differ by a coboundary.

The Affine Action.   In the case of nonequivariant momentum maps, reduction may be carried out by modifying the coadjoint action with a cocycle, as we did above for the Heisenberg group (see, for instance Souriau [1970], [FofM], [MandS], and [HRed]).

Namely, for

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

we consider the modified action

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

In terms of a given σ (the value at which we will do the full reduction, as above), we can define $\nu = \sigma|_\mathfrak{n} \in \mathfrak{n}^*$ and $\rho \in (\mathfrak{m}_\nu/\mathfrak{n}_\nu)^*$ by

$$(r'_\nu)^*(\rho) = \sigma|_{\mathfrak{m}_\nu} - \bar{\nu}, \tag{5.2.19}$$

where $\bar{\nu}$ is an arbitrary extension of $\nu|_{\mathfrak{n}_\nu}$ to $\mathfrak{m}_\nu$. Observe that ρ depends on the choice of the extension $\bar{\nu}$ of ν. Equation (5.2.19) 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, $\sigma|_{\mathfrak{m}_\nu} - \bar{\nu} \in \mathfrak{n}_\nu^0$.

The Second Reduced Space.   In the first stage of reduction, we have reduced P by the action of N at the point ν to obtain the **first reduced space** P_ν. The space P_ν can be further reduced by the action of the Lie group M_ν/N_ν at a value $\rho \in (\mathfrak{m}_\nu/\mathfrak{n}_\nu)^*$. Recall that we have proved that the action of M_ν/N_ν is free and proper, so that this reduced space is regular. Let this **second reduced space** be denoted by

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

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

Compatibility. We also require that the various values at which we are reducing satisfy the **compatibility relation**

$$(r'_\nu)^*(\rho) = k_\nu^* \sigma - \bar\nu. \tag{5.2.20}$$

As above, one may view this relation as defining $\nu = \sigma|_{\mathfrak{n}}$ and ρ in terms of σ and the extension $\bar\nu$.

In the next Proposition, the following general notation is used. Given a group G and two subgroups H and K of G. Define the product $H \cdot K$ of these two subgroups to be the set

$$H \cdot K := \{hk \mid h \in H, k \in K\}.$$

It is an easy general fact that if $H \cdot K = K \cdot H$ (as sets), then $H \cdot K$ is a subgroup of G.

5.2.7 Proposition. *We have $M_\sigma \subset (M_\nu)_{\sigma|_{\mathfrak{m}_\nu}}$. Also,*

$$N_\nu \cdot (M_\nu)_{\sigma|_{\mathfrak{m}_\nu}} = (M_\nu)_{\sigma|_{\mathfrak{m}_\nu}} \cdot N_\nu$$

is a closed Lie subgroup of M_ν containing the subgroup N_ν. In addition,

$$\begin{aligned}(M_\nu/N_\nu)_\rho &= r_\nu\left((M_\nu)_{\sigma|_{\mathfrak{m}_\nu}}\right)\\ &= \left((M_\nu)_{\sigma|_{\mathfrak{m}_\nu}} \cdot N_\nu\right)/N_\nu\\ &= \left(N_\nu \cdot (M_\nu)_{\sigma|_{\mathfrak{m}_\nu}}\right)/N_\nu.\end{aligned}$$

As a consequence, $r_\nu(M_\sigma) \subset (M_\nu/N_\nu)_\rho$.

As the general notation indicates, $N_\nu \cdot (M_\nu)_{\sigma|_{\mathfrak{m}_\nu}}$ means the set of products of the two groups N_ν and $(M_\nu)_{\sigma|_{\mathfrak{m}_\nu}}$.

Proof. First of all, $M_\sigma \subset M_\nu$ by the definition (5.2.1) of the action of M on $\mathfrak{n}^*$. But $m \in M_\sigma$ leaves σ invariant under the coadjoint action of M on $\mathfrak{m}^*$ and therefore it also leaves $\sigma|_{\mathfrak{m}_\nu}$ invariant. Thus, $M_\sigma \subset (M_\nu)_{\sigma|_{\mathfrak{m}_\nu}}$.

We assert that the equality

$$N_\nu \cdot (M_\nu)_{\sigma|_{\mathfrak{m}_\nu}} = (M_\nu)_{\sigma|_{\mathfrak{m}_\nu}} \cdot N_\nu$$

follows because N_ν is normal in M_ν. To show this, let $n \in N_\nu$ and $m \in (M_\nu)_{\sigma|_{\mathfrak{m}_\nu}} \subset M_\nu$, so that $nm \in N_\nu \cdot (M_\nu)_{\sigma|_{\mathfrak{m}_\nu}}$. It follows that $nm = m(m^{-1}nm)$, which lies in $(M_\nu)_{\sigma|_{\mathfrak{m}_\nu}} \cdot N_\nu$. This proves our assertion.

Thus, by our preparatory remarks, $N_\nu \cdot (M_\nu)_{\sigma|_{\mathfrak{m}_\nu}}$ is a group. We shall prove that it is a closed Lie group shortly.

Next we prove that $(M_\nu/N_\nu)_\rho = r_\nu\left((M_\nu)_{\sigma|_{\mathfrak{m}_\nu}}\right)$. To see this, notice that $[m] \in (M_\nu/N_\nu)_\rho$ if and only if

$$[m]\rho = \rho. \tag{5.2.21}$$

Since $r'_\nu : \mathfrak{m}_\nu \to \mathfrak{m}_\nu / \mathfrak{n}_\nu$ is a projection, its dual $(r'_\nu)^* : (\mathfrak{m}_\nu / \mathfrak{n}_\nu)^* \to \mathfrak{m}_\nu^*$ is injective and hence (5.2.21) is equivalent to $(r'_\nu)^*([m]\rho) = (r'_\nu)^*(\rho)$. However, by (5.2.20), $(r'_\nu)^*(\rho) = k_\nu^* \sigma - \bar{\nu} = \sigma|_{\mathfrak{m}_\nu} - \bar{\nu}$. On the other hand, by the definition of the coadjoint action of M_ν / N_ν, (5.2.20), and the formula for $(r'_\nu)^*(\varpi([m]))$ in Proposition 5.2.6, we get

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

since $m \in M_\nu$. Therefore (5.2.21) is equivalent to

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

that is, $m \in (M_\nu)_{\sigma|_{\mathfrak{m}_\nu}}$. It follows that $r_\nu(m) = [m]$ lies in the set $(M_\nu / N_\nu)_\rho$ if and only if $m \in (M_\nu)_{\sigma|_{\mathfrak{m}_\nu}}$ for all representatives m of $[m]$. This proves that $(M_\nu / N_\nu)_\rho = r_\nu \left((M_\nu)_{\sigma|_{\mathfrak{m}_\nu}} \right)$.

Next, observe that because $r_\nu : M_\nu \to M_\nu / N_\nu$ is the quotient map,

$$
\begin{aligned}
r_\nu \left((M_\nu)_{\sigma|_{\mathfrak{m}_\nu}} \right) &= r_\nu^{-1} \left(r_\nu \left((M_\nu)_{\sigma|_{\mathfrak{m}_\nu}} \right) \right) / N_\nu \\
&= \left((M_\nu)_{\sigma|_{\mathfrak{m}_\nu}} \cdot N_\nu \right) / N_\nu.
\end{aligned}
$$

This equals $\left(N_\nu \cdot (M_\nu)_{\sigma|_{\mathfrak{m}_\nu}} \right) / N_\nu$ by the equality of the numerators that we have already proved.

The claim that we have a closed Lie subgroup is a consequence of the equality

$$
(M_\nu)_{\sigma|_{\mathfrak{m}_\nu}} \cdot N_\nu = r_\nu^{-1} \left((M_\nu / N_\nu)_\rho \right).
$$

Indeed, since $r_\nu : M_\nu \to M_\nu / N_\nu$ is a Lie group homomorphism and $(M_\nu / N_\nu)_\rho$ is a closed Lie subgroup of M_ν / N_ν its inverse image is a closed Lie subgroup of M_ν. ∎

Stages Hypothesis I.   Now we shall require a special hypothesis to state the main reduction by stages theorem. This hypothesis will be satisfied in all of our examples and also will be shown to be, to some extent, unnecessary if appropriate alternative topological conditions hold. See the comments following the proof of Theorems 5.2.9 and 5.2.10; this issue is further discussed in Chapter 11.

5.2.8 Definition (Stages Hypothesis I).   *The element $\sigma \in \mathfrak{m}^*$ will be said to **satisfy the stages hypothesis** if for any other element $\sigma' \in \mathfrak{m}^*$ such that*

$$
\sigma|_{\mathfrak{n}} = \sigma'|_{\mathfrak{n}} =: \nu \quad and \quad \sigma|_{\mathfrak{m}_\nu} = \sigma'|_{\mathfrak{m}_\nu},
$$

there exists $n \in N_\nu \cdot (M_\nu)_{\sigma|_{\mathfrak{m}_\nu}}$ such that $\sigma' = \mathrm{Ad}^*_{n^{-1}} \sigma$, where Ad^* is the M-coadjoint action on $\mathfrak{m}^*$.

Notice that the stages hypothesis involves only the structure of the group M and the choice of σ and does not involve the action of M on P.

There are a few useful remarks to be made at this point.

1. The stages hypothesis always holds for $\sigma = 0$ (since this forces the condition $\sigma' = 0$) and hence one can choose $n = e$.

2. The stages hypothesis holds if N is in the center of M. Indeed, in this case, one readily sees that $M_\nu = M$ and so $\mathfrak{m}_\nu = \mathfrak{m}$ and so this forces $\sigma = \sigma'$ and hence one can choose $n = e$. We shall use this fact when we consider reduction by stages for the oscillator group in §9.3.

3. If one can find $n \in N_\nu$ such that $\sigma' = \mathrm{Ad}^*_{n^{-1}} \sigma$, then the stages hypothesis holds. This is clear because $N_\nu \subset N_\nu \cdot (M_\nu)_{\sigma|_{\mathfrak{m}_\nu}}$. In fact, this is the condition that we shall establish in the various examples considered later.

4. There is a sense in which this stages hypothesis is redundant, which we shall explore in greater depth as we proceed to develop the theory in later sections.

Point Reduction by Stages I. We are now ready to state the first reduction by stages theorem, which will assume the preceding stages hypothesis.

5.2.9 Theorem (Point Reduction by Stages I). *Making use of the notations and hypotheses introduced above, there is a symplectic diffeomorphism between the symplectic manifolds P_σ and $(P_\nu)_\rho$.*

This theorem states, in particular, that the spaces $(P_\nu)_\rho$ are mutually symplectically diffeomorphic for different choices of extensions $\bar{\nu}$ and so, correspondingly different choices of ρ.

Proof. Since σ and ν are related by $\nu = \sigma|_{\mathfrak{n}} \in \mathfrak{n}^*$ and $\mathbf{J}_N = i^* \circ \mathbf{J}_M$, there is a natural inclusion map

$$j_\sigma : \mathbf{J}_M^{-1}(\sigma) \to \mathbf{J}_N^{-1}(\nu).$$

Composing this map with π_ν, we get the smooth map

$$\pi_\nu \circ j_\sigma : \mathbf{J}_M^{-1}(\sigma) \to P_\nu.$$

This map takes values in $\mathbf{J}_\nu^{-1}(\rho)$ because of the relations

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

and $(r'_\nu)^*(\rho) = \sigma|_{\mathfrak{m}_\nu} - \bar{\nu}$. Thus, we can regard $\pi_\nu \circ j_\sigma$ as a map

$$\pi_\nu \circ j_\sigma : \mathbf{J}_M^{-1}(\sigma) \to \mathbf{J}_\nu^{-1}(\rho). \tag{5.2.22}$$

We claim that this map drops on the quotient to a map $[\pi_\nu \circ j_\sigma] : P_\sigma \to (P_\nu)_\rho$. We would like to define this map as follows: choose an element $z \in \mathbf{J}_M^{-1}(\sigma)$ and set

$$[\pi_\nu \circ j_\sigma][z] = [(\pi_\nu \circ j_\sigma)(z)]$$

where $[z] \in \mathbf{J}_\mu^{-1}(\sigma)/M_\sigma = P_\sigma$ and

$$[(\pi_\nu \circ j_\sigma)(z)] \in \mathbf{J}_\nu^{-1}(\rho)/(M_\nu/N_\nu)_\rho = (P_\nu)_\rho.$$

We need to show that $[\pi_\nu \circ j_\sigma][z]$ is well-defined; that is, if $z' = h \cdot z$ for $h \in M_\sigma$, then

$$[(\pi_\nu \circ j_\sigma)(z')] = [(\pi_\nu \circ j_\sigma)(z)].$$

Next, we show that the map $\pi_\nu \circ j_\sigma : \mathbf{J}_M^{-1}(\sigma) \to \mathbf{J}_\nu^{-1}(\rho)$ is r_ν-equivariant with respect to the action of M_σ on its domain and $(M_\nu/N_\nu)_\rho$ on the range, where here, r_ν is regarded as a homomorphism of M_σ to $(M_\nu/N_\nu)_\rho$. By the definition of the action of $(M_\nu/N_\nu)_\rho$ on $\mathbf{J}_\nu^{-1}(\rho)$ we have, for $h \in M_\sigma$,

$$(\pi_\nu \circ j_\sigma)(h \cdot z) = \pi_\nu(h \cdot z) = [h \cdot z] = [h] \cdot [z] = r_\nu(h) \cdot (\pi_\nu \circ j_\sigma)(z).$$

Therefore,

$$\begin{aligned}
[(\pi_\nu \circ j_\sigma)(z')] &= [(\pi_\nu \circ j_\sigma)(h \cdot z)] \\
&= [r_\nu(h) \cdot (\pi_\nu \circ j_\sigma)(z)] \\
&= [(\pi_\nu \circ j_\sigma)(z)]
\end{aligned}$$

since $r_\nu(h) \in (M_\nu/N_\nu)_\rho$ by Proposition 5.2.7. This shows that we have constructed a smooth map

$$F := [\pi_\nu \circ j_\sigma] : P_\sigma \to (P_\nu)_\rho.$$

We now show that F is actually the symplectomorphism that we need. We start by showing that:

F is Injective.   Let $z, z' \in \mathbf{J}_M^{-1}(\sigma)$ be such that $F(\pi_\sigma(z)) = F(\pi_\sigma(z'))$. The construction of F implies that $(\pi_\rho \circ \pi_\nu \circ j_\sigma)(z) = (\pi_\rho \circ \pi_\nu \circ j_\sigma)(z')$ and, consequently, there exists an element $[m] \in (M_\nu/N_\nu)_\rho$ such that

$$(\pi_\nu \circ j_\sigma)(z') = [m] \cdot (\pi_\nu \circ j_\sigma)(z) = (\pi_\nu \circ j_\sigma)(m \cdot z).$$

At the same time, this equality implies the existence of an element $n \in N_\nu$ such that $z' = nm \cdot z$. Given that both z and z' lie in $\mathbf{J}_M^{-1}(\sigma)$ equivariance of $\mathbf{J}_M$ gives

$$\sigma = \mathbf{J}_M(z') = \mathbf{J}_M(nm \cdot z) = \mathrm{Ad}^*_{(nm)^{-1}} \mathbf{J}_M(z) = \mathrm{Ad}^*_{(nm)^{-1}} \sigma,$$

that is, $nm \in M_\sigma$ and therefore $\pi_\sigma(z) = \pi_\sigma(z')$, as required.

F **is Surjective.**   Let $\pi_\rho([z]) \in (P_\nu)_\rho$ be arbitrary, where $[z] = \pi_\nu(z) \in \mathbf{J}_\nu^{-1}(\rho)$ and $z \in \mathbf{J}_N^{-1}(\nu)$. Let $\sigma' := \mathbf{J}_M(z)$. Notice that

$$\sigma'|_\mathfrak{n} = i^* \mathbf{J}_M(z) = \mathbf{J}_N(z) = \nu = \sigma|_\mathfrak{n}.$$

Also, for any $\xi \in \mathfrak{m}_\nu$ we have

$$\langle \sigma', \xi \rangle = \langle \mathbf{J}_M(z), \xi \rangle = \langle \mathbf{J}_\nu([z]), r'_\nu(\xi) \rangle + \langle \bar\nu, \xi \rangle = \langle \rho, r'_\nu(\xi) \rangle + \langle \bar\nu, \xi \rangle = \langle \sigma, \xi \rangle.$$

Consequently, $\sigma'|_{\mathfrak{m}_\nu} = \sigma|_{\mathfrak{m}_\nu}$ and therefore, *by the stages hypothesis*, there exists an element $n \in N_\nu(M_\nu)_{\sigma|_{\mathfrak{m}_\nu}}$ such that $\sigma' = \mathrm{Ad}_{n^{-1}}^* \sigma$. By equivariance of $\mathbf{J}_M$, we have $z' := n^{-1} \cdot z \in \mathbf{J}_M^{-1}(\sigma)$. By construction, we have

$$F(\pi_\sigma(z')) = \pi_\rho(\pi_\nu(j_\sigma(z'))) = \pi_\rho(\pi_\nu(j_\sigma(n^{-1} \cdot z)))$$
$$= \pi_\rho\left(r_\nu(n^{-1}) \cdot \pi_\nu(z)\right) = \pi_\rho(\pi_\nu(z)),$$

since $r_\nu(n^{-1}) \in (M_\nu/N_\nu)_\rho$, by Proposition 5.2.7. This proves the surjectivity of F.

F **is Symplectic.**   The strategy for this is as follows: Let the symplectic form on the doubly reduced space $(P_\nu)_\rho$ be denoted $(\Omega_\nu)_\rho$. One is required to show that $F^*(\Omega_\nu)_\rho = \Omega_\sigma$, the reduced symplectic form on P_σ. Recall that Ω_σ is uniquely characterized by the identity $i_\sigma^* \Omega = \pi_\sigma^* \Omega_\sigma$, where $i_\sigma = i_\nu \circ j_\sigma : \mathbf{J}_M^{-1}(\sigma) \to P$ is the inclusion map (see Figure 5.2.1). Thus, one must show that $F^*(\Omega_\nu)_\rho$ satisfies the condition

$$\pi_\sigma^* F^*(\Omega_\nu)_\rho = i_\sigma^* \Omega.$$

However, using the maps in Figure 5.2.1, and the general relation (1.1.15) for reduced symplectic forms, we get

$$\pi_\sigma^* F^*(\Omega_\nu)_\rho = (F \circ \pi_\sigma)^*(\Omega_\nu)_\rho = (\pi_\rho \circ \pi_\nu \circ j_\sigma)^*(\Omega_\nu)_\rho$$
$$= j_\sigma^* \pi_\nu^* \pi_\rho^*(\Omega_\nu)_\rho = j_\sigma^* \pi_\nu^* l_\rho^* \Omega_\nu$$
$$= (l_\rho \circ \pi_\nu \circ j_\sigma)^* \Omega_\nu = (\pi_\nu \circ j_\sigma)^* \Omega_\nu$$
$$= j_\sigma^* \pi_\nu^* \Omega_\nu = j_\sigma i_\nu^* \Omega$$
$$= (i_\nu \circ j_\sigma)^* \Omega = i_\sigma^* \Omega.$$

This shows that F is symplectic.

Remark.   As a consequence of this, it follows that F is an *immersion*. It is a general theorem that if $f : P \to Q$ is a smooth bijective immersion with P Lindelöf or paracompact then f is a diffeomorphism. A proof of this is given later in Lemma 11.1.1. Thus, if our symplectic manifold P_σ were Lindelöf or paracompact, the proof would now be finished. However, we shall continue and construct an inverse without any additional assumptions.

Our next task is to construct an inverse to the map F and prove that this inverse is smooth.

Definition of the Map ϕ. The strategy is to first find a map $\phi :$ $\mathbf{J}_\nu^{-1}(\rho) \to P_\sigma$ that is invariant under $(M_\nu/N_\nu)_\rho$ and which will therefore induce a map $[\phi] : (P_\nu)_\rho \to P_\sigma$ which we will then prove is the inverse of F.

To do this, let $[z] \in \mathbf{J}_\nu^{-1}(\rho) \subset P_\nu$ be a given N_ν-class and let

$$z \in \pi_\nu^{-1}(\mathbf{J}_\nu^{-1}(\rho)) \subset \mathbf{J}_N^{-1}(\nu)$$

be a chosen representative in the fiber over $[z]$. Let $\sigma' := \mathbf{J}_M(z)$. Since $z \in \mathbf{J}_N^{-1}(\nu)$, we have $\sigma'|_\mathfrak{n} = \nu = \sigma|_\mathfrak{n}$. As in the proof that F is surjective, for any $\xi \in \mathfrak{m}_\nu$, by (5.2.7) and (5.2.19), we have

$$\begin{aligned}
\langle \sigma', \xi \rangle &= \langle \mathbf{J}_\nu([z]), [\xi] \rangle + \langle \bar{\nu}, \xi \rangle \\
&= \langle (r_\nu')^* \rho, \xi \rangle + \langle \bar{\nu}, \xi \rangle = \langle \sigma, \xi \rangle .
\end{aligned}$$

Therefore, by the stages hypothesis, there is an $n \in (M_\nu)_{\sigma|_{\mathfrak{m}_\nu}} N_\nu$ such that $\mathrm{Ad}_{n^{-1}}^* \sigma' = \sigma$, and so by equivariance of $\mathbf{J}_M$, $n \cdot z \in \mathbf{J}_M^{-1}(\sigma)$. We then define $\phi([z]) := \pi_\sigma(n \cdot z)$. We next need to show that this map is well-defined.

To do this, choose another representative z' of $[z]$ and an element $n' \in N_\nu$ such that $z' = n' \cdot z$. As above, by the stages hypothesis, there is an $n_0 \in (M_\nu)_{\sigma|_{\mathfrak{m}_\nu}} N_\nu$ such that $n_0 \cdot z' \in \mathbf{J}_M^{-1}(\sigma)$. We need to show that $\pi_\sigma(n \cdot z) = \pi_\sigma(n_0 \cdot z')$. That is, that $n \cdot z$ and $n_0 \cdot z'$ are in the same M_σ-orbit. However, since each of them is in $\mathbf{J}_M^{-1}(\sigma)$, it is enough to show that they are in the same M-orbit by part (ii) of the reduction lemma 1.2.2. However, $n_0 \cdot z' = n_0 n \cdot z$, so that $n_0 \cdot z'$ and $n \cdot z$ are in the same M-orbit.

This completes the construction of ϕ.

Invariance of ϕ and the Inverse of F. Let $[z] \in \mathbf{J}_\nu^{-1}(\rho)$, $[m] \in (M_\nu/N_\nu)_\rho$, and $m \in (M_\nu)_{\sigma|_{\mathfrak{m}_\nu}}$ such that $r_\nu(m) = [m]$. We have

$$\phi([m] \cdot [z]) = \phi(\pi_\nu(m \cdot z)) = \pi_\sigma(n \cdot m \cdot z),$$

where $n \in N_\nu$ is such that $n \cdot m \cdot z \in \mathbf{J}_M^{-1}(\sigma)$. Let $n' \in (M_\nu)_{\sigma|_{\mathfrak{m}_\nu}} N_\nu$ be such that $n' \cdot z \in \mathbf{J}_M^{-1}(\sigma)$. To show that $\phi([z]) = \phi([m] \cdot [z])$, we must show that $n' \cdot z$ and $nm \cdot z$ are on the same M_σ orbit. Again by the reduction lemma, it is enough to show that they are on the same M-orbit. But this follows since $n' \cdot z = (n'm^{-1}n^{-1})nm \cdot z$.

Since ϕ is invariant, it induces a map on the quotient

$$[\phi] : \mathbf{J}_\nu^{-1}(\rho)/(M_\nu/N_\nu)_\rho \to P_\sigma.$$

To see that $[\phi]$ is the inverse of F, note that (see Figure 5.2.1), $F \circ \pi_\sigma = \pi_\rho \circ \pi_\nu \circ j_\sigma$ and hence

$$(F \circ [\phi])([z])) = F(\pi_\sigma(n \cdot z)) = (\pi_\rho \circ \pi_\nu \circ j_\sigma)(n \cdot z) = [z]$$

since $n \in N_\nu$.

Smoothness of the Inverse.   To show smoothness of the inverse, it is enough to show smoothness of ϕ since the inverse $[\phi]$ is the induced quotient map.

Let Γ be a smooth local section of the N_ν bundle $\pi_\nu^{-1}(\mathbf{J}_\nu^{-1}(\rho)) \to \mathbf{J}_\nu^{-1}(\rho)$ defined on a neighborhood U of a given point $[z_0] \in \mathbf{J}_\nu^{-1}(\rho)$. We claim that the smooth map $\mathbf{J}_M \circ \Gamma$ takes values in $(M_\nu)_{\sigma|_{\mathfrak{m}_\nu}} N_\nu \cdot \sigma$, the $(M_\nu)_{\sigma|_{\mathfrak{m}_\nu}} N_\nu$-orbit of σ in $\mathfrak{m}^*$ under the coadjoint action of M restricted to $(M_\nu)_{\sigma|_{\mathfrak{m}_\nu}} N_\nu$, so gives a smooth map

$$\mathbf{J}_M \circ \Gamma : U \to (M_\nu)_{\sigma|_{\mathfrak{m}_\nu}} N_\nu \cdot \sigma \simeq (M_\nu)_{\sigma|_{\mathfrak{m}_\nu}} N_\nu / ((M_\nu)_{\sigma|_{\mathfrak{m}_\nu}} N_\nu)_\sigma.$$

Indeed, the stages hypothesis produces, for each $\Gamma([z])$, as in the construction of ϕ, an element $n \in (M_\nu)_{\sigma|_{\mathfrak{m}_\nu}} N_\nu$ such that $\mathrm{Ad}_{n^{-1}}^* \mathbf{J}_M \Gamma([z]) = \sigma$. Thus, $\mathbf{J}_M \Gamma([z]) = \mathrm{Ad}_n^* \sigma \in (M_\nu)_{\sigma|_{\mathfrak{m}_\nu}} N_\nu \cdot \sigma$, as required.

Next, choose a local smooth lift of the map $\mathbf{J}_M \circ \Gamma$ from a neighborhood of σ (represented by the class of the identity element in $(M_\nu)_{\sigma|_{\mathfrak{m}_\nu}} N_\nu$) in $(M_\nu)_{\sigma|_{\mathfrak{m}_\nu}} N_\nu \cdot \sigma \cong (M_\nu)_{\sigma|_{\mathfrak{m}_\nu}} N_\nu / ((M_\nu)_{\sigma|_{\mathfrak{m}_\nu}} (M_\nu)_{\sigma|_{\mathfrak{m}_\nu}} N_\nu)_\sigma$ to a neighborhood of $n_0 \in N_\nu$ (chosen so that $\mathrm{Ad}_{n_0^{-1}}^* \mathbf{J}_M \Gamma([z_0]) = \sigma$) to give a smooth map

$$\widetilde{\mathbf{J}_M \circ \Gamma} : U \to (M_\nu)_{\sigma|_{\mathfrak{m}_\nu}} N_\nu.$$

We now claim that the local representative of ϕ is given by

$$\phi = \pi_\sigma \circ ((\widetilde{\mathbf{J}_M \circ \Gamma})^{-1} \cdot \Gamma) : U \to P_\sigma,$$

where the inverse is a pointwise inverse in the Lie group $(M_\nu)_{\sigma|_{\mathfrak{m}_\nu}} N_\nu$ (see the last statement in Proposition 5.2.7), which is clearly a smooth map. Explicitly, this equation means that for each $[z] \in U \subset \mathbf{J}_\nu^{-1}(\rho)$, then

$$\phi([z]) = \pi_\sigma((\widetilde{\mathbf{J}_M \circ \Gamma})([z])^{-1} \cdot \Gamma([z])). \tag{5.2.23}$$

To show this, we will prove that the group element $n \in (M_\nu)_{\sigma|_{\mathfrak{m}_\nu}} N_\nu$, chosen such that $n \cdot (\mathbf{J}_M \circ \Gamma)([z]) = \mathrm{Ad}_{n^{-1}}^* \mathbf{J}_M \Gamma([z]) = \sigma$, has the property that

$$\tilde{n} := n((\widetilde{\mathbf{J}_M \circ \Gamma})([z]) \in ((M_\nu)_{\sigma|_{\mathfrak{m}_\nu}} N_\nu)_\sigma. \tag{5.2.24}$$

But having a section of the bundle

$$(M_\nu)_{\sigma|_{\mathfrak{m}_\nu}} N_\nu \longrightarrow (M_\nu)_{\sigma|_{\mathfrak{m}_\nu}} N_\nu / ((M_\nu)_{\sigma|_{\mathfrak{m}_\nu}} N_\nu)_\sigma) \cong (M_\nu)_{\sigma|_{\mathfrak{m}_\nu}} N_\nu \cdot \sigma$$

means that we have

$$((\widetilde{\mathbf{J}_M \circ \Gamma})([z]) \cdot \sigma = (\mathbf{J}_M \circ \Gamma)([z]).$$

Therefore, $n^{-1} \cdot \sigma = ((\widetilde{\mathbf{J}_M \circ \Gamma})([z]) \cdot \sigma$ and so (5.2.24) is valid.

Next we compute the right hand side of (5.2.23) to be

$$\pi_\sigma\left((\widetilde{\mathbf{J}_M \circ \Gamma})([z])^{-1} \cdot \Gamma([z])\right) = \pi_\sigma\left(\tilde{n}^{-1} n \cdot \Gamma([z])\right) = \pi_\sigma\left(n \cdot \Gamma([z])\right)$$

since $\tilde{n} \in ((M_\nu)_{\sigma|_{\mathfrak{m}_\nu}} N_\nu)_\sigma \subset M_\sigma$. By definition of ϕ, the right hand side equals $\phi([z])$.   ∎

Remarks on the Stages Hypothesis.   An inspection of the preceding argument shows that a sufficient condition for the proof to go through (without the stages hypothesis) is the set theoretical equality

$$(\mathbf{J}_\nu \circ \pi_\nu \circ j_\sigma)\left(\mathbf{J}_M^{-1}(\sigma)\right) = \rho. \qquad (5.2.25)$$

Indeed, the stages hypothesis was only used to show that the map F was surjective and to construct the inverse map. This condition for reduction in stages will appear in Part III in the context of a much more general formulation of this problem (see Definition 15.3.1). Let us now show four things:

1. The inclusion
    $$(\pi_\nu \circ j_\sigma)\left(\mathbf{J}_M^{-1}(\sigma)\right) \subset \mathbf{J}_\nu^{-1}(\rho)$$
    holds without making the stages hypothesis.

2. If (5.2.25) holds, then we can construct an inverse to F.

3. The stages hypothesis implies the equality $(\pi_\nu \circ j_\sigma)\left(\mathbf{J}_M^{-1}(\sigma)\right) = \mathbf{J}_\nu^{-1}(\rho)$ and hence (5.2.25).

4. If the element $\sigma \in \mathfrak{m}^*$ satisfies the stages hypothesis so does any element $\mathrm{Ad}_{m^{-1}}^* \sigma$ in its orbit, $m \in M$. This remark will be used in the statement of Theorem 5.3.4 when we say that an *orbit $\mathcal{O}_\sigma$ satisfies the stages hypothesis*.

Proof.   1. This is just a restatement of equation (5.2.22). Notice that this inclusion was used to construct F itself. We are asserting that the opposite inclusion is needed to show that F is surjective and to construct its inverse.

2. The inverse to F was constructed as follows: choose $[z] \in \mathbf{J}_\nu^{-1}(\rho)$. By (5.2.25), we can write $\mathbf{J}_\nu([z]) = \rho = \mathbf{J}_\nu(\pi_\nu(z))$, where $z \in \mathbf{J}_M^{-1}(\sigma)$. Now, similar to the preceding proof, define $\phi([z]) = \pi_\sigma(z)$ and proceed as before.

3. Assume that the stages hypothesis holds. As remarked above, the inclusion $(\pi_\nu \circ j_\sigma)\left(\mathbf{J}_M^{-1}(\sigma)\right) \subset \mathbf{J}_\nu^{-1}(\rho)$ always holds, so we only need to prove the opposite inclusion. Thus, let $[z] \in \mathbf{J}_\nu^{-1}(\rho)$ and let $z \in \pi_\nu^{-1}\left(\mathbf{J}_\nu^{-1}(\rho)\right)$ be a representative of the class $[z]$. Let $\sigma' := \mathbf{J}_M(z)$. Since $z \in \mathbf{J}_N^{-1}(\nu)$, we have $\sigma'|_\mathfrak{n} = \nu = \sigma|_\mathfrak{n}$. Also, for any $\xi \in \mathfrak{m}_\nu$, by (5.2.7) and (5.2.19), we have

$$\langle \sigma', \xi \rangle = \langle \mathbf{J}_\nu([z]), [\xi] \rangle + \langle \bar\nu, \xi \rangle$$
$$= \langle (r_\nu')^* \rho, \xi \rangle + \langle \bar\nu, \xi \rangle = \langle \sigma, \xi \rangle.$$

Therefore, by the stages hypothesis, there is an $n \in (M_\nu)_{\sigma|_{\mathfrak{m}_\nu}} N_\nu$ such that $n \cdot \sigma' = \sigma$, and so, by equivariance, $n \cdot z \in \mathbf{J}_M^{-1}(\sigma)$.

4. It is a straightforward verification.

∎

Now we will give a further discussion of why in an appropriate sense, the stages hypothesis can be expected to hold in examples. In fact, we claim that the equality (5.2.25) is, in some sense, almost automatic. The insight formulated in the following theorem sets the stage for version III of reduction by stages, given later on, in which there are technical hypotheses of a topological nature, but no "algebraic" hypotheses (such as the stages assumption) whatsoever. Here is the key result that makes these remarks precise.

5.2.10 Theorem.   *Under the hypotheses made for the regular reduction by stages theorem, but without the stages hypothesis, $(\pi_\nu \circ j_\sigma)\left(\mathbf{J}_M^{-1}(\sigma)\right)$ is an open subset of the manifold $\mathbf{J}_\nu^{-1}(\rho)$.*

Proof.   First of all, we show that $(\pi_\nu \circ j_\sigma)\left(\mathbf{J}_M^{-1}(\sigma)\right)$ is a smooth manifold. To do so, note that it is the image of the smooth map $f := \pi_\nu \circ j_\sigma : \mathbf{J}_M^{-1}(\sigma) \to \mathbf{J}_\nu^{-1}(\rho)$. The kernel of Tf is just the collection of tangent spaces to the $(N_\nu)_\sigma$ orbits for the action of $(N_\nu)_\sigma$ on $\mathbf{J}_M^{-1}(\sigma)$. Since $(N_\nu)_\sigma$ acts freely, this collection of tangent spaces forms a smooth bundle. Thus, by the Fibration Theorem (see Theorem 3.5.18 of Abraham, Marsden and Ratiu [1988]), and the fact that π_ν is an open map, the image of f is a smooth manifold.

The strategy now is to show that at any point $[z] \in (\pi_\nu \circ j_\sigma)\left(\mathbf{J}_M^{-1}(\sigma)\right) \subset \mathbf{J}_\nu^{-1}(\rho)$, the tangent spaces to the two manifolds are the same. If this holds, then the result is proved. The two tangent spaces are given by

$$T_{[z]}\left(\pi_\nu \circ j_\sigma\right)\left(\mathbf{J}_M^{-1}(\sigma)\right) = T_z \pi_\nu \left(\ker T_z \mathbf{J}_M\right)$$

and

$$T_{[z]}(\mathbf{J}_\nu^{-1}(\rho)) = \ker T_{[z]}\mathbf{J}_\nu$$

Because of the inclusion $(\pi_\nu \circ j_\sigma)\left(\mathbf{J}_M^{-1}(\sigma)\right) \subset \mathbf{J}_\nu^{-1}(\rho)$, we know that

$$T_z \pi_\nu \left(\ker T_z \mathbf{J}_M\right) \subset \ker T_{[z]}\mathbf{J}_\nu$$

and so we must prove the opposite inclusion. To do so, we can represent an element of $\ker T_{[z]}\mathbf{J}_\nu$ as $T_z \pi_\nu(v_z)$ where $v_z \in T_z(\mathbf{J}_N^{-1}(\nu)) = \ker T_z \mathbf{J}_N$. Thus, we have

$$0 = \left(T_{[z]}\mathbf{J}_\nu \circ T_z \pi_\nu\right)(v_z) = T_z \left(\mathbf{J}_\nu \circ \pi_\nu\right)(v_z).$$

Therefore, by equation (5.2.6) (see also Figure 5.2.1), we get

$$0 = (r_\nu')^* \left(T_z \left(\mathbf{J}_\nu \circ \pi_\nu\right)(v_z)\right) = k_\nu^* T_z \mathbf{J}_M(v_z)$$

Since $k_\nu : \mathfrak{m}_\nu \to \mathfrak{m}$ is the inclusion, we see that $T_z \mathbf{J}_M(v_z)$ is zero when applied to elements of $\mathfrak{m}_\nu$. Since $T_z \mathbf{J}_N(v_z) = 0$ and $\mathbf{J}_N = i^* \circ \mathbf{J}_M$, it follows that $T_z \mathbf{J}_M(v_z)$ is also zero when applied to elements of $\mathfrak{n}$.

We now gather a few facts together so we can proceed:

1. The equality

$$T_z(M \cdot z) \cap T_z(\mathbf{J}_N^{-1}(\nu)) = T(M_\nu \cdot z) \qquad (5.2.26)$$

holds because for $\xi_P(z) \in T_z(M \cdot z) \cap T_z(\mathbf{J}_N^{-1}(\nu))$, we obtain $0 = T_z\mathbf{J}_N \cdot \xi_P(z) = -\operatorname{ad}_\xi^* \nu$ by equivariance, so $\xi \in \mathfrak{m}_\nu$. This shows the inclusion $\subset$; the opposite inclusion follows since M_ν leaves $\mathbf{J}_N^{-1}(\nu)$ invariant.

2. If (V, Ω) is a symplectic vector space and $E \subset V$ and $F \subset V$ are subspaces, then
$$(E \cap F)^\Omega = E^\Omega + F^\Omega. \qquad (5.2.27)$$

See, for example [MandS], §2.3.

3. For a momentum map $\mathbf{J} : P \to \mathfrak{g}^*$ for the symplectic action of a group G on a symplectic manifold P, we have

$$\ker T_z\mathbf{J} = (T_z(G \cdot z))^\Omega. \qquad (5.2.28)$$

This was proved in the Reduction Lemma, 1.2.2.

4. We have
$$\ker T_z\mathbf{J}_{M_\nu} = \ker T_z\mathbf{J}_M + T_z(N \cdot z). \qquad (5.2.29)$$

This follows from items 1., 2. and 3.

Now we return to the proof. We decompose our tangent vector v_z according to equation (5.2.29) as

$$v_z = w_z + \eta_P(z) \qquad (5.2.30)$$

where $w_z \in \ker T_z\mathbf{J}_M$ and $\eta \in \mathfrak{n}$. However, applying $T_z\mathbf{J}_N$ to both sides of (5.2.30), and using the facts that $v_z \in \ker T_z\mathbf{J}_N$ and also w_z is in the same kernel since it is even in the kernel of $T_z\mathbf{J}_M$, it follows that $T_z\mathbf{J}_N\eta_P(z) = 0$. By the reduction lemma again, it follows that $\eta \in \mathfrak{n}_\nu$. Finally, this shows that

$$T_z\pi_\nu(v_z) = T_z\pi_\nu(w_z) + T_z\pi_\nu(\eta_P(z)) = T_z\pi_\nu(w_z) \in T_z\pi_\nu\left(\ker T_z\mathbf{J}_M\right).$$

$\blacksquare$

The Stages Hypothesis for Right Actions.   We will need the following observation when we consider right actions of the group M in Chapter 8. Suppose the group M with normal subgroup N satisfies the stages hypothesis. We then know by Theorem 5.2.9 that for any left Hamiltonian action on a symplectic manifold P, that we can reduce by stages. Suppose we now consider a *right* Hamiltonian action of M on some symplectic manifold. There is a corresponding ***right stages hypothesis:***

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

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

there exists $n \in N_\nu \cdot (M_\nu)_\tau$ *such that* $\sigma_2 = \mathrm{Ad}^*_n \sigma_1$.

Like the left hypothesis, this has nothing to do with the action on the manifold. It only depends on the properties of the group M and its action on the normal subgroup N. It is easy to verify that the left stages hypothesis holds if and only if the right stages hypothesis holds since the condition is a statement about two elements of $\mathfrak{m}^*$ lying on the same $(M_\nu)_\tau$ orbit. It is therefore sufficient to check that the group satisfies the left stages hypothesis to conclude that even for a right action, we can reduce by stages.

5.3 Poisson and Orbit Reduction by Stages

In this section we first recall from the introduction, the simple process of Poisson reduction by stages. The symplectic leaves in Poisson reduction are naturally described by orbit reduction, which we recall here. One might think that, therefore, orbit reduction by stages is substantially simpler than point reduction by stages. The main result of this section is to show that this is not the case, but rather, that the orbit reduction point of view gives essentially the same result.

Poisson Reduction by Stages. Recall that if $(P, \{\,,\})$ is a Poisson manifold and if a Lie group G acts in a free and proper way on P by Poisson maps, then the quotient space P/G, which is a smooth manifold, is a Poisson manifold in a natural way obtained by simply declaring that the projection $\pi_{P,G} : P \to P/G$ is a Poisson map.

Assume the same general set up as in the preceding section, namely that we have a Lie group M with a normal Lie subgroup N and that M acts freely and properly on P (and so N also acts freely and properly on P). Then the Lie group M/N acts on the first reduced space P/N. One readily checks that this action is free and proper. Moreover, as in the proof of Lemma 5.2.2, we have

$$\Phi^{P/N,M/N}_{[m]_N} \circ \pi_{P,N} = \pi_{P,N} \circ \Phi^{P,M}_m, \tag{5.3.1}$$

where for the action of a group G on a space S, the action of an element $g \in G$ is denoted $\Phi^{S,G}_g$ and $\pi_{S,G} : S \to S/G$ is the projection map. Also, $[m]_N$ denotes the class of $m \in M$ in the quotient group M/N. It follows from (5.3.1) that the action of M/N on the first Poisson reduced space P/N is Poisson.

Thus, one can form the second Poisson reduced space, $(P/N)/(M/N)$. Then we have

5.3.1 Proposition.   *The Poisson manifolds P/M and $(P/N)/(M/N)$ are Poisson diffeomorphic.*

Proof.   The natural diffeomorphism between these spaces is given by

$$[z]_M \mapsto [[z]_N]_{M/N}.$$

It is easy to check that this is a well-defined map. It is smooth because it is the quotient of a smooth map. It is a diffeomorphism because it has the inverse

$$[[z]_N]_{M/N} \mapsto [z]_M,$$

which is likewise checked to be well-defined and smooth. The second of these maps is Poisson by the following reasoning. The quotient map $P/N \to P/M$ induced by the identity map on P, is Poisson and is M/N invariant, so we can quotient again, showing that the second map is Poisson. The first map, being its inverse, is therefore Poisson as well. ∎

Orbit Reduction by Stages.   Given how easy Poisson reduction by stages is, and the close relation of orbit reduction to its symplectic leaves, one might think that orbit reduction is correspondingly simple. We shall show in this paragraph that this is not so; rather it gives the same results as those of point reduction.

Assume that the Lie group M acts on (P, Ω) in a free, proper, and symplectic fashion. In the setting of orbit reduction that we presented in Theorem 1.2.3, the final reduced space is given by $P_{\mathcal{O}_\sigma} = \mathbf{J}_M(\mathcal{O}_\sigma)/M$, where $\mu \in \mathfrak{m}^*$ and $\mathcal{O}_\sigma \subset \mathfrak{m}^*$ is the M-coadjoint orbit through σ. Its connected components are indeed symplectic leaves of the final Poisson reduced space P/M.

The ***first orbit reduced space*** in this setting is obtained by letting $\nu = i^* \sigma \in \mathfrak{n}^*$ and forming

$$\mathbf{J}_N^{-1}(\mathcal{O}_\nu)/N \subset P/N.$$

Here, $\mathcal{O}_\nu$ denotes the N-coadjoint orbit of ν. Notice that in the definition of the first orbit reduced space *we have made a choice* of an element σ in the orbit $\mathcal{O}_\sigma$. Different choices for this element in $\mathcal{O}_\sigma$ will produce different first and second orbit reduced spaces. The orbit reduction by stages theorem in this section will state that all these doubly orbit reduced spaces are symplectomorphic to $P_{\mathcal{O}_\sigma}$ in the presence of an orbit reduction by stages hypothesis. This choice introduces an additional complication that was not present in the point reduction approach. Recall that in that situation once the element $\sigma \in \mathfrak{m}^*$ had been fixed so was $\nu = i^*(\sigma)$ and the first reduced space.

The quotient group M/N, while acting in a Poisson manner on P/N need not leave $\mathbf{J}_N^{-1}(\mathcal{O}_\nu)/N$ invariant; an example where it does not can be found using semidirect products. As was done in the case of point reduction,

in order to preserve the manifold $\mathbf{J}_N^{-1}(\mathcal{O}_\nu)/N$, we need to consider the stabilizer of the set $\mathcal{O}_\nu$ relative to the M-action (5.2.1) on $\mathfrak{n}^*$; namely, we consider the subset of M defined by

$$M_{\mathcal{O}_\nu} := \{m \in M \mid m \cdot \mathcal{O}_\nu \subset \mathcal{O}_\nu\}.$$

5.3.2 Lemma. $M_{\mathcal{O}_\nu}$ *is a closed Lie subgroup of M containing N as a normal subgroup. The map $\beta : M_\nu/N_\nu \to M_{\mathcal{O}_\nu}/N$ given by $mN_\nu \mapsto mN$, $m \in M_\nu$ is a Lie group isomorphism.*

Proof. Let $\widetilde{M_{\mathcal{O}_\nu}}$ be the maximal closed Lie subgroup contained in $M_{\mathcal{O}_\nu}$. Note that

$$N \subset \widetilde{M_{\mathcal{O}_\nu}} \subset M_{\mathcal{O}_\nu}.$$

By the normality of N in M we have that N_ν is a normal subgroup of $\widetilde{M_{\mathcal{O}_\nu}}$.

We shall prove that the map $\widetilde{\beta} : mN_\nu \in M_\nu/N_\nu \mapsto mN \in \widetilde{M_{\mathcal{O}_\nu}}/N$ is Lie group isomorphism. The map $\widetilde{\beta}$ is well-defined and smooth because it is the quotient of the (N_ν, N)-equivariant smooth inclusion $M_\nu \hookrightarrow \widetilde{M_{\mathcal{O}_\nu}}$.
(i) $\widetilde{\beta}$ is injective: let $m, m' \in M_\nu$ be such that $mN = m'N$. Thus, there is $n \in N$ such that $m' = mn$ and hence $n = m^{-1}m' \in M_\nu \cap N = N_\nu$, which implies that $mN_\nu = m'N_\nu$.
(ii) $\widetilde{\beta}$ is surjective: if $mN \in \widetilde{M_{\mathcal{O}_\nu}}/N$ then there exists $n \in N$ such that $m \cdot \nu = \mathrm{Ad}_{n^{-1}}^* \nu$ and hence $n^{-1}m \in M_\nu$. Therefore for some $n' \in N$ we can write $\widetilde{\beta}(n^{-1}mN_\nu) = n^{-1}mN = mn'N = mN$.

Since $\widetilde{\beta}$ is a bijective Lie group homomorphism it is a Lie group isomorphism (see page 42 in [HRed]).

The argument before also shows that the map β in the statement is a bijection. Hence it induces a unique Lie group structure on $M_{\mathcal{O}_\nu}/N$ relative to which β is a Lie group isomorphism with M_ν/N_ν. Therefore $M_{\mathcal{O}_\nu}/N$ is isomorphic as a Lie group to $\widetilde{M_{\mathcal{O}_\nu}}/N$ via the map $\beta \circ \widetilde{\beta}^{-1} : \widetilde{M_{\mathcal{O}_\nu}}/N \to M_{\mathcal{O}_\nu}/N$, which is the inclusion. In particular

$$\dim\left(\widetilde{M_{\mathcal{O}_\nu}}/N\right) = \dim\left(M_{\mathcal{O}_\nu}/N\right)$$

so $\widetilde{M_{\mathcal{O}_\nu}}/N$ is open in $M_{\mathcal{O}_\nu}/N$ and hence the connected components of the identity of both Lie groups are the same. Furthermore, since $\widetilde{M_{\mathcal{O}_\nu}}/N$ and $M_{\mathcal{O}_\nu}/N$ are isomorphic they are necessarily equal.

In order to conclude the proof let $\pi : M \to M/N$ be the projection. Since both sets $\widetilde{M_{\mathcal{O}_\nu}}$ and $M_{\mathcal{O}_\nu}$ are N-saturated we conclude that

$$\widetilde{M_{\mathcal{O}_\nu}} = \pi^{-1}\left(\widetilde{M_{\mathcal{O}_\nu}}/N\right) = \pi^{-1}\left(M_{\mathcal{O}_\nu}/N\right) = M_{\mathcal{O}_\nu},$$

which proves the claim since $\widetilde{M_{\mathcal{O}_\nu}}$ is a closed Lie subgroup of M. $\blacksquare$

5.3.3 Lemma. *The Lie subgroup $M_{\mathcal{O}_\nu}$ leaves $\mathbf{J}_N^{-1}(\mathcal{O}_\nu)$ invariant and hence the Lie group $M_{\mathcal{O}_\nu}/N$ acts on $\mathbf{J}_N^{-1}(\mathcal{O}_\nu)/N$ in a free, proper, and symplectic fashion.*

Proof. Let $m \in M_{\mathcal{O}_\nu}$ and $z \in \mathbf{J}_N^{-1}(\mathcal{O}_\nu)$. Then, since $\mathbf{J}_N = i^* \circ \mathbf{J}_M$, the definition (5.2.1) of the M-action on $\mathfrak{n}^*$ and M-equivariance of $\mathbf{J}_M$ yields

$$\mathbf{J}_N(m \cdot z) = i^*(\mathbf{J}_M(m \cdot z)) = i^*(\mathrm{Ad}_{m^{-1}}^* \mathbf{J}_M(z)) = m \cdot i^* \mathbf{J}_M(z)$$
$$= m \cdot \mathbf{J}_N(z) \in m \cdot \mathcal{O}_\nu \subset \mathcal{O}_\nu,$$

which shows that $M_{\mathcal{O}_\nu}$ leaves $\mathbf{J}_N^{-1}(\mathcal{O}_\nu)$ invariant. Since the action of M/N on P/N is free and proper, its restriction to the subgroup $M_{\mathcal{O}_\nu}/N$ and the immersed submanifold $\mathbf{J}_N^{-1}(\mathcal{O}_\nu)/N$ is also free and proper. The Lie group $M_{\mathcal{O}_\nu}$ acts symplectically on (P, Ω) so its quotient $M_{\mathcal{O}_\nu}/N$ acts symplectically on the orbit reduced space $\mathbf{J}_N^{-1}(\mathcal{O}_\nu)/N$. ∎

Now we show that the isomorphism $\beta : M_\nu/N_\nu \to M_{\mathcal{O}_\nu}/N$ is compatible with the diffeomorphism between P_ν and $P_{\mathcal{O}_\nu}$. To understand this, we need to use some notations for group actions. Recall that for the action of a group G on a space S, the action of an element $g \in G$ is denoted $\Phi_g^{S,G}$. One readily checks that we have the following commutative diagram for each $m \in M_\nu$:

$$
\begin{array}{ccc}
P_\nu & \xrightarrow{\;[i_\nu]\;} & P_{\mathcal{O}_\nu} \\[2pt]
\Phi_{[m]_{N_\nu}}^{P_\nu, M_\nu/N_\nu} \Big\downarrow & & \Big\downarrow \Phi_{[m]_N}^{P_{\mathcal{O}_\nu}, M_{\mathcal{O}_\nu}/N} \\[2pt]
P_\nu & \xrightarrow{\;[i_\nu]\;} & P_{\mathcal{O}_\nu}
\end{array}
$$

Indeed, start by selecting an element $[z]_{N_\nu} \in P_\nu$ where $z \in \mathbf{J}_\nu^{-1}(\nu)$ and select $m \in M_\nu$; then

$$\left(\Phi_{[m]_N}^{P_{\mathcal{O}_\nu}, M_{\mathcal{O}_\nu}/N} \circ [i_\nu] \right) ([z]_{N_\nu}) = \Phi_{[m]_N}^{P_{\mathcal{O}_\nu}, M_{\mathcal{O}_\nu}/N} ([z]_N) = [m \cdot z]_N.$$

Similarly,

$$\left([i_\nu] \circ \Phi_{[m]_{N_\nu}}^{P_\nu, M_\nu/N_\nu} \right) ([z]_{N_\nu}) = [i_\nu] ([m \cdot z]_{N_\nu}) = [m \cdot z]_N.$$

This establishes the commutative diagram.

Since the M_ν/N_ν-action on P_ν has an associated momentum map $\mathbf{J}_\nu : P_\nu \to (\mathfrak{m}_\nu/\mathfrak{n}_\nu)^*$ and $\beta : M_\nu/N_\nu \to M_{\mathcal{O}_\nu}/N$ is a Lie group isomorphism compatible with the symplectomorphism $[i_\nu] : P_\nu \to P_{\mathcal{O}_\nu}$, the map

$$\mathbf{J}_{\mathcal{O}_\nu} := (T_e \beta^{-1})^* \circ \mathbf{J}_\nu \circ [i_\nu]^{-1} : P_{\mathcal{O}_\nu} \longrightarrow (\mathfrak{m}_{\mathcal{O}_\nu}/\mathfrak{n})^*$$

is a momentum map of the symplectic $M_{\mathcal{O}_\nu}/N$-action on $P_{\mathcal{O}_\nu}$. Using (5.2.7) we can write $\mathbf{J}_{\mathcal{O}_\nu}$ explicitly as

$$\langle \mathbf{J}_{\mathcal{O}_\nu}(\pi_{\mathcal{O}_\nu}(z)), T_e r_{\mathcal{O}_\nu}(\xi) \rangle = \langle \mathbf{J}_M(z), \mathrm{Ad}_n \xi \rangle - \langle \overline{\nu}, \xi \rangle, \qquad (5.3.2)$$

where $\pi_{\mathcal{O}_\nu} : \mathbf{J}^{-1}(\mathcal{O}_\nu) \to P_{\mathcal{O}_\nu}$ and $r_{\mathcal{O}_\nu} : M_{\mathcal{O}_\nu} \to M_{\mathcal{O}_\nu}/N$ are projections, ξ is an arbitrary element of $\mathfrak{m}_\nu$, $\overline{\nu}$ is an arbitrary linear extension of $\nu \in \mathfrak{n}^*$ to $\mathfrak{m}$, and $n \in N$ is some element such that $\mathbf{J}_N(z) = \mathrm{Ad}^*_{n^{-1}} \nu$. Even though the element ξ was taken in $\mathfrak{m}_\nu$, expression (5.3.2) completely determines $\mathbf{J}_{\mathcal{O}_\nu}(\pi_{\mathcal{O}_\nu}(z))$ because by Lemma 5.3.2 $\beta \circ r_\nu = r_{\mathcal{O}_\nu}|_{M_\nu}$ and hence

$$T_e r_{\mathcal{O}_\nu}(\mathfrak{m}_\nu) = T_e(\beta \circ r_\nu)(\mathfrak{m}_\nu) T_e r_{\mathcal{O}_\nu}(\mathfrak{m}_{\mathcal{O}_\nu})$$

since r_ν is onto and β is an isomorphism.

A strategy identical to that followed in Theorem 5.2.9 proves the following statement.

5.3.4 Theorem (Orbit Reduction by Stages Theorem). *If the orbit $\mathcal{O}_\sigma \subset \mathfrak{m}^*$ satisfies the stages hypothesis then there is a symplectic diffeomorphism between $\mathbf{J}_M^{-1}(\mathcal{O}_\sigma)/M$ and $\mathbf{J}_{\mathcal{O}_\nu}^{-1}(\mathcal{O}_\rho)/H_{\mathcal{O}_\nu}$, where $\nu = \sigma|_\mathfrak{n}$, for some $\sigma \in \mathcal{O}_\sigma$, and $\rho \in (\mathfrak{m}_{\mathcal{O}_\nu}/\mathfrak{n})^*$ is defined by the relation*

$$\langle \rho, T_e r_{\mathcal{O}_\nu(\xi)} \rangle = \langle \sigma, \xi \rangle - \langle \overline{\nu}, \xi \rangle,$$

for every $\xi \in \mathfrak{m}_\nu$.

Orbit Reduction by Stages and the Shifting Trick.   The shifting trick that we review below replaces symplectic orbit reduction at $\mathcal{O}_\mu$ by reduction at zero for the symplectic manifold $(P \times \mathcal{O}_\sigma, \Omega - \omega^+_{\mathcal{O}_\sigma})$. The canonical diagonal action of $P \times \mathcal{O}_\sigma$ has a momentum map $\mathcal{J}_M : P \times \mathcal{O}_\sigma \to \mathfrak{g}^*$ given by $\mathcal{J}_M(z, \nu) = \mathbf{J}_M(z) - \nu$. Let $(M \times \mathcal{O}_\sigma)_0$ be the symplectic point reduced space at zero. For the proof of the following classical theorem see e.g. [HRed].

5.3.5 Theorem (Shifting Theorem). *Under the hypotheses of the Symplectic Orbit Reduction Theorem 1.2.4, the symplectic orbit reduced space $P_{\mathcal{O}_\sigma}$, the point reduced space P_σ, and $(P \times \mathcal{O}_\sigma)_0$ are symplectically diffeomorphic.*

If the group M contains a closed normal subgroup N, the Stages Theorem 5.2.9 guarantees that the reduced space $(P \times \mathcal{O}_\sigma)_0 \simeq P_{\mathcal{O}_\sigma}$ can be *always* obtained by a two step reduction process since the stages hypothesis is always verified for the momentum value $\sigma = 0$. However, it should be noticed that the reduced spaces involved in the two steps reduction of $(P \times \mathcal{O}_\sigma)_0$ and $P_{\mathcal{O}_\sigma}$ are *not* the same. Indeed, the first reduced space in the reduction of $(P \times \mathcal{O}_\sigma)_0$ is $\mathcal{J}_N^{-1}(0)/N$, where $\mathcal{J}_N := i^* \mathcal{J}_M$ which is, in general, not symplectomorphic to $\mathbf{J}_N^{-1}(0)/N$.

6

Group Extensions and the Stages Hypothesis

As was discussed in the general setting of reduction by stages, we consider a Lie group M with a normal subgroup N; recall that the goal is to reduce the action of M in two stages, the first stage being reduction by N. The goal of this chapter is to introduce hypotheses under which reduction by stages works—that is, the stages hypothesis (see Definition 5.2.8) is automatically satisfied. The actual reduction by stages procedure for these examples will be carried out in Chapters 8, 9, and 10.

The first class of examples considered in this chapter are Lie group extensions whose underlying manifold is the product of the given Lie group with an Abelian Lie group that extends it. §6.1 reviews the theory for such group extensions, which may be viewed as semidirect products with cocycles. Many formulas are derived that will be useful later when studying the coadjoint orbits of central (Chapter 9) and Abelian extensions with cocycles (Chapter 10) of a given Lie group by appealing to the reduction by stages procedure. Due to its importance, §6.2 presents the special but very important case of *central* extensions by particularizing the statements and formulas derived in §6.1. This section also studies for the first time, the Lie-Poisson structure determined by a centrally extended Lie group by an Abelian Lie group with a cocycle. This is done here directly and will be revisited again in Chapters 8, 9, and 10 from the point of view of reduction by stages. With the theory of extensions in hand, the main result of §6.3 is the statement that Lie groups that are extensions by an Abelian Lie group with a cocycle automatically satisfy the stages hypothesis I given in Definition 5.2.8 so that the main reduction by stages theorem (Theorem 5.2.9) will apply to these cases.

A second class of examples is treated in §6.4, which proves that stages hypothesis I holds for general semidirect products of two Lie groups, neither of which need be Abelian. We emphasize that this is not a special case of the previous section, since the Lie group that extends the given one is not necessarily Abelian. Although our main examples are covered by these results, for completeness, we treat more general group extensions in Chapter 10.

6.1   Lie Group and Lie Algebra Extensions

The reduction by stages theorem provides, for example, a method to classify coadjoint orbits of Lie group extensions (see e.g., Cartan and Eilenberg [1956]) as we shall see later in concrete examples. We shall not present here the theory in its most general form and will review only semidirect products with cocycles. This is a large class of extensions that includes both semidirect products and central extensions. We shall also derive many explicit formulas useful in various computations throughout this book. For more general situations and many examples see Neeb [2002], Guieu and Roger [2003], Maier and Neeb [2003], and Neeb [2004a]. We shall also derive many explicit formulas useful in the reduction by stages computations in Chapters 8, 9, and 10.

Generalities on Group Extensions.   Let G be a (possibly infinite dimensional) Lie group and let A be an Abelian Lie group. It will be convenient to use multiplicative notation in G and additive notation in A. Let $\mathrm{Aut}(A)$ denote the group of automorphisms of A under composition of automorphisms. Consider an extension $\widehat{G}$ of G by A, that is, an exact sequence

$$0 \to A \to \widehat{G} \to G \to 1. \tag{6.1.1}$$

In this section we shall assume that the manifold underlying $\widehat{G}$ equals $G \times A$. Thus, the elements of $\widehat{G}$ are pairs (g, α) with $g \in G$ and $\alpha \in A$. We assume that the group multiplication is given by

$$(g, \alpha)(h, \beta) = (gh, \alpha + \phi(g)(\beta) + B(g, h)), \tag{6.1.2}$$

where $\phi : G \times A \to A$ is a smooth action such that $\phi(g) : A \to A$ is a smooth Abelian group homomorphism for every $g \in G$ and $B : G \times G \to A$ is an A-**valued group 2-cocycle**; that is, the **cocycle condition**

$$\phi(f)(B(g, h)) + B(f, gh) = B(f, g) + B(fg, h) \tag{6.1.3}$$

holds for any f, g, $h \in G$. This relation is equivalent to the associativity of the multiplication (6.1.2).

The following relations are a direct consequence of (6.1.3):

$$B(e, g) = B(e, e), \tag{6.1.4}$$

$$B(g, e) = \phi(g)(B(e, e)), \tag{6.1.5}$$

$$\phi(g^{-1})(B(g, g^{-1})) + B(g^{-1}, e) = B(g^{-1}, g) + B(e, g^{-1}), \tag{6.1.6}$$

for all $g \in G$. A direct verification, using (6.1.4) and (6.1.5), shows that $(e, -B(e, e))$ is the neutral element for the multiplication (6.1.2). From (6.1.4) and (6.1.6) it follows that every $(g, \alpha) \in G \times A$ has a unique inverse given by

$$\begin{aligned}
(g, \alpha)^{-1} &= \left(g^{-1}, -\phi(g^{-1})(B(e, e) + B(g, g^{-1}) + \alpha)\right) \\
&= \left(g^{-1}, -B(e, e) - B(g^{-1}, g) - \phi(g^{-1})(\alpha)\right). \tag{6.1.7}
\end{aligned}$$

Thus $G \times A$ is a group relative to the multiplication (6.1.2). Assuming that the maps $B : G \times G \to A$ and $(g, \alpha) \in G \times A \mapsto \phi(g)(\alpha) \in A$ are smooth, the multiplication (6.1.2) and inversion (6.1.7) are also smooth and $G \times A$ is therefore a Lie group which shall be denoted by $G \times_{\phi, B} A$.

Two special cases, namely semidirect products and central extensions, are noteworthy. In §4.2 we dealt with the case in which A is a vector space V and $B = 0$, which corresponds to the semidirect product $G \times_\phi V$. If B is arbitrary and $\phi(g) = \mathrm{id}$ for all $g \in G$, that is, there is no action of G on A, then the extension $G \times_{\mathrm{id}, B} A$ is central.

The formulas for the neutral element and the inverse in the group extension $G \times_{\phi, B} A$ can be simplified by constructing an isomorphic group from a *normalized cocycle* with coefficients in A (see Cartan and Eilenberg [1956]), that is, a map $B : G \times G \to A$ that, in addition to (6.1.3), satisfies also the identities

$$B(g, e) = B(e, g) = 0, \quad \text{for all} \quad g \in G. \tag{6.1.8}$$

To do this, given the cocycle B satisfying only (6.1.3), define

$$\bar{B}(g, h) := B(g, h) - B(g, e)$$

and verify, using (6.1.3) for B, that $\bar{B}$ also satisfies the cocycle identity (6.1.3) relative to ϕ. Note that $\bar{B}(e, e) = 0$ which, in view of (6.1.4) and (6.1.5), implies that $\bar{B}(g, e) = \bar{B}(e, g) = 0$. Moreover, the map

$$(g, \alpha) \in G \times_{\phi, B} A \mapsto (g, \alpha + B(e, e)) \in G \times_{\phi, \bar{B}} A$$

is easily seen to be a Lie group isomorphism.

This argument shows that *we can assume without loss of generality that the cocycle B is normalized*, that is, it satisfies the identities (6.1.3) and (6.1.8). Thus, the neutral element of $G \times_{\phi, B} A$ is $(e, 0)$, the inverse of (g, α) is

$$\begin{aligned}
(g, \alpha)^{-1} &= \left(g^{-1}, -\phi(g^{-1})(B(g, g^{-1}) + \alpha)\right) \\
&= \left(g^{-1}, -B(g^{-1}, g) - \phi(g^{-1})(\alpha)\right), \tag{6.1.9}
\end{aligned}$$

and the inner automorphisms on $G \times_{\phi,B} V$ are given by

$$I_{(g,\alpha)}(h,\beta) = \big(ghg^{-1}, -\phi(g)\phi(h)\phi(g^{-1})(\alpha + B(g,g^{-1})) + \phi(g)B(h,g^{-1})$$
$$+\phi(g)\beta + B(g,hg^{-1}) + \alpha\big). \qquad (6.1.10)$$

If B is a normalized two-cocycle, we shall also use the identity

$$\phi(g^{-1})(B(g,g^{-1})) = B(g^{-1},g) \qquad (6.1.11)$$

which is a consequence of (6.1.6) and (6.1.8).

From now on we shall always assume that B is normalized and therefore use (6.1.9), (6.1.10), and (6.1.11) in subsequent computations.

The Associated Group One-Cocycle on $G \times_{\phi,B} A$. Let $\mathfrak{g}$ and $\mathfrak{a}$ be the Lie algebras of G and A, respectively. Denote by $L(\mathfrak{g},\mathfrak{a})$ the vector space of linear maps from $\mathfrak{g}$ to $\mathfrak{a}$. The A-valued group two-cocycle $B : G \times G \to A$ induces two $L(\mathfrak{g},\mathfrak{a})$-valued group one-cocycles $\widehat{\sigma}_R, \widehat{\sigma}_L : G \times_{\phi,B} A \to L(\mathfrak{g},\mathfrak{a})$ which we now describe.

We shall use the following notation

$$D_1 B(g,h)(u_g) := \left.\frac{d}{dt}\right|_{t=0} B(g(t),h) \text{ and } D_2 B(g,h)(v_h) := \left.\frac{d}{dt}\right|_{t=0} B(g,h(t)),$$

where $g(t)$ and $h(t)$ are smooth curves in G satisfying $g(0) = g$, $h(0) = h$, $\dot{g}(0) \in u_g \in T_g G$, and $\dot{h}(0) = v_h \in T_h G$. Thus $D_1 B(g,h) : T_g G \to T_{B(g,h)} A$ and $D_2 B(g,h) : T_h G \to T_{B(g,h)} A$ depend smoothly on h and g respectively. In particular, since B is a normalized two-cocycle, we obtain the linear maps $D_1 B(e,h) : \mathfrak{g} \to \mathfrak{a}$ and $D_2 B(g,e) : \mathfrak{g} \to \mathfrak{a}$.

Let $\Lambda_\alpha : A \to A$ denote the translation by $\alpha \in A$ in the Abelian Lie group A. There are two natural commuting right G-representations on $L(\mathfrak{g},\mathfrak{a})$:

- $(S \cdot_1 g)(\xi) := S(\mathrm{Ad}_g \xi)$ and

- $(S \cdot_2 g)(\xi) := T_0 \phi(g^{-1})(S(\xi))$.

where $S \in L(\mathfrak{g},\mathfrak{a})$ and $g \in G$. Therefore,

$$(S \cdot g)(\xi) := T_0 \phi(g^{-1})(S(\mathrm{Ad}_g \xi)) \qquad (6.1.12)$$

also defines a right G-representation on $L(\mathfrak{g},\mathfrak{a})$ which will be used in the next proposition. We shall denote by $g \cdot S := S \cdot g^{-1}$ the associated left action of G on $L(\mathfrak{g},\mathfrak{a})$.

Let $\Upsilon \in \Omega^1(A;\mathfrak{a})$ be the **_Maurer-Cartan one-form_** on A, that is,

$$\Upsilon(\alpha)(v_\alpha) := T_\alpha \Lambda_{-\alpha}(v_\alpha) \in \mathfrak{a} \quad \text{for any} \quad v_\alpha \in T_\alpha A. \qquad (6.1.13)$$

In what follows we shall need three formulas that involve the Maurer-Cartan form and the infinitesimal generator of the action $\phi : G \times A \to A$.

Recall the general formula

$$(\phi(g)_*\xi_A)(\alpha) := T_{\phi(g^{-1})(\alpha)}\phi(g)\left(\xi_A(\phi(g^{-1})(\alpha))\right)$$
$$= (\mathrm{Ad}_g\,\xi)_A(\alpha) \tag{6.1.14}$$

valid for any left action (see e.g. [MandS], [LonM], or [HRed]).

The first formula is

$$\langle \Upsilon, \xi_A\rangle(\alpha + \beta) = \langle \Upsilon, \xi_A(\alpha)\rangle(\alpha) + \langle \Upsilon, \xi_A\rangle(\beta) \tag{6.1.15}$$

for any $\alpha, \beta \in A$ and $\xi \in \mathfrak{g}$. To prove it note that we have

$$\xi_A(\alpha + \beta) = T_\alpha\Lambda_\beta\left(\xi_A(\alpha)\right) + T_\beta\Lambda_\alpha\left(\xi_A(\beta)\right)$$

so that

$$\langle \Upsilon, \xi_A\rangle(\alpha + \beta) = T_{\alpha+\beta}\Lambda_{-\alpha-\beta}\left(\xi_A(\alpha + \beta)\right)$$
$$= T_\alpha\Lambda_{-\alpha}\left(\xi_A(\alpha)\right) + T_\beta\Lambda_{-\beta}\left(\xi_A(\beta)\right)$$
$$= \langle \Upsilon, \xi_A\rangle(\alpha) + \langle \Upsilon, \xi_A\rangle(\beta)$$

as required.

The second formula is a consequence of the identity (6.1.15):

$$\langle \Upsilon, \xi_A\rangle(-\alpha) = -\langle \Upsilon, \xi_A\rangle(\alpha) \tag{6.1.16}$$

for any $\alpha \in A$ and $\xi \in \mathfrak{g}$. This follows from $\xi_A(0) = 0$ for any $\xi \in \mathfrak{g}$ which in turn is a consequence of the fact that the action ϕ is by group homomorphisms and thus $\phi(g)(0) = 0$ for any $g \in G$.

The third formula is

$$T_0\phi(g)\langle \Upsilon, \xi_A\rangle\left(\phi(g^{-1})(\alpha)\right) = \langle \Upsilon, \phi(g)_*\xi_A\rangle(\alpha)$$
$$= \langle \Upsilon, (\mathrm{Ad}_g\,\xi)_A\rangle(\alpha) \tag{6.1.17}$$

for any $g \in G$, $\xi \in \mathfrak{g}$, and $\alpha \in A$. The second equality is a direct consequence of (6.1.14). To prove the first, use the relation $\phi(g) \circ \Lambda_{-\phi(g^{-1})(\alpha)} = \Lambda_{-\alpha} \circ \phi(g)$ in the third equality below to get

$$T_0\phi(g)\langle \Upsilon, \xi_A\rangle\left(\phi(g^{-1})(\alpha)\right)$$
$$= T_0\phi(g)T_{\phi(g^{-1})(\alpha)}\Lambda_{-\phi(g^{-1})(\alpha)}\xi_A\left(\phi(g^{-1})(\alpha)\right)$$
$$= T_{\phi(g^{-1})(\alpha)}\left(\phi(g) \circ \Lambda_{-\phi(g^{-1})(\alpha)}\right)\xi_A\left(\phi(g^{-1})(\alpha)\right)$$
$$= T_{\phi(g^{-1})(\alpha)}\left(\Lambda_{-\alpha} \circ \phi(g)\right)\xi_A\left(\phi(g^{-1})(\alpha)\right)$$
$$= T_\alpha\Lambda_{-\alpha}T_{\phi(g^{-1})(\alpha)}\phi(g)\xi_A\left(\phi(g^{-1})(\alpha)\right)$$
$$= T_\alpha\Lambda_{-\alpha}\left(\phi(g)_*\xi_A\right)(\alpha)$$
$$= \langle \Upsilon, \phi(g)_*\xi_A\rangle(\alpha)$$

as required.

6.1.1 Proposition.   (i) *The smooth map $\sigma_R : G \to L(\mathfrak{g}, \mathfrak{a})$ defined by any of the three formulas below*

$$
\sigma_R(g)(\xi) := \frac{d}{dt}\bigg|_{t=0} \left[\phi(g^{-1}) \left(B(g, \exp(t\xi)g^{-1}) \right) - B(\exp(t\xi)g^{-1}, g) \right]
$$

$$
= T_{B(g^{-1},g)}\Lambda_{-B(g^{-1},g)} \left[T_{B(g,g^{-1})}\phi(g^{-1}) \circ D_2 B(g, g^{-1}) - D_1 B(g^{-1}, g) \right]
$$
$$
\left(T_e R_{g^{-1}} \xi \right)
$$

$$
= \left(T_0 \phi(g^{-1}) \circ T_{B(g,g^{-1})}\Lambda_{-B(g,g^{-1})} \circ D_2 B(g, g^{-1}) \circ T_e R_{g^{-1}} \right)(\xi)
$$
$$
- \left(T_{B(g^{-1},g)}\Lambda_{-B(g^{-1},g)} \circ D_1 B(g^{-1}, g) \circ T_e R_{g^{-1}} \right)(\xi) \qquad (6.1.18)
$$

satisfies the following identity

$$
\sigma_R(gh)(\xi) = \left(\sigma_R(g) \cdot h + \sigma_R(h) \right)(\xi)
$$
$$
+ \frac{d}{dt}\bigg|_{t=0} \left[\phi(\exp(t\xi)h^{-1}g^{-1})(B(g,h)) - \phi(h^{-1}g^{-1})(B(g,h)) \right]
$$
$$
= \left(\sigma_R(g) \cdot h + \sigma_R(h) \right)(\xi) + \langle \Upsilon, \xi_A \rangle \left(\varphi(h^{-1}g^{-1})(B(g,h)) \right) \qquad (6.1.19)
$$

for all $g, h \in G$ and $\xi \in \mathfrak{g}$. An alternative definition of σ_R is given by

$$
\sigma_R(g)(\xi) = \frac{d}{dt}\bigg|_{t=0} \left[-\phi(\exp(t\xi)g^{-1})(B(g, g^{-1})) + \phi(g^{-1})(B(g, \exp(t\xi)) \right.
$$
$$
\left. + \phi(g^{-1})(B(g\exp(t\xi), g^{-1})) \right]. \qquad (6.1.20)
$$

The smooth map $\widehat{\sigma}_R : G \times_{\phi,B} A \to L(\mathfrak{g}, \mathfrak{a})$ defined for any $\xi \in \mathfrak{g}$ by

$$
\widehat{\sigma}_R(g, \alpha)(\xi) := \sigma_R(g)(\xi) - \langle \Upsilon, \xi_A \rangle \left(\phi(g^{-1})(\alpha) \right) \qquad (6.1.21)
$$

defines a $L(\mathfrak{g}, \mathfrak{a})$-valued right one-cocycle on $G \times_{\phi,B} A$, that is, the following identity holds

$$
\widehat{\sigma}_R((g, \alpha)(h, \beta)) = \widehat{\sigma}_R(h, \beta) + \widehat{\sigma}_R(g, \alpha) \cdot (h, \beta)
$$

for any $(g, \alpha), (h, \beta) \in G \times_{\phi,B} A$, where the right action of $G \times_{\phi,B} A$ on $L(\mathfrak{g}, \mathfrak{a})$ is given by

$$
S \cdot (h, \beta) := S \cdot h = T_0 \phi(h^{-1}) \circ S \circ \mathrm{Ad}_h \qquad (6.1.22)
$$

for any $(h, \beta) \in G \times_{\phi,B} A$ and any $S \in L(\mathfrak{g}, \mathfrak{a})$. In particular, $\sigma_R(e) = 0$.
   (ii) *The smooth map $\sigma_L(g) := \sigma_R(g^{-1})$ satisfies*

$$
\sigma_L(gh)(\xi) = \left(\sigma_L(g) + g \cdot \sigma_L(h) \right)(\xi)
$$
$$
+ \frac{d}{dt}\bigg|_{t=0} \left[\phi(\exp(t\xi)gh)(B(h^{-1}, g^{-1})) - \phi(gh)(B(h^{-1}, g^{-1})) \right]
$$
$$
= \left(\sigma_L(g) + g \cdot \sigma_L(h) \right)(\xi) + \langle \Upsilon, \xi_A \rangle \left(\phi(gh)(B(h^{-1}, g^{-1})) \right) \qquad (6.1.23)
$$

*for all $g, h \in G$ and $\xi \in \mathfrak{g}$. The smooth map $\widehat{\sigma}_L : G \times_{\phi, B} A \to L(\mathfrak{g}, \mathfrak{a})$
defined for any $\xi \in \mathfrak{g}$ by*

$$\widehat{\sigma}_L(g, \alpha) := \sigma_L(g) + \langle \Upsilon, \xi_A \rangle \, (\alpha + B(g, g^{-1})) \qquad (6.1.24)$$

*defines a $L(\mathfrak{g}, \mathfrak{a})$-valued left one-cocycle on $G \times_{\phi, B} A$, that is, the following
identity holds*

$$\widehat{\sigma}_L((g, \alpha)(h, \beta)) = \widehat{\sigma}_L(g, \alpha) + (g, \alpha) \cdot \widehat{\sigma}_L(h, \beta)$$

*for any $(g, \alpha), (h, \beta) \in G \times_{\phi, B} A$, where the left action of $G \times_{\phi, B} A$ on
$L(\mathfrak{g}, \mathfrak{a})$ is given by*

$$(g, \alpha) \cdot S := g \cdot S = T_0 \phi(g) \circ S \circ \mathrm{Ad}_{g^{-1}} \qquad (6.1.25)$$

*for any $(g, \alpha) \in G \times_{\phi, B} A$ and any $S \in L(\mathfrak{g}, \mathfrak{a})$. In particular, $\sigma_L(e) = 0$.
Moreover, $\widehat{\sigma}_L(g, \alpha) = \widehat{\sigma}_R((g, \alpha)^{-1})$ for any $(g, \alpha) \in G \times_{\phi, B} A$.*
*(iii) The relations between σ_R and σ_L and the one-cocycles $\widehat{\sigma}_R$ and $\widehat{\sigma}_L$ are
given, respectively, by*

$$- (\sigma_L(g) \cdot g)(\xi) - \langle \Upsilon, \xi_A \rangle \left(B(g^{-1}, g) \right) = \sigma_R(g)(\xi) \qquad (6.1.26)$$

and

$$\widehat{\sigma}_L(g, \alpha) \cdot g = -\widehat{\sigma}_R(g, \alpha) \qquad (6.1.27)$$

for any $g \in G$, $\alpha \in A$, and $\xi \in \mathfrak{g}$.

Proof.   (i) The first equality in (6.1.18) is obtained in the following way.
First note that $t \in \mathbb{R} \mapsto \phi(g^{-1}) \left(B(g, \exp(t\xi)g^{-1}) \right) - B(\exp(t\xi)g^{-1}, g) \in A$
is a curve in A which at $t = 0$ passes through $0 \in A$ due to the identity
(6.1.11). Second, we have

$$\left. \frac{d}{dt} \right|_{t=0} \left[-B(\exp(t\xi)g^{-1}, g) \right]$$
$$= - \left(T_0 \Lambda_{-B(g^{-1}, g)} \circ T_{B(g^{-1}, g)} \Lambda_{-B(g^{-1}, g)} \circ D_1 B(g^{-1}, g) \circ T_e R_{g^{-1}} \right) (\xi).$$

Third, using this formula, the Leibniz identity, (6.1.11), and the chain rule we get

$$
\begin{aligned}
\frac{d}{dt}\bigg|_{t=0} & \left[\phi(g^{-1})\left(B(g,\exp(t\xi)g^{-1})\right) - B(\exp(t\xi)g^{-1},g)\right] \\
= & \left(T_{\phi(g^{-1})(B(g,g^{-1}))}\Lambda_{-B(g^{-1},g)} \circ T_{B(g,g^{-1})}\phi(g^{-1}) \circ D_2 B(g,g^{-1})\right. \\
& \qquad \left. \circ T_e R_{g^{-1}}\right)(\xi) \\
& - \left(T_{-B(g^{-1},g)}\Lambda_{\phi(g^{-1})(B(g,g^{-1}))} \circ T_0\Lambda_{-B(g^{-1},g)} \circ T_{B(g^{-1},g)}\Lambda_{-B(g^{-1},g)}\right. \\
& \qquad \left. \circ D_1 B(g^{-1},g) \circ T_e R_{g^{-1}}\right)(\xi) \\
= & \left(T_{B(g^{-1},g)}\Lambda_{-B(g^{-1},g)} \circ T_{B(g,g^{-1})}\phi(g^{-1}) \circ D_2 B(g,g^{-1}) \circ T_e R_{g^{-1}}\right)(\xi) \\
& - \left(T_{B(g^{-1},g)}\left(\Lambda_{B(g^{-1},g)} \circ \Lambda_{-B(g^{-1},g)} \circ \Lambda_{-B(g^{-1},g)}\right)\right. \\
& \qquad \left. \circ D_1 B(g^{-1},g) \circ T_e R_{g^{-1}}\right)(\xi) \\
= & \; T_{B(g^{-1},g)}\Lambda_{-B(g^{-1},g)}\left[T_{B(g,g^{-1})}\phi(g^{-1}) \circ D_2 B(g,g^{-1}) - D_1 B(g^{-1},g)\right] \\
& \left(T_e R_{g^{-1}}(\xi)\right)
\end{aligned}
$$

which proves the first equality in (6.1.18).

To prove the second equality in (6.1.18) use the identity $\Lambda_{-\alpha} \circ \phi(g^{-1}) = \phi(g^{-1}) \circ \Lambda_{-\phi(g)(\alpha)}$ whose derivative at $B(g,g^{-1})$ for $\alpha = B(g^{-1},g)$ is

$$
T_{B(g^{-1},g)}\Lambda_{-B(g^{-1},g)} \circ T_{B(g,g^{-1})}\phi(g^{-1}) = T_0\phi(g^{-1}) \circ T_{B(g,g^{-1})}\Lambda_{-B(g,g^{-1})}
$$

by using the relation $\phi(g^{-1})(B(g,g^{-1})) = B(g^{-1},g)$.

To prove the identity (6.1.19) we work on the terms in $\sigma_R(gh)(\xi)$ before taking the t-derivative and apply (6.1.3) to the first summand for $f \mapsto g, g \mapsto h, h \mapsto \exp(t\xi)h^{-1}g^{-1}$. We get

$$
\begin{aligned}
\phi(h^{-1}g^{-1})&(B(gh,\exp(t\xi)h^{-1}g^{-1})) - B(\exp(t\xi)h^{-1}g^{-1},gh) \\
& = \phi(h^{-1}g^{-1})\left[\phi(g)(B(h,\exp(t\xi)h^{-1}g^{-1}))\right. \\
& \qquad \left. + B(g,h\exp(t\xi)h^{-1}g^{-1}) - B(g,h)\right] - B(\exp(t\xi)h^{-1}g^{-1},gh) \\
& = \phi(h^{-1})(B(h,\exp(t\xi)h^{-1}g^{-1})) + \phi(h^{-1}g^{-1})(B(g,\exp(t\,\mathrm{Ad}_h\,\xi)g^{-1})) \\
& \qquad - \phi(h^{-1}g^{-1})(B(g,h)) - B(\exp(t\xi)h^{-1}g^{-1},gh).
\end{aligned}
$$

Next, add and subtract the term $\phi(h^{-1})B\left(\exp(t\,\mathrm{Ad}_h\,\xi)g^{-1},g\right)$ and use (6.1.3) on the third and fourth summands below with the choice $f \mapsto$

$h, g \mapsto \exp(t\xi)h^{-1}g^{-1}, h \mapsto g$ to get

$$\phi(h^{-1})\phi(g^{-1})\left(B\left(g, \exp(t\operatorname{Ad}_h \xi)g^{-1}\right)\right) - \phi(h^{-1})\left(B\left(\exp(t\operatorname{Ad}_h \xi)g^{-1}, g\right)\right)$$
$$+ \phi(h^{-1})\left(B\left(h\exp(t\xi)h^{-1}g^{-1}, g\right)\right) + \phi(h^{-1})\left(B(h, \exp(t\xi)h^{-1}g^{-1})\right)$$
$$- \phi(h^{-1}g^{-1})(B(g,h)) - B(\exp(t\xi)h^{-1}g^{-1}, gh)$$
$$= \phi(h^{-1})\phi(g^{-1})\left(B\left(g, \exp(t\operatorname{Ad}_h \xi)g^{-1}\right)\right) - \phi(h^{-1})\left(B\left(\exp(t\operatorname{Ad}_h \xi)g^{-1}, g\right)\right)$$
$$+ \phi(h^{-1})\left[\phi(h)(B(\exp(t\xi)h^{-1}g^{-1}, g)) + B(h, \exp(t\xi)h^{-1})\right]$$
$$- \phi(h^{-1}g^{-1})(B(g,h)) - B(\exp(t\xi)h^{-1}g^{-1}, gh)$$
$$= \phi(h^{-1})\phi(g^{-1})\left(B\left(g, \exp(t\operatorname{Ad}_h \xi)g^{-1}\right)\right) - \phi(h^{-1})\left(B\left(\exp(t\operatorname{Ad}_h \xi)g^{-1}, g\right)\right)$$
$$+ B(\exp(t\xi)h^{-1}g^{-1}, g) + \phi(h^{-1})(B(h, \exp(t\xi)h^{-1}))$$
$$- \phi(h^{-1}g^{-1})(B(g,h)) - B(\exp(t\xi)h^{-1}g^{-1}, gh).$$

Now subtract and add the term $B(\exp(t\xi)h^{-1}, h)$ and use again (6.1.3) on the fifth, sixth, and seventh summands below for $f \mapsto \exp(t\xi)h^{-1}g^{-1}, g \mapsto g, h \mapsto h$. Then the expression above becomes

$$\phi(h^{-1})\phi(g^{-1})\left(B\left(g, \exp(t\operatorname{Ad}_h \xi)g^{-1}\right)\right) - \phi(h^{-1})\left(B\left(\exp(t\operatorname{Ad}_h \xi)g^{-1}, g\right)\right)$$
$$+ \phi(h^{-1})\left(B(h, \exp(t\xi)h^{-1})\right) - B(\exp(t\xi)h^{-1}, h)$$
$$+ B(\exp(t\xi)h^{-1}, h) + B(\exp(t\xi)h^{-1}g^{-1}, g) - B(\exp(t\xi)h^{-1}g^{-1}, gh)$$
$$- \phi(h^{-1}g^{-1})(B(g,h))$$
$$= \phi(h^{-1})\phi(g^{-1})\left(B\left(g, \exp(t\operatorname{Ad}_h \xi)g^{-1}\right)\right) - \phi(h^{-1})\left(B\left(\exp(t\operatorname{Ad}_h \xi)g^{-1}, g\right)\right)$$
$$+ \phi(h^{-1})\left(B(h, \exp(t\xi)h^{-1})\right) - B(\exp(t\xi)h^{-1}, h)$$
$$+ \phi(\exp(t\xi)h^{-1}g^{-1})(B(g,h)) - \phi(h^{-1}g^{-1})(B(g,h)).$$

Taking the t-derivative of this expression at $t = 0$ and taking into account the definitions (6.1.18) of σ_R and of the G-action (6.1.12) on $L(\mathfrak{g}, \mathfrak{a})$ we get the first equality in (6.1.19). The second equality is a direct consequence of the definition of the Maurer-Cartan one-form Υ on A and of the infinitesimal generator ξ_A.

To prove (6.1.20) we use (6.1.3) in the first summand below for $f \mapsto \exp(t\xi)g^{-1}, g \mapsto g, h \mapsto g^{-1}$ and take into account that $B(k, e) = B(e, k') = 0$ for any $k, k' \in G$ to get

$$-\phi(\exp(t\xi)g^{-1})(B(g, g^{-1})) + \phi(g^{-1})(B(g, \exp(t\xi)))$$
$$+ \phi(g^{-1})(B(g\exp(t\xi), g^{-1}))$$
$$= -B(\exp(t\xi)g^{-1}, g) - B(\exp(t\xi), g^{-1})$$
$$+ \phi(g^{-1})\left[B(g, \exp(t\xi)) + B(g\exp(t\xi), g^{-1})\right].$$

Now apply again (6.1.3) to the two terms in the corner bracket for $f \mapsto g, g \mapsto \exp(t\xi), h \mapsto g^{-1}$ to get

$$- B(\exp(t\xi)g^{-1}, g) - B(\exp(t\xi), g^{-1})$$
$$+ \phi(g^{-1}) \left[\phi(g)(B(\exp(t\xi), g^{-1})) + B(g, \exp(t\xi)g^{-1}) \right]$$
$$= \phi(g^{-1})B(g, \exp(t\xi)g^{-1}) - B(\exp(t\xi)g^{-1}, g)$$

which is (6.1.18).

Now we prove that $\widehat{\sigma}_R : G \to L(\mathfrak{g}, \mathfrak{a})$ given by (6.1.21) is a smooth group one-cocycle. Let $\xi \in \mathfrak{g}$, $(g, \alpha), (h, \beta) \in G \times_{\phi, B} A$. We have by (6.1.18)

$$\widehat{\sigma}_R((g, \alpha)(h, \beta))(\xi) = \widehat{\sigma}_R(gh, \alpha + \phi(g)(\beta) + B(g, h))(\xi)$$
$$= \sigma_R(gh)(\xi) - \langle \Upsilon, \xi_A \rangle \left(\phi(h^{-1}g^{-1})(\alpha + \phi(g)(\beta) + B(g, h)) \right)$$
$$= \sigma_R(h)(\xi) + (\sigma_R(g) \cdot h)(\xi) + \langle \Upsilon, \xi_A \rangle \left(\phi(h^{-1}g^{-1})(B(g, h)) \right)$$
$$\quad - \langle \Upsilon, \xi_A \rangle \left(\phi(h^{-1}g^{-1})(\alpha + \phi(g)(\beta) + B(g, h)) \right)$$
$$= \sigma_R(h) + (\sigma_R(g) \cdot h)(\xi) - \langle \Upsilon, \xi_A \rangle \left(\phi(h^{-1}g^{-1})(\alpha + \phi(g)(\beta)) \right),$$

where we have used (6.1.15) in the last equality. Now use (6.1.15) and (6.1.17) in the last summand to get

$$\sigma_R(h)(\xi) + (\sigma_R(g) \cdot h)(\xi) - \langle \Upsilon, \xi_A \rangle \left(\phi(h^{-1}g^{-1})(\alpha) \right) - \langle \Upsilon, \xi_A \rangle \left(\phi(h^{-1})(\beta) \right)$$
$$= \sigma_R(h)(\xi) + (\sigma_R(g) \cdot h)(\xi) - T_0\phi(h^{-1}) \langle \Upsilon, \phi(h)_* \xi_A \rangle \left(\phi(h)\phi(h^{-1}g^{-1})(\alpha) \right)$$
$$\quad - \langle \Upsilon, \xi_A \rangle \left(\phi(h^{-1})(\beta) \right)$$
$$= \sigma_R(h)(\xi) - \langle \Upsilon, \xi_A \rangle \left(\phi(h^{-1})(\beta) \right) + T_0\phi(h^{-1}) \left(\sigma_R(g) \left(\mathrm{Ad}_h \xi \right) \right)$$
$$\quad - T_0\phi(h^{-1}) \langle \Upsilon, (\mathrm{Ad}_h \xi)_A \rangle \left(\phi(g^{-1})(\alpha) \right)$$
$$= \widehat{\sigma}_R(h, \beta)(\xi) + T_0\phi(h^{-1})\widehat{\sigma}_R(g, \alpha) \left(\mathrm{Ad}_h \xi \right)$$
$$= \widehat{\sigma}_R(h, \beta)(\xi) + (\widehat{\sigma}_R(g, \alpha) \cdot h)(\xi),$$

where we have used the definitions (6.1.21) for $\widehat{\sigma}_R$ and (6.1.12) for the G-action on $L(\mathfrak{g}, \mathfrak{a})$. This shows that $\widehat{\sigma}_R((g, \alpha)(h, \beta)) = \widehat{\sigma}_R(h, \beta) + \widehat{\sigma}_R(g, \alpha) \cdot (h, \beta)$ by using the definition (6.1.22) for the $(G \times_{\phi, B} A)$-action on $L(\mathfrak{g}, \mathfrak{a})$. Therefore, $\widehat{\sigma}_R : G \times_{\phi, B} A \to L(\mathfrak{g}, \mathfrak{a})$ is a right smooth group one-cocycle, as claimed.

Finally, since $\xi_A(0) = 0$, formula (6.1.21) implies that $\sigma_R(e) = \widehat{\sigma}_R(e, 0) = 0$, because $\widehat{\sigma}_R$ is a one-cocycle and hence necessarily vanishes on the neutral element of the group. Or, directly, this follows from (6.1.18) since $B(e, g) = B(g, e) = 0$ for all $g \in G$.

(ii) We have by (6.1.19)

$$
\begin{aligned}
\sigma_L(gh) = \sigma_R(h^{-1}g^{-1}) &= \left(\sigma_R(h^{-1}) \cdot g^{-1} + \sigma_R(g^{-1})\right)(\xi) \\
&+ \left.\frac{d}{dt}\right|_{t=0} \left[\phi(\exp(t\xi)gh)(B(h^{-1},g^{-1})) - \phi(gh)(B(h^{-1},g^{-1}))\right] \\
&= \left(g \cdot \sigma_L(h) + \sigma_L(g)\right)(\xi) \\
&+ \left.\frac{d}{dt}\right|_{t=0} \left[\phi(\exp(t\xi)gh)(B(h^{-1},g^{-1})) - \phi(gh)(B(h^{-1},g^{-1}))\right]
\end{aligned}
$$

which proves (6.1.23).

Since $\widehat{\sigma}_R$ is a right one-cocycle, it follows that $\widehat{\sigma}_L(g,\alpha) := \widehat{\sigma}_R((g,\alpha^{-1})$ is a left one-cocycle. Indeed,

$$
\begin{aligned}
\widehat{\sigma}_L((g,\alpha)(h,\beta)) &= \widehat{\sigma}_R((h,\beta)^{-1}(g,\alpha)^{-1}) \\
&= \widehat{\sigma}_R((g,\alpha)^{-1}) + \widehat{\sigma}_R((h,\beta^{-1}) \cdot (g,\alpha)^{-1} \\
&= \widehat{\sigma}_L(g,\alpha) + (g,\alpha) \cdot \widehat{\sigma}_L(h,\beta)
\end{aligned}
$$

which is the left one-cocycle identity.

Moreover, by (6.1.21) and (6.1.16) we get

$$
\begin{aligned}
\widehat{\sigma}_L(g,\alpha) &= \widehat{\sigma}_R((g,\alpha)^{-1}) = \widehat{\sigma}_R(g^{-1}, -\phi(g^{-1})(\alpha + B(g,g^{-1}))) \\
&= \sigma_R(g^{-1}) - \langle \Upsilon, \xi_A \rangle \left(-\phi(g)\phi(g^{-1})(\alpha + B(g,g^{-1}))\right) \\
&= \sigma_L(g) + \langle \Upsilon, \xi_A \rangle (\alpha + B(g,g^{-1})),
\end{aligned}
$$

which proves (6.1.24).

(iii) By (6.1.18) we have for any $g \in G$ and an $\xi \in \mathfrak{g}$

$$
\begin{aligned}
(\sigma_L(g) \cdot g)(\xi) &= T_0\phi(g^{-1})\sigma_R(g^{-1})(\mathrm{Ad}_g \xi) \\
&= \left.\frac{d}{dt}\right|_{t=0} \left[\phi(g^{-1})\phi(g)(B(g^{-1}, \exp(t\,\mathrm{Ad}_g \xi)g)\right. \\
&\qquad\qquad \left. -\phi(g^{-1})(B(\exp(t\,\mathrm{Ad}_g \xi)g, g^{-1}))\right] \\
&= \left.\frac{d}{dt}\right|_{t=0} \left[B(g^{-1}, g\exp(t\xi)) - \phi(g^{-1})(B(g\exp(t\xi), g^{-1}))\right].
\end{aligned}
$$

Now use (6.1.3) on the term $B(g\exp(t\xi), g^{-1})$ with $f \mapsto g, g \mapsto \exp(t\xi), h \mapsto g^{-1}$ to get

$$
\begin{aligned}
B(g\exp(t\xi), g^{-1}) = \phi(g)(B(\exp(t\xi), g^{-1})) &+ B(g, \exp(t\xi)g^{-1}) \\
&- B(g, \exp(t\xi))
\end{aligned}
$$

and hence the previous expression equals

$$
\begin{aligned}
-\left.\frac{d}{dt}\right|_{t=0} & \left[\phi(g^{-1})B(g,\exp(t\xi)g^{-1}) + B(\exp(t\xi),g^{-1})\right. \\
& \left. -\phi(g^{-1})(B(g,\exp(t\xi))) - B(g^{-1}, g\exp(t\xi))\right] \\
= -\left.\frac{d}{dt}\right|_{t=0} & \left[\phi(g^{-1})B(g,\exp(t\xi)g^{-1}) + B(\exp(t\xi),g^{-1}) - B(g^{-1},g)\right]
\end{aligned}
$$

upon using (6.1.3) on the last two summands with $f \mapsto g^{-1}, g \mapsto g, h \mapsto \exp(t\xi)$. Now use again (6.1.3) on the second summand for $f \mapsto \exp(t\xi), g \mapsto g^{-1}, h \mapsto g$ to get

$$
\begin{aligned}
-\left.\frac{d}{dt}\right|_{t=0} & \left[\phi(g^{-1})B(g,\exp(t\xi)g^{-1}) - B(\exp(t\xi)g^{-1},g)\right. \\
& \left. +\phi(\exp(t\xi))B(g^{-1},g) - B(g^{-1},g)\right] \\
= -\sigma_R(g)(\xi) - & \left.\frac{d}{dt}\right|_{t=0} \left[\phi(\exp(t\xi))B(g^{-1},g) - B(g^{-1},g)\right] \\
= -\sigma_R(g)(\xi) - & \langle \Upsilon, \xi_A \rangle \left(B(g^{-1},g)\right)
\end{aligned}
$$

which proves (6.1.26).

Finally, to prove (6.1.27), use (6.1.12), (6.1.26), and (6.1.24) to get for any $\xi \in \mathfrak{g}$, $g \in G$, and $\alpha \in A$

$$
\begin{aligned}
\left(\widehat{\sigma}_L(g,\alpha) \cdot g\right)(\xi) &= T_0\phi(g^{-1})\widehat{\sigma}_L(g,\alpha)\left(\mathrm{Ad}_g\,\xi\right) \\
&= T_0\phi(g^{-1})\sigma_L(g)\left(\mathrm{Ad}_g\,\xi\right) + T_0\phi(g^{-1})\left\langle\Upsilon,(\mathrm{Ad}_g\,\xi)_A\right\rangle\left(\alpha + B(g,g^{-1})\right) \\
&= -\sigma_R(g)(\xi) - \langle\Upsilon,\xi_A\rangle\left(B(g^{-1},g)\right) \\
&\quad + T_0\phi(g^{-1})\left\langle\Upsilon,\phi(g)_*\xi_A\right\rangle\left(\phi(g)\left(\phi(g^{-1})(\alpha) + \phi(g^{-1})(B(g,g^{-1}))\right)\right) \\
&= -\sigma_R(g)(\xi) - \langle\Upsilon,\xi_A\rangle\left(B(g^{-1},g)\right) \\
&\quad + \langle\Upsilon,\xi_A\rangle\left(\phi(g^{-1})(\alpha) + \phi(g^{-1})(B(g,g^{-1}))\right).
\end{aligned}
$$

By (6.1.15) the last term is a sum of two terms. Taking into account (6.1.21) and (6.1.11) this equals

$$
\begin{aligned}
-\widehat{\sigma}_R(g,\alpha)(\xi) &- \langle\Upsilon,\xi_A\rangle\left(B(g^{-1},g)\right) + \langle\Upsilon,\xi_A\rangle\left(\phi(g^{-1})(B(g,g^{-1}))\right) \\
&= -\widehat{\sigma}_R(g,\alpha)(\xi)
\end{aligned}
$$

which proves (6.1.27).   ∎

The Adjoint Action of $G \times_{\phi,B} A$.   The underlying vector space of the Lie algebra of $G \times_{\phi,B} A$ is $\mathfrak{g} \times \mathfrak{a}$ since the underlying manifold of $G \times_{\phi,B} A$ is $G \times A$ and the neutral element is $(e,0)$. Differentiating (6.1.10) at the

neutral element $(e, 0) \in G \times_{\phi, B} A$, we obtain

$$
\mathrm{Ad}_{(g,\alpha)}(\xi, v) = \frac{d}{dt}\Big|_{t=0} \big(g \exp(t\xi) g^{-1}, -\phi(g)\phi(\exp(t\xi))\phi(g^{-1})(\alpha) + \alpha
$$
$$
+ \phi(g)\exp_A(tv) - \phi(g)\phi(\exp(t\xi))\phi(g^{-1})(B(g, g^{-1}))
$$
$$
+ B(g, \exp(t\xi)g^{-1}) + \phi(g)B(\exp(t\xi), g^{-1}) \big),
$$

where $(g, \alpha) \in G \times A$, $(\xi, v) \in \mathfrak{g} \times \mathfrak{a}$, $\exp : \mathfrak{g} \to G$ is the exponential map of the Lie group G, and $\exp_A : \mathfrak{a} \to A$ is the exponential map of the Abelian Lie group A. This becomes

$$
\mathrm{Ad}_{(g,\alpha)}(\xi, v) = \Big(\mathrm{Ad}_g \xi, -\frac{d}{dt}\Big|_{t=0} \big(\phi(g)\phi(\exp(t\xi))\phi(g^{-1})(\alpha) - \alpha \big)
$$
$$
+ T_0\phi(g)(v) + \frac{d}{dt}\Big|_{t=0} \phi(g) \big[-\phi(\exp(t\xi)g^{-1})(B(g, g^{-1})) + B(\exp(t\xi), g^{-1})
$$
$$
+ \phi(g^{-1})(B(g, \exp(t\xi)g^{-1})) \big] \Big).
$$

The first two terms in the corner bracket are equal to $-B(\exp(t\xi)g^{-1}, g)$ by applying (6.1.3) with $f \mapsto \exp(t\xi)g^{-1}, g \mapsto g, h \mapsto g^{-1}$ and taking into account that $B(k, e) = 0$ for any $k \in G$. Thus, using the definition (6.1.18) of σ_R we get

$$
\mathrm{Ad}_{(g,\alpha)}(\xi, v) = \Big(\mathrm{Ad}_g \xi, -\frac{d}{dt}\Big|_{t=0} \big(\phi(g)\phi(\exp(t\xi))\phi(g^{-1})(\alpha) - \alpha \big)
$$
$$
+ T_0\phi(g)(v) + T_0\phi(g)\, (\sigma_R(g)(\xi)) \Big)
$$
$$
= \Big(\mathrm{Ad}_g \xi, -T_\alpha \Lambda_{-\alpha} \big(T_{\phi(g^{-1})(\alpha)}\phi(g) \big(\xi_A(\phi(g^{-1})(\alpha)) \big) \big)
$$
$$
+ T_0\phi(g)(v) + T_0\phi(g)\, (\sigma_R(g)(\xi)) \Big)
$$
$$
= \Big(\mathrm{Ad}_g \xi, - \langle \Upsilon, \phi(g)_* \xi_A \rangle \, (\alpha)
$$
$$
+ T_0\phi(g)(v) + T_0\phi(g)\, (\sigma_R(g)(\xi)) \Big). \tag{6.1.28}
$$

The argument above shows that (6.1.18) and the definition of the adjoint action imply (6.1.28). Conversely, reading the previous computations backwards it follows that (6.1.28) and the definition of the adjoint action imply (6.1.18).

Now we show that (6.1.28) and the identity stating that the adjoint action is a left action is equivalent to (6.1.19). First, we have for any $g, h \in G$,

$\alpha, \beta \in A$, $\xi \in \mathfrak{g}$, and $v \in \mathfrak{a}$,

$$
\begin{aligned}
\mathrm{Ad}_{(g,\alpha)(h,\beta)}(\xi, v) &= \mathrm{Ad}_{(gh,\alpha+\phi(g)(\beta)+B(g,h))}(\xi, v) \\
&= \big(\mathrm{Ad}_{gh}\,\xi, -\langle \Upsilon, \phi(gh)_*\xi_A \rangle \, (\alpha + \phi(g)(\beta) + B(g,h)) \\
&\quad + T_0\phi(gh)(v) + T_0\phi(gh)\,(\sigma_R(gh)(\xi))\,\big).
\end{aligned}
\tag{6.1.29}
$$

On the other hand, by (6.1.28), (6.1.14), (6.1.17), (6.1.12), and (6.1.15) we have

$$
\begin{aligned}
&\mathrm{Ad}_{(g,\alpha)}\,\mathrm{Ad}_{(h,\beta)}(\xi, v) \\
&= \mathrm{Ad}_{(g,\alpha)}\big(\mathrm{Ad}_h\,\xi, -\langle \Upsilon, \phi(h)_*\xi_A \rangle\,(\beta) + T_0\phi(h)(v) + T_0\phi(h)\,(\sigma_R(h)(\xi))\,\big) \\
&= \big(\mathrm{Ad}_{gh}\,\xi, -\langle \Upsilon, \phi(g)_*\,(\mathrm{Ad}_h\,\xi)_A \rangle\,(\alpha) - T_0\phi(g)\,\langle \Upsilon, \phi(h)_*\xi_A \rangle\,(\beta) \\
&\quad + T_0\phi(g)T_0\phi(h)(v) + T_0\phi(g)T_0\phi(h)\,(\sigma_R(h)(\xi)) \\
&\quad + T_0\phi(g)\,(\sigma_R(g)\,(\mathrm{Ad}_h\,\xi))\,\big) \\
&= \big(\mathrm{Ad}_{gh}\,\xi, -\langle \Upsilon, \phi(g)_*\phi(h)_*\xi_A \rangle\,(\alpha) - T_0\phi(g)\,\langle \Upsilon, \phi(h)_*\xi_A \rangle\,(\phi(g^{-1})\phi(g)(\beta)) \\
&\quad + T_0\phi(gh)(v) + T_0\phi(gh)\,(\sigma_R(h)(\xi)) \\
&\quad + T_0\phi(gh)T_0\phi(h^{-1})\,(\sigma_R(g)\,(\mathrm{Ad}_h\,\xi))\,\big) \\
&= \big(\mathrm{Ad}_{gh}\,\xi, -\langle \Upsilon, \phi(gh)_*\xi_A \rangle\,(\alpha) - \langle \Upsilon, \phi(g)_*\phi(h)_*\xi_A \rangle\,(\phi(g)(\beta)) \\
&\quad + T_0\phi(gh)(v) + T_0\phi(gh)\,(\sigma_R(h)(\xi)) \\
&\quad + T_0\phi(gh)\,((\sigma_R(g) \cdot h)\,(\xi))\,\big) \\
&= \big(\mathrm{Ad}_{gh}\,\xi, -\langle \Upsilon, \phi(gh)_*\xi_A \rangle\,(\alpha + \phi(g)(\beta)) + T_0\phi(gh)(v) \\
&\quad + T_0\phi(gh)\,(\sigma_R(h)(\xi) + (\sigma_R(g) \cdot h)\,(\xi))\,\big).
\end{aligned}
\tag{6.1.30}
$$

Thus (6.1.29) and (6.1.30) are equal if and only if

$$
\begin{aligned}
&T_0\phi(gh)\,(\sigma_R(h)(\xi) + (\sigma_R(g) \cdot h)\,(\xi)) \\
&\quad = -\langle \Upsilon, \phi(gh)_*\xi_A \rangle\,(B(g,h)) + T_0\phi(gh)\,(\sigma_R(gh)(\xi)) \\
&\quad = -T_0\phi(gh)\,\langle \Upsilon, \xi_A \rangle\,(\phi(h^{-1}g^{-1})(B(g,h)) + T_0\phi(gh)\,(\sigma_R(gh)(\xi)) \\
&\quad = T_0\phi(gh)\big(-\langle \Upsilon, \xi_A \rangle\,(\phi(h^{-1}g^{-1})(B(g,h)) + (\sigma_R(gh)(\xi))\,\big),
\end{aligned}
$$

that is,

$$
\begin{aligned}
&\sigma_R(h)(\xi) + (\sigma_R(g) \cdot h)\,(\xi) \\
&\quad = -\langle \Upsilon, \xi_A \rangle\,(\phi(h^{-1}g^{-1})(B(g,h)) + \sigma_R(gh)(\xi)
\end{aligned}
$$

which is precisely (6.1.19).

Assume now that $A = V$, a vector space, and $\phi(g)$ is a linear isomorphism of V for every $g \in G$, that is, $\phi : G \to \mathrm{GL}(V)$ is a representation of G on V. Then $\xi_V = \phi'(\xi) \in \mathrm{gl}(V)$, the derivative of ϕ at the identity element e in the direction $\xi \in \mathfrak{g}$, is the induced Lie algebra representation. By (6.1.18) we

get in this case $\sigma_R(g) = \left(\phi(g^{-1})(D_2 B(g, g^{-1})) - D_1 B(g^{-1}, g)\right) \circ T_e R_{g^{-1}}$.
Using this expression and (6.1.20), formula (6.1.28) becomes

$$
\begin{aligned}
\mathrm{Ad}_{(g,\alpha)}(\xi, v) &= \Big(\mathrm{Ad}_g \xi, -(\phi(g) \circ \phi'(\xi) \circ \phi(g^{-1}))(\alpha + B(g, g^{-1})) \\
&\quad + \phi(g)v + D_2 B(g, e)(\xi) + D_1 B(g, g^{-1})(T_e L_g(\xi)) \Big) \\
&= \Big(\mathrm{Ad}_g \xi, -(\phi(g) \circ \phi'(\xi) \circ \phi(g^{-1}))(\alpha) + \phi(g)v \\
&\quad + \left(D_2 B(g, g^{-1}) - \phi(g) \circ D_1 B(g^{-1}, g)\right)(T_e R_{g^{-1}}\xi) \Big) \\
&= \Big(\mathrm{Ad}_g \xi, -\phi'(\mathrm{Ad}_g \xi)(\alpha) + \phi(g)v \\
&\quad + \left(D_2 B(g, g^{-1}) - \phi(g) \circ D_1 B(g^{-1}, g)\right)(T_e R_{g^{-1}}\xi) \Big). \qquad (6.1.31)
\end{aligned}
$$

Note that if $B = 0$, then (6.1.31) is the adjoint action of the semidirect product $G \, \circledS \, V$ and it coincides with (4.2.1).

The Left G- and $\mathfrak{g}$-Representations on $\mathfrak{a}$.   Since the G-action on A is by group homomorphisms, G also acts on on the left on $\mathfrak{a}$ by

$$
g \cdot v := T_0 \phi(g)(v). \qquad (6.1.32)
$$

Note that this formula defines a left G-representation on $\mathfrak{a}$.

The induced left Lie algebra representation of $\mathfrak{g}$ on $\mathfrak{a}$ is therefore given by

$$
\xi \cdot v := \left.\frac{d}{dt}\right|_{t=0} \exp(t\xi) \cdot v = \left.\frac{d}{dt}\right|_{t=0} T_0 \phi(\exp(t\xi))(v) = \xi_{\mathfrak{a}}(v), \qquad (6.1.33)
$$

where $\xi_{\mathfrak{a}} : \mathfrak{a} \to \mathfrak{a}$ denotes the infinitesimal generator of the action (6.1.32) defined by $\xi \in \mathfrak{g}$. Denote in what follows by $\widetilde{\phi} : \mathfrak{g} \to \mathfrak{gl}(\mathfrak{a})$ this representation of $\mathfrak{g}$ on $\mathfrak{a}$, that is, $\widetilde{\phi}(\xi)(v) := \xi \cdot v$ for any $\xi \in \mathfrak{g}$ and $v \in \mathfrak{a}$.

The Associated Lie Algebra Two-Cocycle on $\mathfrak{g}$.   Given the smooth group two-cocycle $B : G \times G \to A$, define the **Lie algebra two-cocycle** $C : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{a}$ associated with B by the formula

$$
C(\xi, \eta) := \left.\frac{d^2}{ds\,dt}\right|_{t=s=0} \left(B(g(t), h(s)) - B(h(s), g(t))\right), \qquad (6.1.34)
$$

where $t \mapsto g(t)$ and $s \mapsto h(s)$ are smooth curves through $e \in G$ with tangent vectors $\left.\frac{dg(t)}{dt}\right|_{t=0} = \xi$ and $\left.\frac{dh(s)}{ds}\right|_{s=0} = \eta \in \mathfrak{g}$. The **Lie algebra two-cocycle identity** states that

$$
\begin{aligned}
&\xi_{\mathfrak{a}}(C(\eta, \zeta)) + \eta_{\mathfrak{a}}(C(\zeta, \xi)) + \zeta_{\mathfrak{a}}(C(\xi, \eta)) \\
&\quad - C([\xi, \eta], \zeta) - C([\eta, \zeta], \xi) - C([\zeta, \xi], \eta) = 0 \qquad (6.1.35)
\end{aligned}
$$

for all $\xi, \eta, \zeta \in \mathfrak{g}$. The easiest proof of this identity, which will be done later, is as a consequence of the Jacobi identity for the bracket on the Lie algebra extension to which we turn next.

The Lie Bracket on $\mathfrak{g} \times_{\widetilde{\phi},C} \mathfrak{a}$.   We now calculate the Lie algebra bracket on $\mathfrak{g} \times_{\widetilde{\phi},C} \mathfrak{a} := T_{(e,0)} (G \times_{\phi,B} A)$, the Lie algebra of the group extension $G \times_{\phi,B} A$. The underlying vector space of $\mathfrak{g} \times_{\widetilde{\phi},C} \mathfrak{a}$ is $\mathfrak{g} \times \mathfrak{a}$. Let $(g(s), \alpha(s))$ be a curve through the identity with tangent vector $(\eta, w) \in \mathfrak{g} \times \mathfrak{a}$. Taking the derivative of $\mathrm{Ad}_{(g(s),\alpha(s))}(\xi, v)$ at $s = 0$ we get from (6.1.28)

$$
\begin{aligned}
[(\eta, w), (\xi, v)] &= \left.\frac{d}{ds}\right|_{s=0} \mathrm{Ad}_{(g(s),\alpha(s))}(\xi, v) \\
&= \left.\frac{d}{ds}\right|_{s=0} \left(\mathrm{Ad}_{g(s)}\, \xi, - \left.\frac{d}{dt}\right|_{t=0} \left[\phi(g(s))\phi(\exp(t\xi))\phi(g(s)^{-1})(\alpha(s)) - \alpha(s) \right] \right. \\
&\qquad \left. + T_0\varphi(g(s))(v) + T_0\varphi(g(s)) \left(\sigma_R(g(s))(\xi) \right) \right).
\end{aligned}
\tag{6.1.36}
$$

The second component of $[(\eta, w), (\xi, v)]$ has three terms. By (6.1.33), the second one equals $\eta_{\mathfrak{a}}(v) = \eta \cdot v$.

Next, we compute the first term of the second component of (6.1.36). We have

$$
\begin{aligned}
\left.\frac{d}{ds}\right|_{s=0} & \left.\frac{d}{dt}\right|_{t=0} \left[\phi(g(s))\phi(\exp(t\xi))\phi(g(s)^{-1})(\alpha(s)) - \alpha(s) \right] \\
&= \left.\frac{d}{dt}\right|_{t=0} T_0\Lambda_{\phi(g(0))\phi(\exp(t\xi)\phi(g(0)^{-1})(\alpha(0))} \left(-T_0\Lambda_{-\alpha(0)} T_0\Lambda_{-\alpha(0)}(w) \right) \\
&\quad + \left.\frac{d}{dt}\right|_{t=0} T_0\Lambda_{-\alpha(0)} \left[\left.\frac{d}{ds}\right|_{s=0} \phi(g(s))\phi(\exp(t\xi))\phi(g(s)^{-1})(\alpha(s)) \right] \\
&= - \left.\frac{d}{dt}\right|_{t=0} w + \left.\frac{d}{dt}\right|_{t=0} \left.\frac{d}{ds}\right|_{s=0} \phi(g(s))\phi(\exp(t\xi))\phi(g(s)^{-1})(\alpha(s)) \\
&= \left.\frac{d}{dt}\right|_{t=0} \left.\frac{d}{ds}\right|_{s=0} \phi(g(s))\phi(\exp(t\xi))\phi(g(s)^{-1})(\alpha(s))
\end{aligned}
$$

since $g(0) = e, \alpha(0) = 0$ and the action ϕ is by group homomorphisms of A. Now use the Leibniz identity for the s-derivative on the operator $\phi(g(s) \exp(t\xi)g(s)^{-1})$ and on $\alpha(s)$ to get

$$
\begin{aligned}
\left.\frac{d}{ds}\right|_{s=0} & \phi(g(s))\phi(\exp(t\xi))\phi(g(s)^{-1})(\alpha(s)) \\
&= \left.\frac{d}{ds}\right|_{s=0} \phi(g(s))\phi(\exp(t\xi))\phi(g(s)^{-1})(\alpha(0)) \\
&\quad + T_0\phi(g(0) \exp(t\xi)g(0)^{-1})(w) \\
&= T_0\phi(\exp(t\xi))(w) = \xi_{\mathfrak{a}}(w) = \xi \cdot w,
\end{aligned}
$$

again since $g(0) = e, \alpha(0) = 0$, and the action ϕ is by group homomorphisms of A.

Summarizing, the first two terms in the second component of (6.1.36) equals

$$\eta_{\mathfrak{a}}(v) - \xi_{\mathfrak{a}}(w) = \eta \cdot v - \xi \cdot w. \tag{6.1.37}$$

The third term of the second component of (6.1.36) is, by the Leibniz rule, equal to

$$\frac{d}{ds}\bigg|_{s=0} T_0\phi(g(s)) \left(\sigma_R(g(s))(\xi)\right)$$

$$= \frac{d}{ds}\bigg|_{s=0} T_0\phi(g(s)) \left(\sigma_R(g(0))(\xi)\right) + T_0\phi(g(0)) \left(\frac{d}{ds}\bigg|_{s=0} \sigma_R(g(s))(\xi)\right)$$

$$= \frac{d}{ds}\bigg|_{s=0} \sigma_R(g(s))(\xi) = (T_e\sigma_R(\eta))(\xi),$$

since $g(0) = e, \sigma_R(e) = 0$, and $\phi(0) = \mathrm{id}_G$. Thus this third term is by (6.1.18) equal to

$$(T_e\sigma_R(\eta))(\xi) = \frac{d}{ds}\bigg|_{s=0} \sigma_R(\exp(s\eta))(\xi)$$

$$= \frac{d}{ds}\bigg|_{s=0} \frac{d}{dt}\bigg|_{t=0} [\varphi(\exp(-s\eta)) \left(B(\exp(s\eta), \exp(t\xi)\exp(-s\eta))\right)$$

$$-B(\exp(t\xi)\exp(-s\eta), \exp(s\eta))] . \tag{6.1.38}$$

The s-derivative of the first term has two summands obtained from the Leibniz identity. The first one vanishes since $B(e, g) = 0$ for any $g \in G$. Since $\varphi(e) = \mathrm{id}_G$ we are left with just

$$\frac{d}{ds}\bigg|_{s=0} \frac{d}{dt}\bigg|_{t=0} B(\exp(s\eta), \exp(t\xi)\exp(-s\eta))$$

$$= \frac{d^2}{dsdt}\bigg|_{t=s=0} \left[B(\exp(s\eta), \exp(t\xi)) + B(e, \exp(t\xi)\exp(-s\eta))\right]$$

$$= \frac{d^2}{dsdt}\bigg|_{t=s=0} B(\exp(s\eta), \exp(t\xi)), \tag{6.1.39}$$

again using the Leibniz identity. A similar computation shows that

$$\frac{d^2}{dsdt}\bigg|_{t=s=0} B(\exp(t\xi)\exp(-s\eta), \exp(s\eta))$$

$$= \frac{d^2}{dsdt}\bigg|_{t=s=0} B(\exp(t\xi), \exp(s\eta)). \tag{6.1.40}$$

Formulas (6.1.38), (6.1.39), and (6.1.40) prove that the third term of the second component of (6.1.36) is equal to

$$(T_e\sigma_R(\eta))(\xi) = C(\eta, \xi). \tag{6.1.41}$$

Since, by definition, $\sigma_L(g) = \sigma_R(g^{-1})$, we also conclude from here that

$$(T_e\sigma_L(\eta))\,(\xi) = -C(\eta, \xi). \qquad (6.1.42)$$

Finally, formulas (6.1.36), (6.1.37), and (6.1.41) yield the following expression for the Lie bracket on the Lie algebra $\mathfrak{g} \times_{\widetilde{\phi},C} \mathfrak{a}$ of $G \times_{\phi,B} A$

$$[(\eta, w), (\xi, v)] = ([\eta, \xi], \eta_{\mathfrak{a}}(v) - \xi_{\mathfrak{a}}(w) + C(\eta, \xi))\,. \qquad (6.1.43)$$

If $A = V$, a vector space, and $\phi : G \times V \to V$ is a representation of G on V, then $\xi_V = \phi'(\xi) \in \mathrm{gl}(V)$. Thus the formula above becomes

$$[(\eta, w), (\xi, v)] = ([\eta, \xi], \phi'(\eta)(v) - \phi'(\xi)(w) + C(\eta, \xi)). \qquad (6.1.44)$$

This bracket reduces to the semidirect product bracket when $B = 0$ and hence $C = 0$. It reduces to the bracket for a central extension when $\phi(g) = \mathrm{id}_G$ for all $g \in G$, a case that will be studied in great detail in §6.2.

Lie Algebra Extensions. Given a Lie algebra $\mathfrak{k}$, a vector space U, a representation $\psi : \mathfrak{k} \to \mathfrak{gl}(U)$, and a Lie algebra two-cocycle $C : \mathfrak{k} \times \mathfrak{k} \to V$, one can form the extension $\mathfrak{k} \times_{\psi,C} V$ as the Lie algebra with underlying vector space $\mathfrak{k} \times V$ and bracket

$$[(\xi, u), (\eta, v)] := ([\xi, \eta], \xi \cdot v - \eta \cdot u + C(\xi, \eta)) \qquad (6.1.45)$$

for any $\xi, \eta \in \mathfrak{g}$ and $u, v \in V$. The Lie algebra two-cocycle identity (6.1.35) is equivalent to the Jacobi identity for the bracket (6.1.45).

Indeed, for $\xi, \eta, \zeta \in \mathfrak{k}$ and $u, v, w \in U$ we have

$$\begin{aligned}
[[(\xi, u), (\eta, v)], (\zeta, w)] &= [([\xi, \eta], \xi \cdot v - \eta \cdot u + C(\xi, \eta)), (\zeta, w)] \\
&= ([[\xi, \eta], \zeta], [\xi, \eta] \cdot w - \zeta \cdot (\xi \cdot v - \eta \cdot u + C(\xi, \eta)) + C([\xi, \eta], \zeta)) \\
&= ([[\xi, \eta], \zeta], [\xi, \eta] \cdot w - \zeta \cdot (\xi \cdot v) + \zeta \cdot (\eta \cdot u) - \zeta \cdot C(\xi, \eta) + C([\xi, \eta], \zeta))\,.
\end{aligned}$$

The sum of this term with the other two obtained by circular permutations yields zero in the first component by the Jacobi identity on $\mathfrak{g}$. The second component of this sum contains two types of terms: those that do not contain C and those that do. The terms that do not contain C in the second component of this sum are

$$\begin{aligned}
[\xi, \eta] \cdot w &- \zeta \cdot (\xi \cdot v) + \zeta \cdot (\eta \cdot u) + [\eta, \zeta] \cdot u - \xi \cdot (\eta \cdot w) + \xi \cdot (\zeta \cdot v) \\
&+ [\zeta, \xi] \cdot v - \eta \cdot (\zeta \cdot u) + \eta \cdot (\xi \cdot w) \\
&= [\xi, \eta] \cdot w - \xi \cdot (\eta \cdot w) + \eta \cdot (\xi \cdot w) + [\eta, \zeta] \cdot u - \eta \cdot (\zeta \cdot u) + \zeta \cdot (\eta \cdot u) \\
&+ [\zeta, \xi] \cdot v - \zeta \cdot (\xi \cdot v) + \xi \cdot (\zeta \cdot v) \\
&= [\xi, \eta] \cdot w - [\xi, \eta] \cdot w + [\eta, \zeta] \cdot u - [\eta, \zeta] \cdot u + [\zeta, \xi] \cdot v - [\zeta, \xi] \cdot v = 0
\end{aligned}$$

since the $\mathfrak{g}$-action on U is a representation. The terms involving C are hence

$$\zeta \cdot C(\xi, \eta) - C([\xi, \eta], \zeta) + \xi \cdot C(\eta, \zeta) - C([\eta, \zeta], \xi) + \eta \cdot C(\zeta, \xi) - C([\zeta, \xi], \eta).$$

Thus the Jacobi identity for the bracket (6.1.45) is equivalent to the vanishing of this expression, that is, to the Lie algebra two-cocycle identity (6.1.35).

In view of this discussion and (6.1.43) we can conclude the following.

6.1.2 Theorem.   *The Lie algebra of the Lie group extension $G \times_{\phi,B} A$ is the extended Lie algebra $\mathfrak{g} \times_{\widetilde{\phi},C} \mathfrak{a}$, where $\widetilde{\phi} : \mathfrak{g} \to \mathfrak{gl}(\mathfrak{a})$ is the naturally induced Lie algebra representation* (6.1.33) *from ϕ and $C : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{a}$ is the Lie algebra two-cocycle* (6.1.34) *defined by B.*

In particular, note that this proves that C defined by (6.1.34) satisfies the Lie algebra two-cocycle identity (6.1.35).

The Coadjoint Action of $G \times_{\phi,B} A$.   For the coadjoint action, let $(\mu, a) \in \mathfrak{g}^* \times \mathfrak{a}^*$ and $(\xi, v) \in \mathfrak{g} \times \mathfrak{a}$. By (6.1.28) we get

$$\left\langle \mathrm{Ad}^*_{(g,\alpha)}(\mu,a), (\xi,v) \right\rangle = \left\langle (\mu,a), \mathrm{Ad}_{(g,\alpha)}(\xi,v) \right\rangle$$
$$= \left\langle (\mu,a), \left(\mathrm{Ad}_g \xi, -T_\alpha \Lambda_{-\alpha} T_{\phi(g^{-1})(\alpha)} \phi(g) \left(\xi_A(\phi(g^{-1})(\alpha)) \right) \right. \right.$$
$$\left. \left. + T_0\phi(g)(v) + T_0\phi(g)\left(\sigma_R(g)(\xi)\right) \right) \right\rangle$$
$$= \left\langle \mu, \mathrm{Ad}_g \xi \right\rangle - \left\langle a, T_\alpha \Lambda_{-\alpha} T_{\phi(g^{-1})(\alpha)} \phi(g) \left(\xi_A(\phi(g^{-1})(\alpha)) \right) \right\rangle$$
$$+ \left\langle a, T_0\phi(g)(v) \right\rangle + \left\langle a, T_0\phi(g)\left(\sigma_R(g)(\xi)\right) \right\rangle.$$

Define $\Gamma^a_{(g,\alpha)} \in \mathfrak{g}^*$ by

$$\left\langle \Gamma^a_{(g,\alpha)}, \xi \right\rangle := \left\langle a, T_\alpha \Lambda_{-\alpha} T_{\phi(g^{-1})(\alpha)} \phi(g) \left(\xi_A(\phi(g^{-1})(\alpha)) \right) \right\rangle$$
$$= \left\langle a, T_\alpha \Lambda_{-\alpha} \left(\mathrm{Ad}_g \xi \right)_A (\alpha) \right\rangle \qquad (6.1.46)$$

for any $\xi \in \mathfrak{g}$, where we have used (6.1.14) in the last equality. Therefore

$$\left\langle \mathrm{Ad}^*_{(g,\alpha)}(\mu,a), (\xi,v) \right\rangle = \left\langle \mathrm{Ad}^*_g \mu, \xi \right\rangle - \left\langle \Gamma^a_{(g,\alpha)}, \xi \right\rangle + \left\langle T_0^* \phi(g)a, v \right\rangle$$
$$+ \left\langle \sigma_R(g)^* T_0^* \phi(g)a, \xi \right\rangle$$

which yields the *right* coadjoint action

$$\mathrm{Ad}^*_{(g,\alpha)}(\mu,a) = \left(\mathrm{Ad}^*_g \mu - \Gamma^a_{(g,\alpha)} + \sigma_R(g)^* T_0^* \phi(g)a, \, T_0^* \phi(g)a \right). \quad (6.1.47)$$

Therefore, from (6.1.9) we get the *left* coadjoint action

$$\mathrm{Ad}^*_{(g,\alpha)^{-1}}(\mu,a) = \left(\mathrm{Ad}^*_{g^{-1}} \mu - \Gamma^a_{(g^{-1},-B(g^{-1},g)-\phi(g^{-1})(\alpha))} \right.$$
$$\left. + \sigma_R(g^{-1})^* T_0^* \phi(g^{-1})a, \, T_0^* \phi(g^{-1})a \right). \qquad (6.1.48)$$

In particular, if $A = V$, a vector space, and $\phi : G \to \mathrm{GL}(V)$, a representation, then (6.1.46) becomes

$$\left\langle \Gamma^a_{(g,\alpha)}, \xi \right\rangle = \left\langle a, \phi(g)\phi'(\xi)(\phi(g^{-1})(\alpha)) \right\rangle = \left\langle \left(\phi'_{\phi(g^{-1})(\alpha)} \right)^* \phi(g)^*(a), \xi \right\rangle$$

where $\phi'_\beta : \mathfrak{g} \to V$ is defined by $\phi'_\beta(\xi) := \phi'(\xi)(\beta)$ for any $\xi \in \mathfrak{g}$ and $\beta \in V$. Therefore

$$-\Gamma^a_{(g^{-1}, -B(g^{-1},g)-\phi(g^{-1})(\alpha))} = -\left(\phi'_{\phi(g)(-\phi(g^{-1})(\alpha)-B(g^{-1},g))}\right)^* \phi(g^{-1})^*(a)$$

$$= (\phi'_\alpha)^* \phi(g^{-1})^*(a) + \left(\phi'_{\phi(g)(B(g^{-1},g))}\right)^* \phi(g^{-1})^*(a).$$

We simplify notation as in §4.2 and denote by concatenation all natural left actions of G on $\mathfrak{g}, \mathfrak{g}^*, V$, and V^*. Thus, $g\xi := \mathrm{Ad}_g \xi$, $g\mu := \mathrm{Ad}^*_{g^{-1}} \mu$, $g\alpha := \phi(g)(\alpha)$, and $ga := \phi(g^{-1})^*(a)$, for $g \in G, \xi \in \mathfrak{g}, \mu \in \mathfrak{g}^*, \alpha \in V$, and $a \in V^*$. Formula (6.1.48) becomes

$$\mathrm{Ad}_{(g,\alpha)^{-1}}(\mu, a)^*$$
$$= \left(g\mu + (\phi'_\alpha)^* ga + \left(\phi'_{\phi(g)(B(g^{-1},g))}\right)^* ga + \sigma_R(g^{-1})^* ga, ga\right). \quad (6.1.49)$$

Recall also that in this case the generalized one-cocycle σ_R has the simpler expression $\sigma_R(g) = \left(\phi(g^{-1})(D_2 B(g, g^{-1})) - D_1 B(g^{-1}, g)\right) \circ T_e R_{g^{-1}}$. Note also that if $B = 0$ and hence also $\sigma_R = 0$, then (6.1.49) gives the formula for the coadjoint action of the semidirect product $G \circledS V$ and coincides with (4.2.2).

Triviality of the Next Extension.   In analogy with (6.1.41), the derivative of $\widehat{\sigma}_R : G \times_{\phi,B} A \to L(\mathfrak{g}, \mathfrak{a})$ at the identity should give a two-cocycle on $\mathfrak{g} \times_{\widehat{\phi},C} \mathfrak{a}$. However, this cannot be the case literally, because $T_{(e,0)}\widehat{\sigma}_R(\xi, u) \in L(\mathfrak{g}, \mathfrak{a})$ does not naturally induce a skew-symmetric bilinear map on $\mathfrak{g} \times_{\widehat{\phi},C} \mathfrak{a}$. To see what the obvious correction to this derivative should be in order to get a two-cocycle, we compute first $T_{(e,0)}\widehat{\sigma}_R(\xi, u)$. For any $\xi, \eta \in \mathfrak{g}$ and $u \in \mathfrak{a}$ we have by (6.1.21) and (6.1.41)

$$T_{(e,0)}\widehat{\sigma}_R(\xi, u)(\eta) = \left.\frac{d}{dt}\right|_{t=0} \widehat{\sigma}_R(\exp(t\xi), \exp_A(tu))(\eta)$$

$$= \left.\frac{d}{dt}\right|_{t=0} \sigma_R(\exp(t\xi))(\eta) - \left.\frac{d}{dt}\right|_{t=0} \langle \Upsilon, \eta_A \rangle \left(\phi(\exp(-t\xi))(\exp_A(tu))\right)$$

$$= T_e \sigma_R(\xi)(\eta) - T_0 \langle \Upsilon, \eta_A \rangle \left(\left.\frac{d}{dt}\right|_{t=0} \phi(\exp(-t\xi))(\exp_A(tu))\right)$$

$$= C(\xi, \eta) - T_0 \langle \Upsilon, \eta_A \rangle (u)$$

since, by the Leibniz rule, we get

$$\left.\frac{d}{dt}\right|_{t=0} \phi(\exp(-t\xi))(\exp_A(tu))$$

$$= \left.\frac{d}{dt}\right|_{t=0} \phi(\exp(-t\xi))(0) + T_0\phi(e)\left(\left.\frac{d}{dt}\right|_{t=0} \exp_A(tu)\right) = u,$$

where we have used the fact that the action ϕ is by group homomorphisms and $\phi(e) = \mathrm{id}_A$.

So we need to compute $T_0 \langle \Upsilon, \eta_A \rangle (u)$. We have

$$
\begin{aligned}
T_0 \langle \Upsilon, \eta_A \rangle (u) &= \left.\frac{d}{dt}\right|_{t=0} \langle \Upsilon, \eta_A \rangle (\exp_A(tu)) \\
&= \left.\frac{d}{dt}\right|_{t=0} T_{\exp_A(tu)} \Lambda_{-\exp_A(tu)} \eta_A(\exp_A(tu)) \\
&= \left.\frac{d}{dt}\right|_{t=0} T_{\exp_A(tu)} \Lambda_{-\exp_A(tu)} \left(\left.\frac{d}{ds}\right|_{s=0} \phi(\exp(s\eta))(\exp_A(tu)) \right) \\
&= \left.\frac{d}{dt}\right|_{t=0} \left.\frac{d}{ds}\right|_{s=0} [\phi(\exp(s\eta))(\exp_A(tu)) - \exp_A(tu)] \\
&= \left.\frac{d}{ds}\right|_{s=0} [T_0\phi(\exp(s\eta))(u) - u] \\
&= \eta_{\mathfrak{a}}(u) = \eta \cdot u.
\end{aligned}
$$

We conclude that $T_{(e,0)}\widehat{\sigma}_R(\xi, u)(\eta) = C(\xi, \eta) - \eta \cdot u$. Since we need a Lie algebra two-cocycle on $\mathfrak{g} \times_{\widetilde{\phi},C} \mathfrak{a}$, we need to add to this expression something that makes it skew-symmetric. So, define

$$
\widehat{C}((\xi, u), (\eta, v)) := T_{(e,0)}\widehat{\sigma}_R(\xi, u)(\eta) + \xi_{\mathfrak{a}}(v) = C(\xi, \eta) - \eta \cdot u + \xi \cdot v. \quad (6.1.50)
$$

The map $\widehat{C}$ is not only bilinear and skew-symmetric, but also an $\mathfrak{a}$-valued coboundary on $\mathfrak{g} \times_{\widetilde{\phi},C} \mathfrak{a}$ relative to the trivial action of $\mathfrak{g} \times_{\widetilde{\phi},C} \mathfrak{a}$ on $\mathfrak{a}$. Indeed, by (6.1.43) we have

$$
\widehat{C}((\xi, u), (\eta, v)) = \lambda([(\xi, u), (\eta, v)]),
$$

where $\lambda : \mathfrak{g} \times \mathfrak{a} \to \mathfrak{a}$ is the projection.

Therefore, *the extension of $\mathfrak{g} \times_{\widetilde{\phi},C} \mathfrak{a}$ by $\mathfrak{a}$ defined by $\widehat{C}$ is central and isomorphic to the product Lie algebra $\left(\mathfrak{g} \times_{\widetilde{\phi},C} \mathfrak{a} \right) \times \mathfrak{a}$.* Indeed, the Lie algebra bracket on the extension $\left(\mathfrak{g} \times_{\widetilde{\phi},C} \mathfrak{a} \right) \times_{\widehat{C}} \mathfrak{a}$ is given by

$$
\begin{aligned}
[(\xi, u_1, u_2), (\eta, v_1, v_2)] &:= \left([(\xi, u_1), (\eta, v_1)], \widehat{C}((\xi, u_1), (\eta, v_1))] \right) \\
&= ([\xi, \eta], \xi \cdot v_1 - \eta \cdot u_1 + C(\xi, \eta), \xi \cdot v_1 - \eta \cdot u_1 + C(\xi, \eta))
\end{aligned}
$$

for any $\xi, \eta \in \mathfrak{g}$ and $u_1, u_2, v_1, v_2 \in \mathfrak{a}$. The map

$$
(\xi, u_1, u_2) \in \left(\mathfrak{g} \times_{\widetilde{\phi},C} \mathfrak{a} \right) \times_{\widehat{C}} \mathfrak{a} \mapsto (\xi, u_1, u_1 - u_2) \in \left(\mathfrak{g} \times_{\widetilde{\phi},C} \mathfrak{a} \right) \times \mathfrak{a}
$$

establishes a Lie algebra isomorphism between the Lie algebra extension $\left(\mathfrak{g} \times_{\widetilde{\phi},C} \mathfrak{a} \right) \times_{\widehat{C}} \mathfrak{a}$ and the product Lie algebra $\left(\mathfrak{g} \times_{\widetilde{\phi},C} \mathfrak{a} \right) \times \mathfrak{a}$.

This shows that *the Lie algebra extension of $\mathfrak{g} \times_{\widetilde{\phi},C} \mathfrak{a}$ determined by the $L(\mathfrak{g}, \mathfrak{a})$-valued one cocycle $\widehat{\sigma}_R$ on $G \times_{\phi,B} A$ is trivial, that is, it is the direct product $\left(\mathfrak{g} \times_{\widetilde{\phi},C} \mathfrak{a} \right) \times \mathfrak{a}$.* Thus, nothing new in terms of extensions is given by $\widehat{\sigma}_R$.

6.2   Central Extensions

The considerations of the previous section simplify considerably if $\phi(g) =$ id$_G$ for every $g \in G$, that is, if we consider a central extension of the Lie group G by the Abelian Lie group $(A, +)$ defined by the smooth normalized group two-cocycle $B : G \times G \to A$. The cocycle identity (6.1.3) becomes in this case

$$B(f, g) + B(fg, h) = B(f, gh) + B(g, h) \quad \text{for all} \quad f, g, h \in G. \quad (6.2.1)$$

In addition, by (6.1.11) we also have $B(g, g^{-1}) = B(g^{-1}, g)$ for any $g \in G$. Proposition 6.1.1 takes on the following form.

6.2.1 Proposition.   (i) *The formula*

$$
\begin{aligned}
\sigma_R(g) &= T_{B(g^{-1},g)}\Lambda_{-B(g^{-1},g)} \circ \left[D_2 B(g, g^{-1}) - D_1 B(g^{-1}, g)\right] \circ T_e R_{g^{-1}} \\
&= \left(T_{B(g,g^{-1})}\Lambda_{-B(g,g^{-1})} \circ D_2 B(g, g^{-1}) \circ T_e R_{g^{-1}}\right)(\xi) \\
&\quad - \left(T_{B(g^{-1},g)}\Lambda_{-B(g^{-1},g)} \circ D_1 B(g^{-1}, g) \circ T_e R_{g^{-1}}\right)(\xi) \\
&= D_2 B(g, e) + T_{B(g,g^{-1})}\Lambda_{-B(g,g^{-1})} \circ D_1 B(g, g^{-1}) \circ T_e L_g \\
&= D_1 B(e, g^{-1}) + T_{B(g,g^{-1})}\Lambda_{-B(g,g^{-1})} \circ D_2 B(g, g^{-1}) \circ T_e R_{g^{-1}} \quad (6.2.2)
\end{aligned}
$$

defines a right $L(\mathfrak{g}, \mathfrak{a})$-valued one-cocycle $\sigma : G \to L(\mathfrak{g}, \mathfrak{a})$ on G, that is, the following identity holds

$$\sigma_R(gh) = \sigma_R(g) \circ \mathrm{Ad}_h + \sigma_R(h) \quad \text{for all} \quad g, h \in G. \quad (6.2.3)$$

(ii) *The formula $\sigma_L(g) := \sigma_R(g^{-1})$, that is*

$$
\begin{aligned}
\sigma_L(g) &= D_1 B(e, g) + T_{B(g^{-1},g)}\Lambda_{-B(g^{-1},g)} \circ D_2 B(g^{-1}, g) \circ T_e R_g \\
&= D_2 B(g^{-1}, e) + T_{B(g^{-1},g)}\Lambda_{-B(g^{-1},g)} \circ D_1 B(g^{-1}, g) \circ T_e L_{g^{-1}} \quad (6.2.4)
\end{aligned}
$$

defines a left $L(\mathfrak{g}, \mathfrak{a})$-valued one-cocycle $\sigma : G \to L(\mathfrak{g}, \mathfrak{a})$ on G, that is, the following identity holds

$$\sigma_L(gh) = \sigma_L(g) + \sigma_L(h) \circ \mathrm{Ad}_{g^{-1}} \quad \text{for all} \quad g, h \in G. \quad (6.2.5)$$

(iii) *The relation between σ_R and σ_L is given by*

$$\sigma_L(g) \circ \mathrm{Ad}_g = -\sigma_R(g) \quad (6.2.6)$$

for any $g \in G$.

The Lie bracket (6.1.44) of $\mathfrak{g} \times \mathfrak{a}$ becomes in this case

$$[(\xi, v), (\eta, w)] = ([\xi, \eta], C(\xi, \eta)), \quad (6.2.7)$$

where the Lie algebra two-cocycle $C : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{a}$ is given, as before, by (6.1.34). The two-cocycle identity (6.1.35) becomes in this case

$$C([\xi, \eta], \zeta) + C([\eta, \zeta], \xi) + C([\zeta, \xi], \eta) = 0 \quad (6.2.8)$$

for all $\xi, \eta, \zeta \in \mathfrak{g}$. Similarly, formula (6.1.41) simplifies to

$$T_e \sigma_R(\xi) = C(\xi, \cdot), \quad \text{for any} \quad \xi \in \mathfrak{g}. \tag{6.2.9}$$

Since $\sigma_L(g) = \sigma_R(g^{-1})$ this immediately yields

$$T_e \sigma_L(\xi) = -C(\xi, \cdot), \quad \text{for any} \quad \xi \in \mathfrak{g}. \tag{6.2.10}$$

It should be noted that up to this point no assumption on the topological character of the central extension was made. In particular, it was not assumed that the underlying manifold of the central extension was the product $G \times A$.

Now, if we assume that the central extension $G \times_B A$ of G by A has underlying manifold $G \times A$, we can further simplify several formulas of the previous section. The group multiplication (6.1.2) in $G \times_B A$ becomes

$$(g, \alpha)(h, \beta) = (gh, \alpha + \beta + B(g, h)), \tag{6.2.11}$$

the identity element is $(e, 0)$ and the inverse (6.1.9) is given by

$$(g, \alpha)^{-1} = (g^{-1}, -\alpha - B(g^{-1}, g)) = (g^{-1}, -\alpha - B(g, g^{-1})). \tag{6.2.12}$$

The inner automorphisms (6.1.10) on $G \times_B V$ are given by

$$I_{(g,\alpha)}(h, \beta) = \left(ghg^{-1}, \beta - B(g, g^{-1}) + B(h, g^{-1}) + B(g, hg^{-1})\right) \tag{6.2.13}$$
$$= \left(ghg^{-1}, \beta - B(g, g^{-1}) + B(g, h) + B(gh, g^{-1})\right). \tag{6.2.14}$$

The adjoint (6.1.28) and coadjoint (6.1.48) actions of $G \times_B A$ on $\mathfrak{g} \times \mathfrak{a}$ and $\mathfrak{g}^* \times \mathfrak{a}^*$ are given for central extensions respectively by

$$\text{Ad}_{(g,\alpha)}(\xi, v) = \left(\text{Ad}_g \xi, v + \sigma_R(g)(\xi)\right), \tag{6.2.15}$$

and

$$\text{Ad}^*_{(g,\alpha)^{-1}}(\mu, a) = \left(\text{Ad}^*_{g^{-1}} \mu + a \circ \sigma_R(g^{-1}), a\right) \tag{6.2.16}$$

for $g \in G, \alpha \in A, \xi \in \mathfrak{g}, \mu \in \mathfrak{g}^*, v \in \mathfrak{a}$ and $a \in \mathfrak{a}^*$. As a consequence we see that the coadjoint orbits of the central extension lie in the affine spaces $\mathfrak{g}^* \times \{a\}$ for all $a \in \mathfrak{a}^*$. We also get

$$\text{ad}^*_{(\xi,t)}(\mu, a) = \left(\text{ad}^*_\xi \mu + a \circ C(\xi, \cdot), 0\right). \tag{6.2.17}$$

The Lie-Poisson bracket and the associated Hamiltonian vector field are now easy to compute. They are

$$\{F, H\}(\mu, a) = \pm \left\langle \mu, \left[\frac{\delta F}{\delta \mu}, \frac{\delta H}{\delta \mu}\right] \right\rangle \pm \left\langle a, C\left(\frac{\delta F}{\delta \mu}, \frac{\delta H}{\delta \mu}\right) \right\rangle \tag{6.2.18}$$

and

$$X_H(\mu, a) = \mp \left(\text{ad}^*_{\frac{\delta F}{\delta \mu}} \mu + a \circ C\left(\frac{\delta F}{\delta \mu}, \cdot\right), 0\right), \tag{6.2.19}$$

which shows that the affine spaces $\mathfrak{g}^* \times \{a\}$ endowed with the Poisson bracket depending on $a \in \mathfrak{a}^*$ given by the right hand side of (6.2.18) are Poisson submanifolds of $\mathfrak{g}^* \times \mathfrak{a}^*$. On each affine space $\mathfrak{g}^* \times \{a\}$ we have an affine G-action given by

$$g \cdot (\mu, a) := \left(\mathrm{Ad}^*_{g^{-1}} \mu + a \circ \sigma_R(g^{-1}), a \right) \qquad (6.2.20)$$

and the orbits of this affine action obviously coincide with the coadjoint orbits of the central extension $G \times A$. Denote by $\mathcal{O}_{(\mu,a)}$ the coadjoint orbit of $G \times A$ containing $(\mu, a) \in \mathfrak{g}^* \times \mathfrak{a}^*$ and by $\widetilde{\mathcal{O}}^a_\mu$ the orbit of the affine G-action through $\mu \in \mathfrak{g}^*$ corresponding to $a \in \mathfrak{a}^*$. We have hence

$$\mathcal{O}_{(\mu,a)} = \widetilde{\mathcal{O}}^a_\mu$$

which proves that $\widetilde{\mathcal{O}}^a_\mu$ is a symplectic manifold relative to the symplectic form

$$\omega^{\pm}_{\widetilde{\mathcal{O}}^a_\mu}(\nu) \left(\mathrm{ad}^*_\xi \mu + a \circ C(\xi, \cdot), \mathrm{ad}^*_\eta \mu + a \circ C(\eta, \cdot) \right)$$
$$= \pm \langle \nu, [\xi, \eta] \rangle \pm a \circ C \left([\xi, \eta], \cdot \right), \qquad (6.2.21)$$

where $\nu = \mathrm{Ad}^*_{g^{-1}} \mu + a \circ \sigma_R(g^{-1}) \in \widetilde{\mathcal{O}}^a_\mu$.

Let us summarize this discussion in the following theorem.

6.2.2 Theorem.   *For any $a \in \mathfrak{a}^*$ there is an affine Poisson structure on $\mathfrak{g}^*$ defined by*

$$\{F, H\}_{a,C,\pm}(\mu) = \pm \left\langle \mu, \left[\frac{\delta F}{\delta \mu}, \frac{\delta H}{\delta \mu} \right] \right\rangle \pm \left\langle a, C \left(\frac{\delta F}{\delta \mu}, \frac{\delta H}{\delta \mu} \right) \right\rangle. \qquad (6.2.22)$$

The associated Hamiltonian vector field is given by

$$X^{a,C,\pm}_H(\mu) = \mp \mathrm{ad}^*_{\frac{\delta F}{\delta \mu}} \mu \mp a \circ C \left(\frac{\delta F}{\delta \mu}, \cdot \right). \qquad (6.2.23)$$

Denote this Poisson manifold by $\mathfrak{g}^*_{a,C,\pm}$. For each $a \in \mathfrak{a}^*$ there is a left affine G-action on $\mathfrak{g}^*$ given by $g \cdot \mu := \mathrm{Ad}^*_{g^{-1}} \mu + a \circ \sigma_R(g^{-1})$. The map $\mu \in \mathfrak{g}^*_{a,C,\pm} \mapsto (\mu, a) \in (\mathfrak{g}^* \times \mathfrak{a}^*)_\pm$ is a Poisson embedding. The symplectic leaves of $\mathfrak{g}^*_{a,C,\pm}$ are the connected components of the affine orbits $\widetilde{\mathcal{O}}^a_\mu$ endowed with the symplectic form (6.2.21). Each affine orbit $\widetilde{\mathcal{O}}^a_\mu$ is symplectically diffeomorphic by the embedding given above to the $G \times A$-coadjoint orbit $\mathcal{O}_{(\mu,a)}$ containing (μ, a) endowed with the $\pm$-orbit symplectic structure.

Remark. That $\widetilde{\mathcal{O}}^a_\mu$ is a symplectic manifold with the given symplectic form is proved *directly*, as opposed to the present approach that uses the coadjoint orbits of the central extension, in [HRed], Theorem 4.5.31.

Part of the previous theorem can be considerably strengthened if we consider only one-dimensional extensions of Lie algebras. Thus, if $C : \mathfrak{g} \times \mathfrak{g} \to \mathbb{R}$ is a Lie algebra two cocycle, Theorem 6.2.2 guarantees that

$$\{F, H\}_C(\mu) = \left\langle \mu, \left[\frac{\delta F}{\delta \mu}, \frac{\delta H}{\delta \mu} \right] \right\rangle + C \left(\frac{\delta F}{\delta \mu}, \frac{\delta H}{\delta \mu} \right) \tag{6.2.24}$$

is an affine Poisson bracket on $\mathfrak{g}^*$. Conversely, if $C : \mathfrak{g} \times \mathfrak{g} \to \mathbb{R}$ is a bilinear skew-symmetric form such that (6.2.24) is a Poisson bracket, then writing the Jacobi identity for linear functionals on $\mathfrak{g}^*$ it follows that C satisfies the Lie algebra two-cocycle identity. The affine Poisson bracket associated to C is then obviously given by (6.2.24). This proves the following result of Bhaskara [1990] (see also Vaisman [1994], Proposition 3.4, page 35).

6.2.3 Corollary.   *The affine Poisson structures on $\mathfrak{g}^*$ are in one-to-one correspondence with the central extensions of $\mathfrak{g}$ by $\mathbb{R}$.*

6.3   Group Extensions Satisfy the Stages Hypotheses

The main goal of this section is to prove that any element of the dual of the Lie algebra of the extension of a Lie group by an Abelian Lie group with a cocycle studied in §6.1 satisfies the stages hypothesis (see Definition 5.2.8).

Let $M = G \times_{\phi,B} A$ be an extension of the Lie group G by the Abelian Lie group A associated to the left smooth action $\phi : G \times A \to A$ by Abelian Lie homomorphisms and the smooth normalized Lie group two-cocycle $B : G \times G \to A$. It is assumed that the underlying manifold of M is the product $G \times A$. Formula (6.1.10) shows that $N := \{e\} \times A$ is a normal subgroup of M. Therefore, (5.2.1) defines an action of M on $\mathfrak{n}^* = \{0\} \times \mathfrak{a}^*$ given therefore by the second component of (6.1.48), that is,

$$(g, \alpha) \cdot a = T_0^* \phi(g^{-1}) a \tag{6.3.1}$$

for $(g, \alpha) \in M$ and $a \in \mathfrak{a}^*$.

6.3.1 Lemma.   *Let $a \in \mathfrak{a}^*$. The isotropy subgroup $(G \times_{\phi,B} A)_a$ for the action (6.3.1) of $G \times_{\phi,B} A$ on $\mathfrak{a}^*$ equals $G_a \times A$, where $G_a = \{g \in G \mid T_0^* \phi(g^{-1}) a = a\}$.*

Proof.   The element $(g, \alpha) \in (G \times_{\phi,B} A)_a$ if and only if $T_0^* \phi(g^{-1}) a = (g, \alpha) \cdot a = a$ which is obviously equivalent to $(g, \alpha) \in G_a \times A$.   ▼

6.3.2 Theorem.   *Any element of the dual $\mathfrak{m}^*$ of the Lie algebra $\mathfrak{m}$ of the Lie group extension $M = G \times_{\phi,B} A$ satisfies the stages hypothesis (see Definition 5.2.8).*

Proof.   We use the notations in the stages hypotheses of Definition 5.2.8. Let $\sigma := (\mu, a) \in \mathfrak{m}^*$ and $\sigma' = (\mu', a')$ be another element of $\mathfrak{m}^*$ satisfying the assumptions of the stages hypothesis. In particular, this means that $\sigma|_{\{0\} \times \mathfrak{a}} = \sigma'|_{\{0\} \times \mathfrak{a}}$ which is equivalent to $a = a'$. Thus $\sigma' = (\mu', a)$.

By the previous lemma, we have $\mathfrak{m}_a = \mathfrak{g}_a \times \mathfrak{a}$, where $\mathfrak{g}_a = \{\xi \in \mathfrak{g} \mid \xi_{\mathfrak{a}^*}(a) = 0\}$ is the Lie algebra of G_a and $\xi_{\mathfrak{a}^*}$ is the infinitesimal generator of the action (6.3.1) given by $\xi \in \mathfrak{g}$. The second assumption of the stages hypothesis is that $\sigma|_{\mathfrak{m}_a} = \sigma'|_{\mathfrak{m}_a}$ which is equivalent to $\mu|_{\mathfrak{g}_a} = \mu'|_{\mathfrak{g}_a}$, that is, $\mu - \mu' \in \mathfrak{g}_a^\circ$.

