High accuracy finite difference approximation to solutions of elliptic partial differential equations

(putential solution to linear differential equations/variable coefficients/nine-point stencil)

ROBERT E. LYNCH AND JOHN R. RICE
Division of Mathematical Sciences, Purdue University, West Lafayette, Indiana 47907

Communicated by Garrett Birkhoff, February 14, 1978

ABSTRACT A flexible finite difference method is described that gives approximate solutions of linear elliptic partial differential equations, $Lu = G$, subject to general linear boundary conditions. The method gives high-order accuracy. The values of the unknown approximation function $U$ are determined at mesh points by solving a system of finite difference equations $L_h U = I_h G$. $L_h U$ is a linear combination of values of $U$ at points of a standard stencil (9-point for two-dimensional problems, 27-point for three-dimensional) and $I_h G$ is a linear combination of values of the given function $G$ at mesh points as well as at other points. A local calculation is carried out to determine the coefficients of the operators $L_h$ and $I_h$ so that the approximation is exact on a specific linear space of functions. Having the coefficients of each difference equation, one solves the resulting system by standard techniques to obtain $U$ at all interior mesh points. Special cases generalize the well-known $O(h^6)$ approximation of smooth solutions of the Poisson equation to $O(h^6)$ approximation for the variable coefficient equation $-(\nabla p \cdot \nabla U) + F = G$. The method can be applied to other than elliptic problems.

1. The HODIE finite difference approximation

We first consider finite difference approximation for the real linear elliptic Dirichlet boundary value problem:

$$L[u] = Au_{xx} + 2Bu_{xy} + Cu_{yy} + Du_x + Eu_y + Fu = G,$$

$$AC > B^2. \quad [1]$$

The coefficients $A$, $B$, ..., $F$, $G$ are assumed given smooth functions on a connected region $R$ with piecewise smooth boundary $\partial R$. For a given function $g$, $u = g$ on the boundary.

We first consider the case of a square mesh with mesh length $h$ and approximation away from the boundary. We approximate $u$ by values of $U$, defined at mesh points in the interior of $R$, as the solution of the linear difference equation

$$L_h u = (1/h^2) \sum_{i=0}^8 a_i U_i = \beta_1 G_1 = I_h G. \quad [2]$$

In Eq. 2, the sum of the $a_i U_i$ is taken over nine mesh points in the interior of $R$ called stencil points that are shown as small circles in Fig. 1. We use $S_h$ to denote the square of side $2h$ with corners labeled 5, 6, 7, and 8 and we call the point labeled 0 the central stencil point.

One key idea of the method we discuss is the use of the right side of Eq. 1 at several points in $S_h$. In Eq. 2, $I_h G$ is a linear combination of values $G_j$ of $G$ at $j$ distinct points; we call these points evaluation points, even if some of them coincide with stencil points. Approximations that use a single value of $G$ are well-known—for example, the usual five-point approximation to the Laplacian:

$$(1/h^2)[-4U(0,0)+U(-h,0)+U(h,0)+U(0,-h)+U(0,0)] = G(0,0).$$

In our terminology, $G(0,0)$ is the value of $G$ at an evaluation point—it coincides with a stencil point. The use of several evaluation points gives the high accuracy of the scheme. Fig. 1 illustrates a simple case; evaluation points are indicated by $x$s and nine of these coincide with stencil points. The other four are nonstencil points.

Rosser (1) analyzes a difference equation of the form $[2]$ which yields an approximation to smooth solutions of the Poisson equation; he uses enough values of $G$ at mesh points to obtain $O(h^6)$ approximation. The Mehrstellenverfahren ("Hermitian") method of Collatz (2) uses linear combinations of $G$ and its derivatives, but only at stencil points. Young and Dauwalder (3) give formulas to determine the coefficients of a Mehrstellenverfahren approximation to [1]; these are obtained by use of Taylor's series expansion about the central stencil point.

A second key idea of the method we discuss is to choose the coefficients $a_1, \beta_1$ to make the approximation exact on some given finite dimensional linear space $S$, such as the space $P_M$ of polynomials of degree at most $M$. That is, when the dimension of the space $S$ is $K + 1 (K + 1 = [M + 1][M + 2]/2$ for $P_M$), then for any basis $s_0, \ldots, s_K$ of $S$, the coefficients satisfy

$$(1/h^2) \sum_{i=0}^8 a_i(s_i)h = \sum_{j=1}^J \beta_j(Ls_k) = 0, \ldots, K. \quad [3]$$

In this respect, this method is different from the Mehrstellenverfahren and the methods in refs. 1 and 3.

