Collocation - A Simple, Flexible and Efficient Method for Integral and Differential Equations

Elias N. Houstis
Purdue University, enh@cs.purdue.edu

Report Number:
78-258
COLLOCATION--A SIMPLE, FLEXIBLE AND EFFICIENT METHOD
FOR INTEGRAL AND DIFFERENTIAL EQUATIONS

E. N. Houstis
Department of Computer Sciences
Purdue University
West Lafayette, Indiana 47907

CSD-TR 258
January 1978

Abstract

We present a collocation method based on piecewise cubic polynomials in C¹ for solving various integral and differential equations. Two new methods for handling curved boundaries are discussed based on collocation and discrete least-squares. Finally, a summary of the properties of the collocation method is given which indicates the simplicity, flexibility, efficiency, and generality of the method.
Introduction. In this report, we present an approximation theoretical method for solving integral and differential equations. The method is the so-called collocation whose idea is analogous to interpolation. Inspite, its simplicity, flexibility, and efficiency, the method is not well known, perhaps because it is difficult to theoretically analyse. We consider a rather specific instance of the method based on piecewise cubic polynomials in \( C^1 \) and apply it to various integral and differential equations. We discuss two new methods for handling curved boundaries based on collocation and discrete least-squares. Finally, we summarize the properties of the collocation method which make it more efficient for general use.

I. COLLOCATION METHOD

We present an abstract formulation of a differential operator problem and an approximation theoretic method for solving it. The problem can be defined in terms of four attributes.

Problem Abstraction

\[
\begin{align*}
\Omega &= \text{domain of } x \\
\Omega_0 &= \text{domain } x \\
Lu(x) &= f(x), \quad x \in \Omega \\
Bu(x) &= g(x), \quad x \in \Omega_0
\end{align*}
\]

operator equation

Auxiliary conditions

These problems are normally associated with two domains: one where an operator equation is to be satisfied and one where certain auxiliary conditions are to be satisfied (such as boundary conditions for differential equations).

An approximation theoretic method for approximating the solution of the above problem consists of three distinct steps.
Step 1 Choose a set of functions with which to approximate \( u(x) \).
Obtain a set of basis functions \( \{B_k\} \) for this choice.

Step 2 Choose a set of linear functionals, say \( \lambda_j \), \( j = 1, 2, \ldots \) that determines the conditions for the approximation. These are usually divided into two groups—one for the operator \( L \) and one for \( B \).

Step 3 Choose two integers \( n \) and \( m \) and determine the coefficients \( \alpha_k \) such that

\[
u(x) = A(x) = \sum_{k=1}^{m} \alpha_k B_k(x)
\]

such that

\[
\lambda_j [L(A(x)) - f(x)] = 0 \quad j = 1, \ldots, m
\]

and

\[
\lambda_j [B(A(x)) - g(x)] = 0 \quad j = m+1, \ldots, n
\]

In the case of collocation, the \( \lambda_j = \delta_{x_j} \) is the point evaluation functional at \( x_j \). In other words, the collocation approximation is forced to satisfy the operator equation and the auxiliary conditions at certain points. Other examples of approximation theoretic methods are Taylor's Series, Fourier Series, Eigenfunction expansion, least-squares, Galerkin and finite element methods.

1.1 \( C^1 \) Collocation Method. We describe a specific instance of the collocation method based on piecewise cubic polynomials in \( C^1 \). We have implemented and theoretically and experimentally studied this method for integral equations of Fredholm type, systems of nonlinear ordinary differential equations, linear and nonlinear elliptic and hyperbolic partial differential equations. (See [5], [3], [7], [4], [8].)
a. **Approximate space**

To approximate the solution of the operator equation problem, we choose the space of piecewise cubic polynomial with first continuous derivatives, which we denote by $\mathcal{H}_p^0(\mathbb{R})$. In one dimension the approximation is defined locally as

$$A(x) = \alpha_1 B_1(x) + \beta_1 B_1'(x) + \alpha_{i+1} B_{i+1}(x) + \beta_{i+1} B_{i+1}'(x), \quad x \in [x_i, x_{i+1}]$$

where the basis functions $B_1, B_1'$ are defined in Figure 1.

In two-dimensions, the approximation is defined locally by

$$A(x) = \sum_{i=1}^{16} \alpha_i B_i(x, y) \quad \text{on } E = \text{rectangular element of } \Omega$$

where

- $B_i$ is the $i$ element of Hermite bicubic basis,
- $\alpha_i$ = coefficient of $B_i$ in the approximate solution
  (the index $i$ refers to one element only)

b. **Interpolation operator.** It is worth noticing that the idea of collocation is analogous to interpolation. In this section, we define an interpolation scheme which is used as a basis for the collocation method to be described. Let $\Delta = \{x_i\}_{i=1}^{N+1}$ partition of $\Omega = [a, b]$, we introduce an interpolating operator
such that

\[(Q_Nf)(\sigma_\lambda) = f(\sigma_\lambda), \quad \lambda = 1, \ldots, 2N+2\]

where

\[\sigma_1 = a, \quad \sigma_2 = b, \quad \sigma_{2j+1} = \xi_{2j+1}, \quad 1 = 1, 2\]

are the two Gaussian points with respect to \([x_1, x_{1+1}]\) for \(j = 1, \ldots, N\) and \(\lambda = 2, \ldots, 2N+1\).

A two-dimensional interpolation operator in \([a, b] \times [c, d]\) is defined as

\[Q_\rho \equiv Q_{N,M}\]

where

\[\rho = \Delta x \Delta y, \quad \Delta_y = \{y_1, \ldots, y_{N+1}\} \text{ partition of } [c, d].\]

The new interpolant can be written as

\[Q_Nf = \sum_{i=1}^{N+1} \left( a_i B_i + b_i B_i^\top \right) f(\sigma_i), \quad \lambda = 1, \ldots, 2N+2\]

The coefficients \(a_i\)'s, \(b_i\)'s are determined by the linear algebraic system

\[\sum_{i=1}^{N+1} \left( a_i B_i(\sigma_\lambda) + b_i B_i^\top(\sigma_\lambda) \right) = f(\sigma_\lambda), \quad \lambda = 1, \ldots, 2N+2\]

or in matrix form

\[
\begin{bmatrix}
\alpha & A \\
B & A \\
B & A \\
B & A \\
\end{bmatrix}
\begin{bmatrix}
a_1 \\
b_1 \\
\vdots \\
a \\
\end{bmatrix}
= 
\begin{bmatrix}
f(\sigma_1) \\
f(\sigma_2) \\
\vdots \\
f(\sigma_\lambda) \\
\end{bmatrix}
\]

where the coefficient matrix is the Gramian (denoted by \(G_N\)) of the interpolation operator \(Q_N\).

