von Karman Institute for Fluid Dynamics

Lecture Series 1990–06

NUMERICAL GRID GENERATION

June 11–15, 1990

THE FUNDAMENTALS OF ADAPTIVE GRID MOVEMENT

P.R. Eiseman

Program Development Co & Columbia U., USA

TABLE OF CONTENTS

	page
INTRODUCTION	1
WEIGHT FUNCTIONS	3
EQUIDISTRIBUTION IN ONE DIMENSION	4
The Differential Statement	5
The Local Integral Statement	5
The Direct Grid Statement	6
The Backward Global Integral Statement	6
The Forward Global Integral Statement	9
The Differential Equation Statement	11
The Direct Finite Difference Statement	13
The Mean-Value Relaxation Statement	13
The Least Squares Statement	15
The Constrained Least Squares Statement	16
Variational Statements	18
Evolutionary Statements	20
THE SPECIFICATION OF COEFFICIENTS IN THE LINEAR WEIGHT	23
Fractions	23
Level of Significance	25
Minimum and Maximum Spacing	36
THE ATTRACTION TO A GIVEN GRID ON A CURVE	27
Inverse Equidistribution	28
Diagonal Dominance	31
The Variational Format	33
EVOLUTIONARY FORCES	34
Globally Determined Weights	34
The Attraction to a Given Grid	36
Minimum and Maximum Spacing	36
Eccentricity	39
METRIC NOTATION	40
CURVE BY CURVE METHODS	41
Metric Specification	42
The Historical Development of Pure Curve by Curve Methods	43
Curve by Curve Adaptation with Derivative Smoothness	49
The Conformal Measure of Smoothness: Basic Elliptic Grid Generation	50
Curve by Curve Adaptation with Conformal Smoothness	53
The Variational Form of Curve by Curve Adaptation	56
FINITE VOLUME METHODS	59
Direct Mean-Value Relaxation	59
Full Mean Value Relaxation	64
Direct Variational Constructions	68
VARIATIONAL METHODS	70
TEMPORAL ASPECTS	71
CONCLUSION	75
ACKNOWLEDGEMENT	75

INTRODUCTION

With an assumed governing system of partial differential equations (PDE's), numerical simulations typically consist of applying a grid generator and a PDE-solver. In each case, a communication link is established from the generated grid to the PDE-solver. This link occurs when the solver is written in terms of a grid which covers the physical region. In the event that a solution should develop large variations on a finer scale than is provided by the local grid spacing, a communication link must also be established from the PDE-solver to the grid generator. The grid is then driven by the solution to meet the resolution requirements and is called adaptive.

While the communication link from the grid to the solver is typically a matter of transferring solution data to the grid and then expressing the PDE's with respect to the grid, the link in the opposite direction is more subtle. To form it, we must use the solution data to determine the locations where the errors might become large. In a formal analytical sense, the locations requiring resolution can be obtained by error estimates that are derived for the PDE solver. Alternatively, the same locations can often be determined directly from a basic knowledge of the problem undergoing simulation. From the direct viewpoint, we note that the solution is generally given as a vector of dependent variables at each point in physical space and geometrically appears as a surface over the physical space. Rather than taking the solution surface itself, we use our knowledge about it to extract the quantities which will experience rapid variations. We then form another surface with those quantities and, to be descriptive, we call it a monitor surface since it must suitably monitor solution behavior. Rather than forming it with a direct vector of the quantities, scalar combinations are employed to achieve simplicity. The simplest situation occurs when a single scalar linear combination is used. By definition, the locations that require resolution are then determined by large gradients of this scalar. In a geometric spirit, the gradients are resolved when the monitor surface is uniformly covered with grid points. Moreover, the solution is more accurately represented by the grid when the bends in the monitor surface are also resolved. This can be achieved with curvature clustering.

With the adaptive data suitably expressed in the form of an error estimate or a monitor surface, the next step is to develop grid generators that

can use the data to cluster points in the desired locations. The manner in which the clustering is performed must also be done automatically and reliably since human intervention is not assumed to be available for corrective actions. In broad terms, the clustering is achieved by the movement of grid points and by changes in the number of grid points.

Most methods which alter the number of points tend to start with a fixed global grid and proceed to locally add points as they are needed. Sometimes a means to remove points when they are not needed is also employed. As a consequence of local point additions and subtractions, the overall grid structure becomes rather complicated in the sense of the accompanying data structures and internal boundary treatment between fine and coarse spacing. In contrast, methods which move the points of a curvilinear grid maintain a simple data structure and have no such internal boundaries. However, with a fixed number of points, local resolution is achieved at the expense of depreciating the resolution in other regions. In many cases the depreciation is quite minor because the other regions often have an overabundance of points that would not significantly contribute to the accuracy of the simulation. The limitation of fixing the number of points then only occurs when there are not enough points to resolve either the local phenomena or the other regions. As a consequence, a reasonable strategy is to globally alter the number of points in each coordinate direction while maintaining grid point movement as the primary adaptive mechanism.

One simple way to make the global decision is to first determine the number of points required for a mildly varying solution and then to estimate the number of extra points required for the severe variations. Along a coordinate curve, the estimate is just the number of severe variations times the number of points required for such variations. To adequately employ all curves in a given coordinate direction, the estimate to be used is then the maximum of the estimates for the curves. By adding it to the number required for a mildly varying solution, the total number of points in a given direction is obtained. With such strategies, the number of points can vary between the requirements of mild and arbitrarily severe solution behavior. The use of points is then somewhat optimal in the sense that there are just enough to resolve everything to our satisfaction without a tremendous excess of modestly contributing points.

To keep our discussion within simple bounds, we will only consider basic

grid point movement schemes with the implicit understanding that any of them can be applied as part of strategy where the number of points in each coordinate direction can also be adaptively adjusted as indicated above. As such, we will refer to the schemes as adaptive grids. This is consistent with the notion of grids as discrete coordinate transformations as opposed to meshes that can assume an unstructured format.

The general topic of grid generation has been the subject of a number of conferences. These occurred at NASA Langley [1] in 1980, at Nashville [2] in 1982, at Houston [3] in 1983, and at Landshut, West Germany [4] in 1986. The latter is the start of a two-year international conference sequence. The topic also has appeared in a conference on adaptive methods in Maryland [5] in 1983 and has assumed a role of growing importance in the AIAA Computational Fluid Dynamics Conference sequence as well as in the international series on Numerical Methods in Fluid Dynamics, published in the numbered sequence of Lecture Notes in Physics by Springer-Verlag.

In addition, there are general surveys by Thompson, Warsi and Mastin [6] in 1982, by Thompson [7] in 1984, by Eiseman [8] in 1985, and by Eiseman and Erlebacher [9] in 1987. More specialized surveys have also appeared in the conferences at NASA Langley and at Nashville. On the topic of adaptive grid generation, the previous reviews were given by Anderson [10] in 1983 and by Thompson [11] in 1985. The present review represents a fairly thorough current account of this active and important subject.

WEIGHT FUNCTIONS

At the root of adaptive grids is the desire to generate good grids with cells that contain equal amounts of a specified weight function. When the generation process is performed in physical space, the error estimates or gradients of the monitor surface are directly inserted into the weight function along with other quantities such as monitor surface curvature or a background distribution. When the process is performed on the monitor surface and the consequent grid is projected down onto the physical region, the gradients are automatically resolved by the use of surface arc lengths. With the weight functions for surfaces, the emphasis is on geometric quantities that allow us to cluster points in a manner that improves the discrete representation of the surface. These weights are applied against surface arc lengths, areas, and

volumes as opposed to the corresponding projected elements in physical space. As a consequence, the main clustering quantities in the weights are the various forms of curvature. This includes normal and mean curvatures for the basic geometry and geodesic curvature for boundaries.

Regardless of where the grid generation process is performed, the basic formation and utilization of weight functions remain the same: only the attracting quantities and the cellular elements are changed. In the most general form, the weight is a positive scalar function that incorporates quantities for a background distribution, for arbitrarily attracting points, and for improving the structure and smoothness of the grid. The simplest and most commonly used form is a linear combination. The general linear expression is given by

$$w = 1 + c_1 M_1 + c_2 M_2 + \ldots + c_m M_m \qquad (1)$$

where the above quantities are represented with nonnegative functions M_i that have nonnegative coefficients c_i to indicate the level of importance attached to them. We now think of the functions M_i as masses which then more strongly attract points when they are large. For the arbitrary attraction to a gradient, it is the magnitude of the gradient or some normalized version of the magnitude. For a background distribution, it is a derivative of that distribution. For an improvement in the structure or smoothness of the grid, it is a pull of the grid towards orthogonality or towards conformal conditions. Altogether, these masses represent a collection of attributes that compete with each other and with the first term of unity which represents uniform conditions. The uniformity is actually achieved if all the coefficients c_i should vanish. The uniformly distributed items are the elements on which the weight is applied. By making any of the coefficients c_i arbitrarily large, the associated attribute can be made to be dominant and to thereby dwarf the effect of the uniformity term. In such a situation, the first term is viewed as only a guarantee that the weight function is never zero. In some circumstances, this term is dropped entirely.

EQUIDISTRIBUTION IN ONE DIMENSION

With the motivation to directly cluster points as the weight becomes

large, the first consideration is to move the grid to get the cells to contain equal amounts of weight. This equally distributes the weight over the cells of a grid and is called an equidistribution process. Upon accomplishment, the equalization results in small cells to accommodate large weights and larger cells to accommodate smaller weights. Since the basic principles involved can be most simply displayed in one dimension and since most of the adaptive grid generation methods have their roots established in that context, we shall first consider one-dimensional developments and then proceed to the higher dimensional versions.

The Differential Statement

In one dimension, equidistribution occurs when the spacing between points is inversely proportional to the weight. Assuming a transformation from a curvilinear coordinate ξ to the arc length s along a given curve, the relationship is given by the differential statement

$$wds = cd\xi \tag{2}$$

where c is the proportionality constant. In the mapping, a uniform distribution in ξ is transformed into a nonuniform distribution in s. The desired spacing is achieved because $d\xi$ is also viewed as a constant which then means that the spacing ds must shrink or expand to accommodate respectively large or small weights.

The Local Integral Statement

In the finite sense of grids, the appropriate conditions come from local integrations that are taken between the successive grid points. For an index i, the integration is from ξ_i to ξ_{i+1} and the corresponding s_i to s_{i+1}. The consequent local integral statement is

$$\int_{s_i}^{s_{i+1}} w\,ds = c(\xi_{i+1} - \xi_i) \tag{3}$$

where

$$\xi_{i+1} - \xi_i = \text{constant}$$

The second part is the requirement for a uniform distribution in ξ that now has assumed the form of a grid with constant spacing. In an intuitive sense, the total weight over each cell is simply required to be a constant.

The Direct Grid Statement

While a variety of schemes could be constructed by evaluating the local integrals in various manners (i.e., distinct quadrature rules), the most common method is to assume a value of weight at the center and to use it as a constant over the interval. This approximation for the integral leads to the explicit grid statement

$$w_{i+1/2}(s_{i+1} - s_i) = \text{constant} \tag{4}$$

where typically

$$w_{i+1/2} = \frac{1}{2}(w_{i+1} + w_i) \tag{5}$$

The effect of the weight on the grid points is now more directly evident than it was in the previous statements. Specifically, as $w_{i+1/2}$ is increased, s_i and s_{i+1} approach each other.

The Backward Global Integral Statement

For a direct uncoupled relationship between grid point locations, global rather than local integrals are considered. When the weight w is given as a function of only the location s along the curve, a direct integration of Eq. 2 leads to the transformation

$$\frac{\xi - \xi_{min}}{\xi_{max} - \xi_{min}} = \frac{F(s)}{F(s_{max})}$$

where (6)

$$F(s) = \int_{s_{min}}^{s} w\,dx$$

The derivation consists of first determining the proportionality constant of Eq. 2 from the total integral as

$$c = \frac{F(s_{max})}{\xi_{max} - \xi_{min}} \tag{7}$$

and then using the integral up to the current location s to get Eq. 6.

Within the context of application, a curve has been given as a function of a parameter s and then the new parameter ξ has also been expressed as a function of s. In the language of mappings, the functional relationships become the map from an interval of values s to the curve and the map from the same interval to the interval for ξ. The latter map, however, is clearly backwards with respect to the straight composition of maps which is from ξ to s and then to the curve. As a consequence, the map of Eq. 6 will be called the backward global integral statement.

An implication from the backwardness is the need to invert the transformation in order to apply it. In terms of grids, the inversion is most simply accomplished by a local linear interpolation within a table of integral values for the previous grid points $a_1 < a_2 < \ldots < a_N$ on the interval $s_{min} \leq s \leq s_{max}$. With midpoint weights determined in the manner of Eq. 5, the integral of Eq. 6 is approximated by the trapezoidal quadrature rule

$$F(a_k) = \sum_{i=1}^{k-1} w_{i+1/2}\left(a_{i+1} - a_i\right) \tag{8}$$

for k = 2, 3, ..., N and by 0 when k = 1. When the previous points a_i are taken as the piecewise linear approximation to curve arc length, the spacing

increments are given by the Euclidian distance norm as

$$a_{i+1} - a_i = ||P_{i+1} - P_i|| \tag{9}$$

where P_1, P_2, ..., P_N are the previous grid points on the curve. Returning to the transformation of Eq. 6, uniform spacing in ξ yields values of $(j-1)/(N-1)$ on the left-hand side and thus the interpolation problem

$$F(s_j) = \frac{j-1}{N-1} F(a_N) \tag{10}$$

for the internal points s_j where $j = 2, 3, \ldots, N-1$. The boundary points are already known and are $s_1 = a_1$ and $s_N = a_N$. For each j, the value on the right-hand side of Eq. 10 is used in a search to find its position within the table of values from Eq. 8. The result is an index $m = m(j)$ for which

$$F(a_m) \leq F(s_j) < F(a_{m+1}) \tag{11}$$

with possible equality on the right if it is the last interval $m = N-1$. Using the fractional amount of entrance into the interval

$$\alpha_j = \frac{F(s_j) - F(a_m)}{F(a_{m+1}) - F(a_m)} \tag{12}$$

the new values of s_j are given by

$$s_j = a_m + \alpha_j (a_{m+1} - a_m) \tag{13}$$

and the corresponding new locations on the curve are given by

$$Q_j = P_m + \alpha_j (P_{m+1} - P_m) \tag{14}$$

In the event that the transformation of Eq. 6 is analytically defined, the inversion can be more accurately executed by a variety of point iterative methods. Moreover, if the curve is also analytically defined, then the grid point locations can indeed be very accurately determined. However, in the general adaptive context, such analytical data is usually not available and

such high accuracy in the grid point locations is rarely critical. As a consequence, the most widely applicable inversion is the simple numerical approach that was just presented here.

The Forward Global Integral Statement

As an alternative, the inversion problem can be avoided altogether by the construction of a forward transformation in place of the previous backward one. The basic assumption is that the weight w must be considered purely as a function of the new curvilinear variable ξ. While the functional dependence can often be readily shifted by a change of variables, this viewpoint is important because w is seen to vary with respect to the number of grid points rather than with respect to specific locations on the curve. In cases where w is explicitly defined as a function of ξ, a change to s becomes a real computational task that requires a global iteration to suitably converge.

To derive the forward global integral statement, Eq. 2 is divided by w and is then integrated in the same manner as before. The result assumes the form

$$\frac{s - s_{min}}{s_{max} - s_{min}} = \frac{G(\xi)}{G(\xi_{max})} \tag{15}$$

where

$$G(\xi) = \int_{\xi_{min}}^{\xi} \frac{d\zeta}{w(\zeta)}$$

and where the proportionality constant in Eq. 2 is given by

$$c = \frac{s_{max} - s_{min}}{G(\xi)} \tag{16}$$

Using the midpoint values of Eq. 5, the integral in Eq. 15 is approximated as in Eq. 8 by the trapezoidal quadrature rule which here appears as

$$G(\xi_k) = \Delta\xi \sum_{i=1}^{k-1} \frac{1}{w_{i+1/2}} \tag{17}$$

By displaying the constant increment $\Delta\xi$ in front of the summation, the fact that ξ represents a counter of points is accented. Since the simplest counter is achieved by taking the index i itself, it is common practice to assume that $\Delta\xi = 1$. When the forward transformation of Eq. 15 is evaluated at the uniform points $\xi_i = i$, the new locations are given by

$$s_j = s_{min} + (s_{max} - s_{min})\frac{G(j)}{G(N)} \tag{18}$$

where

$$G(j) = \sum_{i=1}^{j-1} \frac{1}{w_{i+1/2}}$$

for $j = 2, 3, \ldots, N$. For $j = 1$, the sum $G(1)$ vanishes and the result is $s_1 = s_{min}$.

From the forward character of directly producing new parametric locations s_j, each new grid point location $\mathbf{Q}_j$ is directly obtained from another evaluation. If the curve is only defined by the previous grid points $\mathbf{P}_1, \mathbf{P}_2, \ldots, \mathbf{P}_N$, then the evaluation must be accomplished by interpolation. The simplest and most commonly used approach is to find the interval in the previous parametric grid $a_1 < a_2 < \ldots < a_N$ that contains s_j and then to use local linear interpolation. This proceeds by first searching to get the index $m = m(j)$ such that

$$a_m \leq s_j < a_{m+1} \tag{19}$$

with possible equality on the right if $m = N-1$. Using the fractional distance into the interval

$$\alpha_j = \frac{s_j - a_m}{a_{m+1} - a_m} \tag{20}$$

the new locations on the curve are given by

$$Q_j = P_m + \alpha_j(P_{m+1} - P_m) \tag{21}$$

In contrast to the backward case, the local linear interpolation here occurs with respect to the previous parameter grid.

Altogether, we could reasonably get the impression that the forward transformation is a much more efficient approach than the backward transformation. This is certainly true when the weight function is explicitly defined as a function of the grid point counter ξ. However, in the adaptive context, the variations in the solution are present at distinct locations in the physical region regardless of the grid employed for a simulation. As a consequence, the weights are really given as explicit functions of the spatial length s. This occurs because the weights are formed from solution properties. In the application of the forward transformation, the weights are then only implicitly given as a function of ξ in the form $w(s(\xi))$. The result is the requirement for an iteration. From a given weight function defined with respect to s, the evaluations w_i over an initial grid of points s_i immediately defines a functional relationship between the grid point counter $\xi = i$ and the values w_i. With this data, the forward transformation is then applied to produce new grid point values s_i. A return is made to the weight evaluations and the process is repeated. It is stopped upon reaching the limit of a convergence criterion.

The Differential Equation Statement

While the constant of proportionality in the differential statement of Eq. 2 was explicitly evaluated in the global integral statements, it is removed by differentiation in the differential equation statement only to reappear in the form of a second boundary condition. The first boundary condition was the initial value for the global integrations as applied to the simple first-order equations.

For notational simplicity, a subscript by an independent variable will represent differentiation with respect to the variable. From Eq. 2, a ξ-differentiation of the proportionality constant vanishes and results in the second-order differential equation

$$(ws_\xi)_\xi = 0 \tag{22}$$

Upon expansion and division by w, it assumes the form

$$s_{\xi\xi} + \frac{w_\xi}{w} s_\xi = 0 \tag{23}$$

where the weight term has been isolated. Accenting this isolation, the equation can be written as

$$s_{\xi\xi} + as_\xi = 0 \tag{24}$$

where

$$a = \frac{w_\xi}{w} \tag{25}$$

Conversely, when any equation of the form in Eq. 24 is assumed, the secondary equation, Eq. 25, may be employed to get a weight function. The result is given by

$$w = c_1 e^{\int a d\xi} \tag{25}$$

where c_1 is the constant of integration. This weight is clearly equidistributed since all of the steps leading to the pair Eqs. 24-25 can be readily done in the opposite direction. The important consequence here is that Eq. 24 can be viewed as a basic equidistribution law. This form is called the differential equation statement since it is another statement which is equivalent to the others.

An inverse form may also be given and can be determined either directly from the differential statement of Eq. 2, or from a formal interchange of variables s and ξ in Eq. 23. In the case of Eq. 2, an s-differentiation of ξ_s/w is seen to vanish. Upon simplification, the result is in the same form as in Eq. 24 but with a sign change and an interchange of variables. It also assumes the Poisson equation format for 1-D which is given by

$$\xi_{ss} = \frac{w_s}{w} \xi_s \tag{27}$$

and which in higher dimensions has been extensively used for elliptic grid generation as discussed by Thompson [12] from a general perspective.

The Direct Finite Difference Statement

While various finite difference schemes could be employed to solve the previous statement in terms of a differential equation, a more immediate approach is to use the direct grid statement of Eq. 4. This is an approximation to the differential equation (Eq. 22) that is readily available. By writing Eq. 4 for i-1 and i and noting that the same constant appears on the right-hand side, we have

$$w_{i+1/2}(s_{i+1} - s_i) = w_{i-1/2}(s_i - s_{i-1}) \tag{28}$$

which can be rewritten as

$$w_{i+1/2}s_{i+1} - (w_{i+1/2} + w_{i-1/2})s_i + w_{i-1/2}s_{i-1} = 0 \tag{29}$$

When combined with the boundary conditions

$$\begin{aligned} s_1 &= s_{min} \\ s_N &= s_{max} \end{aligned} \tag{30}$$

we have a simple tridiagonal matrix to invert to get the interior values s_2, s_3, ..., s_{N-1}. The corresponding grid points Q_j along the given curve are then determined in the same manner as in the global forward integral statement. In addition, the need for iteration is also required since the weights appear at unknown locations. Only the index i is known. This is discussed in the paragraph containing Eqs. 19-21.

