JOURNAL OF COMPUTATIONAL PHYSICS 74, 456–476 (1988)

Estimating Limits on Resolution and Predictability in Computations of Hydrodynamically Unstable Flows

JAMES A. VIECELLI

Lawrence Livermore National Laboratory, Livermore, California 94550

Received November 7, 1986; revised April 23, 1987

Unstable hydrodynamic flows containing different materials such as Rayleigh–Taylor instability rapidly evolve to a chaotic state with limited predictability and a distribution of materials having a fractral dimension. Here it is shown how Lagrange points, either existing as part of the hydrodynamic solution method or added as passive tracers, can be used to test for self-similarity and to compute the dimension and spatial resolution scaling for the Rayleigh–Taylor instability mixing region. The efficiency of the integral or point counting methods of estimating dimension is discussed and it is shown how it can be increased by order of magnitude factors using random sampling. The problem of how to compute the rate of loss of memory of initial conditions in the transient case is discussed and a quantity similar to a largest Lyapunov exponent is defined and used to determine the predictability time for the Rayleigh–Taylor problem.

INTRODUCTION

For many types of boundary and initial conditions there are fundamental limitations on how far the Navier–Stokes equations can be integrated forward in time to predict the future flow. Unstable hydrodynamic flows containing different fluids such as Rayleigh–Taylor instability rapidly evolve to turbulence with a chaotic distribution of materials. Some general numerical methods for measuring dimension and the rate of loss of memory of initial conditions have recently become available, but ways of adapting and improving them or creating new methods for hydrodynamic computations are needed.

The theory of nonlinear dynamical systems has yielded a better quantitative understanding of the general mechanism of the transition to chaotic motion [1, 2]. In the ergodic theory there are the Lyapunov exponents to measure how rapidly the trajectories of a dynamical system diverge or converge along the principal axis in phase space, and to specify how long deterministic equations can predict the future motion. There are geometrical types of measures such as the correlation dimension, the Hausdorff or fractal dimension, and an entire spectrum of generalized dimensions that specify how the resolution of the geometrical details of the system under study improve with successively finer scales of linear measurement.

0021-9991/88 $3.00

In the transient case of unstable hydrodynamic flow a quantity similar to a largest Lyapunov exponent can provide useful information defining limits on what can be computed and what cannot, except in terms of statistical probabilities. In this situation where predictability over a limited interval is possible, but with steadily degrading accuracy, such a measure can be used to estimate the limits on predictability due to noise or uncertainty in the initial conditions.

The geometrical configuration of materials in these flows can become fractal and in these cases dimensional measurements can provide upper limits on the spatial resolution improvement that can be achieved with computing mesh refinement. In this way knowledge of dimension and predictability measures can be of help in estimating the limits to what can be learned from deterministic simulation and modeling of a given unstable hydrodynamic flow, as well as providing scaling information for the geometrical properties of the flow. In the following I show how dimension and predictability measures were obtained for the Rayleigh–Taylor instability problem, but these techniques should have application to many other material tracking and chaotic flow problems in computational hydrodynamics.

Rayleigh–Taylor instability occurs at the boundary between two fluids of different density when there is an acceleration of the fluids directed from the lighter density toward the heavier density fluid. Small perturbations of the interface grow exponentially until the flow becomes nonlinear. The early and middle stages of nonlinear growth have been studied extensively using a range of hydrodynamic computing methods [3, 4]. Eventually the flow evolves to an asymptotic stage where the two materials are mixed together in a random pattern of whorls, tendrils, and bubbles. The ability to compute the details of this mix stage, at least the statistics, are of current interest in inertial confinement fusion research, but the problem is also of general computational interest because it is typical of a large class of transient hydrodynamic flows involving the evolution of an unstable interface between different fluids and the tracking of fluid elements in flow evolving toward chaos.

When the memory of the initial conditions is decaying with time and the flow is becoming chaotic, the mixed region between the fluids can become fractal. The geometrical properties of the mixed region are often of physical importance, and one would like to know how the computational accuracy of these features scales with mesh size. Limits on resolution scaling of the geometrical details of the materials in the flow can be determined from the fractal dimension. If the linear mesh resolution is doubled in a two-dimensional problem and the fractal dimension is d, then the spatial resolution of the details of the mixed region are proportional to 2^d. Since d will usually be less than the computing mesh dimension, doubling the mesh will not result in an increase in resolution of the interface details proportional to the increase in total number of mesh points.

Dimensional computations can also provide information that can be used to compensate for the inability of the hydrodynamic computations to resolve geometrical features smaller than the mesh interval. The concept of a fractal dimension by definition means that over some range of length scales the geometry is, at the least, statistically self-similar. The method of computing the dimension

automatically provides a test for self-similarity; the distribution of mixed material must follow a power law over some range of scales. If self-similarity exits in the geometry of the flow near the lower end of the mesh resolution then there is good evidence that the geometrical properties of the flow field just below the mesh resolution are related to those just within the mesh resolution by this self-similar scaling.

Calculation of the dimension is usually not a trivial task and in some cases may not be feasible. Box counting is a straightforward method, but inefficient [5]. In the Rayleigh–Taylor problem the method could be applied by overlaying a fixed area of the mixed region with a square grid of side δ and counting the number of squares occupied by each material. If $N(\delta)$ is the number of squares occupied then the fractal dimension is given by

$$d = \lim_{N \to \infty} \lim_{\delta \to 0} \ln N / \ln(1/\delta) \tag{1}$$

Every square has to be tested for material type, and the mesh has to increase to a very fine net to get convergence. The mesh used in hydrodynamic computations is unlikely to be fine enough to get enough resolution because the materials are stretched and folded into thin filaments that can be much narrower than one mesh width. Box counting methods have been studied by using them on the solutions from small coupled sets of nonlinear differential equations with known properties such as the Lorenz equations and shown to be computationally impractical.

More efficient methods of estimating the dimension have recently become available [6]. These are based on computing distances between pairs of points defining the evolution of the solution vector in m-dimensional phase space, and then creating a plot of the number of pairs of points C^m with separation less then distance r. The slope of the plot of Ln C^m versus Ln r is the correlation dimension.

$$d = \lim_{N \to \infty} \lim_{r \to 0} \mathrm{Ln}\, C^m(r) / \ln r. \tag{2}$$

The correlation dimension is a lower bound on the fractal dimension, but in most cases the difference is quite small so that the computation is also a good estimate of the fractal dimension. The concept of dimension has been generalized to cover a complete spectrum of scaling parameters [16], but here I consider only the correlation dimension since the method of computing it can be applied to get the other dimensional measures if desired. One reason Lagrange point counting is much more efficient than box counting is that it uses only data defining the solution vector, whereas with box counting every cell in space must be examined whether it is occupied by a solution point or not. With the point counting method all pairs of the N points defining the solution are used requiring $N(N-1)/2$ distance computations and bin sorts. Therefore, computation of a dimension estimate can still be costly because the operation count goes like N^2 and because many points are generally required to get reasonably good definition of the solution trajectory—at

least several thousand and, more often, ten thousand or more points. Some speedup can be obtained with special machine programming tricks for the sorting operation [6].

