Mechanics and Symmetry

Reduction Theory

Jerrold E. Marsden and Tudor S. Ratiu

February 3, 1998

Preface

Preface goes here.

Pasadena, CA
Spring, 1998

Jerry Marsden and Tudor Ratiu

Contents

Preface iii

1 Introduction and Overview **1**
1.1 Lagrangian and Hamiltonian Mechanics. 1
1.2 The Euler–Poincaré Equations. 3
1.3 The Lie–Poisson Equations. 8
1.4 The Heavy Top. 10
1.5 Incompressible Fluids. 11
1.6 The Basic Euler–Poincaré Equations. 13
1.7 Lie–Poisson Reduction. 14
1.8 Symplectic and Poisson Reduction. 20

2 Symplectic Reduction **27**
2.1 Presymplectic Reduction 27
2.2 Symplectic Reduction by a Group Action 31
2.3 Coadjoint Orbits as Symplectic Reduced Spaces 38
2.4 Reducing Hamiltonian Systems 40
2.5 Orbit Reduction 42
2.6 Foliation Orbit Reduction 46
2.7 The Shifting Theorem 47
2.8 Dynamics via Orbit Reduction 49
2.9 Reduction by Stages 50

3 Reduction of Cotangent Bundles **53**
3.1 Reduction at Zero 54
3.2 Abelian Reduction 57
3.3 Principal Connections 60
3.4 Cotangent Bundle Reduction—Embedding Version 66
3.5 Cotangent Bundle Reduction—Bundle Version 67
3.6 The Mechanical Connection Revisited 69
3.7 The Poisson Structure on T^*Q/G. 71
3.8 The Amended Potential 71
3.9 Examples 72
3.10 Dynamic Cotangent Bundle Reduction 77
3.11 Reconstruction 77
3.12 Additional Examples 78
3.13 Hamiltonian Systems on Coadjoint Orbits 86
3.14 Energy Momentum Integrators 90
3.15 Maxwell's Equations 91
3.16 Geometric Phases for the Rigid Body 96

3.17 Reconstruction Phases . . . 99
3.18 Dynamics of Coupled Planar Rigid Bodies . . . 100

4 Semidirect Products 117
4.1 Hamiltonian Semidirect Product Theory . . . 117
4.2 Lagrangian Semidirect Product Theory . . . 121
4.3 The Kelvin-Noether Theorem . . . 125
4.4 The Heavy Top . . . 127

5 Semidirect Product Reduction and Reduction by Stages 129
5.1 Semidirect Product Reduction . . . 129
5.2 Reduction by Stages for Semidirect Products . . . 130

Chapter 1

Introduction and Overview

Reduction is of two sorts, Lagrangian and Hamiltonian. In each case one has a group of symmetries and one attempts to pass the structure at hand to an appropriate quotient space. Within each of these broad classes, there are additional subdivisions; for example, in Hamiltonian reduction there is symplectic and Poisson reduction.

These subjects arose from classical theorems of Liouville and Jacobi on reduction of mechanical systems by $2k$ dimensions if there are k integrals in involution. Today, we take a more geometric and general view of these constructions as initiated by Arnold [1966] and Smale [1970] amongst others. The work of Meyer [1973] and Marsden and Weinstein [1974] that formulated symplectic reduction theorems, continued to initiate an avalanch of literature and applications of this theory. Many textbooks appeared that developed and presented this theory, such as Abraham and Marsden [1978], Guillemin and Sternberg [1984], Liberman and Marle [1987], Arnold, Kozlov, and Neishtadt [1988], Arnold [1989], and Woodhouse [1992] to name a few. The present book is intended to present some of the main theoretical and applied aspects of this theory.

With Hamiltonian reduction, the main geometric object one wishes to reduce is the *symplectic or Poisson structure*, while in Lagrangian reduction, the crucial object one wishes to reduce is *Hamilton's variational principle* for the Euler-Lagrange equations.

In this book we assume that the reader is knowledgable of the basic principles in mechanics, as in the authors' book *Mechanics and Symmetry* (Marsden and Ratiu [1998]). We refer to this monograph hereafter as *IMS*.

1.1 Lagrangian and Hamiltonian Mechanics.

Lagrangian Mechanics. The Lagrangian formulation of mechanics can be based on the variational principles behind Newton's fundamental laws of force balance $\mathbf{F} = m\mathbf{a}$. One chooses a configuration space Q (a manifold, assumed to be of finite dimension n to start the discussion) with coordinates denoted $q^i, i = 1, \ldots, n$, that describe the configuration of the system under study. One then forms the velocity phase space TQ (the tangent bundle of Q). Coordinates on TQ are denoted $(q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n)$, and the Lagrangian is regarded as a function $L : TQ \to \mathbb{R}$. In coordinates, one writes $L(q^i, \dot{q}^i, t)$, which is shorthand notation for $L(q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n, t)$. Usually, L is the kinetic minus the potential energy of the system and one takes $\dot{q}^i = dq^i/dt$ to be the system velocity. The variational principle of Hamilton states that the variation of the action is stationary at a solution:

$$\delta\mathfrak{S} = \delta \int_a^b L(q^i, \dot{q}^i, t)\, dt = 0. \tag{1.1.1}$$

In this principle, one chooses curves $q^i(t)$ joining two fixed points in Q over a fixed time interval $[a,b]$, and calculates the action $\mathfrak{S}$, which is the time integral of the Lagrangian, regarded as a function of this curve. Hamilton's principle states that the action $\mathfrak{S}$ has a critical point at a solution in the space of curves. As is well known, Hamilton's principle is equivalent to the Euler–Lagrange equations:

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}^i} - \frac{\partial L}{\partial q^i} = 0, \quad i = 1, \dots, n. \tag{1.1.2}$$

Let $T^{(2)}Q \subset T^2Q$ denote the submanifold which at each point $q \in Q$ consists of second derivatives of curves in Q that pass through q at, say, $t = 0$. We call $T^{(2)}Q$ the ***second order tangent bundle***. The action defines a unique bundle map

$$\mathcal{EL} : T^{(2)}Q \to T^*Q$$

called the ***Euler-Lagrange operator*** such that for a curve $c(\cdot)$ in Q,

$$[\delta\mathfrak{S}(c) \cdot \delta c](t) = \langle \mathcal{EL} \cdot c''(t), \delta c(t) \rangle \,.$$

Thus, the Euler-Lagrange equations can be stated intrinsically as the vanishing of the Euler-Lagrange operator: $\mathcal{EL}(c(\cdot)) = 0$.

If the system is subjected to external forces, these are to be added to the right hand side of the Euler-Lagrange equations. For the case in which L comprises kinetic minus potential energy, the Euler-Lagrange equations reduce to a geometric form of Newton's second law. For Lagrangians that are purely kinetic energy, it was already known in Poincaré's time that the corresponding solutions of the Euler-Lagrange equations are geodesics. (This fact was certainly known to Jacobi by 1840, for example.)

Hamiltonian Mechanics. To pass to the Hamiltonian formalism, one introduces the conjugate momenta

$$p_i = \frac{\partial L}{\partial \dot{q}^i}, \quad i = 1, \dots, n, \tag{1.1.3}$$

and makes the change of variables $(q^i, \dot{q}^i) \mapsto (q^i, p_i)$, by a Legendre transformation. The Lagrangian is called *regular* when this change of variables is invertible. The Legendre transformation introduces the Hamiltonian

$$H(q^i, p_i, t) = \sum_{j=1}^{n} p_j \dot{q}^j - L(q^i, \dot{q}^i, t). \tag{1.1.4}$$

One shows that the Euler–Lagrange equations are equivalent to Hamilton's equations:

$$\frac{dq^i}{dt} = \frac{\partial H}{\partial p_i}, \quad \frac{dp_i}{dt} = -\frac{\partial H}{\partial q^i}, \tag{1.1.5}$$

where $i = 1, \dots, n$. There are analogous Hamiltonian partial differential equations for field theories such as Maxwell's equations and the equations of fluid and solid mechanics.

Hamilton's equations can be recast in Poisson bracket form as

$$\dot{F} = \{F, H\}, \tag{1.1.6}$$

where the canonical Poisson brackets are given by

$$\{F, G\} = \sum_{i=1}^{n} \left(\frac{\partial F}{\partial q^i}\frac{\partial G}{\partial p_i} - \frac{\partial F}{\partial p_i}\frac{\partial G}{\partial q^i} \right). \tag{1.1.7}$$

Associated to any configuration space Q is a phase space T^*Q called the cotangent bundle of Q, which has coordinates $(q^1, \ldots, q^n, p_1, \ldots, p_n)$. On this space, the canonical Poisson bracket is intrinsically defined in the sense that the value of $\{F, G\}$ is independent of the choice of coordinates. Because the Poisson bracket satisfies $\{F, G\} = -\{G, F\}$ and in particular $\{H, H\} = 0$, we see that $\dot{H} = 0$; that is, energy is conserved along solutions of Hamilton's equations. This is the most elementary of many deep and beautiful conservation properties of mechanical systems.

1.2 The Euler–Poincaré Equations.

Poincaré and the Euler equations. Poincaré played an enormous role in the topics treated in this book. His work on the gravitating fluid problem, continued the line of investigation begun by MacLaurin, Jacobi and Riemann. Some solutions of this problem still bear his name today. This work is summarized in Chandrasekhar [1967, 1977] (see Poincaré [1885, 1890, 1892, 1901a] for the original treatments). This background led to his famous paper, Poincaré [1901b], in which he laid out the basic equations of Euler type, including the rigid body, heavy top and fluids as special cases. Abstractly, these equations are determined once one is given a Lagrangian on a Lie algebra. It is because of the paper Poincaré [1901b] that the name ***Euler–Poincaré equations*** is now used for these equations. The work of Arnold [1966a] was very important for geometrizing and developing these ideas.

Euler equations provide perhaps the most basic examples of reduction, both Lagrangian and Hamiltonian. This aspect of reduction is developed in *IMS*, Chapters 13 and 14, but we shall be recalling some of the basic facts here.

To state the Euler–Poincaré equations, let $\mathfrak{g}$ be a given Lie algebra and let $l : \mathfrak{g} \to \mathbb{R}$ be a given function (a Lagrangian), let ξ be a point in $\mathfrak{g}$ and let $f \in \mathfrak{g}^*$ be given forces (whose nature we shall explicate later). Then the evolution of the variable ξ is determined by the Euler–Poincaré equations. Namely,

$$\frac{d}{dt}\frac{\delta l}{\delta \xi} = \mathrm{ad}^*_\xi \frac{\delta l}{\delta \xi} + f.$$

The notation is as follows: $\partial l/\partial \xi \in \mathfrak{g}^*$ (the dual vector space) is the derivative of l with respect to ξ; we use partial derivative notation because l is a function of the vector ξ and because shortly l will be a function of other variables as well. The map $\mathrm{ad}_\xi : \mathfrak{g} \to \mathfrak{g}$ is the linear map $\eta \mapsto [\xi, \eta]$, where $[\xi, \eta]$ denotes the Lie bracket of ξ and η, and where $\mathrm{ad}^*_\xi : \mathfrak{g}^* \to \mathfrak{g}^*$ is its dual (transpose) as a linear map. In the case that $f = 0$, we will call these equations the ***basic Euler–Poincaré equations***.

These equations are valid for either finite or infinite dimensional Lie algebras. For fluids, Poincaré was aware that one needs to use infinite dimensional Lie algebras, as is clear in his paper Poincaré [1910]. He was aware that one has to be careful with the signs in the equations; for example, for rigid body dynamics one uses the equations as they stand, but for fluids, one needs to be careful about the conventions for the Lie algebra operation ad_ξ; cf. Chetayev [1941].

To state the equations in the finite dimensional case in coordinates, one must choose a basis $e_1, \ldots, e_r$ of $\mathfrak{g}$ (so $\dim \mathfrak{g} = r$). Define, as usual, the structure constants C^d_{ab} of the Lie algebra by

$$[e_a, e_b] = \sum_{d=1}^{r} C^d_{ab} e_d, \tag{1.2.1}$$

where a, b run from 1 to r. If $\xi \in \mathfrak{g}$, its components relative to this basis are denoted ξ^a. If $e^1, \ldots, e^n$ is the corresponding dual basis, then the components of the differential of the Lagrangian l are the partial derivatives $\partial l / \partial \xi^a$. The Euler–Poincaré equations in this basis are

$$\frac{d}{dt} \frac{\partial l}{\partial \xi^b} = \sum_{a,d=1}^{r} C_{ab}^d \frac{\partial l}{\partial \xi^d} \xi^a + f_b. \tag{1.2.2}$$

For example, consider the Lie algebra $\mathbb{R}^3$ with the usual vector cross product. (Of course, this is the Lie algebra of the proper rotation group in $\mathbb{R}^3$.) For $l : \mathbb{R}^3 \to \mathbb{R}$, the Euler–Poincaré equations become

$$\frac{d}{dt} \frac{\partial l}{\partial \mathbf{\Omega}} = \frac{\partial l}{\partial \mathbf{\Omega}} \times \mathbf{\Omega} + \mathbf{f},$$

which generalize the Euler equations for rigid body motion.

These equations were written down for a certain class of Lagrangians l by Lagrange [1788, Volume 2, Equation A on p. 212], while it was Poincaré [1901b] who generalized them (without reference to the ungeometric Lagrange!) to an arbitrary Lie algebra. However, it was Lagrange who was grappeling with the derivation and deeper understanding of the nature of these equations. While Poincaré may have understood how to derive them from other principles, he did not reveal this.

Of course, there was a lot of mechanics going on in the decades leading up to Poincaré's work and we shall comment on some of it below. However, it is a curious historical fact that the Euler–Poincaré equations were not pursued extensively until quite recently. While many authors mentioned these equations and even tried to understand them more deeply (see, e.g., Hamel [1904, 1949] and Chetayev [1941]), it was not until the Arnold school that this understanding was at least partly achieved (see Arnold [1966a,c] and Arnold [1988]) and was used for diagnosing hydrodynamical stability (e.g., Arnold [1966b]).

It was already clear in the last century that certain mechanical systems resist the usual canonical formalism, either Hamiltonian or Lagrangian, outlined in the first paragraph. The rigid body provides an elementary example of this. In another example, to obtain a Hamiltonian description for ideal fluids, Clebsch [1857, 1859] found it necessary to introduce certain nonphysical potentials[1].

The Rigid Body. In the absence of external forces, the rigid body equations are usually written as follows:

$$\begin{aligned} I_1 \dot{\Omega}_1 &= (I_2 - I_3)\Omega_2 \Omega_3, \\ I_2 \dot{\Omega}_2 &= (I_3 - I_1)\Omega_3 \Omega_1, \\ I_3 \dot{\Omega}_3 &= (I_1 - I_2)\Omega_1 \Omega_2, \end{aligned} \tag{1.2.3}$$

where $\mathbf{\Omega} = (\Omega_1, \Omega_2, \Omega_3)$ is the body angular velocity vector and I_1, I_2, I_3 are the moments of inertia of the rigid body. Are these equations as written Lagrangian or Hamiltonian in any sense? Since there are an odd number of equations, they cannot be put in canonical Hamiltonian form.

One answer is to reformulate the equations on $T\mathrm{SO}(3)$ or $T^*\mathrm{SO}(3)$, as is classically done in terms of Euler angles and their velocities or conjugate momenta, relative to which the

[1] For modern accounts of Clebsch potentials and further references, see Holm and Kupershmidt [1983], Marsden and Weinstein [1983], Marsden, Ratiu, and Weinstein [1984a,b], Cendra and Marsden [1987], Cendra, Ibort, and Marsden [1987] and Goncharov and Pavlov [1997].

equations *are* in Euler–Lagrange or canonical Hamiltonian form. However, this reformulation answers a different question for a *six* dimensional system. We are interested in these structures for the equations as given above.

The Lagrangian answer is easy: these equations have Euler–Poincaré form on the Lie algebra $\mathbb{R}^3$ using the Lagrangian

$$l(\mathbf{\Omega}) = \frac{1}{2}(I_1\Omega_1^2 + I_2\Omega_2^2 + I_3\Omega_3^2). \tag{1.2.4}$$

which is the (rotational) kinetic energy of the rigid body.

One of our main messages is that the Euler–Poincaré equations possess a natural variational principle. In fact, the Euler rigid body equations are equivalent to the ***rigid body action principle***

$$\delta\mathfrak{S}_{\text{red}} = \delta\int_a^b l\,dt = 0, \tag{1.2.5}$$

where variations of $\mathbf{\Omega}$ are restricted to be of the form

$$\delta\mathbf{\Omega} = \dot{\mathbf{\Sigma}} + \mathbf{\Omega}\times\mathbf{\Sigma}, \tag{1.2.6}$$

in which $\mathbf{\Sigma}$ is a curve in $\mathbb{R}^3$ that vanishes at the endpoints. As before, we regard the ***reduced action*** $\mathfrak{S}_{\text{red}}$ as a function on the space of curves, but only consider variations of the form described. The equivalence of the rigid body equations and the rigid body action principle may be proved in the same way as one proves that Hamilton's principle is equivalent to the Euler–Lagrange equations: Since $l(\mathbf{\Omega}) = \frac{1}{2}\langle\mathbb{I}\mathbf{\Omega},\mathbf{\Omega}\rangle$, and $\mathbb{I}$ is symmetric, we obtain

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

upon integrating by parts and using the endpoint conditions, $\mathbf{\Sigma}(b) = \mathbf{\Sigma}(a) = 0$. Since $\mathbf{\Sigma}$ is otherwise arbitrary, (1.2.5) is equivalent to

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

which are Euler's equations.

Let us explain in concrete terms how to *derive* this variational principle from the *standard* variational principle of Hamilton.

We regard an element $\mathbf{R}\in\mathrm{SO}(3)$ giving the configuration of the body as a map of a reference configuration $\mathcal{B}\subset\mathbb{R}^3$ to the current configuration $\mathbf{R}(\mathcal{B})$; the map $\mathbf{R}$ takes a reference or label point $X\in\mathcal{B}$ to a current point $x = \mathbf{R}(X)\in\mathbf{R}(\mathcal{B})$. When the rigid body is in motion, the matrix $\mathbf{R}$ is time-dependent and the velocity of a point of the body is $\dot{x} = \dot{\mathbf{R}}X = \dot{\mathbf{R}}\mathbf{R}^{-1}x$. Since $\mathbf{R}$ is an orthogonal matrix, $\mathbf{R}^{-1}\dot{\mathbf{R}}$ and $\dot{\mathbf{R}}\mathbf{R}^{-1}$ are skew matrices, and so we can write

$$\dot{x} = \dot{\mathbf{R}}\mathbf{R}^{-1}x = \boldsymbol{\omega}\times x, \tag{1.2.7}$$

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

The corresponding body angular velocity is defined by

$$\boldsymbol{\Omega} = \mathbf{R}^{-1}\boldsymbol{\omega}, \tag{1.2.8}$$

so that $\boldsymbol{\Omega}$ is the angular velocity relative to a body fixed frame. Notice that

$$\begin{aligned} \mathbf{R}^{-1}\dot{\mathbf{R}}X &= \mathbf{R}^{-1}\dot{\mathbf{R}}\mathbf{R}^{-1}x = \mathbf{R}^{-1}(\boldsymbol{\omega} \times x) \\ &= \mathbf{R}^{-1}\boldsymbol{\omega} \times \mathbf{R}^{-1}x = \boldsymbol{\Omega} \times X, \end{aligned} \tag{1.2.9}$$

so that $\boldsymbol{\Omega}$ is given by *left* translation of $\dot{\mathbf{R}}$ to the identity. The kinetic energy is obtained by summing up $m\|\dot{x}\|^2/2$ (where $\|\cdot\|$ denotes the Euclidean norm) over the body:

$$K = \frac{1}{2}\int_{\mathcal{B}} \rho(X)\|\dot{\mathbf{R}}X\|^2\, d^3X, \tag{1.2.10}$$

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

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

K is a quadratic function of $\boldsymbol{\Omega}$. Writing

$$K = \frac{1}{2}\boldsymbol{\Omega}^T\mathbb{I}\boldsymbol{\Omega} \tag{1.2.11}$$

defines the ***moment of inertia tensor*** $\mathbb{I}$, which, provided the body does not degenerate to a line, is a positive-definite (3×3) matrix, or better, a quadratic form. This quadratic form can be diagonalized by a change of basis; thereby defining the principal axes and moments of inertia. In this basis, we write $\mathbb{I} = \mathrm{diag}(I_1, I_2, I_3)$. The function K is taken to be the Lagrangian of the system on $T\mathrm{SO}(3)$ (and by means of the Legendre transformation we obtain the corresponding Hamiltonian description on $T^*\mathrm{SO}(3)$). Notice that K in equation (1.2.10) is *left* (not right) invariant on $T\mathrm{SO}(3)$. It follows that the corresponding Hamiltonian is also *left* invariant.

In the Lagrangian framework, the relation between motion in $\mathbf{R}$ space and motion in body angular velocity (or $\boldsymbol{\Omega}$) space is as follows: The curve $\mathbf{R}(t) \in \mathrm{SO}(3)$ satisfies the Euler-Lagrange equations for

$$L(\mathbf{R}, \dot{\mathbf{R}}) = \frac{1}{2}\int_{\mathcal{B}} \rho(X)|\dot{\mathbf{R}}X|^2\, d^3X, \tag{1.2.12}$$

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

$$\mathbb{I}\dot{\boldsymbol{\Omega}} = \mathbb{I}\boldsymbol{\Omega} \times \boldsymbol{\Omega}. \tag{1.2.13}$$

An instructive proof of this relation involves understanding how to reduce variational principles using their symmetry groups. By Hamilton's principle, $\mathbf{R}(t)$ satisfies the Euler-Lagrange equations, if and only if

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

Let $l(\boldsymbol{\Omega}) = \frac{1}{2}(\mathbb{I}\boldsymbol{\Omega}) \cdot \boldsymbol{\Omega}$, so that $l(\boldsymbol{\Omega}) = L(\mathbf{R}, \dot{\mathbf{R}})$ if $\mathbf{R}$ and $\boldsymbol{\Omega}$ are related as above.

To see how we should transform Hamilton's principle, define the skew matrix $\hat{\boldsymbol{\Omega}}$ by $\hat{\boldsymbol{\Omega}}\mathbf{v} = \boldsymbol{\Omega} \times \mathbf{v}$ for any $\mathbf{v} \in \mathbb{R}^3$, and differentiate the relation $\mathbf{R}^{-1}\dot{\mathbf{R}} = \hat{\boldsymbol{\Omega}}$ with respect to $\mathbf{R}$ to get

$$-\mathbf{R}^{-1}(\delta\mathbf{R})\mathbf{R}^{-1}\dot{\mathbf{R}} + \mathbf{R}^{-1}(\delta\dot{\mathbf{R}}) = \widehat{\delta\boldsymbol{\Omega}}. \tag{1.2.14}$$

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

$$\hat{\boldsymbol{\Sigma}} = \mathbf{R}^{-1}\delta\mathbf{R}, \tag{1.2.15}$$

and define the vector $\boldsymbol{\Sigma}$ by

$$\hat{\boldsymbol{\Sigma}}\mathbf{v} = \boldsymbol{\Sigma} \times \mathbf{v}. \tag{1.2.16}$$

Note that

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

so

$$\mathbf{R}^{-1}\delta\dot{\mathbf{R}} = \dot{\hat{\boldsymbol{\Sigma}}} + \mathbf{R}^{-1}\dot{\mathbf{R}}\hat{\boldsymbol{\Sigma}}\,. \tag{1.2.17}$$

Substituting (1.2.17) and (1.2.15) into (1.2.14) gives

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

that is,

$$\widehat{\delta\boldsymbol{\Omega}} = \dot{\hat{\boldsymbol{\Sigma}}} + [\hat{\boldsymbol{\Omega}}, \hat{\boldsymbol{\Sigma}}]. \tag{1.2.18}$$

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

$$\delta\boldsymbol{\Omega} = \dot{\boldsymbol{\Sigma}} + \boldsymbol{\Omega} \times \boldsymbol{\Sigma}. \tag{1.2.19}$$

These calculations prove the following:

Theorem 1.2.1 *Hamilton's variational principle*

$$\delta\mathfrak{S} = \delta \int_a^b L\,dt = 0 \tag{1.2.20}$$

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

$$\delta\mathfrak{S}_{\mathrm{red}} = \delta \int_a^b l\,dt = 0 \tag{1.2.21}$$

on $\mathbb{R}^3$ *where the variations* $\delta\boldsymbol{\Omega}$ *are of the form (1.2.19) with* $\boldsymbol{\Sigma}(a) = \boldsymbol{\Sigma}(b) = 0$.

This sort of argument applies to any Lie group as we shall see shortly.

1.3 The Lie–Poisson Equations.

Hamiltonian Form of the Rigid Body Equations. If, instead of variational principles, we concentrate on Poisson brackets and drop the requirement that they be in the canonical form, then there is also a simple and beautiful Hamiltonian structure for the rigid body equations that is now well known[2]. To recall this, introduce the angular momenta

$$\Pi_i = I_i \Omega_i = \frac{\partial L}{\partial \Omega_i}, \quad i = 1, 2, 3, \tag{1.3.1}$$

so that the Euler equations become

$$\begin{aligned} \dot{\Pi}_1 &= \frac{I_2 - I_3}{I_2 I_3} \Pi_2 \Pi_3, \\ \dot{\Pi}_2 &= \frac{I_3 - I_1}{I_3 I_1} \Pi_3 \Pi_1, \\ \dot{\Pi}_3 &= \frac{I_1 - I_2}{I_1 I_2} \Pi_1 \Pi_2, \end{aligned} \tag{1.3.2}$$

that is,

$$\dot{\mathbf{\Pi}} = \mathbf{\Pi} \times \mathbf{\Omega}. \tag{1.3.3}$$

Introduce the following rigid body Poisson bracket on functions of the $\mathbf{\Pi}$'s:

$$\{F, G\}(\mathbf{\Pi}) = -\mathbf{\Pi} \cdot (\nabla_\Pi F \times \nabla_\Pi G) \tag{1.3.4}$$

and the Hamiltonian

$$H = \frac{1}{2}\left(\frac{\Pi_1^2}{I_1} + \frac{\Pi_2^2}{I_2} + \frac{\Pi_3^2}{I_3} \right). \tag{1.3.5}$$

One checks that Euler's equations are equivalent to $\dot{F} = \{F, H\}$.

The rigid body variational principle and the rigid body Poisson bracket are special cases of general constructions associated to any Lie algebra $\mathfrak{g}$. Since we have already described the general Euler–Poincaré construction on $\mathfrak{g}$, we turn next to the Hamiltonian counterpart on the dual space.

The Abstract Lie-Poisson Equations. Let F, G be real valued functions on the dual space $\mathfrak{g}^*$. Denoting elements of $\mathfrak{g}^*$ by μ, let the functional derivative of F at μ be the unique element $\delta F/\delta\mu$ of $\mathfrak{g}$ defined by

$$\lim_{\varepsilon \to 0} \frac{1}{\varepsilon}[F(\mu + \varepsilon\delta\mu) - F(\mu)] = \left\langle \delta\mu, \frac{\delta F}{\delta\mu} \right\rangle, \tag{1.3.6}$$

for all $\delta\mu \in \mathfrak{g}^*$, where $\langle \, , \rangle$ denotes the pairing between $\mathfrak{g}^*$ and $\mathfrak{g}$. Define the $(\pm)$ Lie-Poisson brackets by

$$\{F, G\}_\pm(\mu) = \pm \left\langle \mu, \left[\frac{\delta F}{\delta\mu}, \frac{\delta G}{\delta\mu} \right] \right\rangle. \tag{1.3.7}$$

Using the coordinate notation introduced above, the $(\pm)$ Lie-Poisson brackets become

$$\{F, G\}_\pm(\mu) = \pm \sum_{a,b,d=1}^{r} C_{ab}^d \mu_d \frac{\partial F}{\partial \mu_a} \frac{\partial G}{\partial \mu_b}, \tag{1.3.8}$$

[2]See *IMS* for details, references, and the history of this structure.

where $\mu = \sum_{d=1}^{r} \mu_d e^d$.

The Lie-Poisson equations, determined by $\dot{F} = \{F, H\}$ read

$$\dot{\mu}_a = \pm \sum_{b,d=1}^{r} C_{ab}^{d} \mu_d \frac{\partial H}{\partial \mu_b},$$

or intrinsically,

$$\dot{\mu} = \mp \operatorname{ad}^*_{\partial H/\partial \mu} \mu. \tag{1.3.9}$$

This setting of mechanics is a special case of the general theory of systems on Poisson manifolds, for which there is now an extensive theoretical development. (See Guillemin and Sternberg [1984] and Marsden and Ratiu [1998] for a start on this literature.) There is an especially important feature of the rigid body bracket that carries over to general Lie algebras, namely, *Lie-Poisson brackets arise from canonical brackets on the cotangent bundle* (phase space) T^*G associated with a Lie group G which has $\mathfrak{g}$ as its associated Lie algebra.

For a rigid body which is free to rotate about its center of mass, G is the (proper) rotation group $\mathrm{SO}(3)$. The choice of T^*G as the primitive phase space is made according to the classical procedures of mechanics described earlier. For the description using Lagrangian mechanics, one forms the velocity-phase space $T\mathrm{SO}(3)$. The Hamiltonian description on T^*G is then obtained by standard procedures.

The passage from T^*G to the space of $\mathbf{\Pi}$'s (body angular momentum space) is determined by *left* translation on the group. This mapping is an example of a *momentum map*; that is, a mapping whose components are the "Noether quantities" associated with a symmetry group. In this case, the momentum map in question is that associated with *right* translations of the group. Since the Hamiltonian is *left* invariant, this momentum map is not conserved. Indeed, it is the *spatial angular momentum* $\boldsymbol{\pi} = \mathbf{R}\mathbf{\Pi}$ that is conserved, not $\mathbf{\Pi}$.

The map from T^*G to $\mathfrak{g}^*$ being a Poisson (canonical) map *is a general fact about momentum maps*. The Hamiltonian point of view of all this is again a well developed subject.

Geodesic motion. As emphasized by Arnold [1966a], in many interesting cases, the Euler–Poincaré equations on a Lie algebra $\mathfrak{g}$ correspond to *geodesic motion* on the corresponding group G. We shall explain the relationship between the equations on $\mathfrak{g}$ and on G shortly, in theorem 1.6.1. Similarly, on the Hamiltonian side, the preceding paragraphs explained the relation between the Hamiltonian equations on T^*G and the Lie–Poisson equations on $\mathfrak{g}^*$. However, the issue of geodesic motion is simple: if the Lagrangian or Hamiltonian on $\mathfrak{g}$ or $\mathfrak{g}^*$ is purely quadratic, then the corresponding motion on the group is geodesic motion.

More History. The Lie-Poisson bracket was discovered by Sophus Lie (Lie [1890], Vol. II, p. 237). However, Lie's bracket and his related work was not given much attention until the work of Kirillov, Kostant, and Souriau (and others) revived it in the mid-1960s. Meanwhile, it was noticed by Pauli and Martin around 1950 that the rigid body equations are in Hamiltonian form using the rigid body bracket, but they were apparently unaware of the underlying Lie theory. It would seem that while Poincaré was aware of Lie theory, in his work on the Euler equations he was unaware of Lie's work on Lie-Poisson structures. He also seems not to have been aware of the variational structure of the Euler equations.

1.4 The Heavy Top.

Another system important to Poincaré and also for us later when we treat semidirect product reduction theory is the heavy top; that is, a rigid body with a fixed point in a gravitational field. For the Lie-Poisson description, the underlying Lie algebra, surprisingly, consists of the algebra of infinitesimal Euclidean motions in $\mathbb{R}^3$. These do *not* arise as actual Euclidean motions of the body since the body has a fixed point! As we shall see, there is a close parallel with the Poisson structure for compressible fluids.

The basic phase space we start with is again $T^*\mathrm{SO}(3)$. In this space, the equations are in canonical Hamiltonian form. Gravity breaks the symmetry and the system is no longer $\mathrm{SO}(3)$ invariant, so it cannot be written entirely in terms of the body angular momentum $\mathbf{\Pi}$. One also needs to keep track of $\mathbf{\Gamma}$, the "direction of gravity" as seen from the body ($\mathbf{\Gamma} = \mathbf{R}^{-1}\mathbf{k}$ where the unit vector $\mathbf{k}$ points upward and $\mathbf{R}$ is the element of $\mathrm{SO}(3)$ describing the current configuration of the body). The equations of motion are

$$\begin{aligned}
\dot{\Pi}_1 &= \frac{I_2 - I_3}{I_2 I_3}\Pi_2\Pi_3 + Mg\ell\,(\Gamma^2\chi^3 - \Gamma^3\chi^2),\\
\dot{\Pi}_2 &= \frac{I_3 - I_1}{I_3 I_1}\Pi_3\Pi_1 + Mg\ell\,(\Gamma^3\chi^1 - \Gamma^1\chi^3),\\
\dot{\Pi}_3 &= \frac{I_1 - I_2}{I_1 I_2}\Pi_1\Pi_2 + Mg\ell\,(\Gamma^1\chi^2 - \Gamma^2\chi^1),
\end{aligned} \tag{1.4.1}$$

or, in vector notation,

$$\dot{\mathbf{\Pi}} = \mathbf{\Pi} \times \mathbf{\Omega} + Mg\ell\,\mathbf{\Gamma} \times \boldsymbol{\chi}\,, \tag{1.4.2}$$

and

$$\dot{\mathbf{\Gamma}} = \mathbf{\Gamma} \times \mathbf{\Omega}, \tag{1.4.3}$$

where M is the body's mass, g is the acceleration of gravity, $\boldsymbol{\chi}$ is the unit vector on the line connecting the fixed point with the body's center of mass, and ℓ is the length of this segment.

The Lie algebra of the Euclidean group is $\mathfrak{se}(3) = \mathbb{R}^3 \times \mathbb{R}^3$ with the Lie bracket

$$[(\boldsymbol{\xi}, \mathbf{u}), (\boldsymbol{\eta}, \mathbf{v})] = (\boldsymbol{\xi} \times \boldsymbol{\eta}, \boldsymbol{\xi} \times \mathbf{v} - \boldsymbol{\eta} \times \mathbf{u}). \tag{1.4.4}$$

We identify the dual space with pairs $(\mathbf{\Pi}, \mathbf{\Gamma})$; the corresponding $(-)$ Lie-Poisson bracket called the ***heavy top bracket*** is

$$\begin{aligned}
\{F, G\}(\mathbf{\Pi}, \mathbf{\Gamma}) = -\mathbf{\Pi} \cdot (\nabla_\Pi F \times \nabla_\Pi G)&\\
- \mathbf{\Gamma} \cdot (\nabla_\Pi F \times \nabla_\Gamma G - \nabla_\Pi G \times \nabla_\Gamma F).&
\end{aligned} \tag{1.4.5}$$

The above equations for $\mathbf{\Pi}, \mathbf{\Gamma}$ can be checked to be equivalent to

$$\dot{F} = \{F, H\}, \tag{1.4.6}$$

where the ***heavy top Hamiltonian***

$$H(\mathbf{\Pi}, \mathbf{\Gamma}) = \frac{1}{2}\left(\frac{\Pi_1^2}{I_1} + \frac{\Pi_2^2}{I_2} + \frac{\Pi_3^2}{I_3}\right) + Mg\ell\,\mathbf{\Gamma} \cdot \boldsymbol{\chi} \tag{1.4.7}$$

is the total energy of the body (see, for example, Sudarshan and Mukunda [1974]).

The Lie algebra of the Euclidean group has a structure which is a special case of what is called a *semidirect product*. Here it is the product of the group of rotations with the translation group. It turns out that semidirect products occur under rather general circumstances when the symmetry in T^*G is broken. In particular, there are similarities in structure between the Poisson bracket for compressible flow and that for the heavy top. The general theory for semidirect products will be reviewed shortly.

A Kaluza-Klein form for the heavy top. We make a remark about the heavy top equations that is relevant for later purposes. Namely, since the equations have a Hamiltonian that is of the form kinetic plus potential, it is clear that the equations are *not of Lie-Poisson form on* $\mathfrak{so}(3)^*$*, the dual of the Lie algebra of* SO(3) and correspondingly, are not geodesic equations on SO(3). While the equations *are Lie–Poisson* on $\mathfrak{se}(3)^*$, the Hamiltonian is not quadratic, so again the equations are *not geodesic equations on* SE(3).

However, they can be viewed in a different way so that they become Lie-Poisson equations for a different group and with a *quadratic Hamiltonian.* In particular, they are the reduction of geodesic motion. To effect this, one changes the Lie algebra from $\mathfrak{se}(3)$ to the product $\mathfrak{se}(3) \times \mathfrak{so}(3)$. The dual variables are now denoted $\mathbf{\Pi}, \mathbf{\Gamma}, \boldsymbol{\chi}$. We regard the variable $\boldsymbol{\chi}$ as a momentum conjugate to a new variable, namely a *ghost* element of the rotation group in such a way that $\boldsymbol{\chi}$ is a constant of the motion; in Kaluza-Klein theory for charged particles one thinks of the charge this way, as being the momentum conjugate to a (ghost) cyclic variable.

We modify the Hamiltonian by replacing $\mathbf{\Gamma} \cdot \boldsymbol{\chi}$ by, for example, $\mathbf{\Gamma} \cdot \boldsymbol{\chi} + \|\mathbf{\Gamma}\|^2 + \|\boldsymbol{\chi}\|^2$, or any other terms of this sort that convert the potential energy into a positive definite quadratic form in $\mathbf{\Gamma}$ and $\boldsymbol{\chi}$. The added terms, being Casimir functions, do not affect the equations of motion. However, now the Hamiltonian is purely quadratic and hence comes from geodesic motion on the group $SE(3) \times \mathrm{SO}(3)$. Notice that this construction is quite different from that of the well known Jacobi metric method.

Later on in our study of continuum mechanics, we shall repeat this construction to achieve geodesic form for some other interesting continuum models. Of course one can also treat a heavy top that is charged or has a magnetic moment using these ideas.

1.5 Incompressible Fluids.

Arnold [1966a] showed that the Euler equations for an incompressible fluid could be given a Lagrangian and Hamiltonian description similar to that for the rigid body. His approach[3] has the appealing feature that one sets things up just the way Lagrange and Hamilton would have done: one begins with a configuration space Q, forms a Lagrangian L on the velocity phase space TQ and then Legendre transforms to a Hamiltonian H on the momentum phase space T^*Q. Thus, one automatically has variational principles, etc. For ideal fluids, $Q = G$ is the group $\mathrm{Diff}_{\mathrm{vol}}(\mathcal{D})$ of volume preserving transformations of the fluid container (a region $\mathcal{D}$ in $\mathbb{R}^2$ or $\mathbb{R}^3$, or a Riemannian manifold in general, possibly with boundary). Group multiplication in G is composition.

The reason we select $G = \mathrm{Diff}_{\mathrm{vol}}(\mathcal{D})$ as the configuration space is similar to that for the rigid body; namely, each φ in G is a mapping of $\mathcal{D}$ to $\mathcal{D}$ which takes a reference point $X \in \mathcal{D}$ to a current point $x = \varphi(X) \in \mathcal{D}$; thus, knowing φ tells us where each particle of fluid goes and hence gives us the current ***fluid configuration***. We ask that φ be a diffeomorphism to exclude discontinuities, cavitation, and fluid interpenetration, and we ask that φ be volume preserving to correspond to the assumption of incompressibility.

A ***motion*** of a fluid is a family of time-dependent elements of G, which we write as $x = \varphi(X, t)$. The ***material velocity*** field is defined by $\mathbf{V}(X, t) = \partial\varphi(X, t)/\partial t$, and the ***spatial velocity*** field is defined by $\mathbf{v}(x, t) = \mathbf{V}(X, t)$ where x and X are related by $x = \varphi(X, t)$. If we suppress "t" and write $\dot{\varphi}$ for $\mathbf{V}$, note that

$$\mathbf{v} = \dot{\varphi} \circ \varphi^{-1} \quad \text{i.e.,} \quad \mathbf{v}_t = \mathbf{V}_t \circ \varphi_t^{-1}, \tag{1.5.1}$$

[3] Arnold's approach is consistent with what appears in the thesis of Ehrenfest from around 1904; see Klein [1970]. However, Ehrenfest bases his principles on the more sophisticated curvature principles of Gauss and Hertz.

where $\varphi_t(x) = \varphi(X, t)$. We can regard (1.5.1) as a map from the space of $(\varphi, \dot{\varphi})$ (material or Lagrangian description) to the space of $\mathbf{v}$'s (spatial or Eulerian description). Like the rigid body, the material to spatial map (1.5.1) takes the canonical bracket to a Lie-Poisson bracket; one of our goals is to understand this reduction. Notice that if we replace φ by $\varphi \circ \eta$ for a fixed (time-independent) $\eta \in \mathrm{Diff}_{\mathrm{vol}}(\mathcal{D})$, then $\dot{\varphi} \circ \varphi^{-1}$ is independent of η; this reflects the *right* invariance of the Eulerian description ($\mathbf{v}$ is invariant under composition of φ by η on the right). This is also called the ***particle relabeling symmetry*** of fluid dynamics. The spaces TG and T^*G represent the Lagrangian (material) description and we pass to the Eulerian (spatial) description by right translations and use the $(+)$ Lie-Poisson bracket. One of the things we shall explain later is the reason for the switch between right and left in going from the rigid body to fluids.

The ***Euler equations*** for an ideal, incompressible, homogeneous fluid moving in the region $\mathcal{D}$ are

$$\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \cdot \nabla)\mathbf{v} = -\nabla p \tag{1.5.2}$$

with the constraint $\operatorname{div} \mathbf{v} = 0$ and boundary conditions: $\mathbf{v}$ is tangent to $\partial\mathcal{D}$.

The pressure p is determined implicitly by the divergence-free (volume preserving) constraint $\operatorname{div} \mathbf{v} = 0$. The associated Lie algebra $\mathfrak{g}$ is the space of all divergence-free vector fields tangent to the boundary. This Lie algebra is endowed with the *negative* ***Jacobi-Lie bracket*** of vector fields given by

$$[\mathbf{v}, \mathbf{w}]_L^i = \sum_{j=1}^{n} \left(w^j \frac{\partial v^i}{\partial x^j} - v^j \frac{\partial w^i}{\partial x^j} \right). \tag{1.5.3}$$

(The subscript L on $[\cdot\,, \cdot]$ refers to the fact that it is the *left* Lie algebra bracket on $\mathfrak{g}$. The most common convention for the Jacobi-Lie bracket of vector fields, also the one we adopt, has the opposite sign.) We identify $\mathfrak{g}$ and $\mathfrak{g}^*$ by using the pairing

$$\langle \mathbf{v}, \mathbf{w} \rangle = \int_{\mathcal{D}} \mathbf{v} \cdot \mathbf{w} \, d^3x. \tag{1.5.4}$$

Hamiltonian structure for fluids. Introduce the $(+)$ Lie-Poisson bracket, called the ***ideal fluid bracket***, on functions of $\mathbf{v}$ by

$$\{F, G\}(\mathbf{v}) = \int_{\mathcal{D}} \mathbf{v} \cdot \left[\frac{\delta F}{\delta \mathbf{v}}, \frac{\delta G}{\delta \mathbf{v}} \right]_L d^3x, \tag{1.5.5}$$

where $\delta F / \delta \mathbf{v}$ is defined by

$$\lim_{\varepsilon \to 0} \frac{1}{\varepsilon} [F(\mathbf{v} + \varepsilon \delta \mathbf{v}) - F(\mathbf{v})] = \int_{\mathcal{D}} \left(\delta \mathbf{v} \cdot \frac{\delta F}{\delta \mathbf{v}} \right) d^3x. \tag{1.5.6}$$

With the energy function chosen to be the kinetic energy,

$$H(\mathbf{v}) = \frac{1}{2} \int_{\mathcal{D}} |\mathbf{v}|^2 \, d^3x, \tag{1.5.7}$$

one can verify that the Euler equations (1.5.2) are equivalent to the Poisson bracket equations

$$\dot{F} = \{F, H\} \tag{1.5.8}$$

for all functions F on $\mathfrak{g}^*$. For this, one uses the orthogonal decomposition $\mathbf{w} = \mathbb{P}\mathbf{w} + \nabla p$ of a vector field $\mathbf{w}$ into a divergence-free part $\mathbb{P}\mathbf{w}$ in $\mathfrak{g}$ and a gradient. The Euler equations can be written as

$$\frac{\partial \mathbf{v}}{\partial t} + \mathbb{P}(\mathbf{v} \cdot \nabla \mathbf{v}) = 0. \tag{1.5.9}$$

One can also express the Hamiltonian structure in terms of the vorticity as a basic dynamic variable and show that the preservation of coadjoint orbits amounts to Kelvin's circulation theorem. Marsden and Weinstein [1983] show that the Hamiltonian structure in terms of Clebsch potentials fits naturally into this Lie-Poisson scheme, and that Kirchhoff's Hamiltonian description of point vortex dynamics, vortex filaments, and vortex patches can be derived in a natural way from the Hamiltonian structure described above.

Lagrangian structure for fluids. The general framework of the Euler–Poincaré and the Lie-Poisson equations gives other insights as well. For example, this general theory shows that the Euler equations are derivable from the "variational principle"

$$\delta \int_a^b \int_{\mathcal{D}} \frac{1}{2} \|\mathbf{v}\|^2 \, d^3x = 0$$

which should hold for all variations $\delta\mathbf{v}$ of the form

$$\delta \mathbf{v} = \dot{\mathbf{u}} + [\mathbf{u}, \mathbf{v}]_L$$

where $\mathbf{u}$ is a vector field (representing the infinitesimal particle displacement) vanishing at the temporal endpoints. The constraints on the allowed variations of the fluid velocity field are commonly known as "Lin constraints" and their nature was clarified by Newcomb [1962] and Bretherton [1970]. This itself has an interesting history, going back to Ehrenfest, Boltzmann, and Clebsch, but again, there was little if any contact with the heritage of Lie and Poincaré on the subject.

1.6 The Basic Euler–Poincaré Equations.

We now recall the abstract derivation of the "basic" Euler–Poincaré equations (i.e., the Euler–Poincaré equations with no forcing or advected parameters) for left–invariant Lagrangians on Lie groups (see Marsden and Scheurle [1993a,b], Marsden and Ratiu [1998] and Bloch et al. [1996]).

Theorem 1.6.1 *Let G be a Lie group and $L : TG \to \mathbb{R}$ a left (respectively, right) invariant Lagrangian. Let $l : \mathfrak{g} \to \mathbb{R}$ be its restriction to the tangent space at the identity. For a curve $g(t) \in G$, let $\xi(t) = g(t)^{-1}\dot{g}(t)$; i.e., $\xi(t) = T_{g(t)}L_{g(t)^{-1}}\dot{g}(t)$ (respectively, $\xi(t) = \dot{g}(t)g(t)^{-1}$). Then the following are equivalent:*

i *Hamilton's principle*

$$\delta \int_a^b L(g(t), \dot{g}(t)) dt = 0 \tag{1.6.1}$$

holds, as usual, for variations $\delta g(t)$ of $g(t)$ vanishing at the endpoints.

ii *The curve $g(t)$ satisfies the Euler-Lagrange equations for L on G.*

iii *The "variational" principle*

$$\delta \int_a^b l(\xi(t))dt = 0 \tag{1.6.2}$$

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

$$\delta\xi = \dot{\eta} \pm [\xi, \eta], \tag{1.6.3}$$

where η vanishes at the endpoints (+ corresponds to left invariance and $-$ to right invariance).[4]

iv *The* ***basic Euler–Poincaré equations*** *hold*

$$\frac{d}{dt}\frac{\delta l}{\delta \xi} = \pm \operatorname{ad}^*_\xi \frac{\delta l}{\delta \xi}\,. \tag{1.6.4}$$

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

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

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

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

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

To complete the proof, we show the equivalence of **iii** and **iv** in the left-invariant case. Indeed, using the definitions and integrating by parts produces,

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

so the result follows.

There is of course a right invariant version of this theorem in which $\xi = \dot{g}g^{-1}$ and the Euler–Poincaré equations acquire appropriate minus signs as in equation (1.6.4). We shall go into this in detail later.

1.7 Lie–Poisson Reduction.

We now recall from *IMS* some of the key ideas about Lie–Poisson reduction.

Besides the Poisson structure on a symplectic manifold, the Lie–Poisson bracket on $\mathfrak{g}^*$, the dual of a Lie algebra, is perhaps the most fundamental example of a Poisson structure.

[4] Because there are constraints on the variations, this principle is more like a Lagrange d'Alembert principle, which is why we put "variational" in quotes. As we shall explain, such problems are not literally variational.

If P is a Poisson manifold and G acts on it freely and properly, then P/G is also Poisson in a natural way: identify functions on P/G with G-invariant functions on P and use this to induce a bracket on functions on P/G. In the case $P = T^*G$ and G acts on the left by cotangent lift, then $T^*G/G \cong \mathfrak{g}^*$ inherits a Poisson structure. The Lie–Poisson bracket gives an explicit formula for this bracket.

Given two smooth functions F, H on $(\mathfrak{g}^*)$, we extend them to functions, F_L, H_L (respectively, F_R, H_R) on all T^*G by left (respectively, right) translations. The bracket $\{F_L, H_L\}$ (respectively, $\{F_R, H_R\}$) is taken in the canonical symplectic structure Ω on T^*G. The result is then restricted to $\mathfrak{g}^*$ regarded as the cotangent space at the identity; this defines $\{F, H\}$. We shall prove that one gets the Lie–Poisson bracket this way. In *IMS*, Chapter 14, it is shown that the symplectic leaves of this bracket are the coadjoint orbits in $\mathfrak{g}^*$.

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

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

Theorem 1.7.1 (Lie–Poisson Reduction Theorem) *Identifying the set of functions on $\mathfrak{g}^*$ with the set of left (respectively, right) invariant functions on T^*G endows $\mathfrak{g}^*$ with Poisson structures given by*

$$\{F, H\}_{\pm}(\mu) = \pm \left\langle \mu, \left[\frac{\delta F}{\delta \mu}, \frac{\delta H}{\delta \mu} \right] \right\rangle. \tag{1.7.1}$$

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

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

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

$$\xi_L(g) = T_eL_g \cdot \xi \tag{1.7.2}$$

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

$$[\xi, \eta] = [\xi_L, \eta_L](e), \tag{1.7.3}$$

where the bracket on the right-hand side is the Jacobi–Lie bracket on vector fields. The bracket (1.7.3) makes $\mathfrak{g}$ into a Lie algebra, that is, $[\,,]$ is bilinear, antisymmetric, and satisfies Jacobi's identity. For example, if G is a subgroup of $\mathrm{GL}(n)$, the group of invertible $n \times n$ matrices, we identify $\mathfrak{g} = T_eG$ with a vector space of matrices and then as we calculated in *IMS*, Chapter 9,

$$[\xi, \eta] = \xi\eta - \eta\xi, \tag{1.7.4}$$

the usual commutator of matrices.

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

$$F_L \circ T^*L_g = F_L, \tag{1.7.5}$$

where T^*L_g denotes the cotangent lift of L_g, so T^*L_g is the pointwise adjoint of TL_g. Given $F : \mathfrak{g}^* \to \mathbb{R}$ and $\alpha_g \in T^*G$, set

$$F_L(\alpha_g) = F(T_e^*L_g \cdot \alpha_g) \tag{1.7.6}$$

which is the ***left invariant extension*** of F from $\mathfrak{g}^*$ to T^*G. One similarly defines the ***right invariant extension*** by

$$F_R(\alpha_g) = F(T_e^*R_g \cdot \alpha_g). \tag{1.7.7}$$

The main content of the Lie–Poisson reduction theorem is the pair of formulae

$$\{F, H\}_- = \{F_L, H_L\} \,|\mathfrak{g}^* \tag{1.7.8}$$

and

$$\{F, H\}_+ = \{F_R, H_R\} \,|\mathfrak{g}^*, \tag{1.7.9}$$

where $\{\,,\}_\pm$ is the Lie–Poisson bracket on $\mathfrak{g}^*$ and $\{\,,\}$ is the canonical bracket on T^*G. Another way of saying this is that the *map* $\lambda : T^*G \to \mathfrak{g}^*_-$ (respectively, $\rho : T^*G \to \mathfrak{g}^*_+$ on T^*G) given by

$$\alpha_g \mapsto T_e^*L_g \cdot \alpha_g \text{ (respectively, } T_e^*R_g \cdot \alpha_g) \tag{1.7.10}$$

is a Poisson map.

Note that the correspondence between ξ and ξ_L identifies $\mathcal{F}(\mathfrak{g}^*)$ with the left invariant functions on T^*G, which is a subalgebra of $\mathcal{F}(T^*G)$ (since lifts are canonical), so (1.7.1) indeed defines a Poisson structure (although this fact may also be readily verified directly).

To prove the Lie–Poisson reduction theorem, first prove the following.

Lemma 1.7.2 *Let G act on itself by left translations. Then*

$$\xi_G(g) = T_eR_g \cdot \xi. \tag{1.7.11}$$

Proof. By definition of infinitesimal generator,

$$\begin{aligned}\xi_G(g) &= \frac{d}{dt}\Phi_{\exp(t\xi)}(g)\bigg|_{t=0} \\ &= \frac{d}{dt}R_g(\exp(t\xi))\bigg|_{t=0} \\ &= T_eR_g \cdot \xi\end{aligned}$$

by the chain rule. ▼

Proof of the Theorem. Let $J_L : T^*G \to \mathfrak{g}^*$ be the momentum map for the left action. From the formula for the momentum map for a cotangent lift (*IMS*, Chapter 12, we have

$$\begin{aligned}\langle J_L(\alpha_g), \xi\rangle &= \langle \alpha_g, \xi_G(g)\rangle \\ &= \langle \alpha_g, T_eR_g \cdot \xi\rangle \\ &= \langle T_e^*R_g \cdot \alpha_g, \xi\rangle\,.\end{aligned}$$

Thus,

$$J_L(\alpha_g) = T_e^* R_g \cdot \alpha_g,$$

so $J_L = \rho$. Similarly $J_R = \lambda$. However, the momentum maps J_L and J_R are equivariant being the momentum maps for cotangent lifts, and so from *IMS* §12.5, they are Poisson maps. The theorem now follows.

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

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

Note that

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

and so it is now clear that the Lie-Poisson equations (1.3.9) and the Euler–Poincaré equations (1.6.4) are equivalent.

Lie-Poisson Systems on Semidirect Products. As we described above, the heavy top is a basic example of a Lie-Poisson Hamiltonian system defined on the dual of a semidirect product Lie algebra. The *general* study of Lie-Poisson equations for systems on the dual of a semidirect product Lie algebra grew out of the work of many authors including Sudarshan and Mukunda [1974], Vinogradov and Kupershmidt [1977], Ratiu [1980], Guillemin and Sternberg [1980], Ratiu [1981, 1982], Marsden [1982], Marsden, Weinstein, Ratiu, Schmidt and Spencer [1983], Holm and Kupershmidt [1983], Kupershmidt and Ratiu [1983], Holmes and Marsden [1983], Marsden, Ratiu and Weinstein [1984a,b], Guillemin and Sternberg [1984], Holm, Marsden, Ratiu and Weinstein [1985], Abarbanel, Holm, Marsden, and Ratiu [1986] and Marsden, Misiolek, Perlmutter and Ratiu [1997]. As these and related references show, the Lie-Poisson equations apply to a wide variety of systems such as the heavy top, compressible flow, stratified incompressible flow, and MHD (magnetohydrodynamics).

In each of the above examples as well as in the general theory, one can view the given Hamiltonian in the material representation as one that depends on a parameter; this parameter becomes dynamic when reduction is performed; this reduction amounts in many examples to expressing the system in the spatial representation.

Rigid Body in a Fluid. The dynamics of a rigid body in a fluid are often modeled by the classical Kirchhoff equations in which the fluid is assumed to be potential flow, responding to the motion of the body. (For underwater vehicle dynamics we will need to include buoyancy effects.)[5] Here we choose $G = \mathrm{SE}(3)$, the group of Euclidean motions of $\mathbb{R}^3$ and the Lagrangian is the total energy of the body-fluid system. Recall that the Lie algebra of SE(3) is $\mathfrak{se}(3) = \mathbb{R}^3 \times \mathbb{R}^3$ with the bracket

$$[(\Omega, u), (\Sigma, v)] = (\Omega \times \Sigma, \Omega \times v - \Sigma \times u).$$

[5] This model may be viewed inside the larger model of an elastic-fluid interacting system with the constraint of rigidity imposed on the elastic body and with the reduced space for the fluid variables (potential flow is simply reduction at zero for fluids).

The reduced Lagrangian is again quadratic, so has the form

$$l(\Omega, v) = \frac{1}{2}\Omega^T J\Omega + \Omega^T Dv + \frac{1}{2}v^T Mv.$$

The Lie–Poisson equations are computed to be

$$\left.\begin{array}{l} \dot{\Pi} = \Pi \times \Omega + P \times v \\ \dot{P} = P \times \Omega \end{array}\right\}$$

where $\Pi = \partial l/\partial\Omega = J\Omega + Dv$ the "angular momentum" and $P = \partial l/\partial v = Mv + D^T\Omega$, the "linear momentum".

Again, we suggest that the reader work out the reduced variational principle. Relevant references are Lamb [1932], Leonard [1996], Leonard and Marsden [1997], and Holmes, Jenkins and Leonard [1997].

