Nonholonomic Euler-Poincaré Equations and Stability in Chaplygin's Sphere

David Schneider

A dissertation submitted in partial fulfillment of
the requirements for the degree of

Doctor of Philosophy

University of Washington

2000

Program Authorized to Offer Degree: Mathematics

UMI Number: 9983541

Bell & Howell Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

University of Washington

Graduate School

This is to certify that I have examined this copy of a doctoral dissertation by

David Schneider

and have found that it is complete and satisfactory in all respects, and that any and all revisions required by the final examining committee have been made.

Chair of Supervisory Committee:

C. Robin Graham

Reading Committee:

C. Robin Graham

Judith Arms

Tom Duchamp

Date: August 15, 2000

University of Washington

Abstract

Nonholonomic Euler-Poincaré Equations and Stability in Chaplygin's Sphere

by David Schneider

Chair of Supervisory Committee

Professor C. Robin Graham
Department of Mathematics

A method of reducing several classes of nonholonomic mechanical systems that are defined on semidirect products of Lie groups is developed. The method reduces the Lagrange-d'Alembert principle to obtain a reduced constrained principle that determines Euler-Poincaré equations on the reduced space. The theory is demonstrated by deriving the reduced equations of motion for several examples of rigid bodies that roll without slipping. This method of reduction is a generalization to the nonholonomic setting of the method developed in [CHMR] which reduces Hamilton's principle and derives Euler-Poincaré equations for a class of unconstrained systems on Lie Groups.

Our method of reduction is then used as a framework for obtaining several results related to a particular nonholonomic system: Chaplygin's sphere. Chaplygin's sphere is a ball that rolls without slipping on a horizontal plane and has a nonhomogeneous mass distribution. The principal moments of inertia of the ball need not be equal, however the ball's center of mass coincides with its geometric center. The first result determines the stability of the relative equilibria of Chaplygin's sphere: the ball spins stably about its long and short axis and unstably about the middle axis. The next results use two different approaches to stabilize the rotation of the ball about its middle axis by making use of the idea of controlled

Lagrangians introduced in [BLMa]. The first approach controls an internal rotor that has been added to the ball and generalizes the solution presented in [BLMa] to the analogous problem of using an internal rotor to stabilize a free rigid body. The second approach controls the plane the ball rolls upon by forces of horizontal translations. For the system in which the plane is allowed to react to the motion of the ball, we derive the equations of motion, identify the relative equilibria and consider the problem of stabilizing the ball by controlling the plane with forces of horizontal translations.

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGMENTS

I wish to thank several people for their guidance and support during this undertaking. First, thanks to Darryl Holm and Jerry Marsden for suggesting several problems on stability and stabilization in Chaplygin's sphere upon which much of the latter half of this dissertation is based. Special thanks go to Darryl Holm for teaching us Euler-Poincaré equations during his joint course with Tudor Ratiu on Hamiltonian fluid dynamics at UC Santa Cruz during the winter quarter of 1997. The deepest thanks and appreciation go to my advisor, Robin Graham, for the many hours he spent going over my work with me and for his high standards which continually bring forth the best effort from his students. Lastly, I wish to thank my father and family for their never-ending support and pride in my work.

Chapter 1

INTRODUCTION

Constraints which restrict the possible configurations for a mechanical system are called holonomic. Examples are adding rigidity to a system, or restricting a particle in three space to move on a sphere. Constraints on the velocities which cannot be reduced to holonomic constraints are termed nonholonomic. These often arise in rolling systems, such as a ball that rolls without slipping on a surface.

Nonholonomic mechanics has a rich history that dates back to the time of Euler and Lagrange. Understanding the dynamics of physical systems has motivated a great deal of research. Some early work on the problem of a rigid body rolling on a surface can be found in [Ro]. An account of numerous investigations of the same topic is found in [NF]. There is still active research in nonholonomic mechanics. Some recent interest is in developing control applications, such as feedback laws to stabilize or generate locomotion for a given system.

Many problems in nonholonomic mechanics admit some symmetry which can greatly simplify the analysis. Developing methods for effectively using symmetry has continued to be a topic for research. Symmetry is an important component in reduction, a process which allows one to study the dynamics of a mechanical system on a smaller space.

The first results of this dissertation develop a method of reduction for several families of nonholonomic systems with symmetry that are defined on semidirect products. The method proceeds by reducing the Lagrange-d'Alembert principle, the standard principle for determining the equations of motion for nonholonomic systems, and using the group structure to obtain Euler-Poincaré equations on the reduced space. The Euler-Poincaré equations are formulated on a vector space and have a simple and succinct form. The next results pertain to a specific nonholonomic system, Chaplygin's sphere. The method of reduction that we develop is used as a unifying framework to solve several problems pertaining to the system: namely determining the stability of relative equilibria and stabilizing the unstable relative equilibria. These problems, including our formulation of the stabilization problems, are discussed in the latter part of this introduction.

The method of reduction that we develop generalizes work in [CHMR] for reducing unconstrained mechanical systems on Lie groups. The relevant theorem in [CHMR] is a generalization of a simpler case discussed in [MR]. To discuss this simpler case, let G be a Lie group, $\mathfrak{g}$ its Lie algebra, TG the tangent bundle of G and $L : TG \to \mathbb{R}$ the Lagrangian for a mechanical system with configuration space G. Assume that L is invariant under the lift of the left action of G onto TG (an analogous theory holds for right actions). The content of the theorem of [MR] is that Hamilton's variational principle for determining the Euler-Lagrange equations on TG is equivalent to a reduced principle on $\mathfrak{g}$ which leads to reduced equations called the pure Euler-Poincaré equations.

To state the pure Euler-Poincaré equations, let $l : \mathfrak{g} \to \mathbb{R}$ be the restriction of L to $\mathfrak{g} = T_eG$. Denote a vector in $\mathfrak{g}$ by ξ. The differential of l at ξ is denoted by $\partial l/\partial\xi(\xi) \in \mathfrak{g}^*$. Let $ad^*_\xi : \mathfrak{g}^* \to \mathfrak{g}^*$ denote the dual to the adjoint action on $\mathfrak{g}$, where $ad_\xi(\eta) = [\xi, \eta]$ for $\eta \in \mathfrak{g}$. Then the pure Euler-Poincaré equations are

$$\frac{d}{dt}\frac{\partial l}{\partial \xi} - ad^*_\xi \frac{\partial l}{\partial \xi} = 0. \tag{1.1}$$

These are differential equations for a curve $\xi(t) \in \mathfrak{g}$ which typically describes the evolution of the velocity for the system (with respect to a specific frame).

The canonical example used to illustrate this reduction theory is the free rigid body. In this case, the Lie algebra is $\mathfrak{so}(3)$, which we identify with $\mathbb{R}^3$ with the cross product as Lie bracket. Denote vectors in $\mathbb{R}^3$ by Ω and let I denote the inertia tensor for the body. Euler's equations for the evolution of the angular velocity relative to a body frame are

$$I\dot{\Omega} + \Omega \times I\Omega = 0. \tag{1.2}$$

These equations take the form of (1.1) and are a special case of the theorem in [MR] discussed above.

This process of Lagrangian reduction is developed in [CHMR] to apply to a class of systems whose configuration space is a Lie group G but for which the G-invariance has been 'broken' by an advected parameter. (An advected parameter is a quantity, typically a vector, that is constant in the inertial frame.) To describe these systems precisely, we begin with a left representation of a Lie group G on a vector space V and the left action of G on $TG \times V^*$ given by the product of the lifted action of G on TG and the action of G on V^*. Assume that $\bar{L} : TG \times V^* \to \mathbb{R}$ is G-invariant. Let a_0 be a fixed element of V^* and denote an element of TG by the pair $(g, \dot{g})$ where $g \in G$ and $\dot{g} \in T_gG$. The Lagrangian $L : TG \to \mathbb{R}$ is defined by $L(g, \dot{g}) = \bar{L}(g, \dot{g}, a_0)$. Then L is invariant under the lift to TG of the left action of G_{a_0} on G, where G_{a_0} is the isotropy subgroup of a_0. The reduced Lagrangian $l : \mathfrak{g} \times V^* \to \mathbb{R}$ is defined by $l(\xi, a) = \bar{L}(e, \xi, a)$. The reduction theorem proceeds by reducing Hamilton's principle of least action for L on TG to a constrained principle for l on $\mathfrak{g} \times V^*$. The resulting reduced equations are for a curve $(\xi(t), \Gamma(t))$ in $\mathfrak{g} \times V^*$ where $\Gamma(t) = g^{-1}(t)a_0$ (and $g(t)$ is the curve in G describing the motion of the system). Additional notation in the reduced equations, also to be used later on, is as follows. For $v \in V$ and $\Gamma \in V^*$, let ξv and $\xi\Gamma$ denote the induced actions of $\mathfrak{g}$ on V and V^* respectively, where these induced actions refer to the map $\mathfrak{g} \to \mathfrak{gl}(V)$ induced by the left representation of G on V. Let $\rho_v : \mathfrak{g} \to V$ denote the linear map

$$\rho_v(\xi) = \xi v \text{ and } \rho_v^* : V^* \to \mathfrak{g}^* \text{ its dual.} \tag{1.3}$$

For notational convenience, let

$$v \diamond \Gamma \in \mathfrak{g}^* \text{ denote } \rho_v^*\Gamma. \tag{1.4}$$

The reduced equations of motion are:

$$\frac{d}{dt}\frac{\partial l}{\partial \xi} - ad_\xi^* \frac{\partial l}{\partial \xi} = \frac{\partial l}{\partial \Gamma} \diamond \Gamma \tag{1.5}$$

$$\dot{\Gamma} + \xi\Gamma = 0. \tag{1.6}$$

The first equation takes the form of an Euler-Poincaré equation (which is not pure, as the right hand side is nonzero). The second is an advection equation, giving the evolution of a_0

in a frame determined by g^{-1}. The right-invariant version of this theorem covers a number of examples in fluid mechanics (see [HMR]).

An illustrative example of the above theorem is the heavy top. This is a rigid body rotating about a fixed point while under the influence of gravity. The equations are described in a frame carried with the body and centered at the point of rotation. As with the free rigid body, let Ω denote the body's angular velocity and I the inertia tensor (which is constant in the body frame). Let χ denote the fixed vector in the body frame going from the origin to the center of mass. Let e_3 denote the vertical direction in the inertial frame, so that $-e_3$ is the direction of gravity. Let Γ denote e_3 in the body frame. That is, if the attitude of the body is given by a curve $A(t) \in SO(3)$, then $\Gamma = A(t)^{-1}e_3$. Finally, let m denote the mass of the body, and g the constant of acceleration due to gravity. Then the classical Euler-Poisson equations for the heavy top are:

$$I\dot{\Omega} = I\Omega \times \Omega + mg\Gamma \times \chi \tag{1.7}$$

$$\dot{\Gamma} = \Gamma \times \Omega, \tag{1.8}$$

and this is a special case of the above theorem found in [CHMR]. In comparison to the equations of motion on $TSO(3)$ one would obtain by introducing local coordinates for $SO(3)$ and applying Hamilton's principle, equations (1.7) and (1.8) are simpler to compute, more geometric as terms involve the cross product, and are defined globally on the vector space $\mathbb{R}^3 \times \mathbb{R}^3$.

As mentioned above, the first part of this dissertation develops a similar theory of reduction for several families of mechanical systems with linear nonholonomic constraints. A mechanical system with nonholonomic constraints consists of a smooth manifold Q which is the configuration space, a smooth function $L : TQ \to \mathbb{R}$ which is the Lagrangian (typically taken to be the difference between the kinetic energy and potential energy), and a smooth distribution $\mathcal{D} \subset TQ$ which determines the constraints. At a given point q of the configuration space, $\mathcal{D}_q$ parameterizes the allowable directions for the system. This implies constraints that are linear in the velocities. Constraints that are nonlinear in the velocities are unusual, but see [CKS] and [Pa] for studies of such systems. The problem of formulating equations of motion for nonholonomic systems has a history that starts during the time of Euler and Lagrange. The Lagrange-d'Alembert principle has been found to most accurately model physical nonholonomic systems. It states that motions of the system are curves $q(t) \in Q$ satisfying

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

where variations along the curve are taken to be in $\mathcal{D}$, and the curve itself satisfies the constraints $\dot{q} \in \mathcal{D}_{q(t)}$. Nonholonomic systems conserve energy (defined in terms of L and $\mathcal{D}$) similarly to mechanical systems without constraints.

Nonholonomic systems with symmetry may be reduced; following [BKMM], we give a summary of the reduction process. For a given nonholonomic system, let G be a Lie group which acts on the configuration space Q. Denote the lifted action to TQ by $g(q, \dot{q})$ where $g \in G$ and $(q, \dot{q}) \in TQ$. We say that the L and $\mathcal{D}$ are invariant under the group action if

$$L(q, \dot{q}) = L(g(q, \dot{q})) \text{ for all } (q, \dot{q}) \in TQ \text{ and } g \in G,$$

and whenever $(q, \dot{q})$ is in $\mathcal{D}_q$, one has that

$$g(q, \dot{q}) \text{ is in } \mathcal{D}_{gq}, \text{ for all } g \in G.$$

Assume that the action of G on TQ is free and proper, and $\mathcal{D}$ is invariant under the action—then it is known that the quotient spaces TQ/G and $\mathcal{D}/G$ are smooth manifolds. If L and $\mathcal{D}$ are both G-invariant, then the system can be reduced to $\mathcal{D}/G$. The form of the reduced equations on $\mathcal{D}/G$ depends on the method of reduction, of which there are many.

Our reduction theorems develop Euler-Poincaré equations for nonholonomic systems. We reduce families of nonholonomic systems where the configuration space is a semidirect product $S = G \circledS V$ (or an integral submanifold of $\mathcal{D}$ within S) and G is a Lie group with left representation space V. The systems depend on a fixed vector a_0 of V in a manner to be made precise in section 3.1 of Chapter 3. (The manner is similar to that of the [CHMR] case described above.) Let $\mathfrak{s}$ denote the Lie algebra of S and M the orbit space of a_0 under the action of G (M is a submanifold of V). The Lagrange-d'Alembert principle for L and $\mathcal{D}$ on TS is shown to be equivalent to a reduced principle on $\mathfrak{s} \times M$ which then leads to reduced equations on $\mathfrak{g} \times M$. These reduced equations have the same structure as equations (1.5) and (1.6), an Euler-Poincaré equation coupled to an advection equation. One advantage of writing the reduced equations this way is that they take a simple, vector form in terms of the Lie bracket on $\mathfrak{g}$ and action of $\mathfrak{g}$ on V. Another advantage is that they emphasize a momentum quantity which is crucial for understanding systems such as Chaplygin's sphere.

The theory is presented as three theorems: the Basic Theorem, Intermediate Theorem and General Theorem. The distribution in the basic and intermediate cases describes the kinds of constraints found in physical examples. The basic case assumes an S-invariant Lagrangian; the intermediate case assumes less symmetry. The general case assumes the minimal structure required for the reduction and consequently leads to reduced equations that are more complicated than the other two cases.

We provide a brief statement of the Basic Theorem and Intermediate Theorem here; the details are developed in Chapter 3. For the semidirect product $S = G \circledS V$, it is assumed that the Lie group G is a matrix subgroup of $GL(V)$ acting on V in the usual way. The induced action of $\mathfrak{g}$ on V is denoted by ξa for $\xi \in \mathfrak{g}$ and $a \in V$. Let a_0 be a fixed element of V. Denote a point of S by (g, x) with $g \in G$ and $x \in V$ and a point of $T_{(g,x)}S$ by $(\dot{g}, \dot{x})$, where $\dot{g} \in End(V)$. Denote the right and left pullback of tangent vectors in TG to $\mathfrak{g}$ by

$$\omega = \dot{g}g^{-1}, \qquad \xi = g^{-1}\dot{g}.$$

The system reduced by the Basic Theorem is the following. Assume $\bar{L} : TS \to \mathbb{R}$ is an S-invariant function and the distribution $\bar{\mathcal{D}} \subset TS$ is defined by the constraint equation

$$\dot{x} = \omega a_0. \tag{1.9}$$

This distribution generally has integral submanifolds. Our configuration space Q is taken to be one such submanifold. Let $W = \text{range}(\rho_{a_0})$ and take $Q = G \times W$, where the product is as topological spaces, not groups. In the example of a ball rolling on the plane, this corresponds to starting with $S = SE(3)$ while $Q = SO(3) \times \mathbb{R}^2$. The Lagrangian for the system $L : TQ \to \mathbb{R}$ is defined by $L = \bar{L} \mid_{TQ}$. As Q is an integral submanifold of $\bar{\mathcal{D}}$, we may define the distribution for the system by $\mathcal{D} = \bar{\mathcal{D}} \mid_Q \subset TQ$.

This system is reduced to $\mathfrak{g} \times M$. To state the reduced equations, let $l : \mathfrak{s} \to \mathbb{R}$ be the restriction of L to the identity; a vector in $\mathfrak{s}$ will be denoted (ξ, Y) where $\xi \in \mathfrak{g}$ and $Y \in V$.

Taking $\Gamma = g^{-1}a_0$, the reduced constrained Lagrangian, $l_c : \mathfrak{g} \times M \to \mathbb{R}$, is defined by

$$l_c(\xi, \Gamma) = l(\xi, \xi\Gamma);$$

one thinks of l_c as L evaluated on the constraints in the reduced space. The reduced equations are the following differential equations for a curve $(\xi(t), \Gamma(t)) \in \mathfrak{g} \times M$:

$$\frac{d}{dt}\frac{\partial l_c}{\partial \xi} - ad_\xi^* \frac{\partial l_c}{\partial \xi} = -\rho_{\xi\Gamma}^* \frac{\partial l}{\partial Y} \tag{1.10}$$
$$\dot{\Gamma} + \xi\,\Gamma = 0,$$

where $\partial l / \partial Y$ is evaluated at (ξ, Γ).

As previously mentioned, this theorem generalizes the theorem of [CHMR] discussed earlier to the nonholonomic setting. A small difference between the systems reduced is that a_0 has been taken to be in V as opposed to V^* as we've found this provides a more natural framework for describing the constraints found in physical examples.

A natural example of the Basic Theorem is Chaplygin's sphere. Chaplygin's sphere is a ball that is constrained to roll without slipping on a horizontal plane, and whose mass is distributed so that the center of mass is at the center of the ball, but the moments of inertia around the principal axes may differ from one another. In applying the Basic Theorem, S is taken to be $SE(3)$ and a_0 is taken to be e_3, the upward vertical normal to the horizontal plane. Then W is $\mathbb{R}^2$ and the configuration space is $SO(3) \times \mathbb{R}^2$. In the reduced equations, the right hand side of equation (1.10) is zero and one has

$$\dot{M} + \Omega \times M = 0, \qquad \dot{\Gamma} + \Omega \times \Gamma = 0,$$

where here M denotes the momentum vector for the ball in the reference coordinate system which is given by

$$M = \partial l_c / \partial \xi = I\Omega + mr^2 \Gamma \times (\Omega \times \Gamma). \tag{1.11}$$

These are Chaplygin's classical equations for Chaplygin's sphere. The basic theorem is also used to derive the equations of motion for the n dimensional formulation of Chaplygin's sphere.

An additional property of Chaplygin's sphere is that it conserves an $\mathfrak{so}(n)^*$ valued momentum map. Physically this measures the angular momentum of the ball about its contact point. This conservation of angular momentum is crucial for deriving analytical properties of the system. The general system reduced by the basic theorem may conserve an analogous momentum map. In Section 3.1 we will show that if $\rho^*_{\xi\Gamma}\partial l/\partial Y = 0$ in the reduced equation (1.10), then the momentum map $\mathbb{J} : TG \to \mathfrak{g}^*$ defined by $\mathbb{J}(g, \xi) = Ad^*_{g^{-1}}\partial l_c/\partial\xi$ is a conserved quantity for the reduced flow.

Our development of nonholonomic Euler-Poincaré equations continues with the Intermediate Theorem. The setting is the same as the basic case; one starts with a semidirect product $S = G \circledS V$ and a fixed element a_0 of V. The distribution is defined in terms of an arbitrary smooth function $s : V \to V$ by the equation

$$\mathcal{D} = \{(\dot{g}, \dot{x}) \in TS : \dot{x} = \omega g s(g^{-1}a_0)\}.$$

Taking s equal to the identity would recover the distribution defined in the basic case. The Lagrangian satisfies a symmetry condition similar to that of the Lagrangian in the reduction theorem of [CHMR] discussed above. This symmetry condition is described in terms of a function $\bar{L} : TS \times V \to \mathbb{R}$ assumed to be invariant under the action of S on $TS \times V$ given by the product of the action of S on S lifted to TS and the action of G on V. The Lagrangian for the system is taken to be $L = \bar{L}\mid_{a_0}$. Let G_{a_0} be the isotropy subgroup of a_0, then L is invariant under the group $G_{a_0} \circledS V$. As before, the system reduces to $\mathfrak{g} \times M$. The reduced Lagrangian $l : \mathfrak{s} \times M \to \mathbb{R}$ is defined by $l(\xi, Y, \Gamma) = \bar{L}(e, \xi, Y, \Gamma)$. The reduced constrained Lagrangian $l_c : \mathfrak{g} \times M \to \mathbb{R}$ is defined by $l_c(\xi, \Gamma) = l(\xi, \xi s(\Gamma), \Gamma)$. The final reduced equations are

$$\frac{d}{dt}\frac{\partial l_c}{\partial \xi} - ad^*_\xi \frac{\partial l_c}{\partial \xi} + \rho^*_\Gamma \frac{\partial l}{\partial \Gamma} = \rho^*_{\frac{d}{dt}s(\Gamma)} \frac{\partial l}{\partial Y}$$
$$\dot{\Gamma} + \xi\Gamma = 0,$$

where $\partial l/\partial\Gamma$ and $\partial l/\partial Y$ are evaluated at $Y = \xi s(\Gamma)$. Two examples are treated to illustrate this theorem. The first is Chaplygin's top. It differs from Chaplygin's sphere by allowing the center of mass for the ball to be at an arbitrary point within the ball. The second is the rolling disk, a homogeneous disk that rolls without slipping on the plane while under the influence of gravity.

There are numerous classical works that develop the equations of motion for mechanical systems with nonholonomic constraints. We mention several recent papers that, by using different methods, develop equations that are similar (and for some examples the same) as those obtained by our theorems. A class of systems called RL systems are studied in [VV]. These are nonholonomic systems whose configuration space is a group G, Lagrangian is invariant under the lift of the left action of the group on itself to TG and distribution is invariant under the lift of the right action of the group on itself to TG. For such systems, equations of motion are obtained on $\mathfrak{g}^* \times \mathfrak{g}^*$ which take the form of Euler-Poincaré equations. These systems conserve a momentum map defined similarly to the momentum map $\mathbb{J}$ discussed above. This is not a reduction theory as the space $\mathfrak{g}^* \times \mathfrak{g}^*$ is no smaller then the original velocity space TG. In [FK] and [A], a theorem is discussed which states that when the kinetic energy for a nonholonomic system is an invariant of the action of a group G and the action of the group is consistent with the constraints (meaning that the infinitesimal generator for the action corresponding to any $\xi \in \mathfrak{g}$ is in $\mathcal{D}$) then one obtains an equation for the evolution of a $\mathfrak{g}^*$ valued momentum map. This theorem is applied to Chaplygin's sphere and the equations of motion are the same as those we obtain with the Basic Theorem (Theorem 3.1). In [CHK] and [C], one finds another method of obtaining the equations of motion for the rolling disk and Routh's sphere (the special case of Chaplygin's top obtained when two moments of inertia are assumed equal) by starting with the Lagrange-d'Alembert principle and making use of Poincaré's calculation using quasi-velocities as outlined in [A]. The equations obtained by this method are the same as those obtained when our intermediate theorem is applied to these examples.

A general framework for computing the reduced equations of motion for a nonholonomic system with symmetry is developed in [BKMM]. Reduced equations of motion are explicitly calculated on $\mathcal{D}/G$ that emphasize the evolution of a momentum quantity that is based on a study of the geometry of the system. Studying the momentum equation in [BKMM] has led to interesting results in locomotion and optimal control, for example see [O] and [KM].

This is not the same momentum that is developed in our theory or the above mentioned works. In particular, the equations are not as well suited for studying Chaplygin's sphere.

When a mechanical system admits a symmetry group, it will often have solutions that correspond to one parameter subgroups. As is common in the literature, these solutions are called relative equilibria. A fundamental question to ask about the dynamics of a system with symmetry is whether or not the relative equilibria are stable. A relative equilibrium is an equilibrium solution to the reduced equations on the reduced space. We say a relative equilibrium is stable if the corresponding equilibrium solution in the reduced space is stable in the sense of Lyapunov. The framework one uses to identify the symmetry and derive the reduced equations of motion for a given system shapes much of the analysis involved in answering this question. Stability criteria for relative equilibria based upon the framework in [BKMM] are obtained in [ZBM]. However a general theorem about stability is not known when the eigenvalues of the operator obtained from the linearization of the reduced equations about a relative equilibrium are all imaginary. The usual way of proving stability in this case is to construct a Lyapunov function. However, this requires a sufficient number of conserved quantities.

The relative equilibria in Chaplygin's sphere are steady rotations about a vertical principal axis. A straightforward calculation shows that when the moment of inertia about this axis corresponds to the greatest or smallest of the moments of inertia, the eigenvalues of the linearized operator are all imaginary. The next result of the dissertation is to determine the stability of relative equilibria in Chaplygin's sphere. We study Chaplygin's sphere using the Basic Theorem discussed above. This framework illuminates the conservation of the angular momentum of the ball in space, and this conserved momentum leads to a sufficient number of conserved quantities for the construction of a Lyapunov function. The stability result is that if the moment of inertia around the vertical axis is the greatest or smallest of the three principal moments of inertia, then the motion is stable, while if the moment of inertia around the vertical axis is the middle principal moment of inertia, the motion is unstable.

Our final results consider two different means of stabilizing the rotation of Chaplygin's sphere about its middle axis. The first is to add an internal rotor to the ball which is controlled, and the second is to allow the plane the ball rolls upon to move horizontally and apply control forces to it. Our approach is based on the method of controlled Lagrangians developed in [BLM]. This method can produce a control force that conserves an energy like function for a given unconstrained mechanical system. This control force may then be used to construct a Lyapunov function in the usual manner in order to stabilize the system. Chapter 2 discuses an illustrative example of this method, the stabilization of an inverted pendulum on a cart by sliding the cart back and forth with a control force that is a function of the pendulum's position and velocity.

The method of controlled Lagrangians introduces a family of Lagrangians whose Euler-Lagrange equations match the form of the controlled equations for a given mechanical system. In order to construct these controlled Lagrangians, certain "matching" conditions must be satisfied. The construction outlined in [BLM] makes changes to the kinetic energy of the original system. The method may be generalized to the nonholonomic setting by constructing a new Lagrangian and distribution for which the Lagrange-d'Alembert equations match the controlled equations of the original system. However in adapting this method to stabilize Chaplygin's sphere, changing the distribution makes the matching process difficult, and considering modifications to the Lagrangian generally suffices.

Our solution to stabilizing Chaplygin's sphere with an internal rotor follows the solution to stabilizing the free rigid body with an internal rotor which is developed in [BLM]; making analogous modifications to the Lagrangian for Chaplygin's sphere allows us to match the controlled equations with an internal rotor. The resulting control law for Chaplygin's sphere is more involved then the one found for the free rigid body. This is due to the dependence on Γ of the momentum in Chaplygin's sphere, given by equation (1.11).

The second approach to stabilizing the rotation of Chaplygin's sphere about the middle axis is to control the horizontal plane that the ball rolls upon. Two cases are considered; we call the first Chaplygin's sphere on the translating plane, and the second Chaplygin's sphere on the sliding plane. In both cases the plane is controlled by forces that translate it horizontally. The plane is never tilted or rotated by the control force. For Chaplygin's sphere on the translating plane, we assume this is the only force acting on the plane, hence the plane will only move by translations. For Chaplygin's sphere on the sliding plane, the reaction force from the ball's rolling is allowed to affect the plane. This may torque, as well as translate the plane, causing it to rotate about the vertical axis (the plane still remains horizontal). For example, if the ball rolled on a platform which was held and controlled by a machine that is massive relative to the ball, Chaplygin's sphere on the translating plane would provide a good model. If the ball rolled on a platform of comparable mass which rested on a near frictionless surface like ice (and the control mechanism resides on the platform, perhaps in the form of rockets) then Chaplygin's sphere on the sliding plane would provide a good model. These problems become more interesting when the control force is based on measurements of the ball's orientation alone (as opposed to measurements of the ball's orientation and velocity). Using the idea of controlled Lagrangians, a suitable control law is found for Chaplygin's sphere on the translating plane; the modification made to the Lagrangian is to add a potential energy term which generates the control force.