In the case of uniform mesh \(\Delta\), we have shown [5], after lengthy computations, that \(G_N\) is invertible and \(Q_N\) is bounded in the \(L_\infty\) norm. If we denote by \(\mathfrak{g}_H f\) the
piecewise cubic Hermite interpolant the following inequality, obviously holds,

\[ \| f - Q_Nf \|_\infty \leq (1 + \|Q_N\|) \| f - \alpha_Nf \|_{L_\infty}. \]

Based on this relation and the results on Hermite interpolation [5], we have proved (see [5])

**THEOREM**

If \( f \in W^{s \infty}(\Omega) \) \( s = 1, 2, 3, 4 \) and \( \Omega = [0, 1]^2 \), then

\[ \| f - Q_p f \|_\infty \leq c |p|^s \| f \|_{W^{s \infty}} \]

where \( s = p + q \)

and \( p, q \) are integers.

The following numerical data verify the results of the theorem.

<table>
<thead>
<tr>
<th>( N )</th>
<th>( | x^4 - Q_N x^4 |_\infty )</th>
<th>Rate</th>
<th>( | e^x - Q_N e^x |_\infty )</th>
<th>Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>4.155x10^{-4}</td>
<td></td>
<td>3.106x10^{-5}</td>
<td>3.74</td>
</tr>
<tr>
<td>6</td>
<td>2.678x10^{-5}</td>
<td>3.96</td>
<td>2.325x10^{-6}</td>
<td>3.74</td>
</tr>
<tr>
<td>12</td>
<td>1.674x10^{-6}</td>
<td>3.99</td>
<td>1.646x10^{-7}</td>
<td>3.8</td>
</tr>
<tr>
<td>24</td>
<td>1.047x10^{-7}</td>
<td>4.00</td>
<td>1.096x10^{-8}</td>
<td>3.9</td>
</tr>
<tr>
<td>48</td>
<td>6.541x10^{-9}</td>
<td>4.00</td>
<td>7.070x10^{-10}</td>
<td>3.95</td>
</tr>
</tbody>
</table>

**c. \( C^1 \)-Collocation for Fredholm Integral Equations.** We consider the problem of approximating the solution of the integral solution

\[ \Lambda u \equiv u(P) - \lambda \int k(P;Q)u(Q)dQ = f(P) \]

where \( \Omega = [0, 1]^2 \), \( P = (x, y) \), \( Q = (s, t) \) and \( dQ = dsdt \).

We seek an approximation \( u_p \in H_p \) to \( u \) of the form

\[ u_p(P) = \sum_{i=1}^{N+1} \sum_{j=1}^{M+1} \left( a_{ij}B_iB_j + b_{ij}B_iB_j + c_{ij}B_i B_j + d_{ij}B_iB_j \right) \]
such that
\[
\hat{Q}_p u_p - \lambda \int_{\Omega} k(P; Q) u_p(Q) dQ = \hat{Q}_p f
\]
meaning that \( u_p \) satisfies the integral operator at the interpolating points.

For uniform two-dimensional mesh \( p \) the following results hold (see [5]).

**Theorem**

If A1: \( \lambda \) is not an eigenvalue of the kernel \( k(P; Q) \)

A2: \( f, k \) are in \( W^{s, \infty}(\Omega) \), \( s = 1, 2, 3, 4 \)

then,

(11) for sufficiently small \( |\rho| \) the collocation system is uniquely solvable and

(12) for the error of approximation we have

\[
\| u - u_p \|_{L^\infty} \leq c \| u - Q_p u \|_{L^\infty} \leq c |\rho|^s \| u \|_{W^{s, \infty}}
\]

The following data are in good agreement with the results of the above theorem.

The integral equation

\[
u(s) - \int_a^b k(s, t) u(t) dt = f(s), \quad a \leq s \leq b,
\]

is solved for various kernel functions \( k \), right side functions \( f \) and parameters \( \lambda \).

Case (I). \( k(s, t) = \cos(\pi s t), 0 < s, t \leq 1, \lambda = 1 \). The right side \( f \) is chosen so that \( u(s) = e^s \cos(\pi s) \).

<table>
<thead>
<tr>
<th>( N )</th>
<th>Max. Error</th>
<th>Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>4.55 \times 10^{-2}</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>4.24 \times 10^{-3}</td>
<td>3.4</td>
</tr>
<tr>
<td>12</td>
<td>3.37 \times 10^{-4}</td>
<td>3.7</td>
</tr>
<tr>
<td>24</td>
<td>2.59 \times 10^{-5}</td>
<td>3.7</td>
</tr>
<tr>
<td>48</td>
<td>1.75 \times 10^{-6}</td>
<td>3.9</td>
</tr>
</tbody>
</table>
Case (II). \( k(s,t) = e^{8st}, \ 0 \leq s, t \leq 1, \lambda = 1 \). For the numerical example we pick \( f \) so that

\[ u(s) = e^{as}, \quad \alpha = 1, \beta = 5. \]

<table>
<thead>
<tr>
<th>N</th>
<th>Max. Error</th>
<th>Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>1.33 \times 10^{-3}</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>2.17 \times 10^{-5}</td>
<td>5.9</td>
</tr>
<tr>
<td>12</td>
<td>2.72 \times 10^{-7}</td>
<td>6.3</td>
</tr>
<tr>
<td>24</td>
<td>1.1 \times 10^{-8}</td>
<td>4.6</td>
</tr>
<tr>
<td>48</td>
<td>8.57 \times 10^{-10}</td>
<td>3.7</td>
</tr>
</tbody>
</table>

Case (III). \( k(s,t) = t - s, \ 0 \leq s, t \leq 1, \lambda = 1 \). Choose \( f \) so that

\[ u(s) = s^{\alpha/2}, \quad \alpha = 1, 3, 5, 7. \]

\[ \begin{array}{|c|c|c|}
\hline
\alpha = 1 & \alpha = 3 \\
\hline
N  & Max. Error & Rate & N  & Max. Error & Rate \\
\hline
3  & 9.79 \times 10^{-3} & .5   & 3  & 4.21 \times 10^{-4} & 1.5 \\
6  & 6.92 \times 10^{-3} & .5   & 6  & 1.46 \times 10^{-4} & 1.5 \\
12 | 4.87 \times 10^{-3} & .5   & 12 | 5.10 \times 10^{-5} & 1.5 \\
24 | 3.43 \times 10^{-3} & .5   & 24 | 1.79 \times 10^{-5} & 1.5 \\
48 | 2.42 \times 10^{-3} & .5   & 48 | 6.32 \times 10^{-6} & 1.5 \\
\hline
\end{array} \]
d. $C^1$ - Collocation for 2-point Boundary Value Problem

