Controlled Lagrangian and Hamiltonian Systems

Thesis by
Dong Eui Chang

In Partial Fulfillment of the Requirements
for the Degree of
Doctor of Philosophy

California Institute of Technology
Pasadena, California

2002

(Defended April 30, 2002)

To my mother and father

Acknowledgements

First and foremost, I would like to thank my Lord and Savior, Jesus Christ, Who has provided me with His strength, knowledge, and wisdom.

I would like to thank my adviser, Professor Jerrold E. Marsden, who introduced me to geometric mechanics and geometric control theory with enthusiasm and patience. Every conversation with him was itself another lesson in creative thinking and deep insights.

I would also like to thank Professors Joel Burdick, Steven Low, Hideo Mabuchi, and Richard Murray for serving on my thesis committee and offering their constructive comments.

It is with pleasure that I express my gratitude to my collaborators: Professors David Chichka, Anthony Bloch, Naomi E. Leonard, Craig Woolsey, Dmitry Zenkov, Wei Kang, and Arthur Krener.

I am grateful to Professors In-Joong Ha (my Master's thesis adviser), Wook Hyun Kwon, Chong-Ho Choi, Jang Gyu Lee, Sang Uk Lee, Beom Hee Lee, Sang Kyun Cha, Jin Young Choi, Dong-il Choi, and Jaehong Park in the department of control and instrumentation engineering at Seoul National University for my fundamental education in control engineering.

I was fortunate to have so many great friends during my time at Caltech, including Peter Lee, Tim Yoo, Ki Baek Kim, Sean Humbert, Jay Bartroff, Matt "Angel Eyes" Perlmutter, Banavara "Tuco" Shashikanth, and my beloved Jung Ah Woo.

I especially thank my parents, my bother, my sister, and her family for their love and support.

Controlled Lagrangian and Hamiltonian Systems

by

Dong Eui Chang

In Partial Fulfillment of the
Requirements for the Degree of
Doctor of Philosophy

Abstract

Many control systems are mechanical systems. The unique feature of mechanical systems is the notion of energy, which gives much information on the stability of equilibria. Two kinds of forces are associated with the energy: dissipative force and gyroscopic force. A dissipative force is, by definition, a force which decreases the energy, and a gyroscopic force is, by definition, a force that does not change the energy. Gyroscopic forces add couplings to the dynamics. In this thesis, we develop a control design methodology which makes full use of these three physical notions: energy, dissipation, and coupling.

First, we develop the method of controlled Lagrangian systems. It is a systematic procedure for designing stabilizing controllers for mechanical systems by making use of energy, dissipative forces, and gyroscopic forces. The basic idea is as follows: Suppose that we are given a mechanical system and want to design a controller to asymptotically stabilize an equilibrium of interest. We look for a feedback control law such that the closed-loop dynamics can be also described by a new Lagrangian with a dissipative force and a gyroscopic force where the energy of the new Lagrangian has a minimum at the equilibrium. Then we check for asymptotic stability by applying the Lyapunov stability theory with the new energy as a Lyapunov function.

Next, we show that the method of controlled Lagrangian systems and its Hamiltonian counterpart, the method of controlled Hamiltonian systems, are equivalent for simple mechanical systems where the underlying Lagrangian is of the form kinetic minus potential energy. In addition, we extend both the Lagrangian and Hamiltonian sides of this theory to include systems with symmetry and discuss the relevant reduction theory.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Mechanical systems constitute a large part of control systems. Despite this, most control theories have been developed for the generic form of control systems without taking into account particular features of mechanical systems. The unique feature of mechanical systems is the notion of energy. Energy gives much information on the stability of equilibria. Two kinds of forces are associated with the energy: dissipative force and gyroscopic force. A dissipative force is, by definition, a force which decreases the energy, and a gyroscopic force is, by definition, a force which adds couplings to the dynamics without changing the energy. In this thesis, we develop a control design methodology which makes full use of the following physical notions: energy, dissipation, and coupling.

Brockett [1976] first singled out mechanical systems among general control systems. He not only made clear the distinction between mechanics and control theory but also blended these two fields. Since then, the control of mechanical systems has been considerably developed on the Hamiltonian side; see, for example, van der Schaft [1982, 1986], Crouch and van der Schaft [1987], Bloch and Marsden [1990], Bloch, Krishnaprasad, Marsden, and Sánchez De Alvarez [1992], Jalnapurkar and Marsden [1999, 2000]. In particular, van der Schaft [1986] introduced the potential shaping technique by modifying the potential energy part of the Hamiltonian function through feedback, and Bloch and Marsden [1990] initiated the kinetic energy shaping and the Poisson structure modification. These works were continued by Woolsey and Leonard [1999], Ortega and Spong [2000], Ortega, Spong, Gómez-Estern, and Blankenstein [2001]. All of these works were done in coordinates-dependent language under different names and versions. Finally, Chang, Bloch, Leonard, Marsden, and Woolsey [2002] and Chang [2002] gave a definitive intrinsic formulation of the energy shaping method and the Poisson structure modification, naming it the method of controlled Hamiltonian (CH) systems. This new formulation not only gives a foundational setting to the existing theory but also simplifies reduction of CH systems with symmetry.

The development of control of mechanical systems on the Lagrangian side was made in a series of papers by Bloch, Leonard, and Marsden [1997, 1998, 1999a, 1999b, 2000]. This method was named the method of controlled Lagrangian (CL) systems. The kinetic shaping technique and the total energy (both kinetic plus potential energy) shaping technique were launched in Bloch, Leonard, and Marsden [1997] and Bloch, Leonard, and Marsden [1999b], respectively. The introduction of the total energy shaping on the Lagrangian side by Bloch, Leonard, and Marsden [1999b] preceded that on the Hamiltonian side introduced by Ortega and Spong [2000]. Bloch, Chang, Leonard, and Marsden [2001] then found a general class of mechanical systems for which stabilizing controllers can be designed through the CL method. The CL method was also studied from the viewpoint of Riemannian geometry by interpreting mass tensors as metrics in Auckly, Kapitanski, and White [2000], Auckly and Kapitanski [2001], Hamberg [1999, 2000], where the first two developed λ-method to solve systematically the PDE's involved in the CL method. Furthermore, Hamberg noticed that the CL method can be understood as an equivalence relation by feedback transformations. All the papers based on the CL method so far have used the two concepts of energy and dissipation without considering gyroscopic forces. In addition, these works assumed that the originally given system does not have any external forces. Woolsey [2001] studied the effects of dissipative external forces for the first time. Chang, Bloch, Leonard, Marsden, and Woolsey [2002] and Chang [2002] generalized the existing CL method and showed the equivalence of the CL method and the CH method for simple mechanical systems. This generalization allows one to make use of gyroscopic forces as well as energy and dissipation and to extend the λ-method so that one can solve the PDE's more generally. Chang, Bloch, Leonard, Marsden, and Woolsey [2002] and Chang [2002] have kept the Euler-Lagrange formulation, rather than using Riemannian geometry, so that one can perform reduction of CL systems with symmetry by reducing the variational principles.

This thesis is based on Bloch, Chang, Leonard, and Marsden [2001], Chang and Marsden [2000], Bloch, Chang, Leonard, Marsden and Woolsey [2000], Chang, Bloch, Leonard, Marsden, and Woolsey [2002], Chang [2002]. Below, we outline the thesis chapter by chapter.

Chapter 2. We define controlled Lagrangian (CL) systems on a tangent bundle TQ and introduce the CL-equivalence relation to the class of CL systems on TQ. If two CL systems are CL-equivalent, then for any control for one system there exists a control for the other system such that the two closed-loop systems produce the same equations of motion. We then give the usual procedure of applying the CL method to the design of asymptotically stabilizing controllers by making use of the three mechanical notions: energy, dissipation, and coupling. We extend the existing λ-method such that we can not only solve the PDE's involved in the CL method with more freedom than before, but also add gyroscopic forces (coupling forces) into the dynamics. We illustrate the CL method in several examples.

Chapter 3. We define controlled Hamiltonian (CH) systems on a cotangent bundle T^*Q, and the CH equivalence which is analogous to the CL equivalence for CL systems. We prove that the CL method and the CH method are equivalent for simple mechanical systems. The key to the proof is the identification of the failure of the Jacobi identity in terms of gyroscopic forces. We also give the usual procedure of applying the CH method to the design of asymptotically stabilizing controllers.

Chapter 4. We consider CL/CH systems with symmetry and perform the reduction. This reduction leads to the method of reduced CL/CH systems. The reduced CL method and the reduced CH method are equivalent for reduced simple mechanical systems. We apply the reduced CL method to a couple of examples: the satellite with a rotor and the heavy top with two rotors. Our work on the reduction is based on Cendra, Marsden, and Ratiu [2001] and Marsden and Ratiu [1999].

Appendix A. We demonstrate the usefulness of symmetry by considering the problem of orbit transfer between two elliptic orbits around the earth. In the CL method, energy is used to construct Lyapunov functions. In this orbit transfer problem, we employ angular momentum and Laplace(-Runge-Lenz) vectors in the construction of a Lyapunov function, where the angular momentum vector is due to rotational symmetry and the Laplace vector is from hidden rotational symmetry. This work was published in Chang, Chichka, and Marsden [2002].

Chapter 2

The Method of Controlled Lagrangian Systems

In this chapter we develop the method of controlled Lagrangian systems and apply it to a control synthesis for asymptotic stabilization of mechanical systems. We first outline the chapter in the following.

A controlled Lagrangian (CL) system is a triple (L, F, W) of a Lagrangian L on the tangent bundle TQ, an external force $F : TQ \to T^*Q$, and a control bundle $W \subset T^*Q$. Feedback controls are maps of TQ to W. The equations of motion of a CL system (L, F, W) with a control u are the usual Euler-Lagrange equations with the external force F and the control force u as follows:

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}} - \frac{\partial L}{\partial q} = F + u.$$

Once we choose a control law u, we denote the triple (L, F, u) the closed-loop (Lagrangian) system. We call a CL system *simple* if the Lagrangian has the form of kinetic minus potential energy. The incorporation of the external force F into the definition of the CL system is important not only for the sake of generality but also for the possible use of gyroscopic forces for asymptotic stabilization. The importance of the gyroscopic force has been noticed in the control of a coupled rigid body by Wang and Khrishnaprasad [1992].

It is often difficult to design a control to meet specifications with the initial form of the CL system. To transform a given system to a better form such that the design of the controller becomes easier, we introduce an equivalence relation by feedback transformations among the CL systems on TQ as follows:

Two CL systems are CL-equivalent if for an arbitrary control law for one system, there exists a control law for the other system such that the two closed-loop systems produce the same equations of motion.

This is a normal form method in a broad sense. However, the difference is that we allow the dynamics only of the CL system form (L, F, u). Keeping the form of CL systems has a great advantage in designing an asymptotically stabilizing controller for a given mechanical system because the total energy of the system becomes a natural candidate for a Lyapunov function. For a generic control system, there are no such concepts as an energy.

Let us give the usual procedure for applying the concept of CL equivalence relation to asymptotic stabilization of an equilibrium for a given CL system of the form $(L, F = 0, W)$. In most cases, the energy E of the given system does not have a minimum at the equilibrium of interest, so we cannot directly use the energy E as a Lyapunov function. Thus, we find a CL system $(\widehat{L}, \widehat{F}, \widehat{W})$, which is CL-equivalent to the given system $(L, 0, W)$, where the energy $\widehat{E}$ of $\widehat{L}$ has a minimum at the equilibrium and $\widehat{F}$ has the form of gyroscopic force as follows:

$$\widehat{F}(q, \dot{q}) = S(q, \dot{q})\dot{q}, \qquad S^T = -S.$$

Notice that such forces do not change the energy. Since the equilibrium is a minimum point of $\widehat{E}$, we choose a dissipative feedback control $\widehat{u} : TQ \to \widehat{W}$ for the CL system $(\widehat{L}, \widehat{F}, \widehat{W})$ to decrease the energy $\widehat{E}$. A dissipation $\widehat{u}$ is of the form

$$\widehat{u}(q, \dot{q}) = -D(q, \dot{q})\dot{q}, \qquad D^T = D \geq 0.$$

The equilibrium becomes a Lyapunov stable point in the closed-loop system $(\widehat{L}, \widehat{F}, \widehat{u})$. Using CL equivalence, we can derive a stabilizing control u for the original CL system $(L, 0, W)$. Then, we apply LaSalle's theorem to check the asymptotic stability of the equilibrium in the closed-loop system $(\widehat{L}, \widehat{F}, \widehat{u})$, or equivalently, $(L, 0, u)$. In practice, one usually finds a parameterized family of CL systems which are CL-equivalent to the original system, and then one shapes the energy $\hat{E}$ by finding a set of appropriate parameters so that $\hat{E}$ has a minimum at the equilibrium *and* provides a large region of attraction.

Auckly, Kapitanski, and White [2000] and Auckly and Kapitanski [2001] developed the λ-method to systematically solve the PDE's involved in finding simple CL systems which are CL-equivalent to a given simple CL system. However, they only considered simple CL systems of the form $(L, 0, W)$, i.e., systems without external forces. Here, we extend the λ-method to include the full form of simple CL system (L, F, W). This will not only enhance the solvability of the PDE's, but also will make possible the use of gyroscopic forces for asymptotic stabilization. We will apply the extended λ-method to the asymptotic stabilization of the inverted pendulum on a rotor arm system in this thesis.

In the body of this chapter, we will cast the outlined theory above into a mathematically rigorous formulation and illustrate the method by applying it to the following two systems: the inverted pendulum on a cart and the inverted pendulum on a rotor arm. We show that the collocated/noncollocated partial feedback linearization by Spong [1997] can be expressed in the framework of the CL method.

2.1 Controlled Lagrangian (CL) Systems

In this section, we develop the method of CL systems. We first define CL systems and the CL-equivalence relation. We also introduce concepts of energy, dissipation and gyroscopic forces, which are important for the design of asymptotically stabilizing controllers.

2.1.1 Controlled Lagrangian Systems

Review of Lagrangian Mechanics. We briefly review the Lagrangian mechanics. More details can be found in Marsden and Ratiu [1999]. Consider a configuration manifold Q and the tangent space TQ. We consider a function $L : TQ \to \mathbb{R}$ called the ***Lagrangian***. Hamilton's principle of critical action states

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

where we take variations among paths $q^i(t)$ in Q with fixed end points, i.e., $\delta q^i(a) = \delta q^i(b) = 0$. It follows

$$\int \left[\frac{\partial L}{\partial q^i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}^i} \right] \delta q^i dt = 0,$$

where we used the chain rule, integration by parts, and the boundary conditions $\delta q^i(a) = \delta q^i(b) = 0$. Since this holds for all such variations, it follows

$$\frac{\partial L}{\partial q^i} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}^i} = 0$$

which are called the ***Euler-Lagrange equations***. By the chain rule, the Euler-Lagrange equations become

$$\frac{\partial L}{\partial q^i} - \frac{\partial^2 L}{\partial \dot{q}^i \partial q^j}\dot{q}^j - \frac{\partial^2 L}{\partial \dot{q}^i \partial \dot{q}^j}\ddot{q}^j = 0. \tag{2.1}$$

Motivated by this, we define the Euler-Lagrange operator. The ***Euler-Lagrange operator*** $\mathcal{EL}$ assigns to a Lagrangian $L : TQ \to \mathbb{R}$, a bundle map[1] $\mathcal{EL}(L) : T^{(2)}Q \to T^*Q$ which may be written in local coordinates (employing the summation convention) as

$$\mathcal{EL}(L)_i(q,\dot{q},\ddot{q})\mathbf{d}q^i = \left(\frac{d}{dt}\frac{\partial L}{\partial \dot{q}^i}(q,\dot{q}) - \frac{\partial L}{\partial q^i}(q,\dot{q})\right)\mathbf{d}q^i. \tag{2.2}$$

The first term on the right-hand side is regarded as a function on the second order tangent bundle $T^{(2)}Q$ by formally applying the chain rule and then replacing everywhere dq/dt by $\dot{q}$ and $d\dot{q}/dt$ by $\ddot{q}$. Hence the Euler-Lagrange equations of a Lagrangian L may be written as

$$\mathcal{EL}(L)(q,\dot{q},\ddot{q}) = 0.$$

A Lagrangian is called ***regular*** if $\det[\frac{\partial^2 L}{\partial \dot{q}^i \partial \dot{q}^j}] \neq 0$. Hence, a regular Lagrangian gives a second order vector field, $X \in \Gamma(T^{(2)}Q)$, by solving (2.1) for $\ddot{q}^i$ as follows:

$$X = \begin{bmatrix} \dot{q}^i \\ \ddot{q}^i \end{bmatrix} = \begin{bmatrix} \dot{q}^i \\ M^{ij}[\frac{\partial L}{\partial q^j} - \frac{\partial^2 L}{\partial \dot{q}^j \partial q^k}\dot{q}^k] \end{bmatrix},$$

where M^{ij} is the inverse matrix of $\frac{\partial^2 L}{\partial \dot{q}^i \partial \dot{q}^j}$.

[1]The second order tangent bundle $\tau_Q^{(2)} : T^{(2)}Q \to Q$ is defined as follows. For $\bar{q} \in Q$, elements of $T_{\bar{q}}^{(2)}Q$ are equivalence classes of curves in Q, namely, two curves $q_i(t)$, $i = 1, 2$, with $q_1(\bar{t}) = q_2(\bar{t}) = \bar{q}$ are equivalent, by definition, if and only if in any local chart we have $q_1^{(l)}(\bar{t}) = q_2^{(l)}(\bar{t})$ for $l = 1, 2$ where $q^{(l)}(t)$ denotes the derivative of order l.

A Lagrangian L is called ***simple*** if it has the kinetic minus potential energy form:

$$L = \frac{1}{2} m_{ij}(q)\dot{q}^i\dot{q}^j - V(q),$$

where $m = (m_{ij})$ is a Riemannian metric on Q. The Euler-Lagrange operator for a simple Lagrangian is written as

$$\mathcal{EL}(L)(q, \dot{q}, \ddot{q})_k = m_{ki}\ddot{q}^i + [ij,k]\dot{q}^i\dot{q}^j + \frac{\partial V}{\partial q^k} \tag{2.3}$$

with the Christoffel symbol of the first kind

$$[ij,l] = \frac{1}{2}\left[\frac{\partial m_{il}}{\partial q^j} + \frac{\partial m_{jl}}{\partial q^i} - \frac{\partial m_{ij}}{\partial q^l}\right]. \tag{2.4}$$

The Euler-Lagrange equations in (2.1) can also be written equivalently as

$$\nabla_{\dot{q}}\dot{q} + m^{-1}\mathbf{d}V = 0,$$

or, in coordinates

$$\ddot{q}^i + \Gamma^i_{jk}\dot{q}^j\dot{q}^k + m^{ij}\frac{\partial V}{\partial q^j} = 0,$$

where ∇ is the Levi-Civita connection of the metric m, and Γ^i_{jk} is the Christoffel symbol of the second kind defined as

$$\Gamma^i_{jk} = m^{il}[jk,l].$$

Controlled Lagrangian Systems. The concept of the pure Euler-Lagrange equations will be generalized to include external forces and also will be made intrinsic (independent of a specific coordinate representation). This definition is fundamental to the Lagrangian side of this work.

Definition 2.1.1. *A* ***controlled Lagrangian (CL) system*** *is a triple* (L, F, W) *where the function* $L : TQ \to \mathbb{R}$ *is the regular Lagrangian, the fiber-preserving map* $F : TQ \to T^*Q$ *is an external force and* $W \subset T^*Q$ *is a subbundle of* T^*Q *called* **control bundle** *representing the actuation directions.*

Sometimes, we will identify the subbundle W with the set of bundle maps from TQ

to W. The fact that W may be smaller than the whole space corresponds to the system being *underactuated.* The equations of motion of the system (L, F, W) may be written as

$$\mathcal{E}\mathcal{L}(L)(q, \dot{q}, \ddot{q}) = F(q, \dot{q}) + u \tag{2.5}$$

with a control u selected from W. When we choose a specific control map $u : TQ \to W$ (so that u is a function of $(q^i, \dot{q}^i)$), we call the triple (L, F, u) a ***closed-loop (Lagrangian) system.*** If u is a map from $\mathbb{R}$ to W, it can be considered as an open-loop control. We will typically be interested in feedback controls in this thesis.

In the special case when W is integrable (that is, its annihilator $W^o \subset TQ$ is integrable in the usual Frobenius sense) and if we choose coordinates appropriately, equations (2.5) can be locally written in coordinates as

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}^i} - \frac{\partial L}{\partial q^i} = F_i + u_i, \quad i = 1, \ldots, k \tag{2.6}$$

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}^i} - \frac{\partial L}{\partial q^i} = F_i, \quad i = k+1, \ldots, n. \tag{2.7}$$

Here the coordinates $q^1, \ldots, q^k$ are chosen so that $\mathbf{d}q^1, \ldots \mathbf{d}q^k$ span W, so W is k dimensional in this case. The external forces can include gyroscopic forces, friction forces, etc.

Simple CL Systems. In most engineering applications, Lagrangian functions usually have the form of kinetic minus potential energy. Hence, we introduce the following definition.

Definition 2.1.2. *A CL system (L, F, W) is called a* ***simple CL system*** *if the Lagrangian L has the form of kinetic minus potential energy:*

$$L(q, \dot{q}) = \frac{1}{2}m(q)(\dot{q}, \dot{q}) - V(q),$$

where m is a nondegenerate symmetric $(0, 2)$-tensor (or, shortly the mass tensor).

We will sometimes omit the q-dependence of m in the notation keeping this dependence implicitly understood. When L is a simple Lagrangian, the Euler-Lagrange operator is written in a vector form as

$$\mathcal{E}\mathcal{L}(L)(q,\dot{q},\ddot{q}) = m\ddot{q} + (\mathbf{d}m[\dot{q}])\dot{q} - \frac{\partial L}{\partial q}, \tag{2.8}$$

where $(\mathbf{d}m[\dot{q}])_{ij} = \mathbf{d}(m_{ij})[\dot{q}]$. Also see (2.3).

Matching Conditions, CL Equivalence, and CL Inclusion. We are ready to embark on the matching problems of CL systems. We present the result only for simple CL systems for the sake of simplicity and to make the exposition more concrete. One can readily generalize the results to more general forms of Lagrangians (see Remark 2.1.9).

Consider now two *simple* CL systems (L_1, F_1, W_1) and (L_2, F_2, W_2) with

$$L_1(q,\dot{q}) = \frac{1}{2}m_1(\dot{q},\dot{q}) - V_1(q) \quad \text{and} \quad L_2(q,\dot{q}) = \frac{1}{2}m_2(\dot{q},\dot{q}) - V_2(q). \tag{2.9}$$

We denote by $\ddot{q}_{L_\alpha}$ the $\ddot{q}$ equation of the closed-loop system $(L_\alpha, F_\alpha, u_\alpha)$ with $\alpha = 1,2$, which is given in coordinates, using matrix and vector-style notation by

$$\ddot{q}_{L_\alpha} = m_\alpha^{-1}\left[-(\mathbf{d}m_\alpha[\dot{q}])\dot{q} + \frac{\partial L_\alpha}{\partial q} + F_\alpha + u_\alpha\right]. \tag{2.10}$$

We can then formally define matching conditions between the two systems (L_1, F_1, W_1) and (L_2, F_2, W_2).

Definition 2.1.3. *Given the two systems (L_i, F_i, W_i), $i = 1,2$, the **Euler-Lagrange matching conditions** are*

ELM-1 : $W_1 = m_1 m_2^{-1}(W_2)$,

ELM-2 : $\mathrm{Im}\left[(\mathcal{E}\mathcal{L}(L_1) - F_1) - m_1 m_2^{-1}(\mathcal{E}\mathcal{L}(L_2) - F_2)\right] \subset W_1$,

where Im *means the pointwise image of the map in brackets.*

*We say that the two simple CL systems (L_i, F_i, W_i), $i = 1,2$, are **CL-equivalent** if **ELM-1** and **ELM-2** hold. We use the symbol $\overset{L}{\sim}$ for this equivalence relation.*

Claim 2.1.4. *The relation $\overset{L}{\sim}$ is an equivalence relation.*

Proof. The reflexivity and the transitivity are obvious. The symmetry follows if we multiply both sides of **ELM-1** and **ELM-2** by $m_2 m_1^{-1}$. ■

One can easily check by coordinate computation that the map

$$\left[(\mathcal{EL}(L_1) - F_1) - m_1 m_2^{-1}(\mathcal{EL}(L_2) - F_2)\right] \tag{2.11}$$

in **ELM-2** can be regarded as a map defined on TQ because the acceleration terms, $\ddot{q}$, from the two Euler-Lagrange expressions cancel each other.

The following proposition explains the main property of the CL equivalence.

Proposition 2.1.5. *Suppose that two simple CL systems (L_i, F_i, W_i), $i = 1, 2$ are CL-equivalent. Then, for an arbitrary control law for one system, there exists a control law for the other system such that the two closed-loop simple CL systems produce the same equations of motion. The explicit relation between the two control laws u_i, $i = 1, 2$ is given by*

$$u_1 = \mathcal{EL}(L_1) - F_1 - m_1 m_2^{-1}(\mathcal{EL}(L_2) - F_2) + m_1 m_2^{-1} u_2, \tag{2.12}$$

where m_i is the mass tensor of L_i, $i = 1, 2$.

Proof. Denote by $\ddot{q}_{L_i}$ the expression of the acceleration $\ddot{q}$ obtained from the closed-loop CL system (L_i, F_i, u_i), $i = 1, 2$ as in (2.10). Then,

$$m_1(\ddot{q}_{L_1} - \ddot{q}_{L_2}) = u_1 - m_1 m_2^{-1} u_2 - \left[(\mathcal{EL}(L_1) - F_1) - m_1 m_2^{-1}(\mathcal{EL}(L_2) - F_2)\right].$$

The conditions **ELM-1** and **ELM-2** imply that (2.12) holds if and only if $\ddot{q}_{L_1} = \ddot{q}_{L_2}$ if and only if they produce the same equations of motion. Notice that the term

$$\mathcal{EL}(L_1) - m_1 m_2^{-1} \mathcal{EL}(L_2)$$

in (2.12) can be regarded as a map defined on TQ because the acceleration $\ddot{q}$ cancels. ■

There is a more general concept than the CL equivalence relation. We will give this definition for simple Lagrangian systems, but it can be readily generalized for general Lagrangian systems (see Remark 2.1.9).

Definition 2.1.6. *We say that a simple CL system (L_1, F_1, W_1)* ***includes*** *the simple CL system (L_2, F_2, W_2), or simply, $(L_1, F_1, W_1) \supset (L_2, F_2, W_2)$, if the following holds:*

ELI-1 : $W_1 \supset m_1 m_2^{-1}(W_2)$,

ELI-2 : $\operatorname{Im}\left[(\mathcal{EL}(L_1) - F_1) - m_1 m_2^{-1}(\mathcal{EL}(L_2) - F_2)\right] \subset W_1$.

This introduces a partial order in the class of simple controlled Lagrangian systems on TQ. We call this partial order ***CL inclusion***. The following proposition explains the main property of the CL inclusion.

Proposition 2.1.7. *If (L_1, F_1, W_1) includes (L_2, F_2, W_2), then for any choice of control $u_2 : TQ \to W_2$, there is a control $u_1 : TQ \to W_1$ satisfying (2.12) such that the two closed-loop systems (L_1, F_1, u_1) and (L_2, F_2, u_2) produce the same equations of motion. The explicit relation between the two control laws u_i, $i = 1, 2$ is given by*

$$u_1 = \mathcal{EL}(L_1) - F_1 - m_1 m_2^{-1}(\mathcal{EL}(L_2) - F_2) + m_1 m_2^{-1} u_2,$$

where m_i is the mass tensor of L_i, $i = 1, 2$.

Proof. Mimic the proof of Proposition 2.1.5. ■

The CL equivalence can be understood in terms of the CL inclusion.

Proposition 2.1.8. *Two simple CL systems include each other if and only if they are CL-equivalent.*

Proof. Trivial. ■

Remark 2.1.9. *In this thesis we always assume that Lagrangians are regular, i.e.,*

$$\det\left(\frac{\partial^2 L}{\partial \dot{q}^i \partial \dot{q}^j}\right) \neq 0.$$

For simple CL systems this means, of course, that the mass matrix is nonsingular. We can generalize the Euler-Lagrange matching conditions for general Lagrangians which are not necessarily simple as follows. Let $L : TQ \to \mathbb{R}$ be a regular Lagrangian. It induces a globally well defined map $m_L : TQ \to \operatorname{Sym}_2(T^*Q)$ (symmetric two-tensors with indices down) given in tangent bundle charts as follows:

$$m_L(q, \dot{q}) = \frac{\partial^2 L(q, \dot{q})}{\partial \dot{q}^i \partial \dot{q}^j} dq^i \otimes dq^j.$$

Then the (generalized) Euler-Lagrange matching conditions for the general Lagrangian systems (L_1, F_1, W_1) *and* (L_2, F_2, W_2) *are given by*

ELM-a : $W_1 = m_{L_1} m_{L_2}^{-1}(W_2)$,

ELM-b : $\mathrm{Im}\left[(\mathcal{EL}(L_1) - F_1) - m_{L_1} m_{L_2}^{-1}(\mathcal{EL}(L_2) - F_2)\right] \subset W_1$.

The equation in (2.12) is replaced by

$$u_1 = (\mathcal{EL}(L_1) - F_1) - m_{L_1} m_{L_2}^{-1}(\mathcal{EL}(L_2) - F_2) - m_{L_1} m_{L_2}^{-1} u_2$$

so that the two closed-loop systems (L_1, F_1, u_1) *and* (L_2, F_2, u_2) *produce the same equations of motion.*

Coordinate Expressions of the Matching Conditions. Now we express the Euler-Lagrange matching conditions in coordinates. Let (L_i, F_i, W_i), $i = 1, 2$ be two simple CL systems with Lagrangians as in (2.9). Suppose that we are given a decomposition of the forces F_i as

$$F_i = F_i^v + F_i^q \tag{2.13}$$

with $i = 1, 2$, where F_i^q is independent of the velocity $\dot{q}$. Then **ELM-2** can be written in coordinates as

$$\begin{aligned} W_1^\circ \Bigg[&(\mathbf{d}m_1[\dot{q}])\dot{q} - \frac{\partial}{\partial q}\left(\frac{1}{2}\dot{q}^T m_1 \dot{q}\right) - F_1^v(q, \dot{q}) \\ &- m_1 m_2^{-1}\left((\mathbf{d}m_2[\dot{q}])\dot{q} - \frac{\partial}{\partial q}\left(\frac{1}{2}\dot{q}^T m_2 \dot{q}\right) - F_2^v(q, \dot{q})\right)\Bigg] = 0 \end{aligned} \tag{2.14}$$

and

$$W_1^\circ \left[\frac{\partial V_1}{\partial q} - F_1^q - m_1 m_2^{-1}\left(\frac{\partial V_2}{\partial q} - F_2^q\right)\right] = 0, \tag{2.15}$$

where W_1° is a matrix whose rows span the annihilator

$$\{v \in TQ \mid \langle v, \alpha \rangle = 0 \text{ for all } \alpha \in W_1\}$$

of W_1.

The two expressions (2.14) and (2.15) can also be written as follows:

$$(W_1^\circ)_{lk}\left[[ij,k]_1\dot{q}^i\dot{q}^j - (F_1^v)_k - (m_1)_{ka}(m_2)^{ab}([ij,b]_2\dot{q}^i\dot{q}^j - (F_2^v)_b)\right] = 0 \qquad (2.16)$$

and

$$(W_1^\circ)_{lk}\left[\frac{\partial V_1}{\partial q^k} - (F_1^q)_k - (m_1)_{ka}(m_2)^{ab}\left(\frac{\partial V_2}{\partial q^b} - (F_2^q)_b\right)\right] = 0, \qquad (2.17)$$

where $[ij,k]_\alpha$ is the Christoffel symbol of the first kind in (2.4) for the metric m_α, $\alpha = 1, 2$.

2.1.2 Energy, Dissipation, and Gyroscopic Forces

We define the concepts of energy, dissipation, and gyroscopic force. They are key elements for the design of stabilizing controllers. When the energy has its minimum at an equilibrium of interest, we can use the energy as a Lyapunov function. Dissipative forces decrease the energy along the trajectory. Gyroscopic forces add couplings to the dynamics so that the dissipation works more effectively for asymptotic stability. The role of gyroscopic forces in asymptotic stability will be illustrated later.

Energy. We define the energy E of a CL system (L, F, W) by

$$E(q,\dot{q}) = \langle \mathbb{F}L(\dot{q}_q),\ \dot{q}_q\rangle - L(q,\dot{q}),$$

where $\dot{q}_q = (q,\dot{q})$ and $\mathbb{F}L$ is the fiber derivative of L defined by

$$\langle \mathbb{F}L(v),\ w\rangle = \left.\frac{d}{dt}\right|_{t=0} L(v + tw)$$

for $v, w \in T_qQ$. In coordinates,

$$E(q,\dot{q}) = \dot{q}^i\frac{\partial L}{\partial \dot{q}^i} - L(q,\dot{q}).$$

The time derivative of E of a CL system (L, F, W) with a control $u : TQ \to W$, is given by

$$\frac{d}{dt}E = \langle F + u,\ \dot{q}\rangle. \qquad (2.18)$$

In particular, when the Lagrangian $L = \frac{1}{2} m_{ij}\dot{q}^i\dot{q}^j - V(q)$ is simple, the energy is written as

$$E = \frac{1}{2}m(\dot{q}, \dot{q}) + V(q) = \frac{1}{2}m_{ij}\dot{q}^i\dot{q}^j + V(q).$$

Remark 2.1.10. *Notice that for any CL system (L, F, W) the following holds:*

$$(L, F, W) \overset{L}{\sim} (-L, -F, W).$$

The equation (2.12) becomes $u_1 = -u_2$. If an equilibrium is a minimum point of energy of (L, F, W), then it will be a maximum point of energy of $(-L, -F, W)$. Therefore, it is not crucial whether the energy function has a maximum or a minimum at an equilibrium for stability purposes.

Dissipation. We define the dissipative force to be a force map $F : TQ \to T^*Q$ satisfying

$$\langle \dot{q},\, F(q, \dot{q})\rangle \leq 0$$

for all $(q, \dot{q}) \in TQ$. Such forces can be written as

$$F_{\text{diss}}(q, \dot{q}) = -D(q, \dot{q})\dot{q}, \qquad D^T = D \geq 0. \tag{2.19}$$

Physically, dissipation is a force which decreases the total energy; see (2.18). When the energy has a minimum at an equilibrium, dissipation helps the asymptotic stability of the equilibrium. Sometimes, dissipation alone may not be enough for asymptotic stability due to loose couplings in the dynamics. In such cases, gyroscopic forces are needed to create strong couplings in the dynamics.

Sometimes the energy has a maximum at an equilibrium of interest. In such a case, one wants to use an energy pumping force rather than a dissipative force to achieve asymptotic stability of the equilibrium, where the energy pumping force F_{ep} is of the form

$$F_{\text{ep}}(q, \dot{q}) = D(q, \dot{q})\dot{q}, \qquad D^T = D \geq 0.$$

In this case, we can always find an equivalence system whose energy has a minimum at the equilibrium by Remark 2.1.10. This is a matter of choice.

Gyroscopic Forces. Gyroscopic Forces are the forces which do not do any work. Physically, couplings in mechanical systems and gyrators in electrical systems create gyroscopic forces. We now formally define the gyroscopic force. We define a gyroscopic force to be a force F_{gr} of the following form:

$$F_{\mathrm{gr}}(q, \dot{q}) = S(q, \dot{q})\dot{q}, \qquad S^T = -S.$$

One can check

$$\langle \dot{q}, F_{\mathrm{gr}}(q, \dot{q}) \rangle = 0.$$

In the special case that $S = (S_{ij})$ depends only on q, the gyroscopic force F_{gr} can be regarded as an element of $\Gamma(\wedge^2 T^* Q)$ as follows:

$$F_{\mathrm{gr}} = S_{ik} \mathbf{d}q^k \otimes \mathbf{d}q^i = \frac{1}{2} S_{ik} \mathbf{d}q^k \wedge \mathbf{d}q^i.$$

If $\mathbf{d}F_{\mathrm{gr}} = 0$, then by the Poincaré Lemma there is a (locally defined) one form $\alpha = I_i(q)\mathbf{d}q^i$ such that $\mathbf{d}\alpha = F_{\mathrm{gr}}$(see Boothby [1986] for the Poincaré Lemma). It follows that

$$S_{ik}(q) = \frac{\partial I_i(q)}{\partial q^k} - \frac{\partial I_k(q)}{\partial q^i}.$$

Then, one can (locally) incorporate the gyroscopic force $F_{\mathrm{gr}}(q, \dot{q}) = S(q, \dot{q})\dot{q}$ into the Lagrangian as follows: $(L, F_{\mathrm{gr}}, W) \overset{L}{\sim} (L + I_i(q)\dot{q}^i, 0, W)$. However, one does not have to restrict to such special S's because the skew symmetry of S is the only property we need for the gyroscopic force from the energy-conservation point of view. This is essentially the same as using almost Poisson structures by not enforcing the Jacobi identity condition, which will be discussed in the next chapter on controlled Hamiltonian systems.

Example. We illustrate the role of gyroscopic forces in *asymptotic stability*. When there is a loose coupling in the dynamics, dissipation alone sometimes may not be enough for asymptotic stability of an equilibrium. An additional gyroscopic force creates a strong coupling so that the dissipative force together with the gyroscopic force achieves asymptotic stabilization.

Let $Q = \mathbb{R}^2$. Consider a Lagrangian system with Lagrangian

$$L = \frac{1}{2}((\dot{q}^1)^2 + (\dot{q}^2)^2) - \frac{1}{2}((q^1)^2 + (q^2)^2)$$

with an external force

$$F = \underbrace{\begin{bmatrix} 0 \\ -\epsilon\dot{q}^2 \end{bmatrix}}_{\text{dissipative force}} + \underbrace{\begin{bmatrix} 0 & -\delta\dot{q}^1 \\ \delta\dot{q}^1 & 0 \end{bmatrix} \begin{bmatrix} \dot{q}^1 \\ \dot{q}^2 \end{bmatrix}}_{\text{gyroscopic force}}$$

with $\epsilon > 0$, $\delta \in \mathbb{R}$. The equations of the motion are given by

$$\ddot{q}^1 + q^1 = -\delta\dot{q}^1\dot{q}^2 \tag{2.20}$$

$$\ddot{q}^2 + q^2 = \delta(\dot{q}^1)^2 - \epsilon\dot{q}^2 \tag{2.21}$$

If $\delta = 0$, then the two dynamics in (2.20) and (2.21) are *decoupled.* So, the gyroscopic part in F creates a coupling for $\delta \neq 0$. We want to use, as a Lyapunov function, the energy E given by

$$E = \frac{1}{2}((\dot{q}^1)^2 + (\dot{q}^2)^2) + \frac{1}{2}((q^1)^2 + (q^2)^2). \tag{2.22}$$

Then,

$$\frac{d}{dt}E = -\epsilon(\dot{q}^2)^2 \leq 0.$$

Suppose $\delta = 0$. Then the origin is the Lyapunov stable equilibrium but is not asymptotically stable because the dynamics in (2.20) is that of a pure harmonic oscillator. Suppose $\delta \neq 0$. Then one can show the asymptotic stability of the origin using LaSalle's theorem (see Khalil [1996] for LaSalle's theorem). A brief argument goes as follows: Let $(q^1(t), q^2(t), \dot{q}^1(t), \dot{q}^2(t))$ be a trajectory satisfying $\dot{E} = 0$ for all $t \geq 0$. Then it is of the form

$$(q^1(t), q^2(t), \dot{q}^1(t), \dot{q}^2(t)) = (q^1(t), q^2(0), \dot{q}^1(t), 0).$$

Substituting this to (2.21) implies

$$\dot{q}^1(t) = \text{ constant } = \dot{q}^1(0).$$

This implies that $q^1(t) = q^1(0) + \dot{q}^1(0)t$. Since the energy E in (2.22) is constant along the trajectory, it follows

$$\dot{q}^1(t) = \dot{q}^1(0) = 0.$$

Equation (2.20) now implies $q^1(t) = 0$. Hence, we have shown that the only trajectory lying in the set $\dot{E} = 0$ is the origin only. By LaSalle's theorem, the origin is asymptotically stable.

Notice that even though the gyroscopic force chosen here is small near the origin as being quadratic in the velocity, it plays a role for asymptotic stability by creating a coupling between the q^1 dynamics and the q^2 dynamics. See Merkin [1996] for the effect of *linear* gyroscopic forces on the stability. The importance of gyroscopic forces in stabilization has been observed in the control of a coupled rigid body (see Wang and Khrishnaprasad [1992]).

2.1.3 Collocated and Noncollocated Partial Feedback Linearization

The technique of collocated and noncollocated partial feedback linearization was developed to simplify the dynamics of Lagrangian systems through feedback transformation so that it makes easier the design of controllers for mechanical systems (see Spong [1997] and references therein). Here, we will understand it in the framework of controlled Lagrangian systems. It will turn out that the collocated linearization can be understood as a result of CL equivalence and that the noncollocated linearization can be derived through CL inclusion.

For simplicity, take $Q = \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$ as configuration space and use $q = (q_1, q_2) = (q_1^\alpha, q_2^i) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$ as coordinates. Let $(L, F = 0, W)$ be a given CL system (with no external forces and) with $W = \langle \mathbf{d}q_2^i \mid i = 1, \ldots, n_2 \rangle$ and

$$L(q, \dot{q}) = \frac{1}{2} \begin{bmatrix} \dot{q}_1 \\ \dot{q}_2 \end{bmatrix}^T \begin{bmatrix} m_{11}(q) & m_{12}(q) \\ m_{21}(q) & m_{22}(q) \end{bmatrix} \begin{bmatrix} \dot{q}_1 \\ \dot{q}_2 \end{bmatrix} - V(q). \tag{2.23}$$

Then the equations of motion of the closed-loop system $(L, 0, \tau)$ can be written in the following form (as in Spong [1997]):

$$m_{11}\ddot{q}_1 + m_{12}\ddot{q}_2 + h_1(q, \dot{q}) + \phi_1(q) = 0,$$
$$m_{21}\ddot{q}_1 + m_{22}\ddot{q}_2 + h_2(q, \dot{q}) + \phi_2(q) = \tau$$

with control $\tau : TQ \to W$ where h_i includes all $\dot{q}$-dependent terms and ϕ_i contains the terms from the potential energy.

Collocated Partial Feedback Linearization. Define a CL system (L_c, F_c, W_c) as follows:

$$\begin{aligned}
L_c(q, \dot{q}) &= \frac{1}{2}\dot{q}_1^T m_{11}\dot{q}_1 + \frac{1}{2}\dot{q}_2^T \dot{q}_2 \\
F_c(q, \dot{q}) &= \begin{bmatrix} -h_1 - \phi_1 + (\mathbf{d}m_{11}[\dot{q}])\dot{q}_1 - \frac{\partial L_c}{\partial q_1} \\ -\frac{\partial L_c}{\partial q_2} \end{bmatrix} \\
W_c &= \langle \mathbf{d}q_2^i - (m_{12})_{\alpha i}\mathbf{d}q_1^\alpha \mid i = 1, \ldots, n_2 \rangle \\
&= \text{the subbundle spanned by the columns of } \begin{bmatrix} -m_{12}(q) \\ I_{n_2} \end{bmatrix},
\end{aligned}$$

where I_{n_2} is the $n_2 \times n_2$ identity matrix. If a part of F_c is a potential force, i.e., the differential of a function, one can incorporate it into the Lagrangian L_c as a potential function. A control to the system (L_c, F_c, W_c) can be written via a map $u_c : TQ \to \mathbb{R}^{n_2}$ as follows:

$$\begin{bmatrix} -m_{12}(q) \\ I_{n_2} \end{bmatrix} u_c(q, \dot{q}) \in W_c(q). \tag{2.24}$$

One can check that

$$(L, 0, W) \overset{L}{\sim} (L_c, F_c, W_c)$$

and that the equations of motion of (L_c, F_c, W_c) are given by

$$m_{11}\ddot{q}_1 + h_1 + \phi_1 = -m_{12}u_c \tag{2.25}$$
$$\ddot{q}_2 = u_c \tag{2.26}$$

with control $u_c : TQ \to \mathbb{R}^{n_2}$. If we write the control for (L_c, F_c, W_c) in the form (2.24), then equation (2.12) is written as

$$\tau = h_2 + \phi_2 - m_{21}m_{11}^{-1}(h_1 + \phi_1) + (m_{22} - m_{21}m_{11}^{-1}m_{12})u_c$$

so that the closed-loop systems $(L, 0, \tau)$ and (L_c, F_c, u_c) produce the same equations of motion. This coincides with the notion of collocated partial feedback linearization in Spong [1997].

If the control bundle is integrable, one can find a set of local coordinates for the *configuration space* Q, which is useful in doing stability analysis, such as the case in § 2.3.2 and § 2.4.3. The integrability condition of the control bundle W_c is given by (the curvature condition)

$$\frac{\partial A^i_\alpha}{\partial q_2^j} A^j_\beta - \frac{\partial A^i_\beta}{\partial q_2^j} A^j_\alpha + \frac{\partial A^i_\alpha}{\partial q_1^\beta} - \frac{\partial A^i_\beta}{\partial q_1^\alpha} = 0$$

for $1 \leq \alpha, \beta \leq n_1$ and $1 \leq i, j \leq n_2$, where A^i_α is the (i, α)-th element of the matrix $m_{21} = m_{12}^T$ and $(q_1, q_2) = (q_1^\alpha, q_2^i)$; this is seen from the fact that W_c is spanned by the set $\{dq_2^i - A^i_\alpha dq_1^\alpha \mid 1 \leq \alpha \leq n_1, 1 \leq i \leq n_2\}$.

Noncollocated Partial Feedback Linearization. For noncollocated linearization, as in Spong [1997], we make the following assumption on the submatrix , m_{12} in (2.23):

$$\operatorname{rank}(m_{12}(q)) = n_1, \quad \forall q \in Q,$$

i.e., the submatrix m_{12} is onto. Then, there is a pseudo-inverse

$$m_{12}^\dagger := m_{12}^T(m_{12}m_{12}^T)^{-1}$$

such that $m_{12}m_{12}^\dagger = I_{n_1}$ with I_{n_1} the $n_1 \times n_1$ identity matrix. For this assumption to hold, it is necessary that $n_1 \leq n_2$, i.e., the number of actuation degrees of freedom should be at least as big as the number of unactuated degrees of freedom. This same property is put to good use in § 2.3.2.

