Averaging in Hamiltonian systems with slowly varying parameters

Tewfik Sari

▶ **To cite this version:**

Tewfik Sari. Averaging in Hamiltonian systems with slowly varying parameters. RR-4374, INRIA. 2002. inria-00072214

HAL Id: inria-00072214
https://inria.hal.science/inria-00072214v1

Submitted on 23 May 2006

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

INRIA

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Averaging in Hamiltonian systems with slowly varying parameters

Tewfik Sari

N° 4374

January 2002

THÈME 4

ISSN 0249-6399 ISRN INRIA/RR--4374--FR+ENG

INRIA
SOPHIA ANTIPOLIS

Averaging in Hamiltonian systems with slowly varying parameters

Tewfik Sari *

Thème 4 — Simulation et optimisation
de systèmes complexes
Projet Comore

Rapport de recherche n° 4374 — January 2002 — 18 pages

Abstract: The aim of this paper is to describe the general averaging principle and to discuss the particular case of single-frequency systems, the case of systems with constant frequencies and the case of Hamiltonian systems. We show how the stroboscopic method, which is a method of the nonstandard perturbation theory of differential equations, can be used in this kind of problems. We give various examples which illustrate the simplicity and the effectiveness of the method.

Key-words: Perturbation theory, Hamiltonian systems, averaging, stroboscopic method.

This paper was presented as an invited talk at the First Mexican Meeting on Mathematical and Experimental Physics held at El Colegio Nacional, Mexico City, from 10 to 14 September 2001 (`http://www.colegionacional.org.mx/actividad3.htm`). The author wishes to thank the organizers of the meeting for their invitation.

* *Adresse permanente* : Université de Haute Alsace, 4 rue des Frères Lumière, 68093 Mulhouse, `T.Sari@uha.fr`

Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles, BP 93, 06902 Sophia Antipolis Cedex (France)
Téléphone : +33 4 92 38 77 77 — Télécopie : +33 4 92 38 77 65

Moyennisation dans les systèmes hamiltoniens à paramètres lentement variables

Résumé : Le but de cet article est de décrire le principe général de la moyennisation et de discutter le cas particulier des systèmes à une fréquence, celui des systèmes dont les fréquences sont constantes et celui des systèmes hamiltoniens. On montre comment appliquer dans ces problèmes la technique de stroboscopie, qui est une méthode de la théorie non standard des perturbations d'équations différentielles. On donne plusieurs exemples qui illustrent la simplicité de la méthode et son utilité.

Mots-clés : Théorie des perturbations, systèmes hamiltoniens, moyennisation, stroboscopie.

Contents

1 Introduction **3**

2 Averaging in Single-Frequency Systems **4**
2.1 Adiabatic Invariants 5
2.2 The Lorentz pendulum 6

3 The stroboscopic method **6**
3.1 Non Hamiltonian Perturbations 7
3.2 Applications 8
3.3 The evolution of the action variable 10
3.4 Single-frequency systems 11

4 Systems with Constant Frequencies **12**
4.1 KBM Theory 12
4.2 Precession of the planet Mercury 14

5 Averaging in Hamiltonian Systems **15**
5.1 KAM Theory 16
5.2 Arnold diffusion 16

1 Introduction

In nature one often encounters systems which are small perturbations of unperturbed systems. The unperturbed motion is assumed to be completely described. The main problem is to describe the perturbed motion. The problem of the motion of planets around the Sun can be regarded as a perturbation of the integrable problem of the motion of noninteracting planets around a fixed center of attraction. The unperturbed problem decomposes into k Kepler problems, where k is the number of planets. *Perturbation theory* and *Averaging* are the most effective methods for studying such perturbed motions. Historically these methods have arisen in the work of Lagrange and Laplace on secular perturbations of planets.

We consider the unperturbed system defined on $B \times \mathbb{T}^k$

$$\dot{I} = 0, \qquad \dot{\varphi} = \omega(I), \tag{1}$$

where $I \in B$, B is an open subset of $\mathbb{R}^m$, $\varphi \in \mathbb{T}^k$, $\mathbb{T}^k$ is the k-dimensional torus, and $\omega(I) \in \mathbb{R}^k$. The variables $I = (I_1, \cdots, I_m)$ are m first integrals of the system. On each torus I =const, the phases $\varphi = (\varphi_1, \cdots, \varphi_k)$ are coordinates. The frequencies are $\omega(I) = (\omega_1(I), \cdots, \omega_k(I))$. They are constant on each torus, but are varying from torus to torus. The perturbed system is a *slow-fast* system of the form

$$\dot{I} = \varepsilon f(I, \varphi, \varepsilon), \qquad \dot{\varphi} = \omega(I) + \varepsilon g(I, \varphi, \varepsilon), \tag{2}$$

where f and g are 2π-periodic in φ. In Equations (2) the variables I are *slow* and the phases φ are *fast* variables.

The *averaging principle* consists of replacing the perturbed system (2) by the *averaged system*

$$\dot{J} = \varepsilon F(J), \qquad F(J) = \frac{1}{(2\pi)^k} \oint_{\mathbb{T}^k} f(I, \varphi, 0) d\varphi. \tag{3}$$

This principle relies on the belief that the slow variable $I(t)$ and the solution $J(t)$ of the averaged system, with the same initial condition, should be close together over time $1/\varepsilon$. This principle does not always lead to the correct answer and the question of the correspondence between the solutions of both systems is far from being completely solved (for more details see [1] page 293 or [2] page 139).

This paper is organized as follows. In Section 2 we give a justification of the averaging principle for single-frequency systems. As an application we get the adiabatic invariants for Hamiltonian systems, with one degree of freedom. In Section 3 we describe the stroboscopic method and show how to use it to obtain adiabatic invariants. In Section 4 we give a justification of the averaging principle for systems with constant frequencies. In Section 5 we describe the averaging principle in the case where the perturbed system is a small Hamiltonian perturbation of an integrable system.

