Available online at www.sciencedirect.com

ScienceDirect

IFAC PapersOnLine 55-16 (2022) 7–12

Considerations on Two-Phase Averaging of Time-Optimal Control Systems

Lamberto Dell'Elce * **Jean-Baptiste Caillau** * **Jean-Baptiste Pomet** *

* *Université Côte d'Azur, Inria, CNRS, LJAD (e-mail: lamberto.dell-elce@inria.fr)*

Abstract: Averaging is a valuable technique to gain understanding of the long-term evolution of dynamical systems characterized by slow dynamics and fast periodic or quasi-periodic dynamics. Averaging the extremal flow of optimal control problems with one fast variable is possible if a special treatment of the adjoint to this fast variable is carried out. The present work extends these results by tackling averaging of time optimal systems with two fast variables, that is considerably more complex because of resonances. No general theory is presented, but rather a thorough treatement of an example, based on numerical experiments. After providing a justification of the possibility to use averaging techniques for this problem "away from resonances" and discussing compatibility conditions between adjoint variables of the original and averaged systems, we analyze numerically the impact of resonance crossings on the dynamics of adjoint variables. Resonant averaged forms are used to model the effect of resonances and cross them without loosing the accuracy of the averaging prediction.

Keywords: fast-oscillating systems, averaging, optimal control, minimum time

1. 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 on the long-term evolution of the flow.

We recently investigated how trajectories of fast-oscillating control system with a single fast variable converge to their averaged counterpart (Dell'Elce et al. (2021)). The present study is motivated by the need for understanding how to generate "consistent" averaged trajectories of minimum time control systems with two fast variables (*i.e.*, characterized by moderate drift with respect to their original counterpart). For this purpose, we leverage on an academic example to provide evidence that existing theorems on double averaging of dynamical systems (namely, the Neishtadt theorem recalled in Section 3) cannot be directly applied to controlled systems. Hence, we develop a proper near-identity transformation (Section 5) of initial adjoint variables that is sufficient to prevent large drift between averaged and original trajectories in non-resonant regions. Finally, we discuss the impact of resonance crossing on the dynamics of adjoint variables and propose a way to predict this behavior via resonant averaged forms. The considerations of this paper may be of use to generate reliable initial guesses for indirect techniques to solve two-point boundary value problems.

⋆ 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.

2. OPTIMAL CONTROL PROBLEM WITH SLOW AND FAST DYNAMICS

Consider the minimum time maneuvering of a dynamical system characterized by fast and slow dynamics, namely

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

Here, the cost function is t_f, 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 $\boldsymbol{f}_j : \mathcal{I}\times\mathbb{T}^2 \to \mathbb{R}^n$ are periodic with respect to $\boldsymbol{\varphi}$ and analytic on the continuation of $\boldsymbol{\varphi}$ in a non-vanishing complex strip. For the sake of conciseness, ε-small terms are not included in the equation of motion of $\boldsymbol{\varphi}$, and ε is not an argument of $\boldsymbol{f}_j$. Also, initial and final conditions $\boldsymbol{I}_0$ and $\boldsymbol{I}_f$ are prescribed on all components of $\boldsymbol{I}$. These are simplifying assumptions, but all outcomes of this note can be extended to cases where initial and final conditions are only partially prescribed, or dependance on ε or additive perturbations of order 1 with respect to ε are added.

10.1016/j.ifacol.2022.08.073

2.1 Necessary conditions for optimality

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

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

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

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

and H_j, for $j = 0, \dots, m$, are defined as $H_j := \boldsymbol{f}_j(\boldsymbol{I}, \boldsymbol{\varphi}) \cdot \boldsymbol{p_I}$.

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

$$\begin{aligned} \frac{\mathrm{d}\boldsymbol{I}}{\mathrm{d}t} &= \varepsilon \frac{\partial K}{\partial \boldsymbol{p_I}}, & \frac{\mathrm{d}\boldsymbol{p_I}}{\mathrm{d}t} &= -\varepsilon \frac{\partial K}{\partial \boldsymbol{I}} - \boldsymbol{p_\varphi} \frac{\partial \boldsymbol{\omega}}{\partial \boldsymbol{I}}, \\ \frac{\mathrm{d}\boldsymbol{\varphi}}{\mathrm{d}t} &= \boldsymbol{\omega}(\boldsymbol{I}), & \frac{\mathrm{d}\boldsymbol{p_\varphi}}{\mathrm{d}t} &= -\varepsilon \frac{\partial K}{\partial \boldsymbol{\varphi}}, \end{aligned} \tag{3}$$

and of the boundary conditions,

$$\boldsymbol{I}(0) = \boldsymbol{I}_0, \quad \boldsymbol{p_\varphi}(0) = \boldsymbol{0}, \quad \boldsymbol{I}(t_f) = \boldsymbol{I}_f, \quad \boldsymbol{p_\varphi}(t_f) = \boldsymbol{0}. \tag{4}$$