The point counting method is ideally suited for Lagrangian hydrodynamic computations since the Lagrange points are the points defining the solution trajectory in physical and velocity or momentum phase space. Unfortunately, pure Lagrangian hydrodynamic computing methods cannot be used for chaotic flows because of mesh distortion and scrambling, just the situations in which dimensional computations are useful. Free Lagrange methods and some of the early combined Lagrange–Eulerian computing schemes have the advantage of carrying a large number of Lagrange points that can be used for dimension computations while avoiding the scrambling problem [8–10]. The methods discussed here can be used with any Navier–Stokes integration scheme that can continue into the chaotic regime since Lagrange points can always be added to the computations if they are not already present. In the incompressible case if the Lagrange points are restricted to a purely passive tracer role and recent higher order methods are used in solving the underlying Eulerian mesh equations, good accuracy can be obtained [11]. In the section on resolution and dimension I show how the dimension can be computed from the Lagrange points and how the computational efficiency of the dimension calculation can be increased.

Stretching and folding in phase space are necessary conditions for deterministic chaos and hydrodynamic flows evolving towards chaos show Lagrangian mesh scrambling. The breakdown of pure Lagrangian computing methods due to mesh scrambling is a good indication that the flow has limited predictability. In these cases the flow is not predictable from computation for more than a short interval except as one realization drawn from a statistical distribution of possible long time solutions that are independent of the initial conditions. We can, however, hope to get statistical data for these flows from long time hydrodynamic computations.

If only the initial transient stages of the flow are of interest, at least the early instants of the motion can be predicted deterministically, but with decreasing accuracy as time passes. In these cases knowledge of the rate of loss of memory of the initial conditions is useful as a guide as to how much of an improvement in simulation or modeling capability can be gotten from reductions in computational truncation errors, increases in computing mesh size, improvements in definition of the initial conditions, use of adaptive meshes, and other stratagems for increasing computational accuracy.

In chaotic systems the rate at which noise grows and the rate at which information is created is measured by the largest Lyapunov exponent λ_1 [7]. This can be estimated from

$$\lambda_1 = 1/(t_m - t_0) \sum_i^m \log_2[L'(t_i)/L(t_{i-1})]. \tag{3}$$

The summation is a time average over the solution integration time, and L and

L' are the beginning and ending distances between the solution trajectory and a solution for a trajectory starting from a point displaced slightly from the solution trajectory. The definition strictly applies only after transients have completely decayed away.

The largest Lyapunov exponent determines the predictability time for a dynamical system such as the Lorenz equations. If the initial conditions for the set of nonlinear equations defining the system—assumed to be one of the possible long term states of the system—are specified with p bits of precision then the evolution of the system can be predicted only for a time,

$$\tau = p/\lambda_1. \tag{4}$$

In terms of the computational accuracy of integrating the equations this means that if the machine word length is p bits, no matter how accurate the numerical integration method the solution can only be computed for a time τ before all memory of the initial conditions is lost.

Good methods for computing Lyapunov exponents from discrete time series are available that can be applied to the output from hydrodynamic computations; however, these methods and the definition of the Lyapunov exponents assume that the dynamical system under study is ergodic and that all transients have long since decayed away [7]. These conditions are likely to be appropriate for problems such as general circulation modeling of the atmosphere and oceans, but do not apply to transient problems such as the Rayleigh–Taylor instability. If we define a new quantity that is similar to the largest Lyapunov exponent, except that the average is over an ensemble of problems with small random differences in the initial conditions, then we get a statistical quantity that is applicable to the transient case. I call this σ, the predictability exponent, since by definition it measures the divergence of solutions with small random differences in initial conditions, or equivalently, the growth of noise present in the initial conditions,

$$\sigma = t^{-1} 2/N(N-1) \sum_{ij}^{N} \log_2[L_{ij}(t)/L_{ij}(0)], \tag{5}$$

where the summation is over all $N(N-1)/2$ pairs of solutions i and j separated by distance L_{ij}. Thus the average is over an ensemble of solutions evolving over a fixed time interval instead of an average of one solution over many successive time intervals.

With this definition the average distance L between any two pairs of transient solutions starting with small random differences in the initial conditions is given by

$$L = L_0 2^{\sigma t}. \tag{6}$$

This is the rate at which noise or uncertainty in the initial conditions grows with

time. Thus if the initial conditions are specified with p bits of precision then the transient evolution can be predicted only up to a time given by

$$\tau = p/\sigma. \tag{7}$$

In the section on predictability I show how the distances L_{ij} can be computed from hydrodynamic data and how to compute σ.

Rayleigh–Taylor Hydrodynamic Computations

The geometry of the Rayleigh–Taylor problem studied consists of an alternating set of layers of two liquids extending to infinity. In the hydrodynamic computations the domain is restricted to a square region 1.5 cm on a side, with periodic boundary conditions on each edge. A constant body force is applied to one of the liquids which pushes against the other, accelerating the whole mass of liquid. The mean density of the liquids is 1.0 g/cm^3, the viscosity is 0.01 g/cm/s, and the initial thickness of the forced layers is 0.3 cm. Initially the surface of the layers is roughened with a random white noise spectrum. As the forced layer accelerates from rest, the upper surface becomes unstable while the perturbations on the lower surface decay away. Only the limiting case of the Atwood number approaching zero has been studied in these computations since the focus of the work has been on computation of dimension and predictability. Because of this the Boussinesq approximation can be used and the effective body force is given by

$$f = 2\rho_0 A g, \tag{8}$$

where A is the Atwood number, g is the acceleration, and ρ_0 is the mean density.

Widely differing methods have been used to integrate the incompressible Rayleigh–Taylor problem including Marker and Cell [4], Vortex [12], boundary integral [13], and front tracking methods [17]. The dimension and predictability computations described here can be used with any Navier–Stokes integrator that is reasonably accurate; however, an absolute requirement is that the integrator must be able track the flow past the point where Lagrange mesh scrambling occurs. The Navier–Stokes integrator used in these computations is Eulerian and requires no Lagrange grid construction during evolution of the instability. Consequently, problems of mesh scrambling are avoided without sacrificing definition of the layer geometry, and, as a result, the evolution of the instability can be followed into the mixing phase for as long as desired.

This integration method can maintain good accuracy in problems such as the Rayleigh–Taylor instability, provided the Eulerian zone size is kept in the viscous scale range and provided the initial perturbation is not so small as to be quickly swamped by truncation error growth. In more quantitative terms it is desirable to

keep the mesh size smaller than the wave length of the most rapidly growing disturbance predicted from linear analysis of the viscous case,

$$\Delta x < \lambda_m = 4\pi(v^2/Ag)^{1/3} = 4\pi(2\rho v^2/f)^{1/3}, \tag{9}$$

where f is the body force per unit volume, A is the Atwood number, g is the acceleration, v is the kinematic viscosity, and λ_m is the wave length of the most rapidly growing perturbation.

There is always noise from roundoff and truncation error, no matter what the computational method, and in Rayleigh–Taylor unstable flows that noise will rapidly grow to a point where it swamps all memory of the initial conditions. One can still estimate the average growth rate of the noise; it is just necessary to start with a big enough perturbation and keep the mesh size small enough to resolve the viscous scales of motion.

