ANALYSIS OF THE HETEROGENEOUS MULTISCALE METHOD FOR ORDINARY DIFFERENTIAL EQUATIONS*

WEINAN E †

1. Introduction

This paper is concerned with the analysis of a new class of efficient numerical methods for systems of ordinary differential equations (ODEs) with disparate time scales. We will be interested in stiff ODEs of the type that arise in chemical kinetics, for example:

$$\begin{cases} \dot{x} = -\dfrac{1}{\varepsilon}(x - f(y)) \\ \dot{y} = g(x, y) \end{cases} \tag{1.1}$$

as well as oscillatory systems of the type that arise in molecular dynamics and celestial mechanics, a simple example being the system

$$\begin{cases} \dot{\varphi} = \dfrac{1}{\varepsilon}\omega(I) + f(\varphi, I, \varepsilon) \\ \dot{I} = g(\varphi, I, \varepsilon) \end{cases} \tag{1.2}$$

studied in averaging methods [1]. Here f and g are assumed to be periodic in φ and bounded as $\varepsilon \to 0$. In the standard terminology of multiscale analysis, we would call x and φ the fast variables of these systems, y and I the slow variables. But we also have a particular interest on systems for which the fast and slow variables exist but cannot be explicitly identified beforehand.

At the present time, there does not exist a unified strategy for dealing with both problems of type (1.1) and (1.2). There is, however, a large literature on stiff systems of the type (1.1) and oscillatory systems of the type (1.2) separately. Starting from the pioneering work of Dahlquist and Gear [5], there has been extensive work on designing efficient numerical methods for stiff ODEs [7], such as the backward differentiation formula, implicit Runge-Kutta methods [5], extrapolation methods and Rosenbrock methods [7, 8, 6]. There is also extensive work on oscillatory systems, some of which are analytical [1], and some are numerical [9].

We will apply the framework of the heterogeneous multiscale method (HMM) [3], which is a general methodology for problems with multiscales. In the case of two scales, a macroscale and a microscale, HMM consists of two components: selection of a conventional macroscale solver, here a standard ODE solver, and estimating the effective forces used in the macroscale solver by performing numerical experiments using the microscale model and processing the data obtained. The procedure can be iterated if the system has more than two separated scales.

There are a number of related numerical methods for systems with multiple time scales, in particular for stiff ODEs. Even though the most popular numerical methods for stiff ODEs stem from implicit methods such as the backward differentiation formula, a number of explicit methods have also been proposed [8, 2, 7]. In the simplest version, these methods are Runge-Kutta in nature, each stage of which is a forward

Euler step. For a given order, these methods have a variable number of internal stages (Euler steps for the simplest method) chosen to satisfy stability conditions. In the special case when the spectrum of the problem falls into two widely separated clusters on the real axis, such as the case in (1.1), the composite Runge-Kutta method amounts to perform a number of forward Euler steps with small step sizes followed by a few forward Euler steps with large step size, as a consequence of the structure of the optimal stability polynomials. Similar ideas were also proposed more recently in [6]. There are two main differences between HMM and the work in [8, 2, 7, 6]. The first is that HMM makes explicit use of two solvers: a macroscale solver and a microscale solver, whereas the work of Lebedev, Gear et al. uses one solver with different time steps. The other is that HMM uses more sophisticated data processing techniques in order to extract the macroscale data.

Another interesting example of a multiscale methodology is in the work of Vanden-Eijnden, which was started back in 1999, on stochastic systems or systems on long time scales for which the stochastic effects are important [10].

Our analysis of HMM also proceeds along the general principle discussed in [3], which state that HMM is stable if the macroscale solver is stable, and the error comes from standard truncation error of the macroscale solver plus the error in the estimation of the effective forces. The latter depends on the nature of the problem, the microscale solver, as well as the data processing technique used.

This paper is organized as follows. We first present two different versions of HMM. Error analysis of HMM is carried out in Section 3 for stiff systems and in Section 4 for oscillatory systems. A general framework is presented in Section 5.

2. Two versions of HMM

We will discuss two different versions of HMM, depending on whether the slow and fast variables in the system can be explicitly identified.

