Multiscale Computations for Highly Oscillatory Problems

Gil Ariel[1,*], Björn Engquist[1,2,*], Heinz-Otto Kreiss[3], and Richard Tsai[1,*]

[1] The University of Texas at Austin, Austin, TX 78712,
{ariel,engquist,ytsai}@math.utexas.edu
[2] Department of Numerical Analysis, KTH, 100 44 Stockholm, Sweden
[3] Träskö–Storö Institute of Mathematics

Summary. We review a selection of essential techniques for constructing computational multiscale methods for highly oscillatory ODEs. Contrary to the typical approaches that attempt to enlarge the stability region for specialized problems, these lecture notes emphasize how multiscale properties of highly oscillatory systems can be characterized and approximated in a truly multiscale fashion similar to the settings of averaging and homogenization. Essential concepts such as resonance, fast-slow scale interactions, averaging, and techniques for transformations to non-stiff forms are discussed in an elementary manner so that the materials can be easily accessible to beginning graduate students in applied mathematics or computational sciences.

1 Introduction

Oscillatory systems constitute a broad and active field of scientific computations. One of the typical numerical challenges arises when the frequency of the oscillations is high compared to either the time or the spatial scale of interest. In this case, the cost for computations can typically become exceedingly expensive due to the need of sampling oscillations adequately by numerical discretizations over a relatively large domain. Several general strategies for dealing with oscillations can be found in literature, for example, asymptotic analysis [5, 22, 23], averaging [20, 29], envelope tracking [27, 28], explicit solutions to nearby oscillatory problems [25, 30, 31]. These strategies typically utilize some underlying structures, related to the oscillations, which are not oscillatory in the domain of interests. For example, the center or frequency of oscillators may vary slowly in time. Indeed, it is often the case that the quantities of interest are related to these non-oscillatory structures. Reduction in the computational costs is thus possible by avoiding direct resolution of the oscillations. Take geometrical optics [13, 21] for instance. The high frequency solution of the wave equation of the form $A(x,t)\exp(S(x,t)/\varepsilon)$ is computed via solutions of an eikonal equation for the phase S and transport equations for the amplitude A. Since eikonal and transport equations do not depend on the ε-scale oscillations, the cost of

* Ariel, Engquist, and Tsai are partially supported by NSF grant DMS-0714612.

computation is formally independent of the fast scale as well. These current lecture notes focus on building efficient multiscale numerical methods that only sample the fast oscillations. The sampled information is used to describe an effective time evolution for the system at longer time scales. The general approach underlying these methods come from the theory of averaging.

In these notes, we consider systems of ordinary differential equations (ODEs) which evolve on two widely separated time scales. Common examples are

1. Perturbed linear oscillations:
$$\varepsilon x' = Ax + \varepsilon g(x), \tag{1}$$
where A is diagonalizable and has purely imaginary eigenvalues. The class of examples include Newton's equation of motion for perturbed harmonic oscillators
$$\varepsilon x'' = -\Omega^2 x + \varepsilon g(x),$$
which are found in many applications. Here, the parameter ε, $0 < \varepsilon \ll 1$, characterizes the separation of time scales in the system.
2. Fully nonlinear oscillations induced from dissipation in the systems. Examples include Van der Pol oscillators and other relaxation oscillators.
3. Weakly coupled nonlinear oscillators that are close to a slowly varying periodic orbits. Examples include Van der Pol (with small damping) and Volterra-Lotka oscillators.

Efficient and accurate computations of oscillatory problems require significant knowledge about the underlying fast oscillations. Using either analytical or numerical methods, our general underlying principle is to model oscillations and sample their interactions. Very often, analytical methods do not yield explicit solutions, and suitable numerical methods need to be applied.

One of the current major thrusts is in developing numerical methods which allow long time computation of oscillatory solutions to Hamiltonian systems. The interest in such systems comes from molecular dynamics which attempts to simulate some underlying physics on a time scale of interest. These methods typically attempt to approximately preserve some analytical invariance of the solutions; e.g. the total energy of the system, symplectic structures, or the reversibility of the flow. Detailed reviews and further references on this active field of "geometric integration" can be found in [18] and [26].

The Verlet method and other similar geometric integrators are the methods of choice for many highly oscillatory simulations. They require, however, time steps that are shorter than the oscillatory wavelength ε and therefore cannot be used when ε is very small.

Exponential integrators allow for time steps that are longer than the oscillatory wavelength ε but they apply only to restricted classes of differential equations [18]. In a way that resembles the discussion of geometrical optics above since these methods explicitly use the exponential function to represent the leading terms in the oscillations. They work well for problems that are smooth perturbations of problems with constant coefficients.

Another general approach for dealing with multiscale phenomena computationally can be referred to as boosting [33]. The general idea is to artificially "twig" or "boost" the small parameter ε so that the stiffness of the problem is reduced. Known methods that fall into this category are Chorin's artificial compressibility [7] and the Car-Parrinello method used in molecular dynamics [6].

This tutorial deviates from previous texts in that we do not rely or assume some specific properties or a special class of ODEs such as harmonic oscillations or Hamiltonian dynamics. Instead, the multiscale methods discussed here compute the effective behavior of the oscillatory system by integrating the oscillations numerically in *short time windows* and sampling their interactions by suitable averaging. Indeed, one of the main goals of this text is to make the ideas discussed above mathematically meaningful. Subsequent sections will define what we mean by the effective behavior of a given highly oscillatory system, describe the theory of averaging, the structure of our multiscale algorithms and its computational complexity.

The objectives of multiscale computations

One of the major challenges in problems involving multiple scales is that an accurate computations, attempting to resolve the finest scales involved in the dynamics may be computationally infeasible. In the classical numerical analysis for ODEs, the important elements are stability, consistency and ultimately convergence. In the standard theory, any stable consistent method converges to the analytical solution as the step size goes to zero. The errors depend on powers of the eigenvalues of the Jacobian of the ODE's right hand side and the step size. However, in our multiscale setting, similar to the high frequency wave propagation or homogenization, we would like to consider the asymptotic cases when the frequency of the fastest oscillations, which is proportional to $1/\varepsilon$, tends to infinity, *before the step size is sent to zero.* Hence, we need to rethink what consistency and convergence means in the multiscale setting. One possibility is the following: let $E(t;\triangle,\varepsilon)$ denote the error of the numerical approximation at time t, using step size $\triangle$ and for problems with ε^{-1} oscillations, we consider the convergence of E for $0 < \varepsilon < \varepsilon_{\triangle}$

$$\lim_{\triangle \longrightarrow 0} \left(\sup_{0<\varepsilon<\varepsilon_{\triangle}} E(t;\triangle,\varepsilon) \right).$$

In other words, with a prescribed error tolerance E, the same step size Δ can be used for small enough ε. While this notion of convergence may not be possible for the solutions of many problems, we may ask for the convergence of some functions or functionals of the solutions. Throughout the notes we discuss results from the prespective of a few key questions: what is the motivation for constructing a multiscale algorithm? What is being approximated? How does our multiscale approach differ from traditional numerical computations?

A first example, suggested by Germund Dahlquist, is the drift path of a mechanical alarm clock, moving due to fast vibrations when it is set off on a hard surface. If the drift path depends only locally in time on the fast oscillations, then it is reasonable

to design a scheme that evolves the slowly changing averaged drift path by measuring the effects of the fast solutions only locally in time. Herein lies the possibility of reducing the computational complexity.

A second example is Kapitza's pendulum — a rigid pendulum whose pivot is attached to a strong periodic forcing is vibrating vertically with amplitude ε and frequency $1/\varepsilon$. When the oscillations are sufficiently fast, the pendulum swings slowly back and forth, *pointing upwards*, with a slow period that is practically independent of ε. Obviously, in the absence of the oscillatory forcing, the pendulum is only stable pointing downwards. Pyotr Kapitza (physics Nobel Laureate in 1978) used this example to illustrate a general stabilization mechanisms [17]. This, and similar simple dynamical systems are often used as example benchmark problems to study how different methods approximates highly oscillatory problems.

The following assumptions are made throughout these notes: in the fastest scale, the given system exhibits oscillations with amplitudes independent of ε, and that at a larger time scale, some slowly changing quantities can be defined by the oscillatory solutions of the system. To facilitate our discussion, we now present our model scenario described by the following two coupled systems. Consider a highly oscillatory system in $\mathbb{R}^{d_1}$ coupled with a slow system in $\mathbb{R}^{d_2}$:

$$\varepsilon x' = f(x,v,t) + \varepsilon g(x,v,t), \tag{2}$$

$$v' = h(v,x,t), \qquad x(0) = x_0 \in \mathbb{R}^{d_1}, \quad v(0) = v_0 \in \mathbb{R}^{d_2}. \tag{3}$$

We assume that x is highly oscillatory, and v is the slow quantity of interest. However, without proper information about x, v cannot be found. We are also interested in some slowly varying quantity that is being defined along the trajectories of x:

$$\beta' = \psi(\beta,t;x(\,\cdot\,)). \tag{4}$$

In a following section, we shall see that for very special initial conditions, the solutions of (2) may be very smooth and exhibit no oscillations. The problem of initialization, i.e. finding the suitable initial data so that the slowly varying solutions can be computed appear in meteorology. We refer the readers to the paper of Kreiss and Lorenz [25] for further reading on the theory for the slow manifolds. However, in many autonomous equations, e.g. linear equations, the only slowly varying solutions in the system are the equilibria of the system. For problems like the inverted pendulum, it is clear that the slowly varying solutions are not of interest. Then some complicated interactions between the oscillations must take place, and one must look into different strategies in order to characterize the effective influence of the oscillations in x in the evolution of v.

Our objective is to accurately compute the slowly changing quantity v in a long time scale (i.e. $0 \le t \le T$, for some constant T independent of ε). Furthermore, we wish to compute it with a cost that is at least sublinear to (ideally independent of) the cost for resolving all the fast oscillations in this time scale. In general, our objective may be achieved if fast oscillations are computed only in very short time intervals and yet the dynamics for those slowly changing quantities is consistently evolved. Figure 1 depicts two possible schematic structures for such an algorithm. In this section, we

give a few examples of where such type of slowly changing quantities occur; these consist of systems with resonant modes and weakly perturbed Hamiltonians systems so that invariances are changing slowly. Furthermore, we shall see in the next section that these slowly changing quantities may be evaluated conveniently by short time averaging with suitable kernels.

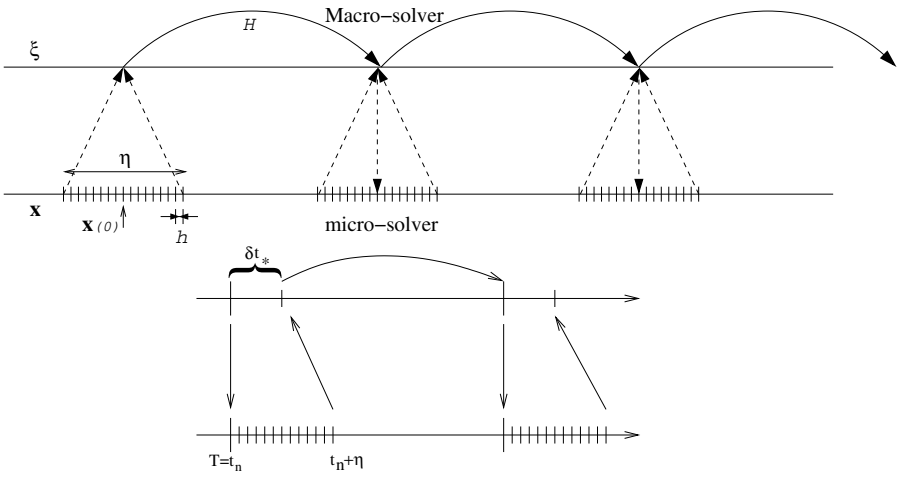


Fig. 1.

Let us briefly comment on the relation of these notes to the standard stiff ODE solvers for multiscale problems with transient solutions, [8, 19]. A typical example of such stiff problems is equation (1) where the eigenvalues of A are either negative or zero. The initial time steps are generally small enough to resolve the transient and of the type of the micro-solver in Fig. 1. After the transient, much longer time steps are possible as in the macro-solver in Fig. 1. Special properties of stiff ODE methods suppress the fast modes and only the slower modes need to be well approximated. Problems with highly oscillatory solutions are much harder to simulate since the fast modes are present for all times and may interact to give contributions to the slower modes.

1.1 Example oscillatory problems

Linear systems with imaginary eigenvalues

$$x' = i\lambda x, \qquad \lambda \in \mathbb{R}.$$

The solution is readily given by $x(t) = x(0)e^{i\lambda t}$. Note that this system is equivalent to the system in $\mathbb{R}^2$:

$$\begin{pmatrix} x \\ y \end{pmatrix}' = \begin{pmatrix} 0 & \lambda \\ -\lambda & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

Hamiltonian systems

Hamiltonian dynamics are defined by the partial derivatives of a Hamiltonian function $H(q,p)$ which represent the total energy of the system. Here, q is a generalized coordinate system and p the associated momentum. The equations of motion are given by

$$\begin{aligned} q' &= H_p(q,p), \\ p' &= -H_q(q,p), \end{aligned} \tag{5}$$

where H_p and H_q denote partial derivatives of H with respect to p and q, respectively. In Hamiltonian mechanics, $H(p,q) = \frac{1}{2}p^2 + V(q)$ and the dynamics defined in (5) yield Newton's equation of motion $q'' = -\nabla_q V(q)$. If $V(x)$ is a convex function then the solutions of this equation are typically oscillatory. An important class of equations of this type appear in molecular dynamics with pairwise potentials

$$H(p,q) = \frac{1}{2}\sum_{i=1}^{N}\frac{1}{m_i}p_i^T p_i + \frac{1}{2}\sum_{i,j=1}^{N} V_{ij}(|q_i - q_j|),$$

where p_i and q_i are components of the vectors p and q.
Notable examples are

$$V_{ij}(r) = \frac{-Gm_i m_j}{r} \qquad \text{(electric or gravitational potential)}$$

and

$$V_{ij}(r) = 4\varepsilon_{ij}\left(\left(\frac{\sigma_{ij}}{r}\right)^{12} - \left(\frac{\sigma_{ij}}{r}\right)^{6}\right), \qquad \text{(Lennard-Jones potential)}$$

for all $i \neq j$, etc.

Volterra–Lotka

This is a simplified model for the predator-prey problem in population dynamics. In this model, x denotes the population of a predator species while y denotes the population of a prey species

$$\begin{aligned} x' &= x\left(1 - \frac{y}{\nu}\right), \\ y' &= \frac{y}{\nu}(x-1). \end{aligned} \tag{6}$$

An example trajectory is depicted in Fig. 2.

Relaxation oscillators

The Van der Pol oscillator is another typical example of nonlinear oscillators. One version of the equation for a Van der Pol oscillator takes the form

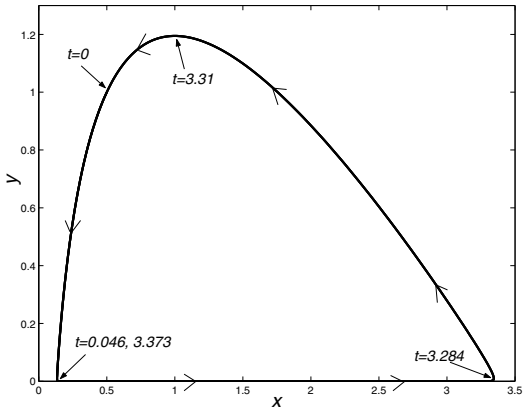


Fig. 2. The trajectory of the Volterra–Lotka oscillator (6) with $\nu = 0.01$, $x(0) = 0.5$ and $y(0) = 1$.

$$x' = y - (x^2 - 1)x,$$
$$y' = -x.$$

This equation can be interpreted as a model of a basic RLC circuit, consisting of a resistor, inductor, and a capacitor; the state variable x corresponds to the current in the inductor and y the voltage in the capacitor. It can be shown that there is a unique periodic solution of this equation and other non-equilibrium solutions approach it as time increases. This periodic solution is called the limit cycle or the invariant manifold of the system. A general result for detecting periodic solutions for such type of systems on a plane is the Poincaré–Bendixson theorem, which says that if a compact limit set in the plane contains no equilibria, it is a closed orbit; i.e. it is a periodic trajectory of a solution.

As a second example, consider [9]

$$x' = -1 - x + 8y^3, \tag{7}$$
$$y' = \frac{1}{\nu}\left(-x + y - y^3\right),$$

where $0 < \nu \ll 1$. For small ν, trajectories quickly come close to the limit cycle defined by $-x + y - y^3 = 0$. The upper and lower branches of this cubic polynomial are stable up to the turning points at which $dx/dy = 0$. For any initial condition, the solution of (7) is rapidly attracted to one of the stable branches on an $\mathcal{O}(\nu)$ time scale. The trajectory then moves closely along the branch until it becomes unstable. At this point the solution is quickly attracted to the other stable branch. The trajectory is depicted in Fig. 3.