The coadjoint isotropy subgroup $A_a = A$ since A is Abelian. The stages hypothesis requires to find an element $(g, \alpha) \in (\{e\} \times A) \cdot (G_a \times A)_{\sigma|_{\mathfrak{m}_a}}$ such that $\sigma' = \mathrm{Ad}^*_{(g,\alpha)} \sigma$, which, by (6.1.48) and $g \in G_a$, is equivalent to

$$\mu' = \mathrm{Ad}^*_g \mu - \Gamma^a_{(g,\alpha)} + \sigma_R(g)^* a.$$

We will search for an element of the form (e, α) that satisfies this relation. Since $\sigma_R(e) = 0$, we need to find an element $\alpha \in A$ such that

$$\mu' - \mu = -\Gamma^a_{(e,\alpha)}. \tag{6.3.2}$$

We begin by finding a more convenient expression for $\Gamma^a_{(e,\alpha)}$. Let $\xi \in \mathfrak{g}$. By (6.1.46) and the obvious identity $\Lambda_{-\alpha} \circ \phi(g) = \phi(g) \circ \Lambda_{-\phi(g^{-1})(\alpha)}$ we get for any $g \in G_a$

$$\begin{aligned}
\left\langle \Gamma^a_{(g,\alpha)}, \xi \right\rangle :&= \left\langle a, T_0\phi(g) T_{\phi(g^{-1})(\alpha)} \Lambda_{-\phi(g^{-1})(\alpha)} \left(\xi_A(\phi(g^{-1})(\alpha)) \right) \right\rangle \\
&= \left\langle T_0^* \phi(g) a, T_{\phi(g^{-1})(\alpha)} \Lambda_{-\phi(g^{-1})(\alpha)} \left(\xi_A(\phi(g^{-1})(\alpha)) \right) \right\rangle \\
&= \left\langle a, T_{\phi(g^{-1})(\alpha)} \Lambda_{-\phi(g^{-1})(\alpha)} \left(\xi_A(\phi(g^{-1})(\alpha)) \right) \right\rangle. \tag{6.3.3}
\end{aligned}$$

To continue this computation we need to link ξ_A and $\xi_\mathfrak{a}$. For an arbitrary $v \in \mathfrak{a}$ and $g \in G$ we have $\phi(g)(\exp_A v) = \exp_A (T_0\phi(g)(v))$. Setting here $g = \exp(t\xi)$ and taking the t-derivative of the resulting relation at $t = 0$ we get

$$\begin{aligned}
\xi_A (\exp_A v) &= \left. \frac{d}{dt} \right|_{t=0} \phi(\exp(t\xi))(\exp_A v) \\
&= \left. \frac{d}{dt} \right|_{t=0} \exp_A \left(T_0\phi(\exp(t\xi))(v) \right) \\
&= T_v \exp_A (\xi_\mathfrak{a}(v)) = T_0\Lambda_{\exp_A v} (\xi_\mathfrak{a}(v)),
\end{aligned}$$

since for Abelian Lie groups the derivative of the exponential map equals the derivative of the translation map at the exponential of the point where the derivative is taken. Since for Abelian Lie groups the exponential map is surjective onto the connected component of the identity, the previous identity shows that

$$\xi_A(\alpha) = T_0\Lambda_\alpha (\xi_\mathfrak{a}(v)), \quad \text{for any} \quad v \in \mathfrak{a} \quad \text{satisfying} \quad \alpha = \exp_A v. \tag{6.3.4}$$

Therefore, if α is in the connected component of the identity, there is some $u \in \mathfrak{a}$ that satisfies $\exp_A u = \phi(g^{-1})(\alpha)$ because $\phi(g^{-1})$ is a Lie group automorphisms of A. Thus, for such α, (6.3.3) becomes

$$\left\langle \Gamma^a_{(g,\alpha)}, \xi \right\rangle := \left\langle a, T_{\phi(g^{-1})(\alpha)} \Lambda_{-\phi(g^{-1})(\alpha)} T_0 \Lambda_{\phi(g^{-1})(\alpha)} \left(\xi_{\mathfrak{a}}(u)\right) \right\rangle$$

$$= \left\langle a, \xi_{\mathfrak{a}}(u) \right\rangle. \tag{6.3.5}$$

Next, note that

$$\left\langle a, \xi_{\mathfrak{a}}(u) \right\rangle = \left.\frac{d}{dt}\right|_{t=0} \left\langle a, T_0 \phi(\exp(t\xi))(u) \right\rangle = \left.\frac{d}{dt}\right|_{t=0} \left\langle T_0^* \phi(\exp(t\xi))a, u \right\rangle$$

$$= -\left\langle \xi_{\mathfrak{a}^*}(a), u \right\rangle$$

which together with (6.3.5) shows that

$$\left\langle \Gamma^a_{(g,\alpha)}, \xi \right\rangle = -\left\langle \xi_{\mathfrak{a}^*}(a), u \right\rangle.$$

Thus, for $g = e$ we get

$$\left\langle \Gamma^a_{(e,\exp_A u)}, \xi \right\rangle = -\left\langle \xi_{\mathfrak{a}^*}(a), u \right\rangle. \tag{6.3.6}$$

Returning to (6.3.2), this shows that it suffices to find $u \in \mathfrak{a}$ such that $\langle \mu' - \mu, \xi \rangle = \langle \xi_{\mathfrak{a}^*}(a), u \rangle$ for all $\xi \in \mathfrak{g}$; the desired group element in A is then $\alpha := \exp_A u$. Since $\mu' - \mu \in \mathfrak{g}_a^\circ$, this relation is proved if we show that the map $u \in \mathfrak{a} \mapsto F(\cdot, u) \in \mathfrak{g}_a^\circ$ is onto, where the bilinear map $F : \mathfrak{g} \times \mathfrak{a} \to \mathbb{R}$ is given by

$$F(\xi, u) := \langle \xi_{\mathfrak{a}^*}(a), u \rangle, \quad \text{for} \quad \xi \in \mathfrak{g} \quad \text{and} \quad u \in \mathfrak{a}.$$

First, note that the range of $F(\cdot, u)$ is indeed in $\mathfrak{g}_a^\circ$. Second, if ξ is in the range of the map $v \mapsto F(\cdot, v)$ then for any $v \in \mathfrak{a}$ we have

$$0 = \langle \xi, F(\cdot, v) \rangle = F(\xi, v) = \langle \xi_{\mathfrak{a}^*}(a), v \rangle$$

which is equivalent to $\xi \in \mathfrak{g}_a$. Now argue as in Lemma 4.2.7, where we used the following fact from linear algebra. Let E and F be vector spaces, and $F_0 \subset F$ 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°. In our case we take $E = \mathfrak{a}$, $F = \mathfrak{g}$, $F_0 = \mathfrak{g}_a$, and $T : \mathfrak{a} \to \mathfrak{g}^*$ given by $T(v) := F(\cdot, v)$. The previous two remarks are precisely the verification of these conditions and so we conclude that $v \mapsto F(\cdot, v)$ is onto $\mathfrak{g}_a^\circ$, as required. ∎

Remarks. (i) Note that the proof shows that the map $\alpha \in A \mapsto \Gamma^a_{(e,\alpha)} \in \mathfrak{g}_a^\circ$ is onto for any $a \in \mathfrak{a}^*$. If A is a torus, the image of this map is a compact subgroup of $(\mathfrak{g}_a^\circ, +)$ and hence must equal the trivial group $\{0\}$, which is

equivalent to the statement that $\mathfrak{g}_a = \mathfrak{g}$ for any $a \in \mathfrak{a}^*$. This is indeed the case for any action of a Lie group G on a torus $\mathbb{T}^n$ by homomorphisms.

To see this, recall that the group of automorphisms of $\mathbb{T}^n$ equals $\mathrm{GL}(n, \mathbb{Z})$. This fact is proved in the following way. It is well known that any endomorphism of the circle $\mathbb{T}^1 := \mathbb{R}/\mathbb{Z}$ has the form $x \mapsto nx \pmod 1$ (see e.g. Adams [1969], Proposition 3.74) and hence $\mathrm{End}(\mathbb{T}^1) \cong \mathbb{Z}$. Also, a standard theorem in group theory states that if A is any Abelian group, then the endomorphism ring $\mathrm{End}(A^n)$ is isomorphic to the matrix ring $M_n(\mathrm{End}(A))$. If $A = \mathbb{T}^1$ this shows that the endomorphism ring of $\mathbb{T}^n$ is isomorphic to $\mathfrak{gl}(n, \mathbb{Z})$. Thus the automorphism group of $\mathbb{T}^n$ is isomorphic to $\mathrm{GL}(n, \mathbb{Z})$.

Next, if G acts smoothly on $\mathbb{T}^n$ by group homomorphisms, since the group of automorphisms of $\mathbb{T}^n$ equals $\mathrm{GL}(n, \mathbb{Z})$, the connected component of G acts trivially on $\mathbb{T}^n$. Thus the induced G-actions on the Lie algebra of $\mathbb{T}^n$ and its dual are also trivial which shows that $\mathfrak{g}_a = \mathfrak{g}$ for any a in the dual.

(ii) Note that if we take $\phi(g) = \mathrm{id}_\mathrm{G}$ for every $g \in G$, then M is the central extension of G, and we conclude that *central extensions satisfy the stages hypothesis of reduction by stages*. Similarly, *this hypothesis holds for semidirect products with Abelian Lie groups*. We shall show in the following section that it holds for general semidirect products.

6.4   The Semidirect Product of Two Groups

In the previous section we have seen, as a particular case of general group extensions, that semidirect products of a Lie group with an Abelian Lie group satisfy the stages hypothesis. In this section we shall prove that this is the case for general semidirect products of two not necessarily Abelian Lie groups. We also show that the curvature for the first stage reduction is zero. Thus, there are no magnetic terms in the first stage reduction, so that the second stage reduction can be carried out using standard cotangent bundle reduction.

Generalities on Semidirect Products.   Let G and H be Lie groups and $\phi : G \times H \to H$ a smooth left action of G on H by Lie group homomorphisms. Denote, as usual, $\phi(g)(h) = g \cdot h$, for $g \in G$ and $h \in H$. The **semidirect product group** $G \circledS H$ is the manifold $G \times H$ endowed with the multiplication

$$(g_1, h_1)(g_2, h_2) := (g_1 g_2, h_1(g_1 \cdot h_2)), \qquad (6.4.1)$$

where $g_1, g_2 \in G$, $h_1, h_2 \in H$. The neutral element is (e_G, e_H), where e_G and e_H are the neutral elements of G and H respectively, and $(g, h)^{-1} = (g^{-1}, g^{-1} \cdot h^{-1})$. The subgroup $\{e_G\} \times H$, which is isomorphic to H, is a normal subgroup of $G \circledS H$.

If $\mathfrak{g}$ and $\mathfrak{h}$ are the Lie algebras of G and H respectively, then the Lie algebra of $G\,\circledS\,H$ is the semidirect product $\mathfrak{g}\,\circledS\,\mathfrak{h}$ whose underlying vector space is $\mathfrak{g} \times \mathfrak{h}$ and its bracket is given by

$$[(\xi_1, \eta_1), (\xi_2, \eta_2)] = ([\xi_1, \xi_2], \xi_1 \cdot \eta_2 - \xi_2 \cdot \eta_1 + [\eta_1, \eta_2]), \tag{6.4.2}$$

where $\xi_1, \xi_2 \in \mathfrak{g}$, $\eta_1, \eta_2 \in \mathfrak{h}$, and $\xi_1 \cdot \eta_2, \xi_2 \cdot \eta_1 \in \mathfrak{h}$ denotes the Lie algebra action of $\xi_1, \xi_2 \in \mathfrak{g}$ on $\eta_2, \eta_1 \in \mathfrak{h}$, respectively.

This Lie algebra action is naturally induced by ϕ in the following manner. For every $g \in G$, $\phi(g) : H \to H$ is a Lie group automorphism of H whose derivative $\tilde{\phi}(g) := T_{e_H}\phi(g) : \mathfrak{h} \to \mathfrak{h}$ is a Lie algebra automorphism. In this way one obtains a Lie group homomorphism $\tilde{\phi} : g \in G \mapsto \tilde{\phi}(g) \in \mathrm{Aut}(\mathfrak{h})$ from G to the Lie group $\mathrm{Aut}(\mathfrak{h})$ of Lie algebra automorphisms of $\mathfrak{h}$, which induces a Lie algebra homomorphism $\tilde{\phi}' := T_{e_G}\tilde{\phi} : \mathfrak{g} \to \mathrm{aut}(\mathfrak{h})$ from $\mathfrak{g}$ to the Lie algebra $\mathrm{aut}(\mathfrak{h})$ of derivations of $\mathfrak{h}$. Now set $\xi \cdot \eta := \tilde{\phi}'(\xi)(\eta)$ if $\xi \in \mathfrak{g}$ and $\eta \in \mathfrak{h}$.

The smooth left G-action $\phi : G \times H \to H$ on H induces a smooth left $\mathfrak{g}$-action $\mathfrak{g} \times H \to TH$ on H given by taking the infinitesimal generator of the action, that is, $(\xi, h) \in \mathfrak{g} \times H \mapsto \xi_H(h) \in T_hH$. Note that the map $\xi \in \mathfrak{g} \mapsto \bar{\phi}(\xi) := \xi_H \in \mathfrak{X}(H)$ is a Lie algebra anti-homomorphism. Since the action ϕ is by Lie group homomorphisms of H, the vector fields ξ_H satisfy additional conditions. We have

$$\xi_H(hh') = T_{h'}L_h\left(\xi_H(h')\right) + T_hR_{h'}\left(\xi_H(h)\right), \text{ for all } h, h' \in H \text{ and } \xi \in \mathfrak{g}.$$

Putting here $h' = e_H$ yields $\xi_H(e_H) = 0$ for all $\xi \in \mathfrak{g}$. Using this and putting $h' = h^{-1}$ gives

$$\xi_H(h^{-1}) = -\left(T_{e_H}L_{h^{-1}} \circ T_hR_{h^{-1}}\right)\left(\xi_H(h)\right) \quad \text{for all} \quad h \in H \quad \text{and} \quad \xi \in \mathfrak{g}.$$

With these notations, the adjoint action of $G\,\circledS\,H$ on $\mathfrak{g}\,\circledS\,\mathfrak{h}$ is given by

$$\mathrm{Ad}_{(g,h)}(\xi, \eta) = \left(\mathrm{Ad}_g \xi, \left(\mathrm{Ad}_h \circ \tilde{\phi}(g) \right)(\eta) + T_{h^{-1}}L_h\left((\mathrm{Ad}_g \xi)_H(h^{-1}) \right) \right),$$

where $\xi \in \mathfrak{g}$, $\eta \in \mathfrak{h}$, $g \in G$, $h \in H$. To compute the coadjoint action one has to introduce one more notation. Given a Lie algebra anti-homomorphism $F : \mathfrak{g} \to \mathfrak{X}(H)$ and $h \in H$, denote $F^{\#}(h) : \mathfrak{g} \to T_hH$ the linear map given by $F^{\#}(h)(\xi) := F(\xi)(h)$, for any $\xi \in \mathfrak{g}$. Let $F^{\#}(h)^* : T_h^*H \to \mathfrak{g}^*$ denote its dual map. With this notation the coadjoint action of $G\,\circledS\,H$ on $(\mathfrak{g}\,\circledS\,\mathfrak{h})^* = \mathfrak{g}^* \times \mathfrak{h}^*$ is given by

$$\mathrm{Ad}^*_{(g,h)^{-1}}(\mu, \nu)$$
$$= \left(\mathrm{Ad}^*_{g^{-1}} \mu + \left(\bar{\phi} \circ \mathrm{Ad}_{g^{-1}}\right)^{\#}(g^{-1} \cdot h)^*\left(T^*_{g^{-1}\cdot h}L_{g^{-1}\cdot h^{-1}}\nu\right), \right.$$
$$\left. \tilde{\phi}(g^{-1})^* \mathrm{Ad}^*_{g^{-1}\cdot h^{-1}} \nu \right), \quad (6.4.3)$$

where $\mu \in \mathfrak{g}^*$, $\nu \in \mathfrak{h}^*$, $g \in G$, and $h \in H$. The coadjoint action of $\mathfrak{g} \circledS \mathfrak{h}$ on $(\mathfrak{g} \circledS \mathfrak{h})^*$ is given by

$$\mathrm{ad}^*_{(\xi,\eta)}(\mu,\nu) = \left(\mathrm{ad}^*_\xi \mu - \left(\tilde{\phi}'_\eta \right)^* \nu, \, \tilde{\phi}'(\xi)^* \nu + \mathrm{ad}^*_\eta \nu \right), \tag{6.4.4}$$

where $\xi \in \mathfrak{g}$, $\eta \in \mathfrak{h}$, $\mu \in \mathfrak{g}^*$, $\nu \in \mathfrak{h}^*$, $\tilde{\phi}'(\xi)^* : \mathfrak{h}^* \to \mathfrak{h}^*$ is the dual of the Lie algebra derivation $\tilde{\phi}'(\xi) : \mathfrak{h} \to \mathfrak{h}$, and $\left(\tilde{\phi}'_\eta \right)^* : \mathfrak{h}^* \to \mathfrak{g}^*$ is the dual of the Lie algebra homomorphism $\tilde{\phi}'_\eta : \mathfrak{g} \to \mathfrak{h}$ given by $\tilde{\phi}'_\eta(\xi) := \tilde{\phi}'(\xi)(\eta) = \xi \cdot \eta$. Note that the action of $\mathfrak{g} \circledS \mathfrak{h}$ on $\{0\} \times \mathfrak{h}^*$ is given by the second component of (6.4.4). Similarly, the action of $G \circledS H$ on $\{e_G\} \times H$ is given by the second component of (6.4.3).

From the formula (6.4.2) of the bracket on $\mathfrak{g} \circledS \mathfrak{h}$, it follows that the $(\pm)$-Lie-Poisson bracket on $(\mathfrak{g} \circledS \mathfrak{h})^*$ is given by

$$\{F, K\} = \pm \left\langle \mu, \left[\frac{\delta F}{\delta \mu}, \frac{\delta K}{\delta \mu} \right] \right\rangle \pm \left\langle \nu, \left[\frac{\delta F}{\delta \nu}, \frac{\delta K}{\delta \nu} \right] \right\rangle$$
$$\mp \left\langle \nu, \frac{\delta K}{\delta \mu} \cdot \frac{\delta F}{\delta \nu} \right\rangle \pm \left\langle \nu, \frac{\delta F}{\delta \mu} \cdot \frac{\delta K}{\delta \nu} \right\rangle \tag{6.4.5}$$

where $F, K \in C^\infty \left((\mathfrak{g} \circledS \mathfrak{h})^* \right)$ and $(\mu, \nu) \in \mathfrak{g}^* \times \mathfrak{h}^*$.

Proof of the Stages Hypothesis.   Now we shall prove that every element $\sigma := (\mu, \nu) \in (\mathfrak{g} \circledS \mathfrak{h})^* = \mathfrak{g}^* \times \mathfrak{h}^*$ satisfies the stages hypothesis in Theorem 5.2.9. In the notation used there we have $\mathfrak{m} = \mathfrak{g} \circledS \mathfrak{h}$ and $\mathfrak{n} = \{0\} \times \mathfrak{h}$. So let $\sigma' := (\mu', \nu')$ be any other element of $(\mathfrak{g} \circledS \mathfrak{h})^*$ such that $\sigma|_\mathfrak{n} = \sigma'|_\mathfrak{n}$. This means, in the present situation, that $\nu = \nu'$. Therefore $\sigma = (\mu, \nu)$ and $\sigma' = (\mu', \nu)$.

Next we compute $\mathfrak{m}_{\sigma|_\mathfrak{n}} = (\mathfrak{g} \circledS \mathfrak{h})_{(0,\nu)}$. By (6.4.4) we have

$$(\mathfrak{g} \circledS \mathfrak{h})_{(0,\nu)} = \{ (\xi, \eta) \in \mathfrak{g} \times \mathfrak{h} \mid \tilde{\phi}'(\xi)^* \nu + \mathrm{ad}^*_\eta \nu = 0 \}.$$

The second assumption in the stages hypothesis is that $\sigma' - \sigma = (\mu' - \mu, 0) \in \left((\mathfrak{g} \circledS \mathfrak{h})_{(0,\nu)} \right)^\circ$ which is equivalent to $\mu' - \mu \in \mathfrak{k}^\circ$, where $\mathfrak{k} := \{ \zeta \in \mathfrak{g} \mid \tilde{\phi}'(\zeta)^* \nu \in \mathfrak{h} \cdot \nu \}$ and $\mathfrak{h} \cdot \nu := \{ \mathrm{ad}^*_\eta \nu \mid \eta \in \mathfrak{h} \}$ is the tangent space to the coadjoint H-orbit through $\nu \in \mathfrak{h}^*$ at ν. We also have

$$(G \circledS H)_{(0,\nu)} = \{ (g, h) \in G \times H \mid \tilde{\phi}(g)^* \mathrm{Ad}^*_h \nu = \nu \}.$$

To prove the stages hypothesis, we need to make a detour and establish a few formulas. First we relate $\xi \cdot \eta$ to $\xi_H(\exp_H \eta)$ for $\xi \in \mathfrak{g}$ and $\eta \in \mathfrak{h}$. To do this, we begin with the obvious relation $\phi(g)(\exp_H \eta) = \exp_H \left(T_{e_H} \phi(g)(\eta) \right)$

in which we set $g = \exp_G(t\xi)$ and then take the t-derivative at $t = 0$ to get

$$
\begin{aligned}
\xi_H(\exp_H \eta) &= \left. \frac{d}{dt} \right|_{t=0} \phi(\exp_G(t\xi))(\exp_H \eta) \\
&= \left. \frac{d}{dt} \right|_{t=0} \exp_H \left(T_{e_H} \phi(\exp_G(t\xi))(\eta) \right) \\
&= T_\eta \exp_H (\xi \cdot \eta) \\
&= T_{e_H} L_{\exp_H \eta} \sum_{n=0}^{\infty} \frac{(-1)^n}{(n+1)!} \operatorname{ad}_\eta^n (\xi \cdot \eta),
\end{aligned}
$$

by using the standard formula for the derivative of the exponential map at any point $\eta \in \mathfrak{h}$ (see, e.g. [HRed]). However, since $\mathfrak{g}$ acts on $\mathfrak{h}$ by derivations, we have

$$
0 = \xi \cdot [\eta, \eta] = [\xi \cdot \eta, \eta] + [\eta, \xi \cdot \eta] = -2[\eta, \xi \cdot \eta] = -2 \operatorname{ad}_\eta (\xi \cdot \eta)
$$

which shows that all terms except the first one in the series above vanish. Therefore we get

$$
T_{\exp_H \eta} L_{\exp_H(-\eta)} \left(\xi_H(\exp_H \eta) \right) = \xi \cdot \eta = \tilde{\phi}'(\xi)(\eta). \tag{6.4.6}
$$

Second, we shall compute the coadjoint action for group elements of the form $(e_G, h) \in G \circledS H$. For $\eta \in \mathfrak{h}$, we have by (6.4.3)

$$
\operatorname{Ad}^*_{(e_G,h)^{-1}}(\mu, \nu) = \left(\mu + \bar{\phi}^\#(h)^* \left(T_h^* L_{h^{-1}} \nu \right), \nu \right).
$$

To compute this explicitly, let $\xi \in \mathfrak{g}$ be arbitrary. We have by (6.4.6)

$$
\begin{aligned}
\left\langle \bar{\phi}^\#(h)^* \left(T_h^* L_{h^{-1}} \nu \right), \xi \right\rangle &= \left\langle T_h^* L_{h^{-1}} \nu, \bar{\phi}^\#(h)(\xi) \right\rangle \\
&= \left\langle T_h^* L_{h^{-1}} \nu, \bar{\phi}(\xi)(h) \right\rangle = \left\langle T_h^* L_{h^{-1}} \nu, \xi_H(h) \right\rangle \\
&= \left\langle \nu, T_h L_{h^{-1}} \left(\xi_H(h) \right) \right\rangle = \left\langle \nu, \tilde{\phi}'(\xi)(\eta) \right\rangle \\
&= \left\langle \left(\tilde{\phi}'_\eta \right)^* \nu, \xi \right\rangle
\end{aligned}
$$

and hence

$$
\operatorname{Ad}^*_{(e_G, \exp_H \eta)^{-1}}(\mu, \nu) = \left(\mu + \left(\tilde{\phi}'_\eta \right)^* \nu, \nu \right). \tag{6.4.7}
$$

Using this formula we shall prove the stages hypothesis by showing that there is an element

$$
(e_G, h) \in \{e_G\} \times H_\nu \subset \left(\{e_G\} \times H_\nu \right) \cdot \left((G \circledS H)_{(0,\nu)} \right)_{\sigma|_{(\mathfrak{g} \circledS \mathfrak{h})_{(0,\nu)}}}
$$

such that $\sigma' = \operatorname{Ad}^*_{(e_G,h)^{-1}} \sigma$. We shall search for $h \in H_\nu$ of the form $h = \exp_H \eta$ for some $\eta \in \mathfrak{h}_\nu$. By (6.4.3) and (6.4.7) this relation is equivalent to

$$
\mu' - \mu = \left(\tilde{\phi}'_\eta \right)^* \nu.
$$

However, by the second assumption in the stages hypothesis, $\mu' - \mu \in \mathfrak{k}°$. Just as a check, let's show that the right hand side is also an element of $\mathfrak{k}°$. Indeed, if $\zeta \in \mathfrak{k}$, that is, $\tilde{\phi}'(\zeta)^* \nu \in \mathfrak{h} \cdot \nu$, there exists some $\eta' \in \mathfrak{h}$ such that $\tilde{\phi}'(\zeta)^* \nu = \mathrm{ad}^*_{\eta'} \nu$ and hence

$$\left\langle \left(\tilde{\phi}'_\eta \right)^* \nu, \zeta \right\rangle = \left\langle \nu, \tilde{\phi}'_\eta(\zeta) \right\rangle = \left\langle \nu, \tilde{\phi}'(\zeta)(\eta) \right\rangle = \left\langle \tilde{\phi}'(\zeta)^* \nu, \eta \right\rangle$$
$$= \left\langle \mathrm{ad}^*_{\eta'} \nu, \eta \right\rangle = - \left\langle \mathrm{ad}^*_\eta \nu, \eta' \right\rangle = 0$$

since $\eta \in \mathfrak{h}_\nu$.

Thus the stages hypothesis is verified if we can show that the map $\eta \in \mathfrak{h}_\nu \mapsto \left(\tilde{\phi}'_\eta \right)^* \nu \in \mathfrak{k}°$ is surjective. To do this, we proceed as in the proofs of Lemma 4.2.7 and Theorem 6.3.2. We use, as before, the following statement. Let E and F be vector spaces, and $F_0 \subset F$ 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* $F_0°$.

In our case, we choose $E = \mathfrak{h}_\nu$, $F = \mathfrak{g}$, $F_0 = \mathfrak{k}$, and we let $T : \mathfrak{h}_\nu \to \mathfrak{g}^*$ be defined by $T(\eta) = \left(\tilde{\phi}'_\eta \right)^* \nu$. We have shown above that the range of T lies in $F_0° = \mathfrak{k}°$. Next, assume that $\xi \in \mathfrak{g}$ annihilates the range of T, that is, for any $\eta \in \mathfrak{h}_\nu$ we have

$$0 = \left\langle \left(\tilde{\phi}'_\eta \right)^* \nu, \xi \right\rangle = \left\langle \tilde{\phi}'(\xi)^* \nu, \eta \right\rangle$$

which is equivalent to $\tilde{\phi}'(\xi)^* \nu \in (\mathfrak{h}_\nu)° = \mathfrak{h} \cdot \nu$. However, this is equivalent to $\xi \in F_0 = \mathfrak{k}$, as required. ∎

We summarize this discussion in the following theorem.

6.4.1 Theorem. *Let G and H be two arbitrary Lie groups with Lie algebras $\mathfrak{g}$ and $\mathfrak{h}$, respectively. Let $G \circledS H$ be their semidirect product defined by a left G-action on H by Lie group homomorphisms. Any element of the dual $(\mathfrak{g} \circledS \mathfrak{h})^*$ of the semidirect product Lie algebra $\mathfrak{g} \circledS \mathfrak{h}$ of $G \circledS H$ satisfies the stages hypothesis of Definition 5.2.8.*

The Mechanical Connection. Let $\langle\!\langle \cdot, \cdot \rangle\!\rangle_{\mathfrak{g}}$ and $\langle\!\langle \cdot, \cdot \rangle\!\rangle_{\mathfrak{h}}$ be two positive definite inner products on the Lie algebras $\mathfrak{g}$ and $\mathfrak{h}$, respectively. Then

$$\langle\!\langle (\xi_1, \eta_1), (\xi_2, \eta_2) \rangle\!\rangle_{\mathfrak{s}} = \langle\!\langle \xi_1, \xi_2 \rangle\!\rangle_{\mathfrak{g}} + \langle\!\langle \eta_1, \eta_2 \rangle\!\rangle_{\mathfrak{h}}, \tag{6.4.8}$$

for any $(\xi_1, \eta_1), (\xi_2, \eta_2) \in \mathfrak{s}$, defines a positive definite inner product on $\mathfrak{s}$. Since the spaces $\mathfrak{g} \times \{0\}$ and $\{0\} \times \mathfrak{h}$ are orthogonal, the orthogonal $\langle\!\langle , \rangle\!\rangle_{\mathfrak{s}}$-projection $\mathcal{P}_{\mathfrak{h}} : \mathfrak{s} = \mathfrak{g} \circledS \mathfrak{h} \to \mathfrak{h}$ is the projection on the second factor.

Extend the inner product (6.4.8) on $\mathfrak{s}$ to a right–invariant Riemannian metric on S by setting

$$\langle\!\langle (X_g, U_h), (Y_g, V_h) \rangle\!\rangle_{(g,v)}$$
$$= \langle\!\langle T_{(g,h)} R_{(g,h)^{-1}}(X_g, U_h), T_{(g,h)} R_{(g,h)^{-1}}(Y_g, V_h) \rangle\!\rangle_{\mathfrak{s}}, \tag{6.4.9}$$

where $(g, h) \in S$, $(X_g, U_h), (Y_g, V_h) \in T_{(g,h)}S$, and $R_{(g,h)}$ is right translation on S.

We are now in the framework of Theorems 2.1.15 and 2.1.16.: $\{e_G\} \times H$ is a closed normal subgroup of $G \circledS H$,

$$\mathrm{Ad}^H_{(g,h)^{-1}} \eta = \left(\mathrm{Ad}_{g^{-1} \cdot h^{-1}} \circ \tilde\phi(g^{-1}) \right) \eta, \qquad \mathrm{ad}_{(\xi, \eta)} \eta' = \xi \cdot \eta' + [\eta, \eta'],$$

for any $g \in G$, $h \in H$, $\xi \in \mathfrak{g}$, and $\eta, \eta' \in \mathfrak{h}$. Since

$$T_{(g,h)}R_{(g,h)^{-1}}(X_g, U_h) = \left(T_g R_{g^{-1}}(X_g), T_h R_{h^{-1}}(U_h) + T_{h^{-1}}L_h \left(\xi_H(h^{-1}) \right) \right),$$

where $g \in G$, $h \in H$, $X_g \in T_g G$, $U_h \in H$, and $\xi := T_g R_{g^{-1}} X_g \in \mathfrak{g}$, the connection one-form (2.1.15) becomes in this case

$$\mathcal{A}^H(g, h)(X_g, U_h)$$
$$= \left(\mathrm{Ad}_{g^{-1} \cdot h^{-1}} \circ \tilde\phi(g^{-1}) \right) \left(T_h R_{h^{-1}}(U_h) + T_{h^{-1}}L_h \left(\xi_H(h^{-1}) \right) \right)$$
$$= \tilde\phi(g^{-1}) \left(T_h L_{h^{-1}} U_h + T_{h^{-1}}R_h \left(\xi_H(h^{-1}) \right) \right) \tag{6.4.10}$$

since $\mathrm{AD}_{g^{-1} \cdot h^{-1}} \circ \phi(g^{-1}) = \phi(g^{-1}) \circ \mathrm{AD}_{h^{-1}}$, where $\mathrm{AD}_k \, k' := kk'k^{-1}$ for any $k, k' \in H$.

Remark. If $H = V$ is a vector space and $\phi : G \to \mathrm{Aut}(V)$ is a (left) representation, then this formula becomes

$$\mathcal{A}^V(g, v)(X_g, u) = \phi(g^{-1})(u - \xi \cdot v),$$

where $(v, u) \in T_v V$, $X_g \in T_g G$, and $\xi := T_g R_{g^{-1}}(X_g)$, which coincides with (4.3.7) derived earlier.

Returning to the general case of a semidirect product of two non-Abelian Lie groups G and H, recall from Theorem 2.1.15 that the connection $\mathcal{A}^H$ is S-equivariant, that is, (2.1.16) holds, which in this case reads $R^*_{(g,h)}\mathcal{A}^H = \mathrm{Ad}^H_{(g,h)^{-1}} \circ \mathcal{A}^H$. This identity can also be checked directly by hand using the expressions for $\mathcal{A}^H$ and of the adjoint action given above, but this verification is somewhat lengthy.

The Curvature. To compute the curvature, we use the general equation (2.1.17). Recall that this formula is

$$\mathrm{curv}_{\mathcal{A}}(X_g, Y_g)$$
$$= \mathrm{Ad}_{g^{-1}} \left(-\mathrm{ad}_\xi \, \mathcal{P}_{\mathfrak{n}}(\eta) + \mathrm{ad}_\eta \, \mathcal{P}_{\mathfrak{n}}(\xi) + \mathcal{P}_{\mathfrak{n}}[\xi, \eta] + [\mathcal{P}_{\mathfrak{n}}(\xi), \mathcal{P}_{\mathfrak{n}}(\eta)] \right)$$

where $\xi := \theta^R(X_g)$, $\eta := \theta^R(Y_g) \in \mathfrak{g}$. This becomes in our case,

$$\mathrm{curv}_{\mathcal{A}}((u_g, v_h), (u'_g, v'_h))$$
$$= \mathrm{Ad}_{(g,h)^{-1}} \left(-\mathrm{ad}_{(\xi, \eta)} \, \mathcal{P}_{\mathfrak{h}}(\xi', \eta') + \mathrm{ad}_{(\xi', \eta')} \, \mathcal{P}_{\mathfrak{h}}(\xi, \eta) \right.$$
$$\left. + \mathcal{P}_{\mathfrak{h}}[(\xi, \eta), (\xi', \eta')] + [\mathcal{P}_{\mathfrak{h}}(\xi, \eta), \mathcal{P}_{\mathfrak{h}}(\xi', \eta')] \right),$$

where $(\xi, \eta) = T_{(g,h)} R_{(g,h)^{-1}}(u_g, v_h)$ and $(\xi', \eta') = T_{(g,h)} R_{(g,h)^{-1}}(u'_g, v'_h)$. We will show that the terms in the parenthesis on the right hand side of this formula add up to zero. These terms are

$$- \mathrm{ad}_{(\xi, \eta)} \eta' + \mathrm{ad}_{(\xi', \eta')} \eta + (\xi \cdot \eta' - \xi' \cdot \eta + [\eta, \eta']) + [\eta, \eta']$$
$$= (-\xi \cdot \eta' - [\eta, \eta']) + (\xi' \cdot \eta + [\eta', \eta]) + \xi \cdot \eta' - \xi' \cdot \eta + 2[\eta, \eta'] = 0$$

which shows that the curvature of $\mathcal{A}$ vanishes. We obtain, therefore, as in §4.3, the following result.

6.4.2 Theorem.   *The mechanical connection $\mathcal{A}^H$ defined on the right principal H-bundle $S \to G$ by formula* (6.4.10) *is flat.*

In Patrick [1999], reduction by stages is carried out directly for an interesting example of the semidirect product of two nonabelian groups that arises in an analysis of the Landau-Lifshitz equation. The symmetry group is the semidirect product of the special Euclidean group on $\mathbb{R}^3$ and the infinite dimensional group consisting of the space of maps from $\mathbb{R}^3$ into $SO(3)$ that tend to the identity at infinity. That reduction by stages is possible for this example is in fact covered by the preceding theorem.

7
Magnetic Cotangent Bundle Reduction

The introductory chapters have discussed the history and the theory of cotangent bundle reduction in some detail. In particular, §2.2 and §2.3 reviewed the basic *standard* theory of the embedding version of cotangent bundle reduction.

However, the story does not end there. If one begins a reduction by stages program with a cotangent bundle, which is of course an important case (perhaps *the most important* case) for reduction theory, then already after the first stage of reduction, one will arrive (in, for instance, the embedding version of cotangent bundle reduction, as in Theorem 2.2.1) at a new cotangent bundle, but *with a magnetic term*. Thus, if one is going to proceed with a second stage of reduction, one must learn how to reduce cotangent bundles that already have magnetic terms. The main goal of this Chapter is to develop such a theory.

Here is a more detailed outline of how we will proceed. Namely, we start with a cotangent bundle with magnetic terms of the type $(T^*Q, \Omega - \pi_Q^* \mathcal{B})$, and having a symmetry group G that acts by cotangent lift and admits a momentum map. As usual, $\pi_Q : T^*Q \to Q$ is the cotangent bundle projection. The (possibly nonequivariant) momentum map for the action of G has the form $\mathbf{J} = \mathbf{J}_{\mathrm{can}} - \pi_Q^* \phi$, where $\phi : Q \to \mathfrak{g}^*$, which we call a $\mathcal{B}\mathfrak{g}$-potential, satisfies $\mathbf{i}_{\xi_Q} \mathcal{B} = \mathbf{d} \langle \phi, \xi \rangle$. We can then adapt the momentum translation used in the embedding version of the cotangent bundle reduction theorem (see Theorem 2.2.1) by means of a *double translation*.

We first translate elements of T^*Q by a 1-form determined by pairing the $\mathcal{B}\mathfrak{g}$-potential with an arbitrary connection on the bundle $Q \to Q/G$. Composing this with the momentum translation used in Theorem 2.2.1 gives a

vector bundle isomorphism $\mathbf{J}^{-1}(\mu) \simeq \mathbf{J}_{\mathrm{can}}^{-1}(0)$. We will see that this map is equivariant with respect to the group $\widetilde{G}_\mu$, the isotropy group of the point μ for the affine action of G on $\mathfrak{g}^*$ determined by the nonequivariance cocycle σ of the momentum map $\mathbf{J}$. This allows us to quotient the vector bundle map to get a diffeomorphism $\mathbf{J}^{-1}(\mu)/\widetilde{G}_\mu \simeq \mathbf{J}_{\mathrm{can}}^{-1}(0)/\widetilde{G}_\mu$. By Theorem 2.2.1, the latter embeds into the cotangent bundle $T^*(Q/\widetilde{G}_\mu)$.

In §7.2, we explore the idea of Lie-Poisson reduction in the presence of cocycles. One again obtains the standard Lie-Poisson bracket but modified with a Lie algebra two-cocycle. We also show that the symplectic leaves in this structure, which are the affine orbits studied in Theorem 6.2.2, are obtained by symplectic reduction of T^*G with a magnetic term.

7.1 Embedding Magnetic Cotangent Bundle Reduction

This section develops the embedding version of cotangent bundle reduction where we start with a cotangent bundle with magnetic terms. We will develop hypotheses in this situation that will be appropriate for later applications to cotangent bundle reduction by stages.

Cotangent Bundles with Magnetic Terms.   We need to quickly review the theory of cotangent bundles with magnetic terms. Given is a manifold Q and a closed two-form $\mathcal{B}$ on it. Endow the cotangent bundle with the symplectic form $\Omega - \pi_Q^*\mathcal{B}$, where Ω is, as usual, the canonical symplectic form on T^*Q. The Poisson bracket $\{F, H\}_{\mathcal{B}}$ of two smooth functions $F, H : T^*Q \to \mathbb{R}$ defined by this symplectic form has the expression

$$\{F, H\}_{\mathcal{B}} = \{F, H\}^{\mathrm{can}} + (\pi_Q^*\mathcal{B}) \left(X_F^{\mathrm{can}}, X_H^{\mathrm{can}} \right) \tag{7.1.1}$$

where $\{F, H\}^{\mathrm{can}}$ denotes the Poisson bracket associated with the canonical symplectic form on T^*Q and X_H^{can} is the Hamiltonian vector field defined by the function H and the canonical symplectic form Ω. The easiest way to see this is to work in local coordinates. In standard cotangent bundle coordinates we have $\Omega - \pi_Q^*\mathcal{B} = \mathbf{d}q^i \wedge \mathbf{d}p_i - \mathcal{B}_{ij}\mathbf{d}q^i \wedge \mathbf{d}q^j$. The Hamiltonian vector field relative to this symplectic form has, as is readily verified, the expression

$$X_H = \frac{\partial H}{\partial p_i} \frac{\partial}{\partial q^i} - \left(\frac{\partial H}{\partial q^i} - 2\mathcal{B}_{ij} \frac{\partial H}{\partial p_j} \right) \frac{\partial}{\partial p_i}.$$

Computing now $\left(\Omega - \pi_Q^*\mathcal{B} \right) (X_F, X_H)$ yields

$$\frac{\partial F}{\partial q^i} \frac{\partial H}{\partial p_i} - \frac{\partial F}{\partial p_i} \frac{\partial H}{\partial q^i} + \mathcal{B}_{ij} \frac{\partial F}{\partial p_i} \frac{\partial H}{\partial p_j} - \mathcal{B}_{ij} \frac{\partial F}{\partial p_j} \frac{\partial H}{\partial p_i}$$

which is the local expression of (7.1.1).

Momentum Maps for Cotangent Bundles with Magnetic Terms.
A first step in our study of magnetic cotangent bundle reduction will be to
develop expressions for momentum maps in this context.

7.1.1 Theorem. *Let $\mathcal{B}$ be a closed two-form on a connected configuration
manifold Q. Let G act freely and properly on Q (on the right) leaving the
form $\mathcal{B}$ invariant. Consider the cotangent lift of the G-action to the sym-
plectic manifold $(T^*Q, \Omega - \pi_Q^* \mathcal{B})$, where $\pi_Q : T^*Q \to Q$ is the cotangent
bundle projection. Suppose that there is a smooth map $\phi : Q \to \mathfrak{g}^*$ that
satisfies*

$$\mathbf{i}_{\xi_Q} \mathcal{B} = \mathbf{d}\langle \phi, \xi \rangle$$

*for all $\xi \in \mathfrak{g}$. We shall call ϕ a $\mathcal{B}\mathfrak{g}$-**potential**. Then the following hold:*

(i) *The map $\mathbf{J} = \mathbf{J}_{\mathrm{can}} - \phi \circ \pi_Q$, where $\mathbf{J}_{\mathrm{can}}$ is the standard momentum
map for the G-action relative to the canonical symplectic form, is a
momentum map for the cotangent lifted action of G on T^*Q with
symplectic form $\Omega - \pi_Q^* \mathcal{B}$.*

(ii) *The momentum map $\mathbf{J}$ is, in general, not equivariant. Its nonequiv-
ariance $\mathfrak{g}^*$-valued group one-cocycle $\sigma^{T^*Q} : G \to \mathfrak{g}^*$ is given by*

$$\sigma^{T^*Q}(g) = -\phi(q \cdot g) + \mathrm{Ad}_g^*(\phi(q)), \tag{7.1.2}$$

*with the right hand side independent of $q \in Q$. The cocycle identity
is in this case*

$$\sigma^{T^*Q}(gh) = \mathrm{Ad}_h^* \sigma^{T^*Q}(g) + \sigma^{T^*Q}(h) \tag{7.1.3}$$

which is the analog of (1.1.8) for right actions.

(iii) *The level sets of $\mathbf{J}$ are affine subbundles of T^*Q.*

Remarks. (i) If $\mathcal{B} = \mathbf{d}\alpha$, where α is G-invariant, then one can take ϕ to
be

$$\langle \phi, \xi \rangle = -\mathbf{i}_{\xi_Q} \alpha$$

which is readily verified using the invariance identity $\pmb{\mathcal{L}}_{\xi_Q} \alpha = 0$ along with
the Cartan formula for the Lie derivative.

(ii) If $Q = G$ and the magnetic symplectic structure comes from a first
step reduction of a central extension given by a smooth group two-cocycle,
then ϕ always exists, as will be proved later on (see equation (8.2.6)).
However, one can consider the general problem of reduction of a magnetic
cotangent bundle of a Lie group G, where the magnetic form on the group
evaluated at the identity is a Lie algebra two-cocycle that does not integrate
to a smooth group two-cocycle. In this case, one has a central extension
$\widehat{G}$ of G which is the connected simply connected Lie group with Lie alge-
bra $\widehat{\mathfrak{g}}$, but the multiplication law is not explicitly known and the manifold
underlying this group is, in general, not the product of G with $\mathbb{R}$. In this

case, the theorem above does not apply and one is forced to work with **cylinder valued momentum maps** and the corresponding regular and singular reduction theory developed in Ortega and Ratiu [2006a,b]. This raises the problem of reduction by stages for cylinder valued momentum maps that can be developed along the lines developed earlier. In the examples presented later on we shall not encounter such situations and the statement above suffices for our present purposes.

Proof.   To prove (i), let $\xi \in \mathfrak{g}$ and let ξ_{T^*Q} be the corresponding infinitesimal generator on T^*Q. By construction, the vector fields ξ_{T^*Q} and ξ_Q are π_Q-related; that is, $T\pi_Q \circ \xi_{T^*Q} = \xi_Q \circ \pi_Q$. By definition of the canonical momentum map,

$$\mathbf{i}_{\xi_{T^*Q}}\Omega = \mathbf{d}\langle \mathbf{J}_{\mathrm{can}}, \xi \rangle. \tag{7.1.4}$$

By the assumed property of ϕ, for $w \in T_{\alpha_q}(T^*Q)$, we have

$$
\begin{aligned}
\left(\mathbf{i}_{\xi_{T^*Q}}\pi_Q^*\mathcal{B}\right)(\alpha_q)(w) &= \mathcal{B}(q)\left((T\pi_Q \circ \xi_{T^*Q})(\alpha_q), T_{\alpha_q}\pi_Q(w)\right) \\
&= \mathcal{B}(q)\left(\xi_Q(q), T_{\alpha_q}\pi_Q(w)\right) \\
&= \mathbf{d}\langle \phi, \xi \rangle(q)\left(T_{\alpha_q}\pi_Q(w)\right) \\
&= \mathbf{d}\left(\langle \phi, \xi \rangle \circ \pi_Q\right)(\alpha_q)(w),
\end{aligned}
$$

that is, $\mathbf{i}_{\xi_{T^*Q}}\pi_Q^*\mathcal{B} = \mathbf{d}\langle \phi \circ \pi_Q, \xi \rangle$. Subtracting this result from equation (7.1.4) proves (i).

To prove (ii), we compute σ^{T^*Q} from the definition (the version for *right* actions of (1.1.6)) as follows:

$$
\begin{aligned}
\sigma^{T^*Q}(g, \alpha_q) &= \mathbf{J}(\alpha_q \cdot g) - \mathrm{Ad}_g^* \mathbf{J}(\alpha_q) \\
&= \left(\mathbf{J}_{\mathrm{can}}(\alpha_q \cdot g) - \mathrm{Ad}_g^*(\mathbf{J}_{\mathrm{can}}(\alpha_q))\right) \\
&\quad - \left(\pi_Q^*\phi(\alpha_q \cdot g) - \mathrm{Ad}_g^*((\pi_Q^*\phi)(\alpha_q))\right) \\
&= -\phi(q \cdot g) + \mathrm{Ad}_g^*(\phi(q));
\end{aligned}
$$

in the last equality we have used the equivariance of $\mathbf{J}_{\mathrm{can}}$ for right actions. By the general theory, $\sigma^{T^*Q}(g, \alpha_q)$ does not depend on α_q because Q and hence T^*Q is connected. Therefore, the expression just given for σ^{T^*Q} is independent of $\alpha_q \in Q$. The cocycle identity (7.1.3) holds since σ^{T^*Q} is a nonequivariance cocycle of a momentum map.

To prove (iii), we first show that $\pi_Q(\mathbf{J}^{-1}(\mu)) = Q$ for any $\mu \in \mathbf{J}(Q)$. Fix $q \in Q$ and $\mu \in \mathbf{J}(T^*Q)$. We must show that there exists an element $\alpha_q \in T_q^*Q$ such that

$$\mathbf{J}_{\mathrm{can}}(\alpha_q) = \mu + \pi_Q^*(\phi)(\alpha_q) = \mu + \phi(q).$$

Because the action of G on Q is free, for all $q \in Q$, the infinitesimal generator map $\Xi_q : \mathfrak{g} \to T_qQ$ defined by $\xi \mapsto \xi_Q(q)$ is injective so that its dual, $\Xi_q^* : T_q^*Q \to \mathfrak{g}^*$ is surjective. However, observe that the restriction of $\mathbf{J}_{\mathrm{can}}$

to any fiber in T^*Q is just the map Ξ_q^*. Since this map is surjective, there must exist some α_q (not necessarily unique) with the desired property.

Let $\mathbf{J}_q := \mathbf{J}|_{T_q^*Q} : T_q^*Q \to \mathfrak{g}^*$. This is an affine map since it is just $\Xi_q^* - \phi(q)$. Therefore $\mathbf{J}_q^{-1}(\mu) = \alpha_q + V_q^0$ where V_q^0 is the annihilator of the space $V_q := \{\xi_Q(q) \mid \xi \in \mathfrak{g}\}$ of vertical vectors in T_qQ and α_q is some covector that satisfies $\mathbf{J}(\alpha_q) = \mu$. ∎

The Affine Action on $\mathfrak{g}^*$. As the general theory states, for nonequivariant momentum maps one needs to work with the affine action on $\mathfrak{g}^*$ induced by the nonequivariance $\mathfrak{g}^*$-valued group one-cocycle σ^{T^*Q}, that is, with the right G-action on $\mathfrak{g}^*$ given by

$$\mu \cdot g := \operatorname{Ad}_g^* \mu + \sigma^{T^*Q}(g) \tag{7.1.5}$$

for $\mu \in \mathfrak{g}^*$ and $g \in G$. Denote by

$$\widetilde{G}_\mu = \{g \in G \mid \mu \cdot g := \operatorname{Ad}_g^* \mu + \sigma^{T^*Q}(g) = \mu\} \tag{7.1.6}$$

the isotropy subgroup for this affine action.

The Lie algebra 2-cocycle. We now compute the real valued Lie algebra 2-cocycle Σ^{T^*Q} associated to the $\mathfrak{g}^*$-valued Lie group 1-cocycle σ^{T^*Q}. We will do this using the general formula

$$\Sigma^{T^*Q}(\xi, \eta) := \langle \mathbf{J}, [\xi, \eta] \rangle + \{\langle \mathbf{J}, \xi \rangle, \langle \mathbf{J}, \eta \rangle\},$$

where $\xi, \eta \in \mathfrak{g}$. This is the *right* action version of (1.1.11).

7.1.2 Lemma. *We have*

$$\Sigma^{T^*Q}(\xi, \eta) = -\langle \phi, [\xi, \eta] \rangle - \mathcal{B}(\xi_Q, \eta_Q). \tag{7.1.7}$$

Proof. In our situation, we compute as follows

$$
\begin{aligned}
\Sigma^{T^*Q}(\xi, \eta) := {}& \langle \mathbf{J}, [\xi, \eta] \rangle + \{\langle \mathbf{J}, \xi \rangle, \langle \mathbf{J}, \eta \rangle\}_{\mathcal{B}} \\
= {}& \langle \mathbf{J}_{\mathrm{can}}, [\xi, \eta] \rangle - \langle \phi \circ \pi_Q, [\xi, \eta] \rangle \\
& + \{\langle \mathbf{J}_{\mathrm{can}}, \xi \rangle - \langle \phi \circ \pi_Q, \xi \rangle, \langle \mathbf{J}_{\mathrm{can}}, \eta \rangle - \langle \phi \circ \pi_Q, \eta \rangle\}_{\mathcal{B}} \\
= {}& \langle \mathbf{J}_{\mathrm{can}}, [\xi, \eta] \rangle - \langle \phi \circ \pi_Q, [\xi, \eta] \rangle \\
& + \{\langle \mathbf{J}_{\mathrm{can}}, \xi \rangle - \langle \phi \circ \pi_Q, \xi \rangle, \langle \mathbf{J}_{\mathrm{can}}, \eta \rangle - \langle \phi \circ \pi_Q, \eta \rangle\}^{\mathrm{can}} \\
& + (\pi_Q^* \mathcal{B}) \left(X^{\mathrm{can}}_{\langle \mathbf{J}_{\mathrm{can}}, \xi \rangle - \langle \phi \circ \pi_Q, \xi \rangle}, X^{\mathrm{can}}_{\langle \mathbf{J}_{\mathrm{can}}, \eta \rangle - \langle \phi \circ \pi_Q, \eta \rangle} \right),
\end{aligned}
$$

where (7.1.1) has been used in the third equality.

Expanding the parentheses in the preceding expression, we get

$$
\begin{aligned}
\Sigma^{T^*Q}(\xi,\eta) \\
= \ & \langle \mathbf{J}_{\mathrm{can}}, [\xi,\eta] \rangle - \langle \phi \circ \pi_Q, [\xi,\eta] \rangle \\
& + \{ \langle \mathbf{J}_{\mathrm{can}}, \xi \rangle, \langle \mathbf{J}_{\mathrm{can}}, \eta \rangle \}^{\mathrm{can}} - \{ \langle \phi \circ \pi_Q, \xi \rangle, \langle \mathbf{J}_{\mathrm{can}}, \eta \rangle \}^{\mathrm{can}} \\
& - \{ \langle \mathbf{J}_{\mathrm{can}}, \xi \rangle, \langle \phi \circ \pi_Q, \eta \rangle \}^{\mathrm{can}} + \{ \langle \phi \circ \pi_Q, \xi \rangle, \langle \phi \circ \pi_Q, \eta \rangle \}^{\mathrm{can}} \\
& + (\pi_Q^* \mathcal{B}) \left(X^{\mathrm{can}}_{\langle \mathbf{J}_{\mathrm{can}},\xi \rangle}, X^{\mathrm{can}}_{\langle \mathbf{J}_{\mathrm{can}},\eta \rangle} \right) - (\pi_Q^* \mathcal{B}) \left(X^{\mathrm{can}}_{\langle \mathbf{J}_{\mathrm{can}},\xi \rangle}, X^{\mathrm{can}}_{\langle \phi \circ \pi_Q,\eta \rangle} \right) \\
& - (\pi_Q^* \mathcal{B}) \left(X^{\mathrm{can}}_{\langle \phi \circ \pi_Q,\xi \rangle}, X^{\mathrm{can}}_{\langle \mathbf{J}_{\mathrm{can}},\eta \rangle} \right) + (\pi_Q^* \mathcal{B}) \left(X^{\mathrm{can}}_{\langle \phi \circ \pi_Q,\xi \rangle}, X^{\mathrm{can}}_{\langle \phi \circ \pi_Q,\eta \rangle} \right) \\
= \ & -\langle \phi \circ \pi_Q, [\xi,\eta] \rangle - \{ \langle \phi \circ \pi_Q, \xi \rangle, \langle \mathbf{J}_{\mathrm{can}}, \eta \rangle \}^{\mathrm{can}} \\
& - \{ \langle \mathbf{J}_{\mathrm{can}}, \xi \rangle, \langle \phi \circ \pi_Q, \eta \rangle \}^{\mathrm{can}} + (\pi_Q^* \mathcal{B}) \left(X^{\mathrm{can}}_{\langle \mathbf{J}_{\mathrm{can}},\xi \rangle}, X^{\mathrm{can}}_{\langle \mathbf{J}_{\mathrm{can}},\eta \rangle} \right).
\end{aligned}
$$

In going from the first equality to the second, we have used right infinitesimal equivariance of $\mathbf{J}_{\mathrm{can}}$ (which guarantees that the sum of the first and the third term is zero), as well as the relation $\{ \langle \phi \circ \pi_Q, \xi \rangle, \langle \phi \circ \pi_Q, \eta \rangle \}^{\mathrm{can}} = 0$ (since $\langle \phi \circ \pi_Q, \xi \rangle$ and $\langle \phi \circ \pi_Q, \eta \rangle$ are functions of only q; this shows that the sixth summand vanishes). For the same reason, we have $T\pi_Q \circ X^{\mathrm{can}}_{\langle \phi \circ \pi_Q,\xi \rangle} = 0$ and $T\pi_Q \circ X^{\mathrm{can}}_{\langle \phi \circ \pi_Q,\eta \rangle} = 0$, which shows that the last three terms vanish.

Since the cocycle does not depend on the point of evaluation, and again using the fact that $\langle \phi \circ \pi_Q, \xi \rangle$ and $\langle \phi \circ \pi_Q, \eta \rangle$ are functions of only q, we can rewrite the above as

$$
\begin{aligned}
\Sigma^{T^*Q}(\xi,\eta) := & -\langle \phi, [\xi,\eta] \rangle - \mathbf{d}\langle \phi, \xi \rangle \left(\mathbb{F}\langle \mathbf{J}_{\mathrm{can}}, \eta \rangle \right) \\
& + \mathbf{d}\langle \phi, \eta \rangle \left(\mathbb{F}\langle \mathbf{J}_{\mathrm{can}}, \xi \rangle \right) + \mathcal{B} \left(\mathbb{F}\langle \mathbf{J}_{\mathrm{can}}, \xi \rangle, \mathbb{F}\langle \mathbf{J}_{\mathrm{can}}, \eta \rangle \right), \quad (7.1.8)
\end{aligned}
$$

where $\mathbb{F}f : T^*Q \to TQ$ denotes the fiber derivative of a function $f : T^*Q \to \mathbb{R}$.

The formula for the canonical momentum map, namely $\langle \mathbf{J}_{\mathrm{can}}, \xi \rangle (\alpha_q) = \langle \alpha_q, \xi_Q(q) \rangle$, implies that $\mathbb{F}\langle \mathbf{J}_{\mathrm{can}}, \xi \rangle = \xi_Q$. Using this, together with the formula $\mathbf{d}\langle \phi, \xi \rangle = \mathbf{i}_{\xi_Q}\mathcal{B}$, we see that the above expression for Σ^{T^*Q} simplifies to the desired formula (7.1.7). ∎

The relationship between σ^{T^*Q} and Σ^{T^*Q} is given by

$$
T_e \sigma^{T^*Q}(\xi) = -\Sigma^{T^*Q}(\xi, \cdot) \tag{7.1.9}
$$

for any $\xi \in \mathfrak{g}$. Note the sign change on the right hand side as compared to (1.1.10) which is due to the fact that here we work with *right* actions.

Formula $(7.1.9)$ is easily proved by taking the g-derivative of the defining relation $\sigma^{T^*Q}(g) = \mathbf{J}(\alpha_q \cdot g) - \mathrm{Ad}_g^* \mathbf{J}(\alpha_q)$ at $g = e$:

$$
\begin{aligned}
\left\langle T_e \sigma^{T^*Q}(\xi), \eta \right\rangle &= \left\langle T_{\alpha_q} \mathbf{J}\left(\xi_{T^*Q}(\alpha_q)\right), \eta \right\rangle - \left\langle \mathrm{ad}_\xi^* \mathbf{J}(\alpha_q), \eta \right\rangle \\
&= \left\langle \mathbf{d}\langle \mathbf{J}, \eta\rangle(\alpha_q), X_{\langle \mathbf{J}, \xi\rangle}(\alpha_q) \right\rangle - \left\langle \mathbf{J}(\alpha_q), [\xi, \eta] \right\rangle \\
&= \{\langle \mathbf{J}, \eta\rangle, \langle \mathbf{J}, \xi\rangle\}(\alpha_q) - \langle \mathbf{J}, [\xi, \eta]\rangle(\alpha_q) \\
&= -\Sigma^{T^*Q}(\xi, \eta),
\end{aligned}
$$

for any $\xi, \eta \in \mathfrak{g}$, where we took into account the fact that the expression on the right hand side is independent of α_q. (It is clear that the argument given above is general and works on any symplectic manifold.)

Affine Isotropy.   From $(7.1.6)$ (the formula for the affine isotropy group $\widetilde{G}_\mu$) and $(7.1.9)$, it follows that the Lie algebra $\widetilde{\mathfrak{g}}_\mu$ of $\widetilde{G}_\mu$ is given by

$$
\begin{aligned}
\widetilde{\mathfrak{g}}_\mu &= \{\xi \in \mathfrak{g} \mid \mathrm{ad}_\xi^* \mu - \Sigma^{T^*Q}(\xi, \cdot) = 0\} \\
&= \{\xi \in \mathfrak{g} \mid \langle \mu + \phi, [\xi, \eta]\rangle + \mathcal{B}(\xi_Q, \eta_Q) = 0 \text{ for all } \eta \in \mathfrak{g}\}, \qquad (7.1.10)
\end{aligned}
$$

where we have used formula $(7.1.7)$ in the second equality.

Shift Maps.   Next, we proceed, by analogy to the embedding version of cotangent bundle reduction, the objective being to identify the reduced spaces as images of an embedding into a cotangent bundle. From this viewpoint we can read off the reduced symplectic form by pulling back, by the embedding, an appropriate form on the cotangent bundle. However, we need to develop a theory that will handle the nonequivariance of the momentum map on T^*Q. As we saw in §2.2 and §2.3, in cotangent bundle reduction theory, a key role is played by momentum shift maps. Naturally, this is also the case in the current setting.

We shall work below with a *right* connection $\mathcal{A} \in \Omega^1(Q; \mathfrak{g})$ on the principal bundle $\pi_{Q,G} : Q \to Q/G$, so the appropriate condition (ii) in Definition 2.1.1 reads

$$
\mathcal{A}(q \cdot g)(v_q \cdot g) = \mathrm{Ad}_{g^{-1}}(\mathcal{A}(q)(v_q)).
$$

for all $q \in Q$, $v_q \in T_q Q$, and $g \in G$. The main result on shift maps is given in the next lemma.

7.1.3 Lemma.   *Let $\mathcal{A} \in \Omega^1(Q; \mathfrak{g})$ be a connection on the right principal bundle $\pi_{Q,G} : Q \to Q/G$ and define the* **one-form associated to $\mathcal{A}$ and ϕ on Q** *by*

$$
\mathcal{A}_\phi(q)(v_q) := \langle \phi(q), \mathcal{A}(q)(v_q)\rangle. \qquad (7.1.11)
$$

Then the induced fiber translation

$$
\mathrm{Shift}_{\mathcal{A},\phi} : T^*Q \to T^*Q \quad \text{given by} \quad \alpha_q \mapsto \alpha_q - \mathcal{A}_\phi(q)
$$

restricted to $\mathbf{J}^{-1}(\mu)$, is an affine bundle isomorphism from $\mathbf{J}^{-1}(\mu)$ to $\mathbf{J}_{\mathrm{can}}^{-1}(\mu)$.

Proof.   Let $\alpha_q \in \mathbf{J}^{-1}(\mu)$. For all $\xi \in \mathfrak{g}$ we have

$$
\begin{aligned}
\langle \mathbf{J}_{\mathrm{can}}(\mathrm{Shift}_{\mathcal{A},\phi}(\alpha_q)), \xi \rangle &= \langle \mathbf{J}_{\mathrm{can}}(\alpha_q - \mathcal{A}_\phi(q)), \xi \rangle \\
&= \langle \alpha_q - \mathcal{A}_\phi(q), \xi_Q(q) \rangle \\
&= \langle \mathbf{J}_{\mathrm{can}}(\alpha_q), \xi \rangle - \langle \mathcal{A}_\phi(q), \xi_Q(q) \rangle \\
&= \langle \mu + \phi(q), \xi \rangle - \langle \phi(q), \mathcal{A}(q)(\xi_Q(q)) \rangle \\
&= \langle \mu + \phi(q), \xi \rangle - \langle \phi(q), \xi \rangle \\
&= \langle \mu, \xi \rangle,
\end{aligned}
\tag{7.1.12}
$$

as required. In the same way one sees that the inverse of the shift map, namely, $\alpha_q \mapsto \alpha_q + \mathcal{A}_\phi(q)$ maps $\mathbf{J}_{\mathrm{can}}^{-1}(\mu)$ into $\mathbf{J}^{-1}(\mu)$. Since both $\mathbf{J}^{-1}(\mu)$ and $\mathbf{J}_{\mathrm{can}}^{-1}(\mu)$ are affine subbundles of T^*Q, this proves that $\mathrm{shift}_{\mathcal{A},\phi} : \mathbf{J}^{-1}(\mu) \to \mathbf{J}_{\mathrm{can}}^{-1}(\mu)$ is an affine bundle isomorphism.   ∎

We next compose this shift map with another one, namely the standard shift map $\mathbf{J}_{\mathrm{can}}^{-1}(\mu) \to \mathbf{J}_{\mathrm{can}}^{-1}(0)$ in cotangent bundle reduction theory, which is given by subtraction with $\mathcal{A}(q)^*(\mu)$ (see Proposition 2.2.5 and (2.2.8)). We will get a momentum fiber translation from $\mathbf{J}^{-1}(\mu)$ to $\mathbf{J}_{\mathrm{can}}^{-1}(0)$.

7.1.4 Theorem.   *Define the **total shift map** $\mathrm{Shift} : T^*Q \to T^*Q$ by*

$$
\mathrm{Shift}(\alpha_q) := \alpha_q - \mathcal{A}_\phi(q) - \mathcal{A}(q)^*\mu.
\tag{7.1.13}
$$

This induces a map $\mathrm{shift} : \mathbf{J}^{-1}(\mu) \to \mathbf{J}_{\mathrm{can}}^{-1}(0)$, *which is a $\widetilde{G}_\mu$-equivariant affine isomorphism and so it induces a smooth diffeomorphism*

$$
\widetilde{\mathrm{shift}} : \mathbf{J}^{-1}(\mu)/\widetilde{G}_\mu \to \mathbf{J}_{\mathrm{can}}^{-1}(0)/\widetilde{G}_\mu
$$

uniquely characterized by the commutativity of the following diagram:

$$
\begin{array}{ccc}
\mathbf{J}^{-1}(\mu) & \xrightarrow{\ \mathrm{shift}\ } & \mathbf{J}_{\mathrm{can}}^{-1}(0) \\[4pt]
\pi_\mu \downarrow & & \downarrow \rho_0 \\[4pt]
\mathbf{J}^{-1}(\mu)/\widetilde{G}_\mu & \xrightarrow[\widetilde{\mathrm{shift}}]{\ } & \mathbf{J}_{\mathrm{can}}^{-1}(0)/\widetilde{G}_\mu.
\end{array}
$$

The vertical arrows in this diagram are quotient maps by the action of $\widetilde{G}_\mu$; note that π_μ is the projection onto the reduced space, whereas the target of ρ_0 is not a reduced space. So the reduced space $\mathbf{J}^{-1}(\mu)/\widetilde{G}_\mu$ is embedded into the smooth manifold $\mathbf{J}_{\mathrm{can}}^{-1}(0)/\widetilde{G}_\mu$.

Proof.   We begin by proving that the one-form $\mathcal{A}_\phi + \mathcal{A}(\cdot)^*\mu \in \Omega^1(Q)$ is $\widetilde{G}_\mu$-invariant. For $q \in Q$, $h \in \widetilde{G}_\mu$, and any $v_{q\cdot h} \in T_{q\cdot h}Q$, we have

$$
\begin{aligned}
\langle(\mathcal{A}_\phi(q) + \mathcal{A}(q)^*\mu)\cdot h, v_{q\cdot h}\rangle &= \langle\mathcal{A}_\phi(q), v_{q\cdot h}\cdot h^{-1}\rangle + \langle\mathcal{A}(q)^*\mu, v_{q\cdot h}\cdot h^{-1}\rangle \\
&= \langle\phi(q), \mathcal{A}(q)(v_{q\cdot h}\cdot h^{-1})\rangle + \langle\mu, \mathcal{A}(q)(v_{q\cdot h}\cdot h^{-1})\rangle \\
&= \langle\phi(q) + \mu, \mathcal{A}((q\cdot h)\cdot h^{-1})(v_{q\cdot h}\cdot h^{-1})\rangle \\
&= \langle\phi(q) + \mu, \mathrm{Ad}_h\,\mathcal{A}(q\cdot h)(v_{q\cdot h})\rangle \\
&= \langle\mathrm{Ad}_h^*\,\phi(q) + \mathrm{Ad}_h^*\,\mu, \mathcal{A}(q\cdot h)(v_{q\cdot h})\rangle. \qquad (7.1.14)
\end{aligned}
$$

On the other hand, by (7.1.2) and taking an arbitrary $\alpha_q \in \mathbf{J}^{-1}(\mu)$, we have

$$
\begin{aligned}
\langle\mathcal{A}_\phi(q\cdot h) + \mathcal{A}(q\cdot h)^*\mu, v_{q\cdot h}\rangle &= \langle\phi(q\cdot h) + \mu, \mathcal{A}(q\cdot h)(v_{q\cdot h})\rangle \\
&= \left\langle\mathrm{Ad}_h^*(\phi(q)) - \sigma^{T^*Q}(h) + \mu, \mathcal{A}(q\cdot h)(v_{q\cdot h})\right\rangle \\
&= \langle\mathrm{Ad}_h^*(\phi(q)) - \mathbf{J}(\alpha_q\cdot h) + \mathrm{Ad}_h^*\,\mathbf{J}(\alpha_q) + \mu, \mathcal{A}(q\cdot h)(v_{q\cdot h})\rangle \\
&= \langle\mathrm{Ad}_h^*(\phi(q)) + \mathrm{Ad}_h^*\,\mu, \mathcal{A}(q\cdot h)(v_{q\cdot h})\rangle \qquad (7.1.15)
\end{aligned}
$$

since $\mathbf{J}(\alpha_q) = \mu$ and $\mathbf{J}^{-1}(\mu)$ is $\widetilde{G}_\mu$-invariant and hence $\alpha_q\cdot h \in \mathbf{J}^{-1}(\mu)$. The two identities (7.1.14) and (7.1.15), arbitrariness of $v_{q\cdot h}$, and surjectivity of $\mathcal{A}(q\cdot h) : T_{q\cdot h}Q \to \mathfrak{g}$ prove the required invariance equality

$$
(\mathcal{A}_\phi(q) + \mathcal{A}(q)^*\mu)\cdot h = \mathcal{A}_\phi(q\cdot h) + \mathcal{A}(q\cdot h)^*\mu.
$$

The proof is finished by noting that $\widetilde{G}_\mu$-equivariance of the map shift is equivalent to $\widetilde{G}_\mu$-invariance of $\mathcal{A}_\phi + \mathcal{A}(\cdot)^*\mu \in \Omega^1(Q)$. ∎

It is worth noting that the $\widetilde{G}_\mu$-equivariance holds for the total shift map, not for the partial one in Lemma 7.1.3; in fact, $\widetilde{G}_\mu$ need not leave $\mathbf{J}_{\mathrm{can}}^{-1}(\mu)$ invariant.

Dropping the two-form $\mathcal{B} + \mathbf{d}(\mathcal{A}_\phi + \mathcal{A}(\cdot)^*\mu)$ to $Q/\widetilde{G}_\mu$.   As we shall see shortly, the two form $\mathcal{B} + \mathbf{d}(\mathcal{A}_\phi + \mathcal{A}(\cdot)^*\mu)$ will generate the magnetic term in magnetic cotangent bundle reduction. The next lemma shows that this two-form on Q in fact always drops to a two-form on $Q/\widetilde{G}_\mu$.

7.1.5 Lemma.   *With the notations and hypotheses of Theorem 7.1.1, the closed two-form $\mathcal{B} + \mathbf{d}(\mathcal{A}_\phi + \mathcal{A}(\cdot)^*\mu)$ on Q drops to a closed two-form $\mathcal{B}_\mu$ on $Q/\widetilde{G}_\mu$; that is,*

$$
\pi_{Q,\widetilde{G}_\mu}^*\,\mathcal{B}_\mu = \mathcal{B} + \mathbf{d}(\mathcal{A}_\phi + \mathcal{A}(\cdot)^*\mu). \qquad (7.1.16)
$$

Proof.   In general, a two-form drops to a quotient space by a group action if it is group invariant and its interior product with any infinitesimal generator vanishes. In our case, observe that $\mathbf{d}(\mathcal{A}_\phi + \mathcal{A}(\cdot)^*\mu)$ is $\widetilde{G}_\mu$-invariant

because $\mathcal{A}_\phi + \mathcal{A}(\cdot)^*\mu$ is $\widetilde{G}_\mu$-invariant by Theorem 7.1.4. We are assuming that $\mathcal{B}$ is G-invariant, and so $\mathcal{B} + \mathbf{d}(\mathcal{A}_\phi + \mathcal{A}(\cdot)^*\mu)$ is clearly $\widetilde{G}_\mu$-invariant.

Thus, the condition for $\mathcal{B} + \mathbf{d}(\mathcal{A}_\phi + \mathcal{A}(\cdot)^*\mu)$ to drop to the quotient is that for all $\eta \in \tilde{\mathfrak{g}}_\mu$

$$\mathbf{i}_{\eta_Q}\left(\mathcal{B} + \mathbf{d}(\mathcal{A}_\phi + \mathcal{A}(\cdot)^*\mu)\right) = 0.$$

From $\boldsymbol{\mathcal{L}}_X = \mathbf{d}\mathbf{i}_X + \mathbf{i}_X\mathbf{d}$ and $\widetilde{G}_\mu$-invariance of $\mathcal{A}_\phi + \mathcal{A}(\cdot)^*\mu$ we get for any $\eta \in \tilde{\mathfrak{g}}_\mu$

$$0 = \boldsymbol{\mathcal{L}}_{\eta_Q}(\mathcal{A}_\phi + \mathcal{A}(\cdot)^*\mu) = \mathbf{i}_{\eta_Q}\mathbf{d}(\mathcal{A}_\phi + \mathcal{A}(\cdot)^*\mu) + \mathbf{d}\mathbf{i}_{\eta_Q}(\mathcal{A}_\phi + \mathcal{A}(\cdot)^*\mu). \quad (7.1.17)$$

However, for any $q \in Q$ we have

$$
\begin{aligned}
\mathbf{i}_{\eta_Q}(\mathcal{A}_\phi + \mathcal{A}(\cdot)^*\mu)(q) &= \langle \mathcal{A}_\phi(q), \eta_Q(q)\rangle + \langle \mathcal{A}(q)^*\mu, \eta_Q(q)\rangle \\
&= \langle \phi(q), \mathcal{A}(\eta_Q)(q)\rangle + \langle \mu, \mathcal{A}(\eta_Q)(q)\rangle \\
&= \langle \phi(q), \eta\rangle + \langle \mu, \eta\rangle,
\end{aligned}
$$

where we have used the fact that $\langle \mu, \mathcal{A}(\eta_Q)(q)\rangle = \langle \mu, \eta\rangle$, which we notice is a constant. Therefore, its differential is

$$\mathbf{d}\mathbf{i}_{\eta_Q}(\mathcal{A}_\phi + \mathcal{A}(\cdot)^*\mu) = \mathbf{d}\langle \phi, \eta\rangle = \mathbf{i}_{\eta_Q}\mathcal{B}. \quad (7.1.18)$$

Substituting (7.1.18) into (7.1.17) gives the desired result. ∎

The Magnetic Symplectic Embedding Theorem. We are now in a position to prove a symplectic embedding theorem for reduction of cotangent bundles with magnetic terms and associated nonequivariant momentum maps for cotangent lifted actions which satisfy the hypotheses of Theorem 7.1.1.

We shall follow the general procedures in Theorem 2.2.1, but since we have connections in our context, we will not make the distinction between the conditions **CBR1** and **CBR2**.

We begin by reviewing the maps that will enter in the proof; this information is recorded in the commutative diagram in Figure 7.1.1.

- Let $\mathbf{J}_{\mathrm{can}}^\mu(\alpha_q) = \mathbf{J}_{\mathrm{can}}(\alpha_q)|\tilde{\mathfrak{g}}_\mu$;

- $i_\mu : \mathbf{J}^{-1}(\mu) \hookrightarrow T^*Q$, $\iota : \mathbf{J}_{\mathrm{can}}^{-1}(0) \hookrightarrow (\mathbf{J}_{\mathrm{can}}^\mu)^{-1}(0)$, $i^\mu : (\mathbf{J}_{\mathrm{can}}^\mu)^{-1}(0) \hookrightarrow T^*Q$, $i_0 = i^\mu \circ \iota : \mathbf{J}_{\mathrm{can}}^{-1}(0) \hookrightarrow T^*Q$ are inclusions;

- $\pi_\mu : \mathbf{J}^{-1}(\mu) \to \mathbf{J}^{-1}(\mu)/\widetilde{G}_\mu$ and $\pi^\mu : (\mathbf{J}_{\mathrm{can}}^\mu)^{-1}(0) \to (\mathbf{J}_{\mathrm{can}}^\mu)^{-1}(0)/\widetilde{G}_\mu$ are projections onto symplectic reduced spaces;

- $\pi_{Q,\widetilde{G}_\mu} : Q \to Q/\widetilde{G}_\mu$ and $\rho_0 : \mathbf{J}_{\mathrm{can}}^{-1}(0) \to \mathbf{J}_{\mathrm{can}}^{-1}(0)/\widetilde{G}_\mu$ are quotient projections; note that $\mathbf{J}_{\mathrm{can}}^{-1}(0)/\widetilde{G}_\mu$ need not be a symplectic manifold since it is not a reduced space;
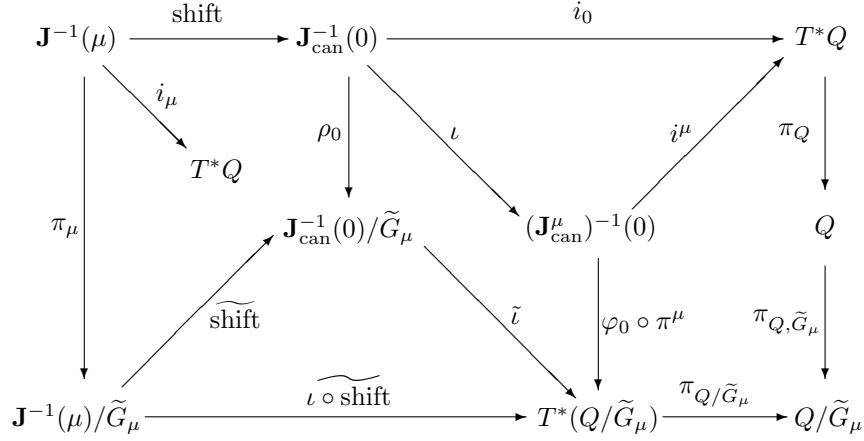
FIGURE 7.1.1. The maps in the proof of Theorem 7.1.6: $\pi_\mu, \rho_0, \pi^\mu, \pi_Q, \pi_{Q,\widetilde{G}_\mu}, \pi_{Q/\widetilde{G}_\mu}$ are projections, i_μ, ι, i^μ, i_0 are inclusions, and a tilde over a map means that it is induced on the quotient.

- $\pi_Q : T^*Q \to Q$ and $\pi_{Q/\widetilde{G}_\mu} : T^*(Q/\widetilde{G}_\mu) \to Q/\widetilde{G}_\mu$ are cotangent bundle projections;

- $\varphi_0 : (\mathbf{J}^\mu_{\mathrm{can}})^{-1}(0)/\widetilde{G}_\mu \to T^*(Q/\widetilde{G}_\mu)$ is a symplectic diffeomorphism, where the reduced space $(\mathbf{J}^\mu_{\mathrm{can}})^{-1}(0)/\widetilde{G}_\mu$ is endowed with its reduced symplectic form and $T^*(Q/\widetilde{G}_\mu)$ with its canonical symplectic form (see §2.2); φ_0 is the $\widetilde{G}_\mu$-quotient of the smooth map $\overline{\varphi}_0 : (\mathbf{J}^\mu_{\mathrm{can}})^{-1}(0) \to T^*(Q/\widetilde{G}_\mu)$ given by (2.2.4), that is, $\overline{\varphi}_0 = \varphi_0 \circ \pi^\mu$ and

$$\langle \overline{\varphi}_0(\alpha_q), T_q \pi_{Q,\widetilde{G}_\mu}(v_q) \rangle := \langle \alpha_q, v_q \rangle,$$

  for any $\alpha_q \in (\mathbf{J}^\mu_{\mathrm{can}})^{-1}(0)$ and $v_q \in T_q Q$;

- shift : $\mathbf{J}^{-1}(\mu) \to \mathbf{J}^{-1}_{\mathrm{can}}(0)$ is the restriction to $\mathbf{J}^{-1}(\mu)$ of the total shift map given by (7.1.13), that is,

$$\mathrm{shift}(\alpha_q) := \alpha_q - \mathcal{A}_\phi(q) - \mathcal{A}(q)^*\mu$$

  for $\alpha_q \in \mathbf{J}^{-1}(\mu)$;

- $\widetilde{\mathrm{shift}} : \mathbf{J}^{-1}(\mu)/\widetilde{G}_\mu \to \mathbf{J}^{-1}_{\mathrm{can}}(0)/\widetilde{G}_\mu$, $\tilde{\iota} : \mathbf{J}^{-1}_{\mathrm{can}}(0)/\widetilde{G}_\mu \to T^*(Q/\widetilde{G}_\mu)$, and $\widetilde{\iota \circ \mathrm{shift}} : \mathbf{J}^{-1}(\mu)/\widetilde{G}_\mu \to T^*(Q/\widetilde{G}_\mu)$ are maps induced on the quotients.

7.1.6 Theorem. *Consider the symplectic reduced space $\mathbf{J}^{-1}(\mu)/\widetilde{G}_\mu$ for the symplectic manifold $(T^*Q, \Omega - \pi_Q^*\mathcal{B})$, where $\mathcal{B}$ is a closed G-invariant*

two-form on the configuration manifold Q satisfying the hypotheses in The-orem 7.1.1. *Let $\mathcal{A} \in \Omega^1(Q; \mathfrak{g})$ be a connection on the right principal bundle $\pi_{Q,G} : Q \to Q/G$ and $\mathcal{B}_\mu \in \Omega(Q/\widetilde{G}_\mu)$ the associated closed two-form de-fined in Lemma* 7.1.5.

(i) *Then $\widetilde{\iota \circ \mathrm{shift}}$ is a symplectic embedding of $\mathbf{J}^{-1}(\mu)/\widetilde{G}_\mu$, with its re-duced symplectic form $\omega_{\mathrm{red}}^\mu$, into the symplectic manifold*

$$\left(T^*(Q/\widetilde{G}_\mu), \Omega_{\mathrm{can}} - \pi_{Q/\widetilde{G}_\mu}^* \mathcal{B}_\mu \right). \tag{7.1.19}$$

(ii) *The image of the map $\widetilde{\iota \circ \mathrm{shift}}$ is the vector subbundle $[T\pi_{Q,\widetilde{G}_\mu}(V)]^\circ$ of $T^*(Q/\widetilde{G}_\mu)$, where $V \subset TQ$ is the vector subbundle consisting of vectors tangent to the G-orbits, that is, its fiber at $q \in Q$ equals $V_q = \{\xi_Q(q) \mid \xi \in \mathfrak{g}\}$, and $^\circ$ denotes the annihilator relative to the natural duality pairing between $T(Q/\widetilde{G}_\mu)$ and $T^*(Q/\widetilde{G}_\mu)$.*

Proof.   To prove (i), we start with the fact that the reduced symplectic form $\omega_{\mathrm{red}}^\mu$ on $\mathbf{J}^{-1}(\mu)/\widetilde{G}_\mu$ is characterized by the identity

$$i_\mu^*(\Omega - \pi_Q^* \mathcal{B}) = \pi_\mu^* \omega_{\mathrm{red}}^\mu, \tag{7.1.20}$$

where Ω is the canonical symplectic form on T^*Q. The characterization of the reduced symplectic form at zero for the $\widetilde{G}_\mu$-action and the symplectic diffeomorphism $\varphi_0 : (\mathbf{J}_{\mathrm{can}}^\mu)^{-1}(0)/\widetilde{G}_\mu \to T^*(Q/\widetilde{G}_\mu)$ yield the identity

$$(\pi^\mu)^* \varphi_0^* \Omega_{\mathrm{can}} = (i^\mu)^* \Omega. \tag{7.1.21}$$

The theorem asserts that

$$\omega_{\mathrm{red}}^\mu = (\widetilde{\iota \circ \mathrm{shift}})^* \left(\Omega_{\mathrm{can}} - \pi_{Q/\widetilde{G}_\mu}^* \mathcal{B}_\mu \right). \tag{7.1.22}$$

To prove this, use the commutativity of the diagram in Figure 7.1.1 and (7.1.21) to compute

$$\pi_\mu^*(\widetilde{\iota \circ \mathrm{shift}})^* \left(\Omega_{\mathrm{can}} - \pi_{Q/\widetilde{G}_\mu}^* \mathcal{B}_\mu \right)$$
$$= \mathrm{shift}^* \iota^* (\pi^\mu)^* \varphi_0^* \left(\Omega_{\mathrm{can}} - \pi_{Q/\widetilde{G}_\mu}^* \mathcal{B}_\mu \right)$$
$$= \mathrm{shift}^* \iota^* (i^\mu)^* \Omega - \mathrm{shift}^* \iota^* (\pi^\mu)^* \varphi_0^* \pi_{Q/\widetilde{G}_\mu}^* (\mathcal{B}_\mu)$$
$$= \mathrm{shift}^* i_0^* \Omega - \mathrm{shift}^* \iota^* (i^\mu)^* \pi_Q^* \pi_{Q,\widetilde{G}_\mu}^* (\mathcal{B}_\mu)$$
$$= \mathrm{shift}^* i_0^* \Omega - \mathrm{shift}^* i_0^* \pi_Q^* \left(\mathcal{B} + \mathbf{d}(\mathcal{A}_\phi + \mathcal{A}(\cdot)^* \mu) \right)$$
$$= - \mathrm{shift}^* i_0^* \mathbf{d}\Theta - i_\mu^* \pi_Q^* \left(\mathcal{B} + \mathbf{d}(\mathcal{A}_\phi + \mathcal{A}(\cdot)^* \mu) \right)$$

since $\pi_Q \circ i_0 \circ \text{shift} = \pi_Q \circ i_\mu$ and where Θ is the canonical one form on T^*Q. Thus we get

$$\pi_\mu^*(\widetilde{\iota \circ \text{shift}})^* \left(\Omega_{\text{can}} - \pi_{Q/\widetilde{G}_\mu}^* \mathcal{B}_\mu \right)$$
$$= -\mathbf{d}\,\text{shift}^*\, i_0^* \Theta - i_\mu^* \pi_Q^* \mathcal{B} - \mathbf{d} i_\mu^* \pi_Q^* (\mathcal{A}_\phi + \mathcal{A}(\cdot)^* \mu). \qquad (7.1.23)$$

We shall compute now the first term of the right hand side explicitly, using the formula for the canonical one-form on T^*Q. For any $\alpha_q \in \mathbf{J}^{-1}(\mu)$ and $V_{\alpha_q} \in T_{\alpha_q} \mathbf{J}^{-1}(\mu) = \ker T_{\alpha_q} \mathbf{J}$, using $\pi_Q \circ i_0 \circ \text{shift} = \pi_Q \circ i_\mu$, we have

$$(\text{shift}^*\, i_0^* \Theta)(\alpha_q)\, (V_{\alpha_q}) = \Theta(\alpha_q - \mathcal{A}_\phi(q) - \mathcal{A}(q)^* \mu) \left(T_{\alpha_q}(i_0 \circ \text{shift})(V_{\alpha_q}) \right)$$
$$= \left\langle \alpha_q - \mathcal{A}_\phi(q) - \mathcal{A}(q)^* \mu, T_{\alpha_q}(\pi_Q \circ i_0 \circ \text{shift})(V_{\alpha_q}) \right\rangle$$
$$= \left\langle \alpha_q - \mathcal{A}_\phi(q) - \mathcal{A}(q)^* \mu, T_{\alpha_q}(\pi_Q \circ i_\mu)(V_{\alpha_q}) \right\rangle$$
$$= \left\langle \alpha_q, T_{\alpha_q} \pi_Q(T_{\alpha_q} i_\mu(V_{\alpha_q})) \right\rangle - \left\langle \mathcal{A}_\phi(q) + \mathcal{A}(q)^* \mu, T_{\alpha_q} \pi_Q(T_{\alpha_q} i_\mu(V_{\alpha_q})) \right\rangle$$
$$= (i_\mu^* \Theta)(\alpha_q)(V_{\alpha_q}) - \left(i_\mu^* \pi_Q^*(\mathcal{A}_\phi + \mathcal{A}(\cdot)^* \mu) \right)(\alpha_q)(V_{\alpha_q}),$$

that is,

$$\text{shift}^*\, i_0^* \Theta = i_\mu^* \Theta - i_\mu^* \pi_Q^*(\mathcal{A}_\phi + \mathcal{A}(\cdot)^* \mu). \qquad (7.1.24)$$

Combining (7.1.23) and (7.1.24) we get

$$\pi_\mu^*(\widetilde{\iota \circ \text{shift}})^* \left(\Omega_{\text{can}} - \pi_{Q/\widetilde{G}_\mu}^* \mathcal{B}_\mu \right) = i_\mu^* \Omega - i_\mu^* \pi_Q^* \mathcal{B}$$
$$= i_\mu(\Omega - \pi_Q^* \mathcal{B}) = \pi_\mu^* \omega_{\text{red}}^\mu$$

by (7.1.20). Since π_μ is a surjective submersion, π_μ^* is injective which proves the statement of the theorem.

Finally we prove the last statement of the theorem. By construction, $V \subset TQ$ is the vertical bundle of the G-principal bundle $\pi_{Q,G} : Q \to Q/G$, that is, V is the vector subbundle of TQ whose fiber at each $q \in Q$ equals $V_q = \{\xi_Q(q) \mid \xi \in \mathfrak{g}\}$. Since $T_q \pi_{Q,\widetilde{G}_\mu} : V_q \to T_{\pi_{Q,\widetilde{G}_\mu}(q)}(Q/\widetilde{G}_\mu)$ has kernel $\{\eta_Q(q) \mid \eta \in \widetilde{\mathfrak{g}}_\mu\}$ whose dimension does not depend on $q \in Q$ (because the G-action is free on Q), it follows that the rank of $T_q \pi_{Q,\widetilde{G}_\mu}|_V$ is constant in q. Therefore, both the kernel and the range of $T\pi_{Q,\widetilde{G}_\mu}|_V$ are subbundles of V and $T(Q/\widetilde{G}_\mu)$ respectively. Thus the annihilator $[T\pi_{Q,\widetilde{G}_\mu}(V)]^\circ$ is a vector subbundle of $T^*(Q/\widetilde{G}_\mu)$.

We shall show that

$$(\widetilde{\iota \circ \text{shift}})((T^*Q)_\mu) = [T\pi_{Q,\widetilde{G}_\mu}(V)]^\circ. \qquad (7.1.25)$$

Indeed, if $\alpha_q \in \mathbf{J}^{-1}(\mu)$, we have, for any $\xi \in \mathfrak{g}$,

$$
\begin{aligned}
\Big\langle (\widetilde{\iota \circ \mathrm{shift}})(\pi_\mu(\alpha_q)), T_q\pi_{Q,\widetilde{G}_\mu}(\xi_Q(q)) \Big\rangle & \\
&= \big\langle (\varphi_0 \circ \pi^\mu \circ \iota \circ \mathrm{shift})(\alpha_q), T_q\pi_{Q,G_\mu}(\xi_Q(q)) \big\rangle \\
&= \big\langle (\varphi_0 \circ \pi^\mu)(\alpha_q - \mathcal{A}_\phi(q) - \mathcal{A}(q)^*\mu), T_q\pi_{Q,G_\mu}(\xi_Q(q)) \big\rangle \\
&= \langle \alpha_q - \mathcal{A}_\phi(q) - \mathcal{A}(q)^*\mu, \xi_Q(q) \rangle \\
&= \langle \mathbf{J}_{\mathrm{can}}(\alpha_q), \xi \rangle - \langle \phi(q), \xi \rangle - \langle \mu, \xi \rangle \\
&= \langle \mathbf{J}(\alpha_q), \xi \rangle - \langle \mu, \xi \rangle = 0
\end{aligned}
\tag{7.1.26}
$$

This shows that the image of $\widetilde{\iota \circ \mathrm{shift}}$ lies in $[T\pi_{Q,\widetilde{G}_\mu}(V)]^\circ$. Conversely, if $\Gamma_{[q]} \in \left[T\pi_{Q,\widetilde{G}_\mu}(V)\right]^\circ$ define $\alpha_q \in T_q^*Q$ by

$$
\langle \alpha_q, v_q \rangle := \langle \Gamma_{[q]}, T_q\pi_{Q,\widetilde{G}_\mu}(v_q) \rangle + \langle \mathcal{A}_\phi(q) + \mathcal{A}(q)^*\mu, v_q \rangle
$$

for any $v_q \in T_qQ$. We claim that $\alpha_q \in \mathbf{J}^{-1}(\mu)$. Indeed, for any $\xi \in \mathfrak{g}$ we have

$$
\begin{aligned}
\langle \mathbf{J}(\alpha_q), \xi \rangle &= \langle \alpha_q, \xi_Q(q) \rangle - \langle \phi(q), \xi \rangle \\
&= \langle \Gamma_{[q]}, T_q\pi_{Q,\widetilde{G}_\mu}(\xi_Q(q)) \rangle + \langle \mathcal{A}_\phi(q) + \mathcal{A}(q)^*\mu, \xi_Q(q) \rangle - \langle \phi(q), \xi \rangle \\
&= 0 + \langle \phi(q), \xi \rangle + \langle \mu, \xi \rangle - \langle \phi(q), \xi \rangle \\
&= \langle \mu, \xi \rangle.
\end{aligned}
$$

From the definition of α_q and a computation similar to that given in (7.1.26), it follows that $(\widetilde{\iota \circ \mathrm{shift}})(\pi_\mu(\alpha_q)) = \Gamma_{[q]}$, which in turn shows that $[T\pi_{Q,G_\mu}(V)]^\circ$ is contained in the range of $\widetilde{\iota \circ \mathrm{shift}}$.   ∎

Remarks.

1. Note that if $\mathcal{B} = 0$ one recovers the embedding cotangent bundle reduction theorem 2.2.1 with α_μ equal to the μ-component of a connection on the principal bundle $\pi_{Q,G} : Q \to Q/G$. Indeed, in this case all hypotheses of Theorem 7.1.6 are trivially satisfied.

2. If, in addition to $\mathcal{B} = 0$, the group G is Abelian, then in this case, $\widehat{G}_\mu = G_\mu = G$ and the embedding is onto; that is, the reduced space is the whole cotangent bundle $T^*(Q/G)$ with the symplectic form $\Omega_{\mathrm{can}} - \pi_{Q/G}^*\beta_\mu$, where β_μ is the two-form on Q/G obtained by dropping $\mathbf{d}\mathcal{A}(\cdot)^*\mu$ to the quotient. This agrees with the result of the standard cotangent bundle reduction theorem for Abelian groups.

3. If $\mathcal{B} \neq 0$ but G is Abelian, the embedding need not be onto because, while $G_\mu = G$, $\widetilde{G}_\mu$ need not equal G.

4. If $Q = G$ (and we are not assuming anything about $\mathcal{B}$), then the bundle V is all of TG and so the image in (iii) is the zero section. Thus, the embedding is onto the zero section of $T^*(G/\widetilde{G}_\mu)$, so that the reduced space in this case is the homogeneous space $G/\widetilde{G}_\mu$ with the symplectic form $-\mathcal{B}_\mu$. This homogeneous space is naturally diffeomorphic to the orbit through μ of the affine action of G on $\mathfrak{g}^*$ given in equation (7.1.5). We will study this orbit in greater detail in the next two sections.

7.2 Magnetic Lie-Poisson and Orbit Reduction

As we have recalled already in the introductory chapters, the ***Lie-Poisson bracket*** on $\mathfrak{g}^*$, the dual of a Lie algebra $\mathfrak{g}$ is the bracket that is naturally induced on the quotient $(T^*G)/G$ by Poisson reduction. This bracket (which comes with a plus or minus sign depending on whether one is doing right or left reduction) is given on two smooth functions on $\mathfrak{g}^*$ by

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

where $\delta f/\delta \mu$ is the derivative of f, but thought of as an element of $\mathfrak{g}$. Recall that Lie–Poisson structures go back to Lie [1890] and are exposed in standard textbooks, such as [MandS].

Another key related reduction theorem is the theorem on reduction to coadjoint orbits given in Theorem 1.2.3. The purpose of this section is to generalize both of these results to the case where we start with the cotangent bundle of a group, but with a magnetic symplectic structure rather than the canonical one. We shall apply these results to the case of central extensions in the following section.

Magnetic Lie-Poisson Reduction.   Our first goal in this section is to derive a magnetic Lie-Poisson reduction theorem. Namely, we shall compute the Poisson structure on $\mathfrak{g}^*$ arising from the *Poisson* reduction of the symplectic manifold

$$(T^*G, \Omega - \pi_G^*\mathcal{B})$$

relative to the cotangent lifted right action of G on itself, and where $\mathcal{B}$ is a closed G-invariant 2-form on G. There is of course a similar statement for the left lifted action. The affine Poisson structure on $\mathfrak{g}^*$ has been already studied in detail in Theorem 6.2.2; our goal here is to recover those results by reduction.

7.2.1 Theorem. *The Poisson reduced space for the right cotangent lifted action of G on $(T^*G, \Omega - \pi_G^*\mathcal{B})$ is $\mathfrak{g}^*$ with Poisson bracket given by*

$$\{f, g\}_{\mathcal{B}}(\mu) = \left\langle \mu, \left[\frac{\delta f}{\delta \mu}, \frac{\delta g}{\delta \mu}\right] \right\rangle + \mathcal{B}(e)\left(\frac{\delta f}{\delta \mu}, \frac{\delta g}{\delta \mu}\right) \qquad (7.2.1)$$

for $f, g \in C^\infty(\mathfrak{g}^*)$.

Proof.   We proceed by computing the Poisson bracket on T^*G for momentum functions generated from right invariant vector fields. Given a vector field X on G, define, as is standard, the **momentum function** $P_X \in C^\infty(T^*G)$, linear on the fibers, by

$$P_X(\alpha_g) = \langle \alpha_g, X(g) \rangle. \qquad (7.2.2)$$

Notice that if X is right invariant on G, then P_X is right invariant on T^*G. The canonical bracket of two momentum functions P_X and P_Y is given by $-P_{[X,Y]}$ as an easy calculation (done in [MandS], §12.1) shows. Using equation (7.1.1), we see that

$$\{P_X, P_Y\}_{\mathcal{B}} = -P_{[X,Y]} + \mathcal{B}(X, Y) \circ \pi_G. \qquad (7.2.3)$$

To prove (7.2.1), it suffices to compute the bracket of two linear functions on $\mathfrak{g}^*$ since the Poisson bracket at any point only depends on the first derivatives. Thus, let $f : \mathfrak{g} \to \mathbb{R}$ be linear, so that $\delta f / \delta \mu$ is constant and $f(\mu) = \langle \mu, \delta f / \delta \mu \rangle$ and similarly for g. For what follows, let

$$X = \left(\frac{\delta f}{\delta \mu}\right)_R \quad \text{and} \quad Y = \left(\frac{\delta g}{\delta \mu}\right)_R$$

denote the right invariant vector fields on G whose values at the identity are $\delta f / \delta \mu$, $\delta g / \delta \mu$ respectively. Recall that the left momentum map for the lifted action using the canonical symplectic form on T^*G is the map $\mathbf{J}_L : T^*G \to \mathfrak{g}^*$ given by $\mathbf{J}_L(\alpha_g) = T_e^* R_g(\alpha_g) =: \alpha_g \cdot g^{-1}$. Then, since

$$(f \circ \mathbf{J}_L)(\alpha_g) = \left\langle \alpha_g \cdot g^{-1}, \frac{\delta f}{\delta \mu} \right\rangle = \left\langle \alpha_g, \frac{\delta f}{\delta \mu} \cdot g \right\rangle$$
$$= \langle \alpha_g, X(g) \rangle = P_X(\alpha_g),$$

the right invariant extension of f is given by $P(X)$. Thus, by definition of the reduced Poisson bracket along with equation (7.2.3), we get

$$\begin{aligned}
\{f, g\}(\mu) &= \{P_X, P_Y\}_{\mathcal{B}}(\mu) \\
&= -P_{[X,Y]}(\mu) + \mathcal{B}(X, Y)(e) \\
&= -\langle \mu, [X, Y](e) \rangle + \mathcal{B}(e)(X(e), Y(e)) \\
&= \left\langle \mu, \left[\frac{\delta f}{\delta \mu}, \frac{\delta g}{\delta \mu}\right] \right\rangle + \mathcal{B}(e)\left(\frac{\delta f}{\delta \mu}, \frac{\delta g}{\delta \mu}\right),
\end{aligned}$$

as required. In the last equality we have used the identity $[\xi_R, \eta_R](e) = -[\xi, \eta]$, valid for the *right* invariant extensions $\xi_R, \eta_R \in \mathfrak{X}(G)$ of $\xi \in \mathfrak{g}$ and $\eta \in \mathfrak{g}$ respectively. ∎

Poisson brackets of this sort arise in a number of contexts, such as those studied in Cendra, Marsden, and Ratiu [2003], Theorem 3.3.1. The present result shows that in fact, the result just quoted may be obtained by magnetic Lie-Poisson reduction.

Affine Orbits as Magnetic Reduced Spaces. In Theorem 1.2.3 it was shown that the reduction of T^*G at $\mu \in \mathfrak{g}^*$ is symplectically diffeomorphic to the coadjoint orbit $\mathcal{O}_\mu \subset \mathfrak{g}^*$ passing through μ. In this paragraph we shall prove an analogous result for magnetic cotangent bundles of Lie groups, namely, the reduction of a magnetic cotangent bundle at a given point $\mu \in \mathfrak{g}^*$ is symplectically diffeomorphic with the affine orbit $\tilde{\mathcal{O}}_\mu$ passing through μ.

We make the following assumptions. Let $\mathcal{B}$ be a closed two-form on a connected Lie group G. Let G act on itself by right translations. Assume that $\mathcal{B}$ is G-invariant. Consider the cotangent lifted action of G to T^*G. This gives a symplectic action of G on the symplectic manifold $(T^*G, \Omega - \pi_G^*\mathcal{B})$, where $\pi_G : T^*G \to G$ is the cotangent bundle projection. Suppose that there is a smooth map $\phi : G \to \mathfrak{g}^*$ that satisfies

$$\mathbf{i}_{\xi_L}\mathcal{B} = \mathbf{d}\langle\phi, \xi\rangle$$

for all $\xi \in \mathfrak{g}$. Recall that the infinitesimal generator of ξ for the right action is the left invariant extension of ξ; that is, $\xi_G = \xi_L$. Recall from Theorem 7.1.1 that

(i) The map $\mathbf{J} = \mathbf{J}_{\mathrm{can}} - \phi \circ \pi_G$, where $\mathbf{J}_{\mathrm{can}} = \mathbf{J}_R$ is the standard right momentum map for the G-action relative to the canonical symplectic form, is a momentum map for the cotangent lifted action of G on T^*G with symplectic form $\Omega - \pi_G^*\mathcal{B}$.

(ii) The momentum map $\mathbf{J}$ is, in general, not equivariant. Its nonequivariance $\mathfrak{g}^*$-valued group one-cocycle $\sigma^{T^*G} : G \to \mathfrak{g}^*$ is given by

$$\sigma^{T^*G}(g) = -\phi(hg) + \mathrm{Ad}_g^*(\phi(h)), \tag{7.2.4}$$

with the right hand side independent of $h \in G$. The cocycle identity is in this case

$$\sigma^{T^*G}(gh) = \mathrm{Ad}_h^* \sigma^{T^*G}(g) + \sigma^{T^*G}(h). \tag{7.2.5}$$

In this equation, taking $g = h = e$ we conclude that $\sigma^{T^*G}(e) = 0$. Therefore, in the same equation, letting $h = g^{-1}$, we get

$$\sigma^{T^*G}(g^{-1}) = -\mathrm{Ad}_{g^{-1}}^* \sigma^{T^*G}(g) \tag{7.2.6}$$

for any $g \in G$.

(iii) The momentum map $\mathbf{J} : T^*G \to \mathfrak{g}^*$ is equivariant relative to the right affine action of G on $\mathfrak{g}^*$ given by

$$\lambda \cdot g := \mathrm{Ad}_g^* \lambda + \sigma^{T^*G}(g) \tag{7.2.7}$$

for $g \in G$ and $\lambda \in \mathfrak{g}^*$. The induced right Lie algebra action is given by

$$\lambda \cdot \xi := \left. \frac{d}{dt} \right|_{t=0} \lambda \cdot \exp(t\xi)$$
$$= \mathrm{ad}_\xi^* \lambda + T_e \sigma^{T^*G}(\xi) = \mathrm{ad}_\xi^* \lambda - \Sigma^{T^*G}(\xi, \cdot) \tag{7.2.8}$$

for $\lambda \in \mathfrak{g}^*$ and $\xi \in \mathfrak{g}$.

7.2.2 Theorem (Reduction to Affine Orbits).   *Let $\mu \in \mathfrak{g}^*$. Then μ is a regular value of $\mathbf{J}$, the (right) action of G on T^*G is free and proper, and the symplectic reduced space $\mathbf{J}^{-1}(\mu)/\widetilde{G}_\mu$ is symplectically diffeomorphic to*

$$\widetilde{\mathcal{O}}_\mu = \left\{ \mu \cdot g = \mathrm{Ad}_g^* \mu + \sigma^{T^*G}(g) \,\middle|\, g \in G \right\},$$

the right affine orbit through μ. The tangent space at $\lambda = \mu \cdot g \in \widetilde{\mathcal{O}}_\mu$ to $\widetilde{\mathcal{O}}_\mu$ is given by

$$T_\lambda \widetilde{\mathcal{O}}_\mu = \left\{ \mathrm{ad}_\xi^* \lambda - \Sigma^{T^*G}(\xi, \cdot) \,\middle|\, \xi \in \mathfrak{g} \right\},$$

where $\Sigma^{T^*G}(\xi, \cdot) = -T_e \sigma^{T^*G}(\xi)$. The symplectic structure on $\widetilde{\mathcal{O}}_\mu$ has the expression

$$\omega_{\mathcal{B}}^+(\lambda) \left(\mathrm{ad}_\xi^* \lambda - \Sigma^{T^*G}(\xi, \cdot), \mathrm{ad}_\eta^* \lambda - \Sigma^{T^*G}(\eta, \cdot) \right)$$
$$= \langle \lambda, [\xi, \eta] \rangle - \Sigma^{T^*G}(\xi, \eta), \tag{7.2.9}$$

*which we call the **magnetic orbit symplectic form**.*

    Equation (7.1.16) *together with Remark 4 of Theorem* 7.1.6 *shows that the affine orbit is symplectomorphic to $G/\widetilde{G}_\mu$ with symplectic form $-\mathcal{B}_\mu$ defined by*

$$\pi_{G,\widetilde{G}_\mu}^* \mathcal{B}_\mu = \mathcal{B} + \mathbf{d}(\mathcal{A}_\phi + \mathcal{A}(\cdot)^* \mu).$$

Remark.   The symplectic form (7.2.9), obtained here by reduction, is a special case of (6.2.21) since equations (1.1.10) and (6.2.10) show that in this case of one dimensional extensions, we have $C = -\Sigma^{T^*G}$.

Proof.   From the fact that $\mathbf{J}_R$ is given by *left* translation to the identity, we have

$$\mathbf{J}(\alpha_g) = T_e^* L_g \alpha_g - \phi(g). \tag{7.2.10}$$

Thus, $\mathbf{J}^{-1}(\mu)$ consists of those $\alpha_g \in T^*G$ such that $\alpha_g = T_g^* L_{g^{-1}}(\mu + \phi(g))$.

It is clear that the G action is free and proper on G and hence on T^*G. Thus, from Proposition 1.1.2, each μ is a regular value.

We claim that the smooth map $\varphi : \mathbf{J}^{-1}(\mu) \to \widetilde{\mathcal{O}}_\mu$ defined for $\alpha_g \in \mathbf{J}^{-1}(\mu)$ by

$$\varphi(\alpha_g) := \mu \cdot g^{-1} = \mathrm{Ad}^*_{g^{-1}} \mu + \sigma^{T^*G}(g^{-1})$$

is $\widetilde{G}_\mu$-invariant. We begin with the following assertion: if $\alpha_g \in \mathbf{J}^{-1}(\mu)$ and $\alpha_g \cdot h = T^*_{gh} R_{h^{-1}} \alpha_g$ for $h \in \widetilde{G}_\mu$, then $\alpha_g \cdot h \in \mathbf{J}^{-1}(\mu)$. To see this, note that

$$\begin{aligned}
\mathbf{J}(\alpha_g \cdot h) &= T^*_e L_{gh} \left(\alpha_g \cdot h \right) - \phi(gh) \\
&= T^*_e L_{gh} \left(T^*_{gh} R_{h^{-1}} \alpha_g \right) - \phi(gh) \\
&= T^*_e L_{gh} \left(T^*_{gh} R_{h^{-1}} T^*_g L_{g^{-1}} (\mu + \phi(g)) \right) - \phi(gh) \\
&= T^*_e L_{gh} \left(T^*_{gh} R_{h^{-1}} T^*_g L_{g^{-1}} (\mu + \phi(g)) \right) - \phi(gh) \\
&= \mathrm{Ad}^*_h (\mu + \phi(g)) - \phi(gh) \\
&= \mathrm{Ad}^*_h \mu + \sigma^{T^*G}(h)
\end{aligned}$$

where, in the last equality, we used (7.2.4). But, since $h \in \widetilde{G}_\mu$, by (7.2.7), the right hand side equals μ, and so $\mathbf{J}(\alpha_g \cdot h) = \mu$ as we asserted.

To prove that φ is is $\widetilde{G}_\mu$-invariant, note that

$$\begin{aligned}
\varphi(\alpha_g \cdot h) &= \mathrm{Ad}^*_{(gh)^{-1}} \mu + \sigma^{T^*G}((gh)^{-1}) \\
&= \mathrm{Ad}^*_{g^{-1}} \mathrm{Ad}^*_{h^{-1}} \mu + \mathrm{Ad}^*_{g^{-1}} \sigma^{T^*G}(h^{-1}) + \sigma^{T^*G}(g^{-1}) \\
&= \mathrm{Ad}^*_{g^{-1}} \left(\mathrm{Ad}^*_{h^{-1}} \mu + \sigma^{T^*G}(h^{-1}) \right) + \sigma^{T^*G}(g^{-1}) \\
&= \mathrm{Ad}^*_{g^{-1}} \mu + \sigma^{T^*G}(g^{-1}) \\
&= \varphi(\alpha_g),
\end{aligned}$$

where we used the cocycle identity (7.2.5).

Note also that $\varphi : \mathbf{J}^{-1}(\mu) \to \widetilde{\mathcal{O}}_\mu$ is surjective for if $\lambda = \mu \cdot g \in \widetilde{\mathcal{O}}_\mu$, then $\alpha_g := T^*_g L_{g^{-1}}(\mu + \phi(g)) \in \mathbf{J}^{-1}(\mu)$ and $\varphi(\alpha_g) = \lambda$.

Thus, φ induces a smooth surjective map $\overline{\varphi} : \mathbf{J}^{-1}(\mu)/\widetilde{G}_\mu \to \widetilde{\mathcal{O}}_\mu$ characterized by the relation $\varphi = \overline{\varphi} \circ \pi_\mu$, where $\pi_\mu : \mathbf{J}^{-1}(\mu) \to \mathbf{J}^{-1}(\mu)/\widetilde{G}_\mu$ is the projection. Let us show that $\overline{\varphi}$ is also injective. Indeed, if

$$\mu \cdot g^{-1} = \varphi(\alpha_g) = \overline{\varphi}(\pi_\mu(\alpha_g)) = \overline{\varphi}(\pi_\mu(\alpha_h)) = \varphi(\alpha_h) = \mu \cdot h^{-1}$$

for some $\alpha_g = T_g^* L_{g^{-1}}(\mu + \phi(g))$ and some $\alpha_h = T_h^* L_{h^{-1}}(\mu + \phi(h)) \in \mathbf{J}^{-1}(\mu)$, then $h^{-1}g \in \widetilde{G}_\mu$, we get

$$
\begin{aligned}
T_g^* R_{g^{-1}h}\alpha_h &= T_g^* R_{g^{-1}h}T_h^* L_{h^{-1}}(\mu + \phi(h)) = T_g^* R_{g^{-1}}\operatorname{Ad}_{h^{-1}}^*(\mu + \phi(h)) \\
&= T_g^* R_{g^{-1}}\operatorname{Ad}_{g^{-1}}^*\operatorname{Ad}_g^*\operatorname{Ad}_{h^{-1}}^*(\mu + \phi(h)) \\
&= T_g^* L_{g^{-1}}\operatorname{Ad}_{h^{-1}g}^*(\mu + \phi(h)) \\
&= T_g^* L_{g^{-1}}\left(\operatorname{Ad}_{h^{-1}g}^*\mu + \sigma^{T^*G}(h^{-1}g)\right) \\
&\quad + T_g^* L_{g^{-1}}\left(\operatorname{Ad}_{h^{-1}g}^*\phi(h) - \sigma^{T^*G}(h^{-1}g)\right).
\end{aligned}
$$

Because $h^{-1}g \in \widetilde{G}_\mu$, the first term in the last equality is $T_g^* L_{g^{-1}}\mu$. Using (7.2.4) (with g replaced by $h^{-1}g$), the second term in the last equality equals $T_g^* L_{g^{-1}}\phi(g)$. Hence,

$$
\begin{aligned}
T_g^* R_{g^{-1}h}\alpha_h &= T_g^* L_{g^{-1}}\mu + T_g^* L_{g^{-1}}\phi(g) \\
&= T_g^* L_{g^{-1}}(\mu + \phi(g)) \\
&= \alpha_g.
\end{aligned}
$$

Therefore, $\pi_\mu(\alpha_g) = \pi_\mu(T_g^* R_{g^{-1}h}\alpha_h) = \pi_\mu(\alpha_h)$, which proves that $\overline{\varphi} : \mathbf{J}^{-1}(\mu)/\widetilde{G}_\mu \to \widetilde{\mathcal{O}}_\mu$ is injective. Thus $\overline{\varphi}$ is a smooth bijective map. To prove that $\overline{\varphi}$ is a diffeomorphism, we need only show that it is a local diffeomorphism; that is, its derivative is bijective.

To accomplish this, we first compute the derivative of φ. Let

$$
\alpha_\xi(t) := T_{\exp(t\xi)g}^* L_{g^{-1}\exp(-t\xi)}\left(\mu + \phi(\exp(t\xi)g)\right) \in \mathbf{J}^{-1}(\mu) \qquad (7.2.11)
$$

be an arbitrary smooth curve in $\mathbf{J}^{-1}(\mu)$ passing through the generic point $\alpha_g := \alpha_\xi(0) = T_g^* L_{g^{-1}}(\mu + \phi(g)) \in \mathbf{J}^{-1}(\mu)$, where $\xi \in \mathfrak{g}$. Then $\varphi(\alpha_\xi(0)) = \mu \cdot g^{-1}$ and we get

$$
\begin{aligned}
T_{\alpha_\xi(0)}\varphi\left(\alpha'_\xi(0)\right) &= \frac{d}{dt}\bigg|_{t=0} \mu \cdot \left(g^{-1}\exp(-t\xi)\right) \\
&= \frac{d}{dt}\bigg|_{t=0}\left((\mu \cdot g^{-1}) \cdot \exp(-t\xi)\right) = -\left(\mu \cdot g^{-1}\right) \cdot \xi.
\end{aligned}
$$
$$(7.2.12)$$

It follows that the derivative of φ, namely $T_{\alpha_g}\varphi$ is surjective at every point α_g since every tangent vector to $\widetilde{\mathcal{O}}_\mu$ at the image point $\varphi(\alpha_g) = \mu \cdot g^{-1}$ is of the form $(\mu \cdot g^{-1}) \cdot \xi$ for some $\xi \in \mathfrak{g}$.

It follows that the derivative of the induced quotient map $\overline{\varphi}$ at every point is also surjective.

Next, note that the two manifolds $\mathbf{J}^{-1}(\mu)/\widetilde{G}_\mu$ and $\widetilde{\mathcal{O}}_\mu$ have the same dimension, namely $\dim G - \dim \widetilde{G}_\mu$. Indeed, observe that the dimension

of $\mathbf{J}^{-1}(\mu)$ is the same as $\dim G$ because the map $G \to \mathbf{J}^{-1}(\mu)$ given by $g \mapsto T_g^* L_{g^{-1}} (\mu + \phi(g))$ is obviously a diffeomorphism. The dimension of the quotient spaces are obtained by subtraction since the actions are free and proper.

Since the derivative $T_{\pi_\mu(\alpha_g)}\overline{\varphi}$ is a surjective linear map between spaces of the same dimension, it is an isomorphism. Thus, the map $\overline{\varphi}$ is a diffeomorphism. It remains to show that it is symplectic.

The reduced symplectic form ω_μ on $\mathbf{J}^{-1}(\mu)/\widetilde{G}_\mu$ is induced by the symplectic form $\Omega - \pi_G^* \mathcal{B}$ on $T^* G$ pulled-back to $\mathbf{J}^{-1}(\mu)$; that is,

$$\pi_\mu^* \omega_\mu = i_\mu^* \left(\Omega - \pi_G^* \mathcal{B} \right).$$

We wish to prove that $\overline{\varphi}^* \omega_\mathcal{B}^+ = \omega_\mu$. To show this, it is sufficient to show that π_μ^* applied to each side is equal; that is,

$$\varphi^* \omega_\mathcal{B}^+ = i_\mu^* \left(\Omega - \pi_G^* \mathcal{B} \right) \tag{7.2.13}$$

Using the notation in (7.2.11) and fixing $g \in G$, let $\alpha_\xi'(0), \alpha_\eta'(0)$ be two tangent vectors to $\mathbf{J}^{-1}(\mu)$ at the point $\alpha_g = T_g^* L_{g^{-1}}(\mu + \phi(g))$. In (7.2.12), we computed that

$$T_{\alpha_g}\varphi(\alpha_\xi'(0)) = -(\mu \cdot g^{-1}) \cdot \xi$$

and similarly for η. Thus, what we are trying to show, namely (7.2.13) is equivalent to

$$\begin{aligned} \omega_\mathcal{B}^+ (\mu \cdot g^{-1}) &\left((\mu \cdot g^{-1}) \cdot \xi, (\mu \cdot g^{-1}) \cdot \eta \right) \\ &= (\Omega - \pi_G^* \mathcal{B}) (\alpha_g) \left(\alpha_\xi'(0), \alpha_\eta'(0) \right) \end{aligned} \tag{7.2.14}$$

We first work on the left hand side of this equation. Recalling from (7.2.8) that tangent vectors to $\widetilde{\mathcal{O}}_\mu$ at the point $\mu \cdot g^{-1}$ have the form

$$(\mu \cdot g^{-1}) \cdot \xi = \operatorname{ad}_\xi^* (\mu \cdot g^{-1}) - \Sigma^{T^*G}(\xi, \cdot),$$

and using (7.2.9), the left hand side of (7.2.14) becomes

$$\left\langle \mu \cdot g^{-1}, [\xi, \eta] \right\rangle + \Sigma^{T^*G}(\xi, \eta). \tag{7.2.15}$$

Next, we work on the right hand side of (7.2.14). We will proceed in a number of steps. The first step will be to find a more convenient expression for $\alpha_\xi(t)$ given in (7.2.11). We do this in order to exploit the right invariance properties of Ω and $\mathcal{B}$.

7.2.3 Lemma.   *The curve $\alpha_\xi(t)$ may be rewritten as*

$$\alpha_\xi(t) = T_{\exp(t\xi)g}^* R_{g^{-1}} T_{\exp(t\xi)}^* L_{\exp(-t\xi)} \left(\mu \cdot g^{-1} + \phi(\exp(t\xi)) \right). \tag{7.2.16}$$

Proof.   We manipulate the right hand side of (7.2.16) by making use of (7.2.7) and (7.2.6) as follows

$$
T^*_{\exp(t\xi)g} R_{g^{-1}} T^*_{\exp(t\xi)} L_{\exp(-t\xi)} \left(\mu \cdot g^{-1} + \phi(\exp(t\xi)) \right)
$$
$$
= T^*_{\exp(t\xi)g} R_{g^{-1}} T^*_{\exp(t\xi)} L_{\exp(-t\xi)} \left(\mathrm{Ad}^*_{g^{-1}} \mu + \sigma^{T^*G}(g^{-1}) + \phi(\exp(t\xi)) \right)
$$
$$
= T^*_{\exp(t\xi)g} R_{g^{-1}} T^*_{\exp(t\xi)} L_{\exp(-t\xi)} \left(\mathrm{Ad}^*_{g^{-1}} \mu - \mathrm{Ad}^*_{g^{-1}} \sigma^{T^*G}(g) + \phi(\exp(t\xi)) \right)
$$

However, by (7.2.4) with $h = \exp(t\xi)$, we have

$$
\phi(\exp(t\xi)) = \mathrm{Ad}^*_{g^{-1}} \left(\sigma^{T^*G}(g) + \phi(\exp(t\xi)g) \right)
$$

and so the right hand side of (7.2.16) becomes

$$
T^*_{\exp(t\xi)g} R_{g^{-1}} T^*_{\exp(t\xi)} L_{\exp(-t\xi)} \, \mathrm{Ad}^*_{g^{-1}} \left(\mu + \phi(\exp(t\xi)g) \right)
$$
$$
= T^*_{\exp(t\xi)g} R_{g^{-1}} T^*_{\exp(t\xi)} L_{\exp(-t\xi)} T^*_e L_{g^{-1}} T^*_{g^{-1}} R_g \left(\mu + \phi(\exp(t\xi)g) \right)
$$
$$
= T^*_{\exp(t\xi)g} L_{g^{-1} \exp(-t\xi)} \left(\mu + \phi(\exp(t\xi)g) \right)
$$

since left and right translations commute. This agrees with (7.2.11) and so the lemma is proved.   ▼

Letting Φ^R_g denote the cotangent lifted action of right translation by g and similarly Φ^L_g that for left translation by g, we can rewrite $\alpha_\xi(t)$ as

$$
\alpha_\xi(t) = \Phi^R_g \Phi^L_{\exp(t\xi)} \left(\mu \cdot g^{-1} + \phi(\exp(t\xi)) \right).
$$

Differentiating this expression with respect to t at $t = 0$, we get

$$
\alpha'_\xi(0) = T\Phi^R_g \left(\xi^L_{T^*G}(\mu \cdot g^{-1} + \phi(e)) + \mathrm{Ver}_{\mu \cdot g^{-1} + \phi(e)} \, T_e \phi(\xi) \right), \qquad (7.2.17)
$$

where $\xi^L_{T^*G}$ denotes the infinitesimal generator of the action Φ^L and where

$$
\mathrm{Ver}_{\alpha_q} \beta_q := \left. \frac{d}{d\epsilon} \right|_{\epsilon=0} (\alpha_q + \epsilon \beta_q) \in T_{\alpha_q} T^*Q \qquad (7.2.18)
$$

denotes the vertical lift of a covector $\beta_q \in T^*_q Q$ on a manifold Q along the covector $\alpha_q \in T^*_q Q$.

Using equation (7.2.17), we can now evaluate the right hand side of (7.2.14). We also make use of the fact that the right action leaves both Ω and $\mathcal{B}$ invariant, and the fact that the right action converts $\alpha_g = T^*_g L_{g^{-1}}(\mu + \phi(g))$ into $\mathrm{Ad}^*_{g^{-1}}(\mu + \phi(g))$, to get

$$
(\Omega - \pi^*_G \mathcal{B})(\mathrm{Ad}^*_{g^{-1}}(\mu + \phi(g)))
$$
$$
\left(\xi^L_{T^*G}(\mu \cdot g^{-1} + \phi(e)) + \mathrm{Ver}_{\mu \cdot g^{-1} + \phi(e)} \, T_e \phi(\xi), \right.
$$
$$
\left. \eta^L_{T^*G}(\mu \cdot g^{-1} + \phi(e)) + \mathrm{Ver}_{\mu \cdot g^{-1} + \phi(e)} \, T_e \phi(\eta) \right). \qquad (7.2.19)
$$

We check now that the base points in this expression, namely the two points $\mathrm{Ad}^*_{g^{-1}}(\mu + \phi(g))$ and $\mu \cdot g^{-1} + \phi(e)$ are in fact equal, as they must be. To see this, we use (7.2.7), (7.2.6), and (7.2.4) with $h = e$ as follows:

$$
\begin{aligned}
\mu \cdot g^{-1} + \phi(e) &= \mathrm{Ad}^*_{g^{-1}} \mu - \mathrm{Ad}^*_{g^{-1}} \sigma^{T^*G}(g) + \phi(e) \\
&= \mathrm{Ad}^*_{g^{-1}} \mu + \mathrm{Ad}^*_{g^{-1}} \phi(g) - \phi(e) + \phi(e) \\
&= \mathrm{Ad}^*_{g^{-1}}(\mu + \phi(g)).
\end{aligned}
$$

We now look at the various terms in (7.2.19). First of all, we consider the terms involving Ω and there are four of these. The first one is

$$
\begin{aligned}
\Omega\left(\mathrm{Ad}^*_{g^{-1}}(\mu + \phi(g))\right) &\left(\xi^L_{T^*G}(\mu \cdot g^{-1} + \phi(e)), \eta^L_{T^*G}(\mu \cdot g^{-1} + \phi(e))\right) \\
&= \left\langle \mu \cdot g^{-1} + \phi(e), [\xi, \eta] \right\rangle
\end{aligned}
\tag{7.2.20}
$$

where we have used the reduction to coadjoint orbits theorem, namely the right invariant version of equation (1.2.6) given in Theorem 1.2.3.

For the cross terms in (7.2.19) corresponding to Ω, we make use of the following.

7.2.4 Lemma. *For a free and proper action of a Lie group G on a manifold Q, cotangent lifted to T^*Q and with Ω the canonical symplectic form, we have*

$$
\Omega(\alpha_q)\left(\xi_{T^*Q}(\alpha_q), \mathrm{Ver}_{\alpha_q}(\beta_q)\right) = \langle \mathbf{J}_{\mathrm{can}}(\beta_q), \xi \rangle.
\tag{7.2.21}
$$

Proof. First of all, using the free and proper assumption, extend the one form β_q at q to a G-invariant one-form β in a neighborhood of q. Define vector fields X and Y on T^*Q in a neighborhood of α_q as follows. First of all, let $X = \xi_{T^*Q}$ and second, let $Y(\gamma_{\bar{q}}) = \mathrm{Ver}_{\gamma_{\bar{q}}} \beta(\bar{q})$.

We now evaluate each of the terms in the identity

$$
\Omega(X, Y) = -X[\Theta(Y)] + Y[\Theta(X)] + \Theta([X, Y]).
$$

We claim that $X[\Theta(Y)] = 0$. Indeed,

$$
\begin{aligned}
\Theta(Y)(\gamma_{\bar{q}}) &= \langle \gamma_{\bar{q}}, T\pi_Q(Y(\gamma_{\bar{q}})) \rangle \\
&= \langle \gamma_{\bar{q}}, T\pi_Q \, \mathrm{Ver}_{\gamma_{\bar{q}}} \beta(\bar{q}) \rangle.
\end{aligned}
$$

However, vertical vectors satisfy $T\pi_Q \, \mathrm{Ver}_{\gamma_{\bar{q}}} \beta(\bar{q}) = 0$. Thus, $X[\Theta(Y)] = 0$.

Second, we claim that $Y[\Theta(X)](\gamma_{\bar{q}}) = \langle \mathbf{J}_{\mathrm{can}}(\beta(\bar{q})), \xi \rangle$. Indeed, first notice that

$$
\begin{aligned}
\Theta(X)(\gamma_{\bar{q}}) &= \langle \gamma_{\bar{q}}, T\pi_Q\left(\xi_{T^*Q}(\gamma_{\bar{q}})\right) \rangle \\
&= \langle \gamma_{\bar{q}}, \xi_Q(\bar{q}) \rangle = \langle \mathbf{J}_{\mathrm{can}}(\gamma_{\bar{q}}), \xi \rangle.
\end{aligned}
$$

Therefore, using the preceding equation and the chain rule, we get

$$
\begin{aligned}
Y[\Theta(X)](\gamma_{\bar{q}}) &= \langle \mathbf{d} \langle \mathbf{J}_{\mathrm{can}}, \xi \rangle (\gamma_{\bar{q}}), Y(\gamma_{\bar{q}}) \rangle \\
&= \left\langle \mathbf{d} \langle \mathbf{J}_{\mathrm{can}}, \xi \rangle (\gamma_{\bar{q}}), \left. \frac{d}{d\epsilon} \right|_{\epsilon=0} (\gamma_{\bar{q}} + \epsilon\beta(\bar{q})) \right\rangle \\
&= \left. \frac{d}{d\epsilon} \right|_{\epsilon=0} \langle \mathbf{J}_{\mathrm{can}}(\gamma_{\bar{q}} + \epsilon\beta(\bar{q})), \xi \rangle \\
&= \left. \frac{d}{d\epsilon} \right|_{\epsilon=0} \langle \gamma_{\bar{q}} + \epsilon\beta(\bar{q}), \xi_Q(\bar{q}) \rangle \\
&= \langle \mathbf{J}_{\mathrm{can}}(\beta(\bar{q})), \xi \rangle .
\end{aligned}
$$

Thus, we have shown our claim, namely $Y[\Theta(X)](\gamma_{\bar{q}}) = \langle \mathbf{J}_{\mathrm{can}}(\beta(\bar{q})), \xi \rangle$.

To prove the lemma, it now suffices to show that $\Theta([X, Y]) = 0$. In fact, we assert that the bracket $[X, Y] = 0$. To see this, recall that $X = \xi_{T^*G}$ and, since we choose β to be G-invariant, it follows that Y, which is given in terms of fiber translation by β, is also G-invariant. Thus, $[X, Y] = 0$.

▼

We use the lemma to evaluate the cross terms in equation (7.2.19) corresponding to Ω as follows:

$$
\Omega(\mathrm{Ad}^*_{g^{-1}}(\mu + \phi(g))) \left(\xi^L_{T^*G}(\mu \cdot g^{-1} + \phi(e)), \mathrm{Ver}_{\mu \cdot g^{-1} + \phi(e)} T_e\phi(\eta) \right)
$$
$$
= \langle \mathbf{J}_L(T_e\phi(\eta)), \xi \rangle = \langle T_e\phi(\eta), \xi \rangle \tag{7.2.22}
$$

There is a similar cross term with ξ and η interchanged with a minus sign. Thus, for the cross terms, we get

$$
\langle T_e\phi(\eta), \xi \rangle - \langle T_e\phi(\xi), \eta \rangle . \tag{7.2.23}
$$

However, from the relation $\mathbf{i}_{\xi_G}\mathcal{B} = \mathbf{d} \langle \phi, \xi \rangle$ we get $\langle T_e\phi(\eta), \xi \rangle = \mathcal{B}(e)(\xi, \eta)$. Thus, the cross terms are

$$
\langle T_e\phi(\eta), \xi \rangle - \langle T_e\phi(\xi), \eta \rangle = 2\mathcal{B}(e)(\xi, \eta). \tag{7.2.24}
$$

The last of the Ω terms is

$$
\Omega(\mathrm{Ad}^*_{g^{-1}}(\mu + \phi(g))) \left(\mathrm{Ver}_{\mu \cdot g^{-1} + \phi(e)} T_e\phi(\xi), \mathrm{Ver}_{\mu \cdot g^{-1} + \phi(e)} T_e\phi(\eta) \right) . \tag{7.2.25}
$$

This is zero, as is seen by a similar argument that was used in Lemma 7.2.4; namely we write the preceding expression as a sum of three terms and since Θ vanishes on vertical vectors, each of the three terms is zero. Notice that the bracket of two vertical vector fields is vertical as each is π_G-related to zero.

In summary, the sum of all the Ω terms in equation (7.2.19) are:

$$
\langle \mu \cdot g^{-1} + \phi(e), [\xi, \eta] \rangle + 2\mathcal{B}(e)(\xi, \eta) \tag{7.2.26}
$$

Now we turn to the $\mathcal{B}$ terms in equation (7.2.19). Because of the presence of π_G^*, all the vertical terms will yield zero and so the only nonzero $\mathcal{B}$ term is

$$- \mathcal{B}(e)\left(\xi_G^L(e), \eta_G^L(e)\right) = -\mathcal{B}(e)(\xi, \eta) \tag{7.2.27}$$

where we have used the fact that $\xi_{T^*G}^L$ is π_G-related to ξ_G^L and so we have the identity $T\pi_G(\xi_{T^*G}^L(\mu \cdot g^{-1} + \phi(e)) = \xi_G^L(e)$.

Adding equations (7.2.26) and (7.2.27), we see that the expression given in equation (7.2.19) is equal to

$$\left\langle \mu \cdot g^{-1} + \phi(e), [\xi, \eta] \right\rangle + \mathcal{B}(e)(\xi, \eta)$$
$$= \left\langle \mu \cdot g^{-1}, [\xi, \eta] \right\rangle + \left\langle \phi(e), [\xi, \eta] \right\rangle + \mathcal{B}(e)(\xi, \eta). \tag{7.2.28}$$

To prove the first part of the theorem (equation (7.2.9)), it suffices to show that

$$\langle \phi(e), [\xi, \eta] \rangle + \mathcal{B}(e)(\xi, \eta) = -\Sigma^{T^*G}(\xi, \eta). \tag{7.2.29}$$

Recall that $-\Sigma^{T^*G}(\xi, \eta) = \left\langle T_e\sigma^{T^*G}(\xi), \eta \right\rangle$. Also, recall equation (7.2.4), namely $\sigma^{T^*G}(g) = -\phi(hg) + \mathrm{Ad}_g^*(\phi(h))$. Setting $h = e$ and taking the derivative with respect to g at $g = e$ in the direction ξ, we get

$$T_e\sigma^{T^*G}(\xi) = -T_e\phi(\xi) + \mathrm{ad}_\xi^* \phi(e) \tag{7.2.30}$$

Pairing both sides with η, we get

$$-\Sigma^{T^*G}(\xi, \eta) = \left\langle T_e\sigma^{T^*G}(\xi), \eta \right\rangle$$
$$= \left\langle -T_e\phi(\xi) + \mathrm{ad}_\xi^* \phi(e), \eta \right\rangle$$
$$= \mathcal{B}(e)(\xi, \eta) + \langle \phi(e), [\xi, \eta] \rangle. \tag{7.2.31}$$

This establishes equation (7.2.29).

To prove the last statement of the theorem, we first observe that the map $\phi : \mathbf{J}^{-1}(\mu) \to \widetilde{\mathcal{O}}_\mu$ in fact projects to a map $\psi : G \to \widetilde{\mathcal{O}}_\mu$ given by $\psi(g) := \mathrm{Ad}_{g^{-1}}^* \mu + \sigma^{T^*G}(g^{-1})$. The map ψ is invariant with respect to the right action of $\widetilde{G}_\mu$ on G and therefore defines a smooth, surjective map $\bar{\psi} : G/\widetilde{G}_\mu \to \widetilde{\mathcal{O}}_\mu$. One proves that $\bar{\psi}$ is a diffeomorphism in exactly the same manner that $\bar{\phi}$ was shown to be a diffeomorphism. We will show that $\bar{\psi}^*\omega_{\mathcal{B}}^+ = -\mathcal{B}_\mu$ for which it suffices to prove that

$$\psi^*\omega_{\mathcal{B}}^+ = -\mathcal{B} - \mathbf{d}(\mathcal{A}_\phi + \mathcal{A}(\cdot)^*\mu). \tag{7.2.32}$$

We begin by computing the right hand side. Since the action of G on G is by right translation, the infinitesimal generators are given by left translation and therefore the connection form on the bundle $G \to \{e\}$ is given by the left Maurer-Cartan one-form on G, which we denote by $\mathcal{A} \in \Omega^1(G; \mathfrak{g})$ and is defined by $\mathcal{A}(g)(v_g) := T_gL_{g^{-1}}v_g$ where $v_g \in T_gG$. We then have $\mathcal{A}_\phi(g)(v_g) = \left\langle \phi(g), T_gL_{g^{-1}}v_g \right\rangle$, where ϕ is the $\mathcal{B}\mathfrak{g}$-potential.

For $v_g, w_g \in T_g G$, define $\xi' := \mathcal{A}(v_g)$ and $\eta' := \mathcal{A}(w_g)$. We compute $\mathbf{d}\mathcal{A}_\phi(v_g, w_g)$ by extending v_g and w_g to left invariant vector fields denoted $\bar{\xi}'$ and $\bar{\eta}'$ respectively. We then have

$$\mathbf{d}\mathcal{A}_\phi(v_g, w_g) = \mathbf{d}\mathcal{A}_\phi(\bar{\xi}', \bar{\eta}')(g) \tag{7.2.33}$$
$$= \bar{\xi}'(\mathcal{A}_\phi(\bar{\eta}'))(g) - \bar{\eta}'(\mathcal{A}_\phi(\bar{\xi}'))(g) - \mathcal{A}_\phi([\bar{\xi}', \bar{\eta}'])(g).$$

For the first term we have,

$$\bar{\xi}'(\mathcal{A}_\phi(\bar{\eta}'))(g) = \frac{d}{dt}\Big|_{t=0} \mathcal{A}_\phi(\bar{\eta}')(g \exp t\xi') = \frac{d}{dt}\Big|_{t=0} \langle \phi(g \exp t\xi'), \eta' \rangle$$
$$= \langle \mathbf{d}\phi(g)(v_g), \eta' \rangle = \iota_{\eta'_G} \mathcal{B}(v_g) = \mathcal{B}(g)(T_e L_g \eta', v_g)$$
$$= \mathcal{B}(g)(w_g, v_g) = \mathcal{B}(e)(\eta, \xi)$$

where $\xi := \mathrm{Ad}_g \xi'$ and $\eta := \mathrm{Ad}_g \eta'$ and satisfy $\xi = T_g R_{g^{-1}} v_g$ and $\eta = T_g R_{g^{-1}} w_g$.

Observing that $[\bar{\xi}', \bar{\eta}'](g) = T_e L_g [\xi', \eta']$ we have

$$\mathcal{A}_\phi([\bar{\xi}', \bar{\eta}']) = \langle \phi(g), [\xi', \eta'] \rangle = \left\langle \phi(g), \mathrm{Ad}_{g^{-1}} [\xi, \eta] \right\rangle,$$

and therefore, putting these together we obtain

$$\mathbf{d}\mathcal{A}_\phi(v_g, w_g) = -2\mathcal{B}(e)(\xi, \eta) - \left\langle \phi(g), \mathrm{Ad}_{g^{-1}} [\xi, \eta] \right\rangle. \tag{7.2.34}$$

Next we compute

$$\mathbf{d}\mathcal{A}(\cdot)^* \mu(g)(v_g, w_g) = \mathbf{d}\mathcal{A}(\cdot)^* \mu(\bar{\xi}', \bar{\eta}')(g)$$
$$= \bar{\xi}'(\mathcal{A}(\cdot)^* \mu(\bar{\eta}'))(g) - \bar{\eta}'(\mathcal{A}(\cdot)^* \mu(\bar{\xi}'))(g) - \mathcal{A}(\cdot)^* \mu([\bar{\xi}', \bar{\eta}'])(g).$$

The first two terms are zero since $\mathcal{A}$, being the left Maurer-Cartan form satisfies $\iota_{\bar{\eta}'} \mathcal{A}(g) = \eta'$, i.e. is a constant function on G since the vector field $\bar{\eta}'$ is left invariant. For the last term we have

$$\mathcal{A}(\cdot)^* \mu([\bar{\xi}', \bar{\eta}']) = \langle \mu, \mathcal{A}(T_e L_g [\xi', \eta']) \rangle$$
$$= \langle \mu, [\xi', \eta'] \rangle = \left\langle \mu, \mathrm{Ad}_{g^{-1}} [\xi, \eta] \right\rangle$$

and so, combining these formulas, the right hand side of equation (7.2.32) is

$$\left[-\mathcal{B} - \mathbf{d}(\mathcal{A}_\phi + \mathcal{A}(\cdot)^* \mu) \right](g)(v_g, w_g) = \tag{7.2.35}$$
$$- \mathcal{B}(e)(\xi, \eta) + 2\mathcal{B}(e)(\xi, \eta) + \left\langle \phi(g), \mathrm{Ad}_{g^{-1}} [\xi, \eta] \right\rangle + \left\langle \mu, \mathrm{Ad}_{g^{-1}} [\xi, \eta] \right\rangle$$

We now work out the left hand side of equation (7.2.32). Given $v_g \in T_g G$ and again let $\xi := T_g R_{g^{-1}} v_g$ and $\xi' := T_g L_{g^{-1}} v_g$. Let $s(t)$ be a curve

through e such that $\frac{d}{dt}\big|_{t=0} s(t) = \xi'$ so that $\frac{d}{dt}\big|_{t=0} s(t)g = v_g$. Denoting $\lambda := \psi(g) = \mathrm{Ad}^*_{g^{-1}}\mu + \sigma^{T^*G}(g^{-1})$, we then have,

$$
\begin{aligned}
T_g\psi(v_g) &= \frac{d}{dt}\bigg|_{t=0} \left(\mathrm{Ad}^*_{g^{-1}s(t)^{-1}}\mu + \sigma^{T^*G}(g^{-1}s(t)^{-1}) \right) \\
&= \frac{d}{dt}\bigg|_{t=0} \left(\mathrm{Ad}^*_{s(t)^{-1}}\mathrm{Ad}^*_{g^{-1}}\mu + \mathrm{Ad}^*_{s(t)^{-1}}\sigma^{T^*G}(g^{-1}) + \sigma^{T^*G}(s(t)^{-1}) \right) \\
&= -\mathrm{ad}^*_\xi\lambda + T_e\sigma^{T^*G}(-\xi),
\end{aligned}
$$

and therefore,

$$
\begin{aligned}
\psi^*\omega^+_{\mathcal{B}}(g)(v_g, w_g) &= \omega(\lambda)\left(\mathrm{ad}^*_{-\xi}\lambda - \Sigma^{T^*G}(-\xi, \cdot), \mathrm{ad}^*_{-\eta}\lambda - \Sigma^{T^*G}(-\eta, \cdot) \right) \\
&= \langle\lambda, [\xi, \eta]\rangle - \Sigma^{T^*G}(\xi, \eta) \\
&= \left\langle \mathrm{Ad}^*_{g^{-1}}\mu + \sigma^{T^*G}(g^{-1}), [\xi, \eta] \right\rangle + \mathcal{B}(e)(\xi, \eta) \\
&\quad + \langle\phi(e), [\xi, \eta]\rangle \\
&= \mathcal{B}(e)(\xi, \eta) + \left\langle \mu, \mathrm{Ad}_{g^{-1}}[\xi, \eta] \right\rangle + \langle\phi(e), [\xi, \eta]\rangle \\
&\quad - \left\langle \sigma^{T^*G}(g), \mathrm{Ad}_{g^{-1}}[\xi, \eta] \right\rangle \\
&= \mathcal{B}(e)(\xi, \eta) + \left\langle \mu, \mathrm{Ad}_{g^{-1}}[\xi, \eta] \right\rangle + \left\langle \phi(g), \mathrm{Ad}_{g^{-1}}[\xi, \eta] \right\rangle \\
&= \left(-\mathcal{B} - \mathbf{d}(\mathcal{A}_\phi + \mathcal{A}(\cdot)^*\mu) \right)(g)(v_g, w_g),
\end{aligned}
$$

where we have used equation (7.2.29) and the following facts about the cocycle:

(i) $\sigma^{T^*G}(g^{-1}) = -\mathrm{Ad}^*_{g^{-1}}\sigma^{T^*G}(g)$

(ii) $\sigma^{T^*G}(g) = -\phi(hg) + \mathrm{Ad}^*_g(\phi(h))$.

Finally, since $\psi = \bar\psi \circ \pi_{G,\widetilde{G}_\mu}$ and $\pi^*_{G,\widetilde{G}_\mu}(-\mathcal{B}_\mu) = -\mathcal{B} - \mathbf{d}(\mathcal{A}_\phi + \mathcal{A}(\cdot)^*\mu)$ we have

$$
\pi^*_{G,\widetilde{G}_\mu}\bar\psi^*\omega^+_{\mathcal{B}} = \psi^*\omega^+_{\mathcal{B}} = \pi^*_{G,\widetilde{G}_\mu}(-\mathcal{B}_\mu), \tag{7.2.36}
$$

from which it follows that $\bar\psi^*\omega^+_{\mathcal{B}} = -\mathcal{B}_\mu$ since $\pi_{G,\widetilde{G}_\mu}$ is a surjective submersion. ∎

Remarks.   Of course there is a *left* invariant version of the theorem as well.

That $\widetilde{\mathcal{O}}_\mu$ is a symplectic manifold with the given symplectic form is proved *directly*, as opposed to the present reduction proof, in [HRed], Theorem 4.5.31.

We also note that, since the point reduced spaces for $(T^*G, \Omega - \pi^*_G\mathcal{B})$ are the symplectic leaves of the Poisson reduced space, namely $\mathfrak{g}^*$ with the magnetic Lie-Poisson bracket given by equation (7.2.1), one can alternatively compute the symplectic structure from this point of view. This approach was taken in Theorem 6.2.2.

8

Stages and Coadjoint Orbits of Central Extensions

This chapter addresses the basic theory of symplectic reduction by stages for central extensions. Examples are given in the following Chapter.

The main feature of this theory is that already after the first reduction, one encounters curvature, or magnetic terms and this complicates the subsequent reductions. To deal with this situation, we use the theory developed in the preceding chapter. The same sort of phenomenon also occurs in Lagrangian reduction by stages, as presented in Cendra, Marsden, and Ratiu [2001a].

An important motivating question is to determine the structure of the coadjoint orbits of a central extension. This question can be viewed as an application of the theory of reduction by stages for the action of a central extension $\widehat{G}$ of a group G on its cotangent bundle $T^*\widehat{G}$ by cotangent lift. In the reduction by stages context, we first reduce by the central subgroup and, by an application of cotangent bundle reduction theory, arrive at T^*G but not with the canonical symplectic structure. Instead we will find the presence of a magnetic term which is related to the Lie algebra two-cocycle on $\mathfrak{g}$. This is the subject of §8.1.

The tools used to carry this out are as follows. We consider the $\mathbb{R}$-principal bundle $\widehat{G} \to G$ and derive a formula for the mechanical connection associated to a given right invariant metric on the group $\widehat{G}$. Computing the curvature of this connection, we obtain a two-form $\mathcal{B}$ on G, which we show is simply the right invariant extension of the Lie algebra two-cocycle C from the identity element $e \in G$ to G.

To carry out the second stage reduction, we shall make use of the results of the preceding Chapter (see especially §7.1), namely, the general question

of cotangent bundle reduction theory for the case in which the symplectic form is modified by a magnetic term. With the magnetic cotangent bundle theory at hand, we proceed with the problem of reducing $T^*\widehat{G}$ by stages in §8.2. Having determined the first stage reduced space in §8.1, recall that the general reduction by stages theory in Chapter 5 guarantees that we can induce a (possibly nonequivariant) momentum map for the symplectic action of G on $(T^*G, \Omega - \pi_G^*\mathcal{B})$. However, in the preceding Chapter, we obtained a more specific formula for this momentum map that is special to this case. In the course of doing this, we also obtain the following interpretation of the triple[1] of cocycles, (B, C, σ) (the group, Lie algebra, and momentum map cocycles respectively). As we have mentioned, C arises as the curvature form $\mathcal{B}$ of a mechanical connection on $\widehat{G} \to G$, evaluated at the identity. As in the general theory, the momentum map at this second stage level can be nonequivariant and we derive in equation (8.2.12), a formula for its cocycle, which is expressed in terms of appropriate duals of first derivatives of the group two-cocycle B. Differentiation of this one-cocycle returns the Lie algebra cocycle C.

The final reduced spaces are of course coadjoint orbits for $\widehat{G}$; the main result in this respect is Theorem 8.2.1, which shows that these coadjoint orbits are symplectomorphic with the affine orbits in $\mathfrak{g}^*$ that were studied in Theorem 6.2.2. As in that result, the symplectic structure on these orbits has the interesting form of the usual coadjoint orbit symplectic structure (see equation (1.1.16)) modified by a Lie algebra two-cocycle. Theorem 8.2.1 also shows that the magnetic two-form on the coadjoint orbits for central extensions is obtainable from a two form on G that drops to a symplectic form on the homogeneous space $G/\widetilde{G}^\nu_\mu$, which is a generalization of the well known corresponding result for coadjoint orbits obtained by equivariant cotangent bundle reduction (see, for instance, Marsden and Ratiu [1999], Corollary 14.3.7).

8.1   Stage One Reduction for Central Extensions

Let G be a group and let

$$0 \to \mathbb{R} \to \widehat{G} \to G \to \{e\} \tag{8.1.1}$$

be a nontrivial central extension of G by $\mathbb{R}$. We assume that topologically $\widehat{G}$ is a product $G \times \mathbb{R}$ with group structure given by

$$(g, \alpha)(h, \beta) = (gh, \alpha + \beta + B(g, h)), \tag{8.1.2}$$

where $B : G \times G \to \mathbb{R}$ is a group two-cocycle.

---

[1] We note that Iglesias [1995] considers the same triple in a different context.

Remark.   There is a similar theory to that developed below for extensions by S^1 rather than by $\mathbb{R}$. There is little change in the theory, but various formulas have to be interpreted correctly. For example the addition $\alpha + \beta + B(g, h)$ needs to be interpreted as addition modulo $\mathbb{Z}$ if one is thinking of S^1 as $\mathbb{R}/\mathbb{Z}$ and as multiplication of complex numbers $e^{i\alpha}$ if one is thinking of S^1 as the unit circle in the plane. In the first case, the exponential is the map $\exp : \alpha \in \mathbb{R} \mapsto [\alpha] \in \mathbb{R}/\mathbb{Z}$ and in the second it is the map $\exp : \alpha \in \mathbb{R} \mapsto e^{i\alpha} \in S^1$.

Recall from §6.2 that B is a group two-cocycle if and only if it satisfies the two-cocycle condition

$$B(g, h) + B(gh, k) = B(g, hk) + B(h, k) \tag{8.1.3}$$

for any $g, h, k \in G$ which is equivalent to the associativity of the group multiplication in $\widehat{G}$. As discussed at the beginning of §6.1, we can assume that B is normalized, that is, it satisfies also $B(g, e) = B(e, g) = 0$ for all $g \in G$. Recall also that (8.1.3) implies that $B(g, g^{-1}) = B(g^{-1}, g)$ for any $g \in G$.

We shall also view $\widehat{G}$ as a smooth principal $\mathbb{R}$-bundle over G. We do this by identifying $\widehat{G}/\mathbb{R}$ with G as follows: we identify the class $[g, \alpha]_{\mathbb{R}} \in \widehat{G}/\mathbb{R}$ with $g \in G$. This identification is a group isomorphism. Note, however, that we do not attempt to realize G as a subgroup of $\widehat{G}$.

Recall also from §6.2 that the Lie algebra of $\widehat{G}$ has as underlying vector space the product $\widehat{\mathfrak{g}} := \mathfrak{g} \times \mathbb{R}$ with commutator given by

$$[(\xi, v), (\eta, w)] = ([\xi, \eta], C(\xi, \eta)), \tag{8.1.4}$$

where $C : \mathfrak{g} \times \mathfrak{g} \to \mathbb{R}$ is the Lie algebra two-cocycle associated with B by the formula

$$C(\xi, \eta) = \left. \frac{d^2}{dsdt} \right|_{t=s=0} (B(g(t), h(s)) - B(h(s), g(t))), \tag{8.1.5}$$

where $t \mapsto g(t)$ and $s \mapsto h(s)$ are curves through $e \in G$ with tangent vectors $\dot{g}(0) = \xi$ and $\dot{h}(0) = \eta$.

Relation to the General Theory.   The notation here corresponds to that of the general reduction by stages theory in §5.2 as follows. The "big group" M is $\widehat{G}$ and the normal subgroup N is $\mathbb{R}$. Thus, in the first reduction, $\nu \in \mathbb{R}$ and, since $\mathbb{R}$ is Abelian, $N_\nu = \mathbb{R}$. Since $\mathbb{R}$ is in the center, conjugation leaves $\mathbb{R}$ pointwise invariant and hence the dual of the derivative at the identity leaves each point ν invariant. Thus, $M_\nu = \widehat{G}$ and therefore, $M_\nu/N_\nu = G$.

Mechanical Connections on Central Extensions.   Recall that to carry out cotangent bundle reduction, as described in §2.2 and §2.3, one needs to have a connection on the relevant principal bundle, which in our

case is $\widehat{G} \to G$. We shall in fact use the mechanical connection associated to a particular group invariant metric that we will choose, with the associated mechanical connection given by the general equation (2.1.4).

In fact, we will choose a (not necessarily Ad-invariant) inner product $\langle \cdot, \cdot \rangle_e$ on the Lie algebra $\mathfrak{g}$ and define one on $\widehat{\mathfrak{g}} = \mathfrak{g} \oplus \mathbb{R}$ by

$$\langle (\xi, v), (\eta, w) \rangle_{(e,0)} = \langle \xi, \eta \rangle_e + vw. \tag{8.1.6}$$

for (ξ, v) and $(\eta, w) \in \widehat{\mathfrak{g}}$.

8.1.1 Theorem. *Let $\mathcal{A}^1$ (the "one" standing for "first stage") be the mechanical connection on the right $\mathbb{R}$-principal bundle $\widehat{G} \to G$ associated to the right invariant metric on $\widehat{G}$, which equals (8.1.6) at the identity.*

(i) *Then*

$$\mathcal{A}^1(g, \alpha)(X_g, a_\alpha) = a_\alpha + D_1 B(g, g^{-1})(X_g), \tag{8.1.7}$$

where $(g, \alpha) \in \widehat{G}$ and $(X_g, a_\alpha) \in T_{(g,\alpha)}\widehat{G}$. This formula may be interpreted as saying that $\mathcal{A}^1$ equals the projection onto the $\mathbb{R}$-component of the right Maurer–Cartan form on $\widehat{G}$ (see Theorem 2.1.14).

(ii) *The one-form $\mathcal{A}^1$ on $\widehat{G}$ is right $\widehat{G}$-invariant.*

(iii) *The curvature of this connection at the identity equals the Lie algebra two-cocycle C.*

Remark. Notice that although the connection $\mathcal{A}^1$ is defined in terms of a metric on $\mathfrak{g}$, remarkably, the formula (8.1.7) does not involve that metric.

Proof. To prove (i) apply Theorem 2.1.15 to the Lie group $\widehat{G}$ and the normal subgroup $\mathbb{R}$. We have the right principal bundle $\widehat{G} \to G$ and we compute the mechanical connection from (2.1.15). Since $\mathbb{R}$ lies in the center of $\widehat{G}$, the action of $\widehat{G}$ on $\mathbb{R}$ is trivial, that is, $\mathrm{Ad}_{\widehat{g}}^{\mathbb{R}} = \mathrm{id}$ for any $\widehat{g} \in \widehat{G}$. The spaces $\mathfrak{g} \times \{0\}$ and $\{0\} \times \mathbb{R}$ are mutually orthogonal and hence $\mathcal{P}_{\mathbb{R}} : \mathfrak{g} \times \mathbb{R} \to \mathbb{R}$ is the projection onto the second factor. By (8.1.2), the derivative of the right translation map is given by

$$T_{(h,\beta)} R_{(g,\alpha)}(Y_h, b_\beta) = (Y_h \cdot g, b_\beta + D_1 B(h, g)(Y_h)). \tag{8.1.8}$$

and hence, since $(g, \alpha)^{-1} = (g^{-1}, -\alpha - B(g, g^{-1}))$, we get

$$T_{(g,\alpha)} R_{(g,\alpha)^{-1}}(X_g, a_\alpha) = \left(T_g R_{g^{-1}}(X_g), a_\alpha + D_1 B(g, g^{-1})(X_g) \right),$$

Formula (2.1.15) then yields

$$\mathcal{A}^1(g, \alpha)(X_g, a_\alpha) = a_\alpha + D_1 B(g, g^{-1})(X_g)$$

which proves (8.1.7).

Next we prove (ii). From Theorem 2.1.15(ii), we have $R_h^* \mathcal{A} = \mathrm{Ad}_{h^{-1}} \circ \mathcal{A}$. But in our case, $\mathcal{A} = \mathcal{A}^1$ takes values in the center $\mathbb{R}$ of $\widehat{\mathfrak{g}}$ and so $\mathrm{Ad}_{h^{-1}} \circ \mathcal{A}^1 = \mathcal{A}^1$ for any $h \in \widehat{G}$. This proves that $\mathcal{A}^1$ is indeed right $\widehat{G}$-invariant.

Finally we turn to the proof of (iii). Since $\mathbb{R}$ is in the center of $\widehat{G}$, we have $\mathrm{ad}_{(\xi,u)}(0,v) = 0$ for all $\xi \in \mathfrak{g}$ and $u, v \in \mathbb{R}$. Thus the first, second, and fourth term in (2.1.17) vanish. By (8.1.4), $\mathcal{P}_{\mathbb{R}}[(\xi, u), (\eta, v)] = C(\xi, \eta)$ and thus (2.1.17) becomes

$$\mathrm{curv}_{\mathcal{A}^1}(g, \alpha) \left((X_g, a_\alpha), (Y_g, b_\alpha) \right) = C \left(T_g R_{g^{-1}}(X_g), T_g R_{g^{-1}}(Y_g) \right)$$

as required. ∎

First Stage Reduction.   Next, starting with the central extension $\widehat{G} \to G$ consider the right cotangent lifted action

$$T^* \widehat{G} \times \widehat{G} \to T^* \widehat{G}.$$

By the standard cotangent bundle reduction theorem (see Corollary 1.1.4), the symplectic reduced spaces for this action are symplectomorphic to coadjoint orbits, that is,

$$\mathbf{J}^{-1}(\mu)/\widehat{G}_\mu \simeq \mathcal{O}_\mu,$$

where $\mathcal{O}_\mu$ denotes the coadjoint orbit through $\mu \in \widehat{\mathfrak{g}}^*$.

The general strategy is to apply the reduction by stages theorem for centrally extended groups (Theorem 6.3.2) to factor the reduction by $\widehat{G}$ through the reduction by the action of $\mathbb{R}$ (done in this section) and then by the action of G (done in §8.2).

8.1.2 Theorem.   *Consider the right cotangent lift of the $\mathbb{R}$-action on $T^* \widehat{G}$ with its canonical symplectic structure. Let $\nu \in \mathbf{J}_{\mathbb{R}}(T^* \widehat{G}) \subset \mathbb{R}$. Then there is a right G-equivariant symplectic diffeomorphism*

$$\mathbf{J}_{\mathbb{R}}^{-1}(\nu)/\mathbb{R} \simeq \left(T^* G, \Omega - \nu \pi_G^* \mathcal{B}^1 \right), \tag{8.1.9}$$

where $\mathcal{B}^1$ is the closed two-form (magnetic term) obtained by dropping $\mathbf{d}\mathcal{A}^1$ to $G = \widehat{G}/\mathbb{R}$ and $\pi_G : T^* G \to G$ is the cotangent bundle projection.

