JOURNAL OF COMPUTATIONAL PHYSICS 119, 231–243 (1995)

Interpolation Schemes for Three-Dimensional Velocity Fields from Scattered Data Using Taylor Expansions

NADEEM A. MALIK* AND TH. DRACOS

Institute of Hydromechanics and Water Resources Management, The Swiss Federal Institute of Technology, ETH Hönggerberg, CH-8093 Zürich, Switzerland

Received March 17, 1994; revised December 5, 1994

We present a numerical scheme that interpolates field data given at randomly distributed locations within a three-dimensional volume to any arbitrary set of points within that volume. The approximation scheme uses local trivariate polynomial interpolants and it is shown to be equivalent to a Taylor expansion of the (velocity) field up to second-order partial derivatives. It is formally a third-order scheme in the (mean) spacing of the data δ; i.e., the errors scale with $(\delta/\lambda)^3$, where λ is the length scale of the flow field. The scheme yields the three-dimensional velocity field (which can be inhomogeneous and anisotropic) and all the 27 first- and second-order partial (spatial) derivatives of the velocity field. It is compared with the adaptive Gaussian window method and shown to be considerably more accurate. The interpolation scheme is local in the sense that it interpolates the data within locally defined volumes defined as the set of points with the same nearest neighbours (which may be set at between 10 and 15 in number). This makes the scheme formally discontinuous in the flow field across neighbouring patches; but by making use of the excess data within a local volume, it is shown that for practical purposes the scheme does yield a continuous flow field throughout the entire interpolation volume. The scheme interpolates the data by an iterative method which is extremely fast in situations where a certain level of error bounds in the data (and, hence, also the solution) is acceptable. Results from sinusoidal and stochastic (turbulent) test flow fields show that the Taylor expansion scheme is widely applicable and highly accurate for the velocity and first derivatives. However, the smallest scale of the (velocity) field λ must be greater than 5δ for the best performance. Second-order derivatives are less accurate. Flow quantities such as the fractal dimension of streamlines can be obtained accurately with much lower data density. Statistics like the power spectrum of the flow can also be obtained accurately. In the presence of noise in the velocity data, small levels of noise have negligible effect on the obtained velocities and a modest effect on the first derivatives. The second derivatives are seriously affected, and only those of the largest scales in a turbulent flow can be adequately resolved.

1. INTRODUCTION

The importance of the Lagrangian picture of particle motion in flows has been appreciated (at least in the context of dispersion phenomena), since Talyor's [18] seminal work on random walk models, but it is only in the last decade or so that a number of research teams have begun to obtain the Lagrangian motion of particles in fluid flows. Computational calculation of Lagrangian quantities using direct numerical simulations (DNS) is very recent, beginning with Yeung and Pope [21] and later by Squires and Eaton [17]. Experimental work began only slightly earlier when suitable cameras became commercially available so that the accurate measurement of particle motion was made possible. See Maas *et al.* [12] and Malik *et al.* [14] (herein after referred to collectively as MM93) and references therein for a summary of recent Lagrangian measurement techniques, known as particle tracking velocimetry (PTV). PTV techniques typically yield sets of particle coordinates within a flow volume V over a large number of consecutive time frames. In the three-dimensional PTV of MM93, the cameras are mounted on a moving carriage so that long particle trajectories can be obtained for up to 1500 particles seeded in the flow.

An important technical problem emerges from PTV measurements—to approximate the randomly distributed velocity field data on to a Cartesian grid. This is necessary when Eulerian quantities are desired and for visualisations of the flow structures, because most graphics packages take regularly spaced data as input.

Because measurements in a Lagrangian frame are relatively new, the problem of interpolation from scattered data on to a regular grid has not received much attention until recently. An important method for two-dimensional problems was introduced by Hardy [9] which uses bivariate quadratic surfaces of different curvatures called multiquadrics. Many workers have found this method to be satisfactory especially in the context of geophysical problems [10]. In geophysical problems the scheme works well probably because the quadratic eigenfunctions are a good leading order approximation to the profile of hills and mountains. (A closely related type of interpolating function is the thin-plate spline [8]. In fluid mechanics, Imaichi and Ohmi [11] describe a two-dimensional interpolation technique based upon an expansion of the velocity field to first-order partial derivatives; their application to vortex flow met with adequate success. This scheme can be extended to any

* Current address: Laboratoire de Mecanique des Fluides et d'Acoustique, Ecole Centrale de Lyon, B.P. 163-69131, Ecully, France.

0021-9991/95 $12.00

number of dimensions and to any order in partial derivatives, although its effectiveness decreases rapidly with increasing order. Agui and Jimenez [2] (AJ86) compared a number of simple techniques, including polynomial interpolators and kriging, but found little advantage over the adaptive Gaussian window (AGW). Kriging is a method that has become popular in the field of mining. It is a linear scheme based upon the autocorrelation of the dependent variable (which may be the concentration of a substance in rocks) which is calculated from the local data values; see [1, Chap. 10].) The AGW can also be extended to any number of dimensions, but is not suitable for obtaining higher than first-order partial derivatives (as discussed later in this paper). However, it is simple and very fast to compute. Recently Spedding and Rignot [16] have developed a two-dimensional thin-shell spline method (which is a variation of the cubic spline) which they compare with the AGW of AJ86 and showed that the former yields results which are twice as accurate within a suitable choice of parameters.

The numerical work of Yeung and Pope [20] [YP88] is closely related to the present problem. They simulate a three-dimensional turbulent flow field using direct numerical simulations (DNS) of the Navier–Stokes equations and obtain the Lagrangian movement (trajectories) of particles in the flow from the information given at the nodes of the regular DNS grid—this is the inverse of the problem addressed in this paper. They present a number of different interpolation schemes, including fourth-order cubic splines and third-order Taylor expansion schemes. (The latter is similar to the idea of Imaichi and Ohmi [11] extended to three dimensions.) The splines were the most accurate, but the Taylor schemes also performed well and are considerably less intensive to compute.

The trivariate polynomial scheme that we have developed is shown to be equivalent to a Taylor expansion scheme. These schemes are local in nature in that they use data from within a neighbourhood of the point at which the interpolation is desired.

We compare our results with the performance of the AGW which seems to have become the standard benchmark for these purposes. The motivation for investigating more advanced schemes stems from the obvious failure of linear schemes like AGW to give acceptably small errors in the interpolated quantities. AJ86 report errors of around 60% in a sinusoidal test flow field, even when the wavelength of the flow field is five times larger than the spacing of the scattered data. With the availability of sophisticated instruments and more powerful computers, it is anticipated that more research effort will be put into Lagrangian properties of flows and it is important to make available accurate interpolation schemes and demonstrate the possibilities and limitations of such schemes.

In Section 2 we describe the technical details of the trivariate polynomial (Taylor) scheme and the AGW method against which we compare the results. In Section 3 the simulation results against two test functions are described; the simple sine wave and the multiple-scales kinematic simulation (turbulent) flow field. In Section 4 we look at the problem of continuity of the interpolated flow field and test our method by computing streamlines, fractal dimensions, and spectra in complex flows. We conclude in Section 5 with an illustration from results from a PTV experiment in a water channel.

2. INTERPOLATION SCHEMES

We assume that the coordinates and velocities of seeded particles, say from a PTV experiment, are given and accurate so that we are interested in the errors due to the interpolation schemes only. (A brief discussion of the reduction in computational cost when high accuracy is not required is contained towards the end of Section 2.2.) The velocity field need not be homogeneous or isotropic since the method is a *local* scheme; interpolation is performed within small local volumes containing a limited number of data points.

Consider a volume $V = [x_{10} - x_{1m}, x_{20} - x_{2m}, x_{30} - x_{3m}]$ within which we are given the velocities at N_v (different) points; i.e., we are given $\{\mathbf{x}\} = \{(\mathbf{x}^n, \mathbf{u}^n) = ((x_1^n, x_2^n, x_3^n), (u_1(\mathbf{x}^n), u_2(\mathbf{x}^n), u_3(\mathbf{x}^n)))\,|\,\mathbf{x}^n \in V, n = 1, ..., N_v\}$. The (mean) spacing of the data is $\delta = (V/N_v)^{1/3}$. The mathematical problem is to estimate the velocities and the 27 first and second (spatial) partial derivatives of the velocity field at an arbitrary set of M (≥ 1) locations $\{\mathbf{x}^*\} \neq \{\mathbf{x}\}$ within V. In particular, we desire to interpolate the data on to a cubic grid of nodes with regular spacing of δ_{1g}, δ_{2g}, δ_{3g}. Without loss of generality we will take N_g nodes on each side so that $M = N_g^3$ and we will take $\delta_{1g} = \delta_{2g} = \delta_{3g} = \delta_g = \delta$.

2.1. *Linear Adaptive Gaussian Window*

For the simple linear adaptive Gaussian window (AGW) the velocity at a point $\mathbf{x}^* \in V$ is given by (AJ86),

$$u_i(\mathbf{x}^*) = \frac{\sum_{n=1}^{N_v} \alpha^n u_i^n}{\sum_{n=1}^{N_v} \alpha^n}, \quad i = 1, 2, 3, \tag{1}$$

where the u^n are the velocities at $\mathbf{x}^n$. The coefficients α^n are given by

$$\alpha^n = \exp\left(\frac{-(\mathbf{x}^* - \mathbf{x}^n)^2}{H^2}\right), \tag{2}$$

where H is the size of the Gaussian window. AJ86 defined two non-dimensional quantities based on the spacing δ, namely $H^* = H/\delta$ and $\lambda^* = \lambda/\delta$, where λ is the length scale of the flow field, say the wavelength if the flow is a sinusoidal velocity field. In more complex flows λ must be estimated from the physics of the flow; for example, in turbulent flows λ would correspond to the Taylor microscale. AJ86 report that the optimum window width was $H = 1.24\delta$.