KdV Equation. Following Ovsienko and Khesin [1987], we will now indicate how the KdV equations may be recast as Euler-Poincaré equations. The ***KdV equation*** is the following equation for a scalar function $u(x,t)$ of the real variables x and t:

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

We let $\mathfrak{g}$ be the Lie algebra of vector fields u on the circle (of length 1) with the standard bracket

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

Let the ***Gelfand-Fuchs cocycle*** be defined by[6]

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

where γ is a constant. Let the ***Virasoro Lie algebra*** be defined by $\tilde{\mathfrak{g}} = \mathfrak{g} \times \mathbb{R}$ with the Lie bracket

$$[(u, a), (v, b)] = ([u, v], \gamma\Sigma(u, v)).$$

This is verified to be a Lie algebra; the corresponding group is called the ***Bott-Virasoro group***. Let

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

Then one checks that the Euler-Poincaré equations are

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

so that for appropriate a and γ and rescaling, we get the KdV equation. Thus, *the KdV equations may be regarded as geodesics on the Bott-Virasoro group.*

Likewise, the Camassa-Holm equation can be recast as geodesics using the H^1 rather than the L^2 metric (see Misiolek [1997] and Holm, Kouranbaeva, Marsden, Ratiu and Shkoller [1998]).

[6] An interesting interpretation of the Gelfand-Fuchs cocycle as the curvature of a mechanical connection is given in Marsden, Misiolek, Perlmutter and Ratiu [1998a,b].

Lie–Poisson Reduction of Dynamics. If H is left G-invariant on T^*G and X_H is its Hamiltonian vector field (recall from *IMS* that it is determined by $\dot{F} = \{F, H\}$), then X_H projects to the Hamiltonian vector field X_h determined by $\dot{f} = \{f, h\}_-$ where $h = H|T_e^*G = H|\mathfrak{g}^*$. We call $\dot{f} = \{f, h\}_-$ the ***Lie-Poisson equations***.

As we have mentioned, if l is regular; *i.e.,* $\xi \mapsto \mu = \partial l/\partial \xi$ is invertible, then the Legendre transformation taking ξ to μ and l to

$$h(\mu) = \langle \xi, \mu \rangle - l(\xi)$$

maps the Euler-Poincaré equations to the Lie-Poisson equations and vice-versa.

The heavy top is an example of a Lie-Poisson system on $\mathfrak{se}(3)^*$. However, its inverse Legendre transformation (using the standard h) is degenerate! This is an indication that something is missing on the Lagrangian side and this is indeed the case. The resolution is found in Holm, Marsden and Ratiu [1998a].

Lie-Poisson systems have a remarkable property: they leave the coadjoint orbits in $\mathfrak{g}^*$ invariant. In fact the coadjoint orbits are the symplectic leaves of $\mathfrak{g}^*$. For each of examples 1 and 3, the reader may check directly that the equations are Lie-Poisson and that the coadjoint orbits are preserved. For example 2, the preservation of coadjoint orbits is essentially Kelvin's circulation theorem. See Marsden and Weinstein [1983] for details. For the rotation group, the coadjoint orbits are the familiar body angular momentum spheres, shown in figure 1.7.1.

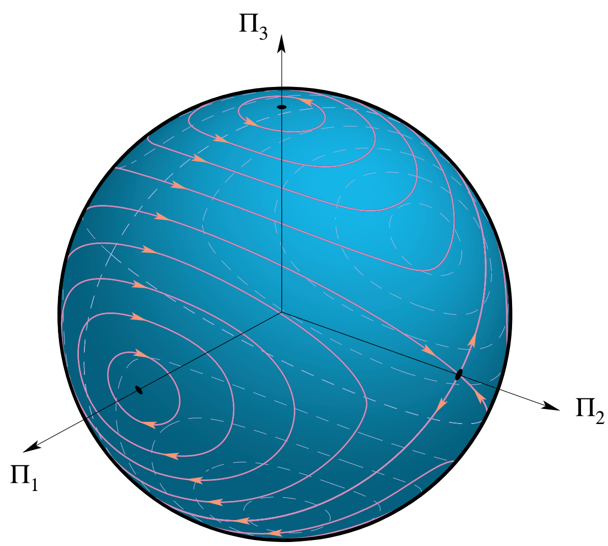

Figure 1.7.1: The rigid body momentum sphere.

History and literature. Lie-Poisson brackets were known to Lie around 1890, but apparently this aspect of the theory was not picked up by Poincaré. The coadjoint orbit symplectic structure was discovered by Kirillov, Kostant and Souriau in the 1960's. They were shown to be symplectic reduced spaces by Marsden and Weinstein [1974]. It is not clear who first observed *explicitly* that $\mathfrak{g}^*$ inherits the Lie-Poisson structure by reduction as in the preceding Lie-Poisson reduction theorem. It is *implicit* in many works such as Lie [1890], Kirillov [1962], Guillemin and Sternberg [1980] and Marsden and Weinstein [1982, 1983], but is *explicit* in Holmes and Marsden [1983] and Marsden, Weinstein, Ratiu, Schmid and Spencer [1983].

1.8 Symplectic and Poisson Reduction.

The ways in which reduction has been generalized and applied has been nothing short of phenomenal. We now sketch just a few of the highlights (eliminating many important references). *We shall be coming back to develop many of these ideas in detail in the text. What follows is an overview that can be returned to later on.*

First of all, in an effort to synthesize coadjoint orbit reduction (suggested by work of Arnold, Kirillov, Kostant and Souriau) with techniques for the reduction of cotangent bundles by Abelian groups of Smale [1970], Marsden and Weinstein [1974] developed symplectic reduction; related results, but with a different motivation and construction were found by Meyer [1973]. The construction is now well known: let (P, Ω) be a symplectic manifold and $J : P \to \mathfrak{g}^*$ be an equivariant momentum map; then avoiding singularities, $J^{-1}(\mu)/G_\mu = P_\mu$ is a symplectic manifold in a natural way. For example, for $P = T^*G$, one gets coadjoint orbits. We shall develop the theory of symplectic reduction in Chapter 2.

Kazhdan, Kostant and Sternberg [1978] showed how P_μ can be realized in terms of orbit reduction $P_\mu \cong J^{-1}(\mathcal{O})/G$ and from this it follows (but not in a totally obvious way) that P_μ are the symplectic leaves in P/G. This paper was also one of the first to notice deep links between reduction and integrable systems, a subject continued by, for example, Bobenko, Reyman and Semenov-Tian-Shansky [1989].

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, for example, Pedroni [1995] and Vanhaecke [1996]).

The mechanical connection. A basic construction implicit in Smale [1970], Abraham and Marsden [1978] and explicit in Kummer [1981] is the notion of the mechanical connection. The geometry of this situation was used to great effect in Guichardet [1984] and Iwai [1987, 1990].

Assume Q is Riemannian (the metric often being the kinetic energy metric) and that G acts on Q freely by isometries, so $\pi : Q \to Q/G$ is a principal bundle. If we declare the horizontal spaces to be metric orthogonal to the group orbits, this uniquely defines a connection called the ***mechanical connection***. There are explicit formulas for it in terms of the locked inertia tensor; see for instance, Marsden [1992] for details. The space Q/G is called ***shape space*** and plays a critical role in the theory.[7]

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

The Abelian version of cotangent bundle reduction was developed by Smale [1970] and Satzer [1975] and was generalized to the nonabelian case in Abraham and Marsden [1978]. It was Kummer [1981] who introduced the interpretations of these results in terms of the mechanical connection.

[7]Shape space and its geometry plays a key role in computer vision. See for example, Le and Kendall [1993].

The Lagrangian analogue of cotangent bundle reduction is called ***Routh reduction*** and was developed by Marsden and Scheurle [1993a,b]. Routh, around 1860 investigated what we would call today the Abelian version.

The "bundle picture" begun by the developments of the cotangent bundle reduction theory was significantly developed by Montgomery, Marsden and Ratiu [1984] and Montgomery [1986] motivated by work of Weinstein and Sternberg on Wong's equations (the equations for a particle moving in a Yang-Mills field).

This bundle picture can be viewed as follows. Choosing a connection, such as the mechanical connection, on $Q \to Q/G$, one gets a natural isomorphism

$$T^*Q/G \cong T^*(Q/G) \oplus \tilde{\mathfrak{g}}^*$$

where the sum is a Whitney sum of vector bundles over Q/G (fiberwise a direct sum) and $\tilde{\mathfrak{g}}^*$ is the associated vector bundle to the co-adjoint action of G on $\mathfrak{g}^*$. The description of the Poisson structure on this bundle (a synthesis of the canonical bracket, the Lie-Poisson bracket and curvature) may be found in Cendra, Marsden and Ratiu [1998].

Lagrangian reduction. The Lagrangian analogue of the bundle picture is the dual isomorphism

$$TQ/G \cong T(Q/G) \oplus \tilde{\mathfrak{g}}$$

whose geometry is developed in Cendra, Marsden and Ratiu [1998]. In particular, the equations and variational principles are developed on this space. For $Q = G$ this reduces to the Euler-Poincaré picture we had previously. For G abelian, it reduces to the Routh procedure.

If we have an invariant Lagrangian on TQ it induces a Lagrangian l on $(TQ)/G$ and hence on $T(Q/G) \oplus \tilde{\mathfrak{g}}$. Calling the variables $r^\alpha, \dot{r}^\alpha$ and Ω^α, the resulting ***reduced Euler-Lagrange*** equations (implicitly contained in Cendra, Ibort and Marsden [1987] and explicitly in Marsden and Scheurle [1993b]) are

$$\begin{aligned} \frac{d}{dt}\frac{\partial l}{\partial \dot{r}^\alpha} - \frac{\partial l}{\partial r^\alpha} &= \frac{\partial l}{\partial \Omega^\alpha}(-B^\alpha_{\alpha\beta}\dot{r}^\beta + \xi^a_{\alpha d}\Omega^d) \\ \frac{d}{dt}\frac{\partial l}{\partial \Omega^b} &= \frac{\partial l}{\partial \Omega^a}(-\xi^a_{\alpha\beta}\dot{r}^\alpha + C^a_{db}\Omega^d) \end{aligned}$$

where $B^a_{\alpha\beta}$ is the curvature of the connection $\mathcal{A}^b_\alpha$, C^a_{bd} are the structure constants of the Lie algebra $\mathfrak{g}$ and where $\xi^a_{\alpha d} = C^a_{bd}\mathcal{A}^b_\alpha$.

Using the geometry of the bundle $TQ/G = T(Q/G) \oplus \tilde{\mathfrak{g}}$, one obtains a nice interpretation of these equations in terms of covariant derivatives. One easily gets the dynamics of particles in a Yang-Mills field (these are called Wong's equations) as a special case; see Cendra, Holm, Marsden and Ratiu [1998] for this example. Methods of Lagrangian reduction and the Wong equations have proven very useful in optimal control problems. It was used in Koon and Marsden [1997] to extend the falling cat theorem of Montgomery [1990] to the case of nonholonomic systems.

Cotangent bundle reduction is very interesting for group extensions, such as the Bott-Virasoro group described earlier, where the Gelfand-Fuchs cocycle may be interpreted as the curvature of a mechanical connection. This is closely related to work of Marsden, Misiolek, Perlmutter and Ratiu [1998a,b] on reduction by stages. This work in turn is an outgrowth of earlier work of Guillemin and Sternberg [1980], Marsden, Ratiu and Weinstein [1984a,b] and many others on systems such as the heavy top, compressible flow and MHD. It also applies to underwater vehicle dynamics as shown in Leonard [1997] and Leonard and Marsden [1997].

Semidirect Product Reduction. In semidirect product reduction, one supposes that G acts on a vector space V (and hence on its dual V^*). From G and V we form the ***semidirect product*** Lie group $S = G \circledS V$, the set $G \times V$ with multiplication

$$(g_1, v_1) \cdot (g_2, v_2) = (g_1 g_2, v_1 + g_1 v_2).$$

The Euclidean group $SE(3) = SO(3) \circledS \mathbb{R}^3$, the semidirect product of rotations and translations is a basic example. Now suppose we have a Hamiltonian on T^*G that is invariant under the isotropy group G_{a_0} for $a_0 \in V^*$. The ***semidirect product reduction theorem*** states that *reduction of T^*G by G_{a_0} gives reduced spaces that are symplectically diffeomorphic to coadjoint orbits in the dual of the Lie algebra of the semi-direct product:* $(\mathfrak{g} \circledS V)^*$.

This is a very important construction in applications where one has "advected quantities" (such as density in compressible flow). Its Lagrangian counterpart, which is *not* simply the Euler-Poincaré equations on $\mathfrak{g} \circledS V$, is developed in Holm, Marsden and Ratiu [1998a] along with applications to continuum mechanics. Cendra, Holm, Hoyle and Marsden [1998] have applied this idea to the Maxwell-Vlasov equations of plasma physics.

If one reduces the semidirect product group $S = G \circledS V$ in two stages, first by V and then by G, one recovers the semidirect product reduction theorem mentioned above.

A far reaching generalization of this semidirect product theory is given in Marsden, Misiolek, Perlmutter and Ratiu [1998a,b] in which one has a group M with a normal subgroup $N \subset M$ and M acts on a symplectic manifold P. One wants to reduce P in two stages, first by N and then by M/N. On the Poisson level this is easy: $P/M \cong (P/N)/(M/N)$ but on the symplectic level it is quite subtle. Cendra, Marsden and Ratiu [1998] have developed a Lagrangian counterpart to reduction by stages.

Singular reduction. Singular reduction starts with the observation of Smale [1970] that $z \in P$ is a regular point of $\mathbf{J}$ iff z has no continuous isotropy. Motivated by this, Arms, Marsden and Moncrief [1981] showed that the level sets $\mathbf{J}^{-1}(0)$ of an equivariant momentum map $\mathbf{J}$ have quadratic singularities at points with continuous symmetry. While easy for compact group actions, their main examples were infinite dimensional! The structure of $J^{-1}(0)/G$ for compact groups was developed in Sjamaar and Lerman [1990], and extended to $J^{-1}(\mu)/G_\mu$ by Bates and Lerman [1996] and Ortega and Ratiu [1997a]. Many specific examples of singular reduction and further references may be found in Bates and Cushman [1997].

The method of invariants. An important method for the reduction construction is called the ***method of invariants.*** This method seeks to parameterize quotient spaces by functions that are invariant under the group action. The method has a rich history going back to Hilbert's *invariant theory* and it has much deep mathematics associated with it. 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] and the book of Bates and Cushman [1997] for more details and references. We shall illustrate the method with a famous system, the three wave interaction, based on Alber, Luther, Marsden and Robbins [1998b].

The three wave interaction. The ***quadratic resonant three wave equations*** are the following ode's on $\mathbb{C}^3$:

$$\begin{aligned}\frac{dq_1}{dt} &= is_1\gamma_1 q_2\bar{q}_3\\ \frac{dq_2}{dt} &= is_2\gamma_2 q_1 q_3\\ \frac{dq_3}{dt} &= is_3\gamma_3\bar{q}_1 q_2\end{aligned}$$

Here, $q_1, q_2, q_3 \in \mathbb{C}$, $i = \sqrt{-1}$, the overbar means complex conjugate, and γ_1, γ_2 and γ_3 are nonzero real numbers with $\gamma_1 + \gamma_2 + \gamma_3 = 0$. The choice $(s_1, s_2, s_3) = (1, 1, -1)$; gives the ***decay interaction***, while $(s_1, s_2, s_3) = (-1, 1, 1)$ gives the ***explosive interaction***.

Resonant wave interactions describe energy exchange among nonlinear modes in contexts involving nonlinear waves (the Benjamin-Feir instability, etc.) in fluid mechanics, plasma physics and other areas. There are other versions of the equations in which coupling associated with phase modulations appears through linear and cubic terms. Much of our motivation comes from nonlinear optics (optical transmission and switching). The three wave equations are discussed in, for example, Whitham [1974] and its dynamical systems aspects are explored in Guckenheimer and Mahalov [1992].

The methods we develop work rather generally for resonances—the rigid body is well known to be intimately connected with the 1:1 resonance (see, for example, Cushman and Rod [1982], Churchill, Kummer and Rod [1983]). The three wave interaction has an interesting Hamiltonian and integrable structure. We shall use a standard Hamiltonian structure and the technique of invariants to understand it. The decay system is ***Lie-Poisson*** for the Lie algebra $\mathfrak{su}(3)$ – this is the one of notable interest for phases (the explosive case is associated with $\mathfrak{su}(2,1)$). This is related to the Lax representation of the equations—the n-wave interaction is likewise related to $\mathfrak{su}(n)$. The general picture developed is useful for many other purposes, such as polarization control (building on work of David, Holm and Tratnik [1989] and David and Holm [1990]) and perturbations of Hamiltonian normal forms (see Kirk, Marsden and Silber [1996]).

The canonical Hamiltonian structure. We describe how the three wave system is Hamiltonian relative to a canonical Poisson bracket. We choose (primarily a matter of convenience) a γ_i-weighted canonical bracket on $\mathbb{C}^3$. This bracket has the real and imaginary parts of each complex dynamical variable q_i as conjugate variables. Correspondingly, we will use a cubic Hamiltonian. The ***scaled canonical Poisson bracket*** on $\mathbb{C}^3$ may be written in complex notation as

$$\{F, G\} = -2i\sum_{k=1}^{3} s_k\gamma_k\left(\frac{\partial F}{\partial q^k}\frac{\partial G}{\partial \bar{q}_k} - \frac{\partial G}{\partial q^k}\frac{\partial F}{\partial \bar{q}_k}\right).$$

The corresponding ***symplectic structure*** can be written

$$\Omega((z_1, z_2, z_3), (w_1, w_2, w_3)) = -\sum_{k=1}^{3}\frac{1}{s_k\gamma_k}\mathrm{Im}(z_k\bar{w}_k).$$

The (cubic) ***Hamiltonian*** is

$$H = -\frac{1}{2}\left(\bar{q}_1 q_2\bar{q}_3 + q_1\bar{q}_2 q_3\right).$$

Hamilton's equations for a Hamiltonian H are

$$\frac{dq_k}{dt} = \{q_k, H\} \ ,$$

and it is straightforward to check that Hamilton's equations are given in complex notation by

$$\frac{dq_k}{dt} = -2is_k\gamma_k \frac{\partial H}{\partial \bar{q}_k} \ ,$$

One checks that Hamilton's equations in our case coincide with the three wave equations.

Integrals of motion. Besides H itself, there are additional constants of motion, often referred to as the ***Manley-Rowe relations***:

$$\begin{aligned} K_1 &= \frac{|q_1|^2}{s_1\gamma_1} + \frac{|q_2|^2}{s_2\gamma_2} \ , \\ K_2 &= \frac{|q_2|^2}{s_2\gamma_2} + \frac{|q_3|^2}{s_3\gamma_3} \ , \\ K_3 &= \frac{|q_1|^2}{s_1\gamma_1} - \frac{|q_3|^2}{s_3\gamma_3} \ . \end{aligned}$$

The vector function (K_1, K_2, K_3) is the ***momentum map*** for the following symplectic action of the group $T^3 = S^1 \times S^1 \times S^1$ on $\mathbb{C}^3$:

$$\begin{aligned} (q_1, q_2, q_3) &\mapsto (q_1 \exp(i\gamma_1), q_2 \exp(i\gamma_1), q_3), \\ (q_1, q_2, q_3) &\mapsto (q_1, q_2 \exp(i\gamma_2), q_3 \exp(i\gamma_2)), \\ (q_1, q_2, q_3) &\mapsto (q_1 \exp(i\gamma_3), q_2, q_3 \exp(-i\gamma_3)). \end{aligned}$$

The Hamiltonian taken with any two of the K_j are checked to be a complete and independent set of conserved quantities. Thus, the system is ***Liouville-Arnold integrable***.

The K_j clearly give only two independent invariants since $K_1 - K_2 = K_3$. Any combination of two of these actions can be generated by the third reflecting the fact that the K_j are linearly dependent. Another way of saying this is that the group action by T^3 is really captured by the action of T^2.

Integrating the equations. To carry out the integration, one can make use of the Hamiltonian plus two of the integrals, K_j to reduce the system to quadratures. This is often carried out using the transformation $q_j = \sqrt{\rho_j} \exp i\phi_j$ to obtain expressions for the phases ϕ_j. The resulting expressions are nice, but the alternative point of view using invariants is also useful.

Poisson reduction. Symplectic reduction of the above Hamiltonian system uses the symmetries and associated conserved quantities K_k. In ***Poisson reduction***, we replace $\mathbb{C}^3$ with the ***orbit space*** $\mathbb{C}^3/T^2$, which then inherits a Poisson structure. To obtain the ***symplectic leaves*** in this reduction, we use the ***method of invariants***. Invariants for the T^2 action are:

$$\begin{aligned} X + iY &= q_1\bar{q}_2 q_3 \\ Z_1 &= |q_1|^2 - |q_2|^2 \\ Z_2 &= |q_2|^2 - |q_3|^2 \end{aligned}$$

These quantities provide coordinates for the four dimensional orbit space $\mathbb{C}^3/T^2$. The following identity (this is part of the invariant theory game) holds for these invariants and the conserved quantities:

$$X^2 + Y^2 = \beta(\delta - Z_2)(Z_2 + s_3\gamma_3 K_2)(s_2\gamma_2 K_2 - Z_2)$$

where the constants β, δ are given by

$$\beta = \frac{s_1\gamma_1 s_2\gamma_2 s_3\gamma_3}{(s_2\gamma_2 + s_3\gamma_3)^3}, \quad \delta = s_2\gamma_2 K_1 + s_3\gamma_3(K_1 - K_2).$$

This defines a two dimensional surface in (X, Y, Z_2) space, with Z_1 determined by the values of these invariants and the conserved quantities (so it may also be thought of as a surface in (X, Y, Z_1, Z_2) as well). A sample of one of these surfaces is plotted in Figure 1.8.1.

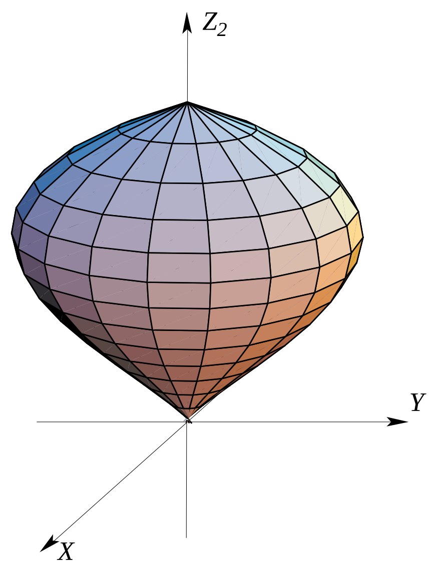

Figure 1.8.1: The reduced phase space for the three-wave equations.

We call these surfaces the ***three wave surfaces.*** They are examples of ***orbifolds***. The evident singularity in the space is typical of orbifolds and comes about from the non-freeness of the group action.

Any trajectory of the original equations defines a curve on each three wave surface, in which the K_j are set to constants. These three wave surfaces are the symplectic leaves in the four dimensional Poisson space with coordinates (X, Y, Z_1, Z_2).

The original equations define a dynamical system in the Poisson reduced space and on the symplectic leaves as well. The ***reduced Hamiltonian*** is

$$H(X, Y, Z_1, Z_2) = -X$$

and indeed, $\dot{X} = 0$ is one of the reduced equations. Thus, the trajectories on the reduced surfaces are obtained by slicing the surface with the planes $X =$ Constant. The ***Poisson structure*** on $\mathbb{C}^3$ drops to a Poisson structure on (X, Y, Z_1, Z_2)-space, and the ***symplectic structure*** drops to one on each three wave surface—this is of course an example of the general procedure of ***symplectic reduction***. Also, from the geometry, it is clear that interesting ***homoclinic orbits*** pass through the singular points—these are cut out by the plane $X = 0$.

A ***control perspective*** allows one to manipulate the plane $H = -X$ and thereby the dynamics. This aspect is explored in Alber, Luther, Marsden and Robbins [1998a].

Chapter 2

Symplectic Reduction

The classical Noether theorem provides conservation laws for mechanical systems with symmetry. The conserved quantities collected as vector valued maps on phase space are called ***momentum maps***. Momentum maps have many wonderful properties; one of these is that they are Poisson maps from the phase space (either a symplectic or a Poisson manifold) to the dual of the Lie algebra of the symmetry group, with its Lie-Poisson structure, as was proved in *IMS*.

The main goal of this chapter is to study the procedure of reducing the size of the phase space by taking advantage of the conserved momentum map and the invariance of the system under the given symmetry group. The results obtained generalize the classical theorems of Liouville and Jacobi on reduction of systems by $2k$ dimensions if there are k integrals in involution. The general reduction method also includes Jacobi's elimination of the node, fixing the center of mass in the n-body problem, as well as the coadjoint orbit symplectic structure. This procedure plays a crucial role in many related constructions, both mathematical and physical. Some key examples are given in the text along with the theory; beside the basic examples using linear and angular momentum, one may treat other more sophisticated examples, such as the Maxwell equations in vacuum and the equations for a charged particle in an electromagnetic field. Appropriately combining them leads to the Maxwell-Vlasov system, as in Marsden and Weinstein [1982]. Another interesting example is the dynamics of coupled rigid bodies, rigid bodies with flexible attachments, etc. Later on, we will also consider reconstruction—the opposite of reduction—and how it can be used to give insight to geometric phases. The basic idea of geometric phases was already discussed in the introduction to *IMS*.

This Chapter begins with the study of reduction in the context of presymplectic structures. Then it goes on to the case of symplectic reduction and then later on links this up with Poisson reduction (the beginnings of Poisson reduction were started in Chapter 10 of *IMS*. Of course the reduction of cotangent bundles is a very important situation and so it is given special attention.

2.1 Presymplectic Reduction

A general setting for symplectic reduction (going back to Cartan [1922]) is the following. Suppose ω is a ***presymplectic form***; *i.e.*, a closed two-form on a manifold N. Let E_ω be the ***characteristic*** (or ***null***) ***distribution*** of ω, defined as the distribution whose fiber at $x \in N$ is

$$E_{\omega,x} = \{u \in T_xN \mid \omega(u,v) = 0 \quad \text{for all} \quad v \in T_xN\}.$$

If X is a vector field on N, note that it takes values in E_ω iff $\mathbf{i}_X\omega = 0$. Call ω ***regular*** if E_ω is a subbundle of TN. The latter condition holds in finite dimensions iff ω has constant rank; see Abraham, Marsden and Ratiu, *Manifolds, Tensor Analysis and Applications*, hereafter referred to as *MTA*, §4.4 for this and a corresponding result for the infinite dimensional case. We assume ω is regular in the following discussion. We begin with the following important observation.

Lemma 2.1.1 *The characteristic distribution E_ω of a regular presymplectic form ω is an involutive distribution; that is, that sections of E_ω are closed under the Jacobi-Lie bracket.*

Proof. We will use the following general identity about differential forms:

$$\mathbf{i}_{[X,Y]}\omega = \pounds_X \mathbf{i}_Y\omega - \mathbf{i}_Y \pounds_X\omega$$

which is proved, for example, in *MTA*. Suppose X and Y are vector fields on N that take their values in E_ω; then,

$$\mathbf{i}_{[X,Y]}\omega = \pounds_X \mathbf{i}_Y\omega - \mathbf{i}_Y \pounds_X\omega = 0 - \mathbf{i}_Y(\mathbf{i}_X\mathbf{d}\omega + \mathbf{d}\mathbf{i}_X\omega) = 0,$$

so $[X, Y]$ also takes its value in E_ω. ■

By Frobenius' theorem (see *MTA*), E_ω, being an integrable distribution, defines a foliation Φ on N, called the ***null foliation*** of ω. Thus, Φ is a disjoint collection of submanifolds whose union is N and whose tangent spaces are, at every point $x, E_{\omega,x}$, the fiber of E_ω over x. Assume, in addition, that Φ is a ***regular foliation***, *i.e.*, the space N/Φ of leaves of the foliation is a smooth manifold and the canonical projection $\pi : N \to N/\Phi$ is a submersion. Necessary and sufficient conditions for this to hold are that the graph of Φ in $N \times N$ is a closed submanifold and the projection p_1 : graph $\Phi \to N$ onto the first factor is a surjective submersion; see *MTA* §3.5 for details. Under these hypotheses, the tangent space to N/Φ at $[x]$, the leaf through $x \in N$, is isomorphic to the vector space quotient $T_xN/E_{\omega,x}$, the isomorphism being implemented by the projection π (see Figure 2.1.1).

We wish to define a two-form ω_Φ on N/Φ by

$$\omega_\Phi([x])([u],[v]) = \omega(x)(u, v) \tag{2.1.1}$$

where $u, v \in T_xN$ and $[u], [v]$ denote their equivalence classes in $T_xN/E_{\omega,x}$. For this definition to make sense, we need to prove that ω_Φ is well-defined, that is, is independent of choices of representatives of equivalence classes. While doing this, refer to Figure 2.1.2.

Lemma 2.1.2 *Formula (2.1.1) defines a two form on N/Φ.*

Proof. If $[x] = [x']$, then (by the proof of Frobenius' theorem) there is a vector field X with values in E_ω such that $x' = \varphi(x)$, where φ is the time one map for X. Let $u' = T_x\varphi \cdot u$, $v' = T_x\varphi \cdot v$, and let $[u''] = [u']$ and $[v''] = [v']$, so that $u'' - u'$ and $v'' - v'$ both belong to $E_{\omega,x'}$. Thus,

$$\begin{aligned}
\omega(x')(u'', v'') &= \omega(x')\left((u'' - u') + u',\, (v'' - v') + v'\right) = \omega(x')(u', v') \\
&= \omega(\varphi(x))(T_x\varphi \cdot u,\, T_x\varphi \cdot v) = (\varphi^*\omega)(x)(u, v).
\end{aligned}$$

Since $\mathbf{d}\omega = 0$ and $\mathbf{i}_X\omega = 0$, we have $\pounds_X\omega = 0$, so that $\varphi^*\omega = \omega$; *i.e.*, $\omega(x')(u'', v'') = \omega(x)(u, v)$, showing that ω_Φ is well-defined. ■

From the construction of ω_Φ, it follows that

$$\omega = \pi^*\omega_\Phi. \tag{2.1.2}$$

Since π is a surjective submersion, ω_Φ is uniquely determined by property (2.1.2).

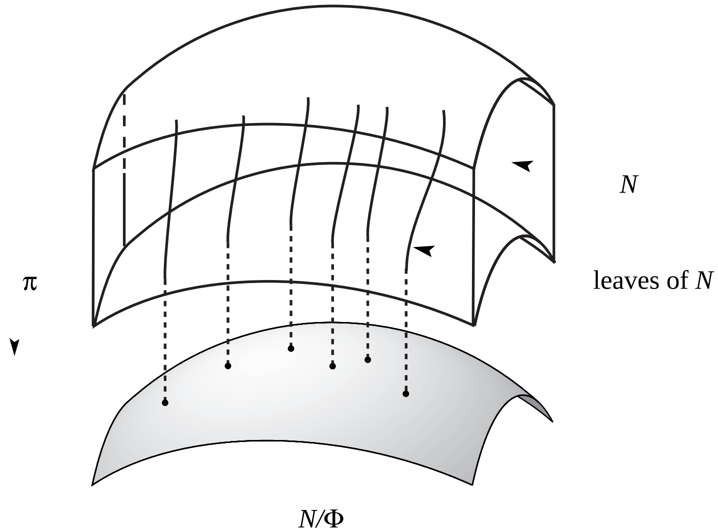

Figure 2.1.1: N is foliated by the leaves of the foliation Φ, the null foliation of a closed two form ω.

Lemma 2.1.3 *The form ω_Φ is closed.*

Proof. From $\mathbf{d}\omega = 0$, we get $0 = \mathbf{d}\pi^*\omega_\Phi = \pi^*\mathbf{d}\omega_\Phi$, and since π is a submersion, $\mathbf{d}\omega_\Phi = 0$. ■

Lemma 2.1.4 *The form ω_Φ is (weakly) non-degenerate.*

Proof. Indeed, if $\omega_\Phi([x])([u],[v]) = \omega(x)(u,v) = 0$ for all $v \in T_xN$, then $u \in E_{\omega,x}$ *i.e.*, $[u] = [0]$. ■

Summarizing these results, we have thus proved the following.

Theorem 2.1.5 (Foliation Reduction Theorem) *Let N be a smooth manifold and ω a closed 2-form on N. Assume that the characteristic distribution $E_\omega \subset TN$ of ω is regular, that the foliation Φ it defines is regular, and let $\pi : N \to N/\Phi$ denote the canonical projection. Then ω induces a unique (weak) symplectic structure ω_Φ on N/Φ by the relation $\pi^*\omega_\Phi = \omega$. The manifold N/Φ of leaves is called the **reduced space**.*

For the next corollary, recall that in a symplectic manifold (P, Ω), the Ω-***orthogonal complement*** of a subbundle $E \subset TP$ is the bundle E^Ω whose fiber at $z \in P$ is the linear space

$$E_z^\Omega = \{v \in T_zP \mid \Omega(v,w) = 0 \quad \text{for all} \quad w \in E_z\}.$$

Corollary 2.1.6 *Let (P, Ω) be a symplectic manifold and $N \subset P$ a submanifold. Suppose that $TN \cap (TN)^\Omega$ is a subbundle of TN. Then*

i *$TN \cap (TN)^\Omega$ is an integrable subbundle of TN;*

ii *$TN/[TN \cap (TN)^\Omega]$ is a symplectic vector bundle over N; that is, each fiber has an induced symplectic structure varying smoothly over N;*

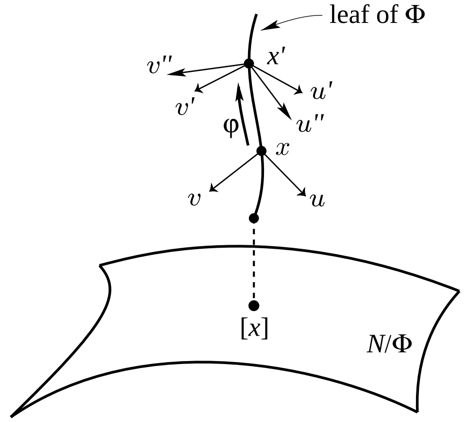

Figure 2.1.2: Showing that there is a well defined two form on the quotient space.

iii *if the foliation Φ determined by $TN \cap (TN)^{\Omega}$ is regular, then N/Φ is a symplectic manifold.*

Proof Let $i : N \to P$ be the inclusion and $\omega = i^*\Omega$. The characteristic distribution of ω at $z \in N$ equals

$$E_{\omega,z} = \{v \in T_zN \mid \Omega(z)(v,u) = 0 \quad \text{for all} \quad u \in T_zN\} = T_zN \cap (T_zN)^{\Omega},$$

so that the assertions of the corollary follow from the foliation reduction theorem. ■

Notice that when N is ***coisotropic***, *i.e.*, when $(TN)^{\Omega} \subset TN$, the foliation Φ is determined by the distribution $(TN)^{\Omega}$.

Exercises

2.1-1

(a) *If (E, Ω) is a symplectic vector space and if $N \subset E$ is a subspace, show directly that $N/(N \cap N^{\Omega})$ is a symplectic vector space.*

(b) *If (N, Ω) is a presymplectic vector space, show that N/N^{Ω} is symplectic.*

2.1-2 *Let $L \subset E$ be a Lagrangian subspace of a symplectic vector space (E, Ω) i.e., $L^{\Omega} = L$. If $N \supset L$, show that N is coisotropic, whereas if $N \subset L$, N is isotropic.*

2.1-3

(a) *Let X be a divergence free vector field in $\mathbb{R}^3$ and let $\mu = dx \wedge dy \wedge dz$ be the standard volume element. Show that $\omega = \mathbf{i}_X \mu$ is a presymplectic form. Describe the characteristic foliation and the reduced phase space for this example.*

(b) *Carry out this construction for the vector field $X(x, y, z) = (1/yz, 1/xz, 1/xy)$.*

2.2 Symplectic Reduction by a Group Action

One of the most important situations in which reduction occurs is when the foliation is determined by a group action, and the set N is a level set of a momentum map. This section deals with this case.

Let $\Phi : G \times P \to P$ be a (left) symplectic action of a Lie group G on the symplectic manifold (P, Ω) with an Ad^*-equivariant momentum map $\mathbf{J} : P \to \mathfrak{g}^*$. Let $\mu \in \mathfrak{g}^*$ be a regular value of $\mathbf{J}$. (As we shall see in the remarks below, this condition can be weakened somewhat and cases where it fails are important, but we assume for the moment that μ is a regular value.) Thus, $\mathbf{J}^{-1}(\mu)$ is a submanifold of P and, if G and P are finite dimensional, then $\dim \mathbf{J}^{-1}(\mu) = \dim P - \dim G$. Let

$$G_\mu = \{g \in G \mid \mathrm{Ad}_g^* \mu = \mu\}$$

be the ***isotropy subgroup*** at μ for the coadjoint action. Its Lie algebra is

$$\mathfrak{g}_\mu = \{\xi \in \mathfrak{g} \mid \mathrm{ad}_\xi^* \mu = 0\}.$$

Lemma 2.2.1 *The set $\mathbf{J}^{-1}(\mu)$ is invariant under the action of G_μ.*

Proof. We are asserting that for $z \in \mathbf{J}^{-1}(\mu)$, then $\Phi_g(z) \in \mathbf{J}^{-1}(\mu)$ for all $g \in G_\mu$. Indeed, this follows from the following calculation in which the appropriate group action is denoted by concatenation and in which equivariance of the momentum map is used in the first equality:

$$\mathbf{J}(gz) = g\mathbf{J}(z) = g\mu = \mu \quad \blacksquare.$$

Definition 2.2.2 *The quotient space $P_\mu = \mathbf{J}^{-1}(\mu)/G_\mu$, is called the **reduced phase space** at $\mu \in \mathfrak{g}^*$.*

The reduced space is a manifold if G_μ acts freely (*i.e.*, if, for each $z \in \mathbf{J}^{-1}(\mu)$, $g \cdot z = z$ implies $g = e$) and properly (*i.e.*, $(g, z) \mapsto (g, g \cdot z)$ is a proper map) on $\mathbf{J}^{-1}(\mu)$. Under these hypotheses, P_μ is a manifold, and the canonical projection $\pi_\mu : \mathbf{J}^{-1}(\mu) \to P_\mu$ is a surjective submersion. We alluded to theorems of this type above and refer to *MTA* for proofs.[1]

The ***symplectic reduction theorem*** states that *P_μ is a symplectic manifold, the symplectic form being naturally induced from Ω.* It has a second part dealing with how a Hamiltonian system drops to the reduced space that will be treated in §**1.3**. The symplectic reduction theorem was formulated in this way by Marsden and Weinstein [1974] (see also Meyer [1973]). Related earlier special but important versions of these theorems were given by Arnold [1966], Smale [1970], and Nehoroshev [1970]. These results are inspired by classical cases of Liouville and Jacobi (see, for example, Whittaker [1925]).

Theorem 2.2.3 (Symplectic Reduction Theorem) *Consider a symplectic manifold (P, Ω) on which there is a Hamiltonian left action of a Lie group G with an equivariant momentum map $\mathbf{J} : P \to \mathfrak{g}^*$. Assume that $\mu \in \mathfrak{g}^*$ is a regular value of $\mathbf{J}$ and that the isotropy group G_μ acts freely and properly[2] on $\mathbf{J}^{-1}(\mu)$. Then the reduced phase space $P_\mu = \mathbf{J}^{-1}(\mu)/G_\mu$ has a unique (weak) symplectic form Ω_μ characterized by*

$$\pi_\mu^* \Omega_\mu = i_\mu^* \Omega, \tag{2.2.1}$$

[1] In the infinite dimensional case one uses special techniques to prove that the quotients are manifolds, based on *slice* theorems (see Ebin [1970], Isenberg and Marsden [1983] and references therein.)

[2] In infinite dimensions, assume that Ω is weakly non-degenerate and add the hypothesis that the map from the group to the orbit is an immersion, or replace the properness and immersion hypothesis by the assumption that there is a slice theorem available to guarantee that the quotient is a manifold.

where $\pi_\mu : \mathbf{J}^{-1}(\mu) \to P_\mu$ is the canonical projection and $i_\mu : \mathbf{J}^{-1}(\mu) \to P$ is inclusion. (See Figure 2.2.1) Finally (in the infinite dimensional case), if Ω is a strong symplectic form, so is Ω_μ.

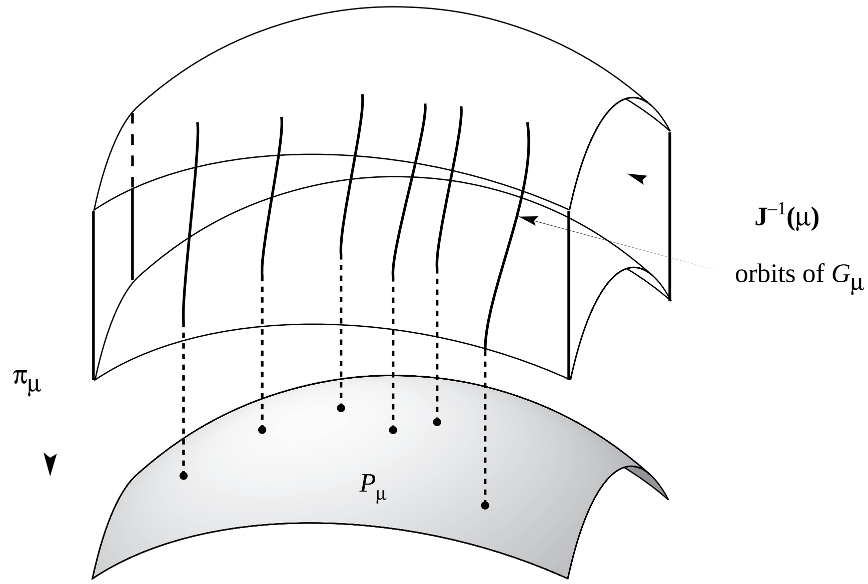

Figure 2.2.1: In the symplectic reduction theorem, the orbits of G_μ give the characteristic foliation of $\mathbf{J}^{-1}(\mu)$.

To prove the symplectic reduction theorem, we prepare a few more lemmas.

Lemma 2.2.4 *Let (V, Ω) be a weak symplectic Banach space and $W \subset V$ be a closed subspace. Then*

$$(W^\Omega)^\Omega = W. \tag{2.2.2}$$

Proof That there is a natural inclusion $W \subset (W^\Omega)^\Omega$ follows directly from the definitions. We first prove the converse inclusion in the finite dimensional situation and prove the general case below.

First we show that

$$\dim V = \dim W + \dim W^\Omega, \tag{2.2.3}$$

even though $V \neq W \oplus W^\Omega$ in general. To prove (2.2.3), 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 non-degenerate, $\Omega^\flat : V \to V^*$ 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 (2.2.3) holds.

Applying (2.2.3) 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 $W = (W^\Omega)^\Omega$. ▼

Optional Proof of Lemma 2.2.4 in the Infinite Dimensional Case
We start by recalling the Hahn-Banach theorem in the setting of locally convex topological spaces. The proof may be found in *MTA*, Choquet [1969, §21], or Yosida [1971].

If V is a locally convex Hausdorff topological vector space, W is a closed subspace, and $v \notin W$, then there is a continuous linear functional $\alpha : V \to \mathbb{R}$ such that $\alpha | W = 0$ and $\alpha(v) = 1$.

Now let (V, Ω) be a weak symplectic Banach space. With respect to the family of seminorms $p_y(x) = |\Omega(x, y)|$, V becomes a locally convex topological vector space; it is verified to be Hausdorff since Ω is nondegenerate. Let us call this the Ω-*topology.* We also require the following result.

The dual of V as a locally convex topological space is $V^*_\Omega = \Omega^\flat(V) \subset V^*$. That is, a linear map $\alpha : V \to \mathbb{R}$ is continuous in the locally convex Ω- topology of V if and only if there exists a (unique) $y \in V$ such that $\alpha(x) = \Omega(x, y)$ for all $x \in V$.

This is proved, as in Choquet [1969, §22] as follows. Uniqueness of y is clear by nondegeneracy of Ω. For existence, note that since $\alpha : V \to \mathbb{R}$ is continuous in the Ω-topology on V, there exist $y_1, \ldots , y_n \in V$ such that $|\alpha(x)| \leq C \max_{1 \leq i \leq n} |\Omega(x, y_i)|$ for C a positive constant. Consequently, α vanishes on

$$D = \bigcap_{i=1}^{n} \ker \Omega^\flat(y_i) = (\operatorname{span}\{y_i, \ldots , y_n\})^\Omega;$$

E is clearly closed and has codimension at most n. Let F be an algebraic complement to E in V; F being finite dimensional, is closed and hence $V = E \oplus F$, a Banach space direct sum (*MTA*, Supplements 2.1B and 3.2C). It is clear that $\Omega^\flat(y_1)|F, \ldots , \Omega^\flat(y_n)|F$ span F^* and thus we can write

$$\alpha \,|\, F = \sum_{k=1}^{n} \alpha_k \Omega^\flat(y_k) \,|\, F = \Omega^\flat \left(\sum_{k=1}^{n} \alpha_k y_k \right) |\, F.$$

Since both sides of this equality vanish on E, we get $\alpha = \Omega^\flat \left(\sum_{k=1}^{n} a_k y_k \right)$, so the claim is proved. ▼

Now we return to Lemma 2.2.4. Suppose that $v \in V \backslash W$. By the Hahn-Banach theorem, there is an $\alpha \in V^*_\Omega$ such that $\alpha = 0$ on W and $\alpha(v) = 1$. By the preceding result applied to α, there exists a unique $u \in V$ such that $\alpha(z) = \Omega(z, u)$ for all $z \in V$. Thus, $\Omega(v, u) \neq 0$ and $\Omega(z, u) = 0$ for all $z \in W$. In other words, $\Omega(v, u) \neq 0$ and $u \in W^\Omega$, *i.e.*, $v \notin (W^\Omega)^\Omega$. Thus we have shown that $(W^\Omega)^\Omega \subset W$; combined with the trivial inclusion $W \subset (W^\Omega)^\Omega$, we get the equality $(W^\Omega)^\Omega = W$. ♦

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 lemma that is useful in a number of situations.

Lemma 2.2.5 (Reduction Lemma) *Let P be a Poisson manifold and let $\mathbf{J} : P \to \mathfrak{g}^*$ be an equivariant momentum map of a Hamiltonian Lie group action 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 $\mathbf{J}^{-1}(G \cdot \mu) = G \cdot \mathbf{J}^{-1}(\mu) = \{g \cdot z \mid g \in G \text{ and } \mathbf{J}(z) = \mu\};$

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*

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

are Ω*-orthogonal complements of each other.*

Refer to Figure 2.2.2 for one way of showing the geometry associated with this lemma.

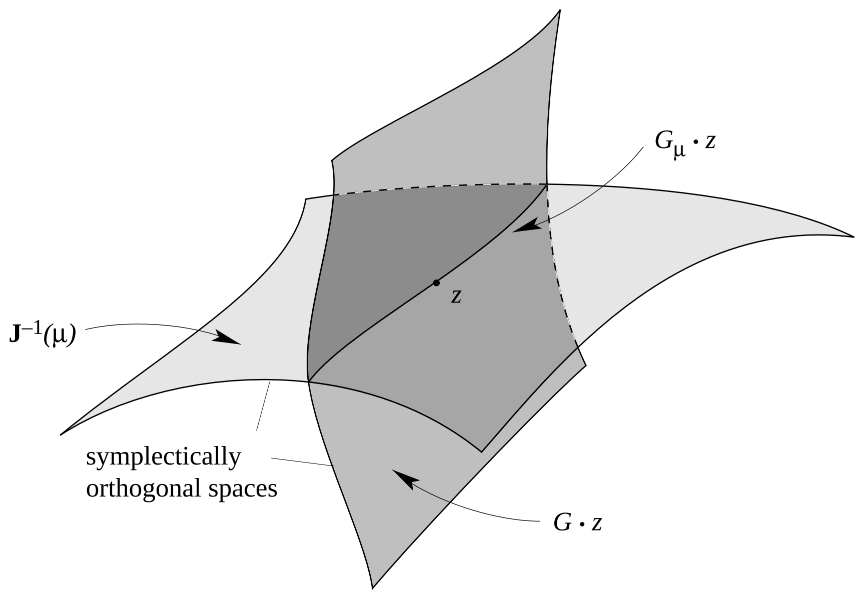

Figure 2.2.2: The geometry of the reduction lemma.

Proof of the Reduction Lemma

i $z \in \mathbf{J}^{-1}(G\cdot\mu)$ iff $\mathbf{J}(z) = \mathrm{Ad}^*_{g^{-1}}\mu$ 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 $g \cdot z \in \mathbf{J}^{-1}(\mu)$ iff $\mu = \mathbf{J}(g \cdot z) = \mathrm{Ad}^*_{g^{-1}}\mathbf{J}(z) = \mathrm{Ad}^*_{g^{-1}}\mu$ iff $g \in G_\mu$.

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 $0 = T_z\mathbf{J}(v_z) = 0$ 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 $v_z \in \ker T_z\mathbf{J} = T_z(\mathbf{J}^{-1}(\mu))$. ■

We notice from **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.

We now give two proofs of the symplectic reduction theorem. The first obtains it as a special case of Corollary **2.1.6**, while the second gives a direct proof that bypasses the terminology and results on foliations.

First Proof of the Symplectic Reduction Theorem (This proof assumes that G_μ is connected.) In Corollary **2.1.6**, let $N = \mathbf{J}^{-1}(\mu)$ and $z \in N$. Then

$$\begin{aligned} T_zN \cap (T_zN)^\Omega &= T_z\mathbf{J}^{-1}(\mu) \cap (T_z\mathbf{J}^{-1}(\mu))^\Omega \\ &= T_z\mathbf{J}^{-1}(\mu) \cap T_z(G \cdot z) \qquad \text{(by the reduction lemma } \mathbf{iv} \text{ and } \mathbf{2.2.4}) \\ &= T_z(G_\mu \cdot z) \qquad \text{(by the reduction lemma } \mathbf{iii}). \end{aligned}$$

Thus the foliation Φ has as leaves the G_μ orbits in $\mathbf{J}^{-1}(\mu)$, since G_μ is connected. Hence N/Φ is just $\mathbf{J}^{-1}(\mu)/G_\mu$, and so the result follows by the corollary to the foliation reduction theorem. ■

Second 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 $T_{[z]}(\mathbf{J}^{-1}(\mu)/G_\mu)$ with $T_z(\mathbf{J}^{-1}(\mu))/T_z(G_\mu \cdot z)$, where $[z] = \pi_\mu(z)$. 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 $y = \Phi_g(z)$, $v' = T_z\Phi_g \cdot v$, and $w' = T_z\Phi_g \cdot w$, where $g \in G_\mu$. If, in addition $[v''] = [v']$ and $[w''] = [w']$, then

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

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 $\mathbf{d}\Omega_\mu = 0$.

For (weak) nondegeneracy of Ω_μ, suppose $\Omega_\mu([z])([v],[w]) = 0$ for all $w \in T_z(\mathbf{J}^{-1}(\mu))$. This means that $\Omega(z)(v,w) = 0$ for all $w \in T_z(\mathbf{J}^{-1}(\mu))$, *i.e.*, that $v \in (T_z(\mathbf{J}^{-1}(\mu)))^\Omega = T_z(G \cdot z)$ by lemma **2.2.4** and the reduction lemma **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 **iii** so that $[v] = 0$, thus proving the weak nondegeneracy of Ω_μ.

Finally, if Ω is a strong symplectic form, let $\overline{\alpha} \in T_{[z]}^*P_\mu$ be represented by the one-form $\alpha : T_z\mathbf{J}^{-1}(\mu) \to \mathbb{R}$ vanishing on the closed subspace $T_z(G_\mu \cdot z)$ *i.e.*, $\overline{\alpha}([w]) = \alpha(w)$ for all $w \in T_z\mathbf{J}^{-1}(\mu)$. Since Ω is strongly nondegenerate, there is a $v \in T_zP$ such that $\Omega(z)(v,w) = \overline{\alpha}(w)$ for all $w \in T_z\mathbf{J}^{-1}(\mu)$. From the reduction lemma **iv**, it follows that $v \in (T_z(G_\mu \cdot z))^\Omega \subset T_z\mathbf{J}^{-1}(\mu)$ and so $\Omega_\mu([z])([v]),[w]) = \alpha([w])$ for all $w \in T_z\mathbf{J}^{-1}(\mu)$. ■

If μ is a regular value of $\mathbf{J}$ then the action is automatically locally free, so the prior construction can be carried out locally. To prove this, consider the ***symmetry algebra*** at $z \in P$ defined by

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

Proposition 2.2.6 *An element $\mu \in \mathfrak{g}^*$ is a regular value of $\mathbf{J}$ iff $\mathfrak{g}_z = 0$ 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 is a remarkably easy way to recognize regular points. For example, in the double spherical pendulum, one can say right away that the only irregular points are those with *both* pendula pointing straight down, or both pointing straight up.

Proof By definition, z is a regular point iff $T_z\mathbf{J}$ is surjective. This is equivalent to saying that its annihilator is zero:[3]

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

But

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

by definition of the momentum map. Therefore, z is a regular point iff

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

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

This result, connecting the symmetry of z with the regularity of μ, suggests that *points with symmetry are bifurcation points of* $\mathbf{J}$. This observation turns out to have many important consequences. A related result is that if G is compact (or if a slice theorem is valid) then the singularities of $\mathbf{J}^{-1}(\mu)$ at symmetric points are necessarily quadratic. We refer to Smale [1970], Abraham and Marsden [1978, §4.5], Arms, Marsden and Moncrief [1981], [1982], Atiyah [1982], Guillemin and Sternberg [1982], [1984] and Kirwan [1984a,b]. See also Arms, Gotay and Cushman [1989] and Sjamaar and Lerman [1991].

Remarks on the Reduction Theorem

1. Even if $\Omega = -\mathbf{d}\theta$ and the action of G leaves θ invariant, Ω_μ *need not be exact.* An explicit example is the coadjoint orbit of $SO(3)$, a sphere. This is shown to be a symplectic reduced space in the next section.

2. 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 will be used later in this section 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]. The general definitions are as follows. 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 $x \in f^{-1}(y)$, $T_xf^{-1}(y) = \ker T_xf$. We say that f intersects a submanifold $L \subset N$ ***cleanly*** if $f^{-1}(L)$ is a submanifold of M and $T_x(f^{-1}(L)) = (T_xf)^{-1}(T_{f(x)}L)$. Note that *regular values of f are clean values and that if f intersects the submanifold L transversally, then it intersects it cleanly.*

[3] In the infinite dimensional case we will not necessarily choose $\mathfrak{g}^*$ to be the Banach space dual, but rather a suitable Sobolev function space in duality with $\mathfrak{g}$. The assertion that A_z^* is onto requires the Fredholm alternative. In many examples, such as gravity and the Yang-Mills equations, this holds because A_z^* is elliptic. See Arms, Marsden and Moncrief [1981], [1982], references therein, and remark **3** following the orbit reduction theorem in §**1.5** below.

3. 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. For example, a slice theorem would often suffice for this hypothesis.

4. Even if μ is a regular value, it need not be a generic point in $\mathfrak{g}^*$, that is a point whose coadjoint orbit is of maximal dimension. Note that the reduction theorem does not require this assumption. 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 generic.

5. If G acts on P on the *right* in a Hamiltonian manner,the same theorem holds; *i.e.*, $\mathbf{J}^{-1}(\mu)/G_\mu$ is still symplectic in a canonical way. Sometimes it is necessary to distinguish the left and right quotients in the notation. When necessary one could use ${}_\mu P = G_\mu \backslash \mathbf{J}^{-1}(\mu)$ for left quotients and $P_\mu = \mathbf{J}^{-1}(\mu)/G_\mu$ for right quotients. Normally there is no danger of confusion, so we simply write $\mathbf{J}^{-1}(\mu)/G_\mu = P_\mu$ in either case.

6. Here is a simple example of reduction. It consists of the manifold of ***solutions of constant energy***. Let P be a symplectic manifold, $H \in \mathcal{F}(P)$ and $e \in \mathbb{R}$ a regular value of H. Then $H^{-1}(e)$ is a codimension one submanifold of P. Assume that the Hamiltonian vector field X_H has a *complete* flow. The flow determines a symplectic action for $\mathbb{R}$ on P with momentum mapping H. Thus by the Symplectic Reduction Theorem, $H^{-1}(e)/\mathbb{R}$, if a manifold, is symplectic. ♦

Exercises

2.2-1

(a) *Construct the symplectic reduced space for the standard action of S^1 on $T^*\mathbb{R}^2$.*

(b) *Construct the symplectic reduced space for the action of S^1 on $T^*\mathbb{R}^3$ obtained by taking the cotangent lift of the action of S^1 on $\mathbb{R}^3$ given by rotations about the z-axis.*

(c) *Examine the construction of the symplectic reduced space for the standard action of $SO(3)$ on $T^*\mathbb{R}^3$ at the zero level of angular momentum.*

2.2-2 *One can imagine carrying out reduction for a nonequivariant momentum map associated with a group G by extending the group to the central extension Σ as was described in "An Introduction to Mechanics and Symmetry" (IMS). In doing so, one would need to compute the isotropy for the action of Σ. Describe this isotropy. Carry out this construction for the action of $SE(2)$ on $\mathbb{R}^2 \times \mathbb{R}^2 \times ... \times \mathbb{R}^2$ given by n copies of the standard action (where n is a given positive integer). Consult §12.6 of IMS for some related remarks.*

2.2-3 *Is it true that if a Lie group G acts on a product of two symplectic manifolds, then the reduced space of the product is the product of the reduced spaces ?*

2.3 Coadjoint Orbits as Symplectic Reduced Spaces

We now show that coadjoint orbits may be realized as reduced spaces. This provides an alternative proof that they are symplectic manifolds (*IMS*, Chapter **14**). The strategy is to show that the (minus) coadjoint symplectic form on a coadjoint orbit $\mathcal{O}$ at a point $\nu \in \mathcal{O}$, namely

$$\omega_\nu^-(\{\mathrm{ad}_\xi^*\nu, \mathrm{ad}_\eta^*\nu) = -\langle \nu, [\xi, \eta]\rangle \tag{2.3.1}$$

may be obtained by means of the symplectic reduction theorem. The following theorem formulates the result for left actions; of course there is a similar one for right actions.