We consider the problem of approximating the solution of a 2-point boundary value problem

$$D^2x = F(x), \quad t \in [a, b]$$

$$x(a) = c_1, \quad x(b) = c_2$$

We seek an approximate solution

$$x_\Delta(t) = (x_{\Delta,1}(t), \ldots, x_{\Delta,H}(t))$$

with

$$x_{\Delta,i}(t) \in H_\Delta, \quad 1 = 1, \ldots, H$$

which is determined by the collocation conditions

$$D^2x_\Delta(\sigma_\varepsilon) = F(x(\sigma_\varepsilon)), \quad \varepsilon = 2, \ldots, 2N+1$$

$$x_\Delta(\sigma_1) = c_1, \quad x_\Delta(\sigma_{2N+2}) = c_2$$
In [3], we show that if $x \in X \subset C^{(2+n)}([a,b]), E$ is smooth enough, the Green's function of a linear problem associated with $x$ exists and its discontinuous derivatives up to $n$th order are bounded then for the error of approximation we have

$$||D^l(x - x_\Delta)|| \leq C |\Delta|^2 + (n^2-1), \quad l = 0, 1, 2.$$  

**$C^1$-Collocation for Elliptic Partial Differential Equations**

We consider the problem of approximating the solution of a general elliptic partial differential equation

$$L u = A_{xx} u_{xx} + A_{xy} u_{xy} + A_{yy} u_{yy} + D u_x + E u_y + F u = G(u, u_x, u_y)$$

defined on a general domain $\Omega$ and subject to mixed type boundary conditions

$$l u = a u_x + b u_y + c u = g \quad \text{on} \quad \partial \Omega = \text{boundary of} \, \Omega.$$  

This collocation method has the following components.

**(C1) Elements:** A rectangular grid is placed over the domain. Rectangular elements whose center is not inside the domain are discarded.

**(C2) Approximation space:** The Hermite bicubic piecewise polynomials.

**(C3) Approximation to the operator:** The approximate solution satisfies the differential equation exactly at the four Gauss points of a rectangular element. For non-rectangular elements near the boundary, the four Gauss points are projected inside the element as indicated by the diagram.
Approximation to the boundary conditions: The boundary conditions are interpolated at a selected set of boundary points for either Dirichlet or Neumann boundary conditions. If the domain is a rectangle and the problem has Dirichlet conditions \( = 0 \), then the Hermite bicubics are selected so as to automatically satisfy the boundary conditions and no boundary approximation equations are used. This is the same procedure as for the Galerkin and least squares methods. The details on how the boundary collocation points are selected are given below.

Equation Solution: The nonlinear algebraic system is linearized using Newton's method. The linear system is solved by Gauss-elimination taking into account the zeroes in the system (profile method), by a sparse Gauss-elimination with partial pivoting and with an almost block diagonal technique. In the case of linear elliptic equations with homogeneous Dirichlet and Neumann conditions, on a rectangular domain, we have shown in [7] that if \( u \in W^{6,\infty}(\Omega) \), the coefficients of \( L \) are in \( C^{(2+n)}(\Omega) \) and the Green's function of the problem exists with discontinuous partial derivatives up to order \( n \) bounded, then the system of linear collocation equations has a unique solution and for the error of approximation we have

\[
\|u-u_p\|_{L^\infty} \leq C \left( (\Delta_y)^4 + (\Delta_x)^{2+\min(n,2)} \right)
\]

with \( C \) constant independent of \( \Delta_x, \Delta_y \).
In the case of linear elliptic equations with homogeneous Dirichlet and Neumann conditions on a rectangular domain, we have shown in [7]

**Theorem.** If $u \in W^{2n_0}(\Omega)$ and the coefficients of $L$ are in $C^{2n_0}(\Omega)$ and the Green's function of the problem exists with discontinuous partial derivatives up to order $n$ bounded, then the system of linear differential equations has a unique solution and for the error of approximation we have

$$
\|u-u^h\|_{L^\infty(\Omega)} \leq C \left( \|L\|_{W^{2n_0}(\Omega)} + \|L\|_{W^{2n_0}(\Omega)} \right)^3
$$

where $C$ is independent of $\Delta x$, $\Delta y$.

Next, we derive similar estimates for the general linear elliptic partial differential equations on a general two-dimensional domain. We denote by $\Omega'$ the discretized region $\Omega$ consisting of rectangular elements with sides $h$, $k$ (for simplicity we assume $h=k$). Let $u_0$ be the bicubic Hermite interpolant of the solution $u$ over the region $\Omega'$. It is easy to show

**Theorem.** If $u \in W^{2n_0}(\Omega')$

then

$$
\|u-u_0\|_{W^{2n_0}(\Omega')} \leq C \|u\|_{W^{2n_0}(\Omega')} \leq C \|u\|_{W^{2n_0}(\Omega')} \leq C \|u\|_{W^{2n_0}(\Omega')} \leq C \|u\|_{W^{2n_0}(\Omega')}
$$

where $S = p+q$ and $p, q$ are integers.
Since, can be expressed as
\[ u = \Omega \, u + h^4 \, R \]
we have
\[ Lu = L \Omega \, u + h^2 \, R_1 \quad \text{and} \quad 2u = 2 \Omega \, u + h^3 \, R_2 \]
where \( R_1, R_2 \) are bounded functions.

Throughout, we will denote by \( u_\Delta \) the piecewise
bicubic Hermite polynomial defined on \( \Omega ' \)
such that in each rectangular element \( E \)
\[ Lu_\Delta = f \quad | \quad (x,y) = p_i \quad i = 1,2,3,4 \] 
\[ 2u_\Delta = q \quad | \quad (x,y) = q_j \quad j = 1,7 \eta (E) \]
with \( p_i \) four Gaussian points with respect to \( E \)
and \( q_j \) \( \eta (E) \) boundary collocation points associated
with \( E \).

The relations
\[ Lu_\Delta (p_i) - L \Omega \, u (p_i) = h^2 \, R_1 \]
and
\[ 2u_\Delta (q_j) - 2 \Omega \, u (q_j) = h^3 \, R_2 \]
held at the collocation points.

Thus, we have
\[ Lu_\Delta (p_i) - L \Omega \, u (p_i) = h^2 \, R_1 \]
and
\[ 2u_\Delta (q_j) - 2 \Omega \, u (q_j) = h^3 \, R_2 \]

We express
\[ u_\Delta = \sum_{i=1}^{N} \beta_i \, B_i \quad \Omega \, u = \sum_{i=1}^{N} b_i \, B_i \]
where \( B_i \) are the basis Bernstein functions of the piecewise Hermite bicubic polynomials.

The above systems of collocation equations can be written in matrix form as

\[
K \beta - K b = h^2 \gamma
\]

where \( K = \left( L B_i(p_j) \right) \) or \( 2B_i(q_{ij}) \) and \( \gamma = \left( R_1(p_j) \right) \) or \( R_2(q_{ij}) \).