In the present work, we have used the above AGW scheme as a benchmark to compare with our trivariate scheme. (We

use a slight modification of (1) in that we do not sum over all n particles in the volume V, but only over those particles, $\mathbf{x}^n$, which are within a correlation length of $\mathbf{x}^*$, where $|\mathbf{x}^n - \mathbf{x}^*| \leq \lambda$. This speeds up the computation and we have found that this does not alter the results because velocities further away from $\mathbf{x}^*$ than a correlation length have no significant effect on the velocity at $\mathbf{x}^*$ by definition.)

2.2. *Trivariate Polynomial (Taylor) Schemes*

There are a number of important differences (summarised in [5, 6]) between univariate interpolation and multivariate interpolation which involves approximating functions of more than one real or complex variables. Practically, though, the most important feature is that multivariate approximations are considerably more expensive to compute—if there are p independent variables and one desires to approximate a function at say q points in each variable, this gives a total of q^p values to compute.

The special case of interpolation on a Cartesian grid presents no special conceptual difficulty because the tensor product of n univariate interpolation operators is an n-variate interpolation operator. Specialising to polynomials, it can be shown that trivariate interpolation on a three-dimensional Cartesian grid

$$\{(r_i, s_j, t_k): 0 \leq i \leq n, 0 \leq j \leq m, 0 \leq k \leq l)\} \tag{3}$$

is *uniquely* possible by a polynomial of the form

$$p(r, s, t) = \sum_{\nu=0}^{n} \sum_{\mu=0}^{m} \sum_{\eta=0}^{l} a_{\nu\mu\eta} r^\nu s^\mu t^\eta; \tag{4}$$

the polynomial subspace that this tensor product involves is $\Pi_n \times \Pi_m \times \Pi_l$, where Π_n is the linear space of all polynomials of degree less than or equal to n. The elements of this subspace are finite sums of the form

$$p(r, s, t) = \sum u_i(r) v_j(s) w_k(t) \quad (u_i \in \Pi_n, v_j \in \Pi_m, w_k \in \Pi_l); \tag{5}$$

these elements are representable as in Eq. (4), and by the general theory of tensor products the dimension of this space is $(l + 1)(m + 1)(n + 1)$. Hence, the uniqueness of the representation in Eq. (4) follows, since the interpolation on a grid of $(l + 1)(m + 1)(n + 1)$ nodes is certainly possible (see [15]).

For polynomial interpolation at n_p nodes not regularly distributed in R^3, a system of linear equations is nonsingular with probability 1. (The case where there are no exploitable patterns in the nodes is called the case of "scattered data.") One can use a trivariate scheme and select a set of n_p monomials $r^\nu s^\mu t^\eta$ and solve for the coefficients in the equation

$$\sum a_{\nu\mu\eta} r_e^\nu s_e^\mu t_e^\eta = \lambda_e, \quad 1 \leq e \leq n_p. \tag{6}$$

For the scattered data problem at hand, where we are given scattered velocity field data, suppose that we want to approximate the velocity $\mathbf{u}$ at a point $\mathbf{x}^* = (x_1^*, x_2^*, x_3^*)$. Then let us define local variables $r = \delta x_1 = (x_1 - x_1^*)$, $s = \delta x_2 = (x_2 - x_2^*)$, $t = \delta x_3 = (x_3 - x_3^*)$. Then the trivariate interpolation scheme represented by (6) for scattered data in three dimensions, using polynomial interpolators from $\Pi_2(R^3)$, is equivalent to the local Taylor expansion of the velocity field about the point $\mathbf{x}^*$ up to second-order derivatives, viz. (for the ith component of the velocity)

$$u_i(\mathbf{x}) = u_i(\mathbf{x}^*) + \sum_j \delta x_j \left(\frac{\partial u_i}{\partial x_j}\right)^* + \sum_{k \leq j}{}' \sum_j \delta x_j \, \delta x_k \left(\frac{\partial^2 u_i}{\partial x_j \, \partial x_k}\right)^* + \mathcal{R}_i, \quad i = 1, 2, 3, \tag{7}$$

where we identify $\lambda_e = (u_i)_e$. The $'$ on the summation is a reminder that there is a factor of 2 in the summation whenever $j \neq k$ due to the symmetry of the terms in these cases. The remainder terms $\mathcal{R}_i$ are formerly proportional to $O|\delta\mathbf{x}|^3$ which makes the above expansion a third-order scheme; in non-dimensional units, $\mathcal{R}_i$ is proportional to $(\delta x/\lambda)^3$. The star on the derivative terms indicates that they are evaluated at $\mathbf{x}^*$. The coefficients $a_{\nu\mu\eta}$ are equal to the various partial derivatives (up to second) of the velocity at $\mathbf{x}^*$.

We now introduce a new set of notations in which each derivative term is regarded as an independent quantity. Let $\boldsymbol{\xi}(\mathbf{x})$ represent a 30-component vector whose entries are as follows: the first three components are the velocity components at $\mathbf{x}$, viz. $(\xi_1, \xi_2, \xi_3) = (u_1, u_2, u_3)$; the next nine terms are the nine first-order partial derivatives of the velocity at $\mathbf{x}$, viz. $(\xi_4, \xi_5, \xi_6, \ldots) = (\partial u_1/\partial x_1, \partial u_2/\partial x_1, \partial u_3/\partial x_1, \partial u_1/\partial x_2, \ldots)$; the next 18 components are the second-order partial derivatives of the velocity field at the point $\mathbf{x}$, viz. $(\xi_{13}, \xi_{14}, \xi_{15}, \xi_{16}, \xi_{17}, \ldots) = (\partial^2 u_1/\partial x_1^2, \partial^2 u_2/\partial x_1^2, \partial^2 u_3/\partial x_1^2, \partial^2 u_1/\partial x_1 \partial x_2, \partial u_2/\partial x_1 \partial x_2, \ldots)$.

In this notation, (7) can be expressed in the simple form (Einstein's summation convention for repeated variables implied)

$$u_i(\mathbf{x}) = a_{ij} \xi_j^* + \mathcal{R}_i, \quad i = 1, 2, 3. \tag{8}$$

a_{ij} is different from that used in Eq. (6); a_{ij} is now a matrix with entries equal to the monomials $\delta x_1^\nu \delta x_2^\mu \delta x_3^\eta$ with $0 \leq \nu + \mu + \eta \leq 2$; and $\boldsymbol{\xi}^* = \boldsymbol{\xi}(\mathbf{x}^*)$. In order to evaluate the 30 independent variables in $\boldsymbol{\xi}$, we need 30 conditions. This is supplied by the 10 *nearest* data points at which we are given $(\mathbf{x}, \mathbf{u})$. Hence (ignoring the remainder terms),

$$u_i^n = a_{ij}^n \boldsymbol{\xi}^*, \quad i = 1, 2, 3; n = 1, 2, \ldots, 10,$$

where n is an index to indicate the nth particle; a^n is thus a 3×30 order matrix. If we bring all components of the velocities

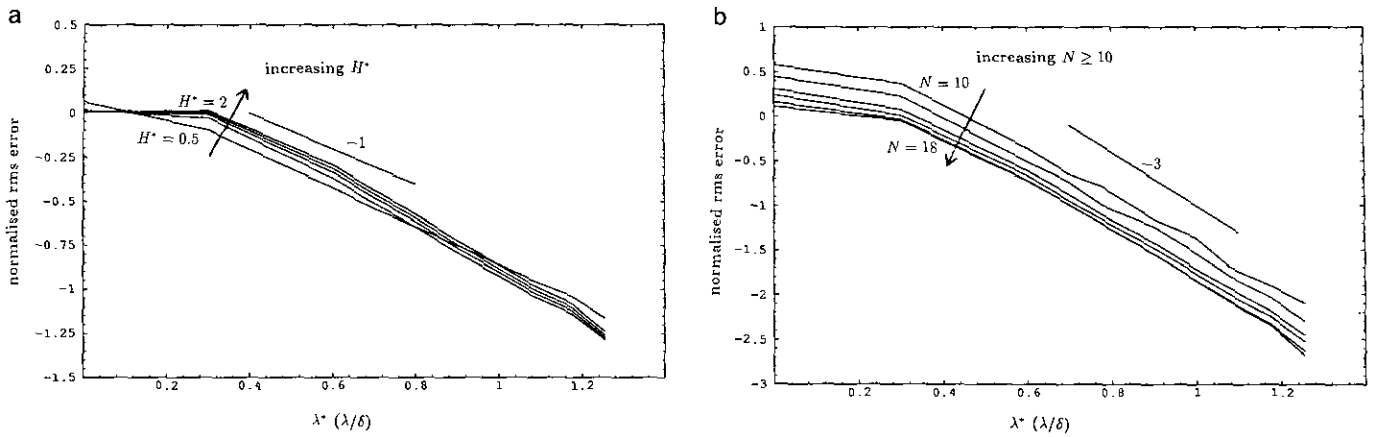

FIG. 1. (a) Log–log plot of rms errors in the velocity, normalised by the rms velocity u', from AGW against $\lambda^* = \lambda/\delta$ for different values of $H^* = H/\delta = 0.5, 1, 1.24, 1.5, 2$. Test flow is a sinusoidal wave with wavelength λ. H is the size of the Gaussian window, and δ is the mean data spacing. A line of slope -1 is shown for comparison. (b) Log–log plot as in (a), but for the Taylor schemes—TN10, TN11, TN12, TN13, TN15, TN18. A line of slope -3 is shown for comparison.