Theorem 2.3.1 (Coadjoint Orbit Reduction) *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 actions of G are 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^-.*

Proof. Recall that $\mathbf{J}_L$ is given by *right* translation to the identity:

$$\mathbf{J}_L(\alpha_g) = T_e^* R_g \cdot \alpha_g \tag{2.3.2}$$

Thus, $\mathbf{J}_L^{-1}(\mu)$ consists of those $\alpha_g \in P = T^*G$ such that $\alpha_g = T_g^* R_{g^{-1}} \cdot \mu$. In other words, if we extend μ to a *right* invariant one 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 (*IMS, 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.,* $\varphi = J_R | J_L^{-1}(\mu)$, 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 $\alpha_{hg} = T_g L_h \cdot \alpha_g$ 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 readilly 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 rght action. By the chain rule,

$$d_\xi'(0) = T\Psi_g \cdot \xi_P(\nu).$$

Thus,

$$\Omega(d_\xi'(0), d_\eta'(0)) = \Omega(\xi_P(\nu), \eta_P(\nu)),$$

since Ψ_g is a symplectic map. Next, recall that $\Omega = -\mathbf{d}\Theta$ where Θ is the canonical one form, so

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

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

$$\Omega(\xi_P(\nu), \eta_P(\nu)) = -\xi_P[\Theta(\eta_P)](\nu) + \eta_P[\Theta(\xi_P)](\nu) + \Theta([\xi_P, \eta_P])(\nu).$$

Since we have a *right* action, $[\xi_P, \eta_P] = [\xi, \eta]_P$. The definition of the canonical one form shows that on infinitesimal generators, it is given by $\Theta(\eta_P) = \langle \mathbf{J}_R, \eta \rangle$, so the preceding displayed expression equals

$$-\{\langle \mathbf{J}_R, \eta \rangle, \langle \mathbf{J}_R, \xi \rangle\}(\nu) + \{\langle \mathbf{J}_R, \xi \rangle, \langle \mathbf{J}_R, \eta \rangle\}(\nu) + \Theta([\xi, \eta]_P)(\nu).$$

By equivarance 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 this result does *not* require μ to be a regular *point* in $\mathfrak{g}^*$; that is, arbitrarily nearby coadjoint orbits may have a different dimension.
2. The form ω^- on the orbit need not be exact even though Ω is. The example of $SO(3)$, whose orbits are spheres shows this. ♦

Exercises.

2.3-1 *Show that (under the assumptions of the symplectic reduction theorem) if $\mathcal{O} \subset \mathfrak{g}^*$ is a coadjoint orbit, then $\mathbf{J}^{-1}(\mathcal{O})$ is coisotropic.*

2.3-2 *Let* SO(3) *act on the space $Q = \mathrm{SO}(3) \times S^1$ by acting by left multiplication on the first factor and acting trivially on the second factor. (This space comes up in the dynamics of a rigid body with a single rotor). Lift the action to $P = T^*Q$ by cotangent lift. Compute the reduced space P_μ at each regular point $\mu \in \mathfrak{so}(3)^*$.*

2.3-3 *Suppose that G acts on a symplectic manifold P and that this action has an equivariant momentum map and that the assumptions of the symplectic reduction theorem hold. Consider the product space $\widetilde{P} = P \times T^*G$ with the product symplectic structure. Let G act on $\widetilde{P}$ with the diagonal action and let $\mathcal{O}$ denote the coadjoint orbit through μ. Show that the symplectic reduced space $\widetilde{P}_\mu$ need not be any of $P_\mu \times T^*G$, $P \times \mathcal{O}$, or $P_\mu \times \mathcal{O}$.*

2.4 Reducing Hamiltonian Systems

Let (P, Ω) be a symplectic manifold, $\Phi : G \times P \to P$ a symplectic Lie group action and $\mathbf{J} : P \to \mathfrak{g}^*$ an equivariant momentum mapping for this action. Let $\mu \in \mathfrak{g}^*$ and G_μ be the isotropy subgroup for μ. In this section, assume that μ is a regular (or weakly regular) value and that G_μ acts freely and properly so that the reduced phase space $P_\mu = \mathbf{J}^{-1}(\mu)/G_\mu$ is a symplectic manifold. Let $\pi_\mu : \mathbf{J}^{-1}(\mu) \to P_\mu$ and $i_\mu : \mathbf{J}^{-1}(\mu) \to P$ denote the canonical projection and inclusion.

Theorem 2.4.1 (Symplectic Reduction of Dynamics) *Let $H : P \to \mathbb{R}$ be a G-invariant Hamiltonian, i.e.,*

$$H \circ \Phi_g = H,$$

for all $g \in G$. Then the flow F_t of X_H leaves the set $\mathbf{J}^{-1}(\mu)$ invariant and commutes with the G_μ-action on $\mathbf{J}^{-1}(\mu)$, so it induces a flow F_t^μ on P_μ that satisfies

$$\pi_\mu \circ F_t = F_t^\mu \circ \pi_\mu.$$

This flow is Hamiltonian on P_μ with Hamiltonian function $H_\mu : P_\mu \to \mathbb{R}$ defined by

$$H_\mu \circ \pi_\mu = H \circ i_\mu.$$

*We call H_μ the **reduced Hamiltonian**. Furthermore, the vector fields X_H and X_{H_μ} are π_μ-related.*

Proof Since $\mathbf{J}$ is conserved by the flow of X_H, $\mathbf{J}^{-1}(\mu)$ is invariant under F_t. In addition, since H is G-invariant, the flow F_t commutes with the action and thus there is a well-defined flow F_t^μ on P_μ characterized by $F_t^\mu \circ \pi_\mu = \pi_\mu \circ F_t$. Using this, the relationships

$$F_t^* \Omega = \Omega, \ i_\mu \circ F_t \,|\, \mathbf{J}^{-1}(\mu) = F_t \circ i_\mu, \quad \text{and} \quad \pi_\mu^* \Omega_\mu = i_\mu^* \Omega,$$

we get

$$\pi_\mu^* (F_t^\mu)^* \Omega_\mu = F_t^* \pi_\mu^* \Omega_\mu = F_t^* i_\mu^* \Omega = i_\mu^* F_t^* \Omega = i_\mu^* \Omega = \pi_\mu^* \Omega_\mu,$$

whence $(F_t^\mu)^* \Omega_\mu = \Omega_\mu$, since π_μ is a surjective submersion. Thus the flow F_t^μ preserves the symplectic form Ω_μ.

The relation $H_\mu \circ \pi_\mu = H \circ i_\mu$ plus G-invariance of H defines $H_\mu : P_\mu \to \mathbb{R}$ uniquely. Let $z \in \mathbf{J}^{-1}(\mu)$, and write its class in P_μ as $[z] = \pi_\mu(z) \in P_\mu$, and similarly for a vector $v \in T_z(\mathbf{J}^{-1}(\mu))$, write $[v] = T_z \pi_\mu(v) \in T_{[z]} P_\mu$. Then

$$\begin{aligned}
\mathbf{d}H_\mu([z]) \cdot [v] &= \mathbf{d}H_\mu(\pi_\mu(z))(T\pi_\mu(v)) = \pi_\mu^*(\mathbf{d}H_\mu)(z)(v) \\
&= \mathbf{d}(H_\mu \circ \pi_\mu)(z)(v) = \mathbf{d}(H \circ i_\mu)(z)(v) = i_\mu^*(\mathbf{d}H)(z)(v) \\
&= i_\mu^*(\mathbf{i}_{X_H} \Omega)_z(v) = (\mathbf{i}_{X_H} \pi_\mu^* \Omega_\mu)_z(v) \\
&= (\pi_\mu^* \Omega_\mu)_z(X_H(z), v),
\end{aligned}$$

since $i_\mu^* X_H = X_H$ and $i_\mu^* \Omega = \pi_\mu^* \Omega_\mu$. Let Y denote the vector field on P_μ whose flow is F_t^μ; Y is, by construction, π_μ-related to X_H, *i.e.*, $T\pi_\mu \circ X_H = Y \circ \pi_\mu$, so that from the last equality above we get

$$\begin{aligned}
\mathbf{d}H_\mu[z]([v]) &= (\pi_\mu^* \Omega_\mu)_z(X_H(z), v) = (\Omega_\mu)_{[z]}(T_z\pi_\mu(X_H(z)), T_z\pi_\mu(v)) \\
&= \Omega_{\mu_{[z]}}(Y[z], [v]),
\end{aligned}$$

and therefore Y is the Hamiltonian vector field whose Hamiltonian function is H_μ. ■

The behavior of Poisson brackets under reduction is given by the following result.

Corollary 2.4.2 *If $H, K : P \to \mathbb{R}$ are G-invariant functions, then $\{H, K\}$ is also G-invariant and*

$$\{H, K\}_\mu = \{H_\mu, K_\mu\}_{P_\mu},$$

where $\{\,,\}_{P_\mu}$ denotes the Poisson bracket on P_μ and $\{H, K\}_\mu$ is the function induced on P_μ by $\{H, K\}$.

Proof. Since the action Φ is symplectic, and H and K are G-invariant functions, we get

$$\Phi_g^*\{H, K\} = \{\Phi_g^* H, \Phi_g^*\} = \{H, K\}.$$

If $z \in \mathbf{J}^{-1}(\mu)$, the symplectic reduction of dynamics theorem, the definition of the Poisson bracket, and the tangency of X_H, X_K to $\mathbf{J}^{-1}(\mu)$, give

$$\begin{aligned}
(\{H_\mu, K_\mu\}_{P_\mu} \circ \pi_\mu)(z) &= \Omega_\mu(\pi_\mu(z))(X_{H_\mu}(\pi_\mu(z)), X_{K_\mu}(\pi_\mu(z)) \\
&= \Omega_\mu(\pi(z))(T_z\pi_\mu(X_H(z)), T_z\pi_\mu(X_K(z))) \\
&= (\pi_\mu^*\Omega_\mu)(z)(X_H(z), X_K(z)) \\
&= (i_\mu^*\Omega)(z)(X_H(z), X_K(z)) = \Omega(X_H, X_K)(i_\mu(z)) \\
&= (\{H, K\} \circ i_\mu)(z),
\end{aligned}$$

i.e., $\{H, K\}_\mu = \{H_\mu, K_\mu\}_{P_\mu}$. $\blacksquare$

This corollary is important because it gives a method for computing reduced Poisson brackets. If the reduced manifold P_μ is determined, even without the explicit computation of Ω_μ, the Poisson brackets on P_μ can be determined directly as follows. Let h and k be functions on P_μ and let $\overline{h} = h \circ \pi_\mu$ and $\overline{k} = k \circ \pi_\mu$ be their lifts to $\mathbf{J}^{-1}(\mu)$. Now extend both $\overline{h}$ and $\overline{k}$ arbitrarily to G-invariant functions (*not* G_μ-invariant ones!) H and K on P. By construction, $H_\mu = h$ and $K_\mu = k$, so that

$$\{h, k\}_{P_\mu} = \{H_\mu, K_\mu\}_{P_\mu} = \{H, K\}_\mu.$$

It follows that the right hand side is independent of the extensions. This says: compute $\{H, K\}$ on P and re-express the results on P_μ to get the function $\{h, k\}_{P_\mu}$.

Exercises.

2.4-1

(a) *Compute the symplectic reduced spaces and the reduced symplectic forms for the action of S^1 on $\mathbb{C}^2$ given by $(z_1, z_2) \mapsto (e^{i\theta} z_1, e^{i\theta} z_2)$. (See §10.7 of IMS). Repeat the problem using the action $(z_1, z_2) \mapsto (e^{i\theta} z_1, e^{-i\theta} z_2)$.*

(b) *Verify directly that the symplectic reduced spaces you found in (a) are the symplectic leaves in the Poisson reduced space $\mathbb{C}^2/S^1$.*

2.4-1 *Give an example of a reduced space P_μ for which the brackets cannot be computed by pulling back functions to $\mathbf{J}^{-1}(\mu)$ and extending them to G_μ invariant functions.*

2.5 Orbit Reduction

Consider a coadjoint orbit $\mathcal{O} = G \cdot \mu$ through a fixed regular element μ and endowed with the "+" orbit symplectic form. If $\mathbf{J} : P \to \mathfrak{g}^*$ is an equivariant momentum map, we claim that

$$P_{\mathcal{O}} := \mathbf{J}^{-1}(\mathcal{O})/G \text{ is in one-to-one correspondence with } P_\mu = \mathbf{J}^{-1}(\mu)/G_\mu.$$

To see this, first note that by equivariance of $\mathbf{J}$, G leaves the set $\mathbf{J}^{-1}(\mathcal{O})$ invariant, so $\mathbf{J}^{-1}(\mathcal{O})/G$ is defined. If we denote elements of $\mathbf{J}^{-1}(\mu)/G_\mu$ by $[z]_\mu$ for $z \in \mathbf{J}^{-1}(\mu)$ and elements of $\mathbf{J}^{-1}(\mathcal{O})/G$ by $[z]$ for $z \in \mathbf{J}^{-1}(\mathcal{O})$, we claim that we have a well defined map

$$L_\mu : [z]_\mu \mapsto [z]$$

of P_μ to $P_{\mathcal{O}}$ determined by the inclusion $l_\mu : \mathbf{J}^{-1}(\mu) \to \mathbf{J}^{-1}(\mathcal{O})$. To see that it is well defined, we must show that the class $[z]$ obtained is independent of the representative $z \in \mathbf{J}^{-1}(\mu)$ chosen. To do this, suppose that $[z]_\mu = [z']_\mu$ for z and z' in $\mathbf{J}^{-1}(\mu)$ implies $z = g \cdot z'$ for some $g \in G_\mu$. This clearly implies that we also have equivalence in G; that is, $[z] = [z']$ and so the map L_μ is well defined.

The map L_μ is one-to-one by a similar argument. Indeed, suppose that $[z] = [z']$ for z and z' in $\mathbf{J}^{-1}(\mu)$. This implies $z = g \cdot z'$ for some $g \in G$; by equivariance of $\mathbf{J}$, $g \in G_\mu$ and so $[z]_\mu = [z']_\mu$. Likewise, by **i** of the Reduction Lemma, this map is onto. Thus, set theoretically, the reduced phase space P_μ equals $P_{\mathcal{O}}$.

The symplectic form on $P_{\mathcal{O}}$ is a little trickier to sort out intrinsically. What we want to do is to find an intrinsic description of the symplectic form on the orbit reduced space $P_{\mathcal{O}}$ and then show that the above map between it and the symplectic reduced space P_μ is a symplectic diffeomorphism. It is characterized by the following result of Marle [1976] and Kazhdan, Kostant, and Sternberg [1978]. We follow the exposition of Marsden [1981].

Theorem 2.5.1 (Orbit Reduction Theorem) *Let μ be 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 in $\mathfrak{g}^*_+$ containing μ. If (in the infinite dimensional case) $(T_z\mathbf{J})^{-1}(T_{\mathbf{J}(z)}\mathcal{O})$ splits in T_zP for all $z \in \mathbf{J}^{-1}(\mathcal{O})$, then*

- **i** $\mathbf{J}$ *is transversal to $\mathcal{O}$ so $\mathbf{J}^{-1}(\mathcal{O})$ is a manifold (if μ is only a clean value of $\mathbf{J}$, assume in addition that $\mathbf{J}$ intersects $\mathcal{O}$ cleanly);*

- **ii** $\mathbf{J}^{-1}(\mathcal{O})/G$ *has a unique differentiable structure such that the canonical projection $\pi_{\mathcal{O}} : \mathbf{J}^{-1}(\mathcal{O}) \to \mathbf{J}^{-1}(\mathcal{O})/G$ is a surjective submersion;*

- **iii** *there is a unique symplectic structure $\Omega_{\mathcal{O}}$ on $\mathbf{J}^{-1}(\mathcal{O})/G$ such that*

$$i_{\mathcal{O}}^*\Omega = \pi_{\mathcal{O}}^*\Omega_{\mathcal{O}} + \mathbf{J}_{\mathcal{O}}^*\omega_{\mathcal{O}}^+$$

  where $i_{\mathcal{O}} : \mathbf{J}^{-1}(\mathcal{O}) \to P$ is the inclusion, $\mathbf{J}_{\mathcal{O}} = \mathbf{J}|\mathbf{J}^{-1}(\mathcal{O})$ regarded as a map of $\mathbf{J}^{-1}(\mathcal{O})$ to $\mathcal{O}$ and $\omega_{\mathcal{O}}^+$ is the "+" orbit symplectic structure on $\mathcal{O}$.

- **iv** *the map L_μ is a symplectic diffeomorphism of P_μ to $P_{\mathcal{O}}$.*

We illustrate these maps in the following diagram:

$$
\begin{tikzcd}
 & P & & \\
\mathbf{J}^{-1}(\mu) \arrow[ur, "i_\mu"] \arrow[rr, "\text{inclusion} = l_\mu"] \arrow[d, "\pi_\mu"] & & \mathbf{J}^{-1}(\mathcal{O}) \arrow[ul, "i_{\mathcal{O}}"'] \arrow[r, "\mathbf{J}_{\mathcal{O}}"] \arrow[d, "\pi_{\mathcal{O}}"] & \mathfrak{g}^*_+ \\
\mathbf{J}^{-1}(\mu)/G_\mu = P_\mu \arrow[rr, "L_\mu"] & & \mathbf{J}^{-1}(\mathcal{O})/G = P_{\mathcal{O}} &
\end{tikzcd}
$$

Proof. **i** Let

$$z \in \mathbf{J}^{-1}(\mathcal{O}), \quad i.e., \quad \mathbf{J}(z) = \mathrm{Ad}^*_{g^{-1}}\mu$$

for some $g \in G$. By equivariance, $g^{-1} \cdot z \in \mathbf{J}^{-1}(\mu)$ and, as μ is a regular value, the mapping $T_{g^{-1}\cdot z}\mathbf{J} : T_{g^{-1}\cdot z}P \to \mathfrak{g}^*$ is onto. Writing $\Phi(g, z) = \Phi_g(z) = g \cdot z$, we find

$$T_{g^{-1}\cdot z}\mathbf{J} \circ T_z\Phi_{g^{-1}} = T_z(\mathbf{J} \circ \Phi_{g^{-1}}) = T_z(\mathrm{Ad}^*_g \circ \mathbf{J}) = \mathrm{Ad}^*_g \circ T_z\mathbf{J} : T_zP \to \mathfrak{g}^*$$

is also onto and thus $T_z\mathbf{J}$ is onto. Thus, necessarily $(T_z\mathbf{J})(T_zP) + T_{\mathrm{J}(z)}(\mathcal{O}) = \mathfrak{g}^*$. Since by hypothesis $(T_z\mathbf{J})^{-1}(T_{\mathrm{J}(z)}(\mathcal{O}))$ splits in T_xP, it follows that $\mathbf{J}$ is transversal to $\mathcal{O}$.

Next we prove **ii**. If there is a differentiable structure on $\mathbf{J}^{-1}(\mathcal{O})/G$ such that the canonical projection $\pi_{\mathcal{O}}$ is a submersion, then as in *MTA* §3.5, the map L_μ is smooth. To show that its inverse is also smooth, consider the map $z \in \mathbf{J}^{-1}(\mathcal{O}) \mapsto [g^{-1} \cdot z]_\mu \in P_\mu$, where $g \in G$ is such that $\mathbf{J}(z) = \mathrm{Ad}^*_{g^{-1}}\mu$. One checks that this map is well-defined, constant on G-orbits and thus induces a map on the quotient $\mathbf{J}^{-1}(\mathcal{O})/G \to P_\mu$; this induced map is L_μ. The following argument shows that this map is also smooth. Locally, $\mathbf{J}^{-1}(\mu)$ is diffeomorphic to $P_\mu \times G_\mu$, and locally $\mathbf{J}^{-1}(\mathcal{O})$ is diffeomorphic to $\mathbf{J}^{-1}(\mu) \times (G/G_\mu)$, *i.e.*, to $P_\mu \times G_\mu \times (G/G_\mu)$ which locally is diffeomorphic to $P_\mu \times G$. These local maps can be chosen to be G_μ-equivariant, so the composition of $z \mapsto [\Phi_{g^{-1}}(z)]_\mu$ with them equals the projection $P_\mu \times G \to P_\mu$. Consequently, the induced map on the quotient is also smooth. Thus the differentiable structure of $\mathbf{J}^{-1}(\mathcal{O})/G$ is uniquely determined by the requirement that L_μ be a diffeomorphism.

For the proof of **iii** we need the following.

Lemma 2.5.2 *Under the hypotheses of* **2.5.1***, 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**. Since $\mathbf{J}$ intersects $\mathcal{O}$ cleanly, $\mathbf{J}^{-1}(\mathcal{O})$ is a submanifold whose tangent space 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 $\mathbf{J}$,

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

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

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

which proves **ii**. ▼

Now we are ready to prove **iii** of the theorem. Using **1.5.2i**, 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 **1.5.2ii**, the defining relation for $\Omega_{\mathcal{O}}$ in the statement of **iii** 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 $\ker(T_z\mathbf{J})$ (see the reduction lemma **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 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 (weakly) nondegenerate as in the symplectic reduction theorem but since this follows from **iv**, we will not give this direct proof.

To prove **iv**, notice that the relation $L^*_\mu\Omega_{\mathcal{O}} = \Omega_\mu$ is equivalent to $\pi^*_\mu L^*_\mu\Omega_{\mathcal{O}} = i^*_\mu\Omega$. Since $L_\mu \circ \pi_\mu = \pi_{\mathcal{O}} \circ l_\mu$, this says that $l^*_\mu\pi^*_{\mathcal{O}}\Omega_{\mathcal{O}} = i^*_\mu\Omega$. By **iii** we have

$$l^*_\mu\pi^*_{\mathcal{O}}\Omega_{\mathcal{O}} = l^*_\mu(i^*_{\mathcal{O}}\Omega - \mathbf{J}^*_{\mathcal{O}}\omega^+_{\mathcal{O}}) = (i_{\mathcal{O}}\circ l_\mu)^*\Omega - (\mathbf{J}_{\mathcal{O}}\circ l_\mu)^*\omega^+_{\mathcal{O}} = i^*_\mu\Omega$$

since $i_{\mathcal{O}} \circ l_\mu = i_\mu$ and $\mathbf{J}_{\mathcal{O}} \circ l_\mu = \mu$ 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 not used. 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. For example, as we remarked earlier, a slice theorem for the G_μ action can be used to show that P_μ is a manifold.
3. The hypothesis that $(T_z\mathbf{J})^{-1}(T_{\mathrm{J}(z)}\mathcal{O}_\mu)$ splits in T_zP for all $z \in \mathbf{J}^{-1}(\mathcal{O})$ can be dropped for Hilbert (and hence finite dimensional) manifolds since this subspace is always closed. In the Banach space case or when $\mathfrak{g}^*$ does not literally mean the Banach space dual, this condition is most naturally obtained via elliptic regularity from a Fredholm alternative theorem. We shall now sketch in an abstract setting, the most

common situation. Let E, F be Banach spaces and assume that on both of them there are continuous bilinear inner products $\langle \cdot\,,\cdot\rangle\,,\langle \cdot\,,\cdot\rangle$. The norm given by these inner products is possibly not complete, so gives a topology weaker than the original one. Consider a continuous linear operator $A : E \to F$. Suppose it has a continuous linear adjoint $A^* : F_1 \subset F \to E$, where F_1 is a Banach space continuously included in F; thus, $\langle Ae, f\rangle = \langle e, A^*f\rangle$ for $e \in E$ and $f \in F_1$.

We say that A is ***E-splitting*** if, whenever $F = S \oplus T$ where $S \perp T$, for S and T closed subspaces of F, we have

$$E = A^{-1}(S) \oplus A^*(T \cap F_1).$$

Likewise, A is ***F-splitting*** if, whenever $E = U \oplus V$ where U and V are closed orthogonal subspaces of E, we have

$$F = A(U) \oplus (A^{*-1})(V).$$

Notice that $A^{-1}(S)$ and $A^*(T \cap F_1)$ are automatically orthogonal; however, it is not automatic that $A^*(T \cap F_1)$ is closed nor that the direct sum is E. In many cases, E and F are spaces of functions or tensors, and $\langle \cdot\,,\cdot\rangle$ is the L^2 inner product. Then the splitting property can be proved if either A of A^* is an elliptic operator. See Marsden and Hughes [1983], p. 320.

4. Although the description of the symplectic structure on $\mathbf{J}^{-1}(\mathcal{O})/G$ is not as simple as it was for $\mathbf{J}^{-1}(\mu)/G$, we shall see later that the description in terms of Poisson brackets is simpler on $\mathbf{J}^{-1}(\mathcal{O})/G$. We shall also see that the symplectic structure depends only on the orbit $\mathcal{O}$ and not on the choice of a point μ on it.

5. (**Homogeneous symplectic manifolds**) Assume that P is a symplectic manifold on which the Lie group G acts transitively and has an equivariant momentum map $\mathbf{J}$. Then $\mathbf{J}(P) = \mathcal{O}_\mu$, where $\mu = \mathbf{J}(x_0)$, and $x_0 \in P$, and $\mathbf{J} : P \to \mathcal{O}_\mu$ is a symplectic local diffeomorphism; if $\mathbf{J}$ is proper, then it is a symplectic covering map. By the Orbit Reduction Theorem, the reduced space $P_\mu = \mathbf{J}^{-1}(\mathcal{O}_\mu)/G = P/G$ is a point.

   We want to investigate the structure of P further, taking advantage of the transitive Lie group action. Fix a point $p \in P$ and define a map $\Phi_p : G/G_p \to P$ by $\Phi_p[g] = g \cdot p$, where $G_p = \{g \in G \mid g \cdot p = p\}$ and $[g] = \gamma_p$ is an element of the quotient G/G_p. The map Φ_p is well-defined, onto, and smooth, since $g \mapsto g \cdot p$ is a smooth map of G onto P. It is easy to check that Φ_p is also one-to-one thus defining a bijective continuous map of G/G_p onto P. If in addition Φ_p is open or closed (e.g. proper), then it is a homeomorphism. It can be shown that Φ_p is in fact always an immersion (See Abraham and Marsden [1978], p. 265) and hence, if in addition Φ_p is open or closed, it is a diffeomorphism.

   In this manner we are led to the study of homogeneous symplectic manifolds, *i.e.*, of manifolds of the type G/H, H a closed Lie subgroup of G, where G/H is symplectic. We refer the reader to Guillemin and Sternberg [1984], Chapter 2, for an account of the known facts about these manifolds. ♦

Exercises.

2.5-1 *Specialize the equation*

$$i_{\mathcal{O}}^*\Omega = \pi_{\mathcal{O}}^*\Omega_{\mathcal{O}} + J_{\mathcal{O}}^*\omega_{\mathcal{O}}^+$$

to the case of Abelian groups.

2.5-2

(a) *Show directly that for $P = T^*G$, the space $P_{\mathcal{O}}$ is diffeomorphic to the orbit $\mathcal{O}$ itself .*

(b) *Specialize the equation*

$$i_{\mathcal{O}}^*\Omega = \pi_{\mathcal{O}}^*\Omega_{\mathcal{O}} + J_{\mathcal{O}}^*\omega_{\mathcal{O}}^+$$

to the case $P = T^*G$.

2.6 Foliation Orbit Reduction

We now discuss how to get the symplectic structure on $\mathbf{J}^{-1}(\mathcal{O})/G$ starting from the foliation reduction theorem or its corollary (**2.1.6**). Let $\mathcal{O}$ be a coadjoint orbit in $\mathfrak{g}^*$ and let $N = \mathbf{J}^{-1}(\mathcal{O})$ and $z \in N$. By **1.5.2**, $T_zN = T_zG \cdot z + \ker(T_z\mathbf{J})$, so

$$\begin{aligned}
(T_zN)^\Omega &= (T_z(G \cdot z) + \ker(T_z\mathbf{J})^\Omega \\
&= (T_z(G \cdot z))^\Omega \cap \ker(T_z\mathbf{J})^\Omega \\
&= \ker T_z\mathbf{J} \cap T_z(G \cdot z) \quad \text{(by } \mathbf{2.2.4} \text{ and the Reduction Lemma } \mathbf{iv}) \\
&= T_z(G_\mu \cdot z) \qquad\qquad \text{(by the Reduction Lemma } \mathbf{iii})
\end{aligned}$$

where $\mu = \mathbf{J}(z)$. Therefore, the characteristic distribution of $j^*\Omega$, where $j : N \to P$ is the inclusion, equals

$$T_zN \cap (T_zN)^\Omega = (T_z(G \cdot z) + \ker T_z\mathbf{J}) \cap T_z(G_\mu \cdot z) = T_z(G_\mu \cdot z)$$

and hence the leaves of the null-foliation Φ are the G_μ-orbits in N. Since the groups G_μ as μ ranges in $\mathcal{O}$ are conjugate, N/Φ is a manifold and the projection $\Pi : N \to N/\Phi$ is a submersion. Thus, N/Φ is the reduced symplectic manifold. Note that it is obtained by identifying all points which are in the same orbit level set of $\mathbf{J}$ and, simultaneously, on the same G- orbit in P. We claim that *there is a symplectic diffeomorphism*

$$\varphi : N/\Phi \to P_{\mathcal{O}} \times \mathcal{O}^+,$$

where $P_{\mathcal{O}} = \mathbf{J}^{-1}(\mathcal{O})/G$. Indeed, denoting the equivalence classes in N/Φ by $[[z]]$, the map $\varphi : N/\Phi \to P_{\mathcal{O}} \times \mathcal{O}^+$, defined by $\varphi([[z]]) = ([z],\ \mathbf{J}(z)$ is well-defined and has inverse

$$\psi : P_{\mathcal{O}} \times \mathcal{O}^+ \to N/\Phi \ \text{ given by } \ \psi([z], \nu) = [[g \cdot z]], \tag{2.6.1}$$

where $g \in G$ is such that $\mathrm{Ad}_{g^{-1}}^*\nu = \mathbf{J}(z)$. By the general submersion arguments (see *MTA*), these maps are smooth.

To show that φ is symplectic, let $\mu \in \mathcal{O}$ so that $\mathcal{O} = \mathcal{O}_\mu = \mathrm{Orb}(\mu)$ and denote by $\Omega_{\mathcal{O}}$ the symplectic form on N/Φ. Since the diagram

$$\begin{tikzcd}
N \arrow[r, "\mathrm{id} \times \mathbf{J}_{\mathcal{O}}"] \arrow[d, "\pi"'] & N \times O \arrow[d, "\pi_{\mathcal{O}} \times \mathrm{id}_{\mathcal{O}}"] \\
N/\Phi \arrow[r, "\varphi"'] & P_{\mathcal{O}} \times \mathcal{O}
\end{tikzcd}$$

is commutative, the formula

$$i_{\mathcal{O}}^* \Omega = \pi_{\mathcal{O}}^* \Omega_{\mathcal{O}} + \mathbf{J}_{\mathcal{O}}^* \omega_{\mathcal{O}}^+ \tag{2.6.2}$$

from the orbit reduction theorem implies

$$\begin{aligned} \pi^* \varphi^* (\Omega_{\mathcal{O}} + \omega_{\mathcal{O}}^+) &= (\pi_{\mathcal{O}} \times \mathbf{J}_{\mathcal{O}})^* (\Omega_{\mathcal{O}} + \omega_{\mathcal{O}}^+) \\ &= \pi_{\mathcal{O}}^* \Omega_{\mathcal{O}} + \mathbf{J}_{\mathcal{O}}^* \omega_{\mathcal{O}}^+ = i_{\mathcal{O}}^* \Omega. \end{aligned} \tag{2.6.3}$$

It follows from the foliation reduction theorem that $\varphi^*(\Omega_{\mathcal{O}} + \omega_{\mathcal{O}}^+) = \Omega_{\mathcal{O}}$ and so φ is symplectic.

Theorem 2.6.1 (Foliation Orbit Reduction Theorem) *Assume that the hypothesis in the symplectic reduction theorem hold and let N/Φ be the foliation reduced symplectic space of $\mathbf{J}^{-1}(\mathcal{O}) \subset P$ as constructed in §1.1. By construction, the map φ defined by $\varphi : N/\Phi \to P_{\mathcal{O}} \oplus \mathcal{O}$, $\varphi([[z]]) = ([z], \mathbf{J}(z))$, where $P_{\mathcal{O}} \oplus \mathcal{O}^+ = P_{\mathcal{O}} \times \mathcal{O}^+$ with the sum symplectic structure, is a symplectic diffeomorphism. Moreover, the symplectic structure on $P_{\mathcal{O}}$ is uniquely determined by the requirement that φ be a symplectic diffeomorphism.*

Proof. We have proved all but the last statement. Let Σ be another symplectic form on $P_{\mathcal{O}}$ such that φ is a symplectic diffeomorphism. Then

$$\varphi^*(\Sigma + \omega^+) = \Omega_{\mathcal{O}}, \tag{2.6.4}$$

so that

$$\begin{aligned} i^* \Omega &= \pi_{\mathcal{O}}^* \Omega_{\mathcal{O}} = \pi^* \varphi^* (\Sigma + \omega_{\mathcal{O}}^+) = (\pi_{\mathcal{O}} \times \mathbf{J}_{\mathcal{O}})^* (\Sigma + \omega_{\mathcal{O}}^+) \\ &= \pi_{\mathcal{O}}^* \Sigma + \mathbf{J}_{\mathcal{O}}^* \omega_{\mathcal{O}}^+. \end{aligned} \tag{2.6.5}$$

By the orbit reduction theorem, we conclude from (1.5.5) that

$$\pi_{\mathcal{O}}^* (\Sigma - \Omega_{\mathcal{O}}^+) = 0$$

which implies $\Sigma = \Omega_{\mathcal{O}}^+$ because $\pi_{\mathcal{O}}$ is a surjective submersion. ■

2.7 The Shifting Theorem

We shall explain how reduction at a general point $\mu \in \mathfrak{g}^*$ can be replaced by reduction at $0 \in \mathfrak{g}^*$ at the expense of enlarging the symplectic manifold. We assume we are in the situation of the Symplectic Reduction Theorem so that we can form the reduced phase space P_μ. Let $\mathcal{O}$ be the coadjoint orbit through μ endowed with the + orbit symplectic structure. The group G acts canonically on the left on $\mathcal{O}$ via the coadjoint action, and the equivariant momentum map of this action is the inclusion map $i : \mathcal{O} \to \mathfrak{g}^*$. Let $P \ominus \mathcal{O}$ denote $P \times \mathcal{O}$ with the symplectic structure $\Omega - \omega_{\mathcal{O}}^+ := \pi_1^* \Omega - \pi_2^* \omega_\mu^+$ where $\pi_1 : P \times \mathcal{O} \to P$ and $\pi_2 : P \times \mathcal{O} \to \mathcal{O}$ are the projections. Then G acts canonically on $P \ominus \mathcal{O}$ by

$$(z, \nu) \mapsto (\Phi_g(z), \mathrm{Ad}_{g^{-1}}^* \nu), \ \text{ where } \ z \in P \ \text{ and } \ \nu \in \mathcal{O}.$$

This action has an equivariant momentum map given by $\mathbf{J} - i : P \ominus \mathcal{O} \to \mathfrak{g}^*$.

Theorem 2.7.1 (Shifting Theorem) *Under the hypotheses of the orbit reduction theorem, the reduced manifolds P_μ and $(P \ominus \mathcal{O})_0$ are symplectically diffeomorphic.*

Proof. The smooth map $\varphi : \mathbf{J}^{-1}(\mathcal{O}) \to (\mathbf{J}-i)^{-1}(0)$ defined by $\varphi(z) = (z, \mathbf{J}(z))$ is easily seen to be G-equivariant, *i.e.*,

$$\varphi(\Phi_g(z)) = (\Phi_g \times \mathrm{Ad}^*_{g^{-1}}\nu)(\varphi(z)) \tag{2.7.1}$$

so that it induces a smooth map $\Phi : \mathbf{J}^{-1}(\mathcal{O})/G \to (\mathbf{J}-i)^{-1}(0)/G$. Similarly, the smooth map $\psi : (\mathbf{J}-i)^{-1}(0) \to \mathbf{J}^{-1}(\mathcal{O})$ defined by $\psi(z,\nu) = z$ is also equivariant so it induces a smooth map $\Psi : (P \ominus \mathcal{O})_0 \to \mathbf{J}^{-1}(\mathcal{O})/G$ which is easily seen to be the inverse of Φ, so Φ is a diffeomorphism. By remark 4 following the symplectic reduction theorem and 2 following the orbit reduction theorem, $(P \ominus \mathcal{O})_0$ is a symplectic manifold.

We summarize the mappings involved in this proof in the following commutative diagram.

$$\begin{tikzcd}
 & & \mathbf{J}^{-1}(\mu) \arrow[d, "i_\mu"] & & & \\
\mathcal{O}^+ \arrow[r, "i"] & \mathfrak{g}^*_+ & \mathbf{J}^{-1}(\mathcal{O}) \arrow[l, "\mathbf{J}_{\mathcal{O}}"'] \arrow[r, shift left, "\varphi"] \arrow[d, "\pi_{\mathcal{O}}"] & (\mathbf{J}-i)^{-1}(0) \arrow[l, shift left, "\psi"] \arrow[r, "i_0"] \arrow[d, "\pi_0"] & P \ominus \mathcal{O} \\
 & & P_{\mathcal{O}} = \mathbf{J}^{-1}(\mathcal{O})/G \arrow[r, shift left, "\Phi"] & (P \ominus \mathcal{O})_0 \arrow[l, shift left, "\Psi"] &
\end{tikzcd}$$

To show that Φ is symplectic, let $i_0 : (\mathbf{J}-i)^{-1}(0) \to P \ominus \mathcal{O}$ be the inclusion and $\pi_0 : (\mathbf{J}-i)^{-1}(0) \to (P \ominus \mathcal{O})_0$ the projection. Let σ_0 be the symplectic form on $(P \ominus \mathcal{O})_0$. With the notations of the orbit reduction theorem, the defining relation for $\Omega_{\mathcal{O}}$ is

$$\pi^*_{\mathcal{O}}\Omega_{\mathcal{O}} = i^*_{\mathcal{O}}\Omega - \mathbf{J}^*_{\mathcal{O}}\omega^+_{\mathcal{O}}.$$

By the symplectic reduction theorem, σ_0 is characterized by

$$\pi^*{}_0\sigma_0 = i^*_0(\Omega - \omega^+_{\mathcal{O}});$$

therefore, since $\pi_0 \circ \varphi = \Phi \circ \pi_{\mathcal{O}}$,

$$\pi^*_{\mathcal{O}}\Phi^*\sigma_0 = (\Phi \circ \pi_{\mathcal{O}})^*\sigma_0 = (\pi_0 \circ \varphi)^*\sigma_0 = \varphi^*\pi^*_0\sigma_0 = \varphi^* i^*_0(\Omega - \omega^+_{\mathcal{O}}),$$

so that Φ is symplectic if and only if

$$i^*_{\mathcal{O}}\Omega - \mathbf{J}^*_{\mathcal{O}}\omega^+_{\mathcal{O}} = \varphi^* i^*_0(\Omega - \omega^+_{\mathcal{O}}). \tag{2.7.2}$$

Formula (2.7.2) is proved in the following way. For $z \in \mathbf{J}^{-1}(\mathcal{O})$ and $u, v \in T_z(\mathbf{J}^{-1}(\mathcal{O})) = (T_z\mathbf{J})^{-1}(T_\mu\mathcal{O})$, where $\mu = J(z)$, we have

$$\begin{aligned}
(i^*_{\mathcal{O}}\Omega - \mathbf{J}^*_{\mathcal{O}}\omega^+_{\mathcal{O}})_z(u,v) &= \Omega_z(u,v) - (\omega^+_{\mathcal{O}})_{\mathrm{J}(z)}(T_z\mathbf{J}(u), T_z\mathbf{J}(v)) \\
&= (\Omega - \omega^+_{\mathcal{O}})_{(z,\mathrm{J}(z))}((u, T_z\mathbf{J}(u)), (v, T_z\mathbf{J}(v))) \\
&= (\varphi^* i^*_0(\Omega - \omega^+_{\mathcal{O}}))_z(u,v).
\end{aligned}$$

The stated result now follows. ■

Since $P \ominus \mathcal{O}^+$ is obviously symplectically diffeomorphic to $P \oplus \mathcal{O}^-$, we have the symplectic diffeomorphisms

$$P_\mu \approx \mathbf{J}^{-1}(\mathcal{O})/G \approx (P \ominus \mathcal{O}^+)_0 \approx (P \oplus \mathcal{O}^-)_0$$

where the reductions are on the left. For right actions and right reductions, this sequence of symplectic diffeomorphisms becomes

$$P_\mu \approx \mathbf{J}^{-1}(\mathcal{O})/G \approx (P \oplus \mathcal{O}^-)_0 \approx (P \ominus \mathcal{O}^-)_0.$$

2.8 Dynamics via Orbit Reduction

The dynamic counterpart of the Orbit Reduction Theorem is the following. Recall that $\pi_{\mathcal{O}} : \mathbf{J}^{-1}(\mathcal{O}^+) \to \mathbf{J}^{-1}(\mathcal{O}^+)/G$ and $i_{\mathcal{O}} : \mathbf{J}^{-1}(\mathcal{O}^+) \to P$ denote the canonical projection and inclusion.

Theorem 2.8.1 (Dynamic Orbit Reduction Theorem) *Let $H : P \to \mathbb{R}$ be G-invariant and $\mathcal{O} \subset \mathfrak{g}^*$ be a coadjoint orbit. Then the flow of X_H leaves $\mathbf{J}^{-1}(\mathcal{O})$ invariant and it commutes with the G-action on $\mathbf{J}^{-1}(\mathcal{O})$, so it induces a flow on $\mathbf{J}^{-1}(\mathcal{O})/G$. With respect to the reduced symplectic structure on $\mathbf{J}^{-1}(\mathcal{O})/G$, this flow is Hamiltonian, with Hamiltonian function $H_{\mathcal{O}}$ determined by $H_{\mathcal{O}} \circ \pi_{\mathcal{O}} = H \circ i_{\mathcal{O}}$. The Hamiltonian vector fields X_H and $X_{H_{\mathcal{O}}}$ are $\pi_{\mathcal{O}}$-related. Moreover, if $K : P \to \mathbb{R}$ is another G-invariant function, then $\{H, K\}$ is also G-invariant and*

$$\{H, K\}_{\mathcal{O}} = \{H_{\mathcal{O}}, K_{\mathcal{O}}\}_{P_{\mathcal{O}}},$$

where $\{\,,\}_{P_{\mathcal{O}}}$ denotes the Poisson bracket on $P_{\mathcal{O}} = \mathbf{J}^{-1}(\mathcal{O})/G$.

Proof. If $\mu \in \mathcal{O} = \mathcal{O}_\mu$, recall that P_μ is symplectically diffeomorphic to $\mathbf{J}^{-1}(\mathcal{O}^+)/G$; a symplectic diffeomorphism is the map $L_\mu : P_\mu \to \mathbf{J}^{-1}(\mathcal{O})/G$ defined by $L_\mu([z]) = [z]$, where $[z]$ denotes an element of $\mathbf{J}^{-1}(\mathcal{O})/G$ and $l_\mu : \mathbf{J}^{-1}(\mu) \to \mathbf{J}^{-1}(\mathcal{O})$ is the inclusion. The theorem is proved once we show that the induced flow on $\mathbf{J}^{-1}(\mathcal{O})/G$ and the function $H_{\mathcal{O}}$ are the push-forwards of F_t^μ and H_μ by L_μ. Referring to the commutative diagram in **2.5.1**, we get

$$H_{\mathcal{O}} \circ L_\mu \circ \pi_\mu = H_{\mathcal{O}} \circ \pi_{\mathcal{O}} \circ l_\mu = H \circ i_\mu = H_\mu \circ \pi_\mu,$$

so that $H_{\mathcal{O}} \circ L_\mu = H_\mu$. Similarly, if the flow F_t of X_H on P induces the flow $F_t^{\mathcal{O}}$ on $\mathbf{J}^{-1}(\mathcal{O})/G$, we have

$$F_t \circ l_\mu = l_\mu \circ F_t|\mathbf{J}^{-1}(\mu)$$

since $F_t(\mathbf{J}^{-1}(\mu)) = \mathbf{J}^{-1}(\mu)$, so that

$$\begin{aligned} F_t^{\mathcal{O}} \circ L_\mu \circ \pi_\mu &= F_t^{\mathcal{O}} \circ \pi_{\mathcal{O}} \circ l_\mu = \pi_{\mathcal{O}} \circ F_t \circ l_\mu \\ &= \pi_{\mathcal{O}} \circ l_\mu \circ F_t = L_\mu \circ \pi_\mu \circ F_t = L_\mu \circ F_t^\mu \circ \pi_\mu, \end{aligned}$$

i.e.,

$$F_t^{\mathcal{O}} \circ L_\mu = L_\mu \circ F_t^\mu. \quad \blacksquare$$

As before, to compute the Poisson bracket on $\mathbf{J}^{-1}(\mathcal{O})/G$, one takes two functions $h, k : \mathbf{J}^{-1}(\mathcal{O})/G \to \mathbb{R}$, lifts them to $\overline{h} = h \circ \pi_{\mathcal{O}}$, $\overline{k} = k \circ \pi_{\mathcal{O}} : \mathbf{J}^{-1}(\mathcal{O}) \to \mathbb{R}$, extends them arbitrarily to G-invariant functions $H, K : P \to \mathbb{R}$, computes their Poisson bracket $\{H, K\}$ on P, and finally re-expresses the result purely in terms of $\mathbf{J}^{-1}(\mathcal{O})/G$; the theorem guarantees that this last step is always possible.
The conclusion in the result about brackets has an important consequence.

Theorem 2.8.2 (Stratification-Reduction Theorem) *Under the conditions of the orbit reduction theorem, the symplectic leaves of the Poisson manifold P/G are given by the orbit reduced spaces $P_{\mathcal{O}} = J^{-1}(\mathcal{O})/G$.*

Proof Obviously $P_{\mathcal{O}}$ embeds naturally into P/G and, as we have seen, $P_{\mathcal{O}}$ is symplectic. To complete the proof, we just need to show that if f, k are functions on $P_{\mathcal{O}}$ and they are extended arbitrarily to functions f, k on P/G, then $\{F, K\} = \{f, k\}$ at points of $P_{\mathcal{O}}$. But this follows directly from the property of brackets given in the dynamic orbit reduction theorem. ■

In *IMS* we gave a direct proof of this result for the case $P = T^*G$ in which case $P/G \cong \mathfrak{g}^*$ is stratified by the coadjoint orbits.

Next we look at dropping dynamics from the point of view of the shifting theorem.

Theorem 2.8.3 (Dynamic Shifting Theorem) *Let $H : P \to \mathbb{R}$ be a G-invariant Hamiltonian and $\mathcal{O} \subset \mathfrak{g}^*$ a coadjoint orbit. Then H induces Hamiltonian functions $\overline{H} : P \ominus \mathcal{O}^+ \to \mathbb{R}$ by $\overline{H}(x, \nu) = H(x)$, for all $x \in P$, $\nu \in \mathcal{O}^+$ and $\overline{H}_0$ on $(P \ominus \mathcal{O}^+)_0$. If F_t is the flow of X_H on P, then $\overline{F}_t(x, \nu) = (F_t(x), \nu)$ is the flow of $X_{\overline{H}}$ on $P \ominus \mathcal{O}^+$ and the flow $\overline{F}_t^\circ$ of $\overline{H}_0$ is determined by $\Phi \circ F_t = \overline{F}_t^0 \circ \Phi$. Moreover, if K is another G-invariant function, then $\{H, K\}$ is G-invariant and $\{\overline{H, K}\}_0 = \{\overline{H}_0, \overline{K}_0\}^0$.*

Proof. From the Shifting Theorem, the map $\Phi : \mathbf{J}^{-1}(\mathcal{O}^+)/G \to (P \ominus \mathcal{O}^+)_0$ is a symplectic diffeomorphism induced by $\phi(z) = (z, \mathbf{J}(z))$. The statements regarding $\overline{H}$ and $X_{\overline{H}}$ are obvious. The theorem is proved if we show that the flow $F_t^{\mathcal{O}}$ and the Hamiltonian $H_{\mathcal{O}}$ are pushed forward to $\overline{F}_t^0$ and $\overline{H}_0$ on $(P \ominus \mathcal{O}^+)_0$. Referring to the commutative diagram in the proof of **2.5.1** and recalling that $F_t^{\mathcal{O}} \circ \pi_{\mathcal{O}} = \pi_{\mathcal{O}} \circ F_t$, we get

$$\Phi \circ F_t^{\mathcal{O}} \circ \pi_{\mathcal{O}} = \Phi \circ \pi_{\mathcal{O}} \circ F_t = \pi_0 \circ \phi \circ F_t,$$

and

$$\overline{F}_t^0 \circ \Phi \circ \pi_{\mathcal{O}} = \overline{F}_t^0 \circ \pi_0 \circ \phi = \pi_0 \circ \overline{F}_t \circ \phi.$$

Thus,

$$\Phi \circ F_t^{\mathcal{O}} = \overline{F}_t^{\mathcal{O}} \circ \Phi,$$

since for any $z \in P$,

$$(\phi \circ F_t)(z) = (F_t(z), \mathbf{J}(F_t(z))) = (F_t(z), \mathbf{J}(z)) = \overline{F}_t(z, \mathbf{J}(z)) = (\overline{F}_t \circ \phi)(z),$$

$\mathbf{J}$ being a conserved quantity. Similarly, if $H : P \to \mathbb{R}$ is G-invariant, then

$$H_0 \circ \Phi \circ \pi_{\mathcal{O}} = \hat{H}_0 \circ \pi_0 \circ \phi = H \circ \phi = H_{\mathcal{O}} \circ \phi,$$

by **1.3.1** and **1.7.1**, so that $\overline{H}_0 \circ \Phi = H_{\mathcal{O}}$. ■

2.9 Reduction by Stages

Theorems on reduction by stages have been given in various special instances by a number of authors, starting with Marsden and Weinstein [1974, p. 127]. This early version, which was a very simple yet basic result stated that for two commuting groups, one could reduce by them in succession and in either order and the result is the same as reducing by the *direct* product group. We state this result in the following theorem.

Theorem 2.9.1 (Commuting Reduction Theorem) *Let P be a symplectic manifold, K be a Lie group acting symplectically on P and having an equivariant momentum map $\mathbf{J}_K : P \to \mathfrak{k}^*$. Assume that $\nu \in \mathfrak{k}^*$ is a regular value of $\mathbf{J}_K$ and that the action of K_ν is free and proper (so that the relevant quotient will be a smooth manifold). Let $P_\nu = \mathbf{J}_K^{-1}(\nu)/K_\nu$ denote the symplectic reduced space. Let G be another group acting on P with an equivariant momentum map $\mathbf{J}_G : P \to \mathfrak{g}^*$. Suppose that the actions of G and K commute. Then*

i $\mathbf{J}_K$ *is invariant under the connected component of the identity of* G *and* $\mathbf{J}_G$ *is invariant under the connected component of the identity of* K*.*

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

iii *The reduced space (assuming it exists) for the action of* G *on* P_ν *at* μ *is symplectically diffeomorphic to the reduction of* P *at* (μ, ν) *by the action of* $G \times K$*.*

For example, in the dynamics of a rigid body with two equal moments of inertia in a gravitational field moving with a fixed point there are two commuting S^1 symmetry groups acting on the space (which are responsible for the complete integrability of the problem). One can reduce these groups either together or one following the other with the same final reduced space.

Proof

i Since the actions commute, we have $[\xi_P, \eta_P] = 0$ for all $\xi \in \mathfrak{g}$ and $\eta \in \mathfrak{k}$, where ξ_P denotes the infinitesimal generator for the action of G on P. Thus, using the definition of the momentum map, we have

$$0 = [\xi_P, \eta_P] = [X_{\langle \mathrm{J}_G, \xi \rangle}, X_{\langle \mathrm{J}_K, \eta \rangle}] = X_{-\{\langle \mathrm{J}_G, \xi \rangle, \langle \mathrm{J}_K, \eta \rangle\}}$$

and hence,

$$0 = \{\langle \mathbf{J}_G, \xi \rangle, \langle \mathbf{J}_K, \eta \rangle\} = \mathbf{d} \langle \mathbf{J}_G, \xi \rangle \cdot X_{\langle \mathrm{J}_K, \eta \rangle} = \mathbf{d} \langle \mathbf{J}_G, \xi \rangle \cdot \eta_P.$$

Thus, $\mathbf{J}_G(\exp_K(t\eta) \cdot p) = \mathbf{J}_G(p)$ for all real t, all $p \in P$, and all $\eta \in \mathfrak{k}$. Since the image of the exponential map generates the connected component of the identity, the result for $\mathbf{J}_G$ follows; a similar argument applies to $\mathbf{J}_K$.

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

Letting $\pi_\nu : \mathbf{J}_K^{-1}(\nu) \to P_\nu$ 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 also from the reduction theorem that $i_\nu^* \Omega = \pi_\nu^* \Omega_\nu$. Therefore,

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

Since π_ν is a surjective submersion, we may conclude that

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

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

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

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

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

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

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.

iii First note that the equivariant momentum map for the action of the product group $G \times K$ is given by $\mathbf{J}_G \times \mathbf{J}_K : P \to \mathfrak{g}^* \times \mathfrak{k}^*$. We begin with the natural inclusion map

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

Composing this map with π_ν gives the map

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

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

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

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

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

Diagram chasing, as above, shows that this map is symplectic.

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

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

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

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

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

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

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

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

Chapter 3

Reduction of Cotangent Bundles

The goal of this chapter is to discuss symplectic reduction for cotangent bundles and to illustrate the theory in numerous examples of mechanical interest. In some cases the reduction gives another cotangent bundle; for example, we will show in the first section below that if G acts on T^*Q by cotangent lift from Q, then the reduced phase space at $\mu = 0$ is

$$(T^*Q)_0 = T^*(Q/G)$$

and $T^*(Q/G)$ carries the canonical symplectic structure. If $\mu \neq 0$, then $(T^*Q)_\mu$ need not equal $T^*(Q/G)$. However, if G is Abelian, then as sets, $(T^*Q)_\mu = T^*(Q/G)$, but even in this case, the symplectic structure need not be the canonical one; rather, it is canonical plus a "magnetic" term. This magnetic term has an interpretation as a magnetic or Coriolis force and can be realized as the curvature of an associated connection often choosen to be a special one we will define called the *mechanical connection*. As we shall also see, for $\mu \neq 0$, $(T^*Q)_\mu$ is a synthesis of the case $\mu = 0$ and Lie-Poisson reduction.

From the point of view of Poisson reduction, the main result on cotangent bundle reduction can be stated this way: the Poisson reduced space $(T^*Q)/G$ is a bundle over the cotangent bundle of shape space $T^*(Q/G)$ with fiber $\mathfrak{g}^*$. That is, we have a bundle:

$$(T^*Q)/G \to T^*(Q/G)$$

where the fibers are copies of $\mathfrak{g}^*$. The description of the Poisson structure on this bundle is quite interesting and as well, the bundle is, in general, not trivial. To enable us to handle this general case, we shall make use of the theory of principal connections; this theory will be summarized in §2.3. The symplectic leaves in this bundle are obtained by restricting to symplectic leaves in the fibers; that is, to coadjoint orbits.

There are special cases of this general theorem that should be kept in mind. The first is the case of Lie-Poisson reduction in which case $Q = G$ and the base is trivial and the bundle is all fiber. Another special case is the case of Abelian groups in which case the bundle "is all base" in the sense that the symplectic leaves are just points in the fiber. Another case in which we get a point in the fiber is when the momentum value at which we are reducing is $\mu = 0$. We shall begin the exposition in this chapter by considering these latter cases.

The cotangent bundle reduction theorem by itself is a powerful tool, but it can also be combined with other results. For example, it may be combined with the semidirect product reduction theory discussed in the last chapter. For example, the semidirect product reduction theory says that the reduction of T^*G by an isotropy subgroup G_a for the action of G on a vector space V is isomorphic to a coadjoint orbit in the dual of the semidirect product $\mathfrak{g} \circledS V$. Suppose, for simplicity, that G_a is abelian. In this case, the cotangent bundle reduction theorem tells us that the reduction of T^*G by G_a is $T^*(G/G_a)$ with a symplectic structure given by the canonical structure plus a magnetic term. Thus, when G_a is abelian, the corresponding coadjoint orbits in the semidirect product are cotangent bundles, possibly with magnetic terms. For example, the generic coadjoint orbits in $SE(3)$ are cotangent bundles of spheres (see IMS, Chapter 14).

3.1 Reduction at Zero

The strategy for understanding the general case is to first deal with the case of reduction at zero and then to treat the general case using a momentum shift. Let $\Phi : G \times Q \to Q$ be a smooth (left) action of G on Q and let $\mathbf{J} : T^*Q \to \mathfrak{g}^*$ be the associated momentum map defined by

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

where $\xi \in \mathfrak{g}$, be the corresponding equivariant momentum map of the (left) cotangent lift of Φ.

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.

Theorem 3.1.1 (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 of G on Q is free, so is the 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.

