Multiscale Modeling and Computation

Weinan E and Bjorn Engquist

Multiscale modeling and computation is a rapidly evolving area of research that will have a fundamental impact on computational science and applied mathematics and will influence the way we view the relation between mathematics and science. Even though multiscale problems have long been studied in mathematics, the current excitement is driven mainly by the use of mathematical models in the applied sciences: in particular, material science, chemistry, fluid dynamics, and biology. Problems in these areas are often multiphysics in nature; namely, the processes at different scales are governed by physical laws of different character: for example, quantum mechanics at one scale and classical mechanics at another.

Emerging from this intense activity is a need for new mathematics and new ways of interacting with mathematics. Fields such as mathematical physics and stochastic processes, which have so far remained in the background as far as modeling and computation is concerned, will move to the frontier. New questions will arise, new priorities will be set as a result of the rapid evolution in the computational fields.

There are several reasons for the timing of the current interest. Modeling at the level of a single scale, such as molecular dynamics or continuum theory, is becoming relatively mature. Our computational capability has reached the stage when serious multiscale problems can be contemplated, and there is an urgent need from science and technology—nano-science being a good example—for multiscale modeling techniques.

It is not an exaggeration to say that almost all problems have multiple scales. We organize our time according to days, months, and years, reflecting the multiple time scales in the dynamics of the solar system. Another example with multiple time scales is that of protein folding. While the time scale for the vibration of the covalent bonds is on the order of femtoseconds (10^{-15} s), folding time for the proteins may very well be on the order of seconds. Well-known examples of problems with multiple length scales include turbulent flows, mass distribution in the universe, and vortical structures on the weather map [1]. In addition, different physical laws may be required to describe the system at different scales. Take the example of fluids. At the macroscale (meters or millimeters), fluids are accurately described by the density, velocity, and temperature fields, which obey the continuum Navier-Stokes equations. On the scale of the mean free path, it is necessary to use kinetic theory (Boltzmann's equation) to get a more detailed description in terms of the one-particle phase-space distribution function. At the nanometer scale, molecular dynamics in the form of Newton's law has to be used to give the actual position and velocity of each individual atom that makes up the fluid. If a liquid such as water is used as the solvent for protein folding, then the electronic structures of the water

Weinan E is professor of mathematics at Princeton University. His email address is weinan@math.princeton.edu.

Bjorn Engquist is professor of mathematics at Princeton University. His email address is engquist@math.princeton.edu.

molecules become important, and these are described by Schrödinger's equation in quantum mechanics. The boundaries between different levels of theories may vary, depending on the system being studied, but the overall trend described above is generally valid. At each finer scale, a more detailed theory has to be used, giving rise to more detailed information on the system.

There is a long history in mathematics for the study of multiscale problems. Fourier analysis has long been used as a way of representing functions according to their components at different scales. More recently, this multiscale, multiresolution representation has been made much more efficient through wavelets. On the computational side, several important classes of numerical methods have been developed which address explicitly the multiscale nature of the solutions. These include multigrid methods, domain decomposition methods, fast multipole methods, adaptive mesh refinement techniques, and multiresolution methods using wavelets.

From a modern perspective, the computational techniques described above are aimed at efficient representation or solution of the fine-scale problem. For many practical problems, full representation or solution of the fine-scale problem is simply impossible for the foreseeable future because of the overwhelming costs. Therefore we must seek alternative approaches that are more efficient. One classical approach is to use analytic techniques to derive effective models at the scale of interest. An early example of such a technique is the averaging method. Consider, for example, a system of ordinary differential equations written in the action-angle variables

$$\begin{aligned} \varphi_t &= \frac{1}{\varepsilon}\omega(I) + f(\varphi, I) \\ I_t &= g(\varphi, I), \end{aligned} \tag{1}$$

where φ is the fast variable, which varies on the time scale of $O(\varepsilon)$, $\varepsilon \ll 1$; I is the slow variable, which mainly varies on the time scale of $O(1)$; and f and g are assumed to be 2π-periodic in φ. The averaging method gives the leading-order behavior ($\bar{I}$) of the slow variable I, which is often the quantity of interest, by an averaged equation [3] (see Figure 2).

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

Another example of mathematical techniques for multiscale problems is the homogenization method. Consider the problem

$$\frac{\partial u^\varepsilon}{\partial t} = \nabla \cdot \left(a\left(x, \frac{x}{\varepsilon}\right) \nabla u^\varepsilon(x,t) \right), \quad x \in \Omega, \tag{3}$$

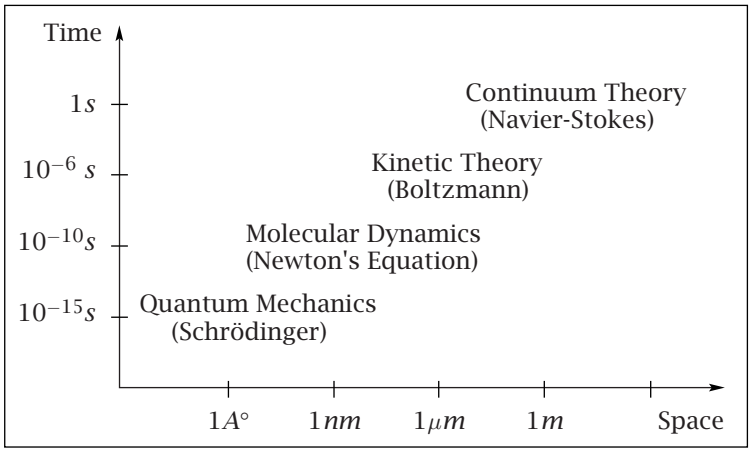

Figure 1. Different laws of physics are required to describe properties and processes of fluids at different scales.