Maximizing control is given by

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

In view of the exploitation of shooting-based techniques, triples $(t_f,\ \boldsymbol{p_{I0}},\ \boldsymbol{\varphi}_0)$ are referred to as candidate solutions for Problem (1) if trajectories of System (3) with initial conditions $\boldsymbol{I}(0) = \boldsymbol{I}_0$, $\boldsymbol{p_I}(0) = \boldsymbol{p_{I0}}$, $\boldsymbol{\varphi}(0) = \boldsymbol{\varphi}_0$, and $\boldsymbol{p_{\varphi_0}} = \boldsymbol{0}$ satisfy the endpoint boundary conditions in Eq. (4), i.e., candidate solutions are zeros of the shooting function

$$S\left(t_f, \boldsymbol{p_{I0}}, \boldsymbol{\varphi}_0\right) := \begin{bmatrix} \boldsymbol{I}(t_f|\boldsymbol{I}_0, \boldsymbol{p_{I0}}, \boldsymbol{\varphi}_0, \boldsymbol{0}) - \boldsymbol{I}_f \\ \boldsymbol{p_\varphi}(t_f|\boldsymbol{I}_0, \boldsymbol{p_{I0}}, \boldsymbol{\varphi}_0, \boldsymbol{0}) - \boldsymbol{0} \\ \|\boldsymbol{p_{I0}}\| - 1 \end{bmatrix}, \tag{5}$$

and are such that $H\left(\boldsymbol{I}_0, \boldsymbol{p_{I0}}, \boldsymbol{\varphi}_0, \boldsymbol{p_{\varphi 0}}, \varepsilon\right) > 0$. Here, the notation $\boldsymbol{I}(t_f|\boldsymbol{I}_0, \boldsymbol{p_{I0}}, \boldsymbol{\varphi}_0, \boldsymbol{p_{\varphi 0}})$ denotes the evaluation of $\boldsymbol{I}(t_f)$ obtained by integrating the equations of motion of $\boldsymbol{I}$ with initial conditions $\boldsymbol{I}_0, \boldsymbol{p_{I0}}, \boldsymbol{\varphi}_0, \boldsymbol{p_{\varphi 0}}$.

For a minimum time problem, the adjoint state cannot vanish and one can normalize it assuming it lies on the unit sphere of the cotangent bundle; since $\boldsymbol{p_{\varphi_0}} = 0$ by transversality, one has $\|\boldsymbol{p_{I0}}\| = 1$.

2.2 Toy problem

A simple case study is introduced to streamline the flow of the discussion. Numerical simulations of this example are used to support discussions and mathematical developments of the paper. The dynamical system and optimal control problem is a special case of (1) with a scalar slow variable, I, and two fast variables, ζ and ψ, i.e. $\boldsymbol{I} = I$, $\boldsymbol{\varphi} = (\zeta, \psi)$. The optimal control problem is

$$\begin{aligned} &\min_{\sqrt{u_1^2+u_2^2} \le 1} t_f \quad \text{subject to :} \\ &\frac{\mathrm{d}I}{\mathrm{d}t} = \varepsilon\left[\cos\zeta + \cos\left(\zeta - \psi\right) u_1 + u_2\right], \\ &\frac{\mathrm{d}\zeta}{\mathrm{d}t} = I, \quad \frac{\mathrm{d}\psi}{\mathrm{d}t} = 1, \\ &I(0) = I_0, \quad I(t_f) = I_f. \end{aligned} \tag{6}$$

In the terms of (1), $\boldsymbol{f}_0 = \cos\zeta$, $\boldsymbol{f}_1 = \cos(\zeta - \psi)$, $\boldsymbol{f}_2 = 1$, $\boldsymbol{\omega} = (I, 1)$. The frequency of ψ is constant; this is not a restrictive assumption: as emphasized in (Lochak and Meunier, 1988, Chap. 4), any problem with two frequencies, where at least one frequency does not vanish on the manifold $\mathcal{I}$ can be recast into a form where one frequency is constant.

We note that the dynamical system of Problem 6 could be recast into a single-input control system. All considerations in this paper indeed hold for any problem in the form (1), *e.g.*, low-thrust orbital transfer, but we preferred to use this simple model to streamline the presentation of results as much as possible.

The Hamiltonian associated to Problem (6) is

$$H = Ip_\zeta + p_\psi + \varepsilon\left[p_I \cos\zeta + |p_I|\sqrt{1 + \cos^2\left(\zeta - \psi\right)}\right]. \tag{7}$$

Numerical values used in all simulations are $\varepsilon = 10^{-3}$ and $I_0 = 2^{-0.5}$. Simulation-specific values are listed in the captions of the figures.