It follows easily that

\[
\| \beta - b \| \leq h^2 \| K^{-1} \| \| x \|
\]

From the nature of the Hermite bicubic local basis functions, we observe that the elements of \( K \)

\[
K_{ij} = P_{ij} h^2 + q_{ij} h^{-1} + s_{ij} + t_{ij} h
\]

or

\[
q_{ij} h^{-1} + s_{ij} + t_{ij} h
\]

Therefore, the coefficient matrix \( K \) can be expressed as

\[
K = \frac{1}{h^2} K'
\]

\[\text{THEOREM}\]

If \( K \) is invertible, then

\[ \| K^{-1} \| \leq \rho \]

where \( \rho = \max \| f \| \) and \( f(x) = x \).
By considering the form of \( K' \), we observe

\[
K' \to K'_{\varepsilon} \quad \text{as} \quad \varepsilon \to 0.
\]

Thus, for \( \varepsilon \) sufficiently small, we have

\[
\inf_{\| K \| = 1} \frac{\| K \|}{\| K \| - \varepsilon} \geq \frac{\beta^2}{\varepsilon} > 0.
\]

By definition of matrix norms

\[
\| (K')^{-1} \| = \frac{1}{\inf_{\| K \| = 1} \| K \|} \leq \frac{1}{\beta^2}
\]

and

\[
\| K^{-1} \| = \frac{\beta}{\inf_{\| K \| = 1} \| K \|} \leq \frac{\beta}{\beta^2} = \frac{1}{\varepsilon}.
\]

Q.E.D.

The following theorem is almost immediate:

Combining the derived bounds we observe

**Theorem** If \( u \in W^{1,\infty}(\mathbb{R}) \),

the system of cell-center equations has a unique solution, and the coefficients of \( L \) are bounded in \( \mathbb{R}^2 \),

then

\[
\| \frac{\partial}{\partial x} \phi - L \phi \| \leq C h^4 \| \frac{\partial f}{\partial x} \| / \beta^2
\]

From the results of interior finite-element interpolating and

boundary layer above, the convergence of the solution and the

coefficients we can show

and

\[
\| u - \phi \|_{L^\infty(\mathbb{R})} \leq h^4 C
\]

where \( C \) is independent of \( h \).
Next, we present a number of numerical data which agree with the theoretical expectations. We consider the equation \( Lu = u_{xx} + u_{yy} + (2 - \cos x \cos y) = -u_{xy} \) defined on unit square and subject to Dirichlet boundary conditions \( u = \text{True} (x, y) \). The true solution is \( \sin x \sin y \). We denote \( \text{maxdif} = \| u^{(n)} - u^{(n-1)} \|_\infty \), \( \text{maxerr} = \| \text{True} - u^{(n)} \|_\infty \). For initial estimate \( u^{(0)} = 0 \), the results are:

RESULTS FOR GRID WITH 4 ELEMENTS

<table>
<thead>
<tr>
<th>ITERATION</th>
<th>MAX.DIF.</th>
<th>MAX. ERR.</th>
<th>TIME</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>9.13E-01</td>
<td>1.27E-02</td>
<td>.38</td>
</tr>
<tr>
<td>2</td>
<td>4.53E-01</td>
<td>1.15E-04</td>
<td>.39</td>
</tr>
<tr>
<td>3</td>
<td>7.45E-04</td>
<td>1.12E-04</td>
<td>.40</td>
</tr>
<tr>
<td>4</td>
<td>2.69E-10</td>
<td>1.12E-04</td>
<td>.40</td>
</tr>
</tbody>
</table>

RESULTS FOR GRID WITH 9 ELEMENTS

<table>
<thead>
<tr>
<th>ITERATION</th>
<th>MAX.DIF.</th>
<th>MAX. ERR.</th>
<th>TIME</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>9.13E-01</td>
<td>1.27E-02</td>
<td>.94</td>
</tr>
<tr>
<td>2</td>
<td>5.04E-01</td>
<td>2.94E-05</td>
<td>.97</td>
</tr>
<tr>
<td>3</td>
<td>6.11E-04</td>
<td>2.50E-05</td>
<td>.96</td>
</tr>
<tr>
<td>4</td>
<td>3.33E-10</td>
<td>2.50E-05</td>
<td>.96</td>
</tr>
</tbody>
</table>

RESULTS FOR GRID WITH 16 ELEMENTS

<table>
<thead>
<tr>
<th>ITERATION</th>
<th>MAX.DIF.</th>
<th>MAX. ERR.</th>
<th>TIME</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>9.13E-01</td>
<td>1.27E-02</td>
<td>1.92</td>
</tr>
<tr>
<td>2</td>
<td>5.42E-01</td>
<td>1.11E-05</td>
<td>1.88</td>
</tr>
<tr>
<td>3</td>
<td>5.45E-04</td>
<td>8.70E-06</td>
<td>1.89</td>
</tr>
<tr>
<td>4</td>
<td>2.81E-10</td>
<td>8.70E-06</td>
<td>1.89</td>
</tr>
</tbody>
</table>

RESULTS FOR GRID WITH 25 ELEMENTS

<table>
<thead>
<tr>
<th>ITERATION</th>
<th>MAX.DIF.</th>
<th>MAX. ERR.</th>
<th>TIME</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>9.13E-01</td>
<td>1.27E-02</td>
<td>3.46</td>
</tr>
<tr>
<td>2</td>
<td>5.71E-01</td>
<td>6.55E-06</td>
<td>3.55</td>
</tr>
<tr>
<td>3</td>
<td>5.12E-04</td>
<td>4.08E-06</td>
<td>3.55</td>
</tr>
<tr>
<td>4</td>
<td>2.50E-10</td>
<td>4.08E-06</td>
<td>3.55</td>
</tr>
</tbody>
</table>

RESULTS FOR GRID WITH 26 ELEMENTS

<table>
<thead>
<tr>
<th>ITERATION</th>
<th>MAX.DIF.</th>
<th>MAX. ERR.</th>
<th>TIME</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>9.13E-01</td>
<td>1.27E-02</td>
<td>5.74</td>
</tr>
<tr>
<td>2</td>
<td>5.95E-01</td>
<td>4.65E-06</td>
<td>5.92</td>
</tr>
<tr>
<td>3</td>
<td>4.92E-04</td>
<td>1.80E-05</td>
<td>5.97</td>
</tr>
<tr>
<td>4</td>
<td>2.89E-10</td>
<td>1.80E-05</td>
<td>5.95</td>
</tr>
</tbody>
</table>