To show that they are symplectically diffeomorphic, first note that from the definition of $\mathbf{J}$,

$$\mathbf{J}^{-1}(0) = \{\alpha_q \in T^*Q \mid \langle \alpha_q, \xi_Q(q) \rangle = 0 \text{ for all } \xi \in \mathfrak{g}\}. \tag{3.1.1}$$

The equivalence relation on $\mathbf{J}^{-1}(0)$ is by the G-action;

$$\alpha_q \sim T^*\Phi_{g^{-1}} \cdot \alpha_q \tag{3.1.2}$$

and we write $[\alpha_q]$ for the equivalence class of α_q.

To understand how to construct a diffeomorphism between $\mathbf{J}^{-1}(0)/G$ and $T^*(Q/G)$ we first need to understand the tangent spaces to $T(Q/G)$. We claim that one can identify $T_{[q]}(Q/G)$ with the equivalence classes of vectors in TQ at points along the orbit $[q] = G \cdot q$, where the equivalence relation is

$$v_q \sim T\Phi_g \cdot (v_q + \xi_Q(q)) \tag{3.1.3}$$

(without the $\xi_Q(q)$, we would be forming $(TQ)/G$).

One way to see this is to consider a curve in Q/G represented by an equivalence class $[q(t)]$ where $q(t)$ is equivalent to $g(t) \cdot q(t)$. Thus, their tangent vectors, $v_q = \dot{q}(0)$ and $T\Phi_g \cdot (v_q + \xi_Q(q))$ where $g = g(0)$ and $\xi = g^{-1}\dot{g}$ represent the same tangent vector to Q/G.

It is convenient to introduce the following maps. Define the map $\pi_{Q,G} : Q \to Q/G$ taking q to its equivalence class. Its tangent is a map

$$T\pi_{Q,G} : TQ \to T(Q/G).$$

At a point $q \in Q$, the kernel of $T\pi_{Q,G}(q)$ is the set of infinitesimal generators

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

at q. With q fixed, we can identify $[v_q]$ with $T\pi_{Q,G} \cdot v_q$. As indicated above, and again with q fixed, we can identify $T_{[q]}(Q/G)$ with $T_qQ/\mathfrak{g}(q)$ and this identification is consistent with the G-action.

We claim that the following equation determines a diffeomorphism $\varphi_0 : \mathbf{J}^{-1}(0)/G \to T^*(Q/G)$:

$$\varphi_0([\alpha_q]) \cdot [v_q] = \langle \alpha_q, v_q \rangle. \tag{3.1.4}$$

To see that φ_0 is well-defined, let

$$\alpha_{q'} = T^*\Phi_{g^{-1}} \cdot \alpha_q \quad \text{and} \quad v_{q'} = T\Phi_g \cdot (v_q + \xi_Q(q)). \tag{3.1.5}$$

Then,

$$\begin{aligned} \langle \alpha_{q'}, v_{q'} \rangle &= \langle T^*\Phi_{g^{-1}} \cdot \alpha_q, T\Phi_g \cdot (v_q + \xi_Q(q)) \rangle \\ &= \langle \alpha_q, v_q + \xi_Q(q) \rangle = \langle \alpha_q, v_q \rangle \end{aligned}$$

Thus, φ_0 is well-defined. Notice that on the right-hand side of (3.1.5), the same point $q \in Q$ is chosen for the representative of each equivalence class.

In terms of the quotient map $\pi_{Q,G} : Q \to Q/G$ we can write

$$\langle \phi_0([\alpha_q]), (T\pi_{Q,G} \cdot v_q) \rangle = \langle \alpha_q, v_q \rangle\,. \tag{3.1.6}$$

The arbitrariness of v_q in (2.1.5) shows that φ_0 is one-to-one and a dimension count (or in the infinite dimensional case the Hahn-Banach theorem) shows φ_0 is onto; it is essentially the vector space isomorphism $V^0 \cong (E/V)^*$, where V is a vector subspace of a Banach space E and $V^0 = \{\alpha \in E^* \mid \alpha(V) = 0\}$ denotes the annihilator of V. (See Abraham, Marsden and Ratiu [1988, Supplement 2.2].) Thus φ_0 is a diffeomorphism. ■ **This argument is not complete ■.**

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

$$\pi_0^*\varphi_0^*\theta = i_0^*\Theta, \tag{3.1.7}$$

where

- θ is the canonical one form on $T^*(Q/G)$,
- Θ is the canonical one form on T^*Q,
- $i_0 : \mathbf{J}^{-1}(0) \to T^*Q$ is the inclusion map, and
- $\pi_0 : \mathbf{J}^{-1}(0) \to \mathbf{J}^{-1}(0)/G$ is the projection.

Taking the exterior derivative gives

$$\pi_0^* \varphi^* \omega = i_0^* \Omega \tag{3.1.8}$$

where ω is the canonical form on $T^*(Q/G)$ and Ω is that on T^*Q. Equation (3.1.8) implies that

$$\varphi_0^* \omega = \Omega_0 \tag{3.1.9}$$

where Ω_0 is the reduced symplectic form at the level $\mathbf{J} = 0$. This is because of the general unique characterization of the reduced form at level $\mathbf{J} = \mu$ given by

$$\pi_\mu^* \Omega_\mu = i_\mu^* \Omega \tag{3.1.10}$$

that we saw in the last chapter.

It remains to prove (3.1.7). To do this, let $\alpha_q \in \mathbf{J}^{-1}(0)$ and let $v \in T_{\alpha_q}\mathbf{J}^{-1}(0)$. Then

$$(\pi_0^* \varphi_0^* \theta)(\alpha_q) \cdot v = \varphi_0^* \theta([\alpha_q]) \cdot (t\pi_0 \cdot v) = \theta(\varphi_0[\alpha_q]) \cdot (T\varphi_0 \cdot T\pi_0 \cdot v) \tag{3.1.11}$$

Letting $\pi_{Q/G} : T^*(Q/G) \to Q/G$ be the projection, and using the definition of θ, equation (3.1.11) becomes

$$\left\langle \varphi[\alpha_q], T\pi_{Q/G} \cdot T\varphi_0 \cdot T\pi_0 \cdot v \right\rangle = \left\langle \varphi_0[\alpha_q], T(\pi_{Q/G} \circ \varphi_0 \circ \pi_0) \cdot v \right\rangle.$$

However,

$$\pi_{Q/G} \circ \varphi_0 \circ \pi_0 = \pi_{Q,G} \circ \pi_Q \circ i_0 \tag{3.1.12}$$

as both sides map a point $\alpha_q \in \mathbf{J}^{-1}(0)$ to $[q] \in Q/G$.

Using (3.1.12) and (3.1.6), we get

$$\begin{aligned} (\pi_0^* \varphi_0^* \theta)(\alpha_q) \cdot v &= \langle \varphi_0([\alpha_q]), T\pi_{Q,G} \cdot T\pi_Q \cdot Ti_0 \cdot v \rangle \\ &= \langle \alpha_q, T\pi_Q \cdot Ti_0 \cdot v \rangle \\ &= i_0^* \Theta(\alpha_q) \cdot v. \end{aligned}$$

Thus, (3.1.7) holds and so the theorem is proved. ∎

Example. Consider the heavy top with a fixed point. That is, choose $P = T^*\mathrm{SO}(3)$ and let $G = S^1$ act on $Q = \mathrm{SO}(3)$ by left multiplication where we regard S^1 as the subgroup of $\mathrm{SO}(3)$ given by rotations about the z-axis. The reduction of P at $\mu = 0$ is thus given by T^*S^2, since the shape space is given by $\mathrm{SO}(3)/S^1 = S^2$.

Exercises

3.1-1 *Calculate the reduction at zero of $T^*\mathrm{SE}(3)$ by the subgroup of translations.*

3.1-2 *If, in the theorem on reduction at zero, there is another group H acting on Q that commutes with the G action, then the H action on T^*Q induces an action on the reduced space (by the theorem on commuting reduction) that is given by the cotangent lift of the induced action of H on Q/G.*

3.2 Abelian Reduction

A second noteworthy situation is when G is Abelian, or more generally, when $G = G_\mu$. In this case, the reduced space $(T^*Q)_\mu$ will, as a manifold be $T^*(Q/G)$, but the symplectic form will be the canonical one plus a magnetic term. For $\mu = 0$, we do have $G = G_\mu$ and in addition, this magnetic term will be zero, so this case agrees with the result on reduction at zero given in the preceding section.

Theorem 3.2.1 (Abelian or Fully Isotropic Cotangent Bundle Reduction). *Let G act freely and properly on Q. Let $\mu \in \mathfrak{g}^*$ and assume that G is Abelian, or more generally that $G_\mu = G$. Also assume that there is a G-invariant one form α_μ on Q which takes values in $\mathbf{J}^{-1}(\mu)$; that is, we have the identity $\mathbf{J}(\alpha_\mu(q)) = \mu$ for all $q \in Q$. Then there is a unique closed 2-form β_μ on Q/G such that*

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

and a symplectic diffeomorphism φ_μ between the symplectic manifolds $((T^*Q)_\mu, \Omega_\mu)$ and $(T^*(Q/G), \omega - B_\mu)$, where $B_\mu = \pi^*_{Q/G}\beta_\mu$.

Before beginning the proof, we point out that there is a specific way to construct α_μ using connections that is given in the next section. This construction will turn out to be very important for us later on.

Proof. First of all, we verify that β_μ, a two form on Q/G, is well defined. Note that α_μ itself is not assumed to drop to a one form on Q/G and indeed, a study of examples shows that it generally will not.

For $\xi \in \mathfrak{g}$ and $q \in Q$, observe that

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

by definition of $\mathbf{J}$ and our assumption that α_μ takes values in $\mathbf{J}^{-1}(\mu)$. Thus, $\mathbf{i}_{\xi_Q}\alpha_\mu$ is a constant function on Q. Since α_μ is G-invariant, $\pounds_{\xi_Q}\alpha_\mu = 0$ and so

$$\mathbf{i}_{\xi_Q}\mathbf{d}\alpha_\mu = \pounds_{\xi_Q}\alpha_\mu - \mathbf{d}\mathbf{i}_{\xi_Q}\alpha_\mu = 0. \tag{3.2.1}$$

The preceding condition and G-invariance implies that $\mathbf{d}\alpha_\mu$ drops as a two form to Q/G. This process is of course similar to what we used to drop symplectic forms to the quotient.

It remains to construct the symplectic diffeomorphism φ_μ. The key idea is to use the given one form α_μ to shift to zero. Indeed, we define the diffeomorphism $\tau^\mu : T^*Q \to T^*Q$ by

$$\tau^\mu(\alpha_q) = \alpha_q - \alpha_\mu(q) \tag{3.2.2}$$

and note that τ^μ gives a diffeomorphism $\tau_\mu : \mathbf{J}^{-1}(\mu) \to \mathbf{J}^{-1}(0)$ such that $i_0 \circ \tau_\mu = \tau^\mu \circ i_\mu$; this is simply the statement that if α_q has momentum μ, then $\tau^\mu(\alpha_q)$ has momentum value zero. Moreover, by the momentum shifting lemma (see *IMS*, §6.6),

$$(\tau^\mu)^*\Omega = \Omega + \pi^*_Q\mathbf{d}\alpha_\mu,$$

which implies that

$$\tau^*_\mu i^*_0\Omega = i^*_\mu(\Omega + \pi^*_Q\mathbf{d}\alpha_\mu). \tag{3.2.3}$$

By G-invariance, we get an induced diffeomorphism

$$\tau_G^\mu : \mathbf{J}^{-1}(\mu)/G \to \mathbf{J}^{-1}(0)/G.$$

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

$$\varphi_\mu = \varphi_0 \circ \tau_G^\mu : \mathbf{J}^{-1}(\mu)/G \to T^*(Q/G).$$

To show that $\varphi_\mu^*(\omega - B_\mu) = \Omega_\mu$, we shall check that $\varphi_\mu^*(\omega - B_\mu)$ satisfies the property *uniquely* characterizing Ω_μ, namely

$$\pi_\mu^* \varphi_\mu^*(\omega - B_\mu) = i_\mu^* \Omega. \tag{3.2.4}$$

Now we shall use the definition of φ_μ and the identity $\tau_G^\mu \circ \pi_\mu = \pi_0 \circ \tau_\mu$, to rewrite the left hand side of (3.2.4) as follows:

$$\begin{aligned}
(\varphi_\mu \circ \pi_\mu)^*(\omega - B_\mu) &= (\varphi_0 \circ \tau_G^\mu \circ \pi_\mu)^*(\omega - B_\mu) \\
&= (\varphi_0 \circ \pi_0 \circ \tau_\mu)^*(\omega - B_\mu) \\
&= \tau_\mu^* \pi_0^* \varphi_0^*(\omega - B_\mu).
\end{aligned} \tag{3.2.5}$$

However, $B_\mu = \pi_{Q/G}^* \beta_\mu$ and $\varphi_0^* \omega = \omega_0$, so (3.2.5) becomes

$$\tau_\mu^* \pi_0^* \omega_0 - \tau_\mu^* \pi_0^* \varphi_0^* \pi_{Q/G}^* \beta_\mu \tag{3.2.6}$$

But

$$\tau_\mu^* \pi_0^* \omega_0 = \tau_\mu^* i_0^* \Omega = i_\mu^* (\Omega + \pi_Q^* \mathbf{d} \alpha_\mu)$$

by (3.2.3). Also, from our work on reduction at zero, we recall the identity $\pi_{Q/G} \circ \pi_0 = \pi_{Q,G} \circ \pi_Q \circ i_0$ and so the last term in (3.2.6) becomes

$$\tau_\mu^* \pi_0^* \pi_{Q/G}^* \beta_\mu = \tau_\mu^* i_0^* \pi_Q^* \pi_{Q,G}^* \beta_\mu = \tau_\mu^* i_0^* \pi_Q^* \mathbf{d} \alpha_\mu.$$

But, $\pi_Q \circ i_\mu = \pi_Q \circ i_0 \circ \tau_\mu$ so (3.2.6) becomes $i_\mu^* \Omega$ as required. ■

Coordinate Form of the Magnetic Terms We now discuss the coordinate form for the reduced symplectic structure for the case in which G is a connected Abelian group. In particular, G must be the product of a torus with a Euclidean space. We introduce a *local trivialization* of Q by writing a G-invariant open set in Q as the product $Q/G \times G$ in which the action of G is in the second factor only and the action in that factor is the standard one of left multiplication—in this case, translation. The phrase "local in the base" Q/G is used since the open set used corresponds to the product of an open set in the base with the fiber for the bundle $\pi_{Q,G} : Q \to Q/G$. Under the conditions of a free and proper action that we have been making, such a local trivialization always exists.

Such a local trivialization thus has the form $U \times G$ in which $U \subset Q/G$ is an open set. Introduce coordinates $(q^1, q^2, \ldots q^n)$ in which $(q^1, \ldots q^m)$ are coordinates for Q/G and for which the remainder are coordinates for G. We shall write the full set of coordinates on Q as $q^i, i = 1, \ldots, n$ and those on Q/G as $q^\gamma, \gamma = 1, \ldots m$. Coordinates on G will be written as $(\theta^1, \ldots \theta^p)$, or $\theta^a, a = 1, \ldots p$, so that $n = m + p$. In summary, we write

$$q^i = (q^\gamma, \theta^a)$$

for the coordinates on Q, with the first batch being for Q/G and the second being for G.

We use the notation θ^a for the coordinates on G since we are thinking primarily of the case in which G is a torus and so $(\theta^1, \dots \theta^p)$ are literally angle variables.

As we have seen above, after reduction, the phase space is

$$T^*(Q/G), \omega - B_\mu).$$

In the local trivialization coordinates just introduced, we can write the given G-invariant one form α_μ as

$$\alpha_\mu = \alpha_\gamma dq^\gamma + \alpha_a d\theta^a.$$

In this expression, the fact that α_μ is G-invariant means that the coefficients α_γ and α_a are independent of θ^a since the group acts simply by translation in the second factor and not in the first factor. Thus, taking the exterior derivative, we get

$$\mathbf{d}\alpha_\mu = \frac{\partial \alpha_\gamma}{\partial q^\delta} dq^\delta \wedge dq^\gamma + \frac{\partial \alpha_a}{\partial q^\gamma} dq^\gamma \wedge d\theta^a$$

Now we want to use the fact that the form α_μ takes values in $\mathbf{J}^{-1}(\mu)$. Since the action is by translation in the angles, the momentum map is given in the above coordinates by

$$\mathbf{J}(q^\gamma, \theta^a, p_\gamma, p_a) = p_a$$

and so the condition that α_μ take values in $\mathbf{J}^{-1}(\mu)$ is simply that $\alpha_a = \mu_a$. In particular, this means that α_a is a constant and so we get

$$\mathbf{d}\alpha_\mu = \frac{\partial \alpha_\gamma}{\partial q^\delta} dq^\delta \wedge dq^\gamma = \sum_{\gamma<\delta} \left(\frac{\partial \alpha_\gamma}{\partial q^\delta} - \frac{\partial \alpha_\delta}{\partial q^\gamma} \right) dq^\delta \wedge dq^\gamma$$

It is evident from these expressions that the one form α_μ need not drop to a one form on the quotient space (with coordinates q^γ) whereas its exterior derivative does. Thus, the coordinate expression for the reduced symplectic form is

$$\omega - B_\mu = dq^\gamma \wedge dp_\gamma - \sum_{\gamma<\delta} B_{\gamma\delta}\, dq^\gamma \wedge dq^\delta \tag{3.2.7}$$

where

$$B_{\gamma\delta} = \frac{\partial \alpha_\gamma}{\partial q^\delta} - \frac{\partial \alpha_\delta}{\partial q^\gamma}.$$

In these coordinates, the space $\mathbf{J}^{-1}(\mu)$ is described by the conditions $p_a = \mu_a$ and the quotient space $\mathbf{J}^{-1}(\mu)/G$ is coordinatized by simply eliminating the angles as coordinates; the projection from $\mathbf{J}^{-1}(\mu)$ to $\mathbf{J}^{-1}(\mu)/G$ is simply

$$(q^\gamma, \theta^a, p_\gamma, \mu_a) \mapsto (q^\gamma, p_\gamma)$$

and the map φ_μ is given by

$$(q^\gamma, p_\gamma) \mapsto (q^\gamma, p_\gamma - \alpha_\gamma(q)).$$

Hamilton's equations for the symplectic form (3.2.7) are given by

$$\begin{aligned} \dot{q}^\gamma &= \frac{\partial H^\mu}{\partial p_\gamma} \\ \dot{p}_\gamma &= -\frac{\partial H^\mu}{\partial q^\gamma} - B_{\gamma,\delta} \frac{\partial H^\mu}{\partial p_\delta} \end{aligned}$$

where H^μ is the reduced Hamiltonian and the Poisson bracket is given by

$$\{F, K\} = \frac{\partial F}{\partial q^\gamma} \frac{\partial K}{\partial p_\gamma} - \frac{\partial K}{\partial q^\gamma} \frac{\partial F}{\partial p_\gamma} - B_{\gamma\delta} \frac{\partial F}{\partial p_\gamma} \frac{\partial K}{\partial p_\delta}.$$

Exercises

2.2-1 *Verify the preceding assertions about Hamilton's equations and the Poisson bracket.*

2.2-2 *Calculate the reduction at a general* μ *of* $T^*\mathrm{SO}(3)$ *by* S^1*, where* S^1 *is regarded as the subgroup of* $\mathrm{SO}(3)$ *given by rotations about the* z*-axis. In this situation, show explicitly that* α_μ *does not drop to the quotient, but that its exterior derivative does.*

2.2-3 *Show that*

a TQ/G *is a* $\mathfrak{g}$ *bundle over* $T(Q/G)$*.*

b $T(Q/G)$ *is diffeomorphic to* $(TQ)/(TG)$*.*

3.3 Principal Connections

In preparation for the general cotangent bundle reduction theorem we now give a review and summary of facts that we shall need about principle connections. The main things to keep in mind are that the magnetic terms in the cotangent bundle reduction theorem will appear as the curvature of a connection and that the theory of connections gives us a useful formalism for constructing the G-invariant one forms α_μ that were used in the preceding section.

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

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

denote the bundle projection from the full configuration space to shape space. As earlier, the Lie algebra of G is denoted $\mathfrak{g}$. We refer to $\pi_{Q,G} : Q \to Q/G$ as a ***principal bundle***.

Of course, one can use right actions too and indeed in the principal bundle literature, it is more common to use right actions. However, following our tradition, we shall use left actions for the main exposition, with the understanding that the reader shall note the needed changes for the case of right actions.

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 $\pi_{Q,G}$.

Definition 3.3.1 *A **connection**, also called a **principal connection** on the bundle* $\pi_{Q,G} : Q \to Q/G$ *is a Lie algebra valued one form*

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

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 element, and*

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

*for all $v \in T_qQ$, where $\Phi_q : Q \to Q$ denotes the given action for $g \in G$ and where Ad_g denotes the adjoint action of G on $\mathfrak{g}$. This second property is referred to as **equivariance** of the connection.*

A remark is noteworthy at this point. Namely, recall the following general fact for infinitesimal generators:

$$T\Phi_g \cdot \xi_Q(q) = (\mathrm{Ad}_g \xi)_Q(gq).$$

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

Associated One Forms. Since $\mathcal{A}$ is a Lie algebra valued one 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 one form:

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

We assert that this one form satisfies the two crucial properties we needed in the last section.

Proposition 3.3.2 *For any connection $\mathcal{A}$ and $\mu \in \mathfrak{g}^*$, the corresponding one form α_μ defined by the preceding equation 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 first property of a connection,

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

Thus, we conclude that $\mathbf{J}(\alpha_\mu(q)) = \mu$ and so α_μ takes values in $\mathbf{J}^{-1}(\mu)$. To show invariance of the form α_μ we compute in the following way using the equivariance of a connection. Let $v \in T_qQ$ and $g \in G$. Then

$$\begin{aligned} (\Phi_g^* \alpha_\mu)(v) &= \alpha_\mu(gq) \cdot (T\Phi_g \cdot v) \\ &= \left\langle \mathcal{A}_{gq}^*(\mu), T\Phi_g \cdot v \right\rangle \\ &= \langle \mu, \mathcal{A}(T\Phi_g \cdot v) \rangle \\ &= \langle \mu, \mathrm{Ad}_g(\mathcal{A}(v)) \rangle \\ &= \left\langle \mathrm{Ad}_g^* \mu, \mathcal{A}(v) \right\rangle , \end{aligned}$$

so that we get the required equivariance property. ■

Notice in particular, if the group is Abelian or if μ is G-invariant, then α_μ is an invariant one form. Thus, in the Abelian case, or the case in which $G = G_\mu$ we have the hypotheses needed for the one form α_μ in the last section.

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

Definition 3.3.3 *Given the conncetion $\mathcal{A}$, 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

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

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

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

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

Because connections map infinitesimal generators of a Lie algebra elements to that 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

$$T_qQ = H_q \oplus V_q$$

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 complementary to V_q at each point, varying smoothly with q and respecting the group action in the sense that $H_{gq} = T\Phi_g H_q$. Clearly this is equivalent to our definition.

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 associated with the kinetic energy of a given mechanical system. We 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***. In the next section we shall develop an explicit formula for the associated Lie algebra valued one form in terms of an inertia tensor and the momentum map.

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

Definition 3.3.4 *The **curvature** of a connection $\mathcal{A}$ is the Lie algebra valued two form on Q defined by*

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

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

When one replaces vectors in the exterior derivative with their horizontal projections, then the result is called the ***exterior covariant derivative*** and one writes it as

$$\mathcal{B} = D\mathcal{A}.$$

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 is called the ***horizontal lift*** to $q \in Q$. Since $\mathcal{B}$ depends only on the horizontal part of the vectors, it defines a Lie algebra valued two form on the base Q/G. The expression on the right hand side of the definition of the curvature is independent of the point representing $[q]$ because of G-equivariance.

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

Proposition 3.3.5 *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 of Cartan relating the exterior derivative and the Lie bracket of vector fields:

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

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

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.

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

Theorem 3.3.6 (Cartan Structure Equations) *For any vector fields u, v on Q we have*

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

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

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

To prove this theorem we prepare a lemma.

Lemma 3.3.7 *We have the identity* $\mathbf{d}\mathcal{A}(\mathrm{hor}(u), \mathrm{ver}(v)) = 0$ *for any two vector fields u, v on Q .*

Proof. Since this identity depens 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 so the lemma follows. ■

Proof of the Cartan structure equations 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 $\mathrm{hor}(u) = \xi_Q$ and that $\mathrm{hor}(v) = \eta_Q$ for $\xi \in \mathfrak{g}$ and $\eta \in \mathfrak{g}$. Then

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

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

Theorem 3.3.8 (Mauer-Cartan Equations). *Let G be a Lie group and let $\rho : TG \to \mathfrak{g}$ be the map that right translates vectors to the identity:*

$$\rho(v_g) = T_g R_{g^{-1}} \cdot v_g.$$

Then

$$\mathbf{d}\rho - [\rho, \rho] = 0.$$

Proof. Note that ρ 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 $\xi_Q(q) = T_e R_g \cdot \xi$. 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. ■

Another proof of this result is given in *IMS*, §9.1. Of course there is a similar result for the left trivialization λ and we get the identity

$$\mathbf{d}\lambda + [\lambda, \lambda] = 0.$$

Bianchi Identities. In geometry the Bianchi identities are a famous set of identities for the Riemann curvature tensor of a given Riemannian metric. In fact, this set of identities is valid for more general notions of connection as well, such as affine connections. The relation between the Levi-Cevita connection with the present formalism is to use the frame bundle as the bundle Q and think of it as a principal bundle over the underlying manifold M and the group SO(n) as the structure group. Then the curvature as defined here coincides with the Riemann curvature tensor. We will not go into this in detail here as it is not needed for our present purposes, and instead we refer to Spivak [19xx] or Kobayashi and Nomizu [1964] for an exposition of this. It is interesting that in the context of principal connections, the general proof is rather easy.

Theorem 3.3.9 (Bianchi Identities.) *We have the identity $D\mathcal{B} = 0$, that is, for any vector fields u, v, w on Q*

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

Proof. From the structure equations and the fact that $\mathbf{d}^2\mathcal{A} = 0$ we find that $\mathbf{d}\mathcal{B} = \mathbf{d}[\mathcal{A}, \mathcal{A}]$. Using the identity relating the exterior derivative and the Jacobi-Lie bracket of vector fields, we get

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

But all the terms in this expression are zero since $\mathcal{A}$ vanishes on horizontal vectors. ■

Coordinate Formulae. Needs to be filled in.

Exercises

2.3-1 *Recall from Exercise 2.2-3 that TQ/G is a $\mathfrak{g}$ bundle over $T(Q/G)$. Put a natural G-connection on this bundle and compute its curvature.*

2.3-2 *In the rigid body, show that $\mathbf{J}^{-1}(\mu) \cong \mathrm{SO}(3)$ is a circle bundle over S^2. Show that the canonical one form defines a principal S^1 connection on this bundle. Compute its curvature.*

2.3-3 *In the Kaluza Klein approach to the dynamics of a particle in a magnetic field (IMS, §7.6), show that if the magnetic potential is A, then $\omega = d\theta + A$ can be interpreted as a connection and that the magnetic field is its curvature. Show that the Kaluza-Klein metric is determined by requiring that the associated mechanical connection be ω.*

3.4 Cotangent Bundle Reduction—Embedding Version

Assume that $\mu \in \mathfrak{g}^*$ is a regular (or clean) value of $\mathbf{J}$, that $Q_\mu = Q/G_\mu$ is a smooth manifold and that the canonical projection $\pi : Q \to Q_\mu$ is a surjective submersion. Recall that this is assured if we assume that the coadjoint isotropy group G_μ acts freely and properly on Q; then G_μ acts freely and properly on $\mathbf{J}^{-1}(\mu)$ so that the reduced phase space $((T^*Q)_\mu, \Omega_\mu)$ is a symplectic manifold.

For $\mu \in \mathfrak{g}^*$, let $\mu' := \mu|\mathfrak{g}_\mu \in \mathfrak{g}^*$ be the restriction of μ to $\mathfrak{g}_\mu$ and consider the G_μ-action on Q and its lift to T^*Q. The equivariant momentum map of this action is the map $\mathbf{J}^\mu : T^*Q \to \mathfrak{g}_\mu$ obtained by restricting $\mathbf{J}$, *i.e.*, $\mathbf{J}^\mu(\alpha_q) = \mathbf{J}(\alpha_q)|\mathfrak{g}_\mu$. Assume there is a G_μ-invariant one-form α_μ on Q with values in $(\mathbf{J}^\mu)^{-1}(\mu')$. We saw how to construct α_μ in terms of connections in the last section.

For $\xi \in \mathfrak{g}_\mu$ and $q \in Q$,

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

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

$$\mathbf{i}_{\xi_Q}\mathbf{d}\alpha_\mu = \pounds_{\xi_Q}\alpha_\mu - \mathbf{d}\mathbf{i}_{\xi_Q}\alpha_\mu = 0,$$

since the Lie derivative is zero by G_μ-invariance. It follows that there is a unique two form β_μ on Q_μ such that $\pi^*\beta_\mu = \mathbf{d}\alpha_\mu$. We saw this sort of argument when we studied the Abelian cotangent bundle reduction theorem. Since π 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 the cotangent bundle projection.

Theorem 3.4.1 (Cotangent Bundle Reduction Theorem) *Under the above hypotheses, there is a symplectic embedding*

$$\varphi_\mu : ((T^*Q)_\mu, \Omega_\mu) \to (T^*Q_\mu, \omega - B_\mu), \tag{3.4.1}$$

where ω is the canonical symplectic structure on the cotangent bundle T^*Q_μ, onto a vector subbundle over Q_μ. The map φ_μ is onto T^*Q_μ if and only if $\mathfrak{g} = \mathfrak{g}_\mu$.

Remarks.

1. This version of the theorem is due to Satzer [1977] for abelian groups and Marsden (Abraham and Marsden [1978, §4.2]) in the general case. It is an outgrowth of the work of Smale [1970] and the reduction theory of Marsden and Weinstein [1974]. An important interpretation in terms of principal bundles due to Kummer [1981] and its relation to the bundle picture of Montgomery, Marsden and Ratiu [1984] and Montgomery [1986] will be discussed briefly below.

2. Note that the version of the theorem with the hypothesis $G = G_\mu$ we proved earlier in this chapter is a special case of this result.

Proof. To prove the general case, we will reduce the problem to reduction at zero using the momentum shifting lemma. We shall use the fact that there is a natural inclusion

$$\mathbf{J}^{-1}(\mu) \subset (\mathbf{J}^{\mu})^{-1}(\mu') \tag{3.4.2}$$

and μ' is G_μ-invariant. Thus, we get a symplectic inclusion of reduced spaces

$$\mathbf{J}^{-1}(\mu)/G_\mu \subset (\mathbf{J}^{\mu})^{-1}(\mu')/G_\mu. \tag{3.4.3}$$

By G_μ-invariance of both $\mathbf{J}^{-1}(\mu)$ and $(\mathbf{J}^{\mu})^{-1}(\mu')$, this inclusion is an equality precisely when $\mathfrak{g} = \mathfrak{g}_\mu$.

Since we can think of the right hand side of the preceding inclusion as reduction at the value μ' which is G_μ invariant, we are reduced to the proof in the case where μ is G-invariant and *in this case φ_μ will be a symplectic diffeomorphism of* $(\mathbf{J}^{-1}(\mu)/G, \Omega_\mu)$ *with* $(T^*(Q/G), \Omega_0 - B_\mu)$. However, we have already proved this case in §3.2. ■

Example. Consider the reduction of a general T^*Q by $G = \mathrm{SO}(3)$. Here $G_\mu = S^1$ and so the reduced space gets embedded into the cotangent bundle $T^*(Q/S^1)$. A specific example is the case of $Q = \mathrm{SO}(3)$. Then the reduced space is S^2, a coadjoint orbit in $\mathfrak{so}(3)$. In this case, $Q/G_\mu = \mathrm{SO}(3)/S^1 = S^2$ and the embedding of S^2 into T^*S^2 is the embedding into the zero section. The symplectic form in this case is "all magnetic". ♦

Using the results of the preceding section, we can interpret the magnetic term B_μ as the curvature of a connection on a principal bundle. Following this, we give a little more information on the bundle point of view following Montgomery, Marsden and Ratiu [1984]; see also Montgomery [1986] and references therein.

We saw in the preamble to the Cotangent Bundle Reduction Theorem 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 says that $\mathbf{d}\alpha_\mu$ is *horizontal* and thus the covariant exterior derivative of α_μ coincides with $\mathbf{d}\alpha_\mu$. Thus, $\mathbf{d}\alpha_\mu$ is the *μ-component of the curvature two-form*. We summarize:

Proposition 3.4.2 *If the principal bundle $Q \to Q_\mu$ with structure group G_μ has a connection $\mathcal{A}$, then $\alpha_\mu(q)$ can be taken to equal $\mathcal{A}_q^*\mu'$ and B_μ is induced on T^*Q_μ by the μ-component $\mathbf{d}\alpha_\mu$ of the curvature of $\mathcal{A}$.*

The term B_μ on T^*Q is usually called a ***magnetic term***. This terminology 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 $T^*\mathbb{R}$ but with the non-standard symplectic structure $dq^i \wedge dp_i - \frac{e}{c}B$, $i = 1, 2, 3$, where c is the speed of light and B is regarded as a closed two form: $B = B_x dy \wedge dz - B_y dx \wedge dz + B_z dx \wedge dy$ (see *IMS*.

The Cotangent Bundle Reduction Theorem gives a realization of the reduced space $P_\mu = \mathbf{J}^{-1}(\mu)/G$ in case $P = T^*Q$. The left side of the following diagram summarizes the situation:

3.5 Cotangent Bundle Reduction—Bundle Version

The version of the theorem presented above says that P_μ *embeds* as a vector subbundle of $T^*(Q/G_\mu)$ — this is the injection in the above figure. The orbit reduction theorem can be viewed as saying that $P_\mu \cong P_{\mathcal{O}}$ is a *coadjoint bundle* over $T^*(Q/G)$ with fiber the coadjoint

$$\begin{tikzcd}
T^*Q \arrow[ddd] & \supset & \mathbf{J}^{-1}(\mu) \arrow[d, "G_\mu"] & \subset & \mathbf{J}^{-1}(\mathcal{O}) \arrow[d, "G"'] & \subset & T^*Q \arrow[ddd] \\
& & P_\mu \arrow[d, "\text{injection}"'] & \cong & P_{\mathcal{O}} \arrow[d, "\text{surjection}"] & & \\
& & T^*(Q/G_\mu) \arrow[d] & & T^*(Q/G) \arrow[d] & & \\
Q \arrow[rr] & & Q/G_\mu \arrow[rr] & & Q/G & & Q \arrow[ll]
\end{tikzcd}$$

orbit $\mathcal{O}$ through μ. We state this version as follows. As above, we choose a one form α_μ that is induced by a choice of connection $\mathcal{A}$.

Theorem 3.5.1 (Cotangent Bundle Reduction—Bundle Version) *The reduced space P_μ is a bundle over $T^*(Q/G)$ with fiber $\mathcal{O}$.*

Proof.

Step 1 Reduction at zero Recall that we have already shown that reduction at zero is given by

$$(T^*Q)_0 \cong T^*(Q/G). \tag{3.5.1}$$

Here $(T^*Q)_0 = \mathbf{J}^{-1}(0)/G$ and the symplectic form on $T^*(Q/G)$ is, as before, the canonical one.

Step 2 Orbit reduction We have also shown in the preceding chapter 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 $\Sigma : T^*Q \to T^*Q$ defined by $\Sigma(z) = z - \alpha_{\mathbf{J}(z)}(q)$ where $q = \pi_Q(z)$ to map $\mathbf{J}^{-1}(\mathcal{O})$ to $\mathbf{J}^{-1}(0)$.

Notice that the map Σ is nothing more than the horizontal projection map for the connection. Letting $\alpha_z \in T_q^*Q$ be defined by $\alpha_z = \alpha_{\mathbf{J}(z)}(q)$, we can write $\Sigma(z) = z - \alpha_z$. We claim that Σ is equivariant with respect to the G-action. To see this, let $h \in G$ and $v_q \in T_qQ$ and note that

$$\begin{aligned}
\langle T_q^*\Phi_h \cdot \alpha_{h\cdot z}, v_q \rangle &= \langle \alpha_{h\cdot z}, T_q\Phi_h \cdot v_q \rangle = \langle \mathbf{J}(h\cdot z), \mathcal{A}(T_q\Phi_h \cdot v_q)\rangle \\
&= \langle \mathrm{Ad}^*_{h^{-1}}\mathbf{J}(z), \mathrm{Ad}_h \mathcal{A}(v_q)\rangle = \langle \alpha_z, v_q \rangle.
\end{aligned}$$

Thus, $T_q^*\Phi_h \cdot \alpha_{h\cdot z} = \alpha_z$ and so Σ is equivariant and hence drops to the quotient, producing the desired map

$$(T^*Q)_\mu = \mathbf{J}^{-1}(\mathcal{O})/G \to \mathbf{J}^{-1}(0)/G = T^*(Q/G). \tag{3.5.2}$$

This map has fiber $\mathcal{O}$, *i.e.*, $\Sigma \circ \alpha_\mu = 0$ for all $\mu \in \mathcal{O}$, so our assertion is proved. ■

This same type of arguement shows the following

Theorem 3.5.2 *The Poisson reduced space* (T^*Q) *is diffeomorphic to a* $\mathfrak{g}^*$ *bundle over* $T^*(Q/G)$.

Below we will calculate the Poisson structure on $(T^*Q)/G$ following Marsden, Montgomery, and Ratiu [1984] and Montgomery [1986]. (See also Lewis, Marsden, and Ratiu [1987] for an application to the dynamics of systems with free boundaries.)

3.6 The Mechanical Connection Revisited

In this section we amplify in the connection used in the cotangent bundle reduction theorem in the context of a simple mechanical G-system. We assume that G acts freely on Q so we can regard $Q \to Q/G$ as a principal G-bundle.

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

$$\langle \mathbb{I}(q)\eta, \zeta \rangle = \langle\langle \eta_Q(q), \zeta_Q(q) \rangle\rangle \,. \tag{3.6.1}$$

Since the action is free, $\mathbb{I}(q)$ is nondegenerate, so (3.6.1) defines an inner product. The terminology 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,

$$I_{ab} = g_{ij} K^i_a K^j_b. \tag{3.6.2}$$

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

Define the map $\mathcal{A} : T_Q \to \mathfrak{g}$ which assigns to each (q, v) the corresponding *angular velocity of the locked system*:

$$\mathcal{A}(q, v) = \mathbb{I}^{-1}(J(\mathbb{F}L(q, v))), \tag{3.6.3}$$

where L is the kinetic energy Lagrangian. In coordinates,

$$\alpha^a = I^{ab} g_{ij} K^i_b v^j \tag{3.6.4}$$

since $J_a(q, p) = p_i K^i_a(q)$.

Earlier we defined the mechanical connection by declaring its horizontal space to be the metric orthogonal to the vertical space. We claim that the above definition coincides with that.

Proposition 3.6.1 *The map defined by (3.6.3) coincides with the mechanical connection on the principal* G*-bundle* $Q \to Q/G$.

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; *i.e.,* equivariance of $\mathbb{F}L : TQ \to T^*Q$, equivariance of $\mathbf{J} : T^*Q \to \mathfrak{g}^*$, and equivariance of $\mathbb{I}$ in the sense of a map $\mathbb{I} : Q \times \mathfrak{g} \to \mathfrak{g}^*$; $\mathbb{I}(gq) \cdot \mathrm{Ad}_g \xi = \mathrm{Ad}^*_{g^{-1}} \mathbb{I}(q) \cdot \xi$. Thus, $\mathcal{A}$ is a connection.

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

$$\mathrm{hor}_q = \{(q, v) \mid \mathbf{J}(\mathbb{F}L(q, v)) = 0\}; \tag{3.6.5}$$

i.e., the space of states with zero momentum. This space is, from the definition of the momentum map clearly the same as the orthogonal space to the infinitesimal generators; *i.e.,* the vertical space. However, any two connections with the same horizontal spaces are equal. ■

Proposition 3.6.2 *The one form α_μ associated with the mechanical connection is characterized by*

$$K(\alpha_\mu(q)) = \inf\{K(q,\beta) \mid \beta \in \mathbf{J}_q^{-1}(\mu)\} \tag{3.6.6}$$

where $\mathbf{J}_q = \mathbf{J}|T_q^*Q$ *and* $K(q,p) = \frac{1}{2}\|p\|_q^2$ *is the kinetic energy function. See Figure 3.6.1.*

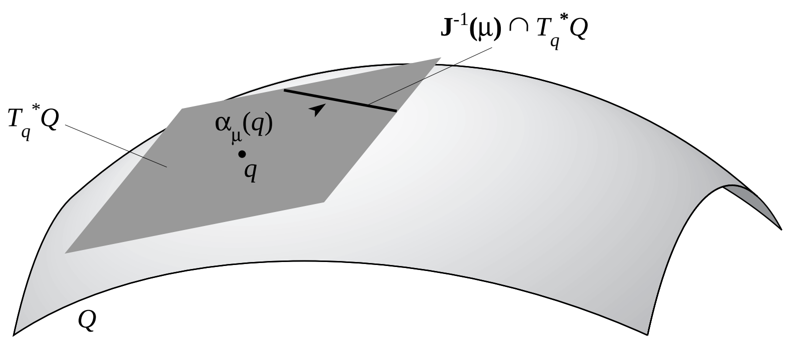

Figure 3.6.1: The extremal characterization of the mechanical connection

The horizontal-vertical decomposition of a vector $(q,v) \in T_qQ$ is given by the general prescription

$$v = \mathrm{hor}_q\, v + \mathrm{ver}_q v \tag{3.6.7}$$

where

$$\mathrm{ver}_q\, v = [\alpha(q,v)]_Q(q) \text{ and } \mathrm{hor}_q\, v = v - \mathrm{ver}_q\, v.$$

In terms of T^*Q rather than TQ, we define a map $\omega : T^*Q \to \mathfrak{g}$ by

$$\omega(q,p) = I(q)^{-1}J(q,p) \tag{3.6.8}$$

i.e.,

$$\omega^a = I^{ab}A_b^i p_i,$$

and, using a slight abuse of notation, a projection $\mathrm{hor} : T^*Q \to \mathbf{J}^{-1}(0)$ by

$$\mathrm{hor}(q,p) = p - \alpha_{\mathrm{J}(q,p)}(q) \tag{3.6.9}$$

i.e.,

$$(\mathrm{hor}(q,p))_i = p_i - g_{ij}A_b^j p_k A_a^k I^{ab}.$$

This map hor plays a fundamental role in what follows. We also refer to hor as the ***shifting map***. This map was used in the proof of the cotangent bundle reduction theorem; it is also an essential ingredient in the description of a particle in a Yang-Mills field via the Kaluza-Klein construction, generalizing the electromagnetic case $p - \frac{e}{c}A$

The ***curvature*** $\mathrm{curv}\alpha$ of the connection α is the covariant exterior derivative of α; also, $\mathrm{curv}\,\alpha$ measures the lack of integrability of the horizontal subbundle. At $q \in Q$,

$$(\mathrm{curv}\,\alpha)(v,w) = \mathbf{d}\alpha(\mathrm{hor}\,v, \mathrm{hor}\,w) = -\alpha([\mathrm{hor}\,v, \mathrm{hor}\,w]), \tag{3.6.10}$$

where the Jacobi-Lie bracket is computed using arbitrary extensions of v, w to vector fields. One can compute curv α by substituting $(2.3.2)'$ and $(2.3.9)'$ into (2.3.10). Taking the μ-component of (2.3.10), we get a 2-form $\langle \mu, \operatorname{curv} \alpha \rangle$ on Q given by

$$\begin{aligned} \langle \mu, \operatorname{curv} \alpha \rangle (v, w) &= \alpha_\mu([\operatorname{hor} v, \operatorname{hor} w]) = -\alpha_\mu([v - \alpha(v)_Q, w - \alpha(w)_Q]) \\ &= -\alpha_\mu([v, \alpha(w)_Q] - \alpha_\mu([w, \alpha(v)_Q] \\ &\quad -\alpha_\mu([v, w]) + \langle \mu, [\alpha(v), \alpha(w)] \rangle \end{aligned} \tag{3.6.11}$$

In (2.3.11) we may choose v and w to be extended by G-invariance. Then $\alpha(w)_Q = \zeta_Q$ for a *fixed* ζ, so $[v, \alpha(w)_Q] = 0$, so we can replace this term by $v[\alpha_\mu(w)], v[\langle \mu, \xi \rangle] = 0$, as $\langle \mu, \xi \rangle$ is constant. Thus, (2.3.11) gives

$$\langle \mu, \operatorname{curv} \alpha \rangle = \mathbf{d}\alpha_\mu + [\alpha, \alpha]_\mu \tag{3.6.12}$$

where $[\alpha, \alpha]_\mu(v, w) = \langle \mu, [\alpha(v), \alpha(w)] \rangle$, the bracket being the Lie algebra bracket. Formula (2.3.12) is, of course, standard for curvatures of principal connections.

3.7 The Poisson Structure on T^*Q/G.

Insert here from Marsden, Montgomery and Ratiu [1984]

3.8 The Amended Potential

The ***ammended potential*** V_μ of Smale [1970] is defined by

$$V_\mu = H \circ \alpha_\mu; \tag{3.8.1}$$

this function also plays a crucual role in what follows. In coordinates,

$$V_\mu(q) = V(q) + \frac{1}{2} I^{ab}(q) \mu_a \mu_b \tag{3.8.2}$$

or, intrinsically

$$V_\mu(q) = V(q) + \frac{1}{2} \left\langle \mu, \mathbb{I}(q)^{-1} \mu \right\rangle. \tag{3.8.3}$$

In the setting of the cotangent bundle reduction theorem, given a Hamiltonian of the form, kinetic plus potential, we get a reduced Hamiltonian system on $P_\mu \cong P_\mathcal{O}$ obtained by restricting H to $H^{-1}(\mu)$ or $\mathbf{J}^{-1}(\mathcal{O})$ and passing to the quotient producing an induced hamiltonian H_μ. Let us compute H_μ in each of the pictures P_μ and $P_\mathcal{O}$. In either case the shift by the map hor is basic, so we first compute the function on $\mathbf{J}^{-1}(0)$ given by

$$H_{\alpha,\mu}(q, p) = H(q, p + \alpha_\mu(q)). \tag{3.8.4}$$

Indeed,

$$\begin{aligned} H_{\alpha,\mu}(q, p) &= \frac{1}{2} \left\langle\left\langle p + \alpha_\mu, p + \alpha_\mu \right\rangle\right\rangle_q + V(q) \\ &= \frac{1}{2} \|p\|_q^2 + \left\langle\left\langle p, \alpha_\mu \right\rangle\right\rangle_q + \frac{1}{2} \|\alpha_\mu\|_q^2 + V(q) \end{aligned} \tag{3.8.5}$$

If $p = \mathbb{F}L \cdot v$, then $\left\langle\left\langle p, \alpha_\mu \right\rangle\right\rangle_q = \langle \alpha_\mu, v \rangle = \langle \mu, \alpha(q, v) \rangle = \langle \mu, I(q)\mathbf{J}(p) \rangle = 0$ since $\mathbf{J}(p) = 0$. Thus, on $J^{-1}(0)$,

$$H_{\alpha,\mu}(q, p) = \frac{1}{2} \|p\|_q^2 + V_\mu(q). \tag{3.8.6}$$

3.9 Examples

Here we consider a few examples of cotangent bundle reduction.

1. The Spherical Pendulum

Here $Q = S^2$, the sphere on which the bob moves, the metric is the standard one, the potential is the gravitational potential and $G = S^1$ acts on S^2 by rotations about the vertical axis. The momentum map is simply the angular momentum. See Figure 2.4.1.

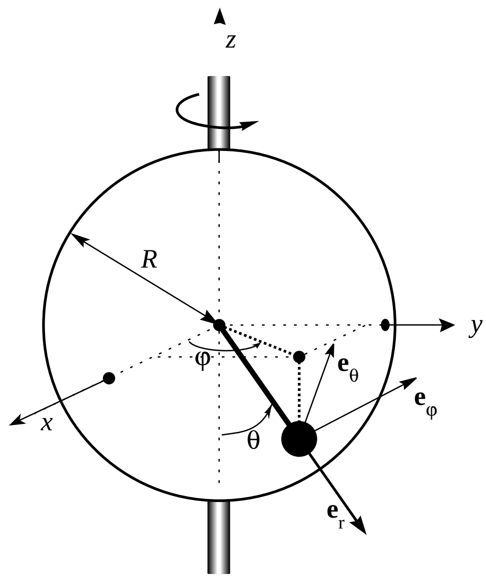

Figure 3.9.1: **The Spherical Pendulum**

Relative to coordinates θ, φ as in Figure 2.4.1, we have $V(\theta, \varphi) = -mgR\cos\theta$. The mechanical connection $\alpha : TQ \to \mathbb{R}$ is given by (2.3.2). We claim that

$$\alpha(\theta, \varphi, \dot{\theta}, \dot{\varphi}) = \dot{\varphi} \tag{3.9.1}$$

To see this, note that

$$\xi_Q(\theta, \varphi) = (\theta, \varphi, 0, \xi) \tag{3.9.2}$$

since $G = \mathbb{S}^1$ acts by rotations about the z-axis: $(\theta, \varphi) \mapsto (\theta, \varphi + \psi)$. The metric is

$$\langle\langle(\theta, \varphi, \dot{\theta}_1, \dot{\varphi}_1), (\theta, \varphi, \dot{\theta}_2, \dot{\varphi}_2)\rangle\rangle = mR^2\dot{\theta}_1\dot{\theta}_2 + mR^2 \sin^2\theta \dot{\varphi}_1\dot{\varphi}_2 \tag{3.9.3}$$

which is m times the standard inner product of the corresponding vectors in $\mathbb{R}^3$.

The momentum map is

$$\mathbf{J} : T^*Q \to \mathbb{R};\ \mathbf{J}(\theta, \varphi, p_\theta, p_\varphi) = p_\varphi \tag{3.9.4}$$

and the Legendre transformation is

$$p_\theta = mR^2\dot{\theta},\ p_\varphi = (mR^2 \sin^2\theta)\dot{\theta}. \tag{3.9.5}$$

$$\langle I(\theta, \varphi)\eta, \zeta\rangle = \langle\langle(\theta, \varphi, 0, \eta)(\theta, \varphi, 0, \zeta)\rangle\rangle = (mP^2 \sin^{-2}\theta)\eta\zeta \tag{3.9.6}$$

Thus, (2.3.2) gives (2.4.1) as claimed.

In this example, we identify Q/S^1 with the interval $[0, \pi]$; *i.e.*, the θ-variable. In fact, it is convenient to regard Q/S^1 as S^1 mod $\mathbb{Z}_2$, where $\mathbb{Z}_2$ acts by reflection $\theta \mapsto -\theta$. This helps to desingularize the quotient space.

For $\mu \in \mathfrak{g}^* \cong \mathbb{R}$, the one form α_μ is given by (2.3.5) as

$$\alpha_\mu(\theta, \varphi) = \mu d\varphi. \tag{3.9.7}$$

From (2.4.4), note that Proposition **2.3.2** is clear. The shifting map is given by

$$\mathrm{hor}(\theta, \varphi, p_\theta, p_\varphi) = (\theta, \varphi, p_\theta, 0). \tag{3.9.8}$$

The curvature of the connection α is zero in this example. The amended potential is

$$V_\mu(\theta) = V(\theta, \varphi) + \frac{1}{2}\langle \mu, I(\theta, \varphi)^{-1}\mu\rangle = -mgR\cos\theta + \frac{1}{2}\frac{\mu^2}{mR^2\sin^2\theta}$$

and so the reduced Hamiltonian on T^*S^1 is

$$H_\mu(\theta, p_\theta) = \frac{1}{2}\frac{p_\theta^2}{mR^2} + V_\mu(\theta). \tag{3.9.9}$$

The reduced Hamiltonian equations are

$$\dot{\theta} = \frac{p_\theta}{mR^2},$$

and

$$\dot{p} = mgR\sin\theta - \frac{\cos\theta}{\sin^2\theta}\frac{\mu^2}{mR^2}.$$

Note that the extra term is singular at $\theta = 0$. ♦

2. Coupled Rigid Bodies

Here we have two rigid bodies in $\mathbb{R}^3$ coupled by a ball in socket joint. We choose $Q = \mathbb{R}^3 \times SO(3) \times SO(3)$ describing the joint position and the attitude of each of the rigid bodies relative to a reference configuration $\mathcal{B}$. The Hamiltonian is the kinetic energy, which defines a metric on Q. Here $G = SE(3)$ which acts on the left in the obvious way by transforming the positions of the particles $x_i = A_iX + w$, $i = 1, 2$ by Euclidean motions. The momentum map is the total linear and angular momentum. If the bodies have additional material symmetry, G is correspondingly enlarged. The cotangent bundle reduction theorem states that $(T^*Q)_\mu$ is a sphere bundle over $T^*(\mathbb{R}^3 \times SO(3))$. We refer to Patrick [1990] for details. ♦

3. Coadjoint Orbits

Let G be a Lie group. The lift to T^*G of the left action of G on itself has the equivariant momentum map $\mathbf{J}_L : T^*G \to \mathfrak{g}^*_+$, $\mathbf{J}(\alpha_g) = T_e^*R_g(\alpha_g)$. Each $\mu \in \mathfrak{g}^*$ is a regular value of $\mathbf{J}_L$ and $\mathbf{J}_L^{-1}(\mu)$ is the graph of the *right* invariant one-form $\tilde{\mu}$ whose value at e is μ, *i.e.*, $\tilde{\mu}(g) = \mu \circ T_gR_L^{-1}$. It is easy to see that $G_\mu = \{g \in G \mid L_g^*\tilde{\mu} = \tilde{\mu}\}$, so G_μ acts on $\mathbf{J}_L^{-1}(\mu)$ by $(g, \tilde{\mu}(h)) \mapsto \tilde{\mu}(gh)$, *i.e.*, by left translation on the base point. Thus, formally $(T^*G)_\mu = \mathbf{J}_L^{-1}(\mu)/G_\mu \cong G/G_\mu \cong \mathcal{O}_\mu \subset \mathfrak{g}^*$. In fact, taking into account that $\mathbf{J}_L^{-1}(\mu)$ is identified with G as the graph of a *right* invariant one-form and the action of G on $\mathbf{J}_L^{-1}(\mu)$ is on the *left*, the prior string of diffeomorphisms can be traced in the following way. The map $\tilde{\mu}(g) \mapsto \mathrm{Ad}_g^*\mu$ is onto $\mathcal{O}_\mu$ and is G_μ-*left* invariant so it factors, inducing the smooth map $\phi_\mu : (T^*G)_\mu \to \mathcal{O}_\mu$, $\phi_\mu[\tilde{\mu}(g)] = \mathrm{Ad}_g^*\mu$, where $[\tilde{\mu}(g)]$ denotes the G_μ-orbit through $\tilde{\mu}(g)$

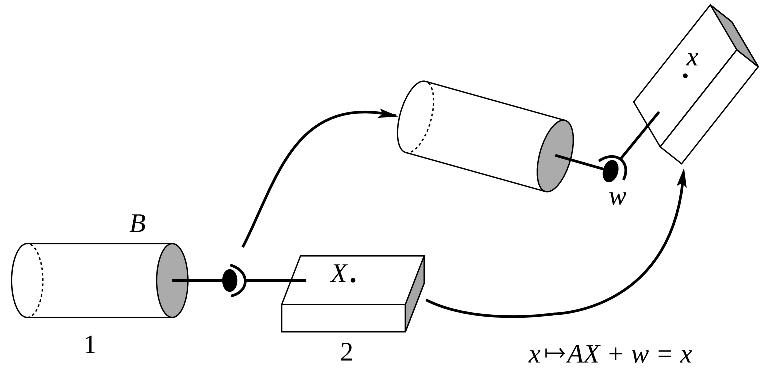

Figure 3.9.2: **Coupled Rigid Bodies**

in T^*G. It is easily checked that $\mathrm{Ad}^*_g\mu \in \mathcal{O}_\mu \mapsto [\tilde{\mu}(g)] \in (T^*G)_\mu$ is the inverse to ϕ_μ and that ϕ_μ is a local diffeomorphism, proving that ϕ_μ is a diffeomorphism of $(T^*G)_\mu$ with $\mathcal{O}_\mu$. A direct calculation shows that the push-forward of the symplectic form Ω_μ of $(T^*G)_\mu$ by ϕ_μ gives the "$-$" coadjoint orbit symplectic form on $\mathcal{O}_\mu$. In the next section, after we have studied Poisson brackets on reduced manifolds, we will give a simple proof by showing that ϕ_μ is the equivariant momentum map of the right G-action on $(T^*G)_\mu$. ♦

4. Reduction to Center of Mass Coordinates

Consider the configuration space for $N+1$ particles in $\mathbb{R}^3$, namely $Q = \mathbb{R}^{3(N+1)}$ regarded as $(N+1)$-tuples $(\mathbf{q}_O, \ldots, \mathbf{q}_N)$ of vectors in $\mathbb{R}^3$. The abelian group $G = \mathbb{R}^3$ acts on Q by $(x, \mathbf{q}_O, \ldots, \mathbf{q}_N) \mapsto (x + \mathbf{q}_O, \ldots, x + \mathbf{q}_N)$. The cotangent lift to $T^*\mathbb{R}^{3(N+1)} = \mathbb{R}^{3(N+1)} \times \mathbb{R}^{3(N+1)} = \{(\mathbf{q}_j, \mathbf{p}^j) \mid (\mathbf{q}_j, \mathbf{p}^j \in \mathbb{R}^3\}$ of this $\mathbb{R}^3$-action has an equivariant momentum map

$$\mathbf{J} : \mathbb{R}^{3(N+1)} \times \mathbb{R}^{3(N+1)} \to \mathbb{R}^3$$

given by the total linear momentum

$$\mathbf{J}(\mathbf{q}_j, \mathbf{p}^j) = \sum_{j=0}^{N} \mathbf{p}^j.$$

This formula shows that any value of $\mathbb{R}^3$ is a regular value of $\mathbf{J}$ and since $\mathbb{R}^3$ acts freely and properly on $\mathbb{R}^{3(N+1)} \times \mathbb{R}^{3(N+1)}$, the reduced phase space $\mathbf{J}^{-1}(\mu)/R^3$ is a symplectic manifold for any $\mu \in \mathbb{R}^3$. This manifold coincides with $T^*(\mathbb{R}^{3(N+1)}/\mathbb{R}^3)$ with values in

$$\mathbf{J}^{-1}(\mu) = \left\{ (\mathbf{q}_j,\ \mathbf{p}^j) \mid \mathbf{q}_j, \mathbf{p}^j \in \mathbb{R}^3,\ \sum_{j=0}^{N} \mathbf{p}^j = \mu \right\}.$$

Clearly $\alpha_\mu(\mathbf{q}_0, \ldots, \mathbf{q}_N) = (\mathbf{q}_0, \ldots, \mathbf{q}_N, \mu/(N+1), \ldots, \mu/(N+1))$ is such a choice, and thus $\mathbf{d}\alpha_\mu = 0$, *i.e.*, $T^*(\mathbb{R}^{3(N+1)}/\mathbb{R}^3)$ is endowed with the canonical symplectic structure.

