SOME CRITICAL ISSUES FOR THE "EQUATION-FREE" APPROACH TO MULTISCALE MODELING

WEINAN E AND ERIC VANDEN-EIJNDEN

ABSTRACT. The "equation-free" approach has been proposed in recent years as a general framework for developing multiscale methods for efficiently capturing the macroscale behavior of a system using only the microscale models. In this paper, we take a close look at some of the algorithms proposed under the "equation-free" umbrella, the projective integrators and the patch dynamics. We discuss some very simple examples in the context of the "equation-free" approach. These examples seem to indicate that while its general philosophy is quite attractive and indeed similar to many other approaches in concurrent multiscale modeling, there are severe limitations to the specific implementation proposed by this approach.

1. INTRODUCTION

The purpose of this note is to examine some of the basic issues surrounding the "equation-free" approach proposed in [17], which has been pursued in recent years as a general tool for multiscale, multi-physics modeling. To begin with, the equation-free approach is an example of concurrent coupling techniques. In contrast to sequential coupling techniques which require establishing the macroscale equations through precomputing, concurrent coupling techniques compute the required macroscale quantities "on-the-fly" from microscopic models [1, 2]. The most well-known example of such concurrent coupling techniques is perhaps the Car-Parrinello molecular dynamics which computes the atomic interaction forces "on-the-fly" by solving the electronic structure problem [5]. Other algorithms, such as the extended multi-grid method [4] and the heterogeneous multiscale method (HMM) [8] are all example of the concurrent coupling approach.

At a technical level, a key idea in the "equation-free" approach is to make use of scale separation in the system. There are many different ways of exploiting scale separation. In [6] and [5], time scale separation was used to artificially slow down the time scale of the microscopic system. As for spatial scales, homogenization-based methods (such as the ones that use representative averaging volumes [3]) and the quasicontinuum methods [16] are all examples of algorithms that explore the separation of spatial scales. Most closely related to the "equation-free" approach is perhaps the extended multi-grid method [4]. In his review article [4], Achi Brandt described ideas that can be used to extend multi-grid techniques to deal with multiscale, multi-physics problems in order to capture the macroscale behavior of a system using microscopic models such as molecular dynamics. As is common in multi-grid methods, the ideas of Brandt rely heavily on mapping back and forth between the macro- and micro-states of the system, through prolongation and restriction operators (which are called respectively reconstruction and compression operators in HMM, and lifting and restriction operators in the "equation-free" approach). Brandt realized that central to the efficiency of these algorithms is the possibility of only performing microscopic simulations in small samples for short periods of times, as a result of the scale separation in the system. These ideas are later adopted by both HMM and

the "equation-free" approach. In fact, HMM and the "equation-free" approach are both alternative approaches with the same motivation and objective.

	Macro to micro	micro to Macro
Extended multi-grid	interpolation	restriction
HMM	reconstruction	compression
Equation-free	lifting	restriction

While the general philosophy of the "equation-free" approach is very similar to the extended multi-grid method and HMM, the "equation-free" approach proposes its own ways of implementing such a philosophy, in particular, ways of dealing with scale separation. The basic idea is to use extrapolation in time and interpolation in space. More precisely, two important building blocks of the "equation-free" approach are:

i) The projective integrators: (An ensemble of) the microscale problems are solved for a short period of time using small time steps. The time derivative of the macro variable is computed from the results of the last few steps and then used to advance the macro variable over a macro time step. It is easy to see that such a procedure amounts to extrapolation, and indeed the authors state in [13]: *"The reader might think that these should be called 'extrapolation methods,' but that name has already been used [...]. Hence we call the proposed methods projective integration methods."*

ii) The gap-tooth scheme: The microscopic problem is solved in small domains (the teeth) separated by large gaps. The solution is averaged over each domain and then interpolated to give the prediction over the gaps.

The combination of these two ideas gives directly the so-called "patch dynamics" [17].