2 Averaging in Single-Frequency Systems

Consider the system of $(m+1)$-equations (2) where $k = 1$. Assume that the frequency $\omega(I)$ does not vanish. Assume that the solution $J(t)$ of the averaged system (3) is defined for $0 \leq t \leq 1/\varepsilon$.

Theorem 1 *Let $I(t,\varepsilon)$ be the slow variable of the perturbed system (2). Let $J(t)$ be the solution of the averaged system (3) with the same initial condition. Then $I(t,\varepsilon)$ remains close to $J(t)$ over time $1/\varepsilon$, that is to say, for any $\nu > 0$, there exists $\delta > 0$ such that $\|I(t,\varepsilon) - J(t)\| < \nu$ for $0 < \varepsilon < \delta$ and $0 \leq t \leq 1/\varepsilon$.*

Proof. Consider the change of variable

$$P = I + \varepsilon h(I,\varphi), \qquad h(P,\varphi) = \frac{1}{\omega(P)} \int_0^{\varphi} (F(P) - f(P,\phi,0))\, d\phi$$

This change of variable takes the perturbed system (2) into the averaged system to which a small perturbation of order ε^2 is added :

$$\dot{P} = \varepsilon F(P) + \varepsilon^2 u(P,\varphi,\varepsilon)$$

Over time $1/\varepsilon$, the solution $P(t,\varepsilon)$ of this system is close to the solution $J(t)$ of the averaged system (3). Since $P(t,\varepsilon)$ is close to $I(t,\varepsilon)$, the desired result follows (for more details see [1] page 294). ■

2.1 Adiabatic Invariants

Consider a Hamiltonian system with one degree of freedom. Assume that the Hamiltonian $H(p,q,\lambda)$ depends on some parameter λ. The equations of motion have the form

$$\dot{q} = \frac{\partial H}{\partial p}(p,q,\lambda), \qquad \dot{p} = -\frac{\partial H}{\partial q}(p,q,\lambda). \tag{4}$$

If λ is constant then the total energy $E = H(p(t),q(t),\lambda)$ is conserved and the phase trajectories are curves $C(E,\lambda)$ of equation $H(p,q,\lambda) = E$. We introduce action-angle variables for the Hamiltonian system (4) where λ is considered as a fixed parameter :

$$I = I(p,q,\lambda), \qquad \varphi = \varphi(p,q,\lambda) \bmod 2\pi.$$

The action variable is defined as follows. Assume that the phase trajectory $C(E,\lambda)$ is closed. Consider the closed phase trajectory passing through a point (p,q). It bounds some region in the phase plane. The area of this region is denoted by $2\pi I(p,q,\lambda)$. In the variables I, φ the Hamiltonian system (4) is described by the Hamiltonian $H(p,q,\lambda) = H(I,\lambda)$:

$$\dot{I} = 0, \qquad \dot{\varphi} = \frac{\partial H}{\partial I}(I,\lambda) \tag{5}$$

Assume that $\lambda = \varepsilon t$. System (4) becomes

$$\dot{q} = \frac{\partial H}{\partial p}(p, q, \varepsilon t), \qquad \dot{p} = -\frac{\partial H}{\partial q}(p, q, \varepsilon t). \tag{6}$$

Theorem 2 *If the frequency $\omega(I, \lambda) = \partial H/\partial I(I, \lambda)$ of system (5) is nowhere zero, then $I(p, q, \lambda)$ is an adiabatic invariant of the system (6), that is to say, $I(p(t), q(t), \varepsilon t)$ is nearly constant over time $1/\varepsilon$.*

Proof. In the variables I, φ the system with slowly varying Hamiltonian (6) is described by the Hamiltonian $H(I, \lambda) + \varepsilon K(I, \varphi, \lambda)$:

$$\dot{I} = -\varepsilon \frac{\partial K}{\partial \varphi}(I, \varphi, \lambda), \quad \dot{\lambda} = \varepsilon, \quad \dot{\varphi} = \frac{\partial H}{\partial I}(I, \lambda) + \varepsilon \frac{\partial K}{\partial I}(I, \varphi, \lambda) \tag{7}$$

where K is 2π-periodic in φ. Since $\omega(I, \lambda) \neq 0$, Theorem 1 is applicable. Since the mean value of the derivative $\partial K/\partial \varphi$ of a periodic function is equal to zero, the averaged system has the form $\dot{J} = 0$. Thus $I(t, \varepsilon)$ remains nearly constant over time $1/\varepsilon$. ■

Notice that the form (7) of system (6) is not obtained through straightforward calculations. It needs the Hamiltonian formalism of time dependent canonical transformations (for more details see [1] page 299).

2.2 The Lorentz pendulum

The small oscillations of a pendulum are described by the differential equation

$$\ddot{x} + \omega^2 x = 0, \qquad \omega^2 = g/l \tag{8}$$

where l is the length of the pendulum, g the constant of gravitation and x the angular deviation. The associated vector field to (8) in phase space $(q = x, p = \dot{x})$ is a Hamiltonian system with slowly varying Hamiltonian $H(p, q, \lambda) = \left(p^2 + \omega^2(\lambda) q^2\right)/2$. The phase trajectories are ellipses of equation $H(p, q, \lambda) = E$. The semi-axes of this ellipse are $a = \sqrt{2E}/\omega(\lambda)$ and $b = \sqrt{2E}$. The area of the region bounded by this ellipse is $\pi ab = 2\pi E/\omega(\lambda)$. Thus the action variable is $I = E/\omega(\lambda)$.

If the length l of the pendulum is slowly varying, that is, ω is a function of εt where ε is small, by Theorem 2, the ratio $E(t)/\omega(\varepsilon t)$ of the total energy $E = (\dot{x}^2 + \omega^2 x^2)/2$ to the frequency remains nearly constant over time $1/\varepsilon$, despite of the fact that $E(t)$ and $\omega(\varepsilon t)$ may change by quantities of order 1. This model was proposed by Einstein and Lorentz at the Soloway Congress in Bruxelles in 1911 to explain how in the Bohr atomic model, the ration of the energy to the frequency of an electron is constant, even if the electron moves in a varying electromagnetic field. Their explanation was based on the fact that the variation of the surrounding electromagnetic field is very slow compared to the high frequency of the electron. Shortly afterwards it appeared that quantum mechanics was more suitable to understand atomic behavior.