We first discuss a seamless version of HMM proposed in [4]. Consider the system of ODEs

$$\dot{Z} = F(Z, \varepsilon) \tag{2.1}$$

As the macroscale solver, we may select a conventional explicit ODE solver such as a Runge-Kutta scheme or a linear multi-step method. We may also select special purpose solvers such as the symplectic integrators. For illustration, let us assume that we will use forward Euler as the macroscale solver. We can then express it as

$$Z^{n+1} = Z^n + \Delta t \tilde{F}(Z^n) \tag{2.2}$$

The time step Δt resolves the macroscale dynamics of interest, but not the small scales.

To estimate the effective forces on the macroscale, here the $\tilde{F}(Z^n)$, we choose a microscale solver for (2.1) using a time step δt that resolves the small scale. Again a number of choices can be made. In the simplest case we may simply pick forward Euler and we can formulate the microscale solver at the macroscale time step n as follows

$$
\begin{aligned}
&Z^{n,m+1} = Z^{n,m} + \delta t\, F(Z^{n,m}, \varepsilon), m = 0, 1, \ldots, M-1 \\
&Z^{n,0} \text{ suitably chosen.}
\end{aligned}
\tag{2.3}
$$

We may even choose different microscale solvers at different n. Having $\{Z^{n,m}\}_{m=0}^{M-1}$,

we can estimate $\tilde{F}(Z^n)$ according to

$$\tilde{F}(Z^n) = \frac{1}{M} \sum_{m=0}^{M-1} K_{m,M} F(Z^{n,m}, \varepsilon) \qquad (2.4)$$

where the weights $\{K_{m,M}\}$ should satisfy the constraint

$$\frac{1}{M} \sum_{m=0}^{M-1} K_{m,M} = 1 \qquad (2.5)$$

The specific choice of $\{K_{m,M}\}$ will affect the overall accuracy of HMM. M should be chosen such that the right hand side of (2.4) has converged to a (quasi)stationary value with desired accuracy. (2.4) is called an F-estimator [3]. Extensive numerical results using this methodology can be found in [4].

For some systems such as (1.1) or (1.2), we can identify explicitly the slow and fast variables in the system. In this case, we can make use of such information and formulate an alternative version of HMM. For (1.1), this can be done as follows. Since y is the slow variable, we pick a macroscale scheme for y, e.g. forward Euler

$$y^{n+1} = y^n + \Delta t \tilde{g}(y^n)$$

To estimate $\tilde{g}(y^n)$, we proceed as before and select a conventional ODE solver as the microscale solver with some initial data $x^{n,0}$. In the case of forward Euler, we have

$$x^{n,m+1} = x^{n,m} - \frac{\delta t}{\varepsilon}(x^{n,m} - f(y^n))$$

$m = 0, 1, \cdots, M-1$. Notice that we have constrained the slow variable y^n to be fixed. We then use

$$\tilde{g}(y^n) = \frac{1}{M} \sum_{m=0}^{M-1} K_{m,M} g(x^{n,m}, y^n)$$

As we will see below, better accuracy can be obtained by constraining the slow variable for the microscale solver.

The method for (1.2) is similar. In that case, the slow variable is I which is kept fixed in the microscale solver.

Extensions to higher order macro- and micro-solvers are straightforward.

For the analysis of these methods, we proceed along the lines of a general principle discussed in [3]. In particular, we have as a consequence of Theorem 5.5 in [3]:

THEOREM 2.1. *HMM is stable if the macroscale solver is stable. Moreover for any given constant T there exists a constant C such that*

$$|E^n| \leq C((\Delta t)^k + e(HMM))$$

if $n\Delta t \leq T$, where E^n is the error between the HMM solution and the slow component of the exact solution (a precise definition is given below), k is the order of the macroscale solver, $e(HMM)$ is the error in the estimation of the macroscale data.

The key to the analysis is to estimate $e(\text{HMM})$, which depends on a number of important components of HMM, in particular the filter K.

3. Error estimates for dissipative systems

For clarity, we will consider the example of (1.1). Other dissipative systems with exponential convergence to the slow manifold can be reduced to essential systems of the form (1.1) in a neighborhood of the slow manifold. We will further assume that f and g are smooth and bounded. We will also assume that $\{Z^{n,0}\}$ is bounded.

