# Periodic averaging with a second order integral error

## Thierry Dargent, Florentina Nicolau, Jean-Baptiste Pomet

**HAL Id: hal-01351613**

**https://hal.archives-ouvertes.fr/hal-01351613v2**

Submitted on 22 Nov 2016

# Periodic averaging
# with a second order integral error [⋆]

**Thierry Dargent** [*] **Florentina Nicolau** [**] **Jean-Baptiste Pomet** [***]

*\* Thales Alenia Space, 5 Alle des Gabians, 06150 Cannes, France,
thierry.dargent@thalesaleniaspace.com
\*\* Quartz Laboratory, Ensea, 6 Avenue du Ponceau, 95014 Cergy-Pontoise,
France, florentina.nicolau@ensea.fr
\*\*\* Inria Sophia Antipolis Mditerrane, 2004 Route des Lucioles, 06902
Valbonne, France, jean-baptiste.pomet@inria.fr*

**Abstract:** In this paper, we study a variation of second order periodic averaging that allows an asymptotic reconstruction of the fast variable ($\varphi$). Providing a good estimation of $\varphi$ has clear practical interest, for instance, if we use averaging in control problems where the values of $\varphi$ at initial and final time are prescribed. Our goal is to present a simpler method (than the classical high order averaging results) to obtain a second order estimate (with respect to the small parameter $\epsilon$) based on changing the initial condition of the averaged system in order to approach an average of the solutions. Moreover, our approach preserves as well the first order estimate of the exact solutions. The idea on which our results are based was already present in a publication by the first author, in the context of solving an optimal control problem with averaging techniques and application to low thrust orbit transfer. Namely, adding well suited second order terms to the averaged system and properly choosing the initial condition for the averaged system yield an error of order $\epsilon^2$ "in the mean" (instead of point-wise) on the slow variables and of order $\epsilon$ for the fast variable.

## 1. INTRODUCTION

In this paper, we study periodic averaging for the class of systems of the form

$$(S_\epsilon): \quad \begin{aligned} \dot{I} &= \epsilon f(I, \varphi) \\ \dot{\varphi} &= \omega(I) \end{aligned}$$

where $\epsilon$ is a small parameter, the slow variable $I$ lives in some open subset of $\mathbb{R}^n$ and the fast variable $\varphi$ is an angle. There may also be terms of order higher than 1 with respect to $\epsilon$ in $\dot{I}$ and higher than zero in $\dot{\varphi}$, this is not considered here. The systems of the form $(S_\epsilon)$ are called "one-frequency integrable" systems in Lochak and Meunier (1988). Indeed, $(S_\epsilon)$ is a perturbation of the integrable Hamiltonian system $\dot{I} = 0$, $\dot{\varphi} = \omega(I)$ called "one-frequency" because $\varphi$ has dimension 1, while $\varphi$ usually has the same dimension as $I$ in the reduction of a generic integrable Hamiltonian system to a system represented in action-angle variables (Liouville's theorem, see, e.g., Arnold (1978)).

The (first order) averaging principle for the system $(S_\epsilon)$ consists in eliminating the fast variable $\varphi$ and determining the approximate behavior of the slow variables $I$ by constructing macroscopic evolution equations (called the averaged system) which handle only certain average characteristics of the small-scale motion, and which give a good approximation to the true evolution of the slow variables on a certain time interval, see, for instance, Arnold (1978); Bogoliubov and Mitropolsky

(1961); Lochak and Meunier (1988); Sanders et al. (2007). These methods yield approximation results that are by nature asymptotic (as $\epsilon$ tends towards 0): typically, these results provide error estimates of order $\epsilon$ for the slow variables on an interval of time whose length has order $1/\epsilon$. Although the results are asymptotic, averaging methods are very useful in practice when $\epsilon$ is finite "reasonably small". In the more recent paper Morosi and Pizzocchero (2006), the first-order approximation was replaced with a fully quantitative estimate for a time interval of order $1/\epsilon$.

Averaging methods have attracted a lot of attention because of their applications to many interesting physical problems (work and heat in thermodynamics, hydrodynamical and molecular dynamical quantities in fluid mechanics, electronics, celestial mechanics, quantum mechanics, etc.). They are also used in optimal control, see the pioneering work Edelbaum (1965) and the more recent Chaplais (1987); Geffroy (1997); Dargent (2015).

The results of the present paper stem out of ideas, remarks and numerical experiments of the first author, published in Dargent (2015), and based on the idea that the trajectory of the averaged system should approach the (oscillating) trajectory of the original system $(S_\epsilon)$ but also, and at a higher order, its average on the approximate duration of one oscillation. However, the paper Dargent (2015) deals with averaging applied to a two point boundary value problem arising from optimal control and the Pontryagin maximum principle rather than a Cauchy problem. Here, we develop the theory for Cauchy problems because it is simpler and interesting in its own

right. When we talk about initial conditions, it should be translated into initial and final conditions in the context of Dargent (2015), with the precaution that, of course, they do not define a unique solution there.

The paper is organized as follows. We recall in Section 2 the classical first order periodic averaging principle and discuss the importance of the fast variable reconstruction, higher order averaging and the motivation of our approach. Section 3 constructs a corrected averaged system and states that, with the right initialization, this corrected averaged system provides an average error over one oscillation that is of order $\epsilon^2$ and also reconstructs the fast variable, with an error of order $\epsilon$. Section 4 is devoted to the proofs of our results.

## 2. PROBLEM STATEMENT AND MOTIVATION

### 2.1 Notations and Assumptions

As stated in the introduction, we consider systems of the form

$$(S_\epsilon): \quad \begin{array}{ll} \dot{I} = \epsilon f(I, \varphi), & I \in O \subset \mathbb{R}^n \\ \dot{\varphi} = \omega(I), & \varphi \in S^1 \end{array} \quad (1)$$

where $f$ is defined on $O \times S^1$, $\omega$ on $O$, with $O$ an open subset of $\mathbb{R}^n$ and $S^1$ the circle $\mathbb{R}/2\pi\mathbb{Z}$. Notice that $\varphi \in S^1$ amounts to saying that $f$ and $g$ are $2\pi$-periodic with respect to $\varphi$.

We make the following assumptions on $O$, $f$ and $\omega$:

(A1)  $f$ and $\omega$ are $\mathcal{C}^1$-smooth in all their arguments, and $\omega(I) > \omega_{min}$ for some positive $\omega_{min}$ and all $I \in O$;

(A2)  $f$, $\frac{\partial f}{\partial I}$, $\omega$, $\frac{\partial \omega}{\partial I}$ are uniformly bounded on $O \times S^1$;

(A3)  $\frac{\partial f}{\partial I}$ has a global Lipschitz constant on $O$.