One convenient way to put coordinates on this reduced manifold is with respect to its center of mass. If $m_0, \ldots, m_N$ are the masses of the $(N+1)$ particles and $m = m_0 + \ldots + m_N$ is the total mass, the vector

$$\mathbf{c} = \frac{1}{m}\left(\sum_{i=0}^{N} m_i \mathbf{q}_i\right) \in \mathbb{R}^3$$

is the center of mass of the system. With respect to $\mathbf{c}$, introduce the new coordinates $\hat{\mathbf{q}}_i = \mathbf{q}_i - \mathbf{c}$ and note that $(\hat{\mathbf{q}}_0, \ldots, \hat{\mathbf{q}}_N)$ lies on the codimension 3 hyperplane

$$M = \left\{ (\mathbf{q}_0, ..., \mathbf{q}_N) \,\middle|\, \sum_{j=0}^{N} m_i q_i = 0 \right\},$$

which can be identified with $\mathbb{R}^{3(N+1)}/\mathbb{R}^3$ since every $\mathbb{R}^3$ orbit intersects M in exactly one point. Consequently, the reduced manifold is $T^*M = \{((\hat{\mathbf{q}}_j, \hat{\mathbf{p}}^j)) \mid \sum_{i=0}^{N} m_i \hat{\mathbf{q}}_i = 0,$ and $\sum_{i=0}^{N} \hat{\mathbf{p}}^i = 0\}$ with its canonical symplectic structure; the relation between $\mathbf{p}^i$ and $\hat{\mathbf{p}}^i$ is given by $\hat{\mathbf{p}^i} = \mathbf{p}^i - m/(N+1)$. Note that the reduced manifold T^*M makes precise the intuitive idea that during the motion it is not the position and the momenta that really count, but only the relative positions and relative momenta. The advantage of coordinatizing everything relative to the center of mass become apparent when the dynamical aspects of reduction are discussed for the $(N+1)$-body problem. It should be noted however that any other coordinate system, e.g. singling out a particular particle such as $\mathbf{q}_0$, $\mathbf{p}^0$, would have worked, but the formulas will then lose their symmetry.

An even better way of coordinatizing $T^*(\mathbb{R}^{3(N+1)}/\mathbb{R}^3)$ is due to Jacobi [1843]; we shall present it following Haretu's [1878] generalization. Let $\mathbf{Q}_i$, $i = 1, \ldots, N$ denote the coordinates of the i-th particle relative to the coordinate system whose axes are parallel to the original one and whose origin is at the center of mass of the first $i-1$ particles. We shall express the $\mathbf{q}_j$, $j = 0, \ldots, N$ in terms of the $\mathbf{Q}_i$, $i = 1, \ldots, N$ and $\mathbf{c}$. For this purpose, let c_i, $i = 0, \ldots, N$ be the center of mass of the first i particles in the original coordinate system and put $\mu_i = m_0 + \ldots + m_i$; note that $\mu_N = m, \mathbf{c}_0 = \mathbf{q}_O$, and $\mathbf{c}_N = \mathbf{c}$. Substituting the values of $\mathbf{c}_i$ given by $\mu_i c_i = m_0\mathbf{q}_0 + \ldots + m_i\mathbf{q}_i$ in $\mathbf{q}_i = \mathbf{c}_{i-1} + \mathbf{Q}_i$, we get

$$\mathbf{q}_i = \mathbf{q}_0 + \frac{m_1}{\mu_1}\mathbf{Q}_1 + \ldots + \frac{m_{i-1}}{\mu_{i-1}}\mathbf{c}_{i-1} + \mathbf{Q}_i,\ i = 1, \ldots, N.$$

Substituting this last relation in $\mu_n \mathbf{c} = m_0\mathbf{q}_0 + \ldots + m_N\mathbf{q}_N$ yields

$$\mathbf{q}_0 = \mathbf{c} - \left(\frac{m_1}{\mu_1}\mathbf{Q}_1 + \ldots + \frac{m_{i-1}}{\mu_{i-1}}\mathbf{Q}_N \right),$$

which together with the foregoing expression for $\mathbf{q}_i$ gives the linear invertible change of coordinates

$$\begin{array}{llllll}
\mathbf{q}_0 = \mathbf{c} - \frac{m_1}{\mu_1}\mathbf{Q}_1 & -\frac{m_2}{\mu_2}\mathbf{Q}_2 & \cdots\cdots & -\frac{m_{N-1}}{\mu_{N-1}}\mathbf{Q}_{N-1} & -\frac{m_N}{\mu_N}\mathbf{Q}_N \\
\mathbf{q}_1 = \mathbf{c} - \frac{m_0}{\mu_1}\mathbf{Q}_1 & -\frac{m_2}{\mu_2}\mathbf{Q}_2 & \cdots\cdots & -\frac{\mu_{N-1}}{\mu_{N-1}}\mathbf{Q}_{N-1} & -\frac{m_N}{\mu_N}\mathbf{Q}_N \\
\mathbf{q}_2 = \mathbf{c} & +\frac{\mu_0}{\mu_1}\mathbf{Q}_2 & \cdots\cdots & -\frac{m_{N-1}}{\mu_{N-1}}\mathbf{Q}_{N-1} & -\frac{m_N}{\mu_N}\mathbf{Q}_N \\
\quad\vdots & & & \quad\vdots & \quad\vdots \\
\mathbf{q}_{N-1} = \mathbf{c} & & & +\frac{\mu_{N-2}}{\mu_{N+1}}\mathbf{Q}_{N-1} & -\frac{m_N}{\mu_N}\mathbf{Q}_N \\
\mathbf{q}_N = \mathbf{c} & & & & +\frac{\mu_{N-1}}{\mu_N}\mathbf{Q}_N
\end{array}$$

It can be easily checked that every $\mathbb{R}^3$-orbit intersects the codimension 3 subspace $\mathbf{c} = 0$ in exactly one point, so that $\mathbb{R}^{3(N+1)}/\mathbb{R}^3$ is identified with $\mathbb{R}^{3N} = \{(\mathbf{Q}_1, \cdots, \mathbf{Q}_N)\}$. Moreover, if $\mathbf{d}, \mathbf{P}^1, \cdots, \mathbf{P}^N$ denote the conjugate variables, representing the momenta of $\mathbf{c}$ and of the last N particles in the corresponding systems, a similar calculation shows that

$$\begin{array}{lllllll}
\mathbf{p}_0 = \frac{m_0}{\mu_N}\mathbf{d} - \frac{m_0}{\mu_1}\mathbf{P}^1 - \frac{m_0}{\mu_2}\mathbf{P}^2 - & \cdots\cdots & -\frac{m_0}{\mu_{N-1}}\mathbf{P}^{N-1} & -\frac{m_0}{\mu_N}\mathbf{P}^N \\
\mathbf{p}^1 = \frac{m_1}{\mu_N}\mathbf{d} + \frac{m_0}{\mu_1}\mathbf{P}^1 - \frac{m_1}{\mu_2}\mathbf{P}^2 - & \cdots\cdots & -\frac{m_1}{\mu_{N-1}}\mathbf{P}^{N-1} & -\frac{m_1}{\mu_N}\mathbf{P}^N \\
\mathbf{p}^2 = \frac{m_2}{\mu_N}\mathbf{d} + \qquad\quad \frac{\mu_1}{\mu_2}\mathbf{P}^2 - & \cdots\cdots & -\frac{m_1}{\mu_{N-1}}\mathbf{P}^{N-1} & -\frac{m_2}{\mu_N}\mathbf{P}^N \\
\quad\vdots & & \quad\vdots & \quad\vdots \\
\mathbf{p}^{N-1} = \frac{m_{N-1}}{\mu_N}\mathbf{d} & & +\frac{\mu_{N-2}}{\mu_{N-1}}\mathbf{P}^{N-1} & -\frac{m_{N-1}}{\mu_N}\mathbf{P}^N \\
\mathbf{p}^N = \frac{m_N}{\mu_N}\mathbf{d} & & & +\frac{\mu_{N-1}}{\mu_N}\mathbf{P}^N
\end{array}$$

Note that $\sum_{i=0}^N \mathbf{p}^i = \mathbf{d} = \mu, s$ so that, as expected, the constraint manifold $\mathbf{J}^{-1}(\mu)$ represents the total momentum of the center of mass. Thus, the Jacobi-Haretu change of variables not only gives the reduced manifold, but also its canonical coordinates. ♦

5. Jacobi's Elimination of the Nodes in the One Body Problem

The lift to $T^*\mathbb{R}^3 = \mathbb{R}^3 \times \mathbb{R}^3$ of the usual $SO(3)$ action on $\mathbb{R}^3$, namely $(A, \mathbf{q}) \mapsto A\mathbf{q}$, has the equivariant momentum map $\mathbf{J} : T^*\mathbb{R}^3 \to \mathfrak{so}(3)^* = \mathbb{R}^3$ given by $\mathbf{J}(\mathbf{q}, \mathbf{p}) = \mathbf{q} \times \mathbf{p}$. The singular set of J is $\sigma(\mathbf{J}) = \{(\mathbf{q}, \mathbf{p}) \in \mathbb{R}^6 \mid \mathbf{q}$ and $\mathbf{p}$ are collinear$\}$ and its projection on the $\mathbf{q}$-space is the entire space; this is in accordance with Remark **C** following the Cotangent Bundle Reduction Theorem since the dimension of the isotropy group $SO(3)_{\mathbf{q}}$ equals one if $\mathbf{q} \neq \mathbf{0}$ and equals three if $\mathbf{q} = \mathbf{0}$. Consequently, any $\mu = 0$ is a regular value of $\mathbf{J}$ and the level set $\mathbf{J}^{-1}(\mu)$ for $\mu \neq 0$ is the codimension three submanifold in $\mathbb{R}^2 \times \mathbb{R}^3$ formed by vectors $\mathbf{q}$ and $\mathbf{p}$ such that $\mathbf{q} \times \mathbf{p} = \mu$. Note that $SO(3)_\mu$ equals the circle group fixing $\mu \in \mathbb{R}^3$ and that the projection of $\mathbf{J}^{-1}(\mu)$ on the $\mathbf{q}$-space equals $\mathbb{R}^3 \backslash \{\mathbf{0}\}$. Since the $SO(3)_\mu$ action on $\mathbb{R}^3$ is not free, the cotangent Bundle Reduction Theorem is not literally applicable. However, Montgomery's generalization does apply. His result shows that the reduced space is the cotangent bundle of a ray. In what follows we shall show this explicitly (we follow an argument of Montgomery); see Exercise **2.3-1** for $\mu = 0$.

Note that if $(\mathbf{q}, \mathbf{p}) \in \mathbf{J}^{-1}(\mu)$ then $\mathbf{q}$ and $\mathbf{p}$ regarded as vectors in $\mathbb{R}^3$ must lie in the plane in $\mathbb{R}^3$ perpendicular to μ. The isotropy group $SO(3)_\mu = S^1$ acts freely and properly on this plane (minus the origin). Moreover, $\mathbf{q} \times \mathbf{p} = \mu$ implies that $\|\mathbf{q}\|\,\|\mathbf{p}\| \sin(\theta(\mathbf{q}, \mathbf{p})) = \|\mu\|$, where $\theta(\mathbf{q}, \mathbf{p})$ denotes the angle between $\mathbf{q}$ and $\mathbf{p}$. Thus $\|\mathbf{q}\|$ and $\mathbf{q} \cdot \mathbf{p}$ determine $(\mathbf{q}, \mathbf{p}) \in \mathbf{J}^{-1}(\mu)$ up to the angle between $\mathbf{q}$ and $\mathbf{p}$. This suggests that the map $\varphi : \mathbf{J}^{-1}(\mu) \to]0, \infty[\times \mathbb{R}, \phi(\mathbf{q}, \mathbf{p}) = (\|\mathbf{q}\|, \mathbf{q} \cdot \mathbf{p})$, which is clearly invariant under the $SO(3)_\mu = S^1$-action, induces a diffeormorphism $\hat{\phi} : (T^*\mathbb{R}^3)_\mu \to]0, \infty[\times \mathbb{R}$. The inverse of $\hat{\phi}$ is easily seen to be given by $\pi_\mu \circ \psi$ where $\psi :]0, \infty[\times \mathbb{R} \to \mathbf{J}^{-1}(\mu)$ is defined by $\psi(r, x) = (r\mathbf{e}_1, (x\mathbf{e}_1 + \|\mu\|\mathbf{e}_2)/r)$, for $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ an orthonormal basis of $\mathbb{R}^3$ such that $\mu = \|\mu\|\mathbf{e}_3$ and $\mathbf{e}_1 \times \mathbf{e}_2 = \mathbf{e}_3$. Denoting as usual the reduced symplectic form by Ω_μ and the canonical one and two-forms on $T^*\mathbb{R}^3$ by $\theta_0 = \sum_{i=1}^3 p_i q^i$, $\Omega_0 = -d\theta_0 = \sum_{i=1}^3 dq^i \wedge dp_i$, the symplectic form on $]0, \infty[\times \mathbb{R}$ is given by $\phi_*\Omega_\mu = \psi^*\pi_\mu^*\Omega_\mu = \psi^* i^*_\mu \Omega_0 = -d\psi^* i_\mu^* \theta_0$. Since $\psi^* q^1 = r$, $\psi^* q^i = 0$, $i = 2, 3$, and

$\psi^* p_1 = x/r$, it follows that $\psi^* i^*{}_\mu \theta_0 = x dr/r = x\, d(\log r)$, so that $\hat{\phi}^* \Omega_\mu = -d(x d(\log r)) = d(\log r) \wedge dx$.

Summarizing, we have shown that the reduced phase space $(T^*\mathbb{R}^3)_\mu$, $\mu \neq 0$ by the lift of the natural $SO(3)$ action on $\mathbb{R}^3$ is $]0, \infty[\times \mathbb{R}$ with the symplectic form $d(\log r) \wedge dx$. ♦

Exercises.

2.4-1 Elimination of the nodes for $\mu = 0$.

3.10 Dynamic Cotangent Bundle Reduction

Finally, the dynamic analog of the cotangent bundle reduction Theorem is obtained by taking a G-invariant Hamiltonian on T^*Q, reducing at $\mu \in \mathfrak{g}^*$ to obtain H_μ on $(T^*Q)_\mu$ and then pushing forwards H_μ, X_{H_μ} and F_t^μ by ϕ_μ onto its image in $(T^*Q_\mu, \omega_0 - \hat{\beta}_\mu)$. All these push-forwards are in the correct relationships, since the embedding ϕ_μ is symplectic.

♡ Needs to be completed with reduced $H = K.E. + V_\mu$ calculation. ♡

3.11 Reconstruction

Here we consider the general reconstruction procedure; *i.e.*, the determination of orbits in the original phase space in terms of those in the reduced space.

Let P be a Poisson manifold on which a Lie group acts in a Hamiltonian manner and which has a momentum map $\mathbf{J} : P \to \mathfrak{g}^*$. For a weakly regular value $\mu \in \mathfrak{g}^*$ of $\mathbf{J}$, assuming that $\mathbf{J}^{-1}(\mu)/G_\mu$ is a smooth manifold with the canonical projection a surjective submersion, $P_\mu := \mathbf{J}^{-1}(\mu)/G_\mu$ is a Poisson manifold. Given $f, h : P_\mu \to \mathbb{R}$, lift them to $\mathbf{J}^{-1}(\mu)$ by π_μ, then extend them to G-invariant functions on $\mathbf{J}^{-1}(\mathcal{O}_\mu)$, where $\mathcal{O}_\mu$ is the coadjoint orbit of μ in $\mathfrak{g}^*$, and then extend these functions arbitrarily to $\overline{f}, \overline{h} : P \to \mathbb{R}$. The Poisson bracket of f and h in the Poisson structure of P_μ is given by $\{f, h\} \circ \pi_\mu = \{\overline{f}|\mathcal{O}_\mu, \overline{h}|\mathcal{O}_\mu\}$. If P is symplectic, then P_μ is also symplectic. If $H : P \to \mathbb{R}$ is a G-invariant Hamiltonian, it induces a Hamiltonian $H_\mu : P_\mu \to \mathbb{R}$ and the flow of the Hamiltonian vector field X_{H_μ} on P_μ is the G_μ-quotient of the flow of X_H on $\mathbf{J}^{-1}(\mu)$, as we have seen.

Assume that the flow F_t^μ of X_{H_μ} on P_μ is known. If $c_\mu(t) = F_t^\mu([p_0])$, for $p_0 \in \mathbf{J}^{-1}(\mu)$, we search for the integral curve $c(t) = F_t(p_0)$ of X_H such that $\pi_\mu(c(t)) = c_\mu(t)$, where $\pi_\mu : \mathbf{J}^{-1}(\mu) \to P_\mu$ is the projection and F_t is the flow of X_H. Pick a smooth curve $d(t)$ in $\mathbf{J}^{-1}(\mu)$ such that $d(0) = p_0$ and $\pi_\mu(d(t)) = c_\mu(t)$. Write $c(t) = \Phi_{g(t)}(d(t))$ for some curve $g(t)$ in G_μ to be determined. We have

$$\begin{aligned} X_H(c(t)) &= c'(t) \\ &= T_{d(t)}\Phi_{g(t)}(d'(t)) \\ &\quad + T_{d(t)}\Phi_{g(t)} \cdot \left(T_{g(t)}L_{g(t)^{-1}}(g'(t))\right)_P (d(t)). \end{aligned} \tag{3.11.1}$$

Since $\Phi_g^* X_H = X_{\Phi_g^* H} = X_H$, (2.6.1) gives

$$\begin{aligned} d'(t) + (T_{g(t)}L_{g(t)^{-1}}(g')(t))_P(d(t)) &= T\Phi_{g(t)^{-1}} X_H \left(\Phi_{g(t)}(d(t))\right) \\ &= (\Phi_{g(t)}^* X_H)(d(t)) = X_H(d(t)). \end{aligned} \tag{3.11.2}$$

This is an equation for $g(t)$ written in terms of $d(t)$ only. We solve it in two steps:

Step 1 *Find $\xi(t) \in \mathfrak{g}_\mu$ such that*

$$\xi(t)_P(d(t)) = X_H(d(t)) - d'(t); \tag{3.11.3}$$

Step 2 *With $\xi(t)$ determined, solve the following non-autonomous ordinary differential equation on G_μ:*

$$g'(t) = T_e L_{g(t)}(\xi(t)), \text{ with } g(0) = e. \tag{3.11.4}$$

We call (2.6.4) *the* ***reconstruction equation***. Step **1** is typically of an algebraic nature. In fact, as in Marsden, Montgomery and Ratiu [1990], $\xi(t)$ can be computed if a connection is given on $\mathbf{J}^{-1}(\mu) \to P_\mu$. Step **2** gives an answer "in quadratures" and represents the main technical difficulty in the reconstruction method.

Proposition 3.11.1 *With $g(t)$ determined by the reconstruction equation, the desired integral curve $c(t)$ is given by $c(t) = \Phi_{g(t)}(d(t))$.*

The same construction works on P/G, even if the G-action does not admit a momentum map.

A particular case in which Step **2** can be carried out explicitly is when G is abelian. Here the connected component of the identity of G is a cylinder $\mathbb{R}^p \times \mathbb{T}^{k-p}$ and the exponential map

$$\exp(\xi_1, \ldots, \xi_k) = (\xi_1, \ldots, \xi_p, \, \xi_{p+1}(\text{mod } 2\pi), \ldots, \xi_k(\text{mod } 2\pi)$$

is onto, so we can write $g(t) = \exp \eta(t)$, $\eta(0) = 0$. Therefore $\xi(t) = T_{g(t)} L_{g(t)^{-1}}(g'(t)) = \eta'(t)$ since η' and η commute, *i.e.*, $\eta(t) = \int_0^t ds$. Thus the solution of (2.5.4) in Step **2** when G is abelian is

$$g(t) = \exp \int_0^t \xi(s) \, ds. \tag{3.11.5}$$

Often, $\xi(t)$ has the interpretation as the angular velocity, so $g(t)$ is the expected accumulated phase due to it. One calls it the ***dynamic phase***. However, as we shall see, an additional contribution can arise from d itself and this will be of geometric origin.

3.12 Additional Examples

To illustrate the methods of reduction and reconstruction, we consider some classical examples. In some cases there is no particular economy in using reduction over the classical "bare hands" approach. However, there is an advantage when we come to more sophisticated cases. Here we do the reconstruction of dynamics directly.

1. Energy Surface

Let $H^{-1}(e)/\mathbb{R}$, if a manifold, be the symplectic manifold obtained by collapsing every orbit of X_H on P to a point. The induced Hamiltonian H_e is the constant function equal everywhere to e. This is consistent with the fact that the quotient operation has removed the dynamics of H. Notice that at saddle points or other critical points of H, the set $H^{-1}(e)$ need not be a manifold and taking the quotient and $H^{-1}(e)/\mathbb{R}$ may also have singularities. ♦

2. Jacobi-Liouville Theorem

Let $H_1, \ldots, H_k$ be smooth functions defined on a symplectic manifold P that are in involution, *i.e.*, $\{H_i, H_j\} = 0$, for all $i, j = 1, \ldots, k$. If all the Hamiltonian vector fields X_{H_i} are complete, then since the flows F_s^i of X_{H_i} all commute,

$$(t, \ldots, t_k) \cdot p = (F_{t_1}^1 \circ \cdots \circ F_{t_k}^k)(p), \text{ where } p \in P,$$

defines a symplectic action of $\mathbb{R}^k$ on P. This action has a momentum map given by $\mathbf{J}(p) = (H_1(p), \ldots, H_k(p))$ where $\mathbb{R}^{k*}$ is identified with $\mathbb{R}^k$. Let $\mu \in \mathbb{R}^k$ be the range of $\mathbf{J} : P \to \mathbb{R}^k$ and assume that the differentials $\mathbf{d}H_1, \ldots, \mathbf{d}H_k$ are linearly independent at every point of $\mathbf{J}^{-1}(\mu)$ *i.e.*, assume that μ is a regular value of $\mathbf{J}$, so that, $\mathbf{J}^{-1}(\mu)$ is a submanifold of P. If the quotient $P_\mu = \mathbf{J}^{-1}(\mu)/\mathbb{R}^k$ is a manifold, it is symplectic (by the symplectic reduction theorem) and has dimension equal to $\dim P - 2k$. Obviously the induced functions $(H_i)_\mu$ are all constant on P_μ. Moreover, if H is another Hamiltonian on P such that $\{H, H_i\} = 0$ for all $i = 1, \ldots, k$, it induces a Hamiltonian H_μ on P_μ. Therefore, *k independent integrals involution for a Hamiltonian system determine a new Hamiltonian system in which $2k$ variables have been eliminated*. This is the statement of the classical ***Jacobi-Liouville Theorem***. Of course, the full *global* assertion assumes not only that μ is a regular value, but that the quotient $\mathbf{J}^{-1}(\mu)/\mathbb{R}^k$ is a manifold if, e.g., the action is free and proper.

If $2k = \dim P$, then P_μ is zero dimensional and the system is called ***completely integrable***. In general,if G is a Lie group acting symplectically on a manifold P, the action is called *completely integrable* if the reduced manifolds P_μ are zero dimensional for almost all $\mu \in \mathfrak{g}^*$. We shall not study completely integrable systems or associated questions involving action-angle variables to any great extent in this book. For this, we direct the reader to Abraham and Marsden [1978], Chapter 5, Arnold [1978], Mischenko and Fomenko [1978], Duistermaat[1979], Cushman and Knörer [1983], Guillemin and Sternberg [1984], Knörer [1985], Fomenko and Trofimov [1988], Bobenko, Reyman and Semenov-Tian- Shansky [1989] and references therein. ♦

3. The Kepler-Newton Central Force Problem in the Plane

The problem of the motion of two bodies in space, under the influence of mutual gravitational attraction is reduced to the problem of one body moving in the plane in a central force potential by utilizing conservation of angular momentum and the reduction to the center of mass (see Example **4** below). The equivalent problem is that of a body of mass in moving the plane in the field of a mass M fixed at the origin with the force field

$$F(\mathbf{q}) = -GMm\frac{\mathbf{q}}{\|\mathbf{q}\|^3} = -Cm\frac{\mathbf{q}}{\|\mathbf{q}\|^3} = \nabla\left(\frac{Cm}{\|\mathbf{q}\|}\right),$$

where $C = GM$ and G is Newton's universal gravitational constant. By Newton's Second law, the equations of motion are

$$m\ddot{\mathbf{q}} = \nabla\left(\frac{Cm}{\|\mathbf{q}\|}\right),$$

which are the Euler-Lagrange equations, for the Lagrangian

$$L(\mathbf{q}, \dot{\mathbf{q}}) = \frac{1}{2}m\|\dot{\mathbf{q}}\|^2 + \frac{CM}{\|\mathbf{q}\|}.$$

The Legendre transformation $(q^i, \dot{q}^i) \mapsto (q^i, \partial L/\partial \dot{q}^i) = (q^i, p_i = m\dot{q}^i)$ transforms these equations to Hamilton's equations on $T^*(\mathbb{R}^2 \setminus \{0\})$;

$$\begin{cases} \dot{q}^i = \dfrac{p_i}{m} \\ \dot{p}_i = Cm\dfrac{\partial}{\partial q^i}\left(\dfrac{1}{\|\mathbf{q}\|}\right) \end{cases}$$

with Hamiltonian

$$H(\mathbf{q},\mathbf{p}) = \dot{\mathbf{q}}\cdot\mathbf{p} - L(\mathbf{q},\dot{\mathbf{q}}) = \frac{\|\mathbf{q}\|^2}{2m} - \frac{Cm}{\|\mathbf{q}\|}.$$

Let $S^1 = SO(2)$ act on $\mathbb{R}^2 \setminus \{0\}$ by counterclockwise rotation:

$$\theta \cdot \mathbf{q} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \mathbf{q}.$$

The lift of this action to $T^*(\mathbb{R}^2 \setminus \{0\})$ is given by

$$\theta \cdot (\mathbf{q},\mathbf{p}) = (\theta\cdot\mathbf{q},\, \theta\cdot\mathbf{p}).$$

Note that H is invariant under this S^1-action. Since our final goal is to determine the motion of the body of mass m by utilizing reduction, taking quotients by S^1 is greatly simplified by the passage to polar coordinates.

The diffeomorphism $\rho :]0,\infty[\, \times S^1 \to \mathbb{R}^2 \setminus \{0\}$ given by $\rho = \rho(r,\theta) = (r\cos\theta, r\sin\theta)$ induces the symplectic diffeomorphism of $T(\,]0,\infty[\,\times S^1) = P \to T^*(\mathbb{R}^2 \setminus \{0\})$ given by $(r,\theta,p_r,p_\theta) \mapsto (\mathbf{q},\mathbf{p})$, where

$$\begin{aligned} p_r &= p_1\cos\theta + p_2\sin\theta;\ p_1 = p_r\cos\theta - \frac{1}{r}p_\theta\sin\theta \\ p_\theta &= -p_1 r\sin\theta + p_2 r\cos\theta;\ p_2 = p_r\sin\theta + \frac{1}{r}p_\theta\cos\theta. \end{aligned}$$

The symplectic form on P is given by $dr\wedge dp_r + d\theta\wedge dp_\theta$, the S^1 action on $]0,\infty[\,\times S^1$ is given by

$$\theta\cdot(r,\phi) = (r,\theta+\phi)$$

and its cotangent lift to $P = T^*(\,]0,\infty[\,\times S^1)$ is

$$\theta\cdot(r,\phi,p_r,p_\theta) = (r,\theta+\phi,p_r,p_\theta).$$

If $\xi\in\mathbb{R}$, the infinitesimal generator of the S^1-action on $]\,0,\infty[\,\times S^1$ is given by

$$(r,\theta)\mapsto(r,\theta,0,\xi),$$

so that the momentum map $\mathbf{J} : P \to \mathbb{R}$ has the expression

$$\mathbf{J}(r,\theta,p_r,p_\theta) = p_\theta.$$

Clearly $\mathbf{J}$ is a submersion and for $\mu\in\mathbb{R}$, $\mathbf{J}^{-1}(\mu) = \{(r,\theta,p_r,p_\theta)\in P \mid p_\theta = \mu\}$. Moreover, the reduced phase space $P_\mu = \mathbf{J}^{-1}(\mu)/S^1$ equals $T^*(]\,0,\infty[\,) = \{(r,p_r) \mid r\in]\,0,\infty[\,,\ p_r\in\mathbb{R}\}$, a fact to be expected by the Cotangent Bundle Reduction Theorem. The symplectic form is easily seen to be the standard one, $dr\wedge dp_r$, either by using the definition or by noting that $]0,\infty[$ is one-dimensional, so β_μ must be zero. We can compute $\alpha_\mu(r,\theta)$ as the minimum of

$$\frac{1}{2m}\left(p_r^2 + \frac{p^2}{r^2}\right)$$

subject to the constraint $p_\theta = \mu$. This minimum is achieved for $p_r = 0$, so that $\alpha_\mu(r,\theta) = \mu d\theta$ which is closed, so once again, $\beta_\mu = 0$.

In polar coordinates, the Hamiltonian has the expression

$$H(r,\theta,p_r,p_\theta) = \frac{1}{2m}\left(p_r^2 + \frac{p_\theta^2}{r^2}\right) - \frac{Cm}{r},$$

so that the reduced Hamiltonian equals

$$H_\mu(r, p_r) = \frac{1}{2m}\left(p_r^2 + \frac{\mu^2}{r^2}\right) - \frac{Cm}{r}.$$

Hamilton's equations on $P_\mu = T^*\,]0, \infty[$ are

$$\begin{aligned}\dot{r} &= \frac{\partial H_\mu}{\partial p_r} = \frac{p_r}{m}\\ \dot{p}_r &= \frac{-\partial H_\mu}{\partial r} = \frac{\mu^2}{mr^3} - \frac{Cm}{r^2}.\end{aligned}$$

The first of these equations combined with conservation of energy $H = E$ yields

$$\dot{r}^2 = (2mEr^2 + 2Cmr - \mu^2)m^2r^2,$$

i.e.,

$$t - t_0 = m\int_{r_0}^{r(t)} \frac{s\,ds}{\sqrt{2mEs^2 + 2Cms - \mu^2}} \tag{3.12.1}$$

which is an elliptic integral. Inverting it gives r as a function of t. Therefore, $p_r(t) = m\dot{r}(t)$, with $r(t)$ determined implicitly by (2.7.1).

To solve the original differential equations in $(r, \theta, p_r, p_\theta)$, consider initial conditions $(r_0, \theta_0, p_r^0, p_\theta^0) \in J^{-1}(\mu)$, *i.e.*, $p_\theta^0 = \mu$ and let the corresponding integral curve of the original Hamiltonian system be denoted by $(r(t), \theta(t), p_r(t), p_\theta(t))$. Choose a curve $d(t) = (\overline{r}(t), \overline{\theta}(t), \overline{p}_r(t), \overline{p}_\theta(t)) \in J^{-1}(\mu)$ which projects to the integral curve $(r(t), p_r(t))$ of H_μ and which has initial conditions $(r_0, \theta_0, p_r^0, p_\theta^0)$. This implies that $\overline{r}(t) = r(t)$, $\overline{p}_r(t) = p_r(t)$, and $\overline{p}_\theta(t) = \mu$. Since no conditions are imposed on $\overline{\theta}(t)$, we may choose it to equal θ_0 for convenience, so there is an angle $\phi(t)$ such that

$$\begin{aligned}(r(t), \theta(t), p_r(t), p_\theta(t)) &= \phi(t) \cdot \left(\overline{r}(t), \overline{\theta}(t), \overline{p}_r(t), \overline{p}_\theta(t)\right)\\ &= (r(t), \theta_0 + \phi(t), p_r(t), \mu)\,,\end{aligned}$$

i.e.,

$$\theta(t) = \theta_0 + \phi(t),$$

with $\phi(0) = 0$. Next, we solve the differential equation on S^1 given by conservation of angular momentum:

$$\phi'(t) = \frac{\mu}{mr(t)^2}, \text{ where } \phi(0) = 0.$$

This has the solution

$$\phi(t) = \frac{\mu}{m}\int_0^t r(s)^{-2}\,ds. \tag{3.12.2}$$

The solution of the Kepler-Newton planar central force problem with initial conditions $(r_0, \theta_0, p_r^0, p_\theta^0 = \mu)$ is given by inverting (2.7.1) to find $r(t)$ and letting $p_r(t) = m\dot{r}(t)$, $\theta(t) = \theta_0 + \phi(t)$ with $\phi(t)$ given by (2.7.2), and $p_\theta(t) = p_\theta^0 = \mu$.

To describe the trajectory of the body of mass m geometrically, we need the relation between $r(t)$ and $\theta(t)$. To obtain it, we proceed as in the classical treatments: intersect the energy surface $H(r, \theta, p_r, p_\theta) = E$ with the level set of the momentum map $p_\theta = \mu$. This requires us to eliminate t and p_r in $H(r, \theta, p_r, \mu) = E$. From $\dot{\theta} = \mu/mr^2$ and $\dot{r} = p_r/m$, we get

$$p_r = m\dot{r} = m\frac{dr}{r^2}\dot{\theta} = \frac{\mu}{r^2}\frac{dr}{d\theta},$$

so that $H(r, \theta, p_r, \mu) = E$ becomes

$$\frac{1}{2m}\left[\frac{\mu^2}{r^4}\left(\frac{dr}{d\theta}\right)^2 + \frac{\mu^2}{r^2}\right] - \frac{Cm}{r} = E,$$

or changing the variable r to $\rho = 1/r$, we get

$$\left(\frac{d\rho}{d\theta}\right)^2 + \rho^2 = \frac{2m}{\mu^2}(Cm\rho + E). \tag{3.12.3}$$

Taking the derivative of this relation with respect to θ and canceling $d\rho/d\theta$ yields the second order constant coefficient equation

$$\frac{d^2\rho}{d\theta^2} + \rho = \frac{Cm^2}{\mu^2}$$

whose solution is given by

$$\rho(\theta) = \frac{Cm^2}{\mu^2} + A\cos(\theta + \theta_0) \tag{3.12.4}$$

with A an arbitrary constant. Substitution of (2.7.4) into (2.7.3) determines the constant A

$$A = \pm\frac{1}{\mu^2}(2m\mu^2 E + C^2 m^4)^{\frac{1}{2}},$$

so that again by (2.7.4) we have

$$\rho(\theta) = \frac{Cm^2}{\mu^2}\left[1 \pm \left(1 + \frac{2E\mu^2}{C^2m^3}\right)^{\frac{1}{2}}\cos(\theta + \theta_0)\right]. \tag{3.12.5}$$

We need not consider signs in (2.7.5) since this is accomplished by the phase shift $\theta \mapsto \theta + \pi$. Replacing θ by $\theta - \theta_0$, we can put (2.7.5) into the form

$$\rho(\theta) = \frac{Cm^2}{\mu^2}\left[1 + \left(1 + \frac{2E\mu^2}{C^2m^3}\right)^{\frac{1}{2}}\cos\theta\right],$$

i.e.,

$$r(\theta) = \frac{\mu^2}{Cm^2}\frac{1}{1 + \left(1 + \frac{2E\mu^2}{C^2m^3}\right)^{\frac{1}{2}}\cos\theta}. \tag{3.12.6}$$

The equation of a conic in polar coordinates with one of the foci at the origin is

$$r(\theta) = \frac{\ell}{1 + e\cos\theta},$$

where e is the eccentricity and ℓ is the distance from a focus to the conic in the direction transverse to the line joining the foci. The cases $e > 1$, $e = 1$, and $0 \le e < 1$ correspond to a hyperbola, parabola, and ellipse, respectively. Thus, we get the classical result that *the orbit of the body of mass M around the body of mass M moving under the influence of the Newtonian gravitational potential is a conic of eccentricity*

$$e = \left(1 + \frac{2E\mu^2}{C^2 m^3}\right)^{\frac{1}{2}}.$$

Note that the sign of $e - 1$ is determined by the sign of E. For example, the conic is an ellipse if and only if $E < 0$, in which case $-e^2 m^3/mu^2 \le E < 0$. ♦

4. Reduction to Center of Mass Coordinates

Let $\mathbb{R}^3$ act on $\mathbb{R}^{3(N+1)}$ by translation on every factor. The lift of this action to $T^*\mathbb{R}^{3(N+1)}$ is given by

$$\mathbf{x} \cdot (\mathbf{q}_0, \ldots, \mathbf{q}_N, \mathbf{p}^0, \ldots, \mathbf{p}^N) = (\mathbf{q}_0 + \mathbf{x}, \ldots, \mathbf{q}_N + \mathbf{x}, \mathbf{p}^0, \ldots, \mathbf{p}^N)$$

so that the infinitesimal generator determined by $\xi \in \mathbb{R}^3$ has the expression

$$\xi_{T^*\mathbb{R}^{3(N+1)}}(\mathbf{q}_0, \ldots, \mathbf{q}_N, \mathbf{p}^0, \ldots, \mathbf{p}^N) = \xi \cdot \left(\frac{\partial}{\partial \mathbf{q}_o} + \cdots + \frac{\partial}{\partial \mathbf{q}_N}\right).$$

The momentum map of this lifted action is thus given by

$$\mathbf{J}(\mathbf{q}_0, \ldots, \mathbf{q}_N, \mathbf{p}^0, \ldots, \mathbf{p}^N) = \mathbf{p}_0 + \cdots + \mathbf{p}_N.$$

The reduced manifold $(T^*\mathbb{R}^{3(N+1)})_\mu$ is symplectically diffeomorphic to $T^*\mathbb{R}^{3N}$ endowed with the canonical symplectic structure. We take

$$(\mathbf{Q}_1, \ldots, \mathbf{Q}_N, \mathbf{P}^1, \ldots, \mathbf{P}^N)$$

to be canonically conjugate coordinates on $T^*\mathbb{R}^{3N}$ given by the Jacobi-Haretu change of variables. Therefore, the projection $\mathbf{J}^{-1}(\mu) \to T^*\mathbb{R}^{3N}$ is given by

$$\begin{aligned}
\mathbf{Q}_i &= \frac{m_0}{\mu_{i-1}}\mathbf{q}_0 - \cdots - \frac{m_{i-1}}{\mu_{i-1}}\mathbf{q}_{i-1} + \mathbf{q}_i, \\
\mathbf{P}^i &= -\frac{m_i}{\mu_{i-1}}(\mathbf{p}^0 + \cdots + \mathbf{p}^{i-1}) + \mathbf{p}^i, \\
& \mathbf{p}^0 + \cdots + \mathbf{p}^N = \mu.
\end{aligned}$$

Consider on $T^*\mathbb{R}^{3(N+1)}$ a Hamiltonian H invariant under the $\mathbb{R}^3 -$ *action*, *i.e.*,

$$H(\mathbf{q}_0 + \mathbf{x}, \cdots, \mathbf{q}_N + \mathbf{x}, \mathbf{p}^0, \cdots, \mathbf{p}^N) = H(\mathbf{q}_0, \cdots, \mathbf{q}_N, \mathbf{p}^0, \cdots, \mathbf{p}^N)$$

for any $\mathbf{x} \in \mathbb{R}^3$. Denote by $(\mathbf{q}_0(t), \ldots, \mathbf{q}_N(t), p^0(t), \ldots, \mathbf{p}^N(t))$ a solution of the Hamiltonian system with Hamiltonian H and with initial condition $\mathbf{q}_i(0) = \overline{\mathbf{q}}_i$, $\mathbf{p}^i(0) = \overline{\mathbf{p}}^i$, $i = 0, 1, \ldots, N$. By conservation of linear momentum,

$$\mathbf{p}^0(t) + \cdots + \mathbf{p}^N(t) = \mu,$$

where $\mu \in \mathbb{R}^3$ is a constant vector. Since $\mathbf{p}^i = m_i \dot{\mathbf{q}}_i$, it follows that

$$m_0 \mathbf{q}_0(t) + \cdots + m_N \mathbf{q}_N(t) = \mu t + (m_0 \overline{\mathbf{q}}_0 + \cdots + m_N \overline{\mathbf{q}}_N),$$

so *the center of mass moves in a straight line during motion.* On the reduced manifold, the center of mass is fixed at the origin.

We will determine the solution $(\mathbf{q}_O(t), \ldots, \mathbf{q}_N(t),\, \mathbf{p}^0(t), \ldots, \mathbf{p}^N(t))$ in terms of a known solution $(\mathbf{Q}_1(t), \ldots, \mathbf{Q}_N(t),\, \mathbf{P}^1(t), \ldots, \mathbf{P}^N(t))$ of the reduced Hamiltonian system given by H_μ on $T^*\mathbb{R}^{3N}$. To do so, let

$$(\mathbf{u}_0(t), \ldots, \mathbf{u}_N(t),\, \mathbf{v}^0(t), \ldots, \mathbf{v}^N(t))$$

be a curve in $\mathbf{J}^{-1}(\mu)$ projecting onto

$$(\mathbf{Q}_1(t), \ldots, \mathbf{Q}_N(t),\, \mathbf{P}^1(t), \ldots, \mathbf{P}^N(t)),$$

i.e.,

$$\begin{aligned}
\mathbf{Q}_i(t) &= -\frac{m_0}{\mu_{i-1}}\mathbf{u}_0(t) - \cdots - \frac{m_{i-1}}{\mu_{i-1}}\mathbf{u}_{i-1}(t) + \mathbf{u}_i(t),\\
\mathbf{P}^i(t) &= -\frac{m_i}{\mu_{i-1}}(\mathbf{v}^0(t) + \cdots + \mathbf{v}^{i-1}(t)) + \mathbf{v}^i(t),
\end{aligned}$$

and

$$\mathbf{v}^0(t) + \cdots + \mathbf{v}^N(t) = \mu,$$

for $i = 1, \ldots, N$. The last $N+1$ equations can be solved for $\mathbf{v}^j(t),\ j = 0, 1, \ldots, N$ to yield

$$\mathbf{v}^j(t) = \frac{m_i}{\mu_N}\mu + \frac{\mu_{j-1}}{\mu_j}\mathbf{P}^j(t) - \frac{m_j}{\mu_{j+1}}\mathbf{P}^{j+1}(t) - \cdots - \frac{m_j}{\mu_N}\mathbf{P}^N(t).$$

The right hand side equals $\mathbf{p}^j(t)$ and thus $\mathbf{v}^j(t) = \mathbf{p}^j(t)$ for all $j = 0, 1, \ldots, N$. We seek a vector $\xi(t) \in \mathbb{R}^3$ such that

$$\xi(t) = \frac{\partial H}{\partial \mathbf{p}^j}(\mathbf{u}_i(t), \mathbf{v}^i(t)) - \mathbf{u}_j(t)',\ \ j = 1, \ldots, N.$$

and

$$0 = -\frac{\partial H}{\partial \mathbf{q}_j}(\mathbf{u}_i(t), \mathbf{v}^i(t)) - \mathbf{v}^j(t)'.$$

The first equation says that $\xi(t)$ is the velocity of the center of mass and that we have Hamilton's equations relative to this motion. The second group of $N+1$ equations is automatically satisfied since $\mathbf{v}^j(t) = \mathbf{p}^j(t)$. A solution of

$$\mathbf{x}(t)' = \xi(t),\ \ \mathbf{x}(0) = 0,$$

is $\mathbf{x}(t) = \int_0^t \xi(s)ds$. Then the solution $(\mathbf{q}_j(t), \mathbf{p}^j(t)),\ j = 0, 1, \ldots, N$ is given by

$$\mathbf{q}_j(t) = \mathbf{x}(t) + \mathbf{u}_j(t),\ \ \mathbf{p}^j(t) = \mathbf{v}^j(t).$$

This process of recovering $\mathbf{q}$ and $\mathbf{p}$ in terms of reduced variables is a special instance of the reconstruction process. ◆

5. Coupled Harmonic Oscillations (*cf.* Ermentrout and Kopell [1984])**.**
We prove that in a chain of coupled oscillators one can average over a single S^1-action representing the "fast" variables and that the resulting averaged equations are still Hamiltonian in the "slow" variables. The results are related to the important averaging procedure and the realization of averaged equations as equations on a center or sub-center manifold;

see Arnold [1978], §3.8, Forest, McLaughlin and Montgomery [1986] and Mielke [1991] for more information.

Start with the phase space of $N+1$ oscillators

$$T^*Q = T^*(S^1 \times \cdots \times S^1) \quad (N+1 \text{ factors})$$

with (action angle) coordinates $(\theta_1, I_1), \ldots, (\theta_{N+1}, I_{N+1})$, and a Hamiltonian of the form

$$H_\varepsilon = H_0 + \varepsilon H_1 + O(\varepsilon^2), \ \ \varepsilon \text{ a small parameter}$$

where

$$H_0 = G_1(I_1) + \cdots + G_{N+1}(I_{N+1})$$

and $G_1, \ldots, G_{N+1}$ and functions of one variable. Thus if $\varepsilon = 0$, we have $N+1$ independent oscillators and $H_1(\theta_1, \ldots, \theta_{N+1}, I_1, \ldots, I_{N+1})$ is a general coupling term. The dynamics on T^*Q is given by Hamilton's equations:

$$\begin{aligned} \dot\theta_i &= \frac{\partial H_\varepsilon}{\partial I_i} = G_i'(I_i) + \varepsilon \frac{\partial H_1}{\partial I_i} + O(\varepsilon^2) \\ \dot I_i &= -\frac{\partial H_\varepsilon}{\partial \theta_i} = -\varepsilon \frac{\partial H_1}{\partial \theta_i} + O(\varepsilon^2). \end{aligned}$$

Let $\omega_i := \partial G_i / \partial I_i$, the (uncoupled) frequency of the ith oscillator. Assume that the frequency differences are $O(\varepsilon)$; for simplicity, assume that

$$\omega_{i+1} - \omega_i = \varepsilon \Delta_i.$$

Thus, while the θ_i evolve as "fast" variables, the differences are "slow": let

$$\phi_i = \theta_{i+1} - \theta_i \ \text{ so that } \ \dot\phi_i = \varepsilon\left(\Delta_i + \frac{\partial H_1}{\partial I_{i+1}} - \frac{\partial H_1}{\partial I_i}\right).$$

We will find the equations of motion for the slow variables by averaging over the fast variables. This is accomplished by considering the following S^1-action on Q

$$(\theta, \theta_1, \ldots, \theta_N) \in S^1 \times Q \mapsto (\theta_1 + \theta, \ldots, \theta_{N+1} + \theta) \in Q$$

and then lifting it to T^*Q to get

$$\theta \cdot (\theta_1, I_1, \ldots, \theta_N, I_N) = (\theta_1 + \theta, I_1, \ldots, \theta_{N+1} + \theta, I_{N+1}).$$

Coordinatize Q/S^1 by $(\phi_1, \ldots, \phi_N)$, the projection $Q \to Q/S^1$ being given by $\phi_i = \theta_{i+1} - \theta_i$, $i = j, \ldots, N$. Let

$$\overline{H}_\varepsilon(\theta_i, I_i) = \frac{1}{2\pi} \int_0^{2\pi} H_\varepsilon(\theta_i + \theta, I_i)\, d\theta,$$

the average of H_ε over the S^1-action. (The Averaging Theorem guarantees that we get a "good" approximation to the H_ε dynamics from that of $\overline{H}_\varepsilon$ for ε small — see Arnold [1978]).

Can one express the dynamics given by $\overline{H}_\varepsilon$ only in terms of the "slow" variables $\phi_1, \ldots, \phi_N$ and their (as yet to be determined) conjugate variables $p_1, \ldots, p_N$? The answer is affirmative and this is accomplished by reduction in the following manner. Since $\overline{H}_\varepsilon$ is S^1-invariant, the momentum map of the S^1-action

$$\mathbf{J} : T^*Q \to \mathbb{R} \ \text{ given by } \ J(\theta_i, I_k) = I_1 + \cdots + I_{N+1}$$

is a constant of the motion for the $\overline{H}_\varepsilon$-dynamics. By the Cotangent Bundle Reduction Theorem, the reduced manifold $J^{-1}(I)/S^1$ is diffeomorphic to $T^*(Q/S^1)$ since S^1 is abelian. The symplectic structure on $T^*(Q/S^1)$ is the canonical one since the one-form

$$(\theta_1, \ldots, \theta_{N+1}) \mapsto \left(\theta_1, \ldots, \theta_{N+1}, \frac{1}{N+1}I, \ldots, \frac{1}{N+1}I\right)$$

has values $\mathbf{J}^{-1}(I)$, is S^1 -invariant, and being constant, is closed. By reduction and reconstruction, no information is lost by dropping to $T^*(Q/S^1)$. To find the conjugate variables p_j of ϕ_j, we set $\sum_{k=1}^N d\phi_\kappa \wedge dp_k = \sum_{i=1}^{N+1} d\theta_i \wedge dI_i$ on $\mathbf{J}^{-1}(I)$. This gives coordinates analogous to the Jacobi-Haretu coordinates for center of mass:

$$\begin{aligned}
p_1 &= -I_1, \;\; so \;\; I_1 = -p_1 \\
p_2 &= -I_1 - I_2, \;\; so \;\; I_2 = p_1 - p_2 \\
p_3 &= -I_1 - I_2 - I_3, \;\; so \;\; I_3 = p_2 - p \\
\vdots & \qquad\qquad \vdots \\
p_N &= -I_1 - \cdots - I_N, \;\; so \;\; I_N = p_{N-1} - p_N
\end{aligned}$$

Now write $\overline{H}_\varepsilon$ as a function of $\phi_1, \ldots, \phi_N$, $p_1, \ldots, p_N$ and call the resulting function $\overline{H}^I$, the reduction of $\overline{H}_\varepsilon$. Reduction tells us we should end up with Hamilton's equations. This can be checked explicitly: a chain rule argument shows that

$$\dot{\phi}_k = \frac{\partial \overline{H}^I_\varepsilon}{\partial p_k}, \quad \dot{p}_k = -\frac{\partial \overline{H}^I_\varepsilon}{\partial \phi_k}, \quad k = 1, \ldots, N$$

which is the Hamiltonian version of the equations for the evolution of phase differences. Reconstruction can be carried out as in the preceding example. ♦

3.13 Hamiltonian Systems on Coadjoint Orbits

For a Lie group G, the reduced phase space $(T^*G)_\mu$ (using the *left* action by G) coincides with the coadjoint orbit $\mathcal{O}_\mu$ with the "−" orbit symplectic structure. Let H be a left G-invariant Hamiltonian on T^*G. Since the projection $\pi_\mu : \mathbf{J}_L^{-1}(\mu) \to (T^*G)_\mu$ is given by $\pi_\mu(T_g^*R_{g^{-1}}(\mu)) = \mathrm{Ad}_g^*\mu$, the induced Hamiltonian $H_\mu : \mathcal{O}_\mu \to \mathbb{R}$ is given by $H_\mu(\mathrm{Ad}_g^*\mu) = H(T_g^*R_{g^{-1}}(\mu)) = H(\mathrm{Ad}_g^*\mu)$ by left invariance of H; *i.e.*, $H_\mu = H|\mathcal{O}_\mu$. The Hamiltonian vector field defined by H_μ on $\mathcal{O}_\mu$ has the expression

$$X_{H_\mu}(\nu) = \mathrm{ad}\left(\frac{\delta H}{\delta \nu}\right)^* \nu$$

for any $\nu \in \mathcal{O}_\mu$. We now determine the flow of X_H on

$$\mathbf{J}_L^{-1}(\mu) = \{\text{graph of the one-form}\, g \mapsto \mu \circ T_g R_{g^{-1}}\},$$

in terms of the flow of X_{H_μ} on $\mathcal{O}_\mu$.

To do this, trivialize T^*G by $\lambda : \alpha_g \mapsto (g, T_e^*L_g(\alpha_g))$ mapping T^*G to $G \times \mathfrak{g}^*$ and note that the induced action of G on $G \times \mathfrak{g}^*$ is given by $g \cdot (h, \mu) = (gh, \mu)$. The momentum map $\mathbf{J}_L$ induces an equivariant momentum map $\mathbf{J}_\lambda : G \times \mathfrak{g}^* \to \mathfrak{g}^*$ given for each $\xi \in \mathfrak{g}$ by $\mathbf{J}_\lambda(\xi) = \mathbf{J}_L(\xi) \circ \lambda^{-1}$. Therefore, for any $g \in G,\ \mu \in \mathfrak{g}^*$ we have

$$\begin{aligned}
\langle \mathbf{J}_\lambda(g, \mu), \xi \rangle &= \mathbf{J}_\lambda(\xi)(g, \mu) = \mathbf{J}_L(\xi)(\lambda^{-1}(g, \mu)) = \langle \mathbf{J}_L(T_c^*L_{g^{-1}}(\mu)), \xi \rangle \\
&= \langle T_g^*R_g \circ T_e^*L_{g^{-1}})\mu, \xi \rangle = \langle \mathrm{Ad}_{g^{-1}}^*\mu, \xi \rangle,
\end{aligned}$$

so that

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

Therefore, $\mathbf{J}_\lambda^{-1}(\mu) = G\times\mathcal{O}_\mu$ and the projection on the manifold $\mathcal{O}_\mu$ is given by the projection on the second factor. Next, we explicitly determine $\lambda_* X_H = X_{H\circ\lambda^{-1}} \in \mathfrak{X}(G\times\mathfrak{g}^*)$; the latter Hamiltonian vector field is with respect to the symplectic form $\lambda_*\Omega_0$ on $G\times\mathfrak{g}^*$, Ω_0 being the canonical sympletic form on T^*G. We have

$$\lambda_* X_H(g,\mu) = (X(g,\mu), \mu, Y(g,\mu)) \in T_gG \times \{\mu\} \times \mathfrak{g}^*$$

with $g \mapsto X(g,\mu)$ a vector field on G for each $\mu \in \mathfrak{g}^*$. Since the derivative of left translation by g on $G\times\mathfrak{g}^*$ is given by

$$g\cdot(v_h,\mu,\nu) = (T_hL_g(v_h),\mu,\nu),$$

left invariance of X_H implies left invariance of $\lambda_* X_H$, *i.e.*, the vector field $g \mapsto X(g,\mu)$ is left invariant on G for any $\mu \in \mathfrak{g}^*$ and $Y(g^{-1}h,\mu) = Y(g,\mu)$ for any $g,h \in G$. Therefore $Y(g,\mu) = Y(e,\mu)$, *i.e.*, Y is independent of G and thus letting $X(g,\mu) = X^\mu(g)$, we have

$$(\lambda_* X_H)(g,\mu) = (X^\mu(g), \mu, Y(\mu)),$$

with $X^\mu \in \mathfrak{X}(G)$ left invariant and $Y : \mathfrak{g}^* \to \mathfrak{g}^*$.

