Averaging and Control of Nonlinear Systems

(with Application to Biomimetic Locomotion)

Thesis by
Patricio Antonio Vela

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

California Institute of Technology
Pasadena, California

2003

(Defended May 5, 2003)

Acknowledgements

I would like to acknowledge my parents and family for the support and guidance they have provided me throughout the years. In many ways they have helped me navigate the turbulent waters of life and Caltech. I would also like to thank my advisor, Professor Joel Burdick, for accepting me as a graduate student and guiding me through the maze of literature regarding my research. No matter how vague the idea, he has been able to provide me with enough references to bring the idea to maturation.

Professors Jerrold Marsden and Richard Murray have also been quite influential in molding me academically, by providing a healthy philosophy and background regarding control and dynamical systems theory. When traveling to conferences and communicating with others, I especially appreciate their and Professor John Doyle's efforts regarding the CDS department. The extent of preparation and training I received is manifest on these occasions.

While at Caltech, I have had the priviledge and benefit of meeting many great people. Being senior to me, Jim Radford and Todd Murphey have been invaluable for their insights into academia and research. A fortunate accident introduced me to Michelle Plug, from whom I have learned much. Ann-Marie Polsenberg has been a great office-mate over the past year; our office certainly has rocked. Chris Boxe, Ebony Blanchard, and Aristo Asimokopolous have helped maintain my sanity over the years.

I would also like to thank Charmaine Boyd and Maria Koeper for having my back. In all, I simply would like to thank those who have made Caltech a great experience, academic and otherwise.

Averaging and Control of Nonlinear Systems
(with Application to Biomimetic Locomotion)

by
Patricio Antonio Vela

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

Abstract

This dissertation investigates three principal areas regarding the dynamics and control of nonlinear systems: averaging theory, controllability of mechanical systems, and control of underactuated nonlinear systems. The most effective stabilizing controllers for underactuated nonlinear systems are time-periodic, which leads to the study of averaging theory for understanding the nonlinear effect generated by resonant oscillatory inputs.

The research on averaging theory generalizes averaging theory to arbitrary order by synthesizing series expansion methods for nonlinear time-varying vector fields and their flows with nonlinear Floquet theory. It is shown that classical averaging theory is the application of perturbation methods in conjunction with nonlinear Floquet theory. Many known properties and consequences of averaging theory are placed within a single framework.

The generalized averaging theory is merged with controllability analysis of underactuated nonlinear systems to derive exponentially stabilizing controllers. Although small-time local controllability (STLC) is easily demonstrated for driftless systems via the Lie algebra rank condition, STLC for systems with drift is more complicated. This thesis exploits notions of geometric homogeneity to show that STLC results for a large class of mechanical systems with drift can be recovered by considering a class of nonlinear dynamical systems satisfying certain homogeneity conditions. These theorems generalize the controllability results for simple mechanical control systems found in Lewis and Murray [85]. Most nonlinear controllability results for classes of mechanical systems may be obtained using these methods.

The stabilizing controllers derived using the generalized averaging theory and STLC analysis can be used to stabilize both systems with and without drift. Furthermore, they result in a set of tunable gains and oscillatory parameters for modification and improvement of the feedback strategy. The procedure can not only derive known controllers from the literature, but can also be used to improve them. Examples demonstrate the diversity of controllers constructed using the generalized averaging theory.

This dissertation concludes with a chapter devoted to biomimetic and biomechanical locomotive control systems that have been stabilized using the generalized averaging theory and the controller construction procedure. The locomotive control systems roll, wriggle, swim, and walk, demonstrating the universal nature of the control strategy proposed.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The analysis of time-varying or time-periodic nonlinear systems is challenging, requiring a degree of sophistication not needed for autonomous nonlinear systems. Most analysis methods for time-varying nonlinear dynamical systems are coordinate-based approaches [161, 17, 50] that do not utilize the power and generality afforded to a description using the language of differential geometry. Such an oversight has important consequences in the area of control of underactuated nonlinear systems due to the increasing prevalence of geometric techniques for the study of nonlinear control systems. This dissertation aims to develop a systematic technique to understand and analyze time-varying and, more specifically, time-periodic nonlinear dynamical systems in a geometric setting.

One immediate outcome is an improved understanding of control techniques for stabilization of underactuated nonlinear control systems. The techniques have practical utility when applied to biomimetic and biomechanical locomotive control systems. Biomimetic and biomechanical locomotive control systems and strategies are becoming increasingly important as engineers seek to push the limits of traditional engineering design. This especially holds in the area of unmanned and autonomous navigation. Biological systems operate in diverse environments, ranging from aquatic and terrestrial to aerial, with some adapted to multiple environments. Within each environment, biological systems can be found operating in different regimes, from the micro-scale to the macro-scale. Often, the biological systems found within a particular environment locomote optimally given their morphology and environmental needs. In spite of the obvious benefits of utilizing biologically inspired locomotive systems, they have yet to be considered practical, standard solutions to the locomotive problem for engineering systems. This failure is in part due to the diversity of locomotive mechanisms, which are difficult to encapsulate into a unified framework, let alone to understand. The references [38, 55, 102, 138, 163] provide a good introduction into the diversity and difficulty of biologically-inspired locomotive strategies. Perhaps the most important reason, to date, is the fact that biomimetic and biomechanical locomotive systems are excellent examples of underactuated nonlinear control systems.

The geometric and mechanical analysis of biomimetic locomotive systems and, more generally, of mechanical control systems, has led to much success regarding the governing equations of motion for such systems [38, 66]. There has been a push to understand the geometry and calculus of mechanical systems, with the goal of creating a general paradigm capable of deriving the governing equations for Lagrangian mechanical systems [15, 29, 30, 82]. Concurrently, research on differential geometry and geometric mechanics has resulted in general and systematic analysis tools for such dynamical systems [98, 99]. The tools for dynamical analysis have then been utilized towards control-theoretic purposes [21, 47, 56, 82, 126]. In spite of these important contributions, the problem of controller synthesis for underactuated nonlinear systems is still an open problem.

Underactuated nonlinear control systems are challenging from many perspectives. Besides being uncontrollable in the linearization, underactuated nonlinear control systems cannot be exponentially stabilized

using continuous state feedback [16, 20], and any exponentially stabilizing solutions are necessarily non-Lipschitz [108]. Combined with the complexity inherent to nonlinear systems, the above results concerning stabilization lead to the realization that construction of exponentially stabilizing controllers for underactuated nonlinear systems is a difficult problem. Support of this statement lies in the fact that few general solutions exist.

It has been shown, however, that time-varying controllers can be used to achieve exponential stabilization [33]. Additional analysis of underactuated nonlinear control systems has further refined the importance of time-varying and time-periodic controllers for locomotion and exponential stabilization [19, 119, 122, 108]. Consequently, the analysis of this thesis will focus on time-varying methods for exponential feedback stabilization of underactuated nonlinear control systems. Importantly, biological systems exhibit time-periodic behavior in their underlying locomotive strategy, meaning that the stabilizing controllers will also apply to such systems.

In addition to the above challenges of underactuated nonlinear control systems, biomimetic and biomechanical locomotive control systems have their unique difficulties when compared to canonical underactuated nonlinear control systems. The majority of nonlinear locomotive control systems examined to date are mechanical systems of the engineer's design. They typically incorporate mechanisms whose actuation directly results in locomotion, e.g., wheels, propellers, jets, etc. Most biological systems do not have such mechanisms at their disposal, and must rely on other strategies for locomotion. More specifically, they must rely on the interaction between their body and the environment within which they locomote. The geometry and mechanics underlying the locomotive strategy results in a system that is not only uncontrollable in the linearization, but is often quite difficult to linearize along along feasible trajectories. For the majority of these systems, the nonlinear coupling of inputs implies that standard linearization techniques will not be capable of capturing the appropriate input-output relationship needed for more refined control. Anholonomy or geometric phase is essential to the basic locomotive strategy of biological systems. Once achieved, nonholonomic effects are again required to obtain controllability in the classical sense (small-time local controllability). Averaging theory will be an indispensible tool for obtaining realizations that are minimally, linearly controllable along feasible trajectories. When viewed from a control-theoretic perspective, these averaged realizations will have multiple control strategies, some of which recover known techniques, and some of which are novel.

While many of the theoretical elements needed for the systematic derivation of a theory for underactuated nonlinear control systems are known, they have not been widely and methodically studied from a practical perspective. Consequently, control of underactuated nonlinear systems remains an endeavour of the theoretician and not a practical tool of the engineer. The work in this thesis can be seen as a large step towards the development of a systematic procedure for understanding the control-theoretic and dynamical properties of underactuated nonlinear systems and, more specifically, of mechanical and biologically-inspired locomotive control systems. Intuitive and systematic controller synthesis is a net result of this improved understanding.

1.1 Review of Relevant Prior Work

The contents of this dissertation build on a large body of literature covering the dynamics and control of nonlinear systems. Prior to describing our contributions and methods, we first review the relevant literature.

1.1.1 Series Expansions

The foundation of this dissertation utilizes series expansion methods to understand nonlinear, time-varying control and dynamical systems. The series expansions methods are rooted in the works of Magnus [96] and

Chen [32]. Agračhev and Gamkrelidze [2] made those series expansions more amenable to nonlinear time-varying differential equations. They also provided a means to better understand the nature of convergence for the series, and to better formulate the series expansions for systematic computation. The resulting series expansion theory for time-varying nonlinear systems has been termed the *chronological calculus*. At heart is an awareness that truncated series expansions have practical use both as an approximation to the true time-varying nonlinear flow, and as a perturbation method for time-varying vector fields.

Within the field of nonlinear control, series expansions have received considerable attention [3, 4]. Kaswki [63, 64, 65] has studied the mathematical structure underlying the series expansions and their connection to controllability and reachability. Leonard and Krishnaprasad have integrated series expansions with averaging theory to construct open-loop controls for drift-free systems evolving on Lie groups [80]. Radford and Burdick [137] have successfully applied these ideas to kinematic nonholonomic systems with symmetry within the context of averaging. Some of the strongest results tying controllability analysis to control design is found in the work of Lafferriere, Liu, and Sussman [76, 154]. They construct extended control systems using the Jacobi-Lie brackets required for controllability, and use ideas similar to averaging in order to prove open-loop trajectory tracking. The extended control systems are based on the work of Kurzweil and Jarník [74, 75], who looked at the nonlinear response of a system to highly oscillatory inputs. An ϵ paremeter is introduced that drives the oscillations to increasing frequency as ϵ tends to the limit $\epsilon \to 0$. In the limit, a new system is derived, whose construction corresponds to an averaging procedure [92, 93].

The framework of the chronological calculus may be utilized to better understand averaging in general. Particularly, how the series expansions fit within the method of averaging, and how they may be used to obtain arbitrary orders of approximation. Sarychev [144, 145] has published in this arena with regards to vibrational control. In [144, 145], important ideas are presented regarding the use of the chronological calculus and nonlinear Floquet theory as a foundation for averaging theory. Due to the scope of the investigations in [144, 145], Sarychev did not examine in detail the consequences of his observations regarding the connection between Floquet theory and averaging. Together with [101], the Sarychev's work can be seen as a step torwards generalizing the vibrational control reasearch of Ballieul, Bullo, Bellman, Bentsman, and Meerkov [7, 8, 9, 10, 11, 23].

1.1.2 Averaging Theory

Sanders and Verhulst [161] give a detailed treatment of the method of averaging and the related theorems that comprise averaging theory. Of particular note, they give formulas for the average of a time-periodic vector field up to second order, as well as theorems concerning the stability of the flow of time-periodic vector fields with regards to stability of the flow due to the averaged vector field. For low-order, averaging is shown to be applicable to time-varying differential equations that do not necessarily exhibit time-periodic behavior. Bogoliubov and Mitropolsky [17] cover the same topics, and also give a general algorithm for calculating higher-order averages. Guckenheimer and Holmes [50] provide a background on stability analysis of time-periodic differential equations via averaging and *Poincaré maps*. The Poincaré map is a powerful tool for proving nonlinear Floquet theory. Sarychev [144, 145] lays down the basic elements required for full development of nonlinear Floquet theory, but does not deeply investigate the consequences due to the scope of his investigation.

Bogoliubov and Mitropolsky were aware that the averaged equations of a time-dependent differential equation gave the Poincaré map, thereby allowing for stability analysis, and also posited that their method was able to recover higher orders of averaging. This method of averaging is known as the Krylov-Bogoliubov-Mitropolsky (KBM) method of averaging. The higher-order methods proposed by Bogoliubov and Mitropolsky have been studied and extended by several researchers as they are the most general and most powerful of the known averaging methods. Perko [132] was able to obtain higher-order averaged expansions for periodic and quasi-periodic differential equations. Higher-order averages are obtained by building upon

the lower-order averages in a systematic manner. Ellison et al. [42] have taken the basic algorithm and developed an improved N^{th}-order estimate for higher-order averaging theory, in the process demonstrating how averaging theory may be used to prove the existence of flows of time-dependent differential equations. A follow-up by Sáenz [143] extended averaging theory to non-periodic systems. Montgomery [111] examined the stability of time-periodic differential equations, by examining the stabillity of the averaged differential equations. Leung and Zhang [81] have demonstrated the link between higher-order averaging and the Poincaré normal form for oscillatory systems. This dissertation will demonstrate how the chronological calculus may be used to obtain the above results, all within the same framework. The use of nonlinear Floquet theory as a cornerstone of averaging theory will also demonstrate why averaging and Poincaré normal forms should be related.

While not immediately addressed by this thesis, it should be noted that averaging theory has been extended in many directions. For example, Lehman and Weibel study the use of averaging theory for control of systems with time delay [79]. In Verduyn-Lunel and Hale [52], averaging theory is extended to infinite-dimensional spaces. Averaging theory also includes two-timing methods, which involve fast and slow timescale; only the fast time scale is averaged over. In many cases averaging over multiple dimensions will introduce resonance [162]. There also exist other papers detailing higher-order averaing theory, however they focus on particular classes of time-periodic systems. In Yagasaki and Ichikawa [164], a computational algorithm is given for weakly nonlinear time-periodic systems. In this dissertation, a general averaging strategy is sought.

Averaging theory in its most general form applies to a larger class of time-dependent vector fields, not necessarily time periodic. Although part of the analysis will focus on strictly time-periodic vector fields, a large portion of the concepts developed herein do not require this property. For the flows of vector fields, it will be possible to obtain a series expansion representing the flow or the average of the vector field in question. The classical first- and second-order averaging theorems are then appropriately truncated versions of the series expansions. Time-periodicity allows for stability analysis using the autonomous averaged vector field alone.

1.1.3 Controllability of Mechanical and Nonlinear Systems

In order to provide a general strategy for stabilization of systems with drift, a general, working definition of controllability must exist from which to build upon. The work in this dissertation touches upon issues regarding the description of mechanical systems, as well as nonlinear controllability tests for mechanical systems.

Mechanical Systems. In order to address specific control issues, prior work has proposed a large number of specialized descriptions for mechanical systems. Baillieul [7, 8] begins with a very general class of mechanical systems, but then restricts elements of the description to be polynomial in the velocities. Morgansen and Brockett [112] consider classes of second-order nonlinear control systems consisting of interconnected blocks, each of which has a very simple structure. Using the structure of the blocks and interconnections, a controllability test is found for the various possibilities. The most general block is excluded from the analysis.

Using ideas from geometric mechanics and control theory, Murray [121] reviewed advances in nonlinear control for previously studied classes of mechanical systems: simple mechanical systems [85], principal kinematic systems [67], and differentially flat systems [45]. Investigations by Bloch, Reyhanoglu and others [16, 139, 141] look at specific forms of nonholonomically constrained mechanical systems and related controllability issues for low-order Jacobi-Lie brackets. The inherent structure of mechanical systems plays an important role in reducing the complexity of the associated control analyses. Sontag and Sussmann [148]

examine the simplifying structure of mechanical systems with respect to manipulators. They also suggest that geometric homogeneity is an important factor to consider.

One class of mechanical systems that generalizes much of the above work is the class of *simple mechanical systems* (SMS). Simple mechanical systems are Lagrangian systems with kinetic and potential energy only. Bullo and Lewis have adroitly used the properties of affine connections to obtain control-theoretic results for this class of systems. However, there are many useful mechanical systems that do not fit into the SMS framework. The papers by Bullo and Cortés et al. [23, 34] consider some possible extensions of the affine connection framework to variations on the SMS class.

Simple mechanical systems which possess Lie group symmetries and the associated process of reduction form an important subclass of simple mechanical systems. In Cortés et al. [35], a complicated structure termed the "connection within a connection" was studied. The mixture of the affine connection and principle bundle connection nomenclature from the two different mathematical perspectives leads to a complex theory that provides the equivalent configuration controllability results for the reduced structures. The approach to nonholonomic mechanical systems with symmetries found in Bloch et al. [15] and more recently in Cendra et al. [29, 30] emphasizes the principal bundle part by explicitly deriving the equations of motion from a reduced perspective, avoiding the "connection within a connection." Recent controllability analysis for such nonholonomic mechanical systems with symmetry by Ostrowski and Burdick [125, 126] suggests that much of the same controllability structure exists without explicitly requiring the affine connection approach. One could also view the work of Kelly and Murray [66, 67] as demonstrating configuration controllability results for principal kinematic systems.

Controllability and Nonlinear Systems. Sussmann's work [152] on sufficient conditions for small-time local controllability forms the backbone of many subsequent analyses of mechanical system controllability. This work introduced two key concepts that we also build upon: the use of free Lie algebras and exponential series representations of flows. The use of the free Lie algebra formalism clarifies the algebraic properties that are inherent in the nonlinear controllability problem. The exponential series representation allows for a systematic means to analyze and compute the flow of a nonlinear system. More recent work by Kawski and Sussmann has expanded upon these ideas [63, 65].

Many prior investigations on the controllability of mechanical systems base their analyses on Sussmann's original work. For example, Bloch and Reyhanoglu et al. [16, 139, 141] examine mechanical systems that are controllable with lower order Lie bracketing. Their analysis is limited to what corresponds to, in this dissertation, the first level symmetric product and the first and second level Lie brackets. Morgansen and Brockett [112, 113, 114] generalize this work to a limited class of systems and still only examine Lie brackets to finite order. The contributions are nevertheless important for they detail the relevant Lie brackets to first examine for systems controllable to low order.

The conditions provided by Sussmann are only sufficient. Furthermore, the small-time local controllability conditions are only applicable in the vicinity of a fixed point. Bridging the gap between necessary and sufficient has been the focus of many investigations [59, 60, 13]. Bianchini and Stefani [12, 14] have extended Sussmann's original results to incorporate controllability away from an equilibrium point. More precisely, Bianchini and Stefani examine controllability along a reference trajectory. This is not the primary goal of this disseration, although suitable extensions may be realized with additional work.

Many of the controllability results summarized above do not take advantage of any inherent structure in the governing equations, although some do rely on the structure obtained by restricting to special cases of mechanical systems. However, when general structure can be shown to exist, it may simplify the properties of the free Lie algebras computed by Sussmann [152]. This is the premise behind the configuration controllability analysis of Lewis and Murray [85, 86] for the case of simple mechanical systems. Subsequent work by Lewis showed how mechanical constraints may be included into the analysis without destroying

the underlying geometric structure [84]. Furthermore Cortés et al. [35] demonstrate how symmetries may lead to configuration controllability tests using reduced structures. Moving beyond the simple mechanical system paradigm is a challenge as the simple mechanical system formalism does not easily incorporate a number of effects that commonly arise in practical mechanical systems. Cortés et al. [34] have shown that the addition of isotropic dissipation to the equations of motion does not severely affect the configuration controllability tests. However, they do not address the common case of configuration dependent dissipation.

Instead of assuming a specific class of mechanical systems, this dissertation focuses on the basic properties of geometric homogeneity, which hold for the great majority of mechanical systems that have been studied to date, including all of the mechanical systems referenced above. The emphasis is not on the derivation of the equations of motion, but on universal properties common to the equations of motion for all mechanical systems. Our results will be compatible with additional mechanical structure such as an affine connection, Lie group symmetries, or nonholonomic constraints, and will allow contemplation of additional elements, such as dissipation, not considered in prior work on the formal control theoretic properties of mechanical systems.

1.1.4 Control of Underactuated Nonlinear Systems

We unite recent results in our generalized averaging theory with nonlinear controllability analysis to synthesize the new control methods for underactuated nonlinear systems. Some of the strongest results tying controllability analysis to control design can be found in the work of Sussmann and Liu [153, 154, 155, 92, 93]. Building upon the work of Kurzweil and Jarnik [74, 75], they studied the limiting behavior of highly oscillatory inputs to driftless control systems. The limiting process leads to a fully controllable autonomous system. Sussmann and Liu are able to demonstrate that, in the limit, the flow of the fully controllable autonomous system is faithful to the flow of the actual underactuated time-varying system. Although the contributions have been enormously important in furthering control for underactuated nonlinear systems, the work is limited to open-loop analysis of driftless systems.

In general, the analysis of open-loop response to sinusoidal and oscillatory inputs has been well studied [122, 19]. Methods for specific classes of control systems have also been developed. Teel et al. [157] have demonstrated how sinusoids may be used to obtain stabilizing controllers for the restricted class of chained form systems. Leonard and Krishnaprasad have used averaging methods, together with series expansions, to approximate the flow of driftless systems evolving on Lie groups [80]. The averaged response is computed up to third order, and is used for open-loop motion planning. Lafferrierre and Sussmann [76, 77] proposed an open-loop path planning methodology for driftless systems based on the averaging process of Sussmann and Liu, which requires highly oscillatory control inputs. Motions planned for the fully controlled averaged system are transformed to oscillatory control inputs for the actual control system.

The closed loop problem for time-varying control of general underactuated driftless systems is difficult and few general solutions exist. Generically, smooth state-feedback will not exponentially stabilize these systems [20], leading to the use of time-varying or non-smooth feedback techniques. Coron [33] has demonstrated that time-varying controllers can asymptotically stabilize underactuated nonlinear control systems. Work on non-smooth stabilizing controllers has varied from continuous asymptotically stabilizing controllers to discontinuous and hybrid controllers; see [71] and references therein. The focus of this thesis is on deriving exponentially stabilizing controllers, for which the time-varying approach appears most promising.

Driftless Systems. One of most general methods with strong results utilizes homogeneous approximations of underactuated driftless control systems. The resulting controllers are exponentially stabilizing with regards to a homogeneous norm, and are termed ρ-exponential stabilizing controllers. In M'Closkey and Murray [109], a method to transform asymptotically stabilized homogeneous systems into ρ-exponentially

stabilized homogeneous systems is given. Although the method is quite general, it does require a pre-existing controller; the pre-existing controller is generally constructed using a method that does not result in exponential stabilization. In Pomet [136], homogeneous controllers for driftless systems are algorithmically constructed by solving a linear partial differential equation. The method is quite general, though often difficult to apply in practice. Pomet addresses how additional problem-specific heuristics may streamline the controller design process and improve the resulting controller. Further work by M'Closkey [108] extended Pomet's algorithm for computing ρ-exponentially stabilizing feedbacks, thereby avoiding the *a priori* assumptions of M'Closkey and Murray [109].

The open-loop path planning ideas of Lafferrierre, Liu, and Sussmann were successfully extended to obtain closed loop ρ-exponentially stabilizing controllers in Morin et al. [117]. Feedback solutions requiring the use of homogeneous approximations were found for stabilization, however the algorithm is computationally difficult to apply. Additional investigation by Morin and Samson has sought to better understand the properties of the time-periodic control inputs required for stabilization [119], and how to derive a constructive controller method that is proveably robust [118]. A drawback to following the approach pioneered by Sussman and Liu is the requirement of highly oscillatory inputs. The oscillatory controllers are designed with a frequency inversely proportional to an ϵ parameter. The averaging process requires for $\epsilon \to 0$, leading to highly oscillatory inputs, although in practice it is understood that the method may work for ϵ small. The generalized averaging methods used in this dissertation build upon the work of Agrachev and Gamkrelidze [2], who consider the effect of a small ϵ parameter, as opposed to an infinitesimally small ϵ parameter that tends to zero. Estimates for the error due to a truncatation of the infinite series expansion, for ϵ small, may be obtained [2]. Consequently, the need to approach the limit $\epsilon \to 0$ is avoided, while it is still possible to obtain stabilization with no error. The results found in the aforementioned citations of Sussmann, Liu, Morin, and Samson can be given a new interpretation by examining them using the generalized averaging theory [92, 93, 118, 119, 117, 153, 154, 155].

In a technique reminiscent of the approach found herein, Struemper and Krishnaprasad [151] utilize Floquet theory to prove linear stability of systems evolving on simple Lie groups. The control inputs, found through approximate inversion, are thereby shown to be exponentially stabilizing. Additional control methods exist, however, these solutions are largely found in specific application domains. Canudas de Wit and Sordalen [28, 41] give exponentially stabilizing controllers for mobile robots or underwater vehicles evolving on the Lie groups $SE(2)$ and $SE(3)$, respectively. Sørensen and Egeland discuss exponential stabilization of chained systems [149]. The focus of this dissertation is on a general and systematic methodology.

Systems with Drift Similar to the case of driftless systems, moving from analysis to the design of stabilizing feedback laws for systems with drift is a challenge. There has been recent success on the use of motion control algorithms and series expansion methods to obtain (exponentially) stabilizing control laws for simple mechanical systems. Bullo [22] has developed series expansions for the flows of simple mechanical systems. Further research led to a constructive algorithm for computing the series expansions and approximate inversions under kinematic actuation of systems with drift [21, 24]. Martínez and Cortés [100] have extended these initial results to arbitrary motions, both in the series expansions and in the necessary approximate inversions. One important element of analysis is missing: although time-periodic inputs are required to generate many of the motion primitives, averaging theory was not explicitly used. By understanding the role of averaging within this context, we can make stronger conclusions regarding the motion control algorithms. Therefore, elements of this thesis can be seen as an extension of the recent work on motion control algorithms for simple mechanical systems [21, 22, 24, 100].

Given the difficulties associated with the control of underactuated nonlinear dynamical systems with drift, many efforts for feedback stabilization of systems with drift tend to focus on specific canonical control

forms, [16, 112], which are actually special instances of simple mechanical or 1-homogeneous control systems. Morgansen and Brockett [112, 113] consider trajectory tracking, via approximate inversion, of specific classes of second-order nonlinear control systems consisting of interconnected blocks, each of which has a very simple structure (e.g., an integrator block and a canonical nonholonomic system block). Work by Reyhanoglu et al. [16, 140, 141] has focused on mechanical systems whose dynamic structure is equivalent to the canonical forms examined by Morgansen and Brockett. The papers emphasize controllability and stabilizability calculations over general constructive controllability for the class of mechanical systems with nonholonomic constraints. Additional investigations have focused on specific subclasses of mechanical systems: Reyhanoglu et al. [139] construct a discontinuous controller for systems underactuated by one degree, and Kolmanovsky and McClamroch [72] study open-loop planning from controllability analysis of nonholonomic mechanical systems. In [72], averaging methods were used to obtain approximations to time-periodic open-loop flows.

The closed loop problem of feedback stabilization of underactuated nonlinear systems with drift is difficult and few general solutions exist. Many of the solutions found for driftless systems fail to extend to the case of systems with drift. One of the few investigations into exponential stabilization of underactuated systems that has been extended from driftless systems to systems with drift is [107]. In [107], M'Closkey and Morin provide two methods for constructing stabilizing controllers, which are exponentially stabilizing with respect to a homogeneous norm. The first method requires a pre-existing stabilizing controller and a Lyapunov function, which is subsequently improved upon to achieve exponential convergence. The second method uses lower-order averaging methods, and is an extension of prior work for stabilization of driftless systems [117]. The averaging methods are based on the work of Sussmann and Liu involving limit of highly oscillatory functions. The controller design often utilizes problem-specific intuition for simplification or transformation of the system with drift into a driftless system.

We note that Floquet theory and averaging have recently been applied to the problem of stabilizing nonlinear systems. However, many of these methods have been restricted to special application domains. For example, [156] utilize feedback techniques and Floquet analysis to stabilize a free joint manipulator. There are additional stabilization techniques to be found, however they, again, are typically restricted to special application domains, or do not result in exponential stabilization. In contrast to these specialized applications, our methods are quite general.

1.2 Contributions of This Thesis

The contributions of this dissertation are multiple. In order to derive a paradigm for understanding nonlinear and biologically inspired locomotive control systems, several distinct theories need to be synthesized in order to construct a framework capable of encompassing all known results. In particular, the emphasis of this dissertation is on providing coherent frameworks to: averaging theory, controllability theory for mechanical systems, and nonlinear control theory of underactuated systems.

Averaging Theory The goal of this disseration, with regards to averaging theory, is to synthesize many of the ideas of classical averaging theory and to develop an intuitively appealing and easily implementable averaging theory for the analysis of time-dependent (time-periodic) nonlinear ordinary differential equations. In doing so, a new approach to averaging theory is developed that encompasses the classical results of Sanders and Verhulst [161], and Bogoliubov and Mitropolsky [17]. It is shown that averaging theory, as classically studied, is derived from nonlinear Floquet theory and perturbation theory. The approach found in this thesis can be seen as complementary to that found in the literature [17, 132], as the analysis involves two important distinctions, (1) there are strict requirements on the vector fields, i.e., smoothness, and (2) the proximity results are conservative. Both can be improved via a more detailed analysis. Nevertheless, the im-

portant elements comprising averaging theory are placed within a single framework. Calculations are made for up to fourth-order averaging. A general algorithm and philosophy are given to aid in the calculation of higher orders of averaging.

Controllability of Mechanical Systems. Controllability for underactuated driftless nonlinear systems typically revolves around the Lie algebra rank condition (LARC). For systems without drift satisfaction of the LARC implies that the control system is small-time locally controllable (STLC), in principal. Underactuated nonlinear systems with drift are more complicated since satisfaction of the LARC does not necessarily imply STLC [152]. This dissertation exploits geometric homogeneity to develop controllability results for a large class of mechanical systems with drift. In particular, we generalize the results of Lewis and Murray [85] to an arbitrary vector bundle. By choosing the vector bundle to be the tangent bundle, the homogeneous structure implies the existence of an affine connection and our results collapse to those found in Lewis and Murray [85]. Selecting the cotangent bundle, one may obtain equivalent controllability results for Hamiltonian systems. Furthermore, geometric homogeneity is preserved under Lie group symmetry reduction, so reduced systems can utilize the same properties for controllability analysis. Thus, our results tie together and extend much prior work in the literature.

No attempt is made to examine the subtle issue of necessary and sufficient conditions. Instead we analyze the interplay between geometric homogeneity of vector fields and the free Lie algebras that are used for controllability analysis. Geometric homogeneity may serve to prune the search process and to provide a heuristic search for "preferable" Jacobi-Lie brackets.

Control of Underactuated Nonlinear Systems. General strategies for feedback stabilization of underactuated driftless nonlinear control systems using time-varying methods utilize many of the same ideas: homogeneity, averaging, and series expansions. The goal of this dissertation is to synthesize these ideas in a coherent fashion; in the process, we develop practical, systematic, and intuitive feedback control strategies for stabilization of underactuated nonlinear control systems. The strategy works for driftless systems and for systems with drift. Importantly, the algorithm does not use nonlinear transformations of state to canonical control forms and does not require a homogeneous norm for proof of exponential stability. Neither are there *a priori* requirements, such as pre-existing controllers or Lyapunov functions, although such knowledge may be helpful.

Linear state-space control is dependent on understanding the exponential representation of flows. Nonlinear time-varying control systems can also be given an exponential representation, meaning that all of the intuition and analysis from linear control theory may provide the control engineer with the needed background to construct and analyze stabilizing controllers for nonlinear systems. Consequently, the main contribution of this thesis, vis-à-vis its control-theoretic goals, is a systematic and intuitive procedure for the design and analysis of stabilizing feedback controllers for underactuated nonlinear systems. The technique will culminate in a set of control functions with tunable gains and oscillatory parameters that can be well understood using concepts from linear control theory and the generalized averaging theory. The systematic nature of the calculations easily lends itself to algorithmic computation.

The material here gives calculations for up to fourth-order averaging for driftless systems, and up to third-order for systems with drift. The structure of the controller synthesis derivation results in a general algorithm which may be followed to aid in the calculation of feedback controller for higher levels of averaging. By following the procedure implicitly given, it is possible to construct a stabilizing controller without detailed understanding of the underlying mathematics of our averaging methods.

1.3 Outline of Dissertation

The thesis is separated into five main chapters and one concluding chapter. Each of the chapters builds upon an extensive literature, some of which the reader is presumed to know for the sake of brevity of exposition. In particular, the reader is assumed to have working knowledge of differential geometry and geometric mechanics [18, 69, 70, 99]. Mastery of the Riemannian and Lagrangian mechanical techniques in addition to the theory of Lie group reduction, both within the context of the control of nonlinear mechanical systems, is also assumed [15, 21, 29, 30, 82, 126].

Chapter 2. Chapter 2 details the generalized averaging theory. Classical averaging theory is first reviewed for presentation of the basic components to which the generalized averaging theory must remain faithful. The exponential representation of nonlinear time-varying flows is discussed and linear Floquet theory is extended to nonlinear systems. Nonlinear Floquet theory demonstrates that time-periodic systems may be represented exactly by a periodic mapping and the flow of an autonomous vector field. The autonomous vector field is the exact average of the time-periodic system. Unfortunately, a closed form solution for the autonomous averaged vector field is generally difficult to compute, necessitating the introduction of perturbation methods. The perturbation methods provide a way to systematically determine the dominant components and the ignorable, perturbative components of an infinite series approximation. In this manner, it is shown that averaging theory is the synthesis of two distinct theories: (1) nonlinear Floquet theory, and (2) perturbation theory. All known results concerning averaging theory can now be systematically proven and analyzed using these two theories in conjunction.

Chapter 3. In Chapter 3, the problem of controller design for underactuated driftless nonlinear systems is covered. Using the generalized averaging theory, it is possible to derive an algorithm for the construction of exponentially stabilizing controllers. The algorithm consists of several simple steps. In some cases, existing controllers are enhanced by recasting them within the framework of the generalized averaging theory. Most controller design methods found in the literature do not result in a set of tunable control parameters, and the appropriate gains must be found by chance. In contrast, the method found herein results in controllers with tunable control gains. The modified controllers demonstrate how the generalized averaging theory may replace the fixed controllers with tunable parameters for performance improvement.

Chapter 4. Prior to extending the results of Chapter 3 to the case of systems with drift, a general description of systems with drift is required. By examining the structure of most nonlinear control systems with drift studied in the literature, a new class of nonlinear control systems can be defined. In Chapter 4 the notion of a *1-homogeneous control system* is introduced. Chapter 4 then concerns itself with the controllability analysis of 1-homogeneous control systems. The idea of geometric homogeneity so critical to the control of underactuated driftless nonlinear control systems is extended to the case of systems with drift. Geometric homogeneity is shown to be a powerful analysis tool for mechanical systems with drift. Since 1-homogeneous systems are an abstraction of simple mechanical systems, the notion of configuration controllability is generalized beyond the simple mechanical systems framework.

Chapter 5. Given the general controllability results of Chapter 4, it is possible to extend the controller synthesis methods of Chapter 3 to 1-homogeneous control systems with drift. The same procedure is developed with attention paid to the particular structures that appear for systems with drift. The net result is also a constructive method for obtaining stabilizing controllers with tunable control parameters.

Chapter 6. Although many investigations have considered how to construct stabilizing feedback control laws for nonlinear systems, it is rare that the control design be applied to more than the token nonlinear control problems. To demonstrate the assertion that the technique discussed in this dissertation is truly of a general nature, a complete chapter, Chapter 6, is devoted solely to the design of locomotive controllers for many distinct biomimetic control systems. The biomimetic control systems vary from those without drift to those with drift, and even include systems whose natural dynamics involve discontinuous evolution. The biomimetic control systems swim, walk, wriggle, and roll, demonstrating the universal nature of the control strategy found in this dissertation. The results are not surprising given the common property of time-periodic actuation found in most biological locomotive systems.

Chapter 7. Finally, Chapter 7, concludes with a summary of the results from the perspectives of nonlinear, and biomimetic and biomechanical control systems. Future research that may be synthesized into a unified control theory for biomimetic locomotive systems is discussed. The research contained in each Chapter may also be furthered, although it does not immediately apply towards the specific context and goals of this dissertation. Therefore, at the end of each Chapter can be found specific comments on possible extensions of the analysis found in the chapter.

Chapter 2

A Generalized Averaging Theory

The method of averaging provides a useful means to study the behavior of nonlinear dynamical systems under periodic forcing. The method of averaging has a long history, which, with its canon of related theorems, has been sufficiently developed into an averaging theory. In spite of its maturity, many of the elements of this theory have not been cast into a comprehensive framework. The goal of this chapter is to provide a more coherent structure to the theory of averaging and in the process consolidate many results that currently appear to stand alone.

This dissertation highlights the use of series expansions of time-varying vector fields. An important tool in the development of this averaging theory, series expansions approximate time-varying vector fields by infinite series' of autonomous vector fields. The series expansions provide alternative expressions for flows or vector fields, and are an important means for proving results in nonlinear Floquet theory. It is shown that Floquet theory places time-periodic vector fields into an autonomous averaged form. For systems that cannot be analyzed in closed form using nonlinear Floquet theory, the use of perturbation methods in conjunction with series expansion methods will lead to an approximate averaging theory, which is how averaging is typically applied. As a consequence, we demonstrate that averaging theory is really the synthesis of two distinct mathematical theories: Floquet theory and perturbation theory. All known results of classical averaging theory can be better understood when viewed from this perspective. This approach also allows for the systematic development of higher-order averaging methods.

Organization of this chapter. Section 2.1 reviews relevant results from classical averaging theory that are subsequently generalized. Section 2.2 contains a synopsis of the chronological calculus as derived by Agrachev and Gamkrelidze [2]. It reviews relevant results regarding the exponential representation of flows, and the approximation of flows and vector fields by series expansions. These expansions are key to the development of our generalized averaging theory via nonlinear Floquet theory in Section 2.3. Section 2.4 shows how to calculate averages to higher-order and also points out the connection between these formulas and those of classical averaging theory. The implications of the averaging theory are detailed for a class of periodically forced systems in Section 2.5. Several examples of dynamical systems that can be analyzed using averaging theory are given in section 2.6. One of the examples demonstrates how averaging via Floquet theory may result in an exact average. The example is important because it shows that averaging theory may be exact; a common misunderstanding is that averaging theory will always result in some quantifiable error. The applicability and exactness of averaging theory is certainly an important issue. Section 2.7 discusses related literature concerning the approximation of time-varying vector fields by autonomous vector fields. Section 2.8 concludes with a synopsis of the current results and ideas for future work.

2.1 Classical Averaging Theory

Before developing the elements of the generalized averaging theory, it is important to understand the basic structure and theorems of classical averaging theory. This section reviews the key ideas that will subsequently be generalized.

The standard form of the equations of motion for averaging are

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \epsilon X(x,t), \quad x(0) = x_0, \tag{2.1}$$

where X is T-periodic, i.e., $X(x,t) = X(x,t+T)$, and $x \in M \subset \mathbb{R}^n$. The average of X is typically given as

$$\overline{X(x,t)} = \int_t^{t+T} X(x,\tau)\,\mathrm{d}\tau\,, \tag{2.2}$$

where the evaluation point x is considered fixed. Often, $\overline{X(\cdot,t)}$ will be written as $\overline{X}(\cdot)$[1]. The average then defines new autonomous equations of motion,

$$\frac{\mathrm{d}y}{\mathrm{d}t} = \epsilon\overline{X}(y), \quad y(0) = x_0. \tag{2.3}$$

The goals of averaging theory are multiple. The primary goal is to determine conditions under which the two flows coincide and to what degree they coincide. The parameter, ϵ, will provide a means to determine this coincidence.

Theorem 1 (first-order averaging) *[161] Consider the initial value problems (2.1) and (2.3) with $x, y, x_0 \in M \subset \mathbb{R}^n,\ t \in [t_0, \infty),\ \epsilon \in (0, \epsilon_0]$. Suppose that the following all hold,*

1. *$X(x,t)$ is Lipschitz-continuous in x on M, $t \geq 0$, continuous in x and t on $M \times \mathbb{R}^+$, and*
2. *$y(t)$ belongs to an interior subset of M on the timescale $\frac{1}{\epsilon}$.*

Then,

$$x(t) - y(t) = O(\epsilon)$$

as $\epsilon \downarrow 0$ on the timescale $\frac{1}{\epsilon}$.

Theorem 1 determines the conditions under which the flow of the autonomous vector field (2.3) remains close to the flow of the original vector field (2.1). Proximity holds on a finite timescale, and few additional results can be derived unless additional constraints are placed on the vector fields.

Theorem 2 *[161] Assume that in addition to the conditions of Theorem 1, the following are met:*

1. *$y = 0$ is an asymptotically stable fixed point.*
2. *$\overline{X}$ is continuously differentiable in M, and has a domain of attraction $M^* \subset M$.*

If $x_0 \in M^*$,then,

$$x(t) - y(t) = O(\delta(\epsilon))\,, \quad 0 \leq t < \infty$$

with $\delta(\epsilon) = o(1)$.

[1] Via a change of coordinates, the average can instead be written: $\overline{X(x,t)} = \int_0^T X(x,\theta)\,\mathrm{d}\theta$. The change of coordinates only holds in the time-periodic case.

An asymptotically stable fixed point in the average (2.1) renders the averaged approximation valid for all time. For a dynamical system satisfying Theorem 2, there are important implications: the restriction of a flow to a neighborhood of a fixed point implies an orbit.

Theorem 3 *[50] Consider the mapping,*

$$y(t) = z(t) + \epsilon w(z(t), t)$$

and the initial value problem

$$\dot{z} = \epsilon \overline{X}(z) + \epsilon^2 Y(z, t),$$

where

$$\begin{aligned} w(x,t) &= \int_0^t \left(X(x,\tau) - \overline{X}(x) \right) \, \mathrm{d}\tau \\ Y(x,t) &= DX(x,t) \cdot w(x,t) - Dw(x,t) \cdot \overline{X}(x). \end{aligned}$$

Suppose that $X(x,t)$ is C^r, $r \geq 2$, and bounded on bounded sets. Then the following statements hold:

1. *$|x(t) - z(t)| = O(\epsilon)$ on the timescale $\frac{1}{\epsilon}$.*
2. *If z^* is a hyperbolic fixed point of (2.3), then there exists an ϵ_0 such that for all $\epsilon \in (0, \epsilon_0]$, (2.1) possesses a unique hyperbolic periodic orbit $\gamma_\epsilon(t) = z^* + O(\epsilon)$ of the same stability type as z^*.*
3. *If $x^s(t)$ is a solution of (2.1) lying in the stable manifold of the hyperbolic periodict orbit γ_ϵ, and $z^s(t)$ is a solution of (2.3) lying in the stable manifold of the hyperbolic fixed point z^*, then $|x^s(t) - z^s(t)| = O(\epsilon)$ for $t \in [0, \infty)$. Similar results apply to solutions lying in the unstable manifolds on the time interval $t \in (-\infty, 0]$.*

The proof of the portion of this theorem related to stability is based on the use of the average as a Poincaré map of the actual flow. Theorem 3 does not preclude the case where the orbit is degenerate, i.e., that it is in fact the point z^*.

Theorem 4 *[68] Assume that the conditions of Theorem 3 hold and that both X and its average $\overline{X}$ share the same equilibrium point. If the equilibrium point is exponentially stable for the averaged system, then the equiblirium point is exponentially stable for the original system.*

First-order averaging is not always sufficient to approximate the dynamics of a system. In these cases, second- or higher-order averaging techniques are needed.

Theorem 5 (second-order averaging) *[161] Consider the mapping,*

$$y(t) = z(t) + \epsilon w(z(t), t)$$

and the initial value problem

$$\dot{z} = \epsilon \overline{X}(z) + \epsilon^2 \overline{Y}(z),$$

where

$$\begin{aligned} w(x,t) &= \int_0^t \left(X(x,\tau) - \overline{X}(x) \right) \, \mathrm{d}\tau + a(x) \\ Y(x,t) &= DX(x,t) \cdot w(x,t) - Dw(x,t) \cdot \overline{X}(x), \end{aligned}$$

and $a(x)$ is chosen so that the time average of $w(x,t)$ vanishes. Suppose that the following all hold,

1. *$X(x,t)$ has a Lipschitz-continuous first derivative in x and is continuous on its domain of definition, and*

2. *$z(t)$ belongs to an interior subset of M on the timescale $\frac{1}{\epsilon}$.*

Then,

$$x(t) = y(t) + O\left(\epsilon^2\right)$$

on the timescale $\frac{1}{\epsilon}$.

When moving to second- and higher-order averaging methods, Theorems 2, 3, and 4 are no longer applicable. The explicit ϵ dependence of the higher-order terms complicates matters, and care must be taken to ensure that even higher-order approximations do not alter stability [111, 161].

Note that by using the coordinate-based definition of the Jacobi-Lie bracket and the integration of products formula, the time average of $Y(z,t)$ may be rewritten as

$$\overline{Y}(z) = \frac{1}{2}\overline{\left[\int_0^t X(z,\tau)\,\mathrm{d}\tau\,, X(z,t)\right]} + \left[a(z), \overline{X}(z)\right], \tag{2.4}$$

meaning that the second-order term, $\overline{Y}$ from Theorem 5, has an intrinsic description within the space of vector fields and has a mathematical interpretation[2]. This points to the fact that averaging may be a natural operation, and that there must exist a natural procedure to obtain Equation (2.4) without the use of coordinate calculations.

The procedure should be general enough to allow for an arbitrary order of approximation while preserving the intrinsic description of the averages. Although the KBM method of averaging works to arbitrary order, the averages do not explicitly involve Jacobi-Lie brackets. Much like Theorem 5, the KBM method does demonstrate that higher order averaging will consist of two elements,

1. an autonomous vector field approximation of the time-periodic vector field in (2.1), and

2. a time-dependent compensation mapping to approximate the flow of (2.1).

The goal of this investigation is to find an algorithm for the computation of higher-order averages involving Jacobi-Lie brackets. The analysis should allow for the use of Poincaré maps for inference of stability, and should recover the known properties of averaged expansions [17, 42, 111, 132, 143]. Additionally, the derivation should illuminate the mathematical context and role of averaging.

2.2 Evolution of Systems and the Chronological Calculus

This section reviews the important concepts of the *chronological calculus* found in Agrachev and Gamkrelidze [2]. The chronological calculus demonstrates that the exponential representation for flows of linear autonomous differential equations can be applied to nonlinear time-varying differential equations. If the chronological calculus is well understood, many concepts and theorems proven for linear (time-varying) differential equations may hold for nonlinear (time-varying) differential equations with little additional work. Section 2.3 demonstrates this beneficial property by extending linear Floquet theory to the nonlinear setting.

The chronological calculus focuses on a series representation of the exponential flow. The series representation is used to determine the consequences of a finite series approximation to the infinite series representation, thereby developing the notion of perturbation methods for nonlinear dynamical systems. From

[2] See Appendix 2.A for derivation.

there, the series expansions and their related approximations can be used to derive a generalized averaging theory.

The general form for the equations of motion are

$$\frac{\mathrm{d}x}{\mathrm{d}t} = X(x,t), \quad x(0) = x_0, \tag{2.5}$$

where $x \in \mathbb{R}^n$, $t \in \mathbb{R}$, X is in $C^r\left(\mathbb{R}^n, \mathbb{R}^n\right)$ as a function of the state x only, and it is absolutely continuous as a function of time t only. Sometimes we will write X_t for $X(\cdot, t)$. The ultimate goal is to understand the flow, denoted by $\Phi_{0,t}^X$. This will require notions of proximity for flows and functions.

Define the semi-norm of a function $\varphi \in C^r\left(\mathbb{R}^n, \mathbb{R}\right)$,

$$||\varphi||_{s,M} \equiv \sup_{x \in M} \sum_{\alpha=0}^{s} \frac{1}{\alpha!} \sup_{|\chi|=1} \left|\chi^\alpha \varphi(x)\right|, \tag{2.6}$$

where $M \subset \mathbb{R}^n$, $s \in [0, r]$ is an arbitrary integer, and χ is a constant vector field. When the second index is absent, then it is assumed to be $M = \mathbb{R}^n$, i.e., $||\varphi||_s = ||\varphi||_{s,\mathbb{R}^n}$. The semi-norm (2.6) induces the following semi-norm on matrix valued functions $A \in C^r\left(\mathbb{R}^n, \mathbb{R}^{m_1 \times m_2}\right)$,

$$||A||_{s,M} = \sum_{\beta=1}^{m_2} \max_{1 \le \alpha \le m_1} \left|\left|a_\beta^\alpha\right|\right|_{s,M}, \quad A = [a_\beta^\alpha],\ a_\beta^\alpha \in C^r\left(\mathbb{R}^n, \mathbb{R}\right),$$

with the semi-norms on column and row vector valued functions defined as

$$||X||_{s,M} = \max_{1 \le \alpha \le n} ||X||_{s,M},$$
$$||\xi||_{s,M} = \sum_{\beta=1}^{n} ||\xi_\beta||_{s,M},$$

respectively. All topologies, unless specified otherwise, will be defined with respect to the semi-norms $||\cdot||_{s,K}$, where K is an arbitrary compact set in $\mathbb{R}^n$. It is called the *topology of compact convergence with respect to all derivatives* [2]. Hereafter, we will only concern ourselves with smooth mappings of state, i.e., $r = \infty$.

The flow of equation (2.5) satisfies

$$\frac{\mathrm{d}}{\mathrm{d}t}\Phi_{0,t}^X = X_t \circ \Phi_{0,t}^X, \quad \Phi_{\tau,\tau}^X = \mathrm{Id}$$

or in pull-back form,

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\Phi_{0,t}^X\right)^\star = \left(\Phi_{0,t}^X\right)^\star \circ X_t, \quad \left(\Phi_{\tau,\tau}^X\right)^\star = \mathrm{Id}^\star, \tag{2.7}$$

where, on the right-hand side of Equation (2.7), the pull-back is not of vector-fields but of functions, i.e.,

$$\left(\Phi_{0,t}^X\right)^\star \circ X_t = X_t \circ \Phi_{0,t}^X = X(\Phi_{0,t}^X(\cdot), t).$$

The pull-back flow is continuous and invertible [2]. The pull-back of vector fields is given by the adjoint,

$$\mathrm{Ad}\left(\Phi_{0,t}^X\right)^\star \cdot X = \left(\Phi_{0,t}^X\right)^\star \circ X \circ \left(\left(\Phi_{0,t}^X\right)^\star\right)^{-1},$$

where the following identity is in effect,

$$\left(\left(\Phi_{0,t}^{X}\right)^{\star}\right)^{-1} = \left(\left(\Phi_{0,t}^{X}\right)^{-1}\right)^{\star}.$$

In [4], Agrachev and Gamkrelidze conceptualize the extension of their results to manifolds. This is done by providing $C^{\infty}(M)$ with the topology given to $\mathbb{R}^n$ using the semi-norms. We restrict the current investigation to $\mathbb{R}^n$, however, the ideas naturally extend to systems evolving on manifolds.

2.2.1 Series Expansions

Series expansion methods for finding the solution to a differential equation are based on the Volterra series expansion. This is also known as a Picard iteration, which is often used to prove existence of solutions to differential equations. A solution, $\mathcal{F}_{t_0,t}$, to the differential equation in (2.7) is

$$\mathcal{F}_{t_0,t} = \mathrm{Id}^{\star} + \int_{t_0}^{t} \mathcal{F}_{t_0,\tau} \circ X_{\tau} \, \mathrm{d}\tau \,,$$

which via successive substitutions results in the Volterra series,

$$\begin{aligned}\mathcal{F}_{t_0,t} &= \mathrm{Id}^{\star} + \int_{t_0}^{t} X_{\tau} \, \mathrm{d}\tau \, + \int_{t_0}^{t} \mathrm{d}\tau_1 \int_{t_0}^{\tau_1} \mathrm{d}\tau_2 \, X_{\tau_2} \circ X_{\tau_1} + \dots \\ &= \mathrm{Id}^{\star} + \sum_{m=1}^{\infty} \int_{t_0}^{t} \mathrm{d}\tau_1 \int_{t_0}^{\tau_1} \mathrm{d}\tau_2 \dots \int_{t_0}^{\tau_{m-1}} \mathrm{d}\tau_m X_{\tau_m} \circ \dots \circ X_{\tau_1}.\end{aligned}$$

Under the appropriate conditions, the series summation converges and represents the true flow of the system [2]; the analysis of this chapter is restricted to smooth functions, therefore, the convergence conditions are satisfied for finite time. The convergent Volterra series summation can be written as

$$\overrightarrow{\mathcal{V}}_{t_0,t}\left(X_{\tau}\right) = \mathrm{Id}^{\star} + \sum_{m=1}^{\infty} \int_{t_0}^{t} \mathrm{d}\tau_1 \int_{t_0}^{\tau_1} \mathrm{d}\tau_2 \dots \int_{t_0}^{\tau_{m-1}} \mathrm{d}\tau_m X_{\tau_m} \circ \dots \circ X_{\tau_1},$$

where $\overrightarrow{\mathcal{V}}_{t_0,t}\left(X_{\tau}\right)$ is the Volterra series expansion of the flow corresponding to the time-varying vector field X. When the vector fields at different points in time commute, i.e., $[X_{t'}, X_{t''}] = 0\,,\ \forall\, t', t'' \in \mathbb{R}$, then the Volterra series expansion of the flow reduces to the time-independent exponential,

$$\overrightarrow{\mathcal{V}}_{t_0,t}\left(X_{\tau}\right) \mathrm{Id} = \sum_{m=1}^{\infty} \frac{1}{m!} \left(\int_{t_0}^{t} X_{\tau} \mathrm{d}\tau\right)^{m} = \exp\left(\int_{t_0}^{t} X_{\tau} \mathrm{d}\tau\right),$$

recovering the classical notion of an exponential for autonomous ordinary differential equations. In the more general time-dependent case, the convergent Volterra series is still an exponential series. It is, therefore, called the *chronological exponential in* X_{τ}, and is written

$$\mathcal{F}_{t_0,t} \cong \overrightarrow{\mathcal{V}}_{t_0,t}\left(X_{\tau}\right) \equiv \overrightarrow{\exp}\left(\int_{t_0}^{t} X_{\tau} \mathrm{d}\tau\right), \tag{2.8}$$

where the flows in Equation (2.8) are pull-back flows. The first equality in Equation (2.8) is an asymptotic equality as per the proposition below.

Proposition 1 *[2] Let X_t, $t \in \mathbb{R}$, be a bounded (locally integrable) vector field, and $\left(\Phi^X_{t_0,t}\right)^\star$ the flow of X_t. Then, for all φ in $C^\infty(\mathbb{R}^n, \mathbb{R})$, $s \geq 0$, $m \geq 2$, and any compact set $K \subset \mathbb{R}^n$,*

$$\left|\left|\left(\left(\Phi^X_{t_0,t}\right)^\star - \left(\mathrm{Id}^\star + \sum_{\alpha=1}^{m-1} \int_{t_0}^{t}\int_{t_0}^{\tau_1} \dots \int_{t_0}^{\tau_{\alpha-1}} X_{\tau_\alpha} \circ \dots \circ X_{\tau_1}\, \mathrm{d}\tau_\alpha \dots\, \mathrm{d}\tau_2\, \mathrm{d}\tau_1\right)\right)\varphi\right|\right|_{s,K}$$
$$\leq C_1 e^{C_2\left|\int_{t_0}^{t} ||X_\tau||_s\, \mathrm{d}\tau\right|} \left(3n(2s+2m)^{(s+m)}\right)^m \frac{1}{m!}\left|\int_{t_0}^{t} ||X_\tau||_{s+m-1}\, \mathrm{d}\tau\right|^m ||\varphi||_{s+m,M}\,,$$

where

$$C_1 = (1+s)(2ns)^s\left(1+n+\operatorname{diam} K + 2\int_0^t ||X_\tau||_0\, \mathrm{d}\tau\right)^s$$
$$C_2 = 3n(2s+2)^{s+1}C_1,$$

and $M = O_R(K)$ is a neighborhood of radius $R = \int_{t_0}^{t} ||X_\tau||_0\, \mathrm{d}\tau$ of the compact set K.

Given the existence of the chronological exponential it is natural to expect likewise the existence of an inverse to the chronological exponential. It will be called the *chronological logarithm*. This is a mapping from

$$\overrightarrow{\exp}\left(\int_{t_0}^{t} X_\tau\, \mathrm{d}\tau\right) \mapsto X_t,$$

and will be written as

$$\overrightarrow{\log}_{t_0} \overrightarrow{\exp}\left(\int_{t_0}^{t} X_\tau\, \mathrm{d}\tau\right) = X_t, \tag{2.9}$$

as per [2]. The most important use of Equation (2.9) for the purposes of this thesis is in the time-independent case, where it will be called the *logarithm*.

The Logarithm. Averaging theory seeks to find an autonomous vector field approximating the vector field of Equation (2.5). This is equivalent to the question posed by Agrachev and Gamkrelidze regarding the existence of a vector field $\overrightarrow{\mathrm{V}}_{t_0,t}(X_\tau)$, such that the asymptotic equality holds:

$$\overrightarrow{\exp}\int_{t_0}^{t} X_\tau\, \mathrm{d}\tau \cong \exp \overrightarrow{\mathrm{V}}_{t_0,t}(X_\tau)\,, \tag{2.10}$$

where the asymptotic equality uses Proposition 2. In other words, does there exist an autonomous differential equation, given by $\overrightarrow{\mathrm{V}}_{t_0,t}(X_\tau)$, whose flow after unit time is equal to the flow of the non-autonomous differential equation given by X_t at time t? A series expansion for this vector field exists, and assuming convergence [2] it is called the *logarithm*, i.e.,

$$\overrightarrow{\mathrm{V}}_{t_0,t}(X_\tau) \cong \ln \overrightarrow{\exp}\int_{t_0}^{t} Y_\tau\, \mathrm{d}\tau\,. \tag{2.11}$$

The vector field, $\overrightarrow{\mathrm{V}}_{t_0,t}(X_\tau)$, is an autonomous vector field, but is parametrized by time, t. It is the autonomous vector field whose flow after unit time maps to the same point in space that the actual time-dependent flow will at time t. Variation of the final time will result in a new autonomous vector field.

The logarithm is at the heart of a series expansion solution to the approximation of the flow $\Phi^X_{t_0,t}$. The

logarithm vector field is an infinite series consisting of *variations*,

$$\overrightarrow{\mathrm{V}}_{t_0,t}\left(X_\tau\right) = \sum_{m=1}^{\infty} \overrightarrow{\mathrm{V}}_{t_0,t}^{(m)}\left(X_\tau\right), \tag{2.12}$$

where

$$\overrightarrow{\mathrm{V}}_{t_0,t}^{(m)}\left(X_\tau\right) = \int_{t_0}^{t}\int_{t_0}^{\tau_1}\cdots\int_{t_0}^{\tau_{m-1}} \mathrm{v}^{(m)}(X_{\tau_1},\ldots,X_{\tau_m})\,\mathrm{d}\tau_m\ldots\ \mathrm{d}\tau_2\ \mathrm{d}\tau_1 \tag{2.13}$$

is the m^{th}*-variation of the identity flow corresponding to the perturbation field* X_τ[2]. The integrands of the variations in Equation (2.13) are the sum of iterated Jacobi-Lie brackets, and thereby reside within the space of vector fields. An algorithm to compute the integrands can be found in Section 2.2.2.

In order to prove the asymptotic equality of Equation (2.10), Agrachev and Gamkrelidze utilize the following proposition,

Proposition 2 *[2] If*

$$\int_{t_0}^{t} ||X_\tau||_{s+m}\ \mathrm{d}\tau\ \leq 1,$$

then

$$\left|\left|\left(\overrightarrow{\exp}\int_{t_0}^{t} X_\tau\,\mathrm{d}\tau\ -\exp\sum_{\alpha=1}^{m}\overrightarrow{V}_{t_0,t}^{(\alpha)}\left(X_\tau\right)\right)\varphi\right|\right|_{s,K} \leq C_1\left|\int_{t_0}^{t}||X_\tau||_{s+2m}\ \mathrm{d}\tau\right|^{m+1}||\varphi||_{s+m+1,N_{C_2}(K)}$$

where the constants C_1 *and* C_2 *depend only on* s, m, *and* $\operatorname{diam} K$, *and where* $N_{C_2}(K)$ *is a* C_2*-neighborhood of the compact set* K.

Proposition 2 is important when using a finite series expansion of the variations from Equation (2.12) in lieu of the full infinite series expansions. In general, the series $\overrightarrow{\mathrm{V}}_{t_0,t}\left(X_\tau\right)$ diverges, but under certain conditions it can converge. Assume that Δ is a sub-algebra of the space of vector fields, with a norm such that $||[X,Y]|| \leq ||X||\ ||Y||\,,\ \forall\, X,\ Y\ \in \Delta$. This makes Δ a Banach Lie algebra.

Proposition 3 *[2] Let* X *be a time dependent vector field, and let* $X_t \in \Delta$ *for all* $t \in \mathbb{R}$. *If* $\int_{t_0}^{t}||X_\tau||\ \mathrm{d}\tau\ \leq 0.44$, *then* $\overrightarrow{V}_{t_0,t}\left(X_\tau\right)$ *converges absolutely in* Δ.

According to Agrachev and Gamkrelidze [3], S. Vakhrameev gave an improved estimate for convergence. It is

$$\int_{0}^{t}||X_\tau||\ \mathrm{d}\tau\ \leq \int_{0}^{\pi}\frac{2}{\theta(1-\cot(\theta))+2}\,\mathrm{d}\theta\,,$$

and is a sharp estimate.

2.2.2 Elements of the Series Expansion

Calculation of the integrands of the variations (2.13) is not simple; Sarychev [144, 145] has an abbreviated discussion of the derivation procedure. The basic idea is very similar to the expansion described by Magnus [96]. Using the time-dependence of the logarithm together with the identity,

$$\exp\left(\overrightarrow{\mathrm{V}}_{t_0,t}\left(X_\tau\right)\right) = \overrightarrow{\exp}\left(\int_{t_0}^{t} X_\tau\,\mathrm{d}\tau\right), \tag{2.14}$$

it is possible to derive the following equation,

$$X_t = \int_0^1 \exp\left(-\tau \operatorname{ad} \overrightarrow{\mathrm{V}}_{t_0,t}(X_\tau)\right) \mathrm{d}\tau \left(\frac{\mathrm{d}}{\mathrm{d}t} \overrightarrow{\mathrm{V}}_{t_0,t}(X_\tau)\right), \tag{2.15}$$

via time differentiation of Equation (2.14), [2]. The integral on the right hand side of Equation (2.15) can be inverted. Define the following function,

$$\varphi(z) \equiv \frac{1}{\int_0^1 \exp(-\tau z)\ \mathrm{d}\tau} = \frac{-z}{\exp(-z) - 1},$$

which has a series expansion,

$$\varphi(z) = 1 + \frac{1}{2}z + \sum_{\alpha=2}^{\infty} \frac{B_\alpha}{\alpha!} z^\alpha,$$

where the B_α are Bernoulli numbers. Equation (2.15) may be rewritten as

$$\frac{\mathrm{d}}{\mathrm{d}t} \overrightarrow{\mathrm{V}}_{t_0,t}(X_\tau) = \varphi\left(\operatorname{ad} \overrightarrow{\mathrm{V}}_{t_0,t}(X_\tau)\right) \cdot X_t. \tag{2.16}$$

There are several equivalent ways to calculate the series expansion of the right-hand side of Equation (2.16). The one we shall adopt comes from [2]; most others are an attempt to extract more structure from derivation that follows [4, 145, 65]. Define a grammar whose constitutive alphabet is $\{\operatorname{ad}, \xi_1, \ldots, \xi_m\}$.

Definition 1 *[3] A* regular word *in the grammar given by the alphabet* $\{\operatorname{ad}, \xi_1, \ldots, \xi_m\}$ *is a word that, by suitable parenthetication, can be expressed as a commutator polynomial.*

The commutator is defined to be the Jacobi-Lie bracket, here given by the letter ad. Note that this choice of definition of a regular word places a requirement on the number of times that the ad letter must appear in a word. As an example, we have the word below:

$$\operatorname{ad}\xi_1 \operatorname{ad}\xi_2\xi_3 = (\operatorname{ad}\xi_1)((\operatorname{ad}\xi_2)\xi_3) = [\xi_1, [\xi_2, \xi_3]]$$

The set consisting of the words of the grammar form a free associative algebra over $\mathbb{R}$ with generators in the alphabet. The free associative algebra is the same one used by Sussmann [152] to prove his general theorem on local controllability. The calculation of the logarithm vector field will also require the definition of the depth of a letter in a word.

Definition 2 *[3] The* depth *of a letter ξ_i in a word $w = w_1\xi_i w_2$, where w_1 or w_2 may be empty, is the number of regular words that can be made with contiguous subsequences of the letters in $w_1\xi_i$ containing ξ_i in them.*

Example
To see this more clearly two examples of regular words are given below where the regular words that can be formed from subsequences have been delineated.

$$\overline{\operatorname{ad}\underline{\operatorname{ad}\xi_3\xi_2}\xi_1} \quad , \quad \overline{\operatorname{ad}\xi_3\underline{\operatorname{ad}\xi_2\xi_1}}$$

For the first word, the depths of the letters (ξ_1, ξ_2, ξ_3) are $(1, 1, 0)$, respectively. For the second word, the depths of the letters (ξ_1, ξ_2, ξ_3) are $(2, 0, 0)$, respectively.

□

To this algebra can be added a differentiation operator.

Definition 3 *[3] The differentiation operator $\mathbf{D}(w)$ is a linear operator associated with each word w acting according to the following rules on the alphabet:*

1. $\mathbf{D}(w)\,\mathrm{ad} = w\,\mathrm{ad}$
2. $\mathbf{D}(w)\,\xi_i = w\,\xi_i$
3. $\mathbf{D}(w)\,w_1\,w_2 = (\mathbf{D}(w)\,w_1)\,w_2 + w_1\,(\mathbf{D}(w)\,w_2)$ *(Leibniz rule)*

Each variation of the series, Equation (2.13), consists of a linear combination of the $(2m-3)!!$ words[3] formed by the operation

$$\mathbf{D}(\mathrm{ad}\xi_m)\,\ldots\,\mathbf{D}(\mathrm{ad}\xi_2)\,\xi_1 = w_1 + \cdots + w_{(2m-3)!!}. \tag{2.17}$$

Define, $\nu_{i\alpha}$ to be the depth of the letter ξ_i in the word w_α from Equation (2.17). The integrand of the m^{th} variation of the identity flow corresponding to the perturbation field X_t is

$$\mathrm{v}^{(m)}(\xi_1,\ldots,\xi_m) = \sum_{\alpha=1}^{(2m-3)!!}\left(\prod_{i=1}^{m} b_{\nu_{i\alpha}}\right) w_\alpha \tag{2.18}$$

where $b_\nu = (B_\nu/\nu!)$, and the B_ν are Bernoulli numbers. The first three variations are given below:

$$\begin{aligned}
\mathrm{v}^{(1)}(\xi_1) &= \xi_1 \\
\mathrm{v}^{(2)}(\xi_1,\xi_2) &= \frac{1}{2}\mathrm{ad}\xi_2\xi_1 \\
\mathrm{v}^{(3)}(\xi_1,\xi_2,\xi_3) &= \frac{1}{12}\mathrm{ad}\,\xi_3\,\mathrm{ad}\,\xi_2\,\xi_1 + \frac{1}{4}\mathrm{ad}\,\mathrm{ad}\,\xi_3\,\xi_2\,\xi_1 + \frac{1}{12}\mathrm{ad}\,\xi_2\,\mathrm{ad}\,\xi_3\,\xi_1.
\end{aligned} \tag{2.19}$$

By taking advantage of the Jacobi-Lie identity the third variation can be simplified. The same must be done to render the fourth variation tractable.[4]

$$\begin{aligned}
\mathrm{v}^{(3)}(\xi_1,\xi_2,\xi_3) &= \frac{1}{6}\,(\mathrm{ad}\,\xi_3\,\mathrm{ad}\,\xi_2\,\xi_1 + \mathrm{ad}\,\mathrm{ad}\,\xi_3\,\xi_2\,\xi_1) \\
\mathrm{v}^{(4)}(\xi_1,\xi_2,\xi_3,\xi_4) &= -\frac{1}{12}\,(\mathrm{ad}\,\mathrm{ad}\,\xi_4\,\xi_3\,\mathrm{ad}\,\xi_2\,\xi_1 + \mathrm{ad}\,\mathrm{ad}\,\mathrm{ad}\,\xi_4\,\xi_3\,\xi_2\,\xi_1 \\
&\qquad + \mathrm{ad}\,\xi_4\,\mathrm{ad}\,\mathrm{ad}\,\xi_3\,\xi_2\,\xi_1 + \mathrm{ad}\,\xi_3\,\mathrm{ad}\,\mathrm{ad}\,\xi_4\,\xi_2\,\xi_1)
\end{aligned} \tag{2.20}$$

The integrands $\mathrm{v}^{(m)}(\cdot)$ are then integrated to give the m^{th}-variations from Equation (2.13). To see how the integrands of the m^{th}-variations are integrated, refer to Equation (2.33).

2.2.3 Exponential Representation of a Flow and Perturbation Theory

This section is concluded by giving the exponential representation of a flow as described by the chronological calculus, under the addition of a smallness parameter ϵ. The logarithm operator utilizes the flow, $\Phi_{0,t}^{X}$, of the time-varying vector field X to compute an autonomous vector field, Z. Consquently, the flow

[3] This is the recursively defined product $k!! = k\,(k-2)!!$, where $1!! = 1$ and $0!! = 1$.

[4] See Appendix 2.B for calculations.

of the autonomous vector field Z can be used in lieu of flow of the time-varying vector field X, however, the autonomous vector field Z is given by an infinite series. If the vector field Z naturally has the (small) parameter ϵ as a factor, Proposition 2 can be used to determine how close the flow of a finite series expansion for Y is to the actual flow of X. Consider the differential equation,

$$\dot{x} = \epsilon X(x,t), \quad x(0) = x_0. \tag{2.21}$$

The actual flow of (2.21), represented by the chronological exponential, $\overrightarrow{\exp}\left(\int_0^t X_\tau \, \mathrm{d}\tau\right)$, is asymptotically approximated by the exponential of the logarithm of the original vector field as per Equation (2.10), rewritten below,

$$\overrightarrow{\exp}\left(\int_{t_0}^t \epsilon X_\tau \, \mathrm{d}\tau\right) \cong \exp\left(\overrightarrow{\mathrm{V}}_{t_0,t}\left(\epsilon X_\tau\right)\right) = \mathrm{Id} + \sum_{m=1}^{\infty} \frac{1}{m!}\left(\overrightarrow{\mathrm{V}}_{t_0,t}\left(\epsilon X_\tau\right)\right)^m.$$

Explicit calculations can be complicated, since both the variation, $\overrightarrow{\mathrm{V}}_{t_0,t}\left(\epsilon X_\tau\right)$, and the exponent, $\exp(\cdot)$, are infinite series. The parameter ϵ can be used to determine how the error of a finite series approximation scales, see Proposition 2.

The expansion of only the first element of the variation is trivial,

$$\overrightarrow{\mathrm{V}}^{(1)}_{t_0,t}\left(\epsilon X_\tau\right) = \int_{t_0}^t \epsilon X_\tau \, \mathrm{d}\tau \,,$$

meaning that the expansion to first-order is

$$\overrightarrow{\exp}\left(\int_{t_0}^t \epsilon X_\tau \, \mathrm{d}\tau\right) \approx \mathrm{Id} + \int_{t_0}^t \epsilon X_\tau \, \mathrm{d}\tau \, + O\left(\epsilon^2\right),$$

naturally recovering the first element of the Volterra series expansion. Continuing the expansion to include the second variation introduces the new term,

$$\overrightarrow{\mathrm{V}}^{(2)}_{t_0,t}\left(\epsilon X_\tau\right) = \frac{1}{2}\epsilon^2 \int_{t_0}^t \left[\int_{t_0}^\tau X_s \, \mathrm{d}s \, , X_\tau\right] \mathrm{d}\tau \,.$$

Incorporation into the exponent gives

$$\overrightarrow{\exp}\left(\int_{t_0}^t \epsilon X_\tau \, \mathrm{d}\tau\right) \approx \mathrm{Id} + \int_{t_0}^t \epsilon X_\tau \, \mathrm{d}\tau + \frac{1}{2}\epsilon^2 \int_{t_0}^t \int_{t_0}^\tau \left[X_s, X_\tau\right] \mathrm{d}s \; \mathrm{d}\tau + \frac{1}{2}\epsilon^2 \int_{t_0}^t X_\tau \, \mathrm{d}\tau \circ \int_{t_0}^t X_\tau \, \mathrm{d}\tau + O\left(\epsilon^3\right).$$

Decomposing the Jacobi-Lie bracket and using integration by parts one can retrieve the original Volterra series expansion to second-order,

$$\overrightarrow{\exp}\left(\int_{t_0}^t \epsilon X_\tau \, \mathrm{d}\tau\right) \approx \mathrm{Id} + \int_{t_0}^t \epsilon X_\tau \, \mathrm{d}\tau \, + \epsilon^2 \int_{t_0}^t \int_{t_0}^\tau X_s \circ X_\tau \, \mathrm{d}s \; \mathrm{d}\tau \, + O\left(\epsilon^3\right).$$

The expansion of the exponent of the logarithm can be done up to any desired order, and will coincide with the Volterra series to the same order (in ϵ). Higher-order expansions are prohibitively expensive from a computational perspective, due to the power series representations of the exponent and the logarithm. Once performed, and the terms determined, the computation need not be repeated. We shall soon see the role played by these expansions in the generalized averaging theory.

Under the condition that the first $(m-1)$ elements of the variations vanish, choosing only the first non-vanishing element to represent the flow gives the following theorem.

Proposition 4 *[2] If at the point $x \in \mathbb{R}^n$, the following holds:*

$$\overrightarrow{V}_{t_0,t}^{(\alpha)}(X_\tau) = 0, \ \forall\, \alpha = 1, \ldots, m-1,$$

then

$$\overrightarrow{\exp} \int_{t_0}^{t} X_\tau \, \mathrm{d}\tau = x + \overrightarrow{V}_{t_0,t}^{(m)}(X_\tau) + O\left(\left(\int_{t_0}^{t} ||X_\tau||_{s+2m} \, \mathrm{d}\tau \right)^{m+1} \right).$$

Notice that if the magnitude of the integral inside the $O(\cdot)$ were to be less than unity, the error would be dimished according to the number of vanishing initial terms. The ϵ term in perturbation theory, which scales the vector field X, will play an important role in the asymptotics of the missing terms using Proposition 4. Certainly, all of the prior theorems of this treatise can be viewed from the perspective of perturbation theory; a fact well known to Agrachev and Gamkrelidze.

The Variation of Constants. Many systems may not come in the form of Equation (2.21), but may come in the form

$$\dot{x} = X(x,t) + \epsilon \widetilde{X}(x,t). \tag{2.22}$$

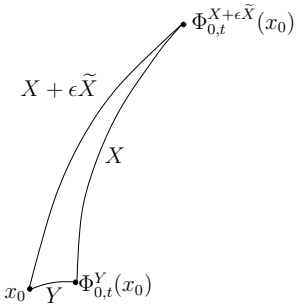


Figure 2.1: Depiction of the geometry of the variation of constants.

The variation of constants method can be used to transform the perturbed system into the form required by Equation (2.21). The variation of constants transformation operates as follows,

$$\Phi_{0,t}^{X+\epsilon\widetilde{X}} = \Phi_{0,t}^{X} \circ \Phi_{0,t}^{Y},$$

where $\Phi_{0,t}^{Y}$ is the flow corresponding to the differential equation,

$$\dot{y} = Y(y,t) \equiv \left(\Phi_{0,t}^{X}\right)^* \left(\epsilon \widetilde{X}\right) = \epsilon \left(\Phi_{0,t}^{X}\right)^* \widetilde{X}, \quad y(0) = x_0;$$

see [2, §4] for more details. Therefore, $\Phi_{0,t}^{Y}$ can be interpreted as the effect of the pertubation $\epsilon\widetilde{X}$ to X on the unperturbed flow $\Phi_{0,t}^{X}$, c.f. Figure 2.1. Using Equation (2.11), the time-varying vector field Y may be approximated by an autonomous vector field, Z,

$$Z = \ln \overrightarrow{\exp} \int_0^t Y_\tau \, \mathrm{d}\tau \,.$$

The ϵ parameter will carry through to the expression for Z, whereby one may obtain truncations of the infinite series representing Z. The variation of constants is a key transformation for proving many of the averaging results of Section 2.3. It is discussed further in Section 2.5, and is utilized in several of the examples found in Section 2.6.

In conclusion, the exponential representation for flows and vector fields leads to infinite series expansions for both. If it is possible to calculate the full series expansions, then solutions may be found in closed form. If, instead, the expansions prove to be computationally intractable, then perturbation methods can be used to obtain finite series approximations. The standard form for application of perturbation methods is found in Equation (2.21). The variation of constants may be used if the system under inspection is not *a priori* in the standard form of Equation (2.21).

2.3 Averaging via Floquet Theory

This section reformulates averaging theory in terms of nonlinear Floquet theory. The basic theorems of Floquet theory are first reviewed in the linear setting, and then extended to the nonlinear setting. Averaging theory is then shown to be the synthesis of nonlinear Floquet theory with the perturbation methods of Section 2.2.3. Consider the flow of the differential equation (2.5) rewritten below,

$$\dot{x} = X(x,t)\,, \quad x(0) = x_0, \tag{2.23}$$

where X is T-periodic, i.e., $X(x,t) = X(x,t+T)$.

2.3.1 Linear Floquet Theory

The case of a homogeneous linear periodic system,

$$\dot{x} = A(t)x, \qquad A(t) = A(t+T), \tag{2.24}$$

is well understood using linear Floquet theory. For convenience, we repeat here many of the basic theorems of linear Floquet theory. For linear systems, the flow corresponding to Equation (2.24), written as $\Phi^A_{0,t}$, can be represented by a matrix, called the *fundamental matrix solution*. Consequently, $\Phi^A_{0,t}$ will be used interchangably for both, as they represent the same object.

Lemma 1 *[51] If C is an invertible matrix, then there is a matrix B such that $C = exp(B)$.*

The proof of this Lemma involves using the logarithm function for matrices [51, pg. 118], consequently Lemma 1 means that $B = \ln(C)$. Suppose that C is the fundamental matrix solution, $\Phi^A_{0,t}$, of the system in Equation (2.24) for a fixed time, t. Then, in the notation of the chronological calculus,

$$B = \ln\left(\Phi^A_{0,t}\right) = \ln \overrightarrow{\exp} \int_0^t A(\tau)\,\mathrm{d}\tau\,,$$

for t fixed, c.f. Equation (2.11).

Theorem 6 (Linear Floquet Theorem) *[51] Every fundamental matrix solution $\Phi^A_{0,t}$ of (2.24) has the form*

$$\Phi^A_{0,t} = P(t)\exp(Bt),$$

where P(t) is T-periodic, $P(t+T) = P(t)$, and B is constant.

Definition 4 *The* monodromy matrix *corresponding to the flow of (2.24) is the fundamental solution matrix of (2.24) after one period of evolution, $M = \Phi^A_{0,T}$.*

The monodromy matrix is the Poincare map for the periodic system (2.24), therefore its eigenvalues may be used to determine stability of system (2.24). The eigenvalues are typically called the *Floquet multipliers*, or the *characteristic multipliers*.

Theorem 7 *[51] A necessary and sufficient condition that the system (2.24) be uniformly stable is that all Floquet multipliers of (2.24) have modulii less than unity.*

Theorem 6 and Lemma 1 imply that the logarithm of the monodromy matrix may be used to determine stability. From the proof of the linear Floquet theorem [51, pg. 118], it is clear the autonomous linear homogeneous system that represents the average of (2.24) is

$$\dot{y} = By, \tag{2.25}$$

where

$$B = \frac{1}{T}\ln(M) = \frac{1}{T}\ln\overrightarrow{\exp}\int_0^T A(\tau)\,\mathrm{d}\tau\,. \tag{2.26}$$

Therefore, an alternative to determining stability of the mondramy matrix, M, is determining stability of the averaged system (2.25) via B. The eigenvalues of B are called the *Floquet exponents*, or the *characteristic exponents*.

Corollary 1 *[51] A necessary and sufficient condition that the system (2.24) be uniformly stable is that all characteristic exponents of (2.25) lie in the complex left half-plane.*

Non-uniqueness of Monodromy Matrix. Given two fundamental matrix solutions, $\Phi_{0,t}^{A_1}$ and $\Phi_{0,t}^{A_2}$, to (2.24) there must exist a transformation of state, D, taking A_1 to A_2, i.e., $A_2 = D^{-1}A_1D$, [51]. Given the monodromy matrix for the flow of A_1, denoted M_1, the monodromy matrix for the flow of A_2 is $M_2 = D^{-1}M_1D$. In developing Floquet theory for nonlinear systems, we shall see this same freedom occur. The freedom is obtained by noting that the time-periodicity of A in Equation (2.24) implies that the time shift $\tau \mapsto t+T$, does not change the differential equation in Equation (2.24). Consequently, there exist a multitude of solutions, each shifted by one period of time, T. For more details, see [51].

2.3.2 Nonlinear Floquet Theory

By using the exponential representation of the flow of nonlinear systems, linear Floquet theory nicely extends to the nonlinear time-periodic case, Equation (2.23). While hinted at in prior work [144], the extension of Floquet theory to nonlinear systems is fully developed in this section.

Theorem 8 (Nonlinear Floquet Theorem) *Let $\Phi_{0,t}^X$ be the flow generated by the time-periodic differential equation (2.23). If the monodromy map has a logarithm, then the flow $\Phi_{0,t}^X$ can be represented as a composition of flows $\Phi_{0,t}^X = P(t)\circ\exp(Zt)$, where P is T-periodic, and Z is an autonomous vector field.*

proof
The proof exactly follows the proof of the linear Floquet theorem [51]. It is assumed that the T-periodic vector field in (2.23) determines a flow, therefore the time shifted version, $\tau = t+T$, does so also. The flow of the time-shifted version is the solution to the differential equation,

$$\frac{\mathrm{d}x}{\mathrm{d}\tau} = X(x,\tau).$$

The flows $\Phi_{0,t}^X$ and $\Phi_{0,\tau}^X$ differ by an invertible mapping, Ψ,

$$\Phi_{0,\tau}^X = \Phi_{0,t+T}^X = \Phi_{0,t}^X\circ\Psi.$$

Assume, for now, that there exists an autonomous flow denoted by $\Phi_{0,t}^Z$ equaling Ψ at time T. Consider,

$$P(t) \equiv \Phi_{0,t}^X\circ\left(\Phi_{0,t}^Z\right)^{-1}. \tag{2.27}$$

This defined flow, $P(t)$, is T-periodic.

$$\begin{aligned} P(t+T) &= \Phi_{0,t+T}^X\circ\left(\Phi_{0,t+T}^Z\right)^{-1} = \Phi_{0,t}^X\circ\Psi\circ\left(\Phi_{0,t}^Z\circ\Phi_{0,T}^Z\right)^{-1} \\ &= \Phi_{0,t}^X\circ\Psi\circ\left(\Phi_{0,T}^Z\right)^{-1}\circ\left(\Phi_{0,t}^Z\right)^{-1} = \Phi_{0,t}^X\circ\left(\Phi_{0,t}^Z\right)^{-1} \\ &= P(t). \end{aligned}$$

The T-periodicity of the original vector field X ensures that

$$\Phi^X_{0,t+T} = \Phi^X_{0,t} \circ \Phi^X_{0,T},$$

implying that

$$\Phi^Z_{0,T} = \Psi \equiv \Phi^X_{0,T}, \tag{2.28}$$

i.e., Ψ is the monodromy map of the flow. If we rewrite the respective flows in the chronological calculus formalism, then

$$\exp(ZT) = \overrightarrow{\exp}\left(\int_0^T X_\tau \, \mathrm{d}\tau\right).$$

Inverting via the logarithm,

$$Z = \frac{1}{T} \ln \overrightarrow{\exp}\left(\int_0^T X(x,\theta) \, \mathrm{d}\theta\right). \tag{2.29}$$

This is precisely the average of the T-periodic vector field; a connection that will be made more explicit in the sequels. Conditions under which this logarithm exists are given in [2]. For now, the theorem statement assumes the existence of the logarithm, and by extension the autonomous flow assumed in the proof of the theorem.[5]

■

The mapping $P(t)$ is called the *Floquet mapping*, and the vector field Z is called the *autonomous averaged vector field corresponding to* X. When the vector field X is known through context, then Z will simply be called the *autonomous averaged vector field*.

The monodromy map, Ψ, plays an important role in nonlinear Floquet theory, just as it did in linear Floquet theory. The proposition below relates the stability of the monodromy map to the stability of the original system.

Theorem 9 *If the monodromy map, Ψ of the system (2.23) has a fixed point, then the actual flow, $\Phi^X_{0,t}$, has a periodic orbit whose stability is determined by the stability of the monodromy map.*

proof

Suppose that the monodromy map has a fixed point denoted x^*, then the actual flow periodically returns to this fixed point, as can be seen by the following equalities:

$$\begin{aligned}
\Phi^X_{0,0}(x^*) &= (P(0) \circ \exp(Z\,0))(x^*) = x^*, \text{ and} \\
\Phi^X_{0,T}(x^*) &= (P(T) \circ \exp(Z\,T))(x^*) = (P(0) \circ \Psi)(x^*) = x^*,
\end{aligned}$$

implying that x^* lies on a periodic solution to the flow, $\Phi^X_{0,t}$, of the system (2.23). The orbit through x^* is the orbit of the theorem statement. The monodromy map is the Poincaré map of (2.23), evaluated on the orbit through x^* [162].

■

If the monodromy map is asymptotically (exponentially) stable, then the correspoding orbit is asymptotically (exponentially) stable.

Corollary 2 *If the flow, $\Phi^X_{0,t}$, of system (2.23) has a fixed point x^*, as does the monodromy map, Ψ, then stability of the fixed point under the flow $\Phi^X_{0,t}$ can be determined from the stability of the monodromy map. In particular an asymptotically (exponentially) stable fixed point for the monodromy map implies an asymptotically (exponentially) stable fixed point for the actual system (2.23).*

[5] In Section 2.2 we restricted the vector fields to be smooth. Smoothness implies existence of the logarithm for finite time.

proof
To see that stability of the monodromy map implies stability of the actual flow, it will be necessary to appeal to notions of D-stability and C-stability in per Pars [131]. C-stability is the traditional notion of stability; the "C" stands for continuous time. The "D" in D-stability refers to discrete time and involves taking discrete samples of the system response at fixed intervals. Pars shows that the two notions are equivalent when stability of a fixed point is sought. If the monodromy map is stable, then that implies D-stability for the actual system. Since D-stability and C-stability are equivalent under the conditions of interest, the system is stable. The same holds for D-instability.
■
Explicitly calculating the monodromy map may be problematic if not impossible. Equivalent to stability of the monodromy map is stability of the autonomous vector field Z whose flow gives the monodromy map.

Corollary 3 *[144] The stability properties of the logarithm of the monodromy map may be used to infer the stability properties of the monodromy map itself.*

Comment. When applied to the simplified case of linear stability analysis, the above conclusions lead to the following well known fact for Floquet theory: calculation of the Floquet multipliers is equivalent to calculation of the Floquet exponents. In the nonlinear theory, the monodromy map is a nonlinear transformation of state, therefore it's a bit more complicated to determine such a property. In this case, the logarithm provides a useful means to analyze stability.

Non-uniqueness of the Monodromy Map. Given two fundamental matrix solutions, $\Phi_{0,t}^{X_1}$ and $\Phi_{0,t}^{X_2}$, to (2.24) there must exist a transformation of state, Θ, from X_1 to X_2, i.e., $X_2 = \Theta^* X_2$. Given the monodromy matrix for X_1, called Ψ_1, the monodromy matrix for X_2 is $\Psi_2 = \Theta^{-1}\Psi_1\Theta$. As in the linear case the freedom occurs because of the ability to shift time by T.

An additional freedom may occur because of the dependence of the mondramy map calculation on initial conditions. Although the monodromy map was defined to be $\Psi = \Phi_{0,T}^{X}$, any choice of initial time may be used, $\widetilde{\Psi} = \Phi_{t_0,t_0+T}^{X}$, so long as there is flow for one period of time. The two monodromy maps will differ by a transformation of state Θ, i.e., $\Psi = \Theta^{-1}\widetilde{\Psi}\Theta$, corresponding to flow by the difference in initial times.

It is not known *a priori* whether or not a given initial condition will provide the best calculation of the monodromy map or its logarithm, therefore one should not expect for the evolution of the autonomous vector field, Z, to generically result in the best average. It may be possible for the time-periodic mapping, P, to contain a bias that leads to a better representation of the average. Suppose that

$$P(t) = \widetilde{P}(t)P_0, \tag{2.30}$$

with P_0 a time-independent transformation. It is possible to recover a different averaged vector field from this knowledge.

Theorem 10 *Suppose the the Floquet mapping has a time-independent bias, i.e., Equation (2.30) holds. Then the new averaged vector field is*

$$\widetilde{Z} = (P_0)_* Z.$$

proof
The proof is rather straightforward and is basically a change of coordinates. The autonomous flow is

$$z(t) = \exp(Zt)x(0)$$

and the actual flow is

$$x(t) = P(t, z(t)).$$

However, with the bias mapping, we obtain

$$x(t) = \widetilde{P}(t) \circ P_0(z(t)).$$

Define the new variable, $\widetilde{z}(t)$,

$$\widetilde{z}(t) = P_0(z(t)).$$

The evolution of $\widetilde{z}(t)$ obeys the differential equation,

$$\dot{\widetilde{z}} = TP_0\left(Z(P_0^{-1}(z))\right) = (P_0)_* Y(z), \quad \widetilde{z}_0 = P_0(x_0)$$

and the solution evolves according to

$$x(t) = \widetilde{P}(t) \circ \exp(\widetilde{Z}t) \circ P_0(x_0).$$

■

The role played by P_0 is equivalent to Θ in the discussion of the non-uniqueness of the monodromy map. The flexibility inherent to the decomposition P and $\exp(Zt)$ is well known in classical n^{th}-order averaging theory [42]. The opposite may also be desirable: to extract unecessary coordinate transformations from the averaged vector field, the reverse procedure may be performed. This is important when one would like to utilize the stability theorems of Section 2.3, because the bias might prevent the fixed point of the autonomous vector field from corresponding to the fixed point of the actual vector field.

2.3.3 Application to Averaging Theory

Floquet theory is an important tool for averaging theory. The vector field Z from Theorem 8 is the exact average of the time-periodic vector field from Equation (2.23). Since the actual flow oscillates around the trajectory determined by the autonomous vector field Z, the monodromy map gives the turnpike behavior of the system [145]. To check stability, one must calculate the monodromy map or the full series expansion of Z. If Z cannot be found in closed form, then truncations of the series expansions for the vector field Z can be thought of as partial averages of the system on the order of the truncation [144]. In order to compute truncations of the infinite series, it is essential for the relevant expression to be in the standard form for the perturbation methods of Section 2.2.3 to be applied. The standard form, Equation (2.21), is rewritten below,

$$\dot{x} = \epsilon X(x, t), \quad X(\cdot, t + T) = X(\cdot, t). \tag{2.31}$$

Introduction of the ϵ parameter allows for important proximity results and minimizes the error when truncating a series expansion, c.f. Section 2.2.3. The averaged vector field corresponding to (2.31) is the autonomous vector field,

$$Z = \frac{1}{T} \ln \overrightarrow{\exp} \left(\int_0^T \epsilon X_\tau \, \mathrm{d}\tau \right).$$

Recall that the integrand of the m^{th} variation of the identity flow, $\mathrm{v}^{(m)}(\cdot)$, is given by the sum of Lie brackets, c.f. Equation (2.18). The arguments are first-order homogeneous so that the parameter ϵ may be factored.

Therefore, Z, has a power series expansion in ϵ,

$$Z \equiv \sum_{\alpha=1}^{\infty} \epsilon^{\alpha} \Lambda^{(\alpha)}. \tag{2.32}$$

Definition 5 *If the function* F *can be given by a power series expansion in* ϵ*, then* $\mathrm{Trunc}_m(\mathrm{F})$ *is a truncation of the power series, where the* $(m+1)$ *and higher terms are removed.*

Thus,

$$\mathrm{Trunc}_m(Z) = \sum_{\alpha=1}^{m} \epsilon^{\alpha} \Lambda^{(\alpha)}.$$

Definition 6 *A truncated power series expansion in* ϵ *is considered to be a* stabilized expansion with respect to property P *if the inclusion of additional terms to the truncation does not affect the given property of the series expansion, i.e., if property* P *holds for all* $\mathrm{Trunc}_{m+k}(\mathrm{F})$*,* $k > 0$*, when property* P *holds for* $\mathrm{Trunc}_m(\mathrm{F})$.

In this treatise, the desired property is linear stability for vector fields.

Definition 7 *[144] A* stabilized truncated series expansion with respect to linear stability *for the vector field (2.32) is a truncated vector field that has the same linear stability properties as any higher-order truncation of the vector field, and also the full series expansion of the vector field.*

Definitions 5-7 imply that when a series is stabilized with respect to linear stability for a given truncation, $\mathrm{Trunc}_m(\cdot)$, the eigenvalues will not be significantly affected by incorporating additional terms of the series to the truncation. At this stabilized truncation, linear stability can be computed and used to determine the linear stability of the original system as per the previous propositions. See [144, 145] for particular examples involving matrix ODEs. This theorem is a generalization of known problems in classical higher-order averaging theory; see [161] and references therein, and also [111].

If a truncation has not yet stabilized, it may still be capable of approximating the actual flow. The following theorem determines the valid interval of approximation. It is reminiscent of classical averaging theorems, only it holds for arbitrary truncations.

Theorem 11 *The* m^{th}*-order truncation of the logarithm of the monodromy map gives an* $(\mathrm{m}+1)^{th}$*-order approximation of the flow for finite time, i.e.,*

$$\exp(Zt) = \exp(\mathrm{Trunc}_m(Z)\, t) + O\left(\epsilon^{m+1}\right)$$

for time $o(1)$.

proof
The difference in the flows could be understood by decomposing the total logarithm into a truncation and a truncated remainder,

$$Z = Z^m + \widetilde{Z}, \quad \text{where } Z^m = \mathrm{Trunc}_m(Z),$$

and using the variation of constants formula on the flow,

$$\Phi_{0,t}^{Z} = \Phi_{0,t}^{Y} \circ \Phi_{0,t}^{Z^m},$$

where

$$Y = \left(\Phi_{-t,0}^{Z^m}\right)^{*} \widetilde{Z}.$$

Thus the flow, $\Phi_{0,t}^{Y}$, acts as a perturbation to the final point of the flow of the truncated vector field. The size of this perturbation determines how far the flow of the truncated vector field is from the actual flow used to find the monodromy map. Since the pull-back is a linear operator, the vector field Y will scale with ϵ according to its contribution in $\widetilde{Z}$. By definition, $\widetilde{Z}$ has a factor of $\epsilon^{(m+1)}$ that may be factored. Therefore,

$$Y = \epsilon^{(m+1)} \left(\Phi_{-t,0}^{Z^m}\right)^* \widehat{Z} \equiv \epsilon^{(m+1)} \widehat{Y},$$

where

$$\widetilde{Z} \equiv \epsilon^{(m+1)} \widehat{Z}.$$

At this point, invoke Proposition 4 with $m = 0$.

$$\overrightarrow{\exp}\left(\int_0^t Y \,\mathrm{d}\tau\right) = \mathrm{Id} + O\left(\int_0^t ||Y_\tau||_s \,\mathrm{d}\tau\right) = \mathrm{Id} + O\left(\epsilon^{m+1} \int_0^t \left|\left|\widehat{Y}_\tau\right|\right|_s \,\mathrm{d}\tau\right)$$

By taking the maximum of $\widehat{Y}$ over space and time, one can arrive at,

$$\overrightarrow{\exp}\left(\int_0^t Y \,\mathrm{d}\tau\right) = \mathrm{Id} + O\left(\epsilon^{m+1} t^{m+2}\right).$$

The rest follows on the timescale 1.

■

Note that the order of time may be increased at the cost of orders of approximation. For example, choosing time on the order of $O\left(\epsilon^{-1/2}\right)$ moves the order of approximation to $O\left(\epsilon^{(m+1)/(2m+4)}\right)$. If the truncation is linearly stable, and is a stabilized truncation with respect to linear stability, then Theorem 2 can be used to extend the timescale of the approximation.

Theorem 12 *[161] Suppose that the* m^{th}*-order truncation of the logarithm of the monodromy map is stabilized with respect to linear stability. Then the truncation is* $(\mathrm{m}+1)^{th}$*-order close over the time domain of definition, i.e.,*

$$\exp(Zt) = \exp(\mathrm{Trunc}_m\left(Z\right) t) + O\left(\epsilon^{m+1}\right)$$

for $t \in [0, t_f)$*, where* t_f *is the maximal time for which the flow is defined.*

Thus it has been shown that in lieu of the full logarithm series expansion, it may be possible to utilize truncations of the expansion. The truncations calculated as per Section 2.2, c.f. Equation (2.19), consist of the variations from Equation (2.18). The variations up to order four are given by

$$\begin{aligned}
\Lambda^{(1)} &= \overline{X}, \\
\Lambda^{(2)} &= \frac{1}{2}\overline{\left[\int_0^t X_\tau \,\mathrm{d}\tau\,, X_t\right]}, \\
\Lambda^{(3)} &= \frac{1}{2}T\left[\Lambda^{(1)}, \Lambda^{(2)}\right] + \frac{1}{3}\overline{\left[\int_0^\tau X_{\tau_1} \,\mathrm{d}\tau_1\,, \left[\int_0^\tau X_{\tau_1} \,\mathrm{d}\tau_1\,, X_\tau\right]\right]}, \text{ and} \\
\Lambda^{(4)} &= -\frac{1}{12}\overline{\int_0^\tau \left[\int_0^{\tau_1} \left[\int_0^{\tau_2} X_{\tau_3} \,\mathrm{d}\tau_3\,, X_{\tau_2}\right] \,\mathrm{d}\tau_2\,, [X_{\tau_1}, X_\tau]\right] \,\mathrm{d}\tau_1} \\
&\quad -\frac{1}{12}\overline{\left[\int_0^\tau \left[\int_0^{\tau_1} \left[\int_0^{\tau_2} X_{\tau_3} \,\mathrm{d}\tau_3\,, X_{\tau_2}\right] \,\mathrm{d}\tau_2\,, x_{\tau_1}\right] \,\mathrm{d}\tau_1\,, X_\tau\right]} \\
&\quad -\frac{1}{12}\overline{\int_0^\tau \left[\int_0^{\tau_1} X_{\tau_2} \,\mathrm{d}\tau_2\,, \left[\left[\int_0^{\tau_1} X_{\tau_2} \,\mathrm{d}\tau_2\,, X_{\tau_1}\right], X_\tau\right]\right] \,\mathrm{d}\tau_1}.
\end{aligned} \tag{2.33}$$

2.3.4 Truncations of the Floquet Mapping

It was shown in the previous subsection how to obtain truncations of the autonomous Floquet flow for approximation of the infinite series expansion. Here it is likewise shown that one may calculate truncations of the time-periodic mapping $P(t)$. Recall that $P(t)$ is given by

$$P(t) \equiv \Phi_{0,t}^{X} \circ \exp\left(-Zt\right).$$

Theorem 13 *An* m^{th}*-order truncation of the time-periodic Floquet mapping is of order* $(m+1)$*-close to the time-periodic Floquet mapping on the timescale* $o\left(1\right)$*.*

$$P(t) = \mathrm{Trunc}_m\left(P(t)\right) + O\left(\epsilon^{m+1}\right).$$

proof
The proof is similar that of Theorem 11. We show the difference between the two flows is of the order $O\left(\epsilon^{m+1}\right)$. The two flows whose composition results in $P(t)$, c.f. Equation (2.27), can be written as power series expansions in ϵ. Hence, consider the two flows to consist of a truncation and a truncated remainder,

$$\Phi_{0,t}^{X} = \Phi^m + \widetilde{\Phi},$$

and

$$\exp\left(-Zt\right) = \Psi^m + \widetilde{\Psi},$$

where

$$\Phi^m = \mathrm{Trunc}_m\left(\Phi_{0,t}^{X}\right) \quad \text{and} \quad \Psi^m = \mathrm{Trunc}_m\left(\exp\left(-Zt\right)\right).$$

Then,

$$\Phi_{0,t}^{X} \circ \exp\left(-Zt\right) = \left(\Phi^m + \widetilde{\Phi}\right) \circ \left(\Psi^m + \widetilde{\Psi}\right).$$

Expanding,

$$\Phi_{0,t}^{X} \circ \exp\left(-Zt\right) = \mathrm{Trunc}_m\left(\Phi^m \circ \Psi^m\right) + \widetilde{\Theta},$$

where $\mathrm{Trunc}_m\left(\Phi^m \circ \Psi^m\right)$ is the truncation of the Floquet mapping and,

$$\widetilde{\Theta} = \left(\Phi^m \circ \Psi^m - \mathrm{Trunc}_m\left(\Phi^m \circ \Psi^m\right)\right) + \widetilde{\Phi} \circ \Psi^m + \Phi^m \circ \widetilde{\Psi} + \widetilde{\Phi} \circ \widetilde{\Psi},$$

is the truncated remainder of the Floquet mapping. The quantity ϵ^{m+1} can be factored from $\widetilde{\Theta}$,

$$\widetilde{\Theta} \equiv \epsilon^{m+1}\widehat{\Theta},$$

implying that,

$$\left|\left|\Phi_{0,t}^{X} \circ \exp\left(-Zt\right) - \mathrm{Trunc}_m\left(\Phi^m \circ \Psi^m\right)\right|\right|_{s,M} = \left|\left|\widetilde{\Theta}\right|\right|_{s,M} = \epsilon^{m+1}\left|\left|\widehat{\Theta}\right|\right|_{s,M}.$$

The composed flows that result in $P(t)$ exist and are finite on the timescale 1, therefore the truncated remainder of $P(t)$ exists and is finite on the timescale 1.
■
So far, we have the following method of averaging for a time-periodic vector field,

$$x(t) = \mathrm{Trunc}_{m-1}\left(P(t)\right)\left(z(t)\right) + O\left(\epsilon^m\right), \text{ and} \tag{2.34a}$$

$$\dot{z} = \mathrm{Trunc}_m\left(Z\right), \tag{2.34b}$$

where Z is the infinite series expansion for the autonomous vector field from the nonlinear Floquet theorem (Theorem 8). It is not known a priori what properties the truncation of $P(t)$ will have, however one essential ingredient is time-periodicity. Therefore a realistic constraint to add to the truncation is time-periodicity,

$$\mathrm{Trunc}_m\left(P(t)\right) = \mathrm{Trunc}_m\left(P(t+T)\right).$$

The removal of the higher-order terms may result in an aperiodic truncated Floquet mapping. If $\mathrm{Trunc}_m\left(P(t)\right)$ is not periodic, but $\mathrm{Trunc}_m\left(P(0)\right) = \mathrm{Trunc}_m\left(P(T)\right)$, then it is possible to make $\mathrm{Trunc}_m\left(P(t)\right)$ periodic. Take the truncated Floquet mapping as defined on $t \in [0,T)$, then extend it periodically for all time by defining $\mathrm{Trunc}_m\left(P(\tau+kT)\right) \equiv \mathrm{Trunc}_m\left(P(\tau)\right)$, where $\tau \in [0,T)$ and $k \in \mathbb{Z}$. When the domain of definition for $\mathrm{Trunc}_m\left(P\right)$ is adjusted to be periodic, it is called the *amended truncation of the Floquet mapping* or, when the context is clear, the *amended truncation*.