3 The stroboscopic method

The *stroboscopic method* is a method of the nonstandard perturbation theory of differential equations. It was proposed by J. L. Callot and G. Reeb (see [3, 5, 11, 10]). The principle of this method is as follows. Let $x(\tau)$ be a function. Suppose there exists a sequence of points $(\tau_n, x_n = x(\tau_n))$ such that $\tau_{n+1} \simeq \tau_n$, $x(\tau) \simeq x_n$ on the interval $[\tau_n, \tau_{n+1}]$ and

$$\frac{x_{n+1} - x_n}{\tau_{n+1} - \tau_n} \simeq f(\tau_n, x_n)$$

where f is a standard continuous function. Then the function $x(\tau)$ is infinitely close to a solution of the differential equation $x' = f(\tau, x)$.

There is another approach, based on the stroboscopic method, of the adiabatic invariance of the action variable in single frequency slowly varying systems. Let us explain it in the more general case of a non Hamiltonian perturbation of (4)

$$\begin{array}{ll} \dot{q} = \frac{\partial H}{\partial p}(p,q,\lambda) + \varepsilon f(p,q,\lambda), & q \in \mathbb{R} \\ \dot{p} = -\frac{\partial H}{\partial q}(p,q,\lambda) + \varepsilon g(p,q,\lambda), & p \in \mathbb{R} \\ \dot{\lambda} = \varepsilon h(p,q,\lambda), & \lambda \in \mathbb{R}^n. \end{array} \tag{9}$$

Let $\gamma(t,\varepsilon) = (q(t,\varepsilon), p(t,\varepsilon), \lambda(t,\varepsilon))$ be a solution of (9). The total energy

$$E(t,\varepsilon) = H(p(t,\varepsilon), q(t,\varepsilon), \lambda(t,\varepsilon)$$

is slowly varying :

$$\dot{E} = \varepsilon\Omega(p,q,\lambda), \qquad \Omega = \frac{\partial H}{\partial p}g + \frac{\partial H}{\partial q}f + \frac{\partial H}{\partial \lambda}h \tag{10}$$

Over time 1, the quantities $\lambda(t,\varepsilon)$ and $E(t,\varepsilon)$ remain nearly constant, so that the phase trajectory $\gamma(t,\varepsilon)$ passing through the point (q_0, p_0, λ_0) remains close to the curve $C(E_0, \lambda_0)$ defined by $H(p,q,\lambda_0) = E_0$ where $H(p_0,q_0,\lambda_0) = E_0$. The problem is to describe what happens over time $1/\varepsilon$. This question is answered in the following section.

3.1 Non Hamiltonian Perturbations

It is more natural to consider system (9) at the time scale $\tau = \varepsilon t$. Let $' = d/d\tau$ be the derivative with respect to the new time variable τ. The system (9,10) becomes

$$\begin{array}{l} q' = \frac{1}{\varepsilon}\frac{\partial H}{\partial p}(p,q,\lambda) + f(p,q,\lambda), \\ p' = -\frac{1}{\varepsilon}\frac{\partial H}{\partial q}(p,q,\lambda) + g(p,q,\lambda) \\ \lambda' = h(p,q,\lambda), \qquad E' = \Omega(p,q,\lambda). \end{array} \tag{11}$$

A region of oscillations of the Hamiltonian function $H(p,q,\lambda)$ is a domain $\mathcal{D} \subseteq \mathbb{R} \times \mathbb{R}^n$ such that for all $(E,\lambda) \in \mathcal{D}$, the equation $H(p,q,\lambda) = E$ defines a closed curve $C(E,\lambda)$ in the

plane (q,p) which does not contain any singular point where both derivatives $\partial H/\partial p$ and $\partial H/\partial q$ vanish. This closed curve corresponds to a periodic solution $(q(t,\lambda,E),p(t,\lambda,E))$ of (4) of energy E. Let $P(E,\lambda)$ be the period of this solution. We define the functions $G(E,\lambda)$ and $K(E,\lambda)$ on $\mathcal{D}$ by

$$G(E,\lambda)=\frac{R(E,\lambda)}{P(E,\lambda)},\quad R(E,\lambda)=\int_0^{P(E,\lambda)}\Omega(p(t,\lambda,E),q(t,\lambda,E),\lambda)dt,$$

$$K(E,\lambda)=\frac{S(E,\lambda)}{P(E,\lambda)},\quad S(E,\lambda)=\int_0^{P(E,\lambda)}h(p(t,\lambda,E),q(t,\lambda,E),\lambda)dt.$$

Theorem 3 *Let $\gamma(\tau,\varepsilon)=(q(\tau,\varepsilon),p(\tau,\varepsilon),\lambda(\tau,\varepsilon))$ be a solution of system (11) passing through the point (q_0,p_0,λ_0). Assume that (E_0,λ_0) where $E_0=H(p_0,q_0,\lambda_0)$, belongs to a region of oscillation $\mathcal{D}$ and that the closed curve $C(E_0,\lambda_0)$ contains the point (p_0,q_0). Let $E(\tau,\varepsilon)=H(p(\tau,\varepsilon),q(\tau,\varepsilon),\lambda(\tau,\varepsilon))$ be the total energy of $\gamma(\tau,\varepsilon)$. The functions $E(\tau,\varepsilon)$ and $\lambda(\tau,\varepsilon)$ are approximated by the solutions $E_0(\tau)$, $\lambda_0(\tau)$ of the averaged system*

$$E'=G(E,\lambda),\qquad \lambda'=K(E,\lambda) \tag{12}$$

with initial condition (E_0,λ_0) as long as τ is limited and $(E_0(\tau),\lambda_0(\tau))$ is limited and takes values in $\mathcal{D}$.

