YURI KIFER

Averaging in difference equations driven by dynamical systems

Astérisque, tome 287 (2003), p. 103-123

<http://www.numdam.org/item?id=AST_2003__287__103_0>

Astérisque
287, 2003, p. 103–123

AVERAGING IN DIFFERENCE EQUATIONS DRIVEN BY DYNAMICAL SYSTEMS

by

Yuri Kifer

Dedicated to Jacob Palis for his sixtieth birthday

Abstract. — The averaging setup arises in the study of perturbations of parametric families of dynamical systems when parameters start changing slowly in time. Usually, averaging methods are applied to systems of differential equations which combine slow and fast motions. This paper deals with difference equations case which leads to wider class of models and examples. The averaging principle is justified here under a general condition which is verified when unperturbed transformations either preserve smooth measures or they are hyperbolic. The convergence speed in the averaging principle is estimated for some cases, as well.

1. Introduction

In the study of evolution of many real systems we can usually observe only few parameters while other less significant ones are regarded as constant in time. A more precise investigation may reveal that these parameters change, as well, but much slower than the others. These leads to complicated double scale equations describing slow and fast motions which are difficult to solve directly. Such problems were encountered with already long ago in celestial mechanics in the study of perturbations of planetary motion. People noticed that good approximations of the slow motion on long time intervals can be obtained by averaging coefficients of its equation in fast variables. This averaging principle was applied in celestial mechanics long before it was rigourously justified in some cases in the middle of the 20th century (see [**18**] and historical remarks there).

Traditionally, averaging methods were employed in the study of two scale ordinary differential equations describing a continuous time motion. On the other hand, it is well known that the study of discrete time dynamical systems, i.e. of iterates of

2000 ***Mathematics Subject Classification.*** — Primary: 34C05; Secondary: 39A11, 60J05, 37D20.
Key words and phrases. — Averaging, difference equations, dynamical systems.

transformations (not necessarily invertible), enables us to deal with a wider class of models and examples and to reveal new effects. Suppose that an idealized physical system can be described by a transformation F_0 of a $(d+m)$-dimensional space and there exist functions $x_1, \dots, x_d$ which do not change along orbits of F_0 (integrals of motion). Then, generically, F_0 can be written as a transformation of a locally trivial fiber bundle $\mathcal{M} = \{(x,y) : x \in \mathbb{R}^d,\, y \in M_x\}$ with base $\mathbb{R}^d$ and fibers M_x being m-dimensional manifolds acting by the formula $F_0(x,y) = (x, f_x y)$ where $f_x = f(x,\cdot) : M_x \to M_x$ is a transformation of M_x. It is natural to view a real physical system as a perturbation of the above idealized one, and so it should be described by a transformation

$$F_\varepsilon(x,y) = (x + \varepsilon\Phi(x,y,\varepsilon), f(x,y,\varepsilon)) \tag{1.1}$$

where $\Phi(\cdot,\cdot,\varepsilon) : \mathcal{M} \to \mathbb{R}^d$ and $f(x,\cdot,\varepsilon) : M_x \to M_x$. Since locally $\mathcal{M}$ has a product structure $U \times M$, where U is an open subset of $\mathbb{R}^d$ and M is an m-dimensional manifold, and iterates $F_\varepsilon^n(x,y)$ of any point (x,y) in $U \times M$ stay there for all $n \leqslant \delta/\varepsilon$ with small but fixed $\delta = \delta(x) > 0$ we conclude that it suffices to study the evolution on time intervals of order $1/\varepsilon$ only on product spaces and then glue pieces of orbits together.