Corollary 4 *If the (amended) truncation* $\mathrm{Trunc}_m\left(P(t)\right)$ *is periodic with period* T*, and the period is on the timescale* 1*, then the (amended) truncation is order* $\epsilon^{(m+1)}$*-close to* $P(t)$ *for all time.*[6]

proof
From Theorem 13, the flow is order $\epsilon^{(m+1)}$-close on the timescale 1. If the period lies within the timescale 1, then the periodicity of both $P(t)$ and the (amended) truncation $\mathrm{Trunc}_m\left(P(t)\right)$ implies that the two mappings will return to the same values on a periodic basis. If $P(t)$ and the (amended) truncation $\mathrm{Trunc}_m\left(P(t)\right)$ do not diverge over one period, the will not diverge over any number of periods.
■

Improved m^{th}-order averaging. Ellison et al. [42] discuss what they term to be *improved* m^{th}*-order averaging*. Essentially, the improved average comes from the observation that the truncation of the Floquet mapping, $P(t)$, is not of the same order of approximation as the truncation of the autonomous vector field, Z. As compared to Equation (2.34), the improved average incorporates an additional order of expansion in the Floquet mapping,

$$x(t) = \mathrm{Trunc}_m\left(P(t)\right)\left(z(t)\right) + O\left(\epsilon^{m+1}\right) \tag{2.35a}$$

$$\dot{z} = \mathrm{Trunc}_m\left(Z\right). \tag{2.35b}$$

The quasi-periodic and aperiodic cases. The papers by Perko and Sáenz [132, 143] have shown that the KBM method of averaging may be used for quasi-periodic and aperiodic systems. The Floquet decomposition in Theorem 8 is still applicable for quasi-periodic and aperiodic systems, however the periodicity of the Floquet mapping, $P(t)$, no longer holds. The method for finding the autonomous vector field, Z, does not depend on the time-periodicity of the vector field X from (2.23). This fact was already proven by Agravchev and Gamkrelidze [2], and is simply the logarithm function

$$Z = \frac{1}{T}\ln\overrightarrow{\exp}\left(\int_0^T X_\tau \,\mathrm{d}\tau\right)$$

from equation (2.29). The choice of T is restricted to be within the interval of definition for the flow $\Phi_{0,t}^X$.

The ability of the generalized averaging theory to hold for the quasi-periodic and aperiodic cases demonstrates why the averaged expansions have such conservative proximity estimates. Classical averaging also begins with derivations of approximate autonomous vector fields that do not require time-periodicity [161].

[6] The corollary can be modified to get other orders of time at the sacrifice of orders of proximity.

The introduction of time-periodicity is then shown to improve the order of approximation. The same improvement occurs for the generalized averaging theory proven here, however it will not be proven. In the examples of Section 2.6, it can be seen that time-periodicity will improve the order of approximation, whereas aperiodicity will result in the conservative estimates of Theorems 11 and 13.

What is often emphasized in the literature is that averaging theory is a perturbation method that includes the computation of an approximate reconstruction function [42]. This function is the Floquet mapping, and maps the flow of the approximated autonomous system back to the flow of the original system. The approximate autonomous vector field, together with the Floquet mapping, will provide a good approximation to the actual flow for some finite order of time. Ellison et al. [42] argue that averaging is a more powerful perturbation method, compared to other perturbation methods, precisely because it allows for approximate reconstruction of the actual flow using the Floquet mapping $P(t)$.

2.4 A General Averaging Theory

This section demonstrates how the chronological calculus, nonlinear Floquet theory, and truncations of series expansions combine to form a generalized averaging theory. Nonlinear Floquet theory successfully decomposes the flow of a time-periodic system into the composition of a time-periodic mapping and the flow of an autonomous vector field. Classical higher-order averaging theory suggests the form of the autonomous vector field and the compensatory periodic mapping. Hence, the nonlinear Floquet theorem indicates that the compensatory mapping of classical averaging theory is the Floquet mapping, $P(t)$. Here, we demonstrate how series expansions and the chronological calculus are used to construct the compensatory (or Floquet) mapping and the autonomous averaged vector field.

First-order averaging. As a simple application, let's revisit first-order averaging. Unlike second- and higher-order averaging, first-order averaging does not involve a compensatory mapping. This is because

$$\mathrm{Trunc}_0\left(P(t)\right) = \mathrm{Trunc}_0\left(\Phi_{0,t}^X \circ \exp\left(-Zt\right)\right) = \mathrm{Id}. \tag{2.36}$$

Since the compensatory mapping is the identity, the T-periodicity constraint is trivially satisfied. This leaves the autonomous vector field,

$$Z = \frac{1}{T}\int_0^T \epsilon X_\tau \,\mathrm{d}\tau \,= \epsilon\overline{X}, \tag{2.37}$$

as the only important element in performing first-order averaging. Floquet theory can be applied to obtain the standard facts concerning stability of the average vector field flow and its relation to the actual flow.

Second-order averaging. The benefits of the chronological calculus become more apparent when we reconstruct the second-order averaging theorem of Sanders and Verhulst, c.f. Theorem 5. Truncating $P(t)$ results in

$$\mathrm{Trunc}_1\left(P(t)\right) = \mathrm{Id} + \epsilon\int_0^t \left(X_\tau - \overline{X}\right)\,\mathrm{d}\tau \,+ O\left(\epsilon^2\right). \tag{2.38}$$

In other words,

$$x(t) = z(t) + \epsilon\int_0^t \left(X(z(t),\tau) - \overline{X}(z(t))\right)\,\mathrm{d}\tau \,+ O\left(\epsilon^2\right), \tag{2.39}$$

where $z(t)$ is the flow of the autonomous vector field,

$$Z = \epsilon\overline{X} + \frac{1}{2}\epsilon^2\overline{\left[\int_0^t X_\tau\,\mathrm{d}\tau\,, X_t\right]}, \tag{2.40}$$

with the initial condition $z(0) = x_0$. The beauty of the series expansions approach lies in the fact that no new theorems need to be derived to obtain and analyze the consquences of the second-order average.

Equation (2.39) is close to what Sanders and Verhulst obtain for the compensating flow. Sanders and Verhulst also require the integral term in the flow to have a vanishing average [161]. Bogoliubov and Mitropolsky [17] obtain the same zero average assertion. By introducing an integration constant to ensure a zero average for the term $\int_0^t \left(X(y(t),\tau) - \overline{X}(y(t))\right)\,\mathrm{d}\tau$, the integration constant factors out the effect of varying the initial time t_0 for which the monodromy map in computed, c.f. Section 2.3.2. For systems that are asymptotically stable, the transient dynamics related to the initial conditions of the full system (2.21) are neglected. Examples in Sections 3.3 and 5.3 address this issue, and also relate it to Theorem 10.

Remark. There are significant differences between the generalized averaging theory and classical averaging theory in spite of their similarity. In order to utilize the powerful results of the chronological calculus, the vector fields must be smooth or analytic. The pertubation approach of Sanders and Verhulst requires the existence of a first derivative and Lipschitz continuity up to the first derivative affording it a level of generality that does not hold for the former. The same goes for the improved n^{th}-order averages found in Ellison et al. [42]. Truncated expansions may be calculated without regard to the requirements of the chronological calculus. By appealing to the classical averaging theorems it is possible to determine the minimal requirements for the calculated average to hold. Regardless of the requirements, it is still posible to carry out the formal series expansions, whose truncations may provide asymptotic understanding of the actual flow [145].

2.4.1 Higher-order Averaging

Third-order averaging. The truncated Floquet mapping for third-order averaging is

$$\begin{aligned}\mathrm{Trunc}_2\left(P(t)\right) =& \mathrm{Id} + \epsilon\int_0^t \left(X_\tau - \overline{X}\right)\,\mathrm{d}\tau + \frac{1}{2}\epsilon^2\int_0^t\left(\left[\int_0^\tau X_s\,\mathrm{d}s\,,X_\tau\right] - \overline{\left[\int_0^t X_s\,\mathrm{d}s\,,X_t\right]}\right)\,\mathrm{d}\tau \\ &+ \frac{1}{2}\epsilon^2\int_0^t X_\tau\,\mathrm{d}\tau\circ\int_0^t X_\tau\,\mathrm{d}\tau - \epsilon^2\int_0^t X_\tau\,\mathrm{d}\tau\circ\overline{X}\,t + \frac{1}{2}\epsilon^2\overline{X}\circ\overline{X}\,t^2\end{aligned} \tag{2.41}$$

This truncation satisfies $P(0) = P(T)$, but may not periodic for $t > T$. If this is so, then it may be amended according to the discussion in Section 2.3.4. The autonomous averaged vector field is

$$Z = \epsilon\overline{X} + \frac{1}{2}\epsilon^2\overline{\left[\int_0^t X_\tau\,\mathrm{d}\tau\,,X_t\right]} + \frac{1}{4}T\epsilon^3\left[\overline{X},\overline{\left[\int_0^t X_\tau\,\mathrm{d}\tau\,,X_t\right]}\right] + \frac{1}{3}\epsilon^3\overline{\left[\int_0^\tau X_{\tau_1}\,\mathrm{d}\tau_1\,,\left[\int_0^\tau X_{\tau_1}\,\mathrm{d}\tau_1\,,X_\tau\right]\right]} \tag{2.42}$$

Fourth-order averaging. The fourth-order averaging results begin to get complicated due to the rate of growth of the averaging terms with each higher order. Averaging calculations for higher orders increasingly become computationally expensive. This fourth level is the last averaging order that will be computed here. In the sequel is given the general algorithm followed in case fifth- or higher-order averaging is required.

The Floquet mapping for fourth-order averaging can be found in Table 2.1, and the autonomous averaged vector field is in Table 2.2. The truncation is easily amended to be T-periodic.

2.4.2 A General Averaging Algorithm

Although calculating higher-order expansions for averaging is a difficult task, there is a very simple algorithm for doing so. Based on the results from the previous sections, this algorithm will obtain a method

$$
\begin{aligned}
\mathrm{Trunc}_3\left(P(t)\right) =& \mathrm{Id} + \epsilon \int_0^t \left(X_\tau - \overline{X}\right)\, \mathrm{d}\tau \\
&+ \frac{1}{2}\epsilon^2 \int_0^t \left(\left[\int_0^\tau X_s \,\mathrm{d}s \,, X_\tau \right] - \overline{\left[\int_0^\sigma X_s \,\mathrm{d}s \,, X_\sigma \right]} \right) \mathrm{d}\tau \\
&+ \frac{1}{4}\epsilon^3 \left(\left[\int_0^t X_\tau \,\mathrm{d}\tau \,, \int_0^t \left[\int_0^\tau X_s \,\mathrm{d}s \,, X_\tau \right] \mathrm{d}\tau \right] - T \left[\overline{X}, \overline{\left[\int_0^t X_\tau \,\mathrm{d}\tau \,, X_t \right]} \right] t \right) \\
&+ \frac{1}{3}\epsilon^3 \int_0^t \left(\left[\int_0^\tau X_s \,\mathrm{d}s \,, \left[\int_0^\tau X_s \,\mathrm{d}s \,, X_\tau \right] \right] - \overline{\left[\int_0^t X_\tau \,\mathrm{d}\tau \,, \left[\int_0^t X_\tau \,\mathrm{d}\tau \,, X_t \right] \right]} \right) \mathrm{d}\tau \\
&+ \frac{1}{2}\epsilon^2 \int_0^t X_\tau \,\mathrm{d}\tau \circ \int_0^t X_\tau \,\mathrm{d}\tau - \epsilon^2 \int_0^t X_\tau \,\mathrm{d}\tau \circ \overline{X}\, t + \frac{1}{2}\epsilon^2 \overline{X} \circ \overline{X}\, t^2 \\
&+ \frac{1}{4}\epsilon^3 \int_0^t X_\tau \,\mathrm{d}\tau \circ \int_0^t \left[\int_0^\tau X_s \,\mathrm{d}s \,, X_\tau \right] \mathrm{d}\tau - \frac{1}{2}\epsilon^2 \int_0^t X_\tau \,\mathrm{d}\tau \circ \overline{\left[\int_0^t X_\tau \,\mathrm{d}\tau \,, X_t \right]}\, t \\
&- \frac{1}{2}\epsilon^3 \int_0^t \left[\int_0^\tau X_s \,\mathrm{d}s \,, X_\tau \right] \mathrm{d}\tau \circ \overline{X}\, t + \frac{1}{4}\epsilon^3 \overline{X} \circ \overline{\left[\int_0^t X_\tau \,\mathrm{d}\tau \,, X_t \right]}\, t^2 \\
&+ \frac{1}{4}\epsilon^3 \int_0^t \left[\int_0^\tau X_s \,\mathrm{d}s \,, X_\tau \right] \mathrm{d}\tau \circ \int_0^t X_\tau \,\mathrm{d}\tau + \frac{1}{4}\epsilon^3 \overline{\left[\int_0^t X_\tau \,\mathrm{d}\tau \,, X_t \right]} \circ \overline{X}\, t^2 \\
&+ \frac{1}{6}\epsilon^3 \left(\int_0^t X_\tau \,\mathrm{d}\tau \circ \int_0^t X_\tau \,\mathrm{d}\tau \circ \int_0^t X_\tau \,\mathrm{d}\tau \,, -\overline{X} \circ \overline{X} \circ \overline{X}\, t^3 \right)
\end{aligned}
$$

Table 2.1: Truncated Floquet mapping, $\mathrm{Trunc}_3\left(P(t)\right)$, for fourth-order average.

of averaging that contains two dimensions of approximation. The first is in the truncations of the logarithm vector field, $\mathrm{Trunc}_m\left(Z\right)$, which captures up to m^{th}-order the dynamics of the system. The second is the truncations of the compensation map, $\mathrm{Trunc}_{(m-1)}\left(P(t)\right)$, which will give m^{th}-order proximity of the composed system to the actual system.

The two components of the generalized averaging theory lead to a very important distinction that can sometimes be confused when performing higher-order averages. For example, ignoring the Floqut mapping and using only the autonomous flow results in averaged equations of motion that are capable of capturing the m^{th}-order dynamics, but will only be good to 1^{st}-order proximity. The distinction between roles of the mapping $P(t)$ and the autonomous vector field was also made by Bogoliubov and Mitropolsky [17].

The averaging process described by Bogoliubov and Mitropolsky utilizes the m^{th}-order average to find the $\left(\mathrm{m}+1\right)^{th}$-order average. This is done by viewing the solution to the $\left(\mathrm{m}+1\right)^{th}$-order average as coming from a perturbed version of the m^{th} order average. The technique is akin to a Picard iteration on the averages. The series expansions given in this dissertation and elsewhere are an attempt to extract structure from the Picard iterations (a.k.a Volterra series expansions). An important similarity is that the higher-order averages build upon the lower-order averages in a systematic way.

The averaging method provided here is capable of replicating classical averaging results and appears to coincide with known methods for obtaining higher-order averages. The general strategy implied by the averaging theory derived herein can be found in Table 2.3.

$$
\begin{aligned}
Z =& \epsilon \overline{X} + \frac{1}{2}\epsilon^2 \overline{\left[\int_0^t X_\tau \, \mathrm{d}\tau \, , X_t\right]} + \frac{1}{4} T \epsilon^3 \left[\overline{X}, \overline{\left[\int_0^t X_\tau \, \mathrm{d}\tau \, , X_t\right]}\right] \\
&+ \frac{1}{3}\epsilon^3 \overline{\left[\int_0^\tau X_{\tau_1} \, \mathrm{d}\tau_1 \, , \left[\int_0^\tau X_{\tau_1} \, \mathrm{d}\tau_1 \, , X_\tau\right]\right]} \\
&+ \frac{1}{3}\epsilon^3 \overline{\left[\int_0^\tau X_{\tau_1} \, \mathrm{d}\tau_1 \, , \left[\int_0^\tau X_{\tau_1} \, \mathrm{d}\tau_1 \, , X_\tau\right]\right]} + \frac{1}{3}\epsilon^3 \left[a_1, a_{21}\right] \\
&+ \frac{1}{3}\epsilon^3 \left[a_1, \overline{\left[\int_0^\tau X_{\tau_1} \, \mathrm{d}\tau_1 \, , X_\tau\right]}\right] + \frac{1}{3}\epsilon^3 \left[\overline{\int_0^\tau X_{\tau_1} \, \mathrm{d}\tau_1} \, , a_{21}\right] \\
&- \frac{1}{12}\epsilon^4 \overline{\int_0^\tau \left[\int_0^{\tau_1} \left[\int_0^{\tau_2} X_{\tau_3} \, \mathrm{d}\tau_3 \, , X_{\tau_2}\right] \mathrm{d}\tau_2 \, , \left[X_{\tau_1}, X_\tau\right]\right] \mathrm{d}\tau_1} \\
&- \frac{1}{12}\epsilon^4 \overline{\left[\int_0^\tau \left[\int_0^{\tau_1} \left[\int_0^{\tau_2} X_{\tau_3} \, \mathrm{d}\tau_3 \, , X_{\tau_2}\right] \mathrm{d}\tau_2 \, , X_{\tau_1}\right] \mathrm{d}\tau_1 \, , X_\tau\right]} \\
&- \frac{1}{12}\epsilon^4 \overline{\int_0^\tau \left[\int_0^{\tau_1} X_{\tau_2} \, \mathrm{d}\tau_2 \, , \left[\left[\int_0^{\tau_1} X_{\tau_2} \, \mathrm{d}\tau_2 \, , X_{\tau_1}\right], X_\tau\right]\right] \mathrm{d}\tau_1}
\end{aligned}
$$

Table 2.2: Truncated averaged vector field, $\mathrm{Trunc}_4\,(Z)$, for fourth-order average.

1. Calculate the logarithm vector field, $\mathrm{Trunc}_m\,(Z)$.
2. Calculate the truncated version of $\exp\,(-Zt)$.
3. Calculate the truncated version of $\overrightarrow{\exp}\left(\int_0^t X_\tau \, \mathrm{d}\tau\right)$.
4. Use the truncations for $\mathrm{Trunc}_{(m-1)}\,(P(t))$.

Table 2.3: Algorithm for Computing the Average.

2.5 Perturbed Systems with Periodic Properties

In order to apply Floquet theory, time-periodicity of X in (2.23) is required. Some nonlinear systems, however, exhibit time-periodic behavior without having time-periodic vector fields. A transformation of state may rectify this problem and place the system into a form suitable to averaging. For other systems, the infinite series required to calculate both the autonomous vector field Z and the Floquet mapping $P(t)$ in the statement of Theorem 8 may be intractable. In order obtain truncations of the infinite series expansions, the form of X must be compatible with the standard form for application of perturbation methods, see Equation (2.21). Many systems are not *a priori* in this form and require transformation to the standard form for averaging. This section investigates the commonly utilized technique known as the *variation of constants*, or the *variation of parameters*, first discussed in Section 2.2.3, for two special cases of dynamical systems. It is a transformation of state that may result in a dynamical system with periodic behavior and in the form required by perturbation methods.

2.5.1 The Variation of Constants Transformation

Consider the case of a dynamical system described by a nominal vector field, X_0, and an additional vector field $\widetilde{X}$,

$$\dot{x} = X_0 + \widetilde{X} \tag{2.43}$$

where $\widetilde{X}$ is independent of time, and X_0 has no such restriction. The variation of constants provides a way to focus on the vector field, $\widetilde{X}$, and begins with the transformation of state, $x(t) = \Phi_{0,t}^{X_0}(y(t))$. The associated evolution equations for the new state $y(t)$ are,

$$\dot{y} = Y(y,t) \equiv \left(\Phi_{0,t}^{X_0}\right)^* \widetilde{X}, \qquad y(0) = x_0. \tag{2.44}$$

Very often, what makes this technique appropriate is that the evolution of X_0 is "nice," in that it is time-periodic, or results in Y being time-periodic.

Assumption 1 *Assume that one of the following two possibilities holds,*

1. *The evolution of the vector field $X_0(x,t)$ is periodic with period T, i.e., $\Phi_{0,t}^{X_0} = \Phi_{0,t+T}^{X_0}$, or*
2. *The pull-back system, Y, given in equation (2.44) is periodic with period T.*

When $\Phi_{0,t}^{X_0}$ is computable in closed-form and the pull-back computation in Equation (2.44) can also be found in closed-form, then analysis may proceed directly. If the pull-back cannot be found in closed form, a series expansion may be used to approximate the vector field Y in Equation (2.44), [2],

$$Y = \widetilde{X} + \sum_{k=1}^{\infty} \int_{\tau_0}^{\tau} \int_{t_0}^{\tau_1} \cdots \int_{t_0}^{\tau_{k-1}} \mathrm{ad}_{X_0(\tau_k)} \cdots \mathrm{ad}_{X_0(\tau_1)} \widetilde{X} \, \mathrm{d}\tau_k \cdots \, \mathrm{d}\tau_2 \, \mathrm{d}\tau_1. \tag{2.45}$$

If the properties of X_0 and $\widetilde{X}$ are such that the series expansions for the pull-back vanish for some $n > N$, then the series is finite in extent and averaging can be applied to this finite summation. A nilpotent system satisfies this criterion. A finite summation may result even if the system is not nilpotent. If the series is not finite, then it must be truncated to some length, resulting in a loss of accuracy. By choosing to truncate at the same order as the order of averaging, this loss is minimized. Proposition 3.2 of [2] can be used to determine the error of this truncation.

As Assumption 1 implies that Y is time-periodic, Theorem 8 may be applied to decompose the flow of Y into the composition of two parts. When combined with the variation of constants transformation, this decomposition yields a Floquet decomposition of the perturbed system (2.43),

$$\Phi_t^X = \Phi_{0,t}^{X_0} \circ P(t) \circ \exp(Zt),$$

where Z is the autonomous averaged vector field corresponding to the time-periodic vector field Y. The composition

$$\widetilde{P} = \Phi_{0,t}^{X_0} \circ P \tag{2.46}$$

is periodic if possibility 1 of Assumption 1 holds. Therefore, the flow of X can be decomposed into an autonomous vector field flow composed with a T-periodic mapping. The Floquet and averaging analysis of Section 2.3 may then be applied.

If, instead, possibility 2 of Assumption 1 holds, then although the mapping, $P(t)$ is periodic, the transformation of state from y to x, $\Phi_{0,t}^{X_0}$, may not be time-periodic. Consequently, the Floquet mapping $\widetilde{P}(t)$ is not periodic and the Floquet analysis of Section 2.3 does not hold. Nevertheless, the decomposition may be used to emphasize important dynamical system behavior. Examples will be given in Section 2.6.

A Comment on the Transformation. Since the variation of constants method involves solving for the flow of the unperturbed system, the averaging procedure for a system transformed using the variation of constants method may be improved by an order of ϵ. In this case, to obtain the correct truncation of the Floquet mapping, the improved m^{th}-order Floquet mapping may be needed, c.f. Equation (2.34a) versus Equation (2.35a).

2.5.2 Perturbed Systems

The following two classes of perturbed systems, to be studied as examples in Section 2.6, can be analyzed via the variation of constants transformation.

1) Perturbed Conservative Systems. Consider a system where X_0 is the Hamiltonian vector field of a conservative Hamiltonian, H_0, and $\widetilde{X}$ is a perturbation to the Hamiltonian,

$$\dot{x} = X_{H_0} + \widetilde{X}. \tag{2.47}$$

Suppose that the flow of the unperturbed system is periodic. What conclusion can be drawn about the system (2.47) and its stability under perturbation? Stability can be studied via Equation (2.44). Using the Floquet decomposition,

$$\Phi^Y_{0,t} = P(t) \circ \exp(Zt),$$

stability of the flow of (2.47), can be determined from the stability properties of the flow of Z. This can be done in one of two ways, using the Floquet multipliers, or the Floquet exponents.

2) Vibrational control. Second order systems under high-frequency high-amplitude forcing can be studied with these methods. The general form for these systems is

$$\ddot{q} = X + \frac{1}{\epsilon}F(q, \frac{t}{\epsilon}), \tag{2.48}$$

where F is T-periodic. Moving to a first-order representation, $x = (q, \dot{q})$, and transforming time, $t/\epsilon \mapsto \tau$, leads to

$$\frac{\mathrm{d}x}{\mathrm{d}\tau} = \epsilon\widehat{X}(x) + \widehat{F}(x, \tau), \tag{2.49}$$

where

$$\widehat{X} = \left\{ \begin{array}{c} \dot{q} \\ X(x) \end{array} \right\}, \quad \text{and} \quad \widehat{F} = \left\{ \begin{array}{c} 0 \\ F(x, \tau) \end{array} \right\},$$

and it must be emphasized that F depends only on the state configurations and not on the state velocities. The variation of constants formula can be used to render system (2.49) into the form required by averaging theory. The form of F ensures that the resulting flow will be time-periodic as demanded by Assumption 1; see [23].

2.5.3 Highly Oscillatory Systems

There also exist unperturbed systems exhibiting oscillatory behavior that need to be truncated, but are not in the form required for truncation by perturbation methods. One particular class of such systems are high-oscillatory systems,

$$\dot{x} = X(x, t/\epsilon), \tag{2.50}$$

where $X(\cdot, t+T) = X(\cdot, t)$. A transformation of time, $t \mapsto \epsilon\tau$, converts the system to

$$\frac{\mathrm{d}x}{\mathrm{d}\tau} = \epsilon X(x, \tau).$$

The averaging process followed by series truncation may be performed, resulting in the m^{th}-th order Floquet decomposition,

$$x(\tau) = \mathrm{Trunc}_{m-1}\left(P\right)\left(\tau\right) + O\left(\epsilon^m\right),$$
$$\frac{\mathrm{d}z}{\mathrm{d}\tau} = \mathrm{Trunc}_m\left(Z\right).$$

Transforming back to original time, the m^{th}-th order Floquet decomposition of (2.50) becomes,

$$x(t) = \mathrm{Trunc}_{m-1}\left(P\right)\left(t/\epsilon\right) + O\left(\epsilon^{m-1}\right), \tag{2.52a}$$

$$\dot{y} = \frac{1}{\epsilon}\mathrm{Trunc}_m\left(Y\right). \tag{2.52b}$$

Although it is not explicitly noted, the series expansions for Y and P will contain increasing orders of ϵ. Because an order of ϵ is lost in the expansions when transforming back to the original time, t, it may be best to use improved m^{th}-th order averaging to recover the lost order in the Floquet mapping, c.f. Equation (2.35a). The work of Sussmann and Liu [155, 93] is recovered in this case, however they do not discuss the existence of the Floquet mapping, $P(t/\epsilon)$. Their work extends Kurzweil and Jarník [74, 75], where averaging is a limit process by which the autonomous differential equation (2.52b) is obtained from (2.50) as $\epsilon \to 0$.

2.6 Examples

This section contains several examples that are used to highlight different components of the generalized averaging theory. The first example in Section 2.6.1 is an averaged expansion which is finite in extent, meaning that the averaged expansion and Floquet mapping are exact. The Floquet decomposition can be found in closed form. In Section 2.6.2 a charged particle channeling through a crystal lattice is studied. The example demonstrates that the generalized averaging theory still applies to systems that are not strictly time-periodic. Lastly, Section 2.6.3 details how averaging theory may be used for nonlinear observability. The averaging process captures the nonlinear interaction dynamics of two coupled systems and utilizes the coupling to make estimations of state. Higher-order averaging is required to incorporate corrections to the first-order estimates.

2.6.1 Particle in an Electromagnetic Field

This example will consist of a charged particle in $\mathbb{R}^3$ constrained to move in the (x, y)-plane while subject to a magnetic field, B, pointing in the z direction, and an electric field, E, pointing in the x direction, see Figure 2.2. The phase space is $P \cong T^*\mathbb{R}^2$ with coordinates $(x, y, p_x, p_y) \in P$. This problem was examined in Omohundro [124] within the context of symmetry reduction.

The Hamiltonian of the system is

$$H = H_0 + \epsilon H_1 = \frac{1}{2m}\left(p_x^2 + p_y^2\right) - \epsilon e E x,$$

with the noncanonical Poisson bracket,

$$\{f,g\} = \{f,g\}_0 + \frac{eB}{c}\left(\frac{\partial f}{\partial p_x}\frac{\partial g}{\partial p_y} - \frac{\partial f}{\partial p_y}\frac{\partial g}{\partial p_x}\right),$$

where $\{\cdot,\cdot\}_0$ is the canonical Poisson bracket,

$$\{f,g\}_0 = \frac{\partial f}{\partial q^i}\frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p^i}\frac{\partial g}{\partial q^i}.$$

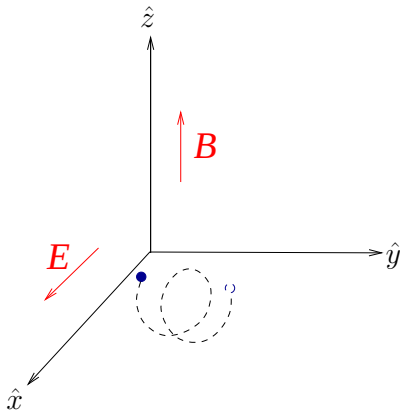


Figure 2.2: Charged particle in an electromagnetic field.

The constants, m, e, and c are the mass of the particle, the charge of the particle, and the speed of light, respectively. Finally, the parameter $\epsilon \in [0,1]$, determines how much of the electric field strength is acting on the particle. In the unperturbed case, $\epsilon = 0$, and in the case of the full field strength, it is $\epsilon = 1$. The dynamical equations of motion are,

$$\begin{aligned} \dot{x} &= \frac{p_x}{m} \qquad \dot{p}_x = \frac{eB}{mc}p_y + \epsilon eE \\ \dot{y} &= \frac{p_y}{m} \qquad \dot{p}_y = -\frac{eB}{mc}p_x. \end{aligned}$$

In the unperturbed case, the system is a charged particle in a constant magnetic field; trajectories are circular orbits. The unperturbed system and perturbative vector fields are,

$$X_0 = \begin{Bmatrix} \frac{1}{m}p_x \\ \frac{1}{m}p_y \\ \frac{eB}{mC}p_y \\ -\frac{eB}{mc}p_x \end{Bmatrix}, \quad \text{and} \quad \widetilde{X} = \epsilon \begin{Bmatrix} 0 \\ 0 \\ eE \\ 0 \end{Bmatrix}, \tag{2.53}$$

where the complete perturbed system is described by the flow of the vector field $X = X_0 + \epsilon\widetilde{X}$. Using the variation of constants formula, the system evolves as

$$\Phi^X_{0,t} = \Phi^{X_0}_{0,t} \circ \Phi^Y_{0,t},$$

where

$$Y = \epsilon\left(\Phi^{X_0}_{0,t}\right)^* \widetilde{X},$$

and Assumption 1 (possibility 1) is satisfied. Via Floquet theory, the flow of Y is decomposed into two parts,

$$\Phi^Y_{0,t} = P(t) \circ \Phi^Z_{0,t},$$

where Z is an autonomous vector field. In total, the flow of X is described by,

$$\Phi^{X_H}_{0,t} = \Phi^{X_0}_{0,t} \circ P(t) \circ \Phi^Z_{0,t}.$$

With $(q_0,p_0) = (x_0,y_0,p_{x,0},p_{y,0})$, the unperturbed flow is

$$\Phi^{X_0}_{0,t}(q_0,p_0) = \begin{Bmatrix} x_0 + \frac{p_{x,0}}{m\omega}\sin(\omega t) + \frac{p_{y,0}}{m\omega}\left(1-\cos(\omega t)\right) \\ y_0 - \frac{p_{x,0}}{m\omega}\left(1-\cos(\omega t)\right) + \frac{p_{y,0}}{m\omega}\sin(\omega t) \\ p_{x,0}\cos(\omega t) + p_{y,0}\sin(\omega t) \\ -p_{x,0}\sin(\omega t) + p_{y,0}\cos(\omega t) \end{Bmatrix},$$

where $\omega = \frac{eB}{mc}$. The vector field, Y, is given by

$$Y = \epsilon e E \left\{ \begin{array}{c} -\frac{1}{m\omega}\sin(\omega t) \\ -\frac{1}{m\omega}(1-\cos(\omega t)) \\ \cos(\omega t) \\ \sin(\omega t) \end{array} \right\},$$

and the average of Y is

$$Z = \epsilon e E \left\{ \begin{array}{c} 0 \\ -\frac{1}{m\omega} \\ 0 \\ 0 \end{array} \right\}.$$

An analysis of the system, shows that this is the exact average, i.e. it has not been truncated. The evolution of Z shows that there is a net drift in the negative y direction, about which the actual system will oscillate. Together, Y and Z can be used to find the partial Floquet mapping,

$$P(t) = \mathrm{Id} + \epsilon e E \left\{ \begin{array}{c} -\frac{1}{m\omega^2}\left(1-\cos(\omega t)\right) \\ \frac{1}{m\omega^2}\sin(\omega t) \\ \frac{1}{\omega}\sin(\omega t) \\ \frac{1}{\omega}(1-\cos(\omega t)) \end{array} \right\}.$$

The true Floquet mapping of the system, according to Equation (2.46), is the periodic mapping

$$\widetilde{P}(t) = \Phi_{0,t}^{X_0} \circ P(t).$$

Thus, using the variation of constants, and Floquet theory, it is possible to obtain a closed form Floquet decomposition of this system. The complete flow of the system is

$$\Phi_{0,t}^{X} = \widetilde{P}(t) \circ \exp(Zt) = \Phi_{0,t}^{X_0} \circ P(t) \circ \exp(Zt).$$

Explicitly composing the flows results in

$$\Phi_{0,t}^{X}(q_0, p_0) = \left\{ \begin{array}{c} x_0 + \frac{\epsilon e E}{m\omega^2}\left(1-\cos(\omega t)\right) + \frac{1}{m\omega}p_{x,0}\sin(\omega t) + \frac{1}{m\omega}p_{y,0}\left(1-\cos(\omega t)\right) \\ y_0 - \frac{\epsilon e E}{m\omega}\left(t - \frac{1}{\omega}\sin(\omega t)\right) - \frac{1}{m\omega}p_{x,0}\left(1-\cos(\omega t)\right) + \frac{1}{m\omega}p_{y,0}\sin(\omega t) \\ p_{x,,0}\cos(\omega t) + p_{y,0}\sin(\omega t) + \frac{\epsilon e E}{\omega}\sin(\omega t) \\ -\frac{\epsilon e E}{m\omega}\left(1-\cos(\omega t)\right) - p_{x,0}\sin(\omega t) + p_{y,0}\cos(\omega t) \end{array} \right\}.$$

By finding the solution in closed form using Floquet theory, we have demonstrated that averaging may be exact. The ϵ parameter is not needed to obtain the averaged vector field and Floquet mapping; there is no restriction on the value of ϵ. For nilpotent systems, the infinite series collapses to a finite series, and the exponential representation is exact if no truncation of the finite series is made. Consequently, the averaging proces is exact. Only in the case that the average vector field requires an infinite series expansion are truncations necessary, at which point the method of averaging becomes an approximation, and ϵ plays a greater role in determining the how faithful the approximation is to the actual system.

The variation of constants transformation that was used in this example is a well-known technique in the area of quantum physics and electrodynamics; it goes by the name of the rotating-wave approximation. Averaging of the oscillatory dynamics due to a magnetic field has also been studied by Littlejohn [89, 90, 91] under the name of *guiding center theory*. Although the resulting averages are the same, the two techniques differ. Littlejohn utilizes a sequence of Lie transforms, based on a proof of the Darboux theorem, to obtain

cyclic coordinates that correspond to the periodic symmetries. Averaging is then reduction of the cyclic symmetries; see also Omohundro [124].

2.6.2 Axial Crystal Lattice Channeling

This example follows from Dumas and Ellison [39], and models a particle channeling through a planar crystal lattice. The three-dimensional case is more involved, but does not provide any additional insight [40]. Understanding the mechanism behind crystal channeling is relevant in particle physics, see Dumas et al. [40] and references therein.

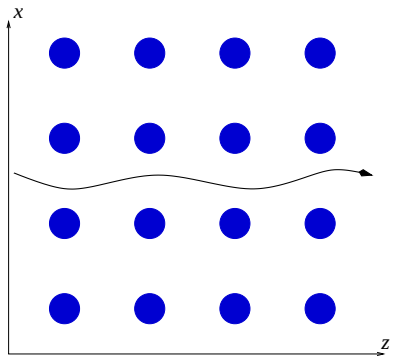


Figure 2.3: Planar channeling in a crystal lattice.

The case of axial planar channeling involves evolution on the cotangent bundle, $T^*\mathbb{R}^2$. The coordinates of the phase space are, $(x, z, p_x, p_z) \in T^*\mathbb{R}^2$. The vector field describing the equations of motion for the simplified planar model is

$$X = \left\{ \begin{array}{c} \epsilon p_x \\ p_z \\ -\frac{1}{2}\epsilon \frac{\partial W}{\partial x}(x,z) \\ -\frac{1}{2}\epsilon^2 \frac{\partial W}{\partial z}(x,z) \end{array} \right\}, \tag{2.54}$$

where

$$W(x,z) = (A + \frac{1}{2}Bx^2) + (A + \frac{1}{2}Cx^2)\cos(2\pi z).$$

The spatial-periodicity of the potential W arises from the spatial regularity of the crystal lattice. The vector field in Equation (2.54), as obtained from [39], has been non-dimensionalized. The non-dimensionalization process introduces the parameter ϵ, which for a $1\mathrm{MeV}$ proton is approximately, $\epsilon = 0.01$.

The equations of motion do not have time-periodic terms, however, the transformation method discussed in Section 2.5.2 may be applied to (2.54). The vector field X can be decomposed into two parts,

$$X = X_1 + \epsilon X_2 = \left\{ \begin{array}{c} 0 \\ p_z \\ 0 \\ 0 \end{array} \right\} + \epsilon \left\{ \begin{array}{c} p_z \\ 0 \\ -\frac{1}{2}\frac{\partial W}{\partial x}(x,z) \\ -\frac{1}{2}\epsilon \frac{\partial W}{\partial z}(x,z) \end{array} \right\}.$$

The flow corresponding to X_1 is

$$\Phi_{0,t}^{X_1} = \mathrm{Id} + \left\{ \begin{array}{c} 0 \\ p_z t \\ 0 \\ 0 \end{array} \right\}.$$

The pull-back flow is found to be

$$Y = \left(\Phi_{0,t}^{X_1}\right)^* \epsilon X_2(q) = \epsilon \left\{ \begin{array}{c} p_x \\ \frac{1}{2}\epsilon t \frac{\partial W}{\partial z}(x, z + p_z t) \\ -\frac{1}{2}\frac{\partial W}{\partial x}(x, z + p_z t) \\ -\frac{1}{2}\epsilon \frac{\partial W}{\partial z}(x, z + p_z t) \end{array} \right\} = \left\{ \begin{array}{c} \epsilon p_x \\ -\pi\epsilon^2(A + \frac{1}{2}Cx^2)t\sin(2\pi z + 2\pi p_z t) \\ -\frac{1}{2}\epsilon(B + C\cos(2\pi z + 2\pi p_z t))x \\ \pi\epsilon^2(A + \frac{1}{2}Cx^2)\sin(2\pi z + 2\pi p_z t) \end{array} \right\}. \tag{2.55}$$

The flow of X is obtained via the composition,

$$\Phi_{0,t}^{X} = \Phi_{0,t}^{X_1} \circ \Phi_{0,t}^{Y}.$$

Due to the spatial regularity of W in the $\hat{z}$-direction, Assumption 1 (possibility 2) is almost satisfied. Although the time-dependent vector field Y has components that are periodic in time, there is one component of the vector field that also has linear growth in time. Although Y is not periodic, the vector field may still be averaged.

After first-order averaging, the flow of the vector field Y is approximated by the vector field

$$Z = \left\{ \begin{array}{c} \epsilon p_x \\ \frac{1}{2}\epsilon^2 \frac{\cos(2\pi z)}{p_z}(A + \frac{1}{2}Cx^2) \\ -\frac{1}{2}\epsilon B x \\ 0 \end{array} \right\}. \tag{2.56}$$

The averaged vector field, Z, predicts nominally stable oscillatory behavior in the (x, p_x) phase-space. It also predicts a slight increase in the z-velocity, with no such increase in the z-momenta. To obtain the Floquet mapping, the improved first-order truncation of $P(t)$ is needed. According to Table 2.3, $\mathrm{Trunc}_1\left(P(t)\right)$ is constructed using the composed truncated flows of $-Z$ and Y, which are

$$\begin{aligned} \overrightarrow{\exp}\left(\int_0^t Y_\tau \, \mathrm{d}\tau\right) \approx\ & \mathrm{Id} \\ & + \left\{ \begin{array}{c} \epsilon p_x t \\ \frac{1}{2p_z}\epsilon^2 \left(A + \frac{1}{2}Cx^2\right)\left(t\cos(2\pi z + 2\pi p_z t) - \frac{1}{2\pi p_z}\left(\sin(2\pi z + 2\pi p_z t) - \sin(2\pi z)\right)\right) \\ -\frac{1}{2}\epsilon B x t - \frac{1}{4\pi p_z}\epsilon C x \sin(2\pi z + 2\pi p_z t) \\ \frac{1}{2p_z}\epsilon^2 \left(A + \frac{1}{2}Cx^2\right)\left(1 - \cos(2\pi z + 2\pi p_z t)\right) \end{array} \right\}, \end{aligned}$$

and

$$\exp\left(-Zt\right) \approx \mathrm{Id} + \left\{ \begin{array}{c} -\epsilon p_x t \\ \frac{1}{2p_z}\left(A + \frac{1}{2}Cx^2\right) t\cos(2\pi z) \\ \frac{1}{2}\epsilon B x t \\ 0 \end{array} \right\}.$$

Combined, they result in a first-order approximation to the Floquet mapping of the almost-periodic vector field (2.55),

$$P(t) \approx \mathrm{Id} + \left\{ \begin{array}{c} 0 \\ \frac{1}{2p_z}\epsilon^2(A + \frac{1}{2}Cx^2)\left(t\cos(2\pi z + 2\pi p_z t) - t\cos(2\pi z) - \frac{1}{2\pi p_z}\left(\sin(2\pi z + 2\pi p_z t) - \sin(2\pi z)\right)\right) \\ -\frac{1}{4\pi p_z}\epsilon C x \sin(2\pi z + 2\pi p_z t) \\ -\frac{1}{2p_z}\epsilon^2(A + \frac{1}{2}Cx^2)\left(1 - \cos(2\pi z + 2\pi p_z t)\right) \end{array} \right\}$$

The actual Floquet mapping for (2.54) is

$$\widetilde{P}(t) = \Phi_{0,t}^{X_1} \circ P(t)$$

Figure 2.4 shows a simulation of the evolution of (2.54) for $A = 1$, $B = 1$, $C = \frac{1}{3}$. These parameters were arbitrarily chosen, and do not necessarily reflect the true values for axial channeling. The ϵ parameter is $\epsilon = 0.01$, corresponding to a $1\mathrm{MeV}$ proton as discused earlier. The initial condition is $q_0 = (10^{-8}, 0, 0, 1)$. Figure 2.5 plots the error of the evolution reconstructed from the flow of the averaged system (2.56) composed with the Floquet mapping $\widetilde{P}(t)$, as compared to the unaverged evolution of Equation (2.54).

The error is largest for the subsystem where the time-dependent vector field Y (2.55) fails to be periodic, e.g. (z, p_z); it is $O\left(\epsilon^2 t\right)$ in the z-state and $O\left(\epsilon^2\right)$ in the p_z state. The time-periodic subsystem (x, p_x) is order $O\left(\epsilon^2\right)$ close (after normalizing by the amplitude of oscillations) for the time course of the simulation. The solid property of some of the plots is due to oscillations. Figure 2.6 shows the drifting z velocity. According to the averaged vector field (2.56), the slope of the drift should be about $5 * 10^{-5}$, whereas in actuality it is $1.56 * 10^{-5}$, a slight over-estimate by a factor of 3.

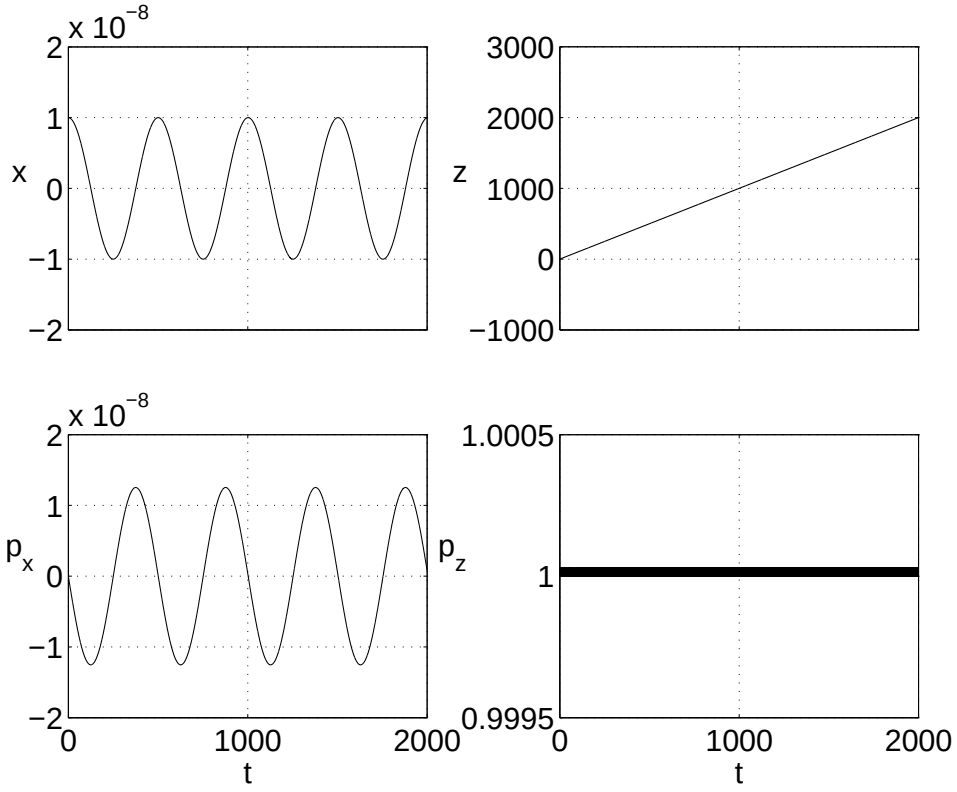


Figure 2.4: Axial channeling plots.

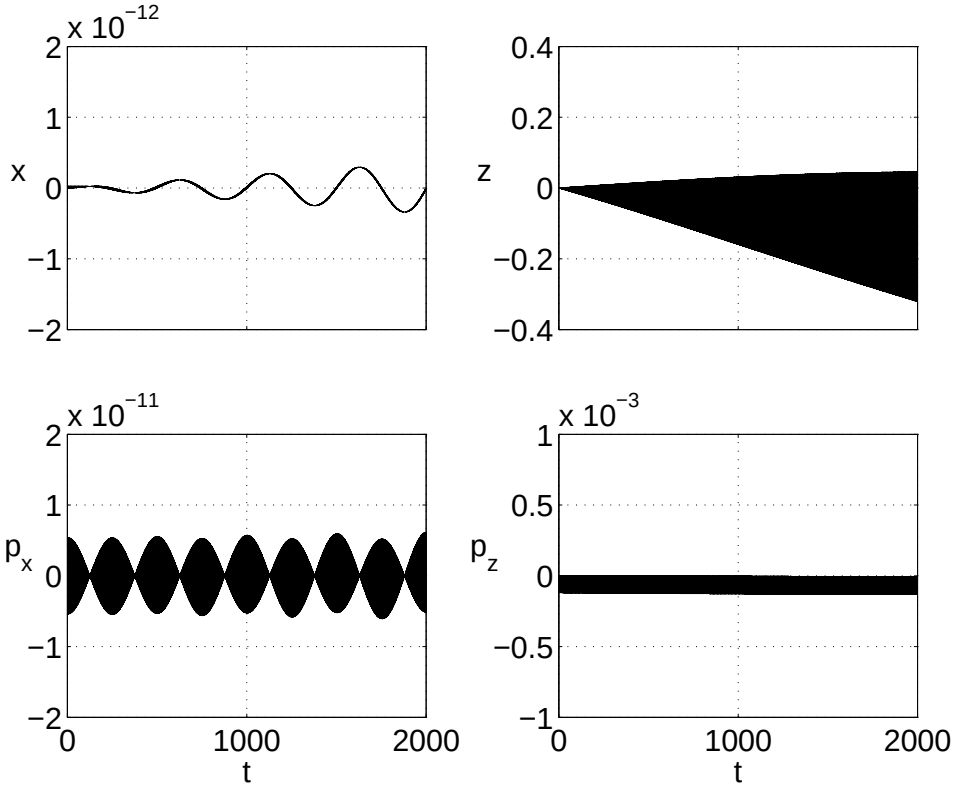


Figure 2.5: Error between evolution of the averaged and actual systems.

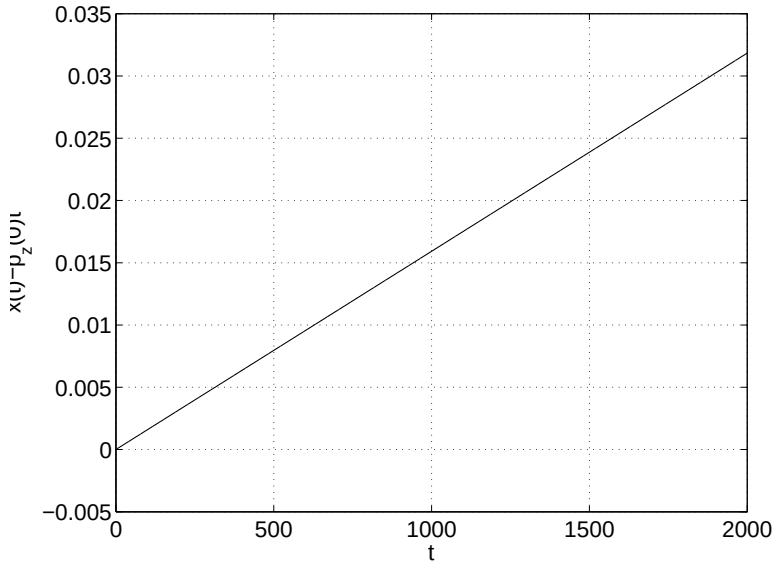


Figure 2.6: Net drift in z-velocity, $z(t) - p_z(0)t$.

2.6.3 Observer Using Coupled Oscillators

This example is a continuous finite dimensional analog of two coupled quantum oscillators, which are idealized as two spring-mass systems. In quantum electrodynamics jargon, the analysis that will be done for this system is known as the rotating-wave approximation, which is obtained by first moving to the interaction picture. The interaction picture corresponds to performing the variation of constants transformation to the system.

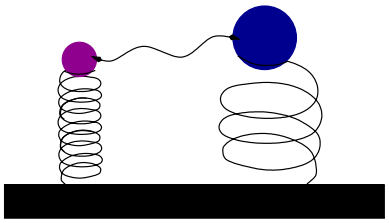


Figure 2.7: Coupled oscillators.

The configuration space is the cotangent bundle, $T^*\mathbb{R}^2$, with coordinates (x_1, x_2, p_1, p_2). The system can be described using a Hamiltonian in canonical coordinates. The Hamiltonian contains two terms, Λ_1 and Λ_2, which couple the dynamics of the two oscillators,

$$H(z) = \frac{1}{2m_1}p_1^2+\frac{1}{2m_2}p_2^2+\frac{1}{2}m_1\omega_1^2x_1^2+\frac{1}{2}m_2\omega_2^2x_2^2+\Lambda_1x_1x_2+\Lambda_2x_1^2x_2^2,$$

where m_1 and m_2 represent mass, and ω_1 and ω_2 represent the natural frequencies of the oscillators. The Hamiltonian vector field is

$$X_H = \begin{Bmatrix} \frac{1}{m_1}p_1 \\ \frac{1}{m_2}p_2 \\ -m_1\omega_1^2x_1 - \Lambda_1x_2 - 2\Lambda_2x_1x_2^2 \\ -m_2\omega_2^2x_2 - \Lambda_1x_1 - 2\Lambda_2x_1^2x_2 \end{Bmatrix}.$$

The system parameters are such that $\omega_1 >> \omega_2$ and $m_1 \leq m_2$, i.e., the time-scale of the dynamics for the first state are much faster than the second state. We'd like to understand how the second state, with much slower oscillatory evolution, is affected by the first state and, quite possibly, vice versa. Decompose the equations of motion into primary and perturbative components,

$$X_H = X_1 + X_2 = \begin{Bmatrix} \frac{1}{m_1}p_1 \\ 0 \\ -m_1\omega_1^2x_1 \\ 0 \end{Bmatrix} + \begin{Bmatrix} 0 \\ \frac{1}{m_2}p_2 \\ -\Lambda_1x_2 - 2\Lambda_2x_1x_2^2 \\ -m_2\omega_2^2x_2 - \Lambda_1x_1 - 2\Lambda_2x_1^2x_2 \end{Bmatrix},$$

both of which are still Hamiltonian flows. Although, there is not an ϵ parameter, we may still perform the decomposition. The flow of the primary vector field, X_1, is

$$\Phi_{0,t}^{X_1} = \begin{Bmatrix} q_1(0)\cos(\omega_1 t) + \frac{1}{m_1\omega_1}p_1(0)\sin(\omega_1 t) \\ q_2(0) \\ p_1(0)\cos(\omega_1 t) + m_1\omega_1 q_1(0)\sin(\omega_1 t) \\ p_2(0) \end{Bmatrix}. \tag{2.57}$$

The time-dependent vector field Y is given by

$$Y = \left(\Phi_{0,t}^{X_1}\right)^* X_2,$$

which is actually time-periodic since the flow $\Phi_{0,t}^{X_1}$ is time-periodic. Therefore, Assumption 1 (posibility 1)

is satisfied. After computing the pull-back,

$$Y = \left\{ \begin{array}{c} \frac{\Lambda_1}{m_1\omega_1} x_2 \sin(\omega_1 t) + \frac{2\Lambda_2}{m_1\omega_1} x_2^2 \left(x_1 \cos(\omega_1 t)\sin(\omega_1 t) + \frac{1}{m_1\omega_1} p_1 \sin^2(\omega_1 t) \right) \\ \frac{1}{m_2} p_2 \\ -\Lambda_1 x_2 \cos(\omega_1 t) - 2\Lambda_2 x_2 \left(x_1 \cos^2(\omega_1) t + \frac{1}{m_1\omega_1} p_1 \cos(\omega_1 t)\sin(\omega_1 t) \right) \\ -m_2\omega_2^2 x_2 - \Lambda_1 \left(x_1 \cos(\omega_1 t) + \frac{1}{m_1\omega_1} p_1 \sin(\omega_1 t) \right) - 2\Lambda_2 x_2 \left(x_1 \cos(\omega_1 t) + \frac{1}{m_1\omega_1} p_1 \sin(\omega_1 t) \right)^2 \end{array} \right\}.$$

Because it is assumed that the oscillations of the first state are fast, the system is a highly oscillatory system as per Section 2.5.3. The frequency ω_1 may be used in lieu of the parameter, ϵ. In other words, $\epsilon = \frac{1}{\omega_1}$. The first-order average is

$$Z = \left\{ \begin{array}{c} \frac{\Lambda_2}{2m_1^2\omega_1^2} x_2^2 p_1 \\ \frac{1}{m_2} p_2 \\ -\frac{\Lambda_2}{2} x_2^2 x_1 \\ -m_2\omega_2^2 x_2 - \frac{2\Lambda_2}{m_1\omega_1^2} \left(\frac{1}{2} m_1\omega_1^2 x_1^2 + \frac{1}{2m_1} p_1^2 \right) x_2 \end{array} \right\}.$$

To first-order, the Λ_1 term has no effect, whereas the Λ_2 term introduces some slow oscillatory behavior into the (x_1, p_1) phase space. Additionally, the natural frequency of the second oscillator, (x_2, p_2), is adjusted by a value on the order of the energy of the first oscillator. By observing the frequency of the (x_2, p_2) phase space oscillations, it should be possible to determine the energy of the first oscillator, (x_1, p_1).

Moving to the second-order averaged expansion, the effect of Λ_1 may be seen. The second-order average is

$$Z = \left\{ \begin{array}{c} \frac{\Lambda_2}{2m_1^2\omega_1^2} x_2^2 p_1 - \frac{\Lambda_1}{m_1 m_2 \omega_1^2} p_2 - \frac{\Lambda_2}{2m_1 m_2 \omega_1^2} x_2 p_2 x_1 - \frac{3\Lambda_2^2}{8m_1^3\omega_1^4} x_2^4 \\ \frac{1}{m_2} p_2 - \frac{\Lambda_1}{m_1 m_2 \omega_1^2} p_1 - \frac{\Lambda_2}{2m_1 m_2 \omega_1^2} x_1 p_1 \\ -\frac{\Lambda_2}{2} x_2^2 x_1 + \frac{\Lambda_2}{2m_1 m_2 \omega_1^2} x_2 p_2 p_1 - \frac{\Lambda_1\Lambda_2}{2m_1\omega_1^2} x_2^3 - \frac{\Lambda_2^2}{8m_1\omega_1^2} x_2^4 x_1 \\ -m_2\omega_2^2 x_2 - \frac{2\Lambda_2}{m_1\omega_1^2} \left(\frac{1}{2} m_1\omega_1^2 x_1^2 + \frac{1}{2m_1} p_1^2 - \frac{\Lambda_1^2}{2\Lambda_2} \right) x_2 - \frac{3\Lambda_1\Lambda_2}{2m_1\omega_1^2} x_2^2 x_1 \\ + \frac{3\Lambda_2^2}{8m_1^3\omega_1^4} (1 + x_2) p_1^2 x_2^2 - \frac{\Lambda_2^2}{8m_1\omega_1^2} (1 + x_2) x_1^2 x_2^2 \end{array} \right\}.$$

A non-vanishing Λ_1 will shift the frequency of the second state by the amount $\Lambda_1^2/(2\Lambda_2)$. When observing the frequency of the second state, it will be necessary to correct for the frequency shift, otherwise the energy estimate will be biased.

Figure 2.8 shows the actual and observed kinetic energies of the fast ocillator, for the parameters, $m_1 = \frac{1}{2}$, $m_2 = \frac{1}{2}$, $\omega_1 = 10$, $\omega_2 = 1$, $\Lambda_1 = 0.25$, and $\Lambda_2 = 1$. At random intervals, the first oscillator receives a kick of energy. The energy observations are made by timing the delay between zero crossings of the second oscillator. The error of the observer is plotted in Figure 2.9. It is of order $O(\epsilon)$, where $\epsilon = \frac{1}{\omega_1} = \frac{1}{10}$, except during the interval when an energy shift occurs. The delay is because the frequency is observed using zero crossings of the state x_2. A better technique capable of continuously observing the frequency, and more robust to noise, would be to examine the evolution of the state x_2 in the frequency domain.

Because the Floquet mapping was not calculated, there are missing parts of the dynamics which account for much of the error in estimating the energy using the second oscillator, (x_2, p_2). Without the Floquet mapping, $P(t)$, it will not be possible to improve the error of the estimate beyond $O(\epsilon)$; see the comments of Section 2.4.2. For quantum systems, knowledge of $P(t)$ is not possible due to the uncertainty in the states of the first oscillator. Consequently, there may be a physical limit to how well the energy can be observed using a coupled oscillator.

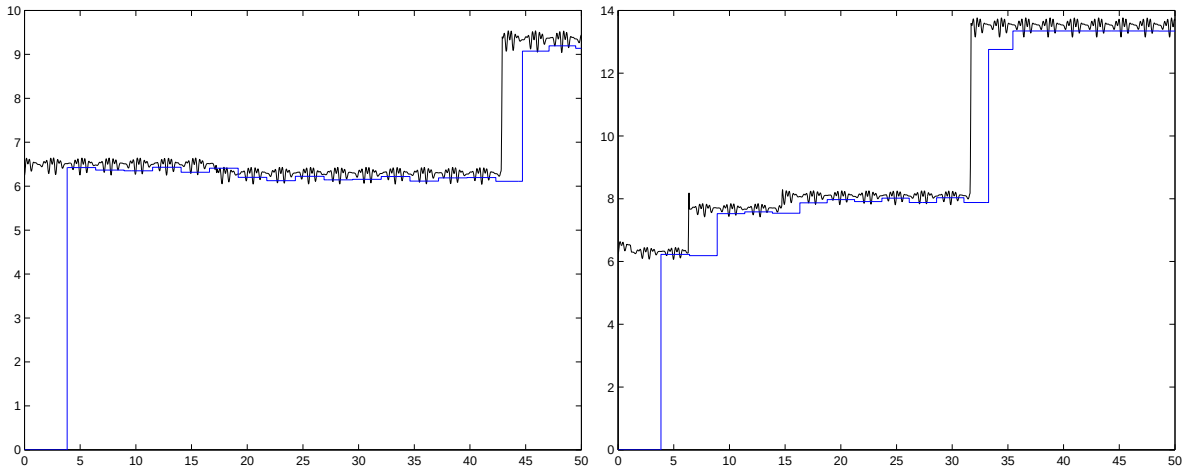


Figure 2.8: Actual and observed energies for two simulations.

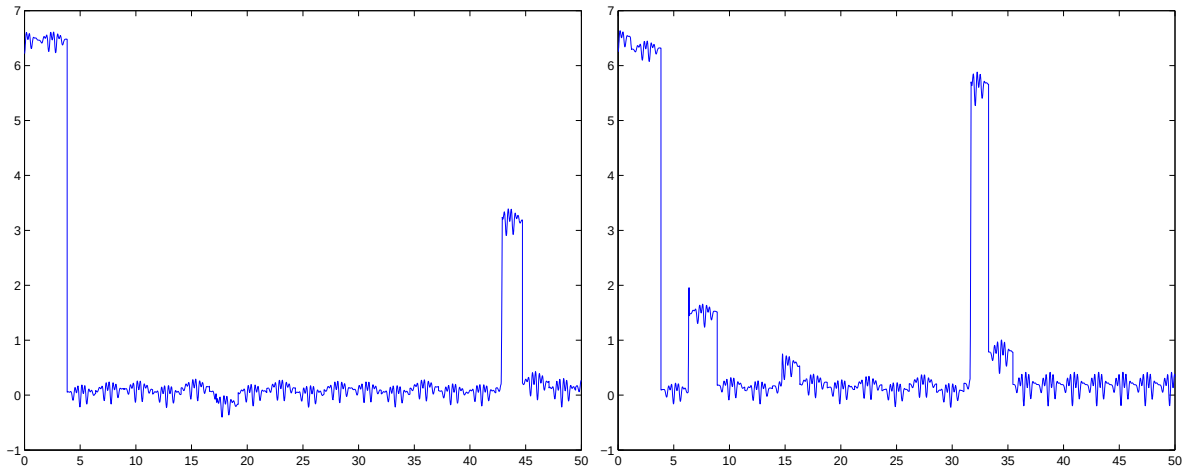


Figure 2.9: Error between observed and actual energy for two simulations.

2.7 Averaging Theory and Dynamical Systems Theory

The averaging theory described in this chapter connects to fundamental theorems concerning dynamical systems theory. Most importantly, it is related to the embedding of (homeo-)diffeomorphisms in flows generated by vector fields. Define $\mathrm{Diff}(M)$ to be the set of diffeomorphisms of M into M, and also define $\mathrm{Diff}^r(M) \subset \mathrm{Diff}(M)$ to consist of C^r diffeomorphisms. It is known that diffeomorphisms embed into the flows of time-varying vector fields [147]. It is more difficult to ascertain conditions for which a diffeomorphism may be embedded into the flow of an autonomous vector field. The logarithm computation of Equation (2.29) provides a means to convert the flow of a smooth time-varying time varying vector field into the flow of a smooth autonomous vector field, providing a mechanism by which smooth diffeomorphisms embed into the flows of smooth autonomous vector fields. Unfortunately, the dynamical systems literature appears to contradict this possibility.

The diffeomorphism, $f \in \mathrm{Diff}(M)$ *embeds into a flow*, if there exists a vector field X and a finite time t_f such that

$$f = \Phi_{0,t_f}^X. \tag{2.58}$$

The *embedding problem* is defined by the question: Does the diffeomprohism $f \in \mathrm{Diff}(M)$ embed into a flow? If a diffeomorpism, f, may be embedded into the flow of an autonomous vector field, then f has a large, non-trivial centralizer. The centralizer of a diffeomorphism, f, is defined to be the set of diffeomorphisms that commute with f,

$$\mathcal{C}(f) \equiv \{\, g \in \mathrm{Diff}(M) \mid gf = fg \,\}.$$

A trivial centralizer has only iterates of f in it, i.e. f^k for $k \in \mathbb{Z}$. Define the trivial centralizer, $\mathcal{C}^0$, to be

$$\mathcal{C}^0(f) \equiv \left\{\, f^k \mid k \in \mathbb{Z} \,\right\}. \tag{2.59}$$

Smale posed the question of whether the elements of $\mathrm{Diff}(M)$ with trivial centralizers form an open and dense subset, i.e., how many $f \in \mathrm{Diff}(M)$ have $\mathcal{C}(f) = \mathcal{C}^0(f)$? Work by Kopell [73] has shown that for diffeomorphisms on the circle, the subset of diffeomorphisms with trivial centralizers is open and dense. Further work by Palis and Yoccoz [128, 129, 130] has demonstrated that this property is not unique to the circle. Although the general problem is still open, these results do imply that the logarithm should not exist for the finite-time flow of most nonlinear time-varying vector fields (in the linear setting, there is no problem).

Given that arbitrary diffeomorphisms may not embed into flows, it is natural to seek conditions under which a diffeomorphism may embed into a flow. It has been shown that if a diffeomorphism embeds in a flow, then it may do so for uncountably many flows [44, 128]. Consequently, there may not be a unique vector field whose flow at a fixed time will result in a given diffeomorphism (this freedom will be seen in future chapters dealing with the construction of stabilizing controllers for underactuated nonlinear control systems). Uniqueness aside, the primary difficulty of the embedding appears to lie near fixed points of diffeomorphisms [129]. Transversality of the flow and non-resonance of the linear eigenvalues of the vector field play an important role in ensuring embeddability [88, 130]. Most proof techniques giving conditions for embeddability revolve around hyperbolic fixed points and involve transversality conditions [97, 128].

For diffeomorphisms that do not have a center manifold in the linearization, there are positive results concerning the embedding problem [5, 88]. The proofs involve extending standard linear Floquet theory to the nonlinear setting; nonresonance of the eigenvalues allows the extension of linear Floquet theory to follow through. In the future chapters of this dissertation, we will explicitly incorporate resonance to excite nonlinear coupling. The goal of the nonlinear coupling is to achieve a response in the center manifold that can be used for control-theoretic purposes; oscillatory forcing is used to create the coupling. Therefore, the approach of this dissertation is in contrast to the prevailing approaches found in the body of literature dealing with the embedding problem via Floquet theory. Recent work by Morin and Samson [119, 120] has demonstrated that oscillatory transverse actuation is important for the stabilization of underactuated nonlinear control systems, which are the focus of this thesis. Although dynamical systems theory indicates that our analysis should not be possible for a large class of systems, the solutions found in this dissertation appear to be inline with the known positive results in the literature.

In spite of the negative results, research suggests that diffeomorphisms may be asymptotically embedded into flows [58, 73], i.e., given an ϵ, there exists a vector field whose flow after a fixed time is a diffeomorphism ϵ-close to the given diffeomorphism. Perhaps, the limit fails to be in the acceptable set of vector fields, but still results in the appropriate flow. The fact that a property may not be preserved by limit processes is not new. Arnold and Khesin [6] discuss a problem whose solutions ϵ-close to the limit solution are smooth, yet in the limit of $\epsilon \to 0$, the solution fails to be smooth. The practical implications are that, for certain cases, the logarithm computation is close in an asymptotic sense, as emphasized by Agrachev and Gamkrelidze [2].

Such a situation is also eluded to by the works of Epstein, Mather, and Thurston [43, 104, 158], who

show that the subgroup of diffeomorphisms that are isotopic to the identity, $\mathrm{Diff}_0^r(M) \subset \mathrm{Diff}^r(M)$, is simple. Here $r >= n+2$, where $n = \dim(M)$. Simplicity is related to the embedding problem [43]. Under weaker hypotheses, or in the appropriate sense, the fixed time maps of flows may generate all diffeomorphisms isotopic to the identity. This supports the idea that the logarithm may be used to obtain an autonomous vector field with the property that its fixed-time flow asymptotically approximates the diffeomorphism obtained from the fixed-time flow of a time-varying vector field sufficiently close to the identity flow, as per Proposition 2.

2.8 Conclusion

One goal of this chapter was to generalize averaging theory to arbitrary order, and in the process prove associated theorems related to stability under averaging. Additionally, a new paradigm for thinking about averaging theory was presented. Averaging theory was shown to come from a synthesis of two distinct theories: (1) Floquet theory, and (2) perturbation theory. By working within this framework many of the properties of classical and KBM averaging methods can be recovered. Understanding how averaging theory is really a synthesis of the two aforementioned theories may lead to a better understanding of what the process of averaging physically means for dynamical systems.

Although the theory was capable of recovering known results in averaging theory, the framework used is restricted to smooth or analytic vector fields, and provided conservative estimates. The works of Ellison et. al. [42] and others have weaker requirements on the vector fields and stronger estimates, but use a perturbative approach. Further investigation into the similarities between the perturbative approach and Volterra series expansions could serve to relax this requirement. The approach of Sussmann and Kawski [65], which provides alternative derivations of the exponential representation, may be favored over the derivation of Agrachev and Gamkrelidze [2].

Given that Floquet theory neatly decomposes time-periodic systems into two elements, a Floquet mapping and an autonomous flow, it may be possible to couch averaging as a reduction technique where the symmetry is time periodicity. Due to the time-dependent nature of composed flows, such an endeavour may not fit squarely within the context of Lie groups, but may rely on the more general concept of Lie groupoids. The autonomous averaged vector field is then the corresponding Lie algebroid.

Additional elements that comprise averaging theory have yet to be explicitly placed within this framework, of which we mention the development of two-timing methods and the analysis of resonance or multiphase averaging. An important extension would also be to partial differential equations. The exponential representation of flows in Agrachev and Gamkrelidze [2] appears to support this goal, although not without difficulty.

The averaging analysis of this chapter leads to several topics within the realm of nonlinear control and dynamics as originally envisioned by Agrachev, Gamkrelidze, and Sarychev [4]. In particular the utilization of this theory for nonlinear time-periodic control systems. The remaining chapters of this dissertation address this issue. A natural consequence is the use of the generalized averaging theory to analyze resonance for nonlinear control systems. Resonant oscillatory forcing of underactuated control systems will lead to flow that would not seem possible from analysis of the linearized system.

2.A Sanders and Verhulst Revisited

Recall that w and $\tilde{f}$ are related to each other via integration in time as per Theorem 5. We obtain an alternative expression for $\bar{g}(z)$,

$$\begin{aligned}
\bar{g}(z) &= \frac{1}{T}\int_0^T \left(\mathrm{D}_z f(z,t)\cdot w(z,t) - \mathrm{D}_z w(z,t)\cdot \bar{f}(z)\right)\mathrm{d}t \\
&= \frac{1}{T}\int_0^T \left(\mathrm{D}_z f(z,t)\cdot\left(\int_0^\tau (f(z,\tau)-\bar{f}(z))\mathrm{d}\tau\right) - \mathrm{D}_z\left(\int_0^t (f(z,\tau)-\bar{f}(z))\mathrm{d}\tau\right)\cdot\bar{f}(z)\right)\mathrm{d}t \\
&= \frac{1}{T}\int_0^T \Bigg(\mathrm{D}_z f(z,t)\cdot\int_0^t f(z,\tau)\mathrm{d}\tau - \mathrm{D}_z f\cdot\int_0^t \bar{f}(z)\mathrm{d}\tau \\
&\qquad\qquad - \mathrm{D}_z\left(\int_0^\tau f(z,\tau)\mathrm{d}\tau\right)\cdot\bar{f}(z) + \mathrm{D}_z\left(\int_0^\tau f(z,\tau)\mathrm{d}\tau\right)\cdot\bar{f}(z)\Bigg)\mathrm{d}t \\
&= \frac{1}{T}\int_0^T \Bigg(\mathrm{D}_z f(z,t)\cdot\int_0^t f(z,\tau)\mathrm{d}\tau - \mathrm{D}_z f(z,t)\cdot\bar{f}(z)t \\
&\qquad\qquad + \mathrm{D}_z f(z,t)\cdot\bar{f}(z)t + \mathrm{D}_z\bar{f}(z)t\cdot\bar{f}(z)\Big) - T\,\mathrm{D}_z\bar{f}(z)\cdot\bar{f}(z) \\
&= \frac{1}{T}\int_0^T \mathrm{D}_z f(z,t)\cdot\int_0^\tau f(z,\tau)\mathrm{d}\tau - \frac{T}{2}\,\mathrm{D}_z\bar{f}(z)\cdot\bar{f}(z) \\
&= \frac{1}{2T}\int_0^T \left(\mathrm{D}_z f(z,t)\cdot\int_0^\tau f(z,\tau)\mathrm{d}\tau - \mathrm{D}_z\int_0^t f(z,\tau)\mathrm{d}\tau\cdot f(z,t)\right)\mathrm{d}t,
\end{aligned}$$

where for now the integration constant is not included. The form of the last equation is the coordinate form of the Jacobi-Lie bracket, meaning that $\bar{g}(z)$ actually takes the simpler form below

$$\bar{g}(z) = \frac{1}{2T}\int_0^T \left[\int_0^t f(z,\tau)\mathrm{d}\tau, f(z,t)\right]\mathrm{d}t = \frac{1}{2}\overline{\left[\int_0^t f(z,\tau)\mathrm{d}\tau, f(z,t)\right]}. \tag{2.60}$$

Using, instead, the actual equation for w from Theorem 5 with the additional integration constant $a(z)$,

$$\bar{g}(z) = \frac{1}{2}\overline{\left[\int_0^t f(z,\tau)\mathrm{d}\tau, f(z,t)\right]} + \left[a(z), \overline{f}(z)\right]. \tag{2.61}$$

Alternatively,

$$\bar{g}(z) = \frac{1}{2}\overline{\left[\int_0^t \left(f(z,\tau)-\bar{f}(z)\right)\mathrm{d}\tau + a(z), f(z,t)\right]}. \tag{2.62}$$

A similar conclusion to (2.60) was found by Kurzweil and Jarník [74] when examining the limit of highly oscillatory differential equations, and subsequently generalized to arbitrary order in [75]. The method used resembles that of a Volterra series expansion, and was given further study by Sussman and Liu [154, 92, 93].

2.B Computation of Logarithm Series Expansions

The integrand of the third variation is obtained from the following derivation

$$\mathbf{D}(\mathrm{ad}\xi_3)\mathbf{D}(\mathrm{ad}\xi_2)\xi_1.$$

It's expansion according to the rule of differentiation is

$$\begin{aligned}
&\mathbf{D}(\mathrm{ad}\xi_3)\mathbf{D}(\mathrm{ad}\xi_3)\xi_1\\
=&\\
&\left(\mathbf{D}(\mathrm{ad}\xi_3)\mathrm{ad}\right)\xi_2\xi_1 + \mathrm{ad}\left(\mathbf{D}(\mathrm{ad}\xi_3)\xi_2\right)\xi_1 + \mathrm{ad}\xi_2\mathbf{D}(\mathrm{ad}\xi_3\xi_1)\\
=&\\
&\mathrm{ad}\xi_3\mathrm{ad}\xi_2\xi_1 + \mathrm{adad}\xi_3\xi_2\xi_1 + \mathrm{ad}\xi_2\mathrm{ad}\xi_3\xi_1.
\end{aligned}$$

The depth of (ξ_1, ξ_2, ξ_3) in each of the regular words is $(2,0,0)$, $(1,1,0)$, and $(2,0,0)$, respectively, implying the coefficients $1/12$, $1/4$, and $1/12$, respectively. Thus the third variation integrand is

$$\mathrm{v}^{(3)}(\xi_1,\xi_2,\xi_3) = \frac{1}{12}\mathrm{ad}\,\xi_3\,\mathrm{ad}\,\xi_2\,\xi_1 + \frac{1}{4}\mathrm{ad}\,\mathrm{ad}\,\xi_3\,\xi_2\,\xi_1 + \frac{1}{12}\mathrm{ad}\,\xi_2\,\mathrm{ad}\,\xi_3\,\xi_1. \tag{2.63}$$

For the fourth term and higher-order terms, the number of words begins to increase at a rapid rate. The fourth term will consist of 15 regular words. It begins as

$$\mathbf{D}(\mathrm{ad}\xi_4)\mathbf{D}(\mathrm{ad}\xi_3)\mathbf{D}(\mathrm{ad}\xi_2)\xi_1.$$

It's expansion according to the rules of differentiation is

$$\begin{aligned}
&\mathbf{D}(\mathrm{ad}\xi_4)\mathbf{D}(\mathrm{ad}\xi_3)\mathbf{D}(\mathrm{ad}\xi_2)\xi_1\\
=&\\
&\mathbf{D}(\mathrm{ad}\xi_4)\left(\mathrm{ad}\xi_3\mathrm{ad}\xi_2\xi_1 + \mathrm{adad}\xi_3\xi_2\xi_1 + \mathrm{ad}\xi_2\mathrm{ad}\xi_3\xi_1\right)\\
=&\\
&\mathrm{ad}\xi_4\mathrm{ad}\xi_3\mathrm{ad}\xi_2\xi_1 + \mathrm{adad}\xi_4\xi_3\mathrm{ad}\xi_2\xi_1 + \mathrm{ad}\xi_3\mathrm{ad}\xi_4\mathrm{ad}\xi_2\xi_1\\
&\quad + \mathrm{ad}\xi_3\mathrm{adad}\xi_4\xi_2\xi_1 + \mathrm{ad}\xi_3\mathrm{ad}\xi_2\mathrm{ad}\xi_4\xi_1\\
&+ \mathrm{ad}\xi_4\mathrm{adad}\xi_3\xi_2\xi_1 + \mathrm{adad}\xi_4\mathrm{ad}\xi_3\xi_2\xi_1 + \mathrm{adadad}\xi_4\xi_3\xi_2\xi_1\\
&\quad + \mathrm{adad}\xi_3\mathrm{ad}\xi_4\xi_2\xi_1 + \mathrm{adad}\xi_3\xi_2\mathrm{ad}\xi_4\xi_1\\
&+ \mathrm{ad}\xi_4\mathrm{ad}\xi_2\mathrm{ad}\xi_3\xi_1 + \mathrm{adad}\xi_4\xi_2\mathrm{ad}\xi_3\xi_1 + \mathrm{ad}\xi_2\mathrm{ad}\xi_4\mathrm{ad}\xi_3\xi_1\\
&\quad + \mathrm{ad}\xi_2\mathrm{adad}\xi_4\xi_3\xi_1 + \mathrm{ad}\xi_2\mathrm{ad}\xi_3\mathrm{ad}\xi_4\xi_1.
\end{aligned}$$

The depths of the fifteen words, in order, are

$$\begin{aligned}
&(3,0,0,0),\ (2,0,1,0),\ (3,0,0,0),\ (2,1,0,0),\ (3,0,0,0)\\
&(2,1,0,0),\ (1,2,0,0),\ (1,1,1,0),\ (1,2,0,0),\ (2,1,0,0)\\
&(3,0,0,0),\ (2,1,0,0),\ (3,0,0,0),\ (2,0,1,0),\ (3,0,0,0)
\end{aligned}$$

Since the Bernoulli numbers vanish for the odd numbers greater than 1, those with depth 3 have vanishing coefficients. Combining the remaining terms gives

$$\begin{aligned}\mathrm{v}^{(4)}(\xi_1,\xi_2,\xi_3\xi_4) = &-\frac{1}{24}\mathrm{ad}(\mathrm{ad}\xi_4\xi_3)(\mathrm{ad}\xi_2\xi_1) - \frac{1}{24}\mathrm{ad}\xi_3(\mathrm{ad}(\mathrm{ad}\xi_4\xi_2)\xi_1)\\ &-\frac{1}{24}\mathrm{ad}\xi_4(\mathrm{ad}(\mathrm{ad}\xi_3\xi_2)\xi_1) - \frac{1}{24}\mathrm{ad}(\mathrm{ad}\xi_4(\mathrm{ad}\xi_3\xi_2)\xi_1\\ &-\frac{1}{8}\mathrm{ad}(\mathrm{ad}(\mathrm{ad}(\xi_4\xi_3)\xi_2)\xi_1 - \frac{1}{24}\mathrm{ad}(\mathrm{ad}\xi_3(\mathrm{ad}\xi_4\xi_2))\xi_1\\ &-\frac{1}{24}\mathrm{ad}(\mathrm{ad}\xi_3\xi_2)(\mathrm{ad}\xi_4\xi_1) - \frac{1}{24}\mathrm{ad}(\mathrm{ad}\xi_4\xi_2)(\mathrm{ad}\xi_3\xi_1)\\ &-\frac{1}{24}\mathrm{ad}\xi_2(\mathrm{ad}(\mathrm{ad}\xi_4\xi_3)\xi_1).\end{aligned}$$

The remaining terms can be simplified by use of the Jacobi-Lie identity,

$$\begin{aligned}\mathrm{v}^{(4)}(\xi_1,\xi_2,\xi_3\xi_4) = -\frac{1}{12}\,&(\mathrm{ad}(\mathrm{ad}\xi_4\xi_3)(\mathrm{ad}\xi_2\xi_1) + \mathrm{ad}(\mathrm{ad}(\mathrm{ad}\xi_4\xi_3)\xi_2)\xi_1\\ &+\mathrm{ad}\xi_4(\mathrm{ad}(\mathrm{ad}\xi_3\xi_2)\xi_1) + \mathrm{ad}\xi_3(\mathrm{ad}(\mathrm{ad}\xi_4\xi_2))\xi_1)\,,\end{aligned}$$

which agrees with Agrachev and Gamkrelidze [2] (except for sign).

2.C The Averaged Expansions

Second-order. The first-order expansions of the Volterra series and the exponential in the periodic Floquet function are,

$$\begin{aligned}\mathrm{Trunc}_1\left(\exp\left(-Yt\right)\right) &= \mathrm{Id} - \epsilon\overline{X}\,t + O\left(\epsilon^2\right)\\ \mathrm{Trunc}_1\left(\overrightarrow{\mathcal{V}}_{0,t}\left(X_\tau\right)\right) &= \mathrm{Id} + \epsilon\int_0^t X_\tau\,\mathrm{d}\tau\, + O\left(\epsilon^2\right).\end{aligned}$$

Taking the composition of the two pull-back flows, and truncating to first order, results in

$$\mathrm{Trunc}_1\left(P(t)\right) = \mathrm{Id} + \epsilon\int_0^t \left(X_\tau - \overline{X}\right)\,\mathrm{d}\tau\, + O\left(\epsilon^2\right).$$

Third-order. The truncations evaluate to

$$\mathrm{Trunc}_2\left(\overrightarrow{\exp}\left(\int_{t_0}^t X_\tau\,\mathrm{d}\tau\right)\right) = \mathrm{Id} + \int_{t_0}^t X_\tau\,\mathrm{d}\tau\, + \frac{1}{2}\int_{t_0}^t\int_{t_0}^\tau \left[X_s, X_\tau\right]\,\mathrm{d}s\,\mathrm{d}\tau\, + \frac{1}{2}\int_{t_0}^t X_\tau\,\mathrm{d}\tau\circ\int_{t_0}^t X_\tau\,\mathrm{d}\tau\,,$$

and

$$\mathrm{Trunc}_2\left(\exp\left(-Yt\right)\right) = \mathrm{Id} + \epsilon\overline{X} + \frac{1}{2}\epsilon^2\overline{\left[\int_0^t X(x,\tau)\,\mathrm{d}\tau\,, X(x,t)\right]} + \frac{1}{2}\epsilon^2\overline{X}\circ\overline{X}t^2,$$

whose composition and truncation to second order gives

$$\begin{aligned}\mathrm{Trunc}_2\left(P(t)\right) =& \mathrm{Id} + \epsilon\int_0^t \left(X_\tau - \overline{X}\right)\,\mathrm{d}\tau + \frac{1}{2}\epsilon^2\int_0^t\left(\left[\int_0^\tau X_s\,\mathrm{d}s\,,X_\tau\right] - \overline{\left[\int_0^t X_s\,\mathrm{d}s\,,X_t\right]}\right)\,\mathrm{d}\tau\\ &+\frac{1}{2}\epsilon^2\int_0^t X_\tau\,\mathrm{d}\tau\circ\int_0^t X_\tau\,\mathrm{d}\tau - \epsilon^2\int_0^t X_\tau\,\mathrm{d}\circ\overline{X}\,t + \frac{1}{2}\epsilon^2\overline{X}\circ\overline{X}\,t^2.\end{aligned}$$

Fourth-order. The fourth-order truncation of the flow is

$$\begin{aligned}\mathrm{Trunc}_3\left(\overrightarrow{\exp}\int_0^t X_\tau\,\mathrm{d}\tau\right) =& \mathrm{Id} + \epsilon\int_0^t X_\tau\,\mathrm{d}\tau + \frac{1}{2}\epsilon^2\int_0^t\left[\int_0^\tau X_s\,\mathrm{d}s\,,X_\tau\right]\,\mathrm{d}\tau\\ &+\frac{1}{4}\epsilon^3\left[\int_0^t X_\tau\,\mathrm{d}\tau\,,\int_0^t\left[\int_0^\tau X_s\,\mathrm{d}s\,,X_\tau\right]\,\mathrm{d}\tau\right]\\ &+\frac{1}{3}\epsilon^3\int_0^t\left[\int_0^\tau X_s\,\mathrm{d}s\,,\left[\int_0^\tau X_s\,\mathrm{d}s\,,X_\tau\right]\right]\,\mathrm{d}\tau\\ &+\frac{1}{2}\int_{t_0}^t X_\tau\,\mathrm{d}\tau\circ\int_{t_0}^t X_\tau\,\mathrm{d}\tau\\ &+\frac{1}{4}\epsilon^3\int_0^t X_\tau\,\mathrm{d}\tau\circ\int_0^t\left[\int_0^\tau X_s\,\mathrm{d}s\,,X_\tau\right]\,\mathrm{d}\tau\\ &+\frac{1}{4}\epsilon^3\int_0^t\left[\int_0^\tau X_s\,\mathrm{d}s\,,X_\tau\right]\,\mathrm{d}\tau\circ\int_0^t X_\tau\,\mathrm{d}\tau\\ &+\frac{1}{6}\epsilon^3\int_0^t X_\tau\,\mathrm{d}\tau\circ\int_0^t X_\tau\,\mathrm{d}\tau\circ\int_0^t X_\tau\,\mathrm{d}\tau\,,\end{aligned}$$

while the fourth order truncation of the autonomous flow is

$$\begin{aligned}\mathrm{Trunc}_3\left(\exp\left(-Yt\right)\right) =& \mathrm{Id} - \epsilon\overline{X}\,t - \frac{1}{2}\epsilon^2\overline{\left[\int_0^t X_\tau\,\mathrm{d}\tau\,,X_t\right]}\,t\\ &-\frac{1}{4}T\epsilon^3\left[\overline{X},\overline{\left[\int_0^t X_\tau\,\mathrm{d}\tau\,,X_t\right]}\right]\,t\\ &-\frac{1}{3}\epsilon^3\overline{\left[\int_0^t X_\tau\,\mathrm{d}\tau\,,\left[\int_0^t X_\tau\,\mathrm{d}\tau\,,X_t\right]\right]}t\\ &+\frac{1}{2}\epsilon^2\overline{X}\circ\overline{X}\,t^2 + \frac{1}{4}\epsilon^2\overline{X}\circ\overline{\left[\int_0^t X_\tau\,\mathrm{d}\tau\,,X_t\right]}\,t^2\\ &+\frac{1}{4}\epsilon^3\overline{\left[\int_0^t X_\tau\,\mathrm{d}\tau\,,X_t\right]}\circ\overline{X}\,t^2\\ &-\frac{1}{6}\epsilon^3\overline{X}\circ\overline{X}\circ\overline{X}\,t^3.\end{aligned}$$

Composing the two will result in the truncated Floquet mapping,

$$
\begin{aligned}
\mathrm{Trunc}_3\left(P(t)\right) =& \mathrm{Id} + \epsilon \int_0^t \left(X_\tau - \overline{X}\right)\,\mathrm{d}\tau \\
&+ \frac{1}{2}\epsilon^2 \int_0^t \left(\left[\int_0^\tau X_s\,\mathrm{d}s\,, X_\tau\right] - \overline{\left[\int_0^\sigma X_s\,\mathrm{d}s\,, X_\sigma\right]} \right)\,\mathrm{d}\tau \\
&+ \frac{1}{4}\epsilon^3 \left(\left[\int_0^t X_\tau\,\mathrm{d}\tau\,, \int_0^t \left[\int_0^\tau X_s\,\mathrm{d}s\,, X_\tau\right]\,\mathrm{d}\tau\right] - T\left[\overline{X}, \overline{\left[\int_0^t X_\tau\,\mathrm{d}\tau\,, X_t\right]}\right]\,t \right) \\
&+ \frac{1}{3}\epsilon^3 \int_0^t \left(\left[\int_0^\tau X_s\,\mathrm{d}s\,, \left[\int_0^\tau X_s\,\mathrm{d}s\,, X_\tau\right]\right] - \overline{\left[\int_0^t X_\tau\,\mathrm{d}\tau\,, \left[\int_0^t X_\tau\,\mathrm{d}\tau\,, X_t\right]\right]} \right)\,\mathrm{d}\tau \\
&+ \frac{1}{2}\epsilon^2 \int_0^t X_\tau\,\mathrm{d}\tau \circ \int_0^t X_\tau\,\mathrm{d}\tau - \epsilon^2 \int_0^t X_\tau\,\mathrm{d}\tau \circ \overline{X}\,t + \frac{1}{2}\epsilon^2 \overline{X} \circ \overline{X}\,t^2 \\
&+ \frac{1}{4}\epsilon^3 \int_0^t X_\tau\,\mathrm{d}\tau \circ \int_0^t \left[\int_0^\tau X_s\,\mathrm{d}s\,, X_\tau\right]\,\mathrm{d}\tau - \frac{1}{2}\epsilon^2 \int_0^t X_\tau\,\mathrm{d}\tau \circ \overline{\left[\int_0^t X_\tau\,\mathrm{d}\tau\,, X_t\right]}\,t \\
&- \frac{1}{2}\epsilon^3 \int_0^t \left[\int_0^\tau X_s\,\mathrm{d}s\,, X_\tau\right]\,\mathrm{d}\tau \circ \overline{X}\,t + \frac{1}{4}\epsilon^3 \overline{X} \circ \overline{\left[\int_0^t X_\tau\,\mathrm{d}\tau\,, X_t\right]}\,t^2 \\
&+ \frac{1}{4}\epsilon^3 \int_0^t \left[\int_0^\tau X_s\,\mathrm{d}s\,, X_\tau\right]\,\mathrm{d}\tau \circ \int_0^t X_\tau\,\mathrm{d}\tau + \frac{1}{4}\epsilon^3 \overline{\left[\int_0^t X_\tau\,\mathrm{d}\tau\,, X_t\right]} \circ \overline{X}\,t^2 \\
&+ \frac{1}{6}\epsilon^3 \left(\int_0^t X_\tau\,\mathrm{d}\tau \circ \int_0^t X_\tau\,\mathrm{d}\tau \circ \int_0^t X_\tau\,\mathrm{d}\tau\,, -\overline{X} \circ \overline{X} \circ \overline{X}\,t^3 \right).
\end{aligned}
$$

Chapter 3

Control of Underactuated Driftless Nonlinear Systems

A tremendous amount of research has gone into understanding controllability and determining conditions under which a system is controllable. There is still a gap, however, between the tests that determine controllability and the actual feedback laws that realize control. This chapter demonstrates how the "generalized averaging theory" may be used in conjunction with controllability tests to realize feedback control of underactuated driftless systems. The control laws will be exponentialy stabilizing in the average. The approach is an easily implementable and understandable strategy for designing exponentially stabilizing controllers for such systems.

The method does not need a homogeneous norm to demonstrate exponential stabilization, does not require complicated coordinate transformations, the construction of Lyapunov functions, nor the pre-existence of stabilizing controllers, although such elements can be used to facilitate the process. The complexity of the nonlinear analysis grows with the order of Lie bracketing. Only lower orders will be discussed. Extensions to higher orders follow the same principles, but may be more involved computationally.

Organization of This Chapter. The control of driftless systems is studied in Section 3.1, with a focus on using the results of averaging theory for control of underactuated driftless affine control systems. The particular class of systems with nonholonomic constraints is studied in Section 3.2, as well as the subclass of systems reducible via Lie group symmetries. Several examples are given in Section 3.3 demonstrating the utility and power of the control strategies derived using the generalized averaging theory. Section 3.4 concludes with a synopsis of the current results and a vision of future work.

3.1 Control of Driftless Systems

The standard form for an underactuated driftless affine control system is

$$\dot{q} = Y_a(q)u^a(q,t), \tag{3.1}$$

where $q \in Q$, with $Q \subset \mathbb{R}^n$ open, and $a = 1 \dots m < n \equiv \dim(Q)$ (full actuation occurs when $m = n$). The inputs u^a reside in $\mathcal{U}$, the set of functions piecewise smooth with respect to q, and absolutely continuous with respect to t. Controllability is defined using the sets of reachable states and accessible states at $q \in Q$. The reachable states at $q \in Q$ are

$$\mathcal{R}^V_{t_f}(q) \equiv \bigcup_{0 \le t \le t_f} A^V_{t_f}(q), \tag{3.2}$$

where,

$$A^V_{t_f}(q) \equiv \left\{ q_f \in Q \mid \exists\, u^a \in \mathcal{U}, a = 1 \dots m,\ q(0) = q,\ \Phi^{Y_a u^a}_{0,t_f}(q) = q_f\ :\ \Phi^{Y_a u^a}_{0,t}(q) \in V, \forall\, 0 \le t \le t_f \right\}, \tag{3.3}$$

is the set of states accessible (at time t_f) from q. To obtain the set of accessible states at $q \in Q$, take the union over all time,

$$A^V(q) \equiv \bigcup_{0 \le T \le \infty} A^V_T(q). \tag{3.4}$$

Definition 8 *The system (3.1) is controllable if $A(q) = Q$, for all $q \in Q$.*

Definitions that further refine the notion of controllability are *local accesibility* and *(small-time) local controllability*.

Definition 9 *The system (3.1) is* locally accessible *if for all $q \in Q$, $\mathcal{R}^V_T(q)$ contains a non-empty open set of Q for all neighborhoods V of q and all $T > 0$.*

The system (3.1) is (small-time) locally controllable (STLC) *if for any neighborhood V, time $T \ge 0$, and $q \in Q$, q is an interior point of $\mathcal{R}^V_T(x)$ for all $T \ge 0$.*

For underactuated driftless affine control systems, local controllability can be determined via the Lie Algebra Rank Condition (LARC), due to the equivalence of local accessibility and local controllability. To determine LARC, it is necessary to construct the involutive closure of the control vector fields. Let the set $\mathcal{C}^\infty$ be recursively constructed as follows

$$\begin{aligned} \mathcal{C}^1 &= \operatorname{span}\{Y_1, ..., Y_m\}, \text{ and} \\ \mathcal{C}^{k+1} &= \mathcal{C}^k + \left[\mathcal{C}^1, \mathcal{C}^k\right]. \end{aligned} \tag{3.5}$$

In practice, the sets $\mathcal{C}^k(q)$ become maximal at $\mathcal{C}^N(q)$ for some $N < \infty$ sufficiently large. The Lie Algebra Rank Condition at $q \in Q$ is satisfied if,

$$\dim \mathcal{C}^\infty(q) = \dim T_q Q. \tag{3.6}$$

Theorem 14 *The system (3.1) is locally accesible if the Lie Algebra Rank Condition (3.6) holds.*

Theorem 15 *[146] The system (3.1) is (small-time) locally controllable if and only if it is locally accessible.*

Definition 10 *The* Lie Algebra Basis Condition *(LABC) is said to hold if there exists a basis of vector fields in $\mathcal{C}$ that spans the tangent space for every point $q \in Q$.*

Definition 11 *[78] An underactuated driftless control system is* well-controllable *if the Lie Algebra Basis Condition holds.*

In general, the Jacobi-Lie brackets used to satisfy the LARC may vary from point to point. If a system is well-controllable then the controller synthesis for global control will be simpler. We assume that Theorem 15 is satisfied, and now focus on relating the terms in $\mathcal{C}^\infty(q)$ to control system design.

3.1.1 Averaging Theory for Control

In Equation (3.1), the functions $u^a(q,t)$ are control functions, which decompose into state feedback terms and time-periodic terms: $u^a(q,t) = f^a(q) + v^a(t/\epsilon)$. The functions $v^a(t)$ are T-periodic functions and

the functions $f^a(q)$ are feedback terms for stabilization of the directly controlled states. Substitution into equation (3.1) gives

$$\dot{q} = Y_a(q)f^a(q) + Y_a(q)v^a(t/\epsilon) = X_S(q) + Y_a(q)v^a(t/\epsilon), \tag{3.7}$$

where $X_S(q) = Y_a(q)f^a(q)$. A transformation of time, $t \mapsto \epsilon\tau$, takes (3.7) into a form compatible with averaging theory,

$$\frac{\mathrm{d}q}{\mathrm{d}\tau} = \epsilon X_S(q) + \epsilon Y_a(q)v^a(t). \tag{3.8}$$

The average of (3.8) is

$$\begin{aligned} q(\tau) &= \mathrm{Trunc}_{m-1}\left(P(\tau)\right)\left(z(\tau)\right) + O\left(\epsilon^m\right), \\ \frac{\mathrm{d}z}{\mathrm{d}\tau} &= \mathrm{Trunc}_m\left(Z\right). \end{aligned}$$

Transformating back to time, t,

$$\begin{aligned} q(t) &= \mathrm{Trunc}_{m-1}\left(P(t/\epsilon)\right)\left(z(t/\epsilon)\right) + O\left(\epsilon^{m-1}\right), \\ \dot{z} &= \frac{1}{\epsilon}\mathrm{Trunc}_m\left(Z\right). \end{aligned}$$

Recall that $P(t)$ and Z can be expressed as power series in ϵ. The time transformation lowers the power series order by one. For $P(t)$, this might pose a problem at lower orders of averaging since $P(t)$ will go from being $O\left(\epsilon^m\right)$-close to $O\left(\epsilon^{m-1}\right)$-close. Choosing the improved m^{th}-order average, with $\mathrm{Trunc}_m\left(P(t)\right)$, will resolve the problem,

$$\begin{aligned} q(t) &= \mathrm{Trunc}_m\left(P(t/\epsilon)\right)\left(z(t/\epsilon)\right) + O\left(\epsilon^m\right), \\ \dot{z} &= \frac{1}{\epsilon}\mathrm{Trunc}_m\left(Z\right). \end{aligned}$$

Averaged Coefficients

The averaged version of (3.8) will contain vector fields that have combinations of time integrals and Jacobi-Lie brackets, see Section 2.4. Since the periodic inputs act as coefficients to the input vector fields, and iterated Jacobi-Lie brackets are multi-linear, the integrals can be factored. The integral terms represent the net effect of the inputs on the Jacobi-Lie bracket terms, and will be called *averaging coefficients*.

