

Analysis of a Block Red-Black Preconditioner Applied to the Hermite Collocation Discretization of a Model Parabolic Equation

Stephen H. Brill

*Department of Mathematics and Computer Science
Boise State University
Boise, Idaho, USA*

George F. Pinder

*Department of Civil and Environmental Engineering
University of Vermont
Burlington, Vermont, USA*

This is a preprint of an article published in *Numerical Methods for Partial Differential Equations*, Vol. 17, No. 6, pp. 584-606, November 2001. © copyright 2001 John Wiley & Sons, Inc.

Received September 1, 2000

We are concerned with the numerical solution of a model parabolic partial differential equation (PDE) in two spatial dimensions, discretized by Hermite collocation. In order to efficiently solve the resulting systems of linear algebraic equations, we choose the Bi-CGSTAB method of van der Vorst (1992) with block Red-Black Gauss-Seidel (RBGS) preconditioner.

In this paper, we give analytic formulae for the eigenvalues that control the rate at which Bi-CGSTAB/RBGS converges. These formulae, which depend on the location of the collocation points, can be utilized to determine where the collocation points should be placed in order to make the Bi-CGSTAB/RBGS method converge as quickly as possible. Along these lines, we discuss issues of choice of time-step size in the context of rapid convergence. A complete stability analysis is also included. © 2000 John Wiley & Sons, Inc.

Keywords: Hermite collocation, Bi-CGSTAB method, eigenvalue formulae, stability

I. INTRODUCTION

In the paper [1], we apply a block Red-Black numbering technique to the Hermite collocation discretization of Poisson's equation. The resulting system of linear algebraic equations is solved using the well known Bi-CGSTAB method [2] with a preconditioner

Numerical Methods for Partial Differential Equations 2, No. 3, 1?? (2000)

© 2000 John Wiley & Sons, Inc.

CCC 1063-8539/94/030117-13

where I represents the identity matrix of appropriate size. For rapid convergence, we want $P^{-1}A$ to be “close” to I [6]. This is clearly equivalent to

$$I - P^{-1}A = \left[\begin{array}{c|c} & -R^{-1}U \\ \hline & B^{-1}LR^{-1}U \end{array} \right]$$

being “close” to the “null” matrix (i.e., the matrix whose entries are all zero). The null matrix has all its eigenvalues equal to zero. Since $I - P^{-1}A$ may be viewed as a block upper-triangular matrix, its spectrum (i.e., set of eigenvalues) is given by the union of the spectra of those matrices on its diagonal blocks, namely the null matrix and

$$J = B^{-1}LR^{-1}U. \quad (2.7)$$

We therefore expect the fastest convergence when the eigenvalues of J are clustered near the origin of the complex plane. Formulae for these eigenvalues are given by complicated expressions found in Theorem 3.5 of [1].

Armed with these eigenvalue formulae, we see that we can enhance the speed of convergence of BiCGSTAB/RBGS in an optimal manner by choosing ξ appropriately. The traditional value for ξ is $\frac{1}{\sqrt{12}}$ which, given certain smoothness conditions, provides greatest accuracy [7]. We refer to $\xi = \frac{1}{\sqrt{12}}$ as the “Gaussian” value for ξ because this choice of ξ is equivalent to placing the collocation points at the points of Gaussian quadrature. For fixed accuracy and fixed problem size, we find that choosing ξ optimally results in speedups (often significantly) greater than 20% when compared to using its Gaussian value. Furthermore, for sufficiently (modestly) large m , the optimal value for ξ is relatively insensitive to m and is approximately equal to 0.154.

III. THE MODEL PARABOLIC EQUATION

We now consider solving the parabolic PDE

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} - H(x, y, t), \quad (3.1)$$

with Dirichlet boundary conditions and appropriate initial conditions, discretized by Hermite collocation on the unit square $[0, 1] \times [0, 1]$ with m finite elements in each of the coordinate directions. We approximate the time derivative by

$$\frac{\partial u}{\partial t} \approx \frac{u^{(n+1)} - u^{(n)}}{\Delta t}, \quad (3.2)$$

where the superscript (n) indicates the value of u after n time steps.

