REPORT DOCUMENTATION PAGE

AFRL-SR-BL-TR-98-
0358

1. AGENCY USE ONLY (Leave Blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	01 Apr 98	Final (01 Mar 96 – 28 Feb 98)

4. TITLE AND SUBTITLE The Nonlinear Control Theory of Complex Mechanical Systems	5. FUNDING NUMBERS F49620-96-1-0059
6. AUTHORS J. Baillieul	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Boston University, College of Engineering 110 Cummington Street Boston, MA 02215	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) AFOSR/NM 110 Duncan Avenue, Room B-115 Bolling Air Force Base, DC 20332-8080	10. SPONSORING/MONITORING AGENCY REPORT NUMBER
11. SUPPLEMENTARY NOTES	
12a. DISTRIBUTION AVAILABILITY STATEMENT Approved for Public Release	12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

This report summarizes a body of research dealing with the nonlinear control theory of complex mechanical systems. The principal focus is on so-called super-articulated (or underactuated) mechanical systems – systems having a mixture of directly controlled as well as uncontrolled degrees of freedom. The twofold goal of the research is: (1) how to control the degrees of freedom which are directly actuated without eliciting undesired behavior in the unactuated degrees of freedom, and (2) how to prescribe motions of the directly actuated degrees of freedom which are not directly controlled.

The report contains a comprehensive list of research published under this grant as well as some predecessor grants. Two papers in particular – "The Geometry of Controlled Mechanical Systems" and "Open-loop Oscillatory Stabilization of an n-Pendulum" comprise the present report.

19980430 091

14. SUBJECT TERMS nonlinear control theory, super-articulated systems, impulsive control, oscillatory inputs	15. NUMBER OF PAGES
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UL

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239.18
Designed using WordPerfect 6.1, AFOSR/XPP, Oct 96

AIR FORCE Grant No. F49620-96-1-0059

FINAL REPORT

4/1/98

Grant Project Title

THE NONLINEAR CONTROL THEORY OF COMPLEX MECHANICAL SYSTEMS

Principal Investigator

John Baillieul
Department of Aerospace/Mechanical Engineering
Boston University
Boston, MA 02215
(617) 353-9848
johnb@enga.bu.edu

ABSTRACT

This report summarizes a body of research dealing with the nonlinear control theory of complex mechanical systems. The principal focus is on so-called *super-articulated* (or underactuated) mechanical systems—systems having a mixture of directly controlled as well as uncontrolled degrees of freedom. The twofold goal of the research is: (1) how to control the degrees of freedom which are directly actuated without eliciting undesired behavior in the unactuated degrees of freedom, and (2) how to prescribe motions of the directly actuated degrees of freedom to achieve motion objectives for the degrees of freedom which are not directly controlled.

The report contains a comprehensive list of research published under this grant as well as some predecessor grants. Two papers in particular—"The Geometry of Controlled Mechanical Systems" and "Open-loop Oscillatory Stabilization of an n–Pendulum" comprise the present report.

The Geometry of Controlled Mechanical Systems

J. Baillieul*

April 13, 1998

Abstract

The control of mechanical systems has become a principal application focus of nonlinear control theory. This development was to a very large extent foreseen in the early work of R.W. Brockett, and a number of his papers from the mid 1970's have been highly influential in establishing the links between nonlinear control theory and geometric mechanics. Taking some of Brockett's early work as a starting point, we study intrinsically second order nonlinear control systems. The theory is developed using the language of differential geometry and affine connections. Velocity and acceleration controlled Lagrangian systems provide a rich class of examples of second order control systems. These arise in modeling the dynamics of *super-articulated* mechanical systems, in which only some of the configuration variables (or degrees of freedom) are directly controlled, with the remaining variables evolving under the dynamic influence of the actuated degrees of freedom. Our goal is to develop a nonlinear control theory which characterizes the way in which the unactuated degrees of freedom in these systems are influenced by the degrees of freedom which can be controlled directly. Central to the theory is a natural Riemannian structure and certain curvature-like quantities which provide a measure of the lack of integrability in the mapping of input trajectories to configuration space trajectories. The report chapter concludes with results connecting the geometry of second order systems with the theory of averaged Lagrangian and Hamiltonian systems.

1 Introduction

The early years (=for the most part the decade of the 1970's) in the development of nonlinear geometric control theory witnessed attempts to extend all of what was known about linear systems into the realm of controlled dynamical systems which could be described by a finite number of ordinary differential equations with inputs. A number of researchers joined the quest, and a sampling of seminal papers includes Brockett, [1972, 1973], Haynes and Hermes, [1970], Hermes and Haynes, [1963], Krener, [1974], and Sussmann and Jurdjevic, [1972]. The simple observation that differentiable manifolds were the natural state space structures for the study of controlled nonlinear differential equations opened a field of research that would be rich and active for many years.

While the field developed initially with the goal of extending linear state space control theory into the nonlinear domain, there was the early discovery of features of nonlinear systems for which no linear analogues exist. Specifically, the geometry of noncommuting vectorfields is a theme which was central in the earliest expositions of geometric nonlinear control theory, and this theme has remained central to many major developments over the past quarter century. In addition to laying down significant parts of the theoretical foundations of modern geometric control theory, Roger Brockett also did seminal work on important applications of the theory to the study of controlled

*The author is very grateful to Dr. Marc Jacobs of the Air Force Office of Scientific Research for many years of support of his research on the nonlinear control theory of mechanical systems. Currently this support is extended through grant number F49620-96-1-0059. Support from the Army Research Office under the ODDR&E MURI97 Program Grant No. DAAG55-97-1-0114 to the Center for Dynamics and Control of Smart Structures (through Harvard University) is also gratefully acknowledged. Finally, it is a pleasure to thank P.S. Krishnaprasad for his insights on the matters discussed herein.

mechanical systems. The 1976 NASA-Ames lectures (Brockett, [1976]) began a school of research that has now grown too large to survey in the present brief article. It is probably not coincidental that during the twenty-five year development of geometric nonlinear control theory, the field of analytical mechanics has itself been developed along increasingly geometric lines. (See, for instance Marsden and Ratiu, [1994].)

The starting point for this article is the global definition of a smooth nonlinear control system which appeared for the first time in Brockett, [1976]. In the next section, we repeat this definition and discuss its extension and specialization to the case of intrinsically second order systems. While such systems can be defined and studied in a purely formal way, we shall show that they arise naturally as Lagrangian models of *super-articulated* (also referred to as *under-actuated*) mechanical systems. In these systems, some of the configuration variables (or degrees of freedom) are directly controlled while the remaining variables evolve under the dynamic influence of the controlled degrees of freedom. We shall study such mechanical systems taking the viewpoint that the directly actuated degrees of freedom are inputs or controls which influence the remaining configuration variables through the given Lagrangian dynamics. Of course, in Lagrangian models, the velocities and accelerations of the directly actuated degrees of freedom also play a role in influencing the dynamics of the unactuated degrees of freedom, and thus it is natural to regard the inputs to these systems as triples comprised of the actuated configuration variables *together with their associated velocities and accelerations.*

In Section 3, we shall associate certain *curvature-like* quantities to our super-articulated Lagrangian control systems. Both *configuration flatness* and *input flatness* will be defined, and it will be shown that input flatness and configuration flatness together are necessary conditions for there to exist a change of coordinates such that in the new coordinates *(i)* the inertia tensor is diagonal, and *(ii)* no input accelerations enter the equations of motion. (In this case, we may say the systems are *velocity controlled.*) It will be noted that the *flatness* discussed in Section 3 is an intrinsically second order concept—related to ideas from classical Riemannian geometry. Thus we must distinguish this from *differential flatness* which has been introduced by Fliess (e.g. [15], and see also [23]).

In Section 4, we consider the same Lagrangian models of super-articulated mechanical systems under the influence of oscillatory inputs. For this class of systems, we wish to answer the question of when high frequency forcing of certain degrees of freedom will cause the total assemblage to respond with a prescribed stable behavior. A growing body of research on this question has been aimed at understanding the extent to which this question can be answered by an appropriate theory of averaging. Section 4 presents recent results on averaging Hamiltonian and Lagrangian systems, and in particular, we describe the role of the *averaged potential* in characterizing the qualitative dynamics of periodically forced mechanical systems. Configuration flatness together with input flatness again play a role in characterizing necessary conditions for the existence of a coordinate system in terms of which the averaged potential is obtained directly by simple averaging of the potential energy terms in the Lagrangian.

2 Second Order *Generalized* Control Systems

A local description of a nonlinear control system typically involves a differential equation of the form

$$\dot{q} = f(q, x), \tag{1}$$

where $f : M \times U \to TM$, with M being an (n-dimensional) *state manifold* and U an (m-dimensional) input manifold. Brockett [1976] pointed out that such models do not provide global descriptions of systems where the available input is state dependent. In the natural global extension of (1), we view f as a bundle map.

Definition 1 A *smooth nonlinear control system* is a quadruple (B, M, π, f) such that

(i) (B, M, π) is a fiber bundle with *total space* B, *base space* M, and canonical projection $\pi : B \to M$, and

(ii) $f : B \to TM$ is a bundle morphism such that for each $q \in M$ and $x \in U_q = \pi^{-1}(q)$, $f(q,x) \in T_qM$.

For a complete introduction to the theory of fiber bundles, the reader is referred to Husemoller [1994] or the classical treatise Steenrod [1951]. Recall that for each $q \in M$, there is a neighborhood V of q and a diffeomorphism ϕ mapping $\pi^{-1}(V)$ onto $V \times \pi^{-1}(q)$. Thus, as Brockett points out (Brockett, [1976], p. 16), by restricting our attention to such neighborhoods, it is always possible to find a *local* representation of a smooth nonlinear control system which is of the form (1).

To study controlled mechanical systems, it will be useful to extend models of the form (1) to include systems which are second order (in the configuration variables q) and in which the controlling effects of the input variables are primarily due to their accelerations, $\ddot{x}(\cdot)$. The descriptions of motions in these systems will involve the notion of parallel displacements of vectorfields in TM along curves in B. Given a vectorfield X on B and any curve $\tau(t) \in B$, the lifting of τ to a curve $X(\tau(t)) \in TB$ is projected by the tangent mapping π_* onto a curve $\pi_*(X(\tau(t))) \in TM$. This association leads to the desired notion of covariant differentiation of vectorfields in TM along curves in B, and we shall prescribe this in terms of local coordinates as follows. Let $(q_1,\dots,q_n,x_1,\dots,x_m)$ denote local coordinates defined in a neighborhood of a point $p \in B$. Let

$$X_i = \frac{\partial}{\partial q_i}, \quad (i = 1,\dots,n),$$

$$\tilde{X}_i = \frac{\partial}{\partial x_i}, \quad (i = 1,\dots,m)$$

be the associated vectorfields in TB. We may also view $(q_1,\dots,q_n)$ as defining local coordinates in a neighborhood of $\pi(p) \in M$. Then we similarly let

$$Y_i = \frac{\partial}{\partial q_i}, \quad (i = 1,\dots,n)$$

be the associated vectorfields in TM. In terms of this local coordinate description, for each X_i and each $(q,x) \in B$, the tangent mapping π_* associates a vector $Y_i^* \in T_qM$ by means of the formula

$$Y_i^*(q,x) = \pi_* X_i(q,x).$$

In general, $Y_i^*(q,x) \neq Y_i(q)$, but we may write

$$Y_i^*(q,x) = \sum_{j=1}^{n} \alpha_{ij}(x) Y_j(q), \quad (i = 1,\dots,n).$$

By restricting to a smaller neighborhood if necessary, there is no loss of generality in assuming that the $n \times n$ matrix $A(x)$, whose ij-th entry is $\alpha_{ij}(x)$, is nonsingular for each x. For any point $q \in M$, $\pi^{-1}(q) \in B$, each vectorfield X_i $(i = 1,\dots,n)$ and $\tilde{X}_j$ $(j = 1\dots,m)$, defines an integral curve passing through $\pi^{-1}(q)$. For each such integral curve, expressed in our coordinate neighborhood as $\tau(t) = (q(t), x(t))$, there is a corresponding set of curves $Y_i^*(q(t),x(t))$ $(i = 1\dots,n)$ in the tangent bundle TM. Tangents to these curves define a *covariant derivative operator* $\nabla : TB \times TM \to TM$ which is prescribed in terms of the given local coordinates by means of a total of $(n+m)n^2$ functions: $\{\Gamma_{ij}^k\}$ $(i,j,k = 1,\dots,n)$ and $\{\tilde{\Gamma}_{ij}^k\}$, $(i = 1,\dots,m;\ j,k = 1,\dots,n)$, and the formulas

$$\nabla_{X_i} Y_j^* = \sum_{k=1}^n \Gamma_{ij}^k Y_k^*, \qquad i,j = 1,\dots,n; \tag{2}$$

$$\nabla_{\tilde{X}_i} Y_j^* = \sum_{k=1}^n \tilde{\Gamma}_{ij}^k Y_k^*, \qquad i = 1,\dots,m;\ j = 1,\dots,n. \tag{3}$$

Since the vectorfields X_i (and $\tilde{X}_i$) lie in the coordinate directions, it is clear what is meant if we write

$$\nabla_{q_i} Y_j^* = \sum_{k=1}^{n} \Gamma_{ij}^k Y_k^*; \quad \nabla_{x_i} Y_j^* = \sum_{k=1}^{n} \tilde{\Gamma}_{ij}^k Y_k^*. \tag{4}$$

The covariant derivative operator defines a generalized affine connection and allows us to study displacement of vectorfields in TM along curves $(q(t), x(t)) \in B$. We refer to Kobayashi and Nomizu, [19], for a classical treatment of affine connections and parallel displacement. In the present development, we shall view points q in the base manifold as states of a nonlinear control system (and eventually as generalized coordinates of a mechanical control system) and the points x in the fiber as control inputs. We shall be interested in lifting curves in the total space B to curves in TB and projecting these liftings onto TM. It is useful to keep in mind the following commutative diagram

$$\begin{array}{ccc} TB & \xrightarrow{\pi_*} & TM \\ \downarrow & & \downarrow \\ B & \xrightarrow{\pi} & M \end{array}$$

where the vertical arrows are the canonical tangent bundle projections. As above, given a curve $(q(t), x(t))$ in B, we lift it to a family of curves $X_i(q(t), x(t))$ in TB $(1 \le i \le n)$ and define corresponding projections $Y_i^*(q(t), x(t)) = \pi_* X_i(q(t), x(t))$ in TM. If $(q(t), x(t))$ $(0 \le t < t_0)$ remains in our coordinate neighborhood in B, we have a valid local representation of the corresponding tangent vector

$$\sum_{i=1}^{n} \alpha_i(t) X_i(q(t), x(t)) + \sum_{i=1}^{m} \beta_i(t) \tilde{X}_i(q(t), x(t))$$

for each t in $0 \le t < t_0$. (This is a curve in TB.) The vectorfields $Y_i^*(q, x) = \pi_* X_i(q, x)$, appearing in the formulas (2)-(3) thus depend on points in the total space B. In terms of the vectorfields Y_i^*, the image curve in TM may be expressed as

$$Y(t) = \sum_{i=1}^{n} \alpha_i(t) Y_i^*(q(t), x(t)). \tag{5}$$

This projection defines a vectorfield (in TM) *along the curve* $\tau(t) = (q(t), x(t))$ in B. Differentiating Y with respect to t and using our covariant derivative operator to project the result back onto TM, we obtain the *covariant derivative in the direction* $\dot{\tau}$, $\nabla_{\dot{\tau}} Y$, for each t in $0 \le t < t_0$:

$$\sum_{i=1}^{n} \Big(\dot{\alpha}_i(t) Y_i^* + \alpha_i(t) (\sum_{j=1}^{n} \nabla_{q_j} Y_i^* \dot{q}_j + \sum_{j=1}^{m} \nabla_{x_j} Y_i^* \dot{x}_j) \Big).$$

(Cf. Kobayashi and Nomizu, [19] p. 114.) Using the relationships (4), this expression may be rendered

$$\sum_{k=1}^{n} \Big(\dot{\alpha}_k(t) + \sum_{i,j=1}^{n} \Gamma_{ji}^{k} \alpha_i \dot{q}_j + \sum_{i=1}^{n} \sum_{j=1}^{m} \tilde{\Gamma}_{ji}^{k} \alpha_i \dot{x}_j \Big) Y_k^*. \tag{6}$$

Using the notion of *parallel displacement*, we can now state conditions under which the curves of the form (5) are trajectories of a second order *generalized* control system. The conditions involve the vanishing of covariant derivatives, and thus these curves are natural generalizations to the input-dependent case of geodesic flows.

Definition 2 We say that a curve

$$V(t) = \sum_{i=1}^{n} v_i(t) Y_i^*(q(t), x(t))$$

is *parallel along* $\tau(t) = (q(t), x(t))$ if the covariant derivative in the direction $\dot{\tau}$, $\nabla_{\dot{\tau}} V$ is zero. Equivalently, in terms of local coordinates, $V(t)$ is parallel along $\tau(t)$ if

$$\sum_{i=1}^{n}\Big(\dot{v}_i(t)Y_i^* + v_i(t)(\sum_{j=1}^{n}[\nabla_{q_j}Y_i^*]\dot{q}_j + \sum_{j=1}^{m}[\nabla_{x_j}Y_i^*]\dot{x}_j)\Big) = 0.\square$$

An even more explicit rendering of the condition for Y in (5) to be parallel along τ is that $\alpha(\cdot)$ is the unique solution to the system of differential equations

$$\dot{\alpha}_k(t) + \sum_{i,j=1}^{n}\Gamma_{ji}^k\alpha_i\dot{q}_j + \sum_{i=1}^{n}\sum_{j=1}^{m}\tilde{\Gamma}_{ji}^k\alpha_i\dot{x}_j = 0 \quad (k = 1,\ldots,n).$$

Definition 3 We say that (5) defines a *smooth generalized second order control system* if *(i)* $\alpha_i(t) = \dot{q}_i(t) + \sum_{j=1}^{m}\gamma_{ij}(q)\dot{x}_j(t)$ where the γ_{ij} are mn smooth functions of q, and *(ii)* the vector field (5) is parallel along the curve $(q(t), x(t))$.$\square$

In particular, if (5) is a trajectory of smooth generalized second order control system, the coordinate functions $q_i(t)$ and $x_i(t)$ satisfy

$$\ddot{q}_k + \sum_{j=1}^{m}\gamma_{ij}(q)\ddot{x}_j \sum_{i,j=1}^{n}\Gamma_{ji}^k\dot{q}_i\dot{q}_j + \sum_{i=1}^{n}\sum_{j=1}^{m}\hat{\Gamma}_{ji}^k\dot{q}_i\dot{x}_j = 0, \quad \text{for } i = 1,\ldots,n, \tag{7}$$

where $\hat{\Gamma}_{ji}^k = \tilde{\Gamma}_{ji}^k + \frac{\partial\gamma_{ij}}{\partial q_i}$. This is a direct extension of the geometric characterization of second-order differential equations found in Abraham and Marsden, 1988, [1]. We employ the word *generalized* to indicate that derivatives of the input curves $x(\cdot)$ influence the dynamics in a direct and significant way. (See Fliess, [1990], for a broader discussion of generalized control systems.) In general, all three components of the triple $(x, \dot{x}, \ddot{x})$ enter the equation (7), but in Sections 4 and 5 we shall be principally concerned with high-frequency periodic inputs, and there the most significant terms will be those involving $\ddot{x}$. (Suppose $x(\cdot)$ is a periodic vector valued function whose fundamental frequency is ω and whose ℓ_∞ norm is $\mathcal{O}(1)$. Then the norm of $\dot{x}(\cdot)$ is $\mathcal{O}(\omega)$ and the norm of $\ddot{x}(\cdot)$ is $\mathcal{O}(\omega^2)$.) When equation (7) depends on $\ddot{x}$ that we shall call these systems *acceleration controlled.* In the next section, we shall study systems in which all $\gamma_{ij}(q) \equiv 0$. We shall also show that the vanishing of curvature-like quantities is a necessary condition for there to be a transformation to coordinates in terms of which x and $\dot{x}$ (but not $\ddot{x}$) enter the equations of motion.

Generalized Lagrangian Control Systems Generalized second order control systems arise in the study of *super-articulated* mechanical systems. (See Seto and Baillieul [1994] for the basic theory. These have also been called *under-actuated* systems in the literature.) The general framework assumes there is given a Lagrangian $L(y, \dot{y})$ defined on the tangent bundle TQ of the configuration space of a mechanical system. Suppose that the generalized coordinates can be partitioned as $y = (r, q)$ and that exogenous generalized forces (control inputs) can be applied to only the coordinates r, while the coordinate variables comprising q evolve freely, subject only to dynamic interactions with the r-variables. In this partitioning, r is an m–tuple and q is an n–tuple of generalized coordinate variables. The equations of motion for the system take the form

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{r}} - \frac{\partial L}{\partial r} = u, \tag{8}$$

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}} - \frac{\partial L}{\partial q} = 0, \tag{9}$$

where u is an m-vector of controls, and we suppose the mapping $\dot{q} \mapsto \frac{\partial L}{\partial \dot{q}}$ is invertible. That is, if the Lagrangian takes the form "kinetic minus potential energy," i.e. $L = \frac{1}{2}\dot{y}^T M\dot{y} - V(y)$, and the inertia matrix is partitioned conformably with (r, q), $M = \begin{pmatrix} \mathcal{N} & \mathcal{A} \\ \mathcal{A}^T & \mathcal{M} \end{pmatrix}$, then $\mathcal{M} = \mathcal{M}(r, q)$

is an $n \times n$ invertible matrix. It is also assumed that $\mathcal{N} = \mathcal{N}(r,q)$ is an invertible matrix, and this ensures a one-one correspondence between control trajectories $u(\cdot)$ and trajectories of of the r–variables. For the purposes of the model, we assume that the components $u_i(\cdot)$ are each piecewise analytic functions on $[0,\infty)$, although more general inputs (e.g. impulse trains) are also of interest and amenable to study.

