A General Averaging Theory via Series Expansions

Patricio A. Vela and Joel W. Burdick

June 18, 2004

Abstract. This paper generalizes averaging theory to arbitrary order by synthesizing series expansion methods for nonlinear time-varying vector fields and their flows with nonlinear Floquet theory. The series expansion methods, based upon the *chronological calculus*, results in series expansions for the flow of time-varying vector fields and also for the approximation of time-varying vector fields by autonomous vector fields. Nonlinear Floquet theory is shown to be the foundation from which the evolution of time-periodic vector fields may be studied via the analysis of autonomous vector fields. For systems too difficult to analyze in closed form using nonlinear Floquet theory, the series expansions naturally allow for the application of perturbation methods. Analysis of the resulting finite series approximation leads to the standard theorems from classical averaging theory. An important distinction is that the averaged vector field has an intrinsic description involving time integrals and Jacobi-Lie brackets of the original time-periodic vector field. Consequently, many known theorems of averaging theory are placed within an intrinsic framework.

1 Introduction

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

This work highlights the use of series expansions for the study of time-varying vector fields and their flows. Series expansions are known to provide alternative expressions for the flows of vector fields and also for autonomous approximations to non-autonomous vector fields. We show that series expansions are an important means for proving results in nonlinear Floquet theory, which places time-periodic vector fields into an autonomous averaged form. For systems that cannot be analyzed in closed form using nonlinear Floquet theory, the use of perturbation methods in conjunction with series expansion methods will lead to approximate averages of m^{th}-order. In the process, we demonstrate that averaging theory is the synthesis of two distinct mathematical theories: Floquet theory and perturbation theory. All known results of classical averaging theory can be better understood when viewed from this perspective. To do so, this paper synthesizes and organizes several fields whose relevant history is summarized below.

Series Expansions. The series expansions used in this paper are rooted in the works of Magnus [27] and Chen [6]. Agračhev and Gamkrelidze [1] made the series expansions more useful for the analysis of nonlinear time-varying differential equations by providing a means to better understand the series' convergence properties and to better formulate the series expansions for systematic computation. The resulting series expansion theory for time-varying nonlinear systems has been termed the *chronological calculus*. The truncated versions of the series expansions have practical use both as an approximation to the true nonlinear flow of a differential equation and as a perturbation method for time-varying vector fields [1].

Within the field of nonlinear control, series expansions have received considerable attention [2, 3]. Kaswki and Sussmann [16, 14, 15, 40] have studied the mathematical structure underlying the series expansions and their connection to controllability and reachability. Controllability analysis and control design were linked by Lafferriere, Liu, and Sussman [21, 41], where extended control systems involving Jacobi-Lie brackets are constructed using series expansions, while ideas similar to averaging are used to prove open-loop trajectory tracking. The extended control systems are based on the work of Kurzweil and Jarník [19, 20], who looked at the nonlinear response of a system to highly oscillatory inputs. The oscillatory inputs depend inversely on the paremeter ϵ, which tends to the limit $\epsilon \to 0$. In the limit, an autonomous system is derived whose construction corresponds to an averaging procedure [25, 26].

The chronological calculus may be utilized to better understand averaging, how the series expansions fit within the method of averaging, and how they may be used to obtain arbitrary orders of approximation. Within the context of vibrational control, Sarychev [37, 38] presented important ideas regarding the use of the chronological calculus and nonlinear Floquet theory as a foundation for averaging theory.

Averaging Theory. Sanders and Verhulst [44] give a detailed treatment of the method of averaging and the related theorems that comprise averaging theory. Of particular note, they give formulas for the average of a time-periodic vector field up to second order, as well as theorems concerning the stability of the flow of time-periodic vector fields with regards to stability of the flow due to the averaged vector field. Bogoliubov and Mitropolsky [5] cover the same topics, and also give a general algorithm for calculating higher-order averages. Guckenheimer and Holmes [10] provide a background on stability analysis of time-periodic differential equations via averaging and *Poincaré maps*. Sarychev [37, 38] lays down the basic elements required for full development of nonlinear Floquet theory, but does not deeply investigate the consequences due to the scope of his investigation.

Bogoliubov and Mitropolsky were aware that the averaged equations of a time-dependent differential equation gave the Poincaré map, thereby allowing for stability analysis, and also posited that their method was able to recover higher orders of averaging by building upon the lower-order averages in a systematic manner. The higher-order methods proposed by Krylov, Bogoliubov, and Mitropolsky (KBM methods) have been studied and extended by several researchers as they are the most general of the known averaging methods. Perko [35] proved the feasibility of higher-order averaged expansions for periodic and quasi-periodic differential equations. Ellison et al. [7] have taken the basic algorithm and developed an improved n^{th}-order estimate for higher-order averaging theory, in the process demonstrating how averaging theory may be used to prove the existence of flows of time-dependent differential equations. Sáenz [36] extended averaging theory to non-periodic systems. Montgomery [30] examined the stability of time-periodic differential equations, by examining the stability of the averaged differential equations. Leung and Zhang [23] have demonstrated the link between higher-order averaging and the Poincaré normal form for oscillatory systems. This paper will demonstrate how the chronological calculus may be used to obtain the above results within an intrinsic framework.

While not immediately relevant to this paper, it should be noted that averaging theory has been extended in many directions. For example, Lehman and Weibel study the use of averaging theory for control of systems with time delay [22]. In Verduyn-Lunel and Hale [12], low-order averaging theory is extended to infinite-dimensional spaces. Averaging theory also includes two-timing methods, which involve fast and slow time scales; only the fast time scale is averaged over. In many cases averaging over multiple dimensions will introduce resonance [45]. The complications due to resonance are often treated separately in the literature, consequently, we leave this topic to a further publication. Other work detailing higher-order averaging theory tends to focus on particular classes of time-periodic systems. In Yagasaki and Ichikawa [46], a computational algorithm is given for weakly nonlinear time-periodic systems. In this paper, a general averaging strategy for time-periodic systems is sought.

Averaging theory in its most general form applies to a larger class of time-dependent vector fields, not necessarily time periodic. Although the analysis will focus on strictly time-periodic vector fields, a large portion of the concepts developed herein do not require this property. For the flows of vector fields, it will be possible to obtain a series expansion representing the flow or the average of the vector field in question. The classical 1^{st}- and 2^{nd}-order averaging theorems are then appropriately truncated versions of the series expansions. Time-periodicity allows for stability analysis using the autonomous averaged vector field alone.

Contribution. This paper synthesizes many of the aforementioned ideas in order to develop an intuitively appealing and easily implementable averaging theory for the analysis of time-dependent (time-periodic) nonlinear ordinary differential equations. The "generalized averaging theory" results from uniting nonlinear Floquet theory with perturbation methods. In doing so, a new approach to averaging theory is developed encompassing the classical results of Sanders and Verhulst [44], and Bogoliubov and Mitropolsky [5]. The approach detailed in this paper contrasts with that found in the literature [5, 35], since there are strict requirements on the vector fields, i.e., smoothness, and the proximity results are conservative. Both can be improved via a more detailed analysis. Nevertheless, the important elements comprising averaging theory are placed within a single framework. Explicit calculations for averages up to 4^{th}-order are given, while a general algorithm and philosophy are given to aid in the calculation of higher orders of averaging.

Organization. Section 2 reviews relevant results from classical averaging theory that are subsequently generalized. Section 3 contains a synopsis of the chronological calculus as derived by Agračhev and Gamkrelidze [1]. It reviews relevant results regarding the exponential representation of flows, and the approximation of flows and vector fields by series expansions. These expansions are key to the development of our generalized averaging theory via nonlinear Floquet theory in Section 4. Section 5 shows how to calculate averages to higher-order and also points out the connection between these formulas and those of classical averaging theory. A special example of averaging via nonlinear Floquet theory is also given, which demonstrates that the averaging process can result in an exact average; a common misunderstanding is that averaging will always result in some quantifiable error. The applicability and exactness of averaging theory is certainly an important issue, and Section 5 also discusses related literature concerning the approximation of time-varying vector fields by autonomous vector fields.

2 Classical Averaging Theory

This section reviews the basic structure and theorems of classical averaging theory that will subsequently be generalized. The standard form of the equations of motion for averaging are

$$\dot{x} = \epsilon X(x,t), \quad x(0) = x_0, \tag{1}$$

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

$$\overline{X(x,t)} \equiv \int_t^{t+T} X(x,\tau)\,\mathrm{d}\tau\,, \tag{2}$$

where the evaluation point x is considered fixed. Often, $\overline{X(\cdot,t)}$ will be written as $\overline{X}(\cdot)$[1]. From the averaged vector field, one can define an autonomous set of equations of motion, termed the "averaged system,"

$$\dot{y} = \epsilon \overline{X}(y), \quad y(0) = x_0. \tag{3}$$

The primary goal of averaging theory is to determine conditions under which the flows of (1) and (3) coincide, and to what degree they coincide. The parameter ϵ provides a means to determine this coincidence[2].

