TECHNICAL RESEARCH REPORT

Averaging and Motion Control On Lie Groups

by N.E. Leonard and P.S. Krishnaprasad

T.R. 93-59

The Institute for Systems Research is supported by the National Science Foundation Engineering Research Center Program (NSFD CD 8803012), the University of Maryland, Harvard University, and Industry

Averaging and Motion Control on Lie Groups [1]

Naomi Ehrich Leonard and P. S. Krishnaprasad

Institute for Systems Research
&
Department of Electrical Engineering
University of Maryland
College Park, MD 20742

Abstract

The deeper investigation of problems of feedback stabilization and constructive controllability has drawn increased attention to the question of structuring control systems. Thus, for instance, it is interesting to know how to combine periodic open loop controls with intermittent feedback corrections to achieve prescribed behavior in robotic motion planning systems. As a first step towards understanding this type of question, it would be useful to obtain some insight into the average behavior of a periodically forced system. In the present paper we are primarily interested in periodic forcing of left-invariant systems on Lie groups such as would arise in spacecraft attitude control. We prove averaging theorems applicable to systems evolving on general matrix Lie groups with particular focus on the attitude control problem. The results of this paper also yield useful formulae for motion planning of a variety of other systems such as an underwater vehicle which can be modelled as a control system evolving on the Lie group $SE(3)$.

Keywords: Averaging, Attitude Control, Left Invariant Systems, Lie Groups.

[1] This research was 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.

1 Introduction

Recent work in nonlinear control has drawn attention to drift-free systems with fewer controls than state variables. These arise in problems of motion planning for wheeled robots subject to nonholonomic constraints [1], models of kinematic drift (or geometric phase) effects in space systems subject to appendage vibrations or articulations [2, 3], and models of self-propulsion of paramecia at low Reynolds numbers [4]. The basic state-space model takes the form,

$$\dot{x} = \sum_{i=1}^{m} F_i(x)u_i, \quad x \in \Re^n, \;\; u_i \in \Re, \;\; n > m \; . \tag{1}$$

It is well known that if the vector fields F_i satisfy a Lie algebra rank condition, then there exists a control $u = (u_1, \ldots, u_m)$ that drives the system to the origin from any initial state. However, 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. As a result, recent research has focused on constructing controls to achieve complete controllability [5, 6, 1, 7, 8]. The success of constructive procedures based on periodically time-varying controls [1, 7, 8] motivates our investigation of periodic forcing on a general family of systems.

In the present paper we are interested in drift-free systems of the form

$$\dot{X} = XU \tag{2}$$

evolving on matrix Lie groups. Here $X(t)$ is a curve in a matrix Lie group G, $U(t)$ is a curve in the Lie algebra $\mathcal{G}$ of G. (For an introduction to matrix Lie groups and Lie algebras see [9]). Equation (2) describes rigid-body kinematics if we interpret $U(t)$ as the time-dependent skew symmetric matrix of body angular velocity such that X evolves on $G = SO(3)$ where

$$SO(3) \triangleq \{A \in \Re^{3\times 3} | A^T A = I, \; det(A) = 1\}.$$

In this case the system is nonholonomic where conservation of (zero) angular momentum is interpreted as the nonholonomic constraint. Alternatively, equation (2) describes the kinematics of a body underwater for $G = SE(3)$, the special Euclidean group, where

$$SE(3) \triangleq \{\begin{bmatrix} A & b \\ 0 & 1 \end{bmatrix} \in \Re^{4\times4} | A \in SO(3), b \in \Re^3\}.$$

In this case $U(t)$ describes body angular and translational velocity.

Our objective is to prescribe means to completely control these types of systems using small-amplitude periodic controls. One strategy is to use periodic controls to provide open loop control of the system and apply intermittent feedback corrections to make finer adjustments in system behavior. 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 accuracy and sensitivity reduction associated with feedback control. (For related ideas see [10]). As part of developing this strategy, we investigate using averaging theory for systems of the form (2) as a means to specify open loop periodic control. The goal of averaging theory in this context is to describe an approximate solution to (2) that remains close to the actual solution to (2), but gives rise to straightforward procedures for achieving complete constructive controllability.

To better identify the systems of interest as control systems with small-amplitude periodic input, we rewrite system (2) as

$$\dot{X} = \epsilon XU, \qquad U(t) = \sum_{i=1}^{n} A_i u_i(t), \tag{3}$$

where again $X(t)$ is a curve in a matrix Lie group G, and $U(t)$ is a curve in the Lie algebra $\mathcal{G}$ of G. The $u_i(\cdot)$ are assumed to be periodic functions of common period T with $n = dim(G)$ and $\epsilon > 0$ a small parameter. $A_1, \ldots, A_n$ form a basis for $\mathcal{G}$. The $\epsilon u_i(\cdot)$ are interpreted as the small-amplitude periodic controls, although some of the $u_i(\cdot)$ may be identically zero.

Returning to the example of a rigid spacecraft ($G = SO(3)$), it is appropriate to consider angular velocity as our control if we assume that (zero) angular momentum of the spacecraft is conserved, i.e., there is no external torque applied to the spacecraft.

Periodic angular velocities can be achieved in practice, for instance, by means of momentum wheels or by means of oscillating appendages. This latter method is motivated by recent results in the study of geometric phase [2, 3, 4] which show that periodic changes in the "shape" of a partially rigid body lead to attitude drift as if the kinematic model were being driven by periodic controls.

Developing attitude control designs using small-amplitude periodic controls is practically motivated in part by the new focus in the space industry on miniaturization and control of small planetary spacecraft. To maintain the low weight and small size of this new generation of spacecraft, small actuators will be needed for control. However, since scaling down of conventional motors has inherent problems and limitations, new means of actuation and control need to be considered. Small-amplitude periodic control is a promising new alternative. Periodic controls are also motivated by recent technological advances in the area of micro-actuation and sensing. For example, new kinds of micro-actuators and micro-sensors such as piezoelectric vibratory actuators and rotation-sensing vibratory sensors work on the principle of oscillation-induced motion, i.e., operate by means of periodic signals. In [11], Brockett examines these types of actuators using nonlinear control concepts and illustrates the role of averaging and holonomy. Additionally, periodic controls can play an important role in stabilization problems. Coron has shown that while systems of the form (1) cannot be stabilized using a smooth feedback law, they can be stabilized with time-varying (periodic) feedback laws [12].

As a second example, for $G = SE(3)$ system (3) captures the kinematics of an underwater vehicle control problem if we interpret the vehicle angular velocities and translational velocities as the periodic controls. A stabilizing control law for an underwater vehicle was developed in [13] using a kinematic model on $SE(3)$. Alternatively, it is important to note that models of the form (3) also arise in treating systems such as spacecraft or underwater vehicles which experience small-amplitude oscillatory disturbances. As a result, our formulae which indicate how to achieve controllability with periodic controls also reveal how to compute drifts in system behavior caused by undesirable vibrations and oscillations. Kinematic drift of a spacecraft caused by thermo-elastically induced vibrations in flexible attachments on the spacecraft is an example.

One of the major difficulties in working with systems of the form (2) (or (3)) is that, in general, there are no explicit global representations of solutions to (2). However, much is known about local representations and certain recursive forms [14, 15, 16]. For example, Wei and Norman show that for any G, there exists the product of exponential representation

$$X(t) = \prod_{i=1}^{n} e^{g_i(t)A_i} \triangleq e^{g_1(t)A_1} e^{g_2(t)A_2} \cdots e^{g_n(t)A_n} , \tag{4}$$

where $A_1, \ldots, A_n$ is a basis for $\mathcal{G}$ and $g_1(t), \ldots, g_n(t)$ are given for $|t| < t_0$ by solving a system of ordinary differential equations (called the Wei-Norman equations). Alternatively, Magnus shows how to express $X(t)$ as a single exponential

$$X(t) = e^{Z(t)} \tag{5}$$

where $Z(t) \in \mathcal{G}$ is given as an infinite series of iterated integrals. Classical averaging theory is equipped to handle systems which evolve on $\Re^n$. As a result, to provide an averaging theory on the group level for systems of the form (3), we make use of representations (4) and (5).

Our goal is to approximate the solution $X(t)$ of (3) by its "average" solution $\bar{X}(t)$ such that $\bar{X}(t)$ provides a "good" approximation to $X(t)$ over a "sufficiently long" time interval. The average solution $\bar{X}$ must be defined so that it makes sense in two important ways. First, for $\bar{X}(t)$ to be a meaningful approximation of $X(t)$, it must be a curve in the matrix Lie group G. Secondly, we must be able to define some notion of the error between the average and actual solutions so that we can give an estimate on the magnitude of this error over the sufficiently long time interval.

The advantage in approximating $X(t)$ by its average $\bar{X}(t)$ is that $\bar{X}(t)$ will satisfy an ordinary differential equation with significantly reduced complexity as compared to (3). Thus, the problem of controlling the system described by (3) is reduced to the problem of controlling a much simpler system. Further, as will be shown most dramatically in Section 3, the averaged system admits a clear geometric interpretation. This geometric interpretation can be used to advantage in the design of open loop and feedback controls for systems of the form (3).

To illustrate the theoretical inspiration for our results, consider a system described by (1) such that $m = 2$, $n = 3$ and $rank([F_1(x) \;\; F_2(x) \;\; [F_1, F_2](x)]) = 3$, $\forall x \in \Re^3$. For $F_1(\cdot)$ and $F_2(\cdot)$ smooth vector fields on $\Re^n$, $[F_1, F_2](\cdot)$ is a new vector field on $\Re^n$ called the Lie bracket and defined by

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

The kinematic motion planning problem for a unicycle with velocity and steering control is an example of a system of this form satisfying the above rank condition [17]. The rank condition in this special case implies that the system is completely controllable since (1) has no drift vector field [17]. Specifically, one can reach any point in $\Re^3$ by flowing along the vector fields F_1, F_2 and $[F_1, F_2]$ which span $\Re^3$ at every point in $\Re^3$.

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 1(a). These controls yield

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

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 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 $\epsilon^2 =$ Area of $\tilde{u}_1$ versus $\tilde{u}_2$ during one period. As Brockett argues in [11], 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 1(c) would produce

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

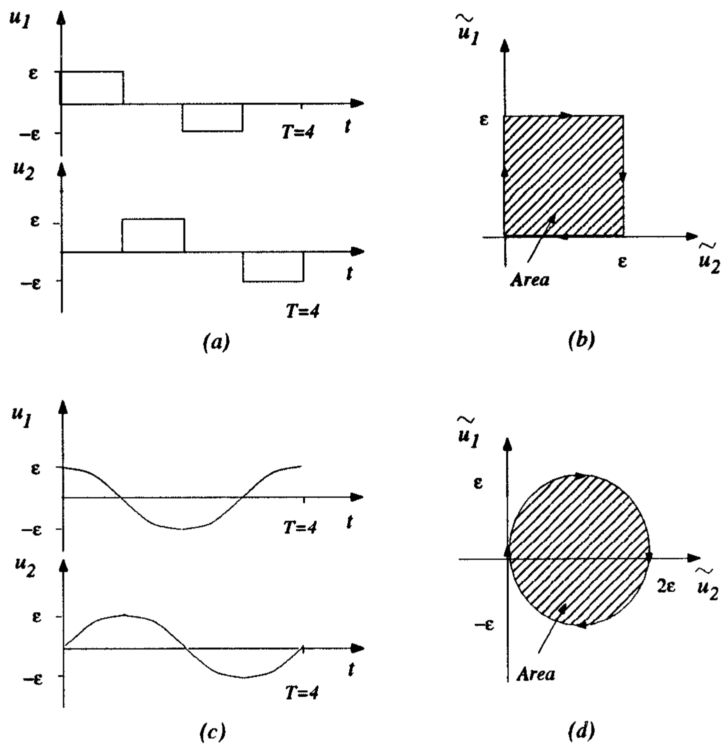

Figure 1: Periodic Controls.

where *Area* is the shaded area of Figure 1(d). Then in some average sense one would predict that

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

Our results generalize this result to the class of systems described by (3). In the present paper we examine both first-order and second-order average terms of the solution to (3). Our results show that the first-order average solution exhibits the contribution of the dc component of the periodic input signal, while the second-order average terms arise from the group level version of depth-one Lie brackets of vector fields. For other recent developments in averaging techniques for systems with highly oscillatory controls, see [7, 8].

In Section 2 we summarize results from our work on first-order averaging [18]. In Section 3 we prove a second-order averaging theorem for general matrix Lie groups. Our main result is an "area rule" for systems on groups. We further examine these results for the Lie group $SO(3)$ in the context of attitude control. The results of Sections 2 and 3 make use of the Wei-Norman product of exponential representation (4) of solutions to (2). We develop analogous first and second-order averaging results using the single exponential representation (5) in Section 4. In Section 5 we describe the geometric interpretation of the results and the consequences for control illustrating how to use the averaging results to achieve constructive controllability. An example is given for achieving complete constructive controllability for the spacecraft attitude control problem when only two controls are available. In Section 6 we explore the spacecraft attitude control problem in terms of practical means of actuation, i.e., via momentum wheels and appended point-mass oscillators, and derive the average approximation formulae in these special cases.

2 First-Order Averaging

We begin by considering the first-order average $\bar{X}$ of the solution X to (3). Since there are no explicit global representations of the solution to (3), we use a local representation. In particular, we work with the local representation given by Wei and Norman [15] for solutions $X(t)$ of (3). In Section 4 we give analogous results based on the single exponential representation of Magnus [14]. However, we choose to develop the Wei-Norman representation first because it has the advantage over the single exponential representation of providing a global representation when the Lie algebra $\mathcal{G}$ is solvable (see comments by Wei and Norman in this connection).

Suppose that $X(0) = I$, then Wei and Norman show that for any G, $\exists t_0 > 0$ such that for $|t| < t_0$ there exists the product of exponential representation

$$X(t) = e^{g_1(t)A_1} e^{g_2(t)A_2} \cdots e^{g_n(t)A_n} \ , \tag{6}$$

where $A_1, \ldots, A_n$ is a basis for $\mathcal{G}$ and $g_1(t), \ldots, g_n(t)$ are given for $|t| < t_0$ by solving a system of ordinary differential equations (called the Wei-Norman equations). It follows from their work in [15], that the form of the Wei-Norman equations is as in the following lemma.