A simple reduction process eliminating the explicit dependence on the input $u(\cdot)$ leads to a system of equations of the desired form in which the r–variables and their derivatives play the role of inputs with the triple $(x,v,a) = (r,\dot{r},\ddot{r})$ controlling the state variables $(q,\dot{q})$ according to equation (9). In this case the dynamical relationship (9) may also be obtained by applying the Euler-Lagrange operator $\frac{d}{dt}\frac{\partial}{\partial \dot{q}} - \frac{\partial}{\partial q}$ to the *reduced Lagarangian*

$$\mathcal{L}(q,\dot{q};x,v) = \frac{1}{2}\dot{q}^T\mathcal{M}(q,x)\dot{q} + v^T\mathcal{A}(q,x)\dot{q} - \mathcal{V}(q;x,v), \tag{10}$$

where $\mathcal{V}(q;x,v) = V(q,x) - \frac{1}{2}v^T\mathcal{N}(q,x)v$. Writing out the equations of motion explicitly, we have

$$\sum_{j=1}^{n} m_{kj}\ddot{q}_j + \sum_{\ell=1}^{m} a_{\ell k}\dot{v}_\ell + \sum_{i,j=1}^{n} \Gamma_{ijk}\dot{q}_i\dot{q}_j + \sum_{j=1}^{n}\sum_{\ell=1}^{m} \hat{\Gamma}_{\ell jk} v_\ell \dot{q}_j = F(t), (k=1,\ldots,n), \tag{11}$$

where

$$\Gamma_{ijk} = \frac{1}{2}\Big(\frac{\partial m_{ki}}{\partial q_j} + \frac{\partial m_{kj}}{\partial q_i} - \frac{\partial m_{ij}}{\partial q_k}\Big),$$

$$\hat{\Gamma}_{\ell jk} = \frac{\partial m_{kj}}{\partial x_\ell} + \frac{\partial a_{\ell k}}{\partial q_j} - \frac{\partial a_{\ell j}}{\partial q_k},$$

and a_{ij} and m_{ij} are the ij-th entries in the $m\times n$ and $n\times n$ matrices $\mathcal{A}(q,x)$ and $\mathcal{M}(q,x)$ respectively. $F(t)$ is a vector of *generalized forces* $F_i(t) = \frac{\partial \mathcal{V}}{\partial q_i} - \sum_{k,\ell=1}^{m} \frac{\partial a_{\ell i}}{\partial x_k} v_\ell v_k$. These may be thought of as coming from a velocity-dependent potential if

$$\frac{\partial^2 a_{\ell i}}{\partial q_j \partial x_k} = \frac{\partial^2 a_{\ell j}}{\partial q_i \partial x_k}$$

for all $k,\ell = 1,\ldots,m$ and $i,j = 1\ldots,n$. In classical mechanics, the quantities Γ_{ijk} defined in this way in terms of the inertia tensor $\mathcal{M}$ are called *Christoffel symbols of the first kind.* To be consistent with this nomenclature, we call the $\hat{\Gamma}_{\ell jk}$ *input symbols of the first kind.*

To define corresponding symbols of the second kind, let m^{ij} denote the ij–th element of $\mathcal{M}^{-1}$. Multiplying both sides of (11) by $m^{\sigma k}$ and summing over the index values $k=1,\ldots,n$, we obtain

$$\ddot{q}_\sigma + \sum_{\ell=1}^{m} \gamma_{\sigma\ell}\dot{v}_\ell + \sum_{i,j=1}^{n} \Gamma_{ij}^{\sigma}\dot{q}_i\dot{q}_j + \sum_{j=1}^{n}\sum_{\ell=1}^{m} \hat{\Gamma}_{\ell j}^{\sigma} v_\ell\dot{q}_j = \tilde{F}_\sigma(t), \quad (\sigma=1,\ldots,n), \tag{12}$$

where

$$\gamma_{\sigma\ell} = \sum_{k=1}^{n} m^{\sigma k} a_{k\ell},$$

$$\Gamma_{ij}^{\sigma} = \sum_{k=1}^{n} m^{\sigma k}\Gamma_{ijk},$$

$$\hat{\Gamma}_{\ell j}^{\sigma} = \sum_{k=1}^{n} m^{\sigma k}\hat{\Gamma}_{\ell jk}, \text{ and}$$

$$\tilde{F}_\sigma(t) = \sum_{k=1}^{n} m^{\sigma k} F_k(t),$$

$\ell = 1, \ldots, m;\ \sigma, i, j = 1, \ldots, n$. The quantities Γ^{σ}_{ij} are called *Christoffel symbols of the second kind*, and the $\hat{\Gamma}^{\sigma}_{\ell j}$'s will be called *input symbols of the second kind.* Modulo a minor change of notation ($\dot{x}_j = v_j$), when the generalize potential forces $F_k(t) = 0$, equations (7) and (12) are the same. To be consistent with the terminology introduced above for abstract second order systems, we shall call systems of the form (12) *generalized Lagrangian control systems.* Whenever at least one $\gamma_{ij} \neq 0,\ i = 1, \ldots, n;\ j = 1, \ldots, m$, these could also be referred as *acceleration-controlled Lagrangian systems.* This usage is quite informal at this point, and it will be shown below that under certain geometric conditions there is a change of coordinates which eliminates the explicit dependence of the dynamics on any of the $\dot{v}_i$ terms. Indeed, one of the main results of the chapter is to provide an invariant characterization of generalized Lagrangian control systems which are intrinsically acceleration controlled.

Remark 1 *Connections with classical Lagrangian reduction.* Marsden [1992] describes Lagrangian reduction for mechanical systems with symmetries which are characterized in terms of a group action that leaves the (unreduced) Lagrangian invariant. For Abelian group actions, this reduction was known to Routh in the nineteenth century, but it has been extended to the non-Abelian case by Marsden and various others cited in Marsden [1992]. The idea is that when a Lagrangian is constant on orbits of the tangent lift of a group action by a symmetry group G acting freely on the configuration manifold Q, these orbits can be "factored out" to yield a reduced configuration space Q/G on which there is defined a reduced Lagrangian system. Symmetries are the key to this reduction. In the above reduction of Lagrangian systems with inputs, we have assumed no particular structure other than a natural partitioning of configuration space variable into two sets: those variables that are directly controlled and those variables whose motions arise solely from dynamic interactions. Without further assumptions, there is little that can be said in general about the dynamics of our reduced Lagrangian systems (12). In the next section, we shall show that curvature-like quantities, defined by the Christoffel symbols and input symbols, and related notions of *flatness* play a role in characterizing key features of systems of the form (12). When these quantities vanish, there is a choice of coordinates with respect to which the inputs and configuration variables are highly decoupled. The "vanishing" of certain additional curvature-like quantities is a necessary condition for there to be a choice of coordinates such that the dynamic effects of input triples $(x, \dot{x}, \ddot{x})$ are completely determined by the configuration-velocity pair $(x, \dot{x})$ alone. These terms will enter the Lagrangian model through velocity dependent potential terms.

Remark 2 *Invariance properties of the reduced Lagrangian.* An important feature of Lagrangian mechanics is the invariance of Lagrange's equations with respect to coordinate changes. Reduced Lagrangians of the form (10) enjoy essentially the same invariance with respect to coordinate transformations of the form $Q = F(q)$ which are independent of the input x. Letting the inverse transformation be written $q = G(Q)$, the reduced Lagrangian in Q-coordinates is expressed as

$$\tilde{\mathcal{L}}(Q, \dot{Q}; x, v) = \frac{1}{2}\dot{Q}^T \tilde{\mathcal{M}}(Q, x)\dot{Q} + v^T \tilde{\mathcal{A}}(Q, x)\dot{Q} - \tilde{\mathcal{V}}(Q; x, v), \tag{13}$$

where

$$\tilde{\mathcal{M}}(Q, x) = \frac{\partial G}{\partial Q}^T \mathcal{M}(G(Q), x)\, \frac{\partial G}{\partial Q}$$

$$\tilde{\mathcal{A}}(Q, x) = \mathcal{A}(G(Q), x)\, \frac{\partial G}{\partial Q}, \quad \text{and}$$

$$\tilde{\mathcal{V}}(Q; x, v) = \mathcal{V}(G(Q); x, v).$$

Invariance with respect to such a coordinate transformation means the equations of motion for the system in Q-coordinates are given by

$$\frac{d}{dt}\frac{\partial \tilde{\mathcal{L}}}{\partial \dot{Q}} - \frac{\partial \tilde{\mathcal{L}}}{\partial Q} = 0.$$

In the following sections, it will be of interest to allow coordinate transformations of the the configuration variables which are input dependent: $Q = F(q,x)$. In this case it is possible to propose an extended notion of invariance for the equations of motion. Specifically, suppose that for some (q_0, x_0) there is a neighborhood $U \times V \subset \mathbb{R}^n \times \mathbb{R}^m$ containing (q_0, x_0) and a mapping $F : U \times V \to \mathbb{R}^n$ such that for each $x \in V$, $Q = F(q,x)$ is a diffeomorphism on U. For each x, write the inverse of this configuration space diffeomorphism as $q = G(Q,x)$. Then in terms of the Q-coordinates, the reduced Lagrangian again has the form (13) where

$$\tilde{\mathcal{M}}(Q,x) = \frac{\partial G}{\partial Q}^T \mathcal{M}(G(Q,x),x)\, \frac{\partial G}{\partial Q} \tag{14}$$

$$\tilde{\mathcal{A}}(Q,x) = \frac{\partial G}{\partial x}^T \mathcal{M}(G(Q,x),x)\, \frac{\partial G}{\partial Q} + \mathcal{A}(G(Q,x),x)\, \frac{\partial G}{\partial Q}, \text{ and} \tag{15}$$

$$\begin{aligned} \tilde{\mathcal{V}}(Q;x,v) \quad = \quad & -\dot{x}^T\Big(\mathcal{A}(G(Q,x),x)\, \tfrac{\partial G}{\partial x} + \tfrac{1}{2}\tfrac{\partial G}{\partial x}^T \mathcal{M}(G(Q,x),x)\, \tfrac{\partial G}{\partial x}\Big)\dot{x} \\ & + \mathcal{V}(G(Q,x);x,v). \end{aligned} \tag{16}$$

Again, we find that our reduced Lagrangian formulation is invariant under this class of input dependent coordinate transformations. This is summarized in the following proposition.

Proposition 1 *Consider a mechanical system with equations of motion prescribed by*

$$\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{q}} - \frac{\partial \mathcal{L}}{\partial q} = 0,$$

where $\mathcal{L}$ is given by (10). Let $Q = F(q,x)$ be an input dependent change of configuration coordinates as described above. Then, in Q-coordinates the equations of motion take the same form:

$$\frac{d}{dt}\frac{\partial \tilde{\mathcal{L}}}{\partial \dot{Q}} - \frac{\partial \tilde{\mathcal{L}}}{\partial Q} = 0, \tag{17}$$

where $\tilde{\mathcal{L}}$ is given by (13), with $\mathcal{M}$, $\mathcal{A}$, and $\mathcal{V}$ given by (14), (15), and (16) respectively.

Proof It is straightforward to show that the reduced Lagrangian has the form claimed when expressed in terms of Q–coordinates. It is a slightly tedious but also straightforward calculation to show that the equations of motion are given by (17). Alternatively, if the reduced Lagrangian (11) arises through the reduction process we have described, it may be shown that the claimed invariance follows from the invariance of the classical Lagrange's equations.□

3 Flat Systems and Systems with Flat Inputs

This section proposes some basic elements of a theory of normal forms for the velocity and acceleration controlled mechanical systems which were introduced in the previous section. We start by considering an uncontrolled mechanical system which is *simple* in the sense that it has a Lagrangian of the form

$$L = \frac{1}{2}\dot{q}^T M(q)\dot{q} - V(q), \tag{18}$$

comprised of a kinetic and potential energy term. (Cf. Abraham and Marsden, 1978.) The results we present are steps in a program aimed at developing an intrinsically second order theory of mechanical control systems.

In the study of first order nonlinear control systems of the form (1), important invariant qualitative features of the dynamics are revealed by computing Lie brackets of vectorfields in the family $\{f(\cdot,x) : x \in U\}$. We refer to Nijmeijer and van der Schaft [1990] for basic results on Lie bracket conditions related to *controllability*, *accessibility*, *feedback linearizability*, etc. The theory of first order control systems also applies to the second order systems treated in the previous section if we represent them in the usual way as first order systems (of twice the dimension) with generalized forces playing the role of control inputs. Our goal in the present section is to describe an intrinsically second order approach in which *curvature* provides an invariant measure of nonintegrability. It is worth remarking that while there has been a great deal written over the past 25 years making connections between analytical mechanics and nonlinear control, a question which remains almost completely unanswered is whether there can be found an intrinsically second order theory of controlled mechanical systems. Almost the last word on this subject in fact appeared in Brockett's NASA Ames lectures.

The section discusses the role of curvature in three different settings. First, we recall the basic result that the vanishing of the curvature tensor associated with the Levi-Civita connection of an uncontrolled simple mechanical system (18) is necessary and sufficient for the existence of a coordinate system with respect to which the inertia tensor is the identity matrix. Second, for the case of a simple mechanical system in which every degree of freedom is independently actuated, we recall that in the Hamiltonian setting the same curvature vanishing implies the existence of a (canonical) coordinate transformation together with a state feedback transformation, not involving conjugate momentum variables, transforming the dynamics into a decoupled set of second order integrators. Finally, we consider generalized Lagrangian control systems of the form introduced in the previous section. For these systems, there is defined an input connection and related curvature-like quantities. The vanishing of both the curvature defined by the Levi-Civita connection and the curvature defined by the input connection is necessary for the existence of an input-dependent change of configuration coordinates under which the inertia tensor is transformed to the identity matrix and all terms depending on input accelerations in the dynamics are eliminated. Thus for the subclass of generalized Lagrangian control systems for which the Levi-Civita connection has zero curvature, the vanishing of the curvature of the input connection is a necessary condition for an acceleration controlled mechanical system to be reducible through a change of coordinates to a velocity controlled mechanical system.

The *simple* Lagrangian (18) gives rise to the Lagrangian dynamics:

$$\sum_{j=1}^{n} m_{ij}(q)\ddot{q}_j + \sum_{j,k=1}^{n} \Gamma_{kji}\dot{q}_k\dot{q}_j + \frac{\partial V}{\partial q_i} = 0, \quad (i = 1, \ldots, n), \tag{19}$$

where the Γ_{kji} are the Christoffel symbols of the first kind introduced in the previous section. The terms they define in equation (19) describe the inertial forces (Coriolis and centrifugal) which affect the system. It is clear from this equation that if M does not depend on q, $\Gamma_{kji} = 0$ for all k, j, i, and these inertial forces will not play a role in the system's dynamics. An interesting question is "When can we find a change of coordinates such that in the new coordinate system the inertia matrix is constant?" Note that the inertia matrix M is constant (i.e. does not depend on the configuration variables q) if and only if there is a system of generalized coordinates $(y_1, \ldots, y_n)$ in terms of which the kinetic energy is expressed as $\frac{1}{2}\sum_{i=1}^{n} \dot{y}_i^2$. (*Proof:* In the case that M is constant, we let Y denote the constant, symmetric, positive definite matrix square root of M and let $y = Yq$. Then $\dot{y} = Y\dot{q}$, and the kinetic energy is $\frac{1}{2}\sum_{i,j} m_{ij}\dot{q}_i\dot{q}_j = \frac{1}{2}\sum_i \dot{y}_i^2$. The reverse implication requires no proof.)

Hence, whether we can find the desired change of coordinates amounts to whether there exists a diffeomorphism $y = F(q)$ such that $\frac{1}{2}\dot{q}^T M\dot{q} = \frac{1}{2}\|\dot{y}\|^2$. Suppose such an F exists. Then $\dot{y} = \frac{\partial F}{\partial q}\dot{q}$, and $\frac{\partial F}{\partial q}^T \frac{\partial F}{\partial q} = M$. For such an M we compute

$$\frac{\partial m_{ij}}{\partial q_k} = \sum_{\ell=1}^{n} \left(\frac{\partial^2 F_\ell}{\partial q_i \partial q_j}\frac{\partial F_\ell}{\partial q_j} + \frac{\partial F_\ell}{\partial q_i}\frac{\partial^2 F_\ell}{\partial q_j \partial q_k}\right).$$

From this it follows from an easy calculation that

$$\begin{aligned}\Gamma_{kji} &= \frac{1}{2}\Big(\frac{\partial m_{ij}}{\partial q_k} + \frac{\partial m_{ik}}{\partial q_j} - \frac{\partial m_{kj}}{\partial q_i}\Big) \\ &= \sum_{\ell=1}^{n} \frac{\partial^2 F_\ell}{\partial q_j \partial q_k}\frac{\partial F_\ell}{\partial q_i}. \end{aligned} \tag{20}$$

As defined in Section 2, the Christoffel symbols (of the first kind) Γ_{kji} are associated with a *connection* (the so-called Levi-Civita connection defined by the inertia tensor M) and a *Riemannian curvature tensor.* It is a classical result that this tensor vanishes if and only if there is a diffeomorphic change of coordinates $Q = F(q)$ such that the inertia tensor may be factored as $M = J_F^T \cdot J_F$ where J_F is the Jacobian of F. To see this, we first write down the corresponding Christoffel symbols of the second kind:

$$\Gamma_{kj}^{\ell} = \sum_{i=1}^{n} m^{\ell i}\Gamma_{kji}.$$

The Riemannian curvature tensor is prescribed by the corresponding Riemann symbols of the first kind:

$$R_{ijk\ell} = \frac{\partial \Gamma_{jki}}{\partial q_\ell} - \frac{\partial \Gamma_{j\ell i}}{\partial q_k} + \sum_{\sigma=1}^{n}\big(\Gamma_{j\ell}^{\sigma}\Gamma_{ik\sigma} - \Gamma_{jk}^{\sigma}\Gamma_{i\ell\sigma}\big). \tag{21}$$

If $F(\cdot)$ is a diffeomorphism (local coordinate transformation) such that $M = J_F^T \cdot J_F$,

$$\frac{\partial \Gamma_{j\ell i}}{\partial q_k} = \sum_{\sigma=1}^{n}\Big(\frac{\partial^3 F_\sigma}{\partial q_k \partial q_j \partial q_\ell}\frac{\partial F_\sigma}{\partial q_i} + \frac{\partial^2 F_\sigma}{\partial q_j \partial q_\ell}\frac{\partial^2 F_\sigma}{\partial q_i \partial q_k}\Big),$$

and a permutation of indices gives a similar expression for $\frac{\partial \Gamma_{jki}}{\partial q_\ell}$. It is straightforward, although a little tedious, to show that

$$\sum_{\sigma=1}^{n}\big(\Gamma_{jk}^{\sigma}\Gamma_{i\ell\sigma} - \Gamma_{j\ell}^{\sigma}\Gamma_{ik\sigma}\big) = \sum_{\beta=1}^{n}\Big(\frac{\partial^2 F_\beta}{\partial q_j \partial q_k}\frac{\partial^2 F_\beta}{\partial q_i \partial q_\ell} - \frac{\partial^2 F_\beta}{\partial q_j \partial q_\ell}\frac{\partial^2 F_\beta}{\partial q_i \partial q_k}\Big).$$

Hence $R_{ijkl} = 0$, proving that the curvature vanishing is a necessary condition for there to be a factorization of the desired type. For a proof of sufficiency and a broader discussion of curvature and its vanishing, the reader is referred to Kobayashi and Nomizu [1963] or Wolf [1972].

Remark 3 The Riemannian curvature may equivalently be characterized in terms of *Riemann symbols of the second kind*:

$$\begin{aligned} R_i{}^j{}_{k\ell} &= \frac{\partial \Gamma_{ki}^{j}}{\partial q_\ell} - \frac{\partial \Gamma_{\ell i}^{j}}{\partial q_k} + \sum_{\sigma=1}^{n}\big(\Gamma_{ki}^{\sigma}\Gamma_{\ell\sigma}^{j} - \Gamma_{\ell i}^{\sigma}\Gamma_{k\sigma}^{j}\big) \\ & \quad i, j, k, \ell = 1, \ldots, n. \end{aligned}$$

It is not difficult to prove the very direct relationship between Riemann symbols of the first and second kinds:

$$R_i{}^j{}_{k\ell} = \sum_{\sigma=1}^{n} m^{j\sigma}R_{\sigma ik\ell}.$$

Remark 4 While there are n^4 components of the curvature tensor, these are not all independent, and the Riemann symbols associated with the Levi-Civita connection can be shown to satisfy:

$$R_{ijk\ell} + R_{ik\ell j} + R_{i\ell jk} = 0,$$

$$R_{ijk\ell} = -R_{ij\ell k}, \quad \text{and}$$
$$R_{ijk\ell} = R_{k\ell ij}.$$

(See Dubrovin *et al.* [1984], Theorem 30.2.1.) It is not difficult to show that these symmetries imply that only $n^2(n^2-1)/12$ of the Riemann symbols of the first kind are independent.

Definition 4 A simple mechanical system of the form (19) will be called *flat* if the corresponding Riemannian curvature tensor (i.e. the set of Riemann symbols of the first kind) is zero.

Bedrosian [1992] and Spong [1998] have noted that this type of flatness points to simplified control designs in the case of fully actuated Lagrangian and Hamiltonian systems. We review the basic ideas and then turn our attention to the more complex (and more interesting) case of super-articulated (or underactuated) systems. Consider a mechanical control system specified by a Lagrangian $L(q,\dot{q}) = \frac{1}{2}\dot{q}^T M(q)\dot{q} - V(q)$ and equations of motion

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = u_i \quad (i = 1, \ldots, n),$$

where each $u_i(\cdot)$ is a piecewise continuous control input (function of time defined on a suitable interval). The conjugate momentum vector for this system is $p = M(q)\dot{q}$ and in terms of the Hamiltonian defined by the Legendre transformation

$$\begin{aligned} H(q,p) &= \dot{q}^T p - L(q,\dot{q}) \\ &= \frac{1}{2}p^T M(q)^{-1} p + V(q), \end{aligned} \tag{22}$$

we define the corresponding Hamiltonian control system

$$\begin{aligned} \dot{q} &= \frac{\partial H}{\partial p} \\ \dot{p} &= -\frac{\partial H}{\partial q} + u. \end{aligned} \tag{23}$$

It has been noted (See, e.g. Bedrosian [1992] and Spong [1998].) that if a system is *flat* there is a canonical transformation together with a control and feedback transformation such that in terms of the new phase space and control variables, the system (23) takes the form of a system of double integrators.

To understand this, suppose the mechanical system is flat. Let $M^{\frac{1}{2}}(q)$ be the $n \times n$ positive definite symmetric square root of the inertia tensor $M(q)$. If there exists a function $F(q)$ such that $\frac{\partial F}{\partial q} = M^{\frac{1}{2}}(q)$ as above, then the coordinate transformation $Q = F(q)$, $P = M^{\frac{1}{2}}(q)\dot{q}$ is canonical—in the sense that it preserves the symplectic form. We refer the reader to Marsden and Ratiu [1994] or Arnold [1989] for a discussion of canonical transformations. That this transformation is canonical is easily seem if we write

$$\begin{aligned} P \cdot dQ &= \dot{q}^T M^{\frac{1}{2}}(q) \cdot \frac{\partial F}{\partial q} dq \\ &= \dot{q}^T M(q) dq \\ &= p \cdot dq. \end{aligned}$$

We rewrite the Hamiltonian in terms of $Q, P-$coordinates:

$$\begin{aligned} H(q,p) &= \frac{1}{2}p^T M(q)^{-1} p + V(q) \\ &= \frac{1}{2}\dot{q}^T M^{\frac{1}{2}}(q) M^{\frac{1}{2}}(q)\dot{q} + V(q) \\ &= \frac{1}{2}\|P\|^2 + \mathcal{V}(Q), \\ &= \mathcal{H}(Q,P), \end{aligned}$$

where $\mathcal{V}$ is defined by $\mathcal{V}(Q) = V(q)$.

Proposition 2 *In terms of the variables Q, P, the equations (23) may be written as*

$$\begin{aligned} \dot{Q} &= P \\ \dot{P} &= -\frac{\partial \mathcal{V}}{\partial Q} + M^{-\frac{1}{2}}u. \end{aligned} \tag{24}$$

Under the feedback $u = M^{\frac{1}{2}}v + M^{\frac{1}{2}}\frac{\partial \mathcal{V}}{\partial Q}$, which is independent of the conjugate momentum variables P, this Hamiltonian control system is transformed into the system of double integrators

$$\begin{aligned} \dot{Q} &= P \\ \dot{P} &= v \end{aligned} \tag{25}$$

Proof: The coordinate change associated with the Legendre transformation may be written explicitly as

$$Q = F(q) \tag{26}$$

$$P = \left(\frac{\partial F}{\partial q}\right)^{-1} p \tag{27}$$

(since $\frac{\partial F}{\partial q}\dot{q} = (\frac{\partial F}{\partial q})^{-1}p$). We wish to find the equations of motion in terms of $Q, P-$coordinates. Differentiating (26), it is clear that $\dot{Q} = P$. Differentiating (27), we find that

$$\begin{aligned} \dot{P} &= [\frac{d}{dt}\left(\frac{\partial F}{\partial q}\right)^{-1}]p + \left(\frac{\partial F}{\partial q}\right)^{-1}\dot{p} \\ &= [\frac{d}{dt}\left(\frac{\partial F}{\partial q}\right)^{-1}]p - \left(\frac{\partial F}{\partial q}\right)^{-1}\frac{\partial H}{\partial q} + \left(\frac{\partial F}{\partial q}\right)^{-1}u. \end{aligned}$$

Using the definition of the coordinate transformation (26)-(27) together with the symmetry properties of $\frac{\partial F}{\partial q} = M^{\frac{1}{2}}(q)$, it can be shown that

$$\begin{pmatrix} \frac{\partial Q}{\partial q} & 0 \\ \frac{\partial P}{\partial q} & \frac{\partial P}{\partial p} \end{pmatrix} \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} \begin{pmatrix} \frac{\partial Q}{\partial q}^T & \frac{\partial P}{\partial q}^T \\ 0 & \frac{\partial P}{\partial p}^T \end{pmatrix} = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}.$$

From this it follows that

$$[\frac{d}{dt}\left(\frac{\partial F}{\partial q}\right)^{-1}]p - \left(\frac{\partial F}{\partial q}\right)^{-1}\frac{\partial H}{\partial q} = -\frac{\partial \mathcal{H}}{\partial Q} = -\frac{\partial \mathcal{V}}{\partial Q},$$

proving the first part of the proposition. The second part is clear. □

More generally, a system of the form (7) or (12) is said to be flat if the Riemannian curvature tensor prescribed by the Christoffel symbols Γ_{ij}^k vanishes. In terms of the Lagrangian system (10), if the inertia tensor $\mathcal{M}$ does not depend on the input x, then the vanishing of the curvature tensor is necessary and sufficient for there to be a change of generalized coordinates $\overline{q} = F(q)$ such that the inertia tensor expressed in terms of the $\bar{q}-$coordinates has entries

$$\delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j. \end{cases}$$

For the case in which $\mathcal{M} = \mathcal{M}(q, x)$, explicitly depends on the input x, it is still possible to have the Riemannian curvature tensor vanish. (Of course, in this case, the Riemann symbols themselves will have no dependence on the variable x.) This vanishing is a necessary and sufficient condition for there to be an $x-$dependent coordinate transformation with respect to which $\mathcal{M}$ is transformed into the identity matrix. Let $\bar{q} = F(q,x)$ be such a coordinate transformation, and let $q = G(\bar{q},x)$ be its x–dependent inverse. Then $\mathcal{M}(q,x) = \frac{\partial F}{\partial q}^T \frac{\partial F}{\partial q}$, and in $\bar{q}$–coordinates, the Lagrangian (10) takes the form

$$\bar{\mathcal{L}}(\bar{q},\dot{\bar{q}};x,\dot{x}) = \frac{1}{2}\|\dot{\bar{q}}\|^2 + \dot{x}^T \bar{\mathcal{A}}(\bar{q},x)\dot{\bar{q}} - \bar{\mathcal{V}}(\bar{q};x,\dot{x}), \tag{28}$$

where

$$\bar{\mathcal{A}}(\bar{q},x) = \mathcal{A}(G(\bar{q},x),x)\frac{\partial G}{\partial \bar{q}} - \frac{\partial F}{\partial x}^T,$$

and

$$\bar{\mathcal{V}}(\bar{q};x,\dot{x}) = \mathcal{V}(G(\bar{q},x);x,\dot{x}) - \dot{x}^T \mathcal{A}\frac{\partial G}{\partial x} - \frac{1}{2}\dot{x}^T \frac{\partial F}{\partial x}^T \frac{\partial F}{\partial x}\dot{x},$$

with all component entries in $\frac{\partial F}{\partial x}$ being evaluated at $q = G(\bar{q},x)$. An interesting question is "When do the coupling terms $\bar{\mathcal{A}}(\bar{q},x)$ vanish?" The following definition is useful in formulating the answer.