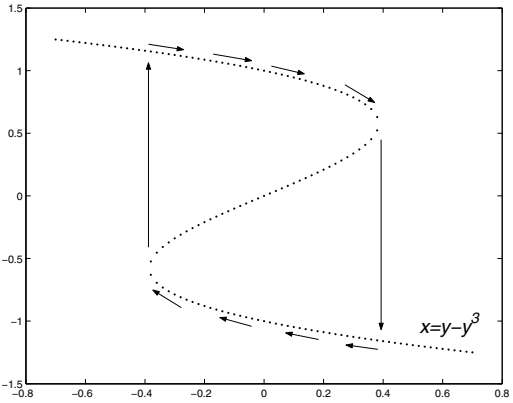


Fig. 3. The trajectory and slow manifold of the relaxation oscillator (7)

1.2 Invariance

Hamiltonian systems:

The Hamiltonian equations of motion (5) admit several invariances. First and foremost is the energy $H(p,q)$

$$\frac{d}{dt}H(p(t),q(t)) = H_p p' + H_q q' = 0$$

$$\Longrightarrow H(p(t),q(t)) = H(p_{0,}q_0) = \text{const.}$$

Let us prove Liouville's theorem on volume preservation of Hamiltonian systems. Consider a smooth Hamiltonian $H(p,q)$. Let

$$\varphi_t(p_0,q_0) = \begin{pmatrix} p(t;p_0,q_0) \\ q(t;p_0,q_0) \end{pmatrix}.$$

Hence,

$$\frac{d}{dt}\frac{\partial \varphi_t}{\partial(p_0,q_0)} = \begin{pmatrix} -H_{pq} & H_{qq} \\ H_{pp} & H_{qp} \end{pmatrix} \begin{pmatrix} \frac{\partial \varphi_t}{\partial(p,q)} \\ (p(t),q(t)) \end{pmatrix}, \qquad q,p \in \mathbb{R}$$

$$\Longrightarrow \frac{d}{dt}\det\frac{\partial \varphi_t}{\partial(p_0,q_0)} = 0. \tag{8}$$

Consider t as a parameter for the family of coordinate changes (diffeomorphisms) $\phi_t : (p_0,q_0) \mapsto (p(t;p_0,q_0),q(t,p_0,q_0))$. Then we have the following change of coordinates formula, for any fixed t,

$$\int_V f(p,q)dqdp = \int_U f(\phi_t(p_0,q_0))Jdp_0dq_0,$$

where $V = \phi_t(U)$ and

$$J := \det \frac{\partial \phi_t}{\partial (p_0, q_0)}.$$

Thus, (8) implies that

$$\frac{dJ}{dt} \equiv 0$$

In particular, taking $f \equiv 1$ implies that U and V have the same volume in the phase (p,q)-space.

Volterra-Lotka oscillators

Let $I(u,v) = \log u - u + 2\log v - v$. Substituting (6) yields $(d/dt)I(u(t), v(t)) = 0$ for $t > 0$.

Relative phase between two linear oscillators

Let $u(t) = (\cos(t), \sin(t))$ and $v(t) = (\cos(t+\phi_0), \sin(t+\phi_0))$ be the solutions of some oscillators. Then

$$\xi(t) = u(t) \cdot v(t) = \cos\phi_0$$

measures the phase difference between $u(t)$ and $v(t)$ and remains constant in time.

In view of the above examples, the following questions naturally appear:

- Can one design numerical schemes so that the important invariances are preserved?
  What is the computational cost or benefit?
  How well do common numerical approximations preserve known invariances of interest and for what time scale?
- What is the importance of preserving invariances? How can this notion be quantified?
- How do small perturbations affect the invariances? For example, in the following linear system which conserves energy for $\varepsilon = 0$: $x'' = -\omega^2 x + \varepsilon\cos(\lambda t)$. How does weak periodic forcing affect the energy? At what time scale does the forcing become important? Can these effects be computed efficiently?

Some aspects of these issues and others are discussed in [18] and [26].

1.3 Resonance

Resonances among oscillations appear in many situations. For example, in pushing a child on a playground swing. It is intuitively clear that unless the swing is pushed at a frequency which is close to the natural oscillation frequency of the swing, the child will be annoyed. However, when the swing is pushed at the right frequency, the amplitude of the swing is gradually increasing. In this subsection, we review a few basic examples of resonance.

Resonance in a forced linear spring

We start with a linear spring under periodic forcing,

$$x'' = -\omega^2 x + \cos\lambda t.$$

Rewriting into a first order system, we obtain

$$\frac{d}{dt}\begin{pmatrix} x \\ x' \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -\omega^2 & 0 \end{pmatrix}\begin{pmatrix} x \\ x' \end{pmatrix} + \begin{pmatrix} 0 \\ \cos\lambda t \end{pmatrix},$$

with initial condition $\begin{pmatrix} x_0 \\ x_0' \end{pmatrix}$. One can show that the solution operator for the homogeneous problem is

$$S_t = \begin{pmatrix} \cos\omega t & \omega^{-1}\sin\omega t \\ -\omega\sin\omega t & \cos\omega t \end{pmatrix}$$

so that the solution for the inhomogeneous problem is

$$\begin{pmatrix} x \\ x' \end{pmatrix} = S_t\begin{pmatrix} x_0 \\ x_0' \end{pmatrix} + \int_0^t S_{t-s}\begin{pmatrix} 0 \\ \cos\lambda s \end{pmatrix} ds$$

$$\Longrightarrow \begin{pmatrix} x \\ x' \end{pmatrix} = S_t\begin{pmatrix} x_0 \\ x_0' \end{pmatrix} + \int_0^t \begin{pmatrix} \omega^{-1}\sin(\omega t-\omega s)\cos\lambda s \\ \cos(\omega t-\omega s)\cos\lambda s \end{pmatrix} ds.$$

When $\lambda = \omega$, resonance happens. More precisely, we see that

$$x(t) = x_0\cos\omega t + \frac{x_0'}{\omega}\sin\omega t + \frac{1}{\omega}\int_0^t \sin(\omega t-\omega s)\cos\omega s\, ds$$

$$\Longrightarrow x(t) = x_0\cos\omega t + \frac{x_0'}{\omega}\sin\omega t + \frac{t}{2}\sin\omega t.$$

In addition,

$$\begin{aligned}
\int_0^t \sin(\omega t-\omega s)\cos\omega s\, ds &= \int_0^t (\sin\omega t\cos\omega s - \sin\omega s\cos\omega t)\cos\omega s\, ds \\
&= \sin(\omega t)\int_0^t \cos^2(\omega s)ds - \cos(\omega t)\int_0^t \sin(\omega s)\cos(\omega s)ds \\
&= \sin(\omega t)\int_0^t \frac{1}{2}(1+\cos 2\omega s)ds - \cos(\omega t)\int_0^t \frac{1}{2}\sin(2\omega s)ds \\
&= \underbrace{\frac{t}{2}\sin\omega t}_{\text{result of resonance}} + \frac{1}{2}\int_0^t (\sin(\omega t)\cos(2\omega s) - \cos(\omega t)\sin(2\omega s))\, ds \\
&= \underbrace{\frac{t}{2}\sin\omega t}_{\text{result of resonance}} + \underbrace{\frac{1}{2}\int_0^t \sin(\omega t-2\omega s)ds}. \\
&= 0
\end{aligned}$$

If $\lambda^2 \neq \omega^2$,we have

$$x = \left(x_0 - \frac{1}{\omega^2-\lambda^2}\right)\cos\omega t + \frac{x_0'}{\omega}\sin\omega t + \frac{1}{\omega^2-\lambda^2}\cos\lambda t.$$

Exercise 1. Compute the solution of the forced oscillation under friction: for $\mu > 0$

$$x'' = -2\mu x' - \omega^2 x + \cos \lambda t.$$

Show that if $\lambda = \omega$, the amplitude of the oscillations remains bounded and is largest when $\lambda = \sqrt{\omega^2 - 2\mu^2}$.

Resonance in first order systems

Consider,

$$x' = \frac{i}{\varepsilon} \Lambda x + f\left(x, \frac{t}{\varepsilon}\right), \qquad x(0,\varepsilon) = x_0,$$

where Λ is a diagonal matrix. We make the substitution:

$$x(t) = e^{\frac{i}{\varepsilon}\Lambda t} w(t), \qquad w(0) = x_0,$$

and obtain the corresponding equation for w:

$$w' = e^{-\frac{i}{\varepsilon}\Lambda t} f\left(e^{\frac{i}{\varepsilon}\Lambda t} w, \frac{t}{\varepsilon}\right), \qquad w(0) = x_0.$$

The simplest type of resonance can be obtained by taking $f(x,t/\varepsilon) = x$, i.e., $w' = w$. We see that the resonance occurs due to the linearity of f in x, and that it results in $|x|$ changing at a rate independent of ε. If $f(x,t/\varepsilon) = f_I(x) + \exp(it/\varepsilon)$ and one of the diagonal elements of Λ is 1, then similar to the resonance in the forced linear spring, the resonance with the forcing term contributes a linear in time growth of $|x|$.

Resonances may occur due to nonlinear interaction. Following the above example, take

$$\Lambda = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \qquad \text{and} \qquad f(x,t) = \begin{pmatrix} 0 \\ -x_1^4 x_2^{-1} \end{pmatrix}.$$

Hence,

$$x = \begin{pmatrix} e^{it/\varepsilon} w_1 \\ e^{2it/\varepsilon} w_2 \end{pmatrix} \Longrightarrow$$

$$w' = \begin{pmatrix} e^{-it/\varepsilon} & 0 \\ 0 & e^{-2it/\varepsilon} \end{pmatrix} \begin{pmatrix} 0 \\ -e^{4it/\varepsilon} e^{-2it/\varepsilon} w_1^4 w_2^{-1} \end{pmatrix} = \begin{pmatrix} 0 \\ -w_1^4 w_2^{-1} \end{pmatrix}.$$

Again, due to the resonance in the system, $|x_2|$ is changing at a rate that is independent of ε.

2 Slowly varying functions of the solutions

In this section we shall study the effect of non-linear interactions. We excerpt important results from [24] and [1, 2, 14].

2.1 Problems with dominant fast linear oscillations and nonlinear interactions

We start with a number of examples.

$$x' = \frac{i\lambda}{\varepsilon}x + x^2, \qquad x(0) = x_0. \tag{9}$$

The solution of (9) can be obtained explicitly. Introducing a new variable x *by*

$$x = e^{\frac{i\lambda}{\varepsilon}t}w$$

gives us a new equation whose right hand side is bounded independent of ε

$$w' = e^{\frac{i\lambda}{\varepsilon}t}w^2, \qquad w(0) = w_0 = x_0.$$

The solution is readily given by

$$w(t) = \frac{1}{\frac{1}{w_0} + \frac{i\varepsilon}{\lambda}(e^{\frac{i\lambda}{\varepsilon}t} - 1)} = \frac{w_0}{1 + \frac{i\varepsilon}{\lambda}w_0(e^{\frac{i\lambda}{\varepsilon}t} - 1)}.$$

As a result,

$$w(t) = w_0\left(1 - \frac{i\varepsilon}{\lambda}w_0(e^{\frac{i\lambda}{\varepsilon}t} - 1)\right) + \mathcal{O}(\varepsilon^2). \tag{10}$$

Thus, the nonlinear term changes the solution only by $\mathcal{O}(\varepsilon)$ in arbitrarily long time intervals. In Sect. 3.2, we will show that $w(t)$ is close to an effective equation from averaging:

$$\bar{w}' = \bar{f}(w), \qquad \bar{f}(w) = \int_0^1 e^{it}w^2 dt = 0.$$

Hence, $\bar{w}(t) = w_0$. Indeed, we see that $|w(t) - \bar{w}(t)| = |w(t) - w_0| \le C_0\varepsilon$ for $0 \le t \le T_1$.

An alternative solution method involves a procedure which is easier to generalize. From (9) we have,

$$\begin{aligned} w(t) - w_0 &= \int_0^t e^{\frac{i\lambda}{\varepsilon}s}w^2 ds = -\frac{i\varepsilon}{\lambda}e^{\frac{i\lambda}{\varepsilon}s}w^2\Big|_0^t + \frac{2i\varepsilon}{\lambda}\int_0^t e^{\frac{i\lambda}{\varepsilon}s}ww' ds \\ &= -\frac{i\varepsilon}{\lambda}\left(e^{\frac{i\lambda}{\varepsilon}t}w^2(t) - w_0^2\right) + \frac{2i\varepsilon}{\lambda}\int_0^t e^{i\frac{2\lambda}{\varepsilon}s}w^3 ds. \end{aligned} \tag{11}$$

Integrating by parts again yields an integral equation for $w(t)$

$$w(t) + \frac{i\varepsilon}{\lambda}e^{\frac{i\lambda}{\varepsilon}t}w^2(t) - \frac{4\varepsilon^2}{\lambda^2}e^{i\frac{2\lambda}{\varepsilon}}w^3(t) = w_0 + \frac{i\varepsilon}{\lambda}w_0^2 - \frac{4\varepsilon^2}{\lambda^2}w_0^3 + \frac{4\varepsilon^2}{\lambda^2}\int_0^t e^{i\frac{3\lambda}{\varepsilon}s}w^4(s)ds.$$

The solution $w(t)$ can then be constructed using fixed point iterations

$$w_{(k+1)}(t) = F(w_{(k)}, w_0, t) + \frac{4\varepsilon^2}{\lambda^2}\int_0^t e^{i\frac{3\lambda}{\varepsilon}s}w_{(k)}^4(s)ds, \qquad k = 0, 1, 2, \cdots, \tag{12}$$

where $w_{(0)} = w_0$ and

$$F(w,w_0,t) = w_0 - \frac{i\varepsilon}{\lambda}e^{\frac{i\lambda}{\varepsilon}t}w^2(t) + \frac{i\varepsilon}{\lambda}w_0^2 + \frac{4\varepsilon^2}{\lambda^2}e^{i\frac{2\lambda}{\varepsilon}}w^3(t) - \frac{4\varepsilon^2}{\lambda^2}w_0^3.$$

By induction, one can show that the iterations converge for $t \geq 0$, $0 \leq \varepsilon \leq \varepsilon_0$ and

$$w(t) + \frac{i\varepsilon}{\lambda}e^{\frac{i\lambda}{\varepsilon}t}w^2(t) = w_0 + \frac{i\varepsilon}{\lambda}w_0^2 + \mathcal{O}(\varepsilon^2).$$

From (11), we have $w(t) = w_0 + \mathcal{O}(\varepsilon)$. Hence,

$$w(t) = w_0\left(1 - \frac{i\varepsilon}{\lambda}w_0\left(e^{\frac{i\lambda}{\varepsilon}t} - 1\right)\right) + \mathcal{O}(\varepsilon^2). \tag{13}$$

Now, consider

$$x' = \frac{i\lambda_1}{\varepsilon}x + y, \qquad y' = \frac{i\lambda_2}{\varepsilon}y + y^2.$$

Changing variables to

$$x = e^{\frac{i\lambda_1}{\varepsilon}t}u, \qquad y = e^{\frac{i\lambda_2}{\varepsilon}t}w$$

yields

$$u' = e^{i(-\lambda_1+\lambda_2)t/\varepsilon}w, \quad w' = e^{\frac{i\lambda_2}{\varepsilon}t}w^2.$$

From (10), we can obtain an asymptotic expansion for w:

$$w = w_0 + \sum_{j=1}^{\infty}\varepsilon^j\beta^j e^{i\frac{j\lambda_2}{\varepsilon}t}.$$

Therefore:

$$u' = e^{\frac{i}{\varepsilon}(\lambda_2-\lambda_1)t}w_0 + \sum_{j=1}^{\infty}\varepsilon^j\beta^j e^{\frac{i}{\varepsilon}((j+1)\lambda_2-\lambda_1)t}.$$

If $\nu\lambda_2 - \lambda_1 \neq 0$ for all $\nu = 1,2,\ldots$, then

$$u(t) = u_0 + \mathcal{O}(\varepsilon).$$

However, if $\nu\lambda_2 = \lambda_1$, then resonance occurs and

$$u(t) = \begin{cases} u_0 + \varepsilon^{\nu-1}\beta^{\nu-1}t, & \text{if } \nu > 1, \\ u_0 + tw_0, & \text{if } \nu = 1. \end{cases}$$

Thus, the solution is not bounded for all time.