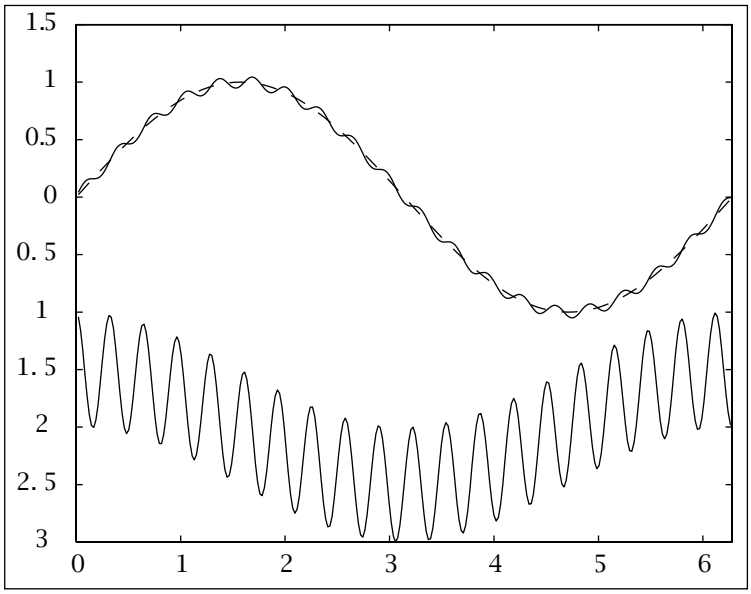

Figure 2. Illustration of the averaging method. The lower curve is φ as a function of t, the upper solid curve is I, and the dashed line is $\bar{I}$.

with the boundary condition $u^\varepsilon|_{\partial\Omega} = 0$. In this problem the multiscale nature comes from the coefficients $a\left(x, \frac{x}{\varepsilon}\right)$, which contain two scales: a scale of $O(\varepsilon)$ and a scale of $O(1)$. Not only is (3) a nice model problem for the homogenization technique, it also describes important physical processes such as heat conduction in a composite material. For simplicity let us assume that $a(x,y)$ is periodic in y. Then it can be shown [4] that for $\varepsilon \ll 1$, $u^\varepsilon(x,t)$ can be expressed in the form

$$\begin{aligned} u^\varepsilon(x,t) = U(x,t) &+ \varepsilon u_1\left(x, \frac{x}{\varepsilon}, t\right) \\ &+ \varepsilon^2 u_2\left(x, \frac{x}{\varepsilon}, t\right) + \cdots, \end{aligned} \tag{4}$$

where U satisfies a homogenized equation

$$\frac{\partial U}{\partial t} = \nabla \cdot (A(x)\nabla U(x,t)) \tag{5}$$

in Ω and $U|_{\partial\Omega} = 0$. Here $A(x)$ may be thought of as being the effective coefficient describing the effective properties of the system on the scale of $O(1)$. Determining $A(x)$ usually requires solving families of so-called cell problems. In the one-dimensional case, however, $A(x)$ is simply given by the harmonic average

$$A(x) = \left(\int_0^1 \frac{1}{a(x,y)}\, dy \right)^{-1}. \tag{6}$$

Many other mathematical techniques have been developed to study multiscale problems, including boundary-layer analysis [12], semiclassical methods [15], geometric theory of diffractions [11], stochastic mode elimination [14], and renormalization group methods [8], [19].

Despite this progress, purely analytical techniques are still very limited when it comes to problems of practical interest. As a result, the overwhelming majority of problems have been approached using empirical techniques to model the small scales in terms of the macroscale variables by empirically derived formulas. In fact, a large part of the progress in physical sciences lies in such empirical modeling. A familiar example is that of the continuum theory of fluid dynamics. To derive the system of equations for fluids, we apply Newton's law to an arbitrary volume of fluid denoted by Ω:

$$\frac{D}{Dt} \int_\Omega \rho u\, dV = F(\Omega), \tag{7}$$

where $\frac{D}{Dt}$ is the material derivative, ρ and u are the density and velocity fields respectively, and $F(\Omega)$ is the total force acting on the volume of fluid in Ω. The forces consist of body forces such as gravity, which we neglect for the present argument, due to the long-range interaction of the molecules that make up the fluid, and forces due to the short-range interaction between the molecules, such as the Van der Waals interaction. In the continuum theory, the short-range forces are represented as a surface integral of the stress tensor τ, which is a macroscopic idealization of the small scale effects,

$$F(\Omega) = \int_{\partial\Omega} (\tau \cdot \hat{n})\, ds, \tag{8}$$

where $\hat{n}$ is the unit outward normal of Ω. The stress τ can be expressed as $\tau = -pI + \tau_d$, where p is the pressure, I is the identity tensor, and τ_d is the dissipative part of the stress. In order to close the system, we need to express τ_d in terms of u. In the simplest empirical approximation, τ_d is assumed to be a linear function of ∇u. This leads to

$$\tau_d = \mu \frac{\nabla u + (\nabla u)^T}{2}, \tag{9}$$

where μ is called the viscosity of the fluid. Substituting this into Newton's law and adding the incompressibility condition gives rise to the well-known Navier-Stokes equation:

$$\rho(u_t + (u \cdot \nabla)u) + \nabla p = \mu \triangle u, \quad \nabla \cdot u = 0. \tag{10}$$

