Averaging Optimal Control Systems with Two Fast Variables

(Working paper)

L. Dell'Elce[1], J.-B. Caillau[2], J.-B. Pomet[1]

[1]Univ. Côte d'Azur & Inria, France
[2]Univ. Côte d'Azur & CNRS/Inria, France
E-mail: lamberto.dell-elce@inria.fr, caillau@unice.fr, jean-baptiste.pomet@inria.fr

Abstract

Averaging is a valuable technique to gain understanding in the long-term evolution of dynamical systems characterized by slow and fast dynamics. Recent contributions proved that averaging can be applied to the extremal flow of optimal control problems. The present work extends these results by tackling averaging of time optimal systems with two fast variables. The first outcome consists of a justification of the application of the averaging principle to this problem. The key role of the adjoints of fast variables is then disclosed, which yields the assessment of a compatibility condition between that their boundary values in the original and averaged systems. A simplified relation is also obtained when a single fast variable is considered. The second outcome is devoted to the *a posteriori* reconstruction of short-period variations. The classical near-identity transformation exploited in dynamical system theory is shown to be inadequate to restore the adjoints of slow variables because of the peculiar form of their equations of motion. Hence, a consistent transformation is developed. Resonance effects are finally discussed. The methodology is applied to a time-optimal low-thrust orbital transfer in the Earth-Moon system.

Keywords: multiphase averaging, slow-fast dynamics, optimal control, near-identity transformation

Introduction

When the state of a dynamical system can be decomposed into slow and fast oscillatory components, averaging the equations of motion over the instantaneous period of the fast variables is a valuable practice to simplify the dynamics of the system and gain understanding of the long-term evolution of the flow. If the estimation of the fast variations of the trajectory is envisaged, a near-identity transformation that restores them can be developed. We provide evidence that existing theorems on double averaging of dynamical systems cannot be directly applied to trajectories of the controlled system. Specifically, we emphasize the key role of the adjoint of fast variables

The paper is organized as follows. We first introduce the slow-fast minimum time control problem that we wish to study in Section 1. After reviewing the application of Pontrjagin maximum principle to such problems, we provide a toy problem whose features help to illustrate the difficulties of averaging control systems with several fast angles. Section 2 is devoted to giving a panorama of existing results on multiphase averaging. The prominent one is Neishtadt theorem for two-phase systems. We discuss in detail the crucial construction of an *ad hoc* near-identity transform. In Section 3, averaging is applied to the extremal flow of minimum time control problems with two angles. One of the main issues is to decide whether the adjoint variables are slow or not. The analysis is then illustrated on the toy problem previously devised. Section 4 addresses another key issue, namely the transformation of the initial state and costate. In Section 5, the effect of resonance is studied. As can be shown either on the toy problem, or on the difficult orbit transfer control problem stemming from space mechanics worked out in the final Section 6, resonances have to be taken into account by means of suitable normal forms when averaging.

1 Optimal control problem with slow and fast dynamics

We consider the time optimal maneuvering of a dynamical system characterized by fast and slow dynamics, namely

$$\begin{aligned}
&\min_{\|\boldsymbol{u}\|\leq 1} \; t_f \quad \text{subject to:}\\
&\qquad \frac{\mathrm{d}\boldsymbol{I}}{\mathrm{d}t} = \varepsilon\left[\boldsymbol{f}_0(\boldsymbol{I},\,\boldsymbol{\varphi}) + \sum_{i=1}^{m}\boldsymbol{f}_i(\boldsymbol{I},\,\boldsymbol{\varphi})\,u_i\right],\\
&\qquad \frac{\mathrm{d}\boldsymbol{\varphi}}{\mathrm{d}t} = \varepsilon\left[\boldsymbol{g}_0(\boldsymbol{I},\,\boldsymbol{\varphi}) + \sum_{i=1}^{m}\boldsymbol{g}_i(\boldsymbol{I},\,\boldsymbol{\varphi})\,u_i\right] + \boldsymbol{\omega}\,(\boldsymbol{I}),\\
&\qquad \boldsymbol{I}(0) = \boldsymbol{I}_0,\\
&\qquad \boldsymbol{I}(t_f) = \boldsymbol{I}_f.
\end{aligned} \tag{1}$$

Here, the cost function, t_f is the maneuvering time, ε is a small parameter, and $\boldsymbol{u}$ denotes the m-dimensional control. Slow variables, $\boldsymbol{I}$, are defined on a smooth n-dimensional manifold $\mathcal{I}$, and are characterized by ε-order dynamics. We limit this study to systems with two fast angle variables, so that $\boldsymbol{\varphi}$ is defined on the two-dimensional torus, $\mathbb{T}^2$. The frequency vector, $\boldsymbol{\omega} : \mathcal{I} \to \mathbb{R}^2\backslash\{0\}$, determines the fast dynamics of $\boldsymbol{\varphi}$. Fields

$\boldsymbol{f}_j : \mathcal{I} \times \mathbb{T}^2 \to \mathbb{R}^n$ and $\boldsymbol{g}_j : \mathcal{I} \times \mathbb{T}^2 \to \mathbb{R}^2$ are assumed to be periodic with respect to $\boldsymbol{\varphi}$ and analytic on the continuation of $\boldsymbol{\varphi}$ in a non-vanishing complex strip. We note that $\boldsymbol{\varepsilon}$ is not included in the arguments of $\boldsymbol{f}_j$ and $\boldsymbol{g}_j$ to relieve the notation. However, all results of the manuscript can be extended to the case where the explicit dependency on this parameter is required. Boundary conditions are enforced on all components of $\boldsymbol{I}$ at the initial and final time, $\boldsymbol{I}_0$ and $\boldsymbol{I}_f$, respectively. These constraints are sufficient for the purpose of this study, so that more general boundary conditions on slow variables are not considered. The possibility to constrain the boundary values of $\boldsymbol{\varphi}$ is discussed in Section 4.1

1.1 Necessary conditions for optimality

Denote by $\boldsymbol{p_I}$ and $\boldsymbol{p_\varphi}$ the adjoints of slow and fast variables, respectively. The application of the Pontryagin maximum principle (PMP) to Problem (1) yields the Hamiltonian of the extremal flow,

$$H = \boldsymbol{p_\varphi} \cdot \boldsymbol{\omega}(\boldsymbol{I}) + \varepsilon K\left(\boldsymbol{I}, \boldsymbol{p_I}, \boldsymbol{\varphi}, \boldsymbol{p_\varphi}\right), \tag{2}$$

where the function $K : T^*\mathcal{I} \times T^*\mathbb{T}^2 \to \mathbb{R}$ that characterizes the slow component of the Hamiltonian is

$$K := H_0 + \sqrt{\sum_{i=1}^{m} H_i^2}, \tag{3}$$

and H_j, for $j = 0, \dots, m$, are defined as

$$H_j := \boldsymbol{f}_j(\boldsymbol{I}, \boldsymbol{\varphi}) \cdot \boldsymbol{p_I} + \boldsymbol{g}_j(\boldsymbol{I}, \boldsymbol{\varphi}) \cdot \boldsymbol{p_\varphi}.$$

Necessary conditions for optimality of Problem (1) consist of the flow associated to the Hamiltonian of Eq. (2),

$$\begin{aligned} \frac{\mathrm{d}\boldsymbol{I}}{\mathrm{d}t} &= \varepsilon \frac{\partial K}{\partial \boldsymbol{p_I}}, & \frac{\mathrm{d}\boldsymbol{p_I}}{\mathrm{d}t} &= -\varepsilon \frac{\partial K}{\partial \boldsymbol{I}} - \boldsymbol{p_\varphi} \frac{\partial \boldsymbol{\omega}}{\partial \boldsymbol{I}}, \\ \frac{\mathrm{d}\boldsymbol{\varphi}}{\mathrm{d}t} &= \varepsilon \frac{\partial K}{\partial \boldsymbol{p_\varphi}} + \boldsymbol{\omega}(\boldsymbol{I}), & \frac{\mathrm{d}\boldsymbol{p_\varphi}}{\mathrm{d}t} &= -\varepsilon \frac{\partial K}{\partial \boldsymbol{\varphi}}, \end{aligned} \tag{4}$$

and of the boundary conditions,

$$\begin{aligned} \boldsymbol{I}(0) &= \boldsymbol{I}_0, & \boldsymbol{p_\varphi}(0) &= \boldsymbol{0}, \\ \boldsymbol{I}(t_f) &= \boldsymbol{I}_f, & \boldsymbol{p_\varphi}(t_f) &= \boldsymbol{0}. \end{aligned} \tag{5}$$

The components of the maximizing control are

$$u_j^{opt}\left(\boldsymbol{I}, \boldsymbol{p_I}, \boldsymbol{\varphi}, \boldsymbol{p_\varphi}\right) = \frac{H_j}{\sqrt{\sum_{i=0}^{m} H_i^2}} \qquad j = 1, \dots, m.$$