Theorem 1 (first-order averaging) *[44] Consider the initial value problems (1) and (3) with $x, y, x_0 \in M \subset \mathbb{R}^n$, $t \in [t_0, \infty)$, $\epsilon \in (0, \epsilon_0]$. Suppose that the following conditions hold:* **(a)** *$X(x,t)$ is Lipschitz-continuous in x on M, $t \geq 0$, continuous in x and t on $M \times \mathbb{R}^+$, and* **(b)** *$y(t)$ belongs to an interior subset of M on the time-scale $\frac{1}{\epsilon}$. Then, $x(t) - y(t) = O(\epsilon)$ as $\epsilon \downarrow 0$ on the time-scale $\frac{1}{\epsilon}$.*

Theorem 1 determines the conditions under which the flow of the autonomous vector field (3) remains close to the flow of the original vector field (1). Proximity holds on a finite time-scale and few additional results can be derived unless further constraints are placed on the vector fields.

Theorem 2 *[44] Assume that in addition to the conditions of Theorem 1, the following are met:* **(a)** *$y = 0$ is an asymptotically stable fixed point, and* **(b)** *$\overline{X}$ is continuously differentiable in M, and has a domain of attraction $M^* \subset M$. If $x_0 \in M^*$, then $x(t) - y(t) = O(\delta(\epsilon))$, for $0 \leq t < \infty$ and $\delta(\epsilon) = o(1)$.*

An asymptotically stable fixed point in the average (1) renders the averaged approximation valid for all time. For a dynamical system satisfying Theorem 2, the restriction of a flow to a neighborhood of a fixed point implies an orbit.

Theorem 3 *[10] Consider the mapping $y(t) = z(t) + \epsilon w(z(t),t)$, and the initial value problem*

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

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

[2] The proximity and asymptotic results will utilize the order notation defined in [44]: **(a)** $\delta_1(\epsilon) = o(\delta_2(\epsilon))$ if, for $\epsilon \to 0$, $\lim_{\epsilon\to 0} \frac{\delta_1(\epsilon)}{\delta_2(\epsilon)} = 0$, and **(b)** $\delta_1(\epsilon) = O(\delta_2(\epsilon))$ if there exists a k such that $\epsilon \to 0$, $\lim_{\epsilon\to 0} \frac{\delta_1(\epsilon)}{\delta_2(\epsilon)} = k$.

where

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

Suppose that $X(x,t)$ is C^r, $r \geq 2$, and bounded on bounded sets. Then $|x(t) - z(t)| = O(\epsilon)$ on the time-scale $\frac{1}{\epsilon}$, and the following hold: **(a)** *If z^* is a hyperbolic fixed point of (3), then there exists an ϵ_0 such that for all $\epsilon \in (0, \epsilon_0]$, the system (1) possesses a unique hyperbolic periodic orbit $\gamma_\epsilon(t) = z^* + O(\epsilon)$ of the same stability type as z^*, and* **(b)** *if $x^s(t)$ is a solution of (1) lying in the stable manifold of the hyperbolic periodic orbit γ_ϵ, and $z^s(t)$ is a solution of (3) lying in the stable manifold of the hyperbolic fixed point z^*, then $|x^s(t) - z^s(t)| = O(\epsilon)$ for $t \in [0, \infty)$. Similar results apply to solutions lying in the unstable manifolds on the time interval $t \in (-\infty, 0]$.*

Theorem 3 does not preclude the case where the orbit is degenerate, i.e., that it is in fact the point z^*.

Theorem 4 *[17] Assume that the conditions of Theorem 3 hold and that both X and its average $\overline{X}$ share the same equilibrium point. If the equilibrium point is exponentially stable for the averaged system, then the equilibrium point is exponentially stable for the original system.*

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

Theorem 5 (second-order averaging) *[44] Consider the mapping, $y(t) = z(t) + \epsilon v(z(t), t)$ and the initial value problem*

$$\dot{z} = \epsilon \overline{X}(z) + \epsilon^2 \overline{Y}(z), \tag{4}$$

where

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

and $a(x)$ is chosen such that $v(x,t)$ has a vanishing average. If the following conditions hold: **(a)** *$X(x,t)$ has a Lipschitz-continuous first derivative in x and is continuous on its domain of definition, and* **(b)** *$z(t)$ belongs to an interior subset of M on the time-scale $\frac{1}{\epsilon}$, then $x(t) = y(t) + O\left(\epsilon^2\right)$ on the time-scale $\frac{1}{\epsilon}$.*

When moving to second- and higher-order averaging methods, Theorems 2, 3, and 4 are no longer applicable. The explicit ϵ dependence of the higher-order terms complicates matters. Care must be taken to ensure that the corrections obtained from higher-order averages are small perturbations incapable of altering the stability properties inferred from a lower order average [30, 44].

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

$$\overline{Y}(z) = \frac{1}{2}\overline{\left[\int_0^t X(z,\tau)\,\mathrm{d}\tau \,, X(z,t)\right]} + \left[a(z), \overline{X}(z)\right], \tag{5}$$

meaning that the second-order term $\overline{Y}$ in Equation (4) has an intrinsic description within the space of vector fields [43]. Equation (5) points to the fact that averaging may be a natural operation, and to the existence of a natural procedure leading to Equation (5) without the coordinate calculations that have been the basis of prior averaging methods.

Although the KBM method of averaging works to arbitrary order, the averages are coordinate based and do not explicitly involve Jacobi-Lie brackets. We seek an algorithm for the computation of higher-order averages in terms of intrinsic quantities, such as Jacobi-Lie brackets. The analysis should involve Poincaré maps for inference of stability, and recover the known properties of averaged expansions [5, 7, 30, 35, 36].

3 Evolution of Systems and the Chronological Calculus

This section reviews the important concepts of the *chronological calculus* found in Agračhev and Gamkrelidze [1]. The chronological calculus generalizes the exponential representation for flows of linear autonomous differential equations to the setting of nonlinear time-varying differential equations via a series representation of the exponential flow. The series representation forms the basis for a perturbation method for nonlinear dynamical systems. We will use the series expansions and their related approximations to derive a generalized averaging theory.

Consider the differential equation

$$\frac{\mathrm{d}x}{\mathrm{d}t} = X(x,t), \quad x(0) = x_0, \tag{6}$$

where $x \in \mathbb{R}^n$, $t \in \mathbb{R}$, X is smooth as a function of state x only, and is absolutely continuous as a function of time t only. The abbreviation X_t will be used to denote $X(\cdot, t)$. The goal is to understand the flow of (6), denoted by $\Phi^X_{0,t}$. This will require notions of proximity for flows and functions.

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

$$||\varphi||_{s,M} \equiv \sup_{x\in M} \sum_{\alpha=0}^{s} \frac{1}{\alpha!} \sup_{|\chi|=1} |\chi^\alpha \varphi(x)|, \tag{7}$$

where $M \subset \mathbb{R}^n$, $s \in [0, \infty]$ is an arbitrary integer, and χ is a constant vector field. When the second index is absent, then it is assumed to be $M = \mathbb{R}^n$, i.e., $||\varphi||_s = ||\varphi||_{s,\mathbb{R}^n}$. The semi-norm (7) induces the following semi-norm on matrix valued functions $A \in C^\infty(\mathbb{R}^n, \mathbb{R}^{m_1\times m_2})$,

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

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

$$||X||_{s,M} \equiv \max_{1\le\alpha\le n} ||X^\alpha||_{s,M} \quad \text{and}$$

$$||\xi||_{s,M} \equiv \sum_{\beta=1}^{n} ||\xi_\beta||_{s,M},$$

respectively. All topologies, unless specified otherwise, will be defined with respect to the semi-norms $||\cdot||_{s,K}$, where K is an arbitrary compact set in $\mathbb{R}^n$.

The flow of equation (6) satisfies

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

or in pull-back form,

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\Phi^X_{0,t}\right)^\star = \left(\Phi^X_{0,t}\right)^\star \circ X_t, \quad \left(\Phi^X_{\tau,\tau}\right)^\star = \mathrm{Id}^\star, \tag{8}$$

where, on the right-hand side of Equation (8), the pull-back is not of vector-fields but of functions, i.e., $\left(\Phi^X_{0,t}\right)^\star \circ X_t = X_t \circ \Phi^X_{0,t}$.

In [3], Agračhev and Gamkrelidze conceptualize the extension of their results to smooth manifolds $C^\infty(M)$. This is done by providing $C^\infty(M)$ with the topology given to $\mathbb{R}^n$ using the semi-norms. We restrict the current investigation to $\mathbb{R}^n$, however, the ideas naturally extend to systems evolving on manifolds.