Lemma 1. For solutions to (3) of the form (6) with initial condition $X(0) = I$, the Wei-Norman equations take the form

$$\dot{g} = \epsilon M(g)u \ , \quad \text{for } |t| < t_0 \ , \tag{7}$$

where $u = (u_1, \ldots, u_n)^T$, $g = (g_1, \ldots, g_n)^T$, $g(0) = 0$ and $M(g)$ is a real analytic matrix-valued function of g. If $\mathcal{G}$ is solvable then there exists a basis of $\mathcal{G}$ and an ordering of this basis for which (7) holds globally, i.e., for all t. $\square$

It is customary to refer to components of g as the canonical coordinates of the second kind for G. Let W be the open neighborhood of $0 \in \Re^n$ such that $\forall g \in W$, $M(g)$ is well-defined. Let $\Phi : \Re^n \to G$ define the mapping

$$\Phi(g) = e^{g_1 A_1} e^{g_2 A_2} \cdots e^{g_n A_n} \tag{8}$$

and define $V \equiv \Phi(W) \subset G$. Then, the Wei-Norman formulation provides a local representation of the solution to (3) for initial condition $X(0) \in V \subset G$. Now let S be the largest neighborhood of $0 \in \Re^n$ contained in W such that $\Psi \equiv \Phi|_S : S \to G$ is one-to-one. Let $Q \equiv \Psi(S) \subset V$. Then $\Psi : S \to Q$ is a diffeomorphism and we can define a metric $\tilde{d} : Q \times Q \to \Re_+$ by

$$\tilde{d}(Y, Z) = d(\Psi^{-1}(Y), \Psi^{-1}(Z)) \tag{9}$$

where $d : \Re^n \times \Re^n \to \Re_+$ is given by

$$d(\alpha, \beta) = \|\alpha - \beta\|_1 = \sum_{i=1}^{n} |\alpha_i - \beta_i|. \tag{10}$$

Our approach to a first-order averaging theory for systems of the form (3) is to apply classical averaging theory to the system (7), obtain estimates for approximations of solutions to (7) over $O(1/\epsilon)$ time intervals and then transfer such estimates to the group level for solutions to (3). The averaged system associated with (7) is defined as

$$\begin{aligned}\dot{\bar{g}} &= \epsilon M(\bar{g})(\frac{1}{T}\int_0^T u(\tau)d\tau) \\ &\triangleq \epsilon M(\bar{g})u_{av}, \quad \bar{g}(0)=\bar{g}_0 \end{aligned} \tag{11}$$

where we assume that $M(\bar{g}_0)$ is well-defined.

Now we define the average solution $\bar{X}(t)$ associated with the solution $X(t)$ of (3) as the solution to

$$\dot{\bar{X}} = \epsilon \bar{X} U_{av}, \quad U_{av} = \sum_{i=1}^{n} A_i u_{avi} \tag{12}$$

where $u_{av} = (u_{av1}, \ldots, u_{avn})^T$. Thus,

$$\bar{X}(t) = \bar{X}(0)e^{\epsilon U_{av} t} \tag{13}$$

and $\bar{X}(t)$ is a curve in the matrix Lie group G as long as $\bar{X}(0) \in G$.

Then from [18] we have the following local representation of $\bar{X}$.

Theorem 1. Consider the solution $\bar{X}(t)$ to (12) with $\bar{X}(0) \in V$. Then $\exists \bar{t} > 0$ such that for $t \in [0, \bar{t}]$,

$$\bar{X}(t) = e^{\bar{g}_1(t)A_1} e^{\bar{g}_2(t)A_2} \cdots e^{\bar{g}_n(t)A_n}, \tag{14}$$

where $\bar{g} = (\bar{g}_1, \ldots, \bar{g}_n)^T$ is the solution to (11) with $\bar{g}(0) = \bar{g}_0$ such that $\Phi(\bar{g}_0) = \bar{X}(0)$. $\square$

Theorem 1 provides the means to transfer classical averaging theory results on the local level back to the group level. We can now state the first order averaging theorem and corollary from [18], which are applications of standard averaging theory (c.f. [19]).

Theorem 2. Let $\epsilon > 0$. Let $D = \{g \in \Re^n \mid \|g\| < r\} \subset S$. Assume that $u(t) \in \Re^n$ is periodic in t with period $T > 0$ and has continuous derivatives up to second order for $t \in [0, \infty)$. Let $X(t)$ be the solution to (3) represented by (6) where $g(t, \epsilon)$ is the solution to (7) with $\Psi(g(0, \epsilon)) = X(0)$ and $g(0, \epsilon) \in D$. Let $\bar{X}(t)$ be the solution to (12) represented by (14) where $\bar{g}(t, \epsilon)$ is the solution to (11) with

$\Psi(\bar{g}(0,\epsilon)) = \bar{X}(0)$.

If $\bar{g}(t,\epsilon) \in D, \quad \forall t \in [0, b/\epsilon]$ and $\|g(0,\epsilon) - \bar{g}(0,\epsilon)\| = O(\epsilon)$

$$\text{then } \|g(t,\epsilon) - \bar{g}(t,\epsilon)\| = O(\epsilon) \text{ on } [0, b/\epsilon],$$

$$\text{i.e., } \|g(t,\epsilon) - \bar{g}(t,\epsilon)\| \leq k\epsilon, \quad \forall \epsilon \in [0, \epsilon_1), \quad \forall t \in [0, b/\epsilon]$$

for some $k > 0$ and $\epsilon_1 > 0$. Further,

$$\tilde{d}(X(t), \bar{X}(t)) \leq k\epsilon, \quad \forall \epsilon \in [0, \epsilon_1), \quad \forall t \in [0, b/\epsilon]. \quad \Box$$

In the next corollary (c.f. [18]), we give a version of Theorem 2 for $G = SO(3)$. In studying the case of $G = SO(3)$, we have in mind applications to rigid-body kinematics and the spacecraft attitude control problem. Instead of using the metric $\tilde{d}$ on $SO(3)$ we use a metric $\bar{\phi}$ on $SO(3)$ and the norm $\|\bar{X}^{-1}X - I\|_1$, both of which give a measure of rotational error. Specifically, following [20] we define $R = \bar{X}^{-1}X$, which may be interpreted as the rotation taking the actual trajectory to the average trajectory. Then we let R define system attitude error with $\phi(R) = cos^{-1}{}_{[0,\pi]}(1/2(tr(R) - 1))$ the magnitude of attitude error (where tr is the trace operator). Then by [20], $\bar{\phi}$, where $\bar{\phi}(A, B) = \phi(AB^T)$, is a metric on $SO(3)$ and can be interpreted as the minimal angular distance between the body coordinates of the actual system and the body coordinates of the average system. We also note that in the corollary since we do not use the metric $\tilde{d}$, we only need to restrict $D \subset W$ and not $D \subset S \subset W$ (see [18] for details).

Corollary 1. Let $\epsilon > 0$. Let $D = \{g \in \Re^3 \mid \|g\| < r \text{ and } |g_2| < \frac{\pi}{2} - \delta, \ \delta > 0\}$. Suppose $u(t) \in \Re^3$ is periodic with period $T > 0$ and has continuous derivatives up to second order for $t \in [0, \infty)$. Let $g(t,\epsilon)$ be the solution to the Wei-Norman equations for $SO(3)$,

$$\begin{bmatrix} \dot{g}_1 \\ \dot{g}_2 \\ \dot{g}_3 \end{bmatrix} = \frac{\epsilon}{cos g_2} \begin{bmatrix} cos g_3 & -sin g_3 & 0 \\ cos g_2 sin g_3 & cos g_2 cos g_3 & 0 \\ -sin g_2 cos g_3 & sin g_2 sin g_3 & cos g_2 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} \equiv \epsilon M(g) u \qquad (15)$$

with $g(0,\epsilon) \in D$.

Let $\bar{g}(t,\epsilon)$ be the solution to these equations for the case in which $u(t)$ is replaced by its average u_{av} over period T.

$$\text{If} \quad \bar{g}(t,\epsilon) \in D \quad \forall t \in [0, b/\epsilon] \quad \text{and} \quad \|g(0,\epsilon) - \bar{g}(0,\epsilon)\| = O(\epsilon) \; ,$$

$$\text{then} \quad \|g(t,\epsilon) - \bar{g}(t,\epsilon)\| = O(\epsilon) \quad \text{on} \quad [0, b/\epsilon],$$

$$\text{i.e.,} \quad \|g(t,\epsilon) - \bar{g}(t,\epsilon)\| \;\leq k\epsilon, \quad \forall \epsilon < \epsilon_1, \quad \forall t \in [0, b/\epsilon]$$

for some $k > 0$ and $\epsilon_1 > 0$.

Further, the corresponding solution X to (3) in the group $SO(3)$ with $X(0) = \Phi(g(0,\epsilon))$ admits the estimate $\bar{X}$ defined by (12) with $\bar{X}(0) = \Phi(\bar{g}(0,\epsilon))$ such that $\forall \epsilon \in [0, \epsilon_2)$, $\epsilon_2 \leq \epsilon_1$, $\forall t \in [0, b/\epsilon]$

$$\phi(\bar{X}^{-1}(t)X(t)) \leq k\epsilon,$$

$$\text{and} \quad \|\bar{X}(t)^{-1}X(t) - I\|_1 \leq 9|cos(k\epsilon) - 1| + 7|sin(k\epsilon)| + |sin(k\epsilon)|^2 \,. \quad \square$$

The choice of basis for $\mathcal{G}$ used in the corollary is $A_i = \hat{e}_i$ where e_i is the ith standard Euclidean basis vector and $\hat{}: \Re^3 \longrightarrow so(3)$ is defined for $x = (x_1, x_2, x_3)^T$ by

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

It was noted in [18] that the time interval $[0, b/\epsilon]$ for which $\bar{X}$ provides an $O(\epsilon)$ approximation of X depends on how long the solution $\bar{g}$ to (11) stays on D. In general, b will depend on u_{av}. This dependence for $G = SO(3)$ is explored further in [18].

In addition to providing an approximate solution to the system of differential equations (7), classical averaging theory can also be used to determine the stability of system equilibria. Since the system of equations (7) is only a valid system of Wei-Norman equations when $M(g)$ is nonsingular, the right hand side of equation (7) is zero only when

$u \equiv 0$. Thus, a nontrivial discussion of equilibria of (7) requires that we introduce feedback into our system.

Consider equation (3) and suppose that $u = u(t, X)$, where u is periodic in t of period T. Then (3) becomes

$$\dot{X} = XU(t, X). \tag{16}$$

Assuming that $X(t) \in Q$ during the time interval of interest, then $g(t) = \Psi^{-1}(X(t))$ is well-defined. So we can write $u = u(t, \Psi(g)) \equiv u(t, g)$, and equation (7) becomes

$$\dot{g} = \epsilon M(g) u(t, g). \tag{17}$$

The average equation associated with (17) is

$$\dot{\bar{g}} = \epsilon M(\bar{g}) \frac{1}{T} \int_0^T u(\tau, \bar{g}) d\tau. \tag{18}$$

So, for example, if $g^* \in S$ and $u(t, g^*) = 0$, $\forall t \geq 0$, then g^* is an equilibrium point of both (17) and (18) $\forall t \geq 0$. As shown in the proposition that follows, we can use averaging theory to draw conclusions about the local stability properties of (17) based on the stability properties of (18). Additionally, the existence of an exponentially stable equilibrium point for the averaged system (18) makes it possible to extend the approximations of Theorem 2 from an $O(1/\epsilon)$ time interval to an infinite time interval.

Proposition 1. Let $\epsilon > 0$. Let $D = \{g \in \Re^n \mid \|g\| < r\} \subset S$. Assume that $u(t, g)$ is periodic in t with period $T > 0$ and suppose that $M(g)u(t, g)$ is continuous and bounded with continuous and bounded derivatives up to second order with respect to both its arguments for $(t, g) \in [0, \infty) \times D$. Let $X(t)$ be the solution to (16) represented by (6) where $g(t, \epsilon)$ is the solution to (17) with $\Psi(g(0, \epsilon)) = X(0)$ and $g(0, \epsilon) \in D$. Let $\bar{X}(t)$ be defined by (14) where $\bar{g}(t, \epsilon)$ is the solution to (18) with $\Psi(\bar{g}(0, \epsilon)) = \bar{X}(0)$.

If $g^* \in D$ is an exponentially stable equilibrium point for (18) then $\exists \rho > 0$ such that

$$\text{if} \quad \|\bar{g}(0, \epsilon) - g^*\| < \rho \text{ and } \|g(0, \epsilon) - \bar{g}(0, \epsilon)\| = O(\epsilon)$$

$$\text{then} \quad \|g(t, \epsilon) - \bar{g}(t, \epsilon)\| \leq k\epsilon, \quad \forall \epsilon \in [0, \epsilon_1), \quad \forall t \in [0, \infty)$$

for some $k > 0$ and $\epsilon_1 > 0$, and

$$\tilde{d}(X(t), \bar{X}(t)) \leq k\epsilon, \quad \forall \epsilon \in [0, \epsilon_1), \quad \forall t \in [0, \infty).$$

Further, $\exists \epsilon^*$ such that for all $0 < \epsilon < \epsilon^*$, (17) has a unique exponentially stable periodic solution of period T in an $O(\epsilon)$ neighborhood of g^*. Similarly, (16) has a unique exponentially stable periodic solution of period T in an $O(\epsilon)$ neighborhood of $X^* = \Psi(g^*) \in Q$.