In such a macroscopic description, all molecular details of the liquid are lumped into a single parameter, the viscosity. Fluids for which (9) gives an accurate description of the small-scale effects are called Newtonian fluids.

This simple derivation illustrates how, in general, continuum models in the form of partial differential equations are derived. One typically starts with some universal conservation laws such as (7). This requires introducing certain currents or flux densities, which are then expressed by some postulated constitutive relations such as (9). In this way, we obtain the heat equation for thermal conduction by postulating Fourier's law, the diffusion equation for mass transport using Fick's law, and the porous-medium equation using Darcy's law.

Such empirical ad hoc descriptions of the small scales are used almost everywhere in science and engineering. Consider, for example, the hierarchy of models depicted in Figure 1. In molecular dynamics, empirical potentials are used to model the forces between atoms, mediated by the electrons. In kinetic theory, empirical collision kernels are used to describe probabilistically the short-range interaction between the atoms and the molecules. Other examples include plasticity, crack propagation, and chemical reactions. While much progress has been made using such empirical approaches, their shortcomings have also been recognized, especially so in recent years, since numerical simulations based on the empirical models are now accurate enough that the modeling error can be clearly identified. Microscale simulation methods such as electronic structure calculations have matured, enabling us to ask more ambitious questions. Moreover, the empirical approach often lacks information about how microstructural changes, such as the conformation of polymers in a polymeric fluid, affect the macroscale properties of the system.

Examples of a New Class of Multiscale Methods

In view of the limitations of the empirical approach, several "first principle"-based multiscale methods have been proposed in recent years. Some of these methods are discussed below.

First Principle Molecular Dynamics

Molecular dynamics describes the behavior of a collection of atoms by their positions and momenta,

denoted by $\{x_j, p_j\}_{j=1}^N$. The dynamics follows Newton's second law:

$$m_j \ddot{x}_j = -\frac{\partial V_0}{\partial x_j},$$

where m_j is the mass of the j-th atom and $V_0(x_1, \ldots, x_N)$ is the potential energy of the system, which is due mainly to the Coulomb interaction between the charges, determined by the positions of the nuclei and the state of the electrons. In this example the macroscale process is the molecular dynamics of the nuclei. The microscale process is the state of the electrons, which determines V_0. Since the electrons are so much lighter than the nuclei, one can assume to a good approximation that they are at the ground state determined by the positions of the nuclei. This is the so-called Born-Oppenheimer approximation. The potential energy surface determined in this way is called the Born-Oppenheimer potential energy surface.

However, determining explicitly the Born-Oppenheimer potential energy surface is a rather daunting task. As a result, most molecular dynamics simulations use an empirical potential such as the Lennard-Jones potential.

In 1985 Car and Parrinello [7] developed a multiscale procedure for probing the Born-Oppenheimer potential energy surface "on the fly" during molecular dynamics simulations. This new method bridges the different temporal and spatial scales in the system, bypassing the need for empirical potentials. It has found wide application in chemistry, material sciences, and biology.

The Quasicontinuum Method

In the continuum theory of nonlinear elasticity, we are often interested in finding the displacement field by solving a variational problem

$$\min_u E(u) = \int_\Omega f(\nabla u)\, dx,$$

where E is the total elastic energy, u is the displacement field, and f is the stored energy functional, subject to certain loading or boundary conditions. This approach takes for granted that the function f is explicitly given. In reality the process of finding f is rather empirical and often even crude.

A different methodology called the quasicontinuum method was proposed in [6], [16] for the analysis of crystalline materials. In this case the microscopic model comprises molecular mechanics of the atoms that make up the crystal. Given a macroscopic triangulation of the material, let V_H be the standard continuous piecewise-linear finite-element space over this triangulation. For $U \in V_H$, ∇U is constant on each element K. Let $E_K(U)$ be the energy of a unit cell in an infinite volume uniformly deformed according to the constant deformation gradient $\nabla U|_K$. In the quasicontinuum approximation, the total energy associated with the trial function U is then given by

$$\tilde{E}(U) = \sum_K n_K E_K(U),$$

where n_K is the number of unit cells in the element K. This approach bypasses the necessity of modeling f empirically. Instead, the effective f is computed on the fly using microscopic models.

What we have described is the simplest version of the quasicontinuum method. There are many improvements, in particular to deal with defects in the crystal [16].

