LBNL-40696
UC-405
Preprint

ERNEST ORLANDO LAWRENCE BERKELEY NATIONAL LABORATORY

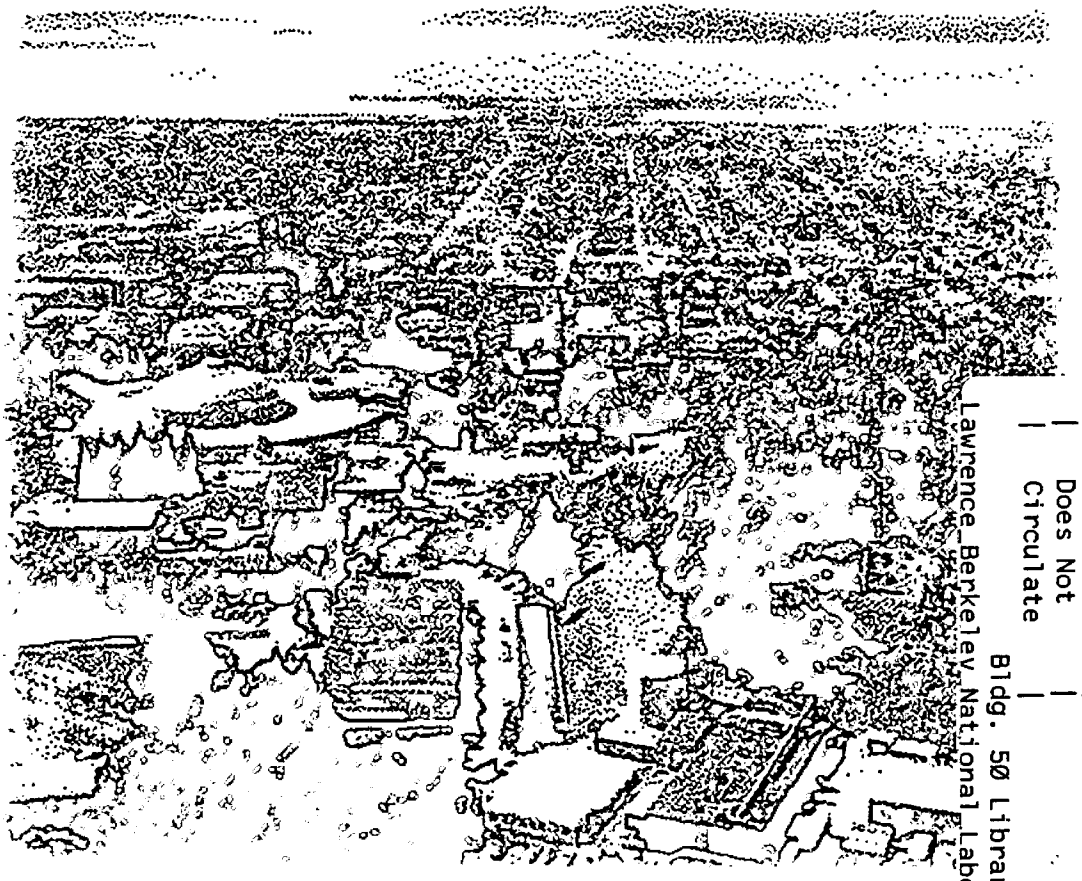
Reducing the Average Error of Underresolved Approximations

Anton Kast and Alexandre J. Chorin

Computing Sciences Directorate
Mathematics Department

August 1997

To be submitted
for publication



REFERENCE COPY
Does Not
Circulate
Bldg. 50 Library - Ref.
Lawrence Berkeley National Laboratory
Copy 1
LBNL-40696

DISCLAIMER

This document was prepared as an account of work sponsored by the United States Government. While this document is believed to contain correct information, neither the United States Government nor any agency thereof, nor the Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or the Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof or the Regents of the University of California.

LBNL-40696

REDUCING THE AVERAGE ERROR OF UNDERRESOLVED APPROXIMATIONS*

Anton Kast
Alexandre J. Chorin

Mathematics Department
Lawrence Berkeley National Laboratory
University of California
Berkeley, CA 94720

August 1997

*This work was supported in part by the Applied Mathematical Sciences subprogram of the Office of Energy Research, U.S. Department of Energy, under Contract DE-AC03-76SF00098.