Proof. Let τ_0 be fixed. Let $p_0=p(\tau_0,\varepsilon)$, $q_0=q(\tau_0,\varepsilon)$, $E_0=E(\tau_0,\varepsilon)$ and $\lambda_0=\lambda(\tau_0,\varepsilon)$. Consider the change of variables

$$r=\frac{\tau-\tau_0}{\varepsilon},\quad F(r,\varepsilon)=\frac{E(\tau_0+\varepsilon r,\varepsilon)-E_0}{\varepsilon},\quad \Lambda(r,\varepsilon)=\frac{\lambda(\tau_0+\varepsilon r)-\lambda_0}{\varepsilon}.$$

The system (11) is taken into the system

$$\begin{aligned}
&dq/dr=\tfrac{\partial H}{\partial p}(p,q,\lambda_0+\varepsilon\Lambda)+\varepsilon f(p,q,\lambda_0+\varepsilon\Lambda),\\
&dp/dr=-\tfrac{\partial H}{\partial q}(p,q,\lambda_0+\varepsilon\Lambda)+\varepsilon g(p,q,\lambda_0+\varepsilon\Lambda)\\
&d\Lambda/dr=h(p,q,\lambda_0+\varepsilon\Lambda),\qquad dF/dr=\Omega(p,q,\lambda_0+\varepsilon\Lambda).
\end{aligned} \tag{13}$$

By continuous dependence of the solutions with respect to the parameters we get

$$q(r,\varepsilon)\simeq q(r,\lambda_0,E_0),\qquad p(r,\varepsilon)\simeq p(r,\lambda_0,E_0), \tag{14}$$

where $(q(r,\lambda_0,E_0),p(r,\lambda_0,E_0))$ is the solution of the unperturbed system (4) with $\lambda=\lambda_0$, of total energy E_0, and starting from (q_0,p_0). From (13) and (14) we have

$$F(r,\varepsilon)\simeq\int_0^r\Omega(p(t,\lambda_0,E_0),q(t,\lambda_0,E_0),\lambda_0)dt,$$

$$\Lambda(r,\varepsilon)\simeq\int_0^r h(p(t,\lambda_0,E_0),q(t,\lambda_0,E_0),\lambda_0)dt.$$

Let $r = P(E_0, \lambda_0)$, $\tau_1 = \tau_0 + \varepsilon r$, $E_1 = E(\tau_1, \varepsilon)$ and $\lambda_1 = \lambda(\tau_1, \varepsilon)$. We have

$$\frac{E_1 - E_0}{\tau_1 - \tau_0} = \frac{F(r,\varepsilon)}{r} \simeq G(E_0, \lambda_0), \qquad \frac{\lambda_1 - \lambda_0}{\tau_1 - \tau_0} = \frac{\Lambda(r,\varepsilon)}{r} \simeq K(E_0, \lambda_0)$$

Using the stroboscopic method we conclude that $E(t,\varepsilon)$ and $\lambda(t,\varepsilon)$ are close to the solution of the averaged system (12) with the same initial condition. ■

3.2 Applications

Let us give in this section some applications of Theorem 3. This theorem applies in many other situations. See [9] for applications to the study of some second, third and n order singularly perturbed differential equations.

Example 1. For the Lorentz Pendulum we have $f = g = 0$, $h = 1$ and $H(p,q,\lambda) = \left(p^2 + \omega^2(\lambda) q^2\right)/2$. Hence the averaged system (12) takes the form :

$$E' = E\omega'(\tau)/\omega(\tau).$$

Hence $E(\tau,\varepsilon) \simeq E(0,\varepsilon)\omega(\tau)/\omega(0)$, that is to say, the ratio $E(\tau,\varepsilon)/\omega(\tau)$ remains nearly constant over time 1.

Example 2. Consider the second order differential equation

$$\ddot{x} + f(x, \varepsilon t) = 0$$

or equivalently at time scale τ

$$\varepsilon^2 x'' + f(x,\tau) = 0$$

The change of variables $q = x$, $p = \varepsilon x'$ transforms this equation into the system

$$q' = p/\varepsilon, \qquad p' = -f(q,\tau)/\varepsilon$$

which is of the form (11) with Hamiltonian $H(p,q,\lambda) = p^2/2 + U(q,\lambda)$, where $\partial U/\partial q = f$. Hence the averaged system (12) takes the form

$$E' = G(E,\tau), \qquad G(E,\tau) = \frac{\int_{q_1(E,\tau)}^{q_2(E,\tau)} \frac{\frac{\partial U}{\partial \tau}(q,\tau)\,dq}{\sqrt{E - U(q,\tau)}}}{\int_{q_1(E,\tau)}^{q_2(E,\tau)} \frac{dq}{\sqrt{E - U(q,\tau)}}}$$

where $q_1(E,\tau)$ and $q_2(E,\tau)$ are the minimum and the maximum of an oscillation. There are the solutions of equation $U(q,\tau) = E$. From Theorem 3 we see that the energy $E(\tau,\varepsilon)$ remains close to the corresponding solution $E_0(\tau)$ of the averaged system.

Example 3. Consider the third order differential equation

$$\varepsilon^2 d^3x/d\tau^3 + f(\tau, x, dx/d\tau) = 0.$$

The change of variables $q = dx/d\tau$, $p = \varepsilon d^2x/d\tau^2$, $\lambda = (\tau, x)$ transforms this equation into the system

$$q' = p/\varepsilon, \qquad p' = -f(\lambda, q, \tau)/\varepsilon, \qquad \lambda' = (1, q)$$

which is of the form (11) with Hamiltonian $H(p, q, \lambda) = p^2/2 + U(\lambda, q)$, where $\partial U/\partial q = f$. Hence the averaged system (12) takes the form

$$E' = G(E, \lambda), \qquad G(E, \lambda) = \frac{\int_{q_1(E,\lambda)}^{q_2(E,\lambda)} \frac{\left(\frac{\partial U}{\partial \tau}(\lambda, q) + q\frac{\partial U}{\partial \lambda}(\lambda, q)\right) dq}{\sqrt{E - U(\lambda, q)}}}{\int_{q_1(E,\lambda)}^{q_2(E,\lambda)} \frac{dq}{\sqrt{E - U(\lambda, q)}}}$$