Define the following notation for the *averaging coefficients*:

$$\mathrm{V}^{(a)}_{(n)}(t) \equiv \int_{t_0}^{t}\int^{s_{n-1}}\cdots\int^{s_2} v^a(s_1)ds_1\cdots ds_{n-1}. \tag{3.12}$$

For the purposes of this paper, the initial time will always be $t_0 = 0$. Cases of multiple upper and lower indices denote products of this type of integral. An example is $\mathrm{V}^{(a,b)}_{(1,1)}(t)$,

$$\mathrm{V}^{(a,b)}_{(1,1)}(t) = \mathrm{V}^{(a)}_{(1)}\,\mathrm{V}^{(b)}_{(1)} = \left(\int_0^t v^a(s_1)ds_1\right)\left(\int_0^t v^b(s_1)ds_1\right).$$

Time-averaged terms are called *averaged coefficients*. The single index averaged coefficients are

$$\overline{\mathrm{V}^{(a)}_{(n)}(\tau)} = \frac{1}{T}\int_0^T \mathrm{V}^{(a)}_{(n)}(\tau)\,\mathrm{d}\tau = \frac{1}{T}\mathrm{V}^{(a)}_{(n+1)}(T).$$

Additionally define the following,

$$\widetilde{\mathrm{V}}^{(a)}_{(n)} \equiv \mathrm{V}^{(a)}_{(n)} - \overline{\mathrm{V}^{(a)}_{(n)}},$$

and for the multi-index version,

$$\widetilde{\mathrm{V}}^{(A)}_{(N)} \equiv \mathrm{V}^{(A)}_{(N)} - \overline{\mathrm{V}^{(A)}_{(N)}},$$

where $(A) = (a_1, a_2, ..., a_{|A|})$ and $(N) = (n_1, n_2, ..., n_{|N|})$. The $\widehat{\cdot}$ symbol will denote integrals within the product structure. For example,

$$\mathrm{V}^{(\widehat{a,b},c)}_{(\widehat{0,0},1)}(t) = \left(\int_0^t \mathrm{V}^{(a,b)}_{(0,0)}(\tau)\,\mathrm{d}\tau\right)\left(\mathrm{V}^{(c)}_{(1)}(t)\right) = \left(\int_0^t (v^a(\tau)\,v^b(\tau))\,\mathrm{d}\tau\right)\left(\int_0^t v^c(\tau)\,\mathrm{d}\tau\right).$$

Another example is the averaged coefficient

$$\overline{\mathrm{V}^{(a,b)}_{(1,0)}(\tau)} = \frac{1}{T}\int_0^T \mathrm{V}^{(a,b)}_{(1,0)}(\tau)\,\mathrm{d}\tau = \frac{1}{T}\mathrm{V}^{(\widehat{a,b})}_{(\widehat{1,0})}(T).$$

The notation will simplify the expressions for the averaged expansions of (3.7).

Comment. Those familiar with approximate inversion techniques for open-loop approximate tracking of nonholonomic systems will notice the indefinite nature of the integrals in Equation (3.12). The generalized averaging theory does lead to definite integrals. However, the lower integral limit can be shown to correspond to integration of the initial conditions of dynamical system (3.7). Although the definite integrals do predict the transient response, under state feedback the transient effects of the initial conditions die out over time. Our generalized averaging theory leads to the use of indefinite integrals, as does classical averaging theory. In the motion control algorithms of Bullo, and Martínez and Cortés [21, 100], the integrals are also indefinite. The examples section, Section 3.3, will more explicitly cover this difference between open-loop approximate inversion and closed-loop approximate inversion.

Averaged Expansions

First-order. The first-order averaged version of system (3.7) is

$$\dot{z} = X_S + \overline{\mathrm{V}^{(a)}_{(0)}(t)}Y_a(z), \tag{3.13}$$

where the upper and lower indicies indicate multiplication and summation. The Floquet mapping is approximated by

$$\mathrm{Trunc}_0\left(P(t/\epsilon)\right) = \mathrm{Id}. \tag{3.14}$$

An improved approximation is

$$\mathrm{Trunc}_1\left(P(t/\epsilon)\right) = \mathrm{Id} + \epsilon\int_0^{t/\epsilon} \widetilde{\mathrm{V}}^{(a)}_{(0)}(\tau)\,\mathrm{d}\tau\ Y_a(\cdot). \tag{3.15}$$

First-order averaging may be useful when the inputs are constrained to lie in an admissable set that is restrictive. In [46], Gallivan et al. achieve a set of transition rates for thin film growth that is unachievable using the inputs that the process is normally restricted to. Although the inputs are time-periodic, the inputs for film-growth are not in control affine form, and the averaging process highlights the nonlinear effect of time-periodic inputs. For systems in the affine control form, a non-vanishing average is equivalent to constant actuation, to first-order. It is the higher-order nonlinear effect of periodic inputs that will be important.

Second-order. Second- and higher-order average approximations are usually sought when the first-order average of the time-periodic inputs vanishes.

Assumption 2 *For higher-order averaging, it is assumed that the time-averages of the oscillatory control inputs vanish, i.e.,* $\overline{\mathrm{V}_{(0)}^{(a)}(t)} = 0$.

The second-order average has the form,

$$\dot{z} = X_S(z) + \epsilon\overline{\mathrm{V}_{(1)}^{(a)}(t)}\,[Y_a(z), X_S(z)] + \frac{1}{2}\epsilon\overline{\mathrm{V}_{(1,0)}^{(a,b)}(t)}\,[Y_a(z), Y_b(z)]\,, \tag{3.16a}$$

$$\mathrm{Trunc}_1\left(P(t/\epsilon)\right) = \mathrm{Id} + \epsilon\mathrm{V}_{(1)}^{(a)}(t/\epsilon)Y_a(\cdot), \tag{3.16b}$$

whose derivation requires integration by parts and Assumption 2. The averaged coefficient notation allows for a quick assessment of the nature of the integral product, as the lower indices determine which terms are integrated and how many times to iterate the integral.

Third-order If the LARC is satisfied via higher levels of iterated Lie brackets, then higher-order averaging is required. The averaged vector field to third order [1] is

$$\begin{aligned}\dot{z} = X_S &+ \epsilon\overline{\mathrm{V}_{(1)}^{(a)}(t)}\,[Y_a, X_S] + \frac{1}{2}\epsilon\overline{\mathrm{V}_{(1,0)}^{(a,b)}(t)}\,[Y_a, Y_b] + \epsilon^2\left(\overline{\mathrm{V}_{(2)}^{(a)}(t)} - \frac{1}{2}T\overline{\mathrm{V}_{(1)}^{(a)}(t)}\right)[X_S,[X_S,Y_a]] \\ &- \frac{1}{3}\epsilon^2\left(\overline{\mathrm{V}_{(\widehat{1,0})}^{(\widehat{a,b})}(t)} + \frac{1}{2}T\overline{\mathrm{V}_{(1,0)}^{(a,b)}(t)}\right)[X_S,[Y_a,Y_b]] \\ &+ \frac{1}{3}\epsilon^2\left(\overline{\mathrm{V}_{(1,1)}^{(a,b)}(t)} + \overline{\mathrm{V}_{(\widehat{1,0})}^{(\widehat{a,b})}(t)} - T\overline{\mathrm{V}_{(1,0)}^{(a,b)}(t)}\right)[Y_a,[Y_b,X_S]] + \frac{1}{3}\epsilon^2\overline{\mathrm{V}_{(1,1,0)}^{(a,b,c)}(t)}\,[Y_a,[Y_b,Y_c]]\,.\end{aligned} \tag{3.17}$$

The Floquet mapping is

$$\begin{aligned}\mathrm{Trunc}_2\left(P(t/\epsilon)\right) = \mathrm{Id} &+ \epsilon\mathrm{V}_{(1)}^{(a)}(t/\epsilon)\;Y_a + \epsilon^2\int_0^{t/\epsilon}\widetilde{\mathrm{V}}_{(1)}^{(a)}(\tau)\,\mathrm{d}\tau\;[Y_a, X_S] \\ &+ \frac{1}{2}\epsilon^2\int_0^{t/\epsilon}\widetilde{\mathrm{V}}_{(1,0)}^{(a,b)}(\tau)\,\mathrm{d}\tau\;[Y_a,Y_b] + \frac{1}{2}\epsilon^2\mathrm{V}_{(1,1)}^{(a,b)}(t/\epsilon)\;Y_a\cdot Y_b\,.\end{aligned} \tag{3.18}$$

Higher-order. To compute fourth- and higher-order averages follow the averaged expansion procedure delineated in Chapter 2, taking into account the affine control decomposition. Appendix 3.A contains the lower-order averaged expansions, to give the reader a sense of the procedure.

The Averages Revisited

The existence of the vector fields $f^a(q)$ complicates matters because they act like drift terms. Suppose that first-order, direct state feedback was not necessary, $f^a(q) = 0$. The simplifications afforded in this case allow for higher-order averages to be computed.

Third-order. The third-order equations simplify to

$$\dot{z} = \frac{1}{2}\epsilon\overline{\mathrm{V}_{(1,0)}^{(a,b)}(t)}\,[Y_a, Y_b] + \frac{1}{3}\epsilon^2\overline{\mathrm{V}_{(1,1,0)}^{(a,b,c)}(t)}\,[Y_a,[Y_b,Y_c]]\,, \tag{3.19}$$

[1] See Appendix 3.A for the relevant calculations

and the reductions actually make a fourth-order expansion tractable.

Fourth-order. The averaged autonomous vector field to fourth-order is

$$\begin{aligned}\dot{z} &= \epsilon\overline{\mathrm{V}^{(a,b)}_{(1,0)}(t)}\,[Y_a,Y_b] + \frac{1}{3}\epsilon^2\overline{\mathrm{V}^{(a,b,c)}_{(1,1,0)}(t)}\,[Y_a,[Y_b,Y_c]] \\ &= \frac{1}{12}\epsilon^3\left(\overline{\mathrm{V}^{(a,b,\widehat{c,d})}_{(1,0,\widehat{1,0})}(t)}\,[[Y_a,Y_b],[Y_c,Y_d]] - \left(\overline{\mathrm{V}^{(a,b,\widehat{c,d})}_{(1,0,\widehat{0,1})}(t)} + \overline{\mathrm{V}^{(a,b,c,d)}_{(1,1,1,0)}(t)}\right)[Y_a,[Y_b,[Y_c,Y_d]]]\right).\end{aligned} \tag{3.20}$$

The Floquet mapping is

$$\begin{aligned}\mathrm{Trunc}_3\left(P(t/\epsilon)\right) = \mathrm{Id} &+ \epsilon\mathrm{V}^{(a)}_{(1)}(t/\epsilon)\,Y_a + \frac{1}{2}\epsilon^2\int_0^{t/\epsilon}\widetilde{\mathrm{V}}^{(a,b)}_{(1,0)}(\tau)\,\mathrm{d}\tau\;[Y_a,Y_b] + \frac{1}{2}\epsilon^2\mathrm{V}^{(a,b)}_{(1,1)}(t/\epsilon)\,Y_a\cdot Y_b \\ &- \frac{1}{2}\epsilon^2\mathrm{V}^{(a)}_{(1)}(t/\epsilon)\overline{\mathrm{V}^{(b,c)}_{(1,0)}(t)}\,t\,Y_a\cdot[Y_b,Y_c] + \frac{1}{6}\epsilon^3\mathrm{V}^{(a,b,c)}_{(1,1,1)}(t)Y_a\cdot Y_b\cdot Y_c \\ &+ \frac{1}{4}\epsilon^3\mathrm{V}^{(a)}_{(1)}(t/\epsilon)\int_0^{t/\epsilon}\mathrm{V}^{(b,c)}_{(1,0)}(\sigma)\,\mathrm{d}\sigma\;\left(Y_a\cdot[Y_b,Y_c]+[Y_b,Y_c]\cdot Y_a\right).\end{aligned} \tag{3.21}$$

For driftless systems, it is the iterated Jacobi-Lie brackets generated using the set $\{\,Y_a\,\}_1^m$ that is important for both determination of controllability, and for realization of control. When the vector field $X_S \neq 0$, the averaged expansions become more complicated, but the Jacobi-Lie brackets that are needed for control do not change. Therefore, it always useful to examine the case where $X_S = 0$ to determine the appropriate averaged coefficients and Jacobi-Lie brackets to analyze.

3.1.2 Sinusoidal Inputs for Indirect Actuation

The work of Murray and Sastry [122], and the subsequent developments by Teel et al. [157] have both demonstrated the use of sinusoidal inputs for motion generation in Lie bracket directions. By approximating the flow using series expansions with averaged coefficients, it is possible to compute the amplitudes of the response to sinusoidal forcing in a given Jacobi-Lie bracket direction. This *approximate inversion* technique is successfully utilized by Bullo [22] and Martínez and Cortés [100] in deriving motion control algorithms for control of underactuated mechanical systems. Recent work on the second-order case has shown how to construct sinusoidal inputs with proper amplitude modulation and frequency spacing so as to isolate the various Lie bracket contributions [159, 160]. More generally, Sussmann and Liu [154, 155, 92] give a strategy for finding the sinusoidal functions and frequencies to obtain individually excited averaged coefficients. This section examines the algorithm for select averaged coefficients with the goal of providing a practical approach to-, and interpretation of-, the Ω-sets found in the work of Sussman and Liu.

Second-order Averaged Coefficients

The process of determining the oscillatory inputs for an arbitrary system to any given order of averaging is difficult. Sinusoidal input pairs in resonance will lead to Jacobi-Lie bracket motion. Each averaged coefficient has algebraic frequency requirements for resonance, which may conflict and cause coupling. Certain Jacobi-Lie brackets can never be actuated independently. Nevertheless, useful guidelines can be established by investigating simple, lower-order cases. Abstraction to higher-order follows naturally.

$\overline{\mathrm{V}^{(a,b)}_{(1,0)}(t)}$	$\alpha^a \sin(\omega_a t)$	$\alpha^a \cos(\omega_a t)$
$\alpha^b \sin(\omega_b t)$	0	$\begin{cases} \frac{\alpha^a \alpha^b}{2\omega_a} & \text{if } \omega_a = \omega_b \\ 0 & \text{otherwise} \end{cases}$
$\alpha^b \cos(\omega_b t)$	$\begin{cases} -\frac{\alpha^a \alpha^b}{2\omega_a} & \text{if } \omega_a = \omega_b \\ 0 & \text{otherwise} \end{cases}$	0

Table 3.1: Averaged coefficients for second-order averaging of driftless systems.

For second-order averaging, the averaged coefficient is

$$\overline{\mathrm{V}^{(a,b)}_{(0,1)}(t)} = \frac{1}{T}\int_0^T \int^t v^a(\tau)\,\mathrm{d}\tau \,.\, v^b(t)\,\mathrm{d}t\,.$$

For each input $v^a(t)$, $a = 1 \ldots m$, one of the following two forcing functions can be chosen,

$$\alpha^a \sin(\omega_a t) \quad \text{or} \quad \alpha^a \cos(\omega_a t),$$

where $\omega_a \in \mathbb{Z}^+$. The four permutations of input function choices result in the averaged coefficients found in Table 3.1. Of the four options, in-phase sinusoidal inputs do not work. Additionally, the algebraic equality

$$\omega_a - \omega_b = 0 \tag{3.22}$$

must hold for the desired out-of-phase sinusoidal input pair and for no other oscillatory inputs. The net result of this analysis is a set of inputs that operate at unique carrier frequencies; a commmonly known fact.

Lemma 2 *Consider the second-order Jacobi-Lie bracket, $[Y_a, Y_b]$. If the associated inputs*

$$v^a(t) = \alpha^a_{ab} \cos(\omega_{ab} t), \qquad v^b(t) = \alpha^b_{ab} \sin(\omega_{ab} t),$$

are chosen for some unique carrier frequency, ω_{ab}, then only the bracket $[Y_a, Y_b]$ will be excited. The corresponding averaged coefficient will evaluate to

$$\overline{\mathrm{V}^{(a,b)}_{(1,0)}(t)} = \frac{\alpha_{ab}}{2\omega_{ab}}, \quad \textit{where } \alpha_{ab} = \alpha^a_{ab}\alpha^b_{ab}.$$

Lemma 2 imposes a unique carrier frequency for each Jacobi-Lie bracket. This will pose a problem for severely underactuated systems, as the number of frequencies needed is directly related to the number of Jacobi-Lie brackets required. For a pair of oscillatory inputs generating motion in the direction of the Jacobi-Lie bracket $[Y_a, Y_b]$, the averaged coefficient is found in Lemma 2 to be $\frac{\alpha_{ab}}{2\omega_{ab}}$. As the frequency increases, the response decreases. If either α^a_{ab} or α^b_{ab} is set equal to ω_{ab}, the response is no longer attenuated by the frequency. Unfortunately, the oscillation amplitude for one of the inputs is on the order of ω_{ab}, which may be large. It is possible to recycle carrier frequencies to lower the total amount of unique frequencies used, while avoiding coupling.

Algorithm 1 (Second-order Selection Algorithm) *The inductive selection algorithm below gives independent Jacobi-Lie bracket contributions.*

1. *All of the brackets that incorporate the input $v^1(t)$ are assigned the frequency $\omega_1 = 1$. The input function for $v^1(t)$ will be $\cos(\omega_1 t)$. Let the number of distinct brackets $[Y_1, Y_a]$ required for STLC be termed the multiplicty, μ_1. For each of the distinct brackets add the term $\alpha_{1a}^a \sin(\omega_1 t)$ to the input function $v^a(t)$.*

2. *Let b = 2.*

3. *Consider next, all of the essential brackets that incorporate the input $v^b(t)$. The corresponding input functions will be assigned the frequency $\omega_b = \omega_{b-1} + 1$. The oscillatory function $\omega_b \cos(\omega_b t)$ will be added to the input function $v^b(t)$. The number of distinct brackets $[Y_b, Y_a]$, $a > b$, required for STLC will be termed the multiplicity μ_b. For each distinct bracket add the term $\alpha_{ba}^a \sin(\omega_b t)$ to the input function $v^a(t)$.*

   If there are no brackets using $[Y_b, Y_a]$, $a > b$, then do no assignments. Set $\omega_b = \omega_{b-1}$ and $\alpha_{ba}^a = 0$, $a > b$.

4. *Repeat Step 3 with $b = b + 1$, until all of the Jacobi-Lie brackets are assigned corresponding cosine and sine functions.*

Mathematically, the input $v^a(t)$ is of the form

$$v^a(t) = \omega_a \delta(\mu_a) \cos(\omega_a t) + \sum_{b=1}^{b \leq a} \alpha_{ba}^a \sin(\omega_a t), \tag{3.23}$$

where

$$\delta(\mu) = \begin{cases} 0 & \text{if } \mu = 0 \\ 1 & \text{otherwise} \end{cases}.$$

According to the algorithm above, the α_{ba}^a terms will be assigned a nonzero value if the corresponding Jacobi-Lie bracket, $[Y_b, Y_a]$ is needed for controllability. The carrier frequency and the choice of sinusoidal functions for $v^a(t)$ and $v^b(t)$ have been chosen to ensure independence of the α values in the averaged coefficient, and therefore independent contributions of the bracket-related terms.

proof
Compare two input functions,

$$v^a(t) = \omega_a \cos(\omega_a t) + \sum_{b=1}^{b<a} \alpha_{ba}^a \sin(\omega_b t), \text{ and}$$
$$v^c(t) = \omega_c \cos(\omega_c t) + \sum_{d=1}^{d<c} \alpha_{dc}^c \sin(\omega_d t).$$

Without loss of generality, assume that $a \leq c$. The averaged coefficient $\overline{\mathrm{V}^{(a,c)}_{(1,0)}(t)}$, can be decomposed into four parts:

$$\overline{\mathrm{V}^{(a,c)}_{(1,0)}(t)} = \underbrace{\overline{\omega_a\omega_c \left(\int_0^t \cos(\omega_a\tau)\,\mathrm{d}\tau\right)\cos(\omega_c t)}}_{1} + \underbrace{\overline{\omega^a \left(\int_0^t \cos(\omega_a\tau)\,\mathrm{d}\tau\right)\sum_{d=1}^{d<c}\alpha^c_{dc}\sin(\omega_d t)}}_{2},$$
$$+ \underbrace{\overline{\int_0^t \sum_{b=1}^{b<a}\alpha^a_{ba}\sin(\omega_b\tau)\,\mathrm{d}\tau\,\omega_c\cos(\omega_c t)}}_{3} + \underbrace{\overline{\int_0^t \sum_{b=1}^{b<a}\alpha^a_{ba}\sin(\omega_b\tau)\,\mathrm{d}\tau\sum_{d=1}^{d<c}\alpha^c_{dc}\sin(\omega_d t)}}_{4}.$$

Using Table 3.2, and the fact that $a < c$, all but the second term vanish. The second term averages to

$$\overline{\mathrm{V}^{(a,c)}_{(1,0)}(t)} = \overline{\omega^a \left(\int_0^t \cos(\omega_a\tau)\,\mathrm{d}\tau\right)\sum_{d=1}^{d<c}\alpha^c_{dc}\sin(\omega_d t)} = \begin{cases} \frac{1}{2}\alpha^c_{ac} & a < c \\ 0 & a = c \end{cases}.$$

Skew-symmetry of the averaged coefficient, $\overline{\mathrm{V}^{(a,c)}_{(1,0)}(t)}$, takes care of the other case, $a > c$. To show skew-symmetry of the averaged coefficient, integrate by parts and use Assumption 2.

■

There are two properties of the selection algorithm worth noting: (1) when all of the α are set to zero, the inputs do not vanish due to the cosine terms, and (2) the selection algorithm requires modification if third- or higher-order averaging is used. The principal modification required when moving to higher-order is the spacing of the carrier frequencies; the current choice $\omega_a = \omega_{a-1} + 1$ will introduce higher-order coupling.

Third-Order Averaged Coefficients

The important factor for the selection of inputs in the previous analysis involved the algebraic equality (3.22). Higher-order expansions will lead to additional algebraic restrictions that keep the effect of the inputs isolated. These restrictions will also affect the previous construction, if both second- and third-order effects are to be simultaneously included. To motivate the use of these additional constraints, consider the following example that involves third-order averaging.

Example (Three inputs, 3^{rd}-order averaging)
Consider the case of a three-input underactuated driftless control system that requires third-order iterated Jacobi-Lie brackets for the LARC to be satisfied. To third-order, the important averaged coefficients are

$$\overline{\mathrm{V}^{(a,b)}_{(1,0)}(t)} \quad \text{and} \quad \overline{\mathrm{V}^{(a,b,c)}_{(1,1,0)}(t)}.$$

Because $\overline{\mathrm{V}^{(a,b)}_{(1,0)}(t)}$ was analyzed in the previous section, the focus will be on the latter averaged coefficient. The possible input choices are

$v^1(t)$	$v^2(t)$	$v^3(t)$
$\alpha^1\sin(\omega_1 t)$	$\alpha^2\sin(\omega_2 t)$	$\alpha^3\sin(\omega_3 t)$
$\alpha^1\cos(\omega_1 t)$	$\alpha^2\cos(\omega_2 t)$	$\alpha^3\cos(\omega_3 t)$

,

whose eight permutations lead to the contributions found in Table 3.2, with potential coupling found in Table 3.3. Based on Table 3.2, the important algebraic equalities that influence the averaged coefficients are as

$v^1(t), v2(t), v^3(t)$		$\overline{\mathrm{V}^{(1,2,3)}_{(1,1,0)}(t)}$
$\alpha^1 \sin(\omega_1 t),\ \alpha^2 \sin(\omega_2 t),\ \alpha^3 \sin(\omega_3 t)$	$=$	0
$\alpha^1 \sin(\omega_1 t),\ \alpha^2 \sin(\omega_2 t),\ \alpha^3 \cos(\omega_3 t)$	$=$	$\begin{cases} & \omega_1 + \omega_2 - \omega_3 = 0, \\ \frac{\alpha^1\alpha^2\alpha^3}{4\omega_1\omega_2} & \text{if } \omega_1 - \omega_2 - \omega_3 = 0, \\ & \omega_1 - \omega_2 + \omega_3 = 0 \\ 0 & \text{otherwise} \end{cases}$
$\alpha^1 \sin(\omega_1 t),\ \alpha^2 \cos(\omega_2 t),\ \alpha^3 \sin(\omega_3 t)$	$=$	$\begin{cases} \frac{\alpha^1\alpha^2\alpha^3}{4\omega_1\omega_2} & \text{if } \omega_1 + \omega_2 - \omega_3 = 0, \\ & \omega_1 - \omega_2 + \omega_3 = 0 \\ -\frac{\alpha^1\alpha^2\alpha^3}{4\omega_1\omega_2} & \text{if } \omega_1 - \omega_2 - \omega_3 = 0 \\ 0 & \text{otherwise} \end{cases}$
$\alpha^1 \sin(\omega_1 t),\ \alpha^2 \cos(\omega_2 t),\ \alpha^3 \cos(\omega_3 t)$	$=$	0
$\alpha^1 \cos(\omega_1 t),\ \alpha^2 \sin(\omega_2 t),\ \alpha^3 \sin(\omega_3 t)$	$=$	$\begin{cases} -\frac{\alpha^1\alpha^2\alpha^3}{4\omega_1\omega_2} & \text{if } \omega_1 - \omega_2 - \omega_3 = 0, \\ & \omega_1 - \omega_2 + \omega_3 = 0 \\ \frac{\alpha^1\alpha^2\alpha^3}{4\omega_1\omega_2} & \text{if } \omega_1 + \omega_2 - \omega_3 = 0 \\ 0 & \text{otherwise} \end{cases}$
$\alpha^1 \cos(\omega_1 t),\ \alpha^2 \sin(\omega_2 t),\ \alpha^3 \cos(\omega_3 t)$	$=$	0
$\alpha^1 \cos(\omega_1 t),\ \alpha^2 \cos(\omega_2 t),\ \alpha^3 \sin(\omega_3 t)$	$=$	0
$\alpha^1 \cos(\omega_1 t),\ \alpha^2 \cos(\omega_2 t),\ \alpha^3 \cos(\omega_3 t)$	$=$	$\begin{cases} \frac{\alpha^1\alpha^2\alpha^3}{4\omega_1\omega_2} & \text{if } \omega_1 - \omega_2 - \omega_3 = 0, \\ & \omega_1 - \omega_2 + \omega_3 = 0 \\ -\frac{\alpha^1\alpha^2\alpha^3}{4\omega_1\omega_2} & \text{if } \omega_1 + \omega_2 - \omega_3 = 0 \\ 0 & \text{otherwise} \end{cases}$

Table 3.2: Averaged coefficients for third-order averaging of driftless systems.

$v^a(t), v^b(t)$		$\overline{\mathrm{V}^{(i,j,k)}_{(1,1,0)}(t)},\ i,j,k \in \{a,b\}$	
$\alpha^a \sin(\omega_a t),\ \alpha^b \sin(\omega_b t)$	$=$	0	
$\alpha^a \sin(\omega_a t),\ \alpha^b \cos(\omega_b t),$	$=$	$\overline{\mathrm{V}^{(a,a,b)}_{(1,1,0)}(t)} = \frac{(\alpha^a)^2\alpha^b}{4\omega_a^2}$	if $2\omega_a = \omega_b$
		$\overline{\mathrm{V}^{(a,b,a)}_{(1,1,0)}(t)} = -\frac{(\alpha^a)^2\alpha^b}{4\omega_a\omega_b}$	if $2\omega_a = \omega_b$
		$\overline{\mathrm{V}^{(b,a,a)}_{(1,1,0)}(t)} = -\frac{(\alpha^a)^2\alpha^b}{4\omega_a\omega_b}$	if $2\omega_a = \omega_b$
		0	otherwise
$\alpha^a \cos(\omega_a t),\ \alpha^b \sin(\omega_b t),$	$=$	$\overline{\mathrm{V}^{(b,b,a)}_{(1,1,0)}(t)} = \frac{(\alpha^b)^2\alpha^a}{4\omega_b^2}$	if $2\omega_b = \omega_a$
		$\overline{\mathrm{V}^{(b,a,b)}_{(1,1,0)}(t)} = -\frac{(\alpha^b)^2\alpha^a}{4\omega_a\omega_b}$	if $2\omega_b = \omega_a$
		$\overline{\mathrm{V}^{(a,b,b)}_{(1,1,0)}(t)} = -\frac{(\alpha^b)^2\alpha^a}{4\omega_a\omega_b}$	if $2\omega_b = \omega_a$
		0	otherwise
$\alpha^a \cos(\omega_a t),\ \alpha^b \cos(\omega_b t),$	$=$	$\overline{\mathrm{V}^{(a,a,b)}_{(1,1,0)}(t)} = -\frac{(\alpha^a)^2\alpha^b}{4\omega_a^2}$	if $2\omega_a = \omega_b$
		$\overline{\mathrm{V}^{(a,b,a)}_{(1,1,0)}(t)} = \frac{(\alpha^a)^2\alpha^b}{4\omega_a\omega_b}$	if $2\omega_a = \omega_b$
		$\overline{\mathrm{V}^{(b,a,a)}_{(1,1,0)}(t)} = \frac{(\alpha^a)^2\alpha^b}{4\omega_a\omega_b}$	if $2\omega_a = \omega_b$
		$\overline{\mathrm{V}^{(b,b,a)}_{(1,1,0)}(t)} = -\frac{(\alpha^b)^2\alpha^a}{4\omega_b^2}$	if $2\omega_b = \omega_a$
		$\overline{\mathrm{V}^{(b,a,b)}_{(1,1,0)}(t)} = \frac{(\alpha^b)^2\alpha^a}{4\omega_a\omega_b}$	if $2\omega_b = \omega_a$
		$\overline{\mathrm{V}^{(a,b,b)}_{(1,1,0)}(t)} = \frac{(\alpha^b)^2\alpha^a}{4\omega_a\omega_b}$	if $2\omega_b = \omega_a$
		0	otherwise

Table 3.3: Coupling of averaged coefficients for third-order averaging of driftless systems.

follows,

$$\omega_1 + \omega_2 - \omega_3 = 0,\ \omega_1 - \omega_2 - \omega_3 = 0,\ \text{and}\ \omega_1 - \omega_2 + \omega_3 = 0. \tag{3.24}$$

In order to avoid second-order coupling between terms when three distinct vector fields are contributing, the following inequality also needs to hold,

$$\omega_i - \omega_j \neq 0, \quad \text{for } i \neq j. \tag{3.25}$$

Failure to satisfy the inequality (3.25) will excite a Jacobi-Lie bracket of the form, $[Y_i, Y_j]$. With the above conditions met, the only non-zero combinations involve an odd number of cosines and an even number of sines. It is important to note that the algebraic inequality,

$$2\omega_i - \omega_j \neq 0, \quad \text{for } i \neq j, \tag{3.26}$$

may also need to hold, according to the coupling of Table 3.3. Failure to satisfy the inequality (3.26) may excite the Jacobi-Lie bracket corresponding to the averaged coefficients found in Table 3.3. These Jacobi-Lie brackets take the form, $[Y_i, [Y_j, Y_i]]$.

Let's examine the possiblities for the 3^{rd} order case. If the Jacobi-Lie bracket is a function of three distinct vector fields, i.e., $[Y_1, [Y_2, Y_3]]$, then satisfying the algebraic relations in Equations (3.24)-(3.26) for the associated input functions, according to the Tables 3.2 and 3.3, will result in at least three contributions from three Jacobi-Lie bracket combinations (obtained by cycling the vector fields) due to the fact that they all satisfy the algebraic equalities. The averaged coefficients are

$$\overline{\mathrm{V}^{(i,j,k)}_{(1,1,0)}(t)} = \begin{bmatrix} \begin{bmatrix}0\\0\\0\end{bmatrix} & \begin{bmatrix}0\\0\\\frac{\alpha^1\alpha^2\alpha^3}{4\omega_1\omega_2}\end{bmatrix} & \begin{bmatrix}0\\-\frac{\alpha^1\alpha^2\alpha^3}{4\omega_1\omega_3}\\0\end{bmatrix} \\ \begin{bmatrix}0\\0\\\frac{\alpha^1\alpha^2\alpha^3}{4\omega_1\omega_2}\end{bmatrix} & \begin{bmatrix}0\\0\\0\end{bmatrix} & \begin{bmatrix}\frac{\alpha^1\alpha^2\alpha^3}{4\omega_2\omega_3}\\0\\0\end{bmatrix} \\ \begin{bmatrix}0\\-\frac{\alpha^1\alpha^2\alpha^3}{4\omega_1\omega_3}\\0\end{bmatrix} & \begin{bmatrix}\frac{\alpha^1\alpha^2\alpha^3}{4\omega_2\omega_3}\\0\\0\end{bmatrix} & \begin{bmatrix}0\\0\\0\end{bmatrix} \end{bmatrix}, \tag{3.27}$$

where $\omega_2 > \omega_3 > \omega_1$ and $i, j, k \in \{1, 2, 3\}$. The matrix representing the averaged coefficient is read as follows: i gives the outer row, j gives the column, and k gives the inner row. When multiplied by the corresponding Jacobi-Lie brackets then summed over all indices,

$$\begin{aligned}
\overline{\mathrm{V}^{(i,j,k)}_{(1,1,0)}(t)}\,[Y_i,[Y_j,Y_k]] &= \frac{\alpha^1\alpha^2\alpha^3}{4\omega_1\omega_2}[Y_1,[Y_2,Y_3]] - \frac{\alpha^1\alpha^2\alpha^3}{4\omega_1\omega_3}[Y_1,[Y_3,Y_2]] + \frac{\alpha^1\alpha^2\alpha^3}{4\omega_2\omega_3}[Y_2,[Y_3,Y_1]] \\
&\quad + \frac{\alpha^1\alpha^2\alpha^3}{4\omega_1\omega_2}[Y_2,[Y_1,Y_3]] - \frac{\alpha^1\alpha^2\alpha^3}{4\omega_1\omega_3}[Y_3,[Y_1,Y_2]] + \frac{\alpha^1\alpha^2\alpha^3}{4\omega_2\omega_3}[Y_3,[Y_2,Y_1]]. \\
&= \frac{\alpha^1\alpha^2\alpha^3}{4}\left(\left(\frac{1}{\omega_1\omega_2}+\frac{1}{\omega_2\omega_3}\right)[Y_1,[Y_2,Y_3]] + \left(\frac{1}{\omega_2\omega_3}-\frac{1}{\omega_1\omega_2}\right)[Y_2,[Y_3,Y_1]]\right. \\
&\qquad \left. - \left(\frac{1}{\omega_1\omega_2}+\frac{1}{\omega_1\omega_3}\right)[Y_3,[Y_1,Y_2]]\right).
\end{aligned} \tag{3.28}$$

Using the Jacobi identity, it is possible to reduce this number.
□

This example highlights a few critical issues when moving to higher-order. The first being the possibility of coupling due to integrally related choices of frequencies failing to satisfy the algebraic inequality (3.25). Additional coupling may occur between the Jacobi-Lie bracket of two of the input vector fields, i.e., $[Y_a, [Y_b, Y_a]]$, if the algebraic inequality (3.26) fails to be satisfied according to Table 3.3. Secondly there is the problem of exciting more than one Jacobi-Lie bracket simultaneously. It will be shown that, when the input functions are restricted to sines and cosine, it may not be possible to simplify the multiple Jacobi-Lie bracket contributions down to a single contribution.

Lemma 3 *For the case of three distinct vector fields entering into the third-order iterated Jacobi-Lie bracket, $[Y_a, [Y_b, Y_c]]$, ($a \neq b$, $a \neq c$, and $b \neq c$), no choice of inputs will result in the excitation of motion along a single bracket direction. If the following inputs are chosen,*

$$v^a(t) = \alpha^a_{abc} \cos(\omega_{abc} t),\ v^b(t) = \alpha^b_{abc} \sin(3\omega_{abc} t),\ v^c(t) = \alpha^c_{abc} \sin(2\omega_{abc} t)$$

for some principle carrier frequency, ω_{abc}, then only the bracket $[Y_a, [Y_b, Y_c]]$ and a cyclicly related bracket, $[Y_c, [Y_a, Y_b]]$ or $[Y_b, [Y_c, Y_a]]$, will be excited. The corresponding averaged coefficients will be,

$$\overline{\mathrm{V}^{(a,b,c)}_{(1,1,0)}(t)} = \frac{3\alpha_{abc}}{8\omega^2_{abc}} \quad \textit{and} \quad \overline{\mathrm{V}^{(c,a,b)}_{(1,1,0)}(t)} = \frac{\alpha_{abc}}{8\omega^2_{abc}}$$

or,

$$\overline{\mathrm{V}^{(a,b,c)}_{(1,1,0)}(t)} = \frac{\alpha_{abc}}{4\omega^2_{abc}} \quad \textit{and} \quad \overline{\mathrm{V}^{(b,c,a)}_{(1,1,0)}(t)} = -\frac{\alpha_{abc}}{8\omega^2_{abc}},$$

where $\alpha_{abc} = \alpha^a_{abc}\alpha^b_{abc}\alpha^c_{abc}$.

proof

Assume for now that, v^a, v^b, and v^c, are the only nonzero inputs to the system. We will examine the corresponding averaged coefficient. Since the other inputs are set to zero, the system behaves similarly to the three input example above. Without loss of generality, assume that $a = 1$, $b = 2$, and $c = 3$.

Let the input functions be

$$v^1(t) = \alpha^1 \cos(\omega_1 t),\ v^2(t) = \alpha^2 \sin(\omega_2 t),\ v^3(t) = \alpha^3 \sin(\omega_3 t).$$

Assume that the algebraic inequality in Equation (3.26) for the coupling found in Table 3.3 is satisfied, and no coupling occurs. The critical elements of the averaged coefficient are the same as in Equation (3.27), where $\omega_2 > \omega_3 > \omega_1$ and $i, j, k \in \{\, 1, 2, 3 \,\}$. From Equation (3.28), the excited Jacobi-Lie brackets are

$$\begin{aligned}
&\overline{\mathrm{V}^{(i,j,k)}_{(1,1,0)}(t)}\,[Y_i, [Y_j, Y_k]] \\
&\quad = \frac{\alpha^1\alpha^2\alpha^3}{4}\left(\left(\frac{1}{\omega_1\omega_2} + \frac{1}{\omega_1\omega_3}\right)[Y_1, [Y_2, Y_3]] + \left(\frac{1}{\omega_2\omega_3} - \frac{1}{\omega_1\omega_2}\right)[Y_2, [Y_3, Y_1]]\right. \\
&\qquad \left. - \left(\frac{1}{\omega_1\omega_3} + \frac{1}{\omega_2\omega_3}\right)[Y_3, [Y_1, Y_2]]\right).
\end{aligned} \tag{3.29}$$

Denote $\hat{\omega}_i$ by

$$\hat{\omega}_1 = \frac{1}{\omega_1}\left(\frac{1}{\omega_2} + \frac{1}{\omega_3}\right),\ \hat{\omega}_2 = \frac{1}{\omega_2}\left(\frac{1}{\omega_3} - \frac{1}{\omega_1}\right),\ \text{and } \hat{\omega}_3 = \frac{1}{\omega_3}\left(\frac{1}{\omega_1} + \frac{1}{\omega_2}\right).$$

With these definitions,

$$\hat{\omega}_1 > 0,\ \hat{\omega}_2 < 0,\ \hat{\omega}_3 > 0.$$

The Jacobi identity may be used to remove one of the Jacobi-Lie brackets, $[Y_2, [Y_3, Y_1]]$ or $[Y_3, [Y_1, Y_2]]$. Select the last Jacobi-Lie bracket from Equation (3.29) to be removed using the Jacobi identity,

$$\overline{\mathrm{V}_{(1,1,0)}^{(i,j,k)}(t)}\,[Y_i, [Y_j, Y_k]] = \frac{\alpha^1\alpha^2\alpha^3}{4}\left((\hat{\omega}_1 + \hat{\omega}_3)\,[Y_1, [Y_2, Y_3]] + (\hat{\omega}_2 + \hat{\omega}_3)\,[Y_2, [Y_3, Y_1]]\right). \tag{3.30}$$

The second Jacobi-Lie bracket in Equation (3.30) can be cancelled only if there exists a selection of ω_i satisfying one of the tensor equalities (3.24), such that

$$\hat{\omega}_2 + \hat{\omega}_3 = 0. \tag{3.31}$$

This requires finding ω_1 and ω_3, achieving the equality,

$$\frac{2}{(\omega_1 + \omega_3)\omega_3} - \frac{1}{(\omega_1 + \omega_3)\omega_1} + \frac{1}{\omega_1\omega_3} = 0.$$

Placing all of the terms under the same denominator gives the equivalent problem,

$$2\omega_1 - \omega_3 + (\omega_1 + \omega_3) = 3\omega_1 = 0,$$

which is not a valid solution. Therefore, there does not exist a selection of $\omega_2 > \omega_3 > \omega_1$, such that only one Jacobi-Lie bracket contribution is achieved.

As for the choice of inputs given above, notice that selecting any $\omega_2 > \omega_3 > \omega_1$ such that the equality $\omega_2 = \omega_1 + \omega_3$ holds will excite only two Jacobi-Lie brackets. Choosing $\omega_3 = 2\omega_1 = 2\omega$ will avoid the coupling in Table 3.3 according to the algebraic inequality in Equation (3.26). The important decision of the frequency assignments is to have the cosine carrier frequency be the smallest to avoid the coupling. The final contribution is

$$\overline{\mathrm{V}_{(1,1,0)}^{(i,j,k)}(t)}\,[Y_i, [Y_j, Y_k]] = \frac{\alpha^1\alpha^2\alpha^3}{8\omega^2}\left(3\,[Y_1, [Y_2, Y_3]] + [Y_2, [Y_3, Y_1]]\right).$$

Suppose that, instead, the second Jacobi-Lie bracket from Equation (3.29) were to be removed using the Jacobi identity. Then the important averaged coefficients are

$$\overline{\mathrm{V}_{(1,1,0)}^{(i,j,k)}(t)}\,[Y_i, [Y_j, Y_k]] = \frac{\alpha^1\alpha^2\alpha^3}{4}\left((\hat{\omega}_1 - \hat{\omega}_2)\,[Y_1, [Y_2, Y_3]] - (\hat{\omega}_2 + \hat{\omega}_3)\,[Y_3, [Y_1, Y_2]]\right). \tag{3.32}$$

The second Jacobi-Lie bracket in Equation (3.32) can be cancelled only if there exists a selection of ω_i satisfying one of the algebraic equalities (3.24), such that Equation (3.31) holds, which is not possible.

Again, choosing $\omega_3 = 2\omega_1 = 2\omega$ will avoid the coupling in Table 3.3 according to the algebraic inequality in Equation (3.26). The final contribution is

$$\overline{\mathrm{V}_{(1,1,0)}^{(i,j,k)}(t)}\,[Y_i, [Y_j, Y_k]] = \frac{\alpha^1\alpha^2\alpha^3}{8\omega^2}\left(2\,[Y_1, [Y_2, Y_3]] - [Y_2, [Y_3, Y_1]]\right).$$

Suppose now that a, b, and c are arbitrary, and the carrier frequency ω_{abc} is chosen to be unique in the sense that the algebraic inequalities of Equation (3.25), and Equation (3.26) according to Table 3.3, hold. Furthermore, the algebraic equalities in Equation (3.24) hold with no undesired coupling. The analysis

implies either,

$$\overline{\mathrm{V}^{(d,e,f)}_{(1,1,0)}(t)}\,[Y_d,[Y_e,Y_f]] = \frac{\alpha_{abc}}{8\omega_{abc}^2}\left(3\,[Y_a,[Y_b,Y_c]] + [Y_b,[Y_c,Y_a]]\right),$$

or,

$$\overline{\mathrm{V}^{(d,e,f)}_{(1,1,0)}(t)}\,[Y_d,[Y_e,Y_f]] = \frac{\alpha_{abc}}{8\omega_{abc}^2}\left(2\,[Y_a,[Y_b,Y_c]] - [Y_b,[Y_c,Y_a]]\right),$$

for summation over all indices $d, e, f \in \{\,1 \ldots m\,\}$, and where $\alpha_{abc} = \alpha^a_{abc}\alpha^b_{abc}\alpha^c_{abc}$.

In the worst case, the frequency $\omega_{abc} = 2\omega_{max} + 1$ may be chosen, where ω_{max} is the maximal carrier frequency of the other sinusoidal inputs. When utilizing higher-order averaged coefficients, the frequency spacing will change to avoid coupling.

■

There is nothing that can be done to resolve the problem of exciting a second Jacobi-Lie bracket in this case. However, one might be fortunate and have a vanishing bracket, i.e., if $[Y_b,[Y_c,Y_a]] = 0$, then it is possible to excite the remaining Lie bracket independently. This is precisely what happens in the two-input, three-bracket case, one of the cycled brackets vanishes.

Lemma 4 *For the case of two distinct vector fields entering into the third-order iterated Jacobi-Lie bracket,* $[Y_b,[Y_a,Y_b]]$*, if the following inputs are chosen,*

$$v^a(t) = \alpha^a_{bab}\cos(2\omega_{bab}t),\ v^b(t) = \alpha^b_{bab}\sin(\omega_{bab}t),$$

for some unique principle carrier frequency, ω_{bab}*, then only the bracket* $[Y_b,[Y_a,Y_b]]$ *will be excited. The corresponding averaged coefficient is*

$$\overline{\mathrm{V}^{(b,a,b)}_{(1,1,0)}(t)} = -\frac{3\alpha_{bab}}{8\omega_{bab}^2},$$

where $\alpha_{bab} = \alpha^a_{bab}\left(\alpha^b_{bab}\right)^2$.

proof

Assume that these are the only nonzero inputs to the system. Without loss of generality, assume that $a = 1$, $b = 2$. Set the input functions to be

$$v^1(t) = \alpha^1\cos(2\omega t),\ v^2(t) = \alpha^2\sin(\omega t). \tag{3.33}$$

The averaged coefficients are

$$\overline{\mathrm{V}^{(i,j,k)}_{(1,1,0)}(t)} = \begin{bmatrix} \begin{bmatrix}0\\0\end{bmatrix} & \begin{bmatrix}0\\ -\frac{\alpha^1(\alpha^2)^2}{8\omega^2}\end{bmatrix} \\ \begin{bmatrix}0\\ -\frac{\alpha^1(\alpha^2)^2}{8\omega^2}\end{bmatrix} & \begin{bmatrix}\frac{\alpha^1(\alpha^2)^2}{4\omega^2}\\0\end{bmatrix} \end{bmatrix}, \tag{3.34}$$

where the averaged coefficient is represented as a tensor with indices $i, j, k \in \{\,1, 2\,\}$. Since these are the only non-zero inputs, the other indices, $i, j, k \in 3, \ldots m$, of the averaged coefficient vanish. The averaged coefficient matrix is read as follows: i choose the outer row, j chooses the column, and k chooses the inner

row. The corresponding contribution to system motion is then,

$$\overline{\mathrm{V}^{(i,j,k)}_{(1,1,0)}(t)}\,[Y_i,[Y_j,Y_k]] = -\frac{\alpha^1(\alpha^2)^2}{8\omega^2}\,[Y_1,[Y_2,Y_2]] - \frac{\alpha^1(\alpha^2)^2}{8\omega_2}\,[Y_2,[Y_1,Y_2]] + \frac{\alpha^1(\alpha^2)^2}{4\omega_2}\,[Y_2,[Y_2,Y_1]]$$
$$= -\frac{3\alpha^1(\alpha^2)^2}{8\omega^2}\,[Y_2,[Y_1,Y_2]]\,, \quad (3.35)$$

for summation over the indices $i,j,k \in \{\,1,2\,\}$. With the choice of $\omega = \omega_{212}$, $\alpha^1 = \alpha^1_{212}$ and $\alpha^2 = \alpha^2_{212}$,

$$\overline{\mathrm{V}^{(i,j,k)}_{(1,1,0)}(t)}\,[Y_i,[Y_j,Y_k]] = -\frac{3\alpha_{212}}{8\omega^2_{212}}\,[Y_2,[Y_1,Y_2]]\,, \quad (3.36)$$

where $\alpha_{212} = \alpha^1_{212}(\alpha^2_{212})^2$. More generally, when $a,b \in \{\,1,\ldots,m\,\}$, this analysis implies

$$\overline{\mathrm{V}^{(c,d,e)}_{(1,1,0)}(t)}\,[Y_c,[Y_d,Y_e]] = -\frac{3\alpha_{bab}}{8\omega^2_{bab}}\,[Y_b,[Y_a,Y_b]]\,, \quad (3.37)$$

where $\alpha_{bab} = \alpha^a_{bab}(\alpha^b_{bab})^2$. In the general case, the algebraic inequalities of Equation (3.25), and Equation (3.26) according to Table 3.3, must hold with respect to other sinusoidal inputs to avoid coupling. The algebraic equalities of Equation (3.24) must only hold for those that combine to form a Jacobi-Lie bracket contribution. Again, the frequency $\omega = 2\omega_{max} + 1$ may be chosen, where ω_{max} is the maximal carrier frequency of the other sinusoidal inputs. When higher-order Jacobi-Lie bracket contributions are desired the frequency spacing will change.

■

Abstracting on the Sinusoidal Inputs

For higher-order expansions, myriad algebraic identities must hold in order to isolate specific bracket related motions. Each averaged coefficient must be examined to determine its contribution, and the limitations arising from the chosen set of input functions. Once a particular calculation is done, it need not be repeated for another problem with the same Jacobi-Lie bracket requirements. The general strategy was proven in [154, 155, 92], and also allows for the frequency choices to be positive rational numbers instead of only positive integers.

Definition 12 *Let B be an iterated Jacobi-Lie bracket. Define $\delta_a(B)$ to be the number of times that the control vector field Y_a appears in the iterated Jacobi-Lie bracket B.*

In [92], it was shown that for any Jacobi-Lie bracket B such that $\delta_i(B) = \delta_j(B)$ for all $i,j = 1\ldots m$, there is no way to obtain unique activation of the Jacobi-Lie bracket B. The cyclicly related Jacobi-Lie brackets will also be excited, as was seen in Lemma 3.

The time-periodic inputs are defined such that the averaged coefficients scale linearly with the parameters α_J. For purposes of feedback, the α_J should be defined as functions of state so as to achieve stabilization for the averaged equations of motion. There is no unique method for picking the α^I_J to obtain the desired product $\alpha_J = \Pi_{i\in I}\alpha^i_J$. Different choices of decomposing the product α_J into α^I_J will result in different controllers, although the averaged evolution will be the same. Both the Theorems of Section 3.1.3 below, and the examples in Section 3.3 will discuss this freedom and its implications.

3.1.3 Stabilization Using Sinusoids

To summarize, we have obtained the system response to a set of oscillatory control inputs for an arbitrary order of averaging. We have also analyzed the effects of the oscillatory control inputs on the averaged expansion, leading to an α-parametrized control form. Now, we must determine a stabilizing feedback strategy.

For convenience, order the Jacobi-Lie brackets as they appear in the series expansion of the averaged vector field. Let $\left\{ \widehat{\mathrm{Y}}_j \right\}$ denote the Jacobi-Lie brackets, and let $\left\{ \mathrm{T}^j \right\}$ be their corresponding averaged coefficients. With this ordering, the averaged equations of motion can be put into the form:

$$\dot{z} = X_S + \mathrm{T}^i(\alpha)\widehat{\mathrm{Y}}_i = X_S + B(z)H(\alpha), \tag{3.38}$$

where the matrices B and H are

$$B(z) = [\widehat{\mathrm{Y}}_1 \dots \widehat{\mathrm{Y}}_N], \quad \text{and} \quad H(\alpha) = [\mathrm{T}^1 \dots \mathrm{T}^N]^{\mathrm{T}}. \tag{3.39}$$

The α parameters will be used to obtain control authority over the indirectly controllable states.

The dependence of the controller design on a local Lie algebra basis is problematic if the Jacobi-Lie brackets used to satisfy the LARC vary pointwise. If the controller is to stabilize to different points, each with a unique basis spanning the Lie algebra, then multiple α-parametrizations are necessary [77]. For systems that are well-controllable, multiple parametrizations can be avoided.

Oscillatory Control via Discretized Feedback

We first consider feedback based on time discretization, while a continuous time version is developed below. State error will be used as feedback to modulate the α-parameters, converting the problem to periodic discrete feedback. It is very similar to the motion control algorithms, based on motion primitives, that use approximate inversion for open loop stabilization and trajectory tracking [21, 100]. In the language of Floquet theory, our goal is to utilize feedback so as to create stable Floquet multipliers.

Theorem 16 *Consider a system of the form (3.7) which satisfies the LARC at $q^* \in Q$. Let $u^a(t)$ be the corresponding set of α parametrized, T-periodic inputs functions where $a = 1 \dots m$ and $\alpha \in \mathbb{R}^{n-m}$. Let $z(t)$ be the averaged system response to the inputs. Given the averaged system (3.38), assuming that the m directly controlled states have been linearly stabilized and that the linearization of H from Equation (3.39) with respect to α at $\alpha = 0$ and $z = q^*$ is invertible on the $(n-m)$ dimensional subspace to control, then there exists a $K \in \mathbb{R}^{(n-m)\times n}$ such that for*

$$\alpha = -\Lambda K z(T\lfloor t/T \rfloor),$$

where $\Lambda^{(n-m)\times(n-m)}$ is invertible and $\lfloor \cdot \rfloor$ denotes the floor function, the average system response is stabilized.

proof
Given the assumptions on the system, the averaged system (3.38) is fully controllable. Linearization of the system with respect to z and α yields

$$\dot{z} = Az + B\frac{\partial H}{\partial \alpha}\alpha \equiv Az + B\Upsilon\alpha. \tag{3.40}$$

Choosing α constant over a period, the above system can be directly integrated:

$$z(T) = e^{AT}z(0) + e^{AT}\int_0^T e^{-At}\mathrm{d}tB\Upsilon\alpha.$$

Now we effectively have the discrete, linear system

$$z(k+1) = \hat{A}z(k) + \hat{B}\alpha,$$

where

$$\hat{A} = e^{AT}, \quad \text{and} \quad \hat{B} = e^{AT}\int_0^T e^{-At}\mathrm{d}tB\Upsilon.$$

The control assumptions on the system imply that the matrix $\hat{B}$ has a pseudo-inverse for the $(n-m)$ dimensional subspace that we would like to stabilize. Denote this pseudo-inverse by Λ. Now choose the matrix K such that the eigenvalues of $\hat{A} - K$ lie within the unit circle. With this choice of Λ and K, the discrete system has been stabilized, and by extension the continuous system with feedback that is piecewise constant over a period.

■

The directly controlled subspace is continuously stabilized, whereas the indirectly controlled subspace is stabilized with time-discretized feedback. The error is sampled at the end of every period, and the error is computed and used to update the α-parameters. The α-parameters then remain remain fixed for the duration of a period.

Using the analysis of Sussman and Liu [155, 93], Morin and Samson [118] derived a discretized feedback strategy which functions the same as the one proven in Theorem 16. Morin and Samson considered an extended system, which consists of an extra copy of the configuration space that evolves discretely in time. The discretely evolving configuration space is used for feedback modulation of sinusoidal inputs. We have more explicitly linked the discretely evolving system to the average of the actual oscillatory system. Additional distinctions are that for the method presented here, (1) the directly controllable subsystem is still continuously stabilized, (2) a homogeneous approximation is not needed to construct the controller, and (3) by appealing to the generalized averaging theory, it is possible to make modifications and improvements to the control strategy. The improvements follow below, some of which recover other known stabilizing controllers for underactuated, driftless nonlinear systems.

Oscillatory Control via Continuous Feedback

The control design procedure outlined above required integration of the linearized dynamical model over a period of actuation, and a discretized feedback strategy. Corollary 3 implies that stability can also be determined from the Floquet exponents, i.e., the stability of the continuous autonomous averaged vector field. A version of Theorem 16 can be proven without discretizing the closed-loop system.

Theorem 17 *Consider a system of the form (3.7) which satisfies the LARC at $q^* \in Q$. Let $u^a(t)$ be the corresponding set of α parametrized, T-periodic inputs functions where $a = 1 \dots m$ and $\alpha \in \mathbb{R}^{n-m}$. Lastly denote by $z(t)$, the averaged system response to the inputs. Given the averaged system (3.38), assuming that the m directly controlled states have been linearly stabilized and that the linearization of H with respect to α at $\alpha = 0$ and $z = q^*$ is invertible on the $(n-m)$ dimensional subspace to control, then there exists a $K \in \mathbb{R}^{(n-m)\times n}$ such that for*

$$\alpha = -\Lambda K z(t),$$

where $\Lambda^{(n-m)\times(n-m)}$ is invertible, we have stabilized the average system response.

proof
Follows the proof of Theorem 16, but without the discretization step.
■

Oscillatory Control via Lyapunov Functions

In Section 2.3, the linearization of the autonomous Floquet vector field, Z, was emphasized as a means to prove stability of the original oscillatory flow. In reality, any method to demonstrate stability of the average can be applied; for example, using a Lyapunov function for the averaged system is permissible. In this section, we demonstrate how ideas from homogeneous control theory [108, 109] can also be used to prove exponential or ρ-exponential stability. The methods provide a way to improve the controllers found in [108, 109, 136] for two reasons: (1) the Lyapunov function for the averaged system will take a simpler form and will not be time-varying, allowing for improved Lyapunov analysis, and (2) in practice, the modification introduces adjustable controller gains that can be used to improve the closed-loop system response.

Definition 13 *A* stabilized truncated series expansion with respect to the Lyapunov function, V, *for the vector field (2.32) is a truncated vector field that has the same stability property using the Lyapunov function, V, as any higher-order truncation of the vector field, and also the full series expansion of the vector field.*

Recall, the averaged system response in Equation (3.38). Consider a parametrization for α, e.g., $\alpha = p(z)$. The system in Equation (3.38) may be expressed as,

$$\dot{z} = Z(z) \equiv X_S(z) + B(z)H \circ p(z). \tag{3.41}$$

Theorem 18 *Consider a system of the form (3.7) which satisfies the LARC at $q^* \in Q$. Let $u^a(t)$ be the corresponding set of α parametrized, T-periodic inputs functions where $a = 1 \dots m$ and $\alpha \in \mathbb{R}^{n-m}$. Lastly denote by $z(t)$, the averaged system response to the inputs. Given the averaged system (3.38), assume that there exists an α-parametrization and a Lyapunov function, V, such that (3.41) is a stabilized truncated series expansion with respect to the Lyapunov function, V. If the system in Equation (3.41) is shown to be asymptotically (exponentially) stable using the Lyapunov function, then the averaged system response is asymptotically (exponentially) stable.*

proof
Follows the proof of 16, but without the discretization step. Instead of the linearization, use the Lyapunov function, V, to demonstrate stability.
■
Since the system (3.7) is nonlinear, the closed-loop system (3.41) will also contain nonlinear terms that complicate the Lyapunov analysis. In a local neighborhood of the equilibrium, some of the nonlinear terms do not affect stability and may be disregarded. Theorem 17 resolved the problem by linearizing the equations of motion; linearization emphasizes the dominant terms within a neighborhood of the equilibrium. A general strategy for extracting the dominant stabilizing terms and removing the unnecessary nonlinear terms is needed for the Lyapunov technique to be useful.

Homogeneous Systems Geometric homogeneity will be used to identify the higher order terms that can be removed. Although what follows can be described in more generality, for brevity we follow the material from [62, 108]. Without loss of generality, the equilibrium point will be translated to the origin of $\mathbb{R}^n$.

Definition 14 *Let $x = (x^1, \ldots, x^n)$ denote a set of coordinates for $\mathbb{R}^n$. A* dilation *is a mapping $\Delta_\lambda^r : \mathbb{R}^+ \times \mathbb{R}^n \to \mathbb{R}^n$ defined by*

$$\Delta_\lambda^r(x) = (\lambda^{r_1} x^1, \ldots, \lambda^{r_n} x^n),$$

where $r = (r_1, \ldots r_n)$ are n positive rationals such that $r_1 = 1 \leq r_i \leq r_j \leq r_n$ for $1 < i < j < n$.

Typically r, called the *scaling*, will be given and the notation Δ_λ will be used in place of Δ_λ^r.

Definition 15 *A continuous function $f : \mathbb{R}^n \to \mathbb{R}$ is homogeneous of degree $l \geq 0$ with respect to Δ_λ if $f(\Delta_\lambda(x)) = \lambda^l f(x)$.*

Definition 16 *The homogeneous space, $\mathcal{H}_k$, is the set of all continuous functions, $f : \mathbb{R}^n \to \mathbb{R}$, with homogeneous degree k.*

Definition 17 *A continuous vector field $X(x,t) = X^i(x,t)\frac{\partial}{\partial x^i}$ is homogeneous of degree $m \leq r_n$ with respect to Δ_λ, if $X^i(\cdot, t) \in \mathcal{H}_{r_i - m}$.*

Definition 18 *A continuous map from $\rho : \mathbb{R}^n \to \mathbb{R}$, is called a* homogeneous norm *with respect to the dilation Δ_λ, when*

1. $\rho(x) \geq 0, \;\; \rho(x) = 0 \iff x = 0.$
2. $\rho(\Delta_\lambda(x)) = \lambda \rho(x), \;\; \forall\, \lambda > 0.$

The notion of homogeneous degree and the homogeneous norm are used to characterize vector fields and truncatable components of vector fields.

Definition 19 *The* m^{th}-degree homogeneous truncation of a vector field, *denoted* $\mathrm{Trunc}_m^\Delta(\cdot)$, *is the truncation obtained by removing all terms of homogeneous degree greater than m.*

The m-degree homogeneous truncation with respect to the dilation Δ results in the differential equation,

$$\dot{z} = Z_m = \mathrm{Trunc}_m^\Delta \left(X_S(z) + B(z) H \circ p(z) \right). \tag{3.42}$$

If there exists a Lyapunov function such that the system in Equation (3.42) is stable, then the averaged system is locally stable.

Theorem 19 *Consider a system of the form (3.7) which satisfies the LARC at $0 \in \mathbb{R}^n$. Let $u^a(t)$ be the corresponding set of α parametrized, T-periodic inputs functions where $a = 1 \ldots m$ and $\alpha \in \mathbb{R}^{n-m}$. Lastly denote by $z(t)$, the averaged system response to the inputs. Given the averaged system (3.38), assuming that there exists an α-parametrization and a Lyapunov function such that the homogeneous truncation in Equation (3.42) is asymptotically (ρ-exponentially) stable with respect to the Lyapunov function, then the actual system response is asymptotically (ρ-exponentially) stable.*

proof
Rosier [142, Theorem 3] has proven that the homogeneous truncation of Equation (3.42) will preserve the dominant terms vis-a-vis stability. If the Lyapunov function can be used to demonstrate stability of the homogeneous truncation (3.42), then locally the Lyapunov function is stabilized with respect to the averaged expansion of Equation (3.38). Invocation of Theorem 18 completes the proof.
■

To achieve ρ-exponential stability for the dynamical system in Equation (3.41), the homogeneous truncation

of order 0 must be stabilizing. The 0-degree homogeneous truncation with respect to the dilation Δ results in the system

$$\dot{z} = Z_0 = \mathrm{Trunc}_0^{\Delta}\left(X_S(z) + B(z)H \circ p(z)\right). \tag{3.43}$$

If there exists a Lyapunov function such that the system in Equation (3.43) is stable, then the averaged system is locally ρ-exponentially stable.

Corollary 5 *Consider a system of the form (3.7) which satisfies the LARC at $0 \in \mathbb{R}^n$. Let $u^a(t)$ be the corresponding set of α parametrized, T-periodic inputs functions where $a = 1 \ldots m$ and $\alpha \in \mathbb{R}^{n-m}$. Lastly denote by $z(t)$, the averaged system response to the inputs. Given the averaged system (3.38), assuming that there exists an α-parametrization and a Lyapunov function such that the 0-degree homogeneous truncation in Equation (3.43) is asymptotically stable with respect to the Lyapunov function, then the actual system response is ρ-exponentially stable.*

If the parametrization is contructed so that the averaged coefficients are of homogeneous order 0 and are stabilizing, then the existence of a Lyapunov function for the homogeneous truncation of the averaged vector field is guaranteed by Rosier [142]. Hence, there is no need to find the Lyapunov function in this case.

Corollary 6 *Consider a system of the form (3.7) which satisfies the LARC at $0 \in \mathbb{R}^n$. Let $u^a(t)$ be the corresponding set of α parametrized, T-periodic inputs functions where $a = 1 \ldots m$ and $\alpha \in \mathbb{R}^{n-m}$. Lastly denote by $z(t)$, the averaged system response to the inputs. Given the averaged system (3.38), assuming that there exists an α-parametrization such that the 0-degree homogeneous truncation in Equation (3.43) is asymptotically stable, then the actual system response is ρ-exponentially stable.*

In M'Closkey [108], dilation is used in the construction of the stabilizing controller. Consequently, the dilation is adapted to the Lie algebra of the underactuated driftless control system[2], which is why ρ-norms must be used. As a consequence, only ρ-exponential stability can be proven. The Lyapunov function used to show ρ-exponential stability in Theorem 19 will typically be a power of the ρ-norm,

$$L(z) = \rho^{2r_n}(z). \tag{3.44}$$

We have provided an alternative method for constructing the time-varying controllers, which only uses the dilation and homogeneous norm to define a homogeneous truncation. Consequently, the dilation need not be adapted to the Lie algebra, but may be the standard dilation where $r_i = 1$. Under the standard dilation, the 2-norm may be used. In this case, Corollary 6 improves upon Theorem 18. It does so, not only by using the standard 2-norm, $||\cdot||$, and feedback of homogeneous order 0 under the 2-norm, but by also proving stabilization of the Lyapunov function with respect truncations of the averaged expansions. In practice the Lyapunov function used to show exponential stability will be the 2-norm squared,

$$L(z) = ||z||^2. \tag{3.45}$$

Just like in the theorems using linearization of the average, the homogeneous truncation focuses on the dominant contributions within a local neighborhood of the equilibrium; local ρ-exponential (exponential) stability is achieved.

Comments on the Stabilizing Controllers

Theorems 16-19 and Corollaries 5-6 stabilize an equilibrium point of the averaged system. If the α-parametrized control input functions do not vanish at the equilibrium q^*, then by Theorem 9, the flow of the

[2] In a dilation adapted to the Lie algebra, the scalings r will not satisfy $r^i = r^j$ for all $i, j \in [1, n]$. See M'Closkey [108] for more details.

actual system stabilizes to an orbit (of size $O(\epsilon)$) around the fixed point. If, on the other hand, the input functions do vanish at the equilibrium, then Corollary 2 implies that the flow of the actual system stabilizes to the fixed point (i.e., the orbit collapses to the fixed point).

The states requiring continuous feedback use the average rather than the instantaneous values of the system for state feedback. Trajectories of the actual flow are related to the average flow by the Floquet mapping,

$$x(t) = P(t)(z(t)). \tag{3.46}$$

We may solve for the average $z(t)$, using the current state $x(t)$. Since $P(t)$ is given by a series expansion, we can easily compute its inverse.

For the discretized feedback strategy, the difference in the averaged and instantaneous indirectly controlled states is not a critical factor to consider due to the fact that $P(t)$ is periodic, i.e., $P(kT) = P(0) = \mathrm{Id},\ k \in \mathbb{Z}^+$. The directly stabilized states do not use discretized feedback, and consequently do require the average to be used as feedback. If the Floquet mapping $P(t)$ is ignored and instantaneous states $x(t)$ are used in the feedback control of the directly controllable states, this averaging method will place an upper bound on the feedback gains. The oscillatory inputs should be faster than the natural stabilizing dynamics of the directly stabilized subsytem, otherwise there will be attenuation of the oscillatory signal (and consequently the feedback signal to the indirectly controlled states). With the exception of [118], this effect has not been discussed in prior presentations of feedback strategies that utilize averaging techniques [108, 136].

In practice, one may utilize averages computed in realtime as continuous feedback. The benefit of this latter approach is that the averaging process may serve to filter out noise in the sensor signals. It will also attenuate the feedback of external disturbances. As the continually computed average, $\bar{x}(t) = \frac{1}{T}\int_{t-T}^{t} x(\tau)\,\mathrm{d}\tau$, is not quite the same as $z(t) = P^{-1}(t)\,(x(t))$, there may be some differences. When performing averaging of sensed measurements, calculate $P(t)$ to determine which states require averaging. Recall that $P(t)$ is a periodic function with a power series representation in ϵ. The states with dominant oscillatory terms will require averaging, whereas the states that have no oscillatory terms, or ignorable, perturbative oscillatory terms, in $P(t)$ will not require averaging.

Trajectory Tracking. To track a trajectory, replace $x(t)$ with $x(t) - x_d(t)$; the system must be locally controllable along the trajectory. As the trajectory evolves, the Jacobi-Lie brackets contributing to the Lie algebra rank condition may vary. The trajectory tracking feedback strategy is more complicated since multiple parametrizations will be needed. If a system is well controllable, then a single parametrization will suffice to track trajectories. Systems with group symmetris are well controllable if and only if there exists a Lie algebra basis at the group identity. Lastly, for the discretized feedback, the Nyquist criteria is a limiting factor in tracking a trajectory for the indirectly controlled states.

Existence and Continuity of Solutions. M'Closkey [108] has shown that any exponentially stabilizing feedback-law is necessarily non-Lipschitz. The feedback laws created by the algorithm in [108] are smooth on $Q \setminus \{ q^* \}$, and continuous at $q^* \in Q$. The stabilization is (ρ-exponentially) asymptotic; although the control system may stabilize to a neighborhood of the origin in finite time, it will not stabilize to the origin in finite time. The stabilizing feedback strategies using the generalized averaging theory are also smooth on $Q \setminus \{ q^* \}$, and continuous at $q^* \in Q$ (some are piecewise smooth and piecewise continuous, respectively). According to [108], solutions to the series expansions exist and are unique. On a related note, conditions for the continuity of solution trajectories with respect to the control inputs may be desired. Liu and Sussman [94, 95] have published work regarding the continuous dependence of trajectories with respect to the input functions.

3.2 Mechanical Systems with Nonholonomic Constraints

This section specializes the above results to systems with nonholonomic constraints; for more background see Bloch et al. [15]. It is assumed that the constraints are linear nonholonomic constraints; the generalization to affine constraints should follow naturally from the analysis of this section and of [15]. Initially, the focus is on systems whose constraints may be defined using an Ehresmann connection form. Lie group symmetries are subsequently added, resulting in a principal connection form. The averaged expansions of the underactuated driftless affine control system, Equation (3.1), may be simplified by incorporating the additional structure imposed by the connections forms. The connection forms may also be used to refine the Lie algebra rank condition, and consequently the controllability verification of mechanical systems with nonholonomic constraints. For more information on the background differential geometry, see the references [69, 70].

3.2.1 Systems with Nonholonomic Constraints

The control system (3.1) evolves on a manifold Q, called the *configuration space* which is assumed to have a bundle structure, $\pi : Q \to R$, where R is a manifold called the base space. In addition to the base projection, π, there is a complementary fiber projection, $\pi_S : Q \to S$, [70]. The mapping, π, may be used to define a vertical subspace to the tangent bundle, TQ. It is given by the kernel of the tangent mapping $T_q\pi : T_qQ \to T_{\pi(q)}R$.

Definition 20 *The* vertical subspace *of T_qQ, denoted, V_qQ, is the kernel of the bundle projection,*

$$V_qQ = \{\, u_q \mid T\pi(u_q) = 0 \,\}.$$

The horizontal subspace is determined by the constraints, which can be described via the vanishing of an Ehresmann connection.

Definition 21 *[15] An* Ehresmann connection, *A, is a vertical-valued one-form on Q that satisfies*

1. *$A_q : T_qQ \mapsto V_qQ$ is a linear map $\forall q \in Q$.*
2. *A is a projection, $A(v_q) = v_q$, $\forall q \in Q$.*

The Ehresmann connection defines a splitting of the tangent space into horizontal and vertical vectors. The horizontal subspace $H_qQ \subset T_qQ$ at $q \in Q$ is given by the kernel of the Ehresmann connection.

In a local trivialization, $Q \cong R \times S$, the constraint equations take the form,

$$\dot{s} + A_{loc}(r, s) \cdot \dot{r} = 0, \tag{3.47}$$

where A_{loc} is the local form of the Ehresmann connection [15]. In local coordinates it is mapping from $Q \times_M TM$ to TS, where $Q \times_M TM$ is the fiber product with base M (see Section 4.1.1). For every point $q \in Q$, there exists a isomorphic mapping taking elements in $T_{\pi(q)}R$ to elements in H_qQ, called the *horizontal lift*. In local coordinates, $q = (r, s)$, it operates as follows,

$$v_r \in T_rR, \quad v_r^h = (v_r, -A_{loc}(r, s) \cdot v_r) \in H_qQ. \tag{3.48}$$

Underactuated Driftless Control Systems. The definitions and operations above extend to vector fields, which are sections on the tangent bundle.

Definition 22 *The horizontal subspace of the space of vector fields on Q, is denoted by $\mathcal{X}^H(Q) \subset \mathcal{X}(Q)$. It is the subspace of vector fields which are horizontal for all points in Q. The vertical subspace of the space of vector fields on Q is denoted by $\mathcal{X}^V(Q) \subset \mathcal{X}(Q)$. It is the subspace of vector field which are vertical for all points in Q.*

Definition 23 *[69] Given an Ehresmann connection and a vector field on R, $X \in \mathcal{X}(R)$, there is a uniquely defined horizontal vector field on Q, denoted by $X^h \in \mathcal{X}^H(Q)$, such that*

$$T\pi \circ X^h = X.$$

The horizontal vector field, X^h is called the horizontal lift of X. *Conversely, given a horizontal vector field $X^h \in \mathcal{X}^H(Q)$, there exists a vector field $X \in \mathcal{X}(R)$, whose horizontal lift is X^h.*

Typically the base space, R, will be directly actuated and fully controllable. Consequently, there exists a collection of control vector fields, $\{ Y_a \}_1^m$, where $m = \dim(R)$ and $Y_a \in \mathcal{X}(M)$. The horizontal lift of the control vector fields, in local coordinates, may be used to give the governing equations for an underactuated driftless affine system,

$$\begin{aligned} \dot{r} &= Y_a u^a \\ \dot{s} &= -A_{loc}(r,s) \cdot Y_a u^a \end{aligned} \tag{3.49}$$

Therefore, under constraints defined by an Ehresmann connection, the standard form for an underactuated driftless affine control system, given in Equation (3.1), may be rewritten as,

$$\dot{q} = Y_a^h u^a. \tag{3.50}$$

The involutive closure at $q \in Q$ is $\mathcal{C}^\infty(q)$, c.f. Equation (3.5), and consists of iterated Jacobi-Lie brackets of horizontally lifted control vector fields. Using the structure of the Ehresmann connection, it is possible to refine the Lie algebra rank condition from Equation (3.6).

The Lie Algebra Rank Condition. Since vertical vector fields in $\mathcal{X}^V(Q)$ are π-related to the zero vector field in $\mathcal{X}(R)$, it is possible to show that

$$\left[X^h, Y^v \right] \in \mathcal{X}^V(Q), \tag{3.51}$$

where $X \in \mathcal{X}(R)$ and $Y^v \in \mathcal{X}^V(Q)$. In other words, the Jacobi-Lie bracket between a horizontal and a vertical vector field is again vertical. Additionally, the curvature of the Ehresmann connection, denoted by B, is given by,

$$B(X^h, Y^h) = \left[X^h, Y^h \right]. \tag{3.52}$$

Importantly, the curvature is vertical valued. Consequently, the only vector fields in $\mathcal{C}^\infty$ with a nontrivial horizontal projection are the horizontal lifts of the control vector fields.

Theorem 20 *Under the assumption that the control vector fields, $Y_a \in \mathcal{X}(R)$, span the tangent space $T_{\pi(q)}R$ for $q \in Q$, the Lie algebra rank condition at q is equivalent to the condition that*

$$\operatorname{span} \left\{ \left[Y_{i_k}^h \ldots, \left[Y_{i_2}^h, \left[Y_{i_1}^h, Y_{i_0}^h \right] \right] \right] (q) \right\} = V_q Q, \tag{3.53}$$

for $k \in \mathbb{Z}^+ \setminus \{ 0 \}$.

The Ehresmann connection can be used to define the notion of parallel transport. Parallel transport is given by flow along horizontal vector fields,

$$\mathrm{Pt}_{0,t}^{X} = \Phi_{0,t}^{X^h}, \quad X \in \mathcal{X}(R). \tag{3.54}$$

Parallel transport naturally leads to the notion of a covariant derivative for elements $f \in C^\infty(R)$,

$$\nabla_X f \equiv \mathcal{L}_{X^h} f \tag{3.55}$$

Corollary 7 *[69] Under the assumption that the control vector fields, $Y_a \in \mathcal{X}(R)$, span the tangent space $T_{\pi(q)}R$ for $q \in Q$, the Lie algebra rank condition at q is equivalent to the condition that*

$$\mathrm{span}\left\{ \nabla_{Y_{i_k}} \cdots \nabla_{Y_{i_2}} B(Y_{i_1}^h, Y_{i_0}^h)(q) \right\} = V_q Q, \tag{3.56}$$

for $k \in \mathbb{Z}^+ \setminus \{0\}$.

Because the curvature and its higher order covariant derivative are all vertical, the local curvature form may be used instead. The local curvature form may be obtained from the complementary projection,

$$B_{loc}(X,Y) \equiv T\pi_S \circ B(X^h, Y^h), \quad \forall\, X, Y \in \mathcal{X}(R). \tag{3.57}$$

The covariant derivative, ∇, induces a covariant derivative, $\widetilde{\nabla}$, on the local curvature form via the identity,

$$\widetilde{\nabla}_X B_{loc}(Y,Z) = T\pi_S \left(\nabla_X B(Y^h, Z^h) \right), \quad \forall\, X, Y, Z \in \mathcal{X}(R). \tag{3.58}$$

Corollary 8 *Under the assumption that the control vector fields, $Y_a \in \mathcal{X}(R)$, span the tangent space $T_{\pi(q)}R$ for $q \in Q$, the Lie algebra rank condition at q is equivalent to the condition that*

$$\mathrm{span}\left\{ \widetilde{\nabla}_{Y_{i_k}} \cdots \widetilde{\nabla}_{Y_{i_2}} B_{loc}(Y_{i_1}, Y_{i_0})(q) \right\} = T_{\pi_S(q)} S, \tag{3.59}$$

for $k \in \mathbb{Z}^+ \setminus \{0\}$.

The Lie algebra rank condition has been simplified to the condition that the local form of the curvature and its covariant derivatives span the tangent to the fiber.

Controllability of Systems with Constraints. From Kelly and Murray [67], we may define the notion of strong and weak fiber controllability.

Definition 24 *A system (3.50) is said to be* strongly fiber controllable *if, for any $q_i, q_f \in Q$, there exists a time $T > 0$ and a curve $r(\cdot) \in R$ connecting $r_i = \pi(q_i)$ and $r_f = \pi(q_f)$ such that the horizontal lift of $r(\cdot)$ passing through q_i satisfies, $r^h(0) = q_i$ and $r^h(T) = q_f$.*

Strong controllability is the standard definition of controllability interpreted using an Ehresmann connection form and its horizontal lift. For systems with constraints, trajectories in the configuration space must be horizontal lifts of trajectories in the base space. Kelly and Murray further noted that some locomotive systems did not require strong controllability, since the evolution of the base space was not of primary importance. By not imposing final conditions on the base space, weak controllability may be defined.

Definition 25 *A system (3.50) is said to be* weakly controllable *if, for any $q_i \in Q$ and $s_f \in S$, there exists a finite time $T > 0$ and a curve in the base space $r(\cdot)$ whose horizontal lift passing through q_i at time $t = 0$ satisfies, $r^h(T) = q_f$, where $s_f = \pi_S(q_f)$.*

Weak controllability does not place restrictions on the final point of the base space. Mobile robots whose base space corresponds to the angles of wheels are systems that may be weakly controlled. By not restricting the final angle of the wheel, greater freedom is allowed. Most importantly, the need to stabilize or control the base space no longer holds for weakly controlled systems. Since controllability of only the fiber is sought, the control brackets may also be used to span the vertical space.

Theorem 21 *For the system (3.50), the following are equivalent:*

1. *the system is small-time weakly controllable,*
2. $\operatorname{span}\left\{ T_q\pi_S\left(\left[Y_{i_k}^h \ldots, \left[Y_{i_3}^h, \left[Y_{i_2}^h, Y_{i_1}^h\right]\right]\right](q)\right)\right\} = T_{\pi_S(q)}S,$
3. $\operatorname{span}\left\{ \left\{ T_q\pi_S\left(\nabla_{Y_{i_k}} \cdots \nabla_{Y_{i_2}} B(Y_{i_1}^h, Y_{i_0}^h)(q)\right)\right\} \bigcup \left\{ T_q\pi_S(Y_{i_0}^h)\right\}\right\} = T_{\pi_S(q)}S,$
4. $\operatorname{span}\left\{ \left\{ \widetilde{\nabla}_{Y_{i_k}} \cdots \widetilde{\nabla}_{Y_{i_2}} B_{loc}(Y_{i_1}, Y_{i_0})(q)\right\} \bigcup \left\{ T_q\pi_S(Y_{i_0}^h)\right\}\right\} = T_{\pi_S(q)}S,$

for $k \in \mathbb{Z}^+ \setminus \{0\}$ *and* $a = 1 \ldots m$.

The theorem follows from the definition of weak controllability, Theorem 20, and Corollaries 7 and 8.

The Averaged Expansions. Suppose that the system (3.50) were to be weakly fiber controllable, and there was no need to stabilize the base space. Then the averaged expansions from Section 3.1 may be simplified. Suppose furthermore that controllability was obtained through the curvature and higher-order covariant derivatives.

Third-order averaging using the averaged expansion of Equation (3.19) is

$$\begin{Bmatrix} \dot{r} \\ \dot{s} \end{Bmatrix} = \begin{Bmatrix} 0 \\ \overline{\frac{1}{2}\mathrm{V}_{(1,0)}^{(\alpha,\beta)}(t)} B_{\alpha\beta}(r,s) + \overline{\frac{1}{3}\mathrm{V}_{(1,1,0)}^{(\alpha,\beta,\gamma)}(t)} B_{\alpha\beta\gamma}(r,s) \end{Bmatrix} \tag{3.60}$$

where $B_{\alpha\beta}$ is the local form of the curvature form, defined as follows,

$$B_{\alpha\beta} \equiv B_{loc}(Y_\alpha, Y_\beta),$$

and $B_{\alpha\beta\gamma}$ is the covariant derivative of the local curvature form,

$$B_{\alpha\beta\gamma} \equiv \nabla_{Y_\alpha} B_{loc}(Y_\beta, Y_\gamma).$$

Due to the structure of the vector fields, there will never be a contribution to the base space unless the first order average is non-vanishing. This means that all expansions will only affect the fiber space. In the average, the evolution of the base space is trivial, a very nice result when dealing with systems of this type. Since, in the average, the base variables are constant, replacing them by their averages in the equations of motion reduces Equation (3.60) to

$$\dot{s} = \frac{1}{2}\overline{\mathrm{V}_{(1,0)}^{(\alpha,\beta)}(t)} B_{\alpha\beta}(\overline{r}, s) + \frac{1}{3}\overline{\mathrm{V}_{(1,1,0)}^{(\alpha,\beta,\gamma)}(t)} B_{\alpha\beta\gamma}(\overline{r}, s) \tag{3.61}$$

The equations have been simplified to evolution of the fiber only, possibly rendering further analysis tractable. Higher-order averages have the same simplifying property.

Many of these results can be found in the literature when the system with nonholonomic constraints also involves Lie groups symmetries [137, 67]. The next section discuss the consequences of reduction under group symmetry.

3.2.2 Systems with Nonholonomic Constraints and Symmetries

This section examines the simplifications that occur when symmetries are added to nonholonomic systems with constraints. The symmetry involved is group invariance with the Lie group G, where G acts on the configuration manifold Q. The action of G will be denoted by Φ_g, and maps $g \mapsto gq$. The action gives the manifold Q a principal bundle structure, $\pi : Q \to Q/G$. The base space is $M = Q/G$. The reader is referred to Marsden and Ratiu [99] for more background on Lie groups and geometric mechanics.

For the Lie group, G, there is a corresponding Lie algebra $\mathfrak{g}$, defined using the tangent space of the group identity, T_eG. The Lie algebra elements act as vector fields on the configuration space Q via the infinitesimal generator,

$$\xi \cdot q \equiv \xi_Q(q) \equiv \frac{\mathrm{d}}{\mathrm{d}t}\Phi_{\exp \xi t}.$$

Assume that the entire system is invariant under the group action. In particular, the constraints are G-invariant, i.e.,

$$T\Phi_g^* A(q, \dot{q}) \equiv A(T\Phi_g(q, \dot{q})) = A(q, \dot{q}) = 0.$$

Group invariance means that the system may be reduced. The Ehresmann connection form reduces to a *principal connection* form.

Definition 26 *[15] A* principal connection *on the principal bundle $\pi : Q \to Q/G$ is a map $\mathcal{A} : TQ \mapsto \mathfrak{g}$ that is linear on each tangent space (e.g., $\mathcal{A}$ is a Lie-algebra valued one-form) and is such that*

1. $\mathcal{A}((_Q q)) = \xi$, $\forall \xi \in \mathfrak{g}$ *and* $q \in Q$.
2. $\mathcal{A}$ *is* Ad*-equivariant:*

$$\mathcal{A}(T_q\Phi_g(v_q)) = \mathrm{Ad}_g \mathcal{A}(v_q),$$

for all $v_q \in T_qQ$ and $g \in G$, where Ad *denotes the adjoint action of G on $\mathfrak{g}$.*

Notions of horizontality and verticality are inherited from the Ehresmann connection, except that now the G-invariant vertical space can be, pointwise, isomorphically identified with the Lie algebra, $\mathfrak{g}$. The equations of motion reduce also to the Lie algebra, and in a trivialization the oscillatory control equation (3.49) is now given by

$$\begin{aligned} \dot{r} &= Y_a(r)u^a \\ \xi &= -\mathcal{A}_{loc}(r)Y_a(r)u^a, \end{aligned} \tag{3.62}$$

where the Lie group evolution is obtained from the *reconstruction equation*,

$$\dot{g} = -g\xi\,, \quad g(0) = g_0. \tag{3.63}$$

The local form of the principal connection is a Lie-algebra valued 1-form, $\mathcal{A}_{loc} : TM \to \mathfrak{g}$.

Lie algebra rank condition. Due to the identification of the vertical vector space with the Lie algebra, Theorem 20 may be reinterpreted for systems evolving on a principal bundle with a principal connection.

Theorem 22 *[137] Under the assumption that the control vector fields, $Y_a \in \mathcal{X}(M)$, span the tangent space $T_{\pi(q)}M$ for $q \in Q$, the Lie algebra rank condition at q is equivalent to*

$$\mathrm{span}\left\{ \mathcal{A}\left(\left[Y_{i_k}^h \ldots, \left[Y_{i_2}^h, \left[Y_{i_1}^h, Y_{i_0}^h \right] \right] \right] \right) \right\} = \mathfrak{g}. \tag{3.64}$$

for $k \in \mathbb{Z}^+ \setminus \{0\}$.

The principal connection can be used to define the notion of parallel transport. Parallel transport is given by flow along horizontal vector fields,

$$\mathrm{Pt}^X_{0,t} = \Phi^{X^h}_{0,t}, \quad X \in \mathcal{X}(M). \tag{3.65}$$

Using the parallel transport, it is possible to define a covariant derivative for elements in $C^\infty M$,

$$\nabla_X f = \mathcal{L}_{X^h} f \tag{3.66}$$

Corollary 9 *[137] Under the assumption that the control vector fields, $Y_a \in \mathcal{X}(R)$, span the tangent space $T_{\pi(q)}R$ for $q \in Q$, the Lie algebra rank condition at q is equivalent to*

$$\mathrm{span}\left\{ \mathcal{A}\left(\nabla_{Y_{i_k}} \cdots \nabla_{Y_{i_2}} \mathcal{B}(Y^h_{i_1}, Y^h_{i_0}) \right) \right\} = \mathfrak{g}, \tag{3.67}$$

for $k \in \mathbb{Z}^+ \setminus \{0\}$.

Because the principal connection is Ad-equivariant and its higher order covariant derivative are all vertical, the curvature form may be identified with an element of the associated adjoint bundle, $\widetilde{\mathfrak{g}}$, of the principal bundle,

$$\widetilde{\mathfrak{g}} \equiv Q \times_G \mathfrak{g} \equiv (Q \times \mathfrak{g})/G. \tag{3.68}$$

Sections of the associated adjoint bundle are in $1-1$ correspondence with Ad-equivariant Lie algebra valued functions. The local form of the principal connection, $\mathcal{A}_{loc}$, is such a section.

The local curvature form may be obtained from the principal connection, via covariant differantiation on the associated adjoint bundle,

$$\mathcal{B}_{loc}(X,Y) \equiv \widetilde{\nabla}_X \mathcal{A}_{loc}(Y), \quad \forall\, X, Y \in \mathcal{X}(R). \tag{3.69}$$

Corollary 10 *[137] Under the assumption that the control vector fields, $Y_a \in \mathcal{X}(M)$, span the tangent space $T_{\pi(q)}M$ for $q \in Q$, the Lie algebra rank condition at q is equivalent to the condition that*

$$\mathrm{span}\left\{ \widetilde{\nabla}_{Y_{i_k}} \cdots \widetilde{\nabla}_{Y_{i_2}} \mathcal{B}_{loc}(Y_{i_1}, Y_{i_0})(r) \right\} = \mathfrak{g}, \tag{3.70}$$

for $k \in \mathbb{Z}^+ \setminus \{0\}$.

The Lie algebra rank condition has been simplified to the condition that the local form of the curvature and its covariant derivatives span the tangent to the fiber.

Controllability of Systems with Constraints. It was shown in the analysis for systems with an Ehresmann connection that weak controllability does not place the same restrictions on the Lie algebra as strong controllability. Analogous results hold for the principal connection.

Theorem 23 *For the system (3.62), the following are equivalent:*

1. *the system is small-time weakly controllable,*
2. $\mathrm{span}\left\{ \left(\left[Y^h_{i_k} \ldots, \left[Y^h_{i_3}, \left[Y^h_{i_2}, Y^h_{i_1} \right] \right] \right] (q) \right) \right\} = V_q Q,$
3. $\mathrm{span}\left\{ \left\{ \left(\widetilde{\nabla}_{Y_{i_k}} \cdots \widetilde{\nabla}_{Y_{i_2}} \mathcal{B}_{loc}(Y_{i_1}, Y_{i_0})(q) \right) \right\} \bigcup \left\{ \mathcal{A}_{loc}(Y_{i_0}) \right\} \right\} = \mathfrak{g},$

for $k \in \mathbb{Z}^+ \setminus \{0\}$ and $a = 1 \ldots m$.

For more information on this derivation, see Cendra et al. [29, 30] or Radford and Burdick [137].

The Averaged Expansions. Suppose that the system (3.62) were to be weakly fiber controllable, and there was no need to stabilize the base space. Then the averaged expansions from Section 3.1 may be simplified. Suppose furthermore that controllability was obtained through the curvature and higher-order covariant derivatives.

Third-order averaging using the averaged expansion of Equation (3.19) is

$$\left\{ \begin{array}{c} \dot{r} \\ \dot{g} \end{array} \right\} = \left\{ \begin{array}{c} 0 \\ -g\left(\frac{1}{2}\overline{\mathrm{V}^{(\alpha,\beta)}_{(1,0)}(t)}\mathcal{B}_{\alpha\beta}(r,s) + \frac{1}{3}\overline{\mathrm{V}^{(\alpha,\beta,\gamma)}_{(1,1,0)}(t)}\mathcal{B}_{\alpha\beta\gamma}(r,s) \right) \end{array} \right\} \tag{3.71}$$

where $\mathcal{B}_{\alpha\beta}$ is the local form of the curvature form, defined as follows,

$$\mathcal{B}_{\alpha\beta} \equiv \mathcal{B}_{loc}(Y_\alpha, Y_\beta),$$

and $\mathcal{B}_{\alpha\beta\gamma}$ is the covariant derivative of the local curvature form,

$$\mathcal{B}_{\alpha\beta\gamma} \equiv \widetilde{\nabla}_{Y_\alpha}\mathcal{B}_{loc}(Y_\beta, Y_\gamma).$$

Due to the structure of the vector fields, there will never be a contribution to the base space unless the first order average is non-vanishing. This means that all expansions will only affect the fiber space. In the average, the evolution of the base space is trivial ince, in the average, the base variables are constant. Replacing them by their averages in the equations of motion reduces Equation (3.71) to

$$\dot{g} = -g\left(\frac{1}{2}\overline{\mathrm{V}^{(\alpha,\beta)}_{(1,0)}(t)}\mathcal{B}_{\alpha\beta}(\overline{r},s) + \frac{1}{3}\overline{\mathrm{V}^{(\alpha,\beta,\gamma)}_{(1,1,0)}(t)}\mathcal{B}_{\alpha\beta\gamma}(\overline{r},s) \right) \tag{3.72}$$

The equations have been simplified to evolution of the fiber only. Higher-order averages have the same simplifying property.

3.3 Examples

Several examples will be examined to demonstrate the applicability of this approach. In some cases, the controllers constructed using the generalized averaging theory are compared to known stabilizing controllers. The first example in Section 3.3.1 is a three-state nonholonomic integrator. Being the simplest type of underactuated nonlinear driftless control system, all of the controller strategies are worked out to demonstrate how different controllers may be derived from different choices of the α-parametrization. Additionally, the difference between averaged coefficients with and without definite integrals will be studied using the nonholonomic integrator.

In Section 3.3.2, the Hilare robot is examined. Because the Hilare robot evolves on the Lie group $SE(2)$, the Lie group exponential is used instead. It is possible to provide a stabilizing controller without the use of coordinate transformations. The essential information needed for stability is found in the averaged coefficients and Jacobi-Lie brackets. Lastly, a discretized control strategy is given for the kinematic car in Section 3.3.3. The controller is quickly derived using the procedure of this chapter. The discretized strategy can be improved via modification to a continuous feedback strategy.

3.3.1 Nonholonomic Integrator

Here, a three-state nonholonomic integrator is presented and worked out in detail for the reader. The various control strategies applicable via averaging theory are demonstrated. The nonholonomic integrator in $\mathbb{R}^3$ is a

driftless system with two control inputs, m = 2,

$$\dot{q} = Y_a(q)u^a(q,t), \quad \text{for } a = 1 \dots m,$$

where,

$$Y_1(q) = \left\{ \begin{array}{c} 1 \\ 0 \\ -q_2 \end{array} \right\} \quad \text{and} \quad Y_2(q) = \left\{ \begin{array}{c} 0 \\ 1 \\ q_1 \end{array} \right\}$$

The system is not linearly controllable at the origin, but does satisfy the LARC. To see this, take the Lie bracket of the two input vector fields,

$$[Y_1(q), Y_2(q)] = \left\{ \begin{array}{c} 0 \\ 0 \\ 2 \end{array} \right\}.$$

Only second-order averaging will be necessary. Nevertheless, this example will serve as an excellent introduction to the capabilities of the theory. For the directly controllable states, define the state-feedback to be

$$f^1(q) = -k_1 q_1, \quad f^2(q) = -k_2 q_2.$$

As per Lemma 2, out of phase sinusoids will be used to excite the critical averaged coefficient. Define

$$v^1(t) = \alpha_{12}^1 \sin(t)\,, \quad v^2(t) = \alpha_{12}^2 \cos(t). \tag{3.73}$$

This choice of control inputs results in the averaged coefficients,

$$\overline{\mathrm{V}_{(1)}^{(a)}(t)} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad , \quad \overline{\mathrm{V}_{(1,0)}^{(a,b)}(t)} = \begin{bmatrix} 0 & -\frac{1}{2}\alpha_{12} \\ \frac{1}{2}\alpha_{12} & 0 \end{bmatrix},$$

where $\alpha_{12} = \alpha_{12}^1 \alpha_{12}^2$. The averaged equations of motion are

$$\frac{\mathrm{d}}{\mathrm{d}t} \left\{ \begin{array}{c} z_1 \\ z_2 \\ z_3 \end{array} \right\} = \left\{ \begin{array}{c} -k_1 z_1 \\ -k_2 z_2 \\ -k_1 z_1 z_2 - k_2 z_1 z_2 \end{array} \right\} + \epsilon \alpha_{12} \left\{ \begin{array}{c} 0 \\ 0 \\ 1 \end{array} \right\} \tag{3.74}$$

The linearized averaged equations of motion are

$$\frac{\mathrm{d}}{\mathrm{d}t} \left\{ \begin{array}{c} z_1 \\ z_2 \\ z_3 \end{array} \right\} = \left\{ \begin{array}{c} -k_1 z_1 \\ -k_2 z_2 \\ 0 \end{array} \right\} + \epsilon \alpha_{12} \left\{ \begin{array}{c} 0 \\ 0 \\ 1 \end{array} \right\}, \tag{3.75}$$

where linearization is with respect to z at 0, and α_{12} at 0. System (3.75) is stabilizable. The first two states are exponentially stable, so the primary concern is the last state. There are several stabilization possibilities.

Discretized Stabilization to an Orbit. The first method is to stabilize the system to an orbit using Theorem 16. To determine the range of feedback gains that will stabilize the system, it suffices to work out the derivation of the feedback law in the proof of Theorem 16. Integrating the linearized system over one time

period, gives the following discrete time system:

$$\begin{aligned} z_1(k+1) &= \exp(-k_1 T) z_1(k), \\ z_2(k+1) &= \exp(-k_2 T) z_2(k), \\ z_3(k+1) &= z_3(k) + \epsilon T \alpha_{12}^1 \alpha_{12}^2. \end{aligned} \tag{3.76}$$

Choosing the parametrization,

$$\alpha_{12}^1 = -K z_3(k), \quad \alpha_{12}^2 = 1,$$

for $K > 0$, the discrete evolution of the last state becomes

$$z_3(k+1) = (1 - \epsilon T K) z_3(k)$$

Exponential stabilization occurs for

$$|1 + \epsilon T K| < 1$$

Solving for K, the range of stabilizing feedback gains is

$$K \in \left(0, \frac{1}{\pi \epsilon^2}\right) \tag{3.77}$$

The most effective gain lies at the midpoint of the range, whereas the boundary coincides with nominal stability; simulations with gains set slightly outside of the boundary points diverge. At this point, there are two options from which to choose: (1) we may feedback instantaneous values of the directly controlled states, or (2) we may use the Floquet mapping to feedback averaged values of the directly controlled state.

(1) Instantaneous Feedback. If Option 1 is chosen, then the oscillatory inputs need to be faster than the natural stabilizing dynamics of the directly controlled states. Otherwise, the closed loop response of the directly controlled states will attenuate the oscillatory signal, and consequently the feedback to the indirectly controlled states. A simulation for this case can be found in Figure 3.1(a). The system parameters are $k_1 = k_2 = 1$, $K = 4$, and $\epsilon = \frac{1}{5}$. The averaged system response does indeed stabilize exponentially, however the actual system stabilizes to an orbit as predicted. For driftless systems this is not a great concern since deactivation of the control inputs halts the system. Stopping at the appropriate moment in the cycle will leave one at the origin.

(2) Averaged Feedback. Option 2 requires the calculation of the Floquet mapping, $P(t)$. To first order, the truncation is

$$P(t) \approx \mathrm{Id} + \left\{ \begin{array}{c} \epsilon \alpha^1 (1 - \cos(t/\epsilon)) \\ \epsilon \alpha^2 \sin(t/\epsilon) \\ -\epsilon \alpha^1 z_2 (1 - \cos(t/\epsilon)) + \epsilon \alpha^2 z_1 \sin(t/\epsilon) \end{array} \right\} + O\left(\epsilon^2\right) \tag{3.78}$$

Since $P(0) = P(kT)$, where T is the period of oscillation, the discretized feedback to the third state may use the instantaneous values $z(kT)$. However, the first two states have oscillatory behavior that must be accounted for. The error values for feedback to the directly controlled subsystem are obtained from the inversion of $P(t)$ on the directly stabilized subsystem,

$$\left\{ \begin{array}{c} z_1 \\ z_2 \end{array} \right\} = \left\{ \begin{array}{c} x_1 - \epsilon \alpha^1 (1 - \cos(t/\epsilon)) \\ x_2 - \epsilon \alpha^2 \sin(t/\epsilon) \end{array} \right\}, \tag{3.79}$$

c.f. Equation (3.46). With this method, not only may one choose any values of ϵ and K that satisfy (3.77), but the feedback values k_1 and k_2 may be arbitrarily chosen. By feeding back averaged values, the controller may give precedence to stabilization of the indirectly controlled state via oscillatory actuation. In the simulation results presented in Figure 3.1(b), the gain choices were, $k_1 = k_2 = 2$, $K = 4$ and $\epsilon = \frac{1}{5}$.