We refer to $I$ as the slow variable, and to $\varphi$ as the fast variable, that has, at each time, an approximate period

$$T(I) = \frac{2\pi}{\omega(I)}. \quad (2)$$

Namely, if $(I_\epsilon(t), \varphi_\epsilon(t))$ is a solution of $(S_\epsilon)$, then $\varphi_\epsilon(t + T(I_\epsilon(t)))$ is equal to $\varphi_\epsilon(t) + 2\pi$ up to a term of order $\epsilon$.

The aim of averaging is to approach in some way the slow variable evolution in $(S_\epsilon)$ by the solutions of an averaged system whose right-hand side depends only on slow variables; it uses the average of $f$ with respect to $\varphi$:

$$\overline{f}(I) = \frac{1}{2\pi} \int_0^{2\pi} f(I, \xi) d\xi. \quad (3)$$

All results concern solutions starting close to a particular initial condition $I_0$ in $O$ on a horizon defined by some fixed $T > 0$. We assume the following on $I_0$ and $T$ (it is always true for $T$ small enough):

(ICA)  The solution $t \mapsto x(t)$ of $\dot{x} = \overline{f}(x)$, with $x(0) = I_0$, remains in $O$ for $0 \leq t \leq T < +\infty$.

Note that assumptions (A1)-(A3) are regularity assumptions, while (ICA) is a non explosion assumption.

### 2.2 First order averaging

The following classical result can be found for instance in Sanders et al. (2007), Lochak and Meunier (1988) or Arnold (1978). It states that the solutions of $(S_\epsilon)$ and those of the averaged system $(AvS_\epsilon)$ remain close (up to order $\epsilon$) for a time interval of order $1/\epsilon$:

*Theorem 1. Let assumptions (A1)-(A2) and (ICA) be satisfied. Let $k$ be some positive constant. Consider a family (indexed by $\epsilon > 0$) of solutions $(I_\epsilon(t), \varphi_\epsilon(t))$ of $(S_\epsilon)$ and a family of solutions $J_\epsilon(t)$ of the averaged system*

$$(AvS_\epsilon): \quad \dot{J} = \epsilon \overline{f}(J) \quad (4)$$

*with $\overline{f}$ given by (3), such that*

$$||I_\epsilon(0) - I_0|| \leq k\epsilon \text{ and } ||J_\epsilon(0) - I_0|| \leq k\epsilon$$

*for all $\epsilon$ small enough. Then there exist two positive constants $\epsilon_0$ and $c$ such that for all $0 \leq \epsilon \leq \epsilon_0$, the solution of the system $(S_\epsilon)$ exists for $0 \leq t \leq \frac{T}{\epsilon}$ and*

$$||I_\epsilon(t) - J_\epsilon(t)|| \leq c\epsilon,$$

*where $c$ depends on $k$ and on the bounds in assumptions (A1)-(A2) only.*

This is usually stated with $I_\epsilon(0) = J_\epsilon(0) = I_0$; allowing a $\epsilon$-discrepancy in the initial conditions does not change the proof. The main idea of the proof is to reduce the original system to the averaged one up to the order $\epsilon^2$ by a near identity change of variables and then apply the Gronwall lemma in order to bound the distance between the new variables and $J_\epsilon$. One shorter and more computational proof, proposed in Sanders et al. (2007), relies on an inequality due to Besjes (1969).

Observe that assumption (A3) is not needed in the above theorem, it will be used later.

### 2.3 Problem statement

The motivation for this paper comes from optimal control applied to low thrust orbital transfer, see references in the introduction. In that case, the slow variables $I$ are the first integrals of the motion with zero control, that describe an ellipse in $\mathbb{R}^3$ with a fixed focus and the "fast" angle $\varphi$ defines the position on the corresponding elliptic orbit. Averaging is relevant because the small control may appear as a perturbation whose effect is averaged with respect to the fast angle. Using averaging in the style of Theorem 1, adapted to optimal control, one may study transfer from one elliptic orbit to another, but the estimate of the position on the target orbit upon arrival is very poor, and a technique to obtain it without re-computing the non averaged problem is very relevant. Another motivation to reconstruct this fast variable is explained below. In optimal control, averaging is in fact applied to a Hamiltonian system obtained through the Pontryagin Maximum principle which also allows to compute the control to be applied at each instant. Although first order averaging is useful to study this system, it does not give a good estimation for the control of the original system, because this control *is* oscillating and depends on the fast variables as well as the slow. Based on Theorem 1, the simplest way to obtain an estimate of $\varphi$, denoted by $\psi$, is to use an equation of the form

$$\dot{\psi} = \omega(J) \quad (5)$$

or some variants. One easily proves that the gap between $\varphi$ and $\psi$ is bounded, but not more.

For the problem of low thrust orbital transfer, it is shown numerically in Dargent (2015) that, when reconstructing the control based on the above estimation $\psi$ of $\varphi$, and recomputing the trajectory for the "real" system, the obtained approximation is not reasonable. It is also noticed that $\varphi$ and $\psi$ may differ by up to 90 degrees, hence totally loosing synchronization.

Reconstruction of the fast variable is usually not tackled in the literature, at least to the best of our knowledge. However, a solution $\psi$ of (5) seems a good candidate. Since $J$ is $\epsilon$-close to $I$, integrating on a time interval whose length is of order $1/\epsilon$ indicates that the error between $\psi$ and $\varphi$ will be bounded and does not necessarily tend to zero with $\epsilon$. If $J$ were $\epsilon^2$-close to $I$, it would be reasonable to say that $\psi$ is $\epsilon$-close to $\varphi$.

In fact, higher order averaging results exist, see for instance Lochak and Meunier (1988); Sanders et al. (2007). The slow component $I$ of the solutions $(I, \varphi)$ of the oscillating system may be approached at arbitrary order with respect to $\epsilon$. More precisely, $k^{\text{th}}$ order averaging, $k > 1$, yields an approximation of order $\epsilon^k$ on a time interval of length $1/\epsilon$ (or an approximation of order $\epsilon^{k-j}$ on a time interval of length $1/\epsilon^{1+j}$, $0 < j < k$). The $k$-th order averaged equation is in the style of (4) with a right-hand side that still depends on $J$ only but has extra terms of order $2, 3, \ldots k$ with respect to $\epsilon$; since $I$ displays "fast" oscillations with an amplitude of order $\epsilon$, it *cannot* be approached at an order larger than 1 with respect to $\epsilon$ by a solution of a differential equation whose right-hand side depends on slow variables only, indeed it is not this $J$ that approaches $I$ in high order averaging, but rather the image of $J$ by a transformation that is close to identity but does have fast oscillations. Although it is not mentioned in the literature, we believe that, as explained above, since second order averaging yields an $\epsilon^2$-estimate of the slow variable, it would yield an $\epsilon$-estimate of the fast variable.