Detailed understanding of the "equation-free" algorithms is made difficult by the fact that the "equation-free" papers are generally quite vague. The present note should be regarded as an attempt to pin down some of these details. Indeed this was initially intended as a regular journal article. But it soon becomes clear that there is still substantial disagreement between our understanding of the "equation-free" approach and that of its developers. However, we believe the simple examples that we discuss here do shed some light on the "equation-free" approach and should be made available to a larger audience in some form. We are grateful to Yannis Kevrekidis for a detailed report on the earlier version of this note. Some of his comments have been taken into account in this revised version. We also welcome any discussion about the issues raised in this note, the most important of which being: What really is the "equation-free" approach? Indeed our primary purpose of presenting this note is to prompt such a discussion.

2. PROJECTIVE INTEGRATORS FOR STOCHASTIC ODES

Projective integrators were proposed as a way of extrapolating the solution of an explicit ODE solver for systems with multiple time scales using large time steps. The basic idea is to run the microscopic solver (using small time steps) for a number of steps, and then estimate the time derivative and use that to extrapolate the solution over a large time step [13]. For stiff ODEs, the extrapolation step is applied to the whole system [13]. For general multiscale problems, the extrapolation step is applied only to the slow variables [17, 15].

In the case of stiff ODEs, projective integrators can give rise to useful numerical schemes, as was demonstrated in [13]. In this case, the idea becomes very close to the ones proposed by Eriksson *et. al* for developing explicit stiff ODE solvers [12]. The objectives

of the two papers are quite different: For Eriksson *et al.*, the objective is to find explicit and efficient stiff ODE solvers. For Gear *et al.*, the objective is to deal with general multiscale, multi-physics problems. However, in the general case such as the case considered in [15], projective integrators have serious limitations, as we now show.

Denote by x the coarse variable of the system. The coarse projective integrators proposed in [15] performs the following steps at each macro time step (of size Δt):

i) Create an ensemble of N microscopic initial conditions consistent with the known coarse variable x^n at time step n.

ii) Run the microscopic solver with these initial conditions for a number of steps, say k, with time step δt. Denote the corresponding values of the coarse variables as $\tilde{x}_j(k\delta t)$ where $j = 1, \ldots, N$.

iii) Perform ensemble averaging to get an approximation to the coarse variable. For example,

$$\bar{x} = \frac{1}{N}\sum_{j=1}^{N} \tilde{x}_j(k\delta t) \tag{1}$$

iv) Use this value to extrapolate the coarse variable to a time step of size Δt.

$$x^{n+1} = x^n + \Delta t \frac{\bar{x} - x^n}{k\delta t} \tag{2}$$

Now consider the simple case when the coarse variable obeys effectively a stochastic ODE:

$$dx(t) = b(x(t))dt + dW(t) \tag{3}$$

Since, to $O(k\delta t)$, we have

$$\tilde{x}_j(k\delta t) - x^n = k\delta t\, b(x^n) + \sqrt{k\delta t}\, \xi_j^n \tag{4}$$

where $\{\xi_j^n\}, j = 1, \cdots, N$ are N independent Gaussian variables with mean 0 and variance 1, (2) becomes, to leading order

$$x^{n+1} = x^n + \Delta t\, b(x^n) + \frac{\Delta t}{\sqrt{k\delta t}} \frac{1}{N}\sum_{j=1}^{N} \xi_j^n. \tag{5}$$

(5) is equivalent in law to

$$x^{n+1} = x^n + \Delta t\, b(x^n) + \frac{\Delta t}{\sqrt{Nk\delta t}}\, \xi^n, \tag{6}$$

where ξ^n is a Gaussian variable with mean 0 and variance 1.

It is obvious from this that the effective dynamics produced by the coarse projective integrator depends on the numerical parameters N, k, δt, and Δt. In particular, if $Nk\delta t \gg \Delta t$, then the noise term in (3) is lost in the limit. If $Nk\delta t \ll \Delta t$, then the noise term overwhelms the drift term. In either case, one obtains a wrong prediction for the effective dynamics of the coarse variable.