3.1 Series Expansions

A solution $\mathcal{V}_{t_0,t}(X_\tau)$ to the differential equation in (8) can be expressed in terms of the Volterra series

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

Under the appropriate conditions, the series summation converges and represents the true flow of the system [1]. Since this paper focuses on smooth vector fields, the convergence conditions are satisfied for finite time. When the vector fields at different points in time commute, i.e., $[X_{t'}, X_{t''}] = 0\,,\ \forall\, t', t'' \in \mathbb{R}$, then the Volterra series expansion of the flow reduces to the time-independent exponential,

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

recovering the classical notion of an exponential for autonomous ordinary differential equations. In the general time-dependent case, the convergent Volterra series is still an exponential series. Agračhev and Gamkrelidze have therefore called this series the *chronological exponential in* X_τ, denoted

$$\overrightarrow{\exp}\left(\int_{t_0}^{t} X_\tau \mathrm{d}\tau\right) \equiv \overrightarrow{\mathcal{V}}_{t_0,t}\,(X_\tau)\,. \tag{9}$$

The chronological exponent $\overrightarrow{\exp}\left(\int_{t_0}^{t} X_\tau \mathrm{d}\tau\right)$ is asymptotically equal to the actual flow, i.e., $\overrightarrow{\exp}\left(\int_{t_0}^{t} X_\tau \mathrm{d}\tau\right) \cong \Phi_{0,t}^{X_t}$, as per Proposition 1 below.

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

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

where

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

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

The constants defined in Proposition 1 will be used in subsequent propositions.

Logarithms of the flow $\Phi_{0,t}^{X_t}$. The inverse of the chronological exponential is called the *chronological logarithm* [1]. It is a mapping from

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

which will be written as

$$\overrightarrow{\log}_{t_0}\overrightarrow{\exp}\left(\int_{t_0}^{t} X_\tau\,\mathrm{d}\tau\right) = X_t, \tag{10}$$

as per [1]. The time-independent logarithm is simply called the *logarithm*, which will be shown to be intimately related to averaging theory. Averaging theory seeks to find an autonomous vector field whose flow approximates the flow of the vector field in Equation (6). This is equivalent to the question posed by Agračhev and Gamkrelidze regarding the existence of a vector field $\overrightarrow{\mathrm{V}}_{t_0,t}\,(X_\tau)$, such that the asymptotic equality holds:

$$\overrightarrow{\exp}\int_{t_0}^{t} X_\tau\,\mathrm{d}\tau \cong \exp\overrightarrow{\mathrm{V}}_{t_0,t}\,(X_\tau)\,, \tag{11}$$

where the asymptotic equality follows Proposition 2. In other words, does there exist an autonomous differential equation, defined by the vector field $\overrightarrow{\mathrm{V}}_{t_0,t}\,(X_\tau)$, whose flow after unit time is equal to the flow of the non-autonomous differential equation given by X_t at time t? A series expansion for this vector field exists, and assuming convergence [1] it is called the *logarithm*, i.e.,

$$\overrightarrow{\mathrm{V}}_{t_0,t}\,(X_\tau) \cong \ln\overrightarrow{\exp}\int_{t_0}^{t} Y_\tau\,\mathrm{d}\tau\,. \tag{12}$$

The vector field $\overrightarrow{\mathrm{V}}_{t_0,t}(X_\tau)$ is an autonomous vector field, but is parametrized by time t. Variation of the final time t will result in a different autonomous vector field.

The logarithm is at the heart of a series expansion solution to the approximation of the flow $\Phi^X_{t_0,t}$. The logarithm vector field is an infinite series consisting of *variations* $\overrightarrow{\mathrm{V}}^{(m)}_{t_0,t}(X_\tau)$,

$$\overrightarrow{\mathrm{V}}_{t_0,t}(X_\tau) = \sum_{m=1}^{\infty} \overrightarrow{\mathrm{V}}^{(m)}_{t_0,t}(X_\tau), \tag{13}$$

where

$$\overrightarrow{\mathrm{V}}^{(m)}_{t_0,t}(X_\tau) \equiv \int_{t_0}^{t}\int_{t_0}^{\tau_1}\cdots\int_{t_0}^{\tau_{m-1}} \mathrm{v}^{(m)}(X_{\tau_1},\ldots,X_{\tau_m})\,\mathrm{d}\tau_m\ldots\,\mathrm{d}\tau_2\,\mathrm{d}\tau_1 \tag{14}$$

is the m^{th}*-variation of the identity flow corresponding to the perturbation field* X_τ[1]. The integrands of the variations in Equation (14) are the sum of iterated Jacobi-Lie brackets, and thereby reside within the space of vector fields. An algorithm to compute the integrands can be found in Section 3.2.

Agračhev and Gamkrelidze use the following proposition to prove the asymptotic equality of Equation (11).

Proposition 2 *[1] If*

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

then

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

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

3.2 Elements of the Series Expansion

Calculation of the integrands of the variations (14) is not simple; Sarychev [37, 38] has an abbreviated discussion of the derivation procedure. The expansion calculation we shall adopt comes from [1]. Alternative methods extract additional structure from the derivation that follows [3, 16, 38].

Define a grammar whose constitutive alphabet is $\{\,\mathrm{ad},\xi_1,\ldots,\xi_m\,\}$.

Definition 1 *[2] A* regular word *in the grammar given by the alphabet* $\{\,\mathrm{ad},\xi_1,\ldots,\xi_m\,\}$ *is a word that, by suitable parenthetication, can be expressed as a commutator polynomial.*

The commutator represents the Jacobi-Lie bracket operator, here given by the letter ad. Note that this definition of a regular word constrains the number of times that the ad letter may appear in a word. An example of a commutator polynomial which is also a word is $\mathrm{ad}\xi_1\mathrm{ad}\xi_2\xi_3 = (\mathrm{ad}\xi_1)((\mathrm{ad}\xi_2)\xi_3) = [\xi_1,[\xi_2,\xi_3]]$. The set consisting of the words of the grammar forms a free associative algebra over $\mathbb{R}$ with generators in the alphabet. The calculation of the logarithm vector field will also require the definition of the depth of a letter in a word.

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

Two examples of regular words, whose regular word subsequences have been delineated by bars, are $\overline{\mathrm{ad}\underline{\mathrm{ad}\xi_3\xi_2}\xi_1}$ and $\overline{\mathrm{ad}\xi_3\underline{\mathrm{ad}\xi_2\xi_1}}$. For the first word, the depths of the letters (ξ_1,ξ_2,ξ_3) are $(1,1,0)$, respectively. For the second word, the depths of the letters (ξ_1,ξ_2,ξ_3) are $(2,0,0)$, respectively.

To this algebra can be added a differentiation operator.

Definition 3 *[2] The differentiation operator $\mathbf{D}(w)$ is a linear operator associated with each word w acting according to the following rules on the alphabet:*

1. $\mathbf{D}(w)\,\mathrm{ad} = w\,\mathrm{ad}$

2. $\mathbf{D}(w)\,\xi_i = w\,\xi_i$

3. $\mathbf{D}(w)\,w_1\,w_2 = (\mathbf{D}(w)\,w_1)\,w_2 + w_1\,(\mathbf{D}(w)\,w_2)$ *(Leibniz rule)*

Each variation of the series, Equation (14), consists of a linear combination of the $(2m-3)!!$ words[3] formed by the operation

$$\mathbf{D}(\mathrm{ad}\xi_m)\,\ldots\,\mathbf{D}(\mathrm{ad}\xi_2)\,\xi_1 = w_1 + \cdots + w_{(2m-3)!!}. \tag{15}$$

Define $\nu_{i\alpha}$ to be the depth of the letter ξ_i in the word w_α from Equation (15). The integrand of the m^{th} variation of the identity flow corresponding to the perturbation field X_t is

$$\mathrm{v}^{(m)}(\xi_1,\ldots,\xi_m) = \sum_{\alpha=1}^{(2m-3)!!}\left(\prod_{i=1}^{m} b_{\nu_{i\alpha}}\right) w_\alpha \tag{16}$$

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

$$\begin{aligned}
\mathrm{v}^{(1)}(\xi_1) &= \xi_1,\\
\mathrm{v}^{(2)}(\xi_1,\xi_2) &= \frac{1}{2}\mathrm{ad}\xi_2\xi_1, \text{ and}\\
\mathrm{v}^{(3)}(\xi_1,\xi_2,\xi_3) &= \frac{1}{12}\mathrm{ad}\,\xi_3\,\mathrm{ad}\,\xi_2\,\xi_1 + \frac{1}{4}\mathrm{ad}\,\mathrm{ad}\,\xi_3\,\xi_2\,\xi_1 + \frac{1}{12}\mathrm{ad}\,\xi_2\,\mathrm{ad}\,\xi_3\,\xi_1.
\end{aligned} \tag{17}$$