$$\lambda' = K(E, \lambda), \qquad K(E, \lambda) = \frac{\int_{q_1(E,\lambda)}^{q_2(E,\lambda)} \frac{q\, dq}{\sqrt{E - U(\lambda, q)}}}{\int_{q_1(E,\lambda)}^{q_2(E,\lambda)} \frac{dq}{\sqrt{E - U(\lambda, q)}}}$$

where $q_1(E, \lambda)$ and $q_2(E, \lambda)$ are the minimum and the maximum of an oscillation. There are the solutions of equation $U(\lambda, q) = E$. From Theorem 3 we see that the energy $E(\tau, \varepsilon)$ and $\lambda(\tau, \varepsilon) = (\tau, x(\tau, \varepsilon))$ remain close to the corresponding solution $E_0(\tau)$, $\lambda_0(\tau)$ of the averaged system.

Example 4. The third order differential equation

$$\varepsilon^2 d^3x/d\tau^3 + dx/d\tau = 1 - x^2 \tag{15}$$

was considered in [4] to explain the behavior of the so-called ghost solutions appearing in numerical schemes. The change of variables $q = x$, $p = \varepsilon dx/d\tau$, $\lambda = x + \varepsilon^2 d^2x/d\tau^2$ transforms equation (15) into the system

$$q' = p/\varepsilon, \quad p' = (\lambda - q)/\varepsilon, \quad \lambda' = 1 - q^2$$

which is of the form (11) with Hamiltonian $H(p, q, \lambda) = \left(p^2 + (\lambda - q)^2\right)/2$. Hence the averaged system (12) takes the form

$$E' = 2E\lambda, \qquad \lambda' = 1 - E - \lambda^2$$

From Theorem 3 we see that a solution $x(t, \varepsilon)$ of (15) is rapidly oscillating and satisfies

$$\lambda_0(\tau) - \sqrt{2E_0(\tau)} \le x(\tau, \varepsilon) \le \lambda_0(\tau) - \sqrt{2E_0(\tau)}$$

where $E_0(\tau)$ and $\lambda_0(\tau)$ are the corresponding solutions of the averaged system.

3.3 The evolution of the action variable

Let I, φ be the action-angle variables for the Hamiltonian system (4) where λ is considered as a fixed parameter. In the variables I, φ the system (11) is equivalent to a system of the form

$$\dot{I} = \varepsilon f_1(I, \varphi, \lambda), \quad \dot{\lambda} = \varepsilon h_1(I, \varphi, \varepsilon), \quad \dot{\varphi} = \omega(I, \lambda) + \varepsilon g_1(I, \varphi, \lambda)$$

This system is of the form of system (2) to which Theorem 1 applies. Here, I do not know how to use the Hamiltonian formalism of canonical transformation to determine the functions f_1, g_1 and h_1, as in the proof of Theorem 2. Thus, I cannot compute the averages of the functions f_1 and g_1 and apply Theorem 1 to obtain the evolution of I. However, from Theorem 3 we can obtain the evolution of I. Let $A(E,\lambda)$ be the area of the region bounded by the closed curve $C(E,\lambda)$:

$$A(E,\lambda) = \oint_{C(E,\lambda)} pdq$$

Recall that $I = A/(2\pi)$ is the action variable for the Hamiltonian system (4) where λ is a fixed parameter.

Theorem 4 *Let $\gamma(\tau,\varepsilon) = (q(\tau,\varepsilon), p(\tau,\varepsilon), \lambda(\tau,\varepsilon))$ be a solution of system (11). Let $E_0(\tau)$, $\lambda_0(\tau)$ of the corresponding solution of the averaged system (12) described in Theorem 3. Let $a(\tau,\varepsilon) = A(E(\tau,\varepsilon),\lambda(\tau,\varepsilon))$. Then*

$$a(\tau,\varepsilon) \simeq a(0) + \int_0^\tau L(E_0(s),\lambda_0(s))ds$$

where $L(E,\lambda) = R(E,\lambda) - Q(E,\lambda)K(E,\lambda)$ and

$$Q(E,\lambda) = \int_0^{P(E,\lambda)} \frac{\partial H}{\partial \lambda}(p(t,\lambda,E), q(t,\lambda,E), \lambda)dt.$$

Proof. A straightforward computation shows that $\partial A/\partial E = P$ and $\partial A/\partial \lambda = -Q$. Hence

$$\frac{da}{d\tau}(\tau,\varepsilon) = P(E(\tau,\varepsilon),\lambda(\tau,\varepsilon))\frac{dE}{d\tau}(\tau,\varepsilon) - Q(E(t,\varepsilon),\lambda(\tau,\varepsilon))\frac{d\lambda}{d\tau}(\tau,\varepsilon).$$

By Theorem 3 we have $da/d\tau(\tau,\varepsilon) \simeq L(E_0(\tau),\lambda_0(\tau))$. ■

As a consequence of this result we get the adiabatic invariance of the action variable in single frequency slowly varying Hamiltonian systems obtained in Theorem 2.

Theorem 5 *Let $\gamma(t,\varepsilon) = (q(t,\varepsilon), p(t,\varepsilon))$ be a solution of system (6) where $\lambda = \varepsilon t$. Let $a(t,\varepsilon) = A(E(t,\varepsilon),\varepsilon t)$. Then $a(t,\varepsilon)$ remains nearly constant over time $1/\varepsilon$.*

Proof. In the particular case of system (6), we have $g = f = 0$ and $h = 1$. Hence $L = 0$. By Theorem 4, $a(\tau,\varepsilon)$ remains nearly constant over time 1. ■