The only way to get a scheme consistent with (3) is to choose the numerical parameters so that they precisely satisfy $Nk\delta t = \Delta t$. The reader should be aware, however, that this choice is not advocated in [15] and is, in fact, quite orthogonal to the original equation-free philosophy since it requires knowing beforehand that (3) is an SODE and not something else. In addition, it is easy to see that using $Nk\delta t = \Delta t$ leads to no gain in efficiency: The total cost is comparable to solving the microscopic problem in a brute force fashion using

δt as the time step, since the size of the ensemble is equal to the number of microscopic simulation time intervals during a time duration of Δt: $N = \Delta t/(k\delta t)$.

For the case when $Nk\delta t \gg \Delta t$, one might think of using the coarse projective integrators (or coarse molecular dynamics) as a way of simulating the dynamics $dx/dt = b(x)$ in the context of molecular dynamics simulations. In this case the unknown drift $b(x)$ is related to the gradient of the free energy and simulating $dx/dt = b(x)$ is then a way to explore this free energy. Indeed, this appears to be how the scheme was actually used in [15]. The problem, however, is that using $Nk\delta t \gg \Delta t$ leads again to a scheme which is no less expensive than a brute force solution of N replica of (3) using δt as the time step.

The problem above seems to be intrinsic to projective integrators in the context of SDEs because it is inherent to the fact that the dynamics (3) is dominated by the noise on short time scales and the extrapolation step in the projective integrators amplifies these fluctuations. Averaging them out can only be done at a cost which is at least comparable to the cost of a direct scheme.

3. PATCH DYNAMICS

Patch dynamics is proposed as a way of analyzing the macroscopic dynamics of a system using microscopic models. Like the extended multi-grid methods [4] and HMM [8, 11], it is formulated in such a way that scale separation can be exploited to reduce computational cost.

The setup is as follows. We have a macroscale grid over the computational domain. The grid size Δx is chosen to resolve the macroscale variations but not the microscale features in the problem. Each grid point is surrounded by a small domain (the "tooth"), the size of which (denoted by h) should be large enough to sample the local microscale variations but can be much smaller than the macroscale grid size if the macro and microscales are very much separated.

Given a set of macroscale values at the macroscale grid points, $\{U_j^n\}$, at the n-th time step $t_n = n\Delta t$ where Δt is the size of the macroscale time step, patch dynamics computes the update of these values at the next macroscale time step, $\{U_j^{n+1}\}$, using the following procedure:

i) *Lifting:* From $\{U_j^n\}$, reconstruct a consistent microscopic initial data, denoted by $\tilde{u}_0$.

ii) *Evolution:* Solve the original microscopic model with this initial data $\tilde{u}_0$ over the small domains (the "teeth") for some time δt: $\tilde{u}_{\delta t} = \mathcal{S}_{\delta t}\tilde{u}_0$.

iii) *Restriction:* Average the microscale solution $\tilde{u}_{\delta t}$ over the small domains. The results are denoted by $\{\tilde{U}_{\delta t}^n\}$.

iv) *Extrapolation:* Compute the approximate derivative and use it to predict $\{U_j^{n+1}\}$:

$$U^{n+1} = U^n + \Delta t \frac{\tilde{U}_{\delta t}^n - U^n}{\delta t} \tag{7}$$

or more generally:

$$U^{n+1} = U^n + \Delta t \frac{\tilde{U}_{\delta t}^n - \tilde{U}_{\alpha\delta t}^n}{(1-\alpha)\delta t} \tag{8}$$

where α is some numerical parameter between 0 and 1.