According to Equation (3.77), the smaller ϵ is, the larger the range of choices for K. As ϵ gets smaller, the discretized feedback occurs more often and approaches the continuous system (3.74). In the limit $\epsilon \to 0$, the averaged (3.75) and discretized (3.76) systems coincide. This is the mathematical perspective of the averaging found in the work of Sussmann and Liu [155, 93].

Discretized Stabilization to a Point. For stabilization to a point to occur, the control law must vanish at the desired equilibrium. If periodic discontinuities in the control equations are admissable, then the following control law will exponentially stabilize the system,

$$\alpha^1_{12} = -\text{sign}(Kz_3)\sqrt{|Kz_3|} \quad , \quad \alpha^2_{12} = \sqrt{|Kz_3|}. \tag{3.80}$$

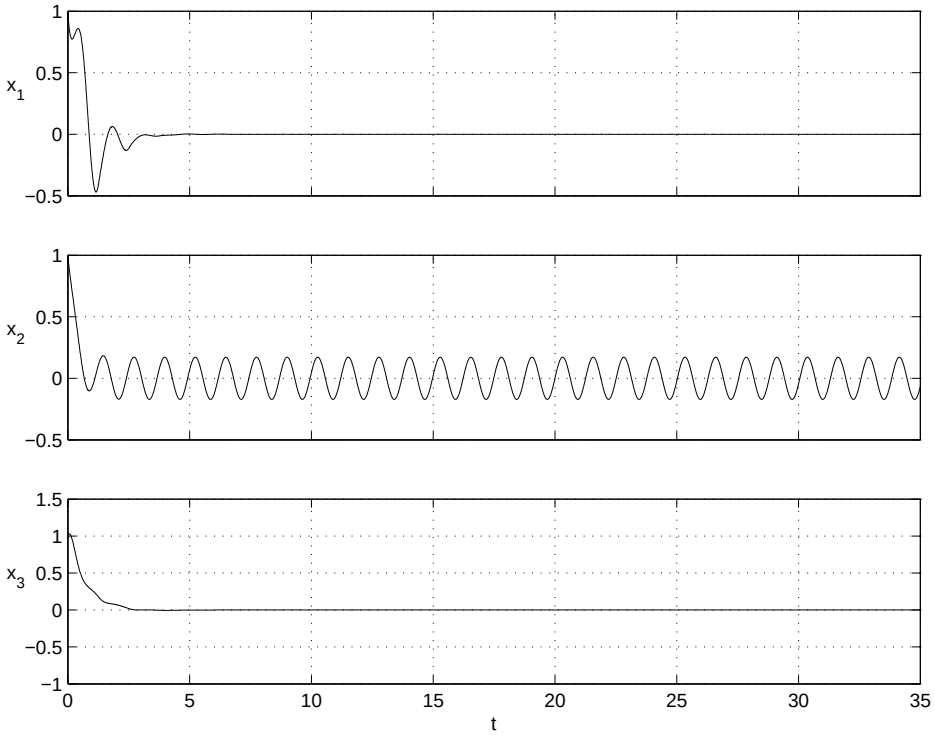
The discontinuity occurs for the second input, $v^2(t)$, at the end of each period, as the feedback gain is adjusted, c.f. Equation (3.73). Again, there are two options concerning the actual implementation, however we will only implement one: feedback of the average of the directly controlled states, Equation (3.79). Simulations can be found in Figure 3.2, using the parameters $k_1 = k_2 = 3$, $K = 4$ and $\epsilon = \frac{1}{5}$.

Continuous Stabilization to a Point. To avoid the discontinuity, one may utilize the inverse of the mapping $P(t)$ to translate the instantaneous state values to the averaged state values for full state-feedback, i.e., $z = P^{-1}(t)(q(t))$, where

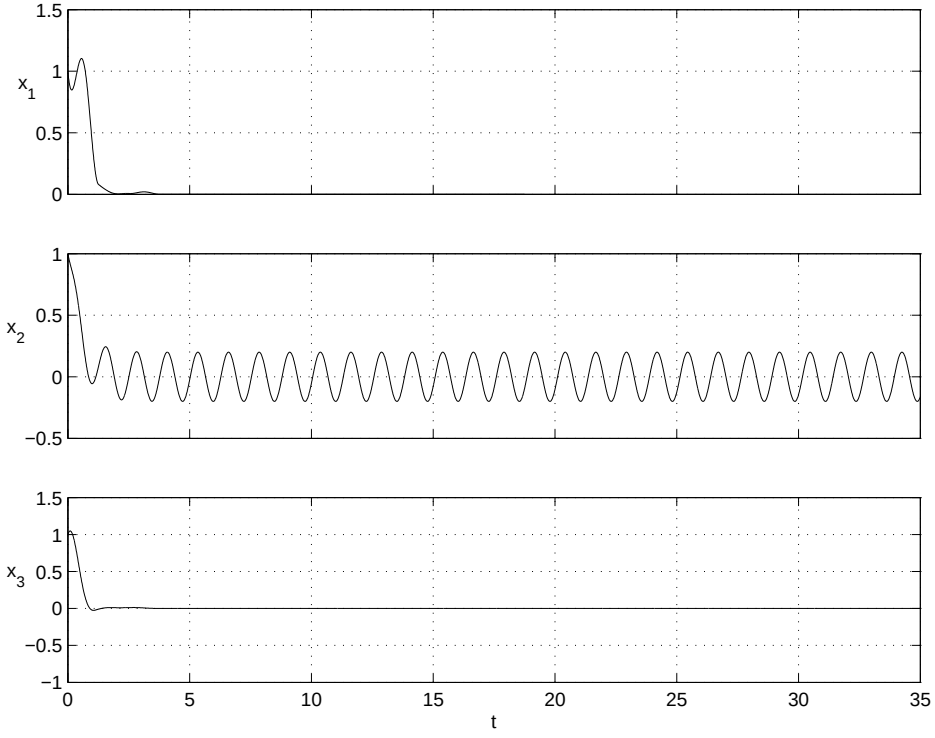
$$P^{-1}(t) \approx \text{Id} - \left\{ \begin{array}{c} \epsilon\alpha^1(1 - \cos(t/\epsilon)) \\ \epsilon\alpha^2 \sin(t/\epsilon) \\ -\epsilon\alpha^1 q_2(1 - \cos(t/\epsilon)) + \epsilon\alpha^2 q_1 \sin(t/\epsilon) \end{array} \right\} - O\left(\epsilon^2\right). \tag{3.81}$$

Continuous feedback is now possible, the discretization step is avoided, and stability of Equation (3.75) is sought. Any choice of the parameters k_1, k_2, K, and ϵ making the system (3.75) stable may be chosen.

Typically, the choice of ϵ is important because ϵ is used to ensure that the series representation may be truncated. The nonholonomic integrator is a nilpotent system, therefore the series is finite in extent. Consequently, there are no restrictions on ϵ and the averaging process is exact. The parameter ϵ will be chosen to coincide with the amount time required to reach a small neighborhood of the origin of the linear averaged system (3.75). Figures 3.1(a), 3.1(b), and 3.2 show that within about 3 seconds the indirectly controlled state is close to the origin compared to the initial condition $q_3 = 1$, so ϵ is chosen to be $\epsilon = \frac{1}{2}$, which corresponds to a period of π seconds. In Figure 3.3(a), can be found the system response to the continuous feedback controller; the gains are $k_1 = k_2 = 7$, and $K = 3$. No matter how high the gains k_1 and k_2 are set, the system will not reach the desired equilibrium until a certain amount of area in the control space has been encircled. This area corresponds to the geometric phase associated with stabilization of the third state. In Figure 3.3(a), it is possible to see that the first state attempts to stabilize to the origin, but is pulled back up so that stabilization of the third state may occur (see also Figures 3.1(b) and 3.2). Figure 3.3(b) is the same controller but with the parameters $k_1 = k_2 = 9$, $K = 1.5$, and $\epsilon = 1$ to demonstrate that ϵ need not be small, nor even less than unity. The complete freedom in ϵ demonstrates a problem with choosing nilpotent approximations. Under the nilpotent approximation, it is not possible to determine what the natural choice of ϵ should be to ensure convergence of the series expansion for the actual system.



(a) Option 1, Instantaneous feedback.



(b) Option 2, Averaged feedback.

Figure 3.1: Discretized orbit stabilization for nonholonomic integrator.

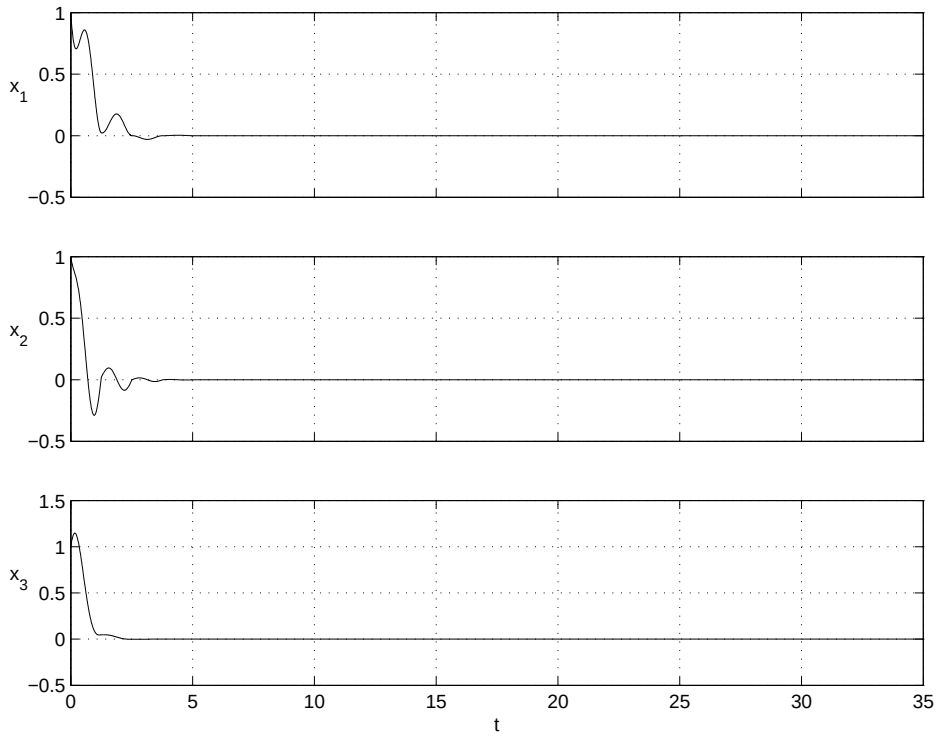


Figure 3.2: Discretized point stabilization for nonholonomic integrator.

Continuous Stabilization via Lyapunov Function The following parametrization choice is provably stable using a Lyapunov function,

$$\alpha_{12}^1 = \frac{-Kz_3}{||z||^{1/2}}, \quad \alpha_{12}^2 = \frac{z_3^2}{||z||^{3/2}}. \tag{3.82}$$

The averaged equations of motion are

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{Bmatrix} z_1 \\ z_2 \\ z_3 \end{Bmatrix} = \begin{Bmatrix} -k_1 z_1 \\ -k_2 z_2 \\ -k_1 z_1 z_2 - k_2 z_1 z_2 \end{Bmatrix} + \epsilon \begin{Bmatrix} 0 \\ 0 \\ -K\frac{z_3^3}{||z||^2} \end{Bmatrix}. \tag{3.83}$$

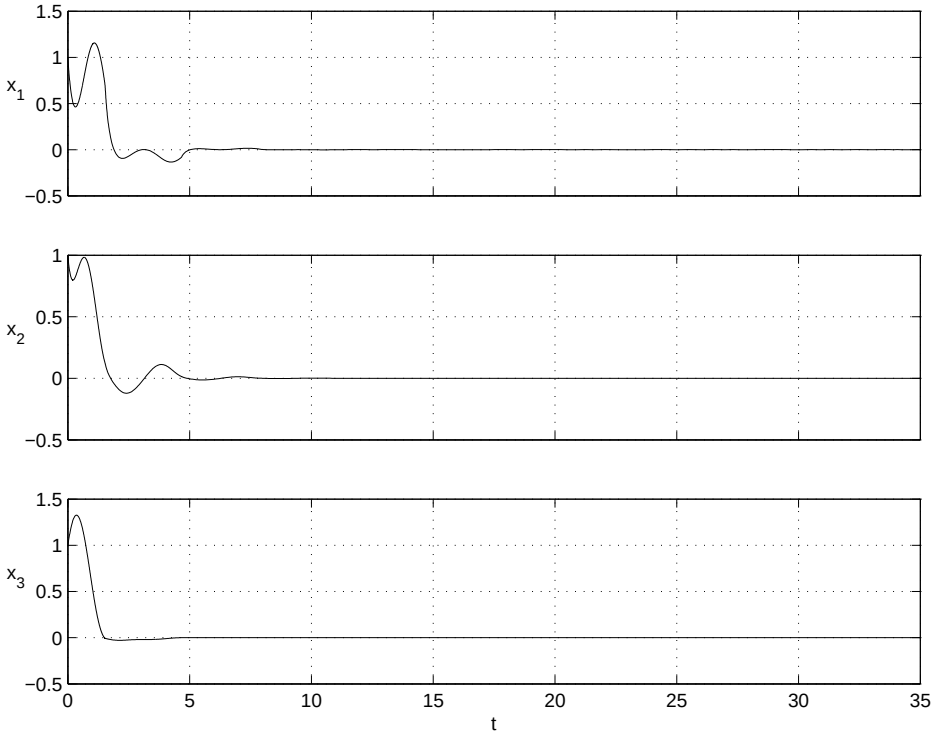
Taking the 0-degree homogeneous truncation,

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{Bmatrix} z_1 \\ z_2 \\ z_3 \end{Bmatrix} = \begin{Bmatrix} -k_1 z_1 \\ -k_2 z_2 \\ 0 \end{Bmatrix} + \epsilon \begin{Bmatrix} 0 \\ 0 \\ -K\frac{z_3^3}{||z||^2} \end{Bmatrix}. \tag{3.84}$$

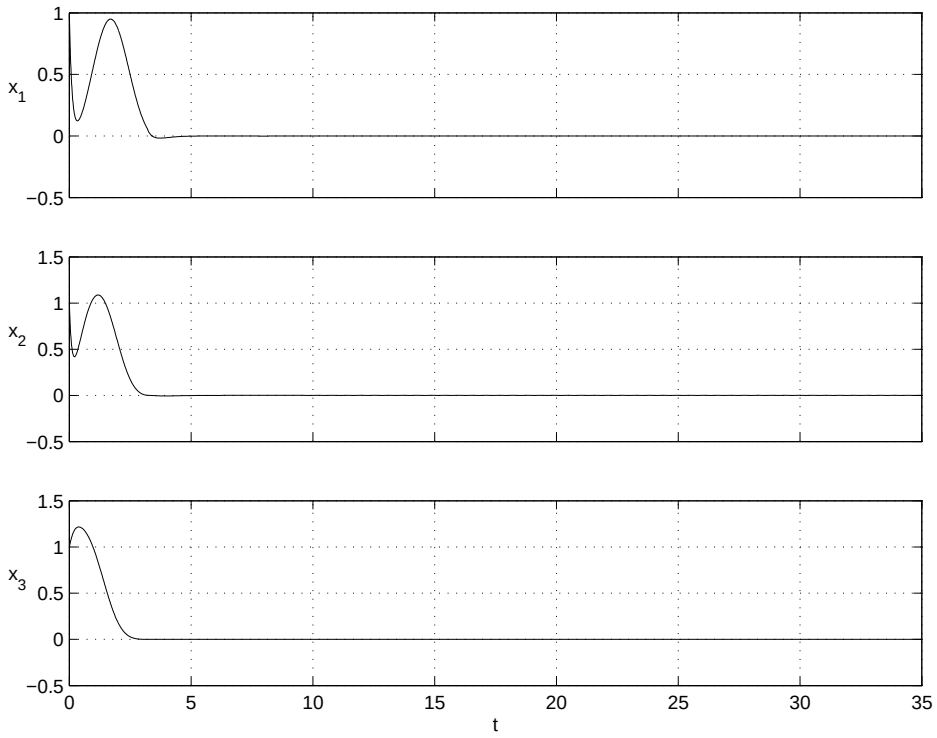
The Lyapunov function

$$V(z) = z_1^2 + z_2^2 + z_3^2 \tag{3.85}$$

trivially demonstrates stability of the system (3.84). Figure 3.4 shows the stabilized system response for $k_1 = k_2 = 3$, $K = 1$, and $\epsilon = 1$.



(a) $\epsilon = \frac{1}{2}$



(b) $\epsilon = 1$

Figure 3.3: Continuous point stabilization of nonholonomic integrator.

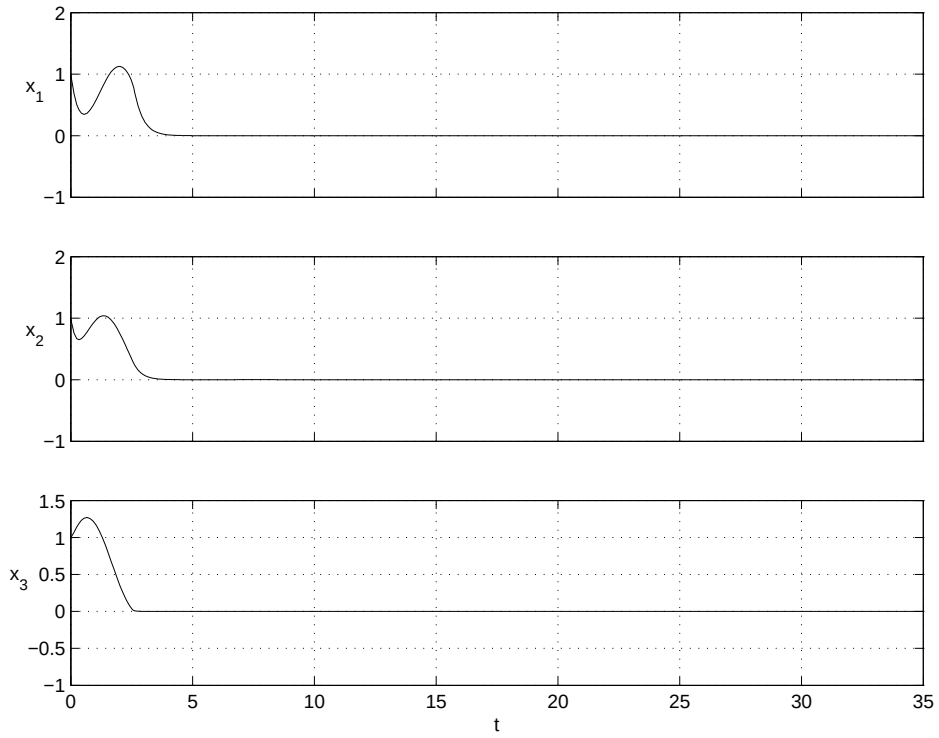


Figure 3.4: Continuous 2-norm stabilization of nonholonomic integrator.

Continuous Homogeneous Stabilization The following parametrization choice leads to the homogeneous controller from M'Closkey and Morin [107],

$$\alpha_{12}^1 = \frac{-Kz_3}{\rho(z)}, \quad \alpha_{12}^2 = \frac{z_3^2}{\rho^3(z)}, \tag{3.86}$$

where

$$\rho(z) = \sqrt[4]{z_1^4 + z_2^4 + z_3^2}. \tag{3.87}$$

The averaged equations of motion are

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{Bmatrix} z_1 \\ z_2 \\ z_3 \end{Bmatrix} = \begin{Bmatrix} -k_1 z_1 \\ -k_2 z_2 \\ -k_1 z_1 z_2 - k_2 z_1 z_2 \end{Bmatrix} + \epsilon \begin{Bmatrix} 0 \\ 0 \\ -K\frac{z_3^3}{\rho^4(z)} \end{Bmatrix}. \tag{3.88}$$

Taking the 0-degree homogeneous truncation,

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{Bmatrix} z_1 \\ z_2 \\ z_3 \end{Bmatrix} = \begin{Bmatrix} -k_1 z_1 \\ -k_2 z_2 \\ 0 \end{Bmatrix} + \epsilon \begin{Bmatrix} 0 \\ 0 \\ -K\frac{z_3^3}{\rho^4(z)} \end{Bmatrix}. \tag{3.89}$$

The Lyapunov function

$$V(z) = z_1^4 + z_2^4 + z_3^2 = \rho^4(z) \tag{3.90}$$

trivially demonstrates stability of the averaged system.

Although the controller is designed using ideas from M'Closkey [108], the resulting controller is different from the homogeneous controllers in [108] because the average values are used as feedback instead of the instantaneous values of the states, affording the feedback strategy more freedom regarding the feedback gains and oscillation frequency. Figure 3.5(a) demonstrate stabilization with the parameters, $k_1 = k_2 = 3$, $K = 1$, and $\epsilon = 1$. Compare this to Figure 3.5(b), which is the ρ-homogeneous stabilizer from Example 4.17 of M'Closkey [108]; the parameters are $k_1 = k_2 = 1$, $K = 1$, and $\epsilon = 1$.

The gains are different for the ρ-homogeneous stabilizer from [108] compared to the ρ-homogeneous stabilizer using the generalized averaging theory, because any adjustment of the gains degrades the system performance. For homogeneous controllers, there is no way to systematically understand the effect of different feedback gains because the nonlinear oscillatory response is not accounted for in the closed loop system analysis. How the oscillatory response and the stabilizing gains interact is not well understood.

The Averaged Coefficients. The averaged coefficients computed in the nonholonomic integrator example used indefinite integrals. If definite integrals were used instead, the averaged coefficients would be

$$\overline{\mathrm{V}^{(a)}_{(1)}(t)} = \begin{bmatrix} \alpha^1_{12} \\ 0 \end{bmatrix} \quad , \quad \overline{\mathrm{V}^{(a,b)}_{(1,0)}(t)} = \begin{bmatrix} 0 & -\frac{1}{2}\alpha_{12} \\ \frac{1}{2}\alpha_{12} & 0 \end{bmatrix},$$

where $\alpha_{12} = \alpha^1_{12}\alpha^2_{12}$. The averaged equations of motion are

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{Bmatrix} z_1 \\ z_2 \\ z_3 \end{Bmatrix} = \begin{Bmatrix} -k_1 z_1 + \alpha^1_{12} \\ -k_2 z_2 \\ (-k_1 z_1 + \alpha^1_{12}) z_2 - k_2 z_1 z_2 \end{Bmatrix} + \epsilon\alpha_{12}\begin{Bmatrix} 0 \\ 0 \\ 1 \end{Bmatrix}. \tag{3.91}$$

The α^1_{12} term may be an obstruction to stability, or to the construction and demonstration of a stabilizing controller. Although the analysis may still work out as desired, we would like to avoid having to do analysis with the α^1_{12} term. If the transformation of state,

$$\Theta(z) = (z_1 + \frac{\alpha^1_{12}}{k_1}, z_2, z_3), \tag{3.92}$$

is defined, then a new averaged vector field may be recovered,

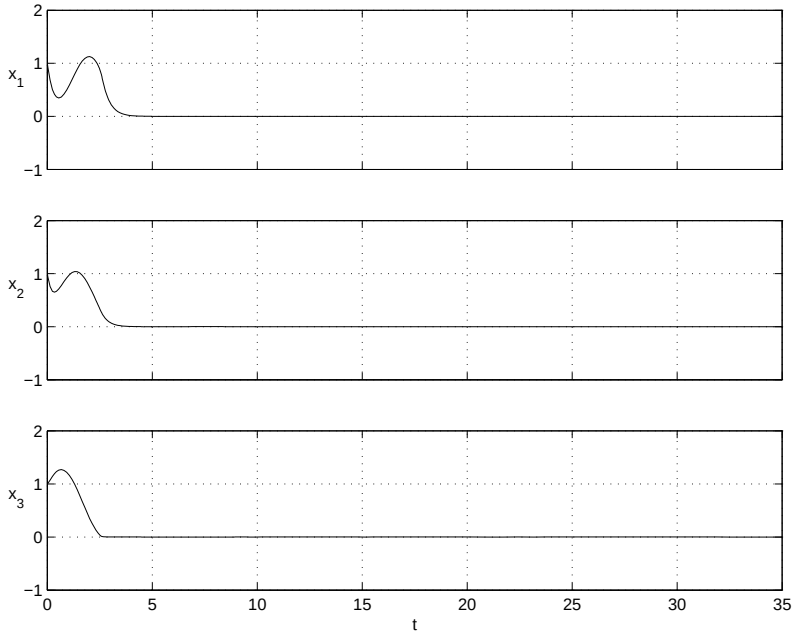
$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{Bmatrix} \tilde{z}_1 \\ \tilde{z}_2 \\ \tilde{z}_3 \end{Bmatrix}\begin{Bmatrix} -k_1 \tilde{z}_1 \\ -k_2 \tilde{z}_2 \\ -k_1 \tilde{z}_1 \tilde{z}_2 - k_2 \tilde{z}_1 \tilde{z}_2 \end{Bmatrix} + \epsilon\alpha_{12}\begin{Bmatrix} 0 \\ 0 \\ 1 \end{Bmatrix}. \tag{3.93}$$

The new system is the same average that would have been obtained had indefinite integrals been used, as in Equation(3.75).

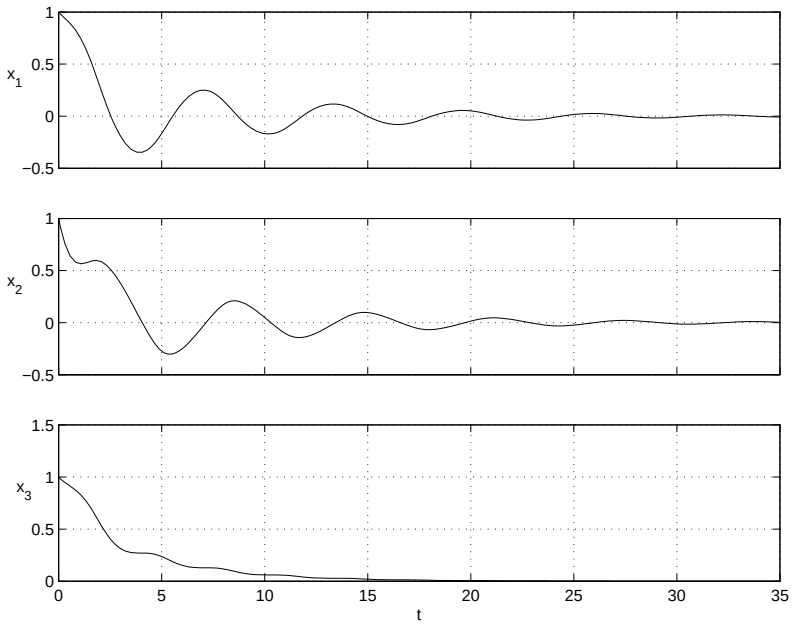
In Section 3.1.1, it was argued that the definite integrals corresponded to integration or inclusion of the initial conditions in the governing equations. Furthermore, this inclusion implied that the averaged expansions with definite integrals were capable of accurately predicting the transient response of the system. The transformation of state Θ gives a clue as to what the transient response will be. For the initial conditions $z_0 = (1, 1, 1)$, the transformation of state Θ is

$$\Theta(z_0) = (1 + \frac{\sqrt{K}}{k_1}, 1, 1), \tag{3.94}$$

impying that the first state will behave as though the initial condition were larger. Analysis of Figure 3.1(b) shows that the initial trajectory of the first state does have a higher average as can be seen by the upwards



(a) ρ-homogeneous controller using averaging methods.



(b) ρ-homogeneous controller from [108].

Figure 3.5: Continuous ρ-homogeneous stabilization of nonholonomic integrator.

bump for early times. The later figures (e.g., Figures 3.2 and 3.3) have larger gains, so this initial bump dies out faster.

A Comment on the Controllers. The purpose behind the averaging methods is to obtain a linearly (or homogeneous order 0) stabilized closed-loop averaged autonomous vector field for local exponential stability. Although the parametrizations differed for the various control strategies, the net averaged closed-cloop system for all of the cases were similar. Consequently, all of the strategies outlined above have the qualitatively similar closed-loop performance in the average, with the exception of the ρ-homogeneous controller from [108]. The parametrization choice does affect how the actual time-varying system will evolve, which explains why some controllers stabilized to an orbit around the equilibrium, whereas others stabilized to the equilibrium.

3.3.2 Hilare Robot

The Hilare robot is a control example that can be transformed into the nonholonomic integrator form just studied. The purpose of this example is to demonstrate that nonlinear transformations to canonical nonlinear control forms may be avoided. In general such transformations are undesirable as they may dilate the natural length scales of the control system [108], involve complicated coordinate transformations, or may have singularities [157]. One consequence may be that a feasible controller for the canonical nonlinear control forms could result in an infeasible controller for the actual system, e.g., the nonlinear transformation to canonical form results in a control law that significantly magnifies a control input, thereby saturating the actual control signal. For homogeneous controllers such a stretching of the manifold space implies a warped homogeneous norm, and hunting may occur [108].

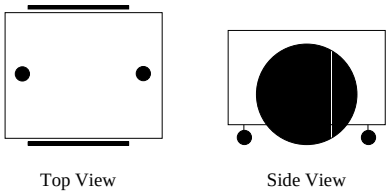


Figure 3.6: Hilare robot.

The Hilare robot is a mechanical system with symmetry, functionally equivalent to a wheelchair, Figure 3.6. The wheels provide sufficient constraints to define a principal connection, c.f. Kelly and Murray [67]. The equations of motion in the principal bundle, $Q \cong M \times G$, are

$$\dot{q} = Y_a^h(q)u^a, \quad q \in Q, \tag{3.95}$$

where $\cdot^h : \mathcal{X}(M) \to \mathcal{X}(Q)$ is the horizontal lift with respect to the principal connection form, denoted $\mathcal{A} : TM \to \mathfrak{g}$. The control input vector fields, $Y_a \in \mathcal{X}(M)$ are

$$Y_1(r) = \frac{\partial}{\partial r^1}, \quad Y_2(r) = \frac{\partial}{\partial r^2}. \tag{3.96}$$

In a local trivialization, $q = (r, g)$, and the horizontal lift takes the form

$$Y_a^h(r,g) = \left\{ \begin{array}{c} Y_a(r) \\ -g\mathcal{A}_{loc}(r)Y_a(r) \end{array} \right\}. \tag{3.97}$$

The local form of the principal connection for the Hilare robot is

$$\mathcal{A}_{loc}(r) = \begin{bmatrix} -\frac{1}{2}\rho \mathbf{d}r^1 - \frac{1}{2}\rho \mathbf{d}r^2 \\ 0 \\ \frac{\rho}{2w}\mathbf{d}r^1 - \frac{\rho}{2w}\mathbf{d}r^2 \end{bmatrix} \tag{3.98}$$

For the robot, the main concern is control of the Lie group, the rotational orientations of the wheels are inconsequential. In the language of Section 3.2, weak fiber controllability is sought. The local connection

form (3.98) spans two unique directions in the three-dimensional Lie algebra $\mathfrak{se}(2)$. The curvature of the connection form is

$$\mathcal{B}_{loc}(r) = \begin{bmatrix} 0 \\ -\frac{\rho^2}{2w}\mathbf{d}r^1 \wedge \mathbf{d}r^2 \\ 0 \end{bmatrix}. \tag{3.99}$$

Together, the local connection and curvature forms span the Lie algebra, therefore the system is (small-time) locally, weakly fiber controllable. Since the curvature form corresponds to the Lie bracket of horizontally lifted control inputs, we can control the system using the controller from the nonholonomic integrator example. The oscillatory controls and α-parametrization will be such that the average system results in a fully controllable system evolving on $SE(2)$. The technique found in Bullo and Murray [26] for control on Lie groups will be used. Define an error function on $SE(2)$,

$$g_{err} = g^{-1}g_{des}, \tag{3.100}$$

where g_{des} is the desired $SE(2)$ position. The controls are

$$u^a(t) = f^a(g; g_{des}) + v^a(t/\epsilon), \tag{3.101}$$

where

$$\begin{aligned} f^1(g; g_{des}) &= k_1 g^1_{err} + k_3 g^3_{err}, \\ f^2(g; g_{des}) &= k_1 g^1_{err} - k_3 g^3_{err}, \end{aligned} \tag{3.102}$$

and the oscillatory terms are obtained from Equations (3.80),

$$v^1(t) = \text{sign}(Kg^2_{err})\sqrt{|Kg^2_{err}|}\sin(t) \quad , \quad v^2(t) = \sqrt{|Kg^2_{err}|}\cos(t). \tag{3.103}$$

The controller construction follows that of the nonholonomic integrator, however we use the matrix Lie group exponential for the flow of $SE(2)$, instead of the $\mathbb{R}^n$ exponential. The averaged equations of motion in the Lie group are

$$\dot{h} = -h\left(X_S + \frac{1}{2}\epsilon\overline{\mathrm{V}^{(a,b)}_{(1,0)}(t)}\mathcal{B}_{ab}\right). \tag{3.104}$$

The truncated Floquet mapping is

$$\text{Trunc}_1\left(P(t/\epsilon)\right) \approx \text{Id} + \hat{\xi}_{avg} + O\left(\epsilon^2\right), \tag{3.105}$$

where

$$\xi_{avg} = \epsilon\alpha^1_{12}(1 - \cos(t/\epsilon))\mathcal{A}_1 + \epsilon\alpha^2_{12}\sin(t/\epsilon)\mathcal{A}_2. \tag{3.106}$$

Therefore, the truncation of $P(t)$ acts in the following manner,

$$\text{Trunc}_1\left(P(t)\right)(g) = g + g\hat{\xi}_{avg} + O\left(\epsilon^2\right). \tag{3.107}$$

The controller (3.101) is shown stabilizing to the origin in Figure 3.7. The simulation parameters are $k_1 = 6$, $K = 4\sqrt{5}$, $k_3 = \frac{3}{2}$, and $\epsilon = \frac{2}{3}$. Additional examples of point stabilization can be found in Figure 3.8. A saturation function was included to prevent the wheel velocities from being too high. The saturation choice of 4.5 radians per second was obtained from M'Closkey [108]. The dimensions of the Hilare robot are $\rho = 0.05$, and $w = 0.2$. A benefit to the averaging theory approach over other methods is the existence of the feedback control parameters. By placing the system within a linear state-feedback framework, it is possible to examine different choices for the feedbacks gains. An additional parameter the control engineer

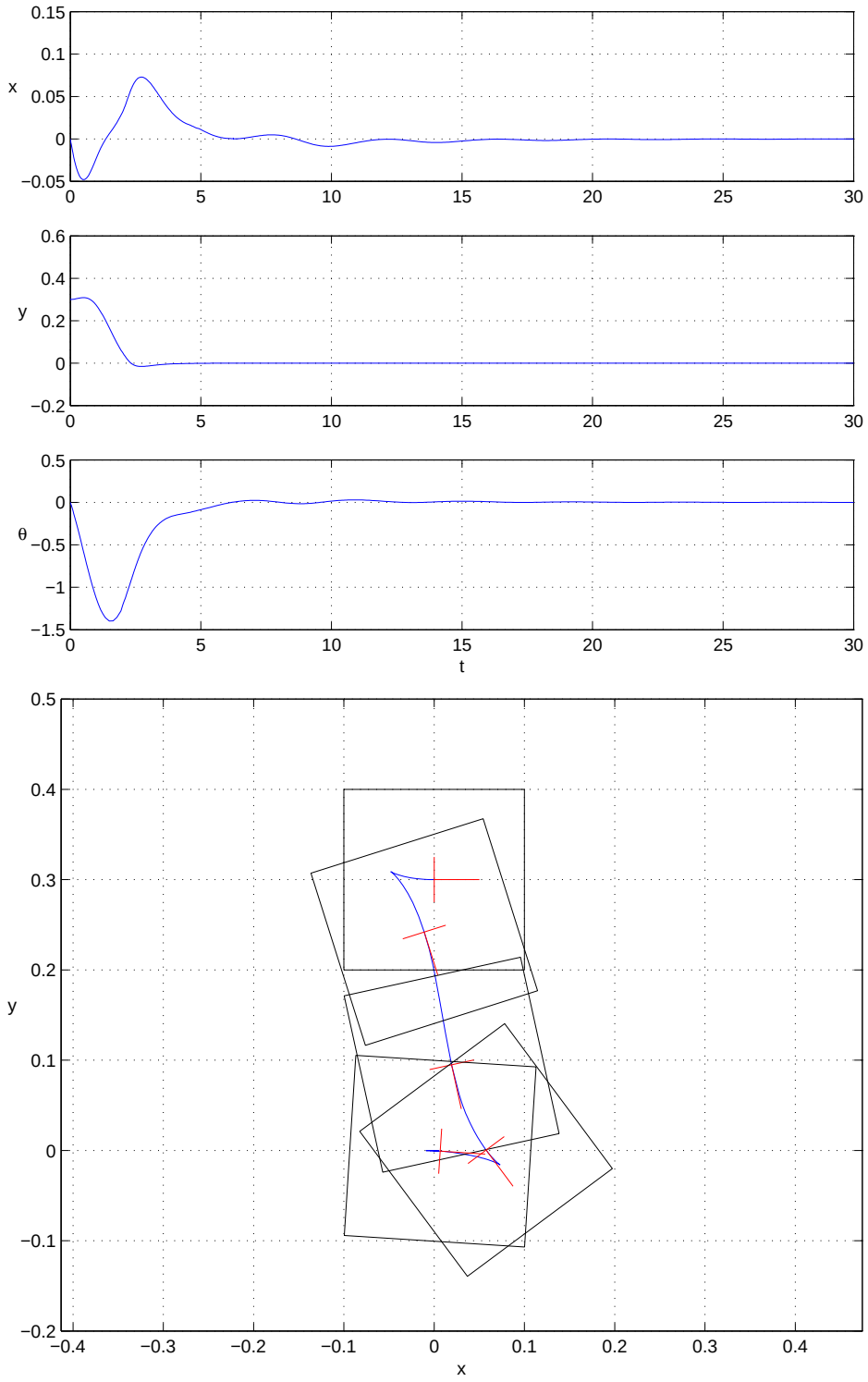
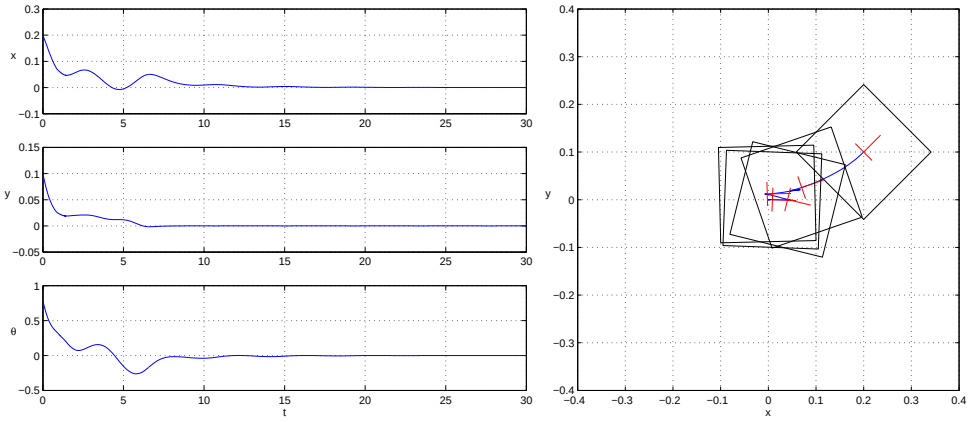
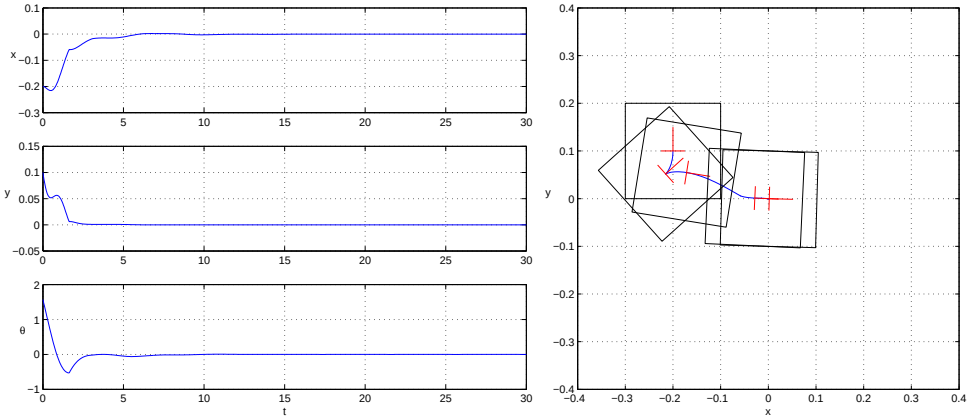


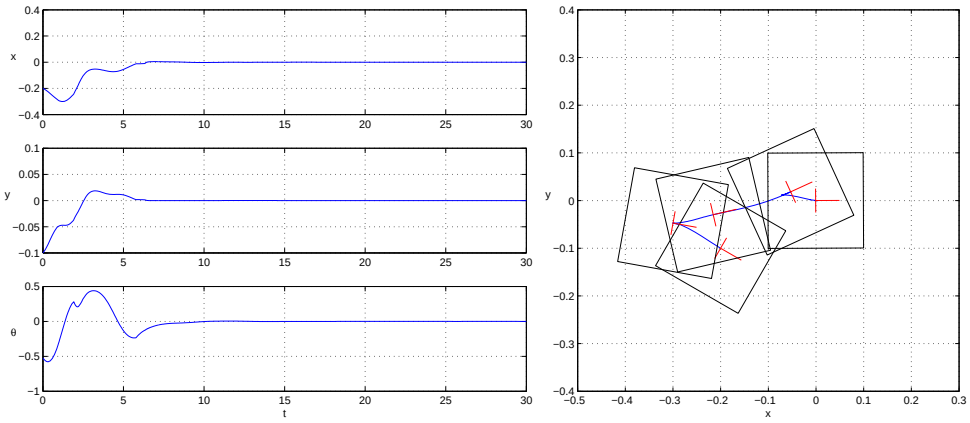
Figure 3.7: Hilare robot point stabilization. Snapshots at $t = 0,\ 1.1,\ 1.8,$ and 5.3 seconds.



(a) Snapshots at $t = 0,\ 0.9,\ 4.3,\ 6.1,$ and 9.5 seconds.



(b) Snapshots at $t = 0,\ 0.4,\ 1.0,\ 2.7,$ and 6.8 seconds.



(c) Snapshots at $t = 0,\ 1.1,\ 2.0,\ 3.4,$ and 11.3 seconds.

Figure 3.8: Additional examples of point stabilization for the Hilare robot.

may modify is the oscillation period. Control of underactuated nonlinear driftless affine control systems boils down to a set of feedback gains plus a set of oscillatory parameters. Such a simplification of the control strategy has obvious benefits.

3.3.3 Kinematic Car

The kinematic car is a driftless model of an automobile. The front two wheels freely spin but may be used for steering, and the rear wheels may be actuated to move the car forward. The kinematic car is a mechanical system with group symmetries. The wheels provide sufficient constraints to define a principal connection, c.f. Kelly and Murray [67]. The principal bundle describing the configuration manifold is $Q \cong M \times G$, where $M \subset \mathbb{T}^2$ describes the steering and driving angles, and $G = SE(2)$ describes the planar position and orientation of the vehicle. In coordinates, the base space is specified by $r = (\psi, \phi)$, where ψ is the wheel rotation angle, and ϕ is the steering angle. The group variables specify planar position and orientation of the center of the rear axle, $g = (x, y, \theta)$. The equations of motion motion for the driftless automobile are

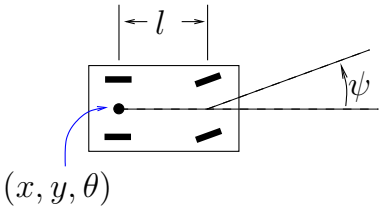


Figure 3.9: Kinematic car.

$$\dot{q} = Y_a^h u^a, \tag{3.108}$$

where $\cdot^h : \mathcal{X}(M) \to \mathcal{X}(Q)$ is the horizontal lift with respect to the principal connection form, $\mathcal{A} : TM \to \mathfrak{g}$. The control input vector fields are

$$Y_1(r) = \frac{\partial}{\partial \psi}, \quad Y_2(r) = \frac{\partial}{\partial \phi}.$$

In a local trivialization, $q = (r, g)$, the horizontal lifts of equation (3.108) are

$$Y_a^h(r, g) = \left\{ \begin{array}{c} Y_a(r) \\ -g\mathcal{A}_{loc}(r) \end{array} \right\}, \tag{3.109}$$

where the local principal connection form for the kinematic car is

$$\mathcal{A}_{loc}(r) = \begin{bmatrix} -\rho\, \mathbf{d}\psi \\ 0 \\ -(\rho/l)\tan(\phi)\, \mathbf{d}\psi \end{bmatrix}. \tag{3.110}$$

The curvature of the local connection form is

$$\mathcal{B}_{loc}(r) = \left\{ \begin{array}{c} 0 \\ 0 \ -(\rho/l)\sec^2(\phi)\, \mathbf{d}\phi \wedge \mathbf{d}\psi \end{array} \right\}, \tag{3.111}$$

and the covariant derivative of the curvature is

$$\nabla\mathcal{B}(r) = \left\{ \begin{array}{c} 0 \\ (\rho^2/l)\sec^2(\phi)\mathbf{d}\psi \wedge \mathbf{d}\phi \wedge \mathbf{d}\psi \\ 2(\rho/l)\sec^2(\phi)\tan(\phi)\mathbf{d}\phi \wedge \mathbf{d}\phi \wedge \mathbf{d}\psi \end{array} \right\}. \tag{3.112}$$

Although it is important that the steering angle remain in a neighborhood of the straight forward configuration ($\phi = 0$), the rotation angle is not restricted in any way. Since the rotation angle, ψ, is not restricted, the local connection can be used for control, and weak fiber controllability will be sufficient [67]. The important

covariant derivative contributions for control are given by the following Jacobi-Lie brackets,

$$\left[Y_1^h, Y_2^h\right], \quad \text{and} \quad \left[Y_1^h, \left[Y_2^h, Y_1^h\right]\right], \tag{3.113}$$

whose corresponding averaged coefficients are

$$\overline{\mathrm{V}_{(1,0)}^{(1,2)}(t)}, \quad \text{and} \quad \overline{\mathrm{V}_{(1,1,0)}^{(1,2,1)}(t)}. \tag{3.114}$$

To define the α-parametrization, an error function is needed. The group error and its logarithm are

$$g_{err} = g^{-1} g_{des}, \quad \text{and} \quad \xi_{err} = \log(g_{err}). \tag{3.115}$$

According to Lemmas 2 and 3, the oscillatory inputs should be,

$$\begin{aligned} v^1(t) &= \alpha_{12}^1 \sin(\omega_{12} t) + \alpha_{121}^1 \sin(2\omega_{121} t), \text{ and} \\ v^2(t) &= \alpha_{12}^2 \cos(\omega_{12} t) + \alpha_{121}^2 \cos(\omega_{121} t). \end{aligned} \tag{3.116}$$

The α-parametrization is

$$\begin{aligned} \alpha_{12}^1 &= 2\sqrt{|k_3 \xi_{err}^3|} & \alpha_{12}^2 &= \frac{1}{2}\mathrm{sign}(k_3 \xi_{err}^3)\sqrt{|k_1 \xi_{err}^1|} \\ \alpha_{121}^1 &= 2\sqrt[3]{|k_2 \xi_{err}^2|} & \alpha_{121}^2 &= \frac{1}{4}\mathrm{sign}(k_2 \xi_{err}^2)\sqrt[3]{|k_2 \xi_{err}^2|}. \end{aligned} \tag{3.117}$$

The constant factors in the oscillatory functions cancel out in the averaged coefficients. They were placed into the oscillatory functions to reshape the controls because the steering wheel cannot turn more than 45 degrees. By driving forward/backwards more, it is possible to compensate for the reduced steering angle. The first state can be stabilized using state-feedback,

$$f^1(g) = -k_1 \xi_{err}^1.$$

Because the first state is continuously stabilized, the inverse Floquet mapping will be needed. To first order, it is

$$\mathrm{Trunc}_1\left(P(t)^{-1}\right) = g - g\mathrm{V}_{(1)}^{(a)}(t/\epsilon)\mathcal{A}_a + O\left(\epsilon^2\right).$$

Snapshots of a simulation can be found in Figure 3.10. The grey line is the trajectory of the vehicle over the previous period of oscillatory actuation. By the end of the third cycle, the car has pretty much stabilized. The system parameters were $\rho = 3/20$, $l = 6/10$, $\epsilon = 1$, $\omega_{12} = 1$, $\omega_{121} = 2$, $k_1 = 2$, $k_2 = 7/(10\rho)$, and $k_3 = 1/(4\rho)$. The error plot can be found in Figure 3.12. The dots represent the error at the end of a period of actuation. The stabilization is discretized feedback to the indirectly controlled states. The stabilizing strategy of the kinematic car was designed in less than a day. With more time, an improved controller can be developed. One improvement would be to employ a continuous feedback strategy. The trajectories would be much smoother and less jagged; more like the Hilare robot example.

The scale is chosen so that 1 unit of distance corresponds to about 10 feet. The kinematic car parameters were set to be on the same order as a standard vehicle. Additional simulation snapshots can be found in Figure 3.11, for which it is possible to see the effects of the oscillatory actuation.

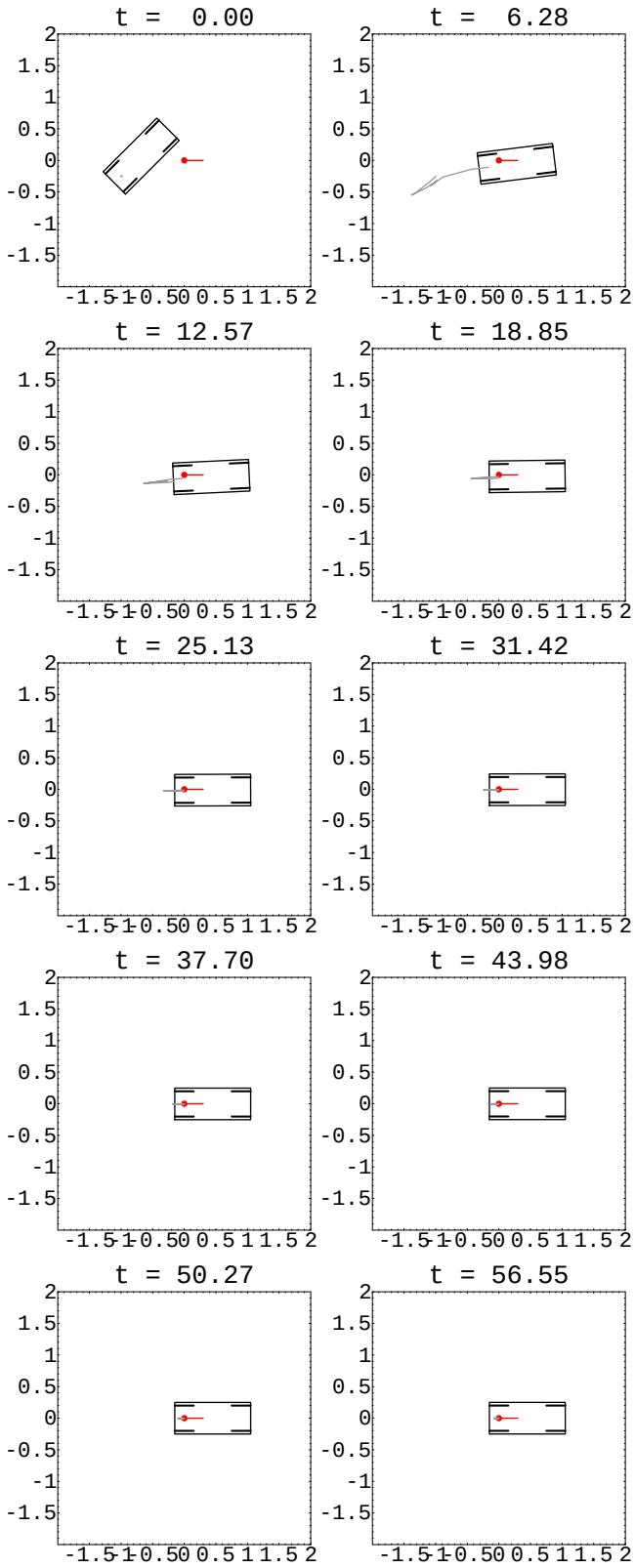
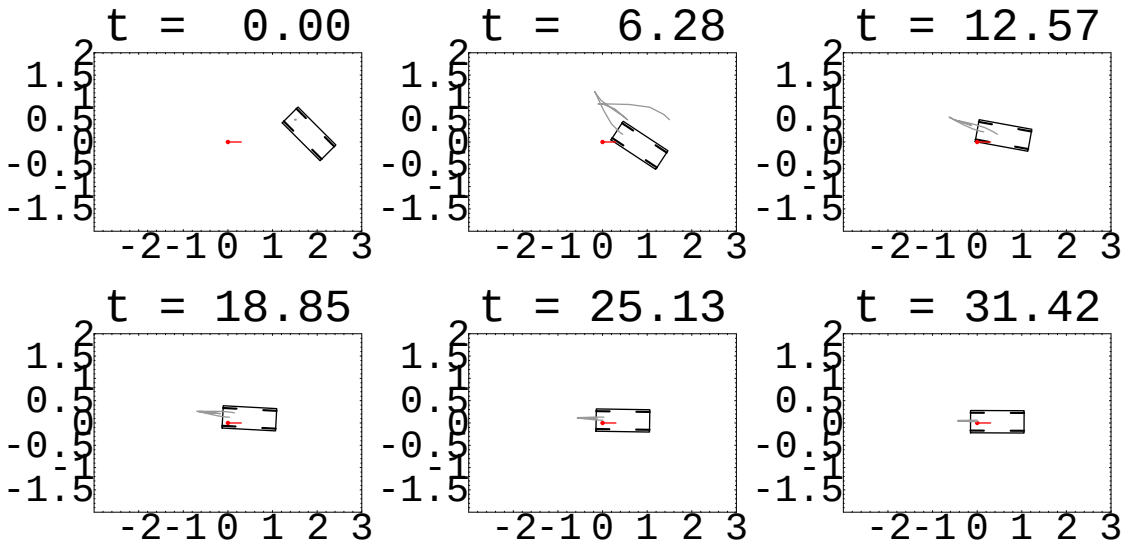
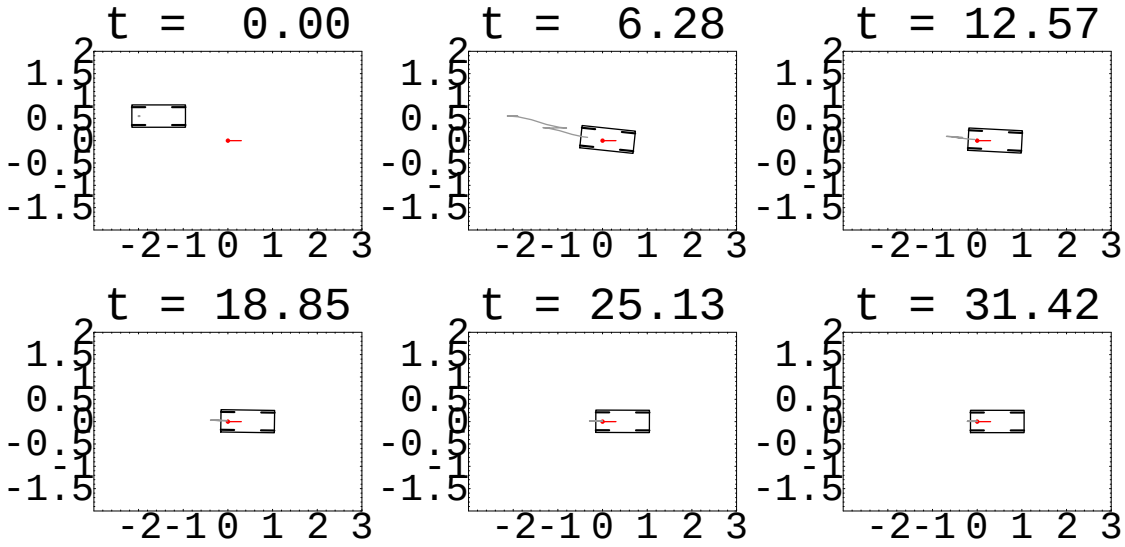


Figure 3.10: Kinematic car point stabilization snapshots.



(a) $g_0 = (1.5, 0.5, -\pi/4)$



(b) $g_0 = (-2, 0.5, 0)$

Figure 3.11: Additional examples of point stabilization for the kinematic car.

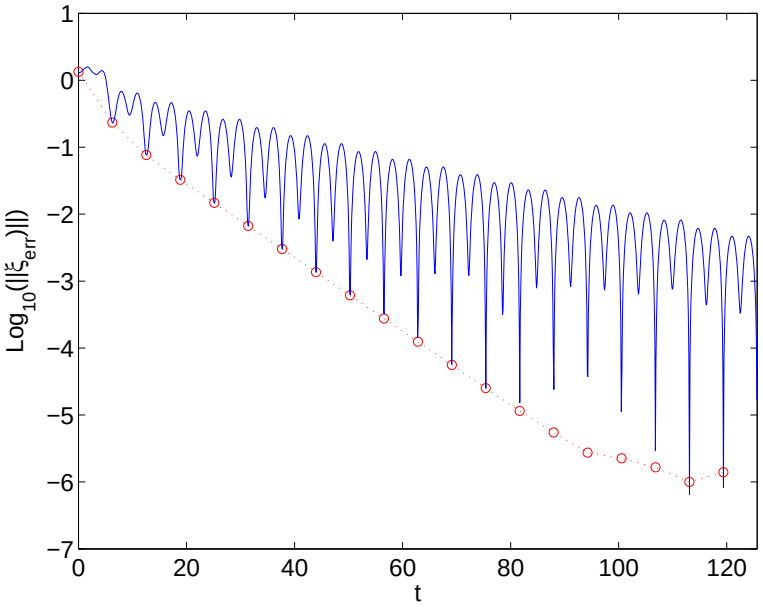


Figure 3.12: Kinematic car stabilization error, $\log_{10}(||\xi_{err}||)$.

3.4 Conclusion

The generalized averaging theory of Chapter 2 was applied to underactuated driftless affine control systems, resulting in exponentially stabilizing control strategies. Since the strategies is constructive, any system satisfying the required conditions can be stabilized. The algorithm itself is decomposed into distinct controller synthesis steps: (1) averaging, (2) sinusoidal actuation design, (3) and feedback parametrization. The only knowledge utilized is the Lie algebra structure of the underactuated driftless control system, thereby avoiding the use of nonlinear transformations, Lyapunov functions, and pre-existing controllers. Although no special knowledge of the system is required, any additional information or structure can be used to facilitate and optimize the process. Most importantly, the strategy results in tunable control parameters reminiscent of the control parameters from linear control theory. The system nonlinearities enter into the control law as additional oscillatory control parameters.

The averaging method takes a time-varying nonlinear driftless control system and determines an autonomous approximation to the control system. It was shown that Lyapunov techniques could be applied to the averaged autonomous system to determine stability. By using the averaged autonomous system, the use of time-varying Lyapunov analysis was avoided. Recently Peuteman and Aeyels have investigated the application of averaging methods to determine stability of nonlinear systems via Lyapunov techniques [133, 134, 135]. The generalized averaging theory can be used not only to recover many of their results, but also to extend them.

The control strategy based on averaging has many other attractive features. Since integration of time dependent portions of the inputs smoothes out discontinuities, the approach can be used for legged locomotion control design. The examples chapter, Chapter 6, will study this idea for a simple bipedal robot model.

Although there are unique problems inherent to systems with drift, the general strategy set forth in this chapter still holds in this case. Prior to developing the ideas for systems with drift, the notion of small-time local controllability must be revisited. Small-time local controllability for systems with drift is more complicated, as not all elements of the involutive closure $\mathcal{C}^\infty$ will be beneficial from a control-theoretic perspective.

3.A Computing the Averaged Expansions

First-order. The first order averaged equations of motion are just the time-average of the original equations of motion,

$$\frac{\mathrm{d}z}{\mathrm{d}\tau} = \epsilon X_S + \epsilon\overline{\mathrm{V}_{(0)}^{(a)}(\tau)}Y_a.$$

The zero order truncation is trivially

$$\mathrm{Trunc}_0\left(P(\tau)\right) = \mathrm{Id}.$$

An improved version involves approximation to an additional order of ϵ,

$$\mathrm{Trunc}_1\left(P(\tau)\right) = \mathrm{Id} + \epsilon\int_0^\tau \left(X_S + \mathrm{V}_{(0)}^{(a)}(\tau)Y_a - X_S - \overline{\mathrm{V}_{(0)}^{(a)}(\tau)}\right)\,\mathrm{d}\tau \;= \mathrm{Id} + \epsilon\int_0^\tau \widetilde{\mathrm{V}}_{(0)}^{(a)}(\tau)\,\mathrm{d}\tau\; Y_a.$$

Second-order. The second-order average has the additional term,

$$\begin{aligned}
\Lambda^{(2)} &= \frac{1}{2}\overline{\left[\int^\tau X\,\mathrm{d}\sigma\,, X\right]} \\
&= \frac{1}{2}\overline{\left[X_S t + \mathrm{V}_{(1)}^{(a)}(\sigma)Y_a, X_S + \mathrm{V}_{(0)}^{(b)}(\tau)Y_b\right]} \\
&= \frac{1}{2}\overline{\left(\left[X_S t, \mathrm{V}_{(0)}^{(b)}(\tau)\right] + \left[\mathrm{V}_{(1)}^{(a)}(\tau)Y_a, X_S\right] + \left[\mathrm{V}_{(1)}^{(a)}(\tau)Y_a, \mathrm{V}_{(0)}^{(b)}(\tau)Y_b\right]\right)} \\
&= \frac{1}{2}\left(\overline{\mathrm{V}_{(0)}^{(a)}(\tau)\tau}\,[X_S, Y_a] + \overline{\mathrm{V}_{(1)}^{(a)}(\tau)}\,[Y_a, X_S] + \overline{\mathrm{V}_{(1,0)}^{(a,b)}(\tau)}\,[Y_a, Y_b]\right) \\
&= \frac{1}{2}\left(2\overline{\mathrm{V}_{(1)}^{(a)}(\tau}\,[Y_a, X_S] + \overline{\mathrm{V}_{(1,0)}^{(a,b)}(\tau)}\,[Y_a, Y_b]\right) \\
&= \overline{\mathrm{V}_{(1)}^{(a)}(\tau)}\,[Y_a, X_S] + \frac{1}{2}\overline{\mathrm{V}_{(1,0)}^{(a,b)}(\tau)}\,[Y_a, Y_b]\,,
\end{aligned}$$

under Assumption 2. Incorporating the second-order term into the averaged equations of motions leads to

$$\dot{z} = X_S(z) + \epsilon\overline{\mathrm{V}_{(1)}^{(a)}(t)}\,[Y_a(z), X_S(z)] + \frac{1}{2}\epsilon\overline{\mathrm{V}_{(1,0)}^{(a,b)}(t)}\,[Y_a(z), Y_b(z)]\,.$$

The truncated Floquet mapping is

$$\mathrm{Trunc}_1\left(P(\tau)\right) = \mathrm{Id} + \epsilon\int_0^\tau \left(X_S + \mathrm{V}_{(0)}^{(a)}(\tau)Y_a - X_S - \overline{\mathrm{V}_{(0)}^{(a)}(\tau)}\right)\,\mathrm{d}\tau \;= \mathrm{Id} + \epsilon\int_0^\tau \mathrm{V}_{(0)}^{(a)}(\tau)\,\mathrm{d}\tau\; Y_a, \tag{3.118}$$

since Assumption 2 means that $\overline{\mathrm{V}_{(0)}^{(a)}(\tau)} = 0$.

Third-order. The third-order contribution to the averaged expansion of equation (3.8) is

$$\begin{aligned}
\Lambda^{(3)} = -\frac{1}{2}T\Bigg[&\epsilon X_S, \epsilon X_S + \epsilon^2\overline{\mathrm{V}_{(1)}^{(a)}(t)}\,[Y_a, X_S] + \frac{1}{2}\epsilon^2\overline{\mathrm{V}_{(1,0)}^{(a,b)}(t)}\,[Y_a, Y_b]\Bigg] \\
&+ \frac{1}{3T}\int_0^T \left[\epsilon X_S t + \mathrm{V}_{(1)}^{(a)}(t)Y_a, \left[\epsilon X_S t + \mathrm{V}_{(1)}^{(b)}(t)Y_b, \epsilon X_S t + \mathrm{V}_{(0)}^{(c)}(t)Y_c\right]\right]\,\mathrm{d}t\,.
\end{aligned} \tag{3.119}$$

Using the bilinear property of the Jacobi-Lie bracket,

$$\begin{aligned}\Lambda^{(\mathcal{3})} =& -\frac{1}{2}T\epsilon^2\,[X_S,X_S] - \frac{1}{2}T\epsilon^3\overline{\mathrm{V}^{(a)}_{(1)}(t)}\,[X_S,[Y_a,X_S]] - \frac{1}{2}T\epsilon^3\overline{\mathrm{V}^{(a,b)}_{(1,0)}(t)}\,[X_S,[Y_a,Y_b]]\\ &+\frac{1}{3T}\epsilon^3\int_0^T t^2\,\mathrm{d}t\;[X_S,[X_S,X_S]] + \frac{1}{3T}\epsilon^3\int_0^T t^2\mathrm{V}^{(a)}_{(0)}(t)\,\mathrm{d}t\;[X_S,[X_S,Y_a]]\\ &+\frac{1}{3T}\epsilon^3\int_0^T t\mathrm{V}^{(a)}_{(1)}(t)\,\mathrm{d}t\;[X_S,[Y_a,X_S]] + \frac{1}{3T}\epsilon^3\int_0^T t\mathrm{V}^{(b,c)}_{(1,0)}(t)\,\mathrm{d}t\;[X_S,[Y_a,Y_b]]\\ &+\frac{1}{3T}\epsilon^3\int_0^T t\mathrm{V}^{(a)}_{(1)}(t)\,\mathrm{d}t\;[Y_a,[X_S,X_S]] + \frac{1}{3T}\epsilon^3\int_0^T t\mathrm{V}^{(a,b)}_{(1,0)}(t)\,\mathrm{d}t\;[Y_a,[X_S,Y_b]]\\ &+\frac{1}{3T}\epsilon^3\int_0^T \mathrm{V}^{(a,b)}_{(1,1)}(t)\,\mathrm{d}t\;[Y_a,[Y_b,X_S]] + \frac{1}{3T}\epsilon^3\int_0^T \mathrm{V}^{(a,b,c)}_{(1,1,0)}(t)\,\mathrm{d}t\;[Y_a,[Y_b,Y_c]]\,.\end{aligned}$$

Removal of the vanishing Jacobi-Lie brackets and integration by parts leads to

$$\begin{aligned}\Lambda^{(\mathcal{3})} =& -\frac{1}{2}T\epsilon^3\overline{\mathrm{V}^{(a)}_{(1)}(t)}\,[X_S,[Y_a,X_S]] - \frac{1}{2}T\epsilon^3\overline{\mathrm{V}^{(a,b)}_{(1,0)}(t)}\,[X_S,[Y_a,Y_b]]\\ &+\frac{1}{3T}\epsilon^3\left(-2T^2\overline{\mathrm{V}^{(a)}_{(1)}(t)} + 2T\overline{\mathrm{V}^{(a)}_{(2)}(t)}\right)[X_S,[X_S,Y_a]]\\ &+\frac{1}{3T}\left(T^2\overline{\mathrm{V}^{(a)}_{(1)}(t)} - T\overline{\mathrm{V}^{(a)}_{(2)}(t)}\right)[X_S,[Y_a,X_S]]\\ &+\frac{1}{3T}\epsilon^3\left(T^2\overline{\mathrm{V}^{(a,b)}_{(1,0)}(t)} - T\overline{\mathrm{V}^{(\widehat{a,b})}_{(\widehat{1,0})}(t)}\right)[X_S,[Y_a,Y_b]]\\ &+\frac{1}{3T}\epsilon^3\left(T^2\overline{\mathrm{V}^{(a,b)}_{(1,0)}(t)} - T\overline{\mathrm{V}^{(\widehat{a,b})}_{(\widehat{1,0})}(t)}\right)[Y_a,[X_S,Y_b]]\\ &+\frac{1}{3}\epsilon^3\overline{\mathrm{V}^{(a,b)}_{(1,1)}(t)}\,[Y_a,[Y_b,X_S]] + \frac{1}{3}\epsilon^3\overline{\mathrm{V}^{(a,b,c)}_{(1,1,0)}(t)}\,[Y_a,[Y_b,Y_c]]\,,\end{aligned}$$

which can further be simplified by consolidating like terms,

$$\begin{aligned}\Lambda^{(\mathcal{3})} =& \epsilon^3\left(\frac{1}{2}T\overline{\mathrm{V}^{(a)}_{(1)}(t)} - \frac{2}{3}T\overline{\mathrm{V}^{(a)}_{(1)}(t)} + \frac{2}{3}\overline{\mathrm{V}^{(a)}_{(2)}(t)} - \frac{1}{3}T\overline{\mathrm{V}^{(a)}_{(1)}(t)} + \frac{1}{3}\overline{\mathrm{V}^{(a)}_{(2)}(t)}\right)[X_S,[X_S,Y_a]]\\ &+\epsilon^3\left(-\frac{1}{2}T\overline{\mathrm{V}^{(a,b)}_{(1,0)}(t)} + \frac{1}{3}T\overline{\mathrm{V}^{(a,b)}_{(1,0)}(t)} - \frac{1}{3}\overline{\mathrm{V}^{(\widehat{a,b})}_{(\widehat{1,0})}(t)}\right)[X_S,[Y_a,Y_b]]\\ &+\frac{1}{3}\epsilon^3\left(\overline{\mathrm{V}^{(a,b)}_{(1,1)}(t)} - T\overline{\mathrm{V}^{(a,b)}_{(1,0)}(t)} + \overline{\mathrm{V}^{(\widehat{a,b})}_{(\widehat{1,0})}(t)}\right)[Y_a,[Y_b,X_S]]\\ &+\frac{1}{3}\epsilon^3\overline{\mathrm{V}^{(a,b,c)}_{(1,1,0)}(t)}\,[Y_a,[Y_b,Y_c]]\,.\end{aligned}$$

The final form of the third variation is

$$\begin{aligned}\Lambda^{(3)} =&\epsilon^3\left(\overline{\mathrm{V}^{(a)}_{(2)}(t)}-\frac{1}{2}T\overline{\mathrm{V}^{(a)}_{(1)}(t)}\right)[X_S,[X_S,Y_a]]\\&-\frac{1}{3}\epsilon^3\left(\overline{\mathrm{V}^{(\widehat{a,b})}_{(\widehat{1,0})}(t)}+\frac{1}{2}T\overline{\mathrm{V}^{(a,b)}_{(1,0)}(t)}\right)[X_S,[Y_a,Y_b]]\\&+\frac{1}{3}\epsilon^3\left(\overline{\mathrm{V}^{(a,b)}_{(1,1)}(t)}+\overline{\mathrm{V}^{(\widehat{a,b})}_{(\widehat{1,0})}(t)}-T\overline{\mathrm{V}^{(a,b)}_{(1,0)}(t)}\right)[Y_a,[Y_b,X_S]]\\&+\frac{1}{3}\epsilon^3\overline{\mathrm{V}^{(a,b,c)}_{(1,1,0)}(t)}\,[Y_a,[Y_b,Y_c]]\,.\end{aligned}$$

The final form of the third order truncation of the averaged autonomous vector field, Z, is

$$\begin{aligned}\mathrm{Trunc}_3\,(Y)&=\epsilon\Lambda^{(1)}+\epsilon^2\Lambda^{(2)}+\epsilon^3\Lambda^{(3)}\\&=\epsilon X_S(z)+\epsilon^2\overline{\mathrm{V}^{(a)}_{(1)}(t)}\,[Y_a(z),X_S(z)]+\frac{1}{2}\epsilon^2\overline{\mathrm{V}^{(a,b)}_{(1,0)}(t)}\,[Y_a(z),Y_b(z)]\\&\quad+\epsilon^3\left(\overline{\mathrm{V}^{(a)}_{(2)}(t)}-\frac{1}{2}T\overline{\mathrm{V}^{(a)}_{(1)}(t)}\right)[X_S,[X_S,Y_a]]\\&\quad-\frac{1}{3}\epsilon^3\left(\overline{\mathrm{V}^{(\widehat{a,b})}_{(\widehat{1,0})}(t)}+\frac{1}{2}T\overline{\mathrm{V}^{(a,b)}_{(1,0)}(t)}\right)[X_S,[Y_a,Y_b]]\\&\quad+\frac{1}{3}\epsilon^3\left(\overline{\mathrm{V}^{(a,b)}_{(1,1)}(t)}+\overline{\mathrm{V}^{(\widehat{a,b})}_{(\widehat{1,0})}(t)}-T\overline{\mathrm{V}^{(a,b)}_{(1,0)}(t)}\right)[Y_a,[Y_b,X_S]]+\frac{1}{3}\epsilon^3\overline{\mathrm{V}^{(a,b,c)}_{(1,1,0)}(t)}\,[Y_a,[Y_b,Y_c]]\,.\end{aligned}$$

A transformation back to time $\tau\mapsto t/\epsilon$ gives

$$\begin{aligned}\mathrm{Trunc}_3\,(Y)&=X_S(z)+\epsilon\overline{\mathrm{V}^{(a)}_{(1)}(t)}\,[Y_a(z),X_S(z)]+\frac{1}{2}\epsilon\overline{\mathrm{V}^{(a,b)}_{(1,0)}(t)}\,[Y_a(z),Y_b(z)]\\&\quad+\epsilon^2\left(\overline{\mathrm{V}^{(a)}_{(2)}(t)}-\frac{1}{2}T\overline{\mathrm{V}^{(a)}_{(1)}(t)}\right)[X_S,[X_S,Y_a]]\\&\quad-\frac{1}{3}\epsilon^2\left(\overline{\mathrm{V}^{(\widehat{a,b})}_{(\widehat{1,0})}(t)}+\frac{1}{2}T\overline{\mathrm{V}^{(a,b)}_{(1,0)}(t)}\right)[X_S,[Y_a,Y_b]]\\&\quad+\frac{1}{3}\epsilon^2\left(\overline{\mathrm{V}^{(a,b)}_{(1,1)}(t)}+\overline{\mathrm{V}^{(\widehat{a,b})}_{(\widehat{1,0})}(t)}-T\overline{\mathrm{V}^{(a,b)}_{(1,0)}(t)}\right)[Y_a,[Y_b,X_S]]+\frac{1}{3}\epsilon^2\overline{\mathrm{V}^{(a,b,c)}_{(1,1,0)}(t)}\,[Y_a,[Y_b,Y_c]]\,.\end{aligned}$$

Fourth-order. The fourth-order terms get very complicated due to the Jacobi-Lie brackets with X_S. Under the assumption that $X_S=0$, fourth order computation is tractable. The additional term introduced by fourth-order averaging is,

$$\begin{aligned}\Lambda^{(4)}=&-\frac{1}{12}\overline{\int^{\tau}\left[\int^{\tau_1}\left[\int^{\tau_2}X_{\tau_3}\,\mathrm{d}\tau_3\,,X_{\tau_2}\right]\,\mathrm{d}\tau_2\,,[X_{\tau_1},X_\tau]\right]\,\mathrm{d}\tau_1}\\&-\frac{1}{12}\overline{\left[\int^{\tau}\left[\int^{\tau_1}\left[\int^{\tau_2}X_{\tau_3}\,\mathrm{d}\tau_3\,,X_{\tau_2}\right]\,\mathrm{d}\tau_2\,,X_{\tau_1}\right]\,\mathrm{d}\tau_1\,,X_\tau\right]}\\&-\frac{1}{12}\overline{\int^{\tau}\left[\int^{\tau_1}X_{\tau_2}\,\mathrm{d}\tau_2\,,\left[\left[\int^{\tau_1}X_{\tau_2}\,\mathrm{d}\tau_2\,,X_{\tau_1}\right],X_\tau\right]\right]\,\mathrm{d}\tau_1}\,.\end{aligned}$$

Substitution of the actual controls gives

$$\begin{aligned}
\Lambda^{(4)} =& -\frac{1}{12}\overline{\int^{\tau}\left[\int^{\tau_1}\left[\int^{\tau_2}\mathrm{V}^{(d)}_{(0)}(\tau_3)Y_d\,\mathrm{d}\tau_3\,,\mathrm{V}^{(c)}_{(0)}(\tau_2)Y_c\right]\,\mathrm{d}\tau_2\,,\left[\mathrm{V}^{(b)}_{(0)}(\tau_1)Y_b,\mathrm{V}^{(a)}_{(0)}(\tau)Y_a\right]\right]\,\mathrm{d}\tau_1}\\
&-\frac{1}{12}\overline{\left[\int^{\tau}\left[\int^{\tau_1}\left[\int^{\tau_2}\mathrm{V}^{(d)}_{(0)}(\tau_3)Y_d\,\mathrm{d}\tau_3\,,\mathrm{V}^{(c)}_{(0)}(\tau_2)Y_c\right]\,\mathrm{d}\tau_2\,,\mathrm{V}^{(b)}_{(0)}(\tau_1)Y_b\right]\,\mathrm{d}\tau_1\,,\mathrm{V}^{(a)}_{(0)}(\tau)Y_a\right]}\\
&-\frac{1}{12}\overline{\int^{\tau}\left[\int^{\tau_1}\mathrm{V}^{(d)}_{(0)}(\tau_2)Y_d\,\mathrm{d}\tau_2\,,\left[\left[\int^{\tau_1}\mathrm{V}^{(c)}_{(0)}(\tau_2)Y_c\,\mathrm{d}\tau_2\,,\mathrm{V}^{(b)}_{(0)}(\tau_1)Y_b\right],\mathrm{V}^{(a)}_{(0)}(\tau)Y_a\right]\right]\,\mathrm{d}\tau_1}\\
=& -\frac{1}{12}\overline{\int^{\tau}\int^{\tau_1}\int^{\tau_2}\mathrm{V}^{(d)}_{(0)}(\tau_3)\,\mathrm{d}\tau_3\,\mathrm{V}^{(c)}_{(0)}(\tau_2)Y_c\,\mathrm{d}\tau_2\,\mathrm{V}^{(b)}_{(0)}(\tau_1)\mathrm{V}^{(a)}_{(0)}(\tau)Y_a\,\mathrm{d}\tau_1}\;[[Y_d,Y_c],[Y_b,Y_a]]\\
&-\frac{1}{12}\overline{\int^{\tau}\int^{\tau_1}\int^{\tau_2}\mathrm{V}^{(d)}_{(0)}(\tau_3)Y_d\,\mathrm{d}\tau_3\,\mathrm{V}^{(c)}_{(0)}(\tau_2)\,\mathrm{d}\tau_2\,\mathrm{V}^{(b)}_{(0)}(\tau_1)\,\mathrm{d}\tau_1\,\mathrm{V}^{(a)}_{(0)}(\tau)}\;[[[Y_d,Y_c],Y_b],Y_a]\\
&-\frac{1}{12}\overline{\int^{\tau}\int^{\tau_1}\mathrm{V}^{(d)}_{(0)}(\tau_2)\,\mathrm{d}\tau_2\int^{\tau_1}\mathrm{V}^{(c)}_{(0)}(\tau_2)\,\mathrm{d}\tau_2\,\mathrm{V}^{(b)}_{(0)}(\tau_1)\mathrm{V}^{(a)}_{(0)}(\tau)\,\mathrm{d}\tau_1}\;[Y_d,[[Y_c,Y_b],Y_a]]\\
=& -\frac{1}{12}\overline{\mathrm{V}^{(\widehat{d,c,b},a)}_{(\widehat{1,0,0},0)}(\tau)}\;[[Y_d,Y_c],[Y_b,Y_a]]-\frac{1}{12}\overline{\mathrm{V}^{(\widehat{d,c,b},a)}_{(\widehat{1,0,0},0)}(\tau)}\;[[[Y_d,Y_c],Y_b],Y_a]\\
&-\frac{1}{12}\overline{\mathrm{V}^{(\widehat{d,c,b},a)}_{(\widehat{1,1,0},0)}(\tau)}\;[Y_d,[[Y_c,Y_b],Y_a]]\\
=& +\frac{1}{12}\overline{\mathrm{V}^{(\widehat{d,c,b,a})}_{(\widehat{1,0},0,1)}(\tau)}\;[[Y_d,Y_c],[Y_b,Y_a]]+\frac{1}{12}\overline{\mathrm{V}^{(\widehat{d,c,b,a})}_{(\widehat{1,0},0,1)}(\tau)}\;[[[Y_d,Y_c],Y_b],Y_a]\\
&+\frac{1}{12}\overline{\mathrm{V}^{(d,c,b,a)}_{(1,1,0,1)}(\tau)}\;[Y_d,[[Y_c,Y_b],Y_a]]\,.
\end{aligned}$$

Rearranging the indices and the Jacobi-Lie brackets,

$$\begin{aligned}
\Lambda^{(4)} =& \frac{1}{12}\overline{\mathrm{V}^{(a,b,\widehat{c,d})}_{(1,0,\widehat{1,0})}(\tau)}\;[[Y_a,Y_b],[Y_c,Y_d]]-\frac{1}{12}\overline{\mathrm{V}^{(a,b,\widehat{c,d})}_{(1,0,\widehat{0,1})}(\tau)}\;[Y_a,[Y_b,[Y_c,Y_d]]]\\
&-\frac{1}{12}\overline{\mathrm{V}^{(a,b,c,d)}_{(1,1,1,0)}(\tau)}\;[Y_a,[Y_b,[Y_c,Y_d]]]\,.
\end{aligned}$$

In summary, the additional fourth-order contribution is

$$\Lambda^{(4)}=\frac{1}{12}\left(\overline{\mathrm{V}^{(a,b,\widehat{c,d})}_{(1,0,\widehat{1,0})}(t)}\;[[Y_a,Y_b],[Y_c,Y_d]]-\left(\overline{\mathrm{V}^{(a,b,\widehat{c,d})}_{(1,0,\widehat{0,1})}(t)}+\overline{\mathrm{V}^{(a,b,c,d)}_{(1,1,1,0)}(t)}\right)[Y_a,[Y_b,[Y_c,Y_d]]]\right).$$

The associated Floquet mapping is,

$$\begin{aligned}
\mathrm{Trunc}_3\left(P(t/\epsilon)\right) =& \mathrm{Id}+\epsilon\mathrm{V}^{(a)}_{(1)}(t/\epsilon)\;Y_a+\frac{1}{2}\epsilon^2\int_0^{t/\epsilon}\widetilde{\mathrm{V}}^{(a,b)}_{(1,0)}(\tau)\,\mathrm{d}\tau\;[Y_a,Y_b]+\frac{1}{2}\epsilon^2\mathrm{V}^{(a,b)}_{(1,1)}(t/\epsilon)\;Y_a\cdot Y_b\\
&-\frac{1}{2}\epsilon^2\mathrm{V}^{(a)}_{(1)}(t/\epsilon)\overline{\mathrm{V}^{(b,c)}_{(1,0)}(t)}\,t\,Y_a\cdot[Y_b,Y_c]+\frac{1}{6}\epsilon^3\mathrm{V}^{(a,b,c)}_{(1,1,1)}(t)Y_a\cdot Y_b\cdot Y_c\\
&+\frac{1}{4}\epsilon^3\mathrm{V}^{(a)}_{(1)}(t/\epsilon)\int_0^{t/\epsilon}\mathrm{V}^{(b,c)}_{(1,0)}(\sigma)\,\mathrm{d}\sigma\;\left(Y_a\cdot[Y_b,Y_c]+[Y_b,Y_c]\cdot Y_a\right).
\end{aligned}$$

Chapter 4

Controllability and 1-Homogeneous Control Systems

Mechanical control systems form a large and important subclass of nonlinear control systems. Besides their practical utility, mechanical systems also have inherent structure that simplifies their analysis. This chapter develops a detailed understanding of this structure in terms of geometric homogeneity. The homogeneous properties of mechanical systems are then used to simplify the analysis of mechanical system controllability.

The problem of determining controllability for nonlinear control systems is in general quite difficult. Sussmann [152] provides sufficient conditions to determine small-time local controllability of general nonlinear systems, however there is factorial growth in the number of elements to test. Lewis and Murray [85] showed how the Riemannian or affine connection description of mechanical systems prunes the tests found in [152] for the particular case of configuration controllability. Their work is applied to *simple mechanical systems*, which form a subset of mechanical systems. Simple mechanical systems are characterized by a Lagrangian with kinetic and potential energy terms only; constraints may be included. In [25], Bullo and Lewis demonstrate that geometric homogeneity is behind much of the simplifications. Geometric homogeneity has the interpretation of being an infinitesimal symmetry of the equations of motion [123]. Bullo and Lewis further argued that geometric homogeneity should be a fruitful avenue of continued study for nonlinear control systems. This is especially true given that analogous controllability results have subsequently been found to hold for other descriptions of mechanical control systems [8, 16, 35, 125]. This chapter develops the concept of geometric homogeneity for vector bundles and demonstrates how geometric homogeneity is the connecting link to all of these seemingly related results.

Organization of this chapter. Section 4.1 reviews and develops notions of geometric homogeneity for application to *1-homogeneous control systems*. This section culminates with results that describe the consequences of homogeneity vis-a-vis the Lie-algebraic structure of dynamical systems evolving on vector bundles. These results are subsequently used in Section 4.2 to extend the configuration controllability theorems of Lewis and Murray to a generic vector bundle. This section closely follows the structure of Lewis and Murray [85] to aid the reader in understanding and comparing the relevant concepts to this prior work. Since dissipation is an important element of realistic control systems, and not accounted for in the simple mechanical system paradigm, Section 4.3 analyzes the effects of including dissipation on the homogeneous nature of mechanical control systems. For some classes of control problems, the configuration controllability results do not differ from the case without dissipation. For others, the inclusion of dissipation complicates the controllability analysis. Finally, Section 4.4, applies our results to various instances of control sytems evolving on vector bundles. Since most of these examples have previously been studied in the literature, our analysis serves to demonstrate how these prior results are all connected and may be placed within the same framework.

4.1 Geometric Homogeneity and Vector Bundles

This section reviews the elementary concepts related to vector bundles and geometric homogeneity. Geometric homogeneity has been a widely, but not uniformly studied aspect of nonlinear dynamics and control. Controllability analysis often involves the use of dilations [53, 54, 14, 110, 152]. Kawski has successfully demonstrated the importance of geometric homogeneity for stabilization [61, 62], as have M'Closkey and Morin et al. [109, 107, 117]. Geometric homogeneity also plays a role in describing the geometry of Lagrangian systems [36]. This section discusses the relevant ideas concerning geometric homogeneity of vector bundles and vector fields in a systematic manner. Doing so will require working knowledge of differential geometry and its attendant terminology; familiarity with the notation found such texts [1, 18, 70] is assumed.

4.1.1 Vector Bundles

The nonlinear control systems analyzed in this paper can be cast into a vector bundle framework. Consequently, we formulate homogeneity in terms of vector bundles. For background material, see standard texts such as [1, 70, 150].

Definition 27 *A (differentiable) fibre bundle (Q, π, M, F, G) consists of the following elements:*

1. *differentiable manifold Q called the total space (configuration space)*
2. *differentiable manifold M called the base space (shape space)*
3. *differentiable manifold F called the fibre (typical fibre).*
4. *A surjection $\pi : Q \to M$ called the projection: $\pi^{-1}(p) = F_p \cong F$ is called the fibre at p.*
5. *A Lie group G called the structure group, which acts on F (on the left).*
6. *A set of open coverings $\{V_i\}$ of M with diffeomorphisms $\phi_i : V_i \times F \to \pi^{-1}(V_i)$ such that $\pi \circ \phi_i(p, f) = p$, where $p \in V_i$ and $f \in F$. The map ϕ_i is called the* local trivialization *since ϕ_i^{-1} maps $\pi^{-1}(V_i)$ onto the direct product $V_i \times F$.*
7. *If defined by $\phi_{i,p}(f) \equiv \phi_i(p, f)$, then the map $\phi_{i,p} : F \to F_p$ is a diffeomorphism. On $U_i \cap U_j \neq \emptyset$, the* transition function *$t_{ij}(p) = \phi_{i,p}^{-1} \circ \phi_{j,p} : F \to F$ must be an element of G. Then ϕ_i and ϕ_j are related by a smooth map $t_{ij} : U_i \cap U_j \to G$ as*

$$\phi_j(p, f) = \phi_i(p, \Phi_{t_{ij}(p)} f).$$

Definition 28 *Given a fibre bundle (Q, π, M, F, G), a* section (cross section) *$\sigma : M \to Q$ is a smooth map satisfying $\pi \circ \sigma = \mathrm{id}_M$. The space defined by the collection of all sections on the fiber bundle Q is denoted by $\Gamma(Q)$. $\Gamma(Q)$ is a fiber bundle.*

The analysis in this paper will focus on vector bundles. Typically the structure group, G, will be the general linear group, $GL(\cdot)$, which acts as a transformation of basis of the vector fiber [70].

Definition 29 *A* vector bundle *(E, π, M, V) is fiber bundle whose typical fiber is a vector space.*

Definition 30 *Given a vector bundle (E, π, M, V), the* zero section *, denoted by σ_0, is a smooth mapping of points in the base space to the zero vector.*

Often, we will refer to a local trivialization of the vector bundle, E, as $E \cong M \times V$. Given a local trivialization, the zero section operates as follows,

$$\sigma_0(x) = 0_x \in E_x, \quad \forall\, x \in M.$$

As a vector bundle, the zero section $M \times \{\, 0 \,\}$ is isomorphic to the base space, M.

Products of Bundles. If two bundles share the same base space, $M = M_1 = M_2$, they may be composed via the *fiber product*, denoted $\times_M$, to form a new bundle $Q_1 \times_M Q_2$, with base space M. Let vector bundles E_1 and E_2 have a common base space. Their fiber product, known as the *Whitney sum*, is denoted $E_1 \oplus E_2$. In a local trivialization, the Whitney sum is $E_1 \oplus E_2 \cong M \times V_1 \times V_2$.

The Tangent Bundle. The tangent bundle is a canonical example of a vector bundle, and will play a prominent role in this paper. In particular, we review the tangent bundle, TE, to the vector bundle, E. A local trivialization for E is given by

$$E \cong M \times V.$$

Let the tangent space, TM, to M, have the local trivialization

$$TM \cong M \times W,$$

where W models the tangent fiber of M. It is known that the fiber of the tangent space to a vector space can be modelled by the vector space itself. This means that,

$$TV \cong V \times V.$$

Logically then, the local trivialization of the tangent space TE, can be given by

$$TE \cong M \times V \times W \times V, \tag{4.1}$$

where $M \times V$ represents the vector bundle, and $W \times V$ represents the tangent fiber to the vector bundle. The trivialization suggests that two natural projection operators exist for TE, ρ_1 and ρ_2 [70]. In a local trivialization, ρ_1 operates as follows,

$$\rho_1 : M \times V \times W \times V \to M \times V \times W \times \{\, 0 \,\}, \quad \rho_1(q, u, w, v) = (q, u, w, 0), \tag{4.2}$$

and ρ_2 is

$$\rho_2 : M \times V \times W \times V \to M \times V \times \{\, 0 \,\} \times V, \quad \rho_2(q, u, w, v) = (q, u, 0, v). \tag{4.3}$$

Note that $\rho_1 = \mathrm{id}_{TE} - \rho_2$.

Very often, the components of a fiber bundle are implicitly known, justifying a shorter notation. When dealing with vector bundles, we will henceforth assume that a vector bundle E consists of the collection (E, π, M, V), although it will not be explicitly noted. Likewise, the tangent bundle TE will be given by $(TE, T\pi, E, W \times V)$. Additionally, all vector bundles are assumed to be finite dimensional.

4.1.2 Geometric Homogeneity

Much of the terminology concerning geometric homogeneity can be found in Crampin [36] or Kawski [62]. Geometric homogeneity is defined using the dilation operator δ_t, which in a local trivialization uniformly dilates the vector fiber,

$$\delta_t : E \to E, \quad (x, u) \mapsto (x, e^t\, u). \tag{4.4}$$

The dilation operator satisfies $(\delta_t)^p = \delta_{pt}$. Corresponding to the dilation is its infinitesimal generator, Δ, a vector field on E. In the coordinates of a local trivialization, the generator is

$$\Delta = u^i \frac{\partial}{\partial u^i}. \tag{4.5}$$

There is an alternative to the dilation operator which exchanges the additive property of the dilation with a multiplicative one,

$$\tilde{\delta}_c \equiv \delta_{ln(c)} : E \to E, \quad (x, u) \mapsto (x, c\, u). \tag{4.6}$$

Our working definition of homogeneity follows.

Definition 31 *A mapping between bundles $\Psi : E_1 \to E_2$ is* homogeneous of order p *if*

$$\Psi \circ \delta_t = \delta_{pt} \circ \Psi.$$

The two bundles need not have the same base nor fiber. We assume that Ψ is a smooth function defined over E, and that all vector bundles are finite dimensional. Below, E_1 and E_2 are vector bundles.

Proposition 5 *All smooth functions, $\Psi : E_1 \to E_2$ of homogeneous order 0 are equivalent to mapping of the zero-section on the vector bundle E_1 into E_2.*

proof
If Ψ is homogeneous of order0, then $\Psi \circ \tilde{\delta}_0 = \Psi$. As the (multiplicative) zero dilation, $\tilde{\delta}_0$, maps arbitrary sections to the zero section, the conclusion naturally follows.
■

Proposition 6 *[70] The only smooth functions, $\Psi : E_1 \to E_2$ of homogeneous order $-p$ for $p > 0$ are mappings of the zero section on the vector bundle E_1 to the zero-section on the vector bundle E_2.*

Proposition 7 *If $\Psi : E_1 \to E_2$ is homogeneous of order $p \geq 1$, then Ψ evaluated on the zero section of E_1 results in the zero section of E_2.*

proof
From Equation (4.6), evaluation of Ψ on the zero section of E_1 is equivalent to $\Psi \circ \tilde{\delta}_0$. Applying Definition 31 yields

$$\Psi \circ \tilde{\delta}_0 = \tilde{\delta}_{0^p} \circ \Psi = \tilde{\delta}_0 \circ \Psi,$$

and we obtain the zero section on E_2.
■
Configuration controllability evaluations are made on the zero section of a vector bundle; the above homogeneity results will allow certain vector fields to be ignored. This is precisely how Lewis and Murray [85] achieved a simplification of Sussmann's controllability theorem.

4.1.3 The Tangent Bundle and Vector Fields

Equally important are the homogeneous properties of vectors in TE, the tangent bundle to E, or equivalently of vector fields in $\mathcal{X}(E)$, the space of vector fields defined on E. We focus now on the properties of the tangent bundle TE. Using the projection operator there is a natural notion of *vertical* for TE.

Definition 32 *The* vertical bundle over E*, denoted by* VE*, is the subbundle of* TE *given by the union of* $T\pi^{-1}(0_q)$ *for all* $q \in Q$*. A vector in* TE *is* vertical *if it lies in the kernel of* $T\pi$*.*

Definition 33 *There is a canonical isomorphism between* $E \oplus E$ *and* VE*, called the* vertical lift*. It is given by*

$$v^{\text{lift}} = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} (u_x + tv_x), \quad u, v \in E. \tag{4.7}$$

A complementary *horizontal* subbundle can be defined using a *connection form*, also called an *Ehresmann connection*. We rewrite below, the definition of an Ehresmann connection,

Definition 34 *An* Ehresmann connection*,* A*, is a vertical valued one form on a manifold* Q *that satisfies,*

1. $A_q : T_qQ \to V_qQ$ *is a linear map for each point* $q \in Q$*.*
2. A *is a projection,* $A \circ A = A$*.*

There is no unique decomposition of the tangent bundle into horizontal and vertical, however there exists a trivial decomposition into horizontal and vertical subbundles. Its associated Ehresmann connection is called the *trivial connection*. On the vector bundle E, the natural projection operator, ρ_2, can be used to define the trivial connection on TE,

$$A^{\text{triv}} : TE \to VE, \quad A^{\text{triv}} = \rho_2. \tag{4.8}$$

Pointwise, horizontal vectors are isomorphic to vectors on the tangent bundle of the base space using the trivial connection and the bundle projection, $T\pi$. In local trivializations for TM and TE, the injective mapping is given by

$$v_x \in T_xM, \quad v_x = T\pi \circ u_{(x,u)}, \text{ for } u_{(x,u)} \in H_{(x,u)}E. \tag{4.9}$$

From (4.2), the horizontal bundle is in 1-1 correspondance with the Whitney sum $E \oplus TM$.

Vector Fields

Properties that hold for tangent vectors will also hold for sections on TE, which give rise to vector fields $\mathcal{X}(E)$. Through the following derivations and propositions, we study how geometric homogeneity applies to vector fields.

Definition 35 *The subspace* $\mathcal{X}^V(E) \subset \mathcal{X}(E)$*, consists of all vector fields that are vertical for all points in* E*.*

Proposition 8 *If* $X, Y \in \mathcal{X}^V(E)$*, then* $[X, Y] \in \mathcal{X}^V(E)$*.*

The notion of homogeneity extends to vector fields via the generator, Δ.

Definition 36 *A vector field* $X \in \mathcal{X}(E)$ *is said to be* homogeneous of order p *if,*

$$[\Delta, X] = pX$$

for $p > -2$*.*

The only vector field of homogeneous order $p \leq -2$ is the zero vector field. Under the trivial splitting of vector fields into horizontal and vertical, the dilation vector field has the following properties,

Proposition 9 *Let* Δ *be the infinitesimal generator corresponding to the dilation action,* δ_t*. Given a vector field* $X \in \mathcal{X}(E)$*, the following properties hold:*

1. $\left[\Delta, X^H\right]$ *is horizontal, and*
2. $\left[\Delta, X^V\right]$ *is vertical,*

for the trivial decomposition, $X = X^H + X^V$, of X into horizontal and vertical components, respectively.

proof
Since horizontal vectors in TE are in 1-1 correspondance with vectors in TM, horizontal vector fields are in 1-1 correspondance with functions in $C^\infty(E, TM)$. We may, therefore, make the identification

$$\rho_1(X) = X^H \mapsto Y \in C^\infty(E, TM).$$

In a local trivialization, the dynamic definition of the Jacobi-Lie bracket is

$$\left[\Delta, X^H\right] \equiv \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} \left((\delta_t)^* X^H\right) = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} (\delta_t)^* (Y, 0) = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} (T\delta_{-t}(Y \circ \delta_t, 0)).$$

Since the dilation operator leaves the base component invariant, the resulting vector field remains horizontal using the trivial decomposition. The second statement is a consequence of Proposition 8, as the dilation vector field, Δ, is vertical.
■

Proposition 10 *Given $X, Y \in \mathcal{X}(E)$ homogeneous of order p and q, respectively, the Jacobi-Lie bracket $[X, Y]$ is homogeneous of order $p + q$.*

Proposition 11 *Any mapping $\Psi : E \to E$ preserving the base, i.e., $\pi = \pi \circ \Psi$, and homogeneous of order p, loses one degree of homogeneity when vertically lifted. That is $\Psi^{\mathrm{lift}} \in \mathcal{X}(E)$ is homogeneous of order $p - 1$.*

proof
By Definition 33, the vertical lift of Ψ is a vector field on E. Homogeneity of Ψ^{lift} is determined by Definition 36,

$$\left[\Delta, \Psi^{\mathrm{lift}}\right] = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} (\delta_t)^* \Psi^{\mathrm{lift}} = \frac{\mathrm{d}}{\mathrm{d}t}\left(T\delta_{-t} \circ \Psi^{\mathrm{lift}} \circ \delta_{-t}\right).$$

In a local trivialization for the tangent bundle,

$$\begin{aligned}
\left[\Delta, \Psi^{\mathrm{lift}}\right] &= \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} \left(0, e^{-t}\Psi \circ \delta_t\right) \\
&= \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} \left(0, e^{(p-1)t}\Psi\right) \\
&= (0, (p-1)\Psi) \\
&= (p-1)\Psi^{\mathrm{lift}}
\end{aligned}$$

■

Corollary 11 *Given a section of the vector bundle E, its vertical lift is homogeneous of order -1.*

As stated next, the converse to the corollary also holds.

