Submitted to Logistics and
Transportation Review

LBL-3360 c.2
Preprint

FINITE ELEMENT APPROACH TO BICUBIC SPLINE CONSTRUCTION

Jonathan D. Young

October 1974

Prepared for the U. S. Atomic Energy Commission
under Contract W-7405-ENG-48

FINITE ELEMENT APPROACH TO BICUBIC SPLINE CONSTRUCTION*

Jonathan D. Young

Lawrence Berkeley Laboratory
University of California
Berkeley, California 94720

October 1974

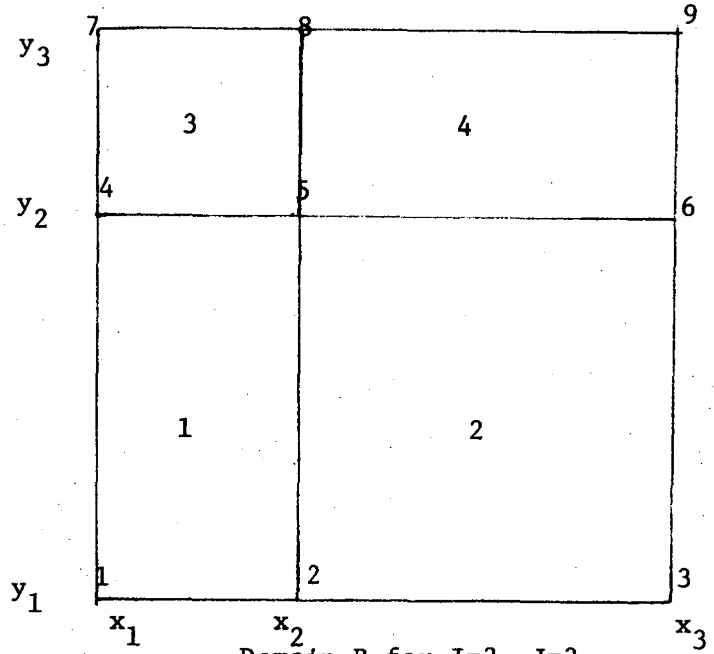
Abstract

A bicubic spline is constructed on a rectangular domain using the finite element approach. The elements are subrectangles of the domain. The process described computes unknown function and derivative values at rectilinear gridpoints for the bicubic spline determined by various specified values at these points.

* This work was done in part under the auspices of the Atomic Energy Commission.

INTRODUCTION

We shall consider a rectangular domain, R, in the XY plane. We assume a rectilinear grid on R formed by the lines $x=x_i$; i=1, I and $y=y_j$; j=1, J. Thus, the vertices of R are (x_1, y_1), (x_I, y_1), (x_I, y_J) and (x_1, y_J). Then we have IJ gridpoints and (I-1) (J-1) subrectangles. We shall call the gridpoints, points and the subrectangles, elements.



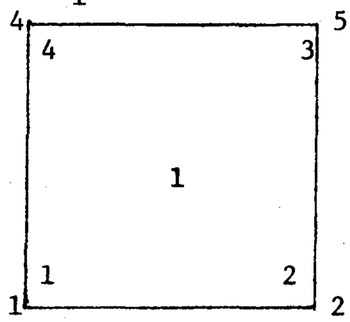
Domain R for I=3, J=3

Note that both I and J must be at least two, but it is not necessary that I=J or that the subintervals in X or Y be of uniform length.

We set N=IJ and M=(I-1) (J-1) and index (number) the gridpoints; with n=1, N and the elements e_m with m=1, M. The indicing may be done in any convenient manner, the one shown in the figure (altho common), is only suggestive. We now abandon our i and j as indices and use

(x_n, y_n) n=1, N

Note that each element, e_m, has four vertices; (x_{n_k}, y_{n_k}) for k=1, 4. For example, e_1, above



$n=1, 2, 5, 4$
$k=1, 2, 3, 4$

Element e_1

We could index the vertices in the manner shown giving $n_1=1$, $n_2=2$, $n_3=5$, $n_4=4$. This indicing (altho common) is only suggestive.