Mean-Value Relaxation Statement

With a view towards direct extensions into higher dimensions, we are motivated to cast the tridiagonal finite difference form of Eq. 29 into a statement of point relaxation that contains an intuitively plausible basis for such extensions. Without this view, there would be no compelling desire to

consider point relaxation since it is substantially more efficient to use the tridiagonal form in the sense of direct Gaussian elimination. In higher dimensions, by contrast, the efficiency can be retrieved by multigrid methods [13] and by the use of parallel computing environments [14].

By solving for the center point s_i in Eq. 29, a form that is immediately suitable for point relaxation arises and is given by

$$s_i = \frac{w_{i+1/2}s_{i+1} + w_{i-1/2}s_{i-1}}{w_{i+1/2} + w_{i-1/2}} \tag{31}$$

In the Gauss-Seidel application, the values on the right-hand side use the available current information while the point Jacobi application ignores the current information in favor of the data from the previous sweep over the field. With either method of relaxation, however, the formula in Eq. 31 does not yield the desired intuitive basis since the evaluations of s are unaligned with the evaluations of w. By adding s_i to each side of Eq. 31 and then dividing the result by 2, the apparent intuitive obstacle can be overcome. Upon appropriately grouping terms, we arrive at

$$s_i = \frac{w_{i+1/2}s_{i+1/2} + w_{i-1/2}s_{i-1/2}}{w_{i+1/2} + w_{i-1/2}} \tag{32}$$

where

$$s_{i+1/2} = \frac{s_{i+1}+s_i}{2}$$

and (33)

$$s_{i-1/2} = \frac{s_i+s_{i-1}}{2}$$

In the present form, we can now intuitively look at the weights over the intervals on either side of s_i. Assuming uniform weights over each side, the weighting centers are at the midpoints as given by Eq. 33 and are in correspondence with the weight values there. In a physical sense, the formula is readily interpreted as a center of mass formula for the interval from s_{i-1} to s_{i+1}.

In the sense of pointwise movement, the center of mass form can be recast to explicitly display the movement. This is accomplished by subtracting the

previous position $\bar{s}_i$ from each side of Eq. 32. The result is given by

$$s_i - \bar{s}_i = \frac{w_{i+1/2} d_+ - w_{i-1/2} d_-}{w_{i+1/2} + w_{i-1/2}} \tag{34}$$

where

$$d_+ = s_{i+1/2} - \bar{s}_i$$

and (35)

$$d_- = \bar{s}_i - s_{i-1/2}$$

are the maximum movement distances in positive and negative directions from $\bar{s}_i$.

The subsequent mapping to the curve is accomplished locally. The localness occurs because the move to a new s_i cannot exceed one-half of the distance from the old s_i to its neighbors s_{i-1} and s_{i+1}. If both neighbors are old locations, then the halfway restriction prevents an interchange of s_i with either s_{i-1} or s_{i+1} when their respective moves are also considered: at best they may simultaneously approach either midpoint of Eq. 33. When the current values of s_{i-1} and s_{i+1} are immediately used upon availability, the halfway restriction is not required to prevent such an interchange. The important fact here is that with either method of point relaxation, the center point cannot move outside of the interval from s_{i-1} to s_{i+1} at any stage of relaxation. This then means that the previous global search used in the direct Gaussian elimination approach would be replaced by a local search between the two adjacent intervals. Beyond this stage, the mapping to the curve proceeds with the same linear interpolation over the selected interval.

The Least Squares Statement

In contrast to the previous statements which evolved directly from the original differential statement, the evolution will now be considered to start from a global expression where the competition among the different locations appears simultaneously. The basic action comes from the minimization of the weighted sum of squares

$$A = \sum_{i=1}^{N-1} w_{i+1/2}(s_{i+1} - s_i)^2 \tag{36}$$

The competition among locations corresponds to the distinct terms in the sum. By viewing the weighting coefficients as a sequence of positive constants, the minimum is determined by setting

$$\frac{\partial A}{\partial s_j} = 2w_{j-1/2}(s_j - s_{j-1}) - 2w_{j+1/2}(s_{j+1} - s_j) = 0 \tag{37}$$

and is thus seen to give the earlier tridiagonal system of Eqs. 29-30.

When the weighting coefficients are considered to be evaluations of a single continuously defined weight function, the points of evaluation are seen to be only a function of the index. This means that the weight is only a function of the curvilinear variable ξ which is then evaluated at uniformly spaced points. In practice, the points are typically chosen to be the index itself.

As a matter of distinction, if the weight had solely been a function of the position s, then the evaluations at each index j would have varied with the current associated location s_j. The result for the minimization equations would have been the inclusion of weight derivatives and a nonlinear dependence upon s. From an iterative point of view, the nonlinear effect would appear as a requirement to reevaluate the weights and/or their derivatives when each s_j assumes a new value.

The Constrained Least Squares Statement

As an alternative to the minimization with respect to the pointwise locations s_j, the previous sum of squares can also be minimized with respect to the spacing between points. In an intuitive sense, the contributions from each term will be separated from its neighbors and will more directly produce the direct grid statement of Eqs. 4-5. The added complication for this desirable separation is the constraint that the sum of all of the spacing increments equals the total length.

With the spacing between successive points being denoted by

$$h_{i+1/2} = s_{i+1} - s_i \tag{38}$$

the previous sum of squares (Eq. 36) is expressed by

$$A = \sum_{i=1}^{N-1} w_{i+1/2} h_{i+1/2}^2 \tag{39}$$

and is subjected to the constraint

$$\sum_{i=1}^{N-1} h_{i+1/2} = s_{max} - s_{min} \tag{40}$$

For the minimization process, a Lagrange multiplier λ is introduced to form

$$B = A - \lambda\left[\sum_{i=1}^{N-1} h_{i+1/2} - (s_{max} - s_{min})\right] \tag{41}$$

from which we proceed to vary each $h_{i+1/2}$ in the same manner as before. The minimum then occurs when

$$\frac{\partial B}{\partial h_{i+1/2}} = 2w_{i+1/2} h_{i+1/2} - \lambda = 0 \tag{42}$$

which is readily recognized as the direct grid statement of Eq. 4 and which can be rewritten as

$$h_{i+1/2} = \frac{\lambda}{2w_{i+1/2}} \tag{43}$$

to express each increment as a function of the single unknown λ. By substituting the expression for each increment (Eq. 43) into the constraint (Eq. 40) of $\partial B/\partial \lambda = 0$, the Lagrange multiplier is found to be

$$\lambda = \frac{2(s_{max} - s_{min})}{\sum_{i=1}^{N-1} \frac{1}{w_{i+1/2}}} \tag{44}$$

Returning to Eq. 43, each increment is now known. Upon also inserting Eq. 38 into Eq. 43, the k-th increment is then given by

$$s_{k+1} - s_k = \frac{\frac{1}{w_{k+1/2}}}{\sum_{i=1}^{N-1} \frac{1}{w_{i+1/2}}} (s_{max} - s_{min}) \tag{45}$$

By forming partial sums starting from k = 1, the forward global integral statement of Eq. 18 is now readily apparent.

The solution to the constrained least squares problem also has a direct geometric interpretation that was observed by Steinberg and Roache [15]. With the increments considered as the Cartesian coordinates of (N-1)-dimensional Euclidean space, the basic sum of squares in Eq. 39 represents a hyper-ellipsoid for each value of A. Clearly, each such Cartesian coordinate $h_{i+1/2}$ is bounded by

$$|h_{i+1/2}| \leq \sqrt{\frac{A}{w_{i+1/2}}} \tag{46}$$

where equality is achieved only along the $h_{i+1/2}$-axis.

From this we see that the family of hyper-ellipsoids varies from the origin when the "radius" $\sqrt{A}$ is zero to arbitrarily large sizes as the radius is increased. Above the origin, however, we have a hyper-plane of constraints defined by Eq. 40. Geometrically, it passes through each axis at values of $s_{max} - s_{min}$. Since hyper-ellipsoids are curved convex surfaces, there then is only one hyper-ellipsoid that intersects the hyper-plane at a single point. This also is a point of tangency. While lesser values of radius $\sqrt{A}$ have no intersection, higher values have infinitely many. But to minimize A, we are only interested in the smallest value with an intersection. This is just the single point where tangency occurs.

Variational Statements

The natural framework to continue the previous least squares statements is provided by the variational statements. In distinction, these formulations

are now considered from a continuum rather than a discrete point of view. The general statement here is to find the extremum of an integral. When the integration is with respect to the curvilinear variable ξ, we shall consider the integrand type

$$F_1 = [w(\xi)s_\xi]^\alpha s_\xi \tag{47}$$

for a weight function dependence on ξ and the integrand type

$$F_2 = [w(s)s_\xi]^\alpha \tag{48}$$

for a weight function dependence on s. When $\alpha = 1$ for F_1 in Eq. 47, the integral becomes the continuum version of the least squares statement. Upon approximation with unit increments in ξ, the relationship becomes

$$\int_1^N w(\xi)s_\xi^2 d\xi = \sum_{i=1}^{N-1} w_{i+1/2}(s_{i+1} - s_i)^2 \tag{49}$$

and provides a direct connection with the previous discrete least squares developments.

Returning to the integral forms $\int F_i d\xi$, the extremum is found as before by setting first derivatives to zero. This is accomplished by using the Calculus of Variations [16]. The result is the Euler equation

$$\left(\frac{\partial}{\partial s} - \frac{d}{d\xi}\frac{\partial}{\partial s_\xi}\right)F_i = 0 \tag{50}$$

By direct computation, it can be readily verified that the Euler equations reduce to a differential equidistribution statement for each choice of integrand. In the case when the weight function depends upon the location s, we also note the appearance of a weight derivative. This was also observed when we examined the basic least squares statement.

As an alternative, we may change the differential in the integral by using $d\xi = \xi_s ds$. We must then, however, also use $s_\xi = 1/\xi_s$ in F_i. By the same argument applied to the interchanged variables, the Euler equation is now given by

$$\left(\frac{\partial}{\partial \xi} - \frac{d}{ds}\frac{\partial}{\partial \xi_s}\right)(F_i \xi_s) = 0 \tag{51}$$

By carrying out the algebraic manipulations for each given F_i, the differential equidistribution statement can be retrieved as earlier. The selection of integration variable is then clearly seen to be a matter of convenience.

Evolutionary Statements

In the previous statements, equidistribution was expressed in various spatial forms that produced a distribution either directly or by iteration. The need for iteration arose in situations where the weight, being known as a function of s, was instead used as a function of $\xi = i$ rather than s. This occurred in a natural fashion when the weights were needed at yet undetermined positions s_i: the only possibility was then to use the currently available values in an iterative cycle. During the course of iteration, no consideration was given to the potential control of the evolutionary iterative cycles. In the present section, such an evolutionary control is examined.

In the most immediate manner, the evolutionary control can be stated by the parabolic partial differential equation

$$Ks_t = (ws_\xi)_\xi \tag{52}$$

where the non-negative function $K = K(\xi)$ represents the rate of evoluation for each point $\xi = i$. This rate varies from an infinite speed when K vanishes to a zero speed when K is infinite. As a consequence, the motion can be greatly reduced or stopped in certain locations and can be gradually increased to large values in other locations. From a physical point of view, the diffusion of points into the equidistribution positions increases with 1/K.

In the case where K vanishes everywhere, the points then instantly move into the locations for the equidistribution of w. The case corresponds to the earlier differential equation statement of Eq. 22 and its various forms in Eqs. 23-27. The infinite speed represented by the instantaneous adjustment comes from the elliptic nature of the consequent differential equation. In practice, however, the elliptic type of differential equation would still have to be solved numerically and, in the given situation, would have to be done by iteration. The instantaneous adjustment is then represented by the direct mo-

tion into the converged positions. In the iterative cycle there is simply a succession of finite speeds. This can be seen, for example, by the mean value relaxation form in Eq. 34.

With a spatial grid, the parabolic partial differential equation of Eq. 52 assumes the form

$$K_i \frac{ds_i}{dt} = w_{i+1/2}(s_{i+1} - s_i) - w_{i-1/2}(s_i - s_{i-1}) \tag{53}$$

where the grid point velocity at i is explicitly displayed. In an algorithm, the positive numbers $w_{i+1/2}$ and $w_{i-1/2}$ are first computed and then the s_i is moved accordingly. When the original s_i lies between s_{i-1} and s_{i+1}, both the i+1/2 term $w_{i+1/2}(s_{i+1} - s_i)$ and the i-1/2 term $w_{i-1/2}(s_i - s_{i-1})$ are positive. If the i+1/2 term exceeds the i-1/2 term, then s_i must move in the positive direction towards s_{i+1} to simultaneously shrink $(s_{i+1} - s_i)$ and expand $(s_i - s_{i-1})$. If the i+1/2 term is less then the i-1/2 term, then the velocity is in the negative direction towards s_{i-1}. If both terms are equal, then the forces balance and thus there is no motion. When the original s_i for an algorithmic step lies outside of the interval $s_{i-1} < s < s_{i+1}$, the motion is directed back towards the interval. For example, if the original s_i exceeds s_{i+1}, then both terms are negative and thus the velocity of eq. 53 is also negative, sending s_i towards or past s_{i+1}. Moreover, when the ordering of the points s_k is monotonically decreasing rather than increasing as in the discussion just given, the same conclusions result with only a change in directions.

As a further observation of Eq. 53, it is noted that when the velocity is zero, the tridiagonal system of Eq. 29 is obtained. That system is rapidly approached when the diffusion rate $1/K_i$ is sufficiently large. In the evolution towards the earlier tridiagonal system, a temporal discretization must be considered. The simplest is the single step from time t_n to time t_{n+1} together with a backward temporal difference over the time step $\Delta t = t_{n+1} - t_n$. With the time indicated by a superscript, the parabolic partial differential equation is approximated by the fully implicit system

$$K_i^n\left(\frac{s_i^{n+1} - s_i^n}{\Delta t}\right) = w_{i+1/2}^n(s_{i+1}^{n+1} - s_i^{n+1}) - w_{i-1/2}^n(s_i^{n+1} - s_{i-1}^{n+1}) \tag{54}$$

which assumes the tridiagonal form

$$w^n_{i+1/2}s^{n+1}_{i+1} - (w^n_{i+1/2} + w^n_{i-1/2} + \frac{K^n_i}{\Delta t})s^{n+1}_i + w^n_{i-1/2}s^{n+1}_{i-1} = -\frac{K^n_i}{\Delta t}s^n_i \qquad (55)$$

Upon inspection, strictly positive values of K^n_i ensure diagonal dominance. Moreover, very large values will force an essentially diagonal form and thereby cause the i-th point to essentially be frozen: that is, an essential reduction to $s^{n+1}_i = s^n_i$. At the opposite extreme, a zero value of K_i will give the earlier tridiagonal form of Eq. 29. As from the earlier form, the same manipulations will lead to a mean value relaxation statement. With a pointwise update from only n-level information (point Jacobi relaxation), the statement becomes

$$s^{n+1/2}_{i+1/2} = \frac{w^n_{i+1/2}s^n_{i+1/2} + \frac{K^n_i}{\Delta t}s^n_i + w^n_{i-1/2}s^n_{i-1/2}}{w_{i+1/2} + \frac{K^n_i}{\Delta t}s^n_i + w_{i-1/2}} \qquad (56)$$

and reduces to Eq. 32 when K^n_i vanishes. In continuation, the least squares statement of Eq. 36 extends into the form

$$A = \sum_{i=1}^{N-1}\{K_i\Delta t(\frac{s^{n+1}_i - s^n_i}{\Delta t})^2 + w^n_{i+1/2}(s^{n+1}_{i+1} - s^{n+1}_i)^2\} \qquad (57)$$

and reproduces the tridiagonal form when it is minimized with respect to variations in s^{n+1}_i. Here, of course, $K_1 = 0$. The pattern of development here clearly proceeds through the more general variational format.

Further accuracy in representing the parabolic partial differential equation can also be developed in a similar manner. This would come from writing the finite difference equation about the (n+1/2)-time level in a Crank-Nicolsen form. The result would be second order rather than first order in time provided that the weights are appropriately linearized about the time level t_n. The associated tridiagonal form would then have more complication in the generation of its coefficients and source term. As a consequence, the simpler first-order accurate form is the likely preferred form.

THE SPECIFICATION OF COEFFICIENTS IN THE LINEAR WEIGHT

Fractions

When linear weights are employed for equidistribution, there is a distinct advantage in using the backward global integral statement as opposed to some of the others. The advantage is the capability to more precisely specify the level of importance of the various clustering quantities in the weight. This comes from the linearity in both the integration and the weight. With the general linear weight of Eq. 1, the integral of Eq. 6 becomes

$$F(s) = H_0(s) + \sum_{i=1}^{m} c_i H_i(s) \tag{58}$$

where

$$H_0(s) = s - s_{min}$$

and

$$H_i(s) = \int_{s_{min}}^{s} M_i d_x$$

Upon evaluation at $s = s_{max}$, the equation becomes a relationship between total amounts. With the total length

$$L = H_0(s_{max}) = s_{max} - s_{min} \tag{59}$$

and the total amount of each quantity

$$I_i = H_i(s_{max}) \tag{60}$$

the total weight integral is given by

$$F(s_{max}) = L + c_1 I_1 + ... + c_m I_m \tag{61}$$

Since each term is positive and represents a part of the total weight, a division by the total weight results in the fractional decomposition

$$1 = f_0 + f_1 + \ldots + f_m \tag{62}$$

where

$$f_0 = \frac{L}{F(s_{max})} \tag{63}$$

is the fractional contribution from the total length and

$$f_i = \frac{c_i I_i}{F(s_{max})} \tag{64}$$

is the fractional contribution from the total amount of the i-th quantity. Since each quantity must have a specified level of importance c_i and since a direct specification would be somewhat ad hoc, a more precise strategy is to specify the fractions of Eq. 64 and then solve for the consequent c_i. With these specifications, the decomposition of Eq. 62 produces a value for the fractional amount of length f_0 which then can be inserted into Eq. 63 to yield the total weight integral in the form

$$F(s_{max}) = \frac{L}{1 - f_1 - \ldots - f_m} \tag{65}$$

Upon substitution into Eq. 64, the weighting coefficients are then given in terms of specified values as

$$c_i = \frac{Lf_i}{(1 - f_1 - \ldots - f_m)I_i} \tag{66}$$

When these coefficients are employed in the integral F(s) of Eq. 58, the corresponding backward global integral statement of Eq. 6 becomes

$$\frac{\xi - \xi_{min}}{\xi_{max} - \xi_{min}} = \sum_{i=0}^{m} f_i \frac{H_i(s)}{H_i(s_{max})} \tag{67}$$

where f_0 has now been reinstated. Upon observing the form of Eq. 67 together with the argument leading up to it, the role of f0 is seen to be assumable by

any of the other fractions appearing in Eq. 62: that latter expression merely provides a partition of unity which determines any fraction when the other m fractions are specified. The fraction so determined then plays the same role as f_0 above. To guard against the possible indeterminate situation should $H_i(s_{max})$ vanish, a safe practice is to replace each $H_i(s_{max})$ by $H_i(s_{max}) + \varepsilon$ for some very small positive number ε. The development of the above fractional specifications was originally derived by Dwyer [17] for cases up to $m = 2$ fractions and was then extended by Eiseman [8] to cover cases with any number of fractions, m.

Level of Significance

When the consideration is shifted from the proportion of the weight $F(s_{max})$ to the proportion of the length L, the specified fractions above must undergo an adjustment to account for multiple effects that occur at the same location. The spirit of the adjustment can be readily examined in the case of just one effect since it must then compete with the uniformity term that produced L. With $m = 1$, the first step is to assign a level α_1 above which M_1 is considered to be significantly large and thus important enough to be counted. The effect here is to remove the minor level activity in M_1 that is typically spread over large regions and to thereby emphasize only the local regions of interest. The assignment of α_1 then defines the set of local regions as $\{s: \ M_1 > \alpha_1\}$. The Lebesque measure [18] of the set represents the length over which M_1 is significant. With

$$L_1 = \text{meas}\{s: \ M_1 > \alpha_1\} \tag{68}$$

the length where M_1 is unimportant is up to the order α_1 just $(L-L_1)$. By contrast, the total length where M_1 is important is counted twice: once by L_1 and once by $c_1 I_1$ in correspondence with 1 and $c_1 M_1$ in w. As a consequence, the total weight is partitioned in the form

$$F(s_{max}) = (L-L_1) + (c_1 I_1 + L_1) \tag{69}$$

and fractions are now defined by

$$f_0 = \frac{L - L_1}{F(s_{max})} \tag{70}$$

and

$$f_1 = \frac{c_1 I_1 + L_1}{F(s_{max})} \tag{71}$$

With $f_0 = 1-f_1$, the equations are solved as before to produce the weighting coefficient which now becomes

$$c_1 = \frac{f_1 L - L_1}{(1-f_1) I_1} \tag{72}$$

The need for the protective ε in Eq. 67 can be dropped here since a vanishing I_1 would mean that $L_1 = L$ and hence it would then be reasonable to set $c_1 = 0$ thereby removing the effect of M_1 altogether. On reflection of the above process, one may also note a parallel resemblance to studies in the area of statistics where it is common practice to assume a level of significance [19] for events.