3. TWO-PHASE AVERAGING OF FAST-OSCILLATING UNCONTROLLED SYSTEMS

This section is aimed at recalling results on the averaging of the uncontrolled form of the dynamical system introduced in Eq. (1).The motion of $\boldsymbol{I}$ is characterized by a slow trend perturbed by ε-small oscillations. Provided that the two frequencies are not in resonance, 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 system

$$\frac{\mathrm{d}\overline{\boldsymbol{I}}}{\mathrm{d}t} = \varepsilon \overline{\boldsymbol{f}_0}\left(\overline{\boldsymbol{I}}\right), \quad \frac{\mathrm{d}\overline{\boldsymbol{\varphi}}}{\mathrm{d}t} = \boldsymbol{\omega}\left(\overline{\boldsymbol{I}}\right), \tag{8}$$

where averaged vector field, $\overline{\boldsymbol{f}_0} : \mathcal{I} \to \mathbb{R}$ is defined as

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

Averaged fields are independent of fast variable by definition. As such, the motion of $\overline{\boldsymbol{I}}$ is decoupled by $\overline{\boldsymbol{\varphi}}$, which can be eventually evaluated *a posteriori*. 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 introduced by using double averaging inside resonant zones. Section 3.1 recalls the Neishtadt theorem (Neishtadt (2014)) which provides an optimal estimate of the drift between trajectories of System (8) and its original version. Section 3.2 details the near-identity transformation aimed at restoring fast oscillations of an averaged trajectory.

3.1 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] $\overline{\boldsymbol{I}}(t) \in \mathcal{I}$ and

$$\left| \left(\omega_1(\overline{\boldsymbol{I}}) \frac{\partial \omega_2}{\partial \overline{\boldsymbol{I}}} - \omega_2(\overline{\boldsymbol{I}}) \frac{\partial \omega_1}{\partial \overline{\boldsymbol{I}}} \right) \cdot \overline{\boldsymbol{f}_0}\left(\overline{\boldsymbol{I}}\right) \right| > 0, \ \forall t \in [0, \frac{1}{\varepsilon}]. \tag{9}$$

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

$$\sup_{t \in \left[0, \frac{1}{\varepsilon}\right]} \left\| \boldsymbol{I}(t) - \overline{\boldsymbol{I}}(t) \right\| < c_1 \sqrt{\varepsilon} \log\left(\frac{1}{\varepsilon}\right) \tag{10}$$
$$\forall \ (\boldsymbol{I}(0), \boldsymbol{\varphi}(0)) \in \mathcal{V}_1, \ \overline{\boldsymbol{I}}(0) = \boldsymbol{I}(0),$$

and

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

where $\mu(\mathcal{V}_2)$ denotes an ordinary measure on $\mathcal{I} \times \mathbb{T}^2$. Detailed statement and proof of the theorem are available in (Lochak and Meunier, 1988, Chap. 4).

In layman's terms, this theorem states that $\overline{\boldsymbol{I}}(t)$ is a good approximation of $\boldsymbol{I}(t)$, i.e., $\sqrt{\varepsilon}\log(1/\varepsilon) = 0$ 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. (9) 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-long period of time.

3.2 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., (Sanders and Verhulst, 1985, Chap. 7), (Danielson et al., 1995, 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

$$\begin{bmatrix} \hat{\boldsymbol{I}} \\ \hat{\boldsymbol{\varphi}} \end{bmatrix} = \begin{bmatrix} \overline{\boldsymbol{I}} \\ \overline{\boldsymbol{\varphi}} \end{bmatrix} + \varepsilon \boldsymbol{\nu}\left(\overline{\boldsymbol{I}}, \overline{\boldsymbol{\varphi}}\right). \tag{11}$$

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 system, namely:

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

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

$$\begin{aligned} \omega_1\left(\overline{\boldsymbol{I}}\right) \frac{\partial \boldsymbol{\nu}_{\boldsymbol{I}}}{\partial \overline{\varphi}_1} + \omega_2\left(\overline{\boldsymbol{I}}\right) \frac{\partial \boldsymbol{\nu}_{\boldsymbol{I}}}{\partial \overline{\varphi}_2} &= \boldsymbol{f}_0\left(\overline{\boldsymbol{I}}, \overline{\boldsymbol{\varphi}}\right) - \overline{\boldsymbol{f}_0}\left(\overline{\boldsymbol{I}}\right), \\ \omega_1\left(\overline{\boldsymbol{I}}\right) \frac{\partial \boldsymbol{\nu}_{\boldsymbol{\varphi}}}{\partial \overline{\varphi}_1} + \omega_2\left(\overline{\boldsymbol{I}}\right) \frac{\partial \boldsymbol{\nu}_{\boldsymbol{\varphi}}}{\partial \overline{\varphi}_2} &= \left.\frac{\partial \boldsymbol{\omega}}{\partial \boldsymbol{I}}\right|_{\overline{\boldsymbol{I}}} \boldsymbol{\nu}_{\boldsymbol{I}}, \\ \int_{\mathbb{T}^2} \boldsymbol{\nu}\left(\overline{\boldsymbol{I}}, \boldsymbol{\varphi}\right) \mathrm{d}\boldsymbol{\varphi} &= \boldsymbol{0}. \end{aligned} \tag{12}$$