In view of the exploitation of shooting-based techniques, triads $\left(t_f,\ \boldsymbol{p_I}_0,\ \boldsymbol{\varphi}_0\right)$ are referred to as candidate solutions for Problem (1) if trajectories of System (19) with initial conditions $\boldsymbol{I}(0) = \boldsymbol{I}_0$, $\boldsymbol{p_I}(0) = \boldsymbol{p_I}_0$, $\boldsymbol{\varphi}(0) = \boldsymbol{\varphi}_0$, and $\boldsymbol{p_\varphi}_0 = \boldsymbol{0}$ satisfy the endpoint boundary conditions in Eq. (5), i.e., candidate solutions are zeros of the shooting function

$$S\left(t_f, \boldsymbol{p_I}_0, \boldsymbol{\varphi}_0\right) := \left\{ \begin{array}{c} \boldsymbol{I}(t_f|\boldsymbol{I}_0, \boldsymbol{p_I}_0, \boldsymbol{\varphi}_0, \boldsymbol{0}) - \boldsymbol{I}_f \\ \boldsymbol{p_\varphi}(t_f|\boldsymbol{I}_0, \boldsymbol{p_I}_0, \boldsymbol{\varphi}_0, \boldsymbol{0}) - \boldsymbol{0} \\ \|\boldsymbol{p_I}_0\| - 1 \end{array} \right\}, \tag{6}$$

and are such that $H\left(\boldsymbol{I}_0, \boldsymbol{p_I}_0, \boldsymbol{\varphi}_0, \boldsymbol{p_\varphi}_0, \varepsilon\right) > 0$. The last equation of the shooting function is aimed at fixing the arbitrary scaling factor of the initial adjoints, which is due to the homogeneity of the Hamiltonian. The very-specific choice of constraining the norm of $\boldsymbol{p_I}_0$ to 1 is arbitrary and not unique, but it reveals to be practical in Section 3.

1.2 Toy problem

A simple case study is introduced to streamline the flow of the discussion. Numerical simulations of this example are used to provide graphical support to the various discussions and mathematical developments of the paper. The dynamical system consists of a scalar slow variable, I, and two fast variables, ζ and ψ, and the optimal control problem is

$$\begin{aligned} &\min_{\sqrt{u_1^2+u_2^2}\le 1} \quad t_f \quad \text{subject to :} \\ &\qquad \frac{\mathrm{d}I}{\mathrm{d}t} = \varepsilon\left[\cos\zeta + \cos\left(\zeta-\psi\right)u_1 + u_2\right], \quad \frac{\mathrm{d}\zeta}{\mathrm{d}t} = I, \quad \frac{\mathrm{d}\psi}{\mathrm{d}t} = 1, \\ &\qquad I(0) = I_0, \quad I(t_f) = I_f. \end{aligned} \tag{7}$$

We note that the frequency of ψ is constant. This assumption does not yield oversimplified results with respect to more involved examples. In fact, if one of the two frequencies is non-vanishing on the manifold $\mathcal{I}$, any problem with two frequencies can be recast into this form by means of a change of the time variable, as emphasized in [3, Chap. 4].

The Hamiltonian associated to Problem (7) is

$$H = Ip_\zeta + p_\psi + \varepsilon\left[p_I\cos\zeta + |p_I|\sqrt{1+\cos^2\left(\zeta-\psi\right)}\right]. \tag{8}$$

This yields the equations of motion

$$\begin{aligned} &\frac{\mathrm{d}I}{\mathrm{d}t} = \varepsilon\left[\cos\zeta + \frac{p_I}{|p_I|}\sqrt{1+\cos^2\left(\zeta-\psi\right)}\right], && \frac{\mathrm{d}p_I}{\mathrm{d}t} = -p_\zeta, \\ &\frac{\mathrm{d}\zeta}{\mathrm{d}t} = I, && \frac{\mathrm{d}p_\zeta}{\mathrm{d}t} = \varepsilon\left[p_I\sin\zeta + |p_I|\frac{\cos\left(\zeta-\psi\right)\sin\left(\zeta-\psi\right)}{\sqrt{1+\cos^2\left(\zeta-\psi\right)}}\right], \\ &\frac{\mathrm{d}\psi}{\mathrm{d}t} = 1, && \frac{\mathrm{d}p_\psi}{\mathrm{d}t} = -\varepsilon|p_I|\frac{\cos\left(\zeta-\psi\right)\sin\left(\zeta-\psi\right)}{\sqrt{1+\cos^2\left(\zeta-\psi\right)}}, \end{aligned} \tag{9}$$

and the maximizing control

$$u_1^{opt} = \frac{p_I}{|p_I|}\frac{\cos\left(\zeta-\psi\right)}{\sqrt{1+\cos^2\left(\zeta-\psi\right)}}, \quad u_2^{opt} = \frac{p_I}{|p_I|}\frac{1}{\sqrt{1+\cos^2\left(\zeta-\psi\right)}}. \tag{10}$$

Equations (9) and (10) reveal that the sign of p_I determines the direction of the control vector, which, in turn, imposes a secular drift to the slow variable. Numerical values used in all simulations are $\varepsilon = 10^{-3}$ and $I_0 = \frac{\sqrt{2}}{2}$. Simulation-specific values are listed in the captions of the figures.

2 Two-phase averaging of fast-oscillating uncontrolled systems

This section is devoted to the averaging of the uncontrolled counterpart of the dynamical system introduced in Eq. (1), namely

$$\frac{\mathrm{d}\boldsymbol{I}}{\mathrm{d}t} = \varepsilon \boldsymbol{f}_0(\boldsymbol{I}, \boldsymbol{\varphi}), \quad \frac{\mathrm{d}\boldsymbol{\varphi}}{\mathrm{d}t} = \varepsilon \boldsymbol{g}_0(\boldsymbol{I}, \boldsymbol{\varphi}) + \boldsymbol{\omega}(\boldsymbol{I}). \tag{11}$$

The motion of $\boldsymbol{I}$ is characterized by a slow trend perturbed by ε-small oscillations. The gross behavior of $\boldsymbol{I}$ can be understood by filtering out these oscillations via the averaging of the equations of motion with respect to the two-dimensional torus. This yields the averaged version of System (11), namely

$$\frac{\mathrm{d}\overline{\boldsymbol{I}}}{\mathrm{d}t} = \varepsilon \overline{\boldsymbol{f}_0}\left(\overline{\boldsymbol{I}}\right), \quad \frac{\mathrm{d}\overline{\boldsymbol{\varphi}}}{\mathrm{d}t} = \varepsilon \overline{\boldsymbol{g}_0}\left(\overline{\boldsymbol{I}}\right) + \boldsymbol{\omega}\left(\overline{\boldsymbol{I}}\right), \tag{12}$$

where averaged vector fields, $\overline{\boldsymbol{f}_0} : \mathcal{I} \to \mathbb{R}$ and $\overline{\boldsymbol{g}_0} : \mathcal{I} \to \mathbb{R}$, are defined as

$$\overline{\boldsymbol{f}_0}\left(\overline{\boldsymbol{I}}\right) := \frac{1}{4\pi^2}\int_{\mathbb{T}^2} \boldsymbol{f}_0\left(\overline{\boldsymbol{I}}, \boldsymbol{\varphi}\right)\mathrm{d}\boldsymbol{\varphi}, \quad \overline{\boldsymbol{g}_0}\left(\overline{\boldsymbol{I}}\right) := \frac{1}{4\pi^2}\int_{\mathbb{T}^2} \boldsymbol{g}_0\left(\overline{\boldsymbol{I}}, \boldsymbol{\varphi}\right)\mathrm{d}\boldsymbol{\varphi}.$$

Averaged fields are independent of fast variables by definition. As such, the motion of $\overline{\boldsymbol{I}}$, can be integrated before $\overline{\boldsymbol{\varphi}}$, which can be eventually evaluated *a posteriori* if required. When ε explicitly appears as an argument of $\boldsymbol{f}_0$ and $\boldsymbol{g}_0$, the averaging is carried out by considering the limit of these functions as ε approaches zero. It is desirable that trajectories $\boldsymbol{I}(t)$ and $\overline{\boldsymbol{I}}(t)$ emanated from the same point on the manifold $\mathcal{I}$ remain "close" for "long" time. Compared to single-frequency systems, this question is non-trivial because the double average is not a good approximation of the original systems whenever the two frequencies are nearly commensurate. Under arguably restrictive assumptions, the Neishtadt theorem provides an estimate of the drift between trajectories of the original and averaged systems that rigorously accounts for the error due to the use of double averaging inside resonant zones.

Section 2.1 recalls the Neishtadt theorem [4] which provides an optimal estimate of the drift between trajectories of Systems (11) and (12). Section 2.2 details the near-identity transformation aimed at restoring the fast oscillations of an averaged trajectory, $\overline{\boldsymbol{I}}(t), \overline{\boldsymbol{\varphi}}(t)$.