A third key idea is the use of nine stencil points which leads to a block tridiagonal matrix equation for the values of $U$. Such equations are amenable to standard, efficient computational schemes.

A fourth key idea is the ease of approximation of general linear boundary conditions given on curved boundaries (see Section 5). The block tridiagonal structure mentioned above is preserved and evaluation of $A, \ldots, F, G$ outside the closed region $R$ is not needed.

If the coefficients of Eq. 2 are normalized by making the sum of the $\beta_j$ equal to unity, then $I_h G = G_0 + O(h)$. Thus the operator $I_h$ is a perturbation of the identity operator. We call the scheme described above High Order Difference Approximation with Identity Expansion and use the acronym HODIE. A complete analysis of the HODIE method as applied to ordinary differential equations will appear elsewhere (4).

Abbreviation: HODIE, high order difference approximation with identity expansion.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.
2. Examples

As a simple example, we consider a new $O(h^6)$ approximation to the Poisson equation: $\nabla^2 u = u_{xx} + u_{yy} = G$. Using the space $S = P_7$, one obtains the well-known nine-point approximation for the Laplacian (5) for $L_h$. The dimension of $P_7$ is 36 and one would expect to need $J = 28$ evaluation points. But, the Laplacian has enough symmetry so that 13 symmetric evaluation points are sufficient, in particular those indicated by xs in Fig. 1. We obtain

$$\frac{1}{6h^2} \begin{pmatrix} 1 & 4 & 1 \\ 4 & -20 & 4 \\ 1 & 4 & 1 \end{pmatrix} u = \frac{1}{360} \begin{pmatrix} 48 & 48 & 48 \\ 48 & 148 & 48 \\ 48 & 48 & 48 \end{pmatrix} G. \quad [4]$$

The value of the right side requires, on the average, two evaluations of $G$ for each interior mesh point. For a different approximation, see Rosser (1).

If $U$ and $G$ are replaced by $u$ and $\nabla^2 u$, where $u$ is in $C^6(\overline{\Omega})$, the space of functions with continuous eighth derivatives on $\overline{\Omega}$, then Eq. 4 fails to be an equality by terms of order $O(h^6)$. Thus, the truncation error is $O(h^6)$ for $u \in C^6(\overline{\Omega})$. Moreover, the difference operator $L_h$ is of monotone type: for $v$ zero on the boundary, $L_h v \leq 0$ implies that $v \leq 0$. It follows that the discretization error, defined as the maximum of $|U - u|$ at mesh points, is also $O(h^6)$ when $u \in C^6(\overline{\Omega})$ and $\overline{\Omega}$ is the union of squares $S_k$.

With $S$ a space of polynomials, $O(h^6)$ approximation is optimal for the nine-point stencil of Fig. 1. This follows from theorem 11 of Birkhoff and Gulati (6) who display an eighth-degree harmonic polynomial that is nonzero at the central stencil point and zero at the outer eight stencil points.

However, the HODIE method gives $O(h^6)$ approximation not only to sufficiently smooth solutions of the Poisson equation but also to the more general differential equation

$$-\text{div}(p \, \text{grad}(u)) + Fu = G. \quad [5]$$

This equation is important in applications such as nuclear reactor design and petroleum reservoir analysis.

We report on results from one of our test cases:

$$-(\exp(xy)u_x)_x - (\exp(xy)u_y)_y + 2(x^2 + y^2)\exp(xy)u = G, \quad [6]$$

for $R$ the unit square. We did not try to minimize the number of evaluation points and used 20, the 13 of Fig. 1 and 7 of the 8 midpoints of the edges of $S_k$ that join stencil points. To measure the discretization error as a function of mesh length $h$, we chose analytic solutions, $u$, and from them determined $G$ and the boundary function $g$. Specifically, we considered

Case I: $u(x,y) = \exp(xy)$ with $G(x,y) = 0$,

Case II: $u(x,y) = \exp(-xy)$ with $G(x,y) = 2(x^2 + y^2)$, \quad [7]

Case III: $u(x,y) = (3 - x^2 - y^2)^{1/3}$. \quad [8]

[For brevity, we do not display the complicated expression for $G$ of case III obtained by substituting $u = (3 - x^2 - y^2)^{1/3}$ into the left side of Eq. 6.] Table 1 lists values of the discretization error for various values of $h$ and also these errors divided by $h^p$ for $p = 6$ or $p = 7$ which shows that the error is $O(h^7)$ for Case I and $O(h^6)$ for the other two cases.