As a generalization, consider the system

$$x' = \frac{i}{\varepsilon}\Lambda x + P(x), \qquad x(0) = x_0, \tag{14}$$

where

$$\Lambda = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_d \end{pmatrix}, \qquad \lambda_1, \cdots, \lambda_d \in \mathbb{R},$$

and $P = (p_1(x), \cdots, p_d(x))$ is a vector of polynomials in $x = (x_1, \cdots, x_d)$. Let

$$x = e^{\frac{i}{\varepsilon}\Lambda t} w.$$

We then have

$$w' = e^{-\frac{i}{\varepsilon}\Lambda t} P\left(e^{\frac{i}{\varepsilon}\Lambda t} w\right), \qquad w(0) = w_0 = x_0. \tag{15}$$

The right hand side of (15) consists of expressions of the form

$$e^{\frac{i}{\varepsilon}(\Sigma m_j \lambda_j)t} p(w), \tag{16}$$

where the m_j are integers and p is a polynomial in w. There are two possibilities.

1. $\tau = \sum m_j \lambda_j = 0$ for some terms. We call these terms the resonant modes. In this case (15) has the form

$$w' = Q_0(w) + Q_\varepsilon\left(\frac{t}{\varepsilon}, w\right), \tag{17}$$

where Q_0 contains the terms corresponding to resonant modes, and Q_ε the remaining terms involving oscillatory exponentials. One can show that the solution of

$$\bar{w}' = Q_0(\bar{w}), \qquad \bar{w}(0) = w_0 = x_0, \tag{18}$$

is very close to w for a long time; i.e.

$$|w(t) - \bar{w}(t)| \le C_0 \varepsilon, \qquad 0 \le t \le T_1, \tag{19}$$

and in general $w(t)$ does not stay close to the initial value w_0.
2. $\tau = \sum m_j \lambda_j \neq 0$ for all terms. No resonance occurs in the system. The solution stays close to the initial value:

$$w(t) = w_0 + \mathcal{O}(\varepsilon). \tag{20}$$

We remark here that the term

$$f\left(\frac{t}{\varepsilon}, w\right) = e^{-\frac{i}{\varepsilon}\Lambda t} P\left(e^{\frac{i}{\varepsilon}\Lambda t} w\right)$$

is in general not strictly periodic, even though it is composed of many highly oscillatory terms. Nonetheless, the self averaging effect of the highly oscillatory terms can be observed using integration by parts:

$$\begin{aligned} w(t) &= w_0 + \sum_\tau \int_0^t e^{\frac{i\tau}{\varepsilon}\xi} p_\tau(w) d\xi \\ &= w_0 - i\varepsilon \sum_\tau \frac{1}{\tau} e^{\frac{i\tau}{\varepsilon}\xi} p_\tau(w)|_0^t + i\varepsilon \sum_\tau \frac{1}{\tau} \int_0^t e^{\frac{i\tau}{\varepsilon}\xi} \frac{\partial p_\tau}{\partial p} w' d\xi \\ &= w_0 - i\varepsilon \sum_\tau \frac{1}{\tau} e^{\frac{i\tau}{\varepsilon}\xi} p_\tau(w)|_0^t + i\varepsilon \sum_\tau \frac{1}{\tau} \int_0^t e^{\frac{i\tau}{\varepsilon}\xi} \tilde{p}_\tau(w)' d\xi. \end{aligned} \tag{21}$$

The integrals in (21) are over terms of type (16) and we can therefore repeat the above arguments. If some of the terms are not of exponential type, then they will, in general, be of order $\mathcal{O}(\varepsilon t)$. For $\varepsilon t \ll 1$ we can replace (21) by

$$\tilde{w}(t) = w_0 - i\varepsilon \sum_\tau \frac{1}{\tau} e^{\frac{i\tau}{\varepsilon}\xi} p_\tau(\tilde{w})|_0^t,$$

i.e.,

$$\tilde{w}(t) = w_0 + \mathcal{O}(\varepsilon) \qquad \text{for } \varepsilon t \ll 1.$$

A more accurate result is

$$\tilde{w}(t) = w_0 - i\varepsilon \sum_\tau \frac{1}{\tau} \left(e^{\frac{i\tau}{\varepsilon}\xi} - 1 \right) p_\tau(y_0) + i\varepsilon \tilde{p}_0(y_0) t + \mathcal{O}(\varepsilon^2 t^2).$$

If all the terms are of exponential type, then we can use integration by parts to reduce them at least to order $\mathcal{O}(\varepsilon^2 t)$. We obtain the following theorem.

Theorem 1. *Assume that for all integers α_j the linear combinations*

$$\sum \alpha_j \lambda_j \neq 0.$$

Then

$$\tilde{w} = w_0 + \mathcal{O}(\varepsilon)$$

in time intervals $0 \leq t \leq T$. $T = \mathcal{O}(\varepsilon^{-p})$ for any p.

There are no difficulties in generalizing the result and techniques to more general equations

$$x' = \frac{1}{\varepsilon} \Lambda(t) x + P(x,t).$$

Here $\Lambda(t)$ is slowly varying and $P(x,t)$ is a polynomial in x with slowly varying coefficients in time.

Remark 1. We see that without the presence of resonance, the highly oscillatory solution x of system (14) stays closely to

$$e^{i\frac{\Lambda}{\varepsilon}t} x_0,$$

for a very long time. Regarding to our ultimate goal of developing efficient algorithms, we may conclude that if no resonance occurs in the system, no computation is needed, since $e^{i\frac{\Lambda}{\varepsilon}t} x_0$ is already a good approximation to the solution.

However, if resonance occurs, then the "envelop" of the oscillations in the solution, $x(t)$ changes non-trivially. In this case, efficient algorithms can be devised from solving the initial value problem of equation (17):

$$w' = Q_0(w) + Q_\varepsilon\left(\frac{t}{\varepsilon}, w\right), \qquad w(0) = x_0.$$

As we showed above, one may even simply drop the oscillatory term and solve an equation that is completely independent of the fast scale:

$$w' = Q_0(w), \qquad w(0) = x_0$$

and still obtain accurate approximations. We refer the readers to the paper of Scheid [31] for an interesting algorithm that explores this special structure of the right hand side. In a later part of these notes, we shall first show that the term Q_ε can be easily "averaged" out without even using its explicit form. This may prove to be very useful for designing our multiscale algorithms for more complicated systems.

Exercise 2. Show that the fixed-point iterations defined in (12) converge for arbitrary time intervals. Follow the steps:

1. For $0 \le t < T$, the difference $e_k(t) := w_{(k)}(t) - w_{(k-1)}(t)$ converge to 0 as k approaches infinity.
2. Show that $w_{(k)}$ is uniformly bounded for $k = 1, 2, \ldots$.
3. Establish the estimate in (13).
4. Arguments in the previous steps can be repeated to extend the solution to larger time intervals.

2.2 Slowly varying solutions

Consider

$$\begin{aligned} \varepsilon x' &= (A(t) + \varepsilon B(x,t))x + F(t), \\ x(0) &= x_0, \end{aligned} \tag{22}$$

where $0 < \varepsilon \ll 1$ and $A(t)$, $F(t)$ satisfy the same conditions as in the linear case, i.e. A, A^{-1}, F and their derivatives are of order one. Also, for x of order one, B and its derivatives with respect to x and t are of order one, and

$$|B| \le C|x|,$$

for some constant $C > 0$. Formally, taking $\varepsilon = 0$, yields the leading order equation $Ax + f = 0$. Denoting the solution $\Phi_0 = -A^{-1}F$ we substitute

$$x = \Phi_0(t) + x_1,$$

and obtain by Taylor expansion

$$\begin{aligned}\varepsilon x_1' &= (A(t)+\varepsilon B(x_1+\Phi_0,t))\,(x_1+\Phi_0)+F(t)-\varepsilon\Phi_0'\\ &= (A_1(t)+\varepsilon B_1(x_1,t))\,x_1+\varepsilon F_1(t),\end{aligned}$$

where B_1 has the same properties as B and

$$A_1(t)=A(t)+\mathcal{O}(\varepsilon),\qquad F_1=B(\Phi_0,t)\Phi_0-\Phi_0'.$$

Thus the new system is of the same form as the original one with the forcing function reduced to $\mathcal{O}(\varepsilon)$. Repeat the process p times yields

$$\begin{aligned}x &= \sum_{\nu=0}^{p-1}\varepsilon^\nu\Phi_\nu+x_p,\\ \varepsilon x_p' &= (A_p(t)+\varepsilon B_p(x_p,t))\,x_p+\varepsilon^p F_p,\quad A_p=A+\mathcal{O}(\varepsilon),\\ x_p(0) &= x(0)-\sum_{\nu=0}^{p-1}\varepsilon^\nu\Phi_\nu(0).\end{aligned}\tag{23}$$

Therefore, we have

Theorem 2. *The solution of (22) has p derivatives bounded independently of ε if and only if*

$$x(0)=\sum_{\nu=0}^{p-1}\varepsilon^\nu\Phi_\nu(0)+\mathcal{O}(\varepsilon^p),$$

i.e. $x(0)$ is, except for terms of order $\mathcal{O}(\varepsilon^p)$, uniquely determined.

If F and all its derivates vanish at $t=0$, then the initial condition

$$x(0)=0$$

defines a solution for which any number of derivatives are bounded independent of ε. We can construct such a solution even if F and its derivatives do not vanish at t=0, provided we can extend F smoothly to negative t. If the solution operator of the linearized problem,

$$\varepsilon v_p' = A_p(t)v,$$

is bounded, then $x_p=\mathcal{O}(\varepsilon^{p-j})$ in time intervals of length $\mathcal{O}(\varepsilon^{-j})$.

Generalizing (22), consider

$$\begin{aligned}\varepsilon x' &= (A(t)+\varepsilon C(v,t)+\varepsilon B(v,x,t))\,x+F(t),\\ v' &= D(v,x,t)x+G(v,t).\end{aligned}\tag{24}$$

Here A,A^{-1},B,C,D,F,G and their derivatives with respect to x,v,t are of order $\mathcal{O}(1)$, if x,v,t are of order $\mathcal{O}(1)$. Following the same reasoning as before, substitute

$$x=-A^{-1}(t)F(v,t)+x_1$$

to obtain a system of the same form with F replaced by εF_1. Repeating the process p times yields

$$\begin{aligned} \varepsilon x_p' &= (A(t) + \varepsilon C_p(v,t) + \varepsilon B_p(v,x_p,t))x_p + \varepsilon^p F_p(t), \\ v' &= D_p(v,x_p,t)x_p + G_p(v,t). \end{aligned} \tag{25}$$

We conclude the following:

Theorem 3. *The solution of (24) has p time derivative which are bounded independently of ε if we choose*

$$x_p(0) = \mathcal{O}(\varepsilon^p),$$

i.e., $x(0)$ is, except for terms of order $\mathcal{O}(\varepsilon^p)$, uniquely determined by $v(0)$.

Generalizations. More generally, we can consider systems

$$\varepsilon w' = h(w,t).$$

If there is a solution $w(t)$ with $w'(t) = \mathcal{O}(1)$, then $h(w(t),t) = \mathcal{O}(\varepsilon)$. This suggests the existence of a C^∞-function $\phi(t)$ with

$$h(\phi(t),t) = 0, \qquad t \geq 0.$$

Introducing the new variable

$$\tilde{w} = w - \phi,$$

one obtains

$$\begin{aligned} \varepsilon \tilde{w}' &= h(\tilde{w} + \phi, t) - h(\phi,t) - \varepsilon \phi'(t) \\ &= (M(t) + N(\tilde{w},t))\,\tilde{w} - \varepsilon \phi'(t) \end{aligned}$$

where

$$M(t) = h_w(\phi(t),t), \qquad |N(\tilde{w},t)| \leq \text{const.}\ |\tilde{w}|.$$

If we further assume that

$$\tilde{w}(0) = w(0) - \phi(0) = \varepsilon z_0, \qquad z_0 = \mathcal{O}(1),$$

then we can rescale the equation for $\tilde{w}$ by introducing a new variable, $z = \varepsilon^{-1}\tilde{w}$. One obtains for $z(t)$

$$\varepsilon z' = (M(t) + \varepsilon Z(z,t))\,z - \phi'(t).$$

Since we are interested in highly oscillatory problems, assume that $M(t)$ has m purely imaginary eigenvalues which are independent of t. Denote

$$\kappa_j = i\mu_j,\ |\mu_j| \geq \delta > 0, \qquad j = 1,\ldots,m,$$

and n eigenvalues

$$\kappa_{m+1} = \ldots = \kappa_{m+n} = 0.$$

Without loss of generality, assume

$$M = \begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix}, \qquad |A^{-1}| = \mathcal{O}(1),$$

where A is an $m \times m$ matrix with eigenvalues κ_j, $j = 1, \ldots, m$. If we partition z accordingly,

$$z = \begin{pmatrix} x \\ v \end{pmatrix},$$

then we obtain

$$\begin{aligned} \varepsilon x' &= \left(A + \varepsilon Z^I(v,x,t)\right) x - \left(\phi'\right)^I, \\ v' &= Z^{II}(v,x,t) - \frac{1}{\varepsilon}\left(\phi'\right)^{II}. \end{aligned}$$

If $(\phi')^{II} = \mathcal{O}(\varepsilon)$, then the resulting system has the form (24).

2.3 Interaction between the fast and the slow scales

Continuing our discussion and ignoring the terms that are higher order in ε, we consider the following model equation:

$$\begin{aligned} \varepsilon x' &= (A(t) + \varepsilon C(v,t) + \varepsilon B(v,x,t))x, \\ v' &= D(v,x,t)x + G(v,t), \\ x(0) &= x_0, \quad v(0) = v_0. \end{aligned} \tag{26}$$

We obtain the slow solution v_s, if we set $x_0 = 0$, i.e.,

$$v_S' = G(v_S,t), \qquad v_S(0) = v_0, \qquad x \equiv 0. \tag{27}$$

Let us make the following assumption.

Assumption 1. The solution operators S_1,S_2 of

$$v_L' = \frac{\partial G}{\partial v}(v_S)v_L$$

and

$$\varepsilon x_L' = (A(t) + \varepsilon C(v_S,t))x_L,$$

respectively, are uniformly bounded.

Here, the solution operator $S_1(t, s)$ for $t > s$ maps $\mathrm{V}_L(s)$ to $\mathrm{V}_L(t)$, and $\mathrm{S}_2(t, s)$ acts the same way for X_L.

If $x_0 \neq 0$, then the slow solution will be perturbed and we want to estimate $v - v_S$. We start with a rather crude estimate. We assume that x_0 is small and want to show

$$|v - v_S| = \mathcal{O}(|x_0|^2 t + \varepsilon |x_0|)$$

in time intervals $0 \leq t \leq T$ with $T \ll |x_0|^{-1}$. We linearize (26) around $v = v_s$ and $x = 0$. Let $v = v_S + v_L$, $x = x_L$, then the linearized equations have the form

$$\begin{aligned}
\varepsilon x_L' &= (A(t) + \varepsilon C(v_S,t))\, x_L, \\
v_L' &= \tilde{D}(v_S,t)x_L + \frac{\partial G}{\partial v}(v_S)v_L, \quad \tilde{D} = D(v_S,0,t), \\
x_L(0) &= x_0,\ v_L(0) = v_0.
\end{aligned}$$

By assumption

$$|x_L| \leq \text{const.}|x_0|.$$

Duhamel's principle and integration by parts gives us

$$\begin{aligned}
v_L(t) &= \int_0^t S_1(t,\xi)\tilde{D}x_L d\xi \\
&= \varepsilon \int_0^t S_1(t,\xi)\tilde{D}(A+\varepsilon C)^{-1}x_L' d\xi \\
&= \varepsilon S_1(t,\xi)\tilde{D}(A+\varepsilon C)^{-1}x_L|_0^t - \varepsilon \int_0^t \frac{\partial}{\partial \xi}(S_1(t,\xi)\tilde{D}(A+\varepsilon C)^{-1})x_L' d\xi.
\end{aligned}$$

The last integral can be treated in the same way. Therefore, Assumption 1 gives us, for any p,

$$|v_L(t)| \leq \text{const.}\,(\varepsilon|x_0| + \mathcal{O}(\varepsilon^p t)).$$

Assume now that A is constant, has distinct purely imaginary eigenvalues and that B,D are polynomials in x. Our goal is to give conditions such that our estimate will be improved to

$$|v - v_S| = \mathcal{O}(\varepsilon x_0) \text{ in time intervals } 0 \leq t \leq T, T \ll (\varepsilon|x_0|)^{-1}. \tag{28}$$

Without restriction we can assume that the system has the simplified form

$$\varepsilon x' = (i\Lambda + \varepsilon\Lambda_1(t) + \varepsilon B(x,t))\, x, \tag{29}$$

$$v' = D(x,t)x + G(v,t), \tag{30}$$

where Λ, Λ_1 are diagonal matrices and $\Lambda_1 + \Lambda_1^* \leq 0$. We introduce new variables

$$x = e^{\frac{i}{\varepsilon}\Lambda t} z.$$

Then (29) becomes

$$z' = \Lambda_1(t)z + \tilde{B}z, \tag{31}$$

where

$$\tilde{B} = e^{-\frac{i}{\varepsilon}\Lambda t} B\left(e^{\frac{i}{\varepsilon}\Lambda t}z, t\right) e^{\frac{i}{\varepsilon}\Lambda t}.$$

We split

$$\tilde{B} = B_1 + B_2,$$

where B_1 is a polynomial in z without exponentials, and all terms of B_2 contain exponentials. B_2 produces a $\mathcal{O}(\varepsilon|x_0|^2 t)$-change of z and, therefore, we neglect it. Thus, we can simplify (31) to

$$z' = \Lambda_1 z + B_1(z,t)z.$$

If $B_1 \neq 0$, then we can in general not expect that $|z| \leq K|x_0|$ holds in time intervals $T \gg |x_0|^{-1}$. We proved that the solution of (29) is of the form

$$x(t) = e^{\frac{i}{\varepsilon}\Lambda t} z(t), \tag{32}$$

where $z(t)$ is varying slowly. We introduce (32) into (30). We can also split

$$D(x,t)x = D_1 + D_\varepsilon, \tag{33}$$

where D_1 does not contain any exponentials and all terms of D_ε contain exponentials. Observe that D_1 is quadratic in z, i.e. $D_1 = \mathcal{O}(|x_0|^2)$. We can further deduce that $\frac{d}{dt}D_1(x(t),t)$ is independent of ε, while $\frac{d}{dt}D_\varepsilon \sim \mathcal{O}(\varepsilon^{-1})$. We have the following important conclusion:

- If $D_1 \neq 0$, then in general
$$|v - v_S| = \mathcal{O}(|x_0|^2 t).$$
- If $D_1 = 0$, then
$$|v - v_S| = \mathcal{O}(\varepsilon|x_0|),$$
and (28) holds.