By taking advantage of the Jacobi-Lie identity, the third and fourth variations can be simplified [4] to

$$\begin{aligned}
\mathrm{v}^{(3)}(\xi_1,\xi_2,\xi_3) &= \frac{1}{6}\left(\mathrm{ad}\,\xi_3\,\mathrm{ad}\,\xi_2\,\xi_1 + \mathrm{ad}\,\mathrm{ad}\,\xi_3\,\xi_2\,\xi_1\right) \text{ and}\\
\mathrm{v}^{(4)}(\xi_1,\xi_2,\xi_3,\xi_4) &= -\frac{1}{12}\left(\mathrm{ad}\,\mathrm{ad}\,\xi_4\,\xi_3\,\mathrm{ad}\,\xi_2\,\xi_1 + \mathrm{ad}\,\mathrm{ad}\,\mathrm{ad}\,\xi_4\,\xi_3\,\xi_2\,\xi_1\right.\\
&\qquad \left. + \mathrm{ad}\,\xi_4\,\mathrm{ad}\,\mathrm{ad}\,\xi_3\,\xi_2\,\xi_1 + \mathrm{ad}\,\xi_3\,\mathrm{ad}\,\mathrm{ad}\,\xi_4\,\xi_2\,\xi_1\right).
\end{aligned} \tag{18}$$

The integrands $\mathrm{v}^{(m)}(\cdot)$ are integrated, as per Equation (14), to give the m^{th}-variations.

3.3 Exponential Representation of a Flow and Perturbation Theory

This section concludes with the exponential representation of a flow as described by the chronological calculus, under the addition of a smallness parameter ϵ. The logarithm operates on the flow $\Phi^X_{0,t}$ of the time-varying vector field X to yield an autonomous vector field Z. For a fixed time, the flow of an autonomous vector field Z can be used in lieu of the flow of the time-varying vector field X for analysis, however, the autonomous vector field Z is given by an infinite series. If the vector field X naturally contains a (small) parameter ϵ as a factor, Proposition 2 can be used to determine how close the flow of a finite series expansion for Z is to the true flow of X. The flow corresponding to the differential equation (1), represented by the chronological exponential $\overrightarrow{\exp}\left(\int_0^t \epsilon X_\tau\,\mathrm{d}\tau\right)$, is asymptotically approximated by the exponential of the logarithm of the original vector field as per Equation (11),

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

Explicit calculations can be complicated, since both the variation $\overrightarrow{\mathrm{V}}_{t_0,t}(\epsilon X_\tau)$ and the exponent $\exp(\cdot)$ are infinite series. Within this context, finite approximations are useful as their effect can be analyzed via Proposition 2. For example, consider the expansion given by the first two variations

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

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

[4] See [43] for calculations.

Incorporation of the variations into the exponent gives

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

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

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

The expansion of the exponent of the logarithm can be done up to any desired order, and will coincide with the Volterra series to the same order (in $\epsilon(t-t_0)$).

The Variation of Constants. Many systems do not come in the form of Equation (1), but may come in the form

$$\dot{x}=X(x,t)+\epsilon\widetilde{X}(x,t). \tag{19}$$

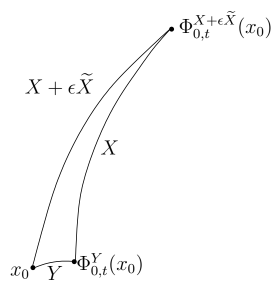

Figure 1: Depiction of the geometry of the variation of constants.

The variation of constants method can be used to transform the perturbed system into the form required by Equation (1). According to the variation of constants,

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

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

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

see [1, §4] for more details. Therefore, $\Phi_{0,t}^{Y}$ can be interpreted as the effect of the pertubation $\epsilon\widetilde{X}$ to X on the unperturbed flow $\Phi_{0,t}^{X}$, c.f. Figure 1. The variation of constants is a key transformation for proving many of the averaging results of Section 4, and is utilized in the example found in Section 5.2.

4 Averaging via Floquet Theory

This section reformulates averaging theory in terms of nonlinear Floquet theory. The basic theorems of Floquet theory are first reviewed in the linear setting, and then extended to the nonlinear setting. Averaging theory is then shown to be the synthesis of nonlinear Floquet theory with the perturbation methods of Section 3.3. Consider the flow of the differential equation (6) rewritten below,

$$\dot{x}=X(x,t)\,,\quad x(0)=x_0, \tag{20}$$

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

4.1 Linear Floquet Theory

The case of a homogeneous linear periodic system,

$$\dot{x}=A(t)x,\qquad A(t)=A(t+T), \tag{21}$$

is well understood using linear Floquet theory. For comparison and convenience, we repeat here the relevant theorems of linear Floquet theory. For linear systems, the flow corresponding to Equation (21), written as $\Phi_{0,t}^{A}$, can be represented by the *fundamental matrix solution*.

Lemma 1 *[11] If C is an invertible matrix, then there exists a matrix B such that $C=\exp(B)$.*

The proof of Lemma 1 uses the logarithm function for matrices [11, pg. 118], therefore $B = \ln(C)$. Suppose that C is the fundamental matrix solution $\Phi^A_{0,t}$ of the system (21) for a fixed time t. Then, in the notation of the chronological calculus,

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

for t fixed, c.f. Equation (12).

Theorem 6 (Linear Floquet Theorem) *[11] Every fundamental matrix solution $\Phi^A_{0,t}$ of (21) takes the form*

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

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

Definition 4 *The* monodromy matrix *corresponding to the flow of (21) is the fundamental solution matrix of (21) after one period of evolution, $M = \Phi^A_{0,T}$.*

The monodromy matrix is a Poincare map for the periodic system (21), therefore its eigenvalues may be used to determine stability of the fixed points of system (21). The eigenvalues are typically called the *Floquet multipliers*, or the *characteristic multipliers*.

Theorem 7 *[11] A necessary and sufficient condition that the system (21) be uniformly stable is that all Floquet multipliers of (21) have modulii less than unity.*

Theorem 7 and Lemma 1 imply that the logarithm of the monodromy matrix may be used to determine stability. From the proof of the linear Floquet theorem [11, pg. 118], it is clear the autonomous linear homogeneous system representing the average of (21) is

$$\dot{y} = By, \tag{22}$$

where

$$B = \frac{1}{T}\ln(M) = \frac{1}{T} \ln \overrightarrow{\exp} \int_0^T A(\tau)\, \mathrm{d}\tau \,. \tag{23}$$

Therefore, an alternative to determining stability of the monodromy matrix M is determining stability of the averaged system (22) via B. The eigenvalues of B are called the *Floquet exponents*, or the *characteristic exponents*.

Corollary 1 *[11] A necessary and sufficient condition that system (21) be uniformly stable is that all characteristic exponents of (22) lie in the complex left half-plane.*

Non-uniqueness of Monodromy Matrix. Given two fundamental matrix solutions $\Phi^1_{0,t}$ and $\Phi^2_{0,t}$ to (21) whose respective generating vector fields are A_1 and A_2, there must exist an invertible transformation of state D taking A_1 to A_2, i.e., $A_2 = D^{-1} A_1 D$ [11]. Denote the monodromy matrix corresponding to the flow of A_1 by M_1. The monodromy matrix for the flow of A_2 is $M_2 = D^{-1} M_1 D$. This non-uniqueness of the monodromy map is equivalent to the freedom of choice regarding the Poincaré section and the conjugacy of different Poincaré sections. In developing Floquet theory for nonlinear systems the same freedom will occur.

4.2 Nonlinear Floquet Theory

By using the exponential representation for the flow of nonlinear systems, linear Floquet theory nicely extends to the nonlinear time-periodic case, Equation (20). While hinted at in prior work [37], the extension of Floquet theory to nonlinear systems is fully developed in this section.

Theorem 8 (Nonlinear Floquet Theorem) *Let $\Phi^X_{0,t}$ be the flow generated by the smooth time-periodic differential equation (20). If the monodromy map has a logarithm, then*

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

where P is T-periodic, $P(t+T) = P(t)$, and Z is an autonomous vector field.

proof
The proof exactly follows the proof of the linear Floquet theorem [11]. It is assumed that the T-periodic vector field in (20) determines a flow, therefore the time shifted version, $\tau = t + T$, does so also. The flow of the time-shifted version is the solution to the differential equation,

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

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

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

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

$$P(t) \equiv \Phi^X_{0,t} \circ \left(\Phi^Z_{0,t}\right)^{-1}. \tag{24}$$

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

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

The T-periodicity of the vector field X ensures that

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

implying that

$$\Phi^Z_{0,T} = \Psi \equiv \Phi^X_{0,T}, \tag{25}$$

i.e., Ψ is the monodromy map of the flow. Written according to the chronological calculus formalism,

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

Inverting via the logarithm,

$$Z = \frac{1}{T} \ln \overrightarrow{\exp}\left(\int_0^T X(x, \theta) \, \mathrm{d}\theta\right), \tag{26}$$

precisely the average of the T-periodic vector field; a connection that will be made more explicit in the sequels. Conditions under which this logarithm exists were briefly discussed in Section 3.2 and are detailed further in [1].
■

The mapping $P(t)$ is called the *Floquet mapping*, and the vector field Z is called the *autonomous averaged vector field corresponding to* X. When the vector field X is known through context, then Z will simply be called the *autonomous averaged vector field*. The monodromy map Ψ found in the proof of the nonlinear Floquet theorem plays an important role in nonlinear Floquet theory just as it did in linear Floquet theory. The proposition below, whose proof can be found in [43], relates the stability of the monodromy map to the stability of the original system.