To state the control force obtained to stabilize the translating plane, let the ball's orientation in space be given by a curve $A(t) \in SO(3)$, which maps from a reference coordinate system for the ball into an inertial coordinate system for space. Let E_3 denote the vertical principal axis of the ball in the reference coordinate system, and I_3 its moment of inertia. As we are stabilizing about the middle axis, assume I_3 is less then one of I_1, I_2 and greater then the other, where I_1, I_2 denote the moments of inertia about the other two principal axes. Let e_3 denote the vertical axis in space. Take m to be the mass of the ball, n the mass of the plane, and r the radius of the ball. The stabilizing control force, applied to the plane in the spatial frame, is

$$f = \frac{\kappa n}{m r} \left\{ AE_3 \times (e_3 \times AE_3) - \langle AE_3 \times (e_3 \times AE_3), e_3 \rangle e_3 \right\}. \tag{1.12}$$

Note that f is a horizontal vector. The feedback parameter κ is chosen so that $\kappa < Z^2(I_3 - I_1)$ and $\kappa < Z^2(I_3 - I_2)$, where Z is the speed of rotation at the equilibrium.

The assumption that the plane is affected by the reaction force of the ball makes Chaplygin's sphere on the sliding plane a much more complicated system than Chaplygin's sphere on the translating plane. The first problem is to obtain manageable equations of motion. The system lacks sufficient symmetry for a direct application of the Basic Theorem. However by generalizing the technique used to obtain the reduced equations for basic and intermediate systems, we obtain equations of motion that are similar to those obtained for Chaplygin's sphere. The equations of motion reveal additional steady rotations of interest, but we confine our attention to the unstable rotations observed in Chaplygin's sphere. Although we do not obtain a stabilizing control law on the sliding plane, we do find that the method used to stabilize Chaplygin's sphere on the translating plane does not generalize. That is, by adding a potential energy to the Lagrangian for Chaplygin's sphere on the sliding plane, one cannot generate a control law depending solely on the ball's orientation which conserves the associated energy and stabilizes the system.

One of the objectives of this work has been deriving equations of motion that are relatively simple to compute and lead to analytic information in a straightforward manner. The Euler-Poincaré equations that result from applying the Basic and Intermediate Theorems (Theorem 3.1 and Theorem 3.3) are stated succinctly on the vector space $\mathfrak{g} \times V$ in terms of the Lie bracket on $\mathfrak{g}$ and the action of $\mathfrak{g}$ on V. The examples will demonstrate that the theorems are well suited for a number of mechanical systems involving rolling rigid bodies. It proves to be a particularly useful framework for studying the stability and stabilization of relative equilibria in Chaplygin's sphere. This is not just because the conservation of angular momentum in the system is simple to deduce, but also because the framework applies to deriving the equations of motion in the stabilization problems. The uncontrolled equations for Chaplygin's sphere with an internal rotor are derived with the Basic Theorem by taking $S = G \circledS \mathbb{R}^3$ where the group G is $SO(3) \times S^1$. Physically G represents the orientation of the ball and the rotor angle. The uncontrolled equations of motion for Chaplygin's sphere on the sliding plane are derived by generalizing the technique used in the nonholonomic reduction theorems and taking $S = G \circledS \mathbb{R}^3$ as well, although in this problem the S^1 factor parameterizes the orientation of the plane. For both problems, the Euler-Poincaré component of the resulting equations of motion can be written in a particularly simple fashion:

$$\dot{M} + [\xi, M] = 0.$$

Here M corresponds to the $(\mathfrak{so}(3) \times \mathbb{R})^*$ valued momentum vector of the system in the body frame, see Section 5.1 and Section 5.3. The controlled equations for Chaplygin's sphere on the sliding plane are derived by starting with Newton's principles and then writing things in terms of the above Euler-Poincaré equation.

Another objective of our work has been stabilizing Chaplygin's sphere. The method of controlled Lagrangians was effectively generalized for the problems of stabilizing Chaplygin's sphere with an internal rotor and stabilizing Chaplygin's sphere on the translating plane. The stabilization problem on the sliding plane will require a different approach than matching the controlled equations through the addition of a potential energy term as was done with the translating plane. Developing the method of controlled Lagrangians and extending it to include controlled distributions to stabilize a more general class of nonholonomic systems will be an interesting problem for the future. Recent work in the area of controlled Lagrangians includes [BLMc], where the method is extended to include potential shaping for mechanical systems without constraints, and [ZBLM], where the matching techniques of controlled Lagrangians are applied to the nonholonomic system consisting of a unicycle with rider.

Chapter 2

BACKGROUND

The results of this dissertation build upon the Lagrangian reduction and resulting Euler-Poincaré equations developed in [CHMR] and the method of controlled Lagrangians presented in [BLMa, BLMb]; theories developed for mechanical systems without constraints. The summary of these topics in the introduction shows how our results, which are developed for mechanical systems with nonholonomic constraints, are related to this prior work. In this section, a more complete discussion of these topics is provided. A summary of the solution to stabilizing a rigid body by using controlled Lagrangians is given at the end of the section. Our solution to stabilizing Chaplygin's sphere with an internal rotor is modeled upon this solution.

2.1 Lagrangian Reduction and Euler-Poincaré Equations

When a mechanical system admits a symmetry group, it may generally be reduced, allowing one to study the dynamics on a smaller space. Reduction for systems in Hamiltonian form, which proceeds by reducing symplectic structures, has been a well developed theory for some time. A more recent approach, for systems in Lagrangian form, is to reduce Hamilton's variational principle of least action. When the configuration space itself is a Lie group, this often leads to Euler-Poincaré equations on the reduced space. We shall review this process and the resulting Euler-Poincaré equations for two different cases. We call these the pure case, which is covered in [MR], and the CHMR case, which is developed in [CHMR].

In what follows, let G denote a Lie group and TG the tangent bundle of G. The pure case applies to a system with configuration space G and Lagrangian $L : TG \rightarrow \mathbb{R}$ that is invariant under the lift of the left or right action of G on itself to TG. The CHMR case also takes the configuration space to be a Lie group, and has a left or right formulation of the invariance condition as well. Typically, the right-invariant version of the theory applies to infinite dimensional groups such as the group of diffeomorphisms of some region of space. Several examples of the right-invariant version of the CHMR case, examples that come from continuum mechanics involving fluid flow, are covered in [HMR]. For our purposes, it suffices to assume that G is a subgroup of $GL(n, V)$, where V is a vector space. This assumption will hold throughout in this paper.

To state the reduced Euler-Poincaré equations for the pure and CHMR cases, denote the Lie algebra of G by $\mathfrak{g}$, the dual to $\mathfrak{g}$ by $\mathfrak{g}^*$ and a vector in $\mathfrak{g}$ by ξ. Let $\partial l/\partial\xi(\xi) \in \mathfrak{g}^*$ denote the differential of a function $l : \mathfrak{g} \to \mathbb{R}$. The dual to the adjoint action on $\mathfrak{g}$ is denoted by $ad_\xi^* : \mathfrak{g}^* \to \mathfrak{g}^*$, where $ad_\xi(\eta) = [\xi, \eta]$ for $\eta \in \mathfrak{g}$. The pure case is the content of the following theorem.

THEOREM 2.1. *Let G be a Lie group and $L : TG \to \mathbb{R}$ a left (right) invariant Lagrangian. Let $l : \mathfrak{g} \to \mathbb{R}$ be the restriction of L to the tangent space at the identity. For a curve $g(t) \in G$, let $\xi(t) = g(t)^{-1}\dot{g}(t)$ (respectively $\xi(t) = \dot{g}(t)g(t)^{-1}$). Then the following are equivalent:*

1. Hamilton's principle

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

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

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

3. The "variational" principle

$$\delta \int_a^b l(\xi(t))dt = 0$$

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

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

where η vanishes at the endpoints, $+$ corresponds to left invariance, and $-$ to right invariance.

*4. The **pure Euler-Poincaré** equations hold:*

$$\frac{d}{dt}\frac{\delta l}{\delta\xi} = \pm ad_\xi^* \frac{\delta l}{\delta\xi}, \tag{2.1}$$

where + corresponds to left invariance and − to right invariance.

The proof of this theorem when G is a matrix group is found in [MR]. The case when G is a general Lie group is treated in [BKMR].

To illustrate the theorem, we derive the classical Euler equations for the motion of a rigid body that freely rotates about its center of mass. In this case the configuration space is the Lie group $SO(3)$. An element $A \in SO(3)$ determines the orientation of the body in space by mapping from a reference coordinate system for the body into an inertial coordinate system for space. Denote a vector in $TSO(3)$ by $(A, \dot{A})$ where $A \in SO(3)$ and $\dot{A} \in M^3(\mathbb{R})$ is obtained by differentiating a curve $A(t) \in SO(3)$ going through A. The angular velocity of the body in the reference coordinate system is denoted by $\Omega \in \mathbb{R}^3$. Ω is defined in terms of $(A, \dot{A}) \in TG$ by the formula $\widehat{\Omega} = A^{-1}\dot{A}$, where$\widehat{\ }$: $\mathbb{R}^3 \to \mathfrak{so}(3)$ denotes the usual isomorphism of the Lie algebras $\mathbb{R}^3$ with the cross product $(\times)$ and $\mathfrak{so}(3)$ with the usual Lie bracket. Denote the inertia tensor for the body by the symmetric matrix $I : \mathbb{R}^3 \to \mathbb{R}^3$. The Lagrangian is

$$L = \frac{1}{2}\langle I\Omega, \Omega\rangle.$$

The lift to $TSO(3)$ of left translation on $SO(3)$ is $B(A, \dot{A}) = (BA, B\dot{A})$. From this we see that L is left-invariant and is determined by its restriction to the identity. Hence Theorem 2.1 may be applied to obtain the following reduced equation on $\mathfrak{so}(3)$:

$$I\dot{\Omega} = I\Omega \times \Omega.$$

Equation (2.1) in Theorem 2.1 is called a pure Euler-Poincaré equation to distinguish it from the case of having additional terms added to the right hand side. The latter form was considered by Poincaré in 1901 [P]. The classical Euler-Poisson equations for the heavy top, the reduced equations of [CHMR], and our own reduced equations are of this more general form.

In [CHMR], the method of reducing variational principles to obtain Euler-Poincaré equations is developed for a class of systems for which the configuration space is also a Lie group. However the invariance property of the system differs from that found in the hypothesis of Theorem 2.1. To describe this class of systems, we begin with a left representation of a Lie group G on a vector space V, and a fixed element a_0 in V^*. Here G is the configuration space. Let $\bar{L} : TG \times V^* \to \mathbb{R}$ be a smooth function invariant under the action of G on $TG \times V^*$ consisting of the product of the lift of the left action of G on itself to TG and the action of G on V^*. The Lagrangian for the system, $L : TG \to \mathbb{R}$, is taken to be $\bar{L}$ evaluated at a_0. The invariance of $\bar{L}$ implies that L is invariant under the isotropy subgroup of a_0 in G. As with Theorem 2.1, curves satisfying Hamilton's principle for the system on TG are equivalent to curves satisfying a constrained principle for a reduced system. The reduced system is described on $\mathfrak{g} \times V^*$ and the reduced Lagrangian $l : \mathfrak{g} \times V^* \to \mathbb{R}$ is defined by $l(\xi, \Gamma) = \bar{L}(e, \xi, \Gamma)$. The identification between the curves $g(t) \in G$ and the curves $(\xi(t), \Gamma(t)) \in \mathfrak{g} \times V^*$ is that $\xi(t) = g(t)^{-1}\dot{g}(t)$ and $\Gamma(t)$ is defined as the unique solution of $\dot{\Gamma}(t) = -\xi(t)\Gamma(t)$ with initial condition $\Gamma(0) = a_0$. [1] The notation $\xi\Gamma$ for $\xi \in \mathfrak{g}$ and $\Gamma \in V$ denotes the induced action of $\mathfrak{g}$ on V^*. The CHMR case is the content of the following theorem.

THEOREM 2.2. *With the preceding notation, the following are equivalent:*

1. *With a_0 held fixed, Hamilton's variational principle*

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

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

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

3. *The constrained variational principle (of d'Alembert type)*

$$\delta \int_a^b l(\xi(t), \Gamma(t))dt = 0$$

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

$$\delta\xi = \dot{\eta} + [\xi, \eta], \quad \delta\Gamma = -\eta\Gamma.$$

where $\eta(t) \in \mathfrak{g}$ vanishes at the endpoints.

4. *The **Euler-Poincaré** equation holds on $\mathfrak{g} \times V^*$:*

$$\frac{d}{dt}\frac{\delta l}{\delta \xi} = ad^*_\xi \frac{\delta l}{\delta \xi} + \frac{\partial l}{\partial \Gamma} \diamond \Gamma \quad \textit{where } \dot{\Gamma} = -\,\xi\Gamma. \tag{2.2}$$

[1] This initial condition for Γ establishes an equivalence with curves $g(t)$ such that $g(0)a_0 = a_0$. Establishing equivalence to other curves is done by taking $\Gamma(0) = ga_0$ for appropriate $g \in G$.

The notation $v \diamond \Gamma$ for $v \in V$, $\Gamma \in V^*$ is as in equation (1.4) in the introduction.

To illustrate this theorem, we shall add to the discussion of the heavy top in the introduction by showing how the heavy top satisfies the requirement of Theorem 2.2. The configuration space is the Lie group $G = SO(3)$, the vector space V is $\mathbb{R}^3$ and the element a_0 in V^* is $e_3 \in \mathbb{R}^3$, represented as an element of $\mathbb{R}^{3^*}$ by identifying $\mathbb{R}^3$ and $\mathbb{R}^{3^*}$ with the Euclidean inner product. Recall that I denotes the inertia tensor for the top, m the mass of the top, g the constant acceleration due to gravity, and χ the vector in a reference coordinate system for the top going from the point about which the body rotates to the center of mass. Then with this notation, the Lagrangian $L : TSO(3) \to \mathbb{R}$ for the heavy top is

$$L(A, \Omega) = \frac{1}{2}\langle I\Omega, \Omega\rangle - mg\langle A\chi, e_3\rangle.$$

The first term is the kinetic energy due to rotation and the second term is the potential energy due to gravity. By letting e_3 vary, one obtains the function $\bar{L} : TG \times V^* \to \mathbb{R}$ described in the notation preceding Theorem 2.2. The action of G on $TG \times V^*$ takes the form $B(A, \Omega, e_3) = (BA, \Omega, Be_3)$ for $B \in SO(3)$ and $(A, \Omega, e_3) \in TSO(3) \times \mathbb{R}^{3^*}$. Noting that

$$\langle BA\chi, Be_3\rangle = \langle A\chi, B^T Be_3\rangle = \langle A\chi, e_3\rangle,$$

shows that $\bar{L}$ is invariant under this action, hence Theorem 2.2 applies. Working out equation (2.2) for the reduced Lagrangian

$$l = \frac{1}{2}\langle I\Omega, \Omega\rangle - mg\langle \chi, \Gamma\rangle,$$

gives the classical Euler-Poisson equations (1.7) for the heavy top.

2.2 *Controlled Lagrangians and Stabilization*

Our stabilization results draw upon the method of controlled Lagrangians presented in [BLMa] and [BLMb]. The method is developed for mechanical systems of the following form. Denote the configuration space of the system by Q and suppose Q takes the form $S \times G$ where S is a smooth manifold and G is a Lie group. Suppose the Lagrangian L is invariant under the action of G on Q where the action is on the G factor alone. One seeks to control the system by using variables in S. This is done by constructing a new Lagrangian, called the controlled Lagrangian, whose Euler-Lagrange equations match the form of the controlled equations for the original Lagrangian.

An example from [BLMa] will demonstrate the theory. The problem is to stabilize the inverted pendulum on a cart. The configuration space is $Q = \mathbb{R} \times S^1$. A point in Q is denoted by (s, θ) with $s \in \mathbb{R}$ being the cart position and $\theta \in S^1$ being the pendulum angle measured relative to the vertical axis. A tangent vector in TQ is denoted as $(s, \theta, \dot{s}, \dot{\theta})$. The following figure illustrates this parameterization and the constants in the system.

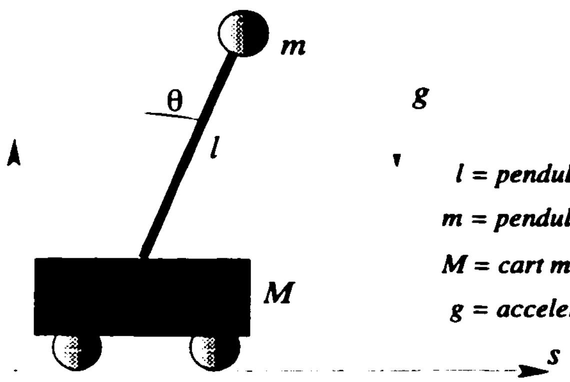

Figure 2.1: Inverted pendulum on a cart. From [BLMa], used with permission.

The Lagrangian is

$$L = \frac{1}{2}\left\{\alpha\dot{\theta}^2 + 2\beta\cos(\theta)\dot{s}\dot{\theta} + \gamma\dot{s}^2\right\} + D\cos(\theta),$$

where $\alpha = ml^2, \beta = ml, \gamma = M + m$ and $D = -mgl$. The Lagrangian is $\mathbb{R}$-invariant as s is a cyclic variable. One may readily check that an equilibrium solution to the equations of motion is when the pendulum stands upright $(\theta = 0, s = s_0)$ and that this solution is unstable. The problem is to find a control law $u(\theta, \dot{\theta})$ such that this equilibrium becomes a stable solution to the following controlled equations:

$$\begin{aligned} \frac{d}{dt}\frac{\partial L}{\partial \dot{s}} &= u \\ \frac{d}{dt}\frac{\partial L}{\partial \dot{\theta}} - \frac{\partial L}{\partial \theta} &= 0. \end{aligned} \tag{2.3}$$

Finding a control law u which induces linear stability is fairly straightforward. The problem solved by the method of controlled Lagrangians is to achieve nonlinear stability. A new Lagrangian, the controlled Lagrangian $L_{k,\sigma} : TQ \to \mathbb{R}$ is obtained by making certain changes to the kinetic energy of L, where k is a function of θ and σ is a scalar as described below. The hope is that the Euler-Lagrange equations corresponding to $L_{k,\sigma}$ will match those of the controlled equations (2.3) and that the control force u is a function only of θ and $\dot{\theta}$ which stabilizes the system. The changes are based on a study of the geometry of the system and has been successful in stabilizing several problems. We refer the reader to [BLMa] for a detailed discussion of these modifications. The new Lagrangian obtained for the inverted pendulum on the cart is

$$L_{k,\sigma} = \frac{1}{2}\left\{\alpha\dot{\theta}^2 + 2\beta\cos(\theta)(\dot{s} + k(\theta)\dot{\theta})\dot{\theta} + \gamma(\dot{s} + k(\theta)\dot{\theta})^2\right\} + \frac{\sigma}{2}\gamma k(\theta)^2\dot{\theta}^2 + D\cos(\theta),$$

where $k(\theta)$ is an undetermined function, and σ is some scalar. The next step in the method, called matching, is to find $k(\theta)$ and σ so that the Euler-Lagrange equations for $L_{k,\sigma}$ match the form of the controlled equations (2.3) for L. This is achieved by taking

$$k(\theta) = \kappa\cos(\theta), \quad \sigma = -\beta/\gamma\kappa,$$

where κ is some scalar. General conditions which guarantee a match are discussed in [BLMb]. As the controlled equations now come from the Lagrangian $L_{k,\sigma}$ the equations conserve the energy corresponding to $L_{k,\sigma}$. Further analysis, using the conserved energy, shows that the equilibrium is now stable when κ is chosen sufficiently positive. The final formula for the control law u is

$$u = -\frac{d}{dt}\left(\gamma k(\theta)\dot{\theta}\right),$$

where $\ddot{\theta}$ is then solved for in terms of θ and $\dot{\theta}$.

Next we show how the method of Controlled Lagrangians is used to stabilize a rigid body with an internal rotor, as discussed in [BLMb]. Our solution to stabilizing Chaplygin's sphere about its middle axis with an internal rotor is modeled after this result. Recall that Euler's equations for a rigid body, discussed above, are

$$I\dot{\Omega} = I\Omega \times \Omega.$$

The eigenvectors and eigenvalues of I are the principal axes and principal moments of inertia for the rigid body. The relative equilibria for the rigid body, which are equilibrium solutions to Euler's equations, occur when Ω coincides with a principal axis. If the moment of inertia for this axis is the middle value of the three principal moments of inertia, the motion is unstable. To stabilize it, a symmetric internal rotor is aligned with the principal axis of the body with the smallest moment of inertia (although choosing the largest would work as well). The problem is to find a control law for this rotor which stabilizes the rotation of the rigid body about its middle axis.

The configuration space for the rigid body with internal rotor is $Q = SO(3) \times S^1$. The first factor being the body's attitude and the second factor being the rotor angle, measured relative to a coordinate system for the body, denoted by θ. The Lagrangian for the system is

$$l = \frac{1}{2}\left(\lambda_1\Omega_1^2 + \lambda_2\Omega_2^2 + I_3\Omega_3^2 + J_3(\Omega_3 + \dot{\theta})^2\right),$$

where $I_1 > I_2 > I_3$ are the principal moments of inertia for the rigid body, J_1, J_2 and J_3 are the moment of inertia for the rotor, and $\lambda_i = I_i + J_i$. The controlled equations, with control torque u acting on the rotor, take the form

$$\lambda_1\dot{\Omega}_1 = \lambda_2\Omega_2\Omega_3 - (\lambda_3\Omega_3 + J_3\dot{\theta})\Omega_2 \tag{2.4}$$

$$\lambda_2\dot{\Omega}_2 = -\lambda_1\Omega_1\Omega_3 + (\lambda_3\Omega_3 + J_3\dot{\theta})\Omega_1 \tag{2.5}$$

$$\lambda_3\dot{\Omega}_3 + J_3\ddot{\theta} = (\lambda_1 - \lambda_2)\Omega_1\Omega_2 \tag{2.6}$$

$$\frac{d}{dt}\left(J_3(\Omega_3 + \dot{\theta})\right) = u. \tag{2.7}$$

As with the pendulum on the cart, changes are made to the kinetic energy to obtain a new Lagrangian, l_r, where r is a scalar. The matching problem here is different then with the inverted pendulum on a cart. One wants to match Euler-Poincaré equations for l_r to the controlled equations above. In [BLMb], conditions are established which guarantee a match for the Euler-Poincaré case. These conditions are met for the rigid body with internal rotor by taking the controlled Lagrangian to be

$$l_r = \frac{1}{2}\left(\lambda_1\Omega_1^2 + \lambda_2\Omega_2^2 + \lambda_3\Omega_3^2 + 2J_3\Omega_3\dot{\theta} + J_3\frac{r}{r-1}\dot{\theta}_3^2\right)$$

The resulting control law is $u = \frac{1}{r}J_3\dot{\Omega}_3$. Using equation (2.6), $\dot{\Omega}_3$ may be eliminated in u. By defining κ by $1/r = \kappa I_3/(1-\kappa)J_3$, one may write the control law as

$$u = \kappa(\lambda_1 - \lambda_2)\Omega_1\Omega_2.$$

Stability is then shown for the middle axis rotation $\Omega = (0, M, 0)$ when $\dot{\theta}(0) = 0$. This is done by first assuming the momentum conjugate to θ for l_r is 0, that is

$$J_3(\Omega_3 + \dot{\theta} - \frac{1}{r}\Omega_3) = 0.$$

Then one can eliminate $\dot{\theta}$ from equations (2.4) and (2.5) and follow the argument used for the free rigid body to show that equations (2.4) to (2.6) are stable in Ω when $\kappa > 1 - J_3/\lambda_2$.

Chapter 3

NONHOLONOMIC EULER-POINCARÉ EQUATIONS

This chapter presents the basic, intermediate and general nonholonomic Euler-Poincaré reduction theorems. These theorems develop the method of reducing variational principles on Lie groups to obtain Euler-Poincaré equations for several families of systems with nonholonomic constraints on semidirect products. As in [CHMR], we use an advected parameter to formulate the symmetry condition of the systems we reduce. Presenting the theory in three cases allows us to identify systems that are more common in practice as well as to describe the most general conditions required for the reduction. The basic theorem is illustrated by deriving the equations of motion for Chaplygin's sphere in three and n dimensions. The equations of motion for Chaplygin's top and the rolling disk are derived using the intermediate theorem. As will be demonstrated by the examples, the distributions in the systems reduced by the basic and intermediate theorems are abstractions of constraints found in physical systems.

This chapter also provides a discussion of energy and momentum in the systems we reduce. Both energy and momentum are important in the study of unconstrained mechanical systems. Energy is also conserved for mechanical systems with nonholonomic constraints (when the constraints are linear and homogenous). The form the energy takes in our setting is stated in Section 3.1 and a proof that it is conserved is given in Section 3.3. Momentum is generally not conserved for nonholonomic systems with symmetry. However there is an important special case in our theory when it is. This special case and the momentum it conserves will be identified in Section 3.1.2. An example of this special case is Chaplygin's sphere; the conservation of momentum is essential to proving our results in the next chapter.

3.1 Basic Case

We start with our notation and definitions for semidirect products, notation which is used in subsequent chapters as well. Next what we call the basic system is described and its symmetry group and reduced space are identified. The Basic Theorem for reducing these systems is then stated, along with the form of the conserved energy in the reduced variables. Following this is a discussion of the special case when momentum is conserved. Finally, considering Chaplygin's sphere in three and n dimensions provides examples of applying the Basic Theorem and the conservation of momentum.

3.1.1 Basic Theorem

Our notation for semidirect products starts with a Lie group G and a left representation space V for G. Let S denote the semidirect product $G \circledS V$. Topologically, S is $G \times V$. The group action on S is:

$$(g_1, y_1) \cdot (g_2, y_2) = (g_1 g_2, g_1 y_2 + y_1),$$

where for $g \in G$ and $y \in V$, the action of G on V is denoted as $g\,y$. Let $\mathfrak{g}$ denote the Lie algebra of G and $\mathfrak{s}$ the Lie algebra of S. As a vector space $\mathfrak{s}$ is $\mathfrak{g} \times V$. The Lie bracket on $\mathfrak{s}$ is:

$$[(\xi_1, Y_1), (\xi_2, Y_2)] = ([\xi_1, \xi_2], \xi_1 Y_2 - \xi_2 Y_1),$$

where for $\xi \in \mathfrak{g}$ and $Y \in V$, the induced action of $\mathfrak{g}$ on V is denoted as ξY. For fixed $a_0 \in V$, define the linear map $\rho_{a_0} : \mathfrak{g} \to V$ by

$$\rho_{a_0}(\xi) = \xi\, a_0,$$

and denote its dual as $\rho_{a_0}^* : V^* \to \mathfrak{g}^*$. For fixed $\xi \in \mathfrak{g}$, define the linear map $\sigma_\xi : V \to V$ by

$$\sigma_\xi(a) = \xi a,$$

with dual map $\sigma_\xi^* : V^* \to V^*$.

It shall be assumed that G can be represented as a subgroup of $GL(n, V)$. Therefore ga and ξa for $g \in G$, $\xi \in \mathfrak{g}$ and $a \in V$ may be represented by matrix multiplication, as may left and right translation of the tangent vectors in TG. In particular, ω and ξ shall represent the right and left pull back of a tangent vector $(g, \dot{g}) \in T_gG$ to $\mathfrak{g}$, that is:

$$\omega = \dot{g}g^{-1}, \qquad \xi = g^{-1}\dot{g}.$$

An element of S shall be denoted as (g, y) with g in G and y in V. A vector in $T_{(g,y)}S$ shall be denoted as $(\dot{g}, \dot{y})$.

The construction of the system reduced by the Basic Theorem starts with a semidirect product S, a smooth function $\bar{L} : TS \to \mathbb{R}$ which is invariant under the lift of the left action of S on itself to TS, and a fixed vector a_0 in V. Using the action of $\mathfrak{g}$ on V and a_0, define $\bar{\mathcal{D}} \subset TS$ to be:

$$\bar{\mathcal{D}}_{(g,y)} = \{(\dot{g}, \dot{y}) : \dot{y} = \omega\, a_0\}. \tag{3.1}$$

Let $W = \text{range } \rho_{a_0}$. The configuration space for the system is $Q = G \times W$. Assuming ρ_{a_0} is not onto, Q is an integral submanifold of $\bar{\mathcal{D}}$. For example, the configuration space for a ball rolling on a table is $Q = SO(3) \times \mathbb{R}^2$ even though $S = SE(3)$. The Lagrangian L and distribution $\mathcal{D}$ for the system are taken to be $L = \bar{L}\,|_{TQ}$ and $\mathcal{D} = \bar{\mathcal{D}}\,|_Q \subset TQ$. The system $(Q, L, \mathcal{D})$ we have constructed from S, a_0 and $\bar{L}$ will be referred to as the basic system.