2.1 The Neishtadt theorem

Assume that there exist $\mathcal{I}_0 \subseteq \mathcal{I}$ such that all trajectories of the averaged slow variables, $\overline{\boldsymbol{I}}(t)$, with initial conditions in $\mathcal{I}_0$ satisfy[1]

$$\overline{\boldsymbol{I}}(t) \in \mathcal{I} \quad \text{and} \quad \left| \left(\omega_1 \left(\overline{\boldsymbol{I}} \right) \frac{\partial \omega_2}{\partial \overline{\boldsymbol{I}}} - \omega_2 \left(\overline{\boldsymbol{I}} \right) \frac{\partial \omega_1}{\partial \overline{\boldsymbol{I}}} \right) \cdot \overline{\boldsymbol{f}_0} \left(\overline{\boldsymbol{I}} \right) \right| > 0 \qquad \forall \, t \in \left[0, \frac{1}{\varepsilon} \right]. \tag{13}$$

Then, there exist a partition $\{\mathcal{V}_1, \mathcal{V}_2\}$ of $\mathcal{I}_0 \times \mathbb{T}^2$ and constants $\{c_1, c_2\} = \mathcal{O}(1)$ such that

$$\sup_{t \in \left[0, \frac{1}{\varepsilon}\right]} \left\| \boldsymbol{I}(t) - \overline{\boldsymbol{I}}(t) \right\| < c_1 \sqrt{\varepsilon} \log \left(\frac{1}{\varepsilon} \right) \quad \forall \left\{ \boldsymbol{I}(0), \boldsymbol{\varphi}(0) \right\} \in \mathcal{V}_1, \; \overline{\boldsymbol{I}}(0) = \boldsymbol{I}(0), \tag{14}$$

and

$$\mu\left(\mathcal{V}_2\right) \leq c_2 \sqrt{\varepsilon},$$

where $\mu\left(\mathcal{V}_2\right)$ denotes an ordinary measure on $\mathcal{I} \times \mathbb{T}^2$. In layman's terms, this theorem states that $\overline{\boldsymbol{I}}(t)$ is a good approximation of $\boldsymbol{I}(t)$, i.e., $\sqrt{\varepsilon} \log \frac{1}{\varepsilon} = 0$ as ε approaches zero, for most initial conditions, since the size of the "bad" set $\mathcal{V}_2$ is bounded by the square root of ε. However, although $\mathcal{V}_2$ vanishes for very small ε, its elements uniformly fill the phase space. The assumption of Eq. (13) guarantees that the frequency ratio of the averaged trajectory, $\omega_1\left(\overline{\boldsymbol{I}}\right)/\omega_2\left(\overline{\boldsymbol{I}}\right)$, evolves monotonically in time. Hence, any resonance is crossed transversally with non-vanishing speed, so that the cumulated error due to the wrong modeling of the motion inside resonant zones is small. Trajectories of the original system emanated from $\mathcal{V}_2$ experience capture into resonance, i.e., they spend long time inside a single resonant zone. The simple double average over the two-dimensional torus ignores the phase lock specific of this resonance, so that the doubly-averaged system is unable to adequately approximate the motion of the original system during this possibly-very-long period of time. A detailed proof of the theorem is available in [3, Chap. 4].

2.2 Near-identity transformation of the initial conditions

Short-period variations of averaged trajectories of slow variables can be restored as a function of the averaged state itself. A large body of literature discusses this process, e.g., [5, Chap. 7], [2, Chap. 2]. Denote by $\hat{\boldsymbol{I}}$ and $\hat{\boldsymbol{\varphi}}$ the reconstructed osculating slow and fast variables, respectively. A transformation, $\boldsymbol{\nu} : \mathcal{I} \times \mathbb{T}^2 \to \mathcal{I} \times \mathbb{T}^2$, can be developed such that

$$\left\{ \begin{array}{c} \hat{\boldsymbol{I}} \\ \hat{\boldsymbol{\varphi}} \end{array} \right\} = \left\{ \begin{array}{c} \overline{\boldsymbol{I}} \\ \overline{\boldsymbol{\varphi}} \end{array} \right\} + \varepsilon \boldsymbol{\nu} \left(\overline{\boldsymbol{I}}, \overline{\boldsymbol{\varphi}} \right). \tag{15}$$

The objective of the transformation is the establishment of second-order matching between the time derivative of the reconstructed variables and the right-hand side of the original

[1] We note that ε could be re-scaled to fit a desired time window of size $\mathcal{O}\left(\frac{1}{\varepsilon}\right)$.

system in Eq. (11):

$$
\begin{aligned}
\frac{\mathrm{d}\hat{\boldsymbol{I}}}{\mathrm{d}t} &= \frac{\mathrm{d}}{\mathrm{d}t}\left[\overline{\boldsymbol{I}} + \varepsilon \boldsymbol{\nu}_{\boldsymbol{I}}\left(\overline{\boldsymbol{I}}, \overline{\boldsymbol{\varphi}}\right)\right] &&= \varepsilon \boldsymbol{f}_0\left(\hat{\boldsymbol{I}}, \hat{\boldsymbol{\varphi}}\right) + O\left(\varepsilon^2\right), \\
\frac{\mathrm{d}\hat{\boldsymbol{\varphi}}}{\mathrm{d}t} &= \frac{\mathrm{d}}{\mathrm{d}t}\left[\overline{\boldsymbol{\varphi}} + \varepsilon \boldsymbol{\nu}_{\boldsymbol{\varphi}}\left(\overline{\boldsymbol{I}}, \overline{\boldsymbol{\varphi}}\right)\right] &&= \boldsymbol{\omega}\left(\hat{\boldsymbol{I}}\right) + \varepsilon \boldsymbol{g}_0\left(\hat{\boldsymbol{I}}, \hat{\boldsymbol{\varphi}}\right) + O\left(\varepsilon^2\right),
\end{aligned}
$$

where $\boldsymbol{\nu}_{\boldsymbol{I}}$ and $\boldsymbol{\nu}_{\boldsymbol{\varphi}}$ denote the components of $\boldsymbol{\nu}$ associated to the slow and fast variables, respectively. In addition, it is desirable that reconstructed trajectories oscillate with zero mean about the averaged ones. These constraints yield the system of partial differential equations (PDE)

$$
\begin{aligned}
\omega_1\left(\overline{\boldsymbol{I}}\right)\frac{\partial \boldsymbol{\nu}_{\boldsymbol{I}}}{\partial \overline{\varphi}_1} + \omega_2\left(\overline{\boldsymbol{I}}\right)\frac{\partial \boldsymbol{\nu}_{\boldsymbol{I}}}{\partial \overline{\varphi}_2} &= \boldsymbol{f}_0\left(\overline{\boldsymbol{I}}, \overline{\boldsymbol{\varphi}}\right) - \overline{\boldsymbol{f}_0}\left(\overline{\boldsymbol{I}}\right) + O(\varepsilon), \\
\left.\frac{\partial \boldsymbol{\omega}}{\partial \boldsymbol{I}}\right|_{\overline{\boldsymbol{I}}} \boldsymbol{\nu}_{\boldsymbol{I}} + \omega_1\left(\overline{\boldsymbol{I}}\right), \frac{\partial \boldsymbol{\nu}_{\boldsymbol{\varphi}}}{\partial \overline{\varphi}_1} + \omega_2\left(\overline{\boldsymbol{I}}\right)\frac{\partial \boldsymbol{\nu}_{\boldsymbol{\varphi}}}{\partial \overline{\varphi}_2} &= \boldsymbol{g}_0\left(\overline{\boldsymbol{I}}, \overline{\boldsymbol{\varphi}}\right) - \overline{\boldsymbol{g}_0}\left(\overline{\boldsymbol{I}}\right) + O(\varepsilon), \\
\int_{\mathbb{T}^2} \boldsymbol{\nu}\left(\overline{\boldsymbol{I}}, \boldsymbol{\varphi}\right)\,\mathrm{d}\boldsymbol{\varphi} &= \boldsymbol{0}.
\end{aligned}
\tag{16}
$$

Equation (16) can be solved by first evaluating $\boldsymbol{\nu}_{\boldsymbol{I}}$ and then $\boldsymbol{\nu}_{\boldsymbol{\varphi}}$. We note that any first-order solution of this problem is valuable.