There are very few examples of how to implement these steps in practice. [17, 24] suggest the following:

For the lifting operator, in the small domain around the macro grid point x_j, use the approximate Taylor expansion:

$$\tilde{u}_0(x) = \sum_{k=0}^{d} \frac{1}{k!} D_k (x - x_j)^k \tag{9}$$

Here D_k is some approximations to the derivatives of the macroscale profile at x_j, for example:

$$D_2 = \frac{U^n_{j+1} - 2U^n_j + U^n_{j-1}}{\Delta x^2}, \quad D_1 = \frac{U^n_{j+1} - U^n_{j-1}}{2\Delta x}, \quad D_0 = U^n_j - \frac{1}{24} h^2 D_2 \tag{10}$$

Below we will consider the case when $d = 2$.

When solving the microscale problem, [24] suggest extending the microscale domain to include some buffer regions in the hope that this would allow the use of *any* boundary conditions for the microscopic solver: By choosing sufficiently large buffers, the effect would be as if the microscale problem is solved in the whole space where and when averaging is performed. This introduces another spatial scale H which is the real size of the region on which microscale problems are solved. (The parameter h, which is (much) smaller than H, is the size of the domain over which averaging is performed). In the following discussion, we will take H to be infinity.

Let us now examine this algorithm in more detail, using some very simple examples. Let us first consider the heat equation

$$\partial_t u = \partial_x^2 u \tag{11}$$

For simplicity, let us assume $x_j = 0$. Denote $\tilde{u}_0 = D_0 + D_1 x + \frac{1}{2} D_2 x^2$. We have

$$\mathcal{S}_{\delta t} \tilde{u}_0(x) = D_0 + D_1 x + D_2 (\frac{1}{2} x^2 + \delta t) \tag{12}$$

Denote by $\mathcal{A}_h$ the averaging operator over the small domain (of size h), we have

$$\tilde{U}^n_{\delta t} = \mathcal{A}_h \mathcal{S}_{\delta t} \tilde{u}_0(x) = D_0 + D_2 \delta t + \frac{1}{24} D_2 h^2 = U^n + D_2 \delta t \tag{13}$$

Inserting this expression in (7) gives the familiar scheme:

$$U^{n+1} = U^n + \Delta t D_2 \tag{14}$$

as was shown in [17]. This is both stable and consistent with the heat equation, which is the right effective model at the large scale.

Now let us turn to the advection equation

$$\partial_t u + \partial_x u = 0 \tag{15}$$

In this case, we have

$$\mathcal{S}_{\delta t} \tilde{u}_0(x) = D_0 + D_1 (x - \delta t) + \frac{1}{2} D_2 (x - \delta t)^2 \tag{16}$$

Hence,

$$\tilde{U}^n_{\delta t} = \mathcal{A}_h \mathcal{S}_{\delta t} \tilde{u}_0(x) = D_0 - D_1 \delta t + \frac{1}{2} D_2 \delta t^2 + \frac{1}{24} D_2 h^2 = U^n - D_1 \delta t + \frac{1}{2} D_2 \delta t^2 \tag{17}$$

and (7) becomes

$$U^{n+1} = U^n + \Delta t (-D_1 + \frac{1}{2} \delta t D_2) \tag{18}$$

Since $\delta t \ll \Delta t$, the last term is much smaller than the other terms, and we are left essentially with a scheme which is unstable under the standard CFL condition that $\Delta t \sim \Delta x$:

$$U^{n+1} = U^n - \Delta t D_1 \tag{19}$$

due to the central character of D_1.

Aside from the stability issue, there can also be problems with consistency. Consider the following example:

$$\partial_t u = -\partial_x^4 u \tag{20}$$

The macroscale model is obviously the same model. However, it is easy to see that if we follow the patch dynamics procedure with $d = 2$, we would be solving $\partial_t U = 0$, which is obviously inconsistent with the correct macroscale model.

