

THESIS REPORT

Ph.D.

Averaging and Motion Control of Systems on Lie Groups

by N.E. Leonard

Advisor: P.S. Krishnaprasad

Ph.D. 94-9



*Sponsored by
the National Science Foundation
Engineering Research Center Program,
the University of Maryland,
Harvard University,
and Industry*

Report Documentation Page				Form Approved OMB No. 0704-0188	
Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.					
1. REPORT DATE 1994		2. REPORT TYPE		3. DATES COVERED 00-00-1994 to 00-00-1994	
4. TITLE AND SUBTITLE Averaging and Motion Control of Systems on Lie Groups				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S)				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) University of Maryland, The Graduate School, 2123 Lee Building, College Park, MD, 20742				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)				10. SPONSOR/MONITOR'S ACRONYM(S)	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution unlimited					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT see report					
15. SUBJECT TERMS					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES 206	19a. NAME OF RESPONSIBLE PERSON
a. REPORT unclassified	b. ABSTRACT unclassified	c. THIS PAGE unclassified			

Abstract

Title of Dissertation: **Averaging and Motion Control
of Systems on Lie Groups**

Naomi Ehrich Leonard, Doctor of Philosophy, 1994

Dissertation directed by: Professor P.S. Krishnaprasad
Department of Electrical Engineering

In this dissertation, we study motion control problems in the framework of systems on finite-dimensional Lie groups. Nonholonomic motion control problems are challenging because nonlinear controllability theory does not provide an explicit procedure for constructing controls and linearization techniques, typically effective for nonlinear system analysis, fail to be useful. Our approach, distinguished from previous motion control research, is to exploit the Lie group framework since it provides a natural and mathematically rich setting for studying nonholonomic systems. In particular, we use the framework to develop explicit, structured formulas that describe system behavior and from these formulas we derive a systematic way of synthesizing controls to achieve desired motion.

As our main tool we derive averaging theory for left-invariant systems on finite-dimensional Lie groups. This theory provides basis-independent formulas

which approximate system behavior on the Lie group to arbitrarily high order in ϵ given small (ϵ) amplitude, periodically time-varying control inputs. We interpret the average formulas geometrically and exploit this interpretation to prove a constructive controllability theorem for the average system. The proof of this theorem provides a constructive control synthesis methodology for drift-free systems which we use to derive algorithms which synthesize sinusoidal open-loop controls. We apply the algorithms to several under-actuated mechanical control problems including problems in spacecraft attitude control, unicycle motion control and autonomous underwater vehicle control. We illustrate the effectiveness of the synthesized controls by simulation and experimentation. We show further that as a consequence of the geometry inherited from the average formulas, our algorithms can be used to produce motion controls that adapt to changes in control authority such as loss of an actuator.

We also apply our theory to synthesize controls for bilinear control systems on $\mathbb{R}^n$ possibly with drift. Our approach is to control the system state by controlling the state transition matrix which evolves on a matrix Lie group. We design and demonstrate a controller for an example system with drift, a simple switched electrical network.

Averaging and Motion Control of Systems on Lie Groups

by

Naomi Ehrich Leonard

Dissertation submitted to the Faculty of the Graduate School
of The University of Maryland in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
1994

Advisory Committee:

Professor P.S. Krishnaprasad, Chairman/Advisor
Professor Eyad H. Abed
Professor Steven I. Marcus
Associate Professor W.P. Dayawansa
Professor Scott A. Wolpert
Professor David L. Elliott

© Copyright by
Naomi Ehrich Leonard
1994

Acknowledgements

It is with great pleasure that I acknowledge my advisor Professor P.S. Krishnaprasad and express my gratitude for his guidance, instruction, encouragement and patience. His enthusiasm for the study of science and engineering has been an inspiration and has proven to be infectious.

I would like to thank Professor W.P. Dayawansa for his critical comments and help and Professors Eyad Abed, W.P. Dayawansa, David Elliott, Steven Marcus and Scott Wolpert for serving on my dissertation committee.

Thanks are also due to fellow and former students in and around the Intelligent Servosystems Laboratory for lots of good conversation and for helping to provide an exciting research environment. In particular, I would like to thank Dimitris Tsakiris, Carole Teolis, Rui Yang, Vikram Manikonda, Ben Bachrach, Reza Shahidi, Narayan Subramanian, Qifeng Wei, Svetlana Vranić-Sowers, Bernie Frankpitt, Reza Ghanadan and Yakup Ózkazanç.

I would also like to acknowledge the members of the Space Systems Laboratory at the University of Maryland. In particular, I thank Rob Cohen for his generous help and cooperation during testing of SCAMP. I am grateful to Professor David Akin for letting me use the SSL facilities and for serving as underwater cameraman during the experiments. I am also grateful to Martin Paredes for his work on the graphics simulation of SCAMP.

During my graduate studies I was supported by fellowships from the University of Maryland Graduate School, the Institute for Systems Research, Litton

Industries and the Zonta International Foundation for which I am grateful. My research was also supported in part by the National Science Foundation's Engineering Research Centers Program: NSFD CDR 8803012, by the AFOSR University Research Initiative Program under grant AFOSR-90-0105, and by the Army Research Office under Smart Structures URI Contract No. DAAL03-92-G-0121.

Finally, I would like to thank my husband, Tim Leonard, for his constant support and for his willingness to learn about my research. I am also deeply indebted to my parents, Joan and Fred Ehrich, who encouraged me from the very beginning.

Table of Contents

<u>Section</u>	<u>Page</u>
List of Figures	vii
1 Introduction	1
2 Preliminaries	11
2.1 Problem Definition	11
2.1.1 Lie Groups and Lie Algebras	12
2.1.2 Control Systems on Lie Groups	16
2.1.3 Problem Statement on Lie Groups	17
2.1.4 Control Systems on Matrix Lie groups	18
2.2 Controllability and Geometry	20
2.2.1 Controllability	20
2.2.2 Geometry	23
2.3 Local Solutions	26
2.3.1 Product of Exponentials	26
2.3.2 Single Exponential	29
3 Examples	36
3.1 Spacecraft Attitude Control	36

3.1.1	Internal Rotors	39
3.1.2	Appended Point Mass Oscillator	41
3.2	Unicycle Motion Control	42
3.3	Autonomous Underwater Vehicle Motion Control	45
3.4	Ball and Plate	48
3.5	Brockett's System	49
3.6	Front-Wheel Drive Car and Other Chained-Form Systems	51
4	Averaging	56
4.1	Averaging with Product of Exponentials	60
4.2	Averaging with Single Exponential	70
4.3	Curvature Form	87
4.3.1	Principal Fiber Bundles and Connections	87
4.3.2	Curvature Form and Area Rule	91
4.4	Stability	98
5	Constructive Controllability and an Algorithm for Control	102
5.1	Constructive Controllability	105
5.2	Algorithms for Control Synthesis	115
5.2.1	Algorithm 0: Depth-Zero Bracket System	118
5.2.2	Algorithm 1: Depth-One Bracket System	119
5.2.3	Algorithm 2: Depth-Two Bracket System	123
5.2.4	Examples	132
5.3	Adaptation to Changes in Control Authority	145
5.4	Underwater Vehicle Experiment	149

6	Application to Bilinear Systems	155
6.1	Switched Electrical Network Problem and Approach	158
6.2	Averaging and Controllability	162
6.3	Controller Design	171
7	Conclusions and Future Research	179
	Bibliography	184

List of Figures

<u>Number</u>	<u>Page</u>
3.1 Spacecraft with Internal Rotors.	40
3.2 Spacecraft with Appended Point Mass Oscillator.	41
3.3 Overhead View of Unicycle.	42
3.4 Autonomous Underwater Vehicle.	45
3.5 Uniform Ball Rolling Between Two Parallel Plates.	48
3.6 Vibratory Actuator on $SO(3)$	50
3.7 Overhead View of Front-Wheel Drive Car.	51
4.1 Periodic Controls.	58
5.1 Control Input Signals for Spacecraft and Unicycle Examples. . . .	134
5.2 Actual (solid lines) and Average (dashed lines) Wei-Norman Pa- rameters for Spacecraft Attitude Control Example.	135
5.3 Response for Unicycle Example.	136
5.4 Control Input Signals for AUV Example with Four Controls. . . .	140
5.5 Response of AUV with Four Controls.	141
5.6 Control Input Signals for AUV Example with Three Controls. . .	145
5.7 Response of AUV with Three Controls.	146
5.8 Control Architecture.	148

5.9	Front View of SCAMP.	149
5.10	Calibration Data for SCAMP.	151
6.1	Network Example	160
6.2	Switching Control	162
6.3	Geometric Interpretation of $Area_{uv}(T')$	167
6.4	Paths in State Space	170
6.5	Control Signal and State Response	176
6.6	State-Space Trajectory	177
6.7	Response of Current Variable x_3	177

Chapter 1

Introduction

Lie groups provide both a natural and a mathematically rich setting for studying a variety of motion control problems. For motion control problems involving rotating and translating bodies, such as space systems, underwater vehicles or mobile robots, the natural appearance of certain Lie groups derives from the fact that these describe the system configuration space or a piece of it. For instance, there is a one-to-one relationship between the matrix Lie group of rotations, $SO(3)$, and all possible orientations of a rigid body. This is not true for local parametrizations such as Euler angles and quaternions. Further, the kinematics of the rotating body can be described by a left (or right) invariant system on $SO(3)$. Left (or right) invariance implies that the kinematic description is independent of how we choose to map orientations to elements in $SO(3)$, i.e., which orientation corresponds to the identity element of $SO(3)$.

Subgroups of $GL(n)$, the general linear Lie group, arise in a natural way in the study of the state transition matrix of a time-varying linear system on $\mathbb{R}^n$. For example, consider a passive electrical network with ideal switches described by a time-varying linear system of state equations. Suppose that energy is conserved

such that the state of the system evolves on an n -dimensional sphere (see Chapter 6 for an example). Then the state transition matrix for the network naturally evolves (“moves”) in the space of n -dimensional rotation matrices, i.e., the matrix Lie group $SO(n)$. This group is the “configuration” space of the state transition matrix because it intrinsically incorporates “configurational” constraints, e.g., that the state transition matrix must be orthogonal.

During the second half of the nineteenth century, Sophus Lie introduced and developed his ideas on continuous transformation groups that leave mathematical systems invariant (later named Lie groups). Lie’s focus was the theory of differential equations [65, 51]. However, since then Lie groups have played an important role in many diverse areas of mathematics and physics including differential geometry, mechanics and particle physics. More recently, Lie groups have found their way into the controls literature.

In his seminal work in the early 1970’s, Roger Brockett put the theory of Lie groups and their associated Lie algebras into the context of nonlinear control theory to express notions such as nonlinear controllability, observability and realization theory for right-invariant systems evolving on matrix Lie groups [11]. Jurdjevic and Sussmann further investigated the controllability properties of these systems on abstract Lie groups [29]. One of the most important insights derived from this work was the recognition that questions about these kinds of systems on Lie groups can be reduced to questions about their associated Lie algebras. Since Lie algebras are vector spaces whereas Lie groups are manifolds, this reduction greatly simplifies the problem.

In this dissertation we study motion control problems in the abstract frame-

work of left (or right) invariant systems on finite-dimensional Lie groups. Our main objectives are

1. To describe the *solutions* of these systems on Lie groups, i.e., to understand the motion of these systems given certain types of input.
2. To derive from the solutions a systematic way of *synthesizing* controls to achieve desired motion.

The Lie group framework is exploited to get coordinate-free descriptions of motion and, consequently, generalized control synthesis algorithms.

The results of Brockett, Jurdjevic and Sussmann provide a simple test to determine for a given motion control problem, expressed as a left (or right) invariant system on a Lie group, whether or not there exists a control law that drives the system from any initial point in the Lie group to any final point in the Lie group. However, the test, which we refer to as the *Lie algebra controllability rank condition*, by itself does not explicitly reveal the control law nor how the system will move. Unlike the linear setting where the controllability Grammian yields constructive controls, here the rank condition does not lead immediately to an explicit procedure for constructing controls. Finding a control law that results in the desired motion is referred to as the *constructive controllability problem*. This problem is considered challenging because linearization techniques, highly effective in the analysis of nonlinear systems, fail to be useful. In fact, the linearized system is generally not controllable.

We approach the constructive controllability problem and our main objectives by considering small-amplitude, periodically time-varying controls and using averaging theory as our main analysis tool. The primary motivation for motion

control based on small-amplitude periodic controls comes from nonlinear systems theory and, in particular, from the theory of systems with nonholonomic constraints. However, the use of small-amplitude periodic controls is also practically motivated. For example, one of the applications of our work is the spacecraft attitude control problem with two internal rotors. We consider a spacecraft for which we can control roll and pitch, but, perhaps due to an actuator failure, we assume no direct control over yaw. The motion control problem is then to show how to orient the spacecraft as desired, e.g., to achieve a net yaw rotation, using small-amplitude periodic changes (wiggles) in roll and pitch velocity. The choice of small-amplitude, low-frequency periodic controls in this application is certainly justified from a practical point of view. Using these small, gentle control inputs, we can avoid both exciting vibrational modes of the system and making large off-course excursions in the vehicle's orientation. Further, because of the nature of these controls, it is possible to consider producing them with micro-actuators, e.g., vibratory actuators. These types of actuators would have the advantage of small size and weight in applications such as the control of small planetary spacecraft.

Control synthesis based on small-amplitude periodic controls can also be applied to the design of micro-actuators and sensors. For example, vibratory actuators, such as piezoelectric actuators, are driven by small-amplitude mechanical oscillations which need to be rectified into linear or rotary motion. Brockett has studied this type of problem [14]. Our results can be applied to interesting types of vibratory-driven sensors or actuators that naturally fit the Lie group framework. We discuss one such possibility, namely an actuator that uses a sphere which rolls between two flat surfaces to convert vibratory motion into rotary

motion.

From the theoretical point of view, it is the observations and results from studies of nonholonomic systems, i.e., systems that satisfy nonholonomic constraints, that provide inspiration for our approach. These studies are relevant to our problem because the mechanical systems that we consider in the Lie group framework satisfy nonholonomic kinematic constraints. Nonholonomic kinematics constraints are nonintegrable velocity constraints, i.e., constraints on velocities that cannot be expressed as configurational constraints.

For example, a unicycle with a wheel that rolls and turns may be subject to a nonholonomic no-slip constraint, i.e., the condition that the wheel does not slip sideways. This constrains the wheel's possible velocities but does not constrain the unicycle's possible positions and orientations. Alternatively, a space system with no spin and no external torques applied will conserve (zero) angular momentum. This conservation law can equivalently be expressed as a nonintegrable constraint on the spacecraft's angular velocity. Yet another example of a nonholonomic system is a low Reynolds number swimmer, e.g., a paramecium which has negligible inertia relative to friction. The swimmer is constrained by a no-slip condition between its body and the fluid, i.e., the tangential velocity between body and fluid is zero. These conditions together with the low Reynolds number limit yield nonintegrable kinematic constraints on the swimmer.

The motion of systems in free fall or in space which conserve angular momentum has been studied in the context of both animate and inanimate bodies. Kane and Scher derived dynamic models and explanations for the “falling cat” problem, i.e., the problem of how a cat when dropped upside down reorients itself while

conserving angular momentum [30]. Frohlich investigated the dynamics of how springboard divers, gymnasts and trampolinists rotate and reorient themselves in free fall [22]. Using the theory of connections in principal bundles, Krishnaprasad, Yang and Dayawansa studied kinematic drift (or geometric phase) effects and related optimal control problems in space systems subject to vibrations or articulations [35, 36, 79]. Montgomery, also using the setting of a principal bundle with connection, investigated the optimal control problem, referred to as the isoholonomic problem, of finding the lowest-energy cyclic “shape” change yielding a desired phase shift [56]. His results were applied to several examples including the falling cat problem. Sreenath studied controllability and control synthesis for reorientation of planar multibody systems in space using reduction [72].

Taylor [74, 75] and Lighthill [52] both made early contributions to the study of how microscopic organisms use cyclic (periodic) shape changes to swim. Purcell investigated these low Reynolds number swimmers further, exploring possible efficient swimming and feeding strategies [63]. Using the familiar (to physicists) formalism of gauge theory, Wilczek and Shapere derived a unifying geometric framework for studying how deformable bodies move. Using this framework they were able to study problems including the falling cat, the diver as well as the low Reynolds number swimmer [68, 69].

In the robotics literature researchers have recently become interested in non-holonomic motion planning for wheeled, mobile robots with wheels subject to the no-slip constraint. The problem is finding feasible trajectories within the set of admissible trajectories. Early work was devoted to path-planning for low-dimensional wheeled mobile robots beginning with Laumond [39, 40] and Bar-

raquand and Latombe [7].

More recently, the constructive controllability problem for nonholonomic systems of a more general form has been studied. Nonholonomic systems can be described in state-space form, at least locally, as drift-free systems on $\mathbb{R}^n$ of the form

$$\dot{x} = \sum_{i=1}^m F_i(x)u_i, \quad x \in \mathbb{R}^n, \quad u_i \in \mathbb{R}, \quad n > m. \quad (1.1)$$

System (1.1) is called drift-free because $\dot{x} = 0$ if $u = (u_1, \dots, u_m) = 0$, i.e., the system does not “drift” when there is no control applied.

Brockett investigated optimal controls to steer a prototype (nilpotent) class of such systems, providing motivation for later work [13]. Murray and Sastry used (suboptimal) sinusoidal controls to steer systems which can be expressed in a special, nilpotent form called “chained” form [58, 60]. They showed when and how systems of the form (1.1) could be transformed into chained form [59]. Lafferiere and Sussmann derived a general strategy for steering controllable systems of the form (1.1) based on an extended system to (1.1) and formal calculations on the free nilpotent Lie algebra generated by the system vector fields [38, 37]. For nilpotent or feedback nilpotentizable systems the generated controls steer the system exactly as desired. Liu and Sussmann developed highly oscillatory control solutions to approximate general paths with feasible trajectories of systems of the form (1.1) [73, 53]. In particular, using averaging theory they show that the approximation converges to the desired path in the limit as the oscillation frequency (and amplitude) goes to infinity. Gurvits and Li also considered highly oscillatory controls to solve the motion planning problem for system (1.1) [24, 25]. Their solution uses averaging theory, but their algorithm is recursive. Brockett

and Dai studied the role of elliptic functions in optimal paths for nonholonomic systems [16].

In Chapter 2 of this dissertation, we formally state the motion control problem in the framework of systems on Lie groups. We discuss left and right-invariant systems with and without drift on matrix and abstract Lie groups. We review definitions and basic theorems from nonlinear controllability theory and define geometric objects that play a key role in the averaging theory of Chapter 4. We also describe two local representations of the solution to the Lie group systems, one a product of exponentials and the other an exponential of a sum, that form the basis for our extension of classical averaging theory to systems on Lie groups.

In Chapter 3 we derive examples that are featured throughout the dissertation including the spacecraft attitude control problem, the unicycle motion planning problem and the motion control problem for an autonomous underwater vehicle (AUV).

Chapter 4 introduces averaging theory for systems on Lie groups. The purpose is to derive an approximation to the Lie group system solution to arbitrarily high order in the small parameter ϵ which represents the small amplitude of the oscillatory control inputs. Although there is in general no analytic solution that describes the motion for a given control input, we show that there is an analytic *average approximation* to the solution that captures the essence of the motion and helps us meet Objective 1 above. The average solution provides a basis-independent formula for the system's motion with an intrinsic geometric interpretation. For instance, the second-order average formula can be interpreted as an “area rule” and higher-order average formulas as “area-moment rules”.

We emphasize the coordinate-free aspect of the solution by illustrating how the important secular term in the area rule derives from the curvature form of a certain principal fiber bundle with connection (see Section 4.3.1 for a definition of principal fiber bundles and connections). We also discuss some implications for system stability that are associated with the averaging theory.

In Chapter 5 we address Objective 2, control synthesis, by relating average formulas of Chapter 4 to system controllability. Specifically, in order to drive a given system's motion as desired, we propose synthesizing controls that drive the *average* motion of the system as desired, knowing from our averaging theory results that the true motion will stay close to the average motion. A critical ingredient in this strategy is determining the minimum order of the average solution that can be controlled as desired, i.e., that captures the controllability of the system. We prove a constructive controllability theorem that shows that this order is one more than the number of Lie bracket iterations needed for the system to satisfy the Lie algebra controllability rank condition. Additionally, the proof of this theorem provides a methodology for constructing controls to direct the average motion as desired.

In the second part of Chapter 5 we use the constructive methodology to derive algorithms that systematically synthesize open-loop, sinusoidal controls for point-to-point system maneuvers of systems which require up to depth-two Lie brackets to satisfy the Lie algebra controllability rank condition. The algorithms are illustrated for the examples of Chapter 3 with simulations for verification. Based on the capabilities of this open-loop control synthesizer, we propose a control architecture for motion control systems which uses a strategic open-loop planner coupled with intermittent feedback plus a second level of feedback which provides

adaptation to changes in control authority such as an actuator failure. We also describe an experiment that was run to test the algorithms on an underwater vehicle in a neutral buoyancy tank at the University of Maryland.

In Chapter 6 we turn to the problem of open-loop control synthesis for bilinear systems on $\mathfrak{R}^n$ with state transition matrices that evolve on matrix Lie groups. We apply the theory of averaging and motion control on Lie groups to the problem of controlling energy transfers between dynamic storage elements in switched electrical networks. In this case the system is a right-invariant system on a matrix Lie group with drift. We show for a system on $SO(3)$ that our techniques can be used to accomplish prescribed energy transfers with a finite number of switchings, and, thus, chattering problems associated with feedback strategies such as sliding mode control can be avoided.

In Chapter 7 we describe our conclusions and suggestions for future work.

The major contributions of this dissertation include the theory of averaging for systems on finite-dimensional Lie groups and the derivation of a systematic method, low in computational burden, for synthesizing open-loop, small-amplitude periodic controls for these types of systems. These contributions are distinguished from past results in nonholonomic motion planning most clearly by the use of averaging in the abstract framework of motion control on Lie groups which yields intuitive, geometric, basis-independent solutions to motion control problems. In essence, we have systematized the means to solve motion control problems that can be represented in our “normal” form, i.e., as left (or right) invariant systems on finite-dimensional Lie groups. For our previous related publications see Leonard and Krishnaprasad [42, 43, 44, 45, 46, 47, 48, 49, 50].

Chapter 2

Preliminaries

In this chapter we present the formal framework of motion control on Lie groups. We begin with a definition of a control system on an abstract finite-dimensional Lie group and specialize to systems on matrix Lie groups. After reviewing definitions and theorems concerning nonlinear controllability, we define geometric objects significant for the motion description formulas in Chapter 4 and the control algorithm of Chapter 5. The chapter ends with a discussion of two important local solutions to systems on Lie groups. These local solutions provide the means to extend classical averaging theory to systems on Lie groups.

2.1 Problem Definition

In this section we define Lie groups and Lie algebras and state the motion control problem in the framework of systems on Lie groups. Useful references are [1, 12, 19, 26, 27, 76].

2.1.1 Lie Groups and Lie Algebras

A *Lie group* G is a differentiable manifold which is also a group, such that multiplication and inversion of elements are smooth maps. Let $h \in G$, and define left translation $L_h : G \rightarrow G$ and right translation $R_h : G \rightarrow G$ by h as

$$L_h(g) = hg, \quad R_h(g) = gh, \quad g \in G,$$

respectively. A vector field Y on G is *left-invariant* if

$$(L_h)_*Y = Y, \quad \forall h \in G,$$

where $(L_h)_*$ is the differential of L_h . Equivalently, $T_g L_h Y(g) = Y(hg)$, $\forall g, h \in G$, where $T_g L_h$ is the linearization of L_h at g . Similarly, Y is *right-invariant* if

$$(R_h)_*Y = Y, \quad \forall h \in G,$$

i.e., $T_g R_h Y(g) = Y(gh)$, $\forall g, h \in G$.

A *Lie algebra* $\mathcal{G}$ over $\mathfrak{R}$ is a real vector space endowed with a binary operation $[\cdot, \cdot] : \mathcal{G} \times \mathcal{G} \rightarrow \mathcal{G}$, called the *Lie bracket*, which satisfies (i)-(iii) below $\forall \xi, \eta, \zeta \in \mathcal{G}$, $\forall \alpha, \beta \in \mathfrak{R}$:

- (i) (Bilinearity) $[\alpha\xi + \beta\eta, \zeta] = \alpha[\xi, \zeta] + \beta[\eta, \zeta],$
- (ii) (Skew Symmetry) $[\xi, \eta] = -[\eta, \xi],$
- (iii) (Jacobi Identity) $[\xi, [\eta, \zeta]] + [\eta, [\zeta, \xi]] + [\zeta, [\xi, \eta]] = 0.$

To each finite-dimensional Lie group G we associate a finite-dimensional Lie algebra $\mathcal{G}$ which can be identified with $T_e G$, the tangent space (i.e., the linear approximation) of G at the identity $e \in G$. G and $\mathcal{G}$ have the same dimension. Further, $T_e G$ is isomorphic to the Lie algebra of left-invariant vector fields on G

(see e.g., [76]). In particular, a vector field Y on G is left-invariant if and only if $Y(g) = T_e L_g \cdot \xi$, $\forall g \in G$ for some fixed $\xi \in \mathcal{G}$. To see this note that for $h \in G$

$$Y(hg) = T_e L_{hg} \cdot \xi = T_e(L_h \circ L_g) \cdot \xi = (T_g L_h)(T_e L_g) \cdot \xi = T_g L_h Y(g).$$

Similarly, Y is right-invariant if $Y(g) = T_e R_g \cdot \xi$, $\forall g \in G$ for some fixed $\xi \in \mathcal{G}$.

For Y_i and Y_j smooth vector fields on a smooth manifold M the Lie bracket $[Y_i, Y_j]$ is defined by

$$[Y_i, Y_j](f) = X(Y(f)) - Y(X(f))$$

where f is a smooth real-valued function on M . Given $\xi_i, \xi_j \in \mathcal{G}$, let $Y_i = (L_g)_* \xi_i$ and $Y_j = (L_g)_* \xi_j$. Then Y_i and Y_j are smooth left-invariant vector fields on G . The Lie bracket on $\mathcal{G} = T_e G$ is defined in terms of the Lie bracket on the associated left-invariant vector fields as

$$[Y_i, Y_j] = [(L_g)_* \xi_i, (L_g)_* \xi_j] = (L_g)_* [\xi_i, \xi_j], \quad \xi_i, \xi_j \in \mathcal{G},$$

where the first two brackets are brackets defined on the vector fields and the bracket on the right is the Lie bracket defined for $\mathcal{G}$.

Given $\xi \in \mathcal{G}$, define $X_\xi = T_e L_g \cdot \xi$, and let the integral curve of X_ξ passing through e at $t = 0$ be denoted by

$$\exp_\xi : \mathbb{R} \rightarrow G : \quad t \mapsto \exp_\xi(t) = \exp(\xi t).$$

Then $\exp_\xi$ is a *one-parameter subgroup* of G , i.e., $\exp_\xi(s + t) = \exp_\xi(s) \exp_\xi(t)$ and $\exp_\xi^{-1}(t) = \exp_\xi(-t)$. We define the *exponential map* of $\mathcal{G}$ into G by

$$\exp : \mathcal{G} \rightarrow G : \quad \xi \mapsto \exp(\xi) \triangleq \exp_\xi(1) \tag{2.1}$$

Then $\exp$ is a diffeomorphism of a neighborhood of $0 \in \mathcal{G}$ onto a neighborhood of $e \in G$ [76].

We define a (*left*) *action* of a Lie group G on a smooth manifold Q by the smooth mapping

$$\Phi : G \times Q \rightarrow Q : (g, q) \mapsto \Phi(g, q) = \Phi_g(q) \triangleq g \cdot q, \quad (2.2)$$

$$\Phi(e, q) = q, \quad \Phi(g, \Phi(h, q)) = \Phi(gh, q), \quad \forall g, h \in G, \forall q \in Q.$$

For example, we define the *adjoint action* of G on $\mathcal{G}$ by

$$\text{Ad} : G \times \mathcal{G} \rightarrow \mathcal{G} : (g, \xi) \mapsto \text{Ad}_g \xi \triangleq T_e(R_{g^{-1}} L_g) \cdot \xi. \quad (2.3)$$

For a group action (2.2), the *infinitesimal generator* of the action corresponding to $\xi \in \mathcal{G}$ is defined by

$$\xi_Q(q) = \left. \frac{d}{dt} \right|_{t=0} \Phi(\exp(\xi t), q), \quad \forall q \in Q. \quad (2.4)$$

For the adjoint action of (2.3),

$$\xi_{\mathcal{G}}(\eta) = \text{ad}_{\xi} \eta \triangleq [\xi, \eta], \quad \forall \eta \in \mathcal{G}.$$

We define the iterated ad operator by

$$\text{ad}_{\xi}^i \eta = \text{ad}_{\xi}(\text{ad}_{\xi}^{i-1} \eta), \quad \text{ad}_{\xi}^0 \eta = \eta, \quad \xi, \eta \in \mathcal{G}.$$

The significance of associating the Lie algebra $\mathcal{G}$ with the Lie group G comes about when G is connected (as a topological manifold). In this case, many *group-theoretic* properties of G can be expressed as *algebraic* properties of $\mathcal{G}$. For example, we define C_G , the *center of G* , and $C_{\mathcal{G}}$, the *center of $\mathcal{G}$* , as

$$C_G = \{h \mid gh = hg, \forall g \in G\}, \quad C_{\mathcal{G}} = \{\eta \mid [\xi, \eta] = 0, \forall \xi \in \mathcal{G}\}.$$

Then we say that the group G is *abelian* if $C_G = G$, and the Lie algebra is *abelian* if $C_{\mathcal{G}} = \mathcal{G}$. One can show that G is an abelian group only if $\mathcal{G}$ is an abelian Lie algebra. Thus, checking to see if G is abelian is easy since checking that $\mathcal{G}$ is abelian amounts to showing that $[\xi, \eta] = 0$ holds for $\xi, \eta \in \{\xi_1, \dots, \xi_n\}$, a basis for $\mathcal{G}$.

We define other important properties of Lie algebras and Lie groups as follows. Let $\mathcal{G}, \mathcal{H}$ be Lie algebras. Denote

$$[\mathcal{G}, \mathcal{H}] = \{[\xi, \eta] \mid \xi \in \mathcal{G}, \eta \in \mathcal{H}\}.$$

A subspace $\mathcal{H} \subset \mathcal{G}$ is called an *ideal* if $[\mathcal{G}, \mathcal{H}] \subseteq \mathcal{H}$. We define the *derived series* of $\mathcal{G}$ to be

$$\mathcal{G} \supseteq \mathcal{G}' = [\mathcal{G}, \mathcal{G}] \supseteq \mathcal{G}'' = [\mathcal{G}', \mathcal{G}'] \supseteq \dots \supseteq \mathcal{G}^{(k)} = [\mathcal{G}^{(k-1)}, \mathcal{G}^{(k-1)}] \supseteq \dots$$

and the *lower central series* of $\mathcal{G}$ to be

$$\mathcal{G} \supseteq \mathcal{G}^2 = \mathcal{G}' = [\mathcal{G}, \mathcal{G}] \supseteq \mathcal{G}^3 = [\mathcal{G}^2, \mathcal{G}] \supseteq \dots \supseteq \mathcal{G}^k = [\mathcal{G}^{k-1}, \mathcal{G}] \supseteq \dots$$

Then we say that $\mathcal{G}$ is *solvable* if $\mathcal{G}^{(k)} = 0$ for some positive integer k . We say that $\mathcal{G}$ is *nilpotent* if $\mathcal{G}^k = 0$ for some positive integer k . The smallest integer l such that $\mathcal{G}^{l+1} = 0$ is called the *order of nilpotency*. A Lie algebra $\mathcal{G}$ is said to be *simple* if it is non-abelian and has no ideals other than 0 and $\mathcal{G}$. It is *semi-simple* if it is non-abelian and has no abelian ideals other than 0. A Lie group G is simple (resp. semi-simple) if its associated Lie algebra is simple (resp. semi-simple). Simplicity implies semi-simplicity.

2.1.2 Control Systems on Lie Groups

We define a control system on the n -dimensional Lie group G as

$$\dot{g} = Y_0(g) + \sum_{i=1}^m u_i(t) Y_i(g), \quad g(t) \in G, \quad u_i(t) \in \mathfrak{R}, \quad m \leq n, \quad (2.5)$$

where $Y_0, \dots, Y_m$ are vector fields on G and $u = (u_1, \dots, u_m)$ are interpreted as the control inputs. If $Y_0, \dots, Y_m$ are left-invariant (resp. right-invariant) on G then we say that (2.5) is a left-invariant (resp. right-invariant) control system on G . If (2.5) is left-invariant we can express it as

$$\dot{g} = T_e L_g (\xi_0 + \sum_{i=1}^m u_i(t) \xi_i), \quad g(t) \in G, \quad u_i(t) \in \mathfrak{R}, \quad m \leq n, \quad (2.6)$$

for some fixed $\xi_0, \xi_1, \dots, \xi_m \in \mathcal{G}$. Similarly, if (2.5) is right-invariant it can be written as

$$\dot{g} = T_e R_g (\xi_0 + \sum_{i=1}^m u_i(t) \xi_i), \quad g(t) \in G, \quad u_i(t) \in \mathfrak{R}, \quad m \leq n, \quad (2.7)$$

for some fixed $\xi_0, \xi_1, \dots, \xi_m \in \mathcal{G}$. The term $T_e L_g \cdot \xi_0$ in (2.6) (similarly $T_e R_g \cdot \xi_0$ in (2.7)) is called the *drift* term since g drifts as $\dot{g} = T_e L_g \cdot \xi_0$ if $u = 0$. Thus, system (2.6) (similarly, (2.7)) is called *drift-free* if the drift term is zero.

In this dissertation we focus on left-invariant systems on Lie groups, but we note that analogous results can be derived for right-invariant systems. In fact, given a right-invariant system (2.7), we can always convert it into a left-invariant system by considering $g^{-1}(t)$ as our state trajectory. Let $g(t) \in G$ satisfy (2.7) and define $\xi(t) = \xi_0 + \sum_{i=1}^m u_i(t) \xi_i$. Since $\frac{d}{dt}(g(t)g^{-1}(t)) = \frac{d}{dt}(e) = 0$, we have that

$$\begin{aligned} 0 = \frac{d}{dt}(gg^{-1}) &= \frac{\partial}{\partial g}(R_{g^{-1}}g)\dot{g} + \frac{\partial}{\partial g^{-1}}(L_g g^{-1})\frac{d}{dt}(g^{-1}) \\ &= (T_g R_{g^{-1}})\dot{g} + (T_{g^{-1}} L_g)\frac{d}{dt}(g^{-1}) \end{aligned}$$

$$\begin{aligned}
&= (T_g R_{g^{-1}})(T_e R_g)\xi(t) + (T_{g^{-1}} L_g) \frac{d}{dt}(g^{-1}) \\
&= \xi(t) + (T_{g^{-1}} L_g) \frac{d}{dt}(g^{-1}).
\end{aligned}$$

This implies $g^{-1}(t)$ satisfies

$$\frac{d}{dt}(g^{-1}) = T_e L_{g^{-1}}(-\xi(t)) \quad (2.8)$$

which is a left-invariant system on G . Thus, there is no loss of generality in specializing to left-invariant systems.

2.1.3 Problem Statement on Lie Groups

The systems of interest in this dissertation take the form

$$\dot{g} = \epsilon T_e L_g \cdot U(t), \quad U(t) = \sum_{i=1}^m u_i(t) \xi_i, \quad m \leq n, \quad (2.9)$$

where $g(t)$ is a curve in the n -dimensional Lie group G and $U(t)$ is a curve in the Lie algebra $\mathcal{G}$ associated with G . Here $\xi_1, \dots, \xi_m$ are fixed, linearly independent elements of $\mathcal{G}$ which can be completed such that $\{\xi_1, \dots, \xi_n\}$ is a basis for $\mathcal{G}$. The $u_i(t)$ are scalars meant to represent control inputs. We will assume that the $u_i(t)$, $i = 1, \dots, m$ are periodic in t of common period T . The factor ϵ is a small parameter representing the small amplitude of the control inputs. That is, $\epsilon u_1(t), \dots, \epsilon u_m(t)$ are our small-amplitude, periodic controls. We will sometimes extend the vector of control components to be of length n by writing $u(t) = (u_1(t), \dots, u_n(t))$ where we set $u_{m+1} = \dots = u_n = 0$. Since 0 is trivially periodic in t of any period T , then the extended vector $u(t)$ is still periodic in t of period T .

We note that (2.9) as described is a drift-free system. However, if at least one of the components u_i is not an adjustable control then (2.9) is a system with

drift. The averaging theory of Chapter 4 does not depend on the system being drift-free, only on the curve $U(t)$ being periodic in t of common period T . In Chapter 6 we address a motion control problem for a system with drift.

We also note that (2.9) is left-invariant. One important implication of left-invariance is that the solution $g(t)$ to (2.9) with any initial condition $g(0) = g_0$ can be expressed as $g(t) = g_0 g_e(t)$ where $g_e(t)$ is the solution to (2.9) with initial condition $g_e(0) = e \in G$. The right-invariant analogue to (2.9) is given by

$$\dot{g} = \epsilon T_e R_g \cdot U(t), \quad U(t) = \sum_{i=1}^m u_i(t) \xi_i, \quad m \leq n. \quad (2.10)$$

The solution $g(t)$ to (2.10) with initial condition $g(0) = g_0 \in G$ is $g(t) = g_e(t) g_0$ where $g_e(t)$ is the solution to (2.10) with $g_e(0) = e \in G$.

The constructive controllability problem for system (2.9) (or (2.10)) can be stated formally as

(P) Given an initial condition $g_i \in G$, a final condition $g_f \in G$ and a time $t_f > 0$,
find $u(t) = (u_1(t), \dots, u_m(t))$, $t \in [0, t_f]$, such that $g(0) = g_i$ and $g(t_f) = g_f$.

2.1.4 Control Systems on Matrix Lie groups

Matrix Lie groups are Lie groups with matrix elements. They are an important special class for our purposes since they are the Lie groups of interest for all of the applications in this dissertation. Every matrix Lie group over $\mathfrak{R}$ is a subgroup of the *general linear* matrix Lie group $GL(n)$ (for some n) which is the space of $n \times n$ nonsingular matrices, under matrix multiplication. Examples of other matrix Lie groups include

$$\begin{aligned}
SL(n) &= \{X \in GL(n) \mid \det X = 1\} \\
O(n) &= \{X \in GL(n) \mid X^T X = I\} \\
SO(n) &= \{X \in GL(n) \mid X \in O(n) \cap SL(n)\} \\
SE(n) &= \left\{ \begin{pmatrix} \mathcal{A} & b \\ 0 & 1 \end{pmatrix} \in GL(n+1) \mid \mathcal{A} \in SO(n), b \in \mathbb{R}^n \right\} \\
\text{Unipotent group} &= \{X = [X_{ij}] \in GL(n) \mid X_{ii} = 1, X_{ij} = 0 \text{ if } i > j\}.
\end{aligned}$$

$SL(n)$ is called the *special linear group*, $O(n)$ the *orthogonal group*, $SO(n)$ the *special orthogonal group*, and $SE(n)$ the *special Euclidean group*.

The Lie algebra $gl(n)$ associated with $GL(n)$ is the space of all $n \times n$ matrices with the Lie bracket operator defined as the matrix commutator, i.e.,

$$[A, B] = AB - BA, \quad \forall A, B \in gl(n).$$

The Lie algebras of other matrix Lie groups are subspaces of $gl(n)$ with the same Lie bracket operation. For example,

$$\begin{aligned}
sl(n) &= \{A \in gl(n) \mid \text{tr}(A) = 0\} \\
so(n) &= \{A \in gl(n) \mid A^T + A = 0\} \\
se(n) &= \left\{ \begin{pmatrix} A & x \\ 0 & 0 \end{pmatrix} \in gl(n+1) \mid A \in so(n), x \in \mathbb{R}^n \right\}
\end{aligned}$$

$$\text{Lie algebra of unipotent group} = \{A = [A_{ij}] \in gl(n) \mid A_{ij} = 0 \text{ if } i \geq j\}.$$

Here, the name of the Lie algebra is the same as the associated group but in lower-case letters. Since $SO(n)$ is the connected component of $O(n)$ containing the identity, $so(n)$ is the Lie algebra associated to both $SO(n)$ and $O(n)$. We note that $SL(k)$, $O(n)$, $SO(n)$ are simple Lie groups (for $k > 1$, $n > 2$) since their associated Lie algebras are simple. Also, $se(2)$ is a solvable Lie algebra, and the Lie algebras of unipotent groups are all nilpotent.

Left-invariant (resp. right-invariant) vector fields on a matrix Lie group G take the form XA (resp. AX) with $X \in G$ and $A \in \mathcal{G}$, where $\mathcal{G}$ is the associated matrix Lie algebra. The exponential map from matrix Lie algebras to matrix Lie groups is equivalent to the matrix exponential defined as $\exp(A) = \sum_{i=0}^{\infty} \frac{1}{i!} A^i$, $A \in \mathcal{G}$. The left-invariant control system described by (2.9) specialized to matrix Lie groups becomes

$$\dot{X} = \epsilon XU(t), \quad U(t) = \sum_{i=1}^m u_i(t) A_i, \quad m \leq n, \quad (2.11)$$

where $X(t)$ is a curve in the n -dimensional Lie group G and $U(t)$ is a curve in the Lie algebra $\mathcal{G}$ associated with G . $A_1, \dots, A_m$ are constant matrices in $\mathcal{G}$ which we can complete as $\{A_1, \dots, A_n\}$ to be a basis for $\mathcal{G}$. The terms $u_i(t)$ and the parameter ϵ are as defined for (2.9).

2.2 Controllability and Geometry

In this section we review basic definitions and theorems on nonlinear controllability with application to systems on Lie groups. We then define geometric quantities that are the essential elements of the basis-independent, average motion control solutions derived in later chapters.

2.2.1 Controllability

We review controllability theory for left-invariant control systems on Lie groups defined by equations (2.6), based on [11, 29]. Although these papers deal with right-invariant systems, by the argument of Section 2.1.2, we can derive analogous results for left-invariant systems. Let G be the system group and let $\mathcal{G}$ be its associated Lie algebra. As will shown below, it is the space spanned

by the system vector fields and their Lie brackets that play a critical role in the determination of controllability.

Following Jurdjevic and Sussmann [29], we make the following definitions. $\mathcal{U}$ is the class of admissible controls where $\mathcal{U}$ is either $\mathcal{U}_u$, $\mathcal{U}_r$ or $\mathcal{U}_b$ and

- (i) $\mathcal{U}_u$ is the class of locally bounded, measurable functions on $[0, \infty)$ taking values in $\mathbb{R}^m$.
- (ii) $\mathcal{U}_r \subset \mathcal{U}_u$ with elements taking values in the unit n -dimensional cube.
- (iii) $\mathcal{U}_b$ is the class of piecewise constant functions on $[0, \infty)$ taking values in $\mathbb{R}^m$ where components of its elements take values in the set $\{-1, 1\}$.

If $u \in \mathcal{U}$ and $g_0 \in G$, we denote the solution g of (2.6) which satisfies $g(0) = g_0$ by $\pi(g_0, u, \cdot)$, i.e., $g(t) = \pi(g_0, u, t)$, $\forall t \in [0, \infty)$. If $\pi(g_0, u, t) = g_1$ for some $t \geq 0$ then we say u steers g_0 into g_1 in t units of time. System (2.6) is said to be *controllable* if there exists an admissible control $u \in \mathcal{U}$ that steers any $g_0 \in G$ into any $g_1 \in G$, with no constraints on how many units of time are required. Let $\mathcal{C}$ denote the set of Lie brackets generated by $\{\xi_1, \dots, \xi_m\}$ defined as

$$\mathcal{C} = \{\eta \mid \eta = [\eta_k, [\eta_{k-1}, [\dots, [\eta_1, \eta_0] \dots]]], \eta_i \in \{\xi_1, \dots, \xi_m\}, i = 0, \dots, k\}. \quad (2.12)$$

Similarly, let $\mathcal{C}_0$ denote the set of Lie brackets generated by $\{\xi_0, \dots, \xi_m\}$, i.e.,

$$\mathcal{C}_0 = \{\eta \mid \eta = [\eta_k, [\eta_{k-1}, [\dots, [\eta_1, \eta_0] \dots]]], \eta_i \in \{\xi_0, \dots, \xi_m\}, i = 0, \dots, k\}. \quad (2.13)$$

We say that $\text{span}(\mathcal{C})$ (resp. $\text{span}(\mathcal{C}_0)$) is the *Lie algebra generated by* $\{\xi_1, \dots, \xi_m\}$ (resp. $\{\xi_0, \dots, \xi_m\}$). We first give the controllability condition for drift-free systems. We refer to this condition as the *Lie algebra controllability rank condition*.

Theorem 2.1 (Jurdjevic and Sussmann) Let (2.6) be a drift-free system on G , a connected Lie group. Then (2.6) is controllable with $u \in \mathcal{U}$ if and only if $\text{span}(\mathcal{C}) = \mathcal{G}$. If $\mathcal{U} = \mathcal{U}_u$ then the system is controllable in arbitrarily short time.

There is a similar condition for systems with drift on compact Lie groups.

Theorem 2.2 (Jurdjevic and Sussmann) Let (2.6) be a system with drift on G , a compact and connected Lie group. Then (2.6) is controllable with $u \in \mathcal{U}$ if and only if $\text{span}(\mathcal{C}_0) = \mathcal{G}$. Further, $\exists t' > 0$ such that for every $g_0, g_1 \in G$ there is a control $u \in \mathcal{U}$ that steers g_0 into g_1 in less than t' units of time. If, further, G is semi-simple, then $\exists t' > 0$ such that for every $g_0, g_1 \in G$ there is a control $u \in \mathcal{U}$ that steers g_0 into g_1 in exactly t' units of time.

We next make definitions to further characterize the controllability of system (2.6). Here, we assume that the system is drift-free, i.e., that $\xi_0 = 0$. We define a *depth- μ Lie bracket* on $\mathcal{G}$ as μ iterated brackets, e.g., $[\eta_\mu, [\eta_{\mu-1}, [\dots, [\eta_1, \eta_0] \dots]]]$. A depth-zero bracket is just an element of the Lie algebra $\mathcal{G}$. We often refer to a depth-one bracket as a *single bracket*, a depth-two bracket as a *double bracket*, etc. If j is the minimum positive integer such that the Lie algebra controllability rank condition is satisfied with $\mathcal{C}$ containing only up to depth- j brackets, i.e., $k \leq j$ in (2.12), then we say that system (2.6) is a *depth- j bracket system*. We refer to a depth-one bracket system as a *single-bracket system*, a depth-two bracket as a *double-bracket system*, etc.

Following [23, 60] we define the growth vector and relative growth vector. Let

$$D_1 \triangleq \text{span}\{\xi_1, \dots, \xi_m\},$$

$$D_i \triangleq D_{i-1} + \text{span}[D_1, D_{i-1}],$$

The *growth vector* for a depth- $(p-1)$ system is defined as

$$r = (r_1, \dots, r_p) \in \mathcal{Z}^p, \text{ where } r_i = \dim D_i.$$

The *relative growth vector* is defined as

$$\sigma = (\sigma_1, \dots, \sigma_p) \in \mathcal{Z}^p, \text{ where } \sigma_i = r_i - r_{i-1}, \quad r_0 = 0. \quad (2.14)$$

Thus, for a single-bracket system, $r = (m, n)$, $\sigma = (m, n - m)$. For a double-bracket system, we define $0 \leq l \leq (n - m)$ such that $r = (m, m + l, n)$, $\sigma = (m, l, n - (m + l))$.

2.2.2 Geometry

Given a basis $\{\xi_1, \dots, \xi_n\}$ for a Lie algebra $\mathcal{G}$, one computes the associated *structure constants* Γ_{ij}^k as

$$[\xi_i, \xi_j] = \sum_{k=1}^n \Gamma_{ij}^k \xi_k, \quad i, j = 1, \dots, n, \quad (2.15)$$

We define the *depth-two structure constants* θ_{ijk}^p associated with basis $\{\xi_1, \dots, \xi_n\}$ by

$$\theta_{ijk}^p \triangleq \sum_{l=1}^n \Gamma_{ij}^l \Gamma_{lk}^p. \quad (2.16)$$

This definition comes from the computation of structure constants for depth-two Lie brackets as follows:

$$[[\xi_i, \xi_j], \xi_k] = \left[\sum_{l=1}^n \Gamma_{ij}^l \xi_l, \xi_k \right] = \sum_{l=1}^n \Gamma_{ij}^l [\xi_l, \xi_k] = \sum_{p=1}^n \sum_{l=1}^n \Gamma_{ij}^l \Gamma_{lk}^p \xi_p = \sum_{p=1}^n \theta_{ijk}^p \xi_p.$$

Skew symmetry of the Lie bracket on $\mathcal{G}$ and the Jacobi identity imply

$$\Gamma_{ij}^k = -\Gamma_{ji}^k, \quad \theta_{ijk}^p = -\theta_{jik}^p, \quad \theta_{ijk}^p + \theta_{jki}^p + \theta_{kij}^p = 0. \quad (2.17)$$

Higher-order structure constants can similarly be defined. Let I_p be the ordered list of indices $I_p = \{i_1, i_2, \dots, i_p\}$ where $i_\nu \in \{1, \dots, n\}$, $\nu = 1, \dots, p$. Then we define the *depth- $(p-1)$ structure constants* $\Gamma_{I_p}^k$ as

$$[\xi_{i_1}, [\xi_{i_2}, \dots, [\xi_{i_{p-1}}, \xi_{i_p}] \dots]] = \sum_{k=1}^n \Gamma_{I_p}^k \xi_k. \quad (2.18)$$

Note that $\theta_{ijk}^p = \Gamma_{kji}^p$.

We make the following definitions given that $u = (u_1, \dots, u_m)$ is periodic in t of common period T . Let the time average of u be $u_{av} = (u_{av1}, \dots, u_{avm})^T$ where

$$u_{avi} = \frac{1}{T} \int_0^T u_i(\tau) d\tau, \quad (2.19)$$

and let the time integral of u be $\tilde{u} = (\tilde{u}_1, \dots, \tilde{u}_m)^T$ where

$$\tilde{u}_i(t) = \int_0^t u_i(\tau) d\tau. \quad (2.20)$$

So $u = \dot{\tilde{u}}$. If $u_{av} = 0$ then $\tilde{u}$ is periodic in t with common period T . We also let

$$U_{av} = \sum_{i=1}^m u_{avi} \xi_i, \quad \tilde{U} = \sum_{i=1}^m \tilde{u}_i \xi_i. \quad (2.21)$$

We next define area and moment terms that are fundamental to our average solution formulas of Chapter 4. Define

$$Area_{ij}(T) = \frac{1}{2} \int_0^T (\tilde{u}_i(\sigma) \dot{\tilde{u}}_j(\sigma) - \tilde{u}_j(\sigma) \dot{\tilde{u}}_i(\sigma)) d\sigma. \quad (2.22)$$

If $u_{av} = 0$, then by Green's Theorem, $Area_{ij}(T)$ is the area bounded by the closed curve described by $\tilde{u}_i$ and $\tilde{u}_j$ over one period, i.e., from $t = 0$ to $t = T$. This area can be interpreted as the projection onto the i - j plane of the area enclosed by the curve $(\tilde{u}_1, \dots, \tilde{u}_m)$ in one period. If $u_{avi} \neq 0$ and/or $u_{avj} \neq 0$ then $Area_{ij}(T)$ can be interpreted as the area bounded by the curve described by $\tilde{u}_i$ and $\tilde{u}_j$ and the straight line described by $u_{avi}t$ and $u_{avj}t$ over one period. Define

$$a_{ij}(t) = \frac{1}{2} \int_0^t (\tilde{u}_i(\sigma) \dot{\tilde{u}}_j(\sigma) - \tilde{u}_j(\sigma) \dot{\tilde{u}}_i(\sigma)) d\sigma. \quad (2.23)$$

For $u_{av} = 0$, $a_{ij}(t)$ is of the form

$$a_{ij}(t) = \frac{t}{T} \text{Area}_{ij}(T) + f(t), \quad (2.24)$$

where $f(t+T) = f(t)$, $f(0) = 0$. So, in particular,

$$a_{ij}(MT) = M \text{Area}_{ij}(T) \quad (2.25)$$

where M is a positive integer. Let

$$m_{ijk}(T) = \frac{1}{3} \int_0^T (\tilde{u}_i(\sigma) \dot{\tilde{u}}_j(\sigma) - \tilde{u}_j(\sigma) \dot{\tilde{u}}_i(\sigma)) \tilde{u}_k(\sigma) d\sigma. \quad (2.26)$$

Assume $u_{av} = 0$ and consider the closed curve C defined by $\tilde{u}_i(t)$, $\tilde{u}_j(t)$ and $\tilde{u}_k(t)$ over one period, i.e., from $t = 0$ to $t = T$. From (2.26) we get that

$$m_{ijk}(T) = \frac{1}{3} \oint_C \tilde{u}_i \tilde{u}_k d\tilde{u}_j - \tilde{u}_j \tilde{u}_k d\tilde{u}_i.$$

Let A be any oriented surface with boundary $\partial A = C$. Then by Stokes' Theorem,

$$m_{ijk}(T) = \frac{1}{3} \int_A -\tilde{u}_i d\tilde{u}_j d\tilde{u}_k - \tilde{u}_j d\tilde{u}_k d\tilde{u}_i + 2\tilde{u}_k d\tilde{u}_i d\tilde{u}_j. \quad (2.27)$$

So $m_{ijk}(T)$ as described by (2.27) can be interpreted as a first moment, i.e., a linear combination of $\tilde{u}_i$ integrated over the area of the projection of A onto the j - k plane, $\tilde{u}_j$ integrated over the area of the projection of A onto the k - i plane and $\tilde{u}_k$ integrated over the area of the projection of A onto the i - j plane. We have the following (skew-symmetric and Jacobi-like) identities:

$$\text{Area}_{ij}(T) = -\text{Area}_{ji}(T), \quad a_{ij}(t) = -a_{ji}(t), \quad (2.28)$$

$$m_{ijk}(T) = -m_{jik}(T), \quad m_{ijk}(T) + m_{jki}(T) + m_{kij}(T) = 0. \quad (2.29)$$

2.3 Local Solutions

Since, in general, there are no explicit global representations of the solution to (2.9) we make use of local representations: the *product of exponentials representation* given by Wei and Norman [77] and the *single exponential representation* given by Magnus [54] and Fomenko and Chakon [21].