In the limit as $\varepsilon \to 0$, one expects that x will be close to $f(y)$, and hence the dynamics of the slow variable y is given by

$$\dot{y} = g(f(y), y) = G(y) \tag{3.1}$$

In fact it can be easily shown that if we denote by y_0 the solution of (3.1), and $(x_\varepsilon, y_\varepsilon)$ the solution of (1.1), such that $y_\varepsilon(0) = y_0(0)$. Then

$$|y_\varepsilon(t) - y_0(t)| \le C_1 \exp\left(-\frac{t}{\varepsilon}\right) + C_2 \varepsilon.$$

Both versions of HMM can be expressed in abstract form as

$$y^{n+1} = y^n + \Delta t \tilde{g}(y^n)$$

We now estimate $E^n = y^n - y_0(t^n)$, where $t^n = n\Delta t$, for $n\Delta t \le T$. We have

$$\frac{E^{n+1} - E^n}{\Delta t} = \tilde{g}(y^n) - G(y_0(t^n)) + O(\Delta t)$$
$$= G(y^n) - G(y_0(t^n)) + \tilde{g}(y^n) - G(y^n) + O(\Delta t)$$
$$= G'(y_0(t^n) + \theta E^n)E^n + \tilde{g}(y^n) - G(y^n) + O(\Delta t)$$

If we let $L = \max |G'(y)|$, then we get

$$|E^n| \le e^{Lt}\left(C\Delta t + \max_{n\Delta t \le T} |\tilde{g}(y^n) - G(y^n)|\right).$$

where $C = \max_{t \le T} |y_0''(t)|$. Here the first term on the right hand side is the standard truncation error at the macroscale. The second term $|\tilde{g}(y^n) - G(y^n)|$ is the error in the F-estimator. To estimate this error, we first have to specify the weights $\{K_{m,M}\}$. For the dissipative case, the simplest working choice is

$$K_{m,M} = 0, \ \text{for} \ m = 0, 1, \dots, M - 1.$$
$$K_{M,M} = M$$

In other words

$$\tilde{g}(y^n) = g(x^{n,M}, y^{n,M})$$

Let us first consider the case when we constrain the slow variable during microscale evolution. In this case, since

$$x^{n,m+1} = x^{n,m} - \frac{\delta t}{\varepsilon}(x^{n,m} - f(y^n))$$

if we let $e^{n,m} = x^{n,m} - f(y^n)$, then

$$e^{n,m+1} = \left(1 - \frac{\delta t}{\varepsilon}\right)e^{n,m}$$

Hence

$$e^{n,M} = \left(1 - \frac{\delta t}{\varepsilon}\right)^{M} e^{n,0}$$

Therefore

$$|\tilde{g}(y^n) - G(y^n)| = |g(x^{n,M}, y^n) - g(f(y^n), y^n)|$$

$$\leq C \left|1 - \frac{\delta t}{\varepsilon}\right|^{M} |e^{n,0}|$$

$$\leq C \left|1 - \frac{\delta t}{\varepsilon}\right|^{M}$$

Hence we obtain

$$|E^n| \leq C \left(\Delta t + \left|1 - \frac{\delta t}{\varepsilon}\right|^{M}\right)$$

For the seamless version, we have

$$x^{n,m+1} = x^{n,m} - \frac{\delta t}{\varepsilon}(x^{n,m} - f(y^{n,m}))$$

$$y^{n,m+1} = y^{n,m} + \delta t g(x^{n,m}, y^{n,m})$$

From this, we get

$$e^{n,m+1} = \left(1 - \frac{\delta t}{\varepsilon}\right) e^{n,m} - \frac{\delta t}{\varepsilon}(f(y^n) - f(y^{n,m}))$$

Thus

$$|e^{n,M}| \leq \left|1 - \frac{\delta t}{\varepsilon}\right|^{M} |e^{n,0}| + \frac{\delta t}{\varepsilon} \sum_{m=0}^{M-1} \left|1 - \frac{\delta t}{\varepsilon}\right|^{M-m} |f(y^n) - f(y^{n,m})|$$

Since

$$|y^n - y^{n,m}| \leq C m \delta t$$

$$\sum_{m=0}^{M-1} \left|1 - \frac{\delta t}{\varepsilon}\right|^{M-m} m \leq M \frac{\varepsilon}{\delta t}$$

we have

$$|e^{n,M}| \leq C \left(\left|1 - \frac{\delta t}{\varepsilon}\right|^{M} + M\delta t\right)$$

$$|\tilde{g}(y^n) - G(y^n)| \leq C \left(\left|1 - \frac{\delta t}{\varepsilon}\right|^{M} + M\delta t\right)$$

and

$$|E^n| \leq C \left(\Delta t + \left|1 - \frac{\delta t}{\varepsilon}\right|^{M} + M\delta t\right)$$