We did not try this because the construction of the oscillating transformation has the drawback that differential equations with fast-varying right-hand side have to be solved and anyway an $\epsilon^2$-estimate of the slow variable is ot needed, or rather it is only needed in an integral sense. The solution proposed in Dargent (2015) is based on the idea that the trajectory $J_\epsilon(t)$ of the averaged system should approach the (oscillating) trajectory $I_\epsilon(t)$ of the original system $(S_\epsilon)$ at order $\epsilon$ and that it should approach at a higher order a filtered version of $I_\epsilon$ that does not display oscillations with an $\epsilon$-amplitude, here we use the "moving average" on the approximate period of the fast variable, see Section 3. This leads to choosing an initial condition $J_\epsilon(0)$ for the averaged system that is not the same as $I_\epsilon(0)$, but rather has an offset of order $\epsilon$ depending on $I_0$ and $\varphi_0$ such that $J_\epsilon(0)$ is $\epsilon^2$ close to the initial value of the above mentioned filtered version of $I_\epsilon$ (this translates formally the fact that, for this precise offset, $J_\epsilon(t)$ seems to lie "in the middle of" the oscillating $I_\epsilon(t)$): the estimate $\|I_\epsilon(t) - J_\epsilon(t)\| \leq c\epsilon$ given by Theorem 1 is then preserved, and it is evidenced numerically in Dargent (2015) that, on the one hand, this particular choice of initial condition results in a solution of the averaged system that approaches the average of the real

solution with a much better accuracy than the one for approaching the solution itself and that, on the other hand, one also obtains an accurate reconstruction of the fast variable that was previously impossible (an error of less than one degree). Moreover, it provides a control (based on the accurate reconstruction of $\varphi$) that, when plugged into the original system, gives a good approximation of the real extremal solution.

This lead us to the following conjecture, in terms of asymptotic results as $\epsilon$ goes to 0. First, with a suitable choice of $J_\epsilon(0)$ and $J_\epsilon$ a solution of a differential solution whose right-hand side depends on slow variables only, one gets not only that $I_\epsilon$ and $J_\epsilon$ are $\epsilon$-close, but also that their "moving averages" are $\epsilon^2$-close on an interval of length of order $1/\epsilon$. Second, it allows a prediction of the fast variable through averaging with an error of order $\epsilon$, which is reasonable because the fast variable more or less integrates a function of the slow variables and an estimate of the difference between moving averages is as useful as a point to point estimate in this integration. The goal of this (preliminary) paper is to formalize and to prove this conjecture. We deal with the Cauchy problem instead of the two point boundary value problem in optimal control because it is simpler and has an interest on its own. Note that we have to add to the first order averaged dynamics some corrective terms of order $\epsilon^2$, that were not explicitly present in Dargent (2015) for the precise system studied there.

The proposed solution is a sort of second order averaging without the oscillating transformation, or rather what plays the role of this transformation is the moving average, and it only has to be formally computed *at initial time* to provide the offset between the two initial conditions that we mentioned above. The terms of order $\epsilon^2$ that appear here are not the same as the (non unique) ones obtained in second order averaging Lochak and Meunier (1988); Sanders et al. (2007), this makes sense since, although it plays the same role, our moving average (also non-unique: it is coordinate-dependent) does not coincide with the oscillatory transformation of second order averaging.

The "filtered" estimate at order 2 on the slow variable may seem a bit abstract at first sight. Intuitively it means that the real solution is oscillating "symmetrically around" its estimate, while Theorem 4 typically provides an estimate that is not centered with respect to the real solution. More interestingly, it allows a reconstruction of the *fast* variable up to an error of order $\epsilon$, which is obviously impossible from classical first order averaging. Obtaining an asymptotic reconstruction of the fast variable via averaging has not yet been proposed in the literature, to the best of our knowledge, and, as we have already explained, this has a clear practical interest.

## 3. MAIN RESULTS

Our main results require the following construction:

*Proposition 2. For any map $f$ satisfying assumptions (A1)-(A3) and (ICA), there is a unique map $B : O \times S^1 \mapsto \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ such that*

- *(i)* $\frac{\partial B}{\partial \varphi}(I, \varphi) = \frac{\partial f}{\partial I}(I, \varphi) - \frac{\partial \overline{f}}{\partial I}(I)$,
- *(ii)* $B(I, \varphi + 2\pi) = B(I, \varphi)$,
- *(iii)* $\overline{B}(I) = \frac{1}{2\pi} \int_0^{2\pi} B(I, \xi) d\xi = 0$.

*and it is given by*

$$B(I, \varphi) = \int_0^\varphi \frac{\partial f}{\partial I}(I, \xi) d\xi - \frac{1}{2\pi} \int_0^{2\pi} (\pi + \varphi - \xi) \frac{\partial f}{\partial I}(I, \xi) d\xi.$$

In Lochak and Meunier (1988), this map $B$ is called the *integral with zero mean value of the fluctuating part of* $\frac{\partial f}{\partial I}$, and is denoted by $\phi[\frac{\partial f}{\partial I}]$.

*Proof.* The map $B(I, \varphi)$ is $2\pi$-periodic in $\varphi$ because $\varphi \mapsto \frac{\partial f}{\partial I}(I, \varphi) - \frac{\partial \overline{f}}{\partial I}(I)$ is $2\pi$-periodic and has zero mean. The other properties are clear and the proof of the above proposition reduces to elementary computations. $\square$

Consider the system $(S_\epsilon)$, given by (1). Instead of the averaged system $(AvS_\epsilon)$ introduced in Theorem 1, we now define the following "corrected averaged system":

$$(CAvS_\epsilon): \quad \begin{aligned} \dot{J} &= \epsilon \overline{f}(J) \left( 1 - \epsilon \frac{\pi}{\omega(J)^2} \omega'(J) \cdot \overline{f}(J) \right) \\ &\quad - \epsilon^2 \frac{1}{2\pi\omega(I)} \int_0^{2\pi} B(I, \xi) f(I, \xi) d\xi, \quad (6) \\ \dot{\psi} &= \omega(J) \left( 1 - \epsilon \frac{\pi}{\omega(J)^2} \omega'(J) \cdot \overline{f}(J) \right). \end{aligned}$$

Note that $(CAvS_\epsilon)$ includes an equation corresponding to the fast variable, contrary to the system $(AvS_\epsilon)$, see (4), where the fast variable has been eliminated. This is because we mean to reconstruct $\varphi$ up to order $\epsilon$.

Our result below will give, like classical first order averaging results, an estimation of $\|I_\epsilon(t) - J_\epsilon(t)\|$ of order $\epsilon$, but it will also give an estimation of order $\epsilon^2$ of the integral error on one period of oscillations, namely

$$\left\| \int_0^{T(I_\epsilon(t))} I_\epsilon(t + \tau) - J_\epsilon(t + \tau) d\tau \right\|. \quad (7)$$

Recall that $T(I)$ was defined by (2). The result is the following.