2.3.1 Product of Exponentials

We begin by defining the product of exponentials representation which is given by Wei and Norman [77]. They state their result for right-invariant systems on matrix Lie groups, but we prove the more general result for a left-invariant system of the form (2.9) on a finite-dimensional abstract Lie group G with Lie algebra $\mathcal{G}$. We use the extended control vector $u = (u_1, \dots, u_n)^T$ where $u_{m+1} = \dots = u_n = 0$.

Lemma 2.3 (Wei and Norman) Let $g(t)$ be the solution to (2.9) with $g(0) = e \in G$. Then $\exists t_0 > 0$ such that for $|t| < t_0$, $g(t)$ can be expressed in the form

$$g(t) = \prod_{i=1}^n e^{\gamma_i(t)\xi_i} \triangleq e^{\gamma_1(t)\xi_1} e^{\gamma_2(t)\xi_2} \dots e^{\gamma_n(t)\xi_n} . \quad (2.30)$$

The Wei-Norman parameters $\gamma = (\gamma_1, \dots, \gamma_n)^T$ satisfy

$$\dot{\gamma} = \epsilon M(\gamma)u , \quad \text{for } |t| < t_0 , \quad (2.31)$$

where $e^{\gamma_i \xi_i} = \exp(\gamma_i \xi_i)$ is the exponential map defined by (2.1), $\gamma(0) = 0$ and $M(\gamma)$ is a real-analytic, matrix-valued function of γ . If $\mathcal{G}$ is solvable then there exists a basis of $\mathcal{G}$ and an ordering of this basis for which (2.31) holds globally, i.e., for all t , and in that case (2.31) can be integrated by quadrature.

Proof: Let $g(t)$ satisfy (2.30). Then

$$\dot{g} = \sum_{i=1}^n \dot{\gamma}_i \frac{\partial}{\partial \gamma_i} (L_{e^{\gamma_1 \xi_1}} \dots L_{e^{\gamma_{i-1} \xi_{i-1}}} R_{e^{\gamma_n \xi_n}} \dots R_{e^{\gamma_{i+1} \xi_{i+1}}} (e^{\gamma_i \xi_i}))$$

$$\begin{aligned}
&= \sum_{i=1}^n \dot{\gamma}_i T_{e^{\gamma_i \xi_i}} (L_{e^{\gamma_1 \xi_1}} \cdots L_{e^{\gamma_{i-1} \xi_{i-1}}} R_{e^{\gamma_n \xi_n}} \cdots R_{e^{\gamma_{i+1} \xi_{i+1}}}) \frac{\partial}{\partial \gamma_i} (e^{\gamma_i \xi_i}) \\
&= \sum_{i=1}^n \dot{\gamma}_i T_{e^{\gamma_i \xi_i}} (L_g L_{e^{-\gamma_n \xi_n}} \cdots L_{e^{-\gamma_i \xi_i}} R_{e^{\gamma_n \xi_n}} \cdots R_{e^{\gamma_{i+1} \xi_{i+1}}}) T_e L_{e^{\gamma_i \xi_i}} \cdot \xi_i \\
&= \sum_{i=1}^n \dot{\gamma}_i (T_e L_g) T_{e^{\gamma_i \xi_i}} (L_{e^{-\gamma_n \xi_n}} \cdots L_{e^{-\gamma_i \xi_i}} R_{e^{\gamma_n \xi_n}} \cdots R_{e^{\gamma_{i+1} \xi_{i+1}}}) T_e L_{e^{\gamma_i \xi_i}} \cdot \xi_i \\
&= T_e L_g \sum_{i=1}^n \dot{\gamma}_i T_{e^{\gamma_i \xi_i}} (R_{e^{\gamma_n \xi_n}} L_{e^{-\gamma_n \xi_n}} \cdots R_{e^{\gamma_{i+1} \xi_{i+1}}} L_{e^{-\gamma_{i+1} \xi_{i+1}}} L_{e^{-\gamma_i \xi_i}}) T_e L_{e^{\gamma_i \xi_i}} \cdot \xi_i \\
&= T_e L_g \sum_{i=1}^n \dot{\gamma}_i T_e (R_{e^{\gamma_n \xi_n}} L_{e^{-\gamma_n \xi_n}} \cdots R_{e^{\gamma_{i+1} \xi_{i+1}}} L_{e^{-\gamma_{i+1} \xi_{i+1}}}) (T_{e^{\gamma_i \xi_i}} L_{e^{-\gamma_i \xi_i}}) T_e L_{e^{\gamma_i \xi_i}} \cdot \xi_i \\
&= T_e L_g \sum_{i=1}^n \dot{\gamma}_i T_e (R_{e^{\gamma_n \xi_n}} L_{e^{-\gamma_n \xi_n}}) \cdots T_e (R_{e^{\gamma_{i+1} \xi_{i+1}}} L_{e^{-\gamma_{i+1} \xi_{i+1}}}) \cdot \xi_i \\
&= T_e L_g \sum_{i=1}^n \dot{\gamma}_i \text{Ad}_{e^{-\gamma_n \xi_n}} (\text{Ad}_{e^{-\gamma_{n-1} \xi_{n-1}}} (\cdots (\text{Ad}_{e^{-\gamma_{i+1} \xi_{i+1}}} \xi_i))) \\
&= T_e L_g \sum_{i=1}^n \dot{\gamma}_i e^{-\gamma_n \text{ad}_{\xi_n}} (e^{-\gamma_{n-1} \text{ad}_{\xi_{n-1}}} (\cdots (e^{-\gamma_{i+1} \text{ad}_{\xi_{i+1}}} \xi_i))), \tag{2.32}
\end{aligned}$$

where

$$e^{-\gamma_j \text{ad}_{\xi_j}} \xi_i = \sum_{l=0}^{\infty} \frac{1}{l!} (-\gamma_j)^l \text{ad}_{\xi_j}^l \xi_i \in \mathcal{G}. \tag{2.33}$$

Premultiplying both (2.32) and (2.9) by $(T_e L_g)^{-1}$ and equating gives

$$\sum_{i=1}^n \dot{\gamma}_i e^{-\gamma_n \text{ad}_{\xi_n}} (e^{-\gamma_{n-1} \text{ad}_{\xi_{n-1}}} (\cdots (e^{-\gamma_{i+1} \text{ad}_{\xi_{i+1}}} \xi_i))) = \sum_{i=1}^n \epsilon u_i \xi_i.$$

By (2.33) and following [77] we can define $\alpha_{ij}(\gamma)$ analytic in γ as

$$e^{-\gamma_n \text{ad}_{\xi_n}} (\cdots (e^{-\gamma_{j+1} \text{ad}_{\xi_{j+1}}} \xi_j)) = \sum_{i=1}^n \alpha_{ij}(\gamma) \xi_i \tag{2.34}$$

such that

$$\sum_{j=1}^n \dot{\gamma}_j \sum_{i=1}^n \alpha_{ij}(\gamma) \xi_i = \sum_{i=1}^n \epsilon u_i \xi_i.$$

Since the ξ_i 's are linearly independent, we equate coefficients of ξ_i to get

$$\hat{M}(\gamma) \dot{\gamma} \triangleq \begin{pmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \vdots & & & \vdots \\ \alpha_{n1} & \cdots & \alpha_{nn} \end{pmatrix} \begin{pmatrix} \dot{\gamma}_1 \\ \vdots \\ \dot{\gamma}_n \end{pmatrix} = \begin{pmatrix} \epsilon u_1 \\ \vdots \\ \epsilon u_n \end{pmatrix}.$$

At $t = 0$, $\gamma(0) = 0$ and so $\hat{M}(0) = I$. Thus, there exists a neighborhood N_0 of $t = 0$ such that $\hat{M}(\gamma(t))$ is nonsingular $\forall t \in N_0$. We define $M(\gamma) = \hat{M}^{-1}(\gamma)$, $\forall t \in N_0$ and the first part of the lemma follows. If $\mathcal{G}$ is solvable, then it can be shown that for a suitable choice of basis, $\hat{M}(\gamma)$ is lower triangular and the above differential equations are integrable by quadrature. $\square$

Remark 2.4 From this proof we observe that we can express $M(\gamma)$ in terms of the structure constants associated to $(\xi_1, \dots, \xi_n)$ defined by (2.15). By (2.33) and (2.34)

$$\begin{aligned}
\sum_{i=1}^n \alpha_{ij}(\gamma) \xi_i &= e^{-\gamma_n \text{ad}_{\xi_n}} (\dots (e^{-\gamma_{j+1} \text{ad}_{\xi_{j+1}}} \xi_j)) \\
&= (I - \gamma_n \text{ad}_{\xi_n} + \dots)(I - \gamma_{n-1} \text{ad}_{\xi_{n-1}} + \dots) \cdots (I - \gamma_{j+1} \text{ad}_{\xi_{j+1}} + \dots) \xi_j \\
&= (I - \sum_{k=j+1}^n \gamma_k \text{ad}_{\xi_k} + O(\gamma^2)) \xi_j \\
&= \xi_j - \sum_{k=j+1}^n \gamma_k [\xi_k, \xi_j] + O(\gamma^2) \\
&= \xi_j - \sum_{i=1}^n (\sum_{k=j+1}^n \gamma_k \Gamma_{kj}^i) \xi_i + O(\gamma^2)
\end{aligned}$$

where $O(\gamma^2)$ is quadratic or higher order in the components of γ . So

$$\alpha_{ij}(\gamma) = \begin{cases} -\sum_{k=j+1}^n \gamma_k \Gamma_{kj}^i + O(\gamma^2), & i \neq j \\ 1 - \sum_{k=j+1}^n \gamma_k \Gamma_{kj}^i + O(\gamma^2), & i = j. \end{cases}$$

Thus, $\hat{M}(\gamma) = I - \tilde{\alpha}(\gamma) + O(\gamma^2)$ where the ij th element of $\tilde{\alpha}$ is

$$\tilde{\alpha}_{ij}(\gamma) = \sum_{k=j+1}^n \gamma_k \Gamma_{kj}^i. \tag{2.35}$$

Since $M(\gamma) = \hat{M}^{-1}(\gamma)$, we have that (for $\|\gamma\|$ small)

$$M(\gamma) = I + \tilde{\alpha}(\gamma) + O(\gamma^2). \tag{2.36}$$

We use this observation in Chapter 4.

It is customary to refer to components of γ as the *canonical coordinates of the second kind* for G . Let W be the open neighborhood of $0 \in \mathfrak{R}^n$ such that $\forall \gamma \in W$, $M(\gamma)$ is well-defined. Let $\text{Pr}: \mathfrak{R}^n \rightarrow G$ define the mapping

$$\text{Pr}(\gamma) = e^{\gamma_1 \xi_1} e^{\gamma_2 \xi_2} \dots e^{\gamma_n \xi_n} \quad (2.37)$$

and define $V = \text{Pr}(W) \subset G$. Then, the Wei-Norman formulation provides a local representation of the solution to (2.9) for initial condition $g(0) \in V \subset G$. Now let S be the largest neighborhood of $0 \in \mathfrak{R}^n$ contained in W such that $\Psi = \text{Pr}|_S : S \rightarrow G$ is injective. Let $Q = \Psi(S) \subset V$. Then $\Psi : S \rightarrow Q$ is a diffeomorphism and we can define a metric $\tilde{d} : Q \times Q \rightarrow \mathfrak{R}_+$ by

$$\tilde{d}(X, Y) = d(\Psi^{-1}(X), \Psi^{-1}(Y)) \quad (2.38)$$

where for $\|\cdot\|$, a norm on $\mathfrak{R}^n$, $d : \mathfrak{R}^n \times \mathfrak{R}^n \rightarrow \mathfrak{R}_+$ is given by

$$d(\alpha, \beta) = \|\alpha - \beta\|. \quad (2.39)$$

2.3.2 Single Exponential

The solution to (2.9) with $g(0) = e$ can also be represented locally as a single exponential,

$$g(t) = e^{Z(t)} = \exp(Z(t)), \quad g(t) \in G, \quad Z(t) \in \mathcal{G}. \quad (2.40)$$

This is irrespective of ϵ in (2.9) being a small parameter. However, we include ϵ in the expressions below since we will need it in later chapters. Feynman represented the solution (2.40) symbolically as $\text{Exp} \int_0^t \epsilon U(\tau) d\tau$ (sometimes referred to as a *T-product of exponents*) and derived a formula for it in terms of a certain formal series [20]. $Z(t)$ is sometimes referred to as the logarithm of a *multiplicative*

integral, represented symbolically as $Z(t) = \ln \prod_{\tau=0}^t e^{\epsilon U(\tau) d\tau}$. Magnus showed for (2.10) (the right-invariant analogue of (2.9)) that $Z(t)$ locally satisfies

$$\frac{dZ}{dt} = \varphi(\text{ad}_{Z(t)})(\epsilon U(t)), \quad Z(0) = 0, \quad (2.41)$$

where $\varphi(x) = x/(e^x - 1)$ [54]. Further, Magnus illustrated that (2.41) can be integrated by iteration to yield an expression for $Z(t)$ as an infinite series of iterated integrals. This formula is the continuous analogue of the Campbell-Baker-Hausdorff formula which gives the solution to C where $e^C = e^A e^B$ and A and B are noncommuting elements of a Lie algebra (close enough to 0). Magnus did not provide general convergence criteria for his infinite series solution.

Karasev and Mosolova addressed the problem of convergence criteria and, using the theory of ordered operators, derived an explicit expression for $Z(t)$ for the right-invariant system (2.10) [31]. The norm $\|\cdot\|$ on $\mathcal{G}$ in their work (and in the work of Fomenko and Chakon that follows) is assumed to be one that makes $\mathcal{G}$ a Banach algebra. So for matrix Lie algebras, $\|\cdot\|$ is the vector space norm. Let $\phi(x) = \ln x/(x - 1)$ and $r = \exp(\int_0^t \|\text{ad}_{\epsilon U(\tau)}\| d\tau)$. Consider the extended control vector $u = (u_1, \dots, u_n)^T$ where $u_{m+1} = \dots = u_n = 0$.

Theorem 2.5 (Karasev and Mosolova) If $r < 2$ then for $\epsilon U(t) = \epsilon \sum_{i=1}^n u_i(t) \xi_i$,

$$\text{Exp} \int_0^t \epsilon U(\tau) d\tau = e^{Z(t)},$$

where $Z(t)$ is an element in the Lie algebra generated by $\{\xi_1, \dots, \xi_m\}$. $Z(t)$ is given by

$$Z(t) = \int_0^t \phi(\text{Exp} \int_0^\tau \text{ad}_{\epsilon U(\sigma)} d\sigma) \epsilon U(\tau) d\tau$$

and satisfies

$$\|Z(t)\| \leq \frac{1}{r-1} \ln \left(\frac{1}{2-r} \right) \int_0^t \|\epsilon U(\tau)\| d\tau.$$

For G a finite-dimensional Lie group, the convergence criterion $r < 2$ is equivalent to

$$\int_0^t \|\Lambda(\epsilon u(\tau))\| d\tau < \ln 2, \quad (2.42)$$

where $\Lambda(\cdot)$ is an $n \times n$ matrix with ij th element $\Lambda_{ij}(\cdot)$ defined by

$$\Lambda_{ij}(u) = \sum_{k=1}^n u_k \Gamma_{kj}^i.$$

In the case that $G = SO(3)$ and $\{A_1, A_2, A_3\}$ (defined in Section 3.1) is the standard basis for $\mathcal{G} = so(3)$, it is easy to compute that $\Lambda(\epsilon u) = \epsilon U$ and so (2.42) is equivalent to

$$\int_0^t \|\epsilon U(\tau)\| d\tau < \ln 2.$$

More recently, Fomenko and Chakon studied the problem of representing $Z(t)$ as a convergent series $Z(t) = \sum_{i=1}^{\infty} Z_i$ [21]. They derived explicit recursion relations satisfied by the terms Z_i . The following theorem states their results (in part) for the right-invariant system (2.10). The theorem uses the universal constant $\hat{b}$ defined as follows. Let $h(x)$ be the solution of the scalar equation $dh/dx = e^x(q(h) + h/2)$ with $h(0) = 0$ and with $q(h) = 1 + \sum_{p=1}^{\infty} |k_{2p}| h^{2p}$. Here, $k_{2p} \cdot (2p)! = \beta_{2p}$ and β_{2p} are Bernoulli numbers given by

$$\frac{x}{e^x - 1} = \sum_{j=0}^{\infty} \frac{\beta_j x^j}{j!}.$$

For example, $\beta_0 = 1$, $\beta_1 = -1/2$, $\beta_2 = 1/6$, $\beta_4 = -1/30$, $\dots$. The universal constant $\hat{b} < 2\pi$ is defined such that $h(x)$ is analytic on a disc $|x| < \hat{b}$.

Theorem 2.6 (Fomenko and Chakon) Let $\delta = \hat{b}/M$, where $M \geq 1$ is a constant such that $\|[\eta, \zeta]\| \leq M\|\eta\|\|\zeta\|$, $\forall \eta, \zeta \in \mathcal{G}$. Suppose that $U(t)$ is a piecewise continuous curve in $\mathcal{G}$ and $\int_0^t \|\epsilon U(\tau)\| d\tau < \delta$. Then $Z(t) = \sum_{i=1}^{\infty} \epsilon^i Z_i(t)$ is a

convergent series. The terms $Z_i(t)$ are uniquely defined by

$$\begin{aligned}
Z_1(t) = T_0(t) &= \int_0^t U(\tau) d\tau, \\
(i+1)Z_{i+1}(t) &= T_i + \sum_{r=1}^i \left\{ \frac{1}{2} [Z_r, T_{i-r}] \right. \\
&\quad \left. + \sum_{p \geq 1, 2p \leq r} k_{2p} \sum_{\sum_{j=1}^{2p} m_j = r, m_j > 0} [Z_{m_1}, [Z_{m_2}, \dots, [Z_{m_{2p}}, T_{i-r}] \dots]] \right\}, \\
T_k(t) &= \int_0^t [U(\tau_1), \int_0^{\tau_1} [U(\tau_2), \dots, \int_0^{\tau_k} U(\tau_{k+1}) d\tau_{k+1}] d\tau_k] \dots d\tau_1.
\end{aligned} \tag{2.43}$$

We note that each term Z_i is composed of depth- $(i-1)$ brackets. The expression for $Z(t)$ as the logarithm of the solution to the *left-invariant* system (2.9) can be computed directly from the results of Theorem 2.6.

Corollary 2.7 Consider the *left-invariant* system (2.9) with $g(0) = e$. For the same assumptions on $U(t)$ as in Theorem 2.6, $g(t) = e^{Z(t)}$, $Z(t) \in \mathcal{G}$ and

$$Z(t) = \sum_{i=1}^{\infty} (-1)^{i+1} \epsilon^i Z_i(t), \tag{2.44}$$

where $Z_i(t)$ are defined by (2.43).

Proof. By (2.8) of Section 2.1.2, (2.9) is equivalent to

$$\frac{d}{dt}(g^{-1}) = \epsilon T_e R_{g^{-1}}(-U(t)).$$

By Theorem 2.6 $g^{-1}(t) = e^{Y(t)}$, $Y(t) = \sum_{i=1}^{\infty} \epsilon^i Y_i(t)$ where

$$Y_i(t) = \begin{cases} -Z_i(t) & i \text{ is odd} \\ Z_i(t) & i \text{ is even} \end{cases}$$

This follows since $Z_i(-U(t)) = -Z_i(U(t))$, for i odd and $Z_i(-U(t)) = Z_i(U(t))$ for i even. Thus, $g(t) = e^{-Y(t)}$ and the corollary follows. $\square$

We show the first few terms of the series:

$$\begin{aligned}
Z_1(t) &= \tilde{U}(t) \\
Z_2(t) &= \frac{1}{2} \int_0^t [U(\tau), \tilde{U}(\tau)] d\tau \\
Z_3(t) &= \frac{1}{3} \int_0^t [U(\tau), \int_0^\tau [U(\sigma), \tilde{U}(\sigma)] d\sigma] d\tau + \frac{1}{12} [\tilde{U}(t), \int_0^t [U(\tau), \tilde{U}(\tau)] d\tau].
\end{aligned} \tag{2.45}$$

Since $Z(t) \in \mathcal{G}$ we can write $Z(t) = \sum_{k=1}^n z_k(t) \xi_k$ where $(\xi_1, \dots, \xi_n)$ is a basis for $\mathcal{G}$. $(z_1, \dots, z_n)$ are referred to as *canonical coordinates of the first kind* for G . Let $\hat{S}$ be the largest neighborhood of $0 \in \mathcal{G}$ such that $\hat{\Psi} = \exp|_{\hat{S}} : \hat{S} \rightarrow G$ is injective. Let $\hat{Q} = \hat{\Psi}(\hat{S}) \subset G$. Then $\hat{\Psi} : \hat{S} \rightarrow \hat{Q}$ is a diffeomorphism and we can define a metric $\hat{d} : \hat{Q} \times \hat{Q} \rightarrow \mathfrak{R}_+$ by

$$\hat{d}(X, Y) = d(\hat{\Psi}^{-1}(X), \hat{\Psi}^{-1}(Y)) \tag{2.46}$$

where d is given by (2.39).

Lazard and Tits showed for a class of Lie groups that the $\exp$ map is injective on a reasonably large domain $\hat{S}$ [41]. Following their work we define an *admissible norm* $\|\cdot\|$ on $\mathcal{G}$ as any norm $\|\cdot\|$ that makes $(\mathcal{G}, \|\cdot\|)$ a Banach space and satisfies

$$\|[A, B]\| \leq \|A\| \|B\|, \quad \forall A, B \in \mathcal{G}.$$

Define $B(\mathcal{G}, \rho) = \{A \in \mathcal{G} \mid \|A\| < \rho\}$. Let C_{G_0} be the connected component of C_G , the center of G , that contains the identity $e \in G$.

Theorem 2.8 (Lazard and Tits) For any choice of admissible norm on $\mathcal{G}$, the restriction of $\exp$ to $\hat{S} = B(\mathcal{G}, \pi)$ is injective if and only if C_{G_0} is simply connected. If G is simply connected then the restriction of $\exp$ to $\hat{S} = B(\mathcal{G}, 2\pi)$ is injective.

Remark 2.9 C_{G_0} is simply connected for finite-dimensional, semi-simple Lie groups such as $SO(n)$, $n > 2$, and $SL(n)$, $n > 1$, since the center of such a group is the identity element which is trivially simply connected. $SE(n)$, $n > 1$, and the unipotent group, $n > 2$, also have a trivial center.

Consider a matrix Lie algebra $\mathcal{G} \subseteq \mathfrak{R}^{n \times n}$ and the induced matrix p -norm $\|\cdot\|_p$ on $\mathfrak{R}^{n \times n}$. We can always construct an admissible norm as $\|\cdot\|_L \triangleq 2\|\cdot\|_p$, since

$$\|[A, B]\|_L = 2\|[A, B]\|_p = 2\|AB - BA\|_p \leq 4\|A\|_p\|B\|_p = \|A\|_L\|B\|_L.$$

Thus, for any G for which C_{G_0} is simply connected, we can let $\hat{S} = \{A \in \mathcal{G} \mid \|A\|_L < \pi\} = \{A \in \mathcal{G} \mid \|A\|_p < \pi/2\}$. Lazard and Tits give an example admissible norm for $so(3)$, $se(2)$, $sl(2)$. Given the appropriate choice of basis $\{A_1, A_2, A_3\}$ for $so(3)$, $se(2)$, $sl(2)$, let $A = \sum_{i=1}^3 a_i A_i$ and $a = (a_1, a_2, a_3)^T$. Then $\|A\|_K \triangleq \|a\|_2$ is an admissible norm. Thus, for $G = SO(3)$, $SE(2)$ or $SL(2)$ and $\mathcal{G}$ the associated Lie algebra, we can let $\hat{S} = \{A \in \mathcal{G} \mid \|A\|_K < \pi\}$.

We say that $\delta_1(\epsilon) \in \mathfrak{R}^n$ is of order $\delta_2(\epsilon)$, denoted $\delta_1(\epsilon) = O(\delta_2(\epsilon))$ if there exists some $k > 0$ and $\epsilon_1 > 0$ such that

$$\|\delta_1(\epsilon)\| \leq k\delta_2(\epsilon), \quad \forall |\epsilon| < \epsilon_1. \quad (2.47)$$

The following lemma shows that for G a matrix Lie group, if $X \in G$ is in a sufficiently small neighborhood of the identity of G then the canonical coordinates of the first and second kind relative to the same basis $\{A_1, \dots, A_n\}$ for the associated Lie algebra $\mathcal{G}$ are close of small order.

Lemma 2.10 Given $X \in Q \cap \hat{Q} \subset G$, let $\gamma = \Psi^{-1}(X)$ and $Z = \hat{\Psi}^{-1}(X)$. Then $\gamma = O(\epsilon^p)$ if and only if $Z = O(\epsilon^p)$, $p \geq 1$. In this case, $\|\gamma_i - z_i\| = O(\epsilon^{2p})$, $i = 1, \dots, n$.

Proof. Given X we expand exponentials in the two local representations to find that

$$\begin{aligned}
X &= e^{\gamma_1 A_1} \dots e^{\gamma_n A_n} \\
&= (I + \gamma_1 A_1 + \gamma_1^2 A_1^2/2 + \dots) \dots (I + \gamma_n A_n + \gamma_n^2 A_n^2/2 + \dots) \\
&= I + \sum_{i=1}^n \gamma_i A_i + \sum_{i=1}^n \gamma_i^2 A_i^2/2 + \sum_{i,j=1}^n \gamma_i \gamma_j A_i A_j + \dots, \tag{2.48}
\end{aligned}$$

and

$$X = e^Z = I + Z + Z^2/2 + \dots = I + \sum_{i=1}^n z_i A_i + (\sum_{i=1}^n z_i A_i)^2/2 + \dots \tag{2.49}$$

The lemma follows by equating (2.48) and (2.49). $\square$

Chapter 3

Examples

In this chapter we describe the motion control applications that we study throughout this dissertation. In particular, we derive the relevant state equations as systems on matrix Lie groups.

3.1 Spacecraft Attitude Control

As our first example, consider the spacecraft attitude control problem. Let $(\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3)$ be an orthonormal frame fixed on the body and let $(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3)$ define an inertial frame with origin coincident with the origin of the body-fixed frame. Then we define $X(t) \in SO(3)$ such that $\mathbf{r}_i = X(t)\mathbf{b}_i$, i.e., $X(t)$ determines the attitude of the spacecraft at time t . Let $e_1 = (1, 0, 0)^T$, $e_2 = (0, 1, 0)^T$ and $e_3 = (0, 0, 1)^T$. Define $\hat{\cdot}: \mathfrak{R}^3 \rightarrow so(3)$ for $x = (x_1, x_2, x_3)^T$, by

$$\hat{x} = \begin{bmatrix} 0 & -x_3 & x_2 \\ x_3 & 0 & -x_1 \\ -x_2 & x_1 & 0 \end{bmatrix}. \quad (3.1)$$

Define $A_i = \hat{e}_i$, $i = 1, 2, 3$. Then $\{A_1, A_2, A_3\}$ is the standard basis for $\mathcal{G} = so(3)$ and $X(t)$ satisfies

$$\dot{X} = X\hat{\Omega}, \quad \hat{\Omega}(t) = \sum_{i=1}^3 \Omega_i(t) A_i \quad (3.2)$$

where $\Omega = (\Omega_1, \Omega_2, \Omega_3)^T$ is the angular velocity of the spacecraft in body-fixed coordinates. If we let $\epsilon u_i = \Omega_i$, $i = 1, 2, 3$, i.e., if we interpret the components of angular velocity as our small-amplitude, periodic controls (see Chapter 1 and Chapter 2), then (3.2) takes the form (2.11) with $G = SO(3)$ and $n = 3$.

We will be most interested in the case when only two components of angular velocity can be controlled, i.e., when $m = 2$. For example, if we can control angular velocity about the $\mathbf{b}_1$ and $\mathbf{b}_2$ axes, then X satisfies

$$\dot{X} = \epsilon X(u_1 A_1 + u_2 A_2). \quad (3.3)$$

By Theorem 2.1 (Brockett/Jurdjevic and Sussmann), (3.3) is controllable and is a single-bracket system since $[A_1, A_2] = A_3$. This means that we can reorient the spacecraft as desired by controlling only two of three angular velocity components (e.g., roll and pitch velocities).

From Section 2.3.1, the solution to (3.3) can locally be expressed as $X(t) = e^{\gamma_1(t)A_1} e^{\gamma_2(t)A_2} e^{\gamma_3(t)A_3}$. The Wei-Norman parameters $(\gamma_1, \gamma_2, \gamma_3)$ satisfy (2.31) which for the standard basis of $so(3)$ defined above becomes

$$\begin{bmatrix} \dot{\gamma}_1 \\ \dot{\gamma}_2 \\ \dot{\gamma}_3 \end{bmatrix} = \epsilon \begin{bmatrix} \sec \gamma_2 \cos \gamma_3 & -\sec \gamma_2 \sin \gamma_3 & 0 \\ \sin \gamma_3 & \cos \gamma_3 & 0 \\ -\tan \gamma_2 \cos \gamma_3 & \tan \gamma_2 \sin \gamma_3 & 1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \epsilon M(\gamma)u, \quad (3.4)$$

where for the case of two controls (3.3) we set $u_3 = 0$. The matrix $M(\gamma)$ is

nonsingular for $|\gamma_2| < \pi/2$, and the Wei-Norman parameters can be interpreted as Euler angles. Let $B = (\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3)$ and $E = (\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3)$ and define the *rotations* $\gamma_1, \gamma_2, \gamma_3$ as follows:

- Rotate E counterclockwise about the $\mathbf{r}_1$ -axis by angle γ_1 to get E' .
- Rotate E' counterclockwise about its y -axis by angle γ_2 to get E'' .
- Rotate E'' counterclockwise about its z -axis by angle γ_3 to get B .

From Section 2.3.2 we can also express the solution to (3.3) locally as a single exponential $X(t) = e^{Z(t)}$, $Z(t) = \sum_{i=1}^3 z_i(t)A_i$. This representation is equivalent to the well-known Euler parametrization for $SO(3)$. Given $X = e^Z \in SO(3)$, $Z = -\phi\hat{c}$ where

$$\phi = \phi(X) = \cos^{-1}_{[0,\pi]}(1/2(\text{tr}(X) - 1)), \quad (3.5)$$

tr is the trace operator and

$$c = c(X) = \frac{1}{2\sin\phi}(X_{23} - X_{32}, X_{31} - X_{13}, X_{12} - X_{21})^T.$$

As an alternative to the metrics $\tilde{d}$ and $\hat{d}$ defined by (2.38) and (2.46), respectively, on a neighborhood of the identity of $SO(3)$, we define a metric $\bar{\phi}$ on $SO(3)$ which gives a more physical measure of rotational distance. Let $X, \bar{X} \in SO(3)$. Following Meyer [55] we let $R = \bar{X}^{-1}X = \bar{X}^T X$ and let $\phi(R)$ be given by (3.5). $R \in SO(3)$ describes the rotation that takes $\bar{X}$ into X and $\phi(R)$ defines the magnitude of this rotation. By [55],

$$\bar{\phi} : SO(3) \times SO(3) \rightarrow \mathfrak{R}_+, \quad \bar{\phi}(A, B) \triangleq \phi(AB^T) \quad (3.6)$$

defines a metric on $SO(3)$ and satisfies

$$\phi(AB^T) \leq \phi(A) + \phi(B). \quad (3.7)$$

So $\bar{\phi}(\bar{X}^T, X^T) = \phi(R)$ describes the distance between X and $\bar{X}$ on $SO(3)$. Further, suppose $X(t)$ and $\bar{X}(t)$ are curves in $SO(3)$ such that $\dot{X} = X\hat{u}$, $\dot{\bar{X}} = \bar{X}\hat{\bar{u}}$. Then $R(t) = \bar{X}^T(t)X(t)$ also is a curve in $SO(3)$ and

$$\dot{R} = R\hat{\Omega}_d, \quad \Omega_d = u - R^T\bar{u}.$$

Ω_d is the difference in body angular velocity between the system defined by X and the system defined by $\bar{X}$. Let $R(0) = I$, $R(t_f) = R_f$, $t_f > 0$. Meyer showed that for any $\Omega_d(t)$,

$$\phi(R_f) \leq \int_0^{t_f} \|\Omega_d(t)\| dt.$$

Thus, $\bar{\phi}(\bar{X}^T, X^T) = \phi(R)$ can be interpreted as the minimal angular distance between the body coordinates of the system defined by X and the body coordinates of the system defined by $\bar{X}$.

The interpretation of the components of spacecraft angular velocity Ω as our control inputs ϵu is reasonable, for example, if we suppose that zero angular momentum of the spacecraft is conserved, i.e., that the spacecraft is not spinning and there is no external torque applied. We describe below two different ways of effecting control over angular velocity under these circumstances.

3.1.1 Internal Rotors

The spacecraft with internal rotors is illustrated in Figure 3.1. Let p be the number of rotors. Assume that the i th rotor spins about the axis d_i (unit vector) which is fixed in the spacecraft such that the center of mass of the i th rotor lies on the d_i axis. Further, d_i is a principal axis for the i th rotor, and the i th rotor is symmetric about d_i . Let μ be the angular momentum of the spacecraft measured in inertial coordinates and assume that $\mu = 0$ is conserved. Define

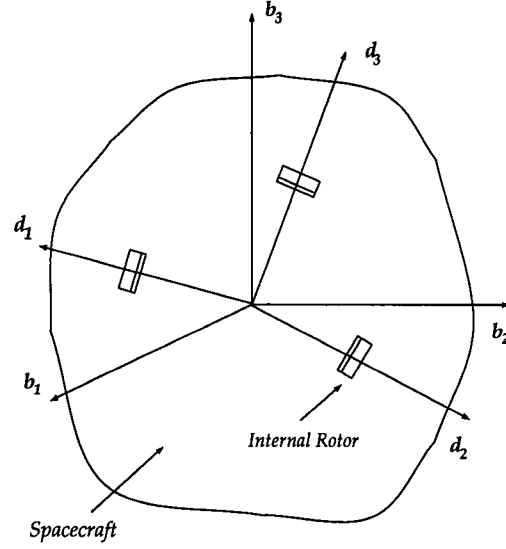


Figure 3.1: Spacecraft with Internal Rotors.

$J^* \triangleq$ inertia matrix of spacecraft without rotors in body-fixed coordinates.

$\bar{J}_i \triangleq$ inertia matrix of i th rotor measured in body-fixed coordinates.

$\nu_i \triangleq \dot{\theta}_i d_i =$ angular velocity of i th rotor in body-fixed coordinates.

$\mu_b \triangleq$ angular momentum of body in body-fixed coordinates.

Then by [55, 18]

$$\mu_b = X^T \mu = \sum_{i=1}^p \bar{J}_i (\Omega + \nu_i) + J^* \Omega$$

which implies that

$$\mu = X((J^* + \sum_{i=1}^p \bar{J}_i) \Omega + \sum_{i=1}^p \bar{J}_i d_i \dot{\theta}_i).$$

Then since $\mu = 0$,

$$\Omega = -J^{-1} \sum_{i=1}^p \bar{J}_i d_i \dot{\theta}_i,$$

where $J \triangleq J^* + \sum_{i=1}^p \bar{J}_i$ is assumed to be nonsingular (which it will be in general).

Thus, assuming that we can control the angular velocities of the internal rotors,

$\dot{\theta}_i$, $i = 1, \dots, p$, we can control some or all of the components of Ω .

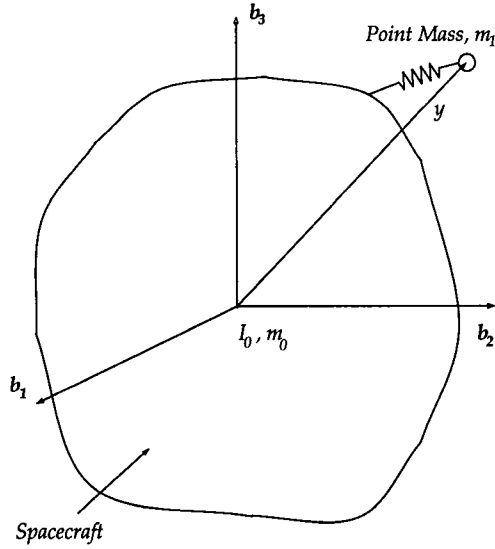


Figure 3.2: Spacecraft with Appended Point Mass Oscillator.

3.1.2 Appended Point Mass Oscillator

The spacecraft with an appended point mass oscillator is illustrated in Figure 3.2. We assume that we can control the oscillation of the point mass in three dimensions. (Alternatively, the problem may involve undesirable oscillations of an appended point mass and, thus, we study the resulting attitude drift of the spacecraft.) We assume that the point mass has no associated inertia (matrix). Let μ be the angular momentum of the spacecraft plus oscillator measured in inertial coordinates and assume that $\mu = 0$ is conserved. Define

$y \triangleq$ distance of the point mass from the spacecraft center of mass in body-fixed coordinates.

$m_0, m_1 \triangleq$ mass of the spacecraft and point mass oscillator, respectively.

$\alpha_m \triangleq m_0 m_1 / (m_0 + m_1)$.

$\mathcal{I}_0 \triangleq$ inertia matrix of the spacecraft without the oscillator.

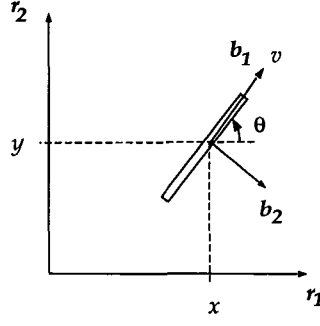


Figure 3.3: Overhead View of Unicycle.

$\mathcal{I}_{lock} \triangleq$ instantaneous locked inertia of body plus oscillator, i.e., the inertia of the system frozen at y in body coordinates. $\mathcal{I}_{lock} = \mathcal{I}_0 + \alpha_m(\|y\|^2 I - yy^T)$ where I is the 3×3 identity matrix.

$\mu_b \triangleq$ angular momentum of body in body-fixed coordinates.

$\dot{y} \triangleq$ velocity of point mass in body-fixed coordinates.

Then by [36]

$$\mu_b = X^T \mu = (\mathcal{I}_{lock} \Omega + \alpha_m y \times \dot{y}).$$

Since $\mu = 0$, we get that

$$\Omega = -\mathcal{I}_{lock}^{-1}(y) \alpha_m (y \times \dot{y}), \quad (3.8)$$

assuming that $\mathcal{I}_{lock}(y)$ is nonsingular (which it will be in general). Thus, Ω is driven by the velocity $\dot{y}$ of the oscillator.

3.2 Unicycle Motion Control

As our second example, we consider the unicycle motion planning problem shown in Figure 3.3. Let the configuration of the unicycle be $(x, y, \theta) \in \mathbb{R}^2 \times S^1$. (x, y)

describes the unicycle's position on a plane relative to an orthonormal inertial frame $(\mathbf{r}_1, \mathbf{r}_2)$. θ describes the orientation of the unicycle, specifically, the angle between the tangent to the wheel on the plane and the $\mathbf{r}_1$ -axis. Assuming the unicycle wheel rolls without slipping and $u_1 = \dot{\theta}$ (steering speed) and $u_2 = v$ (rolling speed) are available as controls, then the system can be described by

$$\begin{aligned}\dot{x} &= u_2 \cos \theta \\ \dot{y} &= u_2 \sin \theta \\ \dot{\theta} &= u_1.\end{aligned}\tag{3.9}$$

Alternatively, let $(\mathbf{b}_1, \mathbf{b}_2)$ be an orthonormal frame in the plane fixed on the vehicle such that $\mathbf{b}_1$ is parallel to the tangent to the wheel on the plane. Define $X(t) \in SE(2)$ by

$$\begin{bmatrix} \mathbf{r}_i \\ 1 \end{bmatrix} = X(t) \begin{bmatrix} \mathbf{b}_i \\ 1 \end{bmatrix}, \quad i = 1, 2.$$

Then $X(t)$ describes the position and orientation of the unicycle at time t and

$$X = \left[\begin{array}{cc|c} \cos \theta & -\sin \theta & x \\ \sin \theta & \cos \theta & y \\ \hline 0 & 0 & 1 \end{array} \right].$$

Let

$$A_1 = \left[\begin{array}{cc|c} 0 & -1 & 0 \\ 1 & 0 & 0 \\ \hline 0 & 0 & 0 \end{array} \right], \quad A_2 = \left[\begin{array}{cc|c} 0 & 0 & 1 \\ 0 & 0 & 0 \\ \hline 0 & 0 & 0 \end{array} \right], \quad A_3 = \left[\begin{array}{cc|c} 0 & 0 & 0 \\ 0 & 0 & 1 \\ \hline 0 & 0 & 0 \end{array} \right]. \tag{3.10}$$

Then $\{A_1, A_2, A_3\}$ defines a basis for $se(2)$, the Euclidean Lie algebra, and $X(t)$ satisfies

$$\dot{X} = \epsilon X(A_1 u_1 + A_2 u_2) \tag{3.11}$$

where we have assumed small-amplitude controls. Equation (3.11) is of the form (2.11) with $G = SE(2)$, $n = 3$ and $m = 2$. By Theorem 2.1, (3.11) is controllable and is a single-bracket system since $[A_1, A_2] = A_3$.

Because $G = SE(2)$ is solvable, Lemma 2.3 ensures that at least one choice of ordered basis for $se(2)$ yields a *global* product of exponentials representation of $X(t)$. It turns out that more than one choice of ordered basis gives a global solution. For the representation $X(t) = e^{\gamma_1(t)A_1}e^{\gamma_2(t)A_2}e^{\gamma_3(t)A_3}$, the Wei-Norman parameters $(\gamma_1, \gamma_2, \gamma_3)$ globally satisfy

$$\begin{aligned}\dot{\gamma}_1 &= u_1 \\ \dot{\gamma}_2 &= \gamma_3 u_1 + u_2 \\ \dot{\gamma}_3 &= -\gamma_2 u_1 + u_3.\end{aligned}$$

u_3 is the velocity in the $\mathbf{b}_2$ direction, i.e., along the wheel axis. For the unicycle problem, u_3 is the slipping speed; thus, the rolling without slipping constraint implies that $u_3 = 0$. Further, one can see that $\theta = \gamma_1$ and

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} \gamma_2 \\ \gamma_3 \end{bmatrix},$$

i.e., (γ_2, γ_3) describes the position of the unicycle with respect to a coordinate frame that rotates with the unicycle but has its origin fixed at the origin of the inertial frame. Coordinates x, y, θ are also Wei-Norman parameters and (3.9) Wei-Norman equations corresponding to the global solution $X(t) = e^{x(t)A_2}e^{y(t)A_3}e^{\theta(t)A_1}$. The parameters z_1, z_2, z_3 in the representation $X(t) = e^{Z(t)}$, $Z(t) = \sum_{i=1}^3 z_i(t)A_i$ are identical to the Wei-Norman parameters $\gamma_1, \gamma_2, \gamma_3$, i.e., $z_i(t) = \gamma_i(t)$, $i = 1, 2, 3$.

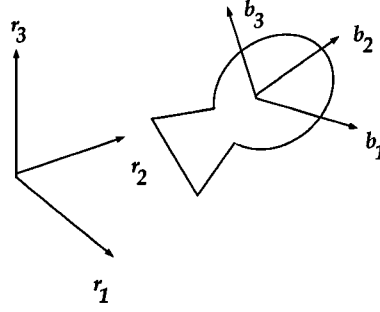


Figure 3.4: Autonomous Underwater Vehicle.

3.3 Autonomous Underwater Vehicle Motion Control

Consider an autonomous underwater vehicle (AUV) as in Figure 3.4 and let $(\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3)$ be an orthonormal frame fixed on the vehicle. Similarly, let $(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3)$ define an inertial frame. Then we define $X(t) \in SE(3)$ where

$$X(t) = \left[\begin{array}{c|c} X_R(t) & x_T(t) \\ \hline 0 & 1 \end{array} \right], \quad X_R(t) \in SO(3), \quad x_T(t) \in \mathbb{R}^3,$$

such that

$$X(t) \begin{bmatrix} \mathbf{b}_i \\ 1 \end{bmatrix} = \left[\begin{array}{c|c} X_R(t) & x_T(t) \\ \hline 0 & 1 \end{array} \right] \begin{bmatrix} \mathbf{b}_i \\ 1 \end{bmatrix} = \begin{bmatrix} X_R(t)\mathbf{b}_i + x_T(t) \\ 1 \end{bmatrix} = \begin{bmatrix} \mathbf{r}_i \\ 1 \end{bmatrix}.$$

Thus, $X(t)$ describes the orientation and position of the AUV at time t . Let

$$A_i = \begin{cases} \left[\begin{array}{c|c} \hat{e}_i & 0 \\ \hline 0 & 0 \end{array} \right] & i = 1, 2, 3 \\ \left[\begin{array}{c|c} 0 & e_{i-3} \\ \hline 0 & 0 \end{array} \right] & i = 4, 5, 6. \end{cases} \quad (3.12)$$

Then $\{A_1, \dots, A_6\}$ defines a basis for $\mathcal{G} = se(3)$. Let $\Omega = (\Omega_1, \Omega_2, \Omega_3)^T$ define the angular velocity of the vehicle and $v = (v_1, v_2, v_3)^T$ the vehicle translational

velocity, all with respect to $(\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3)$. Then $X(t)$ satisfies

$$\dot{X} = X \left(\sum_{i=1}^3 \Omega_i(t) A_i + \sum_{i=4}^6 v_{i-3}(t) A_i \right). \quad (3.13)$$

We assume that we can interpret $\Omega(t)$ and $v(t)$ as small-amplitude controls such that (3.13) is of the form (2.11), e.g., we let $\epsilon u_i = \Omega_i$, $i = 1, 2, 3$ and $\epsilon u_i = v_{i-3}$, $i = 4, 5, 6$. In this case $G = SE(3)$ and $n = 6$.

As in the spacecraft attitude control problem, we are interested here in the case when fewer than n control components are available, i.e., $m < 6$. For example, suppose that we can control angular velocity about $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3$ and translational velocity along $\mathbf{b}_1$. Then $X(t)$ satisfies

$$\dot{X} = \epsilon X \left(\sum_{i=1}^4 u_i(t) A_i \right). \quad (3.14)$$

By Theorem 2.1, (3.14) is controllable and is a single-bracket system since $[A_3, A_4] = A_5$ and $[A_4, A_2] = A_6$. Now suppose that an actuator which controls one component of angular velocity fails. If the failed actuator corresponds to the first component, i.e., $u_1(t) = 0$, then (3.14) is still a single-bracket system. However, if the actuator for one of the other components fails, say $u_3(t) = 0$, then X satisfies

$$\dot{X} = \epsilon X (u_1 A_1 + u_2 A_2 + u_4 A_4). \quad (3.15)$$

System (3.15) is still controllable but now a double-bracket system. This can be seen since $[A_1, A_2] = A_3$, $[A_4, A_2] = A_6$ and $[[A_1, A_2], A_4] = A_5$.

The AUV is controllable even with as few as two controls. Suppose one actuator controls angular velocity (denoted by ϵu_1) about an axis that lies in the $\mathbf{b}_1$ - $\mathbf{b}_2$ plane. For example, suppose that it controls angular velocity about the $\mathbf{b}_1$ and $\mathbf{b}_2$ axes equally. Suppose the other actuator controls angular and

translational velocity about the $\mathbf{b}_2$ axis, simultaneously. This second actuator provides a screw-like velocity that we denote by $\epsilon u_2(t)$. Define $B_1 = A_1 + A_2$ and $B_2 = A_2 + A_5$. Then X satisfies

$$\dot{X} = \epsilon X(u_1 B_1 + u_2 B_2). \quad (3.16)$$

System (3.16) is controllable and is a triple-bracket system since $(B_1, B_2, [B_1, B_2], [B_1, [B_1, B_2]], [B_2, [B_1, B_2]], [B_1, [B_1, [B_1, B_2]]])$ span $se(3)$.

The Wei-Norman parameters $(\gamma_1, \dots, \gamma_6)$ in the product of exponentials representation $X(t) = \prod_{i=1}^6 e^{\gamma_i(t) A_i}$ satisfy

$$\begin{bmatrix} \dot{\gamma}_1 \\ \dot{\gamma}_2 \\ \dot{\gamma}_3 \\ \dot{\gamma}_4 \\ \dot{\gamma}_5 \\ \dot{\gamma}_6 \end{bmatrix} = \begin{bmatrix} \sec \gamma_2 \cos \gamma_3 & -\sec \gamma_2 \sin \gamma_3 & 0 & 0 & 0 & 0 \\ \sin \gamma_3 & \cos \gamma_3 & 0 & 0 & 0 & 0 \\ -\tan \gamma_2 \cos \gamma_3 & \tan \gamma_2 \sin \gamma_3 & 1 & 0 & 0 & 0 \\ 0 & -\gamma_6 & \gamma_5 & 1 & 0 & 0 \\ \gamma_6 & 0 & -\gamma_4 & 0 & 1 & 0 \\ -\gamma_5 & \gamma_4 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \\ u_5 \\ u_6 \end{bmatrix}. \quad (3.17)$$

The parameters $\gamma_1, \gamma_2, \gamma_3$ correspond to the three Euler-angle type parameters of the spacecraft attitude control problem (compare with (3.4)) and parametrize the orientation of the vehicle. The parameters $\gamma_4, \gamma_5, \gamma_6$ parametrize the position of the vehicle. Let $X = \left(\begin{array}{c|c} X_R & x_T \\ \hline 0 & 0 & 0 & 1 \end{array} \right)$. Then $X_R = \prod_{i=1}^3 e^{\gamma_i A_i}$ and $x_T = X_R(\gamma_4, \gamma_5, \gamma_6)^T$. Similarly, consider the single exponential representation $X(t) = e^{Z(t)}$, $Z(t) = \sum_{i=1}^6 z_i(t) A_i$. z_1, z_2, z_3 correspond to the single exponential parameters of the spacecraft problem, i.e., $X_R = e^{\sum_{i=1}^3 z_i(t) A_i}$, and $z_i = \gamma_i$, $i = 4, 5, 6$.

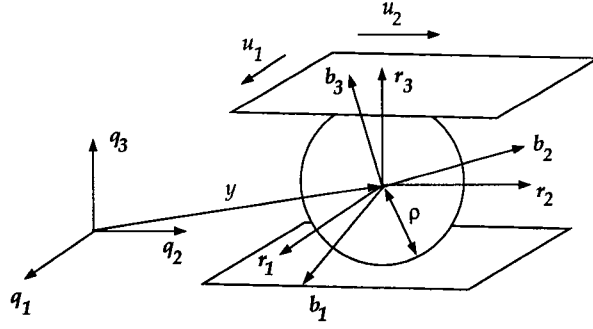


Figure 3.5: Uniform Ball Rolling Between Two Parallel Plates.

3.4 Ball and Plate

Consider a uniform spherical ball rolling without slipping between two parallel, flat, rough plates as in Figure 3.5. Let $(\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3)$ be an orthonormal frame fixed on the ball with origin at the center of the ball. Let $(\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3)$ be an inertial frame. Let $(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3)$ be a frame translating with the ball that has its origin at the center of the ball, but has its orientation coincident with the inertial frame. Let $\mathbf{r}_3$ be perpendicular to the plates. Assume that the bottom plate is stationary and the top plate can be vibrated parallel to the $\mathbf{r}_1$ - $\mathbf{r}_2$ plane. Further, assume that we can control the velocity of the top plate. Let $u_i(t)$ be the velocity component in the $\mathbf{r}_i$ -direction, $i = 1, 2$. Let $y = (y_1, y_2, y_3)^T$ describe the position of the center of the ball with respect to $(\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3)$. Let ρ be the radius of the ball, then $y_3(t) = \rho$. Since the velocity of the center of the ball is equal to the velocity at the top contact point,

$$\dot{y}_1 = u_1(t), \quad \dot{y}_2 = u_2(t), \quad \dot{y}_3 = 0.$$

Now let $X(t) \in SO(3)$ be defined such that $\mathbf{r}_i = X^T \mathbf{b}_i$, i.e., $X(t)$ describes the orientation of the ball at time t . Let Ω be the angular velocity of the ball

expressed with respect to $(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3)$. The rolling without slipping constraint implies that

$$\begin{pmatrix} u_1 \\ u_2 \\ 0 \end{pmatrix} = \Omega \times \begin{pmatrix} 0 \\ 0 \\ \rho \end{pmatrix} = \begin{pmatrix} \rho\Omega_2 \\ -\rho\Omega_1 \\ 0 \end{pmatrix}.$$

Further, assuming zero initial angular momentum about the $\mathbf{r}_3$ axis, by conservation of angular momentum, $\Omega_3 = 0$. So $X(t)$ satisfies

$$\dot{X} = X\hat{\Omega} = \frac{1}{\rho}X(-A_1u_2(t) + A_2u_1(t)), \quad (3.18)$$

where $A_1, A_2 \in so(3)$ are as defined in Section 3.1. System (3.18) is a left-invariant control system on $SO(3)$ of the same form as the spacecraft attitude control problem with only two velocity components available as controls. In particular, this system is controllable and a single-bracket system.

This system could potentially be used in the design of a vibratory actuator [33]. Suppose that we control small-amplitude vibrations of the top plate, e.g., with piezo-actuated flexural members. Since (3.18) is controllable, we can get net rotation of the ball about the $\mathbf{r}_3$ axis. Thus, if we attach a (deformable) gear drive about the equator of the ball, we can convert vibratory motion into rotary motion (see Figure 3.6).