3.4 Single-frequency systems

The stroboscopic method allows also another approach of Theorem 1 (for the details see [3]). With the new time variable $\tau = \varepsilon t$, the perturbed system (2) becomes

$$I' = f(I,\varphi,\varepsilon), \qquad \varphi' = \omega(I)/\varepsilon + g(I,\varphi,\varepsilon), \tag{16}$$

and the averaged system (3) becomes

$$J' = F(J) \tag{17}$$

Let $(I(\tau,\varepsilon),\varphi(\tau,\varepsilon))$ be a solution of (16). Let τ_0 be fixed. Let $I_0 = I(\tau_0,\varepsilon)$ and $\varphi_0 = \varphi(\tau_0,\varepsilon)$. Consider the change of variables

$$r = \frac{\tau-\tau_0}{\varepsilon}, \quad J(r,\varepsilon) = \frac{I(\tau_0+\varepsilon r,\varepsilon)-I_0}{\varepsilon}, \quad \psi(r,\varepsilon) = \varphi(\tau_0+\varepsilon r,\varepsilon)-\varphi_0$$

The perturbed system (16) is taken into the system

$$\begin{array}{ll} dJ/dr = f(I_0,\varphi_0+\psi,0)+\varepsilon f_1(J,\psi,\varepsilon), & J(0)=0, \\ d\psi/dr = \omega(I_0)+\varepsilon g_1(J,\psi,\varepsilon), & \psi(0)=0. \end{array} \tag{18}$$

By continuous dependence of the solutions we get

$$J(r) \simeq \int_0^r f(I_0,\varphi_0+\omega(I_0)s,0)ds = \frac{1}{\omega(I_0)}\int_{\varphi_0}^{\varphi_0+\omega(I_0)r} f(I_0,\psi,0)d\psi$$

Let $r = 2\pi/\omega(I_0)$, $\tau_1 = \tau_0+\varepsilon r$ and $I_1 = I(\tau_1,\varepsilon)$. We have

$$\frac{I_1-I_0}{\tau_1-\tau_0} = \frac{J(r,\varepsilon)}{r} \simeq \frac{1}{2\pi}\int_{\varphi_0}^{\varphi_0+2\pi} f(I_0,\varphi_0+\psi,0)ds = F(I_0)$$

Using the stroboscopic method we conclude that $I(\tau,\varepsilon)$ is close to the solution of the averaged system (17) with the same initial condition.

4 Systems with Constant Frequencies

Consider a system with constant frequencies

$$\dot{I} = \varepsilon f(I,\varphi,\varepsilon), \qquad \dot{\varphi} = \omega+\varepsilon g(I,\varphi,\varepsilon), \tag{19}$$

where the frequency vector ω does not depend on the slow variable I. The change of variable $\theta = \varphi-\omega t$ transforms (19) into

$$\dot{I} = \varepsilon f(I,\theta+\omega t,\varepsilon), \qquad \dot{\xi} = \varepsilon g(I,\theta+\omega t,\varepsilon). \tag{20}$$

Let $x = (I,\theta)$. System (20) becomes

$$\dot{x} = \varepsilon A(t,x,\varepsilon), \qquad x \in \mathbb{R}^{m+k}, \tag{21}$$

where $A(t,x,\varepsilon) = (f(I,\theta+\omega t,\varepsilon), g(I,\theta+\omega t,\varepsilon))$. Since f and g are 2π-periodic in φ, the function A is quasi-periodic in t. Hence system (21) is a *Krylov-Bogoliobov-Mitropolski* (KBM) system.

4.1 KBM Theory

Assume that $A(t,x,0)$ is 2π-periodic in t. Let

$$A_0(x) = \frac{1}{2\pi}\int_0^{2\pi} A(t,x,0)dt.$$

The aim of KBM theory is to approximate the solutions of system (21) by the solutions of the averaged system

$$\dot{x} = \varepsilon A_0(x), \tag{22}$$

Assume that (22) has a unique solution with prescribed initial condition.

Theorem 6 *Let $x(t,\varepsilon)$ be a solution of (21). Let $x_0(t)$ be the solution of (22) with initial condition $x(0,0)$. Assume that $x_0(t)$ is defined on the interval $[0,1/\varepsilon]$. Then $x(t,\varepsilon)$ is defined at least on the same interval and satisfies*

$$x(t,\varepsilon) \simeq x_0(t), \qquad 0 \le t \le 1/\varepsilon.$$

Proof. It is more natural to consider systems (21) and (22) at the time scale $\tau = \varepsilon t$. Let $' = d/d\tau$ be the derivative with respect to the new time variable τ. The system (21) becomes

$$x' = A(\tau/\varepsilon, x, \varepsilon), \tag{23}$$

and the averaged system (22) becomes

$$x' = A_0(x) \tag{24}$$

Let $x(\tau,\varepsilon)$ be a solution of (23). Let τ_0 be fixed. Let $x_0 = x(\tau_0,\varepsilon)$. Consider the change of variables

$$r = \frac{\tau - \tau_0}{\varepsilon}, \quad X(r,\varepsilon) = \frac{x(\tau_0 + \varepsilon r, \varepsilon) - x_0}{\varepsilon}.$$

The system (23) is taken into the system

$$\frac{dX}{dr} = A(s + r, x_0 + \varepsilon X, \varepsilon) \tag{25}$$

where $s = \tau_0/\varepsilon$. By continuous dependence of the solutions we get

$$X(r,\varepsilon) \simeq \int_0^r A(s+r, x_0, 0)dr = \int_s^{s+r} A(t, x_0, 0)dt$$

Let $r = 2\pi$, $\tau_1 = \tau_0 + \varepsilon r$ and $x_1 = x(\tau_1,\varepsilon)$. We have

$$\frac{x_1 - x_0}{\tau_1 - \tau_0} = \frac{X(r,\varepsilon)}{r} \simeq \frac{1}{2\pi}\int_s^{s+2\pi} A(t, x_0, 0)dt = A_0(x_0)$$

Using the stroboscopic method we conclude that $x(\tau,\varepsilon)$ is close to $x_0(\tau)$. ■

Theorem 6 was extended to quasi periodic vector fields and KBM vector fields (see [7] and the references therein). In these cases the average $A_0(x)$ is obtained as follows. If $A(t,x,0)$ is quasi-periodic in t then the limit

$$A_0(x) = \lim_{T\to+\infty} \frac{1}{T}\int_s^{s+T} A(t,x,0)dt$$

exists and is uniform with respect to s. If A is a so-called KBM vector fields, then the limit

$$A_0(x) = \lim_{T\to+\infty} \frac{1}{T}\int_0^{T} A(t,x,0)dt$$

exists. The stroboscopic method allows us to obtain also the KBM theorem of averaging in these more general cases (for more details see [11]).