The basic system is invariant under a group we now identify. Define H to be the isotropy subgroup of a_0 in G, and J to be the subgroup of S given by

$$J = H \circledS W.$$

To see that the product of two members of J is contained in J, it suffices to note that H maps W into W. Let $y \in W$ and $h \in H$. Then $y = \xi a_0$ for some $\xi \in \mathfrak{g}$. Hence

$$hy = h\xi a_0 = h\xi h^{-1} h a_0 = h\xi h^{-1} a_0 = \rho_{a_0} Ad_h \xi \in W.$$

Proposition 3.1. *The Lagrangian L and distribution $\mathcal{D}$ of the basic system are invariant under the action of J on TQ.*

Proof. As $\bar{L} : TS \to \mathbb{R}$ is S-invariant, it is certainly J-invariant as well. As $L = \bar{L}$ on TQ, and J maps TQ to TQ, L is J-invariant as well. To see that $\mathcal{D}$ is invariant, let $(\dot{g}, \dot{y}) \in \mathcal{D}_{(g,y)}$ and $(h, w) \in J$. Then one must show that $\dot{y} = \dot{g}g^{-1}a_0$ implies that $h\dot{y} = h\dot{g}(hg)^{-1}a_0$. This follows by applying h to both sides of $\dot{y} = \dot{g}g^{-1}a_0$ and noting that $ha_0 = a_0$. ☐

The reduced space for the basic system is the quotient space $\mathcal{D}/J$. Let $\pi_{\mathcal{D}} : \mathcal{D} \to \mathcal{D}/J$ denote the natural map taking each $d \in \mathcal{D}$ to its J-orbit. Let $\Phi : J \times \mathcal{D} \to \mathcal{D}$ denote the action of J on $\mathcal{D}$. We may conclude that $\mathcal{D}/J$ is a smooth manifold and that $\pi_{\mathcal{D}}$ is a submersion if Φ is a free and proper action; see [AM p. 266]. For the action to be free means that if $j\,d = d$ where $j \in J$ and $d \in \mathcal{D}$, then $j = e$. The action being proper means that the map $\tilde{\Phi} : J \times \mathcal{D} \to \mathcal{D} \times \mathcal{D}$ defined by $\tilde{\Phi}(j,d) = (d, j\,d)$ is proper: that is if $K \subset \mathcal{D} \times \mathcal{D}$ is compact, then $\tilde{\Phi}^{-1}(K)$ is compact. We will establish this in the next proposition where we also identify $\mathcal{D}/J$ with $\mathfrak{g} \times G/H$.

Denote the orbit space of a_0 in V by M, that is

$$M = \{y \in V : y = g\,a_0 \text{ for some } g \in G\}.$$

A well known fact of Lie theory implies that if M is closed in V, then G/H is diffeomorphic to M, a regular submanifold of V, see [V pp. 80-81]. Henceforth we assume that M is closed in V, as this is always the case in the examples of interest.

Proposition 3.2. *$\mathcal{D}/J$ is a smooth manifold diffeomorphic to $\mathfrak{g} \times M$.*

Proof. Following the remarks above, we first note that Φ is a free action, and next show that $\tilde{\Phi}$ is a proper map. An element of $\mathcal{D}$ will be denoted by $(g, y, \dot{g})$ as the $\dot{y}$ component is determined by the equation $\dot{y} = \omega a_0$. Establishing that $\tilde{\Phi}$ is a proper map is equivalent to showing that if $(g_n, y_n, \dot{g}_n)$ converges in $\mathcal{D}$, and $(h_n, w_n)(g_n, y_n, \dot{g}_n)$ converges in $\mathcal{D}$, where $(h_n, w_n) \in J$, then (h_n, w_n) has a convergent subsequence. Note that since Q is a regular submanifold of S, and as H is a closed Lie subgroup of G, one has that J is also a regular submanifold of S. Now suppose that in $\mathcal{D}$

$$(g_n, y_n, \dot{g}_n) \to (g_0, y_0, \dot{g}_0) \text{ and}$$
$$(h_n, w_n)(g_n, y_n, \dot{g}_n) \to (\tilde{g}_0, \tilde{y}_0, \dot{\tilde{g}}_0).$$

Because of the product action, this implies that $(h_n, w_n)(g_n, y_n) \to (\tilde{g}_0, \tilde{y}_0)$. In S, one has that

$$(h_n, w_n) \to (\tilde{g}_0, \tilde{y}_0)(g_0, y_0)^{-1}.$$

Since J is a regular submanifold of S, this implies (h_n, w_n) is a convergent sequence in J. Hence $\mathcal{D}/J$ is a smooth manifold and $\pi_{\mathcal{D}} : \mathcal{D} \to \mathcal{D}/J$ is a submersion.

To show $\mathcal{D}/J$ is diffeomorphic to $\mathfrak{g} \times M$, first let $\widetilde{F} : \mathcal{D} \to \mathfrak{g} \times V$ be defined by

$$\widetilde{F}(g, y, \dot{g}) = (g^{-1}\dot{g}, g^{-1}a_0).$$

Clearly $\widetilde{F}$ is smooth. Since M is a regular submanifold of V, $\widetilde{F}$ induces a smooth map $F : \mathcal{D} \to \mathfrak{g} \times M$ which is readily seen to be J-invariant. Hence F induces a smooth map $f : \mathcal{D}/J \to \mathfrak{g} \times M$. To see that f is one to one, suppose that $F(d_1) = F(d_2)$ where d_1 and d_2 are in $\mathcal{D}$ and have components $d_1 = (g_1, y_1, \dot{g}_1)$ and $d_2 = (g_2, y_2, \dot{g}_2)$. Let $h = g_2 g_1^{-1}$. We will show that $(h, y_2 - hy_1) \in J$ and that $(h, y_2 - hy_1)d_1 = d_2$, implying that f is one to one. As $F(d_1) = F(d_2)$, $g_1^{-1}a_0 = g_2^{-1}a_0$ and $g_1^{-1}\dot{g}_1 = g_2^{-1}\dot{g}_2$. The former equation implies that $h = g_2 g_1^{-1}$ is in H from which it follows that $(h, y_2 - hy_1) \in J$. The latter equations then implies that

$$(h, y_2 - hy_1)d_1 = d_2.$$

To show that f is onto, take $(\xi, \Gamma) \in \mathfrak{g} \times M$ and choose $(g, \dot{g}) \in TG$ such that $g^{-1}a_0 = \Gamma$ and $g^{-1}\dot{g} = \xi$. Then $F(g, 0, \dot{g}) = (\xi, \Gamma)$.

To see that f^{-1} is smooth, let $(\xi_0, \Gamma_0) \in \mathfrak{g} \times M$ and $g_0 \in G$ such that $g_0^{-1}a_0 = \Gamma_0$. As H is a closed Lie subgroup, G/H is a regular submanifold of G and one has a smooth section $\sigma : V \to \widetilde{U}$, where V is open in G/H, $\widetilde{U}$ is a neighborhood of g_0, and $\pi_G \circ \sigma$ is the identity on V (where π_G is the natural projection from G to G/H). Let (Γ^i) be coordinates on M (and G/H). Then σ may be used to determine coordinates (Γ^i, θ^j) on G such that for $g = (\Gamma^i, 0)$,

$$g^{-1}a_0 = (\Gamma^i). \tag{3.2}$$

We take $(\Gamma^i, \theta^j, y^k, \xi^a)$ to be local coordinates for $\mathcal{D}$ as follows. The components (Γ^i, θ^j) are defined as above, y^k are coordinates for W, and $\dot{g}$ is written as $\xi^a\partial_a$ where ∂_a is a left invariant frame on TG defined by $\partial_a = ge_a$ and e_a is a basis for $\mathfrak{g}$. Note that (ξ^a, Γ^i) provide local coordinates on a neighborhood U of $(\xi_0, \Gamma_0) \in \mathfrak{g} \times M$. Now define $r : U \to \mathcal{D}$ by

$$r(\xi, \Gamma^i) = (\Gamma^i, 0, 0, \xi^a).$$

Clearly r is smooth, hence $\pi_{\mathcal{D}} \circ r$ is smooth, so it suffices to show that $f^{-1} = \pi_{\mathcal{D}} \circ r$ on U. This follows from observing that $F \circ r$ is the identity as a consequence of equation (3.2) above. □

The basic system is reduced from TQ to $\mathfrak{g} \times M$ by first developing a reduced system on $\mathfrak{s} \times M$. On $\mathfrak{s} \times M$ we define a reduced Lagrangian, reduced constraint equation and a reduced constrained principle. From this the reduced equations of motion are obtained on $\mathfrak{g} \times M$ primarily in terms of the reduced constrained Lagrangian. We formulate this reduction process by showing that these principles for determining the curves describing the dynamics on Q and the reduced dynamics on $\mathfrak{g} \times M$ are equivalent to each other, given the following relations between the curves in Q and those in $\mathfrak{s} \times M$.

From $(g(t), y(t)) \in Q$ we obtain $(\xi(t), Y(t), \Gamma(t)) \in \mathfrak{s} \times M$ by setting

$$\xi(t) = g^{-1}(t)\dot{g}(t) \tag{3.3}$$

$$Y(t) = g^{-1}(t)\dot{y}(t) \tag{3.4}$$

$$\Gamma(t) = g^{-1}(t)a_0. \tag{3.5}$$

For a curve $(\xi(t), Y(t), \Gamma(t)) \in \mathfrak{s} \times M$ one first assumes that $\Gamma(t)$ satisfies the differential equation

$$\dot{\Gamma}(t) = -\xi(t)\Gamma(t).,\tag{3.6}$$

and then determines $(g(t), y(t)) \in Q$ by the differential equations (3.3) and (3.4) up to initial conditions $(g(0), y(0))$, such that $g(0)^{-1}a_0 = \Gamma(0)$. We call equation (3.6) an advection equation because it follows from differentiating the equation $\Gamma(t) = g^{-1}(t)a_0$. By assuming that $\Gamma(t)$ satisfies the differential equation (3.6), one has that equation (3.5) is satisfied for some fixed vector $a_0 \in M$.

From a basic system $L, \mathcal{D}$ on TQ, we obtain a reduced Lagrangian $l : \mathfrak{s} \to \mathbb{R}$, reduced constraint equation and reduced constrained Lagrangian $l_c : \mathfrak{g} \times M \to \mathbb{R}$ as follows. First note that the S invariance of $\bar{L}$ implies that both $\bar{L}$ and L are independent of y. Omitting the y dependence, we write $\bar{L}(g, \dot{g}, \dot{y})$ and $L(g, \dot{g}, \dot{y})$. The reduced Lagrangian $l : \mathfrak{s} \to \mathbb{R}$ is defined by

$$l(\xi, Y) = \bar{L}(e, \xi, Y).$$

The S-invariance of $\bar{L}$ implies that

$$\bar{L}(g, \dot{g}, \dot{y}) = l(g^{-1}\dot{g}, g^{-1}\dot{y}).$$

The reduced constraint equation on $\mathfrak{s} \times V$ is obtained by applying g^{-1} to both sides of the constraint equation determining $\mathcal{D}$ ($\dot{y} = \omega a_0$) and writing things in terms of the reduced curves, that is, setting $\xi = g^{-1}\dot{g}$, $Y = g^{-1}\dot{y}$ and $\Gamma = g^{-1}a_0$. In this manner one obtains

$$Y = \xi\Gamma.$$

The J-invariance of L allows us to define the reduced constrained Lagrangian $l_c : \mathfrak{g} \times M \to \mathbb{R}$ by $L|_{\mathcal{D}} = l_c \circ \pi$, where π is the projection from $\mathcal{D}$ to $\mathfrak{g} \times M$. We realize it by restricting l to the constraints:

$$l_c(\xi, \Gamma) = l(\xi, \xi\Gamma).$$

Theorem 3.1. Basic Theorem.
Using the preceding notation, the following are equivalent:

(1) *The curve $(g(t), y(t)) \in Q$ satisfies the Lagrange-d'Alembert principle for the basic system. That is $(\dot{g}, \dot{y}) \in \mathcal{D}_{(g(t),y(t))}$ and*

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

where $\delta g(t)$ is an independent variation vanishing at the endpoints, and $\delta y = (\delta g\, g^{-1})a_0$.

(2) *The curve $(\xi(t), Y(t), \Gamma(t))$, where $\dot{\Gamma} = -\xi(t)\Gamma(t)$, satisfies the following constrained principle on $\mathfrak{s} \times M$. One has $Y(t) = \xi(t)\Gamma(t)$, and*

$$\delta \int l(\xi(t), Y(t))dt = 0,$$

where the variations along the curve take the form:

$$\delta\xi = \dot{\eta} + ad_{\xi}\eta$$
$$\delta Y = \frac{d}{dt}(\eta\,\Gamma) + ad_{\xi}\eta\,\Gamma,$$

and $\eta(t)$ is an independent variation vanishing at the endpoints.

(3) The Euler-Poincaré equation

$$\frac{d}{dt}\frac{\partial l_c}{\partial \xi} - ad^*_\xi \frac{\partial l_c}{\partial \xi} = -\rho^*_{\xi\Gamma}\frac{\partial l}{\partial Y}, \tag{3.7}$$

and advection equation

$$\dot{\Gamma} = -\xi\,\Gamma \tag{3.8}$$

hold on $\mathfrak{g} \times M$, where Y is set equal to $\xi\Gamma$ in the argument of $\partial l/\partial Y$.

The proof for the theorem will be given in Section 3.2 after the Intermediate Theorem is proved.

The energy of the basic system may be written as a function on $\mathfrak{g} \times M$. It takes the form

$$E = \langle \frac{\partial l_c}{\partial \xi}, \xi \rangle - l_c.$$

The energy takes the same form for the intermediate and general systems. We will show it is conserved after stating the reduced equations for the general system.

3.1.2 Momentum

We now turn to the special case when the reduced equation (3.7) is a pure Euler-Poincaré equation, as described in the introduction. In this case, the system conserves a $\mathfrak{g}^*$ valued momentum function $\mathbb{J} : TG \to \mathfrak{g}^*$ defined below. Here $\mathbb{J}$ represents momentum for the system in an inertial frame. One may think of $\partial l_c/\partial \xi$ as this momentum relative to a different frame; for the examples involving a rolling rigid body, this is a frame carried with the body. One may treat $\partial l_c/\partial \xi$ as a function on TG as opposed to $\mathfrak{g} \times M$ by composing it with the map from TG to $\mathfrak{g} \times M$ defined by $(g, \dot{g}) \mapsto (g^{-1}\dot{g}, g^{-1}a_0)$.

Definition 3.1. *For the basic system, the spatial momentum $\mathbb{J} : TG \to \mathfrak{g}^*$ is:*

$$\mathbb{J}(g, \xi) = Ad^*_{g^{-1}}\frac{\partial l_c}{\partial \xi}.$$

The following theorem characterizes when the reduced equation (3.7) is a pure Euler-Poincaré equation.

Theorem 3.2. *For the basic system, suppose that $\rho^*_{\xi\Gamma}\partial l/\partial Y = 0$. Then $\mathbb{J}$ is a conserved quantity for the reduced flow described by equations (3.7) and (3.8).*

Proof. Under the hypothesis, equation (3.7) takes the form

$$\frac{d}{dt}\frac{\partial l_c}{\partial \xi} - ad^*_\xi \frac{\partial l_c}{\partial \xi} = 0.$$

Now take $\frac{d}{dt}$ of $\langle \mathbb{J}(g,\xi), \eta\rangle$ (where $\eta \in \mathfrak{g}$ is arbitrary) to obtain:

$$\frac{d}{dt}\left(\langle Ad^*_{g^{-1}}\frac{\partial l_c}{\partial \xi}, \eta\rangle\right) =$$

$$\langle \frac{d}{dt}\frac{\partial l_c}{\partial \xi}, Ad_{g^{-1}}\eta\rangle + \langle \frac{\partial l_c}{\partial \xi}, \frac{d}{dt}Ad_{g^{-1}}\eta\rangle =$$

$$\langle \frac{d}{dt}\frac{\partial l_c}{\partial \xi}, Ad_{g^{-1}}\eta\rangle + \langle \frac{\partial l_c}{\partial \xi}, -[\xi, Ad_{g^{-1}}(\eta)]\rangle =$$

$$= \langle Ad^*_{g^{-1}}\left(\frac{d}{dt}\frac{\partial l_c}{\partial \xi} - ad^*_\xi \frac{\partial l_c}{\partial \xi}\right), \eta\rangle = 0.$$

□

This case is of interest as $\mathbb{J}$ is conserved, a useful property for analyzing systems.

There are several formulations of momentum for nonholonomic systems in the literature. In general, one finds that momentum is not conserved for nonholonomic systems. One formulation, going back to Lagrange and Jacobi, which is discussed in [FK] is as follows. Let G be a group which acts on the configuration space. Let $\mathfrak{g}$ be the Lie algebra of G. Suppose that the kinetic energy of the system is invariant under the action and the distribution $\mathcal{D}$ is consistent with the action, which means that for any $\xi \in \mathfrak{g}$, the infinitesimal generator for the action corresponding to ξ is contained in $\mathcal{D}$. Then one obtains a $\mathfrak{g}^*$ valued momentum map and an evolution equation for this momentum. Another formulation of momentum found in [BKMM] starts with a group action for which both the Lagrangian and distribution are invariant. The distribution is used to determine a point dependent subspace of $\mathfrak{g}$ and the momentum maps from TQ to the dual of this subspace. The evolution equation that is derived for this momentum has interesting applications in control theory, see [KM] and [O]. The last formulation we discuss, which is the most similar to our own, is presented in [VV] for RL systems. An RL system is one whose configuration space is a group G, constraints are invariant under the right action of G, and Lagrangian is invariant under the left action of G. Spatial momentum is defined for these systems as a map from TG to $\mathfrak{g}^*$. It is shown that RL systems conserve the restriction of this momentum to $\mathcal{D}$. We shall make a few more remarks about the approaches of [FK] and [VV] in the next section.

3.1.3 Examples

We shall demonstrate the Basic Theorem by deriving the equations of motion for Chaplygin's sphere in three and n dimensions. They are both cases which conserve the momentum discussed above in Section 3.1.2. Some consequences of the conserved momentum, namely several conserved quantities in the reduced equations, will be discussed.

Chaplygin's sphere is a ball that is constrained to roll without slipping on a horizontal plane, and whose mass is distributed so that the center of mass is at the center of the ball, but the moments of inertia around the principal axes may differ from one another. Let I denote the inertia tensor, r the ball's radius and m the total mass of the ball. The configuration space is $SO(3) \times \mathbb{R}^2$, but it will be convenient to describe the motion of the ball with a curve $q(t) \in SE(3)$. We write $q \in SE(3)$ as (A, x) where $A \in SO(3)$ and $x \in \mathbb{R}^3$. The curve $q(t)$ maps from a reference coordinate system for the ball with origin at the center of the ball to an inertial coordinate system for space. Let $P \in \mathbb{R}^3$ be a given point in the ball, with coordinates relative to the reference system. The location of P at time t is $A(t)P + x(t)$. The coordinate vectors for the reference system, denoted (E_1, E_2, E_3), are chosen to coincide with principal axes for the ball, thus diagonalizing the inertia tensor. Let I_1, I_2 and I_3 denote the principal moments of inertia of the ball. The coordinate vectors for space, denoted (e_1, e_2, e_3), are chosen so that e_1, e_2 span the horizontal plane upon which the ball rolls, and e_3 is the upward vertical normal vector to this plane. We take the horizontal plane to be located at $x_3 = -r$ so that the ball's center of mass has coordinates $(x_1(t), x_2(t), 0)$ in space. We will sometimes refer to the body frame which is the moving frame $(A(t)E_1, A(t)E_2, A(t)E_3)$ in the spatial coordinate system.

A tangent vector in $T_{(A,x)}SE(3)$ is denoted by $(\dot{A}, \dot{x})$ where $\dot{x} \in \mathbb{R}^3$ and $\dot{A} \in M^3(\mathbb{R})$ is obtained by differentiating a curve in $SO(3)$ going through A. In general we will identify $\mathbb{R}^3$

with the cross product ($\times$) and $\mathfrak{so}(3)$ with the usual Lie bracket with the usual isomorphism of these Lie algebras, the hat map $\hat{}: \mathbb{R}^3 \to \mathfrak{so}(3)$. Specifically, for a vector $\mathbf{v} \in \mathbb{R}^3$ with coordinates (v_1, v_2, v_3),

$$\hat{\mathbf{v}} = \begin{bmatrix} 0 & -v_3 & v_2 \\ v_3 & 0 & -v_1 \\ -v_2 & v_1 & 0 \end{bmatrix}; \tag{3.9}$$

for details see [MR, p. 259]. In particular, the ball's angular velocity about the center of mass in space and in the body are denoted by $\omega \in \mathbb{R}^3$ and $\Omega \in \mathbb{R}^3$. They are defined by

$$\hat{\omega} = \dot{A}A^{-1}$$

$$\hat{\Omega} = A^{-1}\dot{A}.$$

We also identify $\mathfrak{so}(3)^*$ with $\mathbb{R}^3$ using the $\hat{}$ map and Euclidean inner product on $\mathbb{R}^3$. With this identification, one has

$$ad^*_{\Omega}\theta = -\Omega \times \theta,$$

for $\Omega \in \mathfrak{so}(3)$ and $\theta \in \mathfrak{so}(3)^*$. The constraints are derived from the condition that the ball rolls without slipping. This implies that the contact point has zero velocity. The velocity of the contact point is a sum of the velocity of the center of mass, and the rotational velocity of the contact point about the center of mass. As the contact point is $-r\,e_3$, we obtain the constraint equation:

$$\dot{x} + \omega \times (-r\,e_3) = 0. \tag{3.10}$$

Integrating the e_3 component of (3.10) gives the holonomic constraint $x_3 = 0$. The components of (3.10) in the e_1 and e_2 plane give the constraint equations on $SO(3) \times \mathbb{R}^2$.

The Lagrangian for Chaplygin's sphere consists only of kinetic energy. The potential energy is taken to be zero as the center of mass never changes height. The kinetic energy is the sum of rotation about the center of mass and translation of the center of mass. On $TSE(3)$, this is given by substituting the holonomic constraint (in the form $\dot{x}_3 = 0$) into

$$\bar{L} = \frac{1}{2}\{\langle\Omega, I\Omega\rangle + m\|\dot{x}\|^2\}.$$

To apply Theorem 3.1, take $S = SE(3)$, $a_0 = re_3$, $\bar{L}$ as above and $\bar{\mathcal{D}}$ to be the distribution in $TSE(3)$ defined by equation (3.10). Note that $\bar{L}$ is S-invariant and that due to the correspondence between $\mathfrak{so}(3)$ and $\mathbb{R}^3$ given by the hat map, the constraint equation $\dot{x} = \omega \times a_0$ determines a distribution of the appropriate form required for Theorem 3.1. The configuration space is determined by the range of ρ_{re_3}. The range is the horizontal plane with normal vector e_3, so $Q = SO(3) \times \mathbb{R}^2$. The Lagrangian and distribution are then $L = \bar{L}|_{TQ}$ and $\mathcal{D} = \bar{\mathcal{D}}|_Q$ which is contained in TQ as noted in Section 3.1.1. To observe the symmetry group for Chaplygin's sphere, first note that with respect to the action of $SO(3)$ on $\mathbb{R}^3$ the isotropy subgroup of e_3 is S^1, occurring as $SO(2)$. Proposition 3.1 implies that the symmetry group is $SO(2)\circledS\mathbb{R}^2 = SE(2)$. The orbit of e_3 in $\mathbb{R}^3$ with respect to the action of $SO(3)$ is S^2, hence Proposition 3.2 implies that the reduced space is $\mathfrak{so}(3) \times S^2$. To explicate the role of r, we define $\Gamma = A^{-1}e_3$ as opposed to $A^{-1}r\,e_3$. For ease of calculation, $\Gamma \in S^2$ is treated as a vector in $\mathbb{R}^3$.

We proceed to compute the reduced equations in terms of Ω and Γ. The reduced constraint equation is $Y = r\Omega \times \Gamma$. The reduced Lagrangians are

$$l : \mathfrak{se}(3) \to \mathbb{R}, \quad l = \frac{1}{2}\{\langle \Omega, I\Omega\rangle + m\|Y\|^2\}$$

$$l_c : \mathfrak{so}(3) \times S^2 \to \mathbb{R}, \quad l_c = \frac{1}{2}\{\langle \Omega, I\Omega\rangle + mr^2\,\|\Omega \times \Gamma\|^2\}.$$

The derivatives required to compute the equations of motion are

$$\frac{\partial l_c}{\partial \Omega} = I\Omega + mr^2\Gamma \times (\Omega \times \Gamma)$$

$$\frac{\partial l}{\partial Y} = mY = m\,r\Omega \times \Gamma.$$

Let M denote $\partial l_c/\partial\Omega$. This represents the momentum (see Definition 3.1) in the body frame. For $\theta \in \mathbb{R}^{3*}$, and $v \in \mathbb{R}^3$, $\rho_v^*\theta = v \times \theta$, where we identify $\mathbb{R}^{3*}$ and $\mathbb{R}^3$ with the Euclidean inner product. The right hand side of the reduced equation (3.7) is

$$-\rho^*_{\Omega\times r\Gamma}\frac{\partial l}{\partial Y} = -(\Omega \times r\Gamma) \times (m\,r\Omega \times \Gamma) = 0.$$

Hence the reduced equations take the form:

$$\dot{M} + \Omega \times M = 0 \tag{3.11}$$

$$\dot{\Gamma} + \Omega \times \Gamma = 0. \tag{3.12}$$

As the hypothesis for Theorem 3.2 is satisfied, the momentum in space, $Ad^*_{A^{-1}}M = AM$ is conserved. Let $\bar{\mathbf{m}} = AM$; as both $\bar{\mathbf{m}}$ and e_3 are constant vectors in space, one has the following two conserved quantities in the body frame

$$\langle M, M\rangle \quad \text{and} \quad \langle M, \Gamma\rangle.$$

As noted after the statement of Theorem 3.1, a mechanical system with nonholonomic constraints conserves energy. For Chaplygin's sphere, the energy is simply l_c and takes the form $\frac{1}{2}\langle M, \Omega\rangle$, which gives a third conserved quantity.

A special property of Chaplygin's sphere is that the reduced equations have an invariant measure. This is a volume form that is invariant under the flow. Although Hamiltonian systems always have an invariant measure, a nonholonomic system generally does not. To write the volume form for Chaplygin's sphere, let $d\sigma$ denote the surface element on the sphere and $d\Omega$ the canonical volume form induced on $\mathfrak{so}(3)$ by our choice of coordinates Ω for $\mathfrak{so}(3)$. The invariant measure on $\mathfrak{so}(3) \times S^2$ is

$$\left\{\frac{1}{mr^2} - \left(\frac{\Gamma_1^2}{I_1 + mr^2} + \frac{\Gamma_2^2}{I_2 + mr^2} + \frac{\Gamma_3^2}{I_3 + mr^2}\right)\right\}^{1/2} d\Omega d\Gamma.$$

In the next section we will prove that the n dimensional Chaplygin sphere has an invariant measure. The formula obtained for this invariant measure agrees with the above when $n = 3$.

Chaplygin's study of this system is summarized in both [FK] and [A]. The equations of motion are obtained by carefully considering the group of rotations about the contact point. The momentum theorem discussed in [FK] which we summarized in Section 3.1.2, may be applied for this group action. This leads to the conservation of momentum in space and the reduced equations, the same reduced equations that we have obtained with Theorem 3.1. The Euler-Poincaré equation obtained from Theorem 3.1 is an evolution equation, in the body frame, for the same momentum Chaplygin arrived at by considering the group of rotations about the contact point. The conservation of the momentum in space follows from $\rho^*_{\xi\Gamma}\partial l/\partial Y = 0$ as in the hypothesis of Theorem 3.2.

Chaplygin obtained analytic solutions to the reduced equations by using the invariant measure and conserved quantities. Coordinates may be obtained by quadrature construction (which involves computing integrals and inverses to functions) where the reduced flow generically undergoes rectilinear winding on tori; winding that is nonuniform with respect to time. This qualitative picture of the flow is similar to that of an integrable Hamiltonian system. However, in the Hamiltonian picture, the winding is uniform.