Reducing the average error of underresolved approximations

Anton Kast and Alexandre J. Chorin

July 30, 1997

Abstract

There exist problems of practical interest (in particular in turbulence) where the solution is so complex that a resolved numerical calculation is impractical, and where the solution is also chaotic, so that one is interested only in the large scales averaged over the random details. We show by examples that if appropriate statistical information is available, one can calculate accurately the large-scale behavior of the solution even on a crude grid, and that the accuracy of the result can be reliably assessed. The application of the ideas to turbulence is also briefly discussed.

1 Introduction

A necessary condition for the convergence of a numerical scheme is that the solution be sufficiently well-resolved, i.e, that all the degrees of freedom that contribute significantly to the solution be represented in the computation (see e.g [1]). In finite difference schemes, whose vocabulary we adopt for the sake of simplicity, the grid must be fine enough to represent all the relevant

scales (with spectral methods, one must have enough Fourier components, etc.). In turbulence and in related problems it is generally impossible, now or in any foreseeable future, to resolve the solution well enough, and even if one could, one would not necessarily want to; the relevant solutions of the equations of motion are chaotic, i.e, depend sensitively on the initial data, and no specific solution is likely to ever be observed in practice. What one wants is an accurate evaluation of the significant, large scales, accurately averaged over the chaotic details. These obvious observations have led in hydrodynamics to the various "large-eddy simulations", based on heuristic analyses of the physics of turbulence [2]; we suggest better ways of deriving and assessing such methods.

Specifically, we show by example how a probability distribution for the solution of a differential equation can be put to good use. The examples are discretized nonlinear Schrödinger equations and a randomly driven heat equation. The relevance of our constructions to turbulence is briefly discussed in the concluding section.

2 The statistics of numerical error

Suppose one is trying to solve an equation of the form

$$\frac{\partial u}{\partial t} = R(u), \tag{1}$$

where t is the time and $R(u) = R(u, \partial u/\partial x, \dots)$ is a (generally nonlinear) function of u and its spatial derivatives. For definiteness, assume $0 \leq x \leq 1$, with appropriate boundary conditions at 0 and 1. Let equation (1) be approximated by finite differences in the form

$$\frac{du_i}{dt} = R_i^N(u_1, \dots, u_N), \quad i = 1, \dots, N \tag{2}$$

where i is an index which specifies a mesh point $x_i = i/N$, R^N is a difference approximation on the grid and u_i approximates $u(x_i)$. For simplicity, only the space variable has been discretized; this will not cause a loss of generality. Assume further that a measure on the space of solutions of equation (1) is known, and is written symbolically as $F[du]$; given an appropriate set S of functions, the average value of a functional $G[u]$ over the u in the set S is

$$\langle G[u] \rangle = \int_S G[u] F[du], \tag{3}$$

where the integral is in general a function-space integral; specific examples will be given below. If the solution of equation (1) is well-resolved, u is by

definition smooth between grid points and knowing u at the grid points x_i makes it possible to calculate the values of u at other points by interpolation; if the solution is not well-resolved, the function u can vary substantially between mesh points, and the scales invisible to the grid may have a large effect on the values of $R(u)$ at the grid points. The probability density F makes it possible to fill the gaps between mesh points in an average sense.

Suppose the values u_i are given; in the ill-resolved case there are many different functions u that can assume these values. The probability density F induces a (conditional) probability density on this restricted space of functions. An average of a function $G[u]$ with respect to this restricted measure, i.e., an average over all the functions in the set S that take on the given values $u_1, \ldots, u_N$ at the mesh points, can be written in the symbolic form

$$\langle G[u] \rangle_{u_1 \cdots u_N} = \int_S G[u] \prod_{i=1}^{N} \delta\left[u(x_i) - u_i\right] F[du]. \tag{4}$$

In the present paper we discuss only problems where the measure $F[du]$ is invariant in time, when it suffices to discuss how to approximate $R(u)$; the resulting autonomous set of ordinary differential equations can be solved by standard means.