The difference scheme used to integrate the Navier–Stokes equations on the Eulerian grid is similar to the method described by Kim and Moin [11]. This method uses a direct solver to obtain the solution to the Poisson equation for the pressure, and an Adams–Bashforth second-order integration step to advance the time. Kim and Moin also include an implicit treatment of the viscous terms which is not necessary for the Rayleigh–Taylor problem because the mesh spacing is uniform and the viscosity is small. Since an explicit formula for the viscosity terms is faster, it is used instead. The accuracy is further enhanced by using the Lagrange points only to determine which Eulerian cells contain the body forcing term. This is feasible because the Boussinesq approximation eliminates the need to determine a density field for the Eulerian mesh. More details of the method are given in Appendix A.

The numerical method is well known and the accuracy of computation has been documented in [11]. I have tested the method on some other problems as well. The range of nonlinear problems for which analytical solutions exist is limited, however there are a few cases, such as periodic shear flow with a transverse drift velocity, which can be solved analytically and which provide tests of differencing of both the viscous and nonlinear transport terms. The solution for this case is given in Appendix B. This case was tried with good results, an example of which is shown in Fig. 1. Symmetry tests are also useful in checking for program errors, especially problems with boundary conditions. Single mode Rayleigh–Taylor instability is a useful test problem in this respect because any small errors are rapidly amplified by the physical instability.

An example of single mode Rayleigh–Taylor intability on a length scale similar to that for the memory loss rate computations is shown in Fig. 2. This computation was done on a 60×60 Eulerian mesh with 6480 Lagrange points defining the forced layer. The layer surface is initially displaced sinusoidally with a wave length of 12 zones and a peak to peak amplitude $\frac{1}{6}$th of the layer thickness. The body force is 117.6 dynes/cm^3 and the viscosity is 0.01 g/cm/s, approximately the viscosity of water. Although there is no analytical solution available, several features can be

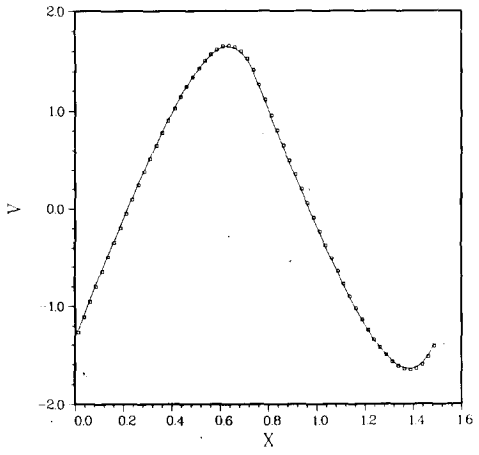

FIG. 1. Comparison of the numerical solution with the analytical solution for a shear flow driven by a spatially alternating body force. The hydrodynamic computations were done on a 60×60 Eulerian grid. The solution is skewed because there is a drift velocity imposed perpendicular to the body force. Solid curve and small squares are the analytic and numerical solutions respectively. $\nu = 0.01$ cm^2/s, $U_0 = 0.0596$ cm/s, $d = 0.75$ cm, and $F = 0.379258$ dynes/cm^3.

compared with results from recent boundary integral solutions of the inviscid problem [15]. In the Boussinesq limit the variables can be put in dimensionless form by dividing lengths by the wave length of the starting perturbation λ, and times by $\sqrt{(\lambda/2Ag)} = \sqrt{(\rho\lambda/f)}$. When this is done the main features of the solution in Fig. 2 are found be very close to those for the inviscid solution, even though the physical viscosity is not small. One characteristic of the Boussinesq limit is that the bubble and spike should have identical shapes and be symmetrical about a line through the mean position of the interface. The solution in Fig. 2 has this symmetry to within the resolution of the Lagrange point density. A second characteristic of the inviscid Boussinesq solution is that the bubble velocity increases rapidly and then levels off at a constant value of approximately 0.22 in dimensionless units. The solution in Fig. 2 reaches this asymptotic stage by $t = 0.15$, leveling off at a dimensionless bubble rise velocity of approximately 0.21 measured relative to the midline of the interface. The one significant difference between the solution in Fig. 2 and the inviscid Boussinesq result is that the singular points where the inviscid flow rolls up are completely smoothed out and removed by the physical viscosity. The RMS diffusion of momentum over the problem time interval, computed from $\sqrt{2\nu t}$ where ν is the kinematic viscosity, covers a circle of diameter approximately 0.14. This is about 6 mesh zones or $\frac{1}{2}$ the perturbation wave length and significantly larger than the rollup region seen in the boundary integral computations.

Since the evolution is chaotic, any errors introduced will grow exponentially with time. As will be shown the noise growth rate σ is approximately equal to 10 bits/s. These computations were done on a computing machine with a 48-bit mantissa so