The generalization of Chaplygin's sphere to n dimensions is a ball in $\mathbb{R}^n$ that rolls on a $n-1$ dimensional plane without slipping. As in the three dimensional case, the center of mass of the ball coincides with the geometric center, however the moments of inertia need not be equal. Let m denote the total mass of the ball and r the radius of the ball. We denote the hyperplane upon which the ball rolls by W and the upward pointing normal vector to W by e_n. The configuration space for the problem is $SO(n) \times \mathbb{R}^{n-1}$; however as with the ball in three space, we describe the motion with a curve $(A(t), x(t))$ in $SE(n)$ which maps from a reference coordinate system for the ball (with origin at the ball's center) into an inertial frame. A tangent vector in $T_{(A,x)}SE(n)$ is denoted by $(\dot{A}, \dot{x})$ where $\dot{x} \in \mathbb{R}^n$ and $\dot{A} \in M^n(\mathbb{R})$ is obtained by differentiating a curve in $SO(n)$ through A. There is no analog of the isomorphism between $\mathbb{R}^3$ and $\mathfrak{so}(3)$ for n dimensions. We denote the angular velocity for the ball with elements of $\mathfrak{so}(n)$. Let $\Omega = A^{-1}\dot{A}$ denote the angular velocity in the body and $\omega = \dot{A}A^{-1}$ the angular velocity in space. The action of $\mathfrak{so}(n)$ on $\mathbb{R}^n$ given by a matrix applied to a vector is denoted Ωx. The Euclidean inner product will be used to identify $\mathbb{R}^n$ with $\mathbb{R}^{n*}$. The following inner product on $\mathfrak{so}(n)$,

$$\langle A, B \rangle = -\frac{1}{2}\mathrm{tr}(AB), \tag{3.13}$$

where $\mathrm{tr}(AB)$ denotes the trace of AB, will be used to identify $\mathfrak{so}(n)$ with its dual $\mathfrak{so}(n)^*$. Note that by using this identification one has $ad^*_\Omega\theta = [\theta, \Omega]$ as the following calculation shows

$$(ad^*_\Omega\theta, \Sigma) = \langle \theta, [\Omega, \Sigma] \rangle = -\frac{1}{2}\mathrm{tr}(\theta\Omega\Sigma - \theta\Sigma\Omega) = -\frac{1}{2}\mathrm{tr}(\theta\Omega\Sigma - \Omega\theta\Sigma) = \langle [\theta, \Omega], \Sigma \rangle.$$

The constraints are determined in the same way as the three dimensional case. Requiring the contact point to have zero velocity leads to the equation

$$\dot{x} = r\omega e_n.$$

The Lagrangian consists only of kinetic energy, which is the sum of rotation about the center of mass and translation of the center of mass. On $TSE(n)$ this is given by

$$\bar{L} = \frac{1}{2}\int_{Q\in\mathrm{Ball}} \zeta(Q)\|\Omega Q\|^2 dQ^n + \frac{1}{2}m\|\dot{x}\|^2,$$

where $\zeta(Q)$ is the mass density of the ball. The kinetic energy due to rotation about the center of mass may be written as

$$\frac{1}{2}\int_{Q\in\mathrm{Ball}} \zeta(Q)\|\Omega Q\|^2 dQ^n = \frac{1}{2}\sum_{i,j,l=1}^{n} \Omega_{ij}\Omega_{il}\int_{Q\in\mathrm{Ball}} \zeta(Q)Q^j Q^l dQ^n. \tag{3.14}$$

The terms in the integral define a symmetric quadratic form on $\mathbb{R}^n$ which can be diagonalized in a suitable orthonormal basis. We take the coordinate axes in the ball to coincide with such a basis and denote the diagonal elements by λ^j. Then (3.14) is equal to

$$\frac{1}{2}\sum_{i,j=1}^{n} \lambda_j(\Omega_{ij})^2 = \frac{1}{2}\sum_{1\leq i<j\leq n} (\lambda_i + \lambda_j)(\Omega_{ij})^2. \tag{3.15}$$

By defining $\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$, and using the inner product (3.13), one may write (3.15) as $\frac{1}{2}\langle\Lambda\Omega + \Omega\Lambda, \Omega\rangle$. Hence the Lagrangian on $TSE(n)$ is

$$\bar{L} = \frac{1}{2}\left\{\langle\Lambda\Omega + \Omega\Lambda, \Omega\rangle + m\|\dot{x}\|^2\right\}.$$

As $\bar{L}$ is left $SE(n)$-invariant and the constraints are of the form $\dot{x} = \omega a_0$, the reduced equations may be obtained by Theorem 3.1. Let (Ω, Y) denote a vector in $\mathfrak{se}(n)$, and define $\Gamma \in \mathbb{R}^n$ by $\Gamma = A^{-1}e_n$. The reduced Lagrangian is

$$l(\Omega, Y) = \frac{1}{2}\left\{\langle\Lambda\Omega + \Omega\Lambda, \Omega\rangle + m\|Y\|^2\right\}.$$

The constraints in the body read $Y = r\Omega\Gamma$, and the constrained reduced Lagrangian is

$$l_c(\Omega, \Gamma) = \frac{1}{2}\left\{\langle\Lambda\Omega + \Omega\Lambda, \Omega\rangle + mr^2\|\Omega\Gamma\|^2\right\}.$$

At this point we calculate the reduced equation:

$$\frac{d}{dt}\frac{\partial l_c}{\partial\Omega} - ad^*_\Omega\frac{\partial l_c}{\partial\Omega} = -r\rho^*_{\Omega\Gamma}\frac{\partial l}{\partial Y}. \tag{3.16}$$

We will write equation (3.16) on $\mathfrak{so}(n)$ by identify $\mathfrak{so}(n)^*$ with $\mathfrak{so}(n)$ with the inner product (3.13). The computation of $\partial l_c/\partial\Omega$ is facilitated by writing l_c as a quadratic form in this inner product.

Proposition 3.3. *The reduced constrained Lagrangian, l_c, may be written as:*

$$l_c(\Omega,\Gamma)=\frac{1}{2}\langle I(\Gamma)\Omega,\Omega\rangle,$$

where $I(\Gamma):\mathfrak{so}(n)\to\mathfrak{so}(n)$ is defined by
$I(\Gamma)\Omega=\Lambda\Omega+\Omega\Lambda+mr^2(P\Omega+\Omega P)$, *and* $P=\Gamma\Gamma^T$.

Proof. It suffices to show that $\|\Omega\Gamma\|^2=\langle P\Omega+\Omega P,\Omega\rangle$.

$$\|\Omega\Gamma\|^2=(\Omega\Gamma)^T\Omega\Gamma=-\Gamma^T\Omega^2\Gamma=-\mathrm{tr}(\Omega^2P),$$

$$\langle P\Omega+\Omega P,\Omega\rangle=-\frac{1}{2}\mathrm{tr}(P\Omega^2+\Omega P\Omega)=-\mathrm{tr}(P\Omega^2).$$

□

As $I(\Gamma)$ is symmetric with respect to the inner product, one has that

$$\frac{\partial l_c}{\partial\Omega}=I(\Gamma)\Omega.$$

Previously we noted that $ad^*_\Omega\theta=[\theta,\Omega]$. It remains to determine the form of $\rho^*_a\beta$ when $a\in\mathbb{R}^n$ and $\beta\in\mathbb{R}^{n*}$, which is identified with $\mathbb{R}^n$ using the Euclidean inner product. To distinguish the two inner products, $\langle,\rangle_e$ is used for the Euclidean inner product on $\mathbb{R}^n$ and $\langle,\rangle$ is used for the inner product (3.13) on $\mathfrak{so}(n)$. Then

$$\langle\rho^*_a\beta,\Omega\rangle=\langle\beta,\Omega\,a\rangle_e=\sum_{i<j}\Omega_{ij}(\beta^ia^j-\beta^ja^i).$$

Therefore $(\rho^*_a\beta)^{ij}=\beta^ia^j-\beta^ja^i$. In particular, $\rho^*_aa=0$. As $\partial l/\partial Y=mr\Omega\Gamma$, the right hand side of (3.16) is zero, and the reduced equations are:

$$\dot M=[M,\Omega] \tag{3.17}$$
$$\dot\Gamma=-\Omega\Gamma$$

where $M=I(\Gamma)\Omega$.

Recall that the equations of motion for Chaplygin's sphere in three dimensions, where $(\Omega,\Gamma)\in\mathbb{R}^3\times\mathbb{R}^3$ are the dynamic variables, conserve the four functions

$$\langle M,M\rangle_e,\quad\langle M,\Gamma\rangle_e,\quad\langle\Gamma,\Gamma\rangle_e=1,\ \text{and}\quad\frac{1}{2}\langle M,\Omega\rangle_e,$$

and preserve an invariant measure. The n dimensional system has a number of conserved quantities which generalize these four, and also has an invariant measure. The energy for the n dimensional system takes the same form, $l_c = 1/2\langle M, \Omega\rangle$. A general description of the other conserved quantities, which is found in [FK], is as follows. The equations $\dot{M} = [M, \Omega]$ and $\dot{P} = [P, \Omega]$ imply that $Ad_A M$ and $Ad_A P$ are conserved. (The conservation of $Ad_A M$ also follows from Theorem 3.2.) The property $\mathrm{tr}(AB) = \mathrm{tr}(BA)$ then implies that any function of the form:

$$\mathrm{tr}(M^{s_1} P M^{s_2} P \ldots M^{s_k} P^l),$$

is conserved (where this last l may be zero or one).

For problems such as integrating a mechanical system or analyzing the stability of relative equilibria, knowing the number of functionally independent conserved quantities is important. The above description gives no indication of what this number is (although one can quickly see that it can't be all functions of the above form). To contribute to the problem of finding the number of functionally independent conserved quantities in the system, we shall next derive our own family of conserved quantities for the n dimensional Chaplygin sphere, done independently from [FK], whereby we obtain

$$\frac{1}{2}(n(n-1)/2 + \lfloor n/2 \rfloor) + 1$$

conserved quantities. Although we do not show that they are functionally independent, we consider them to be a good candidate for the set of functionally independent conserved quantities one could obtain from the family given by [FK]. They shall be derived using Lax pairs in a manner analogous to what was done for the n dimensional free rigid body [R]. The conserved quantities obtained for the n dimensional free rigid body using Lax pairs have been shown to be functionally independent and in involution with one another. This led to the complete integrability, in a Hamiltonian sense, of the system.

A Lax pair for a set of differential equations is a pair of $n \times n$ matrices, typically denoted L and A, satisfying $\dot{L} = [L, A] = LA - AL$. A well known result then implies that the eigenvalues of L are first integrals of the equations [A, p. 129]. Equivalently, one may use $\mathrm{tr}(L^k)$ as the first integrals, where $k \leq n$ and n is the size of the matrices.

Proposition 3.4. *The n dimensional Chaplygin sphere equations (3.17) conserve the functions d_{kj}, $k = 1 \ldots n$, $j = 0 \ldots k$ defined by*

$$\sum_{j=0}^{k} d_{kj}\alpha^j = \operatorname{tr}((M + \alpha P)^k) \quad k = 1 \ldots n,$$

where $P = \Gamma\Gamma^T$ and α is an arbitrary scalar.

Proof. The equation $\dot{M} = [M, \Omega]$ is a Lax pair for M. By differentiating $P = \Gamma\Gamma^T$ and using $\dot{\Gamma} = -\Omega\Gamma$, one has the Lax pair $\dot{P} = [P, \Omega]$ for P. Then for any scalar $\alpha \in \mathbb{R}$, one has the Lax pair

$$\frac{d}{dt}(M + \alpha P) = [M + \alpha P, \Omega],$$

for $M + \alpha P$. Therefore traces of powers of $M + \alpha P$ are conserved. As α is arbitrary, this implies that each of the d_{kj} is conserved. □

As M is skew, P is symmetric, and $\operatorname{tr}(A^T) = \operatorname{tr}(A)$, it follows that $d_{kj} = 0$ when $k - j$ is odd. As $P^2 = P$, $d_{kk} = \operatorname{tr}(P)$ for all k. Our count of the number of conserved quantities comes from counting the number of nonzero functions among the d_{kj}'s when $j \neq k$ and counting the $d'_{kk}s$ once as $P^2 = P$. An inductive argument gives the above count.

The conserved functions $\langle \Gamma, \Gamma \rangle_e = 1$, $\langle M, M \rangle_e$ and $\langle M, \Gamma \rangle_e$ corresponding to the $n = 3$ case are found amongst the d_{kj}'s as follows. Clearly $\operatorname{tr}(P)^2 = \langle \Gamma, \Gamma \rangle_e = 1$ and $d_{20} = 2\langle M, M \rangle_e$. Expanding d_{13}, where $\widehat{M} \in \mathfrak{so}(3)$ and $M \in \mathbb{R}^3$, one finds

$$\begin{aligned} &-\frac{1}{3}d_{31} = -\frac{1}{3}\operatorname{tr}(\widehat{M}^2 P + \widehat{M}P\widehat{M} + P\widehat{M}^2) = -\operatorname{tr}(\widehat{M}^2 P) = \\ &-\widehat{M}_{ij}\widehat{M}_{jk}\Gamma_k\Gamma_i = \langle M \times \Gamma, M \times \Gamma \rangle_e = \langle M, \Gamma \times (M \times \Gamma) \rangle_e = \|M\|_e^2 - \langle M, \Gamma \rangle_e^2. \end{aligned}$$

In calculating the invariant measure for Chaplygin's sphere in n dimensions, we will distinguish between the spaces $\mathfrak{so}(n)$ and $\mathfrak{so}(n)^*$ and express the measure as a form on $\mathfrak{so}(n)^* \times S^{n-1}$. Recall that $M \in \mathfrak{so}(n)^*$ is obtained from $\Omega \in \mathfrak{so}(n)$ by

$$M = I(\Gamma)\Omega = \frac{\partial l_c}{\partial \Omega}.$$

This defines $I(\Gamma)$ as a map from $\mathfrak{so}(n)$ to $\mathfrak{so}(n)^*$. Let dM denote the standard volume form on the vector space $\mathfrak{so}(n)^*$ induced by a choice of coordinates for M and let $d\sigma$ denote the standard area form on $S^{n-1} \subset \mathbb{R}^n$. Then $dM\, d\sigma$ is a volume form on $\mathfrak{so}(n)^* \times S^{n-1}$.

Proposition 3.5. *With the above notation, the reduced equations (3.17) for the n dimensional Chaplygin's sphere on* $\mathfrak{so}(n)^* \times S^{n-1}$ *have the invariant measure*

$$\frac{1}{\sqrt{\det(I(\Gamma))}} dM\, d\sigma.$$

Proof. As the flow conserves $\langle \Gamma, \Gamma \rangle$, it suffices to show that $(\det(I(\Gamma)))^{-1/2} dM\, d\Gamma$ is invariant in $\mathfrak{so}(n)^* \times \mathbb{R}^n$. The calculation will be done in local coordinates on $\mathfrak{so}(n)^* \times \mathbb{R}^n$. Let e_i be an orthonormal basis for $\mathfrak{so}(n)$ relative to the inner product (3.13), and let f_a be a basis for $\mathbb{R}^n$. Write $\Omega \in \mathfrak{so}(n)$ as $\Omega^i e_i$ and $\Gamma \in \mathbb{R}^n$ as $\Gamma^a f_a$. Let de^i be a dual basis to the e_i's which determine coordinates M_i on $\mathfrak{so}(n)^*$. Let c^i_{jk} denote the structure constants for $\mathfrak{so}(n)$ in this basis, that is,

$$[e_j, e_k] = c^i_{jk}\, e_i.$$

Note that $\langle [e_j, e_k], e_k \rangle = \langle e_j, [e_k, e_k] \rangle = 0$. Thus for any k and j, $c^k_{jk} = 0$. Let A^a_{bk} denote the constants for the action of $\mathfrak{so}(n)$ on $\mathbb{R}^n$ in this basis, that is

$$A^a_{bk} = \langle f_a, e_k f_b \rangle.$$

Note that $A^a_{bk} = -A^b_{ak}$. In what follows, we assume the summation convention, that is a repeated index implies a sum. Indices in a, b, c and d go from $1 \ldots n$ and sum over coordinates in $\mathbb{R}^n$. Indices in i, j, k and l sum from $1 \ldots n(n-1)/2$ over coordinates in $\mathfrak{so}(n)$ or $\mathfrak{so}(n)^*$. We write $M = I(\Gamma)\Omega$ in these coordinates as

$$M_i = I_{ij}\Omega^j.$$

From the definition of l_c, one determines the components I_{ij} to be

$$I_{ij} = \lambda_i \delta^i_j + mr^2 A^d_{bi} A_{dcj} \Gamma^b \Gamma^c, \tag{3.18}$$

where indices in A_{dcj} have been lowered with the Euclidean inner product and δ^i_j are Kronecker delta functions. Components of the inverse of the inertia matrix will be denoted I^{ij}, in particular $\Omega^i = I^{ij} M_i$.

Let $\Phi(\Gamma) = (\det I(\Gamma))^{-1/2}$. We must show that the divergence of Φ times the flow is zero. In these coordinates, the flow is

$$\dot{M}_k = c^i_{jk} M_i \Omega^j$$
$$\dot{\Gamma}^a = -A^a_{bi} \Gamma^b \Omega^i.$$

We compute the divergence:

$$\frac{\partial}{\partial M_k}\left(\Phi(\Gamma)\dot{M}_k\right) + \frac{\partial}{\partial \Gamma^a}\left(\dot{\Gamma}^a \Phi(\Gamma)\right) =$$
$$\Phi(\Gamma)\left\{c^k_{jk}\Omega^j + c^i_{jk} M_i I^{jk} - A^a{}_{ai}\Omega^i - A^a{}_{bi}\Gamma^b \frac{\partial \Omega^i}{\partial \Gamma^a}\right\} - A^a{}_{bi}\Gamma^b \Omega^i \frac{\partial \Phi(\Gamma)}{\partial \Gamma^a}. \tag{3.19}$$

The first three terms vanish because $c^k_{jk} = 0$, $c^i_{jk} I^{jk}$ is the trace of a skew and symmetric form, and $A^a{}_{ai} = -A^a{}_{ai}$. The last term is $\frac{d}{dt}\Phi(\Gamma)$, which means (3.19) is equal to

$$\frac{d}{dt}\Phi(\Gamma) - \Phi(\Gamma) A^a{}_{bi}\Gamma^b \frac{\partial \Omega^i}{\partial \Gamma^a}. \tag{3.20}$$

Differentiating $\Omega^i = I^{ij} M_j$ and using (3.18) shows that (3.20) is equal to

$$\frac{d}{dt}\Phi(\Gamma) - mr^2 \Phi(\Gamma) I^{ik}\Omega^l \left\{A^a{}_{bi} A^d{}_{ad} A_{dcl} + A^a{}_{bi} A^d{}_{ck} A_{dal}\right\} \Gamma^b \Gamma^c. \tag{3.21}$$

Using the well known formula for the derivative of the determinant, $\frac{d}{dt}\det(I) = \dot{I}_{ij} I^{ij} \det(I)$, the first term in (3.21) is seen to be

$$-\frac{1}{2}\Phi(\Gamma)\frac{\partial I_{ij}}{\partial \Gamma^a} I^{ij} \dot{\Gamma}^a.$$

After substituting $-A^a_{bi}\Gamma^b\Omega^i$ for $\dot{\Gamma}^a$ and the expression for $\partial I_{ij}/\partial \Gamma^a$ that is obtained by differentiating (3.18), one sees the problem reduces to showing

$$\Gamma^b \Gamma^c \Omega^l I^{ij}\left[\frac{1}{2}\left(A^a{}_{bl} A_{dai} A^d{}_{cj} + A^a{}_{bl} A_{dci} A^d{}_{aj}\right) + A^a{}_{bi} A_{dal} A^d{}_{cj} + A^a{}_{bi} A_{dcl} A^d{}_{aj}\right] = 0.$$

The third term is 0 as it is skew in i and j and contracted with the symmetric terms I^{ij}. By interchanging b with c (as $\Gamma^b\Gamma^c$ is symmetric) and d with a, one obtains:

$$A^a{}_{bi} A_{dal} A^d{}_{cj} = A^d{}_{ci} A_{adl} A^a{}_{bj}.$$

The property $A_{adl} = -A_{dal}$ shows that this third term is skew in i, j.

The first two terms add up to cancel with the fourth term:

$$\frac{1}{2}\left(A^a_{\ bl}A_{dai}A^d_{\ cj} + A^a_{\ bl}A_{dci}A^d_{\ aj}\right) = \frac{1}{2}\left(-A^a_{\ bl}A^d_{ai}A^c_{\ dj} - A^a_{\ bl}A_{cdj}A^d_{\ ai}\right).$$

This follows by using $A_{dci} = -A_{cdi}$ and interchanging i and j in the second term, and using $A^d_{\ cj} = -A^c_{\ dj}$ in the first term. In the fourth term, by interchanging a and d, swapping b and c, and i with j, one obtains

$$A^a_{\ bl}A_{adi}A^d_{\ cj}.$$

This shows that the first two terms cancel with the fourth. □

A different calculation for proving the n dimensional Chaplygin's sphere has this invariant measure is given in [FK]. In [VV], it is shown that RL systems on unimodular groups have an invariant measure. Recall that RL systems have a Lie group for the configuration space, a left-invariant Lagrangian and a right-invariant distribution. In [VV], it is stated that Chaplygin's sphere is an RL system for the group $SE(3)$, however this is not the case. It is true that Chaplygin's sphere is an RL system when the group is $SO(3) \times \mathbb{R}^3$, hence the above invariant measure also follows from this fact.

The n dimensional equations on $\mathfrak{so}(n)$ are equivalent to the three dimensional equations on $\mathbb{R}^3$, when $n = 3$. First one sets $\Omega_{12} = \Omega_3, \Omega_{13} = -\Omega_2$, and $\Omega_{23} = \Omega_1$, as prescribed by the hat map as in equation (3.9). The principal moments of inertia for the ball are related to the λ_i's by

$$I_1 = \lambda_2 + \lambda_3, \quad I_2 = \lambda_1 + \lambda_3, \quad I_3 = \lambda_1 + \lambda_2.$$

An unsolved problem is whether or not Chaplygin's sphere in n dimensions is integrable, in the same way that the three dimensional case is integrable. That is, can coordinates be constructed by quadrature so that the motion undergoes nonuniform rectilinear winding around tori of some dimension. The existence of the invariant measure can be taken as a hopeful sign that this is the case. There are several other problems in nonholonomic mechanics which have an invariant measure and whose solutions perform nonuniform rectilinear winding around tori, but in all these cases the tori are of dimension two [VV]. A method of integrating nonholonomic systems to obtain winding on higher dimensional tori is not known.

3.2 Intermediate Case

In this section we broaden the family of systems reduced by allowing the distribution to depend on the advected parameter a_0 in a more general way and consider Lagrangians whose symmetry is "broken" by a_0 in a manner similar to that introduced in [CHMR]. We will describe this family of systems, identify the symmetry group and reduced space, and then state and prove the reduction theorem. The proof of Theorem 3.1 is obtained by treating the basic case as a special case of the intermediate family. Finally two examples of the theorem are discussed: Chaplygin's top and the falling disk.

3.2.1 Intermediate Reduction Theorem

As with the basic case, the construction of the intermediate system starts with a semidirect product $S = G \circledS V$ where G is a Lie group, V is a vector space, and G can be represented as a subgroup of $Gl(n, V)$. The configuration space Q is taken to be S. Let $\bar{L} : TS \times V \to \mathbb{R}$ be a smooth function which is invariant under the following action of S on $TS \times V$:

$$(h, x) \cdot (g, \dot{g}, y, \dot{y}, a) = (hg, h\dot{g}, hy + x, h\dot{y}, ha). \tag{3.22}$$

We will denote this action by Φ_S. It consists of the product of the usual lifted action of S on TS and the action of G on V. Now let a_0 be a fixed element of V. The Lagrangian $L : TS \to \mathbb{R}$ is defined in terms of $\bar{L}$ as

$$L(g, \dot{g}, y, \dot{y}) = \bar{L}(g, \dot{g}, y, \dot{y}, a_0).$$

Let $\zeta : V \to V$ be a smooth vector valued function. The distribution $\mathcal{D} \subset TS$ is defined by

$$\mathcal{D}_{(g,y)} = \{(\dot{g}, \dot{y}) \in TS : \dot{y} = \omega g \zeta(g^{-1} a_0)\}.$$

We shall call $(S, L, \mathcal{D})$ the intermediate system.

Whether or not $\mathcal{D}$ admits integral submanifolds depends on ζ. For physical examples, $\mathcal{D}$ typically will have integral submanifolds, one being the configuration space. Further remarks on this are made following the examples.

We next identify the symmetry group and reduced space for the intermediate system. As before, we assume the G-orbit of a_0 in V is closed. The proofs of the following propositions are very similar to the proofs of propositions 3.1 and 3.2 and shall be omitted.

Proposition 3.6. *Both L and $\mathcal{D}$ of the intermediate system are invariant under the action of $J = H \circledS V$, where $H \subset G$ is the isotropy subgroup of a_0.*

Proposition 3.7. *The reduced space $\mathcal{D}/J$ is a smooth manifold diffeomorphic to $\mathfrak{g} \times M$, where M is G/H realized as a submanifold of V.*

Similarly to the basic system, from L, $\mathcal{D}$ and the Lagrange d'Alembert principle on TS, we obtain a reduced Lagrangian $l : \mathfrak{s} \times M \to \mathbb{R}$, a reduced constraint equation, and a reduced constrained principle for determining dynamics on the reduced space. We then establish that these are equivalent principles for determining the curves describing the dynamics in S and the curves describing the reduced dynamics, when the curves are related to each other in the same way as with the basic case. That is from the curves $(g(t), y(t)) \in S$ satisfying the Lagrange-d'Alembert principle for the intermediate system one obtains the curves $(\xi(t), Y(t), \Gamma(t)) \in \mathfrak{s} \times M$ satisfying the reduced principle by

$$\xi(t) = g^{-1}(t)\dot{g}(t) \tag{3.23}$$

$$Y(t) = g^{-1}(t)\dot{y}(t), \tag{3.24}$$

$$\Gamma(t) = g^{-1}(t)a_0.$$

Given a curve $(\xi(t), Y(t), \Gamma(t)) \in \mathfrak{s} \times M$ satisfying the constrained principle, which requires

$$\dot{\Gamma}(t) = -\xi(t)\Gamma(t), \tag{3.25}$$

one obtains $(g(t), y(t)) \in Q$ from the differential equations (3.23) and (3.24). Equation (3.25) ensures that $\Gamma(t) = g^{-1}(t)a_0$ for some $a_0 \in V$ when equation (3.23) holds.

We proceed to define the reduced constraint equation and reduced Lagrangians which are used to describe the reduced system. The Φ_S invariance of $\bar{L}$ implies that both $\bar{L}$ and L are independent of y. Omitting the y dependence, we write $L(g, \dot{g}, \dot{y})$ and $\bar{L}(g, \dot{g}, \dot{y}, a_0)$. The reduced Lagrangian $l : \mathfrak{s} \times M \to \mathbb{R}$ is defined by

$$l(\xi, Y, \Gamma) = \bar{L}(e, \xi, Y, \Gamma).$$

Given this, one sees that

$$\bar{L}(g, \dot{g}, \dot{y}, a) = l(g^{-1}\dot{g}, g^{-1}\dot{y}, g^{-1}a), \tag{3.26}$$

for $a \in M$. The reduced constraint equation on $\mathfrak{s} \times V$ is obtained by applying g^{-1} to both sides of $\dot{y} = \omega\, g\, \zeta(g^{-1}a_0)$ and writing things in terms of the reduced curves, that is setting $\xi = g^{-1}\dot{g}$, $Y = g^{-1}\dot{y}$ and $\Gamma = g^{-1}a_0$. In this manner one obtains

$$Y = \xi\, \zeta(\Gamma).$$

The constrained reduced Lagrangian $l_c : \mathfrak{g} \times M \to \mathbb{R}$ is defined by evaluating l on the constraints:

$$l_c(\xi, \Gamma) = l(\xi, \xi\zeta(\Gamma), \Gamma).$$

Theorem 3.3. Intermediate Theorem.

With the preceding notation, the following are equivalent:

(1) The curve $(g(t), y(t)) \in S$ satisfies the Lagrange-d'Alembert principle for the intermediate system. That is $(\dot{g}, \dot{y}) \in \mathcal{D}_{(g(t),y(t))}$ and

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

where $\delta g(t)$ is an independent variation vanishing at the endpoints, and $\delta y = (\delta g\, g^{-1})\, g\, \zeta(g^{-1}a_0)$.

(2) The curve $(\xi(t), Y(t), \Gamma(t))$, where $\dot{\Gamma} = -\xi(t)\Gamma(t)$, satisfies the following constrained principle on $\mathfrak{s} \times M$. One has $Y(t) = \xi(t)\zeta(\Gamma(t))$, and

$$\delta \int l(\xi(t), Y(t), \Gamma(t))dt = 0,$$

where the variations along the curve take the form:

$$\delta\xi = \dot{\eta} + ad_\xi \eta$$

$$\delta\Gamma = -\eta\,\Gamma$$

$$\delta Y = \frac{d}{dt}(\eta\, \zeta(\Gamma)) + ad_\xi \eta\, \zeta(\Gamma).$$

and $\eta(t)$ is an independent variation vanishing at the endpoints.

(3) The Euler-Poincaré equation

$$\frac{d}{dt}\frac{\partial l_c}{\partial \xi} - ad^*_\xi \frac{\partial l_c}{\partial \xi} + \rho^*_\Gamma \frac{\partial l}{\partial \Gamma} = \rho^*_{\frac{d}{dt}\zeta} \frac{\partial l}{\partial Y} \tag{3.27}$$

and advection equation

$$\dot{\Gamma} + \xi\,\Gamma = 0$$

hold on $\mathfrak{g} \times M$, *where* Y *is set equal to* $\xi\,\zeta(\Gamma)$ *in* $\partial l/\partial\Gamma$ *and* $\partial l/\partial Y$ *and the notation* $\frac{d}{dt}\zeta$ *refers to* $\frac{d}{dt}\zeta(\Gamma(t))$.