Let $\boldsymbol{f}_0^{(\boldsymbol{k})}\left(\overline{\boldsymbol{I}}\right)$ be the coefficients of the Fourier series of $\boldsymbol{f}_0\left(\overline{\boldsymbol{I}}, \overline{\boldsymbol{\varphi}}\right) - \overline{\boldsymbol{f}_0}\left(\overline{\boldsymbol{I}}\right)$. The magnitude of $\left|\boldsymbol{f}_0^{(\boldsymbol{k})}\left(\overline{\boldsymbol{I}}\right)\right|$ is bounded by an exponentially-decreasing function of $|\boldsymbol{k}| = |k_1| + |k_2|$ because of the assumptions on the analyticity of $\boldsymbol{f}_0$. Consequently, there is a constant c_3 such that for $N \geq c_3 \log \dfrac{1}{\varepsilon}$

$$
\boldsymbol{f}_0\left(\overline{\boldsymbol{I}}, \overline{\boldsymbol{\varphi}}\right) = \sum_{0 \leq |\boldsymbol{k}| \leq N} \mathrm{e}^{i\boldsymbol{k}\cdot\overline{\boldsymbol{\varphi}}}\, \boldsymbol{f}_0^{(\boldsymbol{k})}\left(\overline{\boldsymbol{I}}\right) + O\left(\varepsilon\right).
$$

Assume that the state of the averaged system is outside of any resonant zone of order smaller than $N \geq c_3 \log \dfrac{1}{\varepsilon}$, i.e., there is a constant c_4 such that

$$
\boldsymbol{k} \cdot \boldsymbol{\omega}\left(\overline{\boldsymbol{I}}\right) \geq c_4\sqrt{\varepsilon} \qquad \forall\ \boldsymbol{k} \in \mathbb{Z}^2,\ 0 < |\boldsymbol{k}| \leq N.
$$

Hence, the transformation $\boldsymbol{\nu}_{\boldsymbol{I}}$ solving Eq. (16) is given by

$$
\boldsymbol{\nu}_{\boldsymbol{I}}\left(\overline{\boldsymbol{I}}, \overline{\boldsymbol{\varphi}}\right) = -i \sum_{0 < |\boldsymbol{k}| \leq N} \frac{\mathrm{e}^{i\boldsymbol{k}\cdot\overline{\boldsymbol{\varphi}}}}{\boldsymbol{k} \cdot \boldsymbol{\omega}\left(\overline{\boldsymbol{I}}\right)}\, \boldsymbol{f}_0^{(\boldsymbol{k})}\left(\overline{\boldsymbol{I}}\right).
\tag{17}
$$

Finally, the solution $\boldsymbol{\nu}_{\boldsymbol{\varphi}}$ can be assessed by means of an equation analogous to Eq. (17), where the coefficients of the series expansion of

$$
\boldsymbol{g}_0\left(\overline{\boldsymbol{I}}, \overline{\boldsymbol{\varphi}}\right) - \overline{\boldsymbol{g}}_0\left(\overline{\boldsymbol{I}}\right) - \left.\frac{\partial \boldsymbol{\omega}}{\partial \boldsymbol{I}}\right|_{\overline{\boldsymbol{I}}} \boldsymbol{\nu}_{\boldsymbol{I}}\left(\overline{\boldsymbol{I}}, \overline{\boldsymbol{\varphi}}\right)
$$

are used instead of $\boldsymbol{f}_0^{(\boldsymbol{k})}\left(\overline{\boldsymbol{I}}\right)$. When frequencies are nearly commensurate, a resonant averaged form of the system needs to be used instead of the double average. This problem is discussed in Section (5). Because the transformation amends ε-small correction to the averaged state. The inverse transformation of Eq. (15) can be approximated by

$$\left\{ \begin{array}{c} \bar{\boldsymbol{I}} \\ \bar{\boldsymbol{\varphi}} \end{array} \right\} = \left\{ \begin{array}{c} \hat{\boldsymbol{I}} \\ \hat{\boldsymbol{\varphi}} \end{array} \right\} + \varepsilon \boldsymbol{\nu}^{-1} \left(\hat{\boldsymbol{I}}, \hat{\boldsymbol{\varphi}}\right) \approx \left\{ \begin{array}{c} \hat{\boldsymbol{I}} \\ \hat{\boldsymbol{\varphi}} \end{array} \right\} - \varepsilon \boldsymbol{\nu} \left(\hat{\boldsymbol{I}}, \hat{\boldsymbol{\varphi}}\right).$$

3 The averaged control system

Applying averaging theory to the extremal flow detailed in Eq. (19) is questionable because the structure of this vector field differs from the conventional fast-oscillating system, e.g., Eq. (11). Specifically, the equation of motion of $\boldsymbol{p_I}$, includes the term $\boldsymbol{p_\varphi} \dfrac{\partial \boldsymbol{\omega}}{\partial \boldsymbol{I}}$ that may possibly be of order larger than ε. Hence, adjoints of slow variables are not necessary slow themselves. This section justifies the application of averaging theory to System (19) by showing that adjoints of fast variables are systematically ε-small for any extremal trajectory with free phases, and, as such, $\dfrac{\mathrm{d}\boldsymbol{p_I}}{\mathrm{d}t} = O(\varepsilon)$ when restrained to these trajectories.

Consider the canonical change of variables $\left\{\boldsymbol{I}, \boldsymbol{p_I}, \boldsymbol{\varphi}, \boldsymbol{p_\varphi}\right\} \to \left\{\boldsymbol{J}, \boldsymbol{p_J}, \boldsymbol{\psi}, \boldsymbol{p_\psi}\right\}$ such that

$$\boldsymbol{J} = \boldsymbol{I}, \qquad \boldsymbol{\psi} = \boldsymbol{\Omega}(\boldsymbol{I})\, \boldsymbol{\varphi},$$

where the matrix-valued function, $\boldsymbol{\Omega} : \mathcal{I} \to \mathbb{R}^{2\times 2}$ is defined as

$$\boldsymbol{\Omega} := \frac{1}{||\boldsymbol{\omega}(\boldsymbol{I})||} \begin{bmatrix} \omega_1(\boldsymbol{I}) & \omega_2(\boldsymbol{I}) \\ -\omega_2(\boldsymbol{I}) & \omega_1(\boldsymbol{I}) \end{bmatrix}.$$

Symplectic constraints yield the transformation of the adjoints

$$\boldsymbol{p_I} = \boldsymbol{p_J} + \boldsymbol{p_\psi} \frac{\partial \boldsymbol{\Omega}}{\partial \boldsymbol{J}} \boldsymbol{\Omega}^T\, \boldsymbol{\psi}, \qquad \boldsymbol{p_\varphi} = \boldsymbol{p_\psi}\, \boldsymbol{\Omega}(\boldsymbol{J}),$$

so that the transformed Hamiltonian is

$$\tilde{H} = ||\boldsymbol{\omega}(\boldsymbol{J})|| p_{\psi_1} + \varepsilon \underbrace{K \left(\boldsymbol{J},\ \boldsymbol{p_J} + \boldsymbol{p_\psi} \frac{\partial \boldsymbol{\Omega}}{\partial \boldsymbol{J}} \boldsymbol{\Omega}^T\, \boldsymbol{\psi},\ \boldsymbol{\Omega}^T\, \boldsymbol{\psi},\ \boldsymbol{p_\psi}\, \boldsymbol{\Omega}\right)}_{:=\ \tilde{K}(\boldsymbol{J}, \boldsymbol{p_J}, \boldsymbol{\psi}, \boldsymbol{p_\psi})}.$$

Boundary conditions on the adjoints of fast variables require that $\boldsymbol{p_\varphi}(0) = \boldsymbol{0}$. Evaluating the Hamiltonian at the initial time and considering the normalization of the initial adjoints proposed in Eq. (6), i.e., $||\boldsymbol{p_I}_0|| = 1$, yields the Hamiltonian level

$$\varepsilon h := \tilde{H}(t=0) = \varepsilon \underbrace{K \left(\boldsymbol{I}_0,\ \boldsymbol{p_I}_0,\ \boldsymbol{\Omega}^T (\boldsymbol{I}_0)\ \boldsymbol{\psi}_0,\ \boldsymbol{0}\right)}_{O(1)}.$$

Hence, p_{ψ_1} can be evaluated at any time by solving the implicit function

$$p_{\psi_1} = \varepsilon \frac{h - \tilde{K}\left(\boldsymbol{J}, \boldsymbol{p_J}, \boldsymbol{\psi}, \boldsymbol{p_\psi}\right)}{||\boldsymbol{\omega}(\boldsymbol{J})||} \approx \frac{h - \tilde{K}\left(\boldsymbol{J}, \boldsymbol{p_J}, \boldsymbol{\psi}, \boldsymbol{0}\right)}{||\boldsymbol{\omega}(\boldsymbol{J})||} \tag{18}$$

Equation (18) reveals that $p_{\psi_1} = O(\varepsilon)$ when evaluated on a candidate optimal trajectory. As a consequence, $\boldsymbol{p_J}$ exhibit ε-slow dynamics, i.e.,

$$\frac{\mathrm{d}\boldsymbol{p_J}}{\mathrm{d}t} = -\underbrace{\frac{\partial \|\boldsymbol{\omega}\|}{\partial \boldsymbol{J}} p_{\psi_1}}_{O(\varepsilon)} - \varepsilon \frac{\partial \tilde{K}}{\partial \boldsymbol{J}} = O(\varepsilon),$$

which justifies the averaging of the extremal flow.