For these simple examples, the difficulties discussed above can be fixed by using different reinitialization procedure for the micro-solvers. For the example of the convection equation, one should use one-sided interpolation schemes in the spirit of upwind schemes. For the last example, one should use piecewise $4th$ order polynomial interpolation. *But in general, finding such a reinitialization procedure seems to be quite a daunting task, since it depends on the nature of the unknown effective macroscale model*. Imagine that the microscopic solver is molecular dynamics. The reinitialization procedure has to take into account not only consistency with the local macrostates of the system (which is the only requirement for the extended multi-grid method and HMM), but also the effective macroscale scale model (which is unknown) such as:

i) The order of the macroscale equation.
ii) The direction of the wind, if the effective macroscale equation turns out to be a first order PDE.
iii) Other unforeseeable factors.

Indeed it is not at all clear how patch dynamics would work if molecular dynamics models are used to model macroscopic gas dynamics.

4. CONCLUSIONS

The idea of interrogating legacy codes as a control system is very attractive and to some extend, has already been commonly used in some disciplines. For example, chemists use packages such as CHARMM and AMBER as legacy codes to perform optimization tasks, e.g. to find free energy surfaces and minimum free energy paths. Optimization techniques such as the Nelder-Mead algorithm that use only function values (not the derivatives) were designed with this kind of problems in mind. One purpose of the work of Keller *et al.* is to extend bifurcation analysis tools to systems that are defined by legacy codes [25]. The "equation-free" approach attempts to extend such practices to another direction, namely the modeling of macroscale spatial/temporal dynamics of systems defined by microscopic models, in the form of legacy codes. While this seems very attractive, the set of tools proposed under this umbrella are quite far from being sufficient for reaching this objective. We have discussed some of the technical difficulties in this note. This discussion is certainly not exhaustive. It is only meant to be illustrative.

ACKNOWLEDGMENTS

We are very grateful to the many people, including Yannis Kevrekidis, who have read the draft of this paper at various stages and gave us their comments. It is also a pleasure to thank Ju Li for discussions that helped to clarify some of the algorithmic details in the

baby-bathwater scheme. The work of W. E was supported in part by ONR grant N00014-01-1-0674 and by DOE grant DOE DE-FG02-03ER25587. The work of E. V.-E. was supported in part by NSF grants DMS02-09959 and DMS02-39625 and by ONR grant N00014-04-1-0565.

REFERENCES

[1] Abraham, F. F.; Broughton, J. Q.; Bernstein, N.; Kaxiras, E. Concurrent Coupling of Length Scales: Methodology and Application, *Phys. Rev. B*, vol. 60, no. 4, pp. 2391–2402, 1999.

[2] Abraham, F. F.; Broughton, J. Q.; Bernstein, N.; Kaxiras, E. Spanning the Continuum to Quantum Length Scales in a Dynamic Simulation of Brittle Fracture, *Europhys. Lett.*, vol. 44, no. 6, pp. 783–787, 1998.

[3] Bear, J. *Dynamics of Fluids in Porous Media*, Dover Publications; New Ed edition (September 1, 1988).

[4] Brandt, A. Multiscale Scientific Computation: Review 2001, *Multiscale and Multiresolution Methods: Theory and Applications*, Yosemite Educational Symposium Conf. Proc., 2000, Lecture Notes in Comp. Sci. and Engrg., T.J. Barth, et.al (eds.), vol. 20, pp. 3–96, Springer-Verlag, 2002.

[5] Car. R.; Parrinello, M. Unified Approach For Molecular Dynamics and Density Functional Theory. *Phys. Rev. Lett.* **55**(22):2471–2475 (1985).

[6] Chorin, A.J. A numerical method for solving incompressible viscous flow problems. *J. Comp. Phys.* **2**: 12–26 (1967).

[7] Chorin, A. J.; Kast, A. P.; Kupferman, R. Optimal Prediction of Under-resolved Dynamics, *Proc. Natl. Acad. Sci. USA*, vol. 95, no. 8, pp. 4094-4098, 1998.

[8] E, W.; Engquist, B. The heterogeneous multi-scale methods. *Comm. Math. Sci.* **1**: 87–133 (2003).