If F_t denotes the flow of X_H, then $\lambda\circ F_t\circ\lambda^{-1}$ is the flow of $\lambda_* X_H$. Letting $g_\mu(t) = (\pi_G\circ F_t)(\mu)$, where $\pi_G : T^*G \to G$ is the cotangent bundle projection, left invariance of F_t gives

$$\begin{aligned}(\lambda\circ F_t\circ\lambda^{-1})(g,\mu) &= (\lambda\circ F_t\circ T_g^*)L_{g^{-1}}(\mu) = (\lambda\circ T_g^*L_{g^{-1}})(F_t(\mu))\\ &= (gg_\mu(t), (T_e^*L_{gg_\mu}(t)\circ T_g^*L_{g^{-1}}\circ F_t)(\mu))\\ &= (gg_\mu(t), F_t(\mu)\circ T_eL_{g_\mu}(t)).\end{aligned}$$

In particular $t\mapsto gg_\mu(t)$ is the flow of X^μ; taking its derivative at $t=0$ yields

$$X^\mu(g) = T_eL_g(T_\mu\pi_G\circ X_H)(\mu).$$

We claim that

$$(T_\mu\pi_G\circ X_H)(\mu) = \frac{\delta H}{\delta\mu}.$$

This is proved in the following manner. Let $i : \mathfrak{g}^* \to T^*G$ denote the inclusion of $\mathfrak{g}^*$ in T^*G as the fiber T_e^*G and let $\alpha_\mu \in T_\mu^*(T^*G)$. Denote by $\alpha_\mu^\sharp \in T_\mu(T^*G)$ the symplectically associated vector ; *i.e.*

$$\Omega_0(\mu)(\alpha_\mu^\sharp, u_\mu) = \langle\alpha_\mu, u_\mu\rangle$$

for any $u_\mu \in T_\mu(T^*G)$. A straightforward local computation shows that

$$T_\mu\pi_G(\alpha\flat_\mu^\sharp) = T_\mu^*i(\alpha_\mu)$$

whence

$$\left[(T_\mu^*\pi_G)(\nu)\right]^\sharp = -T_\mu i(\nu)$$

which in turn implies that for any $\nu \in \mathfrak{g}^*$

$$\begin{aligned}\langle\nu, T_\mu\pi_G(X_H(\mu)\rangle &= \left\langle(T_\mu^*\pi_G)(\nu), X_H(\mu)\right\rangle\\ &= \Omega_0(\mu)([(T_\mu\pi_G)(\nu)]^\flat, X_H(\mu)) = \langle\mathbf{d}H(\mu), T_\mu i(\nu)\rangle\\ &= \langle\mathbf{d}(H\circ i)(\mu),\nu\rangle = \left\langle\nu, \frac{\delta H}{\delta\mu}\right\rangle.\end{aligned}$$

This proves our claim. As a consequence, we have

$$X^{\mu}(g) = T_e L_g \left(\frac{\delta H}{\delta \mu} \right).$$

To compute Y, we use conservation of $\mathbf{J}_L$:

$$\begin{aligned}
\langle d\mathbf{J}_L(\xi) \cdot \mu, X_H(\mu) \rangle &= \left.\frac{d}{dt}\right|_{t=0} \langle \mathbf{J}(\xi), F_t(\mu) \rangle \\
&= \left.\frac{d}{dt}\right|_{t=0} \left\langle F_t(\mu), T_e R_{g_\mu(t)}(\mu) \right\rangle \\
&= \left.\frac{d}{dt}\right|_{t=0} \left\langle F_t(\mu) \circ T_e L_{g_\mu}(t), \mathrm{Ad}_{g_\mu(t)^{-1}}(\xi) \right\rangle \\
&= \langle Y(\mu), \xi \rangle + \left\langle \mu, \left.\frac{d}{dt}\right|_{t=0} \mathrm{Ad}_{g_\mu(t)^{-1}}(\xi) \right\rangle \\
&= \langle Y(\mu), \xi \rangle - \left\langle \mu, \left[\left.\frac{dg_\mu(t)}{dt}\right|_{t=0} \xi \right] \right\rangle
\end{aligned}$$

since $t \mapsto F_t(\mu) \circ T_e L_{g_\mu(t)}$ is the flow of Y. Since

$$\left.\frac{d}{dt}\right|_{t=0} g_\mu(t) = (T_\mu \tau \circ X_H)(\mu) = \frac{\delta H}{\delta \mu},$$

we get

$$Y(\mu) = \mathrm{ad}\left(\frac{\delta H}{\delta \mu} \right)^* \mu.$$

Therefore, we have proved the following proposition.

Proposition 3.13.1 *Let $H : T^*G \to \mathbb{R}$ be left invariant, F_t be the flow of X_H and let $g_\mu(t) = (\pi_G \circ F_t)(\mu)$, where $\pi_G : T^*G \to G$ is the projection. If $\lambda : T^*G \to G \times \mathfrak{g}^*$ defined by $\lambda(\alpha_g) = (g, T_e^* L_g(\alpha_g))$ denotes the left trivialization of T^*G, then*

$$(\lambda_* X_H)(g, \mu) - \left(T_e L_g \mathrm{ad} \left(\frac{\delta H}{\delta \mu} \right)^*, \mu, \mathrm{ad} \left(\frac{\delta H}{\delta \mu} \right)^* \mu \right).$$

The flow of the vector field in the first component is $(t, g) \mapsto g g_\mu(t)$ and in the second component is $(t, \mu) \mapsto F_t(\mu) \circ T_e L_{g_\mu(t)}$.

LAGRANGIAN COUNTERPART TO BE INSERTED.

Now assume that the flow F_t^μ of X_{H_μ} is known. Since X_{H_μ} is the second component of $\lambda_* X_H$, this proposition shows that $F_t^\mu(\nu) = F_t(\nu) \circ T_e L_{g_\nu(t)}$. To reconstruct the flow of $\lambda_* X_H$, fix $(g_0, \mu_0) \in G \times \mathcal{O}_\mu$ and let $(g(t), \mu(t))$ be a curve that projects to the integral curve of $X_{H_\mu} = Y$ with initial condition μ_0. Thus, $\mu(0) = \mu_0$ and $\mu(t)$ is an integral curve of Y. In fact, $g(t)$ is arbitrary satisfying the condition $g(0) = g_0$, so take $g(t) = g_0$ for all t for definiteness. According to the expression for $\lambda_* X_H$, the equation

$$\xi(t)_{G \times \mathfrak{g}^*} (g_0 \mu(t)) = \lambda_* X_H (g_0, \mu(t)) - (0, \mu'(t))$$

is equivalent to

$$\xi(t) = \mathrm{Ad}_{g_0} \left(\frac{\delta H}{\delta \mu(t)} \right).$$

Therefore, the reconstruction equation for $h(t) \in G_\mu$ is

$$h'(t) = T_e L_{h(t)} \mathrm{Ad}_{g_0} \left(\frac{\delta H}{\delta \mu(t)} \right), \quad \text{where } h(0) = e$$

and thus the integral curve of $\lambda_* X_H$ through (g_0, μ_0) is $t \mapsto (h(t)g_0, \mu(t))$.

The Free Rigid Body Let us apply these ideas to the motion of the free rigid body. In this case, $G = SO(3)$ and $Y(\Pi) = \Pi \times \omega$ has the solution discussed in **§16.8**. The reconstruction equation for $h'(t)$ becomes a system of three ordinary differential equations in the Euler angles ϕ, ψ, θ, namely

$$\begin{bmatrix} \dot\theta \cos\psi + \dot\phi \sin\psi \sin\theta \\ -\dot\theta \sin\psi + \dot\phi \cos\psi \sin\theta \\ \dot\theta \cos\theta + \dot\psi \end{bmatrix} = P(\phi_0, \psi_0, \theta_0)\omega(t)$$

where $P(\phi_0, \psi_0, \theta_0)$ is given by §**16.6** and $\omega_i(t) = \Pi_i(t)/I_i$, with $\Pi_i(t)$, $i = 1, 2, 3$, the solution of the Euler equations with given initial conditions. ♦

The analysis above can be repeated on the tangent bundle TG, using an invariant (weakly) nondegenerate symmetric bilinear form $(\cdot\,,\cdot)$. Note that the map $\xi \in \mathfrak{g} \mapsto (\xi, \cdot) \in \mathfrak{g}^*$ is equivariant with respect to the adjoint and coadjoint actions by bi-invariance of $(\cdot\,,\cdot)$, so that by differentiating the equivariance condition, we get

$$([\eta, \xi], \cdot) = -\mathrm{ad}^*(\eta) \cdot \xi.$$

Letting ∇E denote the gradient of the function $E : \mathfrak{g} \to \mathbb{R}$ with respect to $(\cdot\,,\cdot)$ and $\overline{E} : \mathfrak{g}^* \to \mathbb{R}$ denote the function defined by $\overline{E}((\xi, \cdot)) = E(\xi)$, we have

$$\frac{\delta \overline{E}}{\delta(\xi, \cdot)} = \nabla E(\xi).$$

Therefore, the $(-)$ Lie-Poisson bracket and Hamiltonian vector field are given by

$$\{F, H\}(\xi) = -(\xi, [\nabla F(\xi), \nabla H(\xi)])$$
and
$$X_H(\xi) = [\xi, \nabla H(\xi)] = -\mathrm{ad}(\nabla H(\xi)) \cdot \xi.$$

Note the change in sign in the Hamiltonian vector field formula. As before, if $E : TG \to \mathbb{R}$ is a left invariant function, its reduction to $(TG)_{(\xi,\cdot)} = \{\text{the adjoint orbit through } \xi \in \mathfrak{g}\} =: \mathcal{O}_\xi$, is simply the restriction $E|\mathcal{O}_\xi$. Observe that $\mathcal{O}_\xi$ is a (weak) symplectic manifold with weak symplectic form given by

$$\omega_\xi(\eta)([\eta, \zeta_1], [\eta, \zeta_2]) = -(\eta, [\zeta_1, \xi_2])$$

for $\eta \in \mathcal{O}_\xi$, where ζ_1 and $\zeta_2 \in \mathfrak{g}$ are arbitrary.

Trivializing TG by left translation,

$$\lambda : TG \to G \times \mathfrak{g}, \; \lambda(v_g) = (g, T_g L_{g^{-1}}(v_g))$$

we get a left invariant vector field

$$(\lambda_* X_E)(g, \xi) = (X^\xi(g), \xi, Y(\xi)),$$

where $X^\xi \in \mathfrak{X}(g)$ is left invariant and $Y : \mathfrak{g} \to \mathfrak{g}$. But on the tangent bundle $X_E \in \mathfrak{X}(TG)$ has an additional property: it is a second order equation, *i.e.* $(T\tau_G \circ X_E)(v_g) = v_g$ for all $v_g \in T_gG$, where $\tau_G : TG \to G$ is the tangant bundle projection. Finally note that $\rho \circ \lambda = \lambda$, where $\rho : G \times \mathfrak{g} \to G$ is the projection on the first factor. Therefore

$$\begin{aligned} X^\xi(g) &= T\rho(\lambda_*(X_E)(g,\xi) = (T\rho \circ T\lambda \circ X_E \circ \lambda^{-1})(g,\xi) \\ &= (T\tau_G \circ X_E)(\lambda^{-1}(g,\xi)) = T_eL_g(\xi) \end{aligned}$$

so that

$$(\lambda_* X_E)(g,\xi) = (T_eL_g(\xi), \xi, [\xi, \nabla E(\xi)]);$$

the expression for Y is obtained by invoking the fact that on T^*G, Y was the Hamiltonian vector field in the $(-)$ Lie-Poisson structure given by the restriction of the Hamiltonian to $\mathfrak{g}^*$. Finally, note that the form of $\lambda_* X_E$ is already appropriate for the reconstruction of solutions. First one solves the equation $\dot{\xi} = [\xi, \nabla E(\xi)]$ and then the equation $\dot{g} = T_eL_g(\xi)$ with $\xi(t)$ found previously. Thus, we have proved the following:

Proposition 3.13.2 *Let $E : TG \to \mathbb{R}$ be a left invariant function on a Lie group G that has a weak bi-invariant pseudo-metric. If $\lambda : TG \to G \times \mathfrak{g}$ defined by $\lambda(v_g) = (g, T_gL_{g-1}(v_g))$ is the left trivialization of T^*G, we have:*

$$(\lambda_* X_E)(g,\xi) = (T_eL_g(\xi), \xi, [\xi, \nabla E(\xi)]).$$

To solve the differential equations associated to the vector field $\lambda_* X_E$, one first solves the reduced equations $\dot{\xi} = [\xi, \nabla E(\xi)]$, and then the equation on G given by $\dot{g} = T_eL_g(\xi)$.

3.14 Energy Momentum Integrators

Sometimes one may wish to preserve the energy and the momentum map in a scheme rather than the symplectic structure and the momentum map. For the actual implementation of schemes like this, we refer to Simo and Wong [1990] and Austin and Krishnaprasad [1990]. We present here an abstraction of this process.

We now turn to some basic remarks on the construction of algorithms that conserves the Hamiltonian and the momentum map, but will not, in general, conserve the symplectic structure.

A class of algorithms satisfying this requirement can be obtained through the steps outlined below. The geometry of the process is depicted in Figure 2.9.1.

i Formulate an energy-preserving algorithm on the symplectic reduced phase space $P_\mu = \mathbf{J}^{-1}(\mu)/G_\mu$ or the Poisson reduced space P/G. If such an algorithm is interpreted in terms of the primitive phase space P, it becomes an iterative mapping from one orbit of the group action to another.

ii In terms of canonical coordinates (q,p) on P, interpret the orbit-to-orbit mapping described above and if P/G was used, impose the constraint $\mathbf{J}(q_k,p_k) = \mathbf{J}(q_{k+1},p_{k+1})$. The constraint does not uniquely determine the restricted mapping, so we may obtain a large class if iterative schemes.

iii To uniquely determine a map from within the above class, we must determine how points in one G_μ-orbit are mapped to points in another orbit. There is still an ambiguity about how phase space points drift in the G_μ-orbit directions. This drift is closely

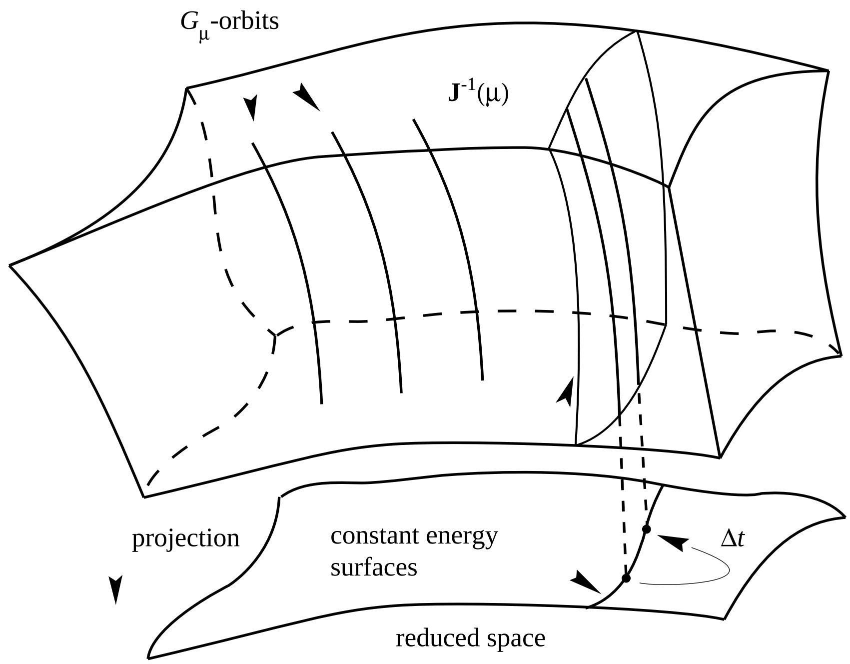

Figure 3.14.1: An energy preserving algorithm is designed on the reduced space and is then lifted to the level set of the momentum map by specifying phase information.

connected with *geometric phases* (Chapter 6)! In fact by discretizing the geometric phase formula for the system under consideration we can specify the shift along each G_μ-orbit associated with each iteration of the map.

The papers of Simo and Wong [1989] and Krishnaprasad and Austin [1990] provide systematic methods for making the choices required in steps **ii** and **iii**. The general construction given above is, in fact, precisely the approach advocated in ref Simo, Tarnow and Wong [1991]. There it is shown that projection from the level set of constant angular momentum onto the surface of constant energy can be performed implicitly or explicitly leading to predictor/corrector type of algorithms. From a numerical analysis standpoint, the nice thing is that the cost involved in the actual construction of the projection reduces to that of a line search (i.e., basically for free). The algorithm advocated in Simo and Wong [1989] is special in the sense that the projection is not needed for $Q = SO(3)$: the discrete flow is shown to lie in the intersection of the level set of angular momentum and the surface of constant energy. This algorithm is singularity–free and integrates the dynamics exactly up to a time reparametrization, consitent with the restrictions on mechanical integrators given in §9.2. Extensions of these schemes to elasticity, rods and shells suitable for large-scale calculation and amenable to paralelization are given in Simo, Fox and Rifai [1991], and Simo and Doblare [1991].

3.15 Maxwell's Equations

In this section we apply the techniques of reduction to the case of Maxwell's equations. Here we shall study the Maxwell equations, possibly with a time-independent charge source; they can be coupled with the Vlasov equation, which provides a time-dependent charge and current source; see Marsden and Weinstein [1982].

The Maxwell equations are

$$\frac{1}{c}\frac{\partial \mathbf{B}}{\partial t} = -\mathrm{curl}\,\mathbf{E} \tag{3.15.1}$$

$$\frac{1}{c}\frac{\partial \mathbf{E}}{\partial t} = \mathrm{curl}\mathbf{B} - \frac{\mathbf{j}}{c} \tag{3.15.2}$$

$$\mathrm{div}\ \mathbf{E} = \rho \tag{3.15.3}$$

$$\mathrm{div}\ \mathbf{B} = 0 \tag{3.15.4}$$

where $\mathbf{E}(\mathbf{x}, t)$ is the electric field, $\mathbf{B}(\mathbf{x}, t)$ is the magnetic field, $\mathbf{x} \in \mathbb{R}^3$, c is the velocity of light $\mathbf{j}(\mathbf{x})$ is a given current and $\rho(\mathbf{x})$ is a given charge density. For simplicity we take the equations to be defined in all of space $\mathbb{R}^3$ with suitable fall off conditions at infinity to justify integration by parts. [Incorporation of boundary conditions, radiation, etc., is very important but is not treated here.] To begin with, we will assume $\mathbf{j} = \mathbf{0}$.

As the configuration space for Maxwell's equations, take the space $\mathfrak{A}$ of one forms $\mathbf{A}$ on $\mathbb{R}^3$. (These are the "vector potentials." In more general situations involving Yang- Mills fields, one can replace $\mathfrak{A}$ by the set of connections on a principal bundle over configuration space.) The corresponding phase space is the cotangent bundle $T^*\mathfrak{A}$, with the canonical symplectic structure. Elements of $T^*\mathfrak{A}$ are identified with pairs (A, Y), where Y is a vector field density on $\mathbb{R}^3$. (We do not distinguish Y and $Y d^3\mathbf{x}$.) The pairing between A's and Y's is given by integration, so that the canonical symplectic structure ω on $T^*\mathfrak{A}$ is given by

$$\omega((\mathbf{A}_1, \mathbf{Y}_1), (\mathbf{A}_2, \mathbf{Y}_2)) = \int (\mathbf{Y}_2 \cdot \mathbf{A}_1 - \mathbf{Y}_1 \cdot \mathbf{A}_2)\ d^3\mathbf{x}, \tag{3.15.5}$$

with associated Poisson bracket

$$\{F, G\} = \int \left(\frac{\delta F}{\delta \mathbf{A}} \frac{\delta G}{\delta \mathbf{Y}} - \frac{\delta F}{\delta \mathbf{Y}} \frac{\delta G}{\delta \mathbf{A}} \right)\ d^3\mathbf{x}. \tag{3.15.6}$$

We will want to set $\mathbf{B} = \nabla \times \mathbf{A}$ in accordance with (2.9.4) and customary practice.

The *Hamiltonian density* for Maxwell's equations is $\mathcal{H} = \frac{1}{2}(\|\mathbf{E}\|^2 + \|\mathbf{B}\|^2)$ and the *Hamiltonian* is

$$H = \int c\mathcal{H}\ d^3x \tag{3.15.7}$$

— we have inserted a factor of c to conform to the dynamical formulation in relativity (see Misner, Thorne and Wheeler [1972], p. 523). Thus we take

$$H(\mathbf{A}, \mathbf{Y}) = \int c \left[\frac{1}{2} \left(\|Y\|^2 + \|\mathrm{curl}\ \mathbf{A}\|^2 \right) \right]\ d^3x. \tag{3.15.8}$$

Hamilton's equations are

$$\frac{\partial \mathbf{A}}{\partial t} = \frac{\delta H}{\delta \mathbf{Y}} = c\mathbf{Y} \tag{3.15.9}$$

and

$$\frac{\partial \mathbf{Y}}{\partial t} = -\frac{\delta H}{\delta \mathbf{A}} = -c\ \mathrm{curl}\ \mathrm{curl}\ \mathbf{A}. \tag{3.15.10}$$

The calculation of $\delta H/\delta \mathbf{A}$ is as follows: first,

$$\mathbf{D}H(\mathbf{A}, Y) \cdot \delta\mathbf{A} = c \int_{\mathbb{R}^3} (\text{curl } \mathbf{A}) \cdot (\text{curl } \mathbf{A}) \; d^3x.$$

The operator "curl" is symmetric, as is seen using the vector identity

$$\text{div } (\mathbf{X} \times \mathbf{Y}) = \mathbf{X} \cdot \text{curl } \mathbf{Y} - \mathbf{Y} \cdot \text{curl } \mathbf{X}$$

and the divergence theorem. Hence

$$\frac{\delta H}{\delta \mathbf{A}} = c \text{ curl curl } \mathbf{A} = c[\text{grad div } \mathbf{A} - \nabla^2 \mathbf{A}],$$

so (2.10.10) follows.

If we let $\mathbf{B} = \text{curl } \mathbf{A}$ and $\mathbf{Y} = -\mathbf{E}$, then (2.10.10) reduces to (2.10.2) and the curl of (2.10.9) gives (2.10.1).

The remaining two Maxwell equations will appear as a consequence of gauge invariance. Notice that we work directly with three-dimensional fields. Four dimensionally, one has an extra "degree" of gauge freedom associated with the time derivative $\partial_t \psi$. We have already eliminated this freedom and the corresponding non-dynamical field A_4 (whose conjugate momentum vanishes). This is the standard Dirac procedure for a relativistic field theory such as Maxwell's equations (see Gimmsy [1991] for details). In the context of principal bundles, $\mathcal{G}$ is defined to be the group of bundle automorphisms (covering the identity). The gauge group $\mathcal{G}$ consists of real valued functions on $\mathbb{R}^3$; with the group operation being addition. An element $\psi \in \mathcal{G}$ acts on $\mathfrak{A}$ by

$$\mathbf{A} \mapsto \mathbf{A} + \nabla\psi. \tag{3.15.11}$$

This "translation" of A extends by cotangent lift to a canonical transformation of $T^*\mathfrak{A}$ given by

$$(\mathbf{A}, \mathbf{Y}) \mapsto (\mathbf{A} + \nabla\psi, \mathbf{Y}). \tag{3.15.12}$$

Since the Hamiltonian is invariant under these transformations, we can use the gauge symmetries to reduce the degrees of freedom of our system. The action of $\mathcal{G}$ on $T^*\mathfrak{A}$ has a momentum map $J : T^*\mathfrak{A} \to \mathfrak{g}^*$, where $\mathfrak{g}$, the Lie algebra of $\mathcal{G}$, is identified with the real valued functions on $\mathbb{R}^3$. This map may be determined by our formula for the momentum map of a cotangent lift: for $\phi \in \mathfrak{g}$,

$$\langle J(\mathbf{A}, \mathbf{Y}), \phi \rangle = \int (\mathbf{Y} \cdot \nabla\phi) \; d^3\mathbf{x} = -\int (\text{div } \mathbf{Y})\phi \; d^3\mathbf{x}. \tag{3.15.13}$$

Thus, we may write

$$J(\mathbf{A}, \mathbf{Y}) = -\text{div } \mathbf{Y}. \tag{3.15.14}$$

If ρ is an element of $\mathfrak{g}^*$ (the densities on $\mathbb{R}^3$), $\mathbf{J}^{-1}(\rho) = \{(\mathbf{A}, \mathbf{Y}) \in T^*\mathfrak{A} \mid \text{div } \mathbf{Y} = -\rho\}$. In terms of $\mathbf{E}$, the condition $\text{div } \mathbf{Y} = -\rho$ becomes the Maxwell equation $\text{div } \mathbf{E} = \rho$, *so we may interpret the elements of* $\mathfrak{g}^*$ *as charge densities.*

Now we apply symplectic reduction to $P = T^*\mathfrak{A}$, with $\mu = \rho$ and $\mathcal{G}$ and J as above.

Theorem 3.15.1 *The reduced symplectic manifold* $\mathbf{J}^{-1}(\rho)/\mathcal{G}$ *can be identified with* $Max_\rho = \{(\mathbf{E}, \mathbf{B}) \mid \text{ div } \mathbf{E} = \rho, \; div \; \mathbf{B} = 0\}$, *and the induced Poisson bracket on* Max_ρ *is given in terms of* $\mathbf{E}$ *and* $\mathbf{B}$ *by the* ***Born-Infeld-Pauli bracket***

$$\{\{F, G\}\} = \int \left(\frac{\delta F}{\delta \mathbf{E}} \text{curl} \frac{\delta G}{\delta \mathbf{B}} - \frac{\delta G}{\delta \mathbf{E}} \text{curl} \frac{\delta F}{\delta \mathbf{B}} \right) \; d^3\mathbf{x}. \tag{3.15.15}$$

Moreover, Maxwell's equations with an ambient charge density ρ *are Hamilton's equations for the Hamiltonian on the space* Max_ρ.

Proof. To each $(\mathbf{A},\mathbf{Y})$ in $J^{-1}(\rho)$, associate the pair $(\mathbf{B},\mathbf{E}) = \operatorname{curl} A, -\mathbf{Y}/c)$ in Max_ρ. Since two vector fields $\mathbf{A}_1$ and $\mathbf{A}_2$ on $\mathbb{R}^3$ have the same curl if and only if they differ by a gradient, and every divergence-free $\mathbf{B}$ is a curl, this association gives a $1-1$ correspondence between $J^{-1}(\rho)/\mathcal{G}$ and Max_ρ.

Let F and G be functions on Max_ρ. To compute their Poisson bracket $\{\{F,G\}\}$ we pull them back to $J^{-1}(\rho)$, extend them to $T^*\mathfrak{A}$ as $\mathcal{G}$-invariant functions, take the canonical Poisson bracket in $T^*\mathfrak{A}$, restrict to $J^{-1}(\rho)$, and "push down" the resulting $\mathcal{G}$-invariant function to Max_ρ. The result does not depend upon the choice of extension made, and in fact we can do the computation without mentioning the extension again. This follows by the general theory, but we can see it directly here. Given $F(\mathbf{B},\mathbf{E})$, define the pull back $\hat{F}(\mathbf{A},\mathbf{Y})$ by

$$\hat{F}(\mathbf{A},\mathbf{Y}) = F(\operatorname{curl}\mathbf{A},\mathbf{Y}).$$

Using the canonical bracket on $T^*\mathfrak{A}$, we have

$$\begin{aligned}\{\{F,G\}\} &= \{\hat{F},\hat{G}\} = \int\left(\frac{\delta\hat{F}}{\delta\mathbf{A}}\frac{\delta\hat{G}}{\delta\mathbf{Y}} - \frac{\delta\hat{G}}{\delta\mathbf{A}}\frac{\delta\hat{F}}{\delta\mathbf{Y}}\right)d\mathbf{x}\\ &= -\int\left(\frac{\delta\hat{F}}{\delta\mathbf{A}}\frac{\delta\hat{G}}{\delta\mathbf{E}} - \frac{\delta\hat{G}}{\delta\mathbf{A}}\frac{\delta\hat{F}}{\delta\mathbf{E}}\right)d\mathbf{x}. \quad (\mathbf{2})\end{aligned} \tag{3.15.16}$$

The chain rule, the definition of functional derivatives, and integration by parts give the identity

$$\int\frac{\delta\hat{F}}{\delta\mathbf{A}}\cdot\mathbf{A}'\,d\mathbf{x} = \int\frac{\delta F}{\delta\mathbf{B}}\cdot\operatorname{curl}\mathbf{A}'\,d\mathbf{x} = \int\mathbf{A}'\cdot\operatorname{curl}\frac{\delta F}{\delta\mathbf{B}}\,d\mathbf{x}. \quad (3) \tag{3.15.17}$$

Substituting (2.10.17) into (2.10.16) produces the desired bracket (2.10.15).

The final statement on dynamics follows from the reduction of dynamics theorem . ■

Remarks.

1. As $c\to\infty$, Maxwell's equations formally turn from electrodynamics to electrostatics; the time dependence becomes undetermined. This is consistent with **(7)** of **§3.5A** in which $H\to\infty$. Alternatively, one could move the c in H to include a factor of c in the Poisson structure.

2. Since $\mathcal{G}$ is abelian, the cotangent bundle reduction theorem implies that Max_ρ is symplectically diffeomorphic to $T^*(\mathfrak{A}/\mathcal{G})$ with the canonical structure (one can check that there is no "magnetic term" in this case). As above, $\mathfrak{A}/\mathcal{G}$ is the space of $\mathbf{B}$'s with $\operatorname{div}\mathbf{B}=0$. A variable conjugate to $\mathbf{B}$, however, is *not* $\mathbf{E}$ (since the bracket (2.10.15) is not canonical). Instead, write $\mathbf{E} = -\nabla\varphi + \frac{1}{c}\nabla\times\pi$ for a fixed function φ determined by $\nabla^2\varphi=\rho$ and a vector field π. Then $\mathbf{B}$ and π are conjugate variables on $T^*(\mathfrak{A}/\mathcal{G})$. Notice that $\mathbf{E}+\nabla\varphi$ acts like the momentum shift that appears in the cotangent bundle reduction theorem. Another procedure is to use a gauge fixing condition to identify $\mathfrak{A}/\mathcal{G}$ with $\mathbf{A}$'s satisfying, for example, $\operatorname{div}\mathbf{A}=0$. Then the canonical variables are $\mathbf{A}$ and $\mathbf{Y}+\frac{1}{c}\nabla\varphi$. However, (2.10.15) seems to be more useful than either of these alternatives.

3. In calculating, for example, $\delta F/\delta\mathbf{B}$, one must remember the $\operatorname{div}\mathbf{B}=0$ constraint. For example, if $F(\mathbf{B}) = \int\mathbf{B}\cdot\nabla\varphi\,d^3\mathbf{x}$ for fixed φ, then as F is zero, $\delta F/\delta\mathbf{B}=0$, so that $\delta F/\delta\mathbf{B}\neq\nabla\varphi$.

4. From the point of view of Poisson reduction one can also construct $Max = (T^*\mathfrak{A})/\mathcal{G} \cong \{(\mathbf{E}, \mathbf{B}) \mid \text{ div } \mathbf{B} = 0\}$ and (2.10.15) defines the induced Poisson structure. The spaces Max_ρ for each ρ define the symplectic leaves. This is a special case of the relation between symplectic and Poisson reduction (see Marsden and Ratiu [1986]). ♦

Poynting's vector is related to the momentum map on $T^*\mathfrak{A}$ for the action of translations. Let $G = \mathbb{R}^3$ and for $\mathbf{a} \in \mathbb{R}^3$, let

$$(\mathbf{a} \cdot \mathbf{A})(\mathbf{x}) = \mathbf{A}(\mathbf{x} - \mathbf{a}). \tag{3.15.18}$$

The momentum map for this action is determined by the cotangent lift formula:

$$\begin{aligned}
\langle \mathbf{J}(\mathbf{A}, \mathbf{Y}), \xi \rangle &= -\int \langle \mathbf{Y}(\mathbf{x}), \mathbf{D}A(\mathbf{x}) \cdot \xi \rangle = \int E^i \frac{\delta A_i}{\delta x^j} \xi^j \, d^3x \\
&= \int E^i \left(\frac{\delta A_i}{\delta x^j} \xi^j - \frac{\delta A_j}{\delta x^i} \xi^j \right) d^3x \\
&\quad - \int \rho A_j \xi^j \, d^3x
\end{aligned} \tag{3.15.19}$$

since div $\mathbf{E} = \rho$. Hence,

$$\langle \mathbf{J}(\mathbf{A}, \mathbf{Y}), \xi \rangle = \int \mathbf{E} \cdot (\mathbf{B} \times \xi) \, d^3x - \int \rho A_i \xi^j d^3x = \int \xi \cdot [(\mathbf{E} \times \mathbf{B}) - \rho \mathbf{A}] \, d^3x.$$

Thus,

$$\langle \mathbf{J}(\mathbf{A}, \mathbf{Y}) \rangle = \int [\mathbf{E} \times \mathbf{B} - \rho \mathbf{A}] \, d^3x. \tag{3.15.20}$$

This is interpreted as the linear momentum of the electromagnetic field and $\mathbf{E} \times \mathbf{B}$ is called the ***Poynting vector***. One can verify directly from Maxwell's equations that $d\mathbf{J}/dt = 0$.

This action does not commute with the $\mathcal{G}$-action, so we must compute the momentum map for the translation action of $\mathbb{R}^3$ on Max separately. We claim that if $\rho = 0$, then

$$\mathbf{J}(\mathbf{E}, \mathbf{B}) = \int \mathbf{E} \times \mathbf{B} \, d^3x. \tag{3.15.21}$$

To check this, write $J_\xi(\mathbf{E}, \mathbf{B}) = \int \xi \cdot (\mathbf{E} \times \mathbf{B}) \, d^3x$ and compute using the Born-Infeld-Pauli bracket:

$$\{F, J_\xi\} = \int \left(\frac{\delta F}{\delta \mathbf{E}} \text{curl } (\xi \times \mathbf{E}) + (\xi \times \mathbf{B}) \cdot \text{curl} \frac{\delta F}{\delta \mathbf{B}} \right) d^3x$$

and

$$\xi_{Max}[F] = \int \left(\frac{\delta F}{\delta \mathbf{E}} \cdot D\mathbf{E}(x) \cdot \xi + \frac{\delta F}{\delta \mathbf{B}} D\mathbf{B}(x) \cdot \xi \right) d^3x.$$

Now

$$\{F, J_\xi\} = \int \frac{\delta F}{\delta \mathbf{E}} \cdot [\xi \cdot \nabla \mathbf{E}] \, d^3x + \int \frac{\delta F}{\delta \mathbf{B}} \cdot [\xi \cdot \nabla \mathbf{B}] = \xi_{Max}[F]$$

using symmetry of "curl," div $\mathbf{E} = 0$ and div $\mathbf{B} = 0$. Thus, (2.10.21) is proved. The reader can work out conservation of angular momentum in a similar way.

Aside If the metric is treated as a variable (whose evolution is not specified) then the entire diffeomorphism group becomes a symmetry group. The corresponding momentum map yields the stress energy tensor and this approach is closely related to the classical Belinfante-Rosenfeld formula: stress = ∂(Lagrangian)/∂(metric); *cf.* Simo and Marsden [1984], Hughes and Marsden [1983] and Hawking and Ellis [1972] for more information, and Gimmsy [1991] for a modern proof of this basic theorem.

3.16 Geometric Phases for the Rigid Body

In this section we derive Montogomery's [1990] formula for the phase of a rigid body. We will do this in two ways. First, it is derived by a "bare hands" method and second, it is obtained as the holonomy of a connection on the bundle $J^{-1}(\mu) \to P_\mu$.

It will be useful to recall some notation concerning the rigid body. Let $A(t) \in SO(3)$ be the configuration of the body and $\alpha_A \in T_A^*SO(3)$ the momentum. The body angular momentum $\Pi \in \mathbb{R}^3$ is related to α_A by

$$\left\langle \alpha_A, \dot{A} \right\rangle = \Pi \cdot V = -\frac{1}{2}\mathrm{tr}(\hat{\Pi}\hat{V})$$

where $\hat{V} = A^{-1}\dot{A}$. Similarly, the spatial angular momentum π satisfies

$$\left\langle \alpha_A, \dot{A} \right\rangle = \pi \cdot v = -\frac{1}{2}\mathrm{tr}(\hat{\pi}\hat{v})$$

where $\hat{v} = \dot{A}A^{-1}$. Thus,

$$\mathrm{tr}(\hat{\Pi}A^{-1}\dot{A}) = \mathrm{tr}(\hat{\pi}\dot{A}A^{-1}) = \mathrm{tr}(A^{-1}\hat{\pi}\dot{A})$$

and since $\dot{A}$ is arbitrary, we get

$$\hat{\Pi}A^{-1} = A^{-1}\hat{\pi}, \quad i.e., \quad \pi = A\Pi.$$

Now suppose that Π undergoes a periodic motion with period T so $\Pi(T) = \Pi(0)$ while π is constant; thus,

$$\pi = A(T)\Pi(T) = A(0)\Pi(0)$$

gives

$$A(T)^{-1}\pi = A(0)^{-1}\pi \quad \text{or} \quad A(T)A(0)^{-1}\pi = \pi.$$

Thus, $A(T)A(0)^{-1}$ is a rotation about the axis π. The angle of rotation, denoted $\Delta\theta$ is the ***geometric phase***. Thus, by definition,

$$A(t)A(0)^{-1} = \exp[(\Delta\theta)\hat{\pi}].$$

Our aim is to derive a formula for $\Delta\theta$. To obtain this formula, we will apply Stokes' theorem to a surface S with bounding curve C and the canonical one form Θ; *i.e.*,

$$\int_S \Omega = -\int_C \Theta.$$

The curve C and the surface S are to be chosen in $J^{-1}(\pi)$. We choose $C = C_1 - C_2$, where C_1 is the curve starting at $\alpha(0) \in T^*_{A(0)}SO(3)$, ending at $\alpha(t) \in T^*_{A(t)}SO(3)$ and given by the dynamics. The curve C_2 is the curve $\sigma_2(\lambda)$ in the set $\{\alpha_A \in T_A^*SO(3) \mid \pi \text{ is fixed and } \Pi = \Pi(0) = \Pi(T)\}$ given as follows. Let

$$A(\lambda) = \exp[(\lambda\Delta\theta)\hat{\pi}]A(0) \quad \text{and} \quad c_2(\lambda) \in T^*_{A(\lambda)}SO(3)$$

be given by

$$\left\langle c_2(\lambda), \dot{A} \right\rangle = -\frac{1}{2}\mathrm{tr}(\hat{\Pi}A(\lambda)^{-1}\dot{A})$$

i.e., $c_2(\lambda)$ is left translation of $\hat{\Pi}$ to the point $A(\lambda)$. Since

$$\begin{aligned}
\mathrm{tr}(\hat{\Pi}A(\lambda)^{-1}\dot{A}) &= \mathrm{tr}(A(0)^{-1}\hat{\pi}A(0)A(\lambda)^{-1}\dot{A}) \\
&= \mathrm{tr}(A(\lambda)^{-1}\exp[-(\lambda\Delta\theta)\hat{\pi}]\hat{\pi}\exp[(\lambda\Delta\theta)\hat{\pi}] \\
&= \mathrm{tr}(A(\lambda)^{-1}\hat{\pi}\dot{A}) = \mathrm{tr}(\hat{\pi}\dot{A}A(\lambda)^{-1})
\end{aligned}$$

Figure 3.16.1: **The curves used in the geometric phase formula**

we see that $c_2(\lambda) \in J^{-1}(\pi)$.

If S is a surface spanning C in $J^{-1}(\pi)$ and projecting to the spherical cap D enclosed by the curve $\Pi(t)$, then

$$\int_S \Omega = \int_D \omega$$

where ω is the reduced symplectic form on S^2. However, ω is $-\|\Pi\|$ times the area element on the sphere, so

$$\int_S \Omega = -\|\Pi\|\Lambda \qquad \mathbb{S}IGN?$$

where Λ is the solid angle determined by D. (If $\Lambda > 2\pi$, one has to be careful with the sign). Next we turn to the evaluation of

$$\int_C \Theta = \int_{C_1} \Theta - \int_{C_2} \Theta.$$

In general, if a Hamiltonian is pure kinetic energy and we integrate the canonical one form along a dynamic curve, we get

$$\int_c p\, dq = \int_c (g_{ij}\dot{q}^i)\dot{q}^j\, dt = 2ET.$$

Thus,

$$\int_{C_1} \Theta = 2ET$$

where E is the energy of the trajectory $\alpha(t)$, *i.e.*, the energy of $\Pi(t)$, and T is the period, as before.

Finally,

$$\begin{aligned}
\int_{C_2} \Theta &= \int_0^1 \langle c_2(\lambda), A'(\lambda)\rangle\, d\lambda \\
&= \int_0^1 -\frac{1}{2}\mathrm{tr}(\hat{\pi} A'(\lambda)A(\lambda)^{-1}) \\
&= \int_0^1 -\frac{1}{2}\mathrm{tr}(\hat{\pi} \cdot \Delta\theta\hat{\pi}) \\
&= \Delta\theta\|\pi\|^2 = \Delta\theta\|\Pi\|^2.
\end{aligned}$$

Putting this together,

$$-\|\Pi\|\Lambda = 2ET - \Delta\theta\|\Pi\|^2$$

so

$$\Delta\theta = \frac{\Lambda}{\|\Pi\|} + \frac{2ET}{\|\Pi\|^2},$$

which is the desired formula.

In the argument above, we assumed that there is a surface that spans the closed curve constructed and to which Stokes theorem was applied. Here we show that there really is such a surface, following an argument of Montgomery.

We note right away that there are other proofs that do not use or need this fact, but get around a similar point in other ways. In particular, the proof given in Marsden, Montgomery and Ratiu [1990] (pages 16 and 41, and also page 47), which is based on formulas from the geometry of bundles for holonomy, uses the fact that the curvature of a connection drops to the base and the fact that there is a local section of the principal bundle over the contractible spherical cap bounding the closed orbit on the sphere (chosen not to cross any separatrices). In this argument, Stokes' theorem is applied to the closed curve on the *base*, so that *this question of spanning does not arise*. We note that in both this proof, as well as in the argument below, one uses that fact from topology that the principle bundle over the contractible spherical cap is trivial.

Proof of spanning for rigid body phases Let J be the standard momentum map (given by right translation to the identity) of the left $SO(3)$ action on T*SO(3), and μ be a fixed nonzero value of J. Let $\pi : J^{-1}(\mu) \to S^2$ be the projection to the two-sphere (given as usual by left translation to the identity).

Let z_1 be a given point in $J^{-1}(\mu)$ whose projection $p_1 = \pi(z_1)$ to the sphere does not lie on a separatrix. Let γ be the orbit of p_1 under the reduced flow (as defined by Euler's equations) and T_1 the period of this orbit. Note that γ bounds a disc $D \subset S^2$ which does not intersect any seperatrix. This disc contains a unique stable fixed point p_0 on the sphere. Let z_0 be such that $\pi(z_0) = p_0$.

Let $G_\theta : T^*SO(3) \to T^*SO(3)$ be the action of the one-parameter subgroup generated by $-n$, where n is the unit vector along the direction of the constant chosen value μ of J. Thus, $\mu = \|\mu\|n$. (Note the minus sign in the definition of G_θ).

Let $F_t : T^*SO(3) \to T^*SO(3)$ be the flow of the rigid body Hamiltonian. Consider the curve $C_1(t) = F_t(z_1)$ for $0 \le t \le T_1$. Since $\pi \circ C_1 = \gamma$ is closed and since the sphere is the quotient of $J^{-1}(\mu)$ by the G action, we know that there exists an **angle** θ_1 such that $G_{\theta_1}(C_1(T_1)) = z_1$. Our goal is to compute this angle. Here, θ_1 is an angle computed mod 2π. Note that $G_\theta = G_{\theta+2\pi}$.

To do this computation, we close up C_1 by adding a curve generated by G and then applying Stokes' theorem to the resulting closed curve. To apply Stokes' theorem, we will choose this closing up curve in such a way that it bounds a disc that lies in $\pi^{-1}(D)$. As we will see, this choice is tantamount to choosing the correct "branch" for the angle θ_1.

Let k be an integer (positive or negative) to be determined below. Let θ_1^* be branch of θ_1 corresponding to this choice of k, *i.e.,* a real number representing the angle θ_1 lying in the interval $(2k\pi, 2(k+1)\pi]$. Thus, if we represent the angle θ_1 by a real number θ_1^0 lying in the interval $(0, 2\pi]$ then $\theta_1^* = \theta_1^0 + 2\pi k$. To emphasize the dependence on k, we will write $\theta_1^* = \theta_k$.

Let $C_{0,k}$ denote the curve $C_{0,k}(t) = G_t(C_1(T_1)), 0 \le t \le \theta_k$. Let $\sigma_k = C_{0,k} * C_1$ denote the concatenation of the curves $C_{0,k}$ and C_1. Thus $\sigma_k(t) = C_1(t)$ for $0 \le t \le T_1$ and $\sigma_k(t) = C_{0,k}(t - T_1)$ when $T_1 \le t \le T_1 + \theta_k$.

By definition of the angle θ_1, all of the loops σ_k are closed. Let $L(t) = G_t(z_1), 0 \le t \le 2\pi$ denote one orbit of the circle through our base point and let kL denote the kth iterate of such an orbit. If k is negative then $kL(t) = G_{-t}, 0 \le t \le |k|2\pi$. Observe that $C_{0,k} = kL * C_{0,0}$. This also gives meaning to $C_{0,k}$ for k a negative integer. It also follows that

$$\sigma_k = kL * \sigma_0.$$

Lemma 3.16.1 *The closed loop σ_k lies in the set $\pi^{-1}(\gamma)$, which is a two-torus. For appropriate choice of the integer k, this closed loop σ_k bounds a disc which lies entirely in the set $M = J^{-1}(\mu)$.*

Proof of the lemma The first statement follows immediately from the definitions of the curve and the commutivity of the flows F and G.

To prove the second fact, let $[\sigma_k] \in H_1(M, Z)$ denote the integer homology class represented by our closed loop σ_k. Recall that the reduction map $\pi : J^{-1}(\mu) \to S^2$ gives $J^{-1}(\mu)$ the structure of a principal circle bundle with the circle action being G. Since D is contractible, it follows that $M = \pi^{-1}(D)$ is equivariantly diffeomorphic to the solid torus $S^1 \times D$. Now the homology of $S^1 \times D$ is isomorphic to the group of integers and any circle orbit $S^1 \times \{\text{point}\}$ represents the generator. Thus the homology class $[L]$ represented by the loop L defined above generates the homology of M. Consequently $[\sigma_0] = -k[L]$ for some integer k. But then $[\sigma_k] = [\sigma_0] + k[L] = 0$. (This equality follows from the relation $\sigma_k = kL * \sigma_0$ and the relation between the first homotopy and homology groups.) But this means that σ_k bounds a disc. ■

Remarks The integer k is interesting and has the physical meaning of how many full revolutions the rigid body made about the axis n during the period T_1. Since its axis of rotation varies with time, it takes a bit of thought to make sense of this statement–to begin, one can imagine p_1 very close to p_0, in which case z_1 is close to z_0, the relative equilibrium. At the relative equilibrium the motion is simply rotation about the axis n with constant angular velocity. Near equilibrium, its instantaneous axes of rotation are near n and so we can view it as approximately rotating about n. Mathematically, one says all this precisely by defining θ_k to be a winding number relative to the solid torus M. Cushman and Levi have shown how to calculate the integer k and hence the **number** θ_k.

3.17 Reconstruction Phases

The general setting is as follows. Consider a symplectic manifold P, a group action with an (equivariant) momentum map $\mathbf{J} : P \to \mathfrak{g}^*$ and a G -invariant Hamiltonian H. Suppose μ is a regular value of $\mathbf{J}$, and the reduced space P_μ is smooth. We consider the problem of reconstructing trajectories in P from those in P_μ.

Assume $\pi_\mu : \mathbf{J}^{-1}(\mu) \to P_\mu$ is a principal G_μ-bundle with connection A, so A is a $\mathfrak{g}_\mu$-valued one form on $\mathbf{J}^{-1}(\mu)$ satisfying

$$\text{(a)} \quad A_p \cdot \xi_P(p) = \xi \quad \text{for} \quad \xi \in \mathfrak{g}_\mu$$

and

$$\text{(b)} \quad \Phi_g^* A = \mathrm{Ad}_g \circ A \quad \text{for} \quad g \in G_\mu$$

where $\Phi_g : \mathbf{J}^{-1}(\mu) \to \mathbf{J}^{-1}(\mu)$ is the G_μ action.

Let $c_\mu(t)$ be an integral curve on P_μ and let $p_0 \in \mathbf{J}^{-1}(\mu)$ project to $c_\mu(0)$. Proceed as follows:

1. *Horizontally lift $c_\mu(t)$ from p_0 to get a curve $d_\mu(t)$ in $\mathbf{J}^{-1}(\mu)$; i.e., $A \cdot \left[\frac{d}{dt} d_\mu(t)\right] = 0$, $d_\mu(t)$ projects to $c_\mu(t)$ and $d_\mu(0) = p_0$.*
2. *Let $\xi(t) = A \cdot X_H(d(t)) \in \mathfrak{g}_\mu$.*
3. *Solve $\dot{g}(t) = g(t) \cdot \xi(t)$; $g(0) = e$.*

Theorem 3.17.1 *$c(t) = g(t) d_\mu(t)$ is the integral curve of X_H on P with initial condition p_0.*

Proof.

$$\begin{aligned} c'(t_0) &= \frac{d}{dt}\left[\Phi_{g(t)} d_\mu(t)\right]\big|_{t=t_0} \\ &= \frac{d}{dt}\left[\Phi_{g(t_0)}\Phi_{g(t_0)^{-1}g(t)} d_\mu(t)\right]\big|_{t=t_0} \\ &= T\Phi_{g(t_0)} \cdot \xi(t_0)_P(d_\mu(t_0)) + T\Phi_{g(t_0)} d'_\mu(t_0) \\ &= \left[\mathrm{Ad}_{g_0^{(t)}}\xi\right]_P (c(t_0)) + T\Phi_{g(t_0)} d'_\mu(t_0) \end{aligned}$$

– $\mathbb{T}$*O BE COMPLETED*
– $\mathbb{E}$*ND WITH A DISCUSSION AND REFERENCE TO MEMOIRS*

3.18 Dynamics of Coupled Planar Rigid Bodies

Here we apply reduction techniques to the study of coupled rigid bodies moving without friction in the plane. We begin with two bodies and then discuss the situation for three bodies. See Sreenath *et al.* [1988], Oh *et al.* [1989], Grossman, Krishnaprasad and Marsden[198♡] and Patrick[1989,1990] for further details and for the three dimensional case. Related references are Krishnaprasad[1985], Krishnaprasad and Marsden[1987] and Bloch, Krishnaprasad, Marsden and Sanchez[1992].

Summary of Results

Refer to Figure 2.13.1 and define the following quantities, for $i = 1, 2$:

d_i		distance from the hinge to the center of mass of body i
ω_i		angular velocity of body i
θ		joint angle from body 1 to body 2
$\lambda(\theta)$	$=$	$d_1 d_2 \cos\theta$
m_1		mass of body i
ε	$=$	$m_1 m_2/(m_1+m_2) =$ reduced mass
I_i		moment of inertia of body i about its center of mass
$\tilde{I}_1$	$=$	$I_1 + \varepsilon d_1^2$; $\tilde{I}_2 = I_2 + \varepsilon d_2^2 =$ augmented moments of inertia
γ	$=$	$\varepsilon\lambda'/(\tilde{I}_1\tilde{I}_2 - \varepsilon^2\lambda^2)$, $' = d/d\theta$.

As we shall see, the dynamics of the system is given by the Euler-Lagrange equations for θ, ω_1, and ω_2 :

$$\begin{aligned} \dot\theta &= \omega_2 - \omega_1, \\ \dot\omega_1 &= -\gamma(\tilde{I}_2\omega_2^2 + \varepsilon\lambda\omega_1^2) \\ \dot\omega_2 &= \gamma(\tilde{I}_1\omega_1^2 + \varepsilon\lambda\omega_2^2) \end{aligned} \tag{3.18.1}$$

For the Hamiltonian structure it is convenient to introduce the momenta

$$\mu_1 = \tilde{I}_1\omega_1 + \varepsilon\lambda\omega_2, \quad \mu_2 = \tilde{I}_2\omega_2 + \varepsilon\lambda\omega_1, \tag{3.18.2}$$

that is,

$$\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} = \mathbf{J}\begin{pmatrix} \omega_1 \\ \omega_2 \end{pmatrix}, \quad \text{where} \quad \mathbf{J} = \begin{pmatrix} \tilde{I}_1 & \varepsilon\lambda \\ \varepsilon\lambda & \tilde{I}_2 \end{pmatrix} \tag{3.18.3}$$

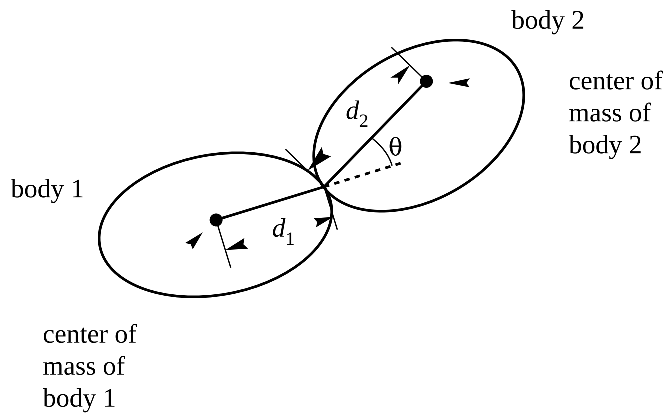

Figure 3.18.1:

(this will be done via the Legendre transform). The evolution equations for μ_i are obtained by solving (2.13.3) for ω_1, ω_2 and substituting into (2.13.1). The Hamiltonian is

$$H = \frac{1}{2}(\omega_1, \omega_2)\mathbf{J}\begin{pmatrix} \omega_1 \\ \omega_2 \end{pmatrix} \tag{3.18.4}$$

that is,

$$H = \frac{1}{2}(\mu_1, \mu_2)\mathbf{J}^{-1}\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \tag{3.18.5}$$

which is the total kinetic energy for the two bodies. The Poisson structure on the (θ, μ_1, μ_2)-space is

$$\{F, H\} = \{F, H\}_2 - \{F, H\}_1, \tag{3.18.6}$$

where

$$\{F, H\}_i = \frac{\partial F}{\partial \theta}\frac{\partial H}{\partial \mu_i} - \frac{\partial H}{\partial \theta}\frac{\partial F}{\partial \mu_i}.$$

The evolution equations (2.13.1) are equivalent to Hamilton's equations $\dot{F} = \{F, H\}$. Casimirs for the bracket (2.13.6) are readily checked to be

$$C = \Phi(\mu_1 + \mu_2) \tag{3.18.7}$$

for Φ any smooth function of one variable; that is, $\{F, C\} = 0$ for any F. One can also verify directly from (2.13.1) that, correspondingly, $(d\mu/dt) = 0$, where $\mu = \mu_1 + \mu_2$ is the total system angular momentum.

The symplectic leaves of (2.13.6) are described by the variables $\nu = (\mu_2 - \mu_1)/2, \theta$ which parametrize a cylinder. The bracket in terms of (θ, ν) is the canonical one on T^*S^1 :

$$\{F, H\}_i = \frac{\partial F}{\partial \theta}\frac{\partial H}{\partial \nu} - \frac{\partial H}{\partial \theta}\frac{\partial F}{\partial \nu}. \tag{3.18.8}$$

This canonical structure on T^*S^1 is consistent with the cotangent bundle reduction theorem (♡♡♡); there is no magnetic term (a *two* form) because the base S^1 is *one* dimensional. XREF

Kinematics for two coupled planar rigid bodies

In this section we set up the phase space for the dynamics of our problem. Refer to Figure 2.13.2 and define the following quantities:

$\mathbf{d}_{12}$		the vector from the center of mass of body 1 to the hinge point in a fixed reference configuration
$\mathbf{d}_{21}$		the vector from the center of mass of body 2 to the hinge point in a fixed reference configuration
$R(\theta_i)$	$=$	$\begin{pmatrix} \cos\theta_i & -\sin\theta_i \\ \sin\theta_i & \cos\theta_i \end{pmatrix}$ the rotation through angle θ_i giving the current orientation of body i (written as a matrix relative to the fixed standard inertial frame)
$\mathbf{r}_i$		current position of the center of mass of body i
$\mathbf{r}$		current position of the system center of mass
$\mathbf{r}_i^0$		the vector from the system center of mass to the center of mass of body i
θ	$=$	$\theta_2 - \theta_1$ joint angle
$R(\theta)$	$=$	$R(\theta_2) \cdot R(-\theta_1)$ joint rotation

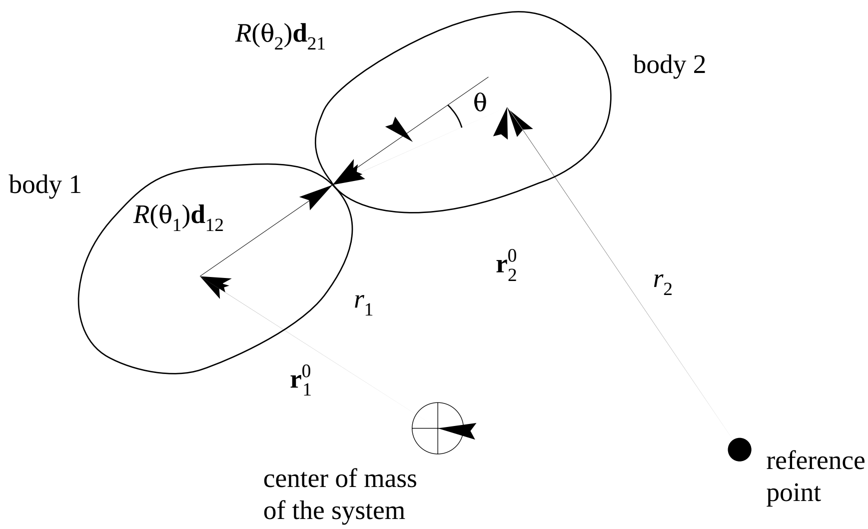

Figure 3.18.2:

The configuration space we start with is Q, the subset of $SE(2)\times SE(2)$ (two copies of the special Euclidean group of the plane) consisting of pairs $(R(\theta_1), \mathbf{r}_1), (R(\theta_2), \mathbf{r}_2))$ satisfying the ***hinge constraint***

$$\mathbf{r}_2 = \mathbf{r}_1 + R(\theta_1)\mathbf{d}_{12} - R(\theta_2)\mathbf{d}_{21}. \tag{3.18.9}$$

Notice that Q is of dimension 4 and is parametrized by θ_1, θ_2 and, say $\mathbf{r}_1$; that is, $Q \approx S^1 \times S^1 \times \mathbb{R}^2$. We form the velocity phase space TQ and momentum phase space T^*Q.