at the 10 nearest data points into a single (ordered) vector $\mathcal{U}$, and, similarly, if all the coefficients a^n are brought into one large matrix $\mathcal{A}$, then this equation can be expressed elegantly as

$$\mathcal{U}_i = \mathcal{A}_{ij}\xi_j^*, \quad i,j = 1, 2, ..., 30. \tag{9}$$

Thus, the mathematical problem posed by Eqs. (3) and (4) for the 30 unknown components of the velocity and its first two derivatives is reduced to a Taylor expansion of the velocity for each of the 10 nearest data points. We need to solve (9), which is a system of 30 independent simultaneous linear equations in 30 unknowns, and it can be solved exactly by standard methods. In fact, we can smooth the velocity field by making use of the excess data that lie within a correlation length of $\mathbf{x}^*$ by extending the number of nearest data points at which (9) is simultaneously solved and a least square solution can then be found. If N ($\geq$10) is the number of data points that are used in solving (9), then $\mathcal{A}$ is a $3N \times 30$ order matrix.

In general, an $n \times n$ system of linear equations requires of the order of $n^3/3$ operations to solve using Gaussian elimination. However, $\mathcal{A}$ is a sparse matrix because two-thirds of the coefficients are zero and it is sometimes more efficient to solve such systems of equations by iterative methods (see [3, Chap. 8.6]). If the number of iterations needed to solve (9) is m, then the number of operations needed is (approximately) of order mn^2, and this method is more efficient than Gaussian elimination provided $m < n/3$. If the exact solution is desired then $m \approx N/3$. If maximum accuracy is not required, then $m < n/3$ and considerable time may be saved. For example, if the relative error (significant figures) in the velocity or the coordinates is smaller than 10^{-4} then from the Taylor scheme TN15, with $N = 15$ nearest data points ($n = 3N = 45$), we have found that $m \approx n/3 = 15$ iterations on average. (We have labeled the Taylor schemes according to the number of data points N ($\geq$10) used in the local interpolation. Thus TN10 is the Taylor scheme using $N = 10$ nearest data points, etc.) However, if the relative

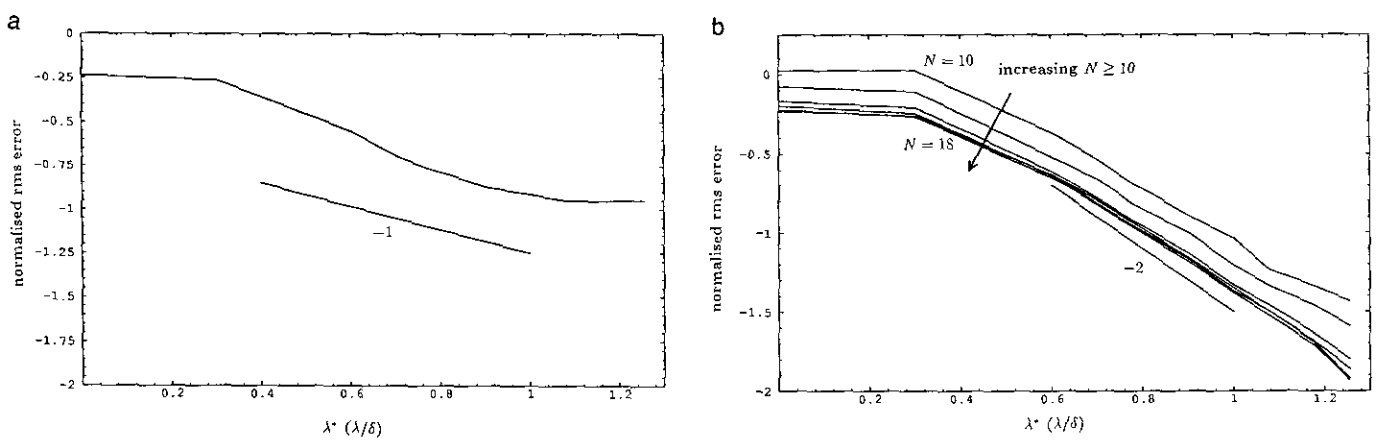

FIG. 2. (a) Log–log plot as in Fig. 1a, but for the first derivatives and for $H^* = 1$. A line of slope -1 is shown for comparison. (b) Log–log plot as in Fig. 1b, but for the first derivatives. A line of slope -2 is shown for comparison.

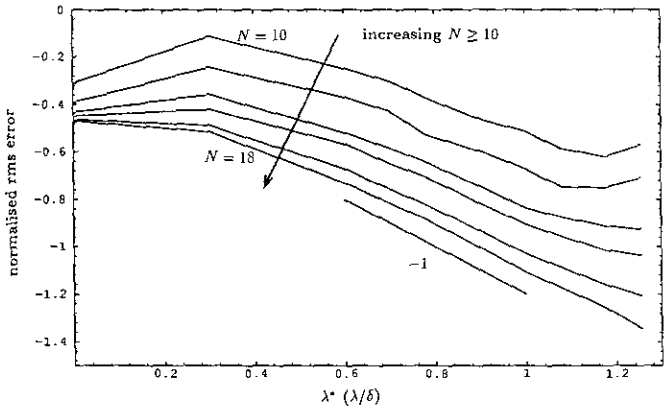

FIG. 3. Log–log plot as in Fig. 1b, but for the second derivatives. A line of slope −1 is shown for comparison.

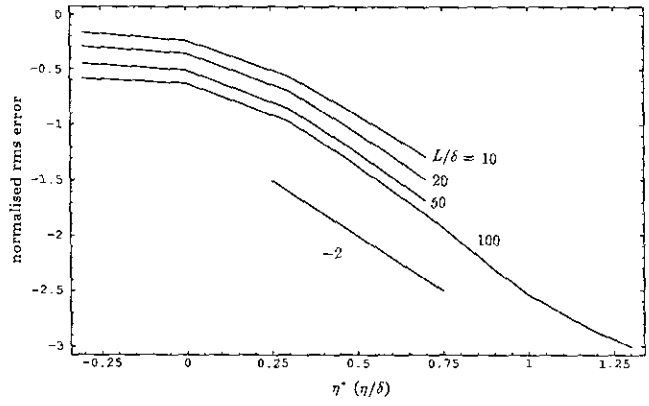

FIG. 5. Log–log plot of normalised rms error curves in the velocity for the stochastic (kinematic simulation) fields against the normalised Kolmogorov scale, using the TN15 scheme. $L = 2\pi/k_1$ and $\eta = 2\pi/k_\eta$. Different values of L/δ as indicated.

errors are, 10^{-3}, 10^{-2}, 10^{-1} then we obtained, respectively, $m \approx$ 8, 2, and 1, respectively. In absolute terms, to give some idea of performance on current machines, the following was obtained on a SUN SPARC 2 station: for a total of 1000 data points in a unit volume, interpolated on to a $10 \times 10 \times 10$ Cartesian grid, the TN10, TN13, and TN15 schemes took, respectively 33s, 45s, and 50s of CPU for a relative accuracy of 10^{-4}. For a relative accuracy of 10^{-1} they took, respectively, 8.7s, 10.5s, and 11s of CPU. These times also include the time for sorting the data near each point of interpolation which is the same for each of the different TN schemes. Nevertheless, a speed up factor of about 5 was obtained from the TN15 scheme for the 10% error level over the case with (almost) no errors. In the present work, to show the limits of the interpolation scheme, we desire the maximum accuracy, so $m \approx n/3 = N$.

Finally, in this section we note that the Taylor scheme presented here is equivalent to YP88 in the special case in which the "scattered" data are forced to be at the nodes of a grid such as in the DNS, (except that YP88 also expanded to one extra third-order term in (7), which is not required here).

3. TEST FUNCTIONS

3.1. *Sine Wave*

We have tested the AGW schemes and the Taylor schemes on two important classes of test flow fields. The first is a single length scale sinusoidal velocity field, similar to that used by AJ86, and the second is a stochastic flow field, similar to turbulence (see Section 3.2).

The sinusoidal flow is given by

$$\mathbf{u}(\mathbf{x}) = 2\mathbf{A}\sin(\boldsymbol{\kappa}\cdot\mathbf{x}). \tag{10}$$

We define $k = |\boldsymbol{\kappa}|$ and the wavelength as $\lambda = 2\pi/k$. The rms

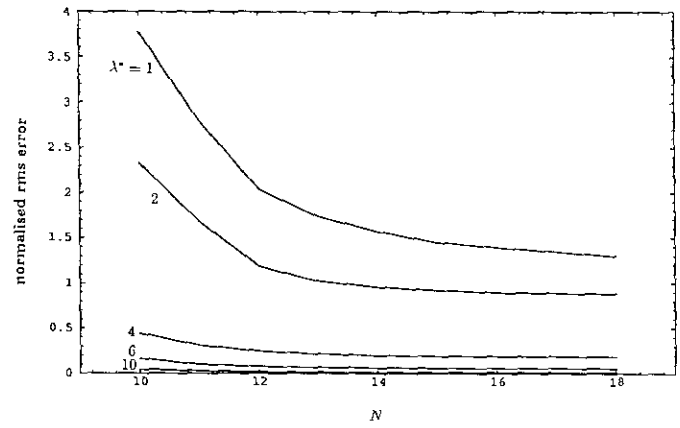

FIG. 4. The rms error in the velocity, normalised by the rms velocity u', against N, the number of points used in the Taylor schemes. Different values of the λ^* as indicated.