Definition 5 *(i)* For the second-order generalized control system (12), we say that the hatted symbols $\hat{\Gamma}^k_{ij}$ define the *input connection.* The input connection is said to be *flat* if the associated Riemann symbols of the second kind

$$\hat{R}_j{}^\gamma{}_{\ell i} = \frac{\partial \hat{\Gamma}^\gamma_{\ell j}}{\partial x_i} - \frac{\partial \hat{\Gamma}^\gamma_{ij}}{\partial x_\ell} + \sum_{k=1}^n (\hat{\Gamma}^k_{\ell j}\hat{\Gamma}^\gamma_{ik} - \hat{\Gamma}^k_{ij}\hat{\Gamma}^\gamma_{\ell k})$$

vanish for all $i,\ell = 1,\dots,m$; $j,\gamma = 1,\dots,n$. *(ii)* Given the second order generalized control system (12), we say that the pair $(x,\dot{x})$ constitutes a set of *flat inputs* if the input connection is flat.□

The desired result on the structure of the Lagrangian control system (12) is the following.

Theorem 1 Consider the Lagrangian control system (10). Let $U \times V \subset \mathbb{R}^n \times \mathbb{R}^m$, and suppose that $F : U \times V \to \mathbb{R}^n$ is an input-dependent change of coordinates $\bar{q} = F(q,x)$ such that for each x the metric tensor $\mathcal{M}$ expressed in $\bar{q}$–coordinates has δ_{ij} as its ij–th entry. Suppose, moreover, that in the $\bar{q}$–coordinates all cross coupling terms $\bar{\mathcal{A}}(\bar{q},x)$ vanish. Then the system (defined by (10)) is flat and has flat inputs, which is to say both the hatted and unhatted Riemann symbols of the second kind,

$$R_j{}^\gamma{}_{\alpha\beta} = \frac{\partial \Gamma^\gamma_{\alpha j}}{\partial q_\beta} - \frac{\partial \Gamma^\gamma_{\beta j}}{\partial q_\alpha} + \sum_{k=1}^n (\Gamma^k_{\alpha j}\Gamma^\gamma_{\beta k} - \Gamma^k_{\beta j}\Gamma^\gamma_{\alpha k})$$
$$j,\gamma,\alpha,\beta = 1,\dots,n,$$

and

$$\hat{R}_j{}^\gamma{}_{\ell i} = \frac{\partial \hat{\Gamma}^\gamma_{\ell j}}{\partial x_i} - \frac{\partial \hat{\Gamma}^\gamma_{ij}}{\partial x_\ell} + \sum_{k=1}^n (\hat{\Gamma}^k_{\ell j}\hat{\Gamma}^\gamma_{ik} - \hat{\Gamma}^k_{ij}\hat{\Gamma}^\gamma_{\ell k})$$
$$i,\ell = 1,\dots,m; \quad j,\gamma = 1,\dots,n,$$

vanish.

The proof of this theorem uses the following lemma.

Lemma 1 Let $U \times V \subset \mathbb{R}^m \times \mathbb{R}^n$ be some neighborhood of $(0,0)$, and let $f_j : U \times V \to \mathbb{R}^n$ be a smooth mapping for $j = 1,\dots,m$. Given $q \in V$, there is defined a neighborhood W of $0 \in \mathbb{R}^m$ and a smooth function $g : W \to \mathbb{R}^n$ such that

(i) $g(0) = q$, and

(ii) $\frac{\partial g}{\partial r_j}(r) = f_j(r, g(r))$ for all $r \in W$, $j = 1, \dots, m$

if and only if there is a neighborhood of $(0, q)$ on which

$$\frac{\partial f_j}{\partial r_i} - \frac{\partial f_i}{\partial r_j} + \sum_{k=1}^{n}\left(\frac{\partial f_j}{\partial q_k} f_i^k - \frac{\partial f_i}{\partial q_k} f_j^k\right) = 0.$$

This lemma is proved in Spivak, [1970]. We can now prove Theorem 1.

Proof (of Theorem 1). Let $\bar{q} = F(q, x)$ be as in the hypothesis of the theorem. Then using Proposition 1 at the end of Section 2, the $\bar{q}$–coordinate rendering of the Lagrangian (10) is

$$\bar{\mathcal{L}}(\bar{q}, \dot{\bar{q}}; x, v) = \frac{1}{2}\dot{\bar{q}}^T \bar{\mathcal{M}}(\bar{q}, x)\dot{\bar{q}} + v^T \bar{\mathcal{A}}(\bar{q}, x)\dot{\bar{q}} - \bar{\mathcal{V}}(\bar{q}; x, v)$$

where the various terms correspond to those given by formulas (14)-(16).

The hypothesis that $\mathcal{M}$ has coordinate entries δ_{ij} is equivalent to writing $\frac{\partial G}{\partial \bar{q}}^T \mathcal{M} \frac{\partial G}{\partial \bar{q}} = I$, and since $\frac{\partial G}{\partial \bar{q}} = \frac{\partial F}{\partial q}^{-1}$, we conclude that $\mathcal{M}(q, x) = \frac{\partial F}{\partial q}^T \frac{\partial F}{\partial q}$. The argument used at the beginning of this section shows that the Reimannian curvature tensor associated with this $\mathcal{M}(q, x)$ is therefore zero.

The second part of the hypothesis is the statement that

$$\frac{\partial G}{\partial x}^T \mathcal{M}(q, x) + \mathcal{A}(q, x) \equiv 0.$$

Using the fact that

$$\frac{\partial G}{\partial \bar{q}}(F(q, x), x) = \frac{\partial F}{\partial q}^{-1}(q, x)$$

together with the assumed factorization $\mathcal{M} = \frac{\partial F}{\partial q}^T \frac{\partial F}{\partial q}$, some straightforward algebra shows the above equation is equivalent to

$$\mathcal{A}(G(\bar{q}, x), x)\frac{\partial G}{\partial \bar{q}}(G(\bar{q}, x), x) \equiv \frac{\partial F}{\partial x}^T (G(\bar{q}, x), x),$$

which is in turn equivalent to writing

$$\mathcal{A}(q, x) = \frac{\partial F}{\partial x}^T \frac{\partial F}{\partial q} \tag{29}$$

in terms of the original coordinates. The componentwise rendering of this equation is

$$a_{\ell i} = \sum_{k=1}^{n} \frac{\partial F_k}{\partial x_\ell}\frac{\partial F_k}{\partial q_i} \qquad \begin{array}{l} \ell = 1, \dots, m \\ i = 1, \dots, n. \end{array}$$

Similarly, under the hypothesis of the theorem,

$$m_{ij} = \sum_{k=1}^{n} \frac{\partial F_k}{\partial q_i}\frac{\partial F_k}{\partial q_j} \quad i, j = 1 \dots, n.$$

It is a straightforward calculation to show that

$$\frac{\partial m_{kj}}{\partial x_\ell} + \frac{\partial a_{\ell k}}{\partial q_j} - \frac{\partial a_{\ell j}}{\partial q_k} = 2\sum_{\sigma} \frac{\partial^2 F_\sigma}{\partial q_j \partial x_\ell}\frac{\partial F_\sigma}{\partial q_k}.$$

We have called this quantity the input symbol of the first kind $\hat{\Gamma}_{\ell jk}$ and we define the associated input symbol of the second kind by

$$\hat{\Gamma}^{\mu}_{\ell j} = \sum_{k=1}^{n} m^{\mu k}\hat{\Gamma}_{\ell jk}$$

where $m^{\mu k}$ is the μk–th entry in $\mathcal{M}^{-1}$. Using these definitions, one can show that

$$\frac{\partial}{\partial x_\ell}\begin{pmatrix} \frac{\partial F_1}{\partial q_j} \\ \vdots \\ \frac{\partial F_n}{\partial q_j} \end{pmatrix} = \sum_{k=1}^{n} \hat{\Gamma}^{k}_{\ell j}\begin{pmatrix} \frac{\partial F_1}{\partial q_k} \\ \vdots \\ \frac{\partial F_n}{\partial q_k} \end{pmatrix}.$$

This is a partial differential equation, conditions for the solution of which are covered by Lemma 1. Specifically, for each F_i we are interested in conditions under which there exists a function $g : \mathbb{R}^m \to \mathbb{R}^n$ satisfying

$$\frac{\partial g}{\partial x_\ell} = f_\ell(x, g(x)) \tag{30}$$

where

$$f^{j}_{\ell}(x,z) = \sum_{\gamma=1}^{n} \hat{\Gamma}^{\gamma}_{\ell j} z_\gamma.$$

According to Lemma 1, a necessary condition for the solution of (30) is that

$$\frac{\partial f_\ell}{\partial x_i} - \frac{\partial f_i}{\partial x_\ell} + \sum_{k=1}^{n}\left(\frac{\partial f_\ell}{\partial z_k} f^{k}_{i} - \frac{\partial f_i}{\partial z_k} f^{k}_{\ell}\right) = 0.$$

This is rendered (componentwise)

$$\sum_{\gamma=1}^{n}\left(\frac{\partial \hat{\Gamma}^{\gamma}_{\ell j}}{\partial x_i} - \frac{\partial \hat{\Gamma}^{\gamma}_{ij}}{\partial x_\ell}\right) z_\gamma + \sum_{k\gamma=1}^{n}\left(\hat{\Gamma}^{k}_{\ell j}\hat{\Gamma}^{\gamma}_{ik} z_\gamma - \hat{\Gamma}^{k}_{ij}\hat{\Gamma}^{\gamma}_{\ell k} z_\gamma\right) = 0.$$

Since this must hold identically in $z = (z_1, \ldots, z_n)$, we have the desired conclusion that

$$\frac{\partial \hat{\Gamma}^{\gamma}_{\ell j}}{\partial x_i} - \frac{\partial \hat{\Gamma}^{\gamma}_{ij}}{\partial x_\ell} + \sum_{k=1}^{n}\left(\hat{\Gamma}^{k}_{\ell j}\hat{\Gamma}^{\gamma}_{ik} - \hat{\Gamma}^{k}_{ij}\hat{\Gamma}^{\gamma}_{\ell k}\right) = 0. \ \square$$

Remark 5 In light of the identity noted in Remark 3, it is clear that the vanishing of both the Riemann Symbols of the first kind and input symbols of the first kind is similarly a necessary condition following from the hypothesis of Theorem 1.

Definition 6 If there is a choice of coordinates $\bar{q}$ such that the inertia tensor expressed in $\bar{q}$–coordinates is the identity and the cross coupling terms $\bar{A}(\bar{q}, x)$ vanish, we refer to the $\bar{q}$–coordinates as *flat coordinates.*

Remark 6 The vanishing of all the Riemann symbols (of the first or second kind) may be shown to be sufficient for the existence of an input-dependent change of coordinates such that the inertia tensor $\mathcal{M}(q, x)$ is transformed into the identity matrix. The argument can be constructed along classical lines, and the reader is referred to Spivak [1970] for details.

When the curvature associated with the Levi-Civita connection of the reduced inertia tensor $\mathcal{M}$ (equation (21) does not vanish, it is not presently known whether the vanishing of the input symbols is sufficient to eliminate the appearance of input accelerations in (12).

Remark 7 Sections 4 and 5 will treat the connection between the geometry of second order control systems and recent results on averaging. Many (indeed most) of the hard results to date which describe close connections between the qualitative dynamics of averaged and nonautonomous versions of Hamiltonian and Lagrangian control systems with oscillatory inputs have been established for the case of single input systems. We note that the input connection is trivially flat for single input systems.

To develop physical intuition regarding the vanishing conditions in Theorem 1, we consider some prototypical controlled mechanical systems, beginning with rotating heavy chains.

Example 1 *(Rotating heavy chains)* The global dynamics of a heavy (hanging) chain undergoing forced rotations about a vertical axis will be crucially dependent on how the joints between the links of the chain constrain the possible relative motions of the links. This dependence has been explored in Baillieul and Levi [1991]. Further insight into the global dynamical effects of relative motion constraints may be developed in terms of the curvatures defined above. Consider the two rotating pendulum systems in Figure 1. Both pendula undergo controlled rotations about the vertical axis. In Figure 1(a), the joint by which the pendulum is suspended is a single degree-of-freedom revolute joint, while in Figure 1(b), the link is suspended by a (two degree-of-freedom) universal joint. The respective Lagrangians are

$$L_1(\phi,\dot\phi;\dot\theta) = \frac{m}{2}(a^2\cos^2\phi + \ell^2\sin^2\phi)\dot\theta^2 + (\ell^2 + \frac{a^2}{2})\dot\phi^2 + mg\ell\cos\phi$$

and

$$\begin{aligned} L_2(\phi,\psi,\dot\phi,\dot\psi;\dot\theta) \;=\;& (l^2 + \tfrac{a^2}{2}(1+\sin^2\psi))\,\dot\phi^2 + (\tfrac{a^2}{2}+l^2)\,\dot\psi^2 \\ +\;& (\tfrac{a^2}{2}-l^2)\,\cos\phi\,\sin(2\,\psi)\,\dot\phi\,\dot\theta + (a^2+2\,l^2)\,\sin\phi\,\dot\psi\,\dot\theta \\ +\;& \Big(a^2\cos^2\phi\cos^2\psi \\ & +(\tfrac{a^2}{2}+l^2)\,(\sin^2\phi + \cos\phi\sin^2\psi)\Big)\,\dot\theta^2 \\ +\;& mg\ell\,\cos\phi\,\cos\psi. \end{aligned}$$

The corresponding dynamics of the controlled equations are

$$\begin{aligned} \tfrac{d}{dt}\left[\tfrac{m}{2}(a^2\cos^2\phi + \ell^2\sin^2\phi)\dot\theta\right] &= u \\ (\ell^2 + \tfrac{a^2}{2})\ddot\phi + m\,(a^2-\ell^2)\,\sin\phi\,\cos\phi\,\dot\theta^2 + mg\ell\,\sin\phi &= 0, \end{aligned} \tag{31}$$

and (in abbreviated form)

$$\frac{d}{dt}\frac{\partial L_2}{\partial\dot\theta} = u,$$
$$\frac{d}{dt}\frac{\partial L_2}{\partial\dot\phi} - \frac{\partial L_2}{\partial\phi} = 0, \qquad \frac{d}{dt}\frac{\partial L_2}{\partial\dot\psi} - \frac{\partial L_2}{\partial\psi} = 0. \tag{32}$$

In both cases a torque u is applied to control the angular velocity $\dot\theta$ of rotation about the vertical axis. (See Baillieul [1987] and Marsden and Scheurle [1993] for more information about rotating heavy chains.) Applying our formal reduction procedure, (10) is rendered respectively

$$\mathcal{L}_1(\phi,\dot\phi;\dot\theta) = \frac{m}{2}(\ell^2 + \frac{a^2}{2})\dot\phi^2 - \mathcal{V}_1(\phi;\dot\theta) \tag{33}$$

where

$$\mathcal{V}_1(\phi;\dot\theta) = -(a^2\cos^2\phi + \ell^2\sin^2\phi)\dot\theta^2 - mg\ell\cos\phi,$$

and

$$\mathcal{L}_2(\phi,\psi,\dot\phi,\dot\psi;\dot\theta) = \frac{1}{2}(\dot\phi,\dot\psi)\,\mathcal{M}_2\begin{pmatrix}\dot\phi\\ \dot\psi\end{pmatrix} + \dot\theta\,\mathcal{A}_2\begin{pmatrix}\dot\phi\\ \dot\psi\end{pmatrix} - \mathcal{V}_2(\phi,\psi;\dot\theta) \tag{34}$$

where

$$\mathcal{M}_2 = \mathcal{M}_2(\phi,\psi) = \begin{pmatrix} l^2\cos^2\psi + \frac{a^2\,(1+\sin^2\psi)}{2} & 0 \\ 0 & \frac{a^2}{2}+l^2 \end{pmatrix},$$

$$\mathcal{A}_2 = \mathcal{A}_2(\phi,\psi) = \begin{pmatrix} \left(\frac{a^2}{2}-l^2\right)\cos\phi\,\cos\psi\,\sin\psi \\ \left(\frac{a^2}{2}+l^2\right)\sin\phi \end{pmatrix},$$

and

$$\begin{aligned}\mathcal{V}_2(\phi,\psi;\dot\theta) &= -\tfrac{1}{2}\left(a^2\cos^2\phi\cos^2\psi + \left(\tfrac{a^2}{2}+l^2\right)\left(\sin^2\phi + \cos\phi\sin^2\psi\right)\right)\dot\theta^2 \\ &\quad -mg\ell\cos\phi\,\cos\psi.\end{aligned}$$

The significant difference between the two rotating pendulum systems is the absence or presence of the coupling terms $\mathcal{A}$ in (33) and (34) respectively. The rotating planar pendulum (31) feels the influence of rotation *only* through velocity terms involving $\dot\theta^2$. The dynamics of the universal joint pendulum (32), however, depend on both the angular velocity $\dot\theta$ and angular acceleration $\ddot\theta$. A natural question which arises is whether there is a change of coordinates which makes the dynamics (32) dependent only on $\dot\theta$ and not on the acceleration $\ddot\theta$. Theorem 1 provides a partial answer, but before considering this in detail, we also examine the Riemann symbols of the first kind for the inertia tensor of the unreduced rotating universal joint system:

$$M_2(\phi,\psi) = \begin{pmatrix} \mathcal{N}_2(\phi,\psi) & \mathcal{A}_2(\phi,\psi) \\ \mathcal{A}_2^T(\phi,\psi) & \mathcal{M}_2(\phi,\psi) \end{pmatrix}$$

where $\mathcal{M}_2$ and $\mathcal{A}_2$ are as above, and

$$\mathcal{N}_2(\phi,\psi) = a^2\cos^2\phi\cos^2\psi + \left(\frac{a^2}{2}+l^2\right)\left(\sin^2\phi + \cos^2\phi\,\sin^2\psi\right).$$

It is a straightforward calculation (which is tedious, if you don't use computer algebra) to show that none of the six independent Riemann symbols of the first kind associated with the 3×3 inertia tensor $M_2(\phi,\psi)$ is zero. We can compare this with the results obtained by evaluating corresponding Riemann symbols and inputs symbols for the reduced Lagrangian system defined by $\mathcal{M}_2(\phi,\psi)$ and $\mathcal{A}_2((\phi,\psi)$. For the 2×2 tensor $\mathcal{M}_2(\phi,\psi)$, the curvature tensor is defined by the single Riemann symbol

$$R_{1212} = (\ell^2 - \frac{a^2}{2})\left(\cos 2\psi - \frac{(\ell^2-\frac{a^2}{2})\sin^2 2\psi}{8(\ell^2\cos^2\psi + \frac{a^2}{2}[1+\sin^2\psi])}\right)$$

The input connection is nonzero in that

$$\hat\Gamma_{112} = (\ell^2 + \frac{a^2}{2})\cos\phi - (\frac{a^2}{2}-\ell^2)\cos\phi(\cos^2\psi - \sin^2\psi).$$

It is obvious from the definition, however, that the input symbols for any single input system, such as this, are going to be zero.

The answer to the question of whether there is a change of coordinates which eliminates the dependence of the equations of motion on input acceleration terms is not completely resolved by these calculations. The nonvanishing of the curvatures associated with both the 3×3 inertia tensor of the unreduced system and the 2×2 inertia tensor of the reduced system imply that there is no change of coordinates such that the dynamical dependence on input accelerations is eliminated and the dynamics of the configuration variables ϕ and ψ are also decoupled. Whether one can simply eliminate the input accelerations by a choice of coordinates in which there is inertial coupling of the ϕ and ψ dynamics remains for the moment an open question.

Example 2 *Controlled planar body with pendulum attachment* Here we consider a pair of rigid bodies which are connected by a simple frictionless single degree-of-freedom hinge as illustrated in Figure 2(a). One of the bodies (the larger one depicted in the figure) is assumed to have actuators allowing its motion in the plane to be controlled to follow any prescribed smooth path. No actuation is applied directly to the pendulum, and hence it moves entirely under the influence of gravity and motion of the controlled body. Take as generalized coordinates for the system x (the horizontal displacement of the large body), y (the vertical displacement of the large body), and θ (the angular displacement of the pendulum from the vertical, downward pointing configuration). The 3×3 inertia matrix M in this case is given by

$$M = \begin{pmatrix} m_a + m_b & 0 & m_b \ell \cos\theta \\ 0 & m_a + m_b & m_b \ell \sin\theta \\ m_b \ell \cos\theta & m_b \ell \sin\theta & I_b \end{pmatrix}.$$

The sub-blocks of the inertia matrix M which are of interest are $\mathcal{A} = \mathcal{A}(\theta) = (m_b \ell \cos\theta, m_b \ell \sin\theta)$ and $\mathcal{M} = I_b$, where m_a is the mass of the controlled body, m_b is the mass of the pendulum, I_b is the pendulum inertia about the hinge point, and ℓ is the distance from the hinge point to the center of mass of the pendulum. Thus in the reduced Lagrangian control system, x and y are the inputs, while θ is the configuration variable whose motion is governed by the reduced Lagrangian dynamics.

The Christoffel symbols of the first kind associated with the Levi-Civita connection of the unreduced inertia tensor M are all zero except for $\Gamma_{331} = -m_b \ell \sin\theta$ and $\Gamma_{332} = m_b \ell \cos\theta$. Nevertheless, the Riemann curvature tensor (21) vanishes, and from this we conclude that there is a choice of coordinates with respect to which the inertia tensor takes the form of the identity matrix. It is also easy to verify that the vanishing conditions of Theorem 1 are satisfied. Indeed, the reduced inertia tensor, $\mathcal{M} = I_b$, is a scalar, and hence the Riemannian curvature is trivially zero. The input connection $\hat{\Gamma}_{1ij}$, $i, j = 1, 2$ is zero, and hence the input curvature is also zero.