Note also that with the exception of the four points which are vertices of the original rectangle, every gridpoint is a common vertex for at least two elements, and may be common to four elements (example, (x_5, y_5) above).

We now consider a bicubic spline function, u, defined on R. By definition, it must have these properties:

(1) On any element, e_m, u is a bicubic in x and y. (1)

$$u = \sum_{\mu=0}^{3} \sum_{\nu=0}^{3} c_{\mu\nu} x^{\mu} y^{\nu} \text{ on } e_m$$

(2) The function, u, has continuous second derivatives (u_{xy}, u_{xx}, u_{yy}) and mixed third derivatives, u_{xxy}, and u_{xyy} on R.

Just as it was useful to replace the double index (i,j) with n, it is now desirable to replace the double index (μ,ν) with a single index, ℓ. This can be done in a number of ways (simply write out the whole expression and number the [sixteen] terms in any order). For convenience, we suggest:

$$u = \sum_{\ell=1}^{16} c_\ell \, x^{(\ell-1)\mathrm{mod}\ 4} \, y^{[(\ell-1)-(\ell-1)\mathrm{mod}\ 4]/4} \tag{2}$$

Since the (x_{n_k}, y_{n_k}); k=1, 4 are known in the element e_m, and in general the c_ℓ; ℓ = 1, 16 are unknown, we write the above in the conventional way:

$$\Sigma t_\ell \, c_\ell = u \tag{3}$$

where

$$t_\ell = x^{(\ell-1)\mathrm{mod}\ 4} \quad y^{[(\ell-1)-(\ell-1)\mathrm{mod}\ 4]/4} \tag{4}$$

Equation (3) may be applied at the four vertices of e_m giving four equations:

$$\Sigma t_{\ell,k} \, c_\ell = u_k \qquad k = 1,4 \tag{5}$$

where

$$t_{\ell,k} = x_k^{(\ell-1)\mathrm{mod}\ 4} \quad y_k^{[(\ell-1)-(\ell-1)\mathrm{mod}\ 4]/4}$$

with

$$u_k = u_{n_k}$$

$$x_k = x_{n_k}, \; y_k = y_{n_k}$$

Since we have 16 unknowns, c_ℓ, ℓ = 1, 16 we need 12 more equations. These can be obtained by differentiating Equation (3) with respect to x, to y, and to both. This gives us:

$$\Sigma \, t_{x_\ell} \, c_\ell = u_x \tag{6}$$

$$\Sigma \, t_{y_\ell} \, c = u_y \tag{7}$$

$$\Sigma \, t_{xy_\ell} \, c_\ell = u_{xy} \tag{8}$$

where

$$t_{x_\ell} = [(\ell-1)\bmod 4]\; x^{(\ell-1)-(\ell-1)\bmod 4]/4} \tag{9}$$

$$t_{y_\ell} = [(\ell-1)-(\ell-1)\bmod 4]/4\; x^{(\ell-1)\bmod 4}\; y^{[(\ell-1)-(\ell-1)\bmod 4]/4-1} \tag{10}$$

$$t_{xy_\ell} = [(\ell-1)\bmod 4][(\ell-1)-(\ell-1)\bmod 4]/4\; x^{(\ell-1)\bmod 4-1}\; y^{[(\ell-1)-(\ell-1)\bmod 4]/4-1}$$