The matrix A of (2.2), whose entries are given in (2.5), can be considered to represent the operator $\frac{\partial^2(\bullet)}{\partial x^2} + \frac{\partial^2(\bullet)}{\partial y^2}$ in (2.1). Analogously, we form the matrix C to represent the operator (\bullet) . (Details may be found in [8].) We find that matrix C has *exactly* the same structure as matrix A , where the entries of C are given below. Note the symmetry $c_{i,j} = c_{j,i}$ and thus the correspondence in the indexing of (2.5) and (3.3). Once again, the scaling procedure of [3], [4], and [5] has been utilized in the generation of the values

$c_{i,j}$, $i, j = 1, 2, 3, 4$.

$$\begin{aligned}
c_{1,1} &= \frac{(1-\xi)^2(1+2\xi)^4}{4} \\
c_{1,2} = c_{2,1} &= \frac{(1-2\xi)(1-\xi)(1+2\xi)^4}{16} \\
c_{1,3} = c_{3,1} &= \frac{(1-2\xi)^2(1-\xi)(1+\xi)(1+2\xi)^2}{4} \\
c_{1,4} = c_{4,1} &= \frac{(1-2\xi)^2(1-\xi)(1+2\xi)^3}{16} \\
c_{2,2} &= \frac{(1-2\xi)^2(1+2\xi)^4}{64} \\
c_{2,3} = c_{3,2} &= \frac{(1-2\xi)^3(1+\xi)(1+2\xi)^2}{16} \\
c_{2,4} = c_{4,2} &= \frac{(1-2\xi)^4(1+\xi)^2}{64} \\
c_{3,3} &= \frac{(1-2\xi)^4(1+\xi)^2}{4} \\
c_{3,4} = c_{4,3} &= \frac{(1+2\xi)(1+\xi)(1-2\xi)^4}{16} \\
c_{4,4} &= \frac{(1+2\xi)^2(1-2\xi)^4}{64}
\end{aligned} \tag{3.3}$$

If we now introduce (3.2) into (3.1) and evaluate the right side of (3.1) at the collocation points at time $\theta t^{(n+1)} + (1-\theta)t^{(n)}$, where $0 \leq \theta \leq 1$, then we obtain the matrix form of the collocation discretization of the parabolic PDE (3.1):

$$\frac{C\mathbf{x}^{(n+1)} - C\mathbf{x}^{(n)}}{\Delta t} = \theta \left[A\mathbf{x}^{(n+1)} - \mathbf{b}^{(n+1)} \right] + (1-\theta) \left[A\mathbf{x}^{(n)} - \mathbf{b}^{(n)} \right]. \tag{3.4}$$

Letting

$$\tau = \theta \Delta t \tag{3.5}$$

and $\bar{\tau} = (1-\theta)\Delta t$, we may express (3.4) as

$$(C - \tau A)\mathbf{x}^{(n+1)} = (C + \bar{\tau} A)\mathbf{x}^{(n)} - \left(\tau \mathbf{b}^{(n+1)} + \bar{\tau} \mathbf{b}^{(n)} \right). \tag{3.6}$$

Equation (3.6) defines how we may move from time step (n) to time step $(n+1)$ because, at time step (n) , all the vectors on the right side of (3.6) contain known values. We may therefore apply to (3.6) the preconditioned Bi-CGSTAB algorithm that we developed for (2.2). That is, at each time step in (3.6), we iterate to convergence using Bi-CGSTAB/RBGS.

We may express (3.6) as

$$A^*\mathbf{x}^{(n+1)} = \mathbf{b}^{*(n+1)}$$

where $A^* = C - \tau A$ and $\mathbf{b}^{*(n+1)}$ is the right side of (3.6). Because matrices A and C have the same structure, we may write, analogously to (2.4) and (2.6) respectively,

$$A^* = \left[\begin{array}{c|c} R^* & U^* \\ \hline L^* & B^* \end{array} \right].$$

and

$$P^* = \left[\begin{array}{c|c} R^* & \\ \hline L^* & B^* \end{array} \right].$$

We now apply the analysis in [1] to determine the eigenvalues of

$$J^* = (B^*)^{-1}L^*(R^*)^{-1}U^*$$

(compare to (2.7)), which govern the rate at which BiCGSTAB/RBGS converges when applied to (3.1). Use of this analysis is permissible because matrices A and C have precisely the same structure.

Theorem 3.1. *The eigenvalues of J^* are given by the following recipe. We first define*

$$\nu = \frac{\tau}{h^2}. \quad (3.7)$$

$$\begin{aligned} \text{Let } a_{\alpha_j} &= [(5 - 64\xi^2 + 288\xi^4 - 512\xi^6 + 256\xi^8) \\ &+ c_j(-1 + 16\xi^2 - 96\xi^4 + 256\xi^6 - 256\xi^8)] \\ &+ [(176 - 2240\xi^2 + 7424\xi^4 - 5120\xi^6) + c_j(-16 + 448\xi^2 - 2816\xi^4 + 5120\xi^6)] \nu \\ &+ [(1280 - 23552\xi^2 + 24576\xi^4) + c_j(-256 + 5120\xi^2 - 24576\xi^4)] \nu^2. \end{aligned}$$