Kinetic-Hydrodynamic Models of Complex Fluids

Consider, for example, polymers in a solvent. The basic equations follow again from that of mass and momentum conservation:

$$\rho(u_t + (u \cdot \nabla)u) + \nabla p = \mu_s \Delta u + \nabla \cdot \tau_p$$
$$\nabla \cdot u = 0.$$

Here we have decomposed the total stress into two parts: one part, τ_p, due to the polymer and the other part due to the solvent, for which we used Newtonian approximation; μ_s is the solvent viscosity. Traditionally, τ_p is modeled empirically using constitutive relations. The most common models are a generalized Newtonian model and various viscoelastic models. It is generally acknowledged that it is an extremely difficult task to construct such empirical models in order to describe the flow under all experimental conditions.

An alternative approach was proposed in the classical work of Kramers, Kuhn, Rouse et al. [5]. Instead of using empirical constitutive relations, this approach uses a simplified kinetic description for the conformation of the polymers. In the simplest situation, the polymers are assumed to be dumbbells, each of which consists of two beads connected by a spring. Its conformation is therefore described by that of the spring. The dumbbells are convected and stretched by the fluid, and at the same time they experience spring and Brownian forces:

$$\gamma\,(Q_t + (u \cdot \nabla)Q - (\nabla u)^T Q)$$
$$= F(Q) + \sqrt{k_B T \gamma}\, \dot{W}(t).$$

Here Q denotes the conformation of the dumbbell, $F(Q)$ is the spring force, γ is the friction coefficient, $\dot{W}(t)$ is temporal white noise, k_B is the Boltzmann constant, T is the temperature, and I is the identity tensor. If we have Q, we can compute the polymer stress τ_p via

$$\tau_p = n k_B T I + \mathbb{E}(F(Q) \otimes Q),$$

where n is the polymer density and $\mathbb{E}$ denotes expectation over the Brownian forces. These equations are valid in the dilute regime when direct interaction between polymers can be neglected.

Traditional Techniques	Recent Techniques
Multigrid Method	Car-Parrinello Method
Domain Decomposition	Quasi-Continuum Method
Multiresolution Methods	Superparametrization
Adaptive Mesh Refinement	Heterogeneous Multiscale Method
Fast Multipole Method	Vanden-Eijnden's Method
Conjugate Gradient Method	Coarse-Grained Monte Carlo Models
	Adaptive Model Refinement
	Patch Dynamics

Table 1. Traditional and modern computational multiscale techniques. Traditional multiscale techniques focus on resolving the fine-scale problem. Modern multiscale techniques try to reduce the computational complexity by using special features in the fine-scale problem, such as scale separation.

The dumbbell model is a very simplified one. In many cases, it has to be improved. This can be done in a number of ways (see [5]).

Many other multiscale methods that are similar to those mentioned above have been developed in the past few years. We mention in particular the work of Abraham et al. on coupling finite element continuum analysis with molecular dynamics and tight binding [2], the work on coupling kinetic equations with hydrodynamic equations, Vanden-Eijnden's method for solving stochastic ordinary differential equations with multiple time scales [18], superparametrization techniques in meteorology in which the parameters for turbulent transport are determined dynamically by local microscale simulations, and the work of Kevrekidis et al. on bifurcation analysis based on microscopic models [17]. *By explicitly taking advantage of the separation of scales, these methods become much more efficient than solving the full fine-scale problem. This is a common feature of the new class of multiscale methods we are interested in.* In contrast, traditional multiscale techniques such as the original multigrid methods are rather blind to the special features of the problem, since they are aimed at solving the full fine-scale problem everywhere in the macroscale domain. Of course many practical problems such as turbulent flows do not have separation of scales. For these problems, other special features, such as self-similarity in scales, must be identified first before we have a way of modeling them more efficiently than simply solving the fine-scale problem by brute force or resorting to ad hoc models.

Before continuing, let us mention that there are a number of semianalytical and seminumerical multiscale techniques. A good example is the coarse-grained Monte Carlo models of Katsoulakis, Majda, and Vlachos [10], in which self-consistency is guaranteed by enforcing detailed balance in the coarse-grained model as well as the fact that the microscopic model and the coarse-grained model share the same mesoscopic limit.

A General Framework for Multiscale Methods

Given such a variety of multiscale methods in many different applications, it is natural to ask whether a general framework can be constructed. The general framework should ideally

- unify existing methods,
- give guidelines on how to design new methods and improve existing ones,
- provide a mathematical theory for stability and accuracy of these methods.

One proposal for such a general framework was made in [9], and it goes under the name of heterogeneous multiscale method (abbreviated HMM). This name was suggested to emphasize the multi-physics nature of the problems that it intends to handle. In contrast, the multigrid method would be called homogeneous, since it uses the same physical and mathematical model at different scales.

The setup of the problem is the following. We are interested in the macroscopic state of a system, with state variables denoted by U. However, we have at our disposal only a microscopic model for the microscale state variable u. The variables U and u are linked together by a compression operator Q such that $Qu = U$.

Before discussing HMM, it is useful to recall the classical Godunov scheme for gas dynamics [13] as a model example for methodology.