Equation (12) can be solved by first evaluating $\boldsymbol{\nu}_{\boldsymbol{I}}$ and then $\boldsymbol{\nu}_{\boldsymbol{\varphi}}$.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 $|\boldsymbol{k}| = |k_1| + |k_2|$, where k_1 and k_2 are the components of $\boldsymbol{k}$, because of the assumptions on the analyticity of $\boldsymbol{f}_0$. Consequently, there is a constant c_3 such that

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

for $N \ge -c_3 \log \varepsilon$. Assume that the state of the averaged system is outside of any resonant zone of order smaller than $N \ge -c_3 \log \varepsilon$, i.e., there is a constant c_4 such that

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

Hence, formal solution of Eq. (12) is

$$\begin{aligned} \boldsymbol{\nu}_{\boldsymbol{I}}\left(\overline{\boldsymbol{I}}, \overline{\boldsymbol{\varphi}}\right) &= -i \sum_{0 < |\boldsymbol{k}| \le N} \frac{\boldsymbol{f}_0^{(\boldsymbol{k})}\left(\overline{\boldsymbol{I}}\right)}{\boldsymbol{k} \cdot \boldsymbol{\omega}\left(\overline{\boldsymbol{I}}\right)} \mathrm{e}^{i\boldsymbol{k}\cdot\overline{\boldsymbol{\varphi}}}, \\ \boldsymbol{\nu}_{\boldsymbol{\varphi}}\left(\overline{\boldsymbol{I}}, \overline{\boldsymbol{\varphi}}\right) &= -\left.\frac{\partial \boldsymbol{\omega}}{\partial \boldsymbol{I}}\right|_{\overline{\boldsymbol{I}}} \left[\sum_{0 < |\boldsymbol{k}| \le N} \frac{\boldsymbol{f}_0^{(\boldsymbol{k})}\left(\overline{\boldsymbol{I}}\right)}{\left(\boldsymbol{k} \cdot \boldsymbol{\omega}\left(\overline{\boldsymbol{I}}\right)\right)^2} \mathrm{e}^{i\boldsymbol{k}\cdot\overline{\boldsymbol{\varphi}}}\right]. \end{aligned} \tag{13}$$

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 (6). Because the transformation amends ε-small correction to the averaged state, the inverse transformation of Eq. (11) can be approximated by

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

4. THE AVERAGED CONTROL SYSTEM

Applying averaging theory to the extremal flow detailed in Eq. (3) is questionable because the structure of this vector field differs from the conventional (uncontrolled) fast-oscillating system. Specifically, the equation of motion of $\boldsymbol{p}_{\boldsymbol{I}}$, includes the term $\boldsymbol{p}_{\boldsymbol{\varphi}} \dfrac{\partial \boldsymbol{\omega}}{\partial \boldsymbol{I}}$ that may possibly be of order larger than ε. Hence, adjoints to slow variables are not necessary slow themselves. This section justifies the application of averaging theory to System (3) by showing that adjoints to fast variables are systematically ε-small for any extremal trajectory with free phases, and, as such, $\dfrac{\mathrm{d}\boldsymbol{p}_{\boldsymbol{I}}}{\mathrm{d}t} = \mathcal{O}(\varepsilon)$ when restrained to these trajectories.

Consider the canonical change of variables $[\boldsymbol{I}, \boldsymbol{p}_{\boldsymbol{I}}, \boldsymbol{\varphi}, \boldsymbol{p}_{\boldsymbol{\varphi}}] \to [\boldsymbol{J}, \boldsymbol{p}_{\boldsymbol{J}}, \boldsymbol{\psi}, \boldsymbol{p}_{\boldsymbol{\psi}}]$ such that $\boldsymbol{J} = \boldsymbol{I}$ and $\boldsymbol{\psi} = \boldsymbol{\Omega}(\boldsymbol{I})\, \boldsymbol{\varphi}$, where $\boldsymbol{\Omega} : \mathcal{I} \to \mathbb{R}^{2\times 2}$ is defined as

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

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

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

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

Boundary conditions on the adjoints to 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. (5), i.e., $\|\boldsymbol{p_I}_0\| = 1$, yields the Hamiltonian level

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

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

$$p_{\psi_1} = \varepsilon\frac{h - \tilde{K}\left(\boldsymbol{J}, \boldsymbol{p_J}, \boldsymbol{\psi}\right)}{||\boldsymbol{\omega}(\boldsymbol{J})||}. \tag{14}$$

Equation (14) reveals that $p_{\psi_1} = \mathcal{O}(\varepsilon)$ when evaluated on a candidate optimal trajectory. As a consequence, $\boldsymbol{p_J}$ exhibit ε-slow dynamics, i.e.,

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

which justifies the averaging of the extremal flow.