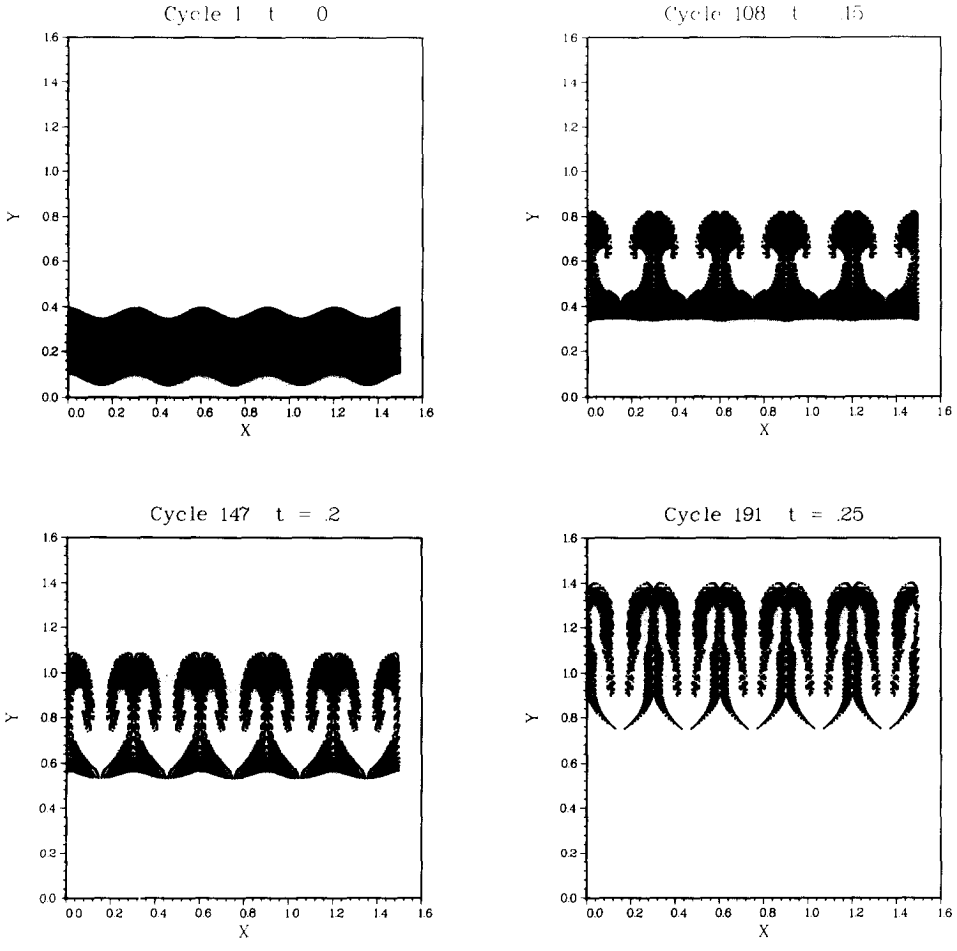

FIG. 2. Single mode Rayleigh–Taylor instability computed on a 60×60 Eulerian mesh with 6480 Lagrange points defining the forced layer Mean density of the liquids is 1 g/cm^3 and the viscosity is 0.01 poise. The body force driving the layer is 117.6 dynes/cm^3. The wave length with the most rapid growth in the linear analysis is $\lambda_m = 0.15$ cm $= 6$ zone widths.

the maximum predictability time is approximately 4.8 s. Beyond this point all memory of the initial conditions is lost. A more realistic limit is the time required for errors to grow to the point that they can be spotted on the plots. The plot resolution is on the order of 10 to 11 bits, hence the maximum time before errors begin to show on the plots can at the most be no longer than about $(48\text{-}11)/10 =$ 3.7 s. This is more than ten times farther than the last step shown in Fig. 2 even though there the single mode distortion is quite large. These test problems show that the chaotic evolution seen in the Rayleigh–Taylor problem is not due to lack of

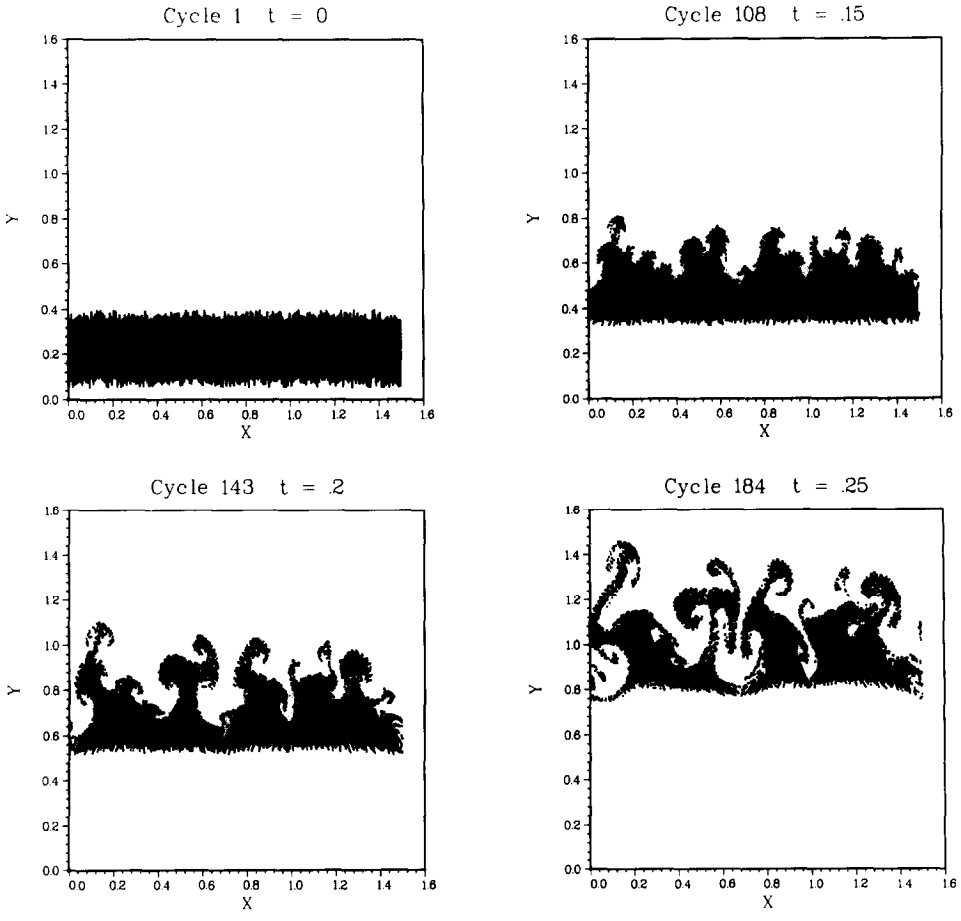

FIG. 3. Rayleigh–Taylor instability with the same physical specifications as the single mode run in Fig. 2, except that the initial condition is random white noise roughening of the surface instead of a single sine wave. The peak to peak maximum amplitude is the same as in Fig. 2. The four snapshots are at the same times.

numerical accuracy or programing errors. When the Rayleigh–Taylor runs are repeated with the same physical parameters and peak to peak perturbation amplitude except that the surface displacement is random white noise, the evolution is typically like that shown in Fig. 3.

PREDICTABILITY

In the theory of ergodic chaos Lyapunov exponents measure the average rate at which trajectories in phase space diverge or converge along the principal axes, and

the largest Lyapunov exponent tells how fast information is created in bits/s, and how fast noise is generated in the system. Lyapunov exponents are not local quantities; they give the average rate over all of the phase space covered by a given chaotic system. They apply to a system which is ergodic, in which all transients have decayed away. However, many problems of interest such as the Rayleigh–Taylor problem are transient. In these cases prediction is possible over some limited time interval, and the question of practical concern is how can we quantitatively determine what the predictability time is.

In the Introduction I defined a new quantity that is like a largest Lyapunov exponent in that it is a statistically averaged quantity measuring a change in distance between two solutions starting with small random differences in the initial conditions. This quantity which I called a predictability exponent is different from the largest Lyapunov exponent in that it is an average over an ensemble of solutions computed over the same time interval, instead of a single solution averaged over many successive time intervals. By definition it is a statistical measure of how rapidly noise is generated during the transient phase of unstable nonlinear evolution and how rapidly any two solutions with small random differences in initial conditions diverge in time.

In experiments the initial conditions are not known exactly. In the Rayleigh–Taylor problem we only have, at best, a statistical description of the roughness of the interface, but we need to estimate the evolution nevertheless. Every slightly different initial configuration for the interface is going to lead to a much different final configuration. In fact, numerical experimentation shows that each possible solution for some small random initial perturbation of the interface tends to diverge exponentially from every other solution, i.e., in Lagrange hydrodynamic computations the mesh configurations diverge. In this context the average we would be interested in would be over the range of all possible final configurations that could evolve, starting from a neighborhood of uncertainty about the initial position. If over this space we had an estimate for a quantity like the predictability exponent as I have defined it, then given a parameter describing the roughness and the predictability exponent, one would be able to determine the limits on how far the interface motion could be computed for a desired range of accuracy.