RESULTS FOR GRID WITH 49 ELEMENTS

<table>
<thead>
<tr>
<th>ITERATION</th>
<th>MAX.DIF.</th>
<th>MAX. ERR.</th>
<th>TIME</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>9.13E-01</td>
<td>1.27E-02</td>
<td>9.16</td>
</tr>
<tr>
<td>2</td>
<td>6.15E-01</td>
<td>4.67E-06</td>
<td>9.34</td>
</tr>
<tr>
<td>3</td>
<td>4.76E-04</td>
<td>7.76E-07</td>
<td>9.33</td>
</tr>
<tr>
<td>4</td>
<td>2.15E-10</td>
<td>7.76E-07</td>
<td>9.33</td>
</tr>
</tbody>
</table>

39.39
II. We consider the equation

\[ Lu = u_{xx} + u_{yy} = \exp(u) + F \text{ on unit square with} \]

\[ F = -\exp(true) + 2(x^2 + y^2)/(xy + 1)^3 \]

and subject to boundary conditions \( u = true. \)

In this example the true has been chosen \( 1/(xy+1). \) The results with initial estimate \( u(0) = 0 \) are:

RESULTS FOR GRID WITH 4 ELEMENTS

<table>
<thead>
<tr>
<th>ITERATION</th>
<th>MAX. DIF.</th>
<th>MAX. ERR.</th>
<th>TIME</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.59E-00</td>
<td>3.21E-02</td>
<td>.36</td>
</tr>
<tr>
<td>2</td>
<td>1.26E-00</td>
<td>5.09E-04</td>
<td>.33</td>
</tr>
<tr>
<td>3</td>
<td>4.68E-04</td>
<td>5.17E-04</td>
<td>.37</td>
</tr>
<tr>
<td>4</td>
<td>6.83E-10</td>
<td>5.17E-04</td>
<td>.38</td>
</tr>
</tbody>
</table>

RESULTS FOR GRID WITH 9 ELEMENTS

<table>
<thead>
<tr>
<th>ITERATION</th>
<th>MAX. DIF.</th>
<th>MAX. ERR.</th>
<th>TIME</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.20E-00</td>
<td>3.16E-02</td>
<td>.90</td>
</tr>
<tr>
<td>2</td>
<td>1.44E-00</td>
<td>1.43E-04</td>
<td>.92</td>
</tr>
<tr>
<td>3</td>
<td>4.94E-04</td>
<td>1.54E-04</td>
<td>.92</td>
</tr>
<tr>
<td>4</td>
<td>7.39E-10</td>
<td>1.54E-04</td>
<td>.92</td>
</tr>
</tbody>
</table>

RESULTS FOR GRID WITH 16 ELEMENTS

<table>
<thead>
<tr>
<th>ITERATION</th>
<th>MAX. DIF.</th>
<th>MAX. ERR.</th>
<th>TIME</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.35E-00</td>
<td>3.16E-02</td>
<td>1.85</td>
</tr>
<tr>
<td>2</td>
<td>1.57E-00</td>
<td>6.37E-05</td>
<td>1.85</td>
</tr>
<tr>
<td>3</td>
<td>5.00E-04</td>
<td>6.77E-05</td>
<td>1.85</td>
</tr>
<tr>
<td>4</td>
<td>7.28E-10</td>
<td>6.77E-05</td>
<td>1.85</td>
</tr>
</tbody>
</table>

RESULTS FOR GRID WITH 25 ELEMENTS

<table>
<thead>
<tr>
<th>ITERATION</th>
<th>MAX. DIF.</th>
<th>MAX. ERR.</th>
<th>TIME</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.46E-00</td>
<td>3.18E-02</td>
<td>3.40</td>
</tr>
<tr>
<td>2</td>
<td>1.69E-00</td>
<td>4.50E-05</td>
<td>3.44</td>
</tr>
<tr>
<td>3</td>
<td>5.01E-04</td>
<td>3.48E-05</td>
<td>3.42</td>
</tr>
<tr>
<td>4</td>
<td>7.24E-10</td>
<td>3.48E-05</td>
<td>3.43</td>
</tr>
</tbody>
</table>

RESULTS FOR GRID WITH 36 ELEMENTS

<table>
<thead>
<tr>
<th>ITERATION</th>
<th>MAX. DIF.</th>
<th>MAX. ERR.</th>
<th>TIME</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.54E-00</td>
<td>3.18E-02</td>
<td>5.67</td>
</tr>
<tr>
<td>2</td>
<td>1.76E-00</td>
<td>4.65E-05</td>
<td>5.77</td>
</tr>
<tr>
<td>3</td>
<td>5.02E-04</td>
<td>1.64E-05</td>
<td>5.74</td>
</tr>
<tr>
<td>4</td>
<td>7.19E-10</td>
<td>1.64E-05</td>
<td>5.73</td>
</tr>
</tbody>
</table>

RESULTS FOR GRID WITH 49 ELEMENTS

<table>
<thead>
<tr>
<th>ITERATION</th>
<th>MAX. DIF.</th>
<th>MAX. ERR.</th>
<th>TIME</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.62E-00</td>
<td>3.16E-02</td>
<td>8.55</td>
</tr>
<tr>
<td>2</td>
<td>1.83E-00</td>
<td>4.56E-05</td>
<td>9.04</td>
</tr>
<tr>
<td>3</td>
<td>5.02E-04</td>
<td>6.60E-06</td>
<td>9.06</td>
</tr>
<tr>
<td>4</td>
<td>7.25E-10</td>
<td>6.60E-06</td>
<td>9.04</td>
</tr>
</tbody>
</table>
III. The equation
\[ Lu = u_{xx} + u_{yy} = u^2 + F \]
is considered on unit square with Dirichlet boundary conditions with
\[ \text{True}(x,y) = x(4-x) + \exp(x^2+y)/20 \]
and \[ F = -(\text{True})^2 + L(\text{True}). \]
The initial estimation \[ u(0) = 0 \] results in:

<table>
<thead>
<tr>
<th>RESULTS FOR GRID WITH 4 ELEMENTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>ITERATION N</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>4</td>
</tr>
<tr>
<td>5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>RESULTS FOR GRID WITH 9 ELEMENTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>ITERATION N</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>4</td>
</tr>
<tr>
<td>5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>RESULTS FOR GRID WITH 16 ELEMENTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>ITERATION N</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>4</td>
</tr>
<tr>
<td>5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>RESULTS FOR GRID WITH 25 ELEMENTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>ITERATION N</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>4</td>
</tr>
<tr>
<td>5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>RESULTS FOR GRID WITH 36 ELEMENTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>ITERATION N</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>4</td>
</tr>
<tr>
<td>5</td>
</tr>
</tbody>
</table>