Define the CL system (L_n, F_n, W_n) as follows:

$$L_n(q,\dot q) = \frac{1}{2}\dot q_1^T \dot q_1 + \frac{1}{2}\dot q_2^T \dot q_2,$$

$$F_n(q,\dot q) = \begin{bmatrix} 0 \\ -m_{12}(q)^\dagger (h_1(q,\dot q) + \phi_1(q)) \end{bmatrix}$$

$$\begin{aligned} W_n &= \langle \mathbf{d}q_1^i - (m_{12}^\dagger m_1 1)_{\alpha i}\mathbf{d}q_2^\alpha \mid i = 1,\dots,n_1\rangle \\ &= \text{subbundle spanned by the columns of } \begin{bmatrix} I_{n_1} \\ -m_{12}(q)^\dagger m_{11}(q) \end{bmatrix}, \end{aligned}$$

where, as above, one can move any potential force parts of F_n into the Lagrangian L_n as a potential function. A control for the system (L_n, F_n, W_n) can be written via a map $u_n : TQ \to \mathbb{R}^{n_1}$ as follows:

$$\begin{bmatrix} I_{n_1} \\ -m_{12}(q)^\dagger m_{11}(q) \end{bmatrix} u_n(q,\dot q) \in W_n(q). \tag{2.27}$$

Notice that dim $W_n(q) \le \dim W(q)$ because $n_1 \le n_2$. Hence, it is appropriate to use the concept of CL inclusion rather than CL equivalence. Indeed, one can easily check that

$$(L_n, F_n, W_n) \subset (L, 0, W)$$

and that the equations of motion of (L_n, F_n, W_n) are written as

$$\begin{aligned} \ddot q_1 &= u_n \\ \ddot q_2 &= -m_{12}^\dagger m_{11}(u_n + h_1 + \phi_1) \end{aligned}$$

with control $u_n : TQ \to \mathbb{R}^{n_2}$. If we write the control u_n for (L_n, F_n, W_n) in the form (2.27), then condition (2.12) is written as

$$\tau = h_2 + \phi_2 - m_{22}m_{12}^\dagger(h_1 + \phi_1) + (m_{21} - m_{22}m_{12}^\dagger m_{11})u_n$$

such that the two closed-loop systems $(L, 0, \tau)$ and (L_n, F_n, u_n) produce the same equations of motion. This coincides with the noncollocated partial feedback linearization in Spong

[1997].

2.2 Control Synthesis via a CL System

We now discuss how one may apply the concept of CL equivalence to designing stabilizing control laws for mechanical systems.

We first consider a simple but motivating example. Consider a CL system $(L_1, F = 0, W)$ with

$$L_1(x, \dot{x}) = \frac{1}{2}\dot{x}^2 + \frac{1}{2}x^2, \quad W = \langle \mathbf{d}x \rangle,$$

where $(x, \dot{x}) \in T\mathbb{R} = \mathbb{R} \times \mathbb{R}$. With a control force u, the equation of motion is given by

$$\ddot{x} - x = u.$$

The control goal is to asymptotically stabilize the equilibrium at the origin $(0, 0)$. We want to use the energy

$$E_1 = \frac{1}{2}\dot{x}^2 - \frac{1}{2}x^2$$

of $(L, 0, W)$ as a Lyapunov function, but we cannot because $(0, 0)$ is a saddle point of E_1. Let

$$L_2 = \frac{1}{2}\dot{x}^2 - \frac{1}{2}\delta x^2,$$

where $\delta \in \mathbb{R}$ is free to choose at present. Then, one can show that

$$(L_1, 0, W) \overset{L}{\sim} (L_2, 0, W).$$

We need to take $\delta > 0$ such that the energy E_2 of $(L_2, 0, W)$

$$E_2 = \frac{1}{2}\dot{x}^2 + \frac{1}{2}\delta x^2 \tag{2.28}$$

has a strict minimum at $(0, 0)$. The energy E_2 becomes a Lyapunov function candidate. Then,

$$\frac{d}{dt}E_2 = \dot{x}u_2$$

with a control u_2 for $(L_2, 0, W)$. Take

$$u_2 = -c\dot{x}$$

with $c > 0$ so that we have $dE_2/dt = -c\dot{x}^2 \leq 0$. The use of LaSalle's theorem implies the asymptotic stability of $(0,0)$ for the closed-loop system $(L_2, 0, u_2)$. Hence, by Proposition 2.1.5, the closed-loop system $(L_1, 0, u_1)$ with

$$u_1 = -x - \delta x + u_2 = -(\delta + 1)x - c\dot{x},$$

which comes from (2.12), has $(0,0)$ as an asymptotically stable equilibrium.

This procedure is traditionally called the potential energy shaping because we *shaped* the original energy $E_1 = \frac{1}{2}\dot{x}^2 - \frac{1}{2}x^2$ into $E_2 = \frac{1}{2}\dot{x}^2 + \frac{1}{2}\delta x^2$ by adding the additional *potential term* $\frac{1}{2}(\delta + 1)x^2$, and by taking $\delta > 0$. In the above example, the potential energy shaping alone was enough for stabilization because the system was fully actuated. When a system is underactuated, one sometimes needs to do the kinetic energy shaping as well. The method of CL systems allows both the kinetic and the potential energy shaping. More detail will follow.

2.2.1 Control Synthesis Procedure

We apply the CL method to the stabilization problem. We want to design a control law to asymptotically stabilize an equilibrium $(q_e, 0) \in TQ$ of a given CL system $(L_1, F_1 = 0, W_1)$ using its energy E_1 as a Lyapunov function, if possible. Usually, the equilibrium $(q_e, 0)$ is not a (strict) minimum of the energy E_1, which prevents us from directly using the energy E_1 as a Lyapunov function.

Here is the usual procedure of applying the method of CL systems to stabilization problems:

1. find a CL system (L_2, F_2, W_2) CL-equivalent to $(L_1, F_1 = 0, W_1)$ where the energy E_2 of L_2 has a strict minimum at the equilibrium $(q_e, 0)$ and F_2 has the form of a gyroscopic force

2. take a dissipative feedback control u_2 for (L_2, F_2, W_2)

3. check the asymptotic stability of $(q_e, 0)$ in the closed-loop system (L_2, F_2, u_2) using its energy E_2 as Lyapunov function

4. if $(q_e, 0)$ is asymptotically stable in the closed-loop dynamics of (L_2, F_2, u_2), then it is also asymptotically stable in the closed-loop system $(L_1, 0, u_1)$ with the control u_1 derived from (2.12).

In practice, item 1 is subdivided into:

1a. find a parameterized family of CL systems (L_2, F_2, W_2), with some set of free parameters, which are CL-equivalent to $(L_1, F_1 = 0, W_1)$

1b. choose a set of appropriate parameters in order for the energy E_2 to have a strict minimum at the equilibrium $(q_e, 0)$ and in order for the force F_2 to be of gyroscopic force form.

Notice that we had a free parameter δ in (2.28) to shape the energy E_2.

We now recall the form of dissipative forces F_{diss} and gyroscopic forces F_{gr} in coordinates:

$$F_{\text{diss}} = -D(q, \dot{q})\dot{q}, \qquad F_{\text{gr}} = S(q, \dot{q})\dot{q}$$

with $D(q, \dot{q}) = D^T(q, \dot{q}) \geq 0$ and $S(q, \dot{q}) = -S^T(q, \dot{q})$.

Remark 2.2.1. *1. When the situation is such that the energy has a maximum at the equilibrium, then one needs to use the energy pumping force F_{ep} instead of the dissipation, where the energy pumping force F_{ep} is of the form:*

$$F_{\text{ep}}(q, \dot{q}) = D(q, \dot{q})\dot{q}, \qquad D^T = D \geq 0.$$

2. One usually takes the gyroscopic force to be quadratic in the velocity $\dot{q}$ for simple CL systems because in the second Euler-Lagrange matching condition **ELM-2**, *all the other terms with $\dot{q}$ in equations of motion are quadratic in $\dot{q}$. In such a case, each element of $S(q, \dot{q})$ should be linear in $\dot{q}$.*

3. When a given CL system (L_1, F_1, W_1) has a non-zero external force $F_1 \neq 0$, one needs to modify the above procedure because this additional force has some effects on the energy.

4. There is a so-called λ-method which systematically solves PDE's involved in application of the CL method. In this thesis, we extend the λ-method to include gyroscopic forces. See § 2.4.2.

2.2.2 Example: Inverted Pendulum on a Cart

Control Synthesis. We now apply the theory to the inverted pendulum on a cart (Figure 2.1). This was solved in Bloch, Chang, Leonard, and Marsden [2001] and Bloch, Leonard, and Marsden [1999b].

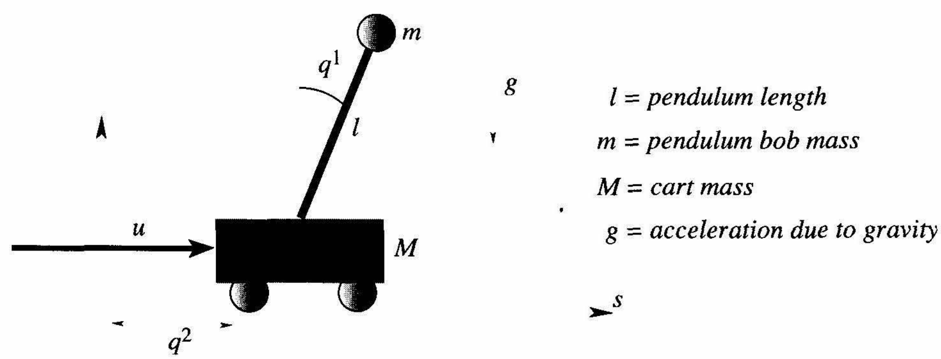

Figure 2.1: Pendulum on a Cart.

The configuration space is $Q = S^1 \times \mathbb{R}$. We shall use $q = (q^1, q^2)$ for the pendulum angle and the displacement of the cart. The control goal is to design a control to asymptotically stabilize the equilibrium at $(0,0,0,0) \in TQ$. Let $(L_1, F_1 = 0, W_1)$ be the inverted pendulum system with

$$L_1(q,\dot{q}) = \frac{1}{2}\alpha(\dot{q}^1)^2 + \beta\cos(q^1)\dot{q}^1\dot{q}^2 + \frac{1}{2}\gamma(\dot{q}^2)^2 - K\cos q^1$$

and

$$W_1 = \langle \mathbf{d}q^2 \rangle,$$

where $\alpha, \beta, \gamma, K > 0$ ($\alpha = ml^2, \beta = ml, \gamma = M + m$ and $K = mgl$ in terms of the data in Figure 2.1). The total energy E_1 is given by

$$E_1 = \frac{1}{2}\alpha(\dot{q}^1)^2 + \beta\cos(q_1)\dot{q}^1\dot{q}^2 + \frac{1}{2}\gamma(\dot{q}^2)^2 + K\cos q_1.$$

The equilibrium is a *saddle point* of the energy E_1. In this case, the potential energy shaping alone is not enough because E_1 has a maximum along the q^1 direction and the q^1 direction is the *unactuated* direction. Hence, we will perform both kinetic and potential energy shaping.

Let (L_2, F_2, W_2) be a candidate CL system equivalent to $(L_1, 0, W_1)$ with Lagrangian

$$L_2(q,\dot{q}) = \frac{1}{2}m_2(\dot{q},\dot{q}) - V_2(q),$$

where

$$m_2(q) = \begin{bmatrix} m_{11} & m_{12} \\ m_{12} & m_{22} \end{bmatrix}$$

and the gyroscopic force

$$F_2(q,\dot{q}) = \begin{bmatrix} 0 & -s_1(q)\dot{q}^1 - s_2(q)\dot{q}^2 \\ s_1(q)\dot{q}^1 + s_2(q)\dot{q}^2 & 0 \end{bmatrix} \begin{bmatrix} \dot{q}_1 \\ \dot{q}_2 \end{bmatrix}$$

where s_1, s_2 are functions to be chosen. Notice that the force decomposition in (2.13) reads in this case as follows:

$$F_2^v = F_2, \qquad F_2^q = 0.$$

It is reasonable to make the force F_2 quadratic in $\dot{q}$ because every term containing $\dot{q}$ in $\mathcal{EL}(L_1)$ and $\mathcal{EL}(L_2)$ is quadratic in $\dot{q}$. Notice $W_1^\circ = \langle [1\ 0] \rangle$.

The second Euler-Lagrange matching condition, **ELM-2**, will give a system of PDE's for m_2, V_2. Namely, equation (2.14) (or, (2.16)) becomes

$$0 = A(q)\frac{\partial m_{11}}{\partial q^1} + B(q)\left(2\frac{\partial m_{12}}{\partial q^1} - \frac{\partial m_{11}}{\partial q^2} - 2s_1\right) \tag{2.29}$$

$$0 = A(q)\left(\frac{\partial m_{11}}{\partial q^2} + s_1\right) + B(q)\left(\frac{\partial m_{22}}{\partial q^1} - s_2\right) \tag{2.30}$$

$$0 = A(q)\left(2\frac{\partial m_{12}}{\partial q^2} - \frac{\partial m_{22}}{\partial q^1} + 2s_2\right) + B(q)\frac{\partial m_{22}}{\partial q^2} \tag{2.31}$$

with

$$A(q) = (\alpha m_{22}(q) - \beta\cos(q^1)m_{12}(q)), \; B(q) = (\beta\cos(q^1)m_{11}(q) - \alpha m_{12}(q)).$$

The equation (2.15) (or, (2.17)) becomes

$$A(q)\frac{\partial V_2}{\partial q^1} + B(q)\frac{\partial V_2}{\partial q^2} = -(m_{11}m_{22} - (m_{12})^2)K\sin(q^1). \qquad (2.32)$$

We first solve (2.29)–(2.31) for $m_{11}, m_{12}, m_{22}, s_1, s_2$, and then solve (2.32) for V_2. Then, W_2 is determined by the first Euler-Lagrange matching condition, **ELM-1** as follows:

$$W_2 = m_2 m_1^{-1} W_1.$$

We do not have to try to find a general solution to (2.29)–(2.32). All we need is *a* particular solution enough to help design a stabilizing controller. Hence, we just set $s_1 = s_2 = 0$ and see if we can get a stabilizing controller (see § 2.4.3 for an example of the use of non-zero gyroscopic forces). One can try to directly solve PDE's in (2.29)–(2.31). However, it is sometimes easier to make some assumptions to reduce the PDE's to a set of ODE's as follows. We assume that m_{11} depends on q^1 only and m_{12} and m_{22} are of the following form just as in the original system:

$$m_{12}(q^1) = b\cos(q^1), \qquad m_{22} = d,$$

with $b, d \in \mathbb{R}$. Then, (2.30) and (2.31) are automatically satisfied and (2.29) becomes

$$(\alpha d - \beta b\cos^2(q^1))m_{11}'(q^1) - (\beta m_{11}(q^1) - \alpha b)2b\cos(q^1)\sin(q^1) = 0$$

which can be solved for $m_{11}(q^1)$:

$$m_{11}(q^1) = \alpha(ad + b/\beta) - ab\beta\cos^2(q^1)$$

with $a \in \mathbb{R}$. With this solution[2], equation (2.32) is simplified to

$$\frac{\partial V_2}{\partial q^1} + a\beta \cos(q^1) \frac{\partial V_2}{\partial q^2} = -K(ad + b/\beta)\sin(q^1)$$

which can be solved for V_2:

$$V_2(q^1, q^2) = K(ad + b/\beta)\cos(q^1) + V_\epsilon(q^2 - a\beta \sin(q^1)),$$

where $V_\epsilon(\cdot)$ is an arbitrary function of $\mathbb{R}$. The control subbundle W_2 is

$$W_2 = \langle (\alpha d - \beta b \cos^2(q^1))(\mathbf{d}q^2 - a\beta \cos(q^1)\mathbf{d}q^1) \rangle \tag{2.33}$$

For simplicity, let us choose a quadratic function for V_ϵ such that V_2 becomes

$$V_2(q^1, q^2) = K(ad + b/\beta)\cos(q^1) + \frac{1}{2}\epsilon(q^2 - a\beta \sin(q^1))^2$$

with $\epsilon \in \mathbb{R}$. The total energy E_2 of the CL system (L_2, F_2, W_2) is given by

$$\begin{aligned} E_2 = {} & \frac{1}{2}(\alpha(ad + b/\beta) - ab\beta \cos^2(q^1))(\dot{q}^1)^2 + b\cos(q^1)\dot{q}^1\dot{q}^2 + \frac{1}{2}d(\dot{q}^2)^2 \\ & + K(ad + b/\beta)\cos(q^1) + \frac{1}{2}\epsilon(q^2 - a\beta \sin(q^1))^2. \end{aligned} \tag{2.34}$$

One can check that the energy E_2 has a minimum at $(0, 0, 0, 0)$ in the set

$$R = \{(q^1, q^2, \dot{q}^1, \dot{q}^2) \mid \alpha d < \beta b \cos^2(q^1)\} \tag{2.35}$$

if the following holds:

$$d > 0, \quad \epsilon > 0, \quad a < 0, \quad \alpha d/\beta < b < -ad\beta. \tag{2.36}$$

---

[2]Here, our notations are different from those in Bloch, Chang, Leonard, and Marsden [2001]. One can recover the result in Bloch, Chang, Leonard, and Marsden [2001] by setting

$$a = -\left(1 - \frac{1}{\sigma} - \frac{1}{\rho}\right)\frac{1}{\gamma}; \quad b = \left(1 - \frac{1}{\sigma}\right)\rho\beta; \quad d = \rho\gamma.$$

Notice that the condition $\alpha d/\beta < b$ guarantees the constant nonzero rank of W_2 in (2.33). One can achieve asymptotic stabilization of the equilibrium by choosing a dissipative feedback control $u_2 \in W_2$ as follows:

$$u_2 = -c(\mathbf{d}q^2 - a\beta\cos(q^1)\mathbf{d}q^1) \otimes (\mathbf{d}q^2 - a\beta\cos(q^1)\mathbf{d}q^1)$$

with $c > 0$ constant. Namely,

$$u_2(q^1, q^2, \dot{q}^1, \dot{q}^2) = -c(\dot{q}^2 - a\beta\cos(q^1)\dot{q}^1)\begin{bmatrix} -a\beta\cos(q^1) \\ 1 \end{bmatrix}.$$

The asymptotic stabilization can be straightforwardly proved by the application of LaSalle's theorem (see Khalil [1996] for LaSalle's theorem). We postpone the proof of asymptotic stabilization. By Proposition 2.1.5, the closed-loop system $(L_1, 0, u_1)$ with u_1 obeying (2.12) has the origin as asymptotically stable equilibrium in the region R. The control u_1 is given by

$$\begin{aligned} &u_1(q^1, q^2, \dot{q}^1, \dot{q}^2) \\ &= \mathbf{d}q^2 \frac{1}{d\alpha - b\beta\cos^2(q^1)}\left[(b\gamma - d\beta)(\alpha(\dot{q}^1)^2 - K\cos(q^1))\sin(q^1)\right. \\ &\qquad \left. + (\alpha\gamma - \beta^2\cos^2(q^1))(c(a\beta\cos(q^1)\dot{q}^1 - \dot{q}^2) + \epsilon(a\beta\sin(q^1) - q^2))\right]. \end{aligned}$$

Remark 2.2.2. *Notice that by choosing appropriate values of a, b, d one can make the set R in (2.35) as close to $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right) \times \mathbb{R}^3$ as possible. Hence, our method provides a large region of attraction, which is a subset of R.*

We now show asymptotic stability of the equilibrium $(0, 0, 0, 0)$ in the closed-loop system. Let us use the following coordinates:

$$x = (x^1, x^2) := (q^1, q^2 - a\beta\sin(q^1)).$$

In these coordinates, the CL system $(L_2, 0, W_2)$ is given by

$$L_2(x,\dot{x}) = \frac{1}{2}(ad + b/\beta)(\alpha + a\beta^2\cos^2(x^1))(\dot{x}^1)^2 + (ad + b/\beta)\beta\cos(x^1)\dot{x}^1\dot{x}^2 + d(\dot{x}^2)^2$$
$$- K(ad + b/\beta)\cos(x^1) - \frac{1}{2}\epsilon(x^2)^2,$$

and $W_2 = \langle(\alpha d - \beta b\cos^2(x^1))\mathbf{d}x^2\rangle$. The feedback control u_2 in the new coordinates is given by

$$u_2(x,\dot{x}) = -c\dot{x}^2\mathbf{d}x^2.$$

Then,

$$\frac{dE_2}{dt}(x,\dot{x}) = -c(\dot{x}^2)^2 \leq 0.$$

Since E_2 has a strict local minimum at the origin, there exists $l \in \mathbb{R}$ such that the set $\Omega_l = \{(x,\dot{x}) = (x^1, x^2, \dot{x}^1, \dot{x}^2) \in (-\pi/2, \pi/2) \times \mathbb{R}^3 \mid E_2 \leq l\}$ is nonempty, compact and positively invariant. Let $\mathcal{M}$ be the largest invariant subset of the set $\{(x,\dot{x}) \mid \dot{E} = 0\} = \{(x,\dot{x}) \mid \dot{x}^2 = 0\}$. Let $(x(t),\dot{x}(t)) = (x^1(t), x^2(t), \dot{x}^1(t), \dot{x}^2(t))$ be an arbitrary trajectory lying in $\mathcal{M}$ for all $t \geq 0$. It follows from the definition of $\mathcal{M}$ that $x^2(t) = x^2(0)$ and $\dot{x}^2(t) = 0$. The equation of motion of the closed-loop systems $(L_2, 0, u_2)$ in the x^2-variable implies

$$(ad\beta + b)\frac{d}{dt}(\cos(x^1)\dot{x}^1) = -\epsilon x^2(0).$$

This implies

$$\sin(x^1(t)) = -\frac{\epsilon}{2(ad\beta + b)}x^2(0)t^2 + c_1 t + c_2 \tag{2.37}$$

for some $c_1, c_2 \in \mathbb{R}$. Since $x^1(t) \in (-\pi/2,\ \pi/2)$, it follows that $x^2(0) = 0$, $c_1 = 0$. The differentiation of (2.37) yields

$$\cos(x^1(t))\dot{x}^1(t) = 0.$$

Since $x^1(t) \in (-\pi/2,\ \pi/2)$, we have $\dot{x}^1(t) = 0$. So far, we have shown that

$$(x^1(t), x^2(t), \dot{x}^1(t), \dot{x}^2(t)) = (x^1(0), 0, 0, 0).$$

The equation of motion of the closed-loop systems $(L_2, 0, u_2)$ in the x^1-variable implies $\sin(x^1(0)) = 0$. It follows that $x^1(0) = 0$ because $x^1(0) \in (-\pi/2, \pi/2)$. Hence, the trajectory lying in $\mathcal{M}$ is the origin only. Therefore, the origin is asymptotically stable by LaSalle's theorem.

Simulations. We show a MATLAB simulation using the control law designed above. Here $m = 0.14$ kg, $M = 0.44$ kg, $l = 0.215$ m. Our goal is to regulate the cart at $q^1 = 0$ and the pendulum at $q^2 = 0$. We choose control gains to be $a = -1.2241 \times 10^2$, $b = 1.2642 \times 10^{-2}$, $d = 1.16 \times 10^{-2}$, $\epsilon = 1.0952 \times 10^{-3}$ and $c = 8.7 \times 10^{-3}$. Figure 2.2 shows plots of pendulum angle and velocity and cart position and velocity for the system subject to our asymptotically stabilizing controller. The pendulum starts from $(q^1(0), q^2(0), \dot{q}^1(0), \dot{q}^2(0)) = (30°, 3, 0, 0)$. Note that the cart comes to rest at the origin with the pendulum in the vertical position.

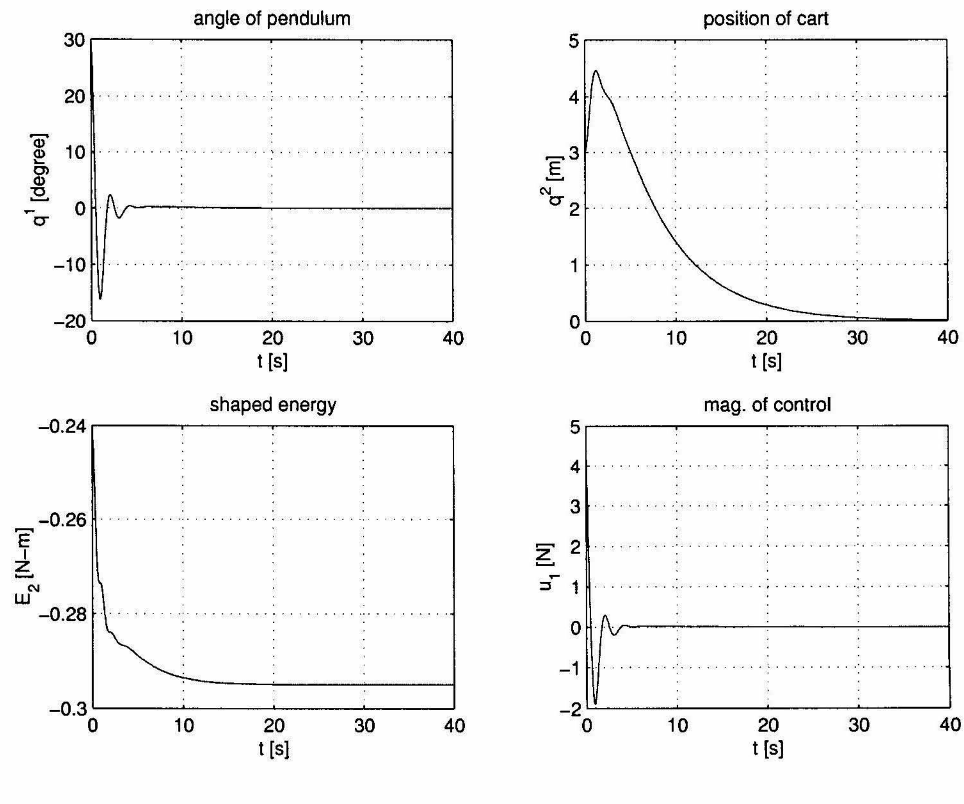

Figure 2.2: Simulation of the controlled pendulum on a cart.

At the bottom of Figure 2.2 we have included a plot of the control law u_1 and the Lyapunov function, i.e., the energy E_2 in (2.34). To keep the pendulum from falling past $90°$, a large initial force is needed. But as the response reaches its steady state, the control law converges to 0 N. The energy E_2 converges to its minimum value at the equilibrium.

Our next simulation is to see our controller enjoying a large region of attraction. Choose the parameters as follows: $a = -6.0517\times 10^2$, $b = 1.8120\times 10^{-1}$, $d = 1.1600\times 10^{-2}$, $\epsilon = 1.0952\times 10^{-3}$ and $c = 8.7\times 10^{-3}$. Figure 2.3 shows the responses for the three different initial conditions. Each row of plots corresponds to a different case. They all converge to the origin demonstrating a large region of attraction for the initial angle of the pendulum. Although we did not plot the force here, we note that we needed a large initial force in the third case. This explains that the large initial translational motion is unavoidable.

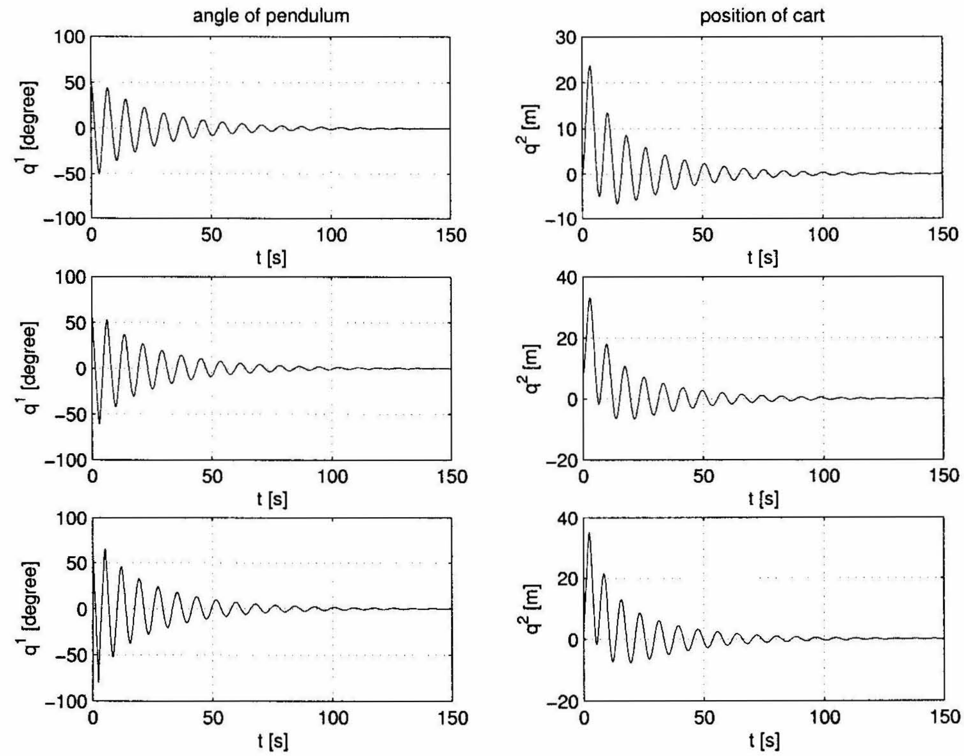

Figure 2.3: Responses to various initial conditions: (a) $z(0) = (51.57°, 0, 0, 0)$, (b) $z(0) = (60°, 8, 0, 0)$, (c) $z(0) = (80°, 5, 0, 0)$.

2.3 Simplified Matching Conditions

We find a class of mechanical systems for which asymptotically stabilizing controllers can be designed using the method of CL systems (see Theorem 2.3.2). This class will include examples such as the inverted pendulum on a cart and the inverted spherical pendulum on a cart. For the same class of systems, we design tracking controllers for constant acceleration reference trajectories. This work was published in Bloch, Chang, Leonard, and Marsden [2001].

2.3.1 Assumptions and Matching Conditions

Assumptions. Let $Q = \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}, (n_1 \leq n_2)$ be the configuration space and $(x, \theta) = (x^\alpha, \theta^a)$ be coordinates for Q with $\alpha = 1, \dots, n_1$, $a = 1, \dots, n_2$[3]. Consider a CL system $(L, 0, W)$ with

$$L = \frac{1}{2} g_{\alpha\beta} \dot{x}^\alpha \dot{x}^\beta + g_{\alpha a} \dot{x}^\alpha \dot{\theta}^a + \frac{1}{2} g_{ab} \dot{\theta}^a \dot{\theta}^b - V(x, \theta) \tag{2.38}$$

and

$$\begin{aligned} W &= \langle \mathbf{d}\theta^a \mid a = 1, \dots, n_2 \rangle \\ &= \text{the subbundle of } T^*Q \text{ spanned by the columns of } \begin{bmatrix} O \\ I_{n_2} \end{bmatrix}, \end{aligned}$$

where I_{n_2} is the $n_2 \times n_2$ identity matrix.

Let $z_e = ((x_e, \theta_e), (0, 0)) \in TQ$ be the equilibrium of interest. We make the following assumptions on the Lagrangian L in (2.38):

A1. $g_{\alpha\beta}$ and $g_{\alpha a}$ depend on x only, and g_{ab} is constant.

A2. $\frac{\partial g_{\alpha a}}{\partial x^\delta} = \frac{\partial g_{\delta a}}{\partial x^\alpha}$.

A3. $g_{a\alpha}(x_e)$ is a 1-1 matrix.

A4. The potential V is of the form $V(x, \theta) = U(x) + \tilde{U}(\theta)$.

A5. $U(x)$ has a local maximum at x_e, i.e.,

$$\mathbf{d}U(x_e) = 0, \qquad \frac{\partial^2 U}{\partial x^\alpha \partial x^\beta}(x_e) < 0.$$

[3]Greek letter indices such as α, β... run from $1, \dots, n_1$ and Latin letter indices such as $a, b, \dots$ run from 1 to n_2.

These are called simplified matching conditions in Bloch, Chang, Leonard, and Marsden [2001]. By **A4**, the energy E is given by

$$E = \frac{1}{2} g_{\alpha\beta}\dot{x}^\alpha \dot{x}^\beta + g_{\alpha a}\dot{x}^\alpha \dot{\theta}^a + \frac{1}{2} g_{ab}\dot{\theta}^a\dot{\theta}^b + U(x) + \tilde{U}(\theta).$$

The equilibrium z_e is neither a minimum point nor a maximum point of E since U has a maximum at x_e by **A4** and the kinetic energy has a minimum at the equilibrium. This cannot be remedied by the potential energy shaping technique through feedback because the control bundle W is generated by $\mathbf{d}\theta^a$, i.e., the actuation is along the θ direction. In this case, one has to perform both kinetic and potential energy shaping.

We consider a second CL system $(L_{\tau,\sigma,\rho,\epsilon}, 0, W_{\tau,\sigma,\rho,\epsilon})$[4] given by

$$\begin{aligned} L_{\tau,\sigma,\rho,\epsilon} = \frac{1}{2}\left(g_{\alpha\beta} + \rho\left(1 - \frac{1}{\sigma}\right)\left(1 - \frac{1}{\sigma} - \frac{1}{\rho}\right) g_{\alpha a} g^{ab} g_{b\beta}\right)\dot{x}^\alpha\dot{x}^\beta \\ + \rho\left(1 - \frac{1}{\sigma}\right) g_{\alpha a}\dot{x}^\alpha\dot{\theta}^a + \frac{1}{2}\rho g_{ab}\dot{\theta}^a\dot{\theta}^b - V(x,\theta) - V_\epsilon(x,\theta), \end{aligned} \tag{2.39}$$

and

$$\begin{aligned} W_{\tau,\sigma,\rho,\epsilon} &= \left\langle \mathbf{d}\theta^a + \left(1 - \frac{1}{\sigma} - \frac{1}{\rho}\right) g^{ac} g_{c\alpha}\mathbf{d}x^\alpha \mid a = 1, \ldots, n_2 \right\rangle \\ &= \text{the subbundle spanned by the columns of } \begin{bmatrix} \left(1 - \frac{1}{\sigma} - \frac{1}{\rho}\right) g_{\alpha d} g^{da} \\ I_{n_2} \end{bmatrix}, \end{aligned}$$

where $\sigma, \rho \in \mathbb{R}$ are free parameters. See Bloch, Chang, Leonard, and Marsden [2001] for the motivation of this form of a CL system.

Matching Conditions. We now examine the Euler-Lagrange matching conditions for the two CL systems. To follow the notation in Definition 2.1.3, let

$$(L_1, 0, W_1) = (L_{\tau,\sigma,\rho,\epsilon}, 0, W_{\tau,\sigma,\rho,\epsilon}), \quad (L_2, 0, W_2) = (L, 0, W).$$

[4]We keep the subscript $\tau, \sigma, \rho, \epsilon$ following Bloch, Chang, Leonard, and Marsden [2001].

First, we examine the condition **ELM-1**. One computes

$$m_2 m_1^{-1} \begin{bmatrix} O \\ I_{n_2} \end{bmatrix} = \begin{bmatrix} \left(1 - \frac{1}{\sigma} - \frac{1}{\rho}\right) g_{\alpha d} g^{da} \\ I_{n_2} \end{bmatrix} \left[g_{ab} - \left(1 - \frac{1}{\sigma}\right) g_{a\alpha} g^{\alpha\beta} g_{\beta b} \right] \left[\rho G^{be} \right] \tag{2.40}$$

where G^{ab} is the inverse matrix of $G_{ab} := g_{ab} - g_{a\alpha} g^{\alpha\beta} g_{\beta b}$. Notice the matrix G_{ab} is always invertible because the kinetic energy part in (2.38) is nondegenerate. For **ELM-1** to hold, the following should hold:

$$\rho \neq 0, \qquad \det \left[g_{ab} - \left(1 - \tfrac{1}{\sigma}\right) g_{a\alpha} g^{\alpha\beta} g_{\beta b} \right] \neq 0. \tag{2.41}$$

Later, we will see that (2.41) is implied by the stability condition in Claim 2.3.1 in § 2.3.2.

Secondly, we examine the condition **ELM-2**. The annihilator $W_1^\circ = \{v \in TQ | \langle v, \alpha \rangle = 0 \text{ for all } \alpha \in W_1\}$ of W_1 is spanned by the row vectors of the matrix

$$K := \left[I_{n_2}, \quad -\left(1 - \tfrac{1}{\sigma} - \tfrac{1}{\rho}\right) g_{\alpha d} g^{da} \right].$$

The condition **ELM-2** can be written as

$$K \left[\mathcal{EL}(L_1) - m_1 m_2^{-1} \mathcal{EL}(L_2) \right] = O. \tag{2.42}$$

By **A1** – **A3**, equation (2.42) holds if and only if

$$-\left(\frac{\partial V}{\partial \theta^a} + \frac{\partial V_\epsilon}{\partial \theta^a} \right) \left(1 - \frac{1}{\sigma} - \frac{1}{\rho} \right) g^{ad} g_{d\alpha} + \frac{\partial V_\epsilon}{\partial x^\alpha} = 0.$$

By **A4**, one can solve this PDE for V_ϵ as follows:

$$V_\epsilon(x, \theta) = -\tilde{U}(\theta) + \tilde{V}_\epsilon(\theta + h(x)), \tag{2.43}$$

where $\tilde{V}_\epsilon$ is an *arbitrary* function on $\mathbb{R}^{n_2}$ and the $\mathbb{R}^{n_2}$-valued function $h = (h^a)$ on $\mathbb{R}^{n_1}$ is defined by

$$\frac{\partial h^a}{\partial x^\alpha}(x) = \left(1 - \frac{1}{\sigma} - \frac{1}{\rho} \right) g^{ac} g_{c\alpha}(x), \quad \text{with} \quad h^a(x_e) = 0. \tag{2.44}$$

The existence of h is guaranteed by **A1**, **A2**, and the Poincaré Lemma. By (2.43), the

Lagrangian $L_{\tau,\sigma,\rho,\epsilon}$ in (2.39) can be written as

$$L_{\tau,\sigma,\rho,\epsilon} = \frac{1}{2}\left(g_{\alpha\beta} + \rho\left(1 - \frac{1}{\sigma}\right)\left(1 - \frac{1}{\sigma} - \frac{1}{\rho}\right) g_{\alpha a} g^{ab} g_{b\beta}\right)\dot{x}^\alpha \dot{x}^\beta$$
$$+ \rho\left(1 - \frac{1}{\sigma}\right) g_{\alpha a}\dot{x}^\alpha\dot{\theta}^a + \frac{1}{2}\rho g_{ab}\dot{\theta}^a\dot{\theta}^b - U(x) - \tilde{V}_\epsilon(\theta + h(x)). \tag{2.45}$$

So far, we have shown that

$$(L, 0, W) \overset{L}{\sim} (L_{\tau,\sigma,\rho,\epsilon}, 0, W_{\tau,\sigma,\rho,\epsilon}). \tag{2.46}$$

2.3.2 Asymptotic Stabilization

We will design a control which asymptotically stabilizes the equilibrium

$$z_e = ((x_e, \theta_e), (0, 0)) \in TQ$$

for the original CL system $(L, 0, W)$. By (2.46), we can equivalently work with the CL system $(L_{\tau,\sigma,\rho,\epsilon}, 0, W_{\tau,\sigma,\rho,\epsilon})$.

We will find a new set of coordinates using the integrability of the control bundle $W_{\tau,\sigma,\rho,\epsilon}$. Consider the following change of coordinates:

$$(x^\alpha, \theta^a) \mapsto (x^\alpha, y^a) := (x^\alpha, \theta^a + h^a(x)),$$

where h^a is defined in (2.44). Denote by $z_e = ((x_e, y_e), (0, 0))$ the equilibrium in the new coordinates. In the following, we will exclusively use the coordinates (x, y) for Q. In (x, y), the CL system $(L_{\tau,\sigma,\rho,\epsilon}, 0, W_{\tau,\sigma,\rho,\epsilon})$ is written as

$$L_{\tau,\sigma,\rho,\epsilon} = \frac{1}{2}\left(g_{\alpha\beta} - \left(1 - \frac{1}{\sigma} - \frac{1}{\rho}\right) g_{\alpha a} g^{ab} g_{b\beta}\right)\dot{x}^\alpha\dot{x}^\beta + g_{\alpha a}\dot{x}^\alpha\dot{y}^a + \frac{1}{2}\rho g_{ab}\dot{y}^a\dot{y}^b$$
$$- U(x) - \tilde{V}_\epsilon(y) \tag{2.47}$$

and

$$W_{\tau,\sigma,\rho,\epsilon} = \langle \mathbf{d}y^a \mid a = 1, \ldots, n_2\rangle.$$

We use, as a Lyapunov function candidate, the energy $E_{\tau,\sigma,\rho,\epsilon}$ of $L_{\tau,\sigma,\rho,\epsilon}$ given by

$$
\begin{aligned}
E_{\tau,\sigma,\rho,\epsilon} &= \frac{1}{2}\left(g_{\alpha\beta} - \left(1 - \frac{1}{\sigma} - \frac{1}{\rho}\right) g_{\alpha a} g^{ab} g_{b\beta}\right) \dot{x}^\alpha \dot{x}^\beta + g_{\alpha a}\dot{x}^\alpha \dot{y}^a + \frac{1}{2}\rho g_{ab}\dot{y}^a\dot{y}^b \\
&\quad + U(x) + \tilde{V}_\epsilon(y) \\
&= K(x, y, \dot{x}, \dot{y}) + U(x) + \tilde{V}_\epsilon(y),
\end{aligned} \tag{2.48}
$$

where K is the kinetic energy part of $E_{\tau,\sigma,\rho,\epsilon}$. Recall that $\tilde{V}_\epsilon$ is still arbitrary. Choose $\tilde{V}_\epsilon$ such that it has a maximum at y_e, i.e.,

$$
\mathbf{d}\tilde{V}_\epsilon(y_e) = 0, \quad \frac{\partial^2 \tilde{V}_\epsilon}{\partial y^a \partial y^b}(y_e) < 0. \tag{2.49}
$$

Claim 2.3.1. *The kinetic energy K is negative definite locally around z_e if $\rho < 0$ and*

$$
1 - \frac{1}{\sigma} > \max\{\, \lambda \mid \det\left(g_{\alpha\beta}(x_e) - \lambda g_{\alpha a}(x_e) g^{ab} g_{b\beta}(x_e)\right) = 0\,\}. \tag{2.50}
$$

Proof. Let A be the following matrix:

$$
A := \begin{bmatrix} I_{n_1} & 0 \\ -\frac{1}{\rho} g^{ab} g_{b\alpha} & I_{n_2} \end{bmatrix}.
$$

Then,

$$
A^T \begin{bmatrix} g_{\alpha\beta} - \left(1 - \frac{1}{\sigma} - \frac{1}{\rho}\right) g_{\alpha a} g^{ab} g_{b\beta} & g_{\alpha b} \\ g_{a\beta} & \rho g_{ab} \end{bmatrix} A = \begin{bmatrix} g_{\alpha\beta} - \left(1 - \frac{1}{\sigma}\right) g_{\alpha a} g^{ab} g_{b\beta} & 0 \\ 0 & \rho g_{ab} \end{bmatrix}.
$$

By **A3** and the positive definiteness of g_{ab}, the matrix $g_{\alpha a} g^{ab} g_{b\beta}$ is a $n_1 \times n_1$ *positive definite* matrix at x_e. Using the standard simultaneous diagonalization technique in linear algebra, one sees that the matrix $g_{\alpha\beta} - \left(1 - \frac{1}{\sigma}\right) g_{\alpha a} g^{ab} g_{b\beta}$ is negative definite at x_e if (2.50) holds. The matrix ρg_{ab} is negative definite if $\rho < 0$. The claim follows by continuity. ■

One can then check that the equilibrium is a (local) strict maximum point of E if $\rho < 0$ and σ satisfies (2.50) because of **A5**, (2.49), and Claim 2.3.1. We now choose an energy-pumping force v as a control for $(L_{\tau,\sigma,\rho,\epsilon}, 0, W_{\tau,\sigma,\rho,\epsilon})$ as follows:

$$v_a = c_{ab}\dot{y}^b,$$

where c_{ab} is a positive definite matrix. Then, z_e is an equilibrium of the closed-loop system $(L_{\tau,\sigma,\rho,\epsilon}, 0, v)$, and the time derivative of the energy $E_{\tau,\sigma,\rho,\epsilon}$ is given by

$$\frac{d}{dt}E_{\tau,\sigma,\rho,\epsilon} = c_{ab}\dot{y}^a\dot{y}^b \geq 0.$$

Thus, z_e becomes a Lyapunov stable equilibrium of the closed-loop system.

We now show that z_e is an *asymptotically stable* equilibrium in the closed-loop system. Since E has a maximum at z_e, there is $c \in \mathbb{R}$ such that the set

$$\Omega_c = \{z = (x, y, \dot{x}, \dot{y}) \in TQ \mid E_{\tau,\sigma,\rho,\epsilon}(z) \geq c\} \tag{2.51}$$

is nonempty, compact and positively invariant. By compactness and positive invariance, integral curves starting in Ω_c are defined and stay in Ω_c for all $t \geq 0$. Define

$$\begin{aligned}\mathcal{E} &= \left\{z \in \Omega_c \mid \frac{d}{dt}E_{\tau,\sigma,\rho,\epsilon}(z) = 0\right\} = \{z \in \Omega_c \mid \dot{y} = 0\},\\ \mathcal{M} &= \text{the largest invariant subset of } \mathcal{E}.\end{aligned}$$

Suppose a trajectory $z(t) = (x(t), y(t), \dot{x}(t), \dot{y}(t))$ is contained in $\mathcal{M}$ for all $t \geq 0$. Then, we have

$$y(t) = y(0) \;\; \forall\, t \geq 0. \tag{2.52}$$

The Euler-Lagrange equations for the y coordinates of $(L_{\tau,\sigma,\rho,\epsilon}, 0, v)$ are given by

$$\frac{d}{dt}(g_{a\alpha}\dot{x}^\alpha) + \rho g_{ab}\ddot{y}^b + \frac{\partial \tilde{V}_\epsilon}{\partial y^a} = c_{ab}\dot{y}^b.$$

By (2.52), this becomes

$$\frac{d}{dt}(g_{a\alpha}\dot{x}^\alpha) = -\frac{\partial \tilde{V}_\epsilon}{\partial y^a}(y(0)),$$

which implies

$$g_{a\alpha}(x(t))\dot{x}^{\alpha} = -\frac{\partial \tilde{V}_{\epsilon}}{\partial y^{a}}(y(0))t + B_{a}$$

for some constant $\mathbb{R}^{n_2}$ vector $B = (B_a)$. By compactness of Ω_c, $g_{a\alpha}(x(t))\dot{x}^{\alpha}(t)$ is bounded. It follows

$$\frac{\partial \tilde{V}_{\epsilon}}{\partial y^{a}}(y(0)) = 0.$$

By (2.49), y_e is the only critical point of $\tilde{V}_{\epsilon}$ in a neighborhood of y_e[5]. Hence,

$$y(t) = y(0) = y_e \;\; \forall\, t \geq 0. \tag{2.53}$$

We now have

$$g_{a\alpha}(x(t))\dot{x}^{\alpha}(t) = B_a. \tag{2.54}$$

By **A2**, there is an $\mathbb{R}^{n_2}$-valued function $l = (l_a)$ on $\mathbb{R}^{n_1}$ satisfying $\frac{\partial l_a}{\partial x^{\alpha}}(x) = g_{a\alpha}(x)$. The equation (2.54) becomes $\frac{d}{dt}l_a(x(t)) = B_a$. Integrating this yields $l_a(x(t)) = B_a t + C_a$ for some $C = (C_a) \in \mathbb{R}^{n_2}$. Again, since Ω_c is compact, $l_a(x(t))$ is bounded for all $t \geq 0$. It follows $B_a = 0$. Substitution of $B_a = 0$ into (2.54) gives

$$g_{a\alpha}(x(t))\dot{x}^{\alpha}(t) = 0.$$

Since $g_{a\alpha}$ is 1-1 in a neighborhood of x_e by **A3**, it follows that $\dot{x}^{\alpha}(t) = 0$ for all $t \geq 0$. So far, we have proved

$$z(t) = (x(t), y(t)\dot{x}(t), \dot{y}(t)) = (x(0), y_e, 0, 0). \tag{2.55}$$

By (2.55), the Euler-Lagrange equations for the x coordinates of the closed-loop system $(L_{\tau,\sigma,\rho,\epsilon}, 0, v)$ is simplified to

$$\frac{\partial U}{\partial x^{\alpha}}(x(0)) = 0.$$

By **A5**, x_e is the only critical point of U in a neighborhood of x_e. It follows $x(0) = x_e$. Therefore, we have proved that every trajectory in the set $\mathcal{M}$ is the equilibrium $z_e =$

[5]This neighborhood depends on our choice of the function $\tilde{V}_{\epsilon}$. For example, if we choose a quadratic function in $y - y_e$, then y_e is *the* critical point of $\tilde{V}_{\epsilon}$.