Proposition 12 *All vector fields of homogeneous order -1 are the vertical lift of a section of E.*

proof
This is a commonly known result, c.f. Crampin [36], but will be proven regardless. Using Proposition 13, the horizontal component must be homogeneous of order -1, but by Proposition 6 the only such vector field is the zero vector field. Thus, all vector fields of homogeneous order -1 must be vertical. The rest follows from Propositions 5 and 11.
■

Proposition 13 *A vector field homogeneous of order p can be decomposed into horizontal and vertical components, $X = X^H + X^V$, that satisfy the following properties,*

1. *X^H is in 1-1 correspondence with a $Y \in C^\infty(E, TM)$ where Y is homogeneous of order p, and*
2. *X^V is in 1-1 correspondence with a Ψ^{lift}, where $\Psi \in C^\infty(E, E)$ is homogeneous of order $p+1$.*

proof
Using Proposition 9, Definition 36, and the linearity of the Jacobi-Lie bracket, the decomposition satisfies

$$\left[\Delta, X^H + X^V\right] = \left[\Delta, X^H\right] + \left[\Delta, X^V\right],$$

That is, the horizontality and verticality of both contributions are preserved. The second statement is a consequence of Proposition 11. The first statement follows by further carrying out the proof of Proposition 9.
■
Proposition 13 is analogous to the discussion following Eq. (3.1) in Bullo and Lewis [25], but extended to hold for arbitrary vector bundles.

Filtrations of Homogeneous Spaces. The subbundles created by examining spaces of constant homogeneous order form a filtration. The graded structure of a filtration means that homogeneous properties may be reduced to algebraic identities on the filtration order. Define the vector subbundle of homogeneous order k to be

$$\mathcal{P}_k(E) \equiv \{\, X \in \mathcal{X}(E) \mid X \text{ is of homogeneous degree } k. \,\}. \tag{4.10}$$

Propositions 6 and 10 imply that

1. $[\mathcal{P}_i(E), \mathcal{P}_j(E)] \subset \mathcal{P}_{i+j}(E)$,
2. $\mathcal{P}_k(E) = \{\, 0 \,\}, \ \forall\, k < -1$.

Accordingly, we may define the following union of homogeneous spaces,

$$\mathcal{M}_{j,k}(E) = \oplus_{i=j}^{i \le k} \mathcal{P}_i(E), \tag{4.11}$$

which inherit the properties of its constitutive sets:

1. $\left[\mathcal{M}_{i,j}(E), \mathcal{M}_{i',j'}(E)\right] \subset \mathcal{M}_{i+i',j+j'}(E)$,
2. $\mathcal{M}_{i,j}(E) = \{\, 0 \,\}\ , \ \forall\, i, j < -1$.

The following notational convenience is used: when the first subscript is absent, then it is assumed to be -1, e.g., $\mathcal{M}_i(E) \equiv \mathcal{M}_{-1,i}(E)$. It can be seen that

$$\mathcal{M}_{-1}(E) = \mathcal{P}_{-1}(E),$$

meaning that this "basic" space is not enlarged. The spaces $\mathcal{M}_k$ form a filtration.

Corollary 12 *If $X, Y \in \mathcal{M}_{-1}(E)$, then $[X, Y] = 0$.*

The Jacobi identity implies a symmetry of the following bracket constructions.

Corollary 13 *Given $X, Y \in \mathcal{M}_{-1}(E)$, then $[X, [\Gamma, Y]] = [Y, [\Gamma, X]]$ for any $\Gamma \in \mathcal{X}(E)$.*

Via Corollary 13, the (2,1)-tensor, $[\cdot, [\Gamma, \cdot]]$, may be used to define a symmetric product for vector fields in $\mathcal{M}_{-1}$.

Definition 37 *The* symmetric product of vector fields in $\mathcal{M}_{-1}(E)$ using the vector field Γ *is defined to be*

$$\langle\, X \;:\; Y \,\rangle^{\Gamma} \equiv [X, [\Gamma, Y]]\,, \tag{4.12}$$

where $X, Y \in \mathcal{M}_{-1}(E)$.

Very often the vector field Γ will be specified and the relevant space of vector fields restricted such that the shortened notation *symmetric product of vector fields*, or even *symmetric product*, will make implicit sense. In this case, we will simply write $\langle\, X \;:\; Y \,\rangle$ without reference to the vector field Γ. By $\mathcal{P}_k$ or $\mathcal{M}_k$, we will mean $\mathcal{P}_k(E)$ or $\mathcal{M}_k(E)$, respectively.

The symmetric product was originally derived and defined using the Riemmanian or affine connection structure of simple mechanical systems [37, 83]. However it is known, and we have shown, that is holds in the more general setting discussed here. What is critical to simple mechanical systems is not the symmetric product alone, but the homogeneous structure implied by the Lagrangian framework from which such systems are derived. The homogeneous structure will imply that, for mechanical systems, it is only Jacobi-Lie brackets with a structure similar to the one above that will be important from both control-theoretic and dynamical systems perspectives.

For the systems that we will study, the homogeneous order of the vector field Γ will not exceed 1, e.g. $\Gamma \in \mathcal{M}_1(E)$. Therefore $\langle\, X \;:\; Y \,\rangle \in \mathcal{M}_{-1}$, and the symmetric product of vector fields in $\mathcal{M}_{-1}$ is again in $\mathcal{M}_{-1}$. Most importantly, this implies that the symmetric product is a vertical lift, and commutes with other vertical lifts in $\mathcal{M}_{-1}$. It also implies that there exists a symmetric product defined on $\mathcal{X}(M)$ such that the following commutative diagram holds,

$$\begin{array}{ccc} X, Y & \xrightarrow{\text{lift}} & X^{\text{lift}}, Y^{\text{lift}} \\ {\scriptstyle \langle\cdot:\cdot\rangle}\downarrow & & \downarrow{\scriptstyle \langle\cdot:\cdot\rangle}, \\ Z & \xrightarrow[\text{lift}]{} & Z^{\text{lift}} \end{array} \tag{4.13}$$

where $X, Y, Z \in \Gamma(E)$ and $X^{\text{lift}}, Y^{\text{lift}}, Z^{\text{lift}} \in \mathcal{M}_{-1}(E)$. Alternatively stated, there exists a symmetric product on sections of E uniquely defined by

$$\langle\, X \;:\; Y \,\rangle^{\text{lift}} \equiv \left\langle\, X^{\text{lift}} \;:\; Y^{\text{lift}} \,\right\rangle, \tag{4.14}$$

where $X, Y \in \Gamma(E)$. Although it is not explicitly noted, the choice of Γ for the symmetric product of vector fields plays an important role in defining the symmetric product of sections, c.f. Equation (4.12). It was

demonstrated in Lewis and Murray [85] that equation (4.14) was responsible for computational reductions in the involutive closure calculations for simple mechanical systems. The structure of vector bundles and geometric homogeneity has been used to show that it holds even when $E \neq TQ$. Again, it is emphasized that the majority of canonical forms for mechanical systems discussed in the introduction satisfy the homogeneous properties assumed above. Systems with these properties will be called *1-homogeneous systems*. We shall further refine the definition of a 1-homogeneous system in Section 4.2.

4.2 Control of Dynamical Systems

This core section of the paper develops the main consequences of geometric homogeneity for controllability of 1-homogeneous systems and presents our main controllability results. Concepts and definitions for controllability of nonlinear affine control systems will be reviewed. Nonlinear control systems are typically written in affine control form,

$$\dot{z} = f(z) + g_a(z)u^a(t), \tag{4.15}$$

where $f \in \mathcal{X}(E)$ is the *drift vector field* and the $g_a \in \mathcal{X}(E)$ are *input* or *control vector fields*. The variable, z, evolves on the vector bundle E, and there are a total of m inputs to the system, $a = 1 \dots m$. The input functions u^a are assumed to be in the class of piecewise constant functions of time. We first review standard notions of controllability.

Definition 38 *The affine nonlinear control system (4.15) is* controllable *if for any two points $z_0, z_f \in E$, there exists controls $u^a(t)$ such that the system evolving via (4.15) and starting at z_0 reaches z_f in finite time.*

Let $\mathcal{R}^U(z_0, T)$ denote the set of reachable points in E at time $T > 0$, starting at point z_0 and using admissable controls, $u^a(t)$, such that the system trajectories remain in the neighborhood U of z_0 for all $t \leq T$. Define

$$\mathcal{R}^U_T(z_0) \equiv \bigcup_{0 \leq t \leq T} \mathcal{R}^U(z_0, t). \tag{4.16}$$

Definition 39 *The nonlinear control system (4.15) is* locally accessible *if for all $z \in E$, $\mathcal{R}^U_T(z)$ contains a non-empty open subset of E for all $T > 0$ and $U \subset E$.*

Local accessibility may be determined using the well-known Lie algebra rank condition (LARC).

Definition 40 *The nonlinear control system (4.15) is* locally controllable *if for all $z \in E$, z is in the interior of $\mathcal{R}^U_T(z)$ for all $T > 0$ and $U \subset E$.*

Definition 41 *A system (4.15) is* small-time locally controllable (STLC) at z_0 *if it is locally accessible at z_0 and if there exists a $T > 0$ such that z_0 is in the interior of $\mathcal{R}^U_t(z_0)$ for every neighborhood U of z_0 and $0 \leq t \leq T$. If this holds for any $z \in E$, the system is* small-time locally controllable.

Configuration Controllability. Suppose that one is not interested in achieving control over all of the vector bundle E, but would instead like to achieve arbitrary control on the base space M. This is equivalent to asking for configuration controllability of the system, where the configurations are elements of M. Let $\mathcal{RC}^U(x_0, T)$ denote the set of reachable configurations in M at time $T > 0$, starting at point $0_{x_0} \in E$ and using admissable controls, $u^a(t)$, such that the trajectories remain in the neighborhood U of x_0 for all $t \leq T$. Define

$$\mathcal{RC}^U_T(x_0) = \bigcup_{0 \leq t \leq T} \mathcal{RC}^U(x_0, t). \tag{4.17}$$

Definition 42 *[85] A system (4.15) is* locally configuration accessible at $x_0 \in M$ *if there exists a $T > 0$ such that $\mathcal{RC}_t^U(x_0)$ contains a non-empty open set of M for all neighborhoods U of x_0 and all $0 \leq t \leq T$. If this holds for any $x_0 \in M$, then the system is* locally configuration accessible.

Definition 43 *[85] A system (4.15) is* small-time locally configuration controllable (STLCC) at x_0 *if it is locally configuration accessible at x_0 and if there exists a $T > 0$ such that x_0 is in the interior of $\mathcal{RC}_t^U(x_0)$ for every neighborhood U of x_0 and $0 \leq t \leq T$. If this holds for any $x \in M$, then the system is called* small-time locally configuration controllable.

Definition 44 *The equilibrium subspace, $\mathfrak{E} \subset E$, of the nonlinear control system (4.15) consists of those elements of the zero section of E where $\dot{z} = 0$.*

Definition 45 *[85] A system (4.15) is* equilibrium controllable *if, for $z_1, z_2 \in \mathfrak{E}$, there exists a solution (c, u) of (4.15) where $c : [0, T] \to M$ is such that $c(0) = \pi(z_1)$, $c(T) = \pi(z_2)$, and $c'(0) = c'(T) = 0$.*

1-Homogeneous Control Systems

The nonlinear affine control system (4.15) will be rewritten using homogeneous and vector bundle notation,

$$\dot{z} = X_S(z) + Y_a^{\text{lift}}(z)u^a, \tag{4.18}$$

where X_S is the drift vector field, and Y_a^{lift} are the control vector fields.

Definition 46 *A* 1-homogeneous control system *is a nonlinear affine control system such that the drift vector field is $X_S \in \mathcal{M}_1$, and the control vector fields are $Y_a^{\text{lift}} \in \mathcal{M}_{-1}$.*

The remaining subsections are devoted to proving sufficient conditions for local configuration controllability.

4.2.1 Configuration Controllability Revisited

The goal of this section is to extend the work of [85] to the broader context of 1-homogeneous control systems. In doing so, the results will now apply to an enlarged class of systems, and will reproduce many of the known results concerning controllability of mechanical control systems. The content and structure found in Sussmann [152], and Lewis and Murray [85], is assumed familiar to the reader. We develop the ideas using the same language and construction. In some cases proofs will be crafted using the language of geometric homogeneity, although the analogous proofs from [85] do not require modification.

We will relate the control system (4.18) to free Lie algebras. Sussmann's work has demonstrated that much progress can be achieved by examining an algebraic abstraction of the vector fields found in the nonlinear affine control system (4.15). By placing the two descriptions in bijective homomorphic correspondence, properties of the free Lie algebra translate to properties of the original control system. Geometric homogeneity fits within this algebraic realization; we need only examine how bracketing affects the homogeneous order of free Lie algebra elements.

Recall the standard 1-homogeneous control form (4.18),

$$\dot{z} = X_S + Y_a^{\text{lift}}(z)u^a, \tag{4.19}$$

where $X_S \in \mathcal{M}_1$ is the drift vector field, the $Y_a^{\text{lift}} \in \mathcal{M}_{-1}$ are control vector fields, and the u^a are inputs, for $a = 1 \ldots m$. Although we will fully develop the case of $X_S \in \mathcal{M}_1$, currently we restrict ourselves to

the *simple homogeneous* case, where $X_S \in \mathcal{P}_1 \oplus \mathcal{P}_{-1}$. In this case, X_S can always be rewritten[1] as

$$X_S = \Gamma - Z^{\text{lift}}, \tag{4.20}$$

where $\Gamma \in \mathcal{P}_1$ and $Z^{\text{lift}} \in \mathcal{P}_{-1}$.

Free Lie algebras

Consider a finite set $\mathbf{X} = \{ X_0, X_1, \ldots, X_l \}$ of elements called indeterminants. $A(\mathbf{X})$ is the set of associative, but not necessarily commutative, products of indeterminants from the set $\mathbf{X}$. Define $I(\mathbf{X})$ to be the two-sided ideal generated by elements of the form $X \cdot X$ and $X \cdot (Y \cdot Z) + Z \cdot (X \cdot Y) + Y \cdot (Z \cdot X)$. The free Lie algebra, denoted $L(\mathbf{X})$, generated by $\mathbf{X}$ is the quotient algebra, $L(\mathbf{X}) = A(\mathbf{X})/I(\mathbf{X})$. The product under consideration, written $[\cdot, \cdot]$, will correspond to the Jacobi-Lie bracket, see Equation (4.24).

Definition 47 *Define $Br(\mathbf{X})$ to be the subset of $L(\mathbf{X})$ containing only products of elements in $\mathbf{X}$.*

The subset $Br(\mathbf{X})$ will generate $L(\mathbf{X})$ as a $\mathbb{R}$-vector space as can be seen using Proposition 14. Every element $X \in L(\mathbf{X})$ has a unique decomposition $X = [X_1, X_2]$. In turn, each of X_1 and X_2 may be uniquely expressed likewise, $X_1 = [X_{11}, X_{12}]$ and $X_2 = [X_{21}, X_{22}]$. This continues until the elements can no longer be decomposed in this manner. The elements $X_{i_1 \ldots i_k}$, $i_a \in 1, 2$, are called *components* of X. The decompositioon of X into components plays an important role in many of the following proofs.

For each element $B \in Br(\mathbf{X})$ let $\delta_a(X)$ be the number of times that the element X_a, $a = 0, \ldots, l$, occurs in B. The *degree of* B, denoted $\delta(\cdot)$, is the sum of the δ_a's,

$$\delta(B) \equiv \delta_0(B) + \sum_{a=1}^{l} \delta_a(B). \tag{4.21}$$

The *relative degree*, is the difference between δ_0 and the sum of the δ_a's for $a = 1 \ldots l$. The relative degree will be denoted by $\tilde{\delta}(\cdot)$, and mathematically it is written

$$\tilde{\delta}(B) \equiv \delta_0(B) - \sum_{a=1}^{l} \delta_a(B). \tag{4.22}$$

Proposition 14 *[85] Every element of $L(\mathbf{X})$ is a linear combination of repeated brackets of the form,*

$$[X_k, [X_{k-1}, [\ldots, [X_2, X_2] \ldots]]]$$

where $X_i \in \mathbf{X},\ i = 1, \ldots, k$.

A consequence of this proposition is that all proofs need only consider the generating set, $Br(\mathbf{X})$ instead of the full set $L(\mathbf{X})$. Henceforth, the focus shall be on $Br(\mathbf{X})$ only.

Connection to Vector Fields. A *family of vector fields* on the manifold Q is a subset $\mathcal{V} \subset \mathcal{X}(Q)$. The family may be used to define a distribution on Q,

$$D_{\mathcal{V}}(x) \equiv \text{span}_{\mathbb{R}} \{ X(x) \mid X \in \mathcal{V} \}. \tag{4.23}$$

The important structure here is the smallest Lie subalgebra of the Lie algebra, $\mathcal{X}(Q)$, which will be the set of vector fields on Q generated by repeated Lie brackets of elements in $\mathcal{V}$. This Lie subalgebra, which will

[1] the minus sign is an artifact of the simple mechanical systems paradigm, and comes from the Euler-Lagrange equations

be used for small-time controllability, is generated by the free Lie algebras previously described. Let $\mathbf{X}$ and $\mathcal{V}$ be in bijective correspondence via the bijection, $\phi : \mathbf{X} \to \mathcal{V}$. Utilizing the tensor algebra $T(\mathbb{R}^{\mathbf{X}})$ and the bijection ϕ, one may define a $\mathbb{R}$-algebra homomorphism from $T(\mathbb{R}^{\mathbf{X}})$ to $\mathcal{X}(Q)$. It is

$$\mathrm{Ev}(\phi) : T(\mathbb{R}^{\mathbf{X}}) \to \mathcal{X}(Q), \quad u_1 \otimes \cdots \otimes u_k \mapsto \phi(u_1) \cdot \ldots \cdot \phi(u_k). \tag{4.24}$$

where the products, $\cdot$, correspond to Jacobi-Lie brackets [85]. The smallest Lie subalgebra of $\mathcal{X}(Q)$ is the image of $\mathrm{Ev}(\phi)|_{L(\mathbf{X})}$, which shall be denoted $\overline{\mathrm{Lie}}(\mathcal{V})$ and is called the *involutive closure of* $\mathcal{V}$.

When evaluated at a point, $q \in Q$, the homomorphism gives a mapping

$$\mathrm{Ev}_q(\phi) : T(\mathbb{R}^{\mathbf{X}}) \to T_qQ, \quad \mathrm{Ev}_q(\phi)(u) = (\mathrm{Ev}(\phi)(u))\,(q).$$

The Lie algebra rank condition (LARC) at q is satisfied if,

$$\mathrm{Ev}_q(\phi)\,(L(\mathbf{X})) = T_qQ,$$

or, equivalently,

$$\mathrm{rank}(\overline{\mathrm{Lie}}(\mathcal{V})) = \dim Q,$$

for $\phi : \mathbf{X} \to \mathcal{V}$, where $\mathcal{V} = \{ f, g_1, \ldots, g_m \}$. LARC is used to determine local accessibility, from which small-time local controllability may follow under additional requirements. The condition still applies if the manifold Q is the vector bundle E.

Properties of and Computations with the Free Lie Algebra. There will be two sets of vector fields that will be placed into 1-1 correspondence with free Lie algebras.

Definition 48 *The bijective mapping* $\phi : \mathbf{X} \to \mathcal{V} = \left\{ \Gamma, Y_1^{\mathrm{lift}}, \ldots, Y_m^{\mathrm{lift}}, Z^{\mathrm{lift}} \right\}$ *maps* X_0 *to* Γ, X_a *to* Y_a^{lift} *for* $a = 1 \ldots m$, *and* X_{m+1} *to* Z^{lift}.

Definition 49 *The bijective mapping* $\phi' : \mathbf{X}' \to \mathcal{V}' = \left\{ X_S, Y_1^{\mathrm{lift}}, \ldots, Y_m^{\mathrm{lift}} \right\}$ *maps* X'_o *to* X_S *and* X_a *to* Y_a^{lift}, *for* $a = 1 \ldots m$.

The bijection ϕ' is obtained by placing the set $\mathbf{X}'$ into 1-1 correspondence with the vector fields appearing in Equation (4.18). This is contrasted to the bijection ϕ, which decomposes the drift term according to Equation (4.20) and considers the potential $Z^{\mathrm{lift}} \in \mathcal{P}_{-1}$ to be an input. The $\mathcal{P}_{-1}$ term contributing to the drift is removed and used as an input vector field because (a) it may contribute to controllability, and (b) it has the same homogeneous order as the control vector fields. Controllability analysis will be done with the free Lie algebra generated by $\mathbf{X}$, then the analysis will be translated to the free Lie algebra generated by $\mathbf{X}'$. The link between the two sets, $\mathbf{X}$ and $\mathbf{X}'$, will be discussed shortly.

Define the following notation,

$$Br^k(\mathbf{X}) \equiv \{ B \in Br(\mathbf{X}) \mid \delta(B) = k \}, \text{ and} \tag{4.25}$$

$$Br_k(\mathbf{X}) \equiv \left\{ B \in Br(\mathbf{X}) \mid \tilde{\delta}(B) = k \right\}. \tag{4.26}$$

Thus, $Br^k(\mathbf{X})$ is the set of elements in $Br(\mathbf{X})$ whose degree is k. The degree of an element will determine how many iterated Jacobi-Lie brackets it consists of under the evaluation map $\mathrm{Ev}(\cdot)$. This is compared to $Br_k(\mathbf{X})$, which is the set of elements in $Br(\mathbf{X})$ whose relative degree is k. Relative degree will correspond to the homogeneous order under the evaluation map $\mathrm{Ev}(\cdot)$.

Definition 50 *[85] Let $B \in Br_0(\mathbf{X}) \bigcup Br_{-1}(\mathbf{X})$ and let $B_1,\ B_2,\ B_{11},\ B_{12},\ B_{21},\ B_{22},\ \ldots$ be the component decomposition of B. We shall say that B is primitive if each of its components is in $Br_{-1}(\mathbf{X}) \bigcup Br_0(\mathbf{X}) \bigcup \{\, X_0 \,\}$.*

Some observations may be made concerning primitive brackets.

Observation 1 *[85] If $B \in Br_{-1}$ is primitive, then, up to sign, we may write $B = [B_1, B_2]$ with $B_1 \in Br_{-1}(\mathbf{X})$ and $B_2 \in Br_0(\mathbf{X})$.*

Observation 2 *[85] If $B \in Br_0(\mathbf{X})$ is primitive, then, up to sign, B may have one of two forms. Either $B = [X_0, B_1]$ with $B_1 \in Br_{-1}(\mathbf{X})$ primitive, or $B = [B_1, B_2]$ with $B_1, B_2 \in Br_0(\mathbf{X})$ primitive.*

Lemma 5 *[85] Let us impose the condition on elements of $Br(\mathbf{X})$ that we shall consider a bracket to be zero if any of its components are in $Br_{-j}(\mathbf{X})$ for $j \geq 2$. Let $B \in Br_0(\mathbf{X}) \bigcup Br_{-1}(\mathbf{X})$. Then B is the finite sum of primitive brackets.*

Observation 3 *If $B \in Br_0(\mathbf{X})$, then B may be written as*

$$[X_0, B_1]\,, \tag{4.27}$$

with $B_1 \in Br_{-1}(\mathbf{X})$, or as the sum of repeated brackets of such forms, i.e., elements that look like

$$[C_i, [C_{i-1}, [\ldots, [C_2, C_1]\ldots]]]$$

where the $C_i \in Br_0(\mathbf{X})$ are, up to sign, of the form $[X_0, C_{i1}]$ with $C_{i1} \in Br_{-1}(\mathbf{X})$.

proof
Characterizing the set $Br_0(\mathbf{X})$ is difficult since it contains elements that have arbitrary orders of products. By examining the intersection of $Br_0(\mathbf{X})$ with $Br^k(\mathbf{X})$ for individual values of k, we may find some useful structure. Consequently, the proof is by induction on k using degrees, i.e., using elements $B \in Br_0(\mathbf{X}) \bigcap Br^k(\mathbf{X})$. For $k = 1$, there are no elements in $Br_0(\mathbf{X}) \bigcap Br^1(\mathbf{X})$, by definition of relative degree, Equation (4.22). For the case $k = 2$, there are only two equivalent ways to express elements in $Br_0(\mathbf{X}) \bigcap Br^2(\mathbf{X})$,

$$[X_a, X_0] \text{ and } [X_0, X_a]\,,$$

for each element of $X_a \in \mathbf{X}$, $a > 0$. In this case, the statement trivially holds.

For clarity, we analyze the next two levels before moving to the inductive case. For $k = 3$, it is not possible to have an element of $Br_0(\mathbf{X}) \bigcap Br^3(\mathbf{X})$. In fact, analysis of Equation (4.22) shows that all odd values, e.g., $Br^{2l+1}(\mathbf{X})$ for $l \in \mathbb{Z}^+$, cannot contribute to $Br_0(\mathbf{X})$. For $k = 4$, there are a two basic possibilities: $[[X_0, X_a]\,, [X_0, X_b]]$, $[X_0, [X_a, [X_0, X_b]]]$, which satisfy the necessary criteria. All others can be transformed into these cases by taking advantage of Jacobi-Lie bracket identities. For example, using the Jacobi-Lie identity and skew-symmetry, the bracket

$$[X_a, [X_0, [X_0, X_b]]]$$

can be expressed as

$$[[X_0, X_b]\,, [X_0, X_a]] + [X_0, [X_a, [X_0, X_b]]]\,.$$

Hypothesize that the statement holds for the k^{th}-case, and must be proven for the $(\mathrm{k}+1)^{th}$-case. Take the element $B \in Br_0(\mathbf{X}) \bigcap Br^{k+1}(\mathbf{X})$, and decompose it according to Observation 2. There will be two options:

(i) $B = [X_0, B_1]$, with $B_1 \in Br_{-1}(\mathbf{X})$ primitive.
(ii) $B = [B_1, B_2]$ with $B_1 \in Br_0(\mathbf{X}) \bigcap Br^{i'}(\mathbf{X})$ and $B_2 \in Br_0(\mathbf{X}) \bigcap Br^{j'}(\mathbf{X})$, where $i' + j' = k + 1$. B_1 and B_2 satisfy the inductive hypothesis. There are three distinct possibilities for the decomposition of B_1 and B_2 into components. They are

1. $B_1 = [X_0, B_{12}]$ and $B_2 = [X_0, B_{22}]$.
2. $B_1 = [X_0, B_{12}]$ and B_2 is the sum of elements resembling $[C_i, [C_{i-1}, [\ldots, [C_2, C_1]\ldots]]]$.
3. B_1 and B_2 are the sum of elements $[C_i, [C_{i-1}, [\ldots, [C_2, C_1]\ldots]]]$.

The first two cases satisfy the conclusion of the observation. The third case requires some manipulation. We obtain the sum of elements that resemble

$$\left[[C_i, [C_{i-1}, [\ldots, [C_2, C_1]\ldots]]], \left[\tilde{C}_j, \left[\tilde{C}_{j-1}, \left[\ldots, \left[\tilde{C}_2, \tilde{C}_1\right]\ldots\right]\right]\right]\right],$$

which we rewrite as

$$\left[[C_i, C_{xx}], \tilde{C}\right].$$

Using Jacobi-Lie bracket identities, the bracket is equivalent to

$$-\left[C_{xx}, \left[C_i, \tilde{C}\right]\right] + \left[C_i, \left[C_{xx}, \tilde{C}\right]\right].$$

Note that the original bracket terms have been redistributed to appear more like the desired elements. The C_{xx} element is decomposed as above (e.g. $C_{xx} = [C_{i-1}, C_{yy}]$) and the bracket manipulated once more. The process is repeated until the C_{xx} element is reduced to the basic component C_1, at which point we are done. ■

Recall that the set $\mathbf{X}$, under the evaluation map $\mathrm{Ev}(\phi)$, did not correspond precisely to the set of vector fields that could be found in equation (4.18). The proper set to use is $\mathbf{X}'$ with the bijection ϕ' from Definition 49. To apply the results to control systems of the form in equation (4.18), a method to equate an element of $Br(\mathbf{X}')$ to elements of $Br(\mathbf{X})$ must be given. Prior to the stating the equation, the set $\mathbf{S}(\cdot) \subset Br(\mathbf{X})$ must be defined.

Definition 51 *Suppose that $B' \in Br(\mathbf{X}')$. It is possible to write B' as an $\mathbb{R}$-linear sum of elements in $Br(\mathbf{X})$. These primitive elements uniquely define the set $\mathbf{S}(B')$.*

Lemma 6 *[82] The mapping $\Phi : Br(\mathbf{X}') \to L(\mathbf{X})$ is*

$$\Phi(B') = \sum_{B \in \mathbf{S}(B')} (-1)^{\delta_{m+1}(B)} B, \quad \textit{for } B' \in Br(\mathbf{X}'). \tag{4.28}$$

Intuitively, the element $B' \in Br(\mathbf{X}')$ is converted to $L \in L(\mathbf{X})$ by replacing X_0' with $X_0 - X_{m+1}$ and replacing X_a' with X_a for $a = 1 \ldots m$. Using Proposition 14, L may be decomposed into a $\mathbb{R}$-linear combination of elements $B_i \in Br(\mathbf{X})$. It is precisely these elements B_i which comprise the set $\mathbf{S}(B') \subset Br(\mathbf{X})$. A simple example is the following,

$$\left[X_1', \left[X_0', X_2'\right]\right] \mapsto [X_1, [X_0 - X_{m+1}, X_2]] = [X_1, [X_0, X_2]] - [X_1, [X_{m+1}, X_2]]. \tag{4.29}$$

The two iterated Jacobi-Lie brackets, $[X_1, [X_0, X_2]]$ and $[X_1, [X_{m+1}, X_2]]$, form the set $\mathbf{S}(B')$, where the sign of their contribution is related to the degree $\delta_{m+1}(\cdot)$. In this manner elements of $Br(\mathbf{X}')$ can be regarding as being in the linear span of elements from $Br(\mathbf{X})$.

Free Symmetric algebras

Consider, now, the algebra $A(\mathbf{Y})$, where $\mathbf{Y} = \{ X_1, \ldots, X_{l+1} \}$. Let $J(\mathbf{Y})$ be the two-sided ideal generated by elements of the form $X \cdot Y - Y \cdot X$. The free symmetric algebra, denoted $S(\mathbf{Y})$, generated by $\mathbf{Y}$ is the quotient algebra, $S(\mathbf{Y}) = A(\mathbf{Y})/J(\mathbf{Y})$. Denote by $Pr(\mathbf{Y})$ the subset of $S(\mathbf{Y})$ consisting of symmetric products using elements from $\mathbf{Y}$. The set $Pr(\mathbf{Y})$ generates $S(\mathbf{Y})$ just as $Br(\mathbf{X})$ generated $L(\mathbf{X})$, see Proposition 14.

For each element $P \in Pr(\mathbf{Y})$, one can define the *degree* of P. Let $\gamma_a(X)$ be the number of times that the element X_a, $a = 1, \ldots, l+1$, occurs in P. The *degree* of $P \in Pr(\mathbf{Y})$, denoted $\gamma(P)$, is the sum of the γ's,

$$\gamma(P) = \sum_{a=1}^{l+1} \gamma_a(P). \tag{4.30}$$

The free symmetric algebra will be used to generate a symmetric subalgebra in a manner equivalent to the free Lie algebras in Section 4.2.1, by considering a family of vector fields $\mathcal{Y} \subset \mathcal{X}(M)$. The construction will use the previously defined symmetric product.

$$\langle X_a \ : \ X_b \rangle^{X_0} \equiv [X_b, [X_0, X_a]] \,, \tag{4.31}$$

where $X_0 \in \mathbf{X}$ maps to $\Gamma \in \mathcal{X}(E)$, c.f. Equation (4.20), under ϕ.

Definition 52 *The bijective mapping $\psi : \mathbf{Y} \to \mathcal{Y}$ maps X_a to Y_a^{lift} for $a = 1 \ldots m$, and X_{m+1} to Z^{lift}.*

The bijection ψ may be extended naturally to obtain a map from $Pr(\mathbf{Y})$ to $\mathcal{X}(E)$, which may be extended further via $\mathbb{R}$-linearity to be defined on $S(\mathbf{Y})$. The map that results will be denoted by $\mathrm{Ev}(\psi)$. When evaluated for $P \in S(\mathbf{Y})$, at a point $x \in E$,

$$\mathrm{Ev}_x(\psi)(P) = (\mathrm{Ev}(\psi)(P))\,(x). \tag{4.32}$$

Consequently, the *symmetric closure* may be defined to be the image under the free symmetric algebra,

$$\overline{\mathrm{Sym}}(\mathcal{Y}) \equiv \mathrm{Im}\,(\mathrm{Ev}(\psi)(S(\mathbf{Y}))) \,. \tag{4.33}$$

Properties of and Computations with the Free Symmetric Algebra. The set, $\mathbf{Y}$, used to create the free symmetric algebra, $S(\mathbf{Y})$, is really a subset of $\mathbf{X}$, the set used to create the free Lie algebra, $L(\mathbf{X})$. By definition of the product used in the free symmetric algebra, $S(\mathbf{Y})$ is a subset of $L(\mathbf{X})$. Any properties derived for the free symmetric algebra $S(\mathbf{Y})$ translate to the appropriat subset of the free Lie algebra $L(\mathbf{X})$. This is desired as the free symmetric algebra can be used to obtain properties of the control vector fields, which will be useful for controllability computations.

Observation 4 *Suppose that B is a primitive. If $B \in Br_{-1}(\mathbf{X})$, then $B \in S(\mathbf{Y})$.*

proof
The proof is by induction on k using degrees, $Br^k(\mathbf{X})$. For the case $k = 1$, the statement holds trivially. For the case $k = 2$, it is not possible since there are not enough individual degree combinations, $\delta_a(B)$, to result in a -1. For the case $k = 3$, the only options are

$$[X_a, [X_0, X_b]] \,, \ [X_b, [X_0, X_a]] \,,$$

but by definition of the symmetric product, there is really only one option,

$$\langle\, Y_a \;:\; Y_b \,\rangle \in Pr(\mathbf{X}).$$

For the inductive case, where it is assumed to hold for the sets $Br^j(\mathbf{X})$ for $j \leq k$, we will need to carefully consider the decomposition of B into its components.

By Observation 1, choose the case where $B = [B_1, B_2]$, with $B_1 \in Br_{-1}(\mathbf{X})$ and $B_2 \in Br_0(\mathbf{X})$. It must hold that $B_1 \in Br^j(\mathbf{X})$ for $j \leq k$, therefore B_1 may be written as a symmetric product, by the inductive assumption.

What remains to be shown is that $[B_1, B_2]$ will be a symmetric product. There are two possibilities for $B_2 \in Br^i(\mathbf{X}) \bigcap Br_0(\mathbf{X})$, $i \leq k$.

(i) By Observation 2, B_2 can be decomposed into $B_2 = [X_0, B_{21}]$, with $B_{21} \in Br_{-1}(\mathbf{X})$. But by the inductive hypothesis, B_{21} can be expressed as a symmetric product. This gives,

$$B = [B_1, [X_0, B_{21}]] = \langle\, B_{21} \;:\; B_1 \,\rangle\,, \tag{4.34}$$

and the first case is complete.

(ii) It may hold, instead, that $B_2 = [B_{21}, B_{22}]$ with $B_{21}, B_{22} \in Br_0(\mathbf{X})$. Observation 3 gives the sum of iterated brackets of the form

$$[C_i, [C_{i-1}, [\ldots, [C_2, C_1]\ldots]]]$$

where the $C_i \in Br_0(\mathbf{X})$ are of the form $C_i = [X_0, C_{i1}]$ and $C_{i1} \in Br_{-1}(\mathbf{X}) \bigcap Br^j(\mathbf{X})$, for some $j \leq k$. By the inductive hypothesis, $C_{i1} \in Pr(\mathbf{X})$. The Lie bracket for B is

$$B = [B_1, [C_i, [C_{i-1}, [\ldots, [C_2, C_1]\ldots]]]]\,.$$

By the Jacobi-Lie identity,

$$B = -\left[[C_{i-1}, [\ldots, [C_2, C_1]\ldots]]\,, [B_1, C_i]\right] - [C_i, [B_1, [C_{i-1}, [\ldots, [C_2, C_1]\ldots]]]]$$

Note that in the first term, the Lie bracket element with B_1 evaluates to a symmetric product. Thus we recuperate a form like,

$$\left[[C_{i-1}, [C_{i-2}, [\ldots, [C_2, C_1]\ldots]]]\,, B_{1i}\right]$$

where B_{1i} is a symmetric product. The Jacobi identity is used repeatedly until all of the C_i have been integrated into symmetric products.

The second term has the structure of the original B, but at a lower degree and with the extra Lie bracket with C_i. Repeated Jacobi identity applications will also result in symmetric products with the outer brackets, i.e., terms like

$$[C_{i-1}, [C_{i-2}, [\ldots, [C_2, B_{1i}]\ldots]]]\,.$$

The brackets really form iterated symmetric products. There will be other terms like

$$\left[C_{i-1}, \left[C_{i-2}, \left[\ldots, \left[B_{1i_1\ldots i_j}, \ldots [C_2, C_1]\right]\right]\right]\right]$$

but via the analysis above, they also become symmetric products. Since the C_i are of the form $[X_0, C_{i1}]$, where C_{i1} is a symmetric product, the recursive simplification must occur.

■

This observation is analogous to the result in Lemma 5.6 part (i) of [85] for simple mechanical systems. Since the observation states that an element $B \in Br_{-1}(\mathbf{X})$ is a symmetric product, this implies for simple mechanical systems that it must be a lift. Recall that the simple mechanical system symmetric product

operates on the unlifted vector fields and is lifted after all of the symmetric products have been calculated (although this need not be the case).

Free Algebras and Homogeneity

Lemma 7 *Let $l \geq 1$ be an integer and let $B \in Br_l(\mathbf{X})$. Then $\mathrm{Ev}(\phi)(B)(0_q) = 0$ for all $q \in Q$.*

proof
An element $B \in Br_l(\mathbf{X})$ contains brackets with l more instances of the vector field X_0 showing up than the vector fields X_a. Supposing that $B \in Br^{(2m+l)}(\mathbf{X})$ for any $m \geq 0$. Then, there are m instances of the elements X_a, and $m + l$ instances of the element, X_0. Since the X_a map to input vector fields, they lie in the space $\mathcal{P}_{-1}$, and since the element X_0 maps to the drift vector field, it is in $\mathcal{P}_1$. The Lie brackets map into $\mathcal{P}_l$. When evaluated on the zero section, homogeneity of order $l \geq 1$ implies that the result is the zero vector, c.f. Proposition 7.
■

Lemma 8 *Let $l \geq 2$ be an integer, and let $B \in Br_{-l}(\mathbf{X})$. Then $\mathrm{Ev}(\phi)(B) = 0$.*

proof
An element $B \in Br_{-l}(\mathbf{X})$ contains brackets with l more instances of the vector fields X_a showing up than the vector field X_0. Supposing that $B \in Br^{(2m+l)}(\mathbf{X})$ for any $m \geq 0$. Then, there are m instances of the element X_0, and $m + l$ instances of the elements, X_a. The Lie brackets map into $\mathcal{P}_{-l}$. For $l \geq 2$, the space only contains the zero vector field, c.f. Propoition 6.
■
The homogeneous structure of the nonlinear system implies that only the brackets that reside in $\mathcal{M}_{-1,0}$ when evaluated on the zero section of E, will give contributions to the controllable Lie subalgebra. The homogeneous structure also implies that the brackets of $\mathcal{P}_{-1}$ will be vertical elements, and the brackets of $\mathcal{P}_0$ evaluated on the zero section will be horizontal elements, using the trivial connection.

STLC in the Algebraic Sense

In order to determine small-time local controllability using the Lie algebra rank condition, it is necessary to define the notion of a good and a bad bracket.

Definition 53 *An element $B \in Br(\mathbf{X})$ is called* bad *if $\delta_a(B)$ is even for all $a = 1 \ldots m$, and if $\delta_0(B)$ is odd. If B is not bad, then it is* good*.*

In addition, Sussmann utilized the permutation group S_m, of permutations using m symbols. An element $\sigma \in S_m$ takes the m elements $X_a \in \mathbf{X}$, and maps them to $X_{\sigma(a)} \in \mathbf{X}$. The set of all possible permutations can be used to define a bracket permutation, $\beta : Br(\mathbf{X}) \to Br(\mathbf{X})$ on the primitive $B \in Br(\mathbf{X})$.

$$\beta(B) = \sum_{\bar{\sigma} \in S_m} \bar{\sigma}(B)$$

Sussman's general derivation, relying on the standard form of the nonlinear control equation (4.15), culminated in the following theorem.

Theorem 24 *[152] Consider the bijection $\phi : \mathbf{X} \rightarrow \{ f, g_1, \ldots, g_m \}$, which sends X_0 to f and X_a to g_a for $a = 1 \ldots m$. Suppose that (4.15) is such that every bad bracket $B \in Br(\mathbf{X})$ has the property that*

$$\mathrm{Ev}_z(\phi)\left(\beta(B)\right) = \sum_{a=1}^{k} \xi_a \mathrm{Ev}_z(\phi)\left(C_a\right),$$

where C_a are good brackets in $Br(\mathbf{X})$ of lower degree than B and $\xi^a \in \mathbb{R}$ for $a = 1 \ldots k$. Also suppose that (4.15) satisfies the LARC at z. Then (4.15) is STLC at z.

Lewis examined the structure of simple mechanical systems to refine this idea for configuration controllability. In particular, only the brackets in the homogeneous space $Br_{-1}(\mathbf{X})$ can result in bad brackets. As they correspond to symmetric products, the search for bad brackets was reduced to examination of the symmetric products only. Define the set $\mathbf{Y} = \{ X_1, \ldots, X_{m+1} \}$ to be bijective to $\mathcal{Y}$ via $\psi : \mathbf{Y} \rightarrow \mathcal{Y}$.

Definition 54 *An element $P \in Pr(\mathbf{Y})$ is called* bad *if $\gamma_a(P)$ is even for each $a = 1 \ldots m$. If P is not bad then it is* good*.*

Theorem 25 *[85] Consider the bijection $\psi : \mathbf{Y} \rightarrow \mathcal{Y}$, which sends X_a to Y_a^{lift} for $a = 1 \ldots m$, and X_{l+1} to Z^{lift}. Suppose that (4.18) is such that every bad symmetric product in $P \in Pr(\mathbf{Y})$ has the property that*

$$\mathrm{Ev}_{0_x}(\psi)(\rho(P)) = \sum_{a=1}^{k} \xi_a \mathrm{Ev}_{0_x}(\psi)(C_a), \tag{4.35}$$

where C_a are good symmetric products in $Pr(\mathbf{Y})$ of lower degree than P and $\xi_a \in \mathbb{R}$ for $a = 1, \ldots, k$. Also, suppose that (4.18) is locally configuration accessible at x. Then (4.18) is STLCC at x.

This theorem was proven for the case that $E = TQ$, and is precisely the theorem that will be recreated for a 1-homogeneous control system evolving on a generic vector bundle[2]. Configuration controllability will imply controllability of the base space, M, of the vector bundle, E.

The Form of the Accessibility Distribution on the Zero Section for 1-Homogeneous Control Systems.

Lemma 9 *Let $x \in M$. Then,*

$$D_{\overline{Lie}(\mathcal{V})}(0_x) \bigcap V_{0_x} E = D_{\overline{Sym}(\mathcal{Y})}(0_x),$$

and

$$D_{\overline{Lie}(\mathcal{V})}(0_x) \bigcap H_{0_x} E = D_{\overline{Lie}\left(\left[\Gamma, \overline{Sym}(\mathcal{Y})\right]\right)}(0_x).$$

proof
The proof mirrors that of Lewis and Murray [85], however, the main steps of the proof will be sketched. Propositions 5 - 7 imply that only the primitive bracket contributions will be essential to this analysis. Of the primitive brackets, the only contributions to the Lie closure are those elements that lie in $Br_{-1}(\mathbf{X})$ and $Br_0(\mathbf{X})$, which correspond to elements of the vertical and horizontal subspaces, respectively (using the trivial decomposition). The space $Br_{-1}(\mathbf{X})$ consists of only symmetric products, via Observation 4, and likewise the space $Br_0(\mathbf{X})$ is generated through the Lie closure of the set $[\Gamma, Br_{-1}(\mathbf{X})]$, using Observation 3.
■

[2]The notation of this theorem is a bit different. The $\mathcal{Y}$ in the theorem statement corresponds to $\mathcal{Y} \bigcup \{ \mathrm{grad} V \}$ from Lewis and Murray [85].

Lemma 9 allows us to compare the involutive closures $\overline{\text{Sym}}(\mathcal{Y})$ and $\overline{\text{Lie}}\left(\left[\Gamma, \overline{\text{Sym}}(\mathcal{Y})\right]\right)$ evaluated on the zero section with the horizontal and vertical distributions of the trivial connection. Since $\overline{\text{Sym}}(\mathcal{Y})$ and $\overline{\text{Lie}}\left(\left[\Gamma, \overline{\text{Sym}}(\mathcal{Y})\right]\right)$ are generated by the evaluation of $Br_{-1}(\mathbf{X})$ and $Br_0(\mathbf{X})$, respectively, we may define horizontal and vertical distributions using the free Lie algebras.

Definition 55 *The horizontal distribution corresponding to $Br^k(\mathbf{X})$ is given by*

$$\mathcal{C}^{(k)}_{hor}(\mathcal{V}) \equiv \left\{ \text{Ev}(\phi)(B) \mid B \in Br^k(\mathbf{X}) \bigcap Br_0(\mathbf{X}) \right\}.$$

The vertical distribution corresponding to $Br^k(\mathbf{X})$ is given by

$$\mathcal{C}^{(k)}_{ver}(\mathcal{V}) \equiv \left\{ \text{Ev}(\phi)(B) \mid B \in Br^k(\mathbf{X}) \bigcap Br_{-1}(\mathbf{X}) \right\}.$$

The union of these spaces over all $k \in \mathbb{Z}^+$ may be used to define the horizontal and vertical distributions.

Definition 56 *The horizontal distribution is defined to be,*

$$\mathcal{C}_{hor}(\mathcal{V}) = \bigcup_{k \in \mathbb{Z}^+} \mathcal{C}^{(k)}_{hor}(\mathcal{V}),$$

and the vertical distribution is defined to be,

$$\mathcal{C}_{ver}(\mathcal{V}) = \bigcup_{k \in \mathbb{Z}^+} \mathcal{C}^{(k)}_{ver}(\mathcal{V}).$$

Certainly in the absence of a potential energy term, $Z^{\text{lift}} = 0$, in the drift vector field, X_S, these definitions and lemma suffice to characterize the Lie algebra of the control system (4.18). Unfortunately, in the presence of a nonvanishing potential, additional work is needed to demonstrate that the Lie algebra $L(\mathbf{X}')$ is represented by $L(\mathbf{X})$, or equivalently $Br(\mathbf{X}')$ is represented by $Br(\mathbf{X})$. This means that the brackets in $Br(\mathbf{X})$ neeed to be equated with brackets in $Br(\mathbf{X}')$, which is done via the replacement operator,

Definition 57 *The replacement operator,* $\text{Rep} : Br(\mathbf{X}) \to Br(\mathbf{X}')$*, operates as follows: given an element $B \in Br(\mathbf{X})$, all instances of $X_0 \in \mathbf{X}$ and $X_{m+1} \in \mathbf{X}$ are replaced with $X'_0 \in \mathbf{X}'$, and all instances of X_a are replaced with X'_a for $a = 1 \ldots m$.*

The term X_{m+1} evaluates to the vector field Z^{lift} which cannot occur on its own, it is part of the drift vector field. The replacement map respects this fundamental fact. The problem with the replacement map is that we have fully characterized the sets $Br(\mathbf{X})$, but not $Br(\mathbf{X}')$, therefore we will need to map the replaced bracket into $L(\mathbf{X})$. This is done by the operator Φ. Define the function $\widetilde{\Phi}$ to map $Br(\mathbf{X})$ into $Br(\mathbf{X}')$, then finally into $L(\mathbf{X})$,

$$\widetilde{\Phi} : Br(\mathbf{X}) \to L(\mathbf{X}), \quad \widetilde{\Phi} \equiv \Phi \circ \text{Rep}. \tag{4.36}$$

Definition 58 *The horizontal distribution corresponding to $Br^k(\mathbf{X}')$ is given by*

$$\mathcal{C}^{(k)}_{hor}(\mathcal{V}') = \left\{ \text{Ev}(\phi)(\widetilde{B}) \mid \widetilde{B} = \widetilde{\Phi}(B),\ \forall\ B \in Br^k(\mathbf{X}) \bigcap Br_0(\mathbf{X}) \right\}.$$

The vertical distribution corresponding to $Br^k(\mathbf{X}')$ is given by

$$\mathcal{C}^{(k)}_{ver}(\mathcal{V}') = \left\{ \text{Ev}(\phi)(\widetilde{B}) \mid \widetilde{B} = \widetilde{\Phi}(B),\ \forall\ B \in Br^k(\mathbf{X}) \bigcap Br_{-1}(\mathbf{X}) \right\}.$$

proof
What will be demonstrated is that not only does the set $\mathbf{X}'$ allow for horizontal and vertical decompositions as did $\mathbf{X}$, but that they in fact coincide. Take an arbitrary primitive bracket $B \in Br^{(k)}(\mathbf{X})$ and calculate its replacement $B' = \mathrm{Rep}(B) \in Br^{(k)}(\mathbf{X}')$, whereby the following identities are obtained: $I' = \delta(B')$, and $J' = \sum_{a=1}^{m} \delta_a(B')$. It turns out that $\delta_0(B') = \delta_0(B) + \delta_{m+1}(B)$. The mapping Φ takes B' and maps it into a sum of brackets in $Br(\mathbf{X})$. Those brackets that linearly combine to give B' form the set $\mathbf{S}(B')$.

We shall characterize an arbitrary element $C \in \mathbf{S}(B')$. The element, C, will have the following degrees, $I = \delta_0(C)$, $J = \sum_{a=1}^{m} \delta_a(C) = J'$, and $K = \delta_{m+1}(C) = I' - I$. Its relative degree will be $\tilde{\delta}(C) = I - J - K$. The only contributions, when evaluated on the zero section will be those with relative degree 0 or -1. There is only one way to contribute in the set $\mathbf{S}(B')$, and it will depend on the parity of the relative degree of the original B', $\tilde{\delta}(B') = I' - J'$.

Take the element with maximal degree $B \in \mathbf{S}(B')$. There can only be one such element. It is the one with $I = I'$. The next highest degree are the terms with $I = I' - 1$. Notice that, if any one of the slots with X_0 in it is replaced by X_{m+1}, then the relative degree changes by two degrees; one degree is lost because X_0 is removed, but one degree is also lost because X_{m+1} is added. Since the degree drops by two degrees each time we move from I to $I - 1$, the relative degree of elements from $\mathbf{S}(B')$ have the same parity, they will be even only, or odd only. For an even parity, the only contributions can be $B \in \mathbf{S}(B') \bigcap Br_0(\mathbf{X})$, whereas for odd parity, the only contributions can be $B \in \mathbf{S}(B') \bigcap Br_{-1}(\mathbf{X})$. But these precisely correspond to the contributions found in $\mathcal{C}_{hor}^{(k)}(\mathcal{V})$ and $\mathcal{C}_{ver}^{(k)}(\mathcal{V})$, respectively, from Definition 55, meaning that the distributions are still horizontal and vertical, accordingly. The only contributing elements turn out to be those that have the same relative degree as the original $B \in Br_0(\mathbf{X}) \bigcup Br_{-1}(\mathbf{X})$.

Elements with the Jacobi-Lie bracket $[X_0, X_{m+1}]$ require special consideration, since X_{m+1} is not an input that we have full control over. Fortunately, under the replacement operator Rep, this bracket becomes $[X_0', X_0'] = 0$, and has no contribution under $\widetilde{\Phi}$. Thus, such primitive elements of $Br(\mathbf{X})$ can be ignored.
■

It should be noted that the proof of the above definition replaces the algorithm found in Lewis and Murray [85]. This deviation was preferred because it will be useful to understand the above proof when dissipation is considered , as will be done in Section 4.3. We may now define the horizontal and vertical distributions by taking the union over all $k \in \mathbb{Z}^+$.

Definition 59 *The horizontal distribution is defined to be,*

$$\mathcal{C}_{hor}(\mathcal{V}') = \bigcup_{k \in \mathbb{Z}^+} \mathcal{C}_{hor}^{(k)}(\mathcal{V}'),$$

and the vertical distribution is defined to be,

$$\mathcal{C}_{ver}(\mathcal{V}') = \bigcup_{k \in \mathbb{Z}^+} \mathcal{C}_{ver}^{(k)}(\mathcal{V}').$$

All of the fundamental work to prove the following statement has been completed.

Proposition 15 *Let $x \in M$. Then*

$$D_{\overline{Lie}(\mathcal{V}')}(0_x) \bigcap H_{0_x} E = \mathcal{C}_{hor}(\mathcal{V}')(0_x),$$

and

$$D_{\overline{Lie}(\mathcal{V}')}(0_x) \bigcap V_{0_x} E = \mathcal{C}_{ver}(\mathcal{V}')(0_x).$$

4.2.2 Conditions for Configuration Controllability

Up to this point, we have given and proven the analogous theorems derived in [85] for use within the context of 1-homogeneous control sytems. Consequently, all of the theorems concerning configuration controllability may now hold for these homogeneous control sytems without modification of the proofs from Lewis and Murray [85].

Theorem 26 *[85] The control system (4.18) is locally configuration accessible at $x \in M$ if $\mathcal{C}_{hor}(\mathcal{V})(0_x) = H_{0_x}E$.*

The partial converse proven for the case when there is no potential also holds.

Theorem 27 *[85] Suppose that $Z^{\text{lift}} = 0$ and (4.18) is locally configuration accessible. Then $\mathcal{C}_{hor}(\mathcal{V})(0_q) = H_{0_q}E$ for Q in an open dense subset of Q.*

Theorem 28 *[85] Suppose that $\mathcal{Y}$ is such that every bad symmetric product in $Pr(\mathbf{Y})$ has the property that*

$$\text{Ev}_{0_x}(\psi)(\rho(P)) = \sum_{a=1}^{m} \xi_a \text{Ev}_{0_x}(\psi)(C_a),$$

where C_a are good symmetric products in $Pr(\mathbf{Y})$ of lower degree that P and $\xi_a \in \mathbb{R}$ for $a = 1 \dots m$. Also, suppose that (4.18) is locally configuration accessible at $x \in M$. Then (4.18) is STLCC at x.

Corollary 14 *[85] Suppose that the hypotheses of Theorem 28 hold for all $x \in M$. Then the system (4.18) is equilibrium controllable.*

4.2.3 Conditions for Small-Time Local Controllability

It may be the case that configuration controllability will not suffice for some systems, and actual controllability is sought. Small-time local controllability for elements in the equilibrium subspace can also be given. Recall that to be an element of the equilibrium subspace, according to Definition 44, the element has to lie in the zero section of E.

Theorem 29 *[85] The control system (4.18) is locally accessible at $0_x \in E$ if $\mathcal{C}_{hor}(\mathcal{V})(0_x) = H_{0_x}E$ and $\mathcal{C}_{ver}(\mathcal{V})(0_x) = V_{0_x}E$.*

Theorem 30 *[85] Suppose that $\mathcal{Y}$ is such that every bad symmetric product in $Pr(\mathbf{Y})$ has the property that*

$$\text{Ev}_{0_x}(\psi)(\rho(P)) = \sum_{a=1}^{m} \xi_a \text{Ev}_{0_x}(\psi)(C_a),$$

where C_a are good symmetric products in $Pr(\mathbf{Y})$ of lower degree that P and $\xi_a \in \mathbb{R}$ for $a = 1 \dots m$. Also, suppose that (4.18) is locally accessible at $0_x \in E$. Then (4.18) is STLC at 0_x.

4.3 Dissipation and Controllability

This section extends the controllability analysis to the case where dissipation is included in the mechanical model. The dissipation modelled is often known as *Rayleigh dissipation*; as a vector field it is vertical and of homogeneous order 0 (implying linearity as a function of the vector fiber).

$$\dot{z} = X_S(z) + Y_a^{\text{lift}}(z)u^a = \Gamma(z) + D^{\text{lift}}(z) - Z^{\text{lift}}(z) + Y_a^{\text{lift}}(z)u^a. \tag{4.37}$$

The vector field $D^{\text{lift}}(z) \in \mathcal{P}_0$ represents the additional dissipation vector field. Cortés et al. [34] studied the effect of isotropic dissipation; more precisely $D^{\text{lift}}(x)$ was restricted to be symmetric and independent of configuration. This section attempts to generalize [34] to configuration dependent, not necessarily symmetric, dissipation. The dissipation vector field is seen as an interfering element in the controllability analysis, and not as a control vector field that may grant additional control authority. Although the free Lie algebras will recover additional control directions, a large set of potential Jacobi-Lie brackets will be ignored by restricting the analysis to the homogeneous spaces that were found to be useful in Section 4.2. Consequently, the results of this section are conservative in nature.

Free Algebras and Homogeneity

Lemma 10 *If* $B \in Br_{-1}(\mathbf{X})$, *then* $\text{Ev}(\phi)(B) \in \mathcal{M}_{-1}$.

proof
If $B \in Br_{-1}(\mathbf{X})$ then $B \in Pr(\mathbf{X})$. Evaluation of the symmetric product gives a vector field in the homogeneous space $\mathcal{M}_{-k-1,-1} = \mathcal{P}_{-1}$.
■
The addition of dissipation, a vertical vector field of homogeneous order 0, does not alter the symmetric product calculation, which can be seen by appealing to the definition of the symetric product, Definition 37. Consequently, the nature of good and bad brackets does not change.

Lemma 11 *Let* $l \geq 2$ *be an integer, and let* $B \in Br^k(\mathbf{X}) \bigcap Br_{-l}(\mathbf{X})$ *for* $k \geq 2$. *Then* $\text{Ev}(\phi)(B) = 0$.

proof
An element $B \in Br_{-l}(\mathbf{X})$ contains brackets with l more instances of the vector fields X_a, $a = 1 \ldots m+1$, than the vector field X_0. Supposing that $B \in Br^{(2k+l)}(\mathbf{X})$ for any $k > 0$. Then, there are k instances of the element X_0, and $k + l$ instances of the elements X_a in the products that make up B. Under the evaluation operator, $\text{Ev}(\phi)$, the X_a map to input vector fields lying in the space $\mathcal{P}_{-1}$, and the element X_0 maps to the drift vector field in $\mathcal{M}_{0,1}$. Therefore, the element B maps to an iterated Jacobi-Lie bracket in the space $\mathcal{M}_{-k-l,-l}$. For $l \geq 2$, the space contains only the zero vector field.
■
The remaining elements requiring analysis for configuration accessability are those in those in $Br_l(\mathbf{X})$, for $l \geq 0$.

Lemma 12 *Let* $l \geq 0$ *be an integer and let* $B \in Br_l(\mathbf{X})$. *Then* $\text{Ev}_{0_q}(\phi)(B) \in \mathcal{M}_{-1,0}(0_q)$ *for all* $q \in Q$.

proof
An element $B \in Br_l(\mathbf{X})$ contains brackets with l more instances of the vector field X_0 than the vector fields X_a, $a = 1 \ldots m+1$. Supposing that $B \in Br^{(2k+l)}(\mathbf{X})$ for any $k > 0$, there are k instances of the elements X_a, and $k + l$ instances of the element, X_0. Under the evaluation operator, $\text{Ev}(\phi)$, the X_a map to input vector fields lying in the space $\mathcal{P}_{-1}$, and the element X_0 maps to $\Gamma + D^{\text{lift}} \in \mathcal{M}_{0,1}$. The element B maps to a iterated Jacobi-Lie bracket in the space $\mathcal{M}_{-k,l} = \mathcal{M}_{-1,l}$. Of course, all vector fields of homogeneous order greater than 0 vanish on the zero section, therefore only the subspace $\mathcal{M}_{-1,0}$ will provide a contribution.
■
Previously, $\text{Ev}(\phi)(Br_0(\mathbf{X})) \subset \mathcal{P}_0$, whereas now $\text{Ev}(\phi)(Br_0(\mathbf{X})) \subset \mathcal{M}_0$. The vertical dissipation vector, introduces new vertical contributions to the Jacobi-Lie bracket computations. By projecting to $\text{Ev}(\phi)(Br_0(\mathbf{X}))$ to the horizontal subspace, we may recover the original (dissipation free) distribution. Equally important, the evaluation $\text{Ev}(\phi)(Br_k(\mathbf{X}))$ is no longer the trivial space $\{0\}$ for $k \geq 0$. Dissipation introduces additional brackets that before had been vanishing. These brackets can provide either new control directions

if needed, or new bad brackets. The bad brackets will only occur for spaces of odd homogeneous order, $Br_{2k+1}(\mathbf{X})$ for $k \geq 0$.

The Form of the Accessiblity Distribution on the Zero Section for 1-Homogeneous Control Systems with Dissipation. The horizontal and vertical decomposition is no longer viable. We shall consider, instead, several distributions. The first consists of only those element that were previously found useful on the zero section.

$$\mathcal{C}_0(\mathcal{V}) = \bigcup_{k \in \mathbb{Z}^+} \mathcal{C}_0^{(k)}(\mathcal{V}), \tag{4.38}$$

where

$$\mathcal{C}_0^{(k)}(\mathcal{V}) \equiv \left\{ \mathrm{Ev}(\phi)(B) \mid B \in Br^k(\mathbf{X}) \bigcap Br_0(\mathbf{X}) \right\}. \tag{4.39}$$

In the case of vanishing dissipation, $\mathcal{C}_0(\mathcal{V})$ is equal to $\mathcal{C}_{hor}(\mathcal{V})$. The remaining two distributions will be used to characterize the good brackets and the bad brackets of $Br(\mathbf{X})$.

$$\mathcal{C}_{\text{even}}(\mathcal{V}) = \bigcup_{k \in \mathbb{Z}^+} \mathcal{C}_{\text{even}}^{(k)}(\mathcal{V}), \tag{4.40}$$

where

$$\mathcal{C}_{\text{even}}^{(k)}(\mathcal{V}) \equiv \left\{ \mathrm{Ev}(\phi)(B) \mid B \in Br^k(\mathbf{X}) \bigcap Br_{2l}(\mathbf{X}),\ \ l \geq 0 \right\}. \tag{4.41}$$

Under the mapping $\widetilde{\Phi}$, we will find that $\mathcal{C}_0(\mathcal{V})$ maps into $\mathcal{C}_{\text{even}}(\mathcal{V})$. Comparing this to the case of no dissipation, where $\widetilde{\Phi}\left(\mathcal{C}_0(\mathcal{V})\right) = \mathcal{C}_0(\mathcal{V})$, we see that dissipation enlarges the set of possible contributions to configuration controllability. The same can be said of $\mathcal{C}_{\text{odd}}(\mathcal{V})$,

$$\mathcal{C}_{\text{odd}}(\mathcal{V}) = \bigcup_{k \in \mathbb{Z}^+} \mathcal{C}_{\text{odd}}^{(k)}(\mathcal{V}), \tag{4.42}$$

where

$$\mathcal{C}_{\text{odd}}^{(k)}(\mathcal{V}) \equiv \left\{ \mathrm{Ev}(\phi)(B) \mid B \in Br^k(\mathbf{X}) \bigcap Br_{2l-1}(\mathbf{X}),\ \ l \geq 0 \right\}. \tag{4.43}$$

In the case of no dissipation, $\widetilde{\Phi}\left(\mathcal{C}_{ver}(\mathcal{V})\right) = \mathcal{C}_{ver}(\mathcal{V})$, whereas in the case with dissipation, $\widetilde{\Phi}\left(\mathcal{C}_{ver}(\mathcal{V})\right) \subset \mathcal{C}_{\text{odd}}(\mathcal{V})$. Consequently, dissipation also enlarges the space of potential bad brackets. The nice thing is that the bad brackets are in a complementary subdistribution to the good brackets.

It is critical to understand how these brackets relate to the actual system, which utilizes the set $\mathcal{V}'$ instead.

Definition 60 *Define the distribution*

$$\mathcal{C}_0(\mathcal{V}') = \bigcup_{k \in \mathbb{Z}^+} \mathcal{C}_0^{(k)}(\mathcal{V}'), \tag{4.44}$$

which consists of

$$\mathcal{C}_0^{(k)}(\mathcal{V}') \equiv \left\{ \mathrm{Ev}(\phi)(\widetilde{B}) \mid \widetilde{B} = \widetilde{\Phi}(B), B \in Br^k(\mathbf{X}) \bigcap Br_0(\mathbf{X}) \right\}. \tag{4.45}$$

Proposition 16 $\mathcal{C}_0(\mathcal{V}') \subset \mathcal{C}_{even}(\mathcal{V})$.

proof

Take an arbitrary primitive bracket $B \in Br^{(k)}(\mathbf{X}) \bigcap Br_0(\mathbf{X})$, and calculate its replacement $B' = \mathrm{Rep}(B) \in Br^{(k)}(\mathbf{X}')$, whereby the following identities are obtained: $I' = \delta(B')$, and $J' = \sum_{a=1}^{m} \delta_a(B')$, where

$\delta_0(B') = \delta_0(B) + \delta_{m+1}(B)$. The mapping Φ takes B' and maps it into a sum of brackets from the set $\mathbf{S}(B') \subset Br(\mathbf{X})$.

We shall characterize an arbitrary element $C \in \mathbf{S}(B')$. The element, C, will have the following degrees, $I = \delta_0(C)$, $J = \sum_{a=1}^{m} \delta_a(C) = J'$, and $K = \delta_{m+1}(C) = I' - I$. Its relative degree will be $\tilde{\delta}(C) = I - J - K$. Whereas before, the only contributions on the zero section were those with relative degree 0 or -1, contributions may be obtained for any relative degree greater than or equal to -1.

Take the element with maximal degree $B \in \mathbf{S}(B')$. There can only be one such element. It is the one with $I = I'$. The next highest degree terms are the terms with $I = I' - 1$. Notice that, if any one of the slots with X_0 in it is replaced by X_{m+1}, then the relative degree changes by two degrees; one degree is lost because X_0 is removed, but one degree is also lost because X_{m+1} is added. Since the degree drops by two degrees each time we move from I to $I-1$, the relative degree of elements from $\mathbf{S}(B')$ have the same parity, they will be even only, or odd only. For an even parity, the only contributions can be $B \in \mathbf{S}(B') \bigcap Br_{2l}(\mathbf{X})$, whereas for odd parity, the only contributions can be $B \in \mathbf{S}(B') \bigcap Br_{2l-1}(\mathbf{X})$, for $l \geq 0$. Since the original parity was even, B' is spanned by elements in $Br_{2k}(\mathbf{X})$ for $k \geq 0$.

As before, elements with the Jacobi-Lie bracket $[X_0, X_{m+1}]$ vanish under the replacement operator.

■

The distribution $\mathcal{C}(\mathcal{V}')$ does not contain brackets of odd relative degree, so it can never contain bad brackets. According to Lemma 12, the distribution $\mathcal{C}(\mathcal{V}')$ will contain vertical contributions. Given that the concern of configuration controllability is only on the horizontal subdistribution, we shall define the projection of $\mathcal{C}(\mathcal{V})$ to the horizontal subdistribution by

$$\mathcal{PC}(\mathcal{V}) = \left(\mathrm{Id} - A^{\mathrm{triv}}\right)(\mathcal{C}(\mathcal{V}))\,. \tag{4.46}$$

Definition 61 *Define the distribution*

$$\mathcal{C}_{-1}(\mathcal{V}') = \bigcup_{k \in \mathbb{Z}^+} \mathcal{C}_{-1}^{(k)}(\mathcal{V}'), \tag{4.47}$$

which consists of

$$\mathcal{C}_{-1}^{(k)}(\mathcal{V}') \equiv \left\{ \mathrm{Ev}(\phi)(\widetilde{B}) \mid \widetilde{B} = \widetilde{\Phi}(B), B \in Br^k(\mathbf{X}) \bigcap Br_{-1}(\mathbf{X}) \right\}. \tag{4.48}$$

Proposition 17 $\mathcal{C}_{-1}(\mathcal{V}') \subset \mathcal{C}_{odd}(\mathcal{V})$

proof

Same as the proof in Proposition 16, except that now the parity is odd.

■

This proposition is used to characterize the space of bad brackets obtained from symmetric products. Just as the space of useful brackets was enlarged, the space of bad brackets is bigger, requiring more work to check configuration controllability. In the case of no potential, $Z^{\mathrm{lift}} = 0$, we can improve our understanding. This is because the mapping $\widetilde{\Phi}$ is no longer needed (it is the identity mapping since the sets $\mathbf{X}$ and $\mathbf{X}'$ map to the same objects under $\mathrm{Ev}(\phi)$), and all of our analysis reduces to understanding the set $Br(\mathbf{X})$.

Proposition 18 *Assume that there does not exist a potential term in the drift vector field, i.e., $Z^{\mathrm{lift}} = 0$. If the space $\mathcal{C}(\mathcal{V}')$ is calculated on the zero section for equation (4.37) according to the algorithm, then its horizontal projection is equal to $\mathcal{C}_{hor}(\mathcal{V}')$, which is calculated on the zero section using equation (4.18), i.e., with the dissipation removed. Equivalently,*

$$\mathcal{PC}(\mathcal{V}')(0_x) = \mathcal{C}_{hor}(\mathcal{V}')(0_x). \tag{4.49}$$

proof
In the case of no potential term, $\mathcal{C}(\mathcal{V}') = \mathcal{C}(\mathcal{V})$. The proof in Lemma 12 reveals that $\mathcal{C}(\mathcal{V}) \subset \mathcal{M}_{-1,0}$ on the zero section. Further analysis of the proof demonstrates that the $\mathcal{P}_{-1}$ terms come from the dissipation vector field only, and there is only one possible bracket that can lie in $\mathcal{P}_0$. This unique bracket is the same bracket that appears in the $\mathcal{C}_{hor}(\mathcal{V}')$ computations for the case without dissipation. Hence, the horizontal projection annihilates the new contributions due to dissipation and preserves the original contribution.
■
This proposition implies that the addition of dissipation does not affect the original configuration controllability calculation in the case that there is no potential term in the drift vector field. When restricted to isotropic dissipation, this replicates the conclusion of Cortés et al. [34].

4.3.1 Conditions for Configuration Controllability with Dissipation

Theorem 31 *[85] The control system (4.18) is locally configuration accessible at $x \in M$ if $\mathcal{PC}(\mathcal{V}')(0_x) = H_{0_x}E$.*

proof
The distribution $\mathcal{C}(\mathcal{V}')(0_x) \subset \mathcal{M}_{-1,0}$ due to geometric homogeneity, c.f. Lemma 12. Do I need this proof?
■

Theorem 32 *[85] Suppose that $\mathbf{X}$ is such that every bad bracket in $B \in \Phi\left(Br_{2l-1}(\mathbf{X})\right)$, for $l \geq 0$, has the property that*

$$\mathrm{Ev}_{0_x}(\psi)(\psi(B)) = \sum_{a=1}^{m} \xi_a \mathrm{Ev}_{0_x}(\psi)(C_a)$$

where C_a are good brackets in $\Phi\left(Br(\mathbf{X})\right)$ of lower degree that P and $\xi_a \in \mathbb{R}$ for $a = 1 \ldots m$. Also, suppose that (4.37) is locally configuration accessible at $x \in M$. Then (4.37) is STLCC at x.

In the absence of a potential term, $Z^{\mathrm{lift}} = 0$, a stronger statement holds.

Theorem 33 *Suppose that the 1-homogeneous control system (4.37) has no potential term, i.e., $Z^{\mathrm{lift}} = 0$. If the 1-homogeneous system with dissipation excluded, eq. (4.37), is STLCC at $x \in M$, then the original system (4.37) is STLCC at $x \in E$.*

Again, this is a generalization of the configuration controllability results found in Cortés et al. [34].

4.4 Examples

In this section, we present several instances of controllability analysis found in the literature. For each case, we define the vector bundle, E, and demonstrate how the instances are special cases of 1-homogeneous control systems.

4.4.1 The Tangent Bundle

Let the vector bundle be given by the tangent bundle to a manifold, $E = TM$. Take a control system whose evolution, $X_S \in \mathcal{X}(TM)$, is a second-order vector field. A second-order vector field has the property that $T\pi \circ X_S(u_q) = u_q$, where $T\pi : TE \to TM$ is the bundle projection for TE. The space of all second-order vector fields will be denoted $\mathcal{X}^{SO}(E) \subset \mathcal{X}(E)$, for more information on second order vector fields, please

see [70, 36, 99, 1]. Given a second-order vector field, $X \in \mathcal{X}^{SO}(TQ)$, if it is of homogeneous order $+1$, e.g., $[\Delta, X] = X$, then it is said to be a *spray* [70]. There exists a bijection between linear symmetric connections and sprays (an affine connection and its related Christoffel symbols form such a bijection); see [70, 36] and references therein.

In a local trivialization the second iterated tangent bundle takes the form $T(TQ) \cong Q \times E \times E \times E$. Thus both the vertical and horizontal subspaces of $T(TQ)$ can be identified with TQ. This is behind much of the simplification that occurs in the paper by Lewis and Murray [85]. In particular, on the zero section,

$$\left[\Gamma, X_a^{\text{lift}}\right] \in HE \quad \text{and} \quad T\pi\left(\left[\Gamma, X_a^{\text{lift}}\right]\right) = -X_a.$$

Consequently, on the zero section,

$$T\pi\left(\left[\left[\Gamma, X_a^{\text{lift}}\right], \left[\Gamma, X_b^{\text{lift}}\right]\right]\right) = [X_a, X_b],$$

meaning that the space $\overline{\text{Lie}}\left(\left[\Gamma, \overline{\text{Sym}}(\mathcal{Y})\right]\right)$, when evaluated on the zero section, can be given by $\overline{\text{Lie}}\left(\overline{\text{Sym}}(\widetilde{\mathcal{Y}})\right)$, where $\widetilde{\mathcal{Y}} = \{ Y_1, \ldots Y_m, Z \}$ is the set of unlifted vector fields from $\mathcal{Y}$. The commutativity of the symmetric product, see Equations (4.13) and (4.14), also has ramifications: the symmetric closure $\overline{\text{Sym}}(\mathcal{Y})$, when evaluated on the zero section, is in 1-1 correspondence with the symmetric closure of the unlifted vector fields $\overline{\text{Sym}}(\widetilde{\mathcal{Y}})$. Essentially, all of the calculations drop to TM, resulting in computational reductions, precisely as emphasized by Lewis and Murray [85]. The important configuration controllability lemmas and theorems are rewritten below using the simplified notation.

Lemma 13 *[85] Let $x \in M$. Then,*

$$D_{\overline{Lie}(\mathcal{V})}(0_x) \bigcap V_{0_x}E = \left(D_{\overline{Sym}(\widetilde{\mathcal{Y}})}(x)\right)^{\text{lift}},$$

and

$$D_{\overline{Lie}(\mathcal{V})}(0_x) \bigcap T_xM = D_{\overline{Lie}\left(\overline{Sym}(\widetilde{\mathcal{Y}})\right)}(x).$$

Theorem 34 *[85] The control system (4.18) is locally configuration accessible at $x \in M$ if $\mathcal{C}_{hor}(\mathcal{V})(x) = T_xM$.*

Theorem 35 *[85] Suppose that $\widetilde{\mathcal{Y}}$ is such that every bad symmetric product in $Pr(\mathbf{Y})$ has the property that*

$$\text{Ev}_x(\psi)(\rho(P)) = \sum_{a=1}^{m} \xi_a \text{Ev}_x(\psi)(C_a),$$

where C_a are good symmetric products in $Pr(\mathbf{Y})$ of lower degree that P and $\xi_a \in \mathbb{R}$ for $a = 1 \ldots m$. Also, suppose that (4.18) is locally configuration accessible at $x \in M$. Then (4.18) is STLCC at x.

The bijection between sprays and linear symmetric connections, means that the controllability analysis of any second-order system such that $X_S \in (\mathcal{P}_1 \oplus \mathcal{P}_{-1}) \bigcap \mathcal{X}^{SO}(TQ)$ is a special case of the affine connection approach; for example [112].

4.4.2 The Cotangent Bundle

As the cotangent bundle is a vector bundle in its own right, controllability analysis can be done on the cotangent bundle just as easily as on the tangent bundle. A notable exception is that the identification of

the horizontal and vertical tangent spaces of $T(T^*Q)$ with the covector bundle T^*Q will not be made. Analysis on the cotangent bundle is useful for Hamiltonian systems, although the authors are unaware of any research that specifically studies how the Lie-algebraic structure of Hamiltonian systems may be used for control. The ability to formulate controllability results on the cotangent bundle will play a vital role when considering the case of nonhonolonomic mechanical systems with symmetry.

4.4.3 Constrained Mechanical Systems

Controllability for systems with nonholonomic velocity constraints can be addressed using this theory. While affine connections and constraints are compatible [84], we focus on alternative approaches that describe contraints via an Ehresmann connection, $A : TQ \to VQ$, on the fiber bundle Q with bundle projection $\tau : Q \to M$ and model fiber S [15]. In a local trivialization the vector bundle is $E \cong M \times S \times V$, where V is the model vector space for the tangent bundle to M, and $M \times S$ is the base space. The bundle projection operator is defined to be $\pi_Q : E \to Q$.

This description applies to the constrained mechanical systems found in Bloch et al. [15], for which the paper by Bloch et al. [16] details configuration controllability for the special case of an Ehresmann connection. The tangent bundle TE cannot be identified with E, so the computational simplifications from Lewis and Murray [85] do not occur. Fortunately, the vertically lifted structure of the control vector fields has fundamental implications beyond this type of identification. The inputs $Y_a^{\text{lift}} \in \mathcal{Y}$ are sections of the vector bundle E, and we can uniquely define the set $\widetilde{\mathcal{Y}}$ using $\widetilde{\mathcal{Y}}^{\text{lift}} = \mathcal{Y}$, just as in the above case.

More importantly, the set $\widetilde{\mathcal{Y}}$ can actually be horizontally lifted into TE, c.f. Section 3.2. The horizontal lift will be done using the constraint Ehresmann connection as opposed to the trivial Ehresmann connection that was defined earlier. It can be shown that the horizontal lift of the constraint Ehresmann connection is still horizontal with regards to the trivial Ehresmann connection. In a local trivialization, the horizontal lift takes an element of E and maps it into a horizontal vector in TE,

$$(r, s, v)^h = (v, -A_{loc}(r, s) \cdot v, 0) \in T_{(r,s,v)}E, \tag{4.50}$$

where A_{loc} is the local form of the Ehresmann connection [15]. It is normally assumed that the base space, M, to Q is fully controllable; (configuration) controllability of the fiber is the unknown. Configuration controllability is found via $\overline{\text{Lie}}(\left[\Gamma, \overline{\text{Sym}}(\mathcal{Y}\right])$, which coincides precisely with what Bloch et al. [16] find, however the analysis found here works to arbitrary orders of bracketing. According to the analysis of Section 3.2, $\overline{\text{Lie}}(\left[\Gamma, \overline{\text{Sym}}(\mathcal{Y}\right])$ corresponds to $\overline{\text{Lie}}(\left(\overline{\text{Sym}}(\widetilde{\mathcal{Y}})\right)^h)$, where $\cdot^h$ is the horizontal lift of sections of E using the constraint Ehresmann connection. It is well known, and was shown in Section 3.2, that the Jacobi-Lie brackets of horizontally lifted vector fields are equivalent to covariant derivatives of the local connection form. The involutive closure $\overline{\text{Lie}}(\left[\Gamma, \overline{\text{Sym}}(\mathcal{Y}\right])$, therefore, corresponds to covariant derivatives of the constraint connection. The configuration controllability test will span $H_{0_q}E$ only if the curvature form, $B(r, s)$, of the Ehresmann connection, and its higher-order covariant derivatives span the tangent fiber of TS. Constraint reduction simplifies the controllability test to analysis of the local connection form only. Thus, we can also show that the constraints and their vector bundle structure will naturally imply a reduction of the configuration controllability tests. Of course, if the base space is not fully controlled, then additional computations are required to determine this.

If the fiber is a Lie group, $S = G$, and the equations of motion are group invariant[3], then we are in the principal kinematic case, which was initially studied by Kelly and Murray [67], and was also covered in Section 3.2.2. All of their notions of fiber controllability (weak and strong) hold for this case. The local form of the principal connection, $\mathcal{A}_{loc}(r) : T_rM \to \mathfrak{g}$, only depends on the base space, meaning that

[3]Typically this arises from group invariant Lagrangian and constraints.

configuration controllability is further reduced to analysis of the associated adjoint bundle $\widetilde{\mathfrak{g}}$, as in Section 3.2.2. See also [30, 67, 137] for more information.

The mechanical systems described by this bundle construction correspond to the dynamic extension of kinematic (i.e. driftless) systems with constraints.

4.4.4 Constrained Mechanical Systems with Symmetry

For some systems, the constraints do not fully constrain the fiber. In this case, it may be possible to accrue momentum in the unconstrained fiber directions. Assume further, that the we have a principal fiber bundle, $Q \cong M \times G$, where G is a Lie group. The Lie algebra, $\mathfrak{g}$, corresponding to the Lie group is decomposed into two parts, (1) one which is kinematically constrained, and (2) a second which is not. The system will evolve on the vector bundle, $E \cong TM \times G \times V$, with base space diffeomorphic to $M \times G$, and vector bundle diffeomorphic to $W \times V$, where W is the model fiber for the tangent space TM. Typically V will be the unconstrained subspace of the Lie algebra, $\mathfrak{g}$, or its dual, $\mathfrak{g}^*$. The controllability analysis may be done for either case.

Choosing $V \subset \mathfrak{g}$, reproduces the conclusions found in Cortés et al. [35]. Cortés et al. combined the affine connection and principal connection approaches to come up with configuration controllability results for constrained mechanical systems with symmetry. By focusing only on the homogeneous structure of the vector bundle E, we may obtain the same conclusions without the complex notation coming from the merger of principal connections with affine connections. This approach corresponds to beginning with the reduced structure; a coordinate invariant derivation may be found in Cendra et al. [29]. Selecting, instead, $V \subset \mathfrak{g}^*$, we recover the equations of motion found in Bloch et al. [15]. Controllability analysis for this case can be found in Ostrowski and Burdick [125, 126], where part of the system evolves on the dual to the Lie algebra instead of the Lie algebra itself. Using $V \subset \mathfrak{g}^*$ corresponds to part of the dynamics having Hamiltonian-like evolution.

In particular, the controllability analysis is reduced to the examination of the curvature of the principal connection as before, and also of the symmetric products. The symmetric products are what allows the system to accrue momentum in the unconstrained fiber directions [126]. Ostrowski and Burdick [125], and Bullo and Lewis [25] were aware that the two approaches are in some way connected, and recommended further study of this connection. This example demonstrates why the link exists, and does so using the geometric homogeneity properties inherent to many mechanical systems as envisioned by Bullo and Lewis [25].

4.5 Conclusion

The description of mechanical systems from the perspective of geometric homogeneity demonstrates that certain system properties are not tied down to any particular instantiation of the equations of motion. Although each mechanical system framework can be used to highlight special properties to varying degrees, certain properties will hold irrespective of the paradigm chosen. Therefore, this leads to the idea of studying second-order mechanical systems from a broader differential geometric perspective. This chapter explored the role of geometric homogeneity on the evolution of control systems on vector bundles, and developed configuration controllability tests analogous to those for simple mechanical systems. The homogeneous structure gives priority to specific Jacobi-Lie bracket combinations when searching for controllability, thus allowing one to heuristically prune the controllability analysis.

A surprising result to come out of this analysis is that the computational simplifications found on $E = TQ$ are not unique to the affine connection case. Utilizing the structure afforded by constraints and symmetries can also drastically reduce the computational costs of configuration controllability tests. In

fact the simplifications may be larger since very often the base space is controllable and what needs to be discovered is configuration controllabilty of the fiber. The computations are reduced not only to the tangent space of the base space of the vector bundle, TQ, but to a submanifold of this tangent space.

The role of dissipation has been examined because of the desire to incorporate linearly dissipative forces into the equations of motion. It has been shown that, for some cases, the original configuration controllability analysis for systems without dissipation is not severely affected by the introduction of dissipation. It has also been shown that dissipation causes many originally vanishing Lie brackets to be non-vanishing. These new contributions may allow for additional control authority if the controllability analysis without dissipation did not give the desired result. Such an example can be found in the work of Ostrowski on the development of an underwater eel-inspired robot [106]. Without external dissipation, the mathematical model of the aquatic robot would not be capable of locomotion.

It would be useful to design control laws based on the controllability analysis of this chapter. Controller synthesis has been done using *motion control algorithms* for simple mechanical systems [21, 24, 100]. The next chapter, Chapter 5, will cover this idea in more detail. The averaging methods of Chapter 2 recover the symmetric product directions, or the Lie bracket directions needed for small-time local controllability. The technique will generalize the body of work on motion control algorithms to arbitrary 1-homogeneous control systems, just as this chapter generalized the body of work on configuration controllability for simple mechanical systems.

Chapter 5

Control of Underactuated 1-Homogeneous Systems with Drift

This chapter introduces a new technique to stabilize a large class of underactuated nonlinear systems with drift. The technique is a direct extension of the stabilization technique for underactuated driftless systems from Chapter 3, and is based on the generalized averaging theory of Chapter 2. We apply this averaging method to *1-homogeneous control systems* with drift. The method exponentially stabilizes 1-homogeneous control systems in the average. As with the driftless case, the method does not need a homogeneous norm to demonstrate exponential stabilization, coordinate transformations to canonical forms, the construction of Lyapunov functions, nor the pre-existence of stabilizing controllers, although such elements can be used to facilitate the process. The complexity of the nonlinear analysis grows with the order of Jacobi-Lie bracketing; only lower order control will be discussed in detail. Extensions to higher-order follow the same principles.

Organization of this chapter. Section 5.1 examines the structure of 1-homogeneous control systems for transformation to the standard form required by the perturbation methods of the generalized averaging theory from Chapter 2. Once in standard form, averaged expansions for the control systems may be computed. Techniques to stabilize 1-homogeneous control systems using the averaged expansions are presented in Section 5.2. Section 5.3 illustrates the method with various examples.

5.1 Averaging of Oscillatory Systems with Drift

The material from Section 2.5 is covered in more detail here, for the case of *vibrational control* of a system with drift. Recall that vibrational control inputs are high amplitude, high frequency, i.e.,

$$\dot{x} = X(x) + (1/\epsilon)F(x, t/\epsilon), \tag{5.1}$$

with ϵ small and where $F(\cdot, t)$ is T-periodic. Although the system (5.1) may be averaged using nonlinear Floquet theory, it is rare that a closed form solution will exist. Consequently, perturbation methods will be required to obtain an approximation. After transformation of time, $t/\epsilon \mapsto \tau$, and application of the variation of constants, the system is transformed to the form required by the perturbation methods of the generalized averaging theory. Transform time, $t/\epsilon \mapsto \tau$, to obtain

$$\frac{\mathrm{d}x}{\mathrm{d}\tau} = \epsilon X(x) + F(x, \tau). \tag{5.2}$$

In the variation of constants method, $\epsilon X(x)$ is a perturbation to the primary vector field $F(x, \tau)$. Define the vector field

$$Y(y, \tau) = \epsilon \left(\left(\Phi^F_{0,\tau} \right)^* X \right)(y), \tag{5.3}$$

where $\Phi^F_{0,\tau}$ is the flow of the vector field F. According to the variation of constants, the solution $x(t)$ is given exactly by

$$x(\tau) = \Phi^F_{0,\tau}(y(\tau)), \tag{5.4}$$

where $y(\tau)$ is the solution to the system

$$\dot{y} = Y(y,\tau), \quad y(0) = x_0, \tag{5.5}$$

as per Eq. (5.3). In order to apply Floquet theory to the vector field Y, a periodicity assumption, Assumption 1, was introduced in Section 2.5. Periodicity of the flow, $\Phi^F_{0,t}$, implies that Y is alo periodic, with the same period, T. Consequently, one can compute the average of Y to obtain the autonomous evolution equation:

$$\frac{\mathrm{d}z}{\mathrm{d}\tau} = Z(z), \tag{5.6}$$

where, Z is the autonomous vector field obtained from application of Theorem 8 to the time-periodic vector field Y.

Theorem 36 *The Floquet decomposition of system (5.2), under Assumption 1, is given by the averaged autonomous vector field of system (5.6), and the Floquet mapping,*

$$\widetilde{P}(\tau) = \Phi^F_{0,\tau} \circ P(t).$$

The mapping $P(\tau)$ and the autonomous vector field Z are obtained via the Floquet decomposition of Y from Equation (5.3).

proof
Floquet theory decomposes the flow of Y into the composition of two parts,

$$\Phi^Y_{0,t} = P(t) \circ \exp(Zt), \tag{5.7}$$

where $P(t)$ is a T-periodic mapping and $\exp(Zt)$ is an autonomous flow. The exponent, or autonomous flow, of Z describes the essential behavior of the system. The flow of the system (5.2) is

$$\Phi^{\epsilon X+F}_{0,\tau} = \Phi^F_{0,\tau} \circ \Phi^Y_{0,\tau}, \tag{5.8}$$

per the variation of constants. Applying Theorem 8,

$$\Phi^{\epsilon X+F}_{0,\tau} = \Phi^F_{0,\tau} \circ P(\tau) \circ \exp(Z\tau) \tag{5.9}$$

As both $\Phi^F_{0,\tau}$ and $P(t)$ are T-periodic, they can both be absorbed into one T-periodic mapping,

$$\widetilde{P}(\tau) \equiv \Phi^F_{0,t} \circ P(t). \tag{5.10}$$

Therefore the Floquet decomposition of (5.2) is

$$\Phi^{\epsilon X+F}_{0,\tau} = \widetilde{P}(\tau) \circ \exp(Z\tau). \tag{5.11}$$

■

The Floquet decomposition of Equation (5.3) is in the form required to obtain truncations of the infinite series expansions for $\widetilde{P}$ and Z. Nonlinear Floquet theory and perturbation methods may be applied to

systems with drift of the form assumed here. In particular the proximity of the flow of truncated averages to the actual flow (Theorems 11 and 13, the determination of stable orbits (Theorem 9), and the stabilization of fixed points (Corollary 2).

Alternative Forms: High Frequency and Small Motions

It is possible to consider instead the case of high-frequency oscillatory control of a system with drift, e.g.,

$$\dot{x} = X(x) + F(x, t/\epsilon). \tag{5.12}$$

with ϵ small and where $F(\cdot, t)$ is T-periodic. A transformation of time converts the system to the standard form required by perturbation theory,

$$\frac{\mathrm{d}x}{\mathrm{d}\tau} = \epsilon X(x) + \epsilon F(x, \tau). \tag{5.13}$$

Equation (5.13) is similar to the case of small motions studied in Bullo et. al. [24]. The case of small-motions takes the form,

$$\dot{x} = X(x) + \epsilon F(x, t), \tag{5.14}$$

with the assumption that $X(x)$ is initially small. The previous theory reviewed above can still be applied to Equations (5.13) and (5.14), however, the dependence of the formulas on ϵ will differ. The control analysis that will be done for the vibrational control case also applies to the two forms in Equations (5.13) and (5.14).

5.1.1 Averaging and Geometric Homogeneity

The vibrational control systems to be examined are assumed to be 1-homogeneous control systems. 1-homogeneous control systems take the form,

$$\dot{x} = X(x) + Y_a^{\text{lift}} u^a(x, t), \tag{5.15}$$

where $X \in \mathcal{M}_1$ and $Y_a^{\text{lift}} \in \mathcal{M}_{-1}$. We consider control inputs that combine state feedback and time-periodic vibrational terms; $u^a(x, t) = f^a(x) + (1/\epsilon)v^a(t/\epsilon)$, with $v^a(\cdot)$ T-periodic. Substituting these controls into (5.15) gives

$$\dot{x} = X(x) + Y_a^{\text{lift}}(x) f^a(x) + \frac{1}{\epsilon} Y_a^{\text{lift}}(x) v^a(t/\epsilon).$$

To meet our requirements, the state feedback, $f^a(x)$, can consist of terms that are at most homogeneous order 2; state feedback which is an arbitrary function of the base space, M, and linear in the vector fiber, V, easily satisfies this requirement. The time-independent terms are absorbed into one term,

$$\dot{x} = X_S(x) + \frac{1}{\epsilon} Y_a^{\text{lift}}(x) v^a(t/\epsilon), \tag{5.16}$$

which is $X_S \in \mathcal{M}_1$. The system may be transformed into the form required by the perturbation methods of averaging theory, according to the previous discussion. After transformation of time, and application of the variation of constants, we obtain the equivalent form of Equation (5.3) for the 1-homogeneous vibrational control systems of Equation (5.16),

$$Y(y, \tau) = \left(\Phi_{0,\tau}^{Y_a^{\text{lift}}(x)v^a(\tau)}\right)^* (\epsilon X_S) = \epsilon \left(\Phi_{0,\tau}^{Y_a^{\text{lift}}(x)v^a(\tau)}\right)^* (X_S). \tag{5.17}$$

Proposition 19 *[23] The flow $\Phi_{0,\tau}^{Y_a^{\text{lift}}(x)v^a(\tau)}$ is T-periodic.*

Proposition 19 implies that Assumption 1 holds, and Floquet analysis of Y may proceed. To compute the pull-back in Equation (5.17), we may utilize the series expansion in Equation (2.45),

$$\left(\Phi_{0,\tau}^{Y_a^{\text{lift}}v^a(\tau)}\right)^* X_S = X_S + \sum_{k=0}^{\infty}\int_0^T \cdots \int_0^{s_{k-1}} \left(\text{ad}_{Y_a^{\text{lift}}v^a(s_k)} \cdots \text{ad}_{Y_a^{\text{lift}}v^a(s_1)} X_S\right) ds_k \dots ds_1 \tag{5.18}$$

where the $\{s_j\}$ represent time and ad operates as follows: $\text{ad}_g f = [g, f]$. Due to the homogeneous structure of our class of systems, only the first two terms of the summation are nonvanishing. Hence, (5.17) takes the form

$$Y = \epsilon X_S + \epsilon \text{V}_{(1)}^{(a)}(t)\left[Y_a^{\text{lift}}, X_S\right] - \frac{1}{2}\epsilon \text{V}_{(1,1)}^{(a,b)}(t)\left\langle Y_a^{\text{lift}} : Y_b^{\text{lift}}\right\rangle, \tag{5.19}$$

where the $\text{V}_{(\cdot)}^{(\cdot)}(t)$ terms are the *averaging coefficients* from Section 3.1.1, and the symmetric product calculations use $\Gamma = X_S$, c.f. Definition 37.

5.2 Averaging and Control

Although the standard form for a 1-homogeneous control system is not in the form required by the perturbation methods of averaging theory, Section 5.1 demonstrated that the variations of constants transformation resulted in a system with a truncatable average, Equation (5.19). The vector field, Y, from Equation (2.1) determines the differential equation,

$$\frac{\text{d}y}{\text{d}\tau} = Y = \epsilon X_S + \epsilon \text{V}_{(1)}^{(a)}(\tau)\left[Y_a^{\text{lift}}, X_S\right] - \frac{1}{2}\epsilon \text{V}_{(1,1)}^{(a,b)}(\tau)\left\langle Y_a^{\text{lift}} : Y_b^{\text{lift}}\right\rangle,$$

which is T-periodic according to Proposition 19. We may now apply averaging theory. Notably, we may use averaged expansions obtained from a truncated approximation according to Theorem 11. Recall from Section 5.1 that the exact Floquet mapping will be

$$\widetilde{P}(\tau) = \Phi_{0,\tau}^{Y_a^{\text{lift}}v^a(\tau)} \circ P(\tau). \tag{5.20}$$

The mapping $P(\tau)$ will be known as the *partial Floquet mapping* to distinguish it from the actual Floquet mapping, $\widetilde{P}(t)$, for the system (5.16). After transformation to regular time, t, the m^{th}-order averaged system corresponding to (5.16) is

$$\begin{aligned} x(t) &= \text{Trunc}_{m-1}\left(\widetilde{P}(t/\epsilon)\right)(z(t)) + O\left(\epsilon^{m-1}\right) \\ \dot{z} &= \frac{1}{\epsilon}\text{Trunc}_m\left(Z\right) + O\left(\epsilon^m\right), \end{aligned}$$

where the m^{th}-th order truncation of the Floquet mapping is

$$\text{Trunc}_m\left(\widetilde{P}(t/\epsilon)\right) = \Phi_{0,t/\epsilon}^{\frac{1}{\epsilon}Y_a^{\text{lift}}v^a(t/\epsilon)} \circ \text{Trunc}_m\left(P(t/\epsilon)\right).$$

The partial Floquet mapping $P(t/\epsilon)$ and the autonomous vector field Z are powers series in ϵ. The transformation back to regular time lowers the power series order by one. This might pose a problem at low orders of averaging since $\text{Trunc}_m\left(P(\tau)\right)$ will go from being $O\left(\epsilon^m\right)$ close to being $\text{Trunc}_m\left(P(t/\epsilon)\right)$, which is $O\left(\epsilon^{m-1}\right)$ close. Choosing the improved m^{th}-th order average, with $\text{Trunc}_m\left(\widetilde{P}(t/\epsilon)\right)$, will resolve the

problem, c.f. Equation (2.35a),

$$x(t) = \mathrm{Trunc}_m\left(\widetilde{P}(t/\epsilon)\right)(z(t)) + O\left(\epsilon^m\right)$$
$$\dot{z} = \frac{1}{\epsilon}\mathrm{Trunc}_m\left(Z\right) + O\left(\epsilon^m\right).$$

5.2.1 Averaged Expansions

The affine form of the control inputs implies the control inputs, u^a, enter linearly. Having oscillatory inputs with non-zero average is equivalent to having constant actuation plus zero average oscillatory inputs, therefore the following assumption will be made:

Assumption 3 *It is assumed that the first order time-averages of the control inputs vanish, i.e.,* $\overline{\mathrm{V}^{(a)}_{(0)}(t)} = 0$.

First-order. First order averaging gives the autonomous differential equation,

$$\dot{z} = X_S + \overline{\mathrm{V}^{(a)}_{(1)}(t)}\left[Y_a^{\mathrm{lift}}, X_S\right] - \frac{1}{2}\overline{\mathrm{V}^{(a,b)}_{(1,1)}(t)}\left\langle Y_a^{\mathrm{lift}} \ : \ Y_b^{\mathrm{lift}}\right\rangle. \tag{5.21}$$

The first order partial Floquet mapping is $\mathrm{Trunc}_0\left(P(t/\epsilon)\right) = \mathrm{Id}$. An improved version capturing an additional order of ϵ is given by

$$\mathrm{Trunc}_1\left(P(t/\epsilon)\right) = \mathrm{Id} + \epsilon\int_0^{t/\epsilon}\widetilde{\mathrm{V}}^{(a)}_{(1)}(\tau)\,\mathrm{d}\tau\,\left[Y_a^{\mathrm{lift}}, X_S\right] - \frac{1}{2}\epsilon\int_0^{t/\epsilon}\widetilde{\mathrm{V}}^{(a,b)}_{(1,1)}(\tau)\,\mathrm{d}\tau\,\left\langle Y_a^{\mathrm{lift}} \ : \ Y_b^{\mathrm{lift}}\right\rangle. \tag{5.22}$$

Second-order. The second-order averaged vector field is

$$\begin{aligned}
\dot{z} = \ & X_S + \overline{\mathrm{V}^{(a)}_{(1)}(t)}\left[Y_a^{\mathrm{lift}}, X_S\right] - \tfrac{1}{2}\overline{\mathrm{V}^{(a,b)}_{(1,1)}(t)}\left\langle Y_a^{\mathrm{lift}} \ : \ Y_b^{\mathrm{lift}}\right\rangle \\
& + \epsilon\left(\overline{\mathrm{V}^{(a)}_{(2)}(t)} - \tfrac{1}{2}T\,\overline{\mathrm{V}^{(a)}_{(1)}(t)}\right)\left[\left[Y_a^{\mathrm{lift}}, X_S\right], X_S\right] \\
& - \tfrac{1}{2}\epsilon\left(\overline{\mathrm{V}^{(a,b)}_{(\widehat{1,1})}(t)} - \tfrac{1}{2}T\,\overline{\mathrm{V}^{(a,b)}_{(1,1)}(t)}\right)\left[\left\langle Y_a^{\mathrm{lift}} \ : \ Y_b^{\mathrm{lift}}\right\rangle, X_S\right] \\
& + \tfrac{1}{2}\epsilon\overline{\mathrm{V}^{(a,b)}_{(2,1)}(t)}\left[\left[Y_a^{\mathrm{lift}}, X_S\right], \left[Y_b^{\mathrm{lift}}, X_S\right]\right] \\
& + \tfrac{1}{2}\epsilon\left(\overline{\mathrm{V}^{(a,b,c)}_{(2,1,1)}(t)} - \tfrac{1}{2}T\,\overline{\mathrm{V}^{(a)}_{(1)}(t)}\,\overline{\mathrm{V}^{(b,c)}_{(1,1)}(t)}\right)\left\langle Y_a^{\mathrm{lift}} \ : \ \left\langle Y_b^{\mathrm{lift}} \ : \ Y_c^{\mathrm{lift}}\right\rangle\right\rangle.
\end{aligned} \tag{5.23}$$

The second-order partial Floquet mapping is given in Equation (5.22).