29.96
Data for solving $u_{xx} + u_{yy} = f$, $u = g$ on $\Omega$ (Figure 1) with $u$ taken as

$$y[(x-2)^2 + y^2 - 1]e^{0.0625x(x-4)(y-2)}/[(3+(x-2)^2)(3+y^2)]$$

**METHOD:** COLLOCATION based on Hermite bicubics ($C^1$)

<table>
<thead>
<tr>
<th>Number of Equations</th>
<th>Matrix Formation</th>
<th>Profile Gauss Elim. Sol.</th>
<th>Maximum Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>56</td>
<td>.146</td>
<td>.507</td>
<td>2.367E-03</td>
</tr>
<tr>
<td>108</td>
<td>.311</td>
<td>1.478</td>
<td>9.307E-04</td>
</tr>
<tr>
<td>164</td>
<td>.496</td>
<td>3.049</td>
<td>2.305E-04</td>
</tr>
<tr>
<td>240</td>
<td>.746</td>
<td>5.646</td>
<td>1.141E-04</td>
</tr>
</tbody>
</table>

**Figure 1** The geometry and boundary conditions for problem in Table 9.

![Geometry and boundary conditions](image)
Data for solving $u_{xx} + u_{yy} = [100 + \cos(3\pi x) + \sin(2\pi y)]u = f$ on unit square with $u$ taken as $(5.4 - \cos(4\pi x))\sin(\pi x)(y^2 - y)(5.4 - \cos(4\pi y))\#(1/(1+4)^{-1/2})$

\[ \phi = 4(x-.5)^2 + 4(y-.5)^2 \]

**METHOD:** COLLOCATION based on Hermite bicubics

<table>
<thead>
<tr>
<th>N</th>
<th>Equation</th>
<th>Half Bandwidth</th>
<th>Matrix Formation</th>
<th>Profile Gauss Solution</th>
<th>Maximum Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>16</td>
<td>10</td>
<td>.002</td>
<td>.139</td>
<td>8.48E-01</td>
</tr>
<tr>
<td>3</td>
<td>36</td>
<td>12</td>
<td>.189</td>
<td>.19</td>
<td>2.10E-01</td>
</tr>
<tr>
<td>4</td>
<td>64</td>
<td>14</td>
<td>.335</td>
<td>.463</td>
<td>1.31E-01</td>
</tr>
<tr>
<td>5</td>
<td>100</td>
<td>16</td>
<td>.518</td>
<td>.921</td>
<td>3.31E-02</td>
</tr>
<tr>
<td>6</td>
<td>144</td>
<td>18</td>
<td>.776</td>
<td>1.710</td>
<td>2.68E-02</td>
</tr>
<tr>
<td>8</td>
<td>256</td>
<td>22</td>
<td>1.367</td>
<td>4.405</td>
<td>1.25E-02</td>
</tr>
<tr>
<td>9</td>
<td>324</td>
<td>24</td>
<td>1.714</td>
<td>6.663</td>
<td>6.88E-03</td>
</tr>
</tbody>
</table>

**Table 3** Data Indicating Collocation equation solution times for BND.SOL, NSPIV, SLVBLK

<table>
<thead>
<tr>
<th>N</th>
<th>Matrix Formation</th>
<th>Equation Solution</th>
<th>Matrix Formation</th>
<th>Equation Solution</th>
<th>Matrix Formation</th>
<th>Equation Solution</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>.033</td>
<td>.036</td>
<td>.036</td>
<td>.054</td>
<td>.036</td>
<td>.061</td>
</tr>
<tr>
<td>3</td>
<td>.089</td>
<td>.151</td>
<td>.081</td>
<td>.216</td>
<td>.086</td>
<td>.199</td>
</tr>
<tr>
<td>4</td>
<td>.178</td>
<td>.419</td>
<td>.143</td>
<td>.584</td>
<td>.159</td>
<td>.477</td>
</tr>
<tr>
<td>5</td>
<td>.308</td>
<td>.924</td>
<td>.223</td>
<td>1.266</td>
<td>.255</td>
<td>.962</td>
</tr>
<tr>
<td>6</td>
<td>.485</td>
<td>1.775</td>
<td>.322</td>
<td>2.391</td>
<td>.368</td>
<td>1.739</td>
</tr>
<tr>
<td>7</td>
<td>.724</td>
<td>3.042</td>
<td>.443</td>
<td>4.055</td>
<td>.5</td>
<td>2.836</td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td>.645</td>
<td>4.451</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 4: Data for solving $u_{xx} + u_{yy} = f$, $\frac{\partial u}{\partial n} = 0$ on unit square with $u$ taken as $\cos \pi x \cos \pi y$.

Max. Residual $\equiv \max |u_\Delta(x,y) - f(x,y)|$ at nodes.

$u_\Delta$ = collocation approximation

<table>
<thead>
<tr>
<th>N</th>
<th>Maximum Error</th>
<th>Rate</th>
<th>Maximum Residual</th>
<th>Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>2.308E-02</td>
<td></td>
<td>3.508E-00</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>5.383E-03</td>
<td>3.59</td>
<td>1.694E-00</td>
<td>1.8</td>
</tr>
<tr>
<td>4</td>
<td>1.822E-03</td>
<td>3.77</td>
<td>9.800E-01</td>
<td>1.9</td>
</tr>
<tr>
<td>5</td>
<td>7.696E-04</td>
<td>3.86</td>
<td>6.332E-01</td>
<td>1.94</td>
</tr>
<tr>
<td>6</td>
<td>3.773E-04</td>
<td>3.91</td>
<td>4.441E-01</td>
<td>1.96</td>
</tr>
<tr>
<td>7</td>
<td>2.057E-04</td>
<td>3.94</td>
<td>3.276E-01</td>
<td>1.97</td>
</tr>
<tr>
<td>8</td>
<td>1.213E-04</td>
<td>3.96</td>
<td>2.515E-01</td>
<td>1.98</td>
</tr>
</tbody>
</table>
We consider the problem of approximating the solution of the nonlinear hyperbolic equation

\[ p(x,t,u)u_{tt} - q(x,t,u)u_{xx} = f(x,t,u,u_x), \quad (x,t) \in (0,1) \times (0,T). \]

subject to the initial conditions

\[ u(x,0) = \alpha_1(x), \quad u_t(x,0) = \alpha_2(x), \quad 0 < t < 1, \]

and the boundary conditions

\[ u(0,t) = 0, \quad u(1,t) = 0, \quad 0 < t \leq T. \]

Assume that the coefficients satisfy

\[ 0 < c_1 \leq p(x,t,u) \leq C_1, \quad c_2 \leq q(x,t,u) \leq C_2, \]

for \( 0 \leq x \leq 1 \), \( 0 \leq t \leq T \) and \( -\infty < u < +\infty \). Also, we assume that \( p, q, f \) are continuously differentiable functions of their arguments and uniformly bounded.