One measure of how far apart two solutions are is the root mean square of the difference in the position of the Lagrange points defining the material regions,

$$L_{ij}(t)=\sqrt{\left[\sum_{n}^{N}\{(x_i(n,t)-x_j(n,t))^2+(y_i(n,t)-y_j(n,t))^2\}\right]}, \tag{10}$$

where $x_i(n, t)$ and $y_i(n, t)$ are the coordinates of the nth Lagrange point of the ith solution at time t, and N is the total number of Lagrange points in each solution.

With the above definition for the distance between any two solutions i and j, we can compute the predictability exponent defined by Eq. (5), where the sum is over all independent pairs of trial solutions, each with a random perturbation of the initial surface roughness. If we had infinite computer time we could run all possible

initial conditions and so carry the average over all solution space. In practice we sample a small number of trial problems with random initial conditions and take the result as an approxiate estimate of the average.

I tested the measure defined by Eq.(5) by plotting the $\log_2$ of the separation between many pairs of solutions L_{ij} versus the time. If there are N solutions, each for a different seed in the random number generator for the initial surface roughness, then there are $N(N-1)/2$ independent pairs of solutions. The separation between any two pairs of solutions has a large random component, but by including many cases on the same semi-log plot, these random fluctuations can be averaged out, leaving an approximation to the statistical limit.

One example of the evolution of the instability and the breakup and rupture of the layer is shown in Fig. 3. Both the top and bottom surfaces of the layer have the same initial roughness; however, only the top surface goes unstable. The bottom surface perturbations decay in time because the computations include viscosity. All 6480 Lagrange particles defining the layer are plotted. These are the coordinates used in Eq. (10) for computing the separation between two pairs of solutions.

The separation between any two pairs of solutions does appear to have a well-defined statistical limiting curve. This can be seen in Fig. 4, where the $\log_2$ of the separation between the fifteen possible pairs of solutions for the case where the body force is 117.6 dynes/cm^3 are plotted. The dead space at the beginning of the runs (where $L_{ij}(t)$ is just equal to the initial separation due to a different seed in the initial roughness generator) is a numerical artifact; before anything can happen the layer has to move a sufficient distance for the initial roughness to show up as a perturbation on the underlying Eulerian mesh. Once the instability starts, the separation between solutions $L_{ij}(t)$ grows rapidly—in these examples at a rate

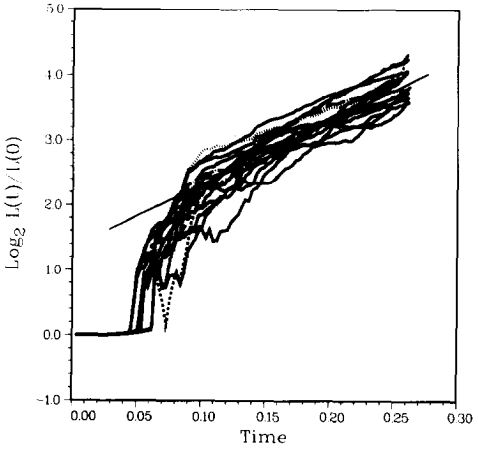

FIG. 4. The separation between pairs of solutions of the Rayleigh–Taylor problem plotted versus time. The separation between two solutions is defined as the square root of the sums of the squares of the differences in coordinates of each of the 6480 Lagrange points defining the layer position. The delay at the start is the time needed for the layer to move one Eulerian mesh width.

of about 125 bits/s—corresponding to the rate of growth derived from the linearized Rayleigh–Taylor theory for the smallest wavelength perturbation that can fit into the Eulerian mesh. This phase lasts only a short time interval and results in a noise growth of about 1.9 bits. After this, $\log_2[L(t)/L(0)]$ approaches a straight line asymptote with a slope of about 10 bits/s. The average slope is the graphical representation of σ, the predictability exponent defined by Eq. (5). Reference to Fig. 3 shows that most of the mixing leading to rupture of the layer takes place during this nonlinear asymptotic phase.

Similar results are obtained when the force is increased to 940.8 dynes/cm^3, corresponding to shortening the wavelength of most rapid growth (in the viscous linear analysis) to 3 zones, or 0.075 cm. The time scale is shortened by a factor of $\sqrt{8}$ and the asymptotic slope increases to 28 bits/s.

The results of the computations can be put in dimensionless form by normalizing the time by

$$t^* = \sqrt{(\rho d/f)} = \sqrt{(d/2Ag)}, \tag{11}$$

where d is the thickness of the layer.

It turns out that the behavior of all of the runs accumulated in this study is summarized by a single $L(t)$ curve if the time is measured in units of t^*. After adjusting the time origin to the takeoff point in Fig. 4 the dimensionless curve of separation between random pairs of solutions is, to a good approximation, given by

$$\log_2[L(t)/L(0)] = 0.5t + 1.9[1 - e^{-3t}]. \tag{12}$$

Equation (12) only predicts the behavior in the sense that it approximates the most likely outcome for the rate of growth of information. Any individual numerical model run or experiment is likely to deviate significantly from Eq. (12) as can be seen from the variation in the runs shown in Fig. 4.

The results summarized in Fig. 4 and by Eqs. (5) and (12) show that any uncertainty in the specification of the initial conditions grows exponentially with time. The average rate of growth can be used to estimate an upper limit on the predictability of computations by dividing the number of bits in the machine word length by the predictability exponent as in Eq. (7).

Resolution and Dimension

It is also of interest to be able to determine limits on the spatial resolution of the Rayleigh–Taylor mixed region as a function of computing mesh refinement. The hydrodynamic runs show an initially roughened layer evolving into a chaotic set of whorls, tendrils, and bubbles. An upper limit on how much the resolution of the details of these shapes can increase with finer mesh spacing is given by the fractal

dimension. An improvement in the method of computing the dimension and results for the Rayleigh–Taylor runs and tests with the Lorenz equations are described in the following.

The method [6] discussed in the Introduction is to count the number of pairs of points C^m closer than distance r in m-dimensional phase space and then plot $\ln C^m$ vs $\ln r$. The slope of the curve is the correlation dimension and a lower bound on the fractal dimension. After doing the computations for a low dimensional example like the Rayleigh–Taylor problem it is easy to see that the sorting and counting of the distances between all pairs of points is equivalent to using each point for the center of an m-dimensional sphere to generate a curve $C^m(r)$ for that point, and then adding all of the individual curves together.

In the two-dimensional hydrodynamic cases, $m=2$, the spheres are circles, the points are the Lagrange points defining the mixed layer, and the correlation dimension is close to the fractal dimension. Two limiting cases illustrate the method. If there is a well-defined thick layer such as the beginning of the two-dimensional Rayleigh–Taylor runs the slope will be $d=2$ because everywhere the number of Lagrange points defining the layer and lying within a circle of radius r will increase propertional to r^2. If instead of a layer there were a thin line of points winding over the grid, the method would yield $d=1$ since the number of points inside a circle of radius r will, on average, be propertional to r. For the 60×60 mesh runs there are 6480 circles with 6480 points to count for each circle. If the method is implemented as described in [6] there will be 2.1×10^7 distance computations and bin sorts. When the same problem is run with a 120×120 Eulerian mesh and 25,920 Lagrange points the number of distance computations and bin sorts required in the point counting method is approximately 3.4×10^8. Even with special attention to programing efficiency using masking and bit shifting operations for the bin sorting, the computations can still be costly; they are more so in hydrodynamic applications because one may want to make many dimension measurements.