Proof.   That the spaces in (8.1.9) are symplectomorphic and that $\mathbf{d}\mathcal{A}^1$ drops to the quotient, follows from the general theory of cotangent bundle reduction for Abelian symmetry groups (see Remark 3 following Theorem 2.2.1 and Theorem 2.2.3).

To show the equivariance it suffices to show that for $\widehat{g}, \widehat{h} \in \widehat{G}$ and $\widehat{p}_{\widehat{g}} \in T_{\widehat{g}}^* \widehat{G}$, we have

$$\mathrm{shift}_\nu \left(\widehat{g} \cdot \widehat{h} \right)(\widehat{p}_{\widehat{g}} \cdot \widehat{h}) = (\mathrm{shift}_\nu \left(\widehat{g} \right)(\widehat{p}_{\widehat{g}})) \cdot \widehat{h},$$

where, following the notation of the general theory from Chapter 2,

$$\mathrm{shift}_\nu \left(\widehat{g} \right) : \mathbf{J}_{\mathbb{R}}^{-1}(\nu) \to \mathbf{J}_{\mathbb{R}}^{-1}(0) \quad \text{is defined by} \quad \mathrm{shift}_\nu \left(\widehat{g} \right)(\widehat{p}_{\widehat{g}}) = \widehat{p}_{\widehat{g}} - \nu \mathcal{A}^1(\widehat{g}).$$

However, from the $\widehat{G}$-invariance of $\mathcal{A}^1$ (i.e., $\mathcal{A}^1(\widehat{g}\cdot\widehat{h})(X_{\widehat{g}}\cdot\widehat{h}) = \mathcal{A}^1(\widehat{g})(X_{\widehat{g}})$), that was proved in Theorem 8.1.1(iii), we have for any $X_{\widehat{g}\widehat{h}} \in T_{\widehat{g}\widehat{h}}\widehat{G}$ and any $\widehat{g}, \widehat{h} \in \widehat{G}$,

$$\mathcal{A}^1(\widehat{g}\cdot\widehat{h})(X_{\widehat{g}\widehat{h}}) = \mathcal{A}^1(\widehat{g}\cdot\widehat{h})((X_{\widehat{g}\widehat{h}}\cdot\widehat{h}^{-1})\cdot\widehat{h})$$
$$= \mathcal{A}^1(\widehat{g})(X_{\widehat{g}\widehat{h}}\widehat{h}^{-1}) = \left\langle \mathcal{A}^1(\widehat{g})\cdot\widehat{h}, X_{\widehat{g}\widehat{h}} \right\rangle,$$

where, on the right hand side of the last equality, we use $\cdot$ to denote the action of $\widehat{G}$ on $T^*\widehat{G}$ by cotangent lift. This shows that shift is $\widehat{G}$-equivariant.

We need to check that the quotient map defined by shift_ν, which we denote by

$$\widetilde{\mathrm{shift}}_\nu : \mathbf{J}_{\mathbb{R}}^{-1}(\nu)/\mathbb{R} \to \mathbf{J}_{\mathbb{R}}^{-1}(0)/\mathbb{R},$$

is equivariant with respect to the action of $G = M_\nu/N_\nu$. From the diagram

$$
\begin{array}{ccc}
\mathbf{J}_{\mathbb{R}}^{-1}(\nu) & \xrightarrow{\;\mathrm{shift}_\nu\;} & \mathbf{J}_{\mathbb{R}}^{-1}(0) \\
\Big\downarrow{\scriptstyle \pi_{\mathbb{R}}^\nu} & & \Big\downarrow{\scriptstyle \pi_{\mathbb{R}}^0} \\
\mathbf{J}_{\mathbb{R}}^{-1}(\nu)/\mathbb{R} & \xrightarrow[\;\widetilde{\mathrm{shift}}_\nu\;]{} & \mathbf{J}_{\mathbb{R}}^{-1}(0)/\mathbb{R}.
\end{array}
$$

that defines $\widetilde{\mathrm{shift}}_\nu$, the $\widehat{G}$-equivariance of the shift map, and writing $[\widehat{g}] = g$ and $[z]_\nu = \pi_{\mathbb{R}}^\nu(z)$, we get

$$\widetilde{\mathrm{shift}}_\nu([z]_\nu \cdot g) = \widetilde{\mathrm{shift}}_\nu([z\cdot\widehat{g}]_\nu) = \pi_{\mathbb{R}}^0(\mathrm{shift}_\nu(z\cdot\widehat{g}))$$
$$= \pi_{\mathbb{R}}^0(\mathrm{shift}_\nu(z)\cdot\widehat{g}) = \pi_{\mathbb{R}}^0(\mathrm{shift}_\nu(z))\cdot g$$
$$= \widetilde{\mathrm{shift}}_\nu([z]_\nu)\cdot g,$$

where we have used the fact that $[z]_\nu \cdot g = [z\cdot\widehat{g}]_\nu$ (and similarly for $\nu = 0$), which defines the action of G on $\mathbf{J}_{\mathbb{R}}^{-1}(\nu)/\mathbb{R}$ (and $\mathbf{J}_{\mathbb{R}}^{-1}(0)/\mathbb{R}$ respectively). This proves the statement. ∎

A result we shall need later is the following.

8.1.3 Proposition. *The two-form $\mathcal{B}^1$ on G is right G-invariant and its value at the identity is given by $\mathcal{B}^1(e) = C$.*

Remark. Notice that we are dealing here with the *right* invariance of $\mathcal{B}^1$ on G, while C is the Lie algebra two-cocycle associated with the *left* (or standard) Lie algebra of G. Of course these theorems have counterparts where one uses right Lie algebras, left actions, etc.

Proof. The fact that the value of $\mathcal{B}^1(e)$ is C follows from part (ii) of Theorem 8.1.1.

Obtain the right invariance statement, we use the following general fact.

8.1.4 Lemma. *Let a Lie group K act freely and properly on a manifold R. Let α be a one form on R such that $\mathbf{d}\alpha$ is K-invariant. Let $L \subset K$ be a closed normal subgroup of K and suppose that $\mathbf{i}_{\xi_S}\mathbf{d}\alpha = 0$ for any $\xi \in \mathfrak{l}$ so that $\mathbf{d}\alpha$ drops to R/L giving a closed two-form β; that is, $\pi_{R,L}^*\beta = \mathbf{d}\alpha$, where $\pi_{R,L} : R \to R/L$ is the projection to the quotient. Recall from Lemma 5.2.1 that K/L acts in a free and proper manner on R/L.*

Then β is K/L-invariant.

Proof. Recall from that the induced action on R/L is characterized as follows. If $\Psi_k : R \to R$ denotes the given action of a group element $k \in K$, and $\Psi_{[k]_L}^L : R/L \to R/L$ denotes the quotient action of an element $[k]_L \in K/L$, then

$$\Psi_{[k]_L}^L \circ \pi_{R,L} = \pi_{R,L} \circ \Psi_k.$$

To show that $\left(\Psi_{[k]_L}^L\right)^* \beta = \beta$, it suffices to show that

$$\pi_{R,L}^* \left(\Psi_{[k]_L}^L\right)^* \beta = \pi_{R,L}^*\beta$$

because $\pi_{R,L}$ is a surjective submersion. But the left hand side equals

$$\left(\Psi_{[k]_L}^L \circ \pi_{R,L}\right)^* \beta = \left(\pi_{R,L} \circ \Psi_k\right)^* \beta = \Psi_k^* \pi_{R,L}^*\beta = \Psi_k^* \mathbf{d}\alpha = \mathbf{d}\alpha$$

since α is K-invariant. By construction of β, the right hand side is also $\mathbf{d}\alpha$.
▼

Returning to the proof of the Proposition, the right invariance is an instance of the above lemma together with the $\widehat{G}$-invariance of $\mathcal{A}^1$ (and hence invariance of $\mathbf{d}\mathcal{A}^1$) that was proved in Theorem 8.1.1(iii). ■

8.2 Reduction by Stages for Central Extensions

We now return to the setting of §8.1 and we will now carry out the second stage reduction by making use of the results of the preceding two sections.

Recall that in the first stage reduction, Theorem 8.1.2 dealt with the reduction by the right cotangent lifted action of the center $\mathbb{R}$ on $T^*\widehat{G}$ with its canonical symplectic structure. We made use of a *first stage* mechanical connection $\mathcal{A}^1 \in \Omega^1(\widehat{G})$ on the principal $\mathbb{R}$-bundle $\widehat{G} \to G$. Recall from equation (8.1.7) that this is given by

$$\mathcal{A}^1(g, \alpha)(X_g, a_\alpha) = a_\alpha + D_1 B(g, g^{-1})(X_g), \qquad (8.2.1)$$

where B is the group two-cocycle defining the group extension.

The reduction was carried out at a point $\nu \in \mathbf{J}_{\mathbb{R}}(T^*\widehat{G}) \subset \mathbb{R}$. It was shown that there is a right G-equivariant symplectic diffeomorphism

$$\mathbf{J}_{\mathbb{R}}^{-1}(\nu)/\mathbb{R} \simeq \left(T^*G, \Omega - \nu\pi_G^*\mathcal{B}^1\right), \qquad (8.2.2)$$

where $\mathcal{B}^1$ is the closed two-form (magnetic term) obtained by dropping $\mathbf{d}\mathcal{A}^1$ to $G = \widehat{G}/\mathbb{R}$ and where $\pi_G : T^*G \to G$ is the cotangent bundle projection. We also showed that $\mathcal{B}^1$ is a G-invariant two form on G.

The goal of this section is to carry out point reduction of the symplectic manifold $\left(T^*G, \Omega - \nu\pi_G^*\mathcal{B}^1\right)$ under the cotangent bundle lifted right action of the connected Lie group G. As we saw in Chapter 2 and the preceding two sections, there are two versions of cotangent bundle reduction, an embedding version and a bundle version. We shall explore the application of both of these versions.

As we shall see, both versions lead to a reduced space which is identified with the affine orbit with its magnetic orbit symplectic form, given by equation (7.2.9). Combining this with the reduction by stages theorem will then give the main result of this section, namely the following.

8.2.1 Theorem.   *The coadjoint orbit $\mathcal{O}_{\mu,\nu} \subset \widehat{\mathfrak{g}}^*$ through the point $(\mu, \nu) \in \widehat{\mathfrak{g}}^* \cong \mathfrak{g}^* \times \mathbb{R}$ for the action of the central extension $\widehat{G}$ with its $+$ coadjoint orbit symplectic structure is symplectically diffeomorphic to the right affine orbit symplectic manifold $(\widetilde{\mathcal{O}}_\mu^\nu, \widetilde{\omega}_\mu^\nu)$. Here,*

$$\widetilde{\mathcal{O}}_\mu^\nu = \left\{\mu \cdot g = \mathrm{Ad}_g^* \mu + \sigma^\nu(g) \mid g \in G\right\},$$

where $\sigma^\nu(g) = \nu \left(D_2 B(g, e) + T_e^* L_g D_1 B(g, g^{-1})\right)$. The magnetic affine orbit symplectic structure $\widetilde{\omega}_\mu^\nu$ is given by

$$\widetilde{\omega}_\mu^\nu(\lambda)\left(\mathrm{ad}_\xi^* \lambda + \nu C(\xi, \cdot), \mathrm{ad}_\eta^* \lambda + \nu C(\eta, \cdot)\right) = \langle\lambda, [\xi, \eta]\rangle + \nu C(\xi, \eta), \quad (8.2.3)$$

where a tangent vector to $\widetilde{\mathcal{O}}_\mu^\nu$ at a point $\lambda = \mu \cdot g \in \widetilde{\mathcal{O}}_\mu^\nu$ has the form $\mathrm{ad}_\xi^* \lambda + \nu C(\xi, \cdot)$ for $\xi \in \mathfrak{g}$.

Furthermore, the right affine orbit symplectic manifold $(\widetilde{\mathcal{O}}_\mu^\nu, \widetilde{\omega}_\mu^\nu)$ is symplectomorphic to the homogeneous space $G/\widetilde{G}_\mu^\nu$ with the symplectic form $-\mathcal{B}_\mu^\nu$ defined by

$$\pi_{G,\widetilde{G}_\mu^\nu}^* \mathcal{B}_\mu^\nu = \nu\mathcal{B}^1 + \mathbf{d}(\mathcal{A}_{\phi^\nu} + \mathcal{A}(\cdot)^*\mu) \qquad (8.2.4)$$

where $\mathcal{A}$ is the left Maurer-Cartan form on G, $\mathcal{A}_{\phi^\nu} := \langle \phi^\nu, \mathcal{A} \rangle$ and ϕ^ν is the $\mathcal{B}\mathfrak{g}$ potential defined below in equation (8.2.5).

Remarks.

1. Notice that the right affine orbit symplectic manifold $(\widetilde{\mathcal{O}}_\mu^\nu, \widetilde{\omega}_\mu^\nu)$, which is obtained here via the reduction by stages methodology, coincides with what we obtained "by hand" in Theorem 6.2.2 and by magnetic cotangent bundle reduction in Theorem 7.2.2.

2. As we discuss in the proof, the results on the homogeneous space $G/\widetilde{G}_\mu^\nu$ are the same as the ones we found in last paragraph of Theorem 7.2.2.

3. A direct calculation of the symplectic structure on the coadjoint orbit $\mathcal{O}_{\mu,\nu}$ through $(\mu, \nu) \in \widehat{\mathfrak{g}}^* \cong \mathfrak{g}^* \times \mathbb{R}$ at a point of the form (λ, ν) using the general coadjoint orbit symplectic form gives:

$$
\begin{aligned}
\omega^+_{\mathcal{O}_{(\mu,\nu)}}(\lambda, \nu)(\mathrm{ad}^*_{(\xi,a)}(\lambda, \nu), \mathrm{ad}^*_{(\eta,b)}(\lambda, \nu)) &= \langle (\mu, \nu), [(\xi, a), (\eta, b)] \rangle \\
&= \langle (\mu, \nu), ([\xi, \eta], C(\xi, \eta)) \rangle \\
&= \langle \mu, [\xi, \eta] \rangle + \nu C(\xi, \eta),
\end{aligned}
$$

which is in agreement with (8.2.3).

Proof. We being with some preliminaries. Recall that we are asserting that one can obtain the preceding theorem in two ways. The first is to apply Theorem 7.1.6 to the right G action on $\left(T^*G, \Omega - \nu\pi_G^*\mathcal{B}^1\right)$. The theorem is applied to the following objects:

1. We take Q to be G.

2. The symbol Ω again is the canonical symplectic form; now on T^*G.

3. We take $\mathcal{B}$ to be $\nu\mathcal{B}^1$, which is a closed right G-invariant two-form by Proposition 8.1.3.

4. We take the connection $\mathcal{A}$ on the bundle $Q \to Q/G$, which in our case is the bundle $G \to \{e\}$, to be the left Maurer–Cartan form; in the notation of Theorem 2.1.14, we take $\mathcal{A} = \theta^R$. Explicitly, $\mathcal{A}(g)(v_g) = T_g L_{g^{-1}} v_g$.

The second method will be to apply Theorem 7.2.2. In either case, we must establish the existence of a $\mathcal{B}\mathfrak{g}$-potential, which we do now.

A Formula for the $\mathcal{B}\mathfrak{g}$-Potential. We must find a smooth map $\phi^\nu : G \to \mathfrak{g}^*$ that satisfies

$$
\nu \mathbf{i}_{\xi_G} \mathcal{B}^1 = \mathbf{d} \langle \phi^\nu, \xi \rangle \tag{8.2.5}
$$

for all $\xi \in \mathfrak{g}$. Note that we write ϕ^ν instead of ϕ because the determining equation depends on ν.

8.2.2 Lemma. *A $\mathcal{B}\mathfrak{g}$-potential $\phi^\nu : G \to \mathfrak{g}^*$ is given by*

$$\phi^\nu(g) = -\nu\left(D_2 B(g,e) + T_e^* L_g D_1 B(g,g^{-1})\right). \qquad (8.2.6)$$

That is,

$$\langle \phi^\nu(g), \xi \rangle = -\nu\left(D_2 B(g,e) \cdot \xi + D_1 B(g,g^{-1}) \cdot T_e L_g(\xi)\right), \qquad (8.2.7)$$

for all $\xi \in \mathfrak{g}$.

Proof. Let $\pi_1 : \widehat{G} = G \times \mathbb{R} \to G$, be defined by $\pi_1(g,\alpha) = g$, the projection onto the first factor. Since this is a surjective submersion, it suffices to show that

$$\nu \pi_1^* \mathbf{i}_{\xi_G} \mathcal{B}^1 = \pi_1^* \mathbf{d} \langle \phi^\nu, \xi \rangle. \qquad (8.2.8)$$

Let us next compute the infinitesimal generator for $(\xi, 0)_{\widehat{G}}$ for the right action of $\widehat{G}$ on itself. By definition, it is given by

$$\begin{aligned}
(\xi, 0)_{\widehat{G}}(g, \alpha) &= \left.\frac{d}{dt}\right|_{t=0} (g, \alpha)(\exp(t\xi), 0) \\
&= \left.\frac{d}{dt}\right|_{t=0} (g\exp(t\xi), \alpha + B(g, \exp(t\xi))) \\
&= (T_e L_g \xi, D_2 B(g,e) \cdot \xi)
\end{aligned}$$

where we used the formula (8.1.2) for multiplication in $\widehat{G}$. Recalling that $\xi_G(g) = T_e L_g \xi$ (since we are using *right actions*), note that $(\xi, 0)_{\widehat{G}}$ and ξ_G are π_1-related. With this, we can compute the left hand side of (8.2.8) as follows:

$$\begin{aligned}
\nu \pi_1^* \mathbf{i}_{\xi_G} \mathcal{B}^1 &= \nu \mathbf{i}_{(\xi,0)_{\widehat{G}}} \pi_1^* \mathcal{B}^1 = \nu \mathbf{i}_{(\xi,0)_{\widehat{G}}} \mathbf{d} \mathcal{A}^1 \\
&= -\nu \mathbf{d} \mathbf{i}_{(\xi,0)_{\widehat{G}}} \mathcal{A}^1
\end{aligned}$$

where we have used the fact that $\mathcal{A}^1$ is $\widehat{G}$-invariant (see Theorem 8.1.1(iii)).
    Thus, equation (8.2.8) reduces to

$$-\nu \mathbf{d} \mathbf{i}_{(\xi,0)_{\widehat{G}}} \mathcal{A}^1 = \mathbf{d} \pi_1^* \langle \phi^\nu, \xi \rangle \qquad (8.2.9)$$

This will hold provided that

$$\pi_1^* \langle \phi^\nu, \xi \rangle = -\nu \mathbf{i}_{(\xi,0)_{\widehat{G}}} \mathcal{A}^1. \qquad (8.2.10)$$

From equation (8.1.7) for $\mathcal{A}^1$, we can compute the right hand side as follows:

$$\begin{aligned}
-\nu \mathbf{i}_{(\xi,0)_{\widehat{G}}} \mathcal{A}^1(g, \alpha) &= -\nu \mathcal{A}^1(g, \alpha)\left(T_e L_g \xi, D_2 B(g,e) \cdot \xi\right) \\
&= -\nu\left(D_2 B(g,e) \cdot \xi + D_1 B(g,g^{-1}) \cdot T_e L_g \xi\right)
\end{aligned}$$

Since the right hand side is independent of α, we can take it to be $\langle \phi^\nu(g), \xi \rangle$, which proves the lemma. ▼

Notice that, since $B(e,g) = B(g,e) = 0$ for all $g \in G$, we have $\phi^\nu(e) = 0$.

The Induced Momentum Map.   According to Theorem 7.1.1(i), the G-action on $\left(T^*G, \Omega - \nu\pi_G^*\mathcal{B}^1\right)$ admits a nonequivariant momentum map $\mathbf{J}^\nu : T^*G \to \mathfrak{g}^*$ given by $\mathbf{J}^\nu = \mathbf{J}_{\text{can}} - \phi^\nu \circ \pi_G$, where $\mathbf{J}_{\text{can}} = \mathbf{J}_R$ is the standard momentum map for the G-action relative to the canonical symplectic form. Again, we append the ν to keep in mind that this momentum map depends on ν.

In view of (8.2.6) and the standard canonical momentum map formula for right actions, namely, $\mathbf{J}_{\text{can}}(\alpha_g) = T_e^*L_g(\alpha_g)$, we obtain

$$\mathbf{J}^\nu(\alpha_g) = T_e^*L_g(\alpha_g + \nu D_1 B(g, g^{-1})) + \nu D_2 B(g, e). \qquad (8.2.11)$$

The Cocycle of $\mathbf{J}^\nu$.   According to equations (7.1.2), (8.2.6), and the third equality in (6.2.2), the nonequivariance one-cocycle of $\mathbf{J}^\nu$ is given by

$$\begin{aligned}
\sigma^\nu(g) &= -\phi^\nu(g) + \text{Ad}_g^*(\phi^\nu(e)) \\
&= \nu\left(D_2 B(g, e) + T_e^*L_g D_1 B(g, g^{-1})\right) = \nu\sigma_R(g). \qquad (8.2.12)
\end{aligned}$$

Here we have used the fact that the cocycle formula (7.1.2) does not depend on the point g since G is connected, and the fact that $\phi^\nu(e) = 0$. Note that this argument shows that

$$\sigma^\nu(g) = -\phi^\nu(g). \qquad (8.2.13)$$

Remark.   The corresponding formula for S^1 extensions, using exponential notation is $\sigma^\nu(g) = \exp\left(-i\phi^\nu(g)\right)$.

The Affine Action.   The next thing we do is to write down the associated affine action of G on $\mathfrak{g}^*$ according to the procedures for reducing by a nonequivariant momentum map. Recall from (7.1.5) that this action is given by

$$\begin{aligned}
\mu \cdot g &= \text{Ad}_g^*\mu + \sigma^\nu(g) \\
&= \text{Ad}_g^*\mu + \nu\left(D_2 B(g, e) + T_e^*L_g D_1 B(g, g^{-1})\right) \qquad (8.2.14)
\end{aligned}$$

for $\mu \in \mathfrak{g}^*$ and $g \in G$. The associated isotropy subgroup (to which we also append a superscript ν) is then

$$\widetilde{G}_\mu^\nu = \left\{g \in G \mid \text{Ad}_g^*\mu + \sigma^\nu(g) = \mu\right\}. \qquad (8.2.15)$$

Remark.   In the case of S^1 extensions, one should use the action $\mu \cdot g = \text{Ad}_g^*\mu - \phi^\nu(g)$.

The Nonequivariance Lie Algebra Two-Cocycle.   We compute this from equation (7.1.7) (and again using a subscript ν to indicate the dependence on ν) and recall that it is independent of the point $g \in G$:

$$\begin{aligned}
\Sigma^\nu(\xi, \eta) &= -\langle\phi^\nu(g), [\xi, \eta]\rangle - \nu\mathcal{B}^1(g)(\xi_G(g), \eta_G(g)) \\
&= -\langle\phi^\nu(e), [\xi, \eta]\rangle - \nu\mathcal{B}^1(e)(\xi_G(e), \eta_G(e)) \\
&= -\nu\mathcal{B}^1(e)(\xi, \eta) = -\nu C(\xi, \eta) \qquad (8.2.16)
\end{aligned}$$

where again we used the fact that $\phi^\nu(e) = 0$ and Theorem 8.1.1(ii) which showed that the curvature $\mathcal{B}^1$ of the connection $\mathcal{A}^1$ at the identity is the Lie algebra 2-cocycle C.

These calculations together with equation (7.1.10) show that the isotropy algebra for the affine action of G on $\mathfrak{g}^*$ is given, for a fixed $\mu \in \mathfrak{g}^*$, by

$$\widetilde{\mathfrak{g}}_\mu^\nu = \{\xi \in \mathfrak{g} \mid \mathrm{ad}_\xi^* \mu + \nu C(\xi, \cdot) = 0\}. \tag{8.2.17}$$

We are now ready to complete the proof of Theorem 8.2.1. According to the right invariant version of reduction to coadjoint orbits given in Theorem 1.2.3, the reduction of $T^*\widehat{G}$ at the point (μ, ν) is the coadjoint orbit $\mathcal{O}_{\mu,\nu}$ with its + coadjoint orbit symplectic form. This is the "one-shot" reduction and the strategy now is to realize the same reduced space in a two-step reduction by stages procedure. Then by the main reduction by stages result, namely Theorem 5.2.9, the two resulting manifolds will be symplectically diffeomorphic. The stages hypothesis, which was assumed in that theorem, hold by virtue of Theorem 6.3.2.

The first stage reduction was carried out in the first section of this chapter and was recalled above; it resulted in the symplectic manifold $T^*G, \Omega - \nu\pi_G^*\mathcal{B}^1$. As we remarked earlier, at the second stage, we are to reduce this by the action of M_ν/N_ν; here, $M = \widehat{G}$ and $N = \mathbb{R}$, and since N is Abelian, $N_\nu = N$ and since it is central, $M_\nu = M$. Thus, in the second stage, we reduce by $M_\nu/N_\nu = \widehat{G}/\mathbb{R} = G$.

To obtain the conclusions of the theorem, we can apply either Theorem 7.1.6 or Theorem 7.2.2. In the preliminaries above, we have computed all of the necessary ingredients to apply these results. In either case, it is clear that one ends up with the affine orbit $\widetilde{\mathcal{O}}_\mu^\nu$ (in the embedding approach, see Remark 4 following Theorem 7.1.6). The symplectic form was computed explicitly in Theorem 7.2.2 and gives the stated result.

For the last statement of the theorem, the proof proceeds exactly as in the analogous statement of Theorem 7.2.2. One simply replaces $\mathcal{B}$ with $\nu\mathcal{B}^1$, ϕ with ϕ^ν and σ^{T^*G} with σ^ν. ∎

Remark. This theorem is also proven in [HRed], Corollary 4.5.32, by using a Poisson point of view together with direct verifications.

9
Examples

We begin this chapter by revisiting the example of the Heisenberg group from §5.1 to illustrate the general theory developed in Chapter 8. As will become evident, now that the general theory is at hand, when it is applied to this example, the results done by hand in §5.1 can be obtained very quickly. This example is interesting because it is a simple, yet nontrivial, example of a group extension that is not a semidirect product.

The second example in this chapter, also a central extension of an Abelian group, is an extension of the loop group of the circle, namely the group of maps of S^1 to S^1 with pointwise multiplication. This group appears in the theory of integrable systems. We also compute the frozen Poisson structures for the extended loop group and show their relevance for the KdV equation and the Camassa-Holm equation.

The third example is the oscillator group, which is a central extension of the nonabelian group SE(2). Both the extension of the loop group and the oscillator group naturally lead into the fourth example, namely the Bott–Virasoro group, which is a central extension of the group of diffeomorphisms of the circle. We briefly indicate how the method of reduction by stages leads to the classification of the coadjoint orbits of the Bott–Virasoro group.

There are many other examples of central extensions that are relevant to physical applications. For example, Holm and Kupershmidt [1982, 1983b, 1988] (see also Cendra, Marsden, and Ratiu [2003]) contain affine Poisson brackets of the type (6.2.22) that are Lie-Poisson brackets of central extensions restricted to the Poisson submanifold which is the dual of the Lie algebra that one is centrally extending by $\mathbb{R}$ (see Theorems 7.2.2 and 8.2.1). We shall not present these examples since they are easy direct applications

of the theory developed Chapters 7 and 8.

9.1   The Heisenberg Group Revisited

In this section, we repeat the direct calculations that were done for the Heisenberg group in §5.1 by making use of the general theory of the preceding chapter. Recall that the Heisenberg group is the central extension of $\mathbb{R}^2$ by its standard symplectic form ω. That is, $B = \omega/2$, where, as usual, B is the Lie group two-cocycle.[1] Therefore, the group extension is $H = \widehat{\mathbb{R}}^2 = \mathbb{R}^2 \oplus \mathbb{R}$ with group multiplication

$$(u, \alpha)(v, \beta) = \left(u + v, \alpha + \beta + \frac{1}{2}\omega(u, v) \right).$$

The Lie algebra bracket is then computed to be

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

which tells us that $C(X, Y) = \omega(X, Y)$.

First Reduced Space.   The first reduced space is, by Theorem 8.1.2,

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

where $\pi_{\mathbb{R}^2} : T^*\mathbb{R}^2 \to \mathbb{R}^2$ is the cotangent bundle projection. By Proposition 8.1.3, $\mathcal{B}^1$ is the right invariant two-form that equals $C = \omega$ at the identity. Being right invariant on a vector group, $\mathcal{B}^1$ is therefore constant and so $\mathcal{B}^1 = \omega$.

Another way to compute $\mathcal{B}^1$ is to use the fact that $\mathcal{B}^1$ is the closed two form obtained by dropping $\mathbf{d}\mathcal{A}^1$ to $\mathbb{R}^2 = H/\mathbb{R}$. By equation (8.1.7) and the fact that $B = \omega/2$, we have

$$\mathcal{A}^1(u, \alpha)(u_1, a_1) = a_1 + D_1 B(u, -u)(u_1)$$

$$= a_1 - \frac{1}{2}\omega(u_1, u).$$

Using the general formula for the exterior derivative of a one-form $\mathcal{A}$ on a vector space V, where $\mathcal{A}$ is thought of as a map of V to V^*, namely

$$\mathbf{d}\mathcal{A}(v)(v_1, v_2) = \mathbf{D}\mathcal{A}(v) \cdot v_1(v_2) - \mathbf{D}\mathcal{A}(v) \cdot v_2(v_1),$$

we get,

$$\mathbf{d}\mathcal{A}^1(u, \alpha)((u_1, a_1), (u_2, a_2)) = \omega(u_1, u_2).$$

Thus, again we get $\mathcal{B}^1 = \omega$. Therefore, the first reduced space is

$$\mathbf{J}_{\mathbb{R}}^{-1}(\nu)/\mathbb{R} \simeq \left(T^*\mathbb{R}^2, \Omega - \nu\pi_{\mathbb{R}^2}^*\omega\right).$$

---

[1] A more general cocycle associated with groups of symplectic transformations is given in Ismagilov, Losik, and Michor [2006].

The Second Reduced Space.   We next compute the $\mathcal{B}\mathfrak{g}$-potential for the induced $\mathbb{R}^2$-action on the base of the first reduced space by appealing to equation (8.2.6). Namely,

$$\langle \phi^\nu(u), \xi \rangle = -\nu(D_2 B(u, 0) \cdot \xi + D_1 B(u, -u) \cdot T_0 L_u \cdot \xi) = -\nu \omega(u, \xi).$$

Thus, $\phi^\nu(u) = \nu(u_y, -u_x)$, where $u = (u_x, u_y)$. By equation (8.2.13), the cocycle is

$$\sigma^\nu(u) = -\nu(u_y, -u_x). \tag{9.1.1}$$

According to Theorem 8.2.1, the coadjoint orbits through the point $(\mu, \nu) \in \mathbb{R}^2 \times \mathbb{R}$ for the Heisenberg group $H = \widehat{\mathbb{R}}^2$ are symplectomorphic to the following submanifolds of $\mathbb{R}^2$:

$$\widetilde{\mathcal{O}}_\mu^\nu = \left\{ \mu + \sigma^\nu(u) \mid u \in \mathbb{R}^2 \right\} = \left\{ \mu - \nu(u_y, -u_x) \mid u \in \mathbb{R}^2 \right\}.$$

Thus, there are two cases:

- A point: $\widetilde{\mathcal{O}}_\mu^{\nu=0} = \{\mu\}$;

- A two dimensional orbit: $\widetilde{\mathcal{O}}_\mu^{\nu \neq 0} = \mathbb{R}^2$.

The symplectic structure on the two-dimensional orbit $\widetilde{\mathcal{O}}_\mu^{\nu \neq 0} = \mathbb{R}^2$ is given by (8.2.3), namely

$$\widetilde{\omega}_\mu^\nu \left(\mu - \nu(v_y, -v_x) \right) (\nu \omega(\xi, \cdot), \nu \omega(\eta, \cdot)) = \nu \omega(\xi, \eta). \tag{9.1.2}$$

In Proposition 5.1.2 the tangent space to the orbit was parametrized by vectors X and Y, which were identified with Lie algebra elements, called here ξ and η. Thus, we see that the symplectic structure here agrees with that given in Proposition 5.1.2.

9.2   A Central Extension of $\mathrm{L}(S^1)$

In this section we shall carry out reduction by stages for a central extension of the loop group of S^1. We begin by recalling the main features of this group.

A Central Extension of the Loop Group of the Circle.   Consider the infinite dimensional Abelian group $\mathrm{L}(S^1)$ of smooth maps of S^1 to S^1 with pointwise multiplication. It is convenient to view its elements as e^{if}, where $f : [0, 2\pi] \to \mathbb{R}$ is a smooth function satisfying $f(2\pi) = f(0) + 2n\pi$ and where $n \in \mathbb{Z}$ is the winding number of e^{if}. In fact, there is an exact sequence

$$0 \to \mathbb{Z} \to \mathrm{L}(\mathbb{R}) \to \mathrm{L}(S^1) \to \mathbb{Z} \to 0$$

where the middle map, given by exponentiation $f \to \exp f = e^{if}$, is preceded by multiplication by 2π and followed by the map which computes

the winding number of a loop. It follows that $\ker \exp = \operatorname{coker} \exp = \mathbb{Z}$ with connected components of $\mathrm{L}(S^1)$ indexed by the winding number.

It is also well known that all central extensions of $\mathrm{L}(S^1)$ by S^1 can be classified (see Pressley and Segal [1986], where a general form of the cocycle depending on the winding number is given). We shall however work only with the connected component of the identity, also denoted $\mathrm{L}(S^1)$, consisting of loops with winding number zero. This has an advantage of making the calculations that follow somewhat simpler.

Formally, the Lie algebra $\mathfrak{L}(S^1)$ of $\mathrm{L}(S^1)$ consists of periodic smooth maps $u : [0, 2\pi] \to \mathbb{R}$ and the exponential map $\exp : \mathfrak{L}(S^1) \to \mathrm{L}(S^1)$ is given by $\exp u = e^{iu}$.

We will now construct a central extension $\widehat{\mathrm{L}}(S^1)$ of $\mathrm{L}(S^1)$. Define the map $B : \mathrm{L}(S^1) \times \mathrm{L}(S^1) \to S^1$ by

$$B(e^{if}, e^{ig}) := \exp\left[i \int_0^{2\pi} f'(x)g(x)\, dx\right],$$

where f' denotes the derivative of f. This map B is easily seen to verify the group two-cocycle condition (8.1.3). On the Cartesian product $\mathrm{L}(S^1) \times S^1$ whose elements are written as $(e^{if}, e^{i\alpha})$, where $\alpha \in \mathbb{R}$ and $f : [0, 2\pi] \to \mathbb{R}$ is a smooth 2π-periodic function, define group multiplication by

$$\left(e^{if}, e^{i\alpha}\right)\left(e^{ig}, e^{i\beta}\right) = \left(e^{i(f+g)}, B(e^{if}, e^{ig})e^{i\alpha}e^{i\beta}\right).$$

The Lie algebra of $\widehat{\mathrm{L}}(S^1)$ has underlying vector space the Cartesian product $\mathfrak{L}(S^1) \times \mathbb{R}$ with commutator

$$[(u, a), (v, b)] = (0, C(u, v)),$$

where

$$C(u, v) = 2 \int_0^{2\pi} u'(x)v(x)\, dx$$

is the Lie algebra two-cocycle and where $u(2\pi) = u(0)$ and $v(2\pi) = v(0)$. One derives the formula for C from that for B using the general equation (6.1.34) by a straightforward calculation.

Remark.   This example is one that can be turned into a rigorous one on the infinite dimensional level in a relatively straightforward way, similar to examples, such as Yang-Mills fields, with gauge group symmetry. One way to do this is to use the context of Fréchet manifolds of smooth maps and invoke the convenient calculus of Kriegl and Michor [1997].

An alternative approach is to use the setting of Hilbert (or Banach) Lie groups, namely to consider the set of periodic maps $f : [0, 2\pi] \to \mathbb{R}$ of Sobolev class H^s, where $s \geq 1$. Of course other function spaces, such as $W^{s,p}$ spaces or Hölder spaces are also possible. Using standard properties of such function spaces (see, for example, Palais [1968]), one checks that this

makes the group $\widehat{\mathrm{L}}(S^1)$ into a Hilbert Lie group (see, for example, Bourbaki [1971] or Neeb [2004b]). Notice, for example, that by the so called *omega lemma*, if $u \in H^s$, then so is e^{iu} and that $C(u,v)$ is well defined and a continuous bilinear (hence smooth) function of (u,v) since u' and v are both in L^2.

The reduction by stages methodology given below can be rigorously carried out in a routine way for either the smooth or the Hilbert-Lie contexts.

The First Reduced Space. Although we do not really require it explicitly, a choice of (weak) metric on the Lie algebra $\mathfrak{L}(S^1)$ is

$$\langle (u,a),(v,b)\rangle = \int_0^{2\pi} u(x)v(x)\,dx + ab.$$

To carry out stage one reduction according to Theorem 8.1.2, we need to compute $\mathcal{B}^1$. By Proposition 8.1.3, $\mathcal{B}^1$ is the right invariant two-form on $\mathrm{L}(S^1)$ that equals C at the identity. To write a formula for $\mathcal{B}^1$, we first compute the derivative of right translation. Let $t \mapsto c(t)$ be a curve in $\mathrm{L}(S^1)$ passing through $c(0) = e^{if}$ in the direction $\dot{c}(0) = u_f \in T_{e^{if}}\mathrm{L}(S^1)$. The derivative of the right translation $R_{e^{ig}}$ in the direction u_f is

$$T_{e^{if}}R_{e^{ig}}u_f = \left.\frac{d}{dt}\right|_{t=0} c(t)e^{ig} = u_f e^{ig},$$

where, by 2π-periodicity, $u_f(2\pi) = u_f(0)$. Thus, by right invariance of $\mathcal{B}^1$, we have

$$\mathcal{B}^1(e^{if})(u_f,v_f) = C(u_f e^{-if}, v_f e^{-if})$$
$$= 2\int_0^{2\pi} \left(u_f e^{-if}\right)' v_f e^{-if}\,dx. \qquad (9.2.1)$$

Thus, by Theorem 8.1.2, reducing $T^*\widehat{\mathrm{L}}(S^1)$ by the central S^1 action at a point $\nu \in \mathbb{R}$ gives

$$\mathbf{J}_{S^1}^{-1}(\nu)/S^1 \simeq \left(T^*\mathrm{L}(S^1), \Omega - \nu\pi^*_{\mathrm{L}(S^1)}\mathcal{B}^1\right),$$

where $\mathcal{B}^1$ is given by equation (9.2.1).

The Second Reduction. We now compute the $\mathcal{B}\mathfrak{g}$-potential $\phi^\nu : \mathrm{L}(S^1) \to \mathfrak{L}(S^1)^*$. We take as dual to $\mathfrak{L}(S^1)$ the same space by using the weakly nondegenerate pairing $\langle\,,\rangle$ given by the L^2 inner product. Formula (8.2.6) gives then

$$\left\langle \phi^\nu(e^{if}), u\right\rangle = -\nu\left(D_2 B(e^{if}, \mathrm{Id})\cdot u + T^*_{\mathrm{Id}}L_{e^{if}}D_1 B(e^{if}, e^{-if})\cdot u\right)$$
$$= -\nu\left(\int_0^{2\pi} f'u\,dx + D_1 B(e^{if}, e^{-if})\cdot ue^{if}\right)$$
$$= -2\nu\int_0^{2\pi} f'u\,dx$$

and so, as the pairing on the left hand side is the L^2-pairing, we get $\phi^\nu(e^{if}) = -2\nu f'$. Since f is of class C^1 and periodic, its derivative f' is also periodic, which shows that $\phi^\nu : L(S^1) \to \mathfrak{L}(S^1)$.

One can also compute ϕ^ν directly from the definition, namely

$$\nu \mathbf{i}_{u_{L(S^1)}} \mathcal{B}^1 = \mathbf{d}\langle \phi^\nu, u \rangle.$$

The infinitesimal generator is

$$u_{L(S^1)}(e^{if}) = T_{id}L_{e^{if}}u = T_{id}R_{e^{if}}u = ue^{if}.$$

Therefore, from (9.2.1), we see that ϕ^ν must satisfy, for all $v_f \in T_{e^{if}} L(S^1)$,

$$\mathbf{d}\langle \phi^\nu, u \rangle (e^{if})(v_f) = 2i\nu \int_0^{2\pi} u'(x)v_f(x)e^{-if(x)}\, dx.$$

Therefore,

$$\langle \phi^\nu, u \rangle (e^{if}) = 2i\nu \int_0^{2\pi} u'(x) \log e^{if(x)}\, dx$$

$$= -2\nu \int_0^{2\pi} u(x)f'(x)\, dx,$$

and thus we get again

$$\phi^\nu(e^{if}) = -2\nu f'.$$

The group 1-cocycle given by (8.2.12) is now just

$$\sigma^\nu(e^{if}) = e^{2i\nu f'}.$$

The affine action (8.2.14) is therefore

$$\mu \cdot e^{if} = \mu + 2\nu f',$$

and hence its isotropy groups are

- $\widetilde{G}^0_\mu = L(S^1)$, if $\nu = 0$;

- $\widetilde{G}^\nu_\mu = \{e^{if} \mid f = \text{const}\} = S^1$, if $\nu \neq 0$.

The affine orbits

$$\widetilde{\mathcal{O}}^\nu_\mu = \left\{ \mu \cdot e^{if} = \mu + 2\nu f' \mid e^{if} \in L(S^1) \right\}$$

are hence

- A point: $\widetilde{\mathcal{O}}^{\nu=0}_\mu = \{\mu\}$;

- An infinite dimensional orbit: $\widetilde{\mathcal{O}}^{\nu\neq0}_\mu \simeq L(S^1)/S^1$.

The (weak) magnetic affine symplectic form on this orbit is given by formula (8.2.3), which in this case becomes

$$\widetilde{\omega}^\nu_\mu(\mu + 2\nu f')\left(\nu C(u, \cdot), \nu C(v, \cdot)\right) = \nu C(u, v) = 2\nu \int_0^{2\pi} u'(x)v(x)\, dx,$$

for any $u, v \in \mathfrak{L}(S^1)$.

Frozen Poisson Structures for the KdV and CH Equations.   In this paragraph, we show how Hamiltonian structures for the KdV equation and the Camassa-Holm equation, which are used as models for shallow water waves, can be obtained in a natural way from the centrally extended loop group $\widehat{L}(S^1)$. More precisely, we consider the dual of the Lie algebra of $\widehat{L}(S^1)$ with a Poisson structure given by the "frozen" Lie-Poisson structure. That is, we fix some point $\mu_0 \in \mathfrak{g}^*$ and define a Poisson structure given by

$$\{f, g\}(\mu) = \langle \mu_0, [\mathbf{d}f(\mu), \mathbf{d}g(\mu)] \rangle \, ;$$

see e.g. [MandS], §10.1, for the proof that this is indeed a Poisson bracket.

Remark.   One normally thinks of the dual of the Bott–Virasoro algebra or just the dual of the Lie algebra of the diffeomorphism group itself as providing the Poisson structure for these equations (see [MandS], Camassa and Holm [1993], Misiołek [1998], and Khesin and Misiołek [2003] for discussions).[2]

Consider the following (weak) metrics on $\mathfrak{L}(S^1) \times \mathbb{R}$

$$\langle (f, a), (g, b) \rangle_{L^2} = \int_{S^1} f(x) g(x) \, dx + ab,$$

and

$$\langle (f, a), (g, b) \rangle_{H^1} = \int_{S^1} f'(x) g'(x) \, dx + \int_{S^1} f(x) g(x) \, dx + ab.$$

Sometimes in what follows, we will use the notation $\partial_x f = f'$ for the derivative, for notational clarity.

Using the two preceding metrics, we identify elements in the dual with elements in the algebra in each case relative to the given metric. We proceed to compute the coadjoint actions relative to the two preceding metrics.

$$\left\langle (\mathrm{ad}^{L^2})^*_{(f,a)}(h, c), (g, b) \right\rangle_{L^2} = \langle (h, c), [(f, a), (g, b)] \rangle_{L^2}$$

$$= \int_{S^1} cf'(x) g(x) \, dx = \langle (cf', 0), (g, b) \rangle_{L^2} \, ,$$

so that

$$(\mathrm{ad}^{L^2})^*_{(h,c)}(f, a) = (cf', 0). \tag{9.2.2}$$

Similarly, we compute

$$\left\langle (\mathrm{ad}^{H^1})^*_{(f,a)}(h, c), (g, b) \right\rangle_{H^1} = \langle (h, c), [(f, a), (g, b)] \rangle_{H^1}$$

$$= \int_{S^1} cf'(x) g(x) \, dx$$

---

[2]There is a large PDE literature about these equations that we do not discuss here.

Now insert the identity written as $(1 - \partial_x^2)(1 - \partial_x^2)^{-1}$ so that the preceding expression becomes

$$\left\langle (\mathrm{ad}^{H^1})^*_{(f,a)}(h,c),(g,b) \right\rangle_{H^1} = \int_{S^1} g(x)(1 - \partial_x^2)(1 - \partial_x^2)^{-1} cf'(x)\, dx$$
$$= \left\langle \left(c(1 - \partial_x^2)^{-1} f', 0 \right), (g,b) \right\rangle_{H^1}.$$

Therefore,

$$(\mathrm{ad}^{H^1})^*_{(f,a)}(f,a) = c\left((1 - \partial_x^2)^{-1} f', 0 \right). \tag{9.2.3}$$

We now proceed to compute "frozen" brackets in each case. Recall that in general, given $(w_0, a_0) \in \mathfrak{L}(S^1) \times \mathbb{R}$, the "frozen" bracket is given by

$$\{F, H\}(w,a) = \left\langle (w_0, a_0), \left[\frac{\delta F}{\delta(w,a)}, \frac{\delta H}{\delta(w,a)} \right] \right\rangle$$
$$= \left\langle -\mathrm{ad}^*_{\frac{\delta H}{\delta(w,a)}}(w_0, a_0), \frac{\delta F}{\delta(w,a)} \right\rangle,$$

where $\langle \cdot, \cdot \rangle$ denotes the L^2 or the H^1 metric. Furthermore, recall that the corresponding equations of motion are given by

$$\frac{d}{dt}(w,a) = -\mathrm{ad}^*_{\left(\frac{\delta H}{\delta w}, \frac{\delta H}{\delta a} \right)}(w_0, a_0),$$

where ad^* denotes the L^2 or the H^1 coadjoint operator. Thus, using (9.2.2) and (9.2.3), we get the following equations for the L^2 and the H^1 metric respectively

$$\frac{d}{dt}(w,a) = \left(-a_0 \partial_x \left(\frac{\delta H}{\delta w} \right), 0 \right),$$

and

$$\frac{d}{dt}(w,a) = \left(-a_0 (1 - \partial_x^2)^{-1} \partial_x \left(\frac{\delta H}{\delta w} \right), 0 \right).$$

Consider the following Hamiltonians

$$H_1(w,a) = \int_{S^1} \left(w^3 + (\partial_x w)^2 \right) dx. \tag{9.2.4}$$

and

$$H_2(w,a) = \int_{S^1} \left(w^3 + w\,(\partial_x w)^2 \right) dx. \tag{9.2.5}$$

The L^2 variational derivative of H_1 is

$$\left(\frac{\delta H_1}{\delta w} \right)_{L^2} = 3w^2 - 2\partial_x^2 w.$$

Similarly, the H^1 variational derivative of H_1 is

$$\left(\frac{\delta H_1}{\delta w} \right)_{H^1} = (1 - \partial_x^2)^{-1} \left(3w^2 - 2\partial_x^2 w \right),$$

and of H_2 is

$$\left(\frac{\delta H_2}{\delta w}\right)_{H^1} = (1 - \partial_x^2)^{-1}\left(3w^2 - (\partial_x w)^2 - 2w\partial_x^2 w\right).$$

We can now write down the corresponding equations. For the L^2 (weak) metric and Hamiltonian H_1 we get the KdV equation:

$$\partial_t w + 6a_0 w\partial_x w - 2a_0\partial_x^3 w = 0.$$

For the H^1 metric and the Hamiltonian H_1 we get

$$\partial_t w - \partial_t\partial_x^2 w + 6a_0 w\partial_x w - 2a_0\partial_x^3 w = 0.$$

Finally, for the H^1 metric and Hamiltonian H_2 we get the Camassa-Holm equation (see Camassa and Holm [1993]),

$$\partial_t w - \partial_t\partial_x^2 w + 6a_0 w\partial_x w - 4a_0\partial_x w\partial_x^2 w - 2a_0 w\partial_x^3 w = 0.$$

9.3   The Oscillator Group

The goal of this section is to describe the coadjoint orbits of the oscillator group, which is a one-dimensional central extension of the Euclidean group SE(2), using the reduction by stages technique described in Theorem 8.2.1.

The Oscillator Group.   We shall denote by

$$\mathcal{R}_\theta := \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

the counter clockwise rotation in the plane by the angle θ. Thus, the group SO(2) := $\{\mathcal{R}_\theta \mid \theta \in \mathbb{R}\}$ is the circle S^1 oriented in the positive trigonometric direction.

A general vector at the identity to SO(2) equals

$$\left.\frac{d}{dt}\right|_{t=0} \mathcal{R}_{t\xi} = -\xi\mathbb{J},$$

where $\xi \in \mathbb{R}$ and

$$\mathbb{J} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}.$$

Note that $\mathbb{J}^T = \mathbb{J}^{-1} = -\mathbb{J}$. The Lie algebra of SO(2) is isomorphic to $\mathbb{R}$ via $\xi \longleftrightarrow -\xi\mathbb{J}$ and the exponential map $\exp : \mathbb{R} \to$ SO(2) is given by $\exp\xi = \mathcal{R}_\xi$.

With these notational conventions we recall that the Euclidean group SE(2) is the semidirect product of SO(2) with $\mathbb{R}^2$ and hence the group operation is given by

$$(\mathcal{R}_\theta, \mathbf{a})(\mathcal{R}_\psi, \mathbf{b}) = (\mathcal{R}_{\theta+\psi}, \mathbf{a} + \mathcal{R}_\theta \mathbf{b}),$$

where $\theta, \psi \in \mathbb{R}$ and $\mathbf{a}, \mathbf{b} \in \mathbb{R}^2$. The neutral element is $(\mathcal{R}_0, \mathbf{0})$ and $(\mathcal{R}_\theta, \mathbf{a})^{-1} = (\mathcal{R}_{-\theta}, -\mathcal{R}_{-\theta}\mathbf{a})$.

Let ω be the standard symplectic form on $\mathbb{R}^2$, that is,

$$\omega(\mathbf{u}, \mathbf{v}) = \mathbf{u} \cdot \mathbb{J}\mathbf{v} \qquad (9.3.1)$$

for $\mathbf{u}, \mathbf{v} \in \mathbb{R}^2$. A real-valued group two-cocycle on SE(2) is given by

$$B\left((\mathcal{R}_\theta, \mathbf{a}), (\mathcal{R}_\psi, \mathbf{b})\right) := \omega(\mathbf{a}, \mathcal{R}_\theta \mathbf{b}) = \mathbf{a} \cdot \mathbb{J}\mathcal{R}_\theta \mathbf{b}. \qquad (9.3.2)$$

An easy calculation, using the fact that the rotations act symplectically on $\mathbb{R}^2$, shows that B verifies the cocycle identity (6.2.1). It is also clear that this cocycle is normalized, that is, $B((\mathcal{R}_0, \mathbf{0}), (\mathcal{R}_\theta, \mathbf{a})) = B((\mathcal{R}_\theta, \mathbf{a}), (\mathcal{R}_0, \mathbf{0})) = 0$, for all $(\mathcal{R}_\theta, \mathbf{a}) \in \text{SE}(2)$. In addition, note that $B((\mathcal{R}_\theta, \mathbf{a}), (\mathcal{R}_\theta, \mathbf{a})^{-1}) = B((\mathcal{R}_\theta, \mathbf{a})^{-1}, (\mathcal{R}_\theta, \mathbf{a})) = 0$. Thus, we can form the central extension $\widehat{\text{SE}(2)}$ of SE(2) by $\mathbb{R}$, called the **oscillator group**, which is the Lie group whose underlying manifold is $\text{SE}(2) \times \mathbb{R}$, the group operation is given by

$$(\mathcal{R}_\theta, \mathbf{a}, s)(\mathcal{R}_\psi, \mathbf{b}, t) = (\mathcal{R}_{\theta+\psi}, \mathbf{a} + \mathcal{R}_\theta \mathbf{b}, s + t + \omega(\mathbf{a}, \mathcal{R}_\theta \mathbf{b})),$$

the neutral element is $(\mathcal{R}_0, \mathbf{0}, 0)$, and $(\mathcal{R}_\theta, \mathbf{a}, s)^{-1} = (\mathcal{R}_{-\theta}, -\mathcal{R}_{-\theta}\mathbf{a}, -s)$ (see (6.2.12)).

The Oscillator Lie Algebra. The Lie algebra $\mathfrak{se}(2)$ of SE(2) has underlying vector space $\mathbb{R}^3$ and Lie bracket

$$[(\xi, \mathbf{u}), (\eta, \mathbf{v})] = (0, \eta u_2 - \xi v_2, \xi v_1 - \eta u_1) = (0, -\xi \mathbb{J}\mathbf{v} + \eta \mathbb{J}\mathbf{u}) \qquad (9.3.3)$$

for all $\xi, \eta \in \mathbb{R}$ and $\mathbf{u} = (u_1, u_2), \mathbf{v} = (v_1, v_2) \in \mathbb{R}^2$.

From the definition (6.1.34) differentiating twice the group two-cocycle B, we get the associated Lie algebra two-cocycle

$$C\left((\xi, \mathbf{u}), (\eta, \mathbf{v})\right) = 2\omega(\mathbf{u}, \mathbf{v}) = 2\mathbf{u} \cdot \mathbb{J}\mathbf{v} \qquad (9.3.4)$$

for all $(\xi, \mathbf{u}), (\eta, \mathbf{v}) \in \mathfrak{se}(2)$. Therefore, the **oscillator Lie algebra** $\widehat{\mathfrak{se}(2)}$, which is the Lie algebra of SE(2) and is a one-dimensional central extension of the Euclidean Lie algebra $\mathfrak{se}(2)$, has underlying vector space $\mathbb{R}^4$ and Lie bracket

$$\begin{aligned}
[(\xi, \mathbf{u}, a), (\eta, \mathbf{v}, b)] &= ([(\xi, \mathbf{u}), (\eta, \mathbf{v})], C\left((\xi, \mathbf{u}), (\eta, \mathbf{v})\right) \\
&= (0, -\xi \mathbb{J}\mathbf{v} + \eta \mathbb{J}\mathbf{u}, 2\mathbf{u} \cdot \mathbb{J}\mathbf{v}) \qquad (9.3.5)
\end{aligned}$$

for any $(\xi, \mathbf{u}, a), (\eta, \mathbf{v}, b) \in \widehat{\mathfrak{se}(2)}$.

Adjoint and Coadjoint action of SE(2)**.** For several computations below we shall need the formulas for the adjoint and coadjoint actions of the Euclidean group SE(2) on its Lie algebra and its dual. These are particular cases of the formulas (4.2.1) and (4.2.2) which in this simple case become

$$\mathrm{Ad}_{(\mathcal{R}_\theta, \mathbf{a})}(\xi, \mathbf{u}) = (\xi, \mathcal{R}_\theta \mathbf{u} + \xi \mathbb{J} \mathbf{a}) \tag{9.3.6}$$

and

$$\mathrm{Ad}^*_{(\mathcal{R}_\theta, \mathbf{a})^{-1}}(x, \boldsymbol{\alpha}) = (x + \mathbb{J} \mathcal{R}_\theta \boldsymbol{\alpha} \cdot \mathbf{a}, \mathcal{R}_\theta \boldsymbol{\alpha}) \tag{9.3.7}$$

for any $(\mathcal{R}_\theta, \mathbf{a}) \in \mathrm{SE}(2)$, $(\xi, \mathbf{u}) \in \mathfrak{se}(2) \simeq \mathbb{R} \times \mathbb{R}^2$, and $(x, \boldsymbol{\alpha}) \in \mathfrak{se}(2)^* \simeq \mathbb{R} \times \mathbb{R}^2$.

First Reduced Space. The first reduced space at $\nu \in \mathbb{R}$ is, by Theorem 8.1.2,

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

where $\pi_{\mathrm{SE}(2)} : T^* \mathrm{SE}(2) \to \mathrm{SE}(2)$ is the cotangent bundle projection and Ω is the canonical symplectic form on $T^* \mathrm{SE}(2)$. By Proposition 8.1.3, $\mathcal{B}^1$ is the right invariant two-form on SE(2) that equals the Lie algebra cocycle C at the identity.

Let $(\xi_{\mathcal{R}_\theta}, \mathbf{u}) \in T_{(\mathcal{R}_\theta, \mathbf{a})} \mathrm{SE}(2) = T_{\mathcal{R}_\theta} \mathrm{SO}(2) \times \mathbb{R}^2$. Then

$$\xi_{\mathcal{R}_\theta} = \left. \frac{d}{dt} \right|_{t=0} \mathcal{R}_{\theta + t\xi} = -\xi \mathcal{R}_\theta \mathbb{J}$$

for some $\xi \in \mathbb{R}$. Therefore, the derivative of right translation is given by

$$\begin{aligned}
T_{(\mathcal{R}_\theta, \mathbf{a})} R_{(\mathcal{R}_\theta, \mathbf{a})^{-1}}(\xi_{\mathcal{R}_\theta}, \mathbf{u}) &= \left. \frac{d}{dt} \right|_{t=0} (\mathcal{R}_{\theta + t\xi}, t\mathbf{u})(\mathcal{R}_{-\theta}, -\mathcal{R}_{-\theta} \mathbf{a}) \\
&= (-\xi \mathbb{J}, \mathbf{u} + \xi \mathbb{J} \mathbf{a}),
\end{aligned}$$

and hence

$$\begin{aligned}
\mathcal{B}^1 &(\mathcal{R}_\theta, \mathbf{a}) \left((\xi_{\mathcal{R}_\theta}, \mathbf{u}), (\eta_{\mathcal{R}_\theta}, \mathbf{v}) \right) \\
&= C \left(T_{(\mathcal{R}_\theta, \mathbf{a})} R_{(\mathcal{R}_\theta, \mathbf{a})^{-1}}(\xi_{\mathcal{R}_\theta}, \mathbf{u}), T_{(\mathcal{R}_\theta, \mathbf{a})} R_{(\mathcal{R}_\theta, \mathbf{a})^{-1}}(\eta_{\mathcal{R}_\theta}, \mathbf{v}) \right) \\
&= C \left((-\xi \mathbb{J}, \mathbf{u} + \xi \mathbb{J} \mathbf{a}), (-\eta \mathbb{J}, \mathbf{v} + \eta \mathbb{J} \mathbf{a}) \right) \\
&= 2 (\mathbf{u} + \xi \mathbb{J} \mathbf{a}) \cdot \mathbb{J} (\mathbf{v} + \eta \mathbb{J} \mathbf{a}) \\
&= 2 (\mathbf{u} + \xi \mathbb{J} \mathbf{a}) \cdot (\mathbb{J} \mathbf{v} - \eta \mathbf{a}) \\
&= 2 \mathbf{u} \cdot \mathbb{J} \mathbf{v} + 2 \mathbf{a} \cdot (\xi \mathbf{v} - \eta \mathbf{u}), \tag{9.3.8}
\end{aligned}$$

where $\xi_{\mathcal{R}_\theta} = -\xi \mathcal{R}_\theta \mathbb{J}$ and $\eta_{\mathcal{R}_\theta} = -\eta \mathcal{R}_\theta \mathbb{J}$, for $\xi, \eta \in \mathbb{R}$.

The $\mathcal{B}\mathfrak{g}$-Potential.   We identify the dual $\mathfrak{se}(2)^*$ with $\mathbb{R}^3$ using the Euclidean inner product. The $\mathcal{B}\mathfrak{g}$-potential $\phi^\nu : \mathrm{SE}(2) \to \mathfrak{se}(2)^* \simeq \mathbb{R}^3$ is given by formula (8.2.6). Thus we get for any $(\mathcal{R}_\theta, \mathbf{a}) \in \mathrm{SE}(2)$ and $(\xi, \mathbf{u}) \in \mathfrak{se}(2)$,

$$
\begin{aligned}
\phi^\nu & (\mathcal{R}_\theta, \mathbf{a}) \cdot (\xi, \mathbf{u}) \\
&= -\nu D_2 B((\mathcal{R}_\theta, \mathbf{a}), (\mathcal{R}_0, \mathbf{0})) \cdot (\xi, \mathbf{u}) \\
&\quad - \nu D_1 B((\mathcal{R}_\theta, \mathbf{a}), (\mathcal{R}_{-\theta}, -\mathcal{R}_{-\theta}\mathbf{a})) \left(T_{(\mathcal{R}_0, \mathbf{0})} L_{(\mathcal{R}_\theta, \mathbf{a})}(\xi, \mathbf{u})\right) \\
&= -\nu \left.\frac{d}{dt}\right|_{t=0} B\left((\mathcal{R}_\theta, \mathbf{a}), (\mathcal{R}_{t\xi}, t\mathbf{u})\right) \\
&\quad - \nu \left.\frac{d}{dt}\right|_{t=0} B\left((\mathcal{R}_\theta, \mathbf{a})(\mathcal{R}_{t\xi}, t\mathbf{u}), (\mathcal{R}_{-\theta}, -\mathcal{R}_{-\theta}\mathbf{a})\right) \\
&= -\nu \left.\frac{d}{dt}\right|_{t=0} \omega(\mathbf{a}, \mathcal{R}_\theta(t\mathbf{u})) - \nu \left.\frac{d}{dt}\right|_{t=0} \omega(\mathcal{R}_\theta(t\mathbf{u}) + \mathbf{a}, -\mathcal{R}_{t\xi}\mathbf{a}) \\
&= -\nu \omega(\mathcal{R}_{-\theta}\mathbf{a}, , \mathbf{u}) - \nu \omega(\mathcal{R}_\theta \mathbf{u}, -\mathbf{a}) - \nu \omega(\mathbf{a}, \xi \mathbb{J} \mathbf{a}) \\
&= -2\nu \omega(\mathcal{R}_{-\theta}\mathbf{a}, \mathbf{u}) + \nu \xi \|\mathbf{a}\|^2 \\
&= 2\nu \mathbb{J} \mathcal{R}_{-\theta}\mathbf{a} \cdot \mathbf{u} + \nu \xi \|\mathbf{a}\|^2 \\
&= \nu(\|\mathbf{a}\|^2, 2\mathcal{R}_{-\pi/2-\theta}\mathbf{a}) \cdot (\xi, \mathbf{u}),
\end{aligned}
$$

since $\mathbb{J} = \mathcal{R}_{-\pi/2}$. Thus we have

$$
\phi^\nu(\mathcal{R}_\theta, \mathbf{a}) = \nu \left(\|\mathbf{a}\|^2, 2\mathcal{R}_{-\pi/2-\theta}\mathbf{a}\right). \tag{9.3.9}
$$

By (8.2.13), the corresponding nonequivariance group one-cocycle σ^ν : $\mathrm{SE}(2) \to \mathfrak{se}(2)^* \simeq \mathbb{R}^3$ of $\mathbf{J}^\nu : T^* \mathrm{SE}(2) \to \mathfrak{se}(2)^* \simeq \mathbb{R}^3$ equals

$$
\sigma^\nu(\mathcal{R}_\theta, \mathbf{a}) = -2\nu \left(\frac{1}{2}\|\mathbf{a}\|^2, \mathcal{R}_{-\pi/2-\theta}\mathbf{a}\right). \tag{9.3.10}
$$

Recall that σ^ν is a *right* group one-cocycle, that is, it satisfies the *right* cocycle identity (6.2.3).

A Second Interpretation of the One-Cocycle σ^ν.   We shall prove below that the cocycle σ^ν also arises naturally as the group one-cocycle obtained from the *nonequivariant momentum map* of the *right* action of $\mathrm{SE}(2)$ on $\mathbb{R}^2$ with symplectic form $-2\nu\omega$ where ω is the standard symplectic structure on $\mathbb{R}^2$ given by (9.3.1).

The right action of $\mathrm{SE}(2)$ on $\mathbb{R}^2$ is given by

$$
\mathbf{x} \cdot (\mathcal{R}_\theta, \mathbf{a}) := \mathcal{R}_{-\theta}(\mathbf{x} - \mathbf{a}) \tag{9.3.11}
$$

for any $\mathbf{x} \in \mathbb{R}^2$ and $(\mathcal{R}_\theta, \mathbf{a}) \in \mathrm{SE}(2)$. Thus, the infinitesimal generator of $(\xi, \mathbf{u}) \in \mathfrak{se}(2)$ has the expression

$$
(\xi, \mathbf{u})_{\mathbb{R}^2}(\mathbf{x}) = \left.\frac{d}{dt}\right|_{t=0} \mathbf{x} \cdot (\mathcal{R}_{t\xi}, t\mathbf{u}) = \left.\frac{d}{dt}\right|_{t=0} \mathcal{R}_{-t\xi}(\mathbf{x} - t\mathbf{u}) = \xi \mathbb{J}\mathbf{x} - \mathbf{u}.
$$

Let $\mathbf{K} : \mathbb{R}^2 \to \mathfrak{se}(2)^* \simeq \mathbb{R}^3$ be the momentum map of this action relative to the symplectic form $-2\nu\omega$. In order to compute it, we use the definition of the momentum map, that is, we require that

$$\mathbf{dK}^{(\xi,\mathbf{u})}(\mathbf{x}) \cdot \mathbf{y} = -2\nu\omega \left((\xi, \mathbf{u})_{\mathbb{R}^2}(\mathbf{x}), \mathbf{y} \right)$$

for any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^2$ and any $(\xi, \mathbf{u}) \in \mathfrak{se}(2)$, where $\mathbf{K}^{(\xi,\mathbf{u})}(\mathbf{x}) := \mathbf{K}(\mathbf{x}) \cdot (\xi, \mathbf{u})$. Since

$$\omega \left((\xi, \mathbf{u})_{\mathbb{R}^2}(\mathbf{x}), \mathbf{y} \right) = \omega \left(\xi\mathbb{J}\mathbf{x} - \mathbf{u}, \mathbf{y} \right) = (\xi\mathbb{J}\mathbf{x} - \mathbf{u}) \cdot \mathbb{J}\mathbf{y} = (\xi\mathbf{x} + \mathbb{J}\mathbf{u}) \cdot \mathbf{y},$$

we get $\mathbf{dK}^{(\xi,\mathbf{u})}(\mathbf{x}) = -2\nu(\xi\mathbf{x} + \mathbb{J}\mathbf{u})$ and hence

$$\mathbf{K}^{(\xi,\mathbf{u})}(\mathbf{x}) = -2\nu \left(\frac{1}{2}\xi\|\mathbf{x}\|^2 + \mathbb{J}\mathbf{u} \cdot \mathbf{x} \right) = -2\nu \left(\frac{1}{2}\|\mathbf{x}\|^2, -\mathbb{J}\mathbf{x} \right) \cdot (\xi, \mathbf{u}),$$

which yields

$$\mathbf{K}(\mathbf{x}) = -2\nu \left(\frac{1}{2}\|\mathbf{x}\|^2, -\mathbb{J}\mathbf{x} \right) \tag{9.3.12}$$

for any $\mathbf{x} \in \mathbb{R}^2$.

Using (9.3.11), (9.3.12), and the fact that the right hand side of the definition of the nonequivariance group one-cocycle $\sigma^{\mathbf{K}} : \mathrm{SE}(2) \to \mathfrak{se}(2)^* \simeq \mathbb{R}^3$ of $\mathbf{K}$ does not depend on the point of evaluation in $\mathbb{R}^2$, we get

$$\sigma^{\mathbf{K}}(\mathcal{R}_\theta, \mathbf{a}) = \mathbf{K}(\mathbf{0} \cdot (\mathcal{R}_\theta, \mathbf{a})) - \mathrm{Ad}^*_{(\mathcal{R}_\theta, \mathbf{a})} \mathbf{K}(\mathbf{0}) = \mathbf{K}(-\mathcal{R}_{-\theta}\mathbf{a})$$

$$= -2\nu \left(\frac{1}{2}\|\mathbf{a}\|^2, \mathbb{J}\mathcal{R}_{-\theta}\mathbf{a} \right) = -2\nu \left(\frac{1}{2}\|\mathbf{a}\|^2, \mathcal{R}_{-\pi/2-\theta}\mathbf{a} \right)$$

which is identical to σ^ν in (9.3.10).

The Second Reduced Space.   We shall apply now Theorem 8.2.1 together with the formulas (9.3.7) and (9.3.10) to characterize the coadjoint orbits of the oscillator group $\widehat{\mathrm{SE}(2)}$. Thus, given $(x, \boldsymbol{\alpha}, \nu) \in \widehat{\mathfrak{se}(2)}^* \simeq \mathbb{R} \times \mathbb{R}^2 \times \mathbb{R} = \mathbb{R}^4$, Theorem 8.2.1 states that the coadjoint orbit $\mathcal{O}_{(x,\boldsymbol{\alpha},\nu)} \subset \widehat{\mathfrak{se}(2)}^*$ of the oscillator group with its usual (plus) orbit symplectic form is symplectically diffeomorphic to the affine orbit

$$\widetilde{\mathcal{O}}^\nu_{(x,\boldsymbol{\alpha})} = \left\{ \mathrm{Ad}^*_{(\mathcal{R}_\theta,\mathbf{a})}(x, \boldsymbol{\alpha}) + \sigma^\nu(\mathcal{R}_\theta, \mathbf{a}) \;\middle|\; (\mathcal{R}_\theta, \mathbf{a}) \in \mathrm{SE}(2) \right\}$$

$$= \left\{ \left(x - \mathcal{R}_{-\pi/2}\boldsymbol{\alpha} \cdot \mathbf{a} - \nu\|\mathbf{a}\|^2, \mathcal{R}_{-\theta}\boldsymbol{\alpha} - 2\nu\mathcal{R}_{-\pi/2-\theta}\mathbf{a} \right) \;\middle|\; (\mathcal{R}_\theta, \mathbf{a}) \in \mathrm{SE}(2) \right\}$$

in $\mathfrak{se}(2)^* \simeq \mathbb{R} \times \mathbb{R}^2$ endowed with the symplectic form given by (8.2.3). Thus we get the following result.

9.3.1 Theorem.   *The Lie-Poisson bracket of $F, H \in C^\infty \left(\widehat{\mathfrak{se}(2)}^* \right)$ is given by*

$$\{F, H\}(x, \boldsymbol{\alpha}, \nu) = -\frac{\partial F}{\partial x}\boldsymbol{\alpha} \cdot \mathbb{J}\nabla_{\boldsymbol{\alpha}} H + \frac{\partial H}{\partial x}\boldsymbol{\alpha} \cdot \mathbb{J}\nabla_{\boldsymbol{\alpha}} F$$

$$+ 2\nu\nabla_{\boldsymbol{\alpha}} F \cdot \mathbb{J}\nabla_{\boldsymbol{\alpha}} H, \tag{9.3.13}$$

where $(x, \boldsymbol{\alpha}, \nu) \in \widehat{\mathfrak{se}(2)}^*$ and $\nabla_{\boldsymbol{\alpha}} F, \nabla_{\boldsymbol{\alpha}} H$ denote the partial gradients of F and H relative to $\boldsymbol{\alpha} \in \mathbb{R}^2$. The Casimir functions of $\widehat{\mathfrak{se}(2)}^*$ are all of the form

$$(x, \boldsymbol{\alpha}, \nu) \mapsto \Phi\left(\nu, x + \frac{1}{4\nu}\|\boldsymbol{\alpha}\|^2\right), \qquad (9.3.14)$$

where Φ is any smooth real valued function of two variables.

For $\nu \in \mathbb{R}$, $(x, \boldsymbol{\alpha}) \in \mathfrak{se}(2)^*$, the coadjoint orbit $\mathcal{O}_{(x,\boldsymbol{\alpha},\nu)} \subset \widehat{\mathfrak{se}(2)}^*$ of the oscillator group with its (plus) orbit symplectic form is symplectically diffeomorphic to the following affine orbit in $\mathfrak{se}(2)^*$:

(i) *If $\nu \neq 0$ then the affine isotropy subgroup equals*

$$\widehat{\mathrm{SE}(2)}^{\nu}_{(x,\boldsymbol{\alpha})} = \left\{ \left(\mathcal{R}_\theta, \frac{1}{2\nu}\mathbb{J}(\mathcal{R}_\theta \boldsymbol{\alpha} - \boldsymbol{\alpha})\right) \;\middle|\; \mathcal{R}_\theta \in \mathrm{SO}(2) \right\} \simeq \mathrm{SO}(2),$$

and hence

$$\mathcal{O}_{(x,\boldsymbol{\alpha},\nu)} = \widetilde{\mathcal{O}}^{\nu}_{(x,\boldsymbol{\alpha})} \simeq \mathrm{SE}(2)/\mathrm{SO}(2).$$

In addition,

$$\mathcal{O}_{(x,\boldsymbol{\alpha},\nu)} = \widetilde{\mathcal{O}}^{\nu}_{(x,\boldsymbol{\alpha})}$$
$$= \left\{ (x', \boldsymbol{\alpha}', \nu) \in \mathbb{R}^4 \;\middle|\; x' + \frac{1}{4\nu}\|\boldsymbol{\alpha}'\|^2 = x + \frac{1}{4\nu}\|\boldsymbol{\alpha}\|^2 \right\} \qquad (9.3.15)$$

is a circular paraboloid in $(x, \boldsymbol{\alpha})$-space whose axis of symmetry is Ox.

The symplectic form on the affine orbit $\widetilde{\mathcal{O}}^{\nu}_{(x,\boldsymbol{\alpha})}$ given in formula (8.2.3) becomes in this case

$$\widetilde{\omega}^{\nu}_{(x,\boldsymbol{\alpha})}(x', \boldsymbol{\alpha}')\left(\mathrm{ad}^*_{(\xi,\mathbf{u})}(x', \boldsymbol{\alpha}') - (0, 2\nu\mathbb{J}\mathbf{u}), \mathrm{ad}^*_{(\eta,\mathbf{v})}(x', \boldsymbol{\alpha}') - (0, 2\nu\mathbb{J}\mathbf{v})\right)$$
$$= \boldsymbol{\alpha}' \cdot (-\xi\mathbb{J}\mathbf{v} + \eta\mathbb{J}\mathbf{u}) + 2\nu\mathbf{u} \cdot \mathbb{J}\mathbf{v} \qquad (9.3.16)$$

where $(x', \boldsymbol{\alpha}') = \left(x - \mathcal{R}_{-\pi/2}\boldsymbol{\alpha} \cdot \mathbf{a} - \nu\|\mathbf{a}\|^2, \mathcal{R}_{-\theta}\boldsymbol{\alpha} - 2\nu\mathcal{R}_{-\pi/2-\theta}\mathbf{a}\right) \in \widetilde{\mathcal{O}}^{\nu}_{(x,\boldsymbol{\alpha})} \subset \mathbb{R}^3$. The area form on the paraboloid (9.3.15) evaluated at $(x', \boldsymbol{\alpha}')$ equals $\sqrt{4\nu^2 + \|\boldsymbol{\alpha}'\|^2}\, \widetilde{\omega}^{\nu}_{(x,\boldsymbol{\alpha})}(x', \boldsymbol{\alpha}')$.

(ii) *If $\nu = 0$ then the affine isotropy subgroup is*

$$\widehat{\mathrm{SE}(2)}^{0}_{(x,\boldsymbol{\alpha})} = \left\{ \begin{array}{ll} \mathrm{SE}(2), & \text{if} \quad \boldsymbol{\alpha} = \mathbf{0} \\ \{1\} \times \mathbb{R}\boldsymbol{\alpha}, & \text{if} \quad \boldsymbol{\alpha} \neq \mathbf{0} \end{array} \right. .$$

The coadjoint orbits are $\mathcal{O}_{(0,x,\boldsymbol{\alpha})} = \{(0, x', \boldsymbol{\alpha}') \mid \|\boldsymbol{\alpha}'\| = \|\boldsymbol{\alpha}\|\}$ which are the concentric cylinders about the second axis of radius $\|\boldsymbol{\alpha}\|$ and, if $\boldsymbol{\alpha} = \mathbf{0}$, points on this axis. The symplectic form on the cylinder passing through the point $(0, x, \boldsymbol{\alpha})$ is the standard area form on the cylinder divided by the constant $\|\boldsymbol{\alpha}\|$.

Proof. The explicit expression (9.3.13) of the Lie-Poisson bracket is a direct consequence of formula (6.2.22). By Theorem 6.2.2 we know that the hyperplanes of constant ν are Poisson submanifolds of $\widehat{\mathfrak{se}(2)}^*$. In addition, it is straightforward to verify that the function $(x, \boldsymbol{\alpha}) \mapsto x + \|\boldsymbol{\alpha}\|^2/4\nu$ is a Casimir function of the bracket (9.3.13). Thus, any function of the form (9.3.14) is a Casimir function. We will show later that these are the only Casimir functions, when studying the coadjoint orbits.

By (7.1.5), (9.3.7), and (9.3.10) it follows that the right action of SE(2) on $\mathfrak{se}(2)^*$ is given by

$$(x, \boldsymbol{\alpha}) \cdot (\mathcal{R}_\theta, \mathbf{a}) = \left(x - \mathbb{J}\boldsymbol{\alpha} \cdot \mathbf{a} - \nu\|\mathbf{a}\|^2, \mathcal{R}_{-\theta}(\boldsymbol{\alpha} - 2\nu\mathbb{J}\mathbf{a})\right). \qquad (9.3.17)$$

Thus $(\mathcal{R}_\theta, \mathbf{a}) \in \widetilde{\mathrm{SE}(2)}^{\nu}_{(x, \boldsymbol{\alpha})}$ if and only if

$$\mathbb{J}\boldsymbol{\alpha} \cdot \mathbf{a} + \nu\|\mathbf{a}\|^2 = 0 \quad \text{and} \quad \boldsymbol{\alpha} - 2\nu\mathbb{J}\mathbf{a} = \mathcal{R}_\theta\boldsymbol{\alpha}.$$

(i) Let $\nu \neq 0$. Solving the second equation for $\mathbf{a}$ we obtain the θ-dependent solution $\mathbf{a} = -\frac{1}{2\nu}\mathbb{J}(\boldsymbol{\alpha} - \mathcal{R}_\theta\boldsymbol{\alpha})$. This value of $\mathbf{a}$ automatically verifies the first equation, which proves the formula for the affine isotropy group in the statement of the theorem.

Using (9.3.17) and $\mathbb{J}^T = \mathbb{J}^{-1} = -\mathbb{J}$ we see that, denoting the function $(x, \boldsymbol{\alpha}) \mapsto x + \frac{1}{4\nu}\|\boldsymbol{\alpha}\|^2$ by $\psi : \mathfrak{se}(2)^* \to \mathbb{R}$,

$$\psi\left((x, \alpha) \cdot (\mathcal{R}_\theta, \mathbf{a})\right) = x - \mathbb{J}\boldsymbol{\alpha} \cdot \mathbf{a} - \nu\|\mathbf{a}\|^2 + \frac{1}{4\nu}\|\mathcal{R}_{-\theta}(\boldsymbol{\alpha} - 2\nu\mathbb{J}\mathbf{a})\|^2$$

$$= x - \mathbb{J}\boldsymbol{\alpha} \cdot \mathbf{a} - \nu\|\mathbf{a}\|^2 + \frac{1}{4\nu}\|\boldsymbol{\alpha}\|^2 - \boldsymbol{\alpha} \cdot \mathbb{J}\mathbf{a} + \nu\|\mathbb{J}\mathbf{a}\|^2$$

$$= x + \frac{1}{4\nu}\|\boldsymbol{\alpha}\|^2,$$

which shows that the generic orbit $\widetilde{\mathcal{O}}^{\nu}_{(x, \boldsymbol{\alpha})} = \mathcal{O}_{(x, \boldsymbol{\alpha}, \nu)}$ is the level set of the Casimir functions (9.3.14). This proves that the functions given by (9.3.14) are all the Casimir functions and that this orbit is given by (9.3.15).

The formula for the symplectic form (9.3.16) follows by applying (9.3.3) and (9.3.4) to the general formula (8.2.3).

Finally, we express the symplectic form on the coadjoint orbit $\mathcal{O}_{(x, \boldsymbol{\alpha}, \nu)} = \widetilde{\mathcal{O}}^{\nu}_{(x, \boldsymbol{\alpha})}$ in terms of the standard area form on the circular paraboloid (9.3.15). In the computation that follows we shall need the following formulas obtained by identifying $\mathfrak{se}(2)^*$ with $\mathbb{R}^3$ via the Euclidean inner product:

$$\mathrm{ad}^*_{(\xi, \mathbf{u})}(x', \boldsymbol{\alpha}') = (-\mathbb{J}\boldsymbol{\alpha}' \cdot \mathbf{u}, \xi\mathbb{J}\boldsymbol{\alpha}')$$

for any $(\xi, \mathbf{u}) \in \mathfrak{se}(2)$ and $(x', \boldsymbol{\alpha}') \in \mathfrak{se}(2)^*$, and

$$(a, \mathfrak{u}) \times (b, \mathfrak{v}) = (\mathfrak{u} \cdot \mathbb{J}\mathfrak{v}, -a\mathbb{J}\mathfrak{v} + b\mathbb{J}\mathfrak{u})$$

where $\times$ denotes the cross product in $\mathbb{R}^3$, $a, b \in \mathbb{R}$ and $\mathfrak{u}, \mathfrak{v} \in \mathbb{R}^2$.

Since

$$\nabla\left(x' + \frac{1}{4\nu}\|\boldsymbol{\alpha}'\|^2\right) = \left(1, \frac{1}{2\nu}\boldsymbol{\alpha}'\right)$$

is normal to the paraboloid, the area element is given by

$$dA(x', \boldsymbol{\alpha}') \left(\mathrm{ad}^*_{(\xi, \mathbf{u})}(x', \boldsymbol{\alpha}') - (0, 2\nu\mathbb{J}\mathbf{u}), \mathrm{ad}^*_{(\eta, \mathbf{v})}(x', \boldsymbol{\alpha}') - (0, 2\nu\mathbb{J}\mathbf{v})\right)$$

$$= \frac{(2\nu, \boldsymbol{\alpha}')}{\sqrt{4\nu^2 + \|\boldsymbol{\alpha}'\|^2}} \cdot \left((-\mathbb{J}\boldsymbol{\alpha}' \cdot \mathbf{u}, \xi\mathbb{J}\boldsymbol{\alpha}' - 2\nu\mathbb{J}\mathbf{u}) \times (-\mathbb{J}\boldsymbol{\alpha}' \cdot \mathbf{v}, \eta\mathbb{J}\boldsymbol{\alpha}' - 2\nu\mathbb{J}\mathbf{v})\right)$$

$$= \frac{(2\nu, \boldsymbol{\alpha}')}{\sqrt{4\nu^2 + \|\boldsymbol{\alpha}'\|^2}} \cdot \left((\xi\mathbb{J}\boldsymbol{\alpha}' - 2\nu\mathbb{J}\mathbf{u}) \cdot (-\eta\boldsymbol{\alpha}' + 2\nu\mathbf{v}),\right.$$

$$\left.(\mathbb{J}\boldsymbol{\alpha}' \cdot \mathbf{u})(-\eta\boldsymbol{\alpha}' + 2\nu\mathbf{v}) - (\mathbb{J}\boldsymbol{\alpha}' \cdot \mathbf{v})(-\xi\boldsymbol{\alpha}' + 2\nu\mathbf{u})\right)$$

$$= \frac{(2\nu, \boldsymbol{\alpha}')}{\sqrt{4\nu^2 + \|\boldsymbol{\alpha}'\|^2}} \cdot \left(2\nu\big(\boldsymbol{\alpha}' \cdot (\eta\mathbb{J}\mathbf{u} - \xi\mathbb{J}\mathbf{v}) + 2\nu\mathbf{u} \cdot \mathbb{J}\mathbf{v}\big),\right.$$

$$\left.(\boldsymbol{\alpha}' \cdot (\eta\mathbb{J}\mathbf{u} - \xi\mathbb{J}\mathbf{v}))\boldsymbol{\alpha}' + 2\nu\big((\mathbb{J}\boldsymbol{\alpha}' \cdot \mathbf{u})\mathbf{v} - (\mathbb{J}\boldsymbol{\alpha}' \cdot \mathbf{v})\mathbf{u}\big)\right)$$

$$= \frac{1}{\sqrt{4\nu^2 + \|\boldsymbol{\alpha}'\|^2}}\left[4\nu^2\big(\boldsymbol{\alpha}' \cdot (\eta\mathbb{J}\mathbf{u} - \xi\mathbb{J}\mathbf{v}) + 2\nu\mathbf{u} \cdot \mathbb{J}\mathbf{v}\big)\right.$$

$$\left.+ \|\boldsymbol{\alpha}'\|^2\big(\boldsymbol{\alpha}' \cdot (\eta\mathbb{J}\mathbf{u} - \xi\mathbb{J}\mathbf{v})\big) + 2\nu\big((\mathbb{J}\boldsymbol{\alpha}' \cdot \mathbf{u})(\mathbf{v} \cdot \boldsymbol{\alpha}') - (\mathbb{J}\boldsymbol{\alpha}' \cdot \mathbf{v})(\mathbf{u} \cdot \boldsymbol{\alpha}')\big)\right].$$

However, a direct computation shows that

$$(\mathbb{J}\boldsymbol{\alpha}' \cdot \mathbf{u})(\mathbf{v} \cdot \boldsymbol{\alpha}') - (\mathbb{J}\boldsymbol{\alpha}' \cdot \mathbf{v})(\mathbf{u} \cdot \boldsymbol{\alpha}') = (\mathbf{u} \cdot \mathbb{J}\mathbf{v})\|\boldsymbol{\alpha}'\|^2$$

so that the above expression becomes

$$\frac{1}{\sqrt{4\nu^2 + \|\boldsymbol{\alpha}'\|^2}}\left[4\nu^2\big(\boldsymbol{\alpha}' \cdot (\eta\mathbb{J}\mathbf{u} - \xi\mathbb{J}\mathbf{v}) + 2\nu\mathbf{u} \cdot \mathbb{J}\mathbf{v}\big)\right.$$

$$\left.+ \|\boldsymbol{\alpha}'\|^2\big(\boldsymbol{\alpha}' \cdot (\eta\mathbb{J}\mathbf{u} - \xi\mathbb{J}\mathbf{v})\big) + 2\nu(\mathbf{u} \cdot \mathbb{J}\mathbf{v})\|\boldsymbol{\alpha}'\|^2\right]$$

$$= \frac{4\nu^2 + \|\boldsymbol{\alpha}'\|^2}{\sqrt{4\nu^2 + \|\boldsymbol{\alpha}'\|^2}}\left[\boldsymbol{\alpha}' \cdot (\eta\mathbb{J}\mathbf{u} - \xi\mathbb{J}\mathbf{v}) + 2\nu\mathbf{u} \cdot \mathbb{J}\mathbf{v}\right].$$

By (9.3.16), this shows that

$$dA(x', \boldsymbol{\alpha}') = \sqrt{4\nu^2 + \|\boldsymbol{\alpha}'\|^2}\; \widetilde{\omega}^{\nu}_{(x, \boldsymbol{\alpha})}(x', \boldsymbol{\alpha}')$$

as stated in the theorem.

(ii) Setting $\nu = 0$ we see that the nonequivariance cocycle $\sigma^0 = 0$ so the orbit $\mathcal{O}_{(x, \boldsymbol{\alpha}, 0)} = \widetilde{\mathcal{O}}^0_{(x, \boldsymbol{\alpha})}$ is the coadjoint orbit of SE(2) containing $(x, \boldsymbol{\alpha})$. All the other results are obtained trivially by setting $\nu = 0$ in the formulas and computations in (i). ∎

9.4   Bott–Virasoro Group

The purpose of this section is to apply reduction by stages to the cotangent bundle of an interesting group called the Bott–Virasoro group. This group is the unique $\mathbb{R}$-central extension of the orientation preserving diffeomorphism group of the circle.

As with earlier examples, such as the Euclidean group and the oscillator group, one of the byproducts of reduction by stages is the classification of coadjoint orbits. In the case of the Bott–Virasoro group, reduction by stages gets one most of the way towards this goal; at a certain point, one must do a significant calculation to finish the job. These calculations were done from a variety of points of view in Balog, Fehér and Palla [1998], Dai [2000], Dai and Pickrell [2003], Kirillov [1962, 1976a, 1982], Lazutkin and Pankratova [1975], Segal [1981], Witten [1988]. We give the final classification following the approach of Balog, Fehér and Palla [1998].

As soon as we introduce the Bott–Virasoro group and its Lie algebra, we recall how the solutions of the KdV equation may be regarded as geodesics on the Bott–Virasoro group.

The Definition of the Bott–Virasoro Group.   This group, denoted by $\mathrm{BVir}(S^1) := \mathrm{Diff}_+(S^1) \times_B \mathbb{R}$, is, as a set, the Cartesian product

$$\mathrm{BVir}(S^1) = \mathrm{Diff}_+(S^1) \times \mathbb{R}$$

where $\mathrm{Diff}_+(S^1)$ is the group of orientation preserving diffeomorphisms of the circle $S^1 := \{e^{ix} \mid x \in \mathbb{R}\} \equiv \mathbb{R}/2\pi\mathbb{Z}$. Thus, $\xi \in \mathrm{Diff}_+(S^1)$ can be thought of as a strictly increasing diffeomorphism of $\mathbb{R}$ satisfying $\xi(x+2\pi) = \xi(x) + 2\pi$ for all $x \in \mathbb{R}$ and the symbol $\partial_x \xi$ means the derivative of ξ on $\mathbb{R}$. Group multiplication on $\mathrm{BVir}(S^1)$ is defined by

$$(\xi, \alpha)(\eta, \beta) = (\xi \circ \eta, \alpha + \beta + B(\xi, \eta)) . \qquad (9.4.1)$$

Here, $B : \mathrm{Diff}_+(S^1) \times \mathrm{Diff}_+(S^1) \to \mathbb{R}$ is the group two-cocycle called the **Bott two-cocycle**, which was introduced in Bott [1977] (and apparently known earlier to Thurston, according to Guieu and Roger [2003]), and is defined by

$$B(\xi, \eta) = \int_{S^1} \log \partial_x(\xi \circ \eta) \, d \log \partial_x \eta. \qquad (9.4.2)$$

Of course, according to the general theory in §6.2, to build a group extension, one needs to verify the group cocycle identity for B. In this case, this is a direct verification using integration by parts.

Formula (9.4.2) for B was derived by studying the problem of finding group extensions of $\mathrm{Diff}_+(S^1)$. This problem was first solved on the Lie algebra level by Gelfand and Fuchs and then elevated to the group level by Bott. We shall next discuss the issue of the extension of the Lie algebra of vector fields on the circle.

According to equation (6.2.7), the (left) Lie algebra $\mathfrak{v}(S^1) := \mathfrak{X}(S^1) \times_C \mathbb{R}$ of $\text{BVir}(S^1)$, called the ***Virasoro algebra***, has underlying vector space $\mathfrak{v}(S^1) = \mathfrak{X}(S^1) \times \mathbb{R}$ and is endowed with the bracket

$$[(V, a), (W, b)] = ([V, W], C(V, W)). \tag{9.4.3}$$

Here $[V, W]$ is the negative Jacobi-Lie bracket of the vector fields $V, W \in \mathfrak{X}(S^1)$ on S^1, which are identified with 2π-periodic real valued functions on $\mathbb{R}$. Therefore $[V, W] := (\partial_x V)W - (\partial_x W)V$. Also, $C : \mathfrak{X}(S^1) \times \mathfrak{X}(S^1) \to \mathbb{R}$ is the Lie algebra two-cocycle called the ***Gelfand-Fuchs two-cocycle*** (see Gelfand and Fuchs [1968]) and is defined by

$$C(V, W) = 2 \int_{S^1} (\partial_x V)(\partial_x^2 W) \, dx. \tag{9.4.4}$$

One derives this formula for C from that for B using the general equation (6.1.34) by a straightforward calculation.

KdV and Bott–Virasoro. We now recall, for the convenience of the reader, how the KdV equation is related to the Bott–Virasoro group; namely we show that the KdV equation is the Euler-Poincaré equation for a certain right invariant quadratic Lagrangian. Thus, the KdV equation can be viewed as the reduction of geodesic flow on the Bott–Virasoro group. The geodesic nature of the solutions of the KdV equation was apparently known for some time, but this fact was written explicitly in Ovsienko and Khesin [1987]. An exposition of the Euler-Poincaré and Lie-Poisson forms of the KdV equation are given in [MandS], §13.5.

Using equation (9.4.3) and the general formula for the Lie-Poisson bracket (see equation (1.3.1)), we see that the bracket for the dual of the Virasoro algebra is given by

$$\{f, h\}(u, a) = \left\langle (u, a), \left[\frac{\delta f}{\delta(u, a)}, \frac{\delta h}{\delta(u, a)} \right] \right\rangle$$

$$= \int \left[u \left(\left(\frac{\delta f}{\delta u} \right)' \frac{\delta h}{\delta u} - \frac{\delta f}{\delta u} \left(\frac{\delta h}{\delta u} \right)' \right) + 2a \left(\frac{\delta f}{\delta u} \right)' \left(\frac{\delta h}{\delta u} \right)'' \right] dx.$$

Using the KdV Hamiltonian

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

along with this bracket, the equations $\dot{f} = \{f, h\}$ give the KdV equation

$$u_t + 3uu_x + 2au_{xxx} = 0$$

together with the fact that a is a constant (which can be chosen to be 1 by using a change of scale). This is one way of viewing the KdV equation as an infinite dimensional Hamiltonian system. It is interesting as it shows how the KdV equation is the reduction of geodesic flow on the tangent bundle of the Bott–Virasoro group.

The First Reduced Space.   To carry out stage one reduction for the right action of $\mathrm{BVir}(S^1)$ on $T^*\mathrm{BVir}(S^1)$, according to Theorem 8.1.2, we need to compute $\mathcal{B}^1$. By Proposition 8.1.3, $\mathcal{B}^1$ is the right invariant two-form on $\mathrm{Diff}_+(S^1)$ that equals C at the identity. To write a formula for $\mathcal{B}^1$, we first compute the derivative of right translation. Recall that the tangent space to $\mathrm{Diff}_+(S^1)$ at a point ξ is the set of vector fields over ξ; that is, maps that have the form $V \circ \xi$, where $V \in \mathfrak{X}(S^1)$. A straightforward computation shows that the derivative of right translation on $\mathrm{Diff}_+(S^1)$ by η, namely R_η, at the point ξ in the direction $V \circ \xi$ equals

$$T_\xi R_\eta(V \circ \xi) = V \circ \xi \circ \eta. \tag{9.4.5}$$

Right invariance of $\mathcal{B}^1$ gives

$$\mathcal{B}^1(\xi)(V \circ \xi, W \circ \xi) = C(V \circ \xi \circ \xi^{-1}, W \circ \xi \circ \xi^{-1})$$

$$= 2\int_{S^1} (\partial_x V)(\partial_x^2 W)\, dx. \tag{9.4.6}$$

Thus, by Theorem 8.1.2, reducing $T^*\mathrm{BVir}(S^1)$ by the central $\mathbb{R}$-action at a point $\nu \in \mathbb{R}$, usually called the **central charge**, gives

$$\mathbf{J}_\mathbb{R}^{-1}(\nu)/\mathbb{R} \simeq \left(T^*\mathrm{Diff}_+(S^1), \Omega - \nu\pi_{\mathrm{Diff}_+(S^1)}^*\mathcal{B}^1 \right),$$

where $\mathcal{B}^1$ is given by equation (9.4.6). Here $T^*\mathrm{BVir}(S^1)$, the cotangent bundle of $\mathrm{BVir}(S^1)$, is identified with $T\,\mathrm{BVir}(S^1)$ using the L^2-right invariant metric whose value at the identity is

$$\langle (V, a), (W, b) \rangle := \int_{S^1} V(x)W(x)\, dx + ab$$

for $V, W \in \mathfrak{X}(S^1)$ and $a, b \in \mathbb{R}$. Note that elements of $T_\xi \mathrm{Diff}_+(S^1)$ are functions on $\mathbb{R}$ of the form $V \circ \xi$, where V is 2π-periodic.

Remark.   The weakly nondegenerate pairing introduced above formally identifies the dual of the Virasoro algebra with itself. There is another way to think of the dual which is sometimes more natural. In general, on an arbitrary manifold, the dual of the Lie algebra of vector fields is formed by (say, compactly supported) one-form densities, the weak pairing being given by the contraction of the one-form with the vector field and then integrated over the manifold relative to the density. In our case, since the manifold is S^1, one-form densities are quadratic differentials. Thus, another realization of the dual of the Virasoro algebra is the product of quadratic differentials with $\mathbb{R}$. For our purposes (with the exception of some comments at the end of this section), it is more convenient to use the L^2-weak pairing that formally identifies the dual of the Virasoro algebra with itself.

The realization of the duals in different manners has important consequences also in other problems. For example, in ideal homogeneous incompressible fluid dynamics, the L^2 pairing realizes the Lie-Poisson equations

for the geodesic spray as the classical Euler equations, whereas the "natural" pairing using one-form densities (and a very specific way to work with this space) realizes the same abstract equations as the vorticity transport equations, as in Marsden and Weinstein [1983].

The preceding discussion together with Theorem 8.1.1 gives the following mechanical and geometric interpretation of the Gelfand-Fuchs cocycle.

9.4.1 Corollary.   *The Gelfand-Fuchs cocycle C given by (9.4.4) is the value at the identity of the curvature $\mathcal{B}^1$ in (9.4.6) of the mechanical connection $\mathcal{A}^1$ whose expression is given in (8.1.7).*

For completeness, we give the explicit expression for the mechanical connection. This is obtained in the following way. Since

$$\partial_x \xi^{-1} = \frac{1}{\partial_x \xi \circ \xi^{-1}}, \tag{9.4.7}$$

we have for any $V \in \mathfrak{X}(S^1)$

$$
D_1 B(\xi, \xi^{-1})(V \circ \xi) = \int_{S^1} (\partial_x V) \, d\log \partial_x \xi^{-1} = \int_{S^1} (\partial_x V) \frac{d(\partial_x \xi^{-1})}{\partial_x \xi^{-1}}
$$

$$
= \int_{S^1} (\partial_x V)(\partial_x \xi \circ \xi^{-1}) \, d\left(\frac{1}{\partial_x \xi \circ \xi^{-1}}\right)
$$

$$
= -\int_{S^1} (\partial_x V)(\partial_x \xi \circ \xi^{-1}) \frac{d(\partial_x \xi \circ \xi^{-1})}{(\partial_x \xi \circ \xi^{-1})^2}
$$

$$
= -\int_{S^1} (\partial_x V) \frac{\partial_x^2 \xi \circ \xi^{-1}}{(\partial_x \xi \circ \xi^{-1})^2} \, dx = -\int_{S^1} (\partial_x V \circ \xi) \partial_x \xi \frac{\partial_x^2 \xi}{(\partial_x \xi)^2} \, dx
$$

$$
= -\int_{S^1} \partial_x (V \circ \xi) \frac{\partial_x^2 \xi}{(\partial_x \xi)^2} \, dx.
$$

Thus, by (8.1.7), the mechanical connection is given by

$$\mathcal{A}^1(\xi, \alpha)(V \circ \xi, a) = a - \int_{S^1} \partial_x (V \circ \xi) \frac{\partial_x^2 \xi}{(\partial_x \xi)^2} \, dx. \tag{9.4.8}$$

The Second Reduction.   We begin by calculating the $\mathcal{B}\mathfrak{g}$-potential given in (8.2.7). In our case, this formula reads

$$\langle \phi^\nu(\xi), V \rangle = -\nu \left(D_2 B(\xi, \mathrm{Id}) \cdot V + D_1 B(\xi, \xi^{-1}) \cdot (T_{\mathrm{Id}} L_\xi(V)) \right), \tag{9.4.9}$$

The first term of (9.4.9) is determined by

$$
D_2 B(\xi, \mathrm{Id}) \cdot V = \int_{S^1} \log \partial_x \xi \cdot \partial_x^2 V \, dx
$$

$$
= \int_{S^1} \partial_x^2 (\log \partial_x \xi) V \, dx
$$

$$
= \left\langle \frac{\partial_x^3 \xi \cdot \partial_x \xi - (\partial_x^2 \xi)^2}{(\partial_x \xi)^2}, V \right\rangle. \tag{9.4.10}
$$

To compute the second term in (9.4.9), we must compute the quantity $D_1 B(\xi, \xi^{-1}) \cdot (T_{\mathrm{Id}} L_\xi(V))$. Making use of (9.4.7) and integrating by parts, we get

$$
\begin{aligned}
D_1 &B(\xi, \xi^{-1}) \cdot (T_{\mathrm{Id}} L_\xi(V)) \\
&= \int_{S^1} \partial_x \left((\partial_x \xi \circ \xi^{-1})(V \circ \xi^{-1}) \right) \, d\log \partial_x \xi^{-1} \\
&= \int_{S^1} \partial_x \left((\partial_x \xi \circ \xi^{-1})(V \circ \xi^{-1}) \right) \partial_x \log \left(\frac{1}{\partial_x \xi \circ \xi^{-1}} \right) dx \\
&= -\int_{S^1} (\partial_x \xi \circ \xi^{-1})(V \circ \xi^{-1}) \partial_x \left[(\partial_x \xi \circ \xi^{-1}) \, \partial_x \left(\frac{1}{\partial_x \xi \circ \xi^{-1}} \right) \right] dx \\
&= \int_{S^1} (\partial_x \xi \circ \xi^{-1})(V \circ \xi^{-1}) \partial_x \left[\frac{\partial_x^2 \xi \circ \xi^{-1}}{(\partial_x \xi \circ \xi^{-1})^2} \right] dx. \qquad (9.4.11)
\end{aligned}
$$

Carrying out the differentiation in the last factor of (9.4.11) and changing variables gives

$$
\begin{aligned}
D_1 B(\xi, \xi^{-1}) \cdot (T_{\mathrm{Id}} L_\xi(V)) &= \int_{S^1} V \frac{(\partial_x \xi)(\partial_x^3 \xi) - 2(\partial_x^2 \xi)^2}{(\partial_x \xi)^2} \, dx \\
&= \left\langle \frac{(\partial_x \xi)(\partial_x^3 \xi) - 2(\partial_x^2 \xi)^2}{(\partial_x \xi)^2}, V \right\rangle \qquad (9.4.12)
\end{aligned}
$$

Adding the results of (9.4.10) and (9.4.12) gives

$$
\phi^\nu(\xi) = -\nu \frac{2\partial_x^3 \xi \cdot \partial_x \xi - 3(\partial_x^2 \xi)^2}{(\partial_x \xi)^2} \qquad (9.4.13)
$$

Therefore, by (8.2.13), we get

$$
\sigma^\nu(\xi) = -\phi^\nu(\xi) = 2\nu S(\xi), \qquad (9.4.14)
$$

where S is the ***Schwarzian derivative***, defined by[3]

$$
S(\xi) = \frac{\partial_x^3 \xi}{\partial_x \xi} - \frac{3}{2} \left[\frac{\partial_x^2 \xi}{\partial_x \xi} \right]^2.
$$

A Schwarzian Excursion.   We can now use this new interpretation of the Schwarzian derivative as a group one-cocycle for a Hamiltonian action to produce a new, easy proof of the following well known properties of the Schwarzian derivative.

---

[3]The Schwarzian derivative is used in a variety of contexts. For example, it comes up in complex variables as a quantity that is invariant under the action of the group of fractional linear transformations.

9.4.2 Theorem.   *The Schwarzian derivative satisfies the following properties:*

(i) $S(\xi \circ \eta) = S(\eta) + (S(\xi) \circ \eta)(\partial_x \eta)^2$,

(ii) $S(\mathrm{id}) = 0$,

(iii) $T_{\mathrm{id}} S \cdot V = \partial_x^3 V$.

Proof.   In fact, we claim that the first part is a consequence of the cocycle identity (7.1.3), which in our case reads

$$\sigma^\nu(\xi \circ \eta) = \mathrm{Ad}_\eta^* \sigma^\nu(\xi) + \sigma^\nu(\eta). \tag{9.4.15}$$

However, an easy calculation using our expressions for the derivatives of left and right translation shows that $\mathrm{Ad}_\eta^* V = (\partial_x \eta)^2 (V \circ \eta)$. Thus, (i) follows.

Part (ii) is obvious from the definition of S and is also consistent with the general fact that $\sigma^\nu(\mathrm{id}) = 0$ for cocycles.

Finally, for the last item, recall from the general equation (6.2.9) that $\langle T_{\mathrm{id}} \sigma_R(V), W \rangle = C(V, W)$. By (8.2.12), we have $\sigma^\nu = \nu \sigma_R$ and hence $\sigma_R = 2S$. Therefore

$$\langle 2 T_{\mathrm{id}} S \cdot V, W \rangle = 2 \int_{S^1} (\partial_x V)(\partial_x^2 W) \, dx = 2 \int_{S^1} (\partial_x^3 V) W \, dx$$

for any $W \in \mathfrak{X}(S^1)$ which proves the formula.   ∎

Remark. For future use it is worth recording the following formulas

$$\mathrm{Ad}_{(\xi,\alpha)}(V, a) = \left((\partial_x \xi \cdot V) \circ \xi^{-1}, a + 2 \int_{S^1} S(\xi) V \, dx \right)$$
$$\mathrm{Ad}_{(\xi,\alpha)}^*(\mu, \nu) = \left((\partial_x \xi)^2 (\mu \circ \xi) + 2\nu S(\xi), \nu \right)$$
$$\mathrm{ad}_{(V,a)}^*(\mu, \nu) = \left(2\mu \partial_x V + V \partial_x \mu + 2\nu \partial_x^3 V, 0 \right),$$

where $V, \mu \in \mathfrak{X}(S^1)$ and $a, \nu \in \mathbb{R}$, which follow from (6.2.15), (6.2.16), and (6.2.17). Consistent with our conventions, here we think of μ as a vector field on S^1. Had we used the "natural" dual, μ would be thought of as a quadratic differential.

To be crystal clear about how to read these formulas, here is a more detailed explanation of the formula for the Ad action. With all the variables in, the first component of this formula reads

$$[(\partial_x \xi \cdot V) \circ \xi^{-1}](x) = \partial_x \xi(\xi^{-1}(x) \cdot V((\xi^{-1}(x)),$$

where the dot means pointwise multiplication. Since

$$\partial_x \xi^{-1}(x) = \frac{1}{\partial_x \xi(\xi^{-1}(x))},$$

the coefficient in front of V may be written $(\partial_x \xi)^{-1}$; note that this coefficient is computed at x whereas V is computed at $\xi^{-1}(x)$.

Coadjoint Orbits for the Bott–Virasoro Group.   At this point one can compute the isotropy algebras for the affine action $\widetilde{\mathfrak{g}}^\nu_\mu$ and thereby classify the coadjoint orbits of the Bott–Virasoro group through various points $(\mu, \nu) \in \mathfrak{X}(S^1) \times \mathbb{R}$. These orbits are well known and their classification from various points of view can be found in the papers cited in the introduction.

Note that for all real numbers $\lambda \neq 0$, we have $\widetilde{G}^{\lambda\nu}_{\lambda\mu} = \widetilde{G}^\nu_\mu$ so it suffices to consider the case $\nu = 0$ and the case $\nu = \nu_0$ for a fixed $\nu_0 \neq 0$. We will see that a natural choice is $\nu_0 = 1/4$.

The case $\nu = 0$. This case corresponds to the coadjoint orbits of $\mathrm{Diff}_+(S^1)$ (without the central extension). A clear presentation of this situation can be found, for example, in Guieu [2000] or Guieu and Roger [2003].

For $\mu \in \mathfrak{X}(S^1)$, consider the compact subset $Z_\mu := \{e^{ix} \in S^1 \,|\, \mu(e^{ix}) = 0\}$ of S^1.

• If $Z_\mu = \varnothing$, then μ belongs to the coadjoint orbit $\mathcal{O}_c$ generated by the non-zero constant vector field $c = \mathrm{sgn}(\mu)\alpha_\mu$ where

$$\alpha_\mu := \int_{S^1} \sqrt{|u|}\, dx$$

and $\mathrm{sgn}(\mu)$ denotes the sign of μ. Note that for different values of c, the corresponding coadjoint orbits are disjoint. The isotropy group of the constant representative vector field is S^1.

• If $Z_\mu \neq \varnothing$ and $\mathrm{Int}(Z_\mu) = \varnothing$ (for example, if Z_μ is finite or consists of a finite number of convergent sequences) then the isotropy group of μ is isomorphic to $\mathbb{Z}_n$ for some integer n. When the number of zeros of μ is finite then n divides this number. If, in addition, the zeros of μ are also simple, then $\mathcal{O}_\mu$ is completely determined (Kirillov [1982], Guieu and Roger [2003]). The other cases are not yet classified.

• If $\mathrm{Int}(Z_\mu) \neq \varnothing$, then the isotropy group of μ is infinite dimensional and contains the infinite dimensional subgroup

$$\mathrm{Diff}_+(S^1)_{Z_\mu} := \{\xi \in \mathrm{Diff}_+(S^1) \,|\, \xi(x) = x \text{ for all } x \notin \mathrm{Int}(Z_\mu)\}.$$

In particular if $\mu = 0$, then the isotropy group is $\mathrm{Diff}_+(S^1)$ and the coadjoint orbit is $\{0\}$. The classification in this case is also not completely known.

The case $\nu \neq 0$. We follow the complete classification as presented in Balog, Fehér and Palla [1998]; for another approach see Dai [2000], Dai and Pickrell [2003]. This classification has been extended to the corresponding group of diffeomorphisms of a suitable Sobolev class; in this case, $\mathrm{Diff}^{s+3}_+(S^1)$ acting on $\mathfrak{X}^s(S^1)$, $s > 1/2$, in Gay-Balmaz [2007].

Here is the idea of how the case $\nu \neq 0$ is dealt with. For $\mu \in \mathfrak{X}(S^1)$, consider the **_Hill equation_** with potential μ, given by

$$\partial_x^2 \psi + \mu\,\psi = 0. \tag{9.4.16}$$

By a **normalized base of solutions** associated to μ, we mean a pair of solutions $\Psi := (\psi_1, \psi_2)$ of (9.4.16) satisfying $\psi_2 \partial_x \psi_1 - \psi_1 \partial_x \psi_2 = 1$. If $\Phi = (\varphi_1, \varphi_2)$ is another normalized base of solutions, then there exists $A \in \mathrm{SL}(2, \mathbb{R})$ such that $\Psi = \Phi A$. For example, for $\tilde{\Psi}(x) := \Psi(x + 2\pi)$, Floquet theory gives the existence of $M_\Psi \in \mathrm{SL}(2, \mathbb{R})$, called the **monodromy matrix** of Ψ, such that

$$\tilde{\Psi} = \Psi M_\Psi.$$

Since $M_{\Psi A} = A^{-1} M_\Psi A$, there exists a well defined map

$$\mathrm{M} : \mu \in \mathfrak{X}^s(S^1) \longmapsto \mathrm{M}(\mu) := [M_\Psi] \in \mathrm{SL}(2, \mathbb{R})/\mathrm{conj}$$

where $\mathrm{SL}(2, \mathbb{R})/\mathrm{conj}$ denotes the set of conjugacy classes of $\mathrm{SL}(2, \mathbb{R})$, Ψ is any normalized base of solutions associated to μ, and $[M_\Psi]$ denotes the conjugacy class of M_Ψ.

A computation shows that M has a remarkable property; it is invariant under the coadjoint action of $\mathrm{Diff}_+(S^1)$ for the central charge $\nu = 1/4$, that is,

$$\mathrm{M}(\mathrm{Ad}^*_\xi \mu + \sigma^{1/4}(\xi)) = \mathrm{M}(\mu), \quad \text{for all} \quad \xi \in \mathrm{Diff}_+(S^1). \tag{9.4.17}$$

A key step in this computation is this: if ψ is a solution of (9.4.16) with potential μ, then

$$\psi^\xi := \frac{1}{\sqrt{\partial_x \xi}} \psi \circ \xi$$

is a solution of (9.4.16) with potential

$$\mathrm{Ad}^*_\xi \mu + \sigma^{1/4}(\xi) = \mathrm{Ad}^*_\xi \mu + \frac{1}{2} S(\xi).$$

Since $M_{\Psi^\xi} = M_\Psi$, where $\Psi^\xi := (\psi_1^\xi, \psi_2^\xi)$, we obtain (9.4.17). If Ψ is normalized, then Ψ^ξ is also normalized. Thus M induces a well-defined map

$$\mathcal{M} : \mathcal{O} \in \left(\mathfrak{X}(S^1) \times \{1/4\}\right)/\mathrm{Diff}_+(S^1) \longmapsto \mathcal{M}(\mathcal{O}) := \mathrm{M}(\mu) \in \mathrm{SL}(2, \mathbb{R})/\mathrm{conj},$$

where $\left(\mathfrak{X}(S^1) \times \{1/4\}\right)/\mathrm{Diff}_+(S^1)$ denotes the set of coadjoint orbits with central charge $\nu = 1/4$ and μ is any element in the coadjoint orbit $\mathcal{O}$.

Now we list the elements of $\mathrm{SL}(2, \mathbb{R})/\mathrm{conj}$:

(i) an **elliptic** conjugacy class is represented by a matrix $\mathrm{Ell}_\pm(\alpha)$ of the form

$$\mathrm{Ell}_\pm(\alpha) = \pm \begin{bmatrix} \cos 2\pi\alpha & -\sin 2\pi\alpha \\ \sin 2\pi\alpha & \cos 2\pi\alpha \end{bmatrix}, \quad \alpha \in (0, 1/2),$$

(ii) a **hyperbolic** conjugacy class is represented by a matrix $\mathrm{Hyp}_\pm(\beta)$ of the form

$$\mathrm{Hyp}_\pm(\beta) = \pm \begin{bmatrix} e^{2\pi\beta} & 0 \\ 0 & e^{-2\pi\beta} \end{bmatrix}, \quad \beta > 0,$$

(iii) a **_parabolic_** conjugacy class is represented by a matrix $\mathrm{Par}_\pm(\varepsilon)$ of the form

$$\mathrm{Par}_\pm(\varepsilon) = \pm \begin{bmatrix} 1 & 0 \\ \varepsilon & 1 \end{bmatrix}, \quad \varepsilon \in \{-1, 0, 1\}.$$

The classification is the following.

(i) If $\mathcal{M}(\mathcal{O}) = [\mathrm{Ell}_\pm(\alpha)]$, then $\mathcal{O}$ is one of the distinct orbits denoted by $\mathrm{Ell}(\alpha, n)$, $n = 0, 1, 2, \ldots$, and is generated by the constant vector field $\left(\alpha + \frac{n}{2}\right)^2 \partial_x$. The monodromy is $\mathcal{M}(\mathrm{Ell}(\alpha, n)) = \mathrm{Ell}_{(-1)^n}(\alpha)$.

(ii) If $\mathcal{M}(\mathcal{O}) = [\mathrm{Hyp}_\pm(\beta)]$, then $\mathcal{O}$ is one of the distinct orbits denoted by $\mathrm{Hyp}(\beta, n)$, $n = 0, 1, 2, \ldots$, and is generated by the vector field

- $-\beta^2 \partial_x$, if $n = 0$, and

- $\left(-\beta^2 - \dfrac{n^2 + 4\beta^2}{2F_{\beta,n}} + \dfrac{3}{4} \dfrac{n^2}{F_{\beta,n}^2} \right) \partial_x$, if $n \neq 0$, where the function $F_{\beta,n}$

    is given by $F_{\beta,n} = \cos^2 \dfrac{nx}{2} + \left(\sin \dfrac{nx}{2} + \dfrac{2\beta}{n} \cos \dfrac{nx}{2} \right)^2$.

    The monodromy is $\mathcal{M}(\mathrm{Hyp}(\beta, n)) = \mathrm{Hyp}_{(-1)^n}(\beta)$.

(iii1) If $\mathcal{M}(\mathcal{O}) = [\mathrm{Par}_\pm(0)]$, then $\mathcal{O}$ is one of the distinct orbits denoted by $\mathrm{Par}(0, n)$, $n \in \mathbb{N}$, $n \neq 0$, and is generated by the constant vector field $\frac{n^2}{4} \partial_x$. The monodromy is $\mathcal{M}(\mathrm{Par}(0, n)) = \mathrm{Par}_{(-1)^n}(0)$.

(iii2) If $\mathcal{M}(\mathcal{O}) = [\mathrm{Par}_\pm(\varepsilon)]$, $\varepsilon = \pm 1$, then $\mathcal{O}$ is one of the distinct orbits denoted by $\mathrm{Par}(\varepsilon, n)$, $n = 0, 1, 2, \ldots$, with the condition $(\varepsilon, n) \neq (-1, 0)$, and is generated by the vector field

$$\left(\frac{3n^2 \left(1 + \frac{\varepsilon}{2\pi}\right)}{4H_{\varepsilon,n}^2} - \frac{n^2}{2H_{\varepsilon,n}} \right) \partial_x \quad \text{where} \quad H_{\varepsilon,n}(x) = 1 + \frac{\varepsilon}{2\pi} \sin^2 \frac{nx}{2}.$$

The monodromy is $\mathcal{M}(\mathrm{Par}(\varepsilon, n)) = \mathrm{Par}_{(-1)^n}(\varepsilon)$.

The coadjoint orbits of the Virasoro algebra can be represented by the "comb" with additional points shown in Figure 9.4.1.

For the orbits $\mathrm{Ell}(\alpha, n), \mathrm{Hyp}(\beta, 0)$, and $\mathrm{Par}(1, 0)$, the isotropy group of the given representative vector field is S^1, acting on S^1 by rotations.

Recall (see, for instance, Knapp [2002]) that $\mathrm{PSU}(1, 1)$ is the group $\mathrm{SU}(1, 1)/\{\pm I\}$, where I is the 2×2 identity and $\mathrm{SU}(1, 1)$ is the group of complex 2×2 matrices of determinant 1 that preserve the Hermitian form on $\mathbb{C}^2$ given by $|z_1|^2 - |z_2|^2$.

For the orbits $\mathrm{Par}(0, n)$, $n \neq 0$, the isotropy group is the n-fold cover $\widetilde{\mathrm{PSU}(1,1)}^n$ of $\mathrm{PSU}(1, 1)$, acting on S^1 as

$$e^{ix} \longmapsto \left(\frac{Ae^{inx} + \overline{B}}{Be^{inx} + \overline{A}} \right)^{\frac{1}{n}} \quad \text{where} \quad \begin{bmatrix} A & \overline{B} \\ B & \overline{A} \end{bmatrix} \in \mathrm{SU}(1, 1),$$
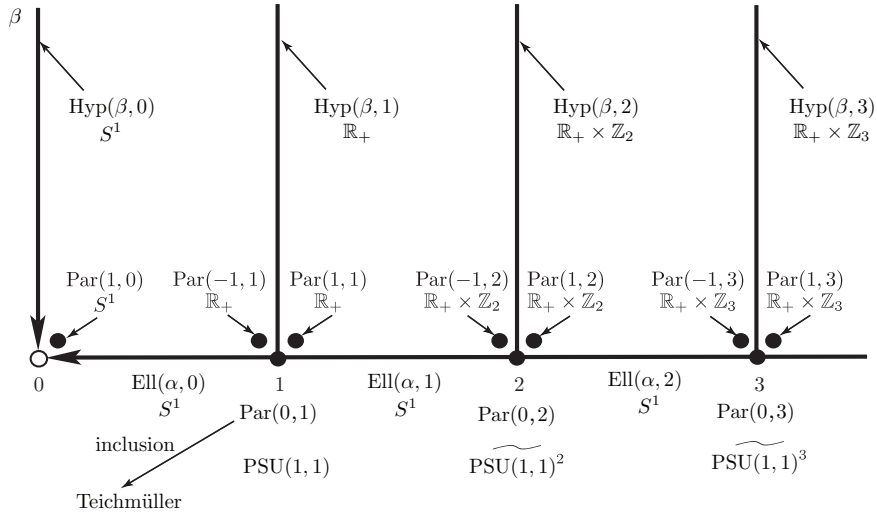
FIGURE 9.4.1. The space of coadjoint orbits of the Bott–Virasoro group for nonzero charge. Points on the "comb" together with the "floating points" represent the space of coadjoint orbits. The vertical lines, each of which is parametrized by β, and labeled by an integer $n = 0, 1, 2, \ldots$, represent the hyperbolic orbits. The non-integer points on the horizontal axis represent the elliptic orbits, while the integer points represent those parabolic orbits $\mathrm{Par}(\varepsilon, n)$ with $\varepsilon = 0$. The parabolic orbits $\mathrm{Par}(\varepsilon, n)$ with $\varepsilon = \pm 1$ and $n = 0, 1, 2, \ldots$ are represented by the "floating points". The open circle at $n = 0, \beta = 0$ is an empty point with no corresponding orbit. Figure adapted from Balog, Fehér and Palla [1998].

with $|A|^2 - |B|^2 = 1$.

For the orbits $\mathrm{Hyp}(\beta, n)$, $n \neq 0$, and $\mathrm{Par}(\varepsilon, n)$, $\varepsilon = \pm 1$, $n \neq 0$, the isotropy group is isomorphic to $\mathbb{R}_+ \times \mathbb{Z}_n$, where $\mathbb{R}_+$ is the multiplicative group of the positive real numbers. See Appendix C of Balog, Fehér and Palla [1998] to understand how it acts on S^1.

In Gay-Balmaz [2007] it is shown that, in the setting of smooth Hilbert manifolds (modeled on Sobolev spaces), the isotropy groups are submanifolds and finite dimensional Lie groups. The coadjoint orbits are injectively immersed closed Hilbert submanifolds.

The symplectic structures. By formula (8.2.3), the symplectic form on a coadjoint orbit $\widetilde{\mathcal{O}}_\mu^\nu$ is given by

$$\tilde{\omega}_\mu^\nu(\lambda) \left(\mathrm{ad}_V^* \lambda + 2\nu \partial_x^3 V, \mathrm{ad}_W^* \lambda + 2\nu \partial_x^3 W\right)$$
$$= \int_{S^1} \lambda((\partial_x V)W - (V\partial_x)W) \, dx + 2\nu \int_{S^1} (\partial_x^3 V)W \, dx,$$

where the symplectic form is evaluated at the point $\lambda \in \widetilde{\mathcal{O}}_\mu^\nu$ and on the pair of tangent vectors $\mathrm{ad}_V^* \lambda + 2\nu\partial_x^3 V, \mathrm{ad}_W^* \lambda + 2\nu\partial_x^3 W \in T_\lambda\widetilde{\mathcal{O}}_\mu^\nu$, for $V, W \in \mathfrak{X}(S^1)$.

Notice that in the case $\nu = 0$ the symplectic forms are simply

$$\tilde{\omega}_\mu^0(\lambda)\,(\mathrm{ad}_V^* \lambda, \mathrm{ad}_W^* \lambda) = \int_{S^1} \lambda(W\partial_x V - V\partial_x W)\, dx.$$

Links with Teichmüller Space.   An object that is closely related to item (iii1) in the above discussion of coadjoint orbits for the Bott–Virasoro group is the universal Teichmüller space $\mathcal{T}$ introduced in Bers [1965, 1970]. For a general introduction to this subject and references, see, for instance, Lehto [1987].

To explain what is going on, we will need to review a few facts from this theory. First of all, recall from complex variables theory that a **_conformal mapping_** in the complex plane on the linearized level is a rotation together with a stretch; thus, a conformal mapping sends, to first order, small discs to other discs. A **_quasiconformal mapping_**, roughly speaking, sends small discs to small ellipses, in which the ratio of major axis to minor axis is bounded. An orientation preserving homeomorphism of the unit circle S^1 is called **_quasisymmetric_** if it extends to a quasiconformal map of the closed unit disc in $\mathbb{C}$. The set of all quasisymmetric homeomorphisms $\mathrm{QS}(S^1)$ is a Banach manifold whose tangent space at the identity is, by definition, the **_classical Zygmund space_** $C_*^1(S^1)$ of maps of S^1 to $\mathbb{R}$, with the norm

$$\|V\|_{C_*^1} = \|V\|_\infty + \sup_{x,h\neq 0}\left|\frac{V(x+h) + V(x-h) - 2V(x)}{h}\right|$$

(see Zygmund [2002]). It is also a group under composition of homeomorphisms. However, unlike $\mathrm{Diff}_+(S^1)$, it is not a topological group (see Gardiner and Sullivan [1992]). The **_universal Teichmüller space_** can now be characterized as the homogeneous space

$$\mathcal{T} = \mathrm{QS}(S^1)/\mathrm{PSU}(1,1) \tag{9.4.18}$$

consisting of those quasisymmetric homeomorphisms which fix the points ± 1 and $-i$ on the circle. As a manifold, $\mathcal{T}$ is a complex Banach manifold. It carries a Kähler metric which is, however, not defined on all tangent vectors (see, for details, Nag and Verjovsky [1990]).

Classical Teichmüller and Moduli Spaces.   There is another approach to Teichmüller space based on moduli spaces of Riemann surfaces. This viewpoint has its origins in the work of Teichmüller [1939] with key contributions from Ahlfors [1961] amongst many others, and its symplectic and Kähler geometry was developed by Atiyah and Bott [1982, 1984] and Goldman [1984], references cited therein along with the work of many others.

The adjective "universal" for the space $\mathcal{T}$ is due to the fact that it contains, as complex submanifolds, all Teichmüller spaces of Riemann surfaces. We recall that the Teichmüller space of a Riemann surface M is the complex manifold whose points represent all complex structures of Riemann surfaces whose underlying topological structure is the same as that of M. In this context, the geometry of the Weil-Petersson metric as well as curvature computations have been much studied; see Ahlfors [1962], Tromba [1986, 1992] and Wolpert [1986, 2003] for further information and additional references.

Teichmüller Space as a Coadjoint Orbit.   Let $\mathcal{T}_0$ denote the **coadjoint Teichmüller space**, defined to be the specific coadjoint orbit $\mathrm{Par}(0,1)$ of $\mathrm{BVir}(S^1)$, which can be identified with $\mathrm{Diff}_+(S^1)/\mathrm{PSU}(1,1)$; that is,

$$\mathcal{T}_0 := \mathrm{Par}(0,1) \cong \mathrm{Diff}_+(S^1)/\mathrm{PSU}(1,1). \qquad (9.4.19)$$

It is known that $\mathcal{T}_0$ is a Fréchet Kähler manifold. In addition, as hinted at in Figure 9.4.1, there is an inclusion map

$$\mathcal{T}_0 \hookrightarrow \mathcal{T},$$

which is complex analytic. The coadjoint orbit symplectic structure on $\mathcal{T}_0$ is the imaginary part of this Kähler structure; in other words, the structure coming from coadjoint orbit reduction agrees with that coming from the Kähler structure. On $\mathcal{T}$ as well as on $\mathcal{T}_0$, the real part of this Kähler structure is the famous **Weil-Petersson metric** (see Nag and Verjovsky [1990] for details).

Takhtajan and Teo [2004, 2006] take a completely different point of view motivated by the fact that in the complex Banach manifold topology, the Weil-Petersson metric on $\mathcal{T}$ is not everywhere defined. They define a new complex Hilbert manifold structure on $\mathcal{T}$ for which the Weil-Petersson metric is strong, is the metric part of a Kähler–Einstein structure and has negative Ricci curvature and negative sectional curvature.

Hamiltonian Flows on Teichmüller Space.   Ignoring functional analytic issues of precise function spaces for the present discussion, we have seen that since the Bott–Virasoro group is a central extension, one can perform reduction (Poisson or symplectic) in either one step or in two stages. From the point of view of Poisson reduction one of course ends up with the dual of the Lie algebra of the Bott–Virasoro group with its Lie-Poisson structure. From the point of view of *symplectic* reduction by stages, things are more intricate and one ends up reducing to a *particular* coadjoint orbit of the Bott–Virasoro group. The stages methodology is helpful when one is carrying this out in detail. The usefulness of the stages technique will be seen in another concrete and nontrivial example in §10.7.

First, a simple but basic remark: The KdV flow regarded as geodesic flow on the Bott–Virasoro group, as described at the beginning of this section,

induces by reduction, a flow on each coadjoint orbit and in particular, on the special coadjoint orbit $\mathcal{T}_0$. This corresponds to imposing special initial conditions for the KdV equation, which are then maintained automatically by the flow. In summary, it follows from basic reduction principles that *the KdV equation induces a flow on coadjoint Teichmüller space $\mathcal{T}_0$.*

But there is something else very special about the coadjoint Teichmüller space $\mathcal{T}_0$, namely, it is a *group*! This is because one can identify $\mathcal{T}_0$ with the subgroup of $\mathrm{Diff}_+(S^1)$ consisting of diffeomorphisms that fix the three points ± 1 and $-i$. Correspondingly, its Lie algebra, namely, $T_e\mathcal{T}_0$ is identified with the Lie subalgebra

$$\mathfrak{h} = \{u \in \mathfrak{X}(S^1) \,|\, u(\pm 1) = u(-i) = 0\}.$$

The group $\mathcal{T}_0$ completes in the topology of Takhtajan and Teo [2004, 2006] to the connected component of the class of the identity of $\mathcal{T}$. This connected component turns out to be a topological group with biholomorphic right translations. Thus, it is a strong Kähler manifold. Moreover, the Kähler structure is right invariant (so the symplectic form and the Weil-Petersson metric are as well).

An interesting issue that was raised in Schonbek, Todorov, and Zubelli [1999] is to compute the Euler-Poincaré equations on $\mathcal{T}_0$ for the Weil-Petersson metric. The resulting **Euler-Weil-Petersson** equations are computed, along with many related facts in Gay-Balmaz, Ratiu, and Marsden [2007].

9.5   Fluids with a Spatial Symmetry

Commuting reduction by stages is already quite interesting in fluid dynamics where it raises several difficult technical issues. We shall present here the general example of a fluid motion invariant under the maximal spatial rigid symmetry, the isometry group of the domain. We shall take the domain to be a compact oriented Riemannian manifold with smooth boundary so the maximal group of spatial symmetries is the group of isometries preserving the orientation. Its natural action on spatial velocities is non-free. Even though non-free reduction by stages is treated in Part III of this book, we include this example here because, from the point of view of Poisson geometry, it is relatively easy; the main technical difficulties lie elsewhere, namely, in the choice of function spaces and in coming to grips with the fact that the group of Sobolev class diffeomorphisms is not a Banach Lie group as well as the non-smoothness of the spatial action.

An example of this situation is the motion of a two-dimensional ideal incompressible fluid in a disc, where the spatial invariance group is the circle. Even in this apparently simple example there are already serious technical difficulties. In this section we shall explain the problem and quote

the main results. For details we refer to Vasylkevych and Marsden [2005] for the first stage reduction and to Gay-Balmaz and Ratiu [2006] for the second stage reduction where all technical details can be found.

The Problem Setting.   Consider the motion of an incompressible ideal fluid in a compact oriented Riemannian manifold M with smooth boundary ∂M. It is well known that a possible choice for the configuration space of this problem is $\mathcal{D}_\mu^s(M)$, the Hilbert manifold of volume preserving H^s-diffeomorphisms of M, where $s > (\dim(M)/2)+1$ and that the appropriate Lagrangian is the kinetic energy of the weak L^2 Riemannian metric

$$\langle\!\langle u_\eta, v_\eta \rangle\!\rangle_\eta = \int_M g(\eta(x))(u_\eta(x), v_\eta(x))\mu(x), \ \ u_\eta, v_\eta \in T_\eta\mathcal{D}_\mu^s(M),$$

where g is the Riemannian metric on M and μ is the volume form induced by g. This Lagrangian is invariant under the following two commuting (right and left) actions

$$R : \mathcal{D}_\mu^s(M) \times T\mathcal{D}_\mu^s(M) \longrightarrow T\mathcal{D}_\mu^s(M), \ \ R(\eta, v_\xi) = R_\eta(v_\xi) := v_\xi \circ \eta$$

$$L : \mathrm{Iso}^+ \times T\mathcal{D}_\mu^s(M) \longrightarrow T\mathcal{D}_\mu^s(M), \ \ L(i, v_\xi) = L_i(v_\xi) := Ti \circ v_\xi,$$

where $\mathrm{Iso}^+ := \mathrm{Iso}^+(M, g)$ denotes the group of Riemannian isometries of (M, g) which preserve the orientation, that is, Iso^+ is the connected component of the identity of the isometry group $\mathrm{Iso} := \mathrm{Iso}(M, g)$ of (M, g). As usual, R is the particle relabeling action whereas the left action L represents spatial motions. Formally, Poisson reduction by stages can be applied to this situation. Here we address the question of how to make this precise and how one deals with the non-free action of Iso^+. Of course it would be nice if one could also deal with point reduction by stages in a similarly rigorous manner, but this would be more difficult.

First Stage Reduction.   We begin with the reduction by the particle relabeling group $\mathcal{D}_\mu^s(M)$ which is well known (see Ebin and Marsden [1970]) and leads to the Euler equations for an ideal incompressible fluid on the first reduced space $\mathfrak{X}_{\mathrm{div}}^s(M) = T\mathcal{D}_\mu^s(M)/\mathcal{D}_\mu^s(M)$ consisting of H^s divergence free vector fields on M that are tangent to the boundary. The most fundamental fact is the existence of the C^∞ geodesic spray $\mathcal{S} \in \mathfrak{X}(T\mathcal{D}_\mu^s(M))$ of the weak Riemannian manifold $(\mathcal{D}_\mu^s(M), \langle\!\langle \, , \rangle\!\rangle)$ which implies that there is no derivative loss in the Lagrangian formulation. Recall that the Euler equations for the spatial velocity u are given by

$$\partial_t u(t) + \nabla_{u(t)} u(t) = -\operatorname{grad} p(t)$$

for some scalar function $p(t) : M \to \mathbb{R}$ called the pressure. An equivalent formulation of these equations is given by

$$\partial_t u(t) = -P_e(\nabla_{u(t)} u(t)),$$

where P_e denotes the projection on the divergence free factor of the Hodge decomposition. A consequence of the smoothness of the geodesic spray is the local existence and uniqueness of the solution u of the Euler equations and their continuous dependence on initial conditions. Moreover, $u \in C^0(I, \mathfrak{X}^s_{\mathrm{div}}(M)) \cap C^1(I, \mathfrak{X}^{s-1}_{\mathrm{div}}(M))$. The dependence on the initial conditions of the integral curves of the geodesic spray $\mathcal{S}$, that is, of the vector field describing the fluid in material representation, is C^∞.

Denoting by $\pi_R : T\mathcal{D}^s_\mu(M) \longrightarrow \mathcal{D}^s_\mu(M)$, $\pi_R(u_\eta) := u_\eta \circ \eta^{-1}$, the projection associated to the reduction by $\mathcal{D}^s_\mu(M)$, we obtain the following commutative diagram

$$
\begin{array}{ccc}
T\mathcal{D}^s_\mu(M) & \xrightarrow{\ F_t\ } & T\mathcal{D}^s_\mu(M) \\
{\scriptstyle \pi_R} \Big\downarrow & & \Big\downarrow {\scriptstyle \pi_R} \\
\mathfrak{X}^s_{\mathrm{div}}(M) & \xrightarrow{\ \widetilde{F}_t\ } & \mathfrak{X}^s_{\mathrm{div}}(M),
\end{array}
$$

where F_t is the flow of $\mathcal{S}$ and $\widetilde{F}_t$ is the flow of the Euler equations. In this diagram, for each fixed t, the map F_t is C^∞ whereas $\widetilde{F}_t$ and π_R are only C^0. Formally, all these maps are Poisson, as it is the case in the standard Lie-Poisson reduction procedure. However, since $\mathcal{D}^s_\mu(M)$ is not really a Lie group and the manifolds are infinite dimensional some technical problems arise. First, the symplectic form on $T\mathcal{D}^s_\mu(M)$ is only weak since the Lagrangian is given by a L^2 metric. Second, $\mathcal{D}^s_\mu(M)$ is not a Lie group since left multiplication and inversion are not smooth; this explains why the projection π_R is only continuous. Vasylkevych and Marsden [2005] have resolved these difficulties by carefully analyzing the function spaces on which Poisson brackets are defined and carrying out a non-smooth Lie-Poisson reduction that takes into account all analytical technical issues. Thus if one evaluates the maps in the diagram above on very precise function spaces, then the statement that all arrows in this diagram are Poisson maps is literally true.

Second Stage Reduction.   To carry out the second stage reduction, that is, reduction by Iso^+, one has to overcome additional difficulties. The action of Iso^+ on $\mathfrak{X}^s_{\mathrm{div}}(M)$ induced by L, given by

$$
l : \mathrm{Iso}^+ \times \mathfrak{X}^s_{\mathrm{div}}(M) \longrightarrow \mathfrak{X}^s_{\mathrm{div}}(M), \ \ l_i(u) := l^u(i) := i_* u,
$$

has two major problems. First, it is not free as there typically will be vector fields that are symmetric (that is, invariant under the action of Iso^+). Second, it is not smooth as a map defined on $\mathrm{Iso}^+ \times \mathfrak{X}^s_{\mathrm{div}}(M)$, so the usual slice theorem for smooth proper finite dimensional Lie group actions is not valid. Recall that the slice theorem is a fundamental tool in the study of the geometric properties of the isotropy type submanifolds and of the orbit set. However, l retains some smoothness properties:

(1) l is a continuous map,

(2) for all $i \in \text{Iso}^+$, $l_i : \mathfrak{X}^s_{\text{div}}(M) \longrightarrow \mathfrak{X}^s_{\text{div}}(M)$ is a smooth map,

(3) for all $u \in \mathfrak{X}^s_{\text{div}}(M)$, $l^u : \text{Iso}^+ \longrightarrow \mathfrak{X}^{s-1}_{\text{div}}(M)$ is of class C^1.

Note that l^u takes value in $\mathfrak{X}^s_{\text{div}}(M)$ but is differentiable only as a map with values in $\mathfrak{X}^{s-1}_{\text{div}}(M)$. Indeed, computing the infinitesimal generator associated to $\xi \in \mathfrak{iso}$, we find

$$\xi_{\mathfrak{X}^s_{\text{div}}(M)}(u) = \left.\frac{d}{dt}\right|_{t=0} l_{\exp(t\xi)}(u) = \left.\frac{d}{dt}\right|_{t=0} \exp(t\xi)_* u = -\boldsymbol{\mathcal{L}}_{X_\xi} u = [u, X_\xi],$$

where X_ξ is the Killing vector field generated by the flow $\exp(t\xi)$ and $[\,,\,]$ is the Jacobi-Lie bracket of vector fields. Thus, for $u \in \mathfrak{X}^s_{\text{div}}(M)$, we have $\xi_{\mathfrak{X}^s_{\text{div}}(M)}(u) = [u, X_\xi] \in \mathfrak{X}^{s-1}_{\text{div}}(M)$, and there is a one derivative loss due to the Jacobi-Lie bracket.

Gay-Balmaz and Ratiu [2006] resolve these difficulties by presenting a slice theorem for any action Φ of a finite dimensional Lie group G on a chain of Banach manifolds $(Q^s)_{s>s_0}$, verifying properties $(1)-(3)$, with $\mathfrak{X}^s_{\text{div}}(M)$ replaced by Q^s. Using this slice theorem, it is then possible to study the geometric properties of the sets $(Q^s)^H$, $(Q^s)_H$, $(Q^s)_{(H)}$, and $Q^s_H/N(H)$ for $H \subset G$ a closed subgroup.

Note that even if the action Φ is free, the orbit set Q^s/G need not be a smooth manifold. To see why this is so, we argue by contradiction. If this were the case, then $\pi : Q^s \to Q^s/G$ would be a principal bundle and we could then take the tangent map $T_q\pi : T_qQ^s \to T_{\pi(q)}(Q^s/G)$ whose kernel is the vertical subspace $V_qQ^s = \ker(T_q\pi)$. We know that the vertical subspace is generated by the infinitesimal generators at q, that is, $V_qQ^s = \{\xi_{Q^s}(q) \mid \xi \in \mathfrak{g}\}$. This is a contradiction because, in general, $\xi_{Q^s}(q) \notin T_qQ^s$ (recall that by property (3) we only know that $\xi_{Q^s}(q) \in T_qQ^{s-1}$).

Since Q^s/G is not a differentiable manifold, the tangent bundle and differentiable curves cannot be defined in the usual way. To overcome this difficulty, one uses the fact that Q^s/G is a topological manifold and that changes of charts (which are not differentiable in the usual sense) are not only homeomorphisms but are also differentiable relative to a weaker topology (the topology of Q^{s-1}). This agrees with the fact that the vertical subbundle is a subset of $TQ^{s-1}|Q^s$, and allows the definition of the notion of weak-tangent bundle, weakly-differentiable curves and weakly-differentiable functions on the orbit space Q^s/G. Roughly speaking, the weakly-differentiable curves in Q^s/G are of the form $\pi \circ c$, where $c \in C^0(I, Q^s) \cap C^1(I, Q^{s-1})$, and the weak-tangent bundle of Q^s/G is defined by taking equivalence classes of such curves tangent at one point (see Gay-Balmaz and Ratiu [2006] for details).

Applying these results to the action l one concludes that $\mathfrak{X}^s_{\text{div}}(M)_H$ is an open set in the H-fixed point subspace $\mathfrak{X}^s_{\text{div}}(M)^H$ of $\mathfrak{X}^s_{\text{div}}(M)$ and

therefore, one can introduce the reduced Poisson bracket for a class of weakly-differentiable functions defined on $\mathfrak{X}_{\mathrm{div}}^s(M)_H/N(H)$. This allows one to carry out in a precise sense the second stage Poisson reduction procedure. As expected, the Euler equations preserve the H-isotropy submanifold $\mathfrak{X}_{\mathrm{div}}^s(M)_H$ and induce a flow $\widetilde{F}_t^H$ on the nonsingular quotient space $\mathfrak{X}_{\mathrm{div}}^s(M)_H/N(H)$. Finally, the following diagram commutes

$$
\begin{array}{ccc}
T\mathcal{D}_\mu^s(M) & \xrightarrow{\ \ F_t\ \ } & T\mathcal{D}_\mu^s(M) \\[4pt]
\pi_R \downarrow & & \downarrow \pi_R \\[4pt]
\mathfrak{X}_{\mathrm{div}}^s(M) & \xrightarrow{\ \ \widetilde{F}_t\ \ } & \mathfrak{X}_{\mathrm{div}}^s(M) \\[4pt]
i_H \uparrow & & \uparrow i_H \\[4pt]
\mathfrak{X}_{\mathrm{div}}^s(M)_H & \xrightarrow{\ \ \widetilde{F}_t^H\ \ } & \mathfrak{X}_{\mathrm{div}}^s(M)_H \\[4pt]
\pi_H \downarrow & & \downarrow \pi_H \\[4pt]
\mathfrak{X}_{\mathrm{div}}^s(M)_H/N(H) & \xrightarrow{\ \ \widetilde{\widetilde{F}_t^H}\ \ } & \mathfrak{X}_{\mathrm{div}}^s(M)_H/N(H)
\end{array}
$$

Here, $i_H : \mathfrak{X}_{\mathrm{div}}^s(M)_H \longrightarrow \mathfrak{X}_{\mathrm{div}}^s(M)_H$ is the smooth inclusion and $\pi_H : \mathfrak{X}_{\mathrm{div}}^s(M)_H \longrightarrow \mathfrak{X}_{\mathrm{div}}^s(M)_H/N(H)$ is the orbit projection. In this diagram all maps are Poisson with respect to a certain class of functions.

The remarkable fact is that the loss of derivative in the solutions of the Euler equations (recall that $u \in C^0(I, \mathfrak{X}_{\mathrm{div}}^s(M)) \cap C^1(I, \mathfrak{X}_{\mathrm{div}}^{s-1}(M))$) and the loss of derivative in property (3) of the action are compatible since the curves in $C^0(I, \mathfrak{X}_{\mathrm{div}}^s(M)) \cap C^1(I, \mathfrak{X}_{\mathrm{div}}^{s-1}(M))$ are exactly those needed to define the weakly differentiable structure on the orbit space.

10

Stages and Semidirect Products with Cocycles

This chapter is concerned with two major themes. The first theme, which is presented in §10.1 and §10.2, deals with the semidirect product M of a group G with an Abelian group A, where the construction of the semidirect product itself involves an A-valued cocycle of G. In this context, A (which is N in the general theory) is still a normal subgroup and the reduction by stages program is fully carried out. In particular, the stages hypothesis holds and so the reduction by stages program for the action of M on a symplectic manifold can be implemented. We focus on the case of the action of M on T^*M (by the lift of right translation) so that the final reduced spaces will be the coadjoint orbits of M.

This generalizes two things that were done previously. The first generalization is of standard semidirect product theory, namely the results in §4.3. In particular, that theory deals with the reduction of a semidirect product $G \circledS V$ acting on its own cotangent bundle, which then gives the structure of the coadjoint orbits of the semidirect product. In the context of this chapter, one arrives at this case by taking A to be a vector space V and the cocycle to be trivial. The second generalization is that of central extensions, which was treated in §8.2. Recall that for central extensions we also considered the structure of coadjoint orbits, such as those of the Bott–Virasoro group. In this case of central extensions, A is the real line and there is a cocycle.

The second theme, which is presented in §10.5 and §10.6, is the case in which M is the semidirect product $G \circledS H$ of two Lie groups, both of which can be nonabelian. Here, G acts on H by group automorphisms and the semidirect product is formed in a way that is analogous to the case $G \circledS V$

with no cocycles. Again, H is a normal subgroup and we carry out the reduction by stages program. In particular, we give a characterization of the coadjoint orbits of $G \circledS H$. In the course of doing this, some needed results that are also of independent interest are obtained. First of all, in §10.3, we develop a useful characterization of the reduction of a Poisson manifold of the form $T^*G \times P$, by a diagonal G-action. Secondly, in §10.4 we verify that Poisson reduction by stages gives the semidirect Poisson structure on $\mathfrak{g} \circledS \mathfrak{h}$, as it should.

The last section §10.7 uses this theory to find and classify all the coadjoint orbits of the group $\mathcal{T} \circledS \mathcal{U}$, where $\mathcal{T}$ is the group of upper triangular 3×3 matrices and where $\mathcal{U}$ is its normal subgroup having ones on the diagonal. The action of $\mathcal{T}$ on $\mathcal{U}$ is by conjugation. This example is interesting because it uses the full power of the theory and has several peculiarities, such as the lack of a sufficient number of Casimirs to distinguish the generic orbits.

10.1   Abelian Semidirect Product Extensions: First Reduction

This section carries out the first stage reduction of an Abelian semidirect product extension of a Lie group acting by cotangent lift by right translation on its cotangent bundle. The Abelian group used to construct the extension is the normal subgroup, so the first stage reduction is by this subgroup. To carry this out, we compute the mechanical connection and its curvature and we analyze their invariance properties. Since a cocycle is involved, the general reduction strategy is modeled after that used in the case of the first stage reduction of central extensions in §8.1.

Preliminaries.   We begin by recalling from §6.1 several relevant definitions and formulas used in this section. Given are a Lie group G (whose composition law is denoted multiplicatively), an Abelian Lie group A (whose composition law is denoted additively), a smooth left action $\phi : G \times A \to A$ by Lie group automorphisms of A, and a smooth normalized group two-cocycle $B : G \times G \to A$. Recall that having a normalized group two-cocycle means that the following identities hold for any $f, g, h \in G$:

$$\phi(f)(B(g,h)) + B(f, gh) = B(f,g) + B(fg, h) \tag{10.1.1}$$

and

$$B(g, e) = B(e, g) = 0.$$

Then the Lie group extension $G \times_{\phi, B} A$ of G by A is defined to be the Lie group whose underlying manifold is the product $G \times A$ and whose multiplication law is given by

$$(g, \alpha)(h, \beta) = (gh, \alpha + \phi(g)(\beta) + B(g, h)), \tag{10.1.2}$$

for any $(g, \alpha), (h, \beta) \in G \times A$. The neutral element is $(e, 0)$, and the inverse of (g, α) has the expression

$$(g, \alpha)^{-1} = \left(g^{-1}, -\phi(g^{-1})(B(g, g^{-1}) + \alpha)\right)$$
$$= \left(g^{-1}, -B(g^{-1}, g) - \phi(g^{-1})(\alpha)\right).$$

The subgroup $\{e\} \times A$ is Abelian, normal, and closed in $G \times_{\phi, B} A$ and the sequence

$$\{0\} \to A \to G \times_{\phi, B} A \to G \to \{e\} \tag{10.1.3}$$

is exact.

The key right principal A-bundle that will be used in the first stage reduction is $G \times_{\phi, B} A \to G$ given by $(g, \alpha) \mapsto g$. Equivalently, we can identify the quotient $(G \times_{\phi, B} A)/A$ with G by identifying the class $[g, \alpha]_A := (g, \alpha)(\{e\} \times A) \in (G \times_{\phi, B} A)/A$ with $g \in G$. This identification is a Lie group isomorphism whose inverse is given by $g \mapsto [g, 0]_A$. Note, however, that we do not attempt to realize G as a subgroup of $G \times_{\phi, B} A$.

Since the G-action on A is by group homomorphisms, G also acts on the left on the Abelian Lie algebra $\mathfrak{a}$ of A by $g \cdot v := T_0 \phi(g)(v)$, where $g \in G$ and $v \in \mathfrak{a}$, and defines therefore a left G-representation on $\mathfrak{a}$. The induced left Lie algebra representation $\widetilde{\phi} : \mathfrak{g} \to \mathfrak{gl}(\mathfrak{a})$ of the Lie algebra $\mathfrak{g}$ of G on $\mathfrak{a}$ is hence given by

$$\widetilde{\phi}(\xi)(v) := \xi \cdot v := \left. \frac{d}{dt}\right|_{t=0} \exp(t\xi) \cdot v$$

$$= \left. \frac{d}{dt}\right|_{t=0} T_0 \phi(\exp(t\xi))(v) = \xi_{\mathfrak{a}}(v), \tag{10.1.4}$$

where $\xi_{\mathfrak{a}} : \mathfrak{a} \to \mathfrak{a}$ denotes the infinitesimal generator of the G-action on $\mathfrak{a}$ defined by $\xi \in \mathfrak{g}$.

As in §6.1 (see equation (6.1.34)), the **Lie algebra two-cocycle** $C : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{a}$ associated with B is given by the formula

$$C(\xi, \eta) := \left. \frac{d^2}{ds\,dt}\right|_{t=s=0} (B(g(t), h(s)) - B(h(s), g(t)))$$

$$= D_1 D_2 B(e, e)(\xi)(\eta) - D_1 D_2 B(e, e)(\eta)(\xi),$$

for any smooth curves $t \mapsto g(t)$ and $s \mapsto h(s)$ through $e \in G$ with tangent vectors $\left.\frac{dg(t)}{dt}\right|_{t=0} = \xi$ and $\left.\frac{dh(s)}{ds}\right|_{s=0} = \eta \in \mathfrak{g}$. The **Lie algebra two-cocycle identity** states that

$$\xi_{\mathfrak{a}}(C(\eta, \zeta)) + \eta_{\mathfrak{a}}(C(\zeta, \xi)) + \zeta_{\mathfrak{a}}(C(\xi, \eta))$$
$$- C([\xi, \eta], \zeta) - C([\eta, \zeta], \xi) - C([\zeta, \xi], \eta) = 0 \tag{10.1.5}$$

for all $\xi, \eta, \zeta \in \mathfrak{g}$.

The extension $\mathfrak{g} \times_{\widetilde{\phi},C} \mathfrak{a}$ of $\mathfrak{g}$ by $\mathfrak{a}$ is defined as the Lie algebra with underlying vector space $\mathfrak{g} \times \mathfrak{a}$ whose bracket is given by

$$[(\xi, u), (\eta, v)] = ([\xi, \eta], \xi_{\mathfrak{a}}(v) - \eta_{\mathfrak{a}}(u) + C(\xi, \eta)), \qquad (10.1.6)$$

for any $(\xi, u), (\eta, v) \in \mathfrak{g} \times \mathfrak{a}$. Then the sequence of Lie algebras

$$0 \to \mathfrak{a} \to \mathfrak{g} \times_{\widetilde{\phi},C} \mathfrak{a} \to \mathfrak{g} \to 0,$$

corresponding to $(10.1.3)$ is exact, $\{0\} \times \mathfrak{a}$ is an Abelian ideal in $\mathfrak{g} \times_{\widetilde{\phi},C} \mathfrak{a}$, and $\mathfrak{g} \times_{\widetilde{\phi},C} \mathfrak{a}$ is the Lie algebra of $G \times_{\phi,B} A$.

Relation to the General Theory. *Throughout this chapter the Lie groups G and A are assumed to be connected.* We shall apply the general reduction by stages theory in §5.2 as follows. The symplectic manifold (P, Ω) is $T^*(G \times A)$ endowed with the canonical symplectic structure. The "big group" M is $G \times_{\phi,B} A$ and the normal subgroup N is $\{e\} \times A$. Thus, in the first reduction, $a := \nu \in \mathfrak{a}^*$ and, since A is Abelian, the coadjoint isotropy subgroup is $A_a = A$. Using $(6.1.48)$, the action of $G \times_{\phi,B} A$ on $\mathfrak{a}^*$ given by $(5.2.1)$ is

$$(g, \alpha) \cdot a = T_0^* \phi(g^{-1}) a.$$

Thus the isotropy subgroup $(G \times_{\phi,B} A)_a$ of this action at $a \in \mathfrak{a}^*$ equals $G_a \times A$, where

$$G_a := \{g \in G \mid T_0^* \phi(g^{-1}) a = a\}$$

is the isotropy subgroup of the G-representation on $\mathfrak{a}^*$. Therefore, the group M_ν/N_ν in the general theory is in this case $(G \times_{\phi,B} A)_a / A_a = G_a$. Note that the Lie algebra $\mathfrak{g}_a$ of G_a is given by

$$\mathfrak{g}_a = \{\xi \in \mathfrak{g} \mid \widetilde{\phi}(\xi)^* a = 0\}.$$

The Mechanical Connection on Group Extensions. Recall that to carry out cotangent bundle reduction, as described in §2.2 and §2.3, one needs to have a connection on $G \times_{\phi,B} A \to G$. As in §8.1, we shall use the mechanical connection associated via $(2.1.4)$ to a particular group invariant metric.

Pick a (not necessarily invariant) inner product $\langle\!\langle \cdot, \cdot \rangle\!\rangle_{\mathfrak{g}}$ on the Lie algebra $\mathfrak{g}$ and another one $\langle\!\langle \cdot, \cdot \rangle\!\rangle_{\mathfrak{a}}$ on $\mathfrak{a}$ and define the positive definite inner product on $\mathfrak{g} \times_{\widetilde{\phi},C} \mathfrak{a}$ by

$$\langle\!\langle (\xi, u), (\eta, v) \rangle\!\rangle_{\mathfrak{g} \times_{\widetilde{\phi},C} \mathfrak{a}} = \langle\!\langle \xi, \eta \rangle\!\rangle_{\mathfrak{g}} + \langle\!\langle u, v \rangle\!\rangle_{\mathfrak{a}}, \qquad (10.1.7)$$

for any (ξ, u) and $(\eta, v) \in \mathfrak{g} \times_{\widetilde{\phi},C} \mathfrak{a}$. Define a right invariant metric on $G \times_{\phi,B} A$ by

$$\langle\!\langle (X_g, u_\alpha), (Y_g, v_\alpha) \rangle\!\rangle_{(g,\alpha)}$$
$$= \langle\!\langle T_{(g,\alpha)} R_{(g,\alpha)^{-1}}(X_g, u_\alpha), T_{(g,\alpha)} R_{(g,\alpha)^{-1}}(Y_g, v_\alpha) \rangle\!\rangle_{\mathfrak{g} \times_{\widetilde{\phi},C} \mathfrak{a}}, \qquad (10.1.8)$$

where (X_g, u_α), $(Y_g, v_\alpha) \in T_{(g,\alpha)}(G \times_{\phi,B} A)$. Note that the derivative of right translation has the expression

$$
\begin{aligned}
T_{(g,\alpha)}R_{(h,\beta)}(X_g, u_\alpha) = \big(&T_g R_h(X_g),\, T_\alpha \Lambda_{\phi(g)\beta + B(g,h)}(u_\alpha) \\
&+ T_{\phi(g)\beta}\Lambda_{\alpha+B(g,h)}T_g\phi^\beta(X_g) \\
&+ T_{B(g,h)}\Lambda_{\alpha+\phi(g)\beta}D_1 B(g,h)(X_g)\big), \quad (10.1.9)
\end{aligned}
$$

where $X_g \in T_g G$, $u_\alpha \in T_\alpha A$, $\Lambda_\gamma : A \to A$ is the translation on A by the element $\gamma \in A$, and $\phi^\beta : G \to A$ is the smooth map given by $\phi^\beta(g) := \phi(g)\beta$ for any $g \in G$ and $\beta \in A$.

By construction, the metric (10.1.8) is right A-invariant and therefore the mechanical connection defined in (2.1.4) is given by

$$
\mathcal{A}^1(g,\alpha)\,(X_g, u_\alpha) = \mathbb{I}(g,\alpha)^{-1}\big(\mathbf{J}_A\left(\langle\!\langle(X_g, u_\alpha)\,,\cdot\rangle\!\rangle\right)\big), \tag{10.1.10}
$$

where $\mathbf{J}_A : T^*(G \times A) \to \mathfrak{a}^*$ is the momentum map of the cotangent lift of right A-translation on $G \times_{\phi,B} A$ and $\mathbb{I}(g,\alpha) : \mathfrak{a} \to \mathfrak{a}^*$ is the value at (g,α) of the locked moment of inertia tensor $\mathbb{I}$ defined by the right invariant metric (10.1.8).