*Theorem 3. Assume that assumptions (A1)-(A3) and (ICA), introduced in Section 2, are satisfied. Let $k$ be some positive constant. Consider a family $(I_\epsilon(t), \varphi_\epsilon(t))$ of solutions of $(S_\epsilon)$ such that $\|I_\epsilon(0) - I_0\| \le k\epsilon^2$ and $\|\varphi_\epsilon(0) - \varphi_0\| \le k\epsilon$, and a family $(J_\epsilon(t), \psi_\epsilon(t))$ of solutions of $(CAvS_\epsilon)$ such that*

$$\left\| J_\epsilon(0) - I_0 \right.$$
$$\left. - \frac{\epsilon}{2\pi\omega(I_0)} \int_0^{2\pi} (\pi - \xi) f(I_0, \varphi_0 + \xi) d\xi \right\| \le k\epsilon^2, \quad (8)$$
$$\|\psi_\epsilon(0) - \varphi_0\| \le k\epsilon.$$

*Then, there exist two positive constants $\epsilon_0$ and $c$ such that for all $0 \le \epsilon \le \epsilon_0$, the solution of the system $(S_\epsilon)$ exists for $0 \le t \le \frac{T}{\epsilon}$, remains in $O \times S^1$ and*

$$\left\| \int_0^{T(I_\epsilon(t))} I_\epsilon(t + \tau) - J_\epsilon(t + \tau) d\tau \right\| \le c\epsilon^2, \quad (9)$$

*and*

$$\|I_\epsilon(t) - J_\epsilon(t)\| \le c\epsilon, \quad \|\varphi_\epsilon(t) - \psi_\epsilon(t)\| \le c\epsilon, \quad (10)$$

*where $c$ depends on $k$ and on the bounds in assumptions (A1)-(A3) only.*

We state (9), but do not claim that $\|I_\epsilon(t) - J_\epsilon(t)\| \le c\epsilon^2$ holds. Indeed, it was noticed in Section 2.3 that $I_\epsilon$ cannot be estimated at an order better than $\epsilon$ by a solution of a differential equation whose right-hand side depends on slow variables only.

Note also that the initial condition $J_\epsilon(0)$ is chosen different from $I_\epsilon(0)$ and such that the expression (7) is small at initial time: indeed, condition (15) is equivalent to

$$\left\| \int_0^{T(I_\epsilon(0))} I_\epsilon(\tau) - J_\epsilon(\tau) d\tau \right\| \le k\epsilon^2.$$

In order to apply Theorem 3, one has to construct the corrected averaged system $(CAvS_\epsilon)$ and its initial conditions. In $(CAvS_\epsilon)$, all the terms can be numerically computed using quadrature formulas. The condition on the initial conditions may obviously be implemented as

$$J_\epsilon(0) = I_0 + \frac{\epsilon}{2\pi\omega(I_0)} \int_0^{2\pi} (\pi - \xi) f(I_0, \varphi_0 + \xi) d\xi,$$
$$\psi_\epsilon(0) = \varphi_0.$$

*Formulation in terms of moving averages.* Theorem 3 can be stated in terms of "moving averages" defined as follows. To a solution $t \mapsto (I_\epsilon(t), \varphi_\epsilon(t))$ of $(S_\epsilon)$, we associate its "moving average" denoted by $(\tilde{I}_\epsilon(t), \tilde{\varphi}_\epsilon(t))$, and defined as follows:

$$\begin{aligned} \tilde{I}_\epsilon(t) &= \frac{1}{T(I_\epsilon(t))} \int_0^{T(I_\epsilon(t))} I_\epsilon(t + \tau) d\tau, \\ \tilde{\varphi}_\epsilon(t) &= \frac{1}{T(I_\epsilon(t))} \int_0^{T(I_\epsilon(t))} \varphi_\epsilon(\varphi_0, t + \tau) d\tau. \end{aligned} \quad (11)$$

Since the average is taken over one period of the fast variable, we may foresee that $\tilde{I}_\epsilon$ is a low pass filtered version of $I_\epsilon$, i.e., it oscillates less. Similarly, we denote by $(\tilde{J}_\epsilon(t), \tilde{\psi}(t))$ the moving average over one period of the solution $t \mapsto (J_\epsilon(t), \psi(t))$ of $(CAvS_\epsilon)$:

$$\begin{aligned} \tilde{J}_\epsilon(t) &= \frac{1}{T(J_\epsilon(t))} \int_0^{T(J_\epsilon(t))} J_\epsilon(t + \tau) d\tau, \\ \tilde{\psi}_\epsilon(t) &= \frac{1}{T(J_\epsilon(t))} \int_0^{T(J_\epsilon(t))} \psi_\epsilon(t + \tau) d\tau. \end{aligned} \quad (12)$$

In terms of moving averages, condition (8) on the initial conditions is equivalent to

$$\|\tilde{I}_\epsilon(0) - \tilde{J}_\epsilon(0)\| \le k\epsilon^2,$$

and the estimate (9) is equivalent to

$$\|\tilde{I}_\epsilon(t) - \tilde{J}_\epsilon(t)\| \le c\epsilon^2.$$

These moving averages are used in the proofs. Theorem 3 could be reformulated using them only but it would be less explicit.

*The case of constant frequency.* To simplify the understanding of the paper, we will re-state our main result for the particular and simpler case where $\dot\varphi$ is constant and, specifically, $\omega(I) = 1$ and $g = 0$ in (1), i.e., the dynamics of the fast variable is given by $\dot\varphi = 1$. With the additional assumption that $\varphi(0) = 0$, this yields $\varphi(t) = t[2\pi]$ and the system $(S_\epsilon)$ reduces to:

$$(S_\epsilon): \quad \dot I = \epsilon f(I, t). \tag{13}$$

The definition (6) of the "corrected averaged system" reduces to

$$(CAvS_\epsilon): \quad \dot J = \epsilon \overline{f}(J) - \epsilon^2 \frac{1}{2\pi} \int_0^{2\pi} B(I, \tau) f(I, \tau) d\tau. \tag{14}$$

Since the fast variable is now time, we do not have an equation for this fast variable in (13). Theorem 3 becomes:

*Theorem 4. Let $(S_\epsilon)$ and $(CAvS_\epsilon)$ be given by (13) and (14), respectively. Assume that assumptions (A1)-(A3) and (ICA) are satisfied. Let $k$ be some positive constant. Consider a family $(I_\epsilon(t), \varphi_\epsilon(t))$ of solutions of $(S_\epsilon)$ such that $||I_\epsilon(0) - I_0|| \leq k\epsilon^2$ and a family $J_\epsilon(t)$ of solutions of $(CAvS_\epsilon)$ such that*

$$||J_\epsilon(0) - I_0 - \frac{\epsilon}{2\pi} \int_0^{2\pi} (\pi - \tau) f(I_0, \tau) d\tau|| \leq k\epsilon^2. \tag{15}$$