Proof. The proposition follows from Theorem 7.4 of [19] and the definition of $\tilde{d}$. □

As an example consider the case $G = SO(3)$ for the attitude control problem. We note that (g_1, g_2, g_3) correspond to a type of Euler angles [18], so suppose that we can measure g. Choose $u_i(t, g) = -2k_i g_i sin^2 t,\ k_i > 0$ for $i = 1, 2, 3$. Since $u(t, 0) = 0,\ \forall t$, then $g^* = 0$ is an equilibrium point of (17) and (18). From (15) and (18) we see that

$$\begin{bmatrix} \dot{\bar{g}}_1 \\ \dot{\bar{g}}_2 \\ \dot{\bar{g}}_3 \end{bmatrix} = \begin{bmatrix} \epsilon sec\bar{g}_2(-k_1\bar{g}_1 cos\bar{g}_3 + k_2\bar{g}_2 sin\bar{g}_3) \\ \epsilon(-k_1\bar{g}_1 sin\bar{g}_3 - k_2\bar{g}_2 cos\bar{g}_3) \\ \epsilon((tan\bar{g}_2)(k_1\bar{g}_1 cos\bar{g}_3 - k_2\bar{g}_2 sin\bar{g}_3) - k_3\bar{g}_3) \end{bmatrix} \triangleq f(\bar{g})$$

and

$$\frac{\partial f}{\partial \bar{g}}|_{\bar{g}=0} = \begin{bmatrix} -\epsilon k_1 & 0 & 0 \\ 0 & -\epsilon k_2 & 0 \\ 0 & 0 & -\epsilon k_3 \end{bmatrix}.$$

Thus, by Lyapunov's indirect method since $k_i > 0,\ \forall i,\ g^* = 0$ is an exponentially stable equilibrium point for (18). From Proposition 1, we can conclude that (17) has a unique exponentially stable periodic solution in an $O(\epsilon)$ neighborhood of $g^* = 0$ and so (16) has a unique exponentially stable periodic solution in an $O(\epsilon)$ neighborhood of $X^* = \Psi(0) = I$. However, since $g^* = 0$ is itself an equilibrium point for (17), the unique exponentially stable periodic solution about $g^* = 0$ must be the trivial solution $g^* = 0$. Thus, $g^* = 0$ is an exponentially stable equilibrium point for (17) and similarly $X^* = I$ is an exponentially stable equilibrium point for (16).

3 Second-Order Averaging

In this section we show the form of the *second-order* average approximation $\bar{\bar{X}}(t)$ to the solution of $X(t)$ of (3). To isolate the second-order averaging effect, we assume that $u_{av} = 0$. Thus, by (13) the first-order approximation $\bar{X}(t)$ from Section 2 is constant, i.e., $\bar{X}(t) = \bar{X}(0)$, $\forall t \geq 0$. We prove that $\bar{\bar{X}}(t)$ is an $O(\epsilon^2)$ approximation over an $O(1/\epsilon)$ time interval.

As in the case of first-order averaging, we use the Wei-Norman local representation (6) of solutions to (3) with associated Wei-Norman equations (7) as a means to do second-order averaging on the group level. Accordingly, we let $\bar{\bar{g}} = (\bar{\bar{g}}_1, \ldots, \bar{\bar{g}}_n)^T$ be the second-order average approximation of the solution $g(t)$ to (7). Then we define the second-order approximation $\bar{\bar{X}}$ on the group level as

$$\bar{\bar{X}}(t) = e^{\bar{\bar{g}}_1(t)A_1} e^{\bar{\bar{g}}_2(t)A_2} \cdots e^{\bar{\bar{g}}_n(t)A_n}, \tag{19}$$

which is well-defined for $\bar{\bar{g}}(t)$ well-defined.

First we make several definitions which will be used in the second-order averaging to follow. We assume that $u(t) = (u_1(t), \ldots, u_n(t))^T$ is periodic in t with common period T and that $u_{avi} = (1/T)\int_0^T u_i(\tau)d\tau = 0$ for all $i = 1, \ldots, n$. Now we define $\tilde{u} = (\tilde{u}_1, \ldots, \tilde{u}_n)^T$ by

$$\tilde{u}_i(t) = \int_0^t u_i(\tau)d\tau. \tag{20}$$

So $u = \dot{\tilde{u}}$ and $\tilde{u}$ is periodic in t with common period T. Next we define $Area_{ij}(T)$ to be 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 $t = T$. By Green's Theorem we can express this area as

$$Area_{ij}(T) = \frac{1}{2}\int_0^T (\tilde{u}_i(\sigma)\dot{\tilde{u}}_j(\sigma) - \tilde{u}_j(\sigma)\dot{\tilde{u}}_i(\sigma))d\sigma. \tag{21}$$

This area can be interpreted as the projection onto the i-j plane of the area enclosed by the curve $(\tilde{u}_1, \ldots, \tilde{u}_n)$ in one period.

Associated to the basis $\{A_1, \ldots, A_n\}$ for the Lie algebra $\mathcal{G}$ of G, let the structure constants be Γ_{ij}^k. Then Γ_{ij}^k are defined by

$$[A_i, A_j] = \sum_{k=1}^{n} \Gamma_{ij}^k A_k, \quad i, j = 1, \ldots, n \tag{22}$$

where $[\cdot, \cdot]$ is the Lie bracket on $\mathcal{G}$ defined by $[A, B] = AB - BA$.

As shown in the work of Wei and Norman, one can express $M(g)$ of (7) in terms of the structure constants above. Specifically, by differentiating (6) with respect to time, equating the result with (3) and pre-multiplying by X^{-1} we get

$$\sum_{j=1}^{n} \dot{g}_j \sum_{i=1}^{n} \xi_{ij}(g) A_i = \sum_{j=1}^{n} \dot{g}_j \left(\prod_{k=n}^{j+1} e^{-g_k ad_{A_k}} A_j\right) = \epsilon \sum_{i=1}^{n} A_i u_i$$

where $ad : \mathcal{G} \to end(\mathcal{G})$ is defined by $ad_X Y = [X, Y]$. So the ijth element of $M(g)^{-1}$ is $\xi_{ij}(g)$ and

$$\begin{aligned}
\sum_{i=1}^{n} \xi_{ij}(g) A_i &= \prod_{k=n}^{j+1} e^{-g_k ad_{A_k}} A_j \triangleq e^{-g_n ad_{A_n}} \cdots e^{-g_{j+1} ad_{A_{j+1}}} A_j \\
&= (I - g_n ad_{A_n} + O(g_n{}^2)) \cdots (I - g_{j+1} ad_{A_{j+1}} + O(g_{j+1}{}^2)) A_j \\
&= (I - g_n ad_{A_n} - \ldots - g_{j+1} ad_{A_{j+1}}) A_j + O(g^2) \\
&= A_j - g_n [A_n, A_j] - \ldots - g_{j+1} [A_{j+1}, A_j] + O(g^2) \\
&= A_j - g_n \sum_{i=1}^{n} \Gamma_{nj}^i A_i - \ldots - g_{j+1} \sum_{i=1}^{n} \Gamma_{(j+1)j}^i A_i + O(g^2) \\
&= A_j - \sum_{i=1}^{n} (g_n \Gamma_{nj}^i + \ldots + g_{j+1} \Gamma_{(j+1)j}^i) A_i + O(g^2) \\
&= A_j - \sum_{i=1}^{n} \left(\sum_{k=j+1}^{n} g_k \Gamma_{kj}^i \right) A_i + O(g^2).
\end{aligned}$$

Let $O(g_i{}^2)$ indicate terms of the form $O(g_i{}^r)$, $r \geq 2$. Then $O(g^2)$ indicates terms of the form $O(g^r)$, for $r \geq 2$ where $O(g^r)$, $r = 1, 2, \ldots$, are linear in terms of the form $g_{i_1} g_{i_2} \cdots g_{i_r}$, $i_\nu \in \{1, \ldots, n\}$, $\nu = 1, \ldots, r$. So

$$\xi_{ij}(g) = \begin{cases} -\sum_{k=j+1}^{n} g_k \Gamma_{kj}^i + O(g^2), & i \neq j \\ 1 - \sum_{k=j+1}^{n} g_k \Gamma_{kj}^i + O(g^2), & i = j \end{cases}$$

Thus, $M(g)^{-1} = I - \tilde{\xi}(g) + O(g^2)$ where the ijth element of ξ is

$$\tilde{\xi}_{ij} = \sum_{k=j+1}^{n} g_k \Gamma_{kj}^{i}. \tag{23}$$

As a consequence we have that (for $\|g\|$ small)

$$M(g) = I + \tilde{\xi}(g) + O(g^2), \tag{24}$$

and in particular M is analytic in g.

In the following lemma we show the special form of the second-order average equation associated with the Wei-Norman equations (7) and prove second-order averaging results for the Wei-Norman parameters using classical averaging theory.

Lemma 2. Let $\epsilon > 0$. Let $D = \{g \in \Re^n \mid \|g\| < r\} \subset W$. Assume that $u(t) \in \Re^n$ is periodic in t with period $T > 0$ and has continuous derivatives up to third order for $t \in [0, \infty)$. Suppose that $u_{av} = 0$. Let $g(t, \epsilon)$ be the solution to (7) with $g(0, \epsilon) = g_0 \in D$ such that $\|g_0\| = O(\epsilon)$. Let $\bar{z}(t, \epsilon)$ be the solution to the equations

$$\dot{\bar{z}}_k = \frac{\epsilon^2}{T} \sum_{i,j=1; i<j}^{n} Area_{ij}(T) \Gamma_{ij}^{k}, \quad \bar{z}_k(0, \epsilon) = 0, \quad k = 1, \ldots, n. \tag{25}$$

Define $\bar{\bar{g}}(t, \epsilon)$ by

$$\bar{\bar{g}} = \bar{z} + \epsilon \tilde{u} + \bar{\bar{g}}_0. \tag{26}$$

Then $\bar{\bar{g}}(0, \epsilon) = \bar{\bar{g}}_0$ and if $(\bar{z}(t, \epsilon) + \bar{\bar{g}}_0) \in D, \ \forall t \in [0, b/\epsilon]$ and $\|g(0, \epsilon) - \bar{\bar{g}}(0, \epsilon)\| = \|g_0 - \bar{\bar{g}}_0\| = O(\epsilon^2)$ then

$$\|g(t, \epsilon) - \bar{\bar{g}}(t, \epsilon)\| = O(\epsilon^2) \text{ on } [0, b/\epsilon],$$

$$\text{i.e., } \|g(t, \epsilon) - \bar{\bar{g}}(t, \epsilon)\| \leq k\epsilon^2, \quad \forall \epsilon \in [0, \epsilon_1), \quad \forall t \in [0, b/\epsilon]$$

for some $k > 0$ and $\epsilon_1 > 0$.

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. Consider the change of variables

$$g = y + \epsilon v(t,y). \tag{27}$$

and note that $y(0) = g_0 = O(\epsilon)$. Differentiating (27) with respect to time and substituting (7) for $\dot{g}$ and (27) for g gives

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

But by (24) and linearity of $M(g)$ in terms $O(g^r)$ we see that

$$\begin{aligned} M(y+\epsilon v) &= I + \tilde{\xi}(y) + O(y^2) + \epsilon\tilde{\xi}(v) + \epsilon O(yv) + \epsilon^2 O(v^2 + yv^2 + y^2v^2) + h.o.t \\ &= M(y) + \epsilon\tilde{\xi}(v) + \epsilon O(yv) + \epsilon^2 p_1(t,y,\epsilon) \end{aligned}$$

where $O(yv)$ is linear in terms of the form $y_i v_j$ with $i,j \in \{1,\ldots,n\}$ and higher order terms are defined similarly. $p_1(t,y,\epsilon)$ is periodic in t with period T and continuous with continuous derivatives up to third order with respect to all its arguments for all $(t,y,\epsilon) \in [0,\infty) \times D \times [0,\epsilon_0]$, $\epsilon_0 > 0$. Now

$$\frac{\partial v}{\partial y}(t,y) = \int_0^t \frac{\partial h}{\partial y}(\tau,y)d\tau,$$

so $\partial v/\partial y$ is periodic since $\partial h/\partial y$ is periodic in t of period T with zero average. Thus, $\partial v/\partial y$ is bounded for all $(t,y) \in [0,\infty) \times D$. So for small enough $\epsilon > 0$, $(I + \epsilon\partial v/\partial y)$ is nonsingular and we can write

$$\begin{aligned} \dot{y} &= (I - \epsilon p_2(t,y,\epsilon))(\epsilon^2\tilde{\xi}(v)u + \epsilon^2 O(yv)u + \epsilon^3 p_1(t,y,\epsilon)u) \\ &= \epsilon^2\tilde{\xi}(v)u + \epsilon^2 O(yv)u + \epsilon^3 p_3(t,y,\epsilon) \end{aligned}$$

where $p_2(t,y,\epsilon)$ and $p_3(t,y,\epsilon)$ are periodic in t with period T and continuous with continuous derivatives up to third order with respect to all their arguments on $[0,\infty) \times D \times [0,\epsilon_0]$.

Now let $s = \epsilon t$, then

$$\frac{dy}{ds} = \epsilon\tilde{\xi}(v)u + \epsilon O(yv)u + \epsilon^2 p_3(\frac{s}{\epsilon}, y, \epsilon). \tag{28}$$

Thus, by standard perturbation theory (c.f. Theorem 7.1 of [19]), the nominal solution to (28) is $y_0(s) = 0 \in D$, $\forall s \geq 0$ and so $\exists \epsilon^* > 0 \ \ni \ \forall |\epsilon| < \epsilon^*$ (28) has the unique solution $y(s, \epsilon)$ defined on $[0, b]$ such that

$$\|y(s, \epsilon) - \epsilon y_1(s, \epsilon)\| = O(\epsilon^2)$$

where b can be chosen arbitrarily large. Here $y_1(s, \epsilon)$ is the solution to

$$\begin{aligned} \frac{dy_1}{ds} &= \tilde{\xi}(v(t, y))u\big|_{y=y_0=0} \\ &= \tilde{\xi}(M(0)\tilde{u})u \\ &= \tilde{\xi}(\tilde{u})u, \quad \epsilon y_1(0, \epsilon) = g_0. \end{aligned}$$

Thus, $\|y(t, \epsilon) - \epsilon y_1(t, \epsilon)\| = O(\epsilon^2)$ on $[0, b/\epsilon]$ and $y_1(t, \epsilon)$ satisfies

$$\dot{y}_1 = \epsilon\tilde{\xi}(\tilde{u})u, \quad \epsilon y_1(0, \epsilon) = g_0.$$

Now let $z \triangleq \epsilon y_1$ then $\|y(t, \epsilon) - z(t, \epsilon)\| = O(\epsilon^2)$ on $[0, b/\epsilon]$ and

$$\dot{z} = \epsilon^2\tilde{\xi}(\tilde{u})u, \quad z(0, \epsilon) = g_0.$$

Let $\bar{z}(t, \epsilon)$ be the solution to

$$\dot{\bar{z}} = \frac{\epsilon^2}{T}\int_0^T \tilde{\xi}(\tilde{u}(\sigma))u(\sigma)d\sigma \triangleq C\epsilon^2, \quad \bar{z}(0, \epsilon) = \bar{\bar{g}}_0. \tag{29}$$

So by classical averaging (Theorem 7.4 [19]), if $\bar{z} \in D$, $\forall t \in [0, b/\epsilon]$ and ϵ is small enough then $\|z(t, \epsilon) - \bar{z}(t, \epsilon)\| = O(\epsilon^2)$ on $[0, b/\epsilon]$. Then by the triangle inequality, $\|y(t, \epsilon) - \bar{z}(t, \epsilon)\| = O(\epsilon^2)$ on $[0, b/\epsilon]$.