Denote by $\overline{K}$ the averaged functional

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

Averaging the extremal flow of Eq. (3) yields

$$\frac{\mathrm{d}\,\overline{\boldsymbol{I}}}{\mathrm{d}\,t} = \varepsilon\frac{\partial\,\overline{K}}{\partial\,\overline{\boldsymbol{p_I}}}, \quad \frac{\mathrm{d}\,\overline{\boldsymbol{p_I}}}{\mathrm{d}\,t} = -\varepsilon\frac{\partial\,\overline{K}}{\partial\,\overline{\boldsymbol{I}}} - \overline{\boldsymbol{p_\varphi}}\ \frac{\partial\,\boldsymbol{\omega}}{\partial\,\overline{\boldsymbol{I}}},$$
$$\frac{\mathrm{d}\,\boldsymbol{\varphi}}{\mathrm{d}\,t} = \boldsymbol{\omega}\left(\overline{\boldsymbol{I}}\right), \quad \frac{\mathrm{d}\,\overline{\boldsymbol{p_\varphi}}}{\mathrm{d}\,t} = \boldsymbol{0}.$$

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

4.1 *Averaging the toy problem*

Averaging Eq. (7) yields

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

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

$$\begin{aligned} &\frac{\mathrm{d}\,\overline{I}}{\mathrm{d}\,t} \approx 1.216\,\varepsilon\frac{\overline{p_I}}{|\overline{p_I}|}, \qquad \frac{\mathrm{d}\,\overline{p_I}}{\mathrm{d}\,t} = -\overline{p_\zeta}, \\ &\frac{\mathrm{d}\,\overline{\zeta}}{\mathrm{d}\,t} = \overline{I}, \quad \frac{\mathrm{d}\,\overline{p_\zeta}}{\mathrm{d}\,t} = 0, \quad \frac{\mathrm{d}\,\overline{\psi}}{\mathrm{d}\,t} = 1, \quad \frac{\mathrm{d}\,\overline{p_\psi}}{\mathrm{d}\,t} = 0, \end{aligned} \tag{15}$$

By definition, the integration of the slowly-changing variables of System (15) is independent of $\overline{\zeta}$ and $\overline{\psi}$. Closed-form solution of the slow averaged flow is

$$\overline{I} \approx \overline{I}_0 + 1.216\,\varepsilon\frac{\overline{p_I}}{|\overline{p_I}|}\,\gamma(t),$$
$$\overline{p_I} = \overline{p_I}_0 - \overline{p_\zeta}_0\,t, \quad \overline{p_\zeta} = \overline{p_\zeta}_0, \quad \overline{p_\psi} = \overline{p_\psi}_0.$$

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

$$\gamma(t) = \begin{cases} t & \text{if } \ \overline{p_I}_0\,\overline{p_\zeta}_0 \leq 0 \ \text{ or } \ t \leq \dfrac{\overline{p_I}_0}{\overline{p_\zeta}_0} \\ 2\,\dfrac{\overline{p_I}_0}{\overline{p_\zeta}_0} - t & \text{otherwise} \end{cases} \tag{16}$$

The slow variable $\overline{I}(t)$ coincides with the frequency ratio of the averaged system, and it evolves as a continuous 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 entire trajectory.

5. NEAR-IDENTITY TRANSFORMATION FOR CONTROLLED SYSTEMS

Changing initial conditions of averaged trajectories by means of the transformation discussed in Section 3.2 reduces 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 for classical fast-oscillating systems when compared to the estimate provided by the Neishtadt theorem, which considers the same initial conditions for the two trajectories. Conversely, the transformation of the initial variables plays a key role for the optimal control problem. To support this claim, consider Problem (6) 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. (16) 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, as required by Neishtadt's theorem. However, Figure 1 shows that p_I and $\overline{p_I}$ exhibit a steady drift that largely exceeds the expected "small" error quantified in Eq. (10) when Systems (**??**) and (15) are integrated with the same initial conditions. In addition, comparing the red and orange curves of Figure 1 reveals that trajectories of the original system strongly depend on the values of initial phases. Section 5.1 shows that transforming the adjoints to fast variables is sufficient to drastically reduce the drift of $\boldsymbol{p_I}$. Section 5.2 describes a new transformation for properly modeling short-periodic variations of the system at hand.