Applying Equations (6), (7) and (8) at the vertices of e_m, provides us with the additional 12 equations required. Our linear system may now be written:

$$T\vec{c} = \vec{u} \tag{11}$$

Where T is a 16 x 16 matrix whose first four rows are obtained from t, the next four from t_x, the next four from t_y and the last four from t_{xy}. The vector $\vec{u}$ has 16 components, the first four are u values, the next four u_x, the next four u_y, and the last four u_{xy}. This ordering of the rows of T and components of u is not essential but it is convenient. It can be easily shown that if the points (x_{n_k}, y_{n_k}) are distinct (as they would be if e_m were a rectangle), then T is non-singular and if the vector $\vec{u}$ is known (u, u_x, u_y, u_{xy} known at the four vertices), the coefficients, c_ℓ; ℓ=1, 16 are obtainable by:

$$\vec{c} = T^{-1}\vec{u} \tag{12}$$

We shall call the values for u, u_x, u_y, and u_{xy} at the vertices the "primary" values. In general, when these are known, we are more interested in finding secondary values u_{xx}, u_{yy}, u_{xxy} and u_{xyy} at the vertices. They are called secondary since u_{xx} can be determined from u and u_x; u_{yy} from u and u_y, u_{xxy} from u_y and u_{xy}; u_{xyy} from

u_x and u_{xy}. They are important since they are continuous [property (2)] hence, at any vertex common to two or four elements, the value computed for one element must agree with those for the other(s).

Differentiating Equation (6) with respect to x and Equation (7) with respect to y, we obtain:

$$\Sigma t_{xx_\ell} c = u_{xx_\ell} \tag{13}$$

and

$$\Sigma t_{yy_\ell} c = u_{yy_\ell} \tag{14}$$

Then, differentiating Equation (13) with respect to y and (14) with respect to x, we have:

$$\Sigma t_{xxy_\ell} c_\ell = u_{xxy} \tag{15}$$

and

$$\Sigma t_{xyy_\ell} c_\ell = u_{xyy} \tag{16}$$

We now apply Equations (13), (14), (15) and (16) at the vertices of e_m and obtain 16 equations for the linear system:

$$D \vec{c} = \vec{v} \tag{17}$$

where D is a 16 x 16 matrix whose first four rows come from t_{xx}, next four from t_{yy}, next four from t_{xxy}, and the last four from t_{xyy}. The vector $\vec{v}$ has sixteen components, the first four are u_{xx} values, the next four u_{yy}, the next four u_{xxy}, and the last four u_{xyy} (at the vertices).

Using the expression for $\vec{c}$ from Equation (12), in Equation (17) we have:

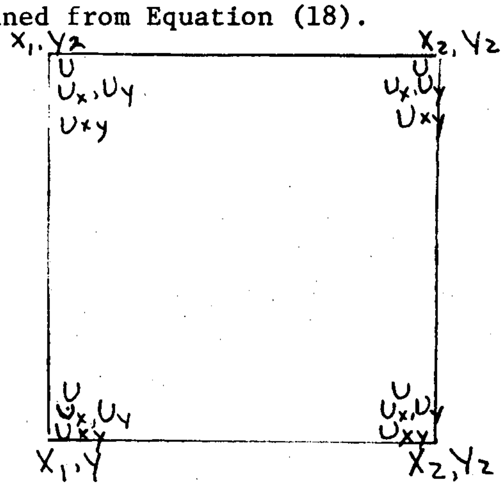
$$DT^{-1} \vec{u} = \vec{v} \tag{18}$$

Thus, if the values for u, u_x, u_y and u_{xy} are known at the vertices of e_m, the values for u_{xx}, u_{yy}, u_{xxy} and u_{xyy} may be readily obtained.

CONSTRUCTION OF A PARTICULAR BICUBIC SPLINE

By construction of a bicubic spline on R we mean finding unknown (if any) primary values (u, u_x, u_y, u_{xy}) and all secondary values (u_{xx}, u_{yy}, u_{xxy}, u_{xyy}) at all the gridpoints of R. Some primary values must be specified.

We have seen in the previous section that if there is only one element (I=2, J=2) then all primary values must be known and that all secondary values can be obtained from Equation (18).

