Oscillatory systems with three separated time scales – analysis and computation

Gil Ariel, Bjorn Engquist and Richard Tsai

Abstract We study a few interesting issues that occur in multiscale modeling and computation for oscillatory dynamical systems that involve three or more separated scales. Averaging in the fastest time scale may indicate a new type of slow variables which do not formally have bounded derivatives. We present a few systems which have such new slow variables and discuss their characterization. The examples motivate a numerical multiscale algorithm that uses nested tiers of integrators which numerically solve the oscillatory system on different time scales. The communication between the scales follows the framework of the Heterogenous Multiscale Method. The applicability and efficiency of the method are demonstrated by examples.

1 Introduction

In this paper we study a few interesting phenomena occurring in oscillatory dynamical systems involving three or more separated time scales. In the typical setting, the fastest time scale is characterized by oscillations whose periods are of the order of a small parameter ε. Classical averaging and multiscale methods consider the effective dynamics of such systems on a time scales which is independent of ε. However, under this scaling, many interesting phenomena, e.g. the nontrivial energy transfer among the linear springs in a Fermi-Pasta-Ulam (FPU) lattice, occur at the $\mathcal{O}(1/\varepsilon)$ or even longer time scales. This motivates our interest in ordinary differential equations (ODEs) with three or more well separated time scales. A good amount of development in numerical methods for long time simulations has been centered

Gil Ariel
Bar-Ilan University, Ramat Gan, 52900, Israel e-mail: arielg@math.biu.ac.il

Bjorn Engquist · Richard Tsai
Department of Mathematics, The University of Texas at Austin, TX 78712, USA e-mail: engquist@ices.utexas.edu, e-mail: ytsai@math.utexas.edu

around the preservation of (approximate) invariances. In the past few years, many numerical algorithms operating on two separated scales have been proposed, see e.g. [17, 25, 23, 14, 15, 27, 12, 8, 13, 11, 1, 2, 3, 24, 7, 26]. To our knowledge, very few algorithms were developed considering directly three or more scales.

For our purpose, it is convenient to rescale time so the slowest time scale of interest is independent of the small parameter ε. Accordingly, the basic assumption underling our discussion is that solutions are oscillatory with periods that are of the order of some powers in ε: $\varepsilon^0, \varepsilon^1, \cdots, \varepsilon^m$. We will study the few issues arising from multiscale modeling and computations for ODEs in the form

$$\dot{\mathbf{x}} = \sum_{i=0}^{m} \varepsilon^{-i} f_i(\mathbf{x}) \;\; ; \quad \mathbf{x}(0) = \mathbf{x}_0 \tag{1}$$

where $0 < \varepsilon \leq \varepsilon_0$, $\mathbf{x} = (x_1, \ldots, x_d) \in \mathbb{R}^d$. We further assumed that the solution of (1) remains in a domain $\mathscr{D}_0 \subset \mathbb{R}^d$ which is bounded independent of ε for all $t \in [0, T]$. For fixed ε and initial condition $\mathbf{x}_0$, the solution of (1) is denoted $\mathbf{x}(t; \varepsilon, \mathbf{x}_0)$. For brevity we will write $\mathbf{x}(t)$ when the dependence on ε and $\mathbf{x}_0$ is not directly relevant to the discussion.

We will focus only on a few model problems involving three time scales. Our goal is to compute the effective dynamics of such a system in a constant, finite time interval $[0, T]$, for the case $0 < \varepsilon < \varepsilon_0 \leq 1$. We will characterize the effective dynamics by some suitable smooth functions $\mathbf{x}$ that change slowly along the trajectories of the solutions, albeit possibly having some fast oscillations with amplitudes that are of the order of $\varepsilon^p, p \geq 1$. Naturally, the invariances of the system will be of interest.

As a simple example, consider the following linear system

$$\begin{cases} \dot{x}_1 &= \frac{1}{\varepsilon} x_2 + x_1 \\ \dot{x}_2 &= -\frac{1}{\varepsilon} x_1 + x_2, \end{cases} \tag{2}$$

with initial conditions $(x_1(0), x_2(0)) = (0, 1)$. The solution is readily given by $(x_1(t), x_2(t)) = (e^t \sin \frac{t}{\varepsilon}, e^t \cos \frac{t}{\varepsilon})$. Taking $I = x_1^2 + x_2^2$, we notice that I has a bounded derivative, i.e., that $\dot{I} := (d/dt) I(x_1(t), x_2(t)) = 2I$ is independent of ε. For this particular example one can easily solve for I, $I(t) = I(0)e^{2t}$. In fact, the uniform bound on $\dot{I}$ indicates the "slow" nature of $I(x_1(t), x_2(t))$ when compared to the fast oscillations in $(x_1(t), x_2(t))$. This type of characterization of the effective dynamics of highly oscillation systems are commonly used in the literatures. See for example [18, 19, 20, 13, 16, 1, 2]. Other approaches to find slow variables includes, e.g. [5, 6]. We formalize this notion with the following definition.

Definition 1. We say that the function $\xi : \mathbf{x} \in \mathscr{A} \mapsto \mathbb{R}$ has a *bounded derivative to order $-k$ for $0 < \varepsilon < \varepsilon_0$ along the flow* $\mathbf{x}(t)$ in $\mathscr{A}$ if

$$\sup_{\mathbf{x} \in \mathscr{A}, \varepsilon \in [0, \varepsilon_0]} |\nabla \xi(\mathbf{x}) \cdot \dot{\mathbf{x}}| \leq C \varepsilon^{-k}, \tag{3}$$

where $\mathscr{A} \subset \mathbb{R}^d$ is an open connected set and C is a constant, both independent of ε. For brevity, we will say that ξ *has a bounded derivative along* $\mathbf{x}(t)$ if (3) holds with $k = 0$. Such functions are commonly referred to as slow variables of the system.

When only two separated time scales are considered, the effective behavior of a highly oscillatory system, $\mathbf{x}(t)$, may be described by a suitably chosen set of variables whose derivative is bounded. In the literature the time dependent function $x_1 = \sin(t)$ with $|\dot{x}_1| = \mathscr{O}(1)$ is naturally regarded as slow and $x_2 = \sin(t/\varepsilon)$ with $|\dot{x}_2| = \mathscr{O}(\varepsilon^{-1})$ is fast. Similarly $x_3 = \sin(t) + \varepsilon \sin(t/\varepsilon)$ is slow. However, when more than two time scales are involved, we also need to consider $x_4 = \sin(t) + \varepsilon \sin(t/\varepsilon^2)$ as slow even if $|\dot{x}_4| = \mathscr{O}(\varepsilon^{-1})$. *It will be regarded as slow because* $|x_4 - \sin t| = \mathscr{O}(\varepsilon)$ *and* $\sin(t)$ *is slow.* As a further example, consider the linear system

$$\begin{cases} \dot{x}_1 = \frac{1}{\varepsilon^2} x_2 + \frac{1}{\varepsilon} + x_1, & x_1(0) = x_{10} \\ \dot{x}_2 = -\frac{1}{\varepsilon^2} x_1 + x_2, & x_2(0) = x_{20} \end{cases} \tag{4}$$

The solution is

$$\begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} = \begin{pmatrix} Ae^t \sin(\varepsilon^{-2} t + \phi) - \frac{\varepsilon^3}{1+\varepsilon^4} \\ Ae^t \cos(\varepsilon^{-2} t + \phi) - \frac{\varepsilon}{1+\varepsilon^4} \end{pmatrix}, \tag{5}$$

where A and ϕ are determined by the initial conditions $A = x_{10}^2 + x_{20}^2$ and $\tan\phi = x_{10}/x_{20}$. As above, we look at the square amplitude $I = x_1^2 + x_2^2$. Its time derivative is bounded to order -1 since

$$\dot{I} = 2\varepsilon^{-1} x_1 + 2I. \tag{6}$$

However, substituting (5) into (6) we find that $I(t) = 2A^2 e^{2t} + O(\varepsilon)$. Hence, even though the derivative of $I(t)$ is not bounded for $0 < \varepsilon < \varepsilon_0$, $I(t)$ consist of a slowly changing part and a small ε-scale perturbation. This example demonstrates that the bounded derivative characterization is not necessary for determining this type of effective property.