5.1 *The fundamental role of the adjoints of fast variables*

The trigger 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 on $\overline{\boldsymbol{p_I}}$, which is of the same order of magnitude, i.e.,

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

In turn, an ε-small error on $\overline{\boldsymbol{p_\varphi}}$ induces an error on the time derivative of $\overline{\boldsymbol{p_I}}$ that is comparable to its slow motion. Transforming initial adjoints to fast variables by means of Eq. (11) is sufficient to greatly mitigate this problem, as

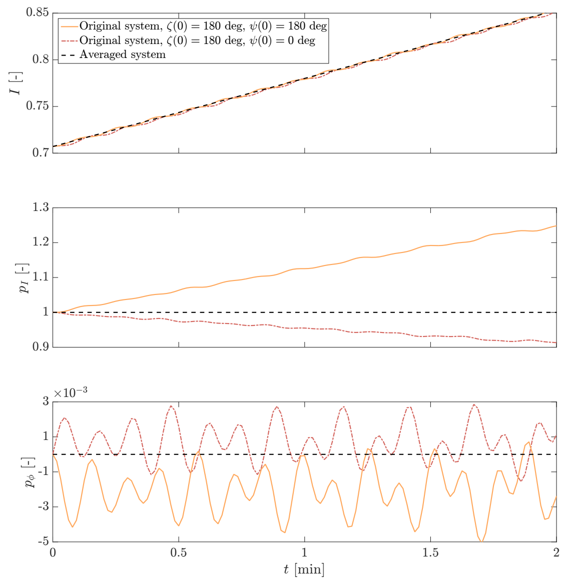

Fig. 1. Numerical integration of the toy problem. Trajectories of the original and averaged system are emanated from the same point of the phase space. Initial adjoints are $p_I(0) = 1$ and $p_\psi(0) = p_\varsigma(0) = 0$.

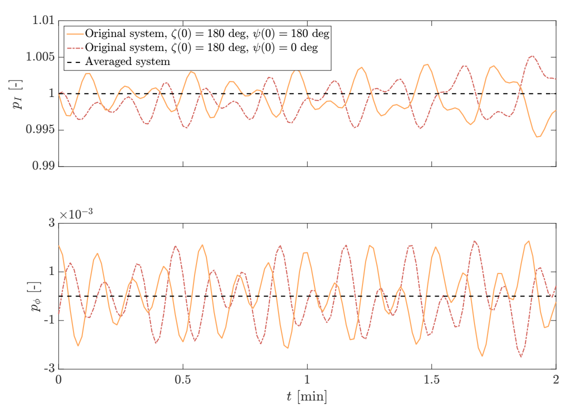

Fig. 2. Numerical integration of the toy problem. Here, initial adjoints to fast variables are transformed by means of Eq. (13).

shown in Figure 2. Here, initial conditions of averaged and original initial value problems (IVP) are mostly the same (specifically $\boldsymbol{I}(0) = \overline{\boldsymbol{I}}(0) = \boldsymbol{I}_0$ and $\boldsymbol{p_I}(0) = \overline{\boldsymbol{p_I}}(0) = \boldsymbol{p_I}_0$, but the adjoints to fast variables are transformed such that $\overline{\boldsymbol{p_\varphi}}(0) = \overline{\boldsymbol{p_\varphi}}_0$ and $\boldsymbol{p_\varphi}(0) = \overline{\boldsymbol{p_\varphi}}_0 + \boldsymbol{\nu}_{\boldsymbol{p_\varphi}}\left(\boldsymbol{I}_0, \boldsymbol{p_I}_0, \boldsymbol{\varphi}_0, \overline{\boldsymbol{p_\varphi}}_0\right)$, where, following Eq. (13) and assuming that $\boldsymbol{I}_0$ is in a non-resonant zone, $\boldsymbol{\nu}_{\boldsymbol{p_\varphi}}$ is given by

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

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, Eq. (17) establishes a mapping between $\boldsymbol{\varphi}$ and $\boldsymbol{p_\varphi}$. Because $\boldsymbol{\nu}_{\boldsymbol{p_\varphi}}$ has zero mean, there exist $\boldsymbol{\varphi}_0 \in \mathbb{T}^2$ such that $\hat{\boldsymbol{p}}_{\boldsymbol{\varphi}} = \overline{\boldsymbol{p_\varphi}} + \boldsymbol{\nu}_{\boldsymbol{p_\varphi}}\left(\overline{\boldsymbol{I}}, \overline{\boldsymbol{p_I}}, \boldsymbol{\varphi}_0\right) = \boldsymbol{0}$.