We should look at the above result together with what we obtained in Sect. 2.1, in particular, the estimate (19) for the case when resonance occurs in the equation for x, and (20) for the case without resonance in the system.

2.4 Slow variables and slow observables

Consider the following ODE system

$$\begin{aligned} x' &= -\varepsilon^{-1}y + x, \quad x(0) = 1, \\ y' &= \varepsilon^{-1}x + y, \quad\ \ y(0) = 0. \end{aligned} \tag{34}$$

The solution of this linear system is $(x(t),y(t)) = (e^t \cos\varepsilon^{-1}t, e^t \sin\varepsilon^{-1}t)$ whose trajectory forms a slowly expanding spiral: i.e. the solution rotates around the origin with a fast frequency $2\pi/\varepsilon$ and the distance to the origin grows in time by e^t. Although both $x(t)$ and $y(t)$ change on the ε time scale, the system can be decomposed into "fast and slow constituents": a fast rotational phase and a slowly changing amplitude. Denoting $\xi = x^2 + y^2$, we have

$$\xi' = \frac{d}{dt}\xi(x(t),y(t)) = 2xx' + 2yy' = 2x^2 + 2y^2 = 2\xi.$$

Three important points call attention:

- ξ' is bounded independent of ε. Accordingly, we refer to the function $\xi(x,y)$ as a slow variable for (34).

- ξ' can be written into a function of the slow variable ξ only. Accordingly, we say that the equation for the slow variable ξ is closed.
- Suppose that there is another (slow) function $\zeta(x,y)$ such that $\frac{d}{dt}\zeta(x(t),y(t)) = \zeta_x x' + \zeta_y y'$ is bounded independent of ε. Then, away from the origin, $\nabla\zeta(x,y)$ is parallel to $\nabla\xi(x,y)$. Otherwise, at every point away from the origin, $\nabla\zeta$ and $\nabla\xi$ form a local basis for the two dimensional vector space. Consequently, the velocity field $\Phi(x,y;\varepsilon) := (-\varepsilon^{-1}y + x, \varepsilon^{-1}x + y)$ can be written as a linear combination of $\nabla\zeta$ and $\nabla\xi$; we write $\Phi = a\nabla\xi + b\nabla\zeta$. From the hypotheses on the slowness of ξ and ζ, $\Phi\cdot\nabla\xi$ and $\phi\cdot\nabla\zeta$ are both bounded independent of ε, implying that the coefficients a and b are also bounded. However, this leads to Φ being bounded which contradicts with the given equation (34).

The three observations described above are essential for building the multiscale numerical methods introduced in the next section.

More generally, consider the ODE systems

$$x' = \varepsilon^{-1} f(x) + g(x), \qquad x(0) = x_0, \tag{35}$$

where $x \in \mathbb{R}^d$. We assume that for $0 < \varepsilon < \varepsilon_0$, and for any x_0 in a region $\mathscr{A} \subset \mathbb{R}^d$, the unique solution of (35), denoted $x(t;\varepsilon,x_0)$, exists in $t \in [0,T]$ and stays in some bounded region D. For brevity, we will omit the explicit dependence of the solution on ε and x_0 whenever it is clear from context.

Definition 1. *Let U be a nonempty open subset of $\mathscr{A}$. A smooth function $\xi : \mathbb{R}^d \mapsto \mathbb{R}$ is said to be slow with respect to (35) in U, if there exists a constant C such that*

$$\max_{x_0\in U,\ t\in[0,T]} \left| \frac{d}{dt}\xi(x(t;\varepsilon,x_0)) \right| \le C.$$

Otherwise, $\xi(x)$ is said to be fast. Similarly, we say that a quantity or constant is of order one if it is bounded independent of ε. It is also no problem generalizing this notion to time dependent slow variables $\xi(x,t)$.

Loosely speaking, $\xi(x)$ being slow means that, to leading order in ε, the quantity $\xi(x(t))$ is evolving on a time scale that is ε independent for all trajectories emanating from a macroscopic domain (radius does not shrink with ε).

Following Definition 1, for systems of the form

$$x' = f\left(\frac{t}{\varepsilon}, x\right), \qquad f \text{ bounded},$$

each scalar component of the state variables x is considered a slow variable. Indeed, for those functions f that are periodic in the first argument, we know from our previous discussion that $x(t)$ stays very close to its initial value for all $0 \le t \le T$. Furthermore, in the case of resonance discussed in Sect. 1.3, $x(t)$ drifts away from the initial value in an average distance that is growing linearly in time. For integrable Hamiltonian systems, the action variables are the slow variables for the systems.

As another example, in Sect. 2.3 function $D_1(y,t)$ may be considered as a slow variable for (29).

At this point, it is natural to ask how many slow variables exist for a given highly oscillatory system? The obvious answer is infinitely many, since any constant multiplication of a found slow variable yield another new one. A more reasonable question is to ask what the dimension of the set of all slow variables is. In preparation to answering this question, we need to make concrete a few more concepts. In the following sections this question is answered for some specific cases.

Definition 2. *Let $\alpha_1, \cdots, \alpha_k : \mathscr{A} \subset \mathbb{R}^n \mapsto \mathbb{R}$ be k smooth functions, $k \leq n$. $\alpha_1(x), \ldots, \alpha_k(x)$ are called functionally independent if the Jacobian has full rank; i.e.*

$$\operatorname{rank}\left(\frac{\partial(\alpha_1, \cdots, \alpha_k)}{\partial x}\right) = k.$$

Let $\alpha(x) = (\alpha_1(x), \ldots, \alpha_k(x))^T$ be a vector containing k functionally independent components.

When coupled with system (35), $\alpha(x)$ is called a maximal vector of functionally independent slow variables if, for any other vector of size ν whose components are functionally independent, then $k \geq \nu$.

Our objective is to use an appropriate set of slow variables together with some other smooth functions to provide a new coordinate system for a subset of the state space of system (35). Such a coordinate system separates the slow behavior from the fast oscillations and provides a way to approximate a large class of slow behavior of (35). See Fig. 4 for an illustration; locally near the trajectory, the space is decomposed into three special directions, $\nabla\phi$ defines the fast direction, and the two slow variables ξ_1 and ξ_2 help gauging the slow behavior of a highly oscillatory system.

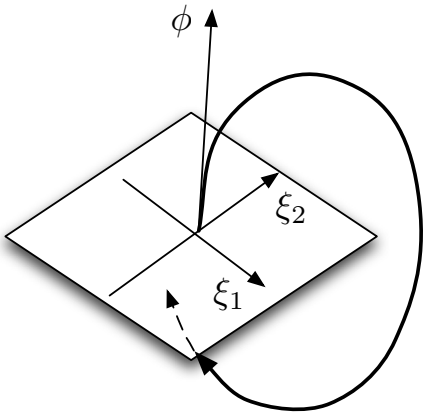


Fig. 4. Illustration of a slow chart. The slow variables ξ_1 and ξ_2 provide a local coordinate system near a trajectory.

Definition 3. *Let $\xi = (\xi_1(x), \ldots, \xi_k(x))$ denote a maximal vector of functionally independent slow variables with respect to (35) in $\mathscr{A}$, and $\phi : \mathscr{A} \subset \mathbb{R}^d \mapsto \mathbb{R}^{d-k}$ be some smooth functions. If the Jacobian matrix $\partial(\xi,\phi)/\partial x$ is nonsingular in $\mathscr{A}$, one obtains a local coordinate systems, i.e., a chart of the states space. We refer to such a chart as a slow chart for $\mathscr{A}$ with respect to the ODE (35).*

In other words, a slow chart is a local coordinate system in which a maximal number of coordinates are slow with respect to (35).

Lemma 1. *Let (ξ,ϕ) denote a slow chart for $\mathscr{A} \subset \mathbb{R}^d$ and $\alpha(x) : \mathscr{A} \to \mathbb{R}$ a slow variable. Then, there exists a function $\tilde{\alpha}(\xi) : \mathbb{R}^k \to \mathbb{R}$ such that $\alpha(x) = \tilde{\alpha}(\xi(x))$.*

Proof. Otherwise, $\alpha(x)$ is a new slow variable that is functionally independent of the coordinates of ξ, in contradiction to the maximal assumption.

Another type of slow behavior can be observed through integrals of the trajectory, referred to as slow observables.

Definition 4. *A bounded functional $\beta : C^1(\mathscr{A} \times [0,T]) \cap L^1(\mathscr{A} \times [0,T]) \mapsto \mathbb{R}$ is called a (global) slow observable if*

$$\beta(t) = \int_0^t \tilde{\beta}(x(\tau;\varepsilon,x_0),\tau)d\tau.$$

Differentiation with respect to time shows that global observables are slow.

From the discussion in Sect. 4.2, we deduce that with an appropriate choice of kernel and η, local averages of the form

$$\beta(t) = \int_{\infty}^{+\infty} \frac{1}{\eta} K\left(\frac{t-\tau}{\eta}\right) \tilde{\beta}(x(\tau;\varepsilon,x_0),\tau)d\tau,$$

can also be slow. We refer to these as local observables.

We observe that along the trajectory passing through y_0, a slow variable defines a slow changing quantity ϑ. We first consider the unperturbed equation

$$\varepsilon y' = f(y,t),$$

and a slow variable α.

$$\frac{d}{dt}\alpha(y(t)) = \nabla\alpha|_{y(t)} \cdot y'(t) = \frac{1}{\varepsilon}\nabla\alpha|_{y(t)} \cdot f =: \phi_{\alpha,f}(t;y_0). \tag{36}$$

Notice that since α is a slow variable, $|\phi_{\alpha,f}(t;y_0)| \leq C_1$. If this bound is valid for $0 < \varepsilon \leq \varepsilon_0$, then $\nabla\alpha \cdot f = 0$ for a neighborhood of $y(t)$. Now consider

$$\varepsilon\tilde{y}' = f_\varepsilon(\tilde{y},t) + \varepsilon g(\tilde{y},t).$$

We may directly consider integrating a slow observable $\vartheta(t)$ satisfying

$$\frac{d}{dt}\vartheta = \phi_{\alpha,f_\varepsilon}(t;y_0), \qquad \vartheta(0) = \vartheta_0.$$

Notice that

$$\phi_{f_\varepsilon} = \frac{1}{\varepsilon}\nabla\alpha|_{\tilde{y}(t)} \cdot f(\tilde{y}(t),t) + \nabla\alpha|_{y(t)} \cdot g(\tilde{y}(t),t) = \nabla\alpha|_{y(t)} \cdot g(\tilde{y}(t),t).$$

So $\vartheta(t)$ is slowly varying in $\mathcal{O}(1)$ time scale.

2.5 Building slow variables by parametrizing time

The time variable may be used to create slow variables that couple different oscillators if we can locally use the coordinates of the state space to parametrize time so that time is treated as a dependent variable. Consider the equations of the form $\varepsilon y' = f(y,t)$ and assume that there exists a function τ, independent of ε, such that $\varepsilon\tau(y(t)) = t$. The function τ by its definition is not slow since

$$\frac{d}{dt}\tau(y(t)) = \frac{t}{\varepsilon}.$$

However, if we have $\varepsilon\tilde{\tau}(z(t)) = t$ for the solutions, $z(t)$, of another oscillatory problem, then the function $\theta(y,z) := \tau(y) - \tilde{\tau}(z)$ is a slow variable since

$$\frac{d}{dt}\theta(y(t),z(t)) \equiv 0.$$

The existence of inverse functions depend on the monotonicity in time of any coordinate of the trajectories. For oscillatory problems, the monotonicity cannot hold globally. In many problems, even though the inverse function τ does not exist globally, its derivative can be defined globally. In this case, we may employ (36) to integrate a slow quantity. For example, the derivative of $\arctan(z)$ is defined on the whole real line. Similarly, on the complex plane, the derivative of the arg function is defined everywhere except at the origin. In the latter case, (36) can be regarded defining a continuous $\theta(t)$ on the Riemann sheet.

One advantage of using time as a slow variable is in defining relative phase between two planar oscillators. Consider

$$\varepsilon z_k' = i\lambda_k z_k, \qquad k = 1,2.$$

We formally define

$$\alpha(z_1,z_2) := \arg(z_1) - \arg(z_2)$$

and obtain the equation for the slow observable. Through this approach, we can define and integrate the slowly changing relative phase between two oscillators.

2.6 Effective closure

Let $U(t) \in \mathbb{R}^n$ and $V(t) \in \mathbb{R}^m$ be two smooth functions. Assume that, for all $0 \le t \le T$, both $U(t)$ and $V(t)$ are bounded above by C_0 and that

$$\frac{dU}{dt} = G(U) + \varepsilon H(U,V,t),$$

for some bounded smooth function $H : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^+ \mapsto [-C_1, C_1]$. We say that the dynamics of U is effectively closed. This means that for $0 \leq t \leq T$ and ε sufficiently small, one can ignore the influence of $V(t)$ and compute instead

$$\frac{d\tilde{U}}{dt} = G(\tilde{U}), \qquad \tilde{U}(0) = U(0),$$

as an approximation of $U(t)$; i.e.

$$|U(t) - \tilde{U}(t)| \leq C_1 \varepsilon.$$

In the spiral example (34), the equation for the single slow variable $\xi = x^2 + y^2$ is effectively closed. The following gives an example of slow variables whose dynamics along the trajectories are not effectively closed. In the complex plane, consider the system

$$\begin{aligned} x' &= \frac{i}{\varepsilon} x + x^* y, \\ y' &= \frac{2i}{\varepsilon} y. \end{aligned} \tag{37}$$

Here the x^* denotes the complex conjugate of x. Evidently, $\xi_1 := xx^*$ and $\xi_2 := yy^*$ are two slow variables. However, the differential equation for ξ_1 along the non-equilibrium trajectories of (37) is given by $\xi_1' = 2\mathrm{Re}((x^*)^2 y)$, which cannot be described in terms of ξ_1 and ξ_2 alone. Hence, the equation for ξ_1 is not effectively closed. In fact, it is easily verified that $\xi_3 = (x^*)^2 y$ is also a slow variable and that $(\xi_1, \xi_2, \xi_2, \arg x)$ is a slow chart.

Later, we will see that in many oscillatory systems, the effective equations for the slow coordinates in a slow chart are effectively closed.

3 Averaging

One of the most important analytic tools for studying highly oscillatory systems are averaging methods, see e.g. [5, 20, 29] . In this section we present a few key results and discuss them using simplified examples.