Denote by $\overline{K}$ the averaged counterpart of the functional defined in Eq. (3), namely

$$\overline{K} := \frac{1}{4\pi^2} \int_{\mathbb{T}^2} K\left(\boldsymbol{I}, \boldsymbol{p_I}, \boldsymbol{\varphi}, \boldsymbol{0}\right) \, \mathrm{d}\boldsymbol{\varphi}.$$

Here, $\boldsymbol{p_\varphi} = \boldsymbol{0}$ because the averaging is carried out by considering the limit of the function as ε approaches zero. Averaging the extremal flow of Eq. (19) yields

$$\begin{aligned} \frac{\mathrm{d}\overline{\boldsymbol{I}}}{\mathrm{d}t} &= \varepsilon \frac{\partial \overline{K}}{\partial \overline{\boldsymbol{p_I}}}, & \frac{\mathrm{d}\overline{\boldsymbol{p_I}}}{\mathrm{d}t} &= -\varepsilon \frac{\partial \overline{K}}{\partial \overline{\boldsymbol{I}}} - \overline{\boldsymbol{p_\varphi}} \frac{\partial \boldsymbol{\omega}}{\partial \overline{\boldsymbol{I}}}, \\ \frac{\mathrm{d}\boldsymbol{\varphi}}{\mathrm{d}t} &= \varepsilon \frac{\partial \overline{K}}{\partial \overline{\boldsymbol{p_\varphi}}} + \boldsymbol{\omega}\left(\overline{\boldsymbol{I}}\right), & \frac{\mathrm{d}\overline{\boldsymbol{p_\varphi}}}{\mathrm{d}t} &= \boldsymbol{0}. \end{aligned} \tag{19}$$

Adjoints of the fast variables are indeed constant along averaged extremal trajectories.

3.1 Averaging the toy problem

The averaged counterpart of the Hamiltonian in Eq. (8) is

$$\overline{H} = \overline{I}\,\overline{p_\zeta} + \overline{p_\psi} + \frac{\sqrt{8}}{\pi} \mathrm{E}\left(\frac{1}{\sqrt{2}}\right) \varepsilon \left|\overline{p_I}\right| \approx \overline{I}\,\overline{p_\zeta} + \overline{p_\psi} + 1.216\varepsilon \left|\overline{p_I}\right|.$$

where $\mathrm{E}(x)$ denotes the complete elliptic integral of the second kind. The vector field associated to this Hamiltonian is

$$\begin{aligned} \frac{\mathrm{d}\overline{I}}{\mathrm{d}t} &\approx 1.216\,\varepsilon \frac{\overline{p_I}}{\left|\overline{p_I}\right|}, & \frac{\mathrm{d}\overline{p_I}}{\mathrm{d}t} &= -\overline{p_\zeta}, \\ \frac{\mathrm{d}\overline{\zeta}}{\mathrm{d}t} &= \overline{I}, & \frac{\mathrm{d}\overline{p_\zeta}}{\mathrm{d}t} &= 0, \\ \frac{\mathrm{d}\overline{\psi}}{\mathrm{d}t} &= 1, & \frac{\mathrm{d}\overline{p_\psi}}{\mathrm{d}t} &= 0, \end{aligned} \tag{20}$$

By definition, the integration of the slowly-changing variables of System (20) is independent of $\overline{\zeta}$ and $\overline{\psi}$. The closed-form solution of the slow averaged flow is

$$\begin{aligned} \overline{I} &\approx \overline{I}_0 + 1.216\,\varepsilon \frac{\overline{p_I}}{\left|\overline{p_I}\right|} \gamma(t), \\ \overline{p_I} &= \overline{p_I}_0 - \overline{p_\zeta}_0\, t, \\ \overline{p_\zeta} &= \overline{p_\zeta}_0, \\ \overline{p_\psi} &= \overline{p_\psi}_0. \end{aligned}$$

where subscript 0 is used to address initial conditions, and $\gamma(t)$ is defined as

$$\gamma(t) = \begin{cases} t & \text{if } \overline{p_I}_0\,\overline{p_\zeta}_0 \leq 0 \text{ or } t \leq \dfrac{\overline{p_I}_0}{\overline{p_\zeta}_0} \\ 2\,\dfrac{\overline{p_I}_0}{\overline{p_\zeta}_0} - t & \text{otherwise} \end{cases} \tag{21}$$

The slow variable $\overline{I}(t)$ coincides with the frequency ratio of the averaged system, and it evolves as a piecewise linear function of time. The slope of $\overline{I}(t)$ switches sign when $\overline{p_I}$ crosses zero. For a given initial condition, this can occur at most one time during the entrire trajectory.

4 Near-identity transformation of the initial state

Changing the initial conditions of averaged trajectories by means of the transformation discussed in Section 2.2 allows to reduce the drift between $\boldsymbol{I}(t)$ and $\overline{\boldsymbol{I}}(t)$. Qualitatively, the transformation shifts the initial point of the averaged trajectory right in the middle of the short-period oscillations of $\boldsymbol{I}(t)$. The improvement obtained with this expedient is possibly negligible when compared to the estimate provided by the Neishtadt theorem, which considers the same initial conditions for the two trajecotries. Nonetheless, the transformation of the initial variables plays a key role for the optimal control problem. To support this claim, consider Problem (7) and assume that initial conditions of $\overline{p_I}$ and $\overline{p_\zeta}$ are restrained to a compact set such that the switching event of $\gamma(t)$ outlined in Eq. (21) is attained not earlier than a desired integration time t_f for any trajectory originated from this set. Then, the frequency ratio of the averaged trajectory evolves monotonically and the Neishtadt theorem recalled in Section 2.1 should apply. However, Figure 1 shows that p_I and $\overline{p_I}$ exhibit a steady derive that largely exceeds the expected "small" drift quantified in Eq. (14) when Systems (9) and (20) are integrated with the same initial conditions. In addition, trajectories of the original system strongly depend on the initial angles. Section 4.1 shows that transforming the adjoints of fast variables is sufficient to drastically reduce the drift of $\boldsymbol{p_I}$. Section 4.2 discusses the transformation of the other variables.

4.1 The fundamental role of the adjoints of fast variables

The trigger at the origin of the drift of $\boldsymbol{p_I}$ is the wrong assessment of the averaged value of $\boldsymbol{p_\varphi}$, as shown in the bottom of Figure 1. This error is of order ε but it induces a steady drift of of $\overline{\boldsymbol{p_I}}$ of the same order of magnitude, i.e.,

$$\frac{\mathrm{d}\,\overline{\boldsymbol{p_I}}}{\mathrm{d}t} = \underbrace{-\overline{\boldsymbol{p_\varphi}}\frac{\partial\,\omega}{\partial\,\overline{\boldsymbol{I}}}}_{\varepsilon\text{-small error}} - \varepsilon\frac{\partial\,\overline{K}}{\partial\,\overline{\boldsymbol{I}}}.$$

In turn, an ε-small error on $\overline{\boldsymbol{p_\varphi}}$ induces a steady drift of $\overline{\boldsymbol{p_I}}$ that is comparable with its slow motion. Transforming the initial adjoints of fast variables by means of Eq. (15) is sufficient to greatly mitigate this problem, as shown in Figure 2. Here, initial conditions