$$\begin{aligned} \text{Let } b_{\alpha_j} &= [(5 + 20\xi - 24\xi^2 - 176\xi^3 - 64\xi^4 + 448\xi^5 + 384\xi^6 - 256\xi^7 - 256\xi^8) \\ &+ c_j(-1 - 4\xi + 8\xi^2 + 48\xi^3 - 192\xi^5 - 128\xi^6 + 256\xi^7 + 256\xi^8)] \\ &+ [(176 + 1024\xi + 448\xi^2 - 5120\xi^3 - 5888\xi^4 + 4096\xi^5 + 5120\xi^6) \\ &+ c_j(-16 - 128\xi + 64\xi^2 + 1536\xi^3 + 1280\xi^4 - 4096\xi^5 - 5120\xi^6)] \nu \\ &+ [(1280 + 10752\xi + 19456\xi^2 - 14336\xi^3 - 24576\xi^4) \\ &+ c_j(-256 - 1536\xi - 1024\xi^2 + 14336\xi^3 + 24576\xi^4)] \nu^2. \end{aligned}$$

$$\begin{aligned} \text{Let } a_{\beta_j} &= [(-5 + 69\xi^2 - 352\xi^4 + 800\xi^6 - 768\xi^8 + 256\xi^{10}) \\ &+ c_j(1 - 17\xi^2 + 112\xi^4 - 352\xi^6 + 512\xi^8 - 256\xi^{10})] \\ &+ [(-96 + 1712\xi^2 - 7872\xi^4 + 11520\xi^6 - 5120\xi^8) \\ &+ c_j(-272\xi^2 + 2496\xi^4 - 6912\xi^6 + 5120\xi^8)] \nu \\ &+ [(-192 + 11520\xi^2 - 36864\xi^4 + 24576\xi^6) \\ &+ c_j(192 - 2304\xi^2 + 18432\xi^4 - 24576\xi^6)] \nu^2. \end{aligned}$$

$$\begin{aligned} \text{Let } b_{\beta_j} &= \\ &+ [(-5 - 30\xi - 21\xi^2 + 184\xi^3 + 320\xi^4 - 288\xi^5 - 800\xi^6 + 128\xi^7 + 768\xi^8 - 256\xi^{10}) \\ &+ c_j(1 + 6\xi + \xi^2 - 56\xi^3 - 80\xi^4 + 160\xi^5 + 352\xi^6 - 128\xi^7 - 512\xi^8 + 256\xi^{10})] \\ &+ [(-96 - 816\xi - 1456\xi^2 + 2688\xi^3 + 7872\xi^4 - 1792\xi^5 - 11520\xi^6 + 5120\xi^8) \\ &+ c_j(48\xi + 16\xi^2 - 1152\xi^3 - 2496\xi^4 + 1792\xi^5 + 6912\xi^6 - 5120\xi^8)] \nu \\ &+ [(-192 - 3456\xi - 11520\xi^2 + 4608\xi^3 + 36864\xi^4 - 24576\xi^6) \\ &+ c_j(192 + 1152\xi + 2304\xi^2 - 4608\xi^3 - 18432\xi^4 + 24576\xi^6)] \nu^2. \end{aligned}$$

$$\text{Now define } \alpha_j^\pm = \frac{a_{\alpha_j} \pm 256q_j\nu^2}{b_{\alpha_j}} \text{ and } \beta_j^\pm = \frac{a_{\beta_j} \pm 192q_j\nu^2}{b_{\beta_j}}, \text{ where}$$

$$q_j = 2\xi\sqrt{(16\xi^4 - 24\xi^2 + 21) + c_j(-32\xi^4 + 18) + c_j^2(16\xi^4 + 24\xi^2 - 3)}.$$

The sets $\{d_i\}_{i=0}^{2m-1}$ and $\{\bar{d}_i\}_{i=0}^{2m-1}$ are defined by

$$\{d_i\}_{i=0}^{2m-1} = \{\alpha_0^+, \alpha_1^+, \alpha_1^-, \dots, \alpha_{m-1}^+, \alpha_{m-1}^-, \alpha_m^+\}$$

$$\{\bar{d}_i\}_{i=0}^{2m-1} = \{\beta_0^-, \beta_1^-, \beta_1^+, \dots, \beta_{m-1}^-, \beta_{m-1}^+, \beta_m^-\}.$$

Then the $2m^2$ eigenvalues of J^* are the elements of the set

$$\{\mu : \mu = d_i^2\} \cup \{\mu : \mu = \bar{d}_i d_i\} \cup \left\{ \mu : \mu = \frac{z_{k,i}^2 + 2\bar{d}_i d_i \pm \sqrt{z_{k,i}^2 (z_{k,i}^2 + 4\bar{d}_i d_i)}}{2} \right\},$$

where $z_{k,i} = (\bar{d}_i - d_i) \cos \theta_k$; $\theta_k = \frac{k\pi}{m}$; $k = 1, 2, \dots, \frac{m}{2} - 1$; $i = 0, 1, \dots, 2m - 1$; $c_j = \cos \theta_j$; $j = 0, 1, \dots, m$.