Proof. We first show that (1) is equivalent to (2). Given the relations between the curves,

$$\xi(t) = g(t)^{-1}\dot{g}(t), \quad Y(t) = g(t)^{-1}\dot{y}(t), \quad \text{and } \Gamma(t) = g(t)^{-1}a_0, \tag{3.28}$$

one sees the condition that $(\dot{g}, \dot{y}) \in \mathcal{D}_{(g(t),y(t))}$, or $\dot{y} = \dot{g}\zeta(g^{-1}a_0)$, is equivalent to $Y = \xi\zeta(\Gamma)$. The Φ_S invariance of $\bar{L}$ in conjunction with the relations (3.28) implies that the integrands $L(g(t), \dot{g}(t), \dot{y}(t))$ and $l(\xi(t), Y(t), \Gamma(t))$ agree. Assuming variations as in (1) variations for the variables ξ, Y and Γ are calculated from the relations (3.28) as follows.

$$\begin{aligned} \delta\xi &= \delta(g^{-1})\dot{g} + g^{-1}\delta(\dot{g}), \\ &= \frac{d}{dt}(g^{-1}\delta g) + g^{-1}\dot{g}g^{-1}\delta g - g^{-1}\delta g g^{-1}\dot{g} \end{aligned} \tag{3.29}$$

By setting $\eta(t) = g^{-1}(t)\delta g(t)$ one has an independent variation in $\mathfrak{g}$ vanishing at the endpoints. Writing equation (3.29) in terms of η and ξ, one has

$$\delta\xi = \dot{\eta} + ad_\xi \eta.$$

Likewise, $\delta\Gamma = -g^{-1}\delta g g^{-1}a_0 = -\eta\Gamma$. Lastly,

$$\delta Y = \frac{d}{dt}(g^{-1}\delta y) + g^{-1}\dot{g}g^{-1}\delta y - g^{-1}\delta g g^{-1}\dot{y}. \tag{3.30}$$

Setting $\delta y = \delta g \zeta(\Gamma)$ and substituting in ξ and η where appropriate, one obtains

$$\delta Y = \frac{d}{dt}\left(\eta\zeta(\Gamma)\right) + ad_\xi \eta\zeta(\Gamma).$$

Now we show the variations in (2) are equivalent to the variations in (1). In this case equation (3.29) is a differential equation which determines $\delta g(t)$. Requiring δg to vanish at an endpoint determines it uniquely to be $\delta g = g\eta$. Hence η being an independent variation vanishing at the endpoints implies that δg is as well. Likewise equation (3.30) is a differential equation determining δy. Taking δy to vanish at an endpoint uniquely determines it to be $\delta g\zeta(g^{-1}a_0)$.

It remains to show that (2) is equivalent to (3). We first apply the reduced principle in (2). Using the independence of η, the fact that η vanishes at the endpoints and integration by parts, one obtains:

$$\delta \int l\,dt =$$

$$\int \left\{ \langle \frac{\partial l}{\partial \xi}, \delta\xi\rangle + \langle \frac{\partial l}{\partial Y}, \delta Y\rangle + \langle \frac{\partial l}{\partial \Gamma}, \delta\Gamma\rangle \right\} dt =$$

$$\int \left\{ \langle \frac{\partial l}{\partial \xi}, \dot\eta + ad_\xi \eta\rangle + \langle \frac{\partial l}{\partial Y}, \dot\eta\,\zeta + ad_\xi\,\eta\,\zeta + \eta\,\dot\zeta\rangle + \langle \frac{\partial l}{\partial \Gamma}, -\eta\,\Gamma\rangle \right\} dt =$$

$$\int \left\{ \langle -\frac{d}{dt}\frac{\partial l}{\partial \xi} + ad_\xi^*\frac{\partial l}{\partial \xi} - \frac{d}{dt}\left(\rho_\zeta^*\frac{\partial l}{\partial Y}\right) + ad_\xi^*\,\rho_\zeta^*\,\frac{\partial l}{\partial Y} + \rho_{\frac{d}{dt}\zeta}^*\frac{\partial l}{\partial Y} - \rho_\Gamma^*\,\frac{\partial l}{\partial \Gamma}\,,\,\eta\rangle \right\} dt.$$

Setting this equal to zero and using the independence of η, we have the equation

$$-\frac{d}{dt}\frac{\partial l}{\partial \xi} + ad_\xi^*\frac{\partial l}{\partial \xi} - \frac{d}{dt}\left(\rho_\zeta^*\frac{\partial l}{\partial Y}\right) + ad_\xi^*\,\rho_\zeta^*\,\frac{\partial l}{\partial Y} + \rho_{\frac{d}{dt}\zeta}^*\frac{\partial l}{\partial Y} - \rho_\Gamma^*\,\frac{\partial l}{\partial \Gamma} = 0 \tag{3.31}$$

on $\mathfrak{s} \times V$. Now write things in terms of the constrained reduced Lagrangian. Using the definition

$$l_c(\xi, \Gamma) = l(\xi, \xi\zeta(\Gamma), \Gamma),$$

one has the relation:

$$\frac{\partial l}{\partial \xi} = \frac{\partial l_c}{\partial \xi} - \rho_\zeta^*\frac{\partial l}{\partial Y}.$$

Therefore (3.31) is equal to:

$$\frac{d}{dt}\frac{\partial l_c}{\partial \xi} - ad_\xi^*\frac{\partial l_c}{\partial \xi} + \rho_\Gamma^*\frac{\partial l}{\partial \Gamma} = \rho_{\frac{d}{dt}\zeta}^*\frac{\partial l}{\partial Y}.$$

Assuming this final equation, the relationships between the derivatives of l_c and l, and the independence of η, one has that (3) implies (2). □

We remark that in calculating the δY variation, one varies $Y = g^{-1}\dot y$ before substituting in the constraint $\dot y = \omega g^{-1}\zeta(g^{-1}a_0)$. This is consistent with the Lagrange-d'Alembert principle which computationally requires variations be taken along the curve before the constraints are imposed.

Proof of Theorem 3.1. Theorem 3.1 is obtained as a special case of Theorem 3.3 as follows. Start with an intermediate system such that the Lagrangian is independent of a_0 and S-invariant, and form $\mathcal{D}$ by taking $\zeta(a) = a$. As discussed in Section 3.1, $\mathcal{D}$ then has integral submanifolds. One was taken to be the configuration space, $Q = G \times W$, where $W = \text{range}(\rho_{a_0})$. Theorem 3.1 reduces L and $\mathcal{D}$ restricted to TQ. This theorem is implied by Theorem 3.3 by observing that the Lagrange-d'Alembert principle for $(L, \mathcal{D})$ on TQ is equivalent to the Lagrange-d'Alembert principle for the system on TS when a restricted class of curves is used in S, namely curves that lie in Q. Variations used in TQ and TS take δg to be free, and $\delta y = \delta g g^{-1} a_0$. However, this is the same set of variations, as $\delta y \in V$ is restricted to lie in W. A simple way to characterize the restricted class of curves is to take the initial condition to lie in Q. Then the condition $\dot{y} = \omega a_0$ implies the solutions stay in Q. Statements (2) and (3) in Theorem 3.3 are equivalent to statements (2) and (3) for Theorem 3.1 when one starts with a S-invariant Lagrangian independent of a_0. □

3.2.2 *Examples*

We shall demonstrate the Intermediate Theorem by deriving the equations of motion for Chaplygin's top and the rolling disk. Both are rigid bodies rolling on the plane with configuration space $SO(3) \times \mathbb{R}^2$. Much of the notation used to describe Chaplygin's sphere in Section 3.1.3 is used for these systems. The motion is given by a curve $(A(t), x(t)) \in SE(3)$ where $A(t) \in SO(3)$ and $x(t) \in \mathbb{R}^3$. An element of $SE(3)$ maps from a reference coordinate system for the ball or disk to an inertial coordinate system for space. The origin of the reference system coincides with the center of mass for the body, so that $x(t)$ gives the location of the center of mass in space. A generic point in the body, given by some $P \in \mathbb{R}^3$ in the reference coordinate system, is located at $A(t)P + x(t)$ at time t. A tangent vector in $T_{(A,x)}SE(3)$ is denoted $(\dot{A}, \dot{x})$ and $\omega, \Omega \in \mathbb{R}^3$ are defined by $\widehat{\omega} = \dot{A}A^{-1}$ and $\widehat{\Omega} = A^{-1}\dot{A}$. The vertical upward unit vector, normal to the plane, is denoted by e_3. Let $z = \langle x, e_3 \rangle$ denote the height of the center of mass of the ball; a holonomic constraint for z is derived in both Chaplygin's top and the rolling disk. We assume the reference coordinate system diagonalizes the inertia tensor for the body, denoted by I, and let I_1, I_2 and I_3 denote the three principal moments of inertia.

Chaplygin's top differs from Chaplygin's sphere in that the center of mass may reside anywhere in the body, as opposed to at the geometric center. Let m denote the total mass of the ball and r the radius of the ball. The constraints are derived as before; the velocity of the contact point is zero and this velocity is the sum of velocity of the center of mass and the rotational velocity of the contact point about the center of mass. This rotational velocity is not simply $\omega \times -re_3$ as with Chaplygin's sphere. Let $\ell\chi$ denote the vector in the reference coordinate system that goes from the center of the ball to the center of mass, where ℓ is the distance between these two points and χ is a unit vector. Then in the inertial frame, the vector going from the center of mass to the contact point is $-\ell A\chi - re_3$, hence the constraint equation is

$$\dot{x} = \omega \times (re_3 + \ell A\chi), \tag{3.32}$$

or equivalently,

$$\dot{x} = \omega \times A(rA^{-1}e_3 + \ell\chi). \tag{3.33}$$

The distribution, $\mathcal{D}$, defined by this constraint is of the form required for Theorem 3.3 by choosing the fixed vector $a_0 \in V$ to be e_3 and

$$\zeta : \mathbb{R}^3 \to \mathbb{R}^3 \text{ to be } \zeta(a) = ra + \ell\chi.$$

The holonomic constraint for z is obtained by dotting both sides of (3.33) with e_3, using the identity $\omega \times v = \dot{A}A^{-1}v$ and then integrating. One obtains

$$z = \langle e_3, \ell A\chi\rangle + K,$$

where K is an arbitrary constants. We take K to be zero which corresponds to having the $z = 0$ plane go through the center of the ball.

As with Chaplygin's sphere, the kinetic energy for Chaplygin's top comes from the translation of the center of mass and the rotation about the center of mass. The potential energy due to gravity is mgz, where g is the constant acceleration due to gravity. Using the holonomic constraint, we have the following Lagrangian

$$L = \frac{1}{2}\{\langle \Omega, I\Omega\rangle + m\|\dot{x}\|^2\} - mg\ell\langle e_3, A\chi\rangle.$$

Both L and $\mathcal{D}$ depend on the fixed parameter $e_3 \in \mathbb{R}^3$. Define $\bar{L} : TSE(3) \times \mathbb{R}^3 \to \mathbb{R}$ by letting e_3 vary in L. We check that $\bar{L}$ is invariant under the action Φ_S described by equation (3.22). For these spaces, the action of $(B, y) \in SE(3)$ on $TSE(3) \times V$ is

$$(B, y) \cdot (\Omega, \dot{x}, A, x, e_3) = (\Omega, B\dot{x}, BA, Bx + y, Be_3).$$

The kinetic energy is invariant under this action. Observing that

$$mg\ell\langle Be_3, BA\chi\rangle = mg\ell\langle e_3, A\chi\rangle$$

shows that the potential energy is invariant as well. Hence Theorem 3.3 applies. The symmetry group acting on $SE(3)$ is $SO(2) \times \mathbb{R}^3$ and the reduced space is $\mathfrak{so}(3) \times S^2$. We proceed to compute the reduced equation

$$\frac{d}{dt}\frac{\partial l_c}{\partial \Omega} - ad^*_\Omega \frac{\partial l_c}{\partial \Omega} + \rho^*_\Gamma \frac{\partial l}{\partial \Gamma} = \rho^*_{\frac{d}{dt}\zeta}\frac{\partial l}{\partial Y}.$$

The reduced Lagrangian is:

$$l(\Omega, Y, \Gamma) = \frac{1}{2}\{\langle \Omega, I\Omega\rangle + m\|Y\|^2\} - mg\ell\langle \Gamma, \chi\rangle,$$

and the constrained reduced Lagrangian is:

$$l_c(\Omega, \Gamma) = \frac{1}{2}\{\langle \Omega, I\Omega\rangle + m\|\Omega \times \zeta\|^2\} - mg\ell\langle \Gamma, \chi\rangle.$$

The relevant derivatives are computed to be:

$$\begin{aligned}\frac{\partial l_c}{\partial \Omega} &= I\Omega + m\zeta \times (\Omega \times \zeta)\\ \frac{\partial l}{\partial \Gamma} &= -gm\ell\chi\\ \frac{\partial l}{\partial Y} &= m\Omega \times \zeta.\end{aligned}$$

Thus the reduced equation is

$$\frac{d}{dt}\{I\Omega + m\zeta \times (\Omega \times \zeta)\} - (-\Omega \times [I\Omega + m\zeta \times (\Omega \times \zeta)]) + (-(-mg\ell\chi \times \Gamma)) =$$

$$-m(\Omega \times \zeta) \times \frac{d}{dt}\zeta.$$

Using the identity $v \times (w \times (v \times w)) = \langle v, w\rangle (w \times v)$ one obtains Routh's classic equations of motion

$$\frac{d}{dt}\{I\Omega + m\zeta \times (\Omega \times \zeta)\} = I\Omega \times \Omega + m(\frac{d}{dt}\zeta) \times (\Omega \times \zeta) + m\langle \zeta, \Omega\rangle \Omega \times \zeta + mg\Gamma \times \ell\chi.$$

See [C] for a different derivation of these equations.

The rolling disk refers to a flat disk with homogeneous mass distribution that rolls without slipping on the plane while under the influence of gravity. Let m denote the total mass of the disk, and r the radius of the disk. Let E_1, E_2, E_3 denote coordinate vectors in the reference system, and e_1, e_2, e_3 coordinate vectors in the inertial system. We assume e_1, e_2 spans the horizontal plane upon which the disk rolls. As before, $z = \langle x, e_3\rangle$ denotes the height of the center of mass of the disk in the inertial system. In the reference coordinate system, we assume the origin is at the center of the disk, and that the disk lies in the E_1, E_2 plane. In the spatial frame, AE_3 will be normal to the disk. If AE_3 is $\pm e_3$, then the disk is no longer rolling and lies flat on the plane. We exclude such cases from the configuration space by omitting those $A \in SO(3)$ such that $AE_3 = \pm e_3$.

The constraint is determined by the condition that the contact point has zero velocity. This velocity is the sum of the velocity of the center of mass and the angular velocity of the contact point about the center of mass. Let ζ denote the vector in the reference coordinate system that goes from the contact point to the center of mass; note that ζ will depend on the orientation of the disk. Let v be a vector tangent to the boundary of the disk at the contact point in the inertial frame. The plane perpendicular to v is spanned by e_3 and AE_3. To obtain an expression for ζ, we note that since the center of mass for the disk is at the geometric center of the disk, $A\zeta$ is perpendicular to v, hence it is in the span of e_3 and AE_3. Furthermore, it is perpendicular to AE_3 and has length r. Define u in the reference coordinate system by

$$Au = e_3 - \langle e_3, AE_3\rangle AE_3.$$

Let $\Gamma = A^{-1}e_3$, then we may write

$$u = \Gamma - \langle \Gamma, E_3 \rangle E_3. \tag{3.34}$$

As ζ has length r,

$$\zeta = r\frac{u}{\|u\|}.$$

Hence the constraints take the form

$$\dot{x} = \omega \times A\zeta(A^{-1}e_3), \tag{3.35}$$

where $\zeta : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ is defined by composition with (3.34). Equation (3.35) defines a distribution of the form required for Theorem 3.3.

To obtain the holonomic constraint for z, one takes the e_3 component of (3.35) and applies the identity $\omega \times v = \dot{A}A^{-1}v$ to obtain

$$\dot{z} = \langle \dot{A}\zeta(\Gamma), e_3 \rangle = \frac{d}{dt}(\langle A\zeta, e_3 \rangle) - \langle A\dot{\zeta}, e_3 \rangle.$$

It is straightforward to show that $\langle \dot{\zeta}, \Gamma \rangle = 0$. This implies that $z = \langle \zeta, \Gamma \rangle + K$ where K is some constant. We take K to be zero which corresponds to having the $z = 0$ plane coincide with the plane upon which the disk rolls.

The Lagrangian takes the form of kinetic energy minus potential energy. The potential energy is proportional to the height of the center of mass. Using the holonomic constraint for z, we have

$$L = \frac{1}{2}\left\{\langle \Omega, I\Omega \rangle + m\|\dot{x}\|^2\right\} - mg\langle \zeta, A^{-1}e_3 \rangle.$$

Let $\bar{L}$ be defined on $TS \times V$ by letting e_3 vary. (Note that e_3 varies in the argument to ζ as well.) By following the argument with Chaplygin's top, $\bar{L}$ is readily seen to be invariant under the action Φ_S in equation 3.22. Therefore L and $\mathcal{D}$ satisfy the hypotheses of Theorem 3.3. The symmetry group acting on $SE(3)$ is $SO(2)\circledS\mathbb{R}^3$ and the reduced space is $\mathfrak{so}(3) \times S^2$. To compute the reduced equations, we first note that the reduced Lagrangian and constrained reduced Lagrangians are

$$l(\Omega, Y, \Gamma) = \frac{1}{2}\left\{\langle \Omega, I\Omega \rangle + m\|Y\|^2\right\} - mg\langle \zeta, \Gamma \rangle, \text{ and}$$

$$l_c(\Omega, \Gamma) = \frac{1}{2}\left\{\langle \Omega, I\Omega\rangle + m\|\Omega \times \zeta\|^2\right\} - mg\langle \zeta, \Gamma\rangle.$$

The derivatives required for the reduced equation are:

$$\frac{\partial l_c}{\partial \Omega} = I\Omega + m\,\zeta \times (\Omega \times \zeta)$$

$$\frac{\partial l}{\partial Y} = mY = m\Omega \times \zeta.$$

To compute $\partial l/\partial \Gamma$, we note that

$$\langle \zeta, \Gamma\rangle = r(\Gamma_1^2 + \Gamma_2^2)^{\frac{1}{2}},$$

therefore

$$\frac{\partial l}{\partial \Gamma} = -mg\frac{r}{(\Gamma_1^2 + \Gamma_2^2)^{\frac{1}{2}}}\begin{bmatrix}\Gamma_1 \\ \Gamma_2 \\ 0\end{bmatrix} = -mg\zeta.$$

Hence the reduced equation is

$$\frac{d}{dt}(I\Omega + m\zeta \times (\Omega \times \zeta)) + \Omega \times I\Omega + m\Omega \times (\zeta \times (\Omega \times \zeta)) - mg\Gamma \times \zeta = m\dot{\zeta} \times (\Omega \times \zeta).$$

For a thorough investigation of the rolling disk, see [CHK]. In particular, these same equations are obtained in a different fashion.

We remark on the difference between the configuration spaces for the basic and intermediate systems. Though both systems are described using a semidirect product S, the configuration space of the basic case is $G \times W$ where W = the range of ρ_{a_0} while the configuration space for the intermediate case is taken to be S. This is a consequence of the distribution defined by the constraint equation $\dot{y} = \omega a_0$ having integral submanifolds determined by the range of ρ_{a_0}. However, the two examples of the intermediate theorem have configuration space $G \times W$ and the distributions have integral submanifolds identified by the holonomic constraint on z. One may ask if any choice of ζ in a distribution defined by the constraint equation $\dot{y} = \omega g\zeta(g^{-1}a_0)$ will lead to integral submanifolds of $\mathcal{D}$. The answer is no: when $S = SE(3)$ and $a_0 = e_3$ as with the examples, one can construct functions ζ such that $\mathcal{D}$ has no integral submanifolds.

3.3 General Case

The distributions constructed for the basic and intermediate systems use the induced action of $\mathfrak{g}$ on V to generalize constraints that are typical for rigid bodies rolling on the plane. The dependence of these distributions on a_0 is such that they are invariant under the action of J and reduce to $\mathfrak{g} \times M$. In this section we formulate the general property required for a distribution on TS to be J-invariant and permit reduction to $\mathfrak{g} \times M$ as with the distributions of the prior cases. Taking a Lagrangian as described for the intermediate case, we form what we call the general case and then state a reduction theorem which gives the reduced Euler-Poincaré equation for the system. The conservation of energy for the basic, intermediate and general systems will then be established. Finally we briefly discuss how these equations compare to the equations in [BKMM] for a general nonholonomic system.

We assume the notation for discussing semidirect products given in Section 3.1. The configuration space is taken to be the semidirect product S. The distribution $\mathcal{D}$ is determined by the constraint equation

$$\dot{y} = A(g, y, a_0)\dot{g}, \tag{3.36}$$

where $A(g, y, a_0) : T_gG \to V$ is linear and satisfies the following invariance property. For any $(h, v) \in S$ one has that

$$A(g, y, a_0)\dot{g} = h^{-1}A(hg, hy + v, ha_0)h\dot{g}. \tag{3.37}$$

at all $(g, y) \in S$ and $a_0 \in V$. This invariance condition implies that $A(g, y, a_0)$ is independent of y, so we will write $A(g, a_0)$. More importantly, the condition implies that the constraints can be reduced to $\mathfrak{s} \times M$.

The Lagrangian for the general system has the same invariance property as in the intermediate case. That is $L : TS \to \mathbb{R}$ is defined by $L(g, \dot{g}, \dot{y}) = \bar{L}(g, \dot{g}, \dot{y}, a_0)$ where $\bar{L} : TS \times V \to \mathbb{R}$ is invariant under the action Φ_S of S on $TS \times V$ given in equation (3.22). We shall call $(S, L, \mathcal{D})$ the general system.

As with the intermediate and basic systems, the general system on TS whose dynamics are determined by the Lagrange d'Alembert principle is used to define a reduced system on $\mathfrak{s} \times M$ for which we have a constrained principle to determine the reduced dynamics. An equivalence is established between the curves $(g(t), y(t)) \in S$ satisfying the Lagrange-d'Alembert principle for the general system and the curves $(\xi(t), Y(t), \Gamma(t)) \in \mathfrak{s} \times M$ satisfying the constrained principle.

The relationship between the curves is as before. A curve $(g(t), y(t)) \in S$ determines $(\xi(t), Y(t), \Gamma(t)) \in \mathfrak{s} \times M$ by

$$\xi(t) = g^{-1}(t)\dot{g}(t) \tag{3.38}$$

$$Y(t) = g^{-1}(t)\dot{y}(t), \tag{3.39}$$

and $\Gamma(t) = g^{-1}a_0$. From a curve $(\xi(t), Y(t), \Gamma(t)) \in \mathfrak{s} \times M$ where $\Gamma(t)$ is assumed to satisfy $\dot{\Gamma}(t) = -\xi(t)\Gamma(t)$, $(g(t), y(t)) \in S$ is determined by the above differential equations.

The reduced constraint equation is determined as follows. For $\Gamma \in V$, define

$$a(\Gamma) : \mathfrak{g} \longrightarrow V,$$

by $a(\Gamma)\xi = A(e, \Gamma)\xi$. Then the invariance condition (3.37) implies that the constraint equation (3.36) is equivalent to

$$Y = a(\Gamma)\xi,$$

where we are using the relationship between the curves above. The reduced Lagrangian $l : \mathfrak{s} \times M \longrightarrow \mathbb{R}$ is defined by

$$l(\xi, Y, \Gamma) = \bar{L}(e, \xi, Y, \Gamma).$$

One then has that

$$\bar{L}(g, \dot{g}, y, \dot{y}, b) = l(g^{-1}\dot{g}, g^{-1}\dot{y}, g^{-1}b), \tag{3.40}$$

for $b \in M$. The constrained reduced Lagrangian $l_c : \mathfrak{g} \times M \longrightarrow \mathbb{R}$ is defined by evaluating l on the constraints:

$$l_c(\xi, \Gamma) = l(\xi, a(\Gamma)\xi, \Gamma).$$

We will make use of the following notation. For $v, \Gamma \in V$ define $D_v a(\Gamma) \in End(\mathfrak{g}, V)$ by

$$(D_v a(\Gamma))\xi = \frac{d}{dt}(a(\Gamma(t))\xi) \,|_{t=0},$$

where $\Gamma(0) = \Gamma$ and $\dot{\Gamma}(0) = v$. For $\xi \in \mathfrak{g}$, let $Da(\Gamma)\xi \in End(V, V)$ be defined similarly as

$$(Da(\Gamma)\xi)v = \frac{d}{dt}\,|_{t=0}\, a(\Gamma(t))\xi,$$

where $\Gamma(0) = \Gamma$ and $\dot{\Gamma}(0) = v$.

Theorem 3.4. General Theorem.

With the preceding notation, the following are equivalent:

(1) The curve $(g(t), y(t)) \in S$ satisfies the Lagrange-d'Alembert principle for the general system. That is $(\dot{g}, \dot{y}) \in \mathcal{D}_{(g(t), y(t))}$ and

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

where $\delta g(t)$ is an independent variation vanishing at the endpoints, and $\delta y = A(g, a_0)\delta g$.

(2) The curve $(\xi(t), Y(t), \Gamma(t))$, where $\dot{\Gamma} = -\xi(t)\Gamma(t)$, satisfies the following constrained principle on $\mathfrak{s} \times M$. One has $Y(t) = a(\Gamma(t))\xi(t)$, and

$$\delta \int l(\xi(t), Y(t), \Gamma(t)) = 0,$$

where the variations along the curve take the form:

$$\delta\xi = \dot{\eta} + ad_\xi \eta$$

$$\delta\Gamma = -\eta\,\Gamma$$

$$\delta Y = \frac{d}{dt}(a(\Gamma)\eta) + \xi\, a(\Gamma)\eta - \eta\, a(\Gamma)\xi,$$

and $\eta(t)$ is an independent variation vanishing at the endpoints.

(3) The Euler-Poincaré equation

$$\frac{d}{dt}\frac{\partial l_c}{\partial \xi} - ad^*_\xi \frac{\partial l_c}{\partial \xi} + \rho^*_\Gamma \frac{\partial l}{\partial \Gamma} = \left\{ a(\Gamma)^* \sigma^*_\xi - \rho^*_{a(\Gamma)\xi} - ad^*_\xi\, a(\Gamma)^* + (D_{(\frac{d}{dt}\Gamma)} a(\Gamma))^* \right\} \frac{\partial l}{\partial Y} \tag{3.41}$$

and advection equation

$$\dot{\Gamma} + \xi\,\Gamma = 0$$

holds on $\mathfrak{g} \times M$, where Y is evaluated at $a(\Gamma)\,\xi$ in the argument to $\partial l/\partial Y$ and $\partial l/\partial \Gamma$.

The proof goes much the same as with Theorem 3.3 and will be omitted.

All nonholonomic systems conserve energy. For the general, basic and intermediate systems we have reduced, the energy may be written as the following function on $\mathfrak{g} \times M$:

$$E = \langle \frac{\partial l_c}{\partial \xi}, \xi \rangle - l_c.$$

As the basic and intermediate systems can be obtained from the general system by choosing $A(g, a_0)$ appropriately, it suffices to check that energy is conserved for the general system. By differentiating $l_c(\xi, \Gamma) = l(\xi, a(\Gamma)\xi, \Gamma)$ with respect to Γ, we have the following equation:

$$\frac{\partial l_c}{\partial \Gamma} = \frac{\partial l}{\partial \Gamma} + (Da(\Gamma)\xi)^* \frac{\partial l}{\partial Y}. \tag{3.42}$$

Using the reduced equation (3.41) and equation (3.42) we now show that $\frac{d}{dt}E = 0$.

$$\begin{aligned} \frac{d}{dt}E &= \langle \frac{d}{dt}\frac{\partial l_c}{\partial \xi}, \xi \rangle - \langle \frac{\partial l_c}{\partial \Gamma}, \dot{\Gamma} \rangle \\ &= \langle ad_\xi^* \frac{\partial l_c}{\partial \xi} - \rho_\Gamma^* \frac{\partial l}{\partial \Gamma} + \left\{ a(\Gamma)^* \sigma_\xi^* - \rho_{a(\Gamma)\xi}^* - ad_\xi^* a(\Gamma)^* + (D_{\frac{d}{dt}\Gamma} a(\Gamma))^* \right\} \frac{\partial l}{\partial Y}, \xi \rangle - \langle \frac{\partial l_c}{\partial \Gamma}, \dot{\Gamma} \rangle. \end{aligned}$$

As $ad_\xi \xi = 0$ and $\sigma(\xi)(a(\Gamma)\xi) = \rho_{a(\Gamma)\xi}\xi$, we are left with

$$\frac{d}{dt}E = \langle -\rho_\Gamma^* \frac{\partial l}{\partial \Gamma} + (D_{\frac{d}{dt}\Gamma} a(\Gamma))^* \frac{\partial l}{\partial Y}, \xi \rangle - \langle \frac{\partial l_c}{\partial \Gamma}, \dot{\Gamma} \rangle.$$

Using equation (3.42) and $\dot{\Gamma} = -\xi\Gamma$, one has that

$$\frac{d}{dt}E = \langle \{ (D_{\frac{d}{dt}\Gamma} a(\Gamma))^* + \rho_\Gamma^* (Da(\Gamma)\xi)^* \} \frac{\partial l}{\partial Y}, \xi \rangle$$

which is zero.

One may rewrite the general reduced equation (3.41) in terms of $\partial l_c / \partial \Gamma$ as opposed to $\partial l / \partial \Gamma$ by using equation (3.42). The resulting equation compares more directly with the equations of motion derived in [BKMM], as the latter equations are mostly in terms of the constrained Lagrangian. We have chose to formulate the reduced equations for the basic, intermediate and general systems in terms of $\partial l / \partial \Gamma$ as it is simpler to compute.