Theorem 9 *If the monodromy map Ψ of the system (20) has a fixed point, then the actual flow $\Phi^X_{0,t}$ has a periodic orbit whose stability is determined by the stability of the monodromy map.*

If the monodromy map is asymptotically (exponentially) stable, then the corresponding orbit is asymptotically (exponentially) stable. Furthermore, it can be shown [43],

Corollary 2 *If the flow $\Phi^X_{0,t}$ of system (20) has a fixed point x^*, as does the monodromy map Ψ then stability of the fixed point under the flow $\Phi^X_{0,t}$ can be determined from the stability of the monodromy map. In particular an asymptotically (exponentially) stable fixed point for the monodromy map implies an asymptotically (exponentially) stable fixed point for the actual system (20).*

Explicitly calculating the monodromy map may be problematic if not impossible. Equivalent to stability of the monodromy map is stability of the autonomous vector field Z whose flow gives the monodromy map.

Corollary 3 *[37] The stability properties of the logarithm of the monodromy map may be used to infer the stability properties of the monodromy map itself.*

Comment. When applied to the simplified case of linear stability analysis, the above conclusions lead to the following well known fact for linear Floquet theory: calculation of the Floquet multipliers is equivalent to calculation of the Floquet exponents. In the nonlinear theory, the monodromy map is a nonlinear transformation of state, complicating the analysis. Li et al. [24] discuss conditions under which the linearization of the nonlinear system is sufficient for analysis. Barring such a tactic, the logarithm provides a useful means to analyze stability.

Non-uniqueness of the Monodromy Map. Given two fundamental matrix solutions $\Phi^1_{0,t}$ and $\Phi^2_{0,t}$ to (20) whose respective generating vector fields are X_1 and X_2, there must exist an invertible transformation of state Θ from X_1 to X_2, i.e., $X_2 = \Theta^* X_1$. Given the monodromy matrix for X_1, denoted Ψ_1, the monodromy matrix for X_2 is $\Psi_2 = \Theta^{-1} \circ \Psi_1 \circ \Theta$. As in the linear case, the freedom occurs because of the ability to shift time when computing the monodromy map (conjugate Poincaré sections). Although the monodromy map was defined to be $\Psi = \Phi^X_{0,T}$, any choice of initial time may be used, $\widetilde{\Psi} = \Phi^X_{t_0,t_0+T}$, so long as there is flow for one period of time. The two monodromy maps will differ by a transformation of state Θ, i.e., $\Psi = \Theta^{-1} \circ \widetilde{\Psi} \circ \Theta$, corresponding to flow by the difference in initial times. Due to this freedom of the monodramy map calculation, the time-periodic mapping P may be composed of a time-periodic component and an autonomous component. Suppose that

$$P(t) = \widetilde{P}(t) P_0, \tag{27}$$

with P_0 a time-independent transformation, then it is possible to recover a different averaged vector field.

Theorem 10 *[43] Suppose the Floquet mapping has a time-independent bias, i.e., Equation (27) holds. Then an alternative Floquet decomposition of the flow $\Phi^{X_t}_{0,t}$ is possible with $\widetilde{P}(t)$ as the Floquet mapping. The alternative averaged vector field is*

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

The role played by P_0 is equivalent to Θ in the discussion of the non-uniqueness of the monodromy map. The flexibility inherent to the decomposition P and $\exp(Zt)$ is well known in classical n^{th}-order averaging theory [7]. The opposite may also be desirable: to extract unecessary coordinate transformations from the averaged vector field, the reverse procedure may be performed. This is important when one would like to utilize the stability theorems of Section 4, because the bias might prevent the fixed point of the autonomous vector field from corresponding to the fixed point of the actual vector field. Such an example is presented in [43].

4.3 Application to Averaging Theory

Floquet theory is an important tool for averaging theory. The vector field Z from Theorem 8 is the exact average of the time-periodic vector field from Equation (20). Since the actual flow oscillates around the trajectory determined by the autonomous vector field Z, the monodromy map gives the turnpike behavior of the system [38]. To check stability, one must calculate the monodromy map or the full series expansion of Z. If Z cannot be found in closed form, then truncations of the series expansions for the vector field Z can be thought of as partial averages of the system on the order of the truncation [37]. To compute truncations of the infinite series, it is essential for the relevant expression to be in the standard form for the perturbation methods of Section 3.3 to be applied; see Equation (1). The averaged vector field corresponding to (1) is the autonomous vector field of (26).

Calculation of Z requires solving for the variations (13). The integrand of the m^{th} variation of the identity flow, $\mathrm{v}^{(m)}(\cdot)$, is given by the sum of Jacobi-Lie brackets, c.f. Equation (16). Due to multilinearity of Jacobi-Lie brackets, the parameter ϵ may be factored from the variations corresponding to (1). Consequently, the average vector field Z has a power series expansion in ϵ,

$$Z \equiv \sum_{\alpha=1}^{\infty} \epsilon^{\alpha} \Lambda^{(\alpha)}. \tag{28}$$

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

Thus,

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

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

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

Definition 7 *[37] A* stabilized truncated series expansion with respect to linear stability *for the vector field (28) is a truncated vector field that has the same linear stability properties as any higher-order truncation of the vector field, and also the full series expansion of the vector field.*

Definitions 5-7 imply that when a series is stabilized with respect to linear stability for a given truncation $\mathrm{Trunc}_m(\cdot)$, the eigenvalues will not be significantly affected by incorporating additional terms of the series to the truncation. At this stabilized truncation linear stability can be computed and used to determine the linear stability of the original system as per Theorem 9 and Corollary 2 via Corollary 3. See [37, 38] for particular examples involving matrix ODEs. This theorem is a generalization of known results in classical higher-order averaging theory; see [44] and references therein, and also [30]. The reference [24] also discusses when the linearization alone is a feasible (local) approximation to full averaged vector field series expansion for stability analysis.

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

Theorem 11 *The m^{th}-order truncation of the logarithm of the monodromy map gives an m^{th}-order approximation of the flow on a compact subset $K \subset M$,*

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

as $\epsilon \downarrow 0$ on the time-scale $\frac{1}{\epsilon}$.

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

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

and using the variation of constants formula on the flow,

$$\Phi^Z_{0,t} = \Phi^Y_{0,t} \circ \Phi^{Z^m}_{0,t}, \text{ where } Y = \left(\Phi^{Z^m}_{-t,0}\right)^* \widetilde{Z}.$$

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

$$Y = \epsilon^{(m+1)} \left(\Phi^{Z^m}_{-t,0}\right)^* \widehat{Z} \equiv \epsilon^{(m+1)} \widehat{Y}, \text{ where } \widetilde{Z} \equiv \epsilon^{(m+1)} \widehat{Z}.$$

To first order, the asymptotics of the flow $\Phi^Y_{0,t}$ are

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

Section 3.2 from Agračhev and Gamkrelidze [1, pg. 760] provides an approximation to the semi-norm of $\widehat{Y}_\tau$ when restricted to a compact set $K \subset \mathbb{R}^n$. Taking into account simplifications due to the fact that Z^m and $\widehat{Z}$ are autonomous vector fields,

$$\left\|\widehat{Y}_\tau\right\|_{s,K} \leq 3n(2s+2)^{s+1} D_1^2 \exp\left(2D_2 \|Z^m\|_{s+1} \tau\right) D_3 \left\|\widehat{Z}\right\|_{s,M},$$

where

$$D_1 = (s+2)(2s+2)^{s+1}(1 + n + \mathrm{diam}\, K + 4\|Z^m\|_0 \tau)^{s+1},$$

$$D_2 = 3n(2s+4)^{s+2}C_1, \quad \text{and} \quad D_3 = d_K + \operatorname{diam} K + 2\,||Z^m||_0\,\tau,$$

with d_K the distance of the compact set K from the origin and $M = O_R(K)$ a neighborhood of the compact set K with radius $R = 2\,||Z^m||_0\,\tau$. Recalling that the vector field Z^m has a factor of ϵ, and taking the maximum of $\left|\left|\widehat{Y}_\tau\right|\right|_s$ over space and time one arrives at

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

as $\epsilon \downarrow 0$ on the time-scale $\frac{1}{\epsilon}$.