Minimum and Maximum Spacing

When there is only one coefficient to be specified, its specification can be translated into the more intuitive and geometric specification of a relationship between the actual minimum and maximum spacing. This is derived from the direct grid statement of Eq. 4 by expressing it under both minimum and maximum conditions. The result is given by

$$w_{max}(\Delta s)_{min} = w_{min}(\Delta s)_{max} \tag{73}$$

When the general linear weight of Eq. 1 is written for a single coefficient ($m = 1$) in the form

$$w = 1 + cM \tag{74}$$

and is inserted into the relationship of Eq. 73, that coefficient is then determined by

$$c = \frac{R - 1}{M_{max} - M_{min} R}$$

where (75)

$$R = \frac{(\Delta s)_{max}}{(\Delta s)_{min}}$$

As a consequence, the ratio R of the actual maximum to minimum spacing can be specified and will determine a valid coefficient c provided that

$$R < \frac{M_{max}}{M_{min}} \tag{76}$$

This condition merely states that the global variation in specified spacing cannot exceed the corresponding variation in the weight. The specification of such ratios R is due to Nakahashi and Diewert [20]. Upon observing the above development, it is also worth noting that the same argument can be applied to the general linear weight when only one coefficient remains to be determined.

THE ATTRACTION TO A GIVEN GRID ON A CURVE

In a number of circumstances there is a need to have some adherence to a specified distribution of points along a curve. These needs will typically arise, for example, when an expanding grid is desired for far-field conditions, when a nearly orthogonal grid is desired for enhanced grid structure, and when a smooth temporal evolution is also desired. While cases such as the expanding grid can be given by either specifying a weight function or a grid, the structural and temporal properties appear only in the form of a grid. The usual grid for the structural aspects comes from the computation of orthogonal trajectories [21]. They evolve from the adjacent and most current coordinate curves to our given curve which here is assumed to be a part of a higher dimensional grid. The grid for the smoothness in temporal evolution comes from the previous grid or grids in a time-like sequence of grids.

Rather than considering further situations in which it is desirable to adhere to given grids, we shall next consider how to accomplish such adherence. This will proceed first in a direct fashion by inverse equidistribution, then by diagonal dominance, and finally in several variational forms.

Inverse Equidistribution

The most direct method of providing attraction to a given grid is to perform the equidistribution process in reverse. That is, we derive the weight function from the given grid. The result is an approximation to the derivative of the desired distribution function had it been available in a smooth continuum form rather than as a grid. Upon applying the weight in a direct sense, the specified grid will be reproduced up to errors in the derivative approximation and in the subsequent quadrature for the direct equidistribution. Assuming that such errors are sufficiently small, the weight derived from the grid is then viewed as only an attracting mass to be inserted into another weight function where it must then compete with other attracting masses. In terms of the general linear weight of Eq. 1, the mass is one of the functions M_i. As the corresponding coefficient c_i increases, the resulting grid approaches the approximate reproduction of the specified grid.

The forces in a weight function, however, are applied to directly control the spacing between the grid points rather than the actual positions of particular grid points. In the general situation, it is the local concentrations of grid points which is achieved, not their actual positions. For any local concentration, grid points are attracted from the entire curve to satisfy the local spacing requirements. This local attraction will cause a shift to appear in all points. As a consequence, that shift will directly appear in the situation when the local attraction is balanced against an attraction to a prescribed grid. While the spacing in the prescribed grid will be rendered in a fair and competetive manner, the shift will lead to unacceptable results for the actual positions. The problem here is that the attraction to a specified grid is position-sensitive while the equidistribution process is position-insensitive. The primary cause of the problem is that local constants of integration are being effectively introduced from each local cluster. The only apparent remedy is then to employ an iterative scheme about the equidistribution process. At each stage, an equidistribution of the current weight would produce results that in turn would lead to a new weight for the next cycle. In this framework, the inverse equidistribution strategy provides a distribution function that would enter the iterative procedure. Although such a procedure has not yet been developed, the process of inverse equidistribution has.

To more closely examine the inverse equidistribution process, we shall explicitly derive weights at the current grid points

$$s_1 < s_2 < \ldots < s_N \tag{77}$$

for the purpose of providing pointwise movement towards a specified grid

$$s_1 = r_1 < r_2 < \ldots < r_N = s_N \tag{78}$$

From a continuum viewpoint, the inverse equidistribution is accomplished by setting

$$w = \frac{1}{r_\xi} \tag{79}$$

for then the constancy of ws_ξ becomes the constancy of s_r and hence a linear relationship between s and r that becomes an equality since $s_1 = r_1$ and $s_N = r_N$.

From a discrete constructive viewpoint, we invert the direct grid statement of Eq. 4 to get

$$w_{i+1/2} = \frac{1}{r_{i+1} - r_i} \tag{80}$$

from the specified grid of Eq. 78. The unity in the numerator was chosen for convenience: any other nonzero constant would produce the same result since it would disappear in the ratio of integrals in Eq. 6. With Eq. 80, the weights are established at $r_{i+1/2}$ and must either be used there or be transferred to the current points s_k. The latter case may proceed either directly or with an intermediate transfer to points r_i. In the direct approach, the derivative profile is obtained by an interpolation between half points $r_{i+1/2}$ and an extrapolation to end points r_1 and r_N.

The simplest profile is given by a piecewise-linear interpolation that is extended by a horizontal extrapolation attached to either end. The piecewise linear transfer of weights is accomplished by searching the points

$$r_1 < r_{1/2} < \ldots. < r_{N-1/2} < r_N \tag{81}$$

to find the interval that contains s_k for each k. If the interval is either the leftmost or rightmost one, then the weight is respectively either $w_{1/2}$ or $w_{N-1/2}$ as determined by Eq. 80. Otherwise, we have $r_{m-1/2} \leq s_k < r_{m+1/2}$ for some $m = m(k)$ and hence the weight

$$\bar{w}_k = w_{m-1/2} + \frac{s_k - r_{m-1/2}}{r_{m+1/2} - r_{m-1/2}}(w_{m+1/2} - w_{m-1/2}) \tag{82}$$

This weight can then be used as an attractive mass at the point s_k within the linear weight of Eq. 1.

To employ it effectively, the values at s_k must represent a fair sampling or else the basic features of the weight could be lost or distorted. A poor representation here is certainly likely to increase the quadrature error upon application. This circumstance can readily arise if there are too many points s_k in any given interval from Eq. 81. There would then be regions where the weight is represented too linearly. As a consequence, it is advisable to use weights from Eq. 82 only when the current (Eq. 77) and the specified (Eq. 78) grids are nearly interlaced. Even still the inaccuracy may be too large. This would then suggest an application only when the r_i and s_i are nearly equal.

Being somewhat forced to consider cases where the s_i and r_i are not too far apart, a further option has arisen. When the backward global integral statement of Eq. 6 is being employed, the curve together with the distribution function is being represented at existing locations s_i of Eq. 77. Upon inspection, those locations could readily be replaced by the locations associated with the specified grid for r_i given by Eq. 78. With that replacement, the quadrature of Eq. 8 produces a backward global transformation from Eq. 6 which precisely moves s_i into r_i. Under such a replacement, other weighting quantities together with the curve description must undergo a transfer to the new positions r_i and thus represents another source for some error. This leads to the enrichment strategy whereby the points r_i are merely included along with the original s_i in the curve representation. In essentially the same manner, the exact reproduction of the given grid locations r_i is possible. The derivation represents a modest extension. These developments leading to weights that reproduce specified grids were jointly constructed with this author and Michael Bockalie.

Diagonal Dominance

In contrast to the inverse equidistribution approach above, we may place the attracting mechanism outside of the weight rather than in it. With this objective in mind, we return to the direct finite difference statement of Eq. 29 and then add $D_i(s_i - r_i)$ to the right-hand side. With the assumption that the numbers D_i are nonnegative, we obtain the diagonally dominant tridiagonal system

$$w_{i+1/2}s_{i+1} - (w_{i+1/2} + D_i + w_{i-1/2})s_i + w_{i-1/2}s_{i-1} = -D_i r_i \qquad (83)$$

which as before is fully defined when boundary conditions such as those given by Eq. 30 are assumed. As D_i increases, the diagonal dominance is intensified. In the limit, the system approaches a simple diagonal form and this causes each current grid point s_i of Eq. 77 to approach the corresponding specified grid point r_i of Eq. 78.

By providing an index to D_i, we have included the possibility to approach the specified grid at a nonuniform rate: a uniform rate would come from setting $D_i = D$. As a consequence, we may also selectively approach only parts of the specified grid by smoothly adjusting the values of D_i from zero to local maximums at the desired locations. The smoothness is viewed by considering a function D of ξ to be smooth: the values at $\xi = i$ being denoted by D_i.

Moreover, we may also consider a sequence of specified grids together with a corresponding sequence of such functions. By adding the associated sequence of terms to the right-hand side of Eq. 29 and following the above reasoning for a single term, the simultaneous attraction to the sequence of grids can be considered. The relative sizes for $D(\xi)$ in each term give the relative importance of that term. In the application, there is then a balancing of importance for the attraction to the separate grids. This is the same type of balancing operation as in the previous case with inverse equidistribution where each specified grid was represented by an attracting mass within the weight.

To explictly examine the relationship between the equidistribution and diagonal dominance approaches, we consider the case of attraction to one prescribed grid of points r_i as displayed in Eq. 78. In the diagonal dominant approach, the attraction is represented by Eq. 83. Since the weights are

functions of s rather than i and since the positions represented by s_i are unknown, iteration is required. As in the earlier tridiagonal approach, weights are evaluated from current values of position s_i and then the tridiagonal system is inverted to obtain new locations s_i from which the process is repeated. Now let the fixed location r_i in Eq. 83 be given by the relative location

$$r_i = \alpha_i s_{i+1/2} + (1 - \alpha_i)s_{i-1/2} \tag{84}$$

through α_i. The half-point values are the same as in Eq. 33 and are determined from the current locations s_i in the iterative cycle. By inserting the local representation of Eq. 84 into the diagonally dominant tridiagonal form of Eq. 83 and regrouping terms we get the new adjusted weights

$$\begin{aligned} \bar{w}_{i+1/2} &= w_{i+1/2} + \frac{1}{2}\alpha_i D_i \\ \bar{w}_{i-1/2} &= w_{i-1/2} + \frac{1}{2}(1 - \alpha_i)D_i \end{aligned} \tag{85}$$

for the original tridiagonal form of straight equidistribution as given in Eq. 29. While this yields a local equidistribution statement for each i, the global form is somewhat altered. That is because the half-point weight evaluations from Eq. 85 will vary with the equation. In particular at i+1/2, the application of Eq. 85 yields

$$\bar{w}_{i+1/2} = w_{i+1/2} + \frac{1}{2}\alpha_i D_i$$

and (86)

$$\bar{w}_{i+1/2} = w_{i+1/2} + \frac{1}{2}(1 - \alpha_{i+1})D_{i+1}$$

respectively for equations about i and i+1. To employ the weights in any of the quadrature forms presented, we need uniqueness and this leads to the requirement that

$$D_{i+1} = \left(\frac{1 - \alpha_{i+1}}{\alpha_i}\right)^{-1} D_i \tag{87}$$

which then determines the diagonal scaling up to a single multiplicative con-

stant. With this determination, we then immediately observe that any of the basic forms for equidistribution can be applied within such an iterative cycle of global updates. Because of the specialized choice for D_i in Eq. 87, it is clear that the diagonal dominance approach has some added and useful generality. The freedom to choose the form of D means that the attraction to the r_i can be locally controlled.

The Variational Format

In correspondence with the previous forms of attraction to a prescribed grid, there are variational statements both with the attraction in the weight and with it outside of the weight. With it in the weight, we can simply use reciprocal derivative of Eq. 79 in the integral of Eq. 49 that represents the continuum version of the discrete least squares statement. This was the starting point for Steinberg and Roache [15] who considered such specified grids to be "reference grids" from which various metrical properties could be extracted and then directly used to form weights.

With the attraction outside of the weight, the minimization of

$$\int [ws_\xi^2 + D(s-r)^2]d\xi \tag{88}$$

provides the same form as in Eq. 83 with the same function $D(\xi)$ that controls the attraction of s to a specified r. The r assumes the previous role of a given grid and is thus a function of the spatial location s rather than the eventual grid point counter ξ. This clearly penalizes the basic equidistribution in favor of attraction to r as D increases in size. In a similar spirit, this statement can also be applied to derivatives. In particular, we have the form

$$\int [ws_\xi^2 + D(s_\xi - r_\xi)^2]d\xi \tag{89}$$

that was considered by Bell and Shubin [22] to enforce temporal smoothness. This was achieved by taking r to be the previous time level of s in the solution of a one-dimensional time-dependent PDE. In the application, the weight was also smoothed: this being done by a spatial filter. With D as a constant, the Euler Equation reduced to the form

$$ws_\xi + D(s-r)_\xi = \text{constant} \tag{90}$$

which then can be simply integrated. This can be contrasted with the more complex integration that comes from Eq. 88.

EVOLUTIONARY FORCES

As in the discussion on the various forms for equidistribution in one dimension, we are now in a position to examine some of the forces which can be applied to some advantage in the evolutionary setting. The forces will appear both in the context of equidistributed weights and outside of such weights. Within the weights, the local application from a global determination will be witnessed. Outside of the weights, the attraction to a specified grid will be examined, the minimum and maximum spacing controls will be displayed, and the control of cell eccentricity will be explored. The role of diagonal dominance, here, is seen to arise again in a combined sense.

Globally Determined Weights

In the earlier consideration of weight functions, the evaluations at half-point locations were typically assumed to come from only the nearest neighbors as given in Eq. 5 of the direct grid statement for equidistribution. Rather than using only that simple local average, we shall consider a global determination for the local weights about the point $\xi = i$ for application about the corresponding s_i. For this purpose, suppose that we have established some error indicator e_i at each point. By using an average error value such as

$$\bar{e} = \frac{1}{N} \sum_{j=1}^{N} e_i \tag{91}$$

the half-point weights about i can be written in the form

$$w_{i+1/2} = c \sum_{k=i+1}^{N} \frac{e_k - \bar{e}}{(k-i)^{\alpha}}$$

$$w_{i-1/2} = c \sum_{k=1}^{i-1} \frac{e_k - \bar{e}}{(i-k)^{\alpha}} \qquad (92)$$

where c is a prescribed positive scaling constant and α is an attenuation control. The effect of α is to give more importance to the error deviations which are closer to i. As α increases, the influence of the local error grows while the more distant points are more attenuated. When the weights of Eq. 92 are inserted into the evolutionary equidistribution statement of Eq. 53, we obtain the grid velocity equation.

$$\frac{K_i}{c} \frac{ds_i}{dt} = (s_{i+1} - s_i) \sum_{k=i+1}^{N} \frac{e_k - \bar{e}}{(k-i)^{\alpha}} - (s_i - s_{i-1}) \sum_{k=1}^{i-1} \frac{e_k - \bar{e}}{(i-k)^{\alpha}} \qquad (93)$$

This equation represents a local equidistribution since the weights about i-1 and i+1 will produce generally distinct values at the respective half points i-1/2 and i+1/2. The circumstance is just a parallel to the case with the diagonally dominant attraction to a given grid as represented in Eq. 86.

With the evolutionary equidistribution of Eq. 93, we are very close to the grid velocity model that was proposed by Rai and Anderson [23] on the basis of physical intuition. They formed their model on the basis of an analogy with gravitational forces. That, for example, is where the attenuation from an inverse power law in distance from i was conceived. In distinction from what is presented here, their model is obtained from Eq. 93 by replacing both $(s_{i+1} - s_i)$ and $(s_i - s_{i-1})$ with a single factor such as $(s_{i+1} - s_{i-1})$ and by combining K_i/c into a single prescribed constant to uniformly scale the right-hand side.

Returning to the choice of weights in Eq. 92, it should be noted that the essential character is the attenuation with respect to distance from the location of application. It is not the particular choice of functional form: many other forms could be employed for the same purpose. The same comment also applies to the implementation of the error indicator: the basic feature is that it can change sign. While this would be undesirable in the nonevolutionary context, it is beneficial here and provides and additional restoring

force.

The Attraction to a Given Grid

With the understanding that a given grid such as that displayed in Eq. 78 comes from a given distribution function r(s), the basic evolutionary parabolic partial differential equation of Eq. 52 can be appropriately adjusted to provide an attraction to the fixed r(s) and to the fixed grid of r_i values. Under the adjustment, it assumes the form

$$Ks_t = (ws_\xi)_\xi + D(s-r) \tag{94}$$

where the attraction is represented by the scaled quantity s-r. When the successive steps leading up to the fully discrete form in Eq. 55 are repeated, another tridiagonal form is obtained and is given by

$$w^n_{i+1/2}s^{n+1}_{i+1} - (w^n_{i+1/2} + w^n_{i-1/2} + \frac{K^n_i}{\Delta t} + D^n_i)s^{n+1}_i + w^n_{i-1/2}s^{n+1}_{i-1} =$$
$$= -\frac{K^n_i}{\Delta t}s^n_i - D^n_i r_i \tag{95}$$

In a combined sense, diagonal dominance is achieved from both the temporal aspect and the attraction to r_i. The respective effects, however, are essentially separated: as s_i converges, the influence of K_i decreases while that from D_i remains. The reason is because the K_i terms converge towards an exact cancellation while the D_i terms generally do not. On comparison with the earlier diagonally dominant attraction of Eq. 83, the evolutionary form presented in Eq. 95 represents a natural extension.

Minimum and Maximum Spacing

At this stage, the use of minimum and maximum spacing has been developed only for a linear weight with only one constant to be determined. That determination appeared only in the form of a ratio between minimum and maximum spacing and was given by Eq. 75. Neither the actual spacing nor the multiple constant environment were considered. In contrast, such a consideration can be readily developed in the temporal setting.

To develop the required forces, the definition of the parabolic partial differential equation is again extended by the inclusion of new terms on the right-hand side. The generic extended form is given by

$$Ks_t = (ws_\xi + A)_\xi + B \tag{96}$$

where forces over the velocity can be inserted as terms in either A or B. In general, A and B represent sums of the forces which do not readily fit into the equidistribution of w with respect to s. While the position of A appears to be somewhat close to that equidistribution, it gives us the opportunity to rationally create similar forces without considering the factor of some Δs as would appear in the weight w. The location B allows for an even more direct insertion of force as has already occurred in Eq. 94 when the (s-r) term appeared. That term provided an attraction of s_i to r_i. In the overall circumstance, the various forces in w, A, and B all act in a concerted fashion to provide a grid with a multitude of desirable attributes.

One such attribute is the bound on minimum and maximum spacing. This is accomplished with terms in the position of A in Eq. 96. The actual construction, however, is more readily accomplished in the finite difference form which is given by

$$K_i \frac{ds_i}{dt} = w_{i+1/2}(s_{i+1} - s_i) - w_{i-1/2}(s_i - s_{i-1}) + A_{i+1/2} - A_{i-1/2} + B_i \tag{97}$$

This represents the extended form of the earlier finite difference equation, Eq. 53. In the case of maximum spacing, the term

$$\Omega_{j+1/2} = c\left\{\frac{s_{j+1} - s_i}{(\Delta s)_{max}}\right\}^\alpha \tag{98}$$

is considered and, within A, yields the contribution

$$\Omega_{i+1/2} - \Omega_{i-1/2} = c\left\{\left(\frac{s_{i+1} - s_i}{(\Delta s)_{max}}\right)^\alpha - \left(\frac{s_i - s_{i-1}}{(\Delta s)_{max}}\right)^\alpha\right\} \tag{99}$$

The positive constants c and α are respectively for scaling and intensity.

Here, we assume the monotonically increasing order for s_i as displayed in Eq. 77. If the spacing $(s_{i+1} - s_i)$ exceeds the specified $(\Delta s)_{max}$, then the corresponding ratio appearing in Eq. 99 exceeds unity and is further amplified by α which is usually about 4. As a consequence, the first term in Eq. 99 gives the velocity of s_i a positive contribution. This is then a force to move s_i in the positive direction towards s_{i+1}. Such a movement, of course, will shrink the interval $(s_{i+1} - s_i)$ which had exceeded the specified value $(\Delta s)_{max}$. Likewise, had $(s_i - s_{i-1})$ exceeded $(\Delta s)_{max}$, then the second term in Eq. 99 would have caused force in the negative direction to push s_i towards s_{i-1} to thereby shorten that interval. If the intervals on both sides of s_i exceed $(\Delta s)_{max}$, then the forces may balance each other. However, in a point iterative sense, a translation in i will eventually produce an unbalanced situation that will successively be propagated to produce the desired result.