Again let $s = \epsilon t$. Since

$$\frac{d\bar{z}}{ds} = C\epsilon, \quad \bar{z}(0, \epsilon) = \bar{\bar{g}}_0$$

$\bar{z}(s) = C\epsilon s + \bar{\bar{g}}_0$ and so for small enough ϵ if $\bar{z}(s) \in D$, $\forall s \in [0, b]$, $\|y(s) - (C\epsilon s + \bar{\bar{g}}_0)\| = O(\epsilon^2)$ on $[0, b]$. This, (23), (24) and the fact that $\bar{\bar{g}}_0 = O(\epsilon)$ imply that $\forall s \in [0, b]$,

$$\begin{aligned} \epsilon v(s, y(s)) &= \epsilon M(y(s))\tilde{u}(s) \\ &= \epsilon(I + \tilde{\xi}(y(s)) + O(y^2))\tilde{u}(s) \\ &= \epsilon(I + \tilde{\xi}(C\epsilon s + O(\epsilon)) + O(\epsilon^2))\tilde{u}(s) \\ &= \epsilon\tilde{u}(s) + O(\epsilon^2). \end{aligned}$$

So $\|\epsilon v(t,y) - \epsilon\tilde{u}(t)\| = O(\epsilon^2)$, $\forall t \in [0, b/\epsilon]$. Thus, if ϵ is small enough and $\bar{z}(t) \in D$, $\forall t \in [0, b/\epsilon]$ then by (27) and the triangle inequality

$$\begin{aligned} \|g(t) - (\bar{z}(t) + \epsilon\tilde{u}(t))\| &= \|y(t) + \epsilon v(t,y) - (\bar{z}(t) + \epsilon\tilde{u}(t))\| \\ &\leq \|y(t) - \bar{z}(t)\| + \|\epsilon v(t,y) - \epsilon\tilde{u}(t))\| \\ &= O(\epsilon^2) \end{aligned}$$

The proof is complete if we let $\bar{\bar{g}} = \bar{z} + \epsilon\tilde{u}$ and show that $\dot{\bar{z}}$ defined by (29) is equivalent to (25).

To do so let $f = [f_1\ f_2\ \cdots\ f_n]^T = \tilde{\xi}(\tilde{u})u$. Then from (23),

$$f_k = \sum_{i=1}^{n} \xi_{ki}(\tilde{u})u_i = \sum_{i=1}^{n}\sum_{j=i+1}^{n} \Gamma_{ji}^k \tilde{u}_j u_i.$$

If we define $f_{av} = [f_{av1}\ f_{av2}\ \cdots\ f_{avn}]^T = (1/T)\int_0^T \tilde{\xi}(\tilde{u}(\sigma))u(\sigma)d\sigma$ then by (29) $\dot{\bar{z}}_k = \epsilon^2 f_{avk}$ and using integration by parts and (21) we get

$$\begin{aligned} f_{avk} &= \frac{1}{T}\sum_{i=1}^{n}\sum_{j=i+1}^{n}\int_0^T \Gamma_{ji}^k \tilde{u}_j(\sigma)\dot{\tilde{u}}_i(\sigma)d\sigma \\ &= \frac{1}{2T}\sum_{i=1}^{n}\sum_{j=i+1}^{n}\int_0^T \Gamma_{ji}^k(\dot{\tilde{u}}_i(\sigma)\tilde{u}_j(\sigma) - \tilde{u}_i(\sigma)\dot{\tilde{u}}_j(\sigma))d\sigma \\ &= \frac{1}{T}\sum_{i=1}^{n}\sum_{j=i+1}^{n} \Gamma_{ji}^k Area_{ji}(T) \\ &= \frac{1}{T}\sum_{i,j=1;i<j}^{n} \Gamma_{ji}^k Area_{ji}(T) \\ &= \frac{1}{T}\sum_{i,j=1;i<j}^{n} \Gamma_{ij}^k Area_{ij}(T) \end{aligned}$$

which completes the proof. $\square$

The form of (25) shows that $\bar{z}$ is a linear function of time where the proportionality constant depends only on ϵ, the period T, the projected areas $Area_{ij}(T)$ bounded by the closed curves described by $\tilde{u}_i$ and $\tilde{u}_j$ over one period, and the structure constants Γ_{ij}^k. We show further in the next lemma that the structure constants are directly related to the Lie brackets of the vector fields defined by the columns of $M(g)$ evaluated at $g = 0$.

Lemma 3. Suppose that $\bar{z}(t)$ is the solution to (25). Let $[f_1\ f_2\ \cdots\ f_n] = M(g)$. Then

$$\dot{\bar{z}} = \frac{\epsilon^2}{T} \sum_{i,j=1;i<j}^{n} Area_{ij}(T)[f_i, f_j]|_{g=0}, \quad \bar{z}(0,\epsilon) = 0. \tag{30}$$

Proof. By (23) and (24) we have that

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

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

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

So for $i < j$,

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

which by (25) completes the proof. $\square$

Using Lemma 2 we can now state the main theorem for second-order averaging on the group level.

Theorem 3 (Area Rule). Let $\epsilon > 0$. Let $D = \{g \in \Re^n \mid \|g\| < r\} \subset S$. Assume that $u(t) \in \Re^n$ is periodic in t with period $T > 0$ and has continuous derivatives up to third order for $t \in [0, \infty)$. Suppose that $u_{av} = 0$. Let $X(t)$ be the solution to (3) represented by (6) where $g(t, \epsilon)$ is the solution to (7) with $g(0, \epsilon) = g_0 \in D$ such that $\Psi(g_0) = X(0)$ and $\|g_0\| = O(\epsilon)$. Define

$$\bar{z}_k(t, \epsilon) = \frac{\epsilon^2 t}{T} \sum_{i,j=1; i<j}^{n} Area_{ij}(T)\Gamma_{ij}^k, \quad k = 1, \ldots, n, \tag{31}$$

$$\bar{\bar{g}} = \bar{z} + \epsilon\tilde{u} + \bar{\bar{g}}_0, \tag{32}$$

$$\bar{\bar{X}}(t) = e^{\bar{\bar{g}}_1(t)A_1} e^{\bar{\bar{g}}_2(t)A_2} \cdots e^{\bar{\bar{g}}_n(t)A_n}, \tag{33}$$

where $Area_{ij}(T)$ and Γ_{ij}^k are as in (21) and (22), respectively, and $\|g_0 - \bar{\bar{g}}_0\| = O(\epsilon^2)$. If $(\bar{z}(t, \epsilon) + \bar{\bar{g}}_0) \in D$, $\forall t \in [0, b/\epsilon]$ then

$$\tilde{d}(X(t), \bar{\bar{X}}(t)) = O(\epsilon^2), \quad \text{on } [0, b/\epsilon],$$

$$\text{i.e.,} \quad \tilde{d}(X(t), \bar{\bar{X}}(t)) \leq k\epsilon^2, \quad \forall \epsilon \in [0, \epsilon_1), \quad \forall t \in [0, b/\epsilon]$$

for some $k > 0$ and $\epsilon_1 > 0$.

Proof. The theorem follows from Lemma 2 and the definition of $\tilde{d}$. □

Remark 1. Theorem 3 gives the formula for the second-order approximation $\bar{\bar{X}}(t)$ to the solution $X(t)$ of (3) assuming that the initial condition $X(0)$ is close to the identity. However, because system (3) is left-invariant, the approximation can be generalized for any other initial condition. Let $X_I(t)$ and $\bar{\bar{X}}_I(t)$ correspond to the actual and approximate solutions, respectively, of (3) with $X_I(0) = I$. Now suppose we wish to find the second-order approximation $\bar{\bar{X}}(t)$ to the solution $X(t)$ of (3) with $X(0) \in G$. Then by left invariance of system (3), $X(t) = X(0)X_I(t)$. It is then easily observed that $\bar{\bar{X}}(t) = X(0)\bar{\bar{X}}_I(t)$ is an $O(\epsilon^2)$ approximation of $X(t)$ on an $O(1/\epsilon)$ time interval since $\bar{\bar{X}}^{-1}(t)X(t) = \bar{\bar{X}}_I^{-1}(t)X_I(t)$.

Thus, the formula of Theorem 3 is useful for specifying open loop controls to get from any initial condition $X(0)$ to any final condition $X(t_f)$, since this reduces to the problem of specifying controls to get from $X_I(0) = I$ to $X_I(t_f) = X^{-1}(0)X(t_f)$. Nonetheless, the fact that the formula of Theorem 3 is valid not just for $X(0) = I$ but for $X(0)$ in a neighborhood of the identity has significant practical advantages. For example, it may be necessary or desired to get from $X(0)$ to $X(t_f)$ in steps, i.e., by specifying different controls during different time intervals of $[0, t_f]$ to reach intermediate goal points. In general, after each interval the control algorithm would have to be restarted, i.e., the next goal point would have to be recomputed relative to the new initial condition. However, if $X(t_f)$ is close enough to the identity then according to Theorem 3 the restarts can be avoided and the control specified according to the original goal point computation. An example of an algorithm for specification of open loop controls in steps without restarts is given in Section 5.2. It should be noted that the analogous area rule of Section 4 for the single exponential representation of $X(t)$ is only valid for $X(0) = I$. In this case, when controls are specified in steps, restart computations will be unavoidable. $\square$

Remark 2. Since $\bar{z}(t)$ is linear in time, it is straightforward to choose controls $u(t)$ such that $(\bar{z}(t) + \bar{\bar{g}}_0) \in D$, $\forall t \in [0, b/\epsilon]$. One simply ensures that $\|\bar{z}(b/\epsilon) + \bar{\bar{g}}_0\| < r$ where

$$\bar{z}_k(b/\epsilon) = \frac{\epsilon b}{T} \sum_{i,j=1;i<j}^{n} Area_{ij}(T)\Gamma_{ij}^k, \quad k = 1, \ldots, n.$$

Further, this justifies calling Theorem 3 an "Area Rule" (c.f. [11] for the area rule on flat spaces that inspired the present result). $\square$

Remark 3. The general solution to system (3) with $X(0) = I$ is given by the Peano-Baker series as

$$X(t) = I + \epsilon \int_0^t U(\sigma)d\sigma + \epsilon^2 \int_0^t \int_0^{\sigma_1} U(\sigma_2)U(\sigma_1)d\sigma_2 d\sigma_1 + \ldots.$$

If we truncate the series and define

$$Y(t) = I + \epsilon \int_0^t U(\sigma)d\sigma + \epsilon^2 \int_0^t \int_0^{\sigma_1} U(\sigma_2)U(\sigma_1)d\sigma_2 d\sigma_1,$$

then $Y(t)$ is not necessarily a curve in the matrix Lie group G. Instead we consider $Y(t)$ as a curve in $\Re^{n^2}$ and the solution to

$$\dot{Y}(t) = \epsilon U(t) + \epsilon^2 \int_0^t U(\sigma)U(t)d\sigma, \quad Y(0) = I.$$

If we treat both $X(t)$ and $Y(t)$ as curves in $\Re^{n^2}$ then it can be shown that $Y(t)$ is an $O(\epsilon^2)$ estimate of $X(t)$ on an $O(1/\epsilon)$ time interval. Additionally, since Y evolves in Euclidean space we can find an average solution $\bar{Y}$ that approximates Y using classical averaging theory. If we define $U_{av} = \sum_{i=1}^{n} A_i u_{avi}(t)$ and $\tilde{U}(t) = \sum_{i=1}^{n} A_i \tilde{u}_i(t)$ then for the assumptions of Theorem 3, $U_{av} = 0$ and $\bar{Y}(t)$ can be defined as the solution to

$$
\begin{aligned}
\dot{\bar{Y}} &= \frac{\epsilon^2}{T}\int_0^T \tilde{U}(\sigma)U(\sigma)d\sigma \\
&= \frac{\epsilon^2}{T}\int_0^T \tilde{U}(\sigma)\dot{\tilde{U}}(\sigma)d\sigma \\
&= \frac{\epsilon^2}{2T}\int_0^T (\tilde{U}(\sigma)\dot{\tilde{U}}(\sigma) - \dot{\tilde{U}}(\sigma)\tilde{U}(\sigma))d\sigma \\
&= \frac{\epsilon^2}{2T}\int_0^T [\tilde{U}, \dot{\tilde{U}}](\sigma)d\sigma, \quad \bar{Y}(0) = I,
\end{aligned}
$$

using integration by parts. Thus, since

$$
\begin{aligned}
\frac{1}{2}\int_0^T [\tilde{U}, \dot{\tilde{U}}](\sigma)d\sigma &= \frac{1}{2}\int_0^T [\sum_{i=1}^{n} \tilde{u}_i(\sigma)A_i, \sum_{j=1}^{n} \dot{\tilde{u}}_j(\sigma)A_j]d\sigma \\
&= \frac{1}{2}\int_0^T \sum_{i=1}^{n}\sum_{j=1}^{n} \tilde{u}_i(\sigma)\dot{\tilde{u}}_j(\sigma)[A_i, A_j]d\sigma \\
&= \frac{1}{2}\int_0^T \sum_{i,j=1;i<j}^{n} (\tilde{u}_i\dot{\tilde{u}}_j - \dot{\tilde{u}}_i\tilde{u}_j)(\sigma)[A_i, A_j]d\sigma \\
&= \sum_{i,j=1;i<j}^{n} Area_{ij}[A_i, A_j] \\
&= \sum_{k=1}^{n}(\sum_{i,j=1;i<j}^{n} Area_{ij}\Gamma_{ij}^{k})A_k, \qquad (34)
\end{aligned}
$$

then

$$
\bar{Y}(t) = I + \frac{\epsilon^2 t}{T}\sum_{k=1}^{n}(\sum_{i,j=1;i<j}^{n} Area_{ij}\Gamma_{ij}^{k})A_k. \qquad (35)
$$