10.1.1 Theorem. *Let $\mathcal{A}^1$ be the mechanical connection (10.1.10) on the right trivial A-principal bundle $G \times_{\phi,B} A \to G$.*

(i) *Then*

$$
\mathcal{A}^1(g,\alpha)(X_g, u_\alpha) = \big(T_0\phi(g^{-1}) \circ \mathcal{P}_\mathfrak{a} \circ T_{(g,\alpha)}R_{(g,\alpha)^{-1}}\big)(X_g, u_\alpha) \tag{10.1.11}
$$

for any $g \in G$, $\alpha \in A$, $X_g \in T_g G$, and $u_\alpha \in T_\alpha A$, where $\mathcal{P}_\mathfrak{a} : \mathfrak{g} \times \mathfrak{a} \to \mathfrak{a}$ denotes the projection onto $\mathfrak{a}$. Explicitly, this formula reads

$$
\begin{aligned}
\mathcal{A}^1(g,\alpha)(X_g, u_\alpha) = T_0\phi(g^{-1})\Big(&T_\alpha \Lambda_{-\alpha}(u_\alpha) \\
&+ T_{-B(g,g^{-1})-\alpha}\Lambda_{\alpha+B(g,g^{-1})}T_g\phi^{-\phi(g^{-1})(B(g,g^{-1})+\alpha)}(X_g) \\
&+ T_{B(g,g^{-1})}\Lambda_{-B(g,g^{-1})}D_1 B(g,g^{-1})(X_g)\Big). \tag{10.1.12}
\end{aligned}
$$

(ii) *The behavior of $\mathcal{A}^1$ under right translations of $G \times_{\phi,B} A$ is given by the following relation*

$$
\big(R^*_{(h,\beta)}\mathcal{A}^1\big)(g,\alpha)\,(X_g, u_\alpha) = T_0\phi(h^{-1})\big[\mathcal{A}^1(g,\alpha)\,(X_g, u_\alpha)\big] \tag{10.1.13}
$$

for any $(g,\alpha), (h,\beta) \in G \times_{\phi,B} A$, $X_g \in T_g G$, and $u_\alpha \in T_\alpha A$.

(iii) *The value of the curvature $\mathrm{curv}_{\mathcal{A}^1}$ of $\mathcal{A}^1$ at $(g,\alpha) \in G \times_{\phi,B} A$ is*

$$
\mathrm{curv}_{\mathcal{A}^1}(g,\alpha)((X_g, u_\alpha), (Y_g, v_\alpha)) = T_0\phi(g^{-1})\big(C(\xi,\eta)\big), \tag{10.1.14}
$$

where $\xi := T_g R_{g^{-1}}(X_g), \eta := T_g R_{g^{-1}}(Y_g) \in \mathfrak{g}$, and $(X_g, u_\alpha), (Y_g, v_\alpha) \in$
$T_g G \times T_\alpha A$. The induced curvature form on the base (see (2.1.13)) is
an $\mathfrak{a}$-valued two-form $\mathcal{B}^1$ on G since the adjoint bundle of the trivial
right bundle $G \times_{\phi,B} A \to G$ is the trivial vector bundle $G \times \mathfrak{a} \to G$;
$\mathcal{B}^1$ is given for any $g \in G$ and $\xi, \eta \in \mathfrak{g}$ by

$$\mathcal{B}^1(g)\left(T_e R_g \xi, T_e R_g \eta\right) = T_0 \phi(g^{-1})\left(C(\xi, \eta)\right).$$

(iv) *Let $G_a = \{g \in G \mid T_0^*(g)a = a\}$ be the isotropy subgroup of $a \in \mathfrak{a}^*$.*
   The one-form $\langle a, \mathcal{A}^1 \rangle \in \Omega^1(G \times_{\phi,B} A)$ is invariant under the right
   translation of the subgroup $G_a \times A$ of $G \times_{\phi,B} A$ and the real-valued
   two-form $\langle a, \mathcal{B}^1 \rangle \in \Omega^2(G)$ is right G_a-invariant.

Proof.   (i) Let us apply Theorem 2.1.15 to the Lie group $G \times_{\phi,B} A$ and
the normal subgroup $N = \{e\} \times A$. We have the right principal bundle
$G \times_{\phi,B} A \to G$ and we compute the mechanical connection from equation
(2.1.15). By (6.1.9), we have $(g, \alpha)^{-1} = (g^{-1}, -\phi(g^{-1})(B(g, g^{-1}) + \alpha))$
so that, using the multiplication law (6.1.2), the action by conjugation of
$G \times_{\phi,B} A$ on the normal subgroup $\{e\} \times A$ is given by

$$(g, \alpha)(e, \beta)(g, \alpha)^{-1}$$
$$= (g, \alpha + \phi(g)\beta)(g^{-1}, -\phi(g^{-1})(B(g, g^{-1}) + \alpha))$$
$$= (e, \alpha + \phi(g)\beta + \phi(g)(-\phi(g^{-1})(B(g, g^{-1}) + \alpha) + B(g, g^{-1})))$$
$$= (e, \phi(g)\beta).$$

Therefore, the derivative of this action relative to β at zero in the direction
$v \in \mathfrak{a}$ is
$$\mathrm{Ad}^A_{(g,\alpha)} v = T_0 \phi(g)v.$$

Formula (2.1.15) for the connection $\mathcal{A}^1$ yields

$$\mathcal{A}^1(g, \alpha)(X_g, u_\alpha) = \left(T_0 \phi(g^{-1}) \circ \mathcal{P}_\mathfrak{a} \circ \theta^R\right)(X_g, u_\alpha)$$
$$= T_0 \phi(g^{-1})\Big(T_\alpha \Lambda_{-\alpha}(u_\alpha)$$
$$\quad + T_{-B(g,g^{-1})-\alpha}\Lambda_{\alpha+B(g,g^{-1})}T_g\phi^{-\phi(g^{-1})(B(g,g^{-1})+\alpha)}(X_g)$$
$$\quad + T_{B(g,g^{-1})}\Lambda_{-B(g,g^{-1})}D_1 B(g, g^{-1})(X_g)\Big).$$

The first equality proves (10.1.11) and the second (10.1.12).

(ii) We use Theorem 2.1.15(ii) which in our case becomes

$$R^*_{(h,\beta)}\mathcal{A}^1 = \mathrm{Ad}^A_{(h,\beta)^{-1}} \circ \mathcal{A}^1.$$

Above we showed that $\mathrm{Ad}^A_{(h,\beta)} v = T_0 \phi(h)v$ and so the result follows.

(iii) The curvature of the connection $\mathcal{A}^1$ is computed using equation (2.1.17). First, note that relative to the inner product (10.1.7), the subspaces $\mathfrak{g} \times \{0\}$ and $\{0\} \times \mathfrak{a}$ are orthogonal complements to each other. Second, by (10.1.9) and (6.1.9) we have

$$
\begin{aligned}
(\xi, u) :&= T_{(g,\alpha)} R_{(g,\alpha)^{-1}}(X_g, u_\alpha) \\
&= \big(T_g R_{g^{-1}}(X_g), T_\alpha \Lambda_\alpha(u_\alpha) \\
&\quad + T_{-B(g,g^{-1})-\alpha} \Lambda_{\alpha+B(g,g^{-1})} T_g \phi^{-\phi(g^{-1})(B(g,g^{-1})+\alpha)}(X_g) \\
&\quad + T_{B(g,g^{-1})} \Lambda_{-B(g,g^{-1})} D_1 B(g, g^{-1})(X_g)\big)
\end{aligned}
$$

and a similar formula for $(\eta, v) := T_{(g,\alpha)} R_{(g,\alpha)^{-1}}(Y_g, v_\alpha)$. Third, for any $\xi \in \mathfrak{g}$ and $u, v \in \mathfrak{a}$, by (6.1.43) we have

$$
\mathrm{ad}_{(\xi,u)}(0, v) = (0, \xi_\mathfrak{a}(v)) .
$$

Therefore, since $\mathrm{Ad}^A_{(g,\alpha)^{-1}} v = T_0 \phi(g^{-1}) v$ for any $v \in \mathfrak{a}$ and $\mathfrak{a}$ is Abelian, formula (2.1.17) gives

$$
\begin{aligned}
\mathrm{curv}_{\mathcal{A}^1}&((X_g, u_\alpha), (Y_g, v_\alpha)) \\
&= T_0 \phi(g^{-1}) \left(-\xi_\mathfrak{a}(v) + \eta_\mathfrak{a}(u) + \xi_\mathfrak{a}(v) - \eta_\mathfrak{a}(u) + C(\xi, \eta)\right) \\
&= T_0 \phi(g^{-1}) \big(C(\xi, \eta)\big),
\end{aligned}
$$

which proves (10.1.14).

(iv) By (10.1.13) we have $R^*_{(h,\beta)} \langle a, \mathcal{A}^1 \rangle = \langle T_0^* \phi(h^{-1}) a, \mathcal{A}^1 \rangle$ for any $(h, \beta) \in G \times_{\phi, B} A$. Taking $h \in G_a$, this proves $(G_a \times A)$-invariance of $\langle a, \mathcal{A}^1 \rangle$. The G_a-right invariance of the ordinary two-form $\langle a, \mathcal{B}^1 \rangle$ on G is a direct consequence of Lemma 8.1.4 or can be easily verified directly from the expression of $\mathcal{B}^1$. ∎

Remark.   Note that if the extension in (10.1.3) is central then $\phi(g) = \mathrm{id}$ for every $g \in G$ and $\mathrm{Ad}_{\{e\} \times A}$ pointwise fixes the set $\mathfrak{g} \times \{0\}$. In this case, Theorem 10.1.1 states the following: first of all, $\mathcal{A}^1$ is the $\mathfrak{a}$-component of the right invariant Maurer–Cartan form on $G \times_{\phi, B} A$, second, it is right $G \times_{\phi, B} A$-invariant, and third, its curvature at the identity is the is the Lie algebra two-cocycle. This statement generalizes Theorem 8.1.1 to arbitrary dimensional central extensions of G.

10.1.2 Corollary.   *The momentum map $\mathbf{J}_A : T^*(G \times_{\phi, B} A) \to \mathfrak{a}^*$ for the cotangent lift of the right translation action of A on $G \times_{\phi, B} A$ is given by*

$$
\mathbf{J}_A(\gamma_g, a_\alpha) = (T_0 \Lambda_\alpha \circ T_0 \phi(g))^* a_\alpha, \tag{10.1.15}
$$

where $(\gamma_g, a_\alpha) \in T^*_{(g,\alpha)}(G \times_{\phi, B} A)$. For any $(h, \beta) \in G \times_{\phi, B} A$ we have

$$
\mathbf{J}_A \circ T^* R_{(h,\beta)^{-1}} = T_0^* \phi(h) \circ \mathbf{J}_A. \tag{10.1.16}
$$

In particular, $G \times_{\phi, B} A$ acts on $\mathbf{J}_A^{-1}(0)$ and the subgroup $G_a \times A \subset G \times_{\phi, B} A$ on $\mathbf{J}_A^{-1}(a)$, where the action is cotangent lift of right translation.

Proof.   We begin by computing the infinitesimal generator of the right translation by A on $G \times_{\phi, B} A$. According to (10.1.2), this right A-action is given by $(g, \alpha) \cdot \beta := (g, \alpha + \phi(g)(\beta))$ for any $g \in G$ and $\alpha, \beta \in A$. Therefore, if $u \in \mathfrak{a}$ we get

$$u_{G \times_{\phi, B} A}(g, \alpha) = \left.\frac{d}{dt}\right|_{t=0} (g, \alpha) \cdot \exp(tu) = \left.\frac{d}{dt}\right|_{t=0} (g, \alpha + \phi(g)(\exp(tu)))$$
$$= (0, (T_0 \Lambda_\alpha \circ T_0 \phi(g)) \, u) \,, \qquad\qquad (10.1.17)$$

where $\Lambda_\alpha : A \to A$ denotes the translation map by α in A.

For $(\gamma_g, a_\alpha) \in T^*_{(g, \alpha)}(G \times_{\phi, B} A)$ and $u \in \mathfrak{a}$, using (10.1.17), the momentum map $\mathbf{J}_A : T^*(G \times A) \to \mathfrak{a}^*$ is computed to be

$$\begin{aligned}
\langle \mathbf{J}_A(\gamma_g, a_\alpha), u \rangle &= \langle (\gamma_g, a_\alpha), u_{G \times_{\phi, B} A}(g, \alpha) \rangle \\
&= \langle (\gamma_g, a_\alpha), (0, (T_0 \Lambda_\alpha \circ T_0 \phi(g))u) \rangle \\
&= \langle a_\alpha, (T_0 \Lambda_\alpha \circ T_0 \phi(g)) \, u \rangle \\
&= \langle (T_0 \Lambda_\alpha \circ T_0 \phi(g))^* a_\alpha, u \rangle
\end{aligned}$$

which proves (10.1.15).

If $(g, \alpha), (h, \beta) \in G \times_{\phi, B} A$, $(\gamma_g, a_\alpha) \in T^*_{(g, \alpha)}(G \times_{\phi, B} A)$, and $u \in \mathfrak{a}$, we have

$$\begin{aligned}
&\left\langle \mathbf{J}_A \left(T^*_{(g, \alpha)(h, \beta)} R_{(h, \beta)^{-1}}(\gamma_g, a_\alpha) \right), u \right\rangle \\
&= \left\langle T^*_{(g, \alpha)(h, \beta)} R_{(h, \beta)^{-1}}(\gamma_g, a_\alpha), u_{G \times_{\phi, B} A}(gh, \alpha + \phi(g)\beta + B(g, h)) \right\rangle \\
&= \langle (\gamma_g, a_\alpha), T_{(gh, \alpha + \phi(g)\beta + B(g, h))} R_{(h^{-1}, -\phi(h^{-1})\beta - \phi(h^{-1})B(h, h^{-1}))} \\
&\qquad\qquad \left(0, (T_0 \Lambda_{\alpha + \phi(g)\beta + B(g, h)} \circ T_0 \phi(gh)) \, u \right) \rangle \\
&= \langle (\gamma_g, a_\alpha), (0, (T_{\alpha + \phi(g)\beta + B(g, h)} \Lambda_{\phi(gh)(-\phi(h^{-1})\beta - \phi(h^{-1})B(h, h^{-1})) + B(gh, h^{-1})} \\
&\qquad\qquad \circ T_0 \Lambda_{\alpha + \phi(g)\beta + B(g, h)} \circ T_0 \phi(gh)) \, u)) \rangle \\
&= \langle a_\alpha, (T_0 \Lambda_{\alpha - \phi(g)B(h, h^{-1}) + B(gh, h^{-1}) + B(g, h)} \circ T_0 \phi(gh)) \, u \rangle \qquad (10.1.18)
\end{aligned}$$

by (10.1.17) and (10.1.9). The element by which one translates in A is given by $\alpha - \phi(g)B(h, h^{-1}) + B(gh, h^{-1}) + B(g, h) = \alpha$ since, by the cocycle identity (10.1.1), we have

$$-\phi(g)B(h, h^{-1}) + B(gh, h^{-1}) + B(g, h) = B(g, e) = 0$$

(using the replacements $f \mapsto g, g \mapsto h, h \mapsto h^{-1}$ in (10.1.1)). Therefore, the expression (10.1.18) becomes

$$\begin{aligned}
\langle a_\alpha, (T_0 \Lambda_\alpha \circ T_0 \phi(g) \circ T_0 \phi(h)) \, u \rangle &= \langle (T_0 \Lambda_\alpha \circ T_0 \phi(g))^* a_\alpha, T_0 \phi(h) u \rangle \\
&= \langle \mathbf{J}_A(\gamma_g, a_\alpha), T_0 \phi(h) u \rangle \\
&= \langle T_0^* \phi(h) \mathbf{J}_A(\gamma_g, a_\alpha), u \rangle
\end{aligned}$$

which proves (10.1.16).   ∎

The First Reduction.   We will apply the reduction by stages theorem to group extensions, since, by §6.3, the stages hypothesis is always verified. The goal is to compute the reduced spaces for the action of $G \times_{\phi,B} A$ on its cotangent bundle, reducing first by the action of A and then by the appropriate subgroup of G. The resulting spaces are the coadjoint orbits of $G \times_{\phi,B} A$. We carry out below the first reduction.

10.1.3 Theorem.   *Consider the cotangent lift of right translation by elements of A on $T^*(G \times_{\phi,B} A)$. Let $a \in \mathfrak{a}^*$. Then there is a right G_a-equivariant symplectic diffeomorphism*

$$\mathbf{J}_A^{-1}(a)/A \simeq \left(T^*G, \Omega - \pi_G^* \left\langle a, \mathcal{B}^1 \right\rangle\right), \qquad (10.1.19)$$

where $\left\langle a, \mathcal{B}^1 \right\rangle \in \Omega^2(G)$ is the closed right G_a-invariant two-form (the magnetic term) given by $\left\langle a, \mathcal{B}^1 \right\rangle(g)(T_e R_g \xi, T_e R_g \eta) = \left\langle a, T_0\phi(g^{-1})\left(C(\xi,\eta)\right)\right\rangle$, where $\xi, \eta \in \mathfrak{g}$, and $\pi_G : T^*G \to G$ is the cotangent bundle projection.

Proof.   The general theory of cotangent bundle reduction for Abelian symmetry groups (see Remark 3 following Theorem 2.2.1, Theorem 2.2.3, and Proposition 2.2.5) guarantees that the spaces in (10.1.19) are symplectomorphic and that $\mathbf{d} \left\langle a, \mathcal{A}^1 \right\rangle \in \Omega^2(G \times A)$ drops to the quotient $(G \times_{\phi,B} A)/A = G$. In addition, Proposition 2.1.13 states, in this case, that this two-form on G induced by $\mathbf{d} \left\langle a, \mathcal{A}^1 \right\rangle \in \Omega^2(G \times A)$ coincides with $\left\langle a, \mathcal{B}^1 \right\rangle$ which is given, according to Theorem 10.1.1(iii), by $\left\langle a, \mathcal{B}^1 \right\rangle(g)(T_e R_g \xi, T_e R_g \eta) = \left\langle a, T_0\phi(g^{-1})\left(C(\xi,\eta)\right)\right\rangle$; this two-form is right G_a-invariant by Theorem 10.1.1(iv).

To prove right G_a-equivariance of the symplectomorphism (10.1.19) we proceed in several steps. First, recall from Theorem 10.1.1(iv) that $\left\langle a, \mathcal{A}^1 \right\rangle$ is a $(G_a \times V)$-invariant one-form on $G \times_{\phi,B} A$; that is, the following relation holds

$$\left\langle a, \mathcal{A}^1 \right\rangle (R_{(h,\beta)}(g,\alpha))T_{(g,\alpha)} R_{(h,\beta)}(X_g, u_\alpha) = \left\langle a, \mathcal{A}^1 \right\rangle (g,\alpha)(X_g, u_\alpha)$$
$$(10.1.20)$$

for any $h \in G_a$ (which means that $T_0^*\phi(h)a = a$), $\beta \in A$, $(g,\alpha) \in G \times_{\phi,B} A$, and $(X_g, u_\alpha) \in T_{(g,\alpha)}(G \times_{\phi,B} A)$.

Second, by Corollary 10.1.2, we know that the subgroup $G_a \times A$ acts on both level submanifolds $\mathbf{J}_A^{-1}(0)$ and $\mathbf{J}_A^{-1}(a)$. We shall show now that the map $\text{shift}_a : \mathbf{J}_A^{-1}(a) \to \mathbf{J}_A^{-1}(0)$ defined in (2.2.8) by

$$\text{shift}_a(\gamma_g, a_\alpha) := (\gamma_g, a_\alpha) - \left\langle a, \mathcal{A}^1(g,\alpha)\right\rangle$$

is $(G_a \times A)$-equivariant, which means that for any $(g,\alpha) \in G \times_{\phi,B} A$, $(h,\beta) \in G_a \times A$, and $(\gamma_g, a_\alpha) \in T^*_{(g,\alpha)}(G \times_{\phi,B} A)$, we have

$$\text{shift}_a(T^*_{(g,\alpha)(h,\beta)} R_{(h,\beta)^{-1}}(\gamma_g, a_\alpha)) = T^*_{(g,\alpha)(h,\beta)} R_{(h,\beta)^{-1}}\left(\text{shift}_a(\gamma_g, a_\alpha)\right).$$

This is equivalent to

$$\big\langle a, \mathcal{A}^1(gh, \alpha + \phi(g)\beta + B(g,h))\big(X_{gh}, u_{\alpha+\phi(g)\beta+B(g,h)}\big)\big\rangle$$
$$= \big\langle a, \mathcal{A}^1(g,\alpha)T_{(g,\alpha)(h,\beta)}R_{(h,\beta)^{-1}}\big(X_{gh}, u_{\alpha+\phi(g)\beta+B(g,h)}\big)\big\rangle$$

for any $(X_{gh}, u_{\alpha+\phi(g)\beta+B(g,h)}) \in T^*_{(g,\alpha)(h,\beta)}(G \times_{\phi,B} A)$ which is in turn equivalent to the true identity (10.1.20) by setting $(X_{gh}, u_{\alpha+\phi(g)\beta+B(g,h)}) = T_{(g,\alpha)}R_{(h,\beta)}(X_g, u_\alpha)$ for some $(X_g, u_\alpha) \in T_{(g,\alpha)}(G\times_{\phi,B} A)$. This shows that $\mathrm{shift}_a : \mathbf{J}_A^{-1}(a) \to \mathbf{J}_A^{-1}(0)$ is $(G_a \times A)$-equivariant.

Third, we check that the quotient map defined by shift_a, denoted by

$$\widetilde{\mathrm{shift}}_a : \mathbf{J}_A^{-1}(a)/A \to \mathbf{J}_A^{-1}(0)/A,$$

is equivariant with respect to the action of $G_a = (G_a \times A)/A$. From the diagram

that defines $\widetilde{\mathrm{shift}}_a$, the $(G_a \times A)$-equivariance of the shift map, identifying $[h, \beta]$ with h for $(h, \beta) \in G_a \times A$, and writing $[g, \alpha]_a = \pi_A^a(g, \alpha)$, we get

$$\widetilde{\mathrm{shift}}_a([g,\alpha]_a \cdot h) = \widetilde{\mathrm{shift}}_a([(g,\alpha)(h,0)]_a) = \pi_A^0(\mathrm{shift}_a((g,\alpha)(h,0)))$$
$$= \pi_A^0(\mathrm{shift}_a(g,\alpha) \cdot (h,0)) = \pi_A^0(\mathrm{shift}_a(g,\alpha)) \cdot h$$
$$= \widetilde{\mathrm{shift}}_a([g,\alpha]_a) \cdot h,$$

where we have used the fact that $[g,\alpha]_a \cdot h = [(g,\alpha)(h,0)]_a$ (and similarly for $a = 0$), which defines the action of G_a on $\mathbf{J}_A^{-1}(a)/A$ (and on $\mathbf{J}_A^{-1}(0)/A$ respectively). This proves the statement. ∎

Group Extensions with Vector Spaces. In the case when the Abelian group A is a vector space V, several of the formulas derived above simplify. Since this is an important case in applications, we summarize here the relevant changes.

Formula (10.1.12) for the connection becomes

$$\mathcal{A}^1(g,v)(X_g, u) = g^{-1} \cdot \Big(u - T_g\phi^{g^{-1}\cdot(B(g,g^{-1})+v)}(X_g) + D_1 B(g, g^{-1})(X_g)\Big)$$

and its equivariance property is $R^*_{(h,w)}\mathcal{A}^1 = h^{-1} \cdot \mathcal{A}^1$. The curvature formulas are

$$\begin{aligned}
\mathrm{curv}_{\mathcal{A}^1}&(g,v)\left((T_e R_g \xi, u), (T_e R_g \eta, w)\right) \\
&= \mathcal{B}^1(g)\left(T_e R_g \xi, T_e R_g \eta\right) = g^{-1} \cdot C(\xi, \eta).
\end{aligned}$$

The momentum map $\mathbf{J}_V : T^*(G \times_{\phi,B} V) \to V^*$ is given by

$$\mathbf{J}_V(\gamma_g, a) = g^{-1} \cdot a,$$

where $\left\langle g^{-1} \cdot a, u \right\rangle = \left\langle a, g \cdot u \right\rangle$, for any $g \in G$, $u \in V$, and $a \in V^*$, is the left contragredient representation of G on V^*. The invariance property of $\mathbf{J}_V$ relative to right translation reads in this case $\mathbf{J}_V \circ T^* R_{(h,w)^{-1}} = h^{-1} \cdot \mathbf{J}_V$.

If, in addition, $B = 0$, then the connection $\mathcal{A}^1$ is flat and all these formulas recover those in §4.3 for the stage one reduction. For example, the formula for the connection given above is identical to that in (4.3.7). The reduced space $\mathbf{J}_V^{-1}(a)/V$ is symplectomorphic in this case to T^*G endowed with the canonical symplectic structure, recovering the result in (4.3.8).

10.2   Abelian Semidirect Product Extensions: Coadjoint Orbits

In this section we carry out the second stage reduction of $T^*(G \times_{\phi,B} A)$ by the cotangent lift of right translation of $G \times_{\phi,B} A$. This gives a description of the coadjoint orbits of $G \times_{\phi,B} A$. As in the case of central extensions, we obtain curvature interpretations for the symplectic forms on the coadjoint orbits. The symplectic reduction by stages for semidirect product extensions is a synthesis of the pure semidirect product case and the pure central extension case.

We already know from §10.1 that the first reduced space $\mathbf{J}_A^{-1}(a)/A$, for $a \in \mathfrak{a}^*$, is right G_a-equivariantly symplectomorphic to the cotangent bundle $\left(T^*G, \Omega - \pi_G^* \left\langle a, \mathcal{B}^1 \right\rangle\right)$. We also have seen that the closed subgroup $G_a = \{g \in G \mid T_0^* \phi(g) a = a\} \subset G$ acts by cotangent lift of right translation on T^*G and leaves the magnetic term $\pi_G^* \left\langle a, \mathcal{B}^1 \right\rangle$ invariant. The magnetic term is the remnant of the group two-cocycle and the residual group which acts in the second stage (namely M_ν/N_ν in the notation of the general theory) is G_a, the remnant of the semidirect product.

The goal of this section then is to implement the point reduction of the symplectic manifold $\left(T^*G, \Omega - \pi_G^* \left\langle a, \mathcal{B}^1 \right\rangle\right)$ under the cotangent bundle lifted right action of the Lie subgroup G_a. As we saw in §2.2, §2.3, 7.1, 7.2, there are two versions of cotangent bundle reduction, an embedding version and a bundle version. We shall explore the application of both of these versions.

The Problem Setting.   We shall apply Theorem 7.1.6 to the right G_a-action on $\left(T^*G, \Omega - \pi_G^* \left\langle a, \mathcal{B}^1 \right\rangle\right)$. The theorem is applied to the following objects:

1. We take Q to be G.

2. The symbol Ω is the canonical symplectic form on $T^*Q = T^*G$.

3. We take the magnetic term $\mathcal{B}$ to be $\left\langle a, \mathcal{B}^1 \right\rangle$, which is a closed right G_a-invariant two-form by Theorem 10.1.1(iv).

4. We take the connection $\mathcal{A}$ on the right principal bundle $\pi_{Q,G} : Q \to Q/G$, which in our case is the right principal G_a-bundle $\pi_{G,G_a} : G \to G/G_a$, to be the mechanical connection associated to an inner product $\langle\!\langle \cdot, \cdot \rangle\!\rangle$ on $\mathfrak{g}$. To construct it explicitly, denote by the same symbol the right invariant metric induced by $\langle\!\langle \cdot, \cdot \rangle\!\rangle$ on G, that is,

$$\langle\!\langle u_g, v_g \rangle\!\rangle_g := \left\langle\!\left\langle T_g R_{g^{-1}} u_g, T_g R_{g^{-1}} v_g \right\rangle\!\right\rangle$$

for any $u_g, v_g \in T_g G$. The vertical bundle $V \subset TG$ has the fiber at $g \in G$ equal to $V_g = \{T_e L_g \xi \mid \xi \in \mathfrak{g}_a\}$ and the horizontal bundle $H \subset TG$ is defined as the orthogonal complement of the vertical one, that is, $H := V^\perp$. Therefore,

$$H_g = \left\{ v_g \mid \left\langle\!\left\langle T_g R_{g^{-1}} v_g, \mathrm{Ad}_g \xi \right\rangle\!\right\rangle = 0 \text{ for all } \xi \in \mathfrak{g}_a \right\} = \left[T_e L_g \left(\mathfrak{g}_a \right) \right]^\perp.$$

By construction, $TG = V \oplus H$ and H is right G_a-invariant, that is, $T_g R_h(H_g) = H_{gh}$ for all $g \in G$ and $h \in G_a$. Therefore, the choice of this horizontal bundle defines a mechanical connection on the right principal G_a-bundle $\pi_{G,G_a} : G \to G/G_a$.

The Mechanical Connection One-Form $\mathcal{A}^{G_a} \in \Omega^1(G; \mathfrak{g}_a)$.   To compute $\mathcal{A}^{G_a}$ we use (2.1.4). Using the right invariance of the metric $\langle\!\langle \cdot, \cdot \rangle\!\rangle$ on G, the locked inertia tensor $\mathbb{I}(g) : \mathfrak{g}_a \to \mathfrak{g}_a^*$ defined by (2.1.2) becomes in this case for $\eta, \zeta \in \mathfrak{g}_a$

$$\langle \mathbb{I}(g)\eta, \zeta \rangle = \langle\!\langle T_e L_g \eta, T_e L_g \zeta \rangle\!\rangle_g = \langle\!\langle \mathrm{Ad}_g \eta, \mathrm{Ad}_g \zeta \rangle\!\rangle = \left\langle\!\left\langle \mathrm{Ad}_g^T \mathrm{Ad}_g \eta, \zeta \right\rangle\!\right\rangle$$

$$= \left\langle\!\left\langle \mathcal{P}_{\mathfrak{g}_a} \mathrm{Ad}_g^T \mathrm{Ad}_g \eta, \zeta \right\rangle\!\right\rangle = \left\langle \left(\mathbb{I}(e) \circ \mathcal{P}_{\mathfrak{g}_a} \circ \mathrm{Ad}_g^T \circ \mathrm{Ad}_g \right) (\eta), \zeta \right\rangle, \quad (10.2.1)$$

where $\mathbb{I}(e) : \mathfrak{g}_a \to \mathfrak{g}_a^*$ is defined by $\langle \mathbb{I}(e)\eta, \zeta \rangle := \langle\!\langle \eta, \zeta \rangle\!\rangle$, $\mathcal{P}_{\mathfrak{g}_a} : \mathfrak{g} \to \mathfrak{g}_a$ is the projection onto $\mathfrak{g}_a$ defined by the orthogonal splitting $\mathfrak{g} = \mathfrak{g}_a \oplus \mathfrak{g}_a^\perp$, and $\mathrm{Ad}_g^T : \mathfrak{g} \to \mathfrak{g}$ is the $\langle\!\langle \cdot, \cdot \rangle\!\rangle$-adjoint of Ad_g for any $g \in G$. Now note that the linear map $\mathbb{I}(g) : \mathfrak{g}_a \to \mathfrak{g}_a^*$ is invertible since $\langle\!\langle \cdot, \cdot \rangle\!\rangle$ is nondegenerate. In addition, it is easy to see that $\mathcal{P}_{\mathfrak{g}_a} \circ \mathrm{Ad}_g^T \circ \mathrm{Ad}_g : \mathfrak{g}_a \to \mathfrak{g}_a$ is invertible and hence we get

$$\mathbb{I}(g)^{-1} = \left(\mathcal{P}_{\mathfrak{g}_a} \circ \mathrm{Ad}_g^T \circ \mathrm{Ad}_g \right)^{-1} \circ \mathbb{I}(e)^{-1} : \mathfrak{g}_a^* \to \mathfrak{g}_a. \quad (10.2.2)$$

Note also that

$$\mathbb{I}(gh) = \mathrm{Ad}_h^* \circ \mathbb{I}(g) \circ \mathrm{Ad}_h$$

for all $g \in G$ and $h \in G_a$.

If L is the Lagrangian on TG associated to the right-invariant metric $\langle\!\langle \cdot, \cdot \rangle\!\rangle$ on G, its fiber derivative $\mathbb{F}L : TG \to T^*G$ is given by $\mathbb{F}L(v_g) = \langle\!\langle v_g, \cdot \rangle\!\rangle$. If $\mathbf{J} : T^*G \to \mathfrak{g}_a^*$ is the momentum map of the right G_a-action, then $\mathbf{J}(\alpha_g) = T_e^* L_g(\alpha_g)|_{\mathfrak{g}_a}$. Therefore, for any $\xi \in \mathfrak{g}_a$, right invariance of the metric gives,

$$\langle \mathbf{J}\left(\mathbb{F}L(v_g)\right), \xi \rangle = \langle\!\langle v_g, T_e L_g \xi \rangle\!\rangle_g = \langle\!\langle T_g R_{g^{-1}} v_g, \mathrm{Ad}_g \xi \rangle\!\rangle$$

$$= \langle\!\langle \mathcal{P}_{\mathfrak{g}_a} \mathrm{Ad}_g^T T_g R_{g^{-1}} v_g, \xi \rangle\!\rangle = \left\langle \left(\mathbb{I}(e) \circ \mathcal{P}_{\mathfrak{g}_a} \circ \mathrm{Ad}_g^T \circ T_g R_{g^{-1}}\right)(v_g), \xi \right\rangle,$$

that is,

$$\mathbf{J}\left(\mathbb{F}L(v_g)\right) = \left(\mathbb{I}(e) \circ \mathcal{P}_{\mathfrak{g}_a} \circ \mathrm{Ad}_g^T \circ T_g R_{g^{-1}}\right)(v_g). \tag{10.2.3}$$

The G_a-equivariance of $\mathbf{J} : T^*G \to \mathfrak{g}_a^*$ means that

$$\mathbf{J}\left(\mathbb{F}L(T_g R_h(v_g)\right) = \mathrm{Ad}_h^* \mathbf{J}\left(\mathbb{F}L(v_g)\right)$$

for any $v_g \in T_g G$ and $h \in G_a$.

Therefore, by (10.2.2) and (10.2.3), the mechanical connection one-form $\mathcal{A}^{G_a}$ defined by (2.1.4) becomes

$$\mathcal{A}^{G_a}(g)(v_g) = \left(\left(\mathcal{P}_{\mathfrak{g}_a} \circ \mathrm{Ad}_g^T \circ \mathrm{Ad}_g\right)^{-1} \circ \mathcal{P}_{\mathfrak{g}_a} \circ \mathrm{Ad}_g^T \circ T_g R_{g^{-1}}\right)(v_g). \tag{10.2.4}$$

Using the equivariance properties of $\mathbb{I}$ and $\mathbf{J}$ mentioned above it is easy to verify that $\mathcal{A}^{G_a}$ satisfies the axioms of a right connection one-form, as expected. This means that $\mathcal{A}^{G_a}(g)\left(T_e L_g \xi\right) = \xi$ for any $g \in G$, $\xi \in \mathfrak{g}_a$, and

$$\mathcal{A}^{G_a}(gh)\left(T_g R_h v_g\right) = \mathrm{Ad}_{h^{-1}}\left(\mathcal{A}^{G_a}(g)\left(v_g\right)\right)$$

for any $g \in G$, $v_g \in T_g G$, and $h \in G_a$.

A Formula for the $\mathcal{B}\mathfrak{g}$-Potential. We must find a smooth map $\psi^a : G \to \mathfrak{g}_a^*$ that satisfies

$$\mathbf{i}_{\xi_G} \left\langle a, \mathcal{B}^1 \right\rangle = \mathbf{d} \left\langle \psi^a, \xi \right\rangle \tag{10.2.5}$$

for all $\xi \in \mathfrak{g}_a$. Note that we write ψ^a instead of ϕ (as was the case in Theorem 7.1.1 where the notion of $\mathcal{B}\mathfrak{g}$-potential was introduced) because the determining equation depends on $a \in \mathfrak{a}^*$.

10.2.1 Lemma. *A $\mathcal{B}\mathfrak{g}$-potential $\psi^a : G \to \mathfrak{g}_a^*$ is given by*

$$\langle \psi^a(g), \xi \rangle = - \Big\langle a, T_0 \phi(g^{-1})\Big(D_2 B(g, e)(\xi)$$

$$+ T_{B(g,g^{-1})} \Lambda_{-B(g,g^{-1})} \left(D_1 B(g, g^{-1})(T_e L_g \xi)\right)\Big)\Big\rangle \tag{10.2.6}$$

for all $g \in G$ and $\xi \in \mathfrak{g}_a$.

Proof. Let $\pi_1 : G \times_{\phi,B} A \to G$, $\pi_1(g, \alpha) := g$, be the projection onto the first factor. Since this is a surjective submersion, it suffices to show that

$$\pi_1^* \mathbf{i}_{\xi_G} \left\langle a, \mathcal{B}^1 \right\rangle = \pi_1^* \mathbf{d} \left\langle \psi^a, \xi \right\rangle. \qquad (10.2.7)$$

Let us first compute the infinitesimal generator for $(\xi, 0)_{G \times_{\phi,B} A}$ for the right action of $G \times_{\phi,B} A$ on itself. By definition, it is given by

$$
\begin{aligned}
(\xi, 0)_{G \times_{\phi,B} A}(g, \alpha) &= \left. \frac{d}{dt} \right|_{t=0} (g, \alpha)(\exp(t\xi), 0) \\
&= \left. \frac{d}{dt} \right|_{t=0} (g \exp(t\xi), \alpha + B(g, \exp(t\xi))) \\
&= (T_e L_g \xi, T_0 \Lambda_\alpha (D_2 B(g, e)(\xi))).
\end{aligned}
$$

Recalling that $\xi_G(g) = T_e L_g \xi$ (since we are using *right* actions), note that $(\xi, 0)_{G \times_{\phi,B} A}$ and ξ_G are π_1-related. In addition, $\pi_1^* \left\langle a, \mathcal{B}^1 \right\rangle = \mathbf{d} \left\langle a, \mathcal{A}^1 \right\rangle$ by Theorem 10.1.3. Hence, the left hand side of (10.2.7) becomes

$$
\begin{aligned}
\pi_1^* \mathbf{i}_{\xi_G} \left\langle a, \mathcal{B}^1 \right\rangle &= \mathbf{i}_{(\xi,0)_{G \times_{\phi,B} A}} \pi_1^* \left\langle a, \mathcal{B}^1 \right\rangle \\
&= \mathbf{i}_{(\xi,0)_{G \times_{\phi,B} A}} \mathbf{d} \left\langle a, \mathcal{A}^1 \right\rangle \\
&= -\mathbf{d} \mathbf{i}_{(\xi,0)_{G \times_{\phi,B} A}} \left\langle a, \mathcal{A}^1 \right\rangle,
\end{aligned}
$$

where we have used the identity $\boldsymbol{\mathcal{L}}_{(\xi,0)_{G \times_{\phi,B} A}} \left\langle a, \mathcal{A}^1 \right\rangle = 0$, which is valid since the one-form $\left\langle a, \mathcal{A}^1 \right\rangle$ is $(G_a \times A)$-invariant by Theorem 10.1.1(iv).

Thus, equation (10.2.7) reduces to

$$- \mathbf{d} \mathbf{i}_{(\xi,0)_{G \times_{\phi,B} A}} \left\langle a, \mathcal{A}^1 \right\rangle = \mathbf{d} \pi_1^* \left\langle \psi^a, \xi \right\rangle. \qquad (10.2.8)$$

This will hold provided that

$$\pi_1^* \left\langle \psi^a, \xi \right\rangle = -\mathbf{i}_{(\xi,0)_{G \times_{\phi,B} A}} \left\langle a, \mathcal{A}^1 \right\rangle. \qquad (10.2.9)$$

From equation (10.1.12) for $\mathcal{A}^1$, we can compute the right hand side of (10.2.9) as follows:

$$
\begin{aligned}
-\mathbf{i}_{(\xi,0)_{G \times_{\phi,B} A}} \left\langle a, \mathcal{A}^1 \right\rangle (g, \alpha) &= - \left\langle a, \mathcal{A}^1(g, \alpha) \big(T_e L_g \xi, T_0 \Lambda_\alpha (D_2 B(g, e)(\xi)) \big) \right\rangle \\
&= - \Big\langle a, T_0 \phi(g^{-1}) \Big(D_2 B(g, e)(\xi) \\
&\quad + T_{-B(g,g^{-1}) - \alpha} \Lambda_{\alpha + B(g,g^{-1})} \left(T_g \phi^{-\phi(g^{-1})(B(g,g^{-1}) + \alpha)} (T_e L_g \xi) \right) \\
&\quad + T_{B(g,g^{-1})} \Lambda_{-B(g,g^{-1})} \left(D_1 B(g, g^{-1})(T_e L_g \xi) \right) \Big) \Big\rangle.
\end{aligned}
$$

However, since

$$\phi^{-\phi(g^{-1})(B(g,g^{-1}) + \alpha)} \circ L_g = \phi^{-(B(g,g^{-1}) + \alpha)} \circ \mathrm{AD}_g,$$

where $\mathrm{AD}_g(h) = ghg^{-1}$ for any $g, h \in G$ denotes conjugation, the preceding expression becomes

$$- \mathbf{i}_{(\xi,0)_{G \times_{\phi,B} A}} \left\langle a, \mathcal{A}^1 \right\rangle (g, \alpha) = - \Big\langle a, T_0 \phi(g^{-1}) \Big(D_2 B(g, e)(\xi)$$

$$+ T_{-B(g,g^{-1})-\alpha} \Lambda_{\alpha + B(g,g^{-1})} \left(T_e \phi^{-(B(g,g^{-1})+\alpha)}(\mathrm{Ad}_g \xi) \right)$$

$$+ T_{B(g,g^{-1})} \Lambda_{-B(g,g^{-1})} \left(D_1 B(g, g^{-1})(T_e L_g \xi) \right) \Big) \Big\rangle. \quad (10.2.10)$$

The second term vanishes because of what we claim to be a general formula:

$$\left\langle a, \left(T_0 \phi(g^{-1}) \circ T_{-\beta} \Lambda_\beta \circ T_e \phi^{-\beta} \circ \mathrm{Ad}_g \right)(\xi) \right\rangle = 0 \qquad (10.2.11)$$

which holds for any $g \in G, \beta \in \mathfrak{a}$, and $\xi \in \mathfrak{g}_a$. Accepting this for the moment, we conclude from (10.2.11) and (10.2.10) that

$$- \mathbf{i}_{(\xi,0)_{G \times_{\phi,B} A}} \left\langle a, \mathcal{A}^1 \right\rangle (g, \alpha) = - \Big\langle a, T_0 \phi(g^{-1}) \Big(D_2 B(g, e)(\xi)$$

$$+ T_{B(g,g^{-1})} \Lambda_{-B(g,g^{-1})} \left(D_1 B(g, g^{-1})(T_e L_g \xi) \right) \Big) \Big\rangle$$

which does not depend on $\alpha \in \mathfrak{a}$. This, together with (10.2.9), shows that we can choose $\psi^a : G \to \mathfrak{g}_a^*$ to be given by

$$\langle \psi^a(g), \xi \rangle = - \Big\langle a, T_0 \phi(g^{-1}) \Big(D_2 B(g, e)(\xi)$$

$$+ T_{B(g,g^{-1})} \Lambda_{-B(g,g^{-1})} \left(D_1 B(g, g^{-1})(T_e L_g \xi) \right) \Big) \Big\rangle$$

for any $g \in G$ and $\xi \in \mathfrak{g}_a$ which proves (10.2.6).

It remains to prove formula (10.2.11). We begin by noting that

$$\left(T_0 \phi(g^{-1}) \circ T_{-\beta} \Lambda_\beta \circ T_e \phi^{-\beta} \circ \mathrm{Ad}_g \right)(\xi)$$

$$= \frac{d}{dt}\bigg|_{t=0} \phi(g^{-1}) \left[-\phi \left(g \exp(t\xi) g^{-1} \right)(\beta) + \beta \right]$$

$$= \frac{d}{dt}\bigg|_{t=0} \left[-\phi(\exp(t\xi)) \phi(g^{-1})(\beta) + \phi(g^{-1})(\beta) \right]$$

$$= \left(T_{-\phi(g^{-1})(\beta)} \Lambda_{\phi(g^{-1})(\beta)} \circ T_e \phi^{-\phi(g^{-1})(\beta)} \right)(\xi).$$

Thus (10.2.11) is equivalent to

$$\left\langle a, \left(T_{-\phi(g^{-1})(\beta)} \Lambda_{\phi(g^{-1})(\beta)} \circ T_e \phi^{-\phi(g^{-1})(\beta)} \right)(\xi) \right\rangle = 0 \qquad (10.2.12)$$

for any $g \in G, \beta \in \mathfrak{a}$, and $\xi \in \mathfrak{g}_a$, which, setting $\gamma := -\phi(g^{-1})(\beta)$, is in turn implied by the general formula

$$\langle a, (T_\gamma \Lambda_{-\gamma} \circ T_e \phi^\gamma)(\xi) \rangle = 0 \qquad (10.2.13)$$

for any $\gamma \in A$ and $\xi \in \mathfrak{g}_a$. To prove (10.2.13), we begin by noting that since A is a connected Abelian Lie group, the exponential map $\exp_A : \mathfrak{a} \to A$ is a surjective group homomorphism and hence any γ is of the form $\gamma = \exp_A u$ for some $u \in \mathfrak{a}$. Therefore, by (10.1.4) we get

$$
\begin{aligned}
T_e \phi^{\exp_A u}(\xi) &= \frac{d}{dt}\bigg|_{t=0} \phi(\exp(t\xi))(\exp_A u) \\
&= \frac{d}{dt}\bigg|_{t=0} \exp_A \left(T_0 \phi(\exp(t\xi)) u \right) \\
&= T_u \exp_A \left(\frac{d}{dt}\bigg|_{t=0} T_0 \phi(\exp(t\xi)) u \right) \\
&= T_0 \Lambda_{\exp_A u} \left(\widetilde{\phi}(\xi)(u) \right),
\end{aligned}
$$

since for an Abelian Lie group A we have $T_u \exp_A = T_0 \Lambda_{\exp_A u}$. Consequently,

$$
\left(T_\gamma \Lambda_{-\gamma} \circ T_e \phi^\gamma \right)(\xi) = \widetilde{\phi}(\xi)(u),
$$

for any $u \in \mathfrak{a}$ satisfying $\exp_A u = \gamma$. Therefore,

$$
\langle a, \left(T_\gamma \Lambda_{-\gamma} \circ T_e \phi^\gamma \right)(\xi) \rangle = \left\langle a, \widetilde{\phi}(\xi)(u) \right\rangle = \left\langle \widetilde{\phi}(\xi)^* a, u \right\rangle = 0
$$

since $\xi \in \mathfrak{g}_a$ which proves (10.2.13) and hence the lemma. ∎

Since $B(e, g) = B(g, e) = 0$ for all $g \in G$ we have $\psi^a(e) = 0$.

The Induced Momentum Map. According to Theorem 7.1.1(i), the right G_a-action on $\left(T^*G, \Omega - \pi_G^* \langle a, \mathcal{B}^1 \rangle \right)$ admits a nonequivariant momentum map $\mathbf{J}_a : T^*G \to \mathfrak{g}_a^*$ given by $\mathbf{J}_a = \mathbf{J}_{\mathrm{can}} - \psi^a \circ \pi_Q$, where $\mathbf{J}_{\mathrm{can}}(\alpha_g) = \mathbf{J}_R(\alpha_g)|_{\mathfrak{g}_a} = T_e^* L_g(\alpha_g)|_{\mathfrak{g}_a}$ is the standard momentum map for the right G_a-action relative to the canonical symplectic form on T^*G. Thus, by (10.2.6), we get

$$
\begin{aligned}
\langle \mathbf{J}_a(\alpha_g), \xi \rangle = \alpha_g \left(T_e L_g(\xi) \right) + \Big\langle a, T_0 \phi(g^{-1}) \Big(&D_2 B(g, e)(\xi) \\
&+ T_{B(g,g^{-1})} \Lambda_{-B(g,g^{-1})} \left(D_1 B(g, g^{-1})(T_e L_g \xi) \right) \Big) \Big\rangle
\end{aligned}
$$
(10.2.14)

for any $\alpha_g \in T_g^* G$ and $\xi \in \mathfrak{g}_a$.

The Cocycle of $\mathbf{J}_a$. According to (7.1.2), the nonequivariance group one-cocycle $\sigma^a : G_a \to \mathfrak{g}_a^*$ of $\mathbf{J}_a$ is given for any $h \in G_a$ by

$$
\sigma^a(h) = -\psi^a(h) + \mathrm{Ad}_h^*(\psi^a(e)) = -\psi^a(h)
$$
(10.2.15)

since $\psi^a(e) = 0$. We have used here the fact that the group one-cocycle does not depend on the point on the connected manifold, in our case G, where it is evaluated, so the point e was chosen in order to get the simplest expression.

The Affine Action.   Next we compute the associated right affine action of G_a on $\mathfrak{g}_a^*$ according to the procedures for reducing by a nonequivariant momentum map. Recall from (7.1.5) that this right action is given by

$$\lambda \cdot h = \mathrm{Ad}_h^* \lambda + \sigma^a(h) = \mathrm{Ad}_h^* \lambda - \psi^a(h) \tag{10.2.16}$$

for $\lambda \in \mathfrak{g}_a^*$ and $h \in G_a$. The associated isotropy subgroup (to which we also append a superscript a) is then

$$\widetilde{G}_\lambda^a = \{h \in G_a \mid \mathrm{Ad}_h^* \lambda - \psi^a(h) = \lambda\}. \tag{10.2.17}$$

The Nonequivariance Lie Algebra Two-Cocycle.   We compute $\Sigma^a :$ $\mathfrak{g}_a \times \mathfrak{g}_a \to \mathbb{R}$ from equation (7.1.7) and use a superscript to indicate its dependence on $a \in \mathfrak{a}^*$. Since the two-cocycle is independent of the point $g \in G$ (by connectedness of G), we can evaluate it at e and use $\psi^a(e) = 0$ to get for any $\xi, \eta \in \mathfrak{g}_a$

$$\begin{aligned}
\Sigma^a(\xi, \eta) &= - \langle \psi^a(e), [\xi, \eta] \rangle - \langle a, \mathcal{B}^1(e)(\xi_G(e), \eta_G(e)) \rangle \\
&= - \langle a, \mathcal{B}^1(e)(\xi, \eta) \rangle \\
&= - \langle a, C(\xi, \eta) \rangle, \tag{10.2.18}
\end{aligned}$$

where we used Theorem 10.1.1(iii) in the last equality.

Therefore, by (7.1.10), the isotropy algebra at $\lambda \in \mathfrak{g}_a^*$ for the affine action of G_a on $\mathfrak{g}_a^*$ is given by

$$\widetilde{\mathfrak{g}}_\lambda^a = \{\xi \in \mathfrak{g}_a \mid \mathrm{ad}_\xi^* \lambda + \langle a, C(\xi, \cdot)|_{\mathfrak{g}_a} \rangle = 0\}. \tag{10.2.19}$$

Coadjoint Orbits of Group Extensions.   By the Reduction by Stages Theorem 5.2.9, the coadjoint orbit $\mathcal{O}_{(\mu, a)}$ of $G \times_{\phi, B} A$ through $(\mu, a) \in \mathfrak{g}^* \times \mathfrak{a}^*$ is symplectically diffeomorphic to the reduced space $\mathbf{J}_a^{-1}(\iota_a^* \mu)/\widetilde{G}_{\iota_a^* \mu}^a$, where $\iota_a : \mathfrak{g}_a \hookrightarrow \mathfrak{g}$ is the inclusion and its dual $\iota_a^* : \mathfrak{g}^* \to \mathfrak{g}_a^*$ is the projection. We summarize this observation in the following statement.

10.2.2 Proposition.   *Any two coadjoint orbits $\mathcal{O}_{(\mu, a)}, \mathcal{O}_{(\mu', a)} \subset \mathfrak{g}^* \times \mathfrak{a}^*$ such that $\mu|_{\mathfrak{g}_a} = \mu'|_{\mathfrak{g}_a}$ are symplectically diffeomorphic to the G_a-reduced space $\mathbf{J}_a^{-1}(\iota_a^* \mu)/\widetilde{G}_{\iota_a^* \mu}^a$.*

By the Magnetic Symplectic Embedding Theorem 7.1.6, the reduced space $\mathbf{J}_a^{-1}(\iota_a^* \mu)/\widetilde{G}_{\iota_a^* \mu}^a$, and hence the coadjoint orbit $\mathcal{O}_{(\mu, a)}$, symplectically embeds into

$$\left(T^* \left(G/\widetilde{G}_{\iota_a^*(\mu)}^a \right), \Omega_{\mathrm{can}} - \pi_{G/\widetilde{G}_{\iota_a^* \mu}^a}^* \mathcal{B}_{\iota_a^* \mu} \right)$$

as the vector subbundle $\left[T\pi_{G, \widetilde{G}_{\iota_a^*(\mu)}}(V) \right]^\circ$. Here, $\pi_{G, \widetilde{G}_{\iota_a^*(\mu)}} : G \to G/\widetilde{G}_{\iota_a^*(\mu)}$ is the projection, $V_g = \{T_e L_g \zeta \mid \zeta \in \mathfrak{g}_a\}$ is the fiber of the vertical subbundle $V \subset TG$ relative to the right G_a-translation action on G,

$\pi_{G/\widetilde{G}^a_{\iota^*_a\mu}} : T^* \left(G/\widetilde{G}^a_{\iota^*_a\mu} \right) \rightarrow G/\widetilde{G}^a_{\iota^*_a\mu}$ is the cotangent bundle projection, Ω_{can} is the canonical symplectic form on this cotangent bundle. Also, $\mathcal{B}_{\iota^*_a\mu}$ is a closed two-form on $G/\widetilde{G}^a_{\iota^*_a\mu}$ naturally induced by the three quantities $\langle a, \mathcal{B}^1 \rangle$, ψ^a, and the connection $\mathcal{A}^{G_a} \in \Omega^2(G; \mathfrak{g}_a)$ as described in Lemma 7.1.5.

The magnetic term $\mathcal{B}_{\iota^*_a\mu} \in \Omega^2 \left(G/\widetilde{G}^a_{\iota^*_a\mu} \right)$ is computed in this concrete case in the following way. By the general formula (7.1.16), we have

$$\pi^*_{G/\widetilde{G}^a_{\iota^*_a\mu}} \mathcal{B}_{\iota^*_a\mu} = \langle a, \mathcal{B}^1 \rangle + \mathbf{d} \left(\mathcal{A}^{G_a}_{\psi^a} + \mathcal{A}^{G_a}(\cdot)^* \iota^*_a \mu \right) \in \Omega^2(G), \qquad (10.2.20)$$

where $\langle a, \mathcal{B}^1 \rangle (g)(T_e R_g \xi, T_e R_g \eta) = \langle a, T_0 \phi(g^{-1}) (C(\xi, \eta)) \rangle$ for any $\xi, \eta \in \mathfrak{g}$ (see Theorem 10.1.1(iii)) and, according to (7.1.11), the one-form $\mathcal{A}^{G_a}_{\psi^a} \in \Omega^1(G)$ is given by

$$\mathcal{A}^{G_a}_{\psi^a}(g)(v_g) = \langle \psi^a(g), \mathcal{A}^{G_a}(g)(v_g) \rangle \qquad (10.2.21)$$

and $\mathcal{A}^{G_a}(\cdot)^* \iota^*_a \mu \in \Omega^1(G)$ by

$$\left(\mathcal{A}^{G_a}(\cdot)^* \iota^*_a \mu \right)(g)(v_g) = \langle \iota^*_a \mu, \mathcal{A}^{G_a}(g)(v_g) \rangle \qquad (10.2.22)$$

for any $g \in G$ and $v_g \in T_g G$. Thus, by (10.2.21), (10.2.22), and (10.2.6), we have

$$\left(\mathcal{A}^{G_a}_{\psi^a} + \mathcal{A}^{G_a}(\cdot)^* \iota^*_a \mu \right)(g)(v_g) = \langle \psi^a(g) + \iota^*_a \mu, \mathcal{A}^{G_a}(g)(v_g) \rangle$$
$$= - \Big\langle a, T_0 \phi(g^{-1}) \Big[D_2 B(g, e) \left(\mathcal{A}^{G_a}(g)(v_g) \right)$$
$$+ T_{B(g,g^{-1})} \Lambda_{-B(g,g^{-1})} \Big(D_1 B(g, g^{-1}) \big(T_e L_g \left(\mathcal{A}^{G_a}(g)(v_g) \right) \big) \Big) \Big] \Big\rangle$$
$$+ \langle \mu, \mathcal{A}^{G_a}(g)(v_g) \rangle \qquad (10.2.23)$$

where $\mathcal{A}^{G_a}(g)(v_g)$ is given by (10.2.4). We summarize these computations in the following statement.

10.2.3 Proposition.   *The reduced space $\mathbf{J}^{-1}_a(\iota^*_a\mu)/\widetilde{G}^a_{\iota^*_a\mu}$ is symplectically diffeomorphic to the coadjoint orbit $\mathcal{O}_{(\mu, a)} \subset \left(\mathfrak{g} \times_{\widetilde{\phi}, C} \mathfrak{a} \right)^*$ and symplectically embeds into*

$$\left(T^* \left(G/\widetilde{G}^a_{\iota^*_a(\mu)} \right), \Omega_{\mathrm{can}} - \pi^*_{G/\widetilde{G}^a_{\iota^*_a\mu}} \mathcal{B}_{\iota^*_a\mu} \right)$$

as the vector subbundle $\left[T\pi_{G, \widetilde{G}_{\iota^*_a(\mu)}}(V) \right]^\circ$, where

- *$\pi_{G, \widetilde{G}_{\iota^*_a(\mu)}} : G \rightarrow G/\widetilde{G}_{\iota^*_a(\mu)}$ is the orbit space projection,*

- $V \subset TG$ is the vertical subbundle associated to the right principal G_a-bundle $G \to G/G_a$ whose fiber at $g \in G$ is therefore given by $V_g = \{T_e L_g \zeta \mid \zeta \in \mathfrak{g}_a\}$,

- $\pi_{G/\widetilde{G}^a_{\iota^*_a \mu}} : T^* \left(G/\widetilde{G}^a_{\iota^*_a \mu} \right) \to G/\widetilde{G}^a_{\iota^*_a \mu}$ is cotangent bundle projection,

- Ω_{can} is the canonical symplectic form on $T^* \left(G/\widetilde{G}^a_{\iota^*_a \mu} \right)$, and

- the magnetic term $\mathcal{B}_{\iota^*_a \mu}$ is a closed two-form on $G/\widetilde{G}^a_{\iota^*_a \mu}$ uniquely determined by (10.2.20) and (10.2.23).

While (10.2.20) and (10.2.23) give, in principle, the explicit formula of the magnetic term $\mathcal{B}_{\iota^*_a \mu} \in \Omega^2 \left(G/\widetilde{G}^a_{\iota^*_a \mu} \right)$, the actual computation is very long and the resulting formula is complicated. On the other hand, in specific examples, this computation can be carried out and gives concrete expressions for the magnetic term as we shall see below.

Group Extensions with Vector Spaces. If the Lie group A in the semidirect product is a vector space V and ϕ is a representation, some of these formulas simplify. The mechanical connection $\mathcal{A}^{G_a} \in \Omega^1(G; \mathfrak{g}_a)$ is still given by (10.2.4) but the $\mathcal{B}\mathfrak{g}$-potential $\psi^a : G \to \mathfrak{g}^*_a$ in (10.2.6) simplifies to

$$\psi^a(g) = - \left[(D_2 B(g,e))^*(g \cdot a) + T^*_e L_g(D_1 B(g,g^{-1}))^*(g \cdot a) \right] |_{\mathfrak{g}_a}$$

for all $g \in G$. Recall that the contragredient representation of G on V^* is given by $\langle g \cdot a, v \rangle := \langle a, g^{-1} \cdot v \rangle$, for any $g \in G$, $v \in V$, and $a \in V^*$. The nonequivariant momentum map $\mathbf{J}_a : T^*G \to \mathfrak{g}^*_a$ given in (10.2.14) becomes

$$\mathbf{J}_a(\alpha_g) = \left[T^*_e L_g \alpha_g + (D_2 B(g,e))^*(g \cdot a) + T^*_e L_g(D_1 B(g,g^{-1}))^*(g \cdot a) \right] |_{\mathfrak{g}_a}$$

for any $\alpha_g \in T^*_g G$ and the nonequivariance group one-cocycle is $-\psi^a|_{G_a}$. The right affine action (10.2.16) of G_a on $\mathfrak{g}^*_a$ relative to which $\mathbf{J}_a$ is equivariant is in this case given by

$$\lambda \cdot h = \mathrm{Ad}^*_h \lambda + \left[(D_2 B(h,e))^*(h \cdot a) + T^*_e L_h(D_1 B(h,h^{-1}))^*(h \cdot a) \right] |_{\mathfrak{g}_a}$$

for $\lambda \in \mathfrak{g}^*_a$ and $h \in G_a$. The infinitesimal nonequivariance Lie algebra two-cocycle is still given by (10.2.18), that is, $\Sigma^a(\zeta_1, \zeta_2) = - \langle a, C(\zeta_1, \zeta_2) \rangle$ for any $\zeta_1, \zeta_2 \in \mathfrak{g}_a$. Finally, the pull back to G of the magnetic term $\mathcal{B}_{\iota^*_a \mu}$ is the sum of the closed two-form $\langle a, \mathcal{B}^1 \rangle$ whose value at $g \in G$ is

$$\langle a, \mathcal{B}^1 \rangle (g)(T_e R_g \xi, T_e R_g \eta) = \langle a, \ g^{-1} \cdot C(\xi, \eta) \rangle$$

for any $\xi, \eta \in \mathfrak{g}$ and the differential of the one-form whose value at $g \in G$ is

$$\left(\mathcal{A}^{G_a}(g) \right)^* \left(\left[\mu - (D_2 B(g,e))^*(g \cdot a) - T^*_e L_g(D_1 B(g,g^{-1}))^*(g \cdot a) \right] |_{\mathfrak{g}_a} \right),$$

where $\mathcal{A}^{G_a}$ is given by (10.2.4).

If, in addition, the cocycle $B = 0$, that is, we are dealing with an ordinary semidirect product $G \circledS V$, then also $C = 0$ and these formulas recover those in the second half of §4.3 dedicated to stage two reduction. For example, the $\mathcal{B}\mathfrak{g}$-potential is zero, formula (10.2.4) of the mechanical connection $\mathcal{A}^{G_a}$ can easily be rewritten as (4.3.11), the momentum map $\mathbf{J}_a : T^*G \to \mathfrak{g}_a^*$ is equivariant, and the pull back to G of the magnetic term is the closed two-form whose value at $g \in G$ is given by $\left(\mathcal{A}^{G_a}(g) \right)^* (\mu|_{\mathfrak{g}_a}) \in T_g^*G$.

10.3 Coupling to a Lie Group

In this section we prove some preliminary results on reduction of a manifold of the form $T^*G \times P$, where P is a Poisson or symplectic manifold, under the action of G. Here, G acts on T^*G by cotangent lift of right (or left) translations and G also acts on P on the right (or left) by Poisson (or symplectic) transformations. The goal is to show that in these circumstances, the problem is equivalent to reducing $T^*G \times P$ by G, where the action of G on P is now trivial.

In the subsequent two sections we will be dealing with reduction by stages for the semidirect product of two groups. The results of this section will be helpful in these subsequent developments.

The Poisson Case.   We begin with a discussion of the Poisson case; in subsequent paragraphs we deal with the symplectic case. The Poisson case has proven to be useful in some applications, such as the stability theory for systems coupling flexible structures to rigid bodies; see Krishnaprasad and Marsden [1987] and Simo, Posbergh, and Marsden [1990].

The setup is as follows. Let G be a Lie group acting by canonical (Poisson) transformations on a Poisson manifold P. We shall carry along both the case of right and of left group actions; the primary results will be given for right actions to be compatible with this and the preceding chapters.

Define $\psi : T^*G \times P \to \mathfrak{g}^* \times P$ by

$$\psi(\alpha_g, x) = \left(T_e^* R_g \alpha_g, x \cdot g^{-1} \right). \tag{10.3.1}$$

In the case of left actions, we would use the map

$$\psi(\alpha_g, x) = \left(T_e^* L_g \alpha_g, g^{-1} \cdot x \right). \tag{10.3.2}$$

Here, $x \cdot g^{-1}$ (resp. $g^{-1} \cdot x$) denotes the right (resp. left) action of g^{-1} on $x \in P$.

For $F, K : \mathfrak{g}^* \times P \to \mathbb{R}$, let $\{F, K\}_\pm$ stand for the plus (resp. minus) Lie–Poisson bracket holding the P variable fixed and let $\{F, K\}_P$ stand for the Poisson bracket on P with the variable $\mu \in \mathfrak{g}^*$ held fixed.

Endow $\mathfrak{g}^* \times P$ with the following bracket (evaluated at the point (μ, x), which we suppress in the notation):

$$\{F, K\} = \{F, K\}_\pm + \{F, K\}_P - \left\langle \mathbf{d}_x F, \left(\frac{\delta K}{\delta \mu}\right)_P \right\rangle$$
$$+ \left\langle \mathbf{d}_x K, \left(\frac{\delta F}{\delta \mu}\right)_P \right\rangle \qquad (10.3.3)$$

where $\mathbf{d}_x F$ means the differential of F with respect to $x \in P$.

10.3.1 Proposition.   *The bracket* (10.3.3) *makes* $\mathfrak{g}^* \times P$ *into a Poisson manifold and* $\psi : T^*G \times P \to \mathfrak{g}^* \times P$ *is a Poisson map, where the Poisson structure on* $T^*G \times P$ *is given by the sum of the canonical bracket on* T^*G *and the bracket on* P. *Moreover,* ψ *is G-invariant and induces a Poisson diffeomorphism* $[\psi]$ *of* $(T^*G \times P)/G$ *with* $\mathfrak{g}^* \times P$.

Proof.   For $F, K : \mathfrak{g}^* \times P \to \mathbb{R}$, let $\bar{F} = F \circ \psi$ and $\bar{K} = K \circ \psi$. We claim that $\{\bar{F}, \bar{K}\}_{T^*G} + \{\bar{F}, \bar{K}\}_P = \{F, K\} \circ \psi$. This will show that ψ is Poisson. Since it is easy to check that ψ is G-invariant and gives a diffeomorphism of $(T^*G \times P)/G$ with $\mathfrak{g}^* \times P$, it follows that (10.3.3) represents the reduced bracket and so defines a Poisson structure.

To prove our claim, write $\psi = \psi_G \times \psi_P$ and let $\mu = T_e^* R_g \alpha_g = \psi_G(\alpha_g, x)$. Since ψ_G does not depend on x and the group action is assumed canonical, $\{\bar{F}, \bar{K}\}_P = \{F, K\}_P \circ \psi$. For the T^*G bracket, note that since ψ_G is a Poisson map of T^*G to $\mathfrak{g}^*_\pm$, the terms involving ψ_G will be $\{F, K\}_\pm \circ \psi$. The terms involving $\psi_P(\alpha_g, x) = x \cdot g^{-1}$ (resp. $\psi_P(\alpha_g, x) = g^{-1} \cdot x$) can be computed as follows. Note that the T^*G-bracket of a function S of g with a function L of α_g is

$$\{S, L\}_{T^*G} = \left\langle \mathbf{d}_g S, \frac{\delta L}{\delta \alpha_g} \right\rangle$$

where $\delta L / \delta \alpha_g$ means the fiber derivative of L regarded as a vector at g. This is paired with the covector $\mathbf{d}_g S$.

Letting $\Psi_x(g) = x \cdot g^{-1}$ (resp. $\Psi_x(g) = g^{-1} \cdot x$), we find by use of the chain rule and the definition of the functional derivative, that the ψ_P terms in the bracket are

$$\left\langle \mathbf{d}_x F, T_g \Psi_x \left(\frac{\delta K}{\delta \mu} \cdot g\right) \right\rangle - \left\langle \mathbf{d}_x K, T_g \Psi_x \left(\frac{\delta F}{\delta \mu} \cdot g\right) \right\rangle. \qquad (10.3.4)$$

However, it is easily checked that

$$T_g \Psi_x(\xi \cdot g) = -\xi_P(x \cdot g^{-1}),$$

so the expression (10.3.4) becomes

$$-\left\langle \mathbf{d}_x F, \left(\frac{\delta K}{\delta \mu}\right)_P (x \cdot g^{-1}) \right\rangle + \left\langle \mathbf{d}_x K, \left(\frac{\delta F}{\delta \mu}\right)_P (x \cdot g^{-1}) \right\rangle, \qquad (10.3.5)$$

where $\mathbf{d}_x F$ and $\mathbf{d}_x K$ are evaluated at $(\mu, x \cdot g^{-1})$.

Thus, the expression (10.3.5) equals the two remaining terms in (10.3.3) composed with ψ. ∎

Notice that Proposition 10.3.1 can be reformulated in the following manner: the Poisson manifold $(\mathfrak{g}^* \times P, \{\,,\,\})$ endowed with the bracket (10.3.3) is Poisson isomorphic via the G-quotient of the map ψ to the reduction of the product Poisson manifold $T^*G \times P$ by the diagonal G-action.

Now suppose that the action of G (which we assume is connected) on P has a momentum map $\mathbf{J} : P \to \mathfrak{g}^*$. Suppose that this momentum map has a nonequivariance one-cocycle $\sigma : G \to L(\mathfrak{g}, \mathcal{C}(P))$ (a linear map from $\mathfrak{g}$ to the space of Casimir functions) defined by

$$\sigma(g) = \mathbf{J}(z \cdot g) - \mathbf{J}(z) \circ \mathrm{Ad}_g;$$

for left actions, we use

$$\sigma(g) = \mathbf{J}(g \cdot z) - \mathbf{J}(z) \circ \mathrm{Ad}_{g^{-1}}.$$

The expressions on the right hand sides are independent of $z \in P$, which follows from connectedness of G (see [MandS], §12.3). The corresponding infinitesimal two-cocycle $\Sigma : \mathfrak{g} \times \mathfrak{g} \to \mathcal{C}(P)$ is defined by $\Sigma(\xi, \eta) = -T_e \sigma_\eta(\xi)$, where $\sigma_\eta : G \to \mathcal{C}(P)$ is defined by $\sigma_\eta(g) = \sigma(g) \cdot \eta$. For left actions the corresponding infinitesimal two-cocycle is defined by $\Sigma(\xi, \eta) = T_e \sigma_\eta(\xi)$.

The infinitesimal two-cocycle satisfies

$$\Sigma(\xi, \eta) = \mathbf{J}^{[\xi, \eta]} + \left\{ \mathbf{J}^\xi, \mathbf{J}^\eta \right\} \tag{10.3.6}$$

and in the case of left actions, it reads

$$\Sigma(\xi, \eta) = \mathbf{J}^{[\xi, \eta]} - \left\{ \mathbf{J}^\xi, \mathbf{J}^\eta \right\}. \tag{10.3.7}$$

Consider the map $\alpha : \mathfrak{g}^* \times P \to \mathfrak{g}^* \times P$ given by

$$\alpha(\mu, x) = (\mu + \mathbf{J}(x), x). \tag{10.3.8}$$

Let the bracket $\{\,,\,\}_0^\Sigma$ on $\mathfrak{g}^* \times P$ be defined by

$$\{F, K\}_0^\Sigma = \{F, K\}_\pm^\Sigma + \{F, K\}_P, \tag{10.3.9}$$

where

$$\{F, K\}_\pm^\Sigma = \{F, K\}_\pm \mp \Sigma\left(\frac{\delta F}{\delta \nu}, \frac{\delta K}{\delta \nu} \right)$$

is the Lie-Poisson bracket with the two-cocycle $C = -\Sigma$ (see Theorem 6.2.2). Thus $\{\,,\,\}_0^\Sigma$ is (10.3.3) with the coupling or interaction terms replaced by a cocycle term. We claim that the map α transforms one bracket to the other.

10.3.2 Proposition. *Assume that the G-action on P admits a momentum map $\mathbf{J} : P \to \mathfrak{g}^*$ with infinitesimal two-cocycle $\Sigma : \mathfrak{g} \times \mathfrak{g} \to \mathcal{C}(P)$. Then the mapping $\alpha : (\mathfrak{g}^* \times P, \{\,,\}) \to (\mathfrak{g}^* \times P, \{\,,\}_0^\Sigma)$ is a Poisson diffeomorphism.*

Proof. For $F, K : \mathfrak{g}^* \times P \to \mathbb{R}$, let $\hat{F} = F \circ \alpha$ and $\hat{K} = K \circ \alpha$. Letting $\nu = \mu + \mathbf{J}(x)$, and dropping the evaluation points, we conclude that

$$\frac{\delta \hat{F}}{\delta \mu} = \frac{\delta F}{\delta \nu} \quad \text{and} \quad \mathbf{d}_x \hat{F} = \left\langle \frac{\delta F}{\delta \nu}, \mathbf{d}_x \mathbf{J} \right\rangle + \mathbf{d}_x F. \tag{10.3.10}$$

Let $B_x : T_x^* P \to T_x P$ be the Poisson tensor on P so that $\{F, K\}(x) = \langle \mathbf{d}_x F, B_x(\mathbf{d}_x K) \rangle$. Substituting this and (10.3.10) into the bracket (10.3.3), we get

$$\begin{aligned}
\{\hat{F}, \hat{K}\} = &\pm \left\langle \mu, \left[\frac{\delta F}{\delta \nu}, \frac{\delta K}{\delta \nu} \right] \right\rangle \\
&+ \left\langle \left\langle \frac{\delta F}{\delta \nu}, \mathbf{d}_x \mathbf{J} \right\rangle + \mathbf{d}_x F, B_x \left(\left\langle \frac{\delta K}{\delta \nu}, \mathbf{d}_x \mathbf{J} \right\rangle + \mathbf{d}_x K \right) \right\rangle \\
&- \left\langle \left\langle \frac{\delta F}{\delta \nu}, \mathbf{d}_x \mathbf{J} \right\rangle + \mathbf{d}_x F, \left(\frac{\delta K}{\delta \nu} \right)_P \right\rangle \\
&+ \left\langle \left\langle \frac{\delta K}{\delta \nu}, \mathbf{d}_x \mathbf{J} \right\rangle + \mathbf{d}_x K, \left(\frac{\delta F}{\delta \nu} \right)_P \right\rangle. \tag{10.3.11}
\end{aligned}$$

However, by definition of the momentum map, we have

$$B_x \left(\langle \xi, \mathbf{d}_x \mathbf{J} \rangle \right) = \xi_P.$$

With $\xi = \delta K / \delta \nu$, substituting this into (10.3.11), gives

$$\begin{aligned}
\{\hat{F}, \hat{K}\} = &\pm \left\langle \mu, \left[\frac{\delta F}{\delta \nu}, \frac{\delta K}{\delta \nu} \right] \right\rangle \\
&+ \left\langle \left\langle \frac{\delta F}{\delta \nu}, \mathbf{d}_x \mathbf{J} \right\rangle + \mathbf{d}_x F, B_x \left(\mathbf{d}_x K \right) \right\rangle \\
&+ \left\langle \left\langle \frac{\delta K}{\delta \nu}, \mathbf{d}_x \mathbf{J} \right\rangle + \mathbf{d}_x K, \left(\frac{\delta F}{\delta \nu} \right)_P \right\rangle \\
= &\pm \left\langle \mu, \left[\frac{\delta F}{\delta \nu}, \frac{\delta K}{\delta \nu} \right] \right\rangle + \{F, K\}_P \\
&- \left\langle B_x \left\langle \frac{\delta F}{\delta \nu}, \mathbf{d}_x \mathbf{J} \right\rangle, \mathbf{d}_x K \right\rangle \\
&+ \left\langle \left\langle \frac{\delta K}{\delta \nu}, \mathbf{d}_x \mathbf{J} \right\rangle + \mathbf{d}_x K, \left(\frac{\delta F}{\delta \nu} \right)_P \right\rangle. \tag{10.3.12}
\end{aligned}$$

Again using $B_x\left(\langle\xi,\mathbf{d}_x\mathbf{J}\rangle\right)=\xi_P$, this time with $\xi=\delta F/\delta\nu$, we get

$$\{\hat{F},\hat{K}\}=\pm\left\langle\mu,\left[\frac{\delta F}{\delta\nu},\frac{\delta K}{\delta\nu}\right]\right\rangle+\{F,K\}_P$$
$$+\left\langle\left\langle\frac{\delta K}{\delta\nu},\mathbf{d}_x\mathbf{J}\right\rangle,\left(\frac{\delta F}{\delta\nu}\right)_P\right\rangle. \qquad (10.3.13)$$

Using the definition of the momentum map and the infinitesimal nonequivariance cocycle, one has, for any $\xi,\eta\in\mathfrak{g}$,

$$\langle\langle\eta,\mathbf{d}_x\mathbf{J}\rangle,\xi_P\rangle=\langle\mathbf{d}_x\mathbf{J}^\eta,X_{\mathbf{J}^\xi}\rangle=\{\mathbf{J}^\eta,\mathbf{J}^\xi\}=\pm\langle\mathbf{J},[\xi,\eta]\rangle\mp\Sigma(\xi,\eta). \qquad (10.3.14)$$

Taking $\xi=\delta F/\delta\nu$ and $\eta=\delta K/\delta\nu$ and substituting (10.3.14) into (10.3.13) gives

$$\{\hat{F},\hat{K}\}(\mu,x)=\pm\left\langle\mu+\mathbf{J}(x),\left[\frac{\delta F}{\delta\nu},\frac{\delta K}{\delta\nu}\right]\right\rangle\mp\Sigma\left(\frac{\delta F}{\delta\nu},\frac{\delta K}{\delta\nu}\right)+\{F,K\}_P(\nu,x)$$
$$=\pm\left\langle\nu,\left[\frac{\delta F}{\delta\nu},\frac{\delta K}{\delta\nu}\right]\right\rangle\mp\Sigma\left(\frac{\delta F}{\delta\nu},\frac{\delta K}{\delta\nu}\right)+\{F,K\}_P(\nu,x)$$
$$=\left(\{F,K\}_0^\Sigma\circ\alpha\right)(\mu,x). \qquad (10.3.15)$$

Thus, $\{F\circ\alpha,K\circ\alpha\}=\{F,K\}_0^\Sigma\circ\alpha$, which proves the claim. ∎

Remarks.

1. Recall (see [MandS] §12.3 that an equivariant momentum map $\mathbf{J}:P\to\mathfrak{g}_\mp^*$ is a Poisson map. While not a direct corollary, an argument similar to the one above, shows that in the nonequivariant case, $\mathbf{J}$ is a Poisson map of P into $\mathfrak{g}^*$ with the bracket $\{\cdot,\cdot\}_\mp^\Sigma$. Compare with Theorems 6.2.2 and 7.2.1 and note the change in sign.

2. Suppose $C(\nu)$ is a Casimir function on $\mathfrak{g}^*$ with the bracket $\{F,K\}_\pm^\Sigma$. Then
$$C(\mu,x)=C(\mu+\mathbf{J}(x))$$
is a Casimir function on $\mathfrak{g}^*\times P$ for the bracket (10.3.3).

3. An argument similar to the one given above provides the Poisson bracket on T^*G written in body coordinates (compare Abraham and Marsden [1978], p. 315). Namely, the map of T^*G to $\mathfrak{g}^*\times G$ given by $\alpha_g\mapsto(T_eR_g^*\alpha_g,g)$ (resp. $\alpha_g\mapsto(T_eL_g^*\alpha_g,g)$) is a Poisson diffeomorphism that takes the canonical bracket to the following bracket on $\mathfrak{g}^*\times G$:

$$\{F,K\}=\{F,K\}_++\mathbf{d}_gF\cdot T_eR_g\frac{\delta K}{\delta\mu}-\mathbf{d}_gK\cdot T_eR_g\frac{\delta F}{\delta\mu}$$

and in the left case,

$$\{F, K\} = \{F, K\}_- + \mathbf{d}_g F \cdot T_e L_g \frac{\delta K}{\delta \mu} - \mathbf{d}_g K \cdot T_e L_g \frac{\delta F}{\delta \mu},$$

where $\mu \in \mathfrak{g}^*$ and $g \in G$. The point is that the canonical bracket in the trivialization of T^*G has terms added to the standard Lie-Poisson terms that have the same form as the interaction terms in equation (10.3.3).

The Symplectic Case. Now assume that P is a (connected) symplectic manifold. Recall that the symplectic leaves of the Poisson reduced space $(T^*G \times P)/G$ are symplectically diffeomorphic to the the symplectic reduced spaces. We have determined the Poisson reduced space above and have shown that its Poisson bracket is the Lie-Poisson bracket (with cocycle) plus the given P bracket. Thus, since the bracket is now "uncoupled", the symplectic leaves are simply the product of the symplectic leaves in each space; but P is symplectic and connected, so it is its own leaf. Thus, we arrive at the following.

10.3.3 Proposition. *In the preceding set up, the symplectic reduced space at the value $\mu \in \mathfrak{g}^*$ for the action of G on $T^*G \times P$ is symplectically diffeomorphic to $\mathcal{O}_\mu^\Sigma \times P$, where $\mathcal{O}_\mu^\Sigma$ is an (plus for right, minus for left) affine action orbit of G in $\mathfrak{g}^*$.*

For a detailed discussion of these affine orbits, see Theorem 6.2.2. Of course one can keep track of exactly what the symplectic diffeomorphism is by using the theory of orbit reduction and its relation to the symplectic leaves.

10.4   Poisson Reduction by Stages: General Semidirect Products

In the next three sections we study the problem of reduction by stages for the semidirect product of two nonabelian Lie groups acting, again, by the cotangent lift by right translation on itself. The semidirect product will be of the form $G \circledS H$, where G acts on H by Lie group automorphisms. In the language of the general theory, the "big group" is $M = G \circledS H$ and the normal subgroup is $N = H$.

This section is devoted to the Poisson case and the following two sections to the symplectic case.

Review of Semidirect Products of Groups. As in §6.4, let $S = G \circledS H$ be the semidirect product of two (in general, nonabelian) Lie groups associated to the smooth left action $\phi : G \times H \to H$ by H-Lie group

homomorphisms. Denote, as usual, $\phi(g)(h) = g \cdot h$, for $g \in G$ and $h \in H$. The group multiplication is given by (6.4.1), that is,

$$(g_1, h_1)(g_2, h_2) = (g_1 g_2, h_1(g_1 \cdot h_2)),$$

where $g_1, g_2 \in G$, $h_1, h_2 \in H$. The neutral element is (e_G, e_H), where e_G and e_H are the neutral elements of G and H respectively, and $(g, h)^{-1} = (g^{-1}, g^{-1} \cdot h^{-1})$. The subgroup $\{e_G\} \times H$, which is isomorphic to H, is a closed normal subgroup of $G \circledS H$.

Recall that in this context, we write $\tilde{\phi}(g) := T_{e_H}\phi(g) : \mathfrak{h} \to \mathfrak{h}$ for each $g \in G$; this defines a Lie group homomorphism $\tilde{\phi} : g \in G \mapsto \tilde{\phi}(g) \in \mathrm{Aut}(\mathfrak{h})$ (the Lie group of Lie algebra automorphisms of $\mathfrak{h}$). The derivative of this map gives the Lie algebra homomorphism $\tilde{\phi}' := T_{e_G}\tilde{\phi} : \mathfrak{g} \to \mathrm{aut}(\mathfrak{h})$ (the Lie algebra of derivations of $\mathfrak{h}$).

Using this notation, we found that the Lie algebra of S is the semidirect product $\mathfrak{s} = \mathfrak{g} \circledS \mathfrak{h}$ with the commutator given by (6.4.2), that is,

$$[(\xi_1, \eta_1), (\xi_2, \eta_2)] = ([\xi_1, \xi_2], \xi_1 \cdot \eta_2 - \xi_2 \cdot \eta_1 + [\eta_1, \eta_2]),$$

where $\xi_1, \xi_2 \in \mathfrak{g}$, $\eta_1, \eta_2 \in \mathfrak{h}$, and $\xi_1 \cdot \eta_2 := \tilde{\phi}'(\xi_1)(\eta_2)$, $\xi_2 \cdot \eta_1 := \tilde{\phi}'(\xi_2)(\eta_1) \in \mathfrak{h}$ denotes the Lie algebra action of $\xi_1, \xi_2 \in \mathfrak{g}$ on $\eta_2, \eta_1 \in \mathfrak{h}$, respectively.

Untangling the H-action.   We now consider the basic map given by

$$\Upsilon : G \times H \to G \times H; \quad (g, h) \mapsto (g, g^{-1} \cdot h) \qquad (10.4.1)$$

This map has the following key property that is easily verified.

10.4.1 Proposition.   *Let $k \in H$ and let $R_k^S : G \times H \to G \times H$ be right translation by (e_G, k) on $G \circledS H$; that is, $R_k^S(g, h) = (g, h(g \cdot k))$. Let $R_k : G \times H \to G \times H$ be right translation on the direct product; that is, $R_k(g, h) = (g, hk)$. Then Υ is an equivariant diffeomorphism for the H actions; that is*

$$\Upsilon \circ R_k^S = R_k \circ \Upsilon. \qquad (10.4.2)$$

First Stage Reduction: The Poisson Case.   Using the map Υ, we obtain the following.

10.4.2 Proposition.   *The Poisson reduced space $T^*(G \circledS H)/H$, where H acts by the cotangent lift of right translations, regarding H as the subgroup $\{e_G\} \times H \subset G \circledS H$, is Poisson diffeomorphic to $T^*G \times \mathfrak{h}^*$ with Poisson structure the sum of the canonical structure on T^*G and the $(+)$-Lie-Poisson structure on $\mathfrak{h}^*$. In fact, the Poisson diffeomorphism*

$$[T^*\Upsilon^{-1}] : (T^*(G \circledS H))/H \to T^*G \times \mathfrak{h}^*$$

induced by Υ is given by

$$[T^*\Upsilon^{-1}]([(\alpha_g, \beta_h)]) = \left(\alpha_g + T_g^* \phi^{g^{-1} \cdot h} \beta_h, \tilde{\phi}(g)^* T_{e_H}^* R_h \beta_h \right)$$

$$= \left(\alpha_g + \left(T_g^* R_{g^{-1}} \circ T_{e_G}^* \phi^h \right) \beta_h, \tilde{\phi}(g)^* T_{e_H}^* R_h \beta_h \right). \qquad (10.4.3)$$

Proof. The first statement is a direct consequence of the fact, given in (10.4.2), that Υ is an equivariant diffeomorphism. To see that this diffeomorphism is explicitly given by (10.4.3), let $U_g = T_{e_G} R_g \xi \in T_g G$ and $V_h = T_{e_H} R_h \eta$, where $\xi \in \mathfrak{g}$ and $\eta \in \mathfrak{h}$, be tangent vectors to G and H respectively at the points $g \in G$ and $h \in H$. Then we have

$$
\begin{aligned}
T_{(g,h)} \Upsilon^{-1}(U_g, V_h) &= \frac{d}{dt}\bigg|_{t=0} \Upsilon^{-1}\Big((\exp t\xi)g, (\exp t\eta)h\Big) \\
&= \frac{d}{dt}\bigg|_{t=0} \Big((\exp t\xi)g, ((\exp t\xi)g) \cdot ((\exp t\eta)h)\Big) \\
&= \left(U_g, \frac{d}{dt}\bigg|_{t=0} ((\exp t\xi)g) \cdot h + \frac{d}{dt}\bigg|_{t=0} g \cdot ((\exp t\eta)h)\right) \\
&= \left(U_g, \frac{d}{dt}\bigg|_{t=0} (\exp t\xi) \cdot (g \cdot h) + \frac{d}{dt}\bigg|_{t=0} (\phi(g) \circ R_h)(\exp t\eta)\right) \\
&= \left(U_g, \left(T_g R_{g^{-1}} U_g\right)_H (g \cdot h) + T_h \phi(g) V_h\right).
\end{aligned} \tag{10.4.4}
$$

The dual of this map is computed as follows. First note that

$$
\begin{aligned}
\Big\langle T^*_{(g,g^{-1} \cdot h)} &\Upsilon^{-1}(\alpha_g, \beta_h), (U_g, V_{g^{-1} \cdot h})\Big\rangle \\
&= \Big\langle (\alpha_g, \beta_h), T_{(g,g^{-1} \cdot h)}\Upsilon^{-1}(U_g, V_{g^{-1} \cdot h})\Big\rangle \\
&= \Big\langle (\alpha_g, \beta_h), \Big(U_g, \left(T_g R_{g^{-1}} U_g\right)_H (h) + T_{g^{-1} \cdot h}\phi(g)V_{g^{-1} \cdot h}\Big)\Big\rangle \\
&= \langle \alpha_g, U_g\rangle + \Big\langle \beta_h, \left(T_{e_G}\phi^h \circ T_g R_{g^{-1}}\right) U_g\Big\rangle + \Big\langle \beta_h, T_{g^{-1} \cdot h}\phi(g)V_{g^{-1} \cdot h}\Big\rangle \\
&= \langle \alpha_g, U_g\rangle + \Big\langle T_g^* R_{g^{-1}} T_{e_G}^* \phi^h \beta_h, U_g\Big\rangle + \Big\langle T_{g^{-1} \cdot h}^* \phi(g)\beta_h, V_{g^{-1} \cdot h}\Big\rangle.
\end{aligned}
$$

Thus,

$$
\begin{aligned}
T^*_{(g,g^{-1} \cdot h)}\Upsilon^{-1}(\alpha_g, \beta_h) &= \left(\alpha_g + T_g^* R_{g^{-1}} T_{e_G}^* \phi^h \beta_h, T_{g^{-1} \cdot h}^* \phi(g)\beta_h\right) \\
&= \left(\alpha_g + T_g^* \phi^{g^{-1} \cdot h} \beta_h, T_{g^{-1} \cdot h}^* \phi(g)\beta_h\right)
\end{aligned} \tag{10.4.5}
$$

since $\phi^h \circ R_{g^{-1}} = \phi^{g^{-1} \cdot h}$.

Therefore, since classes in the image are obtained by simply right translating the second component to the identity, we get

$$
\begin{aligned}
[T^*_{(g,g^{-1} \cdot h)}\Upsilon^{-1}]\left([(\alpha_g, \beta_h)]\right) &= \left(\alpha_g + T_g^* \phi^{g^{-1} \cdot h} \beta_h, T_{e_H}^* R_{g^{-1} \cdot h} T_{g^{-1} \cdot h}^* \phi(g)\beta_h\right) \\
&= \left(\alpha_g + T_g^* \phi^{g^{-1} \cdot h} \beta_h, \tilde{\phi}(g)^* T_{e_H}^* R_h \beta_h\right),
\end{aligned}
$$

since $\phi(g) \circ R_{g^{-1} \cdot h} = R_h \circ \phi(g)$. This proves the first equality in (10.4.3). The second one is a consequence of the identity $T_{e_G}\phi^h \circ T_g R_{g^{-1}} = T_g \phi^{g^{-1} \cdot h}$ which follows from the obvious relation $\phi^{g^{-1} \cdot h} = \phi^h \circ R_{g^{-1}}$. ∎

Note that it is not obvious (but it is true) that the right hand side of equation (10.4.3) does not depend on the representative of the class $[(\alpha_g, \beta_h)]$. This may be verified using the explicit formula for the cotangent lift of right translations, regarding H as the subgroup $\{e_G\} \times H \subset G \circledS H$. This map is a special case of the general formula for the cotangent lift of right translations on the group $G \circledS H$. This general formula may be found by a straightforward, but somewhat tedious calculation. The result is

$$
T^*_{(g_1, h_1)(g_2, h_2)} R_{(g_2, h_2)^{-1}} (\alpha_{g_1}, \beta_{h_1})
$$
$$
= \left(T^*_{g_1 g_2} R_{g_2^{-1}} \alpha_{g_1} + \left(T^*_{g_1 g_2} \phi^{g_2^{-1} \cdot h_2^{-1}} \circ T^*_{g_1 \cdot h_2^{-1}} L_{h_1 (g_1 \cdot h_2)} \right) \beta_{h_1}, \right.
$$
$$
\left. T^*_{h_1 (g_1 \cdot h_2)} R_{g_1 \cdot h_2^{-1}} \beta_{h_1} \right). \tag{10.4.6}
$$

Likewise, the cotangent lift of left translation is given by

$$
T^*_{(g_1, h_1)(g_2, h_2)} L_{(g_1, h_1)^{-1}} (\alpha_{g_2}, \beta_{h_2})
$$
$$
= \left(T^*_{g_1 g_2} L_{g_1^{-1}} \alpha_{g_2}, T^*_{h_1 (g_1 \cdot h_2)} \phi(g_1^{-1}) T^*_{(g_1^{-1} \cdot h_1) h_2} L_{g_1^{-1} \cdot h_1^{-1}} \beta_{h_2} \right) \tag{10.4.7}
$$

Specializing to the cotangent lift of right translation by the normal subgroup $\{e_G\} \times H$, we get

$$
(\alpha_g, \beta_h) \cdot k = \left(\alpha_g + \left(T^*_g \phi^{k^{-1}} \circ T^*_{g \cdot k^{-1}} L_{h(g \cdot k)} \right) \beta_h, \ T^*_{h(g \cdot k)} R_{g \cdot k^{-1}} \beta_h \right). \tag{10.4.8}
$$

Thus, the equivalence relation used in the definition of the class $[(\alpha_g, \beta_h)]$ in the left hand side of (10.4.3) is

$$
(\alpha_g, \beta_h) \sim \left(\alpha_g + \left(T^*_g \phi^{k^{-1}} \circ T^*_{g \cdot k^{-1}} L_{h(g \cdot k)} \right) \beta_h, \ T^*_{h(g \cdot k)} R_{g \cdot k^{-1}} \beta_h \right) \tag{10.4.9}
$$

for any $k \in H$.

The push forward of the right G-action to $T^*G \times \mathfrak{g}^*$. Let us specialize (10.4.6) to the subgroup $G \times \{e_H\}$ of $G \circledS H$. We have for any $l \in G$ and $(\alpha_g, \beta_h) \in T^*_{(g,h)}(G \circledS H)$

$$
(\alpha_g, \beta_h) \cdot l = \left(T^*_{gl} R_{l^{-1}} \alpha_g, \beta_h \right). \tag{10.4.10}
$$

Now notice that, by the definitions of the G- and H-actions on $T^*(G \circledS H)$, we have for any $k \in H$ and $l \in G$

$$
((\alpha_g, \beta_h) \cdot k) \cdot l = ((\alpha_g, \beta_k) \cdot (e_G, k)) \cdot (l, e_H) = (\alpha_g, \beta_k) \cdot ((e_G, k)(l, e_H))
$$
$$
= (\alpha_g, \beta_h) \cdot (l, k) = (\alpha_g, \beta_h) \cdot \left((l, e_H)(e_G, l^{-1} \cdot k) \right)
$$
$$
= ((\alpha_g, \beta_h) \cdot (l, e_H)) \cdot (e_G, l^{-1} \cdot k)
$$
$$
= ((\alpha_g, \beta_h) \cdot l) \cdot (l^{-1} \cdot k).
$$

This identity shows that the right G-action on $T^*(G \circledS H)$ induces a right G-action on $(T^*(G \circledS H))/H$ by

$$[(\alpha_g, \beta_h)] \cdot l := [(\alpha_g, \beta_h) \cdot l] = \left[\left(T^*_{gl}R_{l^{-1}}\alpha_g, \beta_h\right)\right]. \tag{10.4.11}$$

The push forward of this action by $\left[T^*\Upsilon^{-1}\right]$ is given by the following formula.

10.4.3 Proposition. *Using the notations and hypotheses of Proposition 10.4.2, it follows that $\left[T^*\Upsilon^{-1}\right] : (T^*(G \circledS H))/H \to T^*G \times \mathfrak{h}^*$ is G-equivariant, where the G-action on $(T^*(G \circledS H))/H$ is given by (10.4.11) and the G-action on $T^*G \times \mathfrak{h}^*$ is the diagonal action*

$$(\alpha_g, \nu) \cdot l := \left(T^*_{gl}R_{l^{-1}}\alpha_g, \tilde{\phi}(l)^*\nu\right) \tag{10.4.12}$$

for any $g, l \in G$, $\alpha_g \in T^*G$, and $\nu \in \mathfrak{h}^*$.

Proof. The proof is a direct verification. For any $(\alpha_g, \beta_h) \in T^*_{(g,h)}(G \circledS H)$ and any $l \in G$ we have by (10.4.11) and (10.4.3)

$$
\begin{aligned}
\left[T^*\Upsilon^{-1}\right]\left([(\alpha_g, \beta_h)] \cdot l\right) &= \left[T^*\Upsilon^{-1}\right]\left([(\alpha_g, \beta_h) \cdot l]\right) \\
&= \left[T^*\Upsilon^{-1}\right]\left(\left[\left(T^*_{gl}R_{l^{-1}}\alpha_g, \beta_h\right)\right]\right) \\
&= \left(T^*_{gl}R_{l^{-1}}\alpha_g + \left(T^*_{gl}R_{(gl)^{-1}} \circ T^*_{e_G}\phi^h\right)\beta_h, \tilde{\phi}(gl)^*T^*_{e_H}R_h\beta_h\right) \\
&= \left(T^*_{gl}R_{l^{-1}}\left(\alpha_g + \left(T^*_g R_{g^{-1}} \circ T^*_{e_G}\phi^h\right)\beta_h\right), \tilde{\phi}(l)^*\tilde{\phi}(g)^*T^*_{e_H}R_h\beta_h\right) \\
&= \left(\alpha_g + \left(T^*_g R_{g^{-1}} \circ T^*_{e_G}\phi^h\right)\beta_h, \tilde{\phi}(g)^*T^*_{e_H}R_h\beta_h\right) \cdot l \\
&= \left[T^*\Upsilon^{-1}\right]\left([(\alpha_g, \beta_h)]\right) \cdot l
\end{aligned}
$$

as required. ∎

Second Stage Reduction: The Poisson Case. We next apply the results of the preceding section to perform the second stage reduction of the cotangent lift of the right action of $G \circledS H$ on itself. After the first stage reduction, we have arrived at the Poisson manifold $T^*G \times \mathfrak{h}^*$ with the sum Poisson structure and the diagonal G-action. Now we can perform the second stage reduction by making use of Proposition 10.3.1 and formula (10.3.1).

10.4.4 Proposition. *Give $T^*G \times \mathfrak{h}^*$ the sum of the canonical and the plus Lie–Poisson structure on $\mathfrak{h}^*$. The map $\psi : T^*G \times \mathfrak{h}^* \to \mathfrak{g}^* \times \mathfrak{h}^*$ given by $\psi(\alpha_g, \nu) = \left(T^*_{e_G}R_g\alpha_g, \tilde{\phi}(g^{-1})^*\nu\right)$, induces a Poisson diffeomorphism*

$[\psi] : (T^*G \times \mathfrak{h}^*)/G \to \mathfrak{g}^* \times \mathfrak{h}^*$, *where* $\mathfrak{g}^* \times \mathfrak{h}^*$ *has the bracket*

$$\{F, K\} = \{F, K\}_{\mathfrak{g}^*} + \{F, K\}_{\mathfrak{h}^*} - \mathbf{d}_\nu F \cdot \left(\frac{\delta K}{\delta \mu}\right)_{\mathfrak{h}^*} + \mathbf{d}_\nu K \cdot \left(\frac{\delta F}{\delta \mu}\right)_{\mathfrak{h}^*}$$

(10.4.13)

$$= \left\langle \mu, \left[\frac{\delta F}{\delta \mu}, \frac{\delta K}{\delta \mu}\right] \right\rangle + \left\langle \nu, \left[\frac{\delta F}{\delta \nu}, \frac{\delta K}{\delta \nu}\right] \right\rangle$$
$$- \left\langle \nu, \frac{\delta K}{\delta \mu} \cdot \frac{\delta F}{\delta \nu} \right\rangle + \left\langle \nu, \frac{\delta F}{\delta \mu} \cdot \frac{\delta K}{\delta \nu} \right\rangle$$

(10.4.14)

where $(\mu, \nu) \in \mathfrak{g}^* \times \mathfrak{h}^*$, *which is the* (+)*-Lie-Poisson bracket on* $(\mathfrak{g} \, \text{S} \, \mathfrak{h})^*$ *by* (6.4.5).

Proof.   Equation (10.4.13) is a special instance of equation (10.3.3). Thus, to complete the proof, we only need to show how equation (10.4.14) follows from (10.4.13). To see this, note that for $\xi \in \mathfrak{g}$ and $\nu \in \mathfrak{h}^*$,

$$\xi_{\mathfrak{h}^*}(\nu) = \left.\frac{d}{dt}\right|_{t=0} \tilde{\phi}(\exp t\xi)^* \nu = \tilde{\phi}'(\xi)^* \nu$$

and so

$$\mathbf{d}_\nu K \cdot \left(\frac{\delta F}{\delta \mu}\right)_{\mathfrak{h}^*} = \left\langle \left(\frac{\delta F}{\delta \mu}\right)_{\mathfrak{h}^*}(\nu), \frac{\delta K}{\delta \nu} \right\rangle = \left\langle \tilde{\phi}'\left(\frac{\delta F}{\delta \mu}\right)^* \nu, \frac{\delta K}{\delta \nu} \right\rangle.$$

However,

$$\left\langle \tilde{\phi}'(\xi)^* \nu, \eta \right\rangle = \langle \nu, \xi \cdot \eta \rangle,$$

and thus

$$\mathbf{d}_\nu K \cdot \left(\frac{\delta F}{\delta \mu}\right)_{\mathfrak{h}^*} = \left\langle \nu, \frac{\delta F}{\delta \mu} \cdot \frac{\delta K}{\delta \nu} \right\rangle.$$

Substituting this and a similar expression with F and K interchanged into (10.4.13) gives (10.4.14) as required.   ■

Remarks.

1. That this two stage procedure produces the Lie-Poisson bracket on the dual of the Lie algebra of the semidirect product is of course consistent with the general Poisson reduction by stages result. On the other hand, how this comes about through the two-stage process is by an interesting and non-obvious route.

2. Consistent with the stages construction, the composition of the three maps

$$\psi \circ (I \times \rho_H) \circ T^*\Upsilon^{-1} : T^*(G \, \text{S} \, H) \to \mathfrak{g}^* \times \mathfrak{h}^*$$

is simply right translation to the identity on $T^*(G \circledS H)$; that is,

$$\left(\psi \circ (I \times \rho_H) \circ T^*\Upsilon^{-1}\right)(\alpha_g, \beta_h) = T^*_{(e_G, e_H)} R_{(g,h)}(\alpha_g, \beta_h).$$

Here, I is the identity map on T^*G and $\rho_H : T^*H \to \mathfrak{h}^*$ given by $\rho_H(\beta_h) = T^*_{e_H} R_h \beta_h$, is right translation to the identity on T^*H.

10.5 First Stage Reduction: General Semidirect Products

In the case of the action of $G \circledS V$ on its cotangent bundle, where V is a vector space, the first stage reduction gives T^*G endowed with the canonical symplectic form (see Theorem 4.3.8). In the general case of $G \circledS H$ considered in this section, the first stage reduction will yield $T^*G \times \mathcal{O}_\nu$, where $\mathcal{O}_\nu$ is a coadjoint orbit of H, endowed with the product symplectic structure.

The General Strategy.   Recall from Proposition 10.4.2 that we have a Poisson diffeomorphism

$$[T^*\Upsilon^{-1}] : (T^*(G \circledS H))/H \to T^*G \times \mathfrak{h}^*.$$

Being a Poisson diffeomorphism, it must map symplectic leaves to symplectic leaves. Up to connected components, symplectic leaves are the symplectic reduced spaces, and so it is reasonable to expect that this map gives a symplectic diffeomorphism from the first reduced space for $T^*(G \circledS H)$, that is, reduction by the normal subgroup H to the symplectic manifold $T^*G \times \mathcal{O}$, where $\mathcal{O}$ is a coadjoint orbit in $\mathfrak{h}^*$. This is in fact literally true and the diffeomorphism can be given explicitly, as we show next.

Notation.   We next introduce the objects needed to carry out the first stage reduction explicitly. First of all, for the reduction of $T^*(G \circledS H)$, the cotangent lifted action of H is given explicitly by equation (10.4.8). The infinitesimal generator of the H-action on $G \circledS H$ is given for $\eta \in \mathfrak{h}$ by

$$\eta_{G \circledS H}(g, h) = \left.\frac{d}{dt}\right|_{t=0} (g, h) \cdot \exp t\eta = \left.\frac{d}{dt}\right|_{t=0} (g, h(g \cdot \exp t\eta))$$
$$= \left(0, T_{e_H} L_h(\tilde{\phi}(g)\eta)\right) \tag{10.5.1}$$

The corresponding momentum map $\mathbf{J}_H : T^*(G \circledS H) \to \mathfrak{h}^*$ for the cotangent lift of right H-translation on $G \circledS H$ given in (10.4.8) is computed, according to general theory and (10.5.1), to be

$$\langle \mathbf{J}_H(\alpha_g, \beta_h), \eta \rangle = \langle (\alpha_g, \beta_h), \eta_{G \circledS H}(g, h) \rangle = \left\langle \beta_h, T_{e_H} L_h(\tilde{\phi}(g)\eta) \right\rangle.$$

Thus, we get
$$\mathbf{J}_H(\alpha_g, \beta_h) = \tilde{\phi}(g)^* T^*_{e_H} L_h \beta_h. \tag{10.5.2}$$

Since the H-action (10.4.8) that defines this momentum map is free, it follows that all elements in the range of $\mathbf{J}_H$ are regular values. Therefore, if $\nu \in \mathfrak{h}^*$ the level set

$$\mathbf{J}_H^{-1}(\nu) = \left\{ \left(\alpha_g, T^*_h L_{h^{-1}} \tilde{\phi}(g^{-1})^* \nu\right) \, \middle| \, \alpha_g \in T^*_g G, h \in H \right\}. \tag{10.5.3}$$

is a smooth submanifold in $T^*(G \circledS H)$.

The behavior of $\mathbf{J}_H$ under right translation is given by

$$\mathbf{J}_H \circ T^* R_{(l,n)^{-1}} = \tilde{\phi}(l)^* \circ \mathrm{Ad}^*_n \circ \mathbf{J}_H \tag{10.5.4}$$

for all $(l,n) \in G \circledS H$. To prove this formula we use (10.4.6) and (10.5.2) to get for any $(\alpha_g, \beta_h) \in T^*_{(g,h)}(G \circledS H)$

$$\left(\mathbf{J}_H \circ T^* R_{(l,n)^{-1}}\right)(\alpha_g, \beta_h) = \tilde{\phi}(gl)^* T^*_{e_H} L_{h(g \cdot n)} T^*_{h(g \cdot n)} R_{g \cdot n^{-1}} \beta_h$$
$$= \tilde{\phi}(l)^* \left(T_{h(g \cdot n)} R_{g \cdot n^{-1}} \circ T_{e_H} L_{h(g \cdot n)} \circ \tilde{\phi}(g)\right)^* \beta_h$$
$$= \tilde{\phi}(l)^* T^*_{e_H} \left(R_{g \cdot n^{-1}} \circ L_{h(g \cdot n)} \circ \phi(g)\right) \beta_h.$$

However, $R_{g \cdot n^{-1}} \circ L_{h(g \cdot n)} \circ \phi(g) = L_h \circ \phi(g) \circ \mathrm{AD}_n$ as an easy verification shows. Therefore we can continue the computation to get

$$\tilde{\phi}(l)^* T^*_{e_H} \left(L_h \circ \phi(g) \circ \mathrm{AD}_n\right) \beta_h = \tilde{\phi}(l)^* \mathrm{Ad}^*_n \tilde{\phi}(g)^* T^*_{e_H} L_h \beta_h$$
$$= \tilde{\phi}(l)^* \mathrm{Ad}^*_n \mathbf{J}_H(\alpha_g, \beta_h)$$

which proves (10.5.4).

Formulas (10.5.4) and (10.6.3) show that, as expected, $(G \circledS H)_\nu$ acts on $\mathbf{J}_H^{-1}(\nu)$, where the action is cotangent lift of right translation given in (10.4.6). Let H_ν be the isotropy group for the H-coadjoint action on $\mathfrak{h}^*$ and form the corresponding symplectic reduced space $\mathbf{J}_H^{-1}(\nu)/H_\nu$. Let

$$\pi_\nu^H : \mathbf{J}_H^{-1}(\nu) \to \mathbf{J}_H^{-1}(\nu)/H_\nu$$

be the projection to the reduced space.

Let $\mathbf{J}_H^R : T^*H \to \mathfrak{h}^*$, given by $\mathbf{J}_H^R(\beta_h) = T^*_{e_H} L_h \beta_h$, be the momentum map for the cotangent lifted action of H on T^*H given by right translations. Consider the projection associated with the reduction of the H-factor in $T^*G \times T^*H$, namely the map

$$\overline{\pi}_\nu^H : T^*G \times (\mathbf{J}_H^R)^{-1}(\nu) \to T^*G \times \mathcal{O}_\nu$$

defined by

$$\overline{\pi}_\nu^H (\alpha_g, T^*_h L_{h^{-1}} \nu) = (\alpha_g, \mathrm{Ad}^*_{h^{-1}} \nu). \tag{10.5.5}$$

The First Stage Symplectic Reduced Space.   We can now state the main structure theorem for the first stage reduction for the action of $G \, \text{ⓢ} \, H$ on its cotangent bundle, namely the reduction by the H-action.

10.5.1 Theorem.   *Using the preceding notation, the diffeomorphism*

$$T^*\Upsilon^{-1} : T^*(G \, \text{ⓢ} \, H) \to T^*G \times T^*H$$

given by (10.4.5), *restricts to a diffeomorphism*

$$(T^*\Upsilon^{-1})_\nu : \mathbf{J}_H^{-1}(\nu) \to T^*G \times (\mathbf{J}_H^R)^{-1}(\nu).$$

This diffeomorphism is H_ν-equivariant and descends to a symplectic diffeomorphism

$$[T^*\Upsilon^{-1}]_\nu : \mathbf{J}_H^{-1}(\nu)/H_\nu \to T^*G \times \mathcal{O}_\nu$$

given by

$$[T^*\Upsilon^{-1}]_\nu \left[\left(\alpha_g, T_h^* L_{h^{-1}} \tilde{\phi}(g^{-1})^* \nu \right) \right]$$
$$= \left(\alpha_g + T_g^* R_{g^{-1}} T_{e_G}^* \phi^h T_h^* L_{h^{-1}} \tilde{\phi}(g^{-1})^* \nu, \text{Ad}_{g^{-1} \cdot h^{-1}}^* \nu \right), \qquad (10.5.6)$$

where $\alpha_g \in T_g^* G$ and $h \in H$. Thus, the first symplectic reduced space $\mathbf{J}_H^{-1}(\nu)/H_\nu$ is symplectically diffeomorphic to the space $T^*G \times \mathcal{O}_\nu$ with the product symplectic structure $\omega_G \oplus \omega_{\mathcal{O}_\nu}^+$, where ω_G is the canonical symplectic form on T^*G and $\omega_{\mathcal{O}_\nu}^+$ is the plus coadjoint orbit symplectic form on $\mathcal{O}_\nu$.

Proof.   We first claim that $T^*\Upsilon^{-1}$ maps $\mathbf{J}_H^{-1}(\nu)$ to $T^*G \times (\mathbf{J}_H^R)^{-1}(\nu)$. A typical element of $\mathbf{J}_H^{-1}(\nu)$ has, according to equation (10.5.3), the form $\left(\alpha_g, T_h^* L_{h^{-1}} \tilde{\phi}(g^{-1})^* \nu \right)$. Evaluating $T^*\Upsilon^{-1}$, which is given by (10.4.5) on this element gives

$$T_{(g,g^{-1}\cdot h)}^* \Upsilon^{-1} \left(\alpha_g, T_h^* L_{h^{-1}} \tilde{\phi}(g^{-1})^* \nu \right)$$
$$= \left(\alpha_g + T_g^* R_{g^{-1}} T_{e_G}^* \phi^h T_h^* L_{h^{-1}} \tilde{\phi}(g^{-1})^* \nu, T_{g^{-1}\cdot h}^* \phi(g) T_h^* L_{h^{-1}} \tilde{\phi}(g^{-1})^* \nu \right)$$
$$= \left(\alpha_g + T_g^* R_{g^{-1}} T_{e_G}^* \phi^h T_h^* L_{h^{-1}} \tilde{\phi}(g^{-1})^* \nu, T_{g^{-1}\cdot h}^* L_{g^{-1}\cdot h^{-1}} \nu \right), \qquad (10.5.7)$$

since, as is easily checked from the definitions, $\phi(g^{-1}) \circ L_{h^{-1}} \circ \phi(g) = L_{g^{-1}\cdot h^{-1}}$. This proves our claim.

Next from the form of this expression and the arbitrariness of α_g, it is clear that this map is onto. The map is injective because it is the restriction of a diffeomorphism. It has a smooth inverse, namely the restriction of $T^*\Upsilon$, and so it is a diffeomorphism.

From equation (10.4.2), the map Υ is H-equivariant, and therefore, so is $T^*\Upsilon^{-1}$. Thus, its restriction $(T^*\Upsilon^{-1})_\nu$ is H_ν-equivariant. Thus, the induced map on the quotient $[T^*\Upsilon^{-1}]_\nu$ is a diffeomorphism. Since the corresponding

map on the Poisson level is a Poisson diffeomorphism by Proposition 10.4.2, this map on the reduced spaces is symplectic (they are the symplectic leaves up to connected components).

Equation (10.5.6) is proved as follows. By construction, we have $[T^*\Upsilon^{-1}]_\nu \circ \pi_\nu^H = \bar{\pi}_\nu^H \circ (T^*\Upsilon^{-1})_\nu$, which in conjunction with (10.5.5) gives the result. ∎

Example: $\mathcal{T}\,\textcircled{S}\,\mathcal{U}$. Let $\mathcal{T}$ denote the group of $n \times n$ upper triangular matrices and $\mathcal{U}$ the subgroup of upper triangular matrices that have all entries on the diagonal equal to 1. It is readily checked that $\mathcal{U}$ is a normal subgroup of $\mathcal{T}$ and hence $\mathcal{T}$ acts on $\mathcal{U}$ by conjugation. Form the semidirect product $\mathcal{T}\,\textcircled{S}\,\mathcal{U}$; as in the general theory, multiplication and inversion are given by $(t_1, u_1)(t_2, u_2) = (t_1 t_2, u_1 t_1 u_2 t_1^{-1})$ and $(t, u)^{-1} = (t^{-1}, t^{-1}u^{-1}t)$ respectively. The Lie algebras $\mathfrak{t}$ of $\mathcal{T}$ and $\mathfrak{u}$ of $\mathcal{U}$ consist of upper and strictly upper triangular matrices respectively; $\mathfrak{u}$ is an ideal in $\mathfrak{t}$. The Lie algebra of $\mathcal{T}\,\textcircled{S}\,\mathcal{U}$ is the semidirect product Lie algebra $\mathfrak{t}\,\textcircled{S}\,\mathfrak{u}$ whose bracket is given by

$$[(\xi_1, \eta_1), (\xi_2, \eta_2)] = ([\xi_1, \xi_2], [\xi_1, \eta_2] - [\xi_2, \eta_1] + [\eta_1, \eta_2]),$$

for any $\xi_1, \xi_2 \in \mathfrak{t}$ and $\eta_1, \eta_2 \in \mathfrak{u}$.

The nondegenerate pairing $\langle A, B \rangle := \operatorname{trace}(AB^T)$ on $\mathfrak{gl}(n)$ identifies $\mathfrak{gl}(n)$ with its dual $\mathfrak{gl}(n)^*$. The same pairing restricts to a nondegenerate pairing on $\mathfrak{t}$ and on $\mathfrak{u}$. Thus, we identify $\mathfrak{t}^*$ with $\mathfrak{t}$ and $\mathfrak{u}^*$ with $\mathfrak{u}$.

Making use of Theorem 10.5.1, we see that *the first stage reduced Poisson manifold is $T^*\mathcal{T} \times \mathfrak{u}^*$ endowed with the product Poisson structure and, if $\nu \in \mathfrak{u}^* \equiv \mathfrak{u}$, the first stage symplectic reduced space at ν is $(T^*\mathcal{T} \times \mathcal{O}_\nu, \omega_\mathcal{T} \oplus \omega_{\mathcal{O}_\nu}^+)$, where $\mathcal{O}_\nu$ is the coadjoint $\mathcal{U}$-orbit through ν.*

The coadjoint action of $\mathcal{U}$ on $\mathfrak{u}^*$ is computed explicitly as follows. Let $\Pi_{\mathfrak{u}} : \mathfrak{gl}(n) \to \mathfrak{u}$ be the projection that maps a matrix to its strictly upper triangular part. If $u \in \mathcal{U}$, $\nu \in \mathfrak{u}^* \equiv \mathfrak{u}$, and $\eta \in \mathfrak{u}$, we have

$$\langle \operatorname{Ad}_u^* \nu, \eta \rangle = \langle \nu, \operatorname{Ad}_u \eta \rangle = \operatorname{trace}\left(\nu(u\eta u^{-1})^T\right) = \operatorname{trace}\left(\nu(u^{-1})^T \eta^T u^T\right)$$
$$= \operatorname{trace}\left(u^T \nu(u^{-1})^T \eta^T\right) = \operatorname{trace}\left(\Pi_{\mathfrak{u}}\left(u^T \nu(u^{-1})^T\right)\eta^T\right)$$
$$= \left\langle \Pi_{\mathfrak{u}}\left(u^T \nu(u^{-1})^T\right), \eta \right\rangle$$

which gives the formula

$$\operatorname{Ad}_u^* \nu = \Pi_{\mathfrak{u}}\left(u^T \nu(u^{-1})^T\right). \tag{10.5.8}$$

To get a feeling of the coadjoint action, let us classify all coadjoint isotropy subgroups of $\mathcal{U}$ and hence all coadjoint orbits in $\mathfrak{u}^*$ for the case $n = 3$. If

$$u = \begin{bmatrix} 1 & u_{12} & u_{13} \\ 0 & 1 & u_{23} \\ 0 & 0 & 1 \end{bmatrix} \quad \text{then} \quad u^{-1} = \begin{bmatrix} 1 & -u_{12} & u_{12}u_{23} - u_{13} \\ 0 & 1 & -u_{23} \\ 0 & 0 & 1 \end{bmatrix}.$$

Thus, if

$$\nu = \begin{bmatrix} 0 & \nu_{12} & \nu_{13} \\ 0 & 0 & \nu_{23} \\ 0 & 0 & 0 \end{bmatrix} \in \mathfrak{u}^* \equiv \mathfrak{u}$$

then

$$\mathrm{Ad}_u^* \nu = \Pi_{\mathfrak{u}} \left(\begin{bmatrix} 1 & 0 & 0 \\ u_{12} & 1 & 0 \\ u_{13} & u_{23} & 1 \end{bmatrix} \begin{bmatrix} 0 & \nu_{12} & \nu_{13} \\ 0 & 0 & \nu_{23} \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ -u_{12} & 1 & 0 \\ u_{12}u_{23} - u_{13} & -u_{23} & 1 \end{bmatrix} \right)$$

$$= \begin{bmatrix} 0 & \nu_{12} - u_{23}\nu_{13} & \nu_{13} \\ 0 & 0 & \nu_{23} + u_{12}\nu_{13} \\ 0 & 0 & 0 \end{bmatrix}. \tag{10.5.9}$$

This shows that

- if $\nu_{13} \neq 0$ then

$$\mathcal{U}_\nu = \left\{ \begin{bmatrix} 1 & 0 & u_{13} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \,\middle|\, u_{13} \in \mathbb{R} \right\}$$

  and the corresponding coadjoint orbit is two dimensional and equal to

$$\mathcal{O}_\nu = \left\{ \begin{bmatrix} 0 & x & \nu_{13} \\ 0 & 0 & y \\ 0 & 0 & 0 \end{bmatrix} \,\middle|\, x, y \in \mathbb{R} \right\}.$$

- if $\nu_{13} = 0$, then $\mathcal{U}_\nu = \mathcal{U}$ and the corresponding coadjoint orbit is zero dimensional and equal to $\mathcal{O}_\nu = \{\nu\}$.

Similarly, in the 4×4 case, we write

$$u = \begin{bmatrix} 1 & u_{12} & u_{13} & u_{14} \\ 0 & 1 & u_{23} & u_{24} \\ 0 & 0 & 1 & u_{34} \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

and compute that

$$u^{-1} = \begin{bmatrix} 1 & -u_{12} & u_{12}u_{23} - u_{13} & -u_{12}u_{23}u_{34} + u_{12}u_{24} + u_{13}u_{34} - u_{14} \\ 0 & 1 & -u_{23} & u_{23}u_{34} - u_{24} \\ 0 & 0 & 1 & -u_{34} \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Thus, if

$$\nu = \begin{bmatrix} 0 & \nu_{12} & \nu_{13} & \nu_{14} \\ 0 & 0 & \nu_{23} & \nu_{24} \\ 0 & 0 & 0 & \nu_{34} \\ 0 & 0 & 0 & 0 \end{bmatrix} \in \mathfrak{u}^* \equiv \mathfrak{u},$$

then, computing as in the 3×3 case, one finds that

$$
\mathrm{Ad}_u^* \nu =
\begin{bmatrix}
0 & (\mathrm{Ad}_u^* \nu)_{12} & \nu_{13} - \nu_{14} u_{34} & \nu_{14} \\
0 & 0 & (\mathrm{Ad}_u^* \nu)_{23} & \nu_{14} u_{12} + \nu_{24} \\
0 & 0 & 0 & (\mathrm{Ad}_u^* \nu)_{34} \\
0 & 0 & 0 & 0
\end{bmatrix},
$$

where

$$
\begin{aligned}
(\mathrm{Ad}_u^* \nu)_{12} &= \nu_{12} - \nu_{13} u_{23} + \nu_{14}(u_{23} u_{34} - u_{24}) \\
(\mathrm{Ad}_u^* \nu)_{23} &= \nu_{23} + \nu_{13} u_{12} - \nu_{14} u_{12} u_{34} - \nu_{24} u_{34} \\
(\mathrm{Ad}_u^* \nu)_{34} &= \nu_{34} + \nu_{14} u_{13} + \nu_{24} u_{23}.
\end{aligned}
$$

This shows that the conditions for isotropy, that is $\mathrm{Ad}_u^* \nu = \nu$ are

$$
\begin{aligned}
0 &= -\nu_{13} u_{23} + \nu_{14}(u_{23} u_{34} - u_{24}) \\
0 &= \nu_{13} u_{12} - \nu_{14} u_{12} u_{34} - \nu_{24} u_{34} \\
0 &= \nu_{14} u_{13} + \nu_{24} u_{23} \\
0 &= \nu_{14} u_{34} \\
0 &= \nu_{14} u_{12}.
\end{aligned}
$$

We can now classify the coadjoint orbits of $\mathcal{U}$ as follows:

- If $\nu_{14} \neq 0$ then $u_{12} = u_{34} = 0$ and we get the two dimensional isotropy group

$$
\mathcal{U}_\nu =
\left\{
\begin{bmatrix}
1 & 0 & -\frac{\nu_{24}}{\nu_{14}} u_{23} & u_{14} \\
0 & 1 & u_{23} & -\frac{\nu_{13}}{\nu_{14}} u_{23} \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{bmatrix}
\;\middle|\; u_{23}, u_{14} \in \mathbb{R}
\right\}
$$

and the corresponding coadjoint orbit is four dimensional and equal to

$$
\mathcal{O}_\nu =
\left\{
\begin{bmatrix}
0 & x & u & \nu_{14} \\
0 & 0 & \frac{1}{\nu_{14}}(uv - \nu_{13}\nu_{24} + \nu_{23}\nu_{14}) & v \\
0 & 0 & 0 & y \\
0 & 0 & 0 & 0
\end{bmatrix}
\;\middle|\; x, y, u, v \in \mathbb{R}
\right\}.
$$

- If $\nu_{14} = 0$, then the isotropy conditions become

$$0 = -\nu_{13}u_{23}$$
$$0 = \nu_{13}u_{12} - \nu_{24}u_{34}$$
$$0 = \nu_{24}u_{23}.$$

  – If $\nu_{13}\nu_{24} \neq 0$, then $u_{23} = 0$. In this case, the isotropy group is the four dimensional group

$$\mathcal{U}_\nu = \left\{ \begin{bmatrix} 1 & u_{12} & u_{13} & u_{14} \\ 0 & 1 & 0 & u_{24} \\ 0 & 0 & 1 & u_{34} \\ 0 & 0 & 0 & 1 \end{bmatrix} \,\middle|\, \nu_{13}u_{12} - \nu_{24}u_{34} = 0 \right\}.$$

  The corresponding two dimensional coadjoint orbits are

$$\mathcal{O}_\nu = \left\{ \begin{bmatrix} 0 & u & \nu_{13} & 0 \\ 0 & 0 & x & \nu_{24} \\ 0 & 0 & 0 & v \\ 0 & 0 & 0 & 0 \end{bmatrix} \,\middle|\, \nu_{24}u + \nu_{13}v = \nu_{12}\nu_{24} + \nu_{13}\nu_{34} \right\}.$$

  – If $\nu_{13}\nu_{24} = 0$, then $\mathcal{U}_\nu = \mathcal{U}$ and the coadjoint orbit is a point: $\mathcal{O}_\nu = \{\nu\}$.

We shall continue this example in the last section of this chapter, where the second stage reduction is carried out therefore giving the coadjoint orbits of the semidirect product $\mathcal{T} \,\circledS\, \mathcal{U}$.

Remark. We notice that the example $G \,\circledS\, G$, with G acting on itself by conjugation, that is, $\phi(g) = \mathrm{AD}_g$ for any $g \in G$, does not give an interesting example. The reason is that the map $\Phi : (g, h) \in G \,\circledS\, G \mapsto (g, hg) \in G \times G$ is easily checked to be a Lie group isomorphism from the semidirect product to the direct product.

10.6 Second Stage Reduction: General Semidirect Products

In this section we shall carry out the second stage reduction for general semidirect products. As we saw in the preceding section, the first stage reduction of $T^*(G \,\circledS\, H)$ at $\nu \in \mathfrak{h}^*$ by the lift of right translation of H gives the symplectic manifold $(T^*G \times \mathcal{O}_\nu, \omega_G \oplus \omega_{\mathcal{O}_\nu}^+)$. Using the general theory, we will determine the group, denoted G^ν, that acts on this first reduced space. The second stage reduction is possible, in principle, since in §6.4 it was shown that general semidirect products satisfy the stages hypothesis. By the general reduction by stages theorem, when the second stage reduction is carried out, one obtains the result of reduction by the big group $G \,\circledS\, H$, namely the coadjoint orbits in $(\mathfrak{g} \,\circledS\, \mathfrak{h})^*$.

Adjoint and Coadjoint Actions.   The adjoint action of $S = G \circledS H$ on $\mathfrak{g} \circledS \mathfrak{h}$ given in §6.4 is

$$\mathrm{Ad}_{(g,h)}(\xi, \eta) = \Big(\mathrm{Ad}_g \xi, \ \Big(\mathrm{Ad}_h \circ \tilde{\phi}(g) \Big)(\eta) + T_{h^{-1}} L_h \left((\mathrm{Ad}_g \xi)_H (h^{-1}) \right) \Big),$$

where $\xi \in \mathfrak{g}$, $\eta \in \mathfrak{h}$, $g \in G$, $h \in H$. Its restriction to $\mathfrak{h}$ is therefore given by

$$\mathrm{Ad}^H_{(g,h)} \zeta = \Big(\mathrm{Ad}_h \circ \tilde{\phi}(g) \Big)(\zeta) \qquad (10.6.1)$$

for $(g, h) \in G \circledS H$ and $\zeta \in \mathfrak{h}$. The derivative of Ad^H at the identity is $\mathrm{ad}_{(\xi,\eta)} \zeta = [\eta, \zeta] + \xi \cdot \zeta$, for all $(\xi, \eta) \in \mathfrak{s}$ and $\zeta \in \mathfrak{h}$.

The coadjoint action is given by (6.4.3) and, according to (5.2.1), its second component is the (left) dual of the S-action (10.6.1). Thus, the action of S on $\mathfrak{h}^*$ is given by

$$(g, h) \cdot \nu = \tilde{\phi}(g^{-1})^* \mathrm{Ad}^*_{g^{-1} \cdot h^{-1}} \nu = \mathrm{Ad}^*_{h^{-1}} \tilde{\phi}(g^{-1})^* \nu, \qquad (10.6.2)$$

where $(g, h) \in S$ and $\nu \in \mathfrak{h}^*$. Therefore, the isotropy subgroup of S at $\nu \in \mathfrak{h}^*$ equals

$$\begin{aligned} S_\nu &= \{ (g, h) \in S \mid \tilde{\phi}(g^{-1})^* \mathrm{Ad}^*_{g^{-1} \cdot h^{-1}} \nu = \nu \} \\ &= \{ (g, h) \in S \mid \mathrm{Ad}^*_{h^{-1}} \tilde{\phi}(g^{-1})^* \nu = \nu \}. \end{aligned} \qquad (10.6.3)$$

Its Lie algebra is

$$\mathfrak{s}_\nu = \{ (\xi, \eta) \in \mathfrak{s} \mid \mathrm{ad}^*_\eta \nu + \tilde{\phi}'(\xi)^* \nu = 0 \}. \qquad (10.6.4)$$

Indeed, taking the derivative of (10.6.2) relative to (g, h) at the identity, we get the (left) induced Lie algebra action of $\mathfrak{s}$ on $\mathfrak{h}^*$

$$(\xi, \eta) \cdot \nu = -\tilde{\phi}'(\xi)^* \nu - \mathrm{ad}^*_\eta \nu, \qquad (10.6.5)$$

where $\xi \in \mathfrak{g}$, $\eta \in \mathfrak{h}$, and $\nu \in \mathfrak{h}^*$. As expected, this is the second component of the coadjoint action of $\mathfrak{s}$ on its dual $\mathfrak{s}^*$ given in (6.4.4) (the minus sign is necessary since the Lie algebra action is given by the negative of the dual of the adjoint representation). Therefore, the Lie algebra $\mathfrak{s}_\nu$ of S_ν is given by (10.6.4).

Structure of M_ν.   Recall that key ingredients in the general theory of reduction by stages are the groups M_ν and N_ν. In our case, these are the groups S_ν and H_ν. Things would be simplified if S_ν had a simple structure, as it has in the case of a semidirect product with a vector space. We show now that a naive guess for this structure is not true.

Let $H_\nu := \{ h \in H \mid \mathrm{Ad}^*_h \nu = \nu \}$ be the coadjoint isotropy subgroup of H and $G_\nu := \{ g \in G \mid \tilde{\phi}(g^{-1})^* \nu = \nu \}$ the isotropy subgroup of G relative to the action $\tilde{\phi}(g^{-1}) : \mathfrak{h}^* \to \mathfrak{h}^*$ at $\nu \in \mathfrak{h}^*$. We claim that if $h \in H_\nu$ and

$g \in G_\nu$ then $g \cdot h \in H_\nu$. Indeed, since $\mathrm{AD}_{g \cdot h} = \phi(g) \circ \mathrm{AD}_h \circ \phi(g^{-1})$, where $\mathrm{AD}_h(h') := hh'h^{-1}$ for any $h, h' \in H$, we get

$$\mathrm{Ad}^*_{g \cdot h} = \tilde{\phi}(g^{-1})^* \circ \mathrm{Ad}^*_h \circ \tilde{\phi}(g)^* \tag{10.6.6}$$

and hence $\mathrm{Ad}^*_{g \cdot h}\,\nu = \tilde{\phi}(g^{-1})^* \mathrm{Ad}^*_h\, \tilde{\phi}(g)^*\nu = \nu$, since $g \in G_\nu$ and $h \in H_\nu$. Therefore the action $\phi : G \times H \to H$ by H-group homomorphisms naturally restricts to an action, still denoted ϕ, of G_ν on H_ν and so the semidirect product $G_\nu \, \circledS \, H_\nu$ is a closed subgroup of $S = G \circledS H$. From (10.6.3) it immediately follows that $G_\nu \, \circledS \, H_\nu$ is a closed subgroup of S_ν. One might hope that one has equality.

However, in general, $G_\nu \, \circledS \, H_\nu \subsetneqq S_\nu$ as the following counterexample shows. Consider the semidirect product $\mathcal{T} \circledS \mathcal{U}$ introduced in the previous section. It is easily verified that

$$\tilde{\phi}(t^{-1})^*\nu = \Pi_{\mathfrak{u}}\left((t^{-1})^T \nu t^T\right), \tag{10.6.7}$$

where $\nu \in \mathfrak{u}^*$ and $t \in \mathcal{T}$. Since

$$\begin{bmatrix} t_{11} & t_{12} & t_{13} \\ 0 & t_{22} & t_{23} \\ 0 & 0 & t_{33} \end{bmatrix}^{-1} = \begin{bmatrix} \frac{1}{t_{11}} & -\frac{t_{12}}{t_{11}t_{22}} & \frac{t_{12}t_{23}}{t_{11}t_{22}t_{33}} - \frac{t_{13}}{t_{11}t_{33}} \\ 0 & \frac{1}{t_{22}} & -\frac{t_{23}}{t_{22}t_{33}} \\ 0 & 0 & \frac{1}{t_{33}} \end{bmatrix}$$

we get for any

$$\nu = \begin{bmatrix} 0 & \nu_{12} & \nu_{13} \\ 0 & 0 & \nu_{23} \\ 0 & 0 & 0 \end{bmatrix} \in \mathfrak{u}^*$$

the formula

$$\Pi_{\mathfrak{u}}\left((t^{-1})^T \nu t^T\right) = \begin{bmatrix} 0 & \frac{t_{22}}{t_{11}}\nu_{12} + \frac{t_{23}}{t_{11}}\nu_{13} & \frac{t_{33}}{t_{11}}\nu_{13} \\ 0 & 0 & -\frac{t_{12}t_{33}}{t_{11}t_{22}}\nu_{13} + \frac{t_{33}}{t_{22}}\nu_{23} \\ 0 & 0 & 0 \end{bmatrix}. \tag{10.6.8}$$

Therefore, by (10.5.9), we get

$$\mathrm{Ad}^*_{u^{-1}}\, \tilde{\phi}(t^{-1})^*\nu = \begin{bmatrix} 0 & \bar{\nu}_{12} + u_{23}\bar{\nu}_{13} & \bar{\nu}_{13} \\ 0 & 0 & \bar{\nu}_{23} - u_{12}\bar{\nu}_{13} \\ 0 & 0 & 0 \end{bmatrix},$$

where

$$\bar{\nu}_{12} = \frac{t_{22}}{t_{11}}\nu_{12} + \frac{t_{23}}{t_{11}}\nu_{13}$$

$$\bar{\nu}_{13} = \frac{t_{33}}{t_{11}}\nu_{13}$$

$$\bar{\nu}_{23} = -\frac{t_{12}t_{33}}{t_{11}t_{22}}\nu_{13} + \frac{t_{33}}{t_{22}}\nu_{23}.$$

By (10.6.3), $(t, u) \in (\mathcal{T} \circledS \mathcal{U})_\nu$ if and only if

$$\nu_{12} = \frac{t_{22}}{t_{11}}\nu_{12} + \frac{t_{23}}{t_{11}}\nu_{13} + u_{23}\frac{t_{33}}{t_{11}}\nu_{13}$$

$$\nu_{13} = \frac{t_{33}}{t_{11}}\nu_{13}$$

$$\nu_{23} = -\frac{t_{12}t_{33}}{t_{11}t_{22}}\nu_{13} + \frac{t_{33}}{t_{22}}\nu_{23} - u_{12}\frac{t_{33}}{t_{11}}\nu_{13}.$$

If $\nu_{13} \neq 0$, then $t_{11} = t_{33}$ and

$$u_{23} = \frac{(t_{11} - t_{22})\nu_{12}}{t_{11}\nu_{13}} - \frac{t_{23}}{t_{11}}$$

$$u_{12} = \frac{(t_{11} - t_{22})\nu_{23}}{t_{22}\nu_{13}} - \frac{t_{12}}{t_{11}}.$$

Therefore $\dim((\mathcal{T} \circledS \mathcal{U})_\nu) = 6$.

On the other hand, since $\mathcal{T}_\nu = \{t \in \mathcal{T} \mid \tilde{\phi}(t^{-1})^* \nu = \nu\}$, we conclude from (10.6.8) that $t \in \mathcal{T}_\nu$ if and only if

$$\nu_{12} = \frac{t_{22}}{t_{11}}\nu_{12} + \frac{t_{23}}{t_{11}}\nu_{13}$$

$$\nu_{13} = \frac{t_{33}}{t_{11}}\nu_{13}$$

$$\nu_{23} = -\frac{t_{12}t_{33}}{t_{11}t_{22}}\nu_{13} + \frac{t_{33}}{t_{22}}\nu_{23}.$$

If $\nu_{13} \neq 0$, then $t_{11} = t_{33}$ and

$$t_{12} = (t_{11} - t_{22})\frac{\nu_{23}}{\nu_{13}}$$

$$t_{23} = (t_{11} - t_{22})\frac{\nu_{12}}{\nu_{13}}.$$

Therefore $\dim \mathcal{T}_\nu = 3$. We have seen in the previous section that if $\nu_{13} \neq 0$, then $\dim \mathcal{U}_\nu = 1$ and thus we have $\dim(\mathcal{T}_\nu \circledS \mathcal{U}_\nu) = 4$ which shows that $\mathcal{T}_\nu \circledS \mathcal{U}_\nu \subsetneq (\mathcal{T} \circledS \mathcal{U})_\nu$.

Determination of M_ν/N_ν.   Recall that a key step in reduction by stages was to reduce (in the second stage) by M_ν/N_ν. In the present case, this means $S_\nu/(\{e_G\} \times H_\nu)$. We determine this group in the following proposition.

10.6.1 Proposition.   *If $p_G : G \circledS H \to G$ is the projection onto the first factor define $G^\nu := p_G(S_\nu)$. Let $\mathcal{O}_\nu \subset \mathfrak{h}^*$ be the H-coadjoint orbit through $\nu \in \mathfrak{h}^*$. Then*

(i) $G^\nu = \{g \in G \mid \tilde{\phi}(g^{-1})^* \nu \in \mathcal{O}_\nu\}$ *is an immersed Lie subgroup of G;*

(ii) *if $\mathcal{O}_\nu$ is closed in $\mathfrak{h}^*$, then G^ν is a closed Lie subgroup of G;*

(iii) *$S_\nu/(\{e_G\} \times H_\nu)$ and G^ν are isomorphic Lie groups;*

(iv) *the Lie algebra $\mathfrak{g}^\nu$ of G^ν equals $\{\xi \in \mathfrak{g} \mid \tilde{\phi}'(\xi)^* \nu \in T_\nu \mathcal{O}_\nu\}$.*

Proof.   (i) Since p_G is a Lie group homomorphism, its range G^ν is an immersed Lie subgroup of G. Now note that $\tilde{\phi}(g^{-1})^* \nu \in \mathcal{O}_\nu$ if and only if there is some $h \in H$ such that $\tilde{\phi}(g^{-1})^* \nu = \mathrm{Ad}_h^* \nu$ which is equivalent to $\mathrm{Ad}_{h^{-1}}^* \tilde{\phi}(g^{-1})^* \nu = \nu$. By (10.6.3) this happens if and only if $(g, h) \in S_\nu$ which proves the equality in the statement.

(ii) From (i) it immediately follows that if $g_n \to g$ and $g_n \in G^\nu$ then $\tilde{\phi}(g_n^{-1})^* \nu \to \tilde{\phi}(g^{-1})^* \nu \in \mathrm{cl}(\mathcal{O}_\nu) = \mathcal{O}_\nu$. Thus $g \in G^\nu$ which proves that G^ν is closed.

(iii) From (10.6.3) and the normality of H in G it follows that $\{e_G\} \times H_\nu$ is a closed normal subgroup of S_ν and hence $S_\nu/(\{e_G\} \times H_\nu)$ is a Lie group. Since $\ker(p_G|_{S_\nu}) = (\{e_G\} \times H) \cap S_\nu = \{e_G\} \times H_\nu$, it follows that $S_\nu/(\{e_G\} \times H_\nu)$ and G^ν are isomorphic Lie groups.

(iv) Recall from the general theory of Lie subgroups that the Lie algebra of G^ν equals $\mathfrak{g}^\nu = \{\xi \in \mathfrak{g} \mid \exp t\xi \in G^\nu \text{ for all } t \in \mathbb{R}\}$. Thus if $\xi \in \mathfrak{g}$ satisfies the condition $\tilde{\phi}(\exp(-t\xi))^* \nu \in \mathcal{O}_\nu$ for all $t \in \mathbb{R}$, then

$$-\tilde{\phi}'(\xi)^* \nu = \left.\frac{d}{dt}\right|_{t=0} \tilde{\phi}(\exp(-t\xi))^* \nu \in T_\nu \mathcal{O}_\nu$$

which shows that $\mathfrak{g}^\nu \subset \{\xi \in \mathfrak{g} \mid \tilde{\phi}'(\xi)^* \nu \in T_\nu \mathcal{O}_\nu\}$. Note that the right hand side of this inclusion is a vector subspace of $\mathfrak{g}$. We need to prove that this inclusion is an equality. To do this we shall show that the dimensions of both sides are the same. Since $\dim \mathfrak{g}^\nu = \dim S_\nu - \dim H_\nu$ by (iii), we need to prove that $\dim\{\xi \in \mathfrak{g} \mid \tilde{\phi}'(\xi)^* \nu \in T_\nu \mathcal{O}_\nu\} = \dim S_\nu - \dim H_\nu$.

To do this, recall from (10.6.4) that $\mathfrak{s}_\nu = \{(\xi, \eta) \in \mathfrak{s} \mid \tilde{\phi}'(\xi)^* \nu + \mathrm{ad}_\eta^* \nu = 0\}$ and hence $p_{\mathfrak{g}}(\mathfrak{s}_\nu) = \{\xi \in \mathfrak{g} \mid \tilde{\phi}'(\xi)^* \nu \in T_\nu \mathcal{O}_\nu\}$, where $p_{\mathfrak{g}} : \mathfrak{s} \to \mathfrak{g}$ is the projection onto the first factor. Since

$$\ker(p_{\mathfrak{g}}|_{\mathfrak{s}_\nu}) = \mathfrak{s}_\nu \cap (\{0\} \times \mathfrak{h}) = \{\eta \in \mathfrak{h} \mid (0, \eta) \in \mathfrak{s}_\nu\} = \mathfrak{h}_\nu,$$

we get $\dim\{\xi \in \mathfrak{g} \mid \tilde{\phi}'(\xi)^* \nu \in T_\nu \mathcal{O}_\nu\} = \dim \mathfrak{s}_\nu - \dim \mathfrak{h}_\nu$, as required.   ■

10.6.2 Corollary.   *With the notations of Proposition 10.6.1 we have*

(i) *$g \in G^\nu$ if and only if $\tilde{\phi}(g^{-1})^* \rho \in \mathcal{O}_\nu$ for all $\rho \in \mathcal{O}_\nu$;*

(ii) *$\xi \in \mathfrak{g}^\nu$ if and only if $\tilde{\phi}'(\xi)^* \rho \in T_\rho \mathcal{O}_\nu$ for all $\rho \in \mathcal{O}_\nu$.*

Proof.   (i) Since $\mathrm{AD}_{h^{-1}} \circ \phi(g^{-1}) = \phi(g^{-1}) \circ \mathrm{AD}_{g \cdot h^{-1}}$, we get for $\rho := \mathrm{Ad}_{h^{-1}}^* \nu$, the relation $\tilde{\phi}(g^{-1})^* \rho = \tilde{\phi}(g^{-1})^* \mathrm{Ad}_{h^{-1}}^* \nu = \mathrm{Ad}_{g \cdot h^{-1}}^* \tilde{\phi}(g^{-1})^* \nu$. If $g \in G^\nu$ then $\tilde{\phi}(g^{-1})^* \nu \in \mathcal{O}_\nu$ and hence $\tilde{\phi}(g^{-1})^* \rho \in \mathrm{Ad}_{g \cdot h^{-1}}^* (\mathcal{O}_\nu) \subset \mathcal{O}_\nu$.

(ii) Note that for any $\xi \in \mathfrak{g}$ and $\eta \in \mathfrak{h}$ we have

$$\frac{d}{dt}\bigg|_{t=0} \mathrm{Ad}_{(\exp t\xi)\cdot h^{-1}}\, \eta = \mathrm{Ad}_{h^{-1}} \left[T_{h^{-1}} L_h \xi_H(h^{-1}), \eta \right]$$

(see, e.g., formula (9.3.4) in [MandS]), that is,

$$\frac{d}{dt}\bigg|_{t=0} \mathrm{Ad}_{(\exp t\xi)\cdot h^{-1}} = \mathrm{Ad}_{h^{-1}} \circ \mathrm{ad}_{T_{h^{-1}} L_h \xi_H(h^{-1})}\,.$$

Next, use this relation in the computation of the t-derivative at $t = 0$ of the identity $\mathrm{Ad}_{h^{-1}} \circ \tilde{\phi}(\exp(-t\xi)) = \tilde{\phi}(\exp(-t\xi)) \circ \mathrm{Ad}_{(\exp t\xi)\cdot h^{-1}}$ proved in (i) to get

$$- \mathrm{Ad}_{h^{-1}} \circ \tilde{\phi}'(\xi) = -\tilde{\phi}'(\xi) \circ \mathrm{Ad}_{h^{-1}} + \mathrm{Ad}_{h^{-1}} \circ \mathrm{ad}_{T_{h^{-1}} L_h \xi_H(h^{-1})}\,.$$

Therefore, since $\rho = \mathrm{Ad}^*_{h^{-1}}\, \nu$ we get

$$\tilde{\phi}'(\xi)^* \rho = \tilde{\phi}'(\xi)^* \mathrm{Ad}^*_{h^{-1}}\, \nu = \mathrm{Ad}^*_{h^{-1}}\, \tilde{\phi}'(\xi)^* \nu - \mathrm{ad}^*_{T_{h^{-1}} L_h \xi_H(h^{-1})}\, \rho.$$

If $\xi \in \mathfrak{g}^\nu$ then there is some $\zeta \in \mathfrak{h}$ such that $\tilde{\phi}'(\xi)^* \nu = -\mathrm{ad}^*_\zeta \nu$ and hence

$$\tilde{\phi}'(\xi)^* \rho = -\mathrm{Ad}^*_{h^{-1}} \mathrm{ad}^*_\zeta \nu - \mathrm{ad}^*_{T_{h^{-1}} L_h \xi_H(h^{-1})}\, \rho$$

$$= -\mathrm{ad}^*_{\mathrm{Ad}_h \zeta + T_{h^{-1}} L_h \xi_H(h^{-1})}\, \rho \in T_\rho \mathcal{O}_\nu$$

since $\mathrm{ad}_\zeta \circ \mathrm{Ad}_{h^{-1}} = \mathrm{Ad}_{h^{-1}} \circ \mathrm{ad}_{\mathrm{Ad}_h \zeta}$. ∎

In what follows we shall identify $\{e_G\} \times H_\nu$ with H_ν and write S_ν/H_ν instead of $S_\nu/(\{e_G\} \times H_\nu)$.

Note that in the example $\mathcal{T} \,\circledS\, \mathcal{U}$, the computation done previously shows that if $\nu_{13} \neq 0$ then $\dim \mathcal{T}^\nu = 5$ since $\mathcal{T}^\nu \cong (\mathcal{T} \,\circledS\, \mathcal{U})_\nu/\mathcal{U}_\nu$, $\dim((\mathcal{T} \,\circledS\, \mathcal{U})_\nu = 6$, and $\dim \mathcal{U}_\nu = 1$.

Where We Stand. The framework just described fits into the general theory as follows: $M = S$, $N = \{e_G\} \times H$, $P = T^*S$, stage one reduction is relative to the cotangent lift of right H-translation on T^*S, and stage two reduction is relative to the naturally induced action of $M_\nu/N_\nu = S_\nu/(\{e_G\} \times H_\nu) \cong G^\nu$ on the first reduced space. As in the case of semidirect products with vector spaces given by a representation (studied in §4.2), the quotient group $S_\nu/(\{e_G\} \times H_\nu)$ has the explicit realization as the immersed Lie subgroup G^ν of G and hence of S (this quotient group was identified explicitly in §4.2 to be G_a). So, even though the coadjoint isotropy subgroup S_ν does not have a very satisfactory description, the quotient Lie group $S_\nu/(\{e_G\} \times H_\nu)$ does, namely by Proposition 10.6.1 it is isomorphic to the immersed Lie subgroup G^ν of G, and this allows one to carry out the second stage reduction.

Two extreme cases when S_ν is determined are worth mentioning. First, if $H = V$ is a vector space and the G-action is a representation, then $S_\nu = G_\nu \times V$ and $G^\nu = G_\nu$ for any $\nu \in V^*$, since the coadjoint orbits in V^* are just points. In §4.2, ν was denoted by $a \in V^*$. Second, if all automorphisms $\phi(g)$ of H are *inner*, then $G^\nu = G$ for any $\nu \in \mathfrak{h}^*$ and so $S_\nu = G \times H_\nu$.

The Action of G^ν on the First Reduced Space.   To set the scene for the second stage reduction, and being guided by the general theory, we now determine the G^ν action on the first reduced space.

10.6.3 Theorem.   *The lift of right translation of S on itself induces a symplectic right action of G^ν on the first reduced space $(T^*S)_\nu$. The symplectic diffeomorphism*

$$[T^*\Upsilon^{-1}]_\nu : (T^*S)_\nu \to T^*G \times \mathcal{O}_\nu$$

from Theorem 10.5.1 induces a symplectic action of G^ν on $(T^*G \times \mathcal{O}_\nu, \omega_G \oplus \omega_{\mathcal{O}_\nu}^+)$ that is given by

$$(\alpha_g, \rho) \cdot l := \left(T^*_{gl} R_{l^{-1}} \alpha_g, \tilde{\phi}(l)^* \rho \right), \tag{10.6.9}$$

where $\alpha_g \in T^*G$, $\rho \in \mathcal{O}_\nu$ and $l \in G^\nu$.

Proof.   That S_ν/H_ν acts freely, properly and symplectically on the first reduced space $(T^*S)_\nu$ was proven in general in Lemma 5.2.2. Since G^ν is isomorphic to S_ν/H_ν, it also acts freely, properly and symplectically on the first reduced space $(T^*S)_\nu$. This proves the first statement.

From Proposition 10.6.2 (i), it follows that $\tilde{\phi}(l)^* \rho \in \mathcal{O}_\nu$, and so it is clear that equation (10.6.9) defines a right action. The G^ν-action on $T^*G \times \mathcal{O}_\nu$ is given by the lift of right translation by $l \in G^\nu$ on the first factor and the the dual of the Lie algebra isomorphism $\tilde{\phi}(l) \in \text{Aut}(\mathfrak{h})$ on the second factor. Both are symplectic maps. Since the symplectic form on $T^*G \times \mathcal{O}_\nu$ is the sum of the symplectic forms on T^*G and on $\mathcal{O}_\nu$, it immediately follows that this G^ν-action is symplectic. This G^ν-action is obviously free and proper.

The proof will be complete if we can show that the map $[T^*\Upsilon^{-1}]_\nu$ is equivariant, with S_ν/H_ν acting on the domain and with G^ν acting on the range and with the isomorphism between S_ν/H_ν and G^ν being $[(l, k)] \mapsto l$. That is, what we must check is that

$$[T^*\Upsilon^{-1}]_\nu \left(\left[\left(\alpha_g, T^*_h L_{h^{-1}} \tilde{\phi}(g^{-1})^* \nu \right) \right] \cdot [(l, k)] \right)$$
$$= \left([T^*\Upsilon^{-1}]_\nu \left[\left(\alpha_g, T^*_h L_{h^{-1}} \tilde{\phi}(g^{-1})^* \nu \right) \right] \right) \cdot l, \tag{10.6.10}$$

for all $g \in G$, $h \in H$, $l \in G^\nu$, $(l, k) \in S_\nu$, and $\alpha_g \in T^*G$. The left hand side of (10.6.10) equals

$$
[T^*\Upsilon^{-1}]_\nu \left(\left[\left(\alpha_g, T_h^* L_{h^{-1}} \tilde\phi(g^{-1})^* \nu \right) \right] \cdot [(l, k)] \right)
$$
$$
= [T^*\Upsilon^{-1}]_\nu \left(\left[\left(\alpha_g, T_h^* L_{h^{-1}} \tilde\phi(g^{-1})^* \nu \right) \cdot (l, k) \right] \right). \tag{10.6.11}
$$

However, by (10.4.6) and (10.5.3), we have

$$
\left(\alpha_g, T_h^* L_{h^{-1}} \tilde\phi(g^{-1})^* \nu \right) \cdot (l, k)
$$
$$
= T^*_{(g,h)(l,k)} R_{(l,k)^{-1}} \left(\alpha_g, T_h^* L_{h^{-1}} \tilde\phi(g^{-1})^* \nu \right)
$$
$$
= \left(T^*_{gl} R_{l^{-1}} \alpha_g + \left(T^*_{gl} \phi^{l^{-1} \cdot k^{-1}} \circ T^*_{g \cdot k^{-1}} L_{h(g \cdot k)} \circ T_h^* L_{h^{-1}} \circ \tilde\phi(g^{-1})^* \right) \nu, \right.
$$
$$
\left. \left(T^*_{h(g \cdot k)} R_{g \cdot k^{-1}} \circ T_h^* L_{h^{-1}} \circ \tilde\phi(g^{-1})^* \right) \nu \right). \tag{10.6.12}
$$

By (10.6.3), $(l, k) \in S_\nu$ is equivalent to

$$
\mathrm{Ad}^*_{k^{-1}} \tilde\phi(l^{-1})^* \nu = \nu. \tag{10.6.13}
$$

Thus, the second component of (10.6.12) can be simplified to

$$
\left(T^*_{h(g \cdot k)} R_{g \cdot k^{-1}} \circ T_h^* L_{h^{-1}} \circ \tilde\phi(g^{-1})^* \right) \nu
$$
$$
= \left(T^*_{h(g \cdot k)} R_{g \cdot k^{-1}} \circ T_h^* L_{h^{-1}} \circ \tilde\phi(g^{-1})^* \circ \mathrm{Ad}^*_{k^{-1}} \circ \tilde\phi(l^{-1})^* \right) \nu
$$
$$
= T^*_{h(g \cdot k)} L_{(g \cdot k^{-1}) h^{-1}} \tilde\phi(l^{-1} g^{-1})^* \nu
$$

since $\phi(l^{-1}) \circ \mathrm{AD}_{k^{-1}} \circ \phi(g^{-1}) \circ L_{h^{-1}} \circ R_{g \cdot k^{-1}} = \phi(l^{-1} g^{-1}) \circ L_{(g \cdot k^{-1}) h^{-1}}$. Thus, (10.6.12) is given by

$$
\left(\alpha_g, T_h^* L_{h^{-1}} \tilde\phi(g^{-1})^* \nu \right) \cdot (l, k)
$$
$$
= \left(T^*_{gl} R_{l^{-1}} \left(\alpha_g + \left(T_g^* L_{g^{-1}} \circ T^*_{e_G} \phi^{k^{-1}} \circ T^*_{k^{-1}} L_k \right) \nu \right), \right.
$$
$$
\left. T^*_{h(g \cdot k)} L_{[h(g \cdot k)]^{-1}} \tilde\phi(l^{-1} g^{-1})^* \nu \right) \in \mathbf{J}_H^{-1}(\nu) \tag{10.6.14}
$$

since $\phi(g^{-1}) \circ L_{h^{-1}} \circ L_{h(g \cdot k)} \circ \phi^{l^{-1} \cdot k^{-1}} = L_k \circ \phi^{k^{-1}} \circ L_{g^{-1}} \circ R_{l^{-1}}$.