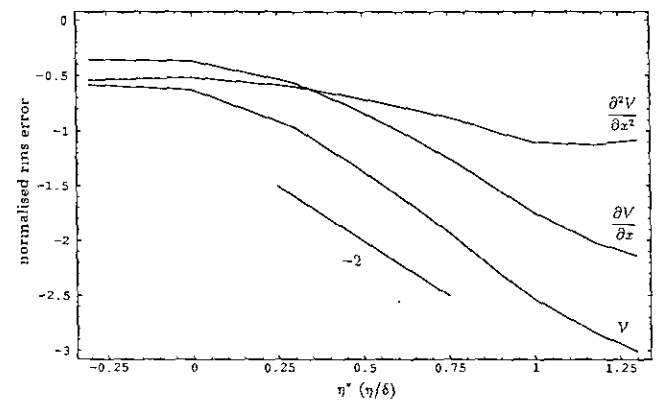

FIG. 6. Log–log plot of the normalised rms error curves in the velocity, and first and second derivatives against the normalised Kolmogorov scale. From the stochastic test flow with $L/\delta = 100$ and using the TN15 scheme.

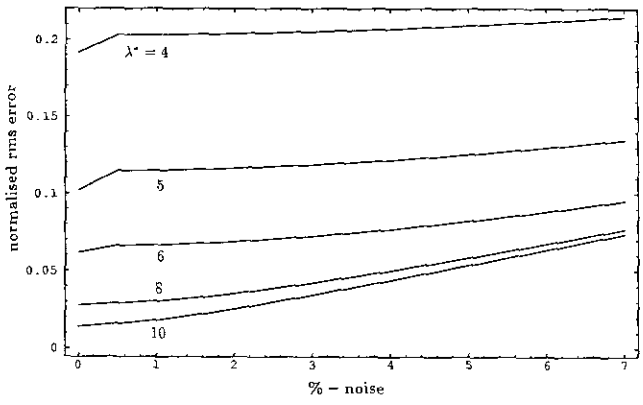

FIG. 7. *Rms error in the velocity, normalised by the rms velocity u', against the percentage of noise in u', using the TN15 scheme on sinusoidal test flows. Different values of λ^* as indicated.*

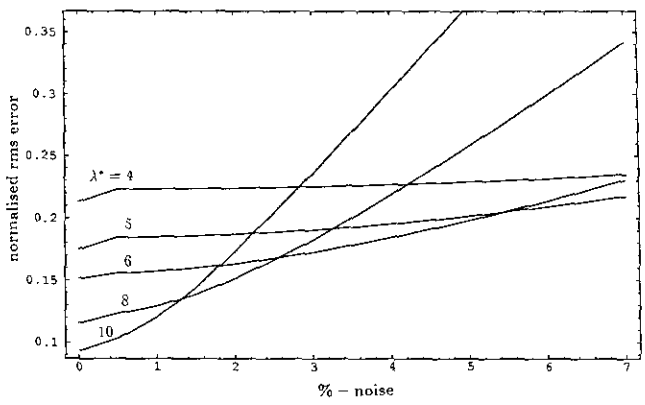

FIG. 9. As Fig. 7, but for the second derivatives.

power (velocity) of this flow is $|A|$ which is chosen to be equal to unity.

To test the accuracy of the interpolation schemes, we take the velocities at randomly located points uniformly distributed within a cubic test volume V of unit dimensions as the input data. In each simulation we generated 15^3 random points, uniformly distributed, within the volume V, which gives a spacing of $\delta = (V/15^3)^{1/3} = 1/15$. We then interpolated the data on to a regular grid of spacing equal to δ. The rms errors were evaluated between the interpolated and true quantities (obtained from Eq. (10)) at the *inner* grid points which do not lie on the outer surface of the volume (because these are extrapolations and can be serious in error), which gives an ensemble of $M = 14^3 = 2744$ points.

The rms errors are normalised by the rms power of the corresponding signals, and they are presented as functions of the non-dimensional parameter $\lambda^* = \lambda/\delta$. This quantity is a measure of the density of the data relative to the length scale of the flow; sparse data corresponds to low values of λ^* and dense data gives large values for λ^*.

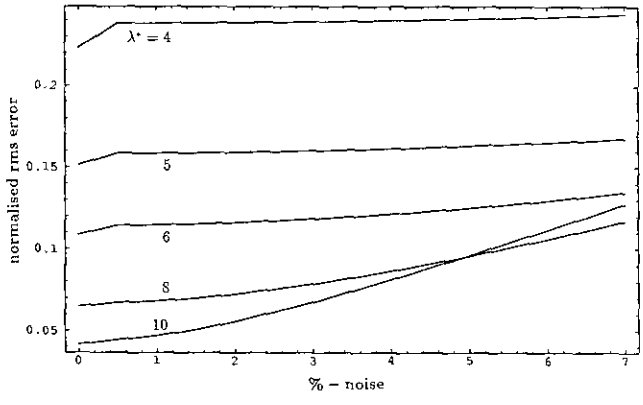

FIG. 8. As Fig. 7, but for the first derivatives.

In the case of the AGW, we have also re-examined schemes for different sizes of the window $H^* = H/\delta$. Figure 1a shows the log–log plot of the normalised rms errors in the velocity from the AGW schemes against λ^* for different values of H^*. The schemes scale linearly with λ^* as expected. However, we cannot confirm AJ86's result that $H^* = 1.24$ is the optimum value that minimises the interpolation errors. Our results show that $H^* = 1$ gives more accurate results, although the difference is small. $H^* = 1$ does indeed seem to be a more natural value to take. The behaviour for $H^* = 0.5$ is anomalous; for small values of λ^* it is slightly more accurate than the other values, but deteriorates for large values of λ^*. This is because such a small value of H^* brings the Gaussian scheme close to the behaviour of a nearest neighbour scheme.

Applying the Taylor schemes to the same problem as above, we obtain Fig. 1b for the velocity. In Fig. 1b we compare TN10, TN11, TN12, TN13, TN15, and TN18. The schemes are third order in λ^* and the curves asymptote between TN13 and TN15. But the inclusion of just a single extra data point produces a striking reduction in the error between TN10, TN11, and TN12—the error is roughly halved. This is probably due to the large factor increase in the number of sets of solutions which are smoothed to obtain the least squares estimate: from TN10 we get one (unique) set of solutions; from TN11 we smooth over 11 sets of values (a factor increase of 11 over TN10); from TN12 we smooth over 66 sets of values (a factor increase of 6 over TN11). Also, there is diminishing accuracy as the next included data point is further away from the point of interpolation. We will adopt TN15 as the standard scheme.

In both Figs. 1a and 1b, the error curves flatten off for small $\lambda^* \leq 2.5$ ($\log(\lambda^*) \leq 0.4$). In this limit the Taylor series expansion, Eq. (7), is not accurate because the non-dimensional data spacing δ/λ is large. AGW and TN schemes produce similar errors in this limit in excess of 100%, so interpolation is not efficient when the spacing of the data is of the same order as the length scale of the flow. Efficient interpolation can only be performed for cases when $\delta < 0.4\lambda$.

For 10% rms error in the velocity, the AGW scheme

($H^* = 1$) requires a data density corresponding to $\lambda^* \approx 10$ ($\log(\lambda^*) \approx 1$), while the TN15 scheme requires only $\lambda^* \approx 5$ ($\log(\lambda^*) \approx 0.7$). For 1% error, the AGW scheme requires unacceptably high data density of around $\lambda^* \approx 50$ ($\log(\lambda^*) \approx 1.7$), whereas the TN15 gives this accuracy with $\lambda^* \approx 10$ ($\log(\lambda^*) \approx 1$).

Figure 2a shows the log–log error curve for the first derivative from the AGW scheme which is equal to the average of the rms errors in $\partial u_1/\partial x_1$, $\partial u_2/\partial x_2$, and $\partial u_3/\partial x_3$ normalised by the rms power of the derivative which is equal to $k|A| = k$. The curve follows no particular pattern. For large λ^* the curve flattens because in this limit the interpolation scheme is zeroth order which means that the error does not depend on the spacing of the data values. The derivatives are obtained by finite difference of the interpolated velocity, viz.

$$\frac{\partial u_i}{\partial x_j} \approx \frac{u_i(x_j + \delta x_j) - u_i(x_j - \delta x_j)}{2\delta x_j}, \tag{11}$$

where $\delta x_j = \delta$ is the grid spacing of the nodes at which the velocity is calculated. The curve flattens off at about the 10% error level and the AGW scheme cannot yield better accuracy than this.