It can then be shown using classical averaging theory that $\bar{Y}(t) + \epsilon\tilde{U}(t)$ is an $O(\epsilon^2)$ approximation of $X(t)$ on an $O(1/\epsilon)$ time interval. It is interesting to check that if we consider $\bar{\bar{X}}(t)$ with $\bar{\bar{X}}(0) = I$ as a series expansion, its truncation to first-order in the A_k's is identical to $\bar{Y}(t) + \epsilon\tilde{U}(t)$. That is,

$$
\begin{aligned}
\bar{\bar{X}}(t) &= e^{(\bar{z}_1+\epsilon\tilde{u}_1)A_1} \cdots e^{(\bar{z}_n+\epsilon\tilde{u}_n)A_n} \\
&= (I + (\bar{z}_1 + \epsilon\tilde{u}_1)A_1 + \ldots) \cdots (I + (\bar{z}_n + \epsilon\tilde{u}_n)A_n + \ldots)
\end{aligned}
$$

$$\begin{aligned} &\approx I + \sum_{k=1}^{n} \bar{z}_k A_k + \sum_{k=1}^{n} \epsilon \tilde{u}_k A_k \\ &= I + \frac{\epsilon^2 t}{T} \sum_{k=1}^{n} \Big(\sum_{i,j=1; i<j}^{n} Area_{ij} \Gamma_{ij}^{k} \Big) A_k + \epsilon \tilde{U}(t). \end{aligned} \tag{36}$$

This confirms that the *answer* given by the area rule is basis independent. □

Corollary 2. Let $\epsilon > 0$. Let $D = \{g \in \Re^3 \mid \|g\| < r \text{ and } |g_2| < \frac{\pi}{2} - \delta,\ \delta > 0\}$. Suppose $u(t) \in \Re^3$ is periodic in t with period $T > 0$ with continuous derivatives up to third order $\forall t \in [0, \infty)$ and let $u_{av} = 0$. Let $G = SO(3)$ and let $X(t)$ be the solution to (3) represented by (6) where $g(t,\epsilon)$ is the solution to the Wei-Norman equations for $SO(3)$ as given by (15) with $g(0,\epsilon) = g_0$ such that $\Phi(g_0) = X(0)$ and $\|g_0\| = O(\epsilon)$. Let $\bar{z}(t,\epsilon)$ be defined by

$$\bar{z}(t,\epsilon) = \frac{\epsilon^2 t}{T} \begin{bmatrix} Area_{23}(T) \\ Area_{31}(T) \\ Area_{12}(T) \end{bmatrix},$$

and define

$$\bar{\bar{X}}(t) = e^{(\bar{z}_1 + \epsilon \tilde{u}_1 + \bar{\bar{g}}_{01}) A_1} e^{(\bar{z}_2 + \epsilon \tilde{u}_2 + \bar{\bar{g}}_{02}) A_2} e^{(\bar{z}_3 + \epsilon \tilde{u}_3 + \bar{\bar{g}}_{03}) A_3}$$

where $A_1 = \hat{e}_1,\ A_2 = \hat{e}_2,\ A_3 = \hat{e}_3$ and $\bar{\bar{g}}_0 = (\bar{\bar{g}}_{01}, \bar{\bar{g}}_{02}, \bar{\bar{g}}_{03})^T$ is such that $\|g_0 - \bar{\bar{g}}_0\| = O(\epsilon^2)$. If $|(\epsilon b/T) Area_{31}(T) + \bar{\bar{g}}_{02}| < \pi/2 - \delta$ then $\forall \epsilon \in [0, \epsilon_1)$ and $\forall t \in [0, b/\epsilon]$,

$$\phi(\bar{\bar{X}}^{-1}(t) X(t)) \leq k\epsilon^2$$

$$\text{and} \quad \|\bar{\bar{X}}(t)^{-1} X(t) - I\|_1 \leq 9|cos(k\epsilon^2) - 1| + 7|sin(k\epsilon^2)| + |sin(k\epsilon^2)|^2$$

for some $k > 0$ and $\epsilon_1 > 0$.

Proof. The structure constants for $\mathcal{G} = so(3)$ can be computed as ${\Gamma_{12}}^3 = {\Gamma_{23}}^1 = {\Gamma_{31}}^2 = 1$, since $[A_1, A_2] = A_3,\ [A_2, A_3] = A_1,\ [A_3, A_1] = A_2$. Thus, the bounds follow from Theorem 3 and the computations of Corollary 1 (see [18] for details). □

Remark 4. It is clear that if we choose $u(t)$ such that $Area_{31} = 0$ then b can be chosen arbitrarily large. □

Remark 5. In their work [4], Wilczek and Shapere use a formula analogous to (35) to estimate the displacement achieved by a paramecium through shape changes in a low Reynolds number limit. The present paper provides a justification via averaging theory for their formula. □

4 Averaging Using Single Exponential Representation

As an alternative to using the Wei-Norman representation of solutions to (3), we consider defining first and second-order averages of solutions to (3) using Magnus' single exponential representation [14]. By Theorem III of [14] under an unspecified condition of convergence, the solution to (3) with $X(0) = I$ can be expressed as

$$X(t) = e^{Z(t)} \tag{37}$$

where $Z(t) \in \mathcal{G}$ is given by the infinite series (we show terms up to $O(\epsilon^3)$):

$$\begin{aligned} Z(t) &= \epsilon \int_0^t U(\tau)d\tau + \frac{\epsilon^2}{2}\int_0^t [\tilde{U}(\tau), U(\tau)]d\tau \\ &+ \frac{\epsilon^3}{4}\int_0^t [\int_0^\tau [\tilde{U}(\sigma), U(\sigma)]d\sigma, U(\tau)]d\tau + \frac{\epsilon^3}{12}\int_0^t [\tilde{U}(\tau), [\tilde{U}(\tau), U(\tau)]]d\tau + \ldots \end{aligned} \tag{38}$$

While the convergence criterion for (38) is not given explicitly in [14], two different sufficient conditions are provided in [21] and [16], respectively. Karasev and Mosolova [21] show that (38) converges if

$$\int_0^t \|ad_{\epsilon U(\tau)}\| d\tau < ln2. \tag{39}$$

For G a finite-dimensional Lie group, the convergence condition (39) is equivalent to

$$\int_0^t \|\Lambda(\epsilon u(\tau))\| d\tau < ln2, \tag{40}$$

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

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

In the case that $G = SO(3)$ and $A_i = \hat{e}_i$ is the basis for $\mathcal{G} = so(3)$, it is easy to compute that $\Lambda(\epsilon u) = \epsilon U$ and so (40) is equivalent to

$$\int_0^t \|\epsilon U(\tau)\| d\tau < ln2. \tag{41}$$

The convergence criterion given by Fomenko and Chakon [16] takes the form

$$\int_0^t \|\epsilon U(\tau)\| d\tau < \frac{\hat{b}}{M}, \tag{42}$$

where $M \geq 1$ is defined such that $\|[A, B]\| \leq M\|A\|\|B\|$ for all $A, B \in \mathcal{G}$ and $\hat{b}$ is the constant radius of a disk over which a scalar differential equation, defined in [16], is analytic.

Satisfying the convergence criterion in any one of the above forms means limiting the length of time of validity of the single exponential representation (37). In this section as in others we express our results over time intervals that depend on the constant b, e.g., approximations are given over $O(1/\epsilon)$ time intervals of the form $[0, b/\epsilon]$. Thus, it is the constant b that we limit to meet the convergence condition.

Now assuming the convergence requirement is met, suppose we differentiate Z as defined by (38) with respect to t, then $Z(t)$ is the solution to

$$\dot{Z} = \epsilon U + \frac{\epsilon^2}{2}[\tilde{U}, U] + \frac{\epsilon^3}{4}[\int_0^t [\tilde{U}(\tau), U(\tau)] d\tau, U] + \frac{\epsilon^3}{12}[\tilde{U}, [\tilde{U}, U]] + \ldots, \quad Z(0) = 0. \tag{43}$$

Now since $\mathcal{G}$ is an n-dimensional vector space, we can identify it with $\Re^n$ and then use classical averaging theory to derive an average approximation of the solution $Z(t)$ to (43).

Lemma 4. Let $\epsilon > 0$. Let $D = \{Z \in \mathcal{G} | \|Z\| < r\}$. Assume that $u(t) \in \Re^n$ is periodic in t with period $T > 0$ and has continuous derivatives up to third order for $t \in [0, \infty)$. Let $b > 0$ be such that the convergence requirement for (38) is met $\forall t \in [0, b/\epsilon]$. Let $Z(t, \epsilon) \in \mathcal{G}$ be the solution to (43). Let $\bar{Z}(t, \epsilon)$ be the solution to

$$\dot{\bar{Z}} = \epsilon U_{av}, \quad \bar{Z}(0, \epsilon) = 0, \tag{44}$$

and let $\bar{\bar{Z}}(t,\epsilon)$ be the solution to

$$\dot{\bar{\bar{Z}}} = \epsilon U_{av} + \frac{\epsilon^2}{2T}\int_0^T [\tilde{U}(\tau), U(\tau)]d\tau, \quad \bar{\bar{Z}}(0,\epsilon) = 0. \tag{45}$$

Define

$$\begin{aligned}
\hat{Z} &= \bar{\bar{Z}} + \epsilon(\tilde{U} - U_{av}t) \\
&= \frac{\epsilon^2 t}{2T}\int_0^T [\tilde{U}(\tau), U(\tau)]d\tau + \epsilon\tilde{U}.
\end{aligned} \tag{46}$$

If $\bar{Z}(t,\epsilon) \in D, \ \forall t \in [0, b/\epsilon]$ then

$$\|Z(t,\epsilon) - \bar{Z}(t,\epsilon)\| = O(\epsilon) \text{ on } [0, b/\epsilon],$$

and if $\bar{\bar{Z}}(t,\epsilon) \in D, \ \forall t \in [0, b/\epsilon]$ then

$$\|Z(t,\epsilon) - \hat{Z}(t,\epsilon)\| = O(\epsilon^2) \text{ on } [0, b/\epsilon].$$

Proof. Let $s = \epsilon t$. Then

$$\frac{dZ}{ds} = U + \frac{\epsilon}{2}[\tilde{U}, U] + \frac{\epsilon^2}{4}[\int_0^s [\tilde{U}(\tau), U(\tau)]d\tau, U] + \frac{\epsilon^2}{12}[\tilde{U}, [\tilde{U}, U]] + \ldots \tag{47}$$

Let $Z_0(s,\epsilon)$ be the solution to

$$\frac{dZ_0}{ds} = U(s), \quad Z_0(0,\epsilon) = 0 \tag{48}$$

and $Z_1(s,\epsilon)$ the solution to

$$\frac{dZ_1}{ds} = \frac{1}{2}[\tilde{U}, U](s), \quad Z_1(0,\epsilon) = 0. \tag{49}$$

By standard perturbation theory (c.f. Theorem 7.1 of [19]), if $Z_0(s,\epsilon) \in D, \ \forall s \in [0, b]$, then $\exists \epsilon^* > 0 \ni \forall |\epsilon| < \epsilon^*$ (47) has the unique solution $Z(s,\epsilon)$ defined on $[0, b]$ such that

$$\|Z(s,\epsilon) - Z_0(s,\epsilon)\| = O(\epsilon), \quad \forall s \in [0, b],$$

$$\|Z(s,\epsilon) - (Z_0(s,\epsilon) + \epsilon Z_1(s,\epsilon))\| = O(\epsilon^2), \quad \forall s \in [0, b].$$

This implies that

$$\|Z(t,\epsilon) - \tilde{Z}_0(t,\epsilon)\| = O(\epsilon), \quad \forall t \in [0, b/\epsilon],$$

$$\|Z(t,\epsilon)-\tilde{Z}_1(t,\epsilon)\| = O(\epsilon^2), \quad \forall t \in [0, b/\epsilon].$$

where

$$\dot{\tilde{Z}}_0 = \epsilon U, \quad \tilde{Z}_0(0,\epsilon) = 0, \tag{50}$$

$$\dot{\tilde{Z}}_1 = \epsilon U + \frac{\epsilon^2}{2}[\tilde{U}, U], \quad \tilde{Z}_1(0,\epsilon) = 0. \tag{51}$$

The average equations associated with (50) and (51) are equations (44) and (45), respectively. So using classical averaging theory, if $\bar{Z}(t,\epsilon) \in D$, $\forall t \in [0, b/\epsilon]$ then $\|\tilde{Z}_0(t,\epsilon) - \bar{Z}(t,\epsilon)\| = O(\epsilon)$ on $[0, b/\epsilon]$ (c.f. Theorem 7.4 [19]) and if $\bar{\bar{Z}}(t,\epsilon) \in D, \forall t \in [0, b/\epsilon]$ then $\|\tilde{Z}_1(t,\epsilon) - \hat{Z}(t,\epsilon)\| = O(\epsilon^2)$ (following the steps of the proof for Lemma 2). The lemma follows from the triangle inequality. □

Let $\hat{\Phi} : \mathcal{G} \to G$ define the mapping

$$\hat{\Phi}(Z) = e^Z. \tag{52}$$

Let $\hat{S}$ be the largest neighborhood of $0 \in \mathcal{G}$ such that $\hat{\Psi} \equiv \hat{\Phi}|_{\hat{S}} : \hat{S} \to G$ is one-to-one. Let $\hat{Q} \equiv \hat{\Psi}(\hat{S}) \subset G$. Then $\hat{\Psi} : \hat{S} \to \hat{Q}$ is a diffeomorphism and we can define a metric $\hat{d} : \hat{Q} \times \hat{Q} \to \Re_+$ by

$$\hat{d}(X,Y) = d(\hat{\Psi}^{-1}(X), \hat{\Psi}^{-1}(Y)) \tag{53}$$

where d is given by (10). We can now state the area rule using the single exponential representation (analogous to Theorem 3).