In other words, we have

$$|e(\text{HMM})| \leq C\left(\left|1 - \frac{\delta t}{\varepsilon}\right|^M + M\delta t\right)$$

Unlike the previous case for which the slow variables were explicitly constrained for the microscale solver, the modeling error as a function of M first decreases and then slowly increases due to the term $M\delta t$ caused by the slow drift of the slow variable y in the microscale solver. It is easy to see that the optimal choice for M is given by

$$M\delta t = \varepsilon \log \frac{1}{\varepsilon}$$

In this case, one has

$$|E^n| \leq C\left(\Delta t + \varepsilon \log \frac{1}{\varepsilon}\right)$$

In the general case when other schemes are used as the macro and micro solvers, we can follow the same analysis. We then obtain the following result.

THEOREM 3.1. *Assume that the macroscale solver is k-th order accurate and stable, i.e. for linear multi-step methods the root condition is satisified. Let $A(\lambda\delta t)$ be the amplification factor of the microscale solver for the ODE $\dot{y} = -\lambda y$. Then there exists a constant C, independent of $\Delta t, \varepsilon, \delta t$ and M, such that*

$$|E^n| \leq C\left(\Delta t^k + \left|A\left(\frac{\delta t}{\varepsilon}\right)\right|^M + M\delta t\right)$$

for the seamless HMM, and

$$|E^n| \leq C\left(\Delta t^k + \left|A\left(\frac{\delta t}{\varepsilon}\right)\right|^M\right)$$

for HMM when the slow variable is constrained for the microscale solver.

4. Error estimates for the oscillatory case

As before we will discuss a simple yet canonical example, namely (1.2). The dynamics for the slow variable I obeys the equation

$$\dot{I} = \frac{1}{2\pi} \int_0^{2\pi} g(\varphi, I)d\varphi = G(I) \tag{4.1}$$

See for example [1].

We will restrict ourselves to the situation when the weights $\{K_{m,M}\}$ are given by

$$K_{m,M} = K_0\left(1 - \frac{m}{M}\right) \tag{4.2}$$

for $m = 0, 1, \ldots, M$, where K_0 is a kernel that satisfies certain conditions to be specified below. We will also assume that $\omega(I)$ never vanishes:

$$\omega(I) \geq \omega_0 > 0 \tag{4.3}$$

First we consider the case when the slow variable is constrained for the micro-solver. In this case,

$$\tilde{g}(I^n) = \frac{1}{M} \sum_{m=1}^{M} K_0 \left(1 - \frac{m}{M}\right) g(\varphi^{n,m}, I^n) \tag{4.4}$$

Denote by $\tilde{\varphi}(t)$ the solution of the problem

$$\dot{\tilde{\varphi}} = \frac{1}{\varepsilon} \omega(I^n) + f(\tilde{\varphi}, I^n, \varepsilon) \tag{4.5}$$

$$\tilde{\varphi}(t^n) = \varphi^{n,0}$$

on the time interval $[t^n, t^{n+1}]$ and let $\tilde{\varphi}^m = \tilde{\varphi}(t^n + m\delta t)$. Then

$$\tilde{g}(I^n) - G(I^n) = \frac{1}{M} \sum_{m=1}^{M} K_0 \left(1 - \frac{m}{M}\right) (g(\varphi^{n,m}, I^n) - g(\tilde{\varphi}^m, I^n))$$

$$+ \frac{1}{M} \sum_{m=1}^{M} K_0 \left(1 - \frac{m}{M}\right) g(\tilde{\varphi}^m, I^n)$$

$$- \frac{1}{M\delta t} \int_0^{M\delta t} K_0 \left(1 - \frac{\tau}{M\delta t}\right) g(\tilde{\varphi}(t^n + \tau), I^n) d\tau \tag{4.6}$$

$$+ \frac{1}{M\delta t} \int_0^{M\delta t} K_0 \left(1 - \frac{\tau}{M\delta t}\right) g(\tilde{\varphi}(t^n + \tau), I^n) d\tau - G(I^n)$$