Example 3 *Controlled planar body coupled with extending pendulum attachment* The concluding example treats another system for which the curvature vanishing conditions of Theorem 1 are not satisfied. The mechanism depicted in Figure 2(b) is similar to the previous example, with the significant difference being that we assume there is an internal mechanism which causes the length of the pendulum attachment to depend on the x-coordinate of the body. Specifically, we assume the length of the pendulum is αx, where $\alpha > 0$ is some fixed constant. We assume again that the pendulum is attached by a frictionless hinge. To simplify the discussion (and with no loss of generality) we idealize the model so that the pendulum is comprised of a point mass m_b and a massless linkage between the point mass and hinge. The inertia matrix M in this case is given by

$$\begin{pmatrix} m_a + m_b + m_b\alpha^2 + 2m_b\alpha\sin\theta & -m_b\alpha\cos\theta & m_b\alpha x\cos\theta \\ -m_b\alpha\cos\theta & m_a + m_b & m_b\alpha x\sin\theta \\ m_b\alpha x\cos\theta & m_b\alpha x\sin\theta & m_b\alpha^2 x^2 \end{pmatrix}.$$

While this example is admittedly somewhat artificial, it serves to illustrate some of the features of the flatness properties we have discussed. As in the previous example, the inputs here are taken to be the x and y positions of the controlled body. We analyze "nonintegrability" in the dynamical relationship between the input variables and the configuration θ of the attachment in three different ways. First, in computing the components of the Riemann curvature tensor (21), we find that all six independent components are zero except

$$R_{1313} = \frac{N(\theta)}{8\,((1-2\,\alpha)\,\cos^2\theta\,\sin\theta\, m_b{}^2 + (1+\alpha^2+\sin\theta)\, m_b\, m_a + m_a{}^2)}$$

where

$$\begin{aligned} N(\theta) \quad = \quad & -\Big((-1+2\,\alpha)\, m_b \Big(-2\,(-1+2\,\alpha)\,\cos^4\theta\, m_b{}^2 - \\ & (-3+\cos(2\,\theta) - 4\,\sin\theta - 4\,\alpha^2\,\sin\theta)\, m_b\, m_a \\ & + 4\,\sin\theta\, m_a{}^2\Big)\Big). \end{aligned}$$

This means of course there is no change of (x,y,θ)–coordinates such that the unreduced inertia tensor M has the form of the identity matrix. Secondly, we analyze the reduced Lagrangian in light of Theorem 1. Since in this case the reduced inertia tensor $\mathcal{M}(\theta;x,y) = m_b\alpha^2x^2$ is of dimension 1×1, we shall only need to analyze the input connection to verify the hypotheses of the theorem. The input connection has two components: $\hat{\Gamma}^1_{11} = \frac{2}{x}$ and $\hat{\Gamma}^1_{21} = 0$. That $\hat{\Gamma}^1_{11}$ is unbounded as $x \to 0$ reflects the fact that the inertia tensor is singular in the $x = 0$ limit. Although the input connection is nonzero, the vanishing of all the relevant curvatures shows that it is flat in the dense of Definition 4. Thus the necessary conditions of the theorem are satisfied. Since we know there is no choice of coordinates in terms of which M is rendered as the identity matrix, the necessary conditions of Theorem 1 thus fail to be sufficient. The theory at present does not answer the question of whether or not the system can be rewritten in such a way that the dynamics do not depend on accelerations of the inputs.

There is a third and final sense in which the influence of the inputs (x,y) in this example is nonintegrable, and this will be discussed in the following section.

4 Averaging Lagrangian and Hamiltonian Systems with Oscillatory Inputs

For systems such as Example 1(a) where we may express the Lagrangian as the sum of a simple kinetic energy term plus a time-dependent potential, it is natural to conjecture that given small-amplitude high-frequency periodic inputs $x(\cdot)$, their influence is approximately equal to the effect of a potential force determined by the time-averaged potential. There is a growing body of theory concerning the conditions under which this conjecture is true. (See Weibel [1997] and Weibel and Baillieul [1998] for details.) The role of *flatness* and *flat inputs* in determining the dynamic response of (11) undergoing high frequency oscillatory forcing is not completely understood, but it is clear that the analysis can be simplified considerably for systems which are flat in the sense described. One specific simplification which occurs for the case of flat systems with flat inputs is that the averaged potential may be computed (with respect to flat coordinates) by taking simple averages of the amended potential $\mathcal{V}(q;x,v)$ in (10). For systems for which we cannot find flat coordinates, there seems to be additional complexity and additional steps in the analysis. One approach which works (at least formally) in the general case is to move to the Hamiltonian setting. Here it is always possible to write down the *averaged potential.*

As a general starting point, we return to the Lagrangian control system (10) which we rewrite here:

$$\mathcal{L}(q,\dot{q};x,v) = \frac{1}{2}\dot{q}^T\mathcal{M}(q,x)\dot{q} + v^T\mathcal{A}(q,x)\dot{q} - \mathcal{V}(q;x,v).$$

To transform to a Hamiltonian formulation, we perform a Legendre transformation, writing $\mathcal{H} = p^T\dot{q} - \mathcal{L}$ and expressing $\mathcal{H}$ in terms of (q,p). This $\mathcal{H}$ is a noncanonical Hamiltonian which takes the form

$$\mathcal{H}(q,p;x,v) = \frac{1}{2}(p-\mathcal{A}^Tv)^T\mathcal{M}^{-1}(p-\mathcal{A}^Tv) + \mathcal{V} \tag{35}$$

We expand equation (35) and apply simple averaging to yield the *averaged Hamiltonian*:

$$\overline{\mathcal{H}}(q,p) = \frac{1}{2}p^T\overline{\mathcal{M}^{-1}}p - \overline{v^T\mathcal{A}\mathcal{M}^{-1}}p + \frac{1}{2}\overline{v^T\mathcal{A}\mathcal{M}^{-1}\mathcal{A}^Tv} + \overline{\mathcal{V}}. \tag{36}$$

Here the overbars indicate simple averages over one period (T) have been taken: given any piecewise continuous function $F(q,p,t)$, the simple average is given by $\bar{F} = \int_0^T F(q,p,t)\,dt$, where for the purpose of evaluating this integral q and p are regarded as constants.

For $\bar{\mathcal{H}}$, there is an obvious decomposition into kinetic and potential energy terms in the case that $\overline{v^T\mathcal{A}\mathcal{M}^{-1}} = 0$:

$$\overline{\mathcal{H}}(q,p) = \underbrace{\frac{1}{2}p^T\overline{\mathcal{M}^{-1}}p}_{\text{avg. kin. energy}} + \underbrace{\frac{1}{2}\overline{v^T\mathcal{A}\mathcal{M}^{-1}\mathcal{A}^Tv} + \overline{\mathcal{V}}}_{\text{averaged potential}}. \tag{37}$$

There are a number of conditions under which $\overline{v^T\mathcal{A}\mathcal{M}^{-1}} = 0$. This will trivially be the case when we are able to find a choice of coordinates with respect to which the $\mathcal{A}$ terms do not appear in (10). In Weibel and Baillieul [1998] it was shown that for zero-mean input $v(\cdot)$ exactness of the 1-form $\mathcal{M}^{-1}(q,x)\mathcal{A}(q,x)\,dx$ is also sufficient to guarantee that $\overline{v^T\mathcal{A}\mathcal{M}^{-1}} = 0$. We remark that this exactness condition is satisfied for the systems in Examples 1 and 2 but *not* Example 3 in the previous section.

In the case that $\overline{v^T\mathcal{A}\mathcal{M}^{-1}} \neq 0$, it remains possible to decompose the averaged Hamiltonian (36) into the sum of averaged kinetic and potential energies, although the description becomes more involved. If $\bar{v} \neq 0$, then the corresponding input variable $x(t)$ will not be periodic. There will be a "drift" in the value of $x(t)$ which changes by an amount $\bar{v}\cdot T$ every T-units of time. We may nevertheless study averaged Hamiltonian systems in this context. We rewrite the averaged Hamiltonian (36) as

$$\begin{aligned}
\overline{\mathcal{H}}(q,p) &= \frac{1}{2}p^T\overline{\mathcal{M}^{-1}}p - \overline{v^T\mathcal{A}\mathcal{M}^{-1}}p + \frac{1}{2}\overline{v^T\mathcal{A}\mathcal{M}^{-1}}\left(\overline{\mathcal{M}^{-1}}\right)^{-1}\overline{\mathcal{M}^{-1}\mathcal{A}^Tv} \\
&\quad + \frac{1}{2}\overline{v^T\mathcal{A}\mathcal{M}^{-1}\mathcal{A}v} - \frac{1}{2}\overline{v^T\mathcal{A}\mathcal{M}^{-1}}\left(\overline{\mathcal{M}^{-1}}\right)^{-1}\overline{\mathcal{M}^{-1}\mathcal{A}^Tv} + \overline{\mathcal{V}} \\
&= \underbrace{\frac{1}{2}(\overline{\mathcal{M}^{-1}}p - \overline{\mathcal{M}^{-1}\mathcal{A}^Tv})^T\left(\overline{\mathcal{M}^{-1}}\right)^{-1}(\overline{\mathcal{M}^{-1}}p - \overline{\mathcal{M}^{-1}\mathcal{A}^Tv})}_{\text{averaged kinetic energy}} \\
&\quad \underbrace{+\frac{1}{2}\overline{v^T\mathcal{A}\mathcal{M}^{-1}\mathcal{A}^Tv} - \frac{1}{2}\overline{v^T\mathcal{A}\mathcal{M}^{-1}}\left(\overline{\mathcal{M}^{-1}}\right)^{-1}\overline{\mathcal{M}^{-1}\mathcal{A}^Tv} + \overline{\mathcal{V}}}_{\text{averaged potential}}.
\end{aligned} \tag{38}$$

The formal distinction which appears between the zero-mean and non-zero-mean cases ((38) and (39) respectively) is that in the latter case both the averaged kinetic and potential energy terms are adjusted to reflect the net (average) motions of the input variables $(x(\cdot), v(\cdot))$. There are important relationships which can be established between the dynamics associated with the averaged Hamiltonian (39) and the dynamics of the periodically forced system. The reader is referred to Weibel, 1997, Weibel *et al.*, 1997, and Weibel & Baillieul, 1998 for details. It is important to mention that for mechanical systems in which $\bar{v} \neq 0$ and $\mathcal{M}$ depends explicitly on x in (35), the averaging analysis of this chapter may not provide an adequate description of the dynamics. Indeed, in this case, $\|x(t)\|$ will not remain bounded as $t \to \infty$, and if $\mathcal{M}(x(t), q_2)$ also fails to remain bounded, the averaged potential will inherit a dependence on time which will make it difficult to apply the critical point analysis proposed below. Despite this cautionary remark, we shall indicate how our methods may be applied in many instances where $\bar{v} \neq 0$.

5 Stability and Flatness in Mechanical Systems with Oscillatory Inputs

The Lagrangian (10) gives rise to the Lagrangian dynamics (11), which we rewrite here as

$$\sum_{j=1}^{n} m_{kj}\ddot{q}_j + \sum_{\ell=1}^{m} a_{\ell k}\dot{v}_\ell + \sum_{i,j=1}^{n}\Gamma_{ijk}\dot{q}_i\dot{q}_j + \sum_{j=1}^{n}\sum_{\ell=1}^{m}\hat{\Gamma}_{\ell jk}v_\ell\dot{q}_j = F(t), (k=1,\ldots,n), \tag{39}$$

where Γ_{ijk} and $\hat{\Gamma}_{\ell jk}$ are defined as in Section 2 in terms of the entries a_{ij} and m_{ij} in the $m \times n$ and $n \times n$ matrices $\mathcal{A}(q,x)$ and $\mathcal{M}(q,x)$ respectively. We seek to understand the stability of responses to oscillatory forcing of (11) in terms of the corresponding *averaged potential* which was introduced in the previous section.

$$\mathcal{V}_A(q) = \frac{1}{2}\overline{v^T\mathcal{A}\mathcal{M}^{-1}\mathcal{A}^Tv} - \frac{1}{2}\overline{v^T\mathcal{A}\mathcal{M}^{-1}}\left(\overline{\mathcal{M}^{-1}}\right)^{-1}\overline{\mathcal{M}^{-1}\mathcal{A}^Tv} + \overline{\mathcal{V}}. \tag{40}$$

An answer to this question is provided by the *averaging principle* for Lagrangian systems: **Averaging Principle for Periodically Forced Lagrangian Systems:** The dynamics of the periodically forced system (11) are locally determined in neighborhoods of critical points of the *averaged potential* $\mathcal{V}_A(q)$ as follows:

1. If q^* is a strict local minimum of $\mathcal{V}_A(\cdot)$, then provided the frequency of the periodic forcing $v(\cdot)$ is sufficiently high, (11) will execute motions confined to a neighborhood of q^*.
2. If $(q,p) = (q^*,0)$ is a hyperbolic fixed point of the averaged system (11), then there is a corresponding periodic orbit of (11) such that the asymptotic stability properties of the fixed point $(q^*,0)$ of the properly averaged version of (11) coincide with the asymptotic stability properties of the periodic orbit with respect to (11).

□

Roughly but simply stated, the averaging principle for forced Lagrangian systems says that for periodic forcing functions $v(\cdot)$ of sufficiently high frequency, the dynamics of the averaged system provide a good local approximation of the dynamics of the nonautonomous (forced) system. Unfortunately, this *averaging principle* is less than a theorem, and it has been rigorously established only in special cases. We highlight several of these in the present section and refer to Baillieul [1995] and Weibel and Baillieul [1998] for further details. The main idea, of course, is that for sufficiently high frequencies the effect of forcing (11) with an oscillatory input $v(\cdot)$ will be to produce stable motions confined to neighborhoods of relative minima of $\mathcal{V}_A(\cdot)$. A naive approach to stability analysis is to linearize the dynamics (11) about such relative minima. Unfortunately, it is only in special cases that such linearizations captures the stabilizing effects implied by an analysis of the averaged potential. Indeed, it turns out that in general, the averaged potential depends on second order jets of the coefficient functions $\mathcal{A}(q)$ and $\mathcal{M}(q)$. (We continue to assume that $\mathcal{M}$ and $\mathcal{A}$ depend on the input x in general, but we suppress this variable to simplify notation in the present section.) To further simplify the presentation, we shall restrict our attention to the case of zero-mean oscillatory forcing in which $\overline{\mathcal{M}^{-1}\mathcal{A}^Tv} = 0$.

Suppose q_0 is a strict local minimum of (40). Applying a high-frequency oscillatory input $v(\cdot)$, we shall look for stable motions of (11) in neighborhoods of $(q,\dot{q}) = (q_0,0)$. Of course, even when there are such stable motions, $(q_0,0)$ need not be a rest point of (11) for any choice of forcing function $v(\cdot)$. (Cf. the "cart-pendulum" dynamics treated in Weibel *et al.* [1997].)

To analyze the relationship between (11) and (40) in more detail, let us assume $q_0 = 0$. This assumption is made without loss of generality, since we may always change coordinates to make it true. Write

$$\mathcal{A}(q) = \mathcal{A}_0 + \mathcal{A}_1(q) + \mathcal{A}_2(q) + \text{h.o.t.};$$
$$\mathcal{M}(q) = \mathcal{M}_0 + \mathcal{M}_1(q) + \mathcal{M}_2(q) + \text{h.o.t.},$$

where the entries in the $n \times n$ matrix $\mathcal{M}_k(q)$ are homogeneous polynomials of degree k in the components of the vector q, and similarly for the $m \times n$ matrix $\mathcal{A}_k(q)$. The averaged potential has a similar expansion, which we write:

$$\mathcal{V}_A(q) = \overline{\mathcal{V}}_0 + \overline{\mathcal{V}}_1(q) + \overline{\mathcal{V}}_2(q) + \text{h.o.t.},$$

All terms in all these expressions may depend on the input x, but we do not explicitly display this dependence. We note, however, that each term in the expansion of $\mathcal{V}_A$ is the sum of a term of the same order in the original potential plus an algebraic function of entries in terms up to the given order in the expansions of $\mathcal{A}$ and $\mathcal{M}$.

There are two essentially different cases which can be described in terms of these expansions. The case that $\mathcal{A}_0 \neq 0$ is qualitatively quite different from the case in which $\mathcal{A}_0 = 0$. It is associated with qualitative dynamical behavior we have referred to as *hovering motions.* These are motions confined to neighborhoods of critical points of the averaged potential which are not equilibrium points of the corresponding forced or nonautonomous system. We will not treat this case further, and refer to Weibel and Baillieul [1998] for further details. We suppose in the remainder of the section that $\mathcal{A}_0 = 0$, in which case 0 is both a critical point of the averaged potential and a rest point of the forced system (11). We'll show that in this case the averaged potential depends only on first order jets of the coefficients of (11) when $\mathcal{A}(q_0) = 0$. In Weibel and Baillieul [1998] it has also been shown that the condition $\mathcal{A}(q_0) = 0$ is also necessary and sufficient for the local minimum q_0 of the averaged potential to define a corresponding fixed point (rather than a periodic orbit) of the *forced* (nonautonomous) Hamiltonian system associated with (10).

Slightly refining our notation, let $\mathcal{A}^\ell(q)$ denote the ℓ-th column of the $n\times m$ matrix $\mathcal{A}^T(q)$. Then we have

$$\mathcal{A}^\ell(q) = \mathcal{A}_1^\ell \cdot q + (\text{terms of order } \geq 2), \text{ and}$$
$$\mathcal{M}(q) = \mathcal{M}_0 + (\text{terms of order } \geq 1),$$

where we interpret $\mathcal{M}_0, \mathcal{A}_1^1, \ldots, \mathcal{A}_1^m$ as $n \times n$ coefficient matrices. Using the same notation, we also have an expansion of the (original) potential function:

$$\mathcal{V}(q) = \frac{1}{2} q^T \mathcal{V}_2 q + \text{h.o.t.}$$

We recall the following result.

Proposition 3 *Suppose $v(\cdot)$ is an $\mathbb{R}^m$-valued piecewise continuous periodic function of period $T > 0$ such that $\bar{v} = \frac{1}{T}\int_0^T v(s)\,ds = 0$. Suppose, moreover, that $\mathcal{A}_0 = 0$. Then the averaged potential of the Lagrangian system (11) agrees up to terms of order 2 with the averaged potential associated with the linear Lagrangian system*

$$\mathcal{M}_0 \ddot{q} + \sum_{\ell=1}^{m} \left(\dot{v}_\ell \mathcal{A}_1^\ell q + v_\ell (\mathcal{A}_1^\ell - {\mathcal{A}_1^\ell}^T) \dot{q} \right) + \mathcal{V}_2 \cdot q = 0. \tag{41}$$

Proof: The proof follows from examining a series expansion of $\mathcal{V}_A$ about the point $q = 0$. We refer the reader to Weibel and Baillieul [1998] for details.□

There are several points worth noting here. First, (41) is exactly the usual (tangential) linearization of (11) with respect to the state variables $(q, \dot{q})$ about the rest state $(q, \dot{q}) = (0, 0)$. Also, if contrary to the hypothesis $\mathcal{A}_0 \neq 0$, then such a linearization does not give a valid local approximation to the dynamics of (11), and the Hessian of the averaged potential (40) involves terms of higher order in both $\mathcal{A}$ and $\mathcal{M}$.

A deeper connection with stability is now expressed in terms of the following theorem.

Theorem 2 *Suppose $w(\cdot)$ is an $\mathbb{R}^m$-valued piecewise continuous periodic function of period $T > 0$ such that $\bar{w} = \frac{1}{T}\int_0^T w(s)\,ds = 0$. Consider the linear Lagrangian system (41) with input $v(t) = w(\omega t)$, and suppose ${\mathcal{A}_1^\ell}^T = \mathcal{A}_1^\ell$ for $\ell = 1, \ldots, m$. The averaged potential for this system is given by*

$$\mathcal{V}_A(q) = \frac{1}{2} q^T \left(\mathcal{V}_2 + \sum_{i,j=1}^{m} \sigma_{ij} \mathcal{A}_1^i \mathcal{M}_0^{-1} \mathcal{A}_1^j \right) q, \tag{42}$$

where $\sigma_{ij} = (1/T)\int_0^T w_i(s) w_j(s)\,ds$. If the matrix $\frac{\partial^2 \mathcal{V}_A}{\partial q^2}$ is positive definite, the origin $(q, \dot{q}) = (0, 0)$ of the phase space is stable in the sense of Lyapunov provided ω is sufficiently large.

□

This theorem has been proved in Baillieul [1995]. Its proof apparently requires the assumption that the coefficient matrices A_1^ℓ are symmetric. Suppose $\mathcal{M}(q,x) = \mathcal{M}(q)$ does not depend on x. Then

$$\hat{\Gamma}\ell jk = \frac{\partial a_{\ell k}}{\partial q_j} - \frac{\partial a_{\ell j}}{\partial q_k},$$

and $\hat{\Gamma}_{\ell jk}(0)$ is the jk–th element in the skew symmetric matrix $A_1^{\ell T} - A_1^\ell$. Thus, in this case, the hypothesis that A_1^ℓ is symmetric for $\ell = 1, \ldots, m$ is equivalent to assuming the input connection vanishes at the rest point. If A_1^ℓ is not symmetric, the vanishing of the input symbols $\hat{R}^\gamma_j{}_{\ell i}$ is equivalent to the set of matrices

$$\{(A_1^{\ell T} - A_1^\ell)M_0^{-1} : \ell = 1, \ldots, m\}$$

being commutative. These observations are interesting, because it is precisely in cases where flatness is present that we are able to prove the stability of motions based on an analysis of the averaged potential.

6 Concluding Remarks

Like many of the chapters in this volume, the present one treats the interface between nonlinear control theory and mechanics. Second order control systems are obviously central in this regard, and the intrinsic geometric approach we have presented builds naturally on earlier work of Roger Brockett. We have shown that in using the language of affine connections, it is possible to state conditions involving the vanishing of curvature-like quantities that must be satisfied in order for *acceleration controlled* Lagrangian systems to be rewritten (more simply) as *velocity controlled* Lagrangian systems. These geometric conditions also bear on the simplicity of computing the *averaged potential* and assessing the stability characteristics of such systems under the influence of high-frequency oscillatory inputs. It is of interest to note that the vanishing of the input connection occurs precisely in the case that we have been able to prove Lyapunov stability of forced "linear Lagrangian" systems. Since our vanishing conditions are necessary for the elimination of certain Coriolis terms in the unreduced Lagrangian, they may also bear on questions of dissipation induced instability which have been discussed by Bloch *et al.* [1994].

References

[1] R. ABRAHAM & J.E. MARSDEN, 1978. *Foundations of Mechanics.*

[2] V.I. ARNOLD, 1989. *Mathematical Methods in Classical Mechanics*, Second Edition,Graduate Texts in Mathematics, 60, Springer-Verlag, New York.

[3] J. BAILLIEUL, 1987. "Equilibrium Mechanics of Rotating Systems," *Proc. 26th IEEE Conf. Dec. and Control*, Los Angeles, CA, 1429-34.

[4] J. BAILLIEUL, 1990. "The Behavior of Super-articulated Mechanisms Subject to Periodic Forcing," in *Analysis of Controlled Dynamical Systems*, Proceedings of a Conference held in Lyon, France 3-6 Juillet, 1990, Gauthier, Bride, Bonnard, Kupka, Eds., Birkhauser.

[5] J. BAILLIEUL & M. LEVI, 1991. "Constrained Relative Motions in Rotational Mechanics," *Archive for Rational Mechanics and Analysis*, No. 115, pp. 101-135.

[6] J. BAILLIEUL, 1993. "Stable Average Motions of Mechanical Systems Subject to Periodic Forcing," *Dynamics and Control of Mechanical Systems: The falling cat and related problems*, Fields Institute Communications, **1**, Michael Enos, Ed., American Mathematical Society, Providence, pp. 1-23.

[7] J. BAILLIEUL, 1995. "Energy Methods for Stability of Bilinear Systems with Oscillatory Inputs," *Int'l J. of Robust and Nonlinear Control*, Special Issue on the Control of Mechanical Systems," H. Nijmeijer & A.J. van der Schaft, Guest Eds., Vol. 5, pp. 205-381.

[8] N.S. BEDROSIAN, "Linearizing Coordinate Transformations and Euclidean Systems," *Proc. 1992 IEEE Int'l Conference on Robotics and Automation*, Nice, France.

[9] A.M. BLOCH,, P.S. KRISHNAPRASAD, J.E. MARSDEN, & T.S. RATIU, 1994. "Dissipation Induced Instabilities," Ann. Inst. H. Poincaré, Analyse Nonlineare, Vol. 11, 37–90.

[10] R.W. BROCKETT, 1972. "System Theory on Group Manifolds and Coset Spaces," *SIAM J. Control*, v. 10, pp. 265-284.