*Then, there exist two positive constants $\epsilon_0$ and $c$ such that for all $0 \leq \epsilon \leq \epsilon_0$, the solution of the system $(S_\epsilon)$ exists for $0 \leq t \leq \frac{T}{\epsilon}$, remains in $O \times S^1$ and*

$$|| \int_0^{2\pi} I_\epsilon(t + \tau) - J_\epsilon(t + \tau) d\tau|| \leq c\epsilon^2, \tag{16}$$

*and*

$$||I_\epsilon(t) - J_\epsilon(t)|| \leq c\epsilon, \tag{17}$$

*where $c$ depends on $k$ and on the bounds in assumptions (A1)-(A3) only.*

The remarks for Theorem 3 are still valid. In particular, Theorem 4 can be reformulated in terms of moving averages. To a solution $t \mapsto I_\epsilon(t)$ of $(S_\epsilon)$ and a solution $t \mapsto J_\epsilon(t)$ of $(CAvS_\epsilon)$, we associate their moving averages

$$\tilde I_\epsilon(t) = \frac{1}{2\pi} \int_0^{2\pi} I_\epsilon(t + \tau) d\tau, \tag{18}$$

$$\tilde J_\epsilon(t) = \frac{1}{2\pi} \int_0^{2\pi} J_\epsilon(t + \tau) d\tau. \tag{19}$$

It is immediate that the estimate (16) of Theorem 4 is actually equivalent to

$$||\tilde I_\epsilon(t) - \tilde J_\epsilon(t)|| \leq c\epsilon^2,$$

and it can be shown that (15) is, in fact, equivalent to

$$||\tilde I_\epsilon(0) - \tilde J_\epsilon(0)|| \leq k\epsilon^2.$$

To sum up, the interest of our method is three-fold: firstly, it allows us to give high order approximation without using involved transformations including the fast oscillations or reconstructing the slow variables. Secondly, it enables us to conclude that the solution of the averaged system is very close to the average of the solution of the original problem. Finally, it makes possible the reconstruction of the fast variable up to order $\epsilon$, contrary to the classical results, where $\varphi$ can only be reconstructed up to a difference that is bounded as $\epsilon$ goes to 0.

## 4. PROOFS OF THE MAIN RESULTS

Since the proofs should fit the length of a conference paper, we will present the detailed proof for the simpler particular case and only a sketch of proof for the general result for which $\dot\varphi = \omega(I)$. Its proof uses similar arguments, but it is more involved and will be presented in detail in a future paper.

In Sections 4.1- 4.2, we assume $\dot\varphi = 1$ and $\varphi(0) = 0$ in (1).

*4.1 Notations and useful results*

Throughout, we will denote by $\mathrm{Lip}\, f$ (resp. $\mathrm{Sup}\, f$) the Lipschitz constant (resp. the supremum) of $f$.

By $a = b + \mathcal{O}(\epsilon^j)$, we mean $||a - b|| \leq k\epsilon^j$ for some positive constant $k$. The reader may check that this constant $k$ will always depend only on the constants associated to the system ($\omega_{min}$, $\mathrm{Lip}\, f$, $\mathrm{Sup}\, f$, $\mathrm{Lip}\, \frac{\partial f}{\partial I} \ldots$).

In order to simplify the notation, we will denote the corrective term introduced in $(CAvS_\epsilon)$, given by (14), by

$$E(I) = \frac{1}{2\pi} \int_0^{2\pi} B(I, \tau) f(I, \tau) d\tau.$$

With that notation, $(CAvS_\epsilon)$ reads:

$$(CAvS_\epsilon): \quad \dot J = \epsilon \overline{f}(J) - \epsilon^2 E(J). \tag{20}$$

The Gronwall lemma will allow us to bound the differnce between $\tilde I_\epsilon$ and $\tilde J_\epsilon$. It is very classical, we write down the version from Lochak and Meunier (1988):

*Lemma. (Gronwall lemma, Lochak and Meunier (1988)) Let $x(t)$, $a(t)$ and $\gamma(t)$ be positive continuous functions defined on an interval $[0, t]$. Assume that $\gamma$ is of class $C^1$ and the following inequality holds:*

$$x(t) \leq \gamma(t) + \int_0^t a(s) x(s) ds.$$

*Then we have:*

$$x(t) \leq \gamma(0) e^{\int_0^t a(s) ds} + \int_0^t \gamma'(s) e^{\int_s^t a(u) du} ds.$$

The following technical results (Propositions 5, 6, 8, 10, 11) are needed in the proof of Theorem 4. We suppose in all of them that assumptions (A1)-(A3) and (ICA) hold.

*Proposition 5. Consider the system $(S_\epsilon)$, given by (13), the moving average $\tilde I_\epsilon$, given by (18), of the solution $I_\epsilon$ of $(S_\epsilon)$ and $\tau$ such that $0 \leq \tau \leq 2\pi$. Then $I_\epsilon(t + \tau) = \tilde I_\epsilon(t) + \mathcal{O}(\epsilon)$.*

*Proof.* We have:

$$I_\epsilon(t + \tau) - \tilde I_\epsilon(t) = \frac{1}{2\pi} \int_0^{2\pi} I_\epsilon(t + \tau) - I_\epsilon(t + \tau') d\tau'.$$

Applying the mean value theorem, for $1 \leq i \leq n$ (where $n$ is the dimension of $I$), there exists $\tau_i$ between $\tau$ and $\tau'$ such that $I_\epsilon^i(t + \tau) - I_\epsilon^i(t + \tau') = \epsilon f^i(I_\epsilon(t + \tau_i))(\tau - \tau')$ and we immediately deduce $I_\epsilon(t + \tau) - \tilde I_\epsilon(t) = \mathcal{O}(\epsilon)$ and

$$||I_\epsilon(t + \tau) - \tilde I_\epsilon(t)|| \leq 2\pi\epsilon (\mathrm{Sup}\, f)$$

thus $||\mathcal{O}(\epsilon)|| \leq k\epsilon$, with $k$ depending on $\mathrm{Sup}\, f$ only. $\square$

*Proposition 6. Consider the corrected averaged system $(CAvS_\epsilon)$, given by (14), the moving average $\tilde{J}_\epsilon$, given by (19), of the solution $J_\epsilon$ of $(CAvS_\epsilon)$ and $\tau$ such that $0 \le \tau \le 2\pi$. Then $J_\epsilon(t+\tau) = \tilde{J}_\epsilon(t) + \epsilon\overline{f}(J_\epsilon(t+\pi))(\tau-\pi) + \mathcal{O}(\epsilon^2)$.*

*Remark 7. In some of the proofs, we will need only the weaker result $J_\epsilon(t+\tau) = \tilde{J}_\epsilon(t) + \mathcal{O}(\epsilon)$, with $||\mathcal{O}(\epsilon)|| \le k\epsilon$, where $k$ is a positive constant, depending on $\text{Sup}\,\overline{f}$ only. This can be proven following exactly the same line as that of Proposition 5 .*