IV. ANALYSIS

Consider the matrix

$$A^* = C - \tau A. \quad (4.1)$$

By comparing (2.5) and (3.3), it is clear that the entries of A are significantly larger than those of C , due primarily to the presence of the h^2 term in the denominator of the entries of A . For convenience, let us rewrite (4.1) as

$$A^* = C - \nu A^\diamond,$$

where ν is given in (3.7) and $A^\diamond = h^2 A$. We note that A^\diamond is independent of h and depends only upon ξ . Thus the entries in C and A^\diamond are of the same order.

So, for values of ν that are large, νA^\diamond dominates C and we obtain results very similar to those given in [1], including the characterization of the optimal value of ξ (i.e., the value of $\xi \in (0, \frac{1}{2})$ that produces the fastest convergence of Bi-CGSTAB/RBGS) as approximately 0.154 for sufficiently large m [1]. (The optimal value of ξ is determined using the MINOS optimization software (see [9] for documentation) applied to the objective function of ξ : $\left\| \sigma \left(I - (P^*)^{-1} A^* \right) \right\|_2^2 = \left\| \sigma (J^*) \right\|_2^2$.)

If, however, ν is small, then the presence of the matrix C becomes significant. In Figure 1, we show how $\left\| \sigma \left(I - (P^*)^{-1} A^* \right) \right\|_2$ varies as a function of ξ for various values of τ for $m = 20$. (Here $\sigma(\bullet)$ is the vector whose entries are the eigenvalues of matrix \bullet .) For $\tau > 10^{-1}$, the corresponding curve is visually indistinguishable from that of $\tau = 10^{-1}$; for $0 < \tau < 10^{-6}$, the corresponding curve is visually indistinguishable from that of $\tau = 10^{-6}$. What is particularly interesting is the fact that for sufficiently small τ , we may obtain $\left\| \sigma \left(I - (P^*)^{-1} A^* \right) \right\|_2 \approx 0$, provided ξ is chosen appropriately. Note that this is precisely the condition we desire for optimal clustering of eigenvalues about the origin of the complex plane, which in turn corresponds to fastest convergence of Bi-CGSTAB/RBGS.

In Figure 2, we consider the value of $\left\| \sigma \left(I - (P^*)^{-1} A^* \right) \right\|_2$ that corresponds to optimal ξ as a function of τ for $m = 10, 20, 30$, and 40. We see that by making τ sufficiently small, we may obtain optimal ξ that makes $\left\| \sigma \left(I - (P^*)^{-1} A^* \right) \right\|_2 \approx 0$.

Figure 3 shows the optimal value of ξ as a function of τ for $m = 10, 20, 30$, and 40. We plainly see that as τ decreases, the optimal value of ξ increases for all values of m shown. Indeed, for very small τ , the optimal value of ξ is approximately $\frac{1}{2}$. In particular, should we wish to use the Gaussian value of $\xi = \frac{1}{\sqrt{12}} \approx 0.29$, we may select τ appropriately.

Thus, at each time step, if we seek to solve the collocation discretization (3.6) as quickly as possible, we may set the value of τ to be sufficiently small to achieve

$$\left\| \sigma \left(I - (P^*)^{-1} A^* \right) \right\|_2 \approx 0.$$

Since $\tau = \theta \Delta t$, small values of τ force either small time steps or small θ (which corresponds to a mostly explicit time advance).