We seek an approximation

\[ u^{N+1}_\Delta(x,t) = \sum_{i=1}^{N+1} \{ \alpha_i(t) \beta_i^1(x) + \beta_i^1(t) \beta_i^1(x) \}, \]

such that

\[ (p(u^\Delta) \frac{\partial^2}{\partial t^2} u^\Delta - q(u^\Delta) \frac{\partial^2 u}{\partial x^2} - f(u^\Delta, \frac{\partial u}{\partial x})) \sigma^\Delta, t) = 0 \]

for \( 0 < t \leq T \), \( i = 2, \ldots, 2N+1 \),

and

\[ u^\Delta(\sigma^\Delta, 0) = \alpha_1(\sigma^\Delta), \quad \sigma^\Delta(\sigma^\Delta, 1) = \alpha_2(\sigma^\Delta), \]

\[ u^\Delta(0,t) = u^\Delta(1,t) = 0. \]

In [4], we show the following theorem:

**Theorem**

if

(1) \( p, q, f \) have bounded third derivatives

(2) \( u \in L_\infty(0,T;W^{6,2}), \quad u_t, u_{tt} \in L^2(0,T;W^{6,2}) \)
(iii) \( u_\Delta(x,0), \frac{3}{\partial t} u_\Delta(x,0) \) are Hermite interpolants of
\( u(x,0), \frac{3u}{\partial t}(x,0) \), respectively,
then for the error of approximation we have
\[
\|u-u_\Delta\|_{L^\infty(0,T;L^\infty)} \leq C h^4
\]
where \( C \) is independent of \( h \).

Next, we present an example of a linear hyperbolic equation solved by applying
the \( c^1 \)-collocation method in both directions.

**Table 5** Data for solving the initial boundary value problem
\[
u_{xx} = 4u_{tt} + u_t \text{ with } u(x,0) = \sin \frac{\pi}{2} x,

\text{and } u_t(x,0) = -\frac{1}{4} \sin \frac{\pi}{2} x \text{ in } [0,2] \text{ and } u(0,t) = u(2,t) = 0 \text{ for } 0 \leq t \leq 2.

The exact solution is
\[
e^{-\frac{1}{8} t} \sin \frac{\pi}{2} x \left( \cos \frac{\sqrt{v} t}{8} - \frac{1}{v} \sin \frac{\sqrt{v} t}{8} \right)
\]
where \( v = \sqrt{4\pi^2 - 1} \).

<table>
<thead>
<tr>
<th>N</th>
<th>Maximum Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.051E-01</td>
</tr>
<tr>
<td>2</td>
<td>6.799E-03</td>
</tr>
<tr>
<td>3</td>
<td>1.762E-03</td>
</tr>
<tr>
<td>4</td>
<td>4.700E-04</td>
</tr>
<tr>
<td>5</td>
<td>2.314E-05</td>
</tr>
</tbody>
</table>
II. METHODS FOR CURVED BOUNDARIES

II.1 COLLOCATION

The most sensitive aspect of collocation is the placement of the boundary collocation points for non-rectangular domains. First, one must take care that these points are reasonably separated from interior collocation points. This is not difficult to do even in an automatic way, but the penalty for overlooking this point is an ill-conditioned computation with large errors.

One first overlays the region with a rectangular grid and discards the elements which intersect the domain slightly or not at all. Let $S_b$ be the number of boundary sides of the resulting rectangular partition. Then the number of boundary collocation points required is $2S_b + 4$. We use two basic schemes for distributing the boundary collocation points as illustrated by the diagrams below for a simple rectangle:

![Diagram](image)

**Figure 1.** Two schemes for distributing boundary collocation points. The $x$'s are the systematic collocation points and the $0$'s are the four extra ones.
A theoretical analysis indicates that the 2-point scheme is superior for rectangular regions if the two points used are the Gauss points for each boundary segment. We compared the Gauss points with equally spaced points and found the equally spaced points give slightly better accuracy for rectangular domains.

We made numerous numerical experiments which confirmed that the 2-point scheme is superior for rectangular regions.

The extension of these two schemes to curved domains is illustrated in Figure 2.

![Figure 2](image)

**Figure 2.** The two schemes for a simple curved domain. The lines show how the collocation points are placed on the edge of the rectangular partition and then mapped onto the portions of the boundary intersecting each rectangular element.

The theoretical advantage of the 2-point scheme no longer holds for curved boundaries and our experiments confirm that it has no advantage over the midpoint scheme in this case. In fact it is, on the average, slightly less accurate. Furthermore, the midpoint scheme automatically gives collocation...
of the boundary conditions at any extremities of the domain (for example, for a piecewise rectangular boundary such as in Problems 16 and 17, see Figure 5). It is often essential that collocation of the boundary conditions be made at all exterior corners of the domain. The midpoint scheme naturally provides this.

Our procedure is to use the 2-point scheme for boundaries which are straight (or nearly so) and parallel to a coordinate axis and to use the midpoint scheme otherwise. The two schemes may be used together for a domain such as shown above and we do this as shown in Figure 3.

![Figure 3](image)

Figure 3. The combination of the two schemes for a partially rectangular region. The mapping from the point on the rectangular edges to the curved boundary is indicated.

There seems to be no particularly advantageous method to distribute the 4 extra collocation points beyond putting them in elements with exterior corners and spreading them somewhat evenly around the boundary. We always map the midpoint type collocation points to segments of the curved boundary which are interior to the rectangular partition. The points are placed uniformly on each such segment. At times this may leave rather large segments of a curved boundary "unused", but we have not found a reliable method to place collocation points on the intermediate segments. We do place collocation
points outside the rectangular partition for the 2-point scheme. An example is shown in Figure 4 which illustrates these procedures.

Figure 4. Example which illustrates boundary collocation points for the 2-point scheme which are outside the rectangular partition and collocation for the midpoint scheme are inside. Collocation is not done on two large boundary segments.
11.2 LEAST-SQUARES COLLOCATION

In the computational analysis of the collocation scheme described above, it was observed that the convergence of the method depends on the location of the boundary collocation points. We have described two schemes for the distribution of boundary points that lead to an experimentally stable method. To avoid the dependence on the boundary collocation points, we have developed and implemented two least-squares schemes for the approximation of the boundary conditions. First, we choose the collocation approximation so that

\[ \sum_{p_i \in \partial E} (B_{p_i} - g(p_i))^2 \]

becomes minimum locally in each boundary element. In the second scheme, the discrete least squares norm of boundary residual over the entire boundary is minimized. Theoretical and experimental results will be published elsewhere.