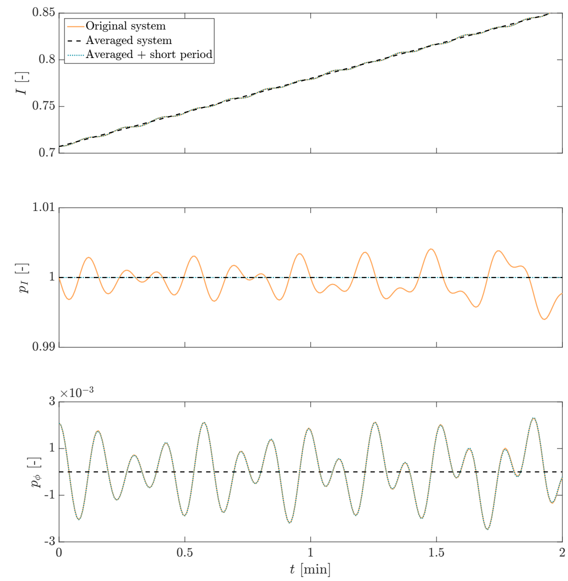

Fig. 3. Reconstruction of short-period variatioins by means of the classical transformation.

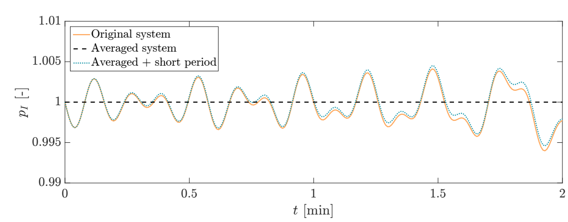

Fig. 4. Reconstruction of short-period variatioins of the adjoints to slow variables by means of the proposed transformation.

5.2 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 critical, but it can further reduce the drift. Direct application of Eq. (13) 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. (13). 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 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 $\boldsymbol{p_\varphi}\partial\boldsymbol{\omega}/\partial\boldsymbol{I}$ in the dynamics of $\boldsymbol{p_I}$ is responsible of the mismatch. In fact, if short-period variations of $\boldsymbol{p_\varphi}$ are neglected, the Fourier expansion of $\mathrm{d}\boldsymbol{p_I}/\mathrm{d}t$ is carried out by introducing ε-small errors in the evaluation of the ε-slow dynamics. The transformation of $\boldsymbol{p_I}$ should be carried out by including $\boldsymbol{\nu}_{\boldsymbol{p_\varphi}}$ in the Fourier expansion, namely

$$\boldsymbol{\nu}_{\boldsymbol{p_I}} = \sum_{0<|\boldsymbol{k}|\leq N} \left\{ \frac{1}{\left(\boldsymbol{k}\cdot\boldsymbol{\omega}\left(\overline{\boldsymbol{I}}\right)\right)^2} \left[\frac{\partial K}{\partial \boldsymbol{\varphi}}\right]^{(\boldsymbol{k})} \frac{\partial \boldsymbol{\omega}}{\partial \boldsymbol{I}} + \frac{i}{\boldsymbol{k}\cdot\boldsymbol{\omega}\left(\overline{\boldsymbol{I}}\right)} \left[\frac{\partial K}{\partial \boldsymbol{I}}\right]^{(\boldsymbol{k})} \right\} \mathrm{e}^{i\boldsymbol{k}\cdot\overline{\boldsymbol{\varphi}}}.$$

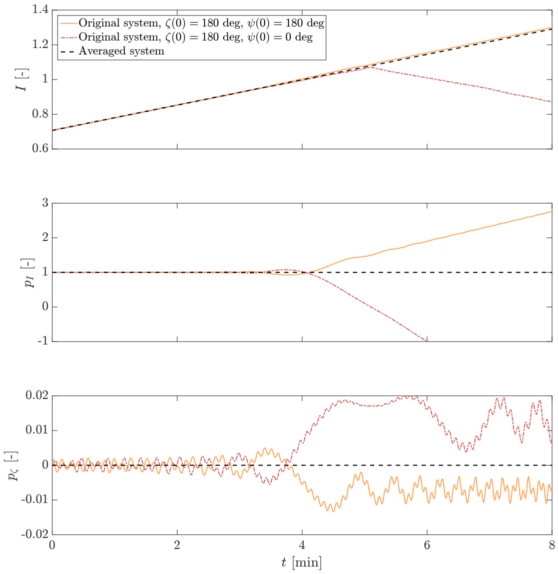

Fig. 5. Example of resonance crossing (1:1). Double averaging is not sufficient to capture the evolution of adjoint variables inside the resonant zone. Hence, large drift occurs after crossing the resonance.