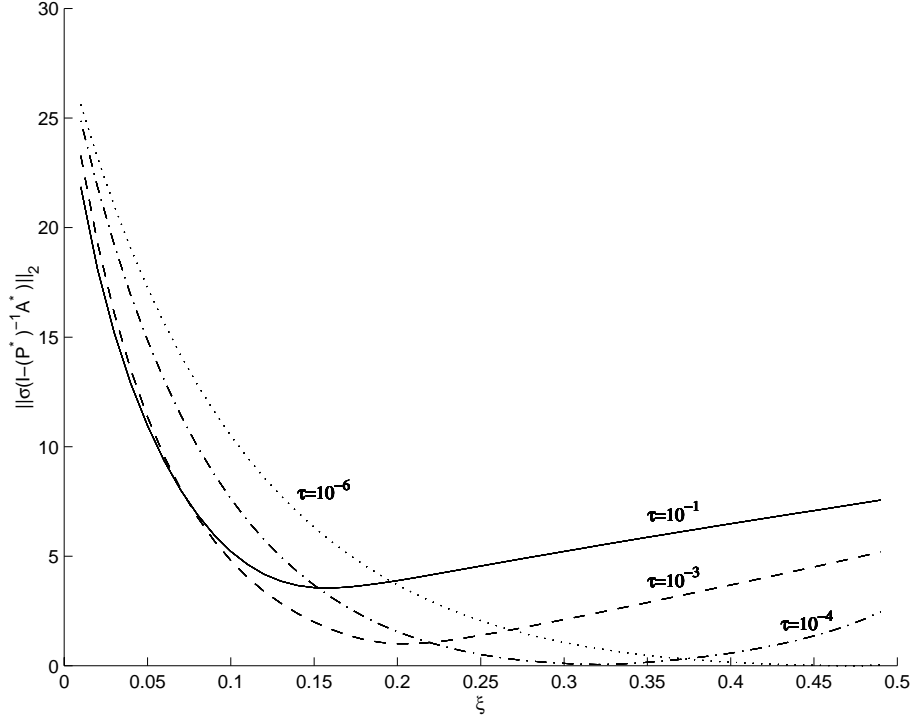


FIG. 1. $\|\sigma(I - (P^*)^{-1}A^*)\|_2$ as a function of ξ for various values of τ for $m = 20$

V. STABILITY

Let us now examine the stability of the time-stepping scheme (3.6), which may be written

$$\mathbf{x}^{(n+1)} = W\mathbf{x}^{(n)} - \mathbf{p}^{(n)}, \quad (5.1)$$

where

$$W = (C - \theta \Delta t A)^{-1} (C + [1 - \theta] \Delta t A)$$

and

$$\mathbf{p}^{(n)} = (C - \theta \Delta t A)^{-1} [\theta \Delta t \mathbf{b}^{(n+1)} + (1 - \theta) \Delta t \mathbf{b}^{(n)}].$$

For stability, we require all the eigenvalues of W to lie inside the unit circle in the complex plane. We may write

$$W = (I - \theta \rho G)^{-1} (I + [1 - \theta] \rho G),$$

where $G = h^2 C^{-1} A = C^{-1} A^\diamond$ (the entries of which depend only on ξ) and where

$$\rho = \frac{\Delta t}{h^2}. \quad (5.2)$$

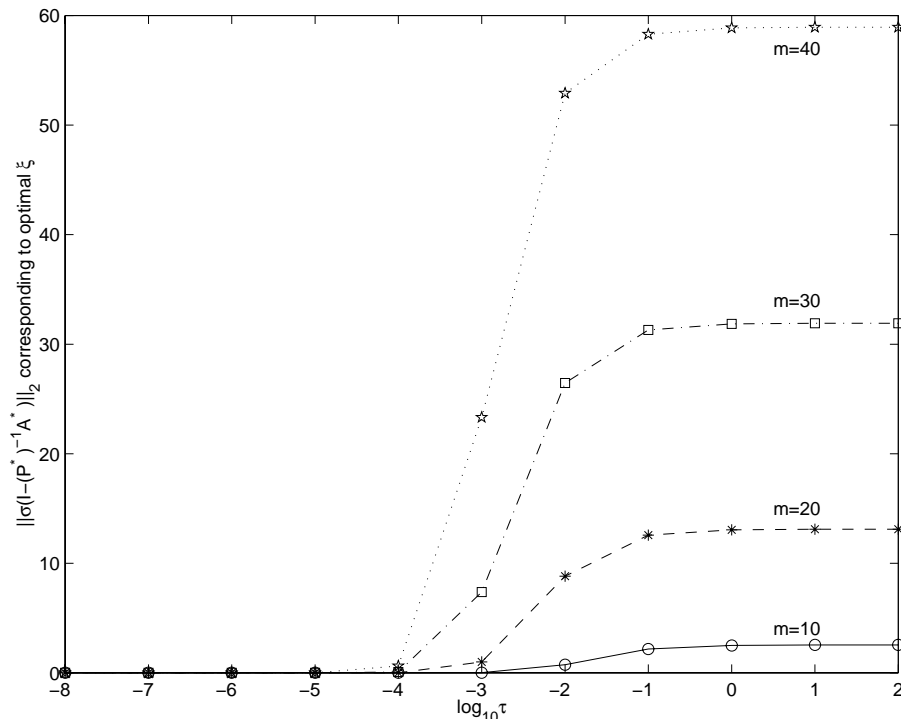


FIG. 2. $\|\sigma(I - (P^*)^{-1}A^*)\|_2$ corresponding to optimal ξ as a function of τ for various values of m

Appealing to well-known linear algebra facts (see, for example, [10]), we see that the eigenvalues of G and those of W are related: If λ is an eigenvalue of G , then $\frac{1+(1-\theta)\rho\lambda}{1-\theta\rho\lambda}$ is an eigenvalue of W .