3.1 Time averaging and integration by parts

We start with a simple example. Let $a(t)$ be a C^1 function whose derivative is bounded on the real line. Consider the integral

$$I(t) = \int_0^t \cos\left(\frac{s}{\varepsilon}\right) a(s) ds = \varepsilon \sin\left(\frac{t}{\varepsilon}\right) a(t) - \varepsilon \int_0^t \sin\left(\frac{s}{\varepsilon}\right) a'(s) ds.$$

Then $|I(t)| \leq C_0(1+t)\varepsilon$ for some constant C_0 coming from the maximum value of a and a' in $[0,t]$. Let $a(t)$ be a λ-periodic function in C^p, $p \geq 1$; and let $a_\infty := \max_{0\leq t\leq\lambda} |a(t)|$. If $a(t)$ has zero average, $\int_0^\lambda a(s)ds = 0$, then for all $T \geq 0$,

$$\left| \int_0^T a(t)dt \right| \leq \lambda a_\infty.$$

We define a particular anti-derivatives of $a(t)$ as follows

$$a^{[0]}(t) = a(t) \text{ and } a^{[k]}(t) = \int_0^t a^{[k-1]}(s)ds + c_k, \quad k = 1,2,3,\ldots \tag{38}$$

where the constant c_k is chosen such that $\int_0^\lambda a^{[k]}(s)ds = 0$. As a result, $a^{[k]}(t)$ are also λ-periodic since

$$a^{[k+1]}(t+\lambda) - a^{[k+1]}(t) = \int_t^{t+\lambda} a^{[k]}(s)ds = \int_0^\lambda a^{[k]}(s)ds = 0, \ \ k = 1,2,3,\cdots. \tag{39}$$

Consequently, all anti-derivatives are uniformly bounded:

$$|a^{[k]}(t)| \leq \lambda a_\infty, \qquad \forall t. \tag{40}$$

If f is differentiable, we can perform integration by parts

$$\int_0^T a\left(\frac{s}{\varepsilon}\right) f(s)ds = \left[\varepsilon a^{[1]}\left(\frac{s}{\varepsilon}\right) f(s)\right]_{s=0}^T - \varepsilon \int_0^T a^{[1]}\left(\frac{s}{\varepsilon}\right) f^{(1)}(s)ds,$$

where $f^{(k)}$ is the k-th derivative of f. The process can be repeated depending on the differentiability of f.

- If $f(T) = f(0) = 0$, and $f \in C^p$, then

$$\left| \int_0^T a\left(\frac{s}{\varepsilon}\right) f(s)ds \right| \leq \sup_{0\leq t\leq T} |f^{(p)}(t)| a_\infty \cdot \varepsilon^p.$$

- If f is in C^∞, we can further obtain a formal asymptotic expansion approximation for the integral

$$\int_0^T a\left(\frac{s}{\varepsilon}\right) f(s)ds = \sum_k \left[(-1)^{k-1}\varepsilon^k a^{[k]}\left(\frac{s}{\varepsilon}\right) f^{(k-1)}(s)\right]_{s=0}^T.$$

- If $\bar{a} := \int_0^\lambda a(\xi)d\xi \neq 0$, then

$$I_\varepsilon := \int_0^T a\left(\frac{s}{\varepsilon}\right) f(s)ds \longrightarrow \bar{I} := \bar{a}\left(\int_0^T f(s)ds\right) \text{ as } \varepsilon \to 0. \tag{41}$$

- Similar averaging results can be obtained for functions $a(t)$ which are not necessarily periodic but whose anti-derivatives are nonetheless bounded.

Exercise 3. Prove (41).

3.2 How does averaging in an oscillatory system appear?

Consider

$$x' = f\left(\frac{t}{\varepsilon}, x\right), \qquad x(0) = x_0, \tag{42}$$

where $f(t,x)$ is Lipschitz in both t and x with constant L and is λ-periodic, $f(t+\lambda,x) = f(t,x)$. In addition, consider

$$y' = \bar{f}(y), \quad y(0) = y_0, \qquad \text{where} \quad \bar{f}(x) = \frac{1}{\lambda}\int_0^{\lambda} f(t,x)dt. \tag{43}$$

We call (43) the averaged, or effective equation derived from (42). The following calculation shows that $|x_\varepsilon(t) - y(t)| \le C_1\varepsilon$ for a long time which is independent of ε. Observe that

$$\begin{aligned} x(t) - x_0 &= \int_0^t f\left(\frac{\tau}{\varepsilon}, x(\tau)\right) d\tau \\ &= \int_{t_M}^t f\left(\frac{\tau}{\varepsilon}, x(\tau)\right) d\tau + \sum_{j=0}^{M-1} \int_{j\varepsilon\lambda}^{(j+1)\varepsilon\lambda} f\left(\frac{\tau}{\varepsilon}, x(\tau)\right) d\tau, \end{aligned}$$

where $t - t_M < \varepsilon\lambda$. In each interval $t_j = j\varepsilon\lambda \le t \le t_{j+1} = (j+1)\varepsilon\lambda$,

$$\begin{aligned} \int_{j\varepsilon\lambda}^{(j+1)\varepsilon\lambda} f\left(\frac{\tau}{\varepsilon}, x(\tau)\right) d\tau &= \int_{j\varepsilon p}^{(j+1)\varepsilon\lambda} f\left(\frac{\tau}{\varepsilon}, x(t_j)\right) + \mathcal{O}(\varepsilon) d\tau \\ &= \int_{j\varepsilon\lambda}^{(j+1)\varepsilon\lambda} f\left(\frac{\tau}{\varepsilon}, x(t_j)\right) d\tau + \mathcal{O}(\varepsilon^2) \\ &= \varepsilon p \cdot \frac{1}{\lambda}\int_0^{\lambda} f(s, x_j) ds + \mathcal{O}(\varepsilon^2) \\ &= \varepsilon\lambda \overline{f}(x_j) + \mathcal{O}(\varepsilon^2). \end{aligned}$$

Hence,

$$\begin{aligned} x(t) - x_0 &= \int_0^t f\left(\frac{\tau}{\varepsilon}, x(\tau)\right) d\tau \\ &= \int_{t_M}^t f\left(\frac{\tau}{\varepsilon}, x(\tau)\right) d\tau + \sum_{j=0}^{M-1} \int_{j\varepsilon\lambda}^{(j+1)\varepsilon\lambda} f\left(\frac{\tau}{\varepsilon}, x(\tau)\right) d\tau \\ &= \int_0^t \overline{f}(x(\tau)) d\tau + \mathcal{O}(\varepsilon). \end{aligned}$$

Now, since

$$y(t) - x_0 = \int_0^t \bar{f}(y(\tau)) d\tau,$$

by Gronwall's lemma we have

$$|x(t) - y(t)| \le L\int_0^t |x(\tau) - y(\tau)| d\tau + C\varepsilon.$$

This result can be generalized as follows. Let

$$x' = \sum_{j=1}^{M} f_j\left(\frac{t}{\varepsilon}, x\right), \qquad x(0) = x_0,$$

where f_j is λ_j-periodic in the time. Then, a calculation similar to the above shows that the solution of

$$y' = \bar{f}(y), \qquad y(0) = x_0,$$

with

$$\bar{f}(x) = \sum_j \frac{1}{\lambda_j} \int_0^{\lambda_j} f_j(\tau, x) d\tau,$$

is close to $x(t)$ on a time segment.

Theorem 4. *For $t \in [0,T]$, $T < \infty$ and independent of ε (assume $x(t), y(t)$ exist in such interval)*

$$|x(t) - y(t)| \leq C_1 \varepsilon.$$

Note that $y' = \overline{f}(y)$ is independent of ε. While $x_\varepsilon(t)$ is highly oscillatory, there are no ε-scale oscillations in $y(t)$. We conclude that the cost of integrating the averaged equation is independent of ε and is in general much more efficient than computing x_ε. If we just pick an arbitrary t^*, $z' = f(t^*, z)$, $z(0) = x_0$ in general we can not expect that $x(t) = z(t) + \mathcal{O}(\varepsilon)$.

Averaging over oscillations may appear in many different ways and should be handled with caution. The following problem presents a case in homogenization, in which harmonic averages are derived as parameters for an effective equation.

Exercise 4. In the following problem, high frequency oscillations in a_ε interact with those in $\frac{d}{dx}u_\varepsilon$ and creates low frequency behavior of $u_\varepsilon(x)$:

$$\begin{cases} \frac{d}{dx}\left(a_\varepsilon(x)\frac{d}{dx}u_\varepsilon\right) = f(x), \quad 0 < x < 1, \\ u_\varepsilon(0) = u_\varepsilon(1) = 0, \\ a_\varepsilon(x) = a(\frac{x}{\varepsilon}) > 0. \end{cases}$$

We derive an effective equation for $u_\varepsilon(x)$ by performing the following steps.

1. Integrate the equation with respect to x and show that

$$\begin{cases} a_\varepsilon \frac{du_\varepsilon}{dx} = \int_0^x f(\xi) d\xi + C, \\ u_\varepsilon(x) = \int_0^x (a_\varepsilon(\xi))^{-1} F(\xi) d\xi, \quad \text{where } F(\xi) = \int_0^\xi f(\eta) d\eta + C. \end{cases}$$

   Determine C from boundary conditions.
2. Show that

$$\lim_{\varepsilon \to 0} \int_0^x a\left(\frac{\xi}{\varepsilon}\right)^{-1} F(\xi) d\xi = \int_0^1 a(y)^{-1} dy \int_0^x F(\xi) d\xi, \quad F \in C[0,1].$$

3. Show that

$$u_\varepsilon \longrightarrow \bar{u} = A^{-1} \int_0^x \left(\int_0^\xi F(\eta) d\eta + C \right) d\xi \text{ as } \varepsilon \longrightarrow 0,$$

where

$$A = \frac{1}{\int_0^1 a(y)^{-1} dy}$$

We conclude that $\bar{u}(x)$ satisfies the effective equation:

$$A \frac{d^2 \bar{u}}{dx^2} = f(x), \qquad 0 < x < 1, \qquad \bar{u}(0) = \bar{u}(1) = 0$$

In summary, we present the following facts about averaging, whose proof can be found, for example, in [20] and [29].

Theorem 5. *Let $x, y, x_0 \in D \subset \mathbb{R}^n$, $\varepsilon \in (0, \varepsilon_0]$. Suppose*

1. *f, g, and $|\nabla f|$ are bounded by M which is independent of ε.*
2. *g is Lipschitz in a bounded domain D.*
3. *$f(t,x)$ is λ-periodic in t, λ independent of ε.*

Then, the solution of

$$x' = f\left(\frac{t}{\varepsilon}, x\right) + \varepsilon g\left(\frac{t}{\varepsilon}, x, \varepsilon\right), \qquad x(0) = x_0 \tag{44}$$

is close to the solution of the averaged equation

$$y' = \bar{f}(x), \qquad y(0) = x_0, \qquad \overline{f}(y) = \frac{1}{\lambda} \int_0^\lambda f(t,y) dt$$

on a time scale of order one. More precisely, for all $t \in [0,T]$, $T < \infty$ independent of ε,

$$|x(t) - y(t)| \leq C \varepsilon T e^{\varepsilon L t},$$

where $C > 0$ and L denotes a Lipschitz constant for $\bar{f}$.

Moreover, equation (44) can be written in the form [20]

$$x' = \bar{f}(x) + \varepsilon f_1\left(\frac{t}{\varepsilon}, x, \varepsilon\right), \qquad x(0) = x_0, \tag{45}$$

where $f_1(t,x,\varepsilon)$ is λ-periodic in t and $f_1 \to 0$ as $\varepsilon \to 0$.

3.3 Effective closure in coupled oscillators

Given the system (35), and in a neighborhood of the trajectory starting from x_0, let (ξ,ϕ) be a slow chart in which ϕ is a fast angular coordinate on the unit circle $\mathbb{S}^1$, i.e., $0 < C_1/\varepsilon < \phi' < C_2/\varepsilon$. Then, by these hypotheses, we know that

$$\begin{aligned} \xi' &= g_I(\xi,\phi), \\ \phi' &= \varepsilon^{-1} g_{II}(\xi,\phi), \end{aligned} \tag{46}$$

where $C_1 < g_{II}(\xi,\phi) < C_2$. Applying the averaging result (45), Equation (46) can be rewritten as

$$\begin{aligned} \xi' &= \int g_I(\xi,\phi)d\phi + \varepsilon g_{III}(\xi,\phi) = \bar{g}_I(\xi) + \varepsilon g_{III}(\xi,\phi), \\ \phi' &= \varepsilon^{-1} g_{II}(\xi,\phi). \end{aligned}$$

Hence, the equation for ξ is effectively closed.

4 Computational considerations

In this section we will describe a few computational methods which gain efficiency by taking into account some of the special properties of the system discussed in previous sections. We will mostly be concerned with equations of the form

$$x'_\varepsilon = g\left(\frac{t}{\varepsilon}, x_\varepsilon\right), \qquad x(0) = x_0,$$

where $g(t,x)$ is λ-periodic, and its averaged form

$$\bar{x}' = \bar{g}(\bar{x}), \qquad \bar{x}(0) = x_0.$$

By the averaging principle, we have that

$$|x_\varepsilon(t) - \bar{x(t)}| \le C\varepsilon, \qquad 0 \le t \le T.$$

4.1 Stability and efficiency

Suppose uniform time stepping is used in the computations.[4] The typical local truncation error of a p'th-order method is $\mathcal{O}((\triangle t L)^p)$, where L is a uniform bound for the $p+1$ derivative of the right hand side. Applied to the two equations above, the error varies tremendously. For x_ε, the error term is

$$E_1 = \mathcal{O}\left(\left[\frac{\triangle t}{\varepsilon}\right]^p\right),$$

[4] With oscillatory systems, variable time step algorithms are not as advantageous in improving efficiency as in stiff, dissipative systems.

while for $\bar{x}$, the truncation error is

$$E_2 = \mathcal{O}(\triangle t^p).$$

Thus, in order for the solution to be reasonably accurate, the step size $\triangle t$ has to be small compared to ε. In addition, typical explicit non-multiscale numerical schemes suffer from linear instabilities when using step sizes that are too large compared to the Lipschitz constant of the right hand size. This constraint restricts the step size of such a method to be of order ε On the other hand, the efficiency of solving an ODE to time T using step size $\triangle t$ is $\mathcal{O}(T/\triangle t)$. Hence, it is clear that it is usually much more efficient to solve the averaged equation for $\bar{x}$ than the original one for x_ε.

In the following we will develop and discuss some of the tools and ideas required to construct a multiscale algorithm that solves the averaged equation without actually deriving it. Instead, the idea of the Heterogenous Multiscale Method is to approximate the averaged equation on the fly using short time integration of the equation for x_ε.

4.2 Averaging kernels

In many numerical calculations involving oscillations with different frequencies, the right hand side may not be strictly periodic. As an example see (15). For this reason, as well as for efficiency considerations, it is convenient to average using some general purpose kernels. In the previous section, we see the need to compute the average of $f(t,x)$ over a period in t

$$\bar{f}(x) := \frac{1}{\lambda}\int_0^\lambda f(\tau,x)d\tau.$$

In this section, we show that $\bar{f}(x)$ can be accurately and efficiently approximated by averaging with respect to a compactly supported kernel whose support is larger, but independent of λ. For simplicity, we shall ignore the x dependence in f.

We will use $\mathbb{K}^{p,q}$ to denote the function space for kernels discussed in this paper.

Definition 5. *Let $K^{p,q}(I)$ denote the space of normalized functions with support in I, q continuous derivatives and p vanishing moments, i.e., $K \in \mathbb{K}^{p,q}(I)$ if $K \in C_c^q(\mathbb{R})$, $supp(K) = I$, and*

$$\int_{\mathbb{R}} K(t)t^r dt = \begin{cases} 1, & r = 0; \\ 0, & 1 \le r \le p. \end{cases}$$

Furthermore, we will use $K_\eta(t)$ to denote a scaling of K as

$$K_\eta(t) := \frac{1}{\eta}K\left(\frac{t}{\eta}\right).$$

For shorthand, we will also use $K^{p,q}$ to denote a function in $\mathbb{K}^{p,q}([-1,1])$.

Most of the numerical examples in this manuscript are obtained using the exponential kernel $K^{\exp} \in \mathbb{K}^{1,\infty}([-1,1])$:

$$K^{\exp}(t) = C_0 \chi_{[-1,1]}(t) \exp(5/(t^2-1)). \tag{47}$$

Here, $\chi_{[-1,1]}$ is the characteristic function of the interval [–1, 1] and C_0 is a normalization constant such that $||K^{\exp}||_{L^1(\mathbb{R})} = 1$. A second commonly used kernel is

$$K^{\cos}(t) = \frac{1}{2}\, \chi_{[-1,1]}(t)(1+\cos(\pi t)) \in K^{1,1}(I).$$

For convenience, we write $f(t) = \bar{f} + g(t)$ where

$$\bar{f} = \frac{1}{\lambda}\int_0^{\lambda} f(s)ds.$$

Hence, $g(t)$ is λ-periodic with zero average.