$$= I_1 + I_2 + I_3$$

To estimate I_1, let us assume that

$$\frac{1}{M} \sum_{m=1}^{M} |K_{m,M}| \le C_0$$

Since

$$|\tilde{\varphi}^m - \varphi^{n,m}| \le C \frac{\delta t}{\varepsilon}$$

hence

$$|I_1| \le C \frac{\delta t}{\varepsilon}$$

I_2 is the error due to quadrature

$$|I_2| \le C \frac{\delta t}{\varepsilon}$$

where the term $\frac{C}{\varepsilon}$ comes from $\max |\frac{d\tilde{\varphi}}{dt}|$. I_3 is the error due to filtering. We write

$$I_3 = \frac{1}{M\delta t} \int_0^{M\delta t} K_0 \left(1 - \frac{\tau}{M\delta t}\right) \left(g(\tilde{\varphi}(t^n + \tau), I^n) - g\left(\varphi^{n,0} + \frac{\omega(I^n)}{\varepsilon}\tau, I^n\right)\right) d\tau$$

$$+ \frac{1}{M\delta t} \int_0^{M\delta t} K_0 \left(1 - \frac{\tau}{M\delta t}\right) g\left(\varphi^{n,0} + \frac{\omega(I^n)}{\varepsilon}\tau, I^n\right) d\tau - G(I^n)$$

$$= I_{3,1} + I_{3,2}$$

$$|I_{3,1}| \leq C \max_{\tau \leq M\delta t} \left| \tilde{\varphi}(t^n + \tau) - \left(\varphi^{n,0} + \frac{\omega(I^n)}{\varepsilon} \tau \right) \right| \leq C \, M\delta t$$

To estimate $I_{3,2}$, consider the following canonical problem: Estimate

$$I = \frac{1}{T} \int_0^T K_0 \left(1 - \frac{\tau}{T} \right) g_0 \left(\omega_0 \frac{\tau}{\varepsilon} \right) d\tau - \frac{1}{2\pi} \int_0^{2\pi} g_0(\varphi) d\varphi$$

where g_0 is a periodic function with period 2π. Without loss of generality, assume $\int_0^{2\pi} g_0(\varphi) d\varphi = 0$. Let $g_1(s) = \int_0^s g_0(s') ds'$. Then $|g_1(s)| \leq$ Const.

$$\int_0^{\frac{T}{\varepsilon}} g(\omega_0 s) K_0 \left(1 - \frac{\varepsilon s}{T} \right) ds$$

$$= \frac{1}{\omega_0} K_0 \left(1 - \frac{\varepsilon s}{T} \right) g_1(\omega_0 s) \Big|_0^{\frac{T}{\varepsilon}} + \frac{1}{\omega_0} \frac{\varepsilon}{T} \int_0^{\frac{T}{\varepsilon}} K_0' \left(1 - \frac{\varepsilon s}{T} \right) g_1(\omega_0 s) ds$$

Assuming

$$K_0(0) = K_0(1) = 0, \quad \int_0^1 |K_0'(\tau)| d\tau \leq \text{ Const.}$$

then

$$|I| \leq \left(\frac{\varepsilon}{T} \right)^2 \frac{1}{\omega_0} \int_0^{\frac{T}{\varepsilon}} \left| K_0' \left(1 - \frac{\varepsilon s}{T} \right) \right| ds$$

$$\leq \frac{C}{\omega_0} \frac{\varepsilon}{T}$$

To get higher order estimates, let $g_2(s) = \int_0^s g_1(s') ds'$. Then $g_2(s) = C_0 s + \tilde{g}_2(s)$, where $\tilde{g}_2(s)$ is periodic. Thus

$$I = \frac{1}{\omega_0^2} \left(\frac{\varepsilon}{T} \right)^2 K_0' \left(1 - \frac{\varepsilon s}{T} \right) g_2(\omega_0 s) \Big|_0^{\frac{T}{\varepsilon}}$$

$$+ \frac{1}{\omega_0^2} \left(\frac{\varepsilon}{T} \right)^3 \int_0^{\frac{T}{\varepsilon}} K_0'' \left(1 - \frac{\varepsilon s}{T} \right) g_2(\omega_0 s) ds$$

Assuming further that

$$K_0'(0) = K_0'(1) = 0, \quad \int_0^1 |K_0''(s)| ds \leq \text{ Const.}$$

then, we obtain

$$I = \frac{1}{\omega_0^2} \left(\frac{\varepsilon}{T} \right)^3 \int_0^{\frac{T}{\varepsilon}} K_0'' \left(1 - \frac{\varepsilon s}{T} \right) (C_0 s + \tilde{g}_2(s)) ds$$

Notice that

$$\int_0^{\frac{T}{\varepsilon}} K_0'' \left(1 - \frac{\varepsilon s}{T} \right) s \, ds = 0$$

therefore

$$|I| \leq \frac{C}{\omega_0^2} \left(\frac{\varepsilon}{T}\right)^2.$$

Proceeding with this argument inductively, we obtain the following.

