Nonlinear Dyn
https://doi.org/10.1007/s11071-019-05085-4

ORIGINAL PAPER

On higher-order averaging of time-periodic systems: reconciliation of two averaging techniques

Marco Maggia · Sameh A. Eisa · Haithem E. Taha

Received: 29 November 2018 / Accepted: 17 June 2019

Abstract In this paper we show how higher-order averaging can be used to remedy serious technical issues with the direct application of the averaging theorem. While doing so, we reconcile two higher-order averaging methodologies that were developed independently using different tools and within different communities: (i) perturbation theory using a near-identity transformation and (ii) chronological calculus using Lie algebraic tools. We provide the underpinning concepts behind each averaging approach and provide a mathematical proof for their equivalence up to the fourth order. Moreover, we provide a higher-order averaging study and analysis for two applications: the classical problem of the Kapitza pendulum and the modern application of flapping flight dynamics of micro-air-vehicles and/or insects.

Keywords Averaging theory · Higher-order averaging · Chronological calculus · Lie transform · Time-periodic systems · Vibrational stabilization · Vibrational control

This work was supported by the National Science Foundation, Grant CMMI-1709746 and its continuation CMMI-1846308.

M. Maggia · S. A. Eisa · H. E. Taha (✉)
Department of Mechanical and Aerospace Engineering, University of California, Irvine, Irvine, CA 92697, USA

H. E. Taha
e-mail: hetaha@uci.edu

M. Maggia
e-mail: mmaggia@uci.edu

S. A. Eisa
e-mail: seisa@uci.edu

1 Introduction

Nature is teeming with periodically forced creatures (e.g., insects, birds, fish). With the current active research trend of mimicking nature's designs, there are plenty of bio-inspired engineering systems that are represented by nonlinear time-periodic (NLTP) models such as swimming robots [1–4] and flapping-wing micro-air-vehicles [5,6]. Additionally, high-frequency periodic control is a common approach for stabilization and path planning for under-actuated mechanical systems (with control inputs less than degrees of freedom) [2,7–9]. Vibrational control is an open-loop stabilization technique via the application of sufficient high-frequency, high-amplitude, periodic inputs. It was first introduced by Meerkov [10] for linear time-varying systems. It is well known for its robustness [11] and elegance (stabilization without feedback). By definition, vibrationally controlled systems are time-periodic systems.

Analyzing the dynamics of these time-varying systems is quite more challenging than that of autonomous (time-invariant) ones. For example, a NLTP system usually attains equilibrium when all of its states undergo periodic solutions. That is, its equilibria may not be represented by fixed points but rather by *periodic orbits*. Moreover, even when the system is linear and

Published online: 22 June 2019

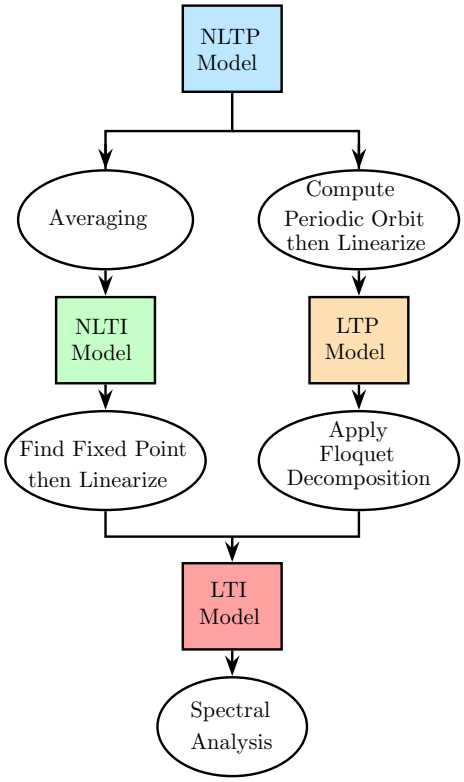

Fig. 1 The two main approaches used to analyze NLTP systems

has a fixed point ($\dot{\mathbf{x}} = [\mathbf{A}(t)]\mathbf{x}$), its stability cannot be assessed through the eigenvalues of the system matrix $[\mathbf{A}(t)]$. That is, in contrast to linear time-invariant systems $\dot{\mathbf{x}} = [\mathbf{A}]\mathbf{x}$ where a Hurwitz matrix (its eigenvalues have strictly negative real parts) implies exponential stability, if all the eigenvalues of the time-varying matrix $[\mathbf{A}(t)]$ lie in the open left half of the complex plane for all times, the system may still be unstable; Markus and Yamabe [12] provided a counterexample.

In general, the stability analysis of NLTP systems is conducted using one of two main approaches: a numerical approach based on the Floquet theorem [13–15] and an analytical approach based on the averaging theorem [16–19]. As shown in Fig. 1, the first approach requires determination of the periodic orbit (i.e., solving the dynamic equations). Then, the Floquet theorem is used to analyze the stability of the linear, time-periodic (LTP) system obtained by linearizing the dynamics about the concerned periodic orbit. For non-trivial applications, this approach cannot be performed analytically. Hence, the inevitable numerical implementation of the Floquet theorem precludes scrutinizing the dynamical behavior of the system on an analytical level. The second approach is to use the averaging theorem to transform the NLTP system into a nonlinear, time-invariant (NLTI), i.e., an autonomous, system in which the periodic orbit of the original system corresponds to a fixed point. Then, the averaging theorem guarantees that, for high enough frequency, exponential stability of this fixed point implies exponential stability of the corresponding periodic orbit. Thus, the averaging approach dodges the determination of the periodic orbit under study and, hence, provides a compact, constructive technique to ensure the periodic orbit and to analyze its stability (by ensuring the corresponding fixed point and studying its stability). In fact, averaging has been the standard and most common technique to design vibrational control systems [7,20].

Despite its convenience to analyze NLTP systems, the application of averaging should be performed carefully. In fact, there might not be a more expressive statement about the subtleties of averaging than that by Sanders and Verhulst in their book on the topic [17]:

> To many physicists and astronomers averaging seems to be such a natural procedure that they do not even bother to justify the process. However, it is important to have a rigorous approximation theory, since it is precisely the fact that averaging seems so natural that obscures the pitfalls and restrictions of the method.

Sticking to the standard form of averaging, its direct application may not be admissible and/or practically useful in many cases. First, the averaging theorem necessitates a special structure (weakly forced systems). Therefore, it cannot be applied directly to vibrational control systems, which are typically high-amplitude, periodically forced systems. The standard remedy for this case is to apply the nonlinear *variation of constants formula* [7,21] to transform the multi-scale system into two companion systems, each of which is amenable to the averaging theorem (see [7,8,22,23]). It is interesting to note that the use of the variation of constants (VOC) formula before averaging goes back to Lagrange in his analysis of the perturbed two-body problem (see [24, pp. 181–184]). However, the VOC formula, which requires computation of flow maps (i.e., solution of part of the differential equation), may not be analytically tractable for complex systems, hence frittering away the analytical advantage of the averag-

ing analysis. Second, even when the system structure permits analytical application of the VOC (e.g., simple mechanical systems [7]), a major issue with the averaging theorem is that it is applicable only for *high enough* oscillation frequency. That is, it only guarantees the existence of a threshold frequency ω^* above which stability of the averaged dynamics implies stability of the corresponding NLTP system. The predicament is that almost none of the averaging theorems tell how much ω^* is. As such, vibrational control is typically achieved by ensuring that the averaged dynamics is stable and then sweeping the frequency (e.g., by simulation) until hopefully the threshold is hit. Then, the averaging theorems guarantee stability for the original time-periodic system for all frequencies above this threshold. The main issue stressed in the recent work of Berg and Wickramasinghe [20] is that finding a stable response of the time-periodic system at a certain frequency does not necessarily mean that the threshold is reached. That is, the time-periodic system may be stable at some ω_1 and unstable at $\omega_2 > \omega_1$ (clearly, $\omega_1 < \omega_2 < \omega^*$). This finding also points to the possibility for vibrational stabilization at a frequency less than the threshold value, hence relaxing the averaging unfeasible requirements.

In this paper, we propose higher-order averaging techniques as potential remedy for the two issues raised above regarding direct (first-order) averaging. In particular, we show how higher-order averaging captures more dynamical features that are typically neglected by direct averaging. Moreover, the threshold frequency may decrease as the order of averaging increases. While doing this, we find it necessary to reconcile different averaging approaches in the literature. Two higher-order averaging techniques can be found in the literature, which are developed separately employing different tools. The first technique is based on the classical averaging theory developed by Krilov et al. [25,26]; some authors such as Vela [27] refer to it as the *KBM* method. This classical approach has its roots in the perturbation theory; the main objective is to find a near-identity transformation that transforms the time-periodic system into a time-invariant system after neglecting terms of arbitrary small order of magnitude.

The other higher-order averaging technique is based on the so-called *chronological calculus*, exploiting Lie algebraic tools. In honor of the 70th birthday of the great Russian mathematician Lev Pontryagin, the founder of the optimal control theory, his students and colleagues Agrachev and Gamkrelidze [28,29] developed a new branch of calculus to study time-varying dynamical systems: the chronological calculus. The main objective was to provide an exponential representation of the flow along a time-varying vector field, in a way that resembles the exponential representation of the flow of autonomous vector fields given by the Fliess functional expansion [30]. Utilizing these mathematical tools, Sarychev [31] and Vela [27] developed higher-order averaging techniques for time-periodic systems, generalizing the classical averaging theorem to cases where the excitation frequency is not high enough. Inspired by the Floquet theorem, the main objective was to determine a time-invariant (averaged) system whose flow after one period matches the flow of the time-periodic system after one period. Hence, the stability characteristics of the NLTP system can be deduced from the stability characteristics of the averaged system, thereby extending the Floquet theorem to nonlinear systems. Because this approach is developed using differential geometric techniques, the higher-order averaged dynamics are given in terms of *Lie brackets* between the time-periodic vector field and its integrals.

Throughout several discussions between the third author and Prof. Ali Nayfeh, the latter always suggested that the two averaging approaches are just the same. Yet, it is not clear in the literature whether such approaches are equivalent or not. Therefore, we find it a good opportunity to reconcile the two higher-order averaging techniques in this special issue in the memory of Prof. Nayfeh, particularly because he made several important contributions to the first classical approach [13–15,32].

The main contribution of this paper is to reconcile these two averaging techniques, under a unified notation, providing explicit formulas for higher-order averaging terms. Moreover, we prove that the two approaches are equivalent up to the fourth order. It is noteworthy to mention that, unlike the classical averaging approach, the one based on chronological calculus does not provide a recursive formula for the computation of higher-order terms. Hence, the equivalence between the two averaging techniques cannot be systematically carried out for an arbitrarily high order. In this paper, the equivalence is mathematically proven only up to the fourth order. Moreover, we show a "conceptual" equivalence between the two approaches which suggests that the mathematical equivalence can

be extended for terms beyond the fourth order. This equivalence will allow each approach to benefit from tools and results developed in the other branch. After proving such an equivalence, we apply these techniques to two applications: (i) a classical application of the Kapitza pendulum (inverted pendulum subjected to vertical oscillation) and (ii) a modern application on the hovering flight of insects and flapping-wing micro-air-vehicles. In these examples, we show how higher-order averaging captures more dynamical features than direct averaging and how it can mitigate the technical issues associated with the direct application of the averaging theorem.

The paper is organized as follows: In Sect. 2, the two approaches (classical averaging and chronological calculus) are described separately. The reconciliation of the two averaging approaches is given in detail in Sect. 3. Lastly, Sect. 4 discusses the two applications mentioned above.

2 Two higher-order averaging techniques

NLTP systems are usually characterized by two timescales: a fast timescale and a slow timescale. The goal of averaging is to filter the slow behavior by capturing the mean effect of the dynamics over a period of the fast timescale. During a period, the slowly varying states (that do not explicitly depend on the fast time) are approximated as constants. As an example, let us consider a hummingbird flapping its wings. The wings move periodically at a frequency ω that is much higher than the natural frequency associated with the bird's body flight ω_n (i.e., $\omega \gg \omega_n$). During one flapping cycle, the body oscillates in all directions. Nevertheless, on the average, it is hovering in the same place (e.g., over a flower). Moreover, the oscillation of the body around the mean position/attitude is minimal. Therefore, averaging seems so natural to analyze such a dynamical system. Analyzing the averaged dynamics can be extremely beneficial because the complexity of the system can be noticeably reduced. In particular, the explicit dependence on time is eliminated and the averaged system becomes autonomous (i.e., it is fast time-invariant). The approximation by averaging is justified by the following theorem, better known simply as the *averaging theorem*.

Theorem 1 Averaging Theorem (Theorem 10.4 in [16]). *Consider the following NLTP system written in averaging-canonical form*

$$\dot{\mathbf{x}} = \epsilon\, \mathbf{f}(\mathbf{x}, t, \epsilon), \tag{1}$$

where $\mathbf{x} \in \mathbb{R}^n$, $\dot{\mathbf{x}}$ *denotes the derivative of* $\mathbf{x}$ *with respect to* t, *and* ϵ *is a small parameter such that* $0 < \epsilon \ll 1$. *The vector field* $\mathbf{f}$ *is smooth in* $\mathbf{x}$, *analytic in* ϵ, *and* T*-periodic in* t. *The averaged system corresponding to* (1) *is*

$$\dot{\overline{\mathbf{x}}} = \epsilon\, \overline{\mathbf{f}}(\overline{\mathbf{x}}) \tag{2}$$

where $\overline{\mathbf{f}}(\mathbf{x}) = \frac{1}{T}\int_0^T \mathbf{f}(\mathbf{x}, \tau, 0)\, \mathrm{d}\tau$.

1. *If* $\mathbf{x}(0) - \overline{\mathbf{x}}(0) = \mathcal{O}(\epsilon)$, *then there exist* $b, \epsilon^* \in \mathbb{R}^+$ *such that* $\mathbf{x}(t) - \overline{\mathbf{x}}(t) = \mathcal{O}(\epsilon)$ *for all* $t \in [0, b/\epsilon]$ *and for all* $\epsilon \in (0, \epsilon^*)$.
2. *If the origin* $\overline{\mathbf{x}} = \mathbf{0}$ *is an exponentially stable equilibrium of system* (2) *and if* $\mathbf{x}(0) - \overline{\mathbf{x}}(0) = \mathcal{O}(\epsilon)$, *then there exists an* ϵ^* *such that* $\mathbf{x}(t) - \overline{\mathbf{x}}(t) = \mathcal{O}(\epsilon)$ *for all* $t \geq 0$ *and for all* $\epsilon \in (0, \epsilon^*)$. *Furthermore, system* (1) *has a unique, exponentially stable,* T*-periodic solution* $\mathbf{x}_T(t)$ *such that* $\|\boldsymbol{x}_T(t)\| \leq k\epsilon$ *for some* $k \in \mathbb{R}^+$.

Sometimes, the first-order averaging approximation is too poor and the averaged system (2) fails to capture some important characteristics of the NLTP system. In particular, Theorem 1 is valid only for weakly forced systems or for systems with large separation between the two time scales (e.g., the ratio of the body's flight natural frequency to the flapping frequency for the hummingbird, $\epsilon \approx \omega_n/\omega$, $\epsilon \ll 1$). Also, no information is given about how large this separation should be for the theorem to be valid. For these reasons, higher-order averaging terms are introduced. The higher-order terms are also time-invariant and are scaled by powers of the small parameter $\epsilon \ll 1$. In the averaging theory, an rth-order averaged system implies neglecting terms of order ϵ^r. As such, this truncation becomes more accurate as the order increases and/or the parameter ϵ decreases.

In this section we briefly summarize the two higher-order averaging approaches that are found in the literature. We start with the classical averaging approach [15,17,18,33] and proceed with the description of the less known approach based on chronological calculus [27,31,34].

2.1 Higher-order averaging via a near-identity transformation (lie transform)

This higher-order averaging approach utilizes a particular change of coordinates to approximate the NLTP

system by a NLTI one. The original NLTP system will be at times referred to as the *exact system*. The methodology has been developed from perturbation theory (see [13,15,17,18]).

We begin our analysis by considering a NLTP dynamical system written in the averaging-canonical form[1] (1). Since $\mathbf{f}$ is smooth in $\mathbf{x}$ and analytical in ϵ, it can be expanded in power series of ϵ as

$$\mathbf{f}(\mathbf{x},t,\epsilon) = \sum_{m=1}^{\infty} \frac{\epsilon^{m-1}}{m!}\mathbf{f}_m(\mathbf{x},t) = \mathbf{f}_1(\mathbf{x},t) + \frac{\epsilon}{2!}\mathbf{f}_2(\mathbf{x},t) + \frac{\epsilon^2}{3!}\mathbf{f}_3(\mathbf{x},t) + \cdots \tag{3}$$

where the functions $\mathbf{f}_m$ are computed as,

$$\mathbf{f}_m(\mathbf{x},t) = m\frac{\partial^{m-1}\mathbf{f}}{\partial\epsilon^{m-1}}\bigg|_{\epsilon=0}. \tag{4}$$

The objective is to find a transformation that converts the $\mathbf{x}$-dynamics into a time-invariant system in the transformed coordinates $\mathbf{y}$. Given a transformation $\mathbf{w}$, we perform the following change of coordinates $\mathbf{x} \to \mathbf{y}$

$$\mathbf{x} = \mathbf{y} + \epsilon\mathbf{w}(\mathbf{y},t,\epsilon), \tag{5}$$

where $\mathbf{y} \in \mathbb{R}^n$ is the transformed state. For $\epsilon = 0$, Eq. (5) reduces to the identity $\mathbf{x} = \mathbf{y}$; because ϵ is small, the transformation in (5) is a quasi-identity, hence the name *near-identity transformation*. This transformation is sometimes referred to as a *Lie transform*. When working with NLTP systems, the function $\mathbf{w}$ is also required to be T-periodic in t. Furthermore, we assume $\mathbf{w}$ to be smooth in $\mathbf{y}$ and analytic in ϵ, so we can expand it in an asymptotic power series of ϵ as

$$\mathbf{w}(\mathbf{y},t,\epsilon) = \sum_{m=1}^{\infty} \frac{\epsilon^{m-1}}{m!}\mathbf{w}_m(\mathbf{y},t)$$

When transformation (5) is applied to system (1), we obtain a new nonlinear smooth system in the form

$$\dot{\mathbf{y}} = \sum_{m=1}^{\infty} \frac{\epsilon^m}{m!}\mathbf{g}_m(\mathbf{y},t), \tag{6}$$

where the functions $\mathbf{g}_m(\mathbf{y},t)$ depend on the particular choice of the near-identity transformation (5) and are in general time-varying vector fields. We now demonstrate the procedure to make $\mathbf{g}_1(\mathbf{y},t)$ time-invariant, and we then extend the reasoning to higher-order terms $\mathbf{g}_m(\mathbf{y},t)$ with $m > 1$.