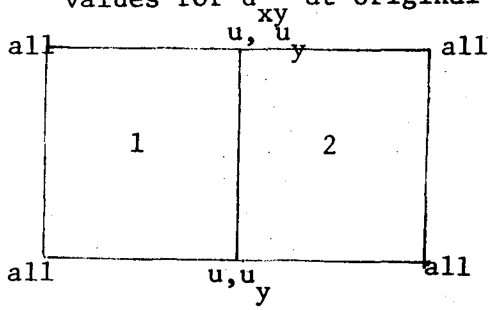


Element E_1

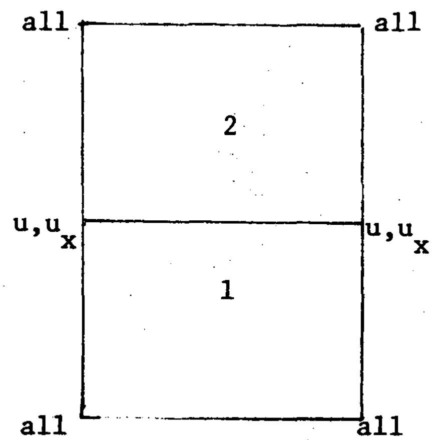
In the figure above, we must know x, y, u, u_x, u_y and u_{xy} at the four vertices, and we can then compute u_{xx}, u_{yy}, u_{xxy} and u_{xyy} at each of these points.

Now, for I=3, J=2 or I=2, J=3 we have two elements, and six gridpoints, but we only have twenty-four primary values, since eight values are shared. The two elements have two <u>common</u> gridpoints and we can use the continuity of u_{xx} and u_{xxy} to get u_x and u_{xy} (for I=3, J=2), or the continuity of u_{yy} and u_{xyy} to get u_y and u_{xy} (for I=2, J=3) at the common points.

Hence, it is sufficient to know values for u at all gridpoints, values for u_x on x boundary points, values for u_y at y boundary points, and values for u_{xy} at original vertices.



I=3, J=2



I=2, J=3

Thus, only 20 primary values need be specified and four can be determined. If fewer than 20 values are specified, a particular cubic spline cannot be constructed, if 20 are specified and they are "independent", we can construct the spline. If more than 20 values are specified, they must be consistent.

Let us examine how u_x and u_{xy} at the "inside" gridpoints on the y boundaries are determined by the continuity of u_{xx} and u_{xxy} for I=3, J=2. Applying the subscripts of e_1 and e_2 to the matrices T and D, vectors u and v in Equation (18)

$$T_1^{-1} D_1 \vec{u}_1 = \vec{v}_1 \tag{19}$$

$$T_2^{-1} D_2 \vec{u}_2 = \vec{v}_2 \tag{20}$$

Let us set:

$$A \equiv T_1^{-1} D_1$$

$$B \equiv T_2^{-1} D_2$$

and denote the elements of A and B by a_{ij} and b_{ij} respectively with i=1 to 16, j=1 to 16. Now, u_{xx} and u_{xxy} at the common gridpoints must have the same value as components of $\vec{v}_1$ as they do in $\vec{v}_2$. Hence, there are four values i* and i** for which:

$$v_{1i*} = v_{2i**}$$

hence

$$\vec{a}_{i*}\ \vec{u}_1 = \vec{b}_{i**}\ \vec{u}_2$$

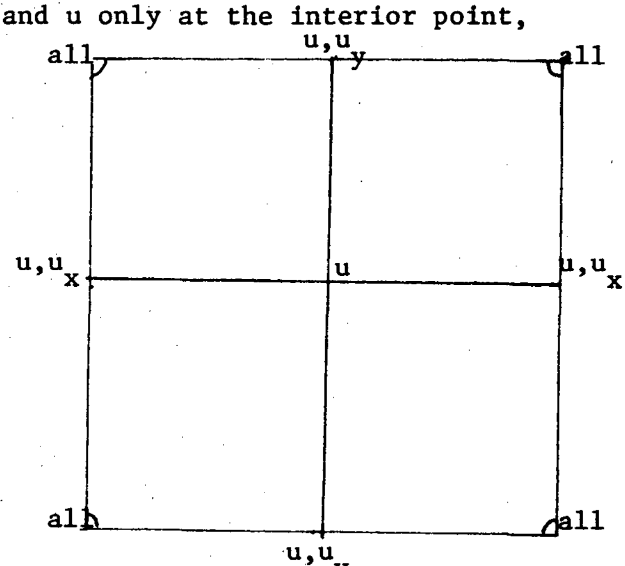
Where vectors a_{i*} and b_{i**} are simply the i* row of $T_1^{-1}\ D_1$ and i** row of $T_2^{-1}\ D_2$ respectively. (These are all known).

Equation (21) can be written:

$$\vec{a}_{i*_k}\ \vec{u}_1 - \vec{b}_{i**_k}\ \vec{u}_2 = 0 \qquad \text{for } k = 1 \text{ to } 4$$

giving us four equations to be solved for the four unknowns u_x and u_{xy} which are common components of u_1 and u_2.

Now let us consider the case I=3, J=3. We have 9 gridpoints and four elements. If we know u, u_x, u_y and u_{xy} at the four vertices of R, u and u_x at "inside" boundary points of x, u and u_y at "inside" boundary points of y and u only at the interior point,



then all the remaining unknown primary values can be determined using continuity of higher derivatives. Hence, twenty-five primary values (if independent) will determine the remaining eleven. The eleven equations required will be obtained by "matching" corresponding rows of the linear systems;

$$T_1^{-1} D_1 \vec{u}_1 = \vec{v}_1$$

$$T_2^{-1} D_2 \vec{u}_2 = \vec{v}_2$$

$$T_3^{-1} D_3 \vec{u}_3 = \vec{v}_3$$

$$T_4^{-1} D_4 \vec{u}_4 = \vec{v}_4$$

pertaining to the four elements.

In general, the number of primary values specified must be $IJ + 2I + 2J + 4$, while the number which can be computed will be $3IJ - 2I - 2J - 4$.

PROBLEM DEFINITION

In general, the following must be known to properly define a bicubic spline or the rectangle with a rectilinear grid, (x_i, y_j).

(1) The number of x_i denoted by I with $I \gtrsim 2$.

(2) The number of y_j denoted by J with $J \gtrsim 2$.

(3) The values of x_i for $i = 1, I$ and y_j for $j = 1, J$.

The number of points, N, and elements, M, are obtained from:

$$N = IJ$$
$$M = (I-1)(J-1)$$

(4) At (x_1, y_1), (x_I, y_1), (x_I, y_J) and (x_1, y_J) we must know the values for u, u_x, u_y, and u_{xy}.

(5) At (x_1, y_j) and (x_I, y_j) for $j = 2, J-1$ we must know u and u_x, or u_y and u_{xy}.

(6) At (x_i, y_1) and (x_i, y_J) for $i = 2, I-1$ we must know u and u_y or u_x and u_{xy}.

(7) At (x_i, y_j) for $i = 2, I-1$ and $j=2, J-1$ we must know u or u_x or u_y or u_{xy}.

In the usual problem the first alternative of 5), 6) and 7) occurs, but the other cases are determinate. If more values are specified than required, they must be consistent.

COMPUTATIONAL PROCEDURE

For each element, e_m, we must know the values of n, n_k, k=1, 4 at the vertices. Then we use the values of x_{n_k} and y_{n_k} to construct the matrices T_m and D_m and invert T_m obtaining T_m^{-1} and multiply $T_m^{-1} D_m$ to obtain a matrix A_m, for m=1,M. (All these matrices are 16 x 16).