Theorem 7 *Assume that the function $A(t,x,0)$ is continuous in x uniformly with respect to t and that the average $A_0(x)$ exists. Let $x(\tau,\varepsilon)$ be a solution of (23). Let $x_0(\tau)$ be the solution of (24) with initial condition $x(0,0)$. Assume that $x_0(\tau)$ is defined on the interval $[0,1]$. Then $x(\tau,\varepsilon)$ is defined at least on this interval and is approximated by $x_0(\tau)$.*

In the classical formulation of the previous theorem (see [7]) it is usually required that the function $A(t,x,\varepsilon)$ is Lipshitz continuous in x. In our formulation this condition has been weakened : only the continuity in x uniformly with respect to t is required. The differential equation

$$x' = \sin(\tau x/\varepsilon) \tag{26}$$

provides an example for which the conclusion of Theorem 7 does not hold. The average of $A(t,x) = \sin(tx)$ is $A_0(x) = 0$. But the function $A(t,x)$ is not continuous in x uniformly with respect to t. The study of (26) was at the origin of the birth of the stroboscopic method (see [3, 10]). Actually, we prove [3, 11] that the solutions of (26) are approximated on the subset $U = \{(x,\tau) \in \mathbb{R}^2 : x > \tau > 0\}$ by the solutions of the averaged differential equation

$$x' = \frac{\sqrt{x^2-\tau^2}-x}{\tau}.$$

4.2 Precession of the planet Mercury

The Einstein equation for the planet Mercury is

$$\frac{d^2u}{d\theta^2} + u = a + \varepsilon u^2, \tag{27}$$

where $u = 1/r$ and r, θ are the polar coordinates of Mercury in the plane centered at the Sun. Here

$$a = \frac{GM\overline{r}}{h^2} \simeq 0.98, \qquad \varepsilon = \frac{3GM}{c^2\overline{r}} \simeq 8.16\times 10^{-8},$$

where G is the Newtonian gravitational constant, M the mass of the Sun, h the angular momentum of Mercury, $\overline{r}$ the mean value of r and c the speed of the light in vacuum. For $\varepsilon = 0$, we have the Newtonian model : the trajectory is an ellipse $u = a + A\sin\theta$. The usual perturbation theory predicts that, under the gravitational effects of the other planets, this orbit should rotate slowly in space, so that the major semi-axis of the ellipse should advance by about 529 seconds of arc per century. The precession is actually observed to be of 572 seconds of arc per century. The difference is explained by the general relativity theory.

Let $x = u - a$, and $t = \theta$, then equation (27) is equivalent to

$$\ddot{x} + x = \varepsilon(x + a)^2$$

which is a particular case of an equation of type

$$\ddot{x} + x = \varepsilon f(x, \dot{x}) \tag{28}$$

The vector field associated to equation (28) is

$$\dot{x} = y, \qquad \dot{y} = -x + \varepsilon f(x, y)$$

The Van der Pol change of variables

$$x = A\sin(t + B), \qquad y = A\cos(t + B)$$

transforms this system into

$$\begin{aligned} \dot{A} &= \varepsilon f(A\sin(t+B), A\cos(t+B))\cos(t+B) \\ \dot{B} &= \tfrac{\varepsilon}{A} f(A\sin(t+B), A\cos(t+B))\sin(t+B) \end{aligned}$$

By Theorem 6 the solutions of this system are approximated by the solutions of the averaged system. For $f(x, y) = (x + a)^2$, the averaged system is

$$\dot{A} = 0, \qquad \dot{B} = a$$

Hence the amplitude A remains nearly constant (there is no secular variation of the major semi-axis of Mercury due to gravitational effects of general relativity) but the phase is slowly varying (there is a precession). More precisely we have $A(t) \simeq A$ and $B(t) \simeq \varepsilon a t$ thus

$$u(t) \simeq a + A\sin((1 + \varepsilon a)t)$$

Hence the precession due to the gravitational effects of general relativity is of

$$2\pi\varepsilon a \text{ radians} = 3600 \times 360\ \varepsilon a \text{ second of arc}$$

per orbit of Mercury. Since the period of Mercury is 88 days, the precession of Mercury due to the relativity theory is of

$$\frac{3600 \times 360 \times (8.16) \times 10^{-8} \times (0.98) \times 100 \times 365}{88} \simeq 43$$

seconds of arc per century. For more details on this subject the reader is referred to [6] page 39, [12] page 29 or [13] page 144.

5 Averaging in Hamiltonian Systems

Consider the case where $k = m = n$ and the perturbed system (2) is a small Hamiltonian perturbation of a completely integrable Hamiltonian system with n degrees of freedom. The Hamiltonian of the perturbed system takes the form

$$H(I, \varphi) = H_0(I) + \varepsilon H_1(I, \varphi, \varepsilon).$$

We have assumed that some domain in phase space is foliated by invariant tori, and that action-angle variables $I = (I_1, \cdots, I_n) \in B \subset \mathbb{R}^n$, $\varphi = (\varphi_1, \cdots, \varphi_n) \in \mathbb{T}^n$ for the unperturbed system are introduced in this domain. The Hamiltonian $H_0(I)$ of the unperturbed system depends only on the action variables I. The equations governing the unperturbed motion have the form :

$$\dot{I} = 0, \qquad \dot{\varphi} = \frac{\partial H_0}{\partial I}(I), \tag{29}$$

The equations governing the perturbed motion have the form :

$$\dot{I} = -\varepsilon \frac{\partial H_1}{\partial \varphi}(I, \varphi, \varepsilon), \qquad \dot{\varphi} = \frac{\partial H_0}{\partial I}(I) + \varepsilon \frac{\partial H_1}{\partial I}(I, \varphi, \varepsilon), \tag{30}$$

Since

$$\oint_{\mathbb{T}^n} \partial H_1 / \partial \varphi(I, \varphi, 0) d\varphi = 0,$$

the averaged system is $\dot{J} = 0$. Using the averaging principle we assert that there is no evolution in the slow variables $I(t)$ of system (30) over time $1/\varepsilon$. A mathematical problem central to Hamiltonian system theory is to establish this result for the perturbed system. Historically this problem gave rise, with the work of Laplace and Lagrange, to the first formulation of the averaging principle and the birth of the perturbation theory. Considerable progress in this problem has been achieved in the framework of the Kolmogorov-Arnold-Moser (KAM) (for more details see [2] page 171).