[1] It is important to notice that, if direct averaging is applied to a system that is *not* written in the averaging-canonical form, the following analysis might not be correct. Averaging a system not written in canonical form is known as *crude averaging* and is discouraged as pointed out in [17].

2.1.1 First-order averaging

Let us rewrite Eqs. (1), (5) and (6) as

$$\dot{\mathbf{x}} = \epsilon\mathbf{f}_1(\mathbf{x},t) + \epsilon^2\hat{\mathbf{f}}(\mathbf{x},t,\epsilon), \tag{7}$$

$$\mathbf{x} = \mathbf{y} + \epsilon\mathbf{w}_1(\mathbf{y},t) + \epsilon^2\hat{\mathbf{w}}(\mathbf{y},t,\epsilon), \tag{8}$$

$$\dot{\mathbf{y}} = \epsilon\mathbf{g}_1(\mathbf{y},t) + \epsilon^2\hat{\mathbf{g}}(\mathbf{y},t,\epsilon), \tag{9}$$

where $\hat{\mathbf{f}}$, $\hat{\mathbf{g}}$ and $\hat{\mathbf{w}}$ denote remainders. Differentiating Eq. (8) with respect to t and substituting by Eq. (9) in the resulting equation, we obtain

$$\dot{\mathbf{x}} = \epsilon\mathbf{g}_1(\mathbf{y},t) + \epsilon^2\hat{\mathbf{g}}(\mathbf{y},t,\epsilon) + \epsilon\left(\frac{\partial\mathbf{w}_1}{\partial\mathbf{y}}\left(\epsilon\mathbf{g}_1(\mathbf{y},t) + \epsilon^2\hat{\mathbf{g}}(\mathbf{y},t,\epsilon)\right) + \frac{\partial\mathbf{w}_1}{\partial t}\right) + \mathcal{O}(\epsilon^2). \tag{10}$$

Substituting Eq. (8) into the argument of $\mathbf{f}_1$ in Eq. (7), we obtain

$$\dot{\mathbf{x}} = \epsilon\mathbf{f}_1\left(\mathbf{y} + \epsilon\mathbf{w}_1(\mathbf{y},t),t\right) + \mathcal{O}(\epsilon^2). \tag{11}$$

Equating the terms in Eqs. (10) and (11) of the same ϵ-power, the ϵ^1 equality implies

$$\mathbf{g}_1(\mathbf{y},t) = \mathbf{f}_1(\mathbf{y},t) - \frac{\partial\mathbf{w}_1(\mathbf{y},t)}{\partial t}. \tag{12}$$

The vector field $\mathbf{g}_1(\mathbf{y},t)$ in Eq. (12) is still time-varying. The only way to make $\mathbf{g}_1$ time-independent, while $\mathbf{w}$ is T-periodic, is to take $\mathbf{g}_1$ equal to the average, over a period T, of the vector field $\mathbf{f}_1$ [17,18]

$$\mathbf{g}_1(\mathbf{y}) = \bar{\mathbf{f}}_1(\mathbf{y}) = \frac{1}{T}\int_0^T \mathbf{f}_1(\mathbf{y},t)\,\mathrm{d}t. \tag{13}$$

Combining Eqs. (12) and (13) we find an expression for the first term of the expansion of the transformation $\mathbf{w}(\mathbf{y},t)$ as

$$\mathbf{w}_1(\mathbf{y},t) = \int_0^t \left(\mathbf{f}_1(\mathbf{y},\tau) - \mathbf{g}_1(\mathbf{y})\right)\mathrm{d}\tau + \mathbf{C}_1(\mathbf{y}), \tag{14}$$

where $\mathbf{C}_1(\mathbf{y})$ is an arbitrary t-independent and $\mathbf{y}$-dependent constant of integration.

2.1.2 Higher-order averaging

The previous reasoning can be easily extended to higher orders. In fact, the terms $\mathbf{g}_m(\mathbf{y},t)$ are obtained by differentiating both sides of Eq. (5) with respect to time, substituting the ϵ-expanded Eqs. (1) and (6) in the resulting expression, and finally equating terms of like power of ϵ. We thus obtain expressions equivalent to Eq. (12)

$$\mathbf{g}_m(\mathbf{y},t) = \mathbf{f}_0^{(m)}(\mathbf{y},t) - \frac{\partial\mathbf{w}_m(\mathbf{y},t)}{\partial t}. \tag{15}$$

Equation (15) is also known as the *homological equation* [17]. Proceeding analogously, we make $\mathbf{g}_m(\mathbf{y}, t)$ time-invariant by taking

$$\mathbf{g}_m(\mathbf{y}) = \bar{\mathbf{f}}_0^{(m)}(\mathbf{y}) = \frac{1}{T}\int_0^T \mathbf{f}_0^{(m)}(\mathbf{y}, t)\,\mathrm{d}t, \tag{16}$$

and the higher-order terms of the transformation $\mathbf{w}$ are given, similarly to Eq. (14) as

$$\mathbf{w}_m(\mathbf{y}, t) = \int_0^t \left(\mathbf{f}_0^{(m)}(\mathbf{y}, \tau) - \mathbf{g}_m(\mathbf{y})\right)\mathrm{d}\tau + \mathbf{C}_m(\mathbf{y}). \tag{17}$$

It is important to highlight that in Eqs. (15), (16) and (17), the functions $\mathbf{f}_0^{(m)}(\mathbf{y}, t)$ are not to be confused with the functions $\mathbf{f}_m(\mathbf{y}, t)$ previously defined in Eq. (4). In fact, as previously described, the functions $\mathbf{f}_0^{(m)}(\mathbf{y}, t)$ stem from equating coefficients of like power of ϵ and can be computed recursively using the following algorithm. Given $\mathbf{f}_0^{(1)} = \mathbf{f}_1^{(0)} = \mathbf{f}_1$ and $\mathbf{g}_1$ computed as in Eq. (13), then the higher-order averaged vector fields $\mathbf{g}_m$ for $m > 1$ can be computed following Algorithm 1. In Algorithm 1, we dropped the argument of $\mathbf{f}$, $\mathbf{g}$ and $\mathbf{w}$ for the sake of notation clarity. Moreover, $A_i^k = k!/((k-i)!i!)$ are the binomial coefficients and $[\cdot,\cdot]$ represents the *Lie bracket* between vector fields. The standard results in the literature [15,33] are written in this section in terms of Lie brackets to facilitate the reconciliation with the other averaging approach based on chronological calculus. The Lie bracket $[\mathbf{X}, \mathbf{Y}]$ between the two vector fields $\mathbf{X}(\mathbf{x}, t), \mathbf{Y}(\mathbf{x}, t) \in \mathbb{R}^n$ is defined as (see [21,35–37])

$$[\mathbf{X}, \mathbf{Y}] = \frac{\partial \mathbf{Y}}{\partial \mathbf{x}}\mathbf{X} - \frac{\partial \mathbf{X}}{\partial \mathbf{x}}\mathbf{Y}. \tag{18}$$

Algorithm 1 Computing $\mathbf{g}_m$ for $m > 1$.

For $k = 2, \ldots, m$:

$$\mathbf{f}_k^{(0)} = \mathbf{f}_k,$$

$$\mathbf{w}_{k-1} = \int_0^t \left(\mathbf{f}_0^{(k-1)} - \mathbf{g}_{k-1}\right)\mathrm{d}\tau + \mathbf{C}_{k-1},$$

$$\mathbf{f}_{k-1}^{(1)} = \mathbf{f}_k^{(0)} + \sum_{i=0}^{k-2} A_i^{k-1}\left[\mathbf{w}_{i+1}, \mathbf{f}_{k-i-1}^{(0)}\right],$$

$$\mathbf{f}_{k-j}^{(j)} = \mathbf{f}_{k-j+1}^{(j-1)} + \sum_{i=0}^{k-j} A_i^{k-j} \times \left[\mathbf{w}_{i+1}, \mathbf{f}_{k-j-i}^{(j-1)} - \mathbf{f}_0^{(k-i-1)} + \mathbf{g}_{k-i-1}\right], \quad j = 2, \ldots, k,$$

$$\mathbf{g}_k = \frac{1}{T}\int_0^T \mathbf{f}_0^{(k)}\,\mathrm{d}t.$$

To compute a certain $\mathbf{g}_m$, one only needs to know the power expansions $\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_m$ in Eq. (4) and to set the integration constants $\mathbf{C}_1, \mathbf{C}_2, \ldots, \mathbf{C}_{m-1}$ in Algorithm 1. The constants of integration $\mathbf{C}_m(\mathbf{y})$ are independent of time, but are dependent on the state $\mathbf{y}$; therefore, different choices of such constants can lead to significantly different results. The selection of $\mathbf{C}_m$ is arbitrary; there are typically three main choices that are found in the literature:

1. *Stroboscopic Averaging*:

$$\mathbf{C}_m = \mathbf{0}. \tag{19}$$

   By choosing the stroboscopic averaging, the near-identity transformation (5) reduces to the identity when $t = nT$, with $n \in \mathbb{Z}$. In other words, if one were to observe, using a stroboscopic light flashing at these times, the solutions of the exact system (1) and the transformed system (21) (with the same initial conditions), they would appear to coincide. For elegance and simplicity of computations, this is the most common choice for higher-order averaging.
2. *Zero-mean* $\mathbf{w}_m$:

$$\mathbf{C}_m = -\frac{1}{T}\int_0^T \left(\int_0^\tau \tilde{\mathbf{f}}_m \mathrm{d}\tau\right)\mathrm{d}t. \tag{20}$$

   By choosing the constants of integration as in Eq. (20), the periodic functions $\mathbf{w}_m$ will have zero mean over the period T. In other words, $\frac{1}{T}\int_0^T \mathbf{w}_m \mathrm{d}t = \mathbf{0}$. This choice is preferable when $\mathbf{f}_m$ are written as Fourier series so that $\mathbf{w}_m$ maintain the same structure (see [18, Ch. 6]).
3. *Reduced Averaging*: $\mathbf{C}_m$ are chosen so that all the functions $\mathbf{g}_m$ vanish for $m > 1$. In other words, the dynamics is not captured by the averaged system, but only via the near-identity transformation. (see [17, Ch. 3]).

The transformed system (6), when all the vector fields $\mathbf{g}_m$ are computed as in Eq. (13) and Algorithm 1 (i.e., $\mathbf{g}_m$ are all time-invariant), becomes a NLTI system

$$\dot{\mathbf{y}} = \sum_{m=1}^{\infty} \frac{\epsilon^m}{m!}\mathbf{g}_m(\mathbf{y}). \tag{21}$$

After proving equivalence with the complete averaging of Sarychev [38] and Vela [27] in the next section, it will be clearly seen that system (21) represents the *complete averaging* of the NLTP system (1). That is, system (21) is an exact transformation of the original system (1): No approximation has been introduced yet; if $\mathbf{y}(t)$ is

the solution of system (21), one can exactly recover the solution $\mathbf{x}(t)$ of the original system (1) via the Lie transformation (5). However, when the summation is truncated after $m = r$,

$$\dot{\overline{\mathbf{y}}} = \sum_{m=1}^{r} \frac{\epsilon^m}{m!} \mathbf{g}_m(\overline{\mathbf{y}}), \tag{22}$$

the following theorem relates the solution of the rth averaged system (22) and the complete averaged system (21).

Theorem 2 (Theorem 6.5.2 in [18]) *Let* $\mathbf{y}(t)$ *and* $\overline{\mathbf{y}}(t)$ *be the solutions of* (21) *and* (22)*, respectively. There exist positive constants* K *and* ϵ_0 *such that, for all* $\epsilon \leq \epsilon_0$

$$|\mathbf{y}(t) - \overline{\mathbf{y}}(t)| = \mathcal{O}(\epsilon^r),$$

for all $0 \leq t \leq K/\epsilon$.

Therefore, we can express the solution of the original system (1) as

$$\mathbf{x}(t) = \overline{\mathbf{y}}(t) + \sum_{m=1}^{r-1} \frac{\epsilon^m}{m!} \mathbf{w}_m(\overline{\mathbf{y}}(t), t) + \mathcal{O}(\epsilon^r). \tag{23}$$

For the convenience of the reader, we report the first four terms of (22).

$$\mathbf{g}_1(\mathbf{y}) = \frac{1}{T} \int_0^T \mathbf{f}_1 \mathrm{d}t, \tag{24}$$

$$\mathbf{g}_2(\mathbf{y}) = \frac{1}{T} \int_0^T \left(\mathbf{f}_2 + [\mathbf{w}_1, \mathbf{f}_1] + \left[\mathbf{w}_1, \mathbf{g}_1\right] \right) \mathrm{d}t, \tag{25}$$

$$\begin{aligned} \mathbf{g}_3(\mathbf{y}) = \frac{1}{T} \int_0^T & \big(\mathbf{f}_3 + [\mathbf{w}_1, \mathbf{f}_2] + 2[\mathbf{w}_2, \mathbf{f}_1] + 2[\mathbf{w}_1, \mathbf{g}_2] \\ & + [\mathbf{w}_1, [\mathbf{g}_1, \mathbf{w}_1]] + [\mathbf{w}_2, \mathbf{g}_1] \big) \mathrm{d}t, \end{aligned} \tag{26}$$

$$\begin{aligned} \mathbf{g}_4(\mathbf{y}) = \frac{1}{T} \int_0^T & \big(\mathbf{f}_4 + [\mathbf{w}_1, \mathbf{f}_3] + 3[\mathbf{w}_2, \mathbf{f}_2] \\ & + [\mathbf{w}_1, [\mathbf{w}_1, [\mathbf{w}_1, \mathbf{g}_1]]] + \\ & - [\mathbf{w}_1, [\mathbf{w}_2, \mathbf{g}_1]] - 2[\mathbf{w}_2, [\mathbf{w}_1, \mathbf{g}_2]] \\ & - 3[\mathbf{w}_1, [\mathbf{w}_1, \mathbf{g}_2]] \\ & + [\mathbf{w}_3, \mathbf{g}_1] + 3[\mathbf{w}_2, \mathbf{g}_2] + 3[\mathbf{w}_1, \mathbf{g}_3] \\ & + 3[\mathbf{w}_3, \mathbf{f}_1]] \big) \mathrm{d}t, \end{aligned} \tag{27}$$

where

$$\begin{aligned} \mathbf{w}_1(\mathbf{y}, t) &= \int_0^t (\mathbf{f}_1 - \mathbf{g}_1) \mathrm{d}\tau + \mathbf{C}_1, \\ \mathbf{w}_2(\mathbf{y}, t) &= \int_0^t (\mathbf{f}_2 + [\mathbf{w}_1, \mathbf{f}_1] \\ &\quad + [\mathbf{w}_1, \mathbf{g}_1] - \mathbf{g}_2) \mathrm{d}\tau + \mathbf{C}_2, \\ \mathbf{w}_3(\mathbf{y}, t) &= \int_0^t \big(\mathbf{f}_3 + [\mathbf{w}_1, \mathbf{f}_2] + 2[\mathbf{w}_2, \mathbf{f}_1] \\ &\quad + 2[\mathbf{w}_1, \mathbf{g}_2] \\ &\quad - [\mathbf{w}_1, [\mathbf{w}_1, \mathbf{g}_1]] + [\mathbf{w}_2, \mathbf{g}_1] - \mathbf{g}_3 \big) \mathrm{d}\tau \\ &\quad + \mathbf{C}_3, \end{aligned} \tag{28}$$

and $\mathbf{f}_1 = \mathbf{f}(\mathbf{y}, t, 0)$, $\mathbf{f}_2 = 2\, \partial \mathbf{f}/\partial \epsilon|_{\epsilon=0}$, $\mathbf{f}_3 = 3\, \partial^2 \mathbf{f}/\partial \epsilon^2|_{\epsilon=0}$, and $\mathbf{f}_4 = 4\, \partial^3 \mathbf{f}/\partial \epsilon^3|_{\epsilon=0}$.

2.2 Higher-order averaging via chronological calculus

In this subsection, we present the higher-order averaging approach based on *chronological calculus*, which was developed in the late 1970s by the Russian mathematicians Agrachev and Gamkrelidze [28,39]. Chronological calculus is concerned with time-varying vector fields and the asymptotic expansion generated by their flows. Sarychev and Vela [27,31,34] utilized the chronological calculus tools to develop a generalized averaging theory (GAT), which gives arbitrarily high-order approximation of the flow along a time-periodic vector field. We provide below a brief description of the technique.