■

If the truncation is linearly stable, and is a stabilized truncation with respect to linear stability, then Theorem 2 can be used to extend the time-scale of the approximation.

Theorem 12 *Suppose that the* m^{th}*-order truncation of the logarithm of the monodromy map is stabilized with respect to linear stability. Then the truncation is* m^{th}*-order close,*

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

for all time $t \in [0, \infty)$.

Thus it has been shown that in lieu of the full logarithm series expansion, it may be possible to utilize truncations of the expansion while maintaining stability properties and a suitable order of approximation. The truncations calculated as per Section 3, c.f. Equation (17), consist of the variations from Equation (16). The variations up to order four are given by

$$\begin{aligned}
\Lambda^{(1)} &= \overline{X}, \\
\Lambda^{(2)} &= \overline{\frac{1}{2}\left[\int_0^t X_\tau\,\mathrm{d}\tau\,, X_t\right]}, \\
\Lambda^{(3)} &= \frac{1}{2}T\left[\Lambda^{(1)}, \Lambda^{(2)}\right] + \frac{1}{3}\overline{\left[\int_0^\tau X_{\tau_1}\,\mathrm{d}\tau_1\,, \left[\int_0^\tau X_{\tau_1}\,\mathrm{d}\tau_1\,, X_\tau\right]\right]}, \text{ and} \\
\Lambda^{(4)} &= -\frac{1}{12}\overline{\int_0^\tau \left[\int_0^{\tau_1}\left[\int_0^{\tau_2} X_{\tau_3}\,\mathrm{d}\tau_3\,, X_{\tau_2}\right]\mathrm{d}\tau_2\,, [X_{\tau_1}, X_\tau]\right]\mathrm{d}\tau_1} \\
&\quad -\frac{1}{12}\overline{\left[\int_0^\tau\left[\int_0^{\tau_1}\left[\int_0^{\tau_2} X_{\tau_3}\,\mathrm{d}\tau_3\,, X_{\tau_2}\right]\mathrm{d}\tau_2\,, x_{\tau_1}\right]\mathrm{d}\tau_1\,, X_\tau\right]} \\
&\quad -\frac{1}{12}\overline{\int_0^\tau\left[\int_0^{\tau_1} X_{\tau_2}\,\mathrm{d}\tau_2\,, \left[\left[\int_0^{\tau_1} X_{\tau_2}\,\mathrm{d}\tau_2\,, X_{\tau_1}\right], X_\tau\right]\right]\mathrm{d}\tau_1}\,.
\end{aligned} \tag{30}$$

4.4 Truncations of the Floquet Mapping

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

$$P(t) \equiv \Phi^X_{0,t} \circ \exp\left(-Zt\right). \tag{31}$$

Theorem 13 *The* m^{th}*-order truncation of the time-periodic Floquet mapping is of order* $\epsilon^{(m+1)}$*-close to the time-periodic Floquet mapping,*

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

as $\epsilon \downarrow 0$ *on the time-scale* 1.

proof
The two flows whose composition results in $P(t)$, c.f. Equation (31), can be written as power series expansions in ϵ. Hence, consider each of the two flows to consist of a truncation, Φ^m and Ψ^m, and a truncated remainder, $\widetilde{\Phi}$ and $\widetilde{\Psi}$:

$$\Phi_{0,t}^X = \Phi^m + \widetilde{\Phi} \quad \text{and} \quad \exp(-Zt) = \Psi^m + \widetilde{\Psi},$$

where

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

Then,

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

Expanding,

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

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

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

is the truncated remainder of the Floquet mapping. From Proposition 1 and Theorem 11, the asymptotics of the terms $\widetilde{\Phi}$ and $\widetilde{\Psi}$ are $O\left(\epsilon^{m+1}t^{m+1}\right)$ and $O\left(\epsilon^{m+1}t\right)$, respectively. By definition of $\mathrm{Trunc}_m\left(\right)$, the asymptotics of the difference $\left(\Phi^m \circ \Psi^m - \mathrm{Trunc}_m\left(\Phi^m \circ \Psi^m\right)\right)$ are $O\left(\epsilon^{m+1}t^{m+1}\right)$. Consequently,

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

from which the theorem statement follows.[5]
■

We have arrived at the following method of averaging for a time-periodic vector field,

$$x(t) = \mathrm{Trunc}_{m-1}\left(P(t)\right)\left(z(t)\right) + O\left(\epsilon^m\right), \text{ and} \tag{32a}$$

$$\dot{z} = \mathrm{Trunc}_m\left(Z\right), \tag{32b}$$

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

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

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

Corollary 4 *[43] If the (amended) truncation* $\mathrm{Trunc}_m\left(P(t)\right)$ *is periodic with period* T*, and the period is on the time-scale* 1*, then the (amended) truncation is order* $\epsilon^{(m+1)}$*-close to* $P(t)$ *for all time.*[6]

Improved m^{th}-order averaging. Ellison et al. [7] discuss what they term to be *improved* m^{th}*-order averaging*. Essentially, the improved average comes from the observation that the truncation of the Floquet mapping $P(t)$ is not of the same order of approximation as the truncation of the autonomous vector field Z. As compared to Equation (32), the improved average incorporates an additional order of expansion in the Floquet mapping,

$$x(t) = \mathrm{Trunc}_m\left(P(t)\right)\left(z(t)\right) + O\left(\epsilon^{m+1}\right) \tag{33a}$$

$$\dot{z} = \mathrm{Trunc}_m\left(Z\right). \tag{33b}$$

[5] The time-scale of approximation can be increased at the cost of decreased order of approximation.
[6] The corollary can be modified to get a larger order of time at the sacrifice of the order of proximity.

The quasi-periodic and aperiodic cases. The papers by Perko and Sáenz [35, 36] have shown that the KBM method of averaging may be used for quasi-periodic and aperiodic systems. The Floquet decomposition in Theorem 8 is still applicable for quasi-periodic and aperiodic systems, however the periodicity of the Floquet mapping $P(t)$ no longer holds. The method for finding the autonomous vector field Z does not depend on the time-periodicity of the vector field X from Equation (20), as proved by Agravchev and Gamkrelidze [1]. The choice of T from (26) is restricted to lie within the interval of definition for the flow $\Phi^X_{0,t}$. For a complete vector field, $T \in [0, \infty]$.

The ability of the generalized averaging theory to hold for the quasi-periodic and aperiodic cases demonstrates why the averaged expansions have such conservative proximity estimates. Classical averaging also begins with derivations of approximate autonomous vector fields that do not require time-periodicity [44]. The introduction of time-periodicity is then shown to improve the order of approximation. The same improvement occurs for the generalized averaging theory, however it will not be proven. See [43] for examples where time-periodicity improves the order of approximation and aperiodicity results in the conservative estimates of Theorems 11 and 13.

5 A General Averaging Theory

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

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

$$\mathrm{Trunc}_0\left(P(t)\right) = \mathrm{Trunc}_0\left(\Phi^X_{0,t} \circ \exp\left(-Zt\right)\right) = \mathrm{Id}. \tag{34}$$

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

$$Z = \frac{1}{T}\int_0^T \epsilon X_\tau \, \mathrm{d}\tau \; = \epsilon \overline{X}, \tag{35}$$

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

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

$$\mathrm{Trunc}_1\left(P(t)\right) = \mathrm{Id} + \epsilon \int_0^t \left(X_\tau - \overline{X}\right) \, \mathrm{d}\tau \; + O\left(\epsilon^2\right). \tag{36}$$

In other words,

$$x(t) = z(t) + \epsilon \int_0^t \left(X(z(t),\tau) - \overline{X}(z(t))\right) \, \mathrm{d}\tau \; + O\left(\epsilon^2\right), \tag{37}$$

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

$$Z = \epsilon \overline{X} + \frac{1}{2}\epsilon^2 \overline{\left[\int_0^t X_\tau \, \mathrm{d}\tau \, , X_t\right]}, \tag{38}$$

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

Equation (37) is close to what Sanders and Verhulst obtain for the compensating flow. Sanders and Verhulst also require the integral term in the flow to have a vanishing average [44]. Bogoliubov and Mitropolsky [5] obtain the same zero average assertion. By introducing an integration constant to ensure a zero average for the term $\int_0^t \left(X(y(t),\tau) - \overline{X}(y(t))\right)\,\mathrm{d}\tau$, the integration constant factors out the effect of varying the initial time t_0 for which the monodromy map is computed, c.f. Section 4.2. For systems that are asymptotically stable, this is akin to neglecting the transient dynamics related to the initial conditions of the full system of Equation (1) [43].