A numerical scheme produces at each point x_i a value $R_i^N(u_1, \ldots, u_N)$. If the numerical scheme is underresolved, this value may be a good or bad

approximation to $R(u)|_{x=x_i}$ for any particular function u that takes on the values u_i at the points x_i; the error can be evaluated only for each specific function u. The sensible thing to look for is an approximation to $\langle R(u)|_{x=x_i}\rangle_{u_1\cdots u_N}$, i.e., an approximation of the average of $R(u)$ over all functions u that take on the given values at the given mesh points. The error, i.e., the amount by which one fails to calculate this average, is

$$e_i = R_i^N(u_1,\ldots,u_N) - \langle R(u)|_{x=x_i}\rangle_{u_1\cdots u_N}. \tag{5}$$

This error depends only on the discrete data u_i. The root mean square numerical error is defined as:

$$E = \sqrt{\left\langle \frac{1}{N}\sum_{i=1}^{N} e_i^2 \right\rangle}, \tag{6}$$

i.e., the average of the norm of the error vector over all solutions of the problem.

Since the average error (5) depends only on the discrete data, and since the choice of numerical scheme $R_i^N(u_1,\ldots,u_N)$ is at our disposal, it is in principle possible to choose the scheme so that the average error is zero regardless of the degree of underresolution. The challenge is to determine how the constrained expectation value in (5) depends on the constraints, and then to encode this dependence into a scheme. We shall show by examples

how this can be done.

3 A randomly driven diffusion equation

Consider a one-dimensional bar located at $0 \leq x \leq 1$, heated at its center by a fluctuating heat source and cooled at its ends by heat baths. The temperature $u(x,t)$ of the bar is described by:

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + J(t)\,\delta\left(x - \frac{1}{2}\right) \tag{7}$$

where $J(t)$ is a random function of time, $\delta(x)$ is a delta function and $x = \frac{1}{2}$ is the location of the source. The boundary conditions are

$$u(0,t) = u(1,t) = 0. \tag{8}$$

Assume that $J(t)$ is white noise, i.e. $J(t)$ is a Gaussian random function of time that has expectation values

$$\begin{aligned} \langle J(t) \rangle &= \overline{J} \\ \langle J(t_1) J(t_2) \rangle &= \overline{J}^2 + \overline{J^2}\,\delta(t_1 - t_2) \end{aligned} \tag{9}$$

for some numbers $\overline{J}$ and $\overline{J^2}$.

The solution $u(x,t)$ of (7) depends on $J(t)$ and on the initial condition for $u(x,0)$. However, after long times it depends only on $J(t)$:

$$u(x,t) = \int_0^\infty dt'\, G(x, t - t')\, J(t'), \qquad (t \gg 0) \tag{10}$$

where $G(x,t)$ is the solution of

$$\frac{\partial G}{\partial t} = \frac{\partial^2 G}{\partial x^2} \tag{11}$$

with initial condition $G(x,0) = \delta(x - \frac{1}{2})$. The heat kernel G is readily computed and so are, after suitably long times, average values of u and of functionals of u. This is enough information to calculate the average errors (5), and to construct a scheme to minimize E.

We denote the values of u at the mesh points i/N by $u_i = u(i/N)$, $i = 1, \ldots, N$. We limit the infinite domain of integration in (10) to $0 \leq t' \leq 1$ and discretize the resulting integral so that the statistics of the u_i are determined by

$$u_i = \frac{1}{T} \sum_{j=1}^{T} G_{ij} J_j \tag{12}$$

where $G_{ij} = G(i/N, j/T)$ and $J_1, \ldots, J_T$ are independent random Gaussian variables with

$$\begin{aligned} \langle J_j \rangle &= \overline{J} \\ \langle J_{j_1} J_{j_2} \rangle &= \overline{J}^2 + T\overline{J^2}\delta_{j_1 j_2}. \end{aligned} \tag{13}$$

We would like a formula for the operator $R^N(u_1, \ldots, u_N)$, which would satisfy, as closely as possible,

$$R_i^N(u_1, \ldots, u_N) \approx \left\langle \frac{\partial^2 u}{\partial x^2}(x_i) \right\rangle_{u_1 \cdots u_N}. \tag{14}$$

We proceed by sampling discrete data $u_1, \ldots, u_N$, computing the constrained average in (14) for each set, and fitting R^N.