III. SUMMARY OF PROPERTIES OF THE COLLOCATION METHOD

a) Simplicity. In the linear case, the collocation equations are easily generated element by element. No numerical integration is required. There are 16 nonzero coefficients per row and the system has an almost block diagonal structure. In the mildly nonlinear case, the Jacobian matrix has the same structure.

b) Flexibility. Once the analogy with interpolation is seen, then one can quickly adopt the method to a wide range of problems.

c) Accuracy. It depends on the choice of basis functions and location of collocation points. This is another analogy with the interpolation.
To avoid the dependence of the geometry of the elements in the region $\Omega$ and make the method applicable for any choice of approximate space, we develop the following least-squares collocation method, coupled with a least-squares procedure for the determination solution of collocation equations.

In each skeletal element $E_j$ of a partition $\Omega'$ of $\Omega$, we pose

$$ L u_\Delta (p_i) = f (p_i) \quad i = 1, \ldots, n(E) $$

where $p_i$ are points of $E$.

For the boundary elements, these collocation approximations is forced to satisfy the boundary conditions

$$ B u_\Delta (q_j) = g (q_j) \quad j = 1, \ldots, n(E) $$

at certain boundary points $q_j$ inside the element $E$.

To ensure existence and unique convergence, we assume that the number of collocation points is larger than the degrees of freedom. The resulting system of collocation equations is over-determined and a least-squares procedure is used for the determination of the solution.

From each element we generate a saddle system of collocation equations

$$ K e \hat{e} = \overline{F e} $$

or

$$ K^T e \hat{e} = K^T \overline{F e} $$

Notice $K e \hat{e} \equiv 0$ is a square symmetric matrix with rank equal to the number of degrees of freedom associated with this element.
d) **Efficiency.** We have systematically evaluated (see [1], [2], [6]) six methods for solving two-dimensional linear elliptic partial differential equations on general domains. The methods are: standard finite difference, collocation, Galerkin and least-squares using Hermite cubic piecewise polynomials, Galerkin using bicubic splines and Galerkin using linear triangular elements. Our test set of problems ranges from simple to moderately complex. The principal conclusion is that collocation is the most efficient method for general use. Standard finite differences is sometimes more efficient for very crude accuracy (where efficiency is not important anyway), but it is also sometimes enormously less efficient even for very modest accuracy. The accuracy of the Galerkin and least-squares $C^1$ methods is sometimes better than collocation, but the extra cost always negates this advantage for our problems. The Galerkin $C^2$ for rectangular domains turns to be competitive to collocation for self-adjoint problems with simple functions in the differential operator and high accuracy requirements.
References


APPENDIX

The piecewise bicubic Hermite element. Given the one-dimensional mesh \( \Delta_x = \{a = x_0 < x_1 < \ldots < x_N = b\} \), let \( H(\Delta_x) \) be the space of piecewise cubic polynomials with respect to \( \Delta_x \) which are continuously differentiable in \([a, b]\). We will denote by \( H_0(\Delta_x) \) the set of functions \( p \in H(\Delta_x) \) which satisfy the boundary conditions \( p(a) = p(b) = 0 \). Given the mesh \( \Delta_y = \{c = y_0 < y_1 < \ldots < y_M = d\} \) the space \( H(\Delta_y) \) is defined analogously.

In order to introduce a representation of a bicubic rectangular Hermite element we consider 8 one-dimensional functions.

\[
\begin{align*}
&x = \frac{s}{a} \quad \text{and} \quad 0 \leq s \leq 1 \\
&t = \frac{y}{b} \quad \text{and} \quad 0 \leq t \leq 1
\end{align*}
\]

\[
\begin{align*}
B_{x1} &= 1 - 3s^2 + 2s^3 \\
B_{x2} &= s^2(3-2s) \\
B_{x3} &= as(s-1)^2 \\
B_{x4} &= as^2(s-1)
\end{align*}
\]

\[
\begin{align*}
B_{y1} &= 1 - 3t^2 + 2t^3 \\
B_{y2} &= t^2(3-2t) \\
B_{y3} &= bt(t-1)^2 \\
B_{y4} &= bt^2(t-1)
\end{align*}
\]

Then the bicubic rectangular element is defined by

\[
U(x, y) = B_{x1} B_{y1} u_1 + B_{x2} B_{y1} u_2 + B_{x2} B_{y2} u_3 + B_{x1} B_{y2} u_4
\]

\[
+ B_{x3} B_{y1} \sigma_{x1} + B_{x4} B_{y1} \sigma_{x2} + B_{x4} B_{y2} \sigma_{x3} + B_{x3} B_{y2} \sigma_{x4}
\]

\[
+ B_{x1} B_{y3} \sigma_{y1} + B_{x2} B_{y3} \sigma_{y2} + B_{x2} B_{y4} \sigma_{y3} + B_{x1} B_{y4} \sigma_{y4}
\]

\[
+ B_{x3} B_{y3} \tau_{xy1} + B_{x4} B_{y3} \tau_{xy2} + B_{x4} B_{y4} \tau_{xy3} + B_{x3} B_{y4} \tau_{xy4}
\]
where \( u_I \) = value at the point \( I \)

\[ \sigma_{xI}, \sigma_{yI} = x \text{ and } y \text{ derivatives at the point } I \]

\[ \tau_{xyI} = xy \text{ (cross) derivative at the point } I. \]

We denote by \( B_I(x,y), I = 1, 16 \) the 16 basis functions in the above representation; i.e.

\[ B_1 \equiv B_{x1} B_{y1}, B_2 \equiv B_{x3} B_{y1}, \ldots, B_{13} \equiv B_{x1} B_{y2}, \ldots, B_{16} \equiv B_{x3} B_{y4} \]

2. The piecewise bicubic Spline Element. Let \( S_0(\Delta_x) \) be the space of functions \( s(x) \) which are cubic polynomials in each subinterval \([x_i, x_{i+1}]\), twice continuously differentiable in \([a,b]\), and satisfy the boundary conditions \( s(a) = s(b) = 0 \). We choose the B-spline basis for the piecewise polynomial space \( S_0(\Delta_x) \) and denote them by \( \{\phi_i(x)\}_{i=0}^N \). The graph of \( \phi_i(x) \) is

![B-spline graph]

The space \( S_0(\Delta_y) \) and the corresponding basis \( \{\phi_j(y)\}_{j=0}^N \) are defined analogously.

Then the bicubic spline is defined in each subrectangle \([x_i, x_{i+1}] \times [y_j, y_{j+1}]\) by

\[ U(x,y) = \sum_{k=1-3}^{1} \sum_{l=j-3}^{1} \sigma_{k,l} \phi_k(x) \phi_l(y) \]

We denote \( B_m(x,y) \equiv \phi_k(x) \phi_l(y) \) for \( m = k + (n+1)l+1, 0 \leq k, \ell \leq n \), \( p = \Delta_x \times \Delta_y \) and \( S_0(p) \) the space of bicubic splines represented by

\[ s(x,y) = \sum_{m=1}^{(N+1)^2} \beta_m B_m(x,y) \]