*Proof.* We have

$$J_\epsilon(t+\tau) = J_\epsilon(t) + \epsilon \int_t^{t+\tau} \overline{f}(J_\epsilon(\tau'))d\tau'$$

$$= J_\epsilon(t) + \epsilon\tau\overline{f}(J_\epsilon(t)) + \mathcal{O}(\epsilon^2)$$

and

$$\tilde{J}_\epsilon(t) = J_\epsilon(t) + \frac{\epsilon}{2\pi} \int_0^{2\pi} \tau\overline{f}(J_\epsilon(t))d\tau + \mathcal{O}(\epsilon^2)$$

$$= J_\epsilon(t) + \epsilon\pi\overline{f}(J_\epsilon(t)) + \mathcal{O}(\epsilon^2).$$

From this, it follows immediately $J_\epsilon(t+\tau) = \tilde{J}_\epsilon(t) + \epsilon\overline{f}(J_\epsilon(t+\pi))(\tau-\pi) + \mathcal{O}(\epsilon^2)$. $\qquad\square$

*Proposition 8. Consider the system $(S_\epsilon)$, given by (13) and the moving average $\tilde{I}_\epsilon$, given by (18), of its solution $I_\epsilon$. Then*

$$\dot{\tilde{I}}_\epsilon(t) - \epsilon F(\tilde{I}_\epsilon(t),\epsilon) = \epsilon^2 \dot{A}_\epsilon(t) + \mathcal{O}(\epsilon^3),$$

*where $||A_\epsilon(t)||$ is bounded by a positive constant and $F(\tilde{I}_\epsilon,\epsilon) = \overline{f}(\tilde{I}_\epsilon) - \epsilon E(\tilde{I}_\epsilon)$.*

*Remark 9. The proof of our main result requires $||A_\epsilon(t)||$ to be bounded, but needs no assumption on $||\dot{A}_\epsilon(t)||$.*

*Proof.* According to the definitions of $\tilde{I}_\epsilon$ and $\overline{f}$, we have:

$$\dot{\tilde{I}}_\epsilon(t) - \epsilon F(\tilde{I}_\epsilon(t),\epsilon) =$$

$$= \frac{\epsilon}{2\pi} \int_0^{2\pi} f(I_\epsilon(t+\tau),t+\tau)d\tau - \epsilon\overline{f}(\tilde{I}_\epsilon) + \epsilon^2 E(\tilde{I}_\epsilon(t))$$

$$= \frac{\epsilon}{2\pi} \int_0^{2\pi} \left(f(I_\epsilon(t+\tau),t+\tau) - f(\tilde{I}_\epsilon(t),t+\tau)\right)d\tau + \epsilon^2 E(\tilde{I}_\epsilon(t)).$$

Applying a Taylor expansion of $f(I_\epsilon(t+\tau),t+\tau)$, with respect to its first entry, around $\tilde{I}_\epsilon(t)$ and since $I_\epsilon(t+\tau) - \tilde{I}_\epsilon(t) = \mathcal{O}(\epsilon)$, see Proposition 5 above, we deduce:

$$\dot{\tilde{I}}_\epsilon(t) - \epsilon F(\tilde{I}_\epsilon(t),\epsilon) =$$

$$= \frac{\epsilon}{2\pi} \int_0^{2\pi} \frac{\partial f}{\partial I}(\tilde{I}_\epsilon(t),t+\tau)(I_\epsilon(t+\tau) - \tilde{I}_\epsilon(t))d\tau$$

$$+ \epsilon^2 E(\tilde{I}_\epsilon(t)) + \mathcal{O}(\epsilon^3).$$

Now recall that $\frac{\partial B}{\partial t}(I,t) = \frac{\partial f}{\partial I}(I,t) - \frac{\partial\overline{f}}{\partial I}(I)$, where $B(I,t)$ is $2\pi$-periodic and of zero mean. Replacing in the above expression $\frac{\partial f}{\partial I}(\tilde{I}_\epsilon(t),t+\tau)$ by $\frac{\partial\overline{f}}{\partial I}(\tilde{I}_\epsilon(t)) + \frac{\partial B}{\partial t}(\tilde{I}_\epsilon(t),t+\tau)$ and observing that $\int_0^{2\pi} I_\epsilon(t+\tau) - \tilde{I}_\epsilon(t)d\tau = 0$ and $\int_0^{2\pi} \frac{\partial B}{\partial t}(\tilde{I}_\epsilon(t),t+\tau)\tilde{I}_\epsilon(t)d\tau = 0$ , we obtain:

$$\dot{\tilde{I}}_\epsilon(t) - \epsilon F(\tilde{I}_\epsilon(t),\epsilon) = \frac{\epsilon}{2\pi} \int_0^{2\pi} \frac{\partial B}{\partial t}(\tilde{I}_\epsilon(t),t+\tau)I_\epsilon(t+\tau)d\tau$$

$$+ \epsilon^2 E(\tilde{I}_\epsilon(t)) + \mathcal{O}(\epsilon^3).$$

Integrating by parts, we have:

$$\dot{\tilde{I}}_\epsilon(t) - \epsilon F(\tilde{I}_\epsilon(t),\epsilon) = \frac{\epsilon}{2\pi} \left[B(\tilde{I}_\epsilon(t),t+\tau)I_\epsilon(t+\tau)\right]_{\tau=0}^{\tau=2\pi}$$

$$- \frac{\epsilon^2}{2\pi} \int_0^{2\pi} B(\tilde{I}_\epsilon(t),t+\tau)f(I_\epsilon(t+\tau),t+\tau)d\tau$$

$$+ \epsilon^2 E(\tilde{I}_\epsilon(t)) + \mathcal{O}(\epsilon^3).$$

Recall that $E(\tilde{I}_\epsilon) = \frac{1}{2\pi} \int_0^{2\pi} B(\tilde{I}_\epsilon,\tau)f(\tilde{I}_\epsilon,\tau)d\tau$ and due to the periodicity of $B$ and $f$, we deduce $E(\tilde{I}_\epsilon(t)) = \frac{1}{2\pi} \int_0^{2\pi} B(\tilde{I}_\epsilon(t),t+\tau)f(\tilde{I}_\epsilon(t),t+\tau)d\tau$. It can be easily shown that $f(I_\epsilon(t+\tau),t+\tau) = f(\tilde{I}_\epsilon(t),t+\tau) + \mathcal{O}(\epsilon)$ and by replacing it in the integral of the above relation, we get

$$\dot{\tilde{I}}_\epsilon(t) - \epsilon F(\tilde{I}_\epsilon(t),\epsilon) = \frac{\epsilon}{2\pi} B(\tilde{I}_\epsilon(t),t)(I_\epsilon(t+2\pi) - I_\epsilon(t))$$

$$+ \mathcal{O}(\epsilon^3).$$

Let $C(I,t)$ be a $2\pi$-periodic primitive of $B(I,t)$. Notice that $C$ is bounded. We set

$$A_\epsilon(t) = \frac{1}{2\pi\epsilon} C(\tilde{I}_\epsilon(t),t)(I_\epsilon(t+2\pi) - I_\epsilon(t)).$$