5.1 KAM Theory

For the unperturbed Hamiltonian system (29) the phase space is foliated by invariant tori I =const. Each torus carries on a conditionally-periodic motion with constant frequencies $\omega(I) = \partial H_0 / \partial I(I)$. A torus on which the frequencies $\omega(I) = (\omega_1(I), \cdots, \omega_n(I))$ are rationally independent is filled densely by phase trajectories. Such a torus is said to be *non-resonant*. The remaining tori I =const on which the frequencies are rationally dependent are said to be *resonant*. Each phase trajectory on such a torus fills densely a torus of smaller dimension.

Consider now the perturbed system (30). Under a supplementary condition on the frequencies, the theorem of Kolmogorov ([2] page 183) asserts that, for sufficiently small Hamiltonian perturbations, most non-resonant invariant tori do not vanish but are only slightly deformed. In the phase space of the perturbed system there are invariant tori

densely filled with phase curves. These tori are called the *Kolmogorov tori*. Their union is the *Kolmogorov set*. The measure of the complement of the Kolmogorov set does not exceed a quantity of order $\sqrt{\varepsilon}$.

5.2 Arnold diffusion

In systems with two degrees of freedom, the phase space is four-dimensional and the energy level manifolds are three-dimensional. Since the Kolmogorov tori are two-dimensional, a phase curve starting in a gap between two Kolmogorov tori remains forever trapped between them. Thus, the slow variables remain forever nearly constant, and the averaging principle leads to the correct answer.

If the number n of degrees of freedom is greater than two, the n-dimensional Kolmogorov tori do not divide the $(2n-1)$-dimensional energy level manifolds, so that the "gaps" between Kolmogorov tori are connected to one another. Arnold conjectured that through an arbitrary small neighborhood of any point there pass phase trajectories along which the slow variables I do not remain nearly constant. Numerical examples show that the evolution of I is not directional but represents a more or less random wandering in the complement of the Kolmogorov set. This behavior is known as *Arnold's diffusion* (for more details see [2] page 189).

References

[1] V.I. Arnold, *Mathematical Methods of Classical Mechanics*, Graduate Texts in Math. 60, Springer-Verlag, New-York 1989, **MR 90c:58046**.

[2] V.I. Arnold, V. V. Kozlov, A. I. Neishtadt, *Dynamical Systems III*, Encyclopedia of Mathematical Sciences, Vol. 3, Springer-Verlag, Berlin 1993, **MR 95d:58043b**.

[3] J. L. Callot, T. Sari, Stroboscopie et moyennisation dans les systèmes d'équations diférentielles à solutions rapidement oscillantes, in : *Mathematical Tools and Models for Control, Systems Analysis and Signal Processing, vol. 3*, CNRS Paris, 1983, 345-353, **MR 783848**.

[4] S. N. Chow, E. M. de Jager, R. Lutz, The ghost solutions of the logistic equation and a singular perturbation problem, in *Advances in Computational Methods for Boundary Layers and Interior Layers*, Dublin (1984), *Adv. Numer. Comput. Ser.*, 6 (1984), 15-20, **MR 86d:65085**.

[5] R. Lutz, L'intrusion de l'analyse non standard dans l'étude des perturbations singulières *Astérisque* **109-110** (1983), 101-140, **MR 86d:58105**.

[6] G. Pascoli, *La gravitation*, Presses Universitaires de France, Coll. Que Sais-je ?, no. 2489, 1989.

[7] J. A. Sanders, F. Verhulst, *Averaging Methods in Nonlinear Dynamical Systems*, Appl. Math. Sciences 58, Springer-Verlag, New-York 1985, **MR 87d:34065**.

[8] T. Sari, Sur la théorie asymptotique des oscillations non stationaires, *Astérisque* **109-110** (1983), 141-158, **MR 86c:34109**.

[9] T. Sari, Systèmes hamiltoniens à paramètres lentement variables, *Cahiers Math. Univ. Oran* **3** (1987), 113-131, **MR 89f:70024**.

[10] T. Sari, Petite histoire de la stroboscopie, in :*Colloque Trajectorien à la Mémoire de J. L. Callot et G. Reeb*, Strasbourg-Obernai 1995, Publication IRMA, Univ. Strasbourg (1995), 5-15, **MR 96k:34118**.

[11] T. Sari, Stroboscopy and Averaging, in :*Colloque Trajectorien à la Mémoire de J. L. Callot et G. Reeb*, Strasbourg-Obernai 1995, Publication IRMA, Univ. Strasbourg (1995), 125-158, **MR 97c:34091**.

[12] D. R. Smith, *Singular-perturbation Theory. An introduction with applications*, Cambridge University Press, Cambridge, 1985, **MR 87d:34001**.

[13] S. Sternberg, *Celestial Mechanics, Part 1*, W. A. Benjamin Inc, New York, 1969.

Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex (France)

Unité de recherche INRIA Lorraine : LORIA, Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex (France)
Unité de recherche INRIA Rennes : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex (France)
Unité de recherche INRIA Rhône-Alpes : 655, avenue de l'Europe - 38330 Montbonnot-St-Martin (France)
Unité de recherche INRIA Rocquencourt : Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex (France)

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, BP 105 - 78153 Le Chesnay Cedex (France)
http://www.inria.fr
ISSN 0249-6399