In a similar manner, the minimum possible spacing is enforced with a term of the form

$$A_{j+1/2} = - c\left(\frac{(\Delta s)_{min}}{s_{j+1} - s_j}\right)^{\alpha} \tag{100}$$

where c and α are the same sort of positive constants and $(\Delta s)_{min}$ is the specified lower bound on the spacing. As above, we get the contribution

$$A_{i+1/2} - A_{i-1/2} = c\left\{\left(\frac{(\Delta s)_{min}}{s_i - s_{i-1}}\right)^{\alpha} - \left(\frac{(\Delta s)_{min}}{s_{i+1} - s_i}\right)^{\alpha}\right\} \tag{101}$$

to the velocity of s_i. Because of the minus sign, the terms are now interchanged. When $(s_{i+1} - s_i)$ falls below $(\Delta s)_{min}$, the velocity contribution in the second term becomes important and pushes s_i in the negative direction towards s_{i-1} and thus enlarges $(s_{i+1} - s_i)$. When $(s_i - s_{i-1})$ falls below $(\Delta s)_{min}$, a significant contribution comes from the first term to push s_i in the positive direction towards s_{i+1} in order to enlarge $(s_i - s_{i-1})$. These terms are due to Winkler, Mihalas, and Norman [24] and were called floors and ceilings on the spacing.

Eccentricity

A certain amount of smoothness in the gradations of the grid can be controlled by a force which tries to equally distribute the ratios of eccentricity in adjacent spacings over the space of indices $\xi = i$. This equidistribution over ξ rather than s is represented by the function A in Eqs. 96-97. At a point j, the spacing eccentricity is given by

$$\varepsilon_j = \frac{s_{j+1} - s_j}{s_j - s_{j-1}} \tag{102}$$

and is not entirely suitable for immediate use because of the directional bias represented by a separation of j-1 and j+1 values between the denominator and the numerator. Otherwise, it would have been a good control to provide smooth gradations in spacing. Fortunately, the eccentricity ratio $\varepsilon_{j+1}/\varepsilon_j$ removes the directional bias and provides the same control. That ratio term is given by

$$\Psi_{j+1/2} = c\,\frac{(s_{j+1} - s_j)^2}{(s_{j+2} - s_{j+1})(s_j - s_{j-1})} \tag{103}$$

where c is a positive scaling constant. Upon equidistribution over the i, we would get the geometric mean

$$\varepsilon_j = \sqrt{\varepsilon_{j-1}\varepsilon_{j+1}} \tag{104}$$

to provide the growth rate. In such a situation, the distribution would be determined solely from the end conditions which would also require an extra point at each end. The extra points are needed for end-point spacings.

When the eccentricity ratio term of Eq. 103 is inserted into the parabolic form of Eq. 97, the force on the velocity of s_i comes from the difference

$$\frac{-(s_{i+1} - s_i)^2}{(s_{i+2} - s_{i+1})(s_i - s_{i-1})} - \frac{(s_i - s_{i-1})^2}{(s_{i+1} - s_i)(s_{i-1} - s_{i-2})} \tag{105}$$

Now suppose that in a point iterative process, s_i were to miss this desired equidistribution by getting too close to s_{i-1}. Then the interval length $(s_i - s_{i-1})$ in Eq. 105 would be too small. The effect would be to enlarge the first term and shrink the second. That enlargement can be substantial since the interval smallness then appears in a denominator. In fact, this is a control by means of a singularity; namely, a pole. As a consequence, the first term contributes a positive contribution to force s_i towards s_{i+1} thus enlarging the offending small interval $(s_i - s_{i-1})$. Likewise, the second term provides a restoring force when the point s_i is too close to s_{i+1}. This is just the opposite situation. The form of smoothness here was developed by Winkler, Mihalas, and Norman [24].

METRIC NOTATION

In the discussion of the various adaptive methods in higher dimensions, the fundamental properties of each object can be more clearly and concisely stated with the use of metric notation than without it. As a consequence, we shall describe just enough of it for our purposes. In two dimensions, we associate 1 with ξ and 2 with η. The differential element of arc length ds is then given in a plane by

$$ds^2 = g_{11}d\xi^2 + 2g_{12}d\xi d\eta + g_{22}d\eta^2 \tag{106}$$

where

$$\begin{aligned} g_{11} &= r_\xi \cdot r_\xi = x_\xi^2 + y_\xi^2 \\ g_{12} &= r_\xi \cdot r_\eta = x_\xi x_\eta + y_\xi y_\eta \\ g_{22} &= r_\eta \cdot r_\eta = x_\eta^2 + y_\eta^2 \end{aligned} \tag{107}$$

and $r = (x,y)^T$ is the position vector. This form of ds^2 can be readily verified by a chain rule expansion of $dx^2 + dy^2$. The matrix of elements g_{ij} is called the metric because it defines the distance measurements with respect to the coordinates (ξ,η). In a straightforward manner, the determinant

$$g = \det(g_{ij}) \tag{108}$$

is also seen to be the square of the Jacobian

$$J = x_\xi y_\eta - x_\eta y_\xi \tag{109}$$

For grids, this gives the cell sizes in the physical (x,y)-space that correspond with the fixed uniformly sized cells in the computational or logical space (ξ,η). When the Jacobian is nonvanishing, the inverse of the metric matrix (g_{ij}) can be computed and has elements that are typically written in the superscripted format g^{ij}. In continuation, the metric notation is equally valid in higher dimensions. For surfaces in three dimensions, we merely extend Eq. 107 by including terms for z that come from the dot products between combinations of the natural curvewise tangents r_ξ and r_η where $r = (x,y,z)$. For volumes in three dimensions, we just associate 3 with a third variable ζ and then repeat the above discussion where, of course, the expressions are slightly longer but the meaning remains the same. For a more compact notation, we shall sometimes use ξ_1, ξ_2, ξ_3 for ξ, η, ζ and X_1, X_2, X_3 for x, y, z. Further discussion can be found in virtually any text on differential geometry. A development specifically aimed at grid generation can be found in either of the monographs by Eiseman [25] and Warsi [26].

CURVE BY CURVE METHODS

With a basic understanding established, we next proceed to consider higher dimensions. The most direct extension into higher dimensions is to apply adaptivity on a curve by curve basis and to cycle through one or more coordinate directions. The methods of this description are called "alternating direction adaptive methods." From a geometric viewpoint, these methods operate by specifying the diagonal part of the metric tensor since each such diagonal entry is inversely proportional to a specified weight. This fundamental geometric framework provides a unified setting for all such methods and was presented by Eiseman [27]. The unity here derives from the fact that no matter how the metric specification is accomplished, the result must be the same up to the distinctive errors of approximation. The completeness in the specification comes from the curvature of space. This is most readily seen in

two dimensions where Gaussian curvature provides a single relationship between all three metric coefficients g_{11}, g_{22} and g_{12}. That means that the coordinates are analytically determined when g_{11} and g_{22} are specified as positive functions that may even involve g_{12}. In other words, the two available degrees of freedom are employed. Likewise, in three dimensions, the three available degrees of freedom are consumed. The typical specifications include gradient magnitudes, various forms for curvature (normal, geodesic, mean, second derivatives), cell properties (eccentricity, aspect ratio, lengths, area or volume), and the attraction to prescribed distributions (uniform, arbitrary, orthogonal alignment, or previous locations). Such specifications usually appear in a weight function. The most commonly utilized form is the linear one given by Eq. 1. Altogether, there is only one weight function employed for each degree of freedom consumed.

Metric Specification

The main distinguishing feature from the previously examined isolated curves is that the curves considered here appear as part of a higher dimensional coordinate system. To appropriately account for this fact, a directional index must first be attached to the earlier equidistribution statements for curves. In particular, the basic differential statement of Eq. 2 becomes

$$w_i ds_i = c_i d\xi_i \tag{110}$$

where i is now the coordinate directional index. For each given curve, c_i is a constant as earlier. However, as we go from curve to curve within the family for a given direction, that constant becomes a function of the transverse variables which govern this progression. As a consequence, the condition on c_i is that it vanish under differentiation with respect to the curve variable ξ_i.

Along the curve, the incremental distance measurement ds_i is just a special case of the measurement along arbitrary curves as indicated in Eq. 106 and its discussed extensions. Quite simply, with the exception of $d\xi_i$, all of the remaining differentials $d\xi_K$ for $k \neq i$ vanish. This leads to the form

$$ds_i = \sqrt{g_{ii}}\, d\xi_i \tag{111}$$

Upon substitution into the equidistribution statement of Eq. 110, a cancellation of $d\xi_i$ leads to the basic metric statement

$$g_{ii} = \left(\frac{c_i}{w_i}\right)^2 \tag{112}$$

which is the direct assertion that a prescribed weight w_i is equivalent to a prescribed metric coefficient g_{ii}.

When the curve constant c_i is isolated from the metric g_{ii} and the weight w_i by solving Eq. 112 for c_i^2, a differentiation in ξ_i then removes that constant and yields the partial differential equation

$$\frac{\partial g_{ii}}{\partial \xi_i} + \frac{2}{w_i}\frac{\partial w_i}{\partial \xi_i} g_{ii} = 0 \tag{113}$$

This corresponds to the earlier ordinary differential equation of Eq. 23. As in the case of Eq. 23, the data from the constant has been effectively transferred into boundary conditions. Upon substitution for g_{ii} in Eq. 113, the partial differential equation is readily observed to be second-order. To be explicit on this matter, we consider the case of the first equation in two-dimensional Euclidean space. There, the metric from Eq. 107 is just $g_{11} = x_\xi^2 + y_\xi^2$ and results in

$$x_\xi x_{\xi\xi} + y_\xi y_{\xi\xi} + \frac{(w_1)_\xi}{w_1}\left(x_\xi^2 + y_\xi^2\right) = 0 \tag{114}$$

The equation for the η-direction has the same form and is obtained similarly. Other forms cover curves on surfaces and higher spatial dimension. All of these come from Eq. 113. In comparison with the earlier case of an isolated curve, the equations here are for coordinate curves, and as a consequence, reflect their embedding within the coordinating system.

The Historical Development of Pure Curve by Curve Methods

To examine the historical roots of alternating direction adaptivity, we return to one dimension and then expand to higher dimensions and to multiple directions. Some of the earlier studies in one dimension were developed by

Winkler [28], Ablow and Schecter [29], and White [30,31]. Winkler [28] considered grid point movement directly in physical space in response to gradients in the monitor surface. All attributes for clustering and most for grid structure were given in the form of a linear weight. By contrast, White [30,31] developed his grid directly on the solution curve and thereby effectively used a weight of unity. When the arc length was expressed as an explicit function of physical space, he obtained the appropriate weight function for arc length. In extrapolating from here, he called the weights monitor functions. Although the term may be appropriately descriptive when everything is combined under a single integral, it is otherwise deficient. A more basic consolidation of the data and one that applies regardless of dimensionality is the concept of monitor surface as discussed in the introduction. Ablow and Schecter [29] preceded White and considered a linear weight with curvature that was applied relative to the arc length of the solution curve.

In the next stage of development, Dwyer et al. [17,32-35] considered the process of adapting the points along each coordinate curve in a fixed direction. In contrast to Winkler, White and Ablow and Schecter, he considered weight functions that depended upon positions in physical space. This was executed in a noniterative fashion by employing the backward global integral statement of Eq. 6 along each coordinate curve in the physical region. Again, linear weights were employed. In terms of Eq. 1, he considered up to two masses consisting of magnitudes for first and second derivatives of the monitor surface. These, however, provided only approximate gradient and curvature data: the variations along transverse coordinate curves were ignored. An advantage that evolved from the choice of linear spatially dependent weights was the capability to more rationally define the coefficients in the linear weights. With the transformation established by the global integrals, Dwyer [17] noted that the fractional contribution of each mass was merely a ratio of that mass integral to the total mass integral. Each integral, of course, was taken over the entire curve. Since each mass integral contains the associated coefficients, he was able to specify the fractions and solve for the coefficients. This was done for up to two masses. The general case was discussed earlier herein. As a consequence, each specified fraction resulted in a weight coefficient that dynamically adjusted from curve to curve.

In progressing from a single direction to multiple directions both Ablow [36] and Gnoffo [37] perfomred bidirectional studies. Ablow considered the

arc length along the curves on a monitor surface and proceeded to solve the equations

$$x_\xi x_{\xi\xi} + y_\xi y_{\xi\xi} + z_\xi z_{\xi\xi} = 0$$
$$x_\xi x_{\xi\xi} + y_\xi y_{\xi\xi} + z_\xi z_{\xi\xi} = 0 \qquad (115)$$

which come from Eq. 113 with constant weights and with $z = z(x,y)$. He employed an ADI procedure for the solution. As a consequence, he automatically clustered to high gradients. In contrast, Gnoffo viewed the monitor surface only from physical space, and thus had to explicitly use gradient information. Like Dwyer [33], he neglected transverse variations and considered derivative magnitudes along the given coordinate curve in physical space. He did not, however, consider derivatives beyond first-order. In the execution phase, the weight was viewed as a function of the grid point index (equivalently ξ or η) and the forward global integral statement of Eq. 15 was employed. That required iteration as discussed earlier. In the trapezoidal quadrature rule for the integral, each increment in ξ and η was unity and the result was a simple sum of reciprocal weights as given by Eq. 18. This is in contrast to the non-iterative statement where the quadrature is a sum of products between midpoint weight values and the corresponding interval arc lengths from Eq. 8. As a matter of terminology, he called this approach a spring analogy. Because of the iterative nature, however, the spring constants are not really constants since they must also change. At best, they may be viewed then as nonlinear springs.

In a more general study, Eiseman [27] consolidated and extended the previous work and presented a mathematical foundation for all such curve by curve methods. To be descriptive, these methods were then called alternating direction adaptive methods. The mathematical foundation was mentioned earlier. In brief terms, the known curvature of a region implies that metric specifications along each coordiante direction (Eq. 112) are enough to completely determine the metric which in turn can be employed to generate the grid by line integration. In the context of orthogonal grid generation, details on the line integration for coordinate positions can be found in Warsi and Thompson [38] and in Eiseman [21].

With the previous work separated between surface grids without weights and planar grids with weights, the more unified approach is to simultaneously

consider surfaces and weights. In the consolidated form, Eiseman [27] stated a preference for generating grids on monitor surfaces while using weights for the resolution of surface properties; albeit, the form applies equally well to viewing the surface from the physical region below it. The stated preference quite naturally derives from the fact that the accurate representation of a surface is more readily apparent in the surface grid than in the corresponding projected grid in the physical region. The reason is that the viewpoint is the location of arc length measurement which is more naturally taken and accurately controlled on the surface. In simple terms, it is reasonable to generate the grid directly upon the very object that must be given a good representation.

With the objective of providing a good representation for the monitor surface, the geometric parameters for the surface must be used along with the parameters which govern the grid quality in the sense of good structure. The basic surface properties are the bends or folds in the surface and are the surface boundaries which may also be bent in some manner. The basic measurement for surface bending in a given direction is the normal curvature. The measurement for the boundaries is the geodesic curvature. Each curvature detects only the desired property. Using the normal curvature in the direction of the current coordinate curve, Eiseman [27] observed the desired clustering effects for surface folds. In addition, the use of normal curvature was seen to provide some alignment of coordinate curves with folds in the surface. Moreover, the formulation used the general linear weight of Eq. 1 where normal and geodesic curvatures are balanced with the unity for uniformity and a mass for orthogonality attraction. The coefficient for geodesic curvature was presented in a form that decayed upon leaving boundaries so that bent interior curves appearing from the iteration would not unnaturally cause clustering. In the application, the backward global integral statement of Eq. 6 was employed along each curve.

The use of fractional specifications initiated by Dwyer [17] were also extended. In a basic sense, it was noted that the resolution of a property which appears over a fixed length could be depreciated by merely increasing the total length (or other masses). Thus a methanism was required to appropriately adjust the clustering intensity to treat the local property in the same manner, regardless of length. For normal curvatures we must somehow detect the likely presence of local clustering regions and the possible exces-

sive consumption of surface arc length. This was done by first forming the ratio R of the actual arc length of a coordinate curve on the surface to the minimum possible length. The latter is computed in the form of a Euclidean distance using the arc length of the projected curve and the change in altitude between endpoints. Upon application, a factor of the form tanh[D(R-1)] was applied to a constant fraction determined in the original manner. The fraction is now seen as a maximum possible fraction that would be employed in the most severe case within the family of all coordinate curves in a given direction. The constant D then gives the rapidity for which the specified maximum fraction is approached. More details are provided in Eiseman [27]. The extension of fractional specifications to include any number of masses in Eq. 1 and to consider what happens when distinct masses appear on the same interval are all discussed in the review by Eiseman [8] and in some detail here in the sections about Eqs. 58-72.

In several subsequent studies, Nakahashi and Deiwert [20,39,40] added a few more items of interest to the development of alternating direction adaptivity, and in addition, presented some rather good examples of adaptive simulation in aeronautics. The main contributed item is their incorporation of an orthogonality control outside of the weights. Given the arc length locations r_i of Eq. 78 corresponding to an orthogonal alignment with the previous curve, they added a term of the form $D_i(s_i - r_i)$ to the right-hand side of the tridiagonal system for the direct finite difference statement of Eq. 29. They arrived at Eq. 83 where an increase in D_i causes the diagonal dominance to grow and thus forces s_i to approach r_i. A variant is to balance the orthogonal locations with the straight line continuations through the previous curve. The purpose was to choose r_i so that the continued transverse curves would be smoother. In the applications, D_i actually varied inversely with respect to the distance from the previous curve at each i. This strengthened the attraction for closely spaced curves.

In three dimensions, there is another similar arc length location t_i and another term $E_i(s_i - t_i)$. There D_i+E_i is added to the diagonal and $D_i r_i + E_i t_i$ forms the right-hand side. In continuing the analogy with springs, they attributed the orthogonality control to torsion springs rather than to diagonal dominance. Although this control might at first appear to make the method distinctive, the fundamental fact still remains that a metric relationship is being specified along curves. In fact, as we saw in Eqs. 84-87,

the diagonal dominance control can be interpreted as a mass for orthogonality attraction as discussed in regard to the general linear weight of Eq. 1. When D_i is chosen as in Eq. 87, the quadrature, Eq. 8, of the backward global integral statement can also be employed because unique half-point weights are then determined.

Rather than a consistent use of linear weights, Nakahashi and Deiwert [40] considered weights where some of the specified constants appeared nonlinearly. This arose mainly because the minimum and maximum spacing could be specified by means of a single algebraic formula over the curve. This was distinct from their use of ratios as discussed in Eqs. 73-76. One constant was an exponent and had to be determined by iteration. By contrast, Winkler, Mihalas and Norman [24] gave upper and lower bounds upon the spacing within the context of the general linear weight of Eq. 1 and did not require an iterative determination of constants. This was possible because the spacing was only required to become close to minimum and maximum spacing along the curve rather than matching it precisely. The development is presented here about Eqs. 97-101.

It should be noted that the objective in setting bounds upon the spacing as represented by Winkler, Mihalas and Norman [24,41] and by Nakahashi and Deiwert [40] is essentially the same objective as specifying the fractional amount of quantities as represented by Dwyer [17] and Eiseman [8,27]. Both prescriptions merely attempt to set limits upon the finite distribution of points. In comparison, the spacing bounds are attractive because of their direct attachment the the actual spacing while the specification of the fractions are attractive because of their flexibility. With the fractions, the constraints upon spacing can be more effectively balanced against the other attributes which quite naturally enter into the same linear format. Moreover, those other attributes can also be precisely separated at the same time as the spacing requirements. This extensibility is not readily apparent in the format of spacing bounds. To consider it, the form presented by Winkler, Mihalas and Norman [24,41] would be preferred because it fits nicely into the context at the general parabolic evolution equation, Eq. 97.

Unlike the study of Eiseman [27], Nakahashi and Deiwert [20,39,40] did not acknowledge any defect in the application of the method. The only indication appears indirectly when they state that some corrective action is required when the orthogonally aligned arc length locations r_i or t_i fail to be

monotonic. Although the action was simply executed, the real problem was covered up. The real problem comes from the errors incurred in the numerous piecewise linear approximations to curves and their parameterizations. Various forms of the problem were illustrated by Eiseman [27] along with appropriate explicit corrective actions. In cases with rapid but not excessive variations, such actions, however, are usually not required.

Curve by Curve Adaptation with Derivative Smoothness

While the straight curve by curve adaptation yields good results in many circumstances, there is a significant underlying limitation on the weights. Namely, the weights cannot be too severe or else the procedure will collapse. This has been observed and while corrective action can be inserted directly into the process, such action is rather detailed and technical.