Third- and higher-order. The third-order truncated partial Floquet mapping and averaged vector field can be found in Tables 5.1 and 5.2, respectively. To compute fourth- and higher-order averages, follow the averaged expansions procedure delineated in Chapter 2, taking into account the control decomposition. In the Appendix 5.A, lower-order averaged expansions are worked out to give the reader a sense of the procedure involved.

5.2.2 Sinusoidal Inputs for Indirect Actuation

From a controls perspective, the averaged coefficients play an important role. By appropriately modulating their values, it is possible to effect controlled flow in the direction of the Jacobi-Lie brackets or symmetric products they multiply. Accordingly, there has been much study of oscillatory input combinations that will

$$
\begin{aligned}
\mathrm{Trunc}_2\left(P(t/\epsilon)\right) = \mathrm{Id} &+ \epsilon \int_0^{t/\epsilon} \widetilde{\mathrm{V}}^{(a)}_{(1)}(\tau)\,\mathrm{d}\tau\, \left[Y_a^{\mathrm{lift}}, X_S\right] - \frac{1}{2}\epsilon \int_0^{t/\epsilon} \widetilde{\mathrm{V}}^{(a,b)}_{(1,1)}(\tau)\,\mathrm{d}\tau\, \left\langle\, Y_a^{\mathrm{lift}} \;:\; Y_b^{\mathrm{lift}} \,\right\rangle \\
&+ \epsilon^2 \left(\int_0^{t/\epsilon} \widetilde{\mathrm{V}}^{(a)}_{(2)}(\tau)\,\mathrm{d}\tau \, - \frac{1}{2}\frac{t}{\epsilon} \int_0^{t/\epsilon} \widetilde{\mathrm{V}}^{(a)}_{(1)}(\tau)\,\mathrm{d}\tau \right. \\
&\qquad \left. - \frac{1}{2}\frac{t}{\epsilon} \left(\int_0^{t/\epsilon} \mathrm{V}^{(a)}_{(1)}(\tau)\,\mathrm{d}\tau \, - T\,\overline{\mathrm{V}^{(a)}_{(1)}(\tau)} \right) \right) \left[\left[Y_a^{\mathrm{lift}}, X_S \right], X_S \right] \\
&- \frac{1}{2}\epsilon^2 \int_0^{t/\epsilon} \widetilde{\mathrm{V}}^{(a,b)}_{(1,1)}(\tau)\,\mathrm{d}\tau \, - \frac{t}{\epsilon} \int_0^{t/\epsilon} \widetilde{\mathrm{V}}^{(a,b)}_{(1,1)}(\tau)\,\mathrm{d}\tau \\
&\qquad - \frac{1}{2}\frac{t}{\epsilon} \left(\int_0^{t/\epsilon} \mathrm{V}^{(a,b)}_{(1,1)}(\tau)\,\mathrm{d}\tau \, - T\,\overline{\mathrm{V}^{(a,b)}_{(1,1)}(\tau)} \right) \left[\left\langle\, Y_a^{\mathrm{lift}} \;:\; Y_b^{\mathrm{lift}} \,\right\rangle, X_S \right] \\
&- \frac{1}{2}\epsilon^2 \int_0^{t/\epsilon} \mathrm{V}^{(a,b)}_{(2,1)}(\tau)\,\mathrm{d}\tau \, \left[\left[Y_a^{\mathrm{lift}}, X_S \right], \left[Y_b^{\mathrm{lift}}, X_S \right] \right] \\
&+ \frac{1}{2}\epsilon^2 \left(\int_0^{t/\epsilon} \widetilde{\mathrm{V}}^{(a,b,c)}_{(2,1,1)}(\tau)\,\mathrm{d}\tau \, - \frac{1}{2} \left(\mathrm{V}^{(a,\widehat{b},c)}_{(2,\widehat{1},1)}(t/\epsilon) - T\frac{t}{\epsilon}\overline{\mathrm{V}^{(a)}_{(1)}(\tau)}\,\overline{\mathrm{V}^{(b,c)}_{(1,1)}(\tau)}\frac{t^2}{\epsilon^2} \right) \right) \\
&\qquad \left\langle\, Y_a^{\mathrm{lift}} \;:\; \left\langle\, Y_b^{\mathrm{lift}} \;:\; Y_c^{\mathrm{lift}} \,\right\rangle \right\rangle \\
&- \epsilon^2 \left(\mathrm{V}^{(a)}_{(2)}(t/\epsilon)\overline{\mathrm{V}^{(b)}_{(1)}(\tau)}\frac{t}{\epsilon} - \frac{1}{2}\mathrm{V}^{(a,b)}_{(2,2)}(t/\epsilon) - \frac{1}{2}\overline{\mathrm{V}^{(a)}_{(1)}(\tau)}\,\overline{\mathrm{V}^{(b,c)}_{(1,1)}(\tau)}\frac{t^2}{\epsilon^2} \right) \\
&\qquad \left[Y_a^{\mathrm{lift}}, X_S \right] \cdot \left[Y_b^{\mathrm{lift}}, X_S \right] \\
&+ \frac{1}{2} \left(\mathrm{V}^{(a)}_{(2)}(t/\epsilon)\overline{\mathrm{V}^{(b,c)}_{(1,1)}(\tau)}\frac{t}{\epsilon} - \frac{1}{2}\mathrm{V}^{(\widehat{a,b},c)}_{(\widehat{1,1},2)}(t/\epsilon) - \frac{1}{2}\overline{\mathrm{V}^{(a)}_{(1)}(\tau)}\,\overline{\mathrm{V}^{(b,c)}_{(1,1)}(\tau)}\frac{t^2}{\epsilon^2} \right) \\
&\qquad \left[Y_a^{\mathrm{lift}}, X_S \right] \cdot \left\langle\, Y_b^{\mathrm{lift}} \;:\; Y_c^{\mathrm{lift}} \,\right\rangle \\
&+ \frac{1}{2}\epsilon^2 \left(\mathrm{V}^{(\widehat{a,b})}_{(\widehat{1,1})}(t/\epsilon)\overline{\mathrm{V}^{(c)}_{(1)}(\tau)}\frac{t}{\epsilon} - \frac{1}{2}\mathrm{V}^{(\widehat{a,b},c)}_{(\widehat{1,1},2)}(t/\epsilon) - \frac{1}{2}\overline{\mathrm{V}^{(a,b)}_{(1,1)}(\tau)}\,\overline{\mathrm{V}^{(c)}_{(1)}(\tau)}\frac{t^2}{\epsilon^2} \right) \\
&\qquad \left\langle\, Y_a^{\mathrm{lift}} \;:\; Y_b^{\mathrm{lift}} \,\right\rangle \cdot \left[Y_c^{\mathrm{lift}}, X_S \right] \\
&- \frac{1}{4}\epsilon^2 \left(\mathrm{V}^{(\widehat{a,b})}_{(\widehat{1,1})}(t/\epsilon)\overline{\mathrm{V}^{(b,c)}_{(1,1)}(\tau)}\frac{t}{\epsilon} - \frac{1}{2}\mathrm{V}^{(\widehat{a,b},\widehat{c,d})}_{(\widehat{1,1},\widehat{1,1})}(t/\epsilon) - \frac{1}{2}\overline{\mathrm{V}^{(a,b)}_{(1,1)}(\tau)}\,\overline{\mathrm{V}^{(c,d)}_{(1,1)}(\tau)}\frac{t^2}{\epsilon^2} \right) \\
&\qquad \left\langle\, Y_a^{\mathrm{lift}} \;:\; Y_b^{\mathrm{lift}} \,\right\rangle \cdot \left\langle\, Y_c^{\mathrm{lift}} \;:\; Y_d^{\mathrm{lift}} \,\right\rangle
\end{aligned}
$$

Table 5.1: Third-order partial Floquet mapping, $\mathrm{Trunc}_2\left(P(t/\epsilon)\right)$, for systems with drift.

$$
\begin{aligned}
\dot{z} = \;& \epsilon X_S + \epsilon \overline{\mathrm{V}^{(a)}_{(1)}(t)}\,\big[Y_a^{\mathrm{lift}}, X_S\big] - \tfrac{1}{2}\epsilon \overline{\mathrm{V}^{(a,b)}_{(1,1)}(t)}\,\big\langle\, Y_a^{\mathrm{lift}} \;:\; Y_b^{\mathrm{lift}} \,\big\rangle \\
& + \epsilon^2 \left(\overline{\mathrm{V}^{(a)}_{(2)}(t)} - \tfrac{1}{2} T\, \overline{\mathrm{V}^{(a)}_{(1)}(t)} \right) \big[\big[Y_a^{\mathrm{lift}}, X_S\big], X_S\big] \\
& - \tfrac{1}{2}\epsilon^2 \left(\overline{\mathrm{V}^{(a,b)}_{(\widehat{1},\widehat{1})}(t)} - \tfrac{1}{2} T\, \overline{\mathrm{V}^{(a,b)}_{(1,1)}(t)} \right) \big[\big\langle\, Y_a^{\mathrm{lift}} \;:\; Y_b^{\mathrm{lift}} \,\big\rangle, X_S\big] \\
& + \tfrac{1}{2}\epsilon^2 \overline{\mathrm{V}^{(a,b)}_{(2,1)}(t)}\,\big[\big[Y_a^{\mathrm{lift}}, X_S\big], \big[Y_b^{\mathrm{lift}}, X_S\big]\big] \\
& + \tfrac{1}{2}\epsilon^2 \left(\overline{\mathrm{V}^{(a,b,c)}_{(2,1,1)}(t)} - \tfrac{1}{2} T\, \overline{\mathrm{V}^{(a)}_{(1)}(t)}\; \overline{\mathrm{V}^{(b,c)}_{(1,1)}(t)} \right) \big\langle\, Y_a^{\mathrm{lift}} \;:\; \big\langle\, Y_b^{\mathrm{lift}} \;:\; Y_c^{\mathrm{lift}} \,\big\rangle \big\rangle \\
& - \epsilon^3 \left(\tfrac{1}{3}\overline{\mathrm{V}^{(a)}_{(2)}(t)\; t} - \tfrac{1}{3}\overline{\mathrm{V}^{(a)}_{(1)}(t)\; t^2} - \tfrac{1}{2} T\, \overline{\mathrm{V}^{(a)}_{(2)}(t)} + \tfrac{1}{4} T^2\, \overline{\mathrm{V}^{(a)}_{(1)}(t)} \right) \big[\big[\big[Y_a^{\mathrm{lift}}, X_S\big], X_S\big], X_S\big] \\
& + \epsilon^3 \left(\tfrac{1}{6}\overline{\mathrm{V}^{(a,\widehat{b})}_{(\widehat{1},\widehat{1})}(t)\; t} - \tfrac{1}{6}\overline{\mathrm{V}^{(a,b)}_{(1,1)}(t)\; t^2} - \tfrac{1}{2} T\, \overline{\mathrm{V}^{(a)}_{(2)}(t)} + \tfrac{1}{4} T^2\, \overline{\mathrm{V}^{(a)}_{(1)}(t)} \right) \big[\big[\big\langle\, Y_a^{\mathrm{lift}} \;:\; Y_b^{\mathrm{lift}} \,\big\rangle, X_S\big], X_S\big] \\
& - \epsilon^3 \left(\tfrac{1}{3}\overline{\mathrm{V}^{(a,b)}_{(2,1)}(t)\; t} - \tfrac{1}{4} T\, \overline{\mathrm{V}^{(a,b)}_{(2,1)}(t)} \right) \big[\big[\big[Y_a^{\mathrm{lift}}, X_S\big], \big[Y_b^{\mathrm{lift}}, X_S\big]\big], X_S\big] \\
& + \epsilon^3 \left(\tfrac{1}{6}\overline{\mathrm{V}^{(a,\widehat{b,c})}_{(1,\widehat{1},\widehat{1})}(t)\; t} - \tfrac{1}{6}\overline{\mathrm{V}^{(a,b,c)}_{(2,1,1)}(t)\; t} + \tfrac{1}{4} T\, \overline{\mathrm{V}^{(a,b,c)}_{(2,1,1)}(t)} - \tfrac{1}{8} T^2\, \overline{\mathrm{V}^{(a)}_{(1)}(t)}\; \overline{\mathrm{V}^{(b,c)}_{(1,1)}(t)} \right) \\
& \qquad \big[\big\langle\, Y_a^{\mathrm{lift}} \;:\; \big\langle\, Y_b^{\mathrm{lift}} \;:\; Y_c^{\mathrm{lift}} \,\big\rangle \big\rangle, X_S\big] \\
& + \epsilon^3 \left(\tfrac{1}{3}\overline{\mathrm{V}^{(a,b)}_{(2,2)}(t)} - -\tfrac{1}{3}\overline{\mathrm{V}^{(a,b)}_{(2,1)}(t)\; t} - \tfrac{1}{2} T\, \overline{\mathrm{V}^{(a)}_{(1)}(t)}\; \overline{\mathrm{V}^{(b)}_{(2)}(t)} + \tfrac{1}{4} T^2\, \overline{\mathrm{V}^{(a)}_{(1)}(t)}\; \overline{\mathrm{V}^{(b)}_{(1)}(t)} \right) \\
& \qquad \big[\big[Y_a^{\mathrm{lift}}, X_S\big], \big[\big[Y_b^{\mathrm{lift}}, X_S\big], X_S\big]\big] \\
& - \epsilon^3 \left(\tfrac{1}{6}\overline{\mathrm{V}^{(a,\widehat{b,c})}_{(2,\widehat{1},\widehat{1})}} - \tfrac{1}{6}\overline{\mathrm{V}^{(a,b,c)}_{(2,1,1)}(t)\; t} + \tfrac{1}{4} T\, \overline{\mathrm{V}^{(a)}_{(1)}(t)}\; \overline{\mathrm{V}^{(\widehat{a,b})}_{(\widehat{1},\widehat{1})}(t)} - \tfrac{1}{8} T^2\, \overline{\mathrm{V}^{(a)}_{(1)}(t)}\; \overline{\mathrm{V}^{(b,c)}_{(1,1)}(t)} \right) \\
& \qquad \big[\big[Y_a^{\mathrm{lift}}, X_S\big], \big[\big\langle\, Y_b^{\mathrm{lift}} \;:\; Y_c^{\mathrm{lift}} \,\big\rangle, X_S\big]\big] \\
& + \epsilon^3 \left(\tfrac{1}{3}\overline{\mathrm{V}^{(a,b,c)}_{(2,2,1)}(t)} - \tfrac{1}{4} T\, \overline{\mathrm{V}^{(a)}_{(1)}(t)}\; \overline{\mathrm{V}^{(b,c)}_{(2,1)}(t)} \right) \big[\big[Y_a^{\mathrm{lift}}, X_S\big], \big[\big[Y_b^{\mathrm{lift}}, X_S\big], \big[Y_c^{\mathrm{lift}}, X_S\big]\big]\big] \\
& + \epsilon^3 \left(\tfrac{1}{6}\overline{\mathrm{V}^{(a,b,\widehat{c,d})}_{(2,1,\widehat{1},\widehat{1})}(t)} - \tfrac{1}{6}\overline{\mathrm{V}^{(a,b,c,d)}_{(2,2,1,1)}(t)} + \tfrac{1}{4} T\, \overline{\mathrm{V}^{(a)}_{(1)}(t)}\; \overline{\mathrm{V}^{(b,c,d)}_{(2,1,1)}(t)} - \tfrac{1}{8} T^2\, \overline{\mathrm{V}^{(a)}_{(1)}(t)}\; \overline{\mathrm{V}^{(b)}_{(1)}(t)}\; \overline{\mathrm{V}^{(c,d)}_{(1,1)}(t)} \right) \\
& \qquad \big\langle\, Y_a^{\mathrm{lift}} \;:\; \big\langle\, Y_b^{\mathrm{lift}} \;:\; \big\langle\, Y_c^{\mathrm{lift}} \;:\; Y_d^{\mathrm{lift}} \,\big\rangle \big\rangle \big\rangle \\
& - \epsilon^3 \left(\tfrac{1}{6}\overline{\mathrm{V}^{(\widehat{a,b},c)}_{(\widehat{1},\widehat{1},2)}(t)} - \tfrac{1}{6}\overline{\mathrm{V}^{(\widehat{a,b},c)}_{(\widehat{1},\widehat{1},1)}(t)\; t} - \tfrac{1}{4} T\, \overline{\mathrm{V}^{(a,b)}_{(1,1)}(t)}\; \overline{\mathrm{V}^{(c)}_{(2)}(t)} + \tfrac{1}{8} T^2\, \overline{\mathrm{V}^{(a,b)}_{(1,1)}(t)}\; \overline{\mathrm{V}^{(c)}_{(2)}(t)} \right) \\
& \qquad \big[\big\langle\, Y_a^{\mathrm{lift}} \;:\; Y_b^{\mathrm{lift}} \,\big\rangle, \big[\big[Y_c^{\mathrm{lift}}, X_S\big], X_S\big]\big] \\
& + \epsilon^3 \left(\tfrac{1}{12}\overline{\mathrm{V}^{(\widehat{a,b},\widehat{c,d})}_{(\widehat{1},\widehat{1},\widehat{1},\widehat{1})}(t)} - \tfrac{1}{12}\overline{\mathrm{V}^{(\widehat{a,b},c,d)}_{(\widehat{1},\widehat{1},1,1)}(t)} - \tfrac{1}{8} T\, \overline{\mathrm{V}^{(a,b)}_{(1,1)}(t)}\; \overline{\mathrm{V}^{(\widehat{c,d})}_{(\widehat{1},\widehat{1})}(t)} + \tfrac{1}{1} T^2\, \overline{\mathrm{V}^{(a,b)}_{(1,1)}(t)}\; \overline{\mathrm{V}^{(c,d)}_{(1,1)}(t)} \right) \\
& \qquad \big\langle\, \big\langle\, Y_a^{\mathrm{lift}} \;:\; Y_b^{\mathrm{lift}} \,\big\rangle \;:\; \big\langle\, Y_c^{\mathrm{lift}} \;:\; Y_d^{\mathrm{lift}} \,\big\rangle \big\rangle \\
& + \epsilon^3 \left(\tfrac{1}{6}\overline{\mathrm{V}^{(a,b,\widehat{c,d})}_{(2,1,\widehat{1},\widehat{1})}(t)} - \tfrac{1}{6}\overline{\mathrm{V}^{(a,b,c,d)}_{(1,2,\widehat{1},\widehat{1})}(t)} - \tfrac{1}{8} T\, \overline{\mathrm{V}^{(a,b)}_{(2,1)}(t)}\; \overline{\mathrm{V}^{(c,d)}_{(1,1)}(t)} - \tfrac{1}{8} T\, \overline{\mathrm{V}^{(a,b)}_{(1,2)}(t)}\; \overline{\mathrm{V}^{(c,d)}_{(1,1)}(t)} \right) \\
& \qquad \big\langle\, Y_a^{\mathrm{lift}} \;:\; \big\langle\, Y_b^{\mathrm{lift}} \;:\; \big\langle\, Y_c^{\mathrm{lift}} \;:\; Y_d^{\mathrm{lift}} \,\big\rangle \big\rangle \big\rangle
\end{aligned}
$$

Table 5.2: Third-order averaged expansion, $\mathrm{Trunc}_3\,(Z)$, for system with drift.

uniquely activate select brackets (i.e. *approximate inversion*), [22, 100]. Sussmann and Liu [154, 155, 92] have provided a general strategy for open-loop approximate inversion of the oscillatory controls for driftless systems. This section examines the algorithm for select averaged coefficients appearing in the averaged expansions of systems with drift, with the goal of providing a practical approach to-, and interpretation of-, the Ω-set notation found in Sussman and Liu.

Remark on bad brackets. For systems with drift, the averaged coefficient analysis is complicated by the existence of bad bracket directions. The bad brackets correspond to symmetric products which have even degrees for each of the control inputs, c.f. Definition 54. In controllability analysis, bad brackets are bad because they are unidirectional; it is not possible to arbitrarily excite bad brackets for controlled flow. The ensuing analysis will also show that bad brackets are activated whenever their constitutive control inputs are actuated by oscillatory controls. Consequently, we operate under the assumption that the bad symmetric products can be neutralized, a critical requirement for most small-time local controllability theorems, and specifically those of Chapter 4. Many of the oscillatory inputs used to modulate specific averaged coefficients will also modulate the bad symmetric product averaged coefficients. When a set of oscillatory inputs is said to uniquely activate an average coefficient, the bad symmetric products are exempt from the statement.

First-order Averaged Coefficients

The process of determining the inputs for an arbitrary system to any given order of averaging is difficult. Sinusoidal input groups in resonance will lead to Jacobi-Lie bracket or symmetric product motion. Each averaged coefficient has requirements for resonance, which may conflict and cause coupling. Certain Jacobi-Lie brackets and symmetric products can never be actuated independently. Nevertheless, useful guidelines can be established by investigating simple, lower order cases. Abstraction to higher-order follows naturally.

For first order averaging, the averaged coefficient is

$$\overline{\mathrm{V}^{(a,b)}_{(1,1)}(t)} = \frac{1}{T}\int_0^T \int^t v^a(\tau)\,\mathrm{d}\tau \int^t v^b(\tau)\,\mathrm{d}\tau\ \mathrm{d}t\,.$$

For each input $v^a(t)$, $a = 1 \ldots m$, one of the following two forcing functions can be chosen,

$$\beta^a \sin(\omega_a t) \quad \text{or} \quad \beta^a \cos(\omega_a t),$$

where $\omega_a \in \mathbb{Z}^+ \setminus \{0\}$. The four permutations of the input function choices result in the averaged coefficient evaluations found in Table 5.3. Of the four options, out-of-phase sinusoidal inputs do not work. Additionally, the algebraic equality

$$\omega_a - \omega_b = 0 \tag{5.24}$$

must hold for the desired in-phase sinusoidal input pair and for no other oscillatory inputs. The net result of this analysis is a set of inputs that operate at unique carrier frequencies; a commonly known fact.

Lemma 14 *In order to excite the symmetric product* $\left\langle\, Y_a^{\mathrm{lift}} \,:\, Y_b^{\mathrm{lift}} \,\right\rangle$*, use the inputs*

$$v^a(t) = \beta^a_{ab}\cos(\omega_{ab}t) \quad \textit{and} \quad v^b(t) = \beta^b_{ab}\cos(\omega_{ab}t),$$

or

$$v^a(t) = \beta^a_{ab}\sin(\omega_{ab}t) \quad \textit{and} \quad v^b(t) = \beta^b_{ab}\sin(\omega_{ab}t),$$

$\overline{\mathrm{V}^{(a,b)}_{(1,1)}(t)}$	$\beta^a \sin(\omega_a t)$	$\beta^a \cos(\omega_a t)$
$\beta^b \sin(\omega_b t)$	$\begin{cases} \frac{\beta^a \beta^b}{2\omega_a \omega_b} & \text{if } \omega_a = \omega_b \\ 0 & \text{otherwise} \end{cases}$	0
$\beta^b \cos(\omega_b t)$	0	$\begin{cases} \frac{\beta\beta}{2\omega_a \omega_b} & \text{if } \omega_a = \omega_b \\ 0 & \text{otherwise} \end{cases}$

Table 5.3: First-order averaged coefficient, $\overline{\mathrm{V}^{(a,b)}_{(1,1)}(t)}$, for systems with drift.

for a unique carrier frequency, ω_{ab}. The averaged coefficient will evaluate to

$$\overline{\mathrm{V}^{(a,b)}_{(1,1)}(t)} = \frac{\beta ab}{2\omega_{ab}^2}, \quad \textit{where } \beta ab = \beta^a_{ab}\beta^b_{ab}.$$

Lemma 14 imposes a unique carrier frequency for each Jacobi-Lie bracket. This will pose a problem for severely underactuated systems, as the number of frequencies needed is directly related to the number of Jacobi-Lie brackets required. For a pair of oscillatory inputs generating motion in the direction of the symmetric product, $\langle Y_a : Y_b \rangle$, the averaged coefficient is found in Lemma 14 to be $\frac{\beta ab}{2\omega_{ab}^2}$. As the frequency increases, the response decreases. If β^a_{ab} and β^b_{ab} include a factor of ω_{ab}, the response is no longer attenuated by the frequency. Unfortunately, the oscillation amplitude for the inputs will scale with ω_{ab}, which may be large. An algorithm analogous to Algorithm 1 for driftless system can be used to recycle carrier frequencies and lower the total amount of unique frequencies used, while avoiding coupling.

Second-order Averaged Coefficients

For second-order averaging, there are two averaged coefficient types that are important with respect to the controllability analysis of Chapter 4. They are the averaged coefficients,

$$\overline{\mathrm{V}^{(a,b)}_{(2,1)}(t)} \qquad \text{and} \qquad \left(\overline{\mathrm{V}^{(a,b,c)}_{(2,1,1)}(t)} - \frac{1}{2}T\overline{\mathrm{V}^{(a)}_{(1)}(t)}\,\overline{\mathrm{V}^{(b,c)}_{(1,1)}(t)}\right).$$

It will be shown that the averaged coefficient $\overline{\mathrm{V}^{(a,b)}_{(2,1)}(t)}$ has the same properties as the averaged coefficient $\overline{\mathrm{V}^{(a,b)}_{(1,0)}(t)}$ from the driftless case, c.f. Section 3.1.2, leaving only the three-input averaged coefficient for serious analysis. Additionally, the use of the sinusoidal functions, sine and cosine, implies that the averaged coefficients $\overline{\mathrm{V}^{(a)}_{(n)}(t)}$ always vanish,

$$\overline{\mathrm{V}^{(a)}_{(n)}(t)} = 0, \quad \forall\, n \in \mathbb{Z}^+. \tag{5.25}$$

Hence analysis of the three-input averaged coefficient terms simplifies to

$$\overline{\mathrm{V}^{(a,b,c)}_{(2,1,1)}(t)}.$$

$\overline{\mathrm{V}^{(a,b)}_{(2,1)}(t)}$	$\alpha^a \sin(\omega_a t)$	$\alpha^a \cos(\omega_a t)$
$\alpha^b \sin(\omega_b t)$	0	$\begin{cases} -\frac{\alpha^a \alpha^b}{2\omega_a^2 \omega_b} & \text{if } \omega_a = \omega_b \\ 0 & \text{otherwise} \end{cases}$
$\alpha^b \cos(\omega_b t)$	$\begin{cases} \frac{\alpha^a \alpha^b}{2\omega_a^2 \omega_b} & \text{if } \omega_a = \omega_b \\ 0 & \text{otherwise} \end{cases}$	0

Table 5.4: Averaged coefficient, $\overline{\mathrm{V}^{(a,b)}_{(2,1)}(t)}$, for second-order averaging of systems with drift.

The averaged coefficient $\overline{\mathrm{V}^{(a,b)}_{(2,1)}(t)}$ is a coefficient to a Jacobi-Lie bracket term, whereas the averaged coefficient $\overline{\mathrm{V}^{(a,b,c)}_{(2,1,1)}(t)}$ is a coefficient to a symmetric product term, c.f. Equation (5.23). When evaluated for points in the equilibrium subspace as defined by Definition 44, the Jacobi-Lie brackets will correspond to motion in the base space only, also known as *kinematic motions*. The symmetric products will correspond to actuation of the vector fiber (e.g. accelerations), and will be called *dynamic motions*. Due to the distinction between the two types of motions, the parametrization labelling will be different. For kinematic motions the parameters will be denoted by α, as in the analysis for driftless systems, since kinematic motions correspond to maneuvers with no net drift. For dynamic motions, the parameters will be denoted by β, as was already seen in the first-order symmetric product case.

The averaged coefficient $\overline{\mathrm{V}^{(a,b)}_{(2,1)}(t)}$ is

$$\overline{\mathrm{V}^{(a,b)}_{(2,1)}(t)} = \frac{1}{T}\int_0^T \int^t \int^\tau v^a(\sigma)\,\mathrm{d}\sigma\ \mathrm{d}\tau\ \int^t v^b(\tau)\,\mathrm{d}\tau\ \mathrm{d}t\,.$$

For each input $v^a(t)$, $a = 1 \ldots m$, one of the following two forcing functions can be chosen,

$$\alpha^a \sin(\omega_a t) \quad \text{or} \quad \alpha^a \cos(\omega_a t),$$

where $\omega_a \in \mathbb{Z}^+$. The four permutations of input function choices result in the averaged coefficient evaluations found in Table 5.4. Of the four options, in-phase sinusoidal inputs do not work. Additionally, the algebraic equality

$$\omega_a - \omega_b = 0, \tag{5.26}$$

must hold for the desired out-of-phase sinusoidal input pair and for no other oscillatory inputs. These are the same conclusions that have been made for the averaged coefficient $\overline{\mathrm{V}^{(a,b)}_{(1,0)}(t)}$ for driftless systems, c.f. Lemma 2.

Lemma 15 *Consider the Jacobi-Lie bracket,* $\left[\left[Y_a^{\mathrm{lift}}, X_S\right], \left[Y_b^{\mathrm{lift}}, X_S\right]\right]$. *If the associated inputs*

$$v^a(t) = \alpha^a_{ab}\cos(\omega_{ab}t), \qquad v^b(t) = \alpha^b_{ab}\sin(\omega_{ab}t),$$

are chosen for some unique carrier frequency, ω_{ab}*, then only the bracket* $\left[\left[Y_a^{\mathrm{lift}}, X_S\right], \left[Y_b^{\mathrm{lift}}, X_S\right]\right]$ *will be*

excited. The corresponding averaged coefficient will evaluate to

$$\overline{\mathrm{V}^{(a,b)}_{(2,1)}(t)} = \frac{\alpha_{ab}}{2\omega^3_{ab}}, \quad \textit{where } \alpha_{ab} = \alpha^a_{ab}\alpha^b_{ab}.$$

The important factor for the selection of inputs in the previous analysis involved the algebraic equality (5.26), as did the averaged coefficient $\overline{\mathrm{V}^{(a,b)}_{(1,1)}(t)}$. For inclusion of averaged coefficients corresponding to higher-order expansions, additional algebraic restrictions are needed to keep the effect of the inputs isolated. These restrictions will also affect the previous construction, if both second and third-order effects are to be simultaneously included. To motivate the use of these additional constraints, consider the following example that involves second-order averaging.

Example
Consider the case of a three-input underactuated driftless control system that requires second-order averaging to recover the iterated Jacobi-Lie brackets or symmetric products required for control. To second-order, the important averaged coefficients are

$$\overline{\mathrm{V}^{(a,b)}_{(1,1)}(t)}, \quad \overline{\mathrm{V}^{(a,b)}_{(2,1)}(t)}, \qquad \text{and} \quad \overline{\mathrm{V}^{(a,b,c)}_{(2,1,1)}(t)}.$$

Because the former two averaged coefficients have been analyzed, the focus will be on the latter averaged coefficient. The possible input choices are

$$\begin{array}{ccc} v^1(t) & v^2(t) & v^3(t) \\ \hline \beta^1\sin(\omega_1 t) & \beta^2\sin(\omega_2 t) & \beta^3\sin(\omega_3 t) \\ \beta^1\cos(\omega_1 t) & \beta^2\cos(\omega_2 t) & \beta^3\cos(\omega_3 t) \end{array},$$

whose eight permutations lead to the contributions found in Table 5.5, with potential coupling found in Table 5.6.

Based on Table 5.5, the important algebraic equalities that influence the averaged coefficients are as follows,

$$\omega_1 + \omega_2 - \omega_3 = 0, \ \omega_1 - \omega_2 - \omega_3 = 0, \ \text{and } \omega_1 - \omega_2 + \omega_3 = 0. \tag{5.27}$$

In order to avoid first-order coupling between terms when three distinct vector fields are contributing, the following inequality needs to hold,

$$\omega_i - \omega_j \neq 0, \quad \text{for } i \neq j. \tag{5.28}$$

Failure to satisfy the inequality (5.28) will excite Jacobi-Lie brackets or symmetric products of the form, $\left[\left[Y_i^{\text{lift}}, X_S\right], \left[Y_j^{\text{lift}}, X_S\right]\right]$ or $\left\langle Y_i^{\text{lift}} \ : \ Y_j^{\text{lift}} \right\rangle$, respectively. With the above conditions met, the only non-zero combinations involve an odd number of cosines and an even number of sines. It is important to note that the algebraic inequality

$$2\omega_i - \omega_j \neq 0, \quad \text{for } i \neq j, \tag{5.29}$$

may also need to hold, according to the coupling of Table 5.6. Failure to satisfy the inequality (5.29) may excite the symmetric products corresponding to the averaged coefficients found in Table 5.6, which are symmetric products of the form, $\left\langle Y_i^{\text{lift}} \ : \ \left\langle Y_j^{\text{lift}} \ : \ Y_i^{\text{lift}} \right\rangle \right\rangle$ or $\left\langle Y_j^{\text{lift}} \ : \ \left\langle Y_i^{\text{lift}} \ : \ Y_i^{\text{lift}} \right\rangle \right\rangle$.

Let's examine the possiblities for the third-order case. If the symmetric product is a function of three distinct vector fields, i.e., $\left\langle Y_1^{\text{lift}} \ : \ \left\langle Y_2^{\text{lift}} \ : \ Y_3^{\text{lift}} \right\rangle \right\rangle$, then satisfying the algebraic relations in Equations (5.27)-(5.29) for the associated input functions, according to the Tables 5.5 and 5.6, will result in at least 3 contributions from 3 symmetric product combinations (obtained by cycling the vector fields) due to the fact

that they all satisfy the algebraic equalities. The averaged coefficients are

$$\overline{\mathrm{V}^{(i,j,k)}_{(1,1,0)}(t)} = \begin{bmatrix} \begin{bmatrix}0\\0\\0\end{bmatrix} & \begin{bmatrix}0\\0\\-\frac{\beta^1\beta^2\beta^3}{4\omega_1^2\omega_2\omega_3}\end{bmatrix} & \begin{bmatrix}0\\-\frac{\beta^1\beta^2\beta^3}{4\omega_1^2\omega_2\omega_3}\\0\end{bmatrix} \\ \begin{bmatrix}0\\0\\\frac{\beta^1\beta^2\beta^3}{4\omega_1\omega_2^2\omega_3}\end{bmatrix} & \begin{bmatrix}0\\0\\0\end{bmatrix} & \begin{bmatrix}\frac{\beta^1\beta^2\beta^3}{4\omega_1\omega_2^2\omega_3}\\0\\0\end{bmatrix} \\ \begin{bmatrix}0\\-\frac{\beta^1\beta^2\beta^3}{4\omega_1\omega_2\omega_3^2}\\0\end{bmatrix} & \begin{bmatrix}-\frac{\beta^1\beta^2\beta^3}{4\omega_1\omega_2\omega_3^2}\\0\\0\end{bmatrix} & \begin{bmatrix}0\\0\\0\end{bmatrix} \end{bmatrix}. \tag{5.30}$$

where $\omega_2 > \omega_3 > \omega_1$ and $i, j, k \in \{1, 2, 3\}$. The matrix representing the averaged coefficient is read as follows: i gives the outer row, j gives the column, and k gives the inner row. When multiplied by the corresponding symmetric products, then summed over all indices,

$$\begin{aligned}
&\overline{\mathrm{V}^{(i,j,k)}_{(1,1,0)}(t)} \left[Y_i^{\mathrm{lift}}, \left[Y_j^{\mathrm{lift}}, Y_k^{\mathrm{lift}}\right]\right] \\
&= \\
&-\frac{\beta^1\beta^2\beta^3}{4\omega_1^2\omega_2\omega_3}\left\langle Y_1^{\mathrm{lift}} \ : \ \left\langle Y_2^{\mathrm{lift}} \ : \ Y_3^{\mathrm{lift}} \right\rangle \right\rangle - \frac{\beta^1\beta^2\beta^3}{4\omega_1^2\omega_2\omega_3}\left\langle Y_1^{\mathrm{lift}} \ : \ \left\langle Y_3^{\mathrm{lift}} \ : \ Y_2^{\mathrm{lift}} \right\rangle \right\rangle \\
&+\frac{\beta^1\beta^2\beta^3}{4\omega_1\omega_2^2\omega_3}\left\langle Y_2^{\mathrm{lift}} \ : \ \left\langle Y_3^{\mathrm{lift}} \ : \ Y_1^{\mathrm{lift}} \right\rangle \right\rangle + \frac{\beta^1\beta^2\beta^3}{4\omega_1\omega_2^2\omega_3}\left\langle Y_2^{\mathrm{lift}} \ : \ \left\langle Y_1^{\mathrm{lift}} \ : \ Y_3^{\mathrm{lift}} \right\rangle \right\rangle \\
&-\frac{\beta^1\beta^2\beta^3}{4\omega_1\omega_2\omega_3^2}\left\langle Y_3^{\mathrm{lift}} \ : \ \left\langle Y_1^{\mathrm{lift}} \ : \ Y_2^{\mathrm{lift}} \right\rangle \right\rangle - \frac{\beta^1\beta^2\beta^3}{4\omega_1\omega_2\omega_3^2}\left\langle Y_3^{\mathrm{lift}} \ : \ \left\langle Y_2^{\mathrm{lift}} \ : \ Y_1^{\mathrm{lift}} \right\rangle \right\rangle \\
&= \\
&\frac{\beta^1\beta^2\beta^3}{2\omega_1\omega_2\omega_3}\left(-\frac{1}{\omega_1}\left\langle Y_1^{\mathrm{lift}} \ : \ \left\langle Y_2^{\mathrm{lift}} \ : \ Y_3^{\mathrm{lift}} \right\rangle \right\rangle + \frac{1}{\omega_2}\left\langle Y_2^{\mathrm{lift}} \ : \ \left\langle Y_3^{\mathrm{lift}} \ : \ Y_1^{\mathrm{lift}} \right\rangle \right\rangle \right. \\
&\qquad\qquad \left. -\frac{1}{\omega_3}\left\langle Y_3^{\mathrm{lift}} \ : \ \left\langle Y_1^{\mathrm{lift}} \ : \ Y_2^{\mathrm{lift}} \right\rangle \right\rangle\right).
\end{aligned} \tag{5.31}$$

It is not possible to simplify this contribution without introducing coupling to other Jacobi-Lie bracket and symmetric product terms.
□

Lemma 16 *For the case of three distinct vector fields entering into the second-order iterated symmetric product, $\left\langle Y_a^{\mathrm{lift}} \ : \ \left\langle Y_b^{\mathrm{lift}} \ : \ Y_c^{\mathrm{lift}} \right\rangle \right\rangle$, ($a \neq b$, $a \neq c$, and $b \neq c$), no choice of inputs will result in the excitation of motion along a single symmetric product direction. If the following inputs are chosen,*

$$v^a(t) = \beta^a_{abc}\cos(\omega_{abc}t),\ v^b(t) = \beta^b_{abc}\sin(3\omega_{abc}t),\ v^c(t) = \beta^c_{abc}\sin(2\omega_{abc}t),$$

for some principle carrier frequency, ω_{abc}, then only the symmetric product $\left\langle Y_a^{\mathrm{lift}} \ : \ \left\langle Y_b^{\mathrm{lift}} \ : \ Y_c^{\mathrm{lift}} \right\rangle \right\rangle$ and its cyclicly related symmetric products, $\left\langle Y_c^{\mathrm{lift}} \ : \ \left\langle Y_a^{\mathrm{lift}} \ : \ Y_b^{\mathrm{lift}} \right\rangle \right\rangle$ and $\left\langle Y_b^{\mathrm{lift}} \ : \ \left\langle Y_c^{\mathrm{lift}} \ : \ Y_a^{\mathrm{lift}} \right\rangle \right\rangle$, will be

$v^1(t), v^2(t), v^3(t)$		$\overline{\mathrm{V}^{(1,2,3)}_{(2,1,1)}(t)}$
$\beta^1 \sin(\omega_1 t),\ \beta^2 \sin(\omega_2 t),\ \beta^3 \sin(\omega_3 t)$	$=$	0
$\beta^1 \sin(\omega_1 t),\ \beta^2 \sin(\omega_2 t),\ \beta^3 \cos(\omega_3 t)$	$=$	$\begin{cases} \frac{\beta^1\beta^2\beta^3}{4\omega_1^2\omega_2\omega_3} & \text{if } \omega_1 - \omega_2 - \omega_3 = 0, \\ & \quad \omega_1 + \omega_2 - \omega_3 = 0 \\ -\frac{\beta^1\beta^2\beta^3}{4\omega_1^2\omega_2\omega_3} & \text{if } \omega_1 - \omega_2 + \omega_3 = 0 \\ 0 & \text{otherwise} \end{cases}$
$\beta^1 \sin(\omega_1 t),\ \beta^2 \cos(\omega_2 t),\ \beta^3 \sin(\omega_3 t)$	$=$	$\begin{cases} \frac{\beta^1\beta^2\beta^3}{4\omega_1^2\omega_2\omega_3} & \text{if } \omega_1 - \omega_2 - \omega_3 = 0, \\ & \quad \omega_1 - \omega_2 + \omega_3 = 0 \\ -\frac{\beta^1\beta^2\beta^3}{4\omega_1^2\omega_2\omega_3} & \text{if } \omega_1 + \omega_2 - \omega_3 = 0 \\ 0 & \text{otherwise} \end{cases}$
$\beta^1 \sin(\omega_1 t),\ \beta^2 \cos(\omega_2 t),\ \beta^3 \cos(\omega_3 t)$	$=$	0
$\beta^1 \cos(\omega_1 t),\ \beta^2 \sin(\omega_2 t),\ \beta^3 \sin(\omega_3 t)$	$=$	$\begin{cases} & \quad \omega_1 + \omega_2 - \omega_3 = 0, \\ -\frac{\beta^1\beta^2\beta^3}{4\omega_1^2\omega_2\omega_3} & \text{if } \omega_1 - \omega_2 - \omega_3 = 0, \\ & \quad \omega_1 - \omega_2 + \omega_3 = 0 \\ 0 & \text{otherwise} \end{cases}$
$\beta^1 \cos(\omega_1 t),\ \beta^2 \sin(\omega_2 t),\ \beta^3 \cos(\omega_3 t)$	$=$	0
$\beta^1 \cos(\omega_1 t),\ \beta^2 \cos(\omega_2 t),\ \beta^3 \sin(\omega_3 t)$	$=$	0
$\beta^1 \cos(\omega_1 t),\ \beta^2 \cos(\omega_2 t),\ \beta^3 \cos(\omega_3 t)$	$=$	$\begin{cases} -\frac{\beta^1\beta^2\beta^3}{4\omega_1^2\omega_2\omega_3} & \text{if } \omega_1 + \omega_2 - \omega_3 = 0, \\ & \quad \omega_1 - \omega_2 + \omega_3 = 0 \\ \frac{\beta^1\beta^2\beta^3}{4\omega_1^2\omega_2\omega_3} & \text{if } \omega_1 - \omega_2 - \omega_3 = 0 \\ 0 & \text{otherwise} \end{cases}$

Table 5.5: Second-order averaged coefficient, $\overline{\mathrm{V}^{(a,b,c)}_{(2,1,1)}(t)}$, for systems with drift.

$v^a(t), v^b(t)$		$\overline{\mathrm{V}^{(i,j,k)}_{(2,2,0)}(t)},\ i,j,k \in \{a,b\}$	
$\beta^a \sin(\omega_a t),\ \beta^b \sin(\omega_b t)$	$=$	0	
$\beta^a \sin(\omega_a t),\ \beta^b \cos(\omega_b t),$	$=$	$\overline{\mathrm{V}^{(a,a,b)}_{(2,1,1)}(t)} = \frac{(\beta^a)^2\beta^b}{4\omega_a^3\omega_b}$	if $2\omega_a = \omega_b$
		$\overline{\mathrm{V}^{(a,b,a)}_{(2,1,1)}(t)} = \frac{(\beta^a)^2\beta^b}{4\omega_a^3\omega_b}$	if $2\omega_a = \omega_b$
		$\overline{\mathrm{V}^{(b,a,a)}_{(2,1,1)}(t)} = -\frac{(\beta^a)^2\beta^b}{4\omega_a^2\omega_b^2}$	if $2\omega_a = \omega_b$
		0	otherwise
$\beta^a \cos(\omega_a t),\ \beta^b \sin(\omega_b t),$	$=$	$\overline{\mathrm{V}^{(b,b,a)}_{(2,1,1)}(t)} = \frac{\beta^a(\beta^b)^2}{4\omega_a\omega_b^3}$	if $2\omega_b = \omega_a$
		$\overline{\mathrm{V}^{(b,a,b)}_{(2,1,1)}(t)} = \frac{\beta^a(\beta^b)^2}{4\omega_a\omega_b^3}$	if $2\omega_b = \omega_a$
		$\overline{\mathrm{V}^{(a,b,b)}_{(2,1,1)}(t)} = -\frac{\beta^a(\beta^b)^2}{4\omega_a^2\omega_b^2}$	if $2\omega_b = \omega_a$
		0	otherwise
$\beta^a \cos(\omega_a t),\ \beta^b \cos(\omega_b t),$	$=$	$\overline{\mathrm{V}^{(a,a,b)}_{(2,1,1)}(t)} = -\frac{(\beta^a)^2\beta^b}{4\omega_a^3\omega_b}$	if $2\omega_a = \omega_b$
		$\overline{\mathrm{V}^{(a,b,a)}_{(2,1,1)}(t)} = -\frac{(\beta^a)^2\beta^b}{4\omega_a^3\omega_b}$	if $2\omega_a = \omega_b$
		$\overline{\mathrm{V}^{(b,a,a)}_{(2,1,1)}(t)} = \frac{(\beta^a)^2\beta^b}{4\omega_a^2\omega_b^2}$	if $2\omega_a = \omega_b$
		$\overline{\mathrm{V}^{(b,b,a)}_{(2,1,1)}(t)} = -\frac{\beta^a(\beta^b)^2}{4\omega_a\omega_b^3}$	if $2\omega_b = \omega_a$
		$\overline{\mathrm{V}^{(b,a,b)}_{(2,1,1)}(t)} = -\frac{\beta^a(\beta^b)^2}{4\omega_a\omega_b^3}$	if $2\omega_b = \omega_a$
		$\overline{\mathrm{V}^{(a,b,b)}_{(2,1,1)}(t)} = \frac{\beta^a(\beta^b)^2}{4\omega_a^2\omega_b^2}$	if $2\omega_b = \omega_a$
		0	otherwise

Table 5.6: Coupling of second-order averaged coefficient, $\overline{\mathrm{V}^{(a,b,c)}_{(2,1,1)}(t)}$, for systems with drift.

excited. The corresponding averaged coefficients will be

$$\overline{\mathrm{V}^{(a,b,c)}_{(2,1,1)}(t)} = -\frac{\beta abc}{24\omega_{abc}^2} \quad \textit{and} \quad \overline{\mathrm{V}^{(b,c,a)}_{(1,1,0)}(t)} = \frac{\beta abc}{36\omega_{abc}^2} \quad \textit{and} \quad \overline{\mathrm{V}^{(c,a,b)}_{(2,1,1)}(t)} = -\frac{\beta abc}{12\omega_{abc}^2}$$

where $\beta abc = \beta^a_{abc}\beta^b_{abc}\beta^c_{abc}$.

proof

Assume for now that, v^a, v^b, and v^c, are the only nonzero inputs to the system. We will examine the corresponding averaged coefficient. Since the other inputs are set to zero, the system behaves similarly to the three-input example above. Without loss of generality, assume that $a = 1$, $b = 2$, and $c = 3$.

Let the input functions be

$$v^1(t) = \beta^1 \cos(\omega_1 t),\ v^2(t) = \beta^2 \sin(\omega_2 t),\ v^3(t) = \beta^3 \sin(\omega_3 t).$$

Assume that the algebraic inequality in Equation (5.29) for the coupling found in Table 5.6 is satisfied, and no coupling occurs. The critical elements of the averaged coefficient are the same as in Equation (5.30), where $\omega_2 > \omega_3 > \omega_1$ and $i, j, k \in \{\, 1, 2, 3 \,\}$. From Equation (5.31), the excited Jacobi-Lie brackets are

$$\begin{aligned}\overline{\mathrm{V}^{(i,j,k)}_{(1,1,0)}(t)} \left[Y_i^{\mathrm{lift}}, \left[Y_j^{\mathrm{lift}}, Y_k^{\mathrm{lift}}\right]\right] = &\frac{\beta^1\beta^2\beta^3}{2\omega_1\omega_2\omega_3}\left(-\frac{1}{\omega_1}\left\langle Y_1^{\mathrm{lift}} \ :\ \left\langle Y_2^{\mathrm{lift}} \ :\ Y_3^{\mathrm{lift}} \right\rangle \right\rangle \right.\\ &+\frac{1}{\omega_2}\left\langle Y_2^{\mathrm{lift}} \ :\ \left\langle Y_3^{\mathrm{lift}} \ :\ Y_1^{\mathrm{lift}} \right\rangle \right\rangle\\ &\left.-\frac{1}{\omega_3}\left\langle Y_3^{\mathrm{lift}} \ :\ \left\langle Y_1^{\mathrm{lift}} \ :\ Y_2^{\mathrm{lift}} \right\rangle \right\rangle\right).\end{aligned} \tag{5.32}$$

The symmetric products cannot be manipulated without introducing coupling to other Jacobi-Lie bracket and symmetric product terms. Choosing $\omega_3 = 2\omega_1 = 2\omega$ will avoid the coupling in Table 5.6 according to the algebraic inequality in Equation (5.29). The final contribution is

$$\begin{aligned}\overline{\mathrm{V}^{(a,b,a)}_{(1,1,0)}(t)}\,[Y_a,[Y_b,Y_a]] = \frac{\beta^1\beta^2\beta^3}{144\omega^4}&\left(-3\left\langle Y_1^{\mathrm{lift}} \ :\ \left\langle Y_2^{\mathrm{lift}} \ :\ Y_3^{\mathrm{lift}} \right\rangle \right\rangle + 2\left\langle Y_2^{\mathrm{lift}} \ :\ \left\langle Y_3^{\mathrm{lift}} \ :\ Y_1^{\mathrm{lift}} \right\rangle \right\rangle \right.\\ &\left.-6\left\langle Y_3^{\mathrm{lift}} \ :\ \left\langle Y_1^{\mathrm{lift}} \ :\ Y_2^{\mathrm{lift}} \right\rangle \right\rangle\right).\end{aligned}$$

Suppose now that a, b, and c are arbitrary, and the carrier frequency ω_{abc} is chosen to be unique in the sense that the algebraic inequalities of Equations (5.28) and (5.29) according to Table 5.6, hold. Furthermore, the algebraic equalities in Equation (5.27) hold with no undesired coupling. The analysis implies

$$\begin{aligned}\overline{\mathrm{V}^{(d,e,f)}_{(1,1,0)}(t)}\,[Y_d,[Y_e,Y_f]] = \frac{\beta_{abc}}{144\omega_{abc}^4}&\left(-3\left\langle Y_a^{\mathrm{lift}} \ :\ \left\langle Y_b^{\mathrm{lift}} \ :\ Y_c^{\mathrm{lift}} \right\rangle \right\rangle + 2\left\langle Y_b^{\mathrm{lift}} \ :\ \left\langle Y_c^{\mathrm{lift}} \ :\ Y_a^{\mathrm{lift}} \right\rangle \right\rangle \right.\\ &\left.-6\left\langle Y_c^{\mathrm{lift}} \ :\ \left\langle Y_a^{\mathrm{lift}} \ :\ Y_b^{\mathrm{lift}} \right\rangle \right\rangle\right),\end{aligned}$$

for summation over all indices $d, e, f \in \{\, 1 \ldots m \,\}$, and where $\beta_{abc} = \beta^a_{abc}\beta^b_{abc}\beta^c_{abc}$.

At worst case, the frequency $\omega_{abc} = 2\omega_{max} + 1$ may be chosen, where ω_{max} is the maximal carrier frequency of the other sinusoidal inputs. When utilizing higher-order averaged coefficients, the frequency spacing will change to avoid coupling.

■

There is nothing that can be done to resolve the problem of exciting the cycled symmetric products in this case. The control parameter β_{abc} must be interpreted as exciting the ensemble of cycled symmetric products.

Lemma 17 *For the case of two distinct vector fields entering into the second-order iterated symmetric product,* $\left\langle Y_b^{\text{lift}} \ : \ \left\langle Y_a^{\text{lift}} \ : \ Y_b^{\text{lift}} \right\rangle \right\rangle$*, if the following inputs are chosen,*

$$v^a(t) = \beta^a_{bab} \cos(2\omega_{bab} t), \ v^b(t) = \beta^b_{bab} \sin(\omega_{bab} t),$$

for some unique principle carrier frequency, ω_{bab}*, then the brackets* $\left\langle Y_b^{\text{lift}} \ : \ \left\langle Y_a^{\text{lift}} \ : \ Y_b^{\text{lift}} \right\rangle \right\rangle$ *and* $\left\langle Y_a^{\text{lift}} \ : \ \left\langle Y_b^{\text{lift}} \ : \ Y_b^{\text{lift}} \right\rangle \right\rangle$ *will be excited. The corresponding averaged coefficients are*

$$\overline{\mathrm{V}^{(b,a,b)}_{(2,1,1)}(t)} = \frac{\beta_{bab}}{4\omega_{bab}^2} \quad \textit{and} \quad \overline{\mathrm{V}^{(a,b,b)}_{(2,1,1)}(t)} = -\frac{\beta_{bab}}{16\omega_{bab}^2},$$

where $\beta_{bab} = \beta^a_{bab} \left(\beta^b_{bab}\right)^2$.

proof
Assume that these are the only nonzero inputs to the system. Without loss of generality, assume that $a = 1$, $b = 2$. Set the input functions to be

$$v^1(t) = \beta^1 \cos(2\omega t), \ v^2(t) = \beta^2 \sin(\omega t).$$

The averaged coefficients are

$$\overline{\mathrm{V}^{(i,j,k)}_{(2,1,1)}(t)} = \begin{bmatrix} \begin{bmatrix} 0 \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ -\frac{\beta^1(\beta^2)^2}{16\omega^4} \end{bmatrix} \\ \begin{bmatrix} 0 & \frac{\beta^1(\beta^2)^2}{8\omega^4} \end{bmatrix} & \begin{bmatrix} \frac{\beta^1(\beta^2)^2}{8\omega^4} \\ 0 \end{bmatrix} \end{bmatrix},$$

where the averaged coefficient is represented as a tensor with indices $i, j, k \in \{1, 2\}$. Since these are the only non-zero inputs, the other indices, $i, j, k \in \{3, \ldots m\}$, of the averaged coefficient vanish. The averaged coefficient matrix is read as follows: i choose the outer row, j chooses the column, and k chooses the inner row. The corresponding contribution to system motion is then,

$$\overline{\mathrm{V}^{(i,j,k)}_{(2,1,1)}(t)} \, [Y_i, [Y_j, Y_k]] = \frac{\beta^1(\beta^2)^2}{16\omega^4} \Big(- \left\langle Y_1^{\text{lift}} \ : \ \left\langle Y_2^{\text{lift}} \ : \ Y_2^{\text{lift}} \right\rangle \right\rangle + 2 \left\langle Y_2^{\text{lift}} \ : \ \left\langle Y_1^{\text{lift}} \ : \ Y_2^{\text{lift}} \right\rangle \right\rangle \\ + 2 \left\langle Y_2^{\text{lift}} \ : \ \left\langle Y_2^{\text{lift}} \ : \ Y_1^{\text{lift}} \right\rangle \right\rangle \Big),$$

which simplifies to

$$\overline{\mathrm{V}^{(i,j,k)}_{(2,1,1)}(t)} \, [Y_i, [Y_j, Y_k]] = \frac{\beta^1(\beta^2)^2}{16\omega^4} \Big(4 \left\langle Y_2^{\text{lift}} \ : \ \left\langle Y_1^{\text{lift}} \ : \ Y_2^{\text{lift}} \right\rangle \right\rangle - \left\langle Y_1^{\text{lift}} \ : \ \left\langle Y_2^{\text{lift}} \ : \ Y_2^{\text{lift}} \right\rangle \right\rangle \Big),$$

for summation over the indices $i, j, k \in \{1, 2\}$. With the choice of $\omega = \omega_{212}$, $\beta^1 = \beta^1_{212}$ and $\beta^2 = \beta^2_{212}$,

$$\overline{\mathrm{V}^{(i,j,k)}_{(2,1,1)}(t)} \, [Y_i, [Y_j, Y_k]] = \frac{\beta_{212}}{16\omega_{212}^4} \Big(4 \left\langle Y_2^{\text{lift}} \ : \ \left\langle Y_1^{\text{lift}} \ : \ Y_2^{\text{lift}} \right\rangle \right\rangle - \left\langle Y_1^{\text{lift}} \ : \ \left\langle Y_2^{\text{lift}} \ : \ Y_2^{\text{lift}} \right\rangle \right\rangle \Big),$$

where $\beta_{212} = \beta^1_{212}(\beta^2_{212})^2$. More generally, when $a, b \in \{1, \ldots, m\}$, this analysis implies

$$\overline{\mathrm{V}^{(c,d,e)}_{(2,1,1)}(t)}\,[Y_c,[Y_d,Y_e]] = \frac{\beta_{bab}}{16\omega^4_{bab}}\left(4\left\langle Y_b^{\text{lift}} \;:\; \left\langle Y_a^{\text{lift}} \;:\; Y_b^{\text{lift}}\right\rangle\right\rangle - \left\langle Y_a^{\text{lift}} \;:\; \left\langle Y_b^{\text{lift}} \;:\; Y_b^{\text{lift}}\right\rangle\right\rangle\right),$$

where $\beta_{bab} = \beta^a_{bab}(\beta^b_{bab})^2$. In the general case the algebraic inequalities of Equation (5.28), and Equation (5.34) according to Table 5.6, must hold with respect to other sinusoidal inputs to avoid coupling. The algebraic equalities of Equation (5.33) must only hold for those that combine to form a Jacobi-Lie bracket contribution. Again, the frequency $\omega = 2\omega_{max} + 1$ may be chosen, where ω_{max} is the maximal carrier frequency of the other sinusoidal inputs. When higher-order Jacobi-Lie bracket contributions are desired the frequency spacing will change.

■

Although the contribution is not unique, it is a result of undesired actuation of a bad symmetric product. The excitation of the bad symmetric product $\left\langle Y_b^{\text{lift}} \;:\; Y_b^{\text{lift}}\right\rangle$ couples to the oscillatory control inputs and also excites the bracket $\left\langle Y_a^{\text{lift}} \;:\; \left\langle Y_b^{\text{lift}} \;:\; Y_b^{\text{lift}}\right\rangle\right\rangle$. Our working assumption is that the bad symmetric products can be neutralized by lower order symmetric products, which in this case would be by a control vector field.

Third-order Averaged Coefficients

The third-order averaged expansion from Table 5.2 is quite complicated. This section will examine only one of the averaged coefficients found in the averaged expansion; the averaged coefficient, $\overline{\mathrm{V}^{(a,b,c)}_{(2,2,1)}(t)}$, associated with the Jacobi-Lie bracket, $\left[\left[Y_a^{\text{lift}}, X_S\right], \left[\left[Y_b^{\text{lift}}, X_S\right], \left[Y_c^{\text{lift}}, X_S\right]\right]\right]$. We shall see that it is analogous to the averaged coefficient $\overline{\mathrm{V}^{(a,b,c)}_{(1,1,0)}(t)}$ from the driftless analysis of Section 3.1.2. It is an averaged coefficient that is sometimes found to be required for controllability. For example, this Jacobi-Lie bracket is needed in the dynamic extension to the kinematic car. The kinematic car was reviewed in Section 3.3.3.

Example (Three inputs, third–order averaging)
Consider the case of a three-input system that requires excitation of a third-order iterated Jacobi-Lie bracket. Suppose further, that the Jacobi-Lie bracket coresponding to the following averaged coefficent was required,

$$\overline{\mathrm{V}^{(a,b,c)}_{(2,2,1)}(t)}.$$

The possible input choices are

$$\begin{array}{ccc} v^1(t) & v^2(t) & v^3(t) \\ \hline \alpha^1\sin(\omega_1 t) & \alpha^2\sin(\omega_2 t) & \alpha^3\sin(\omega_3 t) \\ \alpha^1\cos(\omega_1 t) & \alpha^2\cos(\omega_2 t) & \alpha^3\cos(\omega_3 t) \end{array},$$

whose eight permutations lead to the contributions found in Table 5.7, with potential coupling found in Table 5.8. Based on Table 5.7, the important algebraic equalities that influence the averaged coefficients are as follows,

$$\omega_1 + \omega_2 - \omega_3 = 0,\ \omega_1 - \omega_2 - \omega_3 = 0,\ \text{and}\ \omega_1 - \omega_2 + \omega_3 = 0. \tag{5.33}$$

In order to avoid lower-order coupling between terms when three distinct vector fields are contributing, the inequality (5.28) needs to hold. Failure to satisfy the inequality (5.28) may excite a Jacobi-Lie bracket or symmetric product of the form, $\left[\left[Y_i^{\text{lift}}, X_S\right], \left[Y_j^{\text{lift}}, X_S\right]\right]$ or $\left\langle Y_i^{\text{lift}} \;:\; Y_j^{\text{lift}}\right\rangle$, respectively.

With the above conditions met, the only non-zero combinations involve an even number of cosines and

$v^1(t), v2(t), v^3(t)$		$\overline{\mathrm{V}^{(1,2,3)}_{(2,2,1)}(t)}$
$\alpha^1\sin(\omega_1 t),\ \alpha^2\sin(\omega_2 t),\ \alpha^3\sin(\omega_3 t)$	$=$	$\begin{cases} -\frac{\alpha^1\alpha^2\alpha^3}{4\omega_1^2\omega_2^2\omega_3} & \text{if } \omega_1-\omega_2-\omega_3=0,\ \omega_1-\omega_2+\omega_3=0 \\ \frac{\alpha^1\alpha^2\alpha^3}{4\omega_1^2\omega_2^2\omega_3} & \text{if } \omega_1+\omega_2-\omega_3=0 \\ 0 & \text{otherwise} \end{cases}$
$\alpha^1\sin(\omega_1 t),\ \alpha^2\sin(\omega_2 t),\ \alpha^3\cos(\omega_3 t)$	$=$	0
$\alpha^1\sin(\omega_1 t),\ \alpha^2\cos(\omega_2 t),\ \alpha^3\sin(\omega_3 t)$	$=$	0
$\alpha^1\sin(\omega_1 t),\ \alpha^2\cos(\omega_2 t),\ \alpha^3\cos(\omega_3 t)$	$=$	$\begin{cases} \frac{\alpha^1\alpha^2\alpha^3}{4\omega_1^2\omega_2^2\omega_3} & \text{if } \omega_1-\omega_2-\omega_3=0,\ \omega_1+\omega_2-\omega_3=0 \\ -\frac{\alpha^1\alpha^2\alpha^3}{4\omega_1^2\omega_2^2\omega_3} & \text{if } \omega_1-\omega_2+\omega_3=0 \\ 0 & \text{otherwise} \end{cases}$
$\alpha^1\cos(\omega_1 t),\ \alpha^2\sin(\omega_2 t),\ \alpha^3\sin(\omega_3 t)$	$=$	0
$\alpha^1\cos(\omega_1 t),\ \alpha^2\sin(\omega_2 t),\ \alpha^3\cos(\omega_3 t)$	$=$	$\begin{cases} \frac{\alpha^1\alpha^2\alpha^3}{4\omega_1^2\omega_2^2\omega_3} & \text{if } \omega_1-\omega_2+\omega_3=0,\ \omega_1+\omega_2-\omega_3=0 \\ -\frac{\alpha^1\alpha^2\alpha^3}{4\omega_1^2\omega_2^2\omega_3} & \text{if } \omega_1-\omega_2-\omega_3=0 \\ 0 & \text{otherwise} \end{cases}$
$\alpha^1\cos(\omega_1 t),\ \alpha^2\cos(\omega_2 t),\ \alpha^3\sin(\omega_3 t)$	$=$	$\begin{cases} & \omega_1+\omega_2-\omega_3=0, \\ -\frac{\alpha^1\alpha^2\alpha^3}{4\omega_1\omega_2} & \text{if } \omega_1-\omega_2-\omega_3=0, \\ & \omega_1-\omega_2+\omega_3=0 \\ 0 & \text{otherwise} \end{cases}$
$\alpha^1\cos(\omega_1 t),\ \alpha^2\cos(\omega_2 t),\ \alpha^3\cos(\omega_3 t)$	$=$	0

Table 5.7: Third-order averaged coefficient, $\overline{\mathrm{V}^{(a,b,c)}_{(2,2,1)}(t)}$, for system with drift.

$v^a(t), v^b(t)$		$\overline{\mathrm{V}^{(i,j,k)}_{(2,2,1)}(t)},\ i,j,k \in \{a,b\}$	
$\alpha^a \sin(\omega_a t),\ \alpha^b \sin(\omega_b t)$	$=$	$\overline{\mathrm{V}^{(a,a,b)}_{(2,2,1)}(t)} = \frac{(\alpha^a)^2\alpha^b}{4\omega_a^4\omega_b}$	if $2\omega_a = \omega_b$
		$\overline{\mathrm{V}^{(a,b,a)}_{(2,2,1)}(t)} = -\frac{(\alpha^a)^2\alpha^b}{4\omega_a^3\omega_b^2}$	if $2\omega_a = \omega_b$
		$\overline{\mathrm{V}^{(b,a,a)}_{(2,2,1)}(t)} = -\frac{(\alpha^a)^2\alpha^b}{4\omega_a^3\omega_b^2}$	if $2\omega_a = \omega_b$
		$\overline{\mathrm{V}^{(b,b,a)}_{(2,2,1)}(t)} = -\frac{\alpha^a(\alpha^b)^2}{4\omega_a\omega_b^4}$	if $2\omega_b = \omega_a$
		$\overline{\mathrm{V}^{(b,a,b)}_{(2,2,1)}(t)} = -\frac{\alpha^a(\alpha^b)^2}{4\omega_a^2\omega_b^3}$	if $2\omega_b = \omega_a$
		$\overline{\mathrm{V}^{(a,b,b)}_{(2,2,1)}(t)} = \frac{\alpha^a(\alpha^b)^2}{4\omega_a^2\omega_b^3}$	if $2\omega_b = \omega_a$
		0	otherwise
$\alpha^a \sin(\omega_a t),\ \alpha^b \cos(\omega_b t)$	$=$	$\overline{\mathrm{V}^{(b,b,a)}_{(2,2,1)}(t)} = -\frac{\alpha^a(\alpha^b)^2}{4\omega_a\omega_b^4}$	if $2\omega_b = \omega_a$
		$\overline{\mathrm{V}^{(b,a,b)}_{(2,2,1)}(t)} = \frac{\alpha^a(\alpha^b)^2}{4\omega_a^2\omega_b^3}$	if $2\omega_b = \omega_a$
		$\overline{\mathrm{V}^{(a,b,b)}_{(2,2,1)}(t)} = \frac{\alpha^a(\alpha^b)^2}{4\omega_a^2\omega_b^3}$	if $2\omega_b = \omega_a$
		0	otherwise
$\alpha^a \cos(\omega_a t),\ \alpha^b \sin(\omega_b t)$	$=$	$\overline{\mathrm{V}^{(a,a,b)}_{(2,2,1)}(t)} = -\frac{(\alpha^a)^2\alpha^b}{4\omega_a^4\omega_b}$	if $2\omega_a = \omega_b$
		$\overline{\mathrm{V}^{(a,b,a)}_{(2,2,1)}(t)} = \frac{(\alpha^a)^2\alpha^b}{4\omega_a^3\omega_b^2}$	if $2\omega_a = \omega_b$
		$\overline{\mathrm{V}^{(b,a,a)}_{(2,2,1)}(t)} = \frac{(\alpha^a)^2\alpha^b}{4\omega_a^3\omega_b^2}$	if $2\omega_a = \omega_b$
		0	otherwise
$\alpha^a \cos(\omega_a t),\ \alpha^b \cos(\omega_b t)$	$=$	0	

Table 5.8: Coupling of third-order averaged coefficient, $\overline{\mathrm{V}^{(a,b,c)}_{(2,2,1)}(t)}$, for system with drift.

an odd number of sines. It is important to note that the algebraic inequality

$$2\omega_i - \omega_j \neq 0, \quad \text{for } i \neq j, \tag{5.34}$$

may also need to hold, according to the coupling of Table 5.8. Failure to satisfy the inequality (5.34) may excite the Jacobi-Lie bracket corresponding to the averaged coefficients found in Table 5.8, which take the form, $\left[\left[Y_j^{\text{lift}}, X_S\right], \left[\left[Y_i^{\text{lift}}, X_S\right], \left[Y_j^{\text{lift}}, X_S\right]\right]\right]$.

Let's examine the possiblities for the third-order case. If the Jacobi-Lie bracket is a function of three distinct vector fields, i.e., $\left[\left[Y_1^{\text{lift}}, X_S\right], \left[\left[Y_2^{\text{lift}}, X_S\right], \left[Y_3^{\text{lift}}, X_S\right]\right]\right]$, then satisfying the algebraic relations in Equations (5.28), (5.33), and (5.34) for the associated input functions, according to the Tables 5.7 and 5.8, will result in at least 3 contributions from 3 Jacobi-Lie bracket combinations (obtained by cycling the vector fields) due to the fact that they all satisfy the algebraic equalities. The averaged coefficients are

$$\overline{\mathrm{V}_{(2,2,1)}^{(i,j,k)}(t)} = \begin{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ 0 \\ -\frac{\alpha^1\alpha^2\alpha^3}{4\omega_1^2\omega_2^2\omega_3} \end{bmatrix} & \begin{bmatrix} 0 \\ \frac{\alpha^1\alpha^2\alpha^3}{4\omega_1^2\omega_2\omega_3^2} \\ 0 \end{bmatrix} \\ \begin{bmatrix} 0 \\ 0 \\ -\frac{\alpha^1\alpha^2\alpha^3}{4\omega_1^2\omega_2^2\omega_3} \end{bmatrix} & \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} & \begin{bmatrix} -\frac{\alpha^1\alpha^2\alpha^3}{4\omega_1\omega_2^2\omega_3^2} \\ 0 \\ 0 \end{bmatrix} \\ \begin{bmatrix} 0 \\ \frac{\alpha^1\alpha^2\alpha^3}{4\omega_1^2\omega_2\omega_3^2} \\ 0 \end{bmatrix} & \begin{bmatrix} -\frac{\alpha^1\alpha^2\alpha^3}{4\omega_1\omega_2^2\omega_3^2} \\ 0 \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \end{bmatrix}. \tag{5.35}$$

where $\omega_2 > \omega_3 > \omega_1$ and $i, j, k \in \{1, 2, 3\}$. The matrix representing the averaged coefficient is read as follows: i gives the outer row, j gives the column, and k gives the inner row. When multiplied by the corresponding Jacobi-Lie brackets, then summed over all indices,

$$\begin{aligned}
&\overline{\mathrm{V}_{(2,2,1)}^{(i,j,k)}(t)} \left[\left[Y_i^{\text{lift}}, X_S\right], \left[\left[Y_j^{\text{lift}}, X_S\right], \left[Y_k^{\text{lift}}, X_S\right]\right]\right] \\
&= \\
&\frac{\alpha^1\alpha^2\alpha^3}{4\omega_1^2\omega_2^2\omega_3} \left[\left[Y_1^{\text{lift}}, X_S\right], \left[\left[Y_2^{\text{lift}}, X_S\right], \left[Y_3^{\text{lift}}, X_S\right]\right]\right] - \frac{\alpha^1\alpha^2\alpha^3}{4\omega_1^2\omega_2\omega_3^2} \left[\left[Y_1^{\text{lift}}, X_S\right], \left[\left[Y_3^{\text{lift}}, X_S\right], \left[Y_2^{\text{lift}}, X_S\right]\right]\right] \\
&+ \frac{\alpha^1\alpha^2\alpha^3}{4\omega_1^2\omega_2^2\omega_3} \left[\left[Y_2^{\text{lift}}, X_S\right], \left[\left[Y_1^{\text{lift}}, X_S\right], \left[Y_3^{\text{lift}}, X_S\right]\right]\right] + \frac{\alpha^1\alpha^2\alpha^3}{4\omega_1\omega_2^2\omega_3^2} \left[\left[Y_2^{\text{lift}}, X_S\right], \left[\left[Y_3^{\text{lift}}, X_S\right], \left[Y_1^{\text{lift}}, X_S\right]\right]\right] \\
&- \frac{\alpha^1\alpha^2\alpha^3}{4\omega_1^2\omega_2\omega_3^2} \left[\left[Y_3^{\text{lift}}, X_S\right], \left[\left[Y_1^{\text{lift}}, X_S\right], \left[Y_2^{\text{lift}}, X_S\right]\right]\right] + \frac{\alpha^1\alpha^2\alpha^3}{4\omega_1\omega_2^2\omega_3^2} \left[\left[Y_3^{\text{lift}}, X_S\right], \left[\left[Y_2^{\text{lift}}, X_S\right], \left[Y_1^{\text{lift}}, X_S\right]\right]\right] \\
&= \\
&\frac{\alpha^1\alpha^2\alpha^3}{4\omega_1\omega_2\omega_3} \left(\left(\frac{1}{\omega_1\omega_2} + \frac{1}{\omega_1\omega_3} \right) \left[\left[Y_1^{\text{lift}}, X_S\right], \left[\left[Y_2^{\text{lift}}, X_S\right], \left[Y_3^{\text{lift}}, X_S\right]\right]\right] \right. \\
&\qquad + \left(\frac{1}{\omega_2\omega_3} - \frac{1}{\omega_1\omega_2} \right) \left[\left[Y_2^{\text{lift}}, X_S\right], \left[\left[Y_3^{\text{lift}}, X_S\right], \left[Y_1^{\text{lift}}, X_S\right]\right]\right] \\
&\qquad \left. - \left(\frac{1}{\omega_1\omega_3} + \frac{1}{\omega_2\omega_3} \right) \left[\left[Y_3^{\text{lift}}, X_S\right], \left[\left[Y_1^{\text{lift}}, X_S\right], \left[Y_2^{\text{lift}}, X_S\right]\right]\right] \right).
\end{aligned} \tag{5.36}$$

Using the Jacobi identity, it is possible to reduce this number.
□

Lemma 18 *For the case of three distinct vector fields entering into the third-order iterated Jacobi-Lie bracket,* $\left[\left[Y_a^{\text{lift}}, X_S\right], \left[\left[Y_b^{\text{lift}}, X_S\right], \left[Y_c^{\text{lift}}, X_S\right]\right]\right]$*, (*$a \neq b$*,* $a \neq c$*, and* $b \neq c$*), no choice of inputs will result in the excitation of motion along a single bracket direction. If the following inputs are chosen,*

$$v^a(t) = \alpha^a_{abc} \sin(\omega_{abc} t),\ v^b(t) = \alpha^b_{abc} \cos(3\omega_{abc} t),\ v^c(t) = \alpha^c_{abc} \cos(2\omega_{abc} t),$$

for some principle carrier frequency, ω_{abc}*, then only the bracket* $\left[\left[Y_a^{\text{lift}}, X_S\right], \left[\left[Y_b^{\text{lift}}, X_S\right], \left[Y_c^{\text{lift}}, X_S\right]\right]\right]$ *and a cyclicly related bracket,* $\left[\left[Y_c^{\text{lift}}, X_S\right], \left[\left[Y_a^{\text{lift}}, X_S\right], \left[Y_b^{\text{lift}}, X_S\right]\right]\right]$ *or* $\left[\left[Y_b^{\text{lift}}, X_S\right], \left[\left[Y_c^{\text{lift}}, X_S\right], \left[Y_a^{\text{lift}}, X_S\right]\right]\right]$*, will be excited. The corresponding averaged coefficients will be*

$$\overline{\mathrm{V}^{(a,b,c)}_{(2,2,1)}(t)} = \frac{3\alpha_{abc}}{48\omega^5_{abc}} \quad \textit{and} \quad \overline{\mathrm{V}^{(c,a,b)}_{(2,2,1)}(t)} = \frac{\alpha_{abc}}{48\omega^5_{abc}}$$

or

$$\overline{\mathrm{V}^{(a,b,c)}_{(2,2,1)}(t)} = \frac{\alpha_{abc}}{144\omega^5_{abc}} \quad \textit{and} \quad \overline{\mathrm{V}^{(b,c,a)}_{(2,2,1)}(t)} = -\frac{3\alpha_{abc}}{144\omega^5_{abc}},$$

where $\alpha_{abc} = \alpha^a_{abc}\alpha^b_{abc}\alpha^c_{abc}$.

proof
Assume for now that, v^a, v^b, and v^c, are the only nonzero inputs to the system. We will examine the corresponding averaged coefficient. Since the other inputs are set to zero, the system behaves similarly to the three-input example above. Without loss of generality, assume that $a = 1$, $b = 2$, and $c = 3$.

Let the input functions be

$$v^1(t) = \alpha^1 \sin(\omega_1 t),\ v^2(t) = \alpha^2 \cos(\omega_2 t),\ v^3(t) = \alpha^3 \cos(\omega_3 t).$$

Assume that the algebraic inequality in Equation (5.34) for the coupling found in Table 5.8 is satisfied, and no coupling occurs. The critical elements of the averaged coefficient are the same as in Equation (5.35), where $\omega_2 > \omega_3 > \omega_1$ and $i, j, k \in \{1, 2, 3\}$. From Equation (5.36), the excited Jacobi-Lie brackets are

$$\begin{aligned}\overline{\mathrm{V}^{(a,b,a)}_{(2,2,1)}(t)}\,[Y_a,[Y_b,Y_a]] = \frac{\alpha^1\alpha^2\alpha^3}{4\omega_1\omega_2\omega_3}\Bigg(&\left(\frac{1}{\omega_1\omega_2}+\frac{1}{\omega_1\omega_3}\right)\left[\left[Y_1^{\text{lift}}, X_S\right], \left[\left[Y_2^{\text{lift}}, X_S\right], \left[Y_3^{\text{lift}}, X_S\right]\right]\right] \\ &+\left(\frac{1}{\omega_2\omega_3}-\frac{1}{\omega_1\omega_2}\right)\left[\left[Y_2^{\text{lift}}, X_S\right], \left[\left[Y_3^{\text{lift}}, X_S\right], \left[Y_1^{\text{lift}}, X_S\right]\right]\right] \\ &-\left(\frac{1}{\omega_1\omega_3}+\frac{1}{\omega_2\omega_3}\right)\left[\left[Y_3^{\text{lift}}, X_S\right], \left[\left[Y_1^{\text{lift}}, X_S\right], \left[Y_2^{\text{lift}}, X_S\right]\right]\right]\Bigg).\end{aligned} \tag{5.37}$$

Denote $\hat{\omega}_i$ by

$$\hat{\omega}_1 = \frac{1}{\omega_1}\left(\frac{1}{\omega_2}+\frac{1}{\omega_3}\right),\ \hat{\omega}_2 = \frac{1}{\omega_2}\left(\frac{1}{\omega_3}-\frac{1}{\omega_1}\right),\ \hat{\omega}_3 = \frac{1}{\omega_3}\left(\frac{1}{\omega_1}+\frac{1}{\omega_2}\right).$$

With these definitions,

$$\hat{\omega}_1 > 0,\ \hat{\omega}_2 < 0,\ \hat{\omega}_3 > 0.$$

The Jacobi identity may be used to remove one of the Jacobi-Lie brackets, $[Y_2,[Y_3,Y_1]]$ or $[Y_3,[Y_1,Y_2]]$. Select the last Jacobi-Lie bracket from Equation (5.37) to be removed using the Jacobi identity.

$$\overline{\mathrm{V}_{(2,2,1)}^{(i,j,k)}(t)}\left[\left[Y_i^{\mathrm{lift}},X_S\right],\left[\left[Y_j^{\mathrm{lift}},X_S\right],\left[Y_k^{\mathrm{lift}},X_S\right]\right]\right]= \frac{\alpha^1\alpha^2\alpha^3}{4\omega_1\omega_2\omega_3}\Big((\hat{\omega}_1+\hat{\omega}_3)\left[\left[Y_1^{\mathrm{lift}},X_S\right],\left[\left[Y_2^{\mathrm{lift}},X_S\right],\left[Y_3^{\mathrm{lift}},X_S\right]\right]\right] +(\hat{\omega}_2+\hat{\omega}_3)\left[\left[Y_2^{\mathrm{lift}},X_S\right],\left[\left[Y_3^{\mathrm{lift}},X_S\right],\left[Y_1^{\mathrm{lift}},X_S\right]\right]\right]\Big). \quad (5.38)$$

The second Jacobi-Lie bracket in Equation (5.38) can be cancelled only if there exists a selection of ω_i satisfying one of the tensor equalities (5.33), such that (3.31) holds. It was already shown that (3.31) would not lead to feasible frequency selections.

As for the choice of inputs given above, notice that selecting any $\omega_2>\omega_3>\omega_1$ such that the equality $\omega_2=\omega_1+\omega_3$ holds will excite only two two Jacobi-Lie brackets. Choosing $\omega_3=2\omega_1=2\omega$ will avoid the coupling in Table 5.8 according to the algebraic inequality in Equation (5.34). The important decision of the frequency assignments is to have the sine carrier frequency be the smallest to avoid the coupling. The final contribution is

$$\overline{\mathrm{V}_{(2,2,1)}^{(a,b,c)}(t)}\,[Y_a,[Y_b,Y_a]]\left[\left[Y_a^{\mathrm{lift}},X_S\right],\left[\left[Y_b^{\mathrm{lift}},X_S\right],\left[Y_c^{\mathrm{lift}},X_S\right]\right]\right]= \frac{\alpha^1\alpha^2\alpha^3}{48\omega^5}\Big(3\left[\left[Y_1^{\mathrm{lift}},X_S\right],\left[\left[Y_2^{\mathrm{lift}},X_S\right],\left[Y_3^{\mathrm{lift}},X_S\right]\right]\right] +\left[\left[Y_2^{\mathrm{lift}},X_S\right],\left[\left[Y_3^{\mathrm{lift}},X_S\right],\left[Y_1^{\mathrm{lift}},X_S\right]\right]\right]\Big).$$

Suppose that, instead, the second Jacobi-Lie bracket from Equation (5.37) were to be removed using the Jacobi identity. Then the important averaged coefficients are

$$\overline{\mathrm{V}_{(1,1,0)}^{(a,b,c)}(t)}\left[\left[Y_a^{\mathrm{lift}},X_S\right],\left[\left[Y_b^{\mathrm{lift}},X_S\right],\left[Y_c^{\mathrm{lift}},X_S\right]\right]\right]= \frac{\alpha^1\alpha^2\alpha^3}{4\omega_1\omega_2\omega_3}\Big((\hat{\omega}_1-\hat{\omega}_2)\left[\left[Y_1^{\mathrm{lift}},X_S\right],\left[\left[Y_2^{\mathrm{lift}},X_S\right],\left[Y_3^{\mathrm{lift}},X_S\right]\right]\right] -(\hat{\omega}_2+\hat{\omega}_3)\left[\left[Y_3^{\mathrm{lift}},X_S\right],\left[\left[Y_1^{\mathrm{lift}},X_S\right],\left[Y_2^{\mathrm{lift}},X_S\right]\right]\right]\Big). \quad (5.39)$$

The second JAcobi-Lie bracket in Equation (5.39) can be cancelled only if there exists a selection of ω_i satisfying one of the tensor equalities (5.33), such that Equation (3.31) holds, which is not possible.

Again, choosing $\omega_3=2\omega_1=2\omega$ will avoid the coupling in Table 5.8 according to the algebraic inequality in Equation (5.34). The final contribution is

$$\overline{\mathrm{V}_{(2,2,1)}^{(a,b,a)}(t)}\left[\left[Y_a^{\mathrm{lift}},X_S\right],\left[\left[Y_b^{\mathrm{lift}},X_S\right],\left[Y_c^{\mathrm{lift}},X_S\right]\right]\right]= \frac{\alpha^1\alpha^2\alpha^3}{144\omega^5}\Big(\left[\left[Y_1^{\mathrm{lift}},X_S\right],\left[\left[Y_2^{\mathrm{lift}},X_S\right],\left[Y_3^{\mathrm{lift}},X_S\right]\right]\right] -3\left[\left[Y_2^{\mathrm{lift}},X_S\right],\left[\left[Y_3^{\mathrm{lift}},X_S\right],\left[Y_1^{\mathrm{lift}},X_S\right]\right]\right]\Big).$$

Suppose now that a, b, and c are arbitrary, and the carrier frequency ω_{abc} is chosen to be unique in the sense that the algebraic inequalities of Equation (5.28), and Equation (5.34) according to Table 5.8, hold. Furthermore, the algebraic equalities in Equation (5.33) hold with no undesired coupling. The analysis

implies that either

$$\overline{\mathrm{V}^{(d,e,f)}_{(2,2,1)}(t)}\left[\left[Y_d^{\mathrm{lift}},X_S\right],\left[\left[Y_e^{\mathrm{lift}},X_S\right],\left[Y_f^{\mathrm{lift}},X_S\right]\right]\right] = \frac{\alpha_{abc}}{48\omega^5}\Big(3\left[\left[Y_a^{\mathrm{lift}},X_S\right],\left[\left[Y_b^{\mathrm{lift}},X_S\right],\left[Y_c^{\mathrm{lift}},X_S\right]\right]\right] + \left[\left[Y_b^{\mathrm{lift}},X_S\right],\left[\left[Y_c^{\mathrm{lift}},X_S\right],\left[Y_a^{\mathrm{lift}},X_S\right]\right]\right]\Big),$$

or

$$\overline{\mathrm{V}^{(d,e,f)}_{(2,2,1)}(t)}\left[\left[Y_d^{\mathrm{lift}},X_S\right],\left[\left[Y_e^{\mathrm{lift}},X_S\right],\left[Y_f^{\mathrm{lift}},X_S\right]\right]\right] = \frac{\alpha_{abc}}{144\omega^5}\Big(\left[\left[Y_a^{\mathrm{lift}},X_S\right],\left[\left[Y_b^{\mathrm{lift}},X_S\right],\left[Y_c^{\mathrm{lift}},X_S\right]\right]\right] -3\left[\left[Y_b^{\mathrm{lift}},X_S\right],\left[\left[Y_c^{\mathrm{lift}},X_S\right],\left[Y_a^{\mathrm{lift}},X_S\right]\right]\right]\Big).$$

for summation over all indices $d, e, f \in \{\, 1 \ldots m \,\}$, and where $\alpha_{abc} = \alpha^a_{abc}\alpha^b_{abc}\alpha^c_{abc}$.

At worst case, the frequency $\omega_{abc} = 2\omega_{max} + 1$ may be chosen, where ω_{max} is the maximal carrier frequency of the other sinusoidal inputs. When utilizing higher-order averaged coefficients, the frequency spacing will change to avoid coupling.

■

Lemma 19 *For the case of two distinct vector fields entering into the third-order iterated Jacobi-Lie bracket,* $\left[\left[Y_b^{\mathrm{lift}},X_S\right],\left[\left[Y_a^{\mathrm{lift}},X_S\right],\left[Y_b^{\mathrm{lift}},X_S\right]\right]\right]$*, if the following inputs are chosen,*

$$v^a(t) = \alpha^a_{bab}\cos(\omega_{bab}t),\ v^b(t) = \alpha^b_{bab}\sin(2\omega_{bab}t),$$

for some unique principle carrier frequency, ω_{bab}*, then only the bracket* $\left[\left[Y_b^{\mathrm{lift}},X_S\right],\left[\left[Y_a^{\mathrm{lift}},X_S\right],\left[Y_b^{\mathrm{lift}},X_S\right]\right]\right]$ *will be excited. The corresponding averaged coefficient is*

$$\overline{\mathrm{V}^{(b,a,b)}_{(2,2,1)}(t)} = \frac{3\alpha_{bab}}{16\omega^5_{bab}},$$

where $\alpha_{bab} = \alpha^a_{bab}\left(\alpha^b_{bab}\right)^2$.

proof

Assume that these are the only nonzero inputs to the system. Without loss of generality, assume that $a = 1$, $b = 2$. Set the input functions to be

$$v^1(t) = \alpha^1\sin(2\omega t),\ v^2(t) = \alpha^2\cos(\omega t). \tag{5.40}$$

The averaged coefficients are

$$\overline{\mathrm{V}^{(i,j,k)}_{(1,1,0)}(t)} = \begin{bmatrix} \begin{bmatrix}0\\0\end{bmatrix} & \begin{bmatrix}0\\ \frac{\alpha^1(\alpha^2)^2}{16\omega^5}\end{bmatrix} \\ \begin{bmatrix}0\\ \frac{\alpha^1(\alpha^2)^2}{16\omega^5}\end{bmatrix} & \begin{bmatrix}-\frac{\alpha^1(\alpha^2)^2}{8\omega^5}\\0\end{bmatrix}\end{bmatrix}, \tag{5.41}$$

where the averaged coefficient is represented as a tensor with indices $i, j, k \in \{ 1, 2 \}$. Since these are the only non-zero inputs, the other indices, $i, j, k \in 3, \ldots m$, of the averaged coefficient vanish. The averaged coefficient matrix is read as follows: i choose the outer row, j chooses the column, and k chooses the inner row. The corresponding contribution to system motion is then,

$$\begin{aligned}
\overline{\mathrm{V}^{(i,j,k)}_{(1,1,0)}(t)} \left[\left[Y_i^{\mathrm{lift}}, X_S \right], \left[\left[Y_j^{\mathrm{lift}}, X_S \right], \left[Y_k^{\mathrm{lift}}, X_S \right] \right] \right] \\
= + \frac{\alpha^1 (\alpha^2)^2}{16 \omega^5} \left[\left[Y_1^{\mathrm{lift}}, X_S \right], \left[\left[Y_2^{\mathrm{lift}}, X_S \right], \left[Y_2^{\mathrm{lift}}, X_S \right] \right] \right] \\
+ \frac{\alpha^1 (\alpha^2)^2}{16 \omega_5} \left[\left[Y_2^{\mathrm{lift}}, X_S \right], \left[\left[Y_1^{\mathrm{lift}}, X_S \right], \left[Y_2^{\mathrm{lift}}, X_S \right] \right] \right] \\
- \frac{\alpha^1 (\alpha^2)^2}{8 \omega^5} \left[\left[Y_2^{\mathrm{lift}}, X_S \right], \left[\left[Y_2^{\mathrm{lift}}, X_S \right], \left[Y_1^{\mathrm{lift}}, X_S \right] \right] \right] \\
= \frac{3 \alpha^1 (\alpha^2)^2}{16 \omega^5} \left[\left[Y_2^{\mathrm{lift}}, X_S \right], \left[\left[Y_1^{\mathrm{lift}}, X_S \right], \left[Y_2^{\mathrm{lift}}, X_S \right] \right] \right], \quad (5.42)
\end{aligned}$$

for summation over the indices $i, j, k \in \{ 1, 2 \}$. With the choice of $\omega = \omega_{212}$, $\alpha^1 = \alpha^1_{212}$ and $\alpha^2 = \alpha^2_{212}$,

$$\overline{\mathrm{V}^{(i,j,k)}_{(1,1,0)}(t)} \left[\left[Y_i^{\mathrm{lift}}, X_S \right], \left[\left[Y_j^{\mathrm{lift}}, X_S \right], \left[Y_k^{\mathrm{lift}}, X_S \right] \right] \right] = \frac{3 \alpha_{212}}{16 \omega^5_{212}} \left[\left[Y_2^{\mathrm{lift}}, X_S \right], \left[\left[Y_1^{\mathrm{lift}}, X_S \right], \left[Y_2^{\mathrm{lift}}, X_S \right] \right] \right] \quad (5.43)$$

where $\alpha_{212} = \alpha^1_{212} (\alpha^2_{212})^2$. More generally, when $a, b \in \{ 1, \ldots, m \}$, this analysis implies

$$\overline{\mathrm{V}^{(c,d,e)}_{(1,1,0)}(t)} \left[\left[Y_c^{\mathrm{lift}}, X_S \right], \left[\left[Y_d^{\mathrm{lift}}, X_S \right], \left[Y_e^{\mathrm{lift}}, X_S \right] \right] \right] = \frac{3 \alpha_{bab}}{16 \omega^5_{bab}} \left[\left[Y_b^{\mathrm{lift}}, X_S \right], \left[\left[Y_a^{\mathrm{lift}}, X_S \right], \left[Y_b^{\mathrm{lift}}, X_S \right] \right] \right] \quad (5.44)$$

where $\alpha_{bab} = \alpha^a_{bab} (\alpha^b_{bab})^2$. In the general case the algebraic inequalities of Equation (5.28), and Equation (5.34) according to Table 5.8, must hold with respect to other sinusoidal inputs to avoid coupling. The algebraic equalities of Equation (5.33) must only hold for those that combine to form a Jacobi-Lie bracket contribution. Again, the frequency $\omega = 2\omega_{max} + 1$ may be chosen, where ω_{max} is the maximal carrier frequency of the other sinusoidal inputs. When higher-order Jacobi-Lie bracket contributions are desired the frequency spacing will change.

■

Abstracting on the Sinusoidal Inputs

For higher-order expansions, myriad algebraic identities must hold in order to isolate specific bracket related motions. Each averaged coefficient must be examined to determine its contribution, and the limitations arising from the chosen set of input functions. Once a particular calculation is done, it need not be repeated for another problem with the same Jacobi-Lie bracket requirements. The general strategy for driftless systems was proven in [154, 155, 92], and also allows for the frequency choices to be positive rational numbers instead of only positive integers. One major difference in implementing the strategy for systems with drift is the excitation of bad symmetric products that cannot be avoided. This excitation will cuse additional symmetric products and Jacobi-Lie brackets to be excited unintentionally (as was seen in Lemma 16).

Th sufficient conditions found in Chapter 4 and in most other controllability analyses requires that the unintended excitation of bad brackets can be neutralized. Recall the definition of degree for a Jacobi-Lie bracket.

Definition 62 *Let B be an iterated Jacobi-Lie bracket. Define $\delta_a(B)$ to be the number of times that the*

control vector field Y_a appears in the iterated Jacobi-Lie bracket B.

In [92], for the case of underactuated, driftless nonlinear systems, it was shown that for any Jacobi-Lie bracket B such that $\delta_i(B) = \delta_j(B)$ for all $i, j = 1 \dots m$, there is no way to obtain unique activation of the Jacobi-Lie bracket B. The same property holds for systems with drift; the cyclicly related Jacobi-Lie brackets will also be excited, as was seen in Lemmas 16 and 18.

A system with drift is parametrized by α- and β-parameters. When used in conjunction, they will be denoted by $\gamma = (\alpha, \beta)$ and called the *γ-parametrization*. Occasionally, one of the two of parameter sets will be explicitly referred to, in which case, the complementary parameter set has parameters equal to zero. The time-periodic inputs are defined such that the averaged coefficients scale linearly with the parameters γ_J. For purposes of feedback, the γ_J should be defined as functions of state so as to achieve stabilization for the averaged equations of motion. There is no unique method for picking the γ_J^I to obtain the desired product $\gamma_J = \Pi_{i \in I} \gamma_J^i$. Different choices of decomposing the product γ_J into γ_J^I will result in different controllers, although the averaged evolution will be the same. Both the Theorems found below, and the examples in Section 5.3 will discuss this freedom and its implications.

5.2.3 Stabilization Using Sinusoids

To summarize, we have obtained formulas for the response of 1-homogeneous control systems to an oscillatory control at some arbitrary order. We may analyze the effects of the control inputs on the expansions, leading to an α-parametrized form. Now, we must determine a stabilization feedback strategy. For convenience, order the Jacobi-Lie brackets as they appear in the averaged expansion (symmetric products are also brackets). Let $\widehat{\mathrm{Y}}_j$ denote the Jacobi-Lie brackets, and let $\mathrm{T}^j(\alpha)$ be their corresponding averaged coefficients. With this ordering, the averaged equations can be put into the form:

$$\begin{aligned} \dot{z} &= X_S(z) + \mathrm{T}^j(\alpha)\widehat{\mathrm{Y}}_j(z) \\ &= X_S(z) + B(z)H(\alpha), \end{aligned} \tag{5.45}$$

where the matrices B and H are

$$B(z) = [\widehat{\mathrm{Y}}_1 \dots \widehat{\mathrm{Y}}_N] \text{ and } H(\alpha) = [\mathrm{T}^1 \dots \mathrm{T}^N]. \tag{5.46}$$

It is important to realize that Equation (5.45) will often represent a truncation of the full series expansion. The α parameters will be used to obtain control authority over the indirectly controllable states.

The dependence of the controller design on a local Lie algebra basis is problematic if the Jacobi-Lie brackets used to satisfy the small-time local (configuration) controllability vary pointwise. If the controller is to stabilize to different points, each with a unique basis spanning the Lie algebra, then multiple α-parametrizations are necessary [77]. For systems that are well-controllable, multiple parametizations can be avoided.

Stabilization via Discretized Feedback

We first consider feedback based on discrete sampling, while a continuous time version is developed below. State error will be used as feedback to modulate the γ-parameters, converting the problem to periodic discrete feedback. The technique is very similar to the motion control algorithms, based on motion primitives, that use open-loop approximate inversion for open-loop stabilization and trajectory tracking [21, 24, 100]. However, averaging theory is explicitly used to prove stability. In the language of Floquet theory, our goal is to utilize feedback so as to create stable Floquet multipliers.

Theorem 37 *Consider a system (5.16) which is small-time locally controllable at $x^* \in E$. Let $u^k(t)$ be the set of γ parametrized, T-periodic input functions where $k = \{1, \ldots, m\}$ and $\gamma \in \mathbb{R}^{n-m}$. Let $z(t)$, be the averaged system response to the inputs. Given the averaged system (5.45), assuming that the m directly controlled states have been linearly stabilized and that the linearization of H from Equation (5.46), with respect to γ at $\gamma = 0$ and $z = x^*$, is invertible on the $(n-m)$ dimensional subspace to control, there exists a $K \in \mathbb{R}^{(n-m)\times n}$ such that for*

$$\gamma = -\Lambda K z(T\lfloor t/T \rfloor),$$

where $\Lambda^{(n-m)\times(n-m)}$ is invertible and $\lfloor \cdot \rfloor$ denotes the floor function, the average system response is stabilized.

proof
The proof was given in [160], but will be quickly sketched. Given the assumptions on the system, the averaged system (5.45) is controllable. Linearization with respect to z and γ yields

$$\dot{z} = Az + B \left.\frac{\partial H}{\partial \gamma}\right|_{\gamma=0} \gamma = Az + B\Upsilon\gamma.$$

Choosing γ constant over a period, the above system can be directly integrated to obtain a discrete, linear system

$$z(k+1) = \hat{A}z(k) + \hat{B}\gamma.$$

The assumptions imply that $\hat{B}$ has a pseudo-inverse, Λ, for the $(n-m)$-dimensional subspace to stabilize. Choose K so that the eigenvalues of $\hat{A} - K$ lie in the unit circle. This stabilizes the discrete system (i.e. the monodramy map), and the continuous system with piecewise constant feedback.
■
In the theorem it is assumed that the directly controlled state have been stabilized using continuous feedback, and a discretized feedback strategy is provided to stabilize the remaining states. At the end of each oscillation period, the error is computed and used to update the γ-parameters. The modified γ-parameters are held constant for the duration of the period, at which point a new error estimate is provided.

Stabilization via Continuous Feedback

The control design procedure outlined above required integration of the linearized dynamical model over a period of actuation, and a discretized feedback strategy. Corollary 3 implies that stability can also be determined from the Floquet exponents, i.e., the stability of the continuous autonomous averaged vector field. A version of Theorem 37 can be proven without discretizing the closed-loop system.

Theorem 38 *Consider a system of the form (5.16) which is small-time locally controllable at $x^* \in E$. Let $u^a(t)$ be the corresponding set of γ parametrized, T-periodic inputs functions where $a = 1 \ldots m$ and $\gamma \in \mathbb{R}^{n-m}$. Lastly denote by $z(t)$, the averaged system response to the inputs. Given the averaged system (5.45), assuming that the m directly controlled states have been linearly stabilized and that the linearization of H, with respect to γ at $\gamma = 0$ and $z = x^*$, is invertible on the $(n-m)$ dimensional subspace to control, then there exists a $K \in \mathbb{R}^{(n-m)\times n}$ such that for*

$$\gamma = -\Lambda K z(t)$$

where $\Lambda^{(n-m)\times(n-m)}$ is invertible, we have stabilized the average system response.

proof
Same as Theorem. 37, without discretization.
■

Stabilization via Lyapunov Functions

In Section 2.3, the linearization of the autonomous Floquet vector field, Y, was emphasized as a means to prove stability of the original oscillatory flow. In reality, any method to demonstrate stability of the average can be applied; for example, using a Lyapunov function for the averaged system is permissible.