LEMMA 4.1. *Assume that K_0 satisfies*

$$K_0^\alpha(0) = K_0^\alpha(1) = 0$$

for $\alpha = 0, 1, \ldots, q-1$, and

$$\sum_{\alpha=0}^q \int_0^1 |K_0^\alpha(s)| ds \leq \ Const.$$

for $\alpha = 0, 1, \ldots, q$, then

$$|I| \leq \frac{C}{\omega_0^q} \left(\frac{\varepsilon}{T}\right)^q$$

Consequently, we obtain

$$|I_{3,2}| \leq \frac{C}{\omega_0^q} \left(\frac{\varepsilon}{M\delta t}\right)^q$$

Summarizing, we obtain

$$|E^n| \leq C \left(\Delta t + \frac{\delta t}{\varepsilon} + M\delta t + \frac{1}{\omega_0^q} \left(\frac{\varepsilon}{M\delta t}\right)^q\right).$$

In this result, the term $M\delta t$ comes from $I_{3,1}$ since we compared $\tilde\varphi(t)$ with $\varphi^{n,0} + \frac{\omega_0}{\varepsilon}t$. This term can be avoided if we estimate I_3 directly.

LEMMA 4.2. *Solutions of*

$$\varphi_\tau = \frac{1}{\varepsilon}\omega(I^n) + f(\varphi, I^n)$$

can be expressed in the form

$$\varphi(\tau) = \left(\frac{1}{\varepsilon}\omega(I^n) + a_\varepsilon\right)\tau + \tilde\varphi(\tau)$$

where a_ε is bounded as $\varepsilon \to 0$, $\tilde\varphi$ is periodic with period $\frac{2\pi}{\frac{1}{\varepsilon}\omega(I^n)+a_\varepsilon} = I_\varepsilon = O(\varepsilon)$.

Consequently $g(\varphi(t^n + \tau), I^n)$ is periodic with period I_ε.

Define

$$g_0 = \frac{1}{I_\varepsilon} \int_0^{I_\varepsilon} g(\varphi(t^n + \tau), I^n) d\tau$$

$$g_1(\tau) = \int_0^\tau (g(\varphi(t^n + \tau_1), I^n) - g_0) d\tau_1$$

$$g_2(\tau) = \int_0^\tau g_1(\tau_1) d\tau_1$$

. Then we have

$$|g_1(\tau)| \le C\varepsilon$$
$$g_2(\tau) = C_1\tau + \tilde{g}_2(\tau)$$

where $\tilde{g}_2$ is periodic and

$$|\tilde{g}_2(\tau)| \le C\varepsilon^2$$

This result follows from a simple rescaling argument using the fact that $I_\varepsilon = O(\varepsilon)$.

Let $T = M\delta t$, then

$$I = \frac{1}{T} \int_0^T K_0 \left(1 - \frac{\tau}{T}\right) g(\varphi(t^n + \tau), I^n) d\tau - g_0$$

$$= \frac{1}{T} K_0 \left(1 - \frac{\tau}{T}\right) g_1(\tau)\Big|_0^T + \frac{1}{T^2} \int_0^T K_0' \left(1 - \frac{\tau}{T}\right) g_1(\tau) d\tau$$

Hence

$$|I| \le \frac{\varepsilon}{T^2} \int_0^T \left| K_0' \left(1 - \frac{\tau}{T}\right) \right| d\tau \le C\frac{\varepsilon}{T}$$

To get better estimates, we can proceed in the same way and write

$$I = \frac{1}{T^2} K_0' \left(1 - \frac{\tau}{T}\right) g_2(\tau)\Big|_0^T + \frac{1}{T^3} \int_0^T K_0'' \left(1 - \frac{\tau}{T}\right) g_2(\tau) d\tau$$

$$= \frac{1}{T^3} \int_0^T K_0'' \left(1 - \frac{\tau}{T}\right) \tilde{g}_2(\tau) d\tau$$

Hence

$$|I| \le C\frac{\varepsilon^2}{T^2}$$

if $\int_0^1 |K'(\tau)| d\tau$ is finite. In this way, we obtain:

LEMMA 4.3. *Assume that K_0 satisfies condition K_q, then*

$$|I| \le C_q \left(\frac{\varepsilon}{T}\right)^q$$

When high order microscale solvers are used and/or high order quadrature rules are used in defining the filter, we have instead

$$|I_1| \le C \left(\frac{\delta t}{\varepsilon}\right)^l$$

$$|I_2| \le C \left(\frac{\delta t}{\varepsilon}\right)^p$$

where l is the order of the micro-solver and p is the order of the quadrature rule.