$(x_e, y_e, 0, 0)$ only. By LaSalle's theorem, the equilibrium z_e is asymptotically stable and a region of attraction is Ω_c[6]. Here, we summarize the result:

Theorem 2.3.2. *Let $Q = \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$ be the configuration space. Consider a class of CL systems $(L, 0, W)$ satisfying* **A1**–**A5** *and $W = 0 \times T^*\mathbb{R}^{n_2}$. Let x_e be the maximum point of U in* **A4**. *Then, for any $\theta_e \in \mathbb{R}^{n_2}$ there is a feedback control such that $((x_e, \theta_e), (0, 0)) \in TQ$ becomes an asymptotically stable equilibrium in the closed-loop system.*

Proof. The feedback control is designed by the method of CL systems. The procedure of the design is given in § 2.3.1 and § 2.3.2. ■

Remark 2.3.3. *Here, we made the energy $E_{\tau,\sigma,\rho,\epsilon}$ have a maximum at the equilibrium, not a minimum. This is not at all strange. Recall Remark 2.1.10.*

Example. We apply the above results to the spherical pendulum on a cart that travels on an incline of angle ψ. The system is shown in figure 2.4.

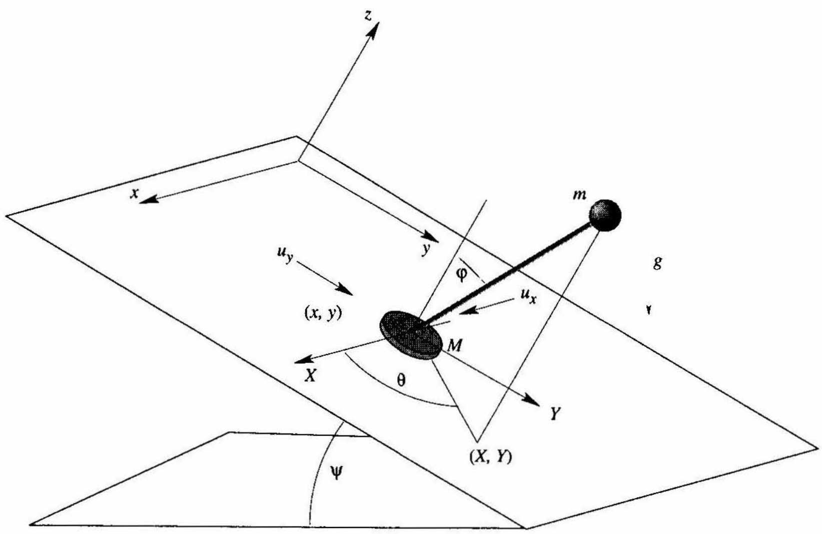

Figure 2.4: Spherical pendulum moving on an incline.

The configuration space for this system is $Q = S \times G = U \times \mathbb{R}^2$, where U is the open hemisphere above the incline which is diffeomorphic to an open subset of $\mathbb{R}^2$. We denote

[6] At this stage, Ω_c can be smaller than the Ω_c defined in (2.51) in the beginning of the proof of asymptotic stability because we might have had to shrink Ω_c in the proof.

by (x, y) the Cartesian coordinates of the cart on the incline and assume that we have independent controls that can move the cart in the x and y directions. Let P be the plane whose origin is attached to the cart and which is parallel to the incline. We will use the projection onto the plane P for a local chart for the open upper hemisphere. Let (X, Y) be the Cartesian coordinates of the bob in the plane P under this local chart. Thus, the velocity phase space TQ has local coordinates $z = (X, Y, x, y, \dot{X}, \dot{Y}, \dot{x}, \dot{y})$.

Let M and m be the masses of the cart and the bob, respectively, and r be the length of the pendulum. The position R of the bob in the inertial frame is given by

$$R = (x + X, y + Y, \sqrt{r^2 - X^2 - Y^2}).$$

The kinetic energy is the sum of the kinetic energies of the cart and the pendulum:

$$\begin{aligned} K(z) = & \frac{1}{2}m\left(\frac{r^2 - Y^2}{r^2 - X^2 - Y^2}\dot{X}^2 + \frac{2XY}{r^2 - X^2 - Y^2}\dot{X}\dot{Y} + \frac{r^2 - X^2}{r^2 - X^2 - Y^2}\dot{Y}^2\right) \\ & + m\left(\dot{x}\dot{X} + \dot{y}\dot{Y}\right) + \frac{1}{2}(m + M)(\dot{x}^2 + \dot{y}^2). \end{aligned} \tag{2.56}$$

The potential energy V is

$$V(X, Y, x, y) = U(X, Y) + \tilde{U}(x, y), \tag{2.57}$$

where

$$\begin{aligned} U(X, Y) &= mg\left(\cos\psi\sqrt{r^2 - X^2 - Y^2} - \sin\psi Y\right), \\ \tilde{U}(x, y) &= -(m + M)gy\sin\psi. \end{aligned}$$

Hence the CL system $(L, 0, W)$ is given by

$$L(z) = K(X, Y, x, y, \dot{X}, \dot{Y}, \dot{x}, \dot{y}) - V(X, Y, x, y).$$

and

$$W = \langle \mathbf{d}x, \mathbf{d}y \rangle.$$

The metric induced from the kinetic energy is

$$\begin{bmatrix} m\left(\frac{r^2-Y^2}{r^2-X^2-Y^2}\right) & m\left(\frac{XY}{r^2-X^2-Y^2}\right) & m & 0 \\ m\left(\frac{XY}{r^2-X^2-Y^2}\right) & m\left(\frac{r^2-X^2}{r^2-X^2-Y^2}\right) & 0 & m \\ m & 0 & m+M & 0 \\ 0 & m & 0 & m+M \end{bmatrix}.$$

It is easy to check that **A1**–**A3** are satisfied. The form of the potential in (2.57) satisfies **A4**. Physically, it is obvious that $U(X,Y)$ has a maximum at $(X,Y) = (0, -r\sin\psi)$ which is, as it should be, the position of the pendulum vertical to the ground, not to the incline, so **A5** is satisfied. The matrix

$$(g_{a\alpha}(0, -r\sin\psi)) = m\mathrm{I}_{2\times 2}$$

is clearly 1-1, so **A3** holds. By Theorem 2.3.2, the vertical position (*relative to the ground*) of the pendulum and any fixed position for the cart on the incline is asymptotically stabilizable.

2.3.3 Tracking

Here we consider a tracking problem for a CL system satisfying **A1**–**A5** in § 2.3.1. We want to let the θ^a variables track a constant acceleration curve in $\mathbb{R}^{n_2}$, while regulating the x^α variables at a fixed point x_e^α in $\mathbb{R}^{n_1}$. This is one of the simplest nontrivial tracking problems.

The configuration space is $Q = \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$ with $n_1 \le n_2$. Use $(x, \theta) = (x^\alpha, \theta^a)$ for coordinates on Q. Consider a CL system $(L, 0, W)$ with L satisfying **A1**–**A5** and $W = \langle \mathbf{d}\theta^a \mid a = 1, \ldots, n_2\rangle$. Let $r(t) \in \mathbb{R}^{n_2}$ be the reference signal satisfying

$$\ddot{r}(t) = c = \text{constant}. \tag{2.58}$$

Consider a moving frame which moves along $(0, r(t))$. Let (x, y) be the coordinates in the moving frame satisfying

$$y^a = \theta^a - r^a(t). \tag{2.59}$$

We will express the CL system $(L, 0, W)$ in the new coordinates (x, y). Let $L_m : TQ \times \mathbb{R} \to \mathbb{R}$ be the Lagrangian in the moving frame defined by

$$L_m(x, y, \dot{x}, \dot{y}, t) = L(x, y + r(t), \dot{x}, \dot{y} + \dot{r}(t)).$$

In coordinates,

$$\begin{aligned} L_m(x, y, \dot{x}, \dot{y}, t) = & \frac{1}{2} g_{\alpha\beta}\dot{x}^\alpha \dot{x}^\beta + g_{\alpha a}\dot{x}^\alpha \dot{y}^a + \frac{1}{2} g_{ab}\dot{y}^a \dot{y}^b + g_{\alpha a}\dot{x}^\alpha \dot{r}^a(t) + g_{ab}\dot{y}^a \dot{r}^b(t) \\ & + \frac{1}{2} g_{ab}\dot{r}^a(t)\dot{r}^b(t) - U(x) - \tilde{U}(y + r(t)). \end{aligned} \tag{2.60}$$

By **A2** and the Poincaré Lemma, there exists a function $l : U \subset \mathbb{R}^{n_1} \to \mathbb{R}^{n_2}$ such that

$$\frac{\partial l_a}{\partial x^\alpha}(x) = g_{\alpha a}(x).$$

Hence (2.60) can be written as

$$\begin{aligned} L_m(x, y, \dot{x}, \dot{y}, t) = & \frac{1}{2} g_{\alpha\beta}\dot{x}^\alpha \dot{x}^\beta + g_{\alpha a}\dot{x}^\alpha \dot{y}^a + \frac{1}{2} g_{ab}\dot{y}^a \dot{y}^b - l_a(x^\alpha)\ddot{r}^a(t) - g_{ab} y^a \ddot{r}^b(t) \\ & + \frac{d}{dt}(l_a(x^\alpha)\dot{r}^a(t)) + \frac{d}{dt}(g_{ab} y^a \dot{r}^b(t)) + \frac{1}{2} g_{ab}\dot{r}^a(t)\dot{r}^b(t) \\ & - U(x) - \tilde{U}(y + r(t)). \end{aligned} \tag{2.61}$$

Since exact time derivatives do not affect the variational principle, we can ignore the following three terms:

$$\frac{d}{dt}(l_a(x^\alpha)\dot{r}^a(t)), \qquad \frac{d}{dt}(g_{ab} y^a \dot{r}^b(t)), \qquad \frac{1}{2} g_{ab}\dot{r}^a(t)\dot{r}^b(t).$$

Hence the Lagrangian L_m in (2.61) can be replaced by the following Lagrangian:

$$\begin{aligned} L_m(x, y, \dot{x}, \dot{y}, t) = & \frac{1}{2} g_{\alpha\beta}\dot{x}^\alpha \dot{x}^\beta + g_{\alpha a}\dot{x}^\alpha \dot{y}^a + \frac{1}{2} g_{ab}\dot{y}^a \dot{y}^b - l_a(x^\alpha)c^a - g_{ab} y^a c^b \\ & - U(x) - \tilde{U}(y + r(t)), \end{aligned} \tag{2.62}$$

where (2.58) was used. The Euler-Lagrange equations in the moving frame are given by

$$\begin{aligned} \frac{d}{dt}\frac{\partial L_m}{\partial \dot{x}^\alpha} - \frac{\partial L_m}{\partial x^\alpha} &= 0 \\ \frac{d}{dt}\frac{\partial L_m}{\partial \dot{y}^a} - \frac{\partial L_m}{\partial y^a} &= v_a, \end{aligned} \tag{2.63}$$

where the input v in the moving frame has the following relationship with the input u in the fixed frame:

$$v(x, y, \dot{x}, \dot{y}, t) = u(x, y + r(t), \dot{x}, \dot{y} + \dot{r}(t), t). \tag{2.64}$$

In other words,

$$(L, 0, W) \overset{L}{\sim} (L_m, 0, W_m)$$

where $W_m = \langle \mathbf{d}y^a \mid a = 1, \ldots .n_2 \rangle$.

Our strategy is as follows. First, design a controller v for $(L_m, 0, W_m)$ which asymptotically stabilizes $((x_e, 0), (0, 0))$ in the moving frame for some $x_e \in \mathbb{R}^{n_1}$. Then, (2.64) and (2.59) will give a controller u for $(L, 0, W)$ which asymptotically tracks the reference signal $((x_e, r(t)), (0, \dot{r}(t))) \in TQ$. Consider the CL system $(\tilde{L}_m, 0, W_m)$ with

$$\tilde{L}_m(x, y, \dot{x}, \dot{y}) = \frac{1}{2} g_{\alpha\beta}\dot{x}^\alpha \dot{x}^\beta + g_{\alpha a}\dot{x}^\alpha \dot{y}^a + \frac{1}{2} g_{ab}\dot{y}^a \dot{y}^b - \tilde{V}(x), \tag{2.65}$$

where

$$\tilde{V}(x) = U(x) + l_a(x)c^a. \tag{2.66}$$

One can check

$$(L_m, 0, W_m) \overset{L}{\sim} (\tilde{L}_m, 0, W_m),$$

and the relation in (2.12) between the control w for $(\tilde{L}_m, 0, W_m)$ and the control v for $(L_m, 0, W_m)$ is given by

$$v_a = g_{ab}c^b + \frac{\partial}{\partial y^a}\tilde{U}(y^a + r^a(t)) + w_a. \tag{2.67}$$

In the above expression, we have performed potential energy shaping. Notice that $\tilde{L}_m$ is time-independent and its kinetic energy is of the form of L. One can check that $\tilde{L}_m$ satisfies **A1**–**A5**. Let x_e be a maximum of $\tilde{V}$. By Theorem 2.3.2, we can design a controller

w such that $((x_e, 0), (0, 0))$ becomes an asymptotically stable equilibrium in the moving frame. From w we can derive the input u for the original system $(L, 0, W)$ by (2.64) and (2.67). The asymptotic stabilization in the moving frame is equal to the tracking in the fixed frame. Thus, u becomes a tracking controller such that $(x(t), \phi(t), \dot{x}(t), \dot{\phi}(t))$ asymptotically converges to $(x_e, r(t), 0, \dot{r}(t))$.

Example. Consider again the inverted pendulum on a cart. In this case, $\tilde{V}$ is given by

$$\tilde{V}(\phi) = mgl\cos\phi + mlc\sin\phi,$$

where ϕ is the pendulum angle and c is the constant acceleration of the reference curve. $\tilde{V}$ has a maximum at $\phi_o = \arctan(c/g)$. This means that the cart will move with acceleration c with the pendulum slanted by angle ϕ_o, and it agrees with physical intuition.

Simulation. Next, we present tracking simulations for the system of an inverted pendulum on a cart. The description of the system is given in § 2.2.2. Our goal is to make the cart track a given curve of constant acceleration a with the pendulum slanted by $\phi_o := \arctan(a_o/g)$. Let $r(t) = \frac{1}{2}a_o t^2$ with $a_o = \frac{\pi}{6}g = 5.13\text{m/s}^2$ be the reference signal for the cart. Then $\phi_o = 0.4824(\text{rad}) = 27.6365°$. First, we choose the following control gains: $a = -1.3966 \times 10^2$, $b = 1.8662 \times 10^{-2}$, $d = 1.1600 \times 10^{-2}$, $\epsilon = 1.0953 \times 10^{-2}$, and $c = 8.7000 \times 10^{-3}$. Let e be the difference between the position of the cart and the reference signal. The first row and the second row of plots in Figure 2.5 are the responses with this controller with the initial conditions $(q^1(0), q^2(0), \dot{q}^1(0), \dot{q}^2(0)) = (0, -2, 0, 0)$ and $(q^1(0), q^2(0), \dot{q}^1(0), \dot{q}^2(0)) = (60°, 2, 0, 0)$, respectively. We can see that the angle of the pendulum converges to ϕ_o and the cart tracks the reference signal. However, this set of gains is not enough to handle a large initial angle difference. Hence we try another controller with $a = -6.0517 \times 10^2$, $b = 1.8120 \times 10^{-1}$, $d = 1.1600 \times 10^{-2}$, $\epsilon = 1.0953 \times 10^{-2}$ and $c = 8.7000 \times 10^{-3}$ which was found earlier in order to get a large region of attraction in the regulation problem. The third row in Figure 2.5 is the response with this controller with the initial condition $(q^1(0), q^2(0), \dot{q}^1(0), \dot{q}^2(0)) = (80°, 0, 0, 0)$. This controller achieves our objective.

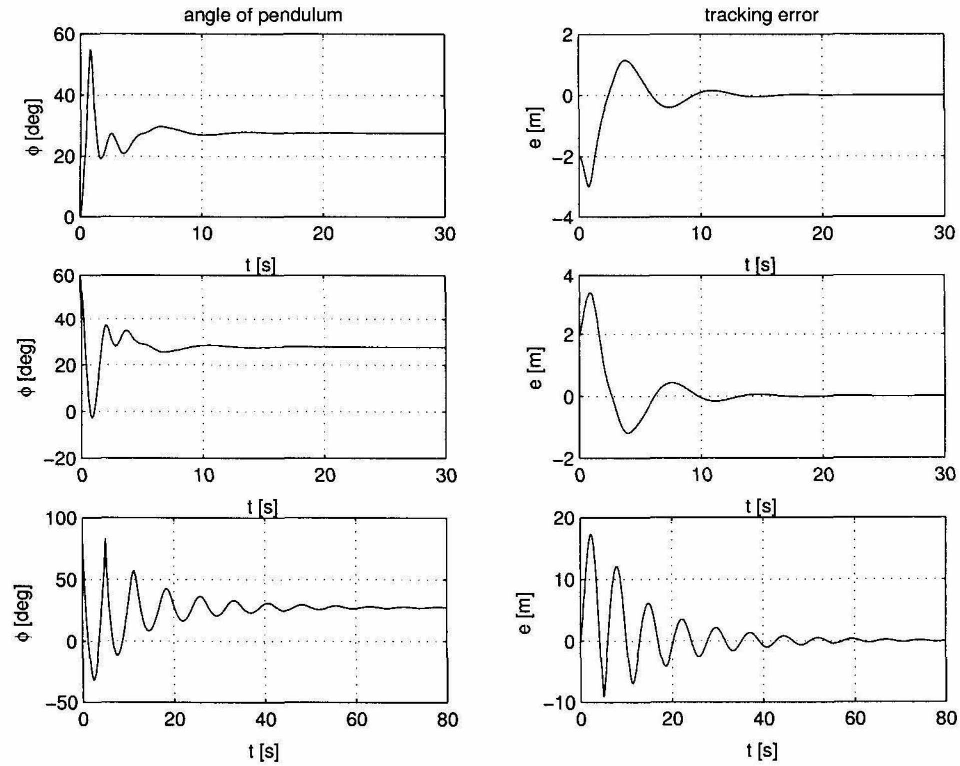

Figure 2.5: Tracking responses: (a) $z(0) = (0, -2, 0, 0)$, (b) $z(0) = (60°, 2, 0, 0)$, (c) $z(0) = (80°, 0, 0, 0)$.

2.4 The Extended λ-Method

Lagrangian mechanics for simple Lagrangians can be understood in terms of the Levi-Civita connection by regarding the mass tensor as a Riemannian metric (see Marsden and Ratiu [1999]). Hence, one can understand the method of controlled Lagrangian systems in the same manner. This approach was taken in Auckly, Kapitanski, and White [2000], Auckly and Kapitanski [2001], Hamberg [1999, 2000]. This has the advantage that the useful tool of tensor analysis can be used. However, it has the disadvantage that the reduction of CL systems with symmetry is difficult in this approach, where the reduction of CL systems with symmetry will be discussed in § 4.

In particular, Auckly, Kapitanski, and White [2000] and Auckly and Kapitanski [2001] developed so called λ-method in order to more efficiently solve the PDE involved in the Euler-Lagrange matching conditions, (2.14) and (2.15), or equivalently, (2.16) and (2.17).

However, they did not consider external forces, i.e., they only dealt with CL systems of the form $(L, 0, W)$. We extend their λ-method by considering the full form of CL systems (L, F, W) as defined in Definition 2.1.1. This extension allows us to use gyroscopic forces for asymptotic stabilization. This will prove useful in the design of a controller for the system of an inverted pendulum on a rotor arm in § 2.4.3. A Hamiltonian analog of the λ-method has not been developed. Refer to Gallot, Hulin, and Lafontaine [1993] for the Riemmanian geometry theory.

2.4.1 Extended λ-Method

We will extend the λ-method presented in Auckly, Kapitanski, and White [2000], by considering the full form of simple CL system (L, F, W) defined in Definition 2.1.1 and 2.1.2.

First, we review how to get equations of motion of a CL system (L, F, W) using the Levi-Civita connection. It is well known that the equation of motion of a CL system (L, F, W) with

$$L(q, \dot{q}) = \frac{1}{2}m(q)(\dot{q}, \dot{q}) - V(q)$$

can be written on TQ as follows:

$$\nabla_{\dot{q}}\dot{q} + m^{-1}\mathbf{d}V = m^{-1}F + m^{-1}u,$$

where $u : TQ \to W$ is the control and ∇ is the Levi-Civita connection of the metric m. Let $P \in \Gamma(T^*Q \otimes TQ)$ be the m-orthogonal projection with $\ker P = m^{-1}W$ where m-orthogonality means

$$m(PX, Y) = m(X, PY). \tag{2.68}$$

This projection P has the complete information of the control bundle W. Hence, the two pairs (L, F, W) and (L, F, P) contain the same information.

We now reformulate the Euler-Lagrange matching conditions in terms of Levi-Civita connections. Consider two CL systems (L_i, F_i, W_i), $i = 1, 2$ with

$$L_i(q, \dot{q}) = \frac{1}{2}m_i(q)(\dot{q}, \dot{q}) - V_i(q).$$

Let ∇^i be the Levi-Civita connection of the metric m_i, and $P_i \in \Gamma(T^*Q \otimes TQ)$ an m_i-orthogonal projection with $\ker P_i = m_i^{-1}W_i$. Then, the Euler-Lagrange matching conditions, **ELM-1** and **ELM-2** in Definition 2.1.3 can be equivalently written as follows:

ELM-1′ : $\ker P_1 = \ker P_2$

ELM-2′ : $P_1[\nabla^1_{\dot q}\dot q - \nabla^2_{\dot q}\dot q + m_1^{-1}\mathbf{d}V_1 - m_2^{-1}\mathbf{d}V_2 - (m_1^{-1}F_1 - m_2^{-1}F_2)] = 0.$

We are interested in the following question:

Given a CL system (L_1, F_1, W_1)(or, equivalently (L_1, F_1, P_1)), find its CL-equivalent CL systems.

One can regard **ELM-1** and **ELM-2** (or, equivalently, **ELM-1′** and **ELM-2′**) as partial differential equations for (L_2, F_2, W_2)(or, (L_2, F_2, P_2)). Without loss of generality, one may assume that $F_1 = 0$ by letting

$$F_2 \mapsto F_2 + m_2 m_1^{-1} F_1. \tag{2.69}$$

We decompose F_2 as follows:

$$F_2(q, \dot q) = F_2^q(q) + F_2^v(q, \dot q),$$

where $F_2^q(q)$ contains all the terms in F_2 which do not depend on the velocity $\dot q$. Then, by collecting the terms dependent on $\dot q$ and those independent of $\dot q$, **ELM-2′** is split into the two conditions:

$$P_1(\nabla^1_X X - \nabla^2_X X + m_2^{-1}F_2^v(X)) = 0, \tag{2.70}$$

$$P_1(m_1^{-1}\mathbf{d}V_1 - m_2^{-1}\mathbf{d}V_2 + m_2^{-1}F_2^q) = 0. \tag{2.71}$$

Let $\lambda = m_2^{-1}m_1 \in \Gamma(T^*Q \otimes TQ)$. Then λ is m_1 self-adjoint, i.e., $m_1\lambda = \lambda^* m_1 \in \Gamma(T^*Q \otimes T^*Q)$, or

$$m_1(\lambda X, Y) = m_1(X, \lambda Y) \tag{2.72}$$

for $X, Y \in TQ$.

The following is an extension of Proposition 1.1 in Auckly, Kapitanski, and White [2000]:

Proposition 2.4.1. *Let $\lambda = m_2^{-1} m_1$ and assume that m_2 and F_2^v satisfy (2.70). Then, λ satisfies*

$$\nabla^1_Z(m_1\lambda)(PX, PX) + \langle F_2^v(\lambda P_1 X + Z) - F_2^v(\lambda P_1 X) - F_2^v(Z),\ \lambda P_1 X\rangle = 0 \tag{2.73}$$

for $X, Z \in TQ$.

Proof. The following identity is from the proof of Proposition 1.1 in Auckly, Kapitanski, and White [2000]:

$$\nabla^1_Z(m_1\lambda)(P_1X, P_1X) = 2m_1(P_1(\nabla^1_{\lambda P_1 X} Z - \nabla^2_{\lambda P_1 X} Z), X). \tag{2.74}$$

Here we give a simpler and more accessible proof of (2.74) for the sake of completeness. Recall the following property of the Levi-Civita connection:

$$\nabla_X Y - \nabla_Y X = [X, Y]. \tag{2.75}$$

Notice that the three $(0,2)$-tensors m_1, m_2 and $m_1\lambda = m_1 m_2^{-1} m_1$ are symmetric. We have

$$\begin{aligned} m_1(P_1 \nabla^1_{\lambda P_1 X} Z,\ X) &= m_1(\nabla^1_{\lambda P_1 X} Z,\ P_1 X) \\ &= m_1(\nabla^1_Z \lambda P_1 X,\ P_1 X) - m_1([\lambda P_1 X, Z],\ P_1 X) \\ &= Z(m_1\lambda(P_1 X,\ P_1 X)) - (m_1\lambda)(P_1 X,\ \nabla^1_Z P_1 X) \\ &\quad - m_1([\lambda P_1 X, Z],\ P_1 X). \end{aligned} \tag{2.76}$$

We also have

$$\begin{aligned}
m_1(P_1\nabla^2_{\lambda P_1 X}Z,\, X) &= m_1(\nabla^2_{\lambda P_1 X}Z,\, P_1X) \\
&= m_2(\nabla^2_{\lambda P_1 X}Z,\, \lambda P_1X) \\
&= m_2(\nabla^2_Z \lambda P_1X,\, \lambda P_1X) - m_2([\lambda P_1X, Z],\, \lambda P_1X) \\
&= \frac{1}{2}Z(m_2(\lambda P_1X,\, \lambda P_1X)) - m_1([\lambda P_1X, Z],\, P_1X) \\
&= \frac{1}{2}Z(m_1\lambda(P_1X,\, P_1X)) - m_1([\lambda P_1X, Z],\, P_1X). \qquad (2.77)
\end{aligned}$$

By (2.76) and (2.77),

$$\begin{aligned}
2m_1(P_1(\nabla^1_{\lambda P_1 X}Z - \nabla^2_{\lambda P_1 X}Z), X) &= Z(m_1\lambda(P_1,\, P_1X)) - 2m_1\lambda(P_1X,\, \nabla^1_Z P_1X) \\
&= \nabla_Z(m_1\lambda)(P_1X, P_1X). \qquad (2.78)
\end{aligned}$$

This proves (2.74). By (2.75) and (2.70),

$$\begin{aligned}
&2P_1(\nabla^1_X Y - \nabla^2_X Y) \\
&= P_1(\nabla^1_{X+Y}(X+Y) - \nabla^2_{X+Y}(X+Y)) - P_1(\nabla^1_X X - \nabla^2_X X) - P_1(\nabla^1_Y Y - \nabla^2_Y Y) \\
&= -P_1 m_2^{-1} F_2^v(X+Y) + P_1 m_2^{-1} F_2^v(X) + P_1 m_2^{-1} F_2^v(Y). \qquad (2.79)
\end{aligned}$$

By (2.74), (2.79), and (2.68),

$$\nabla^1_Z(m_1\lambda)(PX, PX) + \langle F_2^v(\lambda P_1X + Z) - F_2^v(\lambda P_1 X) - F_2^v(Z),\, \lambda P_1 X\rangle = 0.$$

■

The following is an extension of Proposition 1.2 in Auckly, Kapitanski, and White [2000].

Proposition 2.4.2. *Let $\lambda = m_2^{-1}m_1$, and assume m_2 and F_2^v satisfy (2.70). Then, m_2 satisfies*

$$L_{\lambda P_1 X} m_2(Z, Z) = L_{P_1 X} m_1(Z, Z) - 2\langle F_2^v(Z), \lambda P_1 X\rangle \qquad (2.80)$$

for $X, Z \in TQ$.

Proof.

$$
\begin{aligned}
L_{\lambda P_1 X} m_2(Z,Z) &= 2m_2(\nabla^2_Z \lambda P_1 X, Z) \\
&= 2Zm_2(\lambda P_1 X, Z) - 2m_2(\lambda P_1 X, \nabla^2_Z Z) \\
&= 2Zm_1(P_1 X, Z) - 2m_1(P_1 X, \nabla^2_Z Z) \\
&= 2Zm_1(P_1 X, Z) - 2m_1(X, P_1 \nabla^2_Z Z) \\
&= [2Zm_1(P_1 X, Z) - 2m_1(X, P_1 \nabla^1_Z Z)] - 2m_1(X, P_1 m_2^{-1} F_2^v(Z)) \\
&= L_{P_1 X} m_1(Z,Z) - 2\langle F_2^v(Z), \lambda P_1 X\rangle.
\end{aligned}
$$

■

Proposition 2.4.3. *The condition (2.71) for V_2 is the same as the following:*

$$
L_{\lambda P_1 X} V_2 = L_{P_1 X} V_1 + \langle F_2^q, \lambda P_1 X\rangle \tag{2.81}
$$

for $X \in TQ$.

Proof. The equation (2.71) holds if and only if for all $X \in TQ$

$$
m_1(P_1(m_1^{-1}\mathbf{d}V_1 - m_2^{-1}\mathbf{d}V_2 + m_2^{-1}F_2^q),\, X) = 0
$$

from which (2.81) follows. ■

Proposition 2.4.1 and Proposition 2.4.2 convert the first-order quasi-linear PDE of m_2 into a first-order quasi-linear PDE (2.73) for λ and a first-order linear PDE (2.80) for m_2 where the coefficients of the derivatives of λ do not depend on λ. Hence, this splitting makes it easy to solve (2.70) for m_2. Notice that we are free to choose F_2^v in (2.73) and (2.80), which allows more solutions for m_2 than the original λ-method in Auckly, Kapitanski, and White [2000]. We will make use of this additional F_2^v in § 2.4.3.

The following proposition summarizes the extended λ-method:

Proposition 2.4.4. *Let $\lambda = m_2^{-1}m_1$. Then, m_2 and F_2^v satisfy (2.70) if and only if λ, F_2^v and m_2 satisfy (2.73) and (2.80). The equation (2.71) is equivalent to the following:*

for $X \in TQ$

$$L_{\lambda P_1 X} V_2 = L_{P_1 X} V_1 + \langle F_2^q,\, \lambda P_1 X\rangle. \tag{2.82}$$

Once m_2 *is derived,* W_2 *is given by* $W_2 = m_2 m_1^{-1} W_1$.

Proof. This proof is essentially the same as that in Auckly and Kapitanski [1999] except for the additional term F_2^v. For the sake of completeness we give the proof in the following.

First statement: ($\Rightarrow$) Proposition 2.4.1 and Proposition 2.4.2. ($\Leftarrow$) For any $X, Z \in TQ$,

$$\begin{aligned}
&m_1(P_1(\nabla^1_X X - \nabla^2_X X + m_2^{-1} F_2^v(X)), Z)\\
&= m_1(\nabla^1_X X - \nabla^2_X X + m_2^{-1} F_2^v(X), P_1 Z)\\
&= m_1(\nabla^1_X X, P_1 Z) - m_2(\nabla^2_X X, \lambda P_1 Z) + \langle F_2^v(X), \lambda P_1 Z\rangle\\
&= X m_1(X, P_1 Z) - m_1(X, \nabla^1_X P_1 Z) - X m_2(X, \lambda P_1 Z) + m_2(X, \nabla^2_X \lambda P_1 Z)\\
&\qquad + \langle F_2^v(X), \lambda P_1 Z\rangle\\
&= -m_1(X, \nabla^1_X P_1 Z) + m_2(X, \nabla^2_X \lambda P_1 Z) + \langle F_2^v(X), \lambda P_1 Z\rangle\\
&= -\frac{1}{2} L_{P_1 Z} m_1(X, X) + \frac{1}{2} L_{\lambda P_1 Z} m_2(X, X) + \langle F_2^v(X), \lambda P_1 Z\rangle\\
&= -\langle F_2^v(X), \lambda P_1 Z\rangle + \langle F_2^v(X), \lambda P_1 Z\rangle\\
&= 0.
\end{aligned}$$

Since this holds for all $X, Z \in TQ$, the equation (2.70) follows.

Second statement: by Proposition 2.4.3.

Third statement: by **ELM-1′**. ■

2.4.2 Application to Stabilization Problems

We apply the extended λ-method to stabilization problems. Here, we derive general formulae and in § 2.4.3 we apply them to the problem of stabilization of the inverted pendulum on a rotor arm. The work in this section is a generalization of Auckly and Kapitanski [2001] by taking additional gyroscopic forces into account. We keep most of the notations used in Auckly and Kapitanski [2001].

Denote by Q the configuration space of dimension s. Suppose that we are given a CL system $(L_1, F_1 = 0, W_1) = (L, 0, W)$ with

$$L(q,\dot{q}) = \frac{1}{2}\sum_{i=1}^{s} m_{ij}(q)\dot{q}^i\dot{q}^j - V(q)$$

and

$$W = \langle \mathbf{d}q^a \mid a = r+1, \ldots, s\rangle$$

with $r < s$. In this section, we use indices as follows:

$$\begin{aligned} &i, j, k, l = 1, \ldots, s \\ &\alpha, \beta, \gamma = 1, \ldots, r\,(< s) \\ &a, b, c = r+1, \ldots, s. \end{aligned}$$

The m-orthogonal projection $P \in \Gamma(T^*Q \otimes TQ)$ with $\ker P = m^{-1}W$ is given in coordinates by

$$\begin{cases} P\left(\frac{\partial}{\partial q^\alpha}\right) = \frac{\partial}{\partial q^\alpha} & \text{if } \alpha = 1, \ldots, r, \\ P\left(m^{ia}\frac{\partial}{\partial q^i}\right) = 0 & \text{if } a = r+1, \ldots, s. \end{cases} \tag{2.83}$$

We want to find CL systems $(L_2, F_2, W_2) = (\widehat{L}, \widehat{F}, \widehat{W})$ with $\widehat{F}$ gyroscopic force *quadratic in* $\dot{q}$ of the following form:

$$\widehat{L}(q,\dot{q}) = \frac{1}{2}\widehat{m}_{ij}(q)\dot{q}^i\dot{q}^j - \widehat{V}(q)$$

and

$$\widehat{F}(X) = \widehat{F}(X)_k\mathbf{d}q^k = \widehat{G}_{ijk}(q)X^iX^j\mathbf{d}q^k, \qquad \widehat{G}_{ijk} = -\widehat{G}_{ikj}$$

with $\widehat{G} = \widehat{G}_{ijk}\mathbf{d}q^i \otimes \mathbf{d}q^j \otimes \mathbf{d}q^k \in \Gamma(T^*Q \otimes T^*Q \otimes T^*Q)$, where the skew symmetry in the last two indices of G_{ijk} comes from the fact that $\widehat{F}$ is gyroscopic. Among these equivalent systems, we will use a CL system whose energy has a minimum at the equilibrium of interest in applications.

We apply Proposition 2.4.1, 2.4.2, 2.4.3, and 2.4.4 to the current problem, and then express the extended λ-method in coordinates.

Proposition 2.4.5. *Suppose* $(L, 0, W) \overset{L}{\sim} (\widehat{L}, \widehat{F}, \widehat{W})$*, where*

$$L = \frac{1}{2}m(\dot{q}, \dot{q}) - V(q), \qquad \widehat{L} = \frac{1}{2}\widehat{m}(\dot{q}, \dot{q}) - \widehat{V}(q),$$
$$W = \langle \mathbf{d}q^a \mid a = r+1, \dots, s\rangle, \quad \widehat{F}(X) = \widehat{G}(X, X)$$

for $X \in TQ$ *with* $\dim Q = s$ *and* $\widehat{G}$ *a* $(0,3)$*-tensor satisfying* $\widehat{G}_{ijk} = -\widehat{G}_{ikj}$*. Let* ∇ *be the Levi-Civita connection of the metric* m*, and* P *be the* m*-orthogonal projection with* $\ker = m^{-1}W$*. Then, the following holds: for all* $X, Y, Z \in TQ$

$$\nabla_Z(m\lambda)(PX, PY) - \frac{1}{2}\langle \widehat{G}(\lambda PX, \lambda PY) + \widehat{G}(\lambda PY, \lambda PX),\ Z\rangle = 0, \tag{2.84}$$

$$L_{\lambda PX}\widehat{m}(Y, Z) = L_{PX}m(Y, Z) - \langle \widehat{G}(Y, Z) + \widehat{G}(Z, Y),\ \lambda PX\rangle, \tag{2.85}$$

$$L_{\lambda PX}\widehat{V} = L_{PX}V. \tag{2.86}$$

In coordinates,

$$\frac{\partial(m_{\alpha i}\lambda^i_\beta)}{\partial q^k} - [\alpha k, i]\lambda^i_\beta - [\beta k, i]\lambda^i_\alpha - \frac{1}{2}(\widehat{G}_{ijk} + \widehat{G}_{jik})\lambda^i_\alpha\lambda^j_\beta = 0, \tag{2.87}$$

$$\lambda^k_\alpha \frac{\partial \widehat{m}_{ij}}{\partial q^k} + \frac{\partial \lambda^k_\alpha}{\partial q^i}\widehat{m}_{kj} + \frac{\partial \lambda^k_\alpha}{\partial q^j}\widehat{m}_{ki} = \frac{\partial m_{ij}}{\partial q^\alpha} - (\widehat{G}_{ijk} + \widehat{G}_{jik})\lambda^k_\alpha, \tag{2.88}$$

$$\lambda^k_\alpha \frac{\partial \widehat{V}}{\partial q^k} = \frac{\partial V}{\partial q^\alpha} \tag{2.89}$$

where $i, j, k = 1, \dots, s$*, and* $\alpha, \beta = 1, \dots, r$*.*

Proof. Proof of (2.84) and (2.87) : Using the property $\widehat{G}(X, Y)Z = -\widehat{G}(X, Z)Y$, one can write (2.73) as follows:

$$\nabla_Z(m\lambda)(PX, PX) = \langle \widehat{G}(\lambda PX, \lambda PX), Z\rangle.$$

Notice that $m\lambda$ and $\nabla_Z(m\lambda)$ are symmetric $(0,2)$ tensors. Hence,

$$\begin{aligned}\nabla_Z(m\lambda)(PX,PY) &= \frac{1}{2}[\nabla_Z(m\lambda)(P(X+Y),P(X+Y))\\ &\qquad - \nabla_Z(m\lambda)(PX,PX) - \nabla_Z(m\lambda)(PY,PY)]\\ &= \frac{1}{2}\langle \widehat{G}(\lambda PX,\lambda PY) + \widehat{G}(\lambda PY,\lambda PX),\ Z\rangle\end{aligned}$$

which proves (2.84). In coordinates, for $Z = \frac{\partial}{\partial q^k}$,

$$\begin{aligned}\nabla_Z(m\lambda) &= \nabla_{\frac{\partial}{\partial q^k}}(m_{il}\lambda^l_j \mathbf{d}q^i \otimes \mathbf{d}q^j)\\ &= \left(\frac{\partial(m_{il}\lambda^l_j)}{\partial q^k} - \Gamma^s_{ki}m_{sl}\lambda^l_j - \Gamma^r_{kj}m_{il}\lambda^l_r\right)\mathbf{d}q^i\otimes\mathbf{d}q^j\\ &= \left(\frac{\partial(m_{il}\lambda^l_j)}{\partial q^k} - \Gamma^s_{ki}m_{sl}\lambda^l_j - \Gamma^r_{kj}m_{rl}\lambda^l_i\right)\mathbf{d}q^i\otimes\mathbf{d}q^j\\ &= \left(\frac{\partial(m_{il}\lambda^l_j)}{\partial q^k} - [ik,l]\lambda^l_j - [jk,l]\lambda^l_i\right)\mathbf{d}q^i\otimes\mathbf{d}q^j \qquad (2.90)\end{aligned}$$

and

$$\langle \widehat{G}(\cdot,\cdot),\ \frac{\partial}{\partial q^k}\rangle = \widehat{G}_{ijk}\mathbf{d}q^i\otimes\mathbf{d}q^j. \qquad (2.91)$$

Notice from (2.83) that

$$\operatorname{Im} P = \left\langle \frac{\partial}{\partial q^\alpha} \mid \alpha = 1,\ldots,r\right\rangle. \qquad (2.92)$$

The equation (2.87) follows from (2.84), (2.90), (2.91) and (2.92).

Proof of (2.85) and (2.88): Notice that m and $\widehat{m}$ are symmetric tensors. So, (2.85) follows from the polarization of (2.80) in the same way as in the proof of (2.84). By (2.92), consider the case that $PX = \frac{\partial}{\partial q^\alpha}$, $\alpha = 1,\ldots,r$. Then,

$$L_{\lambda PX}\widehat{m} = \left(\lambda^k_\alpha\frac{\widehat{m}_{ij}}{\partial q^k} + \frac{\partial\lambda^k_\alpha}{\partial q^i}\widehat{m}_{kj} + \frac{\partial\lambda^k_\alpha}{\partial q^j}\widehat{m}_{ki}\right)\mathbf{d}q^i\otimes\mathbf{d}q^j \qquad (2.93)$$

and

$$L_{PX}m = \frac{\partial m_{ij}}{\partial q^\alpha}\mathbf{d}q^i\otimes\mathbf{d}q^j. \qquad (2.94)$$

The equation (2.88) follows from (2.85), (2.93), (2.94) and (2.91).

Proof of (2.86) and (2.89) : Proposition 2.4.3 and (2.92). ■

Remark 2.4.6. *If one sets $\widehat{F}$ or $\widehat{G}$ to zero, the results in Auckly and Kapitanski [2001] are recovered.*

2.4.3 Example: Inverted Pendulum on a Rotor Arm

We apply the extended λ-method to the design of an asymptotically stabilizing feedback control law for the inverted pendulum on a rotor arm shown in Figure 2.6. This example was first handled in Bloch, Leonard, and Marsden [1999a] enjoying the CL method, but they did not accomplish the asymptotic stabilization of the position of the lower arm.

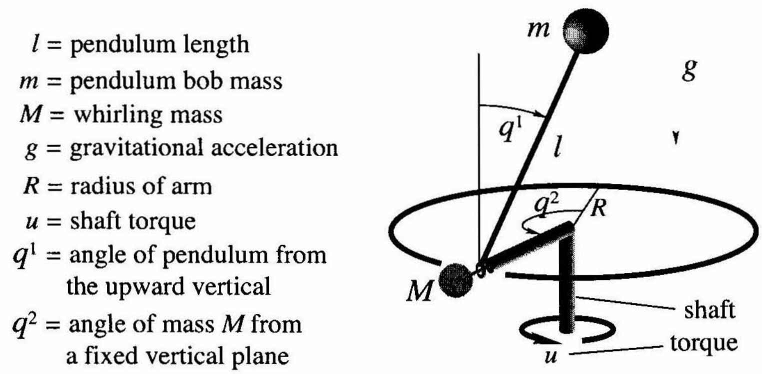

Figure 2.6: Inverted pendulum on a rotor arm.