3.5 Brockett's System

Brockett's prototypical nonholonomic system [13] on $\mathfrak{R}^3$ is defined (after a change of variables) as

$$\dot{x} = u_1$$

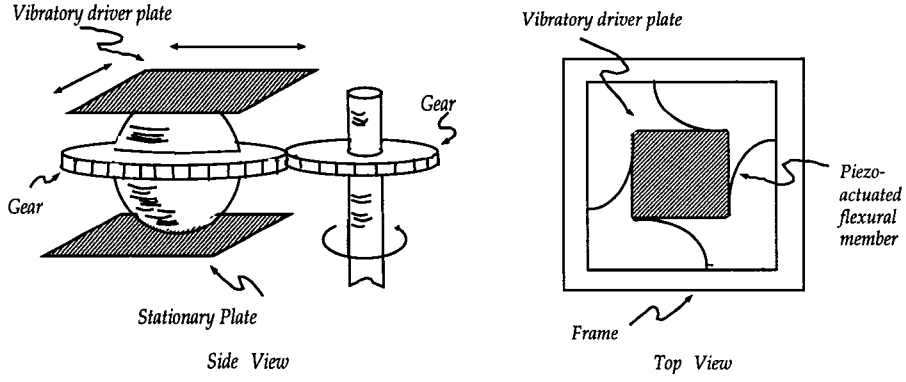


Figure 3.6: Vibratory Actuator on $SO(3)$.

$$\dot{y} = u_2 \quad (3.19)$$

$$\dot{z} = -u_1 y.$$

In the terminology used by Murray and Sastry [60], (3.19) is in *chained form*. However, we can also express this system as a drift-free, left-invariant system on the three-dimensional unipotent matrix Lie group (also called the Heisenberg group) denoted by $H(3)$. Define a basis for the associated (nilpotent of order 2) Lie algebra $\mathfrak{h}(3)$ as

$$A_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

and let

$$X = \begin{pmatrix} 1 & y & -z \\ 0 & 1 & x \\ 0 & 0 & 1 \end{pmatrix}.$$

Then $X(t)$ satisfies

$$\dot{X} = \epsilon X (A_1 u_1(t) + A_2 u_2(t)) \quad (3.20)$$

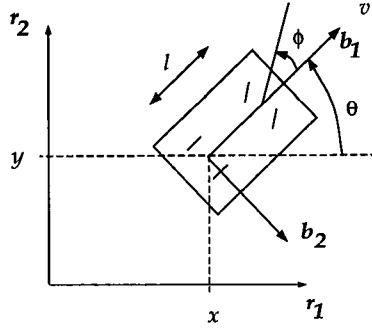


Figure 3.7: Overhead View of Front-Wheel Drive Car.

which is of the form (2.11) with $n = 3$, $m = 2$ where we have assumed small-amplitude controls. By Theorem 2.1, (3.20) is controllable and a single-bracket system since $[A_1, A_2] = A_3$. Since $h(3)$ is nilpotent it is also solvable. Thus, Lemma 2.3 ensures that at least one choice of basis gives a global product of exponentials solution for $X(t)$. The product of exponentials $X(t) = e^{x(t)A_1} e^{y(t)A_2} e^{z(t)A_3}$ is a global solution to (3.20).

3.6 Front-Wheel Drive Car and Other Chained-Form Systems

Consider the front-wheel drive (kinematic) car of length l shown in Figure 3.7. The front-wheel pair and the rear-wheel pair are each modelled as a single wheel located at the midpoint of each axle. We assume that only the front wheels are allowed to turn. The car, like the unicycle, is a nonholonomic system if we assume that the wheels do not slip. The configuration of the unicycle can be described by $(x, y, \theta, \phi) \in \mathbb{R}^2 \times S^1 \times S^1$, where (x, y) describes the car's position on a plane relative to an inertial frame $(\mathbf{r}_1, \mathbf{r}_2)$. $(\mathbf{b}_1, \mathbf{b}_2)$ is an orthonormal frame fixed on the car as shown in Figure 3.7. θ denotes the orientation of the car, i.e., the angle

between the $\mathbf{b}_1$ axis of the car and the $\mathbf{r}_1$ axis, and ϕ denotes the steering angle, i.e., the angle between the $\mathbf{b}_1$ axis of the car and the front wheels. Assuming that we can control $u_1 = \dot{\phi}$ (steering speed) and $u_2 = v$ (rolling speed), then the kinematic state equations are

$$\begin{aligned}\dot{x} &= u_2 \cos \theta \\ \dot{y} &= u_2 \sin \theta \\ \dot{\phi} &= u_1 \\ \dot{\theta} &= u_2 \frac{1}{l} \tan \phi.\end{aligned}\tag{3.21}$$

The configuration of the car is more naturally described by the Lie group $SE(2) \times S^1$. $SE(2)$ describes the position and orientation of the car (as in the unicycle problem) and S^1 describes the angular position of the front wheels. For example, let $X(t) \in SE(2) \times S^1 \approx SE(2) \times SO(2)$ be defined by

$$X = \begin{pmatrix} \cos \theta & -\sin \theta & x & 0 & 0 \\ \sin \theta & \cos \theta & y & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & \cos \phi & -\sin \phi \\ 0 & 0 & 0 & \sin \phi & \cos \phi \end{pmatrix}.$$

Define

$$A_1 = \left(\begin{array}{c|cc} 0 & & 0 \\ \hline & 0 & -1 \\ 0 & 1 & 0 \end{array} \right), \quad A_2 = \left(\begin{array}{ccc|c} 0 & 0 & 1 & \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \\ \hline & 0 & & 0 \end{array} \right), \quad A_3 = \left(\begin{array}{ccc|c} 0 & -1 & 0 & \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & \\ \hline & 0 & & 0 \end{array} \right).$$

Then $\{A_1, A_2, A_3, [A_3, A_2]\}$ is a basis for $se(2) \times so(2)$. Let $\tilde{u}_1(t) = \int_0^t u_1(\tau) d\tau$.

Then $X(t)$ satisfies

$$\dot{X} = X(u_1 A_1 + u_2 A_2 + \frac{1}{l} \tan(\tilde{u}_1) u_2 A_3),$$

which is a left-invariant system on $SE(2) \times S^1$, nonlinear in the controls u_1, u_2 .

An alternative way of describing the kinematics of the car is to convert (3.21) into chained form. This is a *local* transformation about $(x, y, \theta, \phi) = 0$ of the states and controls of (3.21) given by Murray and Sastry [60] to yield the following system on $\mathbb{R}^4$,

$$\begin{aligned} \dot{x}_1 &= v_1 \\ \dot{x}_2 &= v_2 \\ \dot{x}_3 &= x_2 v_1 \\ \dot{x}_4 &= x_3 v_1. \end{aligned} \tag{3.22}$$

Since this form is local it neglects the group structure of the configuration space. However, it has proved useful in the nonholonomic motion planning literature for constructing motion control laws.

System (3.22) can be expressed in our formalism as a left-invariant control system on a four-dimensional subgroup of the group of 4×4 unipotent matrices. This matrix Lie group consists of elements of the form

$$X = \begin{pmatrix} 1 & x_2 & x_3 & x_4 \\ 0 & 1 & x_1 & * \\ 0 & 0 & 1 & x_1 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

where $*$ is arbitrary. A basis for the (nilpotent) Lie algebra of this group is given by $\{A_1, A_2, A_3, A_4\} = \{A_1, A_2, [A_2, A_1], [[A_2, A_1], A_1]\}$ where

$$A_1 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Then $X(t)$ satisfies

$$\dot{X} = \epsilon X(A_1 v_1 + A_2 v_2) \quad (3.23)$$

which is of the form (2.11) with $n = 4$, $m = 2$, where we have assumed small-amplitude controls. The front-wheel drive car described by (3.23) is controllable and a double-bracket system. The product of exponentials $X(t) = \prod_{i=1}^4 e^{x_i(t)A_i}$ is a global solution to (3.23).

Other more general chained-form systems can equivalently be described as left-invariant control systems on unipotent matrix Lie groups. Accordingly, the theory and control synthesis algorithms of this dissertation can be applied to these systems. The connection is that chained-form systems are nilpotent systems in the analogous sense that unipotent matrix Lie groups have nilpotent Lie algebras. For example, consider the two-input chained-form system [59],

$$\begin{aligned} \dot{x}_1 &= u_1 \\ \dot{x}_2 &= u_2 \\ \dot{x}_3 &= x_2 u_1 \\ \dot{x}_4 &= x_3 u_1 \\ &\vdots \\ \dot{x}_k &= x_{k-1} u_1. \end{aligned} \quad (3.24)$$

It has been shown [71] that a kinematic car with $n = k - 3$ trailers can be converted into the form (3.24). System (3.24) can be expressed as a left-invariant control system on a k -dimensional subgroup of the $k \times k$ unipotent matrices, where an element X in this subgroup takes the form

$$X = \begin{pmatrix} 1 & x_2 & x_3 & x_4 & \dots & x_k \\ 0 & 1 & x_1 & * & \dots & * \\ 0 & 0 & 1 & x_1 & \ddots & \vdots \\ 0 & 0 & 0 & 1 & \ddots & * \\ \vdots & & & & \ddots & x_1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Chapter 4

Averaging

In this chapter we address the first main objective of this dissertation: to describe solutions (motion) of systems on Lie groups. To do so we derive averaging theory for left-invariant systems on finite-dimensional Lie groups. These systems are described by (2.9) of Section 2.1.3 on the Lie group G with Lie algebra $\mathcal{G}$. The goal of averaging theory in this context is to describe an approximate solution to (2.9) that remains close to the actual solution *and* that evolves on G . Classical averaging theory on $\mathbb{R}^n$ does not meet this goal because it is restricted to systems which evolve on $\mathbb{R}^n$. However, one can apply classical averaging theory to local solutions of (2.9). We derive averaging theory for systems on Lie groups by finding average solutions to the product of exponentials and single exponential local representations of (2.9) (see Section 2.3).

Averaging theory has a relatively long history. It is linked to the study of oscillations in classical mechanics which began with the description of the motion of a pendulum in the days of Galileo (1564-1642) and Newton (1642-1727). One can find the origins of the theory of small oscillations in the work of Lagrange (1736-1813) and an explicit statement of the principle of averaging in the work

of Gauss (1777-1855) in his study of the perturbations of planets on one another. Ideas related to averaging theory can also be found implicitly in the work of Laplace (1749-1827). Bogoliubov and Mitropolsky published a classic survey of averaging theory in the early 1960's [9]. A more modern treatment is given by Sanders and Verhulst [64]. We make reference to an even more recent treatment by Khalil [32].

In Section 4.1 we present our results based on the product of exponentials representation of the solution $g(t) \in G$ to (2.9). These results include an $O(\epsilon)$ and $O(\epsilon^2)$ approximation to $g(t)$ for $t \in [0, b/\epsilon]$, $b > 0$, where ϵ ($0 < \epsilon < 1$) is the small parameter representing the small amplitude of the periodic forcing function $U(t) \in \mathcal{G}$ (see the end of Section 2.3.2 for the definition of $O(\delta_2(\epsilon))$). In Section 4.2 we present our results based on the single exponential representation of $g(t)$. These results provide a more general $O(\epsilon^q)$, $q \geq 1$, approximation to $g(t)$, $t \in [0, b/\epsilon]$. Additionally, in this section smoothness assumptions on $U(t)$ are relaxed.

In both of these sections we emphasize the geometric interpretation of the average solution formulas. This interpretation is useful for solving the constructive controllability problem as will be shown in Chapter 5. The basic idea can be illustrated by considering the system on $\mathfrak{R}^3$ defined by

$$\dot{x} = F_1(x)u_1 + F_2(x)u_2, \quad x \in \mathfrak{R}^3, \quad (4.1)$$

where $F_1(\cdot)$ and $F_2(\cdot)$ are smooth vector fields on $\mathfrak{R}^3$. We assume that

$$\text{rank}([F_1(x) \ F_2(x) \ [F_1, F_2](x)]) = 3, \quad \forall x \in \mathfrak{R}^3, \quad (4.2)$$

where $[F_1, F_2](\cdot)$ is a new vector field on $\mathfrak{R}^n$ called the Lie bracket of F_1 and F_2

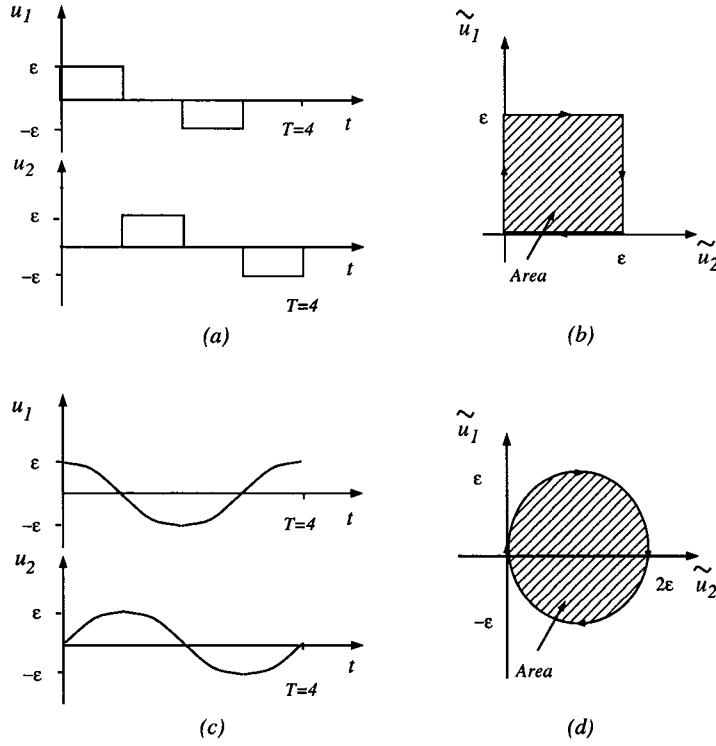


Figure 4.1: Periodic Controls.

and defined by

$$[F_1, F_2](x) = \frac{\partial F_2}{\partial x}(x)F_1(x) - \frac{\partial F_1}{\partial x}(x)F_2(x).$$

By a well-known theorem (e.g., Proposition 3.15 of [61]), analogous to Theorem 2.1 (Jurdjevic and Sussmann), assumption (4.2) implies that system (4.1) is controllable. Specifically, one can reach any point in $\mathfrak{R}^3$ by flowing along the vector fields F_1 , F_2 and $[F_1, F_2]$ which span $\mathfrak{R}^3$ at every point in $\mathfrak{R}^3$.

It is well known that one can generate flow in the direction of $[F_1, F_2]$ using the controls u_1 and u_2 specified in Figure 4.1(a). These controls yield

$$x(T) - x(0) = \epsilon^2 [F_1, F_2](x(0)) + \text{h.o.t.},$$

i.e., after $T = 4$ units of time the state of the system has changed in magnitude by

ϵ^2 in the direction of $[F_1, F_2]$ evaluated at the initial condition $x(0)$. Figure 4.1(b) shows a plot of $\tilde{u}_1 = \int_0^t u_1(\tau) d\tau$ versus $\tilde{u}_2 = \int_0^t u_2(\tau) d\tau$ during T units of time. It is clear that the magnitude of the resulting flow can be expressed as $\epsilon^2 = \text{Area}$ bounded by the curve described by $\tilde{u}_1$ versus $\tilde{u}_2$ during one period. As Brockett argues in [14], one would expect a similar result if the two controls were some other pair of small-amplitude periodic functions. For example, the controls as depicted in Figure 4.1(c) would produce

$$x(T) - x(0) \approx \text{Area}[F_1, F_2](x(0)),$$

where *Area* is the shaded area of Figure 4.1(d), i.e., the area bounded by the closed curve described by $\tilde{u}_1$ and $\tilde{u}_2$ over one period. Then in some *average* sense one would predict that

$$x(t) - x(0) \approx \frac{\text{Area} \cdot t}{T} [F_1, F_2](x(0)).$$

The theory derived in this chapter generalizes this area-rule result to systems on Lie groups and for arbitrarily high-order (in ϵ) approximations.

In Section 4.3 we derive the secular term (linear in t) in the $O(\epsilon^2)$ area rule from the curvature form of a certain principal fiber bundle with connection. This result emphasizes that our solutions are basis-independent. In Section 4.4 we present some stability results for system (2.9) that are derived from the averaging theory of this chapter.

4.1 Averaging with Product of Exponentials

In this section we derive first and second-order averaging theory for systems of the form (2.9). We use the product of exponentials local solution to (2.9)

$$g(t) = e^{\gamma_1(t)\xi_1} \dots e^{\gamma_n(t)\xi_n},$$

given by Lemma 2.3 (Wei and Norman), where

$$\dot{\gamma} = \epsilon M(\gamma)u. \quad (4.3)$$

The definitions of the neighborhoods $S \subset \mathfrak{R}^n$, $Q \subset G$, the diffeomorphism $\Psi : S \rightarrow Q$ and the metric $\tilde{d}$ on Q which we use in the following theorems can be found in Section 2.3.1. u_{av} , U_{av} , $\tilde{u}$ and $\tilde{U}$ are defined by (2.19)-(2.21).

Theorem 4.1 (First-Order Averaging) Consider system (2.9) on the Lie group G with Lie algebra $\mathcal{G}$. Assume that $U(t) \in \mathcal{G}$ is periodic in t with period T and has continuous derivatives up to second order for $t \in [0, \infty)$. Let $D = \{\gamma \in \mathfrak{R}^n \mid \|\gamma\| < r\} \subset S$ (for choice of r see condition on D below). Suppose that $g(0) = g_0 \in Q$. Let $\gamma(t)$ be the solution to (4.3) with $\gamma(0) = \gamma_0 = \Psi^{-1}(g_0)$. Let $\gamma^{(1)}(t) \in \mathfrak{R}^n$ and $g^{(1)}(t) \in G$ be defined by

$$\dot{\gamma}^{(1)} = \epsilon M(\gamma^{(1)})\left(\frac{1}{T} \int_0^T u(\tau) d\tau\right) = \epsilon M(\gamma^{(1)})u_{av}, \quad \gamma^{(1)}(0) = \gamma_0^{(1)}, \quad (4.4)$$

$$g^{(1)}(t) = e^{\gamma_1^{(1)}(t)\xi_1} \dots e^{\gamma_n^{(1)}(t)\xi_n} = g_0^{(1)} e^{\epsilon U_{av}t}, \quad g_0^{(1)} = \Psi(\gamma_0^{(1)}). \quad (4.5)$$

If $\|\gamma_0 - \gamma_0^{(1)}\| = O(\epsilon)$ and if $\gamma^{(1)}(t) \in D$, $\forall t \in [0, b/\epsilon]$ then,

$$\tilde{d}(g(t), g^{(1)}(t)) = O(\epsilon), \quad \forall t \in [0, b/\epsilon].$$

Proof: By classical averaging theory ([32, Theorem 7.4]),

$$\|\gamma(t) - \gamma^{(1)}(t)\| = O(\epsilon), \quad \forall t \in [0, b/\epsilon].$$

For small enough ϵ , since $\gamma^{(1)}(t) \in D \subset S$ then $\gamma(t) \in D \subset S$, $\forall t \in [0, b/\epsilon]$. So by Lemma 2.3,

$$g(t) = e^{\gamma_1(t)\xi_1} \dots e^{\gamma_n(t)\xi_n} = \Psi(\gamma(t)), \quad \forall t \in [0, b/\epsilon].$$

Define

$$g^{(1)}(t) = e^{\gamma_1^{(1)}(t)\xi_1} \dots e^{\gamma_n^{(1)}(t)\xi_n} = \Psi(\gamma^{(1)}(t)), \quad \forall t \in [0, b/\epsilon]. \quad (4.6)$$

Then, by definition of $\tilde{d}$,

$$\tilde{d}(g(t), g^{(1)}(t)) = \|\gamma(t) - \gamma^{(1)}(t)\| = O(\epsilon), \quad \forall t \in [0, b/\epsilon].$$

By Lemma 2.3, $g^{(1)}(t)$ as defined by (4.6) satisfies

$$\dot{g}^{(1)} = \epsilon T_e L_{g^{(1)}} U_{av}, \quad g^{(1)}(0) = g_0^{(1)} = \Psi(\gamma_0^{(1)}). \quad (4.7)$$

Since U_{av} is constant, (4.7) has the global solution

$$g^{(1)}(t) = g_0^{(1)} e^{\epsilon U_{av} t}.$$

□

Remark 4.2 It is clear from (4.7) that $g^{(1)}(t)$ describes the effect of the DC component of $U(t)$ on the system (2.9).

Theorem 4.3 (Second-Order Averaging - Area Rule) Consider system (2.9) on the Lie group G with Lie algebra $\mathcal{G}$. Assume that $U(t) \in \mathcal{G}$ is periodic in t with period T and has continuous derivatives up to third order for $t \in [0, \infty)$ and assume that $U_{av} = 0$. Let $D = \{\gamma \in \mathfrak{R}^n \mid \|\gamma\| < r\} \subset S$ (for choice of r see

condition on D below). Suppose that $g(0) = g_0 \in Q$. Let $\gamma(t)$ be the solution to (4.3) with $\gamma(0) = \gamma_0 = \Psi^{-1}(g_0) = O(\epsilon)$. Let $\gamma_0^{(2)} = (\gamma_{10}^{(2)}, \dots, \gamma_{n0}^{(2)})^T$ and define

$$\bar{w}_k(t) = \epsilon^2 \frac{t}{T} \sum_{i,j=1; i < j}^m Area_{ij}(T) \Gamma_{ij}^k + \gamma_{k0}^{(2)}, \quad (4.8)$$

$$\gamma_k^{(2)}(t) = \epsilon \tilde{u}_k(t) + \bar{w}_k(t), \quad (4.9)$$

$$g^{(2)}(t) = e^{\gamma_1^{(2)}(t)\xi_1} \dots e^{\gamma_n^{(2)}(t)\xi_n}, \quad (4.10)$$

where Γ_{ij}^k and $Area_{ij}(T)$ are defined by (2.15) and (2.22), respectively. If $\|\gamma_0 - \gamma_0^{(2)}\| = O(\epsilon^2)$ and if $\gamma^{(2)}(t) \in D$, $\forall t \in [0, b/\epsilon]$ then,

$$\tilde{d}(g(t), g^{(2)}(t)) = O(\epsilon^2), \quad \forall t \in [0, b/\epsilon].$$

Proof. Following classical averaging theory we define

$$h(t, y) = M(y)u(t),$$

$$v(t, y) = \int_0^t h(\tau, y) d\tau = \int_0^t M(y)u(\tau) d\tau = M(y)\tilde{u}(t),$$

and note that $v(t, y)$ is periodic in t with period T since $u_{av} = 0$. Consider the change of variables

$$\gamma = y + \epsilon v(t, y). \quad (4.11)$$

Note that $y(0) = \gamma_0 = O(\epsilon)$. Differentiating (4.11) with respect to time we get

$$\dot{\gamma} = \dot{y} + \epsilon \frac{\partial v}{\partial y} \dot{y} + \epsilon \frac{\partial v}{\partial t}.$$

Substituting (4.3) for $\dot{\gamma}$, (4.11) for γ and $\partial v / \partial t = M(y)u(t)$ gives

$$\epsilon M(y + \epsilon v)u = (I + \epsilon \frac{\partial v}{\partial y})\dot{y} + \epsilon M(y)u. \quad (4.12)$$

By (2.36) of Remark 2.4, $M(\gamma) = I + \tilde{\alpha}(\gamma) + O(\gamma^2)$ where the ij th element of $\tilde{\alpha}$ is defined by (2.35) and $O(\gamma^2)$ represents terms that are quadratic and higher order in the components of γ . Since, additionally, $M(\gamma)$ is analytic in γ , then

$$\begin{aligned} M(y + \epsilon v) &= I + \tilde{\alpha}(y + \epsilon v) + O((y + \epsilon v)^2) \\ &= I + \tilde{\alpha}(y) + \epsilon \tilde{\alpha}(v) + O(y^2) + \epsilon O(yv) + \epsilon^2 p_1(t, y, \epsilon) \\ &= M(y) + \epsilon \tilde{\alpha}(v) + \epsilon O(yv) + \epsilon^2 p_1(t, y, \epsilon) \end{aligned} \quad (4.13)$$

where $O(yv)$ is quadratic or higher order in the components of y and v . $p_1(t, y, \epsilon)$ is periodic in t with period T . Substituting (4.13) into (4.12) gives

$$\epsilon M(y)u + \epsilon^2 \tilde{\alpha}(v)u + \epsilon^2 O(yv)u + \epsilon^3 p_1(t, y, \epsilon)u = (I + \epsilon \frac{\partial v}{\partial y})\dot{y} + \epsilon M(y)u. \quad (4.14)$$

Since $\partial v / \partial y$ is t -periodic and thus bounded for all $(t, y) \in [0, \infty) \times D$, for small $\epsilon > 0$, $(I + \epsilon \partial v / \partial y)$ is nonsingular and

$$(I + \epsilon \frac{\partial v}{\partial y})^{-1} = I - \epsilon \frac{\partial v}{\partial y} + O(\epsilon^2).$$

Multiplying both sides of (4.14) by $(I + \epsilon \partial v / \partial y)^{-1}$ gives

$$\begin{aligned} \dot{y} &= (I - \epsilon \frac{\partial v}{\partial y} + O(\epsilon^2))(\epsilon^2 \tilde{\alpha}(v)u + \epsilon^2 O(yv)u + \epsilon^3 p_1(t, y, \epsilon)u) \\ &= \epsilon^2 \tilde{\alpha}(v)u(t) + \epsilon^2 O(yv)u(t) + \epsilon^3 p_2(t, y, \epsilon) \end{aligned}$$

where $p_2(t, y, \epsilon)$ is periodic in t with period T .

Now if we let $s = \epsilon t$, then

$$\frac{dy}{ds} = \epsilon \tilde{\alpha}(v)u(s) + \epsilon O(yv)u(s) + \epsilon^2 p_2(\frac{s}{\epsilon}, y, \epsilon) \triangleq F(s, y, \epsilon). \quad (4.15)$$

Define $y_0(s)$ by

$$\frac{dy_0}{ds} = F(s, y_0, 0) = 0, \quad y_0(0) = 0.$$

Then $y_0(s) = 0$, $\forall s > 0$. Define $y_1(s)$ by

$$\begin{aligned}
\frac{dy_1}{ds} &= \left(\frac{\partial F}{\partial y}(s, y_0, 0)y_1(s) + \frac{\partial F}{\partial \epsilon}(s, y_0, 0) \right) \Big|_{y_0=0} \\
&= (\tilde{\alpha}(v(s, y_0))u(s))|_{y_0=0} \\
&= (\tilde{\alpha}(M(y_0)\tilde{u}(s))u(s))|_{y_0=0} \\
&= \tilde{\alpha}(M(0)\tilde{u}(s))u(s) \\
&= \tilde{\alpha}(\tilde{u}(s))u(s), \quad \epsilon y_1(0) = \gamma_0.
\end{aligned}$$

By standard perturbation theory (c.f. [32, Theorem 7.1]), since $y_0(s) \in D$, $\forall s > 0$, $\exists \epsilon^* > 0$ such that $\forall |\epsilon| < \epsilon^*$ (4.15) has the unique solution $y(s)$ defined on $[0, b]$, $b > 0$, such that

$$\|y(s) - (y_0(s) + \epsilon y_1(s))\| = \|y(s) - \epsilon y_1(s)\| = O(\epsilon^2).$$

Now let $w \triangleq \epsilon y_1$ then

$$\|y(t) - w(t)\| = O(\epsilon^2), \quad \forall t \in [0, b/\epsilon], \quad (4.16)$$

$$\dot{w} = \epsilon^2 \tilde{\alpha}(\tilde{u}(t))u(t), \quad w(0) = \gamma_0.$$

Let $\bar{w}(t)$ be the solution to

$$\dot{\bar{w}} = \epsilon^2 \frac{1}{T} \int_0^T \tilde{\alpha}(\tilde{u}(\sigma))u(\sigma) d\sigma, \quad \bar{w}(0) = \gamma_0^{(2)}. \quad (4.17)$$

From the definition of $\tilde{\alpha}$ (2.35), the k th component of the vector $\tilde{\alpha}(\tilde{u})u$ is

$$\sum_{i=1}^m \alpha_{ki}(\tilde{u})u_i = \sum_{i=1}^m \sum_{j=i+1}^m \Gamma_{ji}^k \tilde{u}_j u_i. \quad (4.18)$$

So using integration by parts, the fact that $\Gamma_{ij}^k = -\Gamma_{ji}^k$, $\dot{\tilde{u}}_i = u_i$ and the definition of $Area_{ij}(T)$ (2.22) we get from substituting (4.18) into (4.17) that

$$\bar{w}_k(t) = \epsilon^2 \frac{t}{T} \int_0^T \sum_{i=1}^m \alpha_{ki}(\tilde{u}(\sigma))u_i(\sigma) d\sigma + \gamma_{k0}^{(2)}$$

$$\begin{aligned}
&= \epsilon^2 \frac{t}{T} \sum_{i=1}^m \sum_{j=i+1}^m \int_0^T \Gamma_{ji}^k \tilde{u}_j(\sigma) \dot{\tilde{u}}_i(\sigma) d\sigma + \gamma_{k0}^{(2)} \\
&= \epsilon^2 \frac{t}{T} \sum_{i,j=1; i < j}^m \frac{1}{2} \int_0^T (\tilde{u}_j(\sigma) \dot{\tilde{u}}_i(\sigma) - \tilde{u}_i(\sigma) \dot{\tilde{u}}_j(\sigma)) d\sigma \Gamma_{ji}^k + \gamma_{k0}^{(2)} \\
&= \epsilon^2 \frac{t}{T} \sum_{i,j=1; i < j}^m \frac{1}{2} \int_0^T (\tilde{u}_i(\sigma) \dot{\tilde{u}}_j(\sigma) - \tilde{u}_j(\sigma) \dot{\tilde{u}}_i(\sigma)) d\sigma \Gamma_{ij}^k + \gamma_{k0}^{(2)} \\
&= \epsilon^2 \frac{t}{T} \sum_{i,j=1; i < j}^m \text{Area}_{ij}(T) \Gamma_{ij}^k + \gamma_{k0}^{(2)}.
\end{aligned}$$

Let $\gamma^{(2)}(t) = \bar{w}(t) + \epsilon \tilde{u}(t)$. By assumption, $\gamma^{(2)}(t) \in D$, $\forall t \in [0, b/\epsilon]$. So, for small enough ϵ , $\bar{w}(t) \in D$, $\forall t \in [0, b/\epsilon]$. Thus, by classical averaging (Theorem 7.4 [32]),

$$\|w(t) - \bar{w}(t)\| = O(\epsilon^2), \quad \forall t \in [0, b/\epsilon]. \quad (4.19)$$

So by (4.16) and (4.19)

$$\begin{aligned}
\|y(t) - \bar{w}(t)\| &= \|y(t) - w(t) + w(t) - \bar{w}(t)\| \\
&\leq \|y(t) - w(t)\| + \|w(t) - \bar{w}(t)\| \\
&= O(\epsilon^2), \quad \forall t \in [0, b/\epsilon].
\end{aligned} \quad (4.20)$$

Additionally,

$$\begin{aligned}
\|\epsilon v(t, y) - \epsilon \tilde{u}(t)\| &= \|\epsilon M(y) \tilde{u}(t) - \epsilon \tilde{u}(t)\| \\
&= \|(M(y) - I) \epsilon \tilde{u}(t)\| \\
&= \|O(\epsilon) \epsilon \tilde{u}(t)\| \\
&= O(\epsilon^2), \quad \forall t \in [0, b/\epsilon].
\end{aligned} \quad (4.21)$$

Then by (4.11), (4.20), (4.21),

$$\begin{aligned}
\|\gamma(t) - \gamma^{(2)}(t)\| &= \|y(t) + \epsilon v(t, y) - \bar{w}(t) - \epsilon \tilde{u}(t)\| \\
&\leq \|y(t) - \bar{w}(t)\| + \|\epsilon v(t, y) - \epsilon \tilde{u}(t)\| \\
&= O(\epsilon^2), \quad \forall t \in [0, b/\epsilon].
\end{aligned} \quad (4.22)$$

For small enough ϵ , since $\gamma^{(2)}(t) \in D \subset S$ then $\gamma(t) \in D \subset S$, $\forall t \in [0, b/\epsilon]$. So by Lemma 2.3,

$$g(t) = e^{\gamma_1(t)\xi_1} \dots e^{\gamma_n(t)\xi_n} = \Psi(\gamma(t)), \quad \forall t \in [0, b/\epsilon].$$

Define

$$g^{(2)}(t) = e^{\gamma_1^{(2)}(t)\xi_1} \dots e^{\gamma_n^{(2)}(t)\xi_n} = \Psi(\gamma^{(2)}(t)), \quad \forall t \in [0, b/\epsilon].$$

Then by definition of $\tilde{d}$,

$$\tilde{d}(g(t), g^{(2)}(t)) = \|\gamma(t) - \gamma^{(2)}(t)\| = O(\epsilon^2), \quad \forall t \in [0, b/\epsilon]. \quad \square$$

We show further in the next proposition that the structure constants Γ_{ij}^k associated to a given basis for $\mathcal{G}$ are directly related to the Lie brackets of the vector fields defined by the columns of $M(\gamma)$ evaluated at $\gamma = 0$.

Proposition 4.4 Suppose that $\bar{w}(t)$ is defined by (4.8). Let $[f_1 \ f_2 \ \dots \ f_n] = M(\gamma)$ where f_k is the k th column of the matrix $M(\gamma)$. Then

$$\bar{w}(t) = \frac{\epsilon^2 t}{T} \sum_{i,j=1; i < j}^m \text{Area}_{ij}(T) [f_i, f_j]|_{\gamma=0} + \gamma_0^{(2)}. \quad (4.23)$$

Proof. By (2.35) and (2.36) of Remark 2.4 we have that

$$f_i = \begin{bmatrix} \sum_{k=i+1}^n \gamma_k \Gamma_{ki}^1 + O(\gamma^2) \\ \vdots \\ \sum_{k=i+1}^n \gamma_k \Gamma_{ki}^{(i-1)} + O(\gamma^2) \\ 1 + \sum_{k=i+1}^n \gamma_k \Gamma_{ki}^i + O(\gamma^2) \\ \sum_{k=i+1}^n \gamma_k \Gamma_{ki}^{i+1} + O(\gamma^2) \\ \vdots \\ \sum_{k=i+1}^n \gamma_k \Gamma_{ki}^n + O(\gamma^2) \end{bmatrix}.$$

So $f_i|_{\gamma=0} = e_i$ where e_i is the i th standard basis vector for $\Re^n$ and

$$\left. \frac{\partial f_i}{\partial \gamma} \right|_{\gamma=0} = \begin{bmatrix} 0 & \cdots & 0 & \Gamma_{(i+1)i}^1 & \cdots & \Gamma_{ni}^1 \\ \cdot & & \cdot & \cdot & & \cdot \\ \cdot & & \cdot & \cdot & & \cdot \\ \cdot & & \cdot & \cdot & & \cdot \\ 0 & \cdots & 0 & \Gamma_{(i+1)i}^n & \cdots & \Gamma_{ni}^n \end{bmatrix}.$$

So for $i < j$,

$$\begin{aligned} [f_i, f_j]|_{\gamma=0} &= \left. \frac{\partial f_j}{\partial \gamma} \right|_{\gamma=0} f_i|_{\gamma=0} - \left. \frac{\partial f_i}{\partial \gamma} \right|_{\gamma=0} f_j|_{\gamma=0} \\ &= - \begin{bmatrix} \Gamma_{ji}^1 \\ \vdots \\ \Gamma_{ji}^n \end{bmatrix} = \begin{bmatrix} \Gamma_{ij}^1 \\ \vdots \\ \Gamma_{ij}^n \end{bmatrix}. \end{aligned}$$

which by (4.8) completes the proof. $\square$

Remark 4.5 According to Theorem 4.3, the second-order approximation $g^{(2)}(t)$ to solutions $g(t)$ of (2.9) can be expressed as a product of exponentials where the exponents have an $O(\epsilon)$ periodic term and a secular term (a term linear in t). By (4.8) the secular term is proportional to the structure constants Γ_{ij}^k associated to $\{\xi_1, \dots, \xi_n\}$ and the areas $Area_{ij}(T)$ bounded by the closed curves described by $\tilde{u}_i$ and $\tilde{u}_j$ over one period. This interpretation justifies calling Theorem 4.3 an area rule.

The next corollary illustrates averaging for system (2.11) on the matrix Lie group $G = SO(3)$ where the error estimate is given in terms of the metric $\bar{\phi}$ on $SO(3)$ defined by (3.6) in Section 3.1.

Corollary 4.6 Consider system (2.11), the matrix version of system (2.9), where $G = SO(3)$, $X(t) \in SO(3)$ is the solution to (2.11) and $\{A_1, A_2, A_3\}$ is the standard basis for $\mathcal{G} = so(3)$ as defined in Section 3.1. Assume $U(t) \in so(3)$ is as in Theorem 4.3. Let $D = \{\gamma \in \mathfrak{R}^3 \mid |\gamma_2| < \pi/2 - \delta, \delta > 0\}$ (for choice of δ see condition on D below). Let $X(0) = X_0 \in Q$. Let $\gamma(t) \in \mathfrak{R}^3$ be the solution of the Wei-Norman equations for $SO(3)$ given by (3.4) with $\gamma(0) = \gamma_0 = \Psi^{-1}(X_0) = O(\epsilon)$. Define

$$\begin{aligned}\gamma_1^{(2)}(t) &= \epsilon \tilde{u}_1(t) + \epsilon^2 \frac{t}{T} Area_{23}(T) + \gamma_{10}^{(2)}, \\ \gamma_2^{(2)}(t) &= \epsilon \tilde{u}_2(t) + \epsilon^2 \frac{t}{T} Area_{31}(T) + \gamma_{20}^{(2)}, \\ \gamma_3^{(2)}(t) &= \epsilon \tilde{u}_3(t) + \epsilon^2 \frac{t}{T} Area_{12}(T) + \gamma_{30}^{(2)}, \\ X^{(2)}(t) &= e^{\gamma_1^{(2)}(t)A_1} e^{\gamma_2^{(2)}(t)A_2} e^{\gamma_3^{(2)}(t)A_3}.\end{aligned}\tag{4.24}$$

Then if $\|\gamma_0 - \gamma_0^{(2)}\| = O(\epsilon^2)$ and if $\gamma^{(2)}(t) \in D, \forall t \in [0, b/\epsilon]$,

$$\bar{\phi}(X(t), X^{(2)}(t)) = O(\epsilon^2), \quad \forall t \in [0, b/\epsilon].$$

Proof. Since $\Gamma_{12}^3 = \Gamma_{23}^1 = \Gamma_{31}^2 = 1$ are the nonzero structure constants for the given basis for $so(3)$, $\gamma^{(2)}(t)$ as defined by (4.9) reduces to (4.24). By Theorem 4.3,

$$\|\gamma(t) - \gamma^{(2)}(t)\| = O(\epsilon^2), \quad \forall t \in [0, b/\epsilon].\tag{4.26}$$

We note that D is small enough to ensure that the Wei-Norman equations are well-defined $\forall \gamma \in D$. We need not restrict D to be contained in S (as we do in Theorem 4.3), since we do not use the diffeomorphism Ψ to define a metric on $SO(3)$ in this corollary. By (3.6)

$$\bar{\phi}(X, X^{(2)}) = \phi(XX^{(2)T}) = \cos^{-1}_{[0, \pi]}(1/2(\text{tr}(XX^{(2)T}) - 1)).$$

Since $\text{tr}(ABC) = \text{tr}(BCA) = \text{tr}(CAB)$,

$$\begin{aligned}\text{tr}(XX^{(2)T}) &= \text{tr}(e^{\gamma_1 A_1} e^{\gamma_2 A_2} e^{\gamma_3 A_3} e^{-\gamma_3^{(2)} A_3} e^{-\gamma_2^{(2)} A_2} e^{-\gamma_1^{(2)} A_1}) \\ &= \text{tr}(e^{\gamma_2 A_2} e^{(\gamma_3 - \gamma_3^{(2)}) A_3} e^{-\gamma_2^{(2)} A_2} e^{(\gamma_1 - \gamma_1^{(2)}) A_1}).\end{aligned}$$

Since by (3.7), $\phi(AB^T) \leq \phi(A) + \phi(B)$,

$$\begin{aligned}\phi(XX^{(2)T}) &= \phi(e^{\gamma_2 A_2} e^{(\gamma_3 - \gamma_3^{(2)}) A_3} e^{-\gamma_2^{(2)} A_2} e^{(\gamma_1 - \gamma_1^{(2)}) A_1}) \\ &\leq \phi(e^{\gamma_2 A_2} e^{(\gamma_3 - \gamma_3^{(2)}) A_3} e^{-\gamma_2^{(2)} A_2}) + \phi(e^{(\gamma_1 - \gamma_1^{(2)}) A_1}) \\ &= \phi(e^{(\gamma_3 - \gamma_3^{(2)}) A_3} e^{(\gamma_2 - \gamma_2^{(2)}) A_2}) + \phi(e^{(\gamma_1 - \gamma_1^{(2)}) A_1}) \\ &\leq \phi(e^{(\gamma_3 - \gamma_3^{(2)}) A_3}) + \phi(e^{(\gamma_2 - \gamma_2^{(2)}) A_2}) + \phi(e^{(\gamma_1 - \gamma_1^{(2)}) A_1}).\end{aligned}\quad (4.27)$$

By direct computation, $\text{tr}(\exp(a_i A_i)) = 2 \cos(a_i) + 1$, $i = 1, 2, 3$. Therefore,

$$\begin{aligned}\phi(e^{\pm(\gamma_i - \gamma_i^{(2)}) A_i}) &= \cos_{[0, \pi]}^{-1}\left(\frac{1}{2}(2 \cos(\gamma_i - \gamma_i^{(2)}) + 1 - 1)\right) \\ &= \cos_{[0, \pi]}^{-1}(\cos(\gamma_i - \gamma_i^{(2)})) \\ &= |\gamma_i - \gamma_i^{(2)}|, \quad i = 1, 2, 3.\end{aligned}\quad (4.28)$$

It follows from (4.26), (4.27), (4.28) that

$$\begin{aligned}\bar{\phi}(X(t), X^{(2)}(t)) = \phi(X(t)X^{(2)T}(t)) &\leq \sum_{i=1}^3 |\gamma_i(t) - \gamma_i^{(2)}(t)| \\ &= \|\gamma(t) - \gamma^{(2)}(t)\|_1 \\ &= O(\epsilon^2), \quad \forall t \in [0, b/\epsilon].\end{aligned}\quad \square$$

Remark 4.7 We can draw similar conclusions for other metrics. For example, given the assumption in Corollary 4.6 we can show that $\|X(t)X^{(2)T}(t) - I\|_1 = O(\epsilon^2)$, $\forall t \in [0, b/\epsilon]$. This can be shown by directly computing $X(t)X^{(2)T}(t) - I$. The bound on the norm can then be found using trigonometric identities and the fact that

$$\sum_{i=1}^3 |\gamma_i - \gamma_i^{(2)}| = \|\gamma - \gamma^{(2)}\|_1 = O(\epsilon^2) \implies |\gamma_i - \gamma_i^{(2)}| = O(\epsilon^2), \quad i = 1, 2, 3$$

$$\implies |\sin(\gamma_i - \gamma_i^{(2)})| = O(\epsilon^2), \quad |\cos(\gamma_i - \gamma_i^{(2)}) - 1| = O(\epsilon^2), \quad i = 1, 2, 3,$$

for small enough ϵ .

4.2 Averaging with Single Exponential

In this section we derive averaging theory for systems of the form (2.9) using the single exponential representation of solutions described in Section 2.3.2. This representation is more natural for our purposes than the product of exponentials representation, and we are able to derive a more general q th-order averaging theory ($q \geq 1$). We consider the basic ideas of classical averaging theory to derive the theory of this section, but we do not directly use the results from classical averaging theory. One important consequence of this is that the forcing function $U(t) \in \mathcal{G}$ need only be piecewise continuous rather than smooth as assumed in Theorems 4.1 and 4.3 and generally in classical averaging theory.

Consider the single exponential representation $g(t) = e^{Z(t)}$ for the solution to (2.9), where $Z(t) \in \mathcal{G}$ is given by the infinite series $Z(t) = \sum_{i=1}^{\infty} (-1)^{i+1} \epsilon^i Z_i(t)$ of Corollary 2.7 and $Z_i(t)$ is defined by (2.43) (Fomenko and Chakon). Let $\dot{Z}_i = dZ_i(t)/dt$ which can be computed by differentiating (2.43) as

$$\begin{aligned} \dot{Z}_1(t) = \dot{T}_0(t) &= U(t), \\ (i+1)\dot{Z}_{i+1}(t) &= \dot{T}_i + \sum_{r=1}^i \left\{ \frac{1}{2} [\dot{Z}_r, T_{i-r}] + \frac{1}{2} [Z_r, \dot{T}_{i-r}] \right. \\ &\quad + \sum_{p \geq 1, 2p \leq r} k_{2p} \sum_{\sum_{j=1}^{2p} m_j = r, m_j > 0} ([\dot{Z}_{m_1}, [Z_{m_2}, \dots, [Z_{m_{2p}}, T_{i-r}] \dots]] + \dots \\ &\quad \left. + [Z_{m_1}, [Z_{m_2}, \dots, [\dot{Z}_{m_{2p}}, T_{i-r}] \dots]] + [Z_{m_1}, [Z_{m_2}, \dots, [Z_{m_{2p}}, \dot{T}_{i-r}] \dots]] \right\}, \\ \dot{T}_k(t) &= [U(t), \int_0^t [U(\tau_2), \dots, \int_0^{\tau_k} U(\tau_{k+1}) d\tau_{k+1}] d\tau_k] \dots d\tau_2, \end{aligned} \tag{4.29}$$

The following lemma shows the conditions for which the terms in the series expression for $Z(t)$ are periodic in t .

Lemma 4.8 Let $q \geq 2$ be a positive integer and let $U(t+T) = U(t)$, $\forall t > 0$. If

$$T_k(T) = 0, \quad k = 0, \dots, q-2, \quad (4.30)$$

then $\forall t > 0$,

- (a) $T_k(t+T) = T_k(t)$, $k = 0, \dots, q-2$.
- (b) $\dot{T}_k(t+T) = \dot{T}_k(t)$, $k = 0, \dots, q-1$.
- (c) $Z_i(t+T) = Z_i(t)$, $i = 1, \dots, q-1$.
- (d) $\dot{Z}_i(t+T) = \dot{Z}_i(t)$, $i = 1, \dots, q$.

Proof: Assume (4.30) holds. We prove (a) by induction. For $q = 2$ assumption (4.30) implies $T_0(T) = 0$. Then

$$\begin{aligned} T_0(t+T) &= \int_0^{t+T} U(\tau) d\tau \\ &= \int_0^T U(\tau) d\tau + \int_T^{t+T} U(\tau) d\tau \\ &= T_0(T) + \int_0^t U(\sigma+T) d\sigma \\ &= 0 + \int_0^t U(\sigma) d\sigma \\ &= T_0(t), \end{aligned} \quad (4.31)$$

where we set $\sigma = \tau - T$. Now assume (a) holds for $q = q'$. Then we must show that (a) holds for $q = q' + 1$ assuming $T_k(T) = 0$, $k = 0, \dots, q' - 1$. Since (a) holds for $q = q'$, $T_k(t+T) = T_k(t)$ for $k = 0, \dots, q' - 2$. So we need only show $T_{q'-1}(t+T) = T_{q'-1}(t)$.

$$T_{q'-1}(t+T) = \int_0^{t+T} [U(\tau_1), T_{q'-2}(\tau_1)] d\tau_1$$

$$\begin{aligned}
&= \int_0^T [U(\tau_1), T_{q'-2}(\tau_1)] d\tau_1 + \int_T^{t+T} [U(\tau_1), T_{q'-2}(\tau_1)] d\tau_1 \\
&= T_{q'-1}(T) + \int_0^t [U(\sigma_1 + T), T_{q'-2}(\sigma_1 + T)] d\sigma_1 \\
&= 0 + \int_0^t [U(\sigma_1), T_{q'-2}(\sigma_1)] d\sigma_1 \\
&= T_{q'-1}(t),
\end{aligned}$$

where we set $\sigma_1 = \tau_1 - T$. Thus, (a) holds by induction.

Conclusion (b) follows as a result of (a). By (4.29)

$$\dot{T}_k(t) = [U(t), T_{k-1}(t)],$$

and by (a) $T_k(t+T) = T_k(t)$ for $k = 0, \dots, q-2$. So for $k = 0, \dots, q-1$,

$$\begin{aligned}
\dot{T}_k(t+T) &= [U(t+T), T_{k-1}(t+T)] \\
&= [U(t), T_{k-1}(t)] \\
&= \dot{T}_k(t).
\end{aligned}$$

We prove (c) by induction. Let $q = 2$. Since $Z_1 = T_0$, by (4.31) $Z_1(t+T) = Z_1(t)$. Now assume (c) holds for $q = q'$. We show that it holds for $q = q' + 1$ assuming $T_k(T) = 0$, $k = 0, \dots, q'-1$. Since (c) holds for $q = q'$, $Z_i(t+T) = Z_i(t)$ for $i = 1, \dots, q' - 1$. So we need only show $Z_{q'}(t+T) = Z_{q'}(t)$. However, from (2.43) $Z_{q'}(t)$ is a function only of terms $T_l(t)$ and $Z_l(t)$, $l = 1, \dots, q' - 1$. By (a) $T_l(t+T) = T_l(t)$, $l = 1, \dots, q' - 1$. By the induction hypothesis, $Z_l(t+T) = Z_l(t)$, $l = 1, \dots, q' - 1$. Thus, $Z_{q'}(t+T) = Z_{q'}(t)$, and (c) holds by induction.

Finally, we prove (d) by induction. Let $q = 2$. We have $\dot{Z}_2(t) = \frac{1}{2}\dot{T}_1(t)$. So by (b) $\dot{Z}_2(t+T) = \dot{Z}_2(t)$. Now assume (d) holds for $q = q'$. We show that it holds for $q = q' + 1$ assuming $T_k(T) = 0$, $k = 0, \dots, q' - 1$. Since (d)

holds for $q = q'$, $\dot{Z}_i(t+T) = \dot{Z}_i(t)$ for $i = 1, \dots, q'$. So we need only show that $\dot{Z}_{q'+1}(t+T) = \dot{Z}_{q'+1}(t)$. From (4.29) we have that $\dot{Z}_{q'+1}(t)$ is a function only of terms $T_j(t)$, $j = 1, \dots, q' - 1$ and $\dot{T}_l(t), Z_l(t), \dot{Z}_l(t)$, $l = 1, \dots, q'$. By (a), $T_j(t+T) = T_j(t)$, $j = 1, \dots, q' - 1$. By (b), $\dot{T}_l(t+T) = \dot{T}_l(t)$, $l = 1, \dots, q'$. By (c), $Z_l(t+T) = Z_l(t)$, $l = 1, \dots, q'$. By the induction hypothesis, $\dot{Z}_l(t+T) = \dot{Z}_l(t)$, $l = 1, \dots, q'$. Thus, $\dot{Z}_{q'+1}(t+T) = \dot{Z}_{q'+1}(t)$, and (d) holds by induction. $\square$

The next series of theorems provide increasingly simplified approximations to the single exponential representation of $g(t)$. In Theorem 4.9, the first of the series, we give a condition that ensures a valid local representation on an $O(1/\epsilon)$ time interval. Further, we show for $\mathcal{G}$ a nilpotent Lie algebra, that $g(t)$ can be represented as the exponential of a *finite* sum. The definitions of the neighborhoods $\hat{S} \subset \mathcal{G}$, $\hat{Q} \subset G$, the diffeomorphism $\hat{\Psi} : \hat{S} \rightarrow \hat{Q}$ and the metric $\hat{d}$ on $\hat{Q}$ which we use in the following theorems can be found in Section 2.3.2.

Theorem 4.9 Consider system (2.9) on the Lie group G with Lie algebra $\mathcal{G}$. Assume that $U(t)$ is a piecewise continuous, bounded curve in $\mathcal{G}$. Let $b > 0$ be such that $\int_0^t \|U(\tau)\| d\tau < \delta$, $\forall t \in [0, b]$, where δ is as defined in Theorem 2.6. Then, $g(t) = g_0 e^{Z(t)}$ is the solution to (2.9), $\forall t \in [0, b/\epsilon]$ where $Z(t) = \sum_{i=1}^{\infty} (-1)^{i+1} \epsilon^i Z_i(t)$, $Z_i(t)$ are defined by (2.43) and $g(0) = g_0 \in G$. Suppose that $\mathcal{G}$ is *nilpotent* of order l . Define

$$Z_e^{[l]}(t) = \sum_{i=1}^l (-1)^{i+1} \epsilon^i Z_i(t).$$

Then,

$$g(t) = g_0 e^{Z_e^{[l]}(t)}, \quad \forall t \in [0, b/\epsilon].$$

Proof. Let

$$K \triangleq \sup_{t \geq 0} \|U(t)\|.$$

Let

$$b \leq \frac{\delta}{K}.$$

Then, $\forall t \in [0, b]$,

$$\int_0^t \|U(\tau)\| d\tau \leq \int_0^b \|U(\tau)\| d\tau \leq \int_0^b K d\tau = Kb \leq \delta.$$

Further, $\forall t \in [0, b/\epsilon]$,

$$\int_0^t \epsilon \|U(\tau)\| d\tau \leq \int_0^{b/\epsilon} \epsilon \|U(\tau)\| d\tau \leq \int_0^{b/\epsilon} \epsilon K d\tau = Kb \leq \delta.$$

As a result, by Theorem 2.6 (Fomenko and Chakon) and Corollary 2.7, $g(t) = g_0 e^{Z(t)}$, $\forall t \in [0, b/\epsilon]$, $Z(t) = \sum_{i=1}^{\infty} (-1)^{i+1} \epsilon^i Z_i(t)$.

For $\mathcal{G}$ nilpotent of order l , all Lie brackets on $\mathcal{G}$ of depth greater than $l - 1$ are zero. $Z_i(t)$ is composed of depth- $(i - 1)$ brackets on $\mathcal{G}$; therefore, $Z_i(t) = 0$, $i > l$. So $Z(t)$ is the finite sum $Z_e^{[l]}(t)$. $\square$

In the next theorem we show that for any $\mathcal{G}$, $g(t)$ can be represented *approximately* as the exponential of a finite sum. We also show that if $g(0) \in G$ is in a small enough neighborhood of the identity $e \in G$, then $g(0)$ can be subsumed into the single exponential representation of $g(t)$.