The following analysis shows that the convolution $K_\eta * f$ well approximates the average $\bar{f}$,

$$\begin{aligned}\int_{\mathbb{R}} \frac{1}{\eta} K\left(\frac{t-s}{\eta}\right) f(\frac{s}{\varepsilon})ds &= \int_{t-\eta}^{t+\eta} \frac{1}{\eta} K\left(\frac{t-s}{\eta}\right)\left(\bar{f} + g\left(\frac{s}{\varepsilon}\right)\right) ds \\ &= \bar{f}\int_{t-\eta}^{t+\eta} \frac{1}{\eta} K\left(\frac{t-s}{\eta}\right) ds + \frac{1}{\eta}\int_{t-\eta}^{t+\eta} K\left(\frac{t-s}{\eta}\right) g\left(\frac{s}{\varepsilon}\right) ds \\ &= \bar{f} + \frac{1}{\eta}\int_{t-\eta}^{t+\eta} K\left(\frac{t-s}{\eta}\right) g\left(\frac{s}{\varepsilon}\right) ds.\end{aligned}$$

Integrating by parts, we have

$$\begin{aligned}&\frac{1}{\eta}\int_{t-\eta}^{t+\eta} K\left(\frac{t-s}{\eta}\right) g\left(\frac{s}{\varepsilon}\right) ds \\ &\quad = \frac{\varepsilon}{\eta} K\left(\frac{t-s}{\eta}\right) G\left(\frac{s}{\varepsilon}\right)|_{s=t-\eta}^{t+\eta} - \frac{\varepsilon}{\eta^2}\int_{t-\eta}^{t+\eta} K'\left(\frac{t-s}{\eta}\right) G\left(\frac{s}{\varepsilon}\right) ds \\ &\quad = -\frac{\varepsilon}{\eta^2}\int_{t-\eta}^{t+\eta} K'\left(\frac{t-s}{\eta}\right) G\left(\frac{s}{\varepsilon}\right) ds,\end{aligned}$$

where G is the anti-derivative of g given by (38). Hence,

$$\left|\frac{1}{\eta}\int_{t-\eta}^{t+\eta} K\left(\frac{t-s}{\eta}\right) g\left(\frac{s}{\varepsilon}\right) ds\right| \le \frac{\varepsilon}{\eta}||K'||_\infty ||G||_\infty.$$

Since g is periodic and bounded, its anti-derivative is also a bounded function. For example, taking $\eta = \sqrt{\varepsilon}$, $\bar{f}$ is approximated to order $\sqrt{\varepsilon}$. Repeating this process q times yields

$$\left|\int K_\eta(t-s) f(s) ds - \bar{f}\right| \le C_{K,g}\left(\frac{\varepsilon}{\eta}\right)^q. \tag{48}$$

For convenience, we shall denote $K_\eta * f(t)$ by $<f(t)>_\eta$

4.3 What does a multiscale algorithm approximate?

Loosely speaking, our goal is to construct an algorithm that approximates the slow behaviour of a highly oscillatory ODE system. An important observation is that the slow behavior of a system can be a result of internal mutual cancellation of the oscillations. This, for example, is the case with resonances. Hence, it may not be clear what these slow aspects are. For this reason, we take a wide approach and require that our algorithm approximates all variables and observables which are slow with respect to the ODE.

How is this possible? We now prove that an algorithm which approximates the slow coordinates in a slow chart (ξ,ϕ) approximates all slow variables and observables.

Slow variables: Let $\alpha(x)$ denote a slow variable. From Lemma 1 we have that $\alpha(x)=\tilde{\alpha}(\xi(x))$ for some function $\tilde{\alpha}$. Therefore, values of $\alpha(x)$ depend only on ξ. Furthermore, it is not necessary to know $\tilde{\alpha}$, for suppose $\xi=\xi(x(t))$ at some time t. Then, $\alpha(x(t))=\tilde{\alpha}(\xi)=\tilde{\alpha}(\xi(x(t)))$. In other words, all points x which correspond to the same ξ yield the same value for $\alpha(x)$.

Slow observables – global time averages: We observe that for any smooth functions $\alpha(x,t)$, we have that $\bar{\alpha}(t)=\int_0^t \alpha(x(s),s)ds$ is slow since $|(d/dt)\bar{\alpha}(t)|=|\alpha(x(t),t)|$, which is bounded independent of ε. In ODE form, we have

$$\bar{\alpha}'=\alpha(x,t)$$

which complies to the form required by the averaging theorem. Therefore, $\bar{\alpha}$ can be integrated as a passive variable at the macroscopic level. In other words, it can be approximated by

$$\bar{\alpha}(t)=\int_0^t <\frac{d}{ds}\alpha(x(s),s)>_\eta\, ds$$

Slow observables – local time averages: Consider time averages of the form $<\alpha(x(s))>_\eta$. Since (ξ,ϕ) is a chart, we have that $\alpha(x(s))=\tilde{\alpha}(\xi,\phi)$ for some function $\tilde{\alpha}$. However, as proven in Sect. 4.2, convolution with kernels approximates averaging with respect to the fast angular phase ϕ. Here,

$$<\alpha(x(s))>_\eta=<\tilde{\alpha}(\xi,\phi)>_\eta=\int\tilde{\alpha}(\xi,\phi)d\phi+\text{error}=\bar{\alpha}(\xi(t))+\text{error},$$

where the error is evaluated in Sect. 4.2. Hence, a consistent approximation of ξ implies a consistent approximation of $<\alpha(x(s))>_\eta$. Moreover, in Sect. 5 we will show that the explicit form of $\tilde{\alpha}$ or $\bar{\alpha}$ are not required since all local time averages can be calculated as a by product of micro-solver steps in the algorithm.

4.4 Boosting methods

In the context of averaging, the idea of boosting is particularly simple. Consider, for example, the averaging Theorem 5 which states that, with functions $f(t,x)$, which are 1-periodic in time, the solution of

$$x' = f\left(\frac{t}{\varepsilon}, x\right), \qquad x(0) = x_0 \tag{49}$$

and

$$y' = \bar{f}(y), \qquad y(0) = x_0, \qquad \bar{f}(y) = \int_0^1 f(s,y)ds, \tag{50}$$

are close to order ε:

$$|x(t) - y(t)| < C\varepsilon,$$

on a time scale of order one. Suppose we are interested in solving (49) with a prescribed accuracy Δ which is small, but not as small than ε, i.e., $\varepsilon \ll \Delta \ll 1$. Consider the modified equation

$$z' = f\left(\frac{t}{\Delta}, z\right), \qquad z(0) = x_0. \tag{51}$$

Following the same averaging argument, $z(t)$ is close, to order Δ, to the averaged equation (50). Hence, by the triangle inequality

$$|z(t) - x(t)| \le |z(t) - y(t)| + |y(t) - x(t)| < C(\varepsilon + \Delta) < 2C\Delta. \tag{52}$$

Solving the boosted equation (51) instead of the original one, (49) introduces an error which is of order Δ. On the other hand, the stiffness of the equation is much reduced. The discussion in Sect. 4.1 shows that the efficiency of solving the boosted equation (51) is $\mathcal{O}(\Delta^{-1})$, which can be a considerable improvement over the $\mathcal{O}(\varepsilon^{-1})$ required to solve (49). Moreover, (51) has the exact same form as (49) and preserves the same invariance. For example, if the original system is Hamiltonian, that the boosted version is also Hamiltonian.

Despite their simplicity, boosting suffers from two major drawbacks. The first is related to the nature of the asymptotic expansion used to obtain the averaged equation. Similar to expanding functions in power series, the asymptotic expansion in the averaging Theorem 5 has a "radius of convergence". This implies that the averaged equation may provide a poor approximation for (49) if ε is not small enough. In other words, the proximity between $x(t)$ and $y(t)$ "kicks in" at some value ε_0, which is usually unknown. Hence, the error estimate (52) fails if $\Delta > \varepsilon_0$.

Another drawback is that the efficiency of the method is bound to be $\mathcal{O}(\Delta^{-1})$, no matter what the order of the integrator is. This is not the case with HMM, as will be discussed in the following section. Nonetheless, boosting serves as an important benchmark to test and evaluate the efficiency of our algorithm.

5 Heterogeneous Multiscale Methods

The Heterogeneous Multiscale Method (HMM) is a general framework for systems evolving on multiple, well separated time scales. We will focus on problems with two time scales which are referred to as slow/fast, or macro/micro scales. An HMM consists of two components: a macro-solver, integrating a generally unknown averaged equation, and a micro-solver, approximating the averaged equation using short time integration of the full ODE system.

5.1 A vanilla HMM example

Consider, for example, equations of the form

$$x' = f\left(\frac{t}{\varepsilon}, x\right), \qquad x(0) = x_0,$$

where $x \in \mathbb{R}^d$ and $f(t,x)$ is a smooth function which is 1-periodic in t. We rewrite the system as an homogeneous equation on $\mathbb{R}^d \times [0,1]$,

$$\begin{aligned} x' &= f(\phi, x), \qquad x(0) = x_0, \\ \phi' &= \varepsilon^{-1}, \qquad s(\phi) = 0, \end{aligned} \tag{53}$$

where ϕ is an angular variable defined on the quotient space $\mathbb{R}/[0,1]$. The latter space is isomorphic to the unit circle $\mathbb{S}^1$. By Definition 1, it is clear that all the coordinates in x are slow with respect to (53) while ϕ is fast. Hence, (x,ϕ) is a slow chart for (53). Furthermore, from Sect. 3, the solution for x is close (to order ε) to an effective equation which is effectively closed:

$$\bar{x}' = \bar{f}(\bar{x}), \qquad \bar{x}(0) = x_0,$$

where

$$\bar{f}(z) = \int_0^1 f(\tau, z) d\tau.$$

Earlier we saw that it is much favorable to solve for $\bar{x}$ rather than for x. However, the averaged forcing $\bar{f}(\cdot)$ is usually unknown. For this reason, following Sect. 4.2 we approximate $\bar{f}(\cdot)$ as $\langle f(t,x(t)) \rangle_\eta$. Applying a forward Euler scheme for x with a macroscopic step size H implies taking $x_{n+1} = x_n + H \langle f(t,x(t)) \rangle_\eta$. This is summarized in the following algorithm. Let x_n denote our approximation of (53) at time $t_n = nH$.

1. $n = 0$
2. Micro-simulation: approximate (53) numerically in a reduced time segment $[t_n - \eta, t_n + \eta]$ with step size h and x_0 replaced by x_n Denote the solution $x^n(t)$.
3. Force evaluation: calculate $F_n = \langle f(\frac{\cdot}{\varepsilon}), x^n(\cdot) \rangle_\eta$.
4. Macro-step (forward Euler example): take $x_{n+1} = x_n + HF_n$.
5. $n = n + 1$. Repeat steps 2–4 to time T.

The efficiency of the algorithm is $\mathcal{O}(T\eta/Hh)$. It is further analyzed in Sect. 5.4.

5.2 Systematically constructing heterogeneous multiscale methods

Consider stiff ordinary differential equations (ODEs) of the form

$$\frac{du}{dt} = f_\varepsilon(u,t), \tag{54}$$

where $u : (0,T) \mapsto \mathbb{R}^d$, and a subset of the eigenvalues of $\partial f_\varepsilon / \partial u$ are inversely proportional to a small positive parameter ε. When ε is very small, the complexity of

numerically solving such systems becomes prohibitively high. However, in many situations, one is interested only in a set of quantities U that are derived from the solution of the given stiff system (54), and typically, these quantities change slowly in time; i.e. both U *and* dU/dt *are bounded independent of* ε. For example, U could be the averaged kinetic energy of a particle system u.

Our objective is to construct and analyze ODE solvers that integrate the system

$$\frac{d}{dt}U = F(U,D), \tag{55}$$

where D is the data that can be computed by local solution of (54). U is called the slow (macroscopic) variable that is also some function or functional of u; i.e. $U = U(u,t)$.

If F is well-defined and has a convenient explicit mathematical expression, then there is no need to solve the stiff system (5 4) — one only needs to solve (55). In many situations, the dependence of F on U is not explicitly available. Our proposed strategy involves setting up a formal numerical discretization for (55), and evaluates F from short time history of u with properly chosen initial condition.

We will follow the framework of E and Engquist [11] in constructing efficient multiscale methods. In this framework, one assumes a macroscopic model

$$\Phi(U,D) = 0, \qquad U \in \Omega_{(M)} \tag{56}$$

which may not be explicitly given, but can be evaluated from a given microscopic model,

$$\varphi(u,d) = 0, \qquad u \in \Omega_{(m)} \tag{57}$$

where u are the microscopic variables. $D = D(u)$ and $d = d(U)$ denote the set of data or auxiliary conditions that further couple the macro- and microscopic models. Model (56) is formally discretized at a macroscopic scale, and the adopted numerical scheme dictates when the necessary information $D(u)$ should be acquired from solving (57), locally on the microscopic scale with auxiliary conditions $d(U)$. As part of $d(U)$ and $D(u)$, the macro- and microscopic variables are related by reconstruction and compression operators:

$$\mathscr{R}(U,D_R) = u, \qquad \mathscr{Q}(u) = U, \qquad \mathscr{Q}(\mathscr{R}(U,D_R)) = U,$$

where D_R are the needed data that can be evaluated from u. Errors of this type of schemes generally take the structure [10, 14]

$$\text{Error} = E_H + E_h,$$

where E_H is the error of the macroscopic model (56), and E_h contains the errors from solving (57) and the passing of information through $\mathscr{R}$ and $\mathscr{Q}$. This approach has been used in a number of applications, such as contact line problems, epitaxial growth, thermal expansions, and combustion. See the review article [12].

Figure 1 shows two typical structures of such ODE solvers. An ODE solver for U lies on the upper axis and constructs approximations of U at the grid points depicted

there. The fine meshes on the lower axis depict the very short evolutions of (54) with initial values determined by $\mathscr{R}(U(t_n))$. The reconstruction operator then takes each time evolution of u and evaluates F and U. The algorithms in [10, 16, 15], and [28] are also of a similar structure. As a simple example, the forward Euler scheme applied to (55) would appear to be

$$U_{n+1} = U_n + H \cdot \tilde{F}(U_n), \tag{58}$$

where $\tilde{F}$ contains the passage of $\mathscr{Q}\Phi_t\mathscr{R}(U_n)$ — reconstruction $\mathscr{R}$, evolution Φ_t, and compression $\mathscr{Q}$, and H is the step size. If each evolution of the full scale system (54) is reasonably short, the overall complexity of such type of solvers would be smaller than solving the stiff system (54) for all time. The vanilla HMM presented in Sect. 5.1 uses the identity operator for both $\mathscr{Q}$ and $\mathscr{R}$.

Essential questions that need to be resolved for such a scheme include:

- With the system for u, and a choice of $U(u)$, is F well-defined by the procedure defined above? If not, how can it be properly defined?
- What are $\mathscr{R}$ and $\mathscr{Q}$?
- How long should each evolution be computed?
- What does consistency mean?
- What about stability and convergence?

For a fixed given $\varepsilon > 0$, all well known methods will converge as the step-size $H \to 0$, and there is no difference between stiff and non-stiff problems. In [11], convergence for stiff problems ($\varepsilon \ll H$) is defined by the following error:

$$E(H) = \max_{0 \le t_n \le T} \Big(\sup_{0 < \varepsilon < \varepsilon_0(H)} |U(t_n) - U_n| \Big). \tag{59}$$

Here, $\varepsilon_0(H)$ is a positive function of H, serving as an upper bound for the range of ε, and $U(t_n)$ and U_n denote respectively the analytical solution and the discrete solution at $t_n = nH$. With this notion, it is clear that a sensible method has to utilize the slowly varying property of U and generate accurate approximation with a complexity that is sublinear in ε^{-1}.

The problems we are interested in can be described as follows. A full scale system (54) written in the unknown variable u is given, and the oscillations in u have frequency of order ε^{-1}. We shall also call this system the fine scale system. It is assumed that the fine scale system describes the full behavior of the problem. We want to compute the effective behavior of the given full scale system using a number of slowly changing quantities, $(U, \mathscr{V})$. These slowly changing quantities generally defined as functions or functionals of u, and their governing equations may have no explicit analytical formula. Our approach is to discretize the effective equations for $(U, \mathscr{V})$ formally and use numerical solutions of u to extract the missing information needed to evaluate the formal discretization of the governing equations.

Notation 1. Let $u(t;\alpha)$ denote the solution of the initial value problem:

$$\frac{du}{dt} = f_\varepsilon(u,t), \qquad u(t_*) = \alpha, \tag{60}$$

for some arbitrary initial condition at time t_*.

The notation $u(t)$ or u will be reserved for the solution of the same ODE, Equation (60), for $t > 0$ with the given initial condition u_0.

Definition 6. *Let $\mathscr{G}(\cdot,t)$ be a functional of u. The initial data α is said to be consistent with u under $\mathscr{G}$ if $\mathscr{G}(u(\cdot;\alpha),t) = \mathscr{G}(u,t) + \mathscr{O}(\varepsilon^r)$, for some $r > 0$.*

In Kapitza's pendulum problem, the pivot of a rigid pendulum with length l is attached to a strong periodic forcing, vibrating vertically with period ε. The system has one degree of freedom, and can be described by the angle, θ, between the pendulum arm and the upward vertical direction:

$$l\theta'' = \left(g + \frac{1}{\varepsilon}\sin\left(2\pi\frac{t}{\varepsilon}\right)\right)\sin(\theta), \tag{61}$$

with initial conditions $\theta(0) = \theta_0, \theta'(0) = \omega_0$. With large ε, the only stable equilibria are $\theta_0 = n\pi$, corresponding to the pendulum pointing downward. When ε is sufficiently small and both θ_0 and ω_0 are close to 0, the pendulum will oscillate slowly back and forth, *pointing upward,* with displacement $\theta < \theta_{\max}$. The set up of the pendulum and an example solution are depicted in Fig. 5. The period of the slow oscillation is, to leading order in ε, bounded independent of the forcing period ε. On top of the slow motion around the stable $\theta = 0$ configuration, the trajectory of θ exhibits fast oscillations with amplitude and period proportional to ε.