[11] R.W. BROCKETT, 1973. "Lie Theory and Control Systems Defined on Spheres," *SIAM J. Appl. Math.*, v. 25:2, pp. 213-225.

[12] R.W. BROCKETT, 1976. "Control Theory and Analytical Mechanics," appearing *Geometric Control Theory*, C.F. Martin and R. Hermann, eds., Vol. VII of *Lie Groups: History, Frontiers, and Applications*, Math Sci Press, Brookline, pp. 1-46.

[13] B.A. DUBROVIN, A.T. FOMENKO, & S.P. NOVIKOV, 1984. *Modern Geometry—Methods and Applications. Part I*, Graduate Texts in Mathematics, Vol. 93, Springer-Verlag, New York.

[14] M. FLIESS, 1990. "Generalized Controller Canonical Forms for Linear and Nonlinear Dynamics," *IEEE Trans. on Automatic Control*, **AC 35-9**, September, pp. 994-1001.

[15] M. FLIESS, J. LÉVINE, P. MARTIN, & P. ROUCHON, 1998. "A Lie-Bäcklund Approach to Equivalence and Flatness of Nonlinear Systems," To appear in the *IEEE Transactions on Automatic Control.*

[16] G.W. HAYNES AND H. HERMES, 1970. "Nonlinear Controllability via Lie Theory," *SIAM J. Control*, v. 8, pp. 450-460.

[17] H. HERMES AND G.W. HAYNES, 1963. "On the Nonlinear Control Problem with Control Appearing Linearly," *SIAM J. Control*, . 1, pp. 85-108.

[18] D. HUSEMOLLER, 1994. *Fibre Bundles, 3-rd Ed.*, Springer-Verlag, Graduate texts in mathematics, New York.

[19] S. KOBAYASHI & K. NOMIZU, 1963. *Foundations of Differential Geometry*, Interscience Publishers (Wiley), New York.

[20] A.J. KRENER, 1974. "A Generalization of Chow's Theorem and the Bang-Bang Theorem in Nonlinear Control Problems," *SIAM J. Control*, v. 12:1.

[21] J.E. MARSDEN & T.S. RATIU, *Introduction to Mechanics and Symmetry*, Texts in Applied Mathematics (TAM) 17, Springer-Verlag, New York, 1994.

[22] J.E. MARSDEN, 1992. *Lectures on Mechanics*, London Mathematical Society Lecture Note Series, **174**, Cambridge University Press.

[23] M.J. VAN NIEUWSTADT & R.M. MURRAY, 1998. "Real Time Trajectory Generation for Differentially Flat Systems," CDS Technical Report 96-017 To appear in *Int'l. J. Robust & Nonlinear Control.*

[24] H. NIJMEIJER & A.J. VAN DER SCHAFT, 1990. *Nonlinear Dynamical Control Systems*, Springer-Verlag, New York.

[25] M. SPIVAK, 1970. *A Comprehensive Introduction to Differential Geometry*, vols. 1,2, Publish or Perish, Inc., Boston, MA.

[26] M.W. SPONG, 1998. "On Feedback Linearization of Robot Manipulators and Riemannian Curvature," in *Essays on Mathematical Robotics* Volume 104, IMA Volumes in Mathematics and its Applications, Edited by J. Baillieul, S.S. Sastry, and H.J. Sussmann, Springer-Verlag, New York.

[27] N.E. STEENROD, 1951. *The Topology of Fibre Bundles*, Princeton University Press, Princeton mathematical series, 14, Princeton, NJ.

[28] H.J. SUSSMANN & V. JURDJEVIC, 1972. "Controllability of Nonlinear Systems," *J. Diff. Equs.*, v. 12.

[29] S.P. WEIBEL, 1997. *Applications of Qualitative Methods in the Nonlinear Control of Superarticulated Mechanical Systems*, Boston University Ph.D. Thesis, Department of Aerospace/Mechanical Engineering.

[30] S. WEIBEL & J. BAILLIEUL, 1997. "Oscillatory Control of Bifurcations in Rotating Chains," Proc. of the 1997 American Control Conference, Albuquerque, NM, pp. 2313-2317.

[31] S. WEIBEL, T. KAPER, & J. BAILLIEUL, 1997. "Global Dynamics of a Rapidly Forced Cart and Pendulum," *Nonlinear Dynamics*, **13**: 131-170, July, 1997.

[32] S. WEIBEL, J. BAILLIEUL, & B. LEHMAN, 1997. " Equilibria and Stability of an n-Pendulum Forced by Rapid Oscillations," 1997 *IEEE Conf. on Decision & Control*, San Diego, CA, pp. 1147-1152.

[33] D. SETO & J. BAILLIEUL, 1994. "Control Problems in Super-Articulated Mechanical Systems," *IEEE Transactions on Automatic Control*, Vol 39-12, December, pp. 2442-2453.

[34] S. WEIBEL & J. BAILLIEUL, 1998. "Averaging and Energy Methods for Robust Open-loop Control of Mechanical Systems," in *Essays on Mathematical Robotics* Volume 104, IMA Volumes in Mathematics and its Applications, Edited by J. Baillieul, S.S. Sastry, and H.J. Sussmann, Springer-Verlag, New York.

[35] J.A. WOLF, 1972. *Spaces of Constant Curvature*, 2-nd ed., Publish or Perish, Inc., Boston, MA.

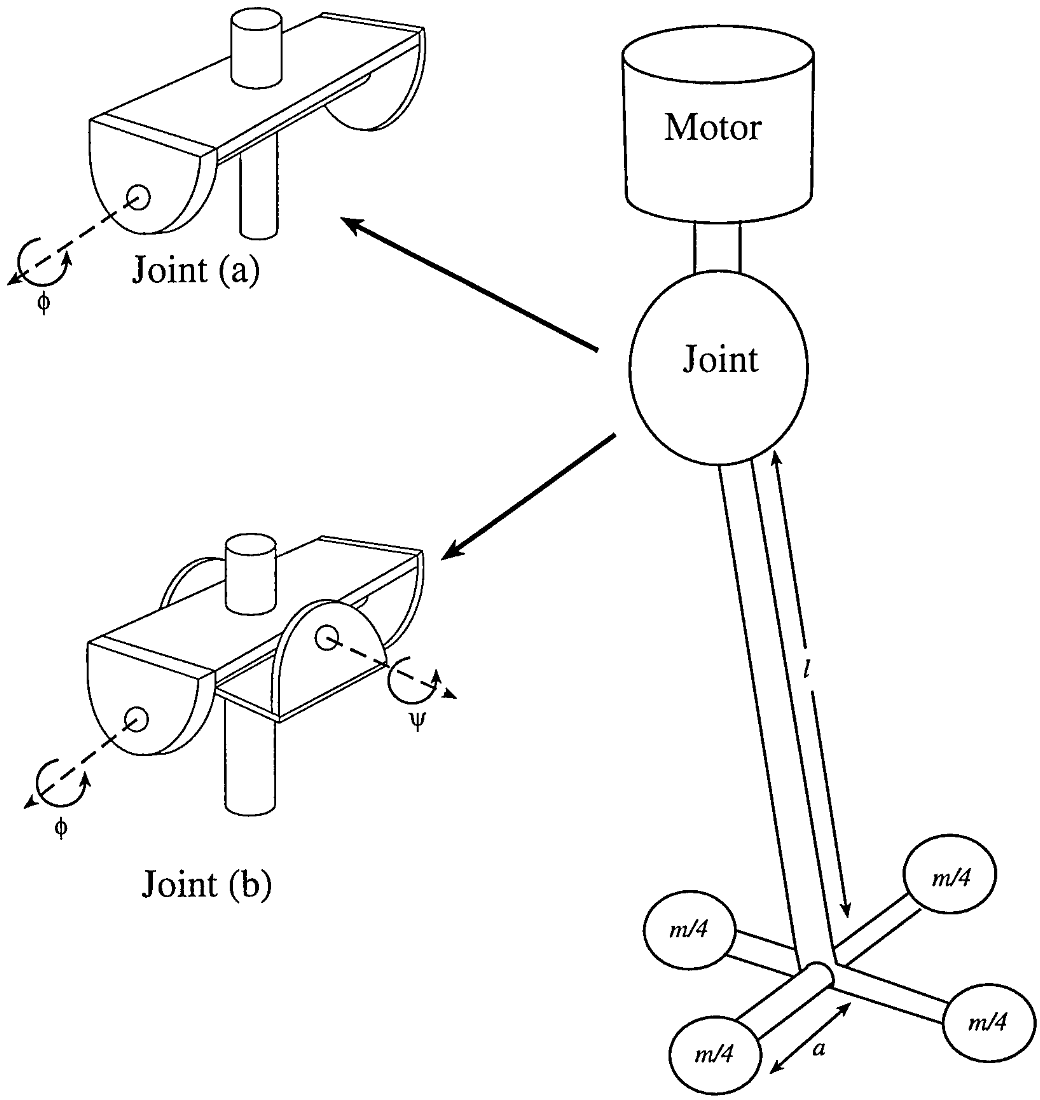

Figure 1: (a) Rotating planar pendulum joint; (b) Rotating universal pendulum joint. We consider a simple rotating pendulum suspended by either type of joint as depicted. The pendulum consists of a link of length ℓ, which we assume to be massless and at the end of which are four equal masses distributed as shown so that the moment of inertia of the pendulum is nonsingular.

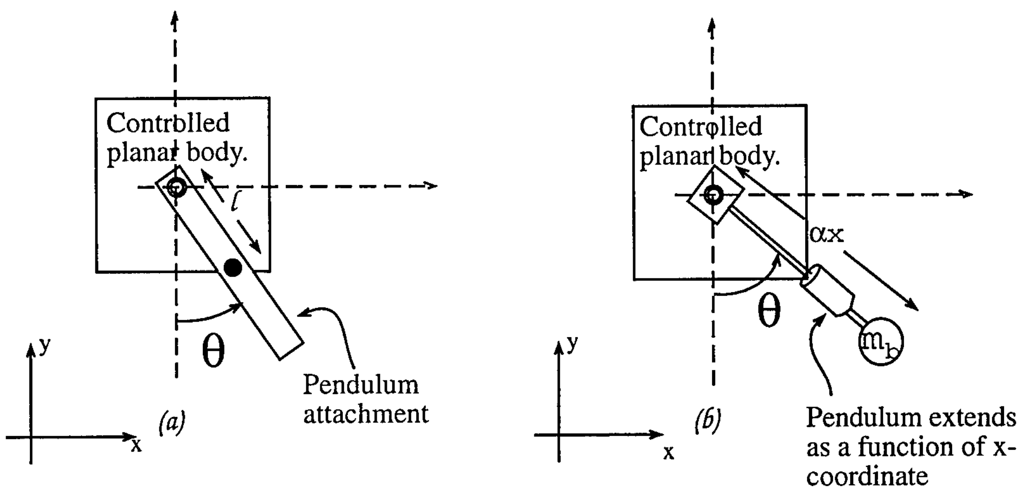

Figure 2: (a) Controlled planar body with a pendulum attachment; (b) Controlled planar body with an *extending* pendulum attachment. The length of the extensible pendulum is αx, where x is the x–coordinate of the hinge point connecting the two bodies.

Open-loop Oscillatory Stabilization of an n-Pendulum

S. Weibel
Dept. of Elec/Comp Engineering
Northeastern University
Boston, MA 02115

J. Baillieul*
Dept. of Aero/Mech Engineering
Boston University
Boston, MA 02215

Abstract

This report presents new results on the behavior of an n–link pendulum which is controlled by means of an oscillatory input forcing one end link. Drawing on recent research on the oscillatory control of second order nonlinear systems, we adopt the *averaged potential* as the primary tool in our equilibrium and stability analysis. While such control of single-degree-of-freedom systems has been studied in fair detail[15], researchers are only beginning to explore the issues which arise in the control of multi-degree-of-freedom systems. Of specific interest is the behavior when the system possesses bifurcations and more than one stable equilibrium. In this report, we present a study of the bifurcations and stability of a periodically-forced cart and n-pendulum on an inclined plane. Periodic open-loop forcing of the single degree-of-freedom pendulum has previously been shown to be a robust and effective means of stabilizing the inverted equilibrium, and this robustness carries over to the multi-degree-of-freedom case provided attention is paid to matching the forcing amplitude to the characteristic length scales of the system. In addition, we present numerical results on the basins of attraction of the vertically forced double pendulum. We present an exact model for the n-degree-of-freedom system, nondimensionalize the model, and compute the averaged potential. Equilibria and their stability characteristics are found through a bifurcation analysis of the averaged potential. For a double pendulum system, we are able to provide a detailed comparison of analytical, numerical, and experimental results—all of which are found to be in close agreement.

1 Introduction

Recent research (e.g. [10] and [6]) has pointed to the geometric nature of stabilization effects produced by oscillatory forcing of controlled mechanical systems. These results togehter with the nondimensionalized theory presented in [18] suggest that high-frequency oscillatory control designs can be effectively applied to mechanical systems over a very wide range of characteristic length scales. In particular, there is great appeal in applying what are essentially robust open-loop methods for the control of micromechanisms where they allow us to avoid the very challenging problem of closing feedback loops using noisy and unreliable sensor technologies. While a great deal of previous work on the oscillatory control of mechanical systems has been aimed at so-called *superarticulated mechanical systems* comprised of multibody chains (e.g. [18],[17], and [15]), recent work suggests interesting applications of the basic concepts to the control of micromechanical systems such as bubble dyanamics ([9]) or the silicon cantilevered beam and plate structures arising in MEMS applications ([7]). A crucial result in applying these methods to systems at small characteristic length scales is that stabilization effects are scale invariant. The present report is aimed at illustrating some of the fundamental questions which are raised in studying this issue, and we have chosen to

*Support from the Army Research Office under the ODDR&E MURI97 Program Grant No. DAAG55-97-1-0114 to the Center for Dynamics and Control of Smart Structures (through Harvard University) is gratefully acknowledged. Partial support from the U.S. Office of Scientific Research through grant F49620-96-1-0059 is also gratefully acknowledged.

place the discussion in the context of the response of heavy chains to rapid oscillatory forcing. This has the threefold advantage of pushing the theory further into the real of multi-degree-of-freedom mechanical systems, giving the questions of length scale a very clear physical interpretation, and at the same time making contact with a body of previously published work.

Our starting point is to consider controlled mechanical systems which can be described by the Lagrangian

$$\mathcal{L}(Q,\dot{Q};v) = \frac{1}{2}\dot{Q}^T\mathcal{M}(Q)\dot{Q} + v\mathcal{A}(Q)^T\dot{Q} - \mathcal{V}_a(Q;v), \tag{1}$$

where $v \in \mathbb{R}$ is the control input, $\mathcal{M}(Q) \in \mathbb{R}^{n\times n}$ is a state-dependent inertia tensor, $\mathcal{A}(Q) \in \mathbb{R}^n$ is a Coriolis coupling vector, and $\mathcal{V}_a(Q;v) \in \mathbb{R}$ is known as the *augmented potential.* Such systems typically arise from $n+1$ degree-of-freedom systems where the velocity of one degree of freedom is viewed as the control input v. Details of the model order reduction process leading to (1) may be found in [4]. The *noncanonical Hamiltonian* $\mathcal{H}(Q,P)$ corresponding to (1) is found through the Legendre transformation $\mathcal{H}(Q,P) = P^T\dot{Q} - \mathcal{L}$:

$$\mathcal{H}(Q,P;v) = \frac{1}{2}(P - v\mathcal{A}(Q))^T\mathcal{M}(Q)^{-1}(P - v\mathcal{A}(Q)) + \mathcal{V}_a(Q;v) \tag{2}$$

where $P = \{P_1, P_2, \ldots, P_n\}^T$ are canonical momenta defined by $P = \partial\mathcal{L}/\partial\dot{Q}$. We refer to (2) as "noncanonical" because for an arbitrary input v, energy is not necessarily conserved (i.e. $\partial\mathcal{H}/\partial t \neq 0$). If we assume now that v is a periodic function of time (i.e. $v(t+T) = v(t)$ for some $T > 0$), in principle $\mathcal{H}(Q,P;v)$ can be averaged over one period of $v(t)$ to obtain the *averaged Hamiltonian* $\overline{\mathcal{H}}$, which is written

$$\overline{\mathcal{H}}(Q,P) = \frac{1}{2}\left(P - \overline{v}\mathcal{A}(Q)\right)^T\mathcal{M}(Q)^{-1}\left(P - \overline{v}\mathcal{A}(Q)\right) + \mathcal{V}_A(Q), \tag{3}$$

where $\overline{v} = \frac{1}{T}\int_0^T v(t)dt$, $\overline{v^2} = \frac{1}{T}\int_0^T v^2(t)dt$, and

$$\mathcal{V}_A(Q) = \frac{1}{2}\left(\overline{v^2} - \overline{v}^2\right)\mathcal{A}(Q)^T\mathcal{M}(Q)^{-1}\mathcal{A}(Q) + \mathcal{V}_a(Q) \tag{4}$$

is known as the *averaged potential.* The resulting averaged Hamiltonian is a proper Hamiltonian (i.e. $\partial\overline{\mathcal{H}}/\partial t = 0$), and the equilibria and stability of the associated dynamics may be studied in terms of the critical points of $\mathcal{V}_A(Q)$. The similarity of the dynamics associated with (3) to the dynamics associated with (2) is largely dependent on the forcing amplitude and frequency. At the moment, this method of averaging is a formal method which has been shown to give meaningful results for many systems, and the development of a more rigorous theory of averaging for such systems is an active research area.

The widely studied vertically-forced pendulum[8, 4, 18] possesses a Hamiltonian of the form (2), and the stabilization of a more general pendulum system using the method of Hamiltonian averaging has recently been presented in [18]. In this work, the system being considered is a pendulum, where the forcing is directed along a line at some angle α measured with respect to the horizontal. The dynamics of the pendulum are given by the second order o.d.e.

$$I\ddot{\theta} + m_b\ell\dot{v}\cos(\theta - \alpha) + m_bg\ell\sin\theta = 0, \tag{5}$$

where $v(t) = \dot{r}(t)$ is the velocity which the pendulum hinge point is forced to follow. The parameters are m_b =mass of the pendulum, I =the planar moment of inertia of the pendulum computed about the hinge point, and ℓ = the distance from the hinge point to the center of mass of the pendulum. The averaged potential for the mechanism is

$$\mathcal{V}_A(\theta) = \frac{m_b^2\ell^2}{2I}\sigma^2\cos^2(\theta - \alpha) - m_bg\ell\cos\theta, \tag{6}$$

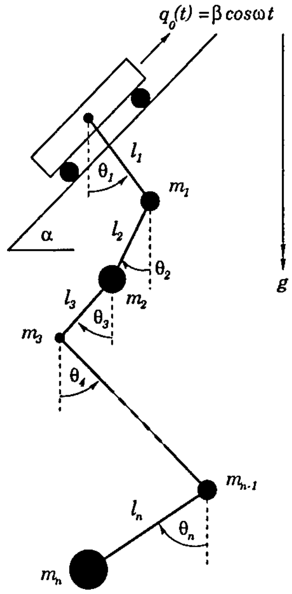

Figure 1: The rapidly forced cart and n-pendulum on an inclined plane.

where $\sigma^2 = (1/T)\int_0^T v(t)^2\,dt$, with $v(\cdot)$ being a piecewise analytic periodic function of period T. Although the exact form of the periodic input is of little importance (as noted in [7]), to focus the discussion, suppose $v(t) = \beta\omega\sin\omega t$, corresponding to having the pendulum hinge point execute the motion $q_0(t) = -\beta\cos\omega t$ along the prescribed axis. In the nondimensional version of this problem (as described in [18]), two dimensionless parameters play a role:

$$\delta \equiv \frac{\beta}{\ell} \quad \text{and} \quad \epsilon \equiv \frac{g}{\ell\omega^2},$$

in terms of which we write down a nondimensional averaged potential $V_A(q)$ as described in the next section. In rescaling the noncanonical Hamiltonian and computing the averaged potential, it is found that for high frequencies (small ϵ), the averaged system tends to stabilize in the direction of oscillation. Further analysis shows that the nonautonomous system executes periodic motions, called *hovering motions* [16, 18, 15], about the averaged fixed point, where the period of motion is the same as that of the forcing, and the amplitude is roughly proportional to the forcing amplitude divided by the forcing frequency squared ($= \delta\epsilon$, and hence the amplitude of the periodic motions tends to zero as $\epsilon \to 0$). Equilibria of the averaged system and periodic orbits of the nonautonomous system are shown to arise as the result of bifurcations at critical values of the parameters δ and ϵ. Furthermore, it has been shown shown that there is a close correlation between the averaged phase space and Poincaré map of the nonautonomous system.

In this report, we show that many of the qualitative features of the single-degree-of-freedom (SDOF) system's dynamics also appear in the dynamics of periodically forced n–link chains. Questions concerning charateristic length scales become more important and more subtle, however. Specifically, the focus of this report is the bifurcation and stability theory of a periodically forced cart and n-pendulum on an inclined plane, illustrated in Figure 1. Roughly speaking, the question we study is how the various links in the chain align themselves when the cart executes high frequency but small amplitude periodic motions. As with the SDOF pendulum, parts of this story have roots dating back to the early part of this century[12]. Stabilization of the inverted n-pendulum has been more recently considered by Acheson[1]. In this work, the Floquet-theoretic approach is applied to a linear model of n-pendulum dynamics in a neighborhood of the fully inverted equilibrium. The main result of this report is a bounding of the forcing amplitude at which the fully inverted equilibrium is stabilized given the minimum and maximum natural frequencies (ω_{min} and ω_{max} in [1], resp.) and the forcing frequency (ω_0). Acheson remarks that a necessary condition for stabilization is $\omega_0^2 \gg \omega_{max}^2$. Furhter results in this direction will be discussed below in Section 5.

An advantage of taking the averaging approach outlined in this introduction is that it provides the means to find, determine the stability, and study bifurcations of all equilibria. While it is, in principle, possible to generally follow the procedure used to study the averaged SDOF system as outlined in [18], the n-DOF system is considerably more complex. Specifically, it possesses many more symmetries arising from mass and length parameters, whose breakings have profound effects on the equilibrium and bifurcation structure of the system. The phase space geometry of the SDOF system lends itself well to a detailed study of the stability and basins of attraction of equilibria of the averaged system dynamics. For many-link pendula, the picture is obscured by the large number of dimensions of the system's phase space. A more important problem, however, is that in the n-DOF system the connection between the averaged system dynamics and the nonautonomous (controlled) system dynamics depends crucially on the characteristic length scale of the smallest system component. This problem is computationally tractable for the double pendulum, whose dynamics we consider in Section 5. In the general case, it may be possible to use the averaged Hamiltonian in estimating the size of the basins of attraction, and we comment on this possibility in the conclusion.

2 The Rapidly Forced Cart and n-Pendulum on an Inclined Plane

2.1 Averaging and Characteristic Length Scale in the SDOF Case

Before treating the general case, it is useful to describe the connection between characteristic length scales and our theory of averaging in the case of a SDOF pendulum on a cart. The goal here is to understand the stability characteristics in terms of the critical points of $\mathcal{V}_A(\cdot)$ given in (6). These points are found as solutions to the critical point equation

$$2Ig\sin\theta - m_b\ell\sigma^2\sin 2(\theta-\alpha) = 0,$$

and when $\alpha = \frac{\pi}{2}$ (the case in which the suspension point is periodically forced along the vertical axis), the inverted configuration $\theta = \pi$ is always a critical point. The averaging theory of [4, 5] states that if a critical point is a local minimum of $\mathcal{V}_A(\cdot)$, then for oscillatory forcing of sufficiently high frequency, the pendulum will execute stable motions in a neighborhood of the point. In the present case, the second derivative $\mathcal{V}''(\pi)$ is positive precisely when $m_b\ell\sigma^2 > gI$. Making the apporoximation $I \approx m_b\ell^2$, and taking the forcing function $v(t) = \beta\omega\sin(\omega t)$, our theory predicts that there will be some value $\omega_s \sim \sqrt{2g\ell}/\beta$ such that the pendulum will execute stable motions very close to the inverted configuration, provided $\omega > \omega_s$. This recovers the classical result on this problem. (C.f. [2], p. 152.)

Since the critical frequency parameter value of the averaged potential is $\omega_{cr} = \sqrt{2g\ell}/\beta$, it is of interest to know how closely ω_{cr} approximates ω_s. Since $\omega_{cr} \sim \mathcal{O}(\ell^{\frac{1}{2}})$, a naive conjecture would be that the stabilization effects of oscillatory forcing are observed at lower frequencies when the pendulum length parameter ℓ becomes small. Unfortunately, this conjecture is false, and we have shown in [7] that $\omega_s \sim \omega_{cr}$ only when $\beta \sim \ell$. This means that to apply averaged potential theory at small characteristic length scales, we need to have $\beta \sim \ell$, and hence $\omega_s \sim \omega_{cr} \sim \sqrt{2g/\ell}$.

For the n–pendulum systems we treat next, there are two characteristic length scales which are important—the total chain length, ℓ, and the length of the shortest link, ℓ_s. Of the two, ℓ_s is the more important for linking the critical point analysis of the averaged potential to the stability of the forced system. Hence in the rescaling discussed below, this length scale will play an important role. A more detailed discussion will be presented in Section 5.

2.2 System Model

Consider the system illustrated in Figure 1, where each m_i is a point mass located at the end of the i-th massless link, ℓ_i is the length of the i-th link, θ_i is the angular deviation of the i-th link with the hanging vertical position, q_0 is the position of the cart along the incline of angle α, and g is the acceleration due to gravity. A kinematic analysis of this system, similar to that presented in [14], results in the derivation of a Lagrangian of the form (1), which is associated with the noncanonical Hamiltonian

$$\mathcal{H}(\theta, P; \dot{q}_0) = \frac{1}{2}(P - \dot{q}_0\mathcal{A}(\theta))^T\mathcal{M}(\theta)^{-1}(P - \dot{q}_0\mathcal{A}(\theta)) + \mathcal{V}_a(\theta; \dot{q}_0), \tag{7}$$

where

$$\begin{aligned} \mathcal{M}_{ij}(\theta) &= \left(\sum_{k=\max(i,j)}^{n} m_k\right)\ell_i\ell_j\cos(\theta_i - \theta_j) \\ \mathcal{A}_i(\theta;\alpha) &= \left(\sum_{k=i}^{n} m_k\right)\ell_i\cos(\theta_i - \alpha) \\ \mathcal{V}_a(\theta;\dot{q}_0) &= \frac{1}{2}\left(\sum_{k=1}^{n} m_k\right)\dot{q}_0^2(t) + \mathcal{V}_g(\theta), \qquad (8) \\ \mathcal{V}_g(\theta) &= \sum_{i=1}^{n}\left[\left(\sum_{k=i}^{n} m_k\right)g\ell_i(1-\cos\theta_i)\right]. \qquad (9) \end{aligned}$$

We immediately note that $\mathcal{H}$ is dependent on the variable q_0 only through the cart velocity $\dot{q}_0$. Suppose that the cart position q_0 is given explicitly as a piecewise analytic periodic function of time—for instance, $q_0(t) \equiv \beta\cos\omega t$.

2.3 Rescaling and Averaging

Let $m = \sum_{i=1}^{n} m_i$, and let ℓ_s be the smallest link length. Then the length of all other links can be written $\ell_i = \lambda_i\ell_s$, where $1 \le \lambda_i \le k$, where k is an $\mathcal{O}(1)$ constant. We nondimensionalize (7) by the change of variables $q_i = \theta_i$, $P = m\omega\ell_s^2 p$, $\mathcal{H} = m\omega^2\ell_s^2 H$, and a change of time scales $\tau = \omega t$. Let $\beta Q(\tau) = q_0(t)$. After making these substitutions and simplifying, we obtain

$$H(q,p;\tau) = \frac{1}{2}(p - \delta v(\tau)A(q))M(q)^{-1}(p - \delta v(\tau)A(q)) + V_a(q;\tau) \tag{10}$$

where

$$\begin{aligned} M_{ij}(q) &= \left(\sum_{k=\max(i,j)}^{n} \mu_k\right)\lambda_i\lambda_j\cos(q_i - q_j) \\ A_i(q;\alpha) &= \left(\sum_{k=i}^{n} \mu_k\right)\lambda_i\cos(q_i - \alpha) \\ V_a(q;\tau) &= \frac{\delta^2}{2}\left(\sum_{k=1}^{n} \mu_k\right)v(\tau)^2 + \epsilon V_g(q), \\ V_g(q) &= \sum_{i=1}^{n}\left[\left(\sum_{k=i}^{n} \mu_k\right)\lambda_i(1-\cos q_i)\right], \qquad (11) \end{aligned}$$

and where $\delta \equiv \beta/\ell_s$, $\epsilon = g/\ell_s\omega^2$, $v(\tau) = \frac{d}{d\tau}Q(\tau)$, and $\mu_i = m_i/m$. The averaged Hamiltonian $\overline{H}$ is defined as $\overline{H}(q,p) = \frac{1}{T}\int_0^T H(q,p;\tau)d\tau$, where T is the fundamental period of $v(\cdot)$. The averaged

Hamiltonian has the general form

$$\overline{H}(q,p) = \frac{1}{2}p^T M(q)^{-1} p + V_A(q;\delta,\epsilon) \tag{12}$$

where the averaged potential $V_A(q;\delta,\epsilon)$ is

$$V_A(q;\delta,\epsilon) = \frac{\delta^2}{4} A(q)^T M(q)^{-1} A(q) + \epsilon V_g(q). \tag{13}$$

Remark 1 The parameters δ and ϵ have very clear physical interpretations. δ represents the ratio of the forcing magnitude β to the chain length ℓ, and ϵ can be interpreted as the ratio of the system "natural frequency" $\omega_n = \sqrt{g/\ell}$ squared to the forcing frequency ω squared. In assuming small amplitude, high-frequency forcing, we are making the assumption that $\beta \ll \ell$ and $\omega_n \ll \omega$. Hence, $0 < \delta \ll 1$ and $0 < \epsilon \ll 1$.

Remark 2 An advantage of using the nondimensional model is that it facilitates the mathmematical discussion concerning equilibria and stability in several physically interesting limits. Of particular interest is the limit in which the forcing frequency becomes infinite while the forcing amplitude remains finite but non-zero. In the nondimensional model, this limit is taken by fixing $\delta > 0$ while letting $\epsilon \to 0$. Another physical interpretation for this limit is that the usual potential forces (i.e. gravity) in V_a become negligible when compared to those potential forces arising from the periodic forcing. While the theory unfolds most easily in terms of the ϵ,δ parameters, it will be useful for physical interpretation to retain also a dimensional version of the averaged potential. The dimensional version of (13) is

$$\mathcal{V}_A(q;\sigma) = \frac{\sigma^2}{2} \mathcal{A}(\theta)^T \mathcal{M}(\theta)^{-1} \mathcal{A}(\theta) + V_g(\theta), \tag{14}$$

where

$$\sigma^2 = \frac{1}{T} \int_0^T v(\tau)^2 \, d\tau,$$

with T being the fundamental period of $v(\cdot)$.

Remark 3 For the sake of explicitness, we may carry out the nondimensionalization of the chain model assuming sinusoidal forcing $q_0(t) = \beta \cos \omega t$, but the assumption of this particular form of the periodic input is not necessary.

3 Behavior of the Averaged System in the Limit $\epsilon \to 0$

In experiments, it is usually the case that the mechanism producing the periodic motion of the pendulum base produces motion of fixed amplitude and variable frequency. This constraint results in fixing the value of β (or, equivalently, δ), while allowing us to vary ω (or ϵ). Studies of stabilization for SDOF systems have shown that stabilization is a high-frequency phenomenon. Therefore, the approach we take is to study the behavior in the limit as $\epsilon \to 0$ and δ is held constant.

We begin by deriving the equilibrium equations, which we write

$$\begin{aligned} M(q)^{-1}p &= 0 \\ \frac{\delta^2}{4}\frac{\partial}{\partial q}\left(A(q)^T M(q)^{-1} A(q)\right) + \epsilon \frac{\partial V_g(q)}{\partial q} &= 0. \end{aligned} \tag{15}$$

Stabilization in the high-frequency $\omega \to \infty$ limit is easily studied by simply letting $\epsilon = 0$. By definition, M^{-1} is nonsingular, from which the first equation of (15) implies that $p = 0$ at an equilibrium. We state the results of our analysis of the second equation in the following theorems.

Theorem 1 *Equilibria associated with (10) in the $\epsilon \to 0$ limit include $q_i = \alpha + k\pi/2$, where $k \in \mathbb{Z}$.*

Proof: Setting $\epsilon = 0$ in (15) leaves us with n equilibrium equations defined by

$$\frac{\partial}{\partial q_i}\left(A(q)^T M(q)^{-1} A(q)\right) = 0. \tag{16}$$

This system of equations may be written in symbolic form as

$$\left(2\frac{\partial A(q)^T}{\partial q} M(q)^{-1} + A(q)^T \frac{\partial M(q)^{-1}}{\partial q}\right) A(q) = 0,$$

where $\frac{\partial A(q)}{\partial q}$ is a diagonal 2-tensor (matrix) and $\frac{\partial M(q)^{-1}}{\partial q}$ is a 3-tensor. Among solutions to this system are those for which one of the following conditions hold:

$$\begin{aligned} A(q) &= 0 & (17)\\ 2\frac{\partial A(q)^T}{\partial q} M(q)^{-1} + A(q)^T \frac{\partial M(q)^{-1}}{\partial q} &= 0, & (18)\end{aligned}$$

since nonzero components of $\frac{\partial A(q)}{\partial q}$ of the form $\sin(q_i - \alpha)$, while nonzero components of $\frac{\partial M(q)^{-1}}{\partial q}$ are of the form $\sin(q_i - q_j)$. This proves the theorem. □

Theorem 2 *Equilibria q_e of (15) with coordinate entries $q_i = \alpha + \frac{(2k_i - 1)\pi}{2}$, $1 \le i \le n$ and $k_i \in \mathbb{Z}$ are stable rest points of the averaged system.*

Proof: We prove the theorem by showing the Hessian matrix $\mathbf{H}(q) = \frac{\partial^2 V_A(q)}{\partial q^2}$ is positive-definite for each q_e of the prescribed form. Positive-definiteness implies that q_e is a local minimum of the averaged potential $V_A(q)$, and therefore q_e is a stable rest point of the averaged Hamiltonian system. Computing partial derivatives, we obtain

$$\begin{aligned}\mathbf{H}(q) &= \left(2\frac{\partial^2 A(q)^T}{\partial q^2} M(q)^{-1} + 3\frac{\partial A(q)^T}{\partial q}\frac{\partial M(q)^{-1}}{\partial q} + A(q)^T \frac{\partial^2 M(q)^{-1}}{\partial q^2}\right) A(q)\\ &\quad + \left(2\frac{\partial A(q)^T}{\partial q} M(q)^{-1} + A(q)^T \frac{\partial M(q)^{-1}}{\partial q}\right)\frac{\partial A(q)}{\partial q}.\end{aligned} \tag{19}$$

At q_s, recall that $A(q_s) = 0$, and therefore (19) simplifies to

$$\mathbf{H}(q_s) = 2\frac{\partial A(q)^T}{\partial q} M(q)^{-1} \frac{\partial A(q)}{\partial q}\bigg|_{q=q_s}.$$

For any vector $x \in \mathbb{R}^n$, we have

$$x^T \left(\frac{\partial A(q)^T}{\partial q} M(q)^{-1} \frac{\partial A(q)}{\partial q}\right)\bigg|_{q=q_s} x = y^T M(q_s)^{-1} y > 0,$$

where we have substituted $y = \left(\frac{\partial A(q)}{\partial q}\right)\Big|_{q=q_s} x$. Because $M(q)^{-1}$ is positive definite by definition, $\mathbf{H}(q_s)$ is also positive definite. □

Theorem 3 *Equilibria $q_i = \alpha + k_i\pi$ experience a bifurcation at $\epsilon = 0$. Further, there exists $\epsilon_1 \ll 1$ such that for $0 < \epsilon < \epsilon_1$, these equilibria are unstable.*

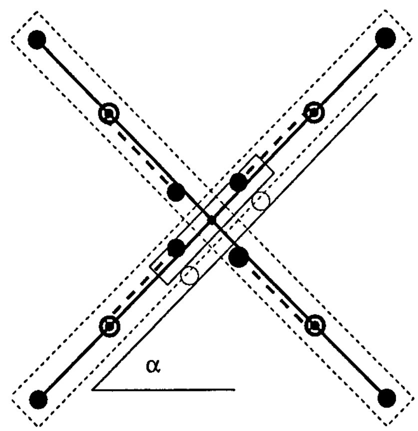

Figure 2: Selected equilibria for the averaged cart and double pendulum in the limit as $\epsilon \to 0$. As stated in Theorem 1, equilibria aligned with the direction of motion are stable, and those aligned normal to the direction of motion are unstable.

Proof: At $q_i = \alpha + n_i\pi$, both $\frac{\partial A(q)}{\partial q} = 0$ and $\frac{\partial M(q)^{-1}}{\partial q} = 0$, in which case $\mathbf{H}(q)$ simplifies to

$$\mathbf{H}(q) = \left(2\frac{\partial^2 A(q)^T}{\partial q^2}M(q)^{-1} + A(q)^T\frac{\partial^2 M(q)^{-1}}{\partial q^2}\right)A(q). \tag{20}$$

We can eliminate either $\frac{\partial^2 A(q)}{\partial q^2}$ or $\frac{\partial^2 M(q)^{-1}}{\partial q^2}$ by computing another partial derivative of (18): solving for $\frac{\partial^2 A(q)^T}{\partial q^2}$ and setting $\frac{\partial A(q)}{\partial q} = 0$ and $\frac{\partial M(q)^{-1}}{\partial q} = 0$, we obtain

$$\frac{\partial^2 A(q)}{\partial q^2} = -\frac{1}{2}A(q)^T\frac{\partial^2 M(q)^{-1}}{\partial q^2}M(q).$$

We substitute this last expression into (20) and find

$$\mathbf{H}(q) = \left(-A(q)^T\frac{\partial^2 M(q)^{-1}}{\partial q^2} + A(q)^T\frac{\partial^2 M(q)^{-1}}{\partial q^2}\right)A(q) = 0.$$

□

Remark 4 Theorem 1 indicates that in the $\epsilon \to 0$ limit, the system stabilizes in configurations where the links are aligned with the cart motion. There are unstable equilibria prescribed by configurations where the links are aligned with the direction normal to the cart motion. This is illustrated for the cart and double pendulum in Figure 2. We consider the equilibria and bifurcations of the cart and double pendulum in Section 5.

4 Bifurcations, Stability, and Behavior of the Vertically Forced n-Pendulum as $n \to \infty$

In this section, we consider stabilization as we pass to the limit in which we let $n \to \infty$ while holding the pendulum length and mass constant. The main result to emerge from this section is that as we add degrees of freedom, the oscillation frequency at which stabilization occurs becomes large, and ultimately tends to infinity.

4.1 Bifurcations of the Vertically Forced n-Pendulum

For the vertically forced pendulum systems ($\alpha = \frac{\pi}{2}$), the link configurations in which all q_i are either 0 or π are stationary points (equilibria) of the forced Hamiltonian system (7) for all values of the input $q_0(\cdot)$. (Cf. Theorem 1 and Theorem 2.) Although these points are always equilibria, their stability characteristics change as we vary the frequency of the oscillatory forcing. This is reflected in bifurcations of the averaged potential described as follows.

Theorem 4 *In the coefficient system (11) associated with the nondimensionalized Hamiltonian, suppose $\alpha = \frac{\pi}{2}$. Let $\{q_s\}$ denote the set of critical points of the averaged potential $V_A(q)$ defined in (13) in which each q_i is either 0 or π, $1 \le i \le n$. For all $q_e \in \{q_s\}$ except the trivial hanging equilibrium $q_e = (0, \ldots, 0)^T$, there exists a critical value of ϵ denoted $\epsilon_{cr} > 0$ such that at $\epsilon = \epsilon_{cr}$, q_e experiences a bifurcation such that for $0 \le \epsilon < \epsilon_{cr}$, q_e is a strict local minimum of V_A.*

Remark 5 *Recall that $\epsilon \sim \frac{1}{\omega^2}$, where ω is the frequency of the oscillatory forcing. Theorem 4 states that for each equilibrium in which all links are aligned vertically, there is a critical frequency ω_{cr} above which the equilibrium is a relative minimum of the averaged potential.*

Proof of Thm. 4: From Theorem 2, equilibria belonging to the set $\{q_s\}$ are stable for the averaged system when $\epsilon = 0$, so what remains to be done to prove the theorem is to show that for each $q_e \in \{q_s\}$ there exists an $\epsilon_{cr} > 0$ at which a bifurcation occurs, changing the stability type of the critical point. Recall that the stability of a critical point of (13) may be determined from the concavity of the averaged potential $V_A(q)$, which is inferred from the Hessian matrix $\mathbf{H}(q)$. Furthermore, the equilibrium experiences a bifurcation if for some set of parameters the averaged potential becomes loses rank, i.e. if the Hessian matrix has one or more zero eigenvalues. For $q_e \in \{q_s\}$, $\mathbf{H}(q)$ can be written

$$\mathbf{H}(q_e) = \left(\frac{\delta^2}{2} \frac{\partial A(q)^T}{\partial q} M(q)^{-1} \frac{\partial A(q)}{\partial q} - \epsilon \frac{\partial^2 V_g(q)}{\partial q^2} \right) \Bigg|_{q=q_e} .$$

Recall from the proof of Theorem 2 that the matrix $\frac{\partial A(q)^T}{\partial q} M(q)^{-1} \frac{\partial A(q)}{\partial q}$ is always positive-definite, and furthermore is symmetric (because $M(q)^{-1}$ is symmetric and positive definite). By definition of $V_g(q)$, $\frac{\partial^2 V_g(q)}{\partial q^2}$ is a diagonal matrix which is easily partitioned into a positive-definite part we shall denote $\mathbf{V}(q)^+$, and a negative-definite part denoted $-\mathbf{V}^-(q)$ where $\mathbf{V}(q)^-$ is itself a positive-definite matrix. Note that for the trivial hanging equilibrium that $\mathbf{V}^-(q) \equiv \{0\}^{n\times n}$, in which case the Hessian is always positive-definite and the equilibrium is always stable. For all other equilibria, the Hessian matrix can be written as the difference of positive-definite matrices as follows:

$$\mathbf{H}(q_e) = \left[\frac{\delta^2}{2} \frac{\partial A(q)^T}{\partial q} M(q)^{-1} \frac{\partial A(q)}{\partial q} + \epsilon \mathbf{V}(q)^+ \right] \Bigg|_{q=q_e} - \epsilon \left[\mathbf{V}(q)^- \right] \big|_{q=q_e} .$$

The theorem will be proved by establishing the following lemma.

Lemma 1 *Let $\mathbf{P}_1$ and $\mathbf{P}_2$ be $n \times n$ positive-definite symmetric matrices. Then there exists a real number $\epsilon_c > 0$ such that $\mathbf{P}_1 - \epsilon \mathbf{P}_2$ is positive definite for all $\epsilon < \epsilon_c$ and $\mathbf{P}_1 - \epsilon \mathbf{P}_2$ is <u>not</u> positive definite for $\epsilon \ge \epsilon_c$.*

Proof of lemma: $\mathbf{P}_1 - \eta \mathbf{P}_2$ is symmetric for all values of the parameter η, and because $\mathbf{P}_1$ is positive definite, for any vector $x \neq 0$ of the appropriate dimension, $x^T \mathbf{P}_1 x > 0$. Given such an x, it is clear that there is some number $\eta_1 > 0$ such that $x^T \mathbf{P}_1 x - \eta\, x^T \mathbf{P}_2 x < 0$ for all $\eta > \eta_1$. Indeed, given x, we can chose $\eta_1 = \eta_1(x) = x^T \mathbf{P}_1 x \,/\, x^T \mathbf{P}_2 x$. The ϵ_c that we seek is the inf of $\eta_1(x)$ over all x such that $\|x\| = 1$. ◇

Making the appropriate substitutions for $\mathbf{P}_1$ and $\mathbf{P}_2$ in the previous lemma and letting ϵ_{cr} be the smallest ϵ_c, there exists for each $q_e \in \{q_s\}$ an $\epsilon_{cr} > 0$ at which $\mathbf{H}(q_e)$ is no longer positive-definite. Given that for $q_e \in \{q_s\}$, the terms $\frac{\partial A(q)^T}{\partial q} M(q)^{-1} \frac{\partial A(q)}{\partial q}$, $\mathbf{V}(q)^+$, and $\mathbf{V}(q)^-$ are constant with respect to q and that the term $\epsilon \mathbf{V}(q)^-$ grows linearly with ϵ, we can conclude that $\mathbf{H}(q_e)$ is never positive definite for $\epsilon > \epsilon_{cr}$. Hence, each q_e is rendered unstable by a bifurcation at some $\epsilon = \epsilon_{cr} > 0$ as ϵ increases in value. □

Remark 6 Note that the type of bifurcation experienced by equilibria in the set $\{q_s\}$ must necessarily be pitchfork, in that exactly two equilibria must bifurcate symmetrically about elements of $\{q_s\}$. Note, however, that this does not imply that other types of bifurcations cannot occur elsewhere. In fact, in our study of the double pendulum in the next section, we shall see that saddle-node bifurcations do occur away from elements of $\{q_s\}$.

4.2 Stabilization of the Fully Inverted Equilibrium as $n \to \infty$

Theorem 5 *Consider the fully inverted equilibrium $q_e = (\pi, \ldots, \pi)^T$ of the n-link, vertically ($\alpha = \pi/2$) forced chain. Assume that all masses and link lengths are equal, and that the total mass and length of the chain is held constant; i.e. $\mu_i = \frac{1}{n}$ and $\lambda_i = 1$ for $1 \leq i \leq n$. Let $\mathbf{H}_n(q_e)$ denote the Hessian matrix associated with this equilibrium, and let ϵ_{cr} be the smallest value of ϵ for which $\det \mathbf{H}_n(q_e) = 0$. Then ϵ_{cr} is strictly decreasing as $n \to \infty$, and the critical frequency ω_{cr} at which the equilibrium is stabilized in the averaged system dynamics tends to infinity (∞) as $n \to$ infty.*

Proof: Our strategy of proof is to convert the determinant calculation to an eigenvalue problem, where the eigenvalues we compute correspond to critical values of ϵ and the smallest of these values is ϵ_{cr}. Because of the special structure of $\mathbf{H}_n(q_e)$, we will be able to show that ϵ_{cr} for the n-DOF system is strictly smaller than for the $(n-1)$-DOF system for all n. We begin by explicitly writing out the determinant as

$$\det \mathbf{H}_n(q_e) = \det\left(\frac{\delta^2}{2}\frac{\partial A(q)^T}{\partial q} M(q)^{-1} \frac{\partial A(q)}{\partial q} + \epsilon \frac{\partial^2 V_g(q)}{\partial q^2}\right) = 0. \tag{21}$$

We now proceed to write each of the terms in the preceding equations in terms of n. First, we note that $\frac{\partial A(q)}{\partial q}$ is a diagonal matrix of the form

$$\begin{aligned}\frac{\partial A(q)}{\partial q} &= \begin{pmatrix} -\left(\sum_{i=1}^n \mu_i\right)\lambda_1 & 0 & \cdots & 0 \\ 0 & -\left(\sum_{i=2}^n \mu_i\right)\lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & -\mu_n\lambda_n \end{pmatrix} \\ &= \frac{1}{n}\begin{pmatrix} -n & 0 & \cdots & 0 \\ 0 & -(n-1) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & -1 \end{pmatrix} = \frac{1}{n}\mathbf{A}.\end{aligned}$$

Similarly, $\frac{\partial^2 V_g(q)}{\partial q^2}$ can be reduced to a diagonal matrix of the form

$$\frac{\partial^2 V_g(q)}{\partial q^2} = \frac{1}{n}\begin{pmatrix} -n & 0 & \cdots & 0 \\ 0 & -(n-1) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & -1 \end{pmatrix} = \frac{1}{n}\mathbf{V}.$$

For $M(q)^{-1}$, we first consider a reduction of $M(q)$: specifically, we have

$$M(q) = \frac{1}{n} \begin{pmatrix} n & n-1 & n-2 & \cdots & 1 \\ n-1 & n-1 & n-2 & \cdots & 1 \\ n-2 & n-2 & n-2 & \cdots & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 1 & \cdots & 1 \end{pmatrix} = \frac{1}{n}\mathbf{M}. \tag{22}$$

The inverse of $\mathbf{M}$ given in (22) can be expressed as

$$\mathbf{M}^{-1} = \begin{pmatrix} 1 & -1 & 0 & \cdots & 0 & 0 & 0 \\ -1 & 2 & -1 & \cdots & 0 & 0 & 0 \\ 0 & -1 & 2 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 2 & -1 & 0 \\ 0 & 0 & 0 & \cdots & -1 & 2 & -1 \\ 0 & 0 & 0 & \cdots & 0 & -1 & 2 \end{pmatrix}$$

for all n. Then $M(q)^{-1} = n\mathbf{M}^{-1}$, and then (21) can be written

$$\det \mathbf{H}_n(q_e) = \det \frac{1}{n}\left(\frac{\delta^2}{2}\mathbf{A}^T\mathbf{M}^{-1}\mathbf{A} + \epsilon\mathbf{V}\right) = \det \frac{1}{n}\overline{\mathbf{H}}_n(q_e) = 0, \tag{23}$$

where

$$\overline{\mathbf{H}}_n(q_e) = \begin{pmatrix} n\left(\frac{\delta^2}{2}n - \epsilon\right) & -\frac{\delta^2}{2}n(n-1) & 0 & \cdots & 0 & 0 \\ -\frac{\delta^2}{2}n(n-1) & 2(n-1)\left(\frac{\delta^2}{2}(n-1) - \epsilon\right) & -\frac{\delta^2}{2}(n-1)(n-2) & \cdots & 0 & 0 \\ 0 & -\frac{\delta^2}{2}(n-1)(n-2) & 2(n-2)\left(\frac{\delta^2}{2}(n-2) - \epsilon\right) & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 4\left(\delta^2 - \epsilon\right) & -\frac{\delta^2}{2} \\ 0 & 0 & 0 & \cdots & -\frac{\delta^2}{2} & 2\left(\frac{\delta^2}{2} - \epsilon\right) \end{pmatrix}.$$

$\overline{\mathbf{H}}_n(q_e)$ is a Jacobi matrix. The theorem can be proven by converting (23) to an eigenvalue problem and making use of the following lemma for Jacobi matrices, which is given in [11].

Lemma 2 *Let $\mathbf{L}_n$ be an $n \times n$ Jacobi matrix of the form*

$$\mathbf{L}_n = \begin{pmatrix} b_1 & c_1 & 0 & \cdots & \cdots & 0 \\ a_2 & b_2 & c_2 & 0 & \cdots & 0 \\ 0 & a_3 & b_3 & c_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & & \vdots \\ 0 & \cdots & \cdots & a_{n-1} & b_{n-1} & c_{n-1} \\ 0 & \cdots & \cdots & 0 & a_n & b_n \end{pmatrix}$$

where $a_i, b_i, c_i \in \mathbb{R}$ for all i and let $\mathbf{L}_r$ denote the principal $r \times r$ submatrix $\mathbf{L}_n[1,\ldots,r|1,\ldots,r]$. If $a_i c_{i-1} > 0$, $i \in 2,\ldots,n$, then

1. *all characteristic roots of $\mathbf{L}_n$ are real and simple, and*
2. *between any two characteristic roots of $\mathbf{L}_n$ lies exactly one characteristic root of $\mathbf{L}_{n-1}$.*

Proof: This lemma appears in [11], and its proof may be found in [3]. ◇

We now premultiply $\overline{\mathbf{H}}_n(q_e)$ by a positive-determinant, $n \times n$ diagonal matrix $\mathbf{C}$ and postmultiply by a positive-determinant, $n \times n$ diagonal matrix $\mathbf{Q}$, where

$$\mathbf{C} = \begin{pmatrix} \frac{1}{n} & 0 & 0 & \cdots & 0 \\ 0 & \frac{1}{n-1} & 0 & \cdots & 0 \\ 0 & 0 & \frac{1}{n-2} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}, \qquad \mathbf{Q} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & \frac{1}{2} & 0 & \cdots & 0 \\ 0 & 0 & \frac{1}{2} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \frac{1}{2} \end{pmatrix} \tag{24}$$

to obtain

$$\det\left[\mathbf{C}\overline{\mathbf{H}}_n(q_e)\mathbf{Q}\right] = \det\overline{\overline{\mathbf{H}}}_n(q_e) = 0, \tag{25}$$

where

$$\overline{\overline{\mathbf{H}}}_n(q_e) = \begin{pmatrix} \frac{\delta^2}{2}n - \epsilon & -\frac{\delta^2}{4}(n-1) & 0 & \cdots & 0 & 0 \\ \frac{\delta^2}{2}n & \frac{\delta^2}{2}(n-1) - \epsilon & -\frac{\delta^2}{4}(n-2) & \cdots & 0 & 0 \\ 0 & -\frac{\delta^2}{4}(n-1) & \frac{\delta^2}{2}(n-2) - \epsilon & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \delta^2 - \epsilon & -\frac{\delta^2}{4} \\ 0 & 0 & 0 & \cdots & -\frac{\delta^2}{2} & \frac{\delta^2}{2} - \epsilon \end{pmatrix}.$$

Equation (25) is now in the form of the eigenvalue problem

$$\det\overline{\overline{\mathbf{H}}}_n(q_e) = \det(\mathbf{L}_n - \xi I) = 0,$$

where $\mathbf{L}_n$ is a positive-definite Jacobi matrix of the form considered in Lemma 2. Note that because $\mathbf{L}_n$ is positive-definite, $\min(\epsilon) \geq 0$. We now note that we can obtain the analogous matrix $\mathbf{L}_{n-1}$ for an $n-1$-DOF system by blocking off the first row and column of $\mathbf{L}_n$. Denote the eigenvalues of $\mathbf{L}_{n-1}$ as $\{\epsilon_{n-1}\}$ and the eigenvalues of $\mathbf{L}_n$ as $\{\epsilon_n\}$. By Lemma 2, the eigenvalues of $\mathbf{L}_{n-1}$ lie between those of $\mathbf{L}_n$. Hence, we have

$$0 \leq \min(\{\epsilon_n\}) < \min(\{\epsilon_{n-1}\})$$

for all n. Therefore, as $n \to \infty$, $\min(\{\epsilon_n\}) \to c$, $c \geq 0$. To complete the proof of the theorem, we note that the critical frequency ω_n is related to ϵ_n by

$$\epsilon_n = \frac{g}{\ell_s \omega_n^2}.$$

Since the length of the shortest link $\ell_s \sim 1/n$, the result of the theorem follows. □

Remark 7 Theorem 5 implies the critical frequency ω_{cr} at which each element of $\{q_s\}$ is rendered stable increases as additional links are added to the pendulum (while holding the total mass and length of the pendulum constant). Consequently, in the limit as the number of links $n \to \infty$, $\omega_{cr} \to \infty$ for each element of $\{q_s\}$. This conclusion suggests that it is therefore impossible to stabilize any equilibria in the set $\{q_s\}$ in the continuum limit; i.e. it is impossible to stabilize an inverted (planar) string.

Theorem 5 can be easily verified using widely available computer algebra software. In this study, we compute the eigenvalues of the $n \times n$ matrix $\mathbf{L}_n$ for $1 \leq n \leq 50$ and $\delta = 0.2$. The results of this calculation are shown in Figure 3, where we see in the left frame that $\min(\xi_n)$ tends toward zero as n becomes large. Consequently, $\epsilon_{cr} \to 0$ quickly, as shown in the right frame.

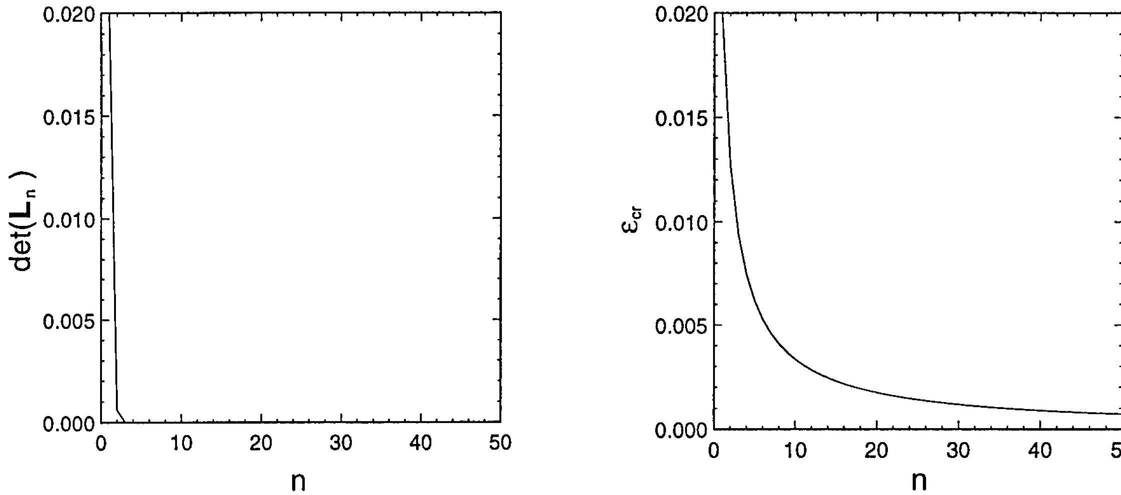

Figure 3: det($\mathbf{L}_n$) vs. n (left) and ϵ_{cr} vs. n for the fully inverted equilibrium of the vertically forced chain. As predicted by Theorem 5, both det($\mathbf{L}_n$) and ϵ_{cr} decrease as n increases.

4.3 Stabilization of the Completely Folded Equilibrium as $n \to \infty$

By employing essentially the same procedure used to prove Theorem 5, we can prove that in the limit as $n \to \infty$, ϵ_{cr} tends to zero for any regular pattern of link configurations. In the next section, we present numerics which suggest that the fully folded configuration is one of the first configurations to stabilize as $\epsilon \to 0$, so it is reasonable to expect that values of ϵ_{cr} for those configurations also tend to zero.

Remark 8 In [18] it was seen in the SDOF system that for $\omega < \omega_{cr}$, the averaged phase portrait generally reflected the Poincaré map of the nonautonomous system, specifically in terms of periodic orbits associated with averaged fixed points and the stability. Periodic orbits associated with averaged fixed points manifest themselves as 1:1 resonances in the Poincaré map. What the averaged phase portrait never captures are subharmonic resonant responses in the Poincaré map. For $\omega > \omega_{cr}$, it was observed in [18] that these subharmonic resonances do not play a significant role in the nonautonomous dynamics. This has also been established rigorously in two limiting cases of the parameters δ and ϵ. For $\omega < \omega_{cr}$, however, subharmonic responses do play a significant role in the nonautonomous dynamics, along with homoclinic tangling of phase space separatrices and the associated bands of stochasticity. It is therefore likely that subharmonic resonances also play a significant role in the dynamics of the periodically forced n-pendulum for $\omega < \omega_{cr}$, and in the "large n" limit dominate the dynamics for all ω. For physical systems, damping and other dissipative forces (e.g. aerodynamic drag) tend to mitigate the chaotic behaviors exhibited by purely conservative systems, and one would expect physical systems to exhibit fairly regular periodic response to forcing.

5 Numerical Studies

5.1 Bifurcations of the Cart and Double Pendulum

The results from Section 2 capture the equilibrium behaviors of the averaged cart and n-pendulum in the case of infinite forcing frequency, which we express in our nondimensional model by allowing $\epsilon \to 0$. As ϵ is perturbed off zero, the equilibria found in Theorem 1 move as the contribution of gravity to potential energy becomes significant, and for distinct critical values of ϵ the system experiences bifurcations. In this section, we focus on obtaining numerical results for the vertically forced double pendulum and comment on bifurcations for the symmetric ($\mu_1 = \mu_2 = \lambda_1 = \lambda_2 = 1/2$)

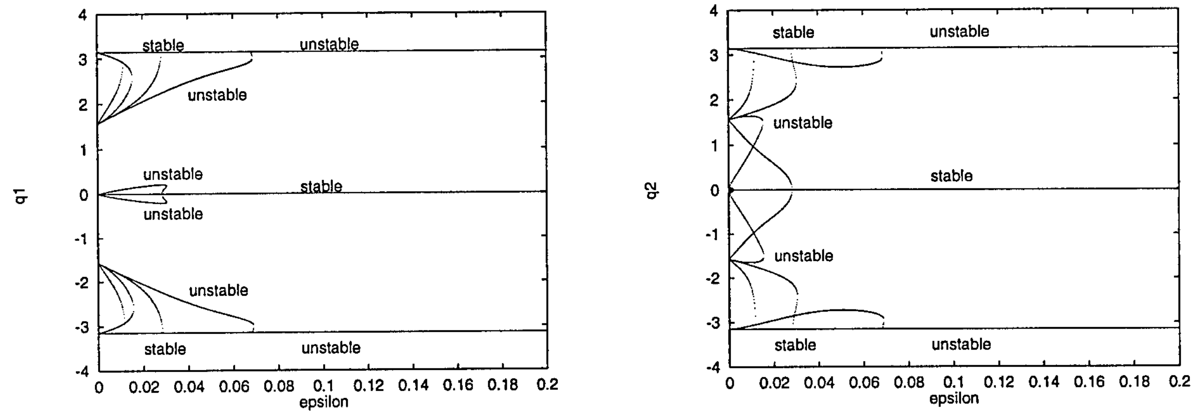

Figure 4: Bifurcation diagram for the averaged cart and double pendulum, where $\mu_1 = \mu_2 = \lambda_1 = \lambda_2 = 1/2$, $\alpha = \pi/2$, and $\delta = 0.2$. In the plots, we see that the inverted and folded equilibria are rendered stable in transcritical pitchfork bifurcations. These bifurcations give rise to the unstable equilibria which are normal to the direction of motion.

system.

The necessary equilibrium equations and Hessian matrix are easily derived from the general form of the averaged potential given in (13). The reader is referred to [13] for full details of this calculation. We now turn our attention to calculations of equilibria and stability for the vertically forced, symmetric ($\mu_1 = \mu_2 = \lambda_1 = \lambda_2 = 1/2$) double pendulum, where we fix $\delta = 0.2$. The algorithm which we use is a two-dimensional bisection algorithm which utilizes the fact that the equilibrium equations for this system define two-dimensional hypersurfaces on the torus. Equilibria are then intersections of the intersections of each surface with zero.

For the vertically forced ($\alpha = \pi/2$) case, we obtain the bifurcation diagram shown in Figure 4. In this figure, we see that for large values of ϵ the system possesses unstable equilibria in configurations aligned with the vertical. As $\epsilon \to 0$, new equilibria are created in a series of transcritical pitchfork bifurcations off the vertically aligned equilibria. These bifurcations have the effect of changing the stability of the inverted and folded equilibria from unstable to stable, and in the process create the unstable equilibria which in the $\epsilon \to 0$ limit correspond to the configurations normal to the direction of motion.

5.2 Stabilization Sequence for Vertically Forced Chains

As the frequency parameter ϵ decreases, the averaged potential for the n–link chain undegoes a rich set of bifurcations, and many equilibria are stabilized by the oscillatory forcing. Some indication of this richness is provided in Figure 5 where for two, three, and four link chains we give the stabilization sequence of equilibria in which successive links have the displayed "up-down" patterns. Here the bifurcation parameter is the r.m.s. value of the forcing $\sigma = \sqrt{\overline{v^2}}$. When $\sigma = 0$, only the hanging equilibrium is stable, and this is the only local minimum of the averaged potential. As σ increases (corresponding to an increase in amplitude or frequency of the forcing), however, the form of the averaged potential changes and other critical points become local minima. There are a number of interesting features in these patterns, including the fact that the case in which all links are inverted is the last to become a relative minimum of the averaged potential (as the forcing frequency is increased) for chains with any number of links.

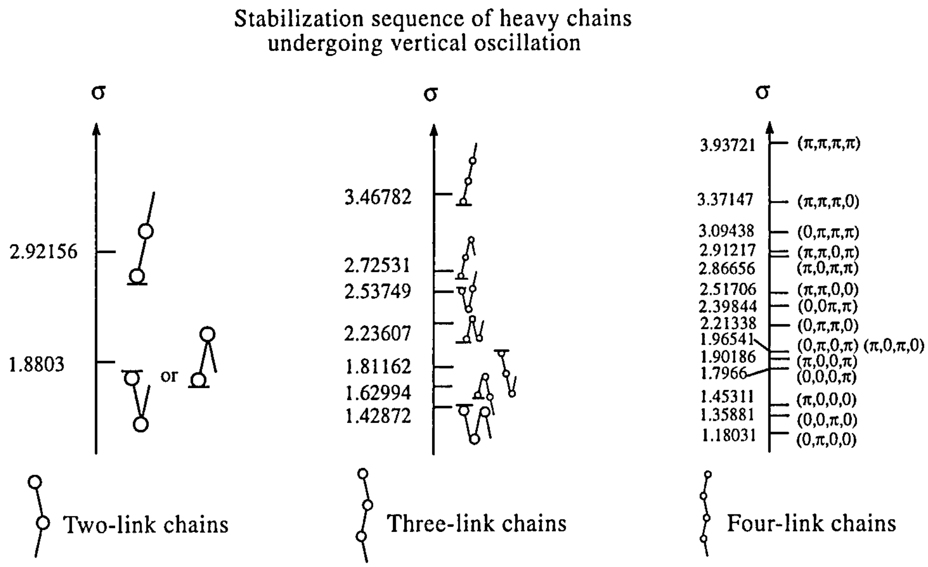

Figure 5: Bifurcation Sequence of Heavy Chains

Another observation worth noting is how the bifurcations of critical points of the averaged potential depend on other physical parameters. Figure 6 illustrates how the number of stable equilibria depends jointly on characteristic link lengths together with r.m.s. value σ. In this figure, the various curves define boundaries in the (ℓ, σ^2) parameter space at which bifurcations occur. Only the first quadrant of the (ℓ, σ^2)–plane is physically menaingful, and this is partitioned into four sectors. In each sector, it is precisely those configurations depicted by the cartoons whose stabilization (under the oscillatory forcing) is predicted by the analysis of the averaged potential. Note that it is only in the topmost sector that stability of *all* vertically aligned configurations is predicted. Here what we see is that the links with shorter lengths tend to become stabilized in the vertically upright configuration before the links with longer lengths. Clearly, if we consider chains with nonuniform links, the bifurcation patterns become very rich as the number of links increases.

These figures raise the question of how precisely the averaged potential analysis predicts the stabilization of an n–pendulum under oscillatory forcing. To some extent this remains an open question. Consider the dimensional version of the averaged potential

$$\mathcal{V}_A(q) = \frac{\sigma^2}{2}\mathcal{A}(q)\mathcal{M}(q)^{-1}\mathcal{A}(q)^T + \mathcal{V}(q).$$

As $v(\cdot)$ ranges over a set of continuous periodic functions of period 1 with

$$\|v\|_2^2 = \int_0^1 v(t)^2\,dt = \beta^2,$$

let $v_\eta(t) = \eta\, v(\eta t)$. Then the r.m.s. value of v_η is

$$\sigma^2 = \int_0^{\frac{1}{\eta}} v_\eta(t)^2\,dt = \beta^2\eta^2,$$

and for a fixed value of β, (as $v(\cdot)$ ranges over a family of periodic functions of L_2–norm β, we ask how much the stabilizing effects of oscillatory forcing by $v_\eta(\cdot)$ can vary.

This question has been answered explicitly for a one-paerameter family of piecewise continuous inputs in the case of a single DOF pendulum in [7]. It is useful to summarize this result before

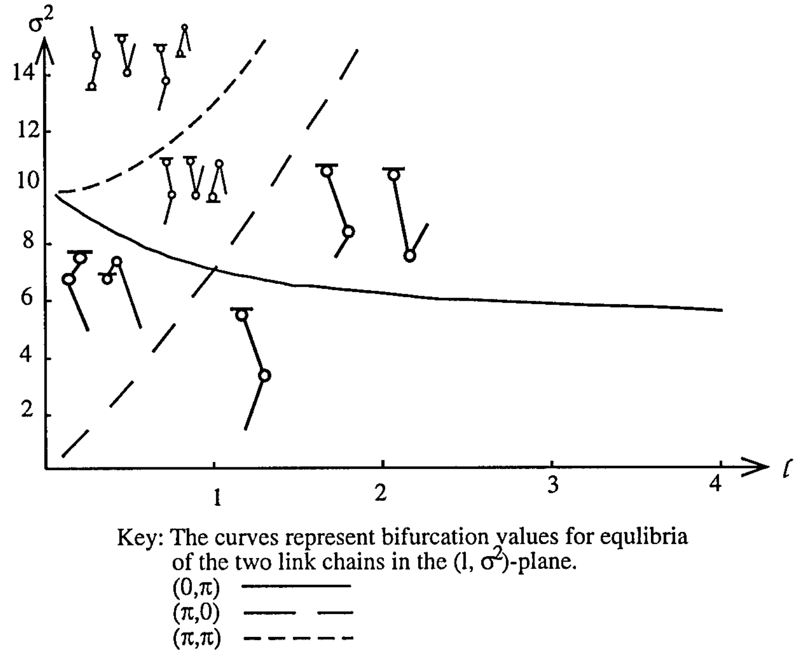

Figure 6: Bifurcation locus of 2-link nonuniform chains. The various curves define boundaries in the (ℓ, σ^2)–parameter space at which bifurcations occur.

addressing the n–DOF case. Take $\dot{v}$ to be a rectangular wave:

$$\dot{v}(t) = \begin{cases} h & \text{if } 0 \le t < w \\ -k & \text{if } w \le t < 1 \end{cases}$$

whose definition we extend to all $t \in [0, \infty)$ by time-1 periodicity: $\dot{v}(t+1) = \dot{v}(t)$. The condition that v is also periodic is satisfied provided that $k = wh/(1-w)$. The dynamics of the system and the averaged potential are given by (5) and (6) respectively with $\alpha = \frac{\pi}{2}$. To study the dependency on frequency of the forcing, let $v_\eta(t) = \eta v(\eta t)$. The corresponding value of σ^2 is

$$\sigma_\eta^2 = \frac{1}{T}\int_0^T v_\eta(t)^2\,dt = \frac{\eta^2 h^2 w^2}{12}.$$

The averaged potential analysis of the stability of the "inverted" equilibrium proceeds as follows. For all η, $\theta = \pi$ satisfies the critical point equation $\mathcal{V}_A'(\theta) = 0$, and since $\mathcal{V}_A''(\pi) = m_b^2\ell^2\eta^2h^2w^2/(12I) - m_b g\ell$, π is a strict local minimum of $\mathcal{V}_A$ whenever $\eta > \sqrt{12Ig/(m_b\ell h w)}$. Consider the linearized dynamics of (5) in a neighborhood of $(\theta, \dot{\theta}) = (\pi, 0)$:

$$I\ddot{\theta} - m_b\ell\dot{v}\theta - m_b g\ell\theta = 0. \tag{26}$$

(See [5] or [15] for information on linearized models of Lagrangian control systems.) We shall apply a Floquet type analysis to the dynamics of (26) in terms of the time-T mapping. ($T = 1/\eta$ is the period of the forcing.) Specifically, letting $(x_k, y_k) = (\theta(k/\eta), \dot{\theta}(k/\eta))$, this map is given explicitly by

$$\begin{pmatrix} x_{k+1} \\ y_{k+1} \end{pmatrix} = e^{A_\eta}e^{B_\eta}\begin{pmatrix} x_k \\ y_k \end{pmatrix} \tag{27}$$

where

$$A_\eta = \frac{1-w}{\eta}\begin{pmatrix} 0 & 1 \\ \frac{m_b\ell}{I}(g - \eta^2\frac{wh}{1-w}) & 0 \end{pmatrix}$$

and

$$B_\eta = \frac{w}{\eta}\begin{pmatrix} 0 & 1 \\ \frac{m_b\ell}{I}(g+\eta^2 h) & 0 \end{pmatrix}.$$

Normalizing the rectangular wave so that $hw = 1$, the second derivative of the averaged potential becomes positive when $\eta > \eta_{cr} = \sqrt{12Ig/(m_b\ell)}$. The averaging theory of mechanical systems ([5]) asserts that there is some value η_s such that for all $\eta > \eta_s$ the linearized model (26) will be Lyapunov stable. Within the class of rectangular wave forms we have specified, η_{cr} will in fact vary even though the r.m.s. value σ on which the averaged potential depends remains constant. More specifically, when $hw = 1$, we shall hold $\sigma_\eta^2 = \eta/12$ constant. Then, letting w range between 0 and 1, the shape of the forcing function $\dot{v}(\cdot)$ changes from a large-amplitude, short-duration up-pulse followed by a relatively shallow negative value (when $w \sim 0$) to a relatively long shallow positive value followed by a large-amplitude down-pulse (when $w \sim 1$). This difference in wave-form shape influences the range of values of η for which (26) is Lyapunov stable. The analysis is carried out in terms of the time-T map, (27), and in particular, we conclude that (26) is stable in the sense of Lyapunov precisely when the eigenvalues of $e^{A_\eta}e^{B_\eta}$ lie on the unit circle. (For more details on this Floquet type of stability analysis, see [5].)

For a single DOF pendulum, it is not difficult to find a fairly explicit relationship between η_s and pulse-width parameter w. In [7] we make the somewhat surprising observation that η_s does not vary a great deal as a function of w, although a waveform characterized by a prominent, large-amplitude downward pulse, with relatively slow return upwards (i.e. in terms of the parameter w we have $w \sim 1$) will stabilize the system at a lower frequency than waveforms for which w is not close to 1. The main result of [7] is that in comparing the bifurcation value η_{cr} with the critical stabilization frequency η_s we have $\eta_s \sim \eta_{cr}$ provided $\beta \sim \ell$.

For a more detailed picture, consider a pendulum with parameter values $m_b = 1/2$, $I = m_b\ell^2$, and take $g = 10$. To examine the case in which β is scaled in direct proportion to the charateristic link length ℓ, we take $\beta = \ell$. Then $\eta_{cr} = \sqrt{120/\ell}$, and in Figure 7 we plot $\eta = \eta_s/\eta_{cr}$ as a function of the pulse-width parameter w. We again see that the stabilizing effect of the oscillatory forcing is relatively insensitive to variation in the waveform. The more remarkable feature of Figure 7, however, is that the relationship between η_s and η_{cr} is totally independent of the characteristic length scale ℓ.

The stabilization effects of oscillatory forcing of n–DOF pendula depend more subtly on characteristic length scales, and it can be shown that the forcing amplitude should scaled with the shortest link length (corresponding to the fastest time constant in the system) in order to ensure that the key qualitative features of the system response are predicted by analyzing the averaged potential. To understand this remark in a bit more detail, we consider a simple two-link system. If we vary the relative link lengths in the chain (as depicted by the cartoon systems in Figure 6), the relationship between the bifurcation values η_{cr} of the averaged potential and the actual critical stabilization frequencies can change a great deal. This is illustrated in Figure 8 where we plot $\eta = \eta_s/\eta_{cr}$ as a function of the pulse-width parameter w for the configuration $(\theta_1, \theta_2) = (\pi, 0)$ in the case of link lengths $\ell_1 = 0.9$ and $\ell_2 = 1$. Here it is no longer the case that the graph remains unchanged as ℓ_1 varies, and in fact we have found that the slope of $\eta(w)$ becomes steeper as ℓ_1 is decreased relative to $\ell_2 = 1$. For this simple family of rectangular wave-form inputs, the stabilizing effect of the periodic forcing is increasingly sensitive to the pulse-width parameter as the relative link lengths become more disproportionate. This suggests there may be a governing robustness principle wherein increasing the degree of nonuniformity in the pendulum system will make it more difficult to apply averaged potential theory to determine stability conditions for the forced system.

5.3 Simple Experiments in Open-Loop Control

Experiments in open-loop oscillatory control for pendulum systems are not terribly difficult to construct and yield compelling demonstrations of the effects of high-frequency forcing in the stabilization of mechanical systems. The apparatus with which we performed our experiments was constructed

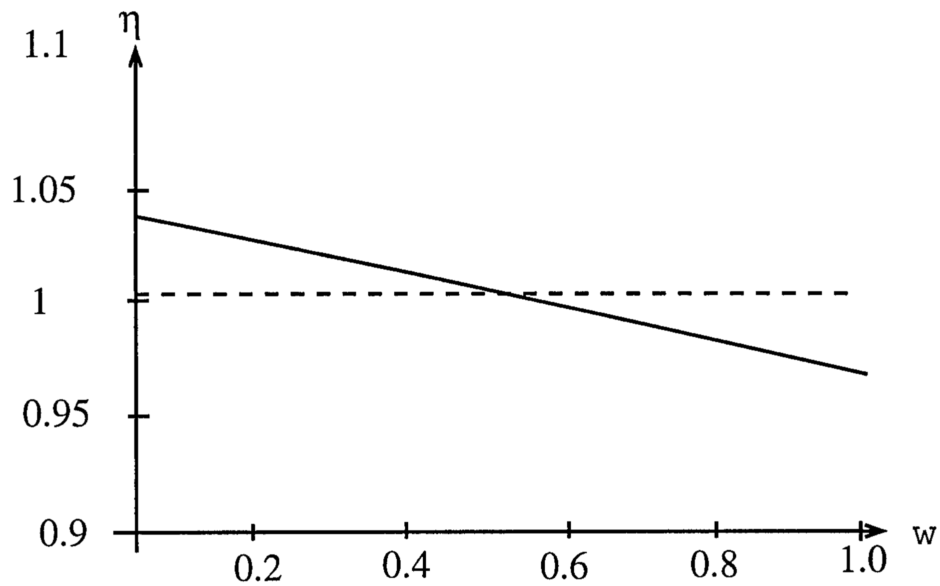

Figure 7: For the single DOF (simple) inverted pendulum, the normalized stabilization frequency $\eta = \eta_s/\eta_{cr}$ is plotted (solid line) as a function of pulsewidth parameter w. *The graph remains unchanged for all values of pendulum length* ℓ.

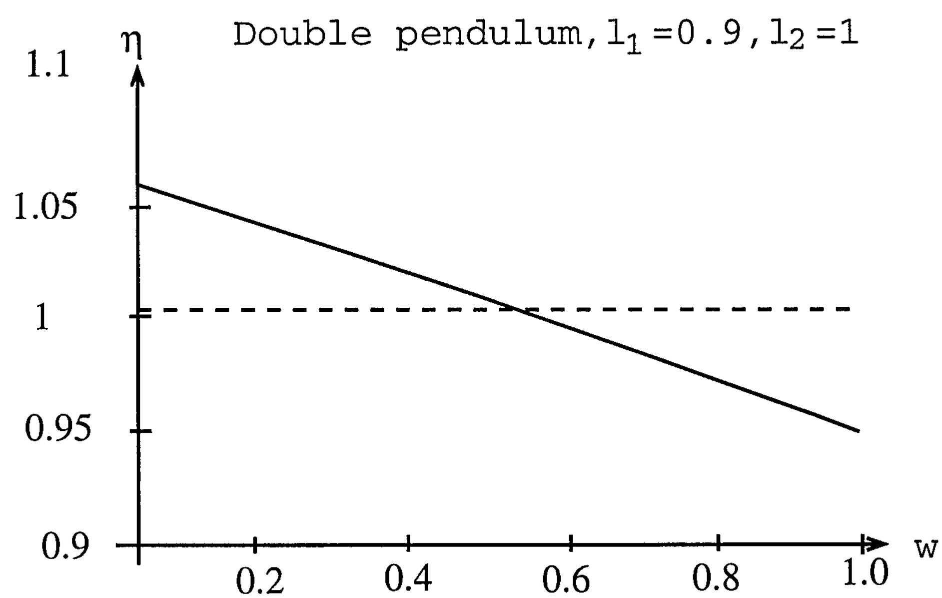

Figure 8: For the 2–DOF pendulum, this graph depicts the normalized stabilization frequency $\eta = \eta_s/\eta_{cr}$ for the configuration $(\theta_1, \theta_2) = (\pi, 0)$. η is plotted (solid line) as a function of pulsewidth parameter w. The link-length parameters are $\ell_1 = 0.9$ and $\ell_2 = 1$. This graph changes with the average slope of the solid line becoming more steep as ℓ_1 is reduced relative to the length $\ell_2 = 1$.

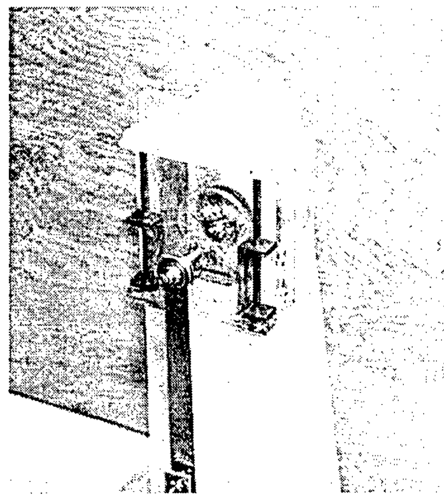

Figure 9: View of the mechanism which rectifies rotary motion to periodic vertical motion.

from readily available plexiglass and plywood stock, with a common toy d.c. motor and strips of aluminum. The mechanism which rectifies the rotary motion of the motor to periodic linear motion is illustrated in Figure 9. Note that the second (outer) link of the pendulum used in the experiment was slightly shorter than the first (inner) link, hence the pendulum is slightly different from the symmetric pendulum considered in many of the previous calculations.

In the experiments conducted, motor (and consequently oscillation frequency) control was open-loop. Hence we only make qualitative observations on the stabilization of equilibria. Using this apparatus, we were able to stabilize all equilibria predicted by our analysis of the averaged potential. These equilibria are depicted in Figure 10. In conducting these experiments, it was observed that the stabilization sequence agreed with the numerical calculations of the last section. Specifically, the hanging folded equilibrium stabilized at roughly the same motor voltage ($\sim$oscillation frequency) as the inverted folded equilibrium, and the fully inverted equilibrium stabilized at a much higher motor voltage. Furthermore, the hanging folded equilibrium was observed to be more robust to disturbances than the inverted folded equilibrium, and much more robust than the fully inverted equilibrium.

5.4 Basins of Attraction for the Cart and Double Pendulum

To this point, we have focused on the stabilization of equilibria and ifurcations of equilibria as the result of variations in primarily the forcing frequency. What arises as an important issue for the n-link chain that is not so important for the single link chain, and is important for the type of open-loop control we propose in general, is the long term dynamics of n-DOF systems with arbitrary initial conditions. The issue of long-term dependence on initial conditions, or basins of attraction, is approachable for SDOF systems because of the existence of powerful analytical tools such as Melnikov's method and invariant manifold theory, and essentially because it is possible to compute and visualize phase space separatrices. This analysis has been performed for the SDOF rapidly forced pendulum[18] and for an entire class of SDOF velocity controlled systems[15]. At present, the theory for higher dimensional systems is highly specialized and much less developed.

Our long term goals, for which we provide preliminary results here, revolve around three questions:

1. To what extent do basins of attraction of the averaged system describe those of the nonautonomous system?

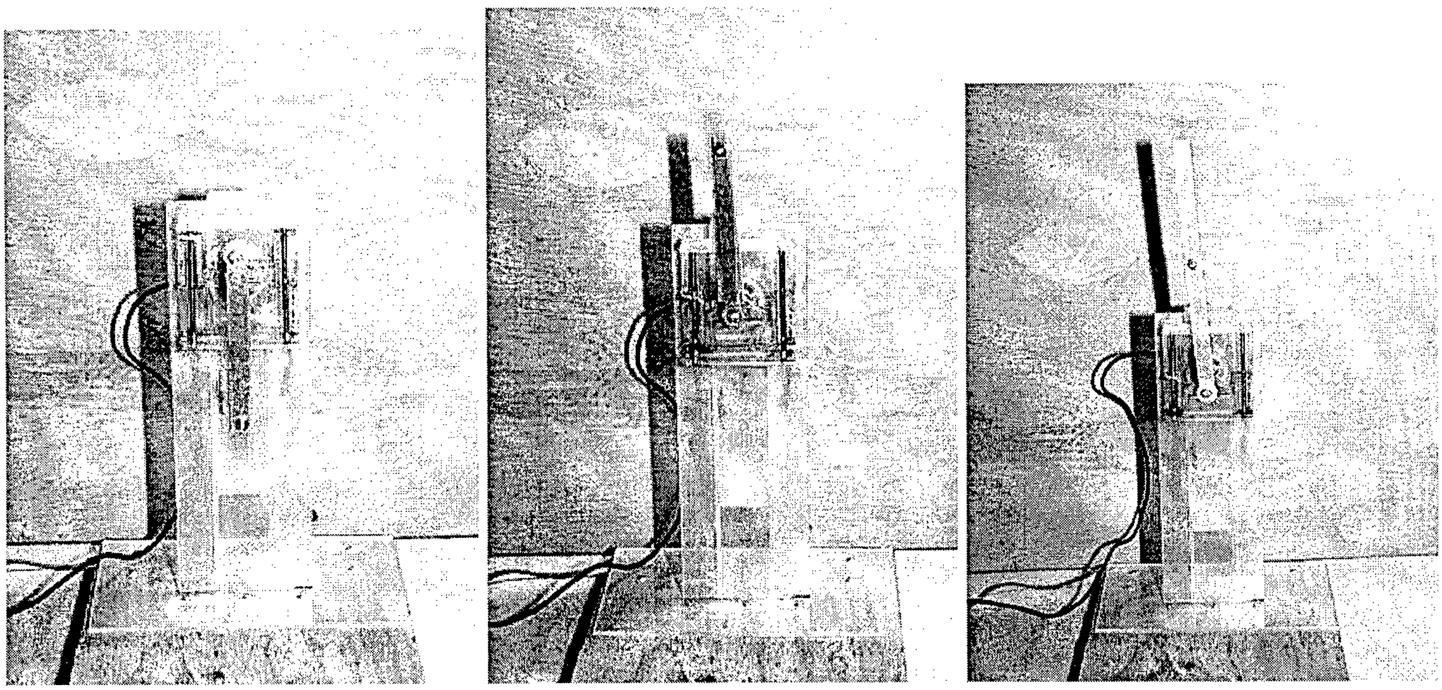

Figure 10: Stabilization of the hanging folded equilibrium(left), the inverted folded equilibrium(center), and the fully inverted equilibrium(right). In conducting these experiments, it was observed that the stabilization sequence agrees with that predicted in the previous section.

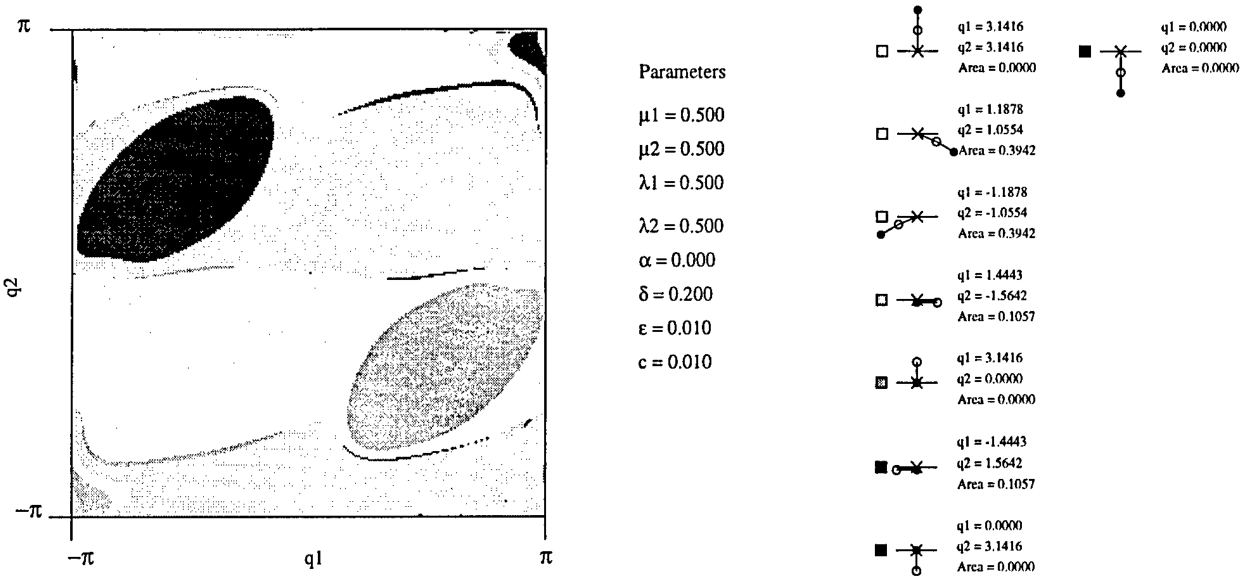

Figure 11: Basins of attraction for the averaged cart and double pendulum.

2. How can forcing or other physical parameters be varied to favor one or more basins of attraction?

3. How can the averaged potential be used directly to estimate either the basins of attraction or the probability that an arbitrary initial condition will be attracted to a specific equilibrium?

In addition, a detailed understanding of the basins of attraction possibly will enable to quantify an equilibrium's robustness to disturbances; i.e. quantify how much of a disturbance will force the system into the basin attraction of another equilibrium.

The dimension of the phase space for even $n = 2$ makes numerical calculations difficult unless we restrict our attention to some two dimensional subset of initial conditions. One reasonable set of initial conditions are those where the system starts at rest. To develop intuition for the first question, we have simulated the double pendulum system with small damping over a 256 × 256 grid of initial conditions which cover the configuration manifold T^2 for a variety of inclinations α, forcing parameters δ and ϵ, and inertial parameters μ_i and λ_i. The results for one set of simulations are shown in Figures 11 and 12, where as indicated in the plot, $\alpha = 0$. In the figures, we see a rough correspondence between the basins of attraction of the averaged and nonautonomous systems, which is typical of these plots for other parameter combinations. Noticeable in the basins of attraction of the nonautonomous systems are the irregular basin boundaries and remnants of the regions of stochasticity which are known to form around basins of regular motion in periodically forced Hamiltonian systems. The basin boundaries represent intersections of unstable manifolds of the various equilibria with the initial condition space, and the irregularities suggest that at the value of damping used in the simulation, significant tangling of stable and unstable manifolds persists. This was observed in the case of the SDOF system in [18], where theorems were presented which establish bounds on the tangling of phase space separatrices. Also note that the folded equilibria dominate the initial condition space, as suggested by the results of the previous section.

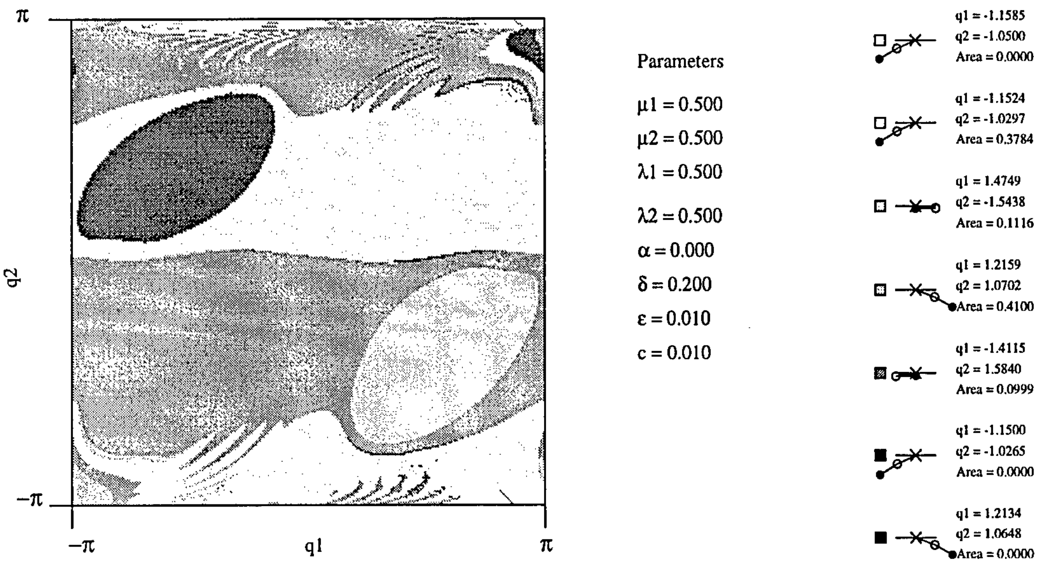

Figure 12: Basins of attraction for the nonautonomous cart and double pendulum.

6 Conclusion

In this report, we have considered the stabilization of a cart and n-pendulum system on an inclined plane by high-frequency periodic forcing. The viewpoint we have taken is that we wish to understand the details of behavior induced by a very simple type of control action applied to a reasonably complex mechanical system. Specifically, we have found that as the forcing frequency becomes large, the n-pendulum system aligns itself with the cart motion. These results were obtained by deriving a dimensionless noncanonical Hamiltonian, and then averaging the noncanonical Hamiltonian to obtain an average Hamiltonian. The results followed from a critical point analysis of the averaged potential. In addition, we saw that the results obtained through the use of the averaged potential may be connected with results obtained through Floquet theory. Finally, we confirmed the results of Section 2 for the vertically forced cart and double through a numerical study of the system's bifurcations and stability. A striking feature distinguishing the n–DOF systems from the more widely studied SDOF case is the sensitivity of the results to nonuniformity in the characteristic length scales of the system components.

References

[1] D.J. Acheson. A pendulum theorem. *Proc. R. Soc. Lond.*, A(114):239–245, 1993.

[2] V.I. Arnol'd, V.V. Kozlov, and A.I. Neishtadt. *Mathematical Aspects of Classical and Celestial Mechanics*, volume 3 of *Encyclopaedia of Mathematical Sciences.* Springer-Verlag, Berlin, 1985.

[3] F. M. Arnscott. Latent roots of tri-diagonal matrices. *Proc. Edinburgh Math. Soc.*, 12:5–7, 1961. Edinburgh Math. Notes No. 44.

[4] J. Baillieul. Stable average motions of mechanical systems subject to periodic forcing. In *Dynamics and Control of Mechanical Systems: The Falling Cat and Related Problems: Fields Institute Communications*, pages 1–23, Providence, R.I., 1993. American Mathematical Society.

[5] J. Baillieul. Energy methods for the stability of bilinear systems with oscillatory inputs. *Intl. J. of Robust and Nonl. Cont.*, pages 285–301, July 1995. Special Issue on the "Control of Nonlinear Mechanical Systems".

[6] J. Baillieul. The geometry of controlled mechanical systems. In J. Baillieul and J.C. Willems, editors, *Essays in Mathematical Control Theory.* Springer-Verlag, New York, 1998. To appear.

[7] J. Baillieul and S. Weibel. Scale dependence in the oscillatory control of micromechanisms. Submitted to the 1998 Conference on Decision and Control., 1998.

[8] N.N. Bogoliubov and Y.A. Mitropolsky. *Asymptotic Methods in the Theory of Nonlinear Oscillators.* International Monographs on Advanced Mathematics and Physics. Gordon and Breach Science Pub., New York, 1961.

[9] Z.C. Feng and L.G. Leal. Nonlinear bubble dynamics. *Annu. Rev. Fluid Mechanics*, 29:201–243, 1997.

[10] M. Levi and W. Weckesser. Stabilization of the inverted pendulum by high frequency vibrations. *SIAM Review*, 27(2):219–223, June 1995.

[11] M. Marcus and H. Minc. *A Survey of Matrix Theory and Matrix Inequalities.* Dover, New York, 1964.

[12] A. Stephenson. On induced stability. *Phil. Mag.*, (17):765–766, 1909.

[13] S. Weibel. *Applications of Qualitative Methods in the Nonlinear Control of Superarticulated Mechanical Systems.* PhD thesis, Boston University, 1997.

[14] S. Weibel and J. Baillieul. Oscillatory control of bifurcations in rotating chains. In *Proc. 1997 ACC, Albuquerque*, pages 2713–2717, 1997.

[15] S. Weibel and J. Baillieul. Averaging and energy methods for robust open-loop control of mechanical systems. In J. Baillieul, S. Sastry, and H. Sussmann, editors, *Essays in Mathematical Robotics*, pages 203–270. Springer-Verlag, Berlin, 1998. To appear.

[16] S. Weibel, J. Baillieul, and T.J. Kaper. Small-amplitude periodic motions of rapidly forced mechanical systems. In *Proc. 34th IEEE CDC, New Orleans*, pages 533–539, 1995.

[17] S. Weibel, J. Baillieul, and B. Lehman. Equilibria and stability of an n-pendulum forced by rapid oscillations. To appear Proc. 36th IEEE CDC, San Diego, 1997.

[18] S. Weibel, T.J. Kaper, and J. Baillieul. Global dynamics of a rapidly forced cart and pendulum. *Nonl. Dyn.*, 13:131–170, 1997.