Consider the NLTP dynamics

$$\dot{\mathbf{x}} = \epsilon \mathbf{f}(\mathbf{x}, t). \tag{29}$$

Equation (29) is similar to (1) with the only difference that $\mathbf{f}$ does not depend explicitly on ϵ.[2] The vector field $\mathbf{f}$ is smooth in $\mathbf{x}$ and T-periodic in t. The solution of the differential Eq. (29) for the initial condition $\mathbf{x}(0) = \mathbf{x}_0$ is denoted by $\mathbf{x}(t) = \boldsymbol{\phi}(\mathbf{x}_0, t) = \boldsymbol{\phi}_t^{\epsilon \mathbf{f}}(\mathbf{x}_0)$, where $\boldsymbol{\phi}_t^{\epsilon \mathbf{f}} : \mathbb{R}^n \to \mathbb{R}^n$ is the t-dependent flow associated with the vector field $\epsilon \mathbf{f}(\mathbf{x}, t)$ and $\boldsymbol{\phi}_0^{\epsilon \mathbf{f}}(\mathbf{x}_0) = \boldsymbol{\phi}(\mathbf{x}_0, 0) = \mathbf{x}_0$. Also, $\boldsymbol{\phi}_t^{\epsilon \mathbf{f}}$ is a diffeomorphism satisfying the differential equation $\dot{\boldsymbol{\phi}}_t^{\epsilon \mathbf{f}} = \boldsymbol{\phi}_t^{\epsilon \mathbf{f}} \circ \epsilon \mathbf{f}(\mathbf{x}, t)$, where "$\circ$" indicates the composition of maps (functions).

[2] Although Sarychev [34] acknowledges the case of explicit dependence of $\mathbf{f}$ on ϵ, it is the authors' opinion that the analysis of such a case needs further investigation. Hence, we restrict our demonstration to systems in the form (29).

We can write the solution of Eq. (29) using the Volterra series expansion as

$$\mathbf{x}(t) = \mathbf{x}_0 + \sum_{m=1}^{\infty} \int_0^t \int_0^{\tau_1} \cdots \int_0^{\tau_{m-1}} \mathcal{Y}(\mathbf{x}_0)\, \mathrm{d}\tau_m \ \ldots \ \mathrm{d}\tau_1 \tag{30}$$

where $\mathcal{Y} = \left(\mathbf{Y}_{\tau_1} \circ \mathbf{Y}_{\tau_2} \circ \cdots \circ \mathbf{Y}_{\tau_m}\right)$. Here, $\mathbf{Y}_\tau = \epsilon \mathbf{f}(\mathbf{x}, \tau)$ and $\boldsymbol{Y}_{\tau_1} \circ \boldsymbol{Y}_{\tau_2} = \frac{\partial \boldsymbol{Y}_{\tau_2}}{\partial \boldsymbol{x}} \boldsymbol{Y}_{\tau_1}$. The conditions of convergence of series (30) are provided by Agrachev and Gamkrelidze [29]. In a similar fashion to the exponential representation of solutions of time-invariant systems [30,40], Agrachev and Gamkrelidze [29] expressed the flow of the time-varying vector field $\epsilon \mathbf{f}(\mathbf{x}, t)$ as

$$\boldsymbol{\phi}_t^{\epsilon \mathbf{f}} = \overrightarrow{\exp} \left(\int_0^t \epsilon \mathbf{f}(\mathbf{x}, \tau)\, d\tau \right). \tag{31}$$

In the chronological calculus, this flow is also known as the *right-chronological exponential*.

Then, the *logarithm* of (31) is defined as

$$\mathbf{V}_t = \ln(\boldsymbol{\phi}_{\mathrm{t}}^{\epsilon \mathbf{f}}).$$

In other words, the flow along the time-invariant vector field $\mathbf{V}_t$ for a unit time is equivalent to the flow along the time-varying vector field $\epsilon \mathbf{f}(\mathbf{x}, t)$ for a time t.

It is noteworthy to mention that although $\mathbf{V}_t$ is an autonomous vector field, it is parametrized by time; if the final time t is changed, the vector field will change. Furthermore, in Ref. [29] it was shown that $\boldsymbol{V}_t = \sum_{m=1}^{\infty} \boldsymbol{V}_t^{(m)}$, with

$$\mathbf{V}_t^{(m)} = \int_0^t \int_0^{\tau_1} \cdots \int_0^{\tau_{m-1}} \mathcal{G}_m(\mathbf{Y}_{\tau_1}, \ldots, \mathbf{Y}_{\tau_m}) \mathrm{d}\tau_m \ldots \mathrm{d}\tau_1,$$

and $\mathcal{G}_m$ are the commutators polynomials. For instance, for $m = 1, 2, 3$, we obtain

$$\mathcal{G}_1(\zeta_1) = \zeta_1, \quad \mathcal{G}_2(\zeta_1, \zeta_2) = \frac{1}{2}[\zeta_2, \zeta_1]$$

$$\mathcal{G}_3(\zeta_1, \zeta_2, \zeta_3) = \frac{1}{6}\left([\zeta_3, [\zeta_2, \zeta_1]] + [[\zeta_3, \zeta_2], \zeta_1]\right),$$

where $[\cdot, \cdot]$ denotes the commutator, which, for vector fields, coincides with the Lie bracket. Therefore, chronological calculus provides an algorithmic approach to determine the logarithm of time-periodic vector fields analytically.

To better understand the concept of GAT, we now recall some key results from the linear Floquet theory. Consider the homogeneous LTP system,

$$\dot{\mathbf{x}} = \mathbf{A}(t)\mathbf{x}, \tag{32}$$

with $\mathbf{A} \in \mathbb{R}^{n \times n}$ and T-periodic in t. The flow associated with (32) can be represented by the so-called *fundamental solution matrix*, written as

$$\boldsymbol{\Phi}_t^{\mathbf{A}} = \overrightarrow{\exp} \left(\int_0^t \mathbf{A}(\tau) \mathrm{d}\tau \right) \in \mathbb{R}^{n \times n}.$$

The fundamental solution matrix of system (32), computed after one period T, is known as the *monodromy matrix* $\mathbf{M} = \boldsymbol{\Phi}_T^{\mathbf{A}}$.

Theorem 3 Linear Floquet Theorem [41, p. 11] *Every fundamental matrix solution $\boldsymbol{\Phi}_t^{\mathbf{A}}$ of the LTP system (32) can be represented as the product $\boldsymbol{\Phi}_t^{\mathbf{A}} = \mathbf{P}(t)e^{\mathbf{R}t}$ of a T-periodic matrix (i.e., $\mathbf{P}(t) = \mathbf{P}(t + T)$) and $\mathbf{R}$ is a constant matrix given by $\mathbf{R} = \frac{1}{T} \ln \mathbf{M}$.*

Theorem 4 *System (32) is uniformly stable if and only if all Floquet multipliers (eigenvalues of $\mathbf{M}$) have moduli less than 1.*

Based on Theorem 3, one can use the transformation $\mathbf{x} = \mathbf{P}\mathbf{y}$ to transform the LTP system (32) to the LTI system

$$\dot{\mathbf{y}} = \mathbf{R}\mathbf{y} \tag{33}$$

whose flow after one period is exactly equivalent to the flow of the original LTP system (32) after one period: $e^{\mathbf{R}T} = \boldsymbol{\Phi}_T^{\mathbf{A}}$. That is, based on our discussion in the previous section on classical averaging, system (33) represents a stroboscopic averaged dynamics of the LTP system (32). Moreover, one can determine the stability of the LTP system (32) from the Floquet multipliers according to Theorem 4, or from the averaged dynamics (33): The solution $\mathbf{x}(t)$ of (32) converges to the origin if and only if the solution $\mathbf{y}(t)$ of (33) converges to zero. Therefore, the LTP system (32) is stable if and only if the averaged dynamics (33) is stable (i.e., the eigenvalues of the $\mathbf{R}$ lie in the open left half of the complex plane); there is no approximation here. It is complete averaging. The eigenvalues of $\mathbf{R}$ are called the Floquet exponents, or the characteristic exponents.

Sarychev [38] utilized the Floquet concept of averaging without approximation to extend the Floquet theory to the nonlinear case, equivalently extending the averaging theorem to higher order. Similar to the

Table 1 Linear versus nonlinear Floquet theory

Linear	Nonlinear
$\dot{\mathbf{x}} = \mathbf{A}(t)\mathbf{x}$	$\dot{\mathbf{x}} = \epsilon\mathbf{f}(\mathbf{x}, t)$
$\boldsymbol{\Phi}_t^{\mathbf{A}} = \overrightarrow{\exp}\left(\int_0^t \mathbf{A}(\tau)\mathrm{d}\tau\right)$	$\boldsymbol{\phi}_t^{\epsilon\mathbf{f}} = \overrightarrow{\exp}\left(\int_0^t \epsilon\mathbf{f}(\mathbf{x}, \tau)\, d\tau\right)$
$\mathbf{M} = \boldsymbol{\Phi}_T^{\mathbf{A}}$	$\mathcal{M} = \phi_T^{\epsilon\mathbf{f}}$
$\mathbf{R} = \frac{1}{T}\ln\mathbf{M}$	$\mathbf{V}_T = \frac{1}{T}\ln\mathcal{M}$
$\dot{\mathbf{y}} = \mathbf{R}\mathbf{y}$	$\dot{\bar{\mathbf{x}}} = (\bar{\mathbf{x}}) = \sum_{m=1}^{\infty} \frac{\epsilon^m}{m!} m(\bar{\mathbf{x}})$

definition of the averaged dynamics $\mathbf{R} = \frac{1}{T}\ln\mathbf{M} = \frac{1}{T}\ln\boldsymbol{\Phi}_T^{\mathbf{A}} = \frac{1}{T}\ln\overrightarrow{\exp}\int_0^T \mathbf{A}(t)\mathrm{d}t$, Sarychev [38] introduced the notion of *complete averaging* of the NLTP system (29) to indicate the autonomous vector field

$$\boldsymbol{\Lambda} = \frac{1}{T}\mathbf{V}_T = \frac{1}{T}\ln\boldsymbol{\phi}_T^{\epsilon\mathbf{f}} = \frac{1}{T}\ln\boldsymbol{\mathcal{M}}, \tag{34}$$

where $\boldsymbol{\mathcal{M}} = \boldsymbol{\phi}_T^{\epsilon\mathbf{f}}$ is the flow of $\epsilon\mathbf{f}(\mathbf{x}, t)$ computed after one period T and is called the *monodromy map*, which maps $\mathbf{x}_0$ to $\mathbf{x}(T)$ (i.e., the solution after a period T). The monodromy map can be seen as the nonlinear analogous to the monodromy matrix $\mathbf{M}$ in the linear Floquet decomposition. The vector field $\boldsymbol{\Lambda}$ is then equivalent to the matrix $\mathbf{R}$. Furthermore, the linear Floquet theorem can also be extended to the nonlinear case as follows

Theorem 5 Nonlinear Floquet Theorem (Theorem 3.2 in [34] and Theorem 8 in [27]) *Assume that $\boldsymbol{\phi}_t^{\epsilon\mathbf{f}}$ is the flow generated by the NLTP system in (29). If the diffeomorphism $\boldsymbol{\mathcal{M}} = \boldsymbol{\phi}_T^{\epsilon\mathbf{f}}$ admits a logarithm $\mathbf{V}_T$, then the flow $\boldsymbol{\phi}_t^{\epsilon\mathbf{f}}$ can be represented as a composition $\boldsymbol{\phi}_t^{\epsilon\mathbf{f}} = e^{\boldsymbol{\Lambda}t} \circ \mathcal{P}_t$ of the flow $e^{\boldsymbol{\Lambda}t}$ of the autonomous vector field $\boldsymbol{\Lambda} = \frac{1}{T}\mathbf{V}_T$ and a T-periodic map $\mathcal{P}_t$ (i.e., $\mathcal{P}_t = \mathcal{P}_{t+T}$).*

Moreover, the stability characteristics of this logarithm $\boldsymbol{\Lambda}$ are exactly related to the stability characteristics of the original NLTP system (29) in a Floquet theorem fashion.

Theorem 6 (Theorem 9 in [27]). *If the monodromy map, $\boldsymbol{\mathcal{M}}$, of system (29) has a fixed point, then the flow, $\boldsymbol{\phi}_t^{\epsilon\mathbf{f}}$, has a periodic orbit whose stability is determined by the stability of the monodromy map.*

Theorem 6 clearly represents the nonlinear extension of Theorem 4. Table 1 summarizes the parallel between the linear and the nonlinear Floquet theorems.

The vector field $\boldsymbol{\Lambda}$ in Eq. (34) can be written as a power series expansion in ϵ such that

$$\dot{\bar{\mathbf{x}}} = \boldsymbol{\Lambda}(\bar{\mathbf{x}}) = \sum_{m=1}^{\infty} \frac{\epsilon^m}{m!}\boldsymbol{\Lambda}_m(\bar{\mathbf{x}}), \tag{35}$$

where $\bar{\mathbf{x}}$ is the average of $\mathbf{x}$. According to Theorem 5, the solution of (29) oscillates around the solution of the averaged NLTI system (35). Finally, if the series in Eq. (35) is truncated after r terms, one arrives at the rth-order averaged version

$$\dot{\mathbf{y}} = \sum_{m=1}^{r} \frac{\epsilon^m}{m!}\boldsymbol{\Lambda}_m(\mathbf{y}). \tag{36}$$

The power of this generalization is that the $\boldsymbol{\Lambda}_m$ terms can be computed analytically using Lie brackets between the vector fields that describe the time-periodic dynamics.

For the convenience of the reader, we report the first four truncations as seen in Refs. [8,27,34].

$$\boldsymbol{\Lambda}_1(\mathbf{y}) = \frac{1}{T}\int_0^T \mathbf{f}\,\mathrm{d}t, \tag{37}$$

$$\boldsymbol{\Lambda}_2(\mathbf{y}) = \frac{1}{T}\int_0^T \left[\int_0^t \mathbf{f}_\tau\,\mathrm{d}\tau\,,\,\mathbf{f}\right]\mathrm{d}t, \tag{38}$$

$$\begin{aligned}\boldsymbol{\Lambda}_3(\mathbf{y}) &= -\frac{3}{2}T\,[\boldsymbol{\Lambda}_1, \boldsymbol{\Lambda}_2] \\ &\quad + \frac{2}{T}\int_0^T \left[\int_0^t \mathbf{f}_\tau\,\mathrm{d}\tau\,,\,\left[\int_0^t \mathbf{f}_\tau\,\mathrm{d}\tau,\,\mathbf{f}\right]\right]\mathrm{d}t \\ &= -\frac{3}{2}T\left[\frac{1}{T}\int_0^T \mathbf{f}\,\mathrm{d}t,\,\frac{1}{T}\int_0^T \left[\int_0^t \mathbf{f}_\tau\,\mathrm{d}\tau\,,\mathbf{f}\right]\mathrm{d}t\right] \\ &\quad + \frac{2}{T}\int_0^T \left[\int_0^t \mathbf{f}_\tau\,\mathrm{d}\tau\,,\,\left[\int_0^t \mathbf{f}_\tau\,\mathrm{d}\tau,\,\mathbf{f}\right]\right]\mathrm{d}t.\end{aligned} \tag{39}$$

$$\begin{aligned}\boldsymbol{\Lambda}_4(\mathbf{y}) &= \frac{2}{T}\int_0^T \Bigg(\int_0^t \left[\int_0^\tau \left[\int_0^{\tau_1} \mathbf{f}_{\tau_2}\mathrm{d}\tau_2, \mathbf{f}_{\tau_1}\right]\mathrm{d}\tau_1, [\mathbf{f}_\tau, \mathbf{f}]\right]\mathrm{d}\tau \\ &\quad + \left[\int_0^t \left[\int_0^\tau \left[\int_0^{\tau_1} \mathbf{f}_{\tau_2}\mathrm{d}\tau_2, \mathbf{f}_{\tau_1}\right]\mathrm{d}\tau_1, \mathbf{f}_\tau\right]\mathrm{d}\tau, \mathbf{f}\right] \\ &\quad + \int_0^t \left[\int_0^\tau \mathbf{f}_{\tau_1}\mathrm{d}\tau_1, \left[\left[\int_0^\tau \mathbf{f}_{\tau_1}\mathrm{d}\tau_1, \mathbf{f}_\tau\right], \mathbf{f}\right]\right]\mathrm{d}\tau\Bigg)\mathrm{d}t,\end{aligned} \tag{40}$$

where $\mathbf{f} = \mathbf{f}(\mathbf{y}, t)$, $\mathbf{f}_\tau = \mathbf{f}(\mathbf{y}, \tau)$, $\mathbf{f}_{\tau_1} = \mathbf{f}(\mathbf{y}, \tau_1)$, and $\mathbf{f}_{\tau_2} = \mathbf{f}(\mathbf{y}, \tau_2)$. It is worth mentioning that the expressions for $\boldsymbol{\Lambda}_3$ and $\boldsymbol{\Lambda}_4$ are not agreed upon between all cited references. In particular, there is a mismatch in the sign of the term $[\boldsymbol{\Lambda}_1, \boldsymbol{\Lambda}_2]$ in Eq. (39) and in the sign of the entire expression of $\boldsymbol{\Lambda}_4$ in Eq. (40). Based on our analysis, we believe that the expressions for $\boldsymbol{\Lambda}_3$ in Eq. (39) and for $\boldsymbol{\Lambda}_4$ in Eq. (40) are the correct ones. We will provide further evidence in the next section.

3 Reconciliation of the two higher-order averaging approaches

In this section we will show that, under some key assumptions, the two higher-order methodologies can be proven equivalent despite the fact that they were developed following very different and independent derivations. We prove equivalence mathematically up to the fourth-order averaging ($r = 4$) under the following assumptions:

Assumption 1 Assume that the NLTP system is written in the form (29)

$$\dot{\mathbf{x}} = \epsilon \mathbf{f}(\mathbf{x}, t),$$

with $\mathbf{f}$ smooth in $\mathbf{x}$ and T-periodic in t.