The configuration space is $Q = S^1 \times S^1$. We will use (q^1, q^2) as local coordinates where q^1 is the angle of the pendulum from the upward vertical and q^2 is the angle of mass M from a fixed vertical plane. This system can be described as a CL system $(L, 0, W)$ with

$$\begin{aligned}L(q^1, q^2, \dot{q}^2, \dot{q}^2) = &\frac{1}{2} ml^2 (\dot{q}^1)^2 + mRl\cos(q^1)\dot{q}^1\dot{q}^2 \\ &+ \frac{1}{2}((m+M)R^2 + ml^2 \sin^2(q^1))(\dot{q}^2)^2 - mgl\cos(q^1),\end{aligned}$$

and

$$W = \langle \mathbf{d}q^2 \rangle.$$

For simplicity, we will use the following Lagrangian instead:

$$L(q^1, q^2, \dot{q}^2, \dot{q}^2) = \frac{1}{2}(\dot{q}^1)^2 + \cos(q^1)\dot{q}^1\dot{q}^2 + \frac{1}{2}(2 + \sin^2(q^1))(\dot{q}^2)^2 - \cos(q^1).$$

Then, the nonzero Christoffel symbols of the first kind are given by

$$[11,2] = -\sin(q^1);\ [12,2] = [21,2] = \sin(q^1)\cos(q^1);\ [22,1] = -\sin(q^1)\cos(q^1).$$

The λ-equations in (2.87) are given by

$$\begin{cases} \frac{\partial\lambda_1^1}{\partial q^1} + \cos(q^1)\frac{\partial\lambda_1^2}{\partial q^1} + \sin(q^1)\lambda_1^2 - \lambda_1^1\lambda_1^2\widehat{G}_{121} - \lambda_1^2\lambda_1^2\widehat{G}_{221} = 0, \\ \\ \frac{\partial\lambda_1^1}{\partial q^2} + \cos(q^1)\frac{\partial\lambda_1^2}{\partial q^2} - 2\sin(q^1)\cos(q^1)\lambda_1^2 - \lambda_1^1\lambda_1^1\widehat{G}_{112} - \lambda_1^1\lambda_1^2\widehat{G}_{212} = 0. \end{cases} \tag{2.95}$$

In stead of dealing with general solutions of (2.95), we will restrict ourselves to a particular solution. To simplify the second equation in (2.95), we choose

$$\widehat{G}_{212} = -\widehat{G}_{221} = -\frac{1}{\lambda_1^1}2\sin(q^1)\cos(q^1);\quad \widehat{G}_{112} = \widehat{G}_{121} = 0.$$

Then, (2.95) becomes

$$\begin{cases} \frac{\partial\lambda_1^1}{\partial q^1} + \cos(q^1)\frac{\partial\lambda_1^2}{\partial q^1} + \sin(q^1)\lambda_1^2 - \frac{\lambda_1^2\lambda_1^2}{\lambda_1^1}2\sin(q^1)\cos(q^1), \\ \\ \frac{\partial\lambda_1^1}{\partial q^2} + \cos(q^1)\frac{\partial\lambda_1^2}{\partial q^2} = 0. \end{cases} \tag{2.96}$$

We will use the following particular solution to (2.96):

$$\lambda_1^1 = -\frac{1}{k_1};\quad \lambda_1^2 = \frac{\cos(q^1)}{k_2 - k_1\cos^2(q^1)}$$

with $k_1, k_2 \in \mathbb{R}$. Hence,

$$\widehat{G}_{212} = -\widehat{G}_{221} = 2k_1\sin(q^1)\cos(q^1);\quad \widehat{G}_{112} = \widehat{G}_{121} = 0.$$

With this solution, the equations (2.88) become

$$\begin{cases} \lambda_1^1 \frac{\partial \widehat{m}_{11}}{\partial q^1} + \lambda_1^2 \frac{\partial \widehat{m}_{11}}{\partial q^2} + 2\frac{\partial \lambda_1^2}{\partial q^1} \widehat{m}_{21} = 0, \\ \\ \lambda_1^1 \frac{\partial \widehat{m}_{12}}{\partial q^1} + \lambda_1^2 \frac{\partial \widehat{m}_{12}}{\partial q^2} + \frac{\partial \lambda_1^2}{\partial q^1} \widehat{m}_{22} = -\sin(q^1) - \widehat{G}_{212}\lambda_1^2, \\ \\ \lambda_1^1 \frac{\partial \widehat{m}_{22}}{\partial q^1} + \lambda_1^2 \frac{\partial \widehat{m}_{22}}{\partial q^2} = -2\sin(q^1)\cos(q^1). \end{cases} \tag{2.97}$$

One can solve the third equation in (2.97) for $\widehat{m}_{22}$ as follows:

$$\widehat{m}_{22}(q^1, q^2) = c_{22}(x^2(q^1, q^2)) - k_1 \cos^2(q^1),$$

where c_{22} is an arbitrary function of x^2 and x^2 is defined by

$$x^2 = x^2(q^1, q^2) := q^2 + \sqrt{\frac{k_1}{k_2 - k_1}} \arctan\left(\frac{2\tan(\frac{q^1}{2})}{1 + \tan^2(\frac{q^1}{2})}\sqrt{\frac{k_1}{k_2 - k_1}}\right).$$

Now, one does not need to solve the first two equations in (2.97) for $\widehat{m}_{12}, \widehat{m}_{11}$. Using the relation $\widehat{m}\lambda = m$, one gets

$$\begin{aligned} \widehat{m}_{12} &= \frac{-k_1(k_2 - c_{22}(x^2))\cos(q^1)}{k_2 - k_1\cos^2(q^1)}, \\ \widehat{m}_{11} &= -k_1\left(1 + \frac{k_1(k_2 - c_{22}(x^2))\cos^2(q^1)}{(k_2 - k_1\cos^2(q^1))^2}\right). \end{aligned}$$

The equation (2.89) is given by

$$\lambda_1^1 \frac{\partial \widehat{V}}{\partial q^1} + \lambda_1^2 \frac{\partial \widehat{V}}{\partial q^2} = -\sin(q^1)$$

whose solution is

$$\widehat{V}(q^1, q^2) = -k_1\cos^2(q^1) + U(x^2(q^1, q^2))$$

where $U(\cdot)$ is an arbitrary function. The control bundle $\widehat{W} = \widehat{m}m^{-1}W$ is given by

$$\widehat{W} = \left\langle \mathbf{d}q^2 + \frac{k_1\cos(q^1)}{k_2 - k_1\cos^2(q^1)}\mathbf{d}q^1 \right\rangle.$$

So far, we have found a CL system $(\widehat{L}, \widehat{F}, \widehat{W})$ parameterized by real numbers k_1, k_2 and functions c_{22}, U, which is CL-equivalent to the original system. Hence, we will equivalently work with $(\widehat{L}, \widehat{F}, \widehat{W})$ to find a controller that asymptotically stabilizes $(0,0,0,0)$. To simplify computation, we use new coordinates (x^1, x^2) defined by

$$x = (x^1, x^2) := \left(q^1, q^2 + \sqrt{\frac{k_1}{k_2 - k_1}} \arctan\left(\frac{2\tan(\frac{q^1}{2})}{1+\tan^2(\frac{q^1}{2})}\sqrt{\frac{k_1}{k_2-k_1}}\right)\right).$$

In the new coordinates (x^1, x^2), the CL system $(\widehat{L}, \widehat{F}, \widehat{W})$ is expressed as

$$\widehat{m}_x = \begin{bmatrix} \frac{-k_1(k_2 - 2k_1\cos^2(x^1))}{k_2 - k_1\cos^2(x^1)} & -k_1\cos(x^1) \\ -k_1\cos(x^1) & c_{22}(x^2) - k_1\cos^2(x^1) \end{bmatrix},$$

$$\widehat{V}(x) = -k_1\cos(x^1) + U(x^2),$$

$$\widehat{F}(x, \dot{x}) = 2k_1 \sin(x^1)\cos(x^1)\left(\dot{x}^2 - \frac{k_1\cos(x^1)\dot{x}^1}{k_2 - k_1\cos^2(x^1)}\right)(-\dot{x}^2\mathbf{d}x^1 + \dot{x}^1\mathbf{d}x^2),$$

$$\widehat{W} = \langle \mathbf{d}x^2 \rangle.$$

Let $\widehat{E} = \frac{1}{2}\widehat{m}(\dot{x}, \dot{x}) + \widehat{V}(x)$ be the energy of the CL system $(\widehat{L}, \widehat{F}, \widehat{W})$. One can check that the energy $\widehat{E}$ has a local minimum at $(0,0,0,0)$ if

$$c_{22}(0) > k_1 > 0, \qquad \frac{k_1(k_1^2 - 2k_1c_{22}(0) + k_2c_{22}(0))}{-k_2 + k_1} > 0; \tag{2.98}$$

$$U'(0) = 0, \qquad U''(0) > 0. \tag{2.99}$$

To make the new coordinates (x^1, x^2) be real-valued coordinates, we need the following additional condition:

$$k_2 > k_1. \tag{2.100}$$

One can always find some parameters $k_1, k_2, c_{22}(x^2), U(x^2)$ satisfying (2.98), (2.99), (2.100). Here, for simplicity, we use the following U:

$$U(x^2) = \frac{1}{2}\epsilon(x^2)^2, \qquad \epsilon > 0.$$

Take the following dissipative force $\widehat{u}$ as a feedback control to the system $(\widehat{L}, \widehat{F}, \widehat{W})$:

$$\widehat{u}(x, \dot{x}) = -c\dot{x}^2 \mathbf{d}x^2, \qquad c > 0.$$

Then,

$$\frac{d\widehat{E}}{dt} = -c(\dot{x}^2)^2 \leq 0.$$

Hence, the equilibrium point $(0,0,0,0)$ is Lyapunov stable in the closed-loop dynamics $(\widehat{L}, \widehat{F}, \widehat{u})$.

We now show the asymptotic stability of the origin $(0,0,0,0)$ in the closed-loop system $(\widehat{L}, \widehat{F}, \widehat{u})$. Since $\widehat{E}$ has a strict local minimum at the origin and $d\widehat{E}/dt \leq 0$, there is a real number l such that the set

$$\Omega_l = \{(x, \dot{x}) \mid \widehat{E}(x, \dot{x}) \leq l\}$$

is non-empty, compact and positively invariant. Define

$$\begin{aligned}
\mathcal{E} &:= \{(x, \dot{x}) \in \Omega_l \mid d\widehat{E}/dt = 0\} = \{(x^1, x^2, x^1, \dot{x}^2) \in \Omega_l \mid \dot{x}^2 = 0\}, \\
\mathcal{M} &:= \text{ the largest invariant subset of } \mathcal{E}.
\end{aligned}$$

Suppose that a trajectory $(x(t), \dot{x}(t)) = (x^1(t), x^2(t), \dot{x}^1(t), \dot{x}^2(t))$ is contained in $\mathcal{M}$ for all $t \geq 0$. By the definition of $\mathcal{M}$, we have

$$(x^1(t), x^2(t), \dot{x}^1(t), \dot{x}^2(t)) = (x^1(t), x^2(0), \dot{x}^1(t), 0) \qquad \forall t \geq 0. \tag{2.101}$$

The trajectory obeys the equations of motion of $(\widehat{L}, \widehat{F}, \widehat{u})$, which along the trajectory (2.101) become as follows:

$$\frac{-k_1(k_2 - 2k_1\cos^2(x^1))}{k_2 - k_1\cos^2(x^1)}\ddot{x}^1 - \frac{k_1^2 k_2 \cos(x^1)\sin(x^1)}{(k_2 - k_1\cos^2(x^1))^2}(\dot{x}^1)^2 + k_1 \sin(x^1) = 0, \tag{2.102}$$

$$-k_1\cos(x^1)\ddot{x}^1 + k_1\sin(x^1)(\dot{x}^1)^2 + \epsilon x^2(0) = -\frac{2k_1^2 \sin(x^1)\cos^2(x^1)}{k_2 - k_1\cos^2(x^1)}(\dot{x}^1)^2. \tag{2.103}$$

If $x^1(t) = 0$ for all $t \geq 0$, (2.103) implies that $x^2(0) = 0$ for all $t \geq 0$. Hence, we will have $(x^1(t), x^2(t)\dot{x}^1(t), \dot{x}^2(t)) = (0,0,0,0)$ for all $t \geq 0$. Now, suppose that $x^1(t) \neq 0$ for some $t \geq 0$. Notice that the following quantity C_o (it is the part of the energy $\widehat{E}$ corresponding to the variables $(x^1, \dot{x}^1)$, or it is the energy of (2.102)) is constant along the trajectory of (2.102):

$$C_o = \frac{1}{2}\frac{-k_1(k_2 - 2k_1\cos^2(x^1))}{k_2 - k_1\cos^2(x^1)}(\dot{x}^1)^2 - k_1\cos(x^1). \tag{2.104}$$

Notice that C_o has a strict local minimum at $(x^1, \dot{x}^1) = (0,0)$. Hence, the trajectory $(x^1(t), \dot{x}^1(t))$ will oscillate about $(0,0)$, so $x^1(t)$ satisfies,

$$\min x^1([0,\infty)) = -a \le x^1(t) \le b = \max x^1([0,\infty)), \quad \forall t \ge 0 \tag{2.105}$$

for some $a, b > 0$. Now, solve (2.102) for $\ddot{x}^1$, substitute it into (2.103), solve the resultant equation for $(\dot{x}^1)^2$, and then substitute it into (2.104). Then, along the trajectory $(x^1(t), x^2(0), \dot{x}^1(t), 0)$ the following holds:

$$\begin{aligned} C_o = -k_1\cos(x^1) &+ \frac{\cos(x^1)(k_1\cos^2(x^1) - k_2)(2k_1\cos^2(x_1) - k_2)}{2(2k_1^2\cos^4(x^1) - k_2^2)} \\ &- \frac{\epsilon x^2(0)(2k_1\cos^2(x^1) - k_2)^2}{2\sin(x^1)(2k_1^2\cos^4(x^1) - k_2^2)}. \end{aligned} \tag{2.106}$$

By (2.105), the identity (2.106) must hold in the interval $[-a, b]$ of x^1 with $a, b > 0$. Since only the term with the factor $\sin(x^1)$ in (2.106) is an odd function, that term must vanish (also notice only the term with the factor $\sin(x^1)$ blows up at 0). So, we have

$$x^2(0) = 0.$$

Then, the following must hold in the interval $[-a, b]$ of x^1:

$$C_o = -k_1\cos(x^1) + \frac{\cos(x^1)(k_1\cos^2(x^1) - k_2)(2k_1\cos^2(x_1) - k_2)}{2(2k_1^2\cos^4(x^1) - k_2^2)}$$

which is impossible unless $x^1(t)$ is constant. So far, we have shown that

$$(x^1(t), x^2(t), \dot{x}^1(t), \dot{x}^2(t)) = (x^1(0), 0, 0, 0).$$

Substitution of this into (2.103) implies $\sin(x^1(0)) = 0$. It follows that $x^1(0) = 0$. Therefore, the only possible trajectory in the set $\mathcal{M}$ is the equilibrium, $(0,0,0,0)$ only. By LaSalle's theorem, the origin is an asymptotically stable equilibrium in the closed-loop system $(\widehat{L}, \widehat{F}, \widehat{W})$. The feedback control u to the original system can be derived from (2.12).

Remark 2.4.7. *Notice that the gyroscopic force $\widehat{F}$ was useful in this example.*

2.5 Summary and Future Work

We have developed the method of controlled Lagrangian (CL) systems and applied it to a couple of systems for stabilization. Here, we summarize this chapter section by section.

§ **2.1.** We first reviewed Lagrangian mechanics. We then defined CL systems on TQ (Definition 2.1.1). Then we defined the CL-equivalence relation among CL systems on TQ and the Euler-Lagrange matching conditions (Definition 2.1.3) for simple CL systems (Definition 2.1.2). If two simple CL systems are CL-equivalent, then for any control for one system there exists a control for the other system such that the two closed-loop systems produce the same equations of motion (Proposition 2.1.5). We then introduced the CL-inclusion relation, which is a partial order in the class of CL systems on TQ (Definition 2.1.6). If one simple CL system includes an another, then for any control for the included system, there exists a control for the including system such that the two closed-loop systems produce the same equations of motion (Proposition 2.1.7). Two simple CL systems are equivalent if and only if they include each other. We illustrated the CL equivalence and the CL inclusion by understanding the well-known collocated/noncollocated partial feedback transformations within the framework of the CL equivalence/inclusion (§ 2.1.3). In addition, we defined notions of energy, dissipative forces, and gyroscopic forces (§ 2.1.2). They are physical quantities which are closely related to stability. Dissipative forces decrease energy. Gyroscopic forces do not change energy but introduce couplings to the dynamics.

§ **2.2.** We gave the usual procedure of applying the method of CL systems to control synthesis for asymptotic stabilization. The basic idea is as follows: Given a CL system of the form $(L, 0, F)$, which is of the usual *ideal* form in applications, find a CL-equivalent system $(\widehat{L}, \widehat{F}_{\mathrm{gr}}, \widehat{W})$, where the energy of the second system has a minimum at the equilibrium of interest and $\widehat{F}_{\mathrm{gr}}$ is a gyroscopic force. Then, add a dissipative feedback force in the direction of $\widehat{W}$. We used this method to design an asymptotically stabilizing control law for the inverted pendulum on a cart with a relatively large region of attraction. In that example, dissipation alone was enough for us to achieve the goal. We did not have to employ any gyroscopic forces. We performed several computer simulations.

§ **2.3.** We found a class of CL systems for which an asymptotically stabilizing control law can be designed with the CL method. This class contains systems such as the inverted pendulum on a cart and the spherical pendulum on a cart. For this same class of systems, we were able to design tracking controllers for a constant acceleration reference signal.

§ **2.4.** We extended the λ-method. The λ-method was originally developed by Auckly, Kapitanski, and White [2000] and Auckly and Kapitanski [2001] to systematically solve PDE's involved in the second Euler-Lagrange matching condition, **ELM-2** for simple CL systems of the form $(L, 0, W)$. Our *extended* λ-method considers the full form of CL systems (L, F, W) such that not only can one solve the PDE's with more freedom, but also more importantly, introduce gyroscopic forces into the dynamics. The usual procedure of applying the extended λ-method to stabilization problems is given in Proposition 2.4.5. We applied it to the design of an asymptotically stabilizing controller for the inverted pendulum on a rotor arm (§ 2.4.3). In this application, we made use of a gyroscopic force as well as a dissipative feedback control.

Future Work.

1. There needs to be a simple criterion, like the controllability rank condition, which tells when the CL method provides an asymptotically stabilizing controller for a given system.

2. When we applied the CL method to stabilization problems, we assumed that a given CL system does not have any external forces, i.e., it is of the form $(L, 0, W)$. Suppose that two simple systems (L_1, F_1, W_1) and (L_2, F_2, W_2) are CL-equivalent and (L_1, F_1, W_1) is the system for which we want to design a control to asymptotically stabilize an equilibrium of interest. Then the force F_1 is transformed to $m_2 m_1^{-1} F_1$ in the second system. However it may be that the force $m_2 m_1^{-1} F_1$ increases the energy E_2 of the second system while the energy E_2 has a minimum at the equilibrium and the force $m_2 m_1^{-1} F$ cannot be cancelled

out by any feedbacks u_2 for the second system. This type of difficulty never happens when $F_1 = 0$. One needs to do more systematic studies of the cases when $F_1 \neq 0$. Woolsey [2001] studied the case when there is a physical dissipation to the original system along the *unactuated* directions.

3. It is reasonable to expect that one can extend the method of CL systems so that it is applied to nonholonomic systems and elastic systems; see Bloch, Krishnaprasad, Marsden, and Murray [1996], Zenkov, Bloch, Leonard, and Marsden [2000] and Zenkov, Bloch, and Marsden [2002].

4. In this chapter, we only considered static feedback laws, i.e., control forces depend on $(q, \dot{q})$. However, one can generalize the CL method such that it includes the dynamic feedback control laws, too.

5. A preliminary work on tracking was done in § 2.3.3. More research on application of the CL method to general tracking problems is necessary. For this purpose, one should generalize the CL method by allowing the explicit time-dependence of CL systems, which is straightforward. Let us make a general comment on a technical point about tracking. When one wants to design a tracking controller, he needs to compare the current state (vector) with the reference state (vector). In general, these two vectors are based at different points. Hence, just naive *subtraction* of these vectors would only hold in local coordinates unless the configuration space is $\mathbb{R}^n$. One might want to make use of any possible geometric structures to compare two vectors; the parallel transport with a Riemannian metric, or the left or right translation for Lie groups. In particular, when the configuration space is a Lie group, the Killing form is very useful whether or not it is nondegenerate.

6. In mechanics, symmetry gives a conserved quantity. For example, the energy is due to the time symmetry and angular momentum is due to the rotation symmetry. In control, we usually make use of them by changing those quantities using control forces. In the CL method, we made use of the energy for stabilization. It would be interesting to consider other possible conserved quantities that may be present in a specific problem. The energy-Casimir method is one way (Bloch, Chang, Leonard, Marsden and Woolsey [2000]). See also Chang, Chichka, and Marsden [2002] or Appendix A in this thesis where they used the angular momentum and Laplace-Runge-Lenz vector to design a feedback

control for transfer between elliptic Keplerian orbits.

Chapter 3

The Method of Controlled Hamiltonian Systems

It is well known in mechanics that most mechanical systems can be described by Hamiltonian mechanics as well as by Lagrangian mechanics. In a similar manner, there is a Hamiltonian counterpart to the method of CL systems. We call it the method of controlled Hamiltonian (CH) systems. Unlike the CL method, the CH method has been used under a couple of different names and in different versions; modification of Hamiltonian and Hamiltonian structures (Bloch, Krishnaprasad, Marsden, and Sánchez De Alvarez [1992] and Woolsey and Leonard [1999]) and interconnection and damping assignment passivity beased control, or IDA-PBC (Ortega and Spong [2000]). In this chapter, we give the intrinsic formulation of the method of CH systems, and then show that the method of CL systems and that of CH systems are equivalent for simple mechanical systems. We outline this chapter in the following.

A controlled Hamiltonian (CH) systems is a quadruple (H, B, F, W) of a Hamiltonian H on the cotangent bundle T^*Q, an almost Poisson tensor B, i.e., a skew-symmetric (2,0)-tensor on T^*Q, an external force $F : T^*Q \to T^*Q$ and a control bundle $W \subset T^*Q$. A feedback control is a map $u : T^*Q \to W$. Once we choose a control u, we call the quadruple (H, B, F, u) the closed-loop (Hamiltonian) system. The equations of motion, or the vector field on T^*Q of the closed-loop system (H, B, F, u) is given by

$$X_{(H,B,F,u)} = B^\sharp \mathbf{d}H + \mathrm{vlift}(F) + \mathrm{vlift}(u),$$

where the operator $\mathrm{vlift}(\cdot)$ denotes the vertical lift. We call a CH system *simple* if the Hamiltonian has the form of kinetic plus potential energy, and the almost Poisson tensor is of the particular form described in this chapter.

We define an equivalence relation by feedback transformation among CH systems on a cotangent space in a similar way to the CL-equivalence relation. We call this equivalence relation the CH-equivalence relation. The method of CH systems is used in a similar manner to that of CL systems.

In mechanics, the Jacobi identity of a Poisson tensor is very important because it is related to an integrability condition. In control problems, however, we sometimes want to design a control so that the trajectory can move from one point to another by departing from any integral submanifold that may be present. This is why we only use almost Poisson tensors in the definition of CH systems without requiring the Jacobi identity. However, the skew symmetry of an almost Poisson tensor is needed because it is related to the energy conservation in the absence of any external or control forces. In this thesis, we also study the relationship between the failure of the Jacobi identity, and gyroscopic forces. We identify the failure of the Jacobi identity in terms of gyroscopic forces and show that gyroscopic forces can be incorporated into the almost Poisson tensor. Hence, in the application of the CH method to stabilization problems, one need not use a gyroscopic force explicitly because it can be encoded into the almost Poisson tensor. This point implies that the introduction of gyroscopic forces to the CL side is crucial in two ways: 1. asymptotic stabilization, and 2. equivalence of the CL method and the CH method.

In applications, most mechanical systems are simple CL/CH systems, that is, simple mechanical systems. We show that the method of CL systems and that of CH systems are equivalent for simple mechanical systems. Although it sounds obvious, the proof requires quite a development of theory. This equivalence implies that one can use either method for applications. Which method to use depends on the specifically given problem, just as the preferred choice of coordinates depends on the given PDE. However, one should remember that on the CL side, the extended λ-method is available to systematically solve the involved PDE's.

3.1 Controlled Hamiltonian (CH) Systems

In this section, we develop the method of CH systems rigorously. This will include all previously known theories such as modification of Hamiltonian and Hamiltonian structures (Bloch, Krishnaprasad, Marsden, and Sánchez De Alvarez [1992] and Woolsey and Leonard [1999]) and IDA-PBC (Ortega and Spong [2000]). In addition, our setting will prove especially useful when we perform reduction of CH systems with symmetry in § 4.

3.1.1 Controlled Hamiltonian Systems

We start this section with a review of classical Hamiltonian mechanics.

Review of Hamiltonian Mechanics. We review the Hamiltonian mechanics. More detail can be found in Marsden and Ratiu [1999]. A ***symplectic manifold*** is a pair (P, Ω), where P is a manifold and Ω is a closed nondegenerate two-form on P called the

symplectic form. The dimension of P becomes even automatically. A vector field X on P is called Hamiltonian if there is a function $H : P \to \mathbb{R}$ such that

$$\mathbf{i}_X \Omega = \mathbf{d}H$$

that is, for all $v \in T_z P$ we have

$$\Omega_z(X(z), v) = \mathbf{d}H(z)v.$$

In this case we write X_H for X. Hamilton's equations are defined by

$$\dot{z} = X_H(z).$$

When $P = T^*Q$ is the cotangent bundle of a manifold Q of dimension n, and $\Omega = \sum_{i=1}^n \mathbf{d}q^i \wedge \mathbf{d}p_i$ is the canonical symplectic form on T^*Q with coordinates (q^i, p_i), Hamilton's equations in (q^i, p_i) coordinates are

$$\dot{q}^i = \frac{\partial H}{\partial p_i}, \qquad \dot{p}_i = -\frac{\partial H}{\partial q^i}.$$

One can check that the Hamiltonian H is constant along the flow of the Hamiltonian vector field X_H since

$$X_H[H] = \mathbf{d}H[X_H] = \Omega(X_H, X_H) = 0 \tag{3.1}$$

by the skew symmetry of Ω.

A symplectic form Ω on a manifold P induces the Poisson bracket $\{\,,\} : \mathcal{F}(P) \times \mathcal{F}(P) \to \mathcal{F}(P)$ defined by

$$\{F, G\}(z) = \Omega(X_F(z), X_G(z)) \tag{3.2}$$

where $\mathcal{F}(P)$ is the space of smooth functions on P. The Poisson bracket $\{\,,\}$ satisfies the following properties:

(i) $\{F, G\} = -\{G, F\}$,

(ii) $\{F + G, H\} = \{F, H\} + \{G, H\}$,

(iii) $\{FG, H\} = F\{G, H\} + G\{F, H\}$,

(iv) $\{F, \{G, H\}\} + \{G, \{H, F\}\} + \{H, \{F, G\}\} = 0$

for $F, G, H \in \mathcal{F}(P)$. In particular, (iv) is called the ***Jacobi identity***. The Jacobi identity is related to the closedness of the symplectic form Ω. One can check

$$X_H(F) = \{F, H\}. \tag{3.3}$$

The Poisson bracket induces a skew symmetric nondegenerate $(2,0)$ tensor B on P as follows:

$$B(\mathbf{d}F, \mathbf{d}G) = \{F, G\}$$

for $F, G \in \mathcal{F}(P)$. By the Jacobi identity of the Poisson Bracket, the Poisson tensor $B = B^{ij}\frac{\partial}{\partial z^i} \otimes \frac{\partial}{\partial z^j}$ satisfies

$$B^{li}\frac{\partial B^{jk}}{\partial z^l} + B^{lj}\frac{\partial B^{ki}}{\partial z^l} + B^{lk}\frac{\partial B^{ij}}{\partial z^l} = 0, \tag{3.4}$$

where the summation over l is implied. For example, when $P = T^*Q$ and $\Omega = \sum_{i=1}^n \mathbf{d}q^i \wedge \mathbf{d}p_i$, then the ***canonical Poisson bracket*** on T^*Q is given by

$$\{F, G\} = \sum_{i=1}^n \frac{\partial F}{\partial q^i}\frac{\partial G}{\partial p_i} - \frac{\partial F}{\partial p_i}\frac{\partial G}{\partial q^i},$$

and the ***canonical Poisson tensor*** B_{can} on T^*Q is given by

$$B_{\text{can}} = \sum_{i=1}^n \frac{\partial}{\partial q^i} \otimes \frac{\partial}{\partial p_i} - \frac{\partial}{\partial p_i} \otimes \frac{\partial}{\partial q^i} = \begin{bmatrix} O & I_n \\ -I_n & O \end{bmatrix}$$

with I_n the $n \times n$ identity matrix, $n = \dim Q$.

Above, we started with a pair (P, Ω). But, one can equivalently start with a pair $(P, \{,\})$. Then, (3.3) defines Hamiltonian vector fields X_H, and (3.2) defines the symplectic form Ω since T_zP is spanned by X_{z_i}, $i = 1, \ldots, n$ where $z = (z_1, \ldots, z_n)$ is local coordinates. One can also start with a pair (P, B).

A ***Poisson manifold*** is a pair $(P, \{,\})$, where P is a manifold and the bracket $\{,\} : \mathcal{F}(P) \times \mathcal{F}(P) \to \mathcal{F}(P)$ is a map satisfying the properties, (i) – (iv). Hence, a symplectic manifold is a Poisson manifold with the Poisson bracket (3.2). But a Poisson manifold may not be a symplectic manifold. One example is the Euclidean 3-space $\mathbb{R}^3$ with the following Poisson bracket:

$$\{F, G\}(\mathbf{q}) = \mathbf{q} \cdot (\nabla F \times \nabla G)$$

for $F, G \in \mathcal{F}(\mathbb{R}^3)$, $\mathbf{q} \in \mathbb{R}^3$ and ∇ is the usual gradient in $\mathbb{R}^3$.

For the purpose of the control of Hamiltonian mechanical systems, we relax the Jacobi identity condition which the Poisson bracket satisfies. The reason is that the conservation of the Hamiltonian along its Hamiltonian flow in (3.1) is due to the skew-symmetry of

$\{,\}$, (or, equivalently, B or Ω).

Almost Poisson Structure. In this chapter, we mainly consider the case that the manifold of interest is T^*Q. Following Cannas Da Silva and Weinstein [1999], we define an ***almost Poisson tensor*** B on T^*Q to be a skew-symmetric $(2,0)$-tensor on T^*Q. Hence, B deos not have to satisfy (3.4). Its ***almost Poisson bracket*** $\{\,,\,\} : \mathcal{F}(T^*Q) \times \mathcal{F}(T^*Q) \to \mathcal{F}(T^*Q)$ is defined as

$$\{F, G\} = B(\mathbf{d}F, \mathbf{d}G)$$

for $F, G \in \mathcal{F}(T^*Q)$. Then $\{\,,\,\}$ satisfies the following properties:

(i) $\{F, G\} = -\{G, F\}$,

(ii) $\{F + G, H\} = \{F, H\} + \{G, H\}$,

(iii) $\{FG, H\} = F\{G, H\} + G\{F, H\}$

for $F, G, H \in \mathcal{F}(T^*Q)$. It is not necessary that the bracket satisfy the Jacobi identity:

$$\{F, \{G, H\}\} + \{G, \{H, F\}\} + \{H, \{F, G\}\} = 0$$

for $F, G, H \in \mathcal{F}(T^*Q)$. In coordinates, the almost Poisson tensor B can be written in terms of its action on the coordinate functions:

$$\begin{aligned} B(q,p) &= \{q^i, q^j\} \frac{\partial}{\partial q^i} \otimes \frac{\partial}{\partial q^j} + \{q^i, p_j\} \frac{\partial}{\partial q^i} \otimes \frac{\partial}{\partial p_j} + \{p_i, q^j\} \frac{\partial}{\partial p_i} \otimes \frac{\partial}{\partial q^j} + \{p_i, p_j\} \frac{\partial}{\partial p_i} \otimes \frac{\partial}{\partial p_j} \\ &= \begin{bmatrix} \{q^i, q^j\} & \{q^i, p_j\} \\ \{p_i, q^j\} & \{p_i, p_j\} \end{bmatrix}. \end{aligned}$$

The induced map $B^\sharp : T^*T^*Q \to TT^*Q$ is defined as

$$B(\alpha_z, \beta_z) = \langle \alpha_z, B^\sharp(z)\beta_z \rangle$$

for $\alpha_z, \beta_z \in T_z^*T^*Q$.

It is well known that almost Poisson structures arise in nonholonomic mechanics (see van der Schaft and Maschke [1994], Koon and Marsden [1998], and references therein).

Vertical Lift. Let V be a vector bundle over a manifold Q. The vertical lift of a vector $w_q \in V_q$ along the vector $v_q \in V_q$ is the vector $\mathrm{vlift}_{v_q}(w_q) \in T_{v_q}V$ defined by

$$\mathrm{vlift}_{v_q}(w_q) = \left.\frac{d}{dt}\right|_{t=0} (v_q + t w_q).$$

In coordinates $(q^i, v^i, \delta q^i, \delta v^i)$ on TV,

$$\mathrm{vlift}_{(q^i,v^i)}((q^i, w^i)) = (q^i, v^i, 0, w^i).$$

The vertical lift $\mathrm{vlift}(F)$ of a fiber-preserving map $F : V \to V$ is the section of V to TV defined by

$$(\mathrm{vlift}(F))(v_q) = \mathrm{vlift}_{v_q}(F(v_q)). \tag{3.5}$$

The vertical lift of a subbundle W of V is defined by

$$\mathrm{vlift}(W) = \{\mathrm{vlift}_{v_q}(w_q) \mid v_q \in V_q, w_q \in W_q, q \in Q\}. \tag{3.6}$$

Controlled Hamiltonian Systems. The Hamiltonian analog of a CL system is defined as follows.

Definition 3.1.1. *A **controlled Hamiltonian system (CH system)** is a quadruple (H, B, F, W) where $H : T^*Q \to \mathbb{R}$ is a function called a Hamiltonian, B is an almost Poisson tensor, $F : T^*Q \to T^*Q$ is the fiber-preserving (external force) map, and $W \subset T^*Q$ is a subbundle of T^*Q and called the **control subbundle**.*

Sometimes, W denotes the set of bundle maps from T^*Q to W. As on the Lagrangian side, when we choose a specific control $u : T^*Q \to W$, we call the quadruple (H, B, F, u) a ***closed-loop Hamiltonian system***. The vector field $X_{(H,B,F,u)}$ of the closed-loop system (H, B, F, u) is given by

$$X_{(H,B,F,u)} = B^\sharp \mathbf{d}H + \mathrm{vlift}(F) + \mathrm{vlift}(u). \tag{3.7}$$

We denote the first term on the right-hand side of (3.7) by $X_{(H,B)}$ as follows:

$$X_{(H,B)} = B^\sharp \mathbf{d}H.$$

This is the same as the classical definition of the Hamiltonian vector field. But in the notation $X_{(H,B)}$, we also make clear the almost Poisson tensor being used because we will deal with two CH systems simultaneously.

Remark 3.1.2. *1. The notion of CH systems is essentially the same as that of port-controlled Hamiltonians in van der Schaft [2000]. We improved the foundational setting for the controlled Hamiltonian method. This improved setting will pay off when we consider systems with symmetry and reduction. The reduction will be explained in § 4.*

2. One could develop the method of CH systems on a general manifold, rather than on T^*Q. In such a case, however, it is vague where to introduce forces into the dynamics.

Matching Conditions, CH Equivalence, and CH Inclusion. Suppose we have two controlled Hamiltonian systems (H_i, B_i, F_i, W_i), $i = 1, 2$.

Definition 3.1.3. *We say that these systems satisfy the* ***matching conditions*** *if*

HM-1 : $W_1 = W_2$,

HM-2 : $\mathrm{Im}[(B_1^\sharp \mathbf{d}H_1 + \mathrm{vlift}(F_1)) - (B_2^\sharp \mathbf{d}H_2 + \mathrm{vlift}(F_2))] \subset \mathrm{vlift}(W_1)$.

In addition, we say that two Hamiltonian systems are ***CH-equivalent*** *if* ***HM-1*** *and* ***HM-2*** *hold for the systems. We use the symbol* $\overset{H}{\sim}$ *for this equivalence relation.*

Proposition 3.1.4. *Suppose that the two controlled Hamiltonian systems (H_i, B_i, F_i, W_i), $i = 1, 2$ are CH-equivalent. Then for an arbitrary control law for one system, there exists a control law for the other system such that the closed-loop systems produce the same equations of motion. The explicit relation between the two control laws u_i, $i = 1, 2$ is given by*

$$\mathrm{vlift}(u_2) = \mathrm{vlift}(u_1) + (B_1^\sharp \mathbf{d}H_1 + \mathrm{vlift}(F_1)) - (B_2^\sharp \mathbf{d}H_2 + \mathrm{vlift}(F_2)). \tag{3.8}$$

Proof. Consider the following equation:

$$X_{(H_1,B_1,F_1,u_1)} = X_{(H_2,B_2,F_2,u_2)}.$$

This leads to (3.8). ■

One can also define a partial order, CH inclusion in the class of Hamiltonian systems as follows.

Definition 3.1.5. *A controlled Hamiltonian system (H_1, B_1, F_1, W_1) is said to* ***include*** *another controlled Hamiltonian system (H_2, B_2, F_2, W_2) if the following hold:*

HI-1 : $W_1 \supset W_2$,

HI-2 : $\mathrm{Im}[(B_1^\sharp \mathbf{d}H_1 + \mathrm{vlift}(F_1)) - (B_2^\sharp \mathbf{d}H_2 + \mathrm{vlift}(F_2))] \subset \mathrm{vlift}(W_1)$.

If (H_1, B_1, F_1, W_1) includes (H_2, B_2, F_2, W_2), then for any choice of control $u_2 : T^*Q \to W_2$, there exists a control $u_1 : T^*Q \to W_1$ satisfying

$$\mathrm{vlift}(u_1) = \mathrm{vlift}(u_2) + (B_2^\sharp \mathbf{d}H_2 + \mathrm{vlift}(F_2)) - (B_1^\sharp \mathbf{d}H_1 + \mathrm{vlift}(F_1))$$

such that the two closed-loop systems with these controls produce the same equations of motion.

3.1.2 Simple Controlled Hamiltonian Systems

In most engineering applications, mechanical systems are described by simple Hamiltonians, i.e., Hamiltonians having the kinetic plus potential energy form. Here, we will make a definition of such systems and study their properties. We will show that the almost Poisson structure (i.e., failure of the Jacobi identity of the Poisson bracket) can be understood as gyroscopic forces for simple CH systems (Proposition 3.1.7 and Proposition 3.1.9).

Simple Hamiltonian Systems. The definition of a simple CH system is slightly more subtle than its Lagrangian counterpart.

Definition 3.1.6. *A CH system (H, B, F, W) is called **simple** when the Hamiltonian function has the form kinetic plus potential energy:*

$$H(q,p) = \frac{1}{2}\langle p, m^{-1}(q)p\rangle + V(q), \tag{3.9}$$

where m is a nondegenerate symmetric $(0,2)$-tensor and the almost Poisson tensor B is nondegenerate and has the form:

$$B(q,p) = \begin{bmatrix} O & K(q)^T \\ -K(q) & J(q,p) \end{bmatrix} \tag{3.10}$$

in cotangent coordinates (q,p) on T^*Q, where K, J are $n \times n$ matrices with $n = \dim Q$.

One can check that the statement that B has the form (3.10) is independent of the choice of cotangent bundle coordinates for T^*Q. We call almost Poisson tensors of form (3.10) with K invertible ***simple***.

Decomposition of Simple Almost Poisson Tensors. Now we define a decomposition of simple almost Poisson tensors in the following way. Let B be a given simple almost Poisson tensor. The relation

$$\mathrm{vlift}(\psi_B) = B \circ \Theta \tag{3.11}$$

defines a unique $\psi_B \in \Gamma(\mathrm{Aut}(T^*Q))$, where Θ is the canonical one form on T^*Q,[1] and B is regarded as a linear map $B : T^*T^*Q \to TT^*Q$. Suppose B is given by (3.10) in coordinates. It implies

$$B(q,p) = -K_{ij}(q)\frac{\partial}{\partial p_i} \otimes \frac{\partial}{\partial q^j} + K_{ij}(q)\frac{\partial}{\partial q^j} \otimes \frac{\partial}{\partial p_i} + J_{ij}(q,p)\frac{\partial}{\partial p_i} \otimes \frac{\partial}{\partial p_j}.$$

[1]The canonical one form on T^*Q is given by $p_i dq^i$ in cotangent coordinates (q^i, p_i) for T^*Q.

Then,

$$B \circ \Theta(q,p) = B(q,p)(p_k \mathrm{d}q^k) = K_{ij}(q)p_j \frac{\partial}{\partial p_i}$$

which is the vertical lift of $K_{ij}(q)p_j\mathrm{d}q^i$ at (q,p). Hence, (3.11) defines a unique $\psi_B \in \Gamma(\mathrm{Aut}(T^*Q))$ and its local expression is given by

$$\psi_B(q,p) = (q, K(q)\,p) \tag{3.12}$$

with B given by (3.10) in coordinates.

Given a simple almost Poisson tensor B, we can uniquely decompose B into the two almost Poisson tensors B_r and B_gr as follows:

$$B = B_\mathrm{r} + B_\mathrm{gr}$$

where

$$B_\mathrm{r} = (\psi_B^{-1})^* B_\mathrm{can}; \quad B_\mathrm{gr} = B - (\psi_B^{-1})^* B_\mathrm{can} \tag{3.13}$$

with B_can the canonical Poisson tensor on T^*Q. When B is given by (3.10) in coordinates, we have the following coordinate expressions:

$$B_\mathrm{r}(q,p) = \begin{bmatrix} O & K(q)^T \\ -K(q) & C_K(q,p) \end{bmatrix} \tag{3.14}$$

$$B_\mathrm{gr}(q,p) = \begin{bmatrix} O & O \\ O & J(q,p) - C_K(q,p) \end{bmatrix} \tag{3.15}$$

where

$$\begin{aligned}(C_K)_{ij}(q,p) &= -\left\langle K(q)^{-1}p, [(K(q)^T)_i, (K(q)^T)_j] \right\rangle \\ &= \left(\frac{\partial K_{is}(q)}{\partial q^r} K_{jr}(q) - K_{ir}(q) \frac{\partial K_{js}(q)}{\partial q^r} \right) K^{sl}(q) p_l \end{aligned} \tag{3.16}$$

where $(K(q)^T)_i$ is the i-th column of the matrix $K(q)^T$ and $[\,,\,]$ is the Lie bracket. The formula (3.16) is essentially the same as the equation (19) in van der Schaft and Maschke [1994]. By (3.10), (3.12) and (3.14), we have

$$\psi_B = \psi_{B_\mathrm{r}}. \tag{3.17}$$

Notice that the Poisson tensor B_r satisfies the Jacobi-identity because it is a pull-back of the canonical Poisson bracket.

Construction of Gyroscopic Forces. Given a simple CH system $(H, B = B_{\mathrm{r}} + B_{\mathrm{gr}}, F, W)$, the almost Poisson tensor B_{gr} and the Hamiltonian H defines a ***gyroscopic force*** $F_{\mathrm{gr}} : T^*Q \to T^*Q$ by the following relation:

$$\mathrm{vlift}(F_{\mathrm{gr}}) = (B_{\mathrm{gr}})^\sharp \mathrm{d}H. \tag{3.18}$$

By (3.15), in coordinates,

$$(B_{\mathrm{gr}})^\sharp \mathrm{d}H(q,p) = \left((J - C_K)_{ij}(q,p)\frac{\partial H(q,p)}{\partial p_j}\right)\frac{\partial}{\partial p_i}$$

which is the vertical lift of $\left((J - C_K)_{ij}(q,p)\frac{\partial H(q,p)}{\partial p_j}\right)\mathrm{d}q^i$ at (q,p). Hence, (3.18) defines the unique force $F_{\mathrm{gr}} : T^*Q \to T^*Q$, which is locally written as

$$F_{\mathrm{gr}}(q,p) = \left(q, (J - C_K)_{ij}(q,p)\frac{\partial H(q,p)}{\partial p_j}\right).$$

The reason we call F_{gr} the gyroscopic force is that it does not change the Hamiltonian H in the following sense

$$\mathrm{vlift}(F_{\mathrm{gr}})[H] = \mathrm{d}H((B_{\mathrm{gr}})^\sharp \mathrm{d}H) = B_{\mathrm{gr}}(\mathrm{d}H, \mathrm{d}H) = 0$$

due to the skew symmetry of B_{gr}. The dynamics with gyroscopic forces still conserve energy.

This decomposition of simple almost Poisson tensors simplifies the class of simple CH systems under the CH-equivalence relation. Suppose that we are given a simple CH system $(H, B = B_{\mathrm{r}} + B_{\mathrm{gr}}, F, W)$. Then (3.18) implies

$$B^\sharp \mathbf{d}H + \mathrm{vlift}(F) = (B_{\mathrm{r}})^\sharp \mathbf{d}H + \mathrm{vlift}(F_{\mathrm{gr}} + F).$$

Therefore the simple CH system $(H, B = B_{\mathrm{r}} + B_{\mathrm{gr}}, F, W)$ is CH-equivalent to the simple CH system $(H, B_{\mathrm{r}}, F_{\mathrm{gr}} + F, W)$, where F_{gr} is given by (3.18). By (3.13) and (3.17),

$$\psi_{B_{\mathrm{r}}}^* B_{\mathrm{r}} = \psi_B^* B_{\mathrm{r}} = B_{\mathrm{can}}. \tag{3.19}$$

This proves the following result.

Proposition 3.1.7. *A given simple CH system $(H, B = B_{\mathrm{r}} + B_{\mathrm{gr}}, F, W)$ is CH-equivalent to the CH system $(H, B_{\mathrm{r}}, F_{\mathrm{gr}} + F, W)$, where $B = B_{\mathrm{r}} + B_{\mathrm{gr}}$ is the decomposition of B into the regular part and the gyroscopic part and $F_{\mathrm{gr}} : T^*Q \to T^*Q$ is determined by the relation $\mathrm{vlift}(F_{\mathrm{gr}}) = B_{\mathrm{gr}}^\sharp \mathrm{d}H$. In particular, B_r satisfies (3.19).*

A consequence is

Corollary 3.1.8. *An arbitrary simple CH system is CH-equivalent to a simple CH system (H, B, F, W) with the Poisson tensor B satisfying the Jacobi identity and $\psi_B^* B = B_{\text{can}}$. Equivalently, one can say $B = \phi^* B_{\text{can}}$ for some $\phi \in \Gamma(\text{Aut}(T^*Q))$. In coordinates, B is always of the form*

$$B(q,p) = \begin{bmatrix} O & K(q)^T \\ -K(q) & C_K(q,p) \end{bmatrix}$$

with C_K in (3.16) when ϕ (or, ψ_B^{-1}) is in coordinates given by

$$\phi(q,p) = (q, K(q)^{-1} p).$$

Proof. A direct consequence of Proposition 3.1.7 and (3.14). ■

We now consider the opposite direction, i.e., one may address the question "can the gyroscopic force be incorporated into the Poisson tensor for a simple CH system?". The answer is yes. Let us consider a simple CH system (H, B, F_{gr}, W) with $H = \frac{1}{2}\langle p, m^{-1}p\rangle + V(q)$, and F_{gr} a gyroscopic force. By the definition of the gyroscopic force, $F_{\text{gr}} = (F_{\text{gr}})_i \mathbf{d}q^i$ satisfies

$$0 = \text{vlift}(F_{\text{gr}})[H] = (F_{\text{gr}})_i m^{ij} p_j,$$

where $m^{-1} = (m^{ij})$. Hence, $(F_{\text{gr}})_i$ should be of the form

$$(F_{\text{gr}})_i(q,p) = p_l S^{lk}(q,p) m_{ki}(q), \quad S^{lk} = -S^{kl}.$$

We have proved the following:

Proposition 3.1.9. *Given a simple CH system $(H, B, F + F_{\text{gr}}, W)$ with F_{gr} a gyroscopic force, the following holds:*

1. *F_{gr} is always written of the form:*

$$F_{\text{gr}}(q,p) = p_l S^{lk}(q,p) m_{ki}(q) \mathbf{d}q^i, \quad S^{lk} = -S^{kl}$$

for some functions $S^{ij}(q,p)$ where m_{ij} is the mass matrix of H.