Chapter 4

STABILITY OF RELATIVE EQUILIBRIA IN CHAPLYGIN'S SPHERE

A fundamental question to ask when investigating the dynamics of a mechanical system with symmetry is what are the relative equilibria and are they stable. Relative equilibria are motions of the dynamical system that correspond to one parameter subgroups of the symmetry group. Equilibrium solutions to the reduced equations are relative equilibria for a given system. We say a relative equilibrium is stable if it is Lyapunov stable in the reduced space. For precise definitions and further background information on stability, see [MR, pp. 25-36]. A framework for determining the stability of relative equilibria for nonholonomic systems, which is based upon the reduced equations of motion developed in [BKMM], is found in [ZBM]. Several general conditions for stability are presented in [ZBM]. However, when the eigenvalues of the linearized equations of motion are all imaginary, a general theorem about stability is not known. An example where this occurs is in the study of Chaplygin's sphere. The main results of this chapter is Theorem 4.1 which determines the stability of relative equilibria in Chaplygin's sphere. A crucial ingredient will be the conserved quantities for the system described in Section 3.1.3.

A description of Chaplygin's sphere and derivation of the reduced equations of motion using the Basic Theorem (Theorem 3.1) is found in Section 3.1.3. We restate the reduced equations before proceeding with the stability analysis. The reduced equations determine a curve $(\Omega(t), \Gamma(t)) \in \mathbb{R}^3 \times \mathbb{R}^3$ where $\Omega(t)$ is the angular velocity of the ball relative to a reference coordinate system for the ball, and $\Gamma(t)$ is the position of e_3 (unit upward vertical vector) in the same reference coordinate system. The reduced equations of motion are

$$\dot{M} + \Omega \times M = 0 \tag{4.1}$$

$$\dot{\Gamma} + \Omega \times \Gamma = 0 \tag{4.2}$$

where

$$M = I\Omega + mr^2\Gamma \times (\Omega \times \Gamma), \tag{4.3}$$

and m, r, and I are the total mass, radius and inertia tensor for the ball. As noted in Section 3.1.3, the equations conserve the following four functions:

$$C_1 = \langle M, M\rangle, \quad C_2 = \langle M, \Gamma\rangle, \quad C_3 = \langle \Gamma, \Gamma\rangle = 1 \quad \text{and} \quad C_4 = \langle M, \Omega\rangle. \tag{4.4}$$

The first two are a consequence of Theorem 3.2 which implies that the angular momentum of the ball in space is conserved. The third follows from Γ being in S^2 and the fourth is twice the energy of the system. By rewriting equation (4.1), one has the following system:

$$I(\Gamma)\dot{\Omega} + \Omega \times I\Omega = 0 \tag{4.5}$$

$$\dot{\Gamma} + \Omega \times \Gamma = 0, \tag{4.6}$$

where $I(\Gamma) : \mathbb{R}^3 \to \mathbb{R}^3$ is defined by $I(\Gamma)v = Iv + mr^2(\Gamma \times (v \times \Gamma))$.

Suppose (Ω_0, Γ_0) is an equilibrium solution to equations (4.5) and (4.6). Then equation (4.5) implies that $I\Omega_0 \times \Omega_0 = 0$, hence Ω_0 is an eigenvector of I, meaning it coincides with a principal axis. Equation (4.6) implies that $\Omega_0 \times \Gamma_0 = 0$, hence Ω_0 coincides with the vertical axis. As relative equilibria, the motion generated are steady rotations about a vertical, principal axis. We assume $AE_3 = e_3$, where $A(t) \in SO(3)$ gives the orientation of the ball in space, E_3 is a principal axis in the body frame, and $e_3 = A\Gamma$ is the upward unit vertical normal to the horizontal plane. We also assume I is diagonalized and denote the principal moments of inertia by I_1, I_2, I_3, where I_3 is the moment corresponding to E_3, the axis of rotation for the ball. The stability result is the following.

Theorem 4.1. Chaplygin's Sphere Stability Theorem. *Let* $(\Omega_0, \Gamma_0) = (ZE_3, E_3)$ *be an equilibrium solution to the reduced equations (4.5) and (4.6) for Chaplygin's sphere, as described above. Assume* $Z \neq 0$. *Let* I_1, I_2 *and* I_3 *denote the moments of inertia around the three principal axes. Assume the third axis coincides with* e_3 *in space. Then in the sense of Lyapunov,*

(1) The solution is unstable if I_3 *is the middle axis, that is* $I_1 < I_3 < I_2$ *or* $I_2 < I_3 < I_1$.

(2) The solution is stable if I_3 is the greatest or smallest of the moments of inertia, that is $I_3 < I_1, I_2$ or $I_3 > I_1, I_2$.

Proof. Statement (1) is shown by linearizing the system around an equilibrium. Some constants that will re-occur in the analysis are:

$$a_1 = I_1 - I_3 \qquad a_2 = I_2 - I_3$$
$$I_1' = I_1 + mr^2 \quad I_2' = I_2 + mr^2 \quad I_3' = I_3 + mr^2.$$

The linearized equations at the equilibrium $(\Omega_0, \Gamma_0) = (ZE_3, E_3)$ take the form

$$\frac{d}{dt}\begin{bmatrix} \Omega_1' \\ \Omega_2' \\ \Omega_3' \\ \Gamma_1' \\ \Gamma_2' \\ \Gamma_3' \end{bmatrix} = \begin{pmatrix} 0 & a_2 Z/I_1' & 0 & 0 & 0 & 0 \\ -a_1 Z/I_2' & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & Z & 0 \\ 1 & 0 & 0 & -Z & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{bmatrix} \Omega_1' \\ \Omega_2' \\ \Omega_3' \\ \Gamma_1' \\ \Gamma_2' \\ \Gamma_3' \end{bmatrix}$$

where $\Omega' = \Omega - ZE_3$ and $\Gamma' = \Gamma - E_3$. The characteristic polynomial for the above matrix is

$$p(\lambda) = \lambda^6 + (Z^2 - Z^2\frac{a_1 a_2}{I_1' I_2'})\lambda^4 - \frac{a_1 a_2}{I_1' I_2'}\lambda^2,$$

from which one finds that the eigenvalues λ satisfy

$$\lambda^2 = 0, \; \lambda^2 = -Z^2, \; \lambda^2 = \frac{-a_1 a_2 Z^2}{I_1' I_2'}. \tag{4.7}$$

If I_3 is the middle principal moment of inertia, a_1 and a_2 have opposite signs. Then λ^2 is positive and there is an eigenvalue with positive real part. Hence the system is unstable and (1) is proved.

Statement (2) is shown by constructing a Lyapunov function out of the conserved quantities. The result of composing an arbitrary function of four variables with the constants of motion is a conserved quantity along the flow. If it also has a non-degenerate minimum at the equilibrium, it will be a Lyapunov function. We shall show such a Lyapunov function exists. To this end, define

$$\tilde{C}_i(\Omega, \Gamma) = C_i(\Omega, \Gamma) - C_i(\Omega_0, \Gamma_0),$$

where the functions C_i are as in equation (4.4). Now consider an arbitrary function on $\mathbb{R}^4$, $F(\widetilde{C}_1, \widetilde{C}_2, \widetilde{C}_3, \widetilde{C}_4)$, such that $F(0) = 0$. Define $H : \mathbb{R}^6 \to \mathbb{R}$ by $H = F \circ \widetilde{C}_i(\Omega, \Gamma)$. We see that H is conserved along the flow. To show that an F may be chosen so that H has a non-degenerate minimum at (Ω_0, Γ_0), H is expanded as a power series around (Ω_0, Γ_0) in the variables $\Omega' = \Omega - \Omega_0$ and $\Gamma' = \Gamma - \Gamma_0$. Denote the expansion of each $\widetilde{C}_i$ by

$$\widetilde{C}_i = 0 + C_i(1) + C_i(2) + \dots ,$$

where $C_i(1)$ is linear in (Ω', Γ'), and $C_i(2)$ quadratic. The expansion of H is

$$H = 0 + h^i C_i(1) + h^i C_i(2) + \frac{1}{2} h^{ij} C_i(1) C_j(1) + \dots ,$$

$$\text{where } h^i = \frac{\partial F}{\partial \widetilde{C}_i}(0) \text{ and } h^{ij} = \frac{\partial^2 F}{\partial \widetilde{C}_i \partial \widetilde{C}_j}(0);$$

here we have adopted the summation convention of summing over a repeated index. In order for H to have a non-degenerate minimum, the h^i and h^{ij} must be chosen so that $h^i C_i(1)$ is 0, and the quadratic form $h^i C_i(2) + \frac{1}{2} h^{i\,j} C_i(1) C_j(1)$ in the variables (Ω', Γ') is positive definite.

The following notation is used in the rest of the expansion. Let I' be defined by:

$$I' = \begin{pmatrix} I_1 + mr^2 & 0 & 0 \\ 0 & I_2 + mr^2 & 0 \\ 0 & 0 & I_3 \end{pmatrix}.$$

For $\mathbf{v} = (v_1, v_2, v_3)$, let $\widetilde{\mathbf{v}}$ denote $E_3 \times (\mathbf{v} \times E_3) = (v_1, v_2, 0)$.

To expand the C_i functions, first expand $M = I\Omega + mr^2\{\Gamma \times (\Omega \times \Gamma)\}$ as $M = M_0 + M_1 + M_2 + \dots$ to obtain the linear and quadratic terms in (Ω', Γ'). Then:

$$\begin{aligned} M_0 &= I_3 Z E_3 \\ M_1 &= I'\Omega' - mr^2 Z \widetilde{\Gamma}' \\ M_2 &= mr^2\{2E_3 \times (\Omega' \times \Gamma') + (Z\Gamma' - \Omega') \times (E_3 \times \Gamma')\}. \end{aligned}$$

Using this, the expansion of the C_i functions are:

$$
\begin{array}{ll}
C_1(1) = 2(I_3)^2 Z\Omega_3'. & C_1(2) = 2I_3Z\langle E_3, M_2\rangle + \|I'\Omega'\|^2 - 2mr^2\langle \tilde{I}'\Omega', Z\tilde{\Gamma}'\rangle + (mr^2Z)^2\|\tilde{\Gamma}'\|^2 \\
C_2(1) = I_3Z\,\Gamma_3' + I_3\Omega_3' & C_2(2) = \langle \Gamma', I'\Omega'\rangle - Zmr^2\|\tilde{\Gamma}'\|^2 + \langle E_3, M_2\rangle \\
C_3(1) = 2\Gamma_3' & C_3(2) = \|\Gamma'\|^2 \\
C_4(1) = 2I_3Z\,\Omega_3' & C_4(2) = Z\,\langle E_3, M_2\rangle + \langle \Omega', I'\Omega'\rangle - Zmr^2\langle \tilde{\Omega}', \tilde{\Gamma}'\rangle.
\end{array}
\tag{4.8}
$$

The conditions for $h^i C_i(1) = 0$ to be true are:

$$
\begin{aligned}
2I_3Zh^1 + h^2 + 2Zh^4 &= 0 \\
I_3Zh^2 + 2h^3 &= 0.
\end{aligned}
$$

Setting $h^2 = -Z$, the conditions become

$$
h^3 = \frac{I_3Z^2}{2} \qquad h^4 = \frac{1}{2} - I_3h^1, \tag{4.9}
$$

where h^1 is a free parameter.

We now turn our attention to showing that h_1 and the h_{ij} may chosen so that $h^iC_i(2) + \frac{1}{2}h^{ij}C_i(1)C_j(1)$ is positive definite. After substituting the expressions for the $C_i(1)$, into $\frac{1}{2}h^{ij}C_i(1)C_j(1)$, it simplifies to be a quadratic form in (Ω_3', Γ_3')

$$
\alpha(\Omega_3')^2 + \beta\,\Omega_3'\cdot\Gamma_3' + \kappa(\Gamma_3')^2, \tag{4.10}
$$

where α, β and κ depend on h^{ij}. By choosing the h^{ij} appropriately, we can make this expression arbitrarily large. In fact, the sum of this second term and terms of the form (4.10) in $h^iC_i(2)$ can be made arbitrarily large by appropriate choices for the h^{ij}.

If h^1 can be chosen so that $h^iC_i(2)$ modulo terms of the form (4.10) is positive definite, then the proof of the theorem will be complete. To see if h^1 can be appropriately chosen, we substitute $h_2 = -Z$, the expressions for h^3 and h^4 determined by equations (4.9), and the expressions for $C_i(2)$ in equations (4.8)), into $h^iC_i(2)$ and then simplify. Ignoring terms of the form (4.10), the result is two quadratic forms in (Ω_1', Γ_1') and (Ω_2', Γ_2'). For $i = 1, 2$, they are:

$$
I_i'\left\{(I_i - I_3 + mr^2)h^1 + \frac{1}{2}\right\}\Omega_i'^2 - 2I_i'Z\left(mr^2h^1 + \frac{1}{2}\right)\Omega_i'\Gamma_i' + Z^2I_3'\left[mr^2h^1 + \frac{1}{2}\right]\Gamma_i'^2, \tag{4.11}
$$

where $I_i' = I_i + mr^2$. It suffices to show that each of the two forms is positive definite.

First assume $I_3 > I_1, I_2$; note that this implies $I_3' > I_i'$. The two forms will be shown to be positive definite by taking $h^1 = 0$. In this case, the forms reduce to:

$$I_i' \frac{1}{2} \Omega_i'^2 - 2 I_i' Z \frac{1}{2} \Omega_i' \Gamma_i' + Z^2 I_3' \frac{1}{2} \Gamma_i'^2.$$

Recall that an expression such as $aX^2 - bXY + cY^2$ is positive definite if the discriminant $4ac - b^2$ is greater then zero. Checking the discriminant gives:

$$I_i' I_3' Z^2 - (I_i')^2 Z^2 > 0 \quad \text{as } I_3' > I_i'.$$

Now assume $I_3 < I_1, I_2$. Choosing h^1 large enough ensures that each form in (4.11) is positive definite. To show this, it suffices to demonstrate that the forms

$$I_i' (I_i - I_3 + mr^2) h^1 \Omega_i'^2 - 2 I_i' Z\, mr^2 h^1 \Omega_i' \Gamma_i' + Z^2 I_3'\, mr^2 h^1 \Gamma_i'^2$$

are positive, we have dropped terms that are not linear in h^1. Factoring out the h^1 and taking the discriminant gives:

$$\begin{aligned} Z^2 mr^2 I_i'[(I_i - I_3 + mr^2) I_3' - I_i' mr^2] &= \\ Z^2 mr^2 I_i'[(I_i - I_3) I_3' + (I_3' - I_i') mr^2] &= \\ Z^2 mr^2 I_i'[(I_i - I_3) I_3' + (I_3 - I_i) mr^2] &= \\ Z^2 mr^2 I_i'[(I_i - I_3)(I_3' - mr^2)] &= \\ Z^2 mr^2 I_i'[(I_i - I_3) I_3], & \end{aligned}$$

which is positive as both $I_1 - I_3$ and $I_2 - I_3$ are positive.

This concludes the proof of theorem. □

Note that the ball is unstable when $Z = 0$. The ball may be perturbed to slowly roll away from a state of rest. For example, near the equilibrium $(0, E_3)$, the initial condition $(\epsilon e_1, E_3)$ with ϵ small will cause the ball to slowly roll around the e_1 axis, so $\Gamma(t)$ will escape a small neighborhood of E_3.

Chapter 5

STABILIZING CHAPLYGIN'S SPHERE

In the previous chapter, we saw that the rotation of Chaplygin's sphere about its middle axis is an unstable relative equilibrium. In this chapter we develop two different methods of stabilizing this motion using control forces. Due to the nature of the problems we consider, the main difficulty is generating a control force which induces stability, as opposed to linear stability. The latter meaning that the linearized equations are stable at the equilibrium. To obtain stability, we seek to construct a Lyapunov function for the relative equilibrium in the presence of a control force. The typical control force will break the conservation of energy in Chaplygin's sphere that was used to construct a Lyapunov function in Theorem 4.1. To find control laws for which the system will conserve an energy-like function, we generalize the method of controlled Lagrangians discussed in Chapter 2. This method proceeds by making changes to the Lagrangian of the uncontrolled system that one hopes will generate equations of motion matching those of the controlled system. Then the controlled system will conserve the energy corresponding to the modified Lagrangian, which may be used in the construction of a Lyapunov function.

The first method of stabilizing Chaplygin's sphere is to control an internal rotor which is added to the system. The coupling between the angular momentum of the ball and of the rotor gives us a means to influence the motion of the ball with that of the rotor. A very similar problem is stabilizing the rotation of the free rigid body about its middle axis. A solution to stabilizing the free rigid body, solved by the method of controlled Lagrangians, is presented in [BLMa] and reviewed in Chapter 2. To stabilize Chaplygin's sphere, we make use of the same modifications to the Lagrangian for the rigid body with internal rotor used in [BLMa]. This gives us a new system to which we may apply the Basic Theorem (Theorem 3.1). In this way, we obtain a control law which conserves an energy-like function for the system. Furthermore, to prove that the control law stabilizes the system, we need only invoke Theorem 4.1 of the previous chapter.

Our second method to stabilizing Chaplygin's sphere is by applying forces to the plane that the ball rolls upon. Using the nonholonomic Euler-Poincaré reduction theory, we formulate and study two problems based on this approach. For both problems, the control force applied to the plane is taken to be a horizontal force that translates but does not rotate the plane. We call the system for the first problem Chaplygin's sphere on the translating plane. The main result of Section 5.2 is a stabilizing control law for the middle axis rotation of Chaplygin's sphere in this setting. We use an intermediate system, as described in the notation for Theorem 3.3. that we use to match the controlled equations on the translating plane. As the intermediate system is similar to the basic system that describes Chaplygin's sphere, the calculation to show that our control law stabilizes the system is quite similar to the calculation showing stability in Theorem 4.1.

It is assumed that the plane can be held steady while the ball is rolling on the translating plane. The plane is only moved in response to the control force which translates it. A control mechanism which can hold the plane steady would have to be massive relative to the ball. Considering a control mechanism whose mass is small, or comparable to that of the ball leads to our second problem, Chaplygin's sphere on the sliding plane. To picture this, imagine the ball rolling on a platform that is free to rotate and translate in the horizontal plane. The platform is not held steady, but rather it is affected by the reaction force of the ball, which may rotate as well as translate the platform. An investigation of this system, which generalizes Theorem 3.3 in order to obtain equations of motion, is presented in Section 5.3. Finding a stabilizing control law in this setting is still an open problem; however we do show that the method used for the translating plane does not generalize.

5.1 Internal Rotor

To describe Chaplygin's sphere with an internal rotor we start with the notation previously outlined in Section 3.1.3. In particular, E_1, E_2, E_3 denote the principal axes in the body reference coordinate system and I_1, I_2, I_3 denote the principal moments of inertia about these axes. The vertical axis about which the ball spins unstably is E_3, so that either

$I_1 > I_3 > I_2$ or $I_1 < I_3 < I_2$. The internal rotor is added so that the ball and rotor are still balanced, meaning the center of mass for the ball and rotor is at the ball's center. For example, the rotor could be placed in the center of the ball, or two rotors which move in tandem could be added towards the top and bottom of an axis running through the ball. We consider the case when the rotor's axis of rotation coincides with E_1.

The configuration space for the ball with rotor is $Q = SO(3) \times S^1 \times \mathbb{R}^2$. The rotor's angle relative to axes in the reference coordinate system is given by $\theta \in S^1$. The constraints are the same as with Chaplygin's sphere:

$$\dot{x} = \omega \times r\, e_3, \tag{5.1}$$

where r is the radius of the ball. The Lagrangian consists only of kinetic energy as the center of mass for the ball with rotor is constant. Let J_1, J_2, J_3 denote the moments of inertia of the rotor about the three principal axes, and let $\lambda_i = I_i + J_i$. The Lagrangian for the uncontrolled system takes the form

$$L = \frac{1}{2}\{\lambda_1\Omega_1^2 + 2J_1\Omega_1\dot{\theta} + J_1\dot{\theta}^2 + \lambda_2\Omega_2^2 + \lambda_3\Omega_3^2 + m\|\dot{x}\|^2\},$$

where m is the combined mass of the ball and rotor.

We will apply the Basic Theorem (Theorem 3.1) to obtain equations of motion for the uncontrolled system. Let S be the semidirect product $G \circledS V$, where $G = SO(3) \times S^1$ and $V = \mathbb{R}^3$. The action of G on V is simply the usual action of $SO(3)$ on $\mathbb{R}^3$. An element of S is denoted by (A, θ, x), where $A \in SO(3)$ and $x \in \mathbb{R}^3$ gives the position and orientation of the ball as usual, and $\theta \in S^1$ measures the rotor angle as discussed above. The induced action of $\mathfrak{g} = \mathfrak{so}(3) \times \mathbb{R}$ on $\mathbb{R}^3$ is $(\Omega, a)\,x = \Omega \times x$. Let (Ω, a) and (Σ, b) be two elements of $\mathfrak{g}$. Then their commutator is

$$[(\Omega, a), (\Sigma, b)] = (\Omega \times \Sigma, 0).$$

The constraint equation (5.1) takes the form required for the Basic Theorem ($\dot{x} = \omega a_0$) if we take $a_0 = re_3$. Then $\omega \times r\, e_3$ corresponds to the induced action of $\mathfrak{g}$ on V given above. The Lagrangian is seen to be S-invariant, so Theorem 3.1 may be applied. The isotropy subgroup of a_0 is $H = SO(2) \times S^1$, hence the symmetry group is $J = H \circledS \mathbb{R}^2$ and the system reduces to $\mathfrak{g} \times S^2$.

Denote a vector in $\mathfrak{s}$ by $(\Omega, \dot{\theta}, Y)$, with $\widehat{\Omega} \in \mathfrak{so}(3), \dot{\theta} \in \mathbb{R}$, and $Y \in \mathbb{R}^3$. Let $\Gamma \in S^2$ denote $A^{-1}e_3$. The reduced constraint equation on $\mathfrak{s} \times S^2$ is

$$Y = r\Omega \times \Gamma,$$

as it was with Chaplygin's sphere. The reduced Lagrangian $l : \mathfrak{s} \to \mathbb{R}$ is

$$l(\Omega, \dot{\theta}, Y) = \frac{1}{2}\{\lambda_1\Omega_1^2 + 2J_1\Omega_1\dot{\theta} + J_1\dot{\theta}^2 + \lambda_2\Omega_2^2 + \lambda_3\Omega_3^2 + m\|Y\|^2\}.$$

The constrained reduced Lagrangian $l_c : \mathfrak{g} \times S^2 \to \mathbb{R}$ is

$$l_c(\Omega, \dot{\theta}, \Gamma) = \frac{1}{2}\left\{\langle \Omega, \lambda\Omega + mr^2\Gamma \times (\Omega \times \Gamma)\rangle + 2J_1\Omega_1\dot{\theta} + J_1\dot{\theta}^2\right\}.$$

Here λ is the diagonal matrix $(\lambda_1, \lambda_2, \lambda_3)$. Next we compute the reduced Euler-Poincaré equation (3.7) of the Basic Theorem. The result is

$$\dot{M} + [\xi, M] = 0.$$

where $\xi \in \mathfrak{g}$ has components $(\Omega, \dot{\theta})$ and M is the four-vector

$$\left(\frac{\partial l_c}{\partial \Omega}, \frac{\partial l_c}{\partial \dot{\theta}}\right).$$

Explicitly, the equations of motion are:

$$I(\Gamma)\dot{\Omega} + \begin{bmatrix} J_1\ddot{\theta} \\ 0 \\ 0 \end{bmatrix} + \Omega \times \left\{ \lambda\Omega + \begin{bmatrix} J_1\dot{\theta} \\ 0 \\ 0 \end{bmatrix} \right\} = 0 \tag{5.2}$$

$$\frac{d}{dt}\{J_1\dot{\theta} + J_1\Omega_1\} = 0 \tag{5.3}$$

$$\dot{\Gamma} + \Omega \times \Gamma = 0.$$

As with Chaplygin's sphere, $I(\Gamma) : \mathbb{R}^3 \to \mathbb{R}^3$ is defined by

$$I(\Gamma)v = \lambda v + mr^2\Gamma \times (v \times \Gamma).$$

The controlled equations are the same as the equations above except that a control torque u is added to the θ equation:

$$\frac{d}{dt}\{J_1\dot{\theta} + J_1\Omega_1\} = u.$$

Now we seek a class of controlled Lagrangians that generate these controlled equations. Because these equations are so similar to the uncontrolled equations for the free rigid body with internal rotor, it is simple to adapt the controlled Lagrangian for the free rigid body to this scenario. We consider the following family of Lagrangians parameterized by $\nu \in \mathbb{R}$:

$$l_\nu = \frac{1}{2}\{\lambda_1\Omega_1^2 + 2J_1\Omega_1\dot{\theta} + \frac{J_1}{1+\nu}\dot{\theta}^2 + \lambda_2\Omega_2^2 + \lambda_3\Omega_3^2 + m\|\dot{x}\|^2\}.$$

The system $(l_\nu, \mathcal{D}, Q)$ is seen to be an example of a basic system as well. By using Theorem 3.1, the reduced equations of motion are found to be

$$I(\Gamma)\dot{\Omega} + \begin{bmatrix} J_1\ddot{\theta} \\ 0 \\ 0 \end{bmatrix} + \Omega \times \left\{ \lambda\Omega + \begin{bmatrix} J_1\dot{\theta} \\ 0 \\ 0 \end{bmatrix} \right\} = 0 \tag{5.4}$$

$$\frac{d}{dt}\{J_1\dot{\theta} + J_1\Omega_1\} = -\nu J_1\dot{\Omega}_1 \tag{5.5}$$

$$\dot{\Gamma} + \Omega \times \Gamma = 0.$$

These match the controlled equations, with the torque on the rotor being $-\nu J_1\dot{\Omega}_1$. We first find conditions on ν which stabilize the system and then express this torque as a function of the velocity and position of the ball in order to have a viable feedback law. In doing so, we assume the θ momentum for l_ν is 0; that is $J_1(\dot{\theta} + (1+\nu)\Omega_1) = 0$. Physically, this means the rotor starts with no momentum of its own relative to the ball (however it does inherit the momentum of the ball). This is consistent with starting the system by applying an outside force to the ball. If the system starts with nonzero momentum p_θ for the rotor, we still expect a control force can be found to stabilize the system, however this control force will depend on p_θ.

When the θ momentum for l_ν is zero, one has

$$\dot{\theta} = -(1+\nu)\Omega_1.$$

Using this, one can eliminate $\dot{\theta}$ and $\ddot{\theta}$ from the equations (5.4) for the ball. Then the equations match those of Chaplygin's sphere (without an internal rotor), except that the moment of inertia around the first principal axis is changed. One has

$$\bar{I}(\Gamma)\dot{\Omega} = \bar{\lambda}\Omega \times \Omega,$$

where $\bar{\lambda}$ is the diagonal matrix $(I_1 - J_1\nu, \lambda_2, \lambda_3)$, and $\bar{I}(\Gamma)$ is defined by

$$\bar{I}(\Gamma)v = \bar{\lambda}v + mr^2\Gamma \times (v \times \Gamma).$$

Hence we can choose ν to force λ_3 to be the largest or smallest of these three inertia and quote the stability theorem for Chaplygin's sphere (Theorem 4.1) to show that these controlled equations are stable. If we are making λ_3 the smallest, that is $\lambda_3 < \lambda_2$, then choose ν so that

$$\nu < \frac{I_1 - \lambda_3}{J_1}.$$

If we are making λ_3 the largest, that is $\lambda_3 > \lambda_2$, then choose

$$\nu > \frac{I_1 - \lambda_3}{J_1}.$$

Either of these conditions on ν stabilizes the system.

To determine the feedback law, we solve for the acceleration $\dot{\Omega}_1$ in terms of Ω and Γ by computing components of the inverse of $\bar{I}(\Gamma)$. In the following formulae, let $\alpha = mr^2$. The control force u is given by

$$u = -\frac{\nu J_1}{\Delta}\{\left[\lambda_3(\lambda_2+\alpha) + \alpha\Gamma_1^2(\lambda_2+\alpha) + \alpha\Gamma_2^2(\lambda_2-\lambda_3)\right](\lambda_2-\lambda_3)\Omega_2\Omega_3 +$$
$$\alpha\Gamma_1\Gamma_2(\lambda_3+\alpha)(\lambda_3 - I_1 + J_1\nu)\Omega_3\Omega_1 + \alpha\Gamma_1\Gamma_3(\lambda_2+\alpha)(I_1 - J_1\nu - \lambda_2)\Omega_1\Omega_2\}, \quad (5.6)$$

where Δ is the determinant of $\bar{I}(\Gamma)$. Let $\bar{\Lambda}_1 = I_1 - J_1\nu + mr^2$, $\Lambda_2 = \lambda_2 + mr^2$, and $\Lambda_3 = \lambda_3 + mr^2$. Then

$$\Delta = \bar{\Lambda}_1\Lambda_2\Lambda_3 - \Lambda_2\Lambda_3\Gamma_1^2 - \Lambda_3\bar{\Lambda}_1\Gamma_2^2 - \bar{\Lambda}_1\Lambda_2\Gamma_3^2.$$

As noted before, if $\nu < \frac{I_1-\lambda_3}{J_1}$ then the rotation is stable. We note that when this control law is evaluated at $\Gamma = (0, 0, 1)^T$, it coincides with the control law obtained for the free rigid body with internal rotor in [BLMa].

5.2 *Translating Plane*

In this section, we derive a control law to stabilize Chaplygin's sphere on the translating plane. This is an idealized model where the plane moves only in response to control forces which translate it. We first derive the equations of motion of the controlled system. Although the controlled equations no longer conserve the energy which Chaplygin's sphere did. they do conserve the vertical component of the angular momentum. This property, in conjunction with an analysis of the linearized equations, allows one to construct stabilizing feedback laws that are functions of the velocity and orientation of the ball. Constructing stabilizing feedback laws based only on the orientation of the ball requires different techniques—this is the problem we consider in detail. A solution is developed based on the method of controlled Lagrangians discussed earlier.