We sample $u_1, \ldots, u_N$ by generating a random noise vector $\mathbf{J}$ whose components are $J_1, \ldots, J_T$ and applying (12). To get the constrained average corresponding to this sample (the right-hand side of (14)), we first find a basis of K vectors $\mathbf{J}^1, \ldots, \mathbf{J}^K$ in the null space of the matrix with components G_{ij}:

$$\sum_{j=1}^{T} G_{ij} J_j^k = 0, \quad k = 1, \ldots, K. \tag{15}$$

(we used the commercially-available Matlab program, which obtains a rational basis in the null space by singular value decomposition using the LINPACK routine ZSVDC). The set of all noise vectors that satisfy equation (12) is spanned by the K vectors $\mathbf{J}^k$ shifted by $\mathbf{J}$. To compute the constrained averages in (14), it is sufficient to sample this space by a Metropolis algorithm, evaluate the relevant functional (in this case $\partial^2 u/\partial x^2$ at $x = x_i$) for each sample, and average the results.

Specifically, the statistics of the noise are characterized by a probability density function,

$$P[\mathbf{J}] = \exp\left[-\frac{1}{2T\overline{J^2}} \sum_{j=1}^{T} \left(\mathbf{J}_j - \overline{J}\right)^2\right]. \tag{16}$$

Let $\mathbf{J}$ be a noise vector consistent with the constraint equations (12) for some given $u_1, \ldots, u_N$. We can obtain a new trial noise vector $\mathbf{J}'$, also consistent with the constraints, by forming an arbitrary linear combination of null vectors and adding them to $\mathbf{J}$; this new noise vector is accepted or rejected by the usual Metropolis recipe based on the probability density $P[\mathbf{J}]$. Thus one can calculate a sequence of noise vectors consistent with the constraints and with the statistics (13).

Now we construct a scheme that uses this information: we approximate the second derivative of u in the form

$$R_i^N(u_1, \ldots, u_N) = \sum_{j=1}^{N} c_{ij} u_j, \tag{17}$$

with undetermined coefficients c_{ij}. For any discrete function $u_1, \ldots, u_N$, we may perform Monte Carlo calculations to obtain $\langle \partial^2 u / \partial x^2(x_i) \rangle_{u_1 \cdots u_N}$ for $i = 1, \ldots, N$. We accumulate such averages for a large number L of discrete functions $u^{(1)}, \ldots, u^{(L)}$, and require

$$\sum_{j=1}^{N} c_{ij} u_j^{(l)} \approx \left\langle \frac{\partial^2 u}{\partial x^2}(x_i) \right\rangle_{u_1^{(l)} \cdots u_N^{(l)}}, \quad l = 1, \ldots, L \tag{18}$$

which constitutes an overdetermined linear system for the c_{ij}. We may get a best approximate solution for the c_{ij} by least-squares fitting. With the c_{ij} in hand, the formula (17) becomes an optimized scheme in the sense that the