Figure 2b shows the corresponding error curves from the TN schemes. In these schemes, the velocity derivatives are evaluated directly as part of the interpolation scheme. The curves show a λ^{*-2} behaviour, as we expect. The 10% error level is achieved when $\lambda^* \approx 6$ ($\log(\lambda^*) \approx 0.78$); 3%, when $\lambda^* \approx 10$ ($\log(\lambda^*) \approx 1$); and 1%, when $\lambda^* \approx 22$ ($\log(\lambda^*) \approx 1.34$). Figure 3 shows the corresponding results for the second derivatives, normalised by $k^2|A| = k^2$. We have from TN15 10% error at around $\lambda^* \approx 10$ ($\log(\lambda^*) \approx 1$) and 3%, at around $\lambda^* \approx 25$ ($\log(\lambda^*) \approx 1.4$). One percent accuracy in the second derivative requires enormous density of data. (The second derivatives can also be obtained in principle from the AGW scheme using finite differences, similar to (11), but they are very inaccurate.)

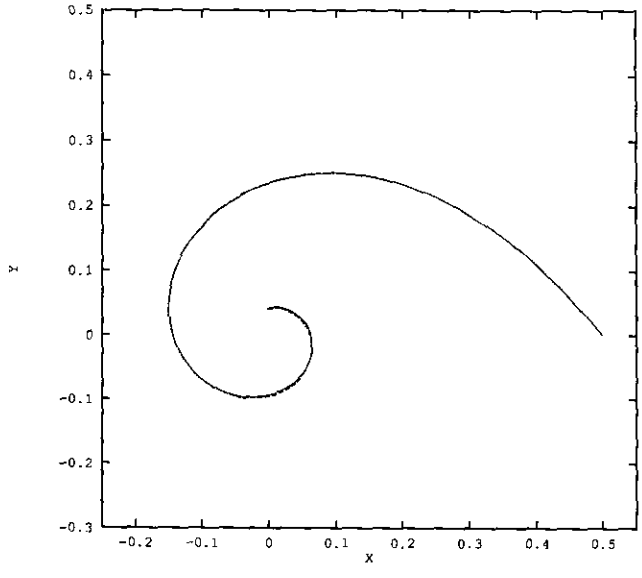

FIG. 10. Streamline of a Burgers vortex tube in the $z = 0$ plane; exact (solid line), interpolated using TN15 (dashed). Vortex parameters are: the rate of strain $a = 1$, circulation $\gamma = 1$, and the kinematic viscosity $\nu = 0.02$. We have $R_c/\delta = 3.2$, where R_c is the vortex core radius.

These results set stringent conditions for the practical applications of interpolation schemes; to obtain acceptable accuracy in the velocity and first derivatives, and qualitative accuracy in the second derivative, one needs to have a spacing of data on average at least 5 times smaller than the wavelength of the velocity field. In other words, if we want to resolve the Kolmogorov scale η, we must seed the flow with particles spaced on average at a distance of 0.2η. However, the situation is not quite as severe as it appears, and in Section 4 we will see that some important quantities can be obtained with good accuracy from interpolation with relatively low data spacing.

The final result in this section is to test the Taylor schemes as the number of neighbouring data points N is varied. Figure 4 shows the error curves in the velocity from the Taylor schemes against $N \geq 10$ for different λ^*. The curves are for $\lambda^* = 1$, 2, 4, 6, 10. Except for $\lambda^* = 1$ all the other curves asymptote at or slightly before $N = 15$. $N = 12$ or 13 are also adequate if accuracy is not the main consideration. This is the optimum number of data points to use in a Taylor scheme. This knowledge can be used to seed the flow with enough particles so that there is almost certainly at least 15 seeded particles within a correlation neighbourhood. It does not help to increase N much more than 15, because the scheme sorts the particles in order of distance from an interpolation point and the extra particles will be further and further away from the interpolation point and there will be diminishing improvement in accuracy as N is increased further.

The TN schemes introduce no extra source of error like the one introduced by the choice of a window in the AGW. Maximum use is made of the raw information and the TN scheme is probably an optimum scheme for the scattered data problem without exploitable pattern. The only way to improve would be if more knowledge is available about the instantaneous flow pattern (which is difficult to obtain in turbulent flows). For example, in separate tests on simple laminar and vortex flows, the Taylor scheme works excellently, achieving accuracies which are orders of magnitude better than that obtained from sinusoidal flows or stochastic flows (described next).

3.2. *Stochastic Turbulent Fields*

The previous section saw the testing of the TN schemes with sinusoidal flows. In the context of turbulence, we deal with stochastic flow fields with a wide range of length scales and it is important to quantify the interpolation errors that we can expect in these types of flow fields.

We have therefore performed a second set of tests on stochas-

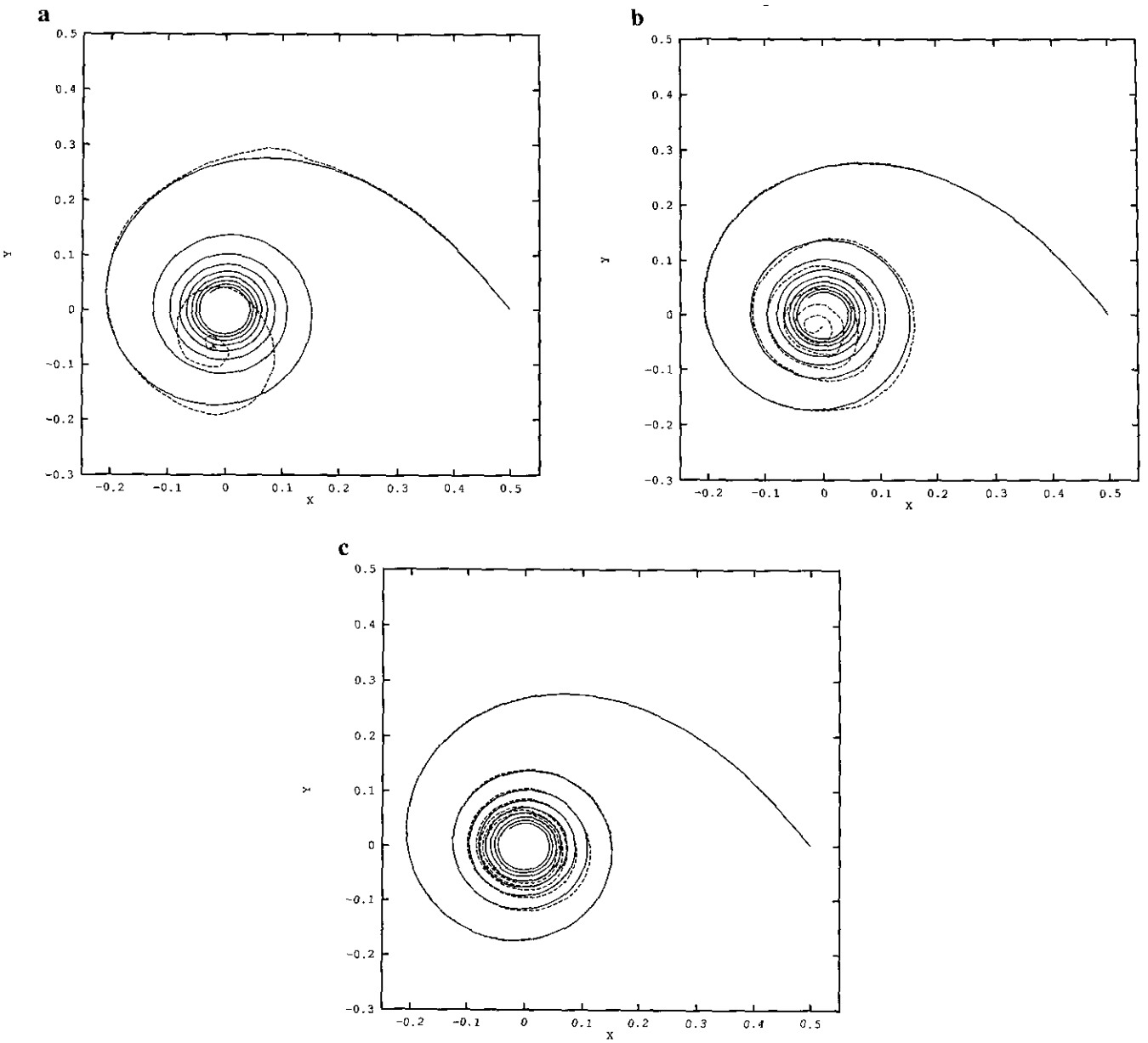

FIG. 11. Streamlines as in Fig. 10, except for $\nu = 0.00125$ and for different data densities: R_c/δ = (a) 0.5; (b) 0.8; (c) 1.13; (d) 2.25. The dashed lines are the interpolated streamlines and the solid line is the exact streamline. (e) The log–log plot of $N(\varepsilon)$, the number of boxes of size ε that cover the interpolated streamlines obtained in (a)–(d), against the box size ε. The solid curve is the case for the exact streamline. The fractal dimension (Kolmogorov capacity) D_k is the negative slope of this graph, which for the Burgers vortex is equal to 4/3 as indicated by a line with this (negative) slope. (The graphs from (c) and (d) nearly overlap.)

tic flow fields generated by kinematic simulations (KS; [7]). KS generates stochastic fields (in any number of dimensions) with a prescribed wavenumber-frequency spectrum $E(k, \omega)$. In high Reynolds number turbulence the wavenumber part of the spectrum is proportional to $k^{-5/3}$ over a wide range of wavenumbers, $k_1 \leq k \leq k_\eta$. This can be simulated using the KS inertial model [13]. An important feature of KS is that flow structures, such as vortical zones, straining zones, and irrotational zones are present in the flow field. The flow field thus retains spatial and temporal correlations which can be matched approximately to those observed in turbulent flows up to second-order two-point statistics, provided that the random Fourier coefficients and the wavenumber-frequency spectrum are chosen appropriately (described in detail in [7]).