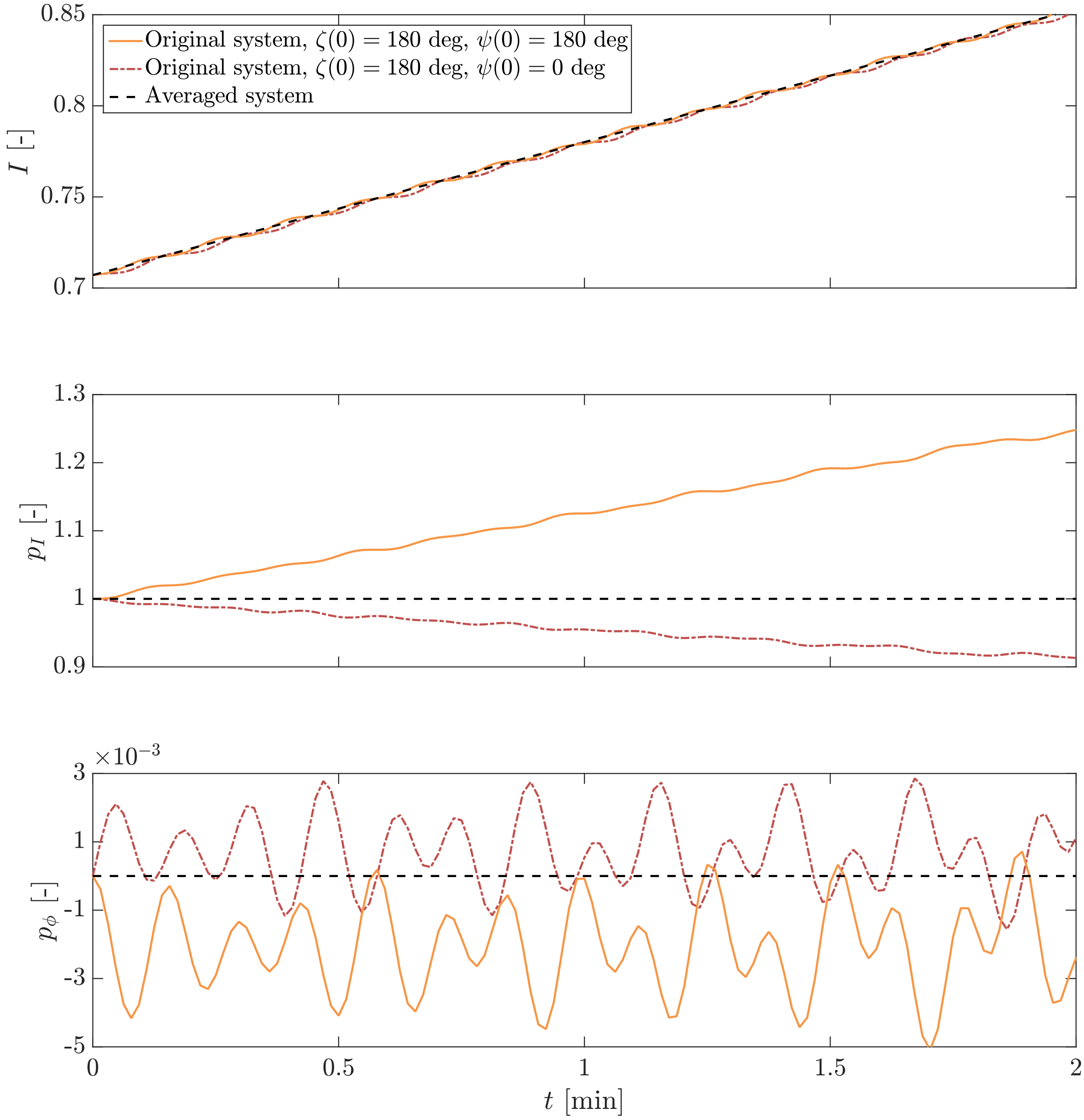


Figure 1: Numerical integration of the toy problem. Trajectories of the original and averaged system are emanated from the same point of the phase space. Initial adjoints are $p_I(0) = 1$ and $p_\psi(0) = p_\zeta(0) = 0$.

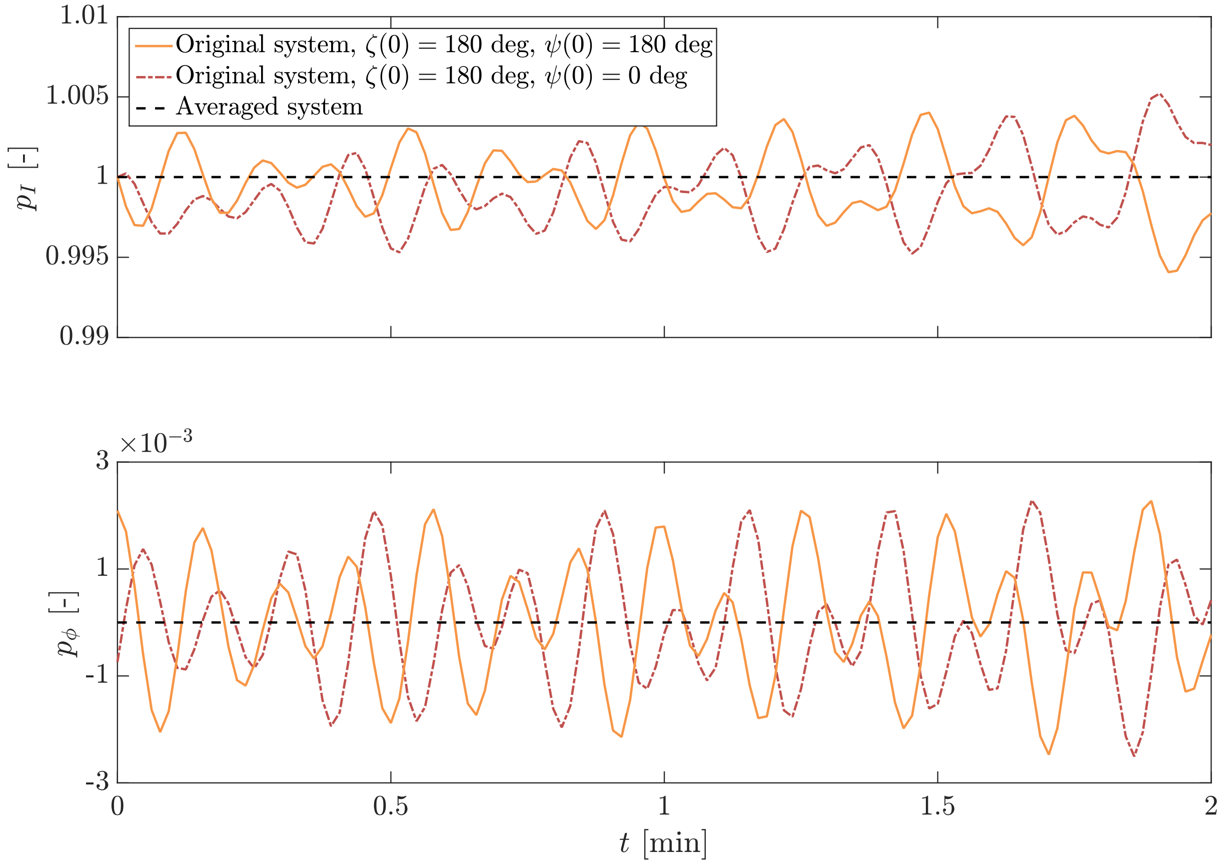


Figure 2: Numerical integration of the toy problem. Here, initial adjoints of fast variables are transformed by means of Eq. (17).

of the averaged and original initial value problems (IVP) are mostly the same, i.e.,

$$\boldsymbol{I}(0) = \overline{\boldsymbol{I}}(0) = \boldsymbol{I}_0, \quad \boldsymbol{p_I}(0) = \overline{\boldsymbol{p_I}}(0) = \boldsymbol{p_I}_0, \quad \boldsymbol{\varphi}(0) = \boldsymbol{\varphi}_0,$$

except for the adjoints of fast variables, which are such that

$$\overline{\boldsymbol{p_\varphi}}(0) = \overline{\boldsymbol{p_\varphi}}_0 \qquad \text{and} \qquad \boldsymbol{p_\varphi}(0) = \overline{\boldsymbol{p_\varphi}}_0 + \boldsymbol{\nu_{p_\varphi}}\left(\boldsymbol{I}_0, \boldsymbol{p_I}_0, \boldsymbol{\varphi}_0, \overline{\boldsymbol{p_\varphi}}_0\right),$$

where, following Eq. (17) and assuming that $\boldsymbol{I}_0$ is in a non-resonant zone, $\boldsymbol{\nu_{p_\varphi}}$ is given by

$$\boldsymbol{\nu_{p_\varphi}} = -i \sum_{0<|\boldsymbol{k}|\leq N} \frac{\mathrm{e}^{i\boldsymbol{k}\cdot\overline{\boldsymbol{\varphi}}}}{\boldsymbol{k}\cdot\boldsymbol{\omega}\left(\overline{\boldsymbol{I}}\right)} \left[(-\frac{\partial K}{\partial \boldsymbol{\varphi}}\right]^{(\boldsymbol{k})}. \tag{22}$$

As a result, $\boldsymbol{p_\varphi}$ oscillates with zero mean about $\overline{\boldsymbol{p_\varphi}}$, and the drift between $\boldsymbol{p_I}(t)$ and $\overline{\boldsymbol{p_I}}(t)$ is drastically reduced.

Given the averaged state, Equation (22) establishes a mapping between $\boldsymbol{\varphi}$ and $\boldsymbol{p_\varphi}$. Because $\boldsymbol{\nu_{p_\varphi}}$ has zero mean, there exist $\boldsymbol{\varphi}_0 \in \mathbb{T}^2$ such that

$$\boldsymbol{p_\varphi} - \overline{\boldsymbol{p_\varphi}} = \boldsymbol{\nu_{p_\varphi}}\left(\overline{\boldsymbol{I}}, \overline{\boldsymbol{p_I}}, \boldsymbol{\varphi}_0\right) = \boldsymbol{0}.$$

Similarly, if the initial conditions on the angles are imposed in Problem (1), it is possible to estimate their adjoints, which are ε-small by construction. We note that Eq. (22) simplifies in the single-frequency case. In fact, the system of PDE outlined in Eq. (16) becomes

$$\begin{aligned} &\omega\left(\overline{\boldsymbol{I}}\right)\frac{\partial \boldsymbol{\nu_I}}{\partial p_\varphi} = -\frac{\partial K}{\partial \boldsymbol{\varphi}} + \frac{\partial \overline{K}}{\partial \boldsymbol{\varphi}} + O(\varepsilon), \\ &\int_0^{2\pi} \boldsymbol{\nu}_{p_\varphi}\,\mathrm{d}\varphi = 0, \end{aligned}$$

which solved for $\boldsymbol{\nu}_{p_\varphi}$ yields

$$\boldsymbol{\nu}_{p_\varphi} = -\frac{K\left(\overline{\boldsymbol{I}}, \overline{\boldsymbol{p_I}}, \varphi, 0\right) - \overline{K}\left(\overline{\boldsymbol{I}}, \overline{\boldsymbol{p_I}}\right)}{\omega(\boldsymbol{I})}. \tag{23}$$

Equation (23) states that in the single-frequency case the adjoint of the fast variables should be such that the Hamiltonians of the averaged and original systems match at first order.