Definition 63 *A* stabilized truncated series expansion with respect to the Lyapunov function, V, *for the vector field (2.32) is a truncated vector field that has the same stability property using the Lyapunov function, V, as any higher-order truncation of the vector field, and also the full series expansion of the vector field.*

Recall, the averaged system response in Equation (5.45). Consider a parametrization for γ, e.g., $\gamma = p(z)$. The system in Equation (5.45) may be expressed as

$$\dot{z} = Z(z) \equiv X_S(z) + B(z)H \circ p(z). \tag{5.47}$$

Theorem 39 *Consider a system of the form (5.16) which is small-time locally controllable at $z^* \in E$. Let $u^a(t)$ be the corresponding set of γ parametrized, T-periodic inputs functions where $a = 1 \dots m$ and $\gamma \in \mathbb{R}^{n-m}$. Lastly, denote by $z(t)$ the averaged system response to the inputs. Given the averaged system (5.45), assume that there exists an γ-parametrization and a Lyapunov function, V, such that (5.47) is a stabilized truncated series expansion with respect to the Lyapunov function, V. If the system in Equation (5.47) is shown to be asymptotically (exponentially) stable using the Lyapunov function, then the averaged system response is asymptotically (exponentially) stable.*

proof
Follows the proof of 37, but without the discretization step. Instead of the linearization, use the Lyapunov function, V, to demonstrate stability.
■

Since the system (3.7) is nonlinear, the closed-loop system (5.47) will also contain nonlinear terms that complicate the Lyapunov analysis. In a local neighborhood of the equilibrium, some of the nonlinear terms do not affect stability and may be disregarded. Theorem 38 resolved the problem by linearizing the equations of motion; linearization emphasizes the dominant terms within a neighborhood of the equilibrium. A general strategy for extracting the dominant stabilizing terms and removing the unnecessary nonlinear terms is needed for the Lyapunov technique to be useful.

Geometric Homogeneity. Geometric homogeneity may be used to identify higher-order terms that act as minor perturbations in a local neighborhood of the desired equilibrium z^*. Recall that equilibria for systems with drift are elements of the zero section of E.

Definition 64 *The* m^{th}-degree homogeneous truncation of a vector field, *denoted* $\mathrm{Trunc}_m^{\Delta}(\cdot)$, *is the truncation obtained by removing all terms of homogeneous degree greater than m.*

In a local neighborhood of the equilibrium $z^* \in E$, the dominant terms will be those of homogeneous order less than or equal to 0. Consequently, the γ-parametrization must result in state error feedback whose

homogeneous truncation of order 0 is stabilizing. The 0-degree homogeneous truncation with respect to the dilation vector field, Δ, results in the vector field,

$$\dot{z} = Z_0 = \mathrm{Trunc}_0^{\Delta}\left(X_S(z) + B(z)H \circ p(z)\right). \tag{5.48}$$

If there exists a Lyapunov function such that the system in Equation (5.48) is stable, then the averaged system is locally stable.

Theorem 40 *Consider a system of the form (5.16) which is small-time locally controllable at $z^* \in E$. Let $u^a(t)$ be the corresponding set of γ parametrized, T-periodic inputs functions where $a = 1 \ldots m$ and $\gamma \in \mathbb{R}^{n-m}$. Lastly denote by $z(t)$, the averaged system response to the inputs. Given the averaged system (5.45), assuming that there exists an γ-parametrization and a Lyapunov function such that the homogeneous truncation in Equation (5.48) is asymptotically (exponentially) stable with respect to the Lyapunov function, then the actual system response is asymptotically (exponentially) stable.*

proof
Same as Theorem 19, followed by invocation of Theorem 39.
■

Comments on the Stabilizing Controllers

Theorems 37-40 stabilize an equilibrium point of the averaged system. If the γ-parametrized control input functions do not vanish at the equilibrium q^*, then by Theorem 9, the flow of the actual system stabilizes to an orbit (of size $O(\epsilon)$) around the fixed point. If, on the other hand, the input functions do vanish at the equilibrium, then Corollary 2 implies that the flow of the actual system stabilizes to the fixed point (i.e. the orbit collapses to the fixed point).

The states requiring continuous feedback use the averaged rather than the instantaneous values of the system for state feedback. Trajectories of the actual flow are related to the averaged flow by the Floquet mapping,

$$x(t) = P(t)(z(t)). \tag{5.49}$$

We may solve for the average $z(t)$, using the current state $x(t)$. Since $P(t)$ is given by a series expansion, we can easily compute its inverse.

For the discretized feedback strategy, the difference in the averaged and instantaneous indirectly controlled states is not a critical factor to consider due to the fact that $P(t)$ is periodic, i.e., $P(kT) = P(0) = \mathrm{Id},\ k \in \mathbb{Z}^+$. The directly stabilized states do not use discretized feedback, and consequently do require the average to be used as feedback. If the Floquet mapping $P(t)$ is ignored and instantaneous states $x(t)$ are used in the feedback control of the directly controllable states, this averaging method will place an upper bound on the feedback gains. The oscillatory inputs should be faster than the natural stabilizing dynamics of the directly stabilized subsytem, otherwise there will be attenuation of the oscillatory signal (and consequently the feedback signal to the indirectly controlled states). With the exception of [118], this effect has not been discussed in prior presentations of feedback strategies that utilize averaging techniques [108, 136].

In practice, one may utilize averages computed in realtime as continuous feedback. The benefit of this latter approach is that the averaging process may serve to filter out noise in the sensor signals. It will also attenuate the feedback of external disturbances. As the continually computed average, $\bar{x}(t) = \frac{1}{T}\int_{t-T}^{t} x(\tau)\,\mathrm{d}\tau$, is not quite the same as $z(t) = P^{-1}(t)\,(x(t))$, there may be some differences. When performing averaging of sensed measurements, calculate $P(t)$ to determine which states require averaging. Recall that $P(t)$ is a periodic function with a power series representation in ϵ. The states with dominant

oscillatory terms will require averaging, whereas the states that have no oscillatory terms, or ignorable, perturbative oscillatory terms, in $P(t)$ will not require averaging.

Trajectory Tracking To track a trajectory, replace $x(t)$ with $x(t) - x_d(t)$; the system must be locally controllable along the trajectory. For the discretized feedback, the Nyquist criteria is a limiting factor in tracking a trajectory for the indirectly controlled states. As the trajectory evolves, the Jacobi-Lie brackets contributing to small-time local controllability may vary. The trajectory tracking feedback strategy is more complicated since multiple parametrizations wil be needed. If a system is well-controllable, then a single parametrization will suffice to track trajectories. Systems with group symmetris are well-controllable if and only if they are small-time locally controllable at the identity.

Configuration Control via Kinematic Motions. To recover the stabilization via kinematics motions from Bullo [21], only the α-parameters are used for feedback. The system must be initially at rest. Kinematic motions are those that return the sysstem to rest and the end of each actuation period.

5.3 Examples

This section examines the control synthesis of this chapter for a couple of standard underactuated systems with drift. The first example, in Section 5.3.1, is a simple second-order nonholonomic integrator. Several control strategies for the nonholonomic integrator are examined. Section 5.3.2 examines vibrational stabilization of a passive arm. The example highlights the use of indefinite integrals for the averaged coefficients, versus the definite integrals used by approximate inversion methods for open-loop path planning. The closed-loop nature of the system is important for ensuring that indefinite integrals may be used instead.

5.3.1 Nonholonomic Integrator

As a simple didactic example, we begin with a second-order nonholonomic integrator in $\mathbb{R}^3$, which only requires first-order averaging techniques. The control vector fields are

$$Y_1(x) = \left\{ \begin{array}{c} 1 \\ 0 \\ q_2 \end{array} \right\}, \quad Y_2(x) = \left\{ \begin{array}{c} 0 \\ 1 \\ 0 \end{array} \right\},$$

and act as accelerations to the second-order system,

$$\ddot{q} = Y_1(q)u^1(q, \dot{q}, t) + Y_2(q, \dot{q}, t)u^2(q, \dot{q}, t).$$

In first order form, $x = (q, \dot{q}) \in TQ$,

$$\dot{x} = Z(x) + Y_a^{\text{lift}}(x)u^a(x, t),$$

where

$$Z = \left\{ \begin{array}{c} \dot{q} \\ 0 \end{array} \right\}.$$

The system is trivially a 1-homogeneous control system with the vector bundle $E = TQ$. The system is fully controllable as can be seen by the fact that the vertically lifted control inputs and vertical lift of the

symmetric product,

$$\langle Y_1 \ : \ Y_2 \rangle^Z = \left\{ \begin{array}{c} 0 \\ 0 \\ 1 \end{array} \right\},$$

span the vertical subbundle of TE. Here, the symmetric product operates on unlifted vector fields due to the identification of vertical vector fields with sections on the base, $\mathcal{X}^V(E) \mapsto \Gamma(E) = \mathcal{X}(Q)$. Decompose the control inputs into state-feedback and oscillatory terms,

$$u^a(x,t) = f^a(x) + \frac{1}{\epsilon} v^a(t/\epsilon) = -(k_p q^a + k_v \dot{q}^a) + \frac{1}{\epsilon} v^a(t/\epsilon).$$

The equations of motion become

$$\dot{x} = X_S + \frac{1}{\epsilon} Y_a^{\text{lift}} v^a(t/\epsilon), \tag{5.50}$$

where $X_S = Z + Y_a^{\text{lift}} f^a(x)$. With the feedback, $X_S \in \mathcal{M}_0$, and the system (5.50) is still a 1-homogeneous control system. The feedback does not affect the symmetric product calculation at all, consequently we use $\Gamma = X_S$ for all symmetric product calculations, c.f. Definition 37. Since the symmetric product can be used to add a third vector to complete the basis for the tangent space, and there are no bad symmetric products up to this level, the system should be locally controllable. Using oscillatory time dependent inputs for $v^a(\tau)$, the averaged form of the equations of motion is

$$\dot{z} = X_S(q) + \overline{\mathrm{V}_{(1)}^{(a)}(t)} \left[Y_a^{\text{lift}}, X_S \right] - \frac{1}{2} \overline{\mathrm{V}_{(1,1)}^{(a,b)}(t)} \langle Y_a \ : \ Y_b \rangle^{\text{lift}}.$$

The corresponding partial Floquet map is

$$\mathrm{Trunc}_0 \left(P(t) \right) = \mathrm{Id},$$

with an improved version being

$$\mathrm{Trunc}_1 \left(P(t/\epsilon) \right) = \mathrm{Id} + \epsilon \int_0^{t/\epsilon} \widetilde{\mathrm{V}}_{(1)}^{(a)}(\tau) \, \mathrm{d}\tau \ \left[Y_a^{\text{lift}}, X_S \right] - \frac{1}{2} \epsilon \int_0^{t/\epsilon} \widetilde{\mathrm{V}}_{(1,1)}^{(a,b)}(\tau) \, \mathrm{d}\tau \ \left\langle \ Y_a^{\text{lift}} \ : \ Y_b^{\text{lift}} \ \right\rangle.$$

Recall that the actual Floquet map also incorporates the flow $\Phi_{0,t/\epsilon}^{Y_a^{\text{lift}} v^a}$, which is

$$\Phi_{0,t/\epsilon}^{Y_a^{\text{lift}} v^a} = \mathrm{Id} + \mathrm{V}_{(1)}^{(a)}(t/\epsilon) Y_a^{\text{lift}}. \tag{5.51}$$

Choosing the inputs

$$v^1(t) = \alpha_{12}^1 \sin(t), \quad v^2(t) = \alpha_{12}^2 \sin(t)$$

the equations of motion become

$$\dot{z} = Z(q) - \alpha_{12} \langle Y_1 \ : \ Y_2 \rangle^{\text{lift}},$$

where $\alpha_{12} = \alpha_{12}^1 \alpha_{12}^2$. Explicitly written, the equations of motion are

$$\frac{\mathrm{d}}{\mathrm{d}t} \left\{ \begin{array}{c} z_1 \\ z_2 \\ z_3 \\ \dot{z}_1 \\ \dot{z}_2 \\ \dot{z}_3 \end{array} \right\} = \left\{ \begin{array}{c} \dot{z}_1 \\ \dot{z}_2 \\ \dot{z}_3 \\ -(k_p z_1 + k_v \dot{z}_1) \\ -(k_p z_2 + k_v \dot{z}_2) \\ -(k_p z_1 + k_v \dot{z}_1) z_2 \end{array} \right\} - \left\{ \begin{array}{c} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{array} \right\} \alpha_{12},$$

whose linearization about the origin is

$$\dot{z} = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ -k_p & 0 & 0 & -k_v & 0 & 0 \\ 0 & -k_p & 0 & 0 & -k_v & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} z - \begin{Bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{Bmatrix} \alpha_{12}. \tag{5.52}$$

System (5.52) is stabilizable. As, the first two directly controlled states are exponentially stable, the primary concern is the last state. There are several options regarding stabilization of control system.

Discrete Orbit Stabilization The first method will stabilize the system to an orbit using Theorem 37. To determine the range of feedback gains that will stabilize the system, it is necessary to work out the derivations of the feedback law in the proof of Theorem 37. The continuous time evolution is

$$\frac{\mathrm{d}}{\mathrm{d}t} \begin{Bmatrix} z_3 \\ \dot{z}_3 \end{Bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{Bmatrix} z_3 \\ \dot{z}_3 \end{Bmatrix} - \begin{Bmatrix} 0 \\ 1 \end{Bmatrix} \alpha_{12}.$$

When integrated over a period,

$$\begin{Bmatrix} z_3(t) \\ \dot{z}_3(t) \end{Bmatrix} = \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix} \begin{Bmatrix} z_3(0) \\ \dot{z}_3(0) \end{Bmatrix} - \begin{bmatrix} T + \frac{1}{2}T^2 \\ T \end{bmatrix} \alpha_{12}.$$

The discretized version is therefore equal to

$$\begin{Bmatrix} z_3(k+1) \\ \dot{z}_3(k+1) \end{Bmatrix} = \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix} \begin{Bmatrix} z_3(k) \\ \dot{z}_3(k) \end{Bmatrix} - \begin{bmatrix} T + \frac{1}{2}T^2 \\ T \end{bmatrix} \alpha_{12}.$$

Choosing α_{12} to be error feedback,

$$\alpha_{12} = K_p\, z(k) + K_v\, \dot{z}(k),$$

the discretized system becomes

$$\begin{Bmatrix} z_3(k+1) \\ \dot{z}_3(k+1) \end{Bmatrix} = \begin{bmatrix} 1 - (T + \frac{1}{2}T^2)K_p & T - (T + \frac{1}{2}T^2)K_v \\ -TK_p & 1 - TK_v \end{bmatrix} \begin{Bmatrix} z_3(k) \\ \dot{z}_3(k) \end{Bmatrix}.$$

Choose K_p and K_v so that the system is stabilized. Note that the value of ϵ is critical because it determines the period T, and consequently will affect the range of K_p and K_v that will stabilize the system. As the limit $\epsilon \to 0$ is approached, the range of feedback gains will increase. In the limit, the discretized averaged system and the continuous system coincide.

At this point, there are two options available: (1) we may feedback instantaneous values of the state, or (2) we may use the Floquet mapping to feedback averaged values of the state. Only the first option will be contemplated. For Option 1, the oscillatory inputs need to be faster then the natural dynamics of the directly controlled states, otherwise the direct state feedback will degrade the oscillatory signal, and consequently the feedback signal to the indirectly controlled states. A plot of the step response if shown in Figure 5.1. Due to the second sinusoidal input, the origin is no longer a fixed point, and the actual time-varying system will have a stable orbit. The parameters are $k_p = 3$, $k_v = 4$, $K_p = 0.5$, $K_v = 1.9$, $\epsilon = 1/7$. The period is $T = 2\pi\epsilon$, which gives a frequency of just over 1Hz. All of the controllers found in the remainder of this

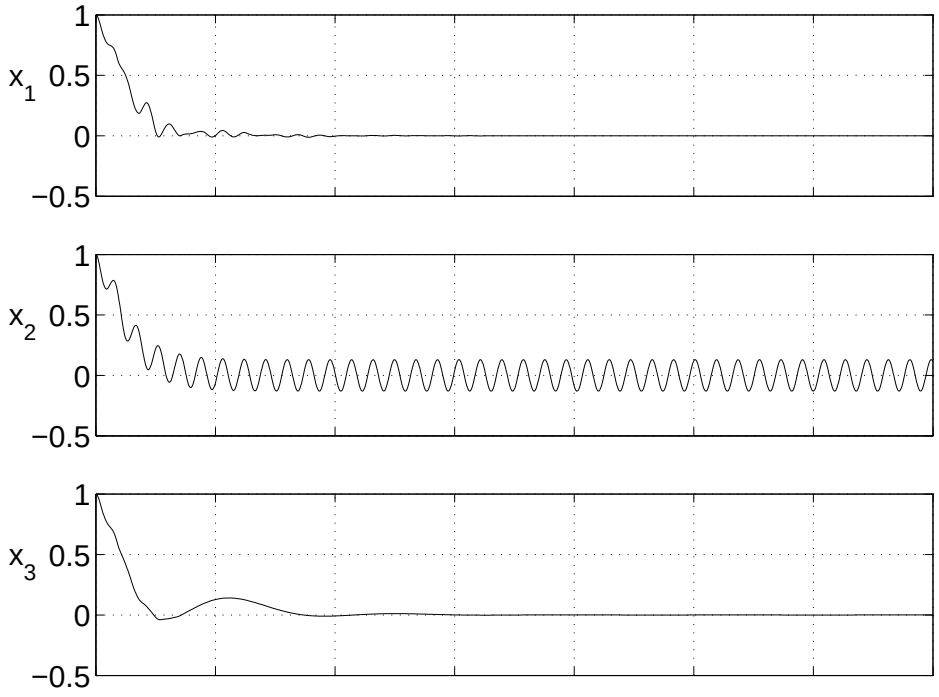


Figure 5.1: Discretized Orbit Stabilization for Nonholonomic Integrator, Option 1.

example use these control parameters also.

Discretized Stabilization to a Point For stabilization to a point to occur, the control law must vanish at the desired equilibrium. The equivalent control law will exponentially stabilize the system:

$$\alpha^1_{12}(t) = \operatorname{sign}(\alpha)\sqrt{|\alpha|}\sin(t) \quad \text{and} \quad \alpha^2_{12}(t) = \sqrt{|\alpha|}\sin(t). \tag{5.53}$$

The origin is a fixed point and the resulting control law will stabilize to the origin precisely, i.e., the orbit collapses to a point. The resulting control law is shown in Figure 5.2, using the same parameters as before, $k_p = 3$, $k_v = 4$, $K_p = 0.5$, $K_v = 1.9$, and $\epsilon = 1/7$.

Continuous Stabilization to a Point We may modify the control strategy to utilize continuous feedback instead of discretized feedback. The α-parametrization the same as Equation (5.53). The inverse of the Floquet mapping, $P(t)$, will be used to translate the instantaneous state values to the averaged state values for full-state feedback. For first-order averaging, $P(t) \approx \mathrm{Id}$, so only the flow $\Phi_{0,t/\epsilon}^{Y_a^{\mathrm{lift}} v^a}$ is needed to compute the Floquet mapping $\widetilde{P}(t)$. Figure 5.3 shows a simulation using continuous feedback of the average.

Improved Continuous Stabilization to a Point The Floquet mapping using the standard averaging method does an decent job, however, by using the Floquet mapping from the improved averaging method, the stabilization can be smoothed out and improved. The improved expansion of $P^{-1}(t)$ is

$$\mathrm{Trunc}_1\left(P^{-1}(t)\right) = \mathrm{Id} - \int_0^t \overline{\widetilde{\mathrm{V}}^{(a)}_{(1)}(\tau)}\,\mathrm{d}\tau\ \left[Y_a^{\mathrm{lift}}, X_S\right] + \int_0^t \overline{\widetilde{\mathrm{V}}^{(a,b)}_{(1,1)}(\tau)}\,\mathrm{d}\tau\ \langle\, Y_a\ :\ Y_b\,\rangle^{\mathrm{lift}} \tag{5.54}$$

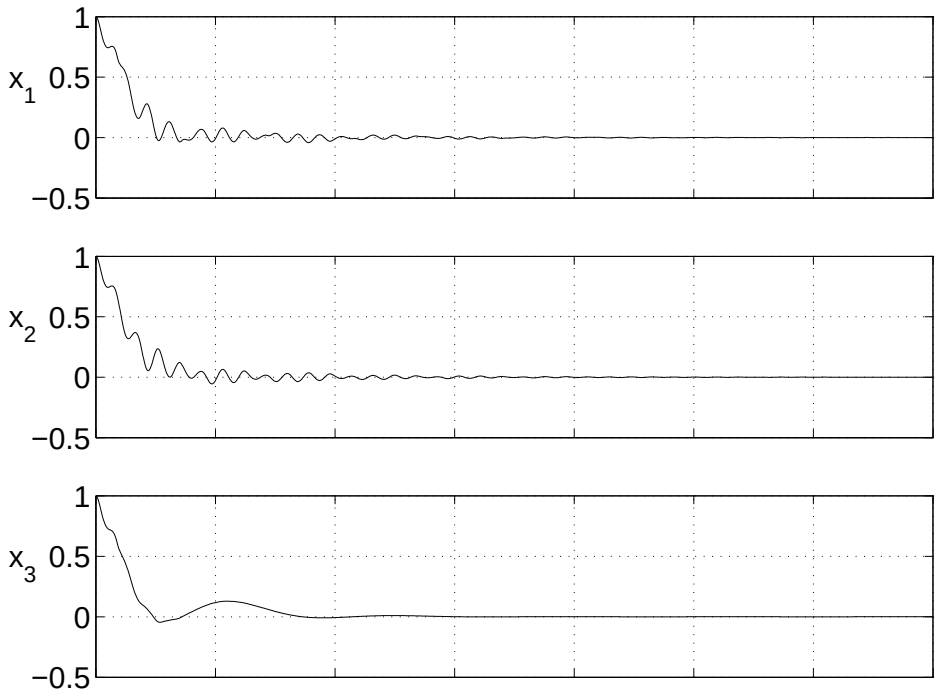


Figure 5.2: Discretized Point Stabilization for Nonholonomic Integrator.

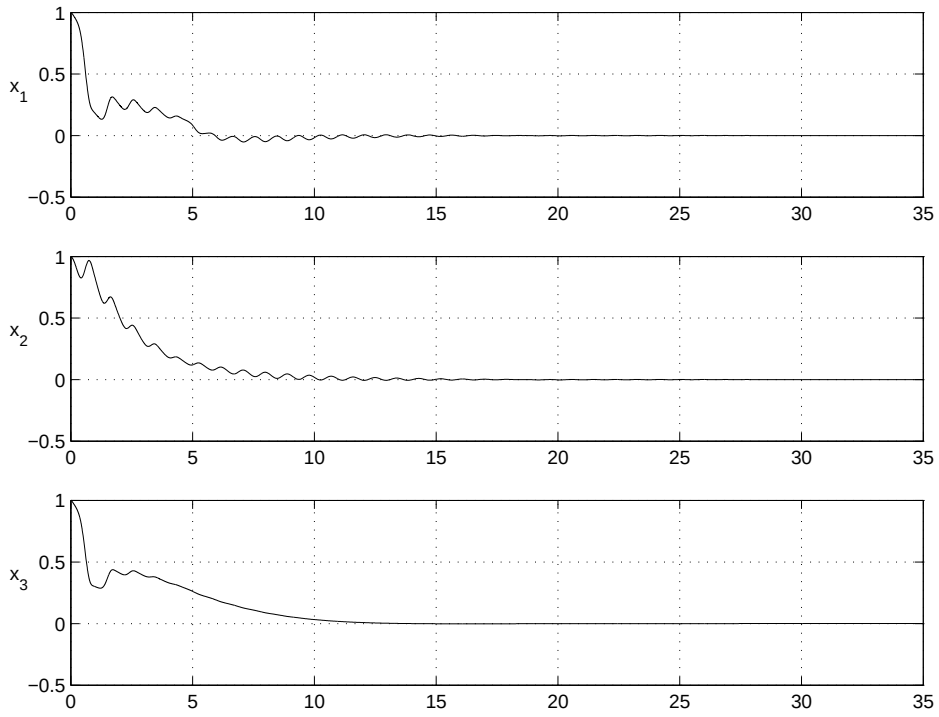


Figure 5.3: Continuous Point Stabilization for Nonholonomic Integrator.

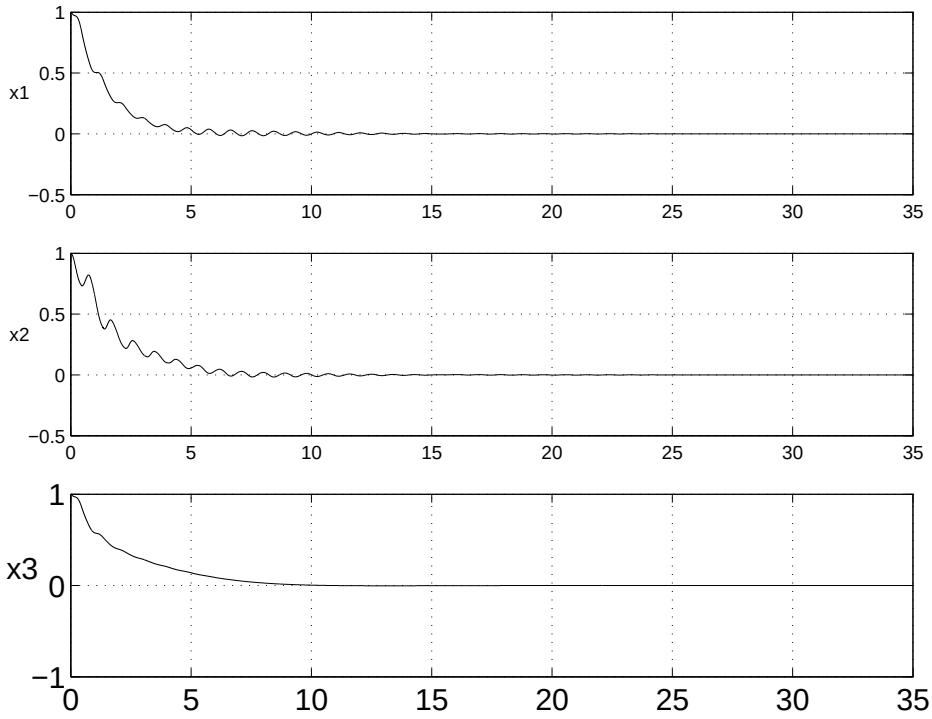


Figure 5.4: Improved Continuous Point Stabilization for Nonholonomic Integrator.

5.3.2 Vibrational Control of a Passive Arm

Consider a two-link manipulator whose second joint is passive. By quickly oscillating the first link, it is possible to stabilize the second link to be at a right angle to the first link. The goal here is to show how some elements of the generalized averaging theory manifest when dealing with an actual oscillatory system. In particular, we will demonstrate how non-uniqueness of the monodramy map is obtained and what it means physically. In the process, the input set restrictions determined in Bullo [23] will be removed. As familiarity with the problem statement is not assumed, there will be some reproduction of material from [23].

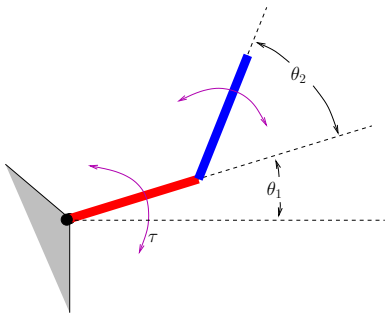


Figure 5.5: Passive arm.

The configuration of the arm can be described by the angle of the first joint, θ_1, and the angle of the second joint with respect to the first link, θ_2. Consequently, the system evolves on the two torus $\mathbb{T}^2$, and is represented by the coordinates $q = (\theta_1, \theta_2)$. The phase space of the system is $x = (q, \dot{q}) = (\theta_1, \theta_2, \omega_1, \omega_2) \in T\mathbb{T}^2$.

The second-order equations of motion can be derived using the Lagrangian of the system, which is kinetic energy minus potential energy,

$$L(q, \dot{q}) = \frac{1}{2}\dot{q}^T M(q)\dot{q} - V_{\text{control}}.$$

The kinetic energy metric is represented by the matrix,

$$M(q) = \begin{bmatrix} \frac{1}{4}\left(1013 + 1968\cos(\theta_2)\right) & 15\left(5 + 6\cos(\theta_2)\right) \\ 16\left(5 + 6\cos(\theta_2)\right) & 80 \end{bmatrix}. \tag{5.55}$$

The gradient corresponding to the kinetic energy metric is

$$\text{grad } \varphi(q) = M^{ij}\frac{\partial \varphi}{\partial q^j}.$$

The potential term consists of integrable control inputs,

$$V_{\text{control}}(q) = \frac{1}{2}k\theta_1^2 + \frac{1}{\epsilon}\sqrt{K}\theta_1 v(t/\epsilon),$$

where $v(t) = v(t+T)$. The equations of motion are

$$\ddot{q}^i = \Gamma^i_{jk}\dot{q}^j\dot{q}^k + R^i_j(q)\dot{q}^j + \text{grad } (V_{\text{control}}), \tag{5.56}$$

where the Γ^i_{jk} are the Christoffel symbols corresponding to the kinetic energy metric (5.55), $R(q)$ is Rayleigh dissipation represented by the matrix,

$$R(q) = \begin{bmatrix} -\frac{2}{10} & 0 \\ 0 & \frac{2}{10} \end{bmatrix}.$$

Although the equations of motion in Equation (5.56) exhibit periodic behavior, additional work is needed to transform them into standard form for averaging. To describe the dynamics of the system in first order form, use $x = (q, \dot{q})$,

$$\dot{x} = \widetilde{\Gamma} + D^{\text{lift}}(x) + (\text{grad } V_{\text{control}})^{\text{lift}},$$

where

$$\widetilde{\Gamma} = \begin{Bmatrix} \dot{q} \\ \Gamma_q(\dot{q}, \dot{q}) \end{Bmatrix}.$$

With the definition,

$$Y_a \equiv \text{grad } q^a,$$

the equations of motion can be expressed as

$$\dot{x} = \widetilde{\Gamma} + D^{\text{lift}}(x) - kY_1^{\text{lift}}\theta_1 + \frac{1}{\epsilon}Y_1^{\text{lift}}v^1(t/\epsilon).$$

A transformation of time, $t \mapsto \epsilon\tau$, leads to

$$\frac{\mathrm{d}x}{\mathrm{d}\tau} = \epsilon\left(\widetilde{\Gamma} + D^{\text{lift}}(x) - kY_1^{\text{lift}}\theta_1\right) + Y_1^{\text{lift}}v^1(\tau) = \epsilon\widetilde{X} + X.$$

The high-amplitude, high-frequency control inputs are the primary vector field, with the remaining dynamics acting as a perturbation. The variation of constants may be used to render the system in the required form for averaging. The vector field Y from Equation (5.17) is computed using the series representation in Equation (5.18),

$$\frac{\mathrm{d}y}{\mathrm{d}\tau} = Y = \epsilon\widetilde{X} + \mathrm{V}^{(1)}_{(1)}(\tau)\left[Y_1^{\text{lift}}, \widetilde{X}\right] - \frac{1}{2}\mathrm{V}^{(1,1)}_{(1,1)}(\tau)\left\langle Y_1^{\text{lift}} \ : \ Y_1^{\text{lift}} \right\rangle,$$

where the Γ used in the symmetric product is $\widetilde{X}$, c.f. Definition 37. The Lagrangian structure of the system implies that it is a 1-homogeneous control system. First order averaging results in

$$\frac{\mathrm{d}y}{\mathrm{d}\tau} = Y = \epsilon\widetilde{X} + \overline{\mathrm{V}^{(1)}_{(1)}(\tau)}\left[Y_1^{\text{lift}}, \widetilde{X}\right] - \frac{1}{2}\overline{\mathrm{V}^{(1,1)}_{(1,1)}(\tau)}\left\langle Y_1^{\text{lift}} \ : \ Y_1^{\text{lift}} \right\rangle.$$

To demonstrate stability of equilibria, the Lagrange-Dirichlet criterion will be used. In order to utilize the Lagrange-Dirichlet criterion, the Lagrangian system must be transformed to a Hamiltonian system using the Legendre transformation [98]. The tangent bundle coordinates $(q, \dot{q})$ are transformed to cotangent bundle coordinates (q, p).

Theorem 41 (Lagrange-Dirichlet criterion) *[98] If the matrix of the second variation $\delta^2 H$ is either positive or negative definite at the equiblirium (q_e, p_e), then it is a stable fixed point.*

For Hamiltonian systems defined on a cotangent bundle, the equilibrium subspace is a subspace of the zero section, c.f. Definition 44. Equilibria will have vanishing momenta, $(q_e, p_e) = (q_e, 0)$. The Lagrange-Dirichlet criterion simplifies to a non-degenerate minimum of the potential energy of the Hamiltonian.

Actuation with Cosine. Using the input function

$$v(\tau) = 2\sqrt{K}\,\cos(\tau),$$

the averaged coefficients evaluate to

$$\overline{\mathrm{V}^{(1)}_{(1)}(\tau)} = 0 \quad \text{and} \quad \overline{\mathrm{V}^{(1,1)}_{(1,1)}(\tau)} = 2K.$$

The averaged equations of motion are

$$\frac{\mathrm{d}y}{\mathrm{d}\tau} = Y = \epsilon\widetilde{X} - K\left\langle\, Y_1^{\mathrm{lift}} \;:\; Y_1^{\mathrm{lift}} \,\right\rangle.$$

The symmetric product term is integrable and may be incorporated into the Lagrangian as an additional potential. The averaged potential is

$$V_{avg} = \frac{1}{2}k\theta_1^2 + K\frac{20}{2313 - 1152\cos(2\theta_2)}, \tag{5.57}$$

and the averaged equations of motion may be derived from the Lagrangian,

$$L_{\mathrm{avg}}(q, \dot{q}) = \frac{1}{2}\dot{q}^T M(q)\dot{q} - V_{\mathrm{avg}}. \tag{5.58}$$

When the system is transformed into Hamiltonian form, the Lagrange-Dirichlet criteria can be used to determine stability of the average, from which orbital stability of the original system (5.56) may hold. Figure 5.6 shows minima at $\pm\frac{\pi}{2}$, which together with dissipation leads to the conclusion of stability. The actual system has an asymptotically stable orbit. The results up to here are known.

Actuation with Sine. What is interesting about this example is that the vibrational control design of [23] restricts the oscillatory actuation to be cosine. Intuitively sine should function equally well, however in developing the theory, the following integral constraints were imposed,

$$\int_0^T v^a(t)\,\mathrm{d}t \;= 0 \quad \text{and} \quad \int_0^T \int_0^t v^a(\tau)\,\mathrm{d}\tau\;\mathrm{d}t \;= 0. \tag{5.59}$$

Both cosine and sine satisfy the first integral constraint, however, only cosine satisfies the second constraint. The double integral constraint is to ensure that $\overline{\mathrm{V}^{(a)}_{(1)}(t)} = 0$, whose vanishing allows the averaged equations of motion to be described by the simple averaged Lagrangian of (5.58). The simple Lagrangian transforms trivially to a Hamiltonian that can be analyzed using the Lagrange-Dirichlet criteria.

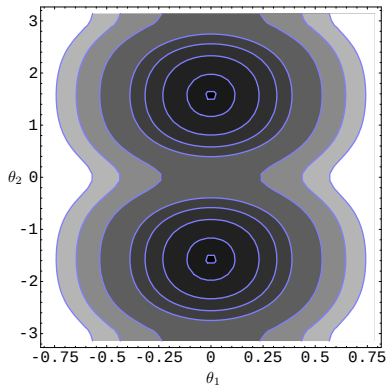


Figure 5.6: Contours for V_{avg}, Equation (5.57).

Choosing the sine function will lead to the same net response: stabilization to the $\frac{\pi}{2}$, with one caveat. The oscillatory inputs will stabilize to one of the two equilibria $\theta_2 = \pm\frac{\pi}{2}$ depending on what region in phase space that the arm begins in. Under the averaging process, the two inputs, sine and cosine, begin in different regions of the phase space for the same initial conditions. This subtle distinction can be seen by examining the structure of the averaged equations. Under oscillatory forcing with cosine, the averaged equations of motion are

$$Z_{\cos} = Z - \langle\, Y_1 \;:\; Y_1 \,\rangle\,, \tag{5.60}$$

whereas, under actuation by sine,

$$Z_{\sin} = Z + 2\sqrt{K}\,[Y_1, Z] - \langle\, Y_1 \;:\; Y_1 \,\rangle\,.$$

The averaged product evaluates to $\overline{\mathrm{V}_{(1)}^{(1)}(t)} = 2\sqrt{K}$ because definite integrals were used to compute the averaged coefficient. The two averaged equations of motion differ by one term, which is very important because it hinders the stability analysis from [23]. The averaged equations of motion are no longer in second-order form and the Lagrange-Dirichlet criterion cannot be applied. The Jacobi-Lie bracket term $[Y_1, Z]$ actually incorporates a bias in the angular velocities. By examining how the term $[Y_1, Z]$ leads to the velocity shift, we may define the following transformation of state,

$$\Theta(z) = \left\{ \begin{array}{c} \theta_1 \\ \theta_2 \\ \omega_1 + \frac{40\sqrt{K}}{2313+1152\cos(2\theta_2)} \\ \omega_2 - \frac{8\sqrt{K}(5+6\cos(\theta_2))}{2313-1152\cos(2\theta_2)} \end{array} \right\}. \tag{5.61}$$

Using the transformation, Θ, it is possible to show that

$$Z_{\cos} = \Theta^* Z_{\sin}.$$

Therefore, the average respresented by $Z_{\cos}$ can be used to represent the averaged evolution using oscillatory inputs with the sine function, only after changing coordinates. In fact, it is possible to show that for any choice of phase, i.e., $v(t) = \cos(t-\varphi)$, there exists a transformation depending on φ, Θ_φ, that will transform the averaged vector field Z_φ, corresponding to the choice of $\cos(t-\varphi)$, to $Z_{\cos}$ from Equation (5.60). Recall from Section 2.3.2 that the monodramy map was used to calculate the averaged autonomous vector field. From Equation (2.28), the monodramy map was defined to be $\Psi = \Phi_{0,T}^X$. The same section, however,

discussed that there was a degree of freedom in calculating the monodramy map, the initial time could be shifted by any amount. Therefore, the choice $\Psi = \Phi^X_{\phi,\phi+T}$ for the monodramy map will give the averaged autonomous vector field Z_{cos}. The mapping Θ above, is precisely the mapping Θ discussed in Section 2.3.2.

What this means is that there will typically exist an averaged vector field with a given phase that does not incorporate the effect of initial conditions on the average. Using the initial conditions, $q_0 = (0, 0, 0, 0)$, the initial averaged conditions for forcing by sine and using the averaged vector field Z_{cos} are

$$q_0' = \Theta(q_0) = (0, 0, \frac{200\sqrt{6}}{1161}, -\frac{440\sqrt{6}}{1161}),$$

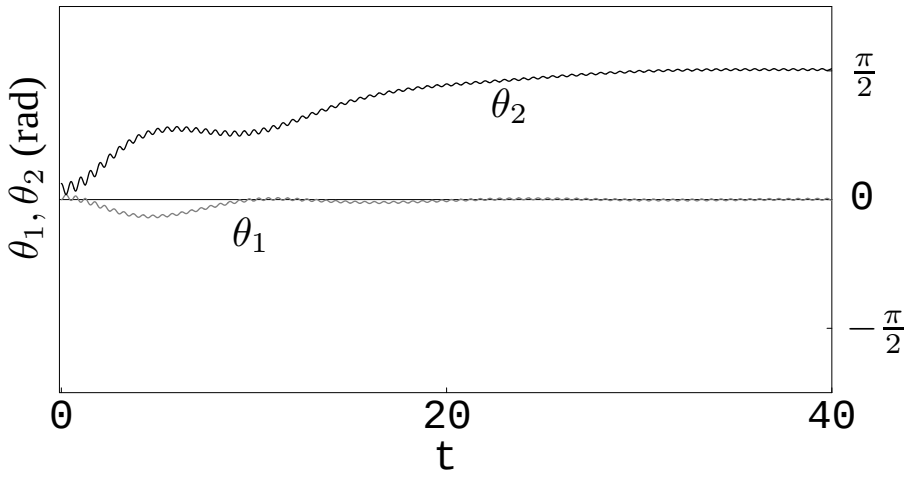
meaning that the "elbow" part of the arm, θ_2, will first be pushed towards the fixed point $-\frac{\pi}{2}$, whereas the "shoulder" part of the arm, θ_1, will be pushed upwards. This behaviour is seen in the initial transient response of the passive arm, c.f. Figure 5.7. Over time, the transient effect due to the initial conditions die out, on account of the stabilizing feedback to the first joint.

In Figure 5.8 the same system is simulated without the use of stabilizing feedback to the first joint. Notice that the transient response is attenuated due to the Rayleigh dissipation, but the angle of the first joint does not return to 0 degrees as in the case with feedback. In both cases, the angle of the second joint still stabilizes to one of the two equilibria, $\pm\frac{\pi}{2}$.

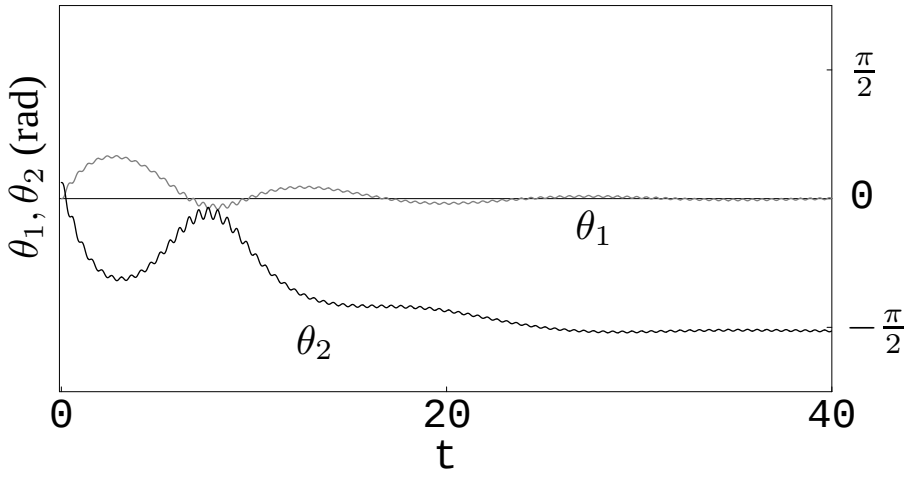
Recall the zero-average assumption that is commonly found in classical averaging theory [161, 17]. The constraints are the same as (5.59), with the exception that the oscillatory input set is not restricted to satisfy the constraints. Instead, the terms are annihilated by introducing an integration constant. The net effect, as we have discussed, is to ignore certain parts of the system dynamics. Consequently, the average may not be a faithful approximation to the system in the open-loop. In particular annihilation of these terms won't capture transients due to initial conditions, but is capable of obtaining the gross characteristic behavior, which might be the more important part. Especially if there exists stabilizing behavior that will attenuate this initial transient response. Under the classical averaging methods, both sine and cosine would result in the same averaged equations of motion. Stability of the $\theta_2 = \frac{\pi}{2}$ position holds for both cases, but we have seen that the domain of attraction is not the same. Thus the classical averaging methods are only valid within a sufficiently small neighborhood of the fixed point. Introduction of integration constants is physically equivalent to averaging around the steady-state oscillatory behavior of the system; interpreted differently, classical averaging theory is akin to a linearization around the ideal average as the constant terms are ignored. Which approach to averaging is more useful depends on the calculation at hand, and also the needs of the problem statement.

5.4 Conclusion

We applied the generalized averaging theory of Chapter 2 to 1-homogeneous control systems, and showed how averaging theory may be used to stabilize a large class of underactuated mechanical systems with drift. By proving feedback stabilization for systems evolving on a vector bundle, E, the theory collapses to known instances from the literature for various choices of the vector bundle E, e.g., for example we recover [16, 21, 24, 112, 100]. These ideas have been successfully used to stabilize systems with drift. In [116], we experimentally verified these ideas with a robotic fish using continuous feedback of the sensed average. In [106], McIsaac and Ostrowski employed discretized feedback based on the sensed average for trajectory tracking of an experimental robotic eel. Although they were unable to theoretically prove stability, the theory in this chapter can be used to do so.

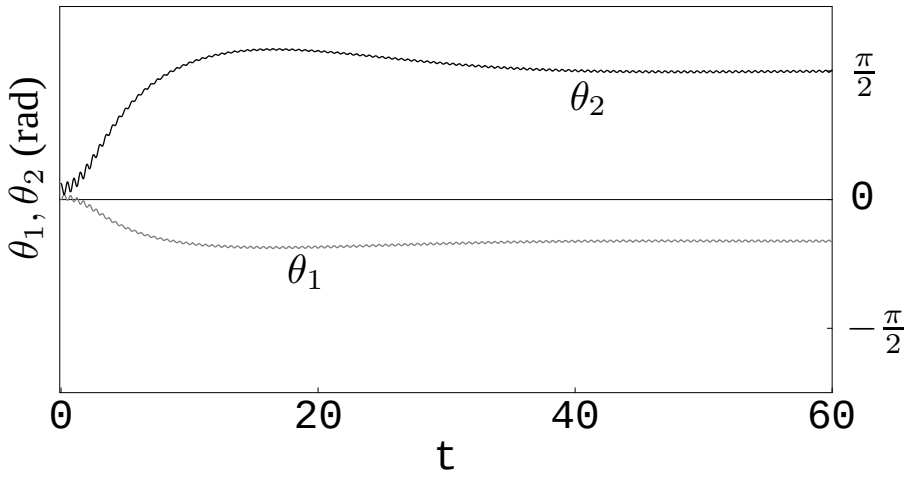


(a) v(t) = cos(t)

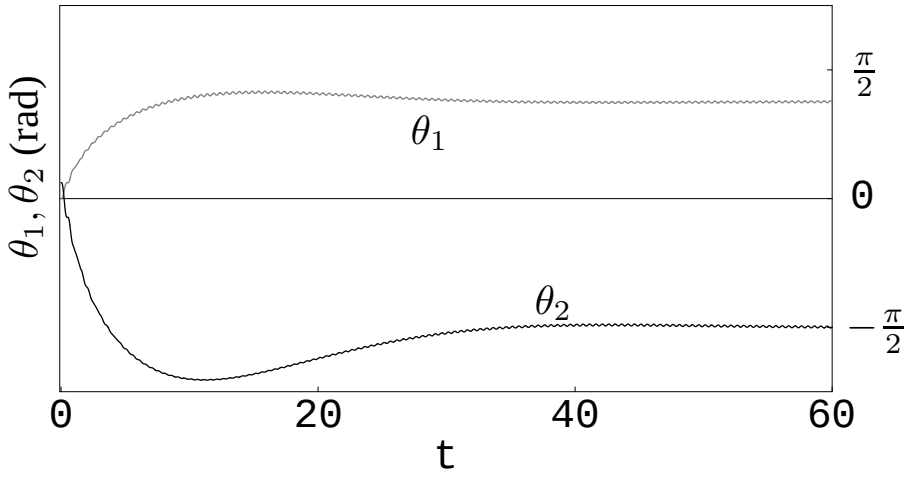


(b) v(t) = sin(t)

Figure 5.7: Closed-loop vibrational stabilization.



(a) v(t) = cos(t)



(b) v(t) = sin(t)

Figure 5.8: Open-loop vibrational stabilization.

5.A Computing the Averaged Expansions

First-order. First order averaging is trivially the time average of the original vector field,

$$\Lambda^{(1)} = X_S + \overline{\mathrm{V}^{(a)}_{(1)}(t)}\,[Y_a, X_S] - \frac{1}{2}\overline{\mathrm{V}^{(a,b)}_{(1,1)}(t)}\,\langle\, Y_a \;:\; Y_b \,\rangle\,.$$

Incorporating $\Lambda^{(1)}$ into the first order truncation of Y,

$$\frac{\mathrm{d}z}{\mathrm{d}\tau} = \epsilon\Lambda^{(1)} = \epsilon X_S + \epsilon\overline{\mathrm{V}^{(a)}_{(1)}(\tau)}\,[Y_a, X_S] - \frac{1}{2}\epsilon\overline{\mathrm{V}^{(a,b)}_{(1,1)}(\tau)}\,\langle\, Y_a \;:\; Y_b \,\rangle\,. \tag{5.62}$$

The truncated partial Floquet mapping is $\mathrm{Trunc}_0\left(P(\tau)\right) = \mathrm{Id}$.

Second-order. The second-order average is obtained by appending the second order correction. The additional term for second order averaging is

$$\Lambda^{(2)} = \frac{1}{2}\overline{\left[\int_0^\tau Y(\sigma)\,\mathrm{d}\sigma\,, Y(\tau)\right]}.$$

The nine terms that result from the correction Jacobi-Lie bracket computation are

$$\begin{aligned}
\left[\int_0^\tau Y(\sigma)\,\mathrm{d}\sigma\,, Y(\tau)\right] =\; & [X_S\;\tau, X_S] + \left[\mathrm{V}^{(a)}_{(2)}(\tau)\,[Y_a, X_S]\,, X_S\right] - \frac{1}{2}\left[\mathrm{V}^{(\widehat{a,b})}_{(\widehat{1,1})}(\tau)\,\langle\, Y_a \;:\; Y_b\,\rangle\,, X_S\right] \\
& + \left[X_S\tau, \mathrm{V}^{(a)}_{(1)}(\tau)\;\,[Y_a, X_S]\right] + \left[\mathrm{V}^{(a)}_{(2)}(\tau)\,[Y_a, X_S]\,, \mathrm{V}^{(b)}_{(1)}(\tau)\,[Y_b, X_S]\right] \\
& - \frac{1}{2}\left[\mathrm{V}^{(\widehat{a,b})}_{(\widehat{1,1})}(\tau)\;\,\langle\, Y_a \;:\; Y_b\,\rangle\,, \mathrm{V}^{(c)}_{(1)}(\tau)\;\,[Y_c, X_S]\right] \\
& - \frac{1}{2}\left[X_S\;\tau, \mathrm{V}^{(a,b)}_{(1,1)}(\tau)\;\,\langle\, Y_a \;:\; Y_b\,\rangle\right] \\
& - \frac{1}{2}\left[\mathrm{V}^{(a)}_{(2)}(\tau)\;\,[Y_a, X_S]\,, \mathrm{V}^{(b,c)}_{(1,1)}(\tau)\;\,\langle\, Y_b \;:\; Y_c\,\rangle\right] \\
& + \frac{1}{4}\left[\mathrm{V}^{(\widehat{a,b})}_{(\widehat{1,1})}(\tau)\;\,\langle\, Y_a \;:\; Y_b\,\rangle\,, \mathrm{V}^{(c,d)}_{(1,1)}(\tau)\;\,\langle\, Y_c \;:\; Y_d\,\rangle\right].
\end{aligned} \tag{5.63}$$

By using the properties of the Jacobi-Lie bracket with regards to geometric homogeneity, the bracket can be written as

$$\begin{aligned}
&\left[\int_0^t Y(\tau)\,\mathrm{d}\tau\,, Y(t)\right] = \mathrm{V}^{(a)}_{(2)}(t)\,[[Y_a, X_S]\,, X_S] - \frac{1}{2}\mathrm{V}^{(\widehat{a,b})}_{(\widehat{1,1})}(t)\,[\langle\, Y_a \;:\; Y_b\,\rangle\,, X_S] + \mathrm{V}^{(a)}_{(1)}(t)\; t\;\; [X_S, [Y_a, X_S]] \\
&\quad + \mathrm{V}^{(a,b)}_{(2,1)}(t)\,[[Y_a, X_S]\,, [Y_b, X_S]] - \frac{1}{2}\mathrm{V}^{(\widehat{a,b,c})}_{(\widehat{1,1},1)}(t)\,\langle\, Y_c \;:\; \langle\, Y_a \;:\; Y_b\,\rangle\,\rangle \\
&\quad - \frac{1}{2}\mathrm{V}^{(a,b)}_{(1,1)}(t)\; t\,[X_S, \langle\, Y_a \;:\; Y_b\,\rangle] + \frac{1}{2}\mathrm{V}^{(a,b,c)}_{(1,1,2)}(t)\,\langle\, Y_c \;:\; \langle\, Y_a \;:\; Y_b\,\rangle\,\rangle \\
=\;& \left(\mathrm{V}^{(a)}_{(2)}(t) - \mathrm{V}^{(a)}_{(1)}(t)\; t\right)[[Y_a, X_S]\,, X_S] \\
&- \frac{1}{2}\left(\mathrm{V}^{(\widehat{a,b})}_{(\widehat{1,1})}(t) - \mathrm{V}^{(a,b)}_{(1,1)}(t)\; t\right)[\langle\, Y_a \;:\; Y_b\,\rangle\,, X_S] \\
&\quad + \mathrm{V}^{(a,b)}_{(2,1)}(t)\,[[Y_a, X_S]\,, [Y_b, X_S]] \\
&- \frac{1}{2}\left(\mathrm{V}^{(a,\widehat{b,c})}_{(1,\widehat{1,1})}(t) - \mathrm{V}^{(a,b,c)}_{(2,1,1)}(t)\right)\langle\, Y_a \;:\; \langle\, Y_b \;:\; Y_c\,\rangle\,\rangle
\end{aligned}$$

The final step is to incorporate the correction into the first order averaged expansion. Using integration of products and Assumption 2, the coefficients that have time multiplication can be converted to integrals of the control inputs only. The final averaged equations of motion are

$$\begin{aligned}
\dot{z} \;&= \epsilon\Lambda^{(1)} + \epsilon^2\Lambda^{(2)} \\
&= \epsilon X_S + \epsilon\overline{\mathrm{V}^{(a)}_{(1)}(t)}\,[Y_a, X_S] - \tfrac{1}{2}\epsilon\overline{\mathrm{V}^{(a,b)}_{(1,1)}(t)}\,\langle\, Y_a \;:\; Y_b\,\rangle \\
&\quad + \epsilon^2\left(\overline{\mathrm{V}^{(a)}_{(2)}(t)} - \tfrac{1}{2}T\;\overline{\mathrm{V}^{(a)}_{(1)}(t)}\;\overline{\mathrm{V}^{(a)}_{(1)}(t)}\right)[[Y_a, X_S]\,, X_S] \\
&\quad - \tfrac{1}{2}\epsilon^2\left(\overline{\mathrm{V}^{(\widehat{a,b})}_{(\widehat{1,1})}(t)} - \tfrac{1}{2}T\;\overline{\mathrm{V}^{(a,b)}_{(1,1)}(t)}\right)[\langle\, Y_a \;:\; Y_b\,\rangle\,, X_S] \\
&\quad + \tfrac{1}{2}\epsilon^2\overline{\mathrm{V}^{(a,b)}_{(2,1)}(t)}\,[[Y_a, X_S]\,, [Y_b, X_S]] \\
&\quad + \tfrac{1}{2}\epsilon^2\left(\overline{\mathrm{V}^{(a,b,c)}_{(2,1,1)}(t)} - \tfrac{1}{2}T\;\overline{\mathrm{V}^{(a)}_{(1)}(t)}\;\overline{\mathrm{V}^{(b,c)}_{(1,1)}(t)}\right)\langle\, Y_a \;:\; \langle\, Y_b \;:\; Y_c\,\rangle\,\rangle\,.
\end{aligned} \tag{5.64}$$

To calculate the partial Floquet mapping, the first order truncation is needed, $\mathrm{Trunc}_1\left(P(t)\right)$. It is

$$\mathrm{Trunc}_1\left(P(t/\epsilon)\right) = \mathrm{Id} + \epsilon\int_0^{t/\epsilon}\widetilde{\mathrm{V}}^{(a)}_{(1)}(\tau)\,\mathrm{d}\tau\;\left[Y_a^{\mathrm{lift}}, X_S\right] - \frac{1}{2}\epsilon\int_0^{t/\epsilon}\widetilde{\mathrm{V}}^{(a,b)}_{(1,1)}(\tau)\,\mathrm{d}\tau\;\left\langle\, Y_a^{\mathrm{lift}} \;:\; Y_b^{\mathrm{lift}}\,\right\rangle.$$

Third-order The additional term for third order averaging is

$$-\frac{1}{2}T\left[\Lambda^{(1)}, \Lambda^{(2)}\right] + \frac{1}{3}\overline{\left[\int_0^t Y(\tau)\,\mathrm{d}\tau\,, \left[\int_0^t Y(\tau)\,\mathrm{d}\tau\,, Y(t)\right]\right]}.$$

The twelve elements of the first Jacobi-Lie bracket are

$$\begin{aligned}
\left[\Lambda^{(1)},\Lambda^{(2)}\right] &= \left(\overline{\mathrm{V}^{(a)}_{(2)}(\tau)} - \frac{1}{2}T\overline{\mathrm{V}^{(a)}_{(1)}(\tau)}\right)[X_S,[[Y_a,X_S],X_S]] \\
&- \left(\frac{1}{2}\overline{\mathrm{V}^{(\widehat{a,b})}_{(\widehat{1,1})}(\tau)} - \frac{1}{4}T\,\overline{\mathrm{V}^{(a,b)}_{(1,1)}(\tau)}\right)[X_S,[\langle\, Y_a\ :\ Y_b\,\rangle,X_S]] \\
&+ \frac{1}{2}\overline{\mathrm{V}^{(a,b)}_{(2,1)}(\tau)}\,[X_S,[[Y_a,X_S],[Y_b,X_S]]] \\
&+ \left(\frac{1}{2}\overline{\mathrm{V}^{(a,b,c)}_{(2,1,1)}(\tau)} - \frac{1}{4}T\,\overline{\mathrm{V}^{(a)}_{(1)}(\tau)}\,\overline{\mathrm{V}^{(b,c)}_{(1,1)}(\tau)}\right)[X_S,\langle\, Y_a\ :\ \langle\, Y_b\ :\ Y_c\,\rangle\,\rangle] \\
&+ \overline{\mathrm{V}^{(a)}_{(1)}(\tau)}\left(\overline{\mathrm{V}^{(b)}_{(2)}(\tau)} - \frac{1}{2}T\,\overline{\mathrm{V}^{(b)}_{(1)}(\tau)}\right)[[Y_a,X_S],[[Y_b,X_S],X_S]] \\
&- \overline{\mathrm{V}^{(a)}_{(1)}(\tau)}\left(\frac{1}{2}\overline{\mathrm{V}^{(\widehat{b,c})}_{(\widehat{1,1})}(\tau)} - \frac{1}{4}T\,\overline{\mathrm{V}^{(b,c)}_{(1,1)}(\tau)}\right)[[Y_a,X_S],[\langle\, Y_a\ :\ Y_b\,\rangle,X_S]] \\
&+ \frac{1}{2}\overline{\mathrm{V}^{(a)}_{(1)}(\tau)}\,\overline{\mathrm{V}^{(b,c)}_{(2,1)}(\tau)}\,[[Y_a,X_S],[[Y_b,X_S],[Y_c,X_S]]] \\
&+ \overline{\mathrm{V}^{(a)}_{(1)}(\tau)}\left(\frac{1}{2}\overline{\mathrm{V}^{(a,b,c)}_{(2,1,1)}(\tau)} - \frac{1}{4}T\,\overline{\mathrm{V}^{(b)}_{(1)}(\tau)}\,\overline{\mathrm{V}^{(c,d)}_{(1,1)}(\tau)}\right)[[Y_a,X_S],\langle\, Y_b\ :\ \langle\, Y_c\ :\ Y_d\,\rangle\,\rangle] \\
&- \frac{1}{2}\overline{\mathrm{V}^{(a,b)}_{(1,1)}(\tau)}\left(\overline{\mathrm{V}^{(c)}_{(2)}t)} - \frac{1}{2}T\,\overline{\mathrm{V}^{(c)}_{(1)}(\tau)}\right)[\langle\, Y_a\ :\ Y_b\,\rangle,[[Y_c,X_S],X_S]] \\
&+ \frac{1}{2}\overline{\mathrm{V}^{(a,b)}_{(1,1)}(\tau)}\left(\frac{1}{2}\overline{\mathrm{V}^{(\widehat{c,d})}_{(\widehat{1,1})}(\tau)} - \frac{1}{4}T\,\overline{\mathrm{V}^{(c,d)}_{(1,1)}(\tau)}\right)[\langle\, Y_a\ :\ Y_b\,\rangle,[\langle\, Y_c\ :\ Y_d\,\rangle,X_S]] \\
&- \frac{1}{4}\overline{\mathrm{V}^{(a,b)}_{(1,1)}(\tau)}\,\overline{\mathrm{V}^{(c,d)}_{(2,1)}(\tau)}\,[\langle\, Y_a\ :\ Y_b\,\rangle,[[Y_c,X_S],[Y_d,X_S]]] \\
&- \frac{1}{4}\overline{\mathrm{V}^{(a,b)}_{(1,1)}(\tau)}\left(\frac{1}{2}\overline{\mathrm{V}^{(c,d,e)}_{(2,1,1)}(\tau)} - \frac{1}{4}T\,\overline{\mathrm{V}^{(c)}_{(1)}(\tau)}\,\overline{\mathrm{V}^{(d,e)}_{(1,1)}(\tau)}\right)[\langle\, Y_a\ :\ Y_b\,\rangle,\langle\, Y_c\ :\ \langle\, Y_d\ :\ Y_e\,\rangle\,\rangle]\,.
\end{aligned}$$

The twelve elements of the second component, prior to averaging, are

$$
\begin{aligned}
&\left[\int_0^\tau Y(\sigma)\,\mathrm{d}\sigma\,,\left[\int_0^\tau F(\sigma)\,\mathrm{d}\sigma\,,F(\tau)\right]\right]\\
=\\
&-\left(\mathrm{V}_{(2)}^{(a)}(\tau)\,\tau-\mathrm{V}_{(1)}^{(a)}(\tau)\,\tau^2\right)[[[Y_a,X_S]\,,X_S]\,,X_S]\\
&+\frac{1}{2}\left(\mathrm{V}_{(\widehat{1},\widehat{1})}^{(\widehat{a,b})}(\tau)\,\tau-\mathrm{V}_{(1,1)}^{(a,b)}(\tau)\,\tau^2\right)[[\langle\,Y_a\;:\;Y_b\,\rangle\,,X_S]\,,X_S]\\
&-\mathrm{V}_{(2,1)}^{(a,b)}(\tau)\,\tau\;[[[Y_a,X_S]\,,[Y_b,X_S]]\,,X_S]\\
&+\frac{1}{2}\left(\mathrm{V}_{(1,\widehat{1},\widehat{1})}^{(a,\widehat{b,c})}(\tau)\,\tau-\mathrm{V}_{(2,1,1)}^{(a,b,c)}(\tau)\,\tau\right)[\langle\,Y_a\;:\;\langle\,Y_b\;:\;Y_c\,\rangle\,\rangle\,,X_S]\\
&+\left(\mathrm{V}_{(2,2)}^{(a,b)}(\tau)-\mathrm{V}_{(2,1)}^{(a,b)}(\tau)\,\tau\right)[[Y_a,X_S]\,,[[Y_b,X_S]\,,X_S]]\\
&-\frac{1}{2}\left(\mathrm{V}_{(2,\widehat{1},\widehat{1})}^{(a,\widehat{b,c})}(\tau)-\mathrm{V}_{(2,1,1)}^{(a,b,c)}(\tau)\,\tau\right)[[Y_a,X_S]\,,[\langle\,Y_b\;:\;Y_c\,\rangle\,,X_S]]\\
&+\mathrm{V}_{(2,2,1)}^{(a,b,c)}(\tau)\,[[Y_a,X_S]\,,[[Y_b,X_S]\,,[Y_c,X_S]]]\\
&+\frac{1}{2}\left(\mathrm{V}_{(2,1,\widehat{1,1})}^{(a,b,\widehat{c,d})}(\tau)-\mathrm{V}_{(2,2,1,1)}^{(a,b,c,d)}(\tau)\right)\langle\,Y_a\;:\;\langle\,Y_b\;:\;\langle\,Y_c\;:\;Y_d\,\rangle\,\rangle\,\rangle\\
&-\frac{1}{2}\left(\mathrm{V}_{(\widehat{1,1},1)}^{(\widehat{a,b},c)}(\tau)-\mathrm{V}_{(\widehat{1,1},1)}^{(\widehat{a,b},c)}(\tau)\,\tau\right)[\langle\,Y_a\;:\;Y_b\,\rangle\,,[[Y_c,X_S]\,,X_S]]\\
&+\frac{1}{4}\left(\mathrm{V}_{(\widehat{1,1},\widehat{1,1})}^{(\widehat{a,b},\widehat{c,d})}(\tau)-\mathrm{V}_{(\widehat{1,1},1,1)}^{(\widehat{a,b},c,d)}(\tau)\,\tau\right)[\langle\,Y_a\;:\;Y_b\,\rangle\,,[\langle\,Y_c\;:\;Y_d\,\rangle\,,X_S]]\\
&-\frac{1}{2}\mathrm{V}_{(\widehat{1,1},2,1)}^{(\widehat{a,b},c,d)}(\tau)\,[\langle\,Y_a\;:\;Y_b\,\rangle\,,[[Y_c,X_S]\,,[Y_d,X_S]]]\\
&+\frac{1}{4}\left(\mathrm{V}_{(\widehat{1,1},\widehat{1,1})}^{(\widehat{a,b},c,\widehat{d,e})}(\tau)-\mathrm{V}_{(\widehat{1,1},2,1,1)}^{(\widehat{a,b},c,d,e)}(\tau)\right)[\langle\,Y_a\;:\;Y_b\,\rangle\,,\langle\,Y_c\;:\;\langle\,Y_d\;:\;Y_e\,\rangle\,\rangle]\,.
\end{aligned}
$$

Properties of the Jacobi-Lie bracket may be used to obtain simplifications of the third-order correction term. Doing so, the third order averaging equations of motion are

$$
\begin{aligned}
\frac{\mathrm{d}z}{\mathrm{d}\tau} =& X_S + \overline{\mathrm{V}^{(a)}_{(1)}(t)}\,[Y_a, X_S] - \frac{1}{2}\overline{\mathrm{V}^{(a,b)}_{(1,1)}(t)}\,\langle\, Y_a \;:\; Y_b \,\rangle \\
&+ \epsilon\left(\overline{\mathrm{V}^{(a)}_{(2)}(t)} - \frac{1}{2}T\,\overline{\mathrm{V}^{(a)}_{(1)}(t)}\right)[[Y_a, X_S], X_S] - \frac{1}{2}\epsilon\left(\overline{\mathrm{V}^{(\widehat{a,b})}_{(\widehat{1,1})}(t)} - \frac{1}{2}T\,\overline{\mathrm{V}^{(a,b)}_{(1,1)}(t)}\right)[\langle\, Y_a \;:\; Y_b \,\rangle, X_S] \\
&+ \frac{1}{2}\epsilon\overline{\mathrm{V}^{(a,b)}_{(2,1)}(t)}\,[[Y_a, X_S], [Y_b, X_S]] + \frac{1}{2}\epsilon\left(\overline{\mathrm{V}^{(a,b,c)}_{(2,1,1)}(t)} - \frac{1}{2}T\,\overline{\mathrm{V}^{(a)}_{(1)}(t)}\,\overline{\mathrm{V}^{(b,c)}_{(1,1)}(t)}\right)\langle\, Y_a \;:\; \langle\, Y_b \;:\; Y_c \,\rangle\,\rangle \\
&- \epsilon^2\left(\frac{1}{3}\overline{\mathrm{V}^{(a)}_{(2)}(t)\,t} - \frac{1}{3}\overline{\mathrm{V}^{(a)}_{(1)}(t)\,t^2} - \frac{1}{2}T\,\overline{\mathrm{V}^{(a)}_{(2)}(t)} + \frac{1}{4}T^2\,\overline{\mathrm{V}^{(a)}_{(1)}(t)}\right)[[[Y_a, X_S], X_S], X_S] \\
&+ \epsilon^2\left(\frac{1}{6}\overline{\mathrm{V}^{(\widehat{a,b})}_{(\widehat{1,1})}(t)\,t} - \frac{1}{6}\overline{\mathrm{V}^{(a,b)}_{(1,1)}(t)\,t^2} - \frac{1}{2}T\,\overline{\mathrm{V}^{(a)}_{(2)}(t)} + \frac{1}{4}T^2\,\overline{\mathrm{V}^{(a)}_{(1)}(t)}\right)[[\langle\, Y_a \;:\; Y_b \,\rangle, X_S], X_S] \\
&- \epsilon^2\left(\frac{1}{3}\overline{\mathrm{V}^{(a,b)}_{(2,1)}(t)\,t} - \frac{1}{4}T\,\overline{\mathrm{V}^{(a,b)}_{(2,1)}(t)}\right)[[[Y_a, X_S], [Y_b, X_S]], X_S] \\
&+ \epsilon^2\left(\frac{1}{6}\overline{\mathrm{V}^{(a,\widehat{b,c})}_{(1,\widehat{1,1})}(t)\,t} - \frac{1}{6}\overline{\mathrm{V}^{(a,b,c)}_{(2,1,1)}(t)\,t} + \frac{1}{4}T\,\overline{\mathrm{V}^{(a,b,c)}_{(2,1,1)}(t)} - \frac{1}{8}T^2\,\overline{\mathrm{V}^{(a)}_{(1)}(t)}\,\overline{\mathrm{V}^{(b,c)}_{(1,1)}(t)}\right) \\
&\qquad\qquad [\langle\, Y_a \;:\; \langle\, Y_b \;:\; Y_c \,\rangle\,\rangle, X_S] \\
&+ \epsilon^2\left(\frac{1}{3}\overline{\mathrm{V}^{(a,b)}_{(2,2)}(t)} - \frac{1}{3}\overline{\mathrm{V}^{(a,b)}_{(2,1)}(t)\,t} - \frac{1}{2}T\,\overline{\mathrm{V}^{(a)}_{(1)}(t)}\,\overline{\mathrm{V}^{(b)}_{(2)}(t)} + \frac{1}{4}T^2\,\overline{\mathrm{V}^{(a)}_{(1)}(t)}\,\overline{\mathrm{V}^{(b)}_{(1)}(t)}\right) \\
&\qquad\qquad [[Y_a, X_S], [[Y_b, X_S], X_S]] \\
&- \epsilon^2\left(\frac{1}{6}\overline{\mathrm{V}^{(\widehat{a,b,c})}_{(2,\widehat{1,1})}(t)} - \frac{1}{6}\overline{\mathrm{V}^{(a,b,c)}_{(2,1,1)}(t)\,t} + \frac{1}{4}T\,\overline{\mathrm{V}^{(a)}_{(1)}(t)}\,\overline{\mathrm{V}^{(\widehat{a,b})}_{(\widehat{1,1})}(t)} - \frac{1}{8}T^2\,\overline{\mathrm{V}^{(a)}_{(1)}(t)}\,\overline{\mathrm{V}^{(b,c)}_{(1,1)}(t)}\right) \\
&\qquad\qquad [[Y_a, X_S], [\langle\, Y_b \;:\; Y_c \,\rangle, X_S]] \\
&+ \epsilon^2\left(\frac{1}{3}\overline{\mathrm{V}^{(a,b,c)}_{(2,2,1)}(t)} - \frac{1}{4}T\,\overline{\mathrm{V}^{(a)}_{(1)}(t)}\,\overline{\mathrm{V}^{(b,c)}_{(2,1)}(t)}\right)[[Y_a, X_S], [[Y_b, X_S], [Y_c, X_S]]] \\
&+ \epsilon^2\left(\frac{1}{6}\overline{\mathrm{V}^{(a,b,\widehat{c,d})}_{(2,1,\widehat{1,1})}(t)} - \frac{1}{6}\overline{\mathrm{V}^{(a,b,c,d)}_{(2,2,1,1)}(t)} + \frac{1}{4}T\,\overline{\mathrm{V}^{(a)}_{(1)}(t)}\,\overline{\mathrm{V}^{(b,c,d)}_{(2,1,1)}(t)} - \frac{1}{8}T^2\,\overline{\mathrm{V}^{(a)}_{(1)}(t)}\,\overline{\mathrm{V}^{(b)}_{(1)}(t)}\,\overline{\mathrm{V}^{(c,d)}_{(1,1)}(t)}\right) \\
&\qquad\qquad \langle\, Y_a \;:\; \langle\, Y_b \;:\; \langle\, Y_c \;:\; Y_d \,\rangle\,\rangle\,\rangle \\
&- \epsilon^2\left(\frac{1}{6}\overline{\mathrm{V}^{(\widehat{a,b,c})}_{(\widehat{1,1,2})}(t)} - \frac{1}{6}\overline{\mathrm{V}^{(\widehat{a,b,c})}_{(\widehat{1,1,1})}(t)\,t} - \frac{1}{4}T\,\overline{\mathrm{V}^{(a,b)}_{(1,1)}(t)}\,\overline{\mathrm{V}^{(c)}_{(2)}(t)} + \frac{1}{8}T^2\,\overline{\mathrm{V}^{(a,b)}_{(1,1)}(t)}\,\overline{\mathrm{V}^{(c)}_{(2)}(t)}\right) \\
&\qquad\qquad [\langle\, Y_a \;:\; Y_b \,\rangle, [[Y_c, X_S], X_S]] \\
&+ \epsilon^2\left(\frac{1}{12}\overline{\mathrm{V}^{(\widehat{a,b,c,d})}_{(\widehat{1,1,1,1})}(t)} - \frac{1}{12}\overline{\mathrm{V}^{(\widehat{a,b,c,d})}_{(\widehat{1,1,1,1})}(t)} - \frac{1}{8}T\,\overline{\mathrm{V}^{(a,b)}_{(1,1)}(t)}\,\overline{\mathrm{V}^{(\widehat{c,d})}_{(\widehat{1,1})}(t)} + \frac{1}{1}T^2\,\overline{\mathrm{V}^{(a,b)}_{(1,1)}(t)}\,\overline{\mathrm{V}^{(c,d)}_{(1,1)}(t)}\right) \\
&\qquad\qquad \langle\,\langle\, Y_a \;:\; Y_b \,\rangle \;:\; \langle\, Y_c \;:\; Y_d \,\rangle\,\rangle \\
&+ \epsilon^2\left(\frac{1}{6}\overline{\mathrm{V}^{(a,b,\widehat{c,d})}_{(2,1,\widehat{1,1})}(t)} - \frac{1}{6}\overline{\mathrm{V}^{(a,b,\widehat{c,d})}_{(1,2,\widehat{1,1})}(t)} - \frac{1}{8}T\,\overline{\mathrm{V}^{(a,b)}_{(2,1)}(t)}\,\overline{\mathrm{V}^{(c,d)}_{(1,1)}(t)} - \frac{1}{8}T\,\overline{\mathrm{V}^{(a,b)}_{(1,2)}(t)}\,\overline{\mathrm{V}^{(c,d)}_{(1,1)}(t)}\right) \\
&\qquad\qquad \langle\, Y_a \;:\; \langle\, Y_b \;:\; \langle\, Y_c \;:\; Y_d \,\rangle\,\rangle\,\rangle\,.
\end{aligned}
$$

The truncation of the partial Floquet mapping remains to be computed. The truncated logarithm of the flow $\Phi^Y_{0,t}$ is

$$\begin{aligned}
\mathrm{Trunc}_2\left(\ln\overrightarrow{\exp}\int_0^\tau Y\,\mathrm{d}\tau\right) &= \epsilon X_S\tau + \epsilon \mathrm{V}^{(a)}_{(2)}(\tau)\left[Y_a^{\mathrm{lift}}, X_S\right] - \frac{1}{2}\epsilon \mathrm{V}^{(\widehat{a,b})}_{(\widehat{1,1})}(\tau)\left\langle Y_a^{\mathrm{lift}} \;:\; Y_b^{\mathrm{lift}}\right\rangle \\
&\quad + \epsilon^2\left(\mathrm{V}^{(a)}_{(3)}(\tau) - \frac{1}{2}\tau\mathrm{V}^{(a)}_{(2)}(\tau)\right)\left[\left[Y_a^{\mathrm{lift}}, X_S\right], X_S\right] \\
&\quad - \frac{1}{2}\epsilon^2\left(\mathrm{V}^{(\widehat{\widehat{a,b}})}_{(\widehat{\widehat{1,1}})}(\tau) - \frac{1}{2}\tau\mathrm{V}^{(\widehat{a,b})}_{(\widehat{1,1})}(\tau)\right)\left[\left\langle Y_a^{\mathrm{lift}} \;:\; Y_b^{\mathrm{lift}}\right\rangle, X_S\right] \\
&\quad + \frac{1}{2}\epsilon^2\mathrm{V}^{(\widehat{a,b})}_{(\widehat{2,1})}(\tau)\left[\left[Y_a^{\mathrm{lift}}, X_S\right], \left[Y_b^{\mathrm{lift}}, X_S\right]\right] \\
&\quad + \frac{1}{2}\epsilon^2\left(\mathrm{V}^{(\widehat{a,b,c})}_{(\widehat{2,1,1})}(\tau) - \frac{1}{2}\mathrm{V}^{(a,\widehat{b,c})}_{(2,\widehat{1,1})}(\tau)\right)\left\langle Y_a^{\mathrm{lift}} \;:\; \left\langle Y_b^{\mathrm{lift}} \;:\; Y_c^{\mathrm{lift}}\right\rangle\right\rangle.
\end{aligned}$$

The truncated exponential of $\mathrm{Trunc}_2\left(\ln\overrightarrow{\exp}\int_0^\tau Y\,\mathrm{d}\tau\right)$ is an approximation to truncation of the actual time varying chronological exponential,

$$\begin{aligned}
\mathrm{Trunc}_2\left(\overrightarrow{\exp}\int_0^\tau Y\,\mathrm{d}\tau\right) &= \mathrm{Id} + \epsilon X_S\tau + \epsilon \mathrm{V}^{(a)}_{(2)}(\tau)\left[Y_a^{\mathrm{lift}}, X_S\right] - \frac{1}{2}\epsilon \mathrm{V}^{(\widehat{a,b})}_{(\widehat{1,1})}(\tau)\left\langle Y_a^{\mathrm{lift}} \;:\; Y_b^{\mathrm{lift}}\right\rangle \\
&\quad + \epsilon^2\left(\mathrm{V}^{(a)}_{(3)}(\tau) - \frac{1}{2}\tau\mathrm{V}^{(a)}_{(2)}(\tau)\right)\left[\left[Y_a^{\mathrm{lift}}, X_S\right], X_S\right] \\
&\quad - \frac{1}{2}\epsilon^2\left(\mathrm{V}^{(\widehat{\widehat{a,b}})}_{(\widehat{\widehat{1,1}})}(\tau) - \frac{1}{2}\tau\mathrm{V}^{(\widehat{a,b})}_{(\widehat{1,1})}(\tau)\right)\left[\left\langle Y_a^{\mathrm{lift}} \;:\; Y_b^{\mathrm{lift}}\right\rangle, X_S\right] \\
&\quad + \frac{1}{2}\epsilon^2\mathrm{V}^{(\widehat{a,b})}_{(\widehat{2,1})}(\tau)\left[\left[Y_a^{\mathrm{lift}}, X_S\right], \left[Y_b^{\mathrm{lift}}, X_S\right]\right] \\
&\quad + \frac{1}{2}\epsilon^2\left(\mathrm{V}^{(\widehat{a,b,c})}_{(\widehat{2,1,1})}(\tau) - \frac{1}{2}\mathrm{V}^{(a,\widehat{b,c})}_{(2,\widehat{1,1})}(\tau)\right)\left\langle Y_a^{\mathrm{lift}} \;:\; \left\langle Y_b^{\mathrm{lift}} \;:\; Y_c^{\mathrm{lift}}\right\rangle\right\rangle \\
&\quad + \frac{1}{2}\epsilon^2 X_S\cdot X_S\,\tau^2 + \frac{1}{2}\epsilon^2\mathrm{V}^{(a)}_{(2)}(\tau)\,\tau X_S\cdot\left[Y_a^{\mathrm{lift}}, X_S\right] \\
&\quad + \frac{1}{2}\epsilon^2\mathrm{V}^{(a)}_{(2)}(\tau)\,\tau\left[Y_a^{\mathrm{lift}}, X_S\right]\cdot X_S \\
&\quad - \frac{1}{4}\epsilon^2\mathrm{V}^{(\widehat{a,b})}_{(\widehat{1,1})}(\tau)\,\tau X_S\cdot\left\langle Y_a^{\mathrm{lift}} \;:\; Y_b^{\mathrm{lift}}\right\rangle \\
&\quad - \frac{1}{4}\epsilon^2\mathrm{V}^{(\widehat{a,b})}_{(\widehat{1,1})}(\tau)\,\tau\left\langle Y_a^{\mathrm{lift}} \;:\; Y_b^{\mathrm{lift}}\right\rangle\cdot X_S \\
&\quad + \frac{1}{2}\epsilon^2\mathrm{V}^{(a)}_{(2)}(\tau)\left[Y_a^{\mathrm{lift}}, X_S\right]\cdot\left[Y_b^{\mathrm{lift}}, X_S\right] \\
&\quad - \frac{1}{4}\epsilon^2\mathrm{V}^{(a,\widehat{b,c})}_{(2,\widehat{1,1})}(\tau)\left[Y_a^{\mathrm{lift}}, X_S\right]\cdot\left\langle Y_b^{\mathrm{lift}} \;:\; Y_c^{\mathrm{lift}}\right\rangle \\
&\quad - \frac{1}{4}\epsilon^2\mathrm{V}^{(\widehat{a,b,c})}_{(\widehat{1,1,2})}(\tau)\left\langle Y_a^{\mathrm{lift}} \;:\; Y_b^{\mathrm{lift}}\right\rangle\cdot\left[Y_c^{\mathrm{lift}}, X_S\right] \\
&\quad + \frac{1}{4}\epsilon^2\mathrm{V}^{(\widehat{a,b,c,d})}_{(\widehat{1,1,1,1})}(\tau)\left\langle Y_a^{\mathrm{lift}} \;:\; Y_b^{\mathrm{lift}}\right\rangle\cdot\left\langle Y_c^{\mathrm{lift}} \;:\; Y_d^{\mathrm{lift}}\right\rangle.
\end{aligned}$$

The truncated exponential of $-Z\tau$ is

$$
\begin{aligned}
\mathrm{Trunc}_2\left(e^{-Z\tau}\right) = \mathrm{Id} &- \epsilon X_S\tau - \epsilon\overline{\mathrm{V}^{(a)}_{(1)}(\tau)}\,\tau\left[Y_a^{\mathrm{lift}}, X_S\right] + \frac{1}{2}\epsilon\overline{\mathrm{V}^{(a,b)}_{(1,1)}(\tau)}\,\tau\left\langle\, Y_a^{\mathrm{lift}} \;:\; Y_b^{\mathrm{lift}}\,\right\rangle \\
&- \epsilon^2\left(\overline{\mathrm{V}^{(a)}_{(2)}(\tau)} - \frac{1}{2}T\overline{\mathrm{V}^{(a)}_{(1)}(\tau)}\right)\tau\left[\left[Y_a^{\mathrm{lift}}, X_S\right], X_S\right] \\
&+ \frac{1}{2}\epsilon^2\left(\overline{\mathrm{V}^{(\widehat{a,b})}_{(\widehat{1,1})}(\tau)} - \frac{1}{2}T\overline{\mathrm{V}^{(a,b)}_{(1,1)}(\tau)}\right)\tau\left[\left\langle\, Y_a^{\mathrm{lift}} \;:\; Y_b^{\mathrm{lift}}\,\right\rangle, X_S\right] \\
&- \frac{1}{2}\epsilon^2\overline{\mathrm{V}^{(a,b)}_{(2,1)}(\tau)}\,\tau\left[\left[Y_a^{\mathrm{lift}}, X_S\right], \left[Y_b^{\mathrm{lift}}, X_S\right]\right] \\
&- \frac{1}{2}\epsilon^2\left(\overline{\mathrm{V}^{(a,b,c)}_{(2,1,1)}(\tau)} - \frac{1}{2}T\overline{\mathrm{V}^{(a)}_{(1)}(\tau)}\,\overline{\mathrm{V}^{(b,c)}_{(1,1)}(\tau)}\right)\left\langle\, Y_a^{\mathrm{lift}} \;:\; \left\langle\, Y_b^{\mathrm{lift}} \;:\; Y_c^{\mathrm{lift}}\,\right\rangle\right\rangle \\
&+ \frac{1}{2}\epsilon^2\tau^2 X_S\cdot X_S + \frac{1}{2}\epsilon^2\overline{\mathrm{V}^{(a)}_{(1)}(\tau)}\,\tau^2 X_S\cdot\left[Y_a^{\mathrm{lift}}, X_S\right] + \frac{1}{2}\epsilon^2\overline{\mathrm{V}^{(a)}_{(1)}(\tau)}\,\tau^2\left[Y_a^{\mathrm{lift}}, X_S\right]\cdot X_S \\
&- \frac{1}{4}\epsilon^2\overline{\mathrm{V}^{(a,b)}_{(1,1)}(\tau)}\,\tau^2 X_S\cdot\left\langle\, Y_a^{\mathrm{lift}} \;:\; Y_b^{\mathrm{lift}}\,\right\rangle - \frac{1}{4}\epsilon^2\overline{\mathrm{V}^{(a,b)}_{(1,1)}(\tau)}\,\tau^2\left\langle\, Y_a^{\mathrm{lift}} \;:\; Y_b^{\mathrm{lift}}\,\right\rangle\cdot X_S \\
&+ \frac{1}{2}\epsilon^2\overline{\mathrm{V}^{(a)}_{(1)}(\tau)}\,\overline{\mathrm{V}^{(b)}_{(1)}(\tau)}\,\tau^2\left[Y_a^{\mathrm{lift}}, X_S\right]\cdot\left[Y_b^{\mathrm{lift}}, X_S\right] \\
&- \frac{1}{4}\epsilon^2\overline{\mathrm{V}^{(a)}_{(1)}(\tau)}\,\overline{\mathrm{V}^{(b,c)}_{(1,1)}(\tau)}\,\tau^2\left[Y_a^{\mathrm{lift}}, X_S\right]\cdot\left\langle\, Y_a^{\mathrm{lift}} \;:\; Y_b^{\mathrm{lift}}\,\right\rangle \\
&- \frac{1}{4}\epsilon^2\overline{\mathrm{V}^{(a,b)}_{(1,1)}(\tau)\mathrm{V}^{(c)}_{(1)}(\tau)}\,\tau^2\left\langle\, Y_a^{\mathrm{lift}} \;:\; Y_b^{\mathrm{lift}}\,\right\rangle\cdot\left[Y_a^{\mathrm{lift}}, X_S\right] \\
&+ \frac{1}{8}\epsilon^2\overline{\mathrm{V}^{(a,b)}_{(1,1)}(\tau)}\,\overline{\mathrm{V}^{(c,d)}_{(1,1)}(\tau)}\,\tau^2\left\langle\, Y_a^{\mathrm{lift}} \;:\; Y_b^{\mathrm{lift}}\,\right\rangle\cdot\left\langle\, Y_c^{\mathrm{lift}} \;:\; Y_d^{\mathrm{lift}}\,\right\rangle.
\end{aligned}
$$

To obtain $\mathrm{Trunc}_2\left(P(t)\right)$, the two truncated flows are composed and truncated. The final truncation keeps only the terms up to order 2 in ϵ,

$$
\begin{aligned}
&\mathrm{Trunc}_2\left(P(\tau)\right)\\
=\;&\\
&\mathrm{Id} + \epsilon \int_0^\tau \widetilde{\mathrm{V}}^{(a)}_{(1)}(\sigma)\,\mathrm{d}\sigma\,\left[Y_a^{\mathrm{lift}}, X_S\right] - \frac{1}{2}\epsilon \int_0^\tau \widetilde{\mathrm{V}}^{(a,b)}_{(1,1)}(\sigma)\,\mathrm{d}\sigma\,\left\langle\, Y_a^{\mathrm{lift}} \;:\; Y_b^{\mathrm{lift}} \,\right\rangle\\
&+ \epsilon^2 \left(\int_0^\tau \widetilde{\mathrm{V}}^{(a)}_{(2)}(\sigma)\,\mathrm{d}\sigma \;- \frac{1}{2}\tau \int_0^\tau \widetilde{\mathrm{V}}^{(a)}_{(1)}(\sigma)\,\mathrm{d}\sigma \;- \frac{1}{2}\tau \left(\mathrm{V}^{(a)}_{(2)}(\tau) - T\,\overline{\mathrm{V}^{(a)}_{(1)}(\tau)} \right) \right) \left[\left[Y_a^{\mathrm{lift}}, X_S \right], X_S \right]\\
&- \frac{1}{2}\epsilon^2 \left(\int_0^\tau \widetilde{\mathrm{V}}^{(\widehat{a,b})}_{(\widehat{1,1})}(\sigma)\,\mathrm{d}\sigma \;- \tau \int_0^\tau \widetilde{\mathrm{V}}^{(a,b)}_{(1,1)}(\sigma)\,\mathrm{d}\sigma \;- \frac{1}{2}\tau \left(\mathrm{V}^{(\widehat{a,b})}_{(\widehat{1,1})}(\tau) - T\,\overline{\mathrm{V}^{(a,b)}_{(1,1)}(\tau)} \right) \right)\\
&- \frac{1}{2}\epsilon^2 \int_0^\tau \widetilde{\mathrm{V}}^{(a,b)}_{(2,1)}(\sigma)\,\mathrm{d}\sigma\,\left[\left[Y_a^{\mathrm{lift}}, X_S \right], \left[Y_b^{\mathrm{lift}}, X_S \right] \right]\\
&+ \frac{1}{2}\epsilon^2 \left(\int_0^\tau \widetilde{\mathrm{V}}^{(a,b,c)}_{(2,1,1)}(\sigma)\,\mathrm{d}\sigma \;- \frac{1}{2} \left(\mathrm{V}^{(a,\widehat{b,c})}_{(2,\widehat{1,1})}(\tau) - T\,\tau\,\overline{\mathrm{V}^{(a)}_{(1)}(\tau)\mathrm{V}^{(b,c)}_{(1,1)}(\tau)} \right) \right) \left\langle\, Y_a^{\mathrm{lift}} \;:\; \langle\, Y_b \;:\; Y_c \,\rangle \,\right\rangle\\
&- \epsilon^2 \left(\mathrm{V}^{(a)}_{(2)}(\tau)\overline{\mathrm{V}^{(b)}_{(1)}(\tau)}\,\tau - \frac{1}{2}\mathrm{V}^{(a,b)}_{(2,2)}(\tau) - \frac{1}{2}\overline{\mathrm{V}^{(a)}_{(1)}(\tau)}\,\mathrm{V}^{(b)}_{(1)}(\tau)\,\tau^2 \right) \left[Y_a^{\mathrm{lift}}, X_S \right] \cdot \left[Y_b^{\mathrm{lift}}, X_S \right]\\
&+ \frac{1}{2}\epsilon^2 \left(\mathrm{V}^{(a)}_{(2)}(\tau)\overline{\mathrm{V}^{(b,c)}_{(1,1)}(\tau)}\,\tau - \frac{1}{2}\mathrm{V}^{(a,\widehat{b,c})}_{(2,\widehat{1,1})}(\tau) - \frac{1}{2}\overline{\mathrm{V}^{(a)}_{(1)}(\tau)}\,\overline{\mathrm{V}^{(b,c)}_{(1,1)}(\tau)}\,\tau^2 \right) \left[Y_a^{\mathrm{lift}}, X_S \right] \cdot \left\langle\, Y_b^{\mathrm{lift}} \;:\; Y_c^{\mathrm{lift}} \,\right\rangle\\
&+ \frac{1}{2}\epsilon^2 \left(\mathrm{V}^{(\widehat{a,b})}_{(\widehat{1,1})}\overline{\mathrm{V}^{(c)}_{(1)}(\tau)}\,\tau - \frac{1}{2}\mathrm{V}^{(\widehat{a,b},c)}_{(\widehat{1,1},2)}(\tau) - \frac{1}{2}\overline{\mathrm{V}^{(a,b)}_{(1,1)}(\tau)} \cdot \overline{\mathrm{V}^{(c)}_{(1)}(\tau)}\,\tau^2 \right) \left\langle\, Y_a^{\mathrm{lift}} \;:\; Y_b^{\mathrm{lift}} \,\right\rangle \cdot \left[Y_c^{\mathrm{lift}}, X_S \right]\\
&- \frac{1}{4}\epsilon^2 \left(\mathrm{V}^{(\widehat{a,b})}_{(\widehat{1,1})}(\tau)\overline{\mathrm{V}^{(b,c)}_{(1,1)}(\tau)}\,\tau - \frac{1}{2}\mathrm{V}^{(\widehat{a,b},\widehat{c,d})}_{(\widehat{1,1},\widehat{1,1})}(\tau) - \frac{1}{2}\overline{\mathrm{V}^{(a,b)}_{(1,1)}(\tau)}\,\overline{\mathrm{V}^{(b,c)}_{(1,1)}(\tau)}\,\tau^2 \right)\\
&\qquad\qquad \left\langle\, Y_a^{\mathrm{lift}} \;:\; Y_b^{\mathrm{lift}} \,\right\rangle \cdot \left\langle\, Y_c^{\mathrm{lift}} \;:\; Y_d^{\mathrm{lift}} \,\right\rangle .
\end{aligned}
$$

Chapter 6

Biomimetic and Biomechanical Control Systems

This chapter is devoted to constructing stabilizing controllers for a variety of biologically-inspired and biomechanical underactuated control systems. Recall that biomimetic and biomechanical control systems are not only uncontrollable in the linearization, but may also be difficult or impossible to linearize along feasible trajectories. Anholonomy or geometric phase is essential to the basic locomotive strategy of biological systems.

6.1 Amoeba

This example studies the locomotive abilities of a macroscopic "robot amoeba," idealized to swim in an irrotational, inviscid, and incompressible fluid. It is assumed that the amoeba is incapable of generating vorticity. The amoeba-like creature that will be studied here has been researched previously [103]. It is a system whose evolution best fits within the framework of nonholonomic systems with symmetry. The position and orientation of the body's center of mass are denoted by $g \in SE(2)$. The boundary of the amoeba body is controlled by the amoeba itself, and is the shape space. The shape corresponds to three deformation modes the amoeba is assumed to have, Figure 6.1. The description of the shape space is denoted by $r = (r^1, r^2, r^3) \in \mathbb{T}^3$. Given a radius of r_0, the amoeba shape is defined by the following function of φ,

Figure 6.1: Deformation modes of amoeba.

$$R(\varphi, r) = r_0[1 + \delta(r^1 \cos(2\varphi) + r^2 \cos(3\varphi) + r^3 \sin(3\varphi))].$$

Traversing around the amoeba's body via $\varphi \in [0, 2\pi)$ gives the radius at distinct points along the circumference. The parameter $\delta \in (0, 1)$ gives the nominal amplitude of the deformations.

The local form of the principle connection is

$$\mathcal{A}_{loc}(r)\dot{r} = \delta^2 \begin{bmatrix} r_0(1-\alpha)r_2 & r_0\, r_1 & 0 \\ r_0(1-\alpha)r_3 & 0 & r_0\, r_1 \\ 0 & \frac{-2\,\pi\, {r_0}^2\, \rho\, r_3}{M} & \frac{2\,\pi\, {r_0}^2\, \rho\, r_2}{M} \end{bmatrix} + O\left(\epsilon^3\right), \tag{6.1}$$

where $\alpha = (2\pi r_0^2 \rho)/(M + \pi r_0^2 \rho)$ [103]. The theory for reduced systems with a principle connection states that the curvature terms are responsible for the second order contribution of motion. Due to skew-symmetry,

only 3 terms are nonzero. They are

$$\begin{aligned}\mathcal{B}_{12}(r) = \delta^2 \begin{bmatrix} r_0\mu + \frac{r_0(1-\mu)2\pi r_0^2\rho}{M}(r_3)^2 \\ \frac{r_0(1-\mu)2\pi r_0^2\rho}{M} r_2\, r_3 \\ 0 \end{bmatrix}, \qquad \mathcal{B}_{13}(r) = \delta^2 \begin{bmatrix} -\frac{r_0(1-\mu)2\pi r_0^2\rho}{M} r_2 r_3 \\ r_0\mu - \frac{\mu r_0(1-\mu)2\pi r_0^2\rho}{M}(r_2)^2 \\ 0 \end{bmatrix}, \\ \mathcal{B}_{23}(r) = \frac{2\pi r_0^2\rho\delta^2}{M} \begin{bmatrix} r_0 r_1 r_3 \\ r_0 r_1 r_2 \\ 2 \end{bmatrix}.\end{aligned} \tag{6.2}$$

For oscillatory inputs with zero average, the average shape will be $r = (0,0,0)$. Entering this into the curvature term results in

$$\mathcal{B}_{12}(\overline{r}) = \left\{ \begin{array}{c} r0\mu\delta^2 \\ 0 \\ 0 \end{array} \right\}, \quad \mathcal{B}_{13}(\overline{r}) = \left\{ \begin{array}{c} 0 \\ r_0\mu\delta^2 \\ 0 \end{array} \right\}, \quad \mathcal{B}_{23}(\overline{r}) = \left\{ \begin{array}{c} 0 \\ 0 \\ \frac{4\pi r_0^2\rho\delta^2}{M} \end{array} \right\}. \tag{6.3}$$

Thus, the second order Lie brackets span the Lie algebra about the zero shape configuration, meaning that the amoeba is locally controllable. In [103] open loop translational controls were given, and in [160] a stabilizing feedback strategy was given for motion in the plane. Here we present a method to control the amoeba in $SE(2)$. A saturation function is included, to prevent the control values from being unrealistic. When the error exceeds a limit, the amoeba travels at a maximum allowable speed until the error falls within this limit. Thereafter, the system exponentially stabilizes to the desired position and orientation, $g_{des} \in SE(2)$. The control inputs are

$$\begin{aligned} v^1(t) &= -\sin(\omega_1 t), \\ v^2(t) &= -\frac{1}{\omega_1}\alpha_{12}^2 \cos(\omega_1 t) + \frac{M}{2\omega_2\pi r_0^2\rho\delta^2}\sin(\omega_2 t), \\ v^3(t) &= -\frac{1}{2}\alpha_{13}^3 \cos(\omega_1 t) + \frac{M}{2\omega_2\pi r_0^2\rho\delta^2}\alpha_{23}^3 \cos(\omega_2 t). \end{aligned} \tag{6.4}$$

With these controls, the amoeba undergoes oscillations of the first and second modes when at equilibrium ($\alpha_J^I = 0$), meaning that stabilization will be orbit stabilization. Notice that the oscillatory functions corresponding to the ω_2 frequency are multiplied by the factor $\frac{M}{2\pi r_0^2\rho\delta^2}$. Rotation for the amoeba is much easier from a control perspective; the curvature terms demonstrate that rotational capability is several orders of magnitude greater than the translational capability for the same magnitude of oscillation. If the control signal is not attenuated then the rotational motion over a period of oscillation will be too large. Systematic examination of the system from the perspective of averaging theory correctly identifies this problem and provides a mechanism to correct for it.

To define the α-parametrization an error function is needed. The group error and Lie algebra error are

$$g_{err} = g_{des}^{-1} g, \quad \text{and} \quad \xi_{err} \equiv \log\left(g_{err}\right), \tag{6.5}$$

where g_{des} is the desired group configuration and the logarithm function is $\log : G \to \mathfrak{se}2$. The α-parametrization is

$$\begin{aligned} \alpha_{12}^2 &= k_1 \xi_{err}^1, \\ \alpha_{13}^3 &= k_2 \xi_{err}^2, \\ \alpha_{23}^3 &= k_3 \xi_{err}^3. \end{aligned} \tag{6.6}$$

Snapshots of simulations demonstrating point stabilization and trajectory tracking can be found in Figure 6.2. The dot and cross located at the center of the amoeba depicts the amoeba's planar position and orientation, respectively. The simulation parameters were $M = 0.01\text{kg}$, $r_0 = 0.2\text{m}$, $\rho = 1000$, $\delta = 0.1$, $\epsilon = \frac{1}{5}$, $\omega_1 = 2$, $\omega_2 = 7$. The choice of ω_2 provides additional attenuation of the rotational control signal as discussed above (otherwise the choice $\omega_2 = 3$ would have been sufficient). The α parameters saturated at unity. The authors were unable to find any information regarding what the natural frequency of robotic amoeba locomotion should be; the choices of ω_1 and ω_2 were arbitrarily chosen, and are presumed to be somewhat plausible (and conservative). Taking into account the value of ϵ, the periods of the oscillations are $\frac{2\pi}{10}$, and $\frac{2\pi}{35}$ for the translational and rotational modes, respectively (about 1.6 Hz and 5.6 Hz). The amoeba is a slow locomotor, requiring approximately 1000 periods of oscillatory actuation to go approximately 10 body lengths.