Remark. There are significant differences between the generalized averaging theory and classical averaging theory in spite of their similarity. In order to utilize the powerful results of the chronological calculus, the vector fields must be smooth or analytic. The pertubation approach of Sanders and Verhulst requires the existence of a first derivative and Lipschitz continuity up to the first derivative affording it a level of generality that does not hold for the former. Similarly relaxed conditions hold for the improved n^{th}-order averages found in Ellison et al. [7]. Regardless of the requirements, it is still possible to carry out the formal series expansions, whose truncations may provide asymptotic understanding of the actual flow [38]. By appealing to the classical averaging theorems it is possible to determine the minimal requirements for the calculated average to be nominally valid.

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

$$\begin{aligned}\mathrm{Trunc}_2\left(P(t)\right) = &\mathrm{Id} + \epsilon\int_0^t \left(X_\tau - \overline{X}\right)\,\mathrm{d}\tau + \frac{1}{2}\epsilon^2\int_0^t \left(\left[\int_0^\tau X_s\,\mathrm{d}s\,,X_\tau\right] - \overline{\left[\int_0^t X_s\,\mathrm{d}s\,,X_t\right]}\right)\,\mathrm{d}\tau \\ &+ \frac{1}{2}\epsilon^2\int_0^t X_\tau\,\mathrm{d}\tau \circ \int_0^t X_\tau\,\mathrm{d}\tau - \epsilon^2\int_0^t X_\tau\,\mathrm{d}\tau \circ \overline{X}\,t + \frac{1}{2}\epsilon^2\overline{X}\circ\overline{X}\,t^2\end{aligned} \tag{39}$$

This truncation satisfies $P(0) = P(T)$, but may not periodic for $t > T$. If this is so, then it may be amended according to the discussion in Section 4.4. The autonomous averaged vector field is

$$Z = \epsilon\overline{X} + \frac{1}{2}\epsilon^2\overline{\left[\int_0^t X_\tau\,\mathrm{d}\tau\,,X_t\right]} + \frac{1}{4}T\epsilon^3\left[\overline{X},\overline{\left[\int_0^t X_\tau\,\mathrm{d}\tau\,,X_t\right]}\right] + \frac{1}{3}\epsilon^3\overline{\left[\int_0^\tau X_{\tau_1}\,\mathrm{d}\tau_1\,,\left[\int_0^\tau X_{\tau_1}\,\mathrm{d}\tau_1\,,X_\tau\right]\right]}. \tag{40}$$

Fourth-order averaging. The Floquet mapping for fourth-order averaging can be found in Table 1, and the autonomous averaged vector field is in Table 2. The truncation of the Floquet mapping is easily amended to be T-periodic. The fourth-order averaging results begin to get complicated due to the rate of growth of the averaging terms with each higher order. Averaging calculations for higher orders increasingly become computationally expensive. This fourth level is the last averaging order that will be computed here. In the sequel is given the general algorithm followed in case fifth- or higher-order averaging is required.

5.1 A General Averaging Algorithm

Although calculating higher-order expansions for averaging is a tedious task, there is a very simple algorithm for doing so. This algorithm provides a method of averaging that contains two dimensions of approximation. The first is in the truncations of the logarithm vector field, $\mathrm{Trunc}_m\left(Z\right)$, which captures up to m^{th}-order the dynamics of the system. The second is the truncations of the compensation map, $\mathrm{Trunc}_{(m-1)}\left(P(t)\right)$, which will give m^{th}-order proximity of the composed system to the actual system.

The two components of the generalized averaging theory lead to a very important distinction that can sometimes be confused when performing higher-order averages. For example, ignoring the Floquet mapping and using only the autonomous flow results in averaged equations of motion that are capable of capturing the m^{th}-order dynamics, but will only be good to 1^{st}-order proximity. The distinction between roles of the mapping $P(t)$ and the autonomous vector field was also made by Bogoliubov and Mitropolsky [5].

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

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

Table 1: Truncated Floquet mapping, $\mathrm{Trunc}_3\left(P(t)\right)$, for fourth-order average.

The averaging method provided here is capable of replicating classical averaging results and appears to coincide with known methods for obtaining higher-order averages. The general strategy implied by the averaging theory derived herein can be found in Table 3.

5.2 Example: Particle in an Electromagnetic Field

The example of this section considers a charged particle in $\mathbb{R}^3$ constrained to move in the (x, y)-plane while subject to a magnetic field of strength B pointing in the z direction, and an electric field of strength E pointing in the x direction, see Figure 2. The phase space is $P \cong T^*\mathbb{R}^2$ with coordinates $(x, y, p_x, p_y) \in P$. This problem was examined in Omohundro [31] within the context of symmetry reduction.

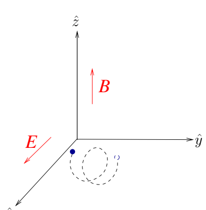

Figure 2: Charged particle in an electromagnetic field.

The Hamiltonian of the system is

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

with the noncanonical Poisson bracket,

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

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

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

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

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

Table 2: Truncated averaged vector field, $\mathrm{Trunc}_4\left(Z\right)$, for fourth-order average.

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

Table 3: Algorithm for Computing the Average.

unperturbed case $\epsilon = 0$, while in the case of the full field strength it is $\epsilon = 1$. The equations of motion are

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

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

$$
X_0 = \left\{ \begin{array}{c} \frac{1}{m}p_x \\ \frac{1}{m}p_y \\ \frac{eB}{mC}p_y \\ -\frac{eB}{mc}p_x \end{array} \right\}, \quad \text{and} \quad \widetilde{X} = \epsilon \left\{ \begin{array}{c} 0 \\ 0 \\ eE \\ 0 \end{array} \right\}, \tag{41}
$$

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

$$
\Phi^X_{0,t} = \Phi^{X_0}_{0,t} \circ \Phi^Y_{0,t}, \quad \text{where} \quad Y = \epsilon\left(\Phi^{X_0}_{0,t}\right)^* \widetilde{X}.
$$

Due to the oscillatory nature of the unperturbed flow $\Phi^{X_0}_{0,t}$, the vector field Y is time periodic. Via Floquet theory, the flow of Y is decomposed into two parts,

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

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

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

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

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

where $\omega = \frac{eB}{mc}$. The vector field Y is calculated to be

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

and the average of Y is

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

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

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

The true Floquet mapping of the system is the composed periodic mapping

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

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

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

Explicitly composing the flows results in

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

By finding the solution in closed form using Floquet theory, we have demonstrated that averaging may be exact. In this example, the ϵ parameter is not needed to obtain the approximate expansions of the averaged vector field and the Floquet mapping. Therefore, the value of the ϵ parameter is not restricted.

5.3 Averaging Theory and Dynamical Systems Theory

The averaging theory described in this paper is related to the embedding of diffeomorphisms into flows generated by vector fields. Define $\mathrm{Diff}(M)$ to be the set of diffeomorphisms of M into M, and also define $\mathrm{Diff}^r(M) \subset \mathrm{Diff}(M)$ to consist of C^r diffeomorphisms. Although it is known that diffeomorphisms embed into the flows of time-varying vector fields [39], it is more difficult to ascertain conditions for which a diffeomorphism may be embedded into the flow of an autonomous vector field. The logarithm computation of Equation (26) provides a means to convert the flow of a smooth time-varying vector field into the flow of a smooth autonomous vector field, providing a mechanism by which smooth diffeomorphisms embed into the flows of smooth autonomous vector fields. Unfortunately, the dynamical systems literature appears to contradict this possibility [18, 32, 34].

Given that arbitrary diffeomorphisms may not embed into flows, it is natural to seek conditions under which a diffeomorphism may embed into a flow. If a diffeomorphism embeds in a flow, then it may do so for uncountably many flows [9, 32]. Uniqueness aside, the primary difficulty of the embedding lies near fixed points of diffeomorphisms [33]. Transversality of the flow and non-resonance of the linear eigenvalues of the vector field play an important role in ensuring embeddability [24, 28, 32, 34]. For diffeomorphisms that do not have a center manifold in the linearization, there are positive results concerning the embedding problem [4, 24]. In addition to these restricted results, research suggests that diffeomorphisms may be asymptotically embedded into flows [13, 18], i.e., given an ϵ, there exists a vector field whose flow after a fixed time is a diffeomorphism ϵ-close to the given diffeomorphism. Such a situation is also suggested by the works of Epstein, Mather, and Thurston [8, 29, 42]. The practical implications are that, for certain cases, the logarithm computation is close in an asymptotic sense, as emphasized by Agračhev and Gamkrelidze [1], supporting the idea that the logarithm may result in an autonomous vector field with the property that its unit-time flow asymptotically approximates the diffeomorphism obtained from the fixed-time flow of a time-varying vector field sufficiently close to the identity flow, as in Proposition 2.