In a further study, Eiseman [42] found that a better course of action is to redefine the directional sweeps by splitting them into two phases: the active phase and the passive phase. In the active phase we just have the original curve by curve strategy in the current direction. This contains the fundamental adaptive forces. In the passive phase, a low pass filter is applied to remove any wiggles or abrupt changes in spacing caused by the active phase but to leave intact the basic results of the intended action. This produces a smooth grid in the sense of derivative continuity. As a consequence, continuous numerical derivatives are available for numerical solution algorithms and for the application of controls in successive sweeps. Such controls include the use of orthogonality and curvature in the weights. As a practical matter, it has been observed that the splitting of sweeps into this predictor-corrector format of active and passive phases has resulted in considerably enhanced stability and a much larger range of severity in the choice of weights.

The simplest form for the passive phase is given by the direct action of a Laplace filter upon the grid point locations. While such action by itself may not be appropriate for the generation of a nonsingular grid, it is certainly suitable for the stated purpose of establishing derivative continuity. The distinction is that the filter is applied at most a few times in each directional sweep rather than being driven towards convergence as is the case when the solution to a system of grid generation equations is sought. The Laplace filter employed by Eiseman [42] was given by the simple Gauss-Seidel

relaxation of

$$\mathbf{r}_{i,j,k}^{n+1} = \mathbf{r}_{i,j,k}^{n} + \frac{1}{12}\left[\mathbf{r}_{i+1,j,k}^{n} + \mathbf{r}_{i-1,j,k}^{n+1} + \mathbf{r}_{i,j+1,k}^{n} + \mathbf{r}_{i,j-1,k}^{n+1} + \mathbf{r}_{i,j,k+1}^{n} + \mathbf{r}_{i,j,k-1}^{n+1} - 6\mathbf{r}_{i,j,k}^{n}\right] \tag{116}$$

At the boundaries, this formula was restricted to provide filtering in only tangential directions. In the application, a three-dimensional monitor surface was defined in four-dimensional Euclidean space by means of the vector (x,y,z,u) where u is considered to be a function of (x,y,z). To obtain the physical space projection, we merely replace u with a constant which is usually 0. In a test of the basic movement, u was taken to have a severe variation across two intersecting ellipsoids that also intersected the boundaries of a Cartesian box. From an initial surface grid with a Cartesian grid projection, an equal arc length grid was rapidly generated on the surface, and accordingly, a smooth gradient clustered grid resulted in the (x,y,z)-projection $(x,y,z,0)$.

The Conformal Measure of Smoothness: Basic Elliptic Grid Generation

While the passive phase of each sweep provided smoothness in the sense of derivative continuity, another measure of smoothness is provided by an attraction to conformal conditions. This latter measure is more demanding and is thus clearly stronger.

The conformal measure of smoothness is most readily observed in a variational setting [8]. In two dimensions, we simply minimize the amount by which the Cauchy-Riemann conditions fail to be satisfied when generally incompatible boundary conditions are employed. At a point in the field, each of the Cauchy-Riemann equations has a residual that in the general circumstance deviates from zero. A measure of nonsatisfaction is clearly given by the sum of squared residuals. For a mapping from (x,y) to (ξ,η), the smallest loss over the whole field is then obtained when the integral

$$I_c = \int \left[(\xi_x - \eta_y)^2 + (\xi_y + \eta_x)^2\right] dxdy \tag{117}$$

is minimized. Although the mapping direction in the formulation is opposite to that of an eventual application, there is a good reason: namely, that the

control is relative to whole coordinate curves that are associated with constant values of ξ and η rather than to horizontal or vertical lines for constant values of x and y. The coordinate curves in the physical space (x,y) are generally not straight lines. More discussion of this choice can be found in a variety of sources [6-9,12].

When the integrand in Eq. 117 is expanded and regrouped, we get

$$I_c = \int [(\xi_x^2 + \xi_y^2) + (\eta_x^2 + \eta_y^2) + 2(\xi_y \eta_x - \xi_x \eta_y)]dxdy \tag{118}$$

where we can identify the newly grouped terms. From Eqs. 107 and 109, we get

$$I_c = \int [g^{11} + g^{22} - \frac{2}{J}]dxdy \tag{119}$$

Since $dxdy = Jd\xi d\eta$, the integral simplifies to be

$$I_c = \int (g^{11} + g^{22})dxdy - 2 \int d\xi d\eta \tag{120}$$

where now the last term is a constant because the domain of curvilinear variables (ξ,η) has a constant area: typically, this is just a rectangle or some collection of them. As a consequence, conformal smoothness is obtained when the constant part of Eq. 120 is dropped and the remaining integral

$$I_s = \int (g^{11} + g^{22})dxdy \tag{121}$$

is minimized. In this form, the extension into three dimensions is clear: we just add a term g^{33} to the integrand and integrate over volumes. As in Eq. 50, the Euler equations for the minimization of I_s are obtained by applying the general operator

$$\frac{\partial}{\partial x_i} - \frac{\partial}{\partial \xi_j} \frac{\partial}{\partial (x_i)_{\xi_j}} \tag{122}$$

to the integrand of Eq. 121 and setting the result equal to zero. For each equation i, the repeated index j in the operator is assumed to be summed from 1 to the number of spatial dimensions. This is just a commonly employed summation convention. For Eq. 121, we get the Laplace system

$$\nabla^2 \xi = 0$$
$$\nabla^2 \eta = 0 \tag{123}$$

that was considered by Winslow [43]. Alternatively, the same system could have been derived by transforming the integral I_s to curvilinear variables $\xi_1 = \xi$ and $\xi_2 = \eta$ and then employing the general operator

$$\frac{\partial}{\partial \xi_i} - \frac{\partial}{\partial x_j} \frac{\partial}{\partial (\xi_i)_{x_j}} \tag{124}$$

to the adjusted integrand. This is just the parallel to Eq. 51. Using the fact that the Laplace formulation of Eq. 123 is directly for the curvilinear variables; Thompson, Thames and Mastin [44] considered the Poisson generalization

$$\nabla^2 \xi = P$$
$$\nabla^2 \eta = Q \tag{125}$$

where the forcing terms P and Q provided separate and active controls on the respective families of coordinate curves. Because of the general effectiveness of this grid generation system, it became thoroughly developed and widely used. The effectiveness arose primarily from the treatment of whole coordinate curves relative to the conformal attraction inherent in Eq. 123. To apply the Poisson system, Eq. 125, the dependent and independent variables must be interchanged. The result is given by

$$\begin{aligned} g^{11} x_{\xi\xi} + 2g^{12} x_{\xi\eta} + g^{22} x_{\eta\eta} + x_\xi P + x_\eta Q = 0 \\ g^{11} y_{\xi\xi} + 2g^{12} y_{\xi\eta} + g^{22} y_{\eta\eta} + y_\xi P + y_\eta Q = 0 \end{aligned} \tag{126}$$

In the general case, we have

$$g^{ij} \mathbf{r}_{\xi_i \xi_j} + P_i \mathbf{r}_{\xi_i} = H\mathbf{N} \tag{127}$$

where $\mathbf{r}$ is the position vector in physical space, P_i is the i-direction forcing function and H is the mean curvature should the application be for a surface with unit normal $\mathbf{N}$. In a Euclidean space, H vanishes and the right-

hand side then also vanishes.

Upon examining Eq. 126, Middlecoff and Thomas [45] noticed that the forcing functions would have to be appropriately scaled to account for potentially disparate variations in the inverse metric elements g^{ij}. As a consequence, they redefined the forcing terms to include a factor of g^{ii} for each direction i. In terms of the general form of Eq. 127, we get

$$g^{ij}\left(r_{\xi_i\xi_j} + \delta_{ij}\Psi_i r_{\xi_i}\right) = HN \tag{128}$$

where the adjusted forcing functions are given by

$$\Psi_i = g^{ii}P_i \tag{129}$$

and the δ_{ij} in the sum on i and j is the Kronecker symbol that is unity when i = j and vanishes otherwise. For more details on the general development of this Poisson system, the reader is referred to the review articles by Thompson, Warsi and Mastin [6], Thompson [7], Eiseman [8], and Eiseman and Erlebacher [9] as well as the text by Thompson, Warsi and Mastin [46].

Curve by Curve Adaptation with Conformal Smoothness

Recognizing the basic need for smoothness, Anderson and Steinbrenner [47,48] brought the process of equidistributing weights along coordinate curves into the format of the Poisson system of Eq. 125. Their development was motivated by the previous work of Middlecoff and Thomas [45] who employed the formulation of Eq. 128 in the planar form

$$g_{22}\left(r_{\xi\xi} + \phi r_{\xi}\right) - 2g_{12}r_{\xi\eta} + g_{11}\left(r_{\eta\eta} + \Psi r_{\eta}\right) = 0 \tag{130}$$

with

$$P = \frac{g_{22}}{g}\phi$$

$$Q = \frac{g_{11}}{g}\Psi \tag{131}$$

for the purpose of converting boundary distributions into forcing functions thereon. While Middlecoff and Thomas interpolated the forcing functions into the field so that boundary distributions could be propagated into the interior of the field, Anderson and Steinbrenner viewed each curve as if it were such a boundary curve. In the boundary application, the assumption of orthogonality and vanishing transverse curvature was employed to get the differential equation for an equidistributed weight. In particular, with $g_{12} = \mathbf{r}_\xi \cdot \mathbf{r}_\eta = 0$ and $\mathbf{r}_{\eta\eta} = 0$, the equidistribution statement of Eqs. 24-25 is obtained from Eq. 130 as

$$s_{\xi\xi} + \phi s_\xi = 0 \tag{132}$$

where s is the curve arc length and

$$\phi = \frac{w_\xi}{w} \tag{133}$$

is the relationship bewteen the forcing function ϕ and the weight function w. In the more general interior application, the same sort of equidistribution statement was established without the previous assumptions. The main distinction is the addition to ϕ of a term for orthogonality and terms for curvature. These orthogonality and curvature terms represent the attraction to the conformal measure of smoothness. Without forces, they yield the desired smoothness. With forces a deviation is obtained.

In the general two-dimensional case represented by Eq. 130, our task is to obtain a reduction into the form of Eqs. 132-133 for each direction. To start, we shall rewrite Eq. 130 in a form where the basic constituent vectors are the unit tangent vectors along coordinate curves and the unit normal vectors perpendicular to coordinate curves. From the metric coefficient definition in Eq. 107, the appropriate normalization factors for the natural tangent directions $\mathbf{r}_\xi$ and $\mathbf{r}_\eta$ are obtained directly. With τ_1 and τ_2 denoting the respective unit tangents, we have

$$\begin{aligned} \mathbf{r}_\xi &= s_\xi \tau_1 \\ \mathbf{r}_\eta &= u_\eta \tau_2 \end{aligned} \tag{134}$$

where $s_\xi = \sqrt{g_{11}}$ and $u_\eta = \sqrt{g_{22}}$ are the respective arc length derivatives. For

Eq. 130, second derivatives are needed and can be obtained by derivatives of Eq. 134. To get them, however, the unit tangents must be differentiated. This is accomplished with the first of the Frenet formulas (cf. [49]) which are given by

$$\begin{aligned}(\tau_1)_s &= k_1 \mathbf{n}_1 \\ (\tau_2)_u &= k_2 \mathbf{n}_2\end{aligned} \tag{135}$$

for the two respective coordinate directions. The unit normal vectors and curve curvatures are given by $\mathbf{n}_1$ and $\mathbf{n}_2$ and by k_1 and k_2. By using the chain rule, the arc length differentiation in Eq. 135 is readily converted into differention with respect to ξ and η. With the result inserted into the derivatives of Eq. 134, we arrive at

$$\begin{aligned}\mathbf{r}_{\xi\xi} &= s_{\xi\xi}\tau_1 + s_\xi k_1 \mathbf{n}_1 \\ \mathbf{r}_{\eta\eta} &= u_{\eta\eta}\tau_2 + u_\eta k_2 \mathbf{n}_2\end{aligned} \tag{136}$$

Upon substitution of the first derivatives in Eq. 134 and the second derivatives in Eq. 136, the grid generation equations of Eq. 130 become

$$\begin{aligned}g_{11}\left[(s_{\xi\xi} + \phi s_\xi)\tau_1 + s_\xi k_1 \mathbf{n}_1\right] - 2g_{12}\mathbf{r}_{\xi\eta} \\ + g_{22}\left[(u_{\eta\eta} + \Psi u_\eta)\tau_2 + u_\eta k_2 \mathbf{n}_2\right] = 0\end{aligned} \tag{137}$$

where the terms have been grouped according to the tangent and normal directions. An immediate observation from the grouping is that the equidistribution process along coordinate curves is now separated by direction: namely, that the representative form appears only in the coefficients of τ_1 and τ_2. To isolate the forcing functions, only a dot product with the respective normal vectors $\mathbf{n}_1$ and $\mathbf{n}_2$ is required. To remove Ψ, a dot product with $\mathbf{n}_2$ yields

$$g_{11}(s_{\xi\xi} + \phi s_\xi)\tau_1 \cdot \mathbf{n}_2 + s_\xi k_1 \mathbf{n}_1 \cdot \mathbf{n}_2 - 2g_{12}\mathbf{r}_{\xi\eta} \cdot \mathbf{n}_2 + g_{22}u_\eta k_2 = 0 \tag{138}$$

or

$$s_{\xi\xi} + \Phi s_{\xi} = 0 \tag{139}$$

where

$$\Phi = \phi + \frac{1}{g_{11}\tau_1 \cdot \mathbf{n}_2}\left[s_{\xi} k_1 \mathbf{n}_1 \cdot \mathbf{n}_2 - 2g_{12}\mathbf{r}_{\xi\eta} \cdot \mathbf{n}_2 + g_{22} u_{\eta} k_2\right] \tag{140}$$

As a matter of interpretation, the deviation from an equidistribution of ϕ is represented by the second term of Eq. 140. It contains an orthogonality part due to g_{12} and $\mathbf{n}_1 \cdot \mathbf{n}_2$ and curvature parts due to k_1 and k_2 for the coordinate curves in the two distinct directions. When the coordinates are orthogonal, that deviation reduces to $g_{22} u_{\eta} k_2 / g_{11}$. On balance, the force towards an equidistribution of ϕ must overcome the forces of conformality that are represented by orthogonality and curvature. A similar development and balance occurs for Ψ.

In the adaptive context, when the local force is sufficiently strong, the equidistribution force overpowers the smoothness conditions to become dominant and, thereby, to provide the desired equidistribution of the weight along each curve. The equidistribution is more localized than the previous derivative continuous measure of smoothness. This occurs because the equidistribution is essentially cut off unless it is sufficiently strong. The value of such a cutoff is that intense local clusters can be formed primarily from curves that are not too far from the given locality. In contrast, the actual equidistribution of weights along curves adjusts all points along curves, and thus, tends to globally propagate adjustments to intense local requirements. This tendency can of course be limited with the explicit use of orthogonality and curvature attraction.

The Variational Form of Curve by Curve Adaptation

In the prior section, the curve by curve controls were injected into the two-dimensional Poisson system of Eq. 125 in a somewhat ad hoc manner. This was done by first casting the system into a form, Eq. 137, that explicitly displayed the appropriate elements of equidistribution and then by making the ad hoc judgment that those elements were dominant. The judgment occurred at the stage represented by Eq. 140. While the same development could be continued into three dimensions, the added detail in getting to the stage of

judgment, not to mention the judgment itself, is much more complex.

In order to simplify the judgment and to gain a higher degree of generality, we shall develop a variational formulation based upon the fundamental statements of equidistribution and conformal smoothness. Upon observing the integrand of the smoothness integral I_s of Eq. 121, it is noted that the i-th coordinate direction is represented by the inverse metric term g^{ii}. In a geometric sense, this represents the spacing between curves as we march along in the i-direction. To control that spacing, we consider a weight function w_i which would be completely effective if g^{ii}/w_i were constant. For orthogonal systems, the basic metric equidistribution statement of Eq. 112 is obtained for $\sqrt{w_i}$. On the overall basis where the various terms in the integrand must compete with each other, we are led to the integral

$$I = \int \left(\frac{g^{11}}{w_1} + \frac{g^{22}}{w_2}\right) dxdy \tag{141}$$

The associated Euler equations derived from the operator of Eq. 122 are given by

$$\nabla^2 \xi = g^{11} \frac{(w_1)_\xi}{w_1} + g^{12} \frac{(w_1)_\eta}{w_1}$$

(142)

$$\nabla^2 \eta = g^{12} \frac{(w_2)_\xi}{w_2} + g^{22} \frac{(w_2)_\eta}{w_2}$$

With the exception of the g^{12} terms, this matches the equations of Anderson and Steinbrenner [47,48] that were discussed in the prior section. The g^{12} terms, here, represent the transverse variations of the weights for each direction. As in the prior case, the inclusion within the Poisson format is clearly quite simple. When $w_1 = w_2 = w$, the conformal forces are scaled by $1/w$ in the variational integral of Eq. 141. Given the conformal tendency to maintain orthogonality, the uniform scaling of $\sqrt{w}$ along each coordinate intuitively suggests an equidistribution of w with respect to volume elements. This equidistribution was analytically verified by Anderson [50] who was motivated by the diffusive form proposed by Winslow [51] rather than the above intuition. The diffusive form comes from the use of $w = 1/D$ in Eqs. 141-142. With the prescribed diffusion function D, the equations proposed by Anderson

[50] are obtained by setting $w_1 = w_2 = 1/D$ in Eq. 142. Here, the logarithmic derivative trades a reciprocal for a minus sign so that the form of the forcing terms in Eq. 142 now appear with minus signs and a replacement of w_1 and w_2 by D.

Unlike the development of Anderson and Steinbrenner, the extension into three dimensions is easily formulated and justified. From the integral

$$I = \int \left(\frac{g^{11}}{w_1} + \frac{g^{22}}{w_2} + \frac{g^{33}}{w_3}\right)dxdydz \tag{143}$$

we get the Euler equations

$$\begin{aligned}
\nabla^2\xi &= g^{11}\frac{(w_1)_\xi}{w_1} + g^{12}\frac{(w_1)_\eta}{w_1} + g^{13}\frac{(w_1)_\zeta}{w_1} \\
\nabla^2\eta &= g^{21}\frac{(w_2)_\xi}{w_2} + g^{22}\frac{(w_2)_\eta}{w_2} + g^{23}\frac{(w_2)_\zeta}{w_2} \\
\nabla^2\zeta &= g^{31}\frac{(w_3)_\xi}{w_3} + g^{32}\frac{(w_3)_\eta}{w_3} + g^{33}\frac{(w_3)_\zeta}{w_3}
\end{aligned} \tag{144}$$

In a more compact form, the two- and three-dimensional equations can be stated as

$$\nabla^2\xi_i = g^{ij}\frac{(w_i)_{\xi_j}}{w_i} \tag{145}$$

Upon returning to the basic statement of Eq. 112 and noting that the derived equidistribution was for each $\sqrt{w_i}$ rather than w_i, we are motivated to replace each w_i by its square and repeat the argument. Even more generally, it is just as easy to consider a replacement by w_i^α for any positive constant α. When the replacement is made, we can simply carry w_i^α through all of the above steps in an undisturbed manner and then perform the final differentiation at the end. Altogether, we obtain the general forcing terms

$$P_i = \alpha g^{ij}\frac{(w_i)_{\xi_j}}{w_i} \tag{146}$$

that can be employed in Eq. 127 which also is equally valid on surfaces. As originally mentioned, when $\alpha = 2$, the equidistribution forces conform most directly to the basic metric equidistribution statement of Eq. 112. This might

actually be the most intuitively plausible application when the linear weight of Eq. 1 is employed. Then, at least, there is a more direct correlation to the previous curve by curve studies that use linear weights.

With the forcing terms of Eq. 146 in the elliptic operator from Eq. 127, a concise evolutionary parabolic partial differential equation can be deduced. In a basic sense, it is the natural extension of the one-dimensional form presented earlier in Eq. 52 and is given by

$$K \frac{\partial \mathbf{r}}{\partial t} = g^{ij}\left(\mathbf{r}_{\xi_i \xi_j} + \alpha \frac{(w_i)_{\xi_j}}{w_i} \mathbf{r}_{\xi_i}\right) - H\mathbf{N} \tag{147}$$

where linear weights retain the original meaning when $\alpha = 2$. The further extension in the spirit of Eq. 96 is also possible. For instance, the attraction to a fixed prescribed grid $\mathbf{q}$ can be inserted as a term within the format of B in Eq. 96. As in Eq. 94, the grid $\mathbf{r}$ is attracted to the fixed $\mathbf{q}$ with the term $D(\mathbf{r} - \mathbf{q})$. The evolutionary equation then becomes

$$K \frac{\partial \mathbf{r}}{\partial t} = g^{ij}\left(\mathbf{r}_{\xi_i \xi_j} + \alpha \frac{(w_i)_{\xi_j}}{w_i} \mathbf{r}_{\xi_i}\right) - H\mathbf{N} - D(\mathbf{r} - \mathbf{q}) \tag{148}$$

FINITE VOLUME METHODS

Finite volume methods are methods where the grid point motion is based upon the volume elements between grid points. The basic format was first established in one dimension with the mean-value relaxation statement of equidistribution. This led to the center of mas form of Eqs. 32-33 and the movement form in Eqs. 34-35. The second format was correspondingly established in the variational setting with Eq. 36.