Then

$$\dot{\tilde{I}}_\epsilon(t) - \epsilon F(\tilde{I}_\epsilon(t),\epsilon) =$$

$$= \epsilon^2[\dot{A}_\epsilon(t) - \frac{1}{2\pi}\frac{\partial C}{\partial I}(\tilde{I}_\epsilon(t),t)\frac{\dot{\tilde{I}}(t)}{\epsilon}(I_\epsilon(t+2\pi) - I_\epsilon(t))$$

$$- \frac{1}{2\pi}C(\tilde{I}_\epsilon(t),t)(f(I_\epsilon(t+2\pi),t) - f(I_\epsilon(t),t))] + \mathcal{O}(\epsilon^3),$$

which gives

$$\dot{\tilde{I}}_\epsilon(t) - \epsilon F(\tilde{I}_\epsilon(t),\epsilon) = \epsilon^2 \dot{A}_\epsilon(t) + \mathcal{O}(\epsilon^3),$$

since $\frac{\dot{\tilde{I}}(t)}{\epsilon} \le \text{Sup}\,f$ and $I_\epsilon(t+2\pi) = I_\epsilon(t) + \mathcal{O}(\epsilon)$ (more precisely, $||I_\epsilon(t+2\pi) - I_\epsilon(t)|| \le 2\pi\epsilon\,\text{Sup}\,f$). From the last inequality and the definition of $A_\epsilon(t)$, we easily deduce that $||A_\epsilon(t)|| \le (\text{Sup}\,C)(\text{Sup}\,f)$. $\qquad\square$

*Proposition 10. Consider the corrected averaged system $(CAvS_\epsilon)$, given by (14), and the moving average $\tilde{J}_\epsilon$, given by (19), of its solution $J_\epsilon$. Then*

$$\epsilon F(\tilde{J}_\epsilon,\epsilon) - \dot{\tilde{J}}_\epsilon = \mathcal{O}(\epsilon^3),$$

*where $F(\tilde{J}_\epsilon,\epsilon) = \overline{f}(\tilde{J}_\epsilon) - \epsilon E(\tilde{J}_\epsilon)$.*

*Proof.* We have

$$\epsilon F(\tilde{J}_\epsilon(t),\epsilon) - \dot{\tilde{J}}_\epsilon(t) =$$

$$= \epsilon\overline{f}(\tilde{J}_\epsilon(t)) - \epsilon^2 E(\tilde{J}_\epsilon(t))$$

$$- \frac{\epsilon}{2\pi} \int_0^{2\pi} \overline{f}(J_\epsilon(t+\tau)) - \epsilon E(J_\epsilon(t+\tau))d\tau$$

$$= \frac{\epsilon}{2\pi} \int_0^{2\pi} \overline{f}(\tilde{J}_\epsilon(t)) - \overline{f}(J_\epsilon(t+\tau))d\tau$$

$$+ \frac{\epsilon^2}{2\pi} \int_0^{2\pi} E(J_\epsilon(t+\tau)) - E(\tilde{J}_\epsilon(t))d\tau.$$

Since $J_\epsilon(t+\tau) = \tilde{J}_\epsilon(t) + \mathcal{O}(\epsilon)$, we immediately have $E(J_\epsilon(t+\tau)) = E(\tilde{J}_\epsilon(t)) + \mathcal{O}(\epsilon)$. It remains to show that $\frac{1}{2\pi} \int_0^{2\pi} \overline{f}(\tilde{J}_\epsilon(t)) - \overline{f}(J_\epsilon(t+\tau))d\tau = \mathcal{O}(\epsilon^2)$. By Proposition 6, $J_\epsilon(t+\tau) = \tilde{J}_\epsilon(t) + \epsilon\overline{f}(J_\epsilon(t+\pi))(\tau-\pi) + \mathcal{O}(\epsilon^2)$.

Thus $\overline{f}(\check{J}_\epsilon(t)) - \overline{f}(J_\epsilon(t+\tau)) = -\epsilon \frac{\partial \overline{f}}{\partial J}(\check{J}_\epsilon(t))\overline{f}(J_\epsilon(t+\pi))(\tau - \pi) + \mathcal{O}(\epsilon^2)$ and it follows that $\frac{1}{2\pi}\int_0^{2\pi} \overline{f}(\check{J}_\epsilon(t)) - \overline{f}(J_\epsilon(t+\tau))\mathrm{d}\tau = \mathcal{O}(\epsilon^2)$. Finally, we obtain $\epsilon F(\check{J}_\epsilon, \epsilon) - \dot{\check{J}}_\epsilon = \mathcal{O}(\epsilon^3)$. $\qquad\square$

*Proposition 11. Consider a family $I_\epsilon(t)$ of solutions of $(S_\epsilon)$, given by ( 14), such that $||I_\epsilon(0) - I_0|| \leq k\epsilon^2$, for all $\epsilon$ small enough, where $k$ is a positive constant, and a family $J_\epsilon(t)$ of solutions of $(CAvS_\epsilon)$, given by (14). We have equivalence between*

$$||\check{J}_\epsilon(0) - \tilde{I}_\epsilon(0)|| = \mathcal{O}(\epsilon^2)$$

*and*

$$||J_\epsilon(0) - I_0 - \frac{\epsilon}{2\pi}\int_0^{2\pi}(\pi - \tau)f(I_0, \tau)\mathrm{d}\tau|| = \mathcal{O}(\epsilon^2).$$

*Proof.* We have:

$$I_\epsilon(t) = I_0 + \epsilon \int_0^t f(I(s), s)\mathrm{d}s + \mathcal{O}(\epsilon^2)$$

$$= I_0 + \epsilon \int_0^t f(I_0, s)\mathrm{d}s + \mathcal{O}(\epsilon^2)$$

and

$$\tilde{I}_\epsilon(0) = I_0 + \frac{\epsilon}{2\pi}\int_0^{2\pi}\int_0^\tau f(I_0, s)\mathrm{d}s\mathrm{d}\tau + \mathcal{O}(\epsilon^2)$$

$$= I_0 + \frac{\epsilon}{2\pi}\int_0^{2\pi}(2\pi - s)f(I_0, s)\mathrm{d}s + \mathcal{O}(\epsilon^2).$$

(21)

Similarly,

$$J_\epsilon(t) = J_\epsilon(0) + \epsilon \int_0^t \overline{f}(J(s)) - \epsilon E(J(s))\mathrm{d}s$$

$$= J_\epsilon(0) + \epsilon \overline{f}(I_0)t + \mathcal{O}(\epsilon^2)$$

and

$$\check{J}_\epsilon(0) = J_\epsilon(0) + \frac{\epsilon}{2\pi}\int_0^{2\pi}\pi f(I_0, s)\mathrm{d}s + \mathcal{O}(\epsilon^2). \qquad (22)$$

It follows from (21) and (22) that $||\check{J}_\epsilon(0) - \tilde{I}_\epsilon(0)|| = \mathcal{O}(\epsilon^2)$ if and only if $||J_\epsilon(0) - I_0 - \frac{\epsilon}{2\pi}\int_0^{2\pi}(\pi - s)f(I_0, s)\mathrm{d}s|| = \mathcal{O}(\epsilon^2)$. $\qquad\square$

*4.2 Proof of Theorem 4*

The estimation (17) is a consequence of Theorem 1, that tolerates an $\epsilon$-discrepancy in the initial conditions (the corrective terms, of order $\epsilon^2$, cannot provoke more than an $\epsilon$ deviation on the time interval). Let us prove (16).