Assumption 2 (Stroboscopic averaging) Assume that the arbitrary constants of integration in Eq. (17) and Algorithm 1 are taken to be null

$$\mathbf{C}_m = \mathbf{0}, \quad \text{for} \quad m \geq 1.$$

With Assumption 1, we reduce the analysis to the class of systems in which the parameter ϵ does not appear explicitly in $\mathbf{f}$. As a result, Eq. (3) becomes $\mathbf{f}(\mathbf{x}, t, \epsilon) = \mathbf{f}(\mathbf{x}, t) = \mathbf{f}_1(\mathbf{x}, t)$ (i.e., $\mathbf{f}_m(\mathbf{x}, t) = \mathbf{0}$ for $m > 1$) and system (1) is in the form (29). Otherwise, the chronological calculus technique would not be even applicable. The second assumption renders the first approach a fundamental concept similar to the chronological calculus approach: The solution of the averaged system coincides with that of the original system stroboscopically (after each cycle). This concept was the basis of the chronological calculus averaging: The goal was to find an autonomous vector field whose flow after one period matches that of the time-periodic vector field after one period. Therefore, while we provide below a mathematical proof for the equivalence between the two approaches up to fourth order, this conceptual matching suggests that the equivalence can be extended to all higher orders.

First, we state the following properties of Lie brackets which will be useful in reconciling the two higher-order averaging techniques. Let $\mathbf{X}(\mathbf{x}, t)$, $\mathbf{Y}(\mathbf{x}, t)$, $\mathbf{Z}(\mathbf{x}, t) \in \mathbb{R}^n$ and $\alpha \in \mathbb{R}$. The Lie brackets, as defined in Eq. (18), satisfy the following properties:

1. $[\mathbf{X}, \mathbf{Y}] = -[\mathbf{Y}, \mathbf{X}]$ (*skew-symmetry*),
2. $[\mathbf{X} + \mathbf{Y}, \alpha\mathbf{Z}] = \alpha[\mathbf{X}, \mathbf{Z}] + \alpha[\mathbf{Y}, \mathbf{Z}]$ (*linearity*),
3. $[[\mathbf{X}, \mathbf{Y}], \mathbf{Z}] + [[\mathbf{Y}, \mathbf{Z}], \mathbf{X}] + [[\mathbf{Z}, \mathbf{X}], \mathbf{Y}] = \mathbf{0}$ (*Jacobi identity*).

Lemma 1 *Given two smooth vector fields* $\mathbf{X}(\mathbf{x}, t)$, $\mathbf{Y}(\mathbf{x}, t) \in \mathbb{R}^n$, *then*

$$\frac{\partial}{\partial t}[\mathbf{X}, \mathbf{Y}] = \left[\mathbf{X}, \frac{\partial \mathbf{Y}}{\partial t}\right] + \left[\frac{\partial \mathbf{X}}{\partial t}, \mathbf{Y}\right], \tag{41}$$

and if $\mathbf{X}$ *is time-invariant, the following are true*

$$\frac{\partial}{\partial t}[\mathbf{X}, \mathbf{Y}] = \left[\mathbf{X}, \frac{\partial \mathbf{Y}}{\partial t}\right], \tag{42}$$

and

$$\int [\mathbf{X}, \mathbf{Y}] \mathrm{d}t = \left[\mathbf{X}, \int \mathbf{Y} \mathrm{d}t\right].$$

Proof Using Eq. (18), it follows that

$$\begin{aligned}
&\frac{\partial}{\partial t}[\mathbf{X}, \mathbf{Y}] \\
&= \frac{\partial}{\partial t}\left(\frac{\partial \mathbf{Y}}{\partial \mathbf{x}}\mathbf{X} - \frac{\partial \mathbf{X}}{\partial \mathbf{x}}\mathbf{Y}\right), \\
&= \frac{\partial}{\partial t}\left(\frac{\partial \mathbf{Y}}{\partial \mathbf{x}}\right)\mathbf{X} + \frac{\partial \mathbf{Y}}{\partial \mathbf{x}}\frac{\partial \mathbf{X}}{\partial t} - \frac{\partial}{\partial t}\left(\frac{\partial \mathbf{X}}{\partial \mathbf{x}}\right)\mathbf{Y} - \frac{\partial \mathbf{X}}{\partial \mathbf{x}}\frac{\partial \mathbf{Y}}{\partial t}, \\
&= \frac{\partial}{\partial \mathbf{x}}\left(\frac{\partial \mathbf{Y}}{\partial t}\right)\mathbf{X} - \frac{\partial \mathbf{X}}{\partial \mathbf{x}}\frac{\partial \mathbf{Y}}{\partial t} + \frac{\partial \mathbf{Y}}{\partial \mathbf{x}}\frac{\partial \mathbf{X}}{\partial t} - \frac{\partial}{\partial \mathbf{x}}\left(\frac{\partial \mathbf{X}}{\partial t}\right)\mathbf{Y}, \\
&= \left[\mathbf{X}, \frac{\partial \mathbf{Y}}{\partial t}\right] + \left[\frac{\partial \mathbf{X}}{\partial t}, \mathbf{Y}\right].
\end{aligned}$$

If $\mathbf{X}$ is time-invariant, then $\frac{\partial \mathbf{X}}{\partial t} = \mathbf{0}$ and Eq. (41) reduces to (42). Moreover,

$$\begin{aligned}
\int [\mathbf{X}, \mathbf{Y}] \mathrm{d}t &= \int \left(\frac{\partial \mathbf{Y}}{\partial \mathbf{x}}\mathbf{X} - \frac{\partial \mathbf{X}}{\partial \mathbf{x}}\mathbf{Y}\right) \mathrm{d}t, \\
&= \int \frac{\partial \mathbf{Y}}{\partial \mathbf{x}}\mathbf{X} \mathrm{d}t - \int \frac{\partial \mathbf{X}}{\partial \mathbf{x}}\mathbf{Y} \mathrm{d}t, \\
&= \frac{\partial \left(\int \mathbf{Y} \mathrm{d}t\right)}{\partial \mathbf{x}}\mathbf{X} - \frac{\partial \mathbf{X}}{\partial \mathbf{x}}\int \mathbf{Y} \mathrm{d}t, \\
&= \left[\mathbf{X}, \int \mathbf{Y} \mathrm{d}t\right].
\end{aligned}$$

□

Theorem 7 *Given Assumptions 1 and 2, the following are true:* $\mathbf{g}_1 = \boldsymbol{\Lambda}_1$, $\mathbf{g}_2 = \boldsymbol{\Lambda}_2$, $\mathbf{g}_3 = \boldsymbol{\Lambda}_3$, *and* $\mathbf{g}_4 = \boldsymbol{\Lambda}_4$.

Proof Given in the following Subsect. 3.1–3.4. □

3.1 First order

By comparing Eqs. (24) and (37), it trivially follows that $\mathbf{g}_1 = \boldsymbol{\Lambda}_1$. This equivalence holds also without Assumption 2.

3.2 Second order

For the second-order terms, we start by expanding $\mathbf{g}_2$ given in Eq. (25) as

$$\begin{aligned}\mathbf{g}_2 =& \frac{1}{T}\int_0^T \Bigg(\mathbf{f}_2 + \left[\int_0^t \mathbf{f}_1\,\mathrm{d}\tau,\ \mathbf{f}_1\right] - \left[\int_0^t \mathbf{g}_1\,\mathrm{d}\tau\ ,\ \mathbf{f}_1\right] \\ &+ \left[\int_0^t \mathbf{f}_1\,\mathrm{d}\tau\ ,\ \mathbf{g}_1\right] - \left[\int_0^t \mathbf{g}_1\,\mathrm{d}\tau\ ,\ \mathbf{g}_1\right] \\ &+ [\mathbf{C}_1, \mathbf{f}_1] + [\mathbf{C}_1, \mathbf{g}_1]\Bigg)\,\mathrm{d}t.\end{aligned} \tag{43}$$

From Assumption 1, we have $\mathbf{f}_2 = \mathbf{0}$ and $\mathbf{f}_1 = \mathbf{f}$, and from Assumption 2, $\mathbf{C}_1 = \mathbf{0}$. Equation (43) then becomes

$$\begin{aligned}\mathbf{g}_2 =& \frac{1}{T}\int_0^T \Bigg(\left[\int_0^t \mathbf{f}\,\mathrm{d}\tau,\ \mathbf{f}\right] - \left[\int_0^t \mathbf{g}_1\,\mathrm{d}\tau\ ,\ \mathbf{f}\right] \\ &+ \left[\int_0^t \mathbf{f}\,\mathrm{d}\tau\ ,\ \mathbf{g}_1\right] - \left[\int_0^t \mathbf{g}_1\,\mathrm{d}\tau\ ,\ \mathbf{g}_1\right]\Bigg)\,\mathrm{d}t.\end{aligned}$$

Since $\mathbf{g}_1$ is time-invariant, Lemma 1 implies

$$\left[\int_0^t \mathbf{g}_1\,\mathrm{d}\tau\ ,\ \mathbf{g}_1\right] = \left[t\,\mathbf{g}_1\ ,\ \mathbf{g}_1\right] = t\left[\mathbf{g}_1\ ,\ \mathbf{g}_1\right] = \mathbf{0}.$$

To recover the second-order averaging using chronological calculus, $\boldsymbol{\Lambda}_2$ in Eq. (38), it remains to show that

$$\frac{1}{T}\int_0^T \left(\left[\int_0^t \mathbf{f}\,\mathrm{d}\tau\ ,\ \mathbf{g}_1\right] - \left[\int_0^t \mathbf{g}_1\,\mathrm{d}\tau\ ,\ \mathbf{f}\right]\right)\,\mathrm{d}t = \mathbf{0}. \tag{44}$$

Using the skew-symmetry property of the Lie bracket (18), the left-hand side of Eq. (44) becomes

$$\frac{1}{T}\int_0^T \left(\left[\int_0^t \mathbf{f}\,\mathrm{d}\tau,\ \mathbf{g}_1\right] + \left[\mathbf{f}, \int_0^t \mathbf{g}_1\,\mathrm{d}\tau,\right]\right)\,\mathrm{d}t.$$

Applying property (41), we obtain

$$\frac{1}{T}\int_0^T \frac{\partial}{\partial t}\left(\left[\int_0^t \mathbf{f}\,\mathrm{d}\tau, \int_0^t \mathbf{g}_1\,\mathrm{d}\tau,\right]\right)\,\mathrm{d}t.$$

Using the fundamental theorem of calculus, the expression above simplifies to

$$\frac{1}{T}\left[\int_0^T \mathbf{f}\,\mathrm{d}t, \int_0^T \mathbf{g}_1\,\mathrm{d}t,\right].$$

Finally, noting that $\int_0^T \mathbf{f}\,\mathrm{d}t = T\mathbf{g}_1$, and $\int_0^T \mathbf{g}_1\,\mathrm{d}t = T\mathbf{g}_1$,

$$\frac{1}{T}\left[T\mathbf{g}_1\ ,\ T\mathbf{g}_1\right] = \mathbf{0}.$$

Hence, the second-order averaging via Lie transform given in Eq. (43) reduces to

$$\mathbf{g}_2 = \frac{1}{T}\int_0^T \left[\int_0^t \mathbf{f}\,\mathrm{d}\tau\ ,\ \mathbf{f}\right]\,\mathrm{d}t. \tag{45}$$

Comparing Eqs. (45) and (38), it is easy to see that $\mathbf{g}_2 = \boldsymbol{\Lambda}_2$.

3.3 Third order

In a similar fashion to the second-order averaging, let us expand $\mathbf{g}_3$ by substituting the values of $\mathbf{w}_1$ and $\mathbf{w}_2$ found in Eq. (28), into Eq. (26).

$$\begin{aligned}\mathbf{g}_3 =& \frac{1}{T}\int_0^T \Bigg(\mathbf{f}_3 + \left[\int_0^t \mathbf{f}_1\mathrm{d}\tau, \mathbf{f}_2\right] - \left[\int_0^t \mathbf{g}_1\mathrm{d}\tau, \mathbf{f}_2\right] \\ &+2\left[\int_0^t \mathbf{f}_2\mathrm{d}\tau, \mathbf{f}_1\right] + \left[\int_0^t \mathbf{f}_2\mathrm{d}\tau, \mathbf{g}_1\right] \\ &+2\left[\int_0^t \left[\int_0^\tau \mathbf{f}_1\mathrm{d}\tau_1, \mathbf{f}_1\right]\mathrm{d}\tau, \mathbf{f}_1\right] \\ &+\left[\int_0^t \left[\int_0^\tau \mathbf{f}_1\mathrm{d}\tau_1, \mathbf{f}_1\right]\mathrm{d}\tau, \mathbf{g}_1\right] + \\ &-2\left[\int_0^t \left[\int_0^\tau \mathbf{g}_1\mathrm{d}\tau_1, \mathbf{f}_1\right]\mathrm{d}\tau, \mathbf{f}_1\right] \\ &-\left[\int_0^t \left[\int_0^\tau \mathbf{g}_1\mathrm{d}\tau_1, \mathbf{f}_1\right]\mathrm{d}\tau, \mathbf{g}_1\right] \\ &+2\left[\int_0^t \left[\int_0^\tau \mathbf{f}_1\mathrm{d}\tau_1, \mathbf{g}_1\right]\mathrm{d}\tau, \mathbf{f}_1\right] \\ &+\left[\int_0^t \left[\int_0^\tau \mathbf{f}_1\mathrm{d}\tau_1, \mathbf{g}_1\right]\mathrm{d}\tau, \mathbf{g}_1\right] \\ &-2\left[\int_0^t \left(\frac{1}{T}\int_0^T \left(\mathbf{f}_2 \right.\right.\right. \\ &\left.\left.\left.+\left[\int_0^t \mathbf{f}_1\mathrm{d}\tau, \mathbf{f}_1\right]\right)\mathrm{d}t\right)\mathrm{d}\tau, \mathbf{f}_1\right] \\ &-\left[\int_0^t \left(\frac{1}{T}\int_0^T \left(\mathbf{f}_2 \right.\right.\right. \\ &\left.\left.\left.+\left[\int_0^t \mathbf{f}_1\mathrm{d}\tau, \mathbf{f}_1\right]\right)\mathrm{d}t\right)\mathrm{d}\tau, \mathbf{g}_1\right] \\ &-\left[\int_0^t \mathbf{f}_1\mathrm{d}\tau, \left[\int_0^t \mathbf{f}_1\mathrm{d}\tau, \mathbf{g}_1\right]\right] \\ &+\left[\int_0^t \mathbf{g}_1\mathrm{d}\tau, \left[\int_0^t \mathbf{f}_1\mathrm{d}\tau, \mathbf{g}_1\right]\right] \\ &+2\left[\int_0^t \mathbf{f}_1\mathrm{d}\tau, \frac{1}{T}\int_0^T \left(\mathbf{f}_2 + \left[\int_0^t \mathbf{f}_1\mathrm{d}\tau, \mathbf{f}_1\right]\right)\mathrm{d}t\right] \\ &-2\left[\int_0^t \mathbf{g}_1\mathrm{d}\tau, \frac{1}{T}\int_0^T \left(\mathbf{f}_2 + \left[\int_0^t \mathbf{f}_1\mathrm{d}\tau, \mathbf{f}_1\right]\right)\mathrm{d}t\right] \\ &+2\left[\int_0^t [\mathbf{C}_1, \mathbf{f}_1]\,\mathrm{d}\tau,\ \mathbf{f}_1\right] - 2\left[\int_0^t [\mathbf{C}_1, \mathbf{g}_1]\,\mathrm{d}\tau, \mathbf{f}_1\right] \\ &+[\mathbf{C}_1, \mathbf{f}_2] + \left[\int_0^t (\mathbf{f}_1 - \mathbf{g}_1)\,\mathrm{d}\tau + 3[\mathbf{C}_1, \mathbf{g}_1]\right] \\ &+\left[\int_0^t [\mathbf{C}_1, \mathbf{g}_1]\,\mathrm{d}\tau\ , \mathbf{g}_1\right] \\ &+\left[\mathbf{C}_1\ , -\left[\int_0^t (\mathbf{f}_1 - \mathbf{g}_1)\,\mathrm{d}\tau,\ \mathbf{g}_1\right]\right] \\ &+2\left[\mathbf{C}_1, \frac{1}{T}\int_0^T \left(\mathbf{f}_2 + \left[\int_0^t \mathbf{f}_1\,\mathrm{d}\tau\ ,\ \mathbf{f}_1\right]\right)\mathrm{d}t\right]\end{aligned}$$

$$+[\mathbf{C}_1, 3[\mathbf{C}_1, \mathbf{g}_1]]\Big)\mathrm{d}t + 3[\mathbf{C}_2, \mathbf{g}_1]. \tag{46}$$