Direct Mean-Value Relaxation

For simplicity, only two dimensions will be considered since the fundamental pattern is established therein. A two-dimensional stencil centered about a grid point $\mathbf{r}_{ij}$ is displayed in Figure 1. The local motion of $\mathbf{r}_{ij}$ comes from the weighted volume elements which are defined by the indicated triangles containing $\mathbf{r}_{ij}$ as a vertex. In each quadrant k, a barycenter $\mathbf{b}_k$ and a weight w_k is obtained as the average of positions and values respectively over the

triangle vertices. With the weight w_k considered to be uniformly distributed over the corresponding triangle of area A_k, the contribution at the barycenter b_k is just $w_k A_k$. From the weighted triangle areas, the direct extension of Eq. 32 is the center of mass for the four quadrants and is given by

$$\mathbf{r}_{ij}^{new} = \frac{\sum_{k=1}^{4} w_k A_k \mathbf{b}_k}{\sum_{k=1}^{4} w_k A_k} \tag{149}$$

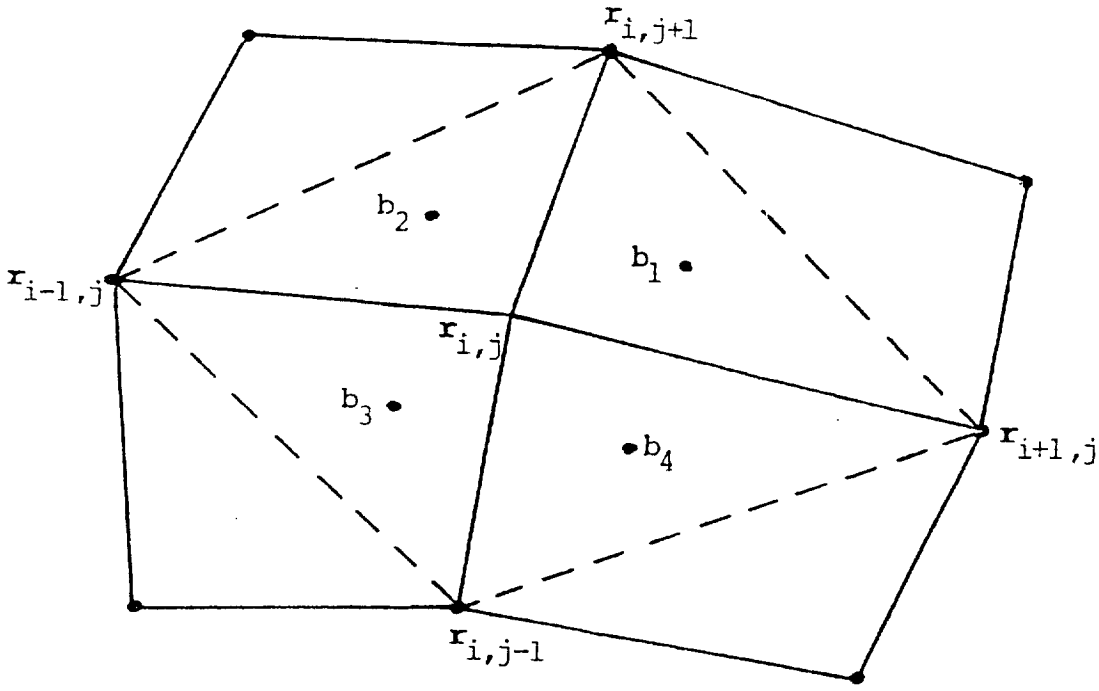


Figure 1: Finite volume stencil in two dimensions

to determine the new position for $\mathbf{r}_{ij}$. The actual movement is given by

$$\mathbf{r}_{ij}^{new} - \mathbf{r}_{ij} = \frac{\sum_{k=1}^{4} w_k A_k (\mathbf{b}_k - \mathbf{r}_{ij})}{\sum_{k=1}^{4} w_k A_k} \tag{150}$$

and is employed in a point iterative cycle. This represents the most direct statement of mean-value relaxation. In the movement form of Eq. 150, the effective origin for the computation has been shifted to the stencil center $\mathbf{r}_{ij}$.

Like the earlier Laplace system of Eq. 123, the motion represents an attraction to conformal conditions when each w_k is unity. In terms of the linear weight form of Eq. 1, the first term which is the number 1 is then the representative of the conformal attraction relative to which the other forces are applied. The analytical indicator for the conformal conditions comes from the converse of the mean-value theorem which is simply presented by Epstein [52] on pages 146-148. While the analytical argument employs the area mean value of functions over circular disks, the finite parallel given by Eq. 149 with $w_k = 1$ is only an approximate form employed within an iterative cycle.

The pointwise relaxation of the center of mass formula of Eq. 149 has been considered by Diaz, Kikuchi and Taylor [53], Oden, Devloo and Strouboulis [54], Schwartz and Connett [55], Connett, Agarwal, and Schwartz [56], Erlebacher and Eiseman [57], Eiseman [58], and Erlebacher [59]. In the studies by Erlebacher and Eiseman, the more general application to unstructured meshes was developed.

The relaxation formula for general connectivity triangular meshes is essentially the same formula as given by Eqs. 149-150. Adjustments, however, must be made to accommodate any number of triangles in the determination of a pointwise move. This requires a careful labeling system. In our discussion, we shall assume a global pointwise index m for $\mathbf{r}_m$ together with a local index j to label the points $P_j = \mathbf{r}_{\alpha(m,j)}$ that are directly joined to $\mathbf{r}_m$. The local index j is attached successively to the positions as we move about $\mathbf{r}_m$ in a circular manner. The range for j is taken to be from 1 to j_m where j_m+1 is identified with 1 (i.e., a reduction modulo j_m). Under these index conditions, the relaxation formula becomes

$$\mathbf{r}_m^{new} = \frac{\sum_{j=1}^{m_j} w_{mj} A_{mj} \mathbf{b}_{mj}}{\sum_{j=1}^{m_j} w_{mj} A_{mj}} \tag{151}$$

where

$$\mathbf{b}_{mj} = \frac{1}{3}\left(\mathbf{r}_m + \mathbf{P}_j + \mathbf{P}_{j+1}\right) \tag{152}$$

In continuation, we next consider the minimization of the sum of squares

$$S = \sum_{m=1}^{N} \sum_{j=1}^{j_m} w_{mj} A_{mj} ||\mathbf{r}_m - \mathbf{b}_{mj}||^2 \tag{153}$$

to establish a variational perspective. The global sum is over all N points which are to be moved. For simplicity, we shall assume Dirichlet data at the boundaries so that only interior points are moved. The norm appearing in Eq. 153 is simply the Euclidean norm (square root of dot product). As a matter of convenient notation, let

$$\mathbf{r}_m = x_{m1}\begin{pmatrix}1\\0\end{pmatrix} + x_{m2}\begin{pmatrix}0\\1\end{pmatrix} = x_{m1}\mathbf{e}_1 + x_{m2}\mathbf{e}_2 \tag{154}$$

To find the condition for minimization, we must set the first derivatives of S to 0. For the k-th Cartesian direction, the general condition is given by

$$\frac{\partial S}{\partial x_{mk}} = 2 \sum_{j=1}^{j_m} w_{mk} A_{mk} \left\{\left(\mathbf{r}_m - \mathbf{b}_{mj}\right) \cdot \left(\frac{2}{3}\mathbf{e}_k\right) + \left(\mathbf{P}_j - \mathbf{b}_{mj}\right) \cdot \left(-\frac{1}{3}\mathbf{e}_k\right) + \left(\mathbf{P}_{j+1} - \mathbf{b}_{mj}\right) \cdot \left(-\frac{1}{3}\mathbf{e}_k\right)\right\} = 0 \tag{155}$$

Of the terms represented by dot products, the first is due to the derivative of the m-th term in S while the second and third terms come respectively from derivatives of the j-th and (j+1)-th terms that appear with distinct global indices in S. This occurs for each triangle j because $\mathbf{r}_m$ appears as the center in the m-th term and as a surrounding term when the center is shifted to $\mathbf{P}_j$ or $\mathbf{P}_{j+1}$. Here, we are assuming that $\mathbf{r}_m$ is at least two connections away

from the boundary so that both P_j and P_{j+1} are varied and accordingly have corresponding sums in S. The basic coupling clearly comes from the barycentric formula of Eq. 152 and gives rise to the factors of 1/3 in Eq. 155. Using the barycentric formula, the minimization condition of Eq. 155 reduces to

$$\sum_{j=1}^{j_m} w_{mj} A_{mj} (\mathbf{r}_m - \mathbf{b}_{mj}) \cdot \mathbf{e}_k = 0 \tag{156}$$

which is a componentwise equality for a vector equation. The vector equation is obtained by just removing the dot product with $\mathbf{e}_k$ in Eq. 156. The immediate consequence is the relaxation formula stated in Eq. 151. In summary, we have just shown that the center of mass formula is the general field equation for the variational problem stated in Eq. 153.

Both the variational statement and the consequent field equation can be brought into the evolutionary context along with a mechanism for attraction to a given grid $\mathbf{q}$. As in the curve by curve case with Eq. 147, we have an underlying parabolic partial differential equation for the Euler equation. As in the one-dimensional evolutionary least squares statement of Eq. 57 and the associated mean-value expression in Eq. 56, we get an extension which also incorporates the attraction to the fixed $\mathbf{q}$ as expressed in Eq. 95. In a directly parallel manner to the previous development, the minimization of

$$S = \sum_{m=1}^{N} \{K_i^n \Delta t \left(\frac{\mathbf{r}_m^{n+1} - \mathbf{r}_m^n}{\Delta t}\right)^2 + D_m ||\mathbf{r}_m - \mathbf{q}_m||^2 + \sum_{j=1}^{j_m} w_{mj} A_{mj} ||\mathbf{r}_m - \mathbf{b}_{mj}||^2\} \tag{157}$$

produces

$$\mathbf{r}_m^{n+1/2} = \frac{\frac{K_i^n}{\Delta t} \mathbf{r}_m^n + D_m \mathbf{q}_m + \sum_{j=1}^{m_j} w_{mj} A_{mj} \mathbf{b}_{mj}}{\frac{K_i^n}{\Delta t} + D_m + \sum_{j=1}^{m_j} w_{mj} A_{mj}} \tag{158}$$

The field equation of Eq. 158, of course, will sepcialize to that of Eq. 149 in successive stages.

Full Mean-Value Relaxation

In the center of mass formula of Eq. 149, the control over the grid comes from the specification of a single weighting quantity that is applied against the triangular elements of area. While this approach yields a control over the elemental area distribution, it does not exercise all of the available degrees of freedom for such control. Yet one more degree of freedom is available. The full utilization of all degrees of freedom was evident in the transition from the Laplace system of Eq. 123 to the Poisson system of Eq. 125. A similar transition is a reasonable expectation for the process of mean-value relaxation.

The utilization of all degrees of freedom in mean-value relaxation was developed by Eiseman [60]. As in the Poisson system of Eq. 125 and in the alternating direction methods of adaptivity, control was established with a coordinate directional bias that provided the required separation into distinct degrees of freedom. In the local stencil of Figure 1, the directional bias was obtained by a projection of the movement vector onto the coordinate curve through $\mathbf{r}_{ij}$ in the appropriate direction. If w is the weight for the i-direction, then the projection is onto the curve from $\mathbf{r}_{i-1,j}$ to $\mathbf{r}_{i+1,j}$. In the implementation, the transverse coordinate curve from $\mathbf{r}_{i,j-1}$ to $\mathbf{r}_{i,j+1}$ was used to divide the weights so that first and fourth quadrants would pull towards $\mathbf{r}_{i+1,j}$ from $\mathbf{r}_{ij}$ while second and third quadrants would pull towards $\mathbf{r}_{i-1,j}$ from $\mathbf{r}_{ij}$. On each side of the transverse coordinate curve, a center of mass was computed and then projected onto the appropriate segment from $\mathbf{r}_{ij}$. Denoting the projected distances by d_+ and d_- for the positive and negative i-directions from $\mathbf{r}_{ij}$, the new position along the curve is given by

$$d = \frac{w_+ d_+ + w_-(-d_-)}{w_+ + w_-} \tag{159}$$

where

$$w_+ = w_1 A_1 + w_4 A_4$$
$$w_- = w_2 A_2 + w_3 A_3 \tag{160}$$

This is the same form as given by Eqs. 34-35 in one dimension. Since the construction is centered at r_{ij}, the sign of d gives the direction of movement: negative is towards $r_{i-1,j}$, while positive is towards $r_{i+1,j}$. In a similar manner, a second weight $\bar{w}$ is employed for the j-direction and a similar distance is determined along that direction. The signs of the two distances then determine the quadrant which contains the new position. That new position is determined by interpolation.

The projected distances appearing in Eq. 159 are constructed from the vectors

$$\begin{aligned} I_+ &= r_{i+1,j} - r_{ij} \\ J_+ &= r_{i,j+1} - r_{ij} \\ I_- &= r_{i-1,j} - r_{ij} \\ J_- &= r_{i,j-1} - r_{ij} \end{aligned} \tag{161}$$

which point from the current position r_{ij} to the nearest neighboring points along coordinate curves. To compute the positive projected distance d_+ in the i-direction, we must first obtain the center of gravity for the side in the direction of I_+. This is given by

$$C_{I_+} = \frac{1}{w_+}\left(w_1 A_1 b_1 + w_4 A_4 b_4\right) \tag{162}$$

and reduces to

$$C_{I_+} = r_{ij} + \frac{1}{3} I_+ + \frac{1}{3w_+}\left(w_1 A_1 J_+ + w_4 A_4 J_-\right) \tag{163}$$

upon employing the definition of barycenters. The projection of $C_{I_+} - r_{ij}$ along the positive i-direction comes from a dot product with $I_+/||I_+||$ and yields the maximum distance

$$d_+ = \frac{1}{3}||I_+|| + \frac{1}{3w_+}\left(w_1 A_1 J_+ + w_4 A_4 J_-\right) \cdot \left(\frac{I_+}{||I_+||}\right) \tag{164}$$

In the original form, Eiseman [60] considered a bilinear interpolation that also included the diametrically opposite point to $\mathbf{r}_{ij}$. In a further study, Schwartz and Connnett [55] and Connett, Agarwal and Schwartz [56] experimentally found that the method became unstable when the weights became sufficiently severe. While this led them to consider the earlier center of mass form of Eq. 149, it lead Eiseman to consider barycentric interpolation in place of the original bilinear interpolation. The intuitive reason was that the asymmetric use of diametrically opposite points would be more restrictive on the stability. In subsequent tests, the intuition was confirmed: the use of barycentric interpolation led to stable results even when the weights had considerable severity.

With the distances along the i- and j-directions determined by Eq. 159 and denoted by d_i and d_j respectively, it is a simple matter to state the barycentric interpolation. If, for example, both d_i and d_j are positive, then the interpolation must be performed in the first quadrant. There, the positively-oriented vectors $\mathbf{I}_+$ and $\mathbf{J}_+$ from Eq. 161 are employed and the interpolation is given by

$$\mathbf{r}_{ij}^{new} = \mathbf{r}_{ij} + d_i \frac{\mathbf{I}_+}{||\mathbf{I}_+||} + d_j \frac{\mathbf{J}_+}{||\mathbf{J}_+||} \tag{165}$$

A substitution from Eq. 161 provides the reduction into the barycentric form

$$\mathbf{r}_{ij}^{new} = \alpha_0 \mathbf{r}_{ij} + \alpha_1 \mathbf{r}_{i+1,j} + \alpha_2 \mathbf{r}_{i,j+1} \tag{166}$$

where

$$\begin{aligned} \alpha_1 &= \frac{d_i}{||\mathbf{I}_+||} \\ \alpha_2 &= \frac{d_j}{||\mathbf{J}_+||} \end{aligned} \tag{167}$$

and

$$\alpha_0 + \alpha_1 + \alpha_2 = 1$$

Since the coefficients presented by Eq. 167 usually lie in the unit interval,

the new position from Eq. 166 usually lies within the triangle determined by the three vectors of Eq. 166. Clearly, the same format applies to the other quadrants.

Upon inspection of the full mean-value relaxation process, we note that the projection was accomplished with a rather strong adherence to the individual coordinate curves. While this is the closest to our objective of separating the action of weights to be along the coordinate directions, a slightly weaker adherence can also be considered. The intent, here, is to get a simpler formulation that still retains the separation feature. The weakness in comparison with the more thorough version is that the center point $\mathbf{r}_{ij}$ of the stencil in Figure 1 is removed for the projection. The consequence is that the algorithm may not be as robust as the stronger version; although, from a practical point of view, experience with the weaker formulation is seen to be fairly competitive. To state the weak form, an i-direction weight w is first employed in the mean-value movement formula of Eq. 150 and is then projected against the i-direction secant to determine a vector increment $(\Delta r)_i$. That is,

$$(\Delta r)_i = [(r_{ij}^{new} - r_{ij}) \cdot \tau_i]\tau_i \tag{168}$$

where

$$\tau_i = \frac{r_{i+1,j} - r_{i-1,j}}{||r_{i+1,j} - r_{i-1,j}||} \tag{169}$$

Similarly, a j-direction weight $\bar{w}$ is used to determine $(\Delta r)_j$. The new position is then determined by the simple sum

$$\Delta r = (\Delta r)_i + (\Delta r)_j \tag{170}$$

for the actual movement from the old $\mathbf{r}_{ij}$.

Upon consideration of the primary motivation to separate the action of the weights to be along coordinate directions, we are led to a variational statement which is a natural extension of that in Eq. 153 for triangles and that in Eq. 36 for curves. With an impulse towards simplicity, we might first examine the minimization of

$$S = \sum_{i,j} \{\Omega_{i+1/2,j}||r_{i+1,j} - r_{ij}||^2 + \bar{\Omega}_{i,j+1/2}||r_{i,j+1} - r_{ij}||^2\} \qquad (171)$$

where

$$\Omega_{i+1/2,j} = w_{i+1/2,j}A_{i+1/2,j}$$

$$\bar{\Omega}_{i,j+1/2} = \bar{w}_{i,j+1/2}A_{i,j+1/2} \qquad (172)$$

The Euler equations, however, yield the relaxation formula

$$r_{ij} = \frac{\Omega_{i+1/2,j}r_{i+1,j} + \Omega_{i-1/2,j}r_{i-1,j} + \bar{\Omega}_{i,j+1/2}r_{i,j+1} + \bar{\Omega}_{i,j-1/2}r_{i,j-1}}{\Omega_{i+1/2,j} + \Omega_{i-1/2,j} + \bar{\Omega}_{i,j+1/2} + \bar{\Omega}_{i,j-1/2}} \qquad (173)$$

which, unfortunately, provides a more modest separation for the coordinate directions. To obtain the strong separation of Eqs. 159-164, we then must consider the statement

$$S = \sum_{i,j} \{(w_+d_+ - w_-d_-)^2_{ij} + (\bar{w}_+\bar{d}_+ - \bar{w}_-\bar{d}_-)^2_{ij}\} \qquad (173)$$

where the constructive elements about each r_{ij} are given in Eqs. 159-164. At equilibrium, we have a balance between the original movement forces and the boundary conditions. As a final note, here, this separation can readily be compared with the curve by curve approaches and extended into the evolutionary format.

Direct Variational Constructions

Rather than a direct construction of a movement formula and a subsequent implied variational statement, the process here will be reversed: the construction will appear in the formlation of the variational statement and the movement action will be the consequence. In so doing, there is the option of directly minimizing the defining sum rather than deriving a movement formula from the Euler equations. Kennon and Dulikravich [61-62] pursued the direct approach by using an optimization technique. The primary motivation for their

minimization problem came from previous variational studies [63-68] along with the desire to place a larger constant at the center location $\mathbf{r}_{ij}$ than would have come from the traditional use of central differences that would have missed $\mathbf{r}_{ij}$. This then led them to consider a finite volume stencil.

The basic attractive forces come from distinct sums of squares that are balanced against each other. Of the forces, the first one pulls towards an equal volume distribution with terms of the form

$$(A_1-A_2)^2 + (A_2-A_3)^2 + (A_3-A_4)^2 + (A_4-A_1)^2 \tag{174}$$

for each point $\mathbf{r}_{ij}$. The second one pulls towards an orthogonal grid with terms of the form

$$(I_+\cdot J_+)^2 + (J_+\cdot I_-)^2 + (I_-\cdot J_-)^2 + (J_-\cdot I_+)^2 \tag{175}$$

where the local vectors about $\mathbf{r}_{ij}$ are defined in Eq. 161. Denoting the volume equalization and orthogonality sums by S_v and S_0 respectively, Kennon and Dulikravich minimized the linear combination

$$(1-\alpha)S_v + \alpha S_0 \tag{176}$$

to generate or improve a grid. The parameter $0 \leq \alpha \leq 1$ was chosen to balance the two forces. The inclusion of adaptive forces comes with the inclusion of yet another sum. This corresponds with the basic pattern of most variational methods. The typical choice for an adaptive sum S_a is the one over weighted volumes which can be assembled with terms of the form

$$w(A_1 + A_2 + A_3 + A_4) \tag{177}$$

for each location $\mathbf{r}_{ij}$. The overall sum then appears in the form

$$\lambda_1 S_v + \lambda_2 S_0 + \lambda_3 S_a \tag{178}$$

which upon minimization provides a competition between the various effects.