2. *We have*

$$(H, B, F + F_{\text{gr}}, W) \overset{H}{\sim} (H, B + \tilde{B}, F, W),$$

where

$$\tilde{B} = m_{ik} S^{kl} m_{lj} \frac{\partial}{\partial p_i} \otimes \frac{\partial}{\partial p_j}.$$

Proof. We have only to check that $B + \tilde{B}$ is a simple almost Poisson tensor, which is readily done. ■

Remark 3.1.10. *Proposition 3.1.7 and Proposition 3.1.9 show that the almost Poisson structure (i.e., failure of the Jacobi identity of the Poisson bracket) can be understood as a gyroscopic force for simple CH systems and vice versa.*

We give an example of Proposition 3.1.9. Consider a simple model for the symmetric flight of a plane where there are only the gravity and the lift acting on the plane (See Figure 3.1). Let $\mathbb{R}^2$ be the configuration space and (x, z) the coordinates of the plane. We

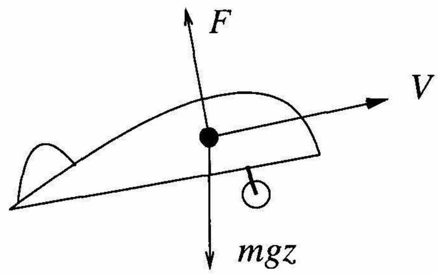

Figure 3.1: Motion in the symmetric plane.

assume that the lift F is of the form

$$F = f(v)(-\dot{z}, \dot{x}),$$

where $v = \sqrt{\dot{x}^2 + \dot{z}^2}$ and $f(\cdot)$ is a real-valued function. The lift is a gyroscopic force because it does not do any works as it is always perpendicular to the velocity $(\dot{x}, \dot{y})$. The dynamics of the plane are given by

$$m\ddot{x} = -f(v)\dot{z}, \qquad m\ddot{z} = f(v)\dot{x} - mg, \tag{3.20}$$

where m is the mass of the plane and g is the gravitation constant. Let $(x, z, p_x, p_z) := (x, z, m\dot{x}, m\dot{z})$ be the coordinates of the cotangent (or, momentum) space $T^*\mathbb{R}^2 = \mathbb{R}^2 \times \mathbb{R}^2$. Then, the plane can be described as a CH system $(H, B_{\text{can}}, F, 0)$ on $T^*\mathbb{R}^2$ with the Hamiltonian

$$H(x, z, p_x, p_z) = \frac{1}{2m}(p_x^2 + p_z^2) + mgz,$$

and the canonical Poisson bracket B_{can} on $T^*\mathbb{R}^2$. The dynamics (3.20) can be equivalently written as

$$\frac{d}{dt}\begin{bmatrix} x \\ z \\ p_x \\ p_z \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix} \mathbf{d}H + \begin{bmatrix} 0 \\ 0 \\ -f(\frac{p}{m})\frac{p_z}{m} \\ f(\frac{p}{m})\frac{p_x}{m} \end{bmatrix} \tag{3.21}$$

with $p = \sqrt{p_x^2 + p_z^2}$ and $\mathbf{d}H = (0, mg, p_x/m, p_z/m)^T$. One can check that the equation of motion in (3.21) is the same as

$$\frac{d}{dt}\begin{bmatrix} x \\ z \\ p_x \\ p_z \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & -f(\frac{p}{m}) \\ 0 & -1 & f(\frac{p}{m}) & 0 \end{bmatrix} \mathbf{d}H. \tag{3.22}$$

Equation (3.22) is the equation of the motion of the CH system $(H, B, 0, 0)$ with

$$B = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & -f(\frac{p}{m}) \\ 0 & -1 & f(\frac{p}{m}) & 0 \end{bmatrix}.$$

Hence,

$$(H, B_{\text{can}}, F, 0) \overset{H}{\sim} (H, B, 0, 0).$$

One can check that this example agrees with Proposition 3.1.9.

Pull-back Systems. The concept of pull-back systems will be useful in § 3.2 for showing the equivalence of the CL system method and the CH system method. This notion is a technical device that is needed for the proofs later and can be omitted on first reading. Consider a CH system (H, B, F, W) and $\phi \in \Gamma(\text{Aut}(T^*Q))$. Then, the pull-back system $\phi^*(H, B, F, W)$ is defined to be the associated CH system $(\phi^*H, \phi^*B, \phi^*F, \phi^*W)$, where

$$\phi^*H = H \circ \phi \quad \text{and} \quad (\phi^*B)_z(\mathbf{d}G_1, \mathbf{d}G_2) = B_{\psi(z)}((\phi^{-1})^*\mathbf{d}G_1, (\phi^{-1})^*\mathbf{d}G_2)$$

for $G_1, G_2 \in \mathcal{F}(T^*Q)$, and $\phi^*F := \phi^{-1} \circ F \circ \phi$. Here, the pull-back notation in ϕ^*F should be regarded as an action of $\Gamma(\text{Aut}(T^*Q))$ on the set of fiber-preserving maps on T^*Q. Notice also that in this thesis, $\phi^*W = \phi^{-1}(W)$ by definition. The notation ϕ^*W should not be confused with the standard notation of pull-back bundles. When we regard W as the set $\{u : T^*Q \to W\}$, then ϕ^*W reads $\phi^*W = \{\phi^*u = \phi^{-1} \circ u \circ \phi \mid u \in W\}$. Hence, we write $\phi^* W$ as $\phi^{-1} W \circ \phi$ to respect both interpretations. We claim

$$\phi^* X_{(H,B,F,u)} = X_{\phi^*(H,B,F,u)}.$$

It is well known (or a straightforward computation) that

$$(\phi^* B)^\sharp \mathbf{d}(\phi^* H) = \phi^*(B^\sharp \mathbf{d} H).$$

We have only to show $\phi^*(\mathrm{vlift}(F)) = \mathrm{vlift}(\phi^* F)$ where one should be careful that pull-back notation in the left hand side is the usual pull-back of a vector field by a diffeomorphism, ϕ, and the pull-back notation on the right-hand side should be understood as $\phi^{-1} \circ F \circ \phi$ as we mentioned before. Indeed, for $w \in T^*Q$, we have

$$\begin{aligned}
(\phi^*(\mathrm{vlift}(F)))(w) &= T_{\phi(w)}\phi^{-1} \cdot \mathrm{vlift}(F)(\phi(w)) \\
&= T_{\phi(w)}\phi^{-1} \left.\frac{d}{ds}\right|_{s=0} (\phi(w) + sF \circ \phi(w)) \\
&= \left.\frac{d}{ds}\right|_{s=0} \phi^{-1}(\phi(w) + sF \circ \phi(w)) \\
&= \left.\frac{d}{ds}\right|_{s=0} (w + s(\phi^* F)(w)) \\
&= \mathrm{vlift}(\phi^* F)(w).
\end{aligned}$$

The same relation holds for u. One can readily show the following:

Proposition 3.1.11. *Let $\phi \in \Gamma(\mathrm{Aut}(T^*Q))$. Then the following hold:*

1. *The pull-back system of a simple CH system via ϕ is also simple.*
2. *Two CH systems (H_1, B_1, F_1, W_1) and (H_2, B_2, F_2, W_2) are CH-equivalent if and only if the corresponding pull-back systems $\phi^*(H_1, B_1, F_1, W_1)$ and $\phi^*(H_2, B_2, F_2, W_2)$ are CH-equivalent.*

In particular, it is useful to have a coordinate expression for $\phi^* B$ when B satisfies $\psi_B^* B = B_{\mathrm{can}}$ for $\psi \in \Gamma(\mathrm{Aut}(T^*Q))$. The almost Poisson tensor B is written in coordinates as in (3.10). Consider $\phi \in \Gamma(\mathrm{Aut}(T^*Q))$ with the local coordinates expression $\phi(q,p) = (q, D(q)^{-1} p)$. Then the pull-back tensor $\phi^* B$ is expressed in coordinates as

$$(\phi^* B)(q,p) = \begin{bmatrix} O & (D(q)K(q))^T \\ -D(q)K(q) & C_{DK} \end{bmatrix}$$

since

$$\phi^* B = ((\psi_B)^{-1} \circ \phi)^* (\psi_B)^* B = ((\psi_B)^{-1} \circ \phi)^* B_{\mathrm{can}}. \tag{3.23}$$

Here $\psi_B^{-1} \circ \phi(q,p) = (q, (D(q)K(q))^{-1}p)$ and we use the formula in Corollary 3.1.8. The equation (3.23) implies that

$$\psi_{\phi^* B} = \phi^{-1} \circ \psi_B \tag{3.24}$$

and $(\psi_{\phi^* B})^* B = B_{\text{can}}$. This proves that a simple CH system (H, B, F, W) with $\psi_B^* B = B_{\text{can}}$ is pulled back by $\phi \in \Gamma(\text{Aut}(T^*Q))$ to the simple CH system $\phi^*(H, B, F, W)$ satisfying

$$(\psi_{\phi^* B})^*(\phi^* B) = B_{\text{can}}; \qquad \psi_{\phi^* B} = \phi^{-1} \circ \psi_B.$$

3.1.3 Control Synthesis via a CH System

We apply the CH method to the stabilization problem. This procedure is almost identical with that for the CL method in § 2.2.1. The application of the CH method has been well known as the interconnection and damping assignment passivity-based control (IDA-PBC) (Ortega and Spong [2000], Ortega, Spong, Gómez-Estern, and Blankenstein [2001], van der Schaft [2000]).

We want to design a control law to asymptotically stabilize an equilibrium $(q_e, 0) \in TQ$ of a given simple CH system $(H_1, B_1, F_1 = 0, W_1)$ using its energy or Hamiltonian H_1 as a Lyapunov function. Usually, the equilibrium $(q_e, 0)$ is not a minimum of the energy H_1, which prevents us from directly using the energy H_1 as a Lyapunov function.

Here is the procedure of applying CH systems:

1. find a simple CH system (H_2, B_2, F_2, W_2) CH-equivalent to $(H_1, B_1, F_1 = 0, W_1)$ where the Hamiltonian H_2 has a strict minimum at the equilibrium $(q_e, 0)$ and F_2 has the form of a gyroscopic force

2. take a dissipative control u_2 for (H_2, B_2, F_2, W_2)

3. check the asymptotic stability of the equilibrium $(q_e, 0)$ in the closed-loop system (H_2, B_2, F_2, u_2) using its Hamiltonian H_2 as Lyapunov function.

4. if the equilibrium $(q_e, 0)$ is asymptotically stable in the closed-loop dynamics of (H_2, B_2, F_2, u_2), so is it in the closed-loop system $(H_1, B_1, 0, u_1)$ with the control u_1 derived from (3.8).

In practice, step 1 is subdivided into:

1a. find a parameterized family of simple CH systems (H_2, B_2, F_2, W_2), with some free parameters, which are CH-equivalent to $(H_1, B_1, F_1 = 0, W_1)$

1b. choose a set of appropriate parameters in order for the Hamiltonian H_2 to have a strict minimum at the equilibrium $(q_e, 0)$ and in order for the force F_2 to be a gyroscopic force.

Recall that a dissipative force F_{diss} for a simple CH (H, B, F, u) system can be written as

$$F_{\text{diss}}(q,p) = -p_i D^{ij}(q,p) m_{jk} \mathbf{d}q^k$$

with D^{ij} a positive-semidefinite symmetric matrix, and a gyroscopic force F_{gr} as

$$F_{\text{gr}}(q,p) = p_i S^{ij}(q,p) m_{jk} \mathbf{d}q^k$$

with $S^{ij} = -S^{ji}$, where m_{ij} is the mass matrix of H.

Remark 3.1.12. *1. By Proposition 3.1.9, without loss of generality, one can just use $F_2 = 0$ in Step 1b.*

2. When a given CH system (H_1, B_1, F_1, W_1) has a non-zero external force $F_1 \neq 0$, one needs to modify the above procedure because this additional force may have some effects on the change of the Hamiltonian.

3. Here, we do not include any examples of applications of the CH method to stabilization problems because several examples have already been worked out in several papers such as Ortega, Spong, Gómez-Estern, and Blankenstein [2001], Ortega and Spong [2000], van der Schaft [2000], and Woolsey and Leonard [1999].

3.2 Equivalence of CL Systems and CH Systems: Simple Mechanical Systems

The goal of this section is to show the equivalence of the method of simple CL systems and that of simple CH systems. A more detailed statement is contained in Theorem 3.2.1 and Corollary 3.2.2. Hence, one can apply either method to control problems. However, the complexity of relevant computation can be different just as a good choice of coordinates can simplify a partial differential equation.

First, we review the Legendre transformations and then tackle the problem of the equivalence of the method of CL systems and that of CH systems.

3.2.1 Legendre Tansformations

Frequently in mechanics, a Hamiltonian system on T^*Q induces a Hamiltonian vector field through a canonical symplectic structure (or, canonical Poisson structure) on T^*Q before any reduction processes. This is because a Hamiltonian system on T^*Q often comes from a Lagrangian system on TQ via a Legendre transformation associated to a given Lagrangian function. Hence, if there is more than one Lagrangian function, there can be multiple transformations between TQ and T^*Q. We will carefully deal with this issue as well.

Fiber Derivatives. We review the definition of the fiber-derivative of a map $f : V \to \mathbb{R}$ with V a vector bundle over a manifold M. Define the fiber derivative $\mathbb{F}f : V \to V^*$ of f as follows:

$$\mathbb{F}f(v_m) \cdot w_m = \left.\frac{d}{dt}\right|_{t=0} f(v_m + tw_m)$$

for $v_m, w_m \in V$. In coordinates, $\mathbb{F}f$ is given by

$$\mathbb{F}f(m, v^i) = \left(m, \frac{\partial f}{\partial v^i}\right).$$

When $\det\left(\frac{\partial^2 f}{\partial v^i \partial v^j}\right) \neq 0$, $\mathbb{F}f$ is locally invertible. We say f is ***regular*** if $\det\left(\frac{\partial^2 f}{\partial v^i \partial v^j}(v)\right) \neq 0$ for all $v \in V$ and ***hyperregular*** if $\mathbb{F}f$ is globally invertible.

Legendre Transformations. Given a Lagrangian L on TQ and a Hamiltonian H on T^*Q, we call their fiber derivatives $\mathbb{F}L : TQ \to T^*Q$ and $\mathbb{F}H : T^*Q \to TQ$ the Legendre transformation and the inverse Legendre transformation, respectively, where the use of word *inverse* will be justified later. Here, we always assume that all Lagrangians and Hamitonians are regular so that $\mathbb{F}L$ and $\mathbb{F}H$ are locally invertible. When a Lagrangian (or, a Hamiltonian) is simple, then it is automatically hyperregular.

It is well known that a given Lagrangian system (L, F^L, W^L) is transformed by the Legendre transformation $\mathbb{F}L$ to the Hamiltonian system $(H, B_{\mathrm{can}}, F^H, W^H)$ (see, for example, Marsden and Ratiu [1999]) where

$$H(\alpha) = \langle \alpha, \mathbb{F}L^{-1}(\alpha)\rangle - L \circ \mathbb{F}L^{-1}(\alpha) \quad \text{for } \alpha \in T^*Q, \tag{3.25}$$

$$F^H = F^L \circ \mathbb{F}L^{-1}, \tag{3.26}$$

$$W^H = W^L \circ \mathbb{F}L^{-1}, \tag{3.27}$$

where $W^L \circ \mathbb{F}L^{-1}$ is understood as W^L as a subbundle of T^*Q, and also understood as the set $\{u \circ \mathbb{F}L^{-1} | u : TQ \to W^L\}$ when we regard $W^L \circ \mathbb{F}L^{-1}$ as a set of fiber-preserving maps from T^*Q to W. Namely, the Euler-Lagrange equation

$$\mathcal{EL}(L) = F^L + u^L$$

with $u^L : TQ \to W^L$ is *equivalent* to the CH vector field

$$X = B^\sharp_{\mathrm{can}} \mathbf{d}H + \mathrm{vlift}(F^L \circ \mathbb{F}L^{-1}) + \mathrm{vlift}(u^L \circ \mathbb{F}L^{-1}).$$

Let us now suppose that we are given a Hamiltonian function H on T^*Q. It is well known (or easy to check) that a given CH system $(H, B_{\text{can}}, F^H, W^H)$ is transformed by the inverse Legendre transformation $\mathbb{F}H$ to the CL system $(L, F^L = F^H \circ \mathbb{F}H^{-1}, W^L = W^H \circ \mathbb{F}H^{-1})$ where $L(v) = \langle v, \mathbb{F}H^{-1}(v)\rangle - H \circ \mathbb{F}H^{-1}(v)$ for $v \in TQ$. The CH vector field $X = B^\sharp_{\text{can}} \mathbf{d}H + \text{vlift}(F^H) + \text{vlift}(u^H)$ is equivalent to the Euler-Lagrange equation $\mathcal{EL}(L) = F^L + u^H \circ \mathbb{F}H^{-1}$.

Suppose that a given CL system (L, F^L, W^L) is transformed by the Legendre transformation $\mathbb{F}L$ to the CH system $(H, B_{\text{can}}, F^H, W^H)$. Then, $(H, B_{\text{can}}, F^H, W^H)$ is transformed back to (L, F^L, W^L) by the inverse Legendre transformation $\mathbb{F}H$ since $\mathbb{F}H = (\mathbb{F}L)^{-1}$ in this case by Proposition 7.4.2 of Marsden and Ratiu [1999]. One can also start this argument from the Hamiltonian system.

3.2.2 CH Equivalence Proves CL Equivalence

We first show that the matching conditions of simple CL systems can be derived from those of simple CH systems; the computation involved in this direction is simpler than that involved in the opposite direction. However, the computation carried out here will also be used in § 3.2.3. A special case of the result in this section (§ 3.2.2) was dealt with in Ortega, Spong, Gómez-Estern, and Blankenstein [2001] and Blankenstein, Ortega, and van der Schaft [2001], where only CL systems of the form $(L, 0, W)$ were considered, i.e., external forces were not considered.

Suppose we have two simple CL systems (L_1, F_1^L, W_1^L) and (L_2, F_2^L, W_2^L) with $L_1(q, \dot q) = \frac{1}{2} m_1(\dot q, \dot q) - V_1(q)$ and $L_2(q, \dot q) = \frac{1}{2} m_2(\dot q, \dot q) - V_2(q)$. They define two Legendre transformations $\mathbb{F}L_1, \mathbb{F}L_2 : TQ \to T^*Q$ as follows

$$(q, p) = \mathbb{F}L_1(q, \dot q) = (q, m_1(q)\dot q), \tag{3.28}$$

$$(q, \tilde p) = \mathbb{F}L_2(q, \dot q) = (q, m_2(q)\dot q). \tag{3.29}$$

The CL system (L_1, F_1^L, W_1^L) is transformed via $\mathbb{F}L_1$ to the CH system

$$(H_1, B_1 = B_{\text{can}}, F_1^H = F_1^L \circ \mathbb{F}L^{-1}, W_1^H = W_1^L \circ \mathbb{F}L^{-1}),$$

and the second CL system (L_2, F_2^L, W_2^L) is transformed, via $\mathbb{F}L_2$, to

$$(\tilde H_2, \tilde B_2 = B_{\text{can}}, \tilde F_2^H = F_2^L \circ \mathbb{F}L_2^{-1}, \tilde W_2^H = W_2^L \circ \mathbb{F}L_2^{-1}),$$

$$\begin{tikzcd} T^*Q \arrow[rr, "m_2 m_1^{-1}"] & & T^*Q \\ & TQ \arrow[ul, "m_1"] \arrow[ur, "m_2"'] & \end{tikzcd}$$

Figure 3.2: Diagram of Legendre transformations.

where

$$\begin{aligned} H_1(q,p) &= \frac{1}{2}\langle p, m_1(q)^{-1}p\rangle + V_1(q), \\ \tilde{H}_2(q,\tilde{p}) &= \frac{1}{2}\langle \tilde{p}, m_2(q)^{-1}\tilde{p}\rangle + V_2(q). \end{aligned}$$

We now would like to show that checking the CL equivalence of (L_1, F_1^L, W_1^L) and (L_2, F_2^L, W_2^L) is the same as checking the CH equivalence of their transformed Hamiltonian systems. Thereby, we show that CH equivalence proves CL-equivalence. Since the two Legendre transformations in (3.28) and (3.29) are not the same in general, we need to pull-back the system $(\tilde{H}_2, \tilde{B}_2, \tilde{F}_2^H, \tilde{W}_2^H)$ via $\mathbb{F}L_2 \circ \mathbb{F}L_1^{-1} = m_2 \circ m_1^{-1} \in \Gamma(\mathrm{Aut}(T^*Q))$, as in the commutative diagram in Figure 3.2.

Let $(H_2, B_2, F_2^H, W_2^H) = (\mathbb{F}L_2 \circ \mathbb{F}L_1^{-1})^*(\tilde{H}_2, \tilde{B}_2, \tilde{F}_2^H, \tilde{W}_2^H)$ where one computes

$$\begin{aligned} H_2(q,p) &= \frac{1}{2}\langle p, m_1(q)^{-1}m_2(q)m_1(q)^{-1}p\rangle + V_2(q) \\ B_2(q,p) &= \begin{bmatrix} O & (m_1(q)m_2(q)^{-1})^T \\ -m_1(q)m_2(q)^{-1} & C_{m_1m_2^{-1}}(q,p) \end{bmatrix} \\ W_2^H &= (\mathbb{F}L_2 \circ \mathbb{F}L_1^{-1})^*(W_2^L \circ \mathbb{F}L_2^{-1}) \\ &= (\mathbb{F}L_2 \circ \mathbb{F}L_1^{-1})^{-1}(W_2^L \circ \mathbb{F}L_2^{-1}) \circ \mathbb{F}L_2 \circ \mathbb{F}L_1^{-1} \\ &= (\mathbb{F}L_1 \circ \mathbb{F}L_2^{-1})(W_2^L \circ \mathbb{F}L_1^{-1}). \end{aligned}$$

We will now show the following

$$\begin{aligned} &(H_1, B_1, F_1^H, W_1^H) \overset{H}{\sim} (H_2, B_2, F_2^H, W_2^H) \\ &\quad \Longleftrightarrow (L_1, F_1^L, W_1^L) \overset{L}{\sim} (L_2, F_2^L, W_2^L). \end{aligned} \tag{3.30}$$

First, **HM-1** reads

$$\begin{aligned} W_1^H = W_2^H \quad &\Longleftrightarrow \quad W_1^L \circ \mathbb{F}L_1^{-1} = m_1 m_2^{-1}(W_2^L \circ \mathbb{F}L_1^{-1}) \\ &\Longleftrightarrow \quad W_1^L = m_1 m_2^{-1} W_2^L \end{aligned}$$

whose right hand side is exactly **ELM-1**. Hence **HM-1** for (H_i, B_i, F_i^H, W_i^H), $i = 1, 2$ is equivalent to **ELM-1** for (L_1, F_1^L, W_1^L), $i = 1, 2$. Second, since $u_1^H, u_2^H \in W_1^H = W_2^H$, **HM-2** can be *equivalently* written in coordinates as

$$\left(\begin{bmatrix} \dot{q} \\ \dot{p} \end{bmatrix}_{H_1} - \begin{bmatrix} \dot{q} \\ \dot{p} \end{bmatrix}_{H_2} \right) \in \operatorname{vlift}(W_1^H(q)) \simeq O \oplus W_1^H(q) \tag{3.31}$$

where

$$\begin{bmatrix} \dot{q} \\ \dot{p} \end{bmatrix}_{H_i} = X_{(H_i, B_i, F_i^H, u_i^H)} = B_i^\sharp \mathrm{d}H_i + \operatorname{vlift}(F_i^H) + \operatorname{vlift}(u_i^H)$$

for each $i = 1, 2$ where the subscript H_i denotes the CH system (H_i, B_i, F_i^H, W_i^H) for simplicity.

Since $W_1^H = W_2^H = W_1^L \circ \mathbb{F}L_1^{-1} = m_1 m_2^{-1} W_2^L \circ \mathbb{F}L_2^{-1}$ under **HM-1**, we can write the controls $u_i^H \in W_1^H$ as

$$\begin{aligned} u_1^H &= u_1^L \circ \mathbb{F}L_1^{-1} \\ u_2^H &= m_1 m_2^{-1} u_2^L \circ \mathbb{F}L_2^{-1} \end{aligned}$$

for $u_i^L \in W_i^L$, which can be considered via the Legendre transformations $\mathbb{F}L_1$ and $\mathbb{F}L_2$ as the controls for (L_i, F_i^L, W_i^L) for $i = 1, 2$, respectively. One can readily check

$$\dot{q}_{H_1} - \dot{q}_{H_2} = m_1(q)^{-1} p - m_1(q)^{-1} p = 0.$$

The equation for $\dot{p}_{H_1}$ can be written in terms of $(q, \dot{q})$ as

$$\dot{p}_{H_1} = \frac{d}{dt}\frac{\partial L_1}{\partial \dot{q}} = \frac{\partial L_1}{\partial q} + F_1^L + u_1^L.$$

Recall from (3.28) and (3.29) that $p = m_1(q) m_2(q)^{-1} \tilde{p}$. The equation $\dot{p}_{H_2}$ can be written in terms of $(q, \dot{q})$ as follows:

$$\begin{aligned}
\dot{p}_{H_2} &= \frac{d}{dt}(m_1 m_2^{-1}\tilde{p}) \\
&= (\mathbf{d}(m_1 m_2^{-1})[\dot{q}])\tilde{p} + m_1 m_2^{-1}\frac{d}{dt}\tilde{p}_{\tilde{H}_2} \\
&= (\mathbf{d}(m_1 m_2^{-1})[\dot{q}])m_2\dot{q} + m_1 m_2^{-1}\frac{d}{dt}\frac{\partial L_2}{\partial \dot{q}} \\
&= (\mathbf{d}m_1[\dot{q}])\dot{q} - m_1 m_2^{-1}(\mathbf{d}m_2[\dot{q}])\dot{q} + m_1 m_2^{-1}\left(\frac{\partial L_2}{\partial q} + F_2^L + u_2^L\right) \\
&= (\mathbf{d}m_1[\dot{q}])\dot{q} - m_1 m_2^{-1}\left((\mathbf{d}m_2[\dot{q}])\dot{q} - \frac{\partial L_2}{\partial q} - F_2^L - u_2^L\right).
\end{aligned}$$

Hence,

$$\begin{aligned}
\dot{p}_{H_2} - \dot{p}_{H_1} &= \left((\mathbf{d}m_1[\dot{q}])\dot{q} - \frac{\partial L_1}{\partial q} - F_1^L - u_1^L\right) \\
&\quad - m_1 m_2^{-1}\left((\mathbf{d}m_2[\dot{q}])\dot{q} - \frac{\partial L_2}{\partial q} - F_2^L - u_2^L\right) \\
&= \left(m_1\ddot{q} + (\mathbf{d}m_1[\dot{q}])\dot{q} - \frac{\partial L_1}{\partial q} - F_1^L - u_1^L\right) \\
&\quad - m_1 m_2^{-1}\left(m_2\ddot{q} + (\mathbf{d}m_2[\dot{q}])\dot{q} - \frac{\partial L_2}{\partial q} - F_2^L - u_2^L\right) \\
&= (\mathcal{EL}(L_1) - F_1^L) - m_1 m_2^{-1}(\mathcal{EL}(L_2) - F_2^L) - u_1^L + m_1 m_2^{-1}u_2^L. \qquad (3.32)
\end{aligned}$$

Therefore, (3.31) holds if and only if

$$\mathrm{Im}[(\mathcal{EL}(L_1) - F_1^L) - m_1 m_2^{-1}(\mathcal{EL}(L_2) - F_2^L)] \subset W_2^L$$

since $u_1^L, m_1 m_2^{-1}u_2^L \in W_1^L = m_1 m_2^{-1}W_2^L$. Therefore, we have shown (3.30). Finally, one can easily show from (3.31) and (3.32) that (2.12) on the controls u_i^L's is equivalent to (3.8) on the controls u_i^H's.

Let us make a remark on an alternative way to compare **HM-2** and **ELM-2**. One can show by a brute-force coordinate computation that

$$\begin{aligned}
&[(B_2^\sharp \mathbf{d}H_2 + \mathrm{vlift}(F_2^H)) - (B_1^\sharp \mathbf{d}H_1 + \mathrm{vlift}(F_1^H))] \\
&\qquad \simeq O \oplus \left[(\mathcal{EL}(L_1) - F_1^L) - m_1 m_2^{-1}(\mathcal{EL}(L_2) - F_2^L)\right].
\end{aligned}$$

This computation is very complicated and it does not directly lead to the equivalence of (2.12) and (3.8). This is why we did not choose this brute-force computational method here.

3.2.3 CL Equivalence Proves CH Equivalence

We now show that the Hamiltonian matching conditions of simple CH systems can be derived from those of simple CL systems. Consider two simple CH systems (H_1, B_1, F_1^H, W_1^H) and (H_2, B_2, F_2^H, W_2^H) with $H_i(q,p) = \frac{1}{2}\langle p, m_i^{-1}(q)p\rangle + V_i(q)$ for $i = 1,2$. By Proposition 3.1.7 and Proposition 3.1.11, without loss of generality, we may assume that

$$B_1 = B_{\text{can}}; \qquad \psi_{B_2}^* B_2 = B_{\text{can}}.$$

In coordinates, we write B_2 and ψ_{B_2} as follows:

$$B_2(q,p) = \begin{bmatrix} O & K_2(q)^T \\ -K_2(q) & C_{K_2}(q,p) \end{bmatrix}$$

and

$$(q,p) = \psi_{B_2}(q,\tilde{p}) = (q, K_2(q)\tilde{p}).$$

Consider the pull-back system

$$(\tilde{H}_2, \tilde{B}_2 = B_{\text{can}}, \tilde{F}_2^H, \tilde{W}_2^H) := (\psi_{B_2})^*(H_2, B_2, F_2^H, W_2^H)$$

with

$$H_2(q,\tilde{p}) = \frac{1}{2}\langle \tilde{p}, \tilde{m}_2^{-1}(q)\tilde{p}\rangle + V_2(q); \quad \tilde{m}_2 := \psi_{B_2}^{-1} m_2 (\psi_{B_2}^*)^{-1}, \tag{3.33}$$

where $\psi_{B_2}^* \in \Gamma(\text{Aut}(TQ))$ is the dual of $\psi_{B_2} \in \Gamma(\text{Aut}(T^*Q))$.

The system (H_1, B_1, F_1^H, W_1^H) is transformed via the inverse Legendre transformation $(q,\dot{q}) = \mathbb{F}H_1(q,p) = (q, m_1(q)^{-1}p)$ to the Lagrangian system (L_1, F_1^L, W_1^L), where

$$L_1(q,\dot{q}) = \frac{1}{2}\langle \dot{q}, m_1\dot{q}\rangle - V_1(q), \qquad F_1^L = F_1^H \circ \mathbb{F}H_1^{-1} \quad \text{and} \quad W_1^L = W_1^H \circ \mathbb{F}H_1^{-1}.$$

The system $(\tilde{H}_2, \tilde{B}_2, \tilde{F}_2^H, \tilde{W}_2^H)$ is transformed via the inverse Legendre transformation $(q,\dot{q}) = \mathbb{F}\tilde{H}_2(q,\tilde{p}) = (q, \tilde{m}_2^{-1}(q)\tilde{p})$ to the Lagrangian system (L_2, F_2^L, W_2^L), where

$$L_2(q,\dot{q}) = \frac{1}{2}\langle \dot{q}, \tilde{m}_2\dot{q}\rangle - V_2(q), \qquad F_2^L = \tilde{F}_2^H \circ \mathbb{F}\tilde{H}_2^{-1} \quad \text{and} \quad W_2^L = \tilde{W}_2^H \circ \mathbb{F}\tilde{H}_2^{-1}.$$

The diagram in Figure 3.3 commutes if and only if $\psi_{B_2} = m_1\tilde{m}_2^{-1}$, which by the definition of $\tilde{m}_2$ in (3.33) is equivalent to

$$\psi_{B_2} = m_2 m_1^{-1}, \tag{3.34}$$

$$\begin{tikzcd}
T^*Q \arrow[rr, "\psi_{B_2}^{-1}"] \arrow[dr, "m_1^{-1}"'] & & T^*Q \arrow[dl, "\tilde{m}_2^{-1}"] \\
& TQ &
\end{tikzcd}$$

Figure 3.3: Diagram of inverse Legendre transformations.

i.e., in the matrix form $K_2(q) = m_2(q)m_1(q)^{-1}$. We show that

$$\begin{aligned}
&(H_1, B_1, F_1^H, W_1^H) \overset{H}{\sim} (H_2, B_2, F_2^H, W_2^H) \\
\iff\ &[(L_1, F_1^L, W_1^L) \overset{L}{\sim} (L_2, F_2^L, W_2^L)] + (3.34).
\end{aligned} \tag{3.35}$$

With (3.34), one computes

$$\begin{aligned}
m_1\tilde{m}_2^{-1}W_2^L &= \psi_{B_2}(\psi_{B_2}^* W_2^H) \circ \mathbb{F}L_2 \\
&= \psi_{B_2}(\psi_{B_2}^{-1} W_2^H \circ \psi_{B_2}) \circ \mathbb{F}L_2 \\
&= W_2^H \circ m_1 = W_2^H \circ \mathbb{F}L_1.
\end{aligned}$$

First, by (3.34), **ELM-1** reads

$$W_1^L = m_1\tilde{m}_2^{-1}W_2^L \iff W_1^H = W_2^H,$$

which is **HM-1** for (H_1, B_1, F_1, W_1) and (H_2, B_2, F_2, W_2). Second, since $u_1^H, u_2^H \in W_1^H = W_2^H$, **HM-2** can be *equivalently* written in coordinates as (3.31). One can show that

$$\dot{q}_{H_1} - \dot{q}_{H_2} = m_1(q)^{-1}p - K_2(q)^T m_2(q)^{-1}p.$$

Hence, the first half of **HM-2** reads

$$\dot{q}_{H_1} - \dot{q}_{H_2} = 0 \iff K_2(q) = m_2(q)m_1(q)^{-1},$$

which is the commutativity condition (3.34). The remaining half of **HM-2** reads $\dot{p}_{H_1} - \dot{p}_{H_2} \in W_2^H$. By a similar computation carried out in § 3.2.2, one can show that

$$\dot{p}_{H_1} - \dot{p}_{H_2} \in W_2^H \iff \textbf{ELM-2}$$

under (3.34). One can also readily show that (2.12) on the u_i^L's is equivalent to (3.8) on the u_i^H's. Therefore we have shown (3.35).

3.2.4 Equivalence of CL/CH methods for Simple Mechanical Systems

We summarize the discussion in § 3.2.2 and § 3.2.3 in the following theorem[2].

Theorem 3.2.1. *The method of controlled Lagrangian systems is equivalent to that of controlled Hamiltonian systems for simple mechanical systems. Namely, the following hold:*

1 *Let (L_i, F_i^L, W_i^L), $i = 1, 2$, be two simple CL systems, and let $(H_i, B_{\mathrm{can}}, F_i^H, W_i^H)$ be the simple CH system Legendre-transformed via $\mathbb{F}L_i$ from the CL system (L_i, F_i^L, W_i^L) for $i = 1, 2$, respectively. Then,*

$$\begin{aligned}&(L_1, F_1^L, W_1^L) \overset{L}{\sim} (L_2, F_2^L, W_2^L)\\ &\quad\Longleftrightarrow\quad (H_1, B_{\mathrm{can}}, F_1^H, W_1^H) \overset{H}{\sim} (\mathbb{F}L_2 \circ \mathbb{F}L_1^{-1})^*(H_2, B_{\mathrm{can}}, F_2^H, W_2^H),\end{aligned}$$

where B_{can} is the canonical Poisson tensor on T^*Q.

2 *Let (H_i, B_i, F_i^H, W_i^H), $i = 1, 2$, be simple CH systems. Decompose B_i into its regular part $B_{\mathrm{r},i}$ and its gyroscopic part $B_{\mathrm{gr},i}$ such that*

$$(H_i, B_i, F_i^H, W_i^H) \overset{H}{\sim} (H_i, B_{\mathrm{r},i}, F_i^H + F_{\mathrm{gr},i}^H, W_i^H),$$

where the gyroscopic force $F_{\mathrm{gr},i}^H$ is defined by $\mathrm{vlift}(F_{\mathrm{gr},i}) = B_{\mathrm{gr},i}^\sharp \mathbf{d}H_i$ (see Proposition 3.1.7). Then there exist $\psi_{B_1}, \psi_{B_2} \in \Gamma(\mathrm{Aut}(T^*Q))$ satisfying $\phi_{B_i}^* B_{\mathrm{r},i} = B_{\mathrm{can}}$, and two simple CH systems $(\hat{H}_i, B_{\mathrm{can}}, \hat{F}_i^H, \hat{W}_i^H)$, $i = 1, 2$ such that

$$\psi_i^*(H_i, B_{\mathrm{r},i}, F_i^H + F_{\mathrm{gr},i}^H, W_i^H) = (\hat{H}_i, B_{\mathrm{can}}, \hat{F}_i^H, \hat{W}_i^H) \qquad i = 1, 2,$$

and finally

$$\begin{aligned}&(H_1, B_1, F_1^H, W_1^H) \overset{H}{\sim} (H_2, B_2, F_2^H, W_2^H) &&(3.36)\\ &\quad\Longleftrightarrow\quad (L_1, F_1^L, W_1^L) \overset{L}{\sim} (L_2, F_2^L, W_2^L) \quad \text{and} \quad \psi_{B_2} \circ \psi_{B_1}^{-1} = m_2 m_1^{-1},\end{aligned}$$

[2]The notation and terminology used in this theorem can be found as follows. Definition 2.1.3 and Definition 3.1.3 give the definitions of CL-equivalence relation, $\overset{L}{\sim}$, and the CH-equivalence relation, $\overset{H}{\sim}$, respectively. Definition 2.1.2 and Definition 3.1.6 give the definitions of simple CL systems and simple CH systems. § 3.2.1 and § 3.1.2 provide the definition of the Legendre transformation and the construction of ψ_1 and ψ_2. Proposition 3.1.11 and the remarks before it discuss the definition of pull-back systems.

where m_i is the mass tensor of H_i, $i = 1, 2$, and (L_i, F_i^L, W_i^L) is the simple CL system inverse-Legendre-transformed via $\mathbb{F}\hat{H}_i$ from the CH system $(\hat{H}_i, B_{\text{can}}, \hat{F}_i^H, \hat{W}_i^H)$ for $i = 1, 2$, respectively.

Proof. We only need to show (3.36) because in (3.35) of § 3.2.2 we showed the relation in (3.36) only in the case that one of the two CH systems has a canonical Poisson tensor. First recall $\psi_{B_i} = \psi_{B_{\mathrm{r},i}}$, $i = 1, 2$. Let us apply (3.35) to the following CH systems:

$$(\hat{H}_1, B_{\text{can}}, \hat{F}_i^H, \hat{W}_i^H) \quad \text{and} \quad \psi_{B_1}^*(H_2, B_{\mathrm{r},2}, F_2^H, W_2^H).$$

The respective mass tensors of $\hat{H}_1$ and $\psi_{B_1}^* H_2$ are

$$\psi_{B_1}^{-1} \circ m_1 \circ (\psi_{B_1}^{-1})^* \quad \text{and} \quad \psi_{B_1}^{-1} \circ m_2 \circ (\psi_{B_1}^{-1})^*.$$

In this case, (3.34) becomes

$$\psi_{\psi_{B_1}^* B_2} = \psi_{B_1}^{-1} \circ m_1 \circ (\psi_{B_1}^{-1})^* \circ (\psi_{B_1}^{-1} \circ m_2 \circ (\psi_{B_1}^{-1})^*)^{-1}$$

which by (3.24) implies $\psi_{B_2} \circ \psi_{B_1}^{-1} = m_2 m_1^{-1}$. Then, apply Proposition 3.1.11. ■

The following corollary compactly summarizes Theorem 3.2.1.

Corollary 3.2.2. *The method of controlled Lagrangian systems is equivalent to that of controlled Hamiltonian systems for simple mechanical systems in the following sense. For any two simple CL systems (L_i, F_i^L, W_i^L), $i = 1, 2$, there exist two associated simple CH systems (H_i, B_i, F_i^H, W_i^H), $i = 1, 2$, such that*

$$\begin{aligned} &(L_1, F_1^L, W_1^L) \overset{L}{\sim} (L_2, F_2^L, W_2^L) \quad \textit{and} \quad \psi_{B_2} \circ \psi_{B_1}^{-1} = m_{H_2}(m_{H_1})^{-1} \\ &\iff (H_1, B_1, F_1^H, W_1^H) \overset{H}{\sim} (H_2, B_2, F_2^H, W_2^H) \end{aligned}$$

with m_{H_i} the mass tensor of H_i and ψ_{B_i} defined in (3.11) for $i = 1, 2$, and vice versa.

Proof. One has to check

$$\psi_{B_2} \circ \psi_{B_1}^{-1} = (\text{mass tensor of } (\mathbb{F}L_2 \circ \mathbb{F}L_1^{-1})^* H_2)\,(\text{mass tensor of } H_1)^{-1} \tag{3.37}$$

in statement 1 of Theorem 3.2.1. Notice

$$\psi_{B_1} = \psi_{\text{can}} = \mathrm{id}_{T^*Q}, \quad \psi_{B_2} = \psi_{(m_2 m_1^{-1})^* B_{\text{can}}} = m_1 m_2^{-1}.$$

The mass tensors of H_1 and $(\mathbb{F}L_2 \circ \mathbb{F}L_1^{-1})^* H_2$ are m_1 and $m_1 m_2^{-1} m_1$, respectively. It follows that (3.37) holds. ■

We now discuss why Theorem 3.2.1 implies the equivalence between the CL method and the CH method. Suppose one is given a CH system (H_1, B_1, F_1^H, W_1^H) and he wants to find a CH-equivalent system. By Propositions 3.1.7, 3.1.9, and 3.1.11, one may assume

$$B_1 = B_{\text{can}}.$$

Let (L_1, F_1^L, W_1^L) be the CL system to which (H_1, B_1, F_1^H, W_1^H) is inverse-Legendre-transformed to via $\mathbb{F}H_1$. Then, find a CL system (L_2, F_2^L, W_2^L) such that

$$(L_1, F_1^L, W_1^L) \overset{L}{\sim} (L_2, F_2^L, W_2^L).$$

Let $(H_2, B_{\text{can}}, F_2^H, W_2^H)$ be the CH system to which (L_2, F_2^L, W_2^L) is Legendre-transformed via $\mathbb{F}L_2$. By statement 1 of Theorem 3.2.1, the following holds

$$(H_1, B_1, F_1^H, W_1^H) \overset{H}{\sim} (\mathbb{F}L_2 \circ \mathbb{F}L_1^{-1})^*(H_2, B_{\text{can}}, F_2^H, W_2^H).$$

Hence, we have found a CH system CH-equivalent to (H_1, B_1, F_1^H, W_1^H) using the CL equivalence relation. We now show that all CH systems CH-equivalent to (H_1, B_1, F_1^H, W_1^H) can be found in such a way. Suppose there is (H_2, B_2, F_2^H, W_2^H) which is CH-equivalent to (H_1, B_1, F_1^H, W_1^H), where we still assume $B_1 = B_{\text{can}}$ without loss of generality. By Propositions 3.1.7, we may assume that $B_2 = B_{\text{r},2}$ because we can always move the gyroscopic part $B_{\text{gr},2}$ to the external force part. Let (L_1, F_1, F_1^L) be the CL system to which (H_1, B_1, F_1^H, W_1^H) is inverse-Legendre-transformed via $\mathbb{F}H_1$. Let (L_2, F_2^L, W_2^L) be the CL system to which $\psi_{B_2}^*(H_2, B_2, F_2^H, W_2^H)$ is inverse-Legendre-transformed via $\mathbb{F}(H_2 \circ \psi_{B_2})$. By statement 2 of Theorem 3.2.1,

$$(L_1, F_1^L, W_1^L) \overset{L}{\sim} (L_2, F_2^L, W_2^L).$$

By statement 1 of Theorem 3.2.1,

$$(H_1, B_1, F_1^H, W_1^H) \overset{H}{\sim} (\mathbb{F}L_2 \circ \mathbb{F}L_1^{-1})^*(\psi_{B_2}^*(H_2, B_2, F_2^H, W_2^H)). \tag{3.38}$$

By the way, one can easily check $\psi_{B_2} \circ (\mathbb{F}L_2 \circ \mathbb{F}L_1^{-1}) = \text{id}$. Hence,

$$(H_2, B_2, F_2^H, W_2^H) = (\mathbb{F}L_2 \circ \mathbb{F}L_1^{-1})^* \psi_{B_2}^*(H_2, B_2, F_2^H, W_2^H).$$

Namely, (H_2, B_2, F_2^H, W_2^H) coincides with the CH system which is derived by the CL equivalence relation.

In the similar manner, one can show that given a CL system, he can find all its CL-equivalent systems by using the CH equivalence relation.

3.3 Summary and Future Work

We have developed the method of CH systems and have shown its equivalence to the method of CL systems for simple mechanical systems. The concept of the CH method has been used under different names and versions. Here, we gave it an intrinsic formulation. In the following, we summarize this chapter.

§ 3.1. We first reviewed the Hamiltonian mechanics. We then defined CH systems on TQ (Definition 3.1.1). Then we defined the CH-equivalence relation among CH systems on TQ and the Hamiltonian matching conditions (Definition 3.1.3). If two CH systems are CH-equivalent, then for any control for one system there exists a control for the other system such that the two closed-loop systems produced the same equations of motion (Proposition 3.1.4). We defined *simple* CH systems (Definition 3.1.6) where the Hamiltonian has the kinetic plus potential energy form and the almost Poisson structure is of the form (3.10). This particular definition was chosen so that we can show the equivalence of the methods of *simple* CL systems and *simple* CH systems. We then interpreted the failure of the Jacobi identity by almost Poisson tensors in terms of gyroscopic forces. (Proposition 3.1.7 and 3.1.9). We developed the concept of pull-back systems (Definition 3.1.11), which was later used in § 3.2.

We gave the usual procedure of applying the method of CH systems to control synthesis for asymptotic stabilization. The basic idea is as follows: Given a CH systems of the form $(H, B, 0, W)$, which is of the usual *ideal* form in application, find a CH-equivalent system $(\widehat{H}, \widehat{B}, \widehat{F}_{\mathrm{gr}}, \widehat{W})$ where $\widehat{H}$ has a minimum at the equilibrium of interest and $\widehat{F}_{\mathrm{gr}}$ is of gyroscopic form. Then, add a dissipative feedback force in the direction of $\widehat{W}$. One can alternatively omit the gyroscopic term because it can be always combined into the almost Poisson structure for simple CH systems by Proposition 3.1.9. We did not give any examples of application of the CH method to stabilization problems because there is already good literature available (Ortega, Spong, Gómez-Estern, and Blankenstein [2001]).

§ 3.2. We showed that the method of simple CL systems and that of simple CH systems are equivalent (Theorem 3.2.1 and Corollary 3.2.2). This equivalence implies that one can use either method for applications. Which method to use depends specifically on the given problem, just as the preferred choice of coordinates depends on the given PDE.