The Lagrangian on TQ is the kinetic energy (relative to the inertial frame) given by summing the kinetic energies of each body. To spell this out, let $\mathbf{X}_i$ denote a position vector in body 1 relative to the center of mass of body 1, and let $\rho_1(\mathbf{X}_1)$ denote the mass density of body 1. Then the current position of the point with material label $\mathbf{X}_1$ is

$$\mathbf{x}_1 = R(\theta_1)\mathbf{X}_1 + \mathbf{r}_1. \tag{3.18.10}$$

Thus $\dot{\mathbf{x}}_1 = \dot{R}(\theta_1)\mathbf{X}_1 + \dot{\mathbf{r}}_1$, and so the kinetic energy of body 1 is

$$\begin{aligned} K_1 &= \frac{1}{2}\int \rho_1(\mathbf{X}_1)\|\dot{\mathbf{x}}_1\|^2 d^2\mathbf{X}_1 \\ &= \frac{1}{2}\int \rho_1(\mathbf{X}_1)\left\langle \dot{R}\mathbf{X}_1 + \dot{\mathbf{r}}_1, \dot{R}\mathbf{X}_1 + \dot{\mathbf{r}}_1\right\rangle d^2\mathbf{X}_1 \\ &= \frac{1}{2}\int \rho_1(\mathbf{X}_1)\left[\left\langle \dot{R}\mathbf{X}_1, \dot{R}\mathbf{X}_1\right\rangle + 2\left\langle \dot{R}\mathbf{X}_1, \dot{\mathbf{r}}_1\right\rangle + \|\dot{\mathbf{r}}_1\|\right] d^2\mathbf{X}_1. \end{aligned} \tag{3.18.11}$$

But

$$\left\langle \dot{R}\mathbf{X}_1, \dot{R}\mathbf{X}_1\right\rangle = \operatorname{tr}(\dot{R}\mathbf{X}_1, \dot{R}\mathbf{X}_1)^T = \operatorname{tr}(\dot{R}\mathbf{X}_1^T\mathbf{X}_1\dot{R}^T) \tag{3.18.12}$$

and

$$\int \rho_1(\mathbf{X}_1)\left\langle \dot{R}\mathbf{X}_1, \dot{\mathbf{r}}_1\right\rangle d^2\mathbf{X}_1 = \left\langle \dot{R}\int \rho_1(\mathbf{X}_1)\mathbf{X}_1 d^2\mathbf{X}_1, \dot{\mathbf{r}}_1\right\rangle = 0 \tag{3.18.13}$$

since $\mathbf{X}_1$ is the vector relative to the center of mass of body 1. Substituting (2.13.12) and (2.13.13) into (2.13.11) and defining the matrix

$$\mathbf{I}^1 = \int \rho(\mathbf{X}_1)\mathbf{X}_1\mathbf{X}_1^T d^2\mathbf{X}_1 \tag{3.18.14}$$

we get

$$K_1 = \frac{1}{2}\operatorname{tr}(\dot{R}(\theta_1)\mathbf{I}^1(\dot{R}(\theta_1)^T) + \frac{1}{2}m_1\|\dot{\mathbf{r}}_1\|^2; \tag{3.18.15}$$

with a similar expression for $\mathbf{K}_2$. Now let

$$L : TQ \to \mathbb{R} \quad \text{be defined by} \quad L = K_1 + K_2. \tag{3.18.16}$$

The equations of motion then are the Euler-Lagrange equations for this L on TQ. Equivalently, they are Hamilton's equations for the corresponding Hamiltonian.

For later convenience, we shall rewrite the energy (2.13.16) in terms of $\omega_1 = \dot{\theta}_1, \omega_2 = \dot{\theta}_2, \mathbf{r}_1^0$ and $\mathbf{r}_2^0$. To do this note that, by definition,

$$m\mathbf{r} = m_1\mathbf{r}_1 + m_2\mathbf{r}_2, \tag{3.18.17}$$

where $m = m_1 + m_2$, and so, using $\mathbf{r}_1 = \mathbf{r} + \mathbf{r}_2^0$,

$$0 = m_1\mathbf{r}_1^0 + m_2\mathbf{r}_2^0 \tag{3.18.18}$$

and, subtracting $\mathbf{r}$ from both sides of (2.13.9),

$$\mathbf{r}_2^0 = \mathbf{r}_1^0 + R(\theta_1)\mathbf{d}_{12} - R(\theta_2)\mathbf{d}_{21}. \tag{3.18.19}$$

From (2.13.18) and (2.13.19) we find that

$$\mathbf{r}_2^0 = \frac{m_1}{m}(R(\theta_1)\mathbf{d}_{12} - R(\theta_2)\mathbf{d}_{21}) \tag{3.18.20}$$

and

$$\mathbf{r}_1^0 = -\frac{m_2}{m}(R(\theta_1)\mathbf{d}_{12} - R(\theta_2)\mathbf{d}_{21}). \tag{3.18.21}$$

Now we substitute

$$\mathbf{r}_1 = \mathbf{r} + \mathbf{r}_1^0 \quad \text{so} \quad \dot{\mathbf{r}}_1 = \dot{\mathbf{r}} + \dot{\mathbf{r}}_1^0 \tag{3.18.22}$$

and

$$\mathbf{r}_2 = \mathbf{r} + \mathbf{r}_2^0 \quad \text{so} \quad \dot{\mathbf{r}}_2 = \dot{\mathbf{r}} + \dot{\mathbf{r}}_2^0 \tag{3.18.23}$$

into (2.13.16) to give

$$\begin{aligned} L &= \frac{1}{2}\mathrm{tr}(\dot{R}(\theta_1)\mathbf{I}^1\dot{R}(\theta_1)^T + \dot{R}(\theta_2)\mathbf{I}^2\dot{R}(\theta_2)^T) \\ &\quad + \frac{1}{2}[m_1\|\dot{\mathbf{r}} + \dot{\mathbf{r}}_1^0\|^2 + m_2\|\dot{\mathbf{r}} + \dot{\mathbf{r}}_2^0\|^2]. \end{aligned} \tag{3.18.24}$$

But $m_1\left\langle \dot{\mathbf{r}}, \dot{\mathbf{r}}_1^0 \right\rangle + m_2\left\langle \dot{\mathbf{r}}, \dot{\mathbf{r}}_2^0 \right\rangle = 0$ since $m_1\dot{\mathbf{r}}_1^0 + m_2\dot{\mathbf{r}}_2^0 = 0$ from (2.13.18). Thus (2.13.22) simplifies to

$$\begin{aligned} L &= \frac{1}{2}\mathrm{tr}(\dot{R}(\theta_1)\mathbf{I}^1\dot{R}(\theta_1)^T + \dot{R}(\theta_2)\mathbf{I}^2\dot{R}(\theta_2)^T) \\ &\quad + \left(\frac{p^2}{2m}\right) + \frac{1}{2}m_1\|\dot{\mathbf{r}}_1^0\|^2 + \frac{1}{2}m_2\|\dot{\mathbf{r}}_2^0\|^2, \end{aligned} \tag{3.18.25}$$

where $p = m\|\dot{\mathbf{r}}\|$ is the magnitude of the system momentum.

Now write

$$\begin{aligned} \dot{R}(\theta_1) &= \frac{d}{dt}\begin{pmatrix} \cos\theta_1 & -\sin\theta_1 \\ \sin\theta_1 & \cos\theta_1 \end{pmatrix} \\ &= \begin{pmatrix} -\sin\theta_1 & -\cos\theta_1 \\ \cos\theta_1 & -\sin\theta_1 \end{pmatrix}\omega_1 \\ &=: R(\theta_1)\begin{pmatrix} 0 & -\omega_1 \\ \omega_1 & 0 \end{pmatrix} \\ &:= R(\theta_1)\hat{\omega}_1, \end{aligned} \tag{3.18.26}$$

so that (2.13.20) gives

$$\dot{\mathbf{r}}_2^0 = \frac{m_1}{m}(R(\theta_1)\hat{\omega}_1\mathbf{d}_{12} - (R(\theta_2)\hat{\omega}_2\mathbf{d}_{21}), \quad \dot{\mathbf{r}}_1^0 = -\frac{m_2}{m}(R(\theta_1)\hat{\omega}_1\mathbf{d}_{12} - (R(\theta_2)\hat{\omega}_2\mathbf{d}_{21}). \tag{3.18.27}$$

Substituting (2.13.25) and (2.13.24) into (2.13.23) gives

$$L = \frac{1}{2}\mathrm{tr}((\hat{\omega}_1\mathbf{I}^1\hat{\omega}_1^T) + \hat{\omega}_1\mathbf{I}^2\hat{\omega}_2^T)) + \frac{p^2}{2m} + \frac{m_1m_2}{m}\|\hat{\omega}_1\mathbf{d}_{12} - R(\theta_2 - \theta_1)\hat{\omega}_2\mathbf{d}_{21}\|^2. \tag{3.18.28}$$

Finally we note that

$$\mathrm{tr}(\hat{\omega}_1\mathbf{I}^1\hat{\omega}_1^T) = \mathrm{tr}(\hat{\omega}_1^T\hat{\omega}_1\mathbf{I}^1) = \mathrm{tr}\left(\begin{pmatrix} \omega_1^2 & 0 \\ 0 & \omega_1^2 \end{pmatrix}\mathbf{I}^1\right) = \omega_1^2\mathrm{tr}\,\mathbf{I}^1 := \omega_1^2 I_1, \tag{3.18.29}$$

where

$$I_1 = \int \rho(X_1, Y_1)(X_1^2 + Y_1^2)dX_1dY_1$$

is the moment of inertia of body 1 about its center of mass. One similarly derives an expression – call it (2.13.27′) – where 1 is replaced by 2 throughout. The final term in (2.13.26) is manipulated as follows: for future reference

$$\begin{aligned} &\|\hat{\omega}_1\mathbf{d}_{12} = -R(\theta)\hat{\omega}_2\mathbf{d}_{21}\|^2 \\ &\quad= \quad \|\hat{\omega}_1\mathbf{d}_{12}\|^2 - 2\langle\hat{\omega}_1\mathbf{d}_{12}, R(\theta)\hat{\omega}_2\mathbf{d}_{21}\rangle + \|\hat{\omega}_2\mathbf{d}_{21}\|^2 \\ &\quad= \quad \omega_1^2 d_1^2 + \omega_2^2 d_2^2 - 2\langle\hat{\omega}_1\mathbf{d}_{12}, \hat{\omega}_2 R(\theta)\mathbf{d}_{21}\rangle \\ &\quad= \quad \omega_1^2 d_1^2 + \omega_2^2 d_2^2 - 2\omega_1\omega_2\langle\mathbf{d}_{12}, R(\theta)\mathbf{d}_{21}\rangle \end{aligned} \tag{3.18.30}$$

Substituting (2.13.27), (2.13.27′) and (2.13.28) into (2.13.26) gives

$$L = \frac{1}{2}[(\omega_1^2\tilde{I}_1 + \omega_2^2\tilde{I}_2 + 2\omega_1\omega_2\varepsilon\lambda(\theta)] + \frac{p^2}{2m}, \tag{3.18.31}$$

where

$$\lambda(\theta) = -\langle\mathbf{d}_{12}, R(\theta)\mathbf{d}_{21}\rangle = -[\mathbf{d}_{12}\cdot\mathbf{d}_{21}\cos\theta - (\mathbf{d}_{12}\times\mathbf{d}_{21})\cdot\mathbf{k}\sin\theta]. \tag{3.18.32}$$

Remarks

i If $\mathbf{d}_{12}$ and $\mathbf{d}_{21}$ are parallel (that is, the reference configuration is chosen with $\mathbf{d}_{12}$ and $\mathbf{d}_{21}$ aligned), then (2.13.30) gives $\lambda(\theta) = d_1d_2\cos\theta$, as in §**3.6A**. how do we reference th

ii The quantities $\tilde{I}_1, \tilde{I}_2$ are the moments of inertia of "augmented" bodies; for example $\tilde{I}_1$ is the moment of inertia of body 1 augmented by putting a mass ε at the hinge point.

Reduction to the center of mass frame

Now we reduce the dynamics by the action of the translation group $\mathbb{R}^2$. This group acts on the original configuration space Q by

$$\mathbf{v}\cdot((R(\theta_1), \mathbf{r}_1), (R(\theta_2), \mathbf{r}_2)) = ((R(\theta_1), \mathbf{r}_1 + \mathbf{v}), (R(\theta_2), \mathbf{r}_2 + \mathbf{v})). \tag{3.18.33}$$

This is well defined since the hinge constraint (2.13.9) is preserved by this action. The induced momentum map on TQ is calculated by the standard formula

$$J_\xi = \frac{\partial L}{\partial \dot{q}_i}\xi_Q^i(q), \tag{3.18.34}$$

or on T^*Q by

$$J_\xi = p_i\xi_Q^i(q), \tag{3.18.35}$$

where ξ_Q^i is the infinitesimal generator of the action on Q (see §2.2). To implement (2.13.32) XREF
we parametrize Q by θ_1, θ_2 and $\mathbf{r}$, with $\mathbf{r}_1$ and $\mathbf{r}_2$ determined by (2.13.20) and (2.13.21). From (2.13.23) we see that the momentum conjugate to $\mathbf{r}$ is

$$\mathbf{p} = \frac{\partial L}{\partial \dot{\mathbf{r}}} = m\dot{\mathbf{r}} \tag{3.18.36}$$

and so (2.13.32) gives

$$J_\xi = \langle \mathbf{p}, \xi \rangle, \quad \xi \in \mathbb{R}^2. \tag{3.18.37}$$

Thus $J = \mathbf{p}$ is conserved since H is cyclic in $\mathbf{r}$ and so H is translation invariant. The corresponding reduced space is obtained by fixing $\mathbf{p} = \mathbf{p}_0$ and letting

$$P_{\mathbf{p}_0} = \mathbf{J}^{-1}(\mathbf{p}_0)/\mathbb{R}^2.$$

But $P_{\mathbf{p}_0}$ is isomorphic to $T^*(S^1 \times S^1)$, that is, to the space of θ_1, θ_2 and their conjugate momenta. The reduced Hamiltonian is simply the Hamiltonian corresponding to (2.13.29) with $\mathbf{p}_0$ regarded as a constant.

In this case the reduced symplectic manifold is a cotangent bundle, in agreement with the Cotangent Bundle Reduction Theorem. The reduced phase space has the canonical symplectic form; one can also check this directly here.

In (2.13.29), we can adjust L by a constant and thus assume that $\mathbf{p}_0 = 0$; this obviously does not affect the equations of motion.

That the reduced system is given by geodesic flow on $S^1 \times S^1$ since (2.13.29) is quadratic in the velocities. Indeed the metric tensor is just the matrix $\mathbf{J}$ given by (1.6) so the conjugate momenta are μ_1 and μ_2 given by (1.6).

We remark, finally, that the reduction to center-of-mass coordinates here is somewhat simpler and more symmetric than the Jacobi-Haretu reduction to center-of-mass coordinates for n point masses. (Just taking the positions relative to the centre of mass does not achieve this since this does not reduce the dimension at all!) What is different here is that the two bodies are hinged, and so by (2.13.20), $\mathbf{r}_1^0$ and $\mathbf{r}_2^0$ are determined by the other data.

Reduction by rotations

To complete the reduction, we reduce by the diagonal action of S^1 on the configuration space $S^1 \times S^1$ that was obtained in §3.6C . The momentum map for this action is

$$J((\theta_1, \mu_1), (\theta_2, \mu_2)) = \mu_1 + \mu_2. \tag{3.18.38}$$

To facilitate stability calculations, form the Poisson reduced space

$$P := T^*(S^1 \times S^1)/S^1 \tag{3.18.39}$$

whose symplectic leaves are the reduced symplectic manifolds

$$P_\mu = J^{-1}(\mu)/S^1 \subset P.$$

We coordinatize P by $\theta = \theta_2 - \theta_1, \mu_1$ and μ_2; topologically, $P = S^1 \times \mathbb{R}^2$. The Poisson structure on P is computed as follows: take two functions $F(\theta, \mu_1, \mu_2)$ and $H(\theta, \mu_1, \mu_2)$. Regard them as functions of $\theta_1, \theta_2, \mu_1, \mu_2$ by substituting $\theta = \theta_2 - \theta_1$ and compute the canonical bracket. The asserted bracket (2.13.6) of is what results. The Casimirs on P are obtained by composing J with Casimirs on the dual of the Lie algebra of S^1; that is, with arbitrary functions of one variable; thus (2.13.7) results. This can be checked directly.

If we parametrize P_μ by θ and $\nu = (\mu_2 - \mu_1)/2$, then the Poisson bracket on P_μ becomes the canonical one. This is consistent with the Cotangent Bundle Reduction Theorem which asserts in this case that the reduction of $T^*(S^1 \times S^1)$ by S^1 is symplectically diffeomorphic to $T^*((S^1 \times S^1)/S^1) \cong T^*S^1$. There are no "magnetic" terms since the reduced configuration space S^1 is one-dimensional, and hence carries no non-zero two-forms.

The realization of P_μ as T^*S^1 is not unique. For example we can parametrize P_μ by (θ_2, μ_2) or by (θ_1, μ_1), each of which also gives the canonical bracket. (In the general theory there can be more than one one-form α_μ (*i.e.,* connection) by which one embeds P_μ into T^*S^1, as well as more than one way to identify $(S^1 \times S^1)/S^1 \cong S^1$. The three listed above correspond to three such choices of α_μ.)

The reduced bracket on $T^*(S^1 \times S^1)/S^1$ can also be obtained from the general formula for the bracket on $(P \times T^*G)/G \cong P \times \mathbf{3}^*$ (see Krishnaprasad and Marsden [1987] and §**3.6G**; it produces one of the variants above, depending on whether we take G to be parametrized by θ_1 or θ_2, or $\theta_2 - \theta_1$.

The reduced Hamiltonian on P is (2.13.5) regarded as a function of μ_1, μ_2 and θ:

$$H = \frac{1}{2\Delta}(\mu_1, \mu_2) \begin{pmatrix} \tilde{I}_2 & -\varepsilon\lambda \\ -\varepsilon\lambda & \tilde{I}_1 \end{pmatrix} \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \tag{3.18.40}$$

where $\Delta = \tilde{I}_1 \tilde{I}_2 - \varepsilon^2 \lambda^2$. Substituting $\mu_1 = (\mu/2) - \nu$ and $\mu_2 = \nu + (\mu/2)$ gives

$$H = \frac{1}{2\Delta}(\tilde{I}_1 + \tilde{I}_2 + 2\varepsilon\lambda)\nu^2 + \frac{1}{2\Delta}\left[\left(\tilde{I}_1 - \tilde{I}_2\right)\mu\right]\nu + \frac{1}{2\Delta}\left(\frac{1}{4}\mu^2\left(\tilde{I}_1 + \tilde{I}_2 - 2\varepsilon\lambda\right)\right). \tag{3.18.41}$$

Thus, the Euler-Lagrange equations (2.13.1) are equivalent to $\dot{F} = \{F, H\}$ for the reduced bracket. We can also obtain a Hamiltonian system on the leaves, parametrized by say (θ, ν) with H given by (2.13.38).

The linear term in ν in (2.13.38) can be eliminated by completion of squares: it is not there in the general theory because reduced coordinates adapted to the metric of the kinetic energy are used; these are produced by the completion of squares. Notice that the Hamiltonian now is of the form kinetic plus potential energy but that the metric now on S^1 is θ-dependent and, unless d_1 or d_2 vanishes, it is a non-trivial dependence. The potential piece is the ***amended potential***.

We summarize as follows:

Theorem 3.18.1 *The reduced phase space for two coupled planar rigid bodies is the three-dimensional Poisson manifold $P = S^1 \times \mathbb{R}$ with the bracket (2.13.6); its symplectic leaves are the cylinders with canonical variables (θ, ν). Casimirs are given by (2.13.7).*

The reduced dynamics are given by $\dot{F} = \{F, H\}$ or equivalently,

$$\dot{\theta} = \frac{\partial H}{\partial \mu_2} - \frac{\partial H}{\partial \mu_1}, \quad \dot{\mu}_1 = \frac{\partial H}{\partial \theta}, \quad \dot{\mu}_2 = -\frac{\partial H}{\partial \theta}, \tag{3.18.42}$$

where H is given by (2.13.5). The equivalent dynamics on the leaves is given by

$$\frac{\partial \theta}{\partial t} = \frac{\partial H}{\partial \nu}, \quad \frac{\partial \nu}{\partial t} = -\frac{\partial H}{\partial \theta}, \tag{3.18.43}$$

where H is given by (2.13.38).

Equilibria and stability by the energy-Casimir method

We now use Arnold's energy-Casimir method, (see the introduction and overview) to determine the equilibrium points and their stability. An equivalent alternative to this method is to look for critical points of H given by (2.13.38) in (θ, ν)-space and test $\delta^2 H$ for definiteness at these equilibria.

To search for equilibria we look directly at Hamilton's equations on P. From (2.13.39), the conditions $\dot{\mu}_1 = \dot{\mu}_2 = 0$ become

$$\frac{\partial H}{\partial \theta} = 0; \tag{3.18.44}$$

that is,

$$-\frac{1}{2}(\mu_1, \mu_2)\mathbf{J}^{-1}\frac{\partial \mathbf{J}}{\partial \theta}\mathbf{J}^{-1}\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} = 0. \tag{3.18.45}$$

Clearly,

$$\frac{d\mathbf{J}}{d\theta} = \begin{pmatrix} 0 & \varepsilon\lambda' \\ \varepsilon\lambda' & 0 \end{pmatrix} \tag{3.18.46}$$

from (2.13.3), so (2.13.42) becomes

$$-\frac{1}{2}(\omega_1, \omega_2)\begin{pmatrix} 0 & \varepsilon\lambda' \\ \varepsilon\lambda' & 0 \end{pmatrix}\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} = 0; \tag{3.18.47}$$

that is,

$$-\varepsilon\lambda'\omega_1\omega_2 = 0. \tag{3.18.48}$$

The equilibrium condition $\dot{\theta} = 0$ becomes $\tilde{I}_1\mu_1 - \varepsilon\lambda\mu_2 = \tilde{I}_2\mu_2 - \varepsilon\lambda\mu_1$ or, equivalently, $\omega_1 = \omega_2$. Thus, the equilibria are given by

i $\omega_1 = \omega_2 = 0$, or

ii $\omega_1 = \omega_2 \neq 0, \lambda' = 0$.

For simplicity, choose the reference configuration so that $\mathbf{d}_{12}$ and $\mathbf{d}_{21}$ are parallel. Then

$$\lambda'(\theta) = \mathbf{d}_{12} \cdot \mathbf{d}_{21} \sin\theta$$

so the equilibria in case **ii** occur when

ii′ either (a) $\mathbf{d}_{12} = 0$ or $\mathbf{d}_{21} = 0$, or (b) $\theta = 0$ or π. The case in which $\theta = \pi$ corresponds to the case of folded bodies, while $\theta = 0$ corresponds to extended (stretched out) bodies.

The first step in the energy-Casimir method is to realize the equilibria as critical points of $H + C$. We calculate that

$$\begin{aligned} \frac{\partial H}{\partial \theta} &= \varepsilon\lambda'\omega_1\omega_2 \\ \frac{\partial H}{\partial \mu_1} = \omega_1; &\quad \frac{\partial H}{\partial \mu_2} - \omega_2, \end{aligned} \tag{3.18.49}$$

where

$$\begin{pmatrix} \omega_1 \\ \omega_2 \end{pmatrix} = \mathbf{J}^{-1}\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} = \frac{1}{\Delta}\begin{pmatrix} \tilde{I}_2\mu_1 - \varepsilon\lambda\mu_2 \\ \tilde{I}_1\mu_2 - \varepsilon\lambda\mu_1 \end{pmatrix}.$$

The first variation is

$$\mathbf{d}(H + C) = \frac{\partial H}{\partial \theta}d\theta + \left(\frac{\partial H}{\partial \mu_1} + \Phi'\right)d\mu_1 + \left(\frac{\partial H}{\partial \mu_2} + \Phi'\right)d\mu_2, \tag{3.18.50}$$

from which we see that critical points of $H + C$ correspond to equilibria provided

$$\Phi'(\mu_e) = -\left(\frac{\partial H}{\partial \mu_1}\right)_e = -\left(\frac{\partial H}{\partial \mu_1}\right)_e, \tag{3.18.51}$$

where the subscript e means evaluation at the equilibrium. As in other examples such as the rigid body and heavy top from the introduction, $\Phi''(\mu_e)$ is arbitrary.

The matrix of the second variation of $H + C$ at equilibrium is

$$\delta^2(H+C) = \begin{pmatrix} \dfrac{\partial^2 H}{\partial \theta^2} & \dfrac{\partial^2 H}{\partial \theta \partial \mu_1} & \dfrac{\partial^2 H}{\partial \theta \partial \mu_2} \\ \dfrac{\partial^2 H}{\partial \theta \partial \mu_1} & \dfrac{\partial^2 H}{\partial \mu_1^2} + \Phi'' & \dfrac{\partial^2 H}{\partial \mu_1 \partial \mu_2} + \Phi'' \\ \dfrac{\partial^2 H}{\partial \theta \partial \mu_2} & \dfrac{\partial^2 H}{\partial \mu_1 \partial \mu_2} + \Phi'' & \dfrac{\partial^2 H}{\partial \mu_2^2} + \Phi'' \end{pmatrix} \tag{3.18.52}$$

where

$$\begin{pmatrix} \dfrac{\partial^2 H}{\partial \mu_1^2} & \dfrac{\partial^2 H}{\partial \mu_1 \partial \mu_2} \\ \dfrac{\partial^2 H}{\partial \mu_1 \partial \mu_2} & \dfrac{\partial^2 H}{\partial \mu_2^2} \end{pmatrix} = \mathbf{J}^{-1} = \frac{1}{\Delta} \begin{pmatrix} \tilde{I}_2 & -\varepsilon\lambda \\ -\varepsilon\lambda & \tilde{I}_1 \end{pmatrix},$$

$$\frac{\partial^2 H}{\partial \theta \partial \mu_1} = -\frac{\varepsilon\lambda'}{\Delta^2}(\tilde{I}_2\omega_2 - \varepsilon\lambda\omega_1), \quad \frac{\partial^2 H}{\partial \theta \partial \mu_2} = -\frac{\varepsilon\lambda'}{\Delta^2}(-\varepsilon\lambda\omega_2 + \tilde{I}_1\omega_1),$$

and

$$\frac{\partial^2 H}{\partial \theta^2} = \frac{\partial}{\partial \theta}\left[-\varepsilon\lambda' \frac{\partial H}{\partial \mu_1}\frac{\partial H}{\partial \mu_2}\right] = -\varepsilon\lambda''\omega_1\omega_2 - \varepsilon\lambda' \frac{\partial^2 H}{\partial \theta \partial \mu_1}\omega_2 - \varepsilon\lambda'\omega_1 \frac{\partial^2 H}{\partial \theta \partial \mu_2}.$$

At equilibrium, $\lambda = \pm d_1 d_2 (+ \text{ if } \theta = 0, - \text{ if } \theta = \pi)$ and so

$$\mathbf{J}^{-1} \frac{1}{(\tilde{I}_1\tilde{I}_2 - \varepsilon^2 d_1^2 d_2^2)} \begin{bmatrix} \tilde{I}_2 & \mp\varepsilon d_1 d_2 \\ \mp\varepsilon d_1 d_2 & \tilde{I}_1 \end{bmatrix},$$

$$\frac{\partial^2 H}{\partial \theta \partial \mu_1} = 0 = \frac{\partial^2}{\partial \theta \partial \mu_2}, \quad \text{and} \quad \frac{\partial^2 H}{\partial \theta^2} = -\varepsilon\lambda''\omega_e^2 = \pm\varepsilon d_1 d_2 \omega_e^2,$$

where $\omega_e = \omega_1 = \omega_2 \neq 0$ at equilibrium. Thus (2.13.48) becomes

$$\delta^2(H+C) = \begin{pmatrix} \pm\varepsilon\lambda d_1 d_2 \omega_e^2 & 0 \\ 0 & \mathbf{J}^{-1} + \Phi'' \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \end{pmatrix}. \tag{3.18.53}$$

This matrix is clearly positive definite if $d_1 \neq 0, d_2 \neq 0$ if $\theta = 0(+ \text{ sign})$ and $\Phi''(\mu_e) \geq 0$ and is indefinite for any choice of $\Phi''(\mu_e)$ if $\theta = \pi$.

Another way to do the stability analysis is to use the reduced Hamiltonian on T^*S^1. After completing squares, H will have the form of kinetic plus potential energy with effective potential given by

$$V(\theta) = \frac{1}{2\Delta}\left[\frac{1}{4}\mu^2(\tilde{I}_1 + \tilde{I}_2 - 2\varepsilon\lambda) + \frac{(\tilde{I}_1 + \tilde{I}_2)^2\mu^2}{4(\tilde{I}_1 + \tilde{I}_2 + 2\varepsilon\lambda)}\right]. \tag{3.18.54}$$

Minima of V are then the stable equilibria while maxima are unstable. The following theorem summarizes the situation.

Theorem 3.18.2 *The dynamics of the* 2*-body problem is completely integrable and contains one stable relative equilibrium solution* ($\theta = 0$ *- the stretched-out case) and one unstable relative equilibrium solution* ($\theta = \pi$ *- the folded-over case). The reduced dynamics contains a homoclinic orbit joining the unstable equilibrium to itself.*

For three or more bodies, this method of looking for minima of the potential will not work in a naive way because the symplectic structures on the symplectic leaves will have magnetic terms. The general theory for dealing with this situation is given in Simo, Lewis and Marsden [1991].

Multibody problems

The Hamiltonian formulation of the previous sections extends to systems of N planar rigid bodies connected to form a ***tree structure***. Since the general statement of this result requires significant additional notation we limit ourselves to the special case of a ***chain*** of N bodies.

Theorem 3.18.3 *The total kinetic energy (Hamiltonian) for an open chain of* N *planar rigid bodies connected together by hinge joints has the form*

$$H = \mu^T \cdot \mathbf{J}^{-1} \cdot \mu \tag{3.18.55}$$

where $\mu = (\mu_1, \mu_2, \dots, \mu_N)^T$ *is the momentum vector and* $\mathbf{J}$ *is the corresponding* $N \times N$ *inertia matrix which is a function of the set of relative (or joint) angles between adjacent bodies. The reduced dynamics takes the form*

$$\begin{aligned}
\dot{\mu}_1 &= \frac{\partial H}{\partial \theta_{2,1}} \\
\dot{\mu}_2 &= \frac{\partial H}{\partial \theta_{3,2}} - \frac{\partial H}{\partial \theta_{2,1}} \\
\dot{\mu}_i &= \frac{\partial H}{\partial \theta_{i+1,i}} - \frac{\partial H}{\partial \theta_{i,i-1}} \\
\dot{\mu}_N &= -\frac{\partial H}{\partial \theta_{N,N-1}} \\
\dot{\theta}_{i+1,i} &= \frac{\partial H}{\partial \theta_{i+1}} - \frac{\partial H}{\partial \mu_i} \quad \text{for} \quad i = 1, \dots, N-1)
\end{aligned} \tag{3.18.56}$$

where $\theta_{i+1,i}$ *is the joint angle between body* $i+1$ *and body* i.

The associated Poisson structure is given by

$$\{f, g\} = \sum_{i=0}^{N-1} \left(\frac{\partial f}{\partial \mu_i} - \frac{\partial f}{\partial \mu_{i+1}} \right) \frac{\partial g}{\partial \theta_{i+1,i}} - \frac{\partial f}{\partial \theta_{i+1,i}} \left(\frac{\partial g}{\partial \mu_i} - \frac{\partial g}{\partial \mu_{i+1}} \right). \tag{3.18.57}$$

This is proven in a way similar to the two-body case. The structure of equilibria and the associated stability analysis become quite complex and interesting as the number of interconnected bodies increases. A mixture of topological and geometric methods may be necessary to extract useful information on the phase portraits.

In the remainder of this section, we illustrate some of the complexities of multibody problems by discussing of the equilibria and stability for a system of three planar rigid bodies connected by hinge joints (see Figure 2.13.3).

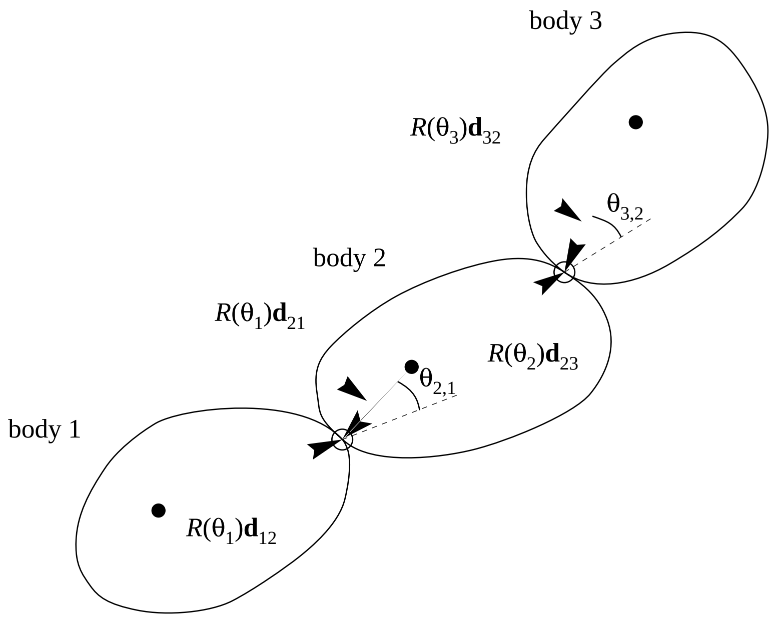

Figure 3.18.3: Planar three-body system

The Hamiltonian of the planar three-body problem is given by equation (2.13.51) with the momentum vector μ and the coefficient of inertia matrix $\mathbf{J}$ being defined as follows:

$$\mu = (\mu_1, \mu_2, \mu_3)^T,$$

$$\mathbf{J} = \begin{pmatrix} \tilde{I}_1 & \tilde{\lambda}_{12}(\theta_{2,1}) & \tilde{\lambda}_{31}(\theta_{2,1} + \theta_{3,2}) \\ \tilde{\lambda}_{12}(\theta_{2,1}) & \tilde{I}_1 & \tilde{\lambda}_{23}(\theta_{3,2}) \\ \tilde{\lambda}_{31}(\theta_{2,1} + \theta_{3,2}) & \tilde{\lambda}_{23}(\theta_{3,2}) & \tilde{I}_3 \end{pmatrix} \qquad (3.18.58)$$

Here $\theta_{2,1}$ and $\theta_{3,2}$ are the relative angles between bodies 2 and 1, and bodies 3 and 2, respectively. The coefficients of inertia $\tilde{I}_i$ and $\tilde{\lambda}_{ij}$ are given by

$$\begin{aligned}
\tilde{I}_1 &= [I_1 + (\varepsilon_{12}\varepsilon_{31}) \langle \mathbf{d}_{12}, \mathbf{d}_{12} \rangle], \\
\tilde{I}_2 &= [I_2 + \varepsilon_{12} \langle \mathbf{d}_{21}, \mathbf{d}_{21} \rangle + \varepsilon_{23} \langle \mathbf{d}_{23} - \mathbf{d}_{23} \rangle \\
&\quad + \varepsilon_{31} \langle (\mathbf{d}_{23} - \mathbf{d}_{21}), (\mathbf{d}_{23} - \mathbf{d}_{21}) \rangle] \\
\tilde{I}_3 &= [I_3 + (\varepsilon_{23} + \varepsilon_{31}) \langle \mathbf{d}_{32}, \mathbf{d}_{32} \rangle] \\
\tilde{\lambda}_{12}(\theta_{2,1}) &= [\varepsilon_{12} \lambda_{(-\mathbf{d}_{21}, \mathbf{d}_{12})}(\theta_{2,1}) + \varepsilon_{31} \lambda_{(\mathbf{d}_{23}, -\mathbf{d}_{21}, \mathbf{d}_{12})})(\theta_{2,1})] \\
\tilde{\lambda}_{23}(\theta_{3,2}) &= [\varepsilon_{23} \lambda_{(-\mathbf{d}_{32}, \mathbf{d}_{23})}(\theta_{3,2}) + \varepsilon_{31} \lambda_{(-\mathbf{d}_{32}, \mathbf{d}_{23}, \mathbf{d}_{21})})(\theta_{3,2})] \\
\tilde{\lambda}_{31}(\theta_{2,1}) + \theta_{3,2}) &= \varepsilon_{31} \lambda_{(\mathbf{d}_{32}, \mathbf{d}_{12})}(\theta_{2,1} + \theta_{3,2}) \\
\varepsilon_{ij} &= \frac{m_i m_j}{m_1 + m_2 + m_3}, \quad i \neq j \quad \text{and} \quad i, j = 1, 2, 3 \\
\lambda_{(x,y)}(\alpha) &= \mathbf{x} \cdot \mathbf{y} \cos \alpha + [\mathbf{x} \times \mathbf{y}] \sin \alpha,
\end{aligned}$$

where the m_i and I_i are the mass and inertia respectively of the body i, and the $\mathbf{d}_{ij}$ are defined as in Figure 2.13.3.

The dynamics of a three-body system of planar, rigid bodies in the Hamiltonian setting is given by:

$$\begin{aligned}
\dot{\mu}_1 &= \frac{\partial H}{\partial \theta_{2,1}} \\
\dot{\mu}_2 &= -\frac{\partial H}{\partial \theta_{2,1}} + \frac{\partial H}{\partial \theta_{3,2}} \\
\dot{\mu}_3 &= -\frac{\partial H}{\partial \theta_{3,2}} \\
\dot{\theta}_{2,1} &= \frac{\partial H}{\partial \mu_2} - \frac{\partial H}{\partial \mu_1} \\
\dot{\theta}_{3,2} &= \frac{\partial H}{\partial \mu_3} - \frac{\partial H}{\partial \mu_2}
\end{aligned} \tag{3.18.59}$$

From (2.13.55) note that the sum $\mu_1 + \mu_2 + \mu_3$ of momentum variables is constant in time.

In Sreenath *et al.* [1988] and Oh *et al.* [1988], it is shown that for three coupled rigid bodies there are either 4 or 6 relative equilibria and the bifurcations between these are determined as a function of the system parameters. It is also shown that near the stable stretched out relative equilibrium, there are relative periodic orbits distinuished by F] symmetry type. This is done using the Stewart et. al. [198♡] symmetric version of the Moser-Weinstein theorem (Weinstein [1973], and Moser [1978]). Also, it is shown that the dynamics is, in general, not integrable by using the Poincaré-Melnikov method.

Coupled Rotating Systems

Some general complements to reduction theory motivated by the dynamics of rotating system are given here. They are taken from Krishnaprasad and Marsden [1987]; this reference and Patrick [1990] should be consulted for further information and applications.

Assume G is a Lie group acting by canonical (Poisson) transfomations on a Poisson manifold P. Define $\phi : T^*G \times P \to \mathbf{3}^* \times P$ by

$$\varphi(\alpha_g, z) = (TL_g^* \cdot \alpha_g, g^{-1} \cdot z) \tag{3.18.60}$$

where $g^{-1} \cdot z$ denotes the action of g^{-1} on $z \in P$. For our example, $G = SO(3)$ and α_g is a momentum variable which is given in coordinates on $T^*SO(3)$ by the momentum variables p_ϕ, p_θ, p_ψ conjugate to the Euler angles ϕ, θ, ψ. The mapping φ in (2.13.56) transforms α_g to body representation and transforms $z \in P$ to $g^{-1} \cdot z$, which represents z relative to the body.

As usual, for $\xi \in \mathbf{3}$, we let ξ_P denote its infinitesimal generator on P, so ξ_P is the vector field on P given by

$$\xi_P(z) = \frac{d}{dt}(\exp(t\xi) \cdot z)\bigg|_{t=0}.$$

For $F, G : \mathbf{3}^* \times P \to \mathbb{R}$, let $\{F, G\}_-$ stand for the minus Lie-Poisson bracket holding the P variable fixed and let $\{F, G\}_P$ stand for the Poisson bracket on P with the variable $\mu \in \mathbf{3}^*$ held fixed.

Endow $\mathbf{3}^* \times P$ with the follwing bracket:

$$\{F, G\} = \{F, G\}_- + \{F, G\}_P - \mathbf{d}_z F \cdot \left(\frac{\delta G}{\delta \mu}\right)_P + \mathbf{d}_z G \cdot \left(\frac{\delta F}{\delta \mu}\right)_P \tag{3.18.61}$$

where $d_z F$ means the differential of F with respect to $z \in P$ and the evaluation point (μ, z) has been suppressed.

Proposition 3.18.4 *The bracket (2.13.57) makes $\mathbf{3}^* \times P$ into a Poisson manifold and $\varphi : T^*G \times P \to \mathbf{3}^* \times P$ 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, ϕ is G-invariant and induces a Poisson diffeomorphism of $(T^*G \times P)/G$ with $\mathbf{3}^* \times P$.*

Proof For $F, G : \mathbf{3}^* \times P \to \mathbb{R}$, let $\bar{F} = F \circ \phi$ and $\bar{G} = G \circ \phi$. We show that $\{\bar{F}, \bar{G}\}_{T^*G} + \{\bar{F}, \bar{G}\}_P = \{\bar{F}, \bar{G}\} \circ \varphi$. This will show φ is canonical. Since it is easy to check that ϕ is G-invariant and gives a diffeomorphism of $(T^*H \times P)/G$ with $\mathbf{3}^* \times P$, it follows that (2.13.57) represents the reduced bracket and so defines a Poisson structre.

To prove our claim, write $\varphi = \varphi_G \times \varphi_P$. Since φ_G does not depend on x and the group action is assumed canonical, $\{\bar{F}, \bar{G}\}_P = \{F, G\} \circ \varphi$. For the T^*G bracket, note that since ϕ_G is a Poisson map of T^*G to $\mathbf{3}^*$, the terms involving φ_G will be $\{F, G\}_{-} \circ \varphi$. The terms involving $\varphi_P(\alpha_\gamma, z) = g^{-1} \cdot z$ are found by noting that the bracket of a function K of g with a function L of α_g is

$$\mathbf{d}_g K \cdot \frac{\delta L}{\delta \alpha_g}$$

where $\delta L/\delta \alpha_g$ means the fiber derivative of L regarded as a vector at g. This is paired with the covector $\mathbf{d}_g K$. Letting $\Psi_z(g) = g^{-1} \cdot z$, we find by use of the chain rule that missing terms in the bracket are

$$\mathbf{d}_z F \cdot T\Psi_z \cdot \frac{\delta G}{\delta \mu} - \mathbf{d}_z G \cdot T\Psi_z \cdot \frac{\delta F}{\delta \mu}.$$

However, $T\Psi_z \cdot (\delta G/\delta \mu) = -(\delta F/\delta \mu)_P \circ \Psi_z$, so the preceeding expression reduces to the last two terms in equation (2.13.57). ■

Suppose the action of G on P has an Ad^*-equivariant momentum map $\mathbf{J} : P \to \mathbf{3}^*$. Consider the map $\alpha : \mathbf{3}^* \times P \to \mathbf{3}^* \times P$ given by

$$\alpha(\mu, z) = (\mu + \mathbf{J}(z), z). \tag{3.18.62}$$

Let the bracket $\{\ ,\ \}_0$ on $\mathbf{3}^* \times P$ be defined by

$$\{F, G\}_0 = \{F, G\}_{-} + \{F, G\}_P. \tag{3.18.63}$$

Thus $\{\ ,\ \}_0$ is (2.13.57) with the coupling or interaction terms dropped. We claim that the map α eliminates the coupling:

Proposition 3.18.5 $\alpha : (\mathbf{3}^* \times P, \{\ ,\ \}) \to (\mathbf{3}^* \times P, \{\ ,\ \}_0)$ *is a Poisson diffeomorphism.*

Proof For $F, G : \mathbf{3}^* \times P \to \mathbb{R}$, let $\hat{F} = F \circ \alpha$ and $\hat{G} = G \circ \alpha$. Letting $\nu = \mu + \mathbf{J}(z)$, and dropping evaluation points, we conclude that

$$\frac{\delta \hat{F}}{\delta \mu} = \frac{\delta F}{\delta \nu} \quad \text{and} \quad \mathbf{d}_z \hat{F} = \left(\frac{\delta F}{\delta \nu}, \mathbf{d}_z \mathbf{J} \right) + \mathbf{d}_z F.$$

Substituting into the bracket (2.6) , we get [XREF]

$$\begin{aligned}
\{\hat{F}, \hat{G}\} \quad = \quad & -\left\{ \mu, \left[\frac{\delta F}{\delta \nu}, \frac{\delta G}{\delta \nu} \right] \right\} + \{F, G\}_P + \left\{ \left(\frac{\delta F}{\delta \nu}, \mathbf{d}_z \mathbf{J} \right), \left\{ \left(\frac{\delta G}{\delta \nu}, \mathbf{d}_z \mathbf{J} \right) \right\} \right\}_P \\
& + \left\{ \left[\frac{\delta F}{\delta \nu}, \mathbf{d}_z \mathbf{J} \right], \mathbf{d}_z G \right\}_P + \left\{ \mathbf{d}_z F, \left[\frac{\delta F}{\delta \nu}, \mathbf{d}_z \mathbf{J} \right] \right\}_P \\
& - \left\{ \frac{\delta F}{\delta \nu}, \mathbf{d}_z \mathbf{J} \cdot \left(\frac{\delta G}{\delta \nu} \right)_P \right\} - \mathbf{d}_z F \cdot \left(\frac{\delta F}{\delta \nu} \right)_P \\
& + \left\{ \frac{\delta G}{\delta \nu}, \mathbf{d}_z \mathbf{J} \cdot \left(\frac{\delta F}{\delta \nu} \right)_P \right\} + \mathbf{d}_z G \cdot \left(\frac{\delta F}{\delta \nu} \right)_P .
\end{aligned} \tag{3.18.64}$$

However, $\{\mathbf{d}_z F, [(\delta G/\delta\nu), \mathbf{d}_z \mathbf{J}]\}_P$ means the pairing of $\mathbf{d}_z F$ with the Hamiltonian vector field associated with the one form $[(\delta F/\delta\nu, \mathbf{d}_z \mathbf{J}]$ which is $(\delta G/\delta\nu)_P$ by definition of the momentum map. Thus the corresponding four terms in (2.13.60) cancel. Let us consider the remaining terms. First of all, consider

$$\left\{ \left[\frac{\delta F}{\delta\nu}, \mathbf{d}_z \mathbf{J} \right], \left[\frac{\delta F}{\delta\nu}, \mathbf{d}_z \mathbf{J} \right] \right\}_P . \tag{3.18.65}$$

Since $\mathbf{J}$ is equivariant, it is a Poisson map to $\mathbf{3}^*_+$ Thus, (2.13.61) becomes $\{\mathbf{J}, [(\delta F/\delta\nu), (\delta G/\delta\nu)]\}$. Similarly each of $-[(\delta F/\delta\nu), \mathbf{d}_z\mathbf{J} \cdot \{(\delta G/\delta\nu\}_P]$ and $-[(\delta G/\delta\nu), \mathbf{d}_z\mathbf{J} \cdot \{(\delta F/\delta\nu)\}_P]$ equals $-\{\mathbf{J}, [(\delta F/\delta\nu), (\delta G/\nu)]\}$, so these three terms collapse to $-\{\mathbf{J}, [(\delta F/\delta\nu), (\delta G/\delta\nu)]\}$ which combines with $-\{\mu, [(\delta F/\delta\nu), (\delta G/\delta\nu)]\}$ to produce $-\{\nu, [(\delta F/\delta\nu), (\delta G/\delta\nu)]\} = \{F, G\}_-$. Thus, (2.13.60) collapses to (2.13.59). ■

Remarks This result is of course intimately related to the momentum shift in the reduction theorem and to the isomorphism between the "Interaction" and "Universal" representations of a reduced principal bundle. (See §3.2F). ♦

Recall that a Casimir function is a function whose Poisson bracket with any other function is zero. From Proposition **2.13.5** we get

Corollary 3.18.6 *Corollary Suppose $C(\nu)$ is a Casimir function on $\mathbf{3}^*$. Then*

$$C(\mu, x) = C(\mu + \mathbf{J}(x))$$

is a Casimir function on $\mathbf{3}^* \times P$ for the bracket (2.13.57).

We conclude with some consequences of the preceding Proposition. The first is a connection with semi-direct products. Namely, we notice that if h is another Lie algebra and G acts on h, we can reduce $T^*G \times h^*$ by G.

Corollary 3.18.7 *Giving $T^*G \times h^*$ the sum of the canonical and the "$-$" Lie-Poisson structure on h^*, the reduced space $(T^*G \times h^*)/G$ is $\mathbf{3}^* \times h^*$ with the bracket*

$$\{F, G\} = \{F, G\}_{\mathbf{3}^*} + \{F, G\}_{\mathbf{3}^*} - d_\nu F \cdot \left(\frac{\delta G}{\delta\mu} \right)_{\mathbf{3}^*} + d_\nu G \cdot \left(\frac{\delta F}{\delta\mu} \right)_{\mathbf{3}^*} \tag{3.18.66}$$

where $(\mu, \nu) \in \mathbf{3}^* \times h^*$, so (2.13.62) is the Lie-Poisson bracket for the semidirect product $\mathbf{3}$Ⓢh.

This is compatible with, and reproduces some of the semidirect product reduction results of Marsden et. al. [1983], and Marsden, Ratiu and Weinstein [1984a, b] (see also Holm, Kupershmidt and Levermore [1983]). Of course such structures are important for examples like a rigid body with a fixed point in the presence of a gravitational field (see Holmes and Marsden [1983]).

Here is another result similar to Proposition **2.13.4** which reproduces the symplectic form on T^*G written in body coordinates (Abraham and Marsden [1978, p. 315}). We phrase the result in terms of brackets.

Corollary 3.18.8 *The map of T^*G to $\mathbf{3}^* \times g$ given by $\alpha_g \mapsto (TL_g^* \alpha_g, g)$ maps the canonical bracket to the following bracket on $\mathbf{3}^* \times G$:*

$$\{F, G\} = \{F, G\}_- + \mathbf{d}_g F \cdot TL_g \left(\frac{\delta G}{\delta\mu} \right) - \mathbf{d}_g G \cdot TL_g \left(\frac{\delta F}{\delta\mu} \right) \tag{3.18.67}$$

where $\mu \in \mathbf{3}^*$ and $g \in G$.

Proof This is proved by the same method as in Proposition **2.13.4**. For $F : \mathbf{3}^* \times G \to \mathbb{R}$, let $\bar{F}(\alpha_g) = G(\mu, g)$ where $\mu = TL_g^* \alpha_g$. The canonical bracket of $\bar{F}$ and $\bar{G}$ will give the $(-)$ Lie-Poisson structure via the μ dependence. The remaining terms are

$$\left(\mathbf{d}_g \bar{F}, \frac{\delta \bar{G}}{\delta p} \right) - \left(\mathbf{d}_g \bar{G}, \frac{\delta \bar{F}}{\delta p} \right)$$

where $(\delta \bar{F}/\delta p)$ means the fiber derivative of $\bar{F}$ regarded as a vector field and $\mathbf{d}_g \bar{F}$ means the derivative holding μ fixed. Using the chain rule, one gets (2.13.63). ■

We now combine this corollary with Proposition **2.13.4** to produce a Poisson structure on $\mathbf{3}^* \times \mathbf{3}^* \times G$. This structure is relevant for the motion of two rigid bodies coupled with a ball in socket joint.

Corollary 3.18.9 *The Poisson reduced space $(T^*G \times T^*G)/G$ may be identified with the space $\mathbf{3}^* \times \mathbf{3}^* \times G$, equipped with the Poisson bracket*

$$\begin{aligned} \{F, G\}(\mu_1, \mu_2, g) \quad = \quad & \{F, G\}_{\mu_1}^- + \{F, G\}_{\mu_2}^- \\ & -\mathbf{d}_g F \cdot TR_g \cdot \left(\frac{\delta G}{\delta \mu_1} \right) + \mathbf{d}_g G \cdot TR_g \cdot \left(\frac{\delta F}{\delta \mu_1} \right) \\ & +\mathbf{d}_g F \cdot TL_g \cdot \left(\frac{\delta G}{\delta \mu_2} \right) - \mathbf{d}_g G \cdot TL_g \cdot \left(\frac{\delta F}{\delta \mu_2} \right) \end{aligned} \tag{3.18.68}$$

where $\{F, G\}_{\mu_1}^-$ is the "$-$" Lie–Poisson bracket with respect to the first variable μ_1, and similarly for $\{F, G\}_{\mu_2}^-$.

Proof The isomorphism of $(T^*G \times T^*G)/G$ with $\mathbf{3}^* \times \mathbf{3}^* \times G$ is implemented by the map

$$(\alpha_g, \beta_h) \mapsto (TL_g^* \alpha_g, TL^* \beta_h, g^{-1}h). \tag{3.18.69}$$

We map this in two steps. First, map

$$T^*G \times T^*G \to T^*G \times \mathbf{3}^* \times G$$

using Corollary **2.13.8**. Now regard G as acting on $\mathbf{3}^* \times G$ by left multiplication on the last factor alone. Then map $T^*G \times (\mathbf{3}^* \times G)$ to $\mathbf{3}^* \times \mathbf{3}^* \times G$ by Proposition **2.13.4**. Noting that at the point (μ_1, μ_2, g)

$$\left(\frac{\delta F}{\delta \mu_1} \right)_{\mathbf{3}^* \times G} = \left(0, TR_g \cdot \frac{\delta F}{\delta \mu_1} \right)$$

we get (2.13.64). This bracket (2.13.64) can also be verified by a direct calculation using the map (2.13.65). ■

Finally, we remark that the theory in this section can be applied to a variety of situations besides those in this paper. For example, Sanchez de Alvarez [1986] uses these ideas to obtain some of the results of Krishnaprasad [1985]; cf §♡♡♡ [REF]

Chapter 4

Semidirect Products

4.1 Hamiltonian Semidirect Product Theory

We first recall how the Hamiltonian theory proceeds for systems defined on semidirect products. We present the abstract theory, but of course historically this grew out of the examples, especially the heavy top and compressible flow.

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 acts *on the left* by linear maps on V (and hence G also acts on on the left on its dual space V^*). As sets, the semidirect product $S = G \circledS V$ is the Cartesian product $S = G \times V$ whose group multiplication is given by

$$(g_1, v_1)(g_2, v_2) = (g_1 g_2, v_1 + g_1 v_2), \tag{4.1.1}$$

where the action of $g \in G$ on $v \in V$ is denoted simply as gv. The identity element is $(e, 0)$ where e is the identity in G. We record for convenience the inverse of an element:

$$(g, v)^{-1} = (g^{-1}, -g^{-1}v). \tag{4.1.2}$$

The Lie algebra of S is the semidirect product Lie algebra, $\mathfrak{s} = \mathfrak{g} \circledS V$, whose bracket has the expression

$$[(\xi_1, v_1), (\xi_2, v_2)] = ([\xi_1, \xi_2], \, \xi_1 v_2 - \xi_2 v_1)\,, \tag{4.1.3}$$

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

Below we will need the formulae for the adjoint and the coadjoint actions for semidirect products. We denote these and other actions by simple concatenation; so they are expressed as (see, e.g., Marsden, Ratiu and Weinstein [1984a,b])

$$(g, v)(\xi, u) = (g\xi, gu - (g\xi)v), \tag{4.1.4}$$

and

$$(g, v)(\mu, a) = (g\mu + \rho_v^*(ga), ga), \tag{4.1.5}$$

where $(g, v) \in S = G \times V$, $(\xi, u) \in \mathfrak{s} = \mathfrak{g} \times V$, $(\mu, a) \in \mathfrak{s}^* = \mathfrak{g}^* \times V^*$, $g\xi = \mathrm{Ad}_g \xi$, $g\mu = \mathrm{Ad}^*_{g^{-1}} \mu$, ga denotes the induced *left* action of g on a (the *left* action of G on V induces a *left* action of G on V^* — the inverse of the transpose of the action on V), $\rho_v : \mathfrak{g} \to V$ is the linear map given by $\rho_v(\xi) = \xi v$, and $\rho_v^* : V^* \to \mathfrak{g}^*$ is its dual.

Important Notation. For $a \in V^*$, we shall write, for notational convenience,

$$\rho_v^* a = v \diamond a \in \mathfrak{g}^* ,$$

which is a bilinear operation in v and a. Using this notation, the above formula for the coadjoint action reads

$$(g, v)(\mu, a) = (g\mu + v \diamond (ga), ga).$$

We shall also denote actions of groups and Lie algebras by simple concatenation. For example, the $\mathfrak{g}$–action on $\mathfrak{g}^*$ and V^*, which is defined as minus the dual map of the $\mathfrak{g}$–action on $\mathfrak{g}$ and V respectively, is denoted by $\xi\mu$ and ξa for $\xi \in \mathfrak{g}$, $\mu \in \mathfrak{g}^*$, and $a \in V^*$.