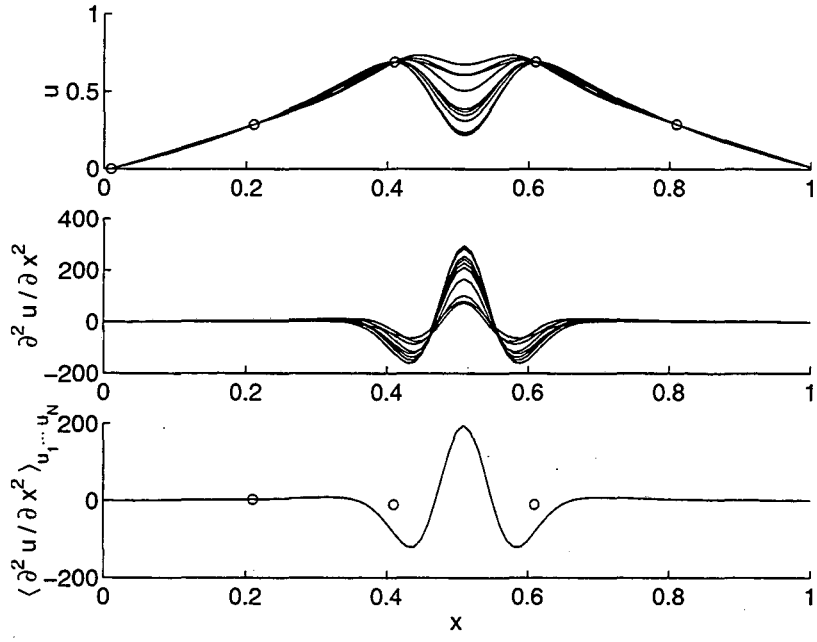


Figure 1: The sampling procedure used to compute average numerical error. 1a: a discretization and continuous functions consistent with it. 1b: the second derivatives of these continuous functions. 1c: the line is the mean of these derivatives, and the circles are naive estimates of the second derivative made by (19) acting on the original discrete data.

mean error (6) has been minimized over all linear schemes.

The optimized scheme may be compared with a naive way of calculating the second derivative of u using finite differences:

$$R_i^N(u_1, \ldots, u_N) = \frac{u_{i+1} - 2u_i + u_{i-1}}{h^2} \tag{19}$$

where $h = 1/N$. Figure 1a shows a family of temperature functions that share a single, poorly-resolved discretization ($N = 5$). In 1b and 1c, the second derivatives and the mean second derivative are shown. Also in 1c, the

Naive:

		j				
		1	2	3	4	5
	1	0	0	0	0	0
	2	25	-50	25	0	0
i	3	0	25	-50	25	0
	4	0	0	25	-50	25
	5	0	0	0	0	0

Improved:

		j				
		1	2	3	4	5
	1	0	-61.8	32.4	0	0
	2	0	31.5	-29.5	0	0
i	3	0	0	0	-29.5	31.5
	4	0	0	0	32.4	-61.8
	5	0	0	0	0	0

Table 1: Table of coefficients c_{ij} for the improved scheme (17) and corresponding coefficients for the naive scheme (19).

result of the naive scheme (19) acting on the discrete data is shown and this result clearly differs from the true average second derivative.

Table 1 shows the matrix of scheme coefficients c_{ij} calculated as described above, with $N = 5$, $T = 100$, $\overline{J} = 1$ and $\overline{J^2} = 1$. These results were obtained using $L = 10^3$ samples and with 10^3 constrained samples used for each constrained average. This table also shows the matrix of coefficients corresponding to the naive scheme (19), for comparison.

Figure (2) shows histograms of errors for the naive scheme (19) and for the improved scheme (17) in this case. Values of e_i have been combined for all

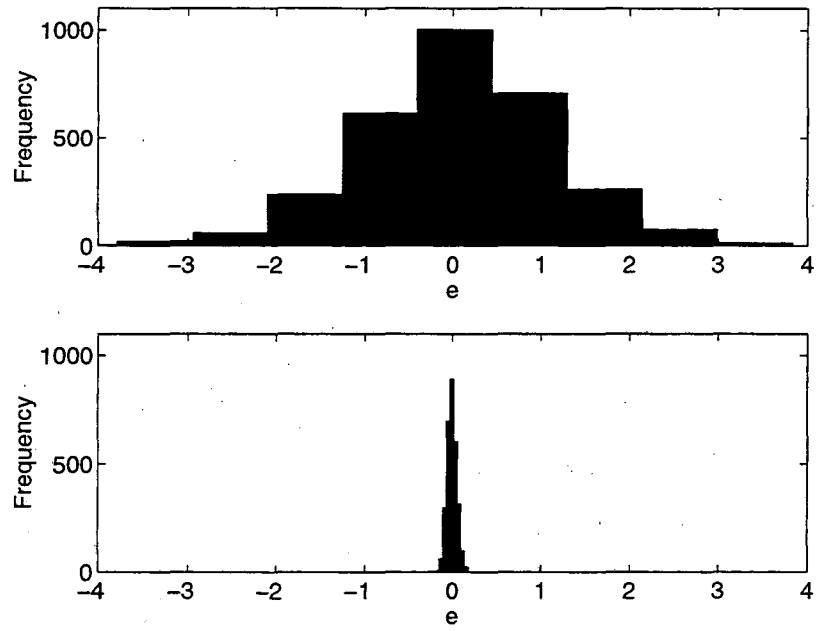


Figure 2: Histograms of errors for the naive scheme and for the optimized scheme.