However, one should remember that on the CL side the extended λ-method is available to systematically solve the involved PDE's.

Future Work.

1. One can develop the theory of CH systems on a general manifold M, not necessarily on cotangent bundles T^*Q. In such a case, however, it is not clear how to introduce external and control forces. In some applications in electrical circuits, the phase space is odd dimensional, i.e., not a cotangent bundle of a manifold (see, for example, Petrović, Ortega, and Stanković [2001]). They are due to the Kirchhoff law, to the degeneracy of Poisson structures, or to the reduction of symmetry. Controlled Hamiltonian systems with symmetry are treated in § 4 in this thesis. Hence, as a future work, it would be interesting to consider the case of degenerate Poisson structures or CH systems with constraints.

2. Hamiltonian normal forms are well developed (Wiggins [1990]) where they only use canonical transformations of phase space variables. It would be interesting to consider the generalized normal form by using feedback transformations as well. It is not clear at the moment how this direction of work is related to the current CH method.

Chapter 4

Reduction of Controlled Lagrangian and Hamiltonian Systems with Symmetry

Symmetry in a system can be regarded as redundancy. To see the essential dynamics of a system with symmetry, one needs to remove the symmetry to reduce the system. Reduction theory in mechanics has been well developed on both the Lagrangian and the Hamiltonian sides. One reduces variational principles on the Lagrangian side and Poisson structures on the Hamiltonian side. In geometric mechanics, external forces or control forces are not usually taken into account whereas forces are important notions in control theory. In the following, we outline this chapter.

In this chapter, we extend the theory of CL/CH methods to include systems with symmetry and the relevant reduction theory. Unlike the traditional mechanics, we take into account both external forces and control forces on both CL/CH sides and use almost Poisson structures on the CH side rather than Poisson structures (namely, we allow for the failure of the Jacobi identity). Our work is based on Lagrangian reduction in Cendra, Marsden, and Ratiu [2001], and Poisson reduction in Marsden and Ratiu [1999]. The method of reduced CL systems was developed *in coordinates for the case that configuration space is the product of two Lie groups* in Bloch, Leonard, and Marsden [1998, 2001]. We will develop the method of reduced CL systems *intrinsically for a general G-principal bundle Q with a free and proper G-action on Q.* The method of reduced CH systems was implicitly used in Krishnaprasad [1985], Bloch, Krishnaprasad, Marsden, and Sánchez De Alvarez [1992], and Woolsey and Leonard [1999]. Here, we improve the foundational setting for both the reduced CL and CH methods and make clear the relationship between reduced systems and G-invariant unreduced systems; see § 4.1 and 4.2. In addition, we show the equivalence of the method of reduced simple CL systems and the method of reduced simple CH systems in § 4.3; see Figure 4.1.

The work in this chapter is critical for many applications including spacecraft control, underwater vehicle control, and many other systems. In fact, this class of reduced systems

$$
\begin{tikzcd}
\text{CL with symmetry} \arrow[r, Leftrightarrow] \arrow[d, "/G"] & \text{CH with symmetry} \arrow[d, "/G"] \\
\text{reduced CL} \arrow[r, Leftrightarrow] & \text{reduced CH}
\end{tikzcd}
$$

Figure 4.1: Equivalence of reduced CL and CH methods.

was recognized early as an important one on both the Hamiltonian and Lagrangian sides by Krishnaprasad [1985], Bloch, Krishnaprasad, Marsden, and Sánchez De Alvarez [1992], Wang and Khrishnaprasad [1992]. We apply the method of reduced CL systems to the following two systems for control synthesis: the satellite with a rotor (§ 4.1.3) and the heavy top with two rotors (§ 4.1.4). The application of the CL method to the first system was started in Bloch, Leonard, and Marsden [1997], and completed in Bloch, Chang, Leonard, Marsden and Woolsey [2000] by showing asymptotic stabilization. The control of the second system was presented in Chang and Marsden [2000]. We refer to Bloch, Leonard, and Marsden [2001] and Woolsey and Leonard [1999] for the application of the reduced CL/CH method to the underwater vehicle system.

Mathematical Notations and General Assumptions. We discuss mathematical notations and general assumptions which will be used in this chapter. Refer to Abraham, Marsden, and Ratiu [1988], Cendra, Marsden, and Ratiu [2001] and Kobayashi and Nomizu [1963] for more details. Let Q be the configuration manifold, and $\tau_Q : TQ \to Q$ and $\pi_Q : T^*Q \to Q$ be the tangent bundle projection and the cotangent bundle projection, respectively. Denote by $\tau_Q^{(2)} : T^{(2)}Q \to Q$ the second order tangent bundle projection. For a manifold M, $\mathcal{F}(M)$ denotes the set of smooth real-valued functions on M.

Let G be a Lie group acting (on the left) on Q freely and properly such that $\pi_G(Q) : Q \to Q/G$ becomes a principal bundle. The tangent (resp. cotangent) lift action of G on TQ (resp. T^*Q) is free and proper and $\tau_{/G} : TQ \to TQ/G$ (resp. $\pi_{/G} : T^*Q \to T^*Q/G$) becomes a principal bundle. When M is a manifold on which G acts, we let $[m]_G$ to denote the equivalence class of $m \in M$ in the quotient space M/G. Even though we do not explicitly specify the manifold M in this notation, it will be clear in the context. The space TQ/G becomes a vector bundle with base Q/G by inheriting the vector bundle structure of TQ as follows:

$$[u_q]_G + \lambda[v_q]_G = [u_q + \lambda v_q]_G$$

where $\lambda \in \mathbb{R}$, $u_q, v_q \in T_qQ$ and $[u_q]_G, [v_q]_G$ are their equivalence classes in the quotient space TQ/G. The fiber $(TQ/G)_x$ is isomorphic, as a vector space, to T_qQ for each $x = [q]_G \in Q/G$, $q \in Q$ (see Lemma 2.4.1 in Cendra, Marsden, and Ratiu [2001]). In the same manner, the space T^*Q/G becomes a vector bundle with base Q/G.

4.1 Reduction of Controlled Lagrangian Systems with Symmetry

Based on the work on the Lagrangian reduction in Cendra, Marsden, and Ratiu [2001], we develop the reduction theory of controlled Lagrangian systems with symmetry. This will draw a clear picture of the relation between CL systems with symmetry and the reduced CL systems. We introduce an equivalence relation, called the reduced-Euler-Lagrange equivalence, by reducing the CL-equivalence for G invariant CL systems. This allows us to apply the *reduced* CL method directly to control problems where we are interested in reduced dynamics. We will apply the reduced CL method to such examples as the satellite with a rotor and the heavy top with two rotors.

4.1.1 Reduction of CL systems with Symmetry

We defined the CL system in Definition 2.1.1. Here, we define G invariant CL systems on TQ and reduced CL systems on TQ/G, where G is a Lie group acting on Q.

Definition 4.1.1. *Let G be a Lie group acting on Q. A G invariant controlled Lagrangian (G-CL) system is a CL system, (L, F, W), where L is a G invariant Lagrangian, F is a G equivariant force map, and W is a G invariant subbundle of T^*Q.*

Definition 4.1.2. *A reduced controlled Lagrangian (RCL) system is a triple (l, f, U) where $l : TQ/G \to \mathbb{R}$ is a smooth function called reduced Lagrangian, $f : TQ/G \to T^*Q/G$ is a fiber-preserving map called reduced force map, and the subbundle U of T^*Q/G is called the reduced control bundle. A feedback control for the RCL system is a (fiber-preserving) map of TQ/G into U.*

Suppose that we are given a G-CL system (L, F, W). The G invariance of L induces the reduced Lagrangian l on TQ/G satisfying

$$l \circ \tau_{/G} = L. \tag{4.1}$$

The G equivariance of F induces a reduced force map $[F]_G : TQ/G \to T^*Q/G$ satisfying

$$[F]_G \circ \tau_{/G} = \pi_{/G} \circ F. \tag{4.2}$$

This leads to the following definition:

Definition 4.1.3. *The RCL system of a G-CL system (L, F, W) is a triple $(l, [F]_G, W/G)$ where l is the reduced Lagrangian satisfying (4.1), and $[F]_G$ is the reduced force satisfying (4.2).*

One may ask if there exists a G-CL system on TQ when one is given a RCL system on TQ/G. The following proposition proves its unique existence.

Proposition 4.1.4. *Given a RCL system (l, f, U) on TQ/G, there is a unique G-CL system (L, F, W) on TQ whose RCL system is (l, f, U).*

Proof. Define L by (4.1). Define a force map F on TQ as follows: for $v_q, w_q \in T_qQ$,

$$\langle F(v_q),\, w_q\rangle = \langle f \circ \tau_{/G}(v_q),\, \tau_{/G}(w_q)\rangle. \tag{4.3}$$

One can check the G equivariance of F. One can also check that relation (4.3) defines a unique *fiber-preserving* map F of TQ to T^*Q. Let $W := \tau_{/G}^{-1}(U)$. By construction, (L, F, W) is the unique G-CL system whose RCL system is (l, f, U). ■

By Proposition 4.1.4, we can, without loss of generality, write an arbitrary RCL system in the form of the RCL system of a G-CL system. In addition, the proof of Proposition 4.1.4 implies the following claim: Given a fiber-preserving map $f : TQ/G \to T^*Q/G$, there exists a unique fiber-preserving map $F : TQ \to T^*Q$ satisfying

$$f \circ \tau_{/G} = \pi_{/G} \circ F.$$

Given a G-CL system (L, F, W), the G invariance of L implies the G-equivariance of the map $\mathcal{EL}(L) : T^{(2)}Q \to T^*Q$ in (2.2), which induces a quotient map

$$\mathcal{REL}(l) := [\mathcal{EL}(L)]_G : T^{(2)}Q/G \to T^*Q/G,$$

which depends only on the reduced Lagrangian l on TQ/G induced from L. The operator $\mathcal{REL}$ is called the reduced Euler-Lagrange operator. The equations of motion of a RCL system $(l, [F]_G, W/G)$ with a choice of control $[u]_G : TQ/G \to W/G$ are given by

$$\mathcal{REL}(l)([q, \dot{q}, \ddot{q}]_G) = [F]_G([q, \dot{q}]) + [u]_G([q, \dot{q}]). \tag{4.4}$$

To write computable equations of $\mathcal{REL}$, one has to choose a principal connection on the principal bundle $Q \to Q/G$ to identify the quotient bundles,

$$TQ/G \text{ with } T(Q/G) \oplus \tilde{\mathfrak{g}}$$
$$T^{(2)}Q/G \text{ with } T^{(2)}(Q/G) \times_{Q/G} 2\tilde{\mathfrak{g}}$$

and

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

where $\tilde{\mathfrak{g}}$ is the adjoint bundle $\mathrm{Ad}(Q)$, $\tilde{\mathfrak{g}}^*$ is the coadjoint bundle $\mathrm{Ad}^*(Q)$, $2\tilde{\mathfrak{g}} := \tilde{\mathfrak{g}} \oplus \tilde{\mathfrak{g}}$, and $\oplus$ is the Whitney sum (see Lemma 2.4.2 and Lemma 3.2.2 in Cendra, Marsden, and Ratiu [2001]). With these identifications, $\mathcal{REL}(l)$ induces the ***Lagrange-Poincaré operator***

$$\mathcal{LP}(l) : T^{(2)}(Q/G) \times_{Q/G} 2\tilde{g} \to T^*(Q/G) \oplus \tilde{g}^*. \tag{4.5}$$

Hence, the reduced Euler-Lagrange operator, $\mathcal{REL}$ may be replaced by the Lagrange-Poincaré operator $\mathcal{LP}$ in the following as long as one chooses a connection on $Q \to Q/G$. More details may be found in Cendra, Marsden, and Ratiu [2001].

We study the relation between trajectories of G-CL systems and trajectories of RCL systems. Let (L, F, W) be a G-CL system and $(l, [F]_G, W/G)$ its RCL system. Choose an arbitrary G *equivariant* feedback control law $u : TQ \to W$ for (L, F, W). The control u induces a reduced map $[u]_G : TQ/G \to W/G$. If $(q(t), \dot{q}(t)) \in TQ$ is a trajectory of the closed-loop system (L, F, u), then $\tau_{/G}(q(t), \dot{q}(t)) \in TQ/G$ is the trajectory of the closed-loop system $(l, [F]_G, [u]_G)$.

4.1.2 Reduced CL Equivalence

Recall the definition of simple CL systems and define G invariant simple CL systems in the following.

Definition 4.1.5. *A CL system (L, F, W) on TQ is called **simple** if its Lagrangian $L : TQ \to \mathbb{R}$ is of the form kinetic minus potential energy as follows:*

$$L(q, \dot{q}) = \frac{1}{2} m_q(\dot{q}, \dot{q}) - V(q), \tag{4.6}$$

*where m is a mass tensor, i.e., a positive definite symmetric $(0, 2)$-tensor. A reduced CL system $(l, [F]_G, W/G)$ is called simple if the reduced Lagrangian l is induced by a G invariant simple Lagrangian L on TQ. The acronym, **(R)SCL**, will denote "(reduced) simple controlled Lagrangian".*

When a simple G invariant Lagrangian L is given by (4.6), its reduced simple Lagrangian $l : TQ/G \to \mathbb{R}$ is given by

$$l([q,\dot{q}]_G) = \frac{1}{2}[m]_G([q,\dot{q}]_G,[q,\dot{q}]_G) - [V]_G([q]_G),$$

where $[m]_G \in \Gamma(Q/G, T^*Q/G \otimes T^*Q/G)$ is the reduced mass tensor induced from the G invariance of the mass tensor $m \in \Gamma(Q, T^*Q \otimes T^*Q)$ and $[V]_G : Q/G \to \mathbb{R}$ is the reduced potential energy.

We defined the Euler-Lagrange matching conditions and the CL-equivalence relation in Definition 2.1.3. We now define an equivalence relation among RSCL systems on TQ/G.

Definition 4.1.6. *Two RSCL systems $(l_i, [F_i]_G, W_i/G)$, $i = 1,2$ are said to be* ***reduced-CL-equivalent (RCL-equivalent)****, or simply, $(l_1, [F_1]_G, W_1/G) \overset{L}{\sim} (l_2, [F_2]_G, W_2/G)$ if the following* ***reduced Euler-Lagrange matching conditions*** *hold:*

RELM-1 : $W_1/G = [m_1]_G[m_2]_G^{-1}(W_2/G)$,

RELM-2 : $\operatorname{Im}[\mathcal{REL}(l_1) - [F_1]_G - [m_1]_G[m_2]_G^{-1}(\mathcal{REL}(l_2) - [F_2]_G)] \subset W_1/G$

where $[m_i]_G$ is the reduced mass tensor of l_i, $i = 1,2$.

The following proposition explains the relationship between the CL-equivalence relation among G-SCL systems and the RCL-equivalence relation among RSCL systems.

Proposition 4.1.7. *Two G-SCL systems are CL-equivalent if and only if their associated RSCL systems are RCL-equivalent.*

Proof. Let (L, F, W) be a G-SCL system, and $(l, [F]_G, W/G)$ its associated RSCL system. Then, the proposition follows from the G invariance of W and the following relations:

$$\mathcal{REL}(l) \circ \tau_{/G}^{(2)} = \pi_{/G} \circ \mathcal{EL}(L); \qquad [F]_G \circ \tau_{/G} = \pi_{/G} \circ F,$$

where $\tau_{/G}^{(2)} : T^{(2)}Q \to T^{(2)}Q/G$ is the G quotient map. ■

Hence, one can check the RCL equivalence of two RSCL systems in two ways; one is to directly check it, and the other is to check the CL equivalence of their associated *unreduced* G-SCL systems. In practice, it is more convenient to check it directly at the reduced level; see § 4.1.3.

The following proposition explains the property of the RCL-equivalence relation:

Proposition 4.1.8. *Suppose that two RSCL systems $(l_i, [F_i]_G, W_i/G)$, $i = 1,2$ are RCL-equivalent. Then, for an arbitrary control law for one system, there exists a control law for the other system such that the two closed-loop RSCL systems produce the same equations* *of motion. The explicit relation between the two feedback control laws $[u_i]_G$, $i = 1,2$ is given by*

$$[u_1]_G = \mathcal{REL}(l_1) - [F_1]_G - [m_1]_G[m_2]_G^{-1}(\mathcal{REL}(l_2) - [F_2]_G) + [m_1]_G[m_2]_G^{-1}[u_2]_G \quad (4.7)$$

where $[m_i]_G$ is the reduced mass tensor of l_i, $i = 1,2$.

Proof. Let $[u_i]_G$ be a feedback control for $(l_i.[F_i]_G, W_i/G)$, $i = 1,2$. Let (L_i, F_i, W_i) be the unreduced G-SCL system of $(l_i, [F_i]_G, W_i/G)$, $i = 1,2$. By Proposition 4.1.7, the two G-SCL systems are CL-equivalent. By Proposition 2.1.5, the two closed-loop G-SCL systems (L_i, F_i, u_i), $i = 1,2$ produce the same equations of motion when u_1 and u_2 satisfy (2.12). Hence, the two closed-loop RSCL systems $(l_i, [F_i]_G, [u_i]_G)$, $i = 1,2$ produce the same equations of motion when $[u_1]_G$ and $[u_2]_G$ satisfy (4.7) because each term in (2.12) is G equivariant. Also notice that for any choice of $[u_i]_G$, one can choose the other $[u_j]_G$ such that (4.7) holds. ■

One can prove Proposition 4.1.8 by comparing the expressions for "accelerations" of both equations as in the proof of Proposition 2.1.5. For this purpose, one needs to choose a connection on $Q \to Q/G$ because one has to split the variations to write down the equations of motion in coordinates as the Euler-Lagrange equations come from the variational principles (see Chapter 3 of Cendra, Marsden, and Ratiu [2001] for more detail). In the current proof of Proposition 4.1.8, we were able to bypass this route by Proposition 4.1.7. In Bloch, Leonard, and Marsden [2001], they choose the trivial connection and then compare the acceleration terms to find the Euler-Poincaré matching conditions, which is a special case of the reduced Euler-Lagrange matching conditions, as will be shown later in § 4.1.3.

Remark 4.1.9. *The application of the reduced CL method to stabilization control problems is similar to that of the CL method in § 2.2.*

4.1.3 Example: Satellite with a Rotor

We use the method of reduced CL systems to design a feedback control law for the system of a satellite with a rotor. This work was published in Bloch, Chang, Leonard, Marsden and Woolsey [2000], Bloch, Leonard, and Marsden [2001], and Bloch, Leonard, and Marsden [1998], where the complete treatment was first given in Bloch, Chang, Leonard, Marsden and Woolsey [2000]. Then, we review the Euler-Poincaré matching conditions presented in Bloch, Leonard, and Marsden [2001] in the framework of this thesis. The satellite with a rotor satisfies the Euler-Poincaré matching conditions.

Satellite with a Rotor. We study the stabilization problem for the system of a satellite with a rotor aligned along the third principal axis of the body within the framework of this thesis; see Figure 4.2. The configuration space is $Q = G \times H = \mathrm{SO}(3) \times S^1$ with the first

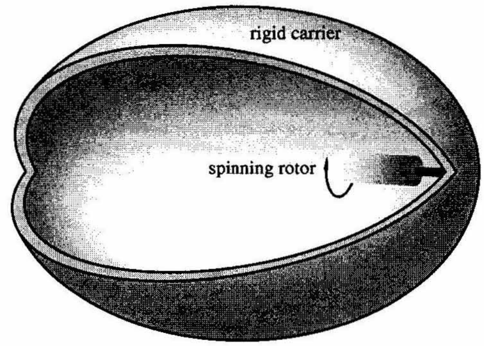

Figure 4.2: Satellite with a rotor along the third body axis.

factor being the satellite attitude and the second factor being the rotor angle. The Lie group $G = \mathrm{SO}(3)$ acts on the first factor of Q only. We take a trivial connection on Q such that $TQ/G \simeq \mathfrak{g} \times TH$. Use $((\Omega_1, \Omega_2, \Omega_3), (\phi, \dot{\phi}))$ as coordinates for $\mathfrak{so}(3) \times TS^1 \simeq \mathbb{R}^3 \times TS^1$. This system is described by the RSCL system $(l_1, [F_1]_G = 0, W_1/G)$ given by

$$l_1(\Omega, \dot{\phi}) = \frac{1}{2}\left(\lambda_1 \Omega_1^2 + \lambda_2 \Omega_2^2 + (I_3 + J_3)\Omega_3^2 + 2J_3 \Omega_3 \dot{\phi} + J_3 \dot{\phi}^2\right) \tag{4.8}$$

and

$$W_1/G = \mathrm{span}\{\mathbf{d}\phi\} = \mathrm{span}\{(0,0,0,1)^t\},$$

where $\lambda_1 > \lambda_2 > \lambda_3 := J_3 + I_3$. Notice that l_1 does not depend on ϕ. Recall that the reduced Euler-Lagrange operator $\mathcal{REL}$ induces the Lagrange-Poincaré operator $\mathcal{LP}$ in (4.5) with respect to the trivial connection. This Lagrange-Poincaré operator $\mathcal{LP}(l_1)$ is given by

$$\mathcal{LP}(l_1) = \begin{bmatrix} \frac{d}{dt}\frac{\partial l_1}{\partial \Omega} - \Omega \times \frac{\partial l_1}{\partial \Omega} \\ \frac{d}{dt}\frac{\partial l_1}{\partial \dot{\phi}} - \frac{\partial l_1}{\partial \phi} \end{bmatrix}. \tag{4.9}$$

See Cendra, Marsden, and Ratiu [2001] or Chapter 13 of Marsden and Ratiu [1999] about more detail on the Lagrange-Poincaré operator, which is sometimes called the Euler-Poincaré operator when $Q = G$. Consider another RCL system $(l_2, 0, W_2/G)$ with $W_2 = W_1$ and

$$l_2(\Omega, \dot{\phi}) = \frac{1}{2}\left(\lambda_1 \Omega_1^2 + \lambda_2 \Omega_2^2 + (I_3 + J_3)\Omega_3^2 + 2J_3 \Omega_3 \dot{\phi} + \rho J_3 \dot{\phi}^2\right)$$

with $\rho \in \mathbb{R}$ a free parameter ρ. The Lagrange-Poincaré operator $\mathcal{LP}(l_2)$ can be written as in (4.9) with the replacement of l_1 by l_2. By definition of W_2, **RELM-1** is automatically satisfied and one can check that **RELM-2** also holds for the two RSCL systems.

The control goal is to design a feedback control which makes asymptotically stable the rotation about the middle axis in the body-fixed frame and gets the rotor asymptotically to rest. Since we are not interested in the angle ϕ of the rotor, we will ignore the ϕ variable in the dynamics. Hence, the phase space will be $\mathfrak{so}(3) \times \mathbb{R}$. The equilibrium of interest is

$$z_e = (\Omega_1, \Omega_2, \Omega_3, \dot{\phi}) = (0, \bar{\Omega}, 0, 0); \qquad \bar{\Omega} \neq 0.$$

By Proposition 4.1.8, we can equivalently work with the system $(l_2, 0, W_2/G)$.

We use the energy-Casimir method to construct a Lyapunov function (see Bloch, Chang, Leonard, Marsden and Woolsey [2000] for more detail on the energy-Casimir method). Consider the following Lyapunov function candidate:

$$E_{\tilde{\Phi}}(\Omega, \dot{\phi}) = l_2(\Omega, \dot{\phi}) + \Phi(C) + \Psi(\tilde{l}),$$

where Φ and Ψ are functions to be chosen and C is half of the square of the total angular momentum given by

$$C = \frac{1}{2}((\lambda_1\Omega_1)^2 + (\lambda_2\Omega_2)^2 + (\lambda_3\Omega_3 + J_3\dot{\phi})^2),$$

and $\tilde{l}$ is defined by

$$\tilde{l} := \frac{\partial l_2}{\partial \dot{\phi}} = J_3(\Omega_3 + \rho\dot{\phi}).$$

Then, the equilibrium z_e is a critical point of $E_{\tilde{\Phi}}$ if and only if

$$\Phi'(z_e) = -\frac{1}{\lambda_2}; \qquad \Psi'(z_e) = 0. \tag{4.10}$$

The second derivative of $E_{\tilde{\Phi}}$ at z_e is given by

$$D^2E_{\tilde{\Phi}}(z_e) = \begin{bmatrix} \lambda_1 - \frac{(\lambda_1)^2}{\lambda_2} & 0 & 0 & 0 \\ 0 & \Phi''(z_e)((\lambda_2)^2\bar{\Omega})^2 & 0 & 0 \\ 0 & 0 & \lambda_3 - \frac{(\lambda_3)^2}{\lambda_2} + \Psi''(z_e)J_3^2 & J_3((1 - \frac{\lambda_3}{\lambda_2}) + \Psi''(z_e)\rho J_3) \\ 0 & 0 & J_3((1 - \frac{\lambda_3}{\lambda_2}) + \Psi''(z_e)\rho J_3) & J_3(\rho - \frac{J_3}{\lambda_2} + \Psi''(z_e)\rho^2 J_3) \end{bmatrix}.$$

Let

$$J = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -\frac{1}{\rho} & \frac{1}{\rho J_3} \end{bmatrix}.$$

Then,

$$
\begin{aligned}
&J^T D^2 E_{\tilde{\Phi}}(z_e) J \\
&= \begin{bmatrix}
\lambda_1 - \frac{(\lambda_1)^2}{\lambda_2} & 0 & 0 & 0 \\
0 & \Phi''(z_e)((\lambda_2)^2\bar{\Omega})^2 & 0 & 0 \\
0 & 0 & \lambda_3 - \frac{J_3}{\rho} - \frac{1}{\lambda_2}(\lambda_3 - \frac{J_3}{\rho})^2 & -\frac{1}{\rho\lambda_2}(\lambda_3 - \frac{J_3}{\rho}) \\
0 & 0 & -\frac{1}{\rho\lambda_2}(\lambda_3 - \frac{J_3}{\rho}) & \frac{1}{\rho J_3} - \frac{1}{\lambda_2\rho^2} + \Psi''(z_e)
\end{bmatrix}.
\end{aligned}
$$

So, the function $E_{\tilde{\Phi}}$ has a local strict maximum[1] at z_e if (4.10) and the following holds

$$
\begin{cases}
\Phi''(z_e) < 0, \\
\lambda_3 - \frac{J_3}{\rho} - \frac{1}{\lambda_2}\left(\lambda_3 - \frac{J_3}{\rho}\right)^2 < 0, \\
\Psi''(z_e) < -\left(\frac{1}{\rho J_3} - \frac{1}{\lambda_2\rho^2}\right) + \frac{\lambda_3 - \frac{J_3}{\rho}}{\rho^2\lambda_2(\lambda_2 - \lambda_3 + \frac{J_3}{\rho})}.
\end{cases} \tag{4.11}
$$

We can always find the parameters ρ, Φ, Ψ satisfying (4.10) and (4.11). For simplicity, we choose

$$\Psi = \frac{1}{2\epsilon J_3}(\tilde{l})^2,$$

with $\epsilon = 1/(J_3\Psi''(z_e))$ satisfying the third equation in (4.11). Take the following feedback control u_2 for $(l_2, 0, W_2/G)$:

$$u_2 = c\left(\frac{1}{\epsilon}\Omega_3 + \left(1 + \frac{\rho}{\epsilon}\right)\dot{\phi}\right)$$

with $c > 0$, so that

$$\frac{dE_{\tilde{\Phi}}}{dt} = c\left(\frac{1}{\epsilon}\Omega_3 + \left(1 + \frac{\rho}{\epsilon}\right)\dot{\phi}\right)^2 \geq 0.$$

Hence, the equilibrium z_e is Lyapunov stable in the closed-loop system $(l_2, 0, u_2)$.

We now show the asymptotic stability of the equilibrium. The equations of motion of the closed-loop system $(l_2, 0, u_2)$ are given by

$$\lambda_1\dot{\Omega}_1 = \lambda_2\Omega_2\Omega_3 - (\lambda_3\Omega_3 + J_3\dot{\phi})\Omega_2, \tag{4.12}$$

$$\lambda_2\dot{\Omega}_2 = -\lambda_1\Omega_1\Omega_3 + (\lambda_3\Omega_3 + J_3\dot{\phi})\Omega_1, \tag{4.13}$$

$$\lambda_3\dot{\Omega}_3 + J_3\ddot{\phi} = (\lambda_1 - \lambda_2)\Omega_1\Omega_2, \tag{4.14}$$

$$\dot{\tilde{l}} = u_2\,. \tag{4.15}$$

Suppose that the flow $(\Omega_1(t), \Omega_2(t), \Omega_3(t), \dot{\phi}(t))$ satisfies $\dot{E}_{\tilde{\Phi}} = 0$, equivalently $u_2 = 0$.

[1] See Remark 2.1.10.

Then, since $\dot{\tilde{l}} = v$, $\tilde{l}(t)$ is constant. This implies that

$$\dot{\phi}(t) = \dot{\phi}(0) = \text{constant},$$
$$\Omega_3(t) = \Omega_3(0) = \text{constant}.$$

Substituting these into (4.14), we get

$$\Omega_1\Omega_2 = 0. \tag{4.16}$$

Since $\Omega_2(t)$ stays near $\bar{\Omega} \neq 0$ by stability, (4.16) implies that

$$\Omega_1(t) = 0$$

for all t. Substitution of this into (4.13) gives

$$\Omega_2(t) = \Omega_2(0) = \text{constant}. \tag{4.17}$$

Substitute these two into (4.12) and we get

$$\left((\lambda_2 - \lambda_3)\Omega_3 - J_3\dot{\phi}\right)\Omega_2(0) = 0$$

or

$$(\lambda_2 - \lambda_3)\Omega_3 - J_3\dot{\phi} = 0 \tag{4.18}$$

since $\Omega_2(0) \neq 0$ by stability. We also have $u_2 = 0$, which is given by

$$\Omega_3 + (\epsilon + \rho)\dot{\phi} = 0. \tag{4.19}$$

All we required on $\epsilon = 1/(J_3\Psi''(z_e))$ was that it satisfy the third inequality in (4.11). We can find ϵ satisfying the following additional condition:

$$(\lambda_2 - \lambda_3)(\epsilon + \rho) + J_3 \neq 0 \tag{4.20}$$

such that the two equations in (4.18) and (4.19) are independent. Then $\Omega_3 = \dot{\phi} = 0$. Thus, the only possible flow satisfying $u_2 = 0$ is

$$\Omega_1(t) = \Omega_3(t) = \dot{\phi}(t) = 0, \qquad \Omega_2(t) = \Omega_2(0).$$

This implies that

$$|\lambda_2\Omega_2(0)|^2 = |\lambda_1\Omega_1|^2 + |\lambda_2\Omega_2|^2 + |\lambda_3\Omega_3 + J_3\dot{\phi}|^2 = |\lambda_2\bar{\Omega}|^2$$

because the magnitude of the angular momentum is conserved. So $\Omega_2(0) = \bar{\Omega}$ by stability. Thus, the only possible flow satisfying $u_2 = 0$ is the equilibrium. By LaSalle's theorem, it is asymptotically stable.

Euler-Poincaré Matching. Here we briefly sketch the proof that the set of the Euler-Poincaré matching conditions in Bloch, Leonard, and Marsden [1998] and Bloch, Leonard, and Marsden [2001] is a special case of the reduced Euler-Lagrange matching conditions. The matching conditions can handle such examples as a satellite with a rotor and underwater vehicles with internal rotors. Let $Q = G \times H$ be the configuration space where G is a Lie group acting trivially on H, and H is an Abelian Lie group[2]. We choose the trivial connection on $Q \to H$ to write down the Lagrange-Poincaré equation on $TQ/G \simeq \mathfrak{g} \times TH$ with the Lie algebra $\mathfrak{g}$ of the Lie group G. We use $\eta = (\eta^\alpha)$ as coordinates for $\mathfrak{g}$ and $(\theta, \dot{\theta}) = (\theta^a, \dot{\theta}^a)$ as coordinates for TH. The Lagrange-Poincaré operator $\mathcal{LP}$ with respect to the trivial connection is given by

$$\mathcal{LP}(l) = \begin{pmatrix} \frac{d}{dt}\frac{\partial l}{\partial \eta^\alpha} - c^\beta_{\alpha\gamma}\eta^\gamma \frac{\partial l}{\partial \eta^\beta} \\ \frac{d}{dt}\frac{\partial l}{\partial \dot{\theta}^a} - \frac{\partial l}{\partial \theta^a} \end{pmatrix} \tag{4.21}$$

for any reduced Lagrangian $l = l(\eta^\alpha, \dot{\theta}^a, \theta^a)$, where $c^\beta_{\alpha\delta}$ are the structure coefficients of the Lie algebra $\mathfrak{g}$. See Cendra, Marsden, and Ratiu [2001] for the derivation of (4.21). One wants to check (4.9) by substituting the structure coefficients, $c^k_{ij} = \delta_{ijk}$, of the Lie algebra $\mathfrak{so}(3)$ into (4.21).

Let $(l, 0, T^*H)$ be the given RSCL system with the reduced Lagrangian

$$l(\eta^\alpha, \dot{\theta}^a) = \frac{1}{2}g_{\alpha\beta}\eta^\alpha\eta^\beta + g_{\alpha a}\eta^\alpha\dot{\theta}^a + \frac{1}{2}g_{ab}\dot{\theta}^a\dot{\theta}^b,$$

where $g_{\alpha\beta}, g_{\alpha a}, g_{ab}$ are constant functions on TQ/G. Notice that this Lagrangian is cyclic in the Abelian variables θ^a and the controls act only on the cyclic variables. Let

[2]In Bloch, Leonard, and Marsden [2001], they used H for the symmetry group. For the sake of consistency, we used G for the symmetry group in this thesis.

$(l_{\tau,\sigma,\rho}, 0, T^*H)$ be another RCL system with the reduced Lagrangian of the following form:

$$\begin{aligned} l_{\tau,\sigma,\rho} &= l(\eta^\alpha, \dot\theta^a + \tau^a_\alpha \eta^\alpha) + \frac{1}{2}\sigma_{ab}\tau^a_\alpha \tau^b_\beta \eta^\alpha \eta^\beta \\ &+ \frac{1}{2}(\rho_{ab} - g_{ab})(\dot\theta^a + g^{ac}g_{c\alpha}\eta^\alpha + \tau^a_\alpha\eta^\alpha)(\dot\theta^b + g^{bc}g_{c\beta}\eta^\beta + \tau^b_\beta\eta^\beta), \end{aligned} \tag{4.22}$$

which is exactly equation (11) in Bloch, Leonard, and Marsden [2001]. See also Bloch, Leonard, and Marsden [2001] for the motivation of this choice of the form in (4.22). Bloch, Leonard, and Marsden [2001] assumes the so-called Euler-Poincaré matching conditions:

EP-1: $\tau^a_\alpha = -\sigma^{ab} g_{b\alpha}$,

EP-2: $\sigma^{ab} + \rho^{ab} = g^{ab}$.

Then, one can show that the two assumptions of **EP-1** and **EP-2** imply the RCL-equivalence of the two RSCL systems $(l, 0, T^*H)$ and $(l_{\tau,\sigma,\rho}, 0, T^*H)$. Hence, one can equivalently work with the second system to design controllers.

4.1.4 Example: Heavy Top with Two Rotors

We apply the Euler-Poincaré matching condition given in § 4.1.3 to the system of a heavy top with two rotors. Strictly speaking, this system does not fall into the class of systems for which the Euler-Poincaré matching was originally developed in Bloch, Leonard, and Marsden [1998, 2001]. We show here that the same matching conditions can be used for more general systems such as a heavy top with two rotors. This work was presented in Chang and Marsden [2000]. For the purpose of concreteness, we keep the original style in Chang and Marsden [2000].

Euler-Poincaré Matching. In this section, we address the method of controlled Lagrangians for the (general) Euler-Poincaré equations. The Euler-Poincaré matching conditions are found in Bloch, Leonard, and Marsden [1998] for pure Euler-Poincaré equations. Here we apply the same conditions to the general Euler-Poincaré equations. See Holm, Marsden, and Ratiu [1998] for more detail about Euler-Poincaré equations.

Assume that there is a left representation of a Lie group G on a vector space V. Let H be an abelian Lie group on which G acts trivially. Let $L : TG \times V^* \times TH \to \mathbb{R}$ be a G-invariant function. We consider the class of mechanical systems whose kinetic energy depends on $TG \times TH$ and potential energy on V^*. The left G-invariance of L allows us to define the reduced Lagrangian $l : \mathfrak{g} \times V^* \times TH \to \mathbb{R}$ by $l(g^{-1}v_g, g^{-1}x, w) = L(v_g, x, w)$

for $(v_g, x, w) \in TG \times V^* \times TH$ and $g \in G$, where $\mathfrak{g}$ is the Lie algebra of G. In coordinates

$$l(\eta^\alpha, x_\alpha, \dot\theta^a) = \frac{1}{2} g_{\alpha\beta}\eta^\alpha\eta^\beta + g_{\alpha a}\eta^\alpha\dot\theta^a + \frac{1}{2} g_{ab}\dot\theta^a\dot\theta^b - U(x_\alpha) \tag{4.23}$$

for $(\eta^\alpha, x_\alpha, \dot\theta^a) \in \mathfrak{g} \times V^* \times TH$. We assume that $g_{\alpha\beta}, g_{\alpha a}$ and g_{ab} are constant and that the controls u_a act in the θ^a directions, i.e., $W/G = T^*H = \langle \mathbf{d}\theta \rangle$. The equations of motion of the closed-loop RCL system $(l, 0, u)$ are given by

$$\frac{d}{dt}\frac{\partial l}{\partial \eta} = \mathrm{ad}^*_\eta \frac{\partial l}{\partial \eta} + \frac{\partial l}{\partial x} \diamond x \tag{4.24}$$

$$\frac{d}{dt}\frac{\partial l}{\partial \dot\theta} = u \tag{4.25}$$

with

$$\dot x(t) = -\eta(t) x(t) \tag{4.26}$$

where the bilinear map $\diamond : V \times V^* \to \mathfrak{g}^*$ is defined by

$$\langle \eta x, v \rangle = -\langle v \diamond x, \eta \rangle$$

for $v \in V, x \in V^*$ and $\eta \in \mathfrak{g}$. See Holm, Marsden, and Ratiu [1998] for the derivation of (4.24) – (4.26).

Consider the following form of RSCL system $(l_{\tau,\sigma,\rho}, 0, W/G)$, where

$$\begin{aligned} l_{\tau,\sigma,\rho}(\eta^\alpha, x_\alpha, \dot\theta^a) = {} & \frac{1}{2}(g_{\alpha\beta} - g_{\alpha a} g^{ab} g_{\beta b} + \sigma_{ab}\tau^a_\alpha \tau^b_\beta)\eta^\alpha\eta^\beta \\ & + \frac{1}{2}\rho_{ab}(\dot\theta^a + (g^{ac} g_{\alpha c} + \tau^a_\alpha)\eta^\alpha)(\dot\theta^b + (g^{bd} g_{\beta d} + \tau^b_\beta)\eta^\beta) \\ & - U(x_\alpha) \end{aligned} \tag{4.27}$$

and

$$W/G = T^*H = \langle \mathbf{d}\theta \rangle.$$

Define the momentum $\tilde J_a$ conjugate to θ^a by

$$\tilde J_a = \frac{\partial l_{\tau,\sigma,\rho}}{\partial \dot\theta^a} = \rho_{ab}(\dot\theta^b + (g^{bd} g_{\beta d} + \tau^b_\beta)\eta^\beta). \tag{4.28}$$

We wish to transform the equations in (4.24)–(4.26), by an appropriate feedback u, to the following controlled Euler-Poincaré equations of $l_{\tau,\sigma,\rho}$:

$$\frac{d}{dt}\frac{\partial l_{\tau,\sigma,\rho}}{\partial \eta} = \mathrm{ad}^*_\eta \frac{\partial l_{\tau,\sigma,\rho}}{\partial \eta} + \frac{\partial l_{\tau,\sigma,\rho}}{\partial x} \diamond x \tag{4.29}$$

$$\frac{d}{dt}\frac{\partial l_{\tau,\sigma,\rho}}{\partial \dot\theta} = v \tag{4.30}$$

with

$$\dot{x}(t) = -\eta(t)x(t). \tag{4.31}$$

In other words, we want to find a condition such that

$$(l, 0, T^*H) \overset{L}{\sim} (l_{\tau,\sigma,\rho}, 0, T^*H).$$

The following is the the Euler-Poincaré matching conditions from Bloch, Leonard, and Marsden [1999a] and Bloch, Leonard, and Marsden [2001]:

EP-1: $\tau^a_\alpha = -\sigma^{ab} g_{b\alpha}$,

EP-2: $\sigma^{ab} + \rho^{ab} = g^{ab}$.

We can then prove the following theorem along the same lines as the proof in Bloch, Leonard, and Marsden [2001].

Proposition 4.1.10. *Under the assumptions* **EP-1** *and* **EP-2**, *the Euler-Poincaré equations in (4.29)-(4.31) coincide with the Euler-Poincaré equations in (4.24)-(4.26) with the following choice of the control u:*

$$u_a = u_a^{\mathrm{cons}} + (g_{ab} - k^\alpha_a g_{\alpha b})\rho^{bc} v_c,$$

where

$$\begin{aligned} u_a^{\mathrm{cons}} &= k^\alpha_a \left(c^\psi_{\alpha\delta}\eta^\delta (g_{\psi\beta}\eta^\beta + g_{\psi b}\dot\theta^b) - d^\psi_{\alpha\delta}\frac{\partial U}{\partial x_\delta} x_\psi \right), \\ k^\alpha_a &= D_{ab}\sigma^{bc} g_{c\beta} B^{\alpha\beta}, \\ B_{\alpha\beta} &= g_{\alpha\beta} - g_{\alpha b} g^{ab} g_{a\beta}, \\ D^{ba} &= g^{ba} + \sigma^{bc} g_{c\beta} B^{\alpha\beta} g_{\alpha e} g^{ae}, \end{aligned}$$

where $c^\psi_{\alpha\delta}$ are the structure constants of the Lie algebra $\mathfrak{g}$ and $d^\psi_{\alpha\delta}$ are the coordinate expression of the bilinear map $\diamond : V \times V^* \to \mathbb{R}$.

Remark 4.1.11. *In applications, we are not usually interested in the θ variables. In such cases, we regard $\mathfrak{g} \times V^* \times \mathfrak{h}$ as a phase space by ignoring H variables and identifying T_hH with the Lie algebra $\mathfrak{h}$ of H for each $h \in H$.*

Asymptotic Stabilization of the Heavy Top. It is well known in mechanics that the upright spinning top is unstable if the angular velocity is small. The motion of the heavy top and the stability of the Lagrange top are well studied in Marsden and Ratiu [1999] and Holm, Marsden, and Ratiu [1998]. In this section, we use the controlled Lagrangian method to asymptotically stabilize the upright spinning motion of the heavy top with small angular velocity, including zero velocity.

We first describe the heavy top with two rotors. We mount two rotors within the top so that each rotor's rotation axis is parallel to the first and the second principal axes of the top; see Figure 4.3. Let I_1, I_2, I_3 be the moments of inertia of the top in the body fixed frame. Let J_1, J_2 be the moments of inertia of the rotors around their rotation axes. Let J_{i1}, J_{i2}, J_{i3} be the moments of inertia of the ith rotor with $i = 1, 2$ around the first, the second and the third principal axes, respectively. Let $\bar{I}_1 = I_1 + J_{11} + J_{21}$, $\bar{I}_2 = I_2 + J_{12} + J_{22}$, and $\bar{I}_3 = I_3 + J_{13} + J_{23}$. Let $\lambda_1 = \bar{I}_1 + J_1$ and $\lambda_2 = \bar{I}_2 + J_2$. Let M be the total mass of the system, g the magnitude of the gravitational acceleration, and h the distance from the origin O to the center of mass of the system.

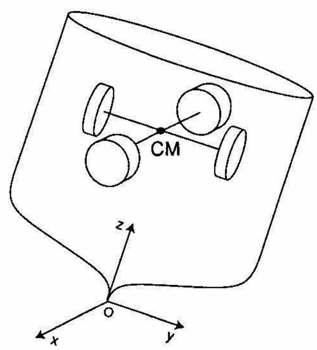

Figure 4.3: Heavy top with two rotors, each consisting of two rigidly coupled disks. The center of mass is at CM.

In this example, we have $G = \mathrm{SO}(3)$, $V^* = \mathbb{R}^3$ and $H = T^2 = S^1 \times S^1$. Let $\Omega = (\Omega_1, \Omega_2, \Omega_3) \in \mathfrak{so}(3) \cong \mathbb{R}^3$ be the angular velocity of the top in the body fixed frame. The vector $\Gamma = (\Gamma_1, \Gamma_2, \Gamma_3)$ represents the motion of the *unit vector* with the *opposite direction* of gravity as seen from the body. The coordinates $\theta = (\theta_1, \theta_2)$ are the rotation angles of rotors around their axes. Then the reduced Lagrangian $l : \mathfrak{so}(3) \times \mathbb{R}^3 \times TT^2 \to \mathbb{R}$

is given by

$$l(\Omega,\Gamma,\dot\theta)=\frac{1}{2}\begin{bmatrix}\Omega_1\\ \Omega_2\\ \Omega_3\\ \dot\theta_1\\ \theta_2\end{bmatrix}^T\begin{bmatrix}\lambda_1 & 0 & 0 & J_1 & 0\\ 0 & \lambda_2 & 0 & 0 & J_2\\ 0 & 0 & \bar I_3 & 0 & 0\\ J_1 & 0 & 0 & J_1 & 0\\ 0 & J_2 & 0 & 0 & J_2\end{bmatrix}\begin{bmatrix}\Omega_1\\ \Omega_2\\ \Omega_3\\ \dot\theta_1\\ \theta_2\end{bmatrix}-Mgh\Gamma_3. \tag{4.32}$$

The angular momentum $\Pi=(\Pi_1,\Pi_2,\Pi_3)$ is

$$\Pi=\frac{\partial l}{\partial\Omega}=(\lambda_1\Omega_1+J_1\dot\theta_1,\ \lambda_2\Omega_2+J_2\dot\theta_2,\ \bar I_3\Omega_3). \tag{4.33}$$

The equations of motion are derived from (4.24)-(4.26) as follows:

$$\dot\Pi=\Pi\times\Omega+Mgh\Gamma\times\chi \tag{4.34}$$

$$\dot\Gamma=\Gamma\times\Omega \tag{4.35}$$

$$\frac{d}{dt}\frac{\partial l}{\partial\dot\theta_i}=u_i \tag{4.36}$$

for $i=1,2$ where $\chi=(0,0,1)$ and u_i's are the control torques acting on the rotors.