Now, if all primary values are known, we obtain the unknown secondary values of e_m by

$$\vec{v}_m = A_m \vec{u}_m$$

where u consists of known values for u, u_x, u_y, and u_{xy} at the vertices of e_m and v_m provides values for u_{xx}, u_{yy}, u_{xxy} and u_{xyy} at these same vertices. To avoid redundancy in computation (vertices are shared by contiguous elements) we simply by-pass the computation if values have already been ascertained.

When some primary values are unknown, some and possibly all, of the vectors $\vec{u}_m$ have unknown components and these must be found before the process of the preceding paragraph can be applied.

If $\vec{u}_m$ is an unknown vector of 16 components, then it is a linear combination of a basis of 16 vectors. For convenience, we shall use the basis

$$\vec{b}_j = (a_{ij}) \text{ with } a_{ji} = 1 \text{ and } a_i = 0 \quad i \neq j \qquad \text{for } j=1 \text{ to } 16$$

(subscripts not to be confused with earlier i, j now abandoned)

When these are arranged suitably as column vectors, they constitute the identity (16 x 16) matrix

$$B = I_{16 \times 16}$$

Hence, we have

$$A_m B = A_m$$

Now, we note that

$$\vec{u}_m = B \vec{u}_m$$

which has the effect of multiplying the columns of B by "rows" of $\vec{u}_m$.

Hence, in the product

$$A_m B \vec{u}_m = A_m \vec{u}_m = B \vec{v}_m$$

in the effect the columns of A_m are multiplied by rows of $\vec{u}_m$.

Thus, for the known components of u_m, we can multiply 'corresponding' columns of A by that known value.

Thus, for any known component u_{m_i} we have

$$u_{m_i} = \sum_{j=1}^{16} a_{m_{i,j}}$$

and v_{m_i} can be computed.

Now, suppose only one u_{m_i} is unknown, and it occurs at m_k (for k=1, 2, 3 or 4).

Then:

$$\sum_{\substack{j=1 \\ j \neq i}}^{16} a_{m_{k,j}} + u_{m_i} a_{m_{ki}} = u_{{xx}_{n_k}}$$

$$\sum_{\substack{j=1 \\ j \neq i}}^{16} a_{m_{k+4,j}} \quad u_{m_i} a_{m_{k+4,i}} = u_{{yy}_{n_k}}$$

$$\sum_{\substack{j=1\\ j\neq i}}^{16} a_{m_{k+8,j}} + u_{m_i}\, a_{m_{k+8,i}} = u_{xxy_{n_k}}$$

$$\sum_{\substack{j=1\\ j\neq i}}^{16} a_{m_{k+12,j}} + u_{m_i}\, a_{m_{k+12,i}} = u_{xyy_{n_k}}$$

But none of the four can be solved for u_{m_i} since all right hand sides are unknown.

But there is at least one other m* in which the unknown value appears as an unknown at a point n_{k*} and as a component $u_{m*_{i*}}$. Now, at this "shared" point of e_m and e_{m*}.

We must have continuity as defined earlier hence.

$$u_{xx_{n_k}} = u_{xx_{n_{k*}}} \qquad u_{yy_{n_k}} = u_{yy_{n_{k*}}}$$

$$u_{xxy_{n_k}} = u_{xxy_{n_{k*}}} \qquad u_{xyy_{n_k}} = u_{xyy_{n_{k*}}}$$

Hence

$$\sum_{\substack{j=1\\ j\neq i}}^{16} a_{m_{\ell,j}} + u_{m_i}\, a_{m_{\ell,i}} = \sum_{\substack{j=1\\ j\neq i*}}^{16} a_{m*_{\ell*,j}} + u_{m*_{i*}}\, a_{m*_{\ell*,i*}}$$

for $\ell=k$, $k+4$, $k+8$, $k+12$ and $\ell*$, respectively, = $k*$, $k*+4$, $k*+8$, $k*+12$.