According to Assumption 1, $\mathbf{f} = \mathbf{f}_1$ and $\mathbf{f}_2 = \mathbf{f}_3 = \mathbf{0}$. Assumption 2 leads to $\mathbf{C}_1 = \mathbf{C}_2 = \mathbf{0}$. Therefore, Eq. (46) simplifies to

$$\begin{aligned}
\mathbf{g}_3 =& \frac{2}{T}\int_0^T \left[\int_0^t \left[\int_0^\tau \mathbf{f}\,\mathrm{d}\tau_1, \mathbf{f}\right]\mathrm{d}\tau, \mathbf{f}\right]\mathrm{d}t \\
&- \frac{2}{T}\int_0^T \left[\int_0^t \left[\int_0^\tau \mathbf{g}_1\mathrm{d}\tau_1, \mathbf{f}\right]\mathrm{d}\tau, \mathbf{f}\right]\mathrm{d}t \\
&+ \frac{1}{T}\int_0^T \left[\int_0^t \left[\int_0^\tau \mathbf{f}\,\mathrm{d}\tau_1, \mathbf{f}\right]\mathrm{d}\tau, \mathbf{g}_1\right]\mathrm{d}t \\
&- \frac{1}{T}\int_0^T \left[\int_0^t \left[\int_0^\tau \mathbf{g}_1\mathrm{d}\tau_1, \mathbf{f}\right]\mathrm{d}\tau, \mathbf{g}_1\right]\mathrm{d}t \\
&+ \frac{2}{T}\int_0^T \left[\int_0^t \left[\int_0^\tau \mathbf{f}\,\mathrm{d}\tau_1, \mathbf{g}_1\right]\mathrm{d}\tau, \mathbf{f}\right]\mathrm{d}t \\
&+ \frac{1}{T}\int_0^T \left[\int_0^t \left[\int_0^\tau \mathbf{f}\,\mathrm{d}\tau_1, \mathbf{g}_1\right]\mathrm{d}\tau, \mathbf{g}_1\right]\mathrm{d}t \\
&- \frac{2}{T}\int_0^T \left[\int_0^t \mathbf{g}_2\,\mathrm{d}\tau, \mathbf{f}\right]\mathrm{d}t \\
&- \frac{1}{T}\int_0^T \left[\int_0^t \mathbf{g}_2\,\mathrm{d}\tau, \mathbf{g}_1\right]\mathrm{d}t \\
&- \frac{1}{T}\int_0^T \left[\int_0^t \mathbf{f}\,\mathrm{d}\tau, \left[\int_0^t \mathbf{f}\,\mathrm{d}\tau, \mathbf{g}_1\right]\right]\mathrm{d}t \\
&+ \frac{1}{T}\int_0^T \left[\int_0^t \mathbf{g}_1\mathrm{d}\tau, \left[\int_0^t \mathbf{f}\,\mathrm{d}\tau, \mathbf{g}_1\right]\right]\mathrm{d}t \\
&+ \frac{2}{T}\int_0^T \left[\int_0^t \mathbf{f}\,\mathrm{d}\tau, \mathbf{g}_2\right]\mathrm{d}t \\
&- \frac{2}{T}\int_0^T \left[\int_0^t \mathbf{g}_1\mathrm{d}\tau, \mathbf{g}_2\right]\mathrm{d}t,
\end{aligned} \tag{47}$$

where $\mathbf{g}_2$ is given by Eq. (45). Using the Lie brackets properties and some other algebraic manipulations,[3] Eq. (47) can be further simplified to

$$\begin{aligned}
\mathbf{g}_3 =& \frac{T}{2}\left[\mathbf{g}_1, \mathbf{g}_2\right] \\
&- \frac{2}{T}\int_0^T \left[\mathbf{f}, \int_0^t \left[\int_0^\tau \mathbf{f}\,\mathrm{d}\tau_1, \mathbf{f}\right]\mathrm{d}\tau\right]\mathrm{d}t.
\end{aligned} \tag{48}$$

Using the fact that $\mathbf{g}_1 = \boldsymbol{\Lambda}_1$ and $\mathbf{g}_2 = \boldsymbol{\Lambda}_2$, we rewrite Eq. (39) as,

$$\boldsymbol{\Lambda}_3 = -\frac{3T}{2}\left[\mathbf{g}_1,\ \mathbf{g}_2\right]$$

$$+\frac{2}{T}\int_0^T \left[\int_0^t \mathbf{f}\,\mathrm{d}\tau\ ,\ \left[\int_0^t \mathbf{f}\,\mathrm{d}\tau,\ \mathbf{f}\right]\right]\mathrm{d}t.$$

In order to prove that $\mathbf{g}_3 = \boldsymbol{\Lambda}_3$, we now need to show that

$$\begin{aligned}
&\frac{T}{2}\left[\mathbf{g}_1, \mathbf{g}_2\right] - \frac{2}{T}\int_0^T \left[\mathbf{f}, \int_0^t \left[\int_0^\tau \mathbf{f}\,\mathrm{d}\tau_1, \mathbf{f}\right]\mathrm{d}\tau\right]\mathrm{d}t = \\
&\quad - \frac{3T}{2}\left[\mathbf{g}_1, \mathbf{g}_2\right] + \frac{2}{T}\int_0^T \left[\int_0^t \mathbf{f}\,\mathrm{d}\tau, \left[\int_0^t \mathbf{f}\,\mathrm{d}\tau, \mathbf{f}\right]\right]\mathrm{d}t,
\end{aligned}$$

or equivalently, that

$$\begin{aligned}
2T\left[\mathbf{g}_1, \mathbf{g}_2\right] =& \frac{2}{T}\int_0^T \left[\mathbf{f}, \int_0^t \left[\int_0^\tau \mathbf{f}\,\mathrm{d}\tau_1, \mathbf{f}\right]\mathrm{d}\tau\right]\mathrm{d}t \\
&+ \frac{2}{T}\int_0^T \left[\int_0^t \mathbf{f}\,\mathrm{d}\tau, \left[\int_0^t \mathbf{f}\,\mathrm{d}\tau, \mathbf{f}\right]\right]\mathrm{d}t.
\end{aligned} \tag{49}$$

Let us consider the right-hand side of Eq. (49)

$$\begin{aligned}
&\frac{2}{T}\int_0^T \left[\mathbf{f}, \int_0^t \left[\int_0^\tau \mathbf{f}\,\mathrm{d}\tau_1, \mathbf{f}\right]\mathrm{d}\tau\right]\mathrm{d}t \\
&\quad + \frac{2}{T}\int_0^T \left[\int_0^t \mathbf{f}\,\mathrm{d}\tau, \left[\int_0^t \mathbf{f}\,\mathrm{d}\tau, \mathbf{f}\right]\right]\mathrm{d}t.
\end{aligned}$$

Using the linearity of integration and the Lie bracket property (41), we obtain

$$\frac{2}{T}\int_0^T \frac{\partial}{\partial t}\left[\int_0^t \mathbf{f}\,\mathrm{d}\tau, \int_0^t \left[\int_0^\tau \mathbf{f}\,\mathrm{d}\tau_1, \mathbf{f}\right]\mathrm{d}\tau\right]\mathrm{d}t,$$

which is further simplified, by applying the fundamental theorem of calculus, to

$$\frac{2}{T}\left[\int_0^T \mathbf{f}\,\mathrm{d}\tau, \int_0^T \left[\int_0^t \mathbf{f}\,\mathrm{d}\tau, \mathbf{f}\right]\mathrm{d}t\right].$$

Finally, recalling the expressions for $\mathbf{g}_1$ and $\mathbf{g}_2$ in Eqs. (24) and (45), respectively, we obtain

$$2T\left[\mathbf{g}_1, \mathbf{g}_2\right].$$

Thus, $\mathbf{g}_3 = \boldsymbol{\Lambda}_3$ and the equivalence of the third-order terms of the two higher-order averaging techniques is proved.

3.4 Fourth order

According to Assumption 1, $\mathbf{f} = \mathbf{f}_1$ and $\mathbf{f}_2 = \mathbf{f}_3 = \mathbf{f}_4 = \mathbf{0}$. Assumption 2 leads to $\mathbf{C}_1 = \mathbf{C}_2 = \mathbf{C}_3 = \mathbf{0}$. Substituting the expressions for $\mathbf{w}_1$, $\mathbf{w}_2$ and $\mathbf{w}_3$ in Eq. (28) into Eq. (27) and after some algebraic manipulations, the expression for $\mathbf{g}_4$ in Eq. (27) reduces to

$$\begin{aligned}
\mathbf{g}_4 =& \frac{6}{T}\int_0^T \left[\int_0^t \left[\int_0^\tau \left[\int_0^{\tau_2} \mathbf{f}\,\mathrm{d}\tau_2\ ,\ \mathbf{f}\right]\mathrm{d}\tau_1, \mathbf{f}\right]\mathrm{d}\tau, \mathbf{f}\right]\mathrm{d}t \\
&- \frac{T^2}{2}\left[\mathbf{g}_1, \left[\mathbf{g}_1, \mathbf{g}_2\right]\right] + T\left[\mathbf{g}_1, \mathbf{g}_3\right],
\end{aligned} \tag{50}$$

[3] The process of reducing Eqs. (47)–(48) is rather lengthy and would take several pages. Thus, the authors decided not to include this part in this paper and refer the reader to our *Mathematica*® file [42] for the complete proof.

where $\mathbf{g}_1$, $\mathbf{g}_2$ and $\mathbf{g}_3$ are given by Eqs. (24), (45) and (48), respectively. □

It can be shown that $\mathbf{g}_4 = \boldsymbol{\Lambda}_4$, or equivalently,

$$\begin{aligned}
&\frac{6}{T}\int_0^T \left[\int_0^t \left[\int_0^\tau \mathbf{h}_{\tau_1}\, \mathrm{d}\tau_1, \mathbf{f}\right] \mathrm{d}\tau, \mathbf{f}\right] \mathrm{d}t \\
&\quad - \frac{T^2}{2}\left[\mathbf{g}_1, \left[\mathbf{g}_1, \mathbf{g}_2\right]\right] + T\left[\mathbf{g}_1, \mathbf{g}_3\right] \\
&\quad = \frac{2}{T}\int_0^T \left(\int_0^t \left[\int_0^\tau \mathbf{h}_{\tau_1}\, \mathrm{d}\tau_1, [\mathbf{f}_\tau, \mathbf{f}]\right] \mathrm{d}\tau \right. \\
&\quad + \left[\int_0^t \left[\int_0^\tau \mathbf{h}_{\tau_1}\, \mathrm{d}\tau_1, \mathbf{f}_\tau\right] \mathrm{d}\tau, \mathbf{f}\right] \\
&\quad + \left. \int_0^t \left[\int_0^\tau \mathbf{f}_{\tau_1} \mathrm{d}\tau_1, [\mathbf{h}_\tau, \mathbf{f}]\right] \mathrm{d}\tau \right) \mathrm{d}t,
\end{aligned} \tag{51}$$

where $\mathbf{h}_{\tau_1} = \left[\int_0^{\tau_1} \mathbf{f}_{\tau_2} \mathrm{d}\tau_1, \mathbf{f}_{\tau_1}\right]$, and $\mathbf{h}_\tau = \left[\int_0^\tau \mathbf{f}_{\tau_1} \mathrm{d}\tau, \mathbf{f}_\tau\right]$. For the details of the proof of the equivalence in Eq. (51) we refer once again to our *Mathematica*® file [42].

With this reconciliation, each approach can benefit from the tools and results developed in the other approach. For example, there are no recursive formulas for higher-order averaging using chronological calculus: Neither Vela [27] nor Sarychev [34] provided such recursive formulas. For a specific desired rth-order averaging, one has to derive it following the approaches presented by Sarychev [34], Vela [27], and the one discussed above. Instead, Algorithm 1 provides a straightforward recursive formula for arbitrary higher-order averaging terms using Lie transform. As such, with the proved equivalence, one can use Algorithm 1, setting $\mathbf{f}_1 = \mathbf{f}$, $\mathbf{f}_m = \mathbf{0}$ for $m > 1$ and $\mathbf{C}_m = \mathbf{0}$ for all m, to determine a recursive formula for the higher-order averaging terms defined through chronological calculus.

On the other hand, Sarychev [34] and Vela [27] showed that system (35) represents *complete* averaging of the NLTP system (29). In other words, the full series in Eq. (35) represents the logarithm of the Monodromy map. Therefore, if this series converges to $\boldsymbol{\Lambda}$, the stability characteristics of this logarithm $\boldsymbol{\Lambda}$ are exactly related to the stability characteristics of the original NLTP system (29). Through the proved equivalence, one can apply Theorem 6 to the full series (21) of the averaged dynamics using Lie transform. This result is important because it has not been proved within the classical averaging formulation.

4 Applications

In this section, we present two applications of higher-order averaging to mechanical systems. First, we apply higher-order averaging to a classical problem, namely the Kapitza inverted pendulum problem (see, for example, [43]). Then we consider a modern application of flapping flight of biological flyers and their man-made counterparts: flapping-wing micro-air-vehicles (see [22,44]).

4.1 Kapitza's inverted pendulum

The Kapitza inverted pendulum problem has been studied for decades because it gives interesting insights into the concept of vibrational stabilization. In fact, it has been shown that high-frequency, high-amplitude, periodic forcing, applied in the vertical direction to the pendulum pivot, stabilizes the pendulum about its inverted position [10,20]. This problem was first analyzed by Stephenson [45] in 1908 using a periodic series solution to the linearized system (Hill's equation) and then by Kapitza [43,46] considering the full nonlinear dynamics using an averaged potential approach. The reader may like to watch the interesting video in Ref. [47] that clearly shows the phenomenon and its associated stabilizing stiffness mechanism.

The pendulum dynamics, linearized about its unstable vertical equilibrium, can be written as

$$\ddot{\theta} + (-\Omega + B_1 \cos(\omega t))\theta = 0, \tag{52}$$

where θ and $\dot{\theta}$ represent the angular position and velocity, respectively, ω is the forcing frequency, B_1 is the forcing amplitude, and $\Omega = g/l$ (g is the gravitational acceleration and l the pendulum length). Figure 2 shows a schematic of the inverted pendulum oscillating vertically about its pivot. Setting $x_1 = \theta$ and $x_2 = \dot{\theta}$, Eq. (52) can be written in state space representation as

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \end{pmatrix} = \begin{pmatrix} x_2 \\ (\Omega - B_1 \cos(\omega t))x_1 \end{pmatrix}.$$

The problem of interest (vibrational control problem) is for high amplitude and high frequency; thus, we introduce the scaling $B_1 = B/\epsilon$ and $\omega = 1/\epsilon$, where ϵ is a small parameter. This yields

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \end{pmatrix} = \begin{pmatrix} x_2 \\ (\Omega - \frac{B}{\epsilon} \cos(\frac{t}{\epsilon}))x_1 \end{pmatrix}. \tag{53}$$

System (53) is not amenable to the averaging theorem: $\mathbf{f}$ is not analytic in epsilon because of the high-amplitude

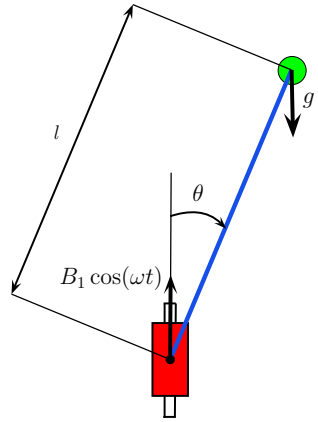

Fig. 2 Kapitza's inverted pendulum

forcing term proportional to $1/\epsilon$. This scenario is typical with vibrational control systems (high-amplitude periodically forced systems). The standard remedy for this issue is performed by applying the variation of constants (VOC) formula [7].

4.1.1 The nonlinear variation of constants formula

The VOC formula is used to decouple the flow along two vector fields. In particular, when applied to the multi-scale system (53), it decouples it into two systems, each of which is amenable to the averaging theorem. The VOC is quite general; it is applicable to systems in the form

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, t) + \mathbf{g}(\mathbf{x}, t)\,. \tag{54}$$

The VOC formula decouples system (54) into the two systems

$$\begin{aligned} \dot{\mathbf{z}} &= \mathbf{F}(\mathbf{z}), \qquad & \mathbf{z}(0) &= \mathbf{x}_0, \\ \dot{\mathbf{y}} &= \mathbf{g}(\mathbf{y}, t), \qquad & \mathbf{y}(0) &= \mathbf{z}(t), \end{aligned} \tag{55}$$

where $\mathbf{F}$ is the *pullback* of $\mathbf{f}$ along the flow of $\mathbf{g}$, defined as

$$\mathbf{F}(\mathbf{z}, t) = \left((\boldsymbol{\phi}_t^{\mathbf{g}})^* \mathbf{f}\right)(\mathbf{z}) = \left(\left[\frac{\partial \boldsymbol{\phi}_t^{\mathbf{g}}}{\partial \mathbf{z}}\right]^{-1} \circ \mathbf{f} \circ \boldsymbol{\phi}_t^{\mathbf{g}}\right)(\mathbf{z}),$$

and $\boldsymbol{\phi}_t^{\mathbf{g}}$ is the flow along the vector field $\mathbf{g}$ for time t. Figure 3 shows the idea behind the VOC formula: The flow along $\mathbf{f}+\mathbf{g}$ corresponds to the flow along the vector field $\mathbf{F}$, and then $\mathbf{g}$. The two systems in Eq. (55) should be amenable to the averaging theorem, perhaps after a simple transformation. Moreover, since $\mathbf{g}$ is a zero-mean, time-periodic vector field, its averaging vanishes

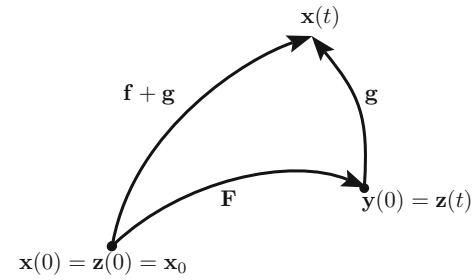

Fig. 3 The flow along $\mathbf{f}+\mathbf{g}$ is equivalent to the flow along the vector fields $\mathbf{F}$ and then $\mathbf{g}$

and the first-order averaged dynamics of the entire system would be obtained by averaging $\mathbf{F}$.