Choose the following diagonal matrix form of σ_{ab} and ρ_{ab} in (4.27):

$$[\sigma_{ab}]=\begin{bmatrix}J_1\sigma_1 & 0\\ 0 & J_2\sigma_2\end{bmatrix},\qquad [\rho_{ab}]=\begin{bmatrix}J_1\rho_1 & 0\\ 0 & J_2\rho_2\end{bmatrix}.$$

In this case, the two matching conditions EP-1 and EP-2 become

$$[\tau^a_\theta]=\begin{bmatrix}-\frac{1}{\sigma_1} & 0 & 0\\ 0 & -\frac{1}{\sigma_2} & 0\end{bmatrix},$$

and

$$1=\frac{1}{\sigma_i}+\frac{1}{\rho_i}$$

for $i=1,2$. The Lagrangian $l_{\tau,\sigma,\rho}$ in (4.27) is computed as

$$\begin{aligned}l_{\tau,\sigma,\rho}(\Omega,\Gamma,\dot\theta)&=\frac{1}{2}\left(\lambda_1-\frac{J_1}{\rho_1}\right)\Omega_1^2+\frac{1}{2}\left(\lambda_2-\frac{J_2}{\rho_2}\right)\Omega_2^2+\frac{1}{2}\bar I_3\Omega_3^2+\frac{1}{2}J_1\rho_1\left(\frac{\Omega_1}{\rho_1}+\dot\theta_1\right)^2\\ &\quad+\frac{1}{2}J_2\rho_2\left(\frac{\Omega_2}{\rho_2}+\dot\theta_2\right)^2-Mgh\Gamma_3.\end{aligned} \tag{4.37}$$

The momentum $\tilde{J} = (\tilde{J}_1, \tilde{J}_2)$ conjugate to $\theta = (\theta_1, \theta_2)$ is

$$\tilde{J}_i = \frac{\partial l_{\tau,\sigma,\rho}}{\partial \dot{\theta}_i} = J_i \Omega_i + J_i \rho_i \dot{\theta}_i \tag{4.38}$$

with $i = 1, 2$. By Proposition 4.1.10, we have only to find an asymptotically stabilizing controller v_i for the following controlled Euler-Poincaré equations:

$$\dot{\Pi} = \Pi \times \Omega + Mgh\Gamma \times \chi \tag{4.39}$$

$$\dot{\Gamma} = \Gamma \times \Omega \tag{4.40}$$

$$\dot{\tilde{J}}_1 = v_1 \tag{4.41}$$

$$\dot{\tilde{J}}_2 = v_2, \tag{4.42}$$

where Π is the same as that in (4.33) by **EP-1** and **EP-2**. We have two constants of motion; $\Pi \cdot \Gamma$ and $||\Gamma||^2$.

Let $\Omega(0), \Gamma(0)$ and $\dot{\theta}(0)$ with $||\Gamma(0)||^2 = 1$ be an initial condition with

$$\Omega_3^\circ := \Pi(0) \cdot \Gamma(0)/\bar{I}_3 < \sqrt{Mgh/\bar{I}_3}. \tag{4.43}$$

As mentioned in Remark 4.1.11, we ignore the θ^a variables and regard $\mathfrak{so}(3) \times \mathbb{R}^3 \times \mathbb{R}^2$ as a phase space. We are interested in the equilibrium $e = (\Omega_e, \Gamma_e, \dot{\theta}_e)$:

$$\Omega_e = (0, 0, \Omega_3^\circ), \;\; \Gamma_e = (0, 0, 1), \;\; \dot{\theta}_e = (0, 0) \tag{4.44}$$

or

$$\Omega_e = (0, 0, \Omega_3^\circ), \;\; \Gamma_e = (0, 0, 1), \;\; \dot{\tilde{J}}_e = (0, 0),$$

which corresponds to the upright spinning top with the rotors at rest. Notice that this equilibrium lies in the same level set of $(\Pi \cdot \Gamma, ||\Gamma||^2)$ as the initial condition.

We construct a Lyapunov function using the energy-Casimir method (see Bloch, Chang, Leonard, Marsden, and Woolsey [2000] for more detail of this method). Set

$$E_{\tilde{\Phi}} = K_{\tau,\sigma,\rho} + U + \Phi(\Pi \cdot \Gamma, ||\Gamma||^2) + \Psi(\tilde{J}_1, \tilde{J}_2) \tag{4.45}$$

where $U(\Gamma) = Mgh\Gamma_3$ and $K_{\tau,\sigma,\rho}$ is given by

$$K_{\tau,\sigma,\rho} = \frac{1}{2}\left(\lambda_1 - \frac{J_1}{\rho_1}\right)\Omega_1^2 + \frac{1}{2}\left(\lambda_2 - \frac{J_2}{\rho_2}\right)\Omega_2^2 + \frac{1}{2}\bar{I}_3\Omega_3^2 + \frac{\tilde{J}_1^2}{2J_1\rho_1} + \frac{\tilde{J}_2^2}{2J_2\rho_2},$$

which is the kinetic energy, consisting of the quadratic terms in (4.37) in the new coordinates $(\Omega, \Gamma, \tilde{J})$. Choose the function Ψ as follows

$$\Psi(\tilde{J}_1, \tilde{J}_2) = \frac{\tilde{J}_1^2}{2\epsilon_1 J_1} + \frac{\tilde{J}_2^2}{2\epsilon_2 J_2}, \tag{4.46}$$

where coefficients ϵ_i will be determined later. Choose the function Φ of the form

$$\begin{aligned}\Phi(x, y) = &-\Omega_3^\circ(x - \bar{I}_3\Omega_3^\circ) + \frac{1}{2}\left(\bar{I}_3(\Omega_3^\circ)^2 - Mgh\right)(y-1)\\ &+ \frac{1}{2}a_1(x - \bar{I}_3\Omega_3^\circ)^2 + \frac{1}{2}a_2(y-1)^2,\end{aligned}$$

where the constants a_1 and a_2 are chosen such that

$$a_1 < -1/\bar{I}_3$$

and

$$4a_2 + a_1(\bar{I}_3\Omega_3^\circ)^2 + \bar{I}_3(\Omega_3^\circ)^2 - Mgh < \frac{\bar{I}_3(a_1\bar{I}_3\Omega_3^\circ - \Omega_3^\circ)^2}{1 + a_1\bar{I}_3}.$$

One can check that the equilibrium e is a critical point of $E_{\tilde{\Phi}}$. We now find conditions under which this critical point is a local maximum.[3] First, choose ρ_i satisfying

$$\frac{\bar{I}_3(\Omega_3^\circ)^2 - Mgh}{(\Omega_3^\circ)^2} < \lambda_i - \frac{J_i}{\rho_i} < 0 \tag{4.47}$$

for $i = 1, 2$, and then we can choose ϵ_1 and ϵ_2 such that the second derivative of $E_{\tilde{\Phi}}$ becomes negative definite at e, which implies that $E_{\tilde{\Phi}}$ has a local maximum at e. For later use, we impose an additional condition on ρ_i and ϵ_i as follows:

$$J_i(\Omega_3^\circ)^2 + (\epsilon_i + \rho_i)\left((\Omega_3^\circ)^2(\bar{I}_3 - \lambda_i) - Mgh\right) \neq 0. \tag{4.48}$$

With (4.48), it is still possible to find ρ_i and ϵ_i to ensure negative definiteness of the second derivative of $E_{\tilde{\Phi}}$ at e.

The following choice of $v = (v_1, v_2)$

$$v_i = c_i\left(\dot{\theta}_i + \frac{\tilde{J}_i}{\epsilon_i J_i}\right) \tag{4.49}$$

with $c_i > 0$ for $i = 1, 2$, implies

$$\frac{d}{dt}E_{\tilde{\Phi}} = \sum_{i=1}^{2} c_i\left(\dot{\theta}_i + \frac{\tilde{J}_i}{\epsilon_i J_i}\right)^2 \geq 0, \tag{4.50}$$

[3] See Remark 2.1.10.

which proves the Lyapunov stability of the equilibrium e in the closed-loop system. The complete control law u can be obtained from Proposition 4.1.10. Asymptotic stabilization will now be shown by using LaSalle's theorem. Since $E_{\tilde{\Phi}}$ has a local maximum at e, it is nondecreasing in time, and $\Pi \cdot \Gamma$ and $||\Gamma||^2$ are conserved, there is a number c such that the set $S = \{x \in \mathfrak{so}(3) \times \mathbb{R}^3 \times \mathbb{R}^2 | E_{\tilde{\Phi}} \geq c,\ \Pi \cdot \Gamma = \Pi_e \cdot \Gamma_e,\ ||\Gamma||^2 = 1\}$ is non-empty, compact, and positively invariant. Define $\mathcal{E} = \{x \in S | \dot{E}_{\tilde{\Phi}} = 0\} = \{x \in S | v = 0\}$. Let $\mathcal{M}$ be the largest invariant subset of $\mathcal{E}$. One can show $\mathcal{M} = \{e\}$ by (4.48) after shrinking the set S if necessary. Thus, by LaSalle's theorem, e is asymptotically stable.

Here is the brief proof of showing $\mathcal{M} = \{e\}$. Let $(\Omega(t), \Gamma(t), \dot{\theta}(t))$ be a trajectory in $\mathcal{M}$. The condition $v_i = 0$, (4.41), and (4.42) imply that $\tilde{J}_i(t)$ is constant. Hence, $\theta_i(t), \Omega_i(t)$ are constant for $i = 1, 2$. By (4.33), $\Pi_i(t)$, $i = 1, 2$ are constant. Then the third component of (4.39) becomes $\lambda_3 \dot{\Omega}_3(t) =$ constant. By the Lyapunov stability of the equilibrium, it follows $\dot{\Omega}_3(t) \equiv 0$. Hence, $\Omega_3(t)$ is constant. The first and second compoent of (4.39) implies that $\Gamma_1(t), \Gamma_2(t)$ are constant. Then, the third component of (4.40) implies that $\Gamma_3(t)$ is constant. So far we have shown that the trajectory $(\Omega(t), \Gamma(t), \dot{\theta}(t))$, or $(\Pi(t), \Gamma(t), \dot{\theta}(t))$ is constant for all $t \geq 0$. Consider the map $f : \mathbb{R}^8 \to \mathbb{R}^{10}$ defined by

$$f(\Omega, \Gamma, \dot{\theta}) = \begin{bmatrix} \Pi \times \Omega + Mgh\Gamma \times \chi \\ \Gamma \times \Omega \\ J_1\Omega_1 + (\epsilon_1 + J_1\rho_1)\dot{\theta}_1 \\ J_2\Omega_2 + (\epsilon_2 + J_2\rho_2)\dot{\theta}_2 \\ \Pi \cdot \Gamma - \Pi_e \cdot \Gamma_e \\ ||\Gamma||^2 - 1 \end{bmatrix} = \begin{bmatrix} \dot{\Pi} \\ \dot{\Gamma} \\ \epsilon_1 v_1 \\ \epsilon_2 v_2 \\ \Pi \cdot \Gamma - \Pi_e \cdot \Gamma_e \\ ||\Gamma||^2 - 1 \end{bmatrix}$$

where Π is expressed in terms of $(\Omega, \dot{\theta})$ as in (4.33). Then, one can see that all the trajectories lying in $\mathcal{M}$ are contained in the set $f^{-1}(O)$. In particular, the equilibrium $(\Omega_e, \Gamma_e, \dot{\theta}_e)$ in (4.44) is also contained in $f^{-1}(O)$. One can check that the rank of the Jacobian matrix Df at the equilibrium is the full rank 8 by (4.48). Thus, f is locally one-to-one around the equilibrium by Theorem 4.12 in Boothby [1986]. Therefore, the only possible trajectory totally lying in $\mathcal{M}$ is the equilibrium only, if necessary, after shrinking the neighborhood of the equilibrium. It follows from LaSalle's theorem that the equilibrium is asymptotically stable.

Remark 4.1.12. *1 . The above procedure shows that the choice of control gains depends on the initial condition. This is unavoidable because we need to know the value of the constant of motion* $\Pi \cdot \Gamma$*, which the internal actuation cannot change; however, our suggested controller is robust to small errors in the measurement of the initial condition. Let* $\tilde{e}$ *be the equilibrium of the form (4.44) with* $\tilde{\Omega}_3^\circ$ *instead of* Ω_3°*. Suppose the* Ω_3° *used in* *constructing the control law is very close to the value* $\tilde{\Omega}_3^\circ$. *Let* $\tilde{E}_{\tilde{\Phi}}$ *be the function of the form (4.45), with* Ω_3° *replaced by* $\tilde{\Omega}_3^\circ$. *Then* $\tilde{e}$ *is a critical point of* $\tilde{E}_{\tilde{\Phi}}$. *By continuity, the second derivative of* $\tilde{E}_{\tilde{\Phi}}$ *at* $\tilde{e}$ *will remain negative definite, proving Lyapunov stability of* $\tilde{e}$.

2. The same form of controller works for the asymptotic stabilization of the upright spinning top with $\Omega_3^\circ > \sqrt{Mgh/\bar{I}_3}$, *which is the opposite of (4.43). All that needs to be done is to choose* ρ_i *and* ϵ_i *to make* $E_{\tilde{\Phi}}$ *have a local minimum at the equilibrium and to choose negative* c_i *such that* $E_{\tilde{\Phi}}$ *decreases in time. The same LaSalle invariance principle argument guarantees asymptotic stability.*

4.2 Reduction of Controlled Hamiltonian Systems with Symmetry

We study the reduction of CH systems with symmetry. We define G invariant CH systems and reduced CH systems. We introduce an equivalence relation, called the RCH-equivalence, among the reduced CH systems by reducing the CH-equivalence relation for G invariant CH systems.

4.2.1 Reduction of CH Systems with Symmetry

We defined controlled Hamiltonian systems on T^*Q in Definition 3.1.1. Here we define G invariant CH systems on T^*Q and reduced CH systems on T^*Q/G.

Definition 4.2.1. *Let* G *be a Lie group acting on* Q. *A* G *invariant controlled Hamiltonian (G-CH) system is a CH system,* (H, B, F, W), *where* H, B, F *and* W *are* G *invariant.*

Definition 4.2.2. *A reduced controlled Hamiltonian (RCH) system is a quadruple*

$$(h, b, f, U),$$

where $h : T^*Q/G \to \mathbb{R}$ *is a smooth function called the reduced Hamiltonian,*

$$b \in \Gamma(\wedge^2 T(T^*Q/G))$$

is called a reduced almost Poisson tensor, $f : T^*Q/G \to T^*Q/G$ *is a fiber-preserving map called the reduced force map, and* U *is a subbundle of* T^*Q/G, *called the reduced control bundle.*

Suppose that we are given a G-CH system (H, B, F, W) on T^*Q. The G invariant

$H : T^*Q \to \mathbb{R}$ induces the reduced Hamiltonian $h : T^*Q/G \to \mathbb{R}$ as follows:

$$H = h \circ \pi_{/G}. \tag{4.51}$$

The G invariance of the Poisson tensor $B \in \wedge^2 TT^*Q$ induces a reduced Poisson tensor $[B]_G \in \wedge^2 T(T^*Q/G)$ as follows: for $f_1, f_2 \in \mathcal{F}(T^*Q/G)$,

$$[B]_{G[q,p]_G}(\mathbf{d}f_1, \mathbf{d}f_2) = B_{(q,p)}(\pi^*_{/G}\mathbf{d}f_1, \pi^*_{/G}\mathbf{d}f_2). \tag{4.52}$$

This is well defined since

$$\begin{aligned} B_{g(q,p)}(\mathbf{d}(f_1 \circ \pi_{/G}), \mathbf{d}(f_2 \circ \pi_{/G})) &= B_{(q,p)}(g^*\mathbf{d}(f_1 \circ \pi_{/G}), g^*\mathbf{d}(f_2 \circ \pi_{/G})) \\ &= B_{(q,p)}(\mathbf{d}(f_1 \circ \pi_{/G} \circ g), \mathbf{d}(f_2 \circ \pi_{/G} \circ g)) \\ &= B_{(q,p)}(\mathbf{d}(f_1 \circ \pi_{/G}), \mathbf{d}(f_2 \circ \pi_{/G})) \end{aligned}$$

for any $g \in G$ where we used the G invariance of B in the first equality. One can easily check that $[B]_G$ is skew-symmetric. The G invariance of F induces the reduced force $[F]_G : T^*Q/G \to T^*Q/G$ satisfying

$$[F]_G \circ \pi_{/G} = \pi_{/G} \circ F. \tag{4.53}$$

This discussion motivates the following definition:

Definition 4.2.3. *The RCH system of a G-CH system (H, B, F, W) is a quadruple $(h, [B]_G, [F]_G, W/G)$ where h is the reduced Hamiltonian defined in (4.51), $[B]_G$ is the reduced almost Poisson tensor defined in (4.52), and $[F]_G$ is the reduced force defined in (4.53).*

Analogous to Proposition 4.1.4, the following proposition explains the relationship between G-CH systems on T^*Q and RCH systems on T^*Q/G.

Proposition 4.2.4. *Given a RCH system (h, b, f, U), there is a (not necessarily unique) G-CH system (H, B, F, W) whose RCH system is (h, b, f, U).*

Proof. Define H by $H = h \circ \pi_{/G}$. Define a force map F on T^*Q as follows: for $\alpha_q \in T^*_qQ$, $v_q \in T_qQ$

$$\langle F(\alpha_q), v_q \rangle = \langle f \circ \pi_{/G}(\alpha_q), \tau_{/G}(v_q) \rangle.$$

Choose a connection on the principal bundle $\tau_{/G} : TQ \to TQ/G$ (see Chapter 2, Theorem 2.1 in Kobayashi and Nomizu [1963] for the proof of the existence). Then, we can split TT^*Q into the vertical space V and the horizontal space H as $TT^*Q = \mathrm{V} \oplus \mathrm{H}$. This induces the decomposition of T^*T^*Q as $T^*T^*Q = \mathrm{H}^\circ \oplus \mathrm{V}^\circ$, where H° and V° are the annihilators of H and V, respectively. Let hor : $T(T^*Q/G) \to \mathrm{H}$ be the horizontal lift. Then, its dual map hor* : $\mathrm{V}^\circ \to T^*(T^*Q/G)$ is an isomorphism. For simplicity, we use H° (resp. V°) as the projection of T^*T^*Q onto H° (resp. V°). Define an almost Poisson tensor B on T^*Q as follows: for $\alpha, \beta \in T_p^*T^*Q$

$$B(\alpha, \beta) := b(\mathrm{hor}^*\, \mathrm{V}^\circ\, \alpha,\, \mathrm{hor}^*\, \mathrm{V}^\circ\, \beta).$$

One can check that this almost Poisson tensor is G invariant. We now show that $\pi_{/G}$: $T^*Q \to T^*Q/G$ is the Poisson map, i.e., $b = [B]_G$. Let h_1, h_2 be two functions on T^*Q/G. Then, $\mathbf{d}(h_i \circ \pi_{/G}) \in V^\circ$, $i = 1, 2$. So, hor* $\mathbf{d}(h_i \circ \pi_{/G}) = \mathbf{d}h_i$, $i = 1, 2$. Hence,

$$\begin{aligned} [B]_G(\mathbf{d}h_1,\, \mathbf{d}h_2) &= B(\mathbf{d}(h_1 \circ \pi_{/G}),\, \mathbf{d}(h_2 \circ \pi_{/G})) \\ &= b(\mathbf{d}h_1,\, \mathbf{d}h_2). \end{aligned}$$

It follows that $[B]_G = b$. Let $W = \pi_{/G}^{-1}(U)$. Then, one can see that (H, B, F, W) is a G-CH system and its RCH system coincides with (h, b, f, U). This completes the proof. ■

By Proposition 4.2.4, we can, without loss of generality, write an arbitrary RCH system in the form of the RCH system of a G-CH system.

Given a G-CH system (H, B, F, W), let $(h, [B]_G, [F]_G, W/G)$ be its RCH system. The (reduced) Hamiltonian vector field of $(h, [B]_G, [F]_G, W/G)$ with a control $[u]_G \in W/G$ is given by

$$X_{(h,[B]_G,[F]_G,[u]_G)} = [B]_G^\sharp \mathbf{d}h + \mathrm{vlift}([F]_G) + \mathrm{vlift}([u]_G),$$

where vlift$([F]_G)$ and vlift$([u]_G)$ are the vertical lifts defined in (3.5). Let $X_{(H,B,F,u)}$ be the vector field of (H, B, F, W) with control $u \in W$. Then, we have

$$X_{(h,[B]_G,[F]_G,[u]_G)} \circ \pi_{/G} = T\pi_{/G} \cdot X_{(H,B,F,u)}. \tag{4.54}$$

4.2.2 Reduced CH Equivalence

First recall that we defined controlled Hamiltonian equivalence relation in Definition 3.1.3. We now introduce an equivalence relation among RCH systems on T^*Q/G.

Definition 4.2.5. *Two RCH systems, $(h_i, [B_i]_G, [F_i]_G, W_i/G)$, $i = 1, 2$, are said to be* ***reduced-CH-equivalent (RCH-equivalent)****, or simply,*

$$(h_1, [B_1]_G, [F_1]_G, W_2/G) \overset{H}{\sim} (h_2, [B_2]_G, [F_2]_G, W_2/G)$$

if the following ***reduced Hamiltonian matching conditions*** *hold:*

RHM-1 : $W_1/G = W_2/G$,

RHM-2 : $\mathrm{Im}[[B_1]^\sharp_G \mathbf{d}h_1 + \mathrm{vlift}([F_1]_G) - [B_2]^\sharp_G \mathbf{d}h_2 - \mathrm{vlift}([F_2]_G)] \subset \mathrm{vlift}(W_1/G)$

where $\mathrm{vlift}(W_1/G)$ *is the vertical lift of the subbundle* W_1/G *defined in (3.6).*

The following proposition explains the relation between the reduced-CH-equivalence relation among RCH systems and the CH-equivalence relation among G-CH systems.

Proposition 4.2.6. *Two G-CH systems are CH-equivalent if and only if their associated RCH systems are RCH-equivalent.*

Proof. Use Definition 3.1.3 and Definition 4.2.5 as well as the relation (4.54). ■

Proposition 4.2.7. *Suppose that two RCH systems $(h_i, [B_i]_G, [F_i]_G, W_i/G)$, $i = 1, 2$, are RCH-equivalent. Then, for an arbitrary control law for one system, there exists a control law for the other system such that the two closed-loop RCH systems produce the same equations of motion. The explicit relation between the two control laws $[u_i]_G$, $i = 1, 2$, is given by*

$$\mathrm{vlift}([u_1]_G) = -[B_1]^\sharp_G \mathbf{d}h_1 - \mathrm{vlift}([F_1]_G) + [B_2]^\sharp_G \mathbf{d}h_2 + \mathrm{vlift}([F_2]_G) + \mathrm{vlift}([u_2]_G).$$

Proof. Mimic the proof of Proposition 4.1.7. However, one has to use Proposition 3.1.4 and Proposition 4.2.6 instead. ■

We defined simple CH systems (or, ***SCH*** systems) in Definition 3.1.6. Unlike the reduced simple controlled Lagrangian systems, the notion of the reduced simple controlled Hamiltonian system is difficult to define because we need to define carefully the reduced simple almost Poisson tensor structure. Now, we adopt the following definition of the reduced simple CH system:

Definition 4.2.8. *A RCH system $(h, [B]_G, [F]_G, W/G)$ is called a reduced simple CH system (or, **RSCH** system) if it is the RCH system of a G invariant simple CH system.*

Recall that we defined simple almost Poisson tensors on T^*Q using local coordinates in Definition 3.1.6 (the definition of the simple almost Poisson tensor is independent of the cotangent bundle coordinates). Here, we characterize the reduced simple almost Poisson tensors using local coordinates, so we may assume that $Q = G \times X$ where G is a Lie group acting on the manifold X trivially. Recall the following identifications by left translation of G:

$$T^*G = G \times \mathfrak{g}^*, \quad TT^*G = (G \times \mathfrak{g}^*) \times (\mathfrak{g} \times \mathfrak{g}^*), \quad T^*T^*G = (G \times \mathfrak{g}^*) \times (\mathfrak{g}^* \times \mathfrak{g}).$$

We use (g_a, μ_a, x_i, p_i) as local coordinates for $T^*Q = G \times \mathfrak{g}^* \times T^*X$, and (μ, x, p) for $T^*Q/G = \mathfrak{g}^* \times T^*X$ where $a = 1, \ldots, \dim G$, and $i = 1, \ldots, \dim X$. We will use $\{e_a\}$ as a basis for $\mathfrak{g}$, and $\{e_a^*\}$ as its dual basis. Let $B \in \wedge^2 TT^*Q$ be a G invariant simple almost Poisson tensor. Then, it is of the following form:

$$\begin{aligned}B(g, \mu, x, p) = {} & A_{ab}(x)(e_a \otimes e_b^* - e_b^* \otimes e_a) + B_{ia}(x)(\partial x_i \otimes e_b^* - e_b^* \otimes \partial x_i) \\ & + C_{ai}(x)(e_a \otimes \partial p_i - \partial p_i \otimes e_a) + D_{ij}(x)(\partial_{x_i} \otimes \partial_{p_j} - \partial_{p_j} \otimes \partial_{x_i}) \\ & + R_{ab}(\mu, x, p) e_a^* \otimes e_b^* + S_{ai}(\mu, x, p)(e_a^* \otimes \partial p_i - \partial p_i \otimes e_a^*) \\ & + U_{ij}(\mu, x, p)\partial_{p_i} \otimes \partial_{p_j}. \end{aligned} \tag{4.55}$$

In the matrix form, B is given by

$$B = \begin{bmatrix} O & O & A(x) & C(x) \\ O & O & B(x) & D(x) \\ -A(x)^T & -B(x)^T & R(\mu, x, p) & S(\mu, x, p) \\ -C(x)^T & -D(x)^T & -S(\mu, x, p)^T & U(\mu, x, p) \end{bmatrix},$$

where we used the basis for $T_z T^*Q$ in the following order:

$$e_a,\ \partial_{x_i},\ e_a^*,\ \partial_{p_i}.$$

The non-degeneracy condition for B is given by

$$\operatorname{rank} \begin{bmatrix} A & C \\ B & D \end{bmatrix} = \dim Q.$$

The reduced simple Poisson tensor $[B]_G$ is given by

$$\begin{aligned}[B]_G(\mu, x, p) = {} & A_{ab}(x)(e_a \otimes e_b^* - e_b^* \otimes e_a) + B_{ia}(x)(\partial x_i \otimes e_b^* - e_b^* \otimes \partial x_i) \\ & + C_{ai}(x)(e_a \otimes \partial p_i - \partial p_i \otimes e_a) + D_{ij}(x)(\partial_{x_i} \otimes \partial_{p_j} - \partial_{p_j} \otimes \partial_{x_i}). \end{aligned} \tag{4.56}$$

In a matrix form,

$$[B]_G = \begin{bmatrix} R(\mu, x, p) & -B(x)^T & S(\mu, x, p) \\ B(x) & O & D(x) \\ -S(\mu, x, p)^T & -D(x)^T & U(\mu, x, p) \end{bmatrix},$$

where we used the basis for T_zT^*Q in the following order:

$$e_a^*,\ \partial_{x_i},\ \partial_{p_i}.$$

The non-degeneracy condition for B induces the following rank condition for $[B]_G$:

$$\operatorname{rank}[B\ \ D] = \dim X. \tag{4.57}$$

Remark 4.2.9. *1. It would be better if we could characterize all the tensors*

$$b \in \Gamma(\wedge^2 T(T^*Q/G))$$

for which there exists a G-invariant simple almost Poisson tensor B such that $b = [B]_G$. Then, we can define reduced simple CH systems without reference to G invariant simple CH systems. This point has to be studied more and we think that the use of connections is crucial; see Montgomery, Marsden, and Ratiu [1984], and Cendra, Marsden, and Ratiu [2001].

2. One does not have to restrict to reduced simple almost Poisson tensors when applying to the method of reduced CH systems to design feedback controllers. One can relax the condition in (4.57) to obtain freedom in choosing reduced almost Poisson tensors.

4.3 Equivalence of CL Systems and CH Systems with Symmetry

We show that the method of reduced simple CL (simply, RSCL) systems is equivalent to that of reduced simple CH (simply, RSCH) systems. Recall that a RSCL/RSCH system is the reduced CL/CH system of a G invariant simple CL/CH system. We will make use of the following lemma:

Lemma 4.3.1 (Corollary 3.2.2). *The method of controlled Lagrangians is equivalent to that of controlled Hamiltonians for simple mechanical systems in the following sense. For given two simple CL systems (L_i, F_i^L, W_i^L), $i = 1, 2$, there exist two simple CH systems (H_i, B_i, F_i^H, W_i^H) such that*

$$\begin{aligned}&(L_1, F_1^L, W_1^L) \overset{L}{\sim} (L_2, F_2^L, W_2^L) \quad \text{and} \quad \psi_{B_2}\circ\psi_{B_1}^{-1} = m_{H_2}(m_{H_1})^{-1}\\ &\iff (H_1, B_1, F_1^H, W_1^H) \overset{H}{\sim} (H_2, B_2, F_2^H, W_2^H)\end{aligned}$$

with m_{H_i} the mass tensor of H_i and ψ_{B_i} defined in (3.11) for $i = 1, 2$, and vice versa.

Now, we apply Lemma 4.3.1 to G-invariant simple systems.

Lemma 4.3.2. *For given two G invariant simple CL systems (L_i, F_i^L, W_i^L), $i = 1,2$, there exist two G-invariant simple CH systems (H_i, B_i, F_i^H, W_i^H) such that*

$$\begin{aligned}&(L_1, F_1^L, W_1^L) \overset{L}{\sim} (L_2, F_2^L, W_2^L) \quad \text{and} \quad \psi_{B_2} \circ \psi_{B_1}^{-1} = m_{H_2}(m_{H_1})^{-1} \\ &\Longleftrightarrow\ (H_1, B_1, F_1^H, W_1^H) \overset{H}{\sim} (H_2, B_2, F_2^H, W_2^H)\end{aligned} \tag{4.58}$$

with m_{H_i} the mass tensor of H_i and ψ_{B_i} defined in (3.11) for $i = 1,2$, and vice versa.

Proof. Keep track of the G invariance in the proof of Theorem 3.2.1. ■

Then, we have the following result.

Theorem 4.3.3. *The method of reduced controlled Lagrangian systems is equivalent to that of reduced controlled Hamiltonian systems for reduced simple mechanical systems in the following sense. For given two reduced simple CL systems $(l_i, [F_i^L]_G, W_i^L/G)$, $i = 1,2$, there exist two reduced simple CH systems*

$$(h_i, [B_i]_G, [F_i^H]_G, W_i^H/G), \quad i = 1,2$$

such that

$$\begin{aligned}&(l_1, [F_1^L]_G, W_1^L/G) \overset{L}{\sim} (l_2, [F_2^L]_G, W_2^L/G) \quad \text{and} \quad [\psi_{B_2} \circ \psi_{B_1}^{-1}]_G = [m_{H_2}]_G [m_{H_1}]_G^{-1} \\ &\Longleftrightarrow\ (h_1, [B_1]_G, [F_1^H]_G, W_1^H/G) \overset{H}{\sim} (h_2, [B_2]_G, [F_2^H]_G, W_2^H/G),\end{aligned} \tag{4.59}$$

with $[m_{H_i}]_G$ the reduced mass tensor of h_i and ψ_{B_i} defined in (3.11) for $i = 1,2$, and vice versa.

Proof. For given two RSCL systems $(l_i, [F_i^L]_G, W_i^L/G)$, $i = 1,2$, consider their unreduced G-SCL systems (L_i, F_i^L, W_i), $i = 1,2$ with $L_i = l_i \circ \tau_{/G}$ (see Proposition 4.1.4). By Lemma 4.3.2, there are two G-SCH systems (H_i, B_i, F_i^H, W_i^H), $i = 1,2$, such that (4.58) holds. Let $(h_i, [B_i]_G, [F_i^H]_G, W_i^H/G)$ be the RSCH system of (H_i, B_i, F_i^H, W_i^H). Then, (4.59) follows from Proposition 4.1.7, Proposition 4.2.6, and (4.58). For the case where one is given two reduced simple CH systems in the beginning, use Proposition 4.2.4 instead of Proposition 4.1.4, and then proceed in a similar manner. ■

Remark 4.3.4. *Notice that $\psi_{B_2} \circ \psi_{B_1}^{-1}$ is G equivariant even though each of ψ_{B_i} may not.*

This equivarience follows from the following commutative diagram:

$$
\begin{array}{ccccc}
TT^*Q & \xleftarrow{B_1} & T^*Q^*Q & \xrightarrow{B_2} & TT^*Q \\
\uparrow{\scriptstyle \text{vlift}} & & \uparrow{\scriptstyle \Theta} & & \uparrow{\scriptstyle \text{vlift}} \\
T^*Q & \xleftarrow{\psi_{B_1}} & T^*Q & \xrightarrow{\psi_{B_2}} & T^*Q
\end{array}
$$

It follows that for $\alpha \in T^*Q$

$$B_2 B_1^{-1}(\text{vlilft}(\alpha)) = \text{vlift}(\psi_{B_2} \circ \psi_{B_1}^{-1}(\alpha)).$$

One can easily check that vlift *is* G *equivariant, i.e.* $\text{vlift}(g\alpha) = g\,\text{vlift}(\alpha)$ *for* $g \in G$. *The* G *equivariance of* $\psi_{B_2} \circ \psi_{B_1}^{-1}$ *follows from the* G *equivariance of* B_1, B_2 *and* vlift, *and the injectivity of* vlift.

4.4 Summary and Future Work

We have studied the reduction of CL/CH systems with symmetry and showed that the method of reduced simple CL systems and that of reduced simple CH systems are equivalent. In the following, we summarize this chapter section by section.

§ 4.1. We defined G invariant CL systems and reduced CL systems (Definition 4.1.1, 4.1.2 and 4.1.3). For a reduced CL system (l, f, U), there exists a unique G invariant CL system (L, F, W) such that the reduced CL system of (L, F, W) is (l, f, U); see Proposition 4.1.4. The equations of motion of a reduced CL system are given in (4.4). We then defined reduced Euler-Lagrange matching conditions and RCL-equivalence relation for reduced simple CL systems (Definition 4.1.5 and 4.1.6). Proposition 4.1.7 shows that the RCL-equivalence relation is induced from the CL-equivalence relation. Proposition 4.1.8 is a reduced version of Proposition 2.1.5; If two reduced simple CL systems are RCL-equivalent, then for any choice of control for one system there exists a control for the other system such that the two closed-loop systems produce the same equations of motion. We applied this RCL equivalence to designing a controller which asymptotically stabilizes the rotation about a middle axis in the dynamics of the satellite with a rotor (§ 4.1.3). We also applied it to asymptotic stabilization of the upright slow rotation of the heavy top with two rotors (§ 4.1.4). Strictly speaking, the heavy top system does not fall into the category of reduced CL systems defined in Definition 4.1.2. So, we developed a new CL method *in coordinates* for the heavy top system. One needs to develop a general theory for this.

§ **4.2.** We defined G invariant CH systems and reduced CH systems (Definition 4.2.1, 4.2.2 and 4.2.3). For a reduced CH system (h, b, f, U), there exists a (not necessarily unique) G invariant CH system (H, B, F, W) such that the reduced CH system of (H, B, F, W) is (h, b, f, U); see Proposition 4.2.4. The non-uniqueness comes from the many possible choices of G invariant almost Poisson tensor B. We then defined reduced Hamiltonian matching conditions and RCH-equivalence relation for reduced CH systems (Definition 4.2.5). Proposition 4.2.6 shows that the RCH-equivalence relation is induced from the CH-equivalence relation. Proposition 4.2.7 is a reduced version of Proposition 2.1.5; if two reduced CH systems are RCH-equivalent, then for any choice of a control for one system there exists a control for the other system such that the two closed-loop systems produce the same equations of motion. The definition of reduced simple CH systems is a bit more subtle than that of reduced simple CL system because on the Hamiltonian side we have to choose an almost Poisson structure to write down the equations of motion whereas on the Lagrangian side, the variational principle uniquely determines them. In this section, we took the definition of reduced simple CH systems in Definition 4.2.8 and expressed reduced simple almost Poisson tensors in local coordinates.

§ **4.3.** We showed that the method of reduced simple CL systems and that of reduced simple CH systems are equivalent (Theorem 4.3.3). This is a reduced version of Theorem 3.2.1 and Corollary 3.2.2.

Future Work.

1. Notice that we could have defined the reduced simple CL system in Definition 4.1.5 without any references to G-invariant simple CL systems. Likewise, one might want to define reduced simple CH systems without any references to G invariant simple CH systems. For this purpose, one needs to address the following problem first: Find all the Poisson tensors $b \in \Gamma(\wedge^2 T(T^*Q/G))$ for which there exists a G *invariant simple* almost Poisson tensor $B \in \Gamma(\wedge^2 TT^*Q)$ such that $[B]_G = b$. We know the solution to this question only locally. One needs to study this globally. To this end, the use of connections will be important (Montgomery, Marsden, and Ratiu [1984]).
2. One needs to extend the reduced CL method to include systems such as the heavy top. Section 7.4 of Cendra, Marsden, and Ratiu [2001] will give a hint.

Chapter 5

Epilogue

We have developed the methods of CL/CH systems and showed the equivalence of the two methods for simple mechanical systems. In addition, we refined both methods to include systems with symmetry and discussed the relevant reduction theory. The CL method was applied to several systems in the thesis: the inverted pendulum on a cart, the spherical pendulum on a cart, the satellite with two rotors, and the heavy top with two rotors. We also found a class of mechanical systems for which the CL method can be applied for designing asymptotically stabilizing controllers. We believe that this method can be applied to various systems and can also be generalized to nonholonomic systems. Possible future directions are discussed at the end of each chapter.

Appendix A

Lyapunov-based Transfer between Elliptic Keplerian Orbits

In mechanics, symmetry gives rise to a conserved quantity. For example, energy is due to the time symmetry and linear momentum is due to the translation symmetry. In the presence of control forces, these quantities may or may not be conserved. However, they can be useful in designing controllers. For example, we made use of the energy for stabilization in the CL method by choosing dissipative feedback control laws to decrease the energy. Another example is the energy-Casimir method used in § 4.1.3 and 4.1.4. In this appendix, we use angular momentum and Laplace-Runge-Lenz vectors to design a feedback control for transfer between two elliptic Keplerian orbits. The angular momentum vector is due to rotational symmetry and the Laplace-Runge-Lenz vector is due to hidden rotational symmetry. This will illuminate the crucial role of geometric mechanics in the control of mechanical systems. This work was published by Chang, Chichka, and Marsden [2002].

A.1 Introduction

Low- and moderate-thrust transfer between satellite orbits in an inverse-square gravity field has been a topic of interest for decades. Some of the earliest work in this field is reviewed and extended by Edelbaum [1964, 1965] where low thrust transfer between elliptic Keplerian orbits was considered. Using variational calculus and considering the effects of thrust to be perturbations about an orbit, Edelbaum derived the optimal thrust histories to effect small changes in orbital elements. His later work extends this to achieve general transfers. More recent work, such as that surveyed in Chbotov [1996], has concentrated on finding optimal trajectories for fixed-time orbit transfer problems between general Keplerian orbits. Generally, the departure and injection points on the respective orbits are defined, as well as the elements of the orbits themselves. Optimal control theory then provides a two-point boundary-value problem, which may be solved to achieve the optimal thrust profile. The resulting calculations are lengthy, and do not lend themselves to closed-form solution or on-line implementation. For the special case of constant acceleration magnitude and fixed transfer time, some simplified results can be obtained.

Here, we present a study of the transfer between elliptic orbits about a spherical Earth, in which the final time is not specified and the injection point is free. We define the orbit at all times through the natural quantities of the angular momentum vector and the Laplace, or eccentricity vector. It is shown that every non-degenerate Keplerian orbit can be uniquely described by these two vectors, and conversely, that every such pair defines a unique orbit. We use the difference between current and desired final values of these vectors to define a Lyapunov function. This Lyapunov function gives an asymptotically stabilizing feedback controller such that the target elliptic Keplerian orbit becomes a locally asymptotically stable periodic orbit. We suggest another Lyapunov function for the transfer to circular orbits using the fact that a circular orbit is uniquely determined by its angular momentum (and energy).

A brief exposition of orbit transfer using a Lyapunov function was presented in Ilgen [1980], where the control is based on a function made up of the squares of the errors between the current and final orbital elements. That paper, however, does not provide a full analysis of the method, and convergence is not shown. Our work does provide a different Lyapunov function as well as a rigorous proof of the validity and convergence for the method presented.

The general method of using Lyapunov functions that are mechanically motivated has appeared in the literature before, such as in Åström and Furuta [1996], Bloch, Leonard, and Marsden [2000], and Bloch, Chang, Leonard, and Marsden [2001]. However, we believe that this paper is the first to apply such a general methodology to the problem of Keplerian orbit transfer.

A.2 Review of the Two-Body Problem

We give a review of some necessary concepts on the two-body problem (see Abraham and Marsden [1978], Cushman and Bates [1997], Goldstein [1980] among many others for more on orbital mechanics). The following is an abridged, modified and improved version of Chapter 2 in Cushman and Bates [1997].

The configuration space is $\mathbb{R}^3_0 := \mathbb{R}^3 - \{0\}$, i.e., $\mathbb{R}^3$ minus the origin. Let $T\mathbb{R}^3_0 = (\mathbb{R}^3 - \{0\}) \times \mathbb{R}^3$ be the tangent space of $\mathbb{R}^3_0$. We use $(\mathbf{r}, \dot{\mathbf{r}})$ as coordinates for $T\mathbb{R}^3_0$, and the over-dot as the derivative with respect to time t. The Keplerian equation of motion is given by

$$\ddot{\mathbf{r}} = -\mu \frac{\mathbf{r}}{\|\mathbf{r}\|^3} \tag{A.1}$$

where μ is the gravitational parameter. We refer to the solutions of (A.1) as *Keplerian flows* or *Keplerian orbits.* The energy $E : T\mathbb{R}^3_0 \to \mathbb{R}$ is defined by

$$E(\mathbf{r}, \dot{\mathbf{r}}) = \frac{1}{2}\|\dot{\mathbf{r}}\|^2 - \frac{\mu}{\|\mathbf{r}\|}. \tag{A.2}$$

Define $\pi = (\mathbf{L}, \mathbf{A}) : T\mathbb{R}^3_0 \to \mathbb{R}^3 \times \mathbb{R}^3$ by

$$\mathbf{L}(\mathbf{r}, \dot{\mathbf{r}}) = \mathbf{r} \times \dot{\mathbf{r}}, \tag{A.3}$$

$$\mathbf{A}(\mathbf{r}, \dot{\mathbf{r}}) = \dot{\mathbf{r}} \times (\mathbf{r} \times \dot{\mathbf{r}}) - \mu \frac{\mathbf{r}}{\|\mathbf{r}\|}, \tag{A.4}$$

where $\mathbf{L}$ is the angular momentum and $\mathbf{A}$ is the Laplace vector. The Laplace vector is occasionally referred to as the eccentricity vector (see Battin [1987]) because the two are identical, other than a scaling by μ. The three quantities E, $\mathbf{L}$, and $\mathbf{A}$ are constants of the motion of (A.1) and satisfy the following relations:

$$\mathbf{L} \cdot \mathbf{A} = 0, \tag{A.5}$$

$$\|\mathbf{A}\|^2 = \mu^2 + 2E\|\mathbf{L}\|^2, \tag{A.6}$$

where $\|\cdot\|$ is the usual Euclidean norm on $\mathbb{R}^3$.

Let $\mathbf{L}$ be the angular momentum of a Keplerian orbit $(\mathbf{r}(t), \dot{\mathbf{r}}(t))$. If $\mathbf{L} = 0$, then $(\mathbf{r}(t), \dot{\mathbf{r}}(t))$ is a degenerate orbit, i.e., $\mathbf{r}(t)$ moves in a straight line. If $\mathbf{L} \neq 0$, then $\mathbf{r}(t)$ traces an ellipse, a parabola, or a hyperbola, depending upon its energy E being negative, zero, or positive, respectively. We will exclude degenerate orbits from consideration. Hence the set

$$\Sigma_e = \{(\mathbf{r}, \dot{\mathbf{r}}) \in T\mathbb{R}^3_0 \mid E(\mathbf{r}, \dot{\mathbf{r}}) < 0, \mathbf{L}(\mathbf{r}, \dot{\mathbf{r}}) \neq 0\} \tag{A.7}$$

becomes the union of all elliptic Keplerian orbits. Define the set

$$D = \{(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^3 \times \mathbb{R}^3 \mid \mathbf{x} \cdot \mathbf{y} = 0, \mathbf{x} \neq 0, \|\mathbf{y}\| < \mu\}. \tag{A.8}$$

By (A.5)–(A.8), it follows that

$$\pi(\Sigma_e) \subset D \text{ and } \pi(T\mathbb{R}^3_0 - \Sigma_e) \cap D = \varnothing, \tag{A.9}$$

which implies

$$\pi^{-1}(D) = \Sigma_e. \tag{A.10}$$

For any $(\mathbf{x}, \mathbf{y}) \in D$, take

$$(\mathbf{r}, \dot{\mathbf{r}}) = \begin{cases} (\frac{-1}{2H}\frac{1-e}{e}\mathbf{y}, \frac{-2H}{\mu^2}\frac{1}{e(1-e)}\mathbf{x} \times \mathbf{y}) & \text{if } \mathbf{y} \neq 0 \\ (\frac{1}{-2H}(\mathbf{p} \times \mathbf{x}), \mathbf{p}) & \text{if } \mathbf{y} = 0, \end{cases}$$

where $H = (||\mathbf{y}||^2 - \mu^2)/(2||\mathbf{x}||^2)$, $e = ||\mathbf{y}||/\mu$, and $\mathbf{p}$ is a vector satisfying $\mathbf{p} \cdot \mathbf{x} = 0$ with $||\mathbf{p}|| = \sqrt{-2H}$. It is simple to show that $(\mathbf{r}, \dot{\mathbf{r}}) \in \Sigma_e$ and $\pi(\mathbf{r}, \dot{\mathbf{r}}) = (\mathbf{x}, \mathbf{y})$. This implies $D \subset \pi(\Sigma_e)$, which with (A.9) implies

$$\pi(\Sigma_e) = D. \tag{A.11}$$

Since $\mathbf{L}$ and $\mathbf{A}$ are constants of the motion of (A.1), equations (A.10) and (A.11) imply that $\pi^{-1}(\mathbf{x}, \mathbf{y})$ consists of a union of elliptic Keplerian orbits for each $(\mathbf{x}, \mathbf{y}) \in D$. Let $(\mathbf{r}(t), \dot{\mathbf{r}}(t))$ be any elliptic Keplerian orbit contained in $\pi^{-1}(\mathbf{L}, \mathbf{A}) \subset T\mathbb{R}^3_0$. Since $\mathbf{L}$ is normal to both $\mathbf{r}(t)$ and $\dot{\mathbf{r}}(t)$, the orbit $(\mathbf{r}(t), \dot{\mathbf{r}}(t))$ is contained in the set $\Pi \times \Pi$, where $\Pi \subset \mathbb{R}^3$ is the plane through the origin normal, to $\mathbf{L}$. The polar equation (r, θ) of the ellipse traced by $\mathbf{r}(t)$ on the plane Π is given by

$$r = \frac{||\mathbf{L}||^2}{\mu + ||\mathbf{A}|| \cos(\theta - \theta_0)} \tag{A.12}$$

where θ_0 is the polar angle of the periapsis when the orbit is a non-circular ellipse, i.e, when $\mathbf{A} \neq 0$. The tangent vector $\dot{\mathbf{r}}$ at $\mathbf{r}$ is derived from (A.3) and (A.4) as follows:

$$\dot{\mathbf{r}} = \frac{\mathbf{L}}{||\mathbf{L}||^2} \times \left(\mathbf{A} + \frac{\mu \mathbf{r}}{||\mathbf{r}||}\right).$$

It follows that $\pi^{-1}(\mathbf{L}, \mathbf{A})$ consists of a unique (oriented) elliptic Keplerian orbit for $(\mathbf{L}, \mathbf{A}) \in D$. Thus, we have proved the following Proposition.