3. Computational techniques and complexity

Appropriate choice of basis elements simplifies the evaluation of the $a_i, b_i$. For $S = P_M$, we use a basis that makes the system 3 reducible. We choose basis elements $s_0, \ldots, s_8$ which span the space of biquadratic polynomials. For the other basis elements, we use polynomials vanishing at all stencil points; each of these has a factor of $x(x^2 - h^2)$ or $y(y^2 - h^2)$.

We first solve the system

$$\beta_1 = 1, \quad [7a]$$

$$\sum_{j=1}^{M} \beta_j(Ls_k) = -(Ls_k)_i, \quad k = 9, 10, \ldots, K, \quad [7b]$$

to obtain the $\beta_i$s. Typically we use $K = J + 7$ so that there are as many equations as unknowns. But in some cases, such as those mentioned in Section 2, the symmetry of the operator $L$ allows system 7b to be solved for some $K$ greater than $J + 7$.

After the $\beta_i$s are evaluated, we solve the system

$$\left( \frac{1}{h^2} \right) \sum_{i=0}^{s} a_i(s_k)_h = \sum_{j=1}^{M} \beta_j(Ls_k)_i, \quad k = 0, \ldots, 8, \quad [8]$$

for the stencil coefficients $a_i$.

Having the coefficients of $Ls_k$ for all $S_k$, we evaluate $I_k G$ and then solve the system of difference Eq. 2 to obtain the estimates $U$ of $u$.

For sufficiently smooth $u$, the discretization error as a function of $h$ depends on the number $J$ of evaluation points and
the number of interior mesh points. For a unit square whose sides are divided into \( N = 1/h \) equal subintervals, there are \( (N-1)^2 \) interior mesh points. Since \( J \) is fixed, the amount of work in determining the \( a_i, \beta_j \) increases with the number of difference equations, \( (N-1)^2 \). But the work involved in solving the system of difference equations to obtain \( U \) increases at a faster rate. For example, if band elimination were used, this work increases as \( (N-1)^3 \). Consequently, the major part of the work occurs in the solution of the system of difference equations, and the work involved in determining the coefficients \( a_i, \beta_j \) that give high accuracy is minor. For a more detailed analysis and comparison with the work involved in other methods, see Lynch and Rice.

4. Outline of theoretical results

For a given space \( Q \) of functions in the domain of \( L \), the truncation operator \( T_h \) is defined by \( T_hq = -L_hq + \sum_{j=0}^{N} q_j \), and the truncation error is defined as the maximum norm of \( T_hq \). When the solution \( u \) is in \( Q \), so that \( Lu = G \), then one obtains an equation for the error \( e = U - u \) in terms of the truncation operator:

\[
L_h e = L_h U - L_h u = I_h G - L_h u = T_h u.
\]

In some cases, such as when \( L_h \) is of monotone type, one can show that an \( O(h^p) \) bound on the truncation error gives an \( O(h^p) \) bound on the discretization error defined as the maximum norm of \( e \).

We begin by considering Eq. 1 with constant coefficients and use lower case letters to denote constants:

\[
Lu = au_{xx} + 2bu_{xy} + cu_{yy} + du_x + ev_y + fu = G.
\]

Consider approximation with \( S = P_M \). For \( p \in P_M \), \( L_p \) is an element of \( P_M \); if \( f \neq 0 \) \( L \) maps \( P_M \) onto \( P_M \). One sees that \( 7b \) with \( K = J + 7 \) has one and only one solution except when \( f \) is an eigenvalue of the linear system. Since \( s_k, k = 0, \ldots, 8 \), in 8 forms a basis for the biquadratic polynomials, there is a unique set of \( a_i \) that satisfies 8. Hence, with few exceptions, there exists a unique set of \( a_i, \beta_j \) that satisfies 7 and 8 provided \( f \neq 0 \).