Consider, for example, a scalar conservation law of the form

$$u_t + f(u)_x = 0. \tag{11}$$

Fix a numerical step size $\Delta x, \Delta t > 0$, and define the cell averages

$$(Qu)(x,t) = U(x,t) = \frac{1}{\Delta x}\int_{x-\frac{\Delta x}{2}}^{x+\frac{\Delta x}{2}} u(y,t)\,dy. \tag{12}$$

Then U satisfies the equation

$$\frac{dU}{dt} + \frac{1}{\Delta x}\left(f(u(x+\tfrac{\Delta x}{2},t)) - f(u(x-\tfrac{\Delta x}{2},t))\right) = 0. \tag{13}$$

Let $x_j = j\Delta x$ and $t^n = n\Delta t$. Denote by U_j^n the numerical approximations of $U(x,t)$ at (x_j, t^n). Given $\{U_j^n\}$ for all j, the Godunov scheme computes $\{U_j^{n+1}\}$ via the following steps:

1. Reconstruction: Let $u_{\Delta x}(x,t^n) = U_j^n$ if $x_j - \frac{\Delta x}{2} \le x \le x_j + \frac{\Delta x}{2}$.
2. Riemann Solver: Solve (11) exactly with the initial condition $u(x,t^n) = u_{\Delta x}(x,t^n)$ to time t^{n+1}. Call that solution $\tilde{u}$.
3. Evaluate the total flux at the cell boundaries (at $x_{j+\frac{1}{2}} = \left(j+\frac{1}{2}\right)\Delta x$) from $t = t^n$ to $t = t^{n+1}$:

$$\Delta t\, f_{j+\frac{1}{2}} = \int_{t^n}^{t^{n+1}} f(\tilde{u}(x_{j+\frac{1}{2}}, s))\,ds.$$

4. Compute $\{U_j^{n+1}\}$ by

$$U_j^{n+1} = U_j^n - \frac{\Delta t}{\Delta x}(f_{j+\frac{1}{2}} - f_{j-\frac{1}{2}}).$$

Here the microscale and macroscale models are respectively (11) and (13). The compression operator is the cell averaging (12). The Godunov scheme is a finite-volume method for U in which the fluxes $f_{j+\frac{1}{2}}$ are computed by solving Riemann problems for the original microscopic model (11).

HMM can be viewed as a substantial generalization of the classical Godunov scheme in which the finite-volume method is replaced by general macroscopic schemes, the Riemann solver is replaced by the solution of constrained microscale models, and the cell averaging is typically replaced by more sophisticated data processing techniques.

There are two main components in HMM. The first is an overall macroscopic scheme for U. The second is to estimate the missing macroscopic data needed for the implementation of the macroscopic scheme by solving locally the microscopic model, subject to appropriate constraints.

For dynamical problems, the estimation of the macroscale data can be done using the following generalization of the Godunov procedure.

1. Reconstruction. From U^n, construct u^n such that $Qu^n = U^n$. This reconstruction is clearly not unique and is problem dependent.
2. Microscale evolution. Evolve the microscale model with initial data $u(x, t^n) = u^n(x)$, subject to appropriate constraints such as boundary conditions. Scale separation should be taken into account to minimize the spatial/temporal size of the computational domain for the microscale solver.
3. Data processing. Process the data obtained in step 2 to extract the needed macroscale data. This step often involves ensemble as well as spatial/temporal averaging.

While some general guildelines exist for steps 1 and 3, step 2 is quite specific to the problem.

One special case of HMM is the gas kinetic scheme [20]. There the microscopic state is the one-particle phase-space distribution function. The microscopic model is the kinetic equation. The macroscopic state is the hydrodynamic field variables, and the macroscale solver is also the finite-volume method. The data that need to be estimated from the microscale model are the fluxes $\{f^n_{j+\frac{1}{2}}\}$. The estimation of this data proceeds by a generalization of the Godunov procedure: namely, at $t = t^n$ the one-particle phase-space distribution function is constructed near each cell boundary using the local macroscale data; the kinetic equation is then solved and the hydrodynamic fluxes evaluated using the local solutions of the kinetic equation.

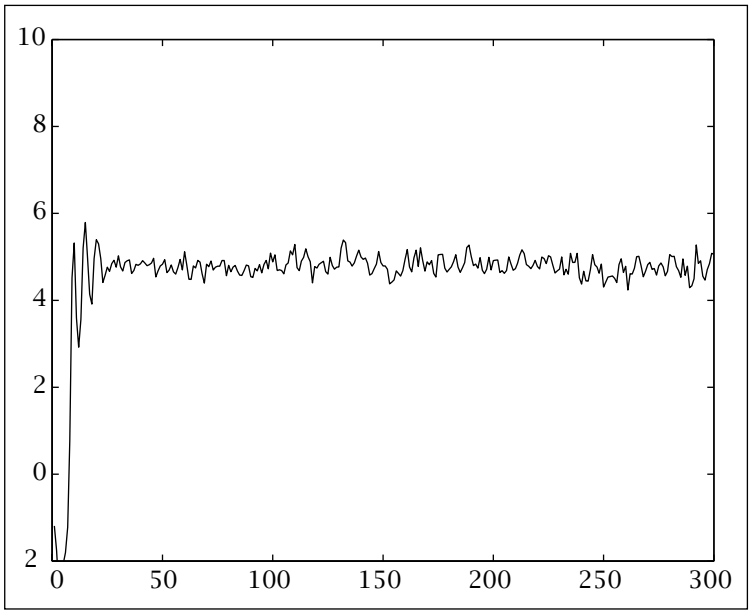

Figure 3. Typical behavior of microscale fluxes as a function of time obtained from molecular dynamics simulation.