Proposition A.2.1. *The following holds:*

1. *Σ_e is the union of all elliptic Keplerian orbits.*

2. *$\pi(\Sigma_e) = D$ and $\Sigma_e = \pi^{-1}(D)$.*

3. *The fiber $\pi^{-1}(\mathbf{x}, \mathbf{y})$ consists of a unique (oriented) elliptic Keplerian orbit for each $(\mathbf{x}, \mathbf{y}) \in D$.*

The following result follows directly from this.

Corollary A.2.2. *D is the space of elliptic Keplerian orbits.*

Another important consequence is the following.

Corollary A.2.3. *The set $\pi^{-1}(K)$ is a compact subset of Σ_e for any compact subset K of D.*

Proof. Take any compact set $K \subset D$. By Proposition A.2.1, $\pi^{-1}(K) \subset \Sigma_e$. Choose any sequence $\{a_k\} \subset \pi^{-1}(K)$. Let $b_k = \pi(a_k)$. Since K is compact, $\{b_k\}$ has a convergent subsequence. By passing to the subindex, we assume that $\{b_k\}$ is convergent to some $b \in K$. Then $\pi^{-1}(b)$ is compact since it is homeomorphic to the unit circle by Proposition A.2.1. By the continuity of π, the sequence $\{a_k\}$ converges to $\pi^{-1}(b)$. Choose a metric on Σ_e. Let $c_k \in \pi^{-1}(b)$ be a closest point from a_k to $\pi^{-1}(b)$ for each k. Since $\pi^{-1}(b)$ is compact and a distance function is continuous, the sequence $\{c_k\}$ is well defined. Since $\pi^{-1}(b)$ is compact $\{c_k\}$ has a convergent subsequence $\{c_{k_j}\}$ with a limit $c \in \pi^{-1}(b)$. One can see that $\{a_{k_j}\}$ converges to $c \in \pi^{-1}(b) \subset \pi^{-1}(K)$. Thus, $\pi^{-1}(K)$ is compact. ■

Remark A.2.4. *For notational simplicity, we will sometimes identify a point $(\mathbf{x}, \mathbf{y}) \in D$ with the set $\pi^{-1}(\mathbf{x}, \mathbf{y}) \subset \Sigma_e$.*

A.3 Main Results

Based on the results in the last section, we design a controller for orbital transfer between two arbitrary elliptic Keplerian orbits by constructing a suitable Lyapunov function. We consider first the case of local transfer, where the initial orbit is within a neighborhood of the target orbit. We then extend the results to transfer between two arbitrary elliptic orbits. Finally, we suggest another Lyapunov function for circular target orbits.

A.3.1 Local Orbit Transfer

We design here a Lyapunov-based controller to achieve asymptotically stable local orbit transfer. The equation of motion with a control force $\mathbf{F}$ is given by

$$\ddot{\mathbf{r}} = -\mu \frac{\mathbf{r}}{\|\mathbf{r}\|^3} + \mathbf{F}. \tag{A.13}$$

Define a metric d_k on $\mathbb{R}^3 \times \mathbb{R}^3$ by

$$\mathrm{d}_k((\mathbf{x}_1, \mathbf{y}_1), (\mathbf{x}_2, \mathbf{y}_2)) = \sqrt{\frac{1}{2} k \|\mathbf{x}_1 - \mathbf{x}_2\|^2 + \frac{1}{2} \|\mathbf{y}_1 - \mathbf{y}_2\|^2}$$

with $k > 0$ a parameter we can choose, and $(\mathbf{x}_1, \mathbf{y}_1), (\mathbf{x}_2, \mathbf{y}_2) \in \mathbb{R}^3 \times \mathbb{R}^3$. Let $B_{\mathrm{d}_k}((\mathbf{x}, \mathbf{y}), r) \subset \mathbb{R}^3 \times \mathbb{R}^3$ be the open ball of radius r centered at $(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^3 \times \mathbb{R}^3$ in d_k-metric and $\bar{B}_{\mathrm{d}_k}((\mathbf{x}, \mathbf{y}), r)$ its closure.

Let $(\mathbf{L}_T, \mathbf{A}_T) \in D$ be the pair of the angular momentum and the Laplace vector of the target elliptic orbit. Define a (Lyapunov) function V on $T\mathbb{R}^3_0$ by

$$V(\mathbf{r}, \dot{\mathbf{r}}) = \frac{1}{2}k\|\mathbf{L}(\mathbf{r}, \dot{\mathbf{r}}) - \mathbf{L}_T\|^2 + \frac{1}{2}\|\mathbf{A}(\mathbf{r}, \dot{\mathbf{r}}) - \mathbf{A}_T\|^2. \tag{A.14}$$

Notice that $V(\mathbf{r}, \dot{\mathbf{r}})$ is the square of the distance between $(\mathbf{L}(\mathbf{r}, \dot{\mathbf{r}}), \mathbf{A}(\mathbf{r}, \dot{\mathbf{r}}))$ and $(\mathbf{L}_T, \mathbf{A}_T)$ in the metric d_k, i.e.,

$$V(\mathbf{r}, \dot{\mathbf{r}}) = \Big[\mathrm{d}_k\big((\mathbf{L}(\mathbf{r}, \dot{\mathbf{r}}), \mathbf{A}(\mathbf{r}, \dot{\mathbf{r}})), \, (\mathbf{L}_T, \mathbf{A}_T)\big)\Big]^2. \tag{A.15}$$

We will find a controller $\mathbf{F}$ whose direction maximally reduces this distance at each moment. Along the trajectories of (A.13),

$$\begin{aligned}
\frac{d}{dt}\mathbf{L}(\mathbf{r}, \dot{\mathbf{r}}) &= \mathbf{r} \times \mathbf{F} \\
\frac{d}{dt}\mathbf{A}(\mathbf{r}, \dot{\mathbf{r}}) &= \mathbf{F} \times \mathbf{L}(\mathbf{r}, \dot{\mathbf{r}}) + \dot{\mathbf{r}} \times (\mathbf{r} \times \mathbf{F}).
\end{aligned}$$

Hence,

$$\frac{d}{dt}V(\mathbf{r}, \dot{\mathbf{r}}) = \mathbf{F} \cdot \Big(k\Delta\mathbf{L} \times \mathbf{r} + \mathbf{L}(\mathbf{r}, \dot{\mathbf{r}}) \times \Delta\mathbf{A} + (\Delta\mathbf{A} \times \dot{\mathbf{r}}) \times \mathbf{r}\Big)$$

where

$$\Delta\mathbf{L} = \mathbf{L}(\mathbf{r}, \dot{\mathbf{r}}) - \mathbf{L}_T; \quad \Delta\mathbf{A} = \mathbf{A}(\mathbf{r}, \dot{\mathbf{r}}) - \mathbf{A}_T. \tag{A.16}$$

Take the controller $\mathbf{F}$ as follows:

$$\mathbf{F}(\mathbf{r}, \dot{\mathbf{r}}; \mathbf{L}_T, \mathbf{A}_T) = -f(\mathbf{r}, \dot{\mathbf{r}})\,(k\Delta\mathbf{L} \times \mathbf{r} + \mathbf{L}(\mathbf{r}, \dot{\mathbf{r}}) \times \Delta\mathbf{A} + (\Delta\mathbf{A} \times \dot{\mathbf{r}}) \times \mathbf{r}) \tag{A.17}$$

with $f(\mathbf{r}, \dot{\mathbf{r}}) > 0$ arbitrary. This choice is such that

$$\frac{dV}{dt}(\mathbf{r}, \dot{\mathbf{r}}) = -f(\mathbf{r}, \dot{\mathbf{r}})\|k\Delta\mathbf{L} \times \mathbf{r} + \mathbf{L}(\mathbf{r}, \dot{\mathbf{r}}) \times \Delta\mathbf{A} + (\Delta\mathbf{A} \times \dot{\mathbf{r}}) \times \mathbf{r}\|^2 \le 0. \tag{A.18}$$

We now use LaSalle's invariance principle to prove asymptotically stable convergence to the target orbit (see Khalil [1996] for an exposition of LaSalle's invariant principle). For notational simplicity, we will suppress the dependence of $\mathbf{L}$ and $\mathbf{A}$ on $(\mathbf{r}, \dot{\mathbf{r}})$ from now on. Let

$$J = \{(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^3 \times \mathbb{R}^3 \mid \mathbf{x} \neq 0, \|\mathbf{y}\| < \mu\}, \tag{A.19}$$

which is open in $\mathbb{R}^3 \times \mathbb{R}^3$. There is an $l > 0$ such that

$$\bar{B}_{\mathrm{d}_k}((\mathbf{L}_T, \mathbf{A}_T), l) \subset J.$$

Let

$$\Omega_l = \pi^{-1}(\bar{B}_{\mathrm{d}_k}((\mathbf{L}_T, \mathbf{A}_T), l)).$$

By (A.15),

$$\Omega_l = \{(\mathbf{r}, \dot{\mathbf{r}}) \in T\mathbb{R}^3_0 \mid V(\mathbf{r}, \dot{\mathbf{r}}) \leq l^2\}. \tag{A.20}$$

Notice that (A.5) implies $\pi(T\mathbb{R}^3_0) \subset I$, where

$$I = \{(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^3 \times \mathbb{R}^3 \mid \mathbf{x} \cdot \mathbf{y} = 0\}.$$

Then $\Omega_l = \pi^{-1}(\bar{B}_{\mathrm{d}_k}((\mathbf{L}_T, \mathbf{A}_T), l) \cap I)$. Notice that the set $\bar{B}_{\mathrm{d}_k}((\mathbf{L}_T, \mathbf{A}_T), l) \cap I$ is a compact subset of D. Hence, Ω_l is a compact subset of Σ_e by Corollary A.2.3. By (A.18) and (A.20), the set Ω_l is a positively invariant compact set. We will show that every trajectory of the closed-loop system starting from Ω_l asymptotically converges to the Keplerian orbit $\pi^{-1}(\mathbf{L}_T, \mathbf{A}_T)$. Define

$$\mathcal{E} = \left\{(\mathbf{r}, \dot{\mathbf{r}}) \in \Omega_l \middle| \frac{dV}{dt}(\mathbf{r}, \dot{\mathbf{r}}) = 0\right\} = \{(\mathbf{r}, \dot{\mathbf{r}}) \in \Omega_l \mid \mathbf{F}(\mathbf{r}, \dot{\mathbf{r}}; \mathbf{L}_T, \mathbf{A}_T) = 0\}$$

$$\mathcal{M} = \text{ the largest invariant subset of } \mathcal{E}.$$

Let $(\mathbf{r}(t), \dot{\mathbf{r}}(t))$ be an arbitrary trajectory contained in $\mathcal{M}$. Since $\mathcal{M} \subset \mathcal{E}$, there is no control force acting on it. Hence, $(\mathbf{r}(t), \dot{\mathbf{r}}(t))$ is an elliptic Keplerian flow. Let E, $\mathbf{L}$, and $\mathbf{A}$ be the respective energy, angular momentum, and Laplace vector of the Keplerian orbit $(\mathbf{r}(t), \dot{\mathbf{r}}(t))$. They are all constant in time t. By the definition of $\mathcal{M}$, $(\mathbf{r}(t), \dot{\mathbf{r}}(t))$ satisfies

$$k\Delta\mathbf{L} \times \mathbf{r}(t) + \mathbf{L} \times \Delta\mathbf{A} + (\Delta\mathbf{A} \times \dot{\mathbf{r}}(t)) \times \mathbf{r}(t) = 0. \tag{A.21}$$

Let Π be the plane through the origin in $\mathbb{R}^3$ which is normal to $\mathbf{L}$, i.e., the plane where the ellipse swept out by $\mathbf{r}(t)$ lies. The inner product of $\mathbf{r}(t)$ and (A.21) gives

$$0 = \mathbf{r}(t) \cdot (\mathbf{L} \times \Delta\mathbf{A}) = \Delta\mathbf{A} \cdot (\mathbf{r}(t) \times \mathbf{L}). \tag{A.22}$$

Notice that

$$\Pi = \mathrm{span}\{\mathbf{r}(t) \times \mathbf{L} \mid t \in \mathbb{R}\}, \tag{A.23}$$

since $\mathbf{r}(t)$ traces an ellipse in Π. By (A.22) and (A.23)

$$\Delta\mathbf{A} = c\mathbf{L} \tag{A.24}$$

for some $c \in \mathbb{R}$. Note that c is constant since both $\Delta\mathbf{A}$ and $\mathbf{L}$ are constant. Substitution of (A.24) into (A.21) gives

$$(k\Delta\mathbf{L} - c(\dot{\mathbf{r}}(t) \times \mathbf{L})) \times \mathbf{r}(t) = 0$$

which by (A.4) gives

$$(k\Delta\mathbf{L} - c\mathbf{A}) \times \mathbf{r}(t) = 0.$$

This implies that the constant vector $(k\Delta\mathbf{L} - c\mathbf{A})$ is parallel to the nonzero vector $\mathbf{r}(t)$ which changes its direction in time since it sweeps an ellipse. It follows that

$$\Delta\mathbf{L} = \frac{c}{k}\mathbf{A}. \tag{A.25}$$

By (A.16), (A.24), and (A.25),

$$\mathbf{L}_T = \mathbf{L} - \frac{c}{k}\mathbf{A}, \qquad \mathbf{A}_T = \mathbf{A} - c\mathbf{L}. \tag{A.26}$$

Since $(\mathbf{L}_T, \mathbf{A}_T)$ and $(\mathbf{L}, \mathbf{A})$ are contained in D, (A.26) implies

$$0 = \mathbf{L}_T \cdot \mathbf{A}_T = -c\left(\|\mathbf{L}\|^2 + \frac{1}{k}\|\mathbf{A}\|^2\right).$$

Since $\|\mathbf{L}\| > 0$ and $k > 0$, it follows that $c = 0$. Substituting $c = 0$ to (A.26) gives

$$\mathbf{L} = \mathbf{L}_T, \qquad \mathbf{A} = \mathbf{A}_T.$$

By Proposition A.2.1, the Keplerian orbit $(\mathbf{r}(t), \dot{\mathbf{r}}(t))$ is the same as the target orbit $\pi^{-1}(\mathbf{L}_T, \mathbf{A}_T)$. Thus, the only trajectory lying in $\mathcal{M}$ is the Keplerian orbit $\pi^{-1}(\mathbf{L}_T, \mathbf{A}_T)$. By LaSalle's invariance principle, the following holds:

Proposition A.3.1. *Let $(\mathbf{L}_T, \mathbf{A}_T) \in D$ be the pair of the angular momentum and the Laplace vector of the target elliptic orbit. Take any closed ball $\bar{B}_{\mathrm{d}_k}((\mathbf{L}_T, \mathbf{A}_T), l)$ of a radius $l > 0$ centered at $(\mathbf{L}_T, \mathbf{A}_T)$ contained in the following open set J*

$$J = \{(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^3 \times \mathbb{R}^3 \mid \mathbf{x} \neq 0, \|\mathbf{y}\| < \mu\}.$$

Then, every trajectory starting in the subset $\pi^{-1}(\bar{B}_{\mathrm{d}_k}((\mathbf{L}_T, \mathbf{A}_T), l))$ of $T\mathbb{R}^3_0$ remains in that subset and asymptotically converges to the target elliptic orbit $\pi^{-1}(\mathbf{L}_T, \mathbf{A}_T)$ in the closed-loop system (A.13) with the control law in (A.17).

Remark A.3.2. *1. Proposition A.3.1 holds for any positive k in the definition of the metric d_k. There are two interpretations of k. One is that k determines the relative weighting between the two quadratic terms in the function V in (A.14). The other is that*

k determines the shape of the region of attraction since k determines the shape of the ball B_{d_k} with the metric d_k.

2. We explain some advantages of using $(\mathbf{L}, \mathbf{A})$ instead of other quantities, such as orbital elements $(a, e, i, \Omega, \omega)$ or equinoctial elements (a, h, k, p, q) (see Battin [1987] for definitions of those elements). First, $(\mathbf{L}, \mathbf{A})$ is globally well defined whereas orbital elements become singular on circular or equatorial orbits. Second, $\mathbf{L}$ and $\mathbf{A}$ are $\mathbb{R}^3$-valued and $\mathbb{R}^3$ has a nice (Lie-)algebraic structure, namely the cross product $\times$ as well as the dot product $\cdot$, and the property

$$(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c} = -\mathbf{b} \cdot (\mathbf{a} \times \mathbf{c}) \tag{A.27}$$

for $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{R}^3$. (It is not accidental that $\mathbf{L}$ and $\mathbf{A}$ are $\mathbb{R}^3$-valued. See Cushman and Bates [1997] for more details). Notice that we have exclusively used the usual Euclidean norm $\|\cdot\|$ on $\mathbb{R}^3$ in the definition of the metric d_k and the Lyapunov function V in order to make use of the algebraic structure of $\mathbb{R}^3$. In particular, the property (A.27) was very useful in the analysis of the set where $dV/dt = 0$ in the application of LaSalle's invariance principle. It will be difficult to analyze $dV/dt = 0$ if one uses orbital elements or equinoctial elements to define a Lyapunov function as a sum of squares of differences of elements, because the elements do not have useful algebraic structures.

A.3.2 Global Orbit Transfer

The basic idea of the global orbit transfer is to use a finite number of intermediate (target) orbits to transfer between two arbitrary elliptic orbits. We will show a way of choosing intermediate target orbits to achieve the global orbit transfer. By proper choice of intermediate orbits we can also avoid undesirable orbits. The essence of the following argument lies in the combination of Proposition A.3.1 and the path-connectivity of the set D defined in (A.8). We first show that D is path-connected. Any two points $(\mathbf{L}_0, \mathbf{A}_0)$ and $(\mathbf{L}_1, \mathbf{A}_1)$ in D can be joined by a path $\mathrm{c} : [0,1] \to D \subset \mathbb{R}^3 \times \mathbb{R}^3$, for example,

$$\mathrm{c}(t) = \begin{cases} (\mathbf{L}_0, (1-3t)\mathbf{A}_0) & 0 \leq t \leq 1/3 \\ (\mathrm{d}(3t-1), 0) & 1/3 \leq t \leq 2/3 \\ (\mathbf{L}_1, (3t-2)\mathbf{A}_1) & 2/3 \leq t \leq 1, \end{cases}$$

where $\mathrm{d} : [0,1] \to \mathbb{R}^3 - \{0\}$ is a path connecting $\mathbf{L}_0$ and $\mathbf{L}_1$. The existence of $\mathrm{d}(t)$ is guaranteed by the path-connectivity of $\mathbb{R}^3 - \{0\}$. Hence, D is path-connected.

Choose two arbitrary elliptic Keplerian orbits $(\mathbf{L}_0, \mathbf{A}_0)$ and $(\mathbf{L}_1, \mathbf{A}_1)$ from D where we want to transfer from $(\mathbf{L}_0, \mathbf{A}_0)$ to $(\mathbf{L}_1, \mathbf{A}_1)$. By the path-connectivity of D, one can choose a path $c : [0,1] \to D \subset \mathbb{R}^3 \times \mathbb{R}^3$ connecting $(\mathbf{L}_0, \mathbf{A}_0)$ and $(\mathbf{L}_1, \mathbf{A}_1)$. Recall that J in (A.19) is open and $D \subset J$. There is $\tilde{l} > 0$ such that $B_{\mathrm{d}_k}(c(s), \tilde{l}) \subset J$ for all $s \in [0,1]$ (for example, take any number less than the distance between the compact set $c([0,1])$ and the boundary of J or just apply the Lebesgue number lemma to $c([0,1])$ and J (see Munkres [1975] for the Lebesgue number lemma). Take any positive number l less than $\tilde{l}$. By the uniform continuity of c, we can find a subdivision of $[0,1]$, say $s_0, \ldots, s_N$ with $s_0 = 0$ and $s_N = 1$ such that for $i = 0, \ldots, N-1$ the set $c([s_i, s_{i+1}])$ is contained in $B_{\mathrm{d}_k}(c(s_{i+1}), l) \cap D$. In particular, $c(s_i) \in B_{\mathrm{d}_k}(c(s_{i+1}), l) \cap D$. Notice that $\bar{B}_{\mathrm{d}_k}(c(s_{i+1}), l) \cap D$ is a region of attraction of $c(s_{i+1})$ with the controller $\mathbf{F}(\,\cdot\,; c(s_{i+1}))$; this follows from Proposition A.3.1 since $\bar{B}_{\mathrm{d}_k}(c(s_i), l) \subset B_{\mathrm{d}_k}(c(s_i), \tilde{l}\,) \subset J$ for each i. Hence, we can drive the trajectory $(\mathbf{r}(t), \dot{\mathbf{r}}(t))$ from the orbit $(\mathbf{L}_0, \mathbf{A}_0)$ to the orbit $(\mathbf{L}_1, \mathbf{A}_1)$ through the intermediate target orbits $\{c(s_i) \mid i = 0, \ldots, N\}$ by using the controllers $\{\mathbf{F}(\,\cdot\,; c(s_i)) \mid i = 1, \ldots, N\}$ of the form (A.17) sequentially. The trajectory lies in $\pi^{-1}(K)$ where

$$K = \left(\bigcup_{i=1}^{N} \bar{B}_{\mathrm{d}_k}(c(s_i), l) \right) \cap D.$$

A lower bound of $\|\mathbf{r}(t)\|$ of the total trajectory $(\mathbf{r}(t), \dot{\mathbf{r}}(t))$ is given by

$$\min \left\{ \frac{\|\mathbf{L}\|^2}{\mu + \|\mathbf{A}\|} \;\middle|\; (\mathbf{L}, \mathbf{A}) \in K \right\} \tag{A.28}$$

and an upper bound is given by

$$\max \left\{ \frac{\|\mathbf{L}\|^2}{\mu - \|\mathbf{A}\|} \;\middle|\; (\mathbf{L}, \mathbf{A}) \in K \right\} \tag{A.29}$$

Remark A.3.3. *Above, we just showed the possibility of global orbit transfer. There can be several ways to achieve global transfer. For example, one can use different radii for each region of attraction, B_{d_k}. Also, one can use different k's for each region of attraction. A discussion on k was given in a remark following Proposition A.3.1.*

A.3.3 Special Transfer: Transfer to Circular Orbits

The Lyapunov function suggested in § A.3.1 is not the only available Lyapunov function for local orbit transfer. We here suggest another Lyapunov function for the transfer to *circular* orbits.

Notice that a circular Keplerian orbit is uniquely determined by its angular momentum $\mathbf{L}$ because the Laplace vector $\mathbf{A}$ is zero for circular orbits. The corresponding energy E is determined by $\mathbf{L}$ since $\mu^2 + 2E\|\mathbf{L}\|^2 = 0$ by (A.6). Let $\mathbf{L}_T$ and E_T be the angular momentum and the energy of a given target circular orbit. Define a function V on $T\mathbb{R}^3_0$

by

$$V(\dot{\mathbf{r}}, \dot{\mathbf{r}}) = \frac{1}{2}k||\mathbf{L}(\dot{\mathbf{r}}, \dot{\mathbf{r}}) - \mathbf{L}_T||^2 + \frac{1}{2}(E(\dot{\mathbf{r}}, \dot{\mathbf{r}}) - E_T)^2 \tag{A.30}$$

with $k > 0$. Then one can compute

$$\frac{dV}{dt}(\mathbf{r}, \dot{\mathbf{r}}) = \mathbf{F} \cdot (k\Delta\mathbf{L} \times \mathbf{r} + \Delta E\dot{\mathbf{r}}),$$

where $\Delta\mathbf{L} := \mathbf{L}(\mathbf{r}, \dot{\mathbf{r}}) - \mathbf{L}_T$ and $\Delta E := E(\mathbf{r}, \dot{\mathbf{r}}) - E_T$. Take the following form of controller

$$\mathbf{F}(\mathbf{r}, \dot{\mathbf{r}}, t) = -f(\mathbf{r}, \dot{\mathbf{r}})(k\Delta\mathbf{L} \times \mathbf{r} + \Delta E\dot{\mathbf{r}}) \tag{A.31}$$

with $f(\mathbf{r}, \dot{\mathbf{r}}) > 0$ an arbitrary positive function. This choice is such that

$$\frac{dV}{dt}(\mathbf{r}, \dot{\mathbf{r}}) = -f(\mathbf{r}, \dot{\mathbf{r}})||k\Delta\mathbf{L} \times \mathbf{r} + \Delta E\dot{\mathbf{r}}||^2 \leq 0. \tag{A.32}$$

One can find $l > 0$ with $l < \frac{k}{2}||\mathbf{L}_T||^2$ such that $\Omega_l := V^{-1}([0, l])$ is a compact subset of Σ_e by (A.6) and Corollary A.2.3. Notice that Ω_l is positively invariant by (A.32). Let $\mathcal{M}$ be the largest invariant subset of the set $\{(\mathbf{r}, \dot{\mathbf{r}}) \in \Omega_l \mid dV/dt = 0\} = \{(\mathbf{r}, \dot{\mathbf{r}}) \in \Omega_l \mid \mathbf{F} = 0\}$. Let $(\mathbf{r}, \dot{\mathbf{r}})$ be an arbitrary trajectory in $\mathcal{M}$. Then it is an elliptic orbit because $\mathbf{F} = 0$. Let $\mathbf{L}$ and E be the angular momentum and the energy, respectively, of the orbit $(\mathbf{r}(t), \dot{\mathbf{r}}(t))$, which are of course, constant in time t. By definition of $\mathcal{M}$, the trajectory $(\mathbf{r}(t), \dot{\mathbf{r}}(t))$ satisfies

$$k\Delta\mathbf{L} \times \mathbf{r}(t) + \Delta E\dot{\mathbf{r}}(t) = 0. \tag{A.33}$$

The constant value ΔE is either zero or nonzero. If $\Delta E = 0$, then $\Delta\mathbf{L} \times \mathbf{r}(t) = 0$ by (A.33), which implies $\Delta\mathbf{L} = 0$ since the constant vector $\Delta\mathbf{L}$ is parallel to the vector $\mathbf{r}(t)$ which sweeps an ellipse. Hence, the trajectory $(\mathbf{r}(t), \dot{\mathbf{r}}(t))$ is the target orbit if $\Delta E = 0$. We now suppose $\Delta E \neq 0$. The inner product of (A.33) with $\mathbf{r}(t)$ gives $\dot{\mathbf{r}}(t) \cdot \mathbf{r}(t) = 0$, which implies that $(\mathbf{r}(t), \dot{\mathbf{r}}(t))$ is a circular orbit. Since $\mathbf{r}(t)$ and $\dot{\mathbf{r}}(t)$ are perpendicular to each other and $\mathbf{r}(t)$ sweeps a circle, it follows from (A.33) that $\Delta\mathbf{L}$ is parallel to $\mathbf{L}$, which implies that $\mathbf{L}$ is parallel to $\mathbf{L}_T$. Since we chose l less than $\frac{k}{2}||\mathbf{L}_T||^2$, the vector $\mathbf{L}$ cannot be in the opposite direction of $\mathbf{L}_T$ by definition of Ω_l. Hence, $\mathbf{L}$ and $\mathbf{L}_T$ have the same directions. Let $\mathbf{e}_L := \mathbf{L}/||\mathbf{L}|| = \mathbf{L}_T/||\mathbf{L}_T||$. Recall the general formulas for energy and the magnitude of the angular momentum for a circular orbit of radius r as follows:

$$E = -\frac{\mu}{2r}, \quad ||\mathbf{L}|| = \sqrt{(\mu r)}, \tag{A.34}$$

where the second formula is derived from (A.12). Let r be the radius of the circular orbit $(\mathbf{r}(t), \dot{\mathbf{r}}(t))$ and r_T be that of the target circular orbit. By (A.34), the equation (A.33) can be written as

$$(\sqrt{r}-\sqrt{r_T})\left(k\sqrt{\mu}(\mathbf{e}_L\times\mathbf{r})+\frac{\mu(\sqrt{r}+\sqrt{r_T})}{2rr_T}\dot{\mathbf{r}}\right)=0. \tag{A.35}$$

Notice that $(\mathbf{e}_L \times \mathbf{r}(t))$ is in the same direction as $\dot{\mathbf{r}}(t)$ and that $r \neq r_T$ since we assumed $\Delta E \neq 0$. The left hand side of (A.35) is not zero, which gives a contradiction. Therefore, the trajectory $(\mathbf{r}(t), \dot{\mathbf{r}}(t))$ is the target circular orbit. We have shown $\mathcal{M}$ consists of the target orbit only. By LaSalle's invariance principle, any trajectory starting in Ω_l remains in Ω_l and asymptotically converges to the target orbit with the control law (A.31). As a remark, we note that the control law (A.31) can be used in the global transfer too.

A.4 Example

For illustrative purposes, we give an example of a transfer from low-Earth orbit (LEO) to geosynchronous orbit (GEO). The initial LEO is a circular orbit with radius 7000 km and inclination 28.5 deg. The target GEO is also circular with radius 42,000 km and inclination 0 deg. The maximum thrust level is 9.8×10^{-5}km/sec^2. These data are from pp. 362–374 in Chbotov [1996]. We use canonical units in simulations; 806.812 sec = 1 canonical time unit, 6378.140 km = 1 canonical distance unit, 9.8×10^{-3}km/sec^2 = 1 canonical acceleration unit, and the gravitational parameter $\mu = 1$. In the following, all units are canonical unless otherwise indicated. The initial point is given by

$$\begin{aligned} x_0 &= (-0.70545852988580, -0.73885031681775, -0.40116299069586), \\ v_0 &= (0.73122658145185, -0.53921753373056, -0.29277123328399), \end{aligned}$$

which corresponds to the initial point in the time-optimal case of Chbotov [1996]. The angular momentum and Laplace vector of the target orbit are given by

$$\mathbf{L}_T = (0, 0, 2.56612389857378); \quad \mathbf{A}_T = (0, 0, 0).$$

We use the Lyapunov function in (A.14) with $k = 2$. To meet the constraint on the magnitude of the thrust, we choose f in (A.17) such that the control law $\mathbf{F}$ becomes

$$\mathbf{F}(\mathbf{r},\dot{\mathbf{r}}) = \begin{cases} \frac{1}{\epsilon}G(\mathbf{r},\dot{\mathbf{r}}) & \text{if } \|G(\mathbf{r},\dot{\mathbf{r}})\| < \epsilon F_{\max} \\ F_{\max}\frac{G(\mathbf{r},\dot{\mathbf{r}})}{\|G(\mathbf{r},\dot{\mathbf{r}})\|} & \text{if } \|G(\mathbf{r},\dot{\mathbf{r}})\| \geq \epsilon F_{\max}, \end{cases}$$

where $F_{\max} = 0.01$, $\epsilon = 0.00001$ and

$$G(\mathbf{r},\dot{\mathbf{r}}) = -\left(k\Delta\mathbf{L}\times\mathbf{r} + \mathbf{L}(\mathbf{r},\dot{\mathbf{r}})\times\Delta\mathbf{A} + (\Delta\mathbf{A}\times\dot{\mathbf{r}})\times\mathbf{r}\right).$$

	time-opt. transfer	Lyap. transfer
sim. time	16.14 hr	18.87 hr
a_f	42,000.001 km	41,974.952 km
e_f	0.00097	0.00462
i_f	0.999359 deg	0.202893 deg

Table A.1: Comparison of the time-optimal transfer and the Lyapunov-based transfer.

One can easily check that $\|\mathbf{F}(\mathbf{r},\dot{\mathbf{r}})\| \leq F_{\max}$. Figure A.1 shows a plot of the simulation results for time $13.4 \times 2\pi$. For comparison of the time-optimal transfer in Chbotov [1996] and our Lyapunov-based transfer, we list the final simulation results in Table A.1 – semi-major axis a_f, eccentricity e_f, and inclination i_f – where all the data are in real units, and the data of the time-optimal transfer are from Chbotov [1996], in which the time-optimal controller has constant magnitude $F_{\max}$ during the entire transfer. When comparing these two results, one should take into account that our controller is in a simple and analytic form, whereas the time-optimal controller is numerical and computationally challenging. Also, we can improve the simulation result by choosing different values of k or inserting intermediate target orbits. Hence, these results are sufficient to show that this simple scheme produces a transfer comparable to those generated by much more complex and numerically intensive approaches.

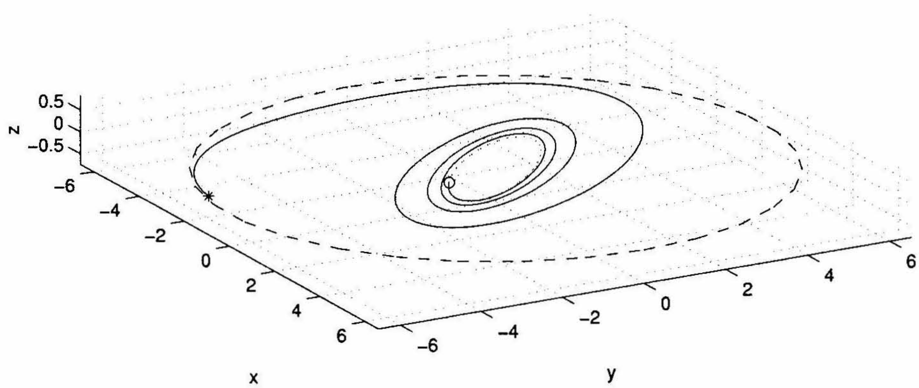

Figure A.1: Lyapunov-based LEO-to-GEO transfer in canonical units. The initial and target orbits are dotted $\cdots$ and dashed $--$, respectively. The initial and final points are marked with o and $*$, respectively.

A.5 Conclusions and Future Work

In this work, we have rigorously shown that mechanically motivated Lyapunov function techniques can be used to systematically produce easily implementable, asymptotically stable controllers for orbit transfers between elliptic Kepler orbits.

For long duration, low-thrust transfers, it may be necessary to take into account the J_2 effect, that is, the effect of the bulge of the earth. We believe that our techniques can be extended to that case, at least in the context of the most important correction terms. This would rely on results on the geometry of the perturbed Kepler problem.

A second direction for future research would be to optimize our method. Although we made no attempt at systematic time or fuel optimization in this paper, it would be interesting to pursue this by exploiting, for example, the freedom in the constant k that appears in the Lyapunov function or the freedom in the choice of the function $f(\mathbf{r}, \dot{\mathbf{r}})$ that appears in the control law.

Bibliography

Abraham, R. and J. E. Marsden [1978] *Foundations of Mechanics,* 2nd ed., Addison-Wesley, Reading, Massachusetts.

Abraham, R., J. E. Marsden, and T. S. Ratiu [1988] *Manifolds, Tensor Analysis, and Applications*, Applied Mathematics Series, **75**, Springer Verlag, New York, Second Edition.

Åström, K. J. and K. Furuta [1996] Swinging up a pendulum by energy control, IFAC Proceedings, **13**.

Auckly, D., L. Kapitanski, and W. White [2000] Control of nonlinear underactuated systems, *Comm. Pure Appl. Math.*, **53**, 354-369.

Auckly, D. and L. Kapitanski [1999] Mathematical problems in the control of underactuated systems, Preprint. `http://www.math.ksu.edu/~dav/`

Auckly, D. and L. Kapitanski [2001] On the λ equations for matching control laws, Preprint. `http://www.math.ksu.edu/~dav/`

Battin, R. H. [1987] *An Introduction to the Mathematics and Methods of Astrodynamics,* in the AIAA Education Series, American Institute of Aeronautics and Astronautics, Inc., New York, New York.

Blankenstein, G., R. Ortega, and A. van Der Schaft [2001] The matching conditions of controlled Lagrangians and interconnection and damping assignment passivity based control, Preprint.

Bloch, A. M., D. E. Chang, N. E. Leonard, and J. E. Marsden [2001] Controlled Lagrangians and the stabilization of mechanical systems II: Potential shaping, *Trans. IEEE on Auto. Control* , **46**, 1156–1171.

Bloch, A. M., D. E. Chang, N. E. Leonard, J. E. Marsden, and C. Woolsey [2000] Asymptotic stabilization of Euler-Poincaré mechanical systems, in *Lagrangian and Hamiltonian Methods for Nonlinear Control: A Proc. Volume from the IFAC Workshop* (N.E. Leonard and R. Ortega eds.), Pergamon, 51–56.

Bloch, A. M. and P. E. Crouch [1998] Representation of Dirac structures on vector space and nonlinear L-C circuits, in *Proc. Symp. on Appl. Math.*, AMS **66**, 103-118.

Bloch, A. M., P. S. Krishnaprasad, J. E. Marsden, and R. Murray [1996] Nonholonomic mechanical systems with symmetry. *Arch. Rational Mech. Anal.*, **136**, 21-99.

A. Bloch, P. S. Krishnaprasad, J. E. Marsden, and T. Ratiu [1994] Dissipation induced instabilities, *Annals Inst. H. Poincaré, Anal. Nonlin.*, **11**, 37–90.

Bloch, A. M., P. S. Krishnaprasad, J. E. Marsden, and G. Sánchez De Alvarez [1992] Stabilization of rigid body dynamics by internal and external torques, *Automatica*, **28**, 745–756.

Bloch, A. M., N. E. Leonard, and J. E. Marsden [1997] Stabilization of mechanical systems using controlled Lagrangians, in *Proc. CDC*, **36**, 2356–2361.

Bloch, A. M., N. E. Leonard, and J. E. Marsden [1998] Matching and stabilization by the method of controlled Lagrangians, *Proc. CDC* **37**, 1446–1451.

Bloch, A. M., N. E. Leonard, and J. E. Marsden [1999a] Stabilization of the pendulum on a rotor arm by the method of controlled Lagrangians, *Proc. of ICRA '99*, IEEE, 500–505.

Bloch, A. M., N. E. Leonard, and J. E. Marsden [1999b] Potential shaping and the method of controlled Lagrangians, *Proc. CDC*, **38**, 1653-1657.

Bloch, A. M., N. E. Leonard, and J. E. Marsden [2000] Controlled Lagrangians and the stabilization of mechanical systems I: The first matching theorem, *IEEE Trans. Automat. Control*, **45**, 2253–2270.

Bloch, A. M., N. E. Leonard, and J. E. Marsden [2001] Controlled Lagrangians and the stabilization of Euler-Poincaré mechanical systems, *Int. J. of Robust and Nonlinear Control*, **11**, 191–214.

Bloch, A. M. and J. E. Marsden [1990] Stabilization of rigid body dynamics by the energy-Casimir method, *Sys. and Cont. Lett.*, **14**, 341–346.

Boothby, W. M. [1986] *An Introduction to Differentiable Manifolds and Riemannian Geometry*, Academic Press.

Brockett, R. W. [1976] Control theory and analytical mechanics, in *1976 Ames Research Center (NASA) Conference on Geometric Control Theory*, (R. Hermann and C. Martin, eds.), Math Sci Press, Brookline, Massachusetts, Lie Groups: History, Frontiers, and Applications., **VII**, 1–46.

Cannas Da Silva, A. and A. Weinstein [1999] *Geometric Models for Noncommutative Alebras*, Amer. Math. Soc., Berkeley Mathematics Lecture Notes.

Cendra, H., J. E. Marsden, and T. S. Ratiu [2001] *Lagrangian Reduction by Stages*, Memoirs of the Amer. Math. Soc., **152**.

Chang, D. E. [2002] Controlled Lagrangian and Hamiltonian Systems, *SIAM student paper competition*, Submitted.

Chang, D. E., A. M. Bloch, N. E. Leonard, J. E. Marsden, and C. Woolsey [2002] The equivalence of controlled Lagrangian and controlled Hamiltonian systems, *ESAIM:COCV*, (to appear).

Chang, D. E., D. F. Chichka, and J. E. Marsden [2002] Lyapunov-based transfer between elliptic Keplerian orbits, *DCDS-Series B*, **2**, No.1, 57–67.

Chang, D. E. and J. E. Marsden [2000] Asymptotic stabilization of the heavy top using controlled Lagrangians, *Proc. CDC* **39**.

Chobotov, V. A. [1996] *Orbital Mechanics,* 2nd ed., in the AIAA Education Series, American Institute of Aeronautics and Astronautics, Inc., New York, New York, 327–402.

Crouch, P. E. and A. J. van der Schaft [1987] *Variational and Hamiltonian Control Systems*, Lecture Notes in Control and Information Sciences, Springer-Verlag, Berlin. **101**.

Cushman, R. H. and L. M. Bates [1997] *Global Aspects of Classical Integrable Systems,* Birkhauser, Boston, 37–82.

Edelbaum, T. N. [1964] Optimum low-thrust rendezvous and station-keeping, **2**, 1196-1201.

Edelbaum, T. N. [1965] Optimum power-limited orbit transfer in strong gravity fields, *AIAA Journal*, **3**, 921–925.

Gallot, S., D. Hulin, and J. Lafontaine [1993] *Riemannian Geometry,* 2nd ed., Springer-Verlag.

Goldstein, H. [1980] *Classical Mechanics,* 2nd ed., Addison-Wesley, Reading, Mass..

Hamberg, J. [1999] General matching conditions in thetheory of controlled Lagrangians, in *Proc. CDC* **38**, 2519–2523.

Hamberg, J. [2000] Controlled Lagrangians, symmetries and conditions for strong matching, in *Lagrangian and Hamiltonian Methods for Nonlinear Control: A Proc. Volume from the IFAC Workshop* (N.E. Leonard and R. Ortega eds.), Pergamon, 62–67.

Holm, D. D., J. E. Marsden, and T. S. Ratiu [1998] The Euler-Poincare equations and semidirect products with applications to continuum theories, *Adv. in Math.* **137**, 1–81

Ilgen, M. R. [1980] Low thrust OTV guidance using Lyapunov optimal feedback control techniques, Advances in the Astronautical Sciences, **85**, 1527–1545

Jalnapurkar, S. M. and J. E. Marsden [1999] Stabilization of relative equilibria II, *Regul. Chaotic Dyn.*, **3**, 161–179.

Jalnapurkar, S. M. and J. E. Marsden [2000] Stabilization of relative equilibria, *IEEE Trans. Automat. Control*, **45**, 1483–1491.

Khalil, H. [1996] *Nonlinear Systems*, Macmillan.

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

Koon, W. S. and J. E. Marsden [1998] The Poisson reduction of nonholonomic mechanical systems, *Reports on Math. Phys.*, **42**, 101–134.

Krishnaprasad, P. S. [1985] Lie–Poisson structures, dual-spin spacecraft and asymptotic stability, *Nonl. Anal. Th. Meth. and Appl.* **9**, 1011–1035.

Marsden, J. E. and T. S. Ratiu [1999] *Introduction to Mechanics and Symmetry*, Springer-Verlag, Texts in Applied Mathematics, vol. 17; Second Edition.

Merkin, D. R. [1996] *Introduction to the Theory of Stability*, Springer-Verlag, Texts in Applied Mathematics, vol. 24.

Montgomery, R., J. E. Marsden, and T. S. Ratiu [1984] Gauged Lie-Poisson structures, *Contemp. Math.*, **28**, 101–114.

Munkres, J. R. [1975] *Topology: A First Course,* Prentice-Hall, Inc., New Jersey, 178–180.

Ortega, R., M. W. Spong, F. Gómez-estern, and G. Blankenstein [2001] Stabilization of underactuated mechanical systems via interconnection and damping assignment, *IEEE Trans. Aut. Control (submitted).*

Ortega, R. and M. W. Spong [2000] Stabilization of underactuated mechanical systems via interconnection and damping assignment, in *Lagrangian and Hamiltonian Methods for Nonlinear Control: A Proc. Volume from the IFAC Workshop* (N.E. Leonard and R. Ortega eds.), Pergamon, 74–79.

Petrović, V., R. Ortega, and A. M. Stanković [2001] Interconnection and damping assignment approach to control of PM synchronous motors, *IEEE Trans. on Contr. Syst. Tech.*, **9**, No. 6, 811–820.

Spong, M. W. [1997] Underactuated mechanical systems, in *Control Problems in Robotics and Automation*, B. Siciliano and K.P. Valavanis (Eds), *Lecture Notes in Control and Information Sciences* **230** Spinger-Verlag. [Presented at the International Workshop on Control Problems in Robotics and Automation: Future Directions Hyatt Regency, San Diego, California, Dec. 1997.]

van der Schaft, A. J. [1982] Hamiltonian dynamics with external forces and observations, *Mathematical Systems Theory*, **15**,145–168.

van der Schaft, A. J. [1986] Stabilization of Hamiltonian systems, *Nonlinear Analysis, Theory, Methods and Applications*, **10**, 1021–1035.

van der Schaft, A. J. [2000] *L_2-Gain and Passivity Techniques in Nonlinear Control*, Springer-Verlag, Communication and Control Engineering Series.

van der Schaft, A. J. and B. Maschke [1994] On the Hamiltonian formulation of nonholonomic mechanical systems, *Rep. Math. Phys.*, **34**, 225–233.

Wang, L. S. and P. S. Khrisnaprasad [1992] Gyroscopic control and stabilization, *J. Nonlinear Sci.*, **2**, 367–415.

Wiggins, S. [1990] *Introduction to Applied Nonlinear Dynamicsl Systems and Chaos*, Texts in Applied Mathematical Science, **2**, Springer-Verlag.

Woolsey, C. A. [2001] *Energy Shaping and Dissipation: Underwater Vehicle Stabilization Using Internal Rotors*, PhD thesis, Princeton University.

Woolsey, C. A. and N. E. Leonard [1999] Underwater vehicle stabilization by internal rotors, *Proc. ACC*, 3417–3421.

Zenkov, D. V., A. M. Bloch, N. E. Leonard, and J. E. Marsden [2000] Matching and stabilization of the unicycle with rider, *Lagrangian and Hamiltonian Methods for Nonlinear Control: A Proc. Volume from the IFAC Workshop* (N.E. Leonard and R. Ortega eds.) Pergamon, 187–188.

Zenkov, D. V. A. M. Bloch, and J. E. Marsden [2002] Flat nonholonomic matching, *Proc. ACC 2002 (to appear)*.