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

6. RESONANCE CROSSING

Approaching a resonance may entail two main effects. The first one is referred to as capture into resonance. It occurs when the state of the system remains in the neighborhood of the resonance for a long amount of time; this defeats any estimate like Eq. (10) of the drift, but, in the framework of Neishtadt theorem, it is proved that the constraint outlined in Eq. (9) prevents such a phenomenon from happening. The second effect consists of the scattering between trajectories of the original and averaged systems due to the rapid crossing of a resonant zone. The drift accumulated by crossing several resonances is somehow small for classical fast oscillating systems (as quantified by Eq. (10)). Conversely, resonance crossing may be detrimental when dealing with trajectories of System (3) because of the very specific form of the equations of motion.

Figure 5 shows that trajectories of the toy problem abruptly drift after the one-to-one resonance is crossed. The mechanism yielding this drift is analogous to the one discussed in Section 5.1: resonance crossing induces a small variation of the averaged value of $\boldsymbol{p_\varphi}$ which is not modeled by the doubly-averaged system. For this purpose, resonant averaged forms can be used to properly model the motion of $\overline{\boldsymbol{p_\varphi}}$ inside important resonances. Specifically, assuming that the resonance identified with the index $\boldsymbol{k}$ is crossed, namely $\overline{\boldsymbol{I}}$ is such that $|\boldsymbol{\omega}(\overline{\boldsymbol{I}})\cdot\boldsymbol{k}| \leq c\sqrt{\varepsilon}$, a change of variables is performed such that $\boldsymbol{\varphi}$ is decomposed into a slow and a fast components, β and α, respectively, namely $\beta = \boldsymbol{k}\cdot\overline{\boldsymbol{\varphi}}$ and $\alpha = \boldsymbol{k}^{\perp}\cdot\overline{\boldsymbol{\varphi}}$, The Hamiltonian is then averaged with respect to α on the period $2\,\pi\,k_1\,k_2$.

A transformation analogous to the one discussed in Section 3.2 can be established to transform the state vector at the interface between the doubly-averaged system and the resonant form. Figure 6 depicts the averaged trajectory obtained by modeling the motion inside the one-to-one resonance crossing of the toy problem.

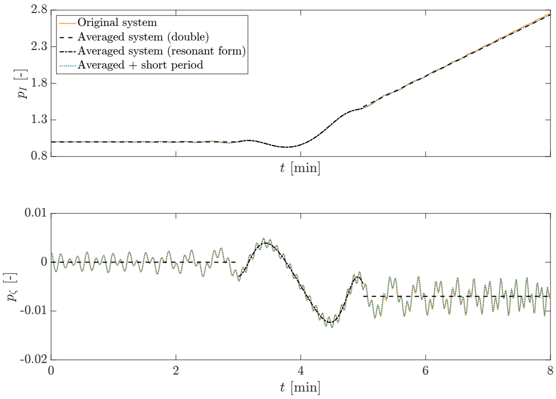

Fig. 6. Example of resonance crossing (1:1). The resonant form efficiently captures the evolution of adjoint variables inside the resonant zone. Double averaging is used outside resonant zones.

Two open questions still need to be addressed before the proposed methodology can be used to automatically generate averaged trajectories. First, important resonances need to be identified by inspection of the averaged trajectory. Second, a quantitative assessment of the width of resonant zones is required.

7. CONCLUSION

This paper is devoted to the averaging of optimal control systems with two fast variables. We showed that existing theorems on multi-frequency averaging are not directly applicable to this problem as trajectories of the original and averaged systems with the same initial conditions quickly drift apart. Hence, we developed a near-identity transformation establishing an equivalence between points of the averaged and original phase spaces, that can be used to generate consistent boundary conditions of the averaged system. A similar transformation, serving as interface between doubly and resonant averaged systems at switching instants, was also introduced to tackle the crossing of important resonances. We believe that this discussion could shed light on some difficulties of the problem.

REFERENCES

Danielson, D.A., Sagovac, C.P., Neta, B., and Early, L.W. (1995). Semianalytic satellite theory. Technical report, Naval Postgraduate School.

Dell'Elce, L., Caillau, J.B., and Pomet, J.B. (2021). On the convergence of time-optimal maneuvers of fast-oscillating control systems. In *2021 European Control Conference (ECC)*. IEEE.

Lochak, P. and Meunier, C. (1988). *Multiphase Averaging for Classical Systems*. Springer New York.

Neishtadt, A.I. (2014). Averaging, passage through resonances, and capture into resonance in two-frequency systems. *Russian Mathematical Surveys*, 69(5), 771–843.

Sanders, J.A. and Verhulst, F. (1985). *Averaging Methods in Nonlinear Dynamical Systems*. Springer New York.