6 Conclusion

This paper generalizes averaging theory to arbitrary order, and in the process proves associated theorems related to stability under averaging. Importantly, the formulation of the averaged expansions is intrinsic as they require Jacobi-Lie bracket calculations instead of coordinate-based calculations. Additionally, a new paradigm for thinking about averaging theory was presented. Averaging theory was shown to be the synthesis of two distinct theories: (1) Floquet theory, and (2) perturbation theory. By working within this framework many of the properties of classical and KBM averaging methods can be recovered. Although the theory is capable of recovering known results in averaging theory, the framework used is restricted to smooth or analytic vector fields, and provides conservative estimates. The work of Ellison et. al. [7] and others have weaker requirements on the vector fields and stronger estimates, but use a perturbative approach. Further investigation into the similarities between the perturbative coordinate-based approach and Volterra series expansions could serve to relax this requirement. Nevertheless, understanding how averaging theory is really a synthesis of the two aforementioned theories may lead to a better understanding of what the process of averaging physically means for dynamical systems.

This analysis leads to several topics within the realm of nonlinear control and dynamics. In particular, we have recently shown how this theory can be used to design exponentially stabilizing feedback controllers for a wide class of nonholonomic mechanical systems [43]. Lastly, many other elements that comprise averaging theory have yet to be placed within the framework of the generalized averaging theory, of which we mention the development of two-timing methods and the analysis of resonance or multi-phase averaging. An important extension would also be to partial differential equations. The derivation of Agračhev and Gamkrelidze [1] appears to support this goal, although not without difficulty.

References

[1] A. Agračhev and R. Gamkrelidze. The exponential representation of flows and the chronological calculus. *USSR Sbornik*, 35(6):727–785, 1978.

[2] A. Agračhev and R. Gamkrelidze. Chronological algebras and nonstationary vector fields. *Journal of Soviet Mathematics*, 17:1650–1675, 1979.

[3] A. Agračhev, R. Gamkrelidze, and A. V. Sarychev. Local invariants of smooth control systems. *Acta Applicandae Mathematicae*, 14:1650–1675, 1989.

[4] J. Arango and A. Gómez. Flows and diffeomorphisms. *Revista Colombiana de Matemáticas*, 32:13–27, 1998.

[5] N. M. Bogoliubov and Y. A. Mitropolsky. *Asymptotic Methods in the Theory of Non-Linear Oscillations*. Hindustan Publishing Co., 1961.

[6] K. Chen. Integration of paths, geometric invariants and a generalized Baker-Hausdorff formula. *Annals of Mathematics*, 65(1):163–178, 1957.

[7] J. A. Ellison, A. W. Saenz, and H. S. Dumas. Improved n^{th}-order averaging for periodic systems. *Journal of Differential Equations*, 84(2):383–403, 1990.

[8] D. B. A. Epstein. The simplicity of certain groups of homeomorphisms. *Compositio Mathematica*, 22(2):165–173, 1970.

[9] N. Fin and G. E. Schweigert. On the group of homeomorphisms of an arc. *Annals of Mathematics*, 62(2):237–253, 1955.

[10] J. Guckenheimer and P. Holmes. *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*, volume 42 of *AMS*. Springer Verlag, N.Y., 1990.

[11] J. K. Hale. *Ordinary Differential Equations*, volume 21 of *Pure and Applied Mathematics*. Wiley-Interscience, 1969.

[12] J. K. Hale and S. M. Verduyn-Lunel. Averaging in infinite dimensions. *Journal of Integral Equations and Applications*, 2(4):463–94, 1990.

[13] Z. Jingzhong and Y. Lu. Embedding of a homeomorphism in a flow and asymptotic embedding. *Scientia Sinica Series A*, 28(7):683–696, 1985.

[14] M. Kawski. Nonlinear control and combinatorics of words. In B. Jakubczyk and W. Respondek, editors, *Geometry of Feedback and Optimal Control*, pages 305–346. Dekker, 1998.

[15] M. Kawski. Controllability via chronological calculus. In *Proceedings of IEEE Conference on Decision and Control*, pages 2920–2925, 1999.

[16] M. Kawski and H. Sussman. Noncommutative power series and formal Lie-algebraic techniques in nonlinear control theory. In U. Helmke, D. Pratzel-Wolters, and Zerz E., editors, *Operators, Systems, and Linear Algebra*, pages 111–128. Teubner, 1997.

[17] H. K. Khalil. *Nonlinear Systems*. MacMillan Pub. Co., NY, 1992.

[18] N. Kopell. Commuting diffeomorphisms. In *Proceedings of the American Mathematical Society Symposium on Pure Math*, volume 14, pages 165–184, 1970.

[19] J. Kurzweil and J. Jarník. Limit processes in ordinary differential equations. *Journal of Applied Mathematics and Physics*, 38, 1987.

[20] J. Kurzweil and J. Jarník. Iterated Lie brackets in limit processes in ordinary differential equations. *Results in Math*, 14, 1988.

[21] G. Lafferriere and H. J. Sussmann. Motion planning for controllable systems without drift. *Proceedings of IEEE International Conference on Robotics and Automation*, pages 1148–1153, 1991.

[22] B. Lehman and S. P. Weibel. Fundamental theorems of averaging for functional differential equations. *Journal of Differential Equations*, 152:160–90, 1999.

[23] A. Y. T. Leung and Q. C. Zhang. Higher-order normal form and period averaging. *Journal of Sound and Vibration*, 217(5):795–806, 1998.

[24] W. Li, J. Llibre, and X. Zhang. Extension of Floquet's theory to nonlinear periodic differential systems and embedding diffeomorphisms in differential flows. *American Journal of Mathemathics*, 124:107–127, 2002.

[25] W. S. Liu. An approximation algorithm for nonholonomic systems. *SIAM Journal of Control and Optimization*, 35(4):1328–1365, 1997.

[26] W. S. Liu. Averaging theorems for highly oscillatory differential equations and iterated Lie brackets. *SIAM Journal of Control and Optimization*, 35(6):1989–2020, 1997.

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

[28] H. Maoan. Conditions for a diffeomorphism to be embedded in a C^r flow. *Acta Mathematicae Sinica*, 4(2):111–123, 1988.

[29] J. N. Mather. Commutators of diffeomorphisms. *Commentarii Mathematici Helvetici*, 49(4):512–528, 1979.

[30] J. T. Montgomery. Existence and stability of periodic motion under higher order averaging. *Journal of Differential Equations*, 64:67–78, 1986.

[31] S. M. Omohundro. *Geometric Perturbation Theory in Physics*. World Scientific, 1986.

[32] J. Palis. On Morse-Smale dynamical systems. *Topology*, 8:385–405, 1969.

[33] J. Palis. Vector fields generate few diffeomorphisms. *Bulletin of the American Mathematical Society*, 80(3):503–505, 1974.

[34] J. Palis and J. C. Yoccoz. Rigidity of centralizers of diffeomorphisms. *Annales Scientifiques de l'École Normale Supérieure*, 22(4):81–98, 1989.

[35] L. M. Perko. Higher order averaging and related methods for perturbed periodic and quasi-periodic systems. *SIAM Journal of Applied Mathematics*, 17(4):469–724, 1968.

[36] A. W. Sáenz. Higher-order averaging for nonperiodic systems. *Journal of Mathematical Physics*, 32(10):2679–94, 1991.

[37] A. V. Sarychev. Lie- and chronologico-algebraic tools for studying stability of time-varying systems. *Systems and Control Letters*, 43(1):59–76, 2001.

[38] A. V. Sarychev. *Stability Criteria for Time-Periodic Systems via Higher-Order Averaging Techniques*. Lecture Notes in Control and Information Sciences. Springer-Verlag, 2001.

[39] S. Smale. Differential dynamical systems. *Bulletin of the American Mathematical Society*, pages 747–817, 1967.

[40] H. J. Sussmann. A general theorem on local controllability. *SIAM Journal of Control and Optimization*, 25(1), 1987.

[41] H. J. Sussmann and W. S. Liu. Limits of highly oscillatory controls and the approximation of general paths by admissible trajectories. In *Proceedings of IEEE Conference on Decision and Control*, pages 437–442, 1991.

[42] W. P. Thurston. Foliations and groups of diffeomorphisms. *Bulletin of the American Mathematical Society*, 80(2):304–307, 1974.

[43] P. A. Vela. *Averaging and Control of Nonlinear Systems*. PhD thesis, California Institute of Technology, 2003.

[44] F. Verhulst and J. A. Sanders. *Averaging Methods in Nonlinear Dynamical Systems*. Springer-Verlag, N. Y., 1985.

[45] S. Wiggins. *Introduction to Applied Nonlinear Dynamical Systems and Chaos*. Number 2 in TAM. Springer, 1996.

[46] K. Yagasaki and T. Ichikawa. Higher-order averaging for periodically forced weakly nonlinear systems. *International Journal of Bifurcation and Chaos*, 9(3):519–31, 1999.