To derive the controlled equations of motion for Chaplygin's sphere on the translating plane we start with Newton's principles. This provides a simple framework for incorporating the control force acting on the plane. Manipulation of the initial equations is guided by our knowledge of the equations of motion for Chaplygin's sphere that were obtained by Theorem 3.1.

We assume the same notation used to describe Chaplygin's sphere in previous chapters. Recall that the ball's configuration in space is determined by $(A, x) \in SE(3)$ which maps from a reference coordinate system into an inertial coordinate system. The angular velocity of the ball is denoted by Ω in the body and by ω in space. The vertical axis in the body is given by $\Gamma = A^{-1}e_3$. Let $a \in \mathbb{R}^3$ denote the location of the plane's center of mass in space. As the plane remains horizontal and does not rotate, this suffices to specify the configuration of the plane. We assume the vertical component of x is 0 and that the vertical component of a is $-r$, where r is the radius of the ball. Let m be the mass of the ball, and let n denote the mass of the plane. The horizontal control force which translates the plane is denoted by $f \in \mathbb{R}^3$, where it is understood that the vertical component of f is 0.

Proposition 5.1. *Assuming the above notation for describing Chaplygin's sphere on the translating plane with control force f, the following equation of motion holds:*

$$\dot{M} + \Omega \times M = -\frac{mr}{n}\Gamma \times A^{-1}f, \tag{5.7}$$

where M is given by $I\Omega + mr^2\Gamma \times (\Omega \times \Gamma)$.

Proof. As we have assumed the plane is unaffected by the reaction force of the ball, its equation of motion is simply

$$n\ddot{a} = f. \tag{5.8}$$

The constraint for this problem is that the ball rolls on this *moving* plane without slipping. This means the velocity of the contact point is equal to the velocity of the plane. One has

$$\dot{x} = \omega \times (re_3) + \dot{a}. \tag{5.9}$$

The outside forces on the ball are gravity (which we may ignore as the height of the center of mass for the ball is constant) and the reaction force from the plane. Let R denote the reaction force. Newton's principles, that the rate of change of linear momentum is equal to the sum of outside forces acting on the center of mass, and that the rate of change of angular momentum (about the center of mass) is equal to the sum of the torques about the center of mass, imply the following:

$$\frac{d}{dt}(m\dot{x}) = R \tag{5.10}$$

$$\frac{d}{dt}(AI\Omega) = -re_3 \times R. \tag{5.11}$$

By substituting equation (5.9) into equation (5.10), and the resulting formula for R into equation (5.11), one has an equation in terms of (Ω, Γ). Using the relation $\dot{\Gamma} = \Gamma \times \Omega$, one can obtain equation (5.7) stated in the proposition. □

Equation (5.7) may be rewritten in the following way

$$I\dot{\Omega} + mr^2\Gamma \times (\dot{\Omega} \times \Gamma) + \Omega \times I\Omega = -\frac{mr}{n}\Gamma \times A^{-1}f. \tag{5.12}$$

Typically a control law is a function of the orientation and velocity of the controlled system. For this problem, it is reasonable to measure the orientation and velocity of the ball from a spatial frame. For example, there may be a device fixed in the inertial frame which scans the movement of the ball. To simplify analysis, we choose f so that $A^{-1}f$ is a function of Ω and Γ. To represent the control force in the body frame, let

$$F(\Omega, \Gamma) = \frac{mr}{n}A^{-1}f. \tag{5.13}$$

One expects the controlled equations to break the conservation of angular momentum in space, however one finds that the vertical component of the momentum is still conserved. Given this symmetry in the controlled equations, one can find a stabilizing control law F depending on Ω and Γ by analyzing the linearized equations, as we now show.

Proposition 5.2. *The controlled equations for Chaplygin's sphere on the translating plane conserve* $\langle M, \Gamma\rangle$.

Proof. The controlled equations take the form:

$$\dot{M} + \Omega \times M = -\Gamma \times F$$
$$\dot{\Gamma} = \Gamma \times \Omega,$$

where F is some vector quantity. Calculating $\frac{d}{dt}\langle M, \Gamma\rangle$, one obtains:

$$\langle \dot{M}, \Gamma\rangle + \langle M, \dot{\Gamma}\rangle =$$

$$\langle -\Omega \times M - \Gamma \times F, \Gamma\rangle + \langle M, \Gamma \times \Omega\rangle =$$

$$\langle M, \Omega \times \Gamma\rangle + \langle M, -\Omega \times \Gamma\rangle = 0.$$

□

Proposition 5.3. *It is possible to find a control force $F(\Omega, \Gamma)$ which stabilizes the middle axis rotation for Chaplygin's sphere on the translating plane.*

Proof. First we linearize the controlled equations in the form

$$I(\Gamma)\dot{\Omega} = I\Omega \times \Omega - \Gamma \times F(\Omega, \Gamma)$$

$$\dot{\Gamma} = \Gamma \times \Omega,$$

at the equilibrium $(\Omega_0, \Gamma_0) = (ZE_3, E_3)$. As (ZE_3, E_3) is an equilibrium, $F(ZE_3, E_3) \times E_3 = 0$. Furthermore, since $F(\Omega, \Gamma) = mr/nA^{-1}f$ where f is horizontal,

$F(ZE_3, E_3) = 0$. For suitable choices of $\partial F^i/\partial\Omega^j$ and $\partial F^i/\partial\Gamma^a$ evaluated at the equilibrium, the linearized equations take the form

$$\frac{d}{dt}\begin{bmatrix} \Omega_1' \\ \Omega_2' \\ \Omega_3' \\ \Gamma_1' \\ \Gamma_2' \\ \Gamma_3' \end{bmatrix} = \begin{pmatrix} a & b & 0 & c & d & 0 \\ e & f & 0 & g & h & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & Z & 0 \\ 1 & 0 & 0 & -Z & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{bmatrix} \Omega_1' \\ \Omega_2' \\ \Omega_3' \\ \Gamma_1' \\ \Gamma_2' \\ \Gamma_3' \end{bmatrix} \tag{5.14}$$

where a, b, c, d, e, f, g and h are arbitrary constants we may choose, and $(\Omega', \Gamma') = (\Omega, \Gamma) - (ZE_3, E_3)$. Let B denote the matrix in the above equation. The constants $a, b, \ldots, h$ may be chosen so that four of the eigenvalues of B are negative, while the remaining two are zero. For instance, the characteristic polynomial of the diagonal matrix $\Lambda = diag[-4, -3, -2, -1, 0, 0]$ is

$$p(\lambda) = \lambda^6 + 10\lambda^5 + 35\lambda^4 + 50\lambda^3 + 24\lambda^2.$$

If we choose $e = f = g = 0$ and $h = 1$, then the characteristic polynomial for B is

$$r(\lambda) = \lambda^6 - a\lambda^5 + (Z^2 - d)\lambda^4 + (-b - Z - cZ - aZ^2)\lambda^3 + (c + aZ)\lambda^2.$$

It is then clear that a, b, c and d may be chosen so that $p(\lambda) = r(\lambda)$.

The above argument shows that we may choose $F(\Omega, \Gamma)$ so that the middle axis rotation becomes linearly stable. Since two of the eigenvalues of B are always 0, the argument does not imply stability. However, as $\langle\Gamma, \Gamma\rangle = 1$ and the controlled equations always conserves $\langle M, \Gamma\rangle$, we may establish stability as follows. A neighborhood of the equilibrium is foliated by four dimensional invariant submanifolds upon which $\langle M, \Gamma\rangle$ is constant. Now on the invariant manifold where $\langle M, \Gamma\rangle = I_3 Z$, the flow is stable (in fact, asymptotically stable) at the equilibrium. This follows from the invariance of the spectrum of the linearized matrix under a change of coordinates. If we can show that the flow is stable on nearby invariant submanifolds where $\langle M, \Gamma\rangle = I_3(Z \pm \epsilon)$ for $\epsilon > 0$ small, then we can conclude stability. Each of these nearby invariant submanifolds contains the equilibrium $(\Omega_0, \Gamma_0) = ((Z \pm \epsilon)E_3, E_3)$. Now $p(\lambda)$ was chosen so that its negative roots are stable under small perturbations. As the coefficients of $p(\lambda)$ depend continuously on Z, there will be four negative real eigenvalues in the spectrum of the linearized matrix at the nearby equilibria $((Z \pm \epsilon)E_3, E_3)$. Hence we can find $F(\Omega, \Gamma)$ such that middle axis rotation becomes stable. □

Clearly there are a number of choices for $F(\Omega, \Gamma)$ that will stabilize the system in this manner. Next we focus our attention on the problem of stabilizing the ball with a control force that is a function of Γ alone. Constructing such a control force requires a different argument than that used in Proposition 5.3. The trace of the matrix obtained by linearizing the controlled equations with a control force $F(\Gamma)$ is 0. Hence if there is an eigenvalue with negative real part, there is also one with positive real part, and the equilibrium would be unstable. To solve this problem, we draw upon the method of controlled Lagrangians to find a control force for which a Lyapunov function may be developed.

The reduced equations for the uncontrolled Chaplygin's sphere on the translating plane agree with the reduced equations for Chaplygin's sphere. Because of this, we will start with the Lagrangian and distribution for Chaplygin's sphere and then make modifications to the Lagrangian with the hope that the resulting reduced equations will match those of the controlled equations for Chaplygin's sphere on the translating plane. The modification we make to the Lagrangian is simply to add a potential energy term. We subtract a function $U(\Gamma)$ from the reduced Lagrangian $l : \mathfrak{se}(3) \to \mathbb{R}$ of Chaplygin's sphere to obtain the new controlled Lagrangian $\tilde{l} : \mathfrak{se}(3) \times \mathbb{R}^3 \to \mathbb{R}$ defined by

$$\tilde{l}(\Omega, Y, \Gamma) = \frac{1}{2}\{\langle I\Omega, \Omega\rangle + m\|Y\|^2\} - U(\Gamma).$$

By treating $\tilde{l}$ as a function on $TSE(3)$ (that is setting $\Gamma = g^{-1}a_0$) and taking the distribution for Chaplygin's sphere, one has an intermediate system to which the Intermediate Theorem (Theorem 3.3) may be applied. The resulting Euler-Poincaré equation and advection equations on $\mathfrak{so}(3) \times S^2$ are

$$\begin{aligned} \dot{M} &= M \times \Omega + \Gamma \times \frac{\partial U}{\partial \Gamma} \\ \dot{\Gamma} &= \Gamma \times \Omega, \end{aligned} \tag{5.15}$$

where $M = I\Omega + mr^2(\Gamma \times (\Omega \times \Gamma))$ as before. These equations match the controlled equations for Chaplygin's sphere on the translating plane with control force

$$F = -\frac{\partial U}{\partial \Gamma} + \nu\Gamma,$$

where ν is some function. For a given U, one chooses ν so that F comes from a horizontal force, that is $\langle AF, e_3\rangle = 0$. Moreover, with such a control force, the controlled equations will conserve the energy corresponding to $\tilde{l}$ and the distribution of Chaplygin's sphere, namely

$$\frac{1}{2}\langle M, \Omega\rangle + U(\Gamma).$$

By studying the linearized equations for $\tilde{l}$ at the equilibrium one can see that if $U(\Gamma)$ satisfies

$$\frac{\partial^2 U}{\partial\Gamma_1\partial\Gamma_1} = \frac{\partial^2 U}{\partial\Gamma_2\partial\Gamma_2} = -\kappa, \text{ and all other } \frac{\partial^2 U}{\partial\Gamma_i\partial\Gamma_j} = 0$$

at the equilibrium, and κ is sufficiently negative, then the equilibrium becomes linearly stable. The function

$$U(\Gamma) = -\kappa\frac{1}{2}\|E_3 \times \Gamma\|^2$$

satisfies these conditions. One then finds that with this $U(\Gamma)$ and the corresponding conserved energy, a Lyapunov function may be developed. The result of these observations is the following.

Theorem 5.1. *The middle axis rotation for Chaplygin's sphere on a translating plane is stabilized by applying the force*

$$f = \frac{\kappa n}{m r}\left\{AE_3 \times (e_3 \times AE_3) - \langle AE_3 \times (e_3 \times AE_3), e_3\rangle e_3\right\} \tag{5.16}$$

to the plane where κ is a constant such that κ is less than both $Z^2(I_3 - I_1)$ and $Z^2(I_3 - I_2)$.

Proof. We first note that $\langle f, e_3\rangle = 0$, hence f is a horizontal force. Substituting this f into the controlled equations (5.7) derived in Proposition 5.1 gives the following equations of motion:

$$\begin{aligned} \dot{M} &= M \times \Omega + \kappa\langle E_3, \Gamma\rangle\Gamma \times E_3 \\ \dot{\Gamma} &= \Gamma \times \Omega. \end{aligned} \tag{5.17}$$

Now consider the nonholonomic system with Lagrangian

$$\tilde{l} = \frac{1}{2}\left\{\langle \Omega, I\Omega\rangle + m\|\dot{x}\|^2\right\} + \kappa\frac{1}{2}\|E_3 \times \Gamma\|^2$$

and distribution defined as with Chaplygin's sphere, namely $\dot{x} = \omega \times re_3$. The equations of motion for this system, derived using Theorem 3.3, agree with the controlled equations (5.17) above. Hence these controlled equations conserve the following function:

$$C_3 = \langle M, \Omega\rangle - \kappa\langle \Gamma, E_3 \times (\Gamma \times E_3)\rangle,$$

as C_3 is twice the energy of the system. In addition, the functions $C_1 = \langle \Gamma, \Gamma\rangle$ and $C_2 = \langle M, \Gamma\rangle$ are conserved (the latter following from Proposition 5.2). Next we construct a Lyapunov function out of these three conserved quantities, in much the same fashion as the Lyapunov function for Theorem 4.1 was constructed.

Let $\tilde{H} : \mathbb{R}^3 \to \mathbb{R}$ be an undetermined function, and define $H : \mathbb{R}^6 \to \mathbb{R}$ by

$$H(\Omega, \Gamma) = \tilde{H}(C_1(\Omega, \Gamma), C_2(\Omega, \Gamma), C_3(\Omega, \Gamma)).$$

Clearly H is conserved along the flow. It remains to show that $\tilde{H}$ may be chosen so that H has a strict local minimum at the equilibrium (ZE_3, E_3). We do so by showing that $\nabla H = 0$, and $\delta^2 H$ is positive definite at the equilibrium. Let

$$C_0 = (C_1(ZE_3, E_3), C_2(ZE_3, E_3), C_3(ZE_3, E_3)),$$

$$h_i = \frac{\partial \tilde{H}}{\partial C_i} \text{ evaluated at } C_0,$$

and

$$h_{ij} = \frac{\partial^2 \tilde{H}}{\partial C_i \partial C_j} \text{ evaluated at } C_0.$$

The condition $\nabla H = 0$ translates into

$$h_2 I_3 + 2h_3 I_3 Z = 0, \quad 2h_1 + h_2 I_3 Z = 0.$$

To satisfy this, take:

$$h_3 = I_3, \quad h_2 = -2I_3 Z, \text{ and } h_1 = (I_3 Z)^2.$$

Let $\Omega' = \Omega - \Omega_0$ and $\Gamma' = \Gamma - \Gamma_0$. We compute the second variation. $\langle [\Omega', \Gamma'], \delta^2 H\, [\Omega', \Gamma']^T \rangle$. This breaks up into the sum $Q_1 + Q_2 + Q_3$, where each Q_i is a quadratic form in the pair (Γ'_i, Ω'_i). Using the constants $A_i = I_i + mr^2$ to simplify the expressions, we obtain:

$$Q_1 = 2\, I_3\, (Z^2\, A_3 - \kappa)\; {\Gamma'_1}^2 - 4\, A_1\, I_3\, Z\, \Gamma'_1\, \Omega'_1 + 2\, A_1\, I_3\, {\Omega'_1}^2,$$

$$Q_2 = 2\, I_3\, (Z^2\, A_3 - \kappa)\; {\Gamma'_2}^2 - 4\, A_2\, I_3\, Z\, \Gamma'_2\, \Omega'_2 + 2\, A_2\, I_3\, {\Omega'_2}^2, \text{ and}$$

$$\begin{aligned} Q_3 = {} & (4\, h_{11} + I_3\, Z\; (4\, h_{21} + (2 + h_{22})\; I_3\, Z))\; {\Gamma'_3}^2 + \\ & 2\, I_3\; (2\, h_{21} + Z\; (4\, h_{31} + I_3\; (-2 + h_{22} + 2\, h_{32}\, Z)))\; \Gamma'_3\, \Omega'_3 + \\ & {I_3}^2\; (2 + h_{22} + 4\, h_{32}\, Z + 4\, h_{33}\, Z^2)\; {\Omega'_3}^2. \end{aligned}$$

The last quadratic form in (Γ'_3, Ω'_3) can be made positive definite by choosing the h_{ij} appropriately. The two terms in (Γ'_1, Ω'_1) and (Γ'_2, Ω'_2) are of the form $aX^2 - bXY + cY^2$. Such a form is positive definite if $4ac - b^2 > 0$. Taking

$$\kappa < Z^2(I_3 - I_1) \text{ and } \kappa < Z^2(I_3 - I_2)$$

ensures that this is so. Hence, we have found a closed loop feedback law that is a function of Γ and stabilizes the middle axis rotation for Chaplygin's sphere on the translating plane. □

5.3 Sliding Plane

When the mass of the plane is comparable to the mass of the ball, the dynamics of the plane are affected by the reaction force of the ball, which will generally torque as well as translate the plane. To study such dynamics, we start with an idealized model where we assume the plane slides on a frictionless surface, moving by horizontal translations and by rotations about the vertical axis. As before, the ball is assumed to roll without slipping on the plane. We call the system Chaplygin's sphere on the sliding plane.

We will start this section by first deriving the equations of motion for Chaplygin's sphere on the sliding plane. Although the system does not have the invariance property required for applying the reduction theorems of Chapter 3, the method used to prove these theorems may be adapted to derive the equations of motion for this system. The resulting equations, although more complicated, are very similar in structure to those derived for Chaplygin's sphere on the translating plane. After some study of the equilibria in these equations, where we identify stationary rotations of the ball, we consider stabilizing the rotation of Chaplygin's sphere about its middle axis by applying forces that translate the plane upon which the ball rolls. The controlled equations are derived by starting with Newton's principles and using the form of the uncontrolled equations to guide the calculation. The question we investigate is whether or not the solutions to the translating plane problem can be generalized to this setting. Although many of the conserved quantities that exist for the system on the translating plane have analogs on the sliding plane, we find that this does not provide enough symmetry to establish stability, at least by employing the methods of the previous chapter. That is studying the linearized equations does not lead to a stability law $F(\Omega, \Gamma)$ nor can we obtain a stabilizing control law $F(\Gamma)$ by adding a potential energy term to the Lagrangian of the uncontrolled system and matching the resulting Euler-Poincaré equations to the equations of motion for the controlled system.

The configuration space for Chaplygin's sphere on the sliding plane is $SO(3) \times \mathbb{R}^2 \times S^1 \times \mathbb{R}^2$. Let a point of this space be denoted by (A, x, θ, a), where $A \in SO(3)$ and $x \in \mathbb{R}^2$ determine the configuration of the ball by mapping from a reference coordinate system for the ball to an inertial coordinate system for space in the usual way. The configuration of the plane is determined by $\theta \in S^1$, and $a \in \mathbb{R}^2$ where a gives the horizontal coordinates of the center of mass of the plane in the inertial frame, and θ gives the orientation of the plane in this frame. As with the previous systems, it is useful to model the motion in $SE(3)$. To this end, we treat x and a as vectors in $\mathbb{R}^3$ whose vertical component is zero. The horizontal plane upon which the ball rolls is taken to be at $x_3 = -r$ (where r is the radius of the ball). The center of mass of the ball always resides in the plane $x_3 = 0$. A tangent vector at a point in the configuration space will be denoted by $(\dot{A}, \dot{x}, \dot{\theta}, \dot{a})$. As usual Ω and ω in $\mathbb{R}^3$ are defined by $\widehat{\Omega} = A^{-1}\dot{A}$ and $\widehat{\omega} = \dot{A}A^{-1}$. Some other constants in the problem are:

I_B - the inertia matrix for the ball	m - the total mass of the ball
n - the total mass of the plane	k - the inertia of the plane.

The constraints are derived from the assumption that the ball rolls without slipping on the moving plane. Hence the velocity of the point on the ball that is in contact with the plane is equal to the velocity of the point of the plane that is in contact with the ball. As usual the contact point on the ball has velocity $\dot{x} - \omega \times re_3$. The velocity of the contact point on the plane is $\dot{a} + \dot{\theta} e_3 \times (x - a)$, which is the sum of the velocity of the center of mass of the plane and the velocity due to rotation of the contact point about the center of mass. Setting the velocities of these two contact points equal to each other gives the constraint equation

$$\dot{x} - \dot{a} = \omega \times re_3 + \dot{\theta} e_3 \times (x - a).$$

Observe that the system is closed: that is there are no outside forces that affects the dynamics of the ball and plane. Hence the center of mass moves uniformly and one has

$$m\dot{x} + n\dot{a} = 0. \tag{5.18}$$

This allows us to reduce the configuration space to $Q = SO(3) \times S^1 \times \mathbb{R}^2$. Let $y = x - a$. Then (5.18) implies that the Lagrangian is

$$L = \frac{1}{2}\left\{\langle \Omega, I_B \Omega\rangle + k\dot{\theta}^2 + \bar{m}\|\dot{y}\|^2\right\}, \text{ where } \bar{m} = \frac{nm}{n+m},$$

and that the constraint equation may be written as

$$\dot{y} = \omega \times re_3 + \dot{\theta} e_3 \times y. \tag{5.19}$$

The symmetry in Chaplygin's sphere on the sliding plane is insufficient for applying any of the nonholonomic reduction theorems of Chapter 3; this is evident from the dependence of the constraints on y. However the techniques used in deriving the Euler-Poincaré equations in Chapter 3 may be generalized to obtain equations for Chaplygin's sphere on the sliding plane that are easily comparable to those for Chaplygin's sphere, and have much of the same structure.

We will derive equations of motion for a certain class of systems defined on the semidirect product $S = G \circledS V$ where $G = SO(3) \times S^1$, $V = \mathbb{R}^3$ and the action of G on V consists of the action of the $SO(3)$ factor on $\mathbb{R}^3$. Let (g, y) denote a point of S, where g has components (A, θ), with $A \in SO(3)$ and $\theta \in S^1$. Let $(\dot{g}, \dot{y})$ denote a tangent vector in $T_{(g,y)}S$, where $\dot{g}$ has components $(\dot{A}, \dot{\theta})$. The distribution for the system will be determined by the constraint equation (5.19) above. Note that the constraint equation implies that $0 = \langle \dot{y}, e_3 \rangle$ so that the configuration space for the problem is $G \times \mathbb{R}^2$. However as with the previous systems, it will be more convenient to work with the semidirect product S.

We will write the equations of motion on $\mathfrak{g} \times S^2 \times \mathbb{R}^3$ after first writing the Lagrangian and distribution on $\mathfrak{s} \times S^2 \times \mathbb{R}^3$. Here $\mathfrak{g}$ is the Lie algebra of $G = SO(3) \times S^1$ and $\mathfrak{s}$ is the Lie algebra of $S = G \circledS \mathbb{R}^3$. The variables we use to parameterize $\mathfrak{s} \times S^2 \times \mathbb{R}^3$ are (ξ, Y, Γ, γ), where $(\xi, Y) \in \mathfrak{s}$ with $\xi \in \mathfrak{g}$ and $Y \in \mathbb{R}^3$, $\Gamma \in S^2$ and $\gamma \in \mathbb{R}^3$. The components of ξ are (Ω, J) with $J \in \mathbb{R}$ and $\Omega \in \mathbb{R}^3$ (so that $\widehat{\Omega} \in \mathfrak{so}(3)$). In terms of a tangent vector in $T_{(g,y)}S$, these variables are defined in the following way:

$$\widehat{\Omega} = A^{-1}\dot{A}, \quad J = \dot{\theta}, \quad Y = A^{-1}\dot{y}, \quad \Gamma = A^{-1}e_3, \quad \gamma = A^{-1}y. \tag{5.20}$$

It is assumed that the Lagrangian for the system, $L : TS \to \mathbb{R}$, may be written as a function on $\mathfrak{s} \times S^2 \times \mathbb{R}^3$ by the change of variables above. Let $l : \mathfrak{s} \times S^2 \times \mathbb{R}^3 \to \mathbb{R}$ denote this function. This effectively assumes an $S^1 \times SO(2)$ invariance in L, with the S^1 factor coming from θ being a cyclic variable, and the $SO(2)$ factor corresponding to the isotropy subgroup of e_3 in $SO(3)$.

The constraint equation can be written on $\mathfrak{s} \times S^2 \times \mathbb{R}^3$ by applying A^{-1} to both sides of equation (5.19). One obtains

$$Y = r\Omega \times \Gamma + J\Gamma \times \gamma. \tag{5.21}$$

For convenience, define $B(\Gamma, \gamma) : \mathfrak{g} \to \mathbb{R}^3$ by

$$B(\Gamma, \gamma)\xi = r\Omega \times \Gamma + J\Gamma \times \gamma,$$

where $\xi = (\Omega, J)$, so that the constraints may be written as $Y = B(\Gamma, \gamma)\xi$. The dual map $B^* : \mathbb{R}^{3^*} \to \mathfrak{g}^*$ takes the form:

$$B^*F = (r\Gamma \times F, -\langle \gamma \times \Gamma, F \rangle). \tag{5.22}$$

The constrained Lagrangian $l_c : \mathfrak{g} \times S^2 \times \mathbb{R}^3$ is defined as

$$l_c(\xi, \Gamma, \gamma) = l(\xi, B(\Gamma, \gamma)\xi, \Gamma, \gamma).$$

We will make use of the following notation when writing the equations of motion. For $\xi, \eta \in \mathfrak{g}$, where ξ has components (Ω, J) and η has components (Σ, K), the commutator on $\mathfrak{g}$ is

$$[\xi, \eta] = (\Omega \times \Sigma, 0).$$

The action of $\mathfrak{g}$ on V is $\xi v = \Omega \times v$, where $v \in \mathbb{R}^3$. (This action refers to the map $\mathfrak{g} \longrightarrow \mathfrak{gl}(\mathbb{R}^3)$ induced by the left representation of G on $\mathbb{R}^3$.) For $v \in \mathbb{R}^3$, this determines the map

$$\rho_v : \mathfrak{g} \longrightarrow V \text{ by } \rho_v(\xi) = \Omega \times v$$

and its dual map

$$\rho_v^* : V^* \longrightarrow \mathfrak{g}^* \text{ by } \rho_v^*(w) = (v \times w, 0).$$

Theorem 5.2. *For the system $(L, \mathcal{D})$ on TS described above, the following equations of motion hold on $\mathfrak{g} \times S^2 \times \mathbb{R}^3$ in terms of the variables $(\Omega, J, \Gamma, \gamma)$.*

$$\frac{d}{dt}\frac{\partial l_c}{\partial \Omega} + \Omega \times \frac{\partial l_c}{\partial \Omega} = -\Gamma \times \frac{\partial l}{\partial \Gamma} - \gamma \times \frac{\partial l}{\partial \gamma} + \Gamma \times \frac{\partial l}{\partial \gamma} + \frac{\partial l}{\partial Y} \times Y \tag{5.23}$$

$$\frac{d}{dt}\frac{\partial l_c}{\partial J} = \langle \frac{\partial l}{\partial \gamma}, \Gamma \times \gamma \rangle + \langle \frac{\partial l}{\partial Y}, \Gamma \times Y \rangle \tag{5.24}$$

$$\dot{\Gamma} = \Gamma \times \Omega$$

$$\dot{\gamma} = \gamma \times \Omega + B(\Gamma, \gamma)\xi,$$

where $Y = B(\Gamma, \gamma)\xi$.