Of particular importance is to see how the interpolation errors behave as the smallest and largest scales of the flow are varied,

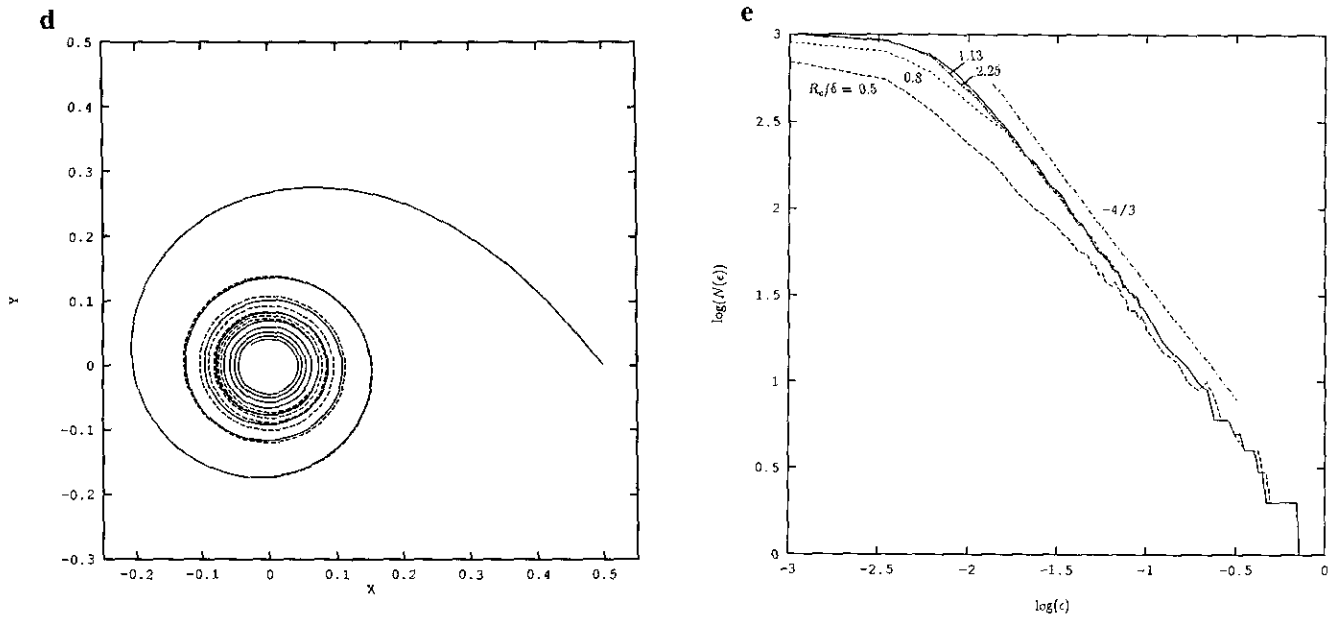

FIG. 11—*Continued*

i.e., as a function of $\eta^* = \eta/\delta$ and $L^* = L/\delta$, where $\eta = 2\pi/k_\eta$ is the Kolmogorov length scale and $L = 2\pi/k_1$ is the integral length scale. The dynamic range of the flow is $\mathcal{L} = L/\eta = k_\eta/k_1$.

Figure 5 shows the log–log error curves in the velocity (normalised by the rms velocity) from TN15 for different $L/\delta =$ 10, 20, 50, 100, against η^*. For a given value of η^* the errors scale uniformly with $\mathcal{L}$. The errors are dominated by the scales which are close to δ, and we have seen from the previous section that scales bigger than about 5δ produce negligible errors. Thus increasing L^* (and so $\mathcal{L}$) will not have a significant effect on the rms errors other than to scale them with the square root of the energy contained in the scales smaller than 5δ. The total energy in the flow is kept constant; hence, increasing the range of wavenumbers merely decreases the proportion of energy in the small scales with which the errors scale. The error curves are close to a -2 slope for large η^*.

Figure 6 shows the log–log normalised error curves in the velocity and the first and second derivatives for $L^* = 100$ obtained from TN15. Curiously, the error curve for the velocity is slightly steeper than -2 and that for the first derivative is slightly shallower than -2, which is surprisingly accurate, since we might reasonably have expected the errors to scale like -1. The unexpected accuracy in the first derivatives is due to the fact that the flow fields are not entirely random and they do exhibit locally correlated regions of structured flow, as explained earlier. Errors at nearby points will also be correlated and taking the difference of velocities at nearby points subtracts off the correlated parts of the errors. Since a derivative is the limit of this difference divided by the distance between them as this distance becomes small, the derivatives will be more accurate than otherwise expected.

The second derivative is nearly zeroth order and remains of the same order of magnitude for all values of η^*.

3.3. *The Influence of Noise*

In experiments there is always a certain amount of error in the data. The influence of errors in the coordinates of the data and in the measured velocity can be important, so we have conducted a number of numerical simulations with imposed random errors in the coordinates of the data points. If the measurements are obtained from PTV experiments of MM93, then these errors are <3% of the rms velocity fluctuations $u' \approx$ 1 mm/s.

We have repeated some of the runs for the sinusoidal test flow with TN15. Noise levels of up to 7% in the velocity were tested (errors in the location of data points are equivalent to putting extra noise in the velocity because in PTV experiments the velocity is obtained from the seeded particle locations and the errors in the two quantities are correlated).

Figure 7 shows the log–log error curves in the velocity against the percentage of noise for $\lambda^* = 4$, 5, 6, 8, and 10. Figures 8 and 9 are the counterparts for the first and second derivatives, respectively. The errors in the velocity are not seriously affected by these low levels of noise and remain at the same order of magnitude.

The error curves from TN15 for the first and second derivatives show interesting behaviour; for large values of λ^* where the accuracy in the noise-free interpolation is good, the effect of noise is severe and the errors increase proportionally to the noise. The errors in the second derivative are particularly severe. This effect can be understood if we look at the derivatives in non-dimensional units. The velocity is non-dimensionalised by

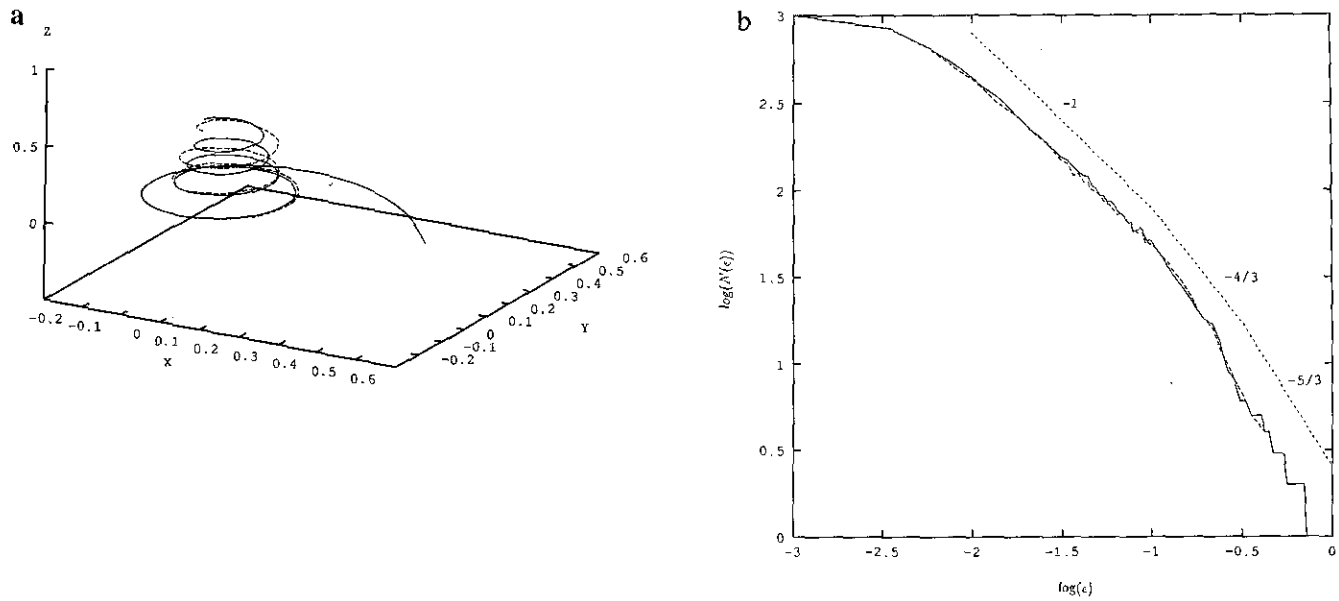

FIG. 12. (a) A streamline of the Burgers vortex, for the same case as in Fig. 11d, except that the streamline begins at $z = 0.01$ and spirals in three dimensions. Dashed line is the interpolated streamline, and the solid line is the exact streamline. $R_c/\delta = 0.63$. (b) The counterpart of figure 11e for the present case. Lines of slope -1, $-4/3$, and -2 are indicated.

the rms velocity and the displacement by the length scale λ of the flow, viz.

$$\frac{\partial u^*}{\partial x^*} \quad \text{and} \quad \frac{\partial^2 u^*}{\partial x^{*2}}.$$

The discrete approximation for a general interpolation scheme with noise is

$$\frac{\Delta u^* + n^*}{\delta^*} \quad \text{and} \quad \frac{\Delta u^* + n^*}{\delta^{*2}}.$$

The errors are thus proportional to

$$\frac{\lambda}{\delta}\frac{n'}{u'} \quad \text{and} \quad \frac{\lambda^2}{\delta^2}\frac{n'}{u'}, \tag{12}$$

since $\delta^* = \delta/\lambda$. u' is the rms fluctuating velocity, and $n^* = n'u'$ is the rms error in u'. Physically, Eq. (12) means that for a given fractional level of noise, $n' = n^*/u'$; the closer the data points are, then, the error in the derivatives will be greater (if the errors are uncorrelated). Greater accuracy can be obtained by increasing the distance between data points δ, but then we lose resolution. Thus, second derivatives can only be accurately determined for the large-scale fields.