Theorem 4.10 Consider system (2.9) on the Lie group G with Lie algebra $\mathcal{G}$. Let $U(t)$ and b be as assumed in Theorem 4.9. Let $q \geq 1$ be an integer. Suppose that $g(0) = g_0 \in \hat{Q} \subset G$ is such that $Z_0 = \hat{\Psi}^{-1}(g_0) = O(\epsilon^{q-1})$ if $q > 1$ and $Z_0 = O(\epsilon)$ if $q = 1$. Define

$$Z^{[q]}(t) = \sum_{i=1}^q (-1)^{i+1} \epsilon^i Z_i(t) + Z_0^{(q)}, \quad (4.32)$$

$$g^{[q]}(t) = e^{Z^{[q]}(t)}, \quad (4.33)$$

where $Z_i(t)$ are defined by (2.43). If $\|Z_0 - Z_0^{(q)}\| = O(\epsilon^q)$ and $Z^{[q]}(t) \in \hat{S}$, $\forall t \in [0, b/\epsilon]$, then

$$\hat{d}(g(t), g^{[q]}(t)) = O(\epsilon^q), \quad \forall t \in [0, b/\epsilon].$$

Proof: As in Theorem 4.9 we have $g(t) = g_0 e^{Z(t)}$, $\forall t \in [0, b/\epsilon]$ where $Z(t) = \sum_{i=1}^{\infty} (-1)^{i+1} \epsilon^i Z_i(t)$. Since $g_0 = e^{Z_0}$, by the Campbell-Baker-Hausdorff formula we have that for small enough ϵ ,

$$g(t) = e^{Z_0} e^{Z(t)} = e^{W(t)}, \quad \forall t \in [0, b/\epsilon].$$

$W(t) \in \mathcal{G}$ can be expressed as the following convergent series:

$$W(t) = Z(t) + Z_0 + \frac{1}{2}[Z_0, Z(t)] + \frac{1}{12}([Z_0, Z(t)], Z(t)) + [[Z(t), Z_0], Z_0] + \dots \quad (4.34)$$

So, substituting in (4.34) the expression

$$Z(t) = Z^{[q]}(t) - Z_0^{(q)} + \sum_{i=q+1}^{\infty} (-1)^{i+1} \epsilon^i Z_i(t)$$

as well as $Z(t) = O(\epsilon)$ and $Z_0 = O(\epsilon^{q-1})$ gives

$$\begin{aligned} W(t) &= Z^{[q]}(t) + (Z_0 - Z_0^{(q)}) + \sum_{i=q+1}^{\infty} (-1)^{i+1} \epsilon^i Z_i(t) + \frac{1}{2}[O(\epsilon^{q-1}), O(\epsilon)] + \dots \\ &= Z^{[q]}(t) + O(\epsilon^q). \end{aligned} \quad (4.35)$$

Since $Z^{[q]}(t) \in \hat{S}$, $\forall t \in [0, b/\epsilon]$, for small enough ϵ , $W(t) \in \hat{S}$, $\forall t \in [0, b/\epsilon]$. Thus, $g(t) = \hat{\Psi}(W(t)) \in \hat{Q}$ and $g^{[q]}(t) = \hat{\Psi}(Z^{[q]}(t)) \in \hat{Q}$, $\forall t \in [0, b/\epsilon]$. It follows from the definition of $\hat{d}$ (2.46) and (4.35) that

$$\hat{d}(g(t), g^{[q]}(t)) = \|W(t) - Z^{[q]}(t)\| = O(\epsilon^q), \quad \forall t \in [0, b/\epsilon]. \quad \square$$

In the next theorem we show that the q th-order finite sum approximation of $Z(t)$ can be further simplified by replacing the q th term in the sum with its time average. This theorem is our general q th-order averaging theorem which we also refer to as the Area-Moment Rule.

Theorem 4.11 (qth-Order Averaging - Area-Moment Rule) Consider system (2.9) on the Lie group G with Lie algebra $\mathcal{G}$. Let $U(t)$ and b be as assumed in Theorem 4.9. Further, assume that $U(t)$ is periodic in t of period T , $\forall t \in [0, \infty)$. Let $q \geq 1$ be an integer. For $q > 1$ assume that $T_k(T) = 0$, $k = 0, \dots, q-2$, where $T_k(t)$ is defined by (2.43). Suppose that $g(0) = g_0 \in \hat{Q} \subset G$ is such that $Z_0 = \hat{\Psi}^{-1}(g_0) = O(\epsilon^{q-1})$ if $q > 1$ and $Z_0 = O(\epsilon)$ if $q = 1$. Define

$$Z^{(q)}(t) = \sum_{i=1}^{q-1} (-1)^{i+1} \epsilon^i Z_i(t) + (-1)^{q+1} \epsilon^q \frac{t}{T} Z_q(T) + Z_0^{(q)}, \quad (4.36)$$

$$g^{(q)}(t) = e^{Z^{(q)}(t)}, \quad (4.37)$$

where $Z_i(t)$ are defined by (2.43). If $\|Z_0 - Z_0^{(q)}\| = O(\epsilon^q)$ and $Z^{(q)}(t) \in \hat{S}$, $\forall t \in [0, b/\epsilon]$, then

$$\hat{d}(g(t), g^{(q)}(t)) = O(\epsilon^q), \quad \forall t \in [0, b/\epsilon]. \quad (4.38)$$

Further, for $t = NT$, N an integer,

$$Z^{(q)}(NT) = (-1)^{q+1} \epsilon^q N Z_q(T) + Z_0^{(q)}. \quad (4.39)$$

Proof. From Theorem 4.10, $g(t) = e^{W(t)}$, $\forall t \in [0, b/\epsilon]$, where $W(t)$ is given by

$$\begin{aligned} W(t) &= Z^{[q]}(t) + O(\epsilon^q) \\ &= \sum_{i=1}^q (-1)^{i+1} \epsilon^i Z_i(t) + Z_0^{(q)} + O(\epsilon^q). \end{aligned} \quad (4.40)$$

Let $Y(t)$ be defined by

$$Y(t) = (-1)^{q+1} \epsilon^q Z_q(t) + Z_0^{(q)}. \quad (4.41)$$

Let $\bar{Y}(t)$ be defined by

$$\begin{aligned} \bar{Y}(t) &= (-1)^{q+1} \epsilon^q \frac{t}{T} \int_0^T \dot{Z}_q(t) dt + Z_0^{(q)} \\ &= (-1)^{q+1} \epsilon^q \frac{t}{T} Z_q(T) + Z_0^{(q)}. \end{aligned} \quad (4.42)$$

Let $V(t)$ be defined by

$$V(t) = Z_q(t) - \frac{t}{T} Z_q(T). \quad (4.43)$$

By Lemma 4.8, since $T_k(T) = 0$, $k = 0, \dots, q-2$, we have that $\dot{Z}_q(t+T) = \dot{Z}_q(t)$, $\forall t > 0$, and so

$$\begin{aligned} V(t+T) &= Z_q(t+T) - \frac{t+T}{T} Z_q(T) \\ &= \int_0^{t+T} \dot{Z}_q(\tau) d\tau - \frac{t}{T} Z_q(T) - Z_q(T) \\ &= \int_0^T \dot{Z}_q(\tau) d\tau + \int_T^{t+T} \dot{Z}_q(\tau) d\tau - \frac{t}{T} Z_q(T) - Z_q(T) \\ &= Z_q(T) + \int_0^t \dot{Z}_q(\sigma+T) d\sigma - \frac{t}{T} Z_q(T) - Z_q(T) \\ &= \int_0^t \dot{Z}_q(\sigma) d\sigma - \frac{t}{T} Z_q(T) \\ &= V(t), \end{aligned}$$

where we set $\sigma = \tau - T$. So $V(t)$ is periodic in t and, thus, bounded $\forall t$. Let M be a constant such that $\|V(t)\| \leq M$, $\forall t \in [0, b/\epsilon]$. Then, by (4.41), (4.42) and (4.43), $\forall t \in [0, b/\epsilon]$,

$$\|Y(t) - \bar{Y}(t)\| = \|(-1)^{q+1} \epsilon^q (Z_q(t) - \frac{t}{T} Z_q(T))\| = |\epsilon^q| \|V(t)\| \leq M |\epsilon^q| = O(\epsilon^q) \quad (4.44)$$

So by (4.32), (4.36), (4.40), (4.41), (4.42), (4.44) and the triangle inequality,

$$\begin{aligned}
\|W(t) - Z^{(q)}(t)\| &= \|W(t) - (\sum_{i=1}^{q-1} (-1)^{i+1} \epsilon^i Z_i(t) + (-1)^{q+1} \epsilon^q \frac{t}{T} Z_q(T) + Z_0^{(q)})\| \\
&= \|W(t) - (Z^{[q]}(t) - (-1)^{q+1} \epsilon^q Z_q(t) + (-1)^{q+1} \epsilon^q \frac{t}{T} Z_q(T))\| \\
&\leq \|W(t) - Z^{[q]}(t)\| + \|Y(t) - \bar{Y}(t)\| \\
&= O(\epsilon^q), \quad \forall t \in [0, b/\epsilon].
\end{aligned} \tag{4.45}$$

Since $Z^{(q)}(t) \in \hat{S}$, $\forall t \in [0, b/\epsilon]$, for small enough ϵ , $W(t) \in \hat{S}$. Thus, $g(t) = \hat{\Psi}(W(t)) \in \hat{Q}$ and $g^{(q)}(t) = \hat{\Psi}(Z^{(q)}(t)) \in \hat{Q}$, $\forall t \in [0, b/\epsilon]$. It follows from the definition of $\hat{d}$ (2.46) and (4.45) that

$$\hat{d}(g(t), g^{(q)}(t)) = \|W(t) - Z^{(q)}(t)\| = O(\epsilon^q), \quad \forall t \in [0, b/\epsilon].$$

Finally, from Lemma 4.8, since $Z_i(0) = 0$ and $Z_i(t+T) = Z_i(t)$, $i = 1, \dots, q-1$, we have that $Z_i(NT) = Z_i(0) = 0$, $i = 1, \dots, q-1$, for N an integer. Thus,

$$\begin{aligned}
Z^{(q)}(NT) &= \sum_{i=1}^{q-1} (-1)^{i+1} \epsilon^i Z_i(NT) + (-1)^{q+1} \epsilon^q \frac{NT}{T} Z_q(T) + Z_0^{(q)} \\
&= (-1)^{q+1} \epsilon^q N Z_q(T) + Z_0^{(q)}
\end{aligned} \quad \square$$

Let I_p be the ordered list of indices $\{i_1, \dots, i_p\}$ where $i_\nu \in \{1, \dots, m\}$, $\nu = 1, \dots, p$. We define

$$\sum_{I_p=1}^m \triangleq \sum_{i_1=1}^m \sum_{i_2=1}^m \dots \sum_{i_p=1}^m.$$

Let

$$t_{I_p}(t) = \int_0^t u_{i_1}(\tau_1) \left(\int_0^{\tau_1} u_{i_2}(\tau_2) \dots \left(\int_0^{\tau_{p-1}} u_{i_p}(\tau_p) d\tau_p \right) \dots d\tau_2 \right) d\tau_1.$$

Then, using the definition of high-order structure constants (equation (2.18)) associated with the basis $\{\xi_1, \dots, \xi_n\}$ for $\mathcal{G}$,

$$\begin{aligned}
T_k(t) &= \sum_{I_{k+1}=1}^m t_{I_{k+1}}(t) [\xi_{i_1}, [\xi_{i_2}, \dots [\xi_{i_k}, \xi_{i_{k+1}}] \dots]] \\
&= \sum_{l=1}^n \sum_{I_{k+1}=1}^m t_{I_{k+1}}(t) \Gamma_{I_{k+1}}^l \xi_l.
\end{aligned} \tag{4.46}$$

In the next theorem and remark we justify the name Area-Moment Rule.

Theorem 4.12 For $q \geq 2$, $Z_q(T)$ of Theorem 4.11 can be expressed as

$$Z_q(T) = \frac{1}{q} \int_0^T [U(\tau), T_{q-2}(\tau)] d\tau.$$

$\tilde{U}(t) = \int_0^t U(\tau) d\tau$ and $T_{q-2}(t)$ are both periodic in t with period T . Consequently, $Z_q(T)$ has the following geometric interpretation. Let $\Delta_{iI_{q-1}}(T)$ be the area bounded by the closed curve described by $\tilde{u}_i(t)$ and $t_{I_{q-1}}(t)$ from $t = 0$ to $t = T$. Then,

$$Z_q(T) = \frac{1}{q} \sum_{p=1}^n \sum_{i=1}^m \sum_{I_{q-1}=1}^m \Delta_{iI_{q-1}}(T) \Gamma_{\{i, I_{q-1}\}}^p \xi_p,$$

where $\{i, I_{q-1}\} = \{i, i_1, \dots, i_{q-1}\}$ and $\Gamma_{\{i, I_{q-1}\}}^p$ are depth- $(q-1)$ structure constants associated with the basis $\{\xi_1, \dots, \xi_n\}$ for $\mathcal{G}$ defined by (2.18).

Proof: Since in Theorem 4.11 we assumed $U(t+T) = U(t)$, $\forall t > 0$ and $T_k(T) = 0$, $k = 0, \dots, q-2$, by Lemma 4.8, $Z_i(t+T) = Z_i(t)$, $\forall t > 0$, $i = 1, \dots, q-1$. Since $Z_i(0) = 0, \forall i$, then $Z_i(T) = 0$, $i = 1, \dots, q-1$. Thus, the only nonzero term in the expression for $Z_q(T)$ ((2.43)) is $T_{q-1}(T)$ and so

$$\begin{aligned} Z_q(T) &= \frac{1}{q} T_{q-1}(T) \\ &= \frac{1}{q} \int_0^T [U(\tau), T_{q-2}(\tau)] d\tau. \end{aligned} \tag{4.47}$$

From Lemma 4.8, $T_{q-2}(t+T) = T_{q-2}(t)$, $\forall t > 0$. Thus, by (4.46) $t_{I_{q-1}}(t+T) = t_{I_{q-1}}(t)$, $\forall t > 0$. Also, $\tilde{U}(t+T) = T_0(t+T) = T_0(t) = \tilde{U}(t)$, $\forall t > 0$ which implies that $\tilde{u}_i(t+T) = \tilde{u}_i(t)$, $\forall t > 0$. Therefore, the plane curve described by $\tilde{u}_i(t)$ and $t_{I_{q-1}}(t)$ from $t = 0$ to $t = T$ is closed. Recalling that $u_i(t) = \dot{\tilde{u}}_i(t)$, by Green's Theorem, $\Delta_{iI_{q-1}}(T) = \int_0^T u_i(\tau) t_{I_{q-1}}(\tau) d\tau$. From (4.46) and (4.47) we have that

$$Z_q(T) = \frac{1}{q} \int_0^T \left[\sum_{i=1}^m u_i(\tau) \xi_i, \sum_{I_{q-1}=1}^m t_{I_{q-1}}(\tau) [\xi_{i_1}, [\xi_{i_2}, \dots [\xi_{i_{q-2}}, \xi_{i_{q-1}}] \dots]] \right] d\tau$$

$$\begin{aligned}
&= \frac{1}{q} \sum_{i=1}^m \sum_{I_{q-1}=1}^m \left(\int_0^T u_i(\tau) t_{I_{q-1}}(\tau) d\tau \right) [\xi_i, [\xi_{i_1}, [\xi_{i_2}, \dots [\xi_{i_{q-2}}, \xi_{i_{q-1}}] \dots]]] \\
&= \frac{1}{q} \sum_{p=1}^n \sum_{i=1}^m \sum_{I_{q-1}=1}^m \Delta_{iI_{q-1}}(T) \Gamma_{\{i, I_{q-1}\}}^p \xi_p.
\end{aligned}
\tag*{$\square$}$$

Remark 4.13 $Z^{(q)}(t)$ of Theorem 4.11 is the q th-order average approximation of $Z(t)$ on $[0, b/\epsilon]$. $Z^{(q)}(t)$ is the sum of a secular term (a term linear in t) and an $O(\epsilon)$ periodic term. The secular term is proportional to $Z_q(T)$ which, according to Theorem 4.12, is proportional to the areas $\Delta_{iI_{q-1}}$ and the depth- $(q-1)$ structure constants associated with $\{\xi_1, \dots, \xi_n\}$. This justifies calling Theorem 4.11 an area rule. Using integration by parts, we can derive equivalent expressions for $Z_q(T)$ that lend an alternative geometric interpretation, i.e.,

$$\begin{aligned}
Z_q(T) &= \frac{1}{q} \int_0^T [U(\tau), T_{q-2}(\tau)] d\tau \\
&= -\frac{1}{q} \int_0^T [\tilde{U}(\tau), [U(\tau), T_{q-3}(\tau)]] d\tau \\
&= \frac{1}{q} \left(\int_0^T [U(\tau), [\tilde{U}(\tau), T_{q-3}(\tau)]] d\tau + \int_0^T [\tilde{U}(\tau), [\tilde{U}(\tau), [U(\tau), T_{q-4}(\tau)]]] d\tau \right) \\
&\vdots
\end{aligned}
\tag{4.48}$$

The terms in these expressions will be proportional to *moments* and depth- $(q-1)$ structure constants. For example, the last term in the last expression will be of the form,

$$\begin{aligned}
&\int_0^T [\tilde{U}(\tau), [\tilde{U}(\tau), \dots [\tilde{U}(\tau), [U(\tau), \tilde{U}(\tau)]] \dots]] d\tau \\
&= \sum_{I_q=1}^m \left(\int_0^T \tilde{u}_{i_1}(\tau) \tilde{u}_{i_2}(\tau) \dots \tilde{u}_{i_{q-2}}(\tau) u_{i_{q-1}}(\tau) \tilde{u}_{i_q}(\tau) d\tau \right) \sum_{p=1}^n \Gamma_{I_q}^p \xi_p,
\end{aligned}$$

where the term in parentheses is interpreted as a moment. Thus, we call Theorem 4.11 an area-moment rule.

In the following corollaries we illustrate q th-order averaging for $q = 1, 2, 3$. For alternative proofs of these corollaries (under more restrictive hypotheses) see [43, 49, 47].

Corollary 4.14 ($q = 1$) Consider system (2.9). Assume $U(t)$ and b are as in Theorem 4.9 and $U(t + T) = U(t)$, $\forall t > 0$. Let $g(0) = g_0 \in \hat{Q} \subset G$ and $Z_0 = \hat{\Psi}^{-1}(g_0) = O(\epsilon)$. Define

$$Z^{(1)}(t) = \epsilon U_{av} t + Z_0^{(1)}, \quad g^{(1)}(t) = e^{Z^{(1)}(t)}, \quad (4.49)$$

where $U_{av} = \frac{1}{T} \int_0^T U(\tau) d\tau$. If $\|Z_0 - Z_0^{(1)}\| = O(\epsilon)$ and $Z^{(1)}(t) \in \hat{S}$, $\forall t \in [0, b/\epsilon]$ then,

$$\hat{d}(g(t), g^{(1)}(t)) = O(\epsilon), \quad \forall t \in [0, b/\epsilon]. \quad (4.50)$$

Proof: Since $Z_1(t) = \int_0^t U(\tau) d\tau$, by (4.36),

$$\begin{aligned} Z^{(1)}(t) &= \epsilon \frac{t}{T} Z_1(T) + Z_0^{(1)} \\ &= \epsilon t \frac{1}{T} \int_0^T U(\tau) d\tau + Z_0^{(1)} \\ &= \epsilon U_{av} t + Z_0^{(1)}. \end{aligned}$$

The corollary follows by Theorem 4.11. □

Remark 4.15 Apart from the different treatment of the initial condition, this first-order approximation of $g(t)$ is identical to that derived in Theorem 4.1 using the product of exponentials representation.

Corollary 4.16 ($q = 2$) Consider system (2.9). Assume $U(t)$ and b are as in Theorem 4.9 and $U(t + T) = U(t)$, $\forall t > 0$. Let $g(0) = g_0 \in \hat{Q} \subset G$ and $Z_0 =$

$\hat{\Psi}^{-1}(g_0) = O(\epsilon)$. Let $Z_0^{(2)} = \sum_{k=1}^n z_{k0}^{(2)} \xi_k$ and assume that $\|Z_0 - Z_0^{(2)}\| = O(\epsilon^2)$.

Let

$$z_k^{[2]}(t) = \epsilon \tilde{u}_k(t) + \epsilon^2 \sum_{i,j=1; i < j}^m a_{ij}(t) \Gamma_{ij}^k + z_{k0}^{(2)}. \quad (4.51)$$

Then,

$$z_k^{[2]}(NT) = \epsilon NT u_{avk} + \epsilon^2 N \sum_{i,j=1; i < j}^m Area_{ij}(T) \Gamma_{ij}^k + z_{k0}^{(2)}, \quad (N \text{ an integer}) \quad (4.52)$$

Γ_{ij}^k , $Area_{ij}(T)$, $a_{ij}(t)$ are defined by (2.15), (2.22), (2.23), respectively. Define

$$Z^{[2]}(t) = \sum_{k=1}^n z_k^{[2]}(t) \xi_k, \quad g^{[2]}(t) = e^{Z^{[2]}(t)}. \quad (4.53)$$

If $Z^{[2]}(t) \in \hat{S}$, $\forall t \in [0, b/\epsilon]$, then

$$\hat{d}(g(t), g^{[2]}(t)) = O(\epsilon^2), \quad \forall t \in [0, b/\epsilon]. \quad (4.54)$$

Suppose that $U_{av} = 0$. Define

$$z_k^{(2)}(t) = \epsilon \tilde{u}_k(t) + \epsilon^2 \frac{t}{T} \sum_{i,j=1; i < j}^m Area_{ij}(T) \Gamma_{ij}^k + z_{k0}^{(2)}, \quad (4.55)$$

$$Z^{(2)}(t) = \sum_{k=1}^n z_k^{(2)}(t) \xi_k, \quad g^{(2)}(t) = e^{Z^{(2)}(t)}. \quad (4.56)$$

If $Z^{(2)}(t) \in \hat{S}$, $\forall t \in [0, b/\epsilon]$, then

$$\hat{d}(g(t), g^{(2)}(t)) = O(\epsilon^2), \quad \forall t \in [0, b/\epsilon]. \quad (4.57)$$

Further, for $t = NT$, N an integer,

$$z_k^{(2)}(NT) = \epsilon^2 N \sum_{i,j=1; i < j}^m Area_{ij}(T) \Gamma_{ij}^k + z_{k0}^{(2)}. \quad (4.58)$$

Proof: From (4.32) of Theorem 4.10

$$Z^{[2]}(t) = \epsilon Z_1(t) - \epsilon^2 Z_2(t) + Z_0^{[2]}.$$

By (2.45), skew-symmetry of the Lie bracket and the definition of $a_{ij}(t)$ (2.23),

$$\begin{aligned}
Z_2(t) &= \frac{1}{2} \int_0^t [U(\tau), \tilde{U}(\tau)] d\tau \\
&= \frac{1}{2} \int_0^t \left[\sum_{i=1}^m u_i(\tau) \xi_i, \sum_{j=1}^m \tilde{u}_j(\tau) \xi_j \right] d\tau \\
&= \sum_{i,j=1}^m \frac{1}{2} \int_0^t u_i(\tau) \tilde{u}_j(\tau) [\xi_i, \xi_j] d\tau \\
&= \sum_{i,j=1; i < j}^m \frac{1}{2} \int_0^t (u_i(\tau) \tilde{u}_j(\tau) - u_j(\tau) \tilde{u}_i(\tau)) d\tau [\xi_i, \xi_j] \\
&= - \sum_{k=1}^n \left(\sum_{i,j=1; i < j}^m a_{ij}(t) \Gamma_{ij}^k \right) \xi_k. \tag{4.59}
\end{aligned}$$

Thus, since $Z_1(t) = \tilde{U}(t)$, $Z^{[2]}(t)$ and $g^{[2]}(t)$ as defined in Theorem 4.10 reduce to (4.51) and (4.53). When $t = NT$, N an integer,

$$\begin{aligned}
Z_1(NT) &= \int_0^{NT} U(\tau) d\tau \\
&= \int_0^T U(\tau) d\tau + \int_T^{2T} U(\tau) d\tau + \dots + \int_{(N-1)T}^{NT} U(\tau) d\tau \\
&= \int_0^T U(\tau) d\tau + \int_0^T U(\sigma_1 + T) d\sigma + \dots + \int_0^T U(\sigma_{N-1} + (N-1)T) d\sigma_{N-1} \\
&= N \int_0^T U(\tau) d\tau = NTU_{av}.
\end{aligned}$$

We show $Z_2(NT) = NZ_2(T)$ by induction. The relation holds trivially for $N = 1$. Suppose $Z_2(MT) = MZ_2(T)$ for M an integer. We need to show that $Z_2((M+1)T) = (M+1)Z_2(T)$.

$$\begin{aligned}
Z_2((M+1)T) &= \frac{1}{2} \int_0^{(M+1)T} [U(\tau_1), \int_0^{\tau_1} U(\tau_2) d\tau_2] d\tau_1 \\
&= \frac{1}{2} \int_0^{MT} [U(\tau_1), \int_0^{\tau_1} U(\tau_2) d\tau_2] d\tau_1 + \frac{1}{2} \int_{MT}^{(M+1)T} [U(\tau_1), \int_0^{\tau_1} U(\tau_2) d\tau_2] d\tau_1 \\
&= Z_2(MT) + \frac{1}{2} \int_0^T [U(\sigma_1 + MT), \int_0^{\sigma_1 + MT} U(\tau_2) d\tau_2] d\sigma_1 \\
&= MZ_2(T) + \frac{1}{2} \int_0^T [U(\sigma_1), \int_0^{MT} U(\tau_2) d\tau_2] d\sigma_1 + \frac{1}{2} \int_0^T [U(\sigma_1), \int_{MT}^{\sigma_1 + MT} U(\tau_2) d\tau_2] d\sigma_1 \\
&= MZ_2(T) + \frac{1}{2} \int_0^T [U(\sigma_1), MTU_{av}] d\sigma_1 + \frac{1}{2} \int_0^T [U(\sigma_1), \int_0^{\sigma_1} U(\sigma_2 + MT) d\sigma_2] d\sigma_1
\end{aligned}$$

$$\begin{aligned}
&= MZ_2(T) + \frac{1}{2}M\left[\int_0^T [U(\sigma_1)d\sigma_1, TU_{av}] + \frac{1}{2}\int_0^T [U(\sigma_1), \int_0^{\sigma_1} U(\sigma_2)d\sigma_2]d\sigma_1\right. \\
&= MZ_2(T) + \frac{1}{2}M[TU_{av}, TU_{av}] + Z_2(T) \\
&= (M+1)Z_2(T).
\end{aligned}$$

Therefore, $Z_2(NT) = NZ_2(T)$ by induction and so

$$Z^{[2]}(NT) = \epsilon NTU_{av} - \epsilon^2 NZ_2(T) + Z_0^{(2)}.$$

From (4.59),

$$\begin{aligned}
Z_2(T) &= -\sum_{k=1}^n \left(\sum_{i,j=1; i < j}^m a_{ij}(T) \Gamma_{ij}^k \right) \xi_k \\
&= -\sum_{k=1}^n \left(\sum_{i,j=1; i < j}^m Area_{ij}(T) \Gamma_{ij}^k \right) \xi_k.
\end{aligned} \tag{4.60}$$

Thus, (4.52) follows. (4.54) follows from Theorem 4.10. For the second part of the corollary we note that by (4.36) and (4.60) in Theorem 4.11,

$$\begin{aligned}
Z^{(2)}(t) &= \epsilon Z_1(t) - \epsilon^2 \frac{t}{T} Z_2(T) + Z_0^{(2)} \\
&= \sum_{k=1}^n (\epsilon \tilde{u}_k(t) + \epsilon^2 \frac{t}{T} \sum_{i,j=1; i < j}^m Area_{ij}(T) \Gamma_{ij}^k + z_{k0}^{(2)}) \xi_k.
\end{aligned}$$

Since $T_0(T) = U_{av} = 0$, the assumptions of Theorem 4.11 are satisfied. (4.57)

and (4.58) follow from Theorem 4.11. $\square$

Remark 4.17 We note that $\gamma_k^{(2)}(t)$ of (4.9) is identical to $z_k^{(2)}(t)$ of (4.55). Thus, both the product of exponentials $e^{z_1^{(2)}(t)\xi_1} \dots e^{z_n^{(2)}(t)\xi_n}$ and the single exponential $e^{z_1^{(2)}(t)\xi_1 + \dots + z_n^{(2)}(t)\xi_n}$ provide an $O(\epsilon^2)$ approximation to $g(t)$.

Corollary 4.18 (q = 3) Consider system (2.9). Assume $U(t)$ and b are as in Theorem 4.9 and $U(t+T) = U(t)$, $\forall t > 0$. Let $g(0) = g_0 \in \hat{Q} \subset G$ and $Z_0 =$

$\hat{\Psi}^{-1}(g_0) = O(\epsilon^2)$. Let $Z_0^{(3)} = \sum_{k=1}^n z_{k0}^{(3)} \xi_k$ and assume that $\|Z_0 - Z_0^{(3)}\| = O(\epsilon^3)$.

Suppose that $U_{av} = 0$ and $Area_{ij}(T) = 0, \forall i, j = 1, \dots, n$. Define

$$z_p^{(3)}(t) = \epsilon \tilde{u}_p(t) + \sum_{i,j=1; i < j}^m (\epsilon^2 a_{ij}(t) \Gamma_{ij}^k - \epsilon^3 \frac{t}{T} \sum_{k=1}^m m_{ijk}(T) \theta_{ijk}^p) + z_{p0}^{(3)} \quad (4.61)$$

$$\begin{aligned} &= \epsilon \tilde{u}_p(t) + \sum_{i,j=1; i < j}^m (\epsilon^2 a_{ij}(t) \Gamma_{ij}^k - \epsilon^3 \frac{t}{T} m_{iji}(T) \theta_{iji}^p - \\ &\quad \sum_{k=i+1}^m (2m_{ijk}(T) - m_{ikj}(T)) \theta_{ijk}^p) + z_{p0}^{(3)}. \end{aligned} \quad (4.62)$$

$$Z^{(3)}(t) = \sum_{p=1}^n z_p^{(3)}(t) \xi_p, \quad g^{(3)}(t) = e^{Z^{(3)}(t)}. \quad (4.63)$$

$\theta_{ijk}^p, Area_{ij}(T), a_{ij}(t), m_{ijk}(T)$ are defined by (2.16), (2.22), (2.23), (2.26), respectively. If $Z^{(3)}(t) \in \hat{S}, \forall t \in [0, b/\epsilon]$ then,

$$\hat{d}(g(t), g^{(3)}(t)) = O(\epsilon^3), \quad \forall t \in [0, b/\epsilon]. \quad (4.64)$$

Further, for $t = NT, N$ an integer,

$$z_p^{(3)}(NT) = -\epsilon^3 N \sum_{i,j=1; i < j}^m (m_{iji}(T) \theta_{iji}^p + \sum_{k=i+1}^m (2m_{ijk}(T) - m_{ikj}(T)) \theta_{ijk}^p) + z_{p0}^{(3)}. \quad (4.65)$$

Proof: By (4.36) of Theorem 4.11,

$$Z^{(3)}(t) = \epsilon Z_1(t) - \epsilon^2 Z_2(t) + \epsilon^3 \frac{t}{T} Z_3(T) + Z_0^{(3)}.$$

$Z_1(t) = \tilde{U}(t)$ and $Z_2(t)$ is given by (4.59). $U_{av} = 0$ implies $T_0(T) = 0$. By (4.60) $Area_{ij}(T) = 0, \forall i, j = 1, \dots, n$, implies $T_1(T) = 2Z_2(T) = 0$, as needed for Theorem 4.11. By (4.48) of Remark 4.13, skew-symmetry of the Lie bracket and the definition of $m_{ijk}(T)$ (2.26),

$$Z_3(T) = -\frac{1}{3} \int_0^T [\tilde{U}(\tau), [U(\tau), \tilde{U}(\tau)]] d\tau$$

$$\begin{aligned}
&= -\frac{1}{3} \int_0^T [[\tilde{U}(\tau), U(\tau)], \tilde{U}(\tau)] d\tau \\
&= -\frac{1}{3} \int_0^T [[\sum_{i=1}^m \tilde{u}_i(\tau) \xi_i, \sum_{j=1}^m u_j(\tau) \xi_j], \sum_{k=1}^m \tilde{u}_k(\tau) \xi_k] d\tau \\
&= -\sum_{i,j,k=1}^m \frac{1}{3} \int_0^T \tilde{u}_i(\tau) u_j(\tau) \tilde{u}_k(\tau) [[\xi_i, \xi_j], \xi_k] d\tau \\
&= -\sum_{i,j=1; i < j}^m \sum_{k=1}^m \frac{1}{3} \int_0^T (\tilde{u}_i(\tau) u_j(\tau) - \tilde{u}_j(\tau) u_i(\tau)) \tilde{u}_k(\tau) d\tau [[\xi_i, \xi_j], \xi_k] d\tau \\
&= -\sum_{p=1}^n \sum_{i,j=1; i < j}^m \sum_{k=1}^m m_{ijk}(T) \theta_{ijk}^p \xi_p.
\end{aligned}$$

From equations (2.17) and (2.29) (skew symmetry of Lie bracket and Jacobi identity),

$$\theta_{jki}^p = \theta_{ikj}^p - \theta_{ijk}^p, \quad m_{jki}(T) = m_{ikj}(T) - m_{ijk}(T).$$

So,

$$\begin{aligned}
&m_{ijk}(T) \theta_{ijk}^p + m_{ikj}(T) \theta_{ikj}^p + m_{jki}(T) \theta_{jki}^p \\
&= m_{ijk}(T) \theta_{ijk}^p + m_{ikj}(T) \theta_{ikj}^p + (m_{ikj}(T) - m_{ijk}(T)) (\theta_{ikj}^p - \theta_{ijk}^p) \\
&= (2m_{ijk}(T) - m_{ikj}(T)) \theta_{ijk}^p + (2m_{ikj}(T) - m_{ijk}(T)) \theta_{ikj}^p.
\end{aligned}$$

As a result, we can eliminate $m_{jki}(T) \theta_{jki}^p$ from our sum as follows:

$$\sum_{k=1}^m \sum_{i,j=1; i < j}^m m_{ijk}(T) \theta_{ijk}^p = \sum_{i,j=1; i < j}^m (m_{iji}(T) \theta_{iji}^p + \sum_{k=i+1}^m (2m_{ijk}(T) - m_{ikj}(T)) \theta_{ijk}^p).$$

Thus, $Z^{(3)}(t)$ is given by (4.61) - (4.63). (4.64) and (4.65) follow from Theorem 4.11. □

Remark 4.19 The formulas in the theorems and corollaries of this section are clearly basis-independent.

4.3 Curvature Form

In this section we show that the secular term in the area rule of Corollary 4.16 derives from the curvature form of a certain principal fiber bundle with connection. The purpose of this derivation is to emphasize that the formulas in our averaging theory are basis-independent. In Section 4.3.1 we define principal fiber bundles, connections and curvature forms. In Section 4.3.2 we discuss a particular principal bundle with connection and derive the secular term of $Z^{(2)}(t)$ from the associated curvature form.

4.3.1 Principal Fiber Bundles and Connections

Principal fiber bundles, connections and curvature forms are defined in this section. Nomizu [62] and Bleeker [8] are our main references. In these references, definitions are made with respect to *right* actions of the structure group on the total space whereas we give definitions based on *left* actions. We follow Yang [79] for these definitions.

A *principal fiber bundle* is denoted by $\mathcal{P} = (Q, B, \pi, G)$. Q and B , smooth manifolds, are the *total space* and *base space*, respectively. G , a Lie group, is the *structure group*. $\mathcal{P}$ satisfies the following conditions:

- (a) There is a (smooth) left action Φ of G on Q as defined by (2.2). Further, Φ is a free action, i.e., if $\Phi(g, q) = q$, $q \in Q$, then $g = e \in G$.
- (b) $B = Q/G$ and $\pi : Q \rightarrow B$ is the *canonical projection map* and is differentiable.
- (c) Q is *locally trivial*. This means that for each $x \in B$ there is a neighborhood W of x such that $\pi^{-1}(W) \subset Q$ is isomorphic to $W \times G$. The map $q \in$

$\pi^{-1}(W) \mapsto (\pi(q), \phi(q)) \subset W \times G$ is a diffeomorphism where $\phi : \pi^{-1}(W) \rightarrow G$ satisfies $\phi(g \cdot q) = g\phi(q)$, $\forall g \in G$. This diffeomorphism, denoted by ϕ_T , is called a *local trivialization*.

For $x \in B$, $\pi^{-1}(x)$ is the *fiber over x* and is isomorphic to G . For $q \in Q$, the *fiber through q* is the fiber over $x = \pi(q)$. If $Q = B \times G$ then Q is a *trivial* principal fiber bundle. In this case, the diffeomorphism ϕ_T of (c) above is called a *global trivialization*. For a trivial bundle the action of G on Q is $\Phi(g, (x, h)) = (x, gh)$ for $x \in B$, $g, h \in G$.

We define a *local section* of $\mathcal{P}$ to be a map $\sigma : W \subset B \rightarrow Q$, where W is open in B , such that $\pi \circ \sigma$ is the identity on W , i.e., $\pi \circ \sigma(x) = x$. If $W = B$ then σ is called a *global section*. There is a natural correspondence between local (global) trivializations and local (global) sections. For instance, in the case of a trivial bundle, given σ define $\phi_T(g\sigma(x)) = (x, g)$. Given ϕ_T define $\sigma(x) = \phi_T^{-1}(x, e)$.

Let V be an r -dimensional vector space with basis $v_1, \dots, v_r$. Let $\Lambda^k(Q)$ be the space of real-valued k -forms on Q . Let $\beta_1, \dots, \beta_r \in \Lambda^k(Q)$. Then, $\sum_{i=1}^r \beta_i v_i$ is a V -valued k -form. We denote $\Lambda^k(Q; V)$ as the space of V -valued k -forms on Q . Let $\mathcal{X}(Q)$ denote the space of vector fields on Q .

There are at least three (equivalent) ways to define a connection on a principal fiber bundle. We give the definition most useful for our purposes. Let $\mathcal{G}$ be the Lie algebra of G . A *connection form* on a principal fiber bundle $\mathcal{P} = (Q, B, \pi, G)$ is a $\mathcal{G}$ -valued 1-form on Q , $\omega \in \Lambda^1(Q; \mathcal{G})$, such that

- (1) $\omega(\xi_Q(q)) = \xi$, $\xi \in \mathcal{G}$;
- (2) $((\Phi_g)^*\omega)(Y) = \text{Ad}_g \omega(Y)$, $Y \in \mathcal{X}(Q)$.

$\xi_Q(q)$ is defined by (2.4), Ad is defined by (2.3) and $(\Phi_g)^*$ is the dual of $(\Phi_g)_* = T\Phi_g$, the differential map of Φ_g .

The *connection* on $\mathcal{P}$ defines a splitting of T_qQ , the tangent space to Q at $q \in Q$. To each point $q \in Q$, the connection assigns a *horizontal subspace* $H_q \subset T_qQ$ which together with the *vertical subspace* defined by $V_q \triangleq \{v \in T_qQ \mid T\pi(v) = 0\}$ satisfies

$$(a) \quad T_qQ = H_q \oplus V_q,$$

$$(b) \quad T_q\Phi_g \cdot H_q = H_{g \cdot q}, \quad \forall g \in G, \quad q \in Q.$$

$$(c) \quad H_q \text{ depends differentiably on } q.$$

A selection of a connection form ω that satisfies (1) and (2) above is equivalent to a unique splitting of T_qQ , $\forall q \in Q$, where $H_q = \text{Ker}(\omega(q))$, $\forall q \in Q$ and Ker stands for kernel. Let $Y \in \mathcal{X}(Q)$, then we can write for any $q \in Q$

$$Y(q) = Y^v(q) + Y^h(q), \quad Y^v(q) \in V_q, \quad Y^h(q) \in H_q.$$

We can then define smooth vector fields Y^v and Y^h such that Y^v maps $q \mapsto Y^v(q)$ and Y^h maps $q \mapsto Y^h(q)$. We say that Y^v and Y^h are the *vertical component* of Y and the *horizontal component* of Y , respectively.

Let $\omega \in \Lambda^k(Q; \mathcal{G})$. The *covariant derivative* of ω , $D\omega \in \Lambda^{k+1}(Q; \mathcal{G})$, is defined by

$$D\omega(Y_0, \dots, Y_k) \triangleq d\omega(Y_0^h, \dots, Y_k^h), \quad Y_i \in \mathcal{X}(Q).$$

The *curvature form* $\Gamma \in \Lambda^2(Q; \mathcal{G})$ of the connection where $\omega \in \Lambda^1(Q; \mathcal{G})$ is the connection form is defined by

$$\Gamma = D\omega. \tag{4.66}$$

The connection form and its corresponding curvature form have the following three important properties for $Y_1, Y_2 \in \mathcal{X}(Q)$:

$$\begin{aligned} ((\Phi_g)^*\Gamma)(Y_1, Y_2) &= \text{Ad}_g \Gamma(Y_1, Y_2). \\ \Gamma(Y_1, Y_2) &= d\omega(Y_1, Y_2) + [\omega(Y_1), \omega(Y_2)]. \\ \Gamma(Y_1, Y_2) &= -\omega([Y_1^h, Y_2^h]). \end{aligned} \tag{4.67}$$

A smooth curve $q(\cdot) = \{q(t), t \in [0, 1]\} \subset Q$ is a *horizontal curve* if $dq(t)/dt \in H_{q(t)}$, $\forall t \in [0, 1]$. Let $x(\cdot) = \{x(t), t \in [0, 1]\} \subset B$ be a piecewise smooth curve. A horizontal curve $q(\cdot)$ is a *horizontal lift* of $x(\cdot)$ if $\pi(q(t)) = x(t)$, $\forall t \in [0, 1]$. An important theorem (c.f. [62]) states that given $x(\cdot)$ and $q_0 \in Q$ where $\pi(q_0) = x(0)$, there is a *unique* horizontal lift $q(\cdot)$ which starts at $q(0) = q_0$. The map $\psi : q_0 = q(0) \mapsto q_1 = q(1)$ is called the *parallel displacement* and is a differentiable isomorphism of the fiber through $q(0)$ onto the fiber through $q(1)$. Consider $x(\cdot)$ where $x(0) = x(1) = x_0$, then the parallel displacement is an automorphism of $\pi^{-1}(x_0)$. Ψ_{x_0} is the set of all such automorphisms and forms a group called the *holonomy group* at x_0 . An element of Ψ_{x_0} is called the *holonomy* at x_0 .

Let $x(\cdot) \subset W \subset B$, W open, be a closed curve, i.e., $x(0) = x(1) = x_0$. Let $q_0 \in \pi^{-1}(x_0) \subset Q$. Let $\sigma : W \rightarrow Q$ be a local section of the bundle. Then $\sigma(x(\cdot))$ is a curve in Q . Let $\omega \in \Lambda^1(Q; \mathcal{G})$ define the connection, and let $q(\cdot)$ be the horizontal lift of $x(\cdot)$ with $q(0) = q_0$. Let $g(\cdot) = \{g(t), t \in [0, 1]\} \subset G$ be defined such that $q(t) = \Phi(g(t), \sigma(x(t)))$, $\forall t \in [0, 1]$. Then $g(t)$ is the solution of the following differential equation on G [79]:

$$\frac{dg(t)}{dt} = -T_e L_g(\sigma^* \omega)(\dot{x}(t)). \tag{4.68}$$

We say that $g(1)$ is the *holonomy* at q_0 with respect to $x(\cdot)$

Remark 4.20 Equation (4.68) is a left-invariant system on the Lie group G of the form (2.9) since $-(\sigma^*\omega)(\dot{x}(t))$ is a curve in $\mathcal{G}$. Accordingly, the averaging theory of this chapter provides an average approximation of the holonomy, i.e., of the solution $g(t)$ to (4.68). For control problems we interpret $\dot{x}$ as our periodically, time-varying control input and $g(t)$ as the resulting drift of the system. For example, recall the spacecraft with an appended point mass oscillator of Section 3.1.2. In this example the control input is the velocity of the oscillator $\dot{y} \in \mathbb{R}^3$. The configuration space of the rigid part of the spacecraft is $G = SO(3)$. We can describe this system by the principal fiber bundle $(\mathbb{R}^3 \times SO(3), \mathbb{R}^3, \pi, SO(3))$ with a (mechanical) connection. For this bundle $\mathbb{R}^3$ is the base space and $\mathbb{R}^3 \times SO(3)$ the total space. Then the spacecraft angular velocity Ω given by (3.8) can be derived from $\hat{\Omega} = -(\sigma^*\omega)(\dot{y}(t))$ and the holonomy equation (4.68) is $\dot{X} = X\hat{\Omega}$, $X(t) \in SO(3)$. In the next section we consider a principal fiber bundle with connection where the control input $\dot{x}$ is given by $\dot{x} = U(t) \in \mathcal{G}$ and $U(t) = -(\sigma^*\omega)(\dot{x})$. In this case, (4.68) specializes to (2.9).

4.3.2 Curvature Form and Area Rule

The tangent bundle TG of a Lie group G is trivially a vector bundle, i.e., $TG \approx G \times \mathcal{G}$, where the global trivialization is given by

$$TG \rightarrow G \times \mathcal{G}, \quad (g, v_g) \mapsto (g, (T_e L_g)^{-1} v_g).$$

One may also view this as the trivial principal fiber bundle $\mathcal{P} = (Q, B, \pi, G) = (\mathcal{G} \times G, \mathcal{G}, \pi, G)$. The action of G on $Q = \mathcal{G} \times G$ is defined by

$$\Phi : G \times \mathcal{G} \times G \rightarrow \mathcal{G} \times G, \quad \Phi(g, (x, h)) = (x, gh).$$

The projection map π is defined by

$$\pi : \mathcal{G} \times G \rightarrow \mathcal{G}, \quad \pi(x, h) = x.$$

The fiber over $x \in \mathcal{G}$ is $\pi^{-1}(x) = (x, h)$, $h \in G$, and the fiber through $(x, g) \in \mathcal{G} \times G$ is (x, h) , $h \in G$, which is isomorphic to G .

Now consider a closed curve $x(\cdot) = \{x(t), t \in [0, T]\} \subset \mathcal{G}$ and a point $q_0 \in \pi^{-1}(x(0)) \subset \mathcal{G} \times G$. Since $\mathcal{P}$ is a trivial bundle, we consider a global section, $\sigma : \mathcal{G} \rightarrow \mathcal{G} \times G$ defined by $\sigma(y) = (y, e)$, $y \in \mathcal{G}$, $e \in G$. Let $\omega \in \Lambda^1(Q; \mathcal{G})$ be the connection form for a connection on $\mathcal{P}$. Then $\omega : T(\mathcal{G} \times G) \rightarrow \mathcal{G}$ and $\sigma^*\omega : T\mathcal{G} \rightarrow \mathcal{G}$. By (4.68) the holonomy at q_0 with respect to $x(\cdot)$ is the solution of the following equation at $t = T$:

$$\dot{g} = -T_e L_g(\sigma^*\omega)(\dot{x}(t)). \quad (4.69)$$

Suppose we let $\sigma^*\omega$ be defined by $(\sigma^*\omega)_x = -dx$, i.e.,

$$\sigma^*\omega : T\mathcal{G} \rightarrow \mathcal{G}, \quad \sigma^*\omega(\xi) = -\xi. \quad (4.70)$$

Then,

$$\dot{g} = T_e L_g \cdot \dot{x}. \quad (4.71)$$

The selection of $\sigma^*\omega$ uniquely defines the connection form ω , which must have the form [57]:

$$\begin{aligned} \omega_{(x,g)} &= T_g R_g^{-1} dg - \text{Ad}_g dx \\ &= T_g R_g^{-1} dg - T_g R_g^{-1} T_e L_g dx \\ &= T_g R_g^{-1} (dg - T_e L_g dx). \end{aligned} \quad (4.72)$$

This can be confirmed by recalling from Section 4.3.1 that horizontal curves must be in the kernel of $\omega_{(x,g)}$. The kernel of $\omega_{(x,g)}$ given by (4.72) is defined by (4.71).

Further, we can show that $\omega_{(x,g)}$ is a valid connection form, i.e., that it satisfies **(1)** and **(2)** of the definition of a connection form in Section 4.3.1. First, we have that for $q = (x, g) \in \mathcal{G} \times G$, $\xi \in \mathcal{G}$,

$$\begin{aligned}
\xi_Q(q) &\triangleq \left. \frac{d}{dt} \right|_{t=0} \Phi(\exp(t\xi), q) \\
&= \left. \frac{d}{dt} \right|_{t=0} (x, \exp(t\xi)g) \\
&= (0, \left. \frac{d}{dt} \right|_{t=0} (R_g \exp(t\xi))) \\
&= (0, T_e R_g \left. \frac{d}{dt} \right|_{t=0} \exp(t\xi)) \\
&= (0, T_e R_g \cdot \xi).
\end{aligned}$$

So,

$$\begin{aligned}
\omega_{(x,g)}(0, T_e R_g \cdot \xi) &= \langle (T_g R_g^{-1} dg - \text{Ad}_g dx), (0, T_e R_g \cdot \xi) \rangle \\
&= -\text{Ad}_g dx \cdot 0 + (T_g R_g^{-1})(T_e R_g \cdot \xi) \\
&= \xi.
\end{aligned}$$

This confirms that $\omega_{(x,g)}$ satisfies **(1)**. Secondly, for $g_1 \in G$ and $Y \in \mathcal{X}(Q)$, we have that

$$\begin{aligned}
(\Phi_{g_1})^* \omega_{(x,g)}(Y) &= \omega_{\Phi_{g_1}(x,g)}((\Phi_{g_1})_* Y) \\
&= \omega_{(x, g_1 g)}((\Phi_{g_1})_* Y) \\
&= (T_{g_1 g} R_{g_1 g}^{-1} d(g_1 g) - \text{Ad}_{g_1 g} dx)((\Phi_{g_1})_* Y) \\
&= (T_{g_1 g} (R_{g_1}^{-1} R_g^{-1})(T_g L_{g_1}) dg - T_e (R_{g_1 g}^{-1} L_{g_1 g}) dx)((\Phi_{g_1})_* Y) \\
&= (T_g (R_{g_1}^{-1} R_g^{-1} L_{g_1}) dg - T_e (R_{g_1}^{-1} R_g^{-1} L_{g_1} L_g) dx)((\Phi_{g_1})_* Y) \\
&= (T_g (R_{g_1}^{-1} L_{g_1} R_g^{-1}) dg - T_e (R_{g_1}^{-1} L_{g_1} R_g^{-1} L_g) dx)((\Phi_{g_1})_* Y) \\
&= (T_e (R_{g_1}^{-1} L_{g_1})(T_g R_g^{-1}) dg - T_e (R_{g_1}^{-1} L_{g_1}) T_e (R_g^{-1} L_g) dx)((\Phi_{g_1})_* Y)
\end{aligned}$$

$$\begin{aligned}
&= \text{Ad}_{g_1}(T_g R_g^{-1} dg - \text{Ad}_g dx)((\Phi_{g_1})_* Y) \\
&= \text{Ad}_{g_1} \omega_{(x,g)}((\Phi_{g_1})_* Y).
\end{aligned}$$

This confirms that $\omega_{(x,g)}$ satisfies **(2)**.

We note that the differential map of our chosen global section is

$$\sigma_*(\xi) = (\xi, 0), \quad \xi \in T\mathcal{G} = \mathcal{G}, \quad 0 \in TG.$$

So we can verify that

$$\sigma^*(\xi) = \omega(\sigma_*(\xi)) = \omega((\xi, 0)) = -\xi.$$

Similarly, we can compute the pullback of the curvature form Γ for $y, z \in T\mathcal{G} = \mathcal{G}$ as

$$\Gamma^*(y, z) = \Gamma(\sigma_* y, \sigma_* z) = \Gamma((y, 0), (z, 0)). \quad (4.73)$$

Based on our definition of connection form, the tangent vectors $(y, 0), (z, 0) \in T\mathcal{G} \times TG$ can be expressed as the sum of horizontal and vertical components, i.e.,

$$(y, 0) = (y, 0)^h + (y, 0)^v = (y, T_e L_g y) + (0, -T_e L_g y),$$

$$(z, 0) = (z, 0)^h + (z, 0)^v = (z, T_e L_g z) + (0, -T_e L_g z),$$

Thus, using the third identity of (4.67) we get

$$\begin{aligned}
\Gamma((y, 0), (z, 0)) &= -\omega([(y, 0)^h, (z, 0)^h]) \\
&= -\omega([(y, T_e L_g y), (z, T_e L_g z)]) \\
&= -\omega([y, z], [T_e L_g y, T_e L_g z]) \\
&= -\omega([y, z], T_e L_g [[y, z]]). \quad (4.74)
\end{aligned}$$

Here we used the notation $[[\cdot, \cdot]]$ for the Lie bracket on $\mathcal{G}$ to distinguish it from the Lie bracket of vector fields. We can express the vector field

$$Y \triangleq ([y, z], T_e L_g [[y, z]]) \in T\mathcal{G} \times TG$$

as the sum of horizontal and vertical components,

$$Y = Y^h + Y^v = ([y, z], T_e L_g[y, z]) + (0, T_e L_g[[y, z]] - T_e L_g[y, z]).$$

By definition, $\omega(Y^h) = 0$. So

$$\begin{aligned} -\omega([y, z], T_e L_g[[y, z]]) &= -\omega((0, T_e L_g[[y, z]] - T_e L_g[y, z])) \\ &= -(T_g R_g^{-1} dg - \text{Ad}_g dx)(0, T_e L_g[[y, z]] - T_e L_g[y, z]) \\ &= -T_g R_g^{-1} T_e L_g([y, z] - [y, z]). \end{aligned}$$

When evaluated at $g = e$,

$$-\omega([y, z], T_e L_g[[y, z]]) = -[[y, z]] + [y, z]. \quad (4.75)$$

Thus, by (4.73) - (4.75)

$$\Gamma^*(y, z) = -[[y, z]] + [y, z]. \quad (4.76)$$

Now let $U(t) = \sum_{i=1}^n u_i(t) \xi_i \in \mathcal{G}$ ($u_{m+1} = \dots = u_n = 0$, $m \leq n$) be a piecewise continuous curve with $U(t+T) = U(t)$, $\forall t > 0$ and $U_{av} = 0$ (defined by (2.21)). Let $\tilde{U}(t) = \sum_{i=1}^n \tilde{u}_i(t) \xi_i$ where $\tilde{u}_i$ are defined by (2.20) and $\{\xi_1, \dots, \xi_n\}$ is a basis for $\mathcal{G}$. Then $\tilde{U}(t+T) = \tilde{U}(t)$, $\forall t > 0$ and $\dot{\tilde{U}} = U$. Let $x(t) = \epsilon \tilde{U}(t) \in \mathcal{G}$, then (4.69) and (4.71) become, respectively,

$$\dot{g} = -T_e L_g(\sigma^* \omega)(\epsilon U(t)), \quad (4.77)$$

$$\dot{g} = \epsilon T_e L_g \cdot U(t). \quad (4.78)$$

Equation (4.78) is our main system of interest on the Lie group G , i.e., it is identical to (2.9).