We show in the Appendix that all eigenvalues of G are real and negative (see Lemma 7.3 and Theorem 7.7). (Actually, the eigenvalues studied in the Appendix are those of $C^{-1}A$. But if δ is an eigenvalue of $C^{-1}A$, then $h^2\delta$ is an eigenvalue of G .) If we then let $\lambda = -\omega^2$, where ω is a real number, then the condition that ensures stability of the time-stepping scheme (5.1) is

$$\left| \frac{1 - \omega^2(1 - \theta)\rho}{1 + \omega^2\theta\rho} \right| \leq 1. \quad (5.3)$$

Algebraic manipulation of (5.3) yields the stability condition

$$\omega^2\rho(1 - 2\theta) \leq 2, \quad (5.4)$$

which must be true when $\frac{1}{2} \leq \theta \leq 1$. That is, if $\frac{1}{2} \leq \theta \leq 1$, we get *unconditional* stability (i.e., stability irrespective of the value of ρ). If, on the other hand, $0 \leq \theta < \frac{1}{2}$, stability is guaranteed only if ρ satisfies (5.4) for all appropriate values of ω^2 , i.e., for all eigenvalues λ of G .

We thus see that if we want *both* unconditional stability and the very rapid convergence that occurs when $\|\sigma(I - (P^*)^{-1}A^*)\|_2 \approx 0$, we must choose very small time steps. For

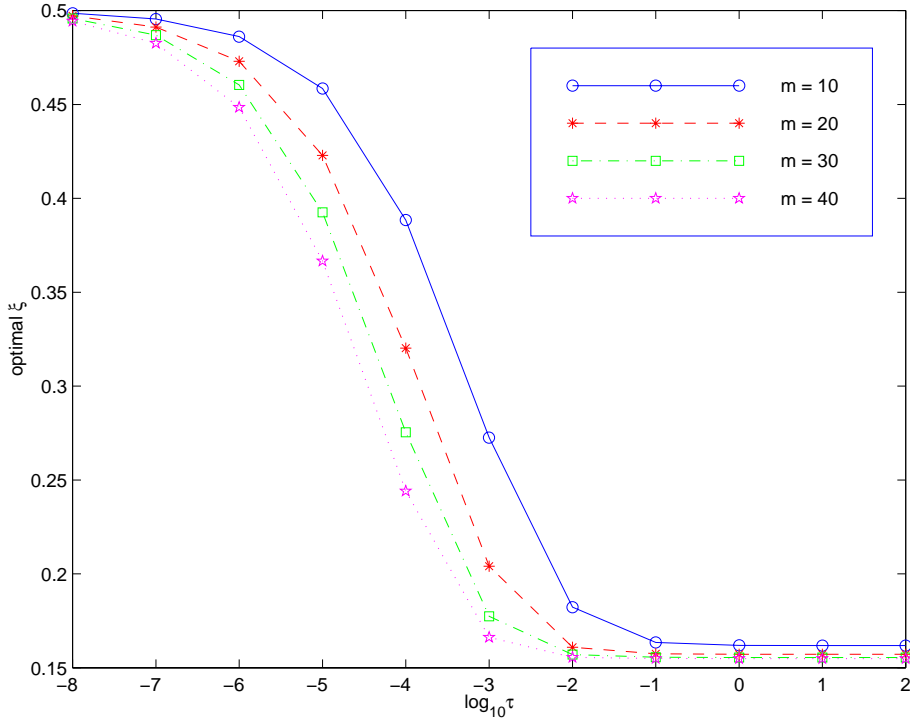


FIG. 3. optimal ξ as a function of τ for various values of m

example, with reference to Figure 1, we achieve

$$\left\| \sigma \left(I - (P^*)^{-1} A^* \right) \right\|_2 \approx 0$$

for $m = 20$ when $\xi \approx 0.32$ and $\tau \approx 10^{-4}$. Recalling (3.5), we see that unconditional stability requires that $10^{-4} < \Delta t < 2 \cdot 10^{-4}$.

Let us compare this restriction on Δt to that imposed by the fully explicit time-stepping scheme corresponding to $\theta = 0$. With reference to (5.4) and the fact that all eigenvalues of G are real and negative, it is clear that if (5.4) holds for the largest value of ω^2 (which corresponds to the eigenvalue greatest in absolute value), then (5.4) holds for *all* relevant values of ω^2 (i.e., for all eigenvalues λ of G). Thus (5.4) may be written

$$\omega_{\max}^2 \rho (1 - 2\theta) \leq 2. \quad (5.5)$$

Using Corollary 7.8, we may easily derive a formula for ω_{\max}^2 :

$$\omega_{\max}^2 = \frac{48}{1 - 4\xi^2}. \quad (5.6)$$

Note that ω_{\max}^2 depends only upon ξ ; in particular, ω_{\max}^2 is independent of problem size m .