Accordingly, Section 2 gives a definition for the time scale on which a certain variable $\alpha(\mathbf{x})$ evolves under the dynamics of ODEs in the form (1). These ideas are further generalized to describe local coordinate systems. In section 3, the dynamics of the variables is analyzed using the operator formalism for homogenization of differential equations, see for example [22]. We focus the discussion to a few example systems in which the singular part of the dynamics is linear. Our observations are discussed in the settings of integrable Hamiltonian systems that can be written in terms of action-angle variables [4].

Another immediate question for problems involving more than two time scales concerns the possible emergence of "stochastic" behaviors. The operator formalism for homogenization is a useful tool in the determination of stochasticity.

Section 4 presents a numerical method that uses nested tiers of integrators which numerically solve the oscillatory system on different time scales. The communication between the scales follows the framework of the Heterogenous Multiscale

Method (HMM) [10, 9]. Section 5 presents a few numerical examples. We give conclusion in section 6.

2 Effective behavior across different time scales

In this section we discuss some of the mathematical notions which we use to study systems containing several well-separated time scales.

2.1 Slowly changing quantities

Definition 2. A smooth time dependent function $\alpha : [0,T] \mapsto \mathbb{R}^n$ is said to *evolve on the ε^k time scale in* $[0,T]$ for some integer k and for $0 < \varepsilon < \varepsilon_0$, if there exists a smooth function $\beta : [0,T] \mapsto \mathbb{R}^n$ and constants C_0 and C_1 such that

$$\sup_{t\in[0,T]} \left| \frac{d}{dt}\beta(t) \right| \leq C_0 \varepsilon^{-k},$$

and

$$\sup_{t\in[0,T]} |\alpha(t) - \beta(t)| \leq C_1 \varepsilon.$$

This motivates the following definition for a variable, $\alpha(\mathbf{x})$, that evolves on the ε^k time scale along the solutions of (1).

Definition 3. A function $\xi(\mathbf{x})$ is said to *evolve on the ε^k time scale along the trajectories of (1) in* $[0,T]$ *and in an open set* $\mathscr{A}$ if, for all initial conditions $\mathbf{x}_0 \in \mathscr{A}$, the time dependent function $\xi(\mathbf{x}(t;\varepsilon,\mathbf{x}_0))$ evolves on the ε^k time scale in $[0,T]$. For brevity, we will refer to quantities and variables that evolve on the ε^0 time scale as slow.

In particular, the above definition suggests that the limit

$$\xi_0(s;\mathbf{x}_0) = \lim_{\varepsilon\to 0} \xi(\mathbf{x}(\varepsilon^k s;\varepsilon,\mathbf{x}_0)) \tag{7}$$

exists for all $s \in [0,T]$ and $\mathbf{x}_0 \in \mathscr{A}$. For instance, in both examples (2) and (4), the square amplitude $I = x_1^2 + x_2^2$ evolve on the ε^0 time scale. The difference is that (according to Definition 3), I has a bounded derivative of order 0 along the flow of (2) but not along the flow of (5). More generally, considering $\alpha(t)$ to be the image of $\xi(\mathbf{x}(t))$, Definitions (2) and (3) allows the inclusion of functions such as $\alpha(t) = \varepsilon \sin(\varepsilon^{-2}t) + \sin(t)$ (with unbounded derivatives) to be characterized as slowly evolving.

Next, we consider the following slightly more general system

$$\frac{dx}{dt} = \frac{i}{\varepsilon^2}x + f_I(x,y,t), \tag{8}$$
$$\frac{dy}{dt} = \frac{1}{\varepsilon}g(x)y + f_{II}(x,y,t). \tag{9}$$

Introducing $x = \exp(it/\varepsilon^2)z$, we obtain

$$\frac{dz}{dt} = \exp(-\frac{it}{\varepsilon^2})f_I(\exp(\frac{it}{\varepsilon^2})z,y,t), \tag{10}$$
$$\frac{dy}{dt} = \frac{1}{\varepsilon}g(\exp(\frac{it}{\varepsilon^2})z)y + f_{II}(\exp(\frac{it}{\varepsilon^2})z,y,t). \tag{11}$$

If

$$\int_0^t g(x(s))ds = \mathscr{O}(\varepsilon), t > 0$$

then y can be bounded. This is possible since the oscillations in x occur on a time scale that is much faster than the ε-scale:

$$\frac{d}{dt}g(x(t)) = \mathscr{O}(\varepsilon^{-2}).$$

Thus, if $g(x-\bar{x_0})$ is an odd function for

$$\bar{x_0} := \lim_{\varepsilon\to 0+}\int_0^t x(s;x_0)ds,$$

then the singular term in the y equation produces only small scale perturbation to the evolution of $y(t)$ as it only produces fast oscillations of $\mathscr{O}(\varepsilon)$ amplitude in the trajectories of y. In this case, $y(t)$ is a slowly changing quantity along the trajectory, i.e, it evolves on the ε^0 scale. Alternatively, if $g(x-\bar{x_0})$ is even then y evolves on the ε time scale. Note that additional restrictions are required to insure that y does not blow up on the ε scale but stays bounded in the domain $\mathscr{D}_0$. This observation also suggests that in determining if y changes slowly in time, we may test if g is odd around a neighborhood of the averages of x. If so, one can simply ignore the term containing g in solving for y.

The above simple example suggests a way to detect functions which are slow but which do not have bounded derivatives. algebraically, without integrating along the fast solutions for long time. Let $\mathbf{x}$ be a quasi-periodic solution of an highly oscillatory system with $\mathscr{O}(\varepsilon^{-2})$ frequencies. Furthermore, assuming that x has an average $\bar{x_0}$ as $\varepsilon \to 0$. Consider $\alpha(t) := \xi(\mathbf{x}(t))$ with

$$\frac{d}{dt}\alpha(t) = \frac{1}{\varepsilon}r(\mathbf{x}(t)).$$

Then α is may be slow if $r(x-\bar{x_0})$ is an odd function.