By equations (10.5.6) and (10.6.14), the expression (10.6.12) becomes

$$[T^*\Upsilon^{-1}]_\nu \left(\left[\left(\alpha_g, T_h^* L_{h^{-1}} \tilde\phi(g^{-1})^* \nu \right) \cdot (l,k) \right] \right)$$

$$= [T^*\Upsilon^{-1}]_\nu \left(\left[T_{gl}^* R_{l^{-1}} \left(\alpha_g + T_g^* L_{g^{-1}} T_{e_G}^* \phi^{k^{-1}} T_{k^{-1}}^* L_k \nu \right), \right. \right.$$

$$\left. \left. T_{h(g\cdot k)}^* L_{[h(g\cdot k)]^{-1}} \tilde\phi(l^{-1} g^{-1})^* \nu \right] \right)$$

$$= \left(T_{gl}^* R_{l^{-1}} \left(\alpha_g + T_g^* L_{g^{-1}} T_{e_G}^* \phi^{k^{-1}} T_{k^{-1}}^* L_k \nu \right) \right.$$

$$+ T_{gl}^* R_{(gl)^{-1}} T_{e_G}^* \phi^{h(g\cdot k)} T_{h(g\cdot k)}^* L_{(h(g\cdot k))^{-1}} \tilde\phi((gl)^{-1})^* \nu,$$

$$\left. \mathrm{Ad}_{(gl)^{-1}\cdot(h(g\cdot k))^{-1}}^* \nu \right)$$

$$= \left(T_{gl}^* R_{l^{-1}} \left(\alpha_g + T_g^* L_{g^{-1}} T_{e_G}^* \phi^{k^{-1}} T_{k^{-1}}^* L_k \nu \right. \right.$$

$$\left. + T_g^* R_{g^{-1}} T_{e_G}^* \phi^{h(g\cdot k)} T_{h(g\cdot k)}^* L_{(h(g\cdot k))^{-1}} \tilde\phi((gl)^{-1})^* \nu \right),$$

$$\left. \tilde\phi(l)^* \mathrm{Ad}_{g^{-1}\cdot h^{-1}}^* \nu \right). \tag{10.6.15}$$

The last equality was obtained as follows. In the first component, we wrote

$$T_{gl}^* R_{l^{-1} g^{-1}} = T_{gl}^* R_{l^{-1}} \circ T_g^* R_{g^{-1}}.$$

The second component was obtained by first noting that

$$(gl)^{-1} \cdot (h(g\cdot k))^{-1} = l^{-1} \cdot (k^{-1}(g^{-1}\cdot h^{-1})).$$

Second, note that

$$\mathrm{AD}_{l^{-1}\cdot(k^{-1}(g^{-1}\cdot h^{-1}))} = \phi(l^{-1}) \circ \mathrm{AD}_{k^{-1}} \circ \mathrm{AD}_{g^{-1}\cdot h^{-1}} \circ\phi(l).$$

Therefore, since $(l,k) \in S_\nu$, we have

$$\mathrm{Ad}_{l^{-1}\cdot(k^{-1}(g^{-1}\cdot h^{-1}))}^* \nu = \tilde\phi(l)^* \mathrm{Ad}_{g^{-1}\cdot h^{-1}}^* \mathrm{Ad}_{k^{-1}} \tilde\phi(l^{-1})^* \nu$$

$$= \tilde\phi(l)^* \mathrm{Ad}_{g^{-1}\cdot h^{-1}}^* \nu$$

where we have used (10.6.13). Thus, (10.6.15) holds.

By (10.6.9), the expression (10.6.15) equals

$$\left(\alpha_g + T_g^* L_{g^{-1}} T_{e_G}^* \phi^{k^{-1}} T_{k^{-1}}^* L_k \nu \right.$$

$$\left. + T_g^* R_{g^{-1}} T_{e_G}^* \phi^{h(g\cdot k)} T_{h(g\cdot k)}^* L_{(h(g\cdot k))^{-1}} \tilde\phi((gl)^{-1})^* \nu, \mathrm{Ad}_{g^{-1}\cdot h^{-1}}^* \nu \right) \cdot l$$

$$= \left(\alpha_g + T_g^* R_{g^{-1}} \left(\mathrm{Ad}_{g^{-1}}^* T_{e_G}^* \phi^{k^{-1}} T_{k^{-1}}^* L_k \nu \right. \right.$$

$$\left. + T_{e_G}^* \phi^{h(g\cdot k)} T_{h(g\cdot k)}^* L_{(h(g\cdot k))^{-1}} \tilde\phi((gl)^{-1})^* \nu \right), \mathrm{Ad}_{g^{-1}\cdot h^{-1}}^* \nu \right) \cdot l$$

$$= \left(\alpha_g + T_g^* R_{g^{-1}} \left(\mathrm{Ad}_{g^{-1}}^* T_{e_G}^* \phi^{k^{-1}} T_{k^{-1}}^* L_k \nu \right. \right.$$

$$\left. + T_{e_G}^* \phi^{h(g\cdot k)} T_{h(g\cdot k)}^* L_{(h(g\cdot k))^{-1}} \tilde\phi(g^{-1})^* \mathrm{Ad}_k^* \nu \right), \mathrm{Ad}_{g^{-1}\cdot h^{-1}}^* \nu \right) \cdot l, \tag{10.6.16}$$

where in the last equality, we again used (10.6.13). We claim that the following identity holds

$$\mathrm{Ad}^*_{g^{-1}} T^*_{e_G} \phi^{k^{-1}} T^*_{k^{-1}} L_k \nu + T^*_{e_G} \phi^{h(g\cdot k)} T^*_{h(g\cdot k)} L_{(h(g\cdot k))^{-1}} \tilde{\phi}(g^{-1})^* \mathrm{Ad}^*_k \nu$$
$$= T^*_{e_G} \phi^h T^*_h L_{h^{-1}} \tilde{\phi}(g^{-1})^* \nu. \tag{10.6.17}$$

This is proved as follows. First of all, it is readily checked that

$$\left(\mathrm{AD}_k \circ \phi(g^{-1}) \circ L_{(g\cdot k^{-1})h^{-1}} \circ \phi^{h(g\cdot k)} \right)(g') \left(L_k \circ \phi^{k^{-1}} \circ \mathrm{AD}_{g^{-1}} \right)(g')$$
$$= \left(\phi(g^{-1}) \circ L_{h^{-1}} \circ \phi^h \right)(g')$$

for all $g, g' \in G$ and $h, k \in H$. Replacing g' by $\exp(t\xi)$ for $\xi \in \mathfrak{g}$, differentiating in t at $t = 0$, and taking duals, one arrives at (10.6.17).

Substituting (10.6.17) in (10.6.16) gives the expression

$$\left(\alpha_g + T^*_g R_{g^{-1}} T^*_{e_G} \phi^h T^*_h L_{h^{-1}} \tilde{\phi}(g^{-1})^* \nu, \mathrm{Ad}^*_{g^{-1}\cdot h^{-1}} \nu \right) \cdot l$$
$$= \left([T^* \Upsilon^{-1}]_\nu \left[\left(\alpha_g, T^*_h L_{h^{-1}} \tilde{\phi}(g^{-1})^* \nu \right) \right] \right) \cdot l$$

by (10.5.6). Thus, we have established (10.6.10) and so the proof of the Theorem is complete. ∎

The G^ν-momentum map on $T^*G \times \mathcal{O}_\nu$. We shall compute the G^ν-momentum map from the general formula (5.2.6) in §5.2. It is tempting to proceed directly since, in this case, both the symplectic manifold $\left(T^*G \times \mathcal{O}_\nu, \omega_G \oplus \omega^+_{\mathcal{O}_\nu} \right)$ and the G^ν-action (10.6.9) are relatively simple: the action is diagonal with the lift of right translation on the first factor and $\tilde{\phi}^*$ on the second. Thus, by general theory, the momentum map $\mathbf{J}_\nu : T^*G \times \mathcal{O}_\nu \to (\mathfrak{g}^\nu)^*$ is the sum of the momentum maps on each factor. The momentum map on the first factor T^*G is given by $\alpha_g \in T^*G \mapsto \left(T^*_{e_G} L_g \alpha_g \right) \big|_{\mathfrak{g}^\nu} \in (\mathfrak{g}^\nu)^*$. To compute the momentum map of the right G^ν-action on $\mathcal{O}_\nu$ given by

$$(\rho, l) \in \mathcal{O}_\nu \times G^\nu \mapsto \tilde{\phi}(l)^* \rho \in \mathcal{O}_\nu \tag{10.6.18}$$

is quite a challenge and it is not a priori clear that this action even admits a momentum map. However, once we have the explicit formula for the momentum map $\mathbf{J}_\nu : T^*G \times \mathcal{O}_\nu \to (\mathfrak{g}^\nu)^*$ obtained by applying (5.2.6), we shall find a candidate for the momentum map of the G^ν-action (10.6.18) on $\mathcal{O}_\nu$ and then we shall check directly that this formula is indeed a G^ν-momentum map on $\mathcal{O}_\nu$. Using this result we shall then compute the nonequivariance cocycle of $\mathbf{J}_\nu$.

In order to calculate the momentum map $\mathbf{J}_\nu : T^*G \times \mathcal{O}_\nu \to (\mathfrak{g}^\nu)^*$ from formula (5.2.6) we need to identify the various spaces, maps, and equivalence classes in our concrete situation. Let us recall that the general formula for $\mathbf{J}_\nu : P_\nu \to (\mathfrak{m}_\nu/\mathfrak{n}_\nu)^*$ is

$$\langle \mathbf{J}_\nu([z]), [\xi] \rangle = \langle \mathbf{J}_M(z), \xi \rangle - \langle \bar{\nu}, \xi \rangle, \tag{10.6.19}$$

where $\nu \in \mathfrak{n}^*$, $\xi \in \mathfrak{m}_\nu$, $z \in \mathbf{J}_N^{-1}(\nu)$, and $\bar{\nu} \in (\mathfrak{m}_\nu)^*$ is an arbitrary linear extension of $\nu|_{\mathfrak{n}_\nu}$ to $\mathfrak{m}_\nu$. In our case the Lie groups, Lie algebras, and symplectic manifolds are $M = S := G \, \textcircled{S} \, H$, $N = H$, $M_\nu = S_\nu$ (given by (10.6.3)), $N_\nu = H_\nu$, $M_\nu/N_\nu = S_\nu/H_\nu \cong G^\nu$ (see Proposition 10.6.1), $\mathfrak{m} = \mathfrak{s} := \mathfrak{g} \, \textcircled{S} \, \mathfrak{h}$, $\mathfrak{n} = \mathfrak{h}$, $\mathfrak{m}_\nu = \mathfrak{s}_\nu$ (given by (10.6.4)), $\mathfrak{n}_\nu = \mathfrak{h}_\nu$, $\mathfrak{m}_\nu/\mathfrak{n}_\nu = \mathfrak{s}_\nu/\mathfrak{h}_\nu \cong \mathfrak{g}^\nu$ (given in Proposition 10.6.1(iv)), $P = (T^*S, \omega_S)$, $P_\nu = \left(T^*G \times \mathcal{O}_\nu, \omega_G \oplus \omega_{\mathcal{O}_\nu}^+\right)$ (by Theorem 10.6.3), where ω_S and ω_G are the canonical symplectic forms on T^*S and T^*G, respectively.

To determine the maps and the various representatives of the equivalence classes we need a few preparatory remarks. We begin by noting that there are many (non-canonical) isomorphisms between $\mathfrak{s}_\nu$ and $\mathfrak{g}^\nu \times \mathfrak{h}_\nu$ as vector spaces. To explain this, recall from (10.6.4) that $(\xi, \eta) \in \mathfrak{s}_\nu$ if and only if $\mathrm{ad}_\eta^* \nu + \tilde{\phi}'(\xi)^* \nu = 0$. Equivalently, $(\xi, \eta) \in \mathfrak{s}_\nu$ if and only if $\xi \in \mathfrak{g}^\nu$ and there is some element $\lambda(\xi) \in \mathfrak{h}$ such that $\tilde{\phi}'(\xi)^* \nu = -\mathrm{ad}_{\lambda(\xi)}^* \nu$. This element $\lambda(\xi)$ is, of course, not uniquely determined but its class in $\mathfrak{h}/\mathfrak{h}_\nu$ is, so this identity defines a linear map $\mathfrak{g}^\nu \to \mathfrak{h}/\mathfrak{h}_\nu$. Let $\lambda : \mathfrak{g}^\nu \to \mathfrak{h}$ be any linear lift of this map. For any such linear map λ we have a vector space isomorphism

$$(\xi, \eta) \in \mathfrak{s}_\nu \mapsto (\xi, \eta - \lambda(\xi)) \in \mathfrak{g}^\nu \times \mathfrak{h}_\nu. \qquad (10.6.20)$$

Note that this map is indeed well defined because the relations $\tilde{\phi}'(\xi)^* \nu = -\mathrm{ad}_\eta^* \nu$ (which holds because $(\xi, \eta) \in \mathfrak{s}_\nu$) and $\tilde{\phi}'(\xi)^* \nu = -\mathrm{ad}_{\lambda(\xi)}^* \nu$ (which holds by definition of λ) imply that $\mathrm{ad}_{\eta - \lambda(\xi)}^* \nu = 0$, that is, $\eta - \lambda(\xi) \in \mathfrak{h}_\nu$. The map (10.6.20) is clearly linear and injective and it is an isomorphism because $\dim \mathfrak{s}_\nu = \dim \mathfrak{g}^\nu + \dim \mathfrak{h}_\nu$ by Proposition 10.6.1(iii). Thus, given $\xi \in \mathfrak{g}^\nu$, a Lie algebra element in $\mathfrak{s}_\nu$ that represents the class in $\mathfrak{s}_\nu/\mathfrak{h}_\nu$ identified with $\xi \in \mathfrak{g}^\nu$ is $(\xi, \eta - \lambda(\xi)) \in \mathfrak{s}_\nu$. Of course, all these considerations are dependent on the choice of the linear map $\lambda : \mathfrak{g}^\nu \to \mathfrak{h}$.

Next, note that in the general formula (10.6.19) the projection map onto the first reduced space $(T^*S)_\nu$ is replaced by the map $\Psi_\nu : \mathbf{J}_H^{-1}(\nu) \to T^*G \times \mathcal{O}_\nu$ defined by $\Psi_\nu = \bar{\pi}_\nu^H \circ \left(T^*\Upsilon^{-1}\right)_\nu$. Thus, if $(\beta_g, \rho) \in T^*G \times \mathcal{O}_\nu$, an element $\left(\alpha_g, T_h^* L_{h^{-1}} \tilde{\phi}(g^{-1})^* \nu\right) \in \mathbf{J}_H^{-1}(\nu)$ satisfying $\Psi_\nu \left(\alpha_g, T_h^* L_{h^{-1}} \tilde{\phi}(g^{-1})^* \nu\right) = (\beta_g, \rho)$ is given by

$$\alpha_g := \beta_g - T_g^* R_{g^{-1}} T_{e_G}^* \phi^h T_h^* L_{h^{-1}} \tilde{\phi}(g^{-1})^* \nu,$$

where $\rho = \mathrm{Ad}_{g^{-1} \cdot h^{-1}}^* \nu$.

We also need the formula for the momentum map of the lift of the right action of $S = G \, \textcircled{S} \, H$ on T^*S, namely,

$$\mathbf{J}_S^R(\alpha_g, \beta_h) = \left(T_{e_G}^* L_g \alpha_g, \tilde{\phi}(g)^* T_{e_H}^* L_h \beta_h\right).$$

Finally, if $\nu \in \mathfrak{h}^*$, all possible linear extensions $\bar{\nu}$ of $\nu|_{\mathfrak{h}_\nu}$ to $\mathfrak{s}_\nu \cong \mathfrak{g}^\nu \times \mathfrak{h}_\nu$ are of the form $\bar{\nu} := (\kappa, \nu|_{\mathfrak{h}_\nu}) \in (\mathfrak{g}^\nu)^* \times (\mathfrak{h}_\nu)^*$, where $\kappa \in (\mathfrak{g}^\nu)^*$ is arbitrary. We shall compute $\mathbf{J}_\nu$ with this arbitrary choice of κ and then choose it in such a way that the resulting formula has the most convenient expression.

With these ingredients in mind we can explicitly compute the momentum map $\mathbf{J}_\nu : T^*G \times \mathcal{O}_\nu \to (\mathfrak{g}^\nu)^*$ of the G^ν-action (10.6.9) given by (10.6.19). If $(\beta_g, \rho) \in T^*G \times \mathcal{O}_\nu$ and $\xi \in \mathfrak{g}^\nu$ we have

$$
\begin{aligned}
\langle \mathbf{J}_\nu(\beta_g, \rho), \xi \rangle \\
= \Big\langle \mathbf{J}_S^R &\left(\beta_g - T_g^* R_{g^{-1}} T_{e_G}^* \phi^h T_h^* L_{h^{-1}} \tilde{\phi}(g^{-1})^* \nu, \right. \\
& \left. T_h^* L_{h^{-1}} \tilde{\phi}(g^{-1})^* \nu \right), (\xi, \eta - \lambda(\xi)) \Big\rangle \\
& - \langle (\kappa, \nu|_{\mathfrak{h}_\nu}), (\xi, \eta - \lambda(\xi)) \rangle \\
= \Big\langle & \left(T_{e_G}^* L_g \beta_g - \mathrm{Ad}_g^* T_{e_G}^* \phi^h T_h^* L_{h^{-1}} \tilde{\phi}(g^{-1})^* \nu, \, \nu \right), (\xi, \eta - \lambda(\xi)) \Big\rangle \\
& - \langle (\kappa, \nu|_{\mathfrak{h}_\nu}), (\xi, \eta - \lambda(\xi)) \rangle \\
= \Big\langle & T_{e_G}^* L_g \beta_g - \mathrm{Ad}_g^* T_{e_G}^* \phi^h T_h^* L_{h^{-1}} \tilde{\phi}(g^{-1})^* \nu, \xi \Big\rangle + \langle \nu, \eta - \lambda(\xi) \rangle \\
& - \langle \kappa, \xi \rangle - \langle \nu, \eta - \lambda(\xi) \rangle .
\end{aligned}
$$

However,

$$
\mathrm{Ad}_g^* T_{e_G}^* \phi^h T_h^* L_{h^{-1}} \tilde{\phi}(g^{-1})^* \nu = T_{e_G}^* \phi^{g^{-1} \cdot h} T_{g^{-1} \cdot h}^* L_{g^{-1} \cdot h^{-1}} \nu
$$

since $\phi(g^{-1}) \circ L_{h^{-1}} \circ \phi^h \circ \mathrm{AD}_g = L_{g^{-1} \cdot h^{-1}} \circ \phi^{g^{-1} \cdot h}$. Thus the expression for $\mathbf{J}_\nu$ becomes

$$
\begin{aligned}
\langle \mathbf{J}_\nu(\beta_g, \rho), \xi \rangle \\
= \Big\langle T_{e_G}^* L_g \beta_g - T_{e_G}^* \phi^{g^{-1} \cdot h} T_{g^{-1} \cdot h}^* L_{g^{-1} \cdot h^{-1}} \nu, \xi \Big\rangle - \langle \kappa, \xi \rangle .
\end{aligned}
$$

Since $\rho = \mathrm{Ad}_{g^{-1} \cdot h^{-1}}^* \nu$ and $g^{-1} \cdot h$ is an arbitrary element of H, this formula yields

$$
\mathbf{J}_\nu \left(\beta_g, \mathrm{Ad}_{h^{-1}}^* \nu \right) = j_\nu^* T_{e_G}^* L_g \beta_g - j_\nu^* T_{e_G}^* \phi^h T_h^* L_{h^{-1}} \nu - \kappa,
$$

where $j_\nu : \mathfrak{g}^\nu \hookrightarrow \mathfrak{g}$ is the inclusion and $\kappa \in (\mathfrak{g}^\nu)^*$ is arbitrary. Thus we can choose the momentum map $\mathbf{J}_\nu : T^*G \times \mathcal{O}_\nu \to (\mathfrak{g}^\nu)^*$ to be

$$
\begin{aligned}
\mathbf{J}_\nu \left(\beta_g, \mathrm{Ad}_{h^{-1}}^* \nu \right) &= j_\nu^* T_{e_G}^* L_g \beta_g - j_\nu^* T_{e_G}^* \phi^h T_h^* L_{h^{-1}} \nu \\
&= j_\nu^* T_{e_G}^* L_g \beta_g + j_\nu^* T_{e_G}^* \phi^{h^{-1}} T_{h^{-1}}^* R_h \nu \qquad (10.6.21)
\end{aligned}
$$

for any $\beta_g \in T^*G$ and any $h \in H$. The second equality above is obtained in the following way. The derivative of the identity

$$
(L_{h^{-1}} \circ \phi^h)(g)(R_h \circ \phi^{h^{-1}})(g) = e_H
$$

relative to $g \in G$ at the value $g = e_G$ yields

$$
T_h L_{h^{-1}} \circ T_{e_G} \phi^h + T_{h^{-1}} R_h \circ T_{e_G} \phi^{h^{-1}} = 0,
$$

which implies, by taking the duals, the second equality in (10.6.21).

The G^ν-momentum map on $\mathcal{O}_\nu$.   Formula (10.6.21) for $\mathbf{J}_\nu$ and the fact that the first term is the momentum map of the lift of right translation of G^ν on T^*G suggests that the G^ν-action (10.6.18) on $\mathcal{O}_\nu$ admits a momentum map given by the second summand in (10.6.21). The next proposition verifies this statement.

10.6.4 Proposition.   *The right G^ν-action (10.6.18) on $\mathcal{O}_\nu$ admits the momentum map $\mathbf{K}^\nu : \mathcal{O}_\nu \to (\mathfrak{g}^\nu)^*$ given by*

$$\mathbf{K}^\nu \left(\mathrm{Ad}^*_{h^{-1}} \nu \right) = -j_\nu^* T^*_{e_G} \phi^h T^*_h L_{h^{-1}} \nu = j_\nu^* T^*_{e_G} \phi^{h^{-1}} T^*_{h^{-1}} R_h \nu. \qquad (10.6.22)$$

The momentum map $\mathbf{K}^\nu$ is, in general, non-equivariant and defines the right $(\mathfrak{g}^\nu)^*$-valued group one-cocycle $\sigma^\nu : G^\nu \to (\mathfrak{g}^\nu)^*$ given by

$$\sigma^\nu(l) = -j_\nu^* T^*_{e_G} \phi^{k(l)} T^*_{k(l)} L_{k(l)^{-1}} \nu, \qquad (10.6.23)$$

where $k(l) \in H$ is any element satisfying $\tilde{\phi}(l)^* \nu = \mathrm{Ad}^*_{k(l)^{-1}} \nu$.

Proof.   The second equality in (10.6.22) was just proved above. In this proof we shall use exclusively the second expression in (10.6.22). We begin by computing the differential of $\mathbf{K}^\nu$ at $\rho := \mathrm{Ad}^*_{h^{-1}} \nu$. For any $\eta \in \mathfrak{h}$ and $\xi \in \mathfrak{g}^\nu$ we get

$$\begin{aligned}
\langle \mathbf{dK}^\nu(\rho) \left(- \mathrm{ad}^*_\eta \rho \right), \xi \rangle &= \frac{d}{dt}\bigg|_{t=0} \left\langle \mathbf{K}^\nu \left(\mathrm{Ad}^*_{\exp(-t\eta)} \rho \right), \xi \right\rangle \\
&= \frac{d}{dt}\bigg|_{t=0} \left\langle \mathbf{K}^\nu \left(\mathrm{Ad}^*_{(\exp(t\eta)h)^{-1}} \nu \right), \xi \right\rangle \\
&= \frac{d}{dt}\bigg|_{t=0} \left\langle j_\nu^* T^*_{e_G} \phi^{(\exp(t\eta)h)^{-1}} T^*_{(\exp(t\eta)h)^{-1}} R_{\exp(t\eta)h} \nu, \xi \right\rangle \\
&= \frac{d}{dt}\bigg|_{t=0} \left\langle \nu, \left(T_{(\exp(t\eta)h)^{-1}} R_{\exp(t\eta)h} \circ T_{e_G} \phi^{(\exp(t\eta)h)^{-1}} \right)(\xi) \right\rangle \\
&= \frac{d}{dt}\bigg|_{t=0} \left\langle \nu, T_{(\exp(t\eta)h)^{-1}} R_{\exp(t\eta)h} \left(\xi_H \left(h^{-1} \exp(-t\eta) \right) \right) \right\rangle.
\end{aligned}$$

Since $\xi_H(hh') = T_h R_{h'} \xi_H(h) + T_{h'} L_h \xi_H(h')$ for any $h, h' \in H$, the above expression equals

$$\begin{aligned}
\frac{d}{dt}\bigg|_{t=0} &\left\langle \nu, T_{(\exp(t\eta)h)^{-1}} R_{\exp(t\eta)h} T_{h^{-1}} R_{\exp(-t\eta)} \xi_H(h^{-1}) \right. \\
&\qquad \left. + T_{(\exp(t\eta)h)^{-1}} R_{\exp(t\eta)h} T_{\exp(-t\eta)} L_{h^{-1}} \xi_H(\exp(-t\eta)) \right\rangle \\
&= \frac{d}{dt}\bigg|_{t=0} \left\langle \nu, T_{h^{-1}} R_h \xi_H(h^{-1}) + \mathrm{Ad}_{h^{-1}} T_{\exp(-t\eta)} R_{\exp(t\eta)} \xi_H(\exp(-t\eta)) \right\rangle \\
&= \left\langle \nu, \mathrm{Ad}_{h^{-1}} \frac{d}{dt}\bigg|_{t=0} T_{\exp(-t\eta)} R_{\exp(t\eta)} \xi_H(\exp(-t\eta)) \right\rangle.
\end{aligned}$$

However,

$$
\frac{d}{dt}\bigg|_{t=0} T_{\exp(-t\eta)} R_{\exp(t\eta)} \xi_H(\exp(-t\eta))
$$

$$
= \frac{d}{dt}\bigg|_{t=0} \frac{d}{ds}\bigg|_{s=0} R_{\exp(t\eta)} \left(\phi(\exp(s\xi))(\exp(-t\eta)) \right)
$$

$$
= \frac{d}{ds}\bigg|_{s=0} \frac{d}{dt}\bigg|_{t=0} \left(\phi(\exp(s\xi))(\exp(-t\eta)) \right) \exp(t\eta)
$$

$$
= \frac{d}{ds}\bigg|_{s=0} \left(\eta + \frac{d}{dt}\bigg|_{t=0} \phi(\exp(s\xi))(\exp(-t\eta)) \right)
$$

$$
= - \frac{d}{ds}\bigg|_{s=0} T_{e_H} \phi(\exp(s\xi))(\eta)
$$

$$
= - \frac{d}{ds}\bigg|_{s=0} \tilde{\phi}(\exp(s\xi))(\eta)
$$

$$
= -\xi \cdot \eta
$$

by the definition of the induced $\mathfrak{g}$-action on $\mathfrak{h}$. Thus, using $\rho = \operatorname{Ad}_{h^{-1}}^* \nu$, we get

$$
\left\langle \mathbf{dK}^\nu(\rho) \left(-\operatorname{ad}_\eta^* \rho \right), \xi \right\rangle = - \left\langle \nu, \operatorname{Ad}_{h^{-1}} (\xi \cdot \eta) \right\rangle = - \left\langle \rho, \xi \cdot \eta \right\rangle . \qquad (10.6.24)
$$

If $\xi \in \mathfrak{g}^\nu$, the infinitesimal generator of the G^ν-action (10.6.18) is

$$
\xi_{\mathcal{O}_\nu}(\rho) = \frac{d}{dt}\bigg|_{t=0} \tilde{\phi}(\exp t\xi)^* \rho = \tilde{\phi}'(\xi)^* \rho \in T_\rho \mathcal{O}_\nu
$$

by Corollary 10.6.2(ii). Thus there is some $\lambda(\xi, \rho) \in \mathfrak{h}$ such that $\tilde{\phi}'(\xi)^* \rho = -\operatorname{ad}_{\lambda(\xi,\rho)}^* \rho$ and hence from (10.6.24) we get

$$
\left\langle \mathbf{dK}^\nu(\rho) \left(-\operatorname{ad}_\eta^* \rho \right), \xi \right\rangle = - \left\langle \rho, \xi \cdot \eta \right\rangle = - \left\langle \rho, \tilde{\phi}'(\xi)(\eta) \right\rangle
$$

$$
= - \left\langle \tilde{\phi}'(\xi)^* \rho, \eta \right\rangle = - \left\langle -\operatorname{ad}_{\lambda(\xi,\rho)}^* \rho, \eta \right\rangle = \left\langle \rho, [\lambda(\xi, \rho), \eta] \right\rangle
$$

$$
= \omega_{\mathcal{O}_\nu}^+(\rho) \left(-\operatorname{ad}_{\lambda(\xi,\rho)}^* \rho, -\operatorname{ad}_\eta^* \rho \right) = \omega_{\mathcal{O}_\nu}^+(\rho) \left(\tilde{\phi}'(\xi)^* \rho, -\operatorname{ad}_\eta^* \rho \right)
$$

$$
= \omega_{\mathcal{O}_\nu}^+(\rho) \left(\xi_{\mathcal{O}_\nu}(\rho), -\operatorname{ad}_\eta^* \rho \right)
$$

which shows that $\mathbf{K}^\nu$ given by (10.6.22) is a momentum map of the G^ν-action (10.6.18) on the H-coadjoint orbit $\mathcal{O}_\nu$.

Next, we compute the non-equivariance one-cocycle of $\mathbf{K}^\nu$. Using the definition, we get for any $l \in G^\nu$ and $h \in H$

$$
\sigma^\nu(l) := \mathbf{K}^\nu \left(\operatorname{Ad}_{h^{-1}}^* \nu \cdot l \right) - \operatorname{Ad}_l^* \mathbf{K}^\nu \left(\operatorname{Ad}_{h^{-1}}^* \nu \right)
$$

$$
= \mathbf{K}^\nu \left(\tilde{\phi}(l)^* \operatorname{Ad}_{h^{-1}}^* \nu \right) - \operatorname{Ad}_l^* \mathbf{K}^\nu \left(\operatorname{Ad}_{h^{-1}}^* \nu \right)
$$

$$
= \mathbf{K}^\nu \left(\operatorname{Ad}_{l^{-1} \cdot h^{-1}}^* \tilde{\phi}(l)^* \nu \right) - \operatorname{Ad}_l^* \mathbf{K}^\nu \left(\operatorname{Ad}_{h^{-1}}^* \nu \right) .
$$

Since $l \in G^\nu$ there is some $k(l) \in H$ such that $\tilde{\phi}(l)^* \nu = \mathrm{Ad}^*_{k(l)^{-1}} \nu$ and hence, using (10.6.22), the expression above becomes

$$
\mathbf{K}^\nu \left(\mathrm{Ad}^*_{[(l^{-1} \cdot h) k(l)]^{-1}} \nu \right) - \mathrm{Ad}^*_l \mathbf{K}^\nu \left(\mathrm{Ad}^*_{h^{-1}} \nu \right)
$$

$$
= j^*_\nu T^*_{e_G} \phi^{[(l^{-1} \cdot h) k(l)]^{-1}} T^*_{[(l^{-1} \cdot h) k(l)]^{-1}} R_{(l^{-1} \cdot h) k(l)} \nu
$$

$$
\qquad - \mathrm{Ad}^*_l \, j^*_\nu T^*_{e_G} \phi^{h^{-1}} T^*_{h^{-1}} R_h \nu
$$

$$
= j^*_\nu \left[T^*_{e_G} \phi^{[(l^{-1} \cdot h) k(l)]^{-1}} T^*_{(l^{-1} \cdot h) k(l)]^{-1}} R_{(l^{-1} \cdot h) k(l)} \nu \right.
$$

$$
\left. \qquad - \mathrm{Ad}^*_l \, T^*_{e_G} \phi^{h^{-1}} T^*_{h^{-1}} R_h \nu \right]
$$

since $j^*_\nu \circ \mathrm{Ad}^*_l = \mathrm{Ad}^*_l \circ j^*_\nu$ for any $l \in G^\nu$. But $\tilde{\phi}(l)^* \nu = \mathrm{Ad}^*_{k(l)^{-1}} \nu$, so from the expression above we get for any $\xi \in \mathfrak{g}^\nu$

$$
\langle \sigma^\nu(l), \xi \rangle = \left\langle \nu, T_{e_G} \left(R_{(l^{-1} \cdot h) k(l)} \circ \phi^{[(l^{-1} \cdot h) k(l)]^{-1}} \right) (\xi) \right\rangle
$$

$$
\qquad - \left\langle \tilde{\phi}(l^{-1})^* \tilde{\phi}(l)^* \nu, T_{e_G} \left(R_h \circ \phi^{h^{-1}} \right) (\mathrm{Ad}_l \, \xi) \right\rangle
$$

$$
= \left\langle \nu, T_{e_G} \left(R_{(l^{-1} \cdot h) k(l)} \circ \phi^{[(l^{-1} \cdot h) k(l)]^{-1}} \right) (\xi) \right\rangle
$$

$$
\qquad - \left\langle \mathrm{Ad}^*_{k(l)^{-1}} \nu, \tilde{\phi}(l^{-1}) T_{e_G} \left(R_h \circ \phi^{h^{-1}} \right) (\mathrm{Ad}_l \, \xi) \right\rangle
$$

$$
= \left\langle \nu, T_{e_G} \left(R_{(l^{-1} \cdot h) k(l)} \circ \phi^{[(l^{-1} \cdot h) k(l)]^{-1}} \right) (\xi) \right\rangle
$$

$$
\qquad - \left\langle \nu, \mathrm{Ad}_{k(l)^{-1}} \tilde{\phi}(l^{-1}) T_{e_G} \left(R_h \circ \phi^{h^{-1}} \right) (\mathrm{Ad}_l \, \xi) \right\rangle. \quad (10.6.25)
$$

Now note that for any $g \in G^\nu$ we have

$$
\left(R_{(l^{-1} \cdot h) k(l)} \circ \phi^{[(l^{-1} \cdot h) k(l)]^{-1}} \right) (g) = \left(g \cdot k(l)^{-1} \right) \left((g l^{-1}) \cdot h^{-1} \right) \left(l^{-1} \cdot h \right) k(l)
$$

and

$$
\left(\mathrm{AD}_{k(l)^{-1}} \circ \phi(l^{-1}) \circ R_h \circ \phi^{h^{-1}} \circ \mathrm{AD}_l \right) (g)
$$

$$
= k(l)^{-1} \left((g l^{-1}) \cdot h^{-1} \right) \left(l^{-1} \cdot h \right) k(l)
$$

$$
= k(l)^{-1} \left(g \cdot k(l) \right) \left(R_{(l^{-1} \cdot h) k(l)} \circ \phi^{[(l^{-1} \cdot h) k(l)]^{-1}} \right) (g).
$$

Replacing in this last identity $g = \exp(t\xi)$, for $\xi \in \mathfrak{g}^\nu$, and taking the t-derivative of the resulting relation at $t = 0$, we get

$$
\mathrm{Ad}_{k(l)^{-1}} \tilde{\phi}(l^{-1}) T_{e_G} \left(R_h \circ \phi^{h^{-1}} \right) (\mathrm{Ad}_l \, \xi)
$$

$$
= \left. \frac{d}{dt} \right|_{t=0} k(l)^{-1} \left(\exp(t\xi) \cdot k(l) \right) + T_{e_G} \left(R_{(l^{-1} \cdot h) k(l)} \circ \phi^{[(l^{-1} \cdot h) k(l)]^{-1}} \right) (\xi)
$$

$$
= T_{k(l)} L_{k(l)^{-1}} T_{e_G} \phi^{k(l)} \xi + T_{e_G} \left(R_{(l^{-1} \cdot h) k(l)} \circ \phi^{[(l^{-1} \cdot h) k(l)]^{-1}} \right) (\xi)
$$

and hence, from (10.6.25) we get

$$\langle \sigma^\nu(l), \xi \rangle = - \left\langle \nu, T_{k(l)} L_{k(l)^{-1}} T_{e_G} \phi^{k(l)} \xi \right\rangle$$
$$= - \left\langle j_\nu^* T_{e_G}^* \phi^{k(l)} T_{k(l)}^* L_{k(l)^{-1}} \nu, \xi \right\rangle,$$

which proves (10.6.23). ■

Remarks. (i) If $H = V$ and the action ϕ is a representation, then $\nu = a \in V^*$, $G^\nu = G_a$, the coadjoint action of V on V^* is trivial, and hence if $l \in G_a$ we have $l \cdot a = 0$ which implies that we can choose $k(l) = 0$. Therefore, $\sigma^a \equiv 0$ and we recovered the result from Lemma 4.2.6 which states that the momentum map $\mathbf{J}_a : T^*G \to \mathfrak{g}_a^*$ is equivariant.

(ii) Let us study the case $G \circledS G$ formally, even though we already know that the Lie groups $G \circledS G$ and $G \times G$ are isomorphic. So take in the previous considerations $G = H$ and ϕ the conjugation. Then $G^\nu = G$ and $\tilde{\phi}(l)^* \nu = \mathrm{Ad}_l^* \nu$ which implies that we can choose $k(l) = l^{-1}$. Thus, for any $l \in G$ and any $\xi \in G$ we have

$$\left(T_{l^{-1}} L_l \circ T_{e_G} \phi^{l^{-1}} \right) (\xi) = \left. \frac{d}{dt} \right|_{t=0} \left(L_l \circ \phi^{l^{-1}} \right) (\exp(t\xi))$$
$$= \left. \frac{d}{dt} \right|_{t=0} \left(l \exp(t\xi) l^{-1} \exp(-t\xi) \right) = \mathrm{Ad}_l \xi - \xi.$$

Therefore, the non-equivariance cocycle in this case equals $\sigma^\nu(l) = \nu - \mathrm{Ad}_l^* \nu$, which is a coboundary. Thus the momentum map $\mathbf{K}^\nu$, and hence $\mathbf{J}_\nu$, can be modified in this case to be equivariant. Concretely, the new momentum maps $\mathbf{K}^\nu - \nu$ and $\mathbf{J}_\nu - \nu$ are equivariant. Of course, this result was expected since the semidirect product is isomorphic in this case to the direct product.

10.6.5 Corollary.   *The non-equivariance $(\mathfrak{g}^\nu)^*$-valued right one-cocycle of the momentum map $\mathbf{J}_\nu : T^*G \times \mathcal{O}_\nu \to (\mathfrak{g}^\nu)^*$ is $\sigma^\nu : G^\nu \to (\mathfrak{g}^\nu)^*$ given by (10.6.23).*

Proof.   Since the momentum map $\mathbf{J}_\nu : T^*G \times \mathcal{O}_\nu \to (\mathfrak{g}^\nu)^*$ is the sum of the equivariant momentum map of the lift of right translation of G^ν on T^*G and the momentum map $\mathbf{K}^\nu$, its non-equivariance one-cocycle coincides with that of $\mathbf{K}^\nu$, that is, with σ^ν. ■

Verification of the Cocycle Identity for σ^ν.   According to general theory, the cocycle σ^ν defined in (10.6.23) verifies the right cocycle identity

$$\sigma^\nu(l_1 l_2) = \mathrm{Ad}_{l_2}^* \sigma^\nu(l_1) + \sigma^\nu(l_2). \tag{10.6.26}$$

Since the definition of σ^ν involves a choice of an element $k(l) \in H$, it is not obvious from the formula of σ^ν why such a relation should hold. We shall verify it now, as a check of our previous computations.

The key observation is that the group elements $k(l) \in H$ satisfying $\tilde{\phi}(l)^*\nu = \mathrm{Ad}^*_{k(l)^{-1}}\nu$ must satisfy a certain relation even though they are determined only up to multiplication by elements of the coadjoint isotropy subgroup H_ν. Indeed, if $l_1, l_2 \in G^\nu$ we have $\mathrm{AD}_{k(l_1)^{-1}} \circ \phi(l_2) = \phi(l_2) \circ \mathrm{AD}_{l_2^{-1} \cdot k(l_1)^{-1}}$ and hence

$$
\begin{aligned}
\mathrm{Ad}^*_{k(l_1 l_2)^{-1}} \nu &= \tilde{\phi}(l_1 l_2)^* \nu = \tilde{\phi}(l_2)^* \tilde{\phi}(l_1)^* \nu = \tilde{\phi}(l_2)^* \mathrm{Ad}^*_{k(l_1)^{-1}} \nu \\
&= \mathrm{Ad}^*_{l_2^{-1} \cdot k(l_1)^{-1}} \tilde{\phi}(l_2)^* \nu = \mathrm{Ad}^*_{l_2^{-1} \cdot k(l_1)^{-1}} \mathrm{Ad}^*_{k(l_2)^{-1}} \nu \\
&= \mathrm{Ad}^*_{[(l_2^{-1} \cdot k(l_1)) k(l_2)]^{-1}} \nu,
\end{aligned}
$$

which is equivalent to the existence of some $h \in H_\nu$ such that

$$
k(l_1 l_2)^{-1} = k(l_2)^{-1} (l_2^{-1} \cdot k(l_1)^{-1}) h.
$$

Using this relation we shall verify now the cocycle identity (10.6.26). For $\xi \in \mathfrak{g}^\nu$ we have by (10.6.23)

$$
\begin{aligned}
\langle \sigma^\nu(l_1 l_2), \xi \rangle &= -\left\langle j_\nu^* T^*_{e_G} \phi^{k(l_1 l_2)} T^*_{k(l_1 l_2)} L_{k(l_1 l_2)^{-1}} \nu, \xi \right\rangle \\
&= -\left\langle \nu, T_{k(l_1 l_2)} L_{k(l_1 l_2)^{-1}} T_{e_G} \phi^{k(l_1 l_2)} \xi \right\rangle \\
&= -\left\langle \nu, T_{h^{-1}(l_2^{-1} \cdot k(l_1)) k(l_2)} L_{k(l_2)^{-1}(l_2^{-1} \cdot k(l_1)^{-1}) h} \xi_H(h^{-1}(l_2^{-1} \cdot k(l_1)) k(l_2)) \right\rangle.
\end{aligned}
$$

Now use the identity $\xi_H(hh') = T_h R_{h'} \xi_H(h) + T_{h'} L_h \xi_H(h')$ for any $h, h' \in H$ with h replaced by $h^{-1}(l_2^{-1} \cdot k(l_1))$ and h' by $k(l_2)$ to get

$$
\begin{aligned}
&T_{h^{-1}(l_2^{-1} \cdot k(l_1)) k(l_2)} L_{k(l_2)^{-1}(l_2^{-1} \cdot k(l_1)^{-1}) h} \xi_H(h^{-1}(l_2^{-1} \cdot k(l_1)) k(l_2)) \\
&= \left(T_{h^{-1}(l_2^{-1} \cdot k(l_1)) k(l_2)} L_{k(l_2)^{-1}(l_2^{-1} \cdot k(l_1)^{-1}) h} \circ T_{h^{-1}(l_2^{-1} \cdot k(l_1))} R_{k(l_2)} \right) \\
&\qquad\qquad \left(\xi_H(h^{-1}(l_2^{-1} \cdot k(l_1))) \right) \\
&+ \left(T_{h^{-1}(l_2^{-1} \cdot k(l_1)) k(l_2)} L_{k(l_2)^{-1}(l_2^{-1} \cdot k(l_1)^{-1}) h} \circ T_{k(l_2)} L_{h^{-1}(l_2^{-1} \cdot k(l_1))} \right) \left(\xi_H(k(l_2)) \right) \\
&= \mathrm{Ad}_{k(l_2)^{-1}} T_{h^{-1}(l_2^{-1} \cdot k(l_1))} L_{(l_2^{-1} \cdot k(l_1)^{-1}) h} \left(\xi_H(h^{-1}(l_2^{-1} \cdot k(l_1))) \right) \\
&\quad + T_{k(l_2)} L_{k(l_2)^{-1}} \left(\xi_H(k(l_2)) \right).
\end{aligned}
$$

In the first summand use again the identity $\xi_H(hh') = T_h R_{h'} \xi_H(h) + T_{h'} L_h \xi_H(h')$ but this time with h replaced by h^{-1} and h' by $l_2^{-1} \cdot k(l_1)$.

Thus the expression above becomes

$$
\begin{aligned}
&\mathrm{Ad}_{k(l_2)^{-1}}\, T_{h^{-1}(l_2^{-1}\cdot k(l_1))} L_{(l_2^{-1}\cdot k(l_1)^{-1})h} T_{h^{-1}} R_{l_2^{-1}\cdot k(l_1)}\left(\xi_H(h^{-1})\right) \\
&\quad + \mathrm{Ad}_{k(l_2)^{-1}}\, T_{h^{-1}(l_2^{-1}\cdot k(l_1))} L_{(l_2^{-1}\cdot k(l_1)^{-1})h} T_{l_2^{-1}\cdot k(l_1)} L_{h^{-1}}\left(\xi_H(l_2^{-1}\cdot k(l_1))\right) \\
&\quad + T_{k(l_2)} L_{k(l_2)^{-1}}\left(\xi_H(k(l_2))\right) \\
&= \mathrm{Ad}_{k(l_2)^{-1}}\, \mathrm{Ad}_{l_2^{-1}\cdot k(l_1)^{-1}}\, T_{h^{-1}} L_h\left(\xi_H(h^{-1})\right) \\
&\quad + \mathrm{Ad}_{k(l_2)^{-1}}\, T_{l_2^{-1}\cdot k(l_1)} L_{l_2^{-1}\cdot k(l_1)^{-1}}\left(\xi_H(l_2^{-1}\cdot k(l_1))\right) \\
&\quad + T_{k(l_2)} L_{k(l_2)^{-1}}\left(\xi_H(k(l_2))\right) \\
&= \mathrm{Ad}_{k(l_2)^{-1}}\, T_{l_2^{-1}\cdot k(l_1)} L_{l_2^{-1}\cdot k(l_1)^{-1}}\left(\xi_H(l_2^{-1}\cdot k(l_1))\right) \\
&\quad + T_{k(l_2)} L_{k(l_2)^{-1}}\left(\xi_H(k(l_2))\right)
\end{aligned}
$$

since the first term vanishes. This is seen in the following way:

$$
T_{h^{-1}} L_h\left(\xi_H(h^{-1})\right) = (L_{h^{-1}}^* \xi_H)(e_H) = (\mathrm{Ad}_h\, \xi)_H(e_H) = 0
$$

because for any $\zeta \in \mathfrak{g}$ we have $\zeta_H(e_H) = 0$. Summarizing, we have shown that

$$
\begin{aligned}
&T_{h^{-1}(l_2^{-1}\cdot k(l_1))k(l_2)} L_{k(l_2)^{-1}(l_2^{-1}\cdot k(l_1)^{-1})h} \xi_H(h^{-1}(l_2^{-1}\cdot k(l_1))k(l_2)) \\
&\quad = \mathrm{Ad}_{k(l_2)^{-1}}\, T_{l_2^{-1}\cdot k(l_1)} L_{l_2^{-1}\cdot k(l_1)^{-1}}\left(\xi_H(l_2^{-1}\cdot k(l_1))\right) \\
&\qquad + T_{k(l_2)} L_{k(l_2)^{-1}}\left(\xi_H(k(l_2))\right)
\end{aligned}
$$

and hence, using $\tilde{\phi}(l_2)^* \nu = \mathrm{Ad}_{k(l_2)^{-1}}^* \nu$, we get

$$
\begin{aligned}
\langle \sigma^\nu(l_1 l_2), \xi \rangle &= -\left\langle \nu, \mathrm{Ad}_{k(l_2)^{-1}}\, T_{l_2^{-1}\cdot k(l_1)} L_{l_2^{-1}\cdot k(l_1)^{-1}}\left(\xi_H(l_2^{-1}\cdot k(l_1))\right) \right\rangle \\
&\quad - \left\langle \nu, T_{k(l_2)} L_{k(l_2)^{-1}}\left(\xi_H(k(l_2))\right)\right\rangle \\
&= -\left\langle \mathrm{Ad}_{k(l_2)^{-1}}^* \nu, T_{l_2^{-1}\cdot k(l_1)} L_{l_2^{-1}\cdot k(l_1)^{-1}}\left(\xi_H(l_2^{-1}\cdot k(l_1))\right)\right\rangle \\
&\quad - \left\langle T_{k(l_2)}^* L_{k(l_2)^{-1}} \nu, T_{e_G} \phi^{k(l_2)} \xi\right\rangle \\
&= -\left\langle \tilde{\phi}(l_2)^* \nu, T_{l_2^{-1}\cdot k(l_1)} L_{l_2^{-1}\cdot k(l_1)^{-1}}\left(\xi_H(l_2^{-1}\cdot k(l_1))\right)\right\rangle \\
&\quad - \left\langle j_\nu^* T_{e_G}^* \phi^{k(l_2)} T_{k(l_2)}^* L_{k(l_2)^{-1}} \nu\right\rangle \\
&= -\left\langle \nu, \tilde{\phi}(l_2) T_{l_2^{-1}\cdot k(l_1)} L_{l_2^{-1}\cdot k(l_1)^{-1}}\left(\xi_H(l_2^{-1}\cdot k(l_1))\right)\right\rangle \\
&\quad + \langle \sigma^\nu(l_2), \xi \rangle.
\end{aligned}
$$

Since $\phi(l_2) \circ L_{l_2^{-1} \cdot k(l_1)^{-1}} = L_{k(l_1)^{-1}} \circ \phi(l_2)$ this becomes

$$
\begin{aligned}
\langle \sigma^\nu(l_1 l_2), \xi \rangle &= -\left\langle \nu, T_{k(l_1)} L_{k(l_1)^{-1}} T_{l_2^{-1} \cdot k(l_1)} \phi(l_2) \left(\xi_H(l_2^{-1} \cdot k(l_1)) \right) \right\rangle \\
&\quad + \langle \sigma^\nu(l_2), \xi \rangle \\
&= -\left\langle T_{k(l_1)}^* L_{k(l_1)^{-1}} \nu, \left(\phi(l_2^{-1})^* \xi_H \right)(k(l_1)) \right\rangle + \langle \sigma^\nu(l_2), \xi \rangle \\
&= -\left\langle T_{k(l_1)}^* L_{k(l_1)^{-1}} \nu, (\mathrm{Ad}_{l_2} \xi)_H (k(l_1)) \right\rangle + \langle \sigma^\nu(l_2), \xi \rangle \\
&= -\left\langle j_\nu^* T_{e_G}^* \phi^{k(l_1)} T_{k(l_1)}^* L_{k(l_1)^{-1}} \nu, \mathrm{Ad}_{l_2} \xi \right\rangle + \langle \sigma^\nu(l_2), \xi \rangle \\
&= \langle \sigma^\nu(l_1), \mathrm{Ad}_{l_2} \xi \rangle + \langle \sigma^\nu(l_2), \xi \rangle \\
&= \left\langle \mathrm{Ad}_{l_2}^* \sigma^\nu(l_1) + \sigma^\nu(l_2), \xi \right\rangle
\end{aligned}
$$

for any $\xi \in \mathfrak{g}^\nu$, which proves the identity (10.6.26).

The coadjoint orbits of $G \, \textcircled{s} \, H$. Theorem 6.4.1 guarantees that any element $(\mu, \nu) \in (\mathfrak{g} \, \textcircled{s} \, \mathfrak{h})^*$ satisfies the Stages Hypothesis of Definition 5.2.8. Therefore we can apply the Point Reduction by Stages Theorem 5.2.9 to obtain the following generalization of Theorem 4.3.2.

10.6.6 Theorem. *Let $(\mu, \nu) \in \mathfrak{g}^* \times \mathfrak{h}^*$, define $\mu_\nu := \mu|_{\mathfrak{g}^\nu} \in (\mathfrak{g}^\nu)^*$, and let $(G^\nu)_{\mu_\nu}$ be the isotropy subgroup of G^ν at μ_ν for the affine action*

$$
\tau \cdot l := \mathrm{Ad}_l^* \tau + \sigma^\nu(l), \quad where \quad \tau \in (\mathfrak{g}^\nu)^*, \quad l \in G^\nu. \tag{10.6.27}
$$

Then the $(G \, \textcircled{s} \, H)$-coadjoint orbit through $(\mu, \nu) \in (\mathfrak{g} \, \textcircled{s} \, \mathfrak{h})^*$, endowed with the plus orbit symplectic form, is symplectically diffeomorphic to the reduced space $\mathbf{J}_\nu^{-1}(\mu_\nu)/(G^\nu)_{\mu_\nu}$.

In particular, note that all coadjoint orbits $\mathcal{O}_{(\mu, \nu)}$ with the same $\mu_\nu \in (\mathfrak{g}^\nu)^*$ are symplectically diffeomorphic.

Reduction in Right Trivialization. Theorem 10.6.6 is useful in classifying orbits of semidirect products. General statements are not possible since the detailed topological and symplectic nature of these orbits depends on (μ, ν). In fact, general statements on the classification of the coadjoint orbits were not possible even in the case when H is a vector space and the G-action is a representation. However, we shall make some remarks regarding the reduced spaces $\mathbf{J}_\nu^{-1}(\mu_\nu)/(G^\nu)_{\mu_\nu}$ that should help in concrete examples.

We begin by recalling that $\mathbf{J}_\nu : T^*G \times \mathcal{O}_\nu \to (\mathfrak{g}^\nu)^*$ has the expression $\mathbf{J}_\nu(\beta_g, \mathrm{Ad}_{h^{-1}}^* \nu) = j_\nu^* T_{e_G}^* L_g \beta_g + j_\nu^* T_{e_G}^* \phi^{h^{-1}} T_{h^{-1}}^* R_h \nu$ (see (10.6.21)), where $j_\nu : \mathfrak{g}^\nu \hookrightarrow \mathfrak{g}$ is the inclusion. To explicitly carry out computations in concrete examples, it is useful to trivialize T^*G. Since we have worked here exclusively with right actions, formula (10.6.9) simplifies considerably if we trivialize T^*G using right translations. Thus we use the diffeomorphism $\alpha_g \in T^*G \mapsto (g, T_{e_G}^* R_g \alpha_g) \in G \times \mathfrak{g}^*$ whose inverse is given by

$(g, \kappa) \in G \times \mathfrak{g}^* \mapsto T_g^* R_{g^{-1}} \kappa \in T^* G$. The G^ν-action (10.6.9) becomes

$$(g, \kappa, \rho) \cdot l = (gl, \kappa, \tilde{\phi}(l)^* \rho) \tag{10.6.28}$$

for any $g \in G$, $\mu \in \mathfrak{g}^*$, $\rho \in \mathcal{O}_\nu$, and $l \in G^\nu$. The canonical cotangent bundle (T^*G, ω_G) is replaced by the symplectic manifold $(G \times \mathfrak{g}^*, \omega_G^R)$, where

$$\omega_G^R(g, \kappa) \left((T_{e_G} R_g \xi_1, \mu_1), (T_{e_G} R_g \xi_2, \mu_2) \right)$$
$$= \langle \mu_2, \xi_1 \rangle - \langle \mu_1, \xi_2 \rangle - \langle \kappa, [\xi_1, \xi_2] \rangle \tag{10.6.29}$$

for any $g \in G$, $\xi_1, \xi_2 \in \mathfrak{g}$, and $\mu, \mu_1, \mu_2 \in \mathfrak{g}^*$ (see, for example, Proposition 4.4.1 in [FofM] or Theorem 6.2.4 in [HRed]). The momentum map $\mathbf{J}_\nu : T^*G \times \mathcal{O}_\nu \to (\mathfrak{g}^\nu)^*$ is replaced by the momentum map $\mathbf{J}_\nu^R : G \times \mathfrak{g}^* \times \mathcal{O}_\nu \to (\mathfrak{g}^\nu)^*$ given by

$$\mathbf{J}_\nu^R(g, \kappa, \mathrm{Ad}_{h^{-1}}^* \nu) = j_\nu^* \mathrm{Ad}_g^* \kappa + j_\nu^* T_{e_G}^* \phi^{h^{-1}} T_{h^{-1}}^* R_h \nu. \tag{10.6.30}$$

Therefore, if $\mu \in \mathfrak{g}^*$ and $\mu_\nu := j_\nu^* \mu \in (\mathfrak{g}^\nu)^*$, we get

$(\mathbf{J}_\nu^R)^{-1}(\mu_\nu) =$
$$\left\{ (g, \kappa, \mathrm{Ad}_{h^{-1}}^* \nu) \in G \times \mathfrak{g}^* \times \mathcal{O}_\nu \,\Big|\, j_\nu^* \left(\mathrm{Ad}_g^* \kappa + T_{e_G}^* \phi^{h^{-1}} T_{h^{-1}}^* R_h \nu - \mu \right) = 0 \right\}$$

and the reduced space $(\mathbf{J}_\nu^R)^{-1}(\mu_\nu)/(G^\nu)_{\mu_\nu}$ is computed using the action (10.6.28). Here $(G^\nu)_{\mu_\nu}$ is the isotropy subgroup of G^ν at μ_ν relative to the affine action (10.6.27).

We compute now explicitly the symplectic diffeomorphism between the second stage reduced space $(\mathbf{J}_\nu^R)^{-1}(\mu_\nu)/(G^\nu)_{\mu_\nu}$ and the coadjoint orbit $\mathcal{O}_{(\mu,\nu)} \subset (\mathfrak{g} \circledS \mathfrak{h})^*$ given in Theorem 10.6.6. In the abstract context of Theorem 5.2.9 (Point Reduction by Stages I) this is the map $[\phi]$. Recall from §5.2 that this symplectic diffeomorphism $[\phi] : \mathbf{J}_\nu^{-1}(\rho)/(M_\nu/N_\nu)_\rho \to P_\sigma$ is the $(M_\nu/N_\nu)_\rho$-quotient of the map $\phi : \mathbf{J}_\nu^{-1}(\rho) \subset P_\nu = \mathbf{J}_N^{-1}(\nu)/N_\nu \to P_\sigma$ defined by $\phi([z]) := \pi_\sigma(n \cdot z)$, where $[z] \in \mathbf{J}_\nu^{-1}(\rho) \subset P_\nu$, $n \in N_\nu \cdot (M_\nu)_{\sigma|_{\mathfrak{m}_\nu}}$ is chosen from the Stages Hypothesis to satisfy $\mathrm{Ad}_{n^{-1}}^* \sigma' = \sigma$, $\sigma' := \mathbf{J}_M(z)$, $\pi_\sigma : \mathbf{J}_M^{-1}(\sigma) \to \mathbf{J}_M^{-1}(\sigma)/M_\sigma = P_\sigma$ is the projection, $(r_\nu')^*(\rho) = \sigma|_{\mathfrak{m}_\nu} - \bar{\nu}$, $r_\nu' : \mathfrak{m}_\nu \to \mathfrak{m}_\nu/\mathfrak{n}_\nu$ is the projection, and $\bar{\nu}$ is an arbitrary chosen extension of $\nu|_{\mathfrak{n}_\nu}$ to $\mathfrak{m}_\nu$. In our case, $M = G \circledS H$, $N = H$, $P = T^*(G \circledS H)$, the group actions are on the right, $\sigma = (\mu, \nu)$, $M_\nu/N_\nu = G^\nu$, and $[T^* \Upsilon^{-1}]_\nu : \mathbf{J}_H^{-1}(\nu)/H_\nu \to T^*G \times \mathcal{O}_\nu$ given by (10.5.6) is a G^ν-equivariant symplectic diffeomorphism relative to the lift of right translation and the action (10.6.9). In addition, in §6.4 it was shown that the Stages Hypothesis holds in this case with the choice of a group element of the form $(e_G, \exp \eta)$, where $\eta \in \mathfrak{h}_\nu$.

So let $(g, \kappa, \mathrm{Ad}_{h^{-1}}^* \nu) \in (\mathbf{J}_\nu^R)^{-1}(\mu_\nu)$, that is,

$$j_\nu^* \left(\mathrm{Ad}_g^* \kappa + T_{e_G}^* \phi^{h^{-1}} T_{h^{-1}}^* R_h \nu - \mu \right) = 0. \tag{10.6.31}$$

A direct computation, using (10.5.6), shows that

$$[T^*\Upsilon^{-1}]_\nu \left(\left[T_g^* R_{g^{-1}}\kappa - T_g^* L_{g^{-1}} T_{e_G}^* \phi^h T_h^* L_{h^{-1}}\nu, \; T_{g\cdot h}^* L_{g\cdot h^{-1}} \tilde{\phi}(g^{-1})^*\nu \right] \right)$$

$$= (g, \kappa, \mathrm{Ad}_{h^{-1}}^* \nu) \in \left(\mathbf{J}_\nu^R \right)^{-1}(\mu_\nu) \subset G \times \mathfrak{g}^* \times \mathcal{O}_\nu.$$

By (10.5.3) we have

$$\left(T_g^* R_{g^{-1}}\kappa - T_g^* L_{g^{-1}} T_{e_G}^* \phi^h T_h^* L_{h^{-1}}\nu, \; T_{g\cdot h}^* L_{g\cdot h^{-1}} \tilde{\phi}(g^{-1})^*\nu \right) \in \mathbf{J}_H^{-1}(\nu)$$

and hence, since the expression of the momentum map for the cotangent lift of right translation on $S = G \, \circledS \, H$ is given by the cotangent lift of left translation, that is,

$$\mathbf{J}_S\left(\alpha_g, \beta_h\right) = \left(T_{e_G}^* L_g \alpha_g, \; \tilde{\phi}(g)^* T_{e_H}^* L_h \beta_h \right),$$

it follows that

$$\mathbf{J}_S\left(T_g^* R_{g^{-1}}\kappa - T_g^* L_{g^{-1}} T_{e_G}^* \phi^h T_h^* L_{h^{-1}}\nu, \; T_{g\cdot h}^* L_{g\cdot h^{-1}} \tilde{\phi}(g^{-1})^*\nu \right)$$

$$= \left(\mathrm{Ad}_g^* \kappa - T_{e_G}^* \phi^h T_h^* L_{h^{-1}}\nu, \; \nu \right)$$

$$= \left(\mathrm{Ad}_g^* \kappa + T_{e_G}^* \phi^{h^{-1}} T_{h^{-1}}^* R_h \nu, \; \nu \right). \tag{10.6.32}$$

The last equality in (10.6.32) is a consequence of the fact that for any $\xi \in \mathfrak{g}$ we have

$$\left(T_{h^{-1}} R_h \circ T_{e_G}\phi^{h^{-1}} + T_h L_{h^{-1}} \circ T_{e_G}\phi^h \right)(\xi)$$

$$= \left. \frac{d}{dt} \right|_{t=0} \left(\left(R_h \circ \phi^{h^{-1}} \right)(\exp(t\xi)) \left(L_{h^{-1}} \circ \phi^h \right)(\exp(t\xi)) \right)$$

$$= \left. \frac{d}{dt} \right|_{t=0} \left((\exp(t\xi) \cdot h^{-1}) h h^{-1} ((\exp(t\xi) \cdot h)) \right) = 0. \tag{10.6.33}$$

As expected from the general theory in §6.4, the second component in (10.6.32) is ν.

By (10.6.31), the difference $\mathrm{Ad}_g^* \kappa + T_{e_G}^* \phi^{h^{-1}} T_{h^{-1}}^* R_h \nu - \mu$ between the first components of (10.6.32) and (μ, ν) annihilates $\mathfrak{g}^\nu$ and thus by the considerations in §6.4 there is some $\eta \in \mathfrak{h}_\nu$ such that

$$\mathrm{Ad}_g^* \kappa + T_{e_G}^* \phi^{h^{-1}} T_{h^{-1}}^* R_h \nu - \mu = \left(\tilde{\phi}'_\eta \right)^* \nu. \tag{10.6.34}$$

(In §6.4 the notation $\mathfrak{g}^\nu$ was not yet introduced and it was called there $\mathfrak{k}$.) Therefore, denoting $l := \exp \eta \in H_\nu$, general theory and (10.4.8) give

$$\mathbf{J}_S^{-1}(\mu, \nu) \ni \left(T_g^* R_{g^{-1}} \kappa - T_g^* L_{g^{-1}} T_{e_G}^* \phi^h T_h^* L_{h^{-1}} \nu, \; T_{g \cdot h}^* L_{g \cdot h^{-1}} \tilde{\phi}(g^{-1})^* \nu \right) \cdot l$$

$$= \left(T_g^* R_{g^{-1}} \kappa - T_g^* L_{g^{-1}} T_{e_G}^* \phi^h T_h^* L_{h^{-1}} \nu \right.$$

$$+ T_g^* \phi^{l^{-1}} T_{g \cdot l^{-1}}^* L_{(g \cdot h)(g \cdot l)} T_{g \cdot h}^* L_{g \cdot h^{-1}} \tilde{\phi}(g^{-1})^* \nu,$$

$$\left. T_{(g \cdot h)(g \cdot l)}^* R_{g \cdot l^{-1}} T_{g \cdot h}^* L_{g \cdot h^{-1}} \tilde{\phi}(g^{-1})^* \nu \right)$$

$$= \left(T_g^* R_{g^{-1}} \kappa - T_g^* L_{g^{-1}} T_{e_G}^* \phi^h T_h^* L_{h^{-1}} \nu + T_g^* L_{g^{-1}} T_{e_G}^* \phi^{l^{-1}} T_{l^{-1}}^* L_l \nu, \right.$$

$$\left. T_{g \cdot (hl)}^* R_{g \cdot l^{-1}} T_{g \cdot h}^* L_{g \cdot h^{-1}} \tilde{\phi}(g^{-1})^* \nu \right). \tag{10.6.35}$$

The last equality involves only the third summand of the first component. This is a consequence of the easily verified identity

$$\phi(g^{-1}) \circ L_{g \cdot h^{-1}} \circ L_{(g \cdot h)(g \cdot l)} \circ \phi^{l^{-1}} = L_l \circ \phi^{l^{-1}} \circ L_g^{-1}.$$

The expression (10.6.35) can be further simplified. We begin with its second component. The identity

$$\mathrm{AD}_{l^{-1}} \circ \phi(g^{-1}) \circ L_{g \cdot h^{-1}} \circ R_{g \cdot l^{-1}} = \phi(g^{-1}) \circ L_{g \cdot (hl)^{-1}}$$

together with the fact that $l \in H_\nu$ implies

$$T_{g \cdot (hl)}^* R_{g \cdot l^{-1}} T_{g \cdot h}^* L_{g \cdot h^{-1}} \tilde{\phi}(g^{-1})^* \nu$$

$$= T_{g \cdot (hl)}^* R_{g \cdot l^{-1}} T_{g \cdot h}^* L_{g \cdot h^{-1}} \tilde{\phi}(g^{-1})^* \mathrm{Ad}_{l^{-1}}^* \nu$$

$$= T_{g \cdot (hl)}^* L_{g \cdot (hl)^{-1}} \tilde{\phi}(g^{-1})^* \nu. \tag{10.6.36}$$

The first component of (10.6.35) will be simplified by showing that

$$T_{e_G}^* \phi^h T_h^* L_{h^{-1}} \nu - T_{e_G}^* \phi^{l^{-1}} T_{l^{-1}}^* L_l \nu = T_{e_G}^* \phi^{hl} T_{hl}^* L_{(hl)^{-1}} \nu. \tag{10.6.37}$$

Using $\mathrm{Ad}_l^* \nu = \nu$ in the right hand side, this is equivalent to

$$T_{e_G}^* \phi^{hl} T_{hl}^* L_{(hl)^{-1}} \mathrm{Ad}_l^* \nu + T_{e_G}^* \phi^{l^{-1}} T_{l^{-1}}^* L_l \nu = T_{e_G}^* \phi^h T_h^* L_{h^{-1}} \nu$$

which is a consequence of the identity

$$\mathrm{Ad}_l \circ T_{hl} L_{(hl)^{-1}} \circ T_{e_G} \phi^{hl} + T_{l^{-1}} L_l \circ T_{e_G} \phi^{l^{-1}} = T_h L_{h^{-1}} \circ T_{e_G} \phi^h.$$

This is in turn proved as follows. Since for any $g' \in G$,

$$\left(\mathrm{AD}_l \circ L_{(hl)^{-1}} \circ \phi^{hl} \right)(g') \left(L_l \circ \phi^{l^{-1}} \right)(g') = \left(L_{h^{-1}} \circ \phi^h \right)(g'),$$

it follows that for any $\xi \in \mathfrak{g}$ we have

$$\left(\mathrm{Ad}_l \circ T_{hl}L_{(hl)^{-1}} \circ T_{e_G}\phi^{hl}\right)(\xi) + \left(T_{l^{-1}}L_l \circ T_{e_G}\phi^{l^{-1}}\right)(\xi)$$

$$= \left.\frac{d}{dt}\right|_{t=0} \left[\left(\mathrm{AD}_l \circ L_{(hl)^{-1}} \circ \phi^{hl}\right)(\exp(t\xi)) \left(L_l \circ \phi^{l^{-1}}\right)(\exp(t\xi))\right]$$

$$= \left.\frac{d}{dt}\right|_{t=0} \left[\left(L_{h^{-1}} \circ \phi^h\right)(\exp(t\xi))\right]$$

$$= \left(T_h L_{h^{-1}} \circ T_{e_G}\phi^h\right)(\xi)$$

as required. Putting now (10.6.35), (10.6.37), and (10.6.36) together, we conclude that

$$\mathbf{J}_S^{-1}(\mu, \nu) \ni \left(T_g^* R_{g^{-1}}\kappa - T_g^* L_{g^{-1}} T_{e_G}^* \phi^{hl} T_{hl}^* L_{(hl)^{-1}}\nu,\right.$$

$$\left. T_{g\cdot(hl)}^* L_{g\cdot(hl)^{-1}} \tilde{\phi}(g^{-1})^* \nu\right). \tag{10.6.38}$$

Finally we recall that the projection map $\pi_{(\mu,\nu)} : \mathbf{J}_S^{-1}(\mu, \nu) \to \mathcal{O}_{(\mu,\nu)}$ relative to the right $S_{(\mu,\nu)}$-action is given by right translation to the identity (that is, by the momentum map of the cotangent lift of left translation on S). Since

$$T_{(e_G, e_H)}^* R_{(g,h)}\left(\alpha_g, \beta_h\right) = \left(T_{e_G}^* R_g \alpha_g + T_{e_G}^* \phi^h \beta_h, \; T_{e_H}^* R_h \beta_h\right)$$

we conclude that

$$\pi_{(\mu,\nu)}\left(T_g^* R_{g^{-1}}\kappa - T_g^* L_{g^{-1}} T_{e_G}^* \phi^{hl} T_{hl}^* L_{(hl)^{-1}}\nu, \; T_{g\cdot(hl)}^* L_{g\cdot(hl)^{-1}}\tilde{\phi}(g^{-1})^* \nu\right)$$

$$= \left(\kappa - \mathrm{Ad}_{g^{-1}}^* T_{e_G}^* \phi^{hl} T_{hl}^* L_{(hl)^{-1}}\nu + T_{e_G}^* \phi^{g\cdot(hl)} T_{g\cdot(hl)}^* L_{g\cdot(hl)^{-1}}\tilde{\phi}(g^{-1})\nu,\right.$$

$$\left. T_{e_H}^* R_{g\cdot(hl)} T_{g\cdot(hl)}^* L_{g\cdot(hl)^{-1}}\tilde{\phi}(g^{-1})^* \nu\right).$$

However,

$$T_{e_H}^* R_{g\cdot(hl)} T_{g\cdot(hl)}^* L_{g\cdot(hl)^{-1}}\tilde{\phi}(g^{-1})^* \nu = \mathrm{Ad}_{g\cdot(hl)^{-1}}^* \tilde{\phi}(g^{-1})^* \nu$$

$$= \tilde{\phi}(g^{-1})^* \mathrm{Ad}_{(hl)^{-1}}^* \nu = \tilde{\phi}(g^{-1})^* \mathrm{Ad}_{h^{-1}}^* \mathrm{Ad}_{l^{-1}}^* \nu = \tilde{\phi}(g^{-1})^* \mathrm{Ad}_{h^{-1}}^* \nu$$

since $l \in H_\nu$. Therefore we get

$$\pi_{(\mu,\nu)}\left(T_g^* R_{g^{-1}}\kappa - T_g^* L_{g^{-1}} T_{e_G}^* \phi^{hl} T_{hl}^* L_{(hl)^{-1}}\nu, \; T_{g\cdot(hl)}^* L_{g\cdot(hl)^{-1}}\tilde{\phi}(g^{-1})^* \nu\right)$$

$$= \left(\kappa - \mathrm{Ad}_{g^{-1}}^* T_{e_G}^* \phi^{hl} T_{hl}^* L_{(hl)^{-1}}\nu + T_{e_G}^* \phi^{g\cdot(hl)} T_{g\cdot(hl)}^* L_{g\cdot(hl)^{-1}}\tilde{\phi}(g^{-1})\nu,\right.$$

$$\left. \tilde{\phi}(g^{-1})^* \mathrm{Ad}_{h^{-1}}^* \nu\right).$$

Now we simplify the first component. The following identity for any $k \in H$ is readily verified

$$\phi(g^{-1}) \circ L_{g\cdot k^{-1}} \circ \phi^{g\cdot k} = L_{k^{-1}} \circ \phi^k \circ \mathrm{AD}_{g^{-1}}.$$

The dual of the derivative of this identity for $k = hl$ gives the cancellation of the last two summands in the formula above. Thus we finally get,

$$\pi_{(\mu,\nu)}\left(T_g^* R_{g^{-1}}\kappa - T_g^* L_{g^{-1}} T_{e_G}^* \phi^{hl} T_{hl}^* L_{(hl)^{-1}}\nu,\; T_{g\cdot(hl)}^* L_{g\cdot(hl)^{-1}}\tilde\phi(g^{-1})^*\nu\right)$$
$$= \left(\kappa,\, \tilde\phi(g^{-1})^*\operatorname{Ad}_{h^{-1}}^*\nu\right).$$

We have proved the following.

10.6.7 Proposition.   *The map*

$$[g,\kappa,\operatorname{Ad}_{h^{-1}}^*\nu] \in \left(\mathbf{J}_\nu^R\right)^{-1}(\mu_\nu)/(G^\nu)_{\mu_\nu}$$
$$\mapsto \left(\kappa,\tilde\phi(g^{-1})^*\operatorname{Ad}_{h^{-1}}^*\nu\right) \in \mathcal{O}_{(\mu,\nu)} \qquad (10.6.39)$$

is the symplectic diffeomorphism in Theorem 10.6.6.

As a non-trivial cross-check, let us show that the right hand side in (10.6.39) is indeed an element of the coadjoint orbit $\mathcal{O}_{(\mu,\nu)}$. This is far from obvious in view of formula (6.4.3) for the coadjoint action of $S := G \,\text{Ⓢ}\, H$ on $(\mathfrak{g}\,\text{Ⓢ}\,\mathfrak{h})^* = \mathfrak{g}^* \times \mathfrak{h}^*$. Thus to prove that the right hand side of (10.6.39) is an element of $\mathcal{O}_{(\mu,\nu)}$ we need to show that there exists some $(\bar g,\bar h) \in G\,\text{Ⓢ}\,H$ such that

$$\left(\kappa,\tilde\phi(g^{-1})^*\operatorname{Ad}_{h^{-1}}^*\nu\right) = \operatorname{Ad}_{(\bar g,\bar h)^{-1}}^*(\mu,\nu)$$
$$= \left(\operatorname{Ad}_{\bar g^{-1}}^*\mu + \left(\bar\phi\circ\operatorname{Ad}_{\bar g^{-1}}\right)^{\#}(\bar g^{-1}\cdot\bar h)^*\left(T_{\bar g^{-1}\cdot\bar h}^* L_{\bar g^{-1}\cdot\bar h^{-1}}\nu\right),\right.$$
$$\left.\tilde\phi(\bar g^{-1})^*\operatorname{Ad}_{\bar g^{-1}\cdot\bar h^{-1}}^*\nu\right).$$

Looking at the second component, this suggests taking $\bar g = g$ and $(hl)^{-1} = g^{-1}\cdot\bar h^{-1}$, that is, $\bar h = g\cdot(hl)$ for some $l \in H_\nu$ to be determined. With these choices the second components are identical, so it remains to show that

$$\kappa = \operatorname{Ad}_{g^{-1}}^*\mu + \left(\bar\phi\circ\operatorname{Ad}_{g^{-1}}\right)^{\#}(hl)^*\left(T_{hl}^* L_{(hl)^{-1}}\nu\right)$$

which is equivalent to

$$\operatorname{Ad}_g^*\kappa - \operatorname{Ad}_g^*\left(\bar\phi\circ\operatorname{Ad}_{g^{-1}}\right)^{\#}(hl)^*\left(T_{hl}^* L_{(hl)^{-1}}\nu\right) - \mu = 0. \qquad (10.6.40)$$

However, using the definitions of $\#$ and $\bar{\phi}$, for any $\xi \in \mathfrak{g}$ we get

$$\left\langle \mathrm{Ad}_g^* \left(\bar{\phi} \circ \mathrm{Ad}_{g^{-1}} \right)^\# (hl)^* \left(T_{hl}^* L_{(hl)^{-1}} \nu \right), \xi \right\rangle$$

$$= \left\langle T_{hl}^* L_{(hl)^{-1}} \nu, \left(\bar{\phi} \circ \mathrm{Ad}_{g^{-1}} \right)^\# (hl) \left(\mathrm{Ad}_g \xi \right) \right\rangle$$

$$= \left\langle T_{hl}^* L_{(hl)^{-1}} \nu, \left(\bar{\phi} \circ \mathrm{Ad}_{g^{-1}} \right) \left(\mathrm{Ad}_g \xi \right) (hl) \right\rangle$$

$$= \left\langle T_{hl}^* L_{(hl)^{-1}} \nu, \bar{\phi}(\xi)(hl) \right\rangle$$

$$= \left\langle T_{hl}^* L_{(hl)^{-1}} \nu, \xi_H(hl) \right\rangle$$

$$= \left\langle \nu, T_{hl} L_{(hl)^{-1}} T_{e_G} \phi^{hl} \xi \right\rangle$$

$$= - \left\langle \nu, T_{(hl)^{-1}} R_{hl} T_{e_G} \phi^{(hl)^{-1}} \xi \right\rangle$$

by (10.6.33). Therefore

$$\mathrm{Ad}_g^* \left(\bar{\phi} \circ \mathrm{Ad}_{g^{-1}} \right)^\# (hl)^* \left(T_{hl}^* L_{(hl)^{-1}} \nu \right) = - T_{e_G}^* \phi^{(hl)^{-1}} T_{(hl)^{-1}}^* R_{hl} \nu$$

and thus the identity (10.6.40) that needs to be verified is equivalent to

$$\mathrm{Ad}_g^* \kappa + T_{e_G}^* \phi^{(hl)^{-1}} T_{(hl)^{-1}}^* R_{hl} \nu - \mu = 0. \tag{10.6.41}$$

Recall, however, that (10.6.31) holds and hence, as we have seen before, there is some $\eta \in \mathfrak{h}_\nu$ such that (10.6.34) is verified. Let us choose now $l := \exp \eta \in H_\nu$. Then (10.6.41) holds, and hence

$$\left(\kappa, \tilde{\phi}(g^{-1})^* \mathrm{Ad}_{h^{-1}}^* \nu \right) \in \mathcal{O}_{(\mu,\nu)},$$

provided that we can show

$$T_{e_G}^* \phi^h T_{h^{-1}}^* R_h \nu - \left(\tilde{\phi}_\eta' \right)^* \nu = T_{e_G}^* \phi^{(hl)^{-1}} T_{(hl)^{-1}}^* R_{hl} \nu. \tag{10.6.42}$$

To prove this identity, we compute for any $\xi \in \mathfrak{g}$

$$\left\langle T_{e_G}^* \phi^{(hl)^{-1}} T_{(hl)^{-1}}^* R_{hl} \nu, \xi \right\rangle$$

$$= \left\langle \nu, T_{(hl)^{-1}} R_{hl} \xi_H(l^{-1} h^{-1}) \right\rangle$$

$$= \left\langle \nu, T_{(hl)^{-1}} R_{hl} T_{h^{-1}} L_{l^{-1}} \xi_H(h^{-1}) + T_{(hl)^{-1}} R_{hl} T_{l^{-1}} R_{h^{-1}} \xi_H(l^{-1}) \right\rangle.$$

The second equality above follows from the identity

$$\xi_H(kk') = T_{k'} L_k \left(\xi_H(k') \right) + T_k R_{k'} \left(\xi_H(k) \right)$$

for all $k, k' \in H$ and $\xi \in \mathfrak{g}$, as was discussed in §6.4. Since left and right translations commute, this becomes

$$\left\langle \nu, \mathrm{Ad}_{l^{-1}} T_{h^{-1}} R_h \xi_H(h^{-1}) \right\rangle + \left\langle \nu, T_{l^{-1}} R_l \xi_H(l^{-1}) \right\rangle$$

$$= \left\langle \mathrm{Ad}_{l^{-1}}^* \nu, T_{h^{-1}} R_h \xi_H(h^{-1}) \right\rangle + \left\langle \nu, T_{l^{-1}} R_l \xi_H(l^{-1}) \right\rangle$$

$$= \left\langle \nu, T_{h^{-1}} R_h \xi_H(h^{-1}) \right\rangle + \left\langle \nu, T_{l^{-1}} R_l \xi_H(l^{-1}) \right\rangle$$

because $l = \exp \eta \in H_\nu$. Therefore we get

$$\left\langle T^*_{e_G} \phi^{(hl)^{-1}} T^*_{(hl)^{-1}} R_{hl} \nu, \, \xi \right\rangle$$
$$= \left\langle \nu, \, T_{h^{-1}} R_h T_{e_G} \phi^{h^{-1}} \xi \right\rangle + \left\langle \nu, \, T_{l^{-1}} R_l \xi_H(l^{-1}) \right\rangle$$

and hence (10.6.42) holds if and only if

$$\left\langle \nu, \, \tilde{\phi}(\xi)(\eta) \right\rangle + \left\langle \nu, \, T_{l^{-1}} R_l \xi_H(l^{-1}) \right\rangle = 0. \tag{10.6.43}$$

In §6.4 we have seen that

$$\xi_H(k^{-1}) = - \left(T_{e_H} L_{k^{-1}} \circ T_k R_{k^{-1}} \right) (\xi_H(k))$$

for all $k \in H$ and $\xi \in \mathfrak{g}$. Therefore (10.6.43) is equivalent to

$$\left\langle \nu, \, \tilde{\phi}(\xi)(\eta) \right\rangle - \left\langle \nu, \, T_l L_{l^{-1}} \xi_H(l) \right\rangle = 0$$

which is true by (6.4.6). This shows that the right hand side of (10.6.39) is indeed an element of the orbit $\mathcal{O}_{(\mu,\nu)}$, as required.

The formulas given above serve as a guideline in concrete examples when one classifies coadjoint orbits of general semidirect products.

The Case $G^\nu = G$.   It is interesting to investigate the situation when $G^\nu = G$ which occurs, for example, if H admits only inner automorphisms. In this case, G acts on $T^*G \times \mathcal{O}_\nu$ and Proposition 10.3.3 can be applied to give the following result.

10.6.8 Proposition.   *Let $(\mu, \nu) \in \mathfrak{g}^* \times \mathfrak{h}^*$, assume that $G^\nu = G$, let $\sigma^\nu : G \to \mathfrak{g}^*$ be the non-equivariance cocycle of the momentum map $\mathbf{J}_\nu : T^*G \times \mathcal{O}_\nu \to \mathfrak{g}^*$, and let*

$$\Sigma^\nu(\xi, \eta) := \mathbf{J}_\nu^{[\xi,\eta]} + \left\{ \mathbf{J}_\nu^\xi, \mathbf{J}_\nu^\eta \right\}$$

be the associated infinitesimal two-cocycle. Denote by $\mathcal{O}_\mu^\Sigma$ the affine orbit of G through μ (relative to the action (10.6.27)) endowed with the plus orbit symplectic form. Then the $(G \circledS H)$-coadjoint orbit through $(\mu, \nu) \in (\mathfrak{g} \circledS \mathfrak{h})^*$, endowed with the plus orbit symplectic form, is symplectically diffeomorphic to the product symplectic manifold $\mathcal{O}_\mu^\Sigma \times \mathcal{O}_\nu$, where both orbits carry their respective plus orbit symplectic forms.

An important special instance of this situation occurs when $\nu = 0$. In this case $G^\nu = G$, $\mathcal{O}_\nu = \mathcal{O}$, and hence G acts on $T^*G \times \{0\}$. Therefore, the second stage reduced spaces or, equivalently, the $(G \circledS H)$-coadjoint orbits through points in $\mathfrak{g}^* \times \mathfrak{h}^*$ whose second component is zero, are symplectically diffeomorphic to the coadjoint orbits of G.

The next section is devoted to an example. Using the full power of the theory developed in the last two sections we compute the coadjoint orbits of a semidirect product of two concrete nonabelian groups.

10.7   Example: The Group $\mathcal{T} \circledS \mathcal{U}$

This section classifies the coadjoint orbits of the semidirect product of the upper triangular group with its normal unipotent subgroup, the group of upper triangular matrices with ones on the diagonal. This will be carried out in detail for 3×3 matrices. To classify the orbits, the full reduction by stages theory developed in the previous two sections will be used. It is interesting to work out this example in detail because it uses the full power of the reduction by stages theory and also has certain unusual properties: even though the generic coadjoint orbit has codimension three, there is only one global Casimir function. Worse, on an open dense Poisson submanifold containing the generic coadjoint orbits, there is a second Casimir globally defined on this open set but, as will be shown, there cannot be a third one, even on this open submanifold. This example also illustrates the complexity of the coadjoint orbits of a general semidirect product. In fact, recall from Proposition 10.6.8 that if $G^\nu = G$, then the coadjoint orbits of the semidirect product $G \circledS H$ are products of affine orbits of G (associated to a non-equivariance group one-cocycle) and coadjoint orbits of H. However, in the present example, $G^\nu \neq G$, and so this complexity also must be dealt with.

The Group $\mathcal{T} \circledS \mathcal{U}$.   In §10.5 we carried out the first stage reduction for the semidirect product Lie group $\mathcal{T} \circledS \mathcal{U}$. We now recall a few of the basic formulas for this example. First of all, recall that $\mathcal{T}$ denotes the group of $n \times n$ upper triangular matrices and that $\mathcal{U}$ denotes its normal subgroup consisting of all elements in $\mathcal{T}$ that have ones on the diagonal. The group $\mathcal{T}$ acts on $\mathcal{U}$ by conjugation: $\phi : \mathcal{T} \to \mathrm{Aut}(\mathcal{U})$ is defined by $\phi(t)(u) := tut^{-1}$ for any $t \in \mathcal{T}$ and $u \in \mathcal{U}$. Thus one can form the semidirect product $\mathcal{T} \circledS \mathcal{U}$ whose multiplication and inverse have the expressions

$$(t_1, u_1)(t_2, u_2) = (t_1 t_2, u_1 t_1 u_2 t_1^{-1}) \qquad \text{and} \qquad (t, u)^{-1} = (t^{-1}, t^{-1} u^{-1} t)$$

for all $t_1, t_2, t \in \mathcal{T}$, and $u_1, u_2, u \in \mathcal{U}$. Note that $\dim(\mathcal{T} \circledS \mathcal{U}) = n^2$.

The Lie algebra $\mathfrak{t}$ of $\mathcal{T}$ consists of all upper triangular matrices and the Lie algebra $\mathfrak{u}$ of $\mathcal{U}$ of all strictly upper triangular matrices; $\mathfrak{u}$ is an ideal in $\mathfrak{t}$. The Lie bracket in $\mathfrak{t} \circledS \mathfrak{u}$ is given by

$$[(\xi_1, \eta_1), (\xi_2, \eta_2)] = ([\xi_1, \xi_2], [\xi_1, \eta_2] - [\xi_2, \eta_1] + [\eta_1, \eta_2])$$

for any $\xi_1, \xi_2 \in \mathfrak{t}$ and $\eta_1, \eta_2 \in \mathfrak{u}$.

The nondegenerate pairing $\langle A, B \rangle := \mathrm{trace}(AB^T)$ for $A, B \in \mathfrak{gl}(n, \mathbb{R})$ identifies $\mathfrak{gl}(n, \mathbb{R})$ with its dual $\mathfrak{gl}(n, \mathbb{R})^*$ and restricts to nondegenerate pairings on $\mathfrak{t}$ and on $\mathfrak{u}$. Thus, we identify $\mathfrak{t}^*$ with $\mathfrak{t}$ and $\mathfrak{u}^*$ with $\mathfrak{u}$. The coadjoint action of $\mathcal{U}$ on $\mathfrak{u}^*$ is given by (10.5.8), that is,

$$\mathrm{Ad}^*_{u^{-1}} \nu = \Pi_\mathfrak{u}\left((u^{-1})^T \nu u^T\right), \quad \text{for} \quad u \in \mathcal{U}, \quad \nu \in \mathfrak{u}^* = \mathfrak{u},$$

where $\Pi_{\mathfrak{u}} : \mathfrak{gl}(n) \to \mathfrak{u}$ is the projection that maps a matrix to its strictly upper triangular part.

Notice that for $n = 3$, if

$$
u = \begin{bmatrix} 1 & u_{12} & u_{13} \\ 0 & 1 & u_{23} \\ 0 & 0 & 1 \end{bmatrix}
$$

then

$$
u^{-1} = \begin{bmatrix} 1 & -u_{12} & u_{12}u_{23} - u_{13} \\ 0 & 1 & -u_{23} \\ 0 & 0 & 1 \end{bmatrix}.
$$

Thus, for

$$
\nu = \begin{bmatrix} 0 & \nu_{12} & \nu_{13} \\ 0 & 0 & \nu_{23} \\ 0 & 0 & 0 \end{bmatrix} \in \mathfrak{u}
$$

we get by (10.5.9)

$$
\mathrm{Ad}^*_{u^{-1}} \nu = \begin{bmatrix} 0 & \nu_{12} + u_{23}\nu_{13} & \nu_{13} \\ 0 & 0 & \nu_{23} - u_{12}\nu_{13} \\ 0 & 0 & 0 \end{bmatrix}. \tag{10.7.1}
$$

We have also seen in §10.5 that for $\nu_{13} \neq 0$ the isotropy subgroup $\mathcal{U}_\nu$ is one-dimensional and the coadjoint orbit $\mathcal{O}_\nu$ is two-dimensional, whereas for $\nu_{13} = 0$, we have $\mathcal{U}_\nu = \mathcal{U}$ and $\mathcal{O}_\nu = \{\nu\}$.

The action of $\mathcal{T}$ on $\mathfrak{u}^*$ is given by (10.6.7), that is,

$$
\tilde{\phi}(t^{-1})^* \nu = \Pi_{\mathfrak{u}}\left((t^{-1})^T \nu t^T\right),
$$

for $\nu \in \mathfrak{u}^*$ and $t \in \mathcal{T}$. Similarly, the coadjoint action of $\mathcal{T}$ on $\mathfrak{t}^*$ has the expression

$$
\mathrm{Ad}^*_{t^{-1}} \mu = \Pi_{\mathfrak{t}}\left((t^{-1})^T \mu t^T\right),
$$

where $t \in \mathcal{T}$, $\mu \in \mathfrak{t}^*$, and $\Pi_{\mathfrak{t}} : \mathfrak{gl}(n, \mathbb{R}) \to \mathfrak{t}$ is the projection that maps a matrix to its upper triangular part.

If $n = 3$ we have

$$
\begin{bmatrix} t_{11} & t_{12} & t_{13} \\ 0 & t_{22} & t_{23} \\ 0 & 0 & t_{33} \end{bmatrix}^{-1} = \begin{bmatrix} \frac{1}{t_{11}} & -\frac{t_{12}}{t_{11}t_{22}} & \frac{t_{12}t_{23}}{t_{11}t_{22}t_{33}} - \frac{t_{13}}{t_{11}t_{33}} \\ 0 & \frac{1}{t_{22}} & -\frac{t_{23}}{t_{22}t_{33}} \\ 0 & 0 & \frac{1}{t_{33}} \end{bmatrix}. \tag{10.7.2}
$$

Thus, if

$$
\mu = \begin{bmatrix} \mu_{11} & \nu_{12} & \nu_{13} \\ 0 & \mu_{22} & \nu_{23} \\ 0 & 0 & \mu_{33} \end{bmatrix} \in \mathfrak{t}^* = \mathfrak{t}
$$

the non-zero entries of $\Pi_{\mathfrak{t}} \left((t^{-1})^T \mu t^T \right)$ are

$$
\left.\begin{array}{l}
[\Pi_{\mathfrak{t}} \left((t^{-1})^T \mu t^T \right)]_{11} = \mu_{11} + \dfrac{t_{12}}{t_{11}} \mu_{12} + \dfrac{t_{13}}{t_{11}} \mu_{13} \\[2mm]
[\Pi_{\mathfrak{t}} \left((t^{-1})^T \mu t^T \right)]_{22} = \mu_{22} - \dfrac{t_{12}}{t_{11}} \mu_{12} - \dfrac{t_{12}t_{23}}{t_{11}t_{22}} \mu_{13} + \dfrac{t_{23}}{t_{22}} \mu_{23} \\[2mm]
[\Pi_{\mathfrak{t}} \left((t^{-1})^T \mu t^T \right)]_{33} = \mu_{33} + \dfrac{t_{12}t_{23}}{t_{11}t_{22}} \mu_{13} - \dfrac{t_{13}}{t_{11}} \mu_{13} - \dfrac{t_{23}}{t_{22}} \mu_{23} \\[2mm]
[\Pi_{\mathfrak{t}} \left((t^{-1})^T \mu t^T \right)]_{12} = \dfrac{t_{22}}{t_{11}} \mu_{12} + \dfrac{t_{23}}{t_{11}} \mu_{13} \\[2mm]
[\Pi_{\mathfrak{t}} \left((t^{-1})^T \mu t^T \right)]_{23} = -\dfrac{t_{12}t_{33}}{t_{11}t_{22}} \mu_{13} + \dfrac{t_{33}}{t_{22}} \mu_{23} \\[2mm]
[\Pi_{\mathfrak{t}} \left((t^{-1})^T \mu t^T \right)]_{13} = \dfrac{t_{33}}{t_{11}} \mu_{13}.
\end{array}\right\} \quad (10.7.3)
$$

In particular, if

$$
\nu = \begin{bmatrix} 0 & \nu_{12} & \nu_{13} \\ 0 & 0 & \nu_{23} \\ 0 & 0 & 0 \end{bmatrix} \in \mathfrak{u}^* = \mathfrak{u}
$$

we get formula (10.6.8), that is,

$$
\Pi_{\mathfrak{u}} \left((t^{-1})^T \nu t^T \right) = \begin{bmatrix} 0 & \frac{t_{22}}{t_{11}} \nu_{12} + \frac{t_{23}}{t_{11}} \nu_{13} & \frac{t_{33}}{t_{11}} \nu_{13} \\[2mm] 0 & 0 & -\frac{t_{12}t_{33}}{t_{11}t_{22}} \nu_{13} + \frac{t_{33}}{t_{22}} \nu_{23} \\[2mm] 0 & 0 & 0 \end{bmatrix}. \quad (10.7.4)
$$

In this section we shall also need the following formulas. Namely, the non-zero entries of $\Pi_{\mathfrak{t}} \left(t^T \mu (t^{-1})^T \right)$ are

$$
\left.\begin{array}{l}
\left[\Pi_{\mathfrak{t}} \left(t^T \mu (t^{-1})^T \right)\right]_{11} = \mu_{11} - \dfrac{t_{12}}{t_{22}} \mu_{12} + \dfrac{t_{12}t_{23}}{t_{22}t_{33}} \mu_{13} - \dfrac{t_{13}}{t_{33}} \mu_{13} \\[2mm]
\left[\Pi_{\mathfrak{t}} \left(t^T \mu (t^{-1})^T \right)\right]_{22} = \mu_{22} + \dfrac{t_{12}}{t_{22}} \mu_{12} - \dfrac{t_{12}t_{23}}{t_{22}t_{33}} \mu_{13} - \dfrac{t_{23}}{t_{33}} \mu_{23} \\[2mm]
\left[\Pi_{\mathfrak{t}} \left(t^T \mu (t^{-1})^T \right)\right]_{33} = \mu_{33} + \dfrac{t_{23}}{t_{33}} \mu_{23} + \dfrac{t_{13}}{t_{33}} \mu_{13} \\[2mm]
\left[\Pi_{\mathfrak{t}} \left(t^T \mu (t^{-1})^T \right)\right]_{12} = \dfrac{t_{11}}{t_{22}} \mu_{12} - \dfrac{t_{11}t_{23}}{t_{22}t_{33}} \mu_{13} \\[2mm]
\left[\Pi_{\mathfrak{t}} \left(t^T \mu (t^{-1})^T \right)\right]_{23} = \dfrac{t_{22}}{t_{33}} \mu_{23} + \dfrac{t_{12}}{t_{33}} \mu_{13} \\[2mm]
\left[\Pi_{\mathfrak{t}} \left(t^T \mu (t^{-1})^T \right)\right]_{13} = \dfrac{t_{11}}{t_{33}} \mu_{13}
\end{array}\right\} \quad (10.7.5)
$$

In §10.5 we have shown that the first stage reduced Poisson manifold of $T^*(\mathcal{T} \circledS \mathcal{U})$ is $T^*\mathcal{T} \times \mathfrak{u}^*$ endowed with the product Poisson structure. If $\nu \in \mathfrak{u}$, the first stage symplectic reduced space at ν is

$$
(T^*\mathcal{T} \times \mathcal{O}_\nu, \omega_{\mathcal{T}} \oplus \omega_{\mathcal{O}_\nu}^+),
$$

where $\mathcal{O}_\nu$ is the coadjoint $\mathcal{U}$-orbit through ν. Thus, by Theorem 10.6.6, the second stage reduction gives the coadjoint orbits of $\mathcal{T} \circledS \mathcal{U}$ in $(\mathfrak{t} \circledS \mathfrak{u})^*$.

The Coadjoint Action.   One could attempt to compute the coadjoint orbits directly. We shall do this calculation now and will see that it does not help much with the classification problem. We shall also point out that the Lie-Poisson space $(\mathfrak{t} \circledS \mathfrak{u})^*$ has a very complicated structure due to the lack of Casimir functions. This then leads one to classify the coadjoint orbits using the reduction by stages program.

To compute the coadjoint orbits of $\mathcal{T} \circledS \mathcal{U}$ directly we use formula (6.4.3) which in this case becomes

$$
\begin{aligned}
\mathrm{Ad}^*_{(t,u)^{-1}} & (\mu, \nu) \\
& = \Big(\mathrm{Ad}^*_{t^{-1}} \mu + \big(\bar{\phi} \circ \mathrm{Ad}_{t^{-1}} \big)^{\#} (t^{-1} \cdot u)^* \big(T^*_{t^{-1} \cdot u} L_{t^{-1} \cdot u^{-1}} \nu \big) , \\
& \qquad\qquad\qquad\qquad \tilde{\phi}(t^{-1})^* \, \mathrm{Ad}^*_{t^{-1} \cdot u^{-1}} \nu \Big), \quad (10.7.6)
\end{aligned}
$$

where $\mu \in \mathfrak{t}^*$, $\nu \in \mathfrak{u}^*$, $t \in \mathcal{T}$, and $u \in \mathcal{U}$. Recall that given a Lie algebra anti-homomorphism $F : \mathfrak{t} \to \mathfrak{X}(\mathcal{U})$ and $u \in \mathcal{U}$, we let

$$
F^{\#}(u) : \mathfrak{t} \to T_u \mathcal{U}
$$

be the linear map given by

$$
F^{\#}(u)(\xi) := F(\xi)(u),
$$

for any $\xi \in \mathfrak{t}$. Then $F^{\#}(u)^* : T^*_u \mathcal{U} \to \mathfrak{t}^*$ denotes its dual map. We shall now explicitly compute this coadjoint action.

We determine first the term

$$
\big(\bar{\phi} \circ \mathrm{Ad}_{t^{-1}} \big)^{\#} (t^{-1} \cdot u)^* \big(T^*_{t^{-1} \cdot u} L_{t^{-1} \cdot u^{-1}} \nu \big).
$$

To do this, recall that

$$
t \cdot u := \phi(t)(u) = tut^{-1}
$$

for any $t \in \mathcal{T}$, $u \in \mathcal{U}$ and hence if $\zeta \in \mathfrak{t}$ and $v \in \mathcal{U}$, we get

$$
\begin{aligned}
T_v L_{v^{-1}} \bar{\phi}(\zeta)(v) :&= \left. \frac{d}{dt} \right|_{t=0} L_{v^{-1}} \left(\exp(t\zeta) \cdot v \right) = \left. \frac{d}{dt} \right|_{t=0} v^{-1} e^{t\zeta} v e^{-t\zeta} \\
& = v^{-1} \zeta v - \zeta. \quad\quad\quad\quad\quad\quad (10.7.7)
\end{aligned}
$$

We have for any $\xi \in \mathfrak{t}$

$$
\begin{aligned}
\Big\langle \big(\bar{\phi} \circ \mathrm{Ad}_{t^{-1}} \big)^{\#} & (t^{-1} \cdot u)^* \big(T^*_{t^{-1} \cdot u} L_{t^{-1} \cdot u^{-1}} \nu \big), \, \xi \Big\rangle \\
& = \Big\langle \nu, \, T_{t^{-1} \cdot u} L_{t^{-1} \cdot u^{-1}} \big(\bar{\phi} \circ \mathrm{Ad}_{t^{-1}} \big)^{\#} (t^{-1} \cdot u)(\xi) \Big\rangle \\
& = \Big\langle \nu, \, T_{t^{-1} \cdot u} L_{t^{-1} \cdot u^{-1}} \bar{\phi} \left(\mathrm{Ad}_{t^{-1}} \xi \right) (t^{-1} \cdot u) \Big\rangle.
\end{aligned}
$$

Applying equation (10.7.7) for $v := t^{-1} \cdot u = t^{-1} u t$ and $\zeta := \mathrm{Ad}_{t^{-1}} \xi = t^{-1} \xi t$ we get

$$\left\langle \nu, \, (t^{-1} u t)^{-1} (t^{-1} \xi t) (t^{-1} u t) - t^{-1} \xi t \right\rangle$$
$$= \left\langle \nu, \, t^{-1} u^{-1} \xi u t - t^{-1} \xi t \right\rangle$$
$$= \mathrm{trace} \left(\nu \left(t^{-1} u^{-1} \xi u t \right)^T - \nu \left(t^{-1} \xi t \right)^T \right)$$
$$= \mathrm{trace} \left(\left((u^{-1})^T (t^{-1})^T \nu t^T u^T - (t^{-1})^T \nu t^T \right) \xi^T \right)$$
$$= \mathrm{trace} \left(\Pi_{\mathfrak{t}} \left((u^{-1})^T (t^{-1})^T \nu t^T u^T - (t^{-1})^T \nu t^T \right) \xi^T \right),$$

where $\Pi_{\mathfrak{t}}$ denotes the projection of a matrix onto its upper triangular part. Therefore we get

$$\left\langle \left(\bar{\phi} \circ \mathrm{Ad}_{t^{-1}} \right)^{\#} (t^{-1} \cdot u)^* \left(T^*_{t^{-1} \cdot u} L_{t^{-1} \cdot u^{-1}} \nu \right), \, \xi \right\rangle$$
$$= \left\langle \Pi_{\mathfrak{t}} \left((u^{-1})^T (t^{-1})^T \nu t^T u^T - (t^{-1})^T \nu t^T \right), \, \xi \right\rangle,$$

for any $\xi \in \mathfrak{t}$ and hence

$$\left(\bar{\phi} \circ \mathrm{Ad}_{t^{-1}} \right)^{\#} (t^{-1} \cdot u)^* \left(T^*_{t^{-1} \cdot u} L_{t^{-1} \cdot u^{-1}} \nu \right)$$
$$= \Pi_{\mathfrak{t}} \left((u^{-1})^T (t^{-1})^T \nu t^T u^T - (t^{-1})^T \nu t^T \right). \qquad (10.7.8)$$

Since

$$\mathrm{Ad}^*_{t^{-1}} \mu = \Pi_{\mathfrak{t}} \left((t^{-1})^T \mu t \right)$$

for any $t \in \mathcal{T}$ and $\mu \in \mathfrak{t}^* = \mathfrak{t}$, the first component of (10.7.6) is

$$\Pi_{\mathfrak{t}} \left((t^{-1})^T \mu t + (u^{-1})^T (t^{-1})^T \nu t^T u^T - (t^{-1})^T \nu t^T \right). \qquad (10.7.9)$$

The second component of (10.7.6) is $\tilde{\phi}(t^{-1})^* \mathrm{Ad}^*_{t^{-1} \cdot u^{-1}} \nu$. Since

$$\tilde{\phi}(t^{-1})^* \lambda = \Pi_{\mathfrak{u}} \left((t^{-1})^T \lambda t^T \right)$$

and

$$\mathrm{Ad}^*_{v^{-1}} \nu = \Pi_{\mathfrak{u}} \left((v^{-1})^T \nu v^T \right),$$

for any $t \in \mathcal{T}$, $u, v \in \mathcal{U}$, and $\lambda, \nu \in \mathfrak{u}^*$, where $\Pi_{\mathfrak{u}}$ is the projection sending any matrix to its strictly upper triangular part, we get

$$\tilde{\phi}(t^{-1})^* \mathrm{Ad}^*_{t^{-1} \cdot u^{-1}} \nu$$
$$= \Pi_{\mathfrak{u}} \left((t^{-1})^T \Pi_{\mathfrak{u}} \left[((t^{-1} u t)^{-1})^T \nu (t^{-1} u t)^T \right] t^T \right)$$
$$= \Pi_{\mathfrak{u}} \left((t^{-1})^T \Pi_{\mathfrak{u}} \left[t^T (u^{-1})^T (t^{-1})^T \nu t^T u^T (t^{-1})^T \right] t^T \right).$$

Since

$$\Pi_{\mathfrak{u}} \left((t^{-1})^T \left(\Pi_{\mathfrak{u}} A \right) t^T \right) = \Pi_{\mathfrak{u}} \left((t^{-1})^T A t^T \right)$$

for any matrix A and any $t \in \mathcal{T}$, the above expression becomes

$$\tilde{\phi}(t^{-1})^* \operatorname{Ad}^*_{t^{-1} \cdot u^{-1}} \nu = \Pi_{\mathfrak{u}} \left((u^{-1})^T (t^{-1})^T \nu t^T u^T \right). \qquad (10.7.10)$$

Therefore, by (10.7.9) and (10.7.10), the coadjoint action of $\mathcal{T} \circledS \mathcal{U}$ on $(\mathfrak{t} \circledS \mathfrak{u})^*$ has the expression

$$\begin{aligned}
\operatorname{Ad}^*_{(t,u)^{-1}} (\mu, \nu) \\
= \Big(\Pi_{\mathfrak{t}} \big((t^{-1})^T \mu t + (u^{-1})^T (t^{-1})^T \nu t^T u^T - (t^{-1})^T \nu t^T \big), \\
\Pi_{\mathfrak{u}} \big((u^{-1})^T (t^{-1})^T \nu t^T u^T \big) \Big). \qquad (10.7.11)
\end{aligned}$$

Note from this formula that if $\nu = 0$ then the $\mathcal{T} \circledS \mathcal{U}$-coadjoint orbit $\mathcal{O}_{(\mu,0)}$ equals $\mathcal{O}_\mu \times \{0\}$, in agreement with the comment following Proposition 10.6.8.

The Coadjoint Orbits. A direct attack on the classification of the coadjoint orbits is relatively involved. If $n = 3$, due to the small size of the matrices, it is still possible to do this by hand. As will be seen below, the classification of the isotropy subgroups by a direct approach quickly degenerates into very long and complicated computations. Later on, we apply the reduction by stages procedure to retrieve the same results with considerably less computations and in the process learn much more about the structure of the orbits. Thus, the reduction by stages method yields automatically more information about the internal structure of the coadjoint orbits and, at the same time, is also computational useful.

If $n = 3$ and $(\overline{\mu}, \overline{\nu}) := \text{Ad}^*_{(t,u)^{-1}}(\mu, \nu)$, formula (10.7.11) becomes

$$
\left.
\begin{aligned}
\overline{\mu}_{11} &= \mu_{11} + \frac{t_{12}}{t_{11}}\mu_{12} + \frac{t_{13}}{t_{11}}\mu_{13} + u_{12}\frac{t_{22}}{t_{11}}\nu_{12} \\
&\quad + u_{12}\frac{t_{23}}{t_{11}}\nu_{13} + u_{13}\frac{t_{33}}{t_{11}}\nu_{13} \\
\overline{\mu}_{22} &= \mu_{22} - \frac{t_{12}}{t_{11}}\mu_{12} + \frac{t_{23}}{t_{22}}\mu_{23} - \frac{t_{12}t_{23}}{t_{11}t_{22}}\mu_{13} \\
&\quad - u_{12}\frac{t_{22}}{t_{11}}\nu_{12} + u_{23}\frac{t_{33}}{t_{22}}\nu_{23} - u_{23}\frac{t_{12}t_{33}}{t_{11}t_{22}}\nu_{13} \\
&\quad - u_{12}\frac{t_{23}}{t_{11}}\nu_{13} - u_{12}u_{23}\frac{t_{33}}{t_{11}}\nu_{13} \\
\overline{\mu}_{33} &= \mu_{33} - \frac{t_{23}}{t_{22}}\mu_{23} - \frac{t_{13}}{t_{11}}\mu_{13} + \frac{t_{12}t_{23}}{t_{11}t_{22}}\mu_{13} - u_{23}\frac{t_{33}}{t_{22}}\nu_{23} \\
&\quad + u_{23}\frac{t_{12}t_{33}}{t_{11}t_{22}}\nu_{13} - u_{13}\frac{t_{33}}{t_{11}}\nu_{13} + u_{12}u_{23}\frac{t_{33}}{t_{11}}\nu_{13} \\
\overline{\mu}_{12} &= \frac{t_{22}}{t_{11}}\mu_{12} + \frac{t_{23}}{t_{11}}\mu_{13} + u_{23}\frac{t_{33}}{t_{11}}\nu_{13} \\
\overline{\mu}_{23} &= \frac{t_{33}}{t_{22}}\mu_{23} - \frac{t_{12}t_{33}}{t_{11}t_{22}}\mu_{13} - u_{12}\frac{t_{33}}{t_{11}}\nu_{13} \\
\overline{\mu}_{13} &= \frac{t_{33}}{t_{11}}\mu_{13} \\
\overline{\nu}_{12} &= \frac{t_{22}}{t_{11}}\nu_{12} + \frac{t_{23}}{t_{11}}\nu_{13} + u_{23}\frac{t_{33}}{t_{11}}\nu_{13} \\
\overline{\nu}_{23} &= \frac{t_{33}}{t_{22}}\nu_{23} - \frac{t_{12}t_{33}}{t_{11}t_{22}}\nu_{13} - u_{12}\frac{t_{33}}{t_{11}}\nu_{13} \\
\overline{\nu}_{13} &= \frac{t_{33}}{t_{11}}\nu_{13}.
\end{aligned}
\right\} \qquad (10.7.12)
$$

These formulas are the result of a lengthy but straightforward computation using (10.7.11), (10.7.3), (10.7.4), and the expression of the product

$$
ut = \begin{bmatrix} t_{11} & t_{12} + u_{12}t_{22} & t_{13} + u_{12}t_{23} + u_{13}t_{33} \\ 0 & t_{22} & t_{23} + u_{23}t_{33} \\ 0 & 0 & t_{33} \end{bmatrix}.
$$

Even though the description of the coadjoint orbit through (μ, ν) by these formulas is complete, we learn very little about its structure. There are only two almost obvious observations:

$$
\overline{\mu}_{11} + \overline{\mu}_{22} + \overline{\mu}_{33} = \mu_{11} + \mu_{33} + \mu_{33} \qquad \text{and} \qquad \frac{\overline{\mu}_{13}}{\overline{\nu}_{13}} = \frac{\mu_{13}}{\nu_{13}}.
$$

The second relation holds, of course, only on the open subset where $\nu_{13} \neq 0$; this subset is invariant under the coadjoint action as the last equation in

(10.7.12) shows. We shall return to these formulas later when discussing the Casimir functions of $(\mathfrak{t} \circledS \mathfrak{u})^*$.

We begin with the calculation of the coadjoint isotropy subgroups. By (10.7.11), $(t, u) \in (\mathcal{T} \circledS \mathcal{U})_{(\mu, \nu)}$ if and only if

$$\mu = \Pi_\mathfrak{t} \left((t^{-1})^T \mu t + (u^{-1})^T (t^{-1})^T \nu t^T u^T - (t^{-1})^T \nu t^T \right)$$
$$\nu = \Pi_\mathfrak{u} \left((u^{-1})^T (t^{-1})^T \nu t^T u^T \right).$$

If $n = 3$, by (10.7.12) this is equivalent to

$$\mu_{11} = \mu_{11} + \frac{t_{12}}{t_{11}} \mu_{12} + \frac{t_{13}}{t_{11}} \mu_{13} + u_{12} \frac{t_{22}}{t_{11}} \nu_{12} + \frac{u_{12} t_{23} + u_{13} t_{33}}{t_{11}} \nu_{13}$$

$$\mu_{22} = \mu_{22} - \frac{t_{12}}{t_{11}} \mu_{12} - \frac{t_{12} t_{23}}{t_{11} t_{22}} \mu_{13} + \frac{t_{23}}{t_{22}} \mu_{23} - u_{12} \frac{t_{22}}{t_{11}} \nu_{12}$$
$$\qquad - \frac{u_{12} t_{22} t_{23} + u_{23} t_{12} t_{33} + u_{12} u_{23} t_{22} t_{33}}{t_{11} t_{22}} \nu_{13} + u_{23} \frac{t_{33}}{t_{22}} \nu_{23}$$

$$\mu_{33} = \mu_{33} + \frac{t_{12} t_{23}}{t_{11} t_{22}} \mu_{13} - \frac{t_{13}}{t_{11}} \mu_{13} - \frac{t_{23}}{t_{22}} \mu_{23}$$
$$\qquad + \frac{u_{12} t_{22} t_{23} + u_{23} t_{12} t_{33} + u_{12} u_{23} t_{22} t_{33}}{t_{11} t_{22}} \nu_{13}$$
$$\qquad - \frac{u_{12} t_{23} + u_{13} t_{33}}{t_{11}} \nu_{13} - u_{23} \frac{t_{33}}{t_{22}} \nu_{23}$$

$$\mu_{12} = \frac{t_{22}}{t_{11}} \mu_{12} + \frac{t_{23}}{t_{11}} \mu_{13} + u_{23} \frac{t_{33}}{t_{11}} \nu_{13}$$

$$\mu_{23} = -\frac{t_{12} t_{33}}{t_{11} t_{22}} \mu_{13} + \frac{t_{33}}{t_{22}} \mu_{23} - u_{12} \frac{t_{33}}{t_{11}} \nu_{13}$$

$$\mu_{13} = \frac{t_{33}}{t_{11}} \mu_{13}$$

$$\nu_{12} = \frac{t_{22}}{t_{11}} \nu_{12} + \frac{t_{23} + u_{23} t_{33}}{t_{11}} \nu_{13}$$

$$\nu_{23} = -\frac{(t_{12} + u_{12} t_{22}) t_{33}}{t_{11} t_{22}} \nu_{13} + \frac{t_{33}}{t_{22}} \nu_{23}$$

$$\nu_{13} = \frac{t_{33}}{t_{11}} \nu_{13}.$$

Case 1: If $\nu_{13} \neq 0$, then $t_{11} = t_{33}$ so that the sixth and ninth equations are satisfied. The seventh and eighth equations give

$$u_{12} = \frac{t_{11}(t_{33} - t_{22})}{t_{22} t_{33}} \frac{\nu_{23}}{\nu_{13}} - \frac{t_{12}}{t_{22}} \qquad \text{and} \qquad u_{23} = \frac{t_{11} - t_{22}}{t_{33}} \frac{\nu_{12}}{\nu_{13}} - \frac{t_{23}}{t_{33}}$$

and the first five equations are

$$
\left.\begin{aligned}
&t_{12}\mu_{12} + t_{13}\mu_{13} + u_{12}t_{22}\nu_{12} + (u_{12}t_{23} + u_{13}t_{33})\nu_{13} = 0 \\
&t_{12}t_{22}\mu_{12} + t_{12}t_{23}\mu_{13} - t_{11}t_{23}\mu_{23} + u_{12}t_{22}^2\nu_{12} \\
&\qquad + (u_{12}t_{22}t_{23} + u_{23}t_{12}t_{33} + u_{12}u_{23}t_{22}t_{33})\nu_{13} \\
&\qquad - u_{23}t_{11}t_{33}\nu_{23} = 0 \\
&t_{12}t_{23}\mu_{13} - t_{13}t_{22}\mu_{13} - t_{23}t_{11}\mu_{23} \\
&\qquad + t_{33}(u_{23}t_{12} + u_{12}u_{23}t_{22} - u_{13}t_{22})\nu_{13} - u_{23}t_{11}t_{33}\nu_{23} = 0 \\
&(t_{22} - t_{11})\mu_{12} + t_{23}\mu_{13} + u_{23}t_{33}\nu_{13} = 0 \\
&t_{11}(t_{33} - t_{22})\mu_{23} - t_{12}t_{33}\mu_{13} - u_{12}t_{22}t_{33}\nu_{13} = 0.
\end{aligned}\right\}
$$
$$(10.7.13)$$

Replacing in the last two equations the values of t_{33}, u_{12}, u_{23} found above gives

$$
\left.\begin{aligned}
(t_{22} - t_{11})(\mu_{12} - \nu_{12}) + t_{23}(\mu_{13} - \nu_{13}) = 0 \\
(t_{22} - t_{11})(\mu_{23} - \nu_{23}) + t_{12}(\mu_{13} - \nu_{13}) = 0.
\end{aligned}\right\}
$$
$$(10.7.14)$$

Case 1A: Thus, if in addition to $\nu_{13} \neq 0$ we assume that $\mu_{13} \neq \nu_{13}$, we get

$$
t_{12} = (t_{11} - t_{22})\frac{\mu_{23} - \nu_{23}}{\mu_{13} - \nu_{13}} \qquad \text{and} \qquad t_{23} = (t_{11} - t_{22})\frac{\mu_{12} - \nu_{12}}{\mu_{13} - \nu_{13}} \, .
$$

and hence

$$
u_{12} = \frac{t_{11} - t_{22}}{t_{22}}\left(\frac{\nu_{23}}{\nu_{13}} - \frac{\mu_{23} - \nu_{23}}{\mu_{13} - \nu_{13}}\right)
$$

$$
u_{23} = \frac{t_{11} - t_{22}}{t_{11}}\left(\frac{\nu_{12}}{\nu_{13}} - \frac{\mu_{12} - \nu_{12}}{\mu_{13} - \nu_{13}}\right)
$$

Plugging in these values in the first equation of (10.7.13) and solving for u_{13} yields

$$
\begin{aligned}
u_{13} = &-\frac{t_{11} - t_{22}}{t_{11}}\frac{\mu_{23} - \nu_{23}}{\mu_{13} - \nu_{13}}\frac{\mu_{12}}{\nu_{13}} - \frac{t_{13}}{t_{11}}\frac{\mu_{13}}{\nu_{13}} \\
&-\frac{t_{11} - t_{22}}{t_{11}}\left(\frac{\nu_{23}}{\nu_{13}} - \frac{\mu_{23} - \nu_{23}}{\mu_{13} - \nu_{13}}\right)\frac{\nu_{12}}{\nu_{13}} \\
&-\frac{(t_{11} - t_{22})^2}{t_{11}t_{22}}\left(\frac{\nu_{23}}{\nu_{13}} - \frac{\mu_{23} - \nu_{23}}{\mu_{13} - \nu_{13}}\right)\frac{\mu_{12} - \nu_{12}}{\mu_{13} - \nu_{13}}
\end{aligned}
$$

A long but straightforward computation shows that with these values for u_{12}, u_{23}, u_{13}, t_{12}, t_{23}, t_{33} the first two equations in (10.7.13) are identically satisfied. Therefore, the coadjoint isotropy subgroup at (μ, ν) with $\nu_{13} \neq 0, \mu_{13} \neq \nu_{13}$, consists of pairs $(t, u) \in \mathcal{T} \circledS \mathcal{U}$ such that $t_{11} = t_{33}$ and $u_{12}, u_{23}, u_{13}, t_{12}, t_{23}$ are given above. Thus $\dim\left((\mathcal{T} \circledS \mathcal{U})_{(\mu,\nu)}\right) = 3$ and hence the dimension of the generic coadjoint orbit is $9 - 3 = 6$, where *generic* means that $\nu_{13} \neq 0$ and $\mu_{13} \neq \nu_{13}$.

Case 1B: We continue with the classification. So let $\nu_{13} \neq 0$ and $\mu_{13} = \nu_{13}$. Then (10.7.14) becomes $(t_{22}-t_{11})(\mu_{12}-\nu_{12}) = 0$ and $(t_{22}-t_{11})(\mu_{23}-\nu_{23}) = 0$. Thus if at least one of $\mu_{12}-\nu_{12}$ or $\mu_{23}-\nu_{23}$ does not vanish, then $t_{11} = t_{22}$ and hence $t_{11} = t_{22} = t_{33}$. The system (10.7.13) simplifies to

$$
\left.
\begin{aligned}
& t_{12}\mu_{12} + t_{13}\nu_{13} + u_{12}t_{11}\nu_{12} + (u_{12}t_{23} + u_{13}t_{11})\nu_{13} = 0 \\
& t_{12}t_{11}\mu_{12} + t_{12}t_{23}\nu_{13} - t_{11}t_{23}\mu_{23} + u_{12}t_{11}^2\nu_{12} \\
& \quad + (u_{12}t_{11}t_{23} + u_{23}t_{12}t_{11} + u_{12}u_{23}t_{11}^2)\nu_{13} \\
& \quad - u_{23}t_{11}^2\nu_{23} = 0 \\
& t_{12}t_{23}\nu_{13} - t_{13}t_{11}\nu_{13} - t_{23}t_{11}\mu_{23} \\
& \quad + t_{11}(u_{23}t_{12} + u_{12}u_{23}t_{11} - u_{13}t_{11})\nu_{13} - u_{23}t_{11}^2\nu_{23} = 0 \\
& t_{23}\nu_{13} + u_{23}t_{11}\nu_{13} = 0 \\
& t_{12}\nu_{13} + u_{12}t_{11}\nu_{13} = 0.
\end{aligned}
\right\} \quad (10.7.15)
$$

From the last two equations we obtain

$$
u_{12} = -\frac{t_{12}}{t_{11}} \qquad \text{and} \qquad u_{23} = -\frac{t_{23}}{t_{11}}
$$

and so the first one gives

$$
u_{13} = -\frac{t_{13}}{t_{11}} - \frac{t_{12}}{t_{11}}\frac{\mu_{12}}{\nu_{13}} + \frac{t_{12}t_{23}}{t_{11}^2} + \frac{t_{12}}{t_{11}}\frac{\nu_{12}}{\nu_{13}}.
$$

With these values for u_{12}, u_{23}, u_{13} the second and third equations in (10.7.15) are identical and give

$$
t_{12}(\mu_{12} - \nu_{12}) = t_{23}(\mu_{23} - \nu_{23})
$$

which can be solved for either t_{12} or t_{23} since at least one of $\mu_{12}-\nu_{12}$ or $\mu_{23}-\nu_{23}$ does not vanish. Thus, in this case, the values of $t_{22}, t_{33}, u_{12}, u_{23}, u_{13}$ and t_{12} (or t_{23}) are determined and hence $\dim \left(\mathcal{T} \circledS \mathcal{U} \right)_{(\mu,\nu)} = 3$. Hence, in this second case, $\dim \mathcal{O}_{(\mu,\nu)} = 9 - 3 = 6$, as before.

Case 1C: So let $\nu_{13} \neq 0$, $\mu_{13} = \nu_{13}$, $\mu_{12} = \nu_{12}$, and $\mu_{23} = \nu_{23}$. Then $t_{11} = t_{33}$ and (10.7.14) is identically satisfied. We get

$$
u_{12} = \frac{t_{11} - t_{22}}{t_{22}}\frac{\nu_{23}}{\nu_{13}} - \frac{t_{12}}{t_{22}} \qquad \text{and} \qquad u_{23} = \frac{t_{11} - t_{22}}{t_{11}}\frac{\nu_{12}}{\nu_{13}} - \frac{t_{23}}{t_{11}}
$$

and the system ($10.7.13$) simplifies to

$$
\left.
\begin{aligned}
&t_{12}\nu_{12} + t_{13}\nu_{13} + u_{12}t_{22}\nu_{12} + (u_{12}t_{23} + u_{13}t_{11})\nu_{13} = 0 \\
&t_{12}t_{22}\nu_{12} + t_{12}t_{23}\nu_{13} - t_{11}t_{23}\nu_{23} + u_{12}t_{22}^2\nu_{12} \\
&\qquad + (u_{12}t_{22}t_{23} + u_{23}t_{12}t_{11} + u_{12}u_{23}t_{22}t_{11})\nu_{13} \\
&\qquad - u_{23}t_{11}^2\nu_{23} = 0 \\
&t_{12}t_{23}\nu_{13} - t_{13}t_{22}\nu_{13} - t_{23}t_{11}\nu_{23} \\
&\qquad + t_{11}(u_{23}t_{12} + u_{12}u_{23}t_{22} - u_{13}t_{22})\nu_{13} - u_{23}t_{11}^2\nu_{23} = 0 \\
&(t_{22} - t_{11})\nu_{12} + t_{23}\nu_{13} + u_{23}t_{11}\nu_{13} = 0 \\
&(t_{11} - t_{22})\nu_{23} - t_{12}\nu_{13} - u_{12}t_{22}\nu_{13} = 0.
\end{aligned}
\right\}
\qquad (10.7.16)
$$

With these values for u_{12} and u_{23} the last two equations in ($10.7.16$) are identically satisfied and the first one gives

$$
u_{13} = \frac{t_{12}t_{23}}{t_{11}t_{22}} - \frac{t_{13}}{t_{11}} - \frac{t_{11} - t_{22}}{t_{11}} \frac{\nu_{23}}{\nu_{13}} \left(\frac{\nu_{12}}{\nu_{13}} + \frac{t_{23}}{t_{22}} \right).
$$

With the values just found for u_{12}, u_{23}, and u_{13}, the first two equations in ($10.7.16$) are identically satisfied.

Therefore, in this case, t_{33}, u_{12}, u_{23}, and u_{13} are determined so that $\dim \big((\mathcal{T} \circledS \mathcal{U})_{(\mu,\nu)} \big) = 5$ and thus $\dim \mathcal{O}_{(\mu,\nu)} = 9 - 5 = 4$.

Case 2: We return to the original conditions describing the isotropy subgroup and assume that $\nu_{13} = 0$. Then the conditions on (t, u) to be in the isotropy subgroup become

$$
\left.
\begin{aligned}
&t_{12}\mu_{12} + t_{13}\mu_{13} + u_{12}t_{22}\nu_{12} = 0 \\
&t_{12}t_{22}\mu_{12} + t_{12}t_{23}\mu_{13} - t_{23}t_{11}\mu_{23} + u_{12}t_{22}^2\nu_{12} - u_{23}t_{33}t_{11}\nu_{23} = 0 \\
&t_{12}t_{23}\mu_{13} - t_{13}t_{22}\mu_{13} - t_{23}t_{11}\mu_{23} - u_{23}t_{33}t_{11}\nu_{23} = 0 \\
&(t_{22} - t_{11})\mu_{12} + t_{23}\mu_{13} = 0 \\
&t_{11}(t_{33} - t_{22})\mu_{23} - t_{12}t_{33}\mu_{13} = 0 \\
&(t_{33} - t_{11})\mu_{13} = 0 \\
&(t_{22} - t_{11})\nu_{12} = 0 \\
&(t_{33} - t_{22})\nu_{23} = 0.
\end{aligned}
\right\}
$$

$$(10.7.17)$$

Case 2A: If $\nu_{13} = 0, \mu_{13} \neq 0$, then the sixth equation in (10.7.17) implies that $t_{33} = t_{11}$ and so the system (10.7.17) becomes

$$
\left.
\begin{aligned}
&t_{12}\mu_{12} + t_{13}\mu_{13} + u_{12}t_{22}\nu_{12} = 0 \\
&t_{12}t_{22}\mu_{12} + t_{12}t_{23}\mu_{13} - t_{23}t_{11}\mu_{23} + u_{12}t_{22}^2\nu_{12} - u_{23}t_{11}^2\nu_{23} = 0 \\
&t_{12}t_{23}\mu_{13} - t_{13}t_{22}\mu_{13} - t_{23}t_{11}\mu_{23} - u_{23}t_{11}^2\nu_{23} = 0 \\
&(t_{22} - t_{11})\mu_{12} + t_{23}\mu_{13} = 0 \\
&(t_{11} - t_{22})\mu_{23} - t_{12}\mu_{13} = 0 \\
&(t_{22} - t_{11})\nu_{12} = 0 \\
&(t_{11} - t_{22})\nu_{23} = 0.
\end{aligned}
\right\}
\qquad (10.7.18)
$$

From the fourth, fifth, and first equations in (10.7.18) we get

$$
\begin{aligned}
t_{12} &= (t_{11} - t_{22})\frac{\mu_{23}}{\mu_{13}} \\
t_{23} &= (t_{11} - t_{22})\frac{\mu_{12}}{\mu_{13}} \\
t_{13} &= -t_{12}\frac{\mu_{12}}{\mu_{13}} - u_{12}t_{22}\frac{\nu_{12}}{\mu_{13}} = -(t_{11} - t_{22})\frac{\mu_{12}\mu_{23}}{\mu_{13}^2} - u_{12}t_{22}\frac{\nu_{12}}{\mu_{13}}.
\end{aligned}
$$

With these values of t_{12}, t_{23}, t_{13}, the second and third equations of (10.7.18) are identical so we are left with

$$
\left.
\begin{aligned}
&u_{12}t_{22}^2\nu_{12} - u_{23}t_{11}^2\nu_{23} = 0 \\
&(t_{22} - t_{11})\nu_{12} = 0 \\
&(t_{11} - t_{22})\nu_{23} = 0.
\end{aligned}
\right\}
\qquad (10.7.19)
$$

Case 2Aα: If in addition to $\nu_{13} = 0, \mu_{13} \neq 0$, at least one of ν_{12} or ν_{23} does not vanish, then $t_{11} = t_{22} = t_{33}$, $t_{12} = 0$, $t_{23} = 0$, $t_{13} = -u_{12}t_{11}\nu_{12}/\mu_{13}$, and the first equation determines either u_{12} or u_{23}. If $\nu_{12} \neq 0$, then $u_{12} = u_{23}\nu_{23}/\nu_{12}$ and $t_{13} = -u_{23}t_{11}\nu_{23}/\mu_{13}$. If $\nu_{23} \neq 0$, then $u_{23} = u_{12}\nu_{12}/\nu_{23}$. Thus, in this case the values of $t_{22}, t_{33}, t_{12}, t_{23}, t_{13}$, and u_{12} (or u_{23}) are determined and hence $\dim\left((\mathcal{T}\circledS\mathcal{U})_{(\mu,\nu)}\right) = 3$. Therefore, $\dim\mathcal{O}_{(\mu,\nu)} = 9 - 3 = 6$.

Case 2Aβ: Assume now that $\nu_{13} = 0, \mu_{13} \neq 0$ and that $\nu_{12} = \nu_{23} = 0$. Then the system (10.7.19) consists of three equations that are identically satisfied. So t_{33}, t_{12}, t_{23}, and t_{13} are determined which means that $\dim\left((\mathcal{T}\circledS\mathcal{U})_{(\mu,\nu)}\right) = 5$. Therefore, in this case $\dim\mathcal{O}_{(\mu,\nu)} = 9 - 5 = 4$.

Case 2B: If $\nu_{13} = \mu_{13} = 0$ then the system (10.7.17) becomes

$$\left.\begin{array}{l} t_{12}\mu_{12} + u_{12}t_{22}\nu_{12} = 0 \\ t_{12}t_{22}\mu_{12} - t_{23}t_{11}\mu_{23} + u_{12}t_{22}^2\nu_{12} - u_{23}t_{33}t_{11}\nu_{23} = 0 \\ t_{23}\mu_{23} + u_{23}t_{33}\nu_{23} = 0 \\ (t_{22} - t_{11})\mu_{12} = 0 \\ (t_{33} - t_{22})\mu_{23} = 0 \\ (t_{22} - t_{11})\nu_{12} = 0 \\ (t_{33} - t_{22})\nu_{23} = 0. \end{array}\right\} \qquad (10.7.20)$$

Case 2Bα: Assume that $\nu_{13} = \mu_{13} = 0$ and $\nu_{12} \neq 0$. Then $t_{22} = t_{11}$ and the system (10.7.20) becomes

$$\left.\begin{array}{l} t_{12}\mu_{12} + u_{12}t_{11}\nu_{12} = 0 \\ t_{12}\mu_{12} - t_{23}\mu_{23} + u_{12}t_{11}\nu_{12} - u_{23}t_{33}\nu_{23} = 0 \\ t_{23}\mu_{23} + u_{23}t_{33}\nu_{23} = 0 \\ (t_{33} - t_{22})\mu_{23} = 0 \\ (t_{33} - t_{22})\nu_{23} = 0. \end{array}\right\} \qquad (10.7.21)$$

From the first equation we get

$$u_{12} = -\frac{t_{12}}{t_{11}}\frac{\mu_{12}}{\nu_{12}}$$

and the system (10.7.21) simplifies further to

$$\left.\begin{array}{l} t_{23}\mu_{23} + u_{23}t_{33}\nu_{23} = 0 \\ (t_{33} - t_{22})\mu_{23} = 0 \\ (t_{33} - t_{22})\nu_{23} = 0. \end{array}\right\} \qquad (10.7.22)$$

Case 2Bαa: Assume that $\nu_{13} = \mu_{13} = 0$, $\nu_{12} \neq 0$, and $\nu_{23} \neq 0$. Then $t_{33} = t_{22} = t_{11}$ and $u_{23} = -t_{23}\mu_{23}/t_{11}\nu_{23}$. Thus t_{22}, t_{33}, u_{12}, and u_{23} are determined which shows that $\dim\left((\mathcal{T} \circledS \mathcal{U})_{(\mu,\nu)}\right) = 5$ and thus $\dim \mathcal{O}_{(\mu,\nu)} = 9 - 5 = 4$.

Case 2Bαb: Assume that $\nu_{13} = \mu_{13} = 0$, $\nu_{12} \neq 0$, and $\nu_{23} = 0$. Then the system (10.7.22) simplifies to

$$\left.\begin{array}{l} t_{23}\mu_{23} = 0 \\ (t_{33} - t_{22})\mu_{23} = 0. \end{array}\right\} \qquad (10.7.23)$$

- If in addition $\mu_{23} \neq 0$, then $t_{23} = 0$ and $t_{33} = t_{22} = t_{11}$. Thus in this case, t_{22}, t_{33}, u_{12}, and t_{23} are determined, so $\dim\left((\mathcal{T} \circledS \mathcal{U})_{(\mu,\nu)}\right) = 5$ and hence $\dim \mathcal{O}_{(\mu,\nu)} = 9 - 5 = 4$.

- If in addition $\mu_{23} = 0$, then only t_{22} and u_{12} are determined, so $\dim\left((\mathcal{T}\,\text{\textcircled{s}}\,\mathcal{U})_{(\mu,\nu)}\right) = 7$ and hence $\dim\mathcal{O}_{(\mu,\nu)} = 9 - 7 = 2$.

Case 2Bβ: Assume that $\nu_{13} = \mu_{13} = 0$ and $\nu_{12} = 0$. Then the system (10.7.20) becomes

$$
\left.
\begin{aligned}
&t_{12}\mu_{12} = 0 \\
&t_{12}t_{22}\mu_{12} - t_{23}t_{11}\mu_{23} - u_{23}t_{33}t_{11}\nu_{23} = 0 \\
&t_{23}\mu_{23} + u_{23}t_{33}\nu_{23} = 0 \\
&(t_{22} - t_{11})\mu_{12} = 0 \\
&(t_{33} - t_{22})\mu_{23} = 0 \\
&(t_{33} - t_{22})\nu_{23} = 0.
\end{aligned}
\right\}
\qquad (10.7.24)
$$

Case 2Bβa: Assume that $\nu_{13} = \mu_{13} = 0$, $\nu_{12} = 0$, and $\nu_{23} \neq 0$. Then $t_{33} = t_{22}$ and the system (10.7.24) becomes

$$
\left.
\begin{aligned}
&t_{12}\mu_{12} = 0 \\
&t_{12}t_{22}\mu_{12} - t_{23}t_{11}\mu_{23} - u_{23}t_{22}t_{11}\nu_{23} = 0 \\
&t_{23}\mu_{23} + u_{23}t_{22}\nu_{23} = 0 \\
&(t_{22} - t_{11})\mu_{12} = 0.
\end{aligned}
\right\}
\qquad (10.7.25)
$$

From the third equation we deduce

$$
u_{23} = -\frac{t_{23}}{t_{22}}\frac{\mu_{23}}{\nu_{23}}
$$

which, when plugged into the second equation, gives $t_{12}t_{22}\mu_{12} = 0$. Since $t_{22} \neq 0$, the system (10.7.25) collapses to

$$
\left.
\begin{aligned}
&t_{12}\mu_{12} = 0 \\
&(t_{22} - t_{11})\mu_{12} = 0.
\end{aligned}
\right\}
$$

- If in addition $\mu_{12} \neq 0$, then $t_{12} = 0$ and $t_{33} = t_{22} = t_{11}$. Thus, in this case, t_{33}, t_{22}, t_{12}, and u_{23} are determined which implies that $\dim\left((\mathcal{T}\,\text{\textcircled{s}}\,\mathcal{U})_{(\mu,\nu)}\right) = 5$ and hence $\dim\mathcal{O}_{(\mu,\nu)} = 9 - 5 = 4$.

- If, in addition, $\mu_{12} = 0$, then only t_{33} and u_{23} are determined which implies that $\dim\left((\mathcal{T}\,\text{\textcircled{s}}\,\mathcal{U})_{(\mu,\nu)}\right) = 7$ and hence $\dim\mathcal{O}_{(\mu,\nu)} = 9 - 7 = 2$.

Case 2Bβb: Assume that $\nu_{13} = \mu_{13} = \nu_{12} = \nu_{23} = 0$. Then the system (10.7.24) becomes

$$
\left.
\begin{aligned}
&t_{12}\mu_{12} = 0 \\
&t_{12}t_{22}\mu_{12} - t_{23}t_{11}\mu_{23} = 0 \\
&t_{23}\mu_{23} = 0 \\
&(t_{22} - t_{11})\mu_{12} = 0 \\
&(t_{33} - t_{22})\mu_{23} = 0.
\end{aligned}
\right\}
\qquad (10.7.26)
$$

- If $\mu_{12} \neq 0$, then $t_{12} = 0$ and $t_{22} = t_{11}$. The system (10.7.26) collapses to

$$\left.\begin{aligned} t_{23}\mu_{23} &= 0 \\ (t_{33} - t_{22})\mu_{23} &= 0. \end{aligned}\right\}$$

  - So, if $\mu_{23} \neq 0$, then $t_{23} = 0$ and $t_{11} = t_{22} = t_{33}$ which says that t_{22}, t_{33}, t_{12}, and t_{23} are determined and hence $\dim\left((\mathcal{T}\,\circledS\,\mathcal{U})_{(\mu,\nu)}\right) = 5$. Thus $\dim\mathcal{O}_{(\mu,\nu)} = 9 - 5 = 4$.

  - If $\mu_{23} = 0$ then only t_{22} and t_{12} are determined so $\dim\left((\mathcal{T}\,\circledS\,\mathcal{U})_{(\mu,\nu)}\right) = 7$ and hence $\dim\mathcal{O}_{(\mu,\nu)} = 9 - 7 = 2$.

- If $\mu_{12} = 0$, system (10.7.26) becomes

$$\left.\begin{aligned} t_{23}\mu_{23} &= 0 \\ (t_{33} - t_{22})\mu_{23} &= 0. \end{aligned}\right\}$$

  - If $\mu_{23} \neq 0$, then $t_{23} = 0$ and $t_{22} = t_{33}$. Therefore $\dim\left((\mathcal{T}\,\circledS\,\mathcal{U})_{(\mu,\nu)}\right) = 7$ and hence $\dim\mathcal{O}_{(\mu,\nu)} = 9 - 7 = 2$.

  - If $\mu_{23} = 0$, then there are no determined group elements and hence $\dim\left((\mathcal{T}\,\circledS\,\mathcal{U})_{(\mu,\nu)}\right) = 9$. Thus $\dim\mathcal{O}_{(\mu,\nu)} = 0$.

This finishes the classification of the coadjoint orbits done by hand.

Summary of the direct classification.   The coadjoint orbits of $\mathcal{T}\,\circledS\,\mathcal{U}$ are

- 6 dimensional if:

  (i) $\nu_{13} \neq 0$ and $\mu_{13} \neq \nu_{13}$, or

  (ii) $\nu_{13} \neq 0$, $\mu_{13} = \nu_{13}$, and at least one of $\mu_{12} - \nu_{12}$ or $\mu_{23} - \nu_{23}$ does not vanish, or

  (iii) $\nu_{13} = 0$, $\mu_{13} \neq 0$, and at least one of ν_{12} or ν_{23} is not zero.

- 4 dimensional if:

  (i) $\nu_{13} \neq 0$, $\mu_{13} = \nu_{13}$, $\mu_{12} = \nu_{12}$, $\mu_{23} = \nu_{23}$, or

  (ii) $\nu_{13} = 0$, $\mu_{13} \neq 0$, $\nu_{12} = 0$, $\nu_{23} = 0$, or

  (iii) $\nu_{13} = 0$, $\mu_{13} = 0$, $\nu_{12} \neq 0$, $\nu_{23} \neq 0$, or

  (iv) $\nu_{13} = 0$, $\mu_{13} = 0$, $\nu_{12} \neq 0$, $\nu_{23} = 0$, $\mu_{23} \neq 0$, or

  (v) $\nu_{13} = 0$, $\mu_{13} = 0$, $\nu_{12} = 0$, $\nu_{23} \neq 0$, $\mu_{12} \neq 0$, or

  (vi) $\nu_{13} = \mu_{13} = \nu_{12} = \nu_{23} = 0$, $\mu_{12} \neq 0$, $\mu_{23} \neq 0$

- 2 dimensional if:

  (i) $\nu_{13} = 0$, $\mu_{13} = 0$, $\nu_{12} \neq 0$, $\nu_{23} = 0$, $\mu_{23} = 0$, or

(ii) $\nu_{13} = 0$, $\mu_{13} = 0$, $\nu_{12} = 0$, $\nu_{23} \neq 0$, $\mu_{12} = 0$, or

(iii) $\nu_{13} = \mu_{13} = \nu_{12} = \nu_{23} = 0$, $\mu_{12} \neq 0$, $\mu_{23} = 0$, or

(iv) $\nu_{13} = \mu_{13} = \nu_{12} = \nu_{23} = 0$, $\mu_{12} = 0$, $\mu_{23} \neq 0$

- 0 dimensional, if $\nu = 0$ and $\mu_{12} = \mu_{23} = \mu_{13} = 0$.

From this point on we shall never use the cases that have previously appeared and refer to the direct classification of the orbits only by the scheme introduced above.

Casimir Functions.   The $(+)$-Lie-Poisson bracket on $(\mathfrak{t}\,\circledS\,\mathfrak{u})^*$ is given for any $F, H \in C^\infty((\mathfrak{t}\,\circledS\,\mathfrak{u})^*)$ by (6.4.5), which in this case becomes

$$
\{F, H\}(\mu, \nu) = \left\langle \mu, \left[\frac{\delta F}{\delta \mu}, \frac{\delta H}{\delta \mu}\right]\right\rangle
$$
$$
+ \left\langle \nu, \left[\frac{\delta F}{\delta \mu}, \frac{\delta H}{\delta \nu}\right] + \left[\frac{\delta F}{\delta \nu}, \frac{\delta H}{\delta \mu}\right] + \left[\frac{\delta F}{\delta \nu}, \frac{\delta H}{\delta \nu}\right]\right\rangle .
$$
$$(10.7.27)$$

Therefore, Hamilton's equations associated to H have the expression

$$
\left.\begin{aligned}
\dot{\mu} &= \Pi_\mathfrak{t}\left[\mu, \left(\frac{\delta H}{\delta \mu}\right)^T\right] + \Pi_\mathfrak{t}\left[\nu, \left(\frac{\delta H}{\delta \nu}\right)^T\right] \\
\dot{\nu} &= \Pi_\mathfrak{u}\left[\nu, \left(\frac{\delta H}{\delta \mu}\right)^T\right] + \Pi_\mathfrak{u}\left[\nu, \left(\frac{\delta H}{\delta \nu}\right)^T\right] .
\end{aligned}\right\}
$$
$$(10.7.28)$$

In coordinates and for the case $n = 3$, these become

$$
\left.
\begin{aligned}
\dot{\mu}_{11} &= \mu_{12}\frac{\partial H}{\partial \mu_{12}} + \mu_{13}\frac{\partial H}{\partial \mu_{13}} + \nu_{12}\frac{\partial H}{\partial \nu_{12}} + \nu_{13}\frac{\partial H}{\partial \nu_{13}} \\
\dot{\mu}_{22} &= \mu_{23}\frac{\partial H}{\partial \mu_{23}} - \mu_{12}\frac{\partial H}{\partial \mu_{12}} + \nu_{23}\frac{\partial H}{\partial \nu_{23}} - \nu_{12}\frac{\partial H}{\partial \nu_{12}} \\
\dot{\mu}_{33} &= -\mu_{13}\frac{\partial H}{\partial \mu_{13}} - \mu_{23}\frac{\partial H}{\partial \mu_{23}} - \nu_{13}\frac{\partial H}{\partial \nu_{13}} - \nu_{23}\frac{\partial H}{\partial \nu_{23}} \\
\dot{\mu}_{12} &= \mu_{12}\left(\frac{\partial H}{\partial \mu_{22}} - \frac{\partial H}{\partial \mu_{11}}\right) + \mu_{13}\frac{\partial H}{\partial \mu_{23}} + \nu_{13}\frac{\partial H}{\partial \nu_{23}} \\
\dot{\mu}_{23} &= \mu_{23}\left(\frac{\partial H}{\partial \mu_{33}} - \frac{\partial H}{\partial \mu_{22}}\right) - \mu_{13}\frac{\partial H}{\partial \mu_{12}} - \nu_{13}\frac{\partial H}{\partial \nu_{12}} \\
\dot{\mu}_{13} &= \mu_{13}\left(\frac{\partial H}{\partial \mu_{33}} - \frac{\partial H}{\partial \mu_{11}}\right) \\
\dot{\nu}_{12} &= \nu_{12}\left(\frac{\partial H}{\partial \mu_{22}} - \frac{\partial H}{\partial \mu_{11}}\right) + \nu_{13}\frac{\partial H}{\partial \mu_{23}} + \nu_{13}\frac{\partial H}{\partial \nu_{23}} \\
\dot{\nu}_{23} &= \nu_{23}\left(\frac{\partial H}{\partial \mu_{33}} - \frac{\partial H}{\partial \mu_{22}}\right) - \nu_{13}\frac{\partial H}{\partial \mu_{12}} - \nu_{13}\frac{\partial H}{\partial \nu_{12}} \\
\dot{\nu}_{13} &= \nu_{13}\left(\frac{\partial H}{\partial \mu_{33}} - \frac{\partial H}{\partial \mu_{11}}\right).
\end{aligned}
\right\}
\tag{10.7.29}
$$

In particular, the smooth function $C : (\mathfrak{t} \circledS \mathfrak{u})^* \to \mathbb{R}$ is a Casimir function if and only if the right hand side of (10.7.28) vanishes. There is an obvious guess for such a function by searching for C satisfying $\delta C/\delta \mu = I$ and $\delta C/\delta \nu = 0$, which gives

$$
C(\mu, \nu) = \mu_{11} + \mu_{22} + \mu_{33}.
$$

This is precisely one of the functions suggested by the coadjoint orbit description (10.7.12), as we have seen before. Another function suggested by this concrete orbit description is

$$
D(\mu, \nu) = \frac{\mu_{13}}{\nu_{13}}
$$

which is, however, only defined on the open Poisson submanifold given by the condition $\nu_{13} \neq 0$. That the open set $\{(\mu, \nu) \in \mathfrak{t}^* \times \mathfrak{u}^* \mid \nu_{13} \neq 0\}$ is a Poisson submanifold of $(\mathfrak{t} \circledS \mathfrak{u})^*$ follows immediately from the last equation in the system (10.7.29). On this open subset, D is also a Casimir function. Indeed a direct computation using

$$
\frac{\delta D}{\delta \mu} = \begin{bmatrix} 0 & 0 & \frac{1}{\nu_{13}} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \qquad \frac{\delta D}{\delta \nu} = \begin{bmatrix} 0 & 0 & -\frac{\mu_{13}}{\nu_{13}^2} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},
$$

and (10.7.29), shows that the Hamiltonian vector field generated by D vanishes. It is clear that on the open Poisson submanifold $\{(\mu, \nu) \in \mathfrak{t}^* \times \mathfrak{u}^* \mid \nu_{13} \neq 0\}$ the two Casimir functions C and D are functionally independent, that is, their differentials at any point of this open set are linearly independent.

Since the dimension of the generic coadjoint orbit is 6, one would expect three globally defined Casimir functions. This is not possible for $(\mathfrak{t} \,\circledS\, \mathfrak{u})^*$ because the generic coadjoint orbit given by the conditions $\nu_{13} \neq 0$ and $\nu_{13} \neq \mu_{13}$ has all elements of the form $(\mu, \nu) = (\text{diagonal matrix}, 0)$ in its closure; each such element is a 0-dimensional orbit. So the generic orbit is not closed. If there would be three functionally independent Casimir functions, the generic coadjoint orbit was necessarily closed. Even on the open Poisson submanifold $\nu_{13} \neq 0$, there cannot exist three functionally independent Casimir functions, because the 4-dimensional orbit given by the conditions $\nu_{13} \neq 0$, $\mu_{13} = \nu_{13}$, $\mu_{12} = \nu_{12}$, $\mu_{23} = \nu_{23}$ is in the closure of the generic 6-dimensional orbit given by the conditions $\nu_{13} \neq 0$ and $\nu_{13} \neq \mu_{13}$. So on the open Poisson submanifold defined by $\nu_{13} \neq 0$, there are at most two globally defined Casimir functions, namely C and D.

Of course, *locally* there are always three Casimir functions. They are given in a Weinstein-Darboux chart by the coordinates transverse to the generic symplectic leaf, that is, to the generic coadjoint orbit. For example, they can be found in the following way. Let (μ, ν) be such that $\nu_{13} \neq 0$ and $\mu_{13} \neq \nu_{13}$ and choose a vector space complement V to $(\mathfrak{t} \,\circledS\, \mathfrak{u})_{(\mu, \nu)}$, that is,

$$\mathfrak{t} \,\circledS\, \mathfrak{u} = (\mathfrak{t} \,\circledS\, \mathfrak{u})_{(\mu, \nu)} \oplus V.$$

Since $\left[(\mathfrak{t} \,\circledS\, \mathfrak{u})_{(\mu, \nu)} \right]^\circ = T_{(\mu, \nu)} \mathcal{O}_{(\mu, \nu)}$ it follows that

$$(\mathfrak{t} \,\circledS\, \mathfrak{u})^* = \left[(\mathfrak{t} \,\circledS\, \mathfrak{u})_{(\mu, \nu)} \right]^\circ \oplus V^\circ = T_{(\mu, \nu)} \mathcal{O}_{(\mu, \nu)} \oplus V^\circ.$$

Therefore, a transverse slice to the generic orbit $\mathcal{O}_{(\mu, \nu)}$ is $(\mu, \nu) + V^\circ$ and hence the coordinate functions on V° provide the three local independent Casimir functions in an open neighborhood of the generic point (μ, ν).

The Reduction by Stages Setup.   Now we shall recover the classification result by using the Reduction by Stages procedure. In this case, the "big group" is $\mathcal{T} \,\circledS\, \mathcal{U}$, the normal subgroup is $\mathcal{U}$, and the phase space is the cotangent bundle $T^* (\mathcal{T} \,\circledS\, \mathcal{U})$. Given is $(\mu, \nu) \in \mathfrak{t}^* \times \mathfrak{u}^*$. The first stage reduced space is the product symplectic manifold $(\mathcal{T} \times \mathfrak{t}^* \times \mathcal{O}_\nu, \omega_{\mathcal{T}}^R \oplus \omega_{\mathcal{O}_\nu}^+)$. By (10.6.28) and (10.6.7), the Lie subgroup $\mathcal{T}^\nu \subset \mathcal{T}$ acts symplectically by

$$(t, \kappa, \rho) \cdot l := \left(tl, \kappa, \Pi_{\mathfrak{u}} \left(l^T \rho (l^{-1})^T \right) \right), \tag{10.7.30}$$

where $t \in \mathcal{T}$, $\kappa \in \mathfrak{t}^*$, $\rho \in \mathcal{O}_\nu$, and $l \in \mathcal{T}^\nu$.

The $\mathcal{T}^\nu$-momentum map $\mathbf{J}_\nu^R : \mathcal{T} \times \mathfrak{t}^* \times \mathcal{O}_\nu \to (\mathfrak{t}^\nu)^*$ is, according to (10.6.30), equal to

$$\mathbf{J}_\nu^R \left(t, \kappa, \mathrm{Ad}_{u^{-1}}^* \nu \right) = j_\nu^* \, \mathrm{Ad}_t^* \, \kappa + j_\nu^* T_I^* \phi^{u^{-1}} T_{u^{-1}}^* R_u \nu,$$

where $j_\nu : \mathfrak{t}^\nu \hookrightarrow \mathfrak{t}$ is the inclusion and $j_\nu^* : \mathfrak{t}^* \to (\mathfrak{t}^\nu)^*$ is the projection relative to the pairing $\langle A, B \rangle = \operatorname{Trace}(AB^T)$. In the first summand we have $\operatorname{Ad}_t^* \kappa = \Pi_{\mathfrak{t}} \left(t^T \kappa (t^{-1})^T \right)$, where $\Pi_{\mathfrak{t}} : \mathfrak{gl}(n) \to \mathfrak{t}$ is the projection that sends a matrix to its upper triangular part. The second summand is the $\mathcal{T}^\nu$-momentum map $\mathbf{K}^\nu : \mathcal{O}_\nu \to (\mathfrak{t}^\nu)^*$ which will be explicitly computed below. Since $(R_u \circ \phi^{u^{-1}})(t) = tu^{-1}t^{-1}u$, we have for any $\xi \in \mathfrak{t}^\nu$

$$\left(T_{u^{-1}} R_u \circ T_I \phi^{u^{-1}} \right)(\xi) = T_I \left(R_u \circ \phi^{u^{-1}} \right)(\xi) = \left. \frac{d}{dt} \right|_{t=0} \left(R_u \circ \phi^{u^{-1}} \right) \left(e^{t\xi} \right)$$

$$= \left. \frac{d}{dt} \right|_{t=0} e^{t\xi} u^{-1} e^{-t\xi} u = \xi - u^{-1} \xi u$$

and hence

$$\left\langle j_\nu^* T_I^* \phi^{u^{-1}} T_{u^{-1}}^* R_u \nu, \xi \right\rangle = \left\langle \nu, \left(T_{u^{-1}} R_u \circ T_I \phi^{u^{-1}} \right)(\xi) \right\rangle = \left\langle \nu, \xi - u^{-1} \xi u \right\rangle$$

$$= \langle \nu, \xi \rangle - \langle \nu, u^{-1} \xi u \rangle = \operatorname{trace}(\nu \xi^T) - \operatorname{trace}(\nu u^T \xi^T (u^{-1})^T)$$

$$= \operatorname{trace}((\nu - (u^{-1})^T \nu u^T)\xi^T) = \langle \nu - (u^{-1})^T \nu u^T, \xi \rangle$$

$$= \left\langle \nu - \Pi_{\mathfrak{t}} \left((u^{-1})^T \nu u^T \right), \xi \right\rangle.$$

Thus the momentum map $\mathbf{J}_\nu^R : \mathcal{T} \times \mathfrak{t}^* \times \mathcal{O}_\nu \to (\mathfrak{t}^\nu)^*$ is given by

$$\mathbf{J}_\nu^R(t, \kappa, \operatorname{Ad}_{u^{-1}}^* \nu) = j_\nu^* \left[\Pi_{\mathfrak{t}} \left(t^T \kappa (t^{-1})^T \right) + \nu - \Pi_{\mathfrak{t}} \left((u^{-1})^T \nu u^T \right) \right].$$
$$(10.7.31)$$

Of course, the projection $j_\nu^* : \mathfrak{t}^* \to (\mathfrak{t}^\nu)^*$ has to be computed on a case by case basis.

The momentum map $\mathbf{J}_\nu^R$ is not equivariant and has a right $(\mathfrak{t}^\nu)^*$-valued group one-cocycle $\sigma^\nu : \mathcal{T}^\nu \to (\mathfrak{t}^\nu)^*$ given by (10.6.23), that is,

$$\sigma^\nu(l) = -j_\nu^* T_I^* \phi^{u(l)} T_{u(l)}^* L_{u(l)^{-1}} \nu,$$

where $u(l) \in \mathcal{U}$ is any element satisfying $\tilde{\phi}(l)^* \nu = \operatorname{Ad}_{u(l)^{-1}}^* \nu$. To explicitly compute σ^ν we note first that for any $\xi \in \mathfrak{t}^\nu$ we have

$$\left(T_{u(l)} L_{u(l)^{-1}} \circ T_I \phi^{u(l)} \right)(\xi) = T_I \left(L_{u(l)^{-1}} \circ \phi^{u(l)} \right)(\xi)$$

$$= \left. \frac{d}{dt} \right|_{t=0} \left(L_{u(l)^{-1}} \circ \phi^{u(l)} \right) \left(e^{t\xi} \right) = \left. \frac{d}{dt} \right|_{t=0} u(l)^{-1} e^{t\xi} u(l) e^{-t\xi}$$

$$= u(l)^{-1} \xi u(l) - \xi.$$

Therefore, for any $\xi \in \mathfrak{t}^\nu$ we get

$$\left\langle j_\nu^* T_I^* \phi^{u(l)} T_{u(l)}^* L_{u(l)^{-1}} \nu, \xi \right\rangle = \left\langle \nu, \left(T_{u(l)} L_{u(l)^{-1}} \circ T_I \phi^{u(l)} \right)(\xi) \right\rangle$$

$$= \left\langle \nu, u(l)^{-1} \xi u(l) - \xi \right\rangle = \left\langle \nu, u(l)^{-1} \xi u(l) \right\rangle - \langle \nu, \xi \rangle$$

$$= \operatorname{trace} \left(\nu u(l)^T \xi^T (u(l)^{-1})^T \right) - \operatorname{trace} \left(\nu \xi^T \right)$$

$$= \operatorname{trace} \left(((u(l)^{-1})^T \nu u(l)^T - \nu) \xi^T \right) = \left\langle (u(l)^{-1})^T \nu u(l)^T - \nu, \xi \right\rangle$$

$$= \left\langle \Pi_{\mathfrak{t}} \left((u(l)^{-1})^T \nu u(l)^T \right) - \nu, \xi \right\rangle$$

and hence the non-equivariance group one cocycle $\sigma^\nu : \mathcal{T}^\nu \to (\mathfrak{t}^\nu)^*$ of $\mathbf{J}_\nu^R$ is given by

$$\sigma^\nu(l) = j_\nu^* \left[\nu - \Pi_{\mathfrak{t}} \left((u(l)^{-1})^T \nu u(l)^T \right) \right], \tag{10.7.32}$$

where $u(l) \in \mathcal{U}$ is any element satisfying $\tilde{\phi}(l)^* \nu = \operatorname{Ad}_{u(l)^{-1}}^* \nu$, that is, $\Pi_{\mathfrak{u}} \left(l^T \nu (l^{-1})^T \right) = \Pi_{\mathfrak{u}} \left((u(l)^{-1})^T \nu u(l)^T \right)$.