One can also replace the microscopic model, here the kinetic equation, by other models such as molecular dynamics. This has been done recently at Princeton by Xiantao Li and Weiqing Ren. The algorithm proceeds in the same way. The macroscale variables and solvers are the same. The data that need to be estimated are again the fluxes at the cell boundaries. Instead of reconstructing the single-particle distribution function f, one reconstructs the positions and momenta of a collection of particles near each cell boundary, consistent with the local hydrodynamic variables. One then evolves molecular dynamics with suitable boundary conditions, at the same time collecting the data of microscopic fluxes. This data is then processed to obtain the macroscopic fluxes necessary for the macroscale solver. The algorithmic details are rather substantial, in particular in the treatment of the boundary condition and the processing of the data from molecular dynamics. A typical example of the data collected from molecular dynamics is illustrated in Figure 3.

It is also straightforward to apply HMM to the parabolic problem (3). One can start by writing (3) in a conservation form

$$u^\varepsilon_t + \nabla \cdot J^\varepsilon = 0, \tag{14}$$

where $J^\varepsilon = -a(x, \frac{x}{\varepsilon})\nabla u^\varepsilon$. Again, for the macroscale solver we will choose the finite-volume method. The macroscale variables are cell averages of u^ε. The macroscale fluxes at the cell boundaries are the data that need to be estimated, and they are computed using the three-step generalized Godunov procedure as follows. From the cell averages, one can make a piecewise-linear reconstruction at the cell boundaries. Equation (3) is then solved with a

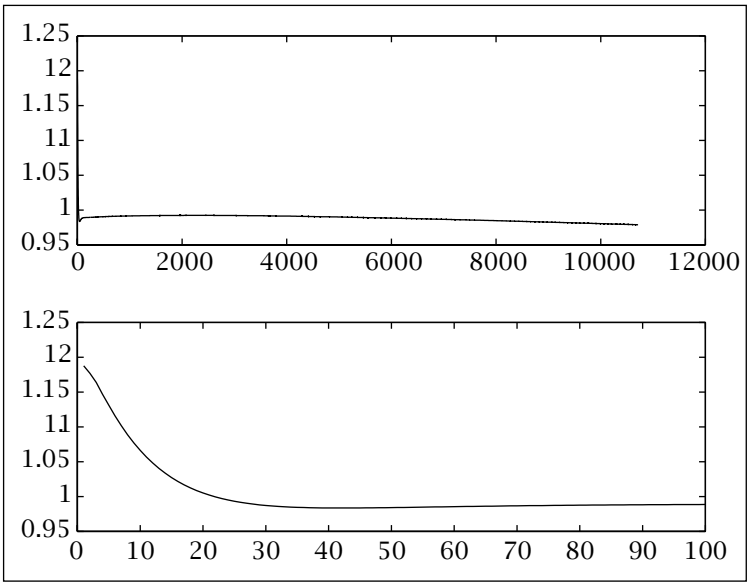

Figure 4. Computed flux $J^\varepsilon(x,t) = a\left(x, \frac{x}{\varepsilon}\right) \nabla u^\varepsilon(x,t)$ as a function of the micro time step over one typical macro time step, for the parabolic homogenization $a\left(x, \frac{x}{\varepsilon}\right) = 2 + \sin 2\pi \frac{x}{\varepsilon}$. The bottom figure is a detailed view of the top figure for small time steps. Notice that $J^\varepsilon(x,t)$ quickly settles down (after about 35 micro time steps) to a quasi-stationary value after a rapid transient.

suitable boundary condition. A typical plot of the microscopic flux J^ε as a function of the microscale time step is shown in Figure 4. Clearly J^ε saturates after a short relaxation time to some quasi-stationary value, and this stationary value is what we use as the macroscale flux.

The two examples, Figures 3 and 4, show the difference between typical data obtained from conservative and from dissipative microscopic processes. Obviously the different data should be processed by different techniques. This is an example of the new techniques that need to be developed and analyzed. It is discussed in detail in the papers that can be found at `http://www.math.princeton.edu/multiscale`.

So far we have discussed cases when the macroscale solver is the finite-volume method. One can also use other methods, such as the finite-element method, as the macroscale solver. In that case the data that need to be estimated are the stiffness and mass matrices. This can be for both elliptic and dynamic problems by solving appropriately formulated microscale problems, as discussed in [9].

Besides the examples discussed above, HMM has also been applied to other homogenization problems, ordinary differential equations with multiple time scales, coupling molecular dynamics with hydrodynamics, coupling molecular dynamics with continuum elasticity, porous-medium flows, and interface dynamics. Along with these applications, a set of tools has been developed for multiscale modeling and computation.

Among the examples of existing multiscale methods we discussed earlier, some versions of the quasicontinuum method, Vanden-Eijnden's method for stochastic differential equations [18], and the multiscale bifurcation analysis method [17] can all be formulated as examples of HMM.