The two-dimensional Rayleigh–Taylor problem is a good test problem for measuring dimension because it is easy to see how the computation works. In graphical terms the idea is to take each Lagrange point as the center of a set of concentric circles forming a set of annular bins around the central point and to count the number of points in these bins. Taking a single point for the center and computing a curve of $\ln C^2$ versus $\ln r$ for this point can give erratic results. With reference to the snapshot of Rayleigh–Taylor mix in Fig. 3 it is easy to see why. If the center point is in the middle of a large chunk of the layer, the point count will go like the square of the radius. If the center point falls on one of the thin linear tendrils then the point count will tend towards proportionally to the first power. What is needed is an average count over the entire set of points—that is what the correlation integral method produces. However, this relatively simple case where the method can be easily visualized suggests that it is not necessary to do a radial count at every point. Instead it should be possible to get a reliable average from a much smaller sample of the total population of points provided this sample is chosen randomly to ensure that it is representative of the entire set.

In these Rayleigh–Taylor problems it is possible to get good statistical convergence with a much smaller sample of points for the centers of the circles if the locations of the sample points are chosen randomly from the full set. In these cases good convergence can be obtained with about 200 random sample centers requiring 200×6480 bin sorting operations versus $6480^2/2$ for the full treatment on a 60×60 mesh, and 200×25920 versus $25920^2/2$ for a 120×120 mesh. The number of bins used does not affect convergence, so it is only necessary to use enough to get a good measurement of the slope of the $\ln C^m(r)$ vs $\ln r$ curve. Fifty bins were used in these computations.

The slope of the curve of $\ln C^m$ vs $\ln r$ usually varies monotonically from infinity to zero with some near-constant stretches along the way. Breaks in the slope define the length scales over which the dimension applies. The slope approaches zero in the limit of large r because the finite computing region begins to look like a single point. In the opposite limit of small r the curve goes to minus infinity because there is a smallest distance separating the nearest pair of Lagrange points.

Figure 5 shows a snapshot taken from a Rayleigh–Taylor run computed on a 120×120 mesh with 25,920 Lagrange points. This picture is at a late stage of mixing where the layer has dispersed into a cloud of material entrained with the ambient liquid. In this case the viscosity, body force, and density ratio are such that the wavelength of most rapid growth in the linearized analysis is six times the hydrodynamic computing mesh width of 0.0125 cm. The run is at a small Atwood number such that the effective body force acceleration $2Ag = 188.2$ cm/s^2. The dimension computations for this picture are shown in Fig. 6. Curves of $\ln C^2$ vs $\ln r$ are plotted for from 1 to 200 randomly drawn center points. In this particular set of

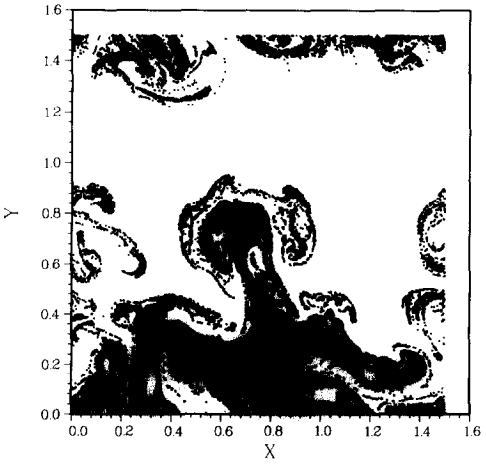

FIG. 5. Late time snapshot of Rayleigh–Taylor instability when the layer has dispersed into a cloud of material entrained with the ambient liquid. This computation was done on a 120×120 mesh with $\Delta x = \Delta y = 0.0125$ cm. The body force is 940.8 dynes/cm^3 and the wave length of most rapid growth in the linear analysis is 0.075 cm = 6 zone widths; time = 0.17 s.

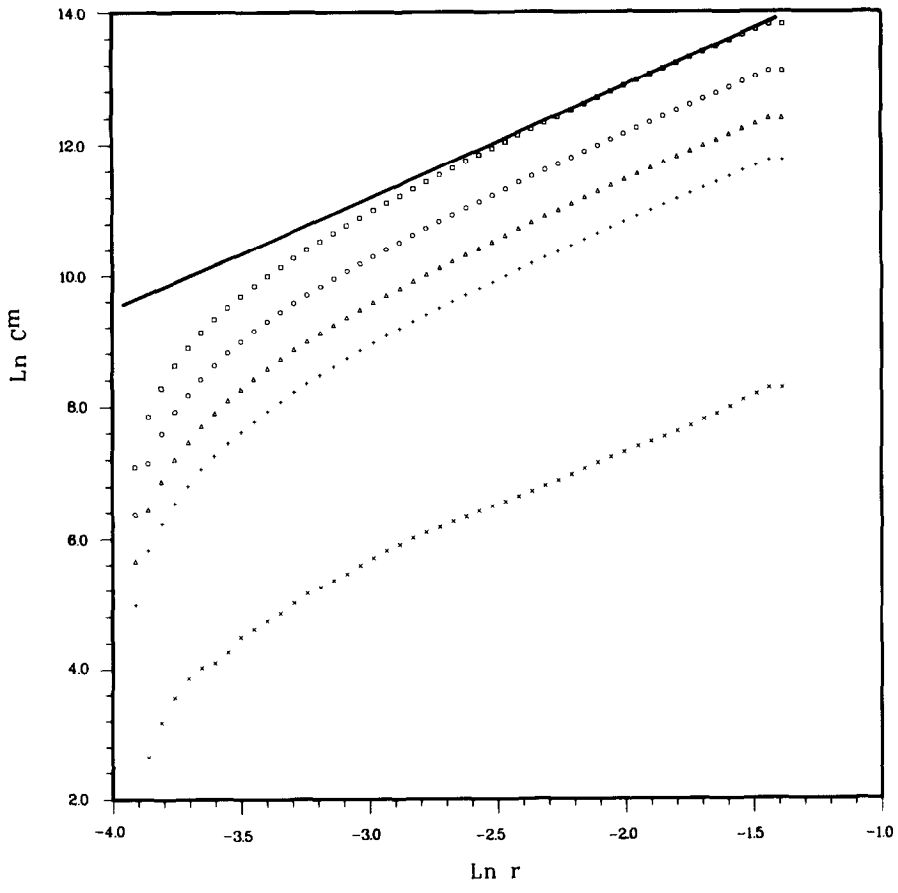

FIG. 6. Plot of the $\log_e$ of the point count C^2 versus the $\log_e$ of the distance from a randomly chosen point for the Lagrange points in Fig. 5. The lower curve denoted by the x's is the point count for a single randomly chosen center point. The curves denoted by the +'s, triangles, circles, and squares are the results of using 25, 50, 100, and 200 randomly chosen centers for the radial count. The right edge corresponds to a circle of diameter 0.5 cm in Fig. 5 or about the scale of the larger features.

trials the curve based on a single sampling is surprizingly good. The right side of the figure corresponds to a circle of diameter 0.5 cm in Fig. 5, and the far left side to a circle with diameter of about 0.04 cm. Thus the slope of the curves at the right side corresponds to the dimension of the larger features in Fig. 5. There the slope of the curves is in the range of 1.7 to 1.8.