If $\mu \in \mathfrak{t}^*$, $\nu \in \mathfrak{u}^*$, Theorem 10.6.6 states that the coadjoint orbit of $\mathcal{T} \,\textcircled{s}\, \mathcal{U}$ through (μ, ν) is symplectically diffeomorphic to the reduced space $\left(\mathbf{J}_\nu^R \right)^{-1} (\mu_\nu)/(\mathcal{T}^\nu)_{\mu_\nu}$. This procedure gives, in principle, all coadjoint orbits in $(\mathfrak{t} \,\textcircled{s}\, \mathfrak{u})^*$. There are various objects that need to be determined explicitly in order for the method to be implemented in a concrete case. We shall do this below for $n = 3$, the first fully nonabelian situation for $\mathcal{T} \,\textcircled{s}\, \mathcal{U}$.

The Second Stage Symmetry Group for $n = 3$.   We begin with the computation of $\mathcal{T}^\nu$. Since

$$t^{-1} = \begin{bmatrix} \frac{1}{t_{11}} & -\frac{t_{12}}{t_{11}t_{22}} & \frac{t_{12}t_{23}}{t_{11}t_{22}t_{33}} - \frac{t_{13}}{t_{11}t_{33}} \\ 0 & \frac{1}{t_{22}} & -\frac{t_{23}}{t_{22}t_{33}} \\ 0 & 0 & \frac{1}{t_{33}} \end{bmatrix}$$

formulas (10.6.7), (10.6.8), and (10.5.9) imply that $t \in \mathcal{T}^\nu$ if and only if the matrix

$$\tilde{\phi}(t^{-1})^* \nu = \Pi_{\mathfrak{u}} \left((t^{-1})^T \nu t^T \right) = \begin{bmatrix} 0 & \frac{t_{22}}{t_{11}} \nu_{12} + \frac{t_{23}}{t_{11}} \nu_{13} & \frac{t_{33}}{t_{11}} \nu_{13} \\ 0 & 0 & \frac{t_{33}}{t_{22}} \nu_{23} - \frac{t_{12}t_{33}}{t_{11}t_{22}} \nu_{13} \\ 0 & 0 & 0 \end{bmatrix}$$

is in the $\mathcal{U}$-coadjoint orbit $\mathcal{O}_\nu$, which means that it is of the form

$$\begin{bmatrix} 0 & \nu_{12} + u_{23}\nu_{13} & \nu_{13} \\ 0 & 0 & \nu_{23} - u_{12}\nu_{13} \\ 0 & 0 & 0 \end{bmatrix}$$

for some $u_{12}, u_{23} \in \mathbb{R}$. Thus, to determine $\mathcal{T}^\nu$ we need to find all $t \in \mathcal{T}$ such that

$$\left. \begin{aligned} \frac{t_{22}}{t_{11}}\nu_{12} + \frac{t_{23}}{t_{11}}\nu_{13} &= \nu_{12} + u_{23}\nu_{13} \\ \frac{t_{33}}{t_{22}}\nu_{23} - \frac{t_{12}t_{33}}{t_{11}t_{22}}\nu_{13} &= \nu_{23} - u_{12}\nu_{13} \\ \frac{t_{33}}{t_{11}}\nu_{13} &= \nu_{13} \end{aligned} \right\} \tag{10.7.33}$$

for some numbers $u_{12}, u_{23} \in \mathbb{R}$. As expected, there are several cases to consider.

Case I: If $\nu_{13} \neq 0$, the third equation in (10.7.33) implies that $t_{11} = t_{33}$ and so the system becomes

$$\left.\begin{array}{l} \dfrac{t_{22}}{t_{33}}\nu_{12} + \dfrac{t_{23}}{t_{33}}\nu_{13} = \nu_{12} + u_{23}\nu_{13} \\[2mm] \dfrac{t_{33}}{t_{22}}\nu_{23} - \dfrac{t_{12}}{t_{22}}\nu_{13} = \nu_{23} - u_{12}\nu_{13} \end{array}\right\}$$

for some $u_{12}, u_{23} \in \mathbb{R}$. This is possible if we choose

$$u_{23} = -\frac{t_{33}-t_{22}}{t_{33}}\frac{\nu_{12}}{\nu_{13}} + \frac{t_{23}}{t_{33}}$$

$$u_{12} = -\frac{t_{33}-t_{22}}{t_{22}}\frac{\nu_{23}}{\nu_{13}} + \frac{t_{12}}{t_{22}}.$$

Therefore

$$\mathcal{T}^\nu = \left\{ \begin{bmatrix} t_{33} & t_{12} & t_{13} \\ 0 & t_{22} & t_{23} \\ 0 & 0 & t_{33} \end{bmatrix} \,\middle|\, t_{ij} \in \mathbb{R}, t_{ii} \neq 0 \right\},$$

which is a closed 5-dimensional Lie subgroup of $\mathcal{T}$. This is in agreement with the computations done in the first part of §10.6 and at the end of §10.5: $\dim \mathcal{T}^\nu = \dim(\mathcal{T} \circledS \mathcal{U})_\nu - \dim \mathcal{U}_\nu = 6 - 1 = 5$.

Next we need to determine the dual of $\mathfrak{t}^\nu$ in terms of the pairing $\langle A, B \rangle = \mathrm{Trace}(AB^T)$. To do this write $\mathfrak{t} = \mathfrak{t}^\nu \oplus \mathfrak{c}_1$, where

$$\mathfrak{t}^\nu = \left\{ \xi = \begin{bmatrix} \xi_{33} & \xi_{12} & \xi_{13} \\ 0 & \xi_{22} & \xi_{23} \\ 0 & 0 & \xi_{33} \end{bmatrix} \,\middle|\, \xi_{12}, \xi_{13}, \xi_{22}, \xi_{23}, \xi_{33} \in \mathbb{R} \right\}$$

and

$$\mathfrak{c}_1 = \left\{ \begin{bmatrix} \zeta & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \,\middle|\, \zeta \in \mathbb{R} \right\}.$$

Therefore, $\mathfrak{t}^* = (\mathfrak{t}^\nu)^\circ \oplus \mathfrak{c}_1^\circ \cong \mathfrak{c}_1^* \oplus (\mathfrak{t}^\nu)^*$. We shall identify in what follows $(\mathfrak{t}^\nu)^\circ = \mathfrak{c}_1^*$ and $\mathfrak{c}_1^\circ = (\mathfrak{t}^\nu)^*$. It is easily verified that

$$\mathfrak{c}_1^* = (\mathfrak{t}^\nu)^\circ = \left\{ \begin{bmatrix} \lambda & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -\lambda \end{bmatrix} \,\middle|\, \lambda \in \mathbb{R} \right\}$$

and

$$(\mathfrak{t}^\nu)^* = \mathfrak{c}_1^\circ = \left\{ \begin{bmatrix} 0 & \mu_{12} & \mu_{13} \\ 0 & \mu_{22} & \mu_{23} \\ 0 & 0 & \mu_{33} \end{bmatrix} \,\middle|\, \mu_{12}, \mu_{13}, \mu_{22}, \mu_{23}, \mu_{33} \in \mathbb{R} \right\}$$

so that $\mu \in \mathfrak{t}^*$ is written in the direct sum $\mathfrak{t}^* = (\mathfrak{t}^\nu)^\circ \oplus \mathfrak{c}_1^\circ$ as

$$\mu := \begin{bmatrix} \mu_{11} & \mu_{12} & \mu_{13} \\ 0 & \mu_{22} & \mu_{23} \\ 0 & 0 & \mu_{33} \end{bmatrix} = \begin{bmatrix} \mu_{11} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -\mu_{11} \end{bmatrix} + \begin{bmatrix} 0 & \mu_{12} & \mu_{13} \\ 0 & \mu_{22} & \mu_{23} \\ 0 & 0 & \mu_{33} + \mu_{11} \end{bmatrix}.$$

Therefore, if $j_\nu : \mathfrak{t}^\nu \hookrightarrow \mathfrak{t}$ is the inclusion, the projection $j_\nu^* : \mathfrak{t}^* \to (\mathfrak{t}^\nu)^*$ is given by

$$\mu_\nu := j_\nu^* \mu = j_\nu^* \begin{bmatrix} \mu_{11} & \mu_{12} & \mu_{13} \\ 0 & \mu_{22} & \mu_{23} \\ 0 & 0 & \mu_{33} \end{bmatrix} = \begin{bmatrix} 0 & \mu_{12} & \mu_{13} \\ 0 & \mu_{22} & \mu_{23} \\ 0 & 0 & \mu_{33} + \mu_{11} \end{bmatrix}. \qquad (10.7.34)$$

Case II: If $\nu_{13} = 0$, the third equation in (10.7.33) disappears and the system becomes

$$\frac{t_{22}}{t_{11}} \nu_{12} = \nu_{12}, \qquad \frac{t_{33}}{t_{22}} \nu_{23} = \nu_{23}.$$

Note also that by (10.5.9), we have $\mathcal{O}_\nu = \{\nu\}$.

Case IIA: If $\nu_{12} = \nu_{23} = 0$, that is, $\nu = 0$, then these equations are identically satisfied and hence $\mathcal{T}^\nu = \mathcal{T}$, $\mathfrak{t}^\nu = \mathfrak{t}$, and j_ν^* is the identity map.

Case IIB: If $\nu_{12} \neq 0$ and $\nu_{23} = 0$, then $t_{11} = t_{22}$ and hence

$$\mathcal{T}^\nu = \left\{ \begin{bmatrix} t_{11} & t_{12} & t_{13} \\ 0 & t_{11} & t_{23} \\ 0 & 0 & t_{33} \end{bmatrix} \middle| t_{ij} \in \mathbb{R}, t_{ii} \neq 0 \right\},$$

which is a closed 5-dimensional Lie subgroup of $\mathcal{T}$.

To determine the dual $\mathfrak{t}^\nu$ using the pairing $\langle A, B \rangle = \mathrm{Trace}(AB^T)$, we proceed as in the previous case writing $\mathfrak{t} = \mathfrak{t}^\nu \oplus \mathfrak{c}_{2B}$, where

$$\mathfrak{t}^\nu = \left\{ \xi = \begin{bmatrix} \xi_{11} & \xi_{12} & \xi_{13} \\ 0 & \xi_{11} & \xi_{23} \\ 0 & 0 & \xi_{33} \end{bmatrix} \middle| \xi_{11}, \xi_{12}, \xi_{13}, \xi_{23}, \xi_{33} \in \mathbb{R} \right\}$$

and

$$\mathfrak{c}_{2B} = \left\{ \begin{bmatrix} 0 & 0 & 0 \\ 0 & \zeta & 0 \\ 0 & 0 & 0 \end{bmatrix} \middle| \zeta \in \mathbb{R} \right\}.$$

Therefore, $\mathfrak{t}^* = (\mathfrak{t}^\nu)^\circ \oplus \mathfrak{c}_{2B}^\circ \cong \mathfrak{c}_{2B}^* \oplus (\mathfrak{t}^\nu)^*$ and we identify, as before, $(\mathfrak{t}^\nu)^\circ = \mathfrak{c}_{2B}^*$ and $\mathfrak{c}_{2B}^\circ = (\mathfrak{t}^\nu)^*$. It is easily verified that

$$\mathfrak{c}_{2B}^* = (\mathfrak{t}^\nu)^\circ = \left\{ \begin{bmatrix} -\lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & 0 \end{bmatrix} \middle| \lambda \in \mathbb{R} \right\}$$

and

$$(\mathfrak{t}^\nu)^* = \mathfrak{c}_{2B}^\circ = \left\{ \begin{bmatrix} \mu_{11} & \mu_{12} & \mu_{13} \\ 0 & 0 & \mu_{23} \\ 0 & 0 & \mu_{33} \end{bmatrix} \, \bigg| \, \mu_{11}, \mu_{12}, \mu_{13}, \mu_{23}, \mu_{33} \in \mathbb{R} \right\}$$

so that $\mu \in \mathfrak{t}^*$ is written in the direct sum $\mathfrak{t}^* = (\mathfrak{t}^\nu)^\circ \oplus \mathfrak{c}_{2B}^\circ$ as

$$\mu := \begin{bmatrix} \mu_{11} & \mu_{12} & \mu_{13} \\ 0 & \mu_{22} & \mu_{23} \\ 0 & 0 & \mu_{33} \end{bmatrix} = \begin{bmatrix} -\mu_{22} & 0 & 0 \\ 0 & \mu_{22} & 0 \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} \mu_{11} + \mu_{22} & \mu_{12} & \mu_{13} \\ 0 & 0 & \mu_{23} \\ 0 & 0 & \mu_{33} \end{bmatrix}.$$

Therefore, if $j_\nu : \mathfrak{t}^\nu \hookrightarrow \mathfrak{t}$ is the inclusion, the projection $j_\nu^* : \mathfrak{t}^* \to (\mathfrak{t}^\nu)^*$ is given by

$$\mu_\nu := j_\nu^* \mu = j_\nu^* \begin{bmatrix} \mu_{11} & \mu_{12} & \mu_{13} \\ 0 & \mu_{22} & \mu_{23} \\ 0 & 0 & \mu_{33} \end{bmatrix} = \begin{bmatrix} \mu_{11} + \mu_{22} & \mu_{12} & \mu_{13} \\ 0 & 0 & \mu_{23} \\ 0 & 0 & \mu_{33} \end{bmatrix}. \quad (10.7.35)$$

Case IIC: If $\nu_{12} = 0$ and $\nu_{23} \neq 0$, then $t_{22} = t_{33}$ and hence

$$\mathcal{T}^\nu = \left\{ \begin{bmatrix} t_{11} & t_{12} & t_{13} \\ 0 & t_{22} & t_{23} \\ 0 & 0 & t_{22} \end{bmatrix} \, \bigg| \, t_{ij} \in \mathbb{R}, t_{ii} \neq 0 \right\},$$

which is a closed 5-dimensional Lie subgroup of $\mathcal{T}$. This case is similar to the previous one. The relevant spaces are

$$\mathfrak{t}^\nu = \left\{ \xi = \begin{bmatrix} \xi_{11} & \xi_{12} & \xi_{13} \\ 0 & \xi_{22} & \xi_{23} \\ 0 & 0 & \xi_{22} \end{bmatrix} \, \bigg| \, \xi_{11}, \xi_{12}, \xi_{13}, \xi_{22}, \xi_{23} \in \mathbb{R} \right\},$$

$$\mathfrak{c}_{2C} = \left\{ \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \zeta \end{bmatrix} \, \bigg| \, \zeta \in \mathbb{R} \right\},$$

$$\mathfrak{c}_{2C}^* = (\mathfrak{t}^\nu)^\circ = \left\{ \begin{bmatrix} 0 & 0 & 0 \\ 0 & -\lambda & 0 \\ 0 & 0 & \lambda \end{bmatrix} \, \bigg| \, \lambda \in \mathbb{R} \right\},$$

$$(\mathfrak{t}^\nu)^* = \mathfrak{c}_{2C}^\circ = \left\{ \begin{bmatrix} \mu_{11} & \mu_{12} & \mu_{13} \\ 0 & \mu_{22} & \mu_{23} \\ 0 & 0 & 0 \end{bmatrix} \, \bigg| \, \mu_{11}, \mu_{12}, \mu_{13}, \mu_{22}, \mu_{23} \in \mathbb{R} \right\}$$

so that $\mu \in \mathfrak{t}^*$ is written in the direct sum $\mathfrak{t}^* = (\mathfrak{t}^\nu)^\circ \oplus \mathfrak{c}_{2C}^\circ$ as

$$\mu := \begin{bmatrix} \mu_{11} & \mu_{12} & \mu_{13} \\ 0 & \mu_{22} & \mu_{23} \\ 0 & 0 & \mu_{33} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -\mu_{33} & 0 \\ 0 & 0 & \mu_{33} \end{bmatrix} + \begin{bmatrix} \mu_{11} & \mu_{12} & \mu_{13} \\ 0 & \mu_{22} + \mu_{33} & \mu_{23} \\ 0 & 0 & 0 \end{bmatrix},$$

and hence if $j_\nu : \mathfrak{t}^\nu \hookrightarrow \mathfrak{t}$ is the inclusion, the projection $j_\nu^* : \mathfrak{t}^* \to (\mathfrak{t}^\nu)^*$ is given by

$$\mu_\nu := j_\nu^* \mu = j_\nu^* \begin{bmatrix} \mu_{11} & \mu_{12} & \mu_{13} \\ 0 & \mu_{22} & \mu_{23} \\ 0 & 0 & \mu_{33} \end{bmatrix} = \begin{bmatrix} \mu_{11} & \mu_{12} & \mu_{13} \\ 0 & \mu_{22} + \mu_{33} & \mu_{23} \\ 0 & 0 & 0 \end{bmatrix}. \quad (10.7.36)$$

Case IID: If $\nu_{12} \neq 0$ and $\nu_{23} \neq 0$, then $t_{11} = t_{22} = t_{33}$ and hence

$$\mathcal{T}^\nu = \left\{ \begin{bmatrix} t_{11} & t_{12} & t_{13} \\ 0 & t_{11} & t_{23} \\ 0 & 0 & t_{11} \end{bmatrix} \bigg| t_{ij} \in \mathbb{R}, t_{11} \neq 0 \right\},$$

which is a closed 4-dimensional Lie subgroup of $\mathcal{T}$. In this case we have

$$\mathfrak{t}^\nu = \left\{ \xi = \begin{bmatrix} \xi_{11} & \xi_{12} & \xi_{13} \\ 0 & \xi_{11} & \xi_{23} \\ 0 & 0 & \xi_{11} \end{bmatrix} \bigg| \xi_{11}, \xi_{12}, \xi_{13}, \xi_{23} \in \mathbb{R} \right\},$$

$$\mathfrak{c}_{2D} = \left\{ \begin{bmatrix} 0 & 0 & 0 \\ 0 & \eta & 0 \\ 0 & 0 & \zeta \end{bmatrix} \bigg| \eta, \zeta \in \mathbb{R} \right\},$$

$$\mathfrak{c}_{2D}^* = (\mathfrak{t}^\nu)^\circ = \left\{ \begin{bmatrix} -\kappa & 0 & 0 \\ 0 & \kappa - \lambda & 0 \\ 0 & 0 & \lambda \end{bmatrix} \bigg| \kappa, \lambda \in \mathbb{R} \right\},$$

$$(\mathfrak{t}^\nu)^* = \mathfrak{c}_{2D}^\circ = \left\{ \begin{bmatrix} \mu_{11} & \mu_{12} & \mu_{13} \\ 0 & \mu_{11} & \mu_{23} \\ 0 & 0 & \mu_{11} \end{bmatrix} \bigg| \mu_{11}, \mu_{12}, \mu_{13}, \mu_{23} \in \mathbb{R} \right\}$$

so that $\mu \in \mathfrak{t}^*$ is written in the direct sum $\mathfrak{t}^* = (\mathfrak{t}^\nu)^\circ \oplus \mathfrak{c}_{2D}^\circ$ as

$$\mu := \begin{bmatrix} \mu_{11} & \mu_{12} & \mu_{13} \\ 0 & \mu_{22} & \mu_{23} \\ 0 & 0 & \mu_{33} \end{bmatrix}$$

$$= \begin{bmatrix} \dfrac{2\mu_{11} - \mu_{22} - \mu_{33}}{3} & 0 & 0 \\ 0 & \dfrac{-\mu_{11} + 2\mu_{22} - \mu_{33}}{3} & 0 \\ 0 & 0 & \dfrac{-\mu_{11} - \mu_{22} + 2\mu_{33}}{3} \end{bmatrix}$$

$$+ \begin{bmatrix} \dfrac{\mu_{11} + \mu_{22} + \mu_{33}}{3} & \mu_{12} & \mu_{13} \\ 0 & \dfrac{\mu_{11} + \mu_{22} + \mu_{33}}{3} & \mu_{23} \\ 0 & 0 & \dfrac{\mu_{11} + \mu_{22} + \mu_{33}}{3} \end{bmatrix},$$

and hence if $j_\nu : \mathfrak{t}^\nu \hookrightarrow \mathfrak{t}$ is the inclusion, , the projection $j_\nu^* : \mathfrak{t}^* \rightarrow (\mathfrak{t}^\nu)^*$ is given by

$$\mu_\nu := j_\nu^* \mu = j_\nu^* \begin{bmatrix} \mu_{11} & \mu_{12} & \mu_{13} \\ 0 & \mu_{22} & \mu_{23} \\ 0 & 0 & \mu_{33} \end{bmatrix}$$

$$= \begin{bmatrix} \dfrac{\mu_{11} + \mu_{22} + \mu_{33}}{3} & \mu_{12} & \mu_{13} \\ 0 & \dfrac{\mu_{11} + \mu_{22} + \mu_{33}}{3} & \mu_{23} \\ 0 & 0 & \dfrac{\mu_{11} + \mu_{22} + \mu_{33}}{3} \end{bmatrix} .$$

$$(10.7.37)$$

The Second Stage Momentum Level Set for $n = 3$.   By (10.7.31), $(\mathbf{J}_\nu^R)^{-1}(\mu_\nu)$ consists of triples $(t, \kappa, \mathrm{Ad}_{u^{-1}}^* \nu) \in \mathcal{T} \times \mathfrak{t}^* \times \mathcal{O}_\nu$ such that

$$j_\nu^* \left[\Pi_{\mathfrak{t}} \left(t^T \kappa (t^{-1})^T \right) + \nu - \Pi_{\mathfrak{t}} \left((u^{-1})^T \nu u^T \right) - \mu \right] = 0. \qquad (10.7.38)$$

By (10.7.5) and (10.7.3), the non-zero entries of $\Pi_{\mathfrak{t}} \left(t^T \kappa (t^{-1})^T \right)$ and $\Pi_{\mathfrak{t}} \left((u^{-1})^T \nu u^T \right)$ are

$$[\Pi_{\mathfrak{t}} \left(t^T \kappa (t^{-1})^T \right)]_{11} = \kappa_{11} - \frac{t_{12}}{t_{22}} \kappa_{12} + \frac{t_{12} t_{23}}{t_{22} t_{33}} \kappa_{13} - \frac{t_{13}}{t_{33}} \kappa_{13}$$

$$[\Pi_{\mathfrak{t}} \left(t^T \kappa (t^{-1})^T \right)]_{22} = \kappa_{22} + \frac{t_{12}}{t_{22}} \kappa_{12} - \frac{t_{12} t_{23}}{t_{22} t_{33}} \kappa_{13} - \frac{t_{23}}{t_{33}} \kappa_{23}$$

$$[\Pi_{\mathfrak{t}} \left(t^T \kappa (t^{-1})^T \right)]_{33} = \kappa_{33} + \frac{t_{13}}{t_{33}} \kappa_{13} + \frac{t_{23}}{t_{33}} \kappa_{23}$$

$$[\Pi_{\mathfrak{t}} \left(t^T \kappa (t^{-1})^T \right)]_{12} = \frac{t_{11}}{t_{22}} \kappa_{12} - \frac{t_{11} t_{23}}{t_{22} t_{33}} \kappa_{13}$$

$$[\Pi_{\mathfrak{t}} \left(t^T \kappa (t^{-1})^T \right)]_{23} = \frac{t_{12}}{t_{33}} \kappa_{13} + \frac{t_{22}}{t_{33}} \kappa_{23}$$

$$[\Pi_{\mathfrak{t}} \left(t^T \kappa (t^{-1})^T \right)]_{13} = \frac{t_{11}}{t_{33}} \kappa_{13}$$

$$[\Pi_{\mathfrak{t}} \left((u^{-1})^T \nu u^T \right)]_{11} = \nu_{12} u_{12} + \nu_{13} u_{13}$$

$$[\Pi_{\mathfrak{t}} \left((u^{-1})^T \nu u^T \right)]_{22} = -\nu_{12} u_{12} - \nu_{13} u_{12} u_{23} + \nu_{23} u_{23}$$

$$[\Pi_{\mathfrak{t}} \left((u^{-1})^T \nu u^T \right)]_{33} = \nu_{13} u_{12} u_{23} - \nu_{13} u_{13} - \nu_{23} u_{23}$$

$$[\Pi_{\mathfrak{t}} \left((u^{-1})^T \nu u^T \right)]_{12} = \nu_{12} + \nu_{13} u_{23}$$

$$[\Pi_{\mathfrak{t}} \left((u^{-1})^T \nu u^T \right)]_{23} = \nu_{23} - \nu_{13} u_{12}$$

$$[\Pi_{\mathfrak{t}} \left((u^{-1})^T \nu u^T \right)]_{13} = \nu_{13}.$$

Therefore the entries of $\Xi := \Pi_{\mathfrak{t}} \left(t^T \kappa (t^{-1})^T \right) + \nu - \Pi_{\mathfrak{t}} \left((u^{-1})^T \nu u^T \right) - \mu$ are

$$
\left.
\begin{aligned}
\Xi_{11} &= \kappa_{11} - \frac{t_{12}}{t_{22}}\kappa_{12} + \frac{t_{12}t_{23}}{t_{22}t_{33}}\kappa_{13} - \frac{t_{13}}{t_{33}}\kappa_{13} \\
&\quad - \nu_{12}u_{12} - \nu_{13}u_{13} - \mu_{11} \\
\Xi_{22} &= \kappa_{22} + \frac{t_{12}}{t_{22}}\kappa_{12} - \frac{t_{12}t_{23}}{t_{22}t_{33}}\kappa_{13} - \frac{t_{23}}{t_{33}}\kappa_{23} \\
&\quad + \nu_{12}u_{12} + \nu_{13}u_{12}u_{23} - \nu_{23}u_{23} - \mu_{22} \\
\Xi_{33} &= \kappa_{33} + \frac{t_{13}}{t_{33}}\kappa_{13} + \frac{t_{23}}{t_{33}}\kappa_{23} - \nu_{13}u_{12}u_{23} \\
&\quad + \nu_{13}u_{13} + \nu_{23}u_{23} - \mu_{33} \\
\Xi_{12} &= \frac{t_{11}}{t_{22}}\kappa_{12} - \frac{t_{11}t_{23}}{t_{22}t_{33}}\kappa_{13} - \nu_{13}u_{23} - \mu_{12} \\
\Xi_{23} &= \frac{t_{12}}{t_{33}}\kappa_{13} + \frac{t_{22}}{t_{33}}\kappa_{23} + \nu_{13}u_{12} - \mu_{23} \\
\Xi_{13} &= \frac{t_{11}}{t_{33}}\kappa_{13} - \mu_{13}.
\end{aligned}
\right\} \qquad (10.7.39)
$$

Now we compute the momentum level set $(\mathbf{J}_\nu^R)^{-1}(\mu_\nu)$ in all cases for $\mathcal{T}^\nu$ identified in the previous paragraph.

Case I: $\nu_{13} \neq 0$. In view of (10.7.34) and (10.7.39), condition (10.7.38) becomes

$$
\begin{aligned}
\mu_{11} + \mu_{33} &= \kappa_{11} - \frac{t_{12}}{t_{22}}\kappa_{12} + \frac{t_{12}t_{23}}{t_{22}t_{33}}\kappa_{13} + \kappa_{33} + \frac{t_{23}}{t_{33}}\kappa_{23} \\
&\quad - \nu_{12}u_{12} - \nu_{13}u_{12}u_{23} + \nu_{23}u_{23} \\
\mu_{22} &= \kappa_{22} + \frac{t_{12}}{t_{22}}\kappa_{12} - \frac{t_{12}t_{23}}{t_{22}t_{33}}\kappa_{13} - \frac{t_{23}}{t_{33}}\kappa_{23} + \nu_{12}u_{12} + \nu_{13}u_{12}u_{23} - \nu_{23}u_{23} \\
\mu_{12} &= \frac{t_{11}}{t_{22}}\kappa_{12} - \frac{t_{11}t_{23}}{t_{22}t_{33}}\kappa_{13} - \nu_{13}u_{23} \\
\mu_{23} &= \frac{t_{12}}{t_{33}}\kappa_{13} + \frac{t_{22}}{t_{33}}\kappa_{23} + \nu_{13}u_{12} \\
\mu_{13} &= \frac{t_{11}}{t_{33}}\kappa_{13}
\end{aligned}
$$

which gives

$$
\begin{aligned}
\kappa_{11} &= \mu_{11} + \mu_{33} + \frac{t_{12}}{t_{22}}\kappa_{12} - \frac{t_{12}t_{23}}{t_{22}t_{11}}\mu_{13} - \frac{t_{23}}{t_{33}}\kappa_{23} - \kappa_{33} \\
&+ \frac{\nu_{12}}{\nu_{13}}\left(\mu_{23} - \frac{t_{12}}{t_{11}}\mu_{13} - \frac{t_{22}}{t_{33}}\kappa_{23}\right) \\
&+ \frac{1}{\nu_{13}}\left(\mu_{23} - \frac{t_{12}}{t_{11}}\mu_{13} - \frac{t_{22}}{t_{33}}\kappa_{23}\right)\left(-\mu_{12} + \frac{t_{11}}{t_{22}}\kappa_{12} - \frac{t_{23}}{t_{22}}\mu_{13}\right) \\
&- \frac{\nu_{23}}{\nu_{13}}\left(-\mu_{12} + \frac{t_{11}}{t_{22}}\kappa_{12} - \frac{t_{23}}{t_{22}}\mu_{13}\right) \\
\kappa_{22} &= \mu_{22} - \frac{t_{12}}{t_{22}}\kappa_{12} + \frac{t_{12}t_{23}}{t_{22}t_{11}}\mu_{13} + \frac{t_{23}}{t_{33}}\kappa_{23} \\
&- \frac{\nu_{12}}{\nu_{13}}\left(\mu_{23} - \frac{t_{12}}{t_{11}}\mu_{13} - \frac{t_{22}}{t_{33}}\kappa_{23}\right) \\
&- \frac{1}{\nu_{13}}\left(\mu_{23} - \frac{t_{12}}{t_{11}}\mu_{13} - \frac{t_{22}}{t_{33}}\kappa_{23}\right)\left(-\mu_{12} + \frac{t_{11}}{t_{22}}\kappa_{12} - \frac{t_{23}}{t_{22}}\mu_{13}\right) \\
&+ \frac{\nu_{23}}{\nu_{13}}\left(-\mu_{12} + \frac{t_{11}}{t_{22}}\kappa_{12} - \frac{t_{23}}{t_{22}}\mu_{13}\right) \\
u_{23} &= \frac{1}{\nu_{13}}\left(-\mu_{12} + \frac{t_{11}}{t_{22}}\kappa_{12} - \frac{t_{23}}{t_{22}}\mu_{13}\right) \\
u_{12} &= -\frac{1}{\nu_{13}}\left(-\mu_{23} + \frac{t_{12}}{t_{11}}\mu_{13} + \frac{t_{22}}{t_{33}}\kappa_{23}\right) \\
\kappa_{13} &= \frac{t_{33}}{t_{11}}\mu_{13}.
\end{aligned}
$$

$$(10.7.40)$$

By (10.5.9) we have

$$
\rho := \mathrm{Ad}^*_{u^{-1}}\nu = \Pi_{\mathfrak{u}}\left((u^{-1})^T \nu u^T\right) = \begin{bmatrix} 0 & \nu_{12} + \nu_{13}u_{23} & \nu_{13} \\ 0 & 0 & -\nu_{13}u_{12} + \nu_{23} \\ 0 & 0 & 0 \end{bmatrix}.
$$

Thus all entries of ρ are determined, namely

$$
\begin{aligned}
\rho_{12} &= \nu_{12} - \mu_{12} + \frac{t_{11}}{t_{22}}\kappa_{12} - \frac{t_{23}}{t_{22}}\mu_{13} \\
\rho_{23} &= \nu_{23} - \mu_{23} + \frac{t_{12}}{t_{11}}\mu_{13} + \frac{t_{22}}{t_{33}}\kappa_{23} \\
\rho_{13} &= \nu_{13}.
\end{aligned}
$$

$$(10.7.41)$$

This shows that $\left(\mathbf{J}^R_\nu\right)^{-1}(\mu_\nu)$ consists of triples

$$
(t, \kappa, \rho := \mathrm{Ad}^*_{u^{-1}}\nu) \in \mathcal{T} \times \mathfrak{t}^* \times \mathcal{O}_\nu
$$

such that (10.7.40) holds. But this shows that in this level set, t is arbitrary and all the other elements depend additionally only on κ_{12}, κ_{23}, and κ_{33}.

Therefore, the dimension of $\left(\mathbf{J}_\nu^R\right)^{-1}(\mu_\nu)$ is $6+3 = 9$. Of course, this agrees with the expected dimension:

$$\dim(\mathbf{J}_\nu^R)^{-1}(\mu_\nu) = \dim(\mathcal{T} \times \mathfrak{t}^* \times \mathcal{O}_\nu) - \dim \mathcal{T}^\nu = 6+6+2-5 = 9.$$

Case IIA: $\nu_{13} = \nu_{12} = \nu_{23} = 0$, that is, $\nu = 0$. Then $\mathcal{O}_\nu = \{0\}$, $\mathcal{T}^\nu = \mathcal{T}$, $\mu_\nu = \mu$, $\mathfrak{t}^\nu = \mathfrak{t}$, and j_ν^* is the identity map. Thus, using (10.7.39), condition (10.7.38) becomes

$$\left.\begin{aligned}
\kappa_{11} &= \mu_{11} + \frac{t_{12}}{t_{11}}\mu_{12} + \frac{t_{13}}{t_{11}}\mu_{13} \\
\kappa_{22} &= \mu_{22} - \frac{t_{12}}{t_{11}}\mu_{12} - \frac{t_{12}t_{23}}{t_{11}t_{22}}\mu_{13} + \frac{t_{23}}{t_{22}}\mu_{23} \\
\kappa_{33} &= \mu_{33} - \frac{t_{13}}{t_{11}}\mu_{13} + \frac{t_{12}t_{23}}{t_{11}t_{22}}\mu_{13} - \frac{t_{23}}{t_{22}}\mu_{23} \\
\kappa_{12} &= \frac{t_{22}}{t_{11}}\mu_{12} + \frac{t_{23}}{t_{11}}\mu_{13} \\
\kappa_{23} &= \frac{t_{33}}{t_{22}}\mu_{23} - \frac{t_{12}t_{33}}{t_{11}t_{22}}\mu_{13} \\
\kappa_{13} &= \frac{t_{33}}{t_{11}}\mu_{13}.
\end{aligned}\right\} \qquad (10.7.42)$$

which shows that $\left(\mathbf{J}_\nu^R\right)^{-1}(\mu_\nu)$ consists of triples $(t, \kappa, 0) \in \mathcal{T} \times \mathfrak{t}^* \times \{0\}$ such that κ is determined by (10.7.42) and $t \in \mathcal{T}$ is arbitrary. Thus the dimension of $\left(\mathbf{J}_\nu^R\right)^{-1}(\mu_\nu)$ is 6, which agrees with the usual computation

$$\dim \left(\mathbf{J}_\nu^R\right)^{-1}(\mu_\nu) = \dim(\mathcal{T} \times \mathfrak{t}^* \times \{0\}) - \dim \mathcal{T}^\nu = 6+6-6 = 6.$$

Case IIB: $\nu_{13} = 0$, $\nu_{12} \neq 0$, $\nu_{23} = 0$. In view of (10.7.35) and (10.7.39), condition (10.7.38) becomes

$$\mu_{11} + \mu_{22} = \kappa_{11} + \kappa_{22} - \frac{t_{13}}{t_{33}}\kappa_{13} - \frac{t_{23}}{t_{33}}\kappa_{23}$$

$$\mu_{33} = \kappa_{33} + \frac{t_{13}}{t_{33}}\kappa_{13} + \frac{t_{23}}{t_{33}}\kappa_{23}$$

$$\mu_{12} = \frac{t_{11}}{t_{22}}\kappa_{12} - \frac{t_{11}t_{23}}{t_{22}t_{33}}\kappa_{13}$$

$$\mu_{23} = \frac{t_{12}}{t_{33}}\kappa_{13} + \frac{t_{22}}{t_{33}}\kappa_{23}$$

$$\mu_{13} = \frac{t_{11}}{t_{33}}\kappa_{13}$$

which gives

$$
\left.
\begin{aligned}
\kappa_{11} &= -\kappa_{22} + \mu_{11} + \mu_{22} + \frac{t_{13}}{t_{11}}\mu_{13} - \frac{t_{12}t_{23}}{t_{11}t_{22}}\mu_{13} + \frac{t_{23}}{t_{22}}\mu_{23} \\
\kappa_{33} &= \mu_{33} - \frac{t_{13}}{t_{11}}\mu_{13} + \frac{t_{12}t_{23}}{t_{11}t_{22}}\mu_{13} - \frac{t_{23}}{t_{22}}\mu_{23} \\
\kappa_{12} &= \frac{t_{22}}{t_{11}}\mu_{12} + \frac{t_{23}}{t_{11}}\mu_{13} \\
\kappa_{23} &= \frac{t_{33}}{t_{22}}\mu_{23} - \frac{t_{12}t_{33}}{t_{11}t_{22}}\mu_{13} \\
\kappa_{13} &= \frac{t_{33}}{t_{11}}\mu_{13}.
\end{aligned}
\right\}
\tag{10.7.43}
$$

By (10.5.9), since $\nu_{13} = 0$, we have

$$
\rho := \mathrm{Ad}^*_{u^{-1}}\nu = \Pi_{\mathfrak{u}}\left((u^{-1})^T \nu u^T\right) = \begin{bmatrix} 0 & \nu_{12} & 0 \\ 0 & 0 & \nu_{23} \\ 0 & 0 & 0 \end{bmatrix} = \nu,
$$

that is, $\mathcal{O}_\nu = \{\nu\}$.

This shows that $\left(\mathbf{J}^R_\nu\right)^{-1}(\mu_\nu)$ consists of triples $(t, \kappa, \nu) \in \mathcal{T} \times \mathfrak{t}^* \times \{\nu\}$ such that κ is given by (10.7.43) and $t \in \mathcal{T}$ and $\kappa_{22} \in \mathbb{R}$ are arbitrary. Therefore, the dimension of $\left(\mathbf{J}^R_\nu\right)^{-1}(\mu_\nu)$ is $6 + 1 = 7$. Of course, this agrees with the expected dimension: $\dim(\mathbf{J}^R_\nu)^{-1}(\mu_\nu) = \dim(\mathcal{T} \times \mathfrak{t}^* \times \{\nu\}) - \dim\mathcal{T}^\nu = 6 + 6 - 5 = 7$.

Case IIC: $\nu_{13} = 0$, $\nu_{12} = 0$, $\nu_{23} \neq 0$. This case is dealt with like the previous one. As before, $\mathcal{O}_\nu = \{\nu\}$. In view of (10.7.36) and (10.7.39), condition (10.7.38) becomes

$$
\begin{aligned}
\mu_{11} &= \kappa_{11} - \frac{t_{12}}{t_{22}}\kappa_{12} + \frac{t_{12}t_{23}}{t_{22}t_{33}}\kappa_{13} - \frac{t_{13}}{t_{33}}\kappa_{13} \\
\mu_{22} + \mu_{33} &= \kappa_{22} + \kappa_{33} + \frac{t_{12}}{t_{22}}\kappa_{12} - \frac{t_{12}t_{23}}{t_{22}t_{33}}\kappa_{13} + \frac{t_{13}}{t_{33}}\kappa_{13} \\
\mu_{12} &= \frac{t_{11}}{t_{22}}\kappa_{12} - \frac{t_{11}t_{23}}{t_{22}t_{33}}\kappa_{13} \\
\mu_{23} &= \frac{t_{12}}{t_{33}}\kappa_{13} + \frac{t_{22}}{t_{33}}\kappa_{23} \\
\mu_{13} &= \frac{t_{11}}{t_{33}}\kappa_{13}
\end{aligned}
$$

which gives

$$
\left.\begin{aligned}
\kappa_{11} &= \mu_{11} + \frac{t_{12}}{t_{11}}\mu_{12} + \frac{t_{13}}{t_{11}}\mu_{13} \\[2mm]
\kappa_{22} &= -\kappa_{33} + \mu_{22} + \mu_{33} - \frac{t_{12}}{t_{11}}\mu_{12} - \frac{t_{13}}{t_{11}}\mu_{13} \\[2mm]
\kappa_{12} &= \frac{t_{22}}{t_{11}}\mu_{12} + \frac{t_{23}}{t_{11}}\mu_{13} \\[2mm]
\kappa_{23} &= \frac{t_{33}}{t_{22}}\mu_{23} - \frac{t_{12}t_{33}}{t_{11}t_{22}}\mu_{13} \\[2mm]
\kappa_{13} &= \frac{t_{33}}{t_{11}}\mu_{13}.
\end{aligned}\right\} \tag{10.7.44}
$$

This shows that $\left(\mathbf{J}_\nu^R\right)^{-1}(\mu_\nu)$ consists of triples $(t,\kappa,\nu)\in\mathcal{T}\times\mathfrak{t}^*\times\{\nu\}$ such that κ is given by (10.7.44) and $t\in\mathcal{T}$ and $\kappa_{33}\in\mathbb{R}$ are arbitrary. Therefore, the dimension of $\left(\mathbf{J}_\nu^R\right)^{-1}(\mu_\nu)$ is $6+1=7$. Of course, this agrees with the expected dimension: $\dim(\mathbf{J}_\nu^R)^{-1}(\mu_\nu) = \dim(\mathcal{T}\times\mathfrak{t}^*\times\{\nu\}) - \dim\mathcal{T}^\nu = 6+6-5=7$.

Case IID: $\nu_{13}=0$, $\nu_{12}\neq 0$, $\nu_{23}\neq 0$. Again, $\mathcal{O}_\nu = \{\nu\}$ since $\nu_{13}=0$ (see (10.5.8)). In view of (10.7.37) and (10.7.39), condition (10.7.38) becomes

$$
\begin{aligned}
\mu_{11} + \mu_{22} + \mu_{33} &= \kappa_{11} + \kappa_{22} + \kappa_{33} \\[2mm]
\mu_{12} &= \frac{t_{11}}{t_{22}}\kappa_{12} - \frac{t_{11}t_{23}}{t_{22}t_{33}}\kappa_{13} \\[2mm]
\mu_{23} &= \frac{t_{12}}{t_{33}}\kappa_{13} + \frac{t_{22}}{t_{33}}\kappa_{23} \\[2mm]
\mu_{13} &= \frac{t_{11}}{t_{33}}\kappa_{13}
\end{aligned}
$$

which gives

$$
\left.\begin{aligned}
\kappa_{11} &= -\kappa_{22} - \kappa_{33} + \mu_{11} + \mu_{22} + \mu_{33} \\[2mm]
\kappa_{12} &= \frac{t_{22}}{t_{11}}\mu_{12} + \frac{t_{23}}{t_{11}}\mu_{13} \\[2mm]
\kappa_{23} &= \frac{t_{33}}{t_{22}}\mu_{23} - \frac{t_{12}t_{33}}{t_{11}t_{22}}\mu_{13} \\[2mm]
\kappa_{13} &= \frac{t_{33}}{t_{11}}\mu_{13}.
\end{aligned}\right\} \tag{10.7.45}
$$

This shows that $\left(\mathbf{J}_\nu^R\right)^{-1}(\mu_\nu)$ consists of triples $(t,\kappa,\nu)\in\mathcal{T}\times\mathfrak{t}^*\times\{\nu\}$ such that κ is given by (10.7.45) and $t\in\mathcal{T}$, $\kappa_{22}\in\mathbb{R}$, and $\kappa_{33}\in\mathbb{R}$ are arbitrary. Therefore, the dimension of $\left(\mathbf{J}_\nu^R\right)^{-1}(\mu_\nu)$ is $6+2=8$. Of course, this agrees with the expected dimension:

$$
\dim(\mathbf{J}_\nu^R)^{-1}(\mu_\nu) = \dim(\mathcal{T}\times\mathfrak{t}^*\times\{\nu\}) - \dim\mathcal{T}^\nu = 6+6-4=8.
$$

Orbit Classification Via Reduction by Stages for $n = 3$.   Now we compute the second reduced space and thus determine the coadjoint orbit of $\mathcal{T} \circledS \mathcal{U}$ through $(\mu, \nu) \in \mathfrak{t}^* \times \mathfrak{u}^*$. This will be done by analyzing each case separately.

Case I: $\nu_{13} \neq 0$. Then $t_{11} = t_{33}$, as was shown before. By Theorem 10.6.6 (see (10.6.27)) and (10.7.32), the elements $t \in (\mathcal{T}^\nu)_{\mu_\nu}$ are characterized by the condition

$$j_\nu^* \left[\Pi_\mathfrak{t} \left(t^T \mu_\nu (t^{-1})^T \right) + \nu - \Pi_\mathfrak{t} (u(t)^{-1})^T \nu u(t)^T \right] = \mu_\nu$$
$$= j_\nu^* \mu, \qquad (10.7.46)$$

where $u(t) \in \mathcal{U}$ is any element satisfying

$$\Pi_\mathfrak{u} t^T \nu (t^{-1})^T = \Pi_\mathfrak{u} (u(t)^{-1})^T \nu u(t)^T$$

and

$$j_\nu^* : \mathfrak{t}^* \to (\mathfrak{t}^\nu)^*$$

is given by (10.7.34).

The element $u(t) \in \mathcal{U}$ is chosen as follows. Since the non-zero entries of $\Pi_\mathfrak{u} \left(t^T \nu (t^{-1})^T \right)$ are (remember that $t_{11} = t_{33}$ since $t \in \mathcal{T}^\nu$)

$$\left[\Pi_\mathfrak{u} t^T \nu (t^{-1})^T \right]_{12} = \frac{t_{11}}{t_{22}} \nu_{12} - \frac{t_{23}}{t_{22}} \nu_{13}$$
$$\left[\Pi_\mathfrak{u} t^T \nu (t^{-1})^T \right]_{23} = \frac{t_{12}}{t_{11}} \nu_{13} + \frac{t_{22}}{t_{11}} \nu_{23}$$
$$\left[\Pi_\mathfrak{u} t^T \nu (t^{-1})^T \right]_{13} = \nu_{13}$$

and those of $\Pi_\mathfrak{u} (u(t)^{-1})^T \nu u(t)^T$ are

$$\left[\Pi_\mathfrak{u} (u(t)^{-1})^T \nu u(t)^T \right]_{12} = \nu_{12} + u(t)_{23} \nu_{13}$$
$$\left[\Pi_\mathfrak{u} (u(t)^{-1})^T \nu u(t)^T \right]_{23} = \nu_{23} - u(t)_{12} \nu_{13}$$
$$\left[\Pi_\mathfrak{u} (u(t)^{-1})^T \nu u(t)^T \right]_{13} = \nu_{13}$$

it follows that we can choose $u(t)$ by the requirement that

$$u(t)_{12} = \frac{t_{11} - t_{22}}{t_{11}} \frac{\nu_{23}}{\nu_{13}} - \frac{t_{12}}{t_{11}}$$
$$u(t)_{23} = \frac{t_{11} - t_{22}}{t_{22}} \frac{\nu_{12}}{\nu_{13}} - \frac{t_{23}}{t_{22}}$$

and $u(t)_{13}$ arbitrary.

Now we return to the relation (10.7.46). The non-zero entries of $\Pi_t\left(t^T\mu_\nu(t^{-1})^T\right)$ and $\Pi_t\left((u^{-1})^T\nu u^T\right)$ are

$$\left[\Pi_t\left(t^T\mu_\nu(t^{-1})^T\right)\right]_{11} = -\frac{t_{12}}{t_{22}}\mu_{12} + \frac{t_{12}t_{23}}{t_{11}t_{22}}\mu_{13} - \frac{t_{13}}{t_{11}}\mu_{13}$$

$$\left[\Pi_t\left(t^T\mu_\nu(t^{-1})^T\right)\right]_{22} = \mu_{22} + \frac{t_{12}}{t_{22}}\mu_{12} - \frac{t_{12}t_{23}}{t_{11}t_{22}}\mu_{13} - \frac{t_{23}}{t_{11}}\mu_{23}$$

$$\left[\Pi_t\left(t^T\mu_\nu(t^{-1})^T\right)\right]_{33} = \mu_{11} + \mu_{33} + \frac{t_{13}}{t_{11}}\mu_{13} + \frac{t_{23}}{t_{11}}\mu_{23}$$

$$\left[\Pi_t\left(t^T\mu_\nu(t^{-1})^T\right)\right]_{12} = \frac{t_{11}}{t_{22}}\mu_{12} - \frac{t_{23}}{t_{22}}\mu_{13}$$

$$\left[\Pi_t\left(t^T\mu_\nu(t^{-1})^T\right)\right]_{23} = \frac{t_{12}}{t_{11}}\mu_{13} + \frac{t_{22}}{t_{11}}\mu_{23}$$

$$\left[\Pi_t\left(t^T\mu_\nu(t^{-1})^T\right)\right]_{13} = \mu_{13}$$

$$\left[\Pi_t\left((u^{-1})^T\nu u^T\right)\right]_{11} = u_{12}\nu_{12} + u_{13}\nu_{13}$$

$$\left[\Pi_t\left((u^{-1})^T\nu u^T\right)\right]_{22} = -u_{12}\nu_{12} - u_{12}u_{23}\nu_{13} + u_{23}\nu_{23}$$

$$\left[\Pi_t\left((u^{-1})^T\nu u^T\right)\right]_{33} = -u_{13}\nu_{13} + u_{12}u_{23}\nu_{13} - u_{23}\nu_{23}$$

$$\left[\Pi_t\left((u^{-1})^T\nu u^T\right)\right]_{12} = \nu_{12} + u_{23}\nu_{13}$$

$$\left[\Pi_t\left((u^{-1})^T\nu u^T\right)\right]_{23} = \nu_{23} - u_{12}\nu_{13}$$

$$\left[\Pi_t\left((u^{-1})^T\nu u^T\right)\right]_{13} = \nu_{13}.$$

Using (10.7.34), we see that (10.7.46) is equivalent to

$$\frac{t_{12}}{t_{22}}\mu_{12} - \frac{t_{12}t_{23}}{t_{11}t_{22}}\mu_{13} - \frac{t_{23}}{t_{11}}\mu_{23} + \nu_{12}u(t)_{12} + \nu_{13}u(t)_{12}u(t)_{23} - \nu_{23}u(t)_{23} = 0$$

$$\frac{t_{11} - t_{22}}{t_{22}}\mu_{12} - \frac{t_{23}}{t_{22}} - u(t)_{23}\nu_{13} = 0$$

$$\frac{t_{12}}{t_{11}}\mu_{13} + u(t)_{12}\nu_{13} - \frac{t_{11} - t_{22}}{t_{11}}\mu_{23} = 0.$$

Plugging in here the values of $u(t)_{12}$ and $u(t)_{23}$ found before we get

$$\left.\begin{array}{l}
t_{11}t_{12}(\mu_{12} - \nu_{12}) - t_{12}t_{23}(\mu_{13} - \nu_{13}) - t_{22}t_{23}(\mu_{23} - \nu_{23}) = 0 \\[2ex]
(t_{11} - t_{22})(\mu_{12} - \nu_{12}) - t_{23}(\mu_{13} - \nu_{13}) = 0 \\[2ex]
t_{12}(\mu_{13} - \nu_{13}) - (t_{11} - t_{22})(\mu_{23} - \nu_{23}) = 0.
\end{array}\right\} \quad (10.7.47)$$

Case IA: If $\nu_{13} \neq 0$ and, in addition, $\mu_{13} - \nu_{13} \neq 0$, the system above gives

$$t_{12} = (t_{11} - t_{22})\frac{\mu_{23} - \nu_{23}}{\mu_{13} - \nu_{13}} \qquad \text{and} \qquad t_{23} = (t_{11} - t_{22})\frac{\mu_{12} - \nu_{12}}{\mu_{13} - \nu_{13}} \ .$$

With these values of t_{12}, t_{23} the first equation of (10.7.47) is identically satisfied. Therefore, if $\mu_{13} \neq \nu_{13}$, the isotropy group $(\mathcal{T}^{\nu})_{\mu_{\nu}}$ is the set of matrices given by

$$\left\{ \begin{bmatrix} t_{11} & (t_{11} - t_{22})\dfrac{\mu_{23} - \nu_{23}}{\mu_{13} - \nu_{13}} & t_{13} \\ 0 & t_{22} & (t_{11} - t_{22})\dfrac{\mu_{12} - \nu_{12}}{\mu_{13} - \nu_{13}} \\ 0 & 0 & t_{11} \end{bmatrix} \, \middle| \, t_{11}, t_{22} \neq 0 \right\}$$

which is three dimensional. Thus the reduced space

$$(\mathbf{J}_{\nu}^{R})^{-1}(\mu_{\nu})/(\mathcal{T}^{\nu})_{\mu_{\nu}},$$

which is diffeomorphic with the generic coadjoint orbit of $\mathcal{T} \, \text{\textcircled{s}} \, \mathcal{U}$ through a point (μ, ν) satisfying $\nu_{13} \neq 0$ and $\nu_{13} \neq \mu_{13}$, has dimension equal to

$$\dim(\mathbf{J}_{\nu}^{R})^{-1}(\mu_{\nu}) - \dim(\mathcal{T}^{\nu})_{\mu_{\nu}} = 9 - 3 = 6.$$

The $(\mathcal{T}^{\nu})_{\mu_{\nu}}$-action is the restriction to the level set $\left(\mathbf{J}_{\nu}^{R}\right)^{-1}(\mu_{\nu})$ of the action (10.7.30), that is,

$$\left(t, \kappa, \Pi_{\mathfrak{u}}\left((u^{-1})^{T}\nu u^{T}\right)\right) \cdot l = \left(tl, \kappa, \Pi_{\mathfrak{u}}\left(((l^{-1}u)^{-1})^{T}\nu(l^{-1}u)^{T}\right)\right).$$

Our next task is to realize this reduced space as a concrete submanifold of $\mathfrak{t}^{*} \times \mathfrak{u}^{*}$, which is the generic coadjoint orbit of $\mathcal{T} \, \text{\textcircled{s}} \, \mathcal{U}$ through (μ, ν) under the hypothesis that $\nu_{13} \neq 0$ and $\mu_{13} \neq \nu_{13}$.

To do this, we make use of the symplectic diffeomorphism

$$\left(\mathbf{J}_{\nu}^{R}\right)^{-1}(\mu_{\nu})/\left(\mathcal{T}^{\nu}\right)_{\mu_{\nu}} \to \mathcal{O}_{(\mu, \nu)}$$

that is given by formula (10.6.39) in Proposition 10.6.7; that is, it is has the expression

$$[t, \kappa, \rho] \mapsto \left(\kappa, \Pi_{\mathfrak{u}}\left((t^{-1})^{T}\rho t^{T}\right)\right) =: (\kappa, \upsilon),$$

where $\left(t, \kappa, \rho := \Pi_{\mathfrak{u}}\left((u^{-1})^{T}\nu u^{T}\right)\right)$ satisfy the systems (10.7.40) and (10.7.41).

Let us compute explicitly $\upsilon := \Pi_{\mathfrak{u}}\left((t^{-1})^{T}\rho t^{T}\right) \in \mathfrak{u}^{*}$. In the expressions

$$\upsilon_{12} = \left[\Pi_{\mathfrak{u}}\left((t^{-1})^{T}\rho t^{T}\right)\right]_{12} = \frac{t_{22}}{t_{11}}\rho_{12} + \frac{t_{23}}{t_{11}}\rho_{13}$$

$$\upsilon_{23} = \left[\Pi_{\mathfrak{u}}\left((t^{-1})^{T}\rho t^{T}\right)\right]_{23} = \frac{t_{33}}{t_{22}}\rho_{23} - \frac{t_{12}t_{33}}{t_{11}t_{22}}\rho_{13}$$

$$\upsilon_{13} = \left[\Pi_{\mathfrak{u}}\left((t^{-1})^{T}\rho t^{T}\right)\right]_{13} = \frac{t_{33}}{t_{11}}\rho_{13}$$

replace ρ by the values given in (10.7.41) to get

$$
\left.
\begin{aligned}
v_{12} &= \frac{t_{22}}{t_{11}} \left[\nu_{12} - \mu_{12} + \frac{t_{11}}{t_{22}} \kappa_{12} - \frac{t_{23}}{t_{22}} \mu_{13} \right] + \frac{t_{23}}{t_{11}} \nu_{13} \\[2ex]
&= \kappa_{12} - \frac{t_{22}}{t_{11}} (\mu_{12} - \nu_{12}) - \frac{t_{23}}{t_{11}} (\mu_{13} - \nu_{13}) \\[2ex]
v_{23} &= \frac{t_{33}}{t_{22}} \left[\nu_{23} - \mu_{23} + \frac{t_{12}}{t_{11}} \mu_{13} + \frac{t_{22}}{t_{33}} \kappa_{23} \right] - \frac{t_{12}t_{33}}{t_{11}t_{22}} \nu_{13} \\[2ex]
&= \kappa_{23} - \frac{t_{33}}{t_{22}} (\mu_{23} - \nu_{23}) + \frac{t_{12}t_{33}}{t_{11}t_{22}} (\mu_{13} - \nu_{13}) \\[2ex]
v_{13} &= \frac{t_{33}}{t_{11}} \nu_{13}.
\end{aligned}
\right\}
\qquad (10.7.48)
$$

We have obtained a very explicit description of the generic coadjoint orbit through $(\mu, \nu) \in \mathfrak{t}^* \times \mathfrak{u}^*$, generic meaning that $\nu_{13} \neq 0$ and $\mu_{13} \neq \nu_{13}$: *the orbit $\mathcal{O}_{(\mu,\nu)}$ is six dimensional and consists of pairs (κ, v) with $\kappa_{11}, \kappa_{22}, \kappa_{13}$ given by the first two and last equations in (10.7.40) and the non-zero entries of v given by (10.7.48), where $t \in \mathcal{T}$ is arbitrary.* This means that the coadjoint orbit $\mathcal{O}_{(\mu,\nu)}$ is explicitly described by all the pairs $(\kappa, v) \in \mathfrak{t}^* \times \mathfrak{u}^*$ subject to the following conditions, for any $t \in \mathcal{T}$ and

any $\kappa_{33}, \kappa_{12}, \kappa_{23} \in \mathbb{R}$:

$$
\begin{aligned}
\kappa_{11} &= -\kappa_{33} + \mu_{11} + \mu_{33} + \frac{t_{12}}{t_{22}}\kappa_{12} - \frac{t_{12}t_{23}}{t_{22}t_{11}}\mu_{13} - \frac{t_{23}}{t_{33}}\kappa_{23} \\
&\quad + \frac{\nu_{12}}{\nu_{13}}\left(\mu_{23} - \frac{t_{12}}{t_{11}}\mu_{13} - \frac{t_{22}}{t_{33}}\kappa_{23}\right) \\
&\quad - \frac{1}{\nu_{13}}\left(\mu_{23} - \frac{t_{12}}{t_{11}}\mu_{13} - \frac{t_{22}}{t_{33}}\kappa_{23}\right)\left(\mu_{12} - \frac{t_{11}}{t_{22}}\kappa_{12} + \frac{t_{23}}{t_{22}}\mu_{13}\right) \\
&\quad + \frac{\nu_{23}}{\nu_{13}}\left(\mu_{12} - \frac{t_{11}}{t_{22}}\kappa_{12} + \frac{t_{23}}{t_{22}}\mu_{13}\right) \\
\kappa_{22} &= \mu_{22} - \frac{t_{12}}{t_{22}}\kappa_{12} + \frac{t_{12}t_{23}}{t_{22}t_{11}}\mu_{13} + \frac{t_{23}}{t_{33}}\kappa_{23} \\
&\quad - \frac{\nu_{12}}{\nu_{13}}\left(\mu_{23} - \frac{t_{12}}{t_{11}}\mu_{13} - \frac{t_{22}}{t_{33}}\kappa_{23}\right) \\
&\quad + \frac{1}{\nu_{13}}\left(\mu_{23} - \frac{t_{12}}{t_{11}}\mu_{13} - \frac{t_{22}}{t_{33}}\kappa_{23}\right)\left(\mu_{12} - \frac{t_{11}}{t_{22}}\kappa_{12} + \frac{t_{23}}{t_{22}}\mu_{13}\right) \\
&\quad - \frac{\nu_{23}}{\nu_{13}}\left(\mu_{12} - \frac{t_{11}}{t_{22}}\kappa_{12} + \frac{t_{23}}{t_{22}}\mu_{13}\right) \\
\kappa_{13} &= \frac{t_{33}}{t_{11}}\mu_{13} \\
\upsilon_{12} &= \kappa_{12} - \frac{t_{22}}{t_{11}}\left(\mu_{12} - \nu_{12}\right) - \frac{t_{23}}{t_{11}}\left(\mu_{13} - \nu_{13}\right) \\
\upsilon_{23} &= \kappa_{23} - \frac{t_{33}}{t_{22}}\left(\mu_{23} - \nu_{23}\right) + \frac{t_{12}t_{33}}{t_{11}t_{22}}\left(\mu_{13} - \nu_{13}\right) \\
\upsilon_{13} &= \frac{t_{33}}{t_{11}}\nu_{13}
\end{aligned} \right\}
$$

$$(10.7.49)$$

for any $t \in \mathcal{T}$. Note that this characterization of the coadjoint orbit involves less parameters and equations than (10.7.12), which was the description of the general coadjoint orbit in the direct approach.

The system (10.7.49) recovers the Casimir functions of $(\mathfrak{t} \circledS \mathfrak{u})^*$ found earlier. Adding the first two equations in (10.7.49) we get

$$\kappa_{11} + \kappa_{22} + \kappa_{33} = \mu_{11} + \mu_{22} + \mu_{33}.$$

Dividing the third by the sixth equation in (10.7.49) we get

$$\frac{\kappa_{13}}{\upsilon_{13}} = \frac{\mu_{13}}{\nu_{13}}.$$

Thus, we have found again the global Casimir function $C(\kappa, \upsilon) = \kappa_{11} + \kappa_{22} + \kappa_{33}$ and the Casimir function $D(\kappa, \upsilon) = \kappa_{13}/\upsilon_{13}$ on the open Poisson submanifold characterized by $\upsilon_{13} \neq 0$. On this open set these two Casimir functions are functionally independent.

Case IB: $\nu_{13} \neq 0$, $\mu_{13} = \nu_{13}$. Then the system (10.7.47) becomes

$$
\left.
\begin{aligned}
t_{11}t_{12}(\mu_{12} - \nu_{12}) - t_{22}t_{23}(\mu_{23} - \nu_{23}) = 0 \\
(t_{11} - t_{22})(\mu_{12} - \nu_{12}) = 0 \\
(t_{11} - t_{22})(\mu_{23} - \nu_{23}) = 0.
\end{aligned}
\right\}
$$

If, in addition, at least one of $\mu_{12} - \nu_{12}$ or $\mu_{23} - \nu_{23}$ does not vanish, then $t_{11} = t_{22}$ and since $t \in \mathcal{T}^{\nu}$ we also have $t_{11} = t_{33}$. Thus, in this case, $t_{11} = t_{22} = t_{33}$ and the first equation implies that

- if $\mu_{12} \neq \nu_{12}$ then

$$
t_{12} = t_{23} \frac{\mu_{23} - \nu_{23}}{\mu_{12} - \nu_{12}}
$$

- if $\mu_{23} \neq \nu_{23}$ then

$$
t_{23} = t_{12} \frac{\mu_{12} - \nu_{12}}{\mu_{23} - \nu_{23}}.
$$

Assuming, for example, that $\mu_{12} \neq \nu_{12}$, the $\mathcal{T}^{\nu}$-affine action isotropy subgroup at μ_{ν} equals

$$
(\mathcal{T}^{\nu})_{\mu_{\nu}} = \left\{ \begin{bmatrix} t_{11} & t_{23}\frac{\mu_{23}-\nu_{23}}{\mu_{12}-\nu_{12}} & t_{13} \\ 0 & t_{11} & t_{23} \\ 0 & 0 & t_{11} \end{bmatrix} \middle| \, t_{11} \neq 0 \right\}
$$

which is three dimensional. Thus the reduced space $(\mathbf{J}_{\nu}^{R})^{-1}(\mu_{\nu})/(\mathcal{T}^{\nu})_{\mu_{\nu}}$, which is diffeomorphic to the non-generic six dimensional coadjoint orbit of $\mathcal{T}\,\circledS\,\mathcal{U}$ through a point (μ, ν) satisfying $\nu_{13} \neq 0$ and $\nu_{13} \neq \mu_{13}$, has dimension equal to $\dim(\mathbf{J}_{\nu}^{R})^{-1}(\mu_{\nu}) - \dim(\mathcal{T}^{\nu})_{\mu_{\nu}} = 9 - 3 = 6$.

The coadjoint orbit is realized as in the previous case by (10.7.49) with $\mu_{13} = \nu_{13}$ which implies that $\kappa_{13} = \upsilon_{13}$.

Case IC: $\nu_{13} \neq 0$, $\mu_{13} = \nu_{13}$, $\mu_{12} = \nu_{12}$, $\mu_{23} = \nu_{23}$. In this case we still have $t_{11} = t_{33}$, but the system (10.7.47) is identically satisfied. Thus, $(\mathcal{T}^{\nu})_{\mu_{\nu}} = \mathcal{T}^{\nu}$ which is five dimensional. Thus the reduced space $(\mathbf{J}_{\nu}^{R})^{-1}(\mu_{\nu})/(\mathcal{T}^{\nu})_{\mu_{\nu}}$, which is diffeomorphic to the coadjoint orbit of $\mathcal{T}\,\circledS\,\mathcal{U}$ through a point (μ, ν) satisfying $\nu_{13} \neq 0$, $\nu_{13} = \mu_{13}$, $\mu_{12} = \nu_{12}$, $\mu_{23} = \nu_{23}$, has dimension equal to $\dim(\mathbf{J}_{\nu}^{R})^{-1}(\mu_{\nu}) - \dim(\mathcal{T}^{\nu})_{\mu_{\nu}} = 9 - 5 = 4$. The coadjoint orbit is again explicitly described by (10.7.49) in which the conditions $\nu_{13} = \mu_{13}$, $\mu_{12} = \nu_{12}$, $\mu_{23} = \nu_{23}$ have been inserted. This implies that $\kappa_{13} = \upsilon_{13}$, $\kappa_{12} = \upsilon_{12}$, and $\kappa_{23} = \upsilon_{23}$. Note that in the last two equalities there are no parameters involved when looking at system (10.7.49); this is where the drop in dimension comes from.

So far we have recovered cases (i) and (ii) for the 6-dimensional orbits and case (i) for the 4- dimensional orbits, when looking at the classification done directly.

Case IIA: $\nu = 0$ and hence $\mathcal{O}_\nu = \{0\}$. Therefore the cocycle is zero, $\mathcal{T}^\nu = \mathcal{T}$, and $\mu_\nu = \mu$. Moreover, by Theorem 10.6.6 and (10.7.32), the elements $t \in (\mathcal{T}^\nu)_{\mu_\nu}$ are now characterized by the condition

$$\Pi_{\mathfrak{t}} \left(t^T \mu (t^{-1})^T \right) = \mu.$$

Using (10.7.3), this is equivalent to

$$\left. \begin{array}{l} \dfrac{t_{12}}{t_{11}} \mu_{12} + \dfrac{t_{13}}{t_{11}} \mu_{13} = 0 \\[2mm] \dfrac{t_{12}}{t_{11}} \mu_{12} + \dfrac{t_{12} t_{23}}{t_{11} t_{22}} \mu_{13} - \dfrac{t_{23}}{t_{22}} \mu_{23} = 0 \\[2mm] \dfrac{t_{12} t_{23}}{t_{11} t_{22}} \mu_{13} - \dfrac{t_{13}}{t_{11}} \mu_{13} - \dfrac{t_{23}}{t_{22}} \mu_{23} = 0 \\[2mm] \mu_{12} = \dfrac{t_{22}}{t_{11}} \mu_{12} + \dfrac{t_{23}}{t_{11}} \mu_{13} \\[2mm] \mu_{23} = -\dfrac{t_{12} t_{33}}{t_{11} t_{22}} \mu_{13} + \dfrac{t_{33}}{t_{22}} \mu_{23} \\[2mm] \mu_{13} = \dfrac{t_{33}}{t_{11}} \mu_{13}. \end{array} \right\} \qquad (10.7.50)$$

Case IIAα: $\nu = 0$, $\mu_{13} \neq 0$. Then $t_{33} = t_{11}$ and the system (10.7.50) becomes

$$t_{12} \mu_{12} + t_{13} \mu_{13} = 0$$
$$t_{12} t_{22} \mu_{12} + t_{12} t_{23} \mu_{13} - t_{11} t_{23} \mu_{23} = 0$$
$$t_{12} t_{23} \mu_{13} - t_{22} t_{13} \mu_{13} - t_{11} t_{23} \mu_{23} = 0$$
$$(t_{11} - t_{22}) \mu_{12} - t_{23} \mu_{13} = 0$$
$$(t_{11} - t_{22}) \mu_{23} - t_{12} \mu_{13} = 0.$$

From the last two and the first equations we get

$$t_{12} = (t_{11} - t_{22}) \frac{\mu_{23}}{\mu_{13}}, \quad t_{23} = (t_{11} - t_{22}) \frac{\mu_{12}}{\mu_{13}}, \quad t_{13} = -(t_{11} - t_{22}) \frac{\mu_{12} \mu_{23}}{\mu_{13}^2} .$$

With these values the second and third equation are identically satisfied. Thus, $\mathcal{T}_\mu$ consists of all $t \in \mathcal{T}$ such that $t_{11} = t_{33}$ and t_{12}, t_{23}, t_{13} are given above. Thus $\dim \mathcal{T}_\mu = 2$ and hence the coadjoint orbit has dimension $6 - 2 = 4$.

The explicit description of the orbit is given by the range of the map (10.6.39). Since in this case $\nu = 0$, follows that $\rho := \Pi_{\mathfrak{u}} \left((u^{-1})^T \nu u^T \right) = 0$ and thus $\upsilon := \Pi_{\mathfrak{u}} \left((t^{-1})^T \rho t^T \right) = 0$. Therefore, the coadjoint orbit consists in this case of points $(\kappa, 0)$ such that κ satisfies the system (10.7.42) for any $t \in \mathcal{T}$.

This recovers case (ii) for the 4-dimensional orbits in the classification done directly.

Case IIAβ: $\nu = 0$, $\mu_{13} = 0$. The system (10.7.50) simplifies to

$$
\left.
\begin{aligned}
t_{12}\mu_{12} &= 0 \\
\frac{t_{12}}{t_{11}}\mu_{12} - \frac{t_{23}}{t_{22}}\mu_{23} &= 0 \\
t_{23}\mu_{23} &= 0 \\
\mu_{12} &= \frac{t_{22}}{t_{11}}\mu_{12} \\
\mu_{23} &= \frac{t_{33}}{t_{22}}\mu_{23}.
\end{aligned}
\right\}
\tag{10.7.51}
$$

- If both $\mu_{12} \neq 0$ and $\mu_{23} \neq 0$ then $t_{12} = t_{23} = 0$ and $t_{11} = t_{22} = t_{33}$. Thus

$$
\mathcal{T}_\mu = \left\{
\begin{bmatrix}
t_{11} & 0 & t_{13} \\
0 & t_{11} & 0 \\
0 & 0 & t_{11}
\end{bmatrix}
\;\middle|\; t_{11} \neq 0
\right\}
$$

which is two dimensional. Thus the reduced space $(\mathbf{J}_\nu^R)^{-1}(\mu)/\mathcal{T}_\mu$ has dimension equal to $\dim(\mathbf{J}_\nu^R)^{-1}(\mu) - \dim \mathcal{T}_\mu = 6 - 2 = 4$. This reduced space is symplectically diffeomorphic to the four dimensional coadjoint orbit through $(\mu, 0)$, where $\mu_{13} = 0$, $\mu_{12} \neq 0$, and $\mu_{23} \neq 0$. This is realized as a subset of $\mathfrak{t}^* \times \mathfrak{u}^*$ by using the symplectic diffeomorphism $\left(\mathbf{J}_\nu^R\right)^{-1}(\mu_\nu)/\left(\mathcal{T}^\nu\right)_{\mu_\nu} \to \mathcal{O}_{(\mu,\nu)}$ given by (10.6.39), which in this case is

$$
[t, \kappa, 0] \mapsto (\kappa, 0),
$$

where κ satisfies the system (10.7.42) with $\mu_{13} = 0$, that is,

$$
\left.
\begin{aligned}
\kappa_{11} &= \mu_{11} + \frac{t_{12}}{t_{11}}\mu_{12} \\
\kappa_{22} &= \mu_{22} - \frac{t_{12}}{t_{11}}\mu_{12} + \frac{t_{23}}{t_{22}}\mu_{23} \\
\kappa_{33} &= \mu_{33} - \frac{t_{23}}{t_{22}}\mu_{23} \\
\kappa_{12} &= \frac{t_{22}}{t_{11}}\mu_{12} \\
\kappa_{23} &= \frac{t_{33}}{t_{22}}\mu_{23} \\
\kappa_{13} &= 0.
\end{aligned}
\right\}
$$

From this description we see again that the orbit is four dimensional; it is parametrized by $t_{12}/t_{11}, t_{23}/t_{22}, t_{22}/t_{11}, t_{33}/t_{22} \in \mathbb{R}$.

This recovers case (vi) of the 4-dimensional orbits in the direct classification.

- If $\mu_{12} \neq 0$ and $\mu_{23} = 0$ system (10.7.51) implies that $t_{12} = 0$ and $t_{11} = t_{22}$. Thus

$$
\mathcal{T}_\mu = \left\{ \begin{bmatrix} t_{11} & 0 & t_{13} \\ 0 & t_{11} & t_{23} \\ 0 & 0 & t_{33} \end{bmatrix} \,\middle|\, t_{11}, t_{33} \neq 0 \right\}
$$

which is four dimensional. Thus the reduced space $(\mathbf{J}_\nu^R)^{-1}(\mu)/\mathcal{T}_\mu$ has dimension equal to $\dim(\mathbf{J}_\nu^R)^{-1}(\mu) - \dim \mathcal{T}_\mu = 6 - 4 = 2$. This reduced space is symplectically diffeomorphic to the two dimensional coadjoint orbit through $(\mu, 0)$, where $\mu_{13} = 0$, $\mu_{12} \neq 0$, and $\mu_{23} = 0$. This is realized as a subset of $\mathfrak{t}^* \times \mathfrak{u}^*$ by using the symplectic diffeomorphism $\left(\mathbf{J}_\nu^R\right)^{-1}(\mu_\nu)/\left(\mathcal{T}^\nu\right)_{\mu_\nu} \to \mathcal{O}_{(\mu,\nu)}$ given by (10.6.39), which in this case is

$$
[t, \kappa, 0] \mapsto (\kappa, 0),
$$

where κ satisfies the system (10.7.42) with $\mu_{13} = 0$ and $\mu_{23} = 0$, that is,

$$
\left.\begin{aligned}
\kappa_{11} &= \mu_{11} + \frac{t_{12}}{t_{11}}\mu_{12} \\
\kappa_{22} &= \mu_{22} - \frac{t_{12}}{t_{11}}\mu_{12} \\
\kappa_{33} &= \mu_{33} \\
\kappa_{12} &= \frac{t_{22}}{t_{11}}\mu_{12} \\
\kappa_{23} &= 0 \\
\kappa_{13} &= 0.
\end{aligned}\right\}
$$

From this description we see again that the orbit is two dimensional; it is parametrized by $t_{12}/t_{11}, t_{22}/t_{11} \in \mathbb{R}$.

This recovers case (iii) of the 2-dimensional orbits in the direct classification.

- The situation $\mu_{12} = 0$ and $\mu_{23} \neq 0$ is analogous to the previous case. Now system (10.7.51) implies that $t_{23} = 0$ and $t_{22} = t_{33}$ so that

$$
\mathcal{T}_\mu = \left\{ \begin{bmatrix} t_{11} & t_{12} & t_{13} \\ 0 & t_{22} & 0 \\ 0 & 0 & t_{22} \end{bmatrix} \,\middle|\, t_{11}, t_{22} \neq 0 \right\}
$$

which is four dimensional. The reduced space $\left(\mathbf{J}_\nu^R\right)^{-1}(\mu_\nu)/\left(\mathcal{T}^\nu\right)_{\mu_\nu}$, and hence the coadjoint orbit $\mathcal{O}_{(\mu,\nu)}$, is now two dimensional and described explicitly as a subset of $\mathfrak{t}^* \times \mathfrak{u}^*$ by the system (10.7.42)

with $\mu_{13} = 0$ and $\mu_{12} = 0$, that is,

$$\left.\begin{aligned}
\kappa_{11} &= \mu_{11} \\
\kappa_{22} &= \mu_{22} + \frac{t_{23}}{t_{22}}\mu_{23} \\
\kappa_{33} &= \mu_{33} - \frac{t_{23}}{t_{22}}\mu_{23} \\
\kappa_{12} &= 0 \\
\kappa_{23} &= \frac{t_{33}}{t_{22}}\mu_{23} \\
\kappa_{13} &= 0.
\end{aligned}\right\}$$

From this description we see again that the orbit is two dimensional; it is parametrized by $t_{23}/t_{22}, t_{33}/t_{22} \in \mathbb{R}$.

This recovers case (iv) of the 2-dimensional orbits in the direct classification.

- If $\mu_{12} = \mu_{23} = 0$ then the system $(10.7.51)$ is identically satisfied and hence $\mathcal{T}_\mu = \mathcal{T}$. The coadjoint orbit $\mathcal{O}_{(\mu,0)}$ is described as a subset of $\mathfrak{t}^* \times \mathfrak{u}^*$ by imposing in any of the systems of the previous three cases the conditions $\mu_{12} = \mu_{23} = 0$, which implies that $\mathcal{O}_{(\mu,0)} = \{(\mu,0)\}$ thereby recovering the 0-dimensional case in the direct classification.

In all the situations covered by the *Case IIA*, we have $\nu = 0$, $\mathcal{T}^\nu = \mathcal{T}$ and vanishing cocycle. Applying Proposition 10.6.8, we conclude that *all these $\mathcal{T}\,\circledS\,\mathcal{U}$-coadjoint orbits $\mathcal{O}_{(\mu,0)}$ are symplectically diffeomorphic to the $\mathcal{T}$-coadjoint orbit $\mathcal{O}_\mu$.*

Case IIB: $\nu_{13} = 0$, $\nu_{12} \neq 0$, $\nu_{23} = 0$. Then $t_{22} = t_{11}$, $\mathcal{O}_\nu = \{\nu\}$, and the seven dimensional level set $\left(\mathbf{J}_\nu^R\right)^{-1}(\mu_\nu)$ is given by triples $(t,\kappa,\nu) \in \mathcal{T} \times \mathfrak{t}^* \times \{\nu\}$ satisfying $(10.7.43)$ for arbitrary $t \in \mathcal{T}$ and $\kappa_{22} \in \mathbb{R}$.

By Theorem 10.6.6 and $(10.7.32)$, the elements $t \in (\mathcal{T}^\nu)_{\mu_\nu}$ are characterized by the condition $(10.7.46)$, that is,

$$j_\nu^* \left[\Pi_{\mathfrak{t}}\left(t^T \mu_\nu (t^{-1})^T\right) + \nu - \Pi_{\mathfrak{t}}(u(t)^{-1})^T \nu u(t)^T\right] = \mu_\nu = j_\nu^* \mu,$$

where $u(t) \in \mathcal{U}$ is any element satisfying $\Pi_{\mathfrak{u}} t^T \nu (t^{-1})^T = \Pi_{\mathfrak{u}}(u(t)^{-1})^T \nu u(t)^T$ and $j_\nu^* : \mathfrak{t}^* \to (\mathfrak{t}^\nu)^*$ is given by $(10.7.35)$.

The element $u(t) \in \mathcal{U}$ is chosen as follows. The non-zero entries of $\Pi_{\mathfrak{u}}\left(t^T \nu (t^{-1})^T\right)$ are

$$\left[\Pi_{\mathfrak{u}} t^T \nu (t^{-1})^T\right]_{12} = \frac{t_{11}}{t_{22}}\nu_{12} - \frac{t_{11}t_{23}}{t_{22}t_{33}}\nu_{13} = \nu_{12}$$

$$\left[\Pi_{\mathfrak{u}} t^T \nu (t^{-1})^T\right]_{23} = \frac{t_{22}}{t_{33}}\nu_{23} + \frac{t_{12}}{t_{33}}\nu_{13} = 0$$

$$\left[\Pi_{\mathfrak{u}} t^T \nu (t^{-1})^T\right]_{13} = \frac{t_{11}}{t_{33}}\nu_{13} = 0$$

since $\nu_{13} = \nu_{23} = 0$ and $t_{11} = t_{22}$ because $t \in \mathcal{T}^\nu$. Similarly, the non-zero entries of $\Pi_{\mathfrak{u}}(u(t)^{-1})^T \nu u(t)^T$ are

$$\left[\Pi_{\mathfrak{u}}(u(t)^{-1})^T \nu u(t)^T\right]_{12} = \nu_{12} + u_{23}\nu_{13} = \nu_{12}$$

$$\left[\Pi_{\mathfrak{u}}(u(t)^{-1})^T \nu u(t)^T\right]_{23} = \nu_{23} - u_{12}\nu_{13} = 0$$

$$\left[\Pi_{\mathfrak{u}}(u(t)^{-1})^T \nu u(t)^T\right]_{13} = \nu_{13} = 0.$$

It follows that we can choose $u(t)$ to be the identity matrix.

Therefore, condition (10.7.46) on $t \in \mathcal{T}^\nu$ becomes in this case

$$j_\nu^* \left[\Pi_{\mathfrak{t}}\left(t^T \mu_\nu (t^{-1})^T\right)\right] = \mu_\nu$$

where $j_\nu : \mathfrak{t}^* \to (\mathfrak{t}^\nu)^*$ is given by (10.7.35) and $\mu_\nu := j_\nu^*\mu$. Using (10.7.5) this is equivalent to

$$\left.\begin{array}{l} t_{13}\mu_{13} + t_{23}\mu_{23} = 0 \\[1.5ex] t_{23}\mu_{13} = 0 \\[1.5ex] (t_{11} - t_{33})\mu_{23} + t_{12}\mu_{13} = 0 \\[1.5ex] (t_{11} - t_{33})\mu_{13} = 0. \end{array}\right\} \qquad (10.7.52)$$

Case IIBα: $\nu_{13} = 0$, $\nu_{12} \neq 0$, $\nu_{23} = 0$, and $\mu_{13} \neq 0$. Then from system (10.7.52) we get $t_{11} = t_{22} = t_{33}$ and $t_{12} = t_{23} = t_{13} = 0$. Therefore,

$$(\mathcal{T}^\nu)_{\mu_\nu} = \left\{ \begin{bmatrix} t_{11} & 0 & 0 \\ 0 & t_{11} & 0 \\ 0 & 0 & t_{11} \end{bmatrix} \,\middle|\, t_{11} \neq 0 \right\}$$

which is one dimensional. The reduced space $\left(\mathbf{J}_\nu^R\right)^{-1}(\mu_\nu)/(\mathcal{T}^\nu)_{\mu_\nu}$ and hence the coadjoint orbit $\mathcal{O}_{(\mu,\nu)}$ is six dimensional. As a subset of $\mathfrak{t}^* \times \mathfrak{u}^*$ this

orbit is the range of the symplectic diffeomorphism $\left(\mathbf{J}_\nu^R\right)^{-1}(\mu_\nu)/\left(\mathcal{T}^\nu\right)_{\mu_\nu} \rightarrow$ $\mathcal{O}_{(\mu,\nu)}$ given by (10.6.39). In this case it is has the expression

$$[t,\kappa,\nu] \mapsto \left(\kappa, \Pi_{\mathfrak{u}}\left((t^{-1})^T \nu t^T\right)\right) =: (\kappa, \upsilon),$$

where κ satisfies the system (10.7.43). Thus, using (10.7.3) to compute υ, the orbit is described by

$$\left.\begin{aligned}
\kappa_{11} &= -\kappa_{22} + \mu_{11} + \mu_{22} + \frac{t_{13}}{t_{11}}\mu_{13} - \frac{t_{12}t_{23}}{t_{11}t_{22}}\mu_{13} + \frac{t_{23}}{t_{22}}\mu_{23}\\
\kappa_{33} &= \mu_{33} - \frac{t_{13}}{t_{11}}\mu_{13} + \frac{t_{12}t_{23}}{t_{11}t_{22}}\mu_{13} - \frac{t_{23}}{t_{22}}\mu_{23}\\
\kappa_{12} &= \frac{t_{22}}{t_{11}}\mu_{12} + \frac{t_{23}}{t_{11}}\mu_{13}\\
\kappa_{23} &= \frac{t_{33}}{t_{22}}\mu_{23} - \frac{t_{12}t_{33}}{t_{11}t_{22}}\mu_{13}\\
\kappa_{13} &= \frac{t_{33}}{t_{11}}\mu_{13}\\
\upsilon_{12} &= \frac{t_{22}}{t_{11}}\nu_{12}\\
\upsilon_{23} &= 0\\
\upsilon_{13} &= 0.
\end{aligned}\right\} \qquad (10.7.53)$$

These formulas show that indeed the coadjoint orbit $\mathcal{O}_{(\mu,\nu)}$ is six dimensional since it is completely described by the parameters $\kappa_{22}, t_{22}/t_{11}, t_{33}/t_{11}, t_{13}/t_{11}, t_{23}/t_{11}, t_{12}/t_{11}$. The other two coefficients can be expressed in terms of these six:

$$\frac{t_{12}}{t_{22}} = \frac{t_{12}}{t_{11}}\left(\frac{t_{22}}{t_{11}}\right)^{-1} \quad \text{and} \quad \frac{t_{23}}{t_{22}} = \frac{t_{23}}{t_{11}}\left(\frac{t_{22}}{t_{11}}\right)^{-1}.$$

This recovers the first third of the case (iii) of the 6-dimensional orbits in the direct classification.

Case IIBβ: $\nu_{13} = 0, \nu_{12} \neq 0, \nu_{23} = 0$, and $\mu_{13} = 0$. Then system (10.7.52) becomes

$$\left.\begin{aligned}
t_{23}\mu_{23} &= 0\\
(t_{11} - t_{22})\mu_{23} &= 0
\end{aligned}\right\}$$

and we distinguish two cases:

• If $\mu_{23} \neq 0$, then $t_{11} = t_{22} = t_{33}$ and $t_{23} = 0$ so that

$$(\mathcal{T}^\nu)_{\mu_\nu} = \left\{\begin{bmatrix} t_{11} & t_{12} & t_{13}\\ 0 & t_{11} & 0\\ 0 & 0 & t_{11}\end{bmatrix} \,\middle|\, t_{11} \neq 0\right\}$$

which is three dimensional. The reduced space $\left(\mathbf{J}_\nu^R\right)^{-1}(\mu_\nu)/(\mathcal{T}^\nu)_{\mu_\nu}$, and hence the coadjoint orbit $\mathcal{O}_{(\mu,\nu)}$, is four dimensional. Reasoning as in the previous case, as a subset of $\mathfrak{t}^* \times \mathfrak{u}^*$, this orbit is described by the system (10.7.53) with $\mu_{13} = 0$, that is,

$$
\left.
\begin{aligned}
\kappa_{11} &= -\kappa_{22} + \mu_{11} + \mu_{22} + \frac{t_{23}}{t_{22}}\mu_{23} \\[6pt]
\kappa_{33} &= \mu_{33} - \frac{t_{23}}{t_{22}}\mu_{23} \\[6pt]
\kappa_{12} &= \frac{t_{22}}{t_{11}}\mu_{12} \\[6pt]
\kappa_{23} &= \frac{t_{33}}{t_{22}}\mu_{23} \\[6pt]
\kappa_{13} &= 0 \\[6pt]
\upsilon_{12} &= \frac{t_{22}}{t_{11}}\nu_{12} \\[6pt]
\upsilon_{23} &= 0 \\[6pt]
\upsilon_{13} &= 0.
\end{aligned}
\right\}
$$

This description of the orbit confirms that it is 4-dimensional since the parameters that completely describe it are κ_{22}, t_{23}/t_{22}, t_{22}/t_{11}, and t_{33}/t_{22}.

This recovers the case (iv) of the 4-dimensional orbits in the direct classification.

- If $\mu_{23} = 0$ then the system (10.7.52) is identically satisfied and hence

$$
(\mathcal{T}^\nu)_{\mu_\nu} = \left\{ \begin{bmatrix} t_{11} & t_{12} & t_{13} \\ 0 & t_{11} & t_{23} \\ 0 & 0 & t_{33} \end{bmatrix} \,\middle|\, t_{11} \neq 0 \right\}
$$

which is five dimensional. The reduced space $\left(\mathbf{J}_\nu^R\right)^{-1}(\mu_\nu)/(\mathcal{T}^\nu)_{\mu_\nu}$, and hence the coadjoint orbit $\mathcal{O}_{(\mu,\nu)}$, is two dimensional. As a subset of $\mathfrak{t}^* \times \mathfrak{u}^*$, this orbit is described by the system in the previous case

with $\mu_{23} = 0$, that is,

$$\left.\begin{aligned}
\kappa_{11} &= -\kappa_{22} + \mu_{11} + \mu_{22} \\
\kappa_{33} &= \mu_{33} \\
\kappa_{12} &= \frac{t_{22}}{t_{11}}\mu_{12} \\
\kappa_{23} &= 0 \\
\kappa_{13} &= 0 \\
\upsilon_{12} &= \frac{t_{22}}{t_{11}}\nu_{12} \\
\upsilon_{23} &= 0 \\
\upsilon_{13} &= 0.
\end{aligned}\right\}$$

The parameters characterizing this orbit are κ_{22} and t_{22}/t_{11}.

This recovers the case (i) of the 2-dimensional orbits in the direct classification.

Case IIC: $\nu_{13} = 0$, $\nu_{12} = 0$, $\nu_{23} \neq 0$. Then $t_{33} = t_{22}$, $\mathcal{O}_\nu = \{\nu\}$, and the seven dimensional level set $\left(\mathbf{J}^R_\nu\right)^{-1}(\mu_\nu)$ is given by triples $(t, \kappa, \nu) \in \mathcal{T} \times \mathfrak{t}^* \times \{\nu\}$ satisfying (10.7.44) for arbitrary $t \in \mathcal{T}$ and $\kappa_{33} \in \mathbb{R}$.

By Theorem 10.6.6 and (10.7.32), the elements $t \in (\mathcal{T}^\nu)_{\mu_\nu}$ are characterized by the condition (10.7.46), that is,

$$j^*_\nu\left[\Pi_{\mathfrak{t}}\left(t^T\mu_\nu(t^{-1})^T\right) + \nu - \Pi_{\mathfrak{t}}(u(t)^{-1})^T\nu u(t)^T\right] = \mu_\nu = j^*_\nu\mu,$$

where $u(t) \in \mathcal{U}$ is any element satisfying $\Pi_{\mathfrak{u}}t^T\nu(t^{-1})^T = \Pi_{\mathfrak{u}}(u(t)^{-1})^T\nu u(t)^T$ and $j^*_\nu : \mathfrak{t}^* \to (\mathfrak{t}^\nu)^*$ is given by (10.7.36). As in *Case 2B* we can take $u(t) = I$ so that $t \in (\mathcal{T}^\nu)_{\mu_\nu}$ if and only if

$$j^*_\nu\left[\Pi_{\mathfrak{t}}\left(t^T\mu_\nu(t^{-1})^T\right)\right] = \mu_\nu,$$

where $\mu_\nu := j^*_\nu\mu$. Using (10.7.5) this is equivalent to

$$\left.\begin{aligned}
t_{12}t_{22}\mu_{12} - t_{12}t_{23}\mu_{13} + t_{13}t_{22}\mu_{13} &= 0 \\
t_{22}(t_{11} - t_{22})\mu_{12} - t_{11}t_{23}\mu_{13} &= 0 \\
t_{12}\mu_{13} &= 0 \\
(t_{11} - t_{22})\mu_{13} &= 0.
\end{aligned}\right\} \tag{10.7.54}$$

Case IICα: $\nu_{13} = 0$, $\nu_{12} = 0$, $\nu_{23} \neq 0$, and $\mu_{13} \neq 0$. Then system (10.7.54) implies that $t_{11} = t_{22} = t_{33}$ and $t_{12} = t_{23} = t_{13} = 0$. Therefore,

$$(\mathcal{T}^\nu)_{\mu_\nu} = \left\{\begin{bmatrix} t_{11} & 0 & 0 \\ 0 & t_{11} & 0 \\ 0 & 0 & t_{11} \end{bmatrix} \,\middle|\, t_{11} \neq 0\right\}$$

which is one dimensional. The reduced space $\left(\mathbf{J}_\nu^R\right)^{-1}(\mu_\nu)/\left(\mathcal{T}^\nu\right)_{\mu_\nu}$, and hence the coadjoint orbit $\mathcal{O}_{(\mu,\nu)}$, is six dimensional. As a subset of the dual $\mathfrak{t}^* \times \mathfrak{u}^*$ this orbit equals the range of the symplectic diffeomorphism $\left(\mathbf{J}_\nu^R\right)^{-1}(\mu_\nu)/\left(\mathcal{T}^\nu\right)_{\mu_\nu} \to \mathcal{O}_{(\mu,\nu)}$ given by (10.6.39). In this case it is has the expression

$$[t, \kappa, \nu] \mapsto \left(\kappa, \Pi_\mathfrak{u}\left((t^{-1})^T \nu t^T\right)\right) =: (\kappa, \upsilon),$$

where κ satisfies the system (10.7.44). Thus, using (10.7.3) to compute υ, the orbit is described by

$$
\left.
\begin{aligned}
\kappa_{11} &= \mu_{11} + \frac{t_{12}}{t_{11}}\mu_{12} + \frac{t_{13}}{t_{11}}\mu_{13} \\
\kappa_{22} &= -\kappa_{33} + \mu_{22} + \mu_{33} - \frac{t_{12}}{t_{11}}\mu_{12} - \frac{t_{13}}{t_{11}}\mu_{13} \\
\kappa_{12} &= \frac{t_{22}}{t_{11}}\mu_{12} + \frac{t_{23}}{t_{11}}\mu_{13} \\
\kappa_{23} &= \frac{t_{33}}{t_{22}}\mu_{23} - \frac{t_{12}t_{33}}{t_{11}t_{22}}\mu_{13} \\
\kappa_{13} &= \frac{t_{33}}{t_{11}}\mu_{13} \\
\upsilon_{12} &= 0 \\
\upsilon_{23} &= \frac{t_{33}}{t_{22}}\nu_{23} \\
\upsilon_{13} &= 0.
\end{aligned}
\right\} \tag{10.7.55}
$$

The six parameters characterizing this orbit are $\kappa_{33}, t_{33}/t_{11}, t_{22}/t_{11}, t_{12}/t_{11}$, t_{13}/t_{11}, and t_{23}/t_{11}. The other two coefficients can be expressed in terms of these six:

$$\frac{t_{12}}{t_{22}} = \frac{t_{12}}{t_{11}}\left(\frac{t_{22}}{t_{11}}\right)^{-1} \quad \text{and} \quad \frac{t_{33}}{t_{22}} = \frac{t_{33}}{t_{11}}\left(\frac{t_{22}}{t_{11}}\right)^{-1}.$$

This recovers the second third of the case (iii) of the 6-dimensional orbits in the direct classification.

Case IICβ: $\nu_{13} = 0$, $\nu_{12} = 0$, $\nu_{23} \neq 0$, and $\mu_{13} = 0$. Then system (10.7.54) becomes

$$
\left.
\begin{aligned}
t_{12}\mu_{12} &= 0 \\
(t_{11} - t_{22})\mu_{12} &= 0
\end{aligned}
\right\}
$$

and we distinguish two cases:

- If $\mu_{12} \neq 0$, then $t_{11} = t_{22} = t_{33}$ and $t_{12} = 0$ so that

$$\left(\mathcal{T}^\nu\right)_{\mu_\nu} = \left\{ \begin{bmatrix} t_{11} & 0 & t_{13} \\ 0 & t_{11} & t_{23} \\ 0 & 0 & t_{11} \end{bmatrix} \,\middle|\, t_{11} \neq 0 \right\}$$

which is three dimensional. The reduced space $\left(\mathbf{J}_\nu^R\right)^{-1}(\mu_\nu)/\left(\mathcal{T}^\nu\right)_{\mu_\nu}$, and hence the coadjoint orbit $\mathcal{O}_{(\mu,\nu)}$, is four dimensional. As a subset of $\mathfrak{t}^* \times \mathfrak{u}^*$, this orbit is described by the system (10.7.55) with the additional condition $\mu_{13} = 0$, that is,

$$
\left.
\begin{aligned}
\kappa_{11} &= \mu_{11} + \frac{t_{12}}{t_{11}}\mu_{12} \\[2mm]
\kappa_{22} &= -\kappa_{33} + \mu_{22} + \mu_{33} - \frac{t_{12}}{t_{11}}\mu_{12} \\[2mm]
\kappa_{12} &= \frac{t_{22}}{t_{11}}\mu_{12} \\[2mm]
\kappa_{23} &= \frac{t_{33}}{t_{22}}\mu_{23} \\[2mm]
\kappa_{13} &= 0 \\[1mm]
\upsilon_{12} &= 0 \\[1mm]
\upsilon_{23} &= \frac{t_{33}}{t_{22}}\nu_{23} \\[2mm]
\upsilon_{13} &= 0.
\end{aligned}
\right\}
$$

The parameters characterizing this orbit are κ_{33}, t_{12}/t_{11}, t_{22}/t_{11}, and t_{33}/t_{22}.

This recovers the case (v) of the 4-dimensional orbits in the direct classification.

- If $\mu_{12} = 0$, then there are no additional conditions on t and hence

$$
\left(\mathcal{T}^\nu\right)_{\mu_\nu} = \left\{
\begin{bmatrix}
t_{11} & t_{12} & t_{13} \\
0 & t_{22} & t_{23} \\
0 & 0 & t_{22}
\end{bmatrix}
\;\middle|\; t_{11}, t_{22} \neq 0
\right\}
$$

which is five dimensional. The reduced space $\left(\mathbf{J}_\nu^R\right)^{-1}(\mu_\nu)/\left(\mathcal{T}^\nu\right)_{\mu_\nu}$, and hence the coadjoint orbit $\mathcal{O}_{(\mu,\nu)}$, is two dimensional. As a subset of $\mathfrak{t}^* \times \mathfrak{u}^*$, this orbit is described by the system in the previous case with the additional condition $\mu_{12} = 0$, that is,

$$
\left.
\begin{aligned}
\kappa_{11} &= \mu_{11} \\[1mm]
\kappa_{22} &= -\kappa_{33} + \mu_{22} + \mu_{33} \\[1mm]
\kappa_{12} &= 0 \\[1mm]
\kappa_{23} &= \frac{t_{33}}{t_{22}}\mu_{23} \\[2mm]
\kappa_{13} &= 0 \\[1mm]
\upsilon_{12} &= 0 \\[1mm]
\upsilon_{23} &= \frac{t_{33}}{t_{22}}\nu_{23} \\[2mm]
\upsilon_{13} &= 0.
\end{aligned}
\right\}
$$

The two parameters on this orbit are κ_{33} and t_{33}/t_{22}.

This recovers the case (ii) of the 2-dimensional orbits in the direct classification.

Case IID: $\nu_{13} = 0$, $\nu_{12} \neq 0$, $\nu_{23} \neq 0$. Then $t_{11} = t_{22} = t_{33}$ (and hence $\mathcal{T}^{\nu}$ is four dimensional), $\mathcal{O}_{\nu} = \{\nu\}$, $j_{\nu}^* : \mathfrak{t}^* \to (\mathfrak{t}^{\nu})^*$ is given by (10.7.37) and hence

$$\mu_{\nu} := j_{\nu}^* \mu = \begin{bmatrix} \frac{\mu_{11}+\mu_{22}+\mu_{33}}{3} & \mu_{12} & \mu_{13} \\ 0 & \frac{\mu_{11}+\mu_{22}+\mu_{33}}{3} & \mu_{23} \\ 0 & 0 & \frac{\mu_{11}+\mu_{22}+\mu_{33}}{3} \end{bmatrix},$$

and the 8-dimensional level set $\left(\mathbf{J}_{\nu}^R\right)^{-1}(\mu_{\nu})$ consists of triples $(t, \kappa, \nu) \in \mathcal{T} \times \mathfrak{t}^* \times \{\nu\}$ such that κ satisfies the system (10.7.45) for $t \in \mathcal{T}$ and $\kappa_{22}, \kappa_{33} \in \mathbb{R}$ arbitrary.

As in *Case IIB* or *Case IIC*, $t \in (\mathcal{T}^{\nu})_{\mu_{\nu}}$ if and only if $j_{\nu}^* \left[\Pi_{\mathfrak{t}} \left(t^T \mu_{\nu} (t^{-1})^T \right) \right] = \mu_{\nu}$ which is equivalent by (10.7.5) to

$$\left. \begin{array}{l} t_{23}\mu_{13} = 0 \\ t_{12}\mu_{13} = 0. \end{array} \right\} \tag{10.7.56}$$

Case IIDα: If, in addition, $\mu_{13} \neq 0$ then $t_{23} = t_{12} = 0$ and hence

$$(\mathcal{T}^{\nu})_{\mu_{\nu}} = \left\{ \begin{bmatrix} t_{11} & 0 & t_{13} \\ 0 & t_{11} & 0 \\ 0 & 0 & t_{11} \end{bmatrix} \,\middle|\, t_{11} \neq 0 \right\}$$

which is two dimensional. The reduced space $\left(\mathbf{J}_{\nu}^R\right)^{-1}(\mu_{\nu})/(\mathcal{T}^{\nu})_{\mu_{\nu}}$, and hence the coadjoint orbit $\mathcal{O}_{(\mu,\nu)}$, is six dimensional. As a subset of $\mathfrak{t}^* \times \mathfrak{u}^*$, this orbit is identical to the range of the symplectic diffeomorphism $\left(\mathbf{J}_{\nu}^R\right)^{-1}(\mu_{\nu})/(\mathcal{T}^{\nu})_{\mu_{\nu}} \to \mathcal{O}_{(\mu,\nu)}$ given by (10.6.39). In this case it is has the expression

$$[t, \kappa, \nu] \mapsto \left(\kappa, \Pi_{\mathfrak{u}} \left((t^{-1})^T \nu t^T \right) \right) =: (\kappa, \upsilon),$$

where κ satisfies the system (10.7.45). Thus, using (10.7.3) to compute υ, the orbit is described by

$$
\left.
\begin{aligned}
\kappa_{11} &= -\kappa_{22} - \kappa_{33} + \mu_{11} + \mu_{22} + \mu_{33}\\[4pt]
\kappa_{12} &= \frac{t_{22}}{t_{11}}\mu_{12} + \frac{t_{23}}{t_{11}}\mu_{13}\\[4pt]
\kappa_{23} &= \frac{t_{33}}{t_{22}}\mu_{23} - \frac{t_{12}t_{33}}{t_{11}t_{22}}\mu_{13}\\[4pt]
\kappa_{13} &= \frac{t_{33}}{t_{11}}\mu_{13}\\[4pt]
\upsilon_{12} &= \frac{t_{22}}{t_{11}}\nu_{12}\\[4pt]
\upsilon_{23} &= \frac{t_{33}}{t_{22}}\nu_{23}\\[4pt]
\upsilon_{13} &= 0.
\end{aligned}
\right\}
\qquad (10.7.57)
$$

The six parameters of this orbit description are κ_{22}, κ_{33}, t_{22}/t_{11}, t_{33}/t_{11}, t_{12}/t_{11}, t_{23}/t_{11}. There is one more coefficient that appears in this system, namely t_{33}/t_{22} which is, however, expressible in terms of the other parameters:

$$
\frac{t_{33}}{t_{22}} = \frac{t_{33}}{t_{11}} = \frac{t_{33}}{t_{11}}\left(\frac{t_{22}}{t_{11}}\right)^{-1}.
$$

This recovers the last third of case (iii) of the 6-dimensional orbits in the direct classification.

Case IIDβ: If, in addition, $\mu_{13} = 0$ then there are no supplementary conditions on $t \in \mathcal{T}^{\nu}$ and hence

$$
(\mathcal{T}^{\nu})_{\mu_{\nu}} = \left\{
\begin{bmatrix}
t_{11} & t_{12} & t_{13}\\
0 & t_{11} & t_{23}\\
0 & 0 & t_{11}
\end{bmatrix}
\;\middle|\; t_{11} \neq 0
\right\}
$$

which is four dimensional. The reduced space $\left(\mathbf{J}_{\nu}^{R}\right)^{-1}(\mu_{\nu})/(\mathcal{T}^{\nu})_{\mu_{\nu}}$, and hence the coadjoint orbit $\mathcal{O}_{(\mu,\nu)}$, is four dimensional. As a subset of $\mathfrak{t}^{*} \times \mathfrak{u}^{*}$, this orbit is described by the system (10.7.57) with the condition $\mu_{13} = 0$

imposed, that is, we have

$$
\left.
\begin{aligned}
\kappa_{11} &= -\kappa_{22} - \kappa_{33} + \mu_{11} + \mu_{22} + \mu_{33} \\
\kappa_{12} &= \frac{t_{22}}{t_{11}}\mu_{12} \\
\kappa_{23} &= \frac{t_{33}}{t_{22}}\mu_{23} \\
\kappa_{13} &= 0 \\
\upsilon_{12} &= \frac{t_{22}}{t_{11}}\nu_{12} \\
\upsilon_{23} &= \frac{t_{33}}{t_{22}}\nu_{23} \\
\upsilon_{13} &= 0.
\end{aligned}
\right\}
\tag{10.7.58}
$$

The four parameters of this orbit description are κ_{22}, κ_{33}, t_{22}/t_{11}, t_{33}/t_{11}. As before, t_{33}/t_{22} is the product of t_{33}/t_{11} and the inverse of t_{22}/t_{11}.

This recovers the last remaining case (v) of the 4-dimensional orbits in the direct classification.

Summary of The Classification and Comparison.   We close this section with a table that summarizes the relationship between the direct method of classification and the one done by the method of reduction by stages. In the left column is the classification obtained by analyzing directly the $\mathcal{T}\circledS\mathcal{U}$-coadjoint orbits and on the right are the labels associated to the classification by the method of reduction by stages.

6-dimensional orbits:

(i) $\longleftrightarrow$ *Case IA*

(ii) $\longleftrightarrow$ *Case IB*

(iii) $\longleftrightarrow$ *Case IIBα, Case IICα,* and *Case IIDα*

4-dimensional orbits:

(i) $\longleftrightarrow$ *Case IC*

(ii) $\longleftrightarrow$ *Case IIAα*

(iii) $\longleftrightarrow$ *Case IIDβ*

(iv) $\longleftrightarrow$ first situation in *Case IIBβ*

(v) $\longleftrightarrow$ first situation in *Case IICβ*

(vi) $\longleftrightarrow$ first situation in *Case IIAβ*

2-dimensional orbits:

(i) $\longleftrightarrow$ second situation in *Case IIBβ*

(ii) $\longleftrightarrow$ second situation in *Case IICβ*

(iii) $\longleftrightarrow$ second situation in *Case IIAβ*

(iv) $\longleftrightarrow$ third situation in *Case IIAβ*

0-dimensional orbits $\longleftrightarrow$ fourth situation in *Case IIAβ*.

In all five situations associated to the *Case IIA*, we have $\nu = 0$ and the $\mathcal{T} \circledS \mathcal{U}$-coadjoint orbit $\mathcal{O}_{(\mu,0)}$ is symplectically diffeomorphic to the $\mathcal{T}$-coadjoint orbit $\mathcal{O}_\mu$. This result is obtained by applying the general theory of reduction by stages and Proposition 10.6.8; it is not obvious in the direct classification.

The same is true about the explicit description of the coadjoint orbits $\mathcal{O}_{(\mu,\nu)}$ as subsets of $\mathfrak{t}^* \times \mathfrak{u}^*$. *The reduction by stages program automatically provides an explicit description in each case (and these were written down in each situation).* The direct classification does not provide such a description and it is very difficult to differentiate between the various cases from the general one obtained from the formula of the coadjoint action.

11

Reduction by Stages via Symplectic Distributions

In the remaining two chapters of this part we are going to take a different technical approach to the reduction by stages problem. It will be mainly based on thinking of the (connected components of the) level sets of the various momentum maps in the setup as the accessible sets of a distribution that we will introduce in the following paragraphs. This point of view has been exploited in Ortega [2002] and Ortega and Ratiu [2002, 2004a] in the context of the so-called *optimal momentum map*.

In this chapter we prove a stages theorem using this approach that forces us to introduce all the technical background on distributions that will be needed in the sequel. The result that we prove is a version of Theorem 5.2.9 that only sees the connected components of the momentum level sets involved. This leads to a modified stages hypothesis. It should be noticed that from the dynamical point of view it is more appropriate to work with these connected spaces since they are preserved by invariant dynamics and are setwise smaller than those in Theorem 5.2.9.

Unlike the Stages Hypothesis I (see Definition 5.2.8) that is purely algebraic in nature, the hypothesis that we will invoke in this chapter involves also topological conditions. As we shall see in Chapter 12 this represents a bridge to a third version of the Reduction by Stages Theorem that requires only topological conditions.

11.1   Reduction by Stages of Connected Components

In our first reduction by stages theorem, we proved that the reduced space $P_\sigma := \mathbf{J}_M^{-1}(\sigma)/M_\sigma$ could be obtained as the end result of two consecutive symplectic reductions performed with the help of a closed normal subgroup N of M. We will show that the same process can be carried out for the smaller symplectic reduced spaces

$$P_\sigma^C := \mathbf{J}_M^{-1}(\sigma)_C/M_\sigma^C, \qquad\qquad (11.1.1)$$

where $\mathbf{J}_M^{-1}(\sigma)_C$ is a connected component of $\mathbf{J}_M^{-1}(\sigma)$, and M_σ^C is the subgroup of M_σ that leaves it invariant, more specifically:

$$M_\sigma^C := \{g \in M_\sigma \mid \Phi_g(z) \in \mathbf{J}_M^{-1}(\sigma)_C \ \text{ for all } \ z \in \mathbf{J}_M^{-1}(\sigma)_C\}.$$

In carrying this out we will have to adapt the Stages Hypothesis I to our new goal. Even though the Stages Theorem I and the one that we are going to prove are closely related, they are not comparable as far as their generality is concerned. The same is true for the two corresponding stages hypotheses.

The General Setup.   Our setup will be identical to the one in §5.2 except for the fact that in all that follows we will assume that the topology of the symplectic manifold P induced by its smooth structure is Lindelöf (every open covering of P has a countable subcovering).

   We recall that any closed subspace of a Lindelöf space is Lindelöf and that if $f : X \to Y$ is a continuous map with X Lindelöf then the subspace $f(X)$ of Y is also Lindelöf. These properties are particularly important in dealing with the spaces that we will encounter in the sequel since all of them will be either closed submanifolds of P or regular quotients of these closed subspaces by a free proper group action. Consequently, the Lindelöf hypothesis on P will ensure the same topological character for all the spaces that we will construct out of it.

   The Lindelöf hypothesis is imposed from now on because of the following Lemma. We provide its proof since it is not readily available in the literature.[1]

11.1.1 Lemma.   *Let $f : P \to Q$ be a smooth bijective immersion between the manifolds P and Q. If P is Lindelöf or paracompact then f is a diffeomorphism.*

Proof.   It suffices to prove that f is a local diffeomorphism, which in turn is implied by the fact that $\dim P = \dim Q$. The immersivity hypotheses implies that $\dim P \leq \dim Q$. Assume, by contradiction, that $\dim P < \dim Q$.

---

[1] The proof of the following Lemma was kindly supplied by T. Schmah.

By Sard's theorem, the set of critical values has dense complement in Q. But if $\dim P < \dim Q$, every point is critical, that is, $f(P)$ has measure zero in Q. In particular, there are points in Q that are not in $f(P)$, which contradicts the bijectivity of f. If P is paracompact, the same proof works by replacing the density argument by the fact that the set of critical values has measure zero in Q.   ∎

Note that, as a set, M_σ^C is closed in M_σ and contains the connected component of the identity of M_σ. It is a subgroup of M_σ since it is clearly closed under the group operation and also, if $g \in M_\sigma^C$, then it follows that $g^{-1} \in M_\sigma^C$; indeed, $g \in M_\sigma^C$ implies that $\Phi_g(\mathbf{J}_M^{-1}(\sigma)_C) \subset \mathbf{J}_M^{-1}(\sigma)_C$. Since the mapping Φ_g is a diffeomorphism, the set $\Phi_g(\mathbf{J}_M^{-1}(\sigma)_C)$ is open and closed in $\mathbf{J}_M^{-1}(\sigma)_C$ and therefore $\Phi_g(\mathbf{J}_M^{-1}(\sigma)_C) = \mathbf{J}_M^{-1}(\sigma)_C$. Applying $\Phi_{g^{-1}}$ to both sides of the previous equality implies that $g^{-1} \in M_\sigma^C$ which proves that M_σ^C is a subgroup of M_σ. The subgroup M_σ^C is clearly closed and is therefore a Lie subgroup of M_σ.

The Two-Stage Reduction Procedure.   We will spell out in more detail how the two-stage reduction procedure is carried out in the connected components framework. Recall that the M–equivariant momentum map for the action of the group N on P is given by $\mathbf{J}_N(z) = i^*(\mathbf{J}_M(z))$, where $i^* : \mathfrak{m}^* \to \mathfrak{n}^*$ is the dual of the inclusion $i : \mathfrak{n} \to \mathfrak{m}$. Let $\sigma \in \mathfrak{m}^*$ and $\nu := i^*(\sigma) \in \mathfrak{n}^*$. In this case the **first symplectic reduced space** is

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

where $\mathbf{J}_N^{-1}(\nu)_C$ is a connected component of $\mathbf{J}_N^{-1}(\nu)$ and N_ν^C is the subgroup of N_ν that leaves it invariant. We will see later on that $\mathbf{J}_N^{-1}(\nu)_C$ fully contains a certain number of connected components $\mathbf{J}_M^{-1}(\sigma)_{C_i}$, $i \in \{1, \ldots, n\}$, of $\mathbf{J}_M^{-1}(\sigma)$. We will denote their union by $\mathbf{J}_M^{-1}(\sigma)_{C_\nu}$, that is,

$$\mathbf{J}_M^{-1}(\sigma)_{C_\nu} := \bigcup_{i \in \{1, \ldots, n\}} \mathbf{J}_M^{-1}(\sigma)_{C_i}, \quad \text{where} \quad \mathbf{J}_M^{-1}(\sigma)_{C_i} \subset \mathbf{J}_N^{-1}(\nu)_C.$$

$$(11.1.2)$$

The symbol $M_\sigma^{C_\nu}$ will denote the subgroup of M_σ that leaves $\mathbf{J}_M^{-1}(\sigma)_{C_\nu}$ invariant and $\pi_\sigma^{C_\nu} : \mathbf{J}_M^{-1}(\sigma)_{C_\nu} \to \mathbf{J}_M^{-1}(\sigma)_{C_\nu}/M_\sigma^{C_\nu}$ the canonical projection. Let now M_ν^C be the subgroup of M_ν, the isotropy subgroup of $\nu \in \mathfrak{n}^*$ for the action of M on $\mathfrak{n}^*$, that leaves $\mathbf{J}_N^{-1}(\nu)_C$ invariant. Since N_ν is a normal subgroup of M_ν, so is N_ν^C with respect to M_ν^C and hence the quotient M_ν^C / N_ν^C is a well-defined Lie group. It is easy to check that M_ν^C / N_ν^C acts symplectically on P_ν^C with a (non equivariant) momentum map $\mathbf{J}_\nu^C : P_\nu^C \to (\mathfrak{m}_\nu^C / \mathfrak{n}_\nu^C)^*$ given by

$$(s_\nu^C)^* \circ \mathbf{J}_\nu^C \circ \pi_\nu^C = (k_\nu^C)^* \circ \mathbf{J}_M \circ i_\nu^C - \bar{\nu} \qquad (11.1.3)$$

where

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

is the canonical projection,

$$s_\nu^C : \mathfrak{m}_\nu^C \to \mathfrak{m}_\nu^C / \mathfrak{n}_\nu^C$$

is the induced Lie algebra homomorphism,

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

is the inclusion,

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

is the projection,

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

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

$$\left\langle \mathbf{J}_\nu^C([z]), [\xi] \right\rangle = \langle \mathbf{J}_M(z), \xi \rangle - \langle \bar{\nu}, \xi \rangle \qquad (11.1.4)$$

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

Consider now the element $\rho \in (\mathfrak{m}_\nu^C / \mathfrak{n}_\nu^C)^*$ constructed out of $\sigma \in \mathfrak{m}^*$ and $\nu \in \mathfrak{n}^*$ by means of the relation

$$\langle \rho, [\xi] \rangle = \langle \sigma, \xi \rangle - \langle \bar{\nu}, \xi \rangle \,, \qquad (11.1.5)$$

for any $\xi \in \mathfrak{m}_\nu^C$. The element $\bar{\nu}$ is the same extension of $\nu|_{\mathfrak{n}_\nu^C}$ to $\mathfrak{m}_\nu^C$ that we used in the construction of the momentum map $\mathbf{J}_\nu^C$. Denote by $(M_\nu^C / N_\nu^C)_\rho$ the isotropy subgroup of ρ with respect to the affine action of M_ν^C / N_ν^C on the dual of its Lie algebra, defined with the help of the non equivariance cocycle of $\mathbf{J}_\nu^C$ (see Proposition 5.2.7). The element $\rho \in (\mathfrak{m}_\nu^C / \mathfrak{n}_\nu^C)^*$ has been constructed in such a way that there is a connected component $(\mathbf{J}_\nu^C)^{-1}(\rho)_C$ of the level set $(\mathbf{J}_\nu^C)^{-1}(\rho) \subset P_\nu^C$ such that

$$\pi_\nu^C(\mathbf{J}_M^{-1}(\sigma)_{C_\nu}) \subset (\mathbf{J}_\nu^C)^{-1}(\rho)_C. \qquad (11.1.6)$$

Let $(M_\nu^C / N_\nu^C)_\rho^C$ be the subgroup of $(M_\nu^C / N_\nu^C)_\rho$ that leaves $(\mathbf{J}_\nu^C)^{-1}(\rho)_C$ invariant and $M_{\nu,\rho}^C$ the subgroup of M_ν^C such that

$$M_{\nu,\rho}^C / N_\nu^C = (M_\nu^C / N_\nu^C)_\rho^C. \qquad (11.1.7)$$

At this point we can introduce what will be the main hypothesis in the reduction by stages theorem to be proved in this chapter:

11.1.2 Definition (Stages Hypothesis II).   *We say that $\sigma \in \mathfrak{m}^*$ satisfies the **stages hypothesis II** if for any other element $\sigma' \in \mathfrak{m}^*$ such that*

$$\sigma'|_{\mathfrak{n}} = \sigma|_{\mathfrak{n}} =: \nu \quad and \quad \sigma'|_{\mathfrak{m}_\nu^C} = \sigma|_{\mathfrak{m}_\nu^C},$$

there exists $n \in M_{\nu,\rho}^C$ such that $\sigma' = \mathrm{Ad}_{n^{-1}}^* \sigma$, where ρ is defined in (11.1.5) and $M_{\nu,\rho}^C$ in (11.1.7).

Even though this stages hypothesis resembles very closely the one in Definition 5.2.8 a direct comparison is not possible, that is, neither implies, in general, the other. This is consistent with the fact that the theorems that we prove using these two hypotheses are not directly comparable.

Reduction by Stages II.   The main goal of this chapter will consist in proving the following theorem whose statement uses the setup and notation introduced above.

11.1.3 Theorem (Reduction by Stages II).   *Suppose that the symplectic manifold (P, Ω) is Lindelöf and that $\sigma \in \mathfrak{m}^*$ satisfies the stages hypothesis II. Then the symplectic reduced spaces*

$$P_\sigma^{C_\nu} := \mathbf{J}_M^{-1}(\sigma)_{C_\nu}/M_\sigma^{C_\nu} \quad and \quad (P_\nu)_\rho^C := (\mathbf{J}_\nu^C)^{-1}(\rho)_C/(M_\nu^C/N_\nu^C)_\rho^C$$

are symplectomorphic.

Before proving this result we will need some technical prerequisites that we introduce in the following paragraphs.

11.2   Momentum Level Sets and Distributions

This section starts with a quick review some well known facts about generalized distributions defined by families of vector fields. The standard references for this topic are Stefan [1974a,b], and Sussman [1973]. We will follow the notation of Libermann and Marle [1987]. We then develop the distributions associated with level sets of the momentum map. This will be used in the following section to give a proof of Theorem 11.1.3.

Some Facts about Generalized Distributions.   Let P be a smooth manifold and $\mathcal{D}$ by an everywhere defined family of vector fields, that is, there is a family of smooth vector fields $\mathcal{D}$ whose elements are vector fields X defined on a open subset $\text{Dom}(X) \subset P$ such that, for any $z \in P$ the generalized distribution D is given by

$$D(z) = \text{span}\{X(z) \in T_z P \mid X \in \mathcal{D} \text{ and } z \in \text{Dom}(X)\}.$$

We will say that D is the generalized distribution **spanned** by $\mathcal{D}$. Note the dimension of D may not be constant; the dimension of D_z is called the **rank** of the distribution D at z. An immersed connected submanifold N of P is said to be an **integral submanifold** of the distribution D if, for every $z \in N$, $T_z i(T_z N) \subset D_z$, where $i : N \to P$ is the canonical injection. The integral submanifold N is said to be of **maximal dimension** at a point $z \in N$ if $T_z i(T_z N) = D_z$. A **maximal integral submanifold** N of D is an integral manifold everywhere of maximal dimension such that any other

integral submanifold of D, which is everywhere of maximal dimension and contains N, is equal to N. The generalized distribution D is said to be **completely integrable** if, for every point $z \in P$, there exists a maximal integral submanifold of D which contains z. This submanifold is usually referred to as the **leaf through** z of the distribution D.

The leaves of an integrable distribution defined by a family of vector fields admit a very convenient description that we briefly review. One of the reasons for our interest in this special case resides in the fact that when these distributions are completely integrable, a very useful characterization of their integral manifolds can be given. In order to describe it we introduce some terminology following Libermann and Marle [1987]. Let X be a vector field defined on an open subset $\mathrm{Dom}(X)$ of M and F_t be its flow. For any fixed $t \in \mathbb{R}$ the domain $\mathrm{Dom}(F_t)$ of F_t is an open subset of $\mathrm{Dom}(X)$ such that $F_t : \mathrm{Dom}(F_t) \to \mathrm{Dom}(F_{-t})$ is a diffeomorphism. If Y is a second vector field defined on the open set $\mathrm{Dom}(Y)$ with flow G_t we can consider, for two fixed values $t_1, t_2 \in \mathbb{R}$, the composition of the two diffeomorphisms $F_{t_1} \circ G_{t_2}$ as defined on the open set $\mathrm{Dom}(G_{t_2}) \cap (G_{t_2})^{-1}(\mathrm{Dom}(F_{t_1}))$ (which may be empty).

The previous prescription allows us to inductively define the composition of an arbitrary number of locally defined flows. We will obviously be interested in the flows associated to the vector fields in $\mathcal{D}$ that define the distribution D. The following sentences describe some important conventions that we will be used throughout the remainder of the book. Let $k \in \mathbb{N}^*$ be a positive natural number, $\mathcal{X}$ be an ordered family $\mathcal{X} = (X_1, \ldots, X_k)$ of k elements of $\mathcal{D}$, and T be a k-tuple $T = (t_1, \ldots, t_k) \in \mathbb{R}^k$ such that $F_{t_i}^i$ denotes the (locally defined) flow of X_i, $i \in \{1, \ldots, k\}$. We will denote by $\mathcal{F}_T$ the locally defined diffeomorphism $\mathcal{F}_T = F_{t_1}^1 \circ F_{t_2}^2 \circ \cdots \circ F_{t_k}^k$ constructed using the above given prescription. Any diffeomorphism from an open subset of P onto another open subset of P that is constructed in the same fashion as $\mathcal{F}_T$ is said to be **generated** by the family $\mathcal{D}$. It can be proven that the composition of diffeomorphisms generated by $\mathcal{D}$ and the inverses of diffeomorphisms generated by $\mathcal{D}$ are themselves diffeomorphisms generated by $\mathcal{D}$ (see Libermann and Marle [1987, Proposition 3.3, Appendix 3]). In other words, the family of diffeomorphisms generated by $\mathcal{D}$ forms a **pseudogroup of transformations** (see page 74 of Paterson [1999]) that will be denoted by G_D. Two points x and y in P are said to be G_D-**equivalent**, if there exists a diffeomorphism $\mathcal{F}_T \in G_D$ such that $\mathcal{F}_T(x) = y$. The relation *being G_D-equivalent* is an equivalence relation whose equivalence classes are called the G_D-**orbits**.

11.2.1 Theorem.   *Let D be a differentiable generalized distribution on the smooth manifold P spanned by an everywhere defined family of vector fields $\mathcal{D}$. The following properties are equivalent:*

(i) *The distribution D is invariant under the pseudogroup of transformations generated by $\mathcal{D}$, that is, for each $\mathcal{F}_T \in G_D$ generated by $\mathcal{D}$*

and for each $z \in P$ in the domain of $\mathcal{F}_T$,

$$T_z \mathcal{F}_T (D_z) = D_{\mathcal{F}_T(z)}.$$

(ii) *The distribution D is completely integrable and its maximal integral manifolds are the G_D-orbits.*

The Momentum Distribution.   In this paragraph we introduce an integrable generalized distribution whose leaves will be the connected components of the level sets of the momentum map. We start with the following proposition whose proof can be found in Ortega [1998] or in Ortega and Ratiu [2002].

11.2.2 Proposition.   *Let M be a Lie group acting properly on the smooth manifold P. Let $z \in P$ be a point with isotropy subgroup $H := M_z$. Then,*

$$\operatorname{span}\{\mathbf{d}f(z) \mid f \in C^{\infty}(P)^M\} = ((T_z(M \cdot z))^{\circ})^H,$$

where $(T_z(M \cdot z))^{\circ}$ is the annihilator of $T_z(M \cdot z)$ in $T_z^* P$ and $((T_z(M \cdot z))^{\circ})^H$ are the elements in $(T_z(M \cdot z))^{\circ}$ fixed by the natural H-action on $T_z^* P$. In particular, if the M-action on P is free, we have

$$
\begin{aligned}
\operatorname{span}\{\mathbf{d}f(z) \mid f \in C^{\infty}(P)^M\} &= (T_z(M \cdot z))^{\circ} \\
&= (T_z(M^0 \cdot z))^{\circ} \\
&= \operatorname{span}\{\mathbf{d}f(z) \mid f \in C^{\infty}(P)^{M^0}\}, \quad (11.2.1)
\end{aligned}
$$

where M^0 denotes the connected component of the identity in M.

The properness hypothesis in the statement of this proposition is essential; the result is actually false, in general, when this condition is absent.

11.2.3 Definition.   *Let (P, Ω) be a symplectic manifold and M be a Lie group acting freely, properly, and canonically on P. Let $\pi_M : P \to P/M$ be the surjective submersion associated to the projection onto the orbit space of the M-action. Let $\mathcal{E} \subset \mathfrak{X}(P)$ be the family of vector fields on P defined by*

$$\mathcal{E} = \{X_{f \circ \pi_M} \mid f \in C_c^{\infty}(P/M)\},$$

*where $C_c^{\infty}(P/M)$ denotes the set of compactly supported smooth functions on P/M. The smooth generalized distribution E spanned by $\mathcal{E}$ is called the M-**characteristic distribution**. If the M-action on P has an associated momentum map, then E is also called the M-**momentum distribution**.*

The terminology that we just introduced is justified by the following theorem.

11.2.4 Theorem.   *Let (P, Ω) be a symplectic manifold and M be a Lie group acting freely, properly, and canonically on P, with associated momentum map $\mathbf{J}_M : P \to \mathfrak{m}^*$. Let E be the M-momentum distribution. Then the following statements hold:*

(i) *The pseudogroup G_E of transformations generated by the family $\mathcal{E}$ is actually a group or, equivalently, the elements of $\mathcal{E}$ are complete vector fields.*

(ii) *The M-momentum distribution E is completely integrable.*

(iii) *For any $z \in P$, $E(z) = \ker T_z \mathbf{J}_M$.*

(iv) *$E = E_C$, where E_C is the distribution spanned by the vector fields*

$$\mathcal{E}_C = \left\{ X_{f \circ \pi_M} \mid f \in C_c^\infty \left(P/M^0 \right) \right\},$$

with M^0 the connected component of the identity of M.

(v) *The integral leaf $\mathcal{L}_z$ of E going through an arbitrary point $z \in P$ is given by*

$$\mathcal{L}_z = G_E \cdot z = (\mathbf{J}_M^{-1}(\sigma))_C,$$

where $\sigma = \mathbf{J}_M(z) \in \mathfrak{m}^*$, $(\mathbf{J}_M^{-1}(\sigma))_C$ is the connected component of $\mathbf{J}_M^{-1}(\sigma)$ containing z, and $G_E \cdot z$ is the G_E-orbit of z, with G_E the group generated by $\mathcal{E}$.

Proof.   To prove (i), we have to show that for any function $f \in C_c^\infty(P/M)$, the Hamiltonian vector field $X_{f \circ \pi_M}$ is complete. We proceed by contradiction: let $z \in P$ be an arbitrary point and $\gamma(t)$, $t \in (T^-(z), T^+(z))$, be the maximal integral curve of $X_{f \circ \pi_M}$ going through z, that is, $\gamma(0) = z$. We assume that the upper bound $T^+(z)$ is finite (the proof dealing with $T^-(z)$ finite is identical). Consider now the maximal integral curve $\rho(t)$ of X_f going through $[z] := \pi_M(z)$. Since the function $f \in C_c^\infty(M/G)$ is compactly supported, so is the vector field X_f and, consequently, the curve $\rho(t)$ is defined for all time t. Let $\rho_0 := \rho(T^+(z))$ and pick an element $y \in P$ in the fiber $\pi_M^{-1}(\rho_0)$. We now call $\sigma(t)$ the maximal integral curve of $X_{f \circ \pi_M}$ that satisfies $\sigma(T^+(z)) = y$ and which is defined on an open neighborhood $(T^+(z) + T^-(y), T^+(z) + T^+(y))$ of $T^+(z)$ ($\sigma(t)$ is just $\sigma(t) := F_{t - T^+(z)}(y)$, with F_t the flow of $X_{f \circ \pi_M}$).

With all these ingredients we now take $0 < \epsilon < |T^-(y)|$ and note that

$$\rho(T^+(z) - \epsilon) = \pi_M(\sigma(T^+(z) - \epsilon)) = \pi_M(\gamma(T^+(z) - \epsilon)),$$

which implies the existence of an element $g \in G$ such that

$$\gamma(T^+(z) - \epsilon) = g \cdot \sigma(T^+(z) - \epsilon).$$

It is easy to verify that the curve $\bar{\gamma}(t)$, $t \in (T^-(z), T^+(z) + T^+(y))$ defined as

$$\bar{\gamma}(t) = \begin{cases} \gamma(t) & \text{when} \quad t \in (T^-(z), T^+(z) - \epsilon] \\ g \cdot \sigma(t) & \text{when} \quad t \in [T^+(z) - \epsilon, T^+(z) + T^+(y)) \end{cases}$$

is an integral curve of $X_{f \circ \pi_M}$ through the point z whose time of existence goes beyond $T^+(z)$, which is a contradiction with the maximality of γ. The time of existence $T^+(z)$ must therefore be infinite.

To prove (ii), we will use Theorem 11.2.1 to prove the complete integrability of E. For simplicity in the exposition, take $\mathcal{F}_T = F_t$, with F_t the flow of $X_{f \circ \pi_M}$, $f \in C_c^\infty(P/M)$ (the general case in which $\mathcal{F}_T$ is the composition of a finite number of flows follows easily by attaching to what we are going to do a straightforward induction argument). The claim holds if we are able to show that

$$T_z F_t(E(z)) = E(F_t(z)).$$

Let $X_{g \circ \pi_M}(z) \in E(z)$, $g \in C_c^\infty(P/M)$. Since any Hamiltonian flow consists of Poisson maps, it follows that

$$T_z F_t(X_{g \circ \pi_M}(z)) = T_z F_t(X_{g \circ \pi_M \circ F_{-t} \circ F_t}(z)) = X_{g \circ \pi_M \circ F_{-t}}(F_t(z)) \in E(F_t(z)),$$

since $g \circ \pi_M \circ F_{-t} \in C^\infty(P)^M$ by the M-equivariance of F_t, and at the same time, it can be written as $g \circ \bar{F}_{-t} \circ \pi_M$, where $\bar{F}_{-t}$ is the diffeomorphism of P/M uniquely determined by the relation $\bar{F}_{-t} \circ \pi_M = \pi_M \circ F_{-t}$. Therefore, $g \circ \bar{F}_{-t} \in C_c^\infty(P/M)$. This implies that $T_z F_t(E(z)) \subset E(F_t(z))$. Conversely, let $X_{g \circ \pi_M}(F_t(z)) \in E(F_t(z))$. Then

$$X_{g \circ \pi_M}(F_t(z)) = T_z F_t(X_{g \circ \pi_M \circ F_t}(z)) = T_z F_t(X_{g \circ \bar{F}_t \circ \pi_M}(z)),$$

which concludes the proof of the integrability of E.

Turning to (iii), for any $z \in P$, let $B^\sharp(z) : T_z^* P \to T_z P$ be the linear isomorphism associated to the symplectic form $\Omega(z)$, evaluated at z. The claim is a consequence of the following chain of equalities:

$$\begin{aligned}
E(z) &= \operatorname{span}\{X_{g \circ \pi_M}(z) \mid g \in C_c^\infty(P/M)\} \\
&= \operatorname{span}\left\{X_f(z) \mid f \in C^\infty(P)^M\right\} = B^\sharp(z)\left(\operatorname{span}\left\{df(z) \mid f \in C^\infty(P)^M\right\}\right) \\
&= B^\sharp(z)\left((T_z(M \cdot z))^\circ\right) \qquad \text{(by Proposition 11.2.2)} \\
&= (T_z(M \cdot z))^\Omega = \ker T_z \mathbf{J}_M.
\end{aligned}$$

Next we prove (iv) As we just saw, for any $z \in P$, we have

$$E(z) = B^\sharp(z)\left((T_z(M \cdot z))^\circ\right) = B^\sharp(z)\left(\left(T_z(M^0 \cdot z)\right)^\circ\right) = E_C(z).$$

To prove (v), note that the integrability of E guarantees, by Theorem 11.2.1, that $\mathcal{L}_z = G_E \cdot z$. Noether's Theorem guarantees that $\mathcal{L}_z = G_E \cdot z \subset (\mathbf{J}_M^{-1}(\sigma))_C$. At the same time, part (ii), implies that $(\mathbf{J}_M^{-1}(\sigma))_C$ is an integral leaf of E, everywhere of maximal dimension. The maximality of $\mathcal{L}_z$ ensures that $\mathcal{L}_z = G_E \cdot z = (\mathbf{J}_M^{-1}(\sigma))_C$. ∎

11.3   Proof: Reduction by Stages II

By now we have all the intermediate results that we need to put together a proof of Theorem 11.1.3. We start by justifying the claim made in (11.1.2) when we said that $\mathbf{J}_N^{-1}(\nu)_C$ fully contains a certain number of connected components $\mathbf{J}_M^{-1}(\sigma)_{C_i}$, $i \in \{1, \ldots, n\}$, of $\mathbf{J}_M^{-1}(\sigma)$. Indeed, let $z \in \mathbf{J}_M^{-1}(\sigma) \subset \mathbf{J}_N^{-1}(\nu)$. Let $\mathbf{J}_N^{-1}(\nu)_C$ be the connected component of $\mathbf{J}_N^{-1}(\nu)$ that contains the point z. We have to show that the connected component $\mathbf{J}_M^{-1}(\sigma)_C$ of $\mathbf{J}_M^{-1}(\sigma)$ that contains z is such that $\mathbf{J}_M^{-1}(\sigma)_C \subset \mathbf{J}_N^{-1}(\nu)_C$. This is actually a corollary of Theorem 11.2.4. Indeed, if we denote by E_M and E_N the momentum distributions associated to the M and N–actions on P, respectively, and by G_{E_M} and G_{E_N} the associated group of transformations we have by construction that $G_{E_M} \subset G_{E_N}$. Consequently,

$$\mathbf{J}_M(\sigma)_C = G_{E_M} \cdot z \subset G_{E_N} \cdot z = \mathbf{J}_N(\nu)_C.$$

We now start the construction of a symplectomorphism between

$$P_\sigma^{C_\nu} := \mathbf{J}_M^{-1}(\sigma)_{C_\nu}/M_\sigma^{C_\nu} \quad \text{and} \quad (P_\nu)_\rho^C := (\mathbf{J}_\nu^C)^{-1}(\rho)_C/(M_\nu^C/N_\nu^C)_\rho^C.$$

As we saw already in (11.1.6), we have

$$\pi_\nu^C(\mathbf{J}_M^{-1}(\sigma)_{C_\nu}) \subset (\mathbf{J}_\nu^C)^{-1}(\rho)_C, \tag{11.3.1}$$

which allows us to define the function $\theta : \mathbf{J}_M^{-1}(\sigma)_{C_\nu} \to (\mathbf{J}_\nu^C)^{-1}(\rho)_C$ as $z \mapsto \pi_\nu^C(z)$. It is easy to check that the smooth mapping θ is equivariant with respect to the $M_\sigma^{C_\nu}$–action on its domain and the $(M_\nu^C/N_\nu^C)_\rho^C$–action on its range, which guarantees the existence of a well-defined smooth function

$$\Theta : \mathbf{J}_M^{-1}(\sigma)_{C_\nu}/M_\sigma^{C_\nu} \to (\mathbf{J}_\nu^C)^{-1}(\rho)_C/(M_\nu^C/N_\nu^C)_\rho^C$$

that makes the diagram

$$
\begin{array}{ccc}
\mathbf{J}_M^{-1}(\sigma)_{C_\nu} & \xrightarrow{\ \theta\ } & (\mathbf{J}_\nu^C)^{-1}(\rho)_C \\
{\scriptstyle \pi_\sigma^{C_\nu}} \downarrow & & \downarrow {\scriptstyle \pi_\rho^C} \\
\mathbf{J}_M^{-1}(\sigma)_{C_\nu}/M_\sigma^{C_\nu} & \xrightarrow{\ \Theta\ } & (\mathbf{J}_\nu^C)^{-1}(\rho)_C/(M_\nu^C/N_\nu^C)_\rho^C.
\end{array}
\tag{11.3.2}
$$

commutative. We will show that the map Θ is the desired symplectic diffeomorphism. We will proceed in several steps.

The Map Θ is Injective.   Let $[z]_\sigma, [z']_\sigma$ be two elements in $\mathbf{J}_M^{-1}(\sigma)_{C_\nu}/M_\sigma^{C_\nu}$ for which $\Theta([z]_\sigma) = \Theta([z']_\sigma)$. By the definition of Θ, this implies that $\pi_\rho^C(\pi_\nu^C(z)) = \pi_\rho^C(\pi_\nu^C(z'))$ hence there exists an element $gN_\nu^C \in (M_\nu^C/N_\nu^C)_\rho^C$ such that $\pi_\nu^C(z') = gN_\nu^C \cdot \pi_\nu^C(z)$ or, equivalently $\pi_\nu^C(z') = \pi_\nu^C(g \cdot z)$. Therefore, there exists a $n \in N_\nu^C$ such that $z' = ng \cdot z$. Given that $z, z' \in \mathbf{J}_M^{-1}(\sigma)_{C_\nu}$, it follows that $ng \in M_\sigma^{C_\nu}$ and therefore $[z]_\sigma = [z']_\sigma$, as required.

The Map Θ is Surjective.   Let $\pi_\rho^C([z]) \in (P_\nu)_\rho^C$ be arbitrary, with $[z] = \pi_\nu^C(z) \in (\mathbf{J}_\nu^C)^{-1}(\rho)$ and $z \in \mathbf{J}_N^{-1}(\nu)_C$. Let $\sigma' := \mathbf{J}_M(z)$. Notice that $\sigma'|_{\mathfrak{n}} = i^* \mathbf{J}_M(z) = \mathbf{J}_N(z) = \nu = \sigma|_{\mathfrak{n}}$. Also, for any $\xi \in \mathfrak{m}_\nu^C$,

$$\langle \sigma', \xi \rangle = \langle \mathbf{J}_M(z), \xi \rangle = \langle \mathbf{J}_\nu^C([z]), [\xi] \rangle + \langle \bar{\nu}, \xi \rangle = \langle \rho, [\xi] \rangle + \langle \bar{\nu}, \xi \rangle = \langle \sigma, \xi \rangle.$$

Consequently, $\sigma'|_{\mathfrak{m}_\nu^C} = \sigma|_{\mathfrak{m}_\nu^C}$ and therefore, by the stages hypothesis II, there exists an element $n \in M_{\nu,\rho}^C$ such that $\sigma' = \mathrm{Ad}_{n^{-1}}^* \sigma$. As $n \in M_{\nu,\rho}^C$, it follows that $[n] := n N_\nu^C \in (M_\nu^C / N_\nu^C)_\rho^C$. Consider now the element $z' := n^{-1} \cdot z \in \mathbf{J}_M^{-1}(\sigma)$. As $n \in M_{\nu,\rho}^C \subset M_\nu^C$, then $z' \in \mathbf{J}_N^{-1}(\sigma)_C$ and therefore $z' \in \mathbf{J}_M^{-1}(\sigma)_{C_\nu}$. By construction,

$$\begin{aligned}
\Theta(\pi_\sigma^{C_\nu}(z')) &= \pi_\rho^C(\theta(z')) \\
&= \pi_\rho^C(\theta(n^{-1} \cdot z)) \\
&= \pi_\rho^C([n^{-1}] \cdot \pi_\nu^C(z)) \\
&= \pi_\rho(\pi_\nu^C(z)) \\
&= \pi_\rho^C([z]),
\end{aligned}$$

which proves the surjectivity of Θ.

The Map Θ is Symplectic and therefore an Immersion.   This assertion is a simple diagram chasing exercise.

Finally, we can affirm that Θ *is a symplectic diffeomorphism* since by Lemma 11.1.1 every bijective immersion is a diffeomorphism once we have assumed that P (and consequently $\mathbf{J}_M^{-1}(\sigma)_{C_\nu} / M_\sigma^{C_\nu}$) is Lindelöf.

12

Reduction by Stages with Topological Conditions

In this chapter we will use the distribution theoretical approach to formulate a reduction by stages theorem that only requires an easily verifiable point set topological condition. This condition is satisfied by a large class of Lie groups, for example, compact ones. Notice that this statement could not have been made had we followed exclusively the purely algebraic approach in § 5.2. Having said that, we will analyze the relation between the stages theorem in this chapter and that in the previous one.

12.1 Reduction by Stages III

In this section we will study a very general condition that also implies the possibility of reducing by stages. We state it in the following result:

12.1.1 Proposition. *Let M^0 be the connected component of the identity of M. Suppose that the symplectic manifold (P, Ω) is Lindelöf and paracompact. Let $\sigma \in \mathbf{J}_M(P) \subset \mathfrak{m}^*$, $\nu := i^*\sigma$, $\mathbf{J}_M^{-1}(\sigma)_C$ be one of the connected components of $\mathbf{J}_M^{-1}(\sigma)$ included in $\mathbf{J}_N^{-1}(\nu)_C$, and θ be the map introduced in (11.3.2). If the orbit $M^0 \cdot \nu \subset \mathfrak{n}^*$ is closed as a subset of $\mathfrak{n}^*$, then*

$$\theta(\mathbf{J}_M^{-1}(\sigma)_C) = (\mathbf{J}_\nu^C)^{-1}(\rho)_C. \tag{12.1.1}$$

Before we proceed with the proof of this proposition we state and prove an important corollary.

12.1.2 Theorem (Reduction by Stages III). *Let M^0 be the connected component of the identity of M. Suppose that the symplectic manifold (P, Ω)*

is Lindelöf and paracompact. Let $\sigma \in \mathbf{J}_M(P) \subset \mathfrak{m}^*$ and $\nu := i^*\sigma$. Suppose that the orbit $M^0 \cdot \nu \subset \mathfrak{n}^*$ is closed as a subset of $\mathfrak{n}^*$. Let $\mathbf{J}_M^{-1}(\sigma)_C$ be one of the connected components of $\mathbf{J}_M^{-1}(\sigma)$ included in $\mathbf{J}_N^{-1}(\nu)_C$. Then the symplectic reduced spaces

$$P_\sigma^C := \mathbf{J}_M^{-1}(\sigma)_C / M_\sigma^C \quad \text{and} \quad (P_\nu)_\rho^C := (\mathbf{J}_\nu^C)^{-1}(\rho)_C / (M_\nu^C / N_\nu^C)_\rho^C$$

are symplectomorphic.

An important particular case in which the closedness hypothesis on the orbit $M^0 \cdot \nu$ in the previous corollary is always satisfied is when the group M is compact. Consequently, whenever the manifold P is Lindelöf and paracompact and the group M is compact, the reduction by stages procedure is always viable.

The closedness hypothesis of the coadjoint orbits in the statement of Proposition 12.1.1 is needed in the proof in relation to the existence of the extensions of certain functions.

Proof of the Theorem.   Consider the following variation of the diagram (11.3.2):

$$
\begin{array}{ccc}
\mathbf{J}_M^{-1}(\sigma)_C & \xrightarrow{\;\theta\;} & (\mathbf{J}_\nu^C)^{-1}(\rho)_C \\[2pt]
{\scriptstyle \pi_\sigma^C} \downarrow & & \downarrow {\scriptstyle \pi_\rho^C} \\[2pt]
\mathbf{J}_M^{-1}(\sigma)_C / M_\sigma^C & \xrightarrow{\;\Theta\;} & (\mathbf{J}_\nu^C)^{-1}(\rho)_C / (M_\nu^C / N_\nu^C)_\rho^C,
\end{array}
$$

where M_σ^C is the subgroup of M_σ that leaves the connected component $\mathbf{J}_M^{-1}(\sigma)_C$ invariant, and $\pi_\sigma : \mathbf{J}_M^{-1}(\sigma)_C \to \mathbf{J}_M^{-1}(\sigma)_C / M_\sigma^C$ is the canonical projection. The equality $\theta(\mathbf{J}_M^{-1}(\sigma)_C) = (\mathbf{J}_\nu^C)^{-1}(\rho)_C$ guaranteed by the previous proposition implies that θ, and consequently Θ, are surjective. Mimicking the proof of Theorem 11.1.3 it can be shown that Θ is also an injective symplectic immersion, and therefore a symplectomorphism.   ■

Proof of Proposition 12.1.1..   We start the proof by stating several lemmas and propositions. For future reference a standard result in manifold theory is stated below. For a proof see Theorem 5.5.9 in [MTA].

12.1.3 Proposition.   *Let P be a paracompact smooth manifold and $A \subset P$ be a closed submanifold of P. Any smooth function $f \in C^\infty(A)$ admits an extension to a smooth function $F \in C^\infty(P)$.*

We now study a distribution that will be of much use.

12.1.4 Lemma.   *Let D be the generalized distribution on P given by $D = A + E_N$, where, for any $z \in P$*

$$A(z) = T_z(M \cdot z) = \{\xi_P(z) \mid \xi \in \mathfrak{m}\},$$

and

$$E_N(z) = \mathrm{span}\{X_{g \circ \pi_N}(z) \mid g \in C_c^\infty(P/N)\},$$

where $\pi_N : P \to P/N$ is the projection onto the orbit space. Then:

(i) *If F_t is the flow of the infinitesimal generator vector field ξ_P, $\xi \in$ $\mathfrak{m}$, and G_t is the flow of the Hamiltonian vector field $X_{g \circ \pi_N}$, $g \in$ $C_c^\infty(P/N)$, then, for any $t_1, t_2 \in \mathbb{R}$, we obtain,*

$$G_{t_1} \circ F_{t_2} = F_{t_2} \circ H_{t_1},$$

where H_t is the flow of the Hamiltonian vector field associated to the function $h := g \circ \pi_N \circ F_{t_2} \in C^\infty(P)^N$ that can also be written as $g \circ \bar{F}_{t_2} \circ \pi_N$. The map $\bar{F}_{t_2}$ is the diffeomorphism of P/N uniquely determined by the relation $\pi_N \circ F_{t_2} = \bar{F}_{t_2} \circ \pi_N$ and $g \circ \bar{F}_{t_2} \in C_c^\infty(P/N)$.

(ii) *D is integrable.*

(iii) *The maximal integral leaves of the distribution D are given by the orbits*

$$G_D \cdot z = G_A \cdot G_{E_N} \cdot z = M^0 \cdot (\mathbf{J}_N^{-1}(\nu)_C),$$

where $\mathbf{J}_N(z) = \nu$, $\mathbf{J}_N^{-1}(\nu)_C$ is the connected component of $\mathbf{J}_N^{-1}(\nu)$ that contains z, and M^0 is the connected component of the identity of M.

Proof.   To prove (i), first note that for any time $t \in \mathbb{R}$ and any $z \in P$, $F_t(z) = \exp t\xi \cdot z$. Also, since N is a normal subgroup of M, for any $n \in N$ and $z \in P$ there exists an element $n' \in N$ such that $F_t(n \cdot z) = \exp t\xi n \cdot z = n' \exp t\xi \cdot z = n' \cdot F_t(z)$. Consequently, the map F_t induces a diffeomorphism $\bar{F}_t$ of P/N uniquely determined by the relation $\bar{F}_t \circ \pi_N = \pi_N \circ F_t$. Also, the function $g \circ \pi_N \circ F_t \in C^\infty(M)$ can be written as $g \circ \bar{F}_t \circ \pi_N$ which guarantees that it is an element of $C^\infty(M)^N$ and that, by Theorem 11.2.4 (i), the Hamiltonian vector field $X_{g \circ \pi_N \circ F_t} = X_{g \circ \bar{F}_t \circ \pi_N}$ is complete. Now, since the M-action on P is canonical the map F_t is Poisson and therefore

$$TF_t \circ X_{g \circ \pi_N \circ F_t} = X_{g \circ \pi_N} \circ F_t.$$

Moreover, if G_t is the flow of $X_{g \circ \pi_N}$ and H_t that of $X_{g \circ \pi_N \circ F_{t_2}}$, then it follows that $G_{t_1} \circ F_{t_2} = F_{t_2} \circ H_{t_1}$. Since all the vector fields involved in this expression are complete, this equality is valid for any $t_1, t_2 \in \mathbb{R}$.

Now we turn to the proof of (ii). According to Theorem 11.2.1 it is enough to show that the distribution D is invariant under the action of the diffeomorphisms group G_D generated by the family of vector fields that spans the distribution D, namely,

$$\{\xi_P \mid \xi \in \mathfrak{m}\} \cup \{X_{g \circ \pi_N} \mid g \in C_c^\infty(P/N)\}.$$

More specifically, we have to show that $T_z \mathcal{F}_T(D(z)) = D(\mathcal{F}_T(z))$, for each $\mathcal{F}_T \in G_D$ and any $z \in P$. Actually, it suffices to show the inclusion

$$T_z \mathcal{F}_T(D(z)) \subset D(\mathcal{F}_T(z)) \tag{12.1.2}$$

given that since (12.1.2) is valid for any element in G_D and any point in P, we get $T_{\mathcal{F}_T(z)}(\mathcal{F}_T)^{-1}(D(\mathcal{F}_T(z))) \subset D(z)$. Applying $T_z \mathcal{F}_T$ to both sides of this inclusion we obtain that $D(\mathcal{F}_T(z)) \subset T_z \mathcal{F}_T(D(z))$, as required.

Hence, we now verify that (12.1.2) holds when $\mathcal{F}_T = F^1_{t_1} \circ \cdots \circ F^n_{t_n}$ with $F^i_{t_i}$ the flow of a vector field either of the form ξ_P, with $\xi \in \mathfrak{m}$, or of the form $X_{g \circ \pi_N}$, with $g \in C^\infty_c(P/N)$. We consider both cases separately.

Firstly, let F_t be the flow of $X_{g \circ \pi_N}$, with $g \in C^\infty_c(P/N)$, and $X_{f \circ \pi_N}$ be another Hamiltonian vector field with $f \in C^\infty_N(P/N)$. Then, since F_t is a Poisson map, we see that for any $z \in P$

$$
\begin{aligned}
T_z F_t(X_{f \circ \pi_N}(z)) &= T_z F_t(X_{f \circ \pi_N \circ F_{-t} \circ F_t}(z)) \\
&= X_{f \circ \pi_N \circ F_{-t}}(F_t(z)) \\
&= X_{f \circ \bar{F}_{-t} \circ \pi_N}(F_t(z)),
\end{aligned}
$$

where $\bar{F}_{-t}$ is the diffeomorphism of P/N uniquely determined by the equality $\bar{F}_{-t} \circ \pi_N = \pi_N \circ F_{-t}$. Given that $f \circ \bar{F}_{-t} \circ \pi_N \in C^\infty(P)^N$ and $f \circ \bar{F}_{-t} \in C^\infty_c(P/N)$, we obtain $X_{f \circ \pi_N \circ F_{-t}}(F_t(z)) \in E_N(F_t(z)) \subset D(F_t(z))$.

Secondly, let ξ_P, be the vector field on P constructed using the infinitesimal generators associated to the element $\xi \in \mathfrak{m}$. The flow of this vector field is given by the map $G_t := \Phi_{\exp t\xi}$. Consequently,

$$
\begin{aligned}
T_z F_t(\xi_P(z)) = \frac{d}{ds}\bigg|_{s=0} F_t(\exp s\xi \cdot z) &= \frac{d}{ds}\bigg|_{s=0} \exp s\xi \cdot F_t^{g \circ \pi_N \circ \Phi_{\exp s\xi}}(z) \\
&= \xi_P(F_t(z)) + \frac{d}{ds}\bigg|_{s=0} F_t^{g \circ \pi_N \circ \Phi_{\exp s\xi}}(z), \quad (12.1.3)
\end{aligned}
$$

where $F_t^{g \circ \pi_N \circ \Phi_{\exp s\xi}}$ is the flow of $X_{g \circ \pi_N \circ \Phi_{\exp s\xi}}$ which, by part (i), is a N-equivariant vector field. Note that the smooth curve $c(s) := F_t^{g \circ \pi_N \circ \Phi_{\exp s\xi}}(z)$ is such that $c(0) = F_t(z)$ and, since $g \circ \pi_N \circ \Phi_{\exp s\xi} \in C^\infty(P)^N$ for all the values of the parameter s then, by Noether's Theorem, $c(s) \in \mathbf{J}_N^{-1}(\nu)$ with $\nu = \mathbf{J}_N(z)$. Therefore,

$$\frac{d}{ds}\bigg|_{s=0} F_t^{g \circ \Phi_{\exp s\xi}}(z) \in \ker T_{F_t(z)}\mathbf{J}_N = E_N(F_t(z))$$

which, substituted in (12.1.3) allows us to conclude that $T_z F_t(\xi_P(z)) \in D(F_t(z))$.

Thirdly, consider the case in which $G_t := \Phi_{\exp t\xi}$ is the flow of ξ_P, $\xi \in \mathfrak{m}$, and let $\eta \in \mathfrak{m}$ be another arbitrary element in the Lie algebra of M. It is

easy to check that

$$T_z G_t(\eta_P(z)) = \eta_P(\exp t\xi \cdot z) + \left.\frac{d}{ds}\right|_{s=0} \exp(t\,\mathrm{Ad}_{\exp -s\eta}\,\xi) \cdot z.$$

Let $g(s) = \exp(t\,\mathrm{Ad}_{\exp -s\eta}\,\xi)$. This curve in M is such that $g(0) = \exp t\xi$, hence, there exists some element $\rho \in \mathfrak{m}$ such that $T_z G_t(\eta_P(z)) = \eta_P(\exp t\xi \cdot z) + \rho_P(\exp t\xi \cdot z) \in A(G_t(z)) \subset E(G_t(z))$, as required.

Finally, let $g \circ \pi_N \in C^\infty(P)^N$, $g \in C_c^\infty(P/N)$, with N-equivariant Hamiltonian flow F_t. Part **(i)** allows us to write

$$T_z G_t(X_{g \circ \pi_N}(z)) = \left.\frac{d}{ds}\right|_{s=0} G_t \circ F_s(z) = \left.\frac{d}{ds}\right|_{s=0} H_s \circ G_t(z),$$

with H_s the flow of the Hamiltonian vector field associated to the N-invariant smooth function $g \circ \pi_N \circ \Phi_{\exp -t\xi}$, hence

$$\begin{aligned}
T_z G_t(X_{g \circ \pi_N}(z)) &= X_{g \circ \pi_N \circ G_{-t}}(G_t(z)) \\
&= X_{g \circ \bar{G}_{-t} \circ \pi_N}(G_t(z)) \in E_N(G_t(z)) \subset D(G_t(z)).
\end{aligned}$$

The four cases studied allow us to conclude that the distribution D is integrable.