Proposition 4.21 Let C be the closed curve in $\mathcal{G}$ traced out by $\epsilon\tilde{u}(t)$ from $t = 0$ to $t = T$. Let D be a surface in $\mathcal{G}$ bounded by the curve C . Then

$$\iint_D \sigma^* \Gamma = - \sum_{k=1}^n \sum_{i,j=1; i < j}^m \text{Area}_{ij}(T) \Gamma_{ij}^k \xi_k,$$

where Γ is the curvature form associated with the connection form ω defined by (4.72).

Proof. Let μ be the parametrization defined by

$$\begin{aligned} \mu : \Delta \subset \mathbb{R}^2 &\rightarrow D \subset \mathcal{G} \\ (x, y) &\mapsto (\epsilon\tilde{u}_1, \dots, \epsilon\tilde{u}_n). \end{aligned}$$

Then,

$$\begin{aligned} \iint_D \sigma^* \Gamma &= \iint_{\Delta} \mu^* \sigma^* \Gamma \\ &= \iint_{\Delta} \sigma^* \Gamma (d\mu \circ \frac{\partial}{\partial x}, d\mu \circ \frac{\partial}{\partial y}) dx \wedge dy \\ &= \iint_{\Delta} \sigma^* \Gamma (\frac{\partial(\epsilon\tilde{U})}{\partial x}, \frac{\partial(\epsilon\tilde{U})}{\partial y}) dx \wedge dy. \end{aligned} \quad (4.79)$$

By (4.76)

$$\begin{aligned} \sigma^* \Gamma (\frac{\partial(\epsilon\tilde{U})}{\partial x}, \frac{\partial(\epsilon\tilde{U})}{\partial y}) &= -[[\frac{\partial(\epsilon\tilde{U})}{\partial x}, \frac{\partial(\epsilon\tilde{U})}{\partial y}]] + [\frac{\partial(\epsilon\tilde{U})}{\partial x}, \frac{\partial(\epsilon\tilde{U})}{\partial y}] \\ &= -[[\frac{\partial(\epsilon\tilde{U})}{\partial x}, \frac{\partial(\epsilon\tilde{U})}{\partial y}]], \end{aligned} \quad (4.80)$$

since coordinate vectors commute. So substituting (4.80) into (4.79) we get

$$\begin{aligned} \iint_D \sigma^* \Gamma &= - \iint_{\Delta} [\frac{\partial(\epsilon\tilde{U})}{\partial x}, \frac{\partial(\epsilon\tilde{U})}{\partial y}] dx \wedge dy \\ &= - \iint_{\Delta} [\sum_{i=1}^n \epsilon \frac{\partial \tilde{u}_i}{\partial x} \xi_i, \sum_{j=1}^n \epsilon \frac{\partial \tilde{u}_j}{\partial y} \xi_j] dx \wedge dy \\ &= - \sum_{i,j=1}^n \epsilon^2 \iint_{\Delta} \frac{\partial \tilde{u}_i}{\partial x} \frac{\partial \tilde{u}_j}{\partial y} dx \wedge dy [\xi_i, \xi_j] \\ &= - \sum_{i,j=1; i < j}^n \epsilon^2 \iint_{\Delta} (\frac{\partial \tilde{u}_i}{\partial x} \frac{\partial \tilde{u}_j}{\partial y} - \frac{\partial \tilde{u}_j}{\partial x} \frac{\partial \tilde{u}_i}{\partial y}) dx \wedge dy [\xi_i, \xi_j], \end{aligned} \quad (4.81)$$

where we returned to the notation $[\cdot, \cdot]$ for the Lie bracket on $\mathcal{G}$ and we used the fact that $[\xi_i, \xi_j] = -[\xi_j, \xi_i]$. Let C_{ij} be the closed curve described by $\tilde{u}_i(t)$ and $\tilde{u}_j(t)$, $t = 0$ to $t = T$. Let Δ_{ij} be the area bounded by C_{ij} . We have

$$\begin{aligned} \iint_{\Delta_{ij}} d\tilde{u}_i \wedge d\tilde{u}_j &= \iint_{\Delta} \left(\frac{\partial \tilde{u}_i}{\partial x} dx + \frac{\partial \tilde{u}_i}{\partial y} dy \right) \wedge \left(\frac{\partial \tilde{u}_j}{\partial x} dx + \frac{\partial \tilde{u}_j}{\partial y} dy \right) \\ &= \iint_{\Delta} \left(\frac{\partial \tilde{u}_i}{\partial x} \frac{\partial \tilde{u}_j}{\partial y} - \frac{\partial \tilde{u}_j}{\partial x} \frac{\partial \tilde{u}_i}{\partial y} \right) dx \wedge dy. \end{aligned} \quad (4.82)$$

From (2.22)

$$\begin{aligned} Area_{ij}(T) &= \frac{1}{2} \int_0^T (\tilde{u}_i(\sigma) \dot{\tilde{u}}_j(\sigma) - \tilde{u}_j(\sigma) \dot{\tilde{u}}_i(\sigma)) d\sigma \\ &= \frac{1}{2} \oint_{C_{ij}} (\tilde{u}_i d\tilde{u}_j - \tilde{u}_j d\tilde{u}_i) \\ &= \frac{1}{2} \iint_{\Delta_{ij}} (d\tilde{u}_i \wedge d\tilde{u}_j - d\tilde{u}_j \wedge d\tilde{u}_i) \\ &= \iint_{\Delta_{ij}} d\tilde{u}_i \wedge d\tilde{u}_j. \end{aligned} \quad (4.83)$$

So using (4.81) - (4.83), we have

$$\begin{aligned} \iint_D \sigma^* \Gamma &= - \sum_{i,j=1; i < j}^n \epsilon^2 \iint_{\Delta_{ij}} d\tilde{u}_i \wedge d\tilde{u}_j [\xi_i, \xi_j] \\ &= - \sum_{i,j=1; i < j}^n \epsilon^2 Area_{ij}(T) [\xi_i, \xi_j] \\ &= - \sum_{k=1}^n \sum_{i,j=1; i < j}^m \epsilon^2 Area_{ij}(T) \Gamma_{ij}^k \xi_k, \end{aligned} \quad (4.84)$$

since $Area_{ij}(T) = 0$ if $i > m$ or $j > m$. □

Remark 4.22 From Corollary 4.16 $g^{(2)}(t) = e^{Z^{(2)}(t)}$ and $Z^{(2)}(t) = \epsilon \tilde{U}(t) + \frac{t}{T} \sum_{k=1}^n \sum_{i,j=1; i < j}^m \epsilon^2 Area_{ij}(T) \Gamma_{ij}^k \xi_k + Z_0^{(2)}$. By Proposition 4.21,

$$Z^{(2)}(t) = \epsilon \tilde{U}(t) - \frac{t}{T} \iint_D \sigma^* \Gamma + Z_0^{(2)}.$$

4.4 Stability

In addition to providing approximate solutions to certain systems of differential equations, classical averaging theory can also be used to determine the stability of system equilibria. In this section we illustrate how to use the averaging theory on Lie groups derived in this chapter for stabilization of equilibria in systems on Lie groups of the form (2.9). We do so using classical averaging theory and the product of exponentials representation of solutions to (2.9) as in Section 4.1.

According to Lemma 2.3, $g(t) = e^{\gamma_1^{(1)}(t)\xi_1} \dots e^{\gamma_n^{(1)}(t)\xi_n}$ is a local representation of the solution to (2.9) on the n -dimensional Lie group G with Lie algebra $\mathcal{G}$. The parameters $\gamma(t) = (\gamma_1(t), \dots, \gamma_n(t))^T \in \mathfrak{R}^n$ satisfy

$$\dot{\gamma} = \epsilon M(\gamma)u(t). \quad (4.85)$$

To determine the stability of equilibria of system (2.9) we determine the stability of equilibria of system (4.85) using classical averaging theory. However, since (4.85) provides a valid representation of the local solution so long as $M(\gamma)$ is nonsingular (see proof of Lemma 2.3), the right hand side of equation (4.85) is zero only when $u(t) = 0$. Thus, a nontrivial discussion of equilibria of (4.85) requires that we introduce feedback into our system.

Consider equation (2.9) and suppose that $u = u(t, g)$, and $u(t + T, g) = u(t, g)$, $\forall t > 0$. Then (2.9) becomes

$$\dot{g} = \epsilon T_e L_g \cdot U(t, g), \quad U(t, g) = \sum_{i=1}^m u_i(t, g)\xi_i, \quad m \leq n, \quad (4.86)$$

where $g(t) \in G$ and $U(t, g) \in \mathcal{G}$. Assuming that $g(t) \in Q$ during the time interval of interest, then $\gamma(t) = \Psi^{-1}(g(t))$ is well-defined. So we can write $u = u(t, \gamma) \triangleq u(t, \Psi(\gamma))$, and equation (4.85) becomes

$$\dot{\gamma} = \epsilon M(\gamma)u(t, \gamma). \quad (4.87)$$

We note that the nonautonomous vector field on G , defined by (4.86) is no longer left-invariant in general. This is a consequence of unrestricted feedback.

Following Theorem 4.1 the first-order average approximation of $\gamma(t)$ is $\gamma^{(1)}(t)$ defined by

$$\dot{\gamma}^{(1)} = \epsilon M(\gamma^{(1)}) \frac{1}{T} \int_0^T u(\tau, \gamma^{(1)}) d\tau. \quad (4.88)$$

If, for example, $\gamma^* \in S$ and $u(t, \gamma^*) = 0$, $\forall t \geq 0$, then γ^* is an equilibrium point of both (4.87) and (4.88) $\forall t \geq 0$. As shown in the theorem that follows, we can use averaging theory to draw conclusions about the local stability properties of (4.87) based on the stability properties of (4.88). Additionally, the existence of an exponentially stable equilibrium point for the averaged system (4.88) makes it possible to extend the approximations of Theorem 4.1 from an $O(1/\epsilon)$ time interval to an infinite time interval.

Theorem 4.23 (First-Order Averaging Extended) Consider system (4.86). Let $D = \{\gamma \in \mathfrak{R}^n \mid \|\gamma\| < r\} \subset S$. Assume that $u(t, \gamma)$ is periodic in t with period T and suppose that $M(\gamma)u(t, \gamma)$ is continuous and bounded with continuous and bounded derivatives up to second order with respect to both its arguments for $(t, \gamma) \in [0, \infty) \times D$. Suppose that $g(0) = g_0 \in Q$. Let $\gamma(t)$ be the solution to (4.87) with $\gamma(0) = \gamma_0 = \Psi^{-1}(g_0)$. Let $\gamma^{(1)}(t)$ be the solution to (4.88) with $\gamma^{(1)}(0) = \gamma_0^{(1)} \in D$. Define

$$g^{(1)}(t) = e^{\gamma_1^{(1)}(t)\xi_1} \dots e^{\gamma_n^{(1)}(t)\xi_n}.$$

If $\gamma^* \in D$ is an exponentially stable equilibrium point for (4.88) then $\exists \rho > 0$ such that if $\|\gamma_0^{(1)} - \gamma^*\| < \rho$ and $\|\gamma_0 - \gamma_0^{(1)}\| = O(\epsilon)$, then

$$\tilde{d}(g(t), g^{(1)}(t)) = O(\epsilon), \quad \forall t \in [0, \infty).$$

Further, $\exists \epsilon^*$ such that for all $0 < \epsilon < \epsilon^*$, (4.87) has a unique exponentially stable periodic solution of period T in an $O(\epsilon)$ neighborhood of γ^* . Similarly, (4.86) has a unique exponentially stable periodic solution of period T in an $O(\epsilon)$ neighborhood of $g^* = \Psi(\gamma^*) \in Q$.

Proof: By classical averaging theory (Theorem 7.4 of [32]),

$$\|\gamma(t) - \gamma^{(1)}(t)\| = O(\epsilon), \quad \forall t \in [0, \infty),$$

and $\exists \epsilon^*$ such that for all $0 < \epsilon < \epsilon^*$, (4.87) has a unique exponentially stable periodic solution of period T in an $O(\epsilon)$ neighborhood of γ^* .

Since $\gamma_0, \gamma^* \in D$ and γ^* is an exponentially stable equilibrium of (4.88), $\gamma^{(1)}(t) \in D, \forall t > 0$. For small enough ϵ this implies that $\gamma(t) \in D, \forall t > 0$. Thus, by Lemma 2.3,

$$g(t) = e^{\gamma_1(t)\xi_1} \dots e^{\gamma_n(t)\xi_n} = \Psi(\gamma(t)), \quad \forall t > 0.$$

Define

$$g^{(1)}(t) = e^{\gamma_1^{(1)}(t)\xi_1} \dots e^{\gamma_n^{(1)}(t)\xi_n} = \Psi(\gamma^{(1)}(t)), \quad \forall t > 0.$$

Then by definition of $\tilde{d}$

$$\tilde{d}(g(t), g^{(1)}(t)) = \|\gamma(t) - \gamma^{(1)}(t)\| = O(\epsilon), \quad \forall t \in [0, \infty).$$

Let $0 < \epsilon < \epsilon^*$ and let $\gamma_p(t)$ be the unique exponentially stable periodic solution of period T in the $O(\epsilon)$ neighborhood $N_\epsilon \subset D$ of γ^* . Let $g_p(t) = \Psi(\gamma_p(t)), \forall t > 0$. Let $g^* = \Psi(\gamma^*)$ and let $\Psi(N_\epsilon) = \{\Psi(x) \mid x \in N_\epsilon\} \subset Q$. Then $g_p(t)$ is a unique exponentially stable periodic solution of period T in the $O(\epsilon)$ neighborhood $\Psi(N_\epsilon)$ of g^* . $\square$

As an example consider the case $G = SO(3)$ for the attitude control problem described in Section 3.1. We recall that $\gamma = (\gamma_1, \gamma_2, \gamma_3)^T$ corresponds to a type of Euler angles. We suppose that we can measure γ . Choose $u_i(t, \gamma) = -2k_i\gamma_i \sin^2 t$, $k_i > 0$ for $i = 1, 2, 3$. Since $u(t, 0) = 0$, $\forall t$, then $\gamma^* = 0$ is an equilibrium point of (4.87) and (4.88). Also, $T = \pi$ and

$$\begin{aligned} \frac{1}{\pi} \int_0^\pi u(\tau, \gamma^{(1)}) d\tau &= \frac{1}{\pi} (-2k_i \gamma_i^{(1)}) \int_0^\pi \sin^2 \tau d\tau \\ &= -k_i \gamma_i^{(1)}. \end{aligned}$$

So, from (3.4) and (4.88) we see that

$$\begin{bmatrix} \dot{\gamma}_1^{(1)} \\ \dot{\gamma}_2^{(1)} \\ \dot{\gamma}_3^{(1)} \end{bmatrix} = \begin{bmatrix} \epsilon \sec \gamma_2^{(1)} (-k_1 \gamma_1^{(1)} \cos \gamma_3^{(1)} + k_2 \gamma_2^{(1)} \sin \gamma_3^{(1)}) \\ \epsilon (-k_1 \gamma_1^{(1)} \sin \gamma_3^{(1)} - k_2 \gamma_2^{(1)} \cos \gamma_3^{(1)}) \\ \epsilon ((\tan \gamma_2^{(1)}) (k_1 \gamma_1^{(1)} \cos \gamma_3^{(1)} - k_2 \gamma_2^{(1)} \sin \gamma_3^{(1)}) - k_3 \gamma_3^{(1)}) \end{bmatrix} \triangleq H(\gamma^{(1)})$$

and

$$\left. \frac{\partial H}{\partial \gamma^{(1)}} \right|_{\gamma^{(1)}=0} = \begin{bmatrix} -\epsilon k_1 & 0 & 0 \\ 0 & -\epsilon k_2 & 0 \\ 0 & 0 & -\epsilon k_3 \end{bmatrix}.$$

Thus, by Lyapunov's indirect method since $k_i > 0$, $\forall i$, $\gamma^* = 0$ is an exponentially stable equilibrium point for (4.88). From Theorem 4.23, we can conclude that (4.87) has a unique exponentially stable periodic solution in an $O(\epsilon)$ neighborhood of $\gamma^* = 0$ and so (4.86) has a unique exponentially stable periodic solution in an $O(\epsilon)$ neighborhood of $g^* = \Psi(0) = I$. However, since $\gamma^* = 0$ is itself an equilibrium point for (4.87), the unique exponentially stable periodic solution about $\gamma^* = 0$ must be the trivial solution $\gamma^* = 0$. Thus, $\gamma^* = 0$ is an exponentially stable equilibrium point for (4.87) and similarly $g^* = I$ is an exponentially stable equilibrium point for (4.86).

Chapter 5

Constructive Controllability and an Algorithm for Control

In this chapter we address the second main objective of this dissertation: to synthesize controls for systems on Lie groups of the form (2.9). In the previous chapter we addressed the first main objective: to describe the solution (motion) of system (2.9) on the Lie group G given periodically time-varying control inputs. We did so by deriving average approximations for the solution to (2.9) on G . Further, we developed a geometric interpretation for the average formulas. In this chapter we use the average formulas and exploit the geometric interpretation to derive a systematic way of synthesizing (periodically time-varying) controls to achieve desired motion on G .

The motion control problem that we consider is the complete constructive controllability problem **(P)** stated in Section 2.1.3 for drift-free systems of the form (2.9). This problem amounts to one in which, given $g_i, g_f \in G$, $t_f > 0$, we must specify control $u(t) = (u_1(t), \dots, u_m(t))$, $t \in [0, t_f]$, such that $g(0) = g_i$ and $g(t_f) = g_f$, where $g(t)$ is the solution to (2.9). This is a “point-to-point” control problem. To follow a certain path in G , e.g., to avoid obstacles, we could choose

target points along the trajectory and then do point-to-point control repeatedly from each target point to the next.

The strategy that we propose for solving **(P)** uses open-loop control and intermittent feedback control. This strategy allows us to take advantage of a priori knowledge of the system and prescribe efficient open-loop controls to drive the system as desired without sacrificing sensitivity reduction associated with feedback control. (For related ideas see [15]). The strategy can be summarized in four steps:

1. Choose intermediate target points $g_1, g_2, \dots, g_r$ between g_i and g_f so that the “distance” between successive target points is small.
2. Specify open-loop, small-amplitude, periodic controls that drive $g(t)$ of (2.9) from g_i to the first target point g_1 *approximately*. To do so, specify controls that drive an $O(\epsilon^q)$ average approximation of $g(t)$ from g_i to g_1 *exactly* (q to be determined).
3. If desired, apply feedback, i.e., make appropriate modifications based on measurement of the new system state. For example, modify selection of intermediate target points.
4. Repeat steps 2 and 3 for each successive target point (letting the previous target point be the new initial position) until done.

Our aim in this chapter is to solve Step 2, i.e., to provide the means to specify open-loop controls that solve **(P)** approximately. We use the formulas for $g^{(q)}(t)$, the $O(\epsilon^q)$ average approximation of $g(t)$ derived in Chapter 4, and the associated averaging theory (e.g., Theorem 4.11) that allows us to conclude that

$g(t)$ will remain close to $g^{(q)}(t)$ (to $O(\epsilon^q)$) over a relatively long ($O(1/\epsilon)$) time interval. Of critical importance is knowing how to choose q such that our $O(\epsilon^q)$ average solution can be controlled to any point in G . In other words, we need to determine how “good” an approximation is needed to capture the controllability of the system.

In Section 5.1 we study this problem of relating the average formulas for $g^{(q)}(t)$ of Chapter 4 to controllability. Given that system (2.9) is a depth- q' bracket system (as defined in Section 2.2.1), we show that $q = q' + 1$ is the minimum positive integer q such that an $O(\epsilon^q)$ average solution can be controlled as desired. This section is of particular significance because the proofs within explicitly describe a step-by-step procedure for constructing controls to drive an $O(\epsilon^q)$ average approximation as desired and thus provide a constructive methodology for solving $(\mathbf{P})$ with $O(\epsilon^q)$ accuracy.

In Section 5.2 we describe algorithms for depth- $(q - 1)$ bracket systems, $q = 1, 2, 3$, that systematically synthesize open-loop, sinusoidal controls to solve $(\mathbf{P})$ with $O(\epsilon^q)$ accuracy. These algorithms are derived according to the constructive procedure of Section 5.1. We illustrate these algorithms for examples described in Chapter 3. In Section 5.3 we propose a control architecture for motion control systems which uses the algorithms of Section 5.2 to provide adaptation to changes in control authority such as an actuator failure. In Section 5.4 we describe an experiment that was run to test the algorithms of Section 5.2 on an underwater vehicle in the neutral buoyancy tank of the Space Systems Laboratory at the University of Maryland.

5.1 Constructive Controllability

Consider system (2.9) with solution $g(t) \in G$. Let $q \geq 1$ be an integer. If we solve the complete constructive controllability problem **(P)** for an $O(\epsilon^q)$ average approximation to $g(t)$ then we will have solved **(P)** for $g(t)$ with $O(\epsilon^q)$ accuracy. The advantage to focusing on the control problem for an $O(\epsilon^q)$ average solution is that we have explicit, geometric formulas for average solutions $g^{(q)}(t)$ from Chapter 4. The formulas for $g^{(q)}(t)$ grow in complexity with increasing q so we would like to consider the minimum q possible. However, if we choose too low a q then our $O(\epsilon^q)$ average approximation might not be controllable. In this section we study the problem of solving **(P)** with $O(\epsilon^q)$ accuracy by deriving open-loop controls that drive an $O(\epsilon^q)$ average approximation as desired. We address the problem of choosing q and we show, by means of the proofs of two lemmas and a theorem, a methodology for synthesizing the open-loop controls.

Let I_p be the length- p ordered list of indices $\{i_1, \dots, i_p\}$ where $i_\nu \in \{1, \dots, m\}$, $\nu = 1, \dots, p$. Define the summation

$$\sum_{I_p=1}^m = \sum_{i_1=1}^m \sum_{i_2=1}^m \dots \sum_{i_p=1}^m.$$

Recall from (2.18) that the depth- $(p-1)$ structure constants $\Gamma_{I_p}^k$ are defined as

$$[\xi_{i_1}, [\xi_{i_2}, \dots, [\xi_{i_{p-1}}, \xi_{i_p}] \dots]] = \sum_{k=1}^n \Gamma_{I_p}^k \xi_k.$$

In the following we denote a set of r length- p ordered lists of indices as $I_p^1, \dots, I_p^r$.

Lemma 5.1 Suppose that system (2.9) on the connected Lie group G with Lie algebra $\mathcal{G}$ is a depth- q' bracket system. Let $q = q' + 1$. Then for any $W \in \mathcal{G}$, there exist n constants

$$\{c_1, \dots, c_m, c_{I_2^1}, \dots, c_{I_2^{p_2}}, c_{I_3^1}, \dots, c_{I_3^{p_3}}, \dots, c_{I_q^1}, \dots, c_{I_q^{p_q}}\},$$

where $(m + \sum_{l=2}^q p_l) = n$ such that

$$W = \sum_{k=1}^m c_k \xi_k + \sum_{k=1}^n \left(\sum_{j=1}^{p_2} c_{I_2^j} \Gamma_{I_2^j}^k + \sum_{j=1}^{p_3} c_{I_3^j} \Gamma_{I_3^j}^k + \dots + \sum_{j=1}^{p_q} c_{I_q^j} \Gamma_{I_q^j}^k \right) \xi_k$$

and $\{\sum_{k=1}^n \Gamma_{I_l^1}^k \xi_k, \dots, \sum_{k=1}^n \Gamma_{I_l^{p_l}}^k \xi_k\}$ are linearly independent for all $l = 2, \dots, q$. In other words, W can be written as the linear combination of $\{\xi_1, \dots, \xi_m\}$ and p_j depth- $(j-1)$ bracket terms, $j = 2, \dots, q$. Further, q is the minimum integer such that this is true.

Proof. In Section 2.2.1 we defined $\mathcal{C}$ to be the set of Lie brackets generated by $\{\xi_1, \dots, \xi_m\}$. Let $\mathcal{C}^{(q')} \subset \mathcal{C}$ be the subset containing Lie brackets up to depth- q' , $\mathcal{C}^{(q')} = \{\eta \mid \eta = [\eta_k, [\eta_{k-1}, [\dots, [\eta_1, \eta_0] \dots]]], \eta_i \in \{\xi_1, \dots, \xi_m\}, i = 0, \dots, k, k \leq q'\}$.

By definition of a depth- q' bracket system, $\text{span}(\mathcal{C}^{(q')}) = \mathcal{G}$. Thus, any element $W \in \mathcal{G}$ can be written as the linear combination of Lie brackets of $\{\xi_1, \dots, \xi_m\}$ of depth less than or equal to q' . q' is the minimum integer for which this is true since $\text{span}(\mathcal{C}^{(q'')}) \neq \mathcal{G}$ for $q'' < q'$. So we can write

$$\begin{aligned} W &= \sum_{k=1}^m c_k \xi_k + \sum_{I_2=1}^m c_{I_2} [\xi_{i_1}, \xi_{i_2}] + \sum_{I_3=1}^m c_{I_3} [[\xi_{i_1}, \xi_{i_2}], \xi_{i_3}] + \dots \\ &\quad + \sum_{I_q=1}^m c_{I_q} [\xi_{i_1}, [\xi_{i_2}, \dots, [\xi_{i_{q-1}}, \xi_{i_q}] \dots]] \\ &= \sum_{k=1}^m c_k \xi_k + \sum_{k=1}^n \left(\sum_{I_2=1}^m c_{I_2} \Gamma_{I_2}^k + \sum_{I_3=1}^m c_{I_3} \Gamma_{I_3}^k + \dots + \sum_{I_q=1}^m c_{I_q} \Gamma_{I_q}^k \right) \xi_k. \end{aligned} \quad (5.1)$$

where $c_k, k = 1, \dots, m, c_{I_2}, I_2 = 1, \dots, m, \dots, c_{I_q}, I_q = 1, \dots, m$ are constants. Since $\mathcal{G}$ is n -dimensional, there exists a choice of these constants such that (5.1) holds with only n of the constants nonzero. Let these n nonzero constants be denoted by

$$\{c_1, \dots, c_m, c_{I_2^1}, \dots, c_{I_2^{p_2}}, c_{I_3^1}, \dots, c_{I_3^{p_3}}, \dots, c_{I_q^1}, \dots, c_{I_q^{p_q}}\},$$

where $(m + \sum_{l=2}^q p_l) = n$. Then W is as in the statement of the lemma, i.e., the sum of m depth-0 bracket terms and p_j depth- $(j-1)$ bracket terms, $j = 2, \dots, q$.

Further we claim that the bracket terms at each depth are linearly independent, i.e., $\{\sum_{k=1}^n \Gamma_{I_l^1}^k \xi_k, \dots, \sum_{k=1}^n \Gamma_{I_l^{p_l}}^k \xi_k\}$ are linearly independent for all $l = 2, \dots, q$. If not then W has been described by fewer than n linearly independent terms. This is a contradiction since $\mathcal{G}$ is n -dimensional. $\square$

In the next lemma we show how to drive $g(t)$ with $O(\epsilon^q)$ accuracy to a point in G that can be expressed as a depth- $(l-1)$ bracket where $l \leq q$. This, together with the previous lemma, will be used in Theorem 5.4 as one step, to be repeated no more than n times, for driving $g(t)$ with $O(\epsilon^q)$ accuracy to *any* point in G . See Section 2.3.2 for definitions of $\hat{Q}, \hat{S}, \hat{\Psi}, \hat{d}$.

Lemma 5.2 Suppose that system (2.9) on the connected Lie group G with Lie algebra $\mathcal{G}$ is a depth- q' bracket system. Let $q = q' + 1$. Suppose that $g_0, g_1 \in \hat{Q} \subset G$ are such that $Z_0 = \hat{\Psi}^{-1}(g_0) = O(\epsilon^{q-1})$, $Y_1 = \hat{\Psi}^{-1}(g_1) = O(\epsilon^{q-1})$. Let $g(0) = g_0$ and let $l \in \{1, \dots, q\}$. Further, suppose that Y_1 can be expressed as a depth- $(l-1)$ bracket, i.e.,

$$Y_1 = \begin{cases} \sum_{k=1}^m c_k \xi_k & l = 1 \\ \sum_{k=1}^n c_{I_{l_0}} \Gamma_{I_{l_0}}^k \xi_k & l > 1 \end{cases}$$

where I_{l_0} is an ordered list of l indices, $\Gamma_{I_{l_0}}^k$ are depth- $(l-1)$ structure constants and $c_k, c_{I_{l_0}}$ are constants. Consider the local representation of $g(t)$ defined by $g(t) = e^{Z(t)}$, $Z(t) = \sum_{i=1}^{\infty} (-1)^{i+1} \epsilon^i Z_i(t)$, where $Z_i(t)$ are given by (2.43). Let $\epsilon_l \triangleq \epsilon^{q/l}$. For N an integer, define

$$\begin{aligned} Z^{(l)}(NT) &= (-1)^{l+1} (\epsilon_l)^l N Z_l(T) + Z_0^{(q)} \\ &= (-1)^{l+1} \epsilon^q N Z_l(T) + Z_0^{(q)}, \\ g^{(l)}(NT) &= e^{Z^{(l)}(t)}. \end{aligned}$$

Then we can choose an integer $M = O(1/\epsilon)$ and ϵ_l -amplitude, periodically time-varying controls $\epsilon_l u(t) = (\epsilon_l u_1(t), \dots, \epsilon_l u_m(t))$, satisfying $T_k(T) = 0$, $k = 0, 1, \dots, l-2$, ($T_k(t)$ defined by (2.43)) such that

$$Z^{(l)}(MT) = Y_1 + Z_0^{(q)}.$$

If $\|Z_0 - Z_0^{(q)}\| = O(\epsilon^q)$ then

$$\begin{aligned} \|Z(MT) - Z^{(l)}(MT)\| &= O(\epsilon^q), \\ \hat{d}(g(MT), g^{(l)}(MT)) &= O(\epsilon^q). \end{aligned}$$

Proof. Let $u(t)$ be piecewise continuous, $u(t+T) = u(t)$, $\forall t > 0$. Further for $l > 1$ let $T_k(T) = 0$, $k = 0, \dots, l-2$. Suppose that $\epsilon_l = \epsilon^{q/l}$ is the amplitude of $u(t)$ (i.e., replace ϵ in (2.9) by ϵ^l) and note that $(\epsilon_l)^l = \epsilon^q$. As in Theorem 4.11 define

$$\begin{aligned} Z^{(l)}(t) &= \sum_{i=1}^{l-1} (-1)^{i+1} (\epsilon_l)^i Z_i(t) + (-1)^{l+1} (\epsilon_l)^l \frac{t}{T} Z_l(T) + Z_0^{(q)}, \\ g^{(l)}(t) &= e^{Z^{(l)}(t)}. \end{aligned}$$

Let $b > 0$ be as defined in Theorem 4.11. Note that

$$Z_0 = O(\epsilon^{q-1}) = O((\epsilon_l^{l/q})^{q-1}) = O((\epsilon_l)^{l-l/q}) = O((\epsilon_l)^{l-1}).$$

If $\|Z_0 - Z_0^{(q)}\| = O((\epsilon_l)^l) = O(\epsilon^q)$ and $Z^{(l)}(t) \in \hat{S}$, $\forall t \in [0, b/\epsilon_l] \supset [0, b/\epsilon]$, then by Theorem 4.11,

$$\begin{aligned} \|Z(t) - Z^{(l)}(t)\| &= O((\epsilon_l)^l) = O(\epsilon^q), \\ \hat{d}(g(t), g^{(l)}(t)) &= O((\epsilon_l)^l) = O(\epsilon^q), \quad \forall t \in [0, b/\epsilon_l]. \end{aligned}$$

Further, for $t = NT$, N an integer,

$$\begin{aligned} Z^{(l)}(NT) &= (-1)^{l+1}(\epsilon_l)^l N Z_l(T) + Z_0^{(q)} \\ &= (-1)^{l+1}\epsilon^q N Z_l(T) + Z_0^{(q)}. \end{aligned} \quad (5.2)$$

The lemma is proved if we show that we can choose ϵ_l -amplitude periodic controls $\epsilon_l u(t)$ and an integer $M = O(1/\epsilon)$ such that $Z^{(l)}(MT) = Y_1 + Z_0^{(q)}$. This is the case because for small enough ϵ , $Z^{(l)}(t) \in \hat{S}$, $t \in [0, MT]$ since $Z^{(l)}(0) = O(\epsilon^{q-1})$ and $Y_1 \in \hat{S}$.

First, consider $l = 1$ and recall that

$$Z_1(T) = T_0(T) = \int_0^T U(\tau) d\tau = T U_{av} = T \sum_{k=1}^m u_{avk} \xi_k. \quad (5.3)$$

So for $l = 1$ we choose an integer $M = O(1/\epsilon)$ and ϵ_1 -amplitude, periodically time-varying controls $\epsilon_1 u(t) = (\epsilon_1 u_1(t), \dots, \epsilon_1 u_m(t))$, $t \in [0, MT]$, such that

$$\frac{1}{T} \int_0^T u_k(\tau) d\tau = u_{avk} = \frac{c_k}{T \epsilon^q M}, \quad k = 1, \dots, m. \quad (5.4)$$

Since $Y_1 = O(\epsilon^{q-1})$ implies $c_k = O(\epsilon^{q-1})$,

$$u_{avk} = \frac{O(\epsilon^{q-1})}{\epsilon^q O(1/\epsilon)} = O(1).$$

Also, from (5.2), (5.3) and (5.4),

$$\begin{aligned} Z^{(1)}(MT) &= \epsilon^q MT \sum_{k=1}^m u_{avk} \xi_k + Z_0^{(q)} \\ &= \epsilon^q MT \sum_{k=1}^m \frac{c_k}{T \epsilon^q M} \xi_k + Z_0^{(q)} \\ &= \sum_{k=1}^m c_k \xi_k + Z_0^{(q)} \\ &= Y_1 + Z_0^{(q)}. \end{aligned}$$

Next we consider $l > 1$. By Theorem 4.12

$$Z_l(T) = \frac{1}{l} \int_0^T [U(\tau_1), \int_0^{\tau_1} [U(\tau_2), \dots, \int_0^{\tau_{l-1}} U(\tau_l) d\tau_l] \dots] d\tau_1$$

$$\begin{aligned}
&= \frac{1}{l} \int_0^T \left[\sum_{i_1=1}^m u_{i_1}(\tau_1) \xi_{i_1}, \int_0^{\tau_1} \left[\sum_{i_2=1}^m u_{i_2}(\tau_2) \xi_{i_2}, \dots, \int_0^{\tau_{l-1}} \sum_{i_l=1}^m u_{i_l}(\tau_l) d\tau_l \xi_{i_l} \right] \dots \right] d\tau_1 \\
&= \frac{1}{l} \sum_{I_l=1}^m \int_0^T u_{i_1}(\tau_1) \int_0^{\tau_1} u_{i_2}(\tau_2) \dots \int_0^{\tau_{l-1}} u_{i_l}(\tau_l) d\tau_l \dots d\tau_1 [\xi_{i_1}, [\xi_{i_2}, \dots, \xi_{i_l}] \dots] \\
&= \frac{1}{l} \sum_{I_l=1}^m \Delta_{I_l}(T) \sum_{k=1}^n \Gamma_{I_l}^k \xi_k,
\end{aligned}$$

where

$$\Delta_{I_l}(T) \triangleq \int_0^T u_{i_1}(\tau_1) \int_0^{\tau_1} u_{i_2}(\tau_2) \dots \int_0^{\tau_{l-1}} u_{i_l}(\tau_l) d\tau_l \dots d\tau_1.$$

Let $\mathcal{I} = \{I_l \mid i_1, \dots, i_l \in \{1, \dots, m\}\}$. The elements of $\mathcal{I}$ are not all linearly independent because of skew symmetry of the Lie bracket and the Jacobi identity. However, we can find a set $\tilde{\mathcal{I}} \subset \mathcal{I}$ with $I_{l_0} \in \tilde{\mathcal{I}}$ such that the elements of $\tilde{\mathcal{I}}$ are linearly independent. For example, we can use a Philip Hall basis of depth- $(l-1)$ brackets [66]. Then, we can compute terms $\tilde{\Delta}_{I_l}(T)$, $I_l \in \tilde{\mathcal{I}}$ such that

$$Z_l(T) = \frac{1}{l} \sum_{I_l \in \tilde{\mathcal{I}}} \tilde{\Delta}_{I_l}(T) \sum_{k=1}^n \Gamma_{I_l}^k \xi_k \quad (5.5)$$

and $\{\sum_{k=1}^n \Gamma_{I_l}^k \xi_k\}_{I_l \in \tilde{\mathcal{I}}}$ are linearly independent. The terms $\tilde{\Delta}_{I_l}(T)$, $I_l \in \tilde{\mathcal{I}}$ are linear combinations of the terms $\Delta_{I_l}(T)$, $I_l \in \mathcal{I}$.

So choose an integer $M = O(1/\epsilon)$ and the controls $u(t)$ such that $\tilde{\Delta}_{I_l}(T)$ is defined by

$$\tilde{\Delta}_{I_l} = \begin{cases} \frac{l}{(-1)^{l+1} \epsilon^q M} c_{I_{l_0}} & I_l = I_{l_0} \\ 0 & I_l \in \tilde{\mathcal{I}}, I_l \neq I_{l_0} \end{cases} \quad (5.6)$$

Since $Y_1 = O(\epsilon^{q-1})$ implies $c_{I_{l_0}} = O(\epsilon^{q-1})$,

$$\tilde{\Delta}_{I_{l_0}}(T) = \frac{O(\epsilon^{q-1})}{\epsilon^q O(1/\epsilon)} = O(1).$$

By (5.5) and (5.6),

$$\begin{aligned}
Z_l(T) &= \frac{1}{l} \sum_{k=1}^n \frac{l}{(-1)^{l+1} \epsilon^q M} c_{I_{l_0}} \Gamma_{I_{l_0}}^k \xi_k \\
&= \frac{1}{(-1)^{l+1} \epsilon^q M} Y_1.
\end{aligned} \quad (5.7)$$

Thus, by (5.2) and (5.7),

$$\begin{aligned} Z^{(l)}(MT) &= (-1)^{l+1} \epsilon^q M \left(\frac{1}{(-1)^{l+1} \epsilon^q M} \right) Y_1 + Z_0^{(q)} \\ &= Y_1 + Z_0^{(q)}. \end{aligned}$$

□

Remark 5.3 Consider $l > 1$. Let $\omega = 2\pi/T$. An example of a control $u(t)$ that satisfies $u(t+T) = u(t)$, $\forall t > 0$, and $T_k(T) = 0$, $k = 0, \dots, q-2$ and that yields a nonzero value for the term $\Delta_{I_l}(T)$ is prescribed by

$$\begin{aligned} u_{i_1}(t) &= \begin{cases} (l-1)\omega \sin(l-1)\omega t, & l \text{ even} \\ (l-1)\omega \cos(l-1)\omega t, & l \text{ odd} \end{cases} \\ u_{i_2}(t) &= \dots = u_{i_l}(t) = \omega \cos \omega t, \quad u_k = 0, \quad k \notin I_l. \end{aligned}$$

In this case,

$$\Delta_{I_l}(T) = \begin{cases} (l-1)\omega^l \int_0^T \sin(l-1)\omega \tau_1 \int_0^{\tau_1} \cos \omega \tau_2 \cdots \int_0^{\tau_{l-1}} \cos \omega \tau_l d\tau_l \dots d\tau_1, & l \text{ even} \\ (l-1)\omega^l \int_0^T \cos(l-1)\omega \tau_1 \int_0^{\tau_1} \cos \omega \tau_2 \cdots \int_0^{\tau_{l-1}} \cos \omega \tau_l d\tau_l \dots d\tau_1, & l \text{ odd} \end{cases}$$

For instance, for $l = 4$, let $I_4 = \{i_1, i_2, i_3, i_4\}$ and

$$u_{i_1}(t) = 3\omega \sin(3\omega t), \quad u_{i_2}(t) = u_{i_3}(t) = u_{i_4}(t) = \omega \cos \omega t.$$

Then $u(t+T) = u(t)$, $\forall t > 0$ and

$$\int_0^T u_k(\tau) d\tau = 0, \quad k \in I_4 \implies T_0(T) = 0.$$

$$\int_0^T u_j(\tau_1) \int_0^{\tau_1} u_k(\tau_2) d\tau_2 d\tau_1 = 0, \quad j, k \in I_4 \implies T_1(T) = 0.$$

$$\int_0^T u_i(\tau_1) \int_0^{\tau_1} u_j(\tau_2) \int_0^{\tau_2} u_k(\tau_3) d\tau_3 d\tau_2 d\tau_1 = 0, \quad i, j, k \in I_4 \implies T_2(T) = 0.$$

Further,

$$\begin{aligned} \Delta_{I_l}(T) &= 3\omega^4 \int_0^T \sin 3\omega \tau_1 \int_0^{\tau_1} \cos \omega \tau_2 \int_0^{\tau_2} \cos \omega \tau_3 \int_0^{\tau_3} \cos \omega \tau_4 d\tau_4 d\tau_3 d\tau_2 d\tau_1 \\ &= -\pi/4. \end{aligned}$$

Other possibilities include

$$u_{i_1}(t) = u_{i_3}(t) = 3\omega \sin(3\omega t), \quad u_{i_2}(t) = u_{i_4}(t) = \omega \cos \omega t.$$

In this case $T_k(T) = 0$, $k = 0, 1, 2$ and $\Delta_{I_l}(T) = -9\pi/16$. For $l = 2$ we recall from (4.60) that

$$Z_2(T) = - \sum_{k=1}^n \left(\sum_{i,j=1; i < j}^m Area_{ij}(T) \Gamma_{ij}^k \right) \xi_k.$$

Thus, comparing with equation (5.5) in the proof of Lemma 5.2, we see that for $l = 2$, $\tilde{\mathcal{I}} = \{i, j = 1, \dots, m; i < j\}$ and $-\frac{1}{2}\tilde{\Delta}_{I_2}(T) = Area_{ij}(T)$. If we choose

$$u_i(t) = \alpha_i \omega \sin \omega t, \quad u_j(t) = \alpha_j \omega \cos \omega t, \quad u_k = 0, \quad k \notin \{i, j\},$$

then $u(t + T) = 0$, $\forall t > 0$, $T_0(T) = TU_{av} = 0$, $Area_{ij}(T) = \alpha_i \alpha_j \pi$ and $Area_{ik}(T) = Area_{jk}(T) = Area_{kk'}(T) = 0$, $k, k' \notin \{i, j\}$.

The next theorem gives the main result of this section.

Theorem 5.4 Suppose that system (2.9) on the connected Lie group G with Lie algebra $\mathcal{G}$ is a depth- q' bracket system. Let $q = q' + 1$. Then the complete constructive controllability problem **(P)** can be solved with $O(\epsilon^q)$ accuracy using the formulas $g^{(r)}(t)$ given by (4.37) for $r = 1, \dots, q$. Further, q is the minimum positive integer such that this is true.

Proof: Without loss of generality we can assume for problem **(P)** that $g(0) = g_i = e \in G$ and $g_f \in \hat{Q} \subset G$ is such that $\hat{d}(g_i, g_f) = O(\epsilon^{q-1})$. This is possible due to the left invariance of the system. The initial and final conditions desired for the local solution defined by the single exponential representation $g(t) = e^{Z(t)}$ are $Z(0) = \hat{\Psi}^{-1}(g_i) = \hat{\Psi}^{-1}(e) = 0 \in \mathcal{G}$ and $Z_f = \hat{\Psi}^{-1}(g_f) \in \mathcal{G}$. As a result,

$$\|Z_f\| = \|Z_f - Z(0)\| = \hat{d}(g_i, g_f) = O(\epsilon^{q-1}).$$

By Lemma 5.1 there exist n constants

$$\{c_1, \dots, c_m, c_{I_2^1}, \dots, c_{I_2^{p_2}}, c_{I_3^1}, \dots, c_{I_3^{p_3}}, \dots, c_{I_q^1}, \dots, c_{I_q^{p_q}}\},$$

where $(m + \sum_{l=2}^q p_l) = n$ and

$$Z_f = \sum_{k=1}^m c_k \xi_k + \sum_{k=1}^n \left(\sum_{j=1}^{p_2} c_{I_2^j} \Gamma_{I_2^j}^k + \sum_{j=1}^{p_3} c_{I_3^j} \Gamma_{I_3^j}^k + \dots + \sum_{j=1}^{p_q} c_{I_q^j} \Gamma_{I_q^j}^k \right) \xi_k.$$

Further, $\{\sum_{k=1}^n \Gamma_{I_l^1}^k \xi_k, \dots, \sum_{k=1}^n \Gamma_{I_l^{p_l}}^k \xi_k\}$ are linearly independent for all $l=2, \dots, q$.

Let

$$Y_1 = \sum_{k=1}^m c_k \xi_k, \quad Y_2 = \sum_{k=1}^n c_{I_2^1} \Gamma_{I_2^1}^k \xi_k, \quad Y_3 = \sum_{k=1}^n c_{I_2^2} \Gamma_{I_2^2}^k \xi_k, \quad \dots, \quad Y_{n-m+1} = \sum_{k=1}^n c_{I_q^{p_q}} \Gamma_{I_q^{p_q}}^k \xi_k.$$

Then

$$Z_f = \sum_{j=1}^{n-m+1} Y_j. \quad (5.8)$$

Let $M = O(1/\epsilon)$ be an integer. We show by induction on j that we can choose periodically time-varying controls $u(t) = (u_1(t), \dots, u_m(t))$ such that

$$\|Z(jMT) - \sum_{i=1}^j Y_i\| = O(\epsilon^q), \quad j = 1, \dots, n-m+1. \quad (5.9)$$

Consider $j = 1$. Let $Y_{0_1} = Z(0) = 0$ and $Y_{0_1}^{(q)} = 0$. Let $\epsilon_1 = \epsilon^q$. By Lemma 5.2 we can choose ϵ_1 -amplitude periodically time-varying controls

$\epsilon_1 u(t) = (\epsilon_1 u_1(t), \dots, \epsilon_1 u_m(t))$, $t \in [0, M]$, such that

$$\|Z(MT) - Y_1\| = \|Z(MT) - (Y_1 + Y_{0_1}^{(q)})\| = O(\epsilon^q).$$

Now suppose (5.9) holds for $j = j' - 1$. Then

$$\|Z((j' - 1)MT) - \sum_{i=1}^{j'-1} Y_i\| = O(\epsilon^q). \quad (5.10)$$

Let $Y_{0_{j'}} = Z((j' - 1)MT)$ and let $Y_{0_{j'}}^{(q)} = \sum_{i=1}^{j'-1} Y_i$. $Z_f = O(\epsilon^{q-1})$ implies by (5.8) that $Y_{0_{j'}}^{(q)} = O(\epsilon^{q-1})$. By (5.10)

$$\|Y_{0_{j'}} - Y_{0_{j'}}^{(q)}\| = O(\epsilon^q), \quad (5.11)$$

and so $Y_{0,j'} = O(\epsilon^{q-1})$. By definition $Y_{j'} = \sum_{k=1}^n c_{I_l^r} \Gamma_{I_l^r}^k \xi_k$, some $l \in \{2, \dots, q\}$, some $r \in \{1, \dots, p_l\}$. By Lemma 5.2 we can choose ϵ_l -amplitude, periodically time-varying controls $\epsilon_l u(t) = (\epsilon_l u_1(t), \dots, \epsilon_l u_m(t))$, $t \in [(j'-1)MT, j'MT]$, such that

$$\|Z(j'MT) - (Y_{j'} + Y_{0,j'}^{(q)})\| = O(\epsilon^q). \quad (5.12)$$

So by (5.12),

$$\begin{aligned} \|Z(j'MT) - \sum_{i=1}^{j'} Y_i\| &= \|Z(j'MT) - (Y_{j'} + \sum_{i=1}^{j'-1} Y_i)\| \\ &= \|Z(j'MT) - (Y_{j'} + Y_{0,j'}^{(q)})\| \\ &= O(\epsilon^q), \end{aligned}$$

i.e., (5.9) holds for $j = j'$. So by induction (5.9) is true for all $j = 1, \dots, n-m+1$.

We can choose the frequency and amplitudes of the controls such that t_f of problem **(P)** is $t_f = (n+m-1)MT$. Then from (5.9) for $j = n-m+1$,

$$\|Z(t_f) - Z_f\| = \|Z(t_f) - \sum_{i=1}^{n-m+1} Y_i\| = O(\epsilon^q).$$

Since $Z_f \in \hat{S}$, then $Z(t_f) \in \hat{S}$ for small enough ϵ . So by definition of $\hat{d}$,

$$\hat{d}(g(t_f), g_f) = \|Z(t_f) - Z_f\| = O(\epsilon^q).$$

This means that **(P)** is solved with $O(\epsilon^q)$ accuracy.

Now suppose $\bar{q} < q$. The $O(\epsilon^{\bar{q}})$ average approximation of $g(t)$ can only be driven to points $e^{Z_f} \in G$ where Z_f can be expressed as the linear combination of depth- $(q-1)$ brackets. However, by Lemma 5.1 q is the minimum integer for which Z_f can be expressed as the linear combination of depth- $(q-1)$ brackets. Thus, the $\bar{q}$ th-order average solution cannot be controlled to arbitrary points in G . $\square$

Remark 5.5 The proofs of Lemma 5.1, Lemma 5.2 and Theorem 5.4 provide a procedure for constructive controllability.

5.2 Algorithms for Control Synthesis

In this section we use the constructive methodology of Section 5.1 to derive algorithms that synthesize open-loop controls to solve the complete constructive controllability problem **(P)** with $O(\epsilon^q)$ accuracy for depth- $(q-1)$ bracket systems (2.9), $q = 1, 2, 3$. These algorithms are notably low in computational burden. The controls are small-amplitude *sinusoids* that are designed to be continuous and to satisfy $u(0) = u(t_f) = 0$. For a drift-free system, $u(0) = u(t_f) = 0$ means the system is initially and finally at rest. Additionally, we allow for flexibility in some of the choices of control parameters so that the period of the controls T can be made relatively large if desired (e.g., to avoid exciting vibrational modes in the system). T should satisfy $t_f/T = O(1/\epsilon)$. Thus, for example, T can be increased if t_f is increased, i.e., if the time constraint on completing the motion control maneuver can be relaxed.

We have chosen to use sinusoids for our control inputs, rather than some other type of periodically time-varying function, for a couple of reasons. First, following Remark 5.3, sinusoids are relatively easily chosen in accordance with the constructive procedure of Section 5.1, i.e., it is clear how to select the resonances between the different control inputs to generate motion as desired. Second, sinusoids are optimal for certain systems (2.9) on certain Lie groups in the sense that they solve the following optimal control problem:

(P') Let $J(u) = \int_0^{t_f} L(u(t))dt$ define a cost function. Find $u(\cdot)$ that minimizes $J(u)$ subject to the condition that $g(0) = g_i$ and $g(t_f) = g_f$.

For example consider the spacecraft attitude control problem of Section 3.1 with only two controls, i.e., $G = SO(3)$ and $m = 2$. Let $L(u) = 1/2(I_1 u_1^2 + I_2 u_2^2)$

where $I_i > 0$, $i = 1, 2$. Then $J(u)$ is the energy expended and problem $(\mathbf{P}')$ is the motion control problem with minimal energy. Baillieul showed that the optimal solutions $u_1^*(t)$ and $u_2^*(t)$ to this problem are sinusoids for $I_1 = I_2 = 1$. [4]. Brockett showed that for the three-dimensional nilpotent system described in Section 3.5, i.e., $G = H(3)$ and $m = 2$, with $L(u) = 1/2(I_1 u_1^2 + I_2 u_2^2)$, problem $(\mathbf{P}')$ is also solved with sinusoids [13]. More recent work on optimal control of systems on Lie groups by Jurdjevic [28] and Krishnaprasad [34] has revealed other types of periodically time-varying controls as optimal solutions. For example, in [34] Krishnaprasad showed that elliptic functions solve the optimal control problem $(\mathbf{P}')$ with $L(u) = u_1^2 + u_2^2$ for the unicycle described in Section 3.2, i.e., $G = SE(2)$ and $m = 2$.

We assume for our algorithms that in problem $(\mathbf{P})$, $g_i = e \in G$ and $g_f \in \hat{Q} \subset G$ is such that $\hat{d}(g_i, g_f) = O(\epsilon^{q-1})$. This is possible due to the left invariance of the system. The assumption implies that the algorithm prescribes controls for small maneuvers. Given a motion control problem which involves a large maneuver, we first break it down into small steps by selecting target points between g_i and g_f and then apply the appropriate algorithm to each step. The selection of these target points is not addressed in this dissertation. However, we note that, for a suitable measure of energy, more energy is expended to move along a depth- j bracket direction for larger values of j . Thus, a more energy efficient choice of target points minimizes motion along higher depth bracket directions.

Recall the single exponential local representation $g(t) = e^{Z(t)}$. Locally, we want to drive $Z(t)$ to $Z_f = \hat{\Psi}^{-1}(g_f) \in \mathcal{G}$. We assume that Z_f is given and $z_f = (z_{f1}, \dots, z_{fn})^T$ computed such that $Z_f = \sum_{i=1}^n z_{fi} \xi_i$. If Z_f is given with respect to some other basis $\{\eta_1, \dots, \eta_n\}$ for $\mathcal{G}$, then a linear transformation will

be needed to compute $(z_{f_1}, \dots, z_{f_n})$.