Proof. The Lagrange d'Alembert principle determines a curve $(g(t), y(t)) \in S$, where $g(t) \in G$ has components $(A(t), \theta(t))$ and $y(t) \in \mathbb{R}^3$, by the following conditions. The curve $(g(t), y(t))$ satisfies the constraints and one has

$$\delta \int_a^b l(g(t), y(t)) = 0, \tag{5.25}$$

where $\delta g(t) = (\delta A(t), \delta\theta(t))$ is an independent variation along the curve vanishing at the endpoints, and

$$\delta y = \delta A\, A^{-1} r e_3 + \delta\theta e_3 \times y. \tag{5.26}$$

By hypothesis, L may be written as a function on $\mathfrak{s} \times S^2 \times \mathbb{R}^3$ denoted by l. Hence equation (5.25) is equivalent to

$$\delta \int_a^b l(\xi(t), Y(t), \Gamma(t), \gamma(t)) = 0, \tag{5.27}$$

where the components of the curve $(\xi(t), Y(t), \Gamma(t), \gamma(t))$ on $\mathfrak{s} \times S^2 \times \mathbb{R}^3$ are determined from $(g(t), y(t))$ as in equation (5.20). Assuming that $(g(t), y(t))$ satisfies the constraints is equivalent to requiring that $Y = B(\Gamma, \gamma)\xi$ as in equation (5.21). We compute variations of the variables (ξ, Y, Γ, γ) in terms of η: an independent variation on $\mathfrak{g}$ (vanishing at the endpoints) defined by $\eta = g^{-1}\delta g$. We denote the components of η by $(\Sigma, \delta\theta)$, where $\Sigma = A^{-1}\delta A$. The computed variations are

$$\begin{aligned}
\delta\xi &= \dot{\eta} + ad_{\xi}\eta \\
\delta Y &= \frac{d}{dt}(B\eta) + \xi(B\eta) - \eta Y \\
\delta\Gamma &= -\eta\Gamma \\
\delta\gamma &= -\eta\gamma + B(\Gamma, \gamma)\eta.
\end{aligned}$$

The $\delta\xi$ and $\delta\Gamma$ variations are computed just as in Theorem 3.3. The components of $\delta\xi$ are $(\dot{\Sigma} + \Omega \times \Sigma, \dot{\delta\theta})$. The δY variation follows from first computing $\delta Y = \delta(A^{-1}\dot{y})$ and then substituting in the δy variation $\delta A\, A^{-1} r e_3 + \delta\theta e_3 \times y$. Likewise the $\delta\gamma$ variation follows from computing $\delta\gamma = \delta(A^{-1}y)$ and then substituting in the δy variation.

Varying the integrand in equation (5.27), substituting in the above variations and using the relation

$$\frac{\partial l_c}{\partial \xi} = \frac{\partial l}{\partial \xi} + B^* \frac{\partial l}{\partial Y}$$

yield the following:

$$\int \{\langle -\frac{d}{dt}\frac{\partial l_c}{\partial \xi} + ad_{\xi}^* \frac{\partial l_c}{\partial \xi} - \rho_{\Gamma}^* \frac{\partial l}{\partial \Gamma} - \rho_{\gamma}^* \frac{\partial l}{\partial \gamma} + B^* \frac{\partial l}{\partial \gamma}, \eta\rangle\} dt + \tag{5.28a}$$

$$\int \{\langle -B^* \frac{\partial l}{\partial Y}, \dot{\eta} + ad_{\xi}\eta\rangle + \langle \frac{\partial l}{\partial Y}, \dot{B}\eta + \xi B\eta - \eta Y\rangle\} \, dt. \tag{5.28b}$$

By integrating by parts and simplifying, The second line, equation (5.28b), can be written as

$$\int \{\langle \frac{\partial l}{\partial Y}, \Sigma \times r\dot{\Gamma} + \delta\theta \frac{d}{dt}(\Gamma \times \gamma) - (\Omega \times \Sigma) \times r\Gamma + \Omega \times (\Sigma \times r\Gamma) + \delta\theta\Omega \times (\Gamma \times \gamma) - \Sigma \times Y\rangle\} dt.$$

After collecting terms involving $\delta\theta$ and noting the two Jacobi identities, equation (5.28b) may be written as

$$\int \{\delta\theta\langle \frac{\partial l}{\partial Y}, \Gamma \times Y\rangle + \langle \frac{\partial l}{\partial Y}, Y \times \Sigma\rangle\} dt.$$

Combining this with equation (5.28a), we set the varied integrand equal to zero as in equation (5.27) to obtain

$$0 = \int_a^b \langle -\frac{d}{dt}\frac{\partial l_c}{\partial \xi} + ad^*_\xi \frac{\partial l_c}{\partial \xi} - \rho^*_\Gamma \frac{\partial l}{\partial \Gamma} - \rho^*_\gamma \frac{\partial l}{\partial \gamma} + B^* \frac{\partial l}{\partial \gamma}, \eta\rangle + \delta\theta\langle \frac{\partial l}{\partial Y}, \Gamma \times Y\rangle + \langle \frac{\partial l}{\partial Y}, Y \times \Sigma\rangle dt.$$

The independence of η implies that the first two equations in Theorem 5.2 hold. The equations $\dot{\Gamma} = \Gamma \times \Omega$ and $\dot{\gamma} = \gamma \times \Omega + B(\Gamma, \gamma)\xi$ follow from differentiating $\Gamma = A^{-1}e_3$ and $\gamma = A^{-1}y$. □

We shall apply Theorem 5.2 to obtain equations of motion for Chaplygin's sphere on the sliding plane. The Lagrangian is written in terms of the variables (ξ, Y, Γ, γ) as

$$l = \frac{1}{2}\left\{\langle \Omega, I_B\Omega\rangle + kJ^2 + \bar{m}\|Y\|^2\right\}, \tag{5.29}$$

and the distribution is determined by the equation

$$Y = r\Omega \times \Gamma + J\Gamma \times \gamma. \tag{5.30}$$

As $\partial l/\partial\Gamma = 0$, $\partial l/\partial\gamma = 0$ and $\partial l/\partial Y = Y$, the terms in the right hand side of equations (5.23) and (5.24) drop out. Let $M = \partial l_c/\partial\xi$. Compactly, the equations are written as

$$\begin{aligned} \dot{M} + [\xi, M] &= 0 \\ \dot{\gamma} &= \gamma \times \Omega + B(\Gamma, \gamma)\xi \\ \dot{\Gamma} &= \Gamma \times \Omega. \end{aligned} \tag{5.31}$$

These equations conserve the four dimensional momentum vector $Ad^*_{g^{-1}}M$. This can be checked by directly differentiating and using the equation $\dot{M} + [\xi, M]$ (compare with Theorem 3.2). The conservation of $Ad^*_{g^{-1}}M$ implies that both $AM_\Omega \in \mathbb{R}^3$ and M_J are conserved. Consequently one has several conserved quantities that we collect in the following proposition.

Proposition 5.4. *Equations (5.31), as a system on $\mathbb{R}^3 \times \mathbb{R} \times \mathbb{R}^3 \times \mathbb{R}^3$ in the variables $(\Omega, J, \Gamma, \gamma)$, have the following six constants of motion*

$$\langle M, \xi\rangle \quad M_J \quad \langle M_\Omega, \Gamma\rangle \quad \langle M_\Omega, M_\Omega\rangle \quad \langle \Gamma, \Gamma\rangle = 1 \quad \langle \gamma, \Gamma\rangle = 0.$$

Proof. The first is twice the energy for the system, given by l_c. That $\frac{d}{dt}M_J = 0$ was noted above. As AM_Ω and e_3 are constant vectors in the inertial frame, one has that $\langle M_\Omega, M_\Omega\rangle$ and $\langle M_\Omega, \Gamma\rangle$ are constant. That $\langle \Gamma, \Gamma\rangle = 1$ and $\langle \gamma, \Gamma\rangle = 0$ follows from Γ being in S^2 and y being a horizontal vector so that $\langle y, e_3\rangle = 0$. □

We shall write the equations in terms of ξ as opposed to M in order to linearize them. M may be written as the product of a (Γ, γ) dependent inertia matrix acting on ξ. Let $P_\Gamma : \mathbb{R}^3 \to \mathbb{R}^3$ denote orthogonal projection onto Γ, that is $P_\Gamma v = \Gamma \times (v \times \Gamma)$. For $v \in \mathbb{R}^3$, v^T represents v written as a row vector. By computing $\partial l_c/\partial \xi$, and noting that $\Gamma \times (\gamma \times \Gamma) = \gamma$ (as $\langle \gamma, \Gamma\rangle = 0$), one finds that

$$M = \frac{\partial l_c}{\partial \xi} = \begin{pmatrix} I_B + \bar{m}r^2 P_\Gamma & -\bar{m}r\gamma \\ \bar{m}r\gamma^T & k + \bar{m}\|\gamma\|^2 \end{pmatrix} \begin{bmatrix} \Omega \\ J \end{bmatrix}.$$

We denote this inertia operator by $\bar{I}$ and write $M = \bar{I}\xi$.

Differentiating the above expression for M in the equations (5.31) leads to the following form of the equations of motion:

$$\begin{pmatrix} I_B + \bar{m}r^2 P_\Gamma & -\bar{m}r\gamma \\ \bar{m}r\gamma^T & k + \bar{m}\|\gamma\|^2 \end{pmatrix} \begin{bmatrix} \dot{\Omega} \\ \dot{J} \end{bmatrix} = \begin{bmatrix} I_B\Omega \times \Omega + \bar{m}r^2 J\Omega \times \Gamma + \bar{m}rJ^2\Gamma \times \gamma \\ -\bar{m}rJ\langle \Omega \times \Gamma, \gamma\rangle \end{bmatrix} \tag{5.32}$$

$$\dot{\Gamma} = \Gamma \times \Omega$$

$$\dot{\gamma} = \gamma \times \Omega + r\Omega \times \Gamma + J\Gamma \times \gamma. \tag{5.33}$$

We will use these equations to identify the steady vertical rotations of the ball on the plane and determine, in one case, if the rotations are linearly stable. Steady vertical rotations are solutions where $\widehat{\omega} = \dot{A}A^{-1}$ is a multiple of $\widehat{e_3}$ and $y(t) = y_0$ is constant. In terms of the variables $(\Omega, J, \Gamma, \gamma)$, this implies that $\Omega(t)$ and $\Gamma(t)$ are constant. However unless the ball rotates about $y_0 = 0$, $\gamma(t)$ is not constant, but rather $\gamma(t) = A(0)^{-1}exp(-tZe_3)y_0$, where Z is some constant. Hence we look for solutions where Γ and Ω are constant, but not necessarily γ or J, although we shall find J must be constant.

Proposition 5.5. *Let E_3 denote the axis of rotation in the ball's coordinate system, so that $A(t)E_3 = e_3$. The steady rotations of Chaplygin's sphere on a sliding plane consist of the following:*

1. *The ball rotates about a vertical principal axis at the center of mass for the plane while the plane also rotates about its center of mass. That is E_3 is a principal axis for the ball, $\Gamma = E_3, \Omega = ZE_3, \gamma = 0$, and $J = P$, where Z and P are constants.*

2. *The ball rotates about a vertical principal axis at an arbitrary point on the plane and the plane does not move. That is E_3 is a principal axis for the ball, $\Gamma = E_3, \Omega = ZE_3, \gamma = \exp(-Z\widehat{E_3}t)A(0)y_0$, and $J = 0$, where Z is some constant.*

Proof. As we have assumed E_3 is the axis of rotation in the ball, $\Omega = ZE_3$ for some constant Z. The Γ equation, $\dot{\Gamma} = \Gamma \times \Omega$, implies that $\Gamma = E_3$. The J equation,

$$\langle \bar{m}r\gamma, \dot{\Omega}\rangle + (k + \bar{m}\|\gamma\|^2)\dot{J} = -\bar{m}rJ\langle \Omega \times \Gamma, \gamma\rangle,$$

then shows that $(k + \bar{m}\|\gamma\|^2)\dot{J} = 0$. As $k + \bar{m}\|\gamma\|^2 > 0$, we must have that $\dot{J} = 0$. Hence we have that J is some constant, call it P. The γ equation then implies that

$$\dot{\gamma} = (P - Z)E_3 \times \gamma,$$

which has the solution

$$\gamma(t) = exp((P - Z)t\widehat{E_3})A(0)^{-1}y_0.$$

Now we substitute this into the Ω equation (5.32) and get

$$0 = (Z^2 I_B E_3 - \bar{m}rP^2 exp((P - Z)t\widehat{E_3})A(0)^{-1}y_0) \times E_3. \tag{5.34}$$

If $y_0 = 0$. then we get the first case: P can be anything, and E_3 must be an eigenvector for I_B. hence a principle axis.
If $y_0 \neq 0$. and $P = 0$ then we get the second case: E_3 must be an eigenvector of I_B.
If $y_0 \neq 0$ and $P \neq 0$. then we have an expression that depends on time. a contradiction, unless $P = Z$. In this case. equation (5.34) is equivalent to

$$I_B E_3 = \bar{m} r A(0)^{-1} y_0 \times E_3.$$

Taking the dot product of both sides with E_3 yields $\langle I_B E_3, E_3 \rangle = 0$. which is a contradiction as we assume none of the three principal moments of inertia are zero. □

Next we continue to investigate the stabilization of the rotation of Chaplygin's sphere about its middle axis. From Proposition 5.5 we see that this motion manifest itself in several ways on the sliding plane. The first set of steady rotations are periodic solutions as opposed to equilibria. We confine our attention to the second set of steady rotations so that we may employ the techniques developed previously. In particular, we consider an equilibrium where the ball rotates about the center of mass of the plane and the plane does not rotate. Stabilizing such equilibria corresponds most closely to the problem we studied on the translating plane. The following proposition shows that the conditions for these equilibria to be linearly stable are the same as those for the equilibria of Chaplygin's sphere.

Proposition 5.6. *The steady rotations of Chaplygin's sphere on the sliding plane which consist of rotations about a vertical principal axis at the center of mass for the plane, while the plane does not rotate, are unstable if the rotation is about the middle axis, and linearly stable if the rotation is about the long or short axis.*

Proof. If one linearizes the equations around an equilibrium solution

$$\Gamma = E_3, \quad \Omega = Z E_3, \quad J = P, \quad \gamma = 0,$$

and computes the characteristic polynomial of the resulting 10×10 matrix, one finds it takes the form

$$p(\lambda) = \frac{\lambda^4}{A_1 A_2} \left(A_1 A_2 \lambda^6 + \beta_4 \lambda^4 + \beta_2 \lambda^2 + \beta_0 \right),$$

where $A_1 = I_1 + \bar{m}r^2$ and $A_2 = I_2 + \bar{m}r^2$. The coefficients β_0, β_2 and β_4 are lengthy expressions in terms of the constants of the problem. In the case when $P = 0$, the characteristic polynomial takes the form:

$$p(\lambda) = \frac{\lambda^4}{A_1 A_2}\left(A_1A_2\lambda^6 + (a_1a_2 + 2A_1A_2)Z^2\lambda^4 + (2a_1a_2 + A_1A_2)Z^4\lambda^2 + a_1a_2Z^6\right),$$

where $a_1 = I_3 - I_1$ and $a_2 = I_2 - I_3$. Let $a = a_1a_2$, $A = A_1A_2$, $\lambda = Z\mu$, and $\mu^2 = x$. Then the equilibrium is linearly stable if the roots of

$$q(x) = Ax^3 + (a + 2A)x^2 + (2a + A)x + a$$

have negative real part. By observing that $q(x)$ has the root $x = -1$, one obtains the factorization

$$q(x) = A(x+1)^2(x + \frac{a}{A}).$$

Hence $q(x)$ has the roots $-1, -a/A$. In the case when I_3 is the middle axis, the product a_1a_2 is negative, so that $q(x)$ has a positive root, implying that the middle axis rotation is unstable. However when $a > 0$, the equilibrium is linearly stable. Hence the short and long axes are linearly stable rotations with eigenvalues:

$$\lambda = \pm iZ, \quad \pm iZ\frac{a}{A}.$$

□

It would be interesting to develop a stability theorem for these equilibria as was done with Chaplygin's sphere. Using the six constants of motion, one may attempt to construct a Lyapunov function for an equilibrium, as was done in Theorem 4.1. However, the same argument will not work in the case when the plane does not rotate. When $J = 0$ expanding the constants of motion to second order will give no control over the growth of γ. When $J \neq 0$, γ will be amongst the second order expansion of the constants of motion, and the approach used in Theorem 4.1 may be fruitful.

We continue studying stabilization on the sliding plane by deriving the equations of motion for the controlled system. Guided by the form of the uncontrolled equations obtained from Theorem 5.2, we apply Newton's principles to obtain the following.

Proposition 5.7. *The equations of motion for Chaplygin's sphere on the sliding plane under the influence of a horizontal force f which translates, but does not torque the plane, take the following form:*

$$
\begin{aligned}
\dot{M} + [\xi, M] &= -B^{*}F \qquad (5.35)\\
\dot{\Gamma} &= \Gamma \times \Omega\\
\dot{\gamma} &= \gamma \times \Omega + r\Omega \times \gamma + J\Gamma \times \gamma,
\end{aligned}
$$

where $M = I(\Gamma, \gamma)\xi$ as in equation (4.3), $F = \frac{m}{n+m}A^{-1}f$ and B^{*} is defined as in equation (5.22).

Proof. We apply Newton's principles to obtain equations of motion for the plane and the ball. Here R is the reaction force of the plane on the ball and f is the outside control force on the plane. The equations for the plane are

$$\frac{d}{dt}\{n\dot{a}\} = -R + f \qquad (5.36)$$

$$\frac{d}{dt}\{k\dot{\theta}\}e_3 = (x - a) \times -R. \qquad (5.37)$$

The equations for the ball are

$$\dot{x} - \dot{a} = \omega \times re_3 + \dot{\theta}e_3 \times (x - a) \qquad (5.38)$$

$$\frac{d}{dt}\{m\dot{x}\} = R \qquad (5.39)$$

$$\frac{d}{dt}\{AI\Omega\} = -re_3 \times R. \qquad (5.40)$$

The equations may be written in terms of $y = x - a$. Take n times equation (5.39) minus m times equation (5.36) to get

$$nm\frac{d}{dt}\{\dot{x} - \dot{a}\} = (n + m)R - mf. \qquad (5.41)$$

Substituting equation (5.38) into (5.41) leads to the following expression for R

$$R = \frac{m}{n+m}f + \frac{d}{dt}\left\{\frac{nmr}{n+m}\omega \times e_3 + \frac{nm}{n+m}\dot{\theta}e_3 \times y\right\}.$$

Substituting this into equations (5.37) and (5.40) leads to the following equations

$$\frac{d}{dt}\left\{AI\Omega + \frac{nmr^2}{n+m}e + 3 \times (\omega \times e_3) - \frac{nmr}{n+m}\dot{\theta}y\right\} = \frac{-rm}{n+m}e_3 \times f$$

$$\frac{d}{dt}\left\{k\dot{\theta} - \frac{nmr}{n+m}\langle \omega, y\rangle + \frac{nm}{n+m}\dot{\theta}\|y\|^2\right\} = \frac{m}{n+m}\langle y, e_3 \times f\rangle$$

$$\dot{y} = \omega \times re_3 + \dot{\theta}e_3 \times y.$$

Writing these in terms of the variables (ξ, Γ, γ) gives the equations stated in the proposition.

□

Recall that for Chaplygin's sphere on the translating plane, it was shown by analyzing the eigenvalues of the linearization that the system could be stabilized by a suitable choice of control force $F(\Omega, \Gamma)$. For a control force $F(\Gamma)$ depending only on the orientation of the ball, a stabilizing control force was found by generalizing the method of controlled Lagrangians. The methods used to obtain these results do not generalize to Chaplygin's sphere on the sliding plane. This is what we shall show in the remainder of this section. One result that does generalize is the conservation of the vertical component of momentum for the controlled system.

Proposition 5.8. *The controlled equations for Chaplygin's sphere on the sliding plane conserve* $\langle M_\Omega, \Gamma\rangle$.

Proof. This follows from computing $\frac{d}{dt}\langle M_\Omega, \Gamma\rangle$, using $\dot{\Gamma} = \Gamma \times \Omega$ and the controlled equation (5.35):

$$\dot{M}_\Omega + \Omega \times M_\Omega = -r\Gamma \times F.$$

□

It seems reasonable to base the control force on measurements of the position and velocity of the ball from a device attached to the plane, or an external device residing in the inertial frame. Taking measurements relative to the platform's frame introduces θ into the problem. Computations show that it is no easier to match an energy to the controlled equations when θ is involved, so we consider the case when f is simply a function of ω and A. We furthermore assume that $A^{-1}f$ is a function of Ω and Γ, as was done with the translating plane.

As with the translating plane, we first consider whether or not a force of the form $F(\Omega, \Gamma)$ can induce stability by following the argument used in Proposition 5.3. Linearizing the controlled equations and computing the characteristic polynomial of the resulting 10 by 10 matrix produces

$$p(\lambda) = \lambda^4(\lambda^6 + a_5\lambda^5 + a_4\lambda^4 + a_3\lambda^3 + a_2\lambda^2 + a_1\lambda^1 + a_0).$$

The choice of F allows for a great deal of control over the coefficients $a_5, \dots, a_0$. However, the best one can hope for is to choose F so that the real parts of six of the roots of $p(\lambda)$ are all negative, while four roots must be zero. The three constants to the controlled equations, $\langle \Gamma, \Gamma \rangle$, $\langle \Gamma, \gamma \rangle$, and $\langle M_\Omega, \Gamma \rangle$, will reduce the flow to seven dimensional submanifolds, where one zero eigenvalue of the linearization remains. In constrast, the two constants to the controlled equations on the translating plane reduced the flow to submanifolds where all eigenvalues were real and negative. Hence the proof used in Proposition 5.3 cannot be applied to the sliding plane unless one can establish more symmetry for the generic control force $F(\Omega, \Gamma)$, or choose one which injects more symmetry into the system.

Finally, we see if our solution to stabilizing Chaplygin's sphere on the translating plane with a control force $F(\Gamma)$ generalizes to the sliding plane. Recall that with the translating plane, any potential energy $U(\Gamma)$ which we added to the Lagrangian could be associated with a horizontal control force $F(\Gamma)$. This made the matching step in the method of controlled Lagrangians simple, and we were able to generate a control force $F(\Gamma)$ which stabilized the system. The analogous step for the sliding plane is to add a potential energy $U(\Gamma, \gamma)$ to the uncontrolled Lagrangian for the sliding plane and to match the resulting Euler-Poincaré equation to the controlled equations with control force $F(\Gamma)$. What we find is that the matching process is much more restrictive than on the sliding plane, and that the only control force that we can match in this manner is $F = 0$. This is the content of the following proposition.

Proposition 5.9. *Let $l, \mathcal{D}$ denote the Lagrangian and distribution for Chaplygin's sphere on the sliding plane as described in equations (5.29) and (5.30). Let $\bar{l} = l - U(\Gamma, \gamma)$ where $U : S^2 \times \mathbb{R}^3 \to \mathbb{R}$ is such that the equations of motion corresponding to $\bar{l}, \mathcal{D}$ match the* *controlled equations for Chaplygin's sphere on the sliding plane as described by Proposition 5.7 for a horizontal control force $f(A^{-1}e_3)$. Then $f = 0$.*

Proof. By applying Theorem 5.2, we obtain the equations of motion for $(\bar{l}, \mathcal{D})$. Ignoring the $\dot{\Gamma}$ and $\dot{\gamma}$ equations, one has

$$\dot{M} + [\xi, M] = -B^{*}\frac{\partial U}{\partial \gamma} + \rho_{\Gamma}^{*}\frac{\partial U}{\partial \Gamma} + \rho_{\gamma}^{*}\frac{\partial U}{\partial \gamma}. \tag{5.42}$$

This is assumed to agree with the controlled equation

$$\dot{M} + [\xi, M] = -B^{*}F. \tag{5.43}$$

where F is a function of Γ alone. Assuming equation (5.42) and (5.43) agree, U must satisfy the following two equations:

$$-r\Gamma \times F = -r\Gamma \times \frac{\partial U}{\partial \gamma} + \Gamma \times \frac{\partial U}{\partial \Gamma} + \gamma \times \frac{\partial U}{\partial \gamma}. \tag{5.44}$$

$$\langle \gamma \times \Gamma, F \rangle = \langle \gamma \times \Gamma, \frac{\partial U}{\partial \gamma} \rangle. \tag{5.45}$$

Taking dot products of both sides of equation (5.44) with Γ shows that

$$\langle \gamma \times \Gamma, \frac{\partial U}{\partial \gamma} \rangle = 0.$$

Equation (5.45) then implies that $\langle \gamma \times \Gamma, F \rangle = 0$. As we have assumed that F comes from the horizontal force f, we also have that $\langle \Gamma, F \rangle = 0$. Therefore since $(\Gamma, \gamma, \Gamma \times \gamma)$ are orthogonal, $F(\Gamma)$ must be in the span of γ. That is the vector equation $F(\Gamma) = \beta(\gamma, \Gamma)\gamma$ holds where $\beta(\gamma, \Gamma)$ is a scalar function, but this can only be true if $F(\Gamma) = 0$. □

This process of using a potential energy to match the controlled equations may be fruitful if one allows for a control force $F(\Gamma, \gamma)$. This would involve measuring y in the inertial frame in such a way that $A^{-1}f$ is a function of Γ and γ. However, to find a stabilizing control force $F(\Gamma)$ by the method of controlled Lagrangians, one will most likely be required to consider modifications to the kinetic energy as well.

REFERENCES

[A] Arnold. V. [1988] *Dynamical Systems III.* Springer-Verlag, New York.

[AM] Abraham. R. and J.E. Marsden [1978] *Foundations of Mechanics.* Second Edition. Benjamin/Cummings Publishing Company, Inc.

[BKMM] Bloch. A.M.. P.S. Krishnaprasad. J.E. Marsden, and R. Murray [1996] Nonholonomic mechanical systems with symmetry. *Arch. Rat. Mech. An.*, **136**, 21–99.

[BKMR] Bloch. A.M., P.S. Krishnaprasad, J.E. Marsden. and T.S. Ratiu [1996] Euler-Poincaré Equations and Double Bracket Dissipation.. *Comm. Math. Phys.*, **175**, 1–42.

[BLMa] Block. A.M., Lenoard, N.M., and J.E. Marsden [1997] Stabilization of Mechanical Systems Using Controlled Lagrangians *Proc CDC*, **36**, 2356–2361.

[BLMb] Block, A.M., Lenoard, N.M., and J.E. Marsden [1998] Matching and Stabilization by the Method of Controlled Lagrangians *Proc CDC*, **37** 1446–1451.

[BLMc] Block. A.M., Lenoard, N.M., and J.E. Marsden [1999] Potential Shaping and the Method of Controlled Lagrangians *Proc CDC*, **38** 1653–1657.

[C] Cushman, R. [1997] Routh's Sphere. Preprint 1032, Department of Mathematics, Utrecht University, 1997, 24 pages. *http://www.math.uu.nl/publications/*

[CKS] Cushman, R., D. Kempainen and J. Śniatychki [1998] A Classical Particle with Spin Realized by Reduction of a Nonlinear Nonholonomic Constraints *Reports on Mathematical Physics*, **41**, no 1. 133-142.

[CHK] Cushman, R., J. Hermans and D. Kempainen [1996] The Rolling Disk. *Progress in Nonlinear Differential Equations and Their Applications*, **19**.

[CHMR] Cendra, H., D.D. Holm, J.E. Marsden, and T.S. Ratiu [1998] Lagrangian Reduction, the Euler-Poincaré Equations, and Semidirect Products. *Amer. Math. Soc. Transl.* (2) Vol. **186**, 1–25.

[FK] Fedorov Yu. N. and V. V. Kozlov [1995] Various Aspects of n-Dimensional Rigid Body Dynamics. *Amer. Math. Soc. Transl.* (2) Vol. **168**, 141–171.

[KM] Koon, W.S. and J.E. Marsden [1997] Optimal control for holonomic and nonholonomic mechanical systems with symmetry and Lagrangian reduction. *SIAM J. Control and Optim.* **35**, 901–929.

[HMR] Holm, D.D., J.E. Marsden, and T.S. Ratiu [1998] The Euler-Poincaré Equations and Semidirect Products with Applications to Continuum Theories. *Adv. in Math.*, **137**, 1-81.

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

[NF] Neimark, Ju. I. & N.A. Fufaev [1972] *Dynamics of Nonholonomic Systems.* Translations of Mathematical Monographs, AMS, vol. **33**.

[O] Ostrowski, J. [1995] *Geometric Perspectives on the Mechanics and Control of Undulatory Locomotion*, PhD dissertation, California Institute of Technology.

[P] Poincarè [1901] *Sur une forme nouvelle des èquations de la Mècanique*, Comptes Rendus des Sèances de L'Acadèmie Des Sciences **18**.

[Pa] Papastavridis, J.G. [1997] Time-Integral Variational Principles for Nonlinear Nonholonomic Systems. *J. Appl. Mech.* Vol. **64**, no. **4**, 985-991.

[R] Ratiu, J. [1980] The N Dimensional Free Rigid Body, *Indiana U. Math. J.*, **29**, 609-627.

[Ro] Routh, E.J. [1860] *Treatise on the Dynamics of a System of Rigid Bodies.* MacMillan, London.

[V] Varadarajan, V.S. [1984] *Lie Groups, Lie Algebras, and Their Representations.* Graduate Texts in Mathematics, **102**, Springer-Verlag.

[VV] Veselov, A.P. and L.E. Veselova [1989] Integrable Nonholonomic Systems on Lie Groups. *Math Notes*, **46** 810-818.

[ZBLM] Zenkov, D. V., Bloch A. M., Leonard, N. E. and J. E. Marsden [2000] Matching and Stabilization of the Unicycle with Rider. *IFAC Proceedings*, Princeton, March 16-18.

[ZBM] Zenkov, D. V., Bloch, A.M., and J.E. Marsden [1998] The Energy-Momentum Method for the Stability of Nonholonomic Systems. *Dyn. Stab. of Systems.* **13**, 123-166.

Vita

David Schneider received his Bachelor of Arts in mathematics at U.C. Santa Cruz in 1991. He then worked for two years as a C++ technical support engineer at Borland International, in Scotts Valley CA. From 1993 to the year 2000, he worked towards and obtained his Ph.D. in mathematics at the University of Washington in Seattle. His interests include art, computer graphics, music, ultimate frisbee and cartooning.