When \( f = 0 \) and \( d \neq 0 \) or \( e \neq 0 \), one also gets unique \( a_i, \beta_j \) that satisfy 7 and 8 with only a few exceptions. If the coefficients \( A, \ldots, F \) in \( L \) were differentiable, then the terms in 7 and 8 for the variable coefficient case would differ by \( O(h) \) from the terms for some constant coefficient operator \( L \). Hence for \( K = J + 7 \) and any sufficiently small \( h \), there are \( a_i, \beta_j \) which satisfy 7 and 8 provided one of \( D, E, F \) is nonzero—again with only a few exceptions.

When, however, in 10 the constants \( d, e, f \) are each zero, then \( L \) maps \( P_M \) into \( P_{M-2} \). Furthermore, there is a subspace \( N_M \) of \( P_M \) that has the dimension 2M + 1 which is also a subspace of the null space of \( L \). For \( L = 0 \), this is the space of harmonic polynomials of degree \( M \). For \( s_k \) in \( N_M \), 3 reduces to

\[
(1/h^2) \sum_{i=0}^{N} a_i (s_i h) = 0
\]

and when this (for all \( s_k \) in \( N_M \)) implies that \( \alpha_i = 0, i = 0, 1, \ldots, 8 \), then the sum in 2 that involves the approximation \( U \) vanishes. In this case one obtains no expression \( U \) of the form

5. Extensions and generalizations

The HODIE method is not limited to second-order operators, to elliptic operators, to operators in two independent variables as in 1, or to a nine-point stencil in two dimensions. For example, there is a three-dimensional \( O(h^6) \) analogue of 4 that uses 27 stencil points and 23 auxiliary points (see Lynch 7).

Although we have only done computer experiments with \( S \) a space of polynomials, other spaces can be used. For example, near a corner where a derivative of \( u \) has a singularity, an appropriate space can be used provided the nature of the singularity is known. Dershem (7) has obtained \( O(h^6) \) approximation to solutions of ordinary differential equations that behave as \( x^* < s < 1 \) by using a single value of \( G \) for each group of three stencil points.

The limit of \( O(h^6) \) for approximation to the Laplacian with \( S \) a space of polynomials is due to the fact that harmonic polynomials of arbitrarily high degree exist. When lower order derivatives also appear, however, there is, in principle, no reason for this limitation. We have experimented with a number of operators and the numerical computations have been stable; the results in Table 1 for Case I illustrate one example in which \( O(h^7) \) discretization error is obtained.

Finally, the difference Eq. 2 can be modified to take into account general linear boundary conditions

\[
L_hu = Pu + Qu_x + Tu_y = g,
\]

on curved portions of the boundary. We assume that \( h \) is sufficiently small so that the portion of the boundary which cuts

\[
*\text{Lynch, R. E. & Rice, J. R. (1975) The HODIE Method, a Brief Introduction with Summary of Computational Properties, Purdue University Department of Computer Science Report CSD-TR 170 (Purdue University, West Lafayette, IN).}
\]

\[
\text{Lynch, R. E. (1977) } O(h^6) \text{ Accurate Finite Difference Approximation to Solutions of the Poisson Equation in Three Variables, Purdue University Department of Computer Science Report CSD-TR 221 (Purdue University, West Lafayette, IN); also, (1977) } O(h^6) \text{ Discretization Error Finite Difference Approximation to Solutions of the Poisson Equation in Three Variables, Report CSD-TR 230.}
\]
As above, $U_i$ denotes the value of $U$ at a stencil point, but here the stencil points are the mesh points in the intersection $S_h \cap R$; $G_j$ denotes values of $G$ at evaluation points in the same intersection. The values $g_m$ are taken at $M$ points on the boundary that cuts through $S_h$ and these values are indicated by small rectangles in Fig. 2. The equations which give the coefficients are

$$
(1/h^2) \sum_{i=0}^{I} \alpha_i s_k U_i = \sum_{j=0}^{J} \beta_j (Ls_k)_j + \sum_{m=1}^{M} \gamma_m (Ls_m)_m, \quad k = 0, \ldots, K.
$$

After the coefficients are evaluated, the value of the right side of 12 is known because $G$ and $g$ are given. Note that the structure of the coefficient matrix that arises from the left side of the difference Eq. 12 is the same structure as the nine-point approximation to $L$ because the only unknowns are $U$ at interior mesh points.

We thank the National Science Foundation for partial support under National Science Foundation Grant MCS 76-10225.