Recall the product of exponentials local representation $g(t) = e^{\gamma_1(t)\xi_1} \dots e^{\gamma_n(t)\xi_n}$. Suppose that $g_f \in Q \subset G$ and that $\gamma_f = \Psi^{-1}(g_f) \in \mathfrak{R}^n$ is given. This is the case, for example, if the desired orientation of the spacecraft (Section 3.1) is given in terms of Euler angles $(\gamma_{f_1}, \gamma_{f_2}, \gamma_{f_3})$. Since we assume $g(0) = g_i = e$ then $Z(0) = \hat{\Psi}^{-1}(g(0)) = 0 \in \mathcal{G}$. By definition of $\hat{d}$ (2.46),

$$\left\| \sum_{i=1}^n z_{f_i} \xi_i \right\| = \|Z_f\| = \|Z_f - Z(0)\| = \hat{d}(g_i, g_f) = O(\epsilon^{q-1}).$$

Thus, $z_{f_i} = O(\epsilon^{q-1})$ and so by Lemma 2.10 for G a matrix Lie group and $q > 1$ it follows that $\gamma_{f_i} = O(\epsilon^{q-1})$ and

$$|z_{f_i} - \gamma_{f_i}| = O(\epsilon^{2(q-1)}) = O(\epsilon^q), \quad i = 1, \dots, n.$$

So suppose we drive $Z(t)$ such that $\|Z(t_f) - \sum_{i=1}^n \gamma_{f_i} \xi_i\| = O(\epsilon^q)$, then

$$\begin{aligned} \|Z(t_f) - \sum_{i=1}^n z_{f_i} \xi_i\| &= \|Z(t_f) - \sum_{i=1}^n z_{f_i} \xi_i + \sum_{i=1}^n \gamma_{f_i} \xi_i - \sum_{i=1}^n \gamma_{f_i} \xi_i\| \\ &\leq \|Z(t_f) - \sum_{i=1}^n \gamma_{f_i} \xi_i\| + \left\| \sum_{i=1}^n \gamma_{f_i} \xi_i - \sum_{i=1}^n z_{f_i} \xi_i \right\| \\ &= O(\epsilon^q). \end{aligned}$$

So we can use the γ_f and z_f interchangeably for G a matrix Lie group, i.e., given γ_f we set $z_f = \gamma_f$. Further let $z(t) = (z_1(t), \dots, z_n(t))^T$ be defined by $Z(t) = \sum_{i=1}^n z_i(t) \xi_i$. Then, for $z(0) = 0$ and $z(t_f) = z_f = O(\epsilon^{q-1})$, we can assume that $z(t) = O(\epsilon^{q-1})$, $t \in [0, t_f]$. So again by Lemma 2.10 for G a matrix Lie group and $q > 1$, $\gamma(t) = O(\epsilon^{q-1})$, $t \in [0, t_f]$, and

$$\|\gamma(t) - z(t)\| = O(\epsilon^{2(q-1)}) = O(\epsilon^q), \quad t \in [0, t_f]. \quad (5.13)$$

We will use this result in the examples at the end of this section.

As in the proof of Theorem 5.4, the algorithms in this section are designed to solve the motion control problem using multiple substeps. The time interval $[0, t_f]$ is divided into subintervals, e.g., $[0, t_f] = [t_0, t_1] \cup [t_1, t_2] \cup \dots \cup [t_{\mu-1}, t_\mu]$, $t_0 = 0$, $t_\mu = t_f$, and controls specified on each subinterval. The controls are specified so that the terms $Area_{ij}(T)$ and $m_{ijk}(T)$ will take on a single constant value on each subinterval. However, these terms may take on different values on different subintervals. Thus, for ease of notation we define the “running total” of the time-varying area terms and moment terms as $Area_{ij}(t)$ and $m_{ijk}(t)$, respectively. Let $Area_{ij}^{(r)}(T)$ and $m_{ijk}^{(r)}(T)$ be the values of the area and moment terms, respectively, during the time interval $[t_{r-1}, t_r]$ and suppose that $t \in [t_\nu, t_{\nu+1}]$, $0 \leq \nu < \mu$. Then define

$$Area_{ij}(t) = \sum_{r=1}^{\nu} \frac{(t_r - t_{r-1})}{T} Area_{ij}^{(r)}(T) + \frac{(t - t_\nu)}{T} Area_{ij}^{(\nu+1)}(T), \quad (5.14)$$

$$m_{ijk}(t) = \sum_{r=1}^{\nu} \frac{(t_r - t_{r-1})}{T} m_{ijk}^{(r)}(T) + \frac{(t - t_\nu)}{T} m_{ijk}^{(\nu+1)}(T). \quad (5.15)$$

5.2.1 Algorithm 0: Depth-Zero Bracket System

The depth-zero bracket system is a trivial case since $m = n$. We are given $g_i = e \in G$, $\sum_{i=1}^n z_{f_i} \xi_i = Z_f = \hat{\Psi}^{-1}(g_f) = O(1)$ and $t_f > 0$. Choose M to be a positive integer such that $M \geq 1/\pi\epsilon$. Let the period T and frequency ω of the controls be

$$T = \frac{t_f}{M}, \quad \omega = \frac{2\pi}{T}.$$

Define controls $\epsilon u_k(\cdot)$ for $k = 1, \dots, n$ by

$$\epsilon u_k(t) = \frac{z_{f_k} \omega}{M\pi} \sin^2 \omega t, \quad 0 \leq t \leq t_f.$$

Then $\epsilon u_{av} t_f = M \int_0^{t_f} \epsilon u_k(\tau) d\tau = z_{f_k}$. This means that $g^{(1)}(t_f) = e^{\epsilon U_{av} t_f} = e^{Z_f} = g_f$, i.e., the first-order average approximation is driven exactly to g_f at $t = t_f$ as

desired.

5.2.2 Algorithm 1: Depth-One Bracket System

For the depth-one bracket (single bracket) system we have $m < n$. We are given $g_i = e \in G$, $\sum_{i=1}^n z_{f_i} \xi_i = Z_f = \hat{\Psi}^{-1}(g_f) = O(\epsilon)$ and $t_f > 0$. Algorithm 1 has two parts, Part A and Part B defined below.

PART A: By Lemma 5.1 there exist constants c_k, c_{ij} such that

$$Z_f = \sum_{k=1}^n z_{f_k} \xi_k = \sum_{k=1}^m c_k \xi_k + \sum_{k=1}^n \sum_{i,j=1; i < j}^m c_{ij} \Gamma_{ij}^k \xi_k. \quad (5.16)$$

Further, these constants can be chosen so that only n of them are nonzero. We compute c_k, c_{ij} such that (5.16) holds, i.e., such that

$$z_{f_k} = \begin{cases} c_k + \sum_{i,j=1; i < j}^m c_{ij} \Gamma_{ij}^k & k = 1, \dots, m \\ \sum_{i,j=1; i < j}^m c_{ij} \Gamma_{ij}^k & k = m+1, \dots, n. \end{cases} \quad (5.17)$$

Consider the matrix

$$\begin{aligned} \hat{\Gamma} &= \begin{bmatrix} \Gamma_{12} & \Gamma_{13} & \dots & \Gamma_{1m} & \Gamma_{23} & \dots & \Gamma_{2m} & \dots & \Gamma_{(m-1)m} \end{bmatrix} \\ &\triangleq \begin{bmatrix} \Gamma_{12}^{m+1} & \Gamma_{13}^{m+1} & \dots & \Gamma_{1m}^{m+1} & \Gamma_{23}^{m+1} & \dots & \Gamma_{2m}^{m+1} & \dots & \Gamma_{(m-1)m}^{m+1} \\ \Gamma_{12}^{m+2} & . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . & . \\ \Gamma_{12}^n & \Gamma_{13}^n & \dots & \Gamma_{1m}^n & \Gamma_{23}^n & \dots & \Gamma_{2m}^n & \dots & \Gamma_{(m-1)m}^n \end{bmatrix}. \end{aligned}$$

Note that $\hat{\Gamma}$ has rank $n - m$. We give two options for choosing the constants c_k, c_{ij} .

Option (1): From the left write down the first $n-m$ linearly independent columns of $\hat{\Gamma}$ as

$$\Gamma \triangleq [\Gamma_{i_{m+1}j_{m+1}} \cdots \Gamma_{i_n j_n}] \in \mathfrak{R}^{(n-m) \times (n-m)}$$

$$i_\nu, j_\nu \in \{1, \dots, m\}, i_\nu < j_\nu, \nu = m+1, \dots, n.$$

Then Γ is nonsingular. So let

$$\begin{bmatrix} c_{i_{m+1}j_{m+1}} \\ \vdots \\ c_{i_n j_n} \end{bmatrix} = \Gamma^{-1} \begin{bmatrix} z_{f_{m+1}} \\ \vdots \\ z_{f_n} \end{bmatrix},$$

$$c_{ij} = 0 \text{ if } \Gamma_{ij} \text{ not a column of } \Gamma,$$

$$c_k = z_{f_k} - \sum_{i,j=1; i < j}^m c_{ij} \Gamma_{ij}^k, \quad k = 1, \dots, m.$$

Option (2): Define the generalized inverse of $\hat{\Gamma}$ to be $\hat{\Gamma}^\dagger = \hat{\Gamma}^T (\hat{\Gamma} \hat{\Gamma}^T)^{-1}$. Then let

$$\begin{bmatrix} c_{12} \\ c_{13} \\ \vdots \\ c_{(m-1)m} \end{bmatrix} = \hat{\Gamma}^\dagger \begin{bmatrix} z_{f_{m+1}} \\ \vdots \\ z_{f_n} \end{bmatrix},$$

$$c_k = z_{f_k} - \sum_{i,j=1; i < j}^m c_{ij} \Gamma_{ij}^k, \quad k = 1, \dots, m.$$

For either option, let

$$S = \{k \mid c_{kj} \neq 0, \text{ some } j > k\}, \quad r = |S| \triangleq \text{number of elements in } S.$$

PART B: Following the proof of Theorem 5.4 we apply controls to drive an $O(\epsilon^2)$ average approximation of $g(t)$ to g_f at $t = t_f$. Component 1(i), described below,

prescribes controls to achieve desired motion in the depth-zero directions. This is done simply by noting that for sinusoidal control inputs *in phase*, area terms $Area_{ij}(T)$ will all be zero. When all area terms are zero and $Z_0^{(2)} = 0$ by (4.36),

$$Z^{(2)}(t) = \epsilon \tilde{U}(t) = \sum_{k=1}^m \epsilon \tilde{u}_k(t) \xi_k.$$

Comparing this with (5.16), we see that to achieve the desired motion in the depth-zero bracket directions, we want $\epsilon \tilde{u}_k(t_f) = c_k$, $k = 1, \dots, m$. According to Lemma 5.2 and Remark 5.3, for $Z_0^{(2)} = 0$ and M an integer,

$$Z^{(2)}(MT) = -\epsilon^2 M Z_2(T) = \epsilon^2 M \sum_{k=1}^n \sum_{i,j=1; i < j}^m Area_{ij}(T) \Gamma_{ij}^k \xi_k.$$

Comparing this to (5.16), it is clear that to achieve the desired motion in the depth-one bracket directions, we want $\epsilon^2 M Area_{ij}(T)$ to be equal to c_{ij} . Component 1(ii) prescribes controls to achieve this depth-one bracket motion.

Choose M to be a positive integer such that $M \geq 1/\pi\epsilon$. Let the period T and frequency ω of the controls be

$$T = \frac{t_f}{r(M+1) + 1/2}, \quad \omega = \frac{2\pi}{T}.$$

Then using the controls defined in Components 1(i) and 1(ii), perform the following iterations:

1. $i = 0$.
2. $i = i + 1$.
3. If $i \notin S$ go to 5.
4. Apply Component 1(ii) for c_{ij} , $j = i + 1, \dots, m$.
5. If $i < m - 1$ go to 2.

6. Apply Component 1(i) for $c_k, k = 1, \dots, m$.

Then we are done and $\hat{d}(g(t_f), g_f) = O(\epsilon^2)$ as desired.

In the following control laws, if a control component is not explicitly prescribed it should be set equal to zero.

Component 1(i)

Given: $c_k, k = 1, \dots, m, T, \omega$ and current time t_0 .

Goal: Let $t_1 = t_0 + T/2$. We define $\epsilon u_k(t), k = 1, \dots, m, t \in [t_0, t_1]$, continuous, such that $\epsilon \tilde{u}_k(t_1) = c_k$ and $\epsilon u_k(t_1) = \epsilon u_k(t_0) = 0$.

Controls:

$$\epsilon u_k(t) = \frac{1}{2} c_k \omega \sin(\omega(t - t_0)), \quad t_0 \leq t \leq t_1.$$

Component 1(ii)

Given: $i < j, c_{ij}$ for $j = i + 1, \dots, m, T, \omega, M$ and current time t_0 .

Goal: Let $t_1 = t_0 + (M + 1)T$. Let

$$\begin{aligned} \alpha_i &= \left(\sum_{j=i+1}^m \frac{c_{ij}^2}{\pi^2 M^2} \right)^{1/4}, \\ \alpha_j &= \frac{c_{ij}}{\alpha_i \pi M}, \quad j = i + 1, \dots, m. \end{aligned}$$

Specify continuous, zero-mean controls $u_i(t), u_j(t), j = i + 1, \dots, m, t \in [t_0, t_1]$ such that $\epsilon^2 \text{Area}_{ij}(t_1) = c_{ij}, \epsilon u_i(t_0) = \epsilon u_i(t_1) = \epsilon u_j(t_0) = \epsilon u_j(t_1) = 0$.

Controls:

$$\left. \begin{array}{l} \epsilon u_i(t) = \alpha_i \omega \sin(\omega(t - t_0)) \\ \epsilon u_j(t) = 0 \end{array} \right\} t_0 \leq t \leq t_0 + \frac{T}{4} = s_1$$

$$\left. \begin{array}{l} \epsilon u_i(t) = \alpha_i \omega \cos(\omega(t - s_1)) \\ \epsilon u_j(t) = \alpha_j \omega \sin(\omega(t - s_1)) \end{array} \right\} s_1 \leq t \leq s_1 + MT = s_2$$

$$\left. \begin{array}{l} \epsilon u_i(t) = \alpha_i \omega \cos(\omega(t - s_2)) \\ \epsilon u_j(t) = 0 \end{array} \right\} s_2 \leq t \leq s_2 + \frac{3T}{4} = t_1$$

Note that $\epsilon^2 \text{Area}_{ij}(t_1) = \alpha_i \alpha_j \pi M = c_{ij}$ and the goal is met.

Remark 5.6 Option (1) can be considered as the time optimal option since it selects the minimum number of nonzero constants c_{ij} , thus, minimizing the number of times Component 1(ii) is called. Option (2) can be thought of as the minimum energy option since it gives a least squares solution to the selection of c_{ij} . In Component 1(ii), given the chosen structure of the controls, the amplitudes α_i, α_j are selected to minimize energy $J(u) = \int_{s_1}^{s_2} ((\epsilon u_i(t))^2 + \sum_{j=i+1}^m (\epsilon u_j(t))^2) dt$ and to satisfy $\alpha_i \alpha_j \pi M = c_{ij}$. This amounts to choosing α_i to minimize $\{\alpha_i^2 + \sum_{j=i+1}^m (c_{ij}/\alpha_i \pi M)^2\}$. If the control inputs are not all equally costly to use, the constants α_i, α_j could be computed instead to minimize a weighted energy function. Finally, we note that Component 1(ii) is energy (and time) saving in that it meets several depth-one bracket conditions simultaneously.

5.2.3 Algorithm 2: Depth-Two Bracket System

For the depth-two bracket (double bracket) system we have $m < n$. We define $0 \leq l \leq (n - m)$ such that the relative growth vector σ defined by equation (2.14) is $\sigma = (m, l, n - (m + l))$. In particular, let

$$\Xi = \{p \in \{m + 1, \dots, n\} \mid \Gamma_{ij}^p \neq 0 \text{ some } i, j \in \{1, \dots, m\}, i < j\}. \quad (5.18)$$

Then $l = |\Xi| = \text{number of elements in } \Xi$. We will assume for the purposes of the algorithm, without loss of generality, that the basis $\{\xi_1, \dots, \xi_n\}$ is ordered such that $\{m+1, \dots, m+l\} \in \Xi$. We are given $g_i = e \in G$, $\sum_{i=1}^n z_{f_i} \xi_i = Z_f = \hat{\Psi}^{-1}(g_f) = O(\epsilon^2)$ and $t_f > 0$. Algorithm 2 has two parts, Part A and Part B defined below.

PART A: By Lemma 5.1 there exist constants c_p, c_{ij}, c_{ijk} such that

$$Z_f = \sum_{p=1}^n z_{f_p} \xi_p = \sum_{p=1}^m c_p \xi_p + \sum_{p=1}^n \sum_{i,j=1; i < j}^m (c_{ij} \Gamma_{ij}^p + c_{iji} \theta_{iji}^p + \sum_{k=i+1}^m c_{ijk} \theta_{ijk}^p) \xi_p. \quad (5.19)$$

Further, these constants can be chosen so that only n of them are nonzero. We compute c_p, c_{ij}, c_{ijk} such that (5.19) holds, i.e., such that

$$z_{f_p} = \begin{cases} c_p + \sum_{i,j=1; i < j}^m (c_{ij} \Gamma_{ij}^p + c_{iji} \theta_{iji}^p + \sum_{k=i+1}^m c_{ijk} \theta_{ijk}^p) & p = 1, \dots, m \\ \sum_{i,j=1; i < j}^m (c_{ij} \Gamma_{ij}^p + c_{iji} \theta_{iji}^p + \sum_{k=i+1}^m c_{ijk} \theta_{ijk}^p) & p = m+1, \dots, m+l \\ \sum_{i,j=1; i < j}^m (c_{iji} \theta_{iji}^p + \sum_{k=i+1}^m c_{ijk} \theta_{ijk}^p) & p = m+l+1, \dots, n \end{cases} \quad (5.20)$$

Consider the matrix

$$\hat{\theta} = \begin{bmatrix} \theta_{121} & \theta_{122} & \dots & \theta_{12m} & \dots & \theta_{1mm} & \theta_{232} & \dots & \theta_{(m-1)mm} \end{bmatrix}$$

$$\triangleq \begin{bmatrix} \theta_{121}^{m+l+1} & \theta_{122}^{m+l+1} & \dots & \theta_{12m}^{m+l+1} & \dots & \theta_{1mm}^{m+l+1} & \theta_{232}^{m+l+1} & \dots & \theta_{(m-1)mm}^{m+l+1} \\ \theta_{121}^{m+l+2} & \theta_{122}^{m+l+2} & & . & & . & . & & . \\ . & . & & . & & . & . & & . \\ . & . & & . & & . & . & & . \\ \theta_{121}^n & \theta_{122}^n & \dots & . & \dots & . & \dots & & \theta_{(m-1)mm}^n \end{bmatrix}$$

and the matrix

$$\hat{\Gamma} = \begin{bmatrix} \Gamma_{12} & \Gamma_{13} & \dots & \Gamma_{1m} & \Gamma_{23} & \dots & \Gamma_{2m} & \dots & \Gamma_{(m-1)m} \end{bmatrix}$$

$$\hat{\Gamma} \triangleq \begin{bmatrix} \Gamma_{12}^{m+1} & \Gamma_{13}^{m+1} & \dots & \Gamma_{1m}^{m+1} & \Gamma_{23}^{m+1} & \dots & \Gamma_{2m}^{m+1} & \dots & \Gamma_{(m-1)m}^{m+1} \\ \Gamma_{12}^{m+2} & \cdot & & \cdot & \cdot & & \cdot & & \cdot \\ \cdot & \cdot & & \cdot & \cdot & & \cdot & & \cdot \\ \cdot & \cdot & & \cdot & \cdot & & \cdot & & \cdot \\ \Gamma_{12}^{m+l} & \Gamma_{13}^{m+l} & \dots & \Gamma_{1m}^{m+l} & \Gamma_{23}^{m+l} & \dots & \Gamma_{2m}^{m+l} & \dots & \Gamma_{(m-1)m}^{m+l} \end{bmatrix}.$$

Let $\alpha \triangleq m+l$. Then $\hat{\theta}$ has rank $n-\alpha$ and $\hat{\Gamma}$ has rank l . We give two options for choosing the constants c_p, c_{ij}, c_{ijk} .

Option (1): From the left write down the first $n-\alpha$ linearly independent columns of $\hat{\theta}$ as

$$\theta \triangleq [\theta_{i_{\alpha+1}j_{\alpha+1}k_{\alpha+1}} \dots \theta_{i_nj_nk_n}] \in \mathfrak{R}^{(n-\alpha) \times (n-\alpha)}$$

$$i_\nu, j_\nu, k_\nu \in \{1, \dots, m\}, i_\nu < j_\nu, i_\nu \leq k_\nu, \nu = \alpha+1, \dots, n.$$

Then θ is nonsingular. So let

$$\begin{bmatrix} c_{i_{\alpha+1}j_{\alpha+1}k_{\alpha+1}} \\ \vdots \\ c_{i_nj_nk_n} \end{bmatrix} = \theta^{-1} \begin{bmatrix} z_{f_{\alpha+1}} \\ \vdots \\ z_{f_n} \end{bmatrix},$$

$$c_{ijk} = 0 \text{ if } \theta_{ijk} \text{ not a column of } \theta.$$

From the left write down the first l linearly independent columns of $\hat{\Gamma}$ as

$$\Gamma \triangleq [\Gamma_{i_1j_1} \dots \Gamma_{i_lj_l}] \in \mathfrak{R}^{l \times l}$$

$$i_\nu, j_\nu \in \{1, \dots, m\}, i_\nu < j_\nu, \nu = 1, \dots, l.$$

Then Γ is nonsingular. So let

$$\begin{bmatrix} c_{i_1j_1} \\ \vdots \\ c_{i_lj_l} \end{bmatrix} = \Gamma^{-1} \begin{bmatrix} z_{f_{m+1}} - \sum_{i,j=1;i < j}^m (c_{iji}\theta_{iji}^{m+1} + \sum_{k=i+1}^m c_{ijk}\theta_{ijk}^{m+1}) \\ \vdots \\ z_{f_{m+l}} - \sum_{i,j=1;i < j}^m (c_{iji}\theta_{iji}^{m+l} + \sum_{k=i+1}^m c_{ijk}\theta_{ijk}^{m+l}) \end{bmatrix},$$

$$c_{ij} = 0 \text{ if } \Gamma_{ij} \text{ not a column of } \Gamma,$$

$$c_p = z_{f_p} - \sum_{i,j=1;i < j}^m (c_{ij}\Gamma_{ij}^p + c_{iji}\theta_{iji}^p + \sum_{k=i+1}^m c_{ijk}\theta_{ijk}^p) \quad p = 1, \dots, m.$$

Option (2): Define the generalized inverse of $\hat{\theta}$ to be $\hat{\theta}^\dagger = \hat{\theta}^T(\hat{\theta}\hat{\theta}^T)^{-1}$. Let

$$\begin{bmatrix} c_{121} \\ c_{122} \\ \vdots \\ c_{(m-1)mm} \end{bmatrix} = \hat{\theta}^\dagger \begin{bmatrix} z_{f_{m+l+1}} \\ \vdots \\ z_{f_n} \end{bmatrix}.$$

Define the generalized inverse of $\hat{\Gamma}$ to be $\hat{\Gamma}^\dagger = \hat{\Gamma}^T(\hat{\Gamma}\hat{\Gamma}^T)^{-1}$. Then let

$$\begin{bmatrix} c_{12} \\ c_{13} \\ \vdots \\ c_{(m-1)m} \end{bmatrix} = \hat{\Gamma}^\dagger \begin{bmatrix} z_{f_{m+1}} - \sum_{i,j=1;i < j}^m (c_{iji}\theta_{iji}^{m+1} + \sum_{k=i+1}^m c_{ijk}\theta_{ijk}^{m+1}) \\ \vdots \\ z_{f_{m+l}} - \sum_{i,j=1;i < j}^m (c_{iji}\theta_{iji}^{m+l} + \sum_{k=i+1}^m c_{ijk}\theta_{ijk}^{m+l}) \end{bmatrix},$$

$$c_p = z_{f_p} - \sum_{i,j=1;i < j}^m (c_{ij}\Gamma_{ij}^p + c_{iji}\theta_{iji}^p + \sum_{k=i+1}^m c_{ijk}\theta_{ijk}^p) \quad p = 1, \dots, m.$$

For either option, let

$$Y = \{c_{ijk} \mid c_{ijk} \neq 0, i < j, i \leq k\},$$

$$Q = \{c_{ijk} \in Y \mid i < j < k\}, \quad \beta = |Q| = \text{number of elements in } Q,$$

$$R = \{c_{ijk} \in Y \mid k = i\} \cup \{c_{ijk} \in Y \mid k = j \text{ and } c_{iji} \notin Y\},$$

$$V = \{c_{ij} \mid c_{ij} \neq 0, i < j\},$$

$$W = \{c_{ijk} \in R \mid c_{ij} \notin V\}, \quad \delta = |W| + |V|.$$

PART B: Following the proof of Theorem 5.4 we apply controls to drive an $O(\epsilon^3)$ average approximation of $g(t)$ to g_f at $t = t_f$. Component 2(i), described

below, prescribes controls to achieve motion in the depth-zero bracket directions as in Algorithm 1. The first portion of Component 2(ii), also described below, achieves motion in the depth-one bracket directions as in Algorithm 1. Here, however, the amplitude of the controls are of order $\epsilon_2 = \epsilon^{3/2}$ as specified in the proof of Theorem 5.4. According to Lemma 5.2 and Corollary 4.18, for $Z_0^{(3)} = 0$ and M an integer,

$$\begin{aligned} Z^{(3)}(MT) &= \epsilon^3 M Z_3(T) \\ &= -\epsilon^3 M \sum_{p=1}^n \sum_{i,j=1;i < j}^m (m_{iji}(T)\theta_{iji}^p + \sum_{k=i+1}^m (2m_{ijk}(T) - m_{ikj}(T))\theta_{ijk}^p)\xi_p. \end{aligned}$$

Comparing this to (5.19) it is clear that to achieve the desired motion in the depth-two bracket directions we want $\epsilon^3 M m_{iji}(T)$ to be equal to $-c_{iji}$ and $\epsilon^3 (2m_{ijk}(T) - m_{ikj}(T))$ to be equal to $-c_{ijk}$. The remainder of Component 2(ii) prescribes controls to satisfy the former while Component 2(iii) prescribes controls to satisfy the latter.

Choose M to be a positive integer such that $M \geq 1/\pi\epsilon$. Let the period T and frequency ω of the controls be

$$T = \frac{t_f}{(6\beta + 3\delta)(M + 1) + 1/2}, \quad \omega = \frac{2\pi}{T}.$$

Then use the controls defined in Components 2(i), 2(ii) and 2(iii) to perform the following iterations:

1. $i = 0$.
2. $i = i + 1, j = i$.
3. $j = j + 1, k = j$.
4. $k = k + 1$. If $k = m + 1$ go to 8.

5. If $c_{ijk} = c_{ikj} = 0$ go to 7.
6. Apply Component 2(iii) for c_{ijk} and c_{ikj} .
7. If $k \leq m - 1$ go to 4.
8. If $c_{ij} = c_{iji} = c_{ijj} = 0$ go to 10.
9. Apply Component 2(ii) for c_{ij} , c_{iji} and c_{ijj} .
10. If $j \leq m - 1$ go to 3.
11. If $i < m - 1$ go to 2.
12. Apply Component 2(i) for c_p for $p = 1, \dots, m$.

Then we are done and $\hat{d}(g(t_f), g_f) = O(\epsilon^3)$ as desired.

In the following control laws, if a control component is not explicitly prescribed it should be set equal to zero.

Component 2(i)

This component is the same as Component 1(i) of Algorithm 1.

Component 2(ii)

Given: $i < j$, c_{ij} , c_{iji} , c_{ijj} , T , ω , M and current time t_0 .

Goal: Let $t_1 = t_0 + 3(M + 1)T$ and

$$\begin{aligned} \alpha_{i1} &= \sqrt{\frac{c_{ij}}{\pi M}}, & \alpha_{i2} &= \left(\frac{32c_{ij1}^2}{\pi^2 M^2} \right)^{1/6}, & \alpha_{i3} &= \left(\frac{32c_{ijj}^2}{\pi^2 M^2} \right)^{1/6}, \\ \alpha_{j1} &= \frac{c_{ij}}{\alpha_{i1}\pi M}, & \alpha_{j2} &= \frac{2c_{iji}}{\alpha_{i2}^2\pi M}, & \alpha_{j3} &= \frac{-2c_{ijj}}{\alpha_{i3}^2\pi M}. \end{aligned}$$

Specify continuous, zero-mean controls $u_i(t)$ and $u_j(t)$, $t \in [t_0, t_1]$, such that $\epsilon^3 \text{Area}_{ij}(t_1) = c_{ij}$, $\epsilon^3 m_{iji}(t_1) = -c_{iji}$, $\epsilon^3 m_{ijj}(t_1) = -c_{ijj}$ and $\epsilon u_i(t_1) = \epsilon u_i(t_0) = \epsilon u_j(t_1) = \epsilon u_j(t_0) = 0$.

Controls:

$$\begin{aligned}
& \left. \begin{aligned} \epsilon u_i(t) &= \alpha_{i1} \omega \sin(\omega(t - t_0)) \\ \epsilon u_j(t) &= 0 \end{aligned} \right\} t_0 \leq t \leq t_0 + \frac{T}{4} = s_1 \\
& \left. \begin{aligned} \epsilon u_i(t) &= \alpha_{i1} \omega \cos(\omega(t - s_1)) \\ \epsilon u_j(t) &= \alpha_{j1} \omega \sin(\omega(t - s_1)) \end{aligned} \right\} s_1 \leq t \leq s_1 + MT = s_2 \\
& \left. \begin{aligned} \epsilon u_i(t) &= \alpha_{i1} \omega \cos(\omega(t - s_2)) \\ \epsilon u_j(t) &= 0 \end{aligned} \right\} s_2 \leq t \leq s_2 + \frac{3T}{4} = s_3 \\
& \left. \begin{aligned} \epsilon u_i(t) &= \alpha_{i2} \omega \sin(\omega(t - s_3)) \\ \epsilon u_j(t) &= 2\alpha_{j2} \omega \sin(\omega(t - s_3)) \end{aligned} \right\} s_3 \leq t \leq s_3 + \frac{T}{4} = s_4 \\
& \left. \begin{aligned} \epsilon u_i(t) &= \alpha_{i2} \omega \cos(\omega(t - s_4)) \\ \epsilon u_j(t) &= 2\alpha_{j2} \omega \cos(2\omega(t - s_4)) \end{aligned} \right\} s_4 \leq t \leq s_4 + MT = s_5 \\
& \left. \begin{aligned} \epsilon u_i(t) &= \alpha_{i2} \omega \cos(\omega(t - s_5)) \\ \epsilon u_j(t) &= 2\alpha_{j2} \omega \cos(\omega(t - s_5)) \end{aligned} \right\} s_5 \leq t \leq s_5 + \frac{3T}{4} = s_6 \\
& \left. \begin{aligned} \epsilon u_i(t) &= 2\alpha_{j3} \omega \sin(\omega(t - s_6)) \\ \epsilon u_j(t) &= \alpha_{i3} \omega \sin(\omega(t - s_6)) \end{aligned} \right\} s_6 \leq t \leq s_6 + \frac{T}{4} = s_7 \\
& \left. \begin{aligned} \epsilon u_i(t) &= 2\alpha_{j3} \omega \cos(2\omega(t - s_7)) \\ \epsilon u_j(t) &= \alpha_{i3} \omega \cos(\omega(t - s_7)) \end{aligned} \right\} s_7 \leq t \leq s_7 + MT = s_8 \\
& \left. \begin{aligned} \epsilon u_i(t) &= 2\alpha_{j3} \omega \cos(\omega(t - s_8)) \\ \epsilon u_j(t) &= \alpha_{i3} \omega \cos(\omega(t - s_8)) \end{aligned} \right\} s_8 \leq t \leq s_8 + \frac{3T}{4} = s_9
\end{aligned}$$

The condition on $Area_{ij}(t)$ is met during the time interval $[0, s_3]$, the condition on $m_{iji}(t)$ is met during $[s_3, s_6]$ and the condition on $m_{ijj}(t)$ is met during $[s_6, s_9]$.

Component 2(iii)

Given: $i < j < k$, c_{ijk} , c_{ikj} , T , ω , M and current time t_0 .

Goal: Let $t_1 = t_0 + 6(M+1)T$. Let $d_1 = 2(\frac{2}{3}c_{ijk} + \frac{1}{3}c_{ikj})/\pi M$ and $d_2 = 2(\frac{1}{3}c_{ijk} + \frac{2}{3}c_{ikj})/\pi M$. Select

$$\begin{aligned} \rho_{j1} &= \left(\frac{d_1}{6}\right)^{1/3}, & \rho_{i1} &= \left(\left|\frac{d_1}{\rho_{j1}}\right|\right)^{1/2}, & \rho_{k1} &= \frac{d_1}{\rho_{i1}\rho_{j1}}, \\ \rho_{j2} &= \left(\frac{d_2}{6}\right)^{1/3}, & \rho_{i2} &= \left(\left|\frac{d_2}{\rho_{j2}}\right|\right)^{1/2}, & \rho_{k2} &= \frac{d_2}{\rho_{i2}\rho_{j2}}. \end{aligned}$$

We specify continuous, zero-mean controls $u_i(t)$, $u_j(t)$ and $u_k(t)$, $t \in [t_0, t_1]$, such that $\epsilon^3(2m_{ijk}(t_1) - m_{ikj}(t_1)) = -c_{ijk}$, $\epsilon^3(2m_{ikj}(t_1) - m_{ijk}(t_1)) = -c_{ikj}$, $\epsilon u_i(t_1) = \epsilon u_i(t_0) = \epsilon u_j(t_1) = \epsilon u_j(t_0) = \epsilon u_k(t_1) = \epsilon u_k(t_0) = 0$. Further, $Area_{ij}(t_1) = Area_{ik}(t_1) = Area_{jk}(t_1) = 0$.

Controls:

$$\left. \begin{aligned} \epsilon u_i(t) &= \rho_{i1}\omega \sin(\omega(t - t_0)) \\ \epsilon u_j(t) &= 2\rho_{j1}\omega \sin(\omega(t - t_0)) \\ \epsilon u_k(t) &= \rho_{k1}\omega \sin(\omega(t - t_0)) \end{aligned} \right\} t_0 \leq t \leq t_0 + \frac{T}{4} = s_1$$

$$\left. \begin{aligned} \epsilon u_i(t) &= \rho_{i1}\omega \cos(\omega(t - s_1)) \\ \epsilon u_j(t) &= 2\rho_{j1}\omega \cos(2\omega(t - s_1)) \\ \epsilon u_k(t) &= \rho_{k1}\omega \cos(\omega(t - s_1)) \end{aligned} \right\} s_1 \leq t \leq s_1 + MT = s_2$$

$$\left. \begin{aligned} \epsilon u_i(t) &= \rho_{i1}\omega \cos(\omega(t - s_2)) \\ \epsilon u_j(t) &= 2\rho_{j1}\omega \cos(\omega(t - s_2)) \\ \epsilon u_k(t) &= \rho_{k1}\omega \cos(\omega(t - s_2)) \end{aligned} \right\} s_2 \leq t \leq s_2 + \frac{3T}{4} = s_3$$

The condition on $m_{ijk}(t)$ is met during the time interval $[0, s_3]$. However, the values of $m_{iji}(t)$ and $m_{jkk}(t)$ at $t = s_3$ may be different from their initial condition.

So repeat the controls above replacing t_0 with s_3 , s_1 with s_4 , s_2 with s_5 and s_3 with s_6 . Also, replace ρ_{j1} by $-\rho_{j1}$ and set $\epsilon u_k(t) = 0$, $t \in [s_3, s_6]$. During $[s_3, s_6]$,

the original value of $m_{iji}(t)$ is restored. Repeat the controls above again, this time replacing t_0 with s_6 , s_1 with s_7 , s_2 with s_8 and s_3 with s_9 . Also, replace ρ_{j_1} by $-\rho_{j_1}$ and set $\epsilon u_i(t) = 0$, $t \in [s_6, s_9]$. Then, during $[s_6, s_9]$ the original value of $m_{jkk}(t)$ is restored.

Finally, rerun the entire set of controls for $t \in [t_0, s_9]$, exchanging the roles of j and k , augmenting the indices of the time intervals appropriately. Also, replace ρ_{i_1} by ρ_{i_2} , ρ_{j_1} by ρ_{j_2} and ρ_{k_1} by ρ_{k_2} . Then $t_1 = s_{18}$ and $\epsilon^3(2m_{ijk}(t_1) - m_{ikj}(t_1)) = -c_{ijk}$, $\epsilon^3(2m_{ikj}(t_1) - m_{ijk}(t_1)) = -c_{ikj}$. Thus, the goal is met.

Remark 5.7 As discussed in Remark 5.6, Option (1) can be interpreted as the time optimal option and Option (2) as the minimum energy option. The constants $\alpha_{i_1}, \alpha_{i_2}, \alpha_{i_3}, \alpha_{j_1}, \alpha_{j_2}, \alpha_{j_3}$ of Component 2(ii) and $\rho_{i_1}, \rho_{j_1}, \rho_{k_1}, \rho_{i_2}, \rho_{j_2}, \rho_{k_2}$ of Component 2(iii) have been selected to minimize energy analogously to that described for Component 1(ii) in Remark 5.6. In general, we note for all of the algorithms that the energy $\int_0^{t_f} \sum_{k=1}^m (\epsilon u_k(t))^2 dt$ is inversely proportional to t_f .

Remark 5.8 Since the solution to system (2.9) depends analytically on the structure constants, the control inputs generated by the algorithms also depend analytically on the relevant structure constants, i.e., the structure constants associated with the basis $\{\xi_1, \dots, \xi_n\}$ for $\mathcal{G}$ that appear in Part A of the algorithms. Small errors in system description are equivalent to small errors in structure constants. Because the specified controls depend analytically on the structure constants, these errors translate into small errors in specified controls. Thus, the open-loop controls derived from these algorithms are robust to modelling errors.

5.2.4 Examples

Spacecraft

Consider the spacecraft attitude control problem of Section 3.1 where only two components of angular velocity (about the $\mathbf{b}_1$ and $\mathbf{b}_2$ axes) can be controlled. Then, from (3.3), $X(t) \in G = SO(3)$ and

$$\dot{X} = \epsilon X(u_1 A_1 + u_2 A_2), \quad (5.21)$$

where $\{A_1, A_2, A_3\}$ is the standard basis for $so(3)$. The only relevant nonzero structure constant is $\Gamma_{12}^3 = 1$ since $[A_1, A_2] = A_3$. Thus, from Theorem 4.3 and Corollary 4.16,

$$\begin{aligned} z_1^{(2)}(t) &= \epsilon \tilde{u}_1(t) + z_{10}^{(2)}, \\ z_2^{(2)}(t) &= \epsilon \tilde{u}_2(t) + z_{20}^{(2)}, \\ z_3^{(2)}(t) &= \epsilon^2 \frac{t}{T} \text{Area}_{12}(T) + z_{30}^{(2)}, \end{aligned} \quad (5.22)$$

$$X^{(2)}(t) = e^{\sum_{i=1}^3 z_i^{(2)}(t) A_i} \quad \text{or} \quad X^{(2)}(t) = e^{z_1^{(2)}(t) A_1} e^{z_2^{(2)}(t) A_2} e^{z_3^{(2)}(t) A_3},$$

and $X^{(2)}(t)$ is an $O(\epsilon^2)$ approximation of $X(t)$ on an $O(1/\epsilon)$ time interval. This result implies that *oscillations* about the $\mathbf{b}_1$ and $\mathbf{b}_2$ axes can produce a net rotation about the $\mathbf{b}_3$ axis. We note that this result and the control algorithm below apply equally to the ball and plate example of Section 3.4 (with the roles of $\mathbf{b}_i$ and $\mathbf{r}_i$ reversed) since the ball and plate problem is also described by (5.21) with $G = SO(3)$.

Since (5.21) is a depth-one bracket system, we use Algorithm 1 to construct open-loop controls to solve $(\mathbf{P})$ with $O(\epsilon^2)$ accuracy. We are given $m = 2$, $n = 3$, $X(0) = g_i = I$ and $\gamma_f = (\gamma_{f1}, \gamma_{f2}, \gamma_{f3})^T = \Psi^{-1}(g_f) = O(\epsilon)$ and $t_f > 0$. We

set $z_f = \gamma_f$. The matrix $\hat{\Gamma}$ is $\hat{\Gamma} = \Gamma_{12} = \Gamma_{12}^3 = 1$. Options (1) and (2) yield $\Gamma = \hat{\Gamma}^\dagger = 1$. Thus, since $\Gamma_{12}^1 = \Gamma_{12}^2 = 0$,

$$c_{12} = \gamma_{f3}/\Gamma = \gamma_{f3}, \quad c_1 = \gamma_{f1} - c_{12}\Gamma_{12}^1 = \gamma_{f1}, \quad c_2 = \gamma_{f2} - c_{12}\Gamma_{12}^2 = \gamma_{f2},$$

and $S = \{1\}$, $r = 1$. We choose an integer $M \geq 1/\pi\epsilon$ and

$$T = \frac{t_f}{M + 3/2}, \quad \omega = \frac{2\pi}{T}.$$

Then we apply Component 1(ii) for c_{12} followed by Component 1(i) for c_1 and c_2 . Let $t_1 = \frac{T}{4}$, $t_2 = t_1 + MT$, $t_3 = t_2 + \frac{3T}{4}$, $t_4 = t_f = t_3 + \frac{T}{2}$. The controls are then defined as follows:

$$\begin{aligned} \epsilon u_1(t) &= \begin{cases} \alpha_1 \omega \sin \omega t & 0 \leq t \leq t_3 \\ \frac{1}{2} c_1 \omega \sin(\omega(t - t_3)) & t_3 \leq t \leq t_4 \end{cases} \\ \epsilon u_2(t) &= \begin{cases} 0 & 0 \leq t \leq t_1 \\ \alpha_2 \omega \sin(\omega(t - t_1)) & t_1 \leq t \leq t_2 \\ 0 & t_2 \leq t \leq t_3 \\ \frac{1}{2} c_2 \omega \sin(\omega(t - t_3)) & t_3 \leq t \leq t_4 \end{cases} \end{aligned} \quad (5.23)$$

where $\alpha_1 = \sqrt{|c_{12}|/\pi M}$ and $\alpha_2 = c_{12}/\alpha_1 \pi M$.

For numerical illustration, let $\epsilon = 0.1$, $\gamma_{f1} = 0.085$, $\gamma_{f2} = 0.03$, $\gamma_{f3} = 0.2$ and $t_f = 23$. Choose $M = 10$, then $T = 2$, $\omega = \pi$, $\alpha_1 = \alpha_2 = 0.08$, $t_1 = 0.5$, $t_2 = 20.5$, $t_3 = 22$ and $t_4 = 23$. Figure 5.1 shows plots of the corresponding controls ϵu_1 and ϵu_2 as a function of time. Figures 5.2(a), 5.2(b) and 5.2(d) show a simulation of the response of the Wei-Norman parameters, γ_1 , γ_2 , γ_3 for the attitude control problem (solid lines). Figure 5.2(c) is a plot of $\tilde{u}_1$ versus $\tilde{u}_2$ for $t \in [t_1, t_2]$, and the area enclosed by the curve is equivalent to $Area_{12}(T)$. The simulation was produced by numerically solving the equations (3.4) using

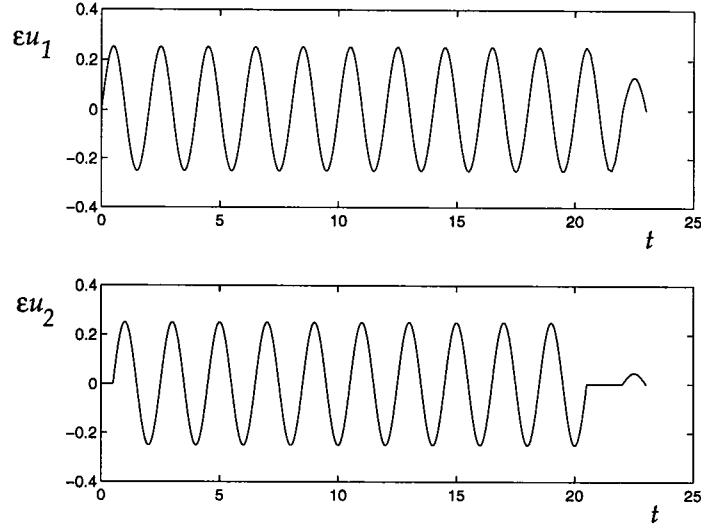


Figure 5.1: Control Input Signals for Spacecraft and Unicycle Examples.

MATLAB^R. The dashed lines of Figure 5.2 represent the $O(\epsilon^2)$ average approximation of the parameters as a function of time computed directly from the average formulas (5.22). Thus, the dashed lines represent the prescribed trajectory of the parameters. Figure 5.2 shows that the actual parameters (solid lines) follow the prescribed average solution (dashed lines) with $O(\epsilon^2)$ accuracy as expected. From Figure 5.2 we see that $\|\gamma(t)\|_2 < \pi$, $\forall t \in [0, t_f]$ which implies that $\|z(t)\|_2 < \pi$, $\forall t \in [0, t_f]$. By Theorem 2.8 (Lazard and Tits) and Remark 2.9 for $\mathcal{G} = so(3)$ we set $\hat{S} = \{\sum_{i=1}^3 a_i A_i \in so(3) \mid \|a\|_2 < \pi\}$. Then, $Z(t) \in \hat{S}$, $\forall t \in [0, t_f]$ and so $\hat{d}(X(t), X^{(2)}(t)) = O(\epsilon^2)$, $\forall t \in [0, t_f]$. In particular, $\hat{d}(X(t_f), g_f) = O(\epsilon^2)$, i.e., the spacecraft has been reoriented as desired with $O(\epsilon^2)$ accuracy.

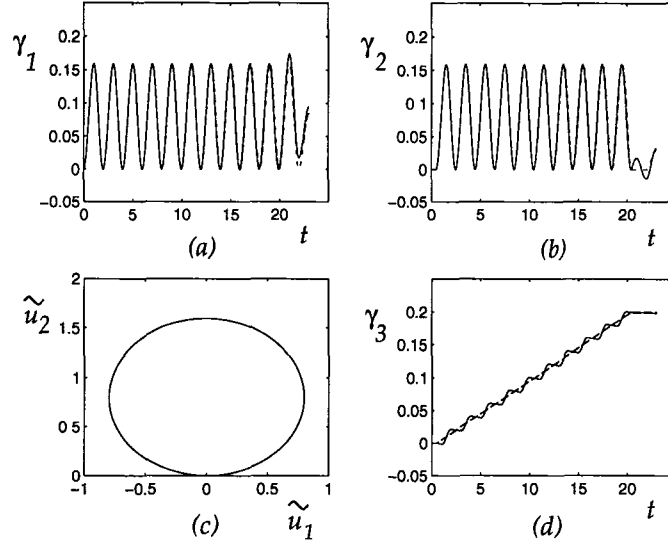


Figure 5.2: Actual (solid lines) and Average (dashed lines) Wei-Norman Parameters for Spacecraft Attitude Control Example.

Unicycle

Consider the unicycle motion control problem of Section 3.2. From (3.11) $X(t) \in G = SE(2)$ describes the position and orientation of the unicycle and satisfies

$$\dot{X} = \epsilon X(A_1 u_1 + A_2 u_2), \quad (5.24)$$

where $\{A_1, A_2, A_3\}$ is the basis for $se(2)$ defined by (3.10). This system is a depth-one bracket system and the only relevant nonzero structure constant is $\Gamma_{12}^3 = 1$ since $[A_1, A_2] = A_3$. The relevant structure constants for this system are the same as those for the spacecraft attitude control problem with two controls. Therefore, $X^{(2)}(t)$ described by equations (5.22), with $\{A_1, A_2, A_3\}$ interpreted as the basis for $se(2)$, is an $O(\epsilon^2)$ average approximation of $X(t) \in SE(2)$ on an $O(1/\epsilon)$ time interval. Moreover, Algorithm 1 applied to the unicycle problem yields the same controls defined by (5.23). Applying the control input of (5.23)

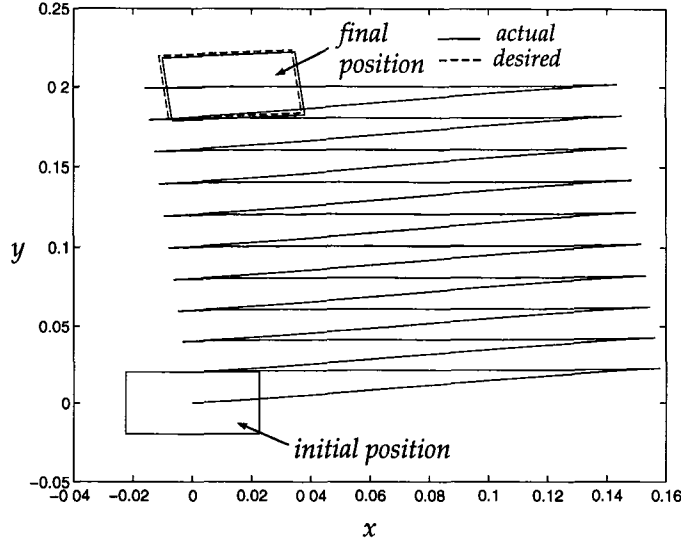


Figure 5.3: Response for Unicycle Example.

to the unicycle problem, with the same values for ϵ , γ_f , t_f , etc., used to illustrate the spacecraft example, yields unicycle Wei-Norman parameter responses similar to those shown in Figure 5.2. Figure 5.3 shows the response of the unicycle in (x, y, θ) coordinates. One can observe that the unicycle has been repositioned and reoriented with $O(\epsilon^2)$ accuracy.

In general, the controls defined by (5.23) solve **(P)** with $O(\epsilon^2)$ accuracy for *any* single-bracket system of the form (2.9) with $n = 3$ and $m = 2$. One simply needs to compute $c_{12} = z_{f3}/\Gamma_{12}^3$, $c_1 = z_{f1} - c_{12}\Gamma_{12}^1$ and $c_2 = z_{f2} - c_{12}\Gamma_{12}^2$. By definition, $\Gamma_{12}^3 \neq 0$. Brockett's system of Section 3.5 is an example. In this case, in addition, $G = H(3)$ is nilpotent of order 2, and, as a consequence of Theorem 4.9, the controls of (5.23) will solve **(P)** for this system exactly. By Theorem 4.9, $X(t) \in H(3)$ is given by $X(t) = e^{Z^{[2]}(t)}$. By Theorems 4.10 and 4.11, for $t = NT$, N an integer, and $U_{av} = 0$, $Z^{[2]}(t) = Z^{(2)}(t)$. So, $X(NT) = e^{Z^{(2)}(NT)} = X^{(2)}(NT)$.

Underwater Vehicle

Consider the autonomous underwater vehicle (AUV) motion control problem of Section 3.3. First, suppose that we can control angular velocity about the $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3$ axes and translational velocity along the $\mathbf{b}_1$ axis. Then from (3.14), $X(t) \in G = SE(3)$ describes the orientation and position of the vehicle and satisfies

$$\dot{X} = \epsilon X \left(\sum_{i=1}^4 u_i(t) A_i \right), \quad (5.25)$$

where $\{A_1, \dots, A_6\}$ is the basis for $se(3)$ defined by (3.12). This system is a depth-one bracket systems. The nonzero structure constants corresponding to our chosen basis for $se(3)$ can easily be computed as $\Gamma_{12}^3 = \Gamma_{23}^1 = \Gamma_{31}^2 = 1$, $\Gamma_{15}^6 = \Gamma_{61}^5 = 1$, $\Gamma_{42}^6 = \Gamma_{26}^4 = 1$, $\Gamma_{34}^5 = \Gamma_{53}^4 = 1$. Thus, from Theorem 4.3 and Corollary 4.16,

$$\begin{aligned} z_1^{(2)}(t) &= \epsilon \tilde{u}_1(t) + \epsilon^2 \frac{t}{T} Area_{23}(T) + z_{10}^{(2)}, & z_4^{(2)}(t) &= \epsilon \tilde{u}_4(t) + z_{40}^{(2)}, \\ z_2^{(2)}(t) &= \epsilon \tilde{u}_2(t) - \epsilon^2 \frac{t}{T} Area_{13}(T) + z_{20}^{(2)}, & z_5^{(2)}(t) &= \epsilon^2 \frac{t}{T} Area_{34}(T) + z_{50}^{(2)}, \\ z_3^{(2)}(t) &= \epsilon \tilde{u}_3(t) + \epsilon^2 \frac{t}{T} Area_{12}(T) + z_{30}^{(2)}, & z_6^{(2)}(t) &= -\epsilon^2 \frac{t}{T} Area_{24}(T) + z_{60}^{(2)}, \end{aligned} \quad (5.26)$$

$$X^{(2)}(t) = e^{\sum_{i=1}^6 z_i^{(2)}(t) A_i} \quad \text{or} \quad X^{(2)}(t) = \prod_{i=1}^6 e^{z_i^{(2)}(t) A_i},$$

and $X^{(2)}(t)$ is an $O(\epsilon^2)$ approximation of $X(t)$ on an $O(1/\epsilon)$ time interval.