4.2 Transformation of the adjoints of slow variables

Changing $\boldsymbol{p_\varphi}$ is mandatory to have consistent trajectories of the averaged and original systems. Transforming the initial value of slow variables and their adjoints is less important, but it can further reduce the drift between these trajectories. Direct application of Eq. (17) is not sufficient to reconstruct short-period variations of $\boldsymbol{p_I}$, as shown in Figure 3. Here, initial conditions of $\boldsymbol{p_I}(t)$ (solid line) are computed by means of Eq. (17). Then, the transformation is evaluated for $t > 0$ to assess if short-period variations are properly modeled. Reconstructed trajectories (dash-dotted lines) of $\boldsymbol{I}$ and $\boldsymbol{p_\varphi}$ well overlap their

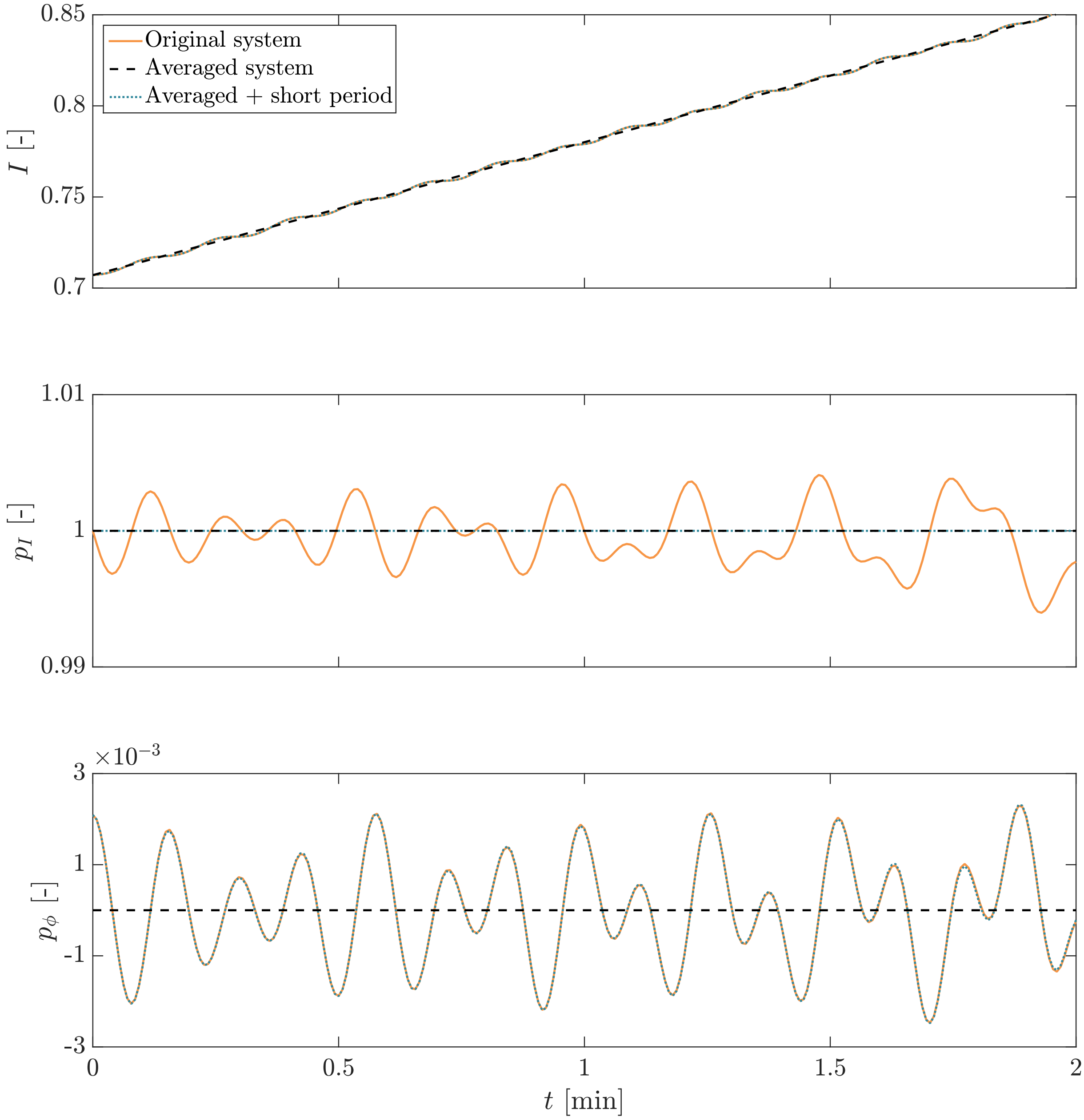


Figure 3: Reconstruction of short-period variatioins by means of the classical transformation.

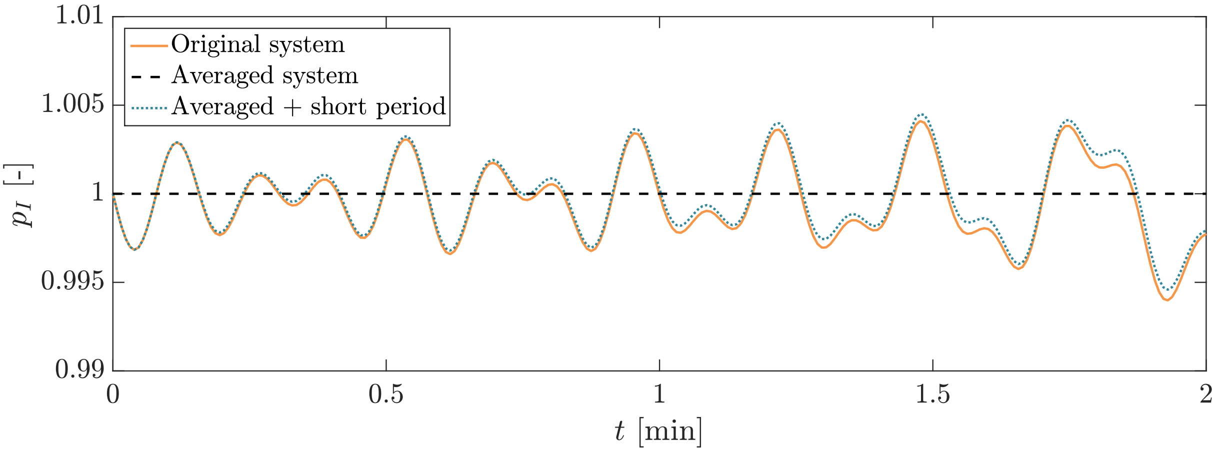


Figure 4: Reconstruction of short-period variatioins of the adjoints of slow variables by means of the proposed transformation.

original counterpart. Conversely, the reconstruction of $\boldsymbol{p_I}$ is wrong (in the very-specific case of the toy problem, $\boldsymbol{\nu}_{p_I} = 0$). Once again, the term $\boldsymbol{p_\varphi}\frac{\partial\,\boldsymbol{\omega}}{\partial\,\boldsymbol{I}}$ in the dynamics of $\boldsymbol{p_I}$ is the responsible. In fact, if short-period variations of $\boldsymbol{p_\varphi}$ are neglected, the Fourier expansion of the right-hand side of is carried out by introducing $\boldsymbol{\varepsilon}$-small errors in the evaluation of the $\boldsymbol{\varepsilon}$-slow dynamics. The transformation of $\boldsymbol{p_I}$ should be carried out by including $\boldsymbol{\nu}_{\boldsymbol{p_\varphi}}$ in the Fourier expansion, namely

$$\boldsymbol{\nu}_{\boldsymbol{p_I}} = -i \sum_{0<|\boldsymbol{k}|\leq N} \frac{\mathrm{e}^{i\boldsymbol{k}\cdot\overline{\boldsymbol{\varphi}}}}{\boldsymbol{k}\cdot\boldsymbol{\omega}\left(\overline{\boldsymbol{I}}\right)} \left[-\left(\overline{\boldsymbol{p_\varphi}} + \boldsymbol{\nu}_{\boldsymbol{p_\varphi}}\right)\frac{\partial\,\boldsymbol{\omega}}{\partial\,\boldsymbol{I}} - \frac{\partial\,K}{\partial\,\boldsymbol{I}}\right]^{(\boldsymbol{k})}.$$

This nested transformation is capable of properly reconstructing short-period variations of the adjoints of slow variables, as shown in Figure 4.