As in all variational constructions, various effects can be readily inserted into the formulation. Here, many such effects can be developed from

the vectors of Eq. 161 which comprise the basic metric structure. In contrast with the ease of formulation, the actual use of such formulations can be complicated; thus, presenting additional issues.

VARIATIONAL METHODS

The basic pattern of developments with variational methods is to gather the desired attributes, form a positive pointwise measure of each such attribute, integrate the measures over the field, form a positive linear combination of the integrals, and then minimize the resultant combination. The minimization process typically follows the standard manipulations from the calculus of variations [16]. This leads to a system of partial differential equations known as Euler equations (Eqs. 50, 51, 122,124) that are then to be solved by the available numerical methods for PDE's. The main attribute in most methods is the attraction to conformality and is given by the integral of Eq. 121 for the Laplace system of Eq. 123. This means that other attributes are balanced against the Laplace system for curvilinear variables that was employed by Winslow [43], Thompson, Thames and Mastin [44] and others. The main adaptive attribute typically comes in the form of a weighted Jacobian so that the minimized integral by itself would produce an equidistribution of the weight over volume elements or powers of them. Along with the attributes of conformality and weighted Jacobians, Yanenko et al [67] included a Lagrangian attraction for fluid motion while Brackbill and Saltzman [63] inlcuded an orthogonality attraction for an improved grid structure. The orthogonality attraction came from integrals of squared cross metrics. In two dimensions, these were either $(g_{12})^2$ or $(g^{12})^2$. With the motivation from the use of orthogonality in Brackbill and Saltzman, Kennon and Dulikravich [61] pursued the finite volume approach discussed earlier, Saunders [68] examined a tensor product of B-splines, Kreis, Thames and Hassan [69] considered scaling problems, and Steinberg and Roache [15] developed a variant with reference grids.

Rather than employ the attraction to the preferred Laplace system for Elliptic grid generation, Morice [70], Oskam and Huizing [71], and Steinberg and Roache [15] used attraction towards the Laplace system for locations in physical space. While this provides conformal smoothness, this is not as robust as the preferred system: it does however offer some simplicity. The main insurance against grid folding then falls upon the volume control or the permis-

sion of boundary point motion. Steinberg and Roache [15] favor volume control while Morice [70] and Oskam and Huizing [71] rely primarily upon boundary point motion.

The idea of reference grids employed by Steinberg and Roache [15] is the same as that employed earlier by Steger and Chaussee [72] in their development of hyperbolic methods. The use, however, is more extensive in that various properties are extracted from the reference grids. They appear essentially in the form of multiplicative factors applied to the standard format of the other investigators. These factors are weights that give appropriate ratios between the current desired grid and the known reference grid. A note of caution to be observed in the extraction of properties is that a desired attribute might not be obtained when the number of available degrees of freedom is exceeded.

A somewhat general development of the variational methods is given in the text by Thompson, Warsi, and Mastin [46]. This was subsequently implemented in a large program [73]. In one dimension, earlier studies were given by Gough, Spiegel, and Toomre [74] and by Pierson and Kutler [75]. Also, a study of temporal smoothness was developed by Bell and Shubin [22]. In the multidimensional context, a study for systematically dealing with the inherent complexity of variational methods was performed by Steinberg and Roache [76,77] who proposed appropriate symbolic manipulation methods.

TEMPORAL ASPECTS

When the temporal accuracy of a simulation is to be enhanced, the evolving grid must closely follow the trajectories of the severe disturbances in the solution. This is in contrast to the situation where a steady-state solution is sought and the grid there eventually settles down into a virtually final configuration. The primary concern in the development of an accurate temporal resolution is that the severe disturbances will not escape from their resolution during the course of any time step in the numerical simulation. With this concern in mind, a number of investigators have felt that the grid equations should be formulated directly for grid point velocities rather than positions. Then, at least, the grid velocities would follow an analytical model for any chosen time level such as the full implicit level or the level halfway between explicit and fully implicit. Clearly, the use of such velocities should be an improvement over using backward differences in time to esti-

mate the same velocities from a grid point movement scheme that only produces pointwise locations from currently available data. From the viewpoint of grid velocities, the schemes which produce only pointwise locations can be referred to as static rather than dynamic (e.g., note Hyman and Naughton [78]). A typical difference is that static methods depend upon data at only one instant of time while dynamic methods often depend upon data over an interval of time. With the advantage of directly obtaining grid velocities, the dynamic methods can possess the disadvantage of a more difficult control over coordinate nonsingularity. This is because grid point locations must ultimately be determined, and if some velocities are too large or change too quickly, then points could be either overrun or artificially encircled: corresopndingly, there would be grid overlap or excessive skewness.

Among the dynamic grid velocity methods, there are methods proposed by Winkler, Mihalas and Norman [24,28,41], and Bell and Shubin [22], Hindman and Spencer [79], Rai and Anderson [23,80,81], Greenberg [82], Piva, DiCarlo, Favini and Gui [83], Ghia, Ghia and Shin [84], and Harten and Hyman [85]. Winkler, Mihalas and Norman [24,41] develop a scheme based upon the equidistribution of weights over grid point indices. Nonsingularity in their one dimensional context is provided primarily by creating singularity barriers which keep the points from interchanging positions. This comes from equidistributing the cell eccentricity as discussed earlier with Eqs. 102-105. In addition, they consider asymmetric time filtering to preserve resolution for the rapid reappearance of salient phenomena such as in shock wave reflections. The chief mechanism is a diffusion coefficient arising from a constructed factor (K in Eq. 97) on the time derivative. The factor contains enough residual memory to slow down the diffusion of the resolution just enough to allow for the rapid reappearance of small structures. Otherwise, resolution would have to rapidly disappear and then reappear, thus causing unnecessary numerical errors because of the temporal jerkiness. In comparison, Bell and Shubin [22] remove temporal jerkiness by balancing the weight function equidistribution against the mangitudes of time derivatives in the variational format given in Eqs. 89 and 90. Their balancing coefficient was, however, only a constant and thus did not contain the residual memory as in the case of Winkler, Mihalas and Norman.

In another direction, Hindman and Spencer [79] converted the Poisson system of Eq. 125 into a grid velocity equation by formal temporal differentia-

tion of the original equation. In addition, they also explored the use of equidistributed weight functions. They found the same relationship to the forcing function in one dimension that Anderson and Steinbrenner [47,48] eventually discovered in two dimensions.

Identifications with alternate metaphors, such as the earlier spring analogy, have also provided the inspiration for several methods. In this spirit, Rai and Anderson [23,80,81] developed an analogy to a gravitational potential while Greenberg [82] related the grid movement to chemical reactions. The gravitational forces decayed according ot an inverse power law for distances in the space of grid point indices. Force magnitudes and directions at each grid point came from the deviation of the error indicator from its average value. This was discussed earlier with Eqs. 91-93. In a similar manner, the chemical reaction rates contained the adaptive data.

An even more direct use of physically based motivation occurred in the somewhat parallel studies of Piva et al. [83] and Ghia, Ghia and Shin [84]. There, idealized two-dimensional momentum equations for viscous flow were transformed into diffusion equations. This was done because diffusion equations are easier to solve. The process basically amounts to a removal of the convective terms which would appear when the equations are expressed in terms of an arbitrary time-dependent coordinate system. The two resulting equations are the grid movement equations which assume the Poisson format. In a sense, this is similar to the pursuit of Hindman and Spencer [79], although there is no consideration of equidistribution.

In addition to the static methods based upon the previous solution step and the dynamic grid velocity methods, there are methods which impose a grid distribution mechanism at some implicit level without the direct determination of a grid velocity. This includes the methods such as that employed by White [31] or Mueller and Carey [86] and that employed with the moving finite-element method investigated by Miller and Miller [87], Miller [88], Miller [89], Gelinas, Doss and Miller [90], Herbst, Mitchell and Schoobie [91], Baines [92], and Baines and Wathen [93]. In the case of White [31] and others like it, the grid equation appears as a time-dependent constraint which is applied in an implicit coupled manner with the evolutionary physical equation. In the moving finite-element method, the coupling and the grid evolution comes directly out of the formal finite-element process when it is directly extended to include grid point motion.

In contrast to the dynamic or implicit temporal treatment of grid point motion, the static methods offer a great amount of simplicity, efficiency and spatial control at the expense of losing the accurate tracking of severe disturbances. The corrective tracking measures taken are typically to either use a smaller time step or preferably to require a broader band of resolution. With the broader band, the idea is that the disturbance will still appear in the high resolution region at the end of the time step. Methods which lead to such breadth typically come from grid smoothness forces and from curvature clustering on the monitor surface.

The static methods also tend to offer more numerical stability for the class of problems where the broad-banded resolution provides an adequate buffer for the containment of the disturbance. In the application, either the grid velocities are employed with backward differences in time or the solution is simply interpolated onto the new grid point locations in what is known as remapping step. While both are commonly used, the remapping approach is more prevalant, particularly in cases where the steady-state convergence is a prime element. In the steady-state cases, the movement may start with direct interlacing between the PDE-solver and the grid generator and then progress to fewer and fewer applications of grid movement until movement is stopped altogether. An example of the interlaced approach is given by Palmerio and Dervieux [94]. In addition, they considered specified grids as in Eq. 83 and a form of time filtering similar to that of Winkler, Mihalas and Norman [41].

With trapezoidal finite elements in one space dimension and time, Davis and Flaherty [95] and Flaherty et al. [96] employed a static grid generation scheme for a PDE-solver that was well-adjusted for temporal evolution. The static grid was generated from data at the explicit time level n for use at the implicit time level n+1. In a sense, this represents a zeroth-order forward extrapolation of the grid in time. Because of the extrapolation, the tracking possibilities are absent. In contrast, Sanz-Serna and Christie [97] and Blom, Sanz-Serna and Verwer [98] consider a predictor step for the application of static grid generation. The essence of their idea is to apply the PDE-solver to get provisional solution values at n+1 and from those values to generate the grid at the implicit level n+1. Then with the grid determined at n+1, they get the actual solution at n+1. Although their development was one-dimensional in space, the idea of first predicting the implicit-level adaptive data is attractive for any application in any number of spatial dimensions.

The extra cost only amounts to an extra application of the PDE-solver. The benefits are the capability to accurately track rapidly moving severe disturbances by using static adaptive grid generators which are known to produce high quality versatile grids.

CONCLUSION

Adaptive grid generation is essential when rapidly varying solutions to partial differential equations are to be simulated in an orderly fashion. The fundamental character of the topic is that the necessary local resolution is dynamically provided while the regular ordering of points is preserved. With the preserved order, many of the best partial differential equation solvers are available and are easily formulated. The main spirit behind the various developments is the creation of effective grid movement strategies that can be coupled into a wide variety of PDE-solvers.

In our discussion, we have attempted to develop the topic of adaptive grid generation in a somewhat structured and coherent manner. The emphasis was on the basic concepts and the interrelationship between the various consequent methods. To maintain a general perspective, basic principles were considered from various viewpoints. This started with a dozen basic ways to view equidistribution in one dimension. It recurred again when the use of prescribed distributions and coefficients in the weights were explored. Upon a foundation of concepts developed in one dimension, the various higher-dimensional methods were then developed. The discussion closed with a consideration of the temporal coupling of PDE-solvers and grid generators.

ACKNOWLEDGEMENT

The support under NASA and Air Force grants NAG3-715 and AFOSR-86-0307 is gratefully acknowledged. The people directing the grants were Capt. John P. Thomas and Dr. James D. Wilson at AFOSR and Bernard H. Anderson at NASA Lewis Research Center.

REFERENCES

(1) Smith, R.E. (Ed.), Numerical Grid Generation Techniques, NASA Conference Publication 2166, NASA Langley Research Center, 1980.

(2) Thompson, J.F. (Ed.), Numerical Grid Generation, North-Holland 1982. (Also published as Vols. 10 and 11 of Applied Mathematics and Computation, 1982).

(3) Ghia, K.N. and Ghia, U. (Eds.), Advances in Grid Generation, FED-Vol. 5, ASME Applied Mechanics, Bioengineering, and Fluids Engineering Conference, Houston, 1983.

(4) Hauser, J. and Taylor, C. (Eds.), Numerical Grid Generation in Computational Fluid Dynamics, Pineridge Press, 1986.

(5) Babuska, I., Chandra, J. and Flaherty, J.E. (Eds.), Adaptive Computational Methods for Partial Differential Equations, SIAM 1983.

(6) Thompson, J.F., Warsi, Z.U.A. and Mastin, C.W., "Boundary-Fitted Coordinate Systems for Numerical Solution of Partial Differential Equations - A Review," Journal of Computational Physics, Vol. 47, 1982, pp. 1-108.

(7) Thompson, J.F., "Grid Generation Techniques in Computational Dynamics," AIAA Journal, Vol. 22, 1984, pp. 1505-1523.

(8) Eiseman, P.R., "Grid Generation for Fluid Mechanics Computations," Annual Review of Fluid Mechanics, Vol. 17, 1985, pp. 487-522.

(9) Eiseman, P.R. and Erlebacher, G., "Grid Generation for the Solution of Partial Differential Equations," State-of-the-Art Surveys on Computational Mechanics, Ed. by A.K. Noor, ASME, to appear June 1987.

(10) Anderson, D.A., "Adaptive Grid Methods for Partial Differential Equations," Advances in Grid Generation, Ed. by K.N. Ghia and U. Ghia, 1983, pp. 1-15.

(11) Thompson, J.F., "A Survey of Dynamically Adaptive Grids in the Numerical Solution of Partial Differential Equations," Applied Numerical Mathematics, Vol. 1, 1985, pp. 3-27.

(12) Thompson, J.F., "Elliptic Grid Generation," Numerical Grid Generation, Ed. by J.F. Thompson, North-Holland, 1982, pp. 79-105.

(13) Stuben, K. and Linden, J., "MultiGrid Methods: An Overview with Emphasis on Grid Generation Processes," Numerical Grid Generation in Computational Fluid Dynamics, Ed. by J. Hauser and C. Taylor, Pineridge Press, 1986, pp. 483-509.

(14) Ortega, J.M. and Voigt, R.G., "Solution of Partial Differential Equations on Vector and Parallel Computers," SIAM Review, Vol. 27, No. 2, 1985, pp. 149-240.

(15) Steinberg, S. and Roache, P.J., "Variational Grid Generation," Numerical Methods in Partial Differential Equations, Vol. 2, 1986, pp. 71-96.

(16) Elsgolic, L.E., Calculus of Variations, Pergamon Press (London) and Addison-Wesley (Reading, MA), 1962.

(17) Dwyer, H.A., "Grid Adaption for Problems in Fluid Dynamics," AIAA Journal, Vol. 22, No. 12, 1984, pp. 1705-1712.

(18) Royden, H.L., Real Analysis, 2nd Ed., Macmillan, New York, 1968.

(19) Lehmann, E.L., Testing Statistical Hypotheses, Wiley, New York, 1966.

(20) Nakahashi, K. and Deiwert, G.S., "Three-Dimensional Adaptive Grid Method," AIAA Journal, Vol. 24, No. 6, 1986, pp. 948-954.

(21) Eiseman, P.R., "Orthogonal Grid Generation," Numerical Grid Generation, Ed. by J.F. Thompson, North-Holland, 1982, pp. 192-233.

(22) Bell, J.B. and Shubin, G.R., "An Adaptive Grid Finite Difference Method for Conservation Laws," Journal of Computational Physics, Vol. 52, 1983, pp. 569-591.

(23) Rai, M.M. and Anderson, D.A., "Grid Evolution in Time Asymptotic Problems," Numerical Grid Generation Techniques, Ed. by R.E. Smith, 1980, pp. 409-430, also in Journal of Computational Physics, Vol. 43, 1981, pp. 327-344.

(24) Winkler, K-H. A., Mihalas, D., and Norman, M.L., "Adaptive Grid Methods with Asymmetric Time-Filtering," Computer Physics Communications, Vol. 36, 1985, pp. 121-140.

(25) Eiseman, P.R., Geometric Methods in Computational Fluid Dynamics, ICASE Report 80-11 and von Karman Institute for Fluid Dynamics Lecture Series Notes, 1980.

(26) Warsi, Z.U.A., Tensors and Differential Geometry Applied to Analytic and Numerical Coordinate Generation, MSSU-EIRS-81-1, Aerospace Engineering, Mississippi State University, 1981.

(27) Eiseman, P.R., "Alternating Direction Adaptive Grid Generation," Proceedings of the 6th AIAA Computational Fluid Dynamics Conference, Danvers, MA, 1983, pp. 339-348. (Also in AIAA Journal, Vol. 23, No. 4, 1985, pp. 551-560).

(28) Winkler, K-H.A., "A Numerical Procedure for the Calculation of Nonsteady Spherical Shock Fronts with Radiation," Ph.D. Thesis, University of Gottingen. Issued as Technical Report MPI-PAE-Astro 90 by the Max-Planck-Institut fur Physik and Astrophysik, Munich; English translation: Lawrence Livermore Laboratory, Report UCRL-Trans-11206, University of California, Livermore, 1976.

(29) Ablow, C.M. and Schecter, S., "Campolytropic Coordinates," Journal of Computational Physics, Vol. 27, 1978, pp. 351-362.

(30) White, A.B., Jr., "On Selection of Equidistributing Meshes for Two-Point Boundary-Value Problems," SIAM Journal of Numerical Analysis, Vol. 16, 1982, pp. 472-502.

(31) White, A.B., Jr., "On the Numerical Solution of Initial Boundary-Value Problems in One Space Dimension," SIAM Journal of Numerical Analysis, Vol. 19, 1982, pp. 683-697.

(32) Dwyer, H.A., Kee, R.J. and Sanders, B.R., "An Adaptive Grid Method for Problems in Fluid Mechanics and Heat Transfer," Proceedings of the 4th AIAA Computational Fluid Dynamics Conference, Ed. by E. Murman, Williamsburg, VA, 1979, pp. 195-203, also in AIAA Journal, Vol. 18, 1980, pp. 205-212.

(33) Dwyer, H.A., Smooke, M.D. and Kee, R.J., "Adaptive Gridding for Finite Difference Solutions to Heat and Mass Transfer Problems," Numerical Grid Generation, Ed. by J.F. Thompson, North-Holland, 1982, pp. 339-356.

(34) Dwyer, H.A., "A Discussion of Some Criteria for the Use of Adaptive Gridding," Adaptive Computational Methods for Partial Differential Equations, Ed. by I. Babuska, J. Chandra and J.E. Flaherty, SIAM, 1983, pp. 111-122.

(35) Dwyer, H.A. and Onyejekwe, O.O., "Generation of Fully Adaptive and/or Orthogonal Grids," Proceedings of the 9th International Conference on Numerical Methods in Fluid Dynamics, Ed. by Soubbaramayer and Boujot, Saclay, France, Lecture Notes in Physics 218, Springer-Verlag, 1985, pp. 203-207.

(36) Ablow, C.M., "Equidistant Mesh for Gas Dynamic Calculations," Numerical Grid Generation, Ed. by J.F. Thompson, 1982, pp. 859-863.

(37) Gnoffo, P.A., "A Vectorized, Finite Volume, Adaptive-Grid Algorithm for Navier-Stokes," Numerical Grid Generation, Ed. by J.F. Thompson, 1982, pp. 819-835.

(38) Warsi, Z.U.A. and Thompson, J.F., "A Noniterative Method for the Generation of Orthogonal Coordinates in Doubly-Connected Regions," Numerical Grid Generation Techniques, Ed. by R.E. Smith, 1980, pp. 519-544, also in Mathematics of Computation, Vol. 38, 1982, pp. 501-516.

(39) Nakahashi, K. and Deiwert, G.S., "A Practical Adaptive Grid Method for Complex Fluid Flow Problems," Proceedings of the 9th International Conference on Numerical Methods in Fluid Dynamics, Ed. by Soubbaramayer and Boujot, Saclay, France, Lecture Notes in Physics 218, Springer-Verlag, 1985, pp. 422-426.

(40) Nakahashi, K. and Deiwert, G.S., "A Self-Adaptive Grid Method with Application to Airfoil Flow," Proceedings of the AIAA 7th Computational Fluid Dynamics Conference, Cincinnati, OH, 1985, pp. 340-350.

(41) Winkler, K-H.A., Norman, M.L. and Mihalas, D., "Implicit Adaptive Grid Radiation Hydrodynamics," Multiple Time Scales, Ed. by J.U. Brackbill and B.I. Cohen, Academic Press, 1985, pp. 141-184.

(42) Eiseman, P.R., "Alternating Direction Adaptive Grid Generation for Three-Dimensional Regions," Proceedings of the 10th International Conference on Numerical Methods in Fluid Dynamics, Lecture Notes in Physics 264, Springer-Verlag, 1986, pp. 258-263.