In this paper we consider difference equations of the form

$$\begin{aligned} X^\varepsilon(n+1) - X^\varepsilon(n) &= \varepsilon\Phi(X^\varepsilon(n), Y^\varepsilon(n), \varepsilon), \quad && X^\varepsilon(0) = x, \\ Y^\varepsilon(n+1) &= f(X^\varepsilon(n), Y^\varepsilon(n), \varepsilon), && Y^\varepsilon(0) = y \end{aligned} \tag{1.2}$$

where $X^\varepsilon(n) = X^\varepsilon_{x,y}(n) \in \mathbb{R}^d$, $Y^\varepsilon(n) = Y^\varepsilon_{x,y}(n)$ runs on a compact m-dimensional Riemannian manifold M, $\Phi = \Phi(x,y,\varepsilon)$ is a Lipschitz in x, y, ε vector function, $f_x(\cdot,\varepsilon) = f(x,\cdot,\varepsilon)$ is a family of smooth maps (usually, endomorphisms or diffeomorphisms) of M close to f_x. Thus $(X^\varepsilon_{x,y}(n), Y^\varepsilon_{x,y}(n)) = F^n_\varepsilon(x,y)$. The equations (1.2) usually cannot be solved explicitly and it is desirable to approximate its solutions for small ε. Returning back to the unperturbed $\varepsilon = 0$ case eliminates the slow motion X^ε completely and gives a rather pure approximation valid only for bounded time intervals. The averaging principle is supposed to give a prescription how to approximate the slow motion X^ε on time intervals of order $1/\varepsilon$. Recurrent relations (1.2) can be regarded as a more general than usual setup for perturbations of dynamical systems where not only the transformation itself is perturbed but also we begin to take into account evolution of some parameters whose change was disregarded before.

We note that the standard continuous time averaging setup (see [**13**]) can be always reduced by discretizing time to a model described by difference equations of type (1.2). On the other hand, an attempt to go the other way around faces substantial difficulties since the standard suspension construction should be implemented now for different transformations f_x and it is not clear how to glue everything together in an appropriate way. Observe, that (1.2) can be generalized adding some randomness in the right hand sides there so that $f_x(\cdot,\varepsilon)$ become random endomorphisms, but we will not discuss this setup here.

Assume, first, that the fast motion $Y^\varepsilon(n)$ is independent of the slow variables, i.e. $f(x,y,\varepsilon)=fy$, and so $Y^\varepsilon_{x,y}(n)=f^ny$. For an ergodic f-invariant probability measure μ the limit

$$\overline{\Phi}_\mu(x)=\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\Phi(x,f^ny)=\int\Phi(x,y)d\mu(y) \tag{1.3}$$

exists for μ-almost all y. For such y's uniformly in n the solution $X^\varepsilon_{x,y}$ of (1.2) is close on any time interval of order $1/\varepsilon$ to the solution $\overline{X}^\varepsilon=\overline{X}^\varepsilon_\mu=\overline{X}^\varepsilon_{x,\mu}$, taken at integer times, of the differential equation

$$\frac{d\overline{X}^\varepsilon(t)}{dt}=\varepsilon\overline{\Phi}(\overline{X}^\varepsilon(t)),\quad \overline{X}^\varepsilon(0)=x \tag{1.4}$$

where $\overline{\Phi}=\overline{\Phi}_\mu$ (see similar continuous time results in [**18**]). Already in this case the averaging principle works only for μ-almost all initial points y and for different y's averaged solutions may be different. In the particular case when f is uniquely ergodic the convergence in (1.3) is uniform in y and for all y, whence the averaged equation (1.4) and its solution are unique and the latter approximates $X^\varepsilon(n)$, $n\in[0,N/\varepsilon]$ uniformly.

The general case (1.2) when the fast and the slow motions are fully coupled is much more complicated. The averaging principle suggests here to approximate X^ε_x by $\overline{X}^\varepsilon_x$ satisfying (1.4) but with $\overline{\Phi}$ given by

$$\overline{\Phi}(x)=\overline{\Phi}_y(x)=\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1}\Phi(x,f^k_xy) \tag{1.5}$$

provided the last limit exists for "most" x and y. If μ_x is an ergodic invariant measure of f_x then the limit (1.5) exists for μ_x-almost all y's and

$$\overline{\Phi}(x)=\overline{\Phi}_{\mu_x}(x)=\int\Phi(x,y)d\mu_x(y). \tag{1.6}$$