5 Resonance effects

6 Low-thrust orbital transfer

Conclusion

This paper was devoted to the averaging of optimal control systems with two fast variables. After emphasizing how the very-specific nature of the equations of motion of the adjoints of slow variables makes this problem fundamentally different of classical averaging non-controlled fast-oscillating dynamical systems, we showed that existing theorems on averaging are not directly applicable so that trajectories of the original and averaged systems with the same initial conditions generally derive since the beginning of the integration. Hence, we developed a near-identity transformation establishing an equivalence between points of the averaged and original phase spaces can be used to generate consistent boundary conditions of the averaged system. A similar transformation

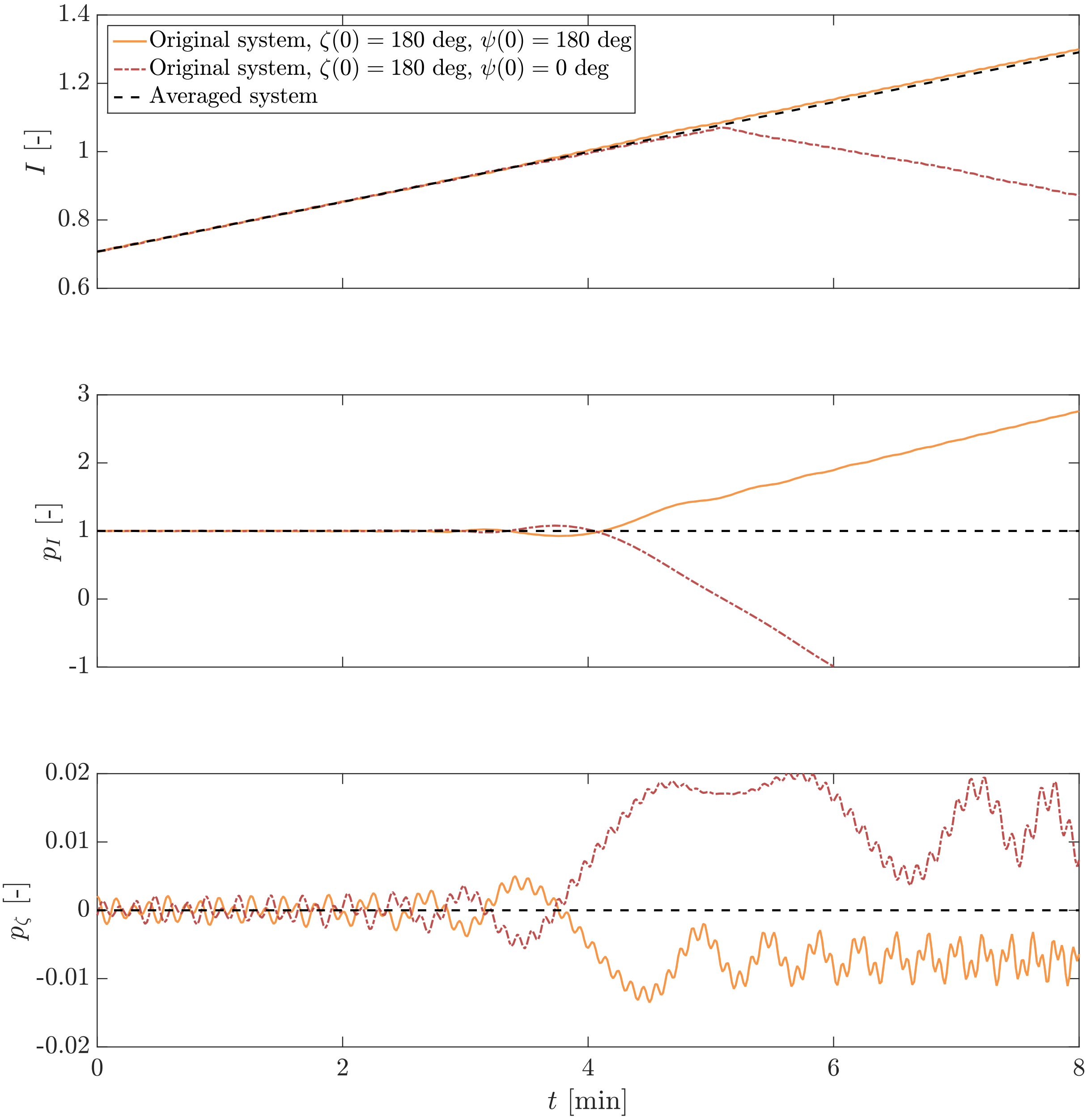
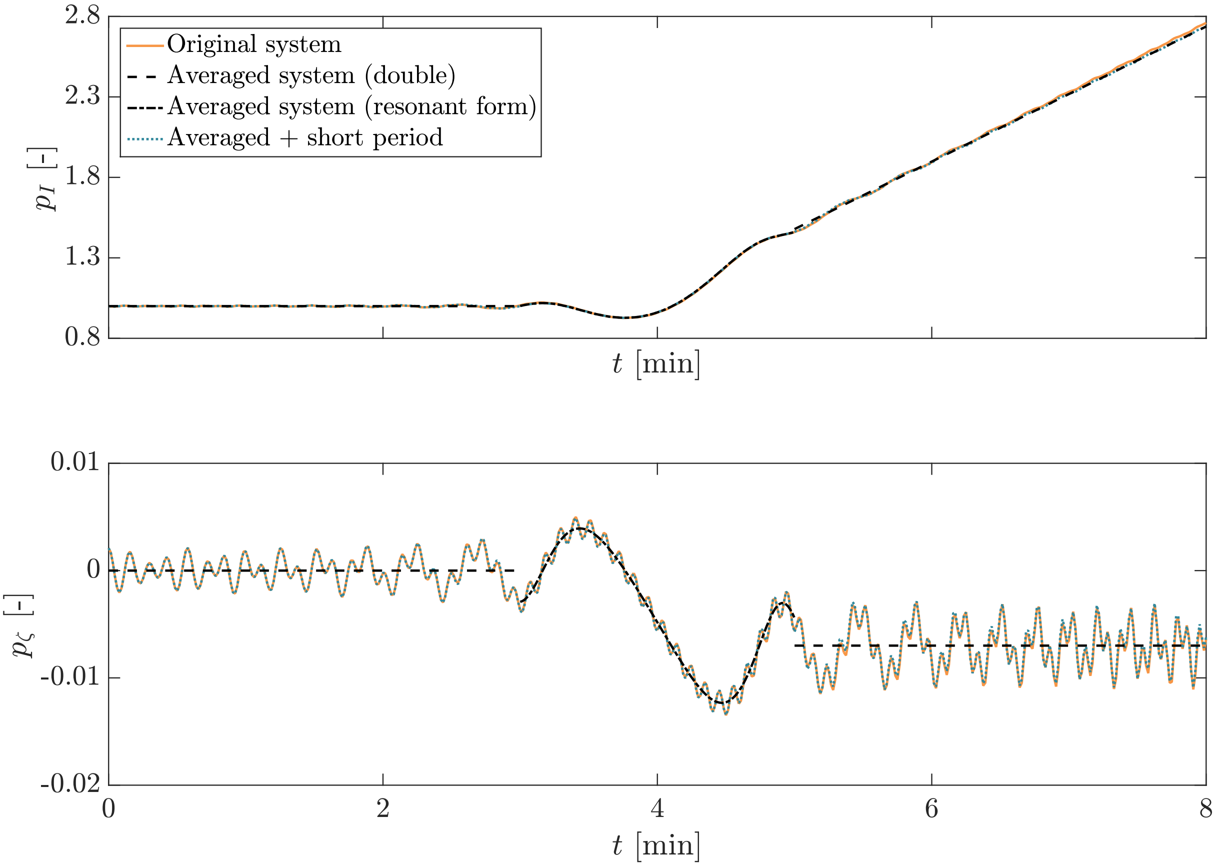


Figure 5: Resonance crossing.



2.8
2.3
1.8
1.3
0.8
p_I [-]
Original system
Averaged system (double)
Averaged system (resonant form)
Averaged + short period
t [min]
0.01
0
-0.01
-0.02
p_ζ [-]
0
2
4
6
8
t [min]

Figure 6: ***I***

was also introduced to tackle the crossing of important resonances, which induce small jumps of the averaged value of the adjoints of fast variables. If neglected, these jumps would generate a steady drift of the adjoints of slow variables that would eventually yield a derive of the entire state. In this context, the transformation served as an interface between the doubly-averaged system and the resonant averaged form, which was used to properly model the aforementioned jumps.

Acknowledgments

This work was partially supported by CNES (contract R-S13/BS-005-012), Thales Alenia Space and Inria (PEPS MSI), PGMO (grant 2016-1753H), and by the French government through the UCA-JEDI Investments in the Future project managed by the National Research Agency (ANR) with the reference number ANR-15-IDEX-01.

References

[1] R. H. Battin, *An Introduction to the Mathematics and Methods of Astrodynamics, Revised Edition*, American Institute of Aeronautics and Astronautics (AIAA), Jan. 1999.

[2] D. A. Danielson, C. P. Sagovac, B. Neta, and L. W. Early, *Semianalytic satellite theory*, tech. rep., Jan. 1995.

[3] P. Lochak and C. Meunier, *Multiphase Averaging for Classical Systems*, Springer New York, 1988.

[4] A. I. Neishtadt, *Averaging, passage through resonances, and capture into resonance in two-frequency systems*, Russian Mathematical Surveys, 69 (2014), pp. 771–843.

[5] J. A. Sanders and F. Verhulst, *Averaging Methods in Nonlinear Dynamical Systems*, Springer New York, 1985.