Theorem 4 (Single Exponential Area Rule). Let $\epsilon > 0$. Let $D = \{Z \in \mathcal{G} | \|Z\| < r\} \subset \hat{S}$. Assume that $u(t) \in \Re^n$ is periodic in t with period $T > 0$ and has continuous derivatives up to third order for $t \in [0, \infty)$. Let $b > 0$ be such that the convergence requirement for (38) is met $\forall t \in [0, b/\epsilon]$. Let $X(t)$ be the solution to (3), with $X(0) = I$, represented by the single exponential (37) where $Z(t,\epsilon) \in \mathcal{G}$ is the solution to (43). Define

$$\bar{Z}(t,\epsilon) = \epsilon U_{av} t, \tag{54}$$

$$\bar{X}_S(t) = e^{\bar{Z}(t)} = e^{\epsilon U_{av} t}. \tag{55}$$

If $\bar{Z}(t,\epsilon) \in D, \ \forall t \in [0, b/\epsilon]$ then

$$\hat{d}(X(t), \bar{X}_S(t)) = O(\epsilon) \text{ on } [0, b/\epsilon].$$

Now suppose additionally that $u_{av} = 0$. Define

$$\bar{\bar{Z}}(t,\epsilon) = \frac{\epsilon^2 t}{T} \sum_{k=1}^{n} (\sum_{i,j=1;i<j}^{n} Area_{ij}(T) \Gamma_{ij}^k) A_k, \tag{56}$$

$$\bar{\bar{X}}_S(t) = e^{\bar{\bar{Z}} + \epsilon \bar{U}}. \tag{57}$$

If $\bar{\bar{Z}}(t,\epsilon) \in D, \ \forall t \in [0, b/\epsilon]$ then

$$\hat{d}(X(t), \bar{\bar{X}}_S(t)) = O(\epsilon^2) \text{ on } [0, b/\epsilon].$$

Proof. The theorem follows from (34), Lemma 4, and the definition of $\hat{d}$. □

Remark 6. The first-order approximations derived using the single exponential representation above and the product of exponential representation (Section 2) are identical, i.e., $\bar{X}_S(t) = \bar{X}(t) = e^{\epsilon U_{av} t}$. A comparison in the second-order case (for $X(0) = I$) shows that $\bar{\bar{X}}_S(t)$ is equal to $\bar{\bar{X}}(t)$ collapsed into a single exponential, i.e.,

$$\bar{\bar{X}}_S(t) = e^{\sum_{k=1}^{n} \bar{\bar{g}}_k(t) A_k}$$

where $\bar{\bar{g}}_k(t)$ are defined by (32). In other words, if we consider $\bar{\bar{X}}(t)$ and $\bar{\bar{X}}_S(t)$ as series expansions then they agree to first order in the A_k's as defined in Remark 3 by (36).

Remark 7. The formulae in Theorem 4 are clearly basis independent.

5 Geometric Interpretation and Control

5.1 First-Order Averaging

The first-order average approximation $\bar{X}(t) = \bar{X}(0) e^{\epsilon U_{av} t}$ of the solution $X(t)$ to (3) exhibits the contribution of the dc component of the periodic control input. This is clear by noting that $\bar{X}(t)$ is the solution to (12) which is identical to (3) except that $U(t)$ is replaced with $U_{av} = (1/T)\int_0^T U(\tau)d\tau$. As a consequence, the first-order approximation alone will not capture the complete controllability picture when the number of controls, m, is less than the dimension of the system, n. Nonetheless, the first-order average result can still be useful for open loop control, particularly if $m = n$. For instance, suppose we are given the problem of specifying periodic open loop controls to drive the solution $X(t)$ of system (3) from a point $Y_i \in G$ at $t = 0$ to a point $Y_f \in G$ at $t = t_f$ with $O(\epsilon)$ accuracy. Specifically, let $G = SO(3)$ and consider the rigid spacecraft control problem. Then Y_i, $Y_f \in SO(3)$ are some specified initial and final orientations of the spacecraft, respectively. From the form of $\bar{X}(t)$ we see that $\bar{X}(t)$ is a trajectory about a fixed axis. So by Euler's theorem it is easy to see that any attitude Y_f can be reached from any attitude Y_i in one step, i.e., with a single choice of U_{av} such that $\bar{X}(0) = Y_i$ and $\bar{X}(t_f) = \bar{X}(0)e^{\epsilon U_{av} t_f} = Y_f$. For instance suppose $Y_i = I$ and Y_f is given such that y_{ij} is the ijth element of Y_f. Then we can find the Euler parametrization of Y_f:

$$Y_f = e^{-\phi\hat{c}} = e^{\epsilon\hat{u}_{av}t_f} \tag{58}$$

where

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

tr is the trace operator and

$$c = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \frac{1}{2sin\phi}\begin{bmatrix} y_{23} - y_{32} \\ y_{31} - y_{13} \\ y_{12} - y_{21} \end{bmatrix}.$$

From this it is clear that we want $\epsilon u_{av} = -\phi c/t_f$. Thus, we could choose our periodic control $\epsilon u(t)$ such that its average is ϵu_{av}. To ensure $O(\epsilon)$ accuracy of the resulting trajectory $X(t)$, however, it may be important to consider reaching Y_f in steps, i.e., by choosing intermediate points as goals along the way. In such a scheme, intermittent feedback control could be used to provide corrections to the open loop control.

5.2 Second-Order Averaging

From Theorem 3 we see that $\bar{\bar{X}}$ provides an $O(\epsilon^2)$ approximation to the solution X of (3) on an $O(1/\epsilon)$ time interval. Further, $\bar{\bar{X}}$ is defined as a product of exponentials where each exponent has a secular term (linear in t) and an $O(\epsilon)$ periodic term. The proportionality constant of the secular term depends only on ϵ, the period T, the projected areas $Area_{ij}(T)$ bounded by the closed curves described by the pairs of components $\tilde{u}_i$ and $\tilde{u}_j$ over one period, and the structure constants Γ^k_{ij} associated with the choice of basis in the Lie algebra $\mathcal{G}$.

This second-order average approximation, in providing more information about the actual solution than the first-order approximation, captures the effect of the group level version of depth-one Lie brackets of vector fields. This means that if the appropriate Lie algebra rank condition is satisfied for a system of the form (3) using only depth-one Lie brackets, then the second-order approximation provides a formula for achieving complete constructive controllability using periodic controls. For example, consider again the spacecraft attitude control problem where $G = SO(3)$. In this case $n = 3$ and suppose there are only $m = 2$ controls available. System (3) becomes

$$\dot{X} = \epsilon X A_1 u_1 + \epsilon X A_2 u_2 \stackrel{\Delta}{=} \epsilon F_1(X) u_1 + \epsilon F_2(X) u_2, \quad X(0) = I, \tag{59}$$

where $A_1 = \hat{e}_1$ and $A_2 = \hat{e}_2$ as in previous sections and F_1 and F_2 are left-invariant vector fields on G. Now, the Lie bracket of left-invariant vector fields on a matrix Lie group can be expressed in terms of the Lie bracket on the associated Lie algebra as $[F_1(X), F_2(X)] = [XA_1, XA_2] = X[A_1, A_2]$. Since $[A_1, A_2] = A_3 = \hat{e}_3$ and $\{A_1, A_2, A_3\}$ forms a basis for $\mathcal{G} = so(3)$, then this system is completely controllable with the Lie algebra rank condition only requiring one depth-one Lie bracket. Thus, the second-order approximation should reveal how to constructively achieve complete control of the spacecraft attitude with only two controls. This is illustrated by the formula for the second-order approximation $\bar{\bar{X}}$ of the solution X to (59) (from Corollary 2)

$$\bar{\bar{X}}(t) \;=\; e^{\bar{\bar{g}}_1 A_1} e^{\bar{\bar{g}}_2 A_2} e^{\bar{\bar{g}}_3 A_3}$$

$$= e^{(\epsilon\tilde{u}_1+\bar{\bar{g}}_{01})A_1} e^{(\epsilon\tilde{u}_2+\bar{\bar{g}}_{02})A_2} e^{(\frac{\epsilon^2 t}{T}Area_{12}(T)+\bar{\bar{g}}_{03})A_3} \qquad (60)$$

which has three nonzero rotational components. That is, oscillations $\tilde{u}_1$ and $\tilde{u}_2$ about their respective two axes produce drift about the third axis, proportional to the area defined by the closed curve described by $\tilde{u}_1$ and $\tilde{u}_2$ over one period.

The second-order approximation formula, e.g., equation (60) in the above example, can then be used to specify open loop control. For the spacecraft attitude control problem, we note again that the Wei-Norman parameters (g_1, g_2, g_3) correspond to a type of Euler angles (c.f. [18]). As a result, given a desired spacecraft orientation $Y_f \in SO(3)$ at time $t = t_f$, we can compute the corresponding desired Euler angles $g_d = (g_{d1}, g_{d2}, g_{d3})^T$. Additionally, we suppose that it is desired to reach the new spacecraft orientation Y_f and stay there, i.e., we would like $\dot{g} = 0$ at $t = t_f$. For the above example, assuming that $X(0) = Y_i = I$, it is clear that we want to choose ϵu_1 and ϵu_2, meeting the constraints of Corollary 2, such that $\bar{\bar{g}}(t_f) = g_d$ and $\epsilon u_1(t_f) = \epsilon u_2(t_f) = 0$. By Corollary 2, we could then conclude that $g(t_f)$ will be $O(\epsilon^2)$ close to g_d with $\dot{g}(t_f) = 0$.

One approach to meeting these requirements is to select controls ϵu_1 and ϵu_2 in steps. Specifically, we divide the time period $[0, t_f]$ into four intervals $[0, t_1]$, $(t_1, t_2]$, $(t_2, t_3]$ and $(t_3, t_f]$, choose controls on each of these intervals and apply Corollary 2 separately to the system over each interval. We assume that $\|g_d\| = O(\epsilon)$ which makes it easy to ensure that the initial condition of g for each of the intervals (i.e., $g(0)$, $g(t_1)$, $g(t_2)$ and $g(t_3)$) will be $O(\epsilon)$ as required in Corollary 2 (i.e., we avoid restart computations as discussed in Remark 1). We also let $\bar{\bar{g}}(0) = g(0)$. Then after application of Corollary 2 to the first interval, we get that $\|g(t_1) - \bar{\bar{g}}(t_1)\| = O(\epsilon^2)$. But this is the necessary initial condition on $\bar{\bar{g}}$ for the second interval. So we can apply Corollary 2 to the second interval which will yield $\|g(t_2) - \bar{\bar{g}}(t_2)\| = O(\epsilon^2)$. Similarly we apply Corollary 2 to the third and fourth intervals to yield $\|g(t_f) - \bar{\bar{g}}(t_f)\| = O(\epsilon^2)$. Also for practical reasons we choose ϵu_1 and ϵu_2 to be continuous throughout, i.e., $\forall t \in [0, t_f]$.

The basic idea for choosing the controls on each of the four intervals comes directly from the geometric interpretation of (60). In the first interval, we select ϵu_1 and ϵu_2 to meet the requirements of the first and second rotational components only, i.e., such that $\bar{\bar{g}}_1(t_1) = \epsilon\tilde{u}_1(t_1) = g_{d1}$, $\bar{\bar{g}}_2(t_1) = \epsilon\tilde{u}_2(t_1) = g_{d2}$ and $Area_{12}(T) = 0$. This last requirement can be satisfied simply by choosing ϵu_1 and ϵu_2 in phase over the first interval. Next during the second interval we choose ϵu_1 and ϵu_2 so that the third rotational component requirement is met without a net effect on the first two components. That is, we choose ϵu_1 and ϵu_2 out of phase such that $Area_{12}(T)$ satisfies

$$\bar{\bar{g}}_3(t_2) = \frac{\epsilon^2}{T}(t_2 - t_1)Area_{12}(T) = g_{d3}.$$

To avoid a net effect on $\bar{\bar{g}}_1$ and $\bar{\bar{g}}_2$, the length of this second time interval should correspond to an integral number q of periods T, i.e., $t_2 - t_1 = qT$. (Note that there is flexibility, subject to the above constraints, in the choice of t_1, t_2, t_3 and T as well as the phase difference between ϵu_1 and ϵu_2 which determines $Area_{12}(T)$).

At the end of the second time interval we will have $\bar{\bar{g}}(t_2) = g_d$. However, $\dot{g}$ will not necessarily be zero since $\epsilon u_1(t_2)$ and $\epsilon u_2(t_2)$ will not necessarily be zero. Thus, in the third and fourth intervals we bring ϵu_1 and ϵu_2 to zero with zero net change to $\bar{\bar{g}}$. It is easy to avoid affecting $\bar{\bar{g}}_3$; this is done by choosing ϵu_1 and ϵu_2 in phase over the third and fourth intervals, i.e., so that $Area_{12}(T) = 0$. To avoid zero net change to $\bar{\bar{g}}_1$ and $\bar{\bar{g}}_2$, we select ϵu_1 and ϵu_2 so that

$$\epsilon\int_{t_2}^{t_3} u_i(\tau)d\tau + \epsilon\int_{t_3}^{t_f} u_i(\tau)d\tau = 0, \quad i = 1,2. \tag{61}$$

In particular, in the third interval we bring ϵu_1 and ϵu_2 to zero and in the fourth interval we bring ϵu_1 and ϵu_2 again to zero while satisfying (61).

The following control sequence employs the above strategy for the system described by (59) with corresponding second-order approximation (60):

$$\epsilon u_1(t) = \begin{cases} g_{d1}\omega sin\omega t & 0 \le t \le t_1 = \frac{\pi}{2\omega} = \frac{T}{4} \\ g_{d1}\omega cos(\omega(t - t_1)) & t_1 < t \le t_2 = t_1 + qT \\ g_{d1}\omega cos(\omega(t - t_2)) & t_2 < t \le t_3 = t_2 + \frac{T}{4} \\ -\frac{g_{d1}}{2}\omega sin(\omega(t - t_3)) & t_3 < t \le t_f = t_3 + \frac{T}{2} \end{cases} \tag{62}$$

$$\epsilon u_2(t) = \begin{cases} g_{d2}\omega sin\omega t & 0 \le t \le t_1 \\ \frac{g_{d2}}{cos(\psi)}\omega cos(\omega(t-t_1)-\psi) & t_1 < t \le t_2 \\ g_{d2}\omega cos(\omega(t-t_2)) & t_2 < t \le t_3 \\ -\frac{g_{d2}}{2}\omega sin(\omega(t-t_3)) & t_3 < t \le t_f \end{cases} \tag{63}$$

where $0 < \psi < \pi/2$ is selected such that q is a positive integer and

$$q = \frac{g_{d3}}{\pi g_{d1} g_{d2} tan(\psi)}.$$

The period of oscillation T and frequency ω can be computed as

$$T = \frac{t_f}{q+1},$$

$$\omega = \frac{2\pi}{T}.$$

The corresponding integrals of these controls are

$$\epsilon \tilde{u}_1(t) = \begin{cases} g_{d1}(1-cos\omega t) & 0 \le t \le t_1 \\ g_{d1}(1+sin(\omega(t-t_1))) & t_1 < t \le t_2 \\ g_{d1}(1+sin(\omega(t-t_2))) & t_2 < t \le t_3 \\ g_{d1}(\frac{3}{2}+\frac{1}{2}cos(\omega(t-t_3))) & t_3 < t \le t_f \end{cases} \tag{64}$$

$$\epsilon \tilde{u}_2(t) = \begin{cases} g_{d2}(1-cos\omega t) & 0 \le t \le t_1 \\ \frac{g_{d2}}{cos(\psi)}(cos(\psi)+sin(\psi)+sin(\omega(t-t_1)-\psi)) & t_1 < t \le t_2 \\ g_{d2}(1+sin(\omega(t-t_2))) & t_2 < t \le t_3 \\ g_{d2}(\frac{3}{2}+\frac{1}{2}cos(\omega(t-t_3))) & t_3 < t \le t_f \end{cases} \tag{65}$$

$Area_{12}(T)$ is only nonzero during the second time interval, i.e., for $t \in (t_1, t_2]$. During this time,

$$Area_{12}(T) = \frac{\pi g_{d1} g_{d2} tan(\psi)}{\epsilon^2}.$$

Let the parameters be selected as $\epsilon = 0.1$, $g_{d1} = 0.1$, $g_{d2} = 0.05$, $g_{d3} = \pi/40$, $\psi = \pi/4$, $t_f = 24$ and, consequently, $q = 5$, $T = 4$, $\omega = \pi/2$, $t_1 = 1$, $t_2 = 21$ and $t_3 = 22$. Figure 2 shows plots of the corresponding controls ϵu_1 and ϵu_2 as a function of time. Note that these controls are continuous and $\epsilon u_1(t_f) = \epsilon u_2(t_f) = \epsilon u_1(0) = \epsilon u_2(0) = 0$.