i. The RMS error E (the widths of these distributions) is reduced by a factor of appproximately 18 using the improved scheme instead of the naive one. Most of the error in the effective scheme is a consequence of imperfections in the numerical Monte Carlo calculation. A small contribution may be due to imperfections in the linear fitting, i.e. the constrained average in (14) may not actually be a linear function of the constraints. (The latter error could be remedied by using an appropriate nonlinear scheme instead of (17).)

The new scheme should be particularly effective near the point $x = \frac{1}{2}$, where the effect of the noise is strongly felt (one should of course avoid

putting a grid point exactly on the singular point).

We would like to emphasize that once we have the improved matrix of coefficients c_{ij} in hand, no more Monte Carlo calculations are needed to achieve the improved accuracy in subsequent applications. Our scheme is (17) with the coefficients in Table 1, and sampling is not performed once the scheme has been determined.

4 A discretized nonlinear Schrödinger equation

Consider the Hamiltonian partial differential equations

$$\begin{aligned} \frac{\partial q}{\partial t} &= -\frac{\partial^2 p}{\partial x^2} + p^3 \\ \frac{\partial p}{\partial t} &= \frac{\partial^2 q}{\partial x^2} - q^3 \end{aligned} \quad . \tag{20}$$

with Hamiltonian density

$$H[q,p] = \frac{1}{2}\int \left[\left(\frac{\partial q}{\partial x}\right)^2 + \left(\frac{\partial p}{\partial x}\right)^2 + \frac{1}{2}\left(q^4 + p^4\right) \right] \tag{21}$$

where the functions $q(x)$ and $p(x)$ are canonical variables.

One can guess that equations (20) have an invariant measure given formally by

$$Z^{-1}\exp(-\beta H) \tag{22}$$

where β is an inverse temperature and Z is a normalization constant. We choose arbitrarily $\beta = 1$ and choose the data from the corresponding distribution.

The differentiated terms in the expression of H characterize the Wiener process [3], and thus the sample functions drawn from this probability measure are almost nowhere differentiable [4]. We do not wish to discuss here how such solutions are to be approximated by finite differences. We start instead with a set of finite difference equations that are formally consistent with the equations (20):

$$\begin{aligned} \frac{dq_i}{dt} &= -\frac{p_{i+1} - 2p_i + p_{i-1}}{h^2} + p_i^3 \\ \frac{dp_i}{dt} &= \frac{q_{i+1} - 2q_i + q_{i-1}}{h^2} - q_i^3 \end{aligned} \tag{23}$$

where as before $h = 1/N$ is the mesh spacing. These are Hamilton's equations of motion for the discretized Hamiltonian

$$H_N = \frac{1}{2}\sum_{i=1}^{N}\left[\left(\frac{q_{i+1} - q_i}{h}\right)^2 + \left(\frac{p_{i+1} - p_i}{h}\right)^2 + \frac{1}{2}\left(q_i^4 + p_i^4\right)\right]. \tag{24}$$

We assume that for some small enough h (large enough N) the approximation (23) is adequate, and do not concern ourselves with how it was arrived at. The system of ordinary differential equations (23) with small h is our system (1). This device allows us to avoid a discussion of the approximation of weak solutions.