In summary, we obtain

THEOREM 4.4. *Under the same assumption as before, we have*

$$|E^n| \le C \left((\Delta t)^k + \left(\frac{\delta t}{\varepsilon} \right)^r + \left(\frac{\varepsilon}{M \delta t} \right)^q \right)$$

where $r = min(l, p)$.

As we expected, the size of δt should be chosen to resolve the small time scale ε, and M should be large enough so that $M \delta t$ is larger than ε in order to sample adequately the dynamics at the small scale.

The seamless version can be analyzed in the same way.

5. A general setting

Consider

$$\dot{Z} = F(Z, \varepsilon) \tag{5.1}$$

such that there exists differmorphisms h:

$$\begin{pmatrix} x \\ y \end{pmatrix} = h(Z) \tag{5.2}$$

where y is the slow variable, x is the fast variable in the sense that the ODE (5.1) transforms to

$$\dot{x} = \frac{1}{\varepsilon} f(x, y, \varepsilon)$$
$$\dot{y} = g(x, y, \varepsilon)$$

where f and g are bounded as $\varepsilon \to 0$. Our main assumption is that for each fixed y, the dynamics of

$$\dot{x} = \frac{1}{\varepsilon} f(x, y, \varepsilon)$$

has a unique invariant measure $\mu_y^\varepsilon(dx)$. The effective equation for y is then

$$\dot{y} = G(y) = \int g(x, y, 0) \mu_y^0(dx)$$

Example 1:

$$f(x, y, \varepsilon) = -x + f(y) \tag{5.3}$$

then

$$\mu_y^0(dx) = \delta(x - f(y))$$

Example 2:

$$\dot{\varphi} = \frac{1}{\varepsilon} \omega(I) + f(\varphi, I) \tag{5.4}$$

φ is on the circle. Then as $\varepsilon \to 0$,

$$\mu_I^\varepsilon(d\varphi) \to \text{Lebesgue measure on the circle.}$$

We will assume that $\mu_y^\varepsilon(dx)$ behaves continuously as ε and y changes

For the micro solver, we will assume that it also has an invariant measure $\mu_y^{\delta t, \varepsilon}(dx)$ if the slow variable y is constrained to be fixed.

The two versions of HMM can be formulated in the same way as before. If we know explicitly what the slow variables are, we can use that information in formulating HMM. Otherwise, we can use the seamless version.

Assume that the macroscale solver is stable and k-th order accurate. Then

$$|E^n| \leq C(\Delta t^k + e(\delta t, \varepsilon, M))$$

where $e(\delta t, \varepsilon, M)$ is the error in the F-estimator.

$$e(\delta t, \varepsilon, M) = \max_{n\Delta t \leq T} |\tilde{g}(y^n) - G(y^n)|$$

$$\begin{aligned}
\tilde{g}(y^n) - G(y^n) =& \frac{1}{M} \sum_{m=0}^{M-1} K_{m,M} g(x^{n,m}, y^n, \varepsilon) - \int g(x, y^n, 0) \mu_{y^n}^{\delta t, \varepsilon}(dx) \\
&+ \int g(x, y^n, \varepsilon)(\mu_{y^n}^{\delta t, \varepsilon}(dx) - \mu_{y^n}^0(dx)) \\
&+ \int (g(x, y^n, \varepsilon) - g(x, y^n, 0)) \mu_{y^n}^0(dx) \\
=& I_1 + I_2 + I_3
\end{aligned}$$

I_1 is controlled by how fast $\{x^{n,m}\}$ converges to the invariant measure $\mu_{y^n}^{\delta t, \varepsilon}(dx)$ and also the filtering. I_2 is controlled by the difference between $\mu_{y^n}^{\delta t, \varepsilon}(dx)$ and $\mu_{y^n}^0(dx)$.