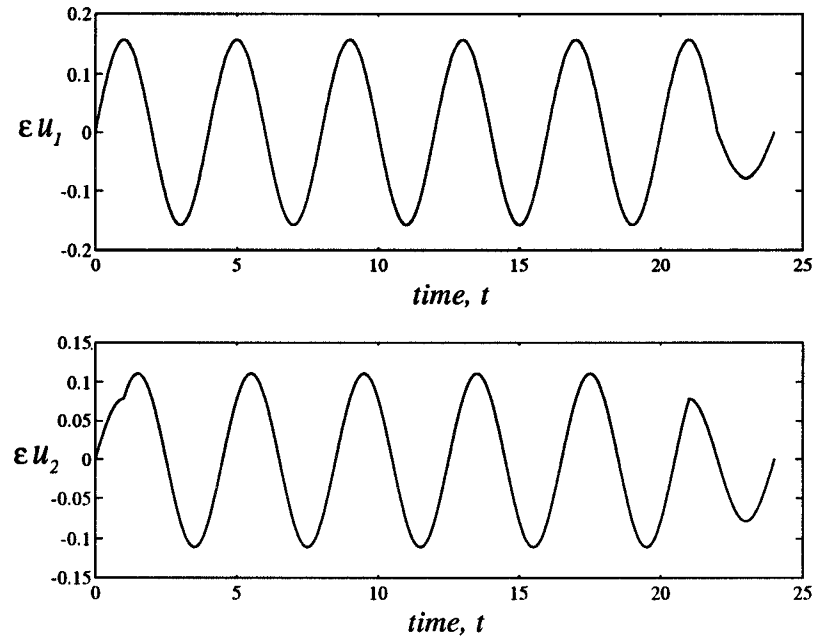

Figure 2: Control Input Signals for Attitude Control Example.

Figures 3(a) and 3(b) show the corresponding plots of $\bar{\bar{g}}_1$ and $\bar{\bar{g}}_2$, respectively (solid lines). Figure 3(c) is a plot of $\tilde{u}_1$ versus $\tilde{u}_2$ and the area enclosed by the curve is equivalent to $Area_{12}(T)$ for $t \in (t_1, t_2]$. Note that for $t \in [0, t_1] \cup (t_2, t_f]$, $\tilde{u}_1$ and $\tilde{u}_2$ are in phase so no area is traced out over this time interval. The trajectory of $\bar{\bar{g}}_3$ is shown in Figure 3(d). It can be seen that $\bar{\bar{g}}_3$ changes only during the second interval, i.e., for $t \in (t_1, t_2]$ when $Area_{12}(T) \neq 0$. Since $(\bar{\bar{g}}_1, \bar{\bar{g}}_2, \bar{\bar{g}}_3)$ reach (g_{d1}, g_{d2}, g_{d3}) at $t = t_f$, $\bar{\bar{X}}(t_f) = Y_f$. By Corollary 2 this means that the controls defined by (62) and (63) should drive the actual solution $X(t)$ of (59) to an $O(\epsilon^2)$ neighborhood of Y_f. To verify this for the above example we have superimposed plots of g_1, g_2 and g_3 (computed by simulation for the controls defined by (62) and (63)) on the plots of $\bar{\bar{g}}_1, \bar{\bar{g}}_2$ and $\bar{\bar{g}}_3$ in Figures 3(a), 3(b) and 3(d), respectively. These plots are the dashed lines in these figures; however, they are difficult to see in Figures 3(a) and 3(b) because of the high accuracy of the second-order approximation. Certainly, the average solution can be observed to be an $O(\epsilon^2)$ approximation of the actual solution.

As an alternative to using the product of exponential second-order approximation $\bar{\bar{X}}$

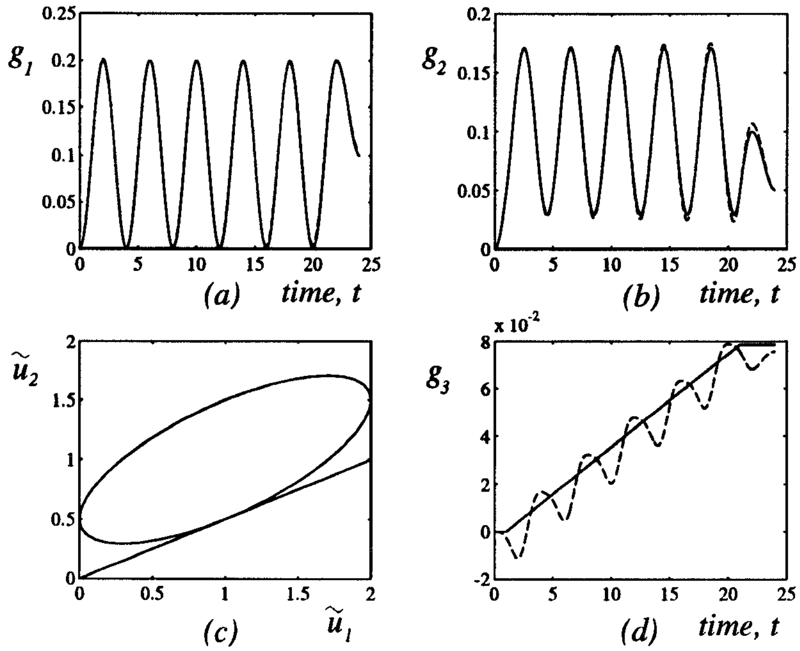

Figure 3: Average (solid lines) and Actual (dashed lines) Wei-Norman Parameters for Attitude Control Example.

to specify open loop controls, we could use the single exponential second-order approximation $\bar{\bar{X}}_S$ from Section 4 which for the two-input spacecraft attitude control example of this section takes the form

$$\bar{\bar{X}}_S(t) = e^{\bar{\bar{Z}}+\epsilon\tilde{U}} = e^{\epsilon\tilde{u}_1 A_1+\epsilon\tilde{u}_2 A_2+\frac{\epsilon^2 t}{T} Area_{12}(T) A_3} \tag{66}$$

Like equation (60) this formula illustrates how to achieve complete controllability since the coefficients of A_1, A_2 and A_3 are all nonzero. In this case, however, instead of trying to match desired Wei-Norman parameters (g_{d1}, g_{d2}, g_{d3}) for a given desired orientation Y_f, we would need to compute and match desired Euler parameters. (Recall the Euler parametrization (58) of Section 5.1.)

In either case to ensure a high order of accuracy we may want to consider reaching Y_f in steps and incorporating feedback corrections as suggested in Section 5.1 for first-order averaging.

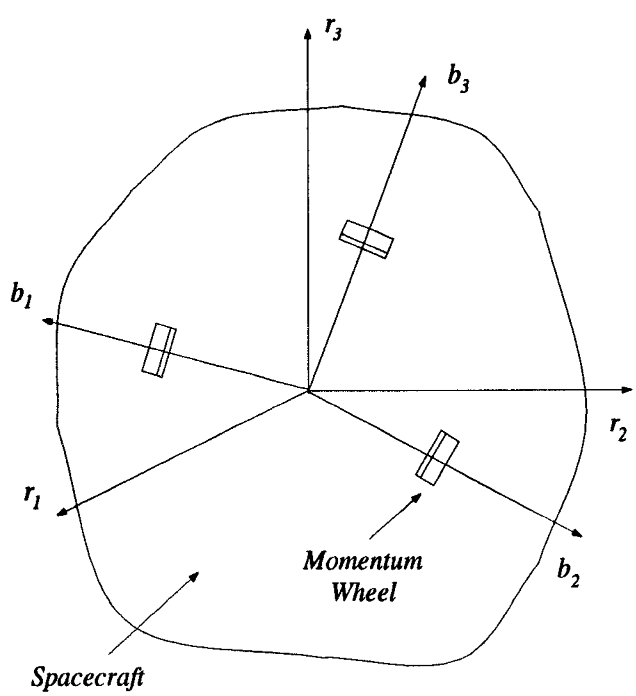

Figure 4: Rigid Spacecraft with Momentum Wheels.

6 Approximation Formulae for Example Actuators

System (3) with $G = SO(3)$ describes the kinematics of a spacecraft in the case that zero angular momentum of the spacecraft is conserved, i.e., there is no external torque applied to the spacecraft. Consider the sketch of Figure 4 which represents the spacecraft (assumed to be a rigid body). Let (r_1, r_2, r_3) be coordinates fixed on the body and let (e_1, e_2, e_3) be inertial coordinates (not shown). Then we define $X(t) \in SO(3)$ such that $e_i = X(t)r_i$, i.e., $X(t)$ describes the attitude of the spacecraft at time t. $X(t)$ satisfies

$$\dot{X} = X\hat{\Omega}, \quad \hat{\Omega}(t) = \sum_{i=1}^{3} \Omega_i(t) A_i \tag{67}$$

where $\Omega = (\Omega_1, \Omega_2, \Omega_3)^T$ is the angular velocity of the spacecraft in body-fixed coordinates and $\{A_1, A_2, A_3\}$ is the basis for $\mathcal{G} = so(3)$ as defined in Section 2.

Equation (67) is in the form of (3) assuming that we can interpret $\Omega(t)$ as the control input, i.e., that we can identify $\epsilon u(t) \equiv \Omega(t)$. This will be the case if zero angular momentum of the spacecraft is conserved. Under this condition there are a variety of practical means of actuation, i.e., a variety of ways of effecting the body angular velocity of the spacecraft as desired. We explore two cases, the first using momentum wheels and the second using a point mass oscillator appended to the spacecraft. Momentum wheels are a typical means of controlling spacecraft attitude. An appended oscillator, on the other hand, would be a novel approach. However, such an approach could potentially provide a lighter weight and less expensive control alternative as would be required, for example, in the case of small planetary spacecraft control described in Section 1. Additionally, by investigating how to control a spacecraft with appended oscillators, we are illustrating as well how to determine the effect of undesirable oscillations or vibrations, e.g., as caused by flexible attachments on the spacecraft.

In this section we derive the $O(\epsilon^2)$ approximation to $X(t)$ explicitly for the momentum wheel case and for the appended point mass oscillator case.

6.1 Momentum Wheels

The spacecraft with a maximum of three momentum wheels is illustrated in Figure 4. Let $p \leq 3$ be the number of wheels. We make the following assumptions about the wheels and the spacecraft:

1. The ith wheel spins about the axis b_i (unit vector) which is fixed in the spacecraft such that the center of mass of the ith wheel lies on the b_i axis. Further, b_i is a principal axis for the ith wheel, and the ith wheel is symmetric about b_i.

2. Let μ be the angular momentum of the spacecraft measured in inertial coordinates. Then $\mu = 0$ is conserved.

We next make the following definitions:

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

$j_i \triangleq$ moment of inertia of ith wheel about b_i axis.

$k_i \triangleq$ moment of inertia of ith wheel about any axis orthogonal to b_i.

$R_i \triangleq [\alpha_i\ \beta_i\ b_i] \in SO(3)$.

$\bar{J}_i \triangleq$ inertia matrix of ith wheel measured in body-fixed coordinates, i.e., $\bar{J}_i = R_i\ \mathrm{diag}(k_i, k_i, j_i)R_i^T$, where $\mathrm{diag}(a, b, c)$ is a diagonal matrix with diagonal elements (a, b, c).

$\nu_i \triangleq \dot{\theta}_i b_i =$ angular velocity of ith wheel in body-fixed coordinates.