An important problem is the stability and accuracy of these multiscale methods. Since they involve more than one mathematical model, the numerical analysis of these methods is quite nonstandard. However, a general principle has been established in [9] for the numerical analysis of heterogeneous multiscale methods for what is referred to in [9] as the type B problems, which include all the problems we've discussed so far. These are problems for which closed macroscale models exist for suitably chosen sets of macroscale variables, but the models are not explicitly available and are very inefficient to use directly in numerical computations. Nevertheless, we can use these models in our analysis. This can be done without knowing the explicit form of the effective equations. For that purpose we first define an effective macroscale method (EMM) corresponding to an HMM that involves only the macroscale model. To define the EMM, one starts with the same macroscale solver as in HMM, except that one replaces the microscale model by the macroscale model (with initial data equivalent to the reconstruction operator) in the data estimation step. Examples of EMM are given in [9].

If this EMM is stable, then we have the following error estimate:

$$\|U_{\text{HMM}} - U_0\| \le C(H^k + e(\text{HMM})),$$

where U_{HMM} is the HMM solution, U_0 is the exact solution of the macroscopic model, k is the order of accuracy of the EMM, H is the step size of the macroscale numerical grid, and $e(\text{HMM})$ is the error in the estimation of data in the three-step procedure discussed earlier. The norm should be chosen according to the specific problem at hand. The first term on the right-hand side is the standard truncation error of EMM. The second term is the new source of error in HMM due to data estimation. The error $e(\text{HMM})$ typically depends on the rate of relaxation of the microscopic model to local equilibrium, the accuracy of the microscopic solver, and the accuracy of the data-processing techniques. This general principle has been applied to the analysis of HMM for ordinary differential equations, quasicontinuum methods, some models of interacting-particle systems, as well as a variety of homogenization problems.

Other general methodologies of multiscale modeling have been proposed. We mention in particular the recent version of the quasicontinuum

method and patch dynamics of Kevrekidis et al. In these methods one starts with local simulations based on the microscale model and performs an effective macroscale computation through spatial interpolation and temporal extrapolation on a macroscale grid. Such a methodology is a bottom-up approach, in the sense that it is based on the microscale model and bootstraps the microscale results to macroscale. In contrast, HMM is a top-down approach in that it is based on the macroscale model, and the microscale model is used only to supplement the data. Of course, in some specific cases they may lead to the same method, even though they are based on different philosophies. There are several advantages of a top-down approach: It enables us to develop a framework for the analysis of the methods, as we discussed above. It allows for a selection of the most appropriate microscopic model according to specific needs; e.g., for the modeling of polymeric liquids, one can choose either the kinetic model or molecular dynamics as the microscopic model. It also gives rise to a simple set of design principles, as illustrated in the examples in [9].

So far we have discussed only what is referred to in [9] as the type B problems. Problems of this type typically exhibit a separation of time scales—the microscopic process relaxes locally to equilibrium on a time scale much shorter than the time scale for the macroscopic dynamics. Another typical class of multiscale problems are referred to as the type A problems. These are problems with localized defects around which microscopic models have to be used; elsewhere one can use some macroscopic models. Classical examples of type A problems include crack propagation in solids and contact-line dynamics in fluids. If there is a separation of time scales between the microscopic and the macroscopic processes, then these problems are also of type B and the principles of HMM can still be used. Otherwise they should be treated using adaptive model refinement techniques, which are a further extension of the well-known adaptive mesh refinement methods and are sometimes called heterogeneous domain decomposition.

Outlook

To summarize, the current excitement in multiscale modeling and computation comes from the prospect that a new class of numerical and analytical modeling techniques can be developed by taking into account the special features, such as scale separation, in a very large class of multiscale problems. These new methods promise to be much more efficient than those more traditional multiscale techniques such as multigrid and multiresolution methods, which are intended to solve the fine-scale problems all over the macroscopic domain.

What will be the impact of this new style of multiscale modeling and computations? From the viewpoint of mathematical analysis, the multiphysics nature of these problems means that numerical analysis will become much closer to mathematical physics [21]. In fact, understanding questions that are traditionally regarded as belonging to mathematical physics will be vital to the progress in multiscale modeling. From the viewpoint of modeling problems of scientific and technological interest, it will allow us to remove the ad hoc procedures that are commonly used in many areas, such as plasticity, non-Newtonian fluids, and crack propagation. It will also allow us to deal with problems that fall in between traditional domains of physical theories. A good example is nano-science.

We conclude this article with a historical note. There are two components in the quantitative study of a scientific or engineering problem: modeling and solution. Before the age of computers, the solutions of mathematical models were obtained by special analytic techniques, such as asymptotic analysis and special functions. This often restricted the study to very simplified equations. The advent of computers has made a paradigmatic change in the way we analyze practical problems. The mathematical models can be more realistic when analytic techniques are replaced by numerical methods. Yet in much of computational mathematics, we are used to taking for granted that the models are given, they are the ultimate truth, and our task is to provide methods to analyze and solve them. This shields us from the frontiers of science where phenomena are analyzed and models are formulated.