Using this concatenation notation for Lie algebra actions provides the following alternative expression of the definition of $v \diamond a \in \mathfrak{g}^*$: For all $v \in V$, $a \in V^*$ and $\eta \in \mathfrak{g}$, we define

$$\langle \eta a, v \rangle = - \langle v \diamond a \, , \eta \rangle \, .$$

Left Versus Right. When working with various models of continuum mechanics and plasmas it is convenient to work with *right* representations of G on the vector space V (as in, for example, Holm, Marsden and Ratiu [1986]). We shall denote the semidirect product by the same symbol $S = G \circledS V$, the action of G on V being denoted by vg. The formulae change under these conventions as follows. Group multiplication (the analog of (4.1.1)) is given by

$$(g_1, v_1)(g_2, v_2) = (g_1 g_2, v_2 + v_1 g_2), \tag{4.1.6}$$

and the Lie algebra bracket on $\mathfrak{s} = \mathfrak{g} \circledS V$ (the analog of (4.1.3)) has the expression

$$[(\xi_1, v_1), (\xi_2, v_2)] = ([\xi_1, \xi_2], \, v_1\xi_2 - v_2\xi_1), \tag{4.1.7}$$

where we denote the induced action of $\mathfrak{g}$ on V by concatenation, as in $v_1\xi_2$. The adjoint and coadjoint actions have the formulae (analogs of (4.1.4) and (4.1.5))

$$(g, v)(\xi, u) = (g\xi, (u + v\xi)g^{-1}), \tag{4.1.8}$$

$$(g, v)(\mu, a) = (g\mu + (vg^{-1}) \diamond (ag^{-1}), ag^{-1}), \tag{4.1.9}$$

where, as usual, $g\xi = \mathrm{Ad}_g \xi$, $g\mu = \mathrm{Ad}^*_{g^{-1}} \mu$, ag denotes the inverse of the dual isomorphism defined by $g \in G$ (so that $g \mapsto ag$ is a *right* action). Note that the adjoint and coadjoint actions are *left* actions. In this case, the $\mathfrak{g}$–actions on $\mathfrak{g}^*$ and V^* are defined as before to be minus the dual map given by the $\mathfrak{g}$–actions on $\mathfrak{g}$ and V and are denoted by $\xi\mu$ (because it is a left action) and $a\xi$ (because it is a right action) respectively.

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

$$\{f, k\}_\pm(\mu, a) \quad = \quad \pm \left\langle \mu, \left[\frac{\delta f}{\delta \mu}, \frac{\delta k}{\delta \mu} \right] \right\rangle \pm \left\langle a, \frac{\delta f}{\delta \mu} \frac{\delta k}{\delta a} - \frac{\delta k}{\delta \mu} \frac{\delta f}{\delta a} \right\rangle \tag{4.1.10}$$

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

$$X_h(\mu, a) = \mp \left(\mathrm{ad}^*_{\delta h/\delta \mu} \, \mu - \frac{\delta h}{\delta a} \diamond a, \, - \frac{\delta h}{\delta \mu} \, a \right) . \tag{4.1.11}$$

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

$$\dot{\mu} \;=\; \mp\, \mathrm{ad}^*_{\delta h/\delta \mu}\, \mu \pm \frac{\delta h}{\delta a} \diamond a\,, \tag{4.1.12}$$

$$\dot{a} \;=\; \pm\, \frac{\delta h}{\delta \mu}\, a\,, \tag{4.1.13}$$

where overdot denotes time derivative. For *right* representations of G on V the above formulae change to:

$$\{f,k\}_{\pm}(\mu,a) \;=\; \pm\left\langle \mu, \left[\frac{\delta f}{\delta \mu}, \frac{\delta k}{\delta \mu}\right]\right\rangle \mp \left\langle a, \frac{\delta k}{\delta a}\frac{\delta f}{\delta \mu} - \frac{\delta f}{\delta a}\frac{\delta k}{\delta \mu}\right\rangle, \tag{4.1.14}$$

$$X_h(\mu,a) = \mp\left(\mathrm{ad}^*_{\delta h/\delta \mu}\, \mu + \frac{\delta h}{\delta a} \diamond a,\; a\,\frac{\delta h}{\delta \mu}\right), \tag{4.1.15}$$

$$\dot{\mu} \;=\; \mp\, \mathrm{ad}^*_{\delta h/\delta \mu}\, \mu \mp \frac{\delta h}{\delta a} \diamond a\,, \tag{4.1.16}$$

$$\dot{a} \;=\; \mp\, a\, \frac{\delta h}{\delta \mu}\,. \tag{4.1.17}$$

Symplectic Actions by Semidirect Products. To avoid a proliferation of signs, in *this section* we consider all semidirect products to come from a left representation. Of course if the representation is from the right, there are similar formulae.

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

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

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

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

$$\mathbf{J}_V(gz) = g\mathbf{J}_V(z) \tag{4.1.18}$$

where we denote the appropriate action of $g \in G$ on an element by concatenation, as before. To prove (4.1.18), one uses the fact that for the coadjoint action of S on $\mathfrak{s}^*$ the second component is just the dual of the given action of G on V.

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

The semidirect product reduction theorem states, roughly speaking, that for the semidirect product $S = G \circledS V$ where G is a group acting on a vector space V and S is the semidirect product, one can first reduce T^*S by V and then by G and thereby obtain the same result as when reducing by S. As above, we let $\mathfrak{s} = \mathfrak{g} \circledS V$ denote the Lie algebra of S. The precise statement is as follows.

Theorem 4.1.1 (Semidirect Product Reduction Theorem.) *Let $S = G \circledS V$, choose $\sigma = (\mu, a) \in \mathfrak{g}^* \times V^*$, and reduce T^*S by the action of S at σ giving the coadjoint orbit $\mathcal{O}_\sigma$ through $\sigma \in \mathfrak{s}^*$. There is a symplectic diffeomorphism between $\mathcal{O}_\sigma$ and the reduced space obtained by reducing T^*G by the subgroup G_a (the isotropy of G for its action on V^* at the point $a \in V^*$) at the point $\mu|\mathfrak{g}_a$ where $\mathfrak{g}_a$ is the Lie algebra of G_a.*

Reduction by Stages. This result is a special case of a theorem on reduction by stages for semidirect products acting on a symplectic manifold (see Marsden, Misiolek, Perlmutter and Ratiu [1997] for this and more general results and see Leonard and Marsden [1997] for an application to underwater vehicle dynamics).

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

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

Theorem 4.1.2 (Reduction by Stages for Semidirect Products.) *The reduced space $(P_a)_{\mu_a}$ is symplectically diffeomorphic to the reduced space P_σ obtained by reducing P by S at the point $\sigma = (\mu, a)$.*

Combined with the cotangent bundle reduction theorem (see Abraham and Marsden [1978] and Marsden [1992] for an exposition and references), the semidirect product reduction theorem is a useful tool. For example, using these tools, 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.

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

- We start with a Hamiltonian H_{a_0} on T^*G that depends parametrically on a variable $a_0 \in V^*$.
- The Hamiltonian, regarded as a map $H : T^*G \times V^* \to \mathbb{R}$ is assumed to be invariant on T^*G under the action of G on $T^*G \times V^*$.
- One shows that this condition is equivalent to the invariance of the function H defined on $T^*S = T^*G \times V \times V^*$ extended to be constant in the variable V under the action of the semidirect product.

- By the semidirect product reduction theorem, the dynamics of H_{a_0} reduced by G_{a_0}, the isotropy group of a_0, is symplectically equivalent to Lie-Poisson dynamics on $\mathfrak{s}^* = \mathfrak{g}^* \times V^*$.

- This Lie-Poisson dynamics is given by the equations (4.1.12) and (4.1.13) for the function $h(\mu, a) = H(\alpha_g, g^{-1}a)$ where $\mu = g^{-1}\alpha_g$.

4.2 Lagrangian Semidirect Product Theory

Despite all the activity in the Hamiltonian theory of semidirect products, little attention has been paid to the corresponding Lagrangian side. Now that Lagrangian reduction is maturing (see Marsden and Scheurle [1993a,b]), it is appropriate to consider the corresponding Lagrangian question. We shall formulate four versions, depending on the nature of the actions and invariance properties of the Lagrangian. (Two of them are relegated to the appendix.)

It should be noted that *none of the theorems below require that the Lagrangian be nondegenerate*. The subsequent theory is entirely based on variational principles with symmetry and is not dependent on any previous Hamiltonian formulation. We shall, however, show that this purely Lagrangian formulation is equivalent to the Hamiltonian formulation on duals of semidirect products, provided an appropriately defined Legendre transformation happens to be a diffeomorphism.

The theorems that follow are modelled after the reduction theorem for the basic Euler–Poincaré equations given earlier. However, as we shall explain, they are *not* literally special cases of it. To distinguish the two types of results, we shall use phrases like *basic* Euler–Poincaré equations for the equations (1.6.4) and simply the Euler–Poincaré equations or the Euler–Poincaré equations *with advection* or the Euler–Poincaré equations *with advected parameters*, for the equations that follow.

The main difference between the left (right) invariant Lagrangians considered in the theorem above and the ones we shall work with below is that L and l depend in addition on another parameter $a \in V^*$, where V is a representation space for the Lie group G and L has an invariance property relative to both arguments. As we shall see below, the resulting ***Euler–Poincaré*** equations are *not* the Euler–Poincaré equations for the semidirect product Lie algebra $\mathfrak{g} \circledS V^*$ or on $\mathfrak{g} \circledS V$, for that matter.

Upcoming Examples. As we shall see in the examples, the parameter $a \in V^*$ acquires dynamical meaning under Lagrangian reduction. For the heavy top, the parameter is the unit vector in the direction of gravity, which becomes a dynamical variable in the body representation. For compressible fluids, the parameter is the density of the fluid in the reference configuration, which becomes a dynamical variable (satisfying the continuity equation) in the spatial representation.

Left Representation and Left Invariant Lagrangian. We begin with the following ingredients:

- There is a *left* representation of Lie group G on the vector space V and G acts in the natural way on the *left* on $TG \times V^*$: $h(v_g, a) = (hv_g, ha)$.

- Assume that the function $L : TG \times V^* \to \mathbb{R}$ is left G–invariant.

- In particular, if $a_0 \in V^*$, define the Lagrangian $L_{a_0} : TG \to \mathbb{R}$ by $L_{a_0}(v_g) = L(v_g, a_0)$. Then L_{a_0} is left invariant under the lift to TG of the left action of G_{a_0} on G, where G_{a_0} is the isotropy group of a_0.

- Left G–invariance of L permits us to define $l : \mathfrak{g} \times V^* \to \mathbb{R}$ by

$$l(g^{-1}v_g, g^{-1}a_0) = L(v_g, a_0).$$

  Conversely, this relation defines for any $l : \mathfrak{g} \times V^* \to \mathbb{R}$ a left G–invariant function $L : TG \times V^* \to \mathbb{R}$.

- For a curve $g(t) \in G$, let

$$\xi(t) := g(t)^{-1}\dot{g}(t)$$

  and define the curve $a(t)$ as the unique solution of the following linear differential equation with time dependent coefficients

$$\dot{a}(t) = -\xi(t)a(t),$$

  with initial condition $a(0) = a_0$. The solution can be written as $a(t) = g(t)^{-1}a_0$.

Theorem 4.2.1 *With the preceding notation, the following are equivalent:*

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

$$\delta \int_{t_1}^{t_2} L_{a_0}(g(t), \dot{g}(t))dt = 0 \tag{4.2.1}$$

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

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

iii *The constrained variational principle*[1]

$$\delta \int_{t_1}^{t_2} l(\xi(t), a(t))dt = 0 \tag{4.2.2}$$

holds on $\mathfrak{g} \times V^*$, using variations of ξ and a of the form

$$\delta\xi = \dot{\eta} + [\xi, \eta], \quad \delta a = -\eta a, \tag{4.2.3}$$

where $\eta(t) \in \mathfrak{g}$ vanishes at the endpoints.

iv *The* ***Euler–Poincaré*** *equations*[2] *hold on $\mathfrak{g} \times V^*$*

$$\frac{d}{dt}\frac{\delta l}{\delta \xi} = \mathrm{ad}^*_\xi \frac{\delta l}{\delta \xi} + \frac{\delta l}{\delta a} \diamond a. \tag{4.2.4}$$

Proof. The equivalence of **i** and **ii** holds for any configuration manifold and so, in particular, it holds in this case.

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

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

Next we show the equivalence of **iii** and **iv**. Indeed, using the definitions, integrating by parts, and taking into account that $\eta(t_1) = \eta(t_2) = 0$, we compute the variation of the integral to be

$$\begin{aligned}
\delta \int_{t_1}^{t_2} l(\xi(t), a(t))dt &= \int_{t_1}^{t_2} \left(\left\langle \frac{\delta l}{\delta \xi}, \delta\xi \right\rangle + \left\langle \delta a, \frac{\delta l}{\delta a} \right\rangle \right) dt \\
&= \int_{t_1}^{t_2} \left(\left\langle \frac{\delta l}{\delta \xi}, \dot{\eta} + \operatorname{ad}_\xi \eta \right\rangle - \left\langle \eta a, \frac{\delta l}{\delta a} \right\rangle \right) dt \\
&= \int_{t_1}^{t_2} \left(\left\langle -\frac{d}{dt} \left(\frac{\delta l}{\delta \xi} \right) + \operatorname{ad}^*_\xi \frac{\delta l}{\delta \xi}, \eta \right\rangle + \left\langle \frac{\delta l}{\delta a} \diamond a, \eta \right\rangle \right) dt \\
&= \int_{t_1}^{t_2} \left\langle -\frac{d}{dt} \left(\frac{\delta l}{\delta \xi} \right) + \operatorname{ad}^*_\xi \frac{\delta l}{\delta \xi} + \frac{\delta l}{\delta a} \diamond a, \eta \right\rangle dt
\end{aligned}$$

and so the result follows.

Finally we show that **i** and **iii** are equivalent. First note that the G–invariance of $L : TG \times V^* \to \mathbb{R}$ and the definition of $a(t) = g(t)^{-1} a_0$ imply that the integrands in (4.2.1) and (4.2.2) are equal. However, all variations $\delta g(t) \in TG$ of $g(t)$ with fixed endpoints induce and are induced by variations $\delta\xi(t) \in \mathfrak{g}$ of $\xi(t)$ of the form $\delta\xi = \dot{\eta} + [\xi, \eta]$ with $\eta(t) \in \mathfrak{g}$ vanishing at the endpoints; the relation between $\delta g(t)$ and $\eta(t)$ is given by $\eta(t) = g(t)^{-1}\delta g(t)$. This is the content of the following lemma proved in Bloch et al. [1996]. [3]

Lemma 4.2.2 *Let $g : U \subset \mathbb{R}^2 \to G$ be a smooth map and denote its partial derivatives by*

$$\xi(t, \varepsilon) = TL_{g(t,\varepsilon)^{-1}}(\partial g(t, \varepsilon)/\partial t)$$

and

$$\eta(t, \varepsilon) = TL_{g(t,\varepsilon)^{-1}}(\partial g(t, \varepsilon)/\partial \varepsilon).$$

Then

$$\frac{\partial \xi}{\partial \varepsilon} - \frac{\partial \eta}{\partial t} = [\xi, \eta]\,. \tag{4.2.5}$$

Conversely, if U is simply connected and $\xi, \eta : U \to \mathfrak{g}$ are smooth functions satisfying (4.2.5) then there exists a smooth function $g : U \to G$ such that $\xi(t, \varepsilon) = TL_{g(t,\varepsilon)^{-1}}(\partial g(t, \varepsilon)/\partial t)$ and $\eta(t, \varepsilon) = TL_{g(t,\varepsilon)^{-1}}(\partial g(t, \varepsilon)/\partial \varepsilon)$.

Thus, if **i** holds, we define $\eta(t) = g(t)^{-1}\delta g(t)$ for a variation $\delta g(t)$ with fixed endpoints. Then if we let $\delta\xi = g(t)^{-1}\dot{g}(t)$, we have by the above proposition $\delta\xi = \dot{\eta} + [\xi, \eta]$. In addition, the variation of $a(t) = g(t)^{-1}a_0$ is $\delta a(t) = -\eta(t)a(t)$. Conversely, if $\delta\xi = \dot{\eta} + [\xi, \eta]$ with $\eta(t)$ vanishing at the endpoints, we define $\delta g(t) = g(t)\eta(t)$ and the above proposition guarantees then that this $\delta g(t)$ is the general variation of $g(t)$ vanishing at the endpoints. From $\delta a(t) = -\eta(t)a(t)$ it follows that the variation of $g(t)a(t) = a_0$ vanishes, which is consistent with the dependence of L_{a_0} only on $g(t), \dot{g}(t)$. ■

Cautionary Remarks. Let us explicitly show that *these Euler–Poincaré equations (4.2.4) are not the Euler–Poincaré equations for the semidirect product Lie algebra $\mathfrak{g} Ⓢ V^*$*. Indeed, by (1.6.4) the basic Euler–Poincaré equations

$$\frac{d}{dt} \frac{\delta l}{\delta(\xi, a)} = \operatorname{ad}^*_{(\xi,a)} \frac{\delta l}{\delta(\xi, a)}\,, \quad (\xi, a) \in \mathfrak{g} Ⓢ V^*$$

[3] This lemma is simple for matrix groups, as in Marsden and Ratiu [1998], but it is less elementary for general Lie groups.

for $l : \mathfrak{g} \circledS V^* \to \mathbb{R}$ become

$$\frac{d}{dt}\frac{\delta l}{\delta \xi} = \operatorname{ad}^*_\xi \frac{\delta l}{\delta \xi} + \frac{\delta l}{\delta a} \diamond a, \quad \frac{d}{dt}\frac{\delta l}{\delta a} = -\xi \frac{\delta l}{\delta a}\,,$$

which is a *different* system from that given by the Euler–Poincaré equation (4.2.4) and $\dot{a} = -\xi a$, even though the first equations of both systems are identical.

The Legendre Transformation. As we explained earlier, one normally thinks of passing from Euler–Poincaré equations on a Lie algebra $\mathfrak{g}$ to Lie–Poisson equations on the dual $\mathfrak{g}^*$ by means of the Legendre transformation. In our case, we start with a Lagrangian on $\mathfrak{g} \times V^*$ and perform a partial Legendre transformation in the variable ξ only, by writing

$$\mu = \frac{\delta l}{\delta \xi}\,, \quad h(\mu, a) = \langle \mu, \xi \rangle - l(\xi, a). \tag{4.2.6}$$

Since

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

and $\delta h/\delta a = -\delta l/\delta a$, we see that (4.2.4) and $\dot{a}(t) = -\xi(t)a(t)$ imply (4.1.11) for the *minus* Lie–Poisson bracket (that is, the sign + in (4.1.11)). If this Legendre transformation is invertible, then we can also pass from the the minus Lie–Poisson equations (4.1.11) to the Euler–Poincaré equations (4.2.4) together with the equations $\dot{a}(t) = -\xi(t)a(t)$.

Right Representation and Right Invariant Lagrangian. There are four versions of the preceding theorem, the given left-left version, a left-right, a right-left and a right-right version. For us, the most important ones are the left-left and the right-right versions. We state the remaining two in the appendix.

Here we make the following assumptions:

- There is a *right* representation of Lie group G on the vector space V and G acts in the natural way on the *right* on $TG \times V^*$: $(v_g, a)h = (v_g h, ah)$.
- Assume that the function $L : TG \times V^* \to \mathbb{R}$ is right G–invariant.
- In particular, if $a_0 \in V^*$, define the Lagrangian $L_{a_0} : TG \to \mathbb{R}$ by $L_{a_0}(v_g) = L(v_g, a_0)$. Then L_{a_0} is right invariant under the lift to TG of the right action of G_{a_0} on G, where G_{a_0} is the isotropy group of a_0.
- Right G–invariance of L permits us to define $l : \mathfrak{g} \times V^* \to \mathbb{R}$ by

  $$l(v_g g^{-1}, a_0 g^{-1}) = L(v_g, a_0).$$

  Conversely, this relation defines for any $l : \mathfrak{g} \times V^* \to \mathbb{R}$ a right G–invariant function $L : TG \times V^* \to \mathbb{R}$.
- For a curve $g(t) \in G$, let $\xi(t) := \dot{g}(t)g(t)^{-1}$ and define the curve $a(t)$ as the unique solution of the linear differential equation with time dependent coefficients $\dot{a}(t) = -a(t)\xi(t)$ with initial condition $a(0) = a_0$. The solution can be written as $a(t) = a_0 g(t)^{-1}$.

Theorem 4.2.3 *The following are equivalent:*

i *Hamilton's variational principle*

$$\delta \int_{t_1}^{t_2} L_{a_0}(g(t), \dot{g}(t))dt = 0 \tag{4.2.7}$$

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

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

iii *The constrained variational principle*

$$\delta \int_{t_1}^{t_2} l(\xi(t), a(t))dt = 0 \tag{4.2.8}$$

holds on $\mathfrak{g} \times V^*$, using variations of the form

$$\delta\xi = \dot{\eta} - [\xi, \eta], \quad \delta a = -a\eta, \tag{4.2.9}$$

where $\eta(t) \in \mathfrak{g}$ vanishes at the endpoints.

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

$$\frac{d}{dt}\frac{\delta l}{\delta \xi} = -\operatorname{ad}^*_\xi \frac{\delta l}{\delta \xi} + \frac{\delta l}{\delta a} \diamond a. \tag{4.2.10}$$

The same partial Legendre transformation (4.2.6) as before maps the Euler–Poincaré equations (4.2.10), together with the equations $\dot{a} = -a\xi$ for a to the plus Lie–Poisson equations (4.1.16) and (4.1.17) (that is, one chooses the overall minus sign in these equations).

Generalizations. The Euler–Poincaré equations are a special case of the reduced Euler-Lagrange equations (see Marsden and Scheurle [1993b] and Cendra, Marsden and Ratiu [1997]). This is shown explicitly in Cendra, Holm, Marsden and Ratiu [1997]. There is, however, an easy generalization that is needed in some of the examples we will consider. Namely, if $L : TG \times V^* \times TQ$ and if the group G acts in a trivial way on TQ, then one can carry out the reduction in the same way as above, carrying along the Euler-Lagrange equations for the factor Q at each step. The resulting reduced equations then are the Euler–Poincaré equations above for the $\mathfrak{g}$ factor, together the Euler-Lagrange equations for the $q \in Q$ factor. The system is coupled through the dependence of L on all variables. (For a full statement, see Cendra, Holm, Hoyle and Marsden [1997], who use this extension to treat the Euler–Poincaré formulation of the Maxwell-Vlasov equations for plasma physics.)

4.3 The Kelvin-Noether Theorem

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

There is a version of the theorem that holds for each of the choices of conventions, but we shall pick the left-left conventions to illustrate the result.

The Kelvin-Noether Quantity. We start with a Lagrangian L_{a_0} depending on a parameter $a_0 \in V^*$ as above. We introduce a manifold $\mathcal{C}$ on which G acts (we assume this is also a left action) and suppose we have an equivariant map $\mathcal{K} : \mathcal{C} \times V^* \to \mathfrak{g}^{**}$.

As we shall see, in the case of continuum theories, the space $\mathcal{C}$ will be a loop space and $\langle \mathcal{K}(c, a), \mu \rangle$ for $c \in \mathcal{C}$ and $\mu \in \mathfrak{g}^*$ will be a circulation. This class of examples also shows why we *do not* want to identify the double dual $\mathfrak{g}^{**}$ with $\mathfrak{g}$.

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

$$I(c, \xi, a) = \left\langle \mathcal{K}(c, a), \frac{\delta l}{\delta \xi}(\xi, a) \right\rangle. \tag{4.3.1}$$

We are now ready to state the main theorem of this section.

Theorem 4.3.1 (Kelvin-Noether.) *Fixing $c_0 \in \mathcal{C}$, let $\xi(t), a(t)$ satisfy the Euler–Poincaré equations and define $g(t)$ to be the solution of $\dot{g}(t) = g(t)\xi(t)$ and, say, $g(0) = e$. Let $c(t) = g(t)^{-1} c_0$ and $I(t) = I(c(t), \xi(t), a(t))$. Then*

$$\frac{d}{dt} I(t) = \left\langle \mathcal{K}(c(t), a(t)), \frac{\delta l}{\delta a} \diamond a \right\rangle. \tag{4.3.2}$$

Proof. First of all, write $a(t) = g(t)^{-1} a_0$ as we did previously and use equivariance to write $I(t)$ as follows:

$$\left\langle \mathcal{K}(c(t), a(t)), \frac{\delta l}{\delta \xi}(\xi(t), a(t)) \right\rangle = \left\langle \mathcal{K}(c_0, a_0), g(t) \left[\frac{\delta l}{\delta \xi}(\xi(t), a(t)) \right] \right\rangle.$$

The g^{-1} pulls over to the right side as g (and not g^{-1}) because of our conventions of always using left representations. We now differentiate the right hand side of this equation. To do so, we use the following well known formula for differentiating the coadjoint action (see Marsden and Ratiu [1998], page 276):

$$\frac{d}{dt} [g(t)\mu(t)] = g(t) \left[-\operatorname{ad}^*_{\xi(t)} \mu(t) + \frac{d}{dt} \mu(t) \right],$$

where, as usual,

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

Using this coadjoint action formula and the Euler–Poincaré equations, we obtain

$$\begin{aligned}
\frac{d}{dt} I &= \frac{d}{dt} \left\langle \mathcal{K}(c_0, a_0), g(t) \left[\frac{\delta l}{\delta \xi}(\xi(t), a(t)) \right] \right\rangle \\
&= \left\langle \mathcal{K}(c_0, a_0), \frac{d}{dt} \left\{ g(t) \left[\frac{\delta l}{\delta \xi}(\xi(t), a(t)) \right] \right\} \right\rangle \\
&= \left\langle \mathcal{K}(c_0, a_0), g(t) \left[-\operatorname{ad}^*_{\xi} \frac{\delta l}{\delta \xi} + \operatorname{ad}^*_{\xi} \frac{\delta l}{\delta \xi} + \frac{\delta l}{\delta a} \diamond a \right] \right\rangle \\
&= \left\langle \mathcal{K}(c_0, a_0), g(t) \left[\frac{\delta l}{\delta a} \diamond a \right] \right\rangle \\
&= \left\langle g(t)^{-1} \mathcal{K}(c_0, a_0), \left[\frac{\delta l}{\delta a} \diamond a \right] \right\rangle \\
&= \left\langle \mathcal{K}(c(t), a(t)), \left[\frac{\delta l}{\delta a} \diamond a \right] \right\rangle.
\end{aligned}$$

where, in the last steps, we used the definitions of the coadjoint action, as well as the Euler–Poincaré equation (4.2.4) and the equivariance of the map $\mathcal{K}$. ■

Corollary 4.3.2 *For the basic Euler–Poincaré equations, the Kelvin quantity $I(t)$, defined the same way as above but with $I : \mathcal{C} \times \mathfrak{g} \to \mathbb{R}$, is conserved.*

For a review of the standard Noether theorem results for energy and momentum conservation in the context of the general theory, see, e.g., Marsden and Ratiu [1998].

4.4 The Heavy Top

In this section we shall use Theorem 4.2.1 to derive the classical Euler–Poisson equations for the heavy top. Our purpose is merely to illustrate the theorem with a concrete example.

The Heavy Top Lagrangian. The heavy top kinetic energy is given by the left invariant metric on $SO(3)$ whose value at the identity is $\langle \mathbf{\Omega}_1, \mathbf{\Omega}_2 \rangle = \mathbb{I}\mathbf{\Omega}_1 \cdot \mathbf{\Omega}_2$, where $\mathbf{\Omega}_1, \mathbf{\Omega}_2 \in \mathbb{R}^3$ are thought of as elements of $\mathfrak{so}(3)$, the Lie algebra of $SO(3)$, via the isomorphism $\mathbf{\Omega} \in \mathbb{R}^3 \mapsto \hat{\mathbf{\Omega}} \in \mathfrak{so}(3)$, $\hat{\mathbf{\Omega}}\mathbf{v} := \mathbf{\Omega} \times \mathbf{v}$, and where $\mathbb{I}$ is the (time independent) moment of inertia tensor in body coordinates, usually taken as a diagonal matrix by choosing the body coordinate system to be a principal axes body frame. This kinetic energy is thus left invariant under the full group $SO(3)$. The potential energy is given by the work done in lifting the weight of the body to the height of its center of mass, with the direction of gravity pointing downwards. If M denotes the total mass of the top, g the magnitude of the gravitational acceleration, $\boldsymbol{\chi}$ the unit vector of the oriented line segment pointing from the fixed point about which the top rotates (the origin of a spatial coordinate system) to the center of mass of the body, and ℓ its length, then the potential energy is given by $-Mg\ell\,\mathbf{R}^{-1}\mathbf{e}_3 \cdot \boldsymbol{\chi}$, where $\mathbf{e}_3$ is the axis of the spatial coordinate system parallel to the direction of gravity but pointing upwards. This potential energy breaks the full $SO(3)$ symmetry and is invariant only under the rotations S^1 about the $\mathbf{e}_3$–axis.

However, for the application of Theorem 4.2.1 we are supposed to think of the Lagrangian of the heavy top as a function on $TSO(3)\times\mathbb{R}^3 \to \mathbb{R}$. That is, we need to think of the potential energy as a function of $(u_{\mathbf{R}}, \mathbf{v}) \in TSO(3) \times \mathbb{R}^3$. This means that we need to replace the vector giving the direction of gravity $\mathbf{e}_3$ by an arbitrary vector $\mathbf{v} \in \mathbb{R}^3$, so that the potential equals

$$U(u_{\mathbf{R}}, \mathbf{v}) = Mg\ell\,\mathbf{R}^{-1}\mathbf{v} \cdot \boldsymbol{\chi}.$$

Thought of this way, the potential is $SO(3)$–invariant. Indeed, if $\mathbf{R}' \in SO(3)$ is arbitrary, then

$$\begin{aligned} U(\mathbf{R}'u_{\mathbf{R}}, \mathbf{R}'\mathbf{v}) &= Mg\ell\,(\mathbf{R}'\mathbf{R})^{-1}\mathbf{R}'\mathbf{v} \cdot \boldsymbol{\chi} \\ &= Mg\ell\,\mathbf{R}^{-1}\mathbf{v} \cdot \boldsymbol{\chi} \\ &= U(u_{\mathbf{R}}, \mathbf{v}) \end{aligned}$$

and the hypotheses of Theorem 4.2.1 are satisfied. Thus, the heavy top equations of motion in the body representation are given by the Euler–Poincaré equations (4.2.4) for the Lagrangian $l : \mathfrak{so}(3) \times \mathbb{R}^3 \to \mathbb{R}$.

The Reduced Lagrangian. To compute the explicit expression of l, denote by $\mathbf{\Omega}$ the angular velocity and by $\mathbf{\Pi} = \mathbb{I}\mathbf{\Omega}$ the angular momentum in the body representation. Let $\mathbf{\Gamma} = \mathbf{R}^{-1}\mathbf{v}$; if $\mathbf{v} = \mathbf{e}_3$, the unit vector pointing upwards on the vertical spatial axis, then $\mathbf{\Gamma}$ is this unit vector viewed by an observer moving with the body. The Lagrangian l :

$\mathfrak{so}(3) \times \mathbb{R}^3 \to \mathbb{R}$ is thus given by

$$\begin{aligned} l(\boldsymbol{\Omega}, \boldsymbol{\Gamma}) &= L(\mathbf{R}^{-1} u_{\mathbf{R}}, \mathbf{R}^{-1}\mathbf{v}) \\ &= \frac{1}{2}\boldsymbol{\Pi} \cdot \boldsymbol{\Omega} + U(\mathbf{R}^{-1} u_{\mathbf{R}}, \mathbf{R}^{-1}\mathbf{v}) \\ &= \frac{1}{2}\boldsymbol{\Pi} \cdot \boldsymbol{\Omega} + Mg\ell\, \boldsymbol{\Gamma} \cdot \boldsymbol{\chi}\,. \end{aligned}$$

The Euler–Poincaré Equations. It is now straightforward to compute the Euler–Poincaré equations. First note that

$$\frac{\delta l}{\delta \boldsymbol{\Omega}} = \boldsymbol{\Pi}, \quad \frac{\delta l}{\delta \boldsymbol{\Gamma}} = Mg\ell\, \boldsymbol{\chi}\,.$$

Since

$$\mathrm{ad}^*_{\boldsymbol{\Omega}}\, \boldsymbol{\Pi} = \boldsymbol{\Pi} \times \boldsymbol{\Omega}\,, \quad \mathbf{v} \diamond \boldsymbol{\Gamma} = -\boldsymbol{\Gamma} \times \mathbf{v}\,,$$

and

$$\boldsymbol{\Omega}\boldsymbol{\Gamma} = -\boldsymbol{\Gamma} \times \boldsymbol{\Omega}\,,$$

the Euler–Poincaré equations are

$$\dot{\boldsymbol{\Pi}} = \boldsymbol{\Pi} \times \boldsymbol{\Omega} + Mg\ell\, \boldsymbol{\Gamma} \times \boldsymbol{\chi}\,,$$

which are coupled to the $\boldsymbol{\Gamma}$ evolution

$$\dot{\boldsymbol{\Gamma}} = \boldsymbol{\Gamma} \times \boldsymbol{\Omega}\,.$$

This system of two vector equations comprises the classical Euler–Poisson equations, which describe the motion of the heavy top in the body representation.

The Kelvin-Noether theorem Let $\mathcal{C} = \mathfrak{g}$ and let $\mathcal{K} : \mathcal{C} \times V^* \to \mathfrak{g}^{**} \cong \mathfrak{g}$ be the map $(\mathbf{W}, \boldsymbol{\Gamma}) \mapsto \mathbf{W}$. Then the Kelvin-Noether theorem gives the statement

$$\frac{d}{dt} \langle \mathbf{W}, \boldsymbol{\Pi} \rangle = Mg\ell \, \langle \mathbf{W}, \boldsymbol{\Gamma} \times \boldsymbol{\chi} \rangle$$

where $\mathbf{W}(t) = \mathbf{R}(t)^{-1}\mathbf{w}$; in other words, $\mathbf{W}(t)$ is the body representation of a space fixed vector. This statement is easily verified directly. Also, note that $\langle \mathbf{W}, \boldsymbol{\Pi} \rangle = \langle \mathbf{w}, \boldsymbol{\pi} \rangle$, with $\boldsymbol{\pi} = \mathbf{R}(t)\boldsymbol{\Pi}$, so the Kelvin-Noether theorem may be viewed as a statement about the rate of change of the momentum map of the system (the spatial angular momentum) relative to the full group of rotations, not just those about the vertical axis.

Chapter 5

Semidirect Product Reduction and Reduction by Stages

5.1 Semidirect Product Reduction

In some applications (such as compressible fluids, the heavy top, MHD, etc.), one has two symmetry groups that don't commute and thus the preceding theorem does not apply. Then it matters in what order one performs the reduction. This occurs, in particular for semidirect products. The main result covering the case of semidirect products is due to Marsden, Ratiu and Weinstein [1984ab] (see this paper for further references, but briefly, important previous versions were due to Vinogradov and Kupershmidt [1977], Ratiu [1980], Guillemin and Sternberg [1980], Ratiu [1981], [1982], Holm and Kupershmidt [1983] and Guillemin and Sternberg [1984]). The semidirect product reduction theorem states, roughly speaking, that for the semidirect product $S = G \circledS V$ where G is a group acting on a vector space V and S is the semidirect product, one can first reduce T^*S by V and then by G and one gets the same result as reducing by S. We will let $\mathfrak{s}$ denote the Lie algebra of S so that $\mathfrak{s} = \mathfrak{g} \circledS V$.

We now state this semidirect product reduction theorem precisely.

Theorem 5.1.1 Semidirect Product Reduction Theorem *Let $S = G \circledS V$ as above and choose $\sigma = (\mu, a) \in \mathfrak{g}^* \times V^*$ and reduce T^*S by the action of S at σ giving the coadjoint orbit $\mathcal{O}_\sigma$ through $\sigma \in \mathfrak{s}^*$. There is a symplectic diffeomeorphism between $\mathcal{O}_\sigma$ and the reduced space obtained by reducing T^*G by the subgroup G_a (the isotropy of G at the point $a \in V^*$) at the point $\mu|\mathfrak{g}_a$ where $\mathfrak{g}_a$ is the Lie algebra of G_a.*

The commuting reduction theorem for the case in which K is a vector space is a special case of this result in which we take the action of G on K to be trivial. This already suggests that there is a generalization of the semidirect product reduction theorem to the case in which V is replaced by a general group. We shall see that this is indeed the case later on. Note that in the commuting reduction theorem, what we called ν is called a in the semidirect product reduction theorem.

We refer to the original papers (Marsden, Ratiu and Weinstein [1984ab]) for a direct proof. In this paper we shall obtain the theorem as a special case of more general results. The main idea linking reduction by stages with semidirect product reduction is the following: we regard the reduction of T^*G by G_a as the successive reduction of T^*S by V followed by reduction by G_a.

Combined with the cotangent bundle reduction theorem, the semidirect product reduction theorem is a very useful tool. For example, using these tools, one sees right away that the generic coadjoint orbits for the Euclidean group are cotangent bundles of spheres with the symplectic structure given by the canonical structure plus a magnetic term. We also point out that the theory of semidirect products was motivated by several examples of physical interest, such as the Poisson structure for compressible fluids and MHD. These examples are discussed in the original papers.

There is a technique for reducing dynamics that is associated with the geometry of the semidirect product reduction theorem. In effect, one can start with a Hamiltonian on either of the phase spaces and induce one (and hence its associated dynamics) on the other space in a natural way. For example, in many applications, one starts with a Hamiltonian H_a on T^*G that depends parametrically on a variable $a \in V^*$; this parametric dependence identifies the space V^* and hence V. The Hamiltonian, regarded as a map $H : T^*G \times V^* \to \mathbb{R}$ should be invariant on T^*G under the action of G on $T^*G \times V^*$. This condition is equivalent to the invariance of the corresponding function on $T^*S = T^*G \times V \times v^*$ extended to be constant in the variable V under the action of the semidirect product.

5.2 Reduction by Stages for Semidirect Products

Although it was not shown explicitly in Marsden, Ratiu and Weinstein [1984ab], the semidirect product reduction theorem generalizes to the setting of semidirect products acting on any symplectic manifold, not merely the natural (left) action on T^*S. We shall concentrate on this case in this section and focus on the problem of generalizing further to the context of a group M with a normal subgroup N in the following sections.

To recap the setting, we start with a semidirect product, $S = G \circledS V$ where V is a vector space and the Lie group G acts on V (and hence on its dual space V^*). Recall that as sets, $S = G \times V$ and that the group multiplication is given by

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

where the action of $g \in G$ on $v \in V$ is denoted simply as gv. The Lie algebra of S is the semidirect product of Lie algebras: $\mathfrak{s} = \mathfrak{g} \circledS V$. The bracket is given by

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

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

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

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

and

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

where $(g, v) \in S = G \times V$, $(\xi, u) \in \mathfrak{s} = \mathfrak{g} \times V$, $(\mu, a) \in \mathfrak{s}^* = \mathfrak{g}^* \times V^*$ and where $\rho_v : \mathfrak{g} \to V$ is the derivative of the map $g \mapsto gv$ at the identity and $\rho_v^* : V^* \to \mathfrak{g}^*$ is its dual. The infinitesimal action of $\mathfrak{g}$ on V will often be denoted by ξv; note that $\xi v = \rho_v(\xi)$.

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

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

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

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

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

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

We can carry out reduction of P by S at a regular value $\sigma = (\mu, a)$ of the momentum map $\mathbf{J}_S$ for S in two stages using the following procedure (see figure 5.2.1).

- First reduce P by V at the value a (assume it to be a regular value) to get the reduced space $P_a = \mathbf{J}_V^{-1}(a)/V$. Here the reduction is by an abelian group, so the quotient is done by the whole of V. We will let the projection to the reduced space be denoted π_a:
$$\pi_a : \mathbf{J}_V^{-1}(a) \to P_a.$$
- Form the group G_a consisting of elements of G that leave the point a fixed using the action of G on V^*. The group G_a acts on P_a and has an induced momentum map $\mathbf{J}_a : P_a \to \mathfrak{g}_a^*$, where $\mathfrak{g}_a$ is the Lie algebra of G_a, as is shown below, so we can reduce P_a at the point $\mu_a := \mu|\mathfrak{g}_a$ to get the reduced space $(P_a)_{\mu_a} = \mathbf{J}_a^{-1}(\mu_a)/(G_a)_{\mu_a}$.

Some comments are in order before we proceed. We need to check that indeed the group G_a acts on the reduced space P_a and that it has a momentum map $\mathbf{J}_a$. We shall do this in a series of steps.

Lemma 5.2.1 *The group G_a leaves the set $\mathbf{J}_V^{-1}(a)$ invariant.*

Proof. First note that the group G_a leaves the set $\mathbf{J}_V^{-1}(a)$ invariant. Indeed, suppose that $\mathbf{J}_V(z) = a$ and that $g \in G$ leaves a invariant. Then by the equivariance relation noted above, we have $\mathbf{J}_V(gz) = g\mathbf{J}_V(z) = ga = a$. Thus, G_a acts on the set $\mathbf{J}_V^{-1}(a)$. We denote this action by $\Psi_g^a : \mathbf{J}_V^{-1}(a) \to \mathbf{J}_V^{-1}(a)$. $\blacksquare$

Lemma 5.2.2 *The action Ψ^a of G_a on $\mathbf{J}_V^{-1}(a)$ induces an action Ψ_a on the quotient space $\mathbf{J}_V^{-1}(a)/V$.*

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

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

Thus, the action Ψ_a of G_a on the V-reduced space P_a is well defined. The action of a group element $g \in G_a$ will be denoted by $\Psi_{g,a} : P_a \to P_a$. $\blacksquare$

Lemma 5.2.3 *The action Ψ_a on the quotient space $\mathbf{J}_V^{-1}(a)/V$ is symplectic.*

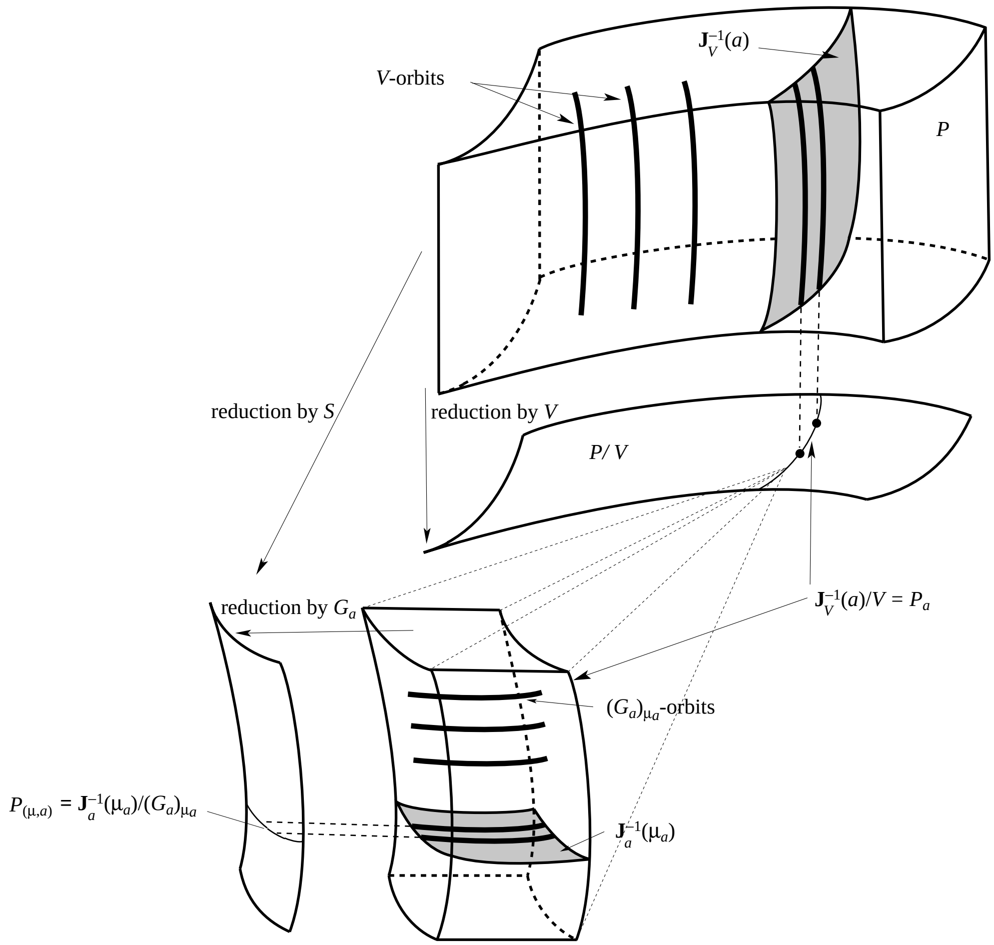

Figure 5.2.1: A schematic of reduction by stages for semidirect products.

Proof. Letting $\pi_a : \mathbf{J}_V^{-1}(a) \to P_a$ denote the natural projection and $i_a : \mathbf{J}_V^{-1}(a) \to P$ be the inclusion. By construction, $\Psi_{g,a} \circ \pi_a = \pi_a \circ \Psi_g^a$ and $\Psi_g \circ i_a = i_\nu \circ \Psi_g^a$, where $\Psi_g : P \to P$ denotes the action of $g \in G$. Recall also from the reduction theorem that $i_a^* \Omega = \pi_a^* \Omega_a$. Therefore,

$$\pi_a^* \Psi_{g,a}^* \Omega_a = (\Psi_g^a)^* \pi_a^* \Omega_a = (\Psi_g^a)^* i_a^* \Omega = i_a^* \Psi_g^* \Omega = i_a^* \Omega = \pi_a^* \Omega_a.$$

Since π_a is a surjective submersion, we may conclude that

$$\Psi_{g,a}^* \Omega_a = \Omega_a.$$

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

Lemma 5.2.4 *The symplectic action Ψ_a on the quotient space $\mathbf{J}_V^{-1}(a)/V$ has an equivariant momentum map.*

Proof. We first show that the momentum map of the G action restricted to $\mathfrak{g}_a$, namely $\mathbf{J}_S$ projected to $\mathfrak{g}_a^*$ induces a well defined map of P_a to $\mathfrak{g}_a^*$. First of all, we restrict $\mathbf{J}_S$ to the set $\mathbf{J}_V^{-1}(a)$ and project it to $\mathfrak{g}_a^*$. We claim that this map drops to the quotient space. To check this, note that for $z \in \mathbf{J}_V^{-1}(a)$, and $\xi \in \mathfrak{g}_a$, equivariance gives us

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

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

$$\begin{aligned} \langle \mathbf{J}_S(vz), \xi \rangle &= \left\langle \mathbf{J}_S(z), (e,v)^{-1}(\xi,0) \right\rangle = \langle \mathbf{J}_S(z), (\xi, \xi v) \rangle \\ &= \langle \mathbf{J}_G(z), \xi \rangle + \langle \mathbf{J}_V(z), \xi v \rangle = \langle \mathbf{J}_G(z), \xi \rangle - \langle \xi a, v \rangle = \langle \mathbf{J}_G(z), \xi \rangle . \end{aligned}$$

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

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

for $\xi \in \mathfrak{g}_a$ is well defined. 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.

To show that the map $\mathbf{J}_a$ is the momentum map, we first note that for all $\xi \in \mathfrak{g}_a$, the vector fields $\xi_P|(\mathbf{J}_a^{-1}(a))$ and ξ_{P_a} are π_a-related. Thus,

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

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

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

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 relation $\mathbf{J}_a \circ \pi_a = \iota_a^* \circ \mathbf{J}_G \circ i_a$ and the relations between the actions of G on P, $\mathbf{J}_V^{-1}(a)$ and on P_a. $\blacksquare$

Having established these preliminary facts, we can state the main reduction by stages theorem for semidirect products.

Theorem 5.2.5 (Reduction by Stages for Semidirect Products) *The reduced space $(P_a)_{\mu_a}$ is symplectically diffeomorphic to the reduced space P_σ obtained by reducing P by S at the point $\sigma = (\mu, a)$.*

Proof. Start with the natural 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 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

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

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

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

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

The map $\pi_a \circ j$ is equivariant with respect to the action of S_σ on the domain and $(G_a)_{\mu_a}$ on the range via the homomorphism ψ. Thus, $\pi_a \circ j$ induces a 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 is a diffeomorphism by finding an inverse. We begin with the construction of a map

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

To do this, we first choose an equivalence class $[p]_a \in \mathbf{J}_a^{-1}(\mu_a) \subset P_a$ for $p \in \mathbf{J}_V^{-1}(a)$. The equivalence relation is that associated with the map π_a; that is, with the action of V. For each such point, we consider a new point vp and will choose v such that $vp \in \mathbf{J}_S^{-1}(\sigma)$. For this to hold, we must have

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

By equivariance, the right hand side equals

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

Thus, we require that

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

This follows from the next lemma.

Lemma 5.2.6 *Denoting the annihilator of $\mathfrak{g}_a$ by $\mathfrak{g}_a^o$, we have*

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

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

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

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

[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 Fredholm type alternative arguments. In the finite dimensional case, the result may be proved by a dimension count.

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

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

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

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

The above argument shows how to construct v so that $vp \in \mathbf{J}_S^{-1}(\sigma)$. We continue with the definition of the map ϕ by mapping vp to $[vp]_\sigma$, its S_σ-equivalence class in P_σ.

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

Next we must show that the map ϕ is invariant under $(G_a)_{\mu_a}$. Thus, let $[p]_a \in \mathbf{J}_a^{-1}(\mu_a)$ and let $g_0 \in (G_a)_{\mu_a}$. Let v be chosen so that $vp \in \mathbf{J}_S^{-1}(\sigma)$ and let u be chosen so that $ug_0 p \in \mathbf{J}_S^{-1}(\sigma)$. We must show that $[vp]_\sigma = [ug_0 p]_\sigma$. In other words, we must find a $(g, w) \in S_\sigma$ such that $(g, w)(e, v)p = (e, u)(g_0, 0)p$. This will hold if we can show that $(g, w) := (e, u)(g_0, 0)(e, v)^{-1} \in S_\sigma$. But we know that $\sigma = \mathbf{J}_S(vp) = \mathbf{J}_S(ug_0 p)$ or, in other words, by equivariance, $\sigma = (e, v)\mathbf{J}_S(p) = (e, u)(g_0, 0)\mathbf{J}_S(p)$. By isolating $\mathbf{J}_S(p)$, this implies that $(e, v)^{-1}\sigma = (g_0, 0)^{-1}(e, u)^{-1}\sigma$ which means that our (g, w) is indeed in S_σ. Hence ϕ is invariant, and so gives a well defined map

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

Chasing the definitions shows that $[\phi]$ is the inverse of the map $[\pi_\nu \circ j]$. Thus, either is a symplectic diffeomorphism. ■

In this framework, one can also reduce the dynamics of a given invariant Hamiltonian; we will discuss this in a later section.

In the preceding theorem, choose $P = T^*S$ where $S = G \circledS V$ is a semidirect product as above, with the cotangent action of S on T^*S induced by left translations of S on itself. Reducing T^*S by the action of V gives a space naturally isomorphic to T^*G. Thus, the reduction by stages theorem gives as a corollary, the semidirect product reduction theorem, 5.1.1. The original proof of this result in Marsden, Ratiu and Weinstein [1984ab] essentially used the map $[\phi]$ constructed above to obtain the required symplectic diffeomorphism. However, the generalization here to get semidirect product reduction by stages requires an essential modification of the original method.

References

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

Abraham, R., J.E. Marsden, and T.S. Ratiu [1988] *Manifolds, Tensor Analysis, and Applications.* Second Edition, Applied Mathematical Sciences **75**, Springer-Verlag.

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

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

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

Arnold, V. I. [1989] Mathematical Methods of Classical Mechanics. Second Edition Graduate Texts in Math **60**, Springer-Verlag.

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.

Bloch, A.M., P.S. Krishnaprasad, J.E. Marsden, and T.S. Ratiu [1994] Dissipation Induced Instabilities, *Ann. Inst. H. Poincaré, Analyse Nonlineare* **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.

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 (to appear).*

Cendra, H., D.D. Holm, J. E. Marsden and T.S. Ratiu [1998] Lagrangian Reduction, the Euler–Poincaré Equations, and Semidirect Products. *Arnold volume II, AMS (to appear).*

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 and T.S. Ratiu [1998] Lagrangian reduction by stages. *preprint.*

Cendra, H., J. E. Marsden and T.S. Ratiu [1997] Lagrangian Reduction, the Euler–Poincaré Equations, and Semidirect Products. *To appear in the AMS Arnold volume II.*

Chetayev, N.G. [1941] On the equations of Poincaré. *J. Appl. Math. Mech.* **5**, 253–262

Chetayev, N.G. [1961] *The Stability of Motion.* Pergamon.

Chetayev, N.G. [1989] *Theoretical Mechanics.* Springer–Verlag.

Gotay, M., J. Isenberg, and J.E. Marsden [1997] *Momentum Maps and the Hamiltonian Structure of Classical Relativistic Field Theories, I.* preprint.

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

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

Holm, D. D., Marsden, J. E., Ratiu, T. and Weinstein, A. [1983] Nonlinear stability conditions and priori estimates for barotropic hydrodynamics. *Phys. Lett.* **98A**, 15–21.

Holm, D. D., Marsden, J. E., Ratiu, T. and Weinstein, A. [1984] Stability of rigid body motion using the energy-Casimir method. *Cont. Math. AMS* **28**, 15–23.

Holm, D.D., J.E. Marsden, T.S. Ratiu, and A. Weinstein [1985] Nonlinear stability of fluid and plasma equilibria, *Phys. Rep.* **123**, 1–116.

Libermann, P. and C.M. Marle [1987] *Symplectic Geometry and Analytical Mechanics.* Kluwer Academic Publishers.

Marsden, J.E. [1981] *Lectures on Geometric Methods in Mathematical Physics.* SIAM, Philadelphia, PA.

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

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

Marsden, J.E., G. Misiolek, M. Perlmutter and T.S. Ratiu [1998] Symplectic Reduction for Semidirect Products and Central Extensions, *Diff. Geometry and Appl. (to appear).*

Marsden, J.E., G. Misiolek, M. Perlmutter and T.S. Ratiu [1998] Reduction by stages and group extensions, *in preparation*.

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

Marsden, J. E., G. W. Patrick, and W. F. Shadwick [1996] *Integration Algorithms and Classical Mechanics.* Fields Institute Communications, **10**, Am. Math. Society.

Marsden, J.E., G.W. Patrick and S. Shkoller [1997] Multisymplectic Geometry, Variational Integrators, and Nonlinear PDEs. *preprint.*

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 [1998] *Introduction to Mechanics and Symmetry.* Texts in Applied Mathematics, **17**, Second Edition, Springer-Verlag.

Marsden, J.E., T.S. Ratiu, and G. Raugel [1991] Symplectic connections and the linearization of Hamiltonian systems, *Proc. Roy. Soc. Ed. A* **117**, 329-380

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

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

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

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

Marsden, J.E. and J. Scheurle [1995] Pattern evocation and geometric phases in mechanical systems with symmetry, *Dyn. and Stab. of Systems.* **10**, 315–338.

Marsden, J.E., J. Scheurle and J. Wendlandt [1996] Visualization of orbits and pattern evocation for the double spherical pendulum, *ICIAM 95: Mathematical Research, Academie Verlag, Ed. by K. Kirchgässner, O. Mahrenholtz and R. Mennicken* **87**, 213–232.

Marsden, J.E. and S. Shkoller [1997] Multisymplectic geometry, covariant Hamiltonians and water waves *Math. Proc. Camb. Phil. Soc., to appear.*

Marsden, J.E., and J.C. Simo [1990] The energy-momentum method, *La "Mecanique Analytique" de Lagrange et son Héritage, Atti della Accademia delle Scienze di Torino*, **124**, 245–268.

Marsden, J.E., J.C. Simo, D.R. Lewis, and T.A. Posbergh [1989] A block diagonalization theorem in the energy momentum method, *Cont. Math. AMS* **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., A. Weinstein, T.S. Ratiu, R. Schmid, and R.G. Spencer [1983] Hamiltonian systems with symmetry, coadjoint orbits and plasma physics, in Proc. IUTAM-IS1MM Symposium on *Modern Developments in Analytical Mechanics*, Torino 1982, *Atti della Acad. della Sc. di Torino* **117**, 289–340.

Marsden, J.E. and J.M. Wendlandt [1997] Mechanical systems with symmetry, variational principles and integration algorithms. *Current and Future Directions in Applied Mathematics*, Edited by M. Alber, B. Hu, and J. Rosenthal, Birkhäuser, 219–261.

McDuff, D. and D. Salamon [1995] *Introduction to Symplectic Topology*, Oxford University Press.

Poincaré, H. [1885] Sur l'équilibre d'une masse fluide animée d'un mouvement de rotation, *Acta. Math.* **7**, 259.

Poincaré, H. [1892–1899], *Les Méthodes Nouvelles de la Mécanique Celeste.* 3 volumes. English translation *New Methods of Celestial Mechanics.* History of Modern Physics and Astronomy **13**, Amer. Inst. Phys., 1993.

Poincaré, H. [1890] *Théorie des tourbillons*, Reprinted by Éditions Jacques Gabay, Paris.

Poincaré, H. [1890] Sur le problème des trois corps et les équations de la dynamique, *Acta Math.* **13**, 1–271.

Poincaré, H. [1892] Les formes d'équilibre d'une masse fluide en rotation, *Revue Générale des Sciences* **3**, 809–815.

Poincaré, H. [1901a] Sur la stabilité de l'équilibre des figures piriformes affectées par une masse fluide en rotation, *Philosophical Transactions A* **198**, 333–373.

Poincaré, H. [1901b] Sur une forme nouvelle des équations de la méchanique, *C.R. Acad. Sci.* **132**, 369–371.

Poincaré, H. [1910] Sur la precession des corps deformables. *Bull Astron* **27**, 321–356.

Woodhouse, N.M.J. [1992] *Geometric Quantization.* Clarendon Press, Oxford University Press, 1980, Second Edition, 1992.