To apply the VOC formula to system (53) with $\mathbf{f}(\mathbf{x}, t) = (x_2, \Omega x_1)^T$ and $\mathbf{g}(\mathbf{x}, t) = (0, -(B/\epsilon)\cos(t/\epsilon)x_1)^T$, we first obtain the flow of $\mathbf{g}$

$$\begin{aligned} \boldsymbol{\phi}_t^{\mathbf{g}}(\mathbf{z}) &= \begin{pmatrix} z_1 \\ z_2 - B\sin\left(\frac{t}{\epsilon}\right) z_1 \end{pmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ -B\sin(\frac{t}{\epsilon}) & 1 \end{bmatrix} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \end{aligned}$$

then

$$\mathbf{f} \circ (\boldsymbol{\phi}_t^{\mathbf{g}})(\mathbf{z}) = \begin{pmatrix} z_2 - B\sin(\frac{t}{\epsilon})z_1 \\ \Omega z_1 \end{pmatrix},$$

and since

$$\left[\frac{\partial \boldsymbol{\phi}_t^{\mathbf{g}}}{\partial \mathbf{z}}\right]^{-1} = \begin{bmatrix} 1 & 0 \\ B\sin(\frac{t}{\epsilon}) & 1 \end{bmatrix},$$

we determine the pullback vector field $\mathbf{F}$ as

$$\begin{aligned} \mathbf{F}(\mathbf{z}, t) &= \begin{bmatrix} 1 & 0 \\ B\sin(\frac{t}{\epsilon}) & 1 \end{bmatrix} \begin{pmatrix} z_1 \\ z_2 - B\sin(\frac{t}{\epsilon})z_1 \end{pmatrix} \\ &= \begin{pmatrix} z_2 - B\sin(\frac{t}{\epsilon})z_1 \\ \left(\Omega - B^2\sin^2(\frac{t}{\epsilon})\right) z_1 + B\sin(\frac{t}{\epsilon})z_2 \end{pmatrix}. \end{aligned} \tag{56}$$

Finally, we perform the scaling $\tau = \omega t$ to system (56), to render it in the averaging-canonical form (weakly forced system with $\mathbf{f}$ analytic in ϵ).

$$\begin{aligned} \mathbf{z}' &= \epsilon \mathbf{F}(\mathbf{z}, \tau) \\ &= \epsilon \begin{pmatrix} z_2 - B\sin(\tau)z_1 \\ \left(\Omega - B^2\sin^2(\tau)\right) z_1 + B\sin(\tau)z_2 \end{pmatrix}, \end{aligned} \tag{57}$$

where $\mathbf{z}' = \mathrm{d}\mathbf{z}/\mathrm{d}\tau$.

We apply higher-order averaging to system (57). In order to assess the accuracy of the higher-order averaging approximations, we compare the obtained results with the Floquet stability map, obtained by numerical application of the Floquet theorem to the LTP

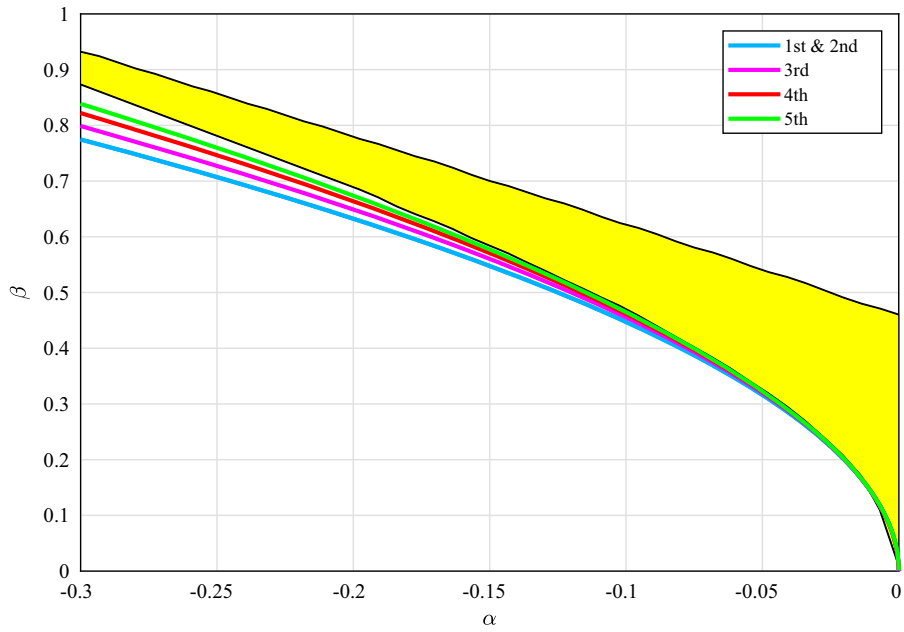

Fig. 4 Stability map for the Kapitza pendulum. The shaded region represents the Floquet stable region. The colored curves represent the higher-order stroboscopic averaging approximations (with VOC formula) of the lower stability boundary

system (53). The approximations are computed up to fifth order using the algorithm described in Sect. 2.1, and by choosing the stroboscopic averaging, namely in Algorithm 1, we set $\mathbf{C}_k = \mathbf{0}$ for $k = 1, 2, 3, 4$. Hence, the two averaging techniques will lead to identical results due to the proved equivalence in this case of $\partial \mathbf{F}/\partial \epsilon = \mathbf{0}$. Figure 4 illustrates the stability map of the Kapitza pendulum in a similar fashion to the work of Berg and Wickramasinghe [20]: The plotted quantities are defined as $\alpha = -\Omega/\omega^2$ and $\beta = B/\omega$, and the region colored in yellow represents the Floquet stable region. In addition, the figure shows the approximations of the lower stability boundary obtained via higher-order averaging. Although only the lower stability boundary is captured, it is evident that the accuracy of the approximations increases with the averaging order. That is, the higher-order averaging stability boundary converges asymptotically to the Floquet boundary as the averaging order increases.

Figure 5 shows the percent error between the Floquet lower stability boundary and the higher-order averaging approximations as the forcing frequency ω is varied. Two important conclusions can be drawn from this figure. First, for a fixed ω, the error decreases as the averaging order is increased, which matches the intuition behind the asymptotically convergent series as more terms are taken [15]. Second, the error associated with all the averaging order approximations decreases as the frequency is increased, which matches the intuition behind averaging. These two conclusions indicate that, for a required level of accuracy, the threshold frequency decreases with the averaging order. That is, higher-order averaging allows stabilization at frequencies lower than that demanded by direct averaging, hence making vibrational control more feasible.

Figure 6 shows the angular displacement and velocity $(\theta, \dot{\theta})$ of the Kapitza pendulum with initial conditions $\theta_0 = 30°$ and $\dot{\theta}_0 = 0°/\text{s}$ and for $\Omega = 2$, $B = 3$, and $\omega = 20$ rad/s. The three curves represent the integration of the complete LTP dynamics (57), the first-order and third-order averaged dynamics. The solutions of the averaged dynamics have been transformed according to Eq. (23). While both the averaged solutions correctly capture the stability of the original system, the third-order approximation is practically indistinguishable from the solution of the full LTP system.

We also present in Figs. 7 and 8 the results for a non-stroboscopic averaging case. In particular, the integration constants $\mathbf{C}_k$ are computed using Eq. (20). Similar conclusions to the stroboscopic case can be drawn.

4.1.2 Higher-order averaging without VOC

As shown above, vibrational control systems are typically forced by high-amplitude periodic inputs and

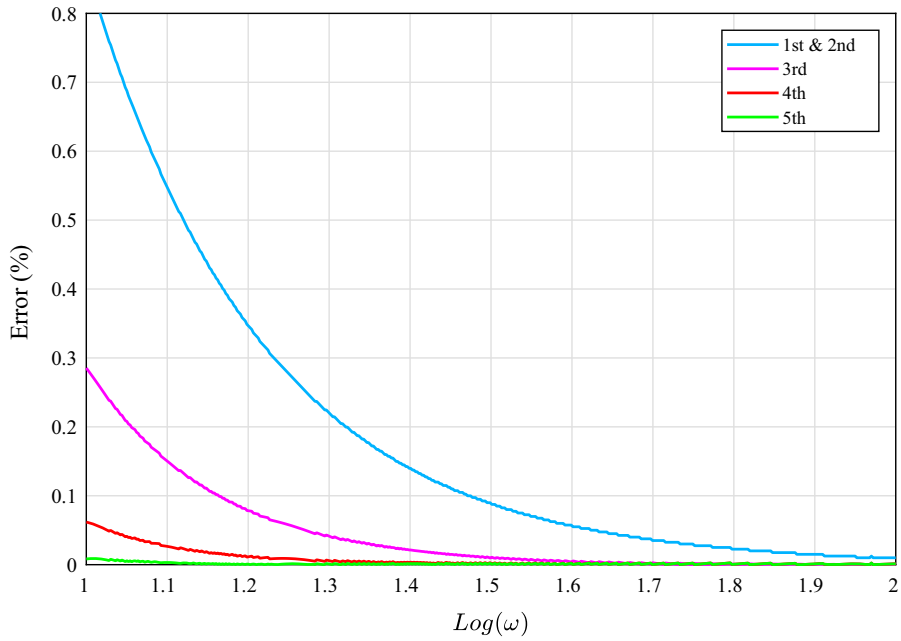

Fig. 5 Percent error in the lower stability boundary for $\Omega = 2$ for the Kapitza pendulum using higher-order stroboscopic averaging with the VOC formula

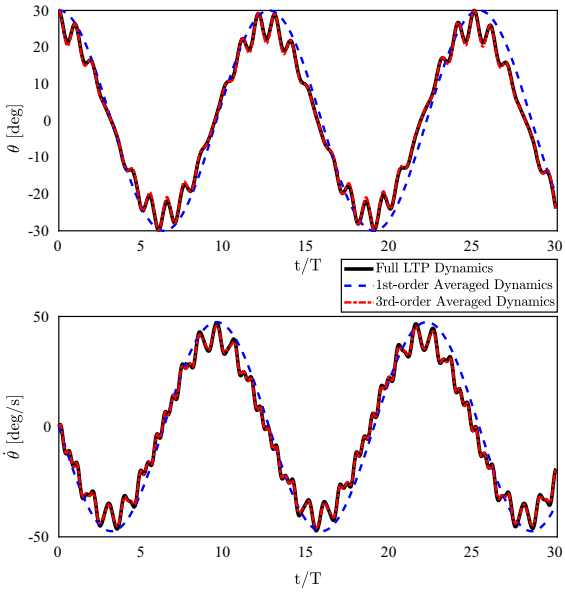

Fig. 6 Integration of the full LTP dynamics, first- and third-order averaged dynamics of the Kapitza pendulum, after applying the VOC formula. The simulations are obtained for $\Omega = 2$, $B = 3$, $\omega = 20$ rad/s, and initial conditions $\theta_0 = 30°$ and $\dot{\theta}_0 = 0°/\text{s}$

therefore are not in the averaging-canonical form; the VOC formula is necessary to ensure a rigorous application of the averaging theorem. However, as shown above, the VOC requires computing the flow map along the time-periodic terms, which may not be analytically tractable for many systems. Here we show that higher-order averaging alone (without the VOC) may capture some of the dynamical features that are lost by disregarding the VOC. In order to demonstrate this point, consider again the system in Eq. (53) after the transformation $\tau = \omega t = t/\epsilon$

$$\begin{pmatrix} x_1' \\ x_2' \end{pmatrix} = \epsilon \begin{pmatrix} x_2 \\ (\Omega - \omega B \cos(\tau)) x_1 \end{pmatrix}, \tag{58}$$

Figure 9 shows the percent error in the stability boundaries obtained by applying higher-order averaging directly (without VOC) to system (58). Similar conclusions to those of Fig. 5 (with the VOC formula) can be drawn: The accuracy increases as ω increases (i.e., ϵ decreases) and/or the averaging order is increased. However, in comparison with the higher-order averaging with VOC, the accuracy of the approximations in Fig. 9 is noticeably lower; a significantly higher-order averaging may be required to attain an acceptable level of accuracy. Finally, Fig. 10 shows the angular displacement and velocity of the Kapitza pendulum with initial conditions $\theta_0 = 30°$ and $\dot{\theta}_0 = 0°/\text{s}$ and for $\Omega = 2$, $B = 3$, and $\omega = 20$ rad/s. The four curves represent the integration of the complete LTP dynamics (58), the first-order, third-order and fifth-order averaged dynamics. The solutions of the averaged dynamics have been adjusted according to Eq. (23). In this case, the solution of the first-order averaged dynamics provides wrong

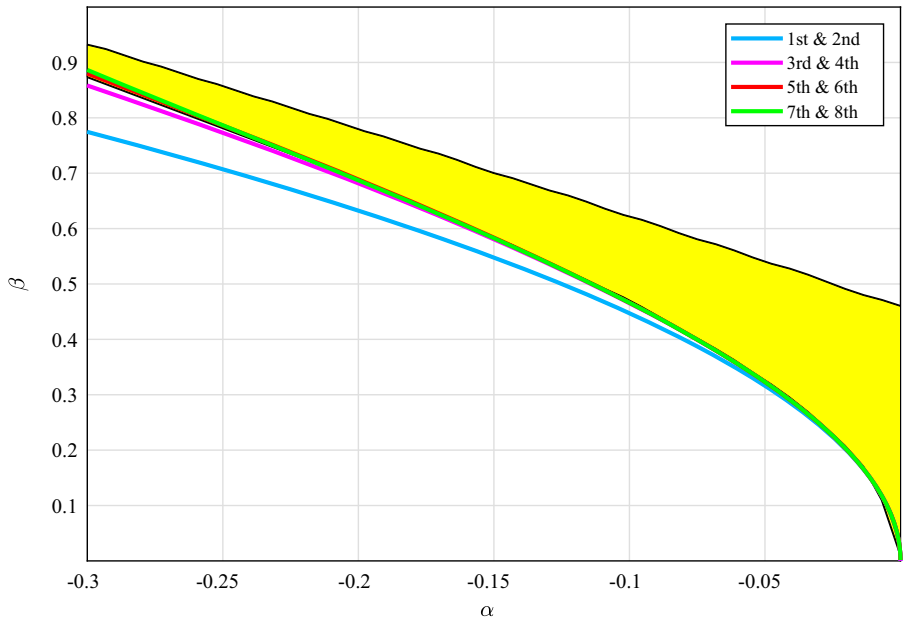

Fig. 7 Stability map for the Kapitza pendulum. The shaded region represents the Floquet stable region. The colored curves represent the higher-order non-stroboscopic averaging approximations (with VOC formula) of the lower stability boundary

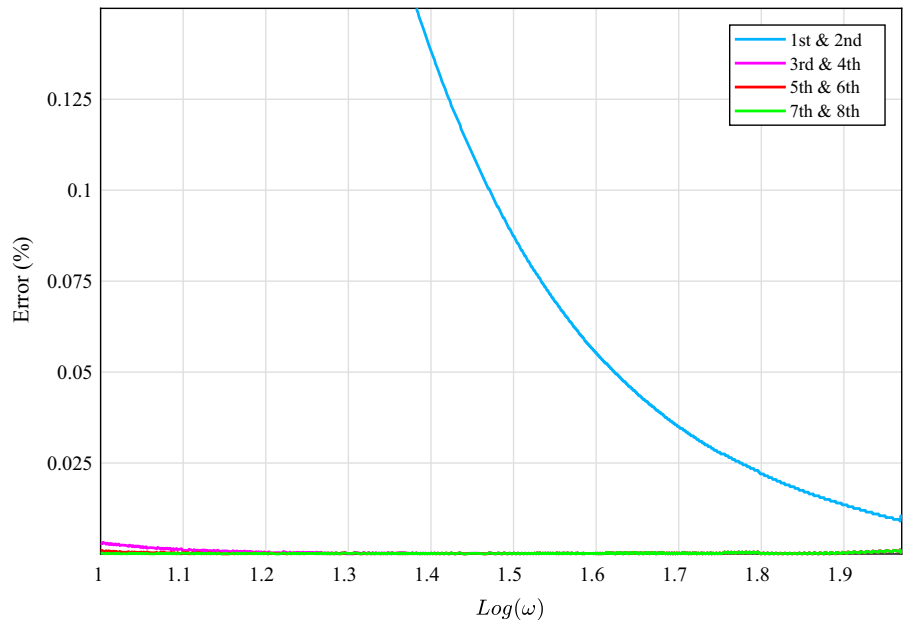

Fig. 8 Percent error in the lower stability boundary for $\Omega = 2$ for the Kapitza pendulum using higher-order non-stroboscopic averaging with the VOC formula

information about the stability of the original system, while the stability characteristics are correctly captured by the higher-order averaged dynamics. As expected, when the VOC formula is not applied, a higher averaging order (fifth order versus third order for the VOC case) may be required to obtain a close approximation of the solution of the original LTP system.

In conclusion, higher-order averaging might be beneficial to the stability analysis of complex systems even when the application of the VOC formula is not possible, such as in the modern example shown below.