Using (5.6), we find that $\omega_{\max}^2 \approx 81.3$ corresponds to $\xi = 0.32$. Setting $h = \frac{1}{m} = 0.05$ and recalling (5.2), we see that (5.5) requires $0 < \Delta t < 6.15 \cdot 10^{-5}$ for stability of the

TABLE I. Number of iterations required for convergence of the model parabolic equation over one time step for various time step sizes

Δt	optimal ξ	iterations
10^{-4}	0.32020	2
10^{-3}	0.20409	4
10^{-2}	0.16094	12
10^{-1}	0.15742	26

fully explicit time stepping scheme. Recall that the time step restriction for unconditional stability and very rapid convergence was $10^{-4} < \Delta t < 2 \cdot 10^{-4}$. Thus for this example, the largest time step permissible for unconditional stability and very rapid convergence is roughly three times the size of time steps imposed by stability requirements for the fully explicit time-stepping problem.

We saw above that if we take very small time steps, then it is possible to achieve very rapid convergence. A natural question to ask is: If we increase the time step size, how will the commensurate increase in iterations required for convergence compare? For example, if we double the time step size and the number of iterations increases by *less* than a factor of two, then we should take one doubled time step as opposed to two smaller ones. On the other hand, if we double the time step size and the number of iterations increases by *more* than a factor of two, then it would be less expensive to take two smaller time steps instead of one doubled larger one.

To examine this question, we solved the model parabolic equation (3.1) over one time step for various time step sizes. We used $m = 20$ and $\theta = \frac{1}{2}$. The boundary conditions, initial conditions (at $t = 0$), and forcing function were chosen to correspond to the exact solution of $u(x, y, t) = x^2 y (1 + e^{-t})$. For each value of Δt , we used the corresponding optimal value of ξ as determined from the MINOS optimization software. The results are given in Table I.

We see from Table I that taking very small step sizes to obtain extremely fast convergence is not a good strategy. For example, if we took ten time steps each of size 10^{-4} , we would require 20 iterations. This is much greater than the 4 iterations required by taking a single time step of size 10^{-3} . Although these reported results occur at the first time step, qualitatively similar results occur at later time steps as well.

VI. SUMMARY

We derived analytical formulae for the eigenvalues that govern the rate at which the Bi-CGSTAB/RBGS method converges to the solution of matrix equations arising from the Hermite collocation discretization of a model parabolic PDE. These formulae depend upon collocation point location, which can be chosen optimally. Strategies for selecting time-step size are discussed. In addition, we derive analytical formulae for eigenvalues that govern the stability of solving the model parabolic PDE.

Since (7.6) and (7.7) hold for all $k = 0, 1, \dots, m-1$, we simplify matters by considering the case $k = 0$. In this case, $x_k = x_0 = 0$, $x_{k+1} = x_1 = h$, and, using trigonometric identities, we obtain

$$\begin{aligned} v_1 &= 0 \\ v_2 &= \pi [m - j + c_2 (m + j)] \\ v_3 &= (1 - c_2) \sin j\pi h \\ v_4 &= -v_2 \cos j\pi h. \end{aligned} \tag{7.8}$$

We now eliminate c_2 from the system (7.6), (7.7), obtaining a quadratic equation in δ :

$$a\delta^2 + b\delta + c = 0, \tag{7.9}$$

where

$$\begin{aligned} a &= \frac{1}{4} (2\xi - 1) (2\xi + 1) (5 - 4\xi^2 + \varsigma - 4\xi^2 \varsigma) \\ b &= \frac{8}{h^2} (-3 + 4\xi^2 + 4\xi^2 \varsigma) \\ c &= \frac{-48}{h^4} (1 + \varsigma). \end{aligned} \tag{7.10}$$

Here $\varsigma = \cos j\pi h$. The discriminant $b^2 - 4ac$ is $\frac{16}{h^4} D^*$, where

$$D^* = (21 - 24\xi^2 + 16\xi^4) + (-18 + 32\xi^4) \varsigma + (-3 + 24\xi^2 + 16\xi^4) \varsigma^2.$$

Lemma 7.1. *The eigenvalues δ that satisfy the quadratic equation (7.9) are all real.*

Proof. It suffices to show that $D^*(\xi, \varsigma) \geq 0$ on the domain $\xi \in (0, \frac{1}{2})$, $\varsigma \in [-1, 1]$. On the boundary of the domain, we see that $D^*(\xi, \varsigma) \geq 0$:

$$\begin{aligned} D^*(0, \varsigma) &= 3(1 - \varsigma)(7 + \varsigma) \geq 0 \\ D^*\left(\frac{1}{2}, \varsigma\right) &= 4(2 - \varsigma)^2 > 0 \\ D^*(\xi, -1) &= 36 > 0 \\ D^*(\xi, 1) &= 64\xi^4 > 0. \end{aligned}$$