We have tested with *uncorrelated* noise. But in an experiment, such as in the PTV experiments for which this interpolation scheme was developed, the errors at nearby points are correlated because the imaging parameters and the spatial configuration are very similar at neighbouring points. This reduces the *un-correlated* part of the noise scheme, an effect that is similar to the one that we observed in the improved accuracy of the first derivatives in Section 3.2. The results above with noise should thus be regarded as the upper bound on the errors.

4. CONTINUITY, STREAMLINES, FRACTALS, AND SPECTRA

Formally, the local N-point Taylor method does not produce continuous velocity fields throughout the domain of interpola-

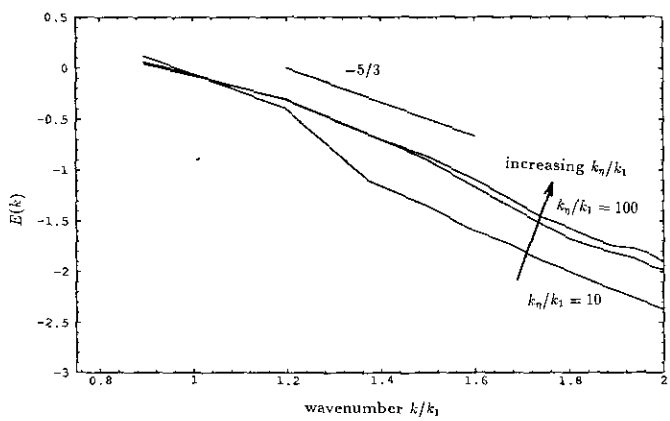

FIG. 13. Log–log plot of the power spectrum from KS fields with varying ranges of wavelengths. The spectra are averaged from ensembles of 154 flow realisations. The Kolmogorov scale η is kept fixed and $\eta/\delta \approx 5$, and $k_\eta/k_1 =$ 10, 50, and 100.

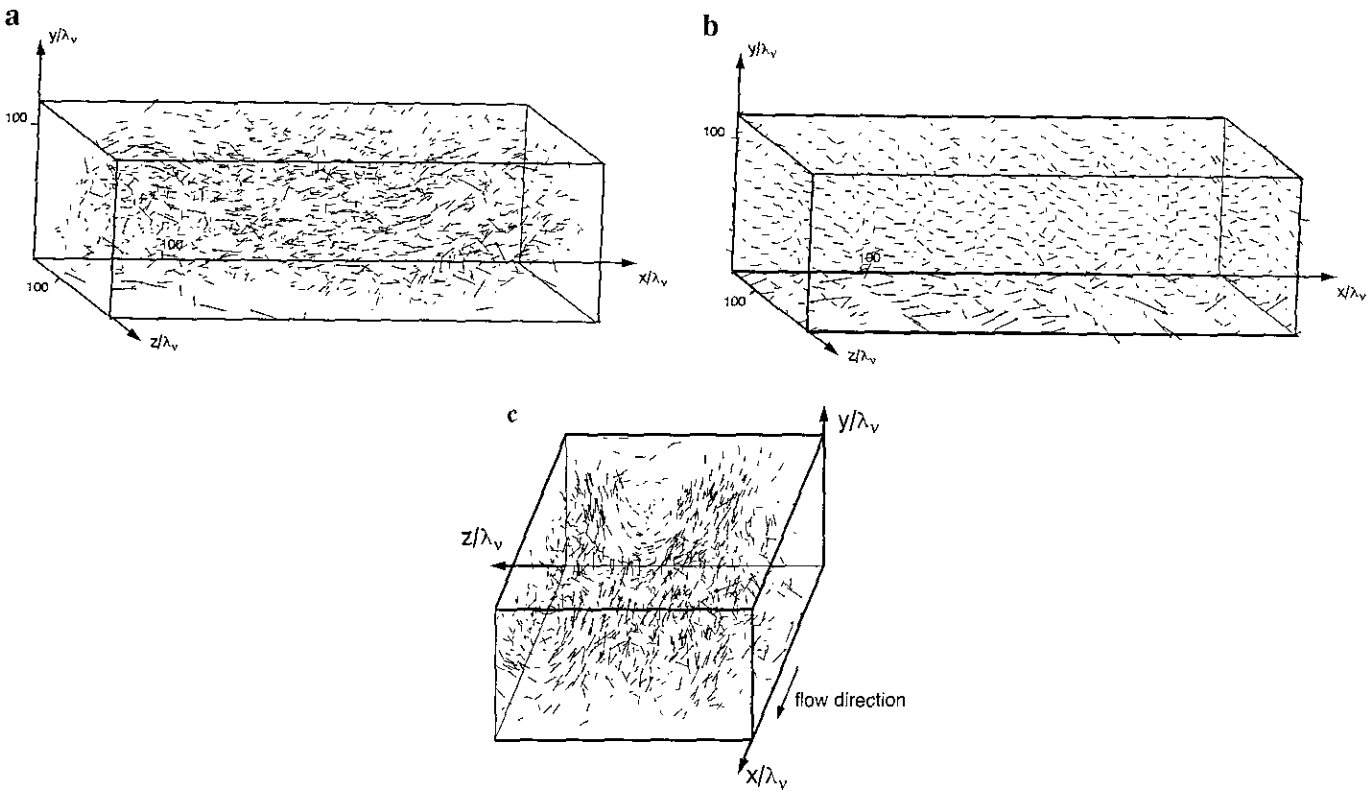

FIG. 14. The velocity vectors obtained from the ETH (IHW) water channel. 1300 velocity vectors were obtained from the PTV method of MM93. (a) The raw data which are the velocities at random positions inside the observation volume which was 60 mm × 45 mm × 20 mm. (b) The interpolated velocity field using TN15 onto a regular grid from the same angle of view as in (a). (c) The interpolated velocity field from a different angle, showing clearly the swirling turbulent motions. In all the plots, the coordinates are normalised by the viscous length scale $\lambda_\nu \approx 0.17$ mm. (The experiments and visualisation were produced by Marko Virant.)

tion—it produces piecewise continuous fields. Each patch of continuous field is defined as the set of points with the same nearest 15 data points (for TN15). We can anticipate that this formal limitation can be reduced by making use of the excess number of data points within a local patch. Using information from the nearest 15 data points, rather than just the 10 nearest points (which gives a unique result), ensures that the discontinuity in the velocity field is minimised because the set of 15 nearest data points changes by just a single point as we move across from one local patch to a neighbouring patch.

Continuity in the interpolated velocity field was tested by computing streamlines in a Burgers vortex [4] which is a steady solution of the Navier–Stokes equations where the action of strain and viscosity balance to give a vortex core of steady finite size, $R_c = \sqrt{2\nu/a}$, which can be interpreted as the Kolmogorov scale of the vortex tube. In cylindrical coordinates (r, ϕ, z), the vorticity is along the z direction with magnitude $\omega(r) = a\,\mathrm{Re}_\gamma/4\pi\, e^{-ar^2/4\nu}$, where ν is the kinematic viscosity, $\mathrm{Re}_\gamma = \gamma/\nu$, and γ is the circulation of the vortex; specifically, $\gamma = 2\pi \int_0^\infty \omega(r) r\, dr$. This flow is sustained by the following straining velocity field, of which a is the strain rate,

$$u_z = az, \quad u_r = -\frac{ar}{2}, \quad u_\phi(r) = \frac{\gamma}{2\pi r}(1 - e^{-ar^2/4\nu}). \tag{13}$$

The streamlines, starting from say $r_0 > 0$ are helical, wind into the vortex core around the z-axis, and are given by

$$\frac{d\phi}{dr} = -\frac{\gamma}{a\pi r^3}(1 - e^{-ar^2/4\nu}). \tag{14}$$

The Burgers vortex is a complex structure exhibiting the effects of strain and circulation and it displays a range of scales defined by the distance between successive windings of its streamlines, as described below. It is thus a good field on which to test the interpolation scheme. In most of our tests the streamlines begin at $z = 0$ so that they remain in the $z = 0$ plane, which reduces the problem to two dimensions. (However, we give an example of a full three-dimensional interpolation case at the end.) We always take the rate of strain $a = 1$ and the circulation $\gamma = 1$, but we take two different cases of the kinematic viscosity ν, and we also vary the density of the data points. We start with the case of $\nu = 0.02$, which gives $R_c = 0.2$. With 500 data points uniformly distributed in an area of 1.4×1.4 centred on the vortex core, the spacing of the data $\delta = 0.063$, which gives $R_c/\delta \approx 3.2$. This is a case of accurate interpolation with the TN15 scheme, and Figure 10 confirms this point. It shows that the interpolated streamline (dashed line) matches the exact

streamline very closely. Figures 11 show the interpolated streamlines (dashed lines) against the exact streamline (solid line) for the case when $\nu = 0.0125$ ($R_c = 0.05$). The density of particles is such that in 11a–11d we have, respectively, $R_c/\delta = 0.5, 0.8, 1.13, 2.25$. The interpolated streamline is quite poor in 11a, but improves progressively as the density of data points increases. In Fig. 11d, the interpolated streamline remains close to the exact one for a number of turns.