Constant slope regions in the curve of $\ln C^m(r)$ vs $\ln r$ indicate self similarity, in the sense that if the details of the flow field inside circles with different radii within the range of constant slope are magnified to the same scale then the patterns will, on average, be identical. In the Rayleigh–Taylor cases the dimension computed

gives information on the scaling of the computational mesh resolution. When the computational mesh size is increased from 60×60 Eulerian mesh zones and 180×36 Lagrange points defining the layer to 120×120 Eulerian zones and 360×72 Lagrange points, the linear mesh resolution is doubled, but the increase in resolution of the boundary details of the whorls, tendrils, and bubbles in the mixed region is only $2^{1.8} = 3.5x$. Although there are $4x$ more mesh points we get less than $4x$ improvement in definition of the mixed layer because of the stretching and twisting of part of the material into thin filaments.

The dimension scaling limit on spatial resolution is a purely geometrical effect. It tells how the resolution changes assuming that everything else including the velocity field remains the same when the mesh size is doubled. In a given run the dimension

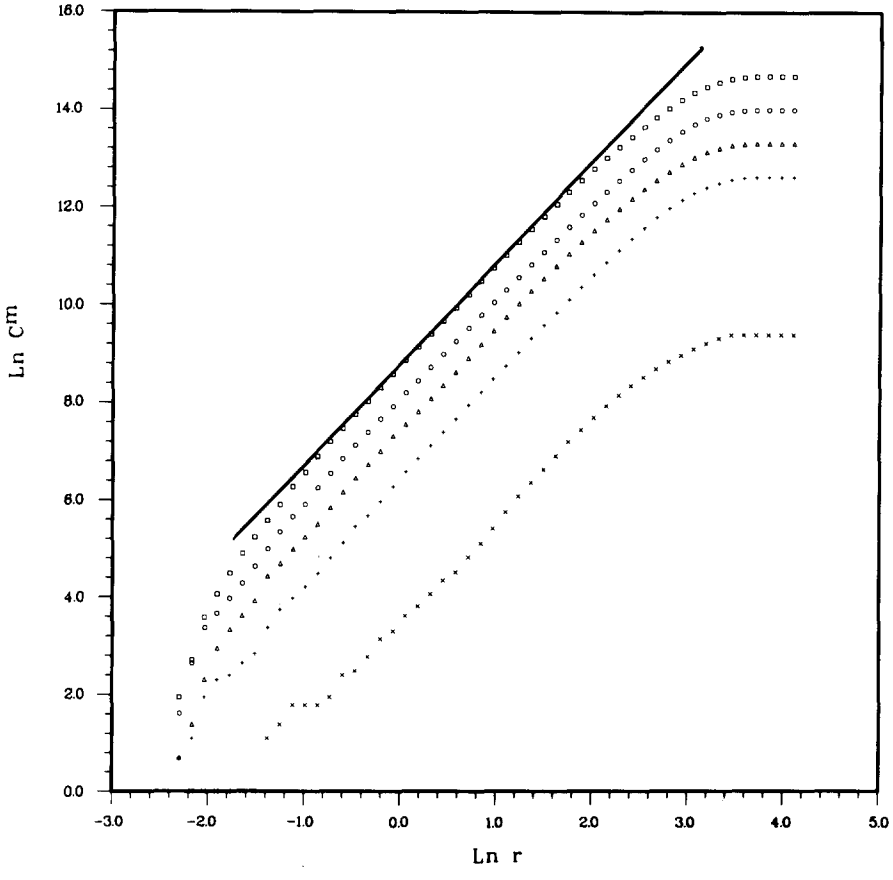

FIG. 7. Plot of $\ln C^3$ versus $\ln r$ for the Lorenz attractor described by Eqs. (14)–(16). The x's, +'s, triangles, circles, and squares denote the curves for random sample sizes of 1, 25, 50, 100, and 200 center points. Over distances in the range $\ln r$ between -1 and 1.5, the slope of the curve for sample size 200 is 2.07 ± 0.01.

changes over time as the mixing of the different fluids progresses. When a 120×120 run is compared with the same run on a 60×60 mesh the curve of dimension versus time is slightly different because the resolution of the velocity field is also changing.

I have tested random sampling on other problems with good results. Figure 7 shows statistical convergence of the $\ln C^3$ vs $\ln r$ curves on the Lorenz equations [14] for the case specified by

$$dx/dt = 10(y - x) \tag{14}$$

$$dy/dt = x(28 - z) - y \tag{15}$$

$$dz/dt = xy - (8/3)\, z. \tag{16}$$

These equations were integrated over a time interval of 240 with data output at intervals of 0.02 yielding 12,000 data points defining the solution. Transients were eliminated from this data by starting with output from a previous integration over an interval long enough for them to decay away. The dimension of the Lorenz attractor has previously been determined to be in the range of 2.06 to 2.07 [6, 7]. The slope over the range $-1 < \ln r < 1.5$ in Fig. 7 is 2.07 with an uncertainty of about ± 0.01 for the 200 random sample case. Above a distance interval of $2e^{1.5}$ the slope starts to decay to zero because that is approximately the maximum width of any cut through the attractor. Although the statistics for small sample sizes are poor, the sample size needed for accurate results is quite small compared to the amount of data in the time series. This example shows that it is not necessary to do a point count at every point in the solution to make accurate correlation dimensional measurements. The amount of computation required for the 200 random sample case is reduced approximately sixty fold over the full correlation integral.

Conclusions

The computations have shown that Lagrange points can be used to measure distance between pairs of solutions to the Rayleigh–Taylor instability problem starting with small random differences in initial conditions and that this information can be used to estimate the rate of loss of memory of initial conditions as the unstable flow evolves toward turbulence. Predictability can be defined in terms of an exponent with the physical dimensions of bits/unit time. This quantity is a statistical measure of the exponential rate of growth of noise during the transient nonlinear phase of the unstable evolution to chaos. The definition resembles that of the largest Lyapunov exponent in an ergodic system except that in the transient case the average is over the same time interval for an ensemble of runs differing by small random perturbations of the initial conditions. Given the accuracy with which the initial conditions are known, which in the case of the Rayleigh–Taylor instability is the uncertainty in the initial interface position, an estimate of the predictability time

is given by the precision with which the initial conditions are specified, in bits, divided by the predictability exponent as defined here. Experiments with the Rayleigh–Taylor instability problem show that as little as six independent runs are sufficient to get a good estimate of this quantity.