Multiscale, multiphysics modeling brings in a new paradigm. Here the problems are given, and a variety of mathematical models at different levels of detail can be considered. The right equation is selected during the process of computation according to the accuracy needs. This brings mathematical analysis and computation closer to the actual scientific and engineering problems. It may no longer be necessary to wait for scientists to develop simplified equations before computational modeling can be done. This is an exciting new opportunity for computational science and for applied mathematics. It will bring applied mathematics closer to other fields of mathematics, as, for example, mathematical physics and probability theory. It will also bring these fields closer to the frontiers of science. One effect of this development should be on education. New courses integrating relevant fields of mathematics and fundamental principles of science are needed for the next generation of computational scientists.

Acknowledgement

It is a pleasure to thank Li-Tien Cheng, Tim Kaxiras, Yannis Kevrekidis, Andy Majda, Xiantao Li, Weiqing Ren, Richard Tsai, Eric Vanden-Eijnden, and Pingwen Zhang for many helpful discussions regarding the topic discussed here. Weinan E's work is supported in part by ONR grant N00014-01-1-0674. Bjorn Engquist's work is supported in part by NSF grant DMS-9973341.

References

[1] Atmospheric and ocean science provides a rich source of multiscale problems. For a review, see A. J. Majda, Real world turbulence and modern applied mathematics, *Mathematics: Frontiers and Perspectives 2000* (V. I. Arnold et al., eds.), Amer. Math. Soc., Providence, RI, 2000, pp. 137–152.

[2] F. F. Abraham, J. Q. Broughton, N. Bernstein, and E. Kaxiras, Spanning the continuum to quantum length scales in a dynamic simulation of brittle fracture, *Europhys. Lett.* **44** (6) (1998), 783–787 .

[3] V. Arnold, *Mathematical Methods of Classical Mechanics,* Springer-Verlag, New York.

[4] A. Bensoussan, J.-L. Lions, and G. Papanicolaou, *Asymptotic Analysis of Periodic Structures,* North-Holland, Amsterdam-New York, 1978.

[5] R. B. Bird, C. F. Curtiss, R. C. Armstrong, and O. Hassager, *Dynamics of Polymeric Liquids, vol. 2: Kinetic Theory*, Wiley, New York, 1987.

[6] A. Brandt, Multigrid methods in lattice field computations, *Nuclear Phys. B Proc. Suppl.* **26** (1992), 137–180.

[7] R. Car and M. Parrinello, Unified approach for molecular dynamics and density-functional theory, *Phys. Rev. Lett.* **55** (1985), 2471–2474.

[8] A. J. Chorin, Conditional expectations and renormalization, *Multiscale Model. Simul.* **1** (2003), 105–118.

[9] W. E and B. Engquist, The heterogeneous multiscale methods, *Comm. Math. Sci.* **1** (2003), 87–133.

[10] M. Katsoulakis, A. J. Majda, and D. G. Vlachos, Coarse-grained stochastic processes for lattice systems, *Proc. Natl. Acad. Sci. U.S.A.* **100** (2003), 782–787.

[11] J. B. Keller, Geometrical theory of diffraction, *J. Opt. Soc. Amer.* **52** (1962), 116–130 .

[12] J. Kevorkian and J. D. Cole, *Multiple Scale and Singular Perturbation Methods*, Springer, New York, 1996.

[13] R. LeVeque, *Numerical Methods for Conservation Laws,* Birkhäuser-Verlag, Basel, 1990.

[14] A. J. Majda, I. Timofeyev, and E. Vanden-Eijnden, Systematic strategies for stochastic mode reduction in climate, *J. Atmos. Sci.,* in press.

[15] V. P. Maslov and M. V. Fedoriuk, *Semiclassical Approximation in Quantum Mechanics*, D. Reidel, Dordrecht-Boston, 1981.

[16] E. B. Tadmor, M. Ortiz, and R. Phillips, Quasicontinuum analysis of defects in crystals, *Phil. Mag.* **A73** (1996), 1529–1563.

[17] C. Theodoropoulos, Y.-H. Qian, and I. G. Kevrekidis, 'Coarse' stability and bifurcation analysis using time-steppers: A reaction-diffusion example, *Proc. Nat. Acad. Sci. U.S.A.* **97** (18) (2000), 9840–9843.

[18] E. Vanden-Eijnden, Numerical techniques for multiscale dynamical systems with stochastic effects, *Comm. Math. Sci.* **1** (2003), 385–391.

[19] K. Wilson, *Phys. Rev.* **B4** (1971), 3174, 3184.

[20] K. Xu and K. H. Prendergast, Numerical Navier-Stokes solutions from gas kinetic theory, *J. Comput. Phys.* **114** (1994), 9–17.

[21] H.-T. Yau, Asymptotic solutions to dynamics of many-body systems and classical continuum equations, in *Current Development in Mathematics*, International Press, Somerville, MA, 1999.

About the Cover

Arizona Center, Phoenix, Arizona. The Arizona Center is a downtown hub that offers dining; a variety of shops; a 24-screen movie theater; and acres of gardens, fountains and pools. Phoenix is the site of the 2004 Joint Mathematics Meetings, January 7–10, 2004.

Photograph courtesy of the Greater Phoenix Convention and Visitor Bureau. Photographer: Babe Sarver.