[9] E, W.; Liu, D.; Vanden-Eijnden, E. Analysis of Multiscale Methods for Stochastic Differential Equations. *Comm. Pure App. Math.* **58**: 1544–1585 (2005).

[10] E, W.; Liu, D.; Vanden-Eijnden, E. Nested stochastic simulation algorithm for chemical kinetic systems with multiple time-scales. *J. Comp. Phys.* **221**: 158–180 (2007).

[11] E, W.; Engquist, B.; Li, X.; Ren, W.; Vanden-Eijnden, E. Heterogeneous Multiscale Methods: A Review. *Comm. Comp. Phys.* **2**(3): 367–450 (2007).

[12] Eriksson, K.; Johnson, C.; Logg, A. Explicit Time-stepping for stiff ODEs. *SIAM J. Sci. Comp.* **25**(4):1142-1157 (2003).

[13] Gear, C. W.; Kevrekidis, I. G. Projective methods for stiff differential equations: problems with gaps in their eigenvalue spectrum. *SIAM J. Sci. Comp.* **24**(4):109-110 (2003).

[14] Givon, D.; Kevrekidis, I. G.; Kupferman, R. Strong Convergence of Projective Integration Schemes for Singularly Perturbed Stochastic Differential Systems. *Comm. Math. Sci.* **4**: 707–729 (2006).

[15] Hummer, G; Kevrekidis, I. G. Coarse molecular dynamics of a peptide fragment: Free energy, kinetics, and long-time dynamics computations. *J. Chem. Phys.* **118**: 10762–10773 (2003).

[16] Knap, J.; Ortiz, M. An Analysis of the Quasicontinuum Method, *J. Mech. Phys. Solids.*, vol. 49, no. 9, pp. 1899–1923, 2001.

[17] Kevrekidis, I. G.; Gear, C. W.; Hyman, J. M.; Panagiotis, G. K.; Runborg, O.; Theodoropoulos, C. Equation-Free, Coarse-Grained Multiscale Computation: Enabling Microscopic Simulators to Perform System-Level Analysis. *Comm. Math. Sci.* **1**: 87–133 (2003).

[18] Khasminsky, R. Z. On Stochastic Processes Defined by Differential Equations with a Small Parameter. *Theory Prob. Applications*, **11**:211–228, 1966.

[19] Kesten, K.; Papanicolaou, G. A limit theorem for stochastic acceleration, *Comm. Math. Phys.*, **78**: 19-63 (1980).

[20] Kurtz, T. G. A limit theorem for perturbed operator semigroups with applications to random evolution. *J. Functional Analysis* **12**: 55–67 (1973).

[21] Maragliano, L.; Vanden-Eijnden, E. A temperature accelerated method for sampling free energy and determining reaction pathways in rare events simulations. *Chem. Phys. Lett.* **426**: 168–175 (2006).

[22] Papanicolaou, G. C. Some probabilistic problems and methods in singular perturbations. *Rocky Mountain J. Math.* **6**:653–674 (1976).

[23] Papavasiliou, A.; Kevrekidis, I. G. Variance Reduction for the Equation-Free Simulation of Multiscale Stochastic Systems *Multiscale Model. Simul.* **6**(1): 70–89 (2007).

[24] Samaey, G.; Kevrekidis, I. G.; Roose, D. Patch dynamics with buffers for homogenization problems, *J. Comput. Phys.*, **213**, 264-287 (2006).

[25] Shroff, G. M.; Keller, H. B. Stablization of unstable procedures: A recursive projection method, *SIAM J. Numer. Anal.*, vol. **30**, pp. 1099-1120, 1993.

[26] Vanden-Eijnden, E. Numerical techniques for multiscale dynamical systems with stochastic effects. *Comm. Math. Sci.*, **1**: 385–391 (2003).
[27] Vanden-Eijnden, E. On HMM-like integrators and projective integration methods for systems with multiple time scales. *Comm. Math. Sci.*, **5**: 495-505 (2007).