In [32], the second order equation is written as a first order system using $u = (\theta, \omega)$, where ω is the derivative of θ. The slow variable $U = (\Theta, \Omega)$ consists of the weak limit of the angle θ and its derivative $\dot{\theta}$, and the effective force for U can be adequately approximated by the time averaging of the right hand side of (61). However, *the reconstruction operator $\mathscr{R}$ can no longer be the identity operator.* The initial values of u at t_n for each fine scale evolution should be carefully constructed such that the averages of θ matches with Θ in order to keep the correct resonance between the terms $\sin(2\pi t/\varepsilon)$ and $\sin(\theta)$. To this end, the reconstruction operator must carry a correction term when setting up ω at t_n :

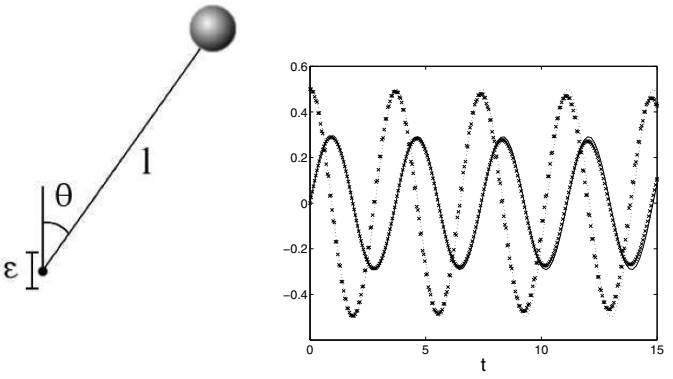
$$\omega_n^0 = \Omega_n - \frac{1}{\varepsilon}\int_{t_n-\varepsilon/2}^{t_n+\varepsilon/2}\int_{t_n}^{t} a_\varepsilon\left(\frac{s}{\varepsilon}\right)\sin(\theta_n(s))dsdt.$$

Consistency of the described multiscale solver to this type of system is thus established.

Problem 1 (Closure). Given V which consists of a set of slow variables or functionals of u, determine the set of extended variables $U = (V,W) : [0,T] \mapsto \mathbb{R}^d \times \mathbb{R}^s$, whose components are functions or functionals of u, so that there exists a function $\mathscr{F}$ independent of ε such that the solution $U(t)$ of the ODE

$$\frac{d}{dt}U = \frac{d}{dt}\begin{pmatrix} V \\ W \end{pmatrix} = \mathscr{F}(U,t) + \mathscr{O}(\varepsilon), \tag{62}$$

is equivalent to its evaluation using the whole scale solution u, i.e,



(a) (b)

Fig. 5. (a) Kapitza's pendulum; (b) The slow scale solutions to equation (61); Three orders of magnitude ($\varepsilon = 10^{-6}$) separate the period of the slow oscillation apparent in the graphs from the fast oscillation.

$$U(t) = U(u,t) = \begin{pmatrix} V(u,t) \\ W(u,t) \end{pmatrix},$$

where $U(t)$ denotes the ODE solution, and $U(u,t)$ the functional evaluation using $u(t)$, and similarly for V and W.

One of our strategies is to look for algebraic functions α and ψ when constructing $W(u,t)$. When these functions are composed of $u(t)$ and viewed as functions of time, $\alpha_\varepsilon(t) = \alpha(u(t))$ and $\psi_\varepsilon(t) = \psi(u(t))$, they should satisfy the following conditions:

1. α and ψ are linear combinations of some simple functions of u;
2. $d\alpha_\varepsilon/dt$ is bounded independent of ε;
3. $d^\nu \langle \psi_\varepsilon \rangle / dt^\nu > \delta > 0$ for some $0 \leq \nu$ and for some δ independent of ε;
4. $\alpha_\varepsilon(t)$ converges pointwise to a smooth function $\bar{\alpha}(t)$, and ψ_ε weakly to a continuous function Ψ.

Here, $\langle \psi_\varepsilon \rangle$ denotes a moving average with respect to a kernel, as described in Sect.4.2. These approaches are motivated by the analysis of resonance, the averaging methods, see e.g. [25, 5, 1, 3], and our previous work on Kapitza's pendulum and a few other model problems. Another interesting point of view makes use of the idea of Young measures [4].

In practice, we do not have $u(t)$, since we do not solve the stiff equation for a long time interval independent of ε. However, the solution U to the closure problem defines an equivalence class for the initial conditions for u. As long as an initial data is selected such that it is consistent to $u(t)$ with respect to $U(t)$, $U(t)$ is properly evolved. Instead, our strategy is to compute the solution $u(\cdot;a)$ for a duration that vanishes with ε, starting from a specified time and using some initial values a. Once $U(t)$ is approximated, we can approximate dU/dt numerically without explicitly

evaluating $\mathscr{F}$. Naturally, this initial value α should be consistent with u with respect to the functional $U(u,t)$:

Problem 2 (Reintialization/reconstruction). Given a functional $U(u,t)$ and its value U_n at t_n that specifies a set of constraints.

$$\text{Find } a_n \in \mathbb{R}^d \text{ such that } U(u(\cdot;a_n),t_n) = U_n + \mathscr{O}(\varepsilon^p) \text{ for some } p > 0.$$

Note that unlike the common constraints of conserved integrals in the computations of Hamiltonian systems, we consider constraints, such as those specified by the components of U, that can be slowly varying in time.

In summary, our multiscale method is outlined as follow: Assuming $U(t_j)$ is known at $t = t_j$,

1. Find a_j that solves $U(u(\cdot;a_j),t_j) \simeq U(t_j)$ (Reinitialization);
2. Solve the given stiff equation and obtain $u(t;a_j)$ for $t_j \le t \le t_j + \eta_\varepsilon$ (Microscale solution);
3. Evaluate $\mathscr{F}(t_j) = dU/dt$ at t_j using $u(t;a_j)$;
4. Use $\mathscr{F}(t_j)$ and $U(t_j)$ to get U at $t_j + \Delta t$. (Macroscale solution)

Note that Δt should be independent of ε and η_ε vanish with ε.

In the following examples the slow behavior is approximated using functions only (the slow chart) and not functionals.

5.3 Example: an expanding spiral

Consider the system (34) describing the expanding spiral

$$\begin{aligned} x' &= -\varepsilon^{-1}y + x, \qquad x(0) = 1, \\ y' &= \varepsilon^{-1}x + y, \qquad y(0) = 0. \end{aligned} \tag{63}$$

Previously, in Sect. 2.4, it was shown that $(\xi,\phi) = (x^2+y^2, \tan^{-1}(y/x))$ is a slow chart for (63). The time evolution of the only slow variable ξ takes the form

$$\xi' = \langle \xi' \rangle_\eta + \mathscr{O}(\varepsilon) = \langle 2xx' + 2yy' \rangle_\eta . \tag{64}$$

This motivates the following multiscale algorithm for approximating $\xi'(t)$. For simplicity, we apply a macroscopic forward Euler solver with step size H. Any consistent and stable integrator can be used as micro-solver. We denote $t_n = nH$ and by x_n, y_n and ξ_n our approximation for $x(t_n)$, $y(t_n)$ and $\xi(t_n)$, respectively. Note that x_n and y_n do not have to be close to $x(t_n)$ and $y(t_n)$. We only require that the slow variable is approximated. The algorithm is depicted in Fig. 6.

1. Initial conditions: $(x(0),y(0)) = (x_0,y_0)$, $n = 0$.
2. Micro-simulation: Solve (63) in $[t_n - \eta/2, t_n + \eta/2]$ with initial conditions

$$(x(t_n),y(t_n)) = (x_n,y_n).$$

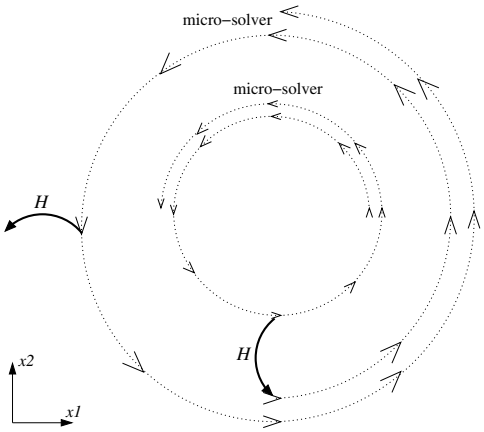


Fig. 6. Two macroscopic steps for the HMM algorithm in the expanding spiral example (63).

3. Force estimation: Approximate ξ' by $\Delta\xi_n = \langle 2xx' + 2yy' \rangle_\eta$. The step involves convoluting $2xx' + 2yy'$ with an averaging kernel as discussed in Sect. 4.2.
4. Macro-step (forward Euler): $\xi_{n+1} = \xi_n + H\Delta\xi_n$.
5. Reconstruction (second order accurate): $(x_{n+1}, y_{n+1}) = (x_n, y_n) + HF_n$, where F_n is the least squares solution of the linear system

$$F_n \cdot \nabla\xi(x_n, y_n) = \Delta\xi_n$$

6. $n = n + 1$. Repeat steps 2–5 to time T.

5.4 HMM using slow charts

Suppose an ODE system of the form (35) admits a slow chart (ξ, ϕ), where $\xi = (\xi^1, \ldots, \xi^k) \in \mathbb{R}^k$ are slow and $\phi \in \mathbb{S}^1$ is fast. In the next section we will see that many highly oscillatory systems indeed admit a slow chart of that form. Then, the algorithm suggested in the previous section can be easily generalized as follows. As before, for simplicity we concentrate on the forward Euler case. Higher order methods are considered in [1]. Approximated quantities at the n'th macroscopic time step are denoted by a subscript n.

1. Initial conditions: $x(0) = x_0$, $n = 0$.
2. Micro-simulation: Solve (35) in $[t_n - \eta/2, t_n + \eta/2]$ with initial conditions $x(t_n) = x_n$
3. Force estimation: Approximate ξ' by $\Delta\xi_n = \langle \nabla\xi \cdot x' \rangle_\eta$ using convolution with an averaging kernel.
4. Macro-step (forward Euler): $\xi_{n+1} = \xi_n + H\Delta\xi_n$.

5. Reconstruction (second order accurate): $x_{n+1} = x_n + HF_n$, where F_n is the least squares solution of the linear system

$$F_n \cdot \nabla \xi(x_n) = \Delta \xi_n. \tag{65}$$

6. $n = n+1$. Repeat steps 2–5 to time T.

Complexity

In this section we analyze the accuracy of the suggested method outlined above. Each step of the approximations preformed in our algorithm introduces a numerical error. In order to optimize performance, the different sources of errors are balanced to a fixed prescribed accuracy Δ. We show how the different parameters: ε, η, h and H scale with Δ in order to have a global accuracy of order Δ. Note that the maximal possible accuracy is $\Delta = \varepsilon$, since this is the error introduced by simulating the averaged equation rather than the original one. We also study the Δ dependence of the complexity of the algorithm.

We begin with estimating the error in our evaluation of the averaged force $\Delta\xi_n$. There are several sources of errors:

- Global error in each micro-simulation. Using an m'th order method with step size h the global error is $\eta h^m/\varepsilon^{m+1}$.
- Quadrature error in $K'_\eta * \xi$: Using a quadrature formula of degree r the error is $\eta h^m/\varepsilon(m+1)$. However, due to the regularity of the kernel used $K \in C^q$, the integrand is smooth and periodic. Hence, the coefficients of its Fourier decomposition decay very fast. As a result, it is advantageous to use the trapezoidal rule, which is exact for $e^{2\pi ikx}$, $k \in \mathbb{N}$. This implies that the quadrature error is typically very small and can be neglected.
- Approximating $\Delta\xi_n$ by $\langle \nabla\xi \cdot x' \rangle_\eta$: Using a kernel $K \in \mathbb{K}^{p,q}$ the error is the larger between η^p and $(\varepsilon/\eta)^q/\eta$. Note that we are losing one order of η compared to (48) since $\Delta\xi_n$ is found through integration by parts (cf. Sect. 5.4). The above two bounds to the averaging error are equal if $\eta^{p+q+1} = \varepsilon^q$, where, for large η, the term η^p dominates, while for small η the other. Since we would like to optimize our complexity, it is always preferable to work in the latter regime. Hence, we can take the averaging error to be $(\varepsilon/\eta)^q/\eta$.

Balancing all terms yields the optimal scaling of the simulation parameters with Δ.

The global accuracy of integrating the original full ODE to time $T = \mathcal{O}(1)$ using a macro-solver of order s with step size H is, assuming errors are accumulative,

$$E \le D \max\left\{ H^s, \frac{\eta h^m}{\varepsilon^{m+1}}, \frac{\varepsilon^q}{\eta^{q+1}} \right\}, \tag{66}$$

For some $D > 0$. For short hand we drop the constant in all following expressions. Balancing the different sources of errors to a prescribed accuracy Δ yields

$$\begin{aligned} \eta &= \varepsilon^{\frac{q}{q+1}} \Delta^{-\frac{1}{q+1}}, \\ H &= \Delta^{\frac{1}{s}}, \\ h &= \varepsilon^{1+\frac{1}{m(q+1)}} \Delta^{\frac{s+1}{sm}+\frac{1}{m(q+1)}}. \end{aligned} \tag{67}$$

The complexity is then

$$C = \frac{\eta}{h}\frac{T}{H} = \varepsilon^{-\frac{m+1}{m(q+1)}} \Delta^{-\frac{1}{s}-\frac{s+1}{sm}-\frac{m+1}{m(q+1)}}. \tag{68}$$

With a smooth kernel we can consider the $q \to \infty$ limit. In this case the complexity estimate is reduced to

$$C(q \to \infty) = \Delta^{-\frac{1}{s}-\frac{s+1}{sm}}. \tag{69}$$

Figure 7 depicts the relative error of the HMM approximation compared to the analytical solution of the expanding spiral example (34). The kernel was constructed from polynomials to have exactly two continuous derivatives and a single vanishing moments, i.e., $q = 2$ and $p = 1$. Fourth order Runge–Kutta schemes were used for both the micro- and the macro-solvers. The simulation parameters are chosen to balance all errors as discussed above.

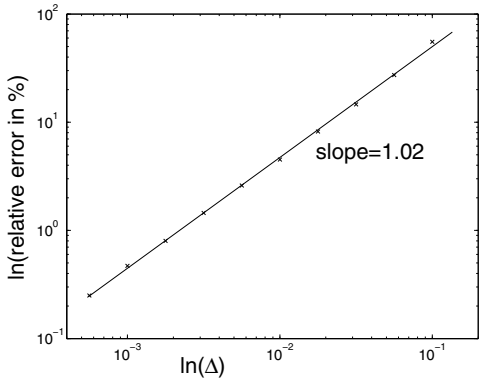


Fig. 7. A log-log plot of the relative error of the HMM approximation to a linear ODE compared to the exact solution: $E = \max_{t_n \in [0,T]} 100 \times |\xi_{\mathrm{HMM}}(t_n) - \xi_{\mathrm{exact}}(t_n)|/|\xi_{\mathrm{exact}}(t_n)|$, as a function of Δ.

From the parameter scaling (67) it is clear that the step size of the macro-solver, H, does not depend on the stiffness ε, but only on the prescribed accuracy Δ. Our algorithm is therefore multiscale is the sense that it converges uniformly for all $\varepsilon < \varepsilon_0$ [11]. More precisely, denote the sample times of the macro-solver by $t_0 = 0, \ldots, t_N = T$ and the corresponding numerical approximations for x by $x_0, \ldots, x_N$. The exact solution is denoted $x(t)$. We have that, for any variable $\alpha(x)$ that is slow with respect to $x(t)$

$$\lim_{H \to 0} \sup_{k=0,\ldots,N} \sup_{\varepsilon < \varepsilon_0} |\alpha(x(t_k)) - \alpha(x_k)| \to 0. \tag{70}$$

Note that the order of the limits is important.

5.5 Almost linear oscillators in resonance

Consider an ODE system of the form

$$x' = \frac{1}{\varepsilon}Ax + f(x), \tag{71}$$

where, $x \in \mathbb{R}^d$ and A is an $d \times d$ real diagonalizable matrix with purely imaginary eigenvalues $\pm i\omega_1, \ldots, \pm i\omega_r$, $2r = d$. In addition, we assume that all oscillatory modes are in resonance. This implies that the ratio of every pair of frequencies is rational, i.e., for all $i, j = 1 \ldots r$, there exist integers m_{ij} and n_{ij}, such that $m_{ij}\omega_i = n_{ij}\omega_j$.