4.2 Hovering flight of insects and flapping-wing micro-air-vehicles (FWMAVs)

Flapping flight dynamics is a very rich dynamical system because it is typically represented by a multi-body, nonlinear, time-varying model. To simplify the analy-

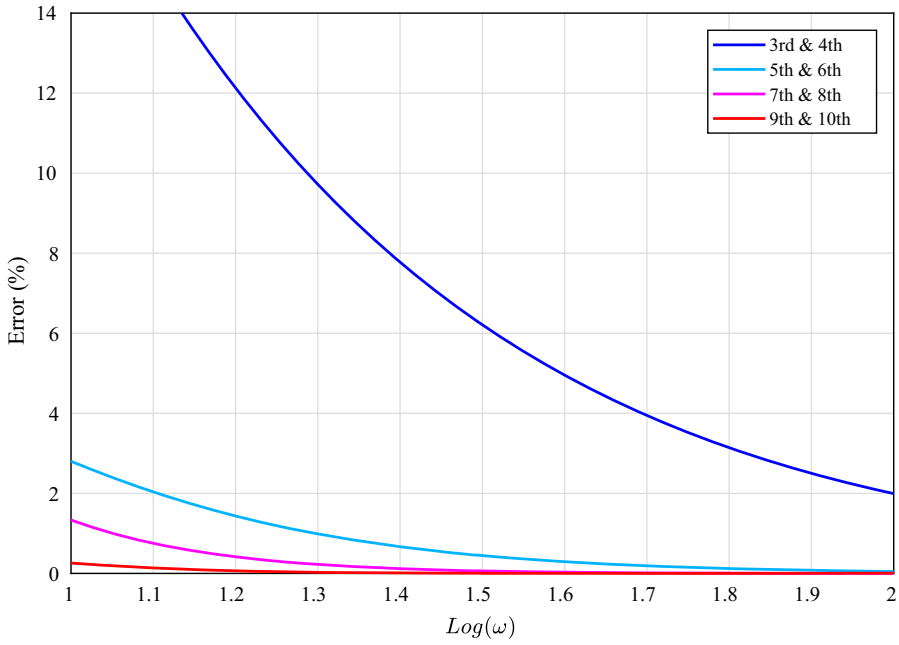

Fig. 9 Percent error in the lower stability boundary for $\Omega = 2$ for the Kapitza pendulum using higher-order averaging without the VOC formula

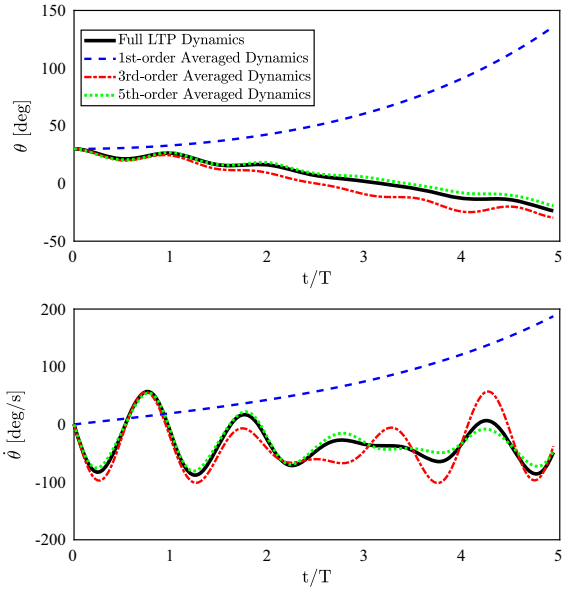

Fig. 10 Integration of the full LTP dynamics, first-, third- and fifth-order averaged dynamics of the Kapitza pendulum, without applying the VOC formula. The simulations are obtained for $\Omega = 2$, $B = 3$, $\omega = 20$ rad/s, and initial conditions $\theta_0 = 30°$ and $\dot{\theta}_0 = 0°/s$

sis of such a complicated system, it may be reasonable to neglect the wing flexibility and inertial effects, even though it is a controversial assumption among the flapping flight dynamics community [6,48]. Doing so, the flapping flight dynamics in the longitudinal plane ($x-z$ plane in Fig. 11) will be governed by the exact same set of equations governing conventional airplanes [49]

$$\begin{pmatrix} \dot{u} \\ \dot{w} \\ \dot{q} \\ \dot{\theta} \end{pmatrix} = \begin{pmatrix} -qw - g\sin\theta \\ qu + g\cos\theta \\ 0 \\ q \end{pmatrix} + \begin{pmatrix} \frac{1}{m}X(\mathbf{x},t) \\ \frac{1}{m}Z(\mathbf{x},t) \\ \frac{1}{I_y}M(\mathbf{x},t) \\ 0 \end{pmatrix}, \quad (59)$$

where I_y and m are the pitch inertia and the body mass, respectively, and g is the gravitational acceleration. Figure 11 illustrates a flapping-wing micro-air-vehicle (FWMAV) in the longitudinal plane ($x-z$ plane) where the state vector $\mathbf{x}$ includes the forward and normal velocity components u, w, the pitching angular velocity q, and the pitching angle θ. In Eq. (59), X, Z are the forward and normal aerodynamic forces, respectively, and M is the aerodynamic pitching moment.

Details of the aerodynamic model (dependence of the aerodynamic forces X, Z and M on the state variables u, w and q) are taken from Ref. [50] and given in "Appendix A."

The main distinction between conventional airplanes and FWMAVs or insect flight is that the aerodynamic forces (X, Z and M) are essentially time-varying in the latter case, which renders system (59) time-varying nature. That is, system (59) can be written as

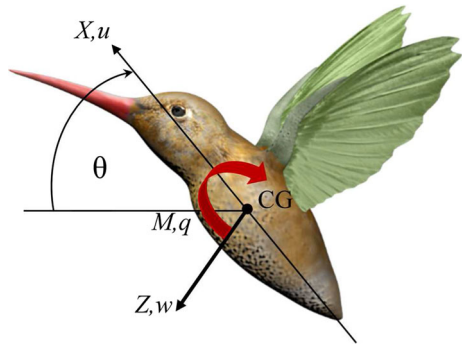

Fig. 11 Longitudinal view of a hovering FWMAV

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \mathbf{g}_a(\mathbf{x}, t), \tag{60}$$

where the autonomous vector field $\mathbf{f}$ represents the inertial and gravitational loads and the non-autonomous vector field $\mathbf{g}_a$ represents the aerodynamic loads. In conclusion, unlike conventional airplanes, flapping flight dynamics (i.e., (59) and (60)) is represented by a nonlinear, time-periodic (NLTP) system whose stability can be studied using one of the two approaches shown in Fig. 1.

Hovering is mostly associated with relatively higher flapping frequencies in comparison with forward flight. The flapping frequencies of hovering insects are usually in the 20–1000 Hz range [51]. Therefore, the dynamics of hovering insects exhibit two timescales. The fast timescale is associated with the flapping motion and the aerodynamic loads, while the slow one is associated with the aggregate motion of the body. The contrast between the fast and slow timescales can be clearly seen if one tries to observe a flying insect: A bare human eye can track the trajectory of an insect body in the space, but can hardly track the motion of its wings, which move much faster. Therefore, we definitely have two timescales. A natural question is to ask about the ratio between these two timescales. One of the slowest flapping insects, the hawkmoth, has its flapping frequency (i.e., forcing frequency) about 30 times its natural frequency of flight dynamics (specifically the short-period flight dynamics) [50,51]. This large separation between the two timescales invokes averaging; the body of an insect hovering over a flower seems to oscillate in all directions. Yet, on average, the insect is hovering over the flower. Therefore, averaging seems a very intuitive tool to analyze flapping flight dynamics.

Over the last two decades, direct (crude) averaging has been the standard analysis technique of the flapping flight dynamics of insects and micro-air-vehicles [52–61]. Using the crude averaging approach, it was shown that the pitch dynamics of hovering insects and FWMAVs are unstable due to lack of pitching stiffness [54–59]. However, we recall the expressive quote from Sanders and Verhulst book [17] presented in Introduction section stressing that averaging has to be rigorously justified. In fact, the flapping flight dynamical system as a high-amplitude periodically forced system is not directly amenable to the averaging theorem. Moreover, its dynamics are too complicated to allow analytical application of the VOC. Therefore, there is a need for higher-order averaging [44].

Prof. Nayfeh had a long career studying energy transfer between high-frequency modes and low-frequency modes [62–67]. It is noteworthy to mention that Prof. Nayfeh opposed direct (crude) averaging of flapping flight dynamics despite the seemingly large separation between the system's two timescales that intuitively invokes averaging. Hence, this whole line of research of higher-order averaging of insect flight dynamics was initiated. In an earlier work [68], the third author and Nayfeh utilized the multiple scales method [15,32] to analyze the hovering flight dynamics of insects and FWMAVs up to second-order accuracy; they indeed highlighted the shortcomings of crude averaging and showed that the parametric excitation due to the oscillatory aerodynamic loads may naturally (without feedback) stabilize the insect flight dynamics.

In this subsection, we present how high-order averaging captures this important dynamical feature in the hovering flight dynamics of insects and FWMAVs [44]. In this demonstration, we adopt the flight dynamic model (i.e., model for the aerodynamic vector field $\mathbf{g}_a$ in terms of the state variables) developed earlier in [50].

Strictly speaking, system (60) (equivalently (59)) is not in the averaging-canonical form because the time-periodic aerodynamic vector field $\mathbf{g}_a(\mathbf{x}, t)$ includes high-amplitude terms. Therefore, a rigorous application of the averaging theorem necessitates application of the VOC formula before averaging. However, it is not possible to analytically compute the flow along the NLTP vector field $\mathbf{g}_a$. This precludes application of direct averaging, which is not clear until a careful scaling is assigned to each term in $\mathbf{g}_a$. Yet, direct (crude) averaging has been ubiquitously applied to analyze flapping flight dynamics over the past two decades, relying on the seemingly large separation between the two timescales [52–61].

Table 2 Ratios of flapping frequency to natural frequency ω/ω_n and eigenvalues assessing stability (using first- and second-order averaging) for five insects

Insect	$\frac{\omega}{\omega_n}$	λ_{1st}	λ_{2nd}
Hawkmoth	28.78	$[-11.89,\ -3.30,\ 0.19 \pm 5.74i]$	$[-10.40,\ -3.09,\ -0.66 \pm 3.72i]$
Crane fly	50.62	$[-47.71,\ -17.31,\ -1.13 \pm 5.53i]$	$[-45.76,\ -16.26,\ -13.16,\ 7.90]$
Bumblebee	144.46	$[-11.63,\ -4.39,\ 1.58 \pm 6.55i]$	$[-11.26,\ -4.37,\ 1.38 \pm 6.17i]$
Dragonfly	145.50	$[-13.11,\ -7.03,\ 1.34 \pm 6.65i]$	$[-12.56,\ -6.98,\ 1.04 \pm 5.99i]$
Hoverfly	113.98	$[-14.01,\ -7.27,\ 2.13 \pm 8.56i]$	$[-13.37,\ -7.24,\ 1.79 \pm 7.92i]$

As explained by Khan and Agrawal [60], the time variable in Eq. (60) is scaled as $\tau = \frac{\omega}{\omega_n} t$, where ω_n is the natural frequency of the body flight dynamics (short-period mode) and ω is the flapping (forcing) frequency. When $\frac{\omega}{\omega_n}$ is large enough and assuming that all the terms in $\mathbf{f}$ and $\mathbf{g}_a$ are equally scaled (which is not true in general), the dynamics expressed in terms of τ are in the averaging-canonical form with $\epsilon = \frac{\omega_n}{\omega}$. It is important to note that, for hovering insects with relatively low flapping frequencies (e.g., for the hawkmoth, $\frac{\omega}{\omega_n} \approx 30$) the ratio $\frac{\omega}{\omega_n}$ has usually been considered high enough to justify direct averaging (see, for example, [50,52,55]).

The averaged dynamics of Eq. (60) are written (see [50]) as

$$\dot{\bar{\mathbf{x}}} = \mathbf{f}(\bar{\mathbf{x}}) + \bar{\mathbf{g}}_a(\bar{\mathbf{x}}) \tag{61}$$

where $\bar{\mathbf{x}}$ is the averaged state vector and $\bar{\mathbf{g}}_a(\bar{\mathbf{x}})$ represents the $\mathbf{g}_a(\mathbf{x}, t)$ averaged aerodynamic loads: $\bar{\mathbf{g}}_a(\mathbf{x}) = \frac{1}{T}\int_0^T \mathbf{g}_a(\mathbf{x}, \tau)\, \mathrm{d}\tau$.

After applying second-order averaging to system (60), we obtain a significant change in the stability characteristics in comparison with crude averaging. Table 2 shows the ratios $\frac{\omega}{\omega_n}$ for five different insects and the eigenvalues λ_{1st} and λ_{2nd}, computed after applying first- and second-order averaging, respectively. The eigenvalues are used to assess stability.

Some important observations can be made about the results in Table 2. First, for very-high-frequency cases, first-order (direct averaging) and second-order averaging give the same information about the averaged system and its stability such as in the cases of the bumblebee, the dragonfly and the hoverfly. In fact, for these insects the ratio $\frac{\omega}{\omega_n}$ is very large (larger than 100). Also, this result is consistent with that obtained by the method of multiple scales [68]. On the other hand, when the ratio $\frac{\omega}{\omega_n}$ is relatively lower, such as in the case of the crane fly and of the hawkmoth, the stability characteristics of the first-order and second-order averaged systems are qualitatively different. Moreover, we can even identify the vibrational stabilization mechanism comparing the matrices of the linearized direct (first-order) and second-order averaged systems for one of the insects (e.g., hawkmoth)

$$D(\epsilon \boldsymbol{\Lambda}_1)(\mathbf{0}) = \begin{bmatrix} -3.59 & 0 & 0 & -9.81 \\ 0 & -3.30 & 0 & 0 \\ 39.95 & 0 & -7.92 & \mathbf{0} \\ 0 & 0 & 1 & 0 \end{bmatrix},$$

$$D(\epsilon \boldsymbol{\Lambda}_1 + \epsilon^2 \boldsymbol{\Lambda}_2)(\mathbf{0}) = \begin{bmatrix} -3.58 & 0 & 0 & -9.81 \\ 0 & -3.09 & 0 & 0 \\ 29.98 & 0 & -8.13 & \mathbf{-28.45} \\ -2.90 & 0 & 0.96 & 0 \end{bmatrix}.$$

As can be seen, the element (3, 4), which corresponds to pitching acceleration (or pitching moment) due to a pitching angle, is changed from 0, in the crude-averaged system, to -28.45 in the second-order averaged system. That is, a significant pitch stiffness (whose lack was the main reason behind the claimed instability) is created due to high-frequency, high-amplitude, zero-mean terms in $\mathbf{g}_a$ which is a typical vibrational stabilization behavior. Figure 12 shows a simulation of the hawkmoth dynamics for an initial θ disturbance of 30°. The curves represent the simulation of the full NLTP system, the first-order averaged dynamics and the second-order averaged dynamics. The solution of the second-order averaged dynamics greatly improves the approximation obtained by applying direct averaging only.

This example shows the power of higher-order averaging in capturing hidden stabilization mechanisms in dynamically rich systems in an analytic manner that allows deep study of the underpinning physics. This scrutiny of the physics behind vibrational stabilization in flapping flight and its experimental validation using motion data from a real hawkmoth is presented in [69].

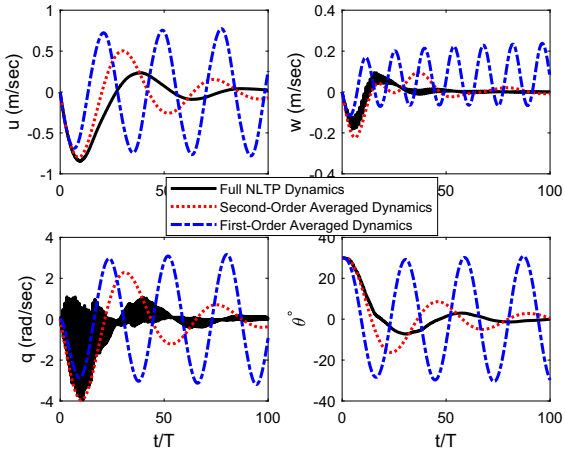

Fig. 12 Integration of the full NLTP dynamics, first- and second-order averaged dynamics for the hawkmoth. The simulations are obtained for an initial disturbance of $\theta_0 = 30^\circ$

5 Concluding remarks

In this paper, we review the two main techniques to analyze nonlinear time-periodic systems: the numerical approach using Floquet theorem and the analytical approach using the averaging theorem with particular focus on the second approach. We discuss the technical issues with the classical averaging theorem: the requirement of a weakly forced system and/or a large separation between the system's two timescales. These requirements confine the applicability of the neat averaging approach to a small class of systems that excludes vibrational control systems, which are typically high-amplitude periodically forced systems. We present higher-order averaging as a remedy for these issues. It is found that higher-order averaging captures more dynamical features than direct (first-order) averaging. In particular, we find that the error in the stability boundary decreases asymptotically as the averaging order increases. We demonstrate this behavior on the classical example of the Kapitza pendulum: inverted pendulum whose pivot is subject to vertical oscillation. We also present a modern example with richer dynamics: flapping flight dynamics of insects and micro-air-vehicles. Over the past two decades, direct averaging has led to the conclusion that insects are unstable at hover. In contrast, we show that higher-order averaging indicates vibrational stabilization in the form of a pitch stiffness. That is, the high-frequency, zero-mean forces, which are typically neglected by direct averaging, lead to a natural stabilizing action without feedback. In addition to these applications which demonstrate the power of higher-order averaging, we show the equivalence between two higher-order averaging approaches that were developed independently in the literature within two different communities and using different analysis tools: the classical averaging approach using perturbation theory and a more modern averaging approach using chronological calculus. We show how the use of chronological calculus and Lie algebraic tools leads to extension of the Floquet theorem to nonlinear systems, equivalently extending the averaging theorem to arbitrarily high orders. We prove equivalence between the two averaging approaches mathematically up to fourth order and also show that the two approaches are conceptually equivalent for all orders: Both rely on the fact that the flows of the averaged system and the time-periodic system match after a complete cycle. Based on this equivalence, we conclude that the full series solution of the classical averaging approach represents *complete averaging* in the sense that its exponential stability implies exponential stability for the original time-periodic system for all values of the "small" parameter (i.e., no large separation between the timescales is needed). Such a result was not developed in the classical averaging literature. On the other hand, the recursive formula developed in the classical averaging literature can be directly used to determine higher-order averaging terms in the chronological calculus approach, which was not presented in its respective literature. This article is considered as a comprehensive introductory reference on the analysis of time-periodic systems that reviews and unifies the classical averaging theorem, the Floquet theorem, the higher-order averaging technique using perturbation theory, and that uses chronological calculus and Lie algebraic tools. We show connections, equivalence, similarity and differences between these approaches.