Lagrange points can also be used to compute generalized dimensional measures during the late stages of Rayleigh–Taylor mixing. That in turn yields an upper limit on how much of an improvement in resolution of the geometrical details of the mixed region can be obtained with an increase in the size of the hydrodynamic computing mesh. The point counting method of estimating dimension can be used with the Lagrange data. In this method the distances between all pairs of points defining the solution are computed and sorted yielding a count of the numbers of pairs within a given distance or radius. The correlation dimension is the slope of the log of the count plotted versus the log of the radius. Other generalized dimensions can be obtained by the point count method if desired. The Rayleigh–Taylor problem is a good example for easily visualizing how point counting works and has been useful in revealing a way to speed up dimension computations by orders of magnitude. The increase is computational efficiency can be obtained by using a small random sampling of fixed center points for computing the radial distribution function instead of computing the distances between all pairs of points. Since the accuracy of sampling depends on the sample size rather than on the total number of points, this reduces the operations count from proportionality to N^2 for the full correlation integral to N for random sampling. On the Rayleigh–Taylor test problem containing 25,920 Lagrange points, the time required to obtain good statistical convergence is reduced by more than a factor of 100.

Any hydrodynamic material tracking problem that has a positive predictability [illegible] scrambling. Thus measurement of dimension and predictability are feasible only with hydrodynamic computing methods that can work in such situations. Even though the problems to which these methods are applied may be fundamentally beyond deterministic predictability; over short time intervals at least some predictive capability is possible.

APPENDIX A: Hydrodynamic Difference Equations

The numerical integration method is essentially that described in [11] except that the viscous terms are explicit in time. Implicit differencing is not needed in these problems because the viscosity is relatively small and the mesh is uniform. An Adams–Bashforth method is used for the time integration of the Navier–Stokes equations,

$$U_i^* = U_i^n + (\Delta t/2)(3H_i^n - H_i^{n-1}) + (1/R)\,\nabla^2 U_i^n \tag{A1}$$

$$U_i^{n+1} = U_i^* - \Delta t\,\partial_{x_i}\phi^{n+1}, \tag{A2}$$

where R is the Reynolds number and

$$H_i = -\partial_{x_j} U_i U_j + f_i. \tag{A3}$$

The body force term f_i is zero unless adjoining zones contain Lagrange points; otherwise it is set equal to the input value for f_y. The average density of points in the adjoining pair of zones is computed and if it is less than the initial point density then the body force is reduced proportionately. Velocity values at the Lagrange points are interpolated from the Eulerian mesh values using area weighting of the four nearest mesh points as in [8]. New point coordinates are obtained by integrating these velocity components with the Adams–Bashforth formula used in Eq. (A1).

The pressure function ϕ is obtained from a direct solution, by cyclic reduction, of the standard five point difference approximation to the Poisson equation

$$\nabla^2 \phi^{n+1} = (1/\Delta t)\, \partial_{x_i} U_i^*. \tag{A4}$$

Periodic boundary conditions are used in both the x and y coordinates. With these boundary conditions ϕ is indeterminate to within an additive constant, since only the gradient of the pressure is used in the solution. The constant is chosen to make the average of ϕ equal to zero. The finite difference form of the velocity divergence in each zone is computed and checked at each cycle. It should always be of order of the machine word roundoff.

The variable placement is the same as in [11] with ϕ located at the centers of the Eulerian zones, with the U velocity components at the left and right side of the zones, and the V velocity components at the top and bottom sides of the zones. Uniform zoning with $\Delta x = \Delta y$ was used in all of the runs. Five-point differencing is used for the Laplacian in the viscous terms, and second-order centered spatial differencing is used for the convective terms. For example,

$$(\partial_x U^2)_k \simeq (U_p^2 - U_m^2)/\Delta x, \tag{A5}$$

where

$$U_p = \tfrac{1}{2}(U_{k+1} + U_k) \tag{A6}$$

and

$$U_m = \tfrac{1}{2}(U_k + U_{k-1}). \tag{A7}$$

All of the computations reported on here were done with a constant time step chosen at the start to be small enough to stay within the Courant condition limit for the duration of the run. The time step can be changed if necessary by interpolating a new value for H^{n-1} at one new time step interval backward in time, but it is better to avoid this if possible.

APPENDIX B: ANALYTIC SOLUTION FOR SHEAR FLOW

The Navier–Stokes equations can be solved by Fourier series methods for the special case of steady shear flow driven by a spatially alternating body force. When a steady cross flow is included, the solution for the velocity component parallel to the driving force is a skewed profile like that shown in Fig. 1. This laminar flow pattern produces a balance between momentum transport by the cross flow and momentum transport by molecular viscosity. If U_0 is the cross flow velocity parallel to the x axis, F is body force parallel to the y axis with a square wave dependence on x periodic in length $2d$, and ν is the kinematic viscosity, then the solution for the flow velocity parallel to the force F is

$$V(x) = \sum_{n=1,3,5,\ldots,\infty} v_n \sin(n\pi x/d + \beta_n), \tag{B1}$$

where

$$\tan \beta_n = -U_0 d/\nu\pi n \tag{B2}$$

and

$$v_n = (4F/\nu\pi n)[(U_0 n\pi/\nu d)\sin \beta_n - (n\pi/d)^2 \cos \beta_n]^{-1}. \tag{B3}$$

ACKNOWLEDGMENT

This work was performed for the U.S. Department of Energy under Contract W-7405-ENG-48.

REFERENCES

1. J.-P. ECKMANN AND D. RUELLE, *Rev. Mod. Phys.* **57**, 617 (1985).
2. P. CVITANOVIC, *Universality in Chaos* (Adam Hilger, Bristol, 1984).
3. D. L. YOUNGS, *Physica D* **12**, 32 (1984).
4. J. DALY, *Phys. Fluids* **10**, 297 (1967).
5. H. S. GREENSIDE, A. WOLF, J. SWIFT, AND T. PIGNATARO, *Phys. Rev. A* **25**, 3453 (1982).
6. P. GRASSBERGER AND I. PROCACCIA, *Physica D*, **9**, 189 (1983).
7. A. WOLF, J. B. SWIFT, H. L. SWINNEY, AND J. A. VASTANO, *Physica D* **16**, 285 (1985).
8. F. H. HARLOW AND J. E. WELCH, *Phys. Fluids* **8**, 2182 (1965).
9. J. A. VIECELLI, *J. Comput. Phys.* **8**, 119 (1971).
10. J. U. BRACKBILL AND H. M. RUPPEL, *J. Comput. Phys.* **65**, 314 (1986).
11. J. KIM AND P. MOIN, *J. Comput. Phys.* **59**, 308 (1985).
12. G. R. BAKER, D. I. MEIRON, AND S. A. ORSZAG, *Phys. Fluids* **23**, 1485 (1980).
13. M. S. LONGUET-HIGGINS AND E. D. COKELET, *Proc. R. Soc. London Ser. A* **364**, 1 (1978).
14. E. N. LORENZ, *J. Atmos. Sci.* **20**, 130 (1963).
15. R. M. KERR, Lawrence Livermore National Laboratory Report No. UCID-20915, 1986; *J. Comput. Phys.*, in press.
16. T. C. HALSEY, M. H. JENSEN, L. P. KADANOFF, I. PROCACCIA, AND B. I. SHRAIMAN, *Phys. Rev. A* **33**, 1141 (1986).
17. J. GLIMM, O. MCBRYAN, AND D. H. SHARP, *SIAM J. Sci. Statist. Comput.* **7**, 230 (1986).