We suppose that we only have computational power to compute the evolution of half of the $2N$ degrees of freedom, and we coarsen the grid: replace h by $h' = 2h$, N by $N' = N/2$, and set

$$q'_i = q_{2i}, \quad p'_i = p_{2i}, \quad i = 1, \ldots, N'. \tag{25}$$

What we want is a scheme to compute the evolution of q', p'; i.e., we want to find $Q_i^N(q', p')$, $P_i^N(q', p')$ which satisfy, as closely as possible,

$$Q_i^N(q', p') = \left\langle \frac{dq'_i}{dt} \right\rangle_{q',p'} \qquad P_i^N(q', p') = \left\langle \frac{dp'_i}{dt} \right\rangle_{q',p'}. \tag{26}$$

The naive scheme to use on the coarsened variables is the simple reapplication of the fine-scale formulas (23) on the coarser scale, i.e.

$$\begin{aligned} \frac{dq'_i}{dt} &= -\frac{p'_{i+1} - 2p'_i + p'_{i-1}}{(h')^2} + (p'_i)^3 \\ \frac{dp'_i}{dt} &= \frac{q'_{i+1} - 2q'_i + q'_{i-1}}{(h')^2} - (q'_i)^3 \end{aligned} \tag{27}$$

where $h' = 1/N' = 2h$. There is, however, a better way.

From the equations (23) one can easily see that

$$\left\langle \frac{dq'_i}{dt} \right\rangle_{q'p'} = \frac{2}{h^2} p'_i + (p'_i)^3 - \frac{1}{h^2} \left\langle p_{2i+1} + p_{2i-1} \right\rangle_{q'p'}, \tag{28}$$

where the subscripts on the brackets have the meaning obvious from the definition (4).

We may exploit our knowledge of the distribution (22) to infer the statistics of the intermediate points (q_{2i+1}, p_{2i+1}) from the known values $q'_i = q_{2i}$, $p'_i = p_{2i}$. In the absence of the quartic terms in (24), the probability distribution of intermediate points would be given by the Levy interpolation formula for the Wiener process [4]:

$$p_{i+1} = \frac{1}{2}(p_i + p_{i+2}) + W, \tag{29}$$

where W is a Gaussian variable of mean zero and variance $\sqrt{h}$, independent of either p_i or p_{i+2}. Thus we could explicitly write (28) as

$$\left\langle \frac{dq'_i}{dt} \right\rangle_{q'p'} = -2\,\frac{p'_{i+1} - 2p'_i + p'_{i-1}}{(h')^2} + (p'_i)^3. \tag{30}$$

Assume temporarily that the statistics of the intermediate points are dominated by the effect of the differentiated terms in the Hamiltonian. Then the mean error, as defined by equation (6), vanishes if

$$\begin{aligned} Q_i^N(q',p') &= -2\,\frac{p'_{i+1} - 2p'_i + p'_{i-1}}{(h')^2} + (q'_i)^3 \\ P_i^N(q',p') &= 2\,\frac{q'_{i+1} - 2q'_i + q'_{i-1}}{(h')^2} - (p'_i)^3 \end{aligned} \tag{31}$$

instead of the naive scheme (27).

The approximation no longer looks consistent with equation (20), as could have been predicted from the inapplicability of Taylor's formula on the small scales; the coefficient of 2 arose out of the fact that the Levy interpolation

formula forces one to use linear interpolation in the evaluation of second derivatives. The recursive use of this formula shows that the derivative term in the Hamiltonian is indeed dominant, as was assumed in the derivation. Note further that the new equations correspond to the Hamiltonian

$$H_{N'} = \frac{1}{2}\sum_{i=1}^{N'}\left[2\left(\frac{q'_{i+1}-q'_i}{h'}\right)^2 + 2\left(\frac{p'_{i+1}-p'_i}{h'}\right)^2 + \frac{1}{2}\left(q_i^4+p_i^4\right)\right]. \tag{32}$$

This is the Hamiltonian (24) renormalized by the change of scale from h to $h' = 2h$ [5].

Note that if the Levy formula were not available, one could instead use Metropolis sampling based on the Hamiltonian to derive the statistics on the finer grid as in the previous example.