Acknowledgements This work is dedicated to the memory of Prof. Ali Hassan Nayfeh whose contributions in the field of nonlinear dynamics and averaging of time-periodic systems are immense. His passion and engineering intuition were always remarkable and inspiring. To honor his memory, we dedicate this effort, which he enlightened its path before leaving. The authors are thankful to the fruitful discussions with Dr. Muhammad Hajj, Dr. Craig Woolsey, Dr. Jordan Berg, Dr. Patricio Vela, Dr. Sevak Tahmasian, Dr. Ahmed Hassan and the support of the National

Science Foundation Grant CMMI-1709746 and its continuation CMMI-1846308.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Appendix A

The higher-order dependence of the non-autonomous aerodynamic vector field $\mathbf{g}_a(\mathbf{x}, t)$ (see Eqs. (59) and (60)) on the state vector $\mathbf{x}$ can be neglected. Thus, we retain only the linear terms and obtain

$$\begin{pmatrix} \dot{u} \\ \dot{w} \\ \dot{q} \\ \dot{\theta} \end{pmatrix} = \begin{pmatrix} -qw - g\sin\theta \\ qu + g\cos\theta \\ 0 \\ q \end{pmatrix} + \begin{pmatrix} \frac{1}{m}X_0(t) \\ \frac{1}{m}Z_0(t) \\ \frac{1}{I_y}M_0(t) \\ 0 \end{pmatrix} + \begin{bmatrix} X_u(t) & X_w(t) & X_q(t) & 0 \\ Z_u(t) & Z_w(t) & Z_q(t) & 0 \\ M_u(t) & M_w(t) & M_q(t) & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{pmatrix} u \\ w \\ q \\ \theta \end{pmatrix}.$$

Assuming a horizontal stroke plane, parameterized by the "back-and-forth" flapping angle φ, and a piecewise constant variation in the wing pitch angle η, one obtains [50]: $X_0(t) = -2K_{21}\dot{\varphi}(t)|\dot{\varphi}(t)|\cos\varphi(t)\sin^2\eta$, $Z_0(t) = -K_{21}\dot{\varphi}(t)|\dot{\varphi}(t)|\sin 2\eta$, $M_0(t) = 2\dot{\varphi}(t)|\dot{\varphi}(t)|\sin\eta\,[K_{22}\Delta\hat{x}\cos\varphi(t) + K_{21}x_h\cos\eta(t) + K_{31}\sin\varphi(t)\cos\eta]$, where x_h is the distance from the vehicle center of mass to the root of the wing hinge line (i.e., the intersection of the hinge line with the x_b-axis) and $\Delta\hat{x}$ is the chord-wise distance from the center of pressure to this same hinge location, normalized by the chord length. Also, ρ is the air density, C_{L_α} is the three-dimensional lift curve slope of the wing, $c(r)$ is the spanwise chord distribution, R is the wing radius, $I_{mn} = 2\int_0^R r^m c^n(r)\,dr$, and $K_{mn} = \frac{1}{4}\rho C_{L_\alpha} I_{mn}$. The time-varying stability derivatives are written directly in terms of the system parameters as in [50]: $X_u = -4\frac{K_{11}}{m}|\dot{\varphi}|\cos^2\varphi\sin^2\eta$, $X_w = -\frac{K_{11}}{m}|\dot{\varphi}|\cos\varphi\sin 2\eta$, $X_q = \frac{K_{21}}{m}|\dot{\varphi}|\sin\varphi\cos\varphi\sin 2\eta - x_h X_w$, $Z_u = 2X_w$, $Z_w = -2\frac{K_{11}}{m}|\dot{\varphi}|\cos^2\eta$, $Z_q = 2\frac{K_{21}}{m}|\dot{\varphi}|\sin\varphi\cos^2\eta - \frac{K_{rot_{12}}}{m}\dot{\varphi}\cos\varphi - x_h Z_w$, $M_u = 4\frac{K_{12}\Delta x}{I_y}|\dot{\varphi}|\cos^2\varphi\sin\eta + \frac{m}{I_y}\left(2X_q - x_h Z_u\right)$, $M_w = 2\frac{K_{12}\Delta x}{I_y}|\dot{\varphi}|\cos\varphi\cos\eta + 2\frac{K_{21}}{I_y}|\dot{\varphi}|\sin\varphi\cos^2\eta - \frac{mx_h}{I_y}Zw$, and

$$\begin{aligned} M_q = &-\frac{2\Delta x}{I_y}|\dot{\varphi}|\cos\varphi\cos\eta\,(K_{12}x_h + K_{22}\sin\varphi) \\ &+\frac{1}{I_y}\dot{\varphi}\cos\varphi\left(K_{rot_{13}}\Delta x\cos\varphi\cos\eta + K_{rot_{22}}\sin\varphi\right) \\ &+-\frac{2}{I_y}|\dot{\varphi}|\cos^2\eta\sin\varphi\,(K_{21}x_h + K_{31}\sin\varphi) \\ &-\frac{K_v\mu_1 f}{I_y}\cos^2\varphi - \frac{mx_h}{I_y}Z_q, \end{aligned}$$

where $K_{rot_{mn}} = \pi\rho(\frac{1}{2} - \Delta\hat{x})I_{mn}$ and $K_v = \frac{\pi}{16}\rho I_{04}$. The hinge line is set at 30% c ($\Delta\hat{x} = 0.05$), and the value of C_{L_α} is computed based on the aspect ratio of the wing, utilizing the extended lifting theory [70]. This flight dynamic model has been developed in Ref. [50], and the resulting eigenvalues of the averaged, linearized dynamics have been validated against numerical simulations of Navier–Stokes equations by Sun et al. [55] and the experimental data of Cheng and Deng [59].

References

1. Vela, P.A., Burdick, J.W.: Control of biomimetic locomotion via averaging theory. In: IEEE International Conference on Robotics and Automation, vol. 1, pp. 1482–1489. IEEE (2003)
2. Vela, P.A., Burdick, J.W.: Control of underactuated mechanical systems with drift using higher-order averaging theory. In: Proceedings of the 42nd IEEE Conference on Decision and Control, vol. 3, pp. 3111–3117. IEEE (2003)
3. Morgansen, K.A., Duidam, V., Mason, R.J., Burdick, J.W., Murray, R.M.: Nonlinear control methods for planar carangiform robot fish locomotion. In: Proceedings 2001 ICRA. IEEE International Conference on Robotics and Automation. vol. 1, pp. 427–434. IEEE (2001)
4. Morgansen, K.A., Vela, P.A., Burdick, J.W.: Trajectory stabilization for a planar carangiform robot fish. In: Proceedings. ICRA'02. IEEE International Conference on Robotics and Automation, vol. 1, pp. 756–762. IEEE (2002)
5. Wood, R.J.: The first takeoff of a biologically inspired at-scale robotic insect. IEEE Trans. Robot. Autom. **24**(2), 341–347 (2008)
6. Taha, H., Hajj, M.R., Nayfeh, A.H.: Flight dynamics and control of flapping-wing mavs: a review. Nonlinear Dyn. **70**(2), 907–939 (2012)
7. Bullo, F.: Averaging and vibrational control of mechanical systems. SIAM J. Control Optim. **41**(2), 542–562 (2002)
8. Taha, H.E., Woolsey, C.A., Hajj, M.R.: Geometric control approach to longitudinal stability of flapping flight. J. Guid. Control Dyn. **39**(2), 214–226 (2015)
9. Tahmasian, S., Woolsey, C.A.: A control design method for underactuated mechanical systems using high-frequency inputs. J. Dyn. Syst. Meas. Control **137**(7), 071,004 (2015)
10. Meerkov, S.: Principle of vibrational control: theory and applications. IEEE Trans. Autom. Control **25**(4), 755–762 (1980)

11. Baillieul, J., Lehman, B.: Open-loop control using oscillatory inputs. In: CRC Control Handbook, pp. 967–980 (1996)
12. Markus, L., Yamabe, H.: Global stability criteria for differential systems. Osaka Math. J. **12**(2), 305–317 (1960)
13. Nayfeh, A.H., Mook, D.T.: Nonlinear Oscillations. Wiley, Hoboken (2008)
14. Nayfeh, A.H., Balachandran, B.: Applied Nonlinear Dynamics: Analytical, Computational, and Experimental Methods. Wiley, Hoboken (2008)
15. Nayfeh, A.H.: Perturbation Methods. Wiley, Hoboken (2008)
16. Khalil, H.K.: Nonlinear Systems. Prentice-Hall, New Jersey **2**(5), 1–5 (1996)
17. Sanders, J.A., Verhulst, F., Murdock, J.: Averaging: the periodic case. In: Averaging Methods in Nonlinear Dynamical Systems, pp. 21–44. Springer (2007)
18. Murdock, J.A.: Perturbations: Theory and Methods, vol. 27. SIAM, Philadelphia (1999)
19. Guckenheimer, J., Holmes, P.: Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields, vol. 42. Springer, Berlin (2013)
20. Berg, J.M., Wickramasinghe, I.M.: Vibrational control without averaging. Automatica **58**, 72–81 (2015)
21. Bullo, F., Lewis, A.D.: Geometric control of mechanical systems. In: Applied Mathematics. Springer, Berlin (2004)
22. Hassan, A.M., Taha, H.E.: Combined averaging-shooting approach for the analysis of flapping flight dynamics. J. Guid. Control Dyn. **41**(2), 542–549 (2017)
23. Hassan, A.M., Taha, H.E.: Differential-geometric-control formulation for flapping flight multi-body dynamics. J. Nonlinear Sci. 1–39 (2018)
24. Sanders, J.A., Verhulst, F.: Averaging Methods in Nonlinear Dynamical Systems. Springer, Berlin (1985)
25. Bogoliubov, N.N., Mitropolsky, Y.A., Gillis, J.: Asymptotic methods in the theory of non-linear oscillations. Phys. Today **16**, 61 (1963)
26. Mitropolsky, I.A.: Averaging method in non-linear mechanics. Int. J. Non-linear Mech. **2**(1), 69–96 (1967)
27. Vela, P.A.: Averaging and Control of Nonlinear Systems. Ph.D. Thesis, California Institute of Technology (2003)
28. Agrachev, A., Gamkrelidze, R.: Chronological algebras and nonstationary vector fields. J. Math. Sci. **17**(1), 1650–1675 (1981)
29. Agračev, A., Gamkrelidze, R.V.: The exponential representation of flows and the chronological calculus. Math. USSR-Sbornik **35**(6), 727 (1979)
30. Fliess, M., Lamnabhi, M., Lamnabhi-Lagarrigue, F.: An algebraic approach to nonlinear functional expansions. IEEE Trans. Circuits Syst. **30**(8), 554–570 (1983)
31. Sarychev, A.: Stability criteria for time-periodic systems via high-order averaging techniques. In: Nonlinear Control in the Year 2000, vol. 2, pp. 365–377. Springer (2001)
32. Nayfeh, A.H.: Introduction to Perturbation Techniques. Wiley, Hoboken (1981)
33. Yagasaki, K., Ichikawa, T.: Higher-order averaging for periodically forced weakly nonlinear systems. Int. J. Bifurc. Chaos **9**(03), 519–531 (1999)
34. Sarychev, A.V.: Lie-and chronologico-algebraic tools for studying stability of time-varying systems. Syst. Control Lett. **43**(1), 59–76 (2001)
35. Nijmeijer, H., Van der Schaft, A.: Nonlinear Dynamical Control Systems, vol. 175. Springer, Berlin (1990)
36. Isidori, A.: Nonlinear Control Systems. Springer, Berlin (2013)
37. Sastry, S.: Nonlinear Systems: Analysis, Stability, and Control, vol. 10. Springer, Berlin (2013)
38. Sarychev, A.: Stability criteria for time-periodic systems via high-order averaging techniques. In: Nonlinear Control in the Year 2000, *Lecture Notes in Control and Information Sciences*, vol. 2, pp. 365–377. Springer (2001)
39. Agrachev, A.A., Gamkrelidze, R., Sarychev, A.: Local invariants of smooth control systems. Acta Appl. Math. **14**(3), 191–237 (1989)
40. Grobner, W., Knapp, H.: Contributions to the method of Lie series (1967)
41. Hale, J.: Ordinary Differential Equations. John Wiley, New York (1969)
42. Mathematica file associated with the third and fourth order proof. http://taha.eng.uci.edu/biblio-html/index.html
43. Kapitza, P.: Dynamic stability of a pendulum with an oscillating point of suspension. J. Exp. Theor. Phys. **21**(5), 588–597 (1951)
44. Taha, H.E., Tahmasian, S., Woolsey, C.A., Nayfeh, A.H., M.R, Hajj: The need for higher-order averaging in the stability analysis of hovering, flapping-wing flight. Bioinspir. Biomim. **10**(1), 016,002 (2015)
45. Stephenson, A.: Xx. On induced stability. Lond. Edinb. Dublin Philos. Mag. J. Sci. **15**(86), 233–236 (1908)
46. Kapitza, P.L.: Dynamical stability of a pendulum when its point of suspension vibrates, and pendulum with a vibrating suspension. Collect. Pap. PL Kapitza **2**, 714–737 (1965)
47. Stabilization of an inverted pendulum under high-frequency excitation (Kapitza pendulum). https://www.youtube.com/watch?v=is_ejYsvAjY
48. Sun, M.: Insect flight dynamics: stability and control. Rev. Mod. Phys. **86**(2), 615 (2014)
49. Nelson, R.C.: Flight Stability and Automatic Control. McGraw-Hill, New York (1989)
50. Taha, H.E., Hajj, M.R., Nayfeh, A.H.: On the longitudinal flight dynamics of hovering MAVs/insects. J. Guid. Control Dyn. **37**(3), 970–978 (2014)
51. Alexander, D.E.: Nature's Flyers: Birds, Insects, and the Biomechanics of Flight. JHU Press (2002)
52. Taylor, G.K., Thomas, A.L.R.: Animal flight dynamics II. Longitudinal stability in flapping flight. J. Theor. Biol. **214**, 351–370 (2002)
53. Taylor, G.K., Thomas, A.L.: Dynamic flight stability in the desert locust schistocerca gregaria. J. Exp. Biol. **206**(16), 2803–2829 (2003)
54. Sun, M., Xiong, Y.: Dynamic flight stability of a hovering bumblebee. J. Exp. Biol. **208**(3), 447–459 (2005)
55. Sun, M., Wang, J., Xiong, Y.: Dynamic flight stability of hovering insects. Acta Mech. Sin. **23**(3), 231–246 (2007)
56. Xiong, Y., Sun, M.: Dynamic flight stability of a bumblebee in forward flight. Acta Mech. Sin. **24**(1), 25–36 (2008)
57. Gao, N., Aono, H., Liu, H.: A numerical analysis of dynamic flight stability of hawkmoth hovering. J. Biomech. Sci. Eng. **4**(1), 105–116 (2009)
58. Faruque, I., Humbert, J.S.: Dipteran insect flight dynamics. Part 1. Longitudinal motion about hover. J. Theor. Biol **264**(2), 538–552 (2010)

59. Cheng, B., Deng, X.: Translational and rotational damping of flapping flight and its dynamics and stability at hovering. IEEE Trans. Robot. **27**(5), 849–864 (2011)
60. Khan, Z.A., Agrawal, S.K.: Control of longitudinal flight dynamics of a flapping-wing micro air vehicle using time-averaged model and differential flatness based controller. In: American Control Conference, 2007. ACC'07, pp. 5284–5289. IEEE (2007)
61. Schenato, L., Campolo, D., Sastry, S.: Controllability issues in flapping flight for biomimetic micro aerial vehicles (MAVs). In: Proceedings of the 42nd IEEE Conference on Decision and Control, vol. 6, pp. 6441–6447. IEEE (2003)
62. Nayfeh, S., Nayfeh, A.: Nonlinear interactions between two widely spaced modes—external excitation. Int. J. Bifurc. Chaos **3**(02), 417–427 (1993)
63. Nayfeh, S., Nayfeh, A.: Energy transfer from high-to low-frequency modes in a flexible structure via modulation. J. Vib. Acoust. **116**(2), 203–207 (1994)
64. Nayfeh, A., Mook, D.: Energy transfer from high-frequency to low-frequency modes in structures. J. Vib. Acoust. **117**(B), 186–195 (1995)
65. Nayfeh, A.H., Chin, C.M.: Nonlinear interactions in a parametrically excited system with widely spaced frequencies. Nonlinear Dyn. **7**(2), 195–216 (1995)
66. Popovic, P., Nayfeh, A.H., Oh, K., Nayfeh, S.A.: An experimental investigation of energy transfer from a high-frequency mode to a low-frequency mode in a flexible structure. Modal Anal. **1**(1), 115–128 (1995)
67. Oh, K., Nayfeh, A.: High-to low-frequency modal interactions in a cantilever composite plate. J. Vib. Acoust. **120**(2), 579–587 (1998)
68. Taha, H.E., Nayfeh, A.H., Hajj, M.R.: Effect of the aerodynamic-induced parametric excitation on the longitudinal stability of hovering MAVs/insects. Nonlinear Dyn. **78**, 2399–2408 (2014). https://doi.org/10.1007/s11071-014-1596-6
69. Taha, H.E., Kiani, M., Hedrick, T.L., Greeter, J.S.M.: Vibrational control: A hidden stabilization mechanism in insect flight (under review). Nat. Commun
70. Taha, H.E., Hajj, M.R., Beran, P.S.: Unsteady nonlinear aerodynamics of hovering MAVs/insects. In: AIAA-Paper 2013–0504 (2013)