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

Then by [20, 22]

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

which implies that

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

Then by assumption 2, $\mu = 0$, so

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

where $J \triangleq J^* + \sum_{i=1}^{p} \bar{J}_i$ is assumed to be nonsingular (which it will be in general). If we let $k \triangleq \sum_{i=1}^{p} k_i$ and $l_i \triangleq j_i - k_i$, $i = 1, \ldots, p$, then it is easy to show that $J = J^* + kI + \sum_{i=1}^{p} l_i b_i b_i^T$. Now since

$$\bar{J}_i b_i = R_i \begin{bmatrix} k_i & 0 & 0 \\ 0 & k_i & 0 \\ 0 & 0 & j_i \end{bmatrix} R_i^T R_i \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = j_i b_i$$

then

$$\Omega = -J^{-1}\sum_{i=1}^{p} j_i b_i \dot{\theta}_i \triangleq C\dot{\theta}. \tag{68}$$

Equation (68) allows us to exactly compute body angular velocity as a function of the angular speeds of the momentum wheels. Assuming that the $\dot{\theta}_i$, $i = 1, \ldots, p$ are small-amplitude, zero-mean periodic functions of time then we can rewrite equation (68) as

$$\epsilon u = \epsilon\Omega = \epsilon C\dot{\theta} \tag{69}$$

where we have redefined Ω and $\dot{\theta}$ by $\epsilon\Omega$ and $\epsilon\dot{\theta}$, respectively. Then, from Corollary 2, the $O(\epsilon^2)$ approximation (on an $O(1/\epsilon)$ time interval) $\bar{\bar{X}}(t)$ to the solution $X(t)$ of $\dot{X} = \epsilon X\hat{\Omega}$, with $X(0)$ close to the identity is given by

$$\bar{\bar{X}}(t) = e^{(\frac{\epsilon^2 t}{T}Area_{23}(T)+\epsilon\tilde{u}_1+\bar{\bar{g}}_{01})A_1}e^{(\frac{\epsilon^2 t}{T}Area_{31}(T)+\epsilon\tilde{u}_2+\bar{\bar{g}}_{02})A_2}e^{(\frac{\epsilon^2 t}{T}Area_{12}(T)+\epsilon\tilde{u}_3+\bar{\bar{g}}_{03})A_3}$$

where $Area_{ij}$ are functions of the $\tilde{u}_i(t)$, $i = 1, 2, 3$, as given by (21). We can then express $\bar{\bar{X}}(t)$ explicitly in terms of our physical controls $\dot{\theta}(t)$. Let $\theta(t) \triangleq \int_0^t \dot{\theta}(\tau)d\tau$ then $\epsilon\tilde{u}(t) = \epsilon C\theta(t)$. Now, define

$$area_{ij}(T) = \frac{1}{2}\int_0^T (\theta_i(\sigma)\dot{\theta}_j(\sigma) - \theta_j(\sigma)\dot{\theta}_i(\sigma))d\sigma,$$

which is the area bounded by the closed curve described by θ_i and θ_j over one period. Let c_{ij} be the ijth element of C (if necessary, add extra columns of zeros to make C a 3×3 matrix). Then

$$Area_{12}(T) = \begin{vmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{vmatrix} area_{12}(T) + \begin{vmatrix} c_{12} & c_{13} \\ c_{22} & c_{23} \end{vmatrix} area_{23}(T) - \begin{vmatrix} c_{11} & c_{13} \\ c_{21} & c_{23} \end{vmatrix} area_{31}(T),$$

$$Area_{23}(T) = \begin{vmatrix} c_{21} & c_{22} \\ c_{31} & c_{32} \end{vmatrix} area_{12}(T) + \begin{vmatrix} c_{22} & c_{23} \\ c_{32} & c_{33} \end{vmatrix} area_{23}(T) - \begin{vmatrix} c_{21} & c_{23} \\ c_{31} & c_{33} \end{vmatrix} area_{31}(T),$$

$$Area_{31}(T) = -\begin{vmatrix} c_{11} & c_{12} \\ c_{31} & c_{32} \end{vmatrix} area_{12}(T) - \begin{vmatrix} c_{12} & c_{13} \\ c_{32} & c_{33} \end{vmatrix} area_{23}(T) + \begin{vmatrix} c_{11} & c_{13} \\ c_{31} & c_{33} \end{vmatrix} area_{31}(T),$$

where $| \cdot |$ indicates determinant.

In the special case when the wheel axes correspond to the body-fixed coordinate axes, then b_i is the ith standard Euclidean basis vector. Additionally, suppose that the

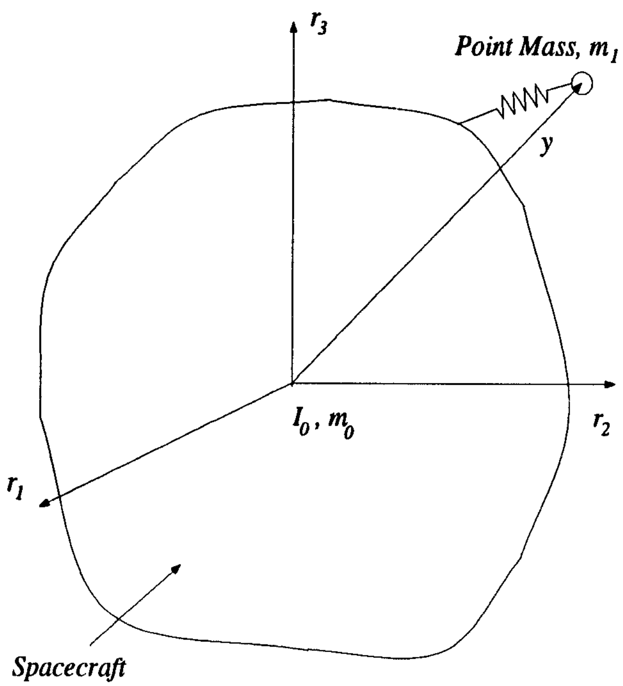

Figure 5: Rigid Spacecraft with Appended Point Mass Oscillator.

body-fixed coordinates have been defined to be the principal axes of the spacecraft. Let $J^* = \mathrm{diag}(i_1, i_2, i_3)$. Then

$$\begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} \Omega_1 \\ \Omega_2 \\ \Omega_3 \end{bmatrix} = - \begin{bmatrix} (j_1/(i_1 + j_1 + k_2 + k_3))\dot{\theta}_1 \\ (j_2/(i_2 + j_2 + k_1 + k_3))\dot{\theta}_2 \\ (j_3/(i_3 + j_3 + k_1 + k_2))\dot{\theta}_3 \end{bmatrix}.$$

In this case if there are only two momentum wheels then the control strategy of Section 5.2 can be used directly to specify $\dot{\theta}_1$ and $\dot{\theta}_2$ for attitude control.

6.2 Appended Point Mass Oscillator

The spacecraft with an appended point mass oscillator is illustrated in Figure 5. The point mass can oscillate in three dimensions. We make the following assumptions about the body plus oscillator:

1. Assume that the point mass has no associated inertia.

2. Let μ be the angular momentum of the spacecraft plus oscillator measured in inertial coordinates. Then $\mu = 0$ is conserved.

3. Let y be the distance of the point mass from the spacecraft center of mass in body-fixed coordinates. For simplicity assume that $y(0) = 0$, i.e., initially the oscillator is located at the center of mass of the spacecraft.

We next make the following definitions:

$m_0 \triangleq$ mass of the spacecraft.

$m_1 \triangleq$ mass of the point mass oscillator.

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

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

$\mathcal{I}_{lock} \triangleq$ instantaneous local 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 (assumed available as control input).

Then by [3]

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

By assumption 2, $\mu = 0$, we get that

$$\Omega = -\mathcal{I}_{lock}^{-1}(y)\alpha_m(y \times \dot{y}), \tag{70}$$

assuming that $\mathcal{I}_{lock}(y)$ is nonsingular (which it will be in general). Suppose that $\dot{y}$ has small amplitude ϵ and is periodic in t of period T and has zero mean. Define $\dot{x}$ by $\dot{y} = \epsilon\dot{x}$

and so $y = \epsilon x$ where $x(t) = \int_0^t \dot{x}(\tau)d\tau$, and x and $\dot{x}$ are periodic in t of period T. Then

$$\begin{aligned}\mathcal{I}_{lock} &= \mathcal{I}_0 + \epsilon^2\alpha_m(\|x\|^2 I - xx^T) \\ &= (I + \epsilon^2\alpha_m(\|x\|^2 I - xx^T)\mathcal{I}_0^{-1})\mathcal{I}_0\end{aligned}$$

So for small enough ϵ,

$$\begin{aligned}\mathcal{I}_{lock}^{-1} &= \mathcal{I}_0^{-1}(I + \epsilon^2\alpha_m(\|x\|^2 I - xx^T)\mathcal{I}_0^{-1})^{-1} \\ &= \mathcal{I}_0^{-1}(I - \epsilon^2\alpha_m(\|x\|^2 I - xx^T)\mathcal{I}_0^{-1} + O(\epsilon^4)) \\ &= \mathcal{I}_0^{-1} - \epsilon^2\alpha_m\mathcal{I}_0^{-1}(\|x\|^2 I - xx^T)\mathcal{I}_0^{-1} + O(\epsilon^4). \qquad (71)\end{aligned}$$

Also

$$y \times \dot{y} = \epsilon^2(x \times x) = \epsilon^2\hat{x}\dot{x}. \qquad (72)$$

Substituting (71) and (72) into (70) gives

$$\begin{aligned}\Omega &= -\alpha_m(\mathcal{I}_0^{-1} - \epsilon^2\alpha_m\mathcal{I}_0^{-1}(\|x\|^2 I - xx^T)\mathcal{I}_0^{-1} + O(\epsilon^4))(\epsilon^2\hat{x}\dot{x}) \\ &= -\epsilon^2\alpha_m\mathcal{I}_0^{-1}\hat{x}\dot{x} + O(\epsilon^4). \qquad (73)\end{aligned}$$

Now let $\bar{\epsilon} \triangleq \epsilon^2$ and suppose we identify $\bar{\epsilon}u$ as the truncation of Ω as follows

$$\bar{\epsilon}u = -\bar{\epsilon}\alpha_m\mathcal{I}_0^{-1}\hat{x}\dot{x}. \qquad (74)$$

Then it is easy to show that the solution to

$$\dot{Y} = \bar{\epsilon}Y\hat{u} \qquad (75)$$

is an $O(\bar{\epsilon})$ approximation on an $O(1/\bar{\epsilon})$ time interval to the solution of

$$\dot{X} = X\hat{\Omega}. \qquad (76)$$

Thus, the $O(\bar{\epsilon})$ *average* approximation to the solution $Y(t)$ of (75) will be an $O(\bar{\epsilon})$ approximation to the solution $X(t)$ of (76) by the triangle inequality. From Corollary 1, the $O(\bar{\epsilon})$ average approximation to the solution $Y(t)$ of (75) on an $O(1/\bar{\epsilon})$ time interval is given by

$$\bar{Y}(t) = \bar{Y}(0)e^{\bar{\epsilon}\hat{u}_{av}t} = \bar{Y}(0)e^{\epsilon^2\hat{u}_{av}t}. \qquad (77)$$

Thus, $\bar{Y}(t)$ is an $O(\epsilon^2)$ approximation of $X(t)$ on an $O(1/\epsilon^2)$ time interval. We then can express $\bar{Y}(t)$ explicitly in terms of our physical controls $\dot{x}$ by computing u_{av} and substituting into (77). By (74)

$$\bar{\epsilon}u = -\bar{\epsilon}\alpha_m \mathcal{I}_0^{-1}\hat{x}\dot{x} = -\epsilon^2 \alpha_m \mathcal{I}_0^{-1} \begin{bmatrix} x_2\dot{x}_3 - x_3\dot{x}_2 \\ x_3\dot{x}_1 - x_1\dot{x}_3 \\ x_1\dot{x}_2 - x_2\dot{x}_1 \end{bmatrix}.$$

Define

$$areax_{ij}(T) = \frac{1}{2}\int_0^T (x_i(\sigma)\dot{x}_j(\sigma) - x_j(\sigma)\dot{x}_i(\sigma))d\sigma.$$

Then

$$u_{av} = -\frac{2\alpha_m}{T}\mathcal{I}_0^{-1} \begin{bmatrix} areax_{23}(T) \\ areax_{31}(T) \\ areax_{12}(T) \end{bmatrix}.$$

Thus, the approximation to the solution of the kinematic equation for the spacecraft plus point mass oscillator is a single exponential with elements proportional to the areas bounded by the closed curves described by x_i and x_j over one period.

7 Acknowledgements

We would like to thank Roger Brockett for useful conversations on the subject of this paper. In particular, his early work [23] has been an inspiration and his paper [24] suggests that averaging may be useful in investigating switched electrical networks. Further, it is a pleasure to acknowledge the helpful and critical comments of W.P. Dayawansa on a draft of this paper.

References

[1] R. M. Murray and S. S. Sastry. Steering nonholonomic systems using sinusoids.

In *Proceedings of the 29th Conference on Decision and Control*, pages 2097–2101, Honolulu, HI, December 1990. IEEE.

[2] 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.

[3] P. S. Krishnaprasad, R. Yang, and W. P. Dayawansa. Control problems on principal bundles and nonholonomic mechanics. In *Proceedings of the 30th Conference on Decision and Control*, pages 1133–1138, Brighton, England, December 1991. IEEE.

[4] A. Shapere and F. Wilczek. Geometry of self-propulsion at low Reynolds number. *Journal of Fluid Mechanics*, 198:557–585, 1989.

[5] R. W. Brockett. Control theory and singular Riemannian geometry. In *New Directions in Applied Mathematics*, pages 13–27. Springer-Verlag, 1982.

[6] 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.

[7] 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.

[8] W. Liu. *Averaging Theorems for Highly Oscillatory Differential Equations and the Approximation of General Paths by Admissible Trajectories for Nonholonomic Systems.* PhD thesis, Rutgers University, New Brunswick, NJ, October 1992.

[9] M. L. Curtis. *Matrix Groups.* Springer-Verlag, 2nd edition, 1984.

[10] R. W. Brockett. Formal language for motion description and map making. In *Proceedings of Symposia in Applied Mathematics*, pages 181–193, Providence, 1990. American Mathematical Society. Volume 41.

[11] R. W. Brockett. On the rectification of vibratory motion. *Sensors and Actuators*, 20(1-2):91–96, 1989.

[12] J.-M. Coron. Global asymptotic stabilization for controllable systems. *Mathematics of Control, Signals and Systems*, 5(3), 1992.

[13] O. J. Sørdalen, M. Dalsmo, and O. Egeland. An exponentially convergent control law for a nonholonomic underwater vehicle. In *Proceedings of the 1993 IEEE International Conference on Robotics and Automation*, Atlanta, GA, May 1993.

[14] W. Magnus. On the exponential solution of differential equations for a linear operator. *Communications on Pure and Applied Mathematics*, VII:649–673, 1954.

[15] 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.

[16] 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.

[17] H. Nijmeijer and A. J. van der Schaft. *Nonlinear Dynamical Control Systems.* Springer-Verlag, New York, 1990.

[18] N. E. Leonard and P. S. Krishnaprasad. Averaging on Lie groups, attitude control and drift. In *Proceedings of the Conference on Information Sciences and Systems*, The Johns Hopkins University, March 1993. (in press).

[19] H. K. Khalil. *Nonlinear Systems.* Macmillan Publishing, New York, 1992.

[20] G. Meyer. Design and global analysis of spacecraft attitude control systems. Technical report TR R-361, NASA, Ames Research Center, March 1971.

[21] M. V. Karasev and M. V. Mosolova. Infinite products and T products of exponentials. *Theoretical and Mathematical Physics*, 28:721–729, 1976.

[22] 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.

[23] R. W. Brockett. System theory on group manifolds and coset spaces. *SIAM Journal of Control*, 10(2):265–284, May 1972.

[24] 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.