Since (5.25) is a depth-one bracket system, we use Algorithm 1 to construct open-loop controls to solve $(\mathbf{P})$ with $O(\epsilon^2)$ accuracy. We are given $m = 4$, $n = 6$, $X(0) = g_i = I$ and $\gamma_f = (\gamma_{f1}, \dots, \gamma_{f6})^T = \Psi^{-1}(g_f) = O(\epsilon)$ and $t_f > 0$. Set $z_f = \gamma_f$. We compute

$$\hat{\Gamma} = \begin{bmatrix} \Gamma_{12} & \Gamma_{13} & \Gamma_{14} & \Gamma_{23} & \Gamma_{24} & \Gamma_{34} \end{bmatrix}$$

$$\begin{aligned}
&= \begin{bmatrix} \Gamma_{12}^5 & \Gamma_{13}^5 & \Gamma_{14}^5 & \Gamma_{23}^5 & \Gamma_{24}^5 & \Gamma_{34}^5 \\ \Gamma_{12}^6 & \Gamma_{13}^6 & \Gamma_{14}^6 & \Gamma_{23}^6 & \Gamma_{24}^6 & \Gamma_{34}^6 \end{bmatrix} \\
&= \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & -1 & 0 \end{bmatrix}.
\end{aligned}$$

Options (1) and (2) of Part A yield

$$\Gamma = \begin{bmatrix} \Gamma_{24}^5 & \Gamma_{34}^5 \\ \Gamma_{24}^6 & \Gamma_{34}^6 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \quad \hat{\Gamma}^\dagger = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & -1 & 0 \end{bmatrix}^T.$$

Therefore, both options give

$$\begin{bmatrix} c_{24} \\ c_{34} \end{bmatrix} = \Gamma^{-1} \begin{bmatrix} z_{f5} \\ z_{f6} \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} z_{f5} \\ z_{f6} \end{bmatrix} = \begin{bmatrix} -z_{f6} \\ z_{f5} \end{bmatrix},$$

$$c_{12} = c_{13} = c_{14} = c_{23} = 0,$$

$$c_1 = z_{f1} - c_{24}\Gamma_{24}^1 - c_{34}\Gamma_{34}^1 = z_{f1}, \quad c_2 = z_{f2} - c_{24}\Gamma_{24}^2 - c_{34}\Gamma_{34}^2 = z_{f2},$$

$$c_3 = z_{f3} - c_{24}\Gamma_{24}^3 - c_{34}\Gamma_{34}^3 = z_{f3}, \quad c_4 = z_{f4} - c_{24}\Gamma_{24}^4 - c_{34}\Gamma_{34}^4 = z_{f4},$$

$$S = \{2, 3\}, \quad r = 2.$$

So we choose an integer $M \geq 1/\pi\epsilon$ and

$$T = \frac{t_f}{2(M+1) + 1/2}, \quad \omega = \frac{2\pi}{T}.$$

Then we apply Component 1(ii) for c_{24} followed by Component 1(ii) for c_{34} followed by Component 1(i) for c_1, c_2, c_3, c_4 . To reduce the time and energy expended by the controls we can instead apply Component 1(ii) just once. To do this let $i = 4, j = 2, 3, c_{42} = -c_{24}, c_{43} = -c_{34}$ and apply Component 1(ii) to match c_{42} and c_{43} . In this case we have $S = \{4\}$ and $r = 1$ so we recompute

$$T = \frac{t_f}{M + 3/2}, \quad \omega = \frac{2\pi}{T}.$$

Let $t_1 = \frac{T}{4}$, $t_2 = t_1 + MT$, $t_3 = t_2 + \frac{3T}{4}$, $t_4 = t_f = t_3 + \frac{T}{2}$. The controls are then defined as follows:

$$\begin{aligned}
\epsilon u_1(t) &= \begin{cases} 0 & 0 \leq t \leq t_3 \\ \frac{1}{2}c_1\omega \sin(\omega(t - t_3)) & t_3 \leq t \leq t_4 \end{cases} \\
\epsilon u_2(t) &= \begin{cases} 0 & 0 \leq t \leq t_1 \\ \alpha_2\omega \sin(\omega(t - t_1)) & t_1 \leq t \leq t_2 \\ 0 & t_2 \leq t \leq t_3 \\ \frac{1}{2}c_2\omega \sin(\omega(t - t_3)) & t_3 \leq t \leq t_4 \end{cases} \\
\epsilon u_3(t) &= \begin{cases} 0 & 0 \leq t \leq t_1 \\ \alpha_3\omega \sin(\omega(t - t_1)) & t_1 \leq t \leq t_2 \\ 0 & t_2 \leq t \leq t_3 \\ \frac{1}{2}c_3\omega \sin(\omega(t - t_3)) & t_3 \leq t \leq t_4 \end{cases} \\
\epsilon u_4(t) &= \begin{cases} \alpha_4\omega \sin \omega t & 0 \leq t \leq t_3 \\ \frac{1}{2}c_4\omega \sin(\omega(t - t_3)) & t_3 \leq t \leq t_4 \end{cases}
\end{aligned} \tag{5.27}$$

where $\alpha_4 = (\frac{c_{42}^2 + c_{43}^2}{\pi^2 M^2})^{1/4}$, $\alpha_2 = c_{42}/\alpha_4\pi M$, $\alpha_3 = c_{43}/\alpha_4\pi M$.

For numerical illustration, let $\epsilon = 0.1$, $z_{fi} = \gamma_{fi} = 0.1$, $i = 1, \dots, 6$, and $t_f = 23$. Choose $M = 10$, then $T = 2$, $\omega = \pi$, $\alpha_4 = 0.067$, $\alpha_2 = 0.047$, $\alpha_3 = -0.047$, $t_1 = 0.5$, $t_2 = 20.5$, $t_3 = 22$ and $t_4 = 23$. Figure 5.4 shows plots of the corresponding controls ϵu_1 and ϵu_2 as a function of time. Figure 5.5 shows a simulation of the response of the Wei-Norman parameters γ as a function of time. The simulation was produced by numerically solving the equations (3.17) using MATLAB. The horizontal dashed lines of Figure 5.5 represent the desired final parameter value γ_f . Figure 5.5 shows that $\gamma(t_f) - \gamma_f = O(\epsilon^2)$ as expected. By Theorem 2.8 (Lazard and Tits) and Remark 2.9 for $\mathcal{G} = se(3)$ we can let

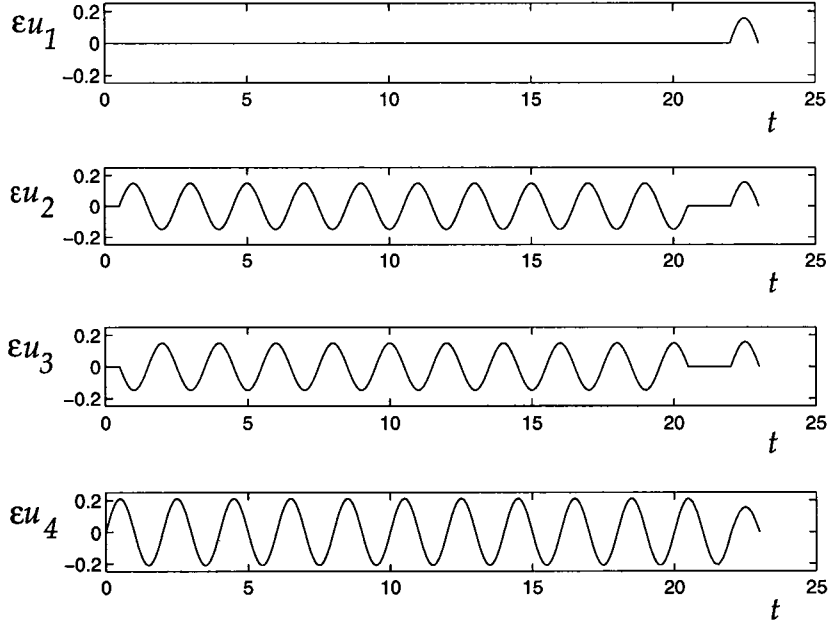


Figure 5.4: Control Input Signals for AUV Example with Four Controls.

$\hat{S} = \{A \in se(3) \mid \|A\|_p < \pi/2\}$ for any p . From Figure 5.5 it is clear that $Z_f \in \hat{S}$ and $Z(t_f) \in \hat{S}$ for some choice of p . Thus, $\|Z(t_f) - Z_f\| = O(\epsilon^2)$ implies $\hat{d}(X(t_f), g_f) = O(\epsilon^2)$, i.e., the AUV has been repositioned and reoriented as desired with $O(\epsilon^2)$ accuracy.

Now consider the case when there are only three controls available: angular velocity control about the $\mathbf{b}_1$ and $\mathbf{b}_2$ axes and translational velocity control along the $\mathbf{b}_1$ axis. By (3.15) $X(t) \in SE(3)$ satisfies

$$\dot{X} = \epsilon X(u_1 A_1 + u_2 A_2 + u_4 A_4), \quad (5.28)$$

which is a depth-two bracket system. The relevant nonzero depth-two structure

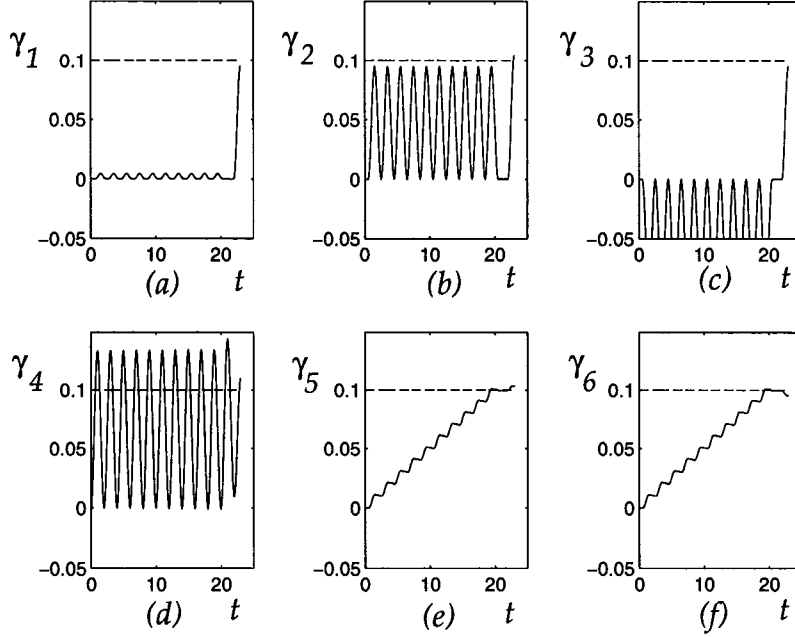


Figure 5.5: Response of AUV with Four Controls.

constants are $\theta_{122}^1 = -1$, $\theta_{121}^2 = \theta_{242}^4 = \theta_{124}^5 = 1$. From Corollary 4.18,

$$\begin{aligned} z_1^{(3)}(t) &= \epsilon \tilde{u}_1(t) + \frac{\epsilon^3 t}{T} m_{122}(T), & z_4^{(3)}(t) &= \epsilon \tilde{u}_4(t) - \frac{\epsilon^3 t}{T} m_{242}(T), \\ z_2^{(3)}(t) &= \epsilon \tilde{u}_2(t) - \frac{\epsilon^3 t}{T} m_{121}(T), & z_5^{(3)}(t) &= -\frac{\epsilon^3 t}{T} (2m_{124}(T) - m_{142}(T)), \\ z_3^{(3)}(t) &= \epsilon^2 a_{12}(t), & z_6^{(3)}(t) &= -\epsilon^2 a_{24}(t). \end{aligned}$$

$$X^{(3)}(t) = e^{\sum_{i=1}^6 z_i^{(3)}(t) A_i},$$

and $X^{(3)}(t)$ is an $O(\epsilon^3)$ approximation of $X(t)$ on an $O(1/\epsilon)$ time interval.

Since (5.28) is a depth-two bracket system, we use Algorithm 2 to construct open-loop controls to solve **(P)** with $O(\epsilon^3)$ accuracy. For the purpose of the algorithm, we reorder our basis for $se(3)$ such that $A_3 \Leftrightarrow A_4$ and $A_5 \Leftrightarrow A_6$ (see the first paragraph of Section 5.2.3). The nonzero structure constants associated with this reordered basis become $\Gamma_{12}^4 = \Gamma_{24}^1 = \Gamma_{41}^2 = 1$, $\Gamma_{16}^5 = \Gamma_{51}^6 = 1$, $\Gamma_{32}^5 = \Gamma_{25}^3 = 1$,

$\Gamma_{43}^6 = \Gamma_{64}^3 = 1$. Further, $\theta_{123}^6 = 1$. Thus, $n = 6$, $m = 3$, and so by (5.18), $\Xi = \{4, 5\}$ and $l = |\Xi| = 2$. Thus, we get

$$\begin{aligned}\hat{\theta} &= [\theta_{121} \ \theta_{122} \ \theta_{123} \ \theta_{131} \ \theta_{132} \ \theta_{133} \ \theta_{231} \ \theta_{232} \ \theta_{233}] \\ &= [\theta_{121}^6 \ \theta_{122}^6 \ \theta_{123}^6 \ \theta_{131}^6 \ \theta_{132}^6 \ \theta_{133}^6 \ \theta_{231}^6 \ \theta_{232}^6 \ \theta_{233}^6] \\ &= [0 \ 0 \ 1 \ 0 \ 0 \ 0 \ -1 \ 0 \ 0],\end{aligned}$$

$$\begin{aligned}\hat{\Gamma} &= [\Gamma_{12} \ \Gamma_{13} \ \Gamma_{23}] \\ &= \begin{bmatrix} \Gamma_{12}^4 & \Gamma_{13}^4 & \Gamma_{23}^4 \\ \Gamma_{12}^5 & \Gamma_{13}^5 & \Gamma_{23}^5 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix}.\end{aligned}$$

Options (1) and (2) yield the same selection of constants. Following Option (1),

$$\theta = \theta_{123} = \theta_{123}^6 = 1.$$

$$c_{123} = \theta^{-1} z_{f_5} = z_{f_5},$$

$$c_{121} = c_{122} = c_{231} = c_{232} = c_{233} = c_{131} = c_{132} = c_{133} = 0.$$

Note that z_{f_5} is based on the original ordering of the basis for $se(3)$ and so it is the coefficient of A_6 in the reordered basis.

$$\Gamma = \begin{bmatrix} \Gamma_{12}^4 & \Gamma_{23}^4 \\ \Gamma_{12}^5 & \Gamma_{23}^5 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

So

$$\begin{bmatrix} c_{12} \\ c_{23} \end{bmatrix} = \Gamma^{-1} \begin{bmatrix} z_{f_3} - c_{123}\theta_{123}^4 \\ z_{f_6} - c_{123}\theta_{123}^5 \end{bmatrix} = \Gamma^{-1} \begin{bmatrix} z_{f_3} \\ z_{f_6} \end{bmatrix} = \begin{bmatrix} z_{f_3} \\ -z_{f_6} \end{bmatrix},$$

$$c_{13} = 0,$$

$$c_1 = z_{f_1} - (c_{12}\Gamma_{12}^1 + c_{23}\Gamma_{23}^1 + c_{123}\theta_{123}^1) = z_{f_1},$$

$$c_2 = z_{f_2} - (c_{12}\Gamma_{12}^2 + c_{23}\Gamma_{23}^2 + c_{123}\theta_{123}^2) = z_{f_2},$$

$$c_3 = z_{f_4} - (c_{12}\Gamma_{12}^3 + c_{23}\Gamma_{23}^3 + c_{123}\theta_{123}^3) = z_{f_4},$$

$$Y = \{c_{123}\}, \quad Q = \{c_{123}\}, \quad R = \emptyset, \quad V = \{c_{12}, c_{23}\}, \quad W = \emptyset, \quad \beta = 1, \quad \gamma = 2.$$

So we choose an integer $M \geq 1/\pi\epsilon$ and

$$T = \frac{t_f}{12(M+1) + 1/2}, \quad \omega = \frac{2\pi}{T}.$$

Then we apply Component 2(iii) for c_{123} , followed by Component 2(ii) for c_{12} , followed by Component 2(ii) for c_{23} , followed by Component 2(i) for c_1, c_2, c_3 . These components will specify controls u_1, u_2 and u_3 . However, u_3 is really our original control u_4 since it is the coefficient of the original A_4 .

For this particular system we note that the algorithm is longer than necessary, i.e., there are steps which have zero net effect on the system. Thus, to save time and energy we eliminate the unnecessary steps of the control algorithm defined above. Specifically, since $c_{233} = 0$ and $c_{132} = 0$ we can leave out the controls defined during $[s_6, s_{18}]$ of Component 2(i) (this requires a change in the computation of d_1 to $d_1 = c_{ijk}/\pi M$). Additionally, since $c_{121} = c_{122} = c_{232} = c_{233} = 0$ we can leave out the controls defined during $[s_3, s_9]$ of Component 2(ii) each time it is applied. Further, we can apply the first part of Component 2(ii) just once for both c_{12} and c_{23} as in Component 1(ii) of Algorithm 1. The total time duration of the parts left out is $9(M+1)T$ so we recompute

$$T = \frac{t_f}{3(M+1) + 1/2}, \quad \omega = \frac{2\pi}{T}.$$

Let $t_1 = \frac{T}{4}$, $t_2 = t_1 + MT$, $t_3 = t_2 + \frac{3T}{4}$, $t_4 = t_3 + \frac{T}{4}$, $t_5 = t_4 + MT$, $t_6 = t_5 + \frac{3T}{4}$, $t_7 = t_6 + \frac{T}{4}$, $t_8 = t_7 + MT$, $t_9 = t_8 + \frac{3T}{4}$, $t_{10} = t_9 + \frac{T}{2}$. The controls are then defined as follows:

$$\epsilon u_1(t) = \begin{cases} \rho_1 \omega \sin \omega t & 0 \leq t \leq t_6 \\ 0 & t_6 \leq t \leq t_7 \\ \alpha_1 \omega \sin(\omega(t - t_7)) & t_7 \leq t \leq t_8 \\ 0 & t_8 \leq t \leq t_9 \\ \frac{1}{2} c_1 \omega \sin(\omega(t - t_9)) & t_9 \leq t \leq t_{10} \end{cases}$$

$$\epsilon u_2(t) = \begin{cases} 2\rho_2\omega \sin \omega t & 0 \leq t \leq t_1 \\ 2\rho_2\omega \cos(2\omega(t - t_1)) & t_1 \leq t \leq t_2 \\ 2\rho_2\omega \cos(\omega(t - t_2)) & t_2 \leq t \leq t_3 \\ -2\rho_2\omega \sin(\omega(t - t_3)) & t_3 \leq t \leq t_4 \\ -2\rho_2\omega \cos(2\omega(t - t_4)) & t_4 \leq t \leq t_5 \\ -2\rho_2\omega \cos(\omega(t - t_5)) & t_5 \leq t \leq t_6 \\ \alpha_2 \sin(\omega(t - t_6)) & t_6 \leq t \leq t_9 \\ \frac{1}{2}c_2\omega \sin(\omega(t - t_9)) & t_9 \leq t \leq t_{10} \end{cases} \quad (5.29)$$

$$\epsilon u_4(t) = \begin{cases} \rho_4\omega \sin \omega t & 0 \leq t \leq t_3 \\ 0 & t_3 \leq t \leq t_7 \\ \alpha_4\omega \sin(\omega(t - t_7)) & t_7 \leq t \leq t_8 \\ 0 & t_8 \leq t \leq t_9 \\ \frac{1}{2}c_4\omega \sin(\omega(t - t_9)) & t_9 \leq t \leq t_{10} \end{cases}$$

where $\rho_2 = (c_{123}/6\pi M)^{1/3}$, $\rho_1 = (|c_{123}/\rho_2\pi M|)^{1/2}$, $\rho_4 = c_{123}/\rho_1\rho_2\pi M$, $c_{21} = -c_{12}$, $\alpha_2 = ((c_{23}^2 + c_{21}^2)/\pi^2 M^2)^{1/4}$, $\alpha_1 = c_{21}/\alpha_2\pi M$, $\alpha_4 = c_{23}/\alpha_2\pi M$.

For numerical illustration, let $\epsilon = 0.2$, $t_f = 37$ and $\gamma_{f_1} = 0.05$, $\gamma_{f_2} = 0.05$, $\gamma_{f_3} = 0.04$, $\gamma_{f_4} = 0.06$, $\gamma_{f_5} = 0.05$, $\gamma_{f_6} = 0.05$. Let $z_f = \gamma_f$. Choose $M = 5$, then $T = 2$, $\omega = \pi$, $\rho_2 = 0.08$, $\rho_1 = 0.20$, $\rho_4 = 0.20$, $\alpha_2 = 0.06$, $\alpha_1 = -0.04$, $\alpha_4 = -0.05$. Figure 5.6 shows plots of the corresponding controls ϵu_1 , ϵu_2 and ϵu_4 as a function of time. Figure 5.7 shows a simulation of the response of the Wei-Norman parameters γ as a function of time. The horizontal dashed lines of Figure 5.7 represent the desired final parameter values γ_f . Figure 5.7 shows that $\gamma(t_f) - \gamma_f = O(\epsilon^3)$ as expected. We conclude that the AUV has been repositioned and reoriented as desired with $O(\epsilon^3)$ accuracy.

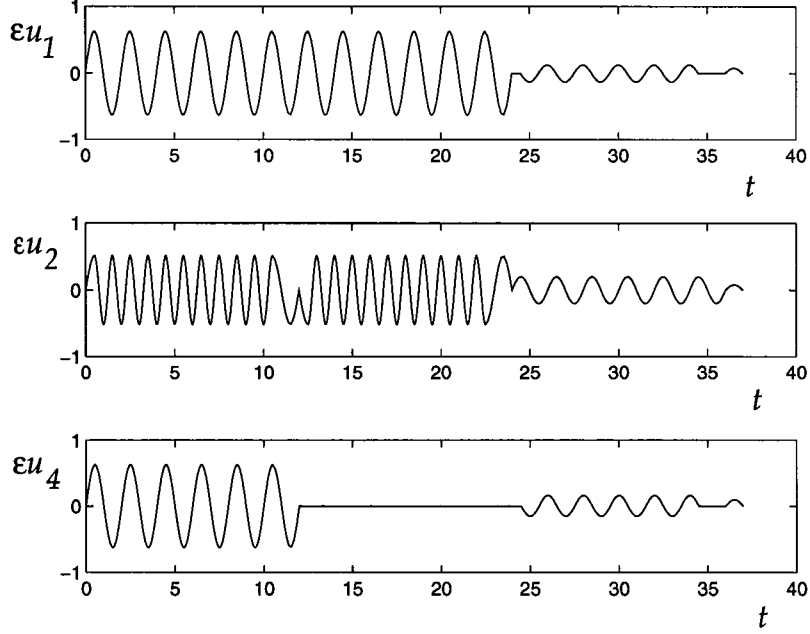


Figure 5.6: Control Input Signals for AUV Example with Three Controls.

5.3 Adaptation to Changes in Control Authority

In this section we describe a control architecture that uses open-loop control generating algorithms such as those of Section 5.2 to provide adaptation to changes in control authority for drift-free systems of the form (2.9) on the Lie group G with Lie algebra $\mathcal{G}$. *Control authority* is defined to be the particular choice of m control inputs which can be actuated, i.e., the choice of $\xi_1, \dots, \xi_m \in \mathcal{G}$ such that $U(t) = \sum_{i=1}^m u_i(t)\xi_i$ and $u_1(t), \dots, u_m(t)$ can be actuated independently.

Given a Lie algebra $\mathcal{G}$, there is a minimal number $m_0 \leq n$ of controls and at least one choice of m_0 controls for which the system is controllable. If the original choice of control authority includes $m > m_0$ controls, then even in the event of a

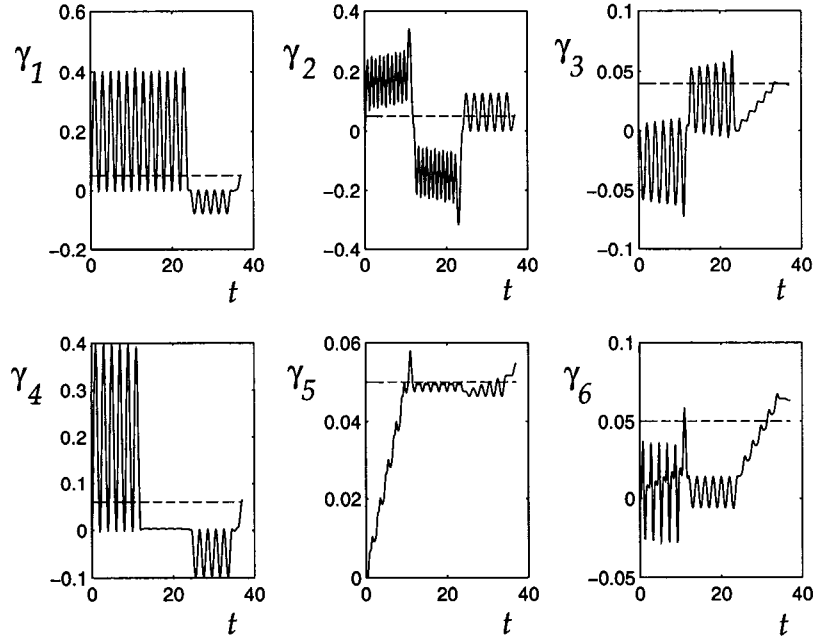


Figure 5.7: Response of AUV with Three Controls.

change in control authority that reduces the number of available controls (such as an actuator failure), the system might still remain controllable. In this case, we can use the algorithms of Section 5.2 (and higher-order algorithms that can be derived according to the procedure of Section 5.1) to maintain control over the system.

As described in Remark 5.8, the open-loop controls depend essentially on the relevant structure constants associated to the basis $\{\xi_1, \dots, \xi_n\}$ and the subset $\{\xi_1, \dots, \xi_m\}$. A change in control authority changes the structure of the system by changing the relevant structure constants since, in particular, the subset $\{\xi_1, \dots, \xi_m\}$ is changed. We can adapt to such a structural change if the system remains controllable by using the new relevant structure constants in the appropriate algorithm to recompute open-loop controls.

For instance, consider a depth-one bracket system with m controls and suppose that during operation there is a change in control authority such that with the remaining $m - 1$ control inputs the system is a depth-two bracket system. Then to adapt to the change in control authority, we switch from the open-loop controls (motion script) generated by Algorithm 1 to open-loop controls (motion script) generated by Algorithm 2. To make the switch on the fly after such a failure, one could first terminate the single-bracket system motion script by bringing the active controls to zero and then begin controlling the system according to the double-bracket system motion script.

The AUV motion control problem of Section 3.3 with control authority including four control inputs is an example of such a system. Suppose the control authority includes the four controls as in the example of Section 5.2, i.e., angular velocity control about the $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3$ axes and translational velocity control along the $\mathbf{b}_1$ axis. The system

$$\dot{X} = \epsilon X \left(\sum_{i=1}^4 u_i(t) A_i \right),$$

where $X(t) \in SE(3)$ and A_i are defined by (3.12), is a depth-one bracket system. During normal operation we can control the AUV with this control authority using the open-loop controls of (5.27). However, in the event of an actuator failure these controls may not be appropriate. If, during the course of operation, one of the actuators for angular velocity control fails, the system is still controllable. As discussed in Section 3.3, if u_1 fails then the system remains a depth-one bracket system. In this case, we can continue to control the AUV with open-loop controls computed from Algorithm 1 based on the new control authority. If u_2 or u_3 fails, the system becomes a depth-two bracket system. In this case, we can continue to control the AUV with open-loop controls computed from Algorithm 2 based

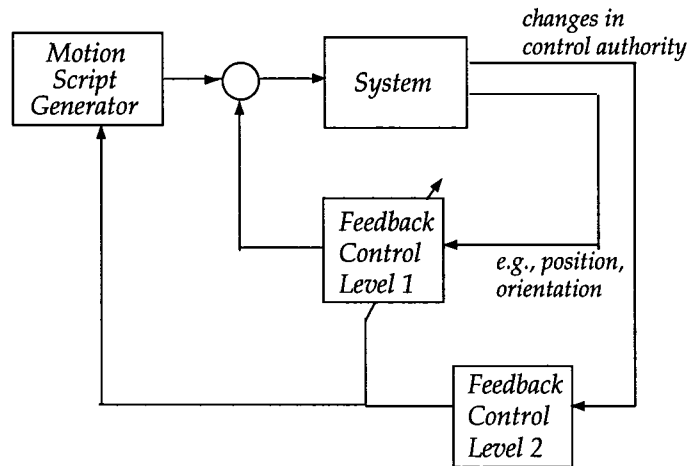


Figure 5.8: Control Architecture.

on the new control authority. For a failure of u_3 we use the controls specified by (5.29).

Figure 5.8 represents the type of overall control architecture that we envision to incorporate open-loop control planning and intermittent feedback (described earlier in the chapter), as well as a second level of feedback to allow for adaptation to changes in control authority. The motion script generator represents algorithms such as those of Section 5.2 which produce open-loop controls (motion scripts) to achieve desired motions. The first level of feedback control complements the open-loop control by adding (possibly intermittent) feedback to make the motion control more robust. The second level of feedback control is used for adaptation to changes in control authority such as an actuator failure. When such a change is sensed the open-loop control generator switches motion scripts as necessary. The first level of feedback control is also updated to account for the change in control authority.

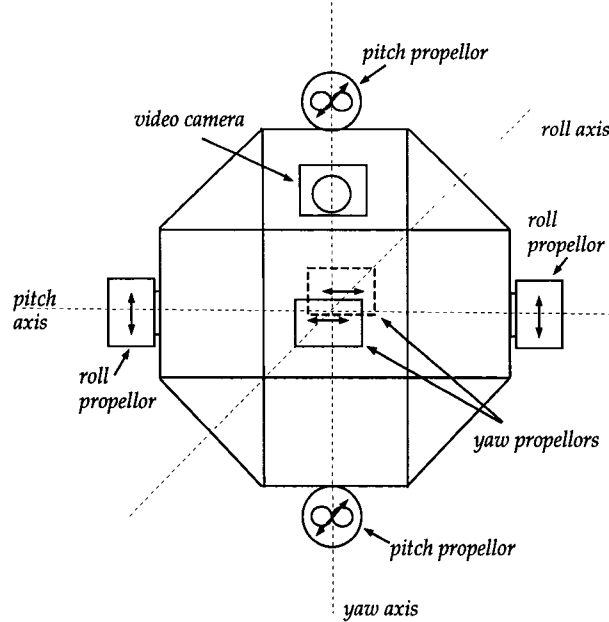


Figure 5.9: Front View of SCAMP.

5.4 Underwater Vehicle Experiment

In this section we describe an experimental implementation of Algorithm 1 on an underwater vehicle built and operated in the Space Systems Laboratory (SSL) at the University of Maryland. The SSL designs, builds and evaluates integrated telerobotic systems, including free-flying vehicles and modular manipulators for space operations such as space structure assembly and satellite servicing [2]. Testing is done in the SSL's Neutral Buoyancy Research Facility, a water tank 50 feet in diameter and 25 feet deep, located at the University of Maryland. The conditions in the tank are intended to simulate zero-gravity conditions of space.

SSL's Supplemental Camera and Maneuvering Platform (SCAMP), illustrated in Figure 5.9, moves freely and carries a video camera on board [3]. For space operations, it provides an external view for an operator performing a difficult task. In the SSL it is used to evaluate how a single person can operate more than

one vehicle at a time. SCAMP has the additional advantage of requiring little maintenance and being easy to prepare for underwater testing.

SCAMP is a 28-inch diameter icosahexahedron (26-sided) object weighing 167 pounds in air. About each of its three axes, SCAMP has a pair of ducted fan propellers, shown in Figure 5.9. A pair of propellers run in the same direction provides translation and run in opposing directions provides rotation about its associated axis. There is an on-board, closed-loop motor controller for each propeller that linearly converts an 8-bit (-128 to +127) command sent to the motor into a propeller speed. All on-board processing is done using a Motorola 68HC11 microcomputer. The Motorola 68HC11 communicates with the control station by means of a message-based serial protocol developed at the SSL. Data is transmitted over a fiber optic link. The control station consists of a MacIntosh IIfx computer and two hand controllers (joysticks). One hand controller is used for translation and the other for rotation. SCAMP also has a 7 lbs lead-weight pendulum that is located below the center of SCAMP along the yaw axis. The pendulum remains fixed with respect to the coordinate axes fixed on the vehicle unless it is actuated to control pitch. But for the pendulum, video camera, and the internal electronics, SCAMP is essentially a symmetric rigid body.

In the experiments described here, SCAMP was used to test the algorithms of Section 5.2 which provide open-loop controls to drive a vehicle without the aid of an operator. For these tests, the hand controllers were bypassed. Instead control signals were computed on the MacIntosh computer using Algorithm 1 and sent directly to the propeller motors. The main objective of the experiments was to show how Algorithm 1 can be used to control the attitude of SCAMP even in the event of an actuator failure. We assumed that the yaw propellers had failed.

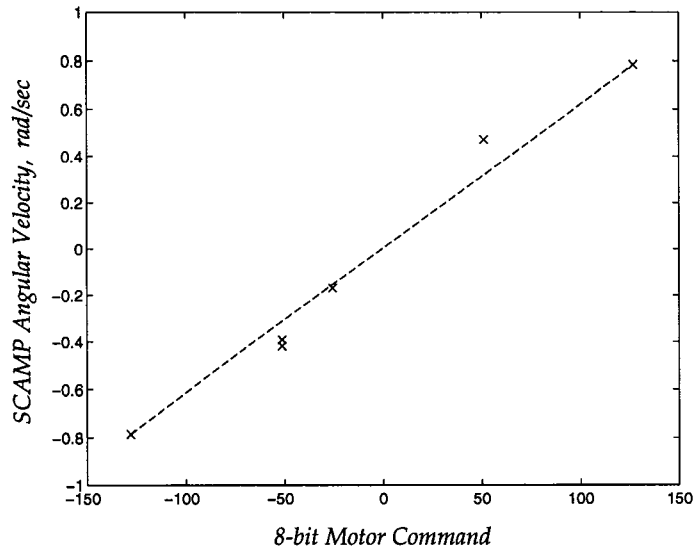


Figure 5.10: Calibration Data for SCAMP.

Then, we applied the open-loop controls from Algorithm 1 to actuate only the roll and pitch propellers to drive SCAMP with a net yaw rotation.

For our experiments we used the fact that due to the drag of the water, a constant propeller speed corresponds to a constant vehicle speed. As a result, a constant 8-bit command sent to the motor corresponds to a constant vehicle speed. Calibration data to determine the relationship between the 8-bit motor command units and vehicle angular velocity is plotted in Figure 5.10. The data for the 8-bit commands of +127 and -128 was provided from previous testing. The other data points were obtained at the time of the experiments described here. A constant 8-bit signal was sent to the yaw propellers for a period of 60 seconds. Angular velocity was measured by counting the number and direction of rotations during the 60 second period. For the purposes of the experiments we assumed that vehicle angular velocity is linearly related to the 8-bit motor command signal. Based on the data in Figure 5.10, we assumed a proportionality

constant of 163 motor command units per rad/sec.

We note that there is a time delay for the vehicle speed to follow the motor command related to the natural frequency of the vehicle. Because of the symmetry of the vehicle, we assumed that this delay is the same for roll, pitch and yaw. The open-loop controls generated by Algorithm 1 are sinusoids. Prescribed secular motion in depth-one bracket directions is a function of the phase difference between the sinusoids. For our experiments we used sinusoidal control signals with frequencies less than the natural frequency of the vehicle. Further, we counted on the fact that equal time delays in all of the sinusoidal control signals should not upset their phase relationships.

The kinematics of SCAMP can be described by equation (3.13) of Section 3.3 for an underwater vehicle. Since we consider only the attitude control problem for SCAMP, ignoring the translational control problem, the kinematics reduce to the system on $G = SO(3)$ defined by (3.2) of Section 3.1. Based on the above discussion we can let $\Omega_i = \epsilon u_i$, i.e., the angular velocities are our small-amplitude controls. When the yaw propellers are inactive, the motion control problem becomes

$$\dot{X} = \epsilon X(u_1 A_1 + u_2 A_2), \quad (5.30)$$

where $X(t) \in SO(3)$, $\{A_1, A_2, A_3\}$ is the standard basis for $so(3)$ and $\epsilon u_1(t)$ and $\epsilon u_2(t)$ are roll and pitch angular velocities (equivalently the 8-bit motor commands to the roll and pitch propellers), respectively.

System (5.30) is a depth-one bracket system on $SO(3)$ where $n = 3$, $m = 2$, identical to the spacecraft example with two controls described in Section 5.2. Thus, we use the controls of (5.23) specified by Algorithm 1 to control the attitude

of SCAMP. For a pure yaw rotation, i.e., $\gamma_{f1} = \gamma_{f2} = 0$, $\gamma_{f3} \neq 0$, these controls become

$$\begin{aligned} \epsilon u_1(t) &= \alpha_1 \omega \sin \omega t \quad 0 \leq t \leq t_3 \\ \epsilon u_2(t) &= \begin{cases} 0 & 0 \leq t \leq t_1 \\ \alpha_2 \omega \sin(\omega(t - t_1)) & t_1 \leq t \leq t_2 \\ 0 & t_2 \leq t \leq t_3 \end{cases} \end{aligned}$$

where $T = 2\pi/\omega$, $t_1 = T/4$, $t_2 = t_1 + MT$, $t_3 = t_2 + 3T/4$. α_1, α_2 should be selected such that

$$\alpha_1 = \sqrt{\left| \frac{\gamma_{f3}}{\pi M} \right|}, \quad \alpha_2 = \frac{\gamma_{f3}}{\alpha_1 \pi M}. \quad (5.31)$$

For the experiments we chose $\epsilon = 0.4$, $|\alpha_1| = 0.4$, $|\alpha_2| = 0.4$, $\omega = \pi/4$ rad/s and $M = 6.25$. From this we can compute $T = 8$ s, $t_1 = 2$ s, $t_2 = 52$ s, $t_3 = 58$ s. Further, from (5.31) $\gamma_{f3} = \alpha_1 \alpha_2 \pi M = \pm \pi$ rad. Thus, we expected to see a net yaw rotation of ± 180 degrees within $O(\epsilon^2)$ accuracy or within about ± 10 degrees accuracy. ϵ was chosen to be relatively large so that the oscillations of the vehicle would be large enough to observe and the test would not be too time consuming, i.e., $t_f = t_3$ relatively small. The fact that M was not chosen to be an integer implies that there should be a small net rotation about the roll and pitch axes at $t = t_f$. However, the passive effect of the pendulum was expected to quickly remove these rotations.

Several repetitions of this experiment were run in May and June 1994. Sensors were not available to take velocity or position measurements during the experiments. However, many of the experiments were recorded on videotape. During these experiments SCAMP was observed to make a net yaw rotation consistently as expected. Gentle oscillations about the roll and pitch axes were clearly visible

throughout the experiments with no significant final net roll or pitch rotation. When $\alpha_1\alpha_2 > 0$, SCAMP was observed to rotate about the yaw axis in the counter clockwise direction (looking from above). When $\alpha_1\alpha_2 < 0$, SCAMP was observed to rotate about the yaw axis in the clockwise direction (looking from above). The net yaw rotation was typically slightly less than 180 degrees but within the accuracy of the predicted motion.

For one series of the experiments with $\alpha_1\alpha_2 < 0$ (i.e., clockwise yaw motion) towards the end of the experimentation, the net yaw rotation was not quite as high as expected (closer to 90 degrees than 180 degrees). This may have been due to error in the approximation of the motor command to velocity conversion or to asymmetry in the roll and pitch propellers in this direction such that the phase difference between these signals was affected. Additionally, there was some difficulty keeping SCAMP neutrally buoyant throughout the test, but this did not seem to have an effect on the attitude motion.

In summary, attitude control of the underwater vehicle SCAMP with inactive yaw propellers was experimentally demonstrated using the open-loop controls derived from Algorithm 1. The experiments were useful in illustrating both the principle of the control design of this chapter as well as the potential for open-loop motion control based on the algorithms of Section 5.2. In particular, the experiment illustrated Lie bracket effects through sinusoidal controls (in this case, roll and pitch propeller thrust reversals). The experiment did not involve precision measurements; however, one could do so in future experimentation.

Chapter 6

Application to Bilinear Systems

In this chapter we consider the problem of control synthesis for bilinear systems on $\mathbb{R}^n$. The theory derived in this dissertation is relevant to this type of control problem because the state transition matrix for a bilinear system on $\mathbb{R}^n$ evolves on a subgroup of $GL(n)$, i.e., on a matrix Lie group G . Further, the state transition matrix equation is a right-invariant system on G (possibly with drift) of the form (2.7). Thus, we can express the control problem for a bilinear system on $\mathbb{R}^n$ equivalently as a motion control problem on G .

Our goal is to generate controls to solve the original control problem on $\mathbb{R}^n$ by solving the equivalent motion control problem on G , i.e., by driving the state transition matrix as necessary. We first examine the evolution of the state transition matrix for periodically time-varying controls by applying the theory of averaging on Lie groups derived in Chapter 4. Then, we consider generating controls based on the formulas for average solutions to the state transition matrix equation. If the matrix equation is not drift-free, as is the case in the example we study in this chapter, we cannot generate open-loop controls using the algorithms of Chapter 5. However, we can use the same basic idea of driving the average

matrix solution exactly in order to drive the actual solution approximately.

To illustrate the application of our theory to bilinear systems, we study a control problem for switched electrical networks. The switched networks of interest have bilinear state-space models in which each component of the control u , representing the position of a switch, takes value in the set $\{0, 1\}$ at any given time. Brockett and Wood [78, 17] have shown that a class of switched electrical networks such as those used in power conversion applications can be modelled as bilinear systems with state transition matrices evolving via right-invariant systems with drift on the Lie group $SO(k)$ or $SE(k)$. The control problem that we study is that of transferring energy between dynamic storage elements in a lossless network according to a prescribed path. We assume that we can control the switching in a periodic way with small (ϵ) period (which is equivalent to small (ϵ) amplitude in scaled time).

Switched electrical networks play an important role in a variety of systems, most notably in power converters such as dc-dc switchmode power converters. These power converters in turn are of great value in many growing application areas such as in communication and data handling systems, portable battery-operated equipment, and uninterruptable power sources (see [67]). New requirements for higher performance, smaller volume and lighter weight power converters make new demands on technology and warrant a fresh look at associated control strategies.

In standard practice, switching strategies (pulse-width-modulated controls) are derived based on state-space averaging methods. The averaging in this context can be interpreted as first-order (i.e., $O(\epsilon)$) averaging, meaning that the

solution to the averaged system approximates the actual system solution with $O(\epsilon)$ accuracy. Thus, increased accuracy is achieved with decreased ϵ which corresponds to a higher switching frequency. In our averaging approach we consider higher-order (i.e., $O(\epsilon^q)$, $q > 1$) averaging. This gives us the means to increase accuracy by increasing q rather than decreasing ϵ . Thus, we can improve performance without requiring excessively high switching frequencies. In their work on switched electrical networks, Brockett and Wood [17] derive high-order approximations of the state-space equations using a truncation of the Campbell-Baker-Hausdorff formula for the logarithm of the product of exponentials. Our work provides a justification via averaging theory on Lie groups for their formula.

Sira-Ramirez [70] has developed a very clear and simple controller design methodology for switched electrical networks which is based on variable structure systems theory and sliding regimes. In his work, control is achieved using feedback based on surprisingly simple laws (requiring as little as one bit of information). Our controller design methodology, on the other hand, focuses primarily on open-loop control, although feedback can be added for robustness. One advantage of open-loop control is that energy transfers in the networks can be accomplished in a predictable, finite number of switchings. This eliminates the possibility of chattering which is sometimes associated with sustaining motion on a sliding regime.

In Section 6.1 we define the general switched electrical network and describe our network example which is modelled as a bilinear system with a state transition matrix that evolves on $SO(3)$. In Section 6.2 we apply averaging theory and study controllability. In Section 6.3 we describe our controller design for the example. Simulation results are also presented.

6.1 Switched Electrical Network Problem and Approach

The (idealized) state-space bilinear equations for switched electrical networks generally take the form [78, 17]:

$$\dot{x} = (A_0 + A_1u_1 + A_2u_2 + \dots + A_mu_m)x + (b_1u_1 + b_2u_2 + \dots + b_mu_m). \quad (6.1)$$

The state $x \in \mathbb{R}^n$ represents inductor currents and capacitor voltages. Each control $u_i(t) \in \{0, 1\}$ represents the position of a switch at time t . Typically, the network will have only one switch or a set of switches that change position synchronously such that control action can be represented by a single switch position u , also taking values in the set $\{0, 1\}$. The vectors $b_1, \dots, b_m$ represent constant power sources. We can always express system (6.1) as a homogeneous equation using the fact that

$$\dot{x} = (Ax + b)u$$

is equivalent to

$$\frac{d}{dt} \begin{bmatrix} x \\ 1 \end{bmatrix} = \begin{bmatrix} A & b \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ 1 \end{bmatrix} u.$$

The state-space equations of interest thus become

$$\dot{x} = (A + Bu)x. \quad (6.2)$$

Alternatively, we can write (6.2) as

$$\dot{x} = (A(1 - u) + (A + B)u)x. \quad (6.3)$$

In changing the switch from the position corresponding to $u = 0$ to the position $u = 1$, the network topology changes. This is reflected in (6.3) as a

change between two different linear system descriptions, i.e., $\dot{x} = Ax$ when $u = 0$ and $\dot{x} = (A + B)x$ when $u = 1$. Thus, these are bilinear systems which always have a drift term, i.e., nontrivial dynamics when $u = 0$.

From linear systems theory it is well known that the state transition matrix $\Phi(t) \in GL(n)$ which describes the evolution of the state according to

$$x(t) = \Phi(t)x(0) \quad (6.4)$$

satisfies the same equation (6.2) as x , i.e.,

$$\dot{\Phi} = (A + Bu)\Phi, \quad \Phi(0) = I. \quad (6.5)$$

It is also well known [11] that $\Phi(t)$ will evolve on the matrix Lie group G associated with the Lie algebra generated by A and B (the connected, simply-connected covering group). Equation (6.5) describes a right-invariant system with drift on the Lie group G . If we define $X = \Phi^{-1} \in G$ then, using the identity

$$\frac{d}{dt}(K^{-1}(t)) = -K^{-1}(t)\dot{K}(t)K^{-1}(t)$$

for $K(t)$ an $n \times n$ matrix, we see that X satisfies

$$\dot{X} = X(-A - Bu). \quad (6.6)$$

Equations (6.6) describes a left-invariant system with drift on the Lie group G . Considering the alternative system description of (6.3) and defining $A_1 = A$, $A_2 = A + B$, $u_1 = -(1 - u)$ and $u_2 = -u$ we get

$$\dot{X} = X(A_1u_1 + A_2u_2), \quad (6.7)$$

which looks like system (2.11) without the ϵ factor. We have disguised the fact that the system has a drift term by defining two controls which, in fact, are

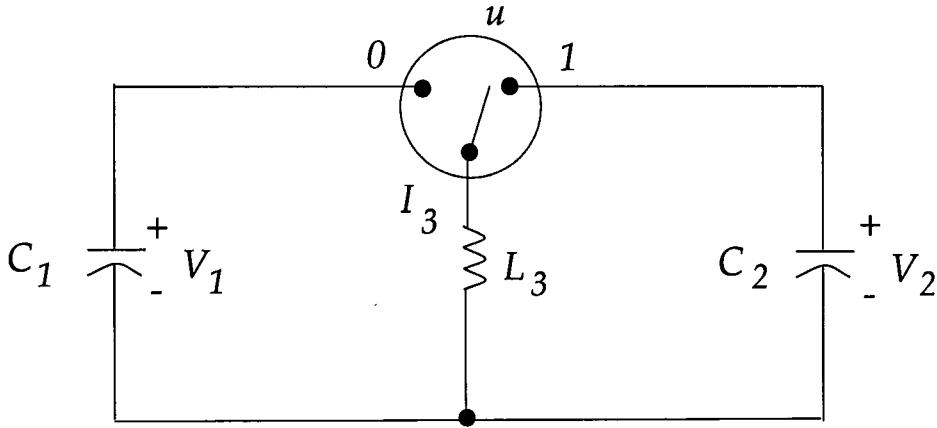


Figure 6.1: Network Example

not independent. However, by transforming our system into the form (6.7), we can apply our averaging theory on Lie groups to determine high-order average approximations to the behavior of X and thus Φ .

We address the problem of specifying switching controls that drive the state of the bilinear system (6.2) from some initial condition x_i along a desired path to a final condition x_f . We approach the switching problem by choosing k target points $x_1, \dots, x_k$ between x_i and x_f along the desired path. We then specify open-loop switching controls to drive the state between successive target points. Feedback can be added (possibly intermittently) for robustness. From equation (6.4), the problem of controlling the evolution of $x(t)$ can be considered as a problem of controlling the evolution of $\Phi(t)$ (or alternatively $X(t)$). As in Chapter 5, we determine the switching control that will drive an average approximation of $\Phi(t)$ exactly as desired in order to drive $\Phi(t)$ approximately as desired.

The network example that we study in this chapter is shown in Figure 6.1. This simple network can be considered as a model of the conversion portion of a dc-dc voltage converter [78]. The control problem is to convert energy from

one capacitor to the other via the inductor. For example, suppose that at $t = 0$, $V_1(0) = V_{10}$, $I_3(0) = 0$ and $V_2(0) = 0$. By appropriate switching, the energy in C_1 can be transferred to C_2 so that at some later time t' , $V_1(t') = 0$, $I_3(t') = 0$ and $V_2(t') = V_2'$. The ratio V_2'/V_{10} will depend on the ratio of the capacitances C_1/C_2 so that voltage scaling up or down can be achieved with an appropriate choice of capacitors. Of particular interest is controlling the energy transfers in this network to meet certain performance criteria. In Section 6.3 we design controls to transfer energy between capacitors while maintaining a constant nonzero current I_3 through the inductor.

The equations for this network are

$$\begin{aligned} C_1 \dot{V}_1 &= (1 - u)I_3 \\ C_2 \dot{V}_2 &= uI_3 \\ L_3 \dot{I}_3 &= -(1 - u)V_1 - uV_2. \end{aligned} \tag{6.8}$$

Define the state vector of the network $x = (x_1, x_2, x_3)^T$ by $x_1 = \sqrt{C_1}V_1$, $x_2 = \sqrt{C_2}V_2$, $x_3 = \sqrt{L_3}I_3$, and let $\omega_1 = 1/\sqrt{C_1 L_3}$ and $\omega_2 = 1/\sqrt{C_2 L_3}$. Then (6.8) can be rewritten as

$$\begin{aligned} \dot{x} &= (\omega_1 A_2 - (\omega_1 A_2 + \omega_2 A_1)u)x \\ &= (\omega_1 A_2(1 - u) - \omega_2 A_1 u)x \\ &= (-u_1 A_1 - u_2 A_2)x \end{aligned} \tag{6.9}$$

where $u_1 = \omega_2 u$, $u_2 = -\omega_1(1 - u)$ and $\{A_1, A_2, A_3\}$ is the standard basis for $so(3)$. Since $[A_1, A_2] = A_3$, the Lie algebra generated by A_1 and A_2 is $so(3)$. Thus, the state transition matrix $\Phi(t)$ evolves on the Lie group $SO(3)$ and the control problem for system (6.9) can equivalently be stated as a motion control problem

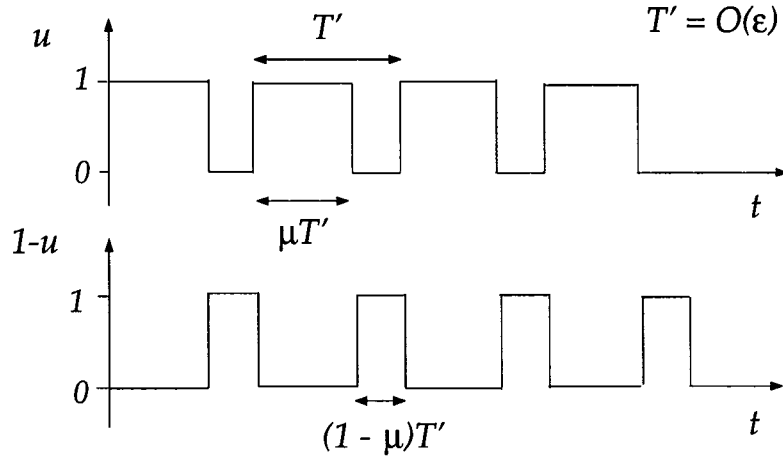


Figure 6.2: Switching Control

on $SO(3)$. Since an orthogonal matrix acting on a vector in $\mathbb{R}^3$ preserves length, $x(t)$ evolves on a sphere in $\mathbb{R}^3$. This confirms that the energy of the system, $\frac{1}{2}x^T x$, is conserved.

The network of Figure 6.1 is representative of a class of interesting networks. With an increased number of capacitors and/or inductors in the network, the state will evolve on a higher dimensional sphere and the transition matrix will evolve on a higher dimensional orthogonal group $SO(k)$. If power sources and/or loads are introduced into the network then the state will no longer evolve on a sphere since energy may be added to or dissipated from the network. In some of these cases it can be shown that the state transition matrix evolves on the Euclidean group $SE(k)$ [78].

6.2 Averaging and Controllability

The relevant controls for switched electrical networks are illustrated in Figure 6.2. These controls have nonzero mean and have $O(1)$ amplitude and $O(\epsilon)$ period

whereas the controls of system (2.9) studied in Chapter 4 have $O(\epsilon)$ amplitude and $O(1)$ period. By scaling time in the averaging theorems of Chapter 4, we can derive the appropriate average formulas for the network models. The following corollary is a time-scaled version of Corollary 4.14 and the first part of Corollary 4.16. It provides $O(\epsilon)$ and $O(\epsilon^2)$ average approximations of the solution to

$$\dot{g} = T_e L_g \cdot U(t), \quad U(t) = \sum_{i=1}^m u_i(t) \xi_i, \quad g(t) \in G, \quad U(t) \in \mathcal{G}, \quad (6.10)$$

where n is the dimension of G , $m \leq n$ and the controls u_i are periodic with $O(\epsilon)$ period and have nonzero mean. We use Γ_{ij}^k , u_{av} , $\tilde{u}$, U_{av} , $\tilde{U}$, $Area_{ij}(T)$ and $a_{ij}(t)$ as defined by (2.15), (2.19)-(2.23), respectively.