We will be done when we show that there are no critical points in the interior of the domain. Indeed, setting $\frac{\partial D^*}{\partial \xi} = 16\xi(1 + \varsigma)(3 + 3\varsigma + 4\xi^2 + 4\varsigma\xi^2)$ equal to zero implies that $\varsigma = -1$ (which is on the boundary of the domain) or $\varsigma = \frac{3-4\xi^2}{3+4\xi^2}$. Setting $\frac{\partial D^*}{\partial \varsigma} = 2(-9 - 3\varsigma + 24\varsigma\xi^2 + 16\xi^4 + 16\varsigma\xi^4)$ equal to zero implies $\varsigma = \frac{9-16\xi^4}{-3+24\xi^2+16\xi^4}$. We thus have critical points where ξ satisfies

$$\frac{3 - 4\xi^2}{3 + 4\xi^2} = \frac{9 - 16\xi^4}{-3 + 24\xi^2 + 16\xi^4}. \tag{7.11}$$

However, it is easy to see that (7.11) reduces to a contradiction. This shows there are no critical points in the interior of our domain, allowing us to conclude that $D^*(\xi, \varsigma)$ is indeed non-negative. \blacksquare

Lemma 7.2. *In (7.10), we have $a < 0$, $b < 0$, and $c \leq 0$.*

Proof. To show $a < 0$, it suffices to show that $a^*(\xi, \varsigma) = 5 - 4\xi^2 + \varsigma - 4\xi^2\varsigma > 0$. Note that a^* is linear in ς and that

$$\begin{aligned} a^*(\xi, -1) &= 4 > 0 \\ a^*(\xi, 1) &= 2(3 - 4\xi^2) > 0. \end{aligned}$$

Therefore $a^*(\xi, \varsigma) > 0$ and thus $a < 0$.

To show $b < 0$, it suffices to show that $b^*(\xi, \varsigma) = -3 + 4\xi^2 + 4\xi^2\varsigma < 0$. Note that b^* is linear in ς and that

$$\begin{aligned} b^*(\xi, -1) &= -3 < 0 \\ b^*(\xi, 1) &= -3 + 8\xi^2 < 0. \end{aligned}$$

Thus $b^*(\xi, \varsigma) < 0$, which implies $b < 0$.

That $c \leq 0$ is obvious. ■

The sum of the roots of (7.9) is $-\frac{b}{a}$, which must be negative. The product of the roots of (7.9) is $\frac{c}{a}$, which must be non-negative. It must therefore be the case that both roots of (7.9) are negative or that one is negative and one is zero. Since the roots of (7.9) are the eigenvalues we seek, we know that all eigenvalues δ are non-positive.

If $\delta = 0$, then the product of the roots of (7.9) is zero, so we have $c = 0$. Thus $\varsigma = -1$ (see (7.10)), which, by the definition of ς , implies that $j = m$. In this case, (7.2) reduces to

$$2m\pi c_2 A \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \\ \vdots \\ 1 \\ 0 \\ 1 \\ 1 \end{bmatrix} = \mathbf{0}.$$

Since A is non-singular, we conclude that $c_2 = 0$. But this implies that \mathbf{x} in (7.1) is a zero eigenvector, which is impossible. So $\delta = 0$ is not an eigenvalue, proving

Lemma 7.3. *All eigenvalues δ of (7.1) are negative.*

According to (7.9), the eigenvalues should occur in pairs given by the quadratic formula. However, we just saw that for $j = m$, one solution of (7.9), namely $\delta = 0$, is not an eigenvalue. If, on the other hand, we have $j = 0$, then (7.8) becomes $v_1 = v_3 = 0$ and $v_2 = -v_4 = \pi m(c_2 + 1)$. In this case, (7.6) and (7.7) are identical, producing one equation in the two unknowns c_2 and ξ :

$$\frac{c_2 + 1}{4h^3} \pi (-8 - h^2\delta + 4h^2\delta\xi^2) = 0.$$

If $c_2 = -1$, then $v_2 = v_4 = 0$, which implies the impossibility that \mathbf{x} is a zero eigenvector. Thus the case $j = 0$ also provides a single eigenvalue.

We can summarize:

Theorem 7.4. *The $2m$ eigenvalues δ of (7.1), which are all negative, are determined by solving the quadratic equation (7.9) for $j = 0, 1, \dots, m$. When $j = 0$ or $j = m$, we produce only one eigenvalue for each of these values of j . For $j = 1, 2, \dots, m