Turning to (iii), the integrability of D proved in the previous point and the general theory summarized in Theorem 11.2.1 establish that the maximal integral leaves of D are given by the G_D-orbits. Clearly, $G_A \cdot G_{E_N} \subset G_D$. Part (i) implies the reverse inclusion and therefore $G_A \cdot G_{E_N} = G_D$. Now, by Theorem 11.2.4, $G_{E_N} \cdot z = \mathbf{J}_N^{-1}(\nu)_C$, where $\mathbf{J}_N(z) = \nu$ and $\mathbf{J}_N^{-1}(\nu)_C$ is the connected component of $\mathbf{J}_N^{-1}(\nu)$ that contains z. Consequently,

$$G_D \cdot z = G_A \cdot G_{E_N} \cdot z = M^0 \cdot (\mathbf{J}_N^{-1}(\nu)_C),$$

as required.   ▼

12.1.5 Lemma. *Let $\nu \in \mathfrak{n}^*$ be an element in $\mathfrak{n}^*$ and M^0 be the connected component of the identity of M. Suppose that ν is such that the orbit $M^0 \cdot \nu \subset \mathfrak{n}^*$ is closed as a subset of $\mathfrak{n}^*$. Then the set $\mathbf{J}_N^{-1}(M^0 \cdot \nu)$ is a closed embedded submanifold of P. Moreover, if $\mathbf{J}_N^{-1}(M^0 \cdot \nu)_C$ is the connected component of $\mathbf{J}_N^{-1}(M^0 \cdot \nu)$ that contains $\mathbf{J}_N^{-1}(\nu)_C$, then*

$$M^0 \cdot \mathbf{J}_N^{-1}(\nu)_C = \mathbf{J}_N^{-1}(M^0 \cdot \nu)_C. \tag{12.1.4}$$

Proof. As we already know, since N is a normal subgroup of M, $\mathfrak{n}^*$ is a M-space, therefore a M^0-space, and hence the orbit $M^0 \cdot \nu$ is an immersed submanifold of $\mathfrak{n}^*$. Moreover, we can think of $M^0 \cdot \nu$ as one of the maximal integral manifolds of the singular integrable distribution D_{M^0} on $\mathfrak{n}^*$ defined by

$$D_{M^0}(\zeta) := \{\mathrm{ad}_\xi^* \zeta \mid \xi \in \mathfrak{m}\}, \quad \text{for all} \quad \zeta \in \mathfrak{n}^*.$$

A standard theorem (see Proposition 2.2 in Dazord [1985]) guarantees that the closed integral leaves of a generalized distribution are always imbedded. Therefore, as $M^0 \cdot \nu$ is closed, it is consequently an embedded submanifold of $\mathfrak{n}^*$. Recall now that since $\mathbf{J}_N$ is the momentum map associated to a free canonical action, it is necessarily a submersion and therefore each point of the orbit $M^0 \cdot \nu$ is one of its regular values. The Transversal Mapping Theorem guarantees in these circumstances that $\mathbf{J}_N^{-1}(M^0 \cdot \nu)$ is an embedded submanifold of P. This result also ensures that, for any $z \in \mathbf{J}_N^{-1}(M^0 \cdot \nu)$,

$$T_z(\mathbf{J}_N^{-1}(M^0 \cdot \nu)) = (T_z \mathbf{J}_N)^{-1}(T_{\mathbf{J}_N(z)}(M^0 \cdot \nu)).$$

The M^0-infinitesimal equivariance of $\mathbf{J}_N$ implies that

$$\begin{aligned}
T_{\mathbf{J}_N(z)}(M^0 \cdot \nu) &= \{-\operatorname{ad}_\xi^* \mathbf{J}_N(z) \mid \xi \in \mathfrak{m}\} \\
&= \{T_z \mathbf{J}_N(\xi_P(z)) \mid \xi \in \mathfrak{m}\} \\
&= T_z \mathbf{J}_N(T_z(M^0 \cdot z)),
\end{aligned}$$

and consequently,

$$T_z(\mathbf{J}_N^{-1}(M^0 \cdot \nu)) = T_z(M^0 \cdot z) + \ker T_z \mathbf{J}_N = T_z(M \cdot z) + E_N(z).$$

This equality implies that the manifold $\mathbf{J}_N^{-1}(M^0 \cdot \nu)_C$ is an integral submanifold of the distribution D introduced in Lemma 12.1.4, everywhere of maximal dimension. In that result we saw that the maximal integral submanifolds are given by the subsets of the form $M^0 \cdot (\mathbf{J}_N^{-1}(\nu)_C)$. It is clear that $M^0 \cdot (\mathbf{J}_N^{-1}(\nu)_C) \subset \mathbf{J}_N^{-1}(M^0 \cdot \nu)_C$. The maximality of $M^0 \cdot (\mathbf{J}_N^{-1}(\nu)_C)$ implies equality (12.1.4).   ▼

We are now in the position to state the result on extensions that we will need in the proof of the proposition.

12.1.6 Proposition. *Let $\nu \in \mathfrak{n}^*$ and M^0 be the connected component of the identity of M. Suppose that ν is such that the orbit $M^0 \cdot \nu \subset \mathfrak{n}^*$ is closed as a subset of $\mathfrak{n}^*$. Then, every function $f \in C^\infty(\mathbf{J}_N^{-1}(\nu)_C)^{M_\nu^C}$ admits an extension to a function $F \in C^\infty(P)^{M^0}$.*

Proof. The natural injection $\varphi : \mathbf{J}_N^{-1}(\nu)_C \hookrightarrow M^0 \cdot \mathbf{J}_N^{-1}(\nu)_C = \mathbf{J}_N^{-1}(M^0 \cdot \nu)_C$ induces a smooth map $\phi : \mathbf{J}_N^{-1}(\nu)_C / M_\nu^C \to \mathbf{J}_N^{-1}(M^0 \cdot \nu)_C / M^0$ that makes the following diagram commutative

$$\begin{array}{ccc}
\mathbf{J}_N^{-1}(\nu)_C & \xrightarrow{\quad\varphi\quad} & \mathbf{J}_N^{-1}(M^0 \cdot \nu)_C \\
{\scriptstyle \pi_{M_\nu^C}} \downarrow & & \downarrow {\scriptstyle \pi_{M^0}} \\
\mathbf{J}_N^{-1}(\nu)_C / M_\nu^C & \xrightarrow{\quad\phi\quad} & \mathbf{J}_N^{-1}(M^0 \cdot \nu)_C / M^0.
\end{array}$$

Since in this case the identity (12.1.4) holds, it is easy to verify that ϕ is a bijection. Moreover, it is an immersion, and therefore a diffeomorphism.

Indeed, let $[z]_{M^C_\nu} \in \mathbf{J}_N^{-1}(\nu)_C/M^C_\nu$ be arbitrary and $v_z \in T_z\mathbf{J}_N^{-1}(\nu)_C$ be such that

$$T_{[z]_{M^C_\nu}}\phi(T_z\pi_{M^C_\nu} \cdot v_z) = 0.$$

This can be rewritten as

$$T_z(\phi \circ \pi_{M^C_\nu}) \cdot v_z = T_z(\pi_{M^0} \circ \varphi) \cdot v_z = T_z\pi_{M^0}(T_z\varphi \cdot v_z) = 0.$$

The last equality implies the existence of an element $\xi \in \mathfrak{m}$ such that $T_z\varphi(v_z) = \xi_P(z)$, hence $v_z \in T_z(M \cdot z) \cap T_z(\mathbf{J}_N^{-1}(\nu)) = T(M_\nu \cdot z)$, and consequently $T_z\pi_{M^C_\nu} \cdot v_z = 0$, as required. The equality

$$T_z(M \cdot z) \cap T_z(\mathbf{J}_N^{-1}(\nu)) = T(M_\nu \cdot z)$$

follows easily after recalling that if $\xi_P(z) \in T_z(M \cdot z) \cap T_z(\mathbf{J}_N^{-1}(\nu))$, then $T_z\mathbf{J}_N \cdot \xi_P(z) = -\operatorname{ad}^*_\xi \nu = 0$.

Now take an arbitrary function $f \in C^\infty(\mathbf{J}_N^{-1}(\nu)_C)^{M^C_\nu}$, which induces a function $\bar{f} \in C^\infty\left(\mathbf{J}_N^{-1}(\nu)_C/M^C_\nu\right)$ uniquely determined by the relation $\bar{f} \circ \pi_{M^C_\nu} = f$. Let now $\bar{g} \in C^\infty\left(\mathbf{J}_N^{-1}(M^0 \cdot \nu)_C/M^0\right)$ be the smooth function defined by $\bar{g} = \bar{f} \circ \phi^{-1}$. This function induces a M^0 invariant function $g \in C^\infty(\mathbf{J}_N^{-1}(M^0 \cdot \nu)_C)^{M^0}$ on $\mathbf{J}_N^{-1}(M^0 \cdot \nu)_C$ via the equality $g = \bar{g} \circ \pi_{M^0}$. Since $\mathbf{J}_N^{-1}(M^0 \cdot \nu)_C$ is a closed embedded submanifold of P, the function g can be extended by Proposition 12.1.3 to a smooth function $F \in C^\infty(P)$. The properness of the M^0-action and the M^0-invariance of the function g and of the submanifold $\mathbf{J}_N^{-1}(M^0 \cdot \nu)_C$ guarantee that F can be chosen M^0 invariant, as required (check for instance with Proposition 2 of Arms, Cushman, and Gotay [1991]). $\blacktriangledown$

12.1.7 Corollary. *Suppose that the coadjoint orbit $M^0 \cdot \sigma \subset \mathfrak{m}^*$ is closed in $\mathfrak{m}^*$. Then every function $f \in C^\infty(\mathbf{J}_M^{-1}(\sigma))^{M^C_\sigma}$ admits an extension to a function $F \in C^\infty(P)^{M^0}$. Also, if the coadjoint orbit $M \cdot \sigma \subset \mathfrak{m}^*$ is closed and embedded in $\mathfrak{m}^*$, then every function $f \in C^\infty(\mathbf{J}_M^{-1}(\sigma))^{M_\sigma}$ admits an extension to a function $F \in C^\infty(P)^M$.*

Proof. For the proof of the first statement just take $N = M$ in the proof of the previous proposition. As to the second one, erase the symbols C that refer to connected components and substitute M^0 by M. As to the hypothesis regarding $M \cdot \sigma \subset \mathfrak{m}^*$ being embedded in $\mathfrak{m}^*$ we need it to reproduce the argument at the very beginning of the proof of Lemma 12.1.5 where we would show, in our case, that $\mathbf{J}_M^{-1}(M \cdot \sigma)$ is a closed embedded submanifold of P. $\blacktriangledown$

We are now ready to prove the relation (12.1.1), that is,

$$\theta(\mathbf{J}_M^{-1}(\sigma)_C) = (\mathbf{J}_\nu^C)^{-1}(\rho)_C.$$

The inclusion $\theta(\mathbf{J}_M^{-1}(\sigma)_C) \subset (\mathbf{J}_\nu^C)^{-1}(\rho)_C$ is already known and is a consequence of (11.3.1). Let $\pi_{M_\nu^C/N_\nu^C} : P_\nu^C \to P_\nu^C/(M_\nu^C/N_\nu^C)$ be the canonical projection onto the orbit space. In order to show the equality take an arbitrary point $\pi_\nu^C(z) \in \theta(\mathbf{J}_M^{-1}(\sigma)_C) \subset P_\nu^C$ and consider the maximal integral leaf of the generalized distribution on P_ν^C defined by

$$E_{M_\nu^C/N_\nu^C} = \left\{ X_f \,\middle|\, f \in C^\infty \left(P_\nu^C\right)^{M_\nu^C/N_\nu^C} \quad \text{with} \right.$$

$$\left. f = F \circ \pi_{M_\nu^C/N_\nu^C},\, F \in C_c^\infty \left(P_\nu^C/(M_\nu^C/N_\nu^C)\right) \right\}$$

that goes through $\pi_\nu^C(z)$ which, by Theorem 11.2.4, is the entire $(\mathbf{J}_\nu^C)^{-1}(\rho)_C$. If we are able to show that for any $\mathcal{F}_T \in G_{E_{M_\nu^C/N_\nu^C}}$, we have

$$\mathcal{F}_T(\pi_\nu^C(z)) \in \theta(\mathbf{J}_M^{-1}(\sigma)_C) = \pi_\nu^C(\mathbf{J}_M^{-1}(\sigma)_C),$$

we will have proved the equality. For the sake of simplicity suppose that $\mathcal{F}_T = F_t$, with F_t the Hamiltonian flow associated to the function $f \in C^\infty \left(P_\nu^C\right)^{M_\nu^C/N_\nu^C}$. Let $\bar{f} \in C^\infty(\mathbf{J}_N^{-1}(\nu)_C)^{M_\nu^C}$ be the smooth function defined by $\bar{f} := f \circ \pi_\nu^C$, and let $g \in C^\infty(P)^{M^0}$ be one of its smooth M^0 invariant extensions to P, whose existence is guaranteed by Proposition 12.1.6. Let G_t be the Hamiltonian flow associated to the function g. Note that

$$F_t(\pi_\nu^C(z)) = \pi_\nu^C(G_t(z)).$$

By Theorem 11.2.4, $G_t(z) \in \mathbf{J}_M^{-1}(\sigma)_C$ and therefore

$$\mathcal{F}_T(\pi_\nu^C(z)) = \pi_\nu^C(G_t(z)) \in \theta(\mathbf{J}_M^{-1}(\sigma)_C),$$

as required.   ∎

12.2   Relation Between Stages II and III

The reader may be wondering if there is a relation between the versions II and III of the reduction by stages theorem. Even though it is true that both results identify sufficient conditions that allow symplectic reduction in two stages, these conditions seem to be nonequivalent. The following proposition shows that the closedness hypothesis in the version III needs to be complemented with an additional condition in order to imply the stages hypothesis II, and therefore the version II of the reduction by stages theorem.

12.2.1 Proposition.   *Suppose that the hypotheses of Proposition 12.1.1 hold and that, additionally, the following condition is satisfied: for any $\sigma' \in \mathfrak{m}^*$ such that $\nu := i^*\sigma'$, there is at least one connected component $\mathbf{J}_M^{-1}(\sigma')_C$ of $\mathbf{J}_M^{-1}(\sigma')$ included in the given connected component $\mathbf{J}_N^{-1}(\nu)_C$. Then, σ satisfies stages hypothesis II.*

Proof.   Let $\sigma' \in \mathfrak{m}^*$ be such that $\sigma'|_{\mathfrak{n}} = \sigma|_{\mathfrak{n}} = \nu$ and $\sigma'|_{\mathfrak{m}_\nu^C} = \sigma|_{\mathfrak{m}_\nu^C}$. By hypothesis, there is a connected component $\mathbf{J}_M^{-1}(\sigma')_C$ of $\mathbf{J}_M^{-1}(\sigma')$ included in $\mathbf{J}_N^{-1}(\nu)_C$. Let $z \in \mathbf{J}_M^{-1}(\sigma')$. Then, for any $[\xi] \in \mathfrak{m}_\nu^C/\mathfrak{n}_\nu^C$ we have

$$\langle \mathbf{J}_\nu^C(\pi_\nu^C(z)), [\xi] \rangle = \langle \sigma', \xi \rangle - \langle \bar{\nu}, \xi \rangle = \langle \sigma, \xi \rangle - \langle \bar{\nu}, \xi \rangle = \langle \rho, [\xi] \rangle.$$

Consequently, $\pi_\nu^C(z)$ belongs to the set $(\mathbf{J}_\nu^C)^{-1}(\rho)_C$ which by (12.1.1) equals $\theta(\mathbf{J}_M^{-1}(\sigma)_C)$. Hence, there exists $z' \in \mathbf{J}_M^{-1}(\sigma)_C$ such that $\pi_\nu^C(z) = \pi_\nu^C(z')$ and therefore $z' = n \cdot z$ for some $n \in N_\nu^C \subset M_{\nu,\rho}^C$. Applying the map $\mathbf{J}_M$ to both sides of this equality we obtain that $\sigma' = \mathrm{Ad}_{n^{-1}}^* \sigma$. Hence, σ satisfies the stages hypothesis II.   ∎

The following example shows that the situation is similar regarding the reverse implication. More specifically, our example will describe a situation where the stages hypothesis II holds but not the closedness hypothesis needed in the version III of the reduction by stages theorem.

Example.   Let M be the subgroup of $SL(2, \mathbb{R})$ given by

$$M = \left\{ \begin{bmatrix} a & b \\ 0 & a^{-1} \end{bmatrix} \middle| a > 0, b \in \mathbb{R} \right\}.$$

Consider now the closed normal subgroup N of M given by

$$N = \left\{ \begin{bmatrix} 1 & c \\ 0 & 1 \end{bmatrix} \middle| c \in \mathbb{R} \right\}.$$

The Lie algebra $\mathfrak{m}$ of M is given by the matrices of the form

$$\mathfrak{m} = \left\{ \begin{bmatrix} \xi_1 & \xi_2 \\ 0 & -\xi_1 \end{bmatrix} \middle| \xi_1, \xi_2 \in \mathbb{R} \right\}.$$

If we choose the matrices

$$\mathbf{e}_1 := \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \quad \text{and} \quad \mathbf{e}_2 := \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

as a basis for $\mathfrak{m}$, we can write its elements as two–tuples of the form $(\xi_1, \xi_2) \in \mathbb{R}^2$. In these coordinates, the adjoint action of M on $\mathfrak{m}$ can be expressed as

$$\mathrm{Ad}_g(\xi_1, \xi_2) = (\xi_1, -2ab\xi_1 + a^2\xi_2),$$

where

$$g = \begin{bmatrix} a & b \\ 0 & a^{-1} \end{bmatrix} \in M \text{ and } (\xi_1, \xi_2) \in \mathfrak{m}.$$

If we identify $\mathfrak{m}^*$ with $\mathfrak{m} \simeq \mathbb{R}^2$ via the Euclidean inner product on $\mathbb{R}^2$, the coadjoint action of M on $\mathfrak{m}^*$ takes the following expression

$$\mathrm{Ad}_{g^{-1}}^*(\alpha_1, \alpha_2) = \left(\alpha_1 + \frac{2b}{a}\alpha_2, \frac{1}{a^2}\alpha_2 \right), \tag{12.2.1}$$

where

$$g = \begin{bmatrix} a & b \\ 0 & a^{-1} \end{bmatrix} \in M \text{ and } (\alpha_1, \alpha_2) \in \mathfrak{m}^*.$$

Analogously, the inclusion $i : \mathfrak{n} \hookrightarrow \mathfrak{m}$ is given by $\eta \longmapsto (0, \eta)$, and the dual projection $i^* : \mathfrak{m}^* \to \mathfrak{n}^*$ by $(\alpha_1, \alpha_2) \longmapsto \alpha_2$. Moreover, the coadjoint action of M on $\mathfrak{n}^*$ is given by

$$\mathrm{Ad}^*_{g^{-1}} \alpha = \frac{1}{a^2} \alpha, \text{ for any } g = \begin{bmatrix} a & b \\ 0 & a^{-1} \end{bmatrix} \in M. \qquad (12.2.2)$$

If we visualize $\mathfrak{m}^*$ as the (α_1, α_2)–plane, expression (12.2.1) implies that the M–coadjoint orbits in $\mathfrak{m}^*$ are the open upper and lower half planes together with the points in the α_1–axis. Analogously, by (12.2.2) we can conclude that the coadjoint action of M on $\mathfrak{n}^*$ exhibits three coadjoint orbits, namely, two open half lines, and the point where they meet. In conclusion, if $\alpha \in \mathfrak{n}^*$ is different from zero, its M–orbit is not closed in $\mathfrak{n}^*$ and, consequently, the hypothesis of version III of the Stages Theorem is not satisfied. Nevertheless, we will now consider a free canonical action of M on a symplectic manifold for which both the Stages Hypotheses I and II hold.

Consider the lifted action of M on its cotangent bundle T^*M. If we trivialize T^*M using right translations (space coordinates) we have the following expressions for this canonical action and for its associated M and N–momentum maps:

$$g \cdot (h, \alpha) = (gh, \mathrm{Ad}^*_{g^{-1}} \alpha), \quad \mathbf{J}_M(g, \alpha) = \alpha, \quad \mathbf{J}_N(g, \alpha) = \alpha_2,$$

for any $g, h \in M$ and $\alpha = (\alpha_1, \alpha_2) \in \mathfrak{m}^*$. As the level sets of $\mathbf{J}_M$ and $\mathbf{J}_N$ are connected, there is no difference between the versions I and II of the stages hypothesis. We now verify that this hypothesis holds: first of all notice that for any $\alpha \in \mathfrak{n}^*$

$$M_\alpha = \begin{cases} M, & if \quad \alpha = 0 \\ N, & if \quad \alpha \neq 0. \end{cases}$$

Therefore, if $\alpha = 0$ the stages hypothesis holds trivially. If $\alpha \neq 0$ and we have two elements $\sigma, \sigma' \in \mathfrak{m}^*$ such that $\sigma|_\mathfrak{n} = \sigma'|_\mathfrak{n} = \sigma|_{\mathfrak{m}_\alpha} = \sigma'|_{\mathfrak{m}_\alpha} = \alpha$ then there exist necessarily $\beta, \gamma \in \mathbb{R}$ such that $\sigma = (\beta, \alpha)$ and $\sigma' = (\gamma, \alpha)$, and consequently:

$$\sigma' = \mathrm{Ad}^*_{g^{-1}} \sigma, \quad \text{with} \quad g = \begin{bmatrix} 1 & \dfrac{\gamma - \beta}{2\alpha} \\ 0 & 1 \end{bmatrix} \in N_\alpha = N,$$

as required.   ◆

12.3 Connected Components of Reduced Spaces

A natural question that arises when making the comparison between the distribution approach to the reduction by stages problem and the one taken in § 5.2 is how the reduced spaces obtained in both cases are related. In the following paragraphs we will show that if the coadjoint orbit $M \cdot \sigma \subset \mathfrak{m}^*$ is a closed and embedded submanifold of $\mathfrak{m}^*$, then P_σ^C is a connected component of $\mathbf{J}_M^{-1}(\sigma)/M_\sigma$. An analogous claim can be made regarding $P_{\nu,\rho}^C$. We state all of this in the following proposition.

12.3.1 Proposition. *Let M^0 be the connected component of the identity of M. Suppose that the symplectic manifold (P, Ω) is Lindelöf and paracompact and that the coadjoint orbit $M \cdot \sigma \subset \mathfrak{m}^*$ is a closed and embedded submanifold of $\mathfrak{m}^*$. Then P_σ^C is a connected component of $\mathbf{J}_M^{-1}(\sigma)/M_\sigma$. The same conclusion holds for $P_{\nu,\rho}^C$ whenever the orbit $(M_\nu^C/N_\nu^C) \cdot \rho$ under the affine action of (M_ν^C/N_ν^C) on $(\mathfrak{m}_\nu^C/\mathfrak{n}_\nu^C)^*$ is closed in $(\mathfrak{m}_\nu^C/\mathfrak{n}_\nu^C)^*$.*

Proof. First of all, notice that since $\mathbf{J}_M^{-1}(\sigma)_C$ is connected, so is

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

and hence the projection of the inclusion $\mathbf{J}_M^{-1}(\sigma)_C \hookrightarrow \mathbf{J}_M^{-1}(\sigma)$ provides us with an injection

$$i_C : \mathbf{J}_M^{-1}(\sigma)_C/M_\sigma^C \longrightarrow \left(\mathbf{J}_M^{-1}(\sigma)/M_\sigma\right)_C$$

of $\mathbf{J}_M^{-1}(\sigma)_C/M_\sigma^C$ into some connected component $\left(\mathbf{J}_M^{-1}(\sigma)/M_\sigma\right)_C$ of P_σ. We will prove that i_C is onto. To do this, we will follow a strategy similar in spirit to the one we used to establish the surjectivity of the map Θ. As $\left(\mathbf{J}_M^{-1}(\sigma)/M_\sigma\right)_C$ is a connected symplectic manifold, any two of its points can be joined by piecewise Hamiltonian paths or, more explicitly, the maximal integral leaf of the distribution on $\left(\mathbf{J}_M^{-1}(\sigma)/M_\sigma\right)_C$

$$D = \left\{ X_{h_\sigma} \mid h_\sigma \in C^\infty\left(\left(\mathbf{J}_M^{-1}(\sigma)/M_\sigma\right)_C\right)\right\}$$

going through any point in $\left(\mathbf{J}_M^{-1}(\sigma)/M_\sigma\right)_C$ is $\left(\mathbf{J}_M^{-1}(\sigma)/M_\sigma\right)_C$ itself. We will show the surjectivity of i_C by proving that for any $h_\sigma \in C^\infty\left(\left(\mathbf{J}_M^{-1}(\sigma)/M_\sigma\right)_C\right)$ with associated Hamiltonian flow F_t^σ, and for any $\pi_\sigma^C(z) \in P_\sigma^C$, we get

$$F_t^\sigma(i_C(\pi_\sigma^C(z))) \in i_C(P_\sigma^C).$$

Let $\bar{h}_\sigma \in C^\infty\left(\mathbf{J}_M^{-1}(\sigma)/M_\sigma\right)$ be an extension of $h_\sigma \in C^\infty\left(\left(\mathbf{J}_M^{-1}(\sigma)/M_\sigma\right)_C\right)$ and let $h \in C^\infty\left(\mathbf{J}_M^{-1}(\sigma)\right)^{M_\sigma}$ be the function defined by $h = \bar{h}_\sigma \circ \pi_\sigma$. By Corollary 12.1.7 the function h admits an extension to a function $H \in C^\infty(P)^M$; let F_t be its associated Hamiltonian flow. Then,

$$F_t^\sigma(i_C(\pi_\sigma^C(z))) = \pi_\sigma(F_t(z)).$$

By Noether's Theorem $F_t(z) \in \mathbf{J}_M^{-1}(\sigma)_C$ and consequently $\pi_\sigma(F_t(z)) \in i_C(P_\sigma^C)$, as required.   ∎

Conclusions for Part II

In this part we have given a thorough treatment of the problem of regular symplectic reduction by stages. There are, however, many things left to do in this area. Amongst these, there is a need for additional concrete physical applications. Another is to make use of the stages structures in numerical applications using, for instance, variational integrators, as in Marsden and West [2001]. Finally, there is a need for additional functional analytic treatments of infinite dimensional cases, some of which were mentioned in the introduction and in the text.

Recall from the introduction that there is a parallel theory of *Lagrangian reduction by stages* developed in Cendra, Marsden, and Ratiu [2001a]. A critical difference is that the theory of Lagrangian reduction by stages is the Lagrangian analog of Poisson reduction in that no imposition of a level set of the momentum map is made. Nevertheless, there are many strong connections and parallels between the results in the present work and those in the theory of Lagrangian reduction by stages. The Lagrangian analog of point reduction in the symplectic context is that of Routh reduction studied in Marsden, Ratiu and Scheurle [2000]. Of course developing a reduction by stages theory in that context would be of interest.

One may also speculate on further relations with group theory along the lines of the orbit method. After all, the orbit method for semidirect products is closely related to the method of induced representations of Mackey. One would imagine that keeping track of representation theory parallel to reduction by stages would also be interesting.

Another important issue is how to properly generalize things to the multi-symplectic context (see, for instance Marsden, Patrick, and Shkoller [1998]). As we have mentioned, in a number of examples in field theory, including complex fluids, one has a cocycle in the associated Poisson structure (Holm and Kupershmidt [1982, 1983b]). The structure of those theories strongly suggests that a reduction by stages approach would be profitable, although the analog of symplectic reduction in field theories is known to be tricky as one normally does *not* impose momentum map constraints until after a $3+1$ (space + time) split has been made. This will complicate any eventual theory. On the other hand, from a Lagrangian reduction by stages standpoint, some interesting progress has been made in this direction (see Holm [2002]).

Part III. Optimal Reduction and Singular Reduction by Stages

by Juan-Pablo Ortega

In this part we will apply the techniques recently introduced by see Ortega and Ratiu [2002] and Ortega [2003a], which are based on the theory of distributions to generalize the results presented in Part II as well as other various symplectic reduction methods due to Marsden, Weinstein, Sjamaar, Bates, Lerman, Marle, Kazhdan, Kostant, and Sternberg. This approach is based on the definition of an object, that will be called the *optimal momentum map*, that to some extent generalizes the standard momentum map that we used previously in this book. One of the advantages of this point of view is its degree of generality which will allow us to construct symplectic point and orbit reduced spaces purely within the Poisson category under hypotheses that do not necessarily imply the existence of a standard or group valued (Alekseev, Malkin, and Meinrenken [1998]) momentum map. All along this part we will refer to the construction of symplectic reduced spaces with the help of the optimal momentum map as **optimal reduction**.

This part is divided into three chapters. The first chapter contains a brief description of the optimal momentum map and explains how to construct symplectic or Marsden–Weinstein reduced spaces in the context in which this object is defined. Most of the definitions and results in this preliminary chapter are contained in Ortega and Ratiu [2002]; Ortega [2002].

The second chapter of this part explores in the context of the optimal momentum map the orbit reduction procedure. Orbit reduction is an approach to symplectic reduction equivalent to the Marsden–Weinstein reduc-

tion that we used in Part II. It consists of considering in the reduction process the inverse images by the momentum map of coadjoint orbits instead of just momentum values; this set is, under certain hypotheses, a smooth G–invariant space whose corresponding orbit space is symplectomorphic to the Marsden–Weinstein reduced space. In order to make a distinction between these two reduction procedures we will talk about **_point_** and **_orbit reduction_**. The expression for the symplectic structure of the orbit reduced space puts into relation the symplectic form of the original manifold with the so called natural Kostant–Kirillov–Souriau symplectic form of the coadjoint orbit that we used to construct it. Our goal in this chapter will be the reproduction of this scheme in the context of the optimal momentum map. This will need the introduction of smooth (pre)–symplectic manifolds that will generalize to this context the coadjoint orbits and their natural symplectic structures. We will refer to these as **_polar reduced spaces_** due to their close ties with the notion of polarity introduced by the author in his study of singular dual pairs (Ortega [2003a]). In particular, we will see that there is an interesting interplay between the so called von Neumann condition for a canonical group action and the polar reduction scheme.

The last chapter of this Part III will extend to the optimal context the reduction by stages procedure that we studied in Part II for the standard momentum map. The advantages in terms of generality presented by the optimal momentum map will allow us to formulate a reduction by stages theorem for any normal subgroup of a Lie group acting canonically on a Poisson manifold for which, in principle, there is no associated standard momentum map. At the end of this chapter, an motivated by its importance in terms of applications we will study the particular case of a Hamiltonian proper action on a symplectic manifold for which we do not impose freeness. These results provide the generalization to the singular context of the reduction by stages theorems in Part II.

13

The Optimal Momentum Map and Point Reduction

Unless the contrary is explicitly stated, all along this part we will work on a Poisson manifold $(M, \{\cdot, \cdot\})$ acted canonically and properly upon by the Lie group G via the left action $\Phi : G \times M \to M$. The group of Poisson transformations associated to this action will be denoted by

$$A_G := \{\Phi_g \mid g \in G\}$$

and the projection of M onto the orbit space by $\pi_{A_G} : M \to M/A_G = M/G$. Given a point $m \in M$ with isotropy subgroup G_m we will denote by $\mathfrak{g} \cdot m := T_m(G \cdot m)$ the tangent space at m to the G–orbit that goes through m. We recall that when the G–action is proper, the connected components of the subset

$$M_{G_m} := \{z \in M \mid G_z = G_m\}$$

made out of the points that have the same isotropy as m, are smooth submanifolds that we will call ***isotropy type submanifolds***.

13.1 Optimal Momentum Map and Space

The optimal momentum map was introduced in Ortega and Ratiu [2002] as a general method to find the conservation laws associated to the symmetries of a Poisson system. We briefly recall its definition and elementary properties. Let A'_G be the distribution on M defined by the relation:

$$A'_G(m) := \{X_f(m) \mid f \in C^\infty(M)^G\}, \qquad \text{for all} \quad m \in M.$$

Depending on the context, the generalized distribution (in the sequel we will omit the adjective "generalized") A'_G is called the G–**characteristic distribution** or the **polar distribution** defined by A_G Ortega [2003a]. The distribution A'_G is smooth and integrable in the sense of Stefan and Sussman Stefan [1974a,b]; Sussman [1973], that is, given any point in M there is an **accessible set** or **maximal integral leaf** of A'_G going through it. We recall (see Ortega [2003a] for the details) that if M is actually a symplectic manifold with form ω then

$$A'_G(m) = (\mathfrak{g} \cdot m)^{\omega} \cap T_m M_{G_m}, \quad \text{for all} \quad m \in M. \tag{13.1.1}$$

Moreover, if the G–action has an associated standard momentum map $\mathbf{J} : M \to \mathfrak{g}^*$ then

$$A'_G(m) = \ker T_m \mathbf{J} \cap T_m M_{G_m}, \quad \text{for all} \quad m \in M. \tag{13.1.2}$$

The **optimal momentum map** $\mathcal{J}$ is defined as the canonical projection onto the leaf space of A'_G, that is,

$$\mathcal{J} : M \longrightarrow M/A'_G.$$

By its very definition, the fibers or levels sets of $\mathcal{J}$ are preserved by the Hamiltonian flows associated to G–invariant Hamiltonian functions and $\mathcal{J}$ is **universal** with respect to this property, that is, any other map whose level sets are preserved by G–equivariant Hamiltonian dynamics factors necessarily through $\mathcal{J}$.

The leaf space M/A'_G is called the **momentum space** of $\mathcal{J}$. When considered as a topological space endowed with the quotient topology, it is easy to see (Ortega [2003a]) that the optimal momentum map is **continuous and open**.

The pair $(C^\infty(M/A'_G), \{\cdot, \cdot\}_{M/A'_G})$ is a Poisson algebra (the term Poisson variety is also frequently used) when we define

$$C^\infty(M/A'_G) := \{f \in C^0(M/A'_G) \mid f \circ \mathcal{J} \in C^\infty(M)\}, \tag{13.1.3}$$

and the bracket $\{\cdot, \cdot\}_{M/A'_G}$ defined by

$$\{f, g\}_{M/A'_G}(\mathcal{J}(m)) = \{f \circ \mathcal{J}, g \circ \mathcal{J}\}(m), \tag{13.1.4}$$

for every $m \in M$ and $f, g \in C^\infty(M/A'_G)$. Note that as $\mathcal{J}$ is open and surjective then, for any real valued function f on M/A'_G such that $f \circ \mathcal{J} \in C^\infty(M)$ we see that $f \in C^0(M/A'_G)$. Hence, in this case, the standard definition (13.1.3) can be rephrased by saying that $f \in C^\infty(M/A'_G)$ if and only if $f \circ \mathcal{J} \in C^\infty(M)$.

The G–action on M naturally induces a smooth action $\Psi : G \times M/A'_G \to M/A'_G$ of G on M/A'_G defined by the expression $\Psi(g, \mathcal{J}(m)) := \mathcal{J}(g \cdot m)$ with respect to which $\mathcal{J}$ is G–equivariant. We recall that the term smooth

in this context means that $\Psi^*C^\infty(M/A'_G) \subset C^\infty(G \times M/A'_G)$. Notice that since M/A'_G is not Hausdorff in general, there is no guarantee that the isotropy subgroups G_ρ of elements $\rho \in M/A'_G$ are closed, and therefore embedded, subgroups of G. However, there is still something that we can say:

13.1.1 Proposition.   *Let G_ρ be the isotropy subgroup of the element $\rho \in M/A'_G$ associated to the G–action on M/A'_G that we just defined. Then:*

(i) *There is a unique smooth structure on G_ρ with respect to which this subgroup is an initial (see below) Lie subgroup of G with Lie algebra $\mathfrak{g}_\rho$ given by*

$$\mathfrak{g}_\rho = \{\xi \in \mathfrak{g} \mid \xi_M(m) \in T_m\mathcal{J}^{-1}(\rho), \text{ for all } m \in \mathcal{J}^{-1}(\rho)\} \quad (13.1.5)$$

or, equivalently

$$\mathfrak{g}_\rho = \{\xi \in \mathfrak{g} \mid \exp t\xi \in G_\rho, \text{ for all } t \in \mathbb{R}\}. \quad (13.1.6)$$

(ii) *With this smooth structure for G_ρ, the left action $\Phi^\rho : G_\rho \times \mathcal{J}^{-1}(\rho) \to \mathcal{J}^{-1}(\rho)$ defined by $\Phi^\rho(g,z) := \Phi(g,z)$ is smooth.*

(iii) *This action has fixed isotropies, that is, if $z \in \mathcal{J}^{-1}(\rho)$ then $(G_\rho)_z = G_z$, and $G_m = G_z$ for all $m \in \mathcal{J}^{-1}(\rho)$.*

(iv) *Let $z \in \mathcal{J}^{-1}(\rho)$ arbitrary. Then,*

$$\mathfrak{g}_\rho \cdot z = A'_G(z) \cap \mathfrak{g} \cdot z = T_z\mathcal{J}^{-1}(\rho) \cap \mathfrak{g} \cdot z. \quad (13.1.7)$$

Proof.   For (i) through (iii), see Ortega [2002]. We prove (iv): the inclusion $\mathfrak{g}_\rho \cdot z \subset A'_G(z) \cap \mathfrak{g} \cdot z$ is a consequence of (13.1.5). Conversely, let $X_f(z) = \xi_M(z) \in A'_G(z) \cap \mathfrak{g} \cdot z$, with $f \in C^\infty(M)^G$ and $\xi \in \mathfrak{g}$. The G–invariance of the function f implies that $[X_f, \xi_M] = 0$, and hence, if F_t is the flow of X_f and G_t is the flow of ξ_M (more explicitly $G_t(m) = \exp t\xi \cdot m$ for any $m \in M$), then $F_t \circ G_s = G_s \circ F_t$. By one of the Trotter product formulas (see [MTA, Corollary 4.1.27]), the flow H_t of $X_f - \xi_M$ is given by

$$H_t(m) = \lim_{n\to\infty} \left(F_{t/n} \circ G_{-t/n}\right)^n (m)$$
$$= \lim_{n\to\infty} \left(F_{t/n}^n \circ G_{-t/n}^n\right)(m) = (F_t \circ G_{-t})(m) = F_t(\exp -t\xi \cdot m),$$

for any $m \in M$. Consequently, as $X_f(z) = \xi_M(z)$, the point $z \in M$ is an equilibrium of $X_f - \xi_M$, hence $F_t(\exp -t\xi \cdot z) = z$ or, analogously $\exp t\xi \cdot z = F_t(z)$. Applying $\mathcal{J}$ on both sides of this equality, and taking into account that F_t is the flow of a G–invariant Hamiltonian vector field, it follows that $\exp t\xi \cdot \rho = \rho$, and hence $\xi \in \mathfrak{g}_\rho$ by (13.1.6) . Thus $\xi_M(z) \in \mathfrak{g}_\rho \cdot z$, as required.   ∎

Recall that we say that N is an ***initial submanifold*** of the smooth manifold M when the inclusion $i : N \rightarrow M$ is a smooth immersion that satisfies that for any manifold Z, a mapping $f : Z \rightarrow N$ is smooth if and only if $i \circ f : Z \rightarrow M$ is smooth. The initial submanifold structure is unique in the sense that if N admits another smooth structure, call it N', that makes it into an initial submanifold of M, then the identity map $id_N : N \rightarrow N'$ is a diffeomorphism. Indeed, as the injection $N \hookrightarrow M$ is smooth and N' is by hypothesis initial then, the identity map $id_N : N \rightarrow N'$ is smooth. As the same argument can be made for $id_{N'} : N' \rightarrow N$, the result follows.

We finish this section by emphasizing that the structure of the momentum space M/A'_G may become very intricate. The following example shows that even when the G–action is very simple and the corresponding orbit space $M/G = M/A_G$ is a quotient regular manifold, the associated momentum space M/A'_G does not need to share those properties.

Example.   Let $M := \mathbb{T}^2 \times \mathbb{T}^2$ be the product of two tori whose elements we will denote by the four–tuples $(e^{i\theta_1}, e^{i\theta_2}, e^{i\psi_1}, e^{i\psi_2})$. We endow M with the symplectic structure ω defined by

$$\omega := \mathbf{d}\theta_1 \wedge \mathbf{d}\theta_2 + \sqrt{2}\,\mathbf{d}\psi_1 \wedge \mathbf{d}\psi_2.$$

We now consider the canonical circle action given by

$$e^{i\phi} \cdot (e^{i\theta_1}, e^{i\theta_2}, e^{i\psi_1}, e^{i\psi_2}) := (e^{i(\theta_1+\phi)}, e^{i\theta_2}, e^{i(\psi_1+\phi)}, e^{i\psi_2}).$$

First of all, notice that since the circle is compact and acts freely on M, the corresponding orbit space M/A_{S^1} is a smooth manifold such that the projection $\pi_{A_{S^1}} : M \rightarrow M/A_{S^1}$ is a surjective submersion. The polar distribution A'_{S^1} does not have that property. Indeed, $C^\infty(M)^{S^1}$ comprises all the functions f of the form $f \equiv f(e^{i\theta_2}, e^{i\psi_2}, e^{i(\theta_1-\psi_1)})$. An inspection of the Hamiltonian flows associated to such functions readily shows that the leaves of A'_{S^1} fill densely the manifold M and that the leaf space M/A'_{S^1} can be identified with the leaf space $\mathbb{T}^2/\mathbb{R}$ of a Kronecker (irrational) foliation of a two–torus $\mathbb{T}^2$. Under these circumstances M/A'_{S^1} cannot possibly be a regular quotient manifold. ♦

13.2   Momentum Level Sets and Associated Isotropies

By construction, the fibers of $\mathcal{J}$ are the leaves of an integrable generalized distribution and thereby ***initial immersed submanifolds*** of M Dazord [1985]. We summarize this and other elementary properties of the fibers of $\mathcal{J}$ in the following proposition.

13.2.1 Proposition.   *Let $(M, \{\cdot, \cdot\})$ be a Poisson manifold and G be a Lie group that acts properly and canonically on M. Let $\mathcal{J} : M \to M/A'_G$ be the associated optimal momentum map. Then for any $\rho \in M/A'_G$ we obtain:*

(i) *The level set $\mathcal{J}^{-1}(\rho)$ is an immersed initial submanifold of M.*

(ii) *There is a unique symplectic leaf $\mathcal{L}$ of $(M, \{\cdot, \cdot\})$ such that $\mathcal{J}^{-1}(\rho) \subset \mathcal{L}$.*

(iii) *Let $m \in M$ be an arbitrary element of $\mathcal{J}^{-1}(\rho)$. Then, $\mathcal{J}^{-1}(\rho) \subset M_{G_m}$, with $M_{G_m} := \{z \in M \mid G_z = G_m\}$.*

In the sequel we will denote by $\mathcal{L}_\rho$ the unique symplectic leaf of M that contains $\mathcal{J}^{-1}(\rho)$. Notice that as $\mathcal{L}_\rho$ is also an immersed initial submanifold of M, the injection $i_{\mathcal{L}_\rho} : \mathcal{J}^{-1}(\rho) \hookrightarrow \mathcal{L}_\rho$ is smooth.

¿From the point of view of the optimal momentum map the existence of a standard ($\mathfrak{g}^*$ or G–valued) momentum map can be seen as an integrability feature of the G–characteristic distribution that makes the fibers of $\mathcal{J}$ particularly well–behaved. Indeed, it can be proved that when M is a symplectic manifold and the G–action has an associated standard momentum map, then the fibers $\mathcal{J}^{-1}(\rho)$ of the optimal momentum map are closed imbedded submanifolds of M. More generally, if $\mathcal{J}^{-1}(\rho)$ is closed as a subset of the isotropy type submanifold M_H in which it is sitting, then (see Ortega and Ratiu [2002])

- $\mathcal{J}^{-1}(\rho)$ is a closed embedded submanifold of M_H and therefore an embedded submanifold of M, and

- the isotropy subgroup G_ρ of $\rho \in A'_G$ is a closed embedded Lie subgroup of G.

13.3   Optimal Momentum Map Dual Pair

It is well known that the standard momentum map can be used to construct a **dual pair** in the sense of Lie [1890] and Weinstein [1983a]: consider a canonical, proper, and **free action** of a connected Lie group G on a symplectic manifold (M, ω). If this action has an associated equivariant momentum map $\mathbf{J} : M \to \mathfrak{g}^*$, the diagram

$$
\begin{array}{ccc}
 & M & \\
\pi_{A_G} \swarrow & & \searrow \mathbf{J} \\
M/G & & \mathfrak{g}^*_\mathbf{J} := \mathbf{J}(M)
\end{array}
$$

is such that the maps π_{A_G} and $\mathbf{J}$ have symplectically orthogonal fibers. This is the defining condition of a dual pair. As as corollary of a result by Weinstein [1983a]; Blaom [2001] in the context of the theory of dual pairs (the so called Symplectic Leaves Correspondence Theorem) we see that if $\mathbf{J}$ has connected fibers then there is a bijective correspondence between the symplectic leaves of M/G, namely the symplectic orbit reduced spaces, and those of $\mathfrak{g}_{\mathbf{J}}^*$, that is, the coadjoint orbits inside $\mathbf{J}(M)$.

The notions of duality and dual pair have been generalized in Ortega [2003a] in such way that in many situations the optimal momentum map provides an example of these newly introduced dual pairs. We now briefly recall some of the notions introduced in that paper. For the details and proofs of the following facts the reader is encouraged to check with Ortega [2003a].

13.3.1 Definition.   *Let M be a smooth manifold. A **pseudogroup of transformations** or **pseudogroup of local diffeomorphisms** A of M is a set of local diffeomorphisms of M that satisfy:*

(i) *Each $\phi \in A$ is a diffeomorphism of an open set (called the domain of ϕ) of M onto an open set (called the range of ϕ) of M.*

(ii) *Let $U = \cup_{i \in I} U_i$, where each U_i is an open set of M. A diffeomorphism ϕ of U onto an open set of M belongs to A if and only if the restrictions of ϕ to each U_i is in A.*

(iii) *For every open set U of M, the identity transformation of U is in A.*

(iv) *If $\phi \in A$, then $\phi^{-1} \in A$.*

(v) *If $\phi \in A$ is a diffeomorphism of U onto V and $\phi' \in A$ is a diffeomorphism of U' onto V' and $V \cap U'$ is nonempty, then the diffeomorphism $\phi' \circ \phi$ of $\phi^{-1}(V \cap U')$ onto $\phi'(V \cap U')$ is in A.*

Let A be a pseudogroup of transformations on a manifold M and $\sim$ be the relation on M defined by: for any $x, y \in M$, $x \sim y$ if and only if there exists $\phi \in A$ such that $y = \phi(x)$. The relation $\sim$ is an equivalence relation whose space of equivalence classes is denoted by M/A.

*If M is a Poisson manifold with Poisson bracket $\{\cdot, \cdot\}$, we say that a pseudogroup of transformations A of M is a **pseudogroup of local Poisson diffeomorphisms** when any diffeomorphism $\phi \in A$ of an open set U of M onto an open set V of M is also a Poisson map between $(U, \{\cdot, \cdot\}_U)$ and $(V, \{\cdot, \cdot\}_V)$. The symbols $\{\cdot, \cdot\}_U$ and $\{\cdot, \cdot\}_V$ denote the restrictions of the bracket $\{\cdot, \cdot\}$ to U and V, respectively.*

13.3.2 Definition.   *Let $(M, \{\cdot, \cdot\})$ be a Poisson manifold and A be a pseudogroup of local Poisson diffeomorphisms of M. Let $\mathcal{A}'$ be the set of Hamiltonian vector fields associated to all the elements of $C^\infty(U)^A$, for all*

the open A–invariant subsets U of M, that is,

$$\mathcal{A}' = \left\{ X_f \mid f \in C^\infty(U)^A, \; with \; U \subset M \; open \; and \; A\text{–}invariant \right\}.$$
$$(13.3.1)$$

*The distribution A' associated to the family $\mathcal{A}'$ will be called the **polar distribution** defined by A (or equivalently the **polar of** A). Any generating family of vector fields for A' will be called a **polar family** of A. The family $\mathcal{A}'$ will be called the **standard polar family** of A. The pseudogroup of local Poisson diffeomorphisms constructed using finite compositions of flows of the vector fields in any polar family of A will be referred to as a **polar pseudogroup** induced by A. The polar pseudogroup $G_{\mathcal{A}'}$ induced by the standard polar family $\mathcal{A}'$ will be called the **standard polar pseudogroup**.*

Remark. We say that the pseudogroup A has the **extension property** when any A–invariant function $f \in C^\infty(U)^A$ defined on any A–invariant open subset U satisfies that: for any $z \in U$, there is a A–invariant open neighborhood $V \subset U$ of z and a A–invariant smooth function $F \in C^\infty(M)^A$ such that $f|_V = F|_V$. If the pseudosubgroup A has the extension property, there is a simpler polar family, we will call it $\mathcal{A}'_{ext}$, that can be used to generate A', namely

$$\mathcal{A}'_{ext} = \left\{ X_f \mid f \in C^\infty(M)^A \right\}.$$

In particular, if $A = A_G$, that is, the Poisson diffeomorphism group associated to a proper canonical G–action, the extension property is always satisfied and hence $\mathcal{A}'_{ext} = A'_G$, the G–characteristic distribution.

13.3.3 Definition.   *Let $(M, \{\cdot, \cdot\})$ be a Poisson manifold and A, B be two pseudogroups of local Poisson diffeomorphisms. We say that the diagram*

$$(M, \{\cdot, \cdot\})$$

$$\pi_A \swarrow \qquad \searrow \pi_B$$

$$(M/A, \{\cdot, \cdot\}_{M/A}) \qquad\qquad (M/B, \{\cdot, \cdot\}_{M/B})$$
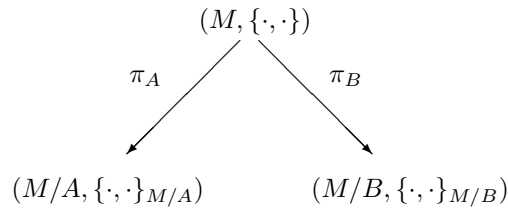
*is a **dual pair** on $(M, \{\cdot, \cdot\})$ when the polar distributions A' and B' are integrable and they satisfy that*

$$M/A' = M/B \; and \; M/B' = M/A. \qquad (13.3.2)$$

We now focus on the dual pairs induced by the optimal momentum map. Hence, let $(M, \{\cdot, \cdot\})$ be a Poisson manifold acted canonically and properly upon by a Lie group G, A_G be the associated group of canonical transformations and $\mathcal{J} : M \to M/A'_G$ be the optimal momentum map. A natural question to ask is when the diagram

$$(M, \{\cdot, \cdot\})$$

$$\pi_{A_G} \swarrow \qquad \searrow \mathcal{J}$$
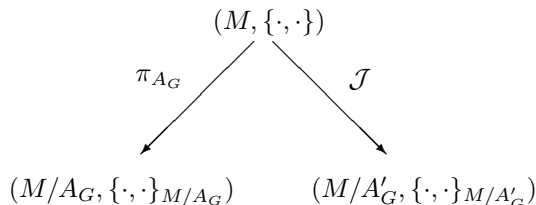
$$(M/A_G, \{\cdot, \cdot\}_{M/A_G}) \qquad (M/A'_G, \{\cdot, \cdot\}_{M/A'_G})$$

is a dual pair in the sense of Definition 13.3.3. Obviously, in this case, condition (13.3.2) is satisfied if and only if the double polar $A''_G := (G_{\mathcal{A}'})'$ of A_G is such that

$$M/A''_G = M/A_G.$$

Poisson subgroups satisfying this condition are referred to as ***von Neumann subgroups***. In our discussion on orbit reduction we will use a slightly less demanding condition, namely, we will need group actions such that

$$\mathfrak{g} \cdot z = T_z(A_G(z)) \subset A''_G(z), \quad \text{for all} \quad z \in M. \tag{13.3.3}$$

A group action that satisfies (13.3.3) is called ***weakly von Neumann***. Obviously, if A_G is von Neumann it is weakly von Neumann. Given that A''_G is spanned by Hamiltonian vector fields, the weak von Neumann condition (13.3.3) implies that for any $z \in M$ sitting in the symplectic leaf $\mathcal{L}_z$, we see that $\mathfrak{g} \cdot z \subset T_z \mathcal{L}_z$, and in particular, if G^0 is the connected component of G containing the identity, then the orbit $G^0 \cdot z$ is contained in the symplectic leaf $\mathcal{L}_z$.

We say that the group A_G is ***weakly Hamiltonian*** when for every element $g \in G$ and any $m \in M$ we can write

$$\Phi_g(m) = F^1_{t_1} \circ F^2_{t_2} \circ \cdots \circ F^k_{t_k}(m),$$

with $F^i_{t_i}$ the flow of a Hamiltonian vector field X_{h_i} associated to a function $h_i \in \left(C^\infty(M)^G\right)^c$ that centralizes the G–invariant functions on M. It is clear that connected Lie group actions that have an associated standard ($\mathfrak{g}^*$ or G–valued) momentum map are weakly Hamiltonian. The importance of this condition in relation to our dual pairs is linked to the fact that ***weakly Hamiltonian proper actions induce von Neumann subgroups***. For a proof of this fact and for other situations where the von Neumann condition is satisfied see Ortega [2003a].

13.4   Dual Pairs, Reduced Spaces, and Symplectic Leaves

Let $(M, \{\cdot, \cdot\})$ be a smooth Poisson manifold, A be a subgroup of its Poisson diffeomorphism group, and $(M/A, \{\cdot, \cdot\}_{M/A})$ be the associated quotient

Poisson variety. Let $V \subset M/A$ be an open subset of M/A and $h \in C^\infty(V)$ be a smooth function defined on it. If we call $U := \pi_A^{-1}(V)$ then, the vector field $X_{h \circ \pi_A|_U}$ belongs to the standard polar family $\mathcal{A}'$ and therefore its flow $(F_t, \mathrm{Dom}(F_t))$ uniquely determines a local Poisson diffeomorphism $(\bar{F}_t, \pi_A(\mathrm{Dom}(F_t)))$ of M/A. We will say that $(\bar{F}_t, \pi_A(\mathrm{Dom}(F_t)))$ is the **Hamiltonian flow** associated to h. The symplectic leaves of M/A will be defined as the accessible sets in this quotient by finite compositions of Hamiltonian flows. It is not clear how to define these flows by projection of A–equivariant flows when A is a pseudogroup of local transformations of M, hence we will restrict in this section to the case in which A is an actual group of Poisson diffeomorphisms.

13.4.1 Definition.   *Let $(M, \{\cdot, \cdot\})$ be a smooth Poisson manifold, A be a subgroup of its Poisson diffeomorphism group, and $(M/A, \{\cdot, \cdot\}_{M/A})$ be the associated quotient Poisson variety. Given a point $[m]_A \in M/A$, the* **symplectic leaf** *$\mathcal{L}_{[m]_A}$ going through it is defined as the (path connected) set formed by all the points that can be reached from $[m]_A$ by applying to it a finite number of Hamiltonian flows associated to functions in $C^\infty(V)$, with $V \subset M/A$ any open subset of M/A, that is,*

$$\mathcal{L}_{[m]_A} := \{ F_{t_1}^1 \circ F_{t_2}^2 \circ \cdots \circ F_{t_k}^k([m]_A) \mid k \in \mathbb{N},$$
$$F_{t_i} \text{ flow of some } X_{h_i}, \ h_i \in C^\infty(V), \ V \subset M/A \text{ open} \}.$$

The relation being in the same symplectic leaf *determines an equivalence relation in M/A whose corresponding space of equivalence classes will be denoted by $(M/A)/\{\cdot, \cdot\}_{M/A}$.*

13.4.2 Theorem (Symplectic leaves correspondence).   *Let $(M, \{\cdot, \cdot\})$ be a smooth Poisson manifold, A, B be two groups of Poisson diffeomorphisms of M, and $G_{\mathcal{A}'}, G_{\mathcal{B}'}$ be the standard polar pseudogroups. If we denote by $(M/A)/\{\cdot, \cdot\}_{M/A}$ and $(M/B)/\{\cdot, \cdot\}_{M/B}$ the space of symplectic leaves of the Poisson varieties $(M/A, \{\cdot, \cdot\}_{M/A})$ and $(M/B, \{\cdot, \cdot\}_{M/B})$, respectively, we have:*

(i) *The symplectic leaves of M/A and M/B are given by the orbits of the $G_{\mathcal{A}'}$ and $G_{\mathcal{B}'}$ actions on M/A and M/B, respectively. As a consequence of this statement, we can write*

$$(M/A)/\{\cdot, \cdot\}_{M/A} = (M/A)/G_{\mathcal{A}'} \text{ and } (M/B)/\{\cdot, \cdot\}_{M/B} = (M/B)/G_{\mathcal{B}'}.$$
$$(13.4.1)$$

(ii) *If the diagram $(M/A, \{\cdot, \cdot\}_{M/A}) \xleftarrow{\pi_A} (M, \{\cdot, \cdot\}) \xrightarrow{\pi_B} (M/B, \{\cdot, \cdot\}_{M/B})$ is a dual pair then the map*

$$
\begin{array}{ccc}
(M/A)/\{\cdot, \cdot\}_{M/A} & \longrightarrow & (M/B)/\{\cdot, \cdot\}_{M/B} \\
\mathcal{L}_{[m]_A} & \longmapsto & \mathcal{L}_{[m]_B}
\end{array}
\qquad (13.4.2)
$$

is a bijection. The symbols $\mathcal{L}_{[m]_A}$ and $\mathcal{L}_{[m]_B}$ denote the symplectic leaves in M/A and M/B, respectively, going through the point $[m]_A$ and $[m]_B$.

One of our goals in the following pages will consist of describing the symplectic leaves of the Poisson varieties in the legs of the diagram

$$(M/G, \{\cdot,\cdot\}_{M/A_G}) \overset{\pi_{A_G}}{\longleftarrow} (M, \{\cdot,\cdot\}) \overset{\mathcal{J}}{\longrightarrow} (M/A'_G, \{\cdot,\cdot\}_{M/A'_G})$$

which, in some situations will coincide with the symplectic reduced spaces that constitute one of the main themes of our work. We emphasize that in order to have well-defined symplectic leaves in the Poisson varieties $(M/A_G, \{\cdot,\cdot\}_{M/A_G})$ and $(M/A'_G, \{\cdot,\cdot\}_{M/A'_G})$ it is very important that A_G is an actual group and not just a local group of Poisson transformations and the same with the polar pseudogroup that generates A'_G. When the manifold M is symplectic and the G–group action is proper it can be proved that there exists a polar family $\mathcal{A}^{c'}_G$ made only of complete vector fields (see Ortega [2003a]) which shows that $M/A'_G = M/G_{\mathcal{A}^{c'}_G}$ is the quotient space by a Poisson group action and that, therefore, its symplectic leaves are well–defined. In general we say that A'_G is **completable** whenever there exists a polar family $\mathcal{A}^{c'}_G$ made only of complete vector fields.

13.5   Optimal Point Reduction

We start by recalling the basics of the classical symplectic or Marsden–Weinstein reduction theory. Let (M, ω) be a symplectic manifold and G be a compact connected Lie group acting freely on (M, ω) by symplectomorphisms. Suppose that this action has an associated standard equivariant momentum map $\mathbf{J} : M \to \mathfrak{g}^*$. There are two equivalent approaches to reduction that can be found in the literature:

- **Point reduction:** it is preferable for applications in dynamics and it is the point of view that we took in Part II. The point reduction theorem says that for any $\mu \in \mathbf{J}(M) \subset \mathfrak{g}^*$, the quotient $\mathbf{J}^{-1}(\mu)/G_\mu$ is a symplectic manifold with symplectic form ω_μ uniquely determined by the equality

$$\pi^*_\mu \omega_\mu = i^*_\mu \omega,$$

  where G_μ is the isotropy subgroup of the element $\mu \in \mathfrak{g}^*$ with respect to the coadjoint action of G on $\mathfrak{g}^*$, $i_\mu : \mathbf{J}^{-1}(\mu) \hookrightarrow M$ is the canonical injection, and $\pi_\mu : \mathbf{J}^{-1}(\mu) \to \mathbf{J}^{-1}(\mu)/G_\mu$ the projection onto the orbit space.

- **Orbit reduction:** this approach is particularly important in the treatment of quantization questions. Let $\mathcal{O}$ be the coadjoint orbit

of some element $\mu \in \mathbf{J}(M)$. The subset $\mathbf{J}^{-1}(\mathcal{O})$ is a smooth submanifold of M and the quotient $\mathbf{J}^{-1}(\mathcal{O})/G$ is a regular symplectic quotient manifold with the symplectic form $\omega_{\mathcal{O}}$ determined by the equality

$$i_{\mathcal{O}}^* \omega = \pi_{\mathcal{O}}^* \omega_{\mathcal{O}} + \mathbf{J}_{\mathcal{O}}^* \omega_{\mathcal{O}}^+, \tag{13.5.1}$$

where $i_{\mathcal{O}} : \mathbf{J}^{-1}(\mathcal{O}) \hookrightarrow M$ is the inclusion, $\pi_{\mathcal{O}} : \mathbf{J}^{-1}(\mathcal{O}) \to \mathbf{J}^{-1}(\mathcal{O})/G$ the projection, $\mathbf{J}_{\mathcal{O}} = \mathbf{J}|_{\mathbf{J}^{-1}(\mathcal{O})}$, and $\omega_{\mathcal{O}}^+$ the "+" orbit symplectic structure on $\mathcal{O}$ (also called Kostant–Kirillov–Souriau —KKS for short— symplectic structure). The use of the orbit reduction approach is particularly convenient when we are interested in the study of the geometry of the orbit space M/G as a Poisson manifold. Indeed, the connected components of $\mathbf{J}^{-1}(\mathcal{O})/G$ constitute the symplectic leaves of M/G and expression (13.5.1) appears as a corollary of a result in the theory of dual pairs. Indeed, as we already said, the diagram $M/G \overset{\pi_{A_G}}{\leftarrow} (M,\omega) \overset{\mathbf{J}}{\to} \mathbf{J}(M) \subset \mathfrak{g}^*$ forms a dual pair and it has been shown (see Weinstein [1983a]; Blaom [2001]) that whenever we have two Poisson manifolds in the legs of a dual pair $(P_1, \{\cdot,\cdot\}_{P_1}) \overset{\pi_1}{\leftarrow} (M,\omega) \overset{\pi_2}{\to} (P_2, \{\cdot,\cdot\}_{P_2})$ (in the sense that $(\ker T\pi_1)^\omega = \ker T\pi_2$) and π_1 and π_2 have connected fibers, its symplectic leaves are in bijection. Moreover, if two symplectic leaves $\mathcal{L}_1 \subset P_1$ and $\mathcal{L}_2 \subset P_2$ are in correspondence, their symplectic structures $\omega_{\mathcal{L}_1}$ and $\omega_{\mathcal{L}_2}$ are linked by the equality

$$i_{\mathcal{K}}^* \omega = \pi_1|_{\mathcal{K}}^* \omega_{\mathcal{L}_1} + \pi_2|_{\mathcal{K}}^* \omega_{\mathcal{L}_2}, \tag{13.5.2}$$

where $\mathcal{K} \subset M$ is the leaf of the integrable distribution $\ker T\pi_1 + \ker T\pi_2$ that contains both $\pi_1^{-1}(\mathcal{L}_1)$ and $\pi_2^{-1}(\mathcal{L}_2)$. Therefore, if we assume that $\mathbf{J}$ has connected fibers, expression (13.5.1) is a particular case of (13.5.2), given that $\mathbf{J}(\mathcal{O})/G$ and $\mathcal{O}$ are symplectic leaves in correspondence of M/G and $\mathbf{J}(M) \subset \mathfrak{g}^*$, respectively.

The use of the optimal momentum map allows the extension of these reduction procedures to far more general situations. Indeed, as we will see in the following paragraphs, the optimal approach allows the construction of symplectically reduced spaces purely within the Poisson category under hypothesis that do not necessarily imply the existence of a standard momentum map. Moreover, we will develop an orbit reduction procedure that in the context of the dual pairs reviewed in Section 13.3 reproduces the beautiful interplay between symplectic reduction and Poisson geometry that we just reviewed.

The study of point reduction has been carried out in Ortega [2002]. We reproduce here the main result in that paper. In the statement we will denote by $\pi_\rho : \mathcal{J}^{-1}(\rho) \to \mathcal{J}^{-1}(\rho)/G_\rho$ the canonical projection onto the orbit space of the G_ρ–action on $\mathcal{J}^{-1}(\rho)$ defined in Proposition 13.1.1.

13.5.1 Theorem (Optimal point reduction by Poisson actions)**.** *Suppose*

that $(M, \{\cdot, \cdot\})$ is a smooth Poisson manifold and G be a Lie group acting canonically and properly on M. Let $\mathcal{J} : M \to M/A'_G$ be the optimal momentum map associated to this action. Then, for any $\rho \in M/A'_G$ whose isotropy subgroup G_ρ acts properly on $\mathcal{J}^{-1}(\rho)$, the orbit space $M_\rho := \mathcal{J}^{-1}(\rho)/G_\rho$ is a smooth symplectic regular quotient manifold with symplectic form ω_ρ defined by:

$$\pi_\rho^* \omega_\rho(m)(X_f(m), X_h(m)) = \{f, h\}(m), \tag{13.5.3}$$

*for any $m \in \mathcal{J}^{-1}(\rho)$ and any $f, h \in C^\infty(M)^G$. We will refer to the pair (M_ρ, ω_ρ) as the **(optimal) point reduced space** of $(M, \{\cdot, \cdot\})$ at ρ.*

Remark. Let $i_{\mathcal{L}_\rho} : \mathcal{J}^{-1}(\rho) \hookrightarrow \mathcal{L}_\rho$ be the natural smooth injection of $\mathcal{J}^{-1}(\rho)$ into the symplectic leaf $(\mathcal{L}_\rho, \omega_{\mathcal{L}_\rho})$ of $(M, \{\cdot, \cdot\})$ in which it is sitting. As $\mathcal{L}_\rho$ is an initial submanifold of M, the injection $i_{\mathcal{L}_\rho}$ is a smooth map. The form ω_ρ can also be written in terms of the symplectic structure of the leaf $\mathcal{L}_\rho$ as

$$\pi_\rho^* \omega_\rho = i_{\mathcal{L}_\rho}^* \omega_{\mathcal{L}_\rho}. \tag{13.5.4}$$

The reader should be warned that this statement does NOT imply that the previous theorem could be obtained by just performing symplectic optimal reduction Ortega and Ratiu [2002] in the symplectic leaves of the Poisson manifold, basically because those leaves are not G–manifolds. Recall that the fact that the G–action is Poisson does not imply that it preserves the symplectic leaves.

In view of this remark we can obtain the standard Symplectic Foliation Theorem of Poisson manifolds as a straightforward corollary of Theorem 13.5.1 by taking the group $G = \{e\}$. In that case the distribution A'_G coincides with the characteristic distribution of the Poisson manifold and the level sets of the optimal momentum map, and thereby the symplectic quotients M_ρ, are exactly the symplectic leaves. We explicitly point this out in our next statement.

13.5.2 Corollary (Symplectic Foliation Theorem). *Assume that $(M, \{\cdot, \cdot\})$ is a smooth Poisson manifold. Then, M is the disjoint union of the maximal integral leaves of the integrable generalized distribution E given by*

$$E(m) := \{X_f(m) \mid f \in C^\infty(M)\}, \quad m \in M.$$

These leaves are symplectic initial submanifolds of M.

Remark. The only extra hypothesis in the statement of Theorem 13.5.1 with respect to the hypotheses used in the classical reduction theorems is the properness of the G_ρ–action on $\mathcal{J}^{-1}(\rho)$. This is a real hypothesis in the sense that the properness of the G_ρ–action is not automatically inherited from the properness of the G–action on M, as it used to be the case in the presence of a standard momentum map (see Ortega and Ratiu [2002]). For

an example illustrating that this is really the case the reader may want to check with Ortega [2002].

The interest of reduction in Poisson dynamics is justified by the following result whose proof is a simple diagram chasing exercise.

13.5.3 Theorem (Optimal Point Reduction: Equivariant Poisson Dynamics). *Let $(M, \{\cdot, \cdot\})$ be a smooth Poisson manifold and G be a Lie group acting canonically and properly on M. Let $\mathcal{J} : M \to M/A_G'$ be the associated optimal momentum map and $\rho \in M/A_G'$ be such that G_ρ acts properly on $\mathcal{J}^{-1}(\rho)$. Let $h \in C^\infty(M)^G$ be a G–invariant function on M and X_h be the associated G–equivariant Hamiltonian vector field on M. Then,*

(i) *The flow F_t of X_h leaves $\mathcal{J}^{-1}(\rho)$ invariant, commutes with the G–action, and therefore induces a flow F_t^ρ on M_ρ uniquely determined by the relation $\pi_\rho \circ F_t \circ i_\rho = F_t^\rho \circ \pi_\rho$, where $i_\rho : \mathcal{J}^{-1}(\rho) \hookrightarrow M$ is the inclusion.*

(ii) *The flow F_t^ρ in (M_ρ, ω_ρ) is Hamiltonian with the Hamiltonian function $h_\rho \in C^\infty(M_\rho)$ given by the equality $h_\rho \circ \pi_\rho = h \circ i_\rho$.*

(iii) *Let $k \in C^\infty(M)^G$ be another G–invariant function on M and $\{\cdot, \cdot\}_\rho$ be the Poisson bracket associated to the symplectic form ω_ρ on M_ρ. Then, $\{h, k\}_\rho = \{h_\rho, k_\rho\}_\rho$.*

13.6 The Symplectic Case and Sjamaar's Principle

In the next few paragraphs we will see that when M is a symplectic manifold with form ω, the optimal point reduction by the G–action on M produces the same results as the reduction of the isotropy type submanifolds by the relevant remaining group actions on them. In the globally Hamiltonian context, that is, in the presence of a G–equivariant momentum map, this idea is usually referred to as **Sjamaar's principle** Sjamaar [1990]; Sjamaar and Lerman [1991].

Let $\mathcal{J} : M \to M/A_G'$ be the optimal momentum map corresponding to the proper G–action on (M, ω). Fix $\rho \in M/A_G'$ a momentum value of $\mathcal{J}$ and let $H \subset G$ be the unique G–isotropy subgroup such that $\mathcal{J}^{-1}(\rho) \subset M_H$ and $G_\rho \subset H$. Recall that the normalizer $N(H)$ of H in G acts naturally. This action induces a **free action** of the quotient group $L := N(H)/H$ on M_H. Let M_H^ρ be the unique connected component of M_H that contains $\mathcal{J}^{-1}(\rho)$ and L^ρ be the closed subgroup of L that leaves it invariant. Obviously, L^ρ can be written as $L^\rho = N(H)^\rho/H$ for some closed subgroup $N(H)^\rho$ of $N(H)$.

The subset M_H^ρ is a symplectic embedded submanifold of M where the group L^ρ acts freely and canonically. We will denote by $\mathcal{J}_{L^\rho} : M_H^\rho \to$

M_H^ρ/A'_{L^ρ} the associated optimal momentum map. The following proposition explains the interest of this construction. We omit the proof since it is a straightforward consequence of the existence of local G–invariant extensions to M for the L^ρ–invariant smooth functions defined in M_H^ρ that has been proved in Lemma 4.4 of Ortega and Ratiu [2002].

13.6.1 Proposition (Optimal Sjamaar's Principle). *Let G be a Lie group that acts properly and canonically on the symplectic manifold (M, ω), with associated optimal momentum map $\mathcal{J} : M \to M/A'_G$. Let $\rho \in M/A'_G$ and $H \subset G$ be the unique G–isotropy subgroup such that $\mathcal{J}^{-1}(\rho) \subset M_H$ and $G_\rho \subset H$. Then, with the notation introduced in the previous paragraphs we get:*

(i) *Let $i_H^\rho : M_H^\rho \hookrightarrow M$ be the inclusion. For any $z \in M_H^\rho$, $T_z i_H^\rho \cdot A'_{L^\rho}(z) = A'_G(z)$.*

(ii) *Let $z \in \mathcal{J}^{-1}(\rho)$ be such that $\mathcal{J}_{L^\rho}(z) = \sigma \in M_H^\rho/A'_{L^\rho}$. Then, $\mathcal{J}^{-1}(\rho) = \mathcal{J}_{L^\rho}^{-1}(\sigma)$.*

(iii) *$L_\sigma^\rho = G_\rho/H$.*

(iv) *$(M_H^\rho)_\sigma = \mathcal{J}_{L^\rho}^{-1}(\sigma)/L_\sigma^\rho = \mathcal{J}^{-1}(\rho)/(G_\rho/H) = \mathcal{J}^{-1}(\rho)/G_\rho = M_\rho$. Moreover, if G_ρ acts properly on $\mathcal{J}^{-1}(\rho)$ this equality is true when we consider M_ρ and $(M_H^\rho)_\sigma$ as symplectic spaces, that is,*

$$(M_\rho, \omega_\rho) = ((M_H^\rho)_\sigma, (\omega|_{M_H^\rho})_\sigma).$$

13.6.2 Definition. *Suppose that we are under the hypotheses of the previous proposition. We will refer to the symplectic reduced space $((M_H^\rho)_\sigma, (\omega|_{M_H^\rho})_\sigma)$ as the **regularization** of the point reduced space (M_ρ, ω_ρ).*

14
Optimal Orbit Reduction

As we already pointed out the main difference between the point and orbit reduced spaces is the invariance properties of the submanifolds out of which they are constructed. More specifically, if we mimic in the optimal context the standard orbit reduction procedure, the optimal orbit reduced space that we should study is $G \cdot \mathcal{J}^{-1}(\rho)/G = \mathcal{J}^{-1}(\mathcal{O}_\rho)/G$, where $\mathcal{O}_\rho := G \cdot \rho \subset M/A'_G$. The following pages constitute an in–depth study of this quotient and its relation with new (pre)–symplectic manifolds that can be used to reproduce the classical orbit reduction program and expressions.

14.1 The Space for Optimal Orbit Reduction

The first question that we have to tackle is: is there a canonical smooth structure for $\mathcal{J}^{-1}(\mathcal{O}_\rho)$ and $\mathcal{J}^{-1}(\mathcal{O}_\rho)/G$ that we can use to carry out the orbit reduction scheme in this framework?

We will first show that there is an affirmative answer for the smooth structure of $\mathcal{J}^{-1}(\mathcal{O}_\rho)$. The main idea that we will prove in the following paragraphs is that *$\mathcal{J}^{-1}(\mathcal{O}_\rho)$ can be naturally endowed with the unique smooth structure that makes it into an initial submanifold of M*. We start with the following proposition.

14.1.1 Proposition. *Let $(M, \{\cdot, \cdot\})$ be a smooth Poisson manifold and G be a Lie group acting canonically and properly on M. Let $\mathcal{J} : M \to M/A'_G$ be the optimal momentum map associated to this action. Then,*

(i). *The generalized distribution D on M defined by $D(m) := \mathfrak{g} \cdot m + A'_G(m)$, for all $m \in M$, is integrable.*

(ii). *Let $m \in M$ be such that $\mathcal{J}(m) = \rho$, then $G^0 \cdot \mathcal{J}^{-1}(\rho)$ is the maximal integral submanifold of D going through the point m. The symbol G^0 denotes the connected component of G containing the identity.*

Proof.   (i). The distribution D can be written as the span of globally defined vector fields on M, that is,

$$D = \operatorname{span}\{\xi_M, \, X_f \mid \xi \in \mathfrak{g} \text{ and } f \in C^\infty(M)^G\}. \qquad (14.1.1)$$

By the Frobenius–Stefan–Sussman Theorem (see Stefan [1974a,b] and Sussman [1973]), the integrability of D can be proved by showing that this distribution is invariant by the flows of the vector fields in (14.1.1) that we used to generate it. Let $f, l \in C^\infty(M)^G$, $\xi, \eta \in \mathfrak{g}$, F_t be the flow of X_l, and H_t be the flow of η_M. Recall that η_M is a complete vector field such that $H_t(m) = \exp t\eta \cdot m$, for all $t \in \mathbb{R}$ and $m \in M$. Now, the integrability of A'_G guarantees that $T_m F_t \cdot X_f(m) \in A'_G(F_t(m)) \subset D(F_t(m))$. Also, the G–equivariance of F_t and the invariance of the function f imply that $T_m F_t \cdot \xi_M(m) = \xi_M(F_t(m))$ and $T_m H_t \cdot X_f(m) = X_f(H_t(m))$. Finally,

$$T_m H_t \cdot \xi_M(m) = \left.\frac{d}{ds}\right|_{s=0} \exp t\eta \exp s\xi \cdot m$$
$$= \left.\frac{d}{ds}\right|_{s=0} \exp t\eta \exp s\xi \exp -t\eta \exp t\eta \cdot m = (\operatorname{Ad}_{\exp t\eta} \xi)_M(\exp t\eta \cdot m),$$

which proves that D is integrable.

(ii) As D is integrable and is generated by the vector fields (14.1.1), its maximal integral submanifolds coincide with the orbits of the action of the pseudogroup constructed by finite composition of flows of the vector fields in (14.1.1), that is, for any $m \in M$, the integral leaf $\mathcal{L}_m$ of D that goes through m is:

$$\mathcal{L}_m = \{F_{t_1} \circ \cdots \circ F_{t_n}(m) \mid \text{with } F_{t_i} \text{ the flow of a vector field in (14.1.1)}\}.$$

Given that $[X_f, \xi_M] = 0$ for all $f \in C^\infty(M)^G$ and $\xi \in \mathfrak{g}$, the previous expression can be rewritten as

$$\mathcal{L}_m = \{H_{t_1} \circ \cdots \circ H_{t_j} \circ G_{s_1} \circ \cdots \circ G_{s_k}(m)$$
$$\mid G_{s_i} \text{ flow of } f_i \in C^\infty(M)^G, \text{ and } H_{t_i} \text{ flow of } \xi^i_M, \, \xi^i \in \mathfrak{g}\}.$$

Therefore, $\mathcal{L}_m = G^0 \cdot \mathcal{J}^{-1}(\rho)$, as required. ∎

As we already said, a general fact about integrable generalized distributions Dazord [1985] states that the smooth structure on a subset of M that

makes it into a maximal integral manifold of a given distribution coincides with the *unique* smooth structure that makes it into an initial submanifold of M. Therefore, the previous proposition shows that the sets $G^0 \cdot \mathcal{J}^{-1}(\rho)$ are initial submanifolds of M.

14.1.2 Proposition.   *Suppose that we have the same setup as in Proposition 14.1.1. If either G_ρ is closed in G or, more generally, G_ρ acts properly on $\mathcal{J}^{-1}(\rho)$, then:*

(i) *The G_ρ action on the product $G \times \mathcal{J}^{-1}(\rho)$ defined by $h \cdot (g, z) := (gh, h^{-1} \cdot z)$ is free and proper and therefore, the corresponding orbit space $G \times \mathcal{J}^{-1}(\rho)/G_\rho =: G \times_{G_\rho} \mathcal{J}^{-1}(\rho)$ is a smooth regular quotient manifold. We will denote by $\pi_{G_\rho} : G \times \mathcal{J}^{-1}(\rho) \to G \times_{G_\rho} \mathcal{J}^{-1}(\rho)$ the canonical surjective submersion.*

(ii) *The mapping $i : G \times_{G_\rho} \mathcal{J}^{-1}(\rho) \to M$ defined by $i([g, z]) := g \cdot z$ is an injective immersion onto $\mathcal{J}^{-1}(\mathcal{O}_\rho)$ such that, for any $[g, z] \in G \times_{G_\rho} \mathcal{J}^{-1}(\rho)$, $T_{[g,z]}i \cdot T_{[g,z]}(G \times_{G_\rho} \mathcal{J}^{-1}(\rho)) = D(g \cdot z)$. On other words $i(G \times_{G_\rho} \mathcal{J}^{-1}(\rho)) = \mathcal{J}^{-1}(\mathcal{O}_\rho)$ is an integral submanifold of D.*

Proof.   (i). It is easy to check that G_ρ is closed in G if and only if the action of G_ρ on G by right translations is proper. Additionally, if G_ρ is closed in G then the G_ρ–action on $\mathcal{J}^{-1}(\rho)$ is proper. In any case, if the action of G_ρ on either G, or on $\mathcal{J}^{-1}(\rho)$, or on both, is proper, so is the action on the product $G \times \mathcal{J}^{-1}(\rho)$ in the statement of the proposition. As to the freeness, it is inherited from the freeness of the G_ρ–action on G.

(ii). First of all, the map i is clearly well-defined and smooth since it is the projection onto the orbit space $G \times_{G_\rho} \mathcal{J}^{-1}(\rho)$ of the G_ρ–invariant smooth map $G \times \mathcal{J}^{-1}(\rho) \to M$ given by $(g, z) \longmapsto g \cdot z$. It is also injective because if $[g, z], [g', z'] \in G \times_{G_\rho} \mathcal{J}^{-1}(\rho)$ are such that $i([g, z]) = i([g', z'])$, then $g \cdot z = g' \cdot z'$ or, analogously, $g^{-1}g' \cdot z' = z$, which implies that $g^{-1}g' \in G_\rho$. Consequently, $[g, z] = [gg^{-1}g', (g')^{-1}g \cdot z] = [g', z']$, as required.

Finally, we check that i is an immersion. Let $[g, z] \in G \times_{G_\rho} \mathcal{J}^{-1}(\rho)$ arbitrary and $\xi \in \mathfrak{g}$, $f \in C^\infty(M)^G$ be such that

$$T_{[g,z]}i \cdot T_{(g,z)}\pi_{G_\rho} \cdot (T_e L_g(\xi), X_f(z)) = 0.$$

If we denote by F_t the flow of X_f we can rewrite this equality as

$$\left. \frac{d}{dt} \right|_{t=0} g \exp t\xi \cdot F_t(z) = 0 \quad \text{or equivalently,} \quad T_z \Phi_g(X_f(z) + \xi_M(z)) = 0.$$

Hence $X_f(z) = -\xi_M(z)$ which by (13.1.7) implies that $\xi \in \mathfrak{g}_\rho$ and therefore $T_{(g,z)}\pi_{G_\rho} \cdot (T_e L_g(\xi), X_f(z)) = T_{(g,z)}\pi_{G_\rho} \cdot (T_e L_g(\xi), -\xi_M(z)) = 0$, as required.

Given that for any $\xi \in \mathfrak{g}$, $f \in C^\infty(M)^G$, and $[g, z] \in G \times_{G_\rho} \mathcal{J}^{-1}(\rho)$ we see that $T_{[g,z]}i \cdot T_{(g,z)}\pi_{G_\rho} \cdot (T_e L_g(\xi), X_f(z)) = (\mathrm{Ad}_g \xi)_M(g \cdot z) + X_f(g \cdot z)$, it is clear that $T_{[g,z]}i \cdot T_{[g,z]}(G \times_{G_\rho} \mathcal{J}^{-1}(\rho)) = D(g \cdot z)$ and thereby $i(G \times_{G_\rho} \mathcal{J}^{-1}(\rho)) = \mathcal{J}^{-1}(\mathcal{O}_\rho)$ is an integral submanifold of D. ∎

By using the previous propositions we will now show that, in the presence of the standard hypotheses for reduction, $\mathcal{J}^{-1}(\mathcal{O}_\rho)$ is an initial submanifold of M whose connected components are the also initial submanifolds $gG^0 \cdot \mathcal{J}^{-1}(\rho)$, $g \in G$. We start with the following definition:

14.1.3 Definition.   *Let $(M, \{\cdot, \cdot\})$ be a smooth Poisson manifold and G be a Lie group acting canonically and properly on M. Let $\mathcal{J} : M \to M/A'_G$ be the optimal momentum map associated to this action and $\rho \in M/A'_G$. Suppose that G_ρ acts properly on $\mathcal{J}^{-1}(\rho)$. In these circumstances, by Proposition 14.1.2, the twist product $G \times_{G_\rho} \mathcal{J}^{-1}(\rho)$ has a canonical smooth structure. Consider in the set $\mathcal{J}^{-1}(\mathcal{O}_\rho)$ the smooth structure that makes the bijection $G \times_{G_\rho} \mathcal{J}^{-1}(\rho) \to \mathcal{J}^{-1}(\mathcal{O}_\rho)$ given by $(g, z) \to g \cdot z$ into a diffeomorphism. We will refer to this structure as the **initial smooth structure** of $\mathcal{J}^{-1}(\mathcal{O}_\rho)$.*

The following theorem justifies the choice of terminology in the previous definition and why we will be able to refer to the smooth structure there introduced as *THE* initial smooth structure of $\mathcal{J}^{-1}(\mathcal{O}_\rho)$.

14.1.4 Theorem.   *Suppose that we are in the same setup as in Definition 14.1.3. Then, the set $\mathcal{J}^{-1}(\mathcal{O}_\rho)$ endowed with the initial smooth structure is an actual initial submanifold of M that can be decomposed as a disjoint union of connected components as*

$$\mathcal{J}^{-1}(\mathcal{O}_\rho) = \dot{\bigcup_{[g] \in G/(G^0 G_\rho)}} gG^0 \cdot \mathcal{J}^{-1}(\rho). \qquad (14.1.2)$$

Each connected component of $\mathcal{J}^{-1}(\mathcal{O}_\rho)$ is a maximal integral submanifold of the distribution D defined in Proposition 14.1.1. If, additionally, the subgroup G_ρ is closed in G, the topology on $\mathcal{J}^{-1}(\mathcal{O}_\rho)$ induced by its initial smooth structure coincides with the initial topology induced by the map $\mathcal{J}_{\mathcal{J}^{-1}(\mathcal{O}_\rho)} : \mathcal{J}^{-1}(\mathcal{O}_\rho) \to \mathcal{O}_\rho$ given by $z \longmapsto \mathcal{J}(z)$, where the orbit $\mathcal{O}_\rho$ is endowed with the smooth structure coming from the homogeneous manifold G/G_ρ. Finally, notice that (14.1.2) implies that $\mathcal{J}^{-1}(\mathcal{O}_\rho)$ has as many connected components as the cardinality of the homogeneous manifold $G/(G^0 G_\rho)$.

Proof.   First of all notice that the sets $gG^0 \cdot \mathcal{J}^{-1}(\rho)$ are clearly maximal integral submanifolds of D by part (ii) in Proposition 14.1.1. As a corollary of this, they are the connected components of $\mathcal{J}^{-1}(\mathcal{O}_\rho)$ endowed with the smooth structure in Definition 14.1.4. Indeed, let S be the connected component of $\mathcal{J}^{-1}(\mathcal{O}_\rho)$ that contains $gG^0 \cdot \mathcal{J}^{-1}(\rho)$, that is, $gG^0 \cdot \mathcal{J}^{-1}(\rho) \subset S \subset \mathcal{J}^{-1}(\mathcal{O}_\rho)$. As $\mathcal{J}^{-1}(\mathcal{O}_\rho)$ is a manifold, it is locally connected, and therefore its connected components are open and closed. In particular, since S is an open connected subset of $\mathcal{J}^{-1}(\mathcal{O}_\rho)$, part (ii) in Proposition 14.1.2 shows that S is a connected integral submanifold of D. By the maximality of $gG^0 \cdot \mathcal{J}^{-1}(\rho)$ as an integral submanifold of D,

$gG^0 \cdot \mathcal{J}^{-1}(\rho) = S$, necessarily. The set $gG^0 \cdot \mathcal{J}^{-1}(\rho)$ is therefore a connected component of $\mathcal{J}^{-1}(\mathcal{O}_\rho)$. As it is a leaf of a smooth integrable distribution on M, it is also an initial submanifold of M Dazord [1985] of dimension $d = \dim \mathcal{J}^{-1}(\mathcal{O}_\rho) = \dim G + \dim \mathcal{J}^{-1}(\rho) - \dim G_\rho$.

We now show that $\mathcal{J}^{-1}(\mathcal{O}_\rho)$ with the smooth structure in Definition 14.1.4 is an initial submanifold of M. First of all part (ii) in Proposition 14.1.2 shows that $\mathcal{J}^{-1}(\mathcal{O}_\rho)$ is an injectively immersed submanifold of M. The initial character can be obtained as a consequence of the fact that its connected components are initial together with the following elementary lemma:

14.1.5 Lemma.   *Let N be an injectively immersed submanifold of the smooth manifold M. Suppose that N can be written as the disjoint union of a family $\{S_\alpha\}_{\alpha \in I}$ of open subsets of N such that each S_α is an initial submanifold of M. Then, N is initial.*

Proof.   Let $i_N : N \hookrightarrow M$ and $i_\alpha : S_\alpha \hookrightarrow N$ be the injections. Let Z be an arbitrary smooth manifold and $f : Z \to M$ be a smooth map such that $f(Z) \subset N$. As the sets S_α are open and partition N, the manifold Z can be written as a disjoint union of open sets $Z_\alpha := f^{-1}(S_\alpha)$, that is

$$Z = \bigcup_{\alpha \in I}^{\cdot} f^{-1}(S_\alpha).$$

Given that for each index α the map $f_\alpha : Z_\alpha \to M$ obtained by restriction of f to Z_α is smooth, the corresponding map $\bar{f}_\alpha : Z_\alpha \to S_\alpha$ defined by the identity $i_\alpha \circ \bar{f}_\alpha = f_\alpha$ is also smooth by the initial character of S_α. Let $\bar{f} : Z \to N$ be the map obtained by union of the mappings $\bar{f}_\alpha$. This map is smooth and satisfies that $i_N \circ \bar{f} = f$ which proves that N is initial.   ▼

We now prove Expression (14.1.2). First of all notice that as G^0 is normal in G, the set $G^0 G_\rho$ is a (possibly non-closed) subgroup of G. We obviously have that

$$\mathcal{J}^{-1}(\mathcal{O}_\rho) = \bigcup_{g \in G} gG^0 \mathcal{J}^{-1}(\rho). \qquad (14.1.3)$$

Moreover, if g and $g' \in G$ are such that $[g] = [g'] \in G/(G^0 G_\rho)$ then we can write that $g' = ghk$ with $h \in G^0$ and $k \in G_\rho$. Consequently, $g' G^0 \mathcal{J}^{-1}(\rho) = ghk G^0 \mathcal{J}^{-1}(\rho) = gh(G^0 k)\mathcal{J}^{-1}(\rho) = g(hG^0)(k\mathcal{J}^{-1}(\rho)) = gG^0 \mathcal{J}^{-1}(\rho)$, which implies that (14.1.3) can be refined to

$$\mathcal{J}^{-1}(\mathcal{O}_\rho) = \bigcup_{[g] \in G/(G^0 G_\rho)} gG^0 \mathcal{J}^{-1}(\rho). \qquad (14.1.4)$$

It only remains to be shown that this union is disjoint: let $gh \cdot z = lh' \cdot z'$ with $h, h' \in G^0$ and $z, z' \in \mathcal{J}^{-1}(\rho)$. If we apply $\mathcal{J}$ to both sides of this equality we obtain that $gh \cdot \rho = lh' \cdot \rho$. Hence, $(h')^{-1}l^{-1}gh \in G_\rho$ and

$l^{-1}g \in h'G_\rho h^{-1} \subset G^0 G_\rho$. This implies that $[l] = [g] \in G/(G^0 G_\rho)$ and $gG^0 \mathcal{J}^{-1}(\rho) = lG^0 \mathcal{J}^{-1}(\rho)$, as required.

We finally show that when G_ρ is closed in G, the topology on $\mathcal{J}^{-1}(\mathcal{O}_\rho)$ induced by its initial smooth structure coincides with the initial topology induced by the map $\mathcal{J}_{\mathcal{J}^{-1}(\mathcal{O}_\rho)} : \mathcal{J}^{-1}(\mathcal{O}_\rho) \to \mathcal{O}_\rho$ on $\mathcal{J}^{-1}(\mathcal{O}_\rho)$. Recall first that this topology is characterized by the fact that for any topological space Z and any map $\phi : Z \to \mathcal{J}^{-1}(\mathcal{O}_\rho)$ the map $\phi : Z \to \mathcal{J}^{-1}(\mathcal{O}_\rho)$ is continuous if and only if $\mathcal{J}_{\mathcal{J}^{-1}(\mathcal{O}_\rho)} \circ \phi$ is continuous. Moreover, as the family

$$\{\mathcal{J}^{-1}_{\mathcal{J}^{-1}(\mathcal{O}_\rho)}(U) \mid U \text{ open subset of } \mathcal{O}_\rho\}$$

is a subbase of this topology, the initial topology on $\mathcal{J}^{-1}(\mathcal{O}_\rho)$ induced by the map $\mathcal{J}_{\mathcal{J}^{-1}(\mathcal{O}_\rho)}$ is first countable. We prove that this topology coincides with the topology induced by the initial smooth structure on $\mathcal{J}^{-1}(\mathcal{O}_\rho)$ by showing that the map

$$f : G \times_{G_\rho} \mathcal{J}^{-1}(\rho) \to \mathcal{J}^{-1}(\mathcal{O}_\rho), \quad \text{where} \quad f([g,z]) := g \cdot z$$

is a homeomorphism when we consider $\mathcal{J}^{-1}(\mathcal{O}_\rho)$ as a topological space with the initial topology induced by $\mathcal{J}_{\mathcal{J}^{-1}(\mathcal{O}_\rho)}$. Indeed, f is continuous if and only if the map $G \times_{G_\rho} \mathcal{J}^{-1}(\rho) \to \mathcal{O}_\rho$ given by $[g,z] \mapsto g \cdot \rho$ is continuous, which in turn is equivalent to the continuity of the map $G \times \mathcal{J}^{-1}(\rho) \to G/G_\rho$ defined by $(g,z) \longmapsto gG_\rho$, which is true. We now show that the inverse

$$f^{-1} : \mathcal{J}^{-1}(\mathcal{O}_\rho) \to G \times_{G_\rho} \mathcal{J}^{-1}(\rho)$$

of f given by $g \cdot z \mapsto [g,z]$ is continuous. Since the initial topology on $\mathcal{J}^{-1}(\mathcal{O}_\rho)$ induced by $\mathcal{J}_{\mathcal{J}^{-1}(\mathcal{O}_\rho)}$ is first countable it suffices to show that for any convergent sequence $\{z_n\} \subset \mathcal{J}^{-1}(\mathcal{O}_\rho) \to z \in \mathcal{J}^{-1}(\mathcal{O}_\rho)$, we have

$$\lim_{n \to \infty} f^{-1}(z_n) = f^{-1}(\lim_{n \to \infty} z_n) = f^{-1}(z).$$

Indeed, as $\mathcal{J}_{\mathcal{J}^{-1}(\mathcal{O}_\rho)}$ is continuous, the sequence $\{\mathcal{J}(z_n) = g_n \cdot \rho\} \subset \mathcal{O}_\rho$ converges in $\mathcal{O}_\rho$ to $\mathcal{J}(z) = g \cdot \rho$, for some $g \in G$. Let $j : \mathcal{O}_\rho \to G/G_\rho$ be the standard diffeomorphism and $\sigma : U_{gG_\rho} \subset G/G_\rho \to G$ be a local smooth section of the submersion $G \to G/G_\rho$ in a neighborhood U_{gG_ρ} of $gG_\rho \in G/G_\rho$. Let $V := \mathcal{J}^{-1}_{\mathcal{J}^{-1}(\mathcal{O}_\rho)}(j^{-1}(U_{gG_\rho}))$. V is an open neighborhood of z in $\mathcal{J}^{-1}(\mathcal{O}_\rho)$ because

$$j \circ \mathcal{J}_{\mathcal{J}^{-1}(\mathcal{O}_\rho)}(z) = j(g \cdot \rho) = gG_\rho \in U_{gG_\rho}.$$

We now notice that for any $m \in V$ we can write

$$f^{-1}(m) = [\sigma \circ j \circ \mathcal{J}_{\mathcal{J}^{-1}(\mathcal{O}_\rho)}(m), (\sigma \circ j \circ \mathcal{J}_{\mathcal{J}^{-1}(\mathcal{O}_\rho)}(m))^{-1} \cdot m].$$

Consequently, since

$$\lim_{n \to \infty} f^{-1}(z_n) = \lim_{n \to \infty} [\sigma \circ j \circ \mathcal{J}_{\mathcal{J}^{-1}(\mathcal{O}_\rho)}(z_n), (\sigma \circ j \circ \mathcal{J}_{\mathcal{J}^{-1}(\mathcal{O}_\rho)}(z_n))^{-1} \cdot z_n]$$

$$= [\sigma \circ j \circ \mathcal{J}_{\mathcal{J}^{-1}(\mathcal{O}_\rho)}(z), (\sigma \circ j \circ \mathcal{J}_{\mathcal{J}^{-1}(\mathcal{O}_\rho)}(z))^{-1} \cdot z] = f^{-1}(z),$$

the continuity of f^{-1} is guaranteed. ∎

14.2 The Symplectic Orbit Reduction Quotient

We will know show that the quotient $\mathcal{J}^{-1}(\mathcal{O}_\rho)/G$ can be endowed with a smooth structure that makes it into a regular quotient manifold, that is, the projection $\pi_{\mathcal{O}_\rho} : \mathcal{J}^{-1}(\mathcal{O}_\rho) \to \mathcal{J}^{-1}(\mathcal{O}_\rho)/G$ is a smooth submersion. We will carry this out under the same hypotheses present in Definition 14.1.3, that is, G_ρ acts properly on $\mathcal{J}^{-1}(\rho)$.

First of all notice that as $\mathcal{J}^{-1}(\mathcal{O}_\rho)$ is an initial G–invariant submanifold of M, the G–action on $\mathcal{J}^{-1}(\mathcal{O}_\rho)$ is smooth. We will prove that $\mathcal{J}^{-1}(\mathcal{O}_\rho)/G$ is a regular quotient manifold by showing that this action is actually proper and satisfies that all the isotropy subgroups are conjugate to a given one. Indeed, recall that the initial manifold structure on $\mathcal{J}^{-1}(\mathcal{O}_\rho)$ is the one that makes it G–equivariantly diffeomorphic to the twist product $G \times_{G_\rho} \mathcal{J}^{-1}(\rho)$ when we take in this space the G–action given by the expression $g \cdot [h, z] := [gh, z]$, $g \in G$, $[h, z] \in G \times_{G_\rho} \mathcal{J}^{-1}(\rho)$. Therefore, it suffices to show that this G–action has the desired properties. First of all this action is proper since a general property about twist products (see [HRed]) says that the G–action on $G \times_{G_\rho} \mathcal{J}^{-1}(\rho)$ is proper if and only if the G_ρ–action on $\mathcal{J}^{-1}(\rho)$ is proper, which we supposed as a hypothesis. We now look at the isotropies of this action: in Proposition 13.2.1 we saw that all the elements in $\mathcal{J}^{-1}(\rho)$ have the same G–isotropy, call it H. As $H \subset G_\rho$, this is also their G_ρ–isotropy. Now, using a standard property of the isotropies of twist products (see [HRed]), we have

$$G_{[g,z]} = g(G_\rho)_z g^{-1} = gHg^{-1},$$

for any $[g, z] \in G \times_{G_\rho} \mathcal{J}^{-1}(\rho)$, as required.

The quotient manifold $\mathcal{J}^{-1}(\mathcal{O}_\rho)/G$ is naturally diffeomorphic to the symplectic point reduced space. Indeed,

$$\mathcal{J}^{-1}(\mathcal{O}_\rho)/G \simeq G \times_{G_\rho} \mathcal{J}^{-1}(\rho)/G \simeq \mathcal{J}^{-1}(\rho)/G_\rho.$$

This diffeomorphism can be explicitly implemented as follows. Let $l_\rho : \mathcal{J}^{-1}(\rho) \to \mathcal{J}^{-1}(\mathcal{O}_\rho)$ be the inclusion. As the inclusion $\mathcal{J}^{-1}(\rho) \hookrightarrow M$ is smooth and $\mathcal{J}^{-1}(\mathcal{O}_\rho)$ is initial l_ρ is smooth. Also, since l_ρ is (G_ρ, G) equivariant it drops to a unique smooth map $L_\rho : \mathcal{J}^{-1}(\rho)/G_\rho \to \mathcal{J}^{-1}(\mathcal{O}_\rho)/G$ that makes the following diagram

$$
\begin{array}{ccc}
\mathcal{J}^{-1}(\rho) & \xrightarrow{\;\; l_\rho \;\;} & \mathcal{J}^{-1}(\mathcal{O}_\rho) \\[2pt]
{\scriptstyle \pi_\rho}\big\downarrow & & \big\downarrow{\scriptstyle \pi_{\mathcal{O}_\rho}} \\[2pt]
\mathcal{J}^{-1}(\rho)/G_\rho & \xrightarrow{\;\; L_\rho \;\;} & \mathcal{J}^{-1}(\mathcal{O}_\rho)/G.
\end{array}
$$

commutative. L_ρ is a smooth bijection. In order to show that its inverse is also smooth we will think of $\mathcal{J}^{-1}(\mathcal{O}_\rho)$ as $G \times_{G_\rho} \mathcal{J}^{-1}(\rho)$. First of all

notice that the projection $G \times \mathcal{J}^{-1}(\rho) \to \mathcal{J}^{-1}(\rho)$ is G_ρ–(anti)equivariant and therefore induces a smooth map $G \times_{G_\rho} \mathcal{J}^{-1}(\rho) \to \mathcal{J}^{-1}(\rho)/G_\rho$ given by $[g, z] \mapsto [z]$, $[g, z] \in G \times_{G_\rho} \mathcal{J}^{-1}(\rho)$. This map is G–invariant and therefore drops to another smooth mapping $G \times_{G_\rho} \mathcal{J}^{-1}(\rho)/G \to \mathcal{J}^{-1}(\rho)/G_\rho$ that coincides with L_ρ^{-1}, the inverse of L_ρ, which is consequently a diffeomorphism.

The orbit reduced space $\mathcal{J}^{-1}(\mathcal{O}_\rho)/G$ can be therefore trivially endowed with a symplectic structure $\omega_{\mathcal{O}_\rho}$ by defining $\omega_{\mathcal{O}_\rho} := (L_\rho^{-1})^* \omega_\rho$. We put together all the facts that we just proved in the following theorem–definition:

14.2.1 Theorem (Optimal Orbit Reduction by Poisson Actions). *Suppose that $(M, \{\cdot, \cdot\})$ is a smooth Poisson manifold and G is a Lie group acting canonically and properly on M. Let $\mathcal{J} : M \to M/A'_G$ be the optimal momentum map associated to this action and $\rho \in M/A'_G$. Suppose that G_ρ acts properly on $\mathcal{J}^{-1}(\rho)$. If we denote $\mathcal{O}_\rho := G \cdot \rho$, then:*

(i) *There is a unique smooth structure on $\mathcal{J}^{-1}(\mathcal{O}_\rho)$ that makes it into an initial submanifold of M.*

(ii) *The G–action on $\mathcal{J}^{-1}(\mathcal{O}_\rho)$ by restriction of the G–action on M is smooth and proper and all its isotropy subgroups are conjugate to a given compact isotropy subgroup of the G–action on M.*

(iii) *The quotient $M_{\mathcal{O}_\rho} := \mathcal{J}^{-1}(\mathcal{O}_\rho)/G$ admits a unique smooth structure that makes the projection $\pi_{\mathcal{O}_\rho} : \mathcal{J}^{-1}(\mathcal{O}_\rho) \to \mathcal{J}^{-1}(\mathcal{O}_\rho)/G$ a surjective submersion.*

(iv) *The quotient $M_{\mathcal{O}_\rho} := \mathcal{J}^{-1}(\mathcal{O}_\rho)/G$ admits a unique symplectic structure $\omega_{\mathcal{O}_\rho}$ that makes it symplectomorphic to the point reduced space M_ρ. We will refer to the pair $(M_{\mathcal{O}_\rho}, \omega_{\mathcal{O}_\rho})$ as the **(optimal) orbit reduced space** of $(M, \{\cdot, \cdot\})$ at $\mathcal{O}_\rho$.*

In this setup we can easily formulate an analog of Theorem 13.5.3.

14.2.2 Theorem (Optimal orbit reduction of G–equivariant Poisson dynamics). *Let $(M, \{\cdot, \cdot\})$ be a smooth Poisson manifold and G be a Lie group acting canonically and properly on M. Let $\mathcal{J} : M \to M/A'_G$ be the optimal momentum map associated and $\rho \in M/A'_G$ be such that G_ρ acts properly on $\mathcal{J}^{-1}(\rho)$. Let $h \in C^\infty(M)^G$ be a G–invariant function on M and X_h be the associated G–equivariant Hamiltonian vector field on M. Then,*

(i) *The flow F_t of X_h leaves $\mathcal{J}^{-1}(\mathcal{O}_\rho)$ invariant, commutes with the G–action, and therefore induces a flow $F_t^{\mathcal{O}_\rho}$ on $M_{\mathcal{O}_\rho}$ uniquely determined by the relation*

$$\pi_{\mathcal{O}_\rho} \circ F_t \circ i_{\mathcal{O}_\rho} = F_t^{\mathcal{O}_\rho} \circ \pi_{\mathcal{O}_\rho},$$

where $i_{\mathcal{O}_\rho} : \mathcal{J}^{-1}(\mathcal{O}_\rho) \hookrightarrow M$ is the inclusion.

(ii) *The flow $F_t^{\mathcal{O}_\rho}$ in $(M_{\mathcal{O}_\rho}, \omega_{\mathcal{O}_\rho})$ is Hamiltonian with the Hamiltonian function $h_{\mathcal{O}_\rho} \in C^\infty(M_{\mathcal{O}_\rho})$ given by the equality $h_{\mathcal{O}_\rho} \circ \pi_{\mathcal{O}_\rho} = h \circ i_{\mathcal{O}_\rho}$.*

(iii) *Let $k \in C^\infty(M)^G$ be another G–invariant function on M and $\{\cdot, \cdot\}_{\mathcal{O}_\rho}$ be the Poisson bracket associated to the symplectic form $\omega_{\mathcal{O}_\rho}$ on $M_{\mathcal{O}_\rho}$. Then, $\{h, k\}_{\mathcal{O}_\rho} = \{h_{\mathcal{O}_\rho}, k_{\mathcal{O}_\rho}\}_{\mathcal{O}_\rho}$.*

We conclude this section with a brief description of the orbit version of the regularized reduced spaces introduced in Definition 13.6.2 for the symplectic case. If we follow the prescription introduced in Section 14.1 using the L^ρ–action on M_H^ρ we are first supposed to study the set $\mathcal{J}_{L^\rho}^{-1}(L^\rho \cdot \sigma)$. The initial smooth structure on this set induced by the twist product $L^\rho \times_{L_\sigma^\rho} \mathcal{J}_{L^\rho}^{-1}(\sigma)$ makes it into an initial submanifold of M_H^ρ. Moreover, if we use the statements in Proposition 13.6.2 it is easy to see that

$$\mathcal{J}_{L^\rho}^{-1}(L^\rho \cdot \sigma) = L^\rho \cdot \mathcal{J}_{L^\rho}^{-1}(\sigma) = N(H)^\rho \cdot \mathcal{J}^{-1}(\rho) = \mathcal{J}^{-1}(\mathcal{N}_\rho),$$

with $\mathcal{N}_\rho := N(H)^\rho \cdot \rho \subset M/A_G'$.

The set $\mathcal{J}_{L^\rho}^{-1}(L^\rho \cdot \sigma) = \mathcal{J}^{-1}(\mathcal{N}_\rho)$ is an embedded submanifold of $\mathcal{J}^{-1}(\mathcal{O}_\rho)$ (since $\mathcal{J}^{-1}(\mathcal{N}_\rho) \simeq N(H)^\rho \times_{G_\rho} \mathcal{J}^{-1}(\rho)$ is embedded in $G \times_{G_\rho} \mathcal{J}^{-1}(\rho) \simeq \mathcal{J}^{-1}(\mathcal{O}_\rho)$). Moreover, a simple diagram chasing shows that the symplectic quotient $(\mathcal{J}_{L^\rho}^{-1}(L^\rho \cdot \sigma)/L^\rho, (\omega|_{M_H^\rho})_{L^\rho \cdot \sigma})$ is naturally symplectomorphic to the orbit reduced space $(\mathcal{J}^{-1}(\mathcal{O}_\rho)/G, \omega_{\mathcal{O}_\rho})$. We will say that $(\mathcal{J}_{L^\rho}^{-1}(L^\rho \cdot \sigma)/L^\rho, (\omega|_{M_H^\rho})_{L^\rho \cdot \sigma})$ is **an orbit regularization** of $(\mathcal{J}^{-1}(\mathcal{O}_\rho)/G, \omega_{\mathcal{O}_\rho})$.

We finally show that

$$\mathcal{J}^{-1}(\mathcal{O}_\rho) = \dot{\bigcup}_{[g] \in G/N(H)^\rho} \mathcal{J}^{-1}(\mathcal{N}_{g \cdot \rho}). \tag{14.2.1}$$

The equality is a straightforward consequence of the fact that for any $g \in G$,

$$M_{gHg^{-1}}^{g\rho} = \Phi_g(M_H^\rho),$$
$$N(gHg^{-1})^{g\rho} = gN(H)^\rho g^{-1},$$
$$\mathcal{J}^{-1}(\mathcal{N}_{g \cdot \rho}) = gN(H)^\rho \mathcal{J}^{-1}(\rho).$$

The last relation implies that if $g, g' \in G$ are such that $[g] = [g'] \in G/N(H)^\rho$, then $\mathcal{J}^{-1}(\mathcal{N}_{g \cdot \rho}) = \mathcal{J}^{-1}(\mathcal{N}_{g' \cdot \rho})$. We now show that the union in (14.2.1) is indeed disjoint: let $gn \cdot z \in \mathcal{J}^{-1}(\mathcal{N}_{g \cdot \rho})$ and $g'n' \cdot z' \in \mathcal{J}^{-1}(\mathcal{N}_{g' \cdot \rho})$ be such that $gn \cdot z = g'n' \cdot z'$, with $g, g' \in G$, $n, n' \in N(H)^\rho$, and $z, z' \in \mathcal{J}^{-1}(\rho)$. Since $gn \cdot z = g'n' \cdot z'$, we necessarily have that $G_{gn \cdot z} = G_{g'n' \cdot z'}$ which implies that $gHg^{-1} = g'H(g')^{-1}$, and hence $g^{-1}g' \in N(H)$. We now recall that M_H^ρ is the accessible set going through z or z' of the integrable generalized distribution B_G' defined by

$$B_G' := \mathrm{span}\{X \in \mathfrak{X}(U)^G \mid U \text{ open } G\text{–invariant set in } M\},$$

where the symbol $\mathfrak{X}(U)^G$ denotes the set of G–equivariant vector fields defined on U. Let $\mathcal{B}_G'$ be the pseudogroup of transformations of M consisting

of the G–equivariant flows of the vector fields that span B'_G. Now, as the points $n \cdot z, n' \cdot z' \in M^\rho_H$, there exists $\mathcal{F}_T \in \mathcal{B}'_G$ such that $n' \cdot z' = \mathcal{F}_T(n \cdot z)$, hence $(g')^{-1}gn \cdot z = \mathcal{F}_T(n \cdot z)$. Moreover, as any element in M^ρ_H can be written as $\mathcal{G}_T(n \cdot z)$ with $\mathcal{G}_T \in \mathcal{B}'_G$, we have

$$(g')^{-1}g \cdot \mathcal{G}_T(n \cdot z) = \mathcal{G}_T((g')^{-1}gn \cdot z) = \mathcal{G}_T(\mathcal{F}_T(n \cdot z)) \in M^\rho_H,$$

which implies that $(g')^{-1}g \in N(H)^\rho$ and therefore $[g] = [g'] \in G/N(H)^\rho$, as required.

14.3   The Polar Reduced Spaces

As we already pointed out, the standard theory of orbit reduction provides a characterization of the symplectic form of the orbit reduced spaces in terms of the symplectic structures of the corresponding coadjoint orbits that, from the dual pairs point of view, play the role of the symplectic leaves of the Poisson manifold in duality, namely $\mathbf{J}(M) \subset \mathfrak{g}^*$.

  We will now show that when the group of Poisson transformations A_G is von Neumann (actually we just need weakly von Neumann), that is, when the diagram

$$(M/G, \{\cdot, \cdot\}_{M/A_G}) \overset{\pi_{A_G}}{\leftarrow} (M, \{\cdot, \cdot\}) \overset{\mathcal{J}}{\to} (M/A'_G, \{\cdot, \cdot\}_{M/A'_G})$$

is a dual pair in the sense of Definition 13.3.3, the classical picture can be reproduced in this context. More specifically, in this section we will show that:

- The symplectic leaves of $(M/A'_G, \{\cdot, \cdot\}_{M/A'_G})$ admit a smooth presymplectic structure that generalizes the Kostant–Kirillov–Souriau symplectic structure in the coadjoint orbits of the dual of a Lie algebra in the sense that they are homogeneous presymplectic manifolds. We will refer to these "generalized coadjoint orbits" as **polar reduced spaces**.

- The presymplectic structure of the polar reduced spaces is related to the symplectic form of the orbit reduced spaces introduced in the previous section via an equality that holds strong resemblance with the classical expression (13.5.1). Also, it is possible to provide a very explicit characterization of the situations in which the polar reduced spaces are actually symplectic.

- When the manifold M is symplectic, the polar reduced space decomposes as a union of embedded symplectic submanifolds that correspond to the polar reduced spaces of the regularizations of the orbit reduced space. Each of these symplectic manifolds is a homogeneous manifold and we will refer to them as the **regularized polar reduced subspaces**.

We start with a proposition that spells out the smooth structure of the polar reduced spaces. ***In this section*** we use a stronger hypothesis on G_ρ with respect to the one we used in the previous section, namely, ***we will assume that G_ρ is closed in G*** which, as we point out in the proof of Proposition 14.1.2, ***implies that the G_ρ action on $\mathcal{J}^{-1}(\rho)$ is proper***.

14.3.1 Proposition.   *Let $(M, \{\cdot, \cdot\})$ be a smooth Poisson manifold and G be a Lie group acting canonically and properly on M. Let $\mathcal{J} : M \to M/A'_G$ be the optimal momentum map associated to this action and $\rho \in M/A'_G$. Suppose that G_ρ is closed in G. Then, the polar distribution A'_G restricts to a smooth integrable regular distribution on $\mathcal{J}^{-1}(\mathcal{O}_\rho)$, that we will also denote by A'_G. The leaf space $M'_{\mathcal{O}_\rho} := \mathcal{J}^{-1}(\mathcal{O}_\rho)/A'_G$ admits a unique smooth structure that makes it into a regular quotient manifold and diffeomorphic to the homogeneous manifold G/G_ρ. With this smooth structure the projection*

$$\mathcal{J}_{\mathcal{O}_\rho} : \mathcal{J}^{-1}(\mathcal{O}_\rho) \to \mathcal{J}^{-1}(\mathcal{O}_\rho)/A'_G$$

*is a smooth surjective submersion. We will refer to $M'_{\mathcal{O}_\rho}$ as the **polar reduced space**.*

Proof.   Let $m \in \mathcal{J}^{-1}(\mathcal{O}_\rho)$. By Proposition 14.1.2 we have $T_m \mathcal{J}^{-1}(\mathcal{O}_\rho) = D(m) = \mathfrak{g} \cdot m + A'_G(m)$, which implies that the restriction of A'_G to $\mathcal{J}^{-1}(\mathcal{O}_\rho)$ is tangent to it. Consequently, as $\mathcal{J}^{-1}(\mathcal{O}_\rho)$ is an immersed submanifold of M, there exists for each Hamiltonian vector field $X_f \in \mathfrak{X}(M)$, $f \in C^\infty(M)^G$, a vector field $X'_f \in \mathfrak{X}(\mathcal{J}^{-1}(\mathcal{O}_\rho))$ such that

$$T i_{\mathcal{O}_\rho} \circ X'_f = X_f \circ i_{\mathcal{O}_\rho},$$

with $i_{\mathcal{O}_\rho} : \mathcal{J}^{-1}(\mathcal{O}_\rho) \hookrightarrow M$ the injection. The restriction $A'_G|_{\mathcal{J}^{-1}(\mathcal{O}_\rho)}$ of A'_G to $\mathcal{J}^{-1}(\mathcal{O}_\rho)$ is generated by the vector fields of the form X'_f and it is therefore smooth. It is also integrable since for any point $m = g \cdot z \in \mathcal{J}^{-1}(\mathcal{O}_\rho)$, $z \in \mathcal{J}^{-1}(\rho)$, the embedded submanifold $\mathcal{J}^{-1}(g \cdot \rho)$ of $\mathcal{J}^{-1}(\mathcal{O}_\rho)$ is the maximal integral submanifold of $A'_G|_{\mathcal{J}^{-1}(\mathcal{O}_\rho)}$. This is so because the flows F_t and F'_t of X_f and X'_f, respectively, satisfy that $i_{\mathcal{O}_\rho} \circ F'_t = F_t \circ i_{\mathcal{O}_\rho}$. It is then clear that $A'_G|_{\mathcal{J}^{-1}(\mathcal{O}_\rho)}$ has constant rank since $\dim A'_G|_{\mathcal{J}^{-1}(\mathcal{O}_\rho)} = \dim \mathcal{J}^{-1}(\rho)$. This all shows that the leaf space $\mathcal{J}^{-1}(\mathcal{O}_\rho)/A'_G$ is well-defined.

In order to show that the leaf space $\mathcal{J}^{-1}(\mathcal{O}_\rho)/A'_G$ is a regular quotient manifold we first notice that

$$\mathcal{J}^{-1}(\mathcal{O}_\rho)/A'_G \simeq (G \times_{G_\rho} \mathcal{J}^{-1}(\rho))/A'_G$$

is in bijection with the quotient G/G_ρ that, by the hypothesis on the closedness of G_ρ is a smooth homogeneous manifold. Take in $M'_{\mathcal{O}_\rho} := \mathcal{J}^{-1}(\mathcal{O}_\rho)/A'_G$ the smooth structure that makes the bijection with G/G_ρ a diffeomorphism. It turns out that that smooth structure is the unique

one that makes $M'_{\mathcal{O}_\rho}$ into a regular quotient manifold since it can be readily verified that the map

$$\mathcal{J}_{\mathcal{O}_\rho}: \quad \mathcal{J}^{-1}(\mathcal{O}_\rho) \simeq G \times_{G_\rho} \mathcal{J}^{-1}(\rho) \quad \longrightarrow \quad \mathcal{J}^{-1}(\mathcal{O}_\rho)/A'_G \simeq G/G_\rho$$
$$[g, z] \quad\qquad \longmapsto \qquad\quad gG_\rho$$

is a surjective submersion. ∎

We now introduce the regularized polar reduced subspaces of $M'_{\mathcal{O}_\rho}$, available when M is symplectic. We retake the ideas and notations introduced just above (14.2.1). Let $(\mathcal{J}_{L^\rho}^{-1}(L^\rho \cdot \sigma)/L^\rho, (\omega|_{M_H^\rho})_{L^\rho \cdot \sigma})$ be an orbit regularization of $(\mathcal{J}^{-1}(\mathcal{O}_\rho)/G, \omega_{\mathcal{O}_\rho})$. A straightforward application of Proposition 13.6.1 implies that the reduced space polar to

$$(\mathcal{J}_{L^\rho}^{-1}(L^\rho \cdot \sigma)/L^\rho, (\omega|_{M_H^\rho})_{L^\rho \cdot \sigma})$$

equals

$$\mathcal{J}_{L^\rho}^{-1}(L^\rho \cdot \sigma)/A'_{L^\rho} = \mathcal{J}^{-1}(\mathcal{N}_\rho)/A'_G$$

which is naturally diffeomorphic to $N(H)^\rho/G_\rho$. We will say that

$$\mathcal{J}^{-1}(\mathcal{N}_\rho)/A'_G$$

is **a regularized polar reduced subspace** of $M'_{\mathcal{O}_\rho}$. We will write

$$M'_{\mathcal{N}_\rho} := \mathcal{J}^{-1}(\mathcal{N}_\rho)/A'_G$$

and denote by $\mathcal{J}_{\mathcal{N}_\rho}: \mathcal{J}^{-1}(\mathcal{N}_\rho) \to \mathcal{J}^{-1}(\mathcal{N}_\rho)/A'_G$ the canonical projection. Notice that the spaces $M'_{\mathcal{N}_\rho}$ are embedded submanifolds of $M'_{\mathcal{O}_\rho}$. Finally, the decomposition (14.2.1) implies that the polar reduced space can be written as the following disjoint union of regularized polar reduced subspaces:

$$M'_{\mathcal{O}_\rho} = \mathcal{J}^{-1}(\mathcal{O}_\rho)/A'_G$$
$$= \dot{\bigcup}_{[g]\in G/N(H)^\rho} \mathcal{J}^{-1}(\mathcal{N}_{g\cdot\rho})/A'_G$$
$$= \dot{\bigcup}_{[g]\in G/N(H)^\rho} M'_{\mathcal{N}_{g\cdot\rho}}. \qquad (14.3.1)$$

Equivalently, we have

$$G/G_\rho = \dot{\bigcup}_{[g]\in G/N(H)^\rho} gN(H)^\rho/G_\rho, \qquad (14.3.2)$$

where the quotient $gN(H)^\rho/G_\rho$ denotes the orbit space of the free and proper action of G_ρ on $gN(H)^\rho$ by $h \cdot gn := gnh$, $h \in G_\rho$, $n \in N(H)^\rho$.

Before we state our next result we need some terminology. We will denote by $C^\infty\left(\mathcal{J}^{-1}(\mathcal{O}_\rho)/A'_G\right)$ the set of smooth real valued functions on $M'_{\mathcal{O}_\rho}$ with

the smooth structure introduced in Proposition 14.3.1. Recall now that, as we pointed out in (13.1.3), there is a notion of smooth function on M/A'_G, namely

$$C^\infty(M/A'_G) := \{f \in C^0(M/A'_G) \mid f \circ \mathcal{J} \in C^\infty(M)^{A'_G}\}.$$

Analogously, for each open A'_G–invariant subset U of M we can define

$$C^\infty(U/A'_G) := \{f \in C^0(U/A'_G) \mid f \circ \mathcal{J}|_U \in C^\infty(U)^{A'_G}\}.$$

We define the set of **Whitney smooth functions** $W^\infty\left(\mathcal{J}^{-1}(\mathcal{O}_\rho)/A'_G\right)$ on $\mathcal{J}^{-1}(\mathcal{O}_\rho)/A'_G$ as

$$
\begin{aligned}
W^\infty &\left(\mathcal{J}^{-1}(\mathcal{O}_\rho)/A'_G\right) \\
&:= \{f : M'_{\mathcal{O}_\rho} \to \mathbb{R} \mid f = F|_{M'_{\mathcal{O}_\rho}}, \text{ with } F \in C^\infty(M/A'_G)\}.
\end{aligned}
$$

The definitions and the fact that $\mathcal{J}_{\mathcal{O}_\rho}$ is a submersion imply that

$$W^\infty\left(\mathcal{J}^{-1}(\mathcal{O}_\rho)/A'_G\right) \subset C^\infty\left(\mathcal{J}^{-1}(\mathcal{O}_\rho)/A'_G\right).$$

Indeed, let $f \in W^\infty\left(\mathcal{J}^{-1}(\mathcal{O}_\rho)/A'_G\right)$ arbitrary. By definition, there exist $F \in C^\infty(M/A'_G)$ such that $f = F|_{M'_{\mathcal{O}_\rho}}$. As $F \in C^\infty(M/A'_G)$ we have $F \circ \mathcal{J} \in C^\infty(M)$. Also, as $\mathcal{J}^{-1}(\mathcal{O}_\rho)$ is an immersed initial submanifold of M, the injection $i_{\mathcal{O}_\rho} : \mathcal{J}^{-1}(\mathcal{O}_\rho) \hookrightarrow M$ is smooth, and therefore so is $F \circ \mathcal{J} \circ i_{\mathcal{O}_\rho} = F \circ \mathcal{J}_{\mathcal{O}_\rho}$. Consequently, $f \circ \mathcal{J}_{\mathcal{O}_\rho} = F \circ \mathcal{J}_{\mathcal{O}_\rho}$ is smooth. As $\mathcal{J}_{\mathcal{O}_\rho}$ is a submersion f is necessarily smooth, that is, $f \in C^\infty\left(\mathcal{J}^{-1}(\mathcal{O}_\rho)/A'_G\right)$, as required.

14.3.2 Definition.   *We say that $M'_{\mathcal{O}_\rho}$ is **Whitney spanned** when the differentials of its Whitney smooth functions span its cotangent bundle, that is,*

$$\operatorname{span}\{\mathbf{d}f(\sigma) \mid f \in W^\infty(M'_{\mathcal{O}_\rho})\} = T^*_\sigma M'_{\mathcal{O}_\rho}, \quad \text{for all} \quad \sigma \in M'_{\mathcal{O}_\rho}.$$

A sufficient (but not necessary!) condition for $M'_{\mathcal{O}_\rho}$ to be Whitney spanned is that $W^\infty(M'_{\mathcal{O}_\rho}) = C^\infty(M'_{\mathcal{O}_\rho})$.

We are now in the position to state the main results of this section.

14.3.3 Theorem (Polar reduction of a Poisson manifold).   *Let $(M, \{\cdot,\cdot\})$ be a smooth Poisson manifold and G be a Lie group acting canonically and properly on M. Let $\mathcal{J} : M \to M/A'_G$ be the optimal momentum map associated to this action and $\rho \in M/A'_G$ be such that G_ρ is closed in G. If A_G is weakly von Neumann then, for each point $z \in \mathcal{J}^{-1}(\mathcal{O}_\rho)$ and vectors $v, w \in T_z\mathcal{J}^{-1}(\mathcal{O}_\rho)$, there exists an open A'_G–invariant neighborhood U of z and two smooth functions $f, g \in C^\infty(U)$ such that $v = X_f(z)$ and $w =$*

$X_g(z)$. *Moreover, there is a unique presymplectic form $\omega'_{\mathcal{O}_\rho}$ on the polar reduced space $M'_{\mathcal{O}_\rho}$ that satisfies*

$$\{f, g\}|_U(z) = \pi^*_{\mathcal{O}_\rho}\omega_{\mathcal{O}_\rho}(z)(v, w) + \mathcal{J}^*_{\mathcal{O}_\rho}\omega'_{\mathcal{O}_\rho}(z)(v, w) \tag{14.3.3}$$

If $M'_{\mathcal{O}_\rho}$ is Whitney spanned then the form $\omega'_{\mathcal{O}_\rho}$ is symplectic.

Remark. It can be proved that when A_G is von Neumann and A'_G satisfies the extension property the symplecticity of $\omega'_{\mathcal{O}_\rho}$ is equivalent to $M'_{\mathcal{O}_\rho}$ being Whitney spanned.

When the Poisson manifold $(M, \{\cdot, \cdot\})$ is actually a symplectic manifold with symplectic form ω the von Neumann condition in the previous result is no longer needed. Moreover, the conditions under which the form $\omega'_{\mathcal{O}_\rho}$ is symplectic can be completely characterized and the regularized polar subspaces appear as symplectic submanifolds of the polar space that contains them.

14.3.4 Theorem (Polar reduction of a symplectic manifold). *Let (M, ω) be a smooth symplectic manifold and G be a Lie group acting canonically and properly on M. Let $\mathcal{J} : M \to M/A'_G$ be the optimal momentum map associated to this action and $\rho \in M/A'_G$ be such that G_ρ is closed in G.*

(i) *There is a unique presymplectic form $\omega'_{\mathcal{O}_\rho}$ on the polar reduced space $M'_{\mathcal{O}_\rho} \simeq G/G_\rho$ that satisfies*

$$i^*_{\mathcal{O}_\rho}\omega = \pi^*_{\mathcal{O}_\rho}\omega_{\mathcal{O}_\rho} + \mathcal{J}^*_{\mathcal{O}_\rho}\omega'_{\mathcal{O}_\rho}. \tag{14.3.4}$$

The form $\omega'_{\mathcal{O}_\rho}$ is symplectic if and only if for one point $z \in \mathcal{J}^{-1}(\mathcal{O}_\rho)$ (and hence for all) we have

$$\mathfrak{g} \cdot z \cap (\mathfrak{g} \cdot z)^\omega \subset T_z M_{G_z} \tag{14.3.5}$$

(ii) *Let $M'_{\mathcal{N}_\rho} = \mathcal{J}^{-1}(\mathcal{N}_\rho)/A'_G \simeq N(H)^\rho/G_\rho$ be a regularized polar reduced subspace of $M'_{\mathcal{O}_\rho}$. Let*

$$j_{\mathcal{N}_\rho} : \mathcal{J}^{-1}(\mathcal{N}_\rho)/A'_G \hookrightarrow \mathcal{J}^{-1}(\mathcal{O}_\rho)/A'_G$$

be the inclusion and $\omega'_{\mathcal{O}_\rho}$ the presymplectic form defined in (i). Then, the form

$$\omega'_{\mathcal{N}_\rho} := j^*_{\mathcal{N}_\rho}\omega'_{\mathcal{O}_\rho} \tag{14.3.6}$$

is symplectic, that is, the regularized polar subspaces are symplectic submanifolds of the polar space that contains them.

Remark. The characterization (14.3.5) of the symplecticity of $\omega'_{\mathcal{O}_\rho}$ admits a particularly convenient reformulation when the G–action on the symplectic manifold (M, ω) admits an equivariant momentum map $\mathbf{J} : M \to \mathfrak{g}^*$.

Indeed, let $z \in M$ be such that $\mathbf{J}(z) = \mu \in \mathfrak{g}^*$ and $G_z = H$. Then, if the symbol G_μ denotes the coadjoint isotropy of μ, (14.3.5) is equivalent to

$$\mathfrak{g} \cdot z \cap (\mathfrak{g} \cdot z)^\omega = \mathfrak{g}_\mu \cdot z \subset T_z M_H,$$

which in turn amounts to $\mathfrak{g}_\mu \cdot z \subset \mathfrak{g}_\mu \cdot z \cap T_z M_H = \mathrm{Lie}(N(H) \cap G_\mu) \cdot z$. Let $N_{G_\mu}(H) := N(H) \cap G_\mu$. With this notation, the condition can be rewritten as $\mathfrak{g}_\mu + \mathfrak{h} \subset \mathrm{Lie}(N_{G_\mu}(H)) + \mathfrak{h} \subset \mathfrak{g}_\mu$ or, equivalently, as

$$\mathfrak{g}_\mu = \mathrm{Lie}(N_{G_\mu}(H)). \qquad (14.3.7)$$

Proof of Theorem 14.3.3..   Since A_G is weakly von Neumann, we see that for any $z \in M$ $\mathfrak{g} \cdot z \subset A_G''(z)$ or, equivalently, that for any $z \in M$ and any $\xi \in \mathfrak{g}$, there is a A_G'–invariant neighborhood U of z and a function $F \in C^\infty(U/A_G')$ such that $\xi_M(z) = X_{F \circ \mathcal{J}}(z)$. Consequently, for any vector $v \in T_z \mathcal{J}^{-1}(\mathcal{O}_\rho)$ there exists $f \in C^\infty(M)^G$ and $F \in C^\infty(U/A_G')$ (shrink U if necessary) such that

$$v = X_f(z) + X_{F \circ \mathcal{J}}(z) = X_{f|_U + F \circ \mathcal{J}}(z).$$

Let $w \in T_z \mathcal{J}^{-1}(\mathcal{O}_\rho)$, $l \in C^\infty(M)^G$, and $L \in C^\infty(U/A_G')$ be such that

$$w = X_l(z) + X_{L \circ \mathcal{J}}(z) = X_{l|_U + L \circ \mathcal{J}}(z).$$

The expression (14.3.3) can then be rewritten as

$$
\begin{aligned}
\mathcal{J}_{\mathcal{O}_\rho}^* \omega_{\mathcal{O}_\rho}'(z)(v, w) &= \mathcal{J}_{\mathcal{O}_\rho}^* \omega_{\mathcal{O}_\rho}'(z)(X_{f|_U + F \circ \mathcal{J}}(z), X_{l|_U + L \circ \mathcal{J}}(z)) \\
&= \{f + F \circ \mathcal{J}, l + L \circ \mathcal{J}\}|_U(z) \\
&\quad -\pi_{\mathcal{O}_\rho}^* \omega_{\mathcal{O}_\rho}(z)(X_{f|_U + F \circ \mathcal{J}}(z), X_{l|_U + L \circ \mathcal{J}}(z)) \\
&= \{F \circ \mathcal{J}, L \circ \mathcal{J}\}|_U(z)
\end{aligned}
$$

$$(14.3.8)$$

We now show that $\omega_{\mathcal{O}_\rho}'$ is well-defined. Indeed, let $z' \in \mathcal{J}^{-1}(\mathcal{O}_\rho)$ and $v', w' \in T_{z'} \mathcal{J}^{-1}(\mathcal{O}_\rho)$ be such that $T_z \mathcal{J}_{\mathcal{O}_\rho} \cdot v = T_{z'} \mathcal{J}_{\mathcal{O}_\rho} \cdot v'$ and $T_z \mathcal{J}_{\mathcal{O}_\rho} \cdot w = T_{z'} \mathcal{J}_{\mathcal{O}_\rho} \cdot w'$. First of all these equalities imply the existence of an element $\mathcal{F}_T$ in the polar pseudogroup of A_G such that $z' = \mathcal{F}_T(z)$. As $\mathcal{F}_T$ is a local diffeomorphism such that $\mathcal{J}_{\mathcal{O}_\rho} \circ \mathcal{F}_T = \mathcal{J}_{\mathcal{O}_\rho}$, we get $T_z \mathcal{J}_{\mathcal{O}_\rho} = T_{z'} \mathcal{J}_{\mathcal{O}_\rho} \cdot T_z \mathcal{F}_T$. Now, we can rewrite the conditions $T_z \mathcal{J}_{\mathcal{O}_\rho} \cdot v = T_{z'} \mathcal{J}_{\mathcal{O}_\rho} \cdot v'$ and $T_z \mathcal{J}_{\mathcal{O}_\rho} \cdot w = T_{z'} \mathcal{J}_{\mathcal{O}_\rho} \cdot w'$ as $T_{z'} \mathcal{J}_{\mathcal{O}_\rho} \cdot T_z \mathcal{F}_T \cdot v = T_{z'} \mathcal{J}_{\mathcal{O}_\rho} \cdot v'$ and $T_{z'} \mathcal{J}_{\mathcal{O}_\rho} \cdot T_z \mathcal{F}_T \cdot w = T_{z'} \mathcal{J}_{\mathcal{O}_\rho} \cdot w'$, respectively, which implies the existence of two functions $f', l' \in C^\infty(M)^G$ such that

$$
\begin{aligned}
v' &= T_z \mathcal{F}_T(X_f(z) + X_{F \circ \mathcal{J}}(z)) + X_{f'}(\mathcal{F}_T(z)) \\
w' &= T_z \mathcal{F}_T(X_l(z) + X_{L \circ \mathcal{J}}(z)) + X_{l'}(\mathcal{F}_T(z))
\end{aligned}
$$

or, equivalently:

$$
\begin{aligned}
v' &= X_{f \circ \mathcal{F}_{-T}}(\mathcal{F}_T(z)) + X_{F \circ \mathcal{J}}(\mathcal{F}_T(z)) + X_{f'}(\mathcal{F}_T(z)) \\
w' &= X_{l \circ \mathcal{F}_{-T}}(\mathcal{F}_T(z)) + X_{L \circ \mathcal{J}}(\mathcal{F}_T(z)) + X_{l'}(\mathcal{F}_T(z)).
\end{aligned}
$$

Therefore, using (14.3.3), we arrive at

$$
\begin{aligned}
\mathcal{J}^*_{\mathcal{O}_\rho}&\omega'_{\mathcal{O}_\rho}(z')(v',w')\\
&= \{f \circ \mathcal{F}_{-T} + F \circ \mathcal{J} + f', l \circ \mathcal{F}_{-T} + L \circ \mathcal{J} + l'\}|_V(\mathcal{F}_T(z))\\
&\quad - \pi^*_{\mathcal{O}_\rho}\omega_{\mathcal{O}_\rho}(\mathcal{F}_T(z))(X_{f \circ \mathcal{F}_{-T}|_V + F \circ \mathcal{J} + f'}(z), X_{l \circ \mathcal{F}_{-T}|_V + L \circ \mathcal{J} + l'}(z))\\
&= \{F \circ \mathcal{J}, L \circ \mathcal{J}\}|_V(\mathcal{F}_T(z)) = \{F \circ \mathcal{J}, L \circ \mathcal{J}\}|_U(z)\\
&= \mathcal{J}^*_{\mathcal{O}_\rho}\omega'_{\mathcal{O}_\rho}(z)(v,w),
\end{aligned}
$$

where $V = U \cap \mathcal{F}_T(\mathrm{Dom}(\mathcal{F}_T)) = \mathcal{F}_T(U \cap \mathrm{Dom}(\mathcal{F}_T))$. Hence, the form $\omega'_{\mathcal{O}_\rho}$ is well-defined. The closedness and skew symmetric character of $\omega'_{\mathcal{O}_\rho}$ is obtained as a consequence of $\mathcal{J}_{\mathcal{O}_\rho}$ being a surjective submersion, $\omega_{\mathcal{O}_\rho}$ being closed and skew symmetric, and the $\{\cdot,\cdot\}$ being a Poisson bracket. An equivalent fashion to realize this is by writing $\omega'_{\mathcal{O}_\rho}$ in terms of the symplectic structure of the leaves of M. Indeed, as A_G is weakly von Neumann, each connected component of $\mathcal{J}^{-1}(\mathcal{O}_\rho)$ lies in a single symplectic leaf of $(M, \{\cdot,\cdot\})$. In order to simplify the exposition suppose that $\mathcal{J}^{-1}(\mathcal{O}_\rho)$ is connected and let $\mathcal{L}_{\mathcal{O}_\rho}$ be the unique symplectic leaf of M that contains it (otherwise one has just to proceed connected component by connected component). Let $i_{\mathcal{L}_{\mathcal{O}_\rho}} : \mathcal{J}^{-1}(\mathcal{O}_\rho) \to \mathcal{L}_{\mathcal{O}_\rho}$ be the natural injection. Given that $i_{\mathcal{O}_\rho} : \mathcal{J}^{-1}(\mathcal{O}_\rho) \to M$ is smooth and $\mathcal{L}_{\mathcal{O}_\rho}$ is an initial submanifold of M, the map $i_{\mathcal{L}_{\mathcal{O}_\rho}}$ is therefore smooth. If we denote by $\omega_{\mathcal{L}_{\mathcal{O}_\rho}}$ the symplectic form of the leaf $\mathcal{L}_{\mathcal{O}_\rho}$, expression (14.3.4) can be rewritten as:

$$
i^*_{\mathcal{L}_{\mathcal{O}_\rho}}\omega_{\mathcal{L}_{\mathcal{O}_\rho}} = \pi^*_{\mathcal{O}_\rho}\omega_{\mathcal{O}_\rho} + \mathcal{J}^*_{\mathcal{O}_\rho}\omega'_{\mathcal{O}_\rho}. \tag{14.3.9}
$$

The antisymmetry and closedness of $\omega'_{\mathcal{O}_\rho}$ appears then as a consequence of the antisymmetry and closedness of $\omega_{\mathcal{O}_\rho}$ and $\omega_{\mathcal{L}_{\mathcal{O}_\rho}}$.

It just remains to be shown that if $M'_{\mathcal{O}_\rho}$ is Whitney spanned then the form $\omega'_{\mathcal{O}_\rho}$ is non degenerate. Let $z \in \mathcal{J}^{-1}(\mathcal{O}_\rho)$ and $v \in T_z\mathcal{J}^{-1}(\mathcal{O}_\rho)$ be such that

$$
\omega'_{\mathcal{O}_\rho}(\mathcal{J}_{\mathcal{O}_\rho}(z))(T_z\mathcal{J}_{\mathcal{O}_\rho} \cdot v, T_z\mathcal{J}_{\mathcal{O}_\rho} \cdot w) = 0, \quad \text{for all} \quad w \in T_z\mathcal{J}^{-1}(\mathcal{O}_\rho). \tag{14.3.10}
$$

Take now $f \in C^\infty(M)^G$ and $F \in C^\infty(U/A'_G)$ such that $v = X_f(z) + X_{F \circ \mathcal{J}}(z)$. Condition (14.3.10) is equivalent to requiring that

$$
\omega'_{\mathcal{O}_\rho}(\mathcal{J}_{\mathcal{O}_\rho}(z))(T_z\mathcal{J}_{\mathcal{O}_\rho} \cdot X_{F \circ \mathcal{J}}(z), T_z\mathcal{J}_{\mathcal{O}_\rho} \cdot X_{L \circ \mathcal{J}}(z)) = 0, \tag{14.3.11}
$$

for all $L \in C^\infty(V/A'_G)$ and all open A'_G–invariant neighborhoods V of z. By (14.3.8) we can rewrite (14.3.11) as

$$
\{F \circ \mathcal{J}, L \circ \mathcal{J}\}|_{U \cap V}(z) = 0. \tag{14.3.12}
$$

Now, notice that for any $h \in W^{\infty}(M'_{\mathcal{O}_{\rho}})$ there exists a function $H \in C^{\infty}(M/A'_G)$ such that $H|_{M'_{\mathcal{O}_{\rho}}} = h$. Moreover, by (14.3.12) we obtain:

$$\mathbf{d}h(\mathcal{J}_{\mathcal{O}_{\rho}}(z)) \cdot (T_z\mathcal{J}_{\mathcal{O}_{\rho}} \cdot X_{F \circ \mathcal{J}}(z))$$
$$= \mathbf{d}(h \circ \mathcal{J}_{\mathcal{O}_{\rho}})(z) \cdot X_{F \circ \mathcal{J}}(z) = \mathbf{d}(H \circ \mathcal{J})(z) \cdot X_{F \circ \mathcal{J}}(z) = 0.$$

Given that the previous equality holds for any $h \in W^{\infty}(M'_{\mathcal{O}_{\rho}})$ and $M'_{\mathcal{O}_{\rho}}$ is Whitney spanned, we obtain

$$T_z\mathcal{J}_{\mathcal{O}_{\rho}} \cdot X_{F \circ \mathcal{J}}(z) = T_z\mathcal{J}_{\mathcal{O}_{\rho}} \cdot v = 0,$$

as required.

Proof of Theorem 14.3.4.   (i) The well–definedness and presymplectic character of $\omega'_{\mathcal{O}_{\rho}}$ in this case can be obtained as a consequence of Theorem 14.3.3. This is particularly evident when we think of $\omega'_{\mathcal{O}_{\rho}}$ as the form characterized by equality (14.3.9) and we recall that in the symplectic case $\omega_{\mathcal{L}_{\mathcal{O}_{\rho}}} = \omega$.

It just remains to be shown that the form $\omega'_{\mathcal{O}_{\rho}}$ is non degenerate if and only if condition (14.3.5) holds. We proceed by showing first that if condition (14.3.5) holds for the point $z \in \mathcal{J}^{-1}(\mathcal{O}_{\rho})$ then it holds for all the points in $\mathcal{J}^{-1}(\mathcal{O}_{\rho})$. We will then prove that (14.3.5) at the point z is equivalent to the non degeneracy of $\omega'_{\mathcal{O}_{\rho}}$ at $\mathcal{J}_{\mathcal{O}_{\rho}}(z)$.

Suppose first that the point $z \in \mathcal{J}^{-1}(\mathcal{O}_{\rho})$ is such that $\mathfrak{g} \cdot z \cap (\mathfrak{g} \cdot z)^{\omega} \subset T_z M_{G_z}$. Notice now that any element in $\mathcal{J}^{-1}(\mathcal{O}_{\rho})$ can be written as $\Phi_g(\mathcal{F}_T(z))$ with $g \in G$ and $\mathcal{F}_T$ in the polar pseudogroup of A_G. It is easy to show that the relation

$$\mathfrak{g} \cdot (\Phi_g(\mathcal{F}_T(z))) \cap (\mathfrak{g} \cdot (\Phi_g(\mathcal{F}_T(z))))^{\omega} \subset T_{\Phi_g(\mathcal{F}_T(z))} M_{G_{\Phi_g(\mathcal{F}_T(z))}}$$

is equivalent to $T_z(\Phi \circ \mathcal{F}_T)(\mathfrak{g} \cdot z \cap (\mathfrak{g} \cdot z)^{\omega}) \subset T_z(\Phi \circ \mathcal{F}_T)M_{G_z}$ and therefore to $\mathfrak{g} \cdot z \cap (\mathfrak{g} \cdot z)^{\omega} \subset T_z M_{G_z}$.

Let now $v \in T_z\mathcal{J}^{-1}(\mathcal{O}_{\rho})$ be such that

$$\omega'_{\mathcal{O}_{\rho}}(\mathcal{J}_{\mathcal{O}_{\rho}}(z))(T_z\mathcal{J}_{\mathcal{O}_{\rho}} \cdot v, T_z\mathcal{J}_{\mathcal{O}_{\rho}} \cdot w) = 0, \quad \text{for all} \quad w \in T_z\mathcal{J}^{-1}(\mathcal{O}_{\rho}).$$
$$(14.3.13)$$

Take now $f \in C^{\infty}(M)^G$ and $\xi \in \mathfrak{g}$ such that $v = X_f(z) + \xi_M(z)$. Condition (14.3.13) is equivalent to having that

$$\omega'_{\mathcal{O}_{\rho}}(\mathcal{J}_{\mathcal{O}_{\rho}}(z))(T_z\mathcal{J}_{\mathcal{O}_{\rho}} \cdot \xi_M(z), T_z\mathcal{J}_{\mathcal{O}_{\rho}} \cdot \eta_M(z)) = 0, \qquad \text{for all } \eta \in \mathfrak{g}$$

which by (14.3.4) can be rewritten as

$$\omega(z)(\xi_M(z), \eta_M(z)) = 0, \qquad \text{for all } \eta \in \mathfrak{g},$$

and thereby amounts to having that $\xi_M(z) \in \mathfrak{g} \cdot z \cap (\mathfrak{g} \cdot z)^{\omega}$. Hence, $\omega'_{\mathcal{O}_{\rho}}(\mathcal{J}_{\mathcal{O}_{\rho}}(z))$ is non degenerate if and only if $\xi_M(z) \in \ker T_z\mathcal{J}_{\mathcal{O}_{\rho}} = A'_G(z)$.

Suppose now that condition (14.3.5) holds; then, as $\xi_M(z) \in \mathfrak{g} \cdot z \cap (\mathfrak{g} \cdot z)^\omega$ we have $\xi_M(z) \in T_z M_{G_z}$. Using (13.1.1) we can conclude that $\xi_M(z) \in A'_G(z)$, as required. Conversely, suppose that $\omega'_{\mathcal{O}_\rho}$ is symplectic. The previous equalities immediately imply that $\mathfrak{g} \cdot z \cap (\mathfrak{g} \cdot z)^\omega \subset A'_G(z) \subset T_z M_{G_z}$, as required.

(ii) The form $\omega'_{\mathcal{N}_\rho}$ is clearly closed and antisymmetric. We now show that it is non degenerate. Recall firs that the tangent space to $T_z \mathcal{J}^{-1}(\mathcal{N}_\rho)$ at a given point $z \in \mathcal{J}^{-1}(\mathcal{N}_\rho)$ is given by the vectors of the form

$$v = X_f(z) + \xi_M(z),$$

with $f \in C^\infty(M)^G$ and $\xi \in \mathrm{Lie}(N(H)^\rho)$. Let $v = X_f(z) + \xi_M(z) \in T_z \mathcal{J}^{-1}(\mathcal{N}_\rho)$ be such that

$$\mathcal{J}^*_{\mathcal{N}_\rho}(j^*_{\mathcal{N}_\rho} \omega'_{\mathcal{O}_\rho})(z)(X_f(z) + \xi_M(z), X_g(z) + \eta_M(z)) = 0,$$

for all $\eta \in \mathrm{Lie}(N(H)^\rho)$ and $g \in C^\infty(M)^G$.

If we plug into the previous expression the definition of the form $\omega'_{\mathcal{O}_\rho}$ we obtain

$$\omega(z)(\xi_M(z), \eta_M(z)) = 0,$$

for all $\eta \in \mathrm{Lie}(N(H)^\rho)$, that is,

$$
\begin{aligned}
\xi_M(z) \in (\mathrm{Lie}(N(H)^\rho) \cdot z) &\cap (\mathrm{Lie}(N(H)^\rho) \cdot z)^\omega \\
&= (\mathrm{Lie}(N(H)^\rho) \cdot z) \cap (\mathrm{Lie}(N(H)^\rho) \cdot z)^{\omega|_{M_H^\rho}} \\
&= (\mathrm{Lie}(N(H)^\rho/H) \cdot z) \cap A'_{N(H)^\rho/H}(z),
\end{aligned}
$$

where the last equality follows from (13.1.1) and the freeness of the natural $N(H)^\rho/H$–action on M_H^ρ. We now recall (see Lemma 4.4 in Ortega and Ratiu [2002]) that any $N(H)^\rho/H$–invariant function on M_H^ρ admits a local extension to a G–invariant function on M, hence $\xi_M(z) \in (\mathrm{Lie}(N(H)^\rho/H) \cdot z) \cap A'_G(z)$, and consequently $T_z \mathcal{J}_{\mathcal{O}_\rho} \cdot \xi_M(z) = T_z \mathcal{J}_{\mathcal{O}_\rho} \cdot v = T_z \mathcal{J}_{\mathcal{N}_\rho} \cdot v = 0$, as required. ∎

14.4   Symplectic Leaves and the Reduction Diagram

Suppose that A'_G is completable so that the symplectic leaves of M/A'_G are well-defined. We recall that this is automatically the case when (M, ω) is symplectic and the G–group action is proper (see Ortega [2003a]). Assume also that A_G is von Neumann so that the diagram $(M/G, \{\cdot, \cdot\}_{M/A_G}) \overset{\pi_{A_G}}{\longleftarrow} (M, \{\cdot, \cdot\}) \overset{\mathcal{J}}{\longrightarrow} (M/A'_G, \{\cdot, \cdot\}_{M/A'_G})$ constitutes a dual pair.

Notice that by Definition 13.4.1, the symplectic leaves of M/A_G and M/A'_G coincide with the connected components of the orbit reduced spaces $M_{\mathcal{O}_\rho}$ and polar reduced spaces $M'_{\mathcal{O}_\rho}$, that we studied in sections 14.2 and 14.3, respectively. We saw that whenever G_ρ is closed in G and the Whitney spanning condition is satisfied these spaces are actual symplectic manifolds. When M is symplectic, the symplecticity of the leaves of M/A'_G is characterized by condition (14.3.5) or even by (14.3.7), provided that the G–action has an associated standard equivariant momentum map $\mathbf{J} : M \to \mathfrak{g}^*$. Moreover, when $M_{\mathcal{O}_\rho}$ and $M'_{\mathcal{O}_\rho}$ are corresponding leaves, their symplectic structures are connected to each other by an identity that naturally generalizes the classical relation that we recalled in (13.5.2).

The following diagram represents all the spaces that we worked with and their relations. The part of the diagram dealing with the regularized spaces refers only to the situation in which M is symplectic.
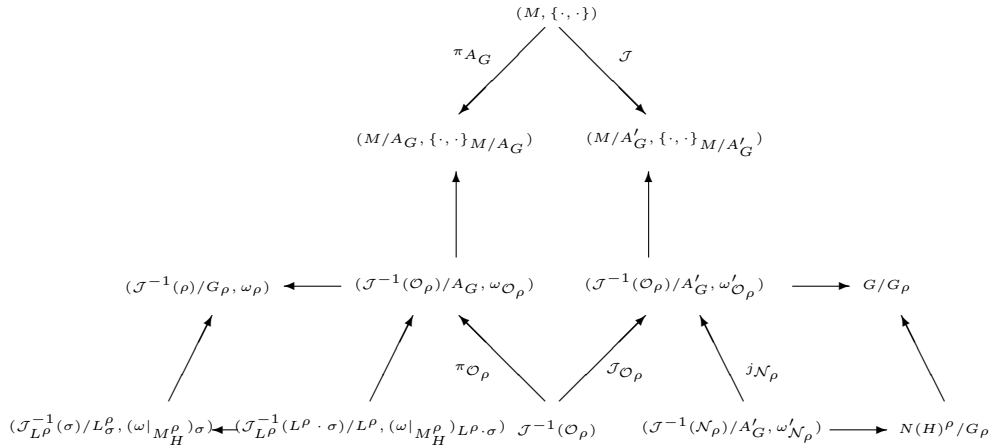
14.5 Orbit Reduction: Beyond Compact Groups

The approach to optimal orbit reduction developed in the last few sections sheds some light on how to carry out orbit reduction with a standard momentum map when the symmetry group is not compact. This absence of compactness poses some technical problems that have been tackled by various people over the years using different approaches. Since these problems already arise in the free actions case we will restrict ourselves to this situation. More specifically we will assume that we have a Lie group G (not necessarily compact) acting freely and canonically on the symplectic manifold (M, ω). We will suppose that this action has an associated coadjoint equivariant momentum map $\mathbf{J} : M \to \mathfrak{g}^*$. For the sake of simplicity in the

exposition and in order to have a better identification with the material presented in the previous sections we will assume that $\mathbf{J}$ has connected fibers. This assumption is not fundamental. The reader interested in the general case with no connectedness hypothesis in the fibers and nonfree actions may want to check with [HRed].

In the presence of the hypotheses that we just stated, the momentum map $\mathbf{J}$ is a submersion that maps M onto an open coadjoint equivariant subset $\mathfrak{g}_{\mathbf{J}}^*$ of $\mathfrak{g}^*$. Moreover, any value $\mu \in \mathfrak{g}_{\mathbf{J}}^*$ of $\mathbf{J}$ is regular and has an associated smooth Marsden–Weinstein symplectic reduced space $\mathbf{J}^{-1}(\mu)/G_\mu$. What about the orbit reduced space $\mathbf{J}^{-1}(\mathcal{O}_\mu)/G$? When the Lie group G is compact there is no problem to canonically endow $\mathbf{J}^{-1}(\mathcal{O}_\mu)/G$ with a smooth structure. Indeed, in this case the coadjoint orbit $\mathcal{O}_\mu$ is an embedded submanifold of $\mathfrak{g}^*$ transverse to the momentum mapping. The Transversal Mapping Theorem ensures that $\mathbf{J}^{-1}(\mathcal{O}_\mu)$ is a G–invariant embedded submanifold of M and hence the quotient $\mathbf{J}^{-1}(\mathcal{O}_\mu)/G$ is smooth and symplectic with the form spelled out in (13.5.1). In the non compact case this argument breaks down due to the non embedded character of $\mathcal{O}_\mu$ in $\mathfrak{g}^*$. In trying to fix this problem this has lead to the assumption of locally closedness on the coadjoint orbits that one can see in a number of papers (see for instance Bates and Lerman [1997]). Nevertheless, this hypothesis is not needed to carry out point reduction, and therefore makes the two approaches non equivalent. The first work where this hypothesis has been eliminated is Cushman and Śniatycki [2001]. In this paper the authors use a combination of distribution theory with Sikorski differential spaces to show that the orbit reduced space is a symplectic manifold. Nevertheless, the first reference where the standard formula (13.5.1) appears at this level of generality is Blaom [2001]. In that paper the author only deals with the free case. Nevertheless the use of a standard technique of reduction to the isotropy type manifolds that the reader can find in Sjamaar and Lerman [1991]; Ortega [1998]; Cushman and Śniatycki [2001]; Ortega and Ratiu [2004a] generalizes the results of Blaom [2001] to singular situations.

In the next few paragraphs we will illustrate Theorem 14.3.3 by showing that the results in Cushman and Śniatycki [2001]; Blaom [2001] can be obtained as a corollary of it.

We start by identifying in this setup all the elements in that result. First of all, we note that the polar distribution satisfies $A_G' = \ker T\mathbf{J}$ (see Ortega and Ratiu [2002]) and the connectedness hypothesis on the fibers of $\mathbf{J}$ implies that the optimal momentum map $\mathcal{J} : M \to M/A_G'$ in this case can be identified with $\mathbf{J} : M \to \mathfrak{g}_{\mathbf{J}}^*$. This immediately implies that for any $\mu \in \mathfrak{g}_{\mathbf{J}}^* \simeq M/A_G'$, the isotropy G_μ is closed in G and, by Theorem 14.1.4 there is a unique smooth structure on $\mathbf{J}^{-1}(\mathcal{O}_\mu)$ that makes it into an initial submanifold of M and, at the same time, an integral manifold of the distribution $D = A_G' + \mathfrak{g} \cdot m = \ker T\mathbf{J} + \mathfrak{g} \cdot m$. This structure coincides with the one given in Blaom [2001]. Also, by Theorem 14.2.1, the quotient $\mathbf{J}^{-1}(\mathcal{O}_\mu)/G$ admits a unique symplectic structure $\omega_{\mathcal{O}_\mu}$ that makes it symplectomor-

phic to the Marsden–Weinstein point reduced space $(\mathbf{J}^{-1}(\mu)/G_\mu, \omega_\mu)$. It remains to be shown that we can use (14.3.4) in this case and that the resulting formula coincides with the standard one (13.5.1) provided by Blaom [2001]. An analysis of the polar reduced space in this setup will provide an affirmative answer to this question.