This scheme can be tested as in the preceding section. Figure 3 shows histograms of errors for the naive scheme (27) and for the "renormalized" scheme (31). These histograms were produced using $N = 16$ and $\beta = 1$. There are 10^3 samples shown, with 10^5 constrained samples used for each sample (values of e_i have been combined for q and p and for all values of i). The RMS error E (the widths of these distributions) is reduced by a factor of approximately 14 using the renormalized scheme instead of (27). The small errors of the renormalized scheme are mostly due to numerical errors of the Monte Carlo calculations. Some additional error may be due to the neglect

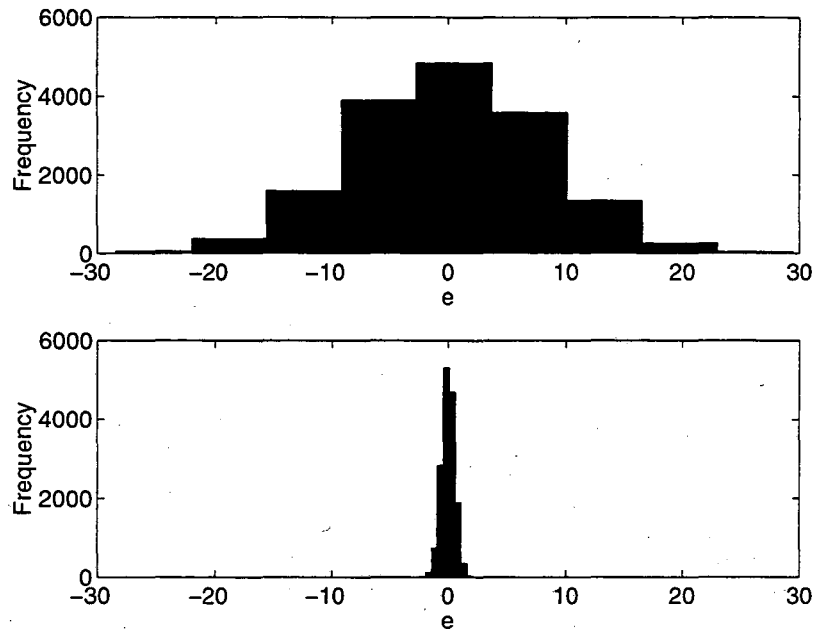


Figure 3: Histograms of errors for the naive scheme and for the renormalized scheme.

of the nonlinear terms in the interpolation formula.

Note that the same transformation (multiplying derivatives by 2) would work with more general potentials, not only the quartic potential in (21).

5 Conclusions

We have shown how to produce and evaluate schemes that calculate correctly the mean solution in underresolved approximations. This is exactly the job that must be carried out in turbulence calculations. The tool we need is a probability distribution for the missing scales; in the turbulence case, such a

probability distribution has been proposed in [6, 7].

6 Acknowledgements

The authors thank Prof. G.I. Barenblatt and Dr. Raz Kupferman for helpful conversations. This work was supported by the Applied Mathematical Sciences Subprogram of the Office of Energy Research, Department of Energy under contract number DE-AC03-76SF 00098.

References

[1] R. Peyret and T. Taylor. *Computational Methods for Fluid Flow.* Springer, New York, 1983.

[2] P. Moin S. Ghosal, T. Lund and K. Akselvoll. A dynamic localization model for large-eddy simulation of turbulent flows. *J. Fluid Mech.*, 286:229–255, 1995.

[3] R. Feynman and A. Hibbs. *Quantum Mechanics and Path Integrals.* McGraw-Hill, New York, 1965.

[4] P. Levy. *Le Mouvement Brownien.* Gauthier-Villars, Paris, 1934.

[5] T. Burkhardt and J.M.J. van Leeuwen, editors. *Real-Space Renormalization*. Springer, 1982.

[6] G.I. Barenblatt and A. J. Chorin. A near-equilibrium theory of turbulence with applications. Manuscript, UC Berkeley, 1997.

[7] G. I. Barenblatt and A.J. Chorin. Scaling laws and vanishing-viscosity limits in turbulence. *AMS Lectures in Applied Mathematics*, 1997. In press.

ERNEST ORLANDO LAWRENCE BERKELEY NATIONAL LABORATORY
ONE CYCLOTRON ROAD | BERKELEY, CALIFORNIA 94720

Prepared for the U.S. Department of Energy under Contract No. DE-AC03-76SF00098