Observe that Lipschitz continuity of $\overline{\Phi}$ cannot be guaranteed now without further assumptions even for smooth Φ, and so we do not have automatically existence and, especially, uniqueness of solutions in (1.4) in these general circumstances. On the other hand, consider the recurrent relation for $\overline{\overline{X}}^\varepsilon(n)=\overline{\overline{X}}^\varepsilon_x(n)$,

$$\overline{\overline{X}}^\varepsilon(n+1)=\overline{\overline{X}}^\varepsilon(n)+\varepsilon\overline{\Phi}(\overline{\overline{X}}^\varepsilon(n)),\quad \overline{\overline{X}}^\varepsilon(0)=x \tag{1.7}$$

which determines $\overline{\overline{X}}^\varepsilon(n)$ without any conditions on $\overline{\Phi}$ and it is easy to see that if $\overline{\Phi}$ is Lipschitz continuous and bounded then

$$\max_{0\leqslant n\leqslant T/\varepsilon}|\overline{X}^\varepsilon_x(n)-\overline{\overline{X}}^\varepsilon_x(n)|\leqslant C_T\varepsilon \tag{1.8}$$

for some $C_T > 0$ independent of ε. Thus we may discuss the approximation of $X^\varepsilon(n)$ by $\overline{\overline{X}}^\varepsilon(n)$ under more general conditions when we even do not have uniquely defined solutions of (1.4).

In general, there exists no natural family of invariant measures μ_x, $x \in \mathbb{R}^d$, since the transformations f_x may have rather different properties for different x's and the averaging principle can be justified here only under substantial restrictions. First, the averaging prescription relies here on existence of a family of probability measures μ_x such that the limit (1.5) exists μ_x-almost everywhere (a.e.) and it is given by (1.6) (at least, Lebesgue a.e. in x). Of course, in addition, we need sufficiently good dependence of Φ and f in (1.2) on ε but still, this does not seem to be enough, in general. The problem here is that the average in (1.5) is taken along orbits of the unperturbed fast motion but in the perturbed evolution (1.2) we cannot disregard now changes in the slow variable parameter of the fast motion, and so we have to study the interplay between unperturbed and perturbed dynamics. Namely, the method of this paper relies on measure estimates of sets of pairs (x, y) which arrive under the action of F_ε^k to sets of points with a specified behavior of averages for the unperturbed evolution. Then we will show that the slow motion is close to the averaged one in certain L^1-sense. Required estimates can be done assuming, for instance, that each f_x is a smooth endomorphism or a diffeomorphism of M preserving a smooth measure μ_x on M which is ergodic for Lebesgue almost all (a.a.) x. This result is a discrete time version of Anosov's theorem [**1**] which is one of few general results about fully coupled averaging. Actually, we prove our result under a general condition which is satisfied in essentially all known cases where the averaging principle holds true and it does not rely on existence of smooth invariant measures as in Anosov's approach.

Recently, quite a few papers dealt with a class of diffeomorphisms called stably ergodic (see, for instance, [**5**]) which are volume preserving ergodic diffeomorphisms having a C^2-neighborhood of volume preserving ergodic diffeomorphisms. If each f_x from our parametric family belongs to such a neighborhood then our results yield an L^1-convergence in the averaging principle. Moreover, we need ergodicity only for almost all x's which suggests to study parametric families of volume preserving diffeomorphisms which are ergodic for almost all parameter values. When convergence in the averaging principle in a fully coupled setup (1.2) holds true for any reasonable Φ we can naturally regard this as a manifestation of compatibility of f_x's or their stability within our parametric family.

Observe that our result works in the case when all f_x's are C^2 expanding transformations of M which always possess fast mixing smooth invariant measures μ_x. On the other hand, close relatives of expanding transformations Anosov and Axiom A diffeomorphism do not possess, generically, smooth invariant measures. Still, relying on specific properties of Axiom A system in a neighborhood of an attractor we will be able to carry out necessary estimates for μ_x being either Lebesgue or corresponding