Thus, we have

$$(a_{m_{\ell,i}} - a_{m^*_{\ell^*,i^*}})\, u_{m_i} = \sum_{\substack{j=1\\ j\neq i\\ j\neq i^*}}^{16} (a_{m^*_{\ell^*,j}} - a_{m_{\ell,j}}) + a_{m^*_{\ell^*,i}} - a_{m_{\ell,i^*}}$$

giving us four possible equations to solve for the unknown u_{m_i}.

In general, some (but not all) of the coefficients $(a_{m_{\ell,i}} - a_{m^*_{\ell^*,i^*}})$ will be zero, and in that case, the equation cannot be used. But we need choose only one equation for which the coefficient is not zero and solve it for u_{m_i}. We now multiply the i^{th} column of A_m by this value and compute v_{m_i} by

$$v_{m_i} = \sum_{j=1}^{16} a_{m_{i,j}}$$

If the above fails (all coeff. of u_{m_i} are zero) we are (hopefully) at a point n shared by four elements. We can then find another m*, and if it fails, a last one, being assured that if the problem is well posed, there is at least one equation which can be solved. (Actually, if the point is in four elements, there are six pairs (m,m*) which can be considered, but we need never try more than three since continuity in the others is redundant.)

If there is no more than one unknown at any point, the above method can be used to solve successively for thoe unknowns.

Now, there may be two (on x or y boundaries) or even three (on interior) unknown at the vertices of e_m and e^*_m. In this case, additional equations from those available must be used. We can thus construct a linear system to be solved simultaneously for all the unknowns. We must, of course, be careful not to include the same equation twice.

COMPUTER CODE

A computer subroutine, FYNLBC, has been written to perform the computation described above. As known data, it required the values of N and M, an array of dimension (5,N) containing indicators (0 if known, if unknown) on the status of u, u_x, u_y and u_{xy}, and higher derivatives, respectively, at the point, n; arrays of dimension (N) of x_n, y_n; an array of dimension (8,N) containing known values of u, u_x, u_y and u_{xy}, respectively, in the first 4 rows; an array of dimension (4,M) containing values of n_k, k=1 to 4 for each element. An array of dimension (3,N) equivalent to x should be specified for storage of powers of x; an array of dimension (16, 16, M) for storage of matrices, A_m; arrays of dimension (MXU, MXU), (MXU) and (3,MXU) where MXU = 3IJ - 2I - 2J -4 must be provided as working space. These latter three are used in constructing the linear system to be solved for the unknown primary values.

Actually, the computer code can handle more general problems than those described above. The indicing in n and m is at the discretion of the user. The indicing 1 for u, 2 for u_x, 3 for u_y, 4 for u_y is essential in the known/unknown indicator and the array of the known values. Unknown values may be entered in this array as blanks, ones or whatever.

If a solution is achieved, the subroutine returns an indicator IFL=0, and the array of dimension (8, N) contains respectively values for u, u_x, u_y, u_{xy}, u_{xx}, u_{yy}, u_{xxy} and u_{xyy} at each n.

Otherwise, IFL≠0, and this array contains "indefinite" values in the array for unknowns it could not compute.

If:

IFL = 1 (Not enough points; must have N ≥ 4)

IFL = 2 (Too many points, exceeds dimension N)

IFL = 3 (Too many elements, exceeds dimension M)

IFL = 4 (Too many unknowns, exceeds dimension MXU)

IFL = 5 (Points are not distinct)

IFL = 6 (For one element, all primary values must be known)

IFL = 7 (Unknown does not appear in two or four elements)

IFL = 8 (Cannot find equation for an unknown)

IFL = 9 (Did not find enough equations)

IFL = 10 (Indeterminate system)

IFL > 10 (IFL -10 is the equation number which could not be found.)

ACKNOWLEDGEMENT

The author wishes to acknowledge the advice and assistance of Professors Paul A. Luger and Frank A. Valente of Seattle University.

TECHNICAL INFORMATION DIVISION
LAWRENCE BERKELEY LABORATORY
UNIVERSITY OF CALIFORNIA
BERKELEY, CALIFORNIA 94720