Importantly, there is no discontinuity in any of the interpolated streamlines that we have computed. But how can we quantify how "close" the interpolated streamline topology is to the real one? One measure is the fractal dimension of the streamline, in the sense of the Kolmogorov capacity, which is obtained by the familiar box counting algorithm. The capacity is a measure of the space fillingness of a geometric object; more particularly, it is sensitive to the accumulating pattern that is displayed by, say a spiral, as it winds in to the centre of the vortex. It is easy to show that for a spiral of the form $r \sim \phi^{-\alpha}$, the capacity (fractal dimension) of the spiral is unique and equal to $D_k = 2/(1 + \alpha)$ in a range of box sizes which is not too large [19]. Figure 11e shows the log–log plots of the number of boxes of size ε that cover the spiral against the box size ε for the interpolated spirals in Figs. 11a–d, and the solid line is the plot from the exact spiral. The capacity D_k is the (negative) of the slope of the plots. It is remarkable that D_k is well defined and equal to the correct value ($\alpha = 1/2$ so $D_k = 4/3$) in all the cases except the one with the sparsest data in Fig. 11a. Thus, the overall spiraling, accumulating topology of a complex object such as a Burgers vortex streamline is well preserved by interpolation in cases where the sparsity of the data points might suggest larger rms errors.

We now give an example of interpolation in three dimensions. We take the same case as in Fig. 11d, except now we start the streamline from $z = 0.01$ which causes it to spiral in the third component. The data points are now distributed in a volume, to give a spacing of 0.079, and hence $R_c/\delta = 0.63$. Figure 12a shows the exact and interpolated streamlines, and Fig. 12b shows the fractal plots from which we see three distinct ranges. The nature of these ranges is outside the scope of this paper, but the capacities in these ranges are correct, and the interpolated spiral streamline retains these regimes very well, as in the two-dimensional case.

Finally, we want to see how accurately statistical quantities can be obtained from the interpolated data. We have obtained the energy spectrum $E(k)$ from the interpolated KS fields averaged from 154 flow realisations. Figure 13 shows the log–log plot of $E(k)$ against the wavenumber k. $E(k)$ asymptote towards the Kolmogorov spectrum $E(k) \sim k^{-5/3}$ as the wavenumber range of the flow increases. The data spacing δ is such that $\eta/\delta \approx 5$, where η is the Kolmogorov scale. This ensures that all the scales are well resolved.

5. APPLICATION AND DISCUSSION

PTV experiments from the ETH open water channel in July 1993 produced a sequence of coordinate data sets which on average yielded about 1300 velocity vectors in an observation volume at any given time. Figure 14a shows the input data to TN15 which consists of the measured velocities at the scattered locations at a particular instant in time. This gave an average data spacing of nearly 3.5 mm in the observation volume. The largest energy-containing turbulent scales are approximately equal to the depth of the flow, namely 35 mm. A measure of the upper end of the inertial subrange is the Taylor micro-scale which in this flow is approximately equal to 10 mm, and the smallest turbulent scale, the Kolmogorov scale, is roughly equal to 0.5 mm. The mean flow in the streamwise direction was 90 mm/s which gives $u' \approx 5$ mm/s and a Reynolds number of around 3150 based on the channel height. Figure 14b shows the interpolated data on a regular Cartesian grid of spacing $\delta_g = 4$ mm. The arrows are the velocity vectors in a frame moving with the carriage velocity which was equal to the mean flow velocity. The interpolation is probably accurate for the largest scales, but, even so, well-defined swirling motion can be seen in the body of the fluid, which is clearly seen in Fig. 14c which gives another angle of view. Interpolation at the channel wall is not performed because, as we have noted, the wall is not in the inner region of data points. (The measurements and visualisation were produced by Marko Virant.)

To summarise, we have developed a scheme for the general problem of interpolating scattered velocity field data in two or three dimensions onto a given set of arbitrary points within the inner region defined by the data points. The method is based on a trivariate polynomial scheme, which is shown to be equivalent to a Taylor expansion, Eq. (7), and it has proved to be highly accurate in the range of parameters for which the length scale of the flow λ is greater than the spacing of the input data δ by a factor of at least 5. The velocities and first derivatives can be obtained with good accuracy. The second derivatives of the velocity field require more dense data input.

However, we have found that certain quantities are accurately preserved even in cases of much lower data density. The accumulating streamline topology of vortex tubes, in both two and three dimensions, was accurately obtained even with data spacing $R_c/\delta \approx 0.8$, and the fractal dimension (Kolmogorov capacity) was extremely well defined. Similarly, statistical quantities like the power spectrum in a stochastic turbulent field were accurately obtained from the interpolated velocity field.

Thus, the seeding density of particles in flows depends upon the application. If all aspects of the turbulent flow are desired down to the smallest scales, η, one would need the spacing $\delta \leq 0.2\eta$ whatever the largest scales. But if the broad features of streamline topology and statistics are required, then we can take $\delta < 1.2\eta$. This is still quite a severe restriction on experiments since it is difficult to resolve down to the Kolmogorov scale in PTV experiments. One would have to look at small observation volumes to obtain such a high seeding density. Nevertheless, scales somewhat larger than the Kolmogorov scale can be accurately interpolated, which is sufficient for many applications.

The first derivatives of the velocity can be obtained also with reasonable accuracy but the second derivatives are not so accurate. One needs quite high densities of data corresponding to more than $\lambda = 10\delta$ for 10% accuracy.

Noise has little effect on the accuracy of the velocity, and only a moderate effect on the first derivatives. But the second derivatives are severely affected by noise in the data, and the results imply that quantities based on second derivatives can only be obtained for the large-scale fields.

We have compared the Taylor scheme with the standard adaptive Gaussian window scheme, Eq. (1), where we found that the Gaussian window with $H = \delta$ was the best. The Taylor schemes performed much better, except for small $\lambda^* < 2$, where the two schemes are comparable. We have found that the 15-point Taylor scheme TN15 is the optimum interpolation scheme. If a certain degree of error in the velocity field is tolerable, then the speed of the computations can be considerably increased; for 10% error bounds, the speed increases by nearly an order of magnitude using the method of iterations to solve the system of linear equations.

ACKNOWLEDGMENTS

This work was supported by the Swiss National Science Foundation Grant 2-26441. We are very grateful to Professor A. Gruen for use of computer facilities in his department, the Institut fuer Geodesie und Photogrammetrie, ETH Zurich. We are also very grateful to Marko Virant for the experimental data and the graphics produced in Fig. 14.

REFERENCES

1. F. P. Agterberg, *Geomathematics. Mathematical Background and Geo-Science Applications,* Developments in Geomathematics, Vol. 1 (Elsevier Scientific, Amsterdam, 1974).
2. J. C. Agui and J. Jimenez, *J. Fluid Mech.* **185,** 447 (1986).
3. K. E. Atkinson, *An Introduction To Numerical Analysis* (Wiley, New York, 1978).
4. J. M. Burgers, in *Advances in Applied Mechanics,* Vol. 1 (Academic Press, New York, 1948).
5. E. W. Cheney, *Multivariate Approximation Theory: Selected Topics* (SIAM, Philadelphia, 1986).
6. C. de Boor, Multivariate piecewise polynomials, *Acta Numer.* (1993), 65.
7. J. C. H. Fung, J. C. R. Hunt, N. A. Malik, and R. J. Perkins, *J. Fluid Mech.* **236,** 281 (1992).
8. R. L. Harder, and R. N. Desmarais, *J. Aircraft* **9,** 189 (1972).
9. R. L. Hardy, *J. Geophys. Res.* **76,** 1905 (1971).
10. R. L. Hardy, *Comput. Math. Appl.* **19**(8/9), 163 (1990).
11. K. Imaichi and K. Ohmi, *J. Fluid Mech.* **129,** 283 (1983).
12. H.-G. Maas, A. Gruen, and D. Papantoniou, *Exp. Fluids* **15**(2), 133 (1993).
13. N. A. Malik, Ph.D. thesis, University of Cambridge, 1991 (unpublished).
14. N. A. Malik, T. Dracos, and D. Papantoniou, *Exp. Fluids* **15**(4/5), 279 (1993).
15. C. A. Micchelli, "Algebraic Aspects of Interpolation," in *Approximation Theory Short Course* (Amer. Math. Soc., Providence, RI, 1986).
16. G. R. Spedding and E. J. M. Rignot, *Exp. Fluids* **15**(6), 417 (1993).
17. K. D. Squires and J. K. Eaton, *J. Fluid Mech.* **226,** 1 (1991).
18. G. I. Talyor, *Proc. R. Soc. London A* **20,** 196 (1921).
19. J. C. Vassilicos and J. C. R. Hunt, *Proc. R. Soc. London A* **435,** 505 (1991).
20. P. K. Yeung and S. B. Pope, *J. Comput. Phys.* **79**(2), 373 (1988).
21. P. K. Yeung and S. B. Pope, *J. Fluid Mech.* **207,** 531 (1989).