(43) Winslow, A., "Numerical Solution of the Quasilinear Poisson Equation in a Nonuniform Triangle Mesh," Journal of Computational Physics, Vol. 2, 1967, pp. 149-172.

(44) Thompson, J.F., Thames, F.C., Mastin, C.W., "Automatic Numerical Generation of Body-Fitted Curvilinear Coordinate System for Field Containing any Number of Arbitrary Two-Dimensional Bodies," Journal of Computational Physics, Vol. 15, 1974, pp. 299-319.

(45) Middlecoff, J.F., Thomas, P.D., "Direct Control of the Grid Point Distribution in Meshes Generated by Elliptic Equations," Proceedings of the 4th AIAA Computational Fluid Dynamics Conference, Williamsburg, VA, 1979, pp. 175-179. (Also in AIAA Journal, Vol. 18, No. 6, 1980, pp. 652-656).

(46) Thompson, J.F., Warsi,, Z.U.A. and Mastin, C.W., Numerical Grid Generation, Foundations and Applications, North-Holland, 1985.

(47) Anderson, D.A. and Steinbrenner, J., "Generating Adaptive Grids with a Conventional Grid Scheme," AIAA 24th Aerospace Sciences Meeting, Reno, NE, 1986, paper AIAA-86-0427.

(48) Anderson, D.A., "Constructing Adaptive Grids with Poisson Grid Generation," Numerical Grid Generation in Computational Fluid Dynamics, Ed. by J. Hauser and C. Taylor, Pineridge Press, 1986, pp. 125-136.

(49) Laugwitz, D., Differential and Riemannian Geometry, Academic Press, New York, 1965.

(50) Anderson, D.A., "Adaptive Grid Scheme Controlling Cell Area/Volume," AIAA 25th Aerospace Sciences Meeting, Reno, NE, 1987, paper AIAA-87-0202.

(51) Winslow, A., "Adaptive Mesh Zoning by the Equipotential Method," UCID-19062, 1981.

(52) Epstein, B., Partial Differential Equations: An Introduction, McGraw-Hill, 1962.

(53) Diaz, A.R., Kikuchi, N. and Taylor, J.E., "A Method of Grid Optimization for Finite-Element Methods," Computational Methods in Applied Mechanics and Engineering, Vol. 41, 1983, pp. 29-45.

(54) Oden, J.T., Devloo, P. and Strouboulis, T., "Adaptive Finite-Element Methods for the Analysis of Inviscid Compressible Flow: I. Fast Refinement/Unrefinement and Moving Mesh Methods for Unstructured Meshes," to appear in Computational Methods in Applied Mechanics and Engineering.

(55) Schwartz, A.L. and Connett, W.C., "Evaluating Algebraic Adaptive Grid Strategies," Numerical Grid Generation in Computational Fluid Dynamics, Ed. by J. Hauser and C. Taylor, Pineridge Press, 1986, pp. 163-174.

(56) Connett, W.C., Agarwal, R.K. and Schwartz, A.L., "An Adaptive Grid Generation Scheme for Flow Field Calculations," AIAA 25th Aerospace Sciences Meeting, Reno, NE, 1987, Paper AIAA-87-0199.

(57) Erlebacher, G. and Eiseman, P.R., "Adaptive Triangular Mesh Generation," AIAA 17th Fluid Dynamics, Plasma Dynamics, and Lasers Conference," 1984, Paper AIAA-84-1607, AIAA Journal, to appear.

(58) Eiseman, P.R., "Solution Adaptivity Using a Triangular Mesh," The Fre-Lagrange Method, Lecture Notes in Physics 238, Springer-Verlag, New York, 1985, pp. 206-235.

(59) Erlebacher, G., "Solution Adaptive Triangular Meshes with Application to Plasma Equilibrium," Doctoral Thesis, Department of Applied Physics and Nuclear Engineering, Columbia University, 1984.

(60) Eiseman, P.R., "Adaptive Grid Generation by Mean-Value Relaxation," Advances in Grid Generation, Ed. by K.N. Ghia, and U. Ghia, ASME-FED-Vol. 5, 1983, pp. 29-34, also in ASME Journal of Fluids Engineering, Vol. 107, 1985, pp. 477-483.

(61) Kennon, S.R. and Dulikravich, G.S., "Generation of Computational Grids Using Optimization," AIAA Journal, Vol. 24, No. 7, 1986, pp. 1069-1073.

(62) Kennon, S.R. and Dulikravich, G.S., "Composite Computational Grid Generation Using Optimization," Numerical Grid Generation in Computational Fluid Dynamics, Ed. J. Hauser and C. Taylor, Proceedings of the 1st International Conference in Landshut, West Germany, July 1986, pp. 217-225.

(63) Brackbill, J.U. and Saltzman, J.S., "Adaptive Zoning for Singular Problems in Two Dimensions," Journal of Computational Physics, Vol. 46, 1982, p. 342.

(64) Brackbill, J.U. and Saltzman, J., "An Adaptive Computation Mesh for the Solution of Singular Perturbation Problems," Numerical Grid Generation Techniques, Ed. by R.E. Smith, 1980, pp. 193-196.

(65) Brackbill, J.U., "Coordinate System Control: Adaptive Meshes," Numerical Grid Generation, Ed. by J.F. Thompson, North-Holland, 1982, pp. 277-294.

(66) Saltzman, J. and Brackbill, J., "Applications and Generalizations of Variational Methods for Generating Adaptive Meshes," Numerical Grid Generation, Ed. by J.F. Thompson, North-Holland, 1982, pp. 865-884.

(67) Yanenko, N.N. et al., "On Some Methods for the Numerical Simulation of Flows with Complex Structure," Proceedings of the 6th International Conference on Numerical Methods in Fluid Dynamics, Lecture Notes in Physics 90, New York, Springer-Verlag, 1978, pp. 565-578.

(68) Saunders, B.V., "Algebraic Grid Generation Using Tensor Product B-Splines," Old Dominion University, Ph.D. Thesis, 1985, also NASA CR 177968, 1985.

(69) Kreis, R.I., Thames, F.C. and Hassan, H.A., "Application of a Variational

Method for Generating Adaptive Grids," AIAA Journal, Vol. 24, No. 3, 1985, pp. 404-410.

(70) Morice, P., "Numerical Generation of Boundary-Fitted Coordinate Systems with Optimal Control of Orthogonality," Advances in Grid Generation, Ed. by K.N. Ghia and U. Ghia, ASME-FED-Vol. 5, 1983, pp. 29-34; also in ASME Journal of Fluids Engineering, Vol. 107, 1985, pp. 71-78; also 1983, ONERA-TP-082.

(71) Oskam, B. and Huizing, G.H., "Flexible Grid Generation for Complex Geometry in Two Space Dimensions Based on Variational Principles," National Aerospace Laboratory NLR, Amsterdam, The Netherlands, 1986, Report NLR MP86024L.

(72) Steger, J.L. and Chaussee, D.S., "Generation of Body-Fitted Coordinates Using Hyperbolic Partial Differential Equations," SIAM Journal of Scientific and Statistical Computing, Vol. 1, 1980, pp. 431-437.

(73) Thompson, J.F., "A Composite Grid Generation Code for General 3-D Regions," AIAA 25th Aerospace Sciences Meeting, Reno, NE, 1987, Paper AIAA-87-0275.

(74) Gough, D.O., Spiegel, E.A. and Toomre, J., "Highly Stretched Mesh as Functionals of Solutions," Proceedings of the 4th International Conference on Numerical Methods in Fluid Dynamics, Lecture Notes in Physics 35, Springer-Verlag, 1975, p. 191.

(75) Pierson, B.L. and Kutler, P., "Optimal Nodal Point Distribution for Improved Accuracy in Computational Fluid Dynamics," AIAA Journal, Vol. 18, No. 1, 1980, pp. 49-54.

(76) Roache, P.J. and Steinberg, S., "A New Approach to Grid Generation Using a Variational Formulation," Proceedings of the AIAA 7th Computational Fluid Dynamics Conference, Ed. by U. Ghia, Cincinnati, OH, 1985, pp. 360-370.

(77) Steinberg, S. and Roache, P.J., "Symbolic Manipulation and Computational Fluid Dynamics," Journal of Computational Physics, Vol. 57, 1985, pp. 251-284.

(78) Hyman, J.M. and Naughton, M.J., "Static Rezone Methods for Tensor Product Grids," Lectures in Applied Mathematics, Vol. 22, American Mathematical Society, 1985, pp. 321-343.

(79) Hindman, R.G. and Spencer, J., "A New Approach to Truly Adaptive Grid Generation," AIAA 21st Aerospace Sciences Meeting, Reno, NE, 1983, Paper AIAA-83-0450.

(80) Rai, M.M. and Anderson, D.A., "Application of Adaptive Grids to Fluid Flow Problems with Asymptotic Solutions," AIAA Journal, Vol. 20, 1982, pp. 496-502.

(81) Anderson, D.A. and Rai, M.M., "The Use of Solution Adaptive Grids in Solving Partial Differential Equations," Numerical Grid Generation, Ed.

by J.F. Thompson, North-Holland, 1982, pp. 317-338.

(82) Greenberg, J.B., "A New Self-Adaptive Grid Method," Proceedings of the 6th AIAA Computational Fluid Dynamics Conference, Ed. by H. Dwyer, Danvers, MA, 1983, pp. 323-332; also in AIAA Journal, Vol. 23, No. 2, 1985, pp. 317-320.

(83) Piva, R. et al., "Adaptive Curvilinear Grids for Large Reynolds Number Viscous Flows," Proceedings of the 8th International Conference on Numerical Methods in Fluid Dynamics, Ed. by E. Krause, Lecture Notes in Physics, Vol. 170, Springer-Verlag, 1982, pp. 414-419.

(84) Ghia, K., Ghia, U. and Shin, C.T., "Adaptive Grid Generation for Flows with Local High Gradient Regions," Advances in Grid Generation, Ed. by Ghia and Ghia, ASME-FED 5, 1983, pp. 35-48.

(85) Harten, A. and Hyman, J.M., "A Self-Adjusting Grid for the Computation of Weak Solutions of Hyperbolic Conservation Laws," Journal of Computational Physics, Vol. 50, 1983, pp. 235-269.

(86) Mueller, A. and Carey, G.F., "Continuously Deforming Finite Elements for Convection Dominated Flows," Advances in Grid Generation, Ed. by Ghia and Ghia, ASME-FED 5, 1983, pp. 135-138.

(87) Miller, K. and Miller, R.N., "Moving Finite Elements I," SIAM Journal of Numerical Analysis, Vol. 18, No. 6, 1981, pp. 1019-1032.

(88) Miller, K., "Moving Finite Elements II," SIAM Journal of Numerical Analysis, Vol. 18, No. 6, 1981, pp. 1033-1057.

(89) Miller, K., "Alternate Modes to Control the Nodes in the Moving Finite Element Method," Adaptive Computational Methods for Partial Differential Equations, Ed. by I. Babuska, J. Chandra and J.E. Flaherty, SIAM, 1983, pp. 165-182.

(90) Gelinas, R., Doss, S. and Miller, K., "The Moving Finite-Element Method: Applications to General Equations with Multiple Large Gradients," Journal of Computational Physics, Vol. 40, 1981, pp. 202-249.

(91) Herbst, B., Mitchell, A. and Schoobmie, S., "A Moving Petrov-Galerkin Method for Transport Equations," International Journal for Numerical Methods in Engineering, Vol. 18, 1982, pp. 1321-1336.

(92) Baines, M.J., "On Approximate Solutions of Time-Dependent Partial Differential Equations by the Moving Finite-Element Method," Numerical Analysis Report 1/85, University of Reading, UK; Numerical Methods in Fluid Dynamics II, Ed. by K.W. Morton and M.J. Baines, Oxford University Press, 1986, p. 1.

(93) Baines, M.J. and Wathen, A.J., "Moving Finite Element Modeling of Compressible Flow," Applied Numerical Mathematics, Vol. 2, No. 6, 1986, pp. 495-514.

(94) Palmerio, B. and Dervieux, A., "Application of a FEM Moving Node Adaptive

Method to Accurate Shock Capturing," Numerical Grid Generation in Computational Fluid Dynamics, Ed. by J. Hauser and C. Taylor, Pineridge Press, 1986, pp. 425-436.

(95) Davis, S.F. and Flaherty, J.E., "An Adaptive Finite-Element Method for Initial-Value Problems for Partial Differential Equations," SIAM Journal of Scientific and Statistical Computing, Vol. 3, 1982, pp. 6-27.

(96) Flaherty, J.E. et al., "Adaptive Finite-Element Methods for Parabolic Partial Differential Equations," Adaptive Computational Methods for Partial Differential Equations, Ed. by I. Babuska, J. Chandra and J.E. Flaherty, SIAM, 1983, pp. 144-164.

(97) Sanz-Serna, J.M. and Christie, I., "A Simple Adaptive Technique for Nonlinear Wave Problems," Journal of Computational Physics, Vol. 67, No. 2, pp. 348-360.

(98) Blom, J.G., Sanz-Serna, J.M. and Verwer, J.G., "On Simple Moving Grid Methods for One-Dimensional Evolutionary Partial Differential Equations," Centrum voor Wiskunder end Informatica, Department of Numerical Mathematics, Report NM-R8620, 1986.

A General Purpose Time Accurate Adaptive Grid Method

Michael J. Bockelie*, Peter R. Eiseman@, and Robert E. Smith*

* NASA Langley Research Center
Hampton, VA 23669

@ Columbia University
New York, NY 10027

Abstract of Paper Submitted to the 12th International Conference on Numerical Methods in Fluid Dynamics, July 9-13, 1990, Oxford, England.

A general purpose adaptive grid method will be presented that has the ability to generate time accurate grids for a wide variety of problems. The solution method consists of three parts: a grid movement scheme; a PDE solver; and a temporal coupling routine that links the dynamic grid and the PDE solver. The ability of the basic scheme to perform time accurate adaptive computations has been previously established [1]. Here, we will present a new innovation for grid control, address efficiency improvements in the temporal coupling and issues pertaining to the transfer of solution data between successive grids, and demonstrate the adaptive solution method on a supersonic flow problem.

The basic grid movement scheme employs a "monitor surface" formed from the solution data to identify regions requiring further grid resolution. The grid points are repositioned on a curve by curve basis by requiring that a weight function be equally distributed over each curve. Including the monitor surface gradient and normal curvature in the weight function ensures that both the gradients and the transition regions of the solution are resolved. The grid movement is interwoven with a smoothing operation to ensure grid smoothness and to alleviate grid skewness. The scheme also contains a means of eliminating "outlier" values in the adaptive data and a grid control that can enforce a prescribed minimum grid cell size.

A new grid control will be presented that can **accurately track** *multiple solution features.* At present, most adaptive grid methods can only track a single solution feature, or at best a linear combination of multiple features. Treating the adaptive data as a *scalar* function can result in an inaccurate grid if the solution features should merge [2]. This limitation can be overcome by combining the individual features to be tracked into a monitor surface which is a *vector* function [2]. With respect to the vector monitor surface, examples of the improved grid resolution it provides and the required modifications for computing its curvature properties will be discussed.

A simple **predictor corrector** scheme is used to couple the adaptive grid to the PDE solver. The scheme treats the time integration as a series of initial value problems in which the solution is first advanced on the existing grid for the purposes of obtaining a new grid, the solution data is then interpolated from the old grid to the new grid, and last, the solution is recomputed on the new grid. This approach creates a time accurate grid because the grid does not lag the solution in time. Furthermore, grid velocity terms are not required because the solution is computed on a static grid. Techniques will be presented that reduce the computational time consumed by the algorithm described in earlier work [1]. In addition, tests to determine the effect on the solution of using a **conservative data transfer** method, as opposed to the **bilinear interpolation** used at present, will be summarized.

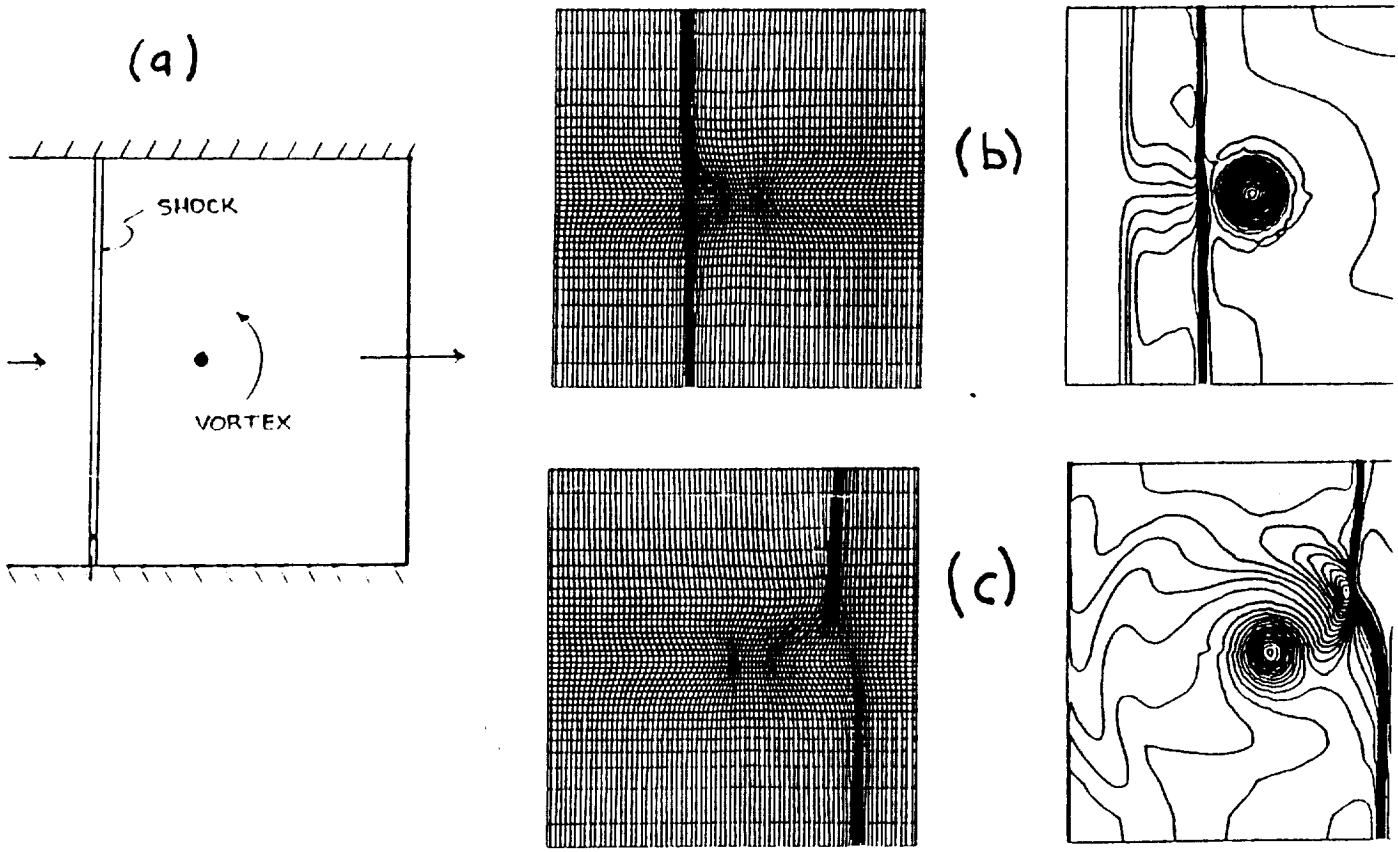
The capabilities of the adaptive method will be demonstrated by computing a time accurate solution for the inviscid unsteady flow field created by a **shock vortex** interaction. The shock vortex problem consists of an initially planar shock wave marching toward and eventually over a solid core vortex lying a short distance upstream. Here, the objective is to generate a grid that will track the shock wave, the vortex, and the sound wave emitted by the interaction;

these three features have substantially different magnitudes, and locations. Thus, the shock vortex problem is a particularly good test of an adaptive method because it requires a grid with locally high resolution whose location changes with time to properly capture both the severe behavior of the shock wave and the much more gentle features of the sound wave and the vortex. In this study we will focus on problems in which the shock wave is propagating at a Mach number slightly greater than 1 and thereby results in more localized behavior that must be refined, as compared to our earlier work in which the shock wave was propagating at Mach 3 [1,2]. A preliminary result in which the shock wave is propagating at Mach 1.1 is contained in Fig. 1.

REFERENCES

1. Eiseman, P.R. and Bockelie, M.J., "Adaptive Grid Solution For Shock Vortex Interaction," *Proceedings of the 11th International Conference on Numerical Methods in Fluid Dynamics*, Dwoyer, Hussaini, and Voigt, eds., Lecture Notes in Physics, Vol. 323, Springer Verlag, 1989, pp. 240-244.

2. Bockelie, M.J., *Adaptive Grid Movement Scheme and the Numerical Simulation of Shock Vortex Interaction*, Doctoral Thesis, Columbia University, 1988.



g. 1. For the shock vortex problem: problem definition (a), grid and pressure contour plots before (b) and after (c) interaction for a grid adapted to the shock wave and the vortex using the vector monitor surface.