At the present time, there is no general theory on estimating I_1 and I_2. From a numerical point of view, the microscale solver should be chosen to minimize I_1 and I_2, namely that its solutions should converge rapidly to the local equilibrium.

6. Summary

In this paper we analyzed the heterogeneous multiscale method for ordinary differential equations with multiple time scales. The analysis is an application of the general principle discussed in [3], which states that HMM is stable if the macroscale solver is stable and the total error is a sum of the standard truncation error of the macroscale solver and the error in the F-estimator. Our analysis gives a fairly detailed understanding of how the error in the F-estimator depends on various components of HMM, such as the microscale solver and the filter, as well as the nature of the microscale problem.

There is still much room for improvement of the results. In our analysis we compared the HMM solution with the slow variable of the original problem in the limit as $\varepsilon \to 0$. It is of interest in future work to compare directly with the slow variable at finite but small ε. Another point to be noted is that we did take into account the moment condition for the averaging kernel. To be specific, let us assume that the microscale data has the form

$$f^\varepsilon(t) = f_0(t) + f_1\left(\frac{t}{\varepsilon}\right)$$

where $f_1(\tau)$ is periodic in τ with period 1. The slow component of $\{f^\varepsilon(t)\}$ is

$$F(t) = f_0(t) + \int_0^1 f_1(\tau)d\tau$$

The macroscale data of interest is $F_0 = F(0)$. Assume that Our F-estimator has the form

$$\tilde{F}_0 = \frac{1}{T}\int_0^T K_0\left(1 - \frac{\tau}{T}\right)f^\varepsilon(\tau)d\tau$$

Then one has

LEMMA 6.1. *Assume that K_0 satisfies the following conditions*
    1. (A1). Moment condition

$$\int_0^1 K_0(\tau)d\tau = 1$$

$$\int_0^1 (1-\tau)^\alpha K_0(\tau)d\tau = 0$$

      for $\alpha = 1, 2, \ldots \ell - 1$
    2. (A2). Regularity condition

$$K_0^{(\alpha)}(0) = K_0^{(\alpha)}(1) = 0$$

    for $\alpha = 0, 1, \ldots, q-1$ and

$$\int_0^1 |K_0^{(\alpha)}(\tau)|\alpha\tau \ < +\infty$$

    for $\alpha = 0, 1, \ldots, q$. Then

$$|F_0 - \tilde{F}_0| \le C\left(T^\ell + \left(\frac{\varepsilon}{T}\right)^q\right)$$

This lemma can be proved by combining the arguments in Section 6 of [3] and Section 4 above. Using the moment condition, we should be able to improve the term $M\delta t$ in $e(\text{HMM})$ to $(M\delta t)^{l+1}$ for the seamless version of HMM. But this remains to be done.

REFERENCES

[1] V. Arnold, *Mathematical Methods in Classical Mechanics,* Springer-Verlag, Berlin, New York, 1978.
[2] A. Abdulle, *Fourth order Chebychev methods with recurrence relations*, SIAM J. Sci. Comput., 23(6):2041–2054, 2002.
[3] W. E and B. Engquist, *The heterogeneous multiscale methods*, Comm. Math. Sci., 1(1):88–134, 2003.

[4] B. Engquist and R. Tsai, *Heterogeneous multiscale methods for stiff ordinary differential equations*, preprint.

[5] C.W. Gear, *Numerical Initial Value Problems in Ordinary Differential Equations*, Prentice Hall, Eaglewood Cliffs, NJ, 1971.

[6] C.W. Gear and I.G. Kevrekidis, *Projective methods for stiff differential equations: problems with gaps in their eigenvalue spectrum,* submitted to SIAM J. Sci. Comp.

[7] E. Hairer, S.P. Norsett, and G. Warner, *Solving Ordinary Differential Equations*, Springer-Verlag, Berlin, New York, 1987.

[8] V.I. Lebedev and S.I. Finogenov, *Explicit methods of second order for the solution of stiff systems of ordinary differential equations*, Zh. Vychisl. Mat. Mat Fiziki., 16(4):895–910, 1976.

[9] L.R. Petzold, L.O. Jay, and J. Yen, *Numerical solution of highly oscillatory ordinary differential equations*, Acta Numerica, A. Iserles ed., Cambridge Univ. Press, Cambridge, 437–484, 1997.

[10] E.V. Eijnden, *Numerical techniques for multiscale dynamical system with stochastic effects*, Comm. Math. Sci., 1(2):385–391, 2003.