We have

$$\dot{\tilde{I}}_\epsilon - \dot{\check{J}}_\epsilon = \dot{\tilde{I}}_\epsilon - \epsilon F(\tilde{I}_\epsilon, \epsilon) + \epsilon(F(\tilde{I}_\epsilon, \epsilon) - F(\check{J}_\epsilon, \epsilon)) + \epsilon F(\check{J}_\epsilon, \epsilon) - \dot{\check{J}}_\epsilon,$$

where $F(\tilde{I}_\epsilon, \epsilon) = \overline{f}(\tilde{I}_\epsilon) - \epsilon E(\tilde{I}_\epsilon)$. By Propositions 8 and 10, we get

$$\dot{\tilde{I}}_\epsilon(t) - \dot{\check{J}}_\epsilon(t) = \epsilon^2 \dot{A}_\epsilon(t) + \epsilon(F(\tilde{I}_\epsilon(t), \epsilon) - F(\check{J}_\epsilon(t), \epsilon)) + \epsilon^3 b(\epsilon, t),$$

where $b$ is bounded. Thus

$$\tilde{I}_\epsilon(t) - \check{J}_\epsilon(t) = \tilde{I}_\epsilon(0) - \check{J}_\epsilon(0) + \epsilon^2(A_\epsilon(t) - A_\epsilon(0))$$

$$+ \int_0^t \epsilon(F(\tilde{I}_\epsilon(\tau), \epsilon) - F(\check{J}_\epsilon(\tau), \epsilon))$$

$$+ \epsilon^3 b(\epsilon, \tau)\mathrm{d}\tau,$$

and we have

$$||\tilde{I}_\epsilon - \check{J}_\epsilon|| \leq ||\tilde{I}_\epsilon(0) - \check{J}_\epsilon(0)||$$

$$+ \epsilon^2(||A_\epsilon|| + ||A_\epsilon(0)||) + \epsilon^3(\text{Sup } b)t$$

$$+ \epsilon \int_0^t (\text{Lip } F)||\tilde{I}_\epsilon - \check{J}_\epsilon||\mathrm{d}\tau.$$

Now, recall that according to the theorem's assumptions $||J_\epsilon(0) - I_0 - \frac{\epsilon}{2\pi}\int_0^{2\pi}(\pi - \tau)f(I_0, \tau)\mathrm{d}\tau|| = \mathcal{O}(\epsilon^2)$. Therefore, by Proposition 11, we have $||\check{J}_\epsilon(0) - \tilde{I}_\epsilon(0)|| = \mathcal{O}(\epsilon^2)$. The result follows immediately by applying Gronwall lemma with $x(t) = \tilde{I}_\epsilon(t) - \check{J}_\epsilon(t)$, $||\tilde{I}_\epsilon(0) - \check{J}_\epsilon(0)|| = \mathcal{O}(\epsilon^2)$ and for $t$ such that $0 \leq t \leq \frac{T}{\epsilon}$. $\qquad\square$

*4.3 Sketch of the proof of Theorem 3*

We now consider the general system $(S_\epsilon)$, given by (1).

Proving the first estimate (9) follows the same line as that of the proof of Theorem 4. Propositions 8-10 remain true for the general case and we can formulate their counterparts for the fast variables $\varphi_\epsilon$ and $\psi_\epsilon$.

In order to obtain the estimate on the fast variable, one has to notice that $\tilde{\varphi}_\epsilon - \tilde{\psi}_\epsilon$ is, roughly speaking, the solution of a differential equation with a forcing term bounded by $||\tilde{I}_\epsilon - \check{J}_\epsilon||$, itself bounded by $c\epsilon^2$, on a time interval of length $\frac{T}{\epsilon}$. This yields an estimate of the form $||\tilde{\varphi}_\epsilon(t) - \tilde{\psi}_\epsilon(t)|| \leq c\epsilon$, i.e.

$$||\int_0^{T(I_\epsilon(t))}(\varphi_\epsilon(t+\tau) - \psi_\epsilon(t+\tau))\,\mathrm{d}\tau||,$$

and this in turn implies $||\varphi_\epsilon(t) - \psi_\epsilon(t)|| \leq c\epsilon$ because the variation of $\varphi_\epsilon - \psi_\epsilon$ in the integral is of order $\epsilon$. $\qquad\square$

REFERENCES

Arnold, V. (1978). *Mathematical Methods of Classical Mechanics*. Graduate Text in Mathematics 60, Springer-Verlag New York.

Besjes, J. (1969). On asymptotic methods for non-linear differential equations. *Journal de Mécanique*, 8, 357–373.

Bogoliubov, N.N. and Mitropolsky, Y. (1961). *Asymptotic methods in the theory of non-linear oscillations*. Gordon and Breach, New York.

Chaplais, F. (1987). Averaging and deterministic optimal control. *SIAM J. Control Optim.*, 25(3), 767–780.

Dargent, T. (2015). Initial and final boundaries transformation when solving optimal control problems with averaging techniques and application to low thrust orbit transfer. In *Proc. 66th International Astronautical Congress IAC 2015, Jerusalem, Israel*.

Edelbaum, T.N. (1965). Optimum power-limited orbit transfer in strong gravity fields. *AIAA J.*, 3, 921–925.

Geffroy, S. (1997). *Généralisation des techniques de moyennation en contrle optimal – Application aux problèmes de rendez-vous orbitaux en poussée faible*. Thèse de doctorat, I. N. P. Toulouse, France.

Lochak, P. and Meunier, C. (1988). *Multiphase averaging for classical systems. With applications to adiabatic theorems*. Appl. Math. Sc. 72. Springer-Verlag.

Morosi, C. and Pizzocchero, L. (2006). On the average principle for one-frequency systems. *J. Phys. A: Math. & General*, 39(14), 3673–3702.

Sanders, J., Verhulst, F., and Murdock, J. (2007). *Averaging methods in nonlinear dynamical systems*. Appl. Math. Sc. 59. Springer-Verlag.