Corollary 6.1 Consider system (6.10). Assume $U(t)$ and b are as in Theorem 4.9 and $U(t + T') = U(t)$, $\forall t > 0$ where $T' = O(\epsilon)$. Let $g(0) = g_0 \in \hat{Q} \subset G$ and $Z_0 = \hat{\Psi}^{-1}(g_0) = O(\epsilon)$. Let $Z_0^{(q)} = \sum_{k=1}^n z_{k0}^{(q)} \xi_k$ and assume that $\|Z_0 - Z_0^{(q)}\| = O(\epsilon^q)$, $q = 1, 2$. Let

$$z_k^{(1)}(t) = u_{av} k t + z_{k0}^{(1)}, \quad Z^{(1)}(t) = \sum_{k=1}^n z_k^{(1)}(t) \xi_k, \quad g^{(1)}(t) = e^{Z^{(1)}(t)}. \quad (6.11)$$

If $Z^{(1)}(t) \in \hat{S}$, $\forall t \in [0, b]$, then

$$\hat{d}(g(t), g^{(1)}(t)) = O(\epsilon), \quad \forall t \in [0, b]. \quad (6.12)$$

Let

$$z_k^{[2]}(t) = \tilde{u}_k(t) + \sum_{i,j=1; i < j}^m a_{ij}(t) \Gamma_{ij}^k + z_{k0}^{(2)}, \quad Z^{[2]}(t) = \sum_{k=1}^n z_k^{[2]}(t) \xi_k, \quad g^{[2]}(t) = e^{Z^{[2]}(t)}. \quad (6.13)$$

For an integer N ,

$$z_k^{[2]}(NT') = NT' u_{av} k + N \sum_{i,j=1; i < j}^m Area_{ij}(T') \Gamma_{ij}^k + z_{k0}^{(2)}. \quad (6.14)$$

Γ_{ij}^k , $Area_{ij}(T')$, $a_{ij}(t)$ are defined by (2.15), (2.22), (2.23), respectively. If $Z^{[2]}(t) \in \hat{S}$, $\forall t \in [0, b]$, then

$$\hat{d}(g(t), g^{[2]}(t)) = O(\epsilon^2), \quad \forall t \in [0, b]. \quad (6.15)$$

Proof: Let $s = t/\epsilon$, then $ds = dt/\epsilon$ and we can write (6.10) as

$$\frac{dg}{ds}(\epsilon s) = \epsilon T_e L_{g(\epsilon s)} \cdot U(\epsilon s), \quad U(\epsilon s) = \sum_{i=1}^m u_i(\epsilon s) \xi_i. \quad (6.16)$$

Define

$$V(s) = U(\epsilon s), \quad v_i(s) = u_i(\epsilon s), \quad h(s) = g(\epsilon s).$$

Then (6.16) becomes

$$\frac{dh}{ds}(s) = \epsilon T_e L_{h(s)} \cdot V(s), \quad V(s) = \sum_{i=1}^m v_i(s) \xi_i.$$

Since $u(t + T') = u(t)$ then $v(s + T) = v(s)$ where $T = T'/\epsilon = O(1)$. Define

$$V_{av} = \frac{1}{T} \int_0^T V(\sigma) d\sigma.$$

Note that

$$\begin{aligned} V_{av} &= \frac{1}{T} \int_0^T V(\sigma) d\sigma \\ &= \frac{\epsilon}{T'} \int_0^{T'/\epsilon} V(\sigma) d\sigma \\ &= \frac{\epsilon}{T'} \int_0^{T'} V\left(\frac{\tau}{\epsilon}\right) \frac{d\tau}{\epsilon} \\ &= \frac{\epsilon}{T'} \int_0^{T'} U(\tau) \frac{d\tau}{\epsilon} \\ &= \frac{1}{T'} \int_0^{T'} U(\tau) d\tau \\ &= U_{av}. \end{aligned}$$

Let

$$Y^{(1)}(s) = \epsilon V_{av}s + Z_0^{(1)}, \quad h^{(1)}(s) = e^{Y^{(1)}(s)}.$$

By Corollary 4.14, if $Y^{(1)}(s) \in \hat{S}$, $\forall s \in [0, b/\epsilon]$ then

$$\hat{d}(h(s), h^{(1)}(s)) = O(\epsilon), \quad \forall s \in [0, b/\epsilon].$$

This implies that

$$\hat{d}(h(\frac{t}{\epsilon}), h^{(1)}(\frac{t}{\epsilon})) = O(\epsilon), \quad \forall t \in [0, b].$$

Recall that $g(t) = h(\frac{t}{\epsilon})$. Let

$$Z^{(1)}(t) \triangleq Y^{(1)}(\frac{t}{\epsilon}) = V_{av}t + Z_0^{(1)} = U_{av}t + Z_0^{(1)},$$

$$g^{(1)}(t) \triangleq h^{(1)}(\frac{t}{\epsilon}) = e^{Y^{(1)}(t/\epsilon)} = e^{Z^{(1)}(t)}.$$

Thus, if $Z^{(1)}(t) \in \hat{S}$, $\forall t \in [0, b]$ then

$$\hat{d}(g(t), g^{(1)}(t)) = O(\epsilon), \quad \forall t \in [0, b],$$

proving the first part of the corollary.

For the second part of the corollary define

$$\tilde{v}_k(s) = \int_0^s v_k(\sigma) d\sigma,$$

$$b_{ij}(s) = \frac{1}{2} \int_0^s (\tilde{v}_i(\sigma) v_j(\sigma) - \tilde{v}_j(\sigma) v_i(\sigma)) d\sigma.$$

Note that

$$\begin{aligned} \tilde{v}_k\left(\frac{t}{\epsilon}\right) &= \int_0^{t/\epsilon} v_k(\sigma) d\sigma \\ &= \int_0^t v_k\left(\frac{\tau}{\epsilon}\right) \frac{d\tau}{\epsilon} \\ &= \frac{1}{\epsilon} \int_0^t u_k(\tau) d\tau \\ &= \frac{\tilde{u}_k(t)}{\epsilon}, \end{aligned}$$

and

$$\begin{aligned}
b_{ij} \left(\frac{t}{\epsilon} \right) &= \frac{1}{2} \int_0^{t/\epsilon} (\tilde{v}_i(\sigma) v_j(\sigma) - \tilde{v}_j(\sigma) v_i(\sigma)) d\sigma \\
&= \frac{1}{2} \int_0^t \left(\tilde{v}_i \left(\frac{\tau}{\epsilon} \right) v_j \left(\frac{\tau}{\epsilon} \right) - \tilde{v}_j \left(\frac{\tau}{\epsilon} \right) v_i \left(\frac{\tau}{\epsilon} \right) \right) \frac{d\tau}{\epsilon} \\
&= \frac{1}{2} \int_0^t \left(\frac{\tilde{u}_i(\tau)}{\epsilon} u_j(\tau) - \frac{\tilde{u}_j(\tau)}{\epsilon} u_i(\tau) \right) \frac{d\tau}{\epsilon} \\
&= \frac{1}{2\epsilon^2} \int_0^t (\tilde{u}_i(\tau) u_j(\tau) - \tilde{u}_j(\tau) u_i(\tau)) d\tau \\
&= \frac{1}{\epsilon^2} a_{ij}(t).
\end{aligned}$$

Let

$$\begin{aligned}
y_k^{[2]}(s) &= \epsilon \tilde{v}_k(s) + \epsilon^2 \sum_{i,j=1; i < j}^m b_{ij}(s) \Gamma_{ij}^k + z_{k0}^{(2)}, \\
Y^{[2]}(s) &= \sum_{k=1}^n y_k^{[2]}(s) \xi_k, \quad h^{[2]}(s) = e^{Y^{[2]}(s)}.
\end{aligned}$$

By Corollary 4.16, if $Y^{[2]}(s) \in \hat{S}$, $\forall s \in [0, b/\epsilon]$, then

$$\hat{d}(h(s), h^{[2]}(s)) = O(\epsilon^2), \quad \forall s \in [0, b/\epsilon].$$

This implies that

$$\hat{d}(h(\frac{t}{\epsilon}), h^{[2]}(\frac{t}{\epsilon})) = O(\epsilon^2), \quad \forall t \in [0, b].$$

Let

$$\begin{aligned}
z_k^{[2]}(t) &\triangleq y_k^{[2]}(\frac{t}{\epsilon}) = \epsilon \tilde{v}_k(\frac{t}{\epsilon}) + \epsilon^2 \sum_{i,j=1; i < j}^m b_{ij}(\frac{t}{\epsilon}) \Gamma_{ij}^k + z_{k0}^{(2)} \\
&= \tilde{u}_k(t) + \sum_{i,j=1; i < j}^m a_{ij}(t) \Gamma_{ij}^k + z_{k0}^{(2)}. \\
Z^{[2]}(t) &\triangleq Y^{[2]}(\frac{t}{\epsilon}) = \sum_{k=1}^n y_k^{[2]}(\frac{t}{\epsilon}) \xi_k = \sum_{k=1}^n z_k^{[2]}(t) \xi_k. \\
g^{[2]}(t) &\triangleq h^{[2]}(\frac{t}{\epsilon}) = e^{Y^{[2]}(t/\epsilon)} = e^{Z^{[2]}(t)}.
\end{aligned}$$

Thus, if $Z^{[2]}(t) \in \hat{S}$, $\forall t \in [0, b]$, then

$$\hat{d}(g(t), g^{[2]}(t)) = O(\epsilon^2), \quad \forall t \in [0, b].$$

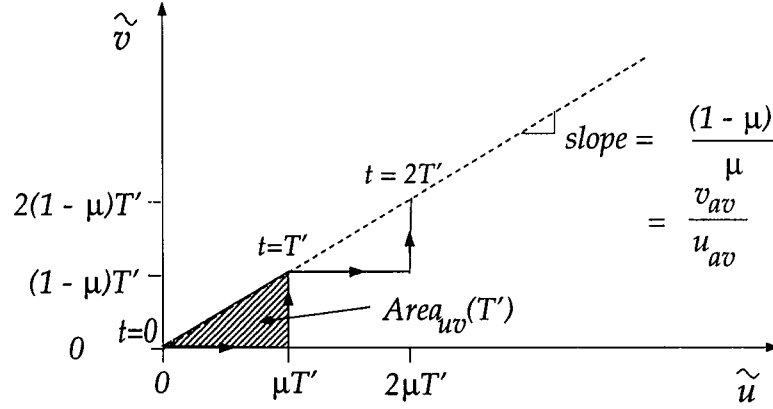


Figure 6.3: Geometric Interpretation of $Area_{uv}(T')$

Following the proof of Corollary 4.16 for N an integer, $\tilde{u}_k(NT') = NT' u_{avk}$ and $a_{ij}(NT') = N Area_{ij}(T')$. So

$$z_k^{[2]}(NT') = NT' u_{avk} + N \sum_{i,j=1; i < j}^m Area_{ij}(T') \Gamma_{ij}^k + z_{k0}^{(2)}. \quad \square$$

It is straightforward to compute the relevant terms in the average formulas of Corollary 6.1 for our controls in Figure 6.2. The variable μ , where $0 \leq \mu \leq 1$, referred to as the *duty ratio*, is defined as the fraction of the period T' for which $u = 1$. If we only concern ourselves with moments of time t when $t = MT'$, M an integer, then it is irrelevant when during a period $u = 1$, i.e., shifts in time have no effect on our results. Letting $v = 1 - u$ we find using definitions of u_{av} (2.19) and $Area_{ij}(T')$ (2.22) that

$$u_{av} = \mu, \quad v_{av} = 1 - \mu, \quad Area_{uv}(T') = \frac{1}{2} T'^2 \mu(1 - \mu). \quad (6.17)$$

Figure 6.3 illustrates the geometric interpretation of the area term $Area_{uv}(T')$.

For the example network of Figure 6.1, we have by (6.9) that

$$\begin{aligned} \dot{\Phi} &= (-u_1 A_1 - u_2 A_2) \Phi, & \Phi(0) &= I \\ \dot{X} &= X(u_1 A_1 + u_2 A_2), & X(0) &= I \end{aligned} \quad (6.18)$$

where $A_1 = \hat{e}_1$, $A_2 = \hat{e}_2$, $u_1 = \omega_2 u$ and $u_2 = -\omega_1(1 - u)$. Thus, from (6.17)

$$u_{av1} = \omega_2 \mu, \quad u_{av2} = -\omega_1(1 - \mu), \quad Area_{12}(T') = -\frac{1}{2}\omega_1\omega_2 T'^2 \mu(1 - \mu).$$

Therefore, from Corollary 6.1, since the relevant structure constant is $\Gamma_{12}^3 = 1$ ($[A_1, A_2] = A_3$) the average state transition matrix solutions are

$$\Phi^{(1)}(MT') = (X^{(1)})^{-1}(MT') = e^{\sum_{k=1}^n -z_k^{(1)}(MT')\xi_k}, \quad (6.19)$$

$$z_1^{(1)}(MT') = \omega_2 MT' \mu, \quad z_2^{(1)}(MT') = -\omega_1 MT'(1 - \mu), \quad z_3^{(1)}(MT') = 0. \quad (6.20)$$

$$\Phi^{[2]}(MT') = (X^{[2]})^{-1}(MT') = e^{\sum_{k=1}^n -z_k^{[2]}(MT')\xi_k}, \quad (6.21)$$

$$\begin{aligned} z_1^{[2]}(MT') &= \omega_2 MT' \mu, & z_2^{[2]}(MT') &= -\omega_1 MT'(1 - \mu), \\ z_3^{[2]}(MT') &= -\frac{1}{2}\omega_1\omega_2 MT'^2 \mu(1 - \mu). \end{aligned} \quad (6.22)$$

Towards the goal of controlling the state transition matrix $\Phi(t)$, we examine the controllability of the system

$$\dot{X} = X(A + Bu) \quad (6.23)$$

on the matrix Lie group G with Lie algebra $\mathcal{G}$. In order to accommodate switching controls, as shown in Figure 6.2, we define the class of controls $\mathcal{U}_s$ as

- $\mathcal{U}_s$ is the class of piecewise constant functions on $[0, \infty)$ taking values in $\mathbb{R}^m$ where components of its elements take values in the set $\{0, 1\}$.

We have the following corollary to Theorem 2.2 of Jurdjevic and Sussmann.

Corollary 6.2 Let G be a compact and connected Lie group. System (6.23) is controllable with $u \in \mathcal{U}_s$ if and only if the Lie algebra L generated by $\{A, B\}$ is equal to $\mathcal{G}$. Further, $\exists t' > 0$ such that for every $X_0, X_1 \in G$ there is a control

$u \in \mathcal{U}_s$ that steers X_0 to X_1 in less than t' units of time. If, further, G is semi-simple, then $\exists t' > 0$ such that for every $X_0, X_1 \in G$ there is a control $u \in \mathcal{U}_s$ that steers X_0 to X_1 in exactly t' units of time.

Proof: Let $w = 2u - 1$. Then $w = 1$ if $u = 1$ and $w = -1$ if $u = 0$. Thus, if $u \in \mathcal{U}_s$ then $w \in \mathcal{U}_b$ ($\mathcal{U}_b$ defined in Section 2.2.1). Since $u = \frac{1}{2}(w + 1)$, system (6.23) can be rewritten as

$$\dot{X} = X(A + B(\frac{1}{2}(w + 1))) = X((A + \frac{1}{2}B) + \frac{1}{2}Bw). \quad (6.24)$$

By Theorem 2.2 (Jurdjevic and Sussmann) system (6.24) is controllable with $w \in \mathcal{U}_b$ if and only if the Lie algebra L_s generated by $\{(A + \frac{1}{2}B), \frac{1}{2}B\}$ is equal to $\mathcal{G}$. Since L_s is L , system (6.24) with $w \in \mathcal{U}_b$ is controllable if and only if $L = \mathcal{G}$. But, system (6.24) with $w \in \mathcal{U}_b$ is equivalent to system (6.23) with $u \in \mathcal{U}_s$. The remaining results follow from Theorem 2.2. $\square$

For our example network (Figure 6.1), the state transition matrix, which satisfies equation (6.18), evolves on the Lie group $G = SO(3)$ which is connected, compact and semi-simple. The Lie algebra L generated by $\{A_1, A_2\}$ is $so(3)$. Thus, by Corollary 6.2 the system described by (6.18) is controllable. Further, there exists a time $t' > 0$ such that for every $\Phi_0, \Phi_1 \in SO(3)$, there is a control $u \in \mathcal{U}_s$ that steers Φ_0 into Φ_1 in exactly t' units of time.

It is important to note that the system (6.18) is not controllable in arbitrarily short periods of time. In fact, steering from some $\Phi_0 \in SO(3)$ to some other arbitrary $\Phi_1 \in SO(3)$ may take a considerably long time. Further, Φ may have to travel far to reach Φ_1 . This implies that we may not be able to choose switching controls to follow any arbitrary path in the system state space.

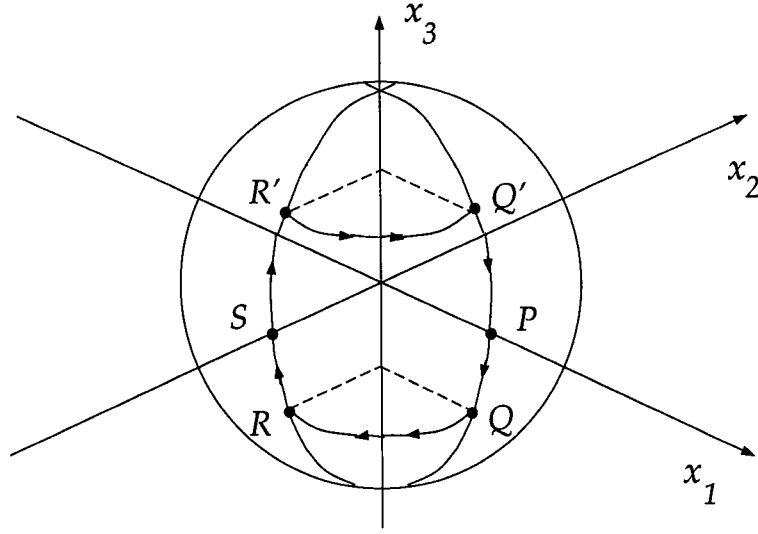


Figure 6.4: Paths in State Space

A qualitative picture of the paths that we can follow closely can be deduced from the $O(\epsilon)$ average formula $\Phi^{(1)}$ for Φ given in equations (6.19) and (6.20). In Figure 6.4 we show the state space for our system, i.e., a sphere in $\mathbb{R}^3$. A curve $\Phi(t)$ corresponds to a rotation of the sphere. While the $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$ axes stay fixed, the system state $x(t)$ rotates with the sphere. Equation (6.19) can be interpreted as an Euler parametrization of the rotation $\Phi^{(1)}$ which approximates the rotation Φ . Let $z^{(1)} = (z_1^{(1)}, z_2^{(1)}, z_3^{(1)})^T$ and note that $z^{(1)}(MT') = Mz^{(1)}(T')$. Let $\phi = \|z^{(1)}(T')\|$. The magnitude of the rotation $\Phi^{(1)}$ is $M\phi$ and the axis of rotation is $-z^{(1)}(T')/\phi$. Since μ is defined such that it satisfies $0 \leq \mu \leq 1$, equations (6.19) and (6.20) imply that $-z_1^{(1)}(T') \leq 0$, $-z_2^{(1)}(T') \geq 0$ and $-z_3^{(1)}(T') = 0$. Thus, the axis of rotation of the $O(\epsilon)$ approximation will point into the quadrant of the $\mathbf{x}_1$ - $\mathbf{x}_2$ plane corresponding to $x_1 \leq 0$, $x_2 \geq 0$. For example, if $\mu = 0$ then $\Phi^{(1)}$ corresponds to a positive rotation about the $\mathbf{x}_2$ -axis. If $\mu = 1$ then $\Phi^{(1)}$ corresponds to a negative rotation about the $\mathbf{x}_1$ -axis.

In general, this type of restriction means that a path that can be followed closely in one direction cannot be followed closely in the opposite direction. In the next section we will illustrate the design of a controller that transfers energy in the example network from C_1 to C_2 following the path $PQRS$ shown in Figure 6.4. The path from P to Q can be achieved using $u = 0$ (i.e., $\mu = 0$). Similarly, the path from R to S can be achieved using $u = 1$ (i.e., $\mu = 1$). The path from Q to R can be followed by successive rotations about different vectors pointing into the second quadrant of the $\mathbf{x}_1$ - $\mathbf{x}_2$ plane. On the other hand, the reverse path $SRQP$ cannot be followed closely due to the limitation on the direction of achievable rotations. Similarly, the path $SR'Q'P$ is achievable while $PQ'R'S$ is not.

6.3 Controller Design

In order to design a controller for the example network of Figure 6.1, we first develop a technique for computing the control switchings that will drive the state $x(t)$ from some initial target point x_a to the next target point x_b . We assume that the path from x_a to x_b is in the “achievable” direction as discussed in Section 6.2. Let the switching period T' be fixed at $T' = \epsilon$. Then the problem becomes one of choosing μ (the duty ratio) and M (the number of switching periods) such that for some initial time t_0 , $x(t_0) = x_a$ and $x(t_0 + MT') = x_b$.

Using our second-order average formula $\Phi^{[2]}$ ((6.21) and (6.22)), we know the state transition matrix solution as a function of μ and M with $O(\epsilon^2)$ accuracy. To determine the desired motion of the state transition matrix, we parametrize the various rotations that will take x_a into x_b as a function of a single parameter.

One rotation that will take x_a into x_b can be computed as

$$\Phi_c \triangleq e^{\theta \hat{x}_c} \text{ where } \theta = \cos^{-1}(x_a^T x_b), \quad x_c = \frac{x_a \times x_b}{\|x_a \times x_b\|} = \frac{\hat{x}_a x_b}{\|\hat{x}_a x_b\|}. \quad (6.25)$$

After this rotation has been performed, the state is invariant to rotations about the vector x_b . A class of rotations Φ_{θ_b} that will take x_a into x_b is described by

$$\Phi_{\theta_b} \triangleq e^{\theta_b \hat{x}_b} e^{\theta \hat{x}_c}, \quad (6.26)$$

where θ_b is a free parameter. Our goal is then to compute μ , M and θ_b such that $\Phi^{[2]}(MT') = \Phi_{\theta_b}$. We describe two methods. The first method is based on quaternions and it produces a set of nonlinear algebraic equations to be solved for the exact values of μ , M and θ_b . The second method uses some approximations to produce a linear set of equations to be solved for μ , M and θ_b .

Method 1 (quaternions). A unit quaternion is a four-tuple that can be used as a representation of a rotation in $\mathbb{R}^3$. Quaternions are easily multiplied making the computation of a product of rotations as in (6.26) straightforward. Let $Q_c = (q_{c0}, q_{c1}, q_{c2}, q_{c3})$, $Q_b = (q_{b0}, q_{b1}, q_{b2}, q_{b3})$, $Q_t = (q_{t0}, q_{t1}, q_{t2}, q_{t3})$, $Q_z = (q_{z0}, q_{z1}, q_{z2}, q_{z3})$ be the quaternion representations of Φ_c , $\Phi_b = e^{\theta_b \hat{x}_b}$, $\Phi_{\theta_b} = \Phi_b \Phi_c$, $\Phi^{[2]}(MT')$, respectively. Then, by definition

$$\begin{aligned} Q_c &= (\cos \frac{\theta}{2}, x_{c1} \sin \frac{\theta}{2}, x_{c2} \sin \frac{\theta}{2}, x_{c3} \sin \frac{\theta}{2}), \\ Q_b &= (\cos \frac{\theta_b}{2}, x_{b1} \sin \frac{\theta_b}{2}, x_{b2} \sin \frac{\theta_b}{2}, x_{b3} \sin \frac{\theta_b}{2}), \\ Q_z &= (\cos \frac{\phi}{2}, -\frac{z_1^{[2]}(MT')}{\phi} \sin \frac{\phi}{2}, -\frac{z_2^{[2]}(MT')}{\phi} \sin \frac{\phi}{2}, -\frac{z_3^{[2]}(MT')}{\phi} \sin \frac{\phi}{2}), \end{aligned} \quad (6.27)$$

where $\phi = \|z^{[2]}(MT')\|$. By the rules of quaternion multiplication [10], $Q_t = Q_b Q_c$

such that

$$Q_t = \begin{pmatrix} q_{t_0} \\ q_{t_1} \\ q_{t_2} \\ q_{t_3} \end{pmatrix} = \begin{pmatrix} q_{b_0}q_{c_0} - q_{b_1}q_{c_1} - q_{b_2}q_{c_2} - q_{b_3}q_{c_3} \\ q_{b_0}q_{c_1} + q_{b_1}q_{c_0} + q_{b_2}q_{c_3} - q_{b_3}q_{c_2} \\ q_{b_0}q_{c_2} + q_{b_2}q_{c_0} + q_{b_3}q_{c_1} - q_{b_1}q_{c_3} \\ q_{b_0}q_{c_3} + q_{b_3}q_{c_0} + q_{b_1}q_{c_2} - q_{b_2}q_{c_1} \end{pmatrix}.$$

To find μ , M and θ_b such that $\Phi^{[2]}(MT') = \Phi_{\theta_b}$, we solve $Q_t = Q_z$. Because we consider unit quaternions, we have three independent equations and three unknowns.

Method 2 (linear approximation). For θ_b and θ small, by the Campbell-Baker-Hausdorff formula [54],

$$\begin{aligned} \Phi_{\theta_b} &= e^{\theta_b \hat{x}_b} e^{\theta \hat{x}_c} = e^{\hat{w}}, \\ \hat{w} &= \theta_b \hat{x}_b + \theta \hat{x}_c + \theta_b \theta [\hat{x}_b, \hat{x}_c] + \frac{1}{12} \theta_b \theta^2 [[\hat{x}_b, \hat{x}_c], \hat{x}_c] + \frac{1}{12} \theta_b^2 \theta [[\hat{x}_c, \hat{x}_b], \hat{x}_b] + \dots \end{aligned}$$

We make an approximation of w linear in θ_b by ignoring the bracket terms of higher-order than single brackets. We use the identity

$$[\hat{a}, \hat{b}] = \hat{a}\hat{b} - \hat{b}\hat{a} = \widehat{a \times b} = \widehat{\hat{a}\hat{b}},$$

and let $x_d = \hat{x}_b \hat{x}_c$. Then we find that

$$w \approx \theta_b x_b + \theta x_c + \theta_b \theta x_d = \theta_b (x_b + \theta x_d) + \theta x_c.$$

Now let $r = M\mu$. From (6.21) and (6.22), $\Phi^{[2]}(MT') \approx e^{\hat{y}}$ where

$$y = \begin{pmatrix} -\omega_2 \epsilon \\ -\omega_1 \epsilon \\ \frac{1}{2} \omega_1 \omega_2 \epsilon^2 \end{pmatrix} r + \begin{pmatrix} 0 \\ \omega_1 \epsilon \\ 0 \end{pmatrix} M.$$

To find μ , M and θ_b we solve $w = y$. This yields three linear equations in the three unknowns r , M and θ_b :

$$\begin{pmatrix} \omega_2\epsilon & 0 & (x_{b_1} + \theta x_{d_1}) \\ \omega_1\epsilon & -\omega_1\epsilon & (x_{b_2} + \theta x_{d_2}) \\ -\frac{1}{2}\omega_1\omega_2\epsilon^2 & 0 & (x_{b_3} + \theta x_{d_3}) \end{pmatrix} \begin{pmatrix} r \\ M \\ \theta_b \end{pmatrix} = \begin{pmatrix} -\theta x_{c_1} \\ -\theta x_{c_2} \\ -\theta x_{c_3} \end{pmatrix}.$$

We can then solve for $\mu = r/M$.

We can, of course, improve the accuracy in the computation of M , μ and θ_b for Method 2 by including nonlinear terms in the expressions for w and y . The linear approximation, however, is useful because it reduces computational complexity. This would be an advantage if the computation were to be done on-line for feedback.

Example Problem. We illustrate a controller design example based on Method 1. Suppose that it is desired to transfer energy from C_1 to C_2 while maintaining a constant current I_3 through the inductor. That is, suppose we wish to drive the state $x(t)$ along the path $PQRS$ as shown in Figure 6.4. Let $x^T x = 1$, i.e., assume that $x(t)$ evolves on the unit sphere. Then we want $x(0) = P = (1, 0, 0)^T$ and $x(t_f) = S = (0, -1, 0)^T$. Let us, for example, choose the constant current such that $x_3 = -1/\sqrt{2}$. Then we want $x(s_0) = Q = (1/\sqrt{2}, 0, -1/\sqrt{2})^T$ and $x(s_f) = R = (0, -1/\sqrt{2}, -1/\sqrt{2})^T$. We choose six target points $x_1, \dots, x_6$ along the path from $x_0 = Q$ to $x_7 = R$ as

$$x_i = e^{-\frac{1}{7} \frac{\pi}{2} A_3} x_{i-1}, \quad i = 1, \dots, 6.$$

We assume the values for the network components are

$$C_1 = 0.1, \quad C_2 = 0.2, \quad L_3 = 0.5.$$

We note that the path from P to Q is a rotation of $\cos^{-1}(1/\sqrt{2}) = \pi/4$ radians about the $\mathbf{x}_2$ -axis. We can follow this path by setting $u = \mu = 0$ for $t \in [0, s_0]$ where s_0 is computed from

$$\omega_1 s_0(1 - \mu) = \omega_1 s_0 = \pi/4.$$

Similarly, the path from R to S is a rotation of $-\pi/4$ radians about the $\mathbf{x}_1$ -axis. This path can be followed by setting $u = \mu = 1$ for $t \in [s_f, t_f]$ where $t_f - s_f$ is computed from

$$-\omega_2(t_f - s_f)\mu = -\omega_2(t_f - s_f) = -\pi/4.$$

To compute the switching control for the path QR , we apply Method 1 to each of the seven steps to be taken between Q and R . Let μ_{ij} and M_{ij} denote the values of μ and M , respectively, to be used to drive $x(t)$ from x_i to x_j . Using $\epsilon = 0.01$, Method 1 (solved numerically using MATLAB) produces

ij	01	12	23	34	45	56	67
μ_{ij}	0.9262	0.8016	0.6924	0.5858	0.4705	0.3310	0.1374
M_{ij}	7.5781	8.3375	8.6737	8.5677	8.0238	7.0701	5.7569

Thus, $s_f = s_0 + \sum_{i=0}^6 M_{i(i+1)}\epsilon = s_0 + 54.0078\epsilon$. We note that the values of M_{ij} are not integers. We account for that by using a fractional period (i.e., a scaled value of T') for the last switching of each step.

The control u computed for the complete path $PQRS$ is shown in Figure 6.5(b). This figure shows that there are a finite number of control switchings. The duty ratio is close to 1 towards the beginning and close to 0 towards the end. Figure 6.5(a) shows the response of the state $x(t)$ as a function of time. The desired energy transfer can be observed, i.e., x_1 goes from 1 to 0 and x_2 goes from 0 to -1. x_3 ramps to $-1/\sqrt{2}$, stays there and then ramps back to 0 as desired.

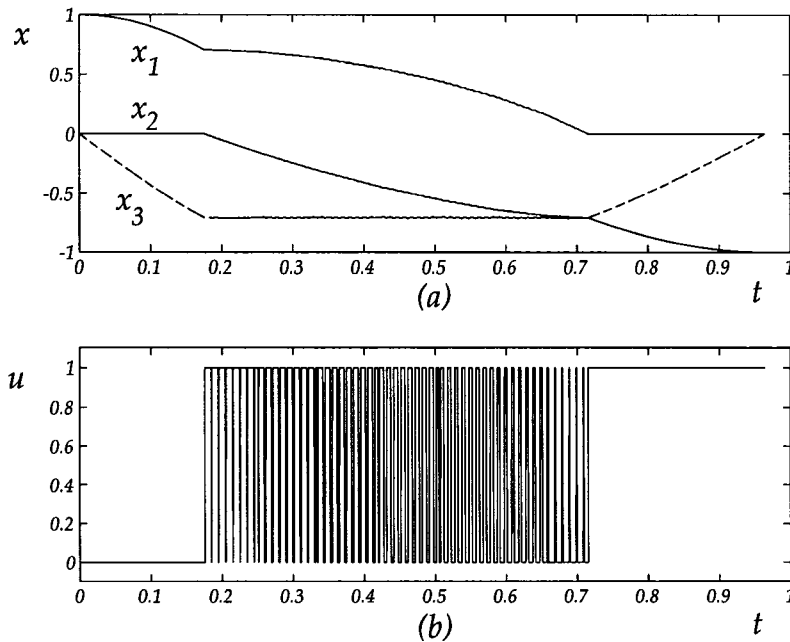


Figure 6.5: Control Signal and State Response

The trajectory of $x(t)$ in the state space is illustrated in Figure 6.6. The simulation was performed using MATLAB. We show a magnified plot of $x_3(t)$ during the path QR in Figure 6.7. The root-mean-square error in $x_3(t)$ during the path QR when it was intended to be constant at $x_3(t) = -1/\sqrt{2}$ was computed to be $0.0037 = 37\epsilon^2$. This is consistent with the $O(\epsilon^2)$ accuracy that we expect.

Increases in accuracy could be achieved either by decreasing ϵ (increasing switching frequency) or by considering higher-order terms in the average approximation of $\Phi(t)$. Feedback could be introduced for robustness by computing μ and M on line after each or every few target points are reached. As explained above, Method 2 might be more suitable for feedback since it requires solving a linear rather than a nonlinear set of algebraic equations.

The network that is the topological dual to this example is also described by

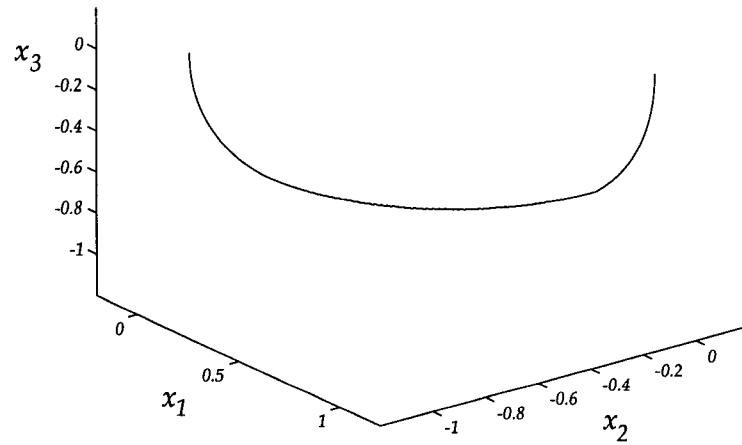


Figure 6.6: State-Space Trajectory

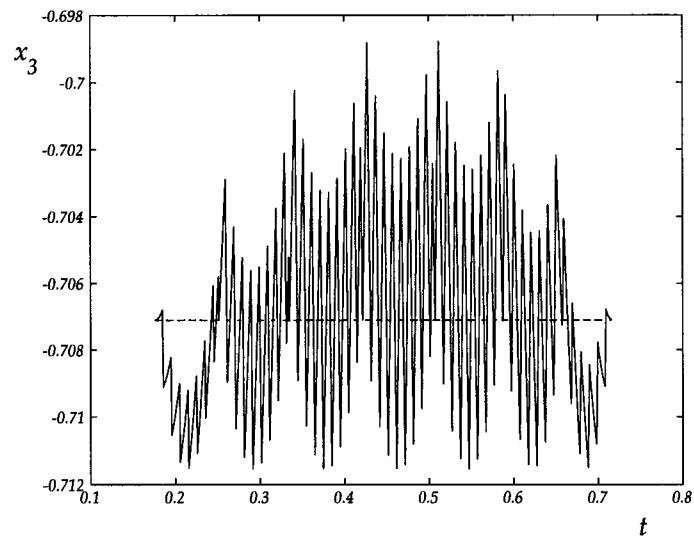


Figure 6.7: Response of Current Variable x_3

a state transition matrix that evolves on $SO(3)$ [70]. Accordingly, the methodology described here can be used to control energy transfers in the dual network. Similarly, it is expected that one could extend the methodology to other more complicated networks such as those with a state transition matrix evolving on $SO(k)$, $k > 3$. The major difficulty with such an extension would be the parametrization of the rotations in $SO(k)$ that drive the state from one point to another on the sphere in $\Re^k$.

Chapter 7

Conclusions and Future Research

Motion control problems were considered in this dissertation in the framework of systems on finite-dimensional Lie groups. One goal of this study was to understand the behavior, or motion, of these systems given small-amplitude, periodically time-varying control inputs. A second goal was to address the problem of control synthesis. Averaging theory proved to be a very useful tool in this investigation. Our success in deriving simplified, geometric formulas for average motion and the generalized nature of our methodology for control synthesis provides evidence for our initial assertion that Lie groups provide both a natural and a rich setting for studying a variety of motion control problems. Some of the contributions of this dissertation are outlined below.

In Chapter 2 we described the abstract framework of motion control on Lie groups. We cited two important local representations of the solution to a left-invariant system on a finite-dimensional Lie group for use in our averaging theory on Lie groups. We extended the applicability of the product of exponentials representation (Wei and Norman) from systems on matrix Lie groups to systems on abstract finite-dimensional Lie groups. We also applied a result of Lazard and

Tits on the injectivity of the exponential map to show that the single exponential representation (Magnus) of solutions on certain Lie groups have reasonably large domains.

In Chapter 3 we showed how to express several example motion control problems as drift-free, left-invariant control systems on matrix Lie groups. The examples are all kinematic control problems for mechanical systems in that the Lie group in each example is the system configuration space, the differential equation on the Lie group describes the kinematics and the system velocity was taken to be the control input. For the spacecraft attitude control problem, we described two different physical arrangements, one using internal rotors and one using an appended point-mass oscillator, for which the motion control problem is a kinematic control problem. Using the example of a ball rolling between two rough, parallel plates, we illustrated how the methodology of this dissertation might be applied to the design of a piezo-actuated vibratory motor. We also showed that, contrary to what is typically done in the nonholonomic motion planning literature, mobile robots such as the unicycle and the front-wheel drive car are naturally modelled as systems on Lie groups. Thus, the methodology of this dissertation is applicable to these systems.

In Chapter 4 we derived averaging theory for left-invariant systems (with or without drift) on finite-dimensional Lie groups. This is one of the most important contributions of this dissertation. Our theory provides explicit formulas for approximating the system solution given small (ϵ) amplitude, periodically time-varying (oscillatory) control inputs to arbitrarily high order in ϵ . We showed that these formulas could be interpreted geometrically as functions of ϵ , time, area and moment-like terms and the structure constants associated to a given basis for the

relevant Lie algebra. The area and moment terms depend only on the path of the control inputs during one period of the oscillation and not on the frequency of the oscillation. We emphasized the intrinsic nature of these formulas by showing that the second-order term in the average formulas derives from the curvature form of a certain principal fiber bundle with connection. We also showed how the results of averaging of systems on Lie groups could be used to derive controls for stabilization of a system equilibrium point.

In Chapter 5 we developed a methodology for open-loop control synthesis to achieve point-to-point system maneuvers. This is the second main contribution of this dissertation. The methodology rests on the idea that by controlling an average solution exactly we will control the system motion approximately. We showed how controllability of the average solution depends on the relationship between the order of the average solution and the depth of the Lie brackets used in the Lie algebra controllability rank condition for the original system. We showed how to generate open-loop controls by exploiting the geometric interpretation of the average solution formulas. Explicit algorithms were derived and illustrated for the examples of Chapter 3. The algorithms depend fundamentally on the structure of the system as described by the structure constants for the basis of $\mathcal{G}$ associated with the system's control authority. If the control authority changes then the algorithm produces different control laws. We showed how this feature of our control synthesis methodology can be used for adaptation to changes in control authority such as an actuator failure. Finally, we described experiments run to test these algorithms on an autonomous underwater vehicle. The experiments demonstrated the potential for accurate motion control based on our methodology.

In Chapter 6 we applied our theory of averaging on Lie groups to address the control synthesis problem for bilinear systems on $\mathbb{R}^n$. This was motivated by the fact that the evolution of the state transition matrix of such a system is described by a right-invariant system on a matrix Lie group. As an example we considered a switched electrical network that models the conversion portion of a dc-dc voltage converter and the problem of controlling energy transfers between dynamic storage elements to meet certain performance criteria. The state transition matrix for this example system is described by a system with drift on the Lie group $SO(3)$. Using average formulas for the state transition matrix evolution we derived a controller design methodology. We illustrated the use of open-loop controls derived from this methodology and successfully showed that the energy in the circuit could be transferred as desired with a finite number of switchings. Open-loop controls have an advantage over purely feedback techniques which sometimes can lead to chattering.

There are several directions for future research related to the work described in this dissertation. One of these will be to investigate the possibility of extending our results to systems on infinite-dimensional Lie groups. Another direction will be to formalize the process of combining open-loop with feedback controls as well as adaptation schemes using a higher level supervisory control strategy. If we use the open-loop controls of this dissertation which solve the motion control problem with a high order of accuracy, then we can consider local feedback control strategies to make the system more robust.

We will also study motion control problems which are dynamic rather than kinematic, i.e., for which the control inputs are forces and torques rather than velocities (e.g., see Baillieul [6, 5]). This will involve, among other things, dealing

with a drift term and motivates future research addressing control synthesis for systems on Lie groups with a drift term (i.e., extending the methodology of Chapter 6). Addressing the motion control problem for systems with drift will allow us to consider control synthesis for bilinear systems more general than the network example of Chapter 6 as well as for nonlinear systems that can be approximated by bilinear systems.

Further work will be done to apply the theory of averaging on Lie groups to the problem of stabilization of a system equilibrium point. For example, we will address the problem of stabilization with periodically time-varying controls when $m < n$ controls are available and n is the dimension of the configuration space of the system (e.g., a Lie group).

Additionally, it is of great interest to continue experimental testing of our algorithms, particularly for a system such as an underwater vehicle where there are sensors which can measure position and velocity.

Bibliography

- [1] R. Abraham and J. E. Marsden. *Foundations of Mechanics*. Addison-Wesley, 2nd edition, 1978.
- [2] D. L. Akin and R. D. Howard. Telerobotic operations testing in neutral buoyancy simulation. In *RPI Conference on Space Robotics*, 1992.
- [3] D. E. Anderson, C. A. Buck, and R. Cohen. Applications of free-flying cameras for space-based operations. In *SAE/AIAA International Conference on Environmental Systems*, Friedrichshafen, Germany, June 1994.
- [4] J. Baillieul. *Some Optimization Problems in Geometric Control Theory*. PhD thesis, Harvard University, 1975.
- [5] J. Baillieul. Energy methods for stability of bilinear systems with oscillatory inputs. Boston University preprint, 1993.
- [6] J. Baillieul. Stable average motions of mechanical systems subject to periodic forcing. In *Dynamics and Control of Mechanical Systems: The Falling Cat and Related Problems*, Fields Institute Communications, **1**, pages 1–23. AMS, Providence, RI, 1993.

- [7] J. Barraquand and J.-C. Latombe. Motion planning with many degrees of freedom and dynamic constraints. In *International Symposium on Robotics Research*, pages 74–83, 1989.
- [8] D. Bleeker. *Gauge Theory and Variational Principles*. Addison-Wesley, Reading, MA, 1981.
- [9] N. N. Bogoliubov and Y. A. Mitropolsky. *Asymptotic Methods in the Theory of Non-Linear Oscillations*. Hindustan Publishing Corp., India, 1961.
- [10] O. Bottema and B. Roth. *Theoretical Kinematics*. North-Holland Publishing Co., 1979.
- [11] R. W. Brockett. System theory on group manifolds and coset spaces. *SIAM Journal of Control*, 10(2):265–284, May 1972.
- [12] R. W. Brockett. Lie algebras and Lie groups in control theory. In D. O. Mayne and R. W. Brockett, editors, *Geometric Methods in Control Theory*, pages 43–82. Reidel, Dordrecht, Holland, 1973.
- [13] R. W. Brockett. Control theory and singular Riemannian geometry. In *New Directions in Applied Mathematics*, pages 13–27. Springer-Verlag, 1982.
- [14] R. W. Brockett. On the rectification of vibratory motion. *Sensors and Actuators*, 20(1-2):91–96, 1989.
- [15] R. W. Brockett. Formal language for motion description and map making. In R. W. Brockett, editor, *Proceedings of Symposia in Applied Mathematics*, pages 181–193, Providence, 1990. American Mathematical Society. Volume 41.

- [16] R. W. Brockett and L. Dai. Non-holonomic kinematics and the role of elliptic functions in constructive controllability. In Z. Li and J. F. Canny, editors, *Nonholonomic Motion Planning*, pages 1–21. Kluwer Academic, 1993.
- [17] R. W. Brockett and J. R. Wood. Electrical networks containing controlled switches. In *Applications of Lie Group Theory to Nonlinear Network Problems*, pages 1–11. Western Periodicals Co., 1974. Supplement to IEEE International Symposium on Circuit Theory.
- [18] P. E. Crouch. Spacecraft attitude control and stabilization: Applications of geometric control theory to rigid body models. *IEEE Transactions on Automatic Control*, AC-29(4):321–331, 1984.
- [19] M. L. Curtis. *Matrix Groups*. Springer-Verlag, 2nd edition, 1984.
- [20] R. P. Feynman. An operator calculus having applications in quantum electrodynamics. *Physical Review*, 84(1):108–128, 1951.
- [21] A. T. Fomenko and R. V. Chakon. Recursion relations for homogeneous terms of a convergent series of the logarithm of a multiplicative integral on Lie groups. *Functional Analysis and its Applications*, 24(1):48–58, January - March 1990. Translated from Russian.
- [22] C. Frohlich. The physics of somersaulting and twisting. *Scientific American*, 263:155–164, March 1980.
- [23] V. Gershkovich and A. Vershik. Nonholonomic manifolds and nilpotent analysis. *J. Geometry Phys.*, 5(3):407–452, 1988.

- [24] L. Gurvits. Averaging approach to nonholonomic motion planning. In *Proceedings of the 1992 IEEE International Conference on Robotics and Automation*, pages 2541–2546, Nice, France, May 1992.
- [25] L. Gurvits and Z. Li. Smooth time-periodic feedback solutions for nonholonomic motion planning. In Z. Li and J. F. Canny, editors, *Nonholonomic Motion Planning*, pages 53–108. Kluwer Academic, 1993.
- [26] R. Hermann. *Lie Groups for Physicists*. Benjamin, 1966.
- [27] N. Jacobson. *Lie Algebras*. John Wiley & Sons, 1962. Also Dover publication, 1979.
- [28] V. Jurdjevic. Optimal control problems on Lie groups: Crossroads between geometry and mechanics. In B. Jakubczyk and W. Respondek, editors, *Geometry of Feedback and Optimal Control*. Marcel-Dekker, 1992. to appear.
- [29] V. Jurdjevic and H. J. Sussmann. Control systems on Lie groups. *Journal of Differential Equations*, 12:313–329, 1972.
- [30] T. R. Kane and M. P. Scher. A dynamical explanation of the falling cat phenomenon. *International Journal on Solids and Structures*, 5:663–670, 1969.
- [31] M. V. Karasev and M. V. Mosolova. Infinite products and T products of exponentials. *Theoretical and Mathematical Physics*, 28:721–729, 1976. Translated from Russian.
- [32] H. K. Khalil. *Nonlinear Systems*. Macmillan Publishing, New York, 1992.

- [33] P. S. Krishnaprasad. Lagrangian and Hamiltonian control systems on bundles. In *Workshop on Mechanics, Holonomy and Control*, IEEE Conference on Decision and Control, San Antonio, TX, December 1993.
- [34] P. S. Krishnaprasad. Optimal control and Poisson reduction. Technical Report 93-87, Institute for Systems Research, University of Maryland, 1993.
- [35] P. S. Krishnaprasad and R. Yang. Geometric phases, anholonomy, and optimal movement. In *Proceedings of the 1991 IEEE International Conference on Robotics and Automation*, pages 2185–2189, Sacramento, CA, April 1991.
- [36] P. S. Krishnaprasad, R. Yang, and W. P. Dayawansa. Control problems on principal bundles and nonholonomic mechanics. In *Proceedings of the 30th IEEE Conference on Decision and Control*, pages 1133–1138, Brighton, England, December 1991.
- [37] G. Lafferriere and H. Sussmann. Motion planning for controllable systems without drift. In *Proceedings of the 1991 IEEE International Conference on Robotics and Automation*, pages 1148–1153, Sacramento, CA, April 1991.
- [38] G. Lafferriere and H. J. Sussmann. Motion planning for controllable systems without drift: A preliminary report. Report SYCON-91-4, Rutgers Center for Systems and Control, June 1990.
- [39] J.-P. Laumond. Feasible trajectories for mobile robots with kinematic and environmental constraints. In *Intelligent Autonomous Systems*, pages 346–354. North-Holland, Amsterdam, The Netherlands, 1987.

- [40] J.-P. Laumond. Finding collision-free smooth trajectories for a nonholonomic mobile robot. In *Proceedings of the 10th International Joint Conference on Artificial Intelligence*, pages 1120–1123, Milano, Italy, 1987.
- [41] M. Lazard and J. Tits. Domaines d’injectivité de l’application exponentielle. *Topology*, 4:315–322, 1966.
- [42] N. E. Leonard. Control synthesis and adaptation for an underactuated autonomous underwater vehicle. Technical Report 94-58, Institute for Systems Research, University of Maryland, 1994.
- [43] N. E. Leonard and P. S. Krishnaprasad. Averaging and motion control on Lie groups. Technical Report 93-59, Institute for Systems Research, University of Maryland, 1993.
- [44] N. E. Leonard and P. S. Krishnaprasad. Averaging for attitude control and motion planning. In *Proceedings of the 32nd IEEE Conference on Decision and Control*, pages 3098–3104, San Antonio, TX, December 1993.
- [45] N. E. Leonard and P. S. Krishnaprasad. Averaging on Lie groups, attitude control and drift. In *Proceedings of the Twenty-Seventh Annual Conference on Information Sciences and Systems*, pages 369–374, The Johns Hopkins University, March 1993.
- [46] N. E. Leonard and P. S. Krishnaprasad. Control of switched electrical networks using averaging on Lie groups. In *Proceedings of the 33rd IEEE Conference on Decision and Control*, Orlando, FL, December 1994. In press.

- [47] N. E. Leonard and P. S. Krishnaprasad. High-order averaging on Lie groups and control of an autonomous underwater vehicle. In *Proceedings of the American Control Conference*, pages 157–162, Baltimore, MD, June 1994.
- [48] N. E. Leonard and P. S. Krishnaprasad. Motion control of autonomous underwater vehicle with an adaptive feature. In *Proceedings of the 1994 Symposium on Autonomous Underwater Vehicle Technology*, pages 283–288. IEEE Oceanic Engineering Society, July 1994.
- [49] N. E. Leonard and P. S. Krishnaprasad. Motion control of drift-free, left-invariant systems on Lie groups, part I: Averaging and controllability. Technical Report 94-8, Institute for Systems Research, University of Maryland, 1994.
- [50] N. E. Leonard and P. S. Krishnaprasad. Motion control of drift-free, left-invariant systems on Lie groups, part II: A general constructive control algorithm. Technical Report 94-9, Institute for Systems Research, University of Maryland, 1994.
- [51] S. Lie. *Sophus Lie's 1884 Differential Invariant Paper*. Math Sci Press, 1976. Comments and additional material by R. Hermann.
- [52] M. J. Lighthill. On the squirming motion of nearly spherical deformable bodies through liquids at very small Reynolds numbers. *Communications on Pure and Applied Mathematics*, V:109–118, 1952.
- [53] W. Liu. *Averaging Theorems for Highly Oscillatory Differential Equations and the Approximation of General Paths by Admissible Trajectories for Non-*

- holonomic Systems*. PhD thesis, Rutgers University, New Brunswick, NJ, October 1992.
- [54] W. Magnus. On the exponential solution of differential equations for a linear operator. *Communications on Pure and Applied Mathematics*, VII:649–673, 1954.
 - [55] G. Meyer. Design and global analysis of spacecraft attitude control systems. Technical report TR R-361, NASA, Ames Research Center, March 1971.
 - [56] R. Montgomery. Isoholonomic problems and some applications. *Communications in Mathematical Physics*, 128:565–592, 1990.
 - [57] R. Montgomery. Nonholonomic control and gauge theory. In Z. Li and J. F. Canny, editors, *Nonholonomic Motion Planning*, pages 343–377. Kluwer Academic, 1993.
 - [58] R. M. Murray and S. S. Sastry. Steering nonholonomic systems using sinusoids. In *Proceedings of the 29th IEEE Conference on Decision and Control*, pages 2097–2101, Honolulu, HI, December 1990.
 - [59] R. M. Murray and S. S. Sastry. Steering nonholonomic systems in chained form. In *Proceedings of the 30th IEEE Conference on Decision and Control*, pages 1121–1126, Brighton, England, December 1991.
 - [60] R. M. Murray and S. S. Sastry. Nonholonomic motion planning: Steering using sinusoids. *IEEE Transactions on Automatic Control*, 38(5):700–716, 1993.
 - [61] H. Nijmeijer and A. J. van der Schaft. *Nonlinear Dynamical Control Systems*. Springer-Verlag, New York, 1990.

- [62] K. Nomizu. *Lie Groups and Differential Geometry*. The Mathematical Society of Japan, 1956.
- [63] E. M. Purcell. Life at low Reynolds number. *American Journal of Physics*, 45(1):3–11, January 1977.
- [64] J. A. Sanders and F. Verhulst. *Averaging Methods in Nonlinear Dynamical Systems*. Springer-Verlag, New York, 1985.
- [65] W. Schmid. Poincaré and Lie groups. In *The Mathematical Heritage of Henri Poincaré*, pages 157–168. American Mathematical Society, 1983. Proceedings of Symposia in Pure Mathematics, Volume 39, Part I.
- [66] J.-P. Serre. *Lie Algebras and Lie Groups*. W. A. Benjamin, New York, 1965.
- [67] R. P. Severns and G. Bloom. *Modern DC-to-DC Switchmode Power Converter Circuits*. Van Nostrand Reinhold, 1985.
- [68] A. Shapere and F. Wilczek, editors. *Geometric Phases in Physics*. World Scientific, Singapore, 1989.
- [69] A. Shapere and F. Wilczek. Geometry of self-propulsion at low Reynolds number. *Journal of Fluid Mechanics*, 198:557–585, 1989.
- [70] H. Sira-Ramirez. Sliding motions in bilinear switched networks. *IEEE Transactions on Circuits and Systems*, 34(8):919–933, August 1987.
- [71] O. J. Sørдалen. *Feedback Control of Nonholonomic Mobile Robots*. PhD thesis, The Norwegian Institute of Technology, 1993.
- [72] N. Sreenath. Nonlinear control of planar multibody systems in shape space. *Mathematics of Control, Signals, and Systems*, 5:343–363, 1992.

- [73] H. J. Sussmann and W. Liu. Limits of highly oscillatory controls and the approximation of general paths by admissible trajectories. In *Proceedings of the 30th Conference on Decision and Control*, pages 437–442, 1991.
- [74] G. Taylor. Analysis of the swimming of microscopic organisms. *Proceedings of the Society of London, Series A*, 209:447–461, 1951.
- [75] G. Taylor. The action of waving cylindrical tails in propelling microscopic organisms. *Proceedings of the Royal Society of London, Series A*, 211:225–239, 1952.
- [76] F. W. Warner. *Foundations of Differentiable Manifolds and Lie Groups*. Springer-Verlag, 1983.
- [77] J. Wei and E. Norman. On global representations of the solution of linear differential equations as a product of exponentials. In *Proceedings of the American Mathematical Society*, pages 327–334, April 1964.
- [78] J. R. Wood. Power conversion in electrical networks. Technical Report NASA Rep. No. CR-120830, (also PhD thesis), Harvard University, 1974.
- [79] R. Yang. *Nonholonomic Geometry, Mechanics and Control*. PhD thesis, University of Maryland, 1992.