Theorem 6. *There exists a slow chart (ξ, ϕ) in $\mathbb{R}^d \setminus \{\xi_i \neq 0, \forall i\}$ for (71) such that all the coordinates of ξ are polynomial in x and $\phi \in \mathbb{S}^1$.*

The theorem is proven in [1]. As an example, consider (71) with

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 \\ 0 & 0 & -2 & 0 \end{pmatrix}.$$

Changing variables so that A is diagonalized yields the complex system

$$z = \frac{1}{\varepsilon}\begin{pmatrix} i & 0 & 0 & 0 \\ 0 & -i & 0 & 0 \\ 0 & 0 & 2i & 0 \\ 0 & 0 & 0 & -2i \end{pmatrix} z + f(x)$$

where $z = (z_1, z_1^*, z_2, z_2^*)^T$ and z^* denotes the complex conjugate of z. It is easily verified that the following are slow variables

$$\begin{aligned} \xi_1 &= z_1 z_1^*, \\ \xi_2 &= z_2 z_2^*, \\ \xi_3 &= z_1^2 z_2^*. \end{aligned}$$

Transforming back to the original coordinates $x = (x_1, v_1, x_2, v_2)$ the slow variables become the real polynomials

$$\begin{aligned} \xi_1 &= x_1^2 + v_1^2, \\ \xi_2 &= x_2^2 + v_2^2, \\ \xi_3 &= x_1 x_2^2 + 2v_1 x_2 v_2 - x_1 v_2^2. \end{aligned}$$

The first two variables, ξ_1 and ξ_2 correspond to the square of the amplitude of the two harmonic oscillators described by (x_1, v_1) and (x_2, v_2), respectively. The third variable, ξ_3, corresponds to the relative propagation of phase in the two oscillators. It is slow because, to leading order in ε, the phase of (x_2, v_2) increases twice as fast as that of (x_1, v_1).

Exercise 5. Verify that $\nabla\xi_1$, $\nabla\xi_2$, and $\nabla\xi_3$ are not linearly dependent in any region in $\mathbb{R}^4 \setminus Q$ where Q is the zeros of ξ_1, ξ_2, and ξ_3.

Fully nonlinear oscillators

Dealing with non-linear oscillators is more complicated than linear ones. However, the slow behavior of weakly coupled systems of oscillators such as Van der Pol, relaxation and Volterra-Lotka can still be described using some generalization of amplitude and relative phase. This is beyond the scope of these notes. We refer to [2] for further reading.

6 Computational exercises

Computer exercise 1. Let $u = (x, y, z)$ and

$$f_\varepsilon(x,y,z) = \begin{pmatrix} a & \frac{1}{\varepsilon} & 0 \\ -\frac{1}{\varepsilon} & b & 0 \\ 0 & 0 & -\frac{1}{10} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ x^2 + cy^2 \end{pmatrix}. \tag{72}$$

The equation for u is

$$u' = f_\varepsilon(u), \qquad u(0) = (1,0,1).$$

Take $\varepsilon = 10^{-4}$, $a = b = 0$ and $c = 1$. Find approximations for $z(t)$ in $0 < t \leq 1$ using the following schemes and compare with the analytical solution. Plot the trajectories of your approximations of $x(t)$ and $y(t)$ on the xy-plane, and the graph $z(t)$ as a function of time. Explain what you observe in each case.

(a) Forward Euler using $\triangle t = \varepsilon/50$.
(b) Backward Euler for x and y and Forward Euler for z, using $\triangle t = 0.1$.
(c) Verlet method or Midpoint rule for x and y, and Forward Euler for z, using $\triangle t = \varepsilon/50$.
(d) Solve this problem by the HMM–FE–fe method (see below), with $\mathscr{Q} = \mathscr{R} = I$ (see Sect. 5.2). $h = \varepsilon/50$, $H = 0.1$, and $hM = 2 \cdot 10^{-3}$.
(e) Derive linear stability criteria on H for HMM–FE–fe, assuming that $h = c_0\varepsilon$.
(f) Let $a = b = 1$ in the system defined above. Solve it by the same HMM–FE–fe scheme with the same parameters as in (d). Does this scheme correctly approximate the behavior of z in the time interval $0 < t \leq 1$? Explain.

HMM–FE–fe scheme for $u' = f_\varepsilon(u)$.

- Macroscale with Forward Euler (FE)

$$U^{n+1} = U^n + HF^n, \qquad U^0 = \mathscr{Q}(u_0)$$

- Microscale with Forward Euler (fe)

$$u_{k+1}^n = u_k^n + hf_\varepsilon(u_k^n), \qquad k = 0, \pm 1, \cdots, \pm M,$$
$$u_0^n = \mathscr{R}(U^n).$$

- Averaging

$$F^n := \frac{1}{2M} \sum_{k=-M}^{M} K^{\cos}\left(\frac{k}{2M}\right) f_\varepsilon(u_k^n),$$
$$K^{\cos}(t) = \frac{1}{2}\chi_{[-1,1]}(t)\,(1+\cos(\pi t))\,,$$
$$\chi_{[-1,1]}(x) = \begin{cases} 1, & -1 \le x \le 1, \\ 0, & \text{otherwise.} \end{cases}$$

Computer exercise 2. Following the previous problem, define the slow variable

$$\xi(x,y) = x^2 + y^2 \text{ and } \xi(t) := x^2(t) + y^2(t),$$

where $x(t)$ and $y(t)$ are defined in (72).

(a) Show that $d\xi/dt$ can be approximated by averaging:

$$\left| \frac{d\xi}{dt}(t_n) - \int_{-\infty}^{\infty} -\frac{d}{dt} K^{\cos}\left(\frac{t_n - t}{2Mh}\right) \left(x^2(t) + y^2(t)\right) dt \right| \le C\eta^p.$$

Find p.

(b) Modify your previous HMM–FE–fe code to HMM–FE–rk4 (see below) as follows and determine if the dynamics of z is accurately approximated by this new scheme. Plot your approximations as in the previous problem. Explain your findings.

(c) Do the same thing as in the previous problem, but with $c = 0$. Does your multiscale algorithm work? Why?

Constrained HMM–FE–rk4 scheme for $u' = f_\varepsilon(u)$.

- Macroscale with Forward Euler

$$U^{n+1} = U^n + HF^n, \qquad U^0 = \mathscr{Q}(u_0).$$

- Microscale with Runge–Kutta 4 (rk4)

$$u_{k+1}^n = \text{rk4}(u_k^n, h), \qquad k = 0, \pm 1, \cdots, \pm M,$$
$$u_0^n = \mathscr{R}(U^n).$$

Here rk4 is an explicit Runge–Kutta 4 routine using step size h.

$$\text{rk4}(y,h) = y + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4),$$

$$k_1 = hf_\varepsilon(y), \quad k_2 = hf_\varepsilon(y + \frac{1}{2}k_1), \quad k_3 = hf_\varepsilon(y + \frac{1}{2}k_2), \quad k_4 = hf_\varepsilon(y + k_3).$$

- Averaging

$$dz^n := \frac{1}{2M}\sum_{k=-M}^{M} K^{\cos}\left(\frac{k}{2M}\right)\left(x_k^n \cdot x_k^n + c y_k^n \cdot y_k^n - \frac{z_k^n}{10}\right).$$

$$d\xi^n := \frac{1}{2M}\sum_{k=-M}^{M} G\left(\frac{k}{2M}\right)(x_k^n \cdot x_k^n + y_k^n \cdot y_k^n),$$

  where $G(\frac{k}{2M}) := \frac{-1}{2Mh}\frac{d}{dt}K^{\cos}(\frac{t}{2Mh})$.
- Evaluate effective force
  Find a unit vector dX^n such that

$$d\xi^n = \nabla_{x,y}\xi|_{x_k^n, y_k^n} \cdot dX^n.$$

$$F^n := \begin{pmatrix} dX^n \\ dz^n \end{pmatrix}.$$

Computer exercise 3. Consider the inverted pendulum equation:

$$l\theta'' = \left(g + \frac{1}{\varepsilon}\sin\left(2\pi\frac{t}{\varepsilon}\right)\right)\sin(\theta). \tag{73}$$

Let $\omega = \theta'$, rewrite it into a system of first order equations for (θ, ω). Let $\Omega_{n+\frac{1}{2}}$ denote the averaged macroscopic angular momentum at time $(n+\frac{1}{2})H$ and Θ_n be the averaged macroscopic angle. Compute the inverted pendulum solutions by using the parameters $\varepsilon = 10^{-6}$, $(\Theta_0, \Omega_0) = (0.0, -0.4)$, $g = 0.1$, $l = 0.05$. Experiment with $\eta = 10\varepsilon$ and 30ε.
This problem is analyzed in [32].

HMM for the inverted pendulum problem.

- Macroscale with Verlet
  Given $U^n = (\Theta^n, \Omega^n)$, for $n = 0, 1, 2, ldots$

$$\begin{aligned}
\Omega^{n+\frac{1}{2}} &= \Omega^n + \frac{H}{2}\cdot\tilde{F}^n, \\
\Theta^{n+1} &= \Theta^n + H\cdot\Omega^{n+\frac{1}{2}}, \\
\Omega^{n+1} &= \Omega^{n+\frac{1}{2}} + \frac{H}{2}\cdot\tilde{F}^{n+1},
\end{aligned}$$

  Here, $\tilde{F}[\theta^n, \omega^n]$ denotes the averaged force using the solutions whose values at $t_n = nH$ are (θ^n, ω^n).
- Microscale evolution
  Solve $l\theta'' = (g + \frac{1}{\varepsilon}\sin(2\pi\frac{t}{\varepsilon}))\sin(\theta)$ for $t_{n-\eta} \le t \le t_{n+\eta}$ with the "reconstructed initial"

$$\theta(t_n) = \Theta^n,$$
$$\omega(t_n) = \theta'(t_n) = \mathscr{R}(\Theta^n, \Omega^n) := \Omega^n - \sin(\Theta^n)\frac{\cos\left(2\pi\frac{t_n}{\varepsilon}\right)}{2\pi l}.$$
$$\left(\omega(t_n) \approx \Omega^n - \left\langle \int_{t_n}^{t} a_\varepsilon\left(\frac{s}{\varepsilon}\right)\sin(\theta(s))ds \right\rangle\right)$$

- Average
  Using the solution computed in the microscale evolutions around t_n. Evaluate

$$\tilde{F}^n = \left\langle (g + \frac{1}{\varepsilon}\sin\left(2\pi\frac{t}{\varepsilon}\right))\sin(\theta) \right\rangle_\eta = K_\eta * f(t_n),$$

  where

$$K_\eta(t) = \frac{422.11}{\eta}\exp\left[5\left(\frac{4t^2}{\eta^2} - 1\right)^{-1}\right],$$

  and

$$f(t) = \left(g + \frac{1}{\varepsilon}\sin\left(2\pi\frac{t}{\varepsilon}\right)\right)\sin(\theta(t)).$$

  Use the Trapezoidal rule to approximate the above convolutions.

Computer exercise 4. The following is a well studied system taken from the theory of stellar orbits in a galaxy

$$\begin{aligned} r_1'' + a^2 r_1 &= \varepsilon r_2^2 \\ r_2'' + b^2 r_2 &= 2\varepsilon r_1 r_2 \end{aligned}.$$

Rewrite the above equation into the standard form (35).

(a) To see how resonances occur, change into polar coordinates and take $a = \pm 2b$.
(b) Let $a = 2$ and $b = 1$. Find a maximal slow chart.
(c) Apply the HMM algorithm described in Sect. 5.4 to approximate the slow behaviour of the system.

References

1. G. Ariel, B. Engquist, and Y.-H. Tsai. A multiscale method for highly oscillatory ordinary differential equations with resonance. *Math. Comp.*, 2008. To appear.
2. G. Ariel, B. Engquist, and Y.-H. Tsai. Numerical multiscale methods for coupled oscillators. *Multiscale Model. Simul.*, 2008. Accepted.
3. V.I. Arnol'd. *Mathematical methods of classical mechanics*. New York, Springer-Verlag, 2 edition, 1989.
4. 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.
5. N. N. Bogoliubov and Yu. A. Mitropolski. *Asymptotic Methods in the Theory of Nonlinear Oscillations*. Gordon and Breach, New York, 1961.
6. R. Car and M. Parrinello. Unified approach for molecular dynamics and density functional theory. *Phys. Rev. Lett.*, 55(22):2471–2475, 1985.

7. A.J. Chorin. A numerical method for solving incompressible viscous flow problems. *J. Comp. Phys.*, 2:12–26, 1967.
8. G. Dahlquist. A special stability problem for linear multistep methods. *Nordisk Tidskr. Informations-Behandling*, 3:27–43, 1963.
9. G. Dahlquist, L. Edsberg, G. Skollermo, and G. Söderlind. Are the numerical methods and software satisfactory for chemical kinetics? In *Numerical Integration of Differential Equations and Large Linear Systems*, volume 968 of *Lecture Notes in Math.*, pages 149–164. Springer-Verlag, 1982.
10. W. E. Analysis of the heterogeneous multiscale method for ordinary differential equations. *Commun. Math. Sci.*, 1(3):423–436, 2003.
11. W. E and B. Engquist. The heterogeneous multiscale methods. *Commun. Math. Sci.*, 1(1):87–132, 2003.
12. W. E, B. Engquist, X. Li, W. Ren, and E. Vanden-Eijnden. Heterogeneous multiscale methods: A review. *Commun. Comput. Phys.*, 2(3):367–450, 2007.
13. B. Engquist and O. Runborg. Computational high frequency wave propagation. *Acta Numerica*, 12:181–266, 2003.
14. B. Engquist and Y.-H. Tsai. Heterogeneous multiscale methods for stiff ordinary differential equations. *Math. Comp.*, 74(252):1707–1742, 2003.
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. D. Ter Haar, editor. *Collected Papers of P.L. Kapitza*, volume II. Pergamon Press, 1965.
18. E. Hairer, C. Lubich, and G. Wanner. *Geometric numerical integration*, volume 31 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, 2002. Structure-preserving algorithms for ordinary differential equations.
19. E. Hairer and G. Wanner. *Solving ordinary differential equations. II*, volume 14 of *Springer Series in Computational Mathematics*. Springer-Verlag, 1996.
20. J. Hale. *Ordinary differential equations*. New York, Wiley-Interscience, 1969.
21. J. B. Keller. Geometrical theory of diffraction. *J. Opt. Soc. Amer.*, 52:116–130, 1962.
22. J. Kevorkian and J. D. Cole. *Perturbation Methods in Applied Mathematics*, volume 34 of *Applied Mathematical Sciences*. Springer-Verlag, New York, Heidelberg, Berlin, 1980.
23. J. Kevorkian and J. D. Cole. *Multiple Scale and Singular Perturbation Methods*, volume 114 of *Applied Mathematical Sciences*. Springer-Verlag, New York, Berlin, Heidelberg, 1996.
24. H.-O. Kreiss. Problems with different time scales. *Acta Numerica*, 1:101–139, 1992.
25. H.-O. Kreiss and J. Lorenz. Manifolds of slow solutions for highly oscillatory problems. *Indiana University Mathematics Journal*, 42(4):1169–1191, 1993.
26. B. Leimkuhler and S. Reich. *Simulating Hamiltonian dynamics*, volume 14 of *Cambridge Monographs on Applied and Computational Mathematics*. Cambridge University Press, 2004.
27. L. R. Petzold. An efficient numerical method for highly oscillatory ordinary differential equations. *SIAM J. Numer. Anal.*, 18(3):455–479, 2003.
28. L. R. Petzold, O. J. Laurent, and Y. Jeng. Numerical solution of highly oscillatory ordinary differential equations. *Acta Numerica*, 6:437–483, 1997.
29. J. A. Sanders and F. Verhulst. *Averaging Methods in Nonlinear Dynamical Systems*, volume 59 of *Applied Mathematical Sciences*. Springer-Verlag, New York, Berlin, Heidelberg, Tokyo, 1985.

30. R. E. Scheid. The accurate numerical solution of highly oscillatory ordinary differential equations. *Math. Comp.*, 41(164):487–509, 1983.
31. R. E. Scheid. Difference methods for problems with different time scales. *Math. Comp.*, 44(169):81–92, 1985.
32. R. Sharp, Y.-H. Tsai, and B. Engquist. Multiple time scale numerical methods for the inverted pendulum problem. In B. Engquist, P. Lötstedt, and O. Runborg, editors, *Multiscale methods in science and engineering*, volume 44 of *Lect. Notes Comput. Sci. Eng.*, pages 241–261. Springer, Berlin, 2005.
33. E. Vanden-Eijnden. On HMM-like integrators and projective integration methods for systems with multiple time scales. *Commun. Math. Sci.*, 5(2):495–505, 2007.