Finally, we point out that the emergence of a slow variable whose time derivative along the oscillatory trajectories may come from a multiscale series expansion of parts of the solution. Consider again equations (8) and (9). The leading order term

comes out naturally when y has an expansion of the form

$$y(t) = y_0(t) + \varepsilon h(x(t)) + \cdots .$$

Hence, we expect that the homogenization approach described in the following Section should capture such type of effective behavior of a dynamical system.

2.2 Multiscale charts

Given an oscillatory dynamical system in $\mathbb{R}^d$, special functions may be used to analyze the structure of the dynamics. For example, the action and angle variables for a given Hamiltonian system provide a coordinate system in the phase space such that the resulting Hamiltonian dynamics is separated into evolutions on certain invariant tori (oscillations) as described by the angle variables, and non-oscillatory evolutions described by the action variables [4]. For example, the function I defined for (2) together with $\arctan(x_2,x_1)$ corresponds to such a situation in which I is non-oscillatory along the dynamics and provides a coordinate perpendicular to the trajectories. In previous work, we propose the use of a similar strategy for a different class of dynamical systems [1, 2].

Consider the oscillatory dynamical system (1), and a family of trajectories $\mathbf{x}(t;\varepsilon,\mathbf{x}_0)$ in a open set $\mathscr{A} \subset \mathbb{R}^d$. Let $\Phi : \mathscr{A} \subset \mathbb{R}^d \to U \subset \mathbb{R}^d$ be a diffeomorphism that is independent of ε. Thus Φ is a local coordinate system (chart) for $\mathscr{A} \subset \mathbb{R}^d$. We denote the vector $\Phi(\mathbf{x})$ by $(\phi^1(\mathbf{x}),\phi^2(\mathbf{x}),\cdots,\phi^d(\mathbf{x}))$, where $\phi^i(\mathbf{x})$ is a real valued function defined in $\mathbb{R}^d$. We shall refer to ϕ^i as the i-th coordinate. Let $n(\Phi,k;\mathbf{x}(\cdot;\varepsilon,\mathbf{x}_0))$ denote the number of coordinates in Φ that evolve along $\mathbf{x}(t;\varepsilon,\mathbf{x}_0)$ on time scales that are smaller or equal to k. We have the following definition:

Definition 4. A chart Φ is said to be maximally slow if for any other chart $\tilde{\Phi}$ defined on $\mathscr{A}$, $n(\Phi,k;\mathbf{x}(\cdot;\varepsilon,\mathbf{x}_0)) \geq n(\tilde{\Phi},k;\mathbf{x}(\cdot;\varepsilon,\mathbf{x}_0))$ for all k.

Loosely speaking, the coordinates of Φ are as slow as possible. A numerical method for identifying a maximally slow chart for the case in which the singular parts of the dynamics is linear is describes in the appendix.

Let ξ denote a maximally slow chart with k time scales, i.e.,

$$\xi = (\xi^1,\ldots,\xi^k)$$

where $\xi^i \in \mathbb{R}^{d_i}$ are the variables evolving on the i-th time scale, $i = 1,\ldots,k$. Using the principle of averaging iteratively for each scale, effective equations for each time scale can be constructed and their solutions that approximates the exact dynamics of the coordinates ξ^i in the corresponding time scale. To obtain these equations, faster time scale components are averaged while keeping the slower ones fixed. Formally, we write

$$\dot{\xi}^i = \varepsilon^{-i} F^i(\xi^i;\xi^0,\ldots,\xi^{i-1}) + \mathscr{O}(\varepsilon),$$

with appropriate initial conditions. The effective equations hold for a time scale which is of the order of ε^i. Furthermore, F^i can be obtained iteratively by averaging over the effective dynamics of the faster $i+1$ scale. Accordingly, we say that the chart ξ is effectively closed.

3 An homogenization approach

The multiscale structure of a system can be analyzed using the operator formalism as presented in [22], which in turn, formally generalizes the work of Papanicolau [21]. Motivated by perturbed integrable systems, in which the dynamics can be written in terms of action-angle variables, we concentrate on example systems in which the singular part of the dynamics is linear.

The analysis motivates a numerical multiscale algorithm along the lines of the HMM framework [1, 9, 12]. The algorithm does not assume that the system is given in the convenient action-angle coordinates, but only that such a transformation exists.

Consider a general ODE system whose right hand side depends on ε

$$\frac{dx}{dt} = f_\varepsilon(x).$$

The associated Liouville equation takes the form

$$\partial_t u^\varepsilon + f_\varepsilon \cdot \nabla_x u^\varepsilon = 0, \tag{12}$$

where ∂_x denotes partial differentiation with respect to X. This is a linear equation whose characteristics coincide with solutions of the ODE. We begin by matching powers of ε in the multiscale expansion of the operator $L_\varepsilon := f_\varepsilon \cdot \nabla_\varepsilon$ and that of the solution u^ε. Formally, we write

$$L_\varepsilon = \frac{1}{\varepsilon^2} L_2 + \frac{1}{\varepsilon} L_1 + L_0, \tag{13}$$

and

$$u^\varepsilon = u_0 + \varepsilon u_1 + \varepsilon^2 u_2 + \dots \tag{14}$$

Substituting the above expansions into (12) yields

$$\partial_t u_0 = \frac{1}{\varepsilon^2} L_2 u + \frac{1}{\varepsilon} (L_2 u_1 + L_1 u_0) + (L_2 u_2 + L_1 u_1 + L_0 u_0) + \mathcal{O}(\varepsilon).$$

Comparing orders of ε, we have

$$\begin{aligned}\frac{1}{\varepsilon^2}&: \; L_2 u_0 = 0,\\ \frac{1}{\varepsilon}&: \; L_2 u_1 = -L_1 u_0,\\ 1&: \; \partial_t u_0 = L_2 u_2 + L_1 u_1 + L_0 u_0.\end{aligned}$$

We see that a closed effective equation for u_0 can be derived if both $L_2 u_2$ and $L_1 u_1$ can be approximated by operations on u_0 only. This closure is typically done by averaging over some invariant manifolds. In the following subsections, we apply this procedure to some model problems.

3.1 A two scales example

For completeness, we recall the application of the operator formalism in a two-scale highly-oscillatory ODE system. Let

$$\begin{cases}\dot{x} = \frac{1}{\varepsilon} y + f(x,y)\\ \dot{y} = -\frac{1}{\varepsilon} x + g(x,y),\end{cases} \tag{15}$$

with some non-zero initial condition. Changing into polar coordinates $(r,\theta) \in \mathbb{R} \times S^1$ yields

$$\begin{cases}\dot{r} = \left(x f(x,y) + y g(x,y)\right)/r\\ \dot{\theta} = -\frac{1}{\varepsilon} + \left(x g(x,y) - y f(x,y)\right)/r^2.\end{cases}$$

It is clear that the amplitude r is a slow variable while the phase ϕ is fast. Hence, we can naively average the right hand side of the equation for r with respect to the fast phase. This yields an effective equation for the amplitude

$$\begin{aligned}\dot{r} &= F(r)\\ F(r) &= \frac{1}{r}\int_{S^1} \left[x(r,\theta) f(x(r,\theta),y(r,\theta)) + y(r,\theta) g(x(r,\theta),y(r,\theta))\right] d\theta.\end{aligned} \tag{16}$$

The effective dynamics can be also derived using the operator formalism [22]. The backwards, or Liouville equation associated with the ODE is

$$\partial_t u = Lu \;\; ; \qquad u(0,x,y) = \psi(x,y), \tag{17}$$

where L is the generator, or the Liouville operator

$$\begin{aligned}L &= \frac{1}{\varepsilon} L_1 + L_0\\ L_1 &= y\partial_x - x\partial_y\\ L_0 &= f(x,y)\partial_x + g(x,y)\partial_y,\end{aligned} \tag{18}$$

Noting that $L_1 = \partial_\theta$, the Null space of $L_1 = -L_1^*$ is

$$\text{Null } L_1 = \text{Null } L_1^* = \{\xi = \xi(x^2+y^2)\},$$

where $[\cdot]^*$ denotes the dual. Let P denote projection on Null L_1, which can be performed by averaging over the fast angle θ, $P[\cdot] = \int_{S^1}[\cdot]d\theta := \langle\cdot\rangle$. Substituting the asymptotic expansion for u, Eq. (14), into the backwards equation (17) yields

$$\begin{cases} L_1 u_0 &= 0 \\ L_1 u_1 &= \partial_t u_0 - L_0 u_0 \end{cases}$$

The equation for u_0 implies that $u_0 \in \text{Null } L_1$, i.e., $u_0 = u_0(t,r)$. Formally, we write

$$Pu_0 = u_0. \tag{19}$$

The solvability condition for u_1 is

$$[\partial_t - L_0]\, u_0 \perp \text{Null } L_1^*.$$

Substituting (19) yields

$$P\,[\partial_t - L_0]\, Pu_0 = 0.$$

This gives the effective equation for u_0

$$\partial_t u_0 = PL_0 Pu_0,$$

where we used the fact that u_0 does not depend on θ, and can therefore be taken out of the averaging. This can be rewritten as

$$\partial_t u_0 = \langle f(x,y)\partial_x u_0 + g(x,y)\partial_y u_0\rangle.$$

Using the chain rule yields

$$\partial_t u_0 = \frac{1}{r}\langle x f(x,y) + y g(x,y)\rangle \partial_r u_0,$$

which is nothing but the Liouville equation associated with the effective ODE (16).

3.2 Three scales - example 1

Consider the following three-scale system which involves slow variables whose derivatives are not bounded.

$$\begin{cases} \dot{x}_1 &= \frac{1}{\varepsilon^2}x_2 + f_I(x_1,x_2,y) \\ \dot{x}_2 &= -\frac{1}{\varepsilon^2}x_1 + f_{II}(x_1,x_2,y) \\ \dot{y} &= \frac{1}{\varepsilon}x_1 + f_{III}(x_1,x_2,y) \end{cases} \tag{20}$$

To get some intuition, consider the unperturbed case $f_I = f_{II} = f_{III} = 0$ with initial conditions $(x_1, x_2, y) = (1, 0, 1)$. The solution is

$$\begin{aligned} x_1(t) &= -\cos(\varepsilon^{-2}t) \\ x_2(t) &= \sin(\varepsilon^{-2}t) \\ y(t) &= 1 - \varepsilon \sin(\varepsilon^{-2}t). \end{aligned}$$

Hence, we can see that the system has two variables which evolve on the $O(1)$ time scale: $I = x_1^2 + x_2^2$ and y. In the unperturbed case, both variables are constants.

Getting back to the full equation (20), the Liouville (backward) equation associated with the ODE is

$$\partial_t u = Lu \ ; \qquad u(0, x_1, x_2, y) = \psi(x_1, x_2, y). \tag{21}$$

where L is the generator of the process

$$\begin{aligned} L &= \frac{1}{\varepsilon^2} L_2 + \frac{1}{\varepsilon} L_1 + L_0 \\ L_2 &= x_2 \partial_{x_1} - x_1 \partial_{x_2} \\ L_1 &= x_1 \partial_y \\ L_0 &= f_I \partial_{x_1} + f_{II} \partial_{x_2} + f_{III} \partial_y. \end{aligned}$$

The Null space of $L_2 = -L_2^*$ is

$$\text{Null } L_2 = \text{Null } L_2^* = \{\xi = \xi(r, y)\},$$

where $r^2 = x_1^2 + x_2^2$. Let P denote projection on Null L_2, which can be performed by averaging over the fast phase $\theta = \arctan x_2/x_1$. Let P denote projection on Null L_2. As before, averaging over the fast phase θ is denoted by $\langle \cdot \rangle$.

Substituting the asymptotic expansion for u, Eq. (14), into the backwards equation (21) and comparing powers of ε yields

$$\begin{cases} L_2 u_0 &= 0 \\ L_2 u_1 &= -L_1 u_0 \\ L_2 u_2 &= \partial_t u_0 - L_1 u_1 - L_0 u_0. \end{cases} \tag{22}$$

Leading order equation:

The equation for u_0 implies that $u_0 \in \text{Null } L_2$, i.e., $u_0 = u_0(t, r, y)$. Formally, we write

$$P u_0 = u_0 \tag{23}$$

Order $1/\varepsilon$ equation:

The solvability condition for u_1 in (22) implies

$$L_1 u_0 \perp \text{Null } L_2^*, \tag{24}$$

which is equivalent to

$$PL_1 u_0 = 0.$$

This holds since

$$\begin{aligned} PL_1 u_0 &= P\,[x_1 \partial_t]\, u_0(x_1^2 + x_2^2, y) = \langle x_1 \partial_y u_0(x_1^2 + x_2^2, y)\rangle = \\ &= \langle x_1 \rangle \partial_y u_0(x_1^2 + x_2^2, y) = 0. \end{aligned}$$

Hence, we formally write

$$u_1 = -L_2^{-1} L_1 u_0. \tag{25}$$

Order $1/\varepsilon$ equation:

The solvability condition for u_2 in (22) is

$$[\partial_t - L_1 u_1 - L_0 u_0] \perp \text{Null } L_2^*.$$

Substituting in the formal solution (25) yields

$$P\left[\partial_t - L_1 L_2^{-1} L_1 - L_0\right] P u_0 = 0.$$

For the example at hand, we have

$$u_1 = -L_2^{-1} x_1 \partial_t u_0.$$

Furthermore, $\partial_y u_0$ has the form $g(x_1^2 + x_2^2, y)$, which implies that $\partial_t u_0 \in \text{Null } L_2^*$. Also, $L_2 x_2 = -x_1$. Hence,

$$L_2^{-1} L_1 u_0 = u_1 = x_2 \partial_y u_0.$$

We conclude that

$$L_1 L_2^{-1} L_1 u_0 = x_1 \partial_y x_2 \partial_y u_0 = x_1 x_2 \partial_{yy} u_0.$$

This yields the effective equation for u_0

$$\partial_t u_0 = \langle x_1 x_2 \partial_y y u_0 \rangle - \langle f_I \partial_{x_1} u_0 \rangle - \langle f_{II} \partial_{x_2} u_0 \rangle - \langle f_{III} \partial_y u_0 \rangle \tag{26}$$

As before, $\partial_{yy} u_0 \in \text{Null } L_2$ and

$$\langle x_1 x_2 \partial_{yy} u_0 \rangle = \langle x_1 x_2 \rangle \partial_{yy} u_0 = O(\varepsilon).$$

The effective equation (26) becomes

$$\partial_t u_0 = -\langle f_I \partial_{x_1} u_0 \rangle - \langle f_{II} \partial_{x_2} u_0 \rangle - \langle f_{III} \partial_y u_0 \rangle,$$

which can be rewritten as

$$\partial_t u_0 = -\frac{1}{r} \langle x_1 f_I + x_2 f_{II} \rangle \frac{\partial u_0}{\partial r} - \langle f_{III} \rangle \partial_y u_0.$$

This equation can be identified as the Liouville equation associated with the ODE

$$\begin{cases} \dot{r} = \langle x_1 f_I + x_2 f_{II} \rangle / r \\ \dot{y} = \langle f_{III} \rangle. \end{cases} \tag{27}$$

We conclude that, to leading order in ε, u_0, and hence r and y, evolve on the $O(1)$ time scale and are deterministic.

3.3 Three scales - Example 2

We consider a simple system involving three time scales

$$\begin{aligned} \frac{d}{dt} \begin{pmatrix} x \\ y \end{pmatrix} &= \frac{1}{\varepsilon^2} \begin{pmatrix} -y \\ x \end{pmatrix} + f(x,y,w,z), \\ \frac{d}{dt} \begin{pmatrix} w \\ z \end{pmatrix} &= \frac{1}{\varepsilon} \begin{pmatrix} -z \\ w \end{pmatrix} + g(x,y,w,z). \end{aligned}$$

If $f = g = 0$, then (x,y) and (w,z) are decoupled harmonic oscillators with frequencies $2\pi/\varepsilon^2$ and $2\pi/\varepsilon$, respectively. The associated Liouville equation (12) has parameters

$$\begin{aligned} L_2 &= \begin{pmatrix} -y \\ x \end{pmatrix} \cdot \nabla_{x,y} \\ L_1 &= \begin{pmatrix} -z \\ w \end{pmatrix} \cdot \nabla_{w,z} \\ L_0 &= f(x,y,w,z) \cdot \nabla_{x,y} + g(x,y,w,z) \cdot \nabla_{w,z} \end{aligned} \tag{28}$$

where $\nabla_{A,B}$ denotes the gradient with respect to the variables A and B, for example, $\nabla_{x,y} = (\partial_x, \partial_y)^T$. We further denote a change of variables into polar coordinates: $(x,y) \mapsto (r,\theta)$ and $(w,z) \mapsto (\rho,\phi)$ and note that $L_2 = \partial_\theta$ and $L_1 = \partial_\phi$. These are also the action-angle variables of this system.

Leading order equation:

$L_2 u_0 = 0$ implies that u_0 must be constant in $\theta(x,y)$, $u_0 = u_0(r,w,z)$. The Jacobian of these variables, which define the null space of $L_2 = -L_2^*$, has full rank at the entire domain of relevance. As before, let P denote projection on the Null L_2^* and $Pu = \langle u \rangle$ denotes averaging with respect to θ.

Order $1/\varepsilon$ equation:

The equation takes the form

$$\partial_\theta u_1 = -\partial_\phi u_0.$$

The solvability condition is $\langle -\partial_\phi u_0 \rangle = 0$. However, since u_0 does not depend on θ, we must have $\partial_\phi u_0 = 0$. In other words, $u_0 = u_0(t,r,\rho)$ We conclude that r and ρ are the only slow variables (evolve on the ε^0 time scale). Both variables have a bounded derivative (of order 0). Further more, the first order perturbation, u_1, vanishes.

Order 1 equation:

Substituting $u_1 = 0$, the equation for u_2 is

$$\partial_\theta u_2 = -\partial_t u_0 + L_0 u_0.$$

The solvability condition is

$$\partial_t u_0 = \langle L_0 \rangle u_0. \tag{29}$$

However, (29) is not closed since $\langle L_0 \rangle$ involves ϕ.

In order to bypass this difficulty we notice that in the example at hand, the dynamics of ϕ and θ are decoupled and the invariance measure for both variables (with r and ρ fixed) is uniform over a 2D torus T^2. This occurrence is not incidental, but actually the general case since the angle variable in systems which are given in action-angle coordinates undergo uniform rotations on a torus [4]. Therefore, (29) can be closed by averaging both sides over both θ and ϕ,

$$\partial_t u_0 = \int_{T^2} L_0 d\theta d\phi u_0, \tag{30}$$

where we used the fact that u_0 is independent of both angles.

3.4 Observations

Following this methodology, we have the following observations:

- The commonly slow variables as defined in Definition 3 with $k = 0$ lie in the Null space of both L_2 and L_1.
- Variables with a bounded derivative need only lie in the Null space of L_2.
- The homogenization approach picks out the new type of slow variables as defined in Definitions 2 and 3 with $k = 0$. See (27) which is derived from the system defined in (20).
- The effective slow dynamics in all the examples presented in this manuscript is found to be deterministic. However, for a large class of equations involving chaotic solutions [22], the effective PDEs are diffusive. This means that the effective behaviors of the original dynamical systems can be approximated by solutions of the corresponding stochastic differential equations weakly.

4 Numerical algorithms

In this section, we discuss an approach which invoke our previous HMM algorithms [1, 2, 3] hiarchically for problems involving more than two time scales. We consider the time scales $\mathscr{O}(\varepsilon^2)$, $\mathscr{O}(\varepsilon)$, and $\mathscr{O}(1)$.

Suppose we obtain a maximal slow chart, for example, using the method described in the Appendix for identifying polynomial variables evolving on different time scales. We denote this system of coordinates $\xi = (\xi^1, \xi^2, \xi^3)$, where $\xi^i = (\xi^i_1, \ldots, \xi^i_{d_i})$ are the variables evolving on the i-th time scale.

The HMM to be constructed should evaluate the rate of change of ξ along the flow of the original oscillatory system. This typically involves numerically averaging over the fast oscillations in the system. For the three scale problems that we consider, we will be required to average over the $\mathscr{O}(\varepsilon^2)$ scale oscillations as well as the $\mathscr{O}(\varepsilon)$ scale oscillations, thus obtaining a numerical approximation for the effective equation for the slow variables ξ^0. The method is schematically illustrated in figure 1.

Our goal is to numerically integrate the effective equation for ξ^0, which is not known. Following HMM, this equation is approximated by averaging the dynamics of the system over a short-time calculation of the slower, $\mathscr{O}(\varepsilon)$ scale. Hence, whenever the algorithm needs to take a coarse, $\mathscr{O}(1)$ step, we call an auxiliary function whose role is to calculate the dynamics on that time scale. This requires approximating the dynamics of ξ^1 for a time segment of order $\mathscr{O}(\varepsilon)$. However, with three scales the dynamics of ξ^1 is given by an effective equation which is itself not known, but can nonetheless be approximated with a second tier of HMM integrator. This second tier approximates effective dynamics of ξ^1 by numerically averaging the dynamics on the $\mathscr{O}(\varepsilon^2)$ scales, namely ξ^2. Note that this is possible because we only require to solve ξ^1 on a time segment of order ε. Longer time scales of order 1 are not accessible as the error of using the effective averaged equation rather than the full one becomes large.

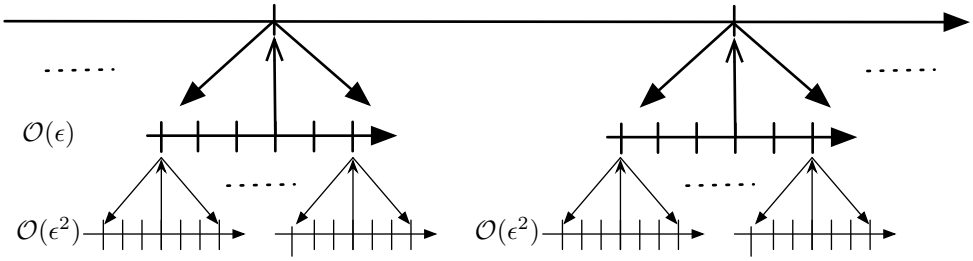

Fig. 1 An illustration of a three scale algorithm.

5 Examples

In this section we demonstrate the new multi-tier HMM algorithm in a few examples.

5.1 Harmonic oscillators

Consider the following system describing two coupled harmonic oscillators in resonance

$$\begin{cases} \dot{x}_1 &= -\frac{1}{\varepsilon^2} y_1 + \frac{1}{\varepsilon} y_2^2 - 3x_1 x_2^2 \\ \dot{y}_1 &= \frac{1}{\varepsilon^2} x_1 + \frac{1}{2} y_1 \\ \dot{x}_2 &= -\left(\frac{1}{\varepsilon^2} + \frac{1}{\varepsilon}\right) y_2 - x_2 \\ \dot{y}_2 &= \left(\frac{1}{\varepsilon^2} + \frac{1}{\varepsilon}\right) x_2 - y_2 + 2x_1^2 y_2 \end{cases} \tag{31}$$

As depicted in Figure 2, all four state variables oscillate with a period which is of the order of ε^2. Hence, x_1, y_1, x_2 and y_2 evolve on the ε^2 time scale.

In order to find a maximally slow coordinate system, we change to polar coordinates $(x_i, y_t) \mapsto (I_i, \varphi_i)$, $i = 1, 2$ and introduce a polynomial variable θ that describes the 1:1 resonance between the oscillators

$$\begin{aligned} I_1 &= x_1^2 + y_1^2 \\ I_2 &= x_2^2 + y_2^2 \\ \theta &= x_1 x_2 + y_1 y_2 \\ \cos\varphi_1 &= x_1 / \sqrt{I_1}. \end{aligned} \tag{32}$$

The corresponding time derivatives are

$$\begin{aligned}
\dot{I}_1 &= \frac{2}{\varepsilon} x_1 y_2^2 - 3x_1 x_2^2 + y_1^2 \\
\dot{I}_2 &= 2x_2 y_2 - 2x_2^2 - 2y_2^2 + 4x_1^2 x_2 y_2 \\
\dot{\theta} &= \frac{1}{\varepsilon}(x_2 y_2^2 + y_1 x_2 - x_1 y_2) + (y_1 x_2 - x_1 x_2 - 3x_1 x_2^3 + 2x_1^2 y_1 y_2) \\
\dot{\varphi}_1 &= \frac{1}{\varepsilon^2}.
\end{aligned} \tag{33}$$

It appears as if $(I_1, I_2, \theta, \varphi_1)$ is a chart in which φ_1 evolves of the ε^2 time scale, I_1 and θ evolve on the ε time scale while I_2, which has a bounded derivative, evolves on the $O(1)$ scale. The dynamics of the three slow variables I_1, I_2 and θ on the $O(\varepsilon)$ scale is depicted in figure 3. The figure suggests that both I_1 and I_2 are practically constant on the ε scale. Indeed, it can be shown that the average of $x_1 y_2^2$ on any segment of length $O(\varepsilon)$ and larger is of order ε^2. Therefore, the averaged $\dot{I}_1$ is bounded independent of ε and I_1 evolves on the $O(1)$ time scale, rather than the expected $O(\varepsilon)$. The time evolution of I_1 and I_2 on the slowest $O(1)$ time scale is depicted in figure 4. In addition, the figure shows the results of the two-tier HMM integrator described in section 4. The HMM algorithm approximates the slow $O(1)$ dynamics using macroscopic steps which are independent of ε. The integration is done using a fourth order method (in the macroscopic step size) and it's efficiency is independent of ε.

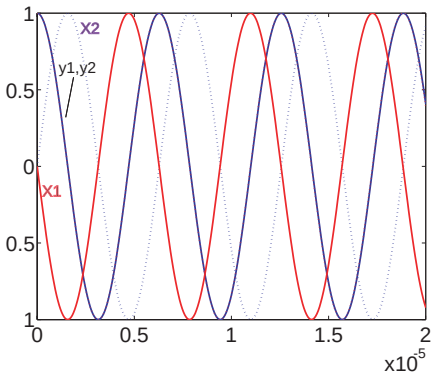

Fig. 2 The dynamics of (31) on the ε^2 time scale. $\varepsilon = 10^{-3}$

5.2 An example motivated by the Fermi-Pasta-Ulam problem

The following example, which consists of three coupled oscillators, is motivated by a version of the FPUα model with periodic boundary conditions. The system is described by the Hamiltonian

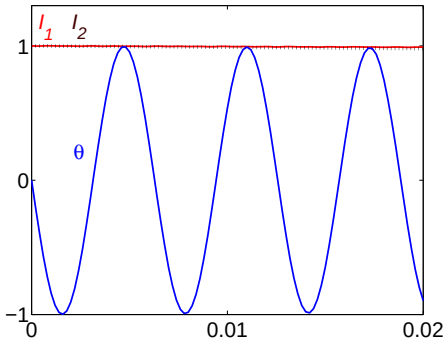

Fig. 3 The dynamics of (31) on the ε^1 time scale. $\varepsilon = 10^{-3}$

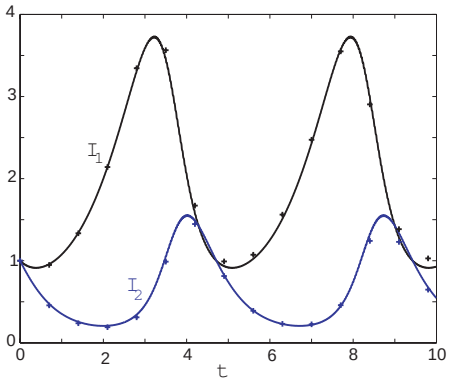

Fig. 4 The dynamics of (31) on the ε^0 time scale. $\varepsilon = 10^{-3}$. plus signs are results of a 2-tier HMM with fourth order RK on all scales.

$$H = \frac{1}{2}\sum_{i=1}^{3} p_i^2 + \sum_{i=1}^{3}\left[\frac{1}{2}(q_{i+1}-q_i)^2 + \frac{\varepsilon}{3}(q_{i+1}-q_i)^3\right], \tag{34}$$

where $q_0 = q_3$ and $q_4 = q_1$. The dynamics, which is given by Hamilton's equations of motion, preserves the Hamiltonian. Accordingly, the purpose of this example is to demonstrate the advantages of the HMM multiscale method for Hamiltonian systems compared to reversible and symplectic integrators.

Rescaling time, $s = \varepsilon^2 t$, and denoting $[\cdot]' = (d/ds)$, the dynamics is given by

$$\begin{cases} q_1' &= \frac{1}{\varepsilon^2} p_1 \\ p_1' &= -\frac{1}{\varepsilon^2}(2q_1 - q_3 - q_2) - \frac{1}{\varepsilon}(q_2 - q_3)(2q_1 - q_3 - q_2) \\ q_2' &= \frac{1}{\varepsilon^2} p_2 \\ p_2' &= -\frac{1}{\varepsilon^2}(2q_2 - q_1 - q_3) - \frac{1}{\varepsilon}(q_3 - q_1)(2q_2 - q_1 - q_3) \\ q_3' &= \frac{1}{\varepsilon^2} p_3 \\ p_3' &= -\frac{1}{\varepsilon^2}(2q_3 - q_2 - q_1) - \frac{1}{\varepsilon}(q_1 - q_2)(2q_3 - q_2 - q_1) \end{cases} \tag{35}$$

Due to the periodic boundaries the total momentum is preserved. Hence, without loss of generality we pick initial conditions such that the center of mass is fixed at the origin, $p_{\text{tot}} = p_0 + p_1 + p_2 = 0$ and $q_{\text{cm}} = q_0 + q_1 + q_2 = 0$. Using the algorithm detailed in Appendix 6, we identify five variables evolving on the 1 time scale:

$$\begin{aligned} p_{\text{tot}} &= p_1 + p_2 + p_3 \\ q_{\text{cm}} &= q_1 + q_2 + q_3 \\ I_1 &= 3q_2^2 + p_2^2 \\ I_2 &= 3q_3^2 + p_3^2 \\ \theta &= 3q_2q_3 + p_2p_3 \end{aligned} \tag{36}$$

Differentiating with respect to the rescaled time and using the fixed center of mass assumption yields

$$\begin{aligned} p_{\text{tot}'} &= 0 \\ q'_{\text{cm}} &= 0 \\ I'_1 &= \frac{18}{\varepsilon} p_2q_2(2q_3 + q_2) \\ I'_2 &= -\frac{18}{\varepsilon} p_3q_3(2q_3 + q_2) \\ \theta' &= \frac{3}{\varepsilon}(p_2q_2 - p_3q_3)(2q_3 + q_2) \end{aligned} \tag{37}$$

Figure 5 compares HMM with Verlet using the initial conditions $(q_1, q_2, q_3) = (0.3, -0.4, 0.1)$ and $(p_1, p_2, p_3) = (0.3, -0.4, 0.1)$. Taking $\varepsilon = 10^{-4}$, parameters are chosen to give an error of about 1%. HMM parameters are $H = 10$, $\eta = 75.1\varepsilon^2$, $h = \varepsilon^2/20$. The micro-solver is Verlet. The macro-solver is the midpoint rule, the kernel is exponential [12]. Solving the system using Verlet with $\varepsilon = 10^{-4}$ is practically impossible. However, in order to achieve the desired accuracy the values of ε can be increased artificially [28]. Since we require a relative error of order 0.01, we take $\varepsilon = 0.01$ and decrease step size until the energy drift is of the same order. This requires $h = \varepsilon/100$. With these parameters HMM runs over 300,000 time faster. It is interesting to note that the efficiency of both methods is independent of ε. Hence, the gain in efficiency does not depend on ε as long as it is small enough. Lastly, the energy drift with HMM is small, but seems to increase. At longer time segments time reversible schemes, as developed in [3], may be advantageous.

6 Summary

In this paper, we investigate several issues related to the design of multiscale algorithms for computing the effective behavior of highly oscillatory dynamical systems involving more than two separated scales. We discuss a type of effective behavior (slowly changing quantities) which do not have bounded derivatives. Consequently,

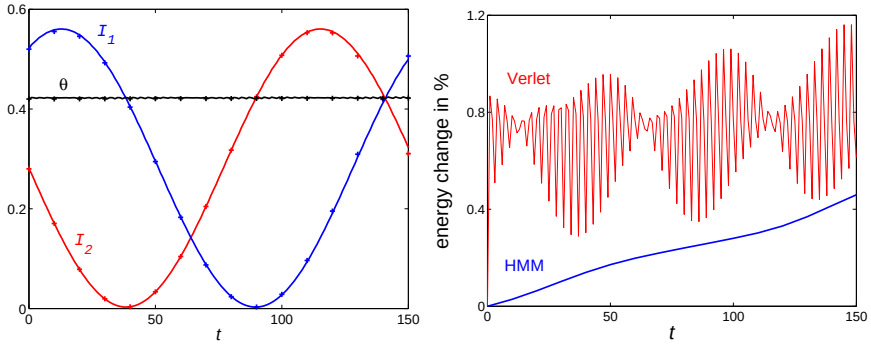

Fig. 5 Left: slow variables. Dotted line - Verlet with an adjusted value of ε, plus signs - HMM. Right: energy drift. With simulation parameters tuned to give comparable errors in the total energy, HMM runs over 300,000 times faster.

they cannot be found by using a common homogenization technique based on multi-scale expansions. We further demonstrate that this type of effective behavior cannot be ignored in our numerical examples.

Acknowledgments

Research are performed under the support of NSF DMS-0714612. Tsai is partially supported by a Sloan Fellowship.

Appendix: Finding polynomial slow variables

In many highly oscillatory physical systems the leading order oscillations are harmonic. It can then be shown [1, 4] that slow variables can be polynomials. In [1] we suggest a variational method to automatically identify the polynomials making up maximally slow charts using the bounded derivative concept. As was demonstrated in this paper, with three or more well-separated scales, the bounded derivative concept is not sufficient to determine the time scale on which a variable evolves. Hence, it cannot be used to construct maximally slow charts. Accordingly, the purpose of this appendix is to modify the variational method of [1] to use the new concept of slow variable, Definition 3. The main idea is to compare two trajectories with different values of ε and find a polynomial that takes similar values on both trajectories.

Let $p(x)$ denote a polynomial in $\mathbb{R}^d$. Following Definition 3, $p(x)$ evolves on the ε^k time scale if, for all $0 \leq i \leq k$, the limit

$$\lim_{\varepsilon \to 0} p(\mathbf{x}(\varepsilon^i s; \varepsilon, \mathbf{x}_0)) \tag{38}$$

exists for all $s \in [0,S]$ and $\mathbf{x}_0 \in \mathscr{A}$, a connected open set. Both $\mathscr{A}$ and S are independent of ε. Changing variables $\tau = \varepsilon^k t$, the general ODE (1) becomes

$$\dot{\mathbf{x}} = \sum_{i=0}^{m-k} \varepsilon^{-i} f_{i+k}(\mathbf{x}) + \mathscr{O}(\varepsilon) \tag{39}$$

Let $\mathbf{x}_0 \in \mathscr{A}$ and consider two solutions of (39) with the same initial condition $\mathbf{x}_0$ but different ε. Using the notations of Section 1, the first solution, obtained with a small parameter ε is denoted by $\mathbf{x}(\tau;\varepsilon,\mathbf{x}_0)$. The second solution, obtained with a small parameter 2ε, is denoted by $\mathbf{x}(\tau;2\varepsilon,\mathbf{x}_0)$. Furthermore, let $\mathbf{x}_1^\varepsilon,\ldots,\mathbf{x}_{2^{m-k}N}^\varepsilon$ denote a numerical solution of $\mathbf{x}(\tau;\varepsilon,\mathbf{x}_0)$ at times $t_j = \varepsilon^{-m+k} H j$ that are computed by some integrator with constant step size $\varepsilon^{m-k}H$. Here, H is a constant which is independent of ε. Similarly, let $\mathbf{x}_1^{2\varepsilon},\ldots,\mathbf{x}_N^{2\varepsilon}$ denote a numerical solution that approximate $\mathbf{x}(\tau;2\varepsilon,\mathbf{x}_0)$ at times $s_j = (2\varepsilon)^{-m+k} H j$. Note that $s_j = t_{2^{m-k}j}$. Then, if $p(\mathbf{x})$ is slow of order ε^k we have that

$$|p(\mathbf{x}_{2^{m-k}j}^\varepsilon) - p(\mathbf{x}_j^{2\varepsilon})| = o(1) \tag{40}$$

for all $j = 1\ldots N$, N independent of ε.

Consider,

$$Q(p) = \sum_{i=0}^{N} \left[p(\mathbf{x}_{2^m j}^\varepsilon) - p(\mathbf{x}_j^{2\varepsilon})\right]^2. \tag{41}$$

Then, a polynomial $p(\mathbf{x})$ that minimizes $Q(p)$ is a good candidate for a variable that evolves on the ε^k scale. Since $Q(p)$ is quadratic in the coefficients of p, the minimizer can be found using least squares.

Finally, the process described above can be repeated, starting with the slowest order ε^0 and gradually moving to faster time scales. Each time a slow variables is identified, one can add a penalty that forces subsequent minimizers to be orthogonal to it in the space of polynomial coefficients. The method is described in [1].

References

1. G. Ariel, B. Engquist, and R. Tsai. A multiscale method for highly oscillatory ordinary differential equations with resonance. *Math. Comp.*, 78:929–956, 2009.
2. G. Ariel, B. Engquist, and R. Tsai. Numerical multiscale methods for coupled oscillators. *Multi. Mod. Simul.*, 7:1387–1404, 2009.
3. G. Ariel, B. Engquist, and R. Tsai. A reversible multiscale integration method. *Comm. Math. Sci.*, 7:595–610, 2009.
4. V.I. Arnol'd. *Mathematical methods of classical mechanics*. Springer-Verlag, New York, 1989.
5. Z. Artstein, I. G. Kevrekidis, M. Slemrod, and E. S. Titi. Slow observables of singularly perturbed differential equations. *Nonlinearity*, 20(11):2463–2481, 2007.
6. Z Artstein, J. Linshiz, and E. S. Titi. Young measure approach to computing slowly advancing fast oscillations. *Multiscale Model. Simul.*, 6(4):1085–1097, 2007.

7. M. P. Calvo and J. M. Sanz-Serna. Instabilities and inaccuracies in the integration of highly oscillatory problems. *SIAM J. Sci. Comput.*, 31(3):1653–1677, 2009.
8. W. E. Analysis of the heterogeneous multiscale method for ordinary differential equations. *Commun. Math. Sci.*, 1(3):423–436, 2003.
9. W. E and B. Engquist. The heterogeneous multiscale methods. *Commun. Math. Sci.*, 1(1):87–132, 2003.
10. W. E, B. Engquist, X. Li, W. Ren, and E. Vanden-Eijnden. Heterogeneous multiscale methods: A review. *Comm. Comput. Phys.*, 2:367–450, 2007.
11. W. E, D. Liu, and E. Vanden-Eijnden. Analysis of multiscale methods for stochastic differential equations. *Commun. on Pure and Applied Math.*, 58:1544–1585, 2005.
12. B. Engquist and Y.-H. Tsai. Heterogeneous multiscale methods for stiff ordinary differential equations. *Math. Comp.*, 74(252):1707–1742, 2005.
13. I. Fatkullin and E. Vanden-Eijnden. A computational strategy for multiscale chaotic systems with applications to Lorenz 96 model. *J. Comp. Phys.*, 200:605–638, 2004.
14. B. García-Archilla, J. M. Sanz-Serna, and R. D. Skeel. Long-time-step methods for oscillatory differential equations. *SIAM J. Sci. Comput.*, 20(3):930–963, 1999.
15. C. W. Gear and I. G. Kevrekidis. Projective methods for stiff differential equations: problems with gaps in their eigenvalue spectrum. *SIAM J. Sci. Comput.*, 24(4):1091–1106 (electronic), 2003.
16. C. W. Gear and I. G. Kevrekidis. Constraint-defined manifolds: a legacy code approach to low-dimensional computation. *J. Sci. Comput.*, 25(1-2):17–28, 2005.
17. C.W. Gear and K.A. Gallivan. Automatic methods for highly oscillatory ordinary differential equations. In *Numerical analysis (Dundee, 1981)*, volume 912 of *Lecture Notes in Math.*, pages 115–124. Springer, 1982.
18. H.-O. Kreiss. Problems with different time scales for ordinary differential equations. *SIAM J. Numer. Anal.*, 16(6):980–998, 1979.
19. H.-O. Kreiss. Problems with different time scales. In *Acta numerica, 1992*, pages 101–139. Cambridge Univ. Press, 1992.
20. H.-O. Kreiss and J. Lorenz. Manifolds of slow solutions for highly oscillatory problems. *Indiana Univ. Math. J.*, 42(4):1169–1191, 1993.
21. G. Papanicolaou. Introduction to the asymptotic analysis of stochastic equations. In *Modern modeling of continuum phenomena*, volume 16 of *Lectures in Applied Mathematics*, pages 47–109. Amer. Math. Soc., Providence, RI, 1977.
22. G. A. Pavliotis and A. M. Stuart. *Multiscale Methods: Averaging and Homogenization*. Number 53 in Texts in Applied Mathematics. Springer-Verlag, New York, 2008.
23. R.L. Petzold, O.J. Laurent, and Y. Jeng. Numerical solution of highly oscillatory ordinary differential equations. *Acta Numerica*, 6:437–483, 1997.
24. J. M. Sanz-Serna. Modulated Fourier expansions and heterogeneous multiscale methods. *IMA J. Numer. Anal.*, 29(3):595–605, 2009.
25. R.E. Scheid. The accurate numerical solution of highly oscillatory ordinary differential equations. *Mathematics of Computation*, 41(164):487–509, 1983.
26. M. Tao, H. Owhadi, and J. Marsden. Non-intrusive and structure preserving multiscale integration of stiff odes, sdes and hamiltonian systems with hidden slow dynamics via flow averaging. *To appear*, 2010.
27. E. Vanden-Eijnden. Numerical techniques for multi-scale dynamical systems with stochastic effects. *Comm. Math. Sci.*, 1:385–391, 2003.
28. E. Vanden-Eijnden. On hmm-like integrators and projective integration methods for systems with multiple time scales. *Commun. Math. Sci.*, 5:495–505, 2007.