By Proposition 14.3.1 the polar reduced space $\mathbf{J}(\mathcal{O}_\mu)/A_G'$ is endowed with the only smooth structure that makes it diffeomorphic to the homogeneous space $G/G_\mu \simeq \mathcal{O}_\mu$. Hence, in this case $\mathbf{J}_{\mathcal{O}_\mu} : \mathbf{J}^{-1}(\mathcal{O}_\mu) \to \mathcal{O}_\mu$ is the map given by $\mathbf{J}_{\mathcal{O}_\mu}(z) := \mathbf{J}(z)$ which is smooth because the coadjoint orbits are always initial submanifolds of $\mathfrak{g}^*$. Therefore we can already compute the polar symplectic form $\omega_{\mathcal{O}_\mu}'$. By (14.3.4) we see for any $\xi, \eta \in \mathfrak{g}$ and any $z \in \mathbf{J}^{-1}(\mathcal{O}_\mu)$ (for simplicity in the exposition we take $\mathbf{J}(z) = \mu$):

$$
\begin{aligned}
\mathbf{J}_{\mathcal{O}_\mu}^* \omega_{\mathcal{O}_\mu}'(z)(\xi_M(z), \eta_M(z)) \\
= i_{\mathcal{O}_\mu}^* \omega(z)(\xi_M(z), \eta_M(z)) - \pi_{\mathcal{O}_\mu}^* \omega_{\mathcal{O}_\mu}(z)(\xi_M(z), \eta_M(z)),
\end{aligned}
$$

or, equivalently:

$$
\omega_{\mathcal{O}_\mu}'(\mu)(\mathrm{ad}_\xi^* \mu, \mathrm{ad}_\eta^* \mu) = \omega(z)(\xi_M(z), \eta_M(z)) = \langle \mathbf{J}(z), [\xi, \eta] \rangle = \langle \mu, [\xi, \eta] \rangle.
$$

In conclusion, in this case the polar reduced form $\omega_{\mathcal{O}_\mu}'$ coincides with the "+"–Kostant–Kirillov–Souriau symplectic form on the coadjoint orbit $\mathcal{O}_\mu$. Therefore, the general optimal orbit reduction formula (14.3.4) coincides with the standard one (13.5.1).

14.6   Examples: Polar Reduction of the Coadjoint Action

We now provide two examples on how we can use the coadjoint action along with Theorems 14.3.3 and 14.3.4 to easily produce symplectic manifolds and symplectically decomposed presymplectic manifolds.

Coadjoint Orbits as Polar Reduced Spaces.   Let G be a Lie group, $\mathfrak{g}$ be its Lie algebra, and $\mathfrak{g}^*$ be its dual considered as a Lie–Poisson space. In this elementary example we show how the coadjoint orbits appear as the polar reduced spaces of the coadjoint G–action on $\mathfrak{g}^*$.

A straightforward computation shows that the coadjoint action of G on the Lie–Poisson space $\mathfrak{g}^*$ is canonical. Moreover, the polar distribution $A_G'(\mu) = 0$ for all $\mu \in \mathfrak{g}^*$ and therefore the optimal momentum map $\mathcal{J} : \mathfrak{g}^* \to \mathfrak{g}^*$ is the identity map on $\mathfrak{g}^*$. This immediately implies that any open set $U \subset \mathfrak{g}^*$ is A_G'–invariant, that $C^\infty(U)^{A_G} = C^\infty(U)$, and that therefore $\mathfrak{g} \cdot \mu \subset A_G''(\mu)$, for any $\mu \in \mathfrak{g}^*$. The coadjoint action on $\mathfrak{g}^*$ is therefore weakly von Neumann (actually, if G is connected A_G is von Neumann).

We now look at the corresponding reduced spaces. On one hand the orbit reduced spaces $\mathcal{J}^{-1}(\mathcal{O}_\rho)/G$ are the quotients $G \cdot \mu/G$ and therefore amount to points. At the same time, we have $\mathcal{J}^{-1}(\mathcal{O}_\rho)/A'_G = \mathcal{O}_\mu/A'_G = \mathcal{O}_\mu$, that is, the polar reduced spaces are the coadjoint orbits which, by Theorem 14.3.3, are symplectic. Indeed, the Whitney spanning condition necessary for the application of this result is satisfied since in this case

$$\mathrm{span}\{\mathbf{d}f(\mu) \mid f \in W^\infty(M'_{\mathcal{O}_\rho})\} = \mathrm{span}\{\mathbf{d}h|_{\mathcal{O}_\mu}(\mu) \mid h \in C^\infty(\mathfrak{g}^*)\} = T^*_\mu \mathcal{O}_\mu.$$

Note that the last equality is a consequence of the immersed character of the coadjoint orbits $\mathcal{O}_\mu$ as submanifolds of $\mathfrak{g}^*$ (the equality is easily proved using immersion charts around the point μ).

Symplectic Decomposition of Presymplectic Homogeneous Manifolds. Let G be a Lie group, $\mathfrak{g}$ be its Lie algebra, and $\mathfrak{g}^*$ be its dual. Let $\mathcal{O}_{\mu_1}$ and $\mathcal{O}_{\mu_2}$ be two coadjoint orbits of $\mathfrak{g}^*$ that we will consider as symplectic manifolds endowed with the KKS–symplectic forms $\omega_{\mathcal{O}_{\mu_1}}$ and $\omega_{\mathcal{O}_{\mu_2}}$, respectively. The cartesian product $\mathcal{O}_{\mu_1} \times \mathcal{O}_{\mu_2}$ is also a symplectic manifold with the sum symplectic form $\omega_{\mathcal{O}_{\mu_1}} + \omega_{\mathcal{O}_{\mu_2}}$. The diagonal action of G on $\mathcal{O}_{\mu_1} \times \mathcal{O}_{\mu_2}$ is canonical with respect to this symplectic structure and, moreover, it has an associated standard equivariant momentum map $\mathbf{J} : \mathcal{O}_{\mu_1} \times \mathcal{O}_{\mu_2} \to \mathfrak{g}^*$ given by $\mathbf{J}(\nu, \eta) = \nu + \eta$. We now suppose that this action is proper and we will study, in this particular case, the orbit and polar reduced spaces introduced in the previous sections.

We start by looking at the level sets of the optimal momentum map

$$\mathcal{J} : \mathcal{O}_{\mu_1} \times \mathcal{O}_{\mu_2} \to \mathcal{O}_{\mu_1} \times \mathcal{O}_{\mu_2}/A'_G.$$

A general result (see Theorem 3.6 in Ortega and Ratiu [2002]) states that in the presence of a standard momentum map the fibers of the optimal momentum map coincide with the connected components of the intersections of the level sets of the momentum map with the isotropy type submanifolds. Hence, in our case, if $\rho = \mathcal{J}(\mu_1, \mu_2)$, we have

$$\mathcal{J}^{-1}(\rho) = (\mathbf{J}^{-1}(\mu_1 + \mu_2) \cap (\mathcal{O}_{\mu_1} \times \mathcal{O}_{\mu_2})_{G_{(\mu_1,\mu_2)}})_c, \qquad (14.6.1)$$

where the subscript c in the previous expression stands for the connected component of $\mathbf{J}^{-1}(\mu_1 + \mu_2) \cap (\mathcal{O}_{\mu_1} \times \mathcal{O}_{\mu_2})_{G_{(\mu_1,\mu_2)}}$ that contains $\mathcal{J}^{-1}(\rho)$. Given that the isotropy $G_{(\mu_1,\mu_2)} = G_{\mu_1} \cap G_{\mu_2}$, with G_{μ_1} and G_{μ_2} the coadjoint isotropies of μ_1 and μ_2, respectively, the expression (14.6.1) can be rewritten as

$$\mathcal{J}^{-1}(\rho) = (\{(\mathrm{Ad}^*_{g^{-1}} \mu_1, \mathrm{Ad}^*_{h^{-1}} \mu_2) \mid g, h \in G, \text{ such that }$$
$$\mathrm{Ad}^*_{g^{-1}} \mu_1 + \mathrm{Ad}^*_{h^{-1}} \mu_2 = \mu_1 + \mu_2, \ gG_{\mu_1}g^{-1} \cap hG_{\mu_2}h^{-1} = G_{\mu_1} \cap G_{\mu_2}\})_c.$$

It is easy to show that in this case

$$G_\rho = N_{G_{\mu_1+\mu_2}}(G_{\mu_1} \cap G_{\mu_2})^c, \qquad (14.6.2)$$

where the superscript c denotes the closed subgroup of

$$N_{G_{\mu_1+\mu_2}}(G_{\mu_1} \cap G_{\mu_2}) := N(G_{\mu_1} \cap G_{\mu_2}) \cap G_{\mu_1+\mu_2}$$

that leaves $\mathcal{J}^{-1}(\rho)$ invariant. Theorems 13.5.1 and 14.2.1 guarantee that the quotients $\mathcal{J}^{-1}(\rho)/G_\rho \simeq \mathcal{J}^{-1}(\mathcal{O}_\rho)/G$ are symplectic. Nevertheless, we will focus our attention in the corresponding polar reduced spaces.

According to Theorem 14.3.4 and to (14.6.2), the polar reduced space corresponding to $\mathcal{J}^{-1}(\mathcal{O}_\rho)/G$ is the homogeneous presymplectic manifold

$$G/N_{G_{\mu_1+\mu_2}}(G_{\mu_1} \cap G_{\mu_2})^c. \tag{14.6.3}$$

Expression (14.3.7) states that $G/N_{G_{\mu_1+\mu_2}}(G_{\mu_1} \cap G_{\mu_2})^c$ is symplectic if and only if

$$\mathfrak{g}_{\mu_1+\mu_2} = \operatorname{Lie}(N_{G_{\mu_1+\mu_2}}(G_{\mu_1} \cap G_{\mu_2})),$$

which is obviously true when, for instance, $G_{\mu_1} \cap G_{\mu_2}$ is a normal subgroup of $G_{\mu_1+\mu_2}$. In any case, using (14.3.2) we can write the polar reduced space (14.6.3) as a disjoint union of its regularized symplectic reduced subspaces that, that in this case are of the form

$$gN(G_{\mu_1} \cap G_{\mu_2})^\rho/N_{G_{\mu_1+\mu_2}}(G_{\mu_1} \cap G_{\mu_2})^c$$

with $g \in G$ and where the superscript ρ denotes the closed subgroup of $N(G_{\mu_1} \cap G_{\mu_2})$ that leaves invariant the connected component of $(\mathcal{O}_{\mu_1} \times \mathcal{O}_{\mu_2})_{G_{\mu_1} \cap G_{\mu_2}}$ that contains $\mathcal{J}^{-1}(\rho)$. More explicitly, we can write the following symplectic decomposition of the polar reduced space:

$$G/N_{G_{\mu_1+\mu_2}}(G_{\mu_1} \cap G_{\mu_2})^c$$
$$= \dot{\bigcup}_{[g] \in G/N(G_{\mu_1} \cap G_{\mu_2})^c} gN(G_{\mu_1} \cap G_{\mu_2})^c/N_{G_{\mu_1+\mu_2}}(G_{\mu_1} \cap G_{\mu_2})^c.$$

What we just did in the previous paragraphs for two coadjoint orbits can be inductively generalized to n orbits. We collect the results of that construction under the form of a proposition.

14.6.1 Proposition.   *Let G be a Lie group, $\mathfrak{g}$ be its Lie algebra, and $\mathfrak{g}^*$ be its dual. Let $\mu_1, \ldots, \mu_n \in \mathfrak{g}^*$. Then, the homogeneous manifold*

$$G/N_{G_{\mu_1+\cdots+\mu_n}}(G_{\mu_1} \cap \ldots \cap G_{\mu_n})^c \tag{14.6.4}$$

has a natural presymplectic structure that is nondegenerate if and only if

$$\mathfrak{g}_{\mu_1+\cdots+\mu_n} = \operatorname{Lie}(N_{G_{\mu_1+\cdots+\mu_n}}(G_{\mu_1} \cap \ldots \cap G_{\mu_n})).$$

Moreover, (14.6.4) can be written as a the following disjoint union of symplectic submanifolds

$$G/N_{G_{\mu_1+\cdots+\mu_n}}(G_{\mu_1} \cap \ldots \cap G_{\mu_n})^c$$
$$= \dot{\bigcup}_{[g] \in G/N(G_{\mu_1} \cap \ldots \cap G_{\mu_n})^\rho} gN(G_{\mu_1} \cap \ldots \cap G_{\mu_n})^\rho/N_{G_{\mu_1+\cdots+\mu_n}}(G_{\mu_1} \cap \ldots \cap G_{\mu_n})^c.$$

15
Optimal Reduction by Stages

As we already saw in Part II, the reduction by stages procedure consists of carrying out reduction in two shots using the normal subgroups of the symmetry group. To be more specific, suppose that we are in the same setup as Theorem 13.5.1 and that the symmetry group G has a closed normal subgroup N. In this chapter we will spell out the conditions under which optimal reduction by G renders the same result as reduction in the following two stages: we first reduce by N; the resulting space inherits symmetry properties coming from the quotient Lie group G/N that can be used to reduce one more time.

The results on reduction by stages that we will obtain in the optimal context will give us, as a byproduct, a generalization to the singular case (nonfree actions) of the reduction by stages theorem in the presence of an standard equivariant momentum map for a canonical action on a symplectic manifold.

15.1 The Polar Distribution of a Normal Subgroup

All along this section we will work on a Poisson manifold $(M, \{\cdot, \cdot\})$ acted properly and canonically upon by a Lie group G. We will assume that G has a closed normal subgroup that we will denote by N. The closedness of N implies that the N–action on M by restriction is still proper and that G/N is a Lie group when considered as a homogenous manifold. We

will denote by A'_G and A'_N the polar distributions associated to the G and N–actions, respectively, and by $\mathcal{J}_G : M \to M/A'_G$ and $\mathcal{J}_N : M \to M/A'_N$ the corresponding optimal momentum maps.

The following proposition provides a characterization of the conditions under which the polar distribution A'_H associated to a closed subgroup H of G is invariant under the lifted action of G to the tangent bundle TM.

15.1.1 Proposition. *Let $(M, \{\cdot, \cdot\})$ be a Poisson manifold acted properly and canonically upon by a Lie group G via the map $\Phi : G \times M \to M$. Let H be a closed Lie subgroup of G. Then:*

(i) *The lifted action of G to the tangent bundle TM leaves the H–polar distribution A'_H invariant if and only if $f \circ \Phi_{g^{-1}} \in C^\infty(M)^H$, for any $f \in C^\infty(M)^H$ and any $g \in G$. This condition holds if and only if for all $g \in G$, $h \in H$, and $m \in M$, there exists an element $h' \in H$ such that*

$$gh \cdot m = h'g \cdot m. \qquad (15.1.1)$$

(ii) *If G acts on A'_H so it does on the corresponding momentum space M/A'_H with a natural action that makes the H–optimal momentum map $\mathcal{J}_H : M \to M/A'_H$ G–equivariant.*

(iii) *For any $m \in M$ we have $A'_G(m) \subset A'_H(m)$. There is consequently a natural projection $\pi_H : M/A'_G \to M/A'_H$ such that*

$$\mathcal{J}_H = \pi_H \circ \mathcal{J}_G. \qquad (15.1.2)$$

Moreover, if G acts on A'_H and consequently on M/A'_H, the map π_H is G–equivariant.

Proof. (i). Since H is closed in G, its action on M by restriction of the G–action is still proper. Therefore, $A'_H = \{X_f \mid f \in C^\infty(M)^H\}$. Given that for any $f \in C^\infty(M)^H$ and any $g \in G$ we have $T\Phi_g \circ X_f = X_{f \circ \Phi_{g^{-1}}} \circ \Phi_g$, we can conclude that the polar distribution A'_H is G–invariant if and only if $f \circ \Phi_{g^{-1}} \in C^\infty(M)^H$, for any $f \in C^\infty(M)^H$ and any $g \in G$. We now check that this condition is equivalent to (15.1.1).

First of all suppose that $f \circ \Phi_g \in C^\infty(M)^H$, for all $f \in C^\infty(M)^H$ and $g \in G$. Consequently, if we take $m \in M$ and $h \in H$ arbitrary, we have $f(gh \cdot m) = f(g \cdot m)$. Since the H–action on M is proper, the set $C^\infty(M)^H$ of H–invariant functions on M separates the H–orbits. Therefore, the points $gh \cdot m$ and $g \cdot m$ are in the same H–orbit and hence there exists an element $h' \in H$ such that $gh \cdot m = h'g \cdot m$.

Conversely, suppose that for all $g \in G$, $h \in H$, and $m \in M$, there exists an element $h' \in H$ such that $gh \cdot m = h'g \cdot m$. Then, if $f \in C^\infty(M)^H$ we get

$$f \circ \Phi_g(h \cdot m) = f(gh \cdot m) = f(h'g \cdot m) = f(g \cdot m) = f \circ \Phi_g(m).$$

Consequently, $f \circ \Phi_g \in C^\infty(M)^H$, as required.

(ii). Suppose that the lifted action of G to the tangent bundle TM leaves the H–polar distribution A'_H invariant. We define the action $G \times M/A'_H \to M/A'_H$ by $g \cdot \mathcal{J}_H(m) := \mathcal{J}_H(g \cdot m)$. It is clearly a left action so all we have to do is showing that it is well-defined. Indeed, let $m' \in M$ be such that $m' = \mathcal{F}_T(m)$, with $\mathcal{F}_T \in G_{A'_H}$. For the sake of simplicity in the exposition, suppose that $\mathcal{F}_T = F_T$ with F_T the Hamiltonian flow associated to $f \in C^\infty(M)^H$. Then, for any $g \in G$ we have

$$
\begin{aligned}
g \cdot \mathcal{J}_H(m') &= \mathcal{J}_H(g \cdot F_T(m)) \\
&= \mathcal{J}_H\left(G_T^{f \circ \Phi_{g^{-1}}}(g \cdot m) \right) \\
&= \mathcal{J}_H(g \cdot m) \\
&= g \cdot \mathcal{J}_H(m),
\end{aligned}
$$

where $G_T^{f \circ \Phi_{g^{-1}}}$ is the Hamiltonian flow associated to the function $f \circ \Phi_{g^{-1}}$ that, by the hypothesis on the G–invariance of A'_H, is H–invariant and guarantees the equality $\mathcal{J}_H\left(G_T^{f \circ \Phi_{g^{-1}}}(g \cdot m) \right) = \mathcal{J}_H(g \cdot m)$.

(iii) The inclusion $A'_G(m) \subset A'_H(m)$ is a direct consequence of the definition of the polar distributions and it implies that each maximal integral leaf of A'_G is included in a single maximal integral leaf of A'_H. This feature constitutes the definition of π_H that assigns to each leaf in M/A'_G the unique leaf in M/A'_H in which it is sitting. With this definition it is straightforward that $\mathcal{J}_H = \pi_H \circ \mathcal{J}_G$. Now, if G acts on A'_H the map $\mathcal{J}_H$ is G–equivariant by part (ii). The G–equivariance of $\mathcal{J}_G$ plus the relation $\mathcal{J}_H = \pi_H \circ \mathcal{J}_G$ implies that π_H is G–invariant. ∎

Remark. If H is normal in G then, condition (15.1.1) is trivially satisfied and therefore G acts on A'_H. Conversely, if G acts on A'_H and the identity element is an isotropy subgroup of the G–action on M then H is necessarily normal in G. Indeed, in that case for any $m \in M$, $g \in G$, and $h \in H$, there exists an element $h' \in H$ such that $gh \cdot m = h'g \cdot m$. In particular, if we take an element $m \in M_{\{e\}}$ we obtain $gh = h'g$ or, equivalently that $gHg^{-1} \subset H$, for all $g \in G$, which implies that H is normal in G.

For future reference we state in the following corollary the claims of Proposition 15.1.1 in the particular case in which H is a normal subgroup of G.

15.1.2 Corollary. *Let $(M, \{\cdot, \cdot\})$ be a Poisson manifold acted properly and canonically upon by a Lie group G. Let N be a closed normal Lie subgroup of G. Then:*

(i) *The group G acts on A'_N and on the corresponding momentum space M/A'_N with a natural action that makes the N–optimal momentum map $\mathcal{J}_N : M \to M/A'_N$ G–equivariant.*

(ii) *There is a natural G–equivariant projection $\pi_N : M/A'_G \to M/A'_N$ such that $\mathcal{J}_N = \pi_N \circ \mathcal{J}_G$.*

15.2 Isotropy Subgroups and Quotient Groups

In this section we introduce the relevant groups and spaces for optimal reduction in two stages.

15.2.1 Lemma.   *Let $(M, \{\cdot, \cdot\})$ be a Poisson manifold acted properly and canonically upon by a Lie group G. Let N be a closed normal Lie subgroup of G. Let $\rho \in M/A'_G$ and $\nu := \pi_N(\rho) \in M/A'_N$.*

(i) *Let G_ρ and G_ν be the isotropy subgroups of $\rho \in M/A'_G$ and $\nu := \pi_N(\rho) \in M/A'_N$ with respect to the G–actions on M/A'_G and M/A'_N, respectively. Then, $G_\rho \subset G_\nu$.*

(ii) *Let N_ν be the N–isotropy subgroup of $\nu \in M/A'_N$. Then $N_\nu = N \cap G_\nu$ and N_ν is normal in G_ν.*

(iii) *Endow N_ν and G_ν with the unique smooth structures that make them into initial Lie subgroups of G. Then, N_ν is closed in G_ν and therefore the quotient $H_\nu := G_\nu/N_\nu$ is a Lie group.*

Proof.   Part (i) is a consequence of the G–equivariance of the projection $\pi_N : M/A'_G \to M/A'_N$ and (ii) is straightforward. To prove (iii), let A and B two subsets of a smooth manifold M such that $A \subset B \subset M$. It can be checked by simply using the definition of initial submanifold that if A and B are initial submanifolds of M then A is an initial submanifold of B. In our setup, this fact implies that N_ν is an initial Lie subgroup of G_ν. We actually check that it is a closed Lie subgroup of G_ν. Indeed, let $g \in G_\nu$ be an element in the closure of N_ν in G_ν. Let $\{g_n\}_{n \in \mathbb{N}} \subset N_\nu$ be a sequence of elements in N_ν that converges to g in the topology of G_ν. As G_ν is initial in G we get $g_n \to g$ also in the topology of G. Now, as $\{g_n\}_{n \in \mathbb{N}} \subset N$ and N is closed in G, $g \in N$ necessarily. Hence $g \in N \cap G_\nu = N_\nu$, as required. ∎

Suppose now that the value $\nu \in M/A'_N$ is such that the action of N_ν on the level set $\mathcal{J}_N^{-1}(\nu)$ is proper. We emphasize that this property is not automatically inherited from the properness of the N–action on M. Theorem 13.5.1 guarantees in that situation that the orbit space $M_\nu := \mathcal{J}_N^{-1}(\nu)/N_\nu$ is a smooth symplectic regular quotient manifold with symplectic form ω_ν defined by:

$$\pi_\nu^* \omega_\nu(m)(X_f(m), X_h(m)) = \{f, h\}(m),$$

for any $\in \mathcal{J}_N^{-1}(\nu)$ and any $f, h \in C^\infty(M)^N$. As customary $\pi_\nu : \mathcal{J}_N^{-1}(\nu) \to \mathcal{J}_N^{-1}(\nu)/N_\nu$ denotes the canonical projection and $i_\nu : \mathcal{J}_N^{-1}(\nu) \hookrightarrow M$ the

inclusion. We will refer to the pair (M_ν, ω_ν) as the **first stage reduced space**.

15.2.2 Proposition.   *Let $(M, \{\cdot, \cdot\})$ be a Poisson manifold acted properly and canonically upon by a Lie group G via the map $\Phi : G \times M \to M$. Let N be a closed normal Lie subgroup of G. Let $\rho = \mathcal{J}_G(m) \in M/A'_G$, for some $m \in M$, and $\nu := \pi_N(\rho) = \mathcal{J}_N(m) \in M/A'_N$.*

(i) *If the Lie group N_ν acts properly on the level set $\mathcal{J}_N^{-1}(\nu)$ then the Lie group $H_\nu := G_\nu/N_\nu$ acts smoothly and canonically on the first stage reduced space $(\mathcal{J}_N^{-1}(\nu)/N_\nu, \omega_\nu)$ via the map*

$$gN_\nu \cdot \pi_\nu(m) := \pi_\nu(g \cdot m), \qquad (15.2.1)$$

for all $gN_\nu \in H_\nu$ and $m \in \mathcal{J}_N^{-1}(\nu)$.

(ii) *Suppose that N_ν and H_ν act properly on $\mathcal{J}_N^{-1}(\nu)$ and M_ν, respectively. Let $\mathcal{J}_{H_\nu} : M_\nu \to M_\nu/A'_{H_\nu}$ be the optimal momentum map associated to the H_ν–action on $M_\nu = \mathcal{J}_N^{-1}(\nu)/N_\nu$ and $\sigma = \mathcal{J}_{H_\nu}(\pi_\nu(m))$. Then,*

$$\mathcal{J}_{H_\nu} \left(\pi_\nu(\mathcal{J}_G^{-1}(\rho)) \right) = \sigma. \qquad (15.2.2)$$

Proof.   (i). We first show that the action given by expression (15.2.1) is well-defined and is smooth. The action $\varphi^\nu : G_\nu \times \mathcal{J}_N^{-1}(\nu) \to \mathcal{J}_N^{-1}(\nu)$ obtained by restriction of the domain and range of Φ is smooth since G_ν and $\mathcal{J}_N^{-1}(\nu)$ are initial submanifolds of G and M, respectively. Also, this map is compatible with the action of $N_\nu \times N_\nu$ on $G_\nu \times \mathcal{J}_N^{-1}(\nu)$ via $(n, n') \cdot (g, z) := (gn^{-1}, n' \cdot z)$, and the N_ν–action on $\mathcal{J}_N^{-1}(\nu)$. Indeed, for any $(n, n') \in N_\nu \times N_\nu$ and any $(g, z) \in G_\nu \times \mathcal{J}_N^{-1}(\nu)$, the point $(gn^{-1}, n' \cdot z)$ gets sent by this map to $gn^{-1}n' \cdot z$. As N_ν is normal in G_ν there exists some $n'' \in N_\nu$ such that $gn^{-1}n' \cdot z = n''g \cdot z$ which is in the same N_ν–orbit as $g \cdot z$. Consequently, the map $\varphi^\nu : G_\nu \times \mathcal{J}_N^{-1}(\nu) \to \mathcal{J}_N^{-1}(\nu)$ drops to a smooth map

$$\phi^\nu : G_\nu/N_\nu \times \mathcal{J}_N^{-1}(\nu)/N_\nu \to \mathcal{J}_N^{-1}(\nu)/N_\nu$$

that coincides with (15.2.1) and therefore satisfies $\phi^\nu_{kN_\nu} \circ \pi_\nu = \pi_\nu \circ \varphi^\nu_k$, for any $kN_\nu \in H_\nu$.

We now show that the action given by the map ϕ^ν is canonical. Let $kN_\nu \in H_\nu$, $m \in \mathcal{J}_N^{-1}(\nu)$, and $f, h \in C^\infty(M)^N$ arbitrary. Then, taking into account that $\phi^\nu_{kN_\nu} \circ \pi_\nu = \pi_\nu \circ \varphi^\nu_k$ and that by part (i) in Proposition 15.1.1

the functions $f \circ \Phi_{k^{-1}}$ and $h \circ \Phi_{k^{-1}}$ are N–invariant, we can write:

$$
\pi_\nu^*((\phi_{kN_\nu}^\nu)^*\omega_\nu)(m)(X_f(m), X_h(m))
$$

$$
\begin{aligned}
&= ((\phi_{kN_\nu}^\nu \circ \pi_\nu)^*\omega_\nu)(m)(X_f(m), X_h(m)) \\
&= ((\pi_\nu \circ \varphi_k^\nu)^*\omega_\nu)(m)(X_f(m), X_h(m)) \\
&= (\varphi_k^\nu)^*(\pi_\nu^*\omega_\nu)(m)(X_f(m), X_h(m)) \\
&= \pi_\nu^*\omega_\nu(k \cdot m)(T_m\Phi_k \cdot X_f(m), T_m\Phi_k \cdot X_h(m)) \\
&= \pi_\nu^*\omega_\nu(k \cdot m)(X_{f \circ \Phi_{k^{-1}}}(k \cdot m), X_{h \circ \Phi_{k^{-1}}}(k \cdot m)) \\
&= \{f \circ \Phi_{k^{-1}}, h \circ \Phi_{k^{-1}}\}(k \cdot m) \\
&= \{f, h\}(m) \\
&= \pi_\nu^*\omega_\nu(m)(X_f(m), X_h(m)).
\end{aligned}
$$

Since the map π_ν is a surjective submersion, this chain of equalities implies that $(\phi_{kN_\nu}^\nu)^*\omega_\nu = \omega_\nu$, as required.

(ii) Let $m' \in \mathcal{J}_G^{-1}(\rho)$ be such that $m' \neq m$. Then, there exists $\mathcal{F}_T \in G_{\mathcal{A}'_G}$ such that $m' = \mathcal{F}_T(m)$. For simplicity in the exposition take $\mathcal{F}_T = F_T$, with F_T the Hamiltonian flow associated to the G–invariant function $f \in C^\infty(M)^G$. Let now $f_\nu \in C^\infty(M_\nu)^{H_\nu}$ be the H_ν–invariant function on M_ν uniquely determined by the relation $f_\nu \circ \pi_\nu = f \circ i_\nu$. The Hamiltonian flow F_T^ν associated to f_ν is related to F_T by the relation $F_T^\nu \circ \pi_\nu = \pi_\nu \circ F_T \circ i_\nu$. Therefore, by Noether's Theorem applied to $\mathcal{J}_{H_\nu}$ we obtain:

$$
\begin{aligned}
\mathcal{J}_{H_\nu}(\pi_\nu(m')) &= \mathcal{J}_{H_\nu}(\pi_\nu(F_T(m))) \\
&= \mathcal{J}_{H_\nu}(F_T^\nu(\pi_\nu(m))) \\
&= \mathcal{J}_{H_\nu}(\pi_\nu(m)) \\
&= \sigma,
\end{aligned}
$$

as required.   ∎

15.3   The Optimal Reduction by Stages Theorem

Now we are ready to prove the optimal optimal reduction by stages theorem. Here is the set up. Let $m \in M$ be such that $\rho = \mathcal{J}_G(m)$. Also, let $\nu = \mathcal{J}_N(m)$ and $\sigma = \mathcal{J}_{H_\nu}(\pi_\nu(m))$. The second part of Proposition 15.2.2 guarantees that the restriction of π_ν to $\mathcal{J}_G^{-1}(\rho)$ gives us a well-defined map

$$
\pi_\nu|_{\mathcal{J}_G^{-1}(\rho)} : \mathcal{J}_G^{-1}(\rho) \longrightarrow \mathcal{J}_{H_\nu}^{-1}(\sigma).
$$

This map is smooth because $\mathcal{J}_{H_\nu}^{-1}(\sigma)$ is an initial submanifold of M_ν and also because $\mathcal{J}_G^{-1}(\rho)$ and $\mathcal{J}_N^{-1}(\nu)$ are initial submanifolds of M such that

$\mathcal{J}_G^{-1}(\rho) \subset \mathcal{J}_N^{-1}(\nu) \subset M$, which implies that $\mathcal{J}_G^{-1}(\rho)$ is an initial submanifold of $\mathcal{J}_N^{-1}(\nu)$ (this argument is a straightforward consequence of the definition of initial submanifold). Denote by $i_{\rho,\nu} : \mathcal{J}_G^{-1}(\rho) \hookrightarrow \mathcal{J}_N^{-1}(\nu)$ the corresponding smooth injection. Let $(H_\nu)_\sigma$ be the H_ν–isotropy subgroup of the element $\sigma \in M_\nu/A'_{H_\nu}$. Then, the map $\pi_\nu|_{\mathcal{J}_G^{-1}(\rho)} = \pi_\nu \circ i_{\rho,\nu} : \mathcal{J}_G^{-1}(\rho) \to \mathcal{J}_{H_\nu}^{-1}(\sigma)$ is smooth and $(G_\rho, (H_\nu)_\sigma)$–equivariant. Indeed, let $g \in G_\rho$ and $m \in \mathcal{J}_G^{-1}(\rho)$ arbitrary. By Lemma 15.2.1 we know that as $G_\rho \subset G_\nu$, then $g \in G_\nu$ and $gN_\nu \in G_\nu/N_\nu$. Using Definition 15.2.1, we get $\pi_\nu(g \cdot m) = gN_\nu \cdot \pi_\nu(m)$. Additionally, by (15.2.2) we have

$$\mathcal{J}_{H_\nu}(gN_\nu \cdot \pi_\nu(m)) = \mathcal{J}_{H_\nu}(\pi_\nu(g \cdot m)) = \sigma,$$

because $g \cdot m \in \mathcal{J}_G^{-1}(\rho)$, which shows that $gN_\nu \in (H_\nu)_\sigma$ and therefore guarantees the $(G_\rho, (H_\nu)_\sigma)$–equivariance of $\pi_\nu|_{\mathcal{J}_G^{-1}(\rho)}$. Consequently, the map $\pi_\nu|_{\mathcal{J}_G^{-1}(\rho)}$ drops to a well-defined map F that makes the following diagram

$$
\begin{array}{ccc}
\mathcal{J}_G^{-1}(\rho) & \xrightarrow{\pi_\nu|_{\mathcal{J}_G^{-1}(\rho)}} & \mathcal{J}_{H_\nu}^{-1}(\sigma) \\
\pi_\rho \downarrow & & \downarrow \pi_\sigma \\
\mathcal{J}_G^{-1}(\rho)/G_\rho & \xrightarrow{\quad F \quad} & \mathcal{J}_{H_\nu}^{-1}(\sigma)/(H_\nu)_\sigma.
\end{array}
$$

commutative. We remind the reader once more that the G_ρ and $(H_\nu)_\sigma$–actions on $\mathcal{J}_G^{-1}(\rho)$ and $\mathcal{J}_{H_\nu}^{-1}(\sigma)$, respectively are not automatically proper as a consequence of the properness of the G–action on M. If that happens to be the case, the map F is smooth. Moreover, in that situation Theorem 13.5.1 guarantees that the quotients $M_\rho := \mathcal{J}_G^{-1}(\rho)/G_\rho$ and $(M_\nu)_\sigma := \mathcal{J}_{H_\nu}^{-1}(\sigma)/(H_\nu)_\sigma$ are symplectic manifolds. We will refer to the symplectic manifold $(\mathcal{J}_{H_\nu}^{-1}(\sigma)/(H_\nu)_\sigma, \omega_\sigma)$ as the **second stage reduced space**. Recall that the symplectic form ω_σ is uniquely determined by the equality $\pi_\sigma^* \omega_\sigma = i_\sigma^* \omega_\nu$, where $i_\sigma : \mathcal{J}_{H_\nu}^{-1}(\sigma) \hookrightarrow \mathcal{J}_N^{-1}(\nu)/N_\nu$ is the injection and $\pi_\sigma : \mathcal{J}_{H_\nu}^{-1}(\sigma) \to \mathcal{J}_{H_\nu}^{-1}(\sigma)/(H_\nu)_\sigma$ the projection.

Our goal in this section will consist of proving a theorem that under certain hypotheses states that the map F is a symplectomorphism between the one–shot reduced space $(\mathcal{J}_G^{-1}(\rho)/G_\rho, \omega_\rho)$ and the reduced space in two shots $(\mathcal{J}_{H_\nu}^{-1}(\sigma)/(H_\nu)_\sigma, \omega_\sigma)$.

Given that the properness assumptions appear profusely we will simplify the exposition by grouping them all in the following definition.

15.3.1 Definition. *Let $(M, \{\cdot, \cdot\})$ be a Poisson manifold acted properly and canonically upon by a Lie group G via the map $\Phi : G \times M \to M$. Let N be a closed normal Lie subgroup of G. Let $\rho \in M/A'_G$, $\nu := \pi_N(\rho)$, $H_\nu := G_\nu/N_\nu$, and $\sigma = \mathcal{J}_{H_\nu}\left(\pi_\nu(\mathcal{J}_G^{-1}(\rho))\right) \in M_\nu/A'_{H_\nu}$. We will say that* **we have a proper action at** ρ *whenever G_ρ acts properly on $\mathcal{J}_G^{-1}(\rho)$, N_ν acts properly on $\mathcal{J}_N^{-1}(\nu)$, H_ν acts properly on $\mathcal{J}_N^{-1}(\nu)/N_\nu$, and $(H_\nu)_\sigma$ acts properly on $\mathcal{J}_{H_\nu}^{-1}(\sigma)$.*

Let $(G_\nu)_\sigma \subset G_\nu$ be the unique subgroup of G_ν such that $(H_\nu)_\sigma = (G_\nu)_\sigma/N_\nu$. We say that the element $\rho \in M/A'_G$ satisfies the **stages hypothesis** when for any other element $\rho' \in M/A'_G$ such that

$$\pi_N(\rho) = \pi_N(\rho') =: \nu \qquad (15.3.1)$$

and

$$\mathcal{J}_{H_\nu}(\pi_\nu(\mathcal{J}_G^{-1}(\rho))) = \mathcal{J}_{H_\nu}(\pi_\nu(\mathcal{J}_G^{-1}(\rho'))) = \sigma \qquad (15.3.2)$$

there exists an element $h \in (G_\nu)_\sigma$ such that $\rho' = h \cdot \rho$.

We say that the element $\nu \in M/A'_N$ has the **local extension property** when any function $f \in C^\infty(\mathcal{J}_N^{-1}(\nu))^{G_\nu}$ is such that for any $m \in \mathcal{J}_N^{-1}(\nu)$ there is an open N–invariant neighborhood U of m and a function $F \in C^\infty(M)^G$ such that $F|_U = f|_U$.

15.3.2 Theorem (Optimal Reduction by Stages). *Suppose that $(M, \{\cdot, \cdot\})$ is a Poisson manifold acted properly and canonically upon by a Lie group G via the map $\Phi : G \times M \to M$. Let N be a closed normal Lie subgroup of G. Let $\rho \in M/A'_G$, $\nu := \pi_N(\rho)$, $H_\nu := G_\nu/N_\nu$, and $\sigma = \mathcal{J}_{H_\nu}\big(\pi_\nu(\mathcal{J}_G^{-1}(\rho))\big) \in M_\nu/A'_{H_\nu}$. Then, if ρ satisfies the stages hypothesis, we have a proper action at ρ, and the quotient manifold $\mathcal{J}_G^{-1}(\rho)/G_\rho$ is either Lindelöf or paracompact, the map*

$$
\begin{array}{cccc}
F : & \big(\mathcal{J}_G^{-1}(\rho)/G_\rho, \omega_\rho\big) & \longrightarrow & \big(\mathcal{J}_{H_\nu}^{-1}(\sigma)/(H_\nu)_\sigma, \omega_\sigma\big) \\
& \pi_\rho(m) & \longmapsto & \pi_\sigma(\pi_\nu(m))
\end{array}
$$

is a symplectomorphism between the one shot reduced space $(\mathcal{J}_G^{-1}(\rho)/G_\rho, \omega_\rho)$ and $(\mathcal{J}_{H_\nu}^{-1}(\sigma)/(H_\nu)_\sigma, \omega_\sigma)$ that was obtained by reduction in two stages.

Proof of the Theorem. First we prove that F is injective. let $\pi_\rho(m)$ and $\pi_\rho(m') \in \mathcal{J}_G^{-1}(\rho)/G_\rho$ be such that $F(\pi_\rho(m)) = F(\pi_\rho(m'))$. By the definition of F this implies that $\pi_\sigma(\pi_\nu(m)) = \pi_\sigma(\pi_\nu(m'))$. Hence, there exists an element $gN_\nu \in (H_\nu)_\sigma$ such that $\pi_\nu(m') = gN_\nu \cdot \pi_\nu(m)$ which, by the definition (15.2.1), is equivalent to $\pi_\nu(m') = \pi_\nu(g \cdot m)$. Therefore, there exists a $n \in N_\nu$ such that $m' = ng \cdot m$. However, since both m and m' sit in $\mathcal{J}_G^{-1}(\rho)$ we have $ng \in G_\rho$, necessarily and, consequently $\pi_\rho(m) = \pi_\rho(m')$, as required.

Next we prove that F is surjective. To do this, let $\pi_\sigma(\bar{z}) \in (M_\nu)_\sigma = \mathcal{J}_{H_\nu}^{-1}(\sigma)/(H_\nu)_\sigma$. Take any $z \in \mathcal{J}_N^{-1}(\nu)$ such that $\pi_\nu(z) = \bar{z}$ and let $\rho' := \mathcal{J}_G(z)$. It is clear that

$$\pi_N(\rho') = \pi_N(\mathcal{J}_G(z)) = \mathcal{J}_N(z) = \nu = \pi_N(\rho)$$

and also, as $\mathcal{J}_{H_\nu}(\pi_\nu(z)) = \sigma$, Proposition 15.2.2 guarantees that

$$\mathcal{J}_{H_\nu}(\pi_\nu(\mathcal{J}_G^{-1}(\rho'))) = \sigma = \mathcal{J}_{H_\nu}(\pi_\nu(\mathcal{J}_G^{-1}(\rho))).$$

By the stages hypothesis, there exists $h \in (G_\nu)_\sigma$ such that $\rho' = h \cdot \rho$. Notice that

$$
\begin{aligned}
F(\pi_\rho(h^{-1} \cdot z)) &= \pi_\sigma(\pi_\nu(h^{-1} \cdot z)) \\
&= \pi_\sigma(h^{-1} N_\nu \cdot \pi_\nu(z)) \\
&= \pi_\sigma(\pi_\nu(z)) \\
&= \pi_\sigma(\bar{z}),
\end{aligned}
$$

which proves the surjectivity of F.

We next turn to the proof that F is a symplectic map. We will do this by showing that $F^* \omega_\sigma = \omega_\rho$. Let $m \in \mathcal{J}_G^{-1}(\rho)$ and $f, g \in C^\infty(M)^G$ arbitrary. Then,

$$
\begin{aligned}
\pi_\rho^*(F^*\omega_\sigma)(m)(X_f(m), X_g(m)) &= (F \circ \pi_\rho)^* \omega_\sigma(m)(X_f(m), X_g(m)) \\
&= \left(\pi_\sigma \circ \pi_\nu|_{\mathcal{J}_G^{-1}(\rho)}\right)^* \omega_\sigma(m)(X_f(m), X_g(m)) \\
&= (\pi_\sigma \circ \pi_\nu \circ i_{\rho,\nu})^* \omega_\sigma(m)(X_f(m), X_g(m)) \\
&= ((\pi_\nu \circ i_{\rho,\nu})^*(\pi_\sigma^*\omega_\sigma))(m)(X_f(m), X_g(m)) \\
&= ((\pi_\nu \circ i_{\rho,\nu})^*(i_\sigma^*\omega_\nu))(m)(X_f(m), X_g(m)) \\
&= (i_\sigma \circ \pi_\nu \circ i_{\rho,\nu})^* \omega_\nu(m)(X_f(m), X_g(m)) \\
&= \pi_\nu^*\omega_\nu(m)(X_f(m), X_g(m)) = \{f, g\}(m) \\
&= \pi_\rho^*\omega_\rho(m)(X_f(m), X_g(m)).
\end{aligned}
$$

This chain of equalities guarantees that $\pi_\rho^*(F^*\omega_\sigma) = \pi_\rho^*\omega_\rho$. Since the map π_ρ is a surjective submersion, we see that $F^*\omega_\sigma = \omega_\rho$, and consequently F is a symplectic map.

Finally, we show that F is a symplectomorphism. Given that F is a bijective symplectic map, it is necessarily an immersion. Since by hypothesis the space $\mathcal{J}_G^{-1}(\rho)/G_\rho$ is either Lindelöf or paracompact, a standard result in manifolds theory guarantees that F is actually a diffeomorphism. ∎

15.3.3 Proposition.   *Let $(M, \{\cdot, \cdot\})$ be a Poisson manifold acted properly and canonically upon by a Lie group G via the map $\Phi : G \times M \to M$. Let N be a closed normal Lie subgroup of G. Let $\rho \in M/A_G'$, $\nu := \pi_N(\rho)$, $H_\nu := G_\nu/N_\nu$, and $\sigma = \mathcal{J}_{H_\nu}\left(\pi_\nu(\mathcal{J}_G^{-1}(\rho))\right) \in M_\nu/A_{H_\nu}'$. If ν has the local extension property and N_ν acts properly on $\mathcal{J}_N^{-1}(\nu)$, then $\pi_\nu(\mathcal{J}_G^{-1}(\rho)) = \mathcal{J}_{H_\nu}^{-1}(\sigma)$ and ρ satisfies the stages hypothesis.*

Proof.   The inclusion $\pi_\nu(\mathcal{J}_G^{-1}(\rho)) \subset \mathcal{J}_{H_\nu}^{-1}(\sigma)$ is guaranteed by (15.2.2). In order to prove the equality take $\pi_\nu(m) \in \pi_\nu(\mathcal{J}_G^{-1}(\rho))$ and $f \in C^\infty(M_\nu)^{H_\nu}$ arbitrary, such that the Hamiltonian vector field X_f on M_ν has flow F_t. Let $\bar{f} \in C^\infty(\mathcal{J}_N^{-1}(\nu))^{G_\nu}$ be the function defined by $\bar{f} := f \circ \pi_\nu$. The H_ν–invariance of f implies that $\bar{f}$ is G_ν–invariant. In principle, the point

$F_T(\pi_\nu(m))$ lies somewhere in $\mathcal{J}_{H_\nu}^{-1}(\sigma)$. However, we will show that it actually stays in $\pi_\nu(\mathcal{J}_G^{-1}(\rho))$, which will prove the desired equality. Indeed, as the curve $\{F_t(\pi_\nu(m))\}_{t\in[0,T]}$ is compact it can be covered by a finite number of open sets $\{U_1,\ldots,U_n\}$. Suppose that we have chosen the neighborhoods U_i such that $\pi_\nu(m) \in U_1$, $F_T(\pi_\nu(m)) \in U_n$, $U_i \cap U_j \neq \varnothing$ if and only if $|j-i| = 1$, and for each open N–invariant set $\pi_\nu^{-1}(U_i)$, there is a $g_i \in C^\infty(M)^G$ such that $\bar{f}|_{\pi_\nu^{-1}(U_i)} = g_i|_{\pi_\nu^{-1}(U_i)}$. where the function $\bar{f}$ admits local extensions to G–invariant functions on M. We call G_t^i the flow of the Hamiltonian vector field X_{g_i} on M associated to $g_i \in C^\infty(M)^G$. The flows G_t^i and F_t are related by the equality

$$F_t \circ \pi_\nu|_{\mathcal{J}_N^{-1}(\nu)\cap\pi_\nu^{-1}(U_i)} = \pi_\nu \circ G_t^i \circ i_\nu|_{\mathcal{J}_N^{-1}(\nu)\cap\pi_\nu^{-1}(U_i)}.$$

Due to the G–invariance of g we have $\mathcal{J}_G \circ G_t^i = \mathcal{J}_G$ and, consequently

$$\{F_t(\pi_\nu(m))\}_{t\in[0,T]} \subset \pi_\nu(\mathcal{J}_G^{-1}(\rho)),$$

as required. This proves that $\pi_\nu(\mathcal{J}_G^{-1}(\rho)) = \mathcal{J}_{H_\nu}^{-1}(\sigma)$.

We conclude by showing that this equality implies that ρ satisfies the stages hypothesis. Indeed, if $\rho' \in M/A_G'$ is such that $\mathcal{J}_{H_\nu}(\pi_\nu(\mathcal{J}_G^{-1}(\rho'))) = \sigma$, then $\pi_\nu(\mathcal{J}_G^{-1}(\rho')) \subset \mathcal{J}_{H_\nu}^{-1}(\sigma) = \pi_\nu(\mathcal{J}_G^{-1}(\rho))$. Consequently, for any $\pi_\nu(z') \in \pi_\nu(\mathcal{J}_G^{-1}(\rho'))$, $z' \in \mathcal{J}_G^{-1}(\rho')$, there exists an element $z \in \mathcal{J}_G^{-1}(\rho)$ such that $\pi_\nu(z') = \pi_\nu(z)$. Hence, there is an element $n \in N_\nu \subset (G_\nu)_\sigma$ available such that $z' = n \cdot z$ which, by applying the map $\mathcal{J}_G$ to both sides of this equality implies that $\rho' = n \cdot \rho$. ∎

15.4 Optimal Orbit Reduction by Stages

The goal of this section is formulating and proving the analog of Theorem 15.3.2 in the context of orbit reduction. The setup is identical to the one in the previous section but this time we will determine under what conditions the orbit reduced space $\mathcal{J}^{-1}(\mathcal{O}_\rho)/G$ can be constructed in two stages. Unlike the situation with point reduction, in the orbit reduction context there will be choices involved in the construction of the first stage reduced space.

Let $\rho \in M/A_G'$ and $\mathcal{O}_\rho := G \cdot \rho$ be its G-orbit in M/A_G'. Let $\nu \in M/A_N'$ be given by $\nu \in \pi_N(\rho)$ and $N\cdot\nu$ its corresponding N-orbit in M/A_N'. Define

$$G_{N\cdot\nu} := \{g \in G \mid g \cdot (N \cdot \nu) \subset N \cdot \nu\}.$$

An argument identical to the one in Lemma 5.3.2 shows that $G_{N\cdot\nu}$ is a closed Lie subgroup of G containing N as a normal subgroup. Suppose that N_ν acts properly on $\mathcal{J}_N^{-1}(\nu)$. Then the symplectic orbit reduced space $M_{N\cdot\nu} := \mathcal{J}_N^{-1}(N\cdot\nu)/N$ is canonically acted upon by the quotient Lie group

$H_{N \cdot \nu} := G_{N \cdot \nu}/N$. Let $\mathcal{J}_{H_{N \cdot \nu}} : M_{N \cdot \nu} \to M_{N \cdot \nu}/A'_{H_{N \cdot \nu}}$ be the corresponding momentum map. An argument identical to the one followed in the proof of part (ii) of Proposition 15.2.2 shows that

$$\mathcal{J}_{H_{N \cdot \nu}} \left(\pi_{N \cdot \nu}(\mathcal{J}_G^{-1}(\rho)) \right) = \sigma,$$

for some $\sigma \in M_{N \cdot \nu}/A'_{H_{N \cdot \nu}}$ and where $\pi_{N \cdot \nu} : \mathcal{J}_N^{-1}(N \cdot \nu) \to \mathcal{J}_N^{-1}(N \cdot \nu)/N$ is the projection.

15.4.1 Definition. *Let $(M, \{\cdot, \cdot\})$ be a Poisson manifold acted properly and canonically upon by a Lie group G via the map $\Phi : G \times M \to M$. Let N be a closed Lie subgroup of G. Let $\rho \in M/A'_G$, $\nu \in \pi_N(\rho)$, $H_{N \cdot \nu} := G_{N \cdot \nu}/N$, and $\sigma := \mathcal{J}_{H_{N \cdot \nu}} \left(\pi_{N \cdot \nu}(\mathcal{J}_G^{-1}(\rho)) \right)$. We say that we have a **proper action at the orbit** $G \cdot \rho$ whenever the group G_ρ acts properly on $\mathcal{J}_G^{-1}(\rho)$, $N \cdot \nu$ acts properly on $\mathcal{J}_N^{-1}(\nu)$, $H_{N \cdot \nu}$ acts properly on $M_{N \cdot \nu}$, and $(H_{N \cdot \nu})_\sigma$ acts properly on $\mathcal{J}_{H_{N \cdot \nu}}^{-1}(\sigma)$.*

Remark. The independence of the previous hypothesis on the choice of the element ρ in the orbit $G \cdot \rho$ is easy to verify by using the equalities that follow. Let $\rho' = g \cdot \rho$ for some $g \in G$ and $\nu' = \pi_N(\rho')$. Then

$$G_{\rho'} = gG_\rho g^{-1}; \quad N_{\nu'} = gN_\nu g^{-1},$$

$$H_{N \cdot \nu'} = (gN)H_{N \cdot \nu}(gN)^{-1},$$

$$\mathcal{J}_G^{-1}(\rho') = g \cdot \mathcal{J}_G^{-1}(\rho); \quad \mathcal{J}_N^{-1}(\nu') = g \cdot \mathcal{J}_N^{-1}(\nu),$$

$$M_{N \cdot \nu'} = gN \cdot M_{N \cdot \nu}; \quad A'_{H_{N \cdot \nu'}} = T\varphi_{gN} \cdot A'_{H_{N \cdot \nu}},$$

and

$$\mathcal{J}_{H_{N \cdot \nu'}} = [\varphi_{gN}] \circ \mathcal{J}_{H_{N \cdot \nu}} \circ \varphi_{(gN)^{-1}},$$

where $\varphi : G/N \times M/N \to M/N$ is the naturally induced G/N-action on M/N by $\Phi : G \times M \to M$ and $[\varphi_{gN}]$ is the unique map that makes the diagram

$$
\begin{array}{ccc}
M_{N \cdot \nu} & \xrightarrow{\varphi_{gN}} & M_{N \cdot \nu'} \\
\mathcal{J}_{H_{N \cdot \nu}} \downarrow & & \downarrow \mathcal{J}_{H_{N \cdot \nu'}} \\
M_{N \cdot \nu}/A'_{H_{N \cdot \nu}} & \xrightarrow{[\varphi_{gN}]} & M_{N \cdot \nu'}/A'_{H_{N \cdot \nu'}}.
\end{array}
$$

commutative.

Additionally, since $\pi_{N \cdot \nu'} \circ \Phi_g = \varphi_{gN} \circ \pi_{N \cdot \nu}$ it is easy to show that the element $\sigma' := \mathcal{J}_{H_{N \cdot \nu'}} \left(\pi_{N \cdot \nu'}(\mathcal{J}_G^{-1}(\rho')) \right)$ is related to σ by $\sigma' = [\varphi_{gN}](\sigma)$. It then follows that

$$(H_{N \cdot \nu'})_{\sigma'} = (gN)(H_{N \cdot \nu})_\sigma(gN)^{-1} \quad \text{and} \quad \mathcal{J}_{H_{N \cdot \nu'}}^{-1}(\sigma') = \varphi_{gN} \left(\mathcal{J}_{H_{N \cdot \nu}}^{-1}(\sigma) \right).$$

15.4.2 Definition.   *We say that the orbit $G \cdot \rho \in M/A'_G$ satisfies the* **orbit stages hypothesis** *when for one (and hence for any) element $\rho \in G \cdot \rho$ the following condition holds: for any other element $\rho' \in M/A'_G$ satisfying*

$$\pi_N(\rho) = \pi_N(\rho') =: \nu \tag{15.4.1}$$

and

$$\mathcal{J}_{H_{N \cdot \nu}}(\pi_{N \cdot \nu}(\mathcal{J}_G^{-1}(\rho))) = \mathcal{J}_{H_{N \cdot \nu}}(\pi_{N \cdot \nu}(\mathcal{J}_G^{-1}(\rho'))) =: \tau \tag{15.4.2}$$

there exists an element $h \in G_{N \cdot \nu}$ such that $\rho' = h \cdot \rho$.

Remarks.
1. The identities in the Remark preceding this definition together with the equality $gG_{N \cdot \nu}g^{-1} = G_{N \cdot g\nu}$ guarantee that the formulation of the stages hypothesis is independent of the choice of the element ρ in the orbit $G \cdot \rho$.

2. If the orbit stages hypothesis holds then the element $h \in G_{N \cdot \nu}$ for which $\rho' = h \cdot \rho$ belongs necessarily to $G_\nu \cap (G_{N \cdot \nu})_\tau$, where $(G_{N \cdot \nu})_\tau$ is uniquely determined by the equality $(G_{N \cdot \nu})_\tau/N = (G_{N \cdot \nu}/N)_\tau$.

15.4.3 Theorem (Optimal Orbit Reduction by Stages).   *Let $(M, \{\cdot, \cdot\})$ be a Poisson manifold acted canonically and properly upon by a Lie group G. Let N be a closed normal Lie subgroup of G. Let $G \cdot \rho \subset M/A'_G$ be an orbit that has a proper action and that satisfies the orbit stages hypothesis, and suppose that the quotient $\mathcal{J}_G^{-1}(G \cdot \rho)/G$ is Lindelöf or paracompact. Then for any $\rho \in G \cdot \rho$, $\nu := \pi_N(\rho)$, $\tau := \mathcal{J}_{H_{N \cdot \nu}}(\pi_{N \cdot \nu}(\mathcal{J}_G^{-1}(\rho)))$, the map*

$$F : \mathcal{J}_G^{-1}(G \cdot \rho)/G \longrightarrow \mathcal{J}_{H_{N \cdot \nu}}^{-1}(H_{N \cdot \nu} \cdot \nu)/H_{N \cdot \nu}$$

defined by $F(\pi_{G \cdot \rho}(z)) = \pi_{H_{N \cdot \nu}}(\pi_{N \cdot \nu}(z))$, where z has been chosen in $\mathcal{J}_G^{-1}(\rho)$, is a symplectomorphism.

Proof.   We will proceed in several steps.
(i). **The map F is well-defined**. Recall that if the point z has been chosen in $\mathcal{J}_G^{-1}(\rho) \subset \mathcal{J}_N^{-1}(\nu)$ then $\mathcal{J}_{H_{N \cdot \nu}}(\pi_{N \cdot \nu}(z)) = \tau$ necessarily and hence

$$F(\pi_{G \cdot \rho}(z)) \in \mathcal{J}_{H_{N \cdot \nu}}^{-1}(H_{N \cdot \nu} \cdot \nu)/H_{N \cdot \nu}.$$

We now show that the definition of the map F does not depend on the element $z \in \mathcal{J}_G^{-1}(\rho)$ used to define it. Indeed let $z' \in \mathcal{J}_G^{-1}(\rho)$ be such that $\pi_{G \cdot \rho}(z) = \pi_{G \cdot \rho}(z')$. This equality implies the existence of an element $g \in G$ such that $z' = g \cdot z$. Since $\rho = \mathcal{J}_G(z') = g \cdot \mathcal{J}_G(z) = g \cdot \rho$ then $g \in G_\rho$ necessarily and hence

$$\pi_{H_{N \cdot \nu}}(\pi_{N \cdot \nu}(z')) = \pi_{H_{N \cdot \nu}}(\pi_{N \cdot \nu}(g \cdot z))$$
$$= \pi_{H_{N \cdot \nu}}(gN \cdot \pi_{N \cdot \nu}(z)) = \pi_{H_{N \cdot \nu}}(\pi_{N \cdot \nu}(z)),$$

as required. Note that in the second equality we have used the inclusion $G_\rho \subset G_{N\cdot\nu}$.

(ii) **The map F is injective.** Let $\pi_{G\cdot\rho}(z)$ and $\pi_{G\cdot\rho}(z')$ be two elements in the orbit space $\mathcal{J}_G^{-1}(G\cdot\rho)/G$ such that $z, z' \in \mathcal{J}_G^{-1}(\rho)$ and $F(\pi_{G\cdot\rho}(z)) = F(\pi_{G\cdot\rho}(z'))$. By definition this is equivalent to the identity $\pi_{H_{N\cdot\nu}}(\pi_{N\cdot\nu}(z)) = \pi_{H_{N\cdot\nu}}(\pi_{N\cdot\nu}(z'))$. This equality implies the existence of an element $gN \in G_{N\cdot\nu}/N$ such that $\pi_{N\cdot\nu}(z') = gN\pi_{N\cdot\nu}(z)$ and hence $\pi_{N\cdot\nu}(z') = \pi_{N\cdot\nu}(g\cdot z)$, which in turn ensures the existence of an element $n \in N$ such that $z' = ng\cdot z$. Since both elements z, z' belong to $\mathcal{J}_G(\rho)$ then $ng \in G_\rho$ and hence $\pi_{G\cdot\rho}(z) = \pi_{G\cdot\rho}(z')$, as required.

(iii) **The map F is surjective.** Let $y = \pi_{H_{N\cdot\nu}}(\pi_{N\cdot\nu}(z))$ be an arbitrary element of the quotient $\mathcal{J}_{H_{N\cdot\nu}}^{-1}(H_{N\cdot\nu}\cdot\nu)/H_{N\cdot\nu}$, for some $z \in \mathcal{J}_N^{-1}(N\cdot\nu)$ such that $\mathcal{J}_{H_{N\cdot\nu}}(\pi_{N\cdot\nu}(z)) = gN\cdot\tau$, for some $g \in G_{N\cdot\nu}$. Notice that

$$\mathcal{J}_{H_{N\cdot\nu}}(\pi_{N\cdot\nu}(g^{-1}\cdot z)) = \mathcal{J}_{H_{N\cdot\nu}}(g^{-1}N\cdot\pi_{N\cdot\nu}(z))$$
$$= g^{-1}N\cdot\mathcal{J}_{H_{N\cdot\nu}}(\pi_{N\cdot\nu}(z)) = \tau. \quad (15.4.3)$$

Since $g^{-1}\cdot z \in \mathcal{J}_N^{-1}(N\cdot\nu)$, there exists an element $n \in N$ for which $\mathcal{J}_N(g^{-1}\cdot z) = n\cdot\nu$. Consequently, the point $x := n^{-1}g^{-1}\cdot z \in \mathcal{J}_N^{-1}(\nu)$. Let $\rho' := \mathcal{J}_G(x)$. Notice that

$$\pi_N(\rho') = \pi_N(\mathcal{J}_G(x)) = \mathcal{J}_N(x) = \nu = \pi_N(\rho).$$

Additionally, using (15.4.3) we can write

$$\mathcal{J}_{H_{N\cdot\nu}}(\pi_{N\cdot\nu}(\mathcal{J}_G^{-1}(\rho'))) = \mathcal{J}_{H_{N\cdot\nu}}(\pi_{N\cdot\nu}(x)) = \mathcal{J}_{H_{N\cdot\nu}}(\pi_{N\cdot\nu}(n^{-1}g^{-1}\cdot z))$$
$$= \mathcal{J}_{H_{N\cdot\nu}}(\pi_{N\cdot\nu}(g^{-1}\cdot z)) = \tau = \mathcal{J}_{H_{N\cdot\nu}}(\pi_{N\cdot\nu}(\mathcal{J}_G^{-1}(\rho))).$$

By the orbit stages hypothesis there exists $h \in G_{N\cdot\nu}$ such that $\rho' = h\cdot\rho$. Consequently, $h^{-1}\cdot x \in \mathcal{J}_G^{-1}(\rho)$ and hence, using the normality of N in G and that $h^{-1}g^{-1} \in G_{N\cdot\nu}$, we have

$$F(\pi_{G\cdot\rho}(h^{-1}\cdot x)) = \pi_{H_{N\cdot\nu}}(\pi_{N\cdot\nu}(h^{-1}n^{-1}g^{-1}\cdot z)) = \pi_{H_{N\cdot\nu}}(\pi_{N\cdot\nu}(h^{-1}g^{-1}\cdot z))$$
$$= \pi_{H_{N\cdot\nu}}(h^{-1}g^{-1}N\cdot\pi_{N\cdot\nu}(z)) = \pi_{H_{N\cdot\nu}}(\pi_{N\cdot\nu}(z)).$$

(iv) **The map F is smooth.** Let $\psi : \mathcal{J}_G^{-1}(G\cdot\rho) \to G\times_{G_\rho}\mathcal{J}_G^{-1}(\rho)$ be the diffeomorphism introduced in Definition 14.1.3. Let $\pi_{G\cdot\rho}(z) \in \mathcal{J}_G^{-1}(G\cdot\rho)/G$ be arbitrary and $s : U \subset \mathcal{J}_G^{-1}(G\cdot\rho)/G \to \mathcal{J}_G^{-1}(G\cdot\rho)$ be a local section of $\pi_{G\cdot\rho}$ in a neighborhood U of $\pi_G(z)$. Let $t : V \subset G\times_{G_\rho}\mathcal{J}_G^{-1}(\rho) \to G\times\mathcal{J}_G^{-1}(\rho)$ be a local section of the projection $\pi_{G_\rho} : G\times\mathcal{J}^{-1}(\rho) \to G\times_{G_\rho}\mathcal{J}_G^{-1}(\rho)$ defined on a local neighborhood V of $\psi(s(\pi_{G\cdot\rho(z)}))$. Denote by $p_{\mathcal{J}_G^{-1}(\rho)} : G\times\mathcal{J}_G^{-1}(\rho) \to \mathcal{J}_G^{-1}(\rho)$ the projection onto the second factor. Since we can locally write

$$p_{\mathcal{J}_G^{-1}(\rho)} \circ t \circ \psi \circ s \circ \pi_{G\cdot\rho} = \pi_{G\cdot\rho}$$

then the map F can be locally expressed as the composition of the smooth functions

$$F(\pi_{G \cdot \rho}(y)) = \pi_{H_{N \cdot \nu}}(\pi_{N \cdot \nu}(p_{\mathcal{J}_G^{-1}(\rho)}(t(\psi(s(\pi_{G \cdot \rho}(y))))))),$$

which implies that F is smooth.

(v) **The map F is a symplectic diffeomorphism.** An argument similar to that in the proof of Theorem 15.3.2 shows that F is a symplectic map and hence and immersion. Since $\mathcal{J}_G^{-1}(G \cdot \rho)/G$ is by hypothesis Lindelöf or paracompact then F is necessarily a symplectic diffeomorphism. ∎

Remark. We now check that even though the point and orbit stages hypotheses look slightly different they are actually equivalent. In order to prove this statement we notice that there exists a unique $(H_\nu, H_{N \cdot \nu})$-equivariant bijection $\mathcal{L}_\nu : M_\nu/A'_{H_\nu} \to M_{N \cdot \nu}/A'_{H_{N \cdot \nu}}$ that makes the diagram

$$
\begin{array}{ccc}
\mathcal{J}_N^{-1}(\nu) & \xrightarrow{\quad l_\nu \quad} & \mathcal{J}_N^{-1}(N \cdot \nu) \\
\pi_\nu \downarrow & & \downarrow \pi_{N \cdot \nu} \\
M_\nu = \mathcal{J}_N^{-1}(\nu)/N \cdot \nu & \xrightarrow{\quad L_\nu \quad} & M_{N \cdot \nu} = \mathcal{J}_N^{-1}(N \cdot \nu)/N \\
\mathcal{J}_{H_\nu} \downarrow & & \downarrow \mathcal{J}_{H_{N \cdot \nu}} \\
M_\nu/A'_{H_\nu} & \xrightarrow{\quad \mathcal{L}_\nu \quad} & M_{N \cdot \nu}/A'_{H_{N \cdot \nu}}.
\end{array}
$$

commutative. Here l_ν is the inclusion and L_ν is the $(H_\nu, H_{N \cdot \nu})$-equivariant symplectic diffeomorphism induced by l_ν. The map $\mathcal{L}_\nu$ is well-defined as a consequence of the compatibility of L_ν with the equivalence relations induced by the distributions A'_{H_ν} and $A'_{H_{N \cdot \nu}}$ on M_ν and $M_{N \cdot \nu}$, respectively. This is indeed so since if $z, z' \in M_\nu$ are such that $z' = F_t(z)$ with F_t the Hamiltonian flow associated to a function $f \in C^\infty(M_\nu)^{H_\nu}$ then $L_\nu(z') = \overline{F_t}(L_\nu(z))$, where $\overline{F_t}$ is the Hamiltonian flow associated to $f \circ L_\nu^{-1} \in C^\infty(M_{N \cdot \nu})^{H_{N \cdot \nu}}$.

Suppose now that the element $\rho \in M/A'_G$ satisfies the point stages hypothesis. We shall prove that the orbit $G \cdot \rho$ satisfies the orbit stages hypothesis. Since by the remark that follows Definition 15.4.2 any element in the orbit $G \cdot \rho$ can be chosen to formulate this hypothesis, we can take ρ without loss of generality. Let $\rho' \in M/A'_G$ be such that (15.4.1) and (15.4.2) hold. Therefore, condition (15.3.1) obviously holds. Additionally, using the diagram where we defined the map $\mathcal{L}_\nu$ and the fact that $\mathcal{J}_G^{-1}(\rho) \subset \mathcal{J}_N^{-1}(\nu) \subset \mathcal{J}_N^{-1}(N \cdot \nu)$, $\mathcal{J}_G^{-1}(\rho') \subset \mathcal{J}_N^{-1}(\nu) \subset \mathcal{J}_N^{-1}(N \cdot \nu)$, we have

$$\mathcal{J}_{H_{N \cdot \nu}}(\pi_{N \cdot \nu}(\mathcal{J}_G^{-1}(\rho))) = \mathcal{L}_\nu \left(\mathcal{J}_{H_\nu}(\pi_\nu(\mathcal{J}_G^{-1}(\rho))) \right), \quad \text{and} \qquad (15.4.4)$$

$$\mathcal{J}_{H_{N \cdot \nu}}(\pi_{N \cdot \nu}(\mathcal{J}_G^{-1}(\rho'))) = \mathcal{L}_\nu \left(\mathcal{J}_{H_\nu}(\pi_\nu(\mathcal{J}_G^{-1}(\rho'))) \right). \qquad (15.4.5)$$

This implies that

$$\mathcal{J}_{H_\nu}(\pi_\nu(\mathcal{J}_G^{-1}(\rho))) = \mathcal{J}_{H_\nu}(\pi_\nu(\mathcal{J}_G^{-1}(\rho'))) = \mathcal{L}_\nu^{-1}(\tau)$$

and hence there exists an element $h \in (G_\nu)_{\mathcal{L}_\nu^{-1}(\tau)} \subset G_{N\cdot\nu}$ such that $\rho' = h \cdot \rho$ which proves that the orbit $G \cdot \rho$ satisfies the orbit stages hypothesis.

Conversely, suppose that $G \cdot \rho$ satisfies the orbit stages hypothesis and let $\rho' \in M/A_G'$ be such that the conditions (15.3.1) and (15.3.2) hold. By (15.4.4) and (15.4.5) the identities (15.4.1) and (15.4.2) hold and hence there exists a group element $h \in G_{N\cdot\nu}$ such that $\rho' = h \cdot \rho$. As we noticed in the second remark that follows Definition 15.4.2, the element h actually belongs to $G_\nu \cap (G_{N\cdot\nu})_{\mathcal{L}_\nu(\sigma)} = (G_\nu)_\sigma$ which proves the implication. The previous equality follows from the bijectivity and the $(H_\nu, H_{N\cdot\nu})$-equivariance of $\mathcal{L}_\nu$.

15.5   Reduction by Stages of Globally Hamiltonian Actions

In this section we will assume that M is a symplectic manifold and that the G–action is proper and canonical, has a standard $\mathfrak{g}^*$–valued equivariant momentum map $\mathbf{J}_G : M \to \mathfrak{g}^*$, and that, as usual, it contains a closed normal subgroup $N \subset G$. Recall that the inclusion $N \subset G$ and the normal character of N in G implies that $\mathfrak{n}$ is an ideal in $\mathfrak{g}$. Let $i : \mathfrak{n} \hookrightarrow \mathfrak{g}$ be the inclusion. As a corollary to these remarks, it is easy to conclude (see Part II) that the N–action on M is also globally Hamiltonian with a G–equivariant momentum map $\mathbf{J}_N : M \to \mathfrak{n}^*$ given by $\mathbf{J}_N = i^*\mathbf{J}_G$.

When the G–action on M is free we fall in the situation studied in Part II. In the following pages we will see how our understanding of the optimal reduction by stages procedure allows us to generalize the results that we obtained in that part the nonfree actions case. More specifically, we will see that the reduced spaces and subgroups involved in the Optimal Reduction by Stages Theorem 15.3.2 admit in this case a very precise characterization in terms of level sets of the standard momentum maps present in the problem, and of various subgroups of G obtained as a byproduct of isotropy subgroups related to the G and N–actions on M and the coadjoint actions on $\mathfrak{g}^*$ and $\mathfrak{n}^*$.

We start our study by looking in this setup at the level sets of the G and N–optimal momentum maps. A basic property of the optimal momentum map whose proof can be found in Ortega and Ratiu [2002], establishes the following characterization: let $m \in M$ be such that $\mathcal{J}_G(m) = \rho$, $\mathbf{J}_G(m) = \mu$, and $G_m =: H$. Then, $\mathcal{J}_G^{-1}(\rho)$ equals the unique connected component of the submanifold $\mathbf{J}_G^{-1}(\mu) \cap M_H$ that contains it. Analogously, if $\mathcal{J}_N(m) = \nu$, $\mathbf{J}_N(m) = \eta$, and $N_m = H \cap N$, then $\mathcal{J}_N^{-1}(\nu)$ equals the unique connected component of the submanifold $\mathbf{J}_N^{-1}(\eta) \cap M_{H\cap N}$ that contains it. Recall that

the symbol M_H denotes the ***isotropy type submanifold*** associated to the isotropy subgroup H and that it is defined by $M_H := \{z \in M \mid G_z = H\}$. All along this section we will assume the following

Connectedness hypothesis. *The two submanifolds* $\mathbf{J}_G^{-1}(\mu) \cap M_H$ *and* $\mathbf{J}_N^{-1}(\eta) \cap M_{H \cap N}$ *are connected.*

While this hypothesis is not entirely satisfactory, it will make the presentation that follows much more clear and accessible. The reduction by stages problem does not differ much, qualitatively speaking, no matter if we assume the connectedness hypothesis or not, however the necessary additions in the notation to accommodate the most general case would make the following pages very difficult to read. In order to adapt to the general situation our results, the reader should just take the relevant connected components of $\mathbf{J}_G^{-1}(\mu) \cap M_H$ and $\mathbf{J}_N^{-1}(\eta) \cap M_{H \cap N}$, and each time that we quotient them by a group that leaves them invariant, the reader should take the closed subgroup that leaves invariant the connected component that he has previously chosen. The notation becomes immediately rather convoluted but the ideas involved in the process are the same.

We continue our characterization of the ingredients for reduction by stages in the following proposition.

15.5.1 Proposition. *Let (M, ω) be a symplectic manifold acted properly and canonically upon by a Lie group G and suppose that this action has an associated standard equivariant momentum map $\mathbf{J}_G : M \to \mathfrak{g}^*$. Let $N \subset G$ be a closed normal subgroup of G. Then, if $m \in M$ is such that $\mathcal{J}_G(m) = \rho$, $\mathbf{J}_G(m) = \mu$, and its isotropy subgroup G_m equals $G_m =: H \subset G$ we obtain*

(i) $\mathcal{J}_G^{-1}(\rho) = \mathbf{J}_G^{-1}(\mu) \cap M_H$.

(ii) $\mathcal{J}_N(m) = \pi_N(\rho) =: \nu$, $\mathbf{J}_N(m) = i^*\mu =: \eta$, *and*

$$\mathcal{J}_N^{-1}(\nu) = \mathbf{J}_N^{-1}(\eta) \cap M_{N_\eta \cap H}.$$

(iii) $G_\rho = N_{G_\mu}(H)$, $N_\nu = N_{N_\eta}(N_\eta \cap H)$, *and* $G_\nu = N_{G_\eta}(N_\eta \cap H)$. *The symbol* $N_{G_\mu}(H) := N(H) \cap G_\mu$ *where $N(H)$ denotes the normalizer of H in G. We will refer to $N_{G_\mu}(H)$ as the normalizer of H in G_μ.*

Proof. The proof of the equalities $\mathcal{J}_G^{-1}(\rho) = \mathbf{J}_G^{-1}(\mu) \cap M_H$ and $G_\rho = N_{G_\mu}(H)$ can be found in Ortega and Ratiu [2002].

We now show that $\mathcal{J}_N^{-1}(\nu) = \mathbf{J}_N^{-1}(\eta) \cap M_{N_\eta \cap H}$. By the results in the paper that we just cited, it suffices to show that $N_m = N_\eta \cap H$. Indeed, as the G–equivariance of $\mathbf{J}_N$ implies that $H = G_m \subset G_\eta$, we get $N_m = H \cap N = H \cap G_\eta \cap N = N_\eta \cap H$. Consequently, the same result that gave us $G_\rho = N_{G_\mu}(H)$, can be applied to the N–action on M to obtain $N_\nu = N_{N_\eta}(N_\eta \cap H)$.

Finally, we prove the identity $G_\nu = N_{G_\eta}(N_\eta \cap H)$ by double inclusion. Let first $g \in G_\nu$. The equality $g \cdot \nu = \nu$ implies that $g \cdot m = \mathcal{F}_T(m)$, with

$\mathcal{F}_T \in G_{\mathcal{A}'_N}$. For simplicity suppose that $\mathcal{F}_T = F_t$, with F_t the Hamiltonian flow associated to a N–invariant function on M. The standard Noether's Theorem implies that $g \cdot m = F_t(m) \in \mathbf{J}_N^{-1}(\eta)$ and therefore $g \in G_\eta$. Also, as the flow F_t is N–equivariant we have

$$N_\eta \cap H = N_m = N_{F_t(m)} = N_{g \cdot m} = g N_m g^{-1} = g(N_\eta \cap H)g^{-1},$$

and consequently $g \in N_{G_\eta}(N_\eta \cap H)$. The reverse inclusion is trivial.   ∎

Remark. A major consequence of the previous proposition is the fact that the subgroups G_ν and N_ν, and those that will derive from them, are automatically closed subgroups. This circumstance implies that the proper action hypothesis given in Definition 15.3.1 and necessary for reduction by stages is automatically satisfied in this setup.

The previous proposition allows us to explicitly write down in our setup the one–shot reduced space:

$$M_\rho := \mathcal{J}_G^{-1}(\rho)/G_\rho = \mathbf{J}_G^{-1}(\mu) \cap M_H/N_{G_\mu}(H), \qquad (15.5.1)$$

as well as the first stage reduced space:

$$M_\nu := \mathcal{J}_N^{-1}(\nu)/N_\nu = \mathbf{J}_N^{-1}(\eta) \cap M_{N_\eta \cap H}/N_{N_\eta}(N_\eta \cap H).$$

We now proceed with the construction of the second stage reduced space. As it was already the case in the general optimal setup, the quotient group

$$\mathcal{H}_\nu := G_\nu/N_\nu = \frac{N_{G_\eta}(N_\eta \cap H)}{N_{N_\eta}(N_\eta \cap H)}$$

acts canonically on the quotient M_ν with associated optimal momentum map $\mathcal{J}_{\mathcal{H}_\nu} : M_\nu \to M_\nu/A'_{\mathcal{H}_\nu}$. In this setup we can say more. Indeed, in this case the $\mathcal{H}_\nu$–action on M_ν is automatically proper and has an associated standard momentum map $\mathbf{J}_{\mathcal{H}_\nu} : M_\nu \to \mathrm{Lie}(\mathcal{H}_\nu)^*$, where the symbol $\mathrm{Lie}(\mathcal{H}_\nu)$ denotes the Lie algebra of the group $\mathcal{H}_\nu$. An explicit expression for $\mathbf{J}_{\mathcal{H}_\nu}$ can be obtained by mimicking the computations made in Part II for the free case. In order to write it down we introduce the following maps: let $\pi_{G_\nu} : G_\nu \to G_\nu/N_\nu$ be the projection, $r_\nu = T_e \pi_{G_\nu} : \mathfrak{g}_\nu \to \mathrm{Lie}(\mathcal{H}_\nu) \simeq \mathfrak{g}_\nu/\mathfrak{n}_\nu$ be its derivative at the identity, and $r_\nu^* : \mathrm{Lie}(\mathcal{H}_\nu)^* \to \mathfrak{g}_\nu^*$ be the corresponding dual map. Then, for any $\pi_\nu(z) \in M_\nu$ and any $r_\nu(\xi) \in \mathrm{Lie}(\mathcal{H}_\nu)$, the momentum map $\mathbf{J}_{\mathcal{H}_\nu}$ is given by the expression

$$\langle \mathbf{J}_{\mathcal{H}_\nu}(\pi_\nu(z)), r_\nu(\xi) \rangle = \langle \mathbf{J}_G(z), \xi \rangle - \langle \bar{\eta}, \xi \rangle, \qquad (15.5.2)$$

where $\bar{\eta} \in \mathfrak{g}_\nu^*$ is some chosen extension of the restriction $\eta|_{\mathfrak{n}_\nu}$ to a linear functional on $\mathfrak{g}_\nu$. This momentum map is not equivariant. Indeed, its non equivariance cocycle $\bar{\omega}$ is given by the expression

$$r_\nu^*(\bar{\omega}(\pi_{G_\nu}(h))) = \mathrm{Ad}_{h^{-1}}^* \bar{\eta} - \bar{\eta},$$

for any $\pi_{G_\nu}(h) \in G_\nu/N_\nu$. The map $\mathbf{J}_{\mathcal{H}_\nu}$ becomes equivariant if we replace the coadjoint action of $\mathcal{H}_\nu$ on the dual of its Lie algebra by the affine action defined by

$$\pi_{G_\nu}(h) \cdot \lambda := \mathrm{Ad}^*_{(\pi_{G_\nu}(h))^{-1}} \lambda + \bar{\omega}(\pi_{G_\nu}(h)), \qquad (15.5.3)$$

for any $\pi_{G_\nu}(h) \in \mathcal{H}_\nu$ and any $\lambda \in \mathrm{Lie}(\mathcal{H}_\nu)^*$. Let now $\tau \in \mathrm{Lie}(\mathcal{H}_\nu)^*$ be the element defined by

$$\langle \tau, r_\nu(\xi) \rangle = \langle \mu, \xi \rangle - \langle \bar{\nu}, \xi \rangle, \qquad (15.5.4)$$

for any $r_\nu(\xi) \in \mathrm{Lie}(\mathcal{H}_\nu)$. A calculation following the lines of Part II shows that

$$(G_\nu)_{\mu|_{\mathfrak{g}_\nu}} = \left(N_{G_\eta}(N_\eta \cap H)\right)_{\mu|_{\mathrm{Lie}\left(N_{G_\eta}(N_\eta \cap H)\right)}}, \qquad (15.5.5)$$

and hence the isotropy subgroup $(\mathcal{H}_\nu)_\tau$ of τ with respect to the affine action (15.5.3) of $\mathcal{H}_\nu$ on the dual of its Lie algebra, is given by

$$(\mathcal{H}_\nu)_\tau = \pi_{G_\nu}\left((G_\nu)_{\mu|_{\mathfrak{g}_\nu}}\right)$$

$$= \frac{\left(\left(N_{G_\eta}(N_\eta \cap H)\right)_{\mu|_{\mathrm{Lie}\left(N_{G_\eta}(N_\eta \cap H)\right)}}\right) N_{N_\eta}(N_\eta \cap H)}{N_{N_\eta}(N_\eta \cap H)} \qquad (15.5.6)$$

Now, for any $m \in \mathcal{J}_G^{-1}(\rho)$, the choice of $\tau \in \mathrm{Lie}(\mathcal{H}_\nu)^*$ in (15.5.4) guarantees that $\mathbf{J}_{\mathcal{H}_\nu}(\pi_\nu(m)) = \tau$ and, moreover, if $\mathcal{J}_{\mathcal{H}_\nu}(\pi_\nu(m)) = \sigma \in M/A'_{\mathcal{H}_\nu}$ then,

$$\mathcal{J}_{\mathcal{H}_\nu}^{-1}(\sigma) = \mathbf{J}_{\mathcal{H}_\nu}^{-1}(\tau) \cap (M_\nu)_{(\mathcal{H}_\nu)_{\pi_\nu(m)}} \qquad (15.5.7)$$

since, by extension of the connectedness hypothesis we will suppose that $\mathbf{J}_{\mathcal{H}_\nu}^{-1}(\tau) \cap (M_\nu)_{(\mathcal{H}_\nu)_{\pi_\nu(m)}}$ is also connected.

We compute the isotropy subgroup $(\mathcal{H}_\nu)_{\pi_\nu(m)}$ in terms of the groups that already appeared in our study. Indeed, we will now show that

$$(\mathcal{H}_\nu)_{\pi_\nu(m)} = \frac{N_{N_\eta}(H \cap N_\eta)H}{N_{N_\eta}(N_\eta \cap H)}. \qquad (15.5.8)$$

Take first an element $\pi_{G_\nu}(g) \in \mathcal{H}_\nu$ such that $\pi_{G_\nu}(g) \cdot \pi_\nu(m) = \pi_\nu(m)$ or, equivalently, $\pi_\nu(g \cdot m) = \pi_\nu(m)$. Hence, there exists a group element $n \in N_\nu = N_{N_\eta}(N_\eta \cap H)$ such that $g \cdot m = n \cdot m$. Given that $G_m = H$, we see that $n^{-1} \cdot g \in H$, necessarily and hence $g \in N_{N_\eta}(N_\eta \cap H)H$ and

$$\pi_{G_\nu}(g) \in N_{N_\eta}(N_\eta \cap H)H/N_{N_\eta}(N_\eta \cap H).$$

Conversely, if

$$\pi_{G_\nu}(g) \in N_{N_\eta}(N_\eta \cap H)H/N_{N_\eta}(N_\eta \cap H),$$

we can write $g = nh$, with $n \in N_{N_\eta}(N_\eta \cap H)$ and $h \in H$ and therefore

$$\pi_{G_\nu}(g) \cdot \pi_\nu(m) = \pi_\nu(nh \cdot m) = \pi_\nu(n \cdot m) = \pi_\nu(m),$$

as required.

In order to write down the second stage reduced space we have to compute the isotropy subgroup $(\mathcal{H}_\nu)_\sigma$. In view of (15.5.6) and (15.5.8), and Proposition 15.5.1 adapted to the optimal momentum map $\mathcal{J}_{\mathcal{H}_\nu}$, we have

$$(\mathcal{H}_\nu)_\sigma = N_{(\mathcal{H}_\nu)_\tau}\left(\frac{N_{N_\eta}(H \cap N_\eta)H}{N_{N_\eta}(N_\eta \cap H)} \right), \qquad (15.5.9)$$

where the group $(\mathcal{H}_\nu)_\tau$ is given by the expression (15.5.6). We now recall a standard result about normalizers that says that if $A \subset B \subset C$ are groups such that A is normal in both B and C, then

$$N_{C/A}(B/A) = N_C(B)/A.$$

If we apply this equality to Expression (15.5.9) we obtain

$$(\mathcal{H}_\nu)_\sigma = \frac{N_{(G_\nu)_{\mu|_{\mathfrak{g}_\nu}} N_{N_\eta}(N_\eta \cap H)}(N_{N_\eta}(H \cap N_\eta)H)}{N_{N_\eta}(N_\eta \cap H)}$$

$$= \frac{N_{\left(\left(N_{G_\eta}(N_\eta \cap H)\right)_{\mu|_{\text{Lie}\left(N_{G_\eta}(N_\eta \cap H)\right)}} N_{N_\eta}(N_\eta \cap H)\right)}(N_{N_\eta}(H \cap N_\eta)H)}{N_{N_\eta}(N_\eta \cap H)}$$

$$(15.5.10)$$

All the computations that we just carried out allow us to explicitly write down the second stage reduced space. Namely, by combination of expressions (15.5.7), (15.5.8), and (15.5.10), we arrive at

$$(M_\nu)_\sigma = \mathcal{J}_{\mathcal{H}_\nu}^{-1}(\sigma)/(\mathcal{H}_\nu)_\sigma = \frac{\mathbf{J}_{\mathcal{H}_\nu}^{-1}(\tau) \cap (M_\nu)_{\frac{N_{N_\eta}(H\cap N_\eta)H}{N_{N_\eta}(N_\eta\cap H)}}}{\frac{N_{(G_\nu)_{\mu|_{\mathfrak{g}_\nu}} N_{N_\eta}(N_\eta\cap H)}(N_{N_\eta}(H\cap N_\eta)H)}{N_{N_\eta}(N_\eta\cap H)}}, \quad (15.5.11)$$

where the group $(G_\nu)_{\mu|_{\mathfrak{g}_\nu}}$ is given by (15.5.5).

The Optimal Reduction by Stages Theorem 15.3.2 guarantees that the second stage reduced space (15.5.11) is symplectomorphic to the one–shot reduced space (15.5.1) in the presence of the stages hypothesis introduced in Definition 15.3.1. In this setup, that hypothesis can be completely reformulated in terms of relations between Lie algebraic elements and isotropy subgroups. More specifically, in the globally Hamiltonian framework, the stages hypothesis is equivalent to the following condition:

Hamiltonian Stages Hypothesis: Let $\mu \in \mathfrak{g}^*$ and $H \subset G$. We say that the pair (μ, H) satisfies the Hamiltonian stages hypothesis whenever for any other similar pair (μ', H') such that

$$\begin{cases} i^*\mu = i^*\mu' =: \eta \in \mathfrak{n}^* \\ N_\eta \cap H = N_\eta \cap H' =: K \end{cases}$$

and

$$
\begin{cases}
\mu|_{\mathrm{Lie}\left(N_{G_\eta}(K)\right)} & = & \mu'|_{\mathrm{Lie}\left(N_{G_\eta}(K)\right)} \\
& =: & \zeta \in \mathrm{Lie}\left(N_{G_\eta}(K)\right)^* \\
N_{N_\eta}(K)H = N_{N_\eta}(K)H' =: L,
\end{cases}
$$

there exists an element $n \in N_{(N_{G_\eta}(K))_\zeta}(L)N_{N_\eta}(K)$ such that

$$
\mu' = \mathrm{Ad}^*_{n^{-1}}\mu \quad \text{and} \quad H' = nHn^{-1}.
$$

A quick inspection shows that when the G–action is free, that is, when all the isotropy subgroups $H = \{e\}$, the previous condition collapses into the stages hypothesis introduced in Part II.

Recall that in the same fashion in which the proper action hypothesis introduced in Definition 15.3.1 is automatically satisfied in this setup, so is the Lindelöf hypothesis on the one–shot reduced space M_ρ if we just assume that M is Lindelöf. This is so because closed subsets and continuous images of Lindelöf spaces are always Lindelöf.

The Optimal Reduction by Stages Theorem together with the ideas that we just introduced implies in this setup the following highly non trivial symplectomorphism that we enunciate in the form of a theorem. The following statement is consistent with the previously introduced notations.

15.5.2 Theorem (Hamiltonian Reduction by Stages).
Hamiltonian reduction Let (M, ω) be a symplectic manifold acted properly and canonically upon by a Lie group G that has a closed normal subgroup N. Suppose that this action has an associated equivariant momentum map $\mathbf{J}_G : M \to \mathfrak{g}^*$. Let $\mu \in \mathfrak{g}^*$ be a value of $\mathbf{J}_G$ and H an isotropy subgroup of the G action on M. If the manifold M is Lindelöf and the pair (μ, H) satisfies the Hamiltonian stages hypothesis, then the symplectic reduced spaces

$$
\frac{\mathbf{J}_G^{-1}(\mu) \cap M_H}{N_{G_\mu}(H)} \quad \text{and} \quad \frac{\dfrac{\mathbf{J}_{\mathcal{H}_\nu}^{-1}(\tau) \cap (M_\nu)_{\frac{N_{N_\eta}(H \cap N_\eta)H}{N_{N_\eta}(N_\eta \cap H)}}}{\dfrac{N_{(G_\nu)_{\mu|_{\mathfrak{g}_\nu}}} N_{N_\eta}(N_\eta \cap H)(N_{N_\eta}(H \cap N_\eta)H)}{N_{N_\eta}(N_\eta \cap H)}}}
$$

are symplectomorphic. In this expression $\eta = i^*\mu$,

$$
M_\nu = \frac{\mathbf{J}_N^{-1}(\eta) \cap M_{N_\eta \cap H}}{N_{N_\eta}(N_\eta \cap H)}, \quad \mathcal{H}_\nu = \frac{N_{G_\eta}(N_\eta \cap H)}{N_{N_\eta}(N_\eta \cap H)},
$$

$$
(G_\nu)_{\mu|_{\mathfrak{g}_\nu}} = \left(N_{G_\eta}(N_\eta \cap H)\right)_{\mu|_{\mathrm{Lie}\left(N_{G_\eta}(N_\eta \cap H)\right)}},
$$

$\mathbf{J}_{\mathcal{H}_\nu} : M_\nu \to \mathrm{Lie}(\mathcal{H}_\nu)^*$ *is the momentum map associated to the $\mathcal{H}_\nu$– action on M_ν defined in (15.5.2), and $\tau \in \mathrm{Lie}(\mathcal{H}_\nu)^*$ the element defined in (15.5.4).*

When the G–action is free, the previous theorem coincides with the Reduction by Stages Theorem 5.2.9 presented in Part II.

A special but very important particular case of Theorem 15.5.2 takes place when the group G is discrete ($\mathfrak{g} = \{0\}$). In that situation, all the standard momentum maps in the construction vanish and the theorem gives us a highly non trivial relation between quotients of isotropy type submanifolds. We start by reformulating the Hamiltonian stages hypothesis in this particular case.

Discrete Reduction by Stages Hypothesis.   *Let G be a discrete group, N a normal subgroup, and H a subgroup. We say that H satisfies the Discrete Reduction by stages hypothesis with respect to N if for any other subgroup H' such that*

$$N \cap H = N \cap H' =: K \quad and \quad N_N(K)H = N_N(K)H' =: L,$$

there exists an element $n \in N_{N_G(K)}(L)N_N(K)$ such that $H' = nHn^{-1}$.

15.5.3 Theorem (Discrete Reduction by Stages).   *Let (M, ω) be a symplectic manifold acted properly and canonically upon by a discrete Lie group G that has a closed normal subgroup N. Let H be an isotropy subgroup of the G action on M and $K := N \cap H$. If the manifold M is Lindelöf and H satisfies the Discrete Reduction by stages hypothesis with respect to N, then the symplectic reduced spaces*

$$\frac{M_H}{N_G(H)} \quad and \quad \frac{\left(\frac{M_K}{N_N(K)}\right)_{\frac{N_N(K)H}{N_N(K)}}}{\frac{N_{N_G(K)}(N_N(K)H)}{N_N(K)}} \tag{15.5.12}$$

are symplectomorphic.

When the G–action on M is free, discrete reduction by stages hypothesis is trivially satisfied and Theorem 15.5.3 produces the straightforward symplectomorphism

$$M/G \simeq (M/N)/(G/N).$$

Hence, it is in the presence of singularities that the relation stablished in (15.5.12) is really visible and non trivial.

Bibliography

Abarbanel, H.D.I. and D.D. Holm [1987], Nonlinear stability analysis of inviscid flows in three dimensions: incompressible fluids and barotropic fluids, *Phys. Fluids*, **30**, 3369–3382.

Abarbanel, H.D.I., D.D. Holm, J.E. Marsden, and T.S. Ratiu [1986], Nonlinear stability analysis of stratified fluid equilibria, *Phil. Trans. Roy. Soc. London A*, **318**, 349–409.

Abraham, R. and J.E. Marsden [1978], *Foundations of Mechanics*. Benjamin-Cummings Publ. Co, Updated 1985 version, reprinted by Perseus Publishing, second edition. Available online at `http://www.cds.caltech.edu/~marsden/books/Foundations_of_Mechanics.html`.

Abraham, R., J.E. Marsden, and T.S. Ratiu [1988], *Manifolds, Tensor Analysis, and Applications*. Applied Mathematical Sciences, volume 75. Springer-Verlag.

Adams, J.F. [1969], *Lectures on Lie Groups*, The University of Chicago Press.

Ahlfors, L. [1961], Some remarks on Teichmüller's space of Riemann surfaces, *Ann. of Math.*, **74**, 171–191.

Ahlfors, L. [1962], Curvature properties of Teichmüller's space, *J. Analyse Math.*, **9**, 161–176.

Alber, M.S., G.G. Luther, J.E. Marsden, and J.M. Robbins [1998], Geometric phases, reduction and Lie-Poisson structure for the resonant three-wave interaction, *Physica D*, **123**, 271–290.

Alekseev, A., A. Malkin, and E. Meinrenken [1998], Lie group valued momentum maps. *J. Differential Geom.*, **48**, 445–495.

Arms, J. [1979], Linearization stability of gravitational and gauge fields, *J. Math. Phys.*, **20**, 443–453.

Arms, J.M., R.H. Cushman, and M.J. Gotay [1991], A universal reduction procedure for Hamiltonian group actions, In *The Geometry of Hamiltonian Systems*, T.S. Ratiu ed., 33–51, Springer Verlag.

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

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

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

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

Arnold, V.I. [1969], On an a priori estimate in the theory of hydrodynamical stability, *Amer. Math. Soc. Transl.*, **79**, 267–269.

Arnold, V.I., V.V. Kozlov, and A.I. Neishtadt [1988], *Mathematical Aspects of Classical and Celestial Mechanics*, in *Dynamical Systems III*, V.I. Arnold, ed., Springer-Verlag, third edition, 2006.

Arnold, V.I. [1989], *Mathematical Methods of Classical Mechanics*, Graduate Texts in Math., volume 60, Springer-Verlag, Second edition.

Atiyah, M. [1982], Convexity and commuting Hamiltonians, *Bull. London Math. Soc.*, **14**, 1–5.

Atiyah, M.F. and R. Bott [1982], The Yang-Mills equations over Riemann surfaces, *Phil. Trans. R. Soc. Lond. A*, **308**, 523–615.

Atiyah, M. and R. Bott [1984], The moment map and equivariant cohomology, *Topology*, **23**, 1–28.

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

Balog, J., L. Fehér and L. Palla [1998], Coadjoint orbits of the Virasoro algebra and the global Liouville equation, *Int. J. Mod. Phys.*, **A 13**, 315–362.

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

Bates, L. and J. Śniatycki [1993], Nonholonomic reduction, *Reports on Math. Phys.*, **32**, 99–115.

Bayen, F., M. Flato, C. Fronsdal, A. Lichnerowicz, and D. Sternheimer [1978], Deformation theory and quantization. I. Deformations of symplectic structures, *Ann. Physics*, **111**, 61–110; part II on pages 111–151.

Benenti, S. and W. Tulczyjew [1985], Momentum relations for Hamiltonian group actions, *Géométrie symplectique et mécanique*, 12–29.

Bers, L. [1965], Automorphic forms and general Teichmüller spaces, *Proc. Conf. Complex Analysis (Minneapolis 1964)*, Springer, Berlin 1965, 109–113.

Bers, L. [1970], Universal Teichmüller space, in *Analytic methods in mathematical physics (Sympos., Indiana Univ., Bloomington, Ind., 1968)*, 65–83. Gordon and Breach, New York.

Bhaskara, K.H. [1990], Affine Poisson structures, *Proc. Indian Academy of Sci.*, **100**, 189–202.

Binz, E., J. Śniatycki, and H. Fischer [1988], *Geometry of Classical Fields*, volume 154 of *North-Holland Mathematics Studies*, North-Holland Publishing Co., Amsterdam.

Birtea, P., M. Puta, T. S. Ratiu, and R. Tudoran [2005], Symmetry breaking for toral actions in simple mechanical systems, *J. Diff. Eqns.*, **216**, 282–323.

Blankenstein, G. and A. J. van der Schaft [2001], Symmetry and reduction in implicit generalized Hamiltonian systems, *Rep. Math. Phys.* **47**, 57–100.

Blaom, A.D. [2000], Reconstruction phases via Poisson reduction, *Diff. Geom. and Appl.*, **12**, 231–252.

Blaom, A.D. [2001], A geometric setting for Hamiltonian perturbation theory, *Mem. Amer. Math. Soc.*, volume 153, number 727. Providence, R.I.

Bleecker, D. [1981], *Gauge Theory and Variational Principles*, *Global Analysis Pure and Applied Series A*, volume **1**, Addison-Wesley Publishing Co., Reading, Mass.

Bloch, A. (with the collaboration of J. Ballieul, P. Crouch, and J.E. Marsden) [2003], *Nonholonomic Mechanics and Control*, Interdisciplinary Applied Mathematics, volume 24. Springer-Verlag.

Bloch, A.M. and P.E. Crouch [1994], Reduction of Euler–Lagrange problems for constrained variational problems and relation with optimal control problems, *Proc. CDC*, **33**, 2584–2590. IEEE.

Bloch, A.M., P.E. Crouch, J.E. Marsden, and T.S. Ratiu [2000], An almost Poisson structure for the generalized rigid body equations, *IFAC Proceedings*, 0-08-043658-7, March 2000.

Bloch, A.M., P. Crouch, J.E. Marsden, and T.S. Ratiu [2002], The symmetric representation of the rigid body equations and their discretization, *Nonlinearity*, **15**, 1309–1341.

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

Bloch, A.M., P.S. Krishnaprasad, J.E. Marsden, and T.S. Ratiu [1994], Dissipation induced instabilities, *Ann. Inst. H. Poincaré, Analyse Nonlinéaire*, **11**, 37–90.

Bloch, A.M., P.S. Krishnaprasad, J.E. Marsden and T.S. Ratiu [1996], The Euler–Poincaré equations and double bracket dissipation, *Comm. Math. Phys.*, **175**, 1–42.

Bloch, A.M. and J.E. Marsden [1990], Stabilization of rigid body dynamics by the energy–Casimir method, *Systems Control Lett.*, **14**, 341–346.

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

Bobenko, A. and Y. Suris [1999], Discrete Lagrangian reduction, discrete Euler-Poincaré equations, and semidirect products, *Letters in Mathematical Physics*, **49**, 79–93.

Born, M. [1927], *The Mechanics of the Atom.* G. Bells and Sons, Ltd., London.

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

Bourbaki, N. [1971], *Lie Groups and Lie Algebras. Chapters 1–3*, Springer-Verlag, Berlin, 1998. Translated from the 1971 French edition.

Bredon, G.E. [1972] *Introduction to Compact Transformation Groups.* Academic Press.

Bretherton, F.P. [1970], A note on Hamilton's principle for perfect fluids, *J. Fluid Mech.*, **44**, 19–31.

Brouwer, D. and G.M. Clemence [1961], *Methods of Celestial Mechanics*, Academic Press, New York.

Camassa, R. and D.D. Holm [1993], An integrable shallow water equation with peaked solitons, *Phys. Rev. Lett.*, **71**, 1661–1664.

Cannas da Silva, A. and A. Weinstein [1999], *Geometric Models for Noncommutative Algebras*, Berkeley Math. Lecture Notes, volume 10. Amer. Math. Soc.

Cartan, É. [1922], *Leçons sur les Invariants Intégraux,* Hermann.

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

Castrillón López, M., P.L. Garcia Pérez, and T.S. Ratiu [2001], Euler-Poincaré reduction on principal bundles, *Lett. Math. Phys.*, **58**, 167–180.

Castrillón López, M. and J.E. Marsden [2003], Some remarks on Lagrangian and Poisson reduction for field theories, *J. Geom. Phys.*, **48**(1), 52–83.

Castrillón López, M. and T.S. Ratiu [2003], Reduction in principal bundles: covariant Lagrange-Poincaré equations, *Comm. Math. Phys.* **236**(2), 223–250.

Castrillón López, M., T.S. Ratiu, and S. Shkoller [2000], Reduction in principal fiber bundles: Covariant Euler-Poincaré equations, *Proc. Amer. Math. Soc.*, **128**, 2155–2164.

Cendra, H., D.D. Holm, M.J.W. Hoyle, and J.E. Marsden [1998], The Maxwell–Vlasov equations in Euler–Poincaré form, *J. Math. Phys.*, **39**, 3138–3157.

Cendra, H., D.D. Holm, J.E. Marsden, and T. S. Ratiu [1998], Lagrangian reduction, the Euler–Poincaré equations and semidirect products, *Amer. Math. Soc. Transl.*, **186**, 1–25.

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

Cendra, H., J.E. Marsden, S. Pekarsky, and T.S. Ratiu [2003], Variational principles for Lie-Poisson and Hamilton-Poincaré equations, *Moscow Mathematics Journal*, **3**, 833–867.

Cendra, H., J.E. Marsden, and T.S. Ratiu [2001a], *Lagrangian Reduction by Stages*, Memoirs of the American Mathematical Society, volume 152, number 722. Providence, R.I.

Cendra, H., J.E. Marsden, and T.S. Ratiu [2001b], Geometric mechanics, Lagrangian reduction and nonholonomic systems. In Engquist, B. and W. Schmid, editors, *Mathematics Unlimited-2001 and Beyond*, pages 221–273. Springer-Verlag, New York.

Cendra, H., J.E. Marsden, and T.S. Ratiu [2003], Cocycles, compatibility, and Poisson brackets for complex fluids. In Capriz, G. and P. Mariano, editors, *Advances in Multifield Theories with Substructures*, Birkhäuser, Boston.

Chang, D., A.M. Bloch, N. Leonard, J.E. Marsden, and C. Woolsey [2002], The equivalence of controlled Lagrangian and controlled Hamiltonian systems, *Control and the Calculus of Variations (special issue dedicated to J.L. Lions)*, **8**, 393–422.

Chang, D. E., D. F. Chichka, and J.E. Marsden [2002], Lyapunov-based transfer between elliptic Keplerian orbits, *Discrete and Continuous Dynamical Systems: Series B*, **2**, 57–67.

Charlier, C. L. [1927], *Die Mechanik des Himmels. Bd.I, II*, 2nd Edition, Walter de Gruyter, Berlin.

Chern, S.J. and J.E. Marsden [1990], A note on symmetry and stability for fluid flows, *Geo. Astro. Fluid. Dyn.*, **51**, 1–4.

Chernoff, P.R. and J.E. Marsden [1974], *Properties of Infinite Dimensional Hamiltonian systems*, Lecture Notes in Math., **425**. Springer, New York.

Chetayev, N.G. [1941], On the equations of Poincaré, *J. Appl. Math. Mech.*, **5**, 253–262.

Chobotov, V.A., ed. [1996], *Orbital Mechanics,* 2nd ed., in the AIAA Education Series, American Institute of Aeronautics and Astronautics, Inc., New York.

Chossat, P., D. Lewis, J.-P. Ortega, and T.S. Ratiu [2003], Bifurcation of relative equilibria in mechanical systems with symmetry, *Adv. in Appl. Math.*, **31**, 10–45.

Chossat, P., J.-P. Ortega, and T.S. Ratiu [2002], Hamiltonian Hopf bifurcation with symmetry, *Arch. Rational Mech. Anal.*, **163**, 1–33; **167**, 83–84.

Coffey, S.L., A. Depit, and B.R. Miller [1986], The Critical Inclination in Artificial Satellite Theory, *Celestial Mechanics*, **39**, 365 – 406.

Condevaux, M. P., P. Dazord, and P. Molino [1988], Géométrie du moment, *Travaux du Séminaire Sud-Rhodanien de Géométrie, I*, Publ. Dép. Math. Nouvelle Sér. B, 88-1, Univ. Claude-Bernard, Lyon, 131–160.

Cortés, J., M. de León, J. C. Marrero, D. Martín de Diego, and E. Martínez [2006], A survey of Lagrangian mechanics and control on Lie algebroids and groupoids, *Int. J. Geom. Methods Mod. Phys.*, **3**, 509–558.

Cushman, R.H. [1991], A Survey of Normalization Techniques applied to Perturbed Keplerian Systems, in K. Jones et al. *Dynamics Reported*, **1**, Springer-Verlag, New York.

Cushman, R.H. and L.M. Bates [1997], *Global Aspects of Classical Integrable Systems*, Birkhhäuser, Boston.

Cushman, R.H. and D. Rod [1982], Reduction of the semi-simple 1:1 resonance, *Physica D*, **6**, 105–112.

Cushman, R.H., D. Kemppainen, and J. Śniatycki [1998], A classical particle with spin realized by reduction of a nonlinear nonholonomic constraint, *Rep. Math. Phys.*, **41**, 133–142.

Cushman, R.H. and J. Śniatycki [1999], Hamiltonian mechanics on principal bundles, *C. R. Math. Acad. Sci. Soc. R. Can.*, **22**, 60–64.

Cushman, R.H. and Śniatycki, J. [2001], Differential structure of orbit spaces, *Canad. J. Math.*, **53**, 715–755, erratum **55**, (2003), 247.

Cushman, R.H. and J. Śniatycki [2002], Nonholonomic reduction for free and proper actions, *Regul. Chaotic Dyn.*, **7**, 61–72.

Dai, J. [2000], *Conjugacy Classes, Characters and Coadjoint Orbits of* $Diff^+(S^1)$, Ph.D. Thesis, University of Arizona.

Dai, J. and D. Pickrell [2003], The orbit method and the Virasoro extension of $Diff^+(S^1)$, I. Orbital integrals *J. Geom. and Phys.*, **44**, 623–653.

Dazord, P. [1985], Feuilletages à singularités, *Nederl. Akad. Wetensch. Indag. Math.*, **47**, 21–39.

de Azcárraga, J.A. and J.M. Izquierdo [1995], *Lie Groups, Lie Algebras, Cohomology and Some Applications in Physics*, Cambridge Monographs on Mathematical Physics, Cambridge University Press.

Delaunay, C. [1860], *Théorie du Mouvement de la Lune*, Mem. **28** (1860); **29** (1867), *Acad. Sci. France*, Paris.

Deprit, A. [1983], Elimination of the nodes in problems of N bodies, *Celestial Mech.*, **30**, 181–195.

Derks, G., D.K. Lewis, and T.S. Ratiu [1995], Approximations with curves of relative equilibria in Hamiltonian systems with dissipation, *Nonlinearity*, **8**, 1087–1113.

Dirac, P.A.M. [1930], *The Principles of Quantum Mechanics.* Oxford University Press.

Dirac, P.A.M. [1950], Generalized Hamiltonian mechanics, *Canad. J. Math.*, **2**, 129–148.

Dirac, P.A.M. [1964], *Lectures on Quantum Mechanics*, Belfer Graduate School of Science, Monograph Series, volume 2. Yeshiva University.

Duistermaat, J.J. [1980], On global action angle coordinates, *Comm. Pure Appl. Math.*, **33**, 687–706.

Duistermaat, J.J. and J.A.C. Kolk [1999], *Lie Groups.* Springer Verlag.

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

Duval, C., J. Elhadad, M.J. Gotay, and G.M. Tuynman [1992], Pukanszky's condition and symplectic induction, *J. Diff. Geom.* **36**, 331–348.

Eardley, D.M. and V. Moncrief [1982], The global existence of Yang-Mills-Higgs fields in 4-dimensional Minkowski space, *Comm. Math. Phys.*, **83**, 171–212.

Ebin, D.G. and J.E. Marsden [1970], Groups of diffeomorphisms and the motion of an incompressible fluid, *Ann. of Math.*, **92**, 102–163.

Emmrich, C. and H. Römer [1990], Orbifolds as configuration spaces of systems with g symmetries, *Comm. Math. Phys.*, **129**, 69–94.

Feynman, R.P. and A.R. Hibbs [1965], *Quantum Mechanics and Path Integrals*, McGraw-Hill.

Fischer, A. E., J. E. Marsden, and V. Moncrief [1980], The structure of the space of solutions of Einstein's equations, I: One Killing field, *Ann. Ins. H. Poincaré*, **33**, 147–194.

Flato, M., A. Lichnerowicz, and D. Sternheimer [1976], Deformations of Poisson brackets, Dirac brackets and applications, *J. Mathematical Phys.*, **17**, 1754–1762.

Fomenko, A.T. and V.V. Trofimov [1989], *Integrable Systems on Lie Algebras and Symmetric Spaces*, Gordon and Breach.

Gardiner, F.P. and D.P. Sullivan [1992], Symmetric structures on a closed curve, *Amer. J. Math.*, **114**, 683–736.

Gay-Balmaz, F. [2007], PhD Thesis, École Polytechnique Fédérale de de Lausanne.

Gay-Balmaz, F. and T. S. Ratiu [2006], Group actions on chains of Banach manifolds and application to fluid dynamics, *Ann. Global Analysis Geom.*, to appear.

Gay-Balmaz, F. and T. S. Ratiu [2007], Reduced Lagrangian and Hamiltonian formulations of Euler-Yang-Mills fluids, preprint.

Gay-Balmaz, F., T. S. Ratiu, and J. E. Marsden [2007], The geometry of Teichmüller space and the Euler-Weil-Petersson equations, in preparation.

Gelfand I. and D. Fuchs [1968], Cohomologies of the Lie algebra of vector fields on the circle. *Funkc. Anal. Prilozh.* , **2**, 92–93.

Goldman, W. M. [1984], The symplectic nature of fundamental groups of surfaces. *Adv. in Math.*, **54**, 200–225.

Goldman, W.M. and J.J. Millson [1990], Differential graded Lie algebras and singularities of level sets of momentum mappings, *Comm. Math. Phys.*, **131**, 495–515.

Goldstein, H. [1980], *Classical Mechanics*, 2nd Edition, Addison-Wesley, Reading, Mass.

Golubitsky, M., I. Stewart, and D. Schaeffer [1988], *Singularities and Groups in Bifurcation Theory*. Vol. II, Applied Mathematical Sciences, volume 69, Springer-Verlag, New York.

Gotay, M., J. Isenberg, and J.E. Marsden [1997], Momentum Maps and the Hamiltonian Structure of Classical Relativistic Field Theories I, II, *Available from* http://www.cds.caltech.edu/~marsden/.

Grabsi, F., J. Montaldi, and J.-P. Ortega [2004], Bifurcation and forced symmetry breaking in Hamiltonian systems, *C. R. Acad. Sci. Paris Sér. I Math.*, **338**, 565–570.

Guichardet, A. [1984], On rotation and vibration motions of molecules, *Ann. Inst. H. Poincaré*, **40**, 329–342.

Guieu, L. [2000], Stabilisateurs cycliques pour la reprsentation coadjointe du groupe des diffeomorphismes du cercle, *Bull. Sci. Math.*, **124**, 125–145.

Guieu, L. and C. Roger [2003], *L'Algèbre et le Groupe de Virasoro: Aspects Géométriques et Algébriques, Généralisations*, preprint.

Guillemin, V. and S. Sternberg [1978], On the equations of motions of a classic particle in a Yang–Mills field and the principle of general covariance, *Hadronic J.*, **1**, 1–32.

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

Guillemin, V. and S. Sternberg [1982a], Convexity properties of the moment map, *Invent. Math.*, **67**, 491–513; also **77**, 533–546.

Guillemin, V. and S. Sternberg [1982b], Geometric quantization and multiplicities of group representations, *Invent. Math.*, **67**, 515–538.

Guillemin, V. and S. Sternberg [1983], The Frobenius reciprocity theorem from a symplectic point of view, *Nonlinear Partial Differential Operators and Quantization Procedures (Clausthal, 1981)*, 242–256, Lecture Notes in Math., volume 1037, Springer-Verlag.

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

Hamel, G. [1904], Die Lagrange–Eulerschen Gleichungen der Mechanik, *Z. für Mathematik u. Physik*, **50**, 1–57.

Hamel, G. [1949], *Theoretische Mechanik*, Springer-Verlag, Berlin.

Hernandez, A. and J. E. Marsden [2005], Regularization of the amended potential and the bifurcation of relative equilibria, *J. of Nonlinear Sci.*, **15**, 93–132.

Hochgerner, S. [2007] Singular cotangent bundle reduction and spin Calogero-Moser systems, *Diff. Geom. Appl.*, to appear.

Holm, D.D. [2002], Euler-Poincaré dynamics of perfect complex fluids, in *Geometry, Dynamics, and Mechanics: 60th Birthday Volume for J.E. Marsden.* P. Holmes, P. Newton, and A. Weinstein, eds., Springer-Verlag.

Holm, D.D. and B.A. Kupershmidt [1982], Poisson structures of superfluids, *Phys. Lett. A*, **91**, 425–430.

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

Holm, D.D. and B.A. Kupershmidt [1983b], Poisson structures and superconductors, *Phys. Lett. A*, **93**, 177–181.

Holm, D.D. and B.A. Kupershmidt [1988], The analogy between spin glasses and Yang-Mills fluids, *J. Math. Phys.*, **29**, 21–30.

Holm, D. D., J.E. Marsden, and T.S. Ratiu [1986], The Hamiltonian structure of continuum mechanics in material, spatial and convective representations, in *Séminaire de Mathématiques supérieurs*, **100**, 11–122, Les Presses de L'Univ. de Montréal, Montréal.

Holm, D.D., J.E. Marsden, and T.S. Ratiu [1998], The Euler–Poincaré equations and semidirect products with applications to continuum theories, *Adv. in Math.*, **137**, 1–8.

Holm, D.D., J.E. Marsden, and T.S. Ratiu [2002], The Euler–Poincaré equations in geophysical fluid dynamics. In Norbury, J. and I. Roulstone, editors, *Large-Scale Atmosphere-Ocean Dynamics II: Geometric Methods and Models*, pages 251–300. Cambridge Univ. Press.

Holm, D. D., J.E. Marsden, T.S. Ratiu, and A. Weinstein [1985], Nonlinear stability of fluid and plasma equilibria, *Phys. Rep.* **123**, 1–6.

Hopf, H. [1931], Über die Abildungen der dreidimensionalen sphäre auf die Kugelfläche, *Math. Annalen*, **104**, 38–63.

Huebschmann, J. [1998], Smooth structures on certain moduli spaces for bundles on a surface, *J. Pure Appl. Algebra* **126**, 183–221.

Huebschmann, J. [2001], Singularities and Poisson geometry of certain representation spaces, in *Quantization of singular symplectic quotients*, *Progr. Math.*, volume 198, 119–135. Birkhäuser Boston, Boston, MA.

Iglésias P. [1995], La trilogie du moment, *Ann. Inst. Fourier (Grenoble)*, **45**(3), 825–857.

Ismagilov, R., M. Losik, and P. Michor [2006], A 2-cocycle on a group of symplectomorphisms, *preprint.*

Iwai, T. [1982], The symmetry group of the harmonic oscillator and its reduction, *J. Math. Phys.*, **23**, 1088–1092.

Iwai, T. [1985], On reduction of two degrees of freedom Hamiltonian systems by an S^1 action and $SO_0(1,2)$ as a dynamical group, *J. Math. Phys.*, **26**, 885–893.

Iwai, T. [1987], A geometric setting for classical molecular dynamics, *Ann. Inst. Henri Poincaré, Phys. Th.*, **47**, 199–219.

Iwai, T. [1990], On the Guichardet/Berry connection, *Phys. Lett. A* **149**, 341–344.

Jacobi, C.G.K. [1843], *J. für Math.*, **26**, 115.

Jacobi, C.G.K. [1866], *Vorlesungen über Dynamik*. Verlag G. Reimer. Based on lectures given in 1842-3, Reprinted by Chelsea, 1969.

Jalnapurkar, S.M. and J.E. Marsden [2000], Reduction of Hamilton's variational principle, *Dynam. Stability Systems*, **15**, 287–318.

Jalnapurkar, S.M., M. Leok, J.E. Marsden and M. West [2006], Discrete Routh reduction, *J. Phys. A: Math. Gen.*, **39**, 5521–5544.

Kane, C., J.E. Marsden, M. Ortiz, and M. West [2000], Variational integrators and the Newmark algorithm for conservative and dissipative mechanical systems, *Int. J. Num. Math. Eng.*, **49**, 1295–1325.

Kanso, E., J.E. Marsden, C.W. Rowley, and J. Melli-Huber [2005], Locomotion of articulated bodies in a perfect fluid, *J. Nonlinear Science*, **15**, 255–289.

Kappeler, T. and J. Pöschel [2003], *KdV & KAM*, Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge, volume 45. Springer-Verlag, Berlin.

Karapetyan, A.V. [2000], On construction of the effective potential in singular cases, *Regul. Chaotic Dyn.*, **5**, 219–224.

Kato, T. [1983], On the Cauchy problem for the (generalized) Korteweg-de Vries equation, in *Studies in applied mathematics*, volume 8 of *Adv. Math. Suppl. Stud.*, pages 93–128. Academic Press, New York.

Kato, T. [1991], Weak solutions of infinite-dimensional Hamiltonian systems, in *Frontiers in Pure and Applied Mathematics*, North Holland, 133–149.

Kazhdan, D., B. Kostant, and S. Sternberg [1978], Hamiltonian group actions and dynamical systems of Calogero type, *Comm. Pure Appl. Math.*, **31**, 481–508.

Khesin, B. and G. Misiołek [2003], Euler equations on homogeneous spaces and Virasoro orbits, *Adv. Math.*, **176**, 116–144.

Kijowski, J. and W. Tulczyjew [1979], *A Symplectic Framework for Field Theories*, Springer Lect. Notes in Physics, volume 107.

Kirillov, A.A. [1962], Unitary representations of nilpotent Lie groups, *Russian Math. Surveys*, **17**, 53–104.

Kirillov, A.A. [1976a], *Elements of the Theory of Representations*, Grundlehren der Math. Wiss., volume 220, Springer-Verlag.

Kirillov, A.A. [1976b], Local Lie Algebras, *Russian Math. Surveys* **31**, 55–75.

Kirillov, A.A. [1982], Infinite dimensional Lie groups; their orbits, invariants and representations. The geometry of moments, in *Lecture Notes in Mathematics*, **970**, 101–123.

Kirk, V., J.E. Marsden, and M. Silber [1996], Branches of stable three-tori using Hamiltonian methods in Hopf bifurcation on a rhombic lattice, *Dyn. and Stab. of Systems*, **11**, 267–302.

Kirwan, F.C. [1984a], *Cohomology Quotients in Symplectic and Algebraic Geometry*, Princeton Math. Notes, volume 31. Princeton University Press.

Kirwan, F.C. [1984b], Convexity properties of the moment map III, *Invent. Math.*, **77**, 547–552.

Knapp, A. W. [2002], *Lie Groups Beyond an Introduction*, Second edition, *Progress in Mathematics*, volume **140**, Birkhäuser, Boston.

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

Koiller, J. [1992], Reduction of some classical nonholonomic systems with symmetry, *Arch. Rational Mech. Anal.*, **118**, 113–148.

Koon, W.S., M. Lo, J.E. Marsden, and S. Ross [2000], Heteroclinic Connections between periodic orbits and resonance transitions in celestial mechanics, *Chaos*, **10**, 427–469.

Koon, W.S. and J.E. Marsden [1997], Optimal control for holonomic and nonholonomic mechanical systems with symmetry and Lagrangian reduction, *SIAM J. Control and Optim.*, **35**, 901–929.

Koon, W.S. and J.E. Marsden [1998], The Poisson reduction of nonholonomic mechanical systems, *Reports on Math. Phys.*, **42**, 101–134.

Kostant, B. [1966], Orbits, symplectic structures and representation theory, in *Proc. US-Japan Seminar on Diff. Geom. (Kyoto)*, Nippon Hyronsha, Tokyo., **77**.

Kostant, B. [1970], Quantization and unitary representations, in *Springer Lect. Notes in Math.*, **570**, 177–306.

Kovalevsky, J. [1967], *Introduction to Celestial Mechanics*, Springer-Verlag, New York.

Kriegl, A. and P.W. Michor [1997], *The Convenient Setting of Global Analysis*, Mathematical Surveys and Monographs, volume 53, American Mathematical Society, Providence, RI.

Krishnaprasad, P.S. [1990], Geometric phases and optimal reconfiguration for multibody systems, in *Proc. Am. Control Conf.*, pages 2440–2444.

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

Krishnaprasad, P.S., J.E. Marsden, and T. Posbergh [1987], Stability analysis of a rigid body with a flexible attachment using the enregy-Casimir method, *Contemp. Math., Amer. Math. Soc.*, **68**, 253–273.

Krishnaprasad, P.S., R. Yang, and W.P. Dayawansa [1993], Chaplygin dynamics and Lagrangian reduction, in *Proc. 2nd Int. Cong. on Nonlinear Mechanics*, W-Z. Chien, Z.H. Guo & Y.Z. Guo, eds., 745–749. Peking University Press.

Krupa, M., M. Schagerl, A. Steindl, P. Szmolyan, and H. Troger [2001], Relative equilibria of tethered satellite systems and their stability for very stiff tethers, *Dynamical Systems: An Int. J.*, **16**, 253–278.

Kummer, M. [1981], On the construction of the reduced phase space of a Hamiltonian system with symmetry, *Indiana Univ. Math. J.*, **30**, 281–291.

Kummer, M. [1990], On resonant classical Hamiltonians with n frequencies, *J. Diff. Eqns.*, **83**, 220–243.

Lagrange, J. L. [1788], *Mécanique Analytique*, Chez la Veuve Desaint, Paris.

Landsman, N.P. [1995], Rieffel induction as generalized quantum Marsden-Weinstein reduction, *J. Geom. Phys.*, **15**, 285–319. (Erratum: *J. Geom. Phys.*, **17**, 298).

Landsman, N.P. [1998], *Mathematical Topics Between Classical and Quantum Mechanics*, Springer Monographs in Mathematics, Springer-Verlag, New York.

Lazutkin, V. and T. Pankratova [1975], Normal forms and versal deformations for Hill's equation, *Funkc. Anal. Prilozh.* , **9**, 41–48.

Lehto, O. [1987], *Univalent Functions and Teichmüller Spaces*, Graduate Texts in Mathematics, **109**, Springer-Verlag.

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.

Lerman, E., Montgomery, R., and Sjamaar, R. [1993] Examples of singular reduction, in *Symplectic Geometry*, 127–155, London Math. Soc. Lecture Note Ser., **192**, Cambridge Univ. Press, Cambridge.

Lerman, E. and S.F. Singer [1998], Stability and persistence of relative equilibria at singular values of the moment map, *Nonlinearity*, **11**, 1637–1649.

Lerman, E. and T.F. Tokieda [1999], On relative normal modes, *C. R. Acad. Sci. Paris Sér. I Math.*, **328**, 413–418.

Lew, A., J.E. Marsden, M. Ortiz, and M. West [2004], Variational time integrators, *Internat. J. Numer. Methods Engrg.*, **60**, 153–212.

Lewis, D. [1989], Nonlinear stability of a rotating planar liquid drop, *Arch. Rational Mech. Anal.*, **106**, 287–333.

Lewis, D. [1992a], Lagrangian block diagonalization, *Dyn. Diff. Eqn's*, **4**, 1–42.

Lewis, D. [1992b], Bifurcation of liquid drops, *Nonlinearity*, **6**, 491–522.

Lewis, D. [1994], Quasi-periodic motions of planar liquid drops, *Contemp. Math.*, **170**, 185–199.

Lewis, D. [1995], Linearized dynamics of symmetric Lagrangian systems. In *Hamiltonian dynamical systems*, IMA Math. Appl., **63**, 195–216. Springer-Verlag, New York.

Lewis, D., J.E. Marsden, R. Montgomery, and T.S. Ratiu [1986], The Hamiltonian structure for dynamic free boundary problems, *Physica D*, **18**, 391–404.

Lewis, D., J.E. Marsden, and T. Ratiu [1986], Formal stability of liquid drops with surface tension. In Shlesinger, M. F., R. Cawley, A. W. Saenz, and W. Zachary, editors, *Perspectives in Nonlinear Dynamics*, pages 71–83. World Scientific.

Lewis, D., J.E. Marsden, and T.S. Ratiu [1987], Stability and bifurcation of a rotating liquid drop, *J. Math. Phys.*, **28**, 2508–2515.

Lewis, D., J.E. Marsden, T.S. Ratiu, and J.C. Simo [1990], Normalizing connections and the energy–momentum method, in *Proceedings of the CRM conference on Hamiltonian systems, Transformation Groups and Spectral Transform Methods*, Harnad and Marsden, eds., , 207–227, CRM Press.

Lewis, D., T.S. Ratiu, J.C. Simo, and J.E. Marsden [1992], The heavy top: a geometric treatment, *Nonlinearity*, **5**, 1–48.

Lewis, D. and J.C. Simo [1990], Nonlinear stability of rotating pseudo-rigid bodies, *Proc. Roy. Soc. Lon. A*, **427**, 281–319.

Libermann, P. and C.-M. Marle [1987], *Symplectic Geometry and Analytical Mechanics*, Kluwer Academic Publishers.

Lichnerowicz, A. [1977], Les variétés de Poisson et leurs algèbres de Lie associées, *J. Diff. Geom.*, **12**, 253–300.

Lie, S. [1890], *Theorie der Transformationsgruppen*, Volumes I-III, Teubner, Leipzig, reprinted by Chelsea (AMS).

Littlejohn, R. and M. Reinch [1997], Gauge fields in the separation of rotations and internal motions in the n-body problem, *Rev. Mod. Phys.* **69**, 213–275.

Mackey, G. [1968] *Induced Representations of Groups and Quantum Mechanics*, W.A. Benjamin, Inc.

Maier, P. and Neeb, K.-H. [2003], Central extensions of current groups, *Math. Ann.*, **326**, 367–415.

Marle, C.-M. [1976], Symplectic manifolds, dynamical groups and Hamiltonian mechanics. In *Differential Geometry and Relativity*. M. Cahen, and M. Flato (eds.). Reidel.

Marsden, J.E. [1976], Well-posedness of the equations of a nonhomogeneous perfect fluid, *Comm. Partial Differential Equations*, **1**, 215–230.

Marsden, J.E. [1981], *Lectures on Geometric Methods in Mathematical Physics*, CBMS series, SIAM, Philadelphia, PA.

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 Math. Soc. Lecture Note Ser., volume 174, Cambridge University Press.

Marsden, J.E. [1999], Park City lectures on mechanics, dynamics and symmetry, in Eliashberg, Y. and L. Traynor, editors, *Symplectic Geometry and Topology*, *IAS/Park City Math. Ser.*, **7**, 335–430, Amer. Math. Soc., Providence, RI.

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

Marsden, J., G. Misiołek, M. Perlmutter, and T.S. Ratiu [1998], Symplectic reduction for semidirect products and central extensions, *Diff. Geom. and its Appl.*, **9**, 173–212.

Marsden, J. E., R. Montgomery, P.J. Morrison, and W.B. Thompson [1986], Covariant Poisson brackets for classical fields, *Annals of Physics*, **169**, 29–48.

Marsden, J.E., R. Montgomery, and T.S. Ratiu [1990], *Reduction, Symmetry and Phases in Mechanics*, Memoirs Amer. Math. Soc., **436**.

Marsden, J.E., P.J. Morrison, and A. Weinstein [1984], The Hamiltonian structure of the BBGKY hierarchy equations, *Contemp. Math.* **28**, 115–124.

Marsden, J.E. and J. Ostrowski [1996], Symmetries in motion: Geometric foundations of motion control, *Nonlinear Sci. Today.* (http://link.springer-ny.com).

Marsden, J.E., G.W. Patrick, and W.F. Shadwick (Eds.) [1996], *Integration Algorithms and Classical Mechanics*, Fields Inst. Commun., **10**, Amer. Math. Soc.

Marsden, J.E., G.W. Patrick, and S. Shkoller [1998], Multisymplectic geometry, variational integrators and nonlinear PDEs, *Comm. Math. Phys.*, **199**, 351–395.

Marsden, J.E., S. Pekarsky, and S. Shkoller [1999], Discrete Euler–Poincaré and Lie–Poisson equations, *Nonlinearity*, **12**, 1647–1662.

Marsden, J.E. and M. Perlmutter [2000], The orbit bundle picture of cotangent bundle reduction, *C. R. Math. Acad. Sci. Soc. R. Can.*, **22**, 33–54.

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 [1999], *Introduction to Mechanics and Symmetry*, Texts in Applied Mathematics, **17**, Springer-Verlag, 1994; Second Edition, 1999, reprinted with corrections, 2003.

Marsden, J.E., T.S. Ratiu, and J. Scheurle [2000], Reduction theory and the Lagrange-Routh equations, Amer. Inst. Phys., Melville, NY, *J. Math. Phys.*, **41**, 3379–3429.

Marsden, J.E., T.S. Ratiu, and A. Weinstein [1984a], Semidirect products and reduction in mechanics, *Trans. Amer. 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, *Contemp. Math., Am. Math. Soc.*, **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 Inst. Comm.*, **1**, 139–164.

Marsden, J.E. and S. Shkoller [1999], Multisymplectic geometry, covariant Hamil-
tonians and water waves, *Math. Proc. Camb. Phil. Soc.*, **125**, 553–575.

Marsden, J.E. and J.C. Simo [1990], The energy–momentum method, *La
"Mécanique Analytique" de Lagrange et son Héritage*, Atti della Accademia
delle Scienze di Torino, **124**, 245–268.

Marsden, J.E., J.C. Simo, D.K. Lewis, and T.A. Posbergh [1989], A block diag-
onalization theorem in the energy–momentum method, *Contemp. Math.* **97**,
297–313.

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

Marsden, J. E. and A. Weinstein [1982], The Hamiltonian structure of the
Maxwell–Vlasov equations, *Physica D*, **4**, 394–406.

Marsden, J.E. and A. Weinstein [1983], Coadjoint orbits, vortices and Clebsch
variables for incompressible fluids, *Physica D*, **7**, 305–323.

Marsden, J.E. and A. Weinstein [2001], Comments on the history, theory,
and applications of symplectic reduction. In Landsman, N., M. Pflaum, and
M. Schlichenmaier, editors, *Quantization of Singular Symplectic Quotients*.
Progress in Mathematics, volume 198. Birkhäuser Boston, pp 1-20.

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

Marsden, J.E. and M. West [2001], Discrete mechanics and variational integrators,
Acta Numerica, **10**, 357–514.

Martin, J.L. [1959], Generalized classical dynamics and the "classical analogue"
of a Fermi oscillation, *Proc. Roy. Soc. A*, **251**, 536–542.

McDuff, D. [1988], The moment map for circle actions on symplectic manifolds,
J. Geom. Phys., **5**, 149–160.

McDuff, D. and D. Salamon [1995], *Introduction to Symplectic Topology*, Oxford
University Press.

Mestdag, T. [2005], Lagrangian reduction by stages for non-holonomic systems
in a Lie algebroid framework, *J. Phys. A*, **38**, 10157–10179.

Meyer, K.R. [1973], Symmetries and integrals in mechanics, in *Dynamical Sys-
tems*, M. Peixoto, ed., 259–273, Academic Press.

Mielke, A. [1991], *Hamiltonian and Lagrangian Flows on Center Manifolds, with
Applications to Elliptic Variational Problems*, Springer Lect. Notes in Math.,
1489.

Mielke, A. [2002a], On the energetic stability of solitary water waves, *Phil. Trans.
Roy. Soc. A.*, **360**, 2337–2358.

Mielke, A. [2002b], Finite elastoplasticity Lie groups and geodesics on SL(d), in *Geometry, mechanics, and dynamics*, 61–90, Springer, New York.

Mikami, K. and A. Weinstein [1988], Moments and reduction for symplectic groupoid actions, *Publ. RIMS Kyoto Univ.*, **24**, 121–140.

Mishchenko, A.S. and A.T. Fomenko [1978], Generalized Liouville method of integration of Hamiltonian systems, *Funct. Anal. Appl.*, **12**, 113–121.

Misiołek, G. [1997], Conjugate points in the Bott–Virasoro group and the KdV equation, *Proc. Amer. Math. Soc.*, **125**, 935–940.

Misiołek, G. [1998], A shallow water equation as a geodesic flow on the Bott–Virasoro group, *J. Geom. Phys.*, **24**, 203–208.

Misner, C., K. Thorne, and J. A. Wheeler [1973], *Gravitation*, W. H. Freeman, San Francisco.

Mittagunta, G.G. [1994] *Reduced Spaces for Coupled Rigid Bodies and their Relation to Relative Equilibria*, Ph.D. Thesis, University of California, Santa Cruz, 1994.

Mittagunta, G.G. [1996], Reduced spaces for coupled rigid bodies, *J. Nonlinear Sci.*, **6**, 293–310.

Montaldi, J. [1997], Persistence and stability of relative equilibria, *Nonlinearity*, **10**, 449–466.

Montaldi, J.A., R.M. Roberts, and I.N. Stewart [1988], Periodic solutions near equilibria of symmetric Hamiltonian systems, *Phil. Trans. R. Soc. Lond. A*, **325**, 237–293.

R. Montgomery [1983], The structure of reduced cotangent phase spaces for non-free group actions, preprint **143** of the U.C. Berkeley Center for Pure and App. Mathematics.

Montgomery, R. [1984], Canonical formulations of a particle in a Yang–Mills field, *Lett. Math. Phys.*, **8**, 59–67.

Montgomery, R. [1986], *The Bundle Picture in Mechanics*, Ph.D. Thesis, University of California, Berkeley.

Montgomery, R. [1988], The connection whose holonomy is the classical adiabatic angles of Hannay and Berry and its generalization to the nonintegrable case, *Comm. Math. Phys.*, **120**, 269–294.

Montgomery, R. [1990], Isoholonomic problems and some applications, *Comm. Math Phys.*, **128**, 565–592.

Montgomery, R. [1991a], Optimal control of deformable bodies and its relation to gauge theory, in *The Geometry of Hamiltonian Systems*, T. Ratiu, ed., Springer-Verlag.

Montgomery, R. [1991b], How much does a rigid body rotate? A Berry's phase from the 18th century, *Amer. J. Phys.*, **59**, 394–398.

Montgomery, R. [1993], Gauge theory of the falling cat, *Fields Inst. Commun.*, **1**, 193–218.

Montgomery, R. [1995], The geometric phase for the three body problem, *Nonlinearity*, **9**, 1341–1360.

Montgomery, R., J.E. Marsden, and T.S. Ratiu [1984], Gauged Lie–Poisson structures, *Contemp. Math., Amer. Math. Soc.*, **28**, 101–114.

Morrison, P.J. [1987], Variational principle and stability of nonmonotone Vlasov–Poisson equilibria, *Z. Naturforsch.*, **42a**, 1115–1123.

Morrison, P.J. and J.M. Greene [1980], Noncanonical Hamiltonian density formulation of hydrodynamics and ideal magnetohydrodynamics, *Phys. Rev. Lett.*, **45**, 790–794; errata **48** (1982), 569.

Moser, J. [1958], New aspects in the theory of stability of Hamiltonian systems, *Comm. Pure Appl. Math.*, **XI**, 81–114.

Moser, J. [1965], On the volume elements on a manifold, *Trans. Am. Math. Soc.*, **120**, 286–294.

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

Moser, J. [1973], *Stable and Random Motions in Dynamical Systems with Special Emphasis on Celestial Mechanics*, Princeton University Press.

Moser, J. [1976], Periodic orbits near an equilibrium and a theorem by Alan Weinstein, *Comm. Pure Appl. Math.*, **29**, 724–747.

Naber, G.L. [1997], *Topology, Geometry, and Gauge Fields: Foundations*, Springer-Verlag, New York.

Nag, S. and A. Verjovsky [1990], Diff(S^1) and the Teichmüller spaces, *Comm. Math. Phys.*, **130**, 123–138.

Nambu, Y. [1973], Generalized Hamiltonian dynamics, *Phys. Rev. D*, **7**, 2405–2412.

Neeb, K.-H. [2002], Central extensions of infinite-dimensional Lie groups, *Ann. Inst. Fourier (Grenoble)*, **52**(5), 1365–1442.

Neeb, K.-H. [2004a], Abelian extensions of infinite-dimensional Lie groups, preprint.

Neeb, K.-H. [2004b], Infinite-dimensional Groups and their Representations. In *Lie theory*, Progress in Mathematics **228**, 213–328. Birkhäuser Boston, Boston, MA.

Nekhoroshev, N.M. [1971], Action angle variables and their generalizations, *Trans. Moscow Math. Soc.*, **26**, 180–198.

Nill, F. [1983], An effective potential for classical Yang–Mills fields as outline for bifurcation on gauge orbit space, *Ann. Phys.*, **149**, 179–202.

Oh, Y.G., N. Sreenath, P.S. Krishnaprasad, and J.E. Marsden [1989], The dynamics of coupled planar rigid bodies Part 2: bifurcations, periodic solutions and chaos, *Dynamics and Diff. Eq.*, **1**, 269–298.

Ortega, J.-P. [1998], *Symmetry, Reduction, and Stability in Hamiltonian Systems*. Ph.D. Thesis. University of California, Santa Cruz. June, 1998.

Ortega, J.-P. [2002], The symplectic reduced spaces of a Poisson action, *C. R. Acad. Sci. Paris Sér. I Math.*, **334**, 999–1004.

Ortega, J.-P. [2003a], Singular dual pairs, *Differential Geom. Appl.*, **19**, 61–95.

Ortega, J.–P. [2003b], Relative normal modes for nonlinear Hamiltonian systems, *Proc. Roy. Soc. Edinburgh Sect. A*, **133**, 665–704.

Ortega, J.-P. and V. Planas-Bielsa [2004], Dynamics on Leibniz manifolds. *J. Geom. Phys.*, **52**(1), 1–27

Ortega, J.-P. and T.S. Ratiu [1997], Persistence and smoothness of critical relative elements in Hamiltonian systems with symmetry, *C.R. Acad. Sci. Paris Sr. I Math.*, **325**, 1107–1111.

Ortega, J.-P. and Ratiu, T.S. [1998], Singular reduction of Poisson manifolds. *Letters in Mathematical Physics*, **46**, 359–372.

Ortega, J.-P. and T.S. Ratiu [1999a], Non-linear stability of singular relative periodic orbits in Hamiltonian systems with symmetry, *J. Geom. Phys.*, **32**, 160–188.

Ortega, J.-P. and T.S. Ratiu [1999b], Stability of Hamiltonian relative equilibria, *Nonlinearity*, **12**(3), 693–720.

Ortega, J.-P. and T.S. Ratiu [2001], A symplectic slice theorem. *Lett. Math. Phys.*, **59**, 81–93.

Ortega, J.-P. and T.S. Ratiu [2002], The optimal momentum map, in *Geometry, Dynamics, and Mechanics: 60th Birthday Volume for J.E. Marsden*, pages 329–362. P. Holmes, P. Newton, and A. Weinstein, eds., Springer-Verlag.

Ortega, J.-P. and T.S. Ratiu [2004a], *Momentum Maps and Hamiltonian Reduction*, volume 222 of Progress in Mathematics. Birkhäuser Boston Inc., Boston, MA.

Ortega, J.-P. and T.S. Ratiu [2004b], Relative equilibria near stable and unstable Hamiltonian relative equilibria, *Proc. R. Soc. Lond. Ser. A*, **460**(2045), 1407–1431.

Ortega, J.-P. and T.S. Ratiu [2006a], The reduced spaces of a symplectic Lie group action, *Annals of Global Analysis and Geometry*, **30**, 335–381.

Ortega, J.-P. and T.S. Ratiu [2006b], The stratified spaces of a symplectic Lie group action, *Reports on Mathematical Physics*, **58**, 51-75.

Otto, M. [1987], A reduction scheme for phase spaces with almost Kähler symmetry. Regularity results for momentum level sets, *J. Geom. Phys.*, **4**, 101–118.

Ovsienko, V.Y. and B.A. Khesin [1987], Korteweg-de Vries superequations as an Euler equation, *Funct. Anal. and Appl.*, **21**, 329–331.

Palais, R.S. [1957], *A Global Formulation of the Lie Theory of Transformation Groups*, Mem. Amer. Math. Soc., volume **22**.

Palais, R.S. [1968], *Foundations of Global Non-Linear Analysis*, Benjamin/Cummins Publishing Co., Reading, MA.

Paterson, A.L.T. [1999], *Groupoids, Inverse Semigroups, and their Operator Algebras*. Progress in Mathematics, volume 170. Birkhäuser.

Patrick, G.W. [1989], The dynamics of two coupled rigid bodies in three space, *Contemp. Math.*, **97**, 315–336.

Patrick, G.W. [1990], *Two Axially Symmetric Coupled Rigid Bodies: Relative Equilibria, Stability, Bifurcations, and a Momentum Preserving Symplectic Integrator*. Ph.D. Thesis, University of California, Berkeley.

Patrick, G.W. [1992], Relative equilibria in Hamiltonian systems: The dynamic interpretation of nonlinear stability on a reduced phase space, *J. Geom. and Phys.*, **9**, 111–119.

Patrick, G.W. [1995], Relative equilibria of Hamiltonian systems with symmetry: linearization, smoothness and drift, *J. Nonlinear Sci.*, **5**, 373–418.

Patrick, G.W. [1999], The Landau-Lifshitz equation by semidirect product reduction, *Lett. Math. Phys.*, **50**, 177–188.

Patrick, G.W., M. Roberts, and C. Wulff [2004], Stability of Poisson equilibria and Hamiltonian relative equilibria by energy methods, *Arch. Ration. Mech. Anal.*, **174**, 301–344.

Pauli, W. [1953], On the Hamiltonian structure of nonlocal field theories, *Il Nuovo Cimento*, **X**, 648–667.

Pedroni, M. [1995], Equivalence of the Drinfeld-Sokolov reduction to a bi-Hamiltonian reduction, *Lett. Math. Phys.*, **35**, 291–302.

Perlmutter, M. [1999], *Symplectic Reduction by Stages*, Ph.D. Thesis, Mathematics, University of California, Berkeley.

Perlmutter, M. and T.S. Ratiu [2005], Gauged Poisson structures, preprint.

Perlmutter, M. and M. Rodríguez-Olmos [2006], On singular Sternberg spaces, in progress.

Perlmutter, M., M. Rodríguez-Olmos, and M.E. Sousa Dias [2006], On the geometry of reduced cotangent bundles at zero momentum, *J. of Geom. and Phys.*, **57**, 571–596.

Perlmutter, M., M. Rodríguez-Olmos, and M.E. Sousa Dias [2007], On the symplectic normal space for cotangent lifted actions, *Diff. Geom. Appl.*, to appear; see http://front.math.ucdavis.edu/math.SG/0501207.

Pflaum, M. J. [2001], *Analytic and Geometric Study of Stratified Spaces*, volume 1768 of Lecture Notes in Mathematics, Springer–Verlag.

Planas-Bielsa, V. [2004], Point reduction in almost symplectic manifolds, *Rep. Math. Phys.*, **54**, 295–308.

Poincaré, H. [1885], Sur l'équilibre d'une masse fluide animée d'un mouvement de rotation, *Acta. Math.*, **7**, 259.

Poincaré, H. [1890], Sur le problème des trois corps et les équations de la dynamique, *Acta Math.*, **13**, 1–27.

Poincaré, H. [1901a], Sur une forme nouvelle des équations de la mécanique, *C. R. Acad. Sci.*, **132**, 369–371.

Poincaré, H. [1901b], Sur la stabilité de l'équilibre des figures piriformes affectées par une masse fluide en rotation, *Philosophical Transactions A*, **198**, 333–373.

Poincaré, H. [1910], Sur la precession des corps deformables, *Bull. Astron.*, **27**, 321–356.

Pressley, A. and G. Segal [1986], *Loop Groups*, Oxford University Press.

Ratiu, T.S. [1980a], *The Euler–Poisson Equations and Integrability*, Ph.D. Thesis, University of California, Berkeley.

Ratiu, T.S. [1980b], The motion of the free n-dimensional rigid body, *Indiana Univ. Math. Journal*, **29**, 609–629.

Ratiu, T.S. [1980c], Involution theorems, in *Geometric Methods in Mathematical Physics*, G. Kaiser and J. Marsden, eds., Springer Lecture Notes, **775**, 219–257.

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, *Amer. J. Math.*, **104**, 409–448, 1337.

Roberts, M. and M.E.R. de Sousa Dias [1997], Bifurcations from relative equilibria of Hamiltonian systems, *Nonlinearity*, **10**, 1719–1738.

Roberts, M., C. Wulff, and J.S.W. Lamb [2002], Hamiltonian systems near relative equilibria, *J. Differential Equations*, **179**(2), 562–604.

Rodríguez-Olmos, M. [2006a], The isotropy lattice of a lifted action, *C. R. Math. Acad. Sci. Paris. Ser. I*, **343**, 41–46.

Rodríguez-Olmos, M. [2006b], Stability of relative equilibria with singular momentum values in simple mechanical systems, *Nonlinearity*, **19**, 853–877.

Routh, E.J. [1860], *Treatise on the Dynamics of a System of Rigid Bodies*, MacMillan, London.

Routh, E.J. [1877], *Stability of a Given State of Motion*, Halsted Press, New York; Reprinted in *Stability of Motion* (1975), A. T. Fuller ed.

Routh, E.J. [1884], *Advanced Rigid Dynamics*, MacMillian and Co., London.

Satake, I. [1956], On a generalization of the notion of manifold, *Proc. Nat. Acad. Sci. U.S.A.*, **42**, 359–363.

Satzer, W.J. [1977], Canonical reduction of mechanical systems invariant under Abelian group actions with an application to celestial mechanics, *Ind. Univ. Math. J.*, **26**, 951–976.

Schmah, T. [2002], *Symmetries of Cotangent Bundles*. Ph.D. thesis, École Polytechnique Fédérale de Lausanne, Switzerland

Schonbek, M. E., Todorov, A. N., and Zubelli, J. P. [1999], Geodesic flows on diffeomorphisms of the circle, Grassmannians, and the geometry of the periodic KdV equation, *Adv. Theor. Math. Phys.*, **3**, 1027–1092.

Segal, G. [1981], Unitary representations of some infinite-dimensional groups, *Comm. Math. Phys.* , **80**, 301–342.

Simo, J.C., D.K. Lewis, and J.E. Marsden [1991], Stability of relative equilibria I: The reduced energy momentum method, *Arch. Rational Mech. Anal.*, **115**, 15–59.

Simo, J., T.A. Posbergh, and J.E. Marsden [1989], Stability analysis of a rigid body with attached geometrically nonlinear rod by the energy-momentum method, *Contemp. Math., Amer. Math. Soc.*, **97**, 371–398.

Simo, J.C., T.A. Posbergh, and J.E. Marsden [1990], Stability of coupled rigid body and geometrically exact rods: block diagonalization and the energy-momentum method, *Physics Reports*, **193**, 280–360.

Simo, J.C., T.A. Posbergh, and J.E. Marsden [1991], Stability of relative equilibria II: Three dimensional elasticity, *Arch. Rational Mech. Anal.*, **115**, 61–100.

Sjamaar, R. [1990], *Singular Orbit Spaces in Riemannian and Symplectic Geometry*, Ph.D. Thesis, Rijksuniversiteit te Utrecht.

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

Smale, S. [1970], Topology and Mechanics, *Inv. Math.*, **10**, 305–331; **11**, 45–64.

Śniatycki, J. [1974], Dirac brackets in geometric dynamics, *Ann. Inst. H. Poincaré*, **20**, 365–372.

Śniatycki, J. [1999], A Hamiltonian analysis of Yang-Mills equations, *Rep. Math. Phys.*, **44**, 205–214.

Śniatycki, J. [2001], Almost Poisson spaces and nonholonomic singular reduction, *Rep. Math. Phys.*, **48**, 235–248.

Śniatycki, J. and W. Tulczyjew [1971], Canonical dynamics of relativistic charged particles, *Ann. Inst. H. Poincaré(Grenoble)* **15**, 177–187.

Souriau, J.M. [1966], Quantification géométrique, *Comm. Math. Phys.*, **1**, 374–398.

Souriau, J.-M. [1970], *Structure des Systèmes Dynamiques*, Dunod, Paris. English translation by R.H. Cushman and G.M Tuynman *Structure of Dynamical Systems*, volume 149 of Progress In Mathematics, Birkhäuser Boston Inc., Boston, MA, 1997.

Spivak, M. [1979], *A Comprehensive Introduction to Differential Geometry*, Publish or Perish, Willmington, Deleware, 5 volumes.

Sreenath, N., Y.G. Oh, P.S. Krishnaprasad, and J.E. Marsden [1988], The dynamics of coupled planar rigid bodies. Part 1: Reduction, equilibria and stability, *Dyn. and Stab. of Systems*, **3**, 25–49.

Stefan, P. [1974a], Accessibility and foliations with singularities, *Bull. Amer. Math. Soc.*, **80**, 1142–1145.

Stefan, P. [1974b], Accessible sets, orbits and foliations with singularities. *Proc. Lond. Math. Soc.*, **29**, 699–713.

Sternberg, S. [1977], Minimal coupling and the symplectic mechanics of a classical particle in the presence of a Yang–Mills field, *Proc. Nat. Acad. Sci.*, **74**, 5253–5254.

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

Sussman, H. [1973], Orbits of families of vector fields and integrability of distributions, *Trans. Amer. Math. Soc.*, **180**, 171–188.

Takhtajan, L. and L. P. Teo [2004], Weil–Petersson metric on the universal Teichmüller space I: curvature properties and Chern forms, *ArXiv, math.CV/0312172, and* Weil-Petersson metric on the universal Teichmüller space II: Kähler potential and period mapping, *ArXiv, math.CV/0406408*.

Takhtajan, L. and L. P. Teo [2005], Weil-Petersson geometry of the universal Teichmüller space. In *Infinite dimensional algebras and quantum integrable systems*, *Progr. Math.*, **237**, 225–233, Birkhuser, Basel.

Takhtajan, L. and L. P. Teo [2006], Weil-Petersson Metric on the Universal Teichmller Space, Memoirs of the AMS, **183**.

Tanimura, S. and T. Iwai [2000], Reduction of quantum systems on Riemannian manifolds with symmetry and application to molecular mechanics, *J. Math. Phys.*, **41**, 1814–1842.

Teichmüller, O. [1939], Extremale quasikonforme Abbildungen und quadratische Differentiale, *Abh. Preuss. Akad. Wiss. Math.-Nat. Kl.*, **22**, 197 pp.

Tromba, A. J. [1986], On a natural algebraic affine connection on the space of almost complex structures and the curvature of Teichmüller space with respect to its Weil-Petersson metric, *Manuscripta Math.*, **56**, 475–497.

Tromba, A. J. [1992], *Teichmüller Theory in Riemannian Geometry,* Lecture notes prepared by Jochen Denzler; Lectures in Mathematics ETH Zrich, Birkhäuser Verlag, Basel.

Tulczyjew, W.M. [1977], The Legendre transformation, *Ann. Inst. Poincaré*, **27**, 101–114.

Tulczyjew, W. M. and P. Urbański [1999], A slow and careful Legendre transformation for singular Lagrangians, *Acta Phys. Polon. B* **30**, 2909–2978. The Infeld Centennial Meeting (Warsaw, 1998).

Vaisman, I. [1994], *Lectures on the Geometry of Poisson Manifolds*, Progress in Mathematics, volume 118, Birkhäuser, Basel.

Vanhaecke, P. [1996], *Integrable Systems in the Realm of Algebraic Geometry*, Lecture Notes in Math., volume 1638, Springer-Verlag, New York.

Vasylkevych, S. and J. E. Marsden [2005], The Lie-Poisson structure of the Euler equations of an ideal fluid, *Dynamics of PDE*, **2** (4), 281–300.

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

Vinti, J. P. [1998], *Orbital and Celestial Mechanics*, AIAA, Virginia.

Wang, L.S. and P.S. Krishnaprasad [1992], Gyroscopic control and stabilization, *J. Nonlinear Sci.*, **2**, 367–415.

Wang, L.S., P.S. Krishnaprasad, and J.H. Maddocks [1991], Hamiltonian dynamics of a rigid body in a central gravitational field, *Cel. Mech. Dyn. Astr.*, **50**, 349–386.

Weinstein, A. [1971], Symplectic manifolds and their Lagrangian submanifolds, *Adv. Math.*, **6**, 329–346; see also *Bull. Am. Math. Soc.*, **75** (1969), pp. 1040–1041.

Weinstein, A. [1973], Normal modes for nonlinear Hamiltonian systems, *Inv. Math.*, **20**, 47–57.

Weinstein, A. [1977], *Lectures on Symplectic Manifolds*, **29**, CBMS Regional Conf. Ser. in Math., Conf. Board Math. Sci., Washington, DC.

Weinstein, A. [1978a], A universal phase space for particles in Yang–Mills fields, *Lett. Math. Phys.*, **2**, 417–420.

Weinstein, A. [1978b], Bifurcations and Hamilton's principle, *Math. Zeit.*, **159**, 235–248.

Weinstein, A. [1983a], The local structure of Poisson manifolds. *J. Differential Geometry*, **18**, 523–557.

Weinstein, A. [1983b], Sophus Lie and symplectic geometry, *Expo. Math.*, **1**, 95–96.

Weinstein, A. [1984], Stability of Poisson–Hamilton equilibria, *Contemp. Math.*, **28**, 3–14.

Weinstein, A. [1990], Connections of Berry and Hannay type for moving Lagrangian submanifolds, *Adv. in Math.*, **82**, 133–159.

Weinstein, A. [1996], Lagrangian mechanics and groupoids, *Fields Inst. Commun.*, **7**, 207–231.

Wendlandt, J.M. and J.E. Marsden [1997], Mechanical integrators derived from a discrete variational principle, *Physica D*, **106**, 223–246.

Whittaker, E.T. [1937], *A Treatise on the Analytical Dynamics of Particles and Rigid Bodies*, Cambridge University Press; First Edition 1904, Fourth Edition, 1937, Reprinted by Dover 1944 and Cambridge University Press, 1988, fourth edition.

Witten, E. [1988], Coadjoint orbits of the Virasoro group. *Comm. Math. Phys* , **114**, 1–53.

Wolpert, S. A. [1983], On the symplectic geometry of deformations of a hyperbolic surface, *Ann. of Math.*, **117**, 207–234.

Wolpert, S. A. [1986], Chern forms and the Riemann tensor for the moduli space of curves, *Invent. Math.*, **85**, 119–145.

Wolpert, S. A. [2003], Geometry of the Weil-Petersson completion of Teichmüller space, in *Surveys in differential geometry, Vol. VIII (Boston, MA, 2002)*, Int. Press, Somerville, MA. Surv. Differ. Geom., VIII, 357–393.

Woodhouse, N.M.J. [1992], *Geometric Quantization*, Clarendon Press, Oxford.

Wulff, C.[2003], Persistence of relative equilibria in Hamiltonian systems with non-compact symmetry, *Nonlinearity*, **16**, 67–91.

Wulff, C. and M. Roberts [2002], Hamiltonian systems near relative periodic orbits, *SIAM J. Appl. Dyn. Syst.*, **1**, 1–43.

Xu, P. [1991], Morita equivalent symplectic groupoids. In *Symplectic Geometry, Groupoids, and Integrable Systems*, Dazord, P. and Weinstein, A. (editors), 291–311, Springer Verlag.

Zaalani, N. [1999], Phase space reduction and Poisson structure, *J. Math. Phys.*, **40**, 3431–3438.

Zakrzewski, S. [1986], Induced representations and induced hamiltonian actions, *J. Geom. Phys.*, **3**, 211–219.

Zenkov, D.V., A.M. Bloch, and J.E. Marsden [1998], The energy momentum method for the stability of nonholonomic systems, *Dyn. Stab. of Systems*, **13**, 123–166.

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

Zygmund, A. [2002], *Trigonometric Series. Vol. I, II*, third edition, Cambridge University Press.

Index