For point stabilization the gains were $k = (75, 75, 6)$. Figure 6.3 has a time plot of the individual state values and also a log plot of the error signal norm, where the norm was chosen to be,

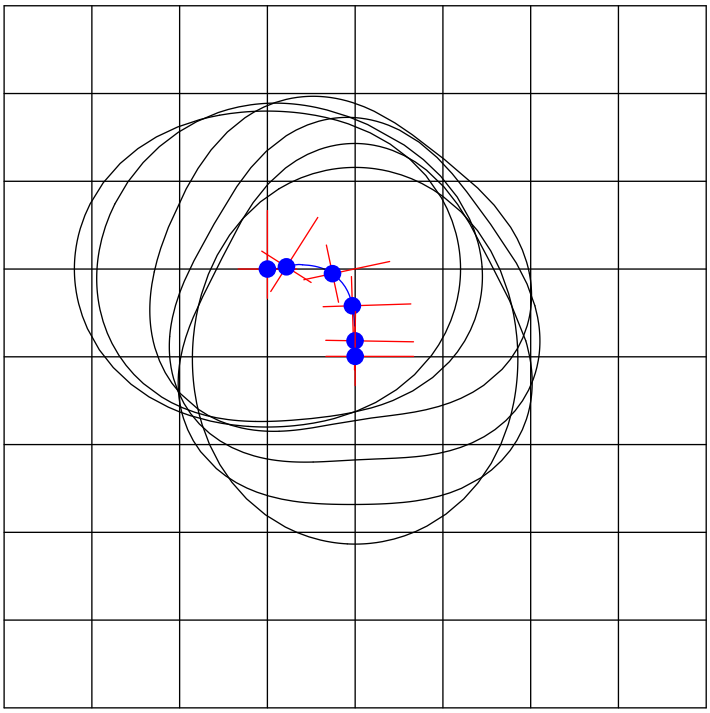
$$|g| = \sqrt{(g^1)^2 + (g^2)^2 + (g^3)^3}. \tag{6.7}$$

The linear slope demonstrates exponential convergence. The trajectory tracking gains were set to $k = (100, 100, 5)$. The trajectory tracking error is shown in Figure 6.4. The error in the translational states oscillates with an amplitude on the order of 2.5% of the body radius, r_0. The oscillating error in θ, which averages to about 1.3 degrees, is steady state error due to the ramp input, and can be improved by increasing the gain k_3, or by incorporating a feedforward term. This highlights another nice feature of the averaging method for control of nonlinear systems. By resulting in a linearized averaged expansion, it is possible to understand elements of the system response in terms of standard linear control ideas.

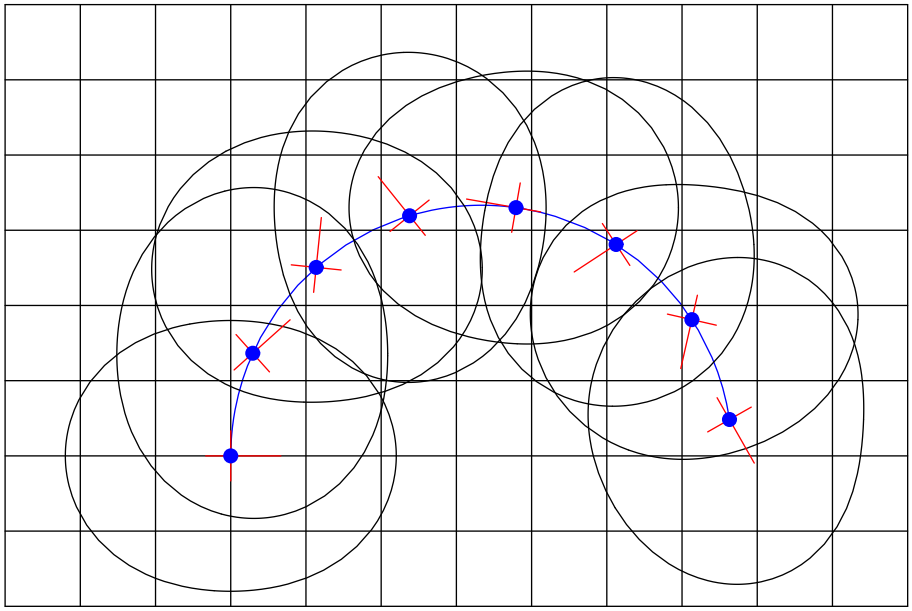
Although the controls are discontinuous for the amoeba example, what is important is the shape space evolution. Integrating the velocities leads to the following shape space evolution,

$$\begin{aligned} r^1(t) &= -\frac{\epsilon}{\omega_1}\cos(\omega_1 t) \\ r^2(t) &= -\frac{\epsilon}{\omega_1^2}\alpha_{12}^2 \sin(\omega_1 t) + \frac{\epsilon M}{2\omega_2^2 \pi r_0^2 \rho \delta^2}\cos(\omega_2 t) \\ r^3(t) &= -\frac{\epsilon}{2\omega_1}\alpha_{13}^3 \sin(\omega_1 t) + \frac{\epsilon M}{2\omega_2^2 \pi r_0^2 \rho \delta^2}\alpha_{23}^3 \sin(\omega_2 t). \end{aligned} \tag{6.8}$$

Discontinuous changes of the α_J^I at the times $t = kT$, $k\mathbb{Z}^+$, are continuous because the sine function evaluates to zero. For biomimetic systems it is important that the controls represent continuous and feasible shape space trajectories. Command of the shape space has been successfully done with an experimental carangiform fish [160].



(a) Snapshots at $t = 0,\ 5,\ 15,\ 25,\ 35,$ and 45 seconds.



(b) Snapshot interval is 75 seconds.

Figure 6.2: Point stabilization and trajectory tracking for amoeba (grid spacing $0.1\mathrm{m}$).

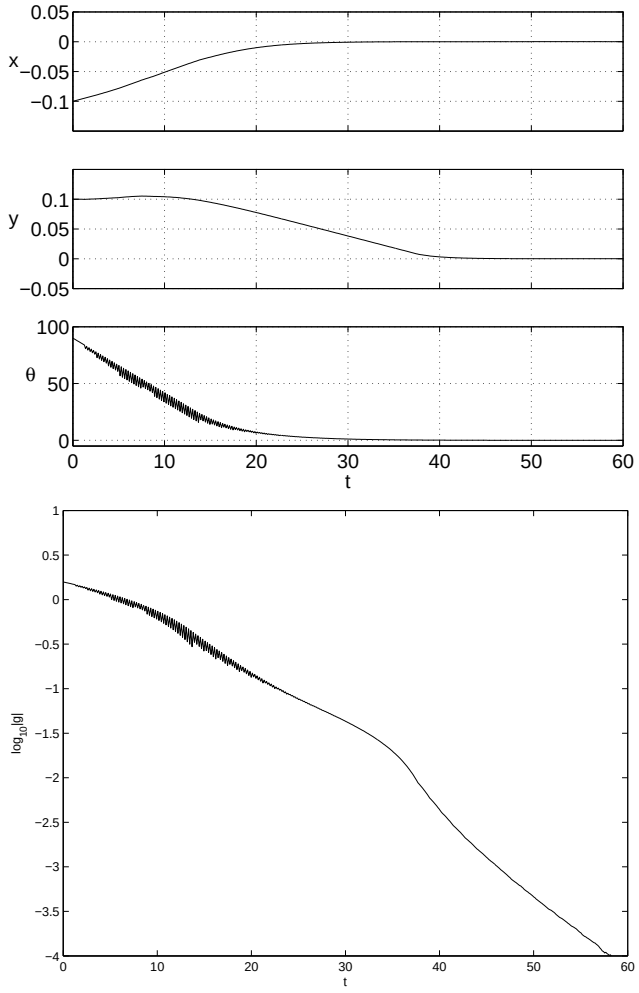


Figure 6.3: Amoeba point stabilization.

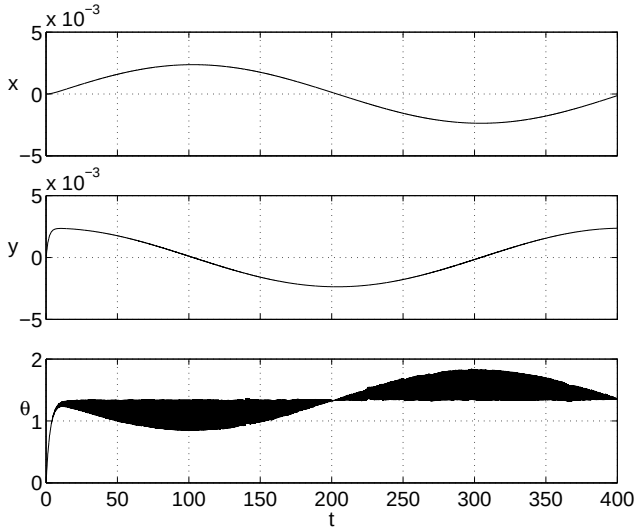


Figure 6.4: Amoeba trajectory tracking error.

6.2 Ball and Plate Control Problem

The ball and plate control problem is an idealization of two-finger manipulation of a sphere. A sphere is located between two plates, with the constraint that it rolls without slipping. The system is symmetric with respect the right action (spatial symmetry). Its derivation is discussed in Bloch et al. [15]. It is assumed that the sphere is initially at rest, and the fingers are at the origin. The goal is to utilize the fingers to manipulate the ball to achieve a desired orientation. The Lie group describing the ball configuration is $SO(3)$, and the shape space is $\mathbb{R}^2$, specifying the planar position of the top finger. The principal bundle is then, $Q = \mathbb{R}^2 \times SO(3)$.

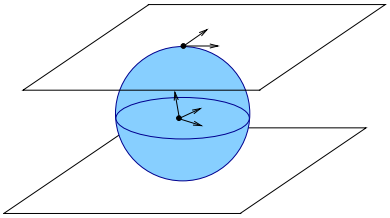


Figure 6.5: Ball and Plate.

The local form of the connection is

$$\mathcal{A}_{loc}(r) = \left\{ \begin{array}{c} \frac{1}{\rho}\mathbf{d}r^1 \\ \frac{1}{\rho}\mathbf{d}r^2 \\ 0 \end{array} \right\}. \tag{6.9}$$

The two fingers must remain near the origin, therefore we seek strong fiber controllability of the control system [67]. A system with group symmetries is strongly fiber controllable if the local curvature form and its higher order covariant derivative span the Lie algebra, c.f. Section 3.2.2. Since the local connection form is constant, Cartan's structure equations imply that the Lie brackets of the local connection will give these covariant derivatives. The curvature is

$$\mathcal{B}_{loc}(r) = \left\{ \begin{array}{c} 0 \\ 0 \\ \frac{1}{\rho^2}\mathbf{d}r^1 \wedge \mathbf{d}r^2 \end{array} \right\}, \tag{6.10}$$

and the covariant derivative of the curvature is

$$\nabla\mathcal{B}_{loc}(r) = \left\{ \begin{array}{c} \frac{1}{\rho^3}\mathbf{d}r^1 \wedge \mathbf{d}r^2 \wedge \mathbf{d}r^2 \\ \frac{1}{\rho^3}\mathbf{d}r^1 \wedge \mathbf{d}r^2 \wedge \mathbf{d}r^1 \\ 0 \end{array} \right\}. \tag{6.11}$$

The higher-order covariant derivatives span the Lie algebra, therefore the system is locally controllable. Translating the curvature calculations to Jacobi-Lie brackets of horizontally lifted vector fields, the important Jacobi-Lie brackets are

$$\left[Y_1^h, Y_2^h\right], \quad \left[Y_1^h, \left[Y_2^h, Y_1^h\right]\right], \quad \text{and} \quad \left[Y_2^h, \left[Y_1^h, Y_2^h\right]\right], \tag{6.12}$$

whose corresponding averaged coefficients are

$$\overline{\mathrm{V}^{(1,2)}_{(1,0)}(t)}, \quad \overline{\mathrm{V}^{(1,2,1)}_{(1,1,0)}(t)}, \quad \text{and} \quad \overline{\mathrm{V}^{(2,1,2)}_{(1,1,0)}(t)}. \tag{6.13}$$

According to Lemmas 2 and 3, the oscillatory input functions take the form,

$$v^1(t) = \alpha^1_{12}\sin(\omega_{12}t) + \alpha^1_{121}\sin(2\omega_{121}t) + \alpha^1_{212}\cos(\omega_{212}t), \text{ and} \tag{6.14}$$

$$v^2(t) = \alpha^2_{12}\cos(\omega_{12}t) + \alpha^2_{121}\cos(\omega_{121}t) + \alpha^2_{212}\sin(2\omega_{212}t). \tag{6.15}$$

The underactuated driftless nonlinear control system,

$$\dot{q} = Y_a^h u^a, \tag{6.16}$$

is averaged to obtain the fully fiber-actuated driftless control system,

$$\dot{z} = \epsilon^2 \overline{\mathrm{V}^{(a,b)}_{(1,0)}(t)} \mathcal{B}_{ab} + \frac{1}{3}\epsilon^3 \overline{\mathrm{V}^{(a,b,c)}_{(1,1,0)}(t)} \mathcal{B}_{abc}, \tag{6.17}$$

as per Equation (3.72). To define the α-parametrization an error function is needed. The group error and group error function are

$$g_{err} = g_{des}^{-1} g, \quad \text{and} \quad \xi_{err} \equiv \log\left(g_{err}\right), \tag{6.18}$$

where g_{des} is the desired group configuration and the logarithm function is $\log : G \to \mathfrak{se}2$. The α-parametrization is

$$\alpha^1_{12} = \mathrm{sign}(k_1 \xi^1_{err}) \sqrt{|k_1 \xi^1_{err}|} \qquad \alpha^2_{12} = \frac{1}{2}\sqrt{|k_1 \xi^1_{err}|} \tag{6.19}$$

$$\alpha^1_{121} = \sqrt{|k_2 \xi^2_{err}|} \qquad \alpha^2_{121} = \mathrm{sign}(k_2 \xi^2_{err}) \sqrt{|k_2 \xi^2_{err}|} \tag{6.20}$$

$$\alpha^1_{212} = \sqrt{|k_3 \xi^3_{err}|} \qquad \alpha^2_{212} = \mathrm{sign}(k_3 \xi^3_{err}) \sqrt{|k_3 \xi^3_{err}|} \tag{6.21}$$

Integrating the control functions, the shape space evolution is

$$r^1(t) = \frac{\alpha^1_{12}\epsilon}{\omega_{12}}(1 - \cos(\omega_{12} t)) + \frac{\alpha^1_{121}\epsilon}{2\omega_{121}}(1 - \cos(2\omega_{121} t)) + \frac{\alpha^1_{212}\epsilon}{\omega_{212}} \sin(\omega_{212} t), \text{ and} \tag{6.22}$$

$$r^2(t) = \frac{\alpha^2_{12}\epsilon}{\omega_{12}} \sin(\omega_{12} t) + \frac{\alpha^2_{121}\epsilon}{\omega_{121}}(1 - \cos(\omega_{121} t)) + \frac{\alpha^2_{212}\epsilon}{2\omega_{212}} \sin(2\omega_{212} t). \tag{6.23}$$

At the end of each period of evolution, the shape space returns to the initial configuration, $r_0 = (0, 0)$.

Snapshots of a point stabilization simulation are given in Figure 6.7. The ball is drawn as a box to help visualize the stabilizing trajectory in $SO(3)$. The static line indicates the desired location of the body $z-$axis. The exponential convergence of the average can be seen in Figure 6.6. The circles are the logarithm of the error feedback after one period of actuation, whereas the continuous oscillatory line is the logarithm of the actual time-varying error feedback. The simulation parameters are $\rho = 1$, $\epsilon = 1$, $\omega_{12} = 2$, $\omega_{121} = 3$, $\omega_{212} = 4$, and $k = (1, 3, 13)$. To achieve trajectory tracking for the discretized strategy, the choice of ϵ must be smaller, or equivalently, the frequencies ω_J must be larger. Higher frequency inputs will also result in smaller orbits, hence less deviation from the desired trajectory. Alternatively, a continuous controller could be derived using the ideas from Chapter 3.

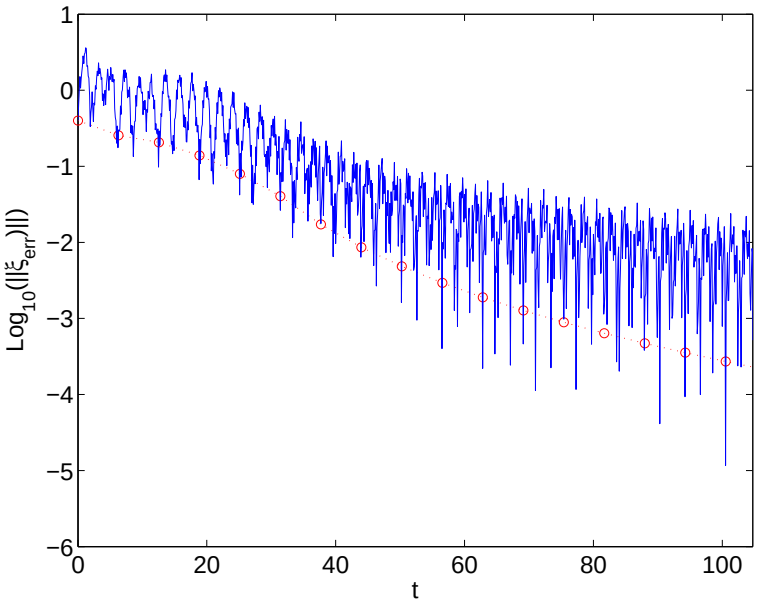


Figure 6.6: Ball and plate stabilization error, $\log_{10}(||\xi_{err}||)$.

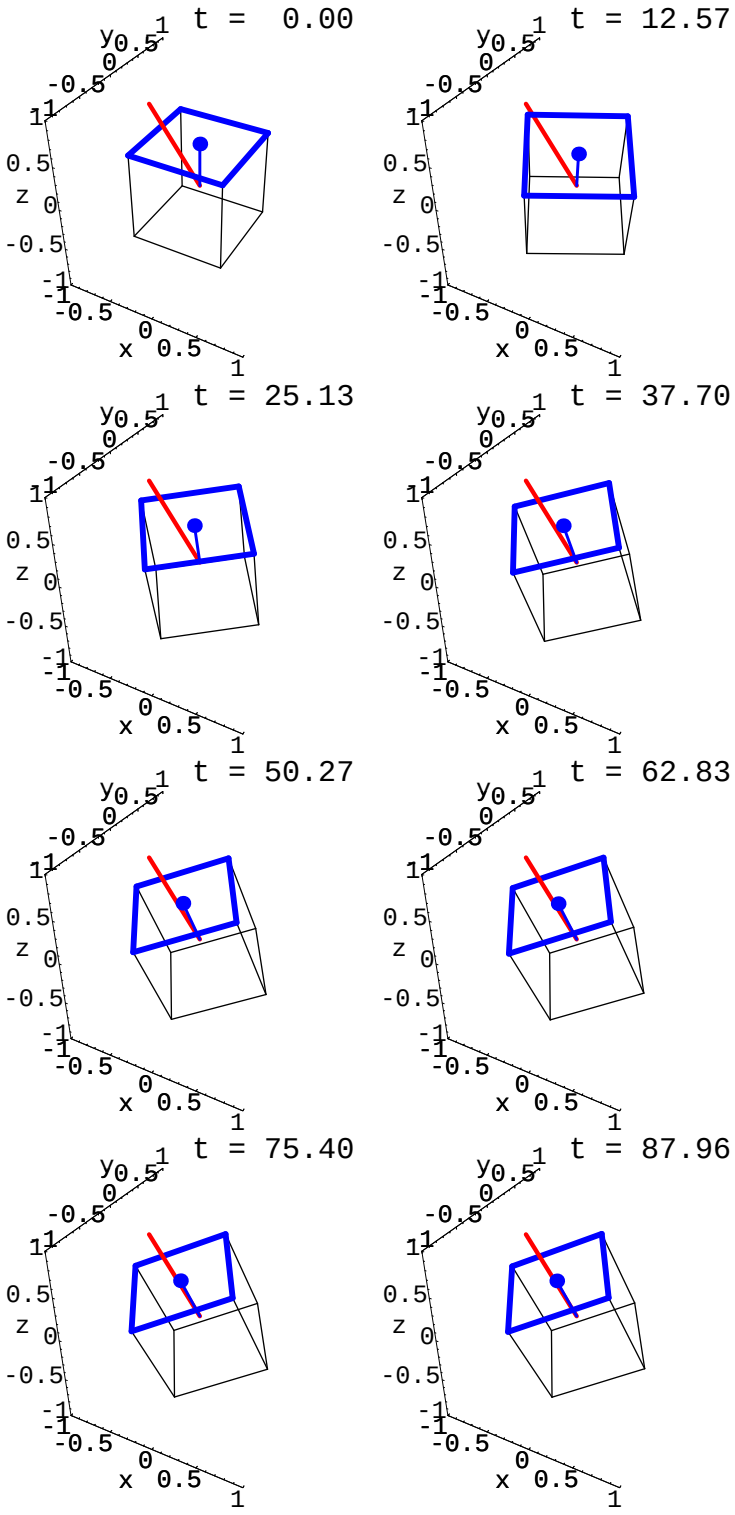


Figure 6.7: Ball and plate point stabilization snapshots.

6.3 Kinematic Biped

As a demonstration of the theory for systems with time-discontinuous evolution, we consider a "kinematic biped" that is motivated by the underactuated crawling robot of Ito et. al. [57]. The model (see Fig. 6.8) operates on flat ground. When either of the two vertically moving peg-legs contact the ground, we assume the crawler pivots about the contact. Casters on the four corners provide balance, thereby planarizing the motion. A rotating momentum wheel is the main motion generator. It is hereafter called the *rotor*. By properly switching the feet contacts in phase with the rotor's spin, the biped can maneuver.

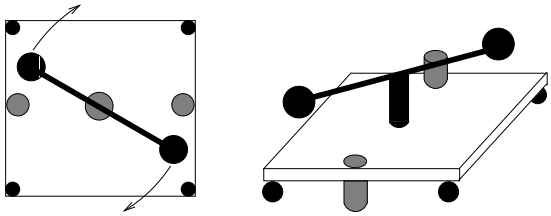


Figure 6.8: Top and side views of the kinematic biped.

In keeping with the language of [127], the configuration space, Q, is the principal bundle $Q = G \times B$, where G and B denote the group and shape spaces. The shape variables are the rotor angle, ψ, and the feet state, $N = \{ L, R \}$. The symbols L and R denote when the left or right foot is put down. The group is $G = SE(2)$, the set of planar rigid body motions.

The equations of motion are derived using the approach of Bloch et al. [15]. This method uses the principal bundle structure and the group invariance of the Lagrangian and constraints to compute the "reduced" Euler-Lagrange equations. The Lagrangian for the system is:

$$L(q, \dot{q}) = \frac{1}{2} M(\dot{x}^2 + \dot{y}^2) + \frac{1}{2} J \dot{\theta}^2 + \frac{1}{2} J_m (\dot{\theta} + \dot{\psi})^2, \tag{6.24}$$

where M is the robot's total mass, J (J_m) is the inertia of the body (rotor). When the right leg is planted, the resulting holonomic constraint on the robot's motion is:

$$\dot{x} = d \sin\theta \, \dot{\theta}, \text{ and } \dot{y} = -d \cos\theta \, \dot{\theta}.$$

Since the Lagrangian and constraints are G-invariant, they may be reduced to the Lie algebra; equivalent to giving a description in body coordinates. The reduced constraints are

$$\xi^1 = 0 \text{ and } \xi^2 = -d\xi^3.$$

Since the constraints span 2 dimensions of the 3-dimensional Lie algebra, $\mathfrak{se}(2)$, Noether's theorem implies a conserved momentum (about the foot's axis). Conservation of momentum is given more abstractly by the equation, $\mathcal{A}^{nhc} \equiv (\mathbb{I}^c)^{-1} \, \mathrm{J}^{nhc} = (\mathbb{I}^c)^{-1} \, p$, where $\mathbb{I}^c$ is the inertia of the system and p is the conserved momentum. More explicitly,

$$M d \, \xi^2 + (J + J_m) \, \xi^3 + J_m \, \dot{\psi} = p. \tag{6.25}$$

This equation is combined with the kinematic constraints to obtain the constraint equations,

$$\mathrm{E}(r)\xi + \mathrm{F}(r)\dot{r} = \mathrm{G}(r)p,$$

where

$$\mathrm{E}(r) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & d \\ 0 & Md & J + J_m \end{bmatrix} \quad \mathrm{F}(r) = \begin{bmatrix} 0 \\ 0 \\ J_m \end{bmatrix} \quad \mathrm{G}(r) = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

The local connection form, describing the system constraints in body coordinates, is found by solving for

the Lie algebra element, ξ, giving the reconstruction equation,

$$\xi = -\mathcal{A}_{loc}(r)\dot{r} + \Gamma_{loc}(r)p, \tag{6.26}$$

with the local connection and the affine forms equal to

$$\mathcal{A}_{loc}(r) = \frac{J_m}{J + J_m + Md^2}\begin{bmatrix} 0 \\ -d \\ 1 \end{bmatrix}, \quad \text{and} \quad \Gamma_{loc}(r) = \frac{1}{J + J_m + Md^2}\begin{bmatrix} 0 \\ -d \\ 1 \end{bmatrix}. \tag{6.27}$$

Continuing with the method of [15], one defines an orthogonal basis (in the Lagrangian metric) for the Lie algebra:

$$e_1(r) = \begin{bmatrix} 0 \\ -d \\ 1 \end{bmatrix}, \quad e_2(r) = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad e_3(r) = \begin{bmatrix} 0 \\ \frac{J+J_m}{Md} \\ 1 \end{bmatrix}.$$

The basis is used to derive the momentum equation,

$$\frac{\mathrm{d}}{\mathrm{d}t}p_1 = 0, \tag{6.28}$$

and also the shape space dynamics,

$$\left(\frac{1}{2}J_m + \frac{3}{2}\frac{J_r^2}{J_b + J_r + Md^2}\right)\ddot{\psi} = \tau, \tag{6.29}$$

where τ is the torque on the rotor. For zero initial momentum, the reconstruction equation becomes

$$\xi = -\mathcal{A}_{loc}(\psi)\dot{\psi}, \tag{6.30}$$

where $\mathcal{A}_{loc}(\psi)$ will be one of two choices, according to which foot [1] is on the ground

$$\begin{aligned} \text{right} : \mathcal{A}_{loc}(\psi) &= \frac{J_m}{J + J_m + Md^2}\begin{bmatrix} 0 \\ d \\ -1 \end{bmatrix}, \\ \text{left} : \mathcal{A}_{loc}(\psi) &= \frac{-J_m}{J + J_m + Md^2}\begin{bmatrix} 0 \\ d \\ 1 \end{bmatrix}. \end{aligned} \tag{6.31}$$

Locomotion and Controllability. By appropriate cycling of the feet and phasing of the rotor, it is possible to (i) rotate in place (ii) move forward (iii) move laterally. The existence of these motions can be shown by our averaging theory. Choosing $\dot{\psi}$ and the foot placements as inputs, Equation (6.26) takes the form

$$\xi = \mathbb{A}_1 v^1(t) + \mathbb{A}_2 v^2(t) \tag{6.32}$$

where

$$\mathbb{A}_1 = (\mathcal{A}_{loc})_{\text{right}} \quad , \quad \mathbb{A}_2 = (\mathcal{A}_{loc})_{\text{left}}\,, \tag{6.33}$$

and the functions $v^1(t)$ and $v^2(t)$ cannot be non-zero simultaneously. Further, we assume that foot placement transitions occur when at the robot is at rest. Since the functions $v^a(t)$ will be chosen to be periodic functions

[1] The calculations for the opposite foot merely changes a sign on the constraints. The trivial dynamics of the momentum state are not modified, nor are the shape space dynamics.

of time, we can use averaging theory to analyze the crawler's motion.

To demonstrate the crawler's controllability, note that the two distinct foot placement possibilities represent two linearly independent vectors in the Lie algebra. A third linearly independent possibility must exists for controllability. From Bullo and Zefran [27], it suffices to show that the Lie bracket between the two motions results in a lineary independent direction. Since

$$[\mathbb{A}_1, \mathbb{A}_2] = -\begin{bmatrix} 2d & 0 & 0 \end{bmatrix}^T, \tag{6.34}$$

the system is configuration controllable, therefore enabling point-to-point repositioning and reorientation. Matsuno et. al. [105] reached the same conclusion for a similar crawling robot, but without utilizing the inherent symmetries of the system to reduce to the Lie algebra. The geometry of this biped also avoids the singularity associated with their robot analysis.

Forward Gait. For forward motion, let $\dot{\psi}(t) = \sin(t)$. The right foot starts in the down position, while the left foot is placed at $t = \pi$, when the rotor reverses direction. The periodic functions $v^a(t)$ decompose accordingly,

$$\begin{aligned} v^1(t) &= \begin{cases} \sin(t) & \text{for } 0 \le t < \pi \\ 0 & \text{for } \pi \le t \le 2\pi \end{cases}, \\ v^2(t) &= \begin{cases} 0 & \text{for } 0 \le t < \pi \\ \sin(t) & \text{for } \pi \le t \le 2\pi \end{cases}. \end{aligned} \tag{6.35}$$

A non-zero average, implies that first order averaging theory will suffice. The average is

$$\eta = \mathbb{A}_1\overline{v^1(t)} + \mathbb{A}_1\overline{v^2(t)} = \mathbb{A}_1\frac{1}{\pi} - \mathbb{A}_2\frac{1}{\pi} = \begin{bmatrix} 0 \\ \frac{2d}{\pi} \\ 0 \end{bmatrix}. \tag{6.36}$$

Rotate Gait. Rotational motion is similar, but does not require a reversal of the rotor's spin direction, i.e., $\dot{\psi} = 1 - \cos(2t)$. The input functions are

$$\begin{aligned} v^1(t) &= \begin{cases} 1 - \cos(2t) & \text{for } 0 \le t < \pi \\ 0 & \text{for } \pi \le t \le 2\pi \end{cases}, \\ v^2(t) &= \begin{cases} 0 & \text{for } 0 \le t < \pi \\ 1 - \cos(2t) & \text{for } \pi \le t \le 2\pi \end{cases}. \end{aligned} \tag{6.37}$$

The first order average is

$$\eta = \mathbb{A}_1\overline{v^1(t)} + \mathbb{A}_1\overline{v^2(t)} = \mathbb{A}_1\frac{1}{2} + \mathbb{A}_2\frac{1}{2} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}. \tag{6.38}$$

Sideways Gait. This gait is subtler, as it arises from a higher order dynamical effect. The rotor follows the function

$$\dot{\psi}(t) = \begin{cases} \sin(2t) & 0 \le t < \pi \\ -\sin(2t) & \pi \le t \le 2\pi \end{cases}, \tag{6.39}$$

with foot placements occuring at the zero crossings of this function. The corresponding input functions are

$$v^1(t) = \begin{cases} \sin(2t) & 0 \le t \le \pi/2 \\ -\sin(2t) & \pi \le t \le 3\pi/2 \\ 0 & \text{otherwise} \end{cases}$$
$$v^2(t) = \begin{cases} \sin(2t) & \pi/2 \le t \le \pi \\ -\sin(2t) & 3\pi \le t \le 2\pi \\ 0 & \text{otherwise} \end{cases} \tag{6.40}$$

The averages of these functions vanish. Hence, second-order averaging methods must be used. The averaged system evolution is given by

$$\xi = \frac{1}{2}\overline{\mathrm{V}^{(a,b)}_{(1,0)}(t)}\,[\mathbb{A}_a, \mathbb{A}_b]\,, \tag{6.41}$$

where the only non-vanishing second order averaged coefficent, $\overline{\mathrm{V}^{(a,b)}_{(1,0)}(t)}$, is:

$$\overline{\mathrm{V}^{(1,2)}_{(1,0)}(t)} = \frac{1}{T}\int_0^T \int_0^t v^1(\tau)\,\mathrm{d}\tau\; v^2(t)\,\mathrm{d}t\; = -\frac{1}{2\pi},$$

and it's skew-symmetric partner, $\overline{\mathrm{V}^{(2,1)}_{(1,0)}(t)}$. The averaged evolution of the system is

$$\eta = \frac{1}{2}\overline{\mathrm{V}^{(1,2)}_{(1,0)}}\,[\mathbb{A}_1, \mathbb{A}_2] + \frac{1}{2}\overline{\mathrm{V}^{(2,1)}_{(1,0)}}\,[\mathbb{A}_2, \mathbb{A}_1] = -\frac{1}{4\pi}\cdot\begin{bmatrix}-2d\\0\\0\end{bmatrix} + \frac{1}{4\pi}\cdot\begin{bmatrix}2d\\0\\0\end{bmatrix} = \begin{bmatrix}\frac{d}{\pi}\\0\\0\end{bmatrix}, \tag{6.42}$$

which is sideways translation for the discrete crawler.

The various gaits are simulated in Fig 6.9. The forward gait moves the kinematic biped forward at an angle due to neglected second-order effects. Feedback control will be added below to correct for existing gait errors and other disturbances.

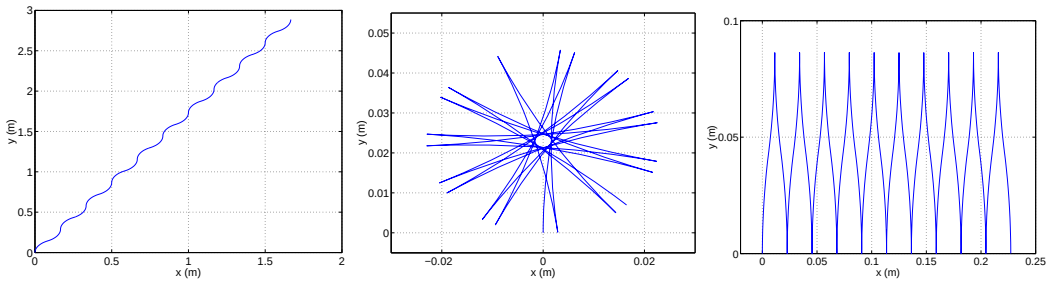


Figure 6.9: Gaits for the kinematic biped.

Stabilization. Define the rotor's control function to be

$$\psi(t) = \alpha_1 \sin(t) + \frac{1}{2}\alpha_2(1 - \cos(2t)), \tag{6.43}$$

with α_1 and α_2 to be chosen. The foot constraints transition at $t = (2k+1)\pi$, $k \in \mathbb{Z}$. The parameter α_1 controls forward motion, while α_2 controls rotational motion, allowing a superposition of the forward and rotate gaits.

The configuration error is defined to be $g_{err} \equiv g^{-1} g_{des}$, where $g_{des} \in SE(2)$ is the goal configuration and $g \in SE(2)$ the current configuration. Define error functions

$$\begin{aligned} \alpha_1 &= k_2\, y_{err} \\ \alpha_2 &= k_1 \tan^{-1}(x_{err}/y_{err}) + k_3 \theta_{err}. \end{aligned} \tag{6.44}$$

where the k_i are chosen according to Theorem 16. Note that when the configuration error is zero, the terms (α_1, α_2) vanish. Hence, both the original and averaged systems share the same equilibria. Therefore, via Corollary 2, the system exponentially stabilizes to the goal, achieving point-to-point repositioning (control of orientation is possible, but more complicated).

Using the gains $k = (3, 8, 0)$ results in the stabilizating trajectory pictured in Fig. 6.10. The biped started at $(x, y, \theta) = (-1, 5, \frac{\pi}{5})$ and was commanded to the origin. Fig. 6.11 shows superimposed snapshots from the simulation. The small dot on either side of the biped's body represents the active foot constraint. The gridlines are spaced 1 unit apart. Orientation control can also be included, but would require a slightly more sophisticated controller.

Our stabilization theorem also allows for trajectory tracking—the goal point is time varying. Fig. 6.12(a) demonstrates trajectory tracking for a straight path. In Fig. 6.12(b) the crawler must track the same trajectory, but starts with a large initial error of $(x, y, \theta) = (-1, 5, \frac{\pi}{5})$. It successfully moves toward the trajectory, then proceeds to track it. There is a small steady state error.

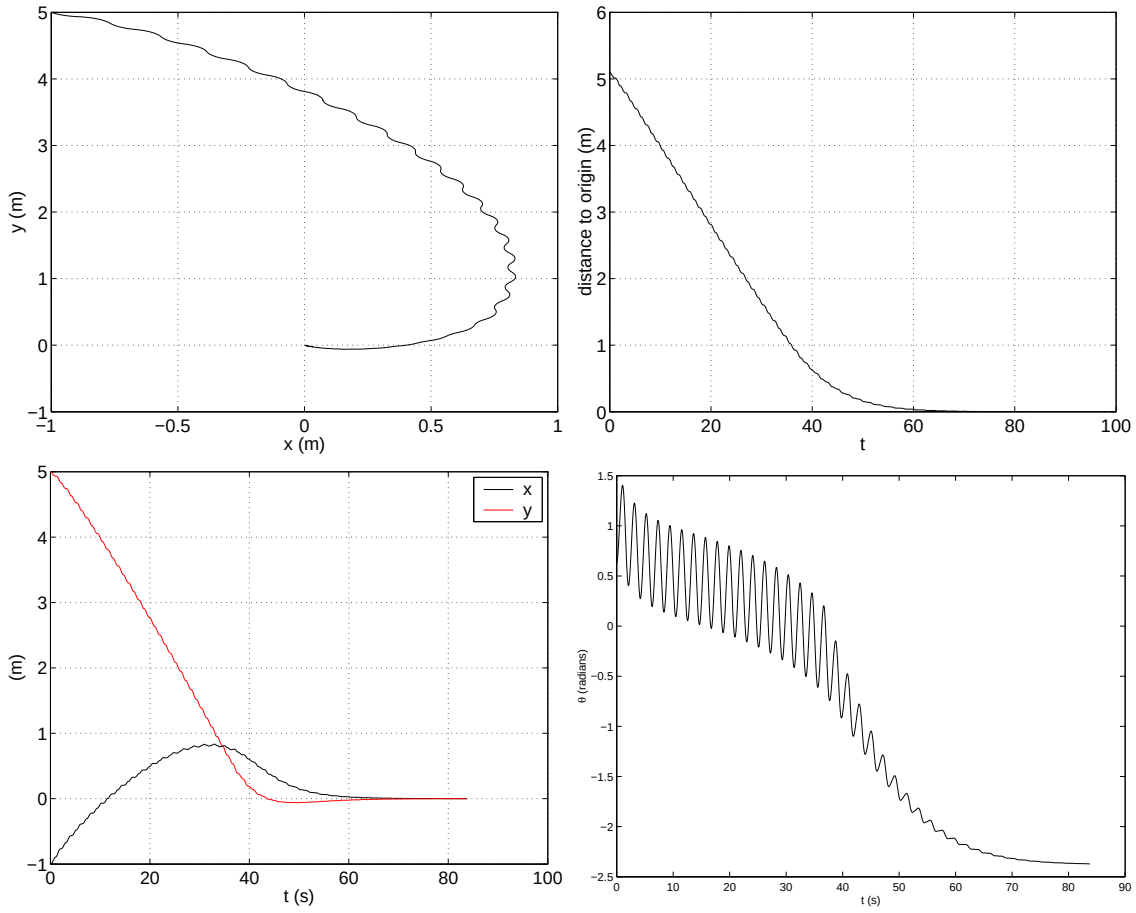


Figure 6.10: Point stabilization; parametric plot, magnitude plot, $x(t)$ and $y(t)$, and $\theta(t)$ plots.

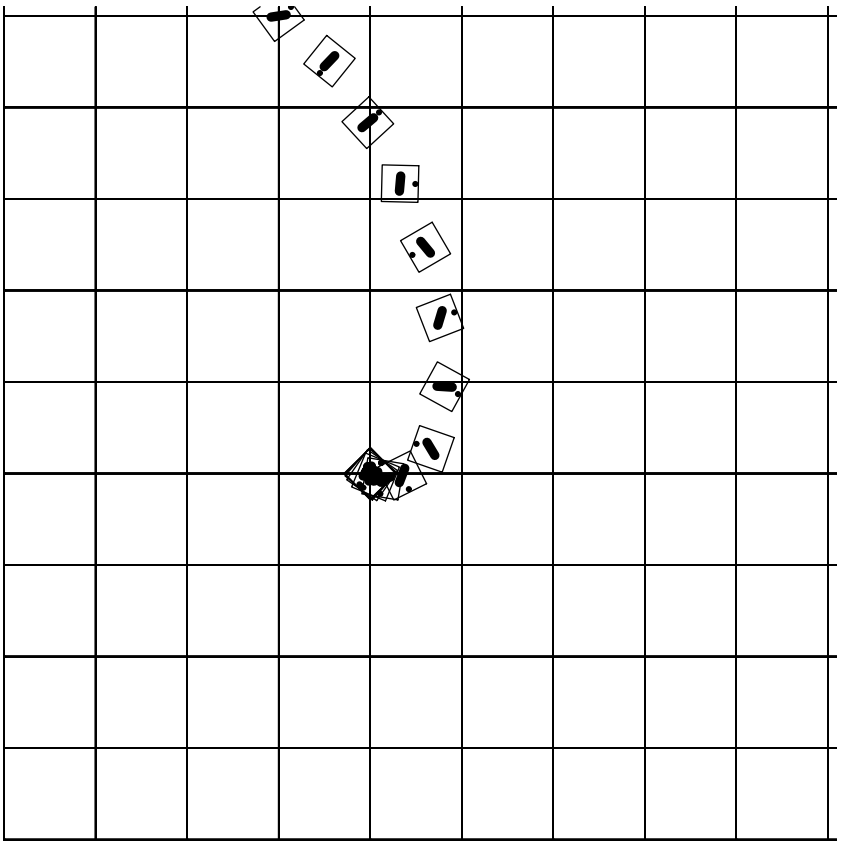
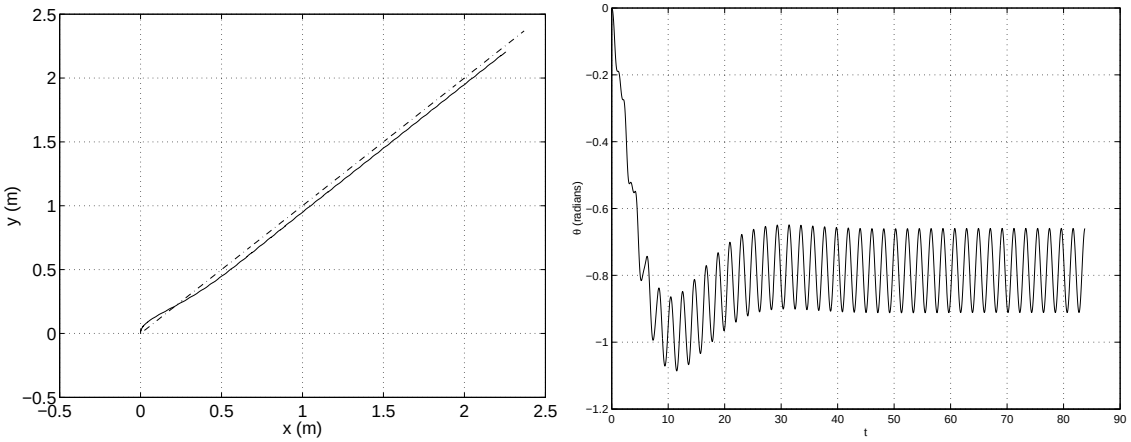
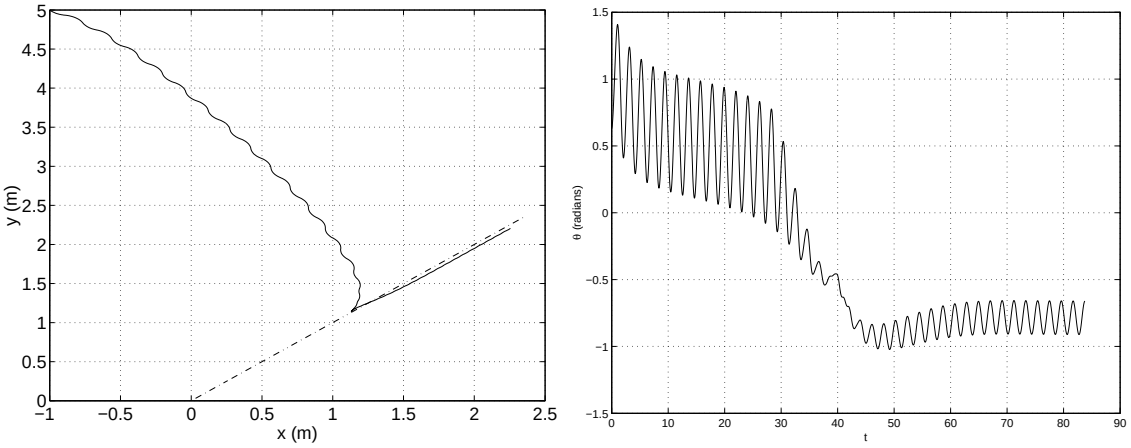


Figure 6.11: Snapshots of point stabilization of kinematic biped.



(a)



(b)

Figure 6.12: Trajectory tracking for kinematic biped.

6.4 The Snakeboard

Since the snakeboard model was introduced and described in detail in [87], we only summarize here its equations. The snakeboard represents an underactuated system with drift—traditionally a difficult control problem. In the simplified model used here, the snakeboard consists of a rigid body to which there is attached a momentum wheel (or rotor), and two pairs of wheels. The wheels axles are coupled to the main body by independently controllable revolute joints (see Fig. 6.13).

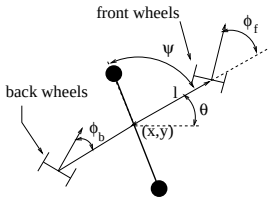


Figure 6.13: Model of the snakeboard.

The configuration space for this system is $[x, y, \theta, \psi, \phi_b, \phi_f] \in SE(2) \times \mathbb{T}^3$ where $g = [x, y, \theta]^T$ is the position and orientation of the board with respect to a fixed reference frame, ψ is the angle of the rotor relative to the board, and ϕ_b and ϕ_f are, respectively, the angles subtended by the back and front wheels (relative to the body). For simplicity, we restrict the front and back wheel motions so that $\phi_b = \phi = -\phi_f$. The rotor and wheels are actuated. The wheels provide nonholonomic constraints on the snakeboard's motions. However, because it is not fully constrained, the snakeboard can build up momentum and coast.

After feedback linearization on the directly controlled states, the equations of motion are

$$\begin{array}{rcl} \ddot{\psi} & = & u_1, \\ \ddot{\phi} & = & u_2, \\ \dot{g} & = & g\left(-\mathcal{A}_{loc}(r)\dot{r} + (\mathbb{I}^c(r))^{-1}\, p\right), \\ \dot{p} & = & 2J_r \cos^2(\phi)\dot{\phi}\dot{\psi} - \tan(\phi)\dot{\phi}p, \end{array} \tag{6.45}$$

where

$$\mathcal{A}_{loc}(r) = \left\{ \begin{array}{c} -\frac{J_r}{2Ml}\cos^2(\phi)\tan(\phi) \\ 0 \\ -\frac{J_r}{2Ml^2}\sin(2\phi)\tan(\phi) \end{array} \right\},$$

$$\mathbb{I}^c(r) = \left\{ \begin{array}{c} \frac{1}{2Ml} \\ 0 \\ -\frac{1}{2Ml}\tan(\phi) \end{array} \right\},$$

and $r = [\psi, \phi_b, \phi_f]^T$.

The control inputs consist of state feedback for the directly controlled states, coupled with periodic terms to influence the indirectly controlled states:

$$u_1 = -2\psi - 3\dot{\psi} + \frac{1}{\epsilon}v^1(\frac{t}{\epsilon}), \;\; u_2 = -2\phi - 3\dot{\phi} + \frac{1}{2}\alpha_2\epsilon + \frac{1}{\epsilon}v^2(\frac{t}{\epsilon}),$$

which results in the autonomous contribution,

$$Y_0 = -\left(2\psi + 3\dot{\psi}\right)\frac{\partial}{\partial\dot{\psi}} - \left(2\phi + 3\dot{\phi} - \frac{1}{2}\alpha_2\epsilon\right)\frac{\partial}{\partial\dot{\phi}}.$$

The time-periodic control input contribution is

$$\begin{array}{rl} v^1(t) & = \alpha_{12}^1\cos(t), \\ v^2(t) & = \alpha_{12}^2\cos(t), \end{array} \tag{6.46}$$

with α_1 a free parameter. The averaged equations become

$$\begin{array}{rcl} \ddot{\psi} & = & -2\psi - 3\dot{\psi}, \\ \ddot{\phi} & = & -2\phi - 3\dot{\phi} + \frac{1}{2}\alpha_2\epsilon, \\ \dot{g} & = & g\left(-\mathcal{A}_{loc}(r)\dot{r} + (\mathbb{I}^c(r))^{-1}\, p\right), \\ \dot{p} & = & 2J_r\cos^2(\phi)\dot{\phi}\dot{\psi} - \tan(\phi)\dot{\phi}p - 2J_r\cos^2(\phi)\alpha_{12}, \end{array} \tag{6.47}$$

with $\alpha_{12} = \alpha_{12}^1\alpha_{12}^2$. The controllable subspace is then stabilized to the point $(0, \frac{1}{2}\alpha_2\epsilon)$, corresponding to a steering bias. The group and velocity error in body coordinates are

$$\begin{array}{ll} g_{err} & = g_{fdbk}^{-1}\, g_{des}, \\ \xi_{err} & = g_{des}^{-1}\, g'_{des} + \mathrm{Ad}_{g_{fdbk}}\, g'_{fdbk}, \end{array} \tag{6.48}$$

where g_{fdbk} is the feedback of the group variables, and g_{des} is the desired trajectory. The error function is

$$\begin{array}{l} err^1 = -k_p^1 g_{err}^1 + k_v^1 \xi_{err}^1, \\ err^2 = k_p^2 g_{err}^2 + k_v^2 \xi_{err}^2 + k_p^3 g_{err}^3 + k_v^3 g_{err}^3. \end{array}$$

Noting that the term err^1 corrects velocity and position along the relative equilibria, and the term err^2 attracts trajectories to the relative equilibria curve, we define the parameters α_{12}^i and α_2 according to Theorem 16,

$$\begin{array}{rcl} \alpha_{12}^1 & = & -2\,\mathrm{sign}(err^1)\sqrt{|err^1|}, \\ \alpha_{12}^2 & = & -\frac{1}{2}\sqrt{|err^1|}, \\ \alpha_2 & = & err^2. \end{array} \tag{6.49}$$

At zero configuration error, the feedback terms (α_1, α_2) and the oscillatory signals vanish. Hence, both the original and averaged systems share the same relative equilibria. Therefore, via Corollary 2, the system exponentially stabilizes to the desired trajectory. Recall that state feedback corrections are determined at discrete times corresponding to the basic control oscillation. With $\epsilon = \frac{1}{3}$ this update occurs every $T = 2\pi/3$ seconds. The feedback gains are $k_p^1 = \frac{3}{16}$, $k_v^1 = \frac{3}{8}$, $k_p^2 = \frac{1}{4}$, $k_v^2 = \frac{3}{4}$, $k_p^3 = 1$, and $k_v^3 = \frac{3}{2}$.

Stabilization is demonstrated in Fig. 6.14 where the system parameters are $M = 8\,\mathrm{kg}$, $J_r = 0.167\,\mathrm{kg\,m^2}$, $l = 0.3\,\mathrm{m}$. The initial condition occurs at the origin, with zero initial velocity. The upper figures depict the evolution of the states (x, y, θ) and forward velocity, v, describing the relative equilibria in the body frame. The desired trajectory is the dash-dotted plot, and the actual trajectory is the solid plot.

The bottom two graphs show the rotor and steering angle histories. Fig. 6.15, shows snapshots (every 7 seconds) of the snakeboard's configuration during the first 50π seconds. The grid lines are spaced 5 units apart. At time $t = 100$, there is a discontinuous change in the desired trajectory's bearing and velocity. The oscillations decrease in amplitude as the error approaches zero. There is a small steady state error.

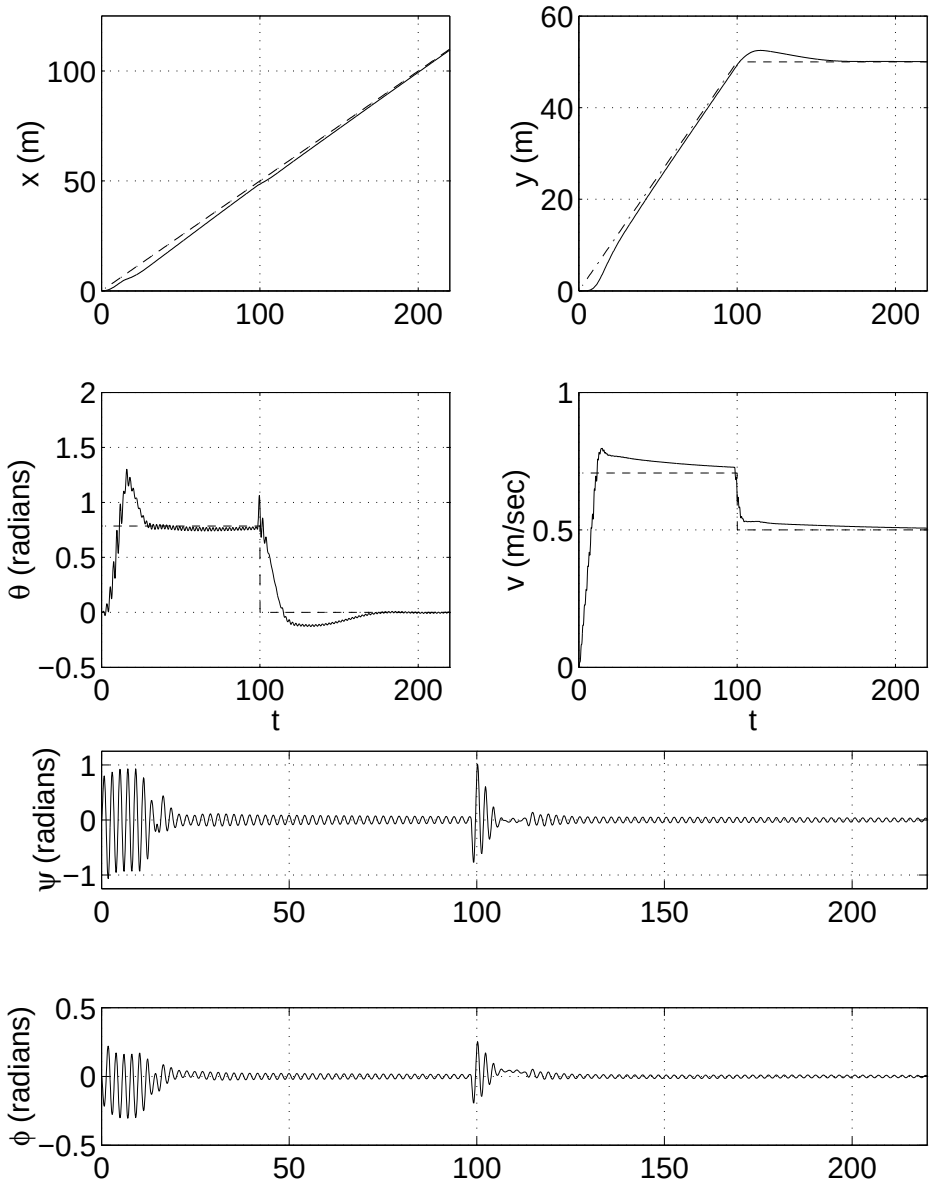


Figure 6.14: Trajectory tracking results for the snakeboard.

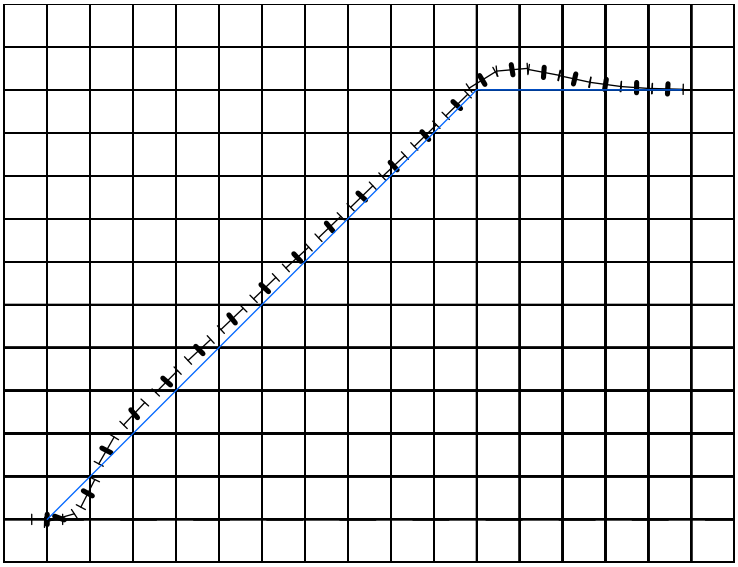


Figure 6.15: Snapshots of trajectory tracking for snakeboard (snakeboard enlarged 6x).

6.5 Carangiform Fish

The model of a carangiform fish here is not the full infinite-dimensional model that incorporates the fluid effects, but is a finite-dimensionl model with a lift vector field that seeks to determine the effect of fluid-body interaction when the caudal fin flaps. The simple fish-like robot discussed in this paper and related earlier works [116, 115] consists of a planar three-link mechanism immersed in water (see Fig. 6.16). The orientation of the peduncle and tail joints are the shape space of the system, and are denoted by $r = [\psi_1, \psi_2]^T \in \mathbb{T}^2$. They are angles measured with respect to the main body reference frame. The fish is restricted to planar motion, meaning that the Lie group will be $SE(2)$. The configuration space is the principal bundle $Q = \mathbb{T}^2 \times SE(2)$.

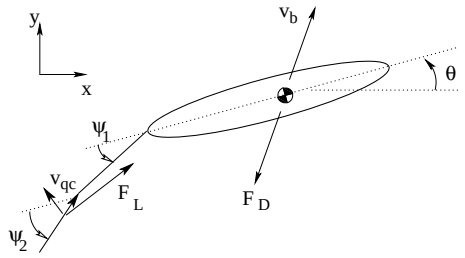


Figure 6.16: Model for carangiform locomotion.

The forces acting on the system are lift on the tail and drag on the body. The comparatively small effects of lift on the body, form drag on the tail, and skin friction are ignored. As discussed in [115], the drag for a translating and rotating plate is taken to be

$$F_D = \frac{1}{2}\rho C_d h \int_{\frac{l}{2}-a}^{\frac{l}{2}+a} \|\xi((a+s)e_1) \times e_1\| \xi((a+s)e_1) ds,$$

and the associated moment to be

$$M_D = \frac{1}{2}\rho C_d h \int_{\frac{l}{2}-a}^{\frac{l}{2}+a} \|\xi((a+s)e_1) \times e_1\| (\xi((a+s)e_1) \times e_1) s ds,$$

where ρ is the fluid density, l is the plate length, C_d is the plate's drag coefficient when its velocity lies in the y direction, h is the plate height, a is the difference in position between the plate's center of mass and center of geometry, and $\xi((a+s)e_1)$ is an infinitesimal generator giving the body-fixed velocity of the plate at the point $a+s$ along the body. The value of s varies from $\frac{l}{2} - a$ to $\frac{l}{2} + a$, and the unit vector e_1 is in the direction of the body-fixed x axis. The lift acting on a flat plate is

$$F_L = \pi \rho A (\xi_{qc} \times e_t) \times \xi_{qc},$$

where ξ_{qc} is the velocity at the plate's quarter chord point as measured in the body frame, e_t is a unit vector pointing along the plate toward its leading edge and A is the plate's area. These equations are a simplification via reduction of those from [115], in recognition of the position invariant nature of the lift and drag forces.

The principal bundle structure of the configuration space can be used to obtain further refinements. The Lagrangian is invariant with respect to the Lie group $SE(2)$. Consequently, one can apply reduction theory [15, 99] to simplify the resulting equations of motion and to expose useful geometric structure in the mechanics. In the absence of constraints, the equations of motion are given by Hamel's equations [15]:

$$\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \xi} = ad^*_\xi \frac{\partial \mathcal{L}}{\partial \xi} + F_\xi, \quad \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{r}} = \frac{\partial \mathcal{L}}{\partial r} + F_r, \tag{6.50}$$

where $\xi = g^{-1}\dot{g}$ is the vehicle's velocity (in body coordinates), $F_\xi = F_\xi(g, r, \xi, \dot{r})$ and $F_r = F_r(r, \dot{r}, t)$ are forces acting on the mechanism in the position and shape directions, and $\mathcal{L}$ is the system Lagrangian. The Lagrangian structure of the system, and the nature of the control inputs means that this system is a

1-homogeneous control system. For more information, please see [116, 160] and references therein.

The control inputs are

$$Y_1 = \frac{\partial}{\partial \psi_1}, \quad \text{and} \quad Y_2 = \frac{\partial}{\partial \psi_2}.$$

The symmetric product $\left\langle Y_1^{\text{lift}} \, : \, Y_2^{\text{lift}} \right\rangle$ can be used to generate forward locomotion, and the Jacobi-Lie bracket between the drift vector field and the peduncle angle control vector field, $\left[Y_1^{\text{lift}}, X_S\right]$ can be used to generate rotation while in motion. The first order symmetric product and Jacobi-Lie bracket for this system evaluated at the reference configuration ($[\psi_1, \psi_2] = [0, 0]$) for typical physical parameters (see [116]) give,

$$\begin{aligned} \left\langle Y_1^{\text{lift}} \, : \, Y_2^{\text{lift}} \right\rangle &\approx (0.15,\ 0,\ 0)^T, \quad \text{and} \\ [Y_1, S + Y_0 - D] &\approx (f(\xi_2, \xi_3),\ 0.27\xi_1,\ -18.70\xi_1)^T, \end{aligned}$$

for the group variables, where f is a function of the lateral and rotational body velocities, ξ_2 and ξ_3, respectively. The symmetric product demonstrates that forward locomotion is theoretically possible. The Lie bracket shows that a turning motion is achievable; it must be mentioned that the lateral translation and the turning do not oppose each other in directing the carangiform fish to reduce the error. Thus, in the plane we have rudimentary control over forward motion and bearing, which can be used for locomotion. Interpreted differently, the above calculations demonstrate that, after averaging, the planar carangiform model is linearly controllable along feasible trajectories (relative equilibria in the plane). Using the average, it is possible to stabilize to trajectories.

The oscillatory control functions are

$$\begin{aligned} v^1(t) &= \beta_{12}^1 \sin(\omega_{12} t) + \frac{1}{2}\alpha_{12}^1 (1 - \cos(\omega_{12} t)), \\ v^2(t) &= \beta_{12}^2 \sin(\omega_{12} t) + \frac{1}{2}\alpha_{12}^1 (1 - \cos(\omega_{12} t)). \end{aligned} \tag{6.51}$$

For second order systems the group error and Lie algebra error must be defined [26],

$$g_{err} = g^{-1} g_{des} \quad \text{and} \quad \xi_{err} = g_{des}^{-1} \dot{g}_{des} - \text{ad}_{g_{err}} \xi,$$

where g_{des} is the desired trajectory and ξ is the carangiform fish's body velocity. The error functions for forward motion and bearing control are

$$\begin{aligned} err^1 &= k_p^1 g_{err}^1 + k_v^1 \xi_{err}^1, \\ err^2 &= \tan^{-1}(g_{err}^1, g_{err}^2), \\ err^3 &= k_p^3 g_{err}^3 + k_v^3 \xi_{err}^3) + k_p^2 err^2 + k_v^2 \xi_{err}^2. \end{aligned} \tag{6.52}$$

The α- and β-parametrizations are

$$\begin{aligned} \alpha_{12}^1 = \text{sign}(err^1)\sqrt{|err^1|}, \quad \alpha_{12}^2 = \sqrt{|err^1|}, \\ \beta_{12}^1 = err^3, \quad \beta_{12}^2 = err^3. \end{aligned} \tag{6.53}$$

The results of tracking a straight line along the x-axis given an initial position error are shown in Fig. 6.17. The carangiform fish model was given an initial position error. Steady state error is within 5% and could be eliminated by including an integral term in the feedback control functions. Actual experimental results can be found in [116]. In Figure 6.18, the fish was commanded to track a trajectory similar to the one given to the snakeboard example from Section 6.4. The closed-loop fish model was able to control to the new

heading and velocity. The control parameters for the simulation were $\omega_{12} = 1$, $\epsilon = \frac{1}{8}$, $k_p^1 = \frac{1}{64}$, $k_v^1 = \frac{1}{16}$, $k_p^2 = \frac{1}{3}$, $k_v^2 = \frac{2}{3}$, $k_p^3 = \frac{1}{3}$, and $k_v^3 = 1$.

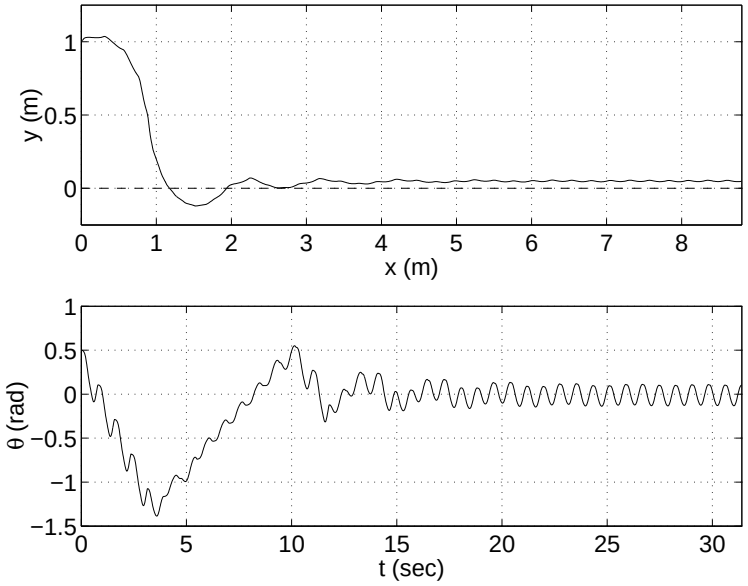


Figure 6.17: Trajectory tracking for carangiform fish.

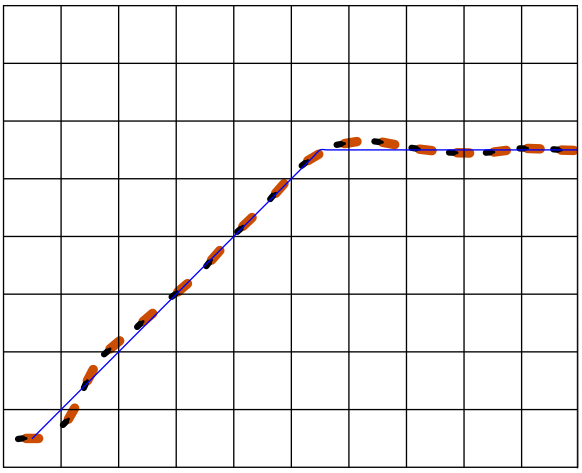


Figure 6.18: Snapshots of trajectory tracking for carangiform fish.

Chapter 7

Conclusion

In this thesis, the foundations underlying several common areas in control and dynamical systems theory have been generalized. Each of the areas studied: averaging theory, controllability theory for mechanical systems, and nonlinear control theory, have been provided with a framework capable of reproducing and extending known results. The most practically important contribution of this work can be found in Chapter 6, where the basic methods of the thesis are synthesized into a paradigm for constructing controllers for nonlinear biomimetic locomotory systems. The α-parametrized controllers correspond to oscillatory shape parameters that are used by biomimetic system to generate directed locomotion; a common theme for biomimetic systems. The fact that the control strategy successfully works for biomimetic control systems that roll, undulate, swim, and walk, has deep implications. Although the thesis contributions are primarily theoretical, the controllers designed by the theory have found practical application for a variety of experimental nonlinear control setups [105, 106, 116, 108], implying that the techniques are viable from a practical engineering perspective.

7.1 Summary of Dissertation

In Chapter 2 the exponential representation of nonlinear time-varying flows was introduced and utilized to derive a generalized averaging theory. It was shown that averaging theory is the synthesis of two distinct theories: (1) nonlinear Floquet theory, and (2) perturbation theory. Nonlinear Floquet theory demonstrates that time-periodic systems may be represented exactly by a periodic mapping and the flow of an autonomous vector field. The perturbation methods provide a way to systematically compute a series expansion approximation when the exact closed-form solution is intractable.

Chapter 3 detailed the problem of controller design for underactuated driftless nonlinear systems. Using the generalized averaging theory, an algorithm was implicitly given for the construction of exponentially stabilizing controllers. In some cases, existing controller design methods may be enhanced by recasting them within the framework of the generalized averaging theory. The controllers result in a set of tunable control parameters some of which are analogous to the feedback gains found in linear control theory.

In order to extend the results of Chapter 3 to the case of systems with drift, a new class of nonlinear control systems was defined. Chapter 4 introduced the notion of a *1-homogeneous control system*, which is an abstraction of simple mechanical systems and most other canonical forms for mechanical systems. It was shown that geometric homogeneity is the unifying concept to many of the related results regarding controllability of mechanical systems.

Chapter 4 then synthesized the controller design method of Chapter 3 with the controllability analysis of Chapter 4 to design stabilizing controllers for 1-homogeneous control systems with drift. The net result is also a constructive method for obtaining stabilizing controllers with tunable control parameters.

Finally, Chapter 6 was devoted to the design of locomotive controllers for a variety biomimetic systems.

The purpose of grouping the biologically-inspired control examples was to emphasize the wide range of dynamical systems controllable using the generalized averaging theory, and also to demonstrate the degree to which biologically-inspired systems utilize fundamentally similar control strategies.

7.2 Future Directions

This dissertation focused on controller synthesis methods and the understanding of dynamical systems exhibiting time-periodic behaviour, both basic contributions to the field of control and dynamical systems. Future work may proceed in two directions: (1) continued investigations into the basic elements of nonlinear dynamics and control, and (2) the synthesis of these concepts with higher-level goal oriented control. Additionally, our understanding of biomimetic and biomechanical control systems may be improved by exploring the connection between these ideas and the mechanics of biologically-inspired locomotive systems.

The Dynamics of Nonlinear Systems. The generalized averaging theory was derived for use in constructing stabilizing controller for underactuated nonlinear systems, however it has potential use within the broader context of dynamical systems theory. Furthermore, there is a large literature dealing with the analysis of time-varying and oscillatory systems via averaging theory, much of which has not been explicitly covered by this dissertation. The examples from Chapter 2 suggest that averaging may be a useful theory in area of physics, especially particle physics.

The Control of Underactuated Nonlinear Systems. Continuous stabilization for underactuated nonlinear control systems was recently proven, consequently many of the examples found in this dissertation do not utilize the continuous controller strategies. The strategy has had success in practical applications [106, 116], and further exploration of this novel control strategy should be pursued. The nonlinear observability example using coupled oscillators is an interesting application that was brought to my attention recently. Given that observability is dual to controllability, it would be interesting to see the application of averaging theory towards observation of nonlinear systems. Since the averaging theory can actually determine the effect of resonance between different oscillatory states, it should be possible to derive a nonlinear identification strategy. The averaged coefficients may be used to estimate coupling, and are what must be measured for system identification.

Robust and Optimal Control. The linear (or homogeneous degree 0) autonomous control model approximating the actual nonlinear time-varying control model may be used to determine stability margins under uncertainty. Morin and Samson [118] have demonstrated robustness of a discretized feedback method for stabilization of driftless systems. The analysis is based on the work of Sussmann and Liu [155, 93], which utilizes averaging methods to obtain a fully controllable representation of the original underactuated system. The discretized feedback method of [118] is one of the control strategies derived using the generalized averaging theory, meaning that the success of [118] is initial validation of this research avenue.

Computation of optimal trajectories for time-varying nonlinear systems is a difficult problem. The difficulty is enhanced when nonholonomic effects are required for locomotion. By utilizing the averaged dynamics it is possible to simplify the optimal control problem to an autonomous, potentially fully-actuated, control problem. Trajectories generated for the autonomous sytem will be feasible trajectories for the actual time-varying, possibly underactuated, system.

Control and Dynamics of Biomimetic Systems. The mathematics derived in this dissertation have practical utility for biomimetic and biomechanical control systems. Many such systems exhibit time-periodic behavior to overcome actuator limitations and nonholonomic constraints. Biological systems often employ

multiple locomotive strategies, or have parametrized gaits whose variations result in a variety of locomotive modes. The averaging theory is capable of recovering oscillatory parameters for modulation to achieve concerted motion. An interesting system to apply averaging methods to would be the housefly. Houseflies are equipped with muscles capable of modulating wing dynamics for directed flight. Perhaps a rigorous analysis of housefly wing kinematics could lead to an α-parametrized model capable of qualitatively reproducing the real flight dynamics.

Secondly, we are trying to connect this work to the stratified configuration space approach for legged locomotion modeling [48, 49]. The success of the kinematic biped suggests that this would be a promising avenue of study. Recent work on the application of Floquet theory to legged locomotion systems by Cham et al. [31] has shown that Floquet analysis can be used to optimally control stride parameters for improved locomotion over unknown terrain.

Bibliography

[1] R. Abraham, J. E. Marsden, and T. Ratiu. *Manifolds, Tensor Analysis, and Applications*. Springer-Verlag, 2 edition, 1988.

[2] A. Agrachev and R. Gamkrelidze. The exponential representation of flows and the chronological calculus. *USSR Sbornik*, 35(6):727–785, 1978.

[3] A. Agrachev and R. Gamkrelidze. Chronological algebras and nonstationary vector fields. *Journal of Soviet Mathematics*, 17:1650–1675, 1979.

[4] A. Agrachev, R. Gamkrelidze, and A. V. Sarychev. Local invariants of smooth control systems. *Acta Applicandae Mathematicae*, 14:1650–1675, 1989.

[5] J. Arango and A. Gómez. Flows and difeomorphisms. *Revista Colombiana de Matemáticas*, 32:13–27, 1998.

[6] V. I. Arnold and B. Khesin. *Topological Methods in Hydrodynamics*, volume 25 of *Applied Mathematical Sciences*. Springer-Verlag, 1998.

[7] J. Baillieul. Stable averaged motions of mechanical systems subject to periodic forcing. *Fields Institute Communications*, 1:1–23, 1993.

[8] J. Baillieul. Energy methods for stabilization of bilinear systems with oscillatory inputs. *International Journal of Robust and Nonlinear Control*, 5:285–301, 1995.

[9] R. E. Bellman, J. Bentsman, and S. M. Meerkov. Vibrational control of nonlinear systems: Vibrational controllability and transient behavior. *IEEE Transaction on Automation and Control*, 31(8):717–724, 1986.

[10] R. E. Bellman, J. Bentsman, and S. M. Meerkov. Vibrational control of nonlinear systems: Vibrational stabilizability. *IEEE Transaction on Automation and Control*, 31(8):710–716, 1986.

[11] J. Bentsman. Vibrational control of a class of nonlinear systems by nonlinear multiplicative vibrations. *IEEE Transaction on Automation and Control*, 32(8):711–716, 1987.

[12] R. M. Bianchini. Controllability along a trajectory: A variational approach. *SIAM Journal of Control and Optimization*, 31(4):900–927, 1993.

[13] R. M. Bianchini and M. Kawski. Needle variations that cannot be summed. *Submitted to SIAM Journal of Control and Optimization*, 2002.

[14] R. M. Bianchini and G. Stefani. Graded approximations and controllability along a trajectory. *SIAM Journal of Control and Optimization*, 28(4):803–924, 1990.

[15] A. M. Bloch, P. S. Krishnaprasad, J. E. Marsden, and R. M. Murray. Nonholonomic mechanical systems with symmetry. *Archives for Rational Mechanics and Analysis*, 136:21–99, 1996.

[16] A. M. Bloch, M. Reyhanoglu, and N. H. McClamroch. Control and stabilization of nonholonomic dynamic systems. *IEEE Transaction on Automation and Control*, 37(11):1746–1757, 1992.

[17] N. M. Bogoliubov and Y. A. Mitropolsky. *Asymptotic Methods in the Theory of Non-Linear Oscillations*. Hindustan Publishing Co., 1961.

[18] W. M. Boothby. *An Introduction to Differentiable Manifolds and Riemmanian Geometry*, volume 120 of *Pure and Applied Mathematics*. Academic Press, second edition, 1986.

[19] R. W. Brocket and L. Dai. Nonholonomic kinematics and the role of elliptice functions in constructive controllability. In Z. X. Li and J. F. Canny, editors, *Progress in Nonholonomic Motion Planning*, pages 1–22. Kluwer, 1992.

[20] R. W. Brockett. Asymptotic stability and feedback stabilization. In *Differential Geometric Control Theory*, pages 181–91. Birkhauser, 1993.

[21] F. Bullo. *Nonlinear Control of Mechanical Systems: A Riemannian Geometry Approach*. Ph.D. thesis, California Institute of Technology, 1999.

[22] F. Bullo. Series expansions for the evolution of mechanical control systems. *SIAM Journal of Control and Optimization*, 40(1):166–190, 2001.

[23] F. Bullo. Averaging and vibrational control of mechanical systems. *SIAM Journal of Control and Optimization*, 41(2):542–62, 2002.

[24] F. Bullo, N. Leonard, and A.D. Lewis. Controllability and motion algorithms for underactuated Lagrangian systems on Lie groups. *IEEE Transaction on Automation and Control*, 45(8):1437–1454, 2000.

[25] F. Bullo and A. D. Lewis. On the homogeneity of the affine connection model for mechanical systems. In *Proceedings of IEEE Conference on Decision and Control*, pages 1260–1265, 2000.

[26] F. Bullo and R. M. Murray. Tracking for fully actuated mechanical systems: A geometric framework. *Automatica*, 35(1):17–34, 1999.

[27] F. Bullo and M. Zefran. On modelling and locomotion of hybrid mechanical systems with impacts. *IEEE Conference on Decision and Control*, 1998.

[28] C. Canudas de Wit and O. J. Sordalen. Exponential stabilization of mobile robots with nonholonomic constraints. *IEEE Transaction on Automation and Control*, 37(11):1791–7, 1992.

[29] H. Cendra, J. Marsden, and T. Ratiu. Geometric mechanics, Lagrangian reduction, and nonholonomic systems. In Enguist B. and Schmid W., editors, *Mathematics Unlimited-2001 and Beyond*, pages 221–273. Springer-Verlag, 2001.

[30] H. Cendra, J. Marsden, and T. Ratiu. Lagrangian reduction by stages. *Memoires of the American Mathematical Society*, 152(722), 2001.

[31] J. G. Cham, J. Karpick, J. E. Clark, and M. R. Cutkosky. Stride period adaptation for a biomimetic running hexapod. In *International Symposium of Robotics Research*, 2001.

[32] K. Chen. Integration of paths, geometric invariants and a generalized Baker-Hausdorff formula. *Annals of Mathematics*, 65(1):163–178, 1957.

[33] J. M. Coron. Global asymptotic stabilization for controllable systems without drift. *Math. Contr., Sign., Sys.*, 5(3):295–312, 1992.

[34] J. Cortés, S. Martínez, and F. Bullo. On nonlinear controllability and series expansions for lagrangian systems with dissipative forces. *IEEE Transaction on Automation and Control*, 48(1):1396–1401, 2002.

[35] J. Cortés, S. Martínez, J. P. Ostrowski, and H. Zhang. Simple mechanical control systems with constraints and symmetry. *SIAM Journal of Control and Optimization*, 41(3):851–874, 2002.

[36] M. Crampin. Tangent bundle geometry for Lagrangian dynamics. *Journal of Physics A*, 16:3755–72, 1983.

[37] P. E. Crouch. Geometric structures in systems theory. *IEE Proc. Part D, Control Theory and Applications*, 128(5), 1981.

[38] M. H. Dickinson. Unsteady mechanisms of force generation in aquatic and aerial locomotion. *American Zoology*, 36:537–554, 1996.

[39] H. S. Dumas and J. A. Ellison. Particle channeling in crystals and the method of averaging. In A. W. Sáenz, W. W. Zachary, and R. Cawley, editors, *Local and Global Methods of Nonlinear Dynamics*, pages 200–230. Springer-Verlag, 1986.

[40] H. S. Dumas, J. A. Ellison, and F. Golse. A mathematical theory of planar particle channeling in crystals. *Physica D*, (2564):1–26, 2000.

[41] O. Egeland, M. Dalsmo, and O. J. Sordalen. Feedback control of a nonholonomic underwater vehicle with a constant desired configuration. *International Journal of Robotics Research*, 15(1):24–35, 1996.

[42] J. A. Ellison, A. W. Saenz, and H. S. Dumas. Improved n^{th}-order averaging for periodic systems. *Journal of Differential Equations*, 84(2):383–403, 1990.

[43] D. B. A. Epstein. The simplicity of certain groups of homeomorphisms. *Compositio Mathematica*, 22(2):165–173, 1970.

[44] N. Fin and G. E. Schweigert. On the group of homeomorphisms of an arc. *Annals of Mathematics*, 62(2):237–253, 1955.

[45] M. Fliess, J. Levine, P. Martin, and P. Rouchon. On differentially flat nonlinear systems. *Comptes Rendues des Séances de l'Academie des Sciences*, 317(10):981–986, 1992.

[46] M. Gallivan, D. G. Goodwin, and R. M. Murray. Modeling and control of thin film morphology with unsteady processing parameters: Problem formulation and initial results. In *Proceedings of IEEE Conference on Decision and Control*, 2001.

[47] B. Goodwine. *Control of Stratified Systems with Robotic Applications*. Ph.D. thesis, California Institute of Technology, 1998.

[48] B. Goodwine and J. W. Burdick. Controllability of kinematic control systems on stratified configuration spaces. *IEEE Transaction on Automation and Control*, 46(3):358–68, 2001.

[49] B. Goodwine and J. W. Burdick. Motion planning for kinematic stratified systems with application to quasi–static legged locomotion and finger gaiting. *IEEE Transaction on Robotics and Automation*, 18(2):209–222, 2002.

[50] J. Guckenheimer and P. Holmes. *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*, volume 42 of *AMS*. Springer Verlag, N.Y., 1990.

[51] J. K. Hale. *Ordinary Differential Equations*, volume 21 of *Pure and Applied Mathematics*. Wiley-Interscience, 1969.

[52] J. K. Hale and S. M. Verduyn-Lunel. Averaging in infinite dimensions. *Journal of Integral Equations and Applications*, 2(4):463–94, 1990.

[53] H. Hermes. Nilpotent and high-order approximations of vector field systems. *SIAM Review*, 33(2):238–264, 1991.

[54] H. Hermes. Large-time local controllability via homogeneous approximation. *SIAM Journal of Control and Optimization*, 34(4):1291–1299, 1995.

[55] S. Hirose. *Biologically Inspired Robots: Snake-like Locomotors and Manipulators*. Oxford University Press, 1993.

[56] A. Isidori. *Nonlinear Control Systems: An Introduction*. Springer Verlag, second edition, 1989.

[57] K. Ito, F. Matsuno, and R. Takahashi. Underactuate crawling robot. *Proceedings of IEEE International Conference on Robotics and Automation*, 2001.

[58] Z. Jingzhong and Y. Lu. Embedding of a homeomorphism in a flow and asymptotic embedding. *Scientia Sinica Series A*, 28(7):683–696, 1985.

[59] M. Kawski. A necessary condition for local controllability. In *Differential Geometry: the Interface Between Pure and Applied Mathematics*, volume 68 of *Contemporary Mathematics*, pages 143–155. American Mathematical Society, Providence, RI, 1987.

[60] M. Kawski. Control variations with an increasing number of switchings. *Bulletin of the American Mathematical Society*, 18(2), 1988.

[61] M. Kawski. Families of dilations and asymptotic stability. In B. Bonnard, B. Bride, J. P. Gauthier, and I. Kupka, editors, *Analysis of Controlled Dynamical Systems*, volume 9 of *Progress in Systems and Control Theory*, pages 285–294. Bikh auser, 1991.

[62] M. Kawski. Geometric homogeneity and applications to stabilization. In A. Krener and D. Mayne, editors, *Proceedings IFAC Symposium on Nonlinear Control Systems Design*, pages 147–152, Tahoe City, CA, 1995.

[63] M. Kawski. Nonlinear control and combinatorics of words. In B. Jakubczyk and W. Respondek, editors, *Geometry of Feedback and Optimal Control*, pages 305–346. Dekker, 1998.

[64] M. Kawski. Controllability via chronological calculus. In *Proceedings of IEEE Conference on Decision and Control*, pages 2920–2925, 1999.

[65] M. Kawski and H. Sussman. Noncommutative power series and formal Lie-algebraic techniques in in nonlinear control theory. In U. Helmke, D. Pratzel-Wolters, and Zerz E., editors, *Operators, Systems, and Linear Algebra*, pages 111–128. Teubner, 1997.

[66] S. D. Kelly. *The Mechanics and Control of Robotic Locomotion with Applications to Aquatic Vehicles*. Ph.D. thesis, California Institute of Technology, 1998.

[67] S. D. Kelly and R. M. Murray. Geometric phases and robotic locomotion. *Journal of Robotic Systems*, 12(6):417–431, 1995.

[68] H. K. Khalil. *Nonlinear Systems*. MacMillan Pub. Co., NY, 1992.

[69] S. Kobayashi and K. Nomizu. *Foundations of Differential Geometry*. John Wiley & Sons, Inc., 1963.

[70] I. Kolář, P. Michor, and J. Slovák. *Natural Operations in Differential Geometry*. Springer-Verlag, Berlin, 1993.

[71] I. Kolmanovsky and N. H. McClamroch. Developments in nonholonomic control problems. *IEEE Control Systems Magazine*, 15(6):20–36, 1995.

[72] I. Kolmanovsky and N. H. McClamroch. Controllability and motion planning for noncatastatic nonholonomic control systems. *Mathematical and Computer Modelling*, 24(1):31–42, 1996.

[73] N. Kopell. Commuting diffeomorphisms. In *Proceedings of the American Mathematical Society Symposium on Pure Math*, volume 14, pages 165–184, 1970.

[74] J. Kurzweil and J. Jarník. Limit processes in ordinary differential equations. *Journal of Applied Mathematics and Physics*, 38, 1987.

[75] J. Kurzweil and J. Jarník. Iterated Lie brackets in limit processes in ordinary differential equations. *Results in Math*, 14, 1988.

[76] G. Lafferriere and H. J. Sussmann. Motion planning for controllable systems without drift. *Proceedings of IEEE International Conference on Robotics and Automation*, pages 1148–1153, 1991.

[77] G. Lafferriere and H. J. Sussmann. A differential geometric approach to motion planning. In Z. X. Li and J. F. Canny, editors, *Progress in Nonholonomic Motion Planning*, pages 235–270. Kluwer, 1992.

[78] J. P. Laumond. Singularities and topological aspects in nonholonomic motion planning. In Z. X. Li and J. F. Canny, editors, *Progress in Nonholonomic Motion Planning*, pages 149–200. Kluwer, 1992.

[79] B. Lehman and S. P. Weibel. Fundamental theorems of averaging for functional differential equations. *Journal of Differential Equations*, 152:160–90, 1999.

[80] N. E. Leonard and P. S. Krishnaprasad. High-order averaging on Lie groups and control of an autonomous underwater vehicle. Technical Report T.R. 93-70, UMD, 1993.

[81] A. Y. T. Leung and Q. C. Zhang. Higher-order normal form and period averaging. *Journal of Sound and Vibration*, 217(5):795–806, 1998.

[82] A. D. Lewis. *Aspects of Geometric Mechanics and Control of Mechanical Systems*. Ph.D. thesis, California Institute of Technology, 1995.

[83] A. D. Lewis. A symmetric product for vector fields and its geometric meaning. Control and Dynamical Systems 96-003, California Institute of Technology, 1996.

[84] A. D. Lewis. Simple mechanical control systems with constraints. *IEEE Transaction on Automation and Control*, 45(8):1420–1436, 2000.

[85] A. D. Lewis and R. M. Murray. Configuration controllability of simple mechanical control systems. *SIAM Journal of Control and Optimization*, 35(3):766–790, 1997.

[86] A. D. Lewis and R. M. Murray. Configuration controllability of simple mechanical control systems. *SIAM Review*, 41(3):555–574, 1999.

[87] A. D. Lewis, J. Ostrowski, R. M. Murray, and J. W. Burdick. Nonholonomic mechanics and locomotion: The snakeboard example. *Proceedings of IEEE International Conference on Robotics and Automation*, 1994.

[88] W. Li, J. Llibre, and X. Zhang. Extension of Floquet's theory to nonlinear periodic differential systems and embedding diffeomorphisms in differential flows. *American Journal of Mathemathics*, 124:107–127, 2002.

[89] R. G. Littlejohn. A guiding center Hamiltonian: A new approach. *Journal of Plasma Physics*, 20(12), 1979.

[90] R. G. Littlejohn. Variational principles of guiding centre motion. *Journal of Plasma Physics*, 29(1), 1983.

[91] R. G. Littlejohn. Geometry and guiding center motion. *Contemporary Mathematics*, 28, 1984.

[92] W. S. Liu. An approximation algorithm for nonholonomic systems. *SIAM Journal of Control and Optimization*, 35(4):1328–1365, 1997.

[93] W. S. Liu. Averaging theorems for highly oscillatory differential equations and iterated Lie brackets. *SIAM Journal of Control and Optimization*, 35(6):1989–2020, 1997.

[94] W. S. Liu and H. J. Sussmann. Continuous dependence of trajectories with respect to the input. In *Proceedings of IEEE Conference on Decision and Control*, Pheonix, AZ, 1999.

[95] W. S. Liu and H. J. Sussmann. Continuous dependence with respect to the input of trajectories of control-affine systems. *SIAM Journal of Control and Optimization*, 37(3):777–803, 1999.

[96] W. Magnus. On the exponential solutions of differential equations for a linear operator. *Comm. Pure Appl. Math.*, VII:649–673, 1954.

[97] H. Maoan. Conditions for a diffeomorphism to be embedded in a c^r flow. *Acta Mathematicae Sinica*, 4(2):111–123, 1988.

[98] J. E. Marsden. *Lectures On Mechanics 2nd Edition*. Springer-Verlag, 1997.

[99] J. E. Marsden and T. S. Ratiu. *Introduction to Mechanics and Symmetry*, volume 17 of *TAM*. Springer-Verlag, 1994.

[100] S. Martínez and J. Cortés. Motion control algorithms for simple mechanical systems with symmetry. *Submitted to Acta Applicandae Mathetmaticae*, 2001.

[101] S. Martínez, J. Cortés, and F. Bullo. On analysis and design of oscillatory control systems. *Submitted to IEEE Transactions Automation and Control*, 2001.

[102] R. Mason. *Fluid Locomotion and Trajectory Planning for Shape-Changing Robots*. Ph.D. thesis, California Institute of Technology, 2003.

[103] R. Mason and J. Burdick. Propulsion and control of deformable bodies in an ideal fluid. *Proceedings of IEEE International Conference on Robotics and Automation*, 1999.

[104] J. N. Mather. Commutators of diffeomorphisms. *Commentarii Mathematici Helvetici*, 49(4):512–528, 1979.

[105] F. Matsuno, K. Ito, and R. Takahashi. Local accessibility and stabilization of an underactuated crawling robot with changing constraints. *IEEE International Conference on Robotics and Automation*, 2001.

[106] K. A. McIsaac and J. P. Ostrowski. Experiments in closed-loop control for an underwater eel-like robot. In *Proceedings of IEEE International Conference on Robotics and Automation*, Washington D.C., 2002.

[107] R. M'Closkey and P. Morin. Time-varying homogeneous feedback: Design tools for the exponential stabilization of systems with drift. *International Journal of Control*, 71(5):837–869, 1998.

[108] R. T. M'Closkey. *Exponential Stabiliation of Driftless Nonlinear Control Systems*. Ph.D. thesis, California Institute of Technology, 1994.

[109] R. T. M'Closkey and R. M. Murray. Exponential stabilization of driftless nonlinear control systems using homogeneous feedback. *IEEE Transaction on Automation and Control*, 42(5):614–28, 1997.

[110] J. Melody, T. Basar, and F. Bullo. On nonlinear controllability of homogeneous systems linear in control. *IEEE Transaction on Automation and Control*, 48(1):139–143, 2003.

[111] J. T. Montgomery. Existence and stability of periodic motion under higher order averaging. *Journal of Differential Equations*, 64:67–78, 1986.

[112] K. A. Morgansen. Controllability and trajectory tracking for classes of cascade-form second order nonholonomic systems. In *Proceedings of IEEE Conference on Decision and Control*, Orlando, FL, 2001.

[113] K. A. Morgansen and R. W. Brocket. Nonholonomic control based on approximate inversion. *Proceedings of the American Control Conference*, pages 3515–3519, 1999.

[114] K. A. Morgansen and R. W. Brocket. Trajectory tracking for driftless homogeneous nonholonomic systems. *preprint*, 2002.

[115] K. A. Morgansen, V Duindam, R. J. Mason, J. W. Burdick, and R. M. Murray. Nonlinear control methods for planar carangiform robot fish locomotion. *Proceedings of IEEE International Conference on Robotics and Automation*, 2001.

[116] K. A. Morgansen, P. A. Vela, and J. W. Burdick. Trajectory stabilization for a planar carangiform robot fish. In *Proceedings of IEEE International Conference on Robotics and Automation*, Washinton D.C., 2002.

[117] P. Morin, J.-B. Pomet, and C. Samson. Design of homogeneous time-varying stabilizing control laws for driftless controllable systems via oscillatory approximation of Lie brackets in closed loop. *SIAM Journal of Control and Optimization*, 38(1):22–49, 1999.

[118] P. Morin and C. Samson. *Robust Point-Stabilization of Nonlinear Affine Control Systems*, volume 216 of *Lecture Notes in Control and Information Sciences*, pages 215–237. Springer-Verlag, 1999.

[119] P. Morin and C. Samson. A characterization of the Lie algebra rank condition by transverse periodic functions. *SIAM Journal of Control and Optimization*, 40(4):1227–1249, 2001.

[120] P. Morin and C. Samson. Practical stabilization of driftless homogeneous systems based on the use of tranverse periodic functions. Technical Report 4184, INRIA, 2001.

[121] R. M. Murray. Nonlinear control of mechanical systems: A Lagrangian perspective. *Annual Reviews in Control*, 21:31–45, 1997.

[122] R. M. Murray and S. S. Sastry. Nonholonommic motion planning: Steering using sinusoids. *IEEE Transaction on Automation and Control*, 38(5):700–716, 1993.

[123] P. J. Olver. *Applications of Lie Groups to Differential Equations*. Number 107 in Graduate Texts in Mathematics. Springer-Verlag, second edition, 1993.

[124] S. M. Omohundro. *Geometric Perturbation Theory in Physics*. World Scientific, 1986.

[125] J. Ostroswki and J. Burdick. Controllability tests for mechanical systems with constraints and symmetries. *Journal of Applied Mathematics and Computer Science*, 7(2):101–127, 1997.

[126] J. Ostrowski. *The Mechanics and Control of Undulatory Robotic Locomotion*. Ph.D. thesis, California Institute of Technology, 1996.

[127] J. Ostrowski and J.W. Burdick. The mechanics and control of undulatory locomotion. *International Journal of Robotics Research*, 17(7):683–701, 1998.

[128] J. Palis. On Morse-Smale dynamical systems. *Topology*, 8:385–405, 1969.

[129] J. Palis. Vector fields generate few diffeomorphisms. *Bulletin of the American Mathematical Society*, 80(3):503–505, 1974.

[130] J. Palis and J. C. Yoccoz. Rigidity of centralizers of diffeomorphisms. *Annales Scientifiques de l'École Normale Supérieure*, 22(4):81–98, 1989.

[131] L. A. Pars. *A Treatise on Analytical Dynamics*. John Wiley & Sons, Inc., 1965.

[132] L. M. Perko. Higher order averaging and related methods for perturbed periodic and quasi-periodic systems. *SIAM Journal of Applied Mathematics*, 17(4):469–724, 1968.

[133] J. Peuteman and D. Aeyels. Averaging results and the study of uniform asymptotic stability of homogeneous differential equations that are not fast time-varying. *SIAM Journal of Control and Optimization*, 37(4):997–1010, 1999.

[134] J. Peuteman and D. Aeyels. Exponential stability of nonlinear time-varying differential equations and partial averaging. *Math. Contr., Sign., Sys.*, 15(1):42–70, 2002.

[135] J. Peuteman and D. Aeyels. Exponential stability of slowly time-varying nonlinear systems. *Math. Contr., Sign., Sys.*, 15(3):202–228, 2002.

[136] J. B. Pomet. Explicit design of time-vaying stabilization control laws for a class of controllable systems with drift. *Systems and Control Letters*, 18(2):147–158, 1992.

[137] J. E. Radford and J. W. Burdick. Local motion planning for nonholonomic control systems evolving on principal bundles. *Proceedings of IEEE International Conference on Robotics and Automation*, 1999.

[138] M. H. Raibert. *Legged Robots that Balance*. MIT Press, Cambridge, 1986.

[139] M. Reyhanoglu, S. Cho, and N. H. McClamroch. Discontinuous feedback control of a special class of underactuated mechanical systems. *International Journal of Robust and Nonlinear Control*, 10(4):265–281, 2000.

[140] M. Reyhanoglu, N. H. McClamroch, and A. M. Bloch. Motion planning for nonholonomic dynamic systems. In Z. X. Li and J. F. Canny, editors, *Progress in Nonholonomic Motion Planning*, pages 201–234. Kluwer, 1992.

[141] M. Reyhanoglu, A. van der Schaft, N. H. McClamroch, and I. Kolmanovsky. Dynamics and control of a class of underactuated mechanical systems. *IEEE Transaction on Automation and Control*, 44(9), 1999.

[142] L. Rosier. Homogeneous Lyapunov function for homogeneous continuous vector field. *Systems and Control Letters*, 19(6):467–473, 1992.

[143] A. W. Sáenz. Higher-order averaging for nonperiodic systems. *Journal of Mathematical Physics*, 32(10):2679–94, 1991.

[144] A. V. Sarychev. Lie- and chronologico-algebraic tools for studying stability of time-varying systems. *Systems and Control Letters*, 43(1):59–76, 2001.

[145] A. V. Sarychev. *Nonlinear Control in the Year 2000, vol. 2*. Lecture Notes in Control and Information Sciences. Springer-Verlag, 2001.

[146] S. S. Sastry. *Nonlinear Systems: Analysis, Stability, and Control*. Springer, 1999.

[147] S. Smale. Differential dynamical systems. *Bulletin of the American Mathematical Society*, pages 747–817, 1967.

[148] E. D. Sontag and H. J. Sussmann. Time-optimal control of manipulators. *Proceedings of IEEE International Conference on Robotics and Automation*, pages 1646–1652, 1986.

[149] O. J. Sordalen and O. Egeland. Exponential stabilization of nonholonomic chained systems. *IEEE Transaction on Automation and Control*, 40(1):35–49, 1995.

[150] N. Steenrod. *The Topology of Fibre Bundles*. Princeton University Press, 1951.

[151] H. Struemper and P. S. Krishnaprasad. Approximate inversion and feedback stabilization for systems on matrix Lie groups. In *Proceedings of the American Control Conference*, Albuquerque, NM, 1997.

[152] H. J. Sussmann. A general theorem on local controllability. *SIAM Journal of Control and Optimization*, 25(1), 1987.

[153] H. J. Sussmann and W. S. Liu. Limiting behavior of trajectories for highly oscillatory controls. Technical Report 02, Rutgers SYCON, 1991.

[154] H. J. Sussmann and W. S. Liu. Limits of highly oscillatory controls and the approximation of general paths by admissible trajectories. In *Proceedings of IEEE Conference on Decision and Control*, pages 437–442, 1991.

[155] H. J. Sussmann and W. S. Liu. Lie bracket extensions and averaging: The single bracket case. In Z. X. Li and J. F. Canny, editors, *Progress in Nonholonomic Motion Planning*, pages 109–148. Kluwer, 1992.

[156] T. Suzuki and Y. Nakamura. Nonlinear control of a nonholonomic free joint manipulator with the averaging method. In *Proceedings of IEEE Conference on Decision and Control*, pages 1694–1699, 1996.

[157] A. R. Teel, R. M. Murray, and G. C. Walsh. Non-holonomic control systems: From steering to stabilization with sinusoids. *International Journal of Control*, 62(4):849–870, 1995.

[158] W. P. Thurston. Foliations and groups of diffeomorphisms. *Bulletin of the American Mathematical Society*, 80(2):304–307, 1974.

[159] P. A. Vela, K. A. Morgansen, and J. W. Burdick. Second-order averaging methods for oscillatory control of underactuated mechanical systems. In *Proceedings of the American Control Conference*, 2002.

[160] P. A. Vela, K. A. Morgansen, and J. W. Burdick. Underwater locomotion from oscillatory shape deformations. In *Proceedings of IEEE Conference on Decision and Control*, Las Vegas, NV, 2002.

[161] F. Verhulst and J. A. Sanders. *Averaging Methods in Nonlinear Dynamical Systems*. Springer-Verlag, N. Y., 1985.

[162] S. Wiggins. *Introduction to Applied Nonlinear Dynamical Systems and Chaos*. Number 2 in TAM. Springer, 1996.

[163] T. W. Wu. On theoretical modeling of aquatic and aerial animal locomotion. *Advances in Applied Mechanics*, 38:292–353, 2001.

[164] K. Yagasaki and T. Ichikawa. Higher-order averaging for periodically forced weakly nonlinear systems. *International Journal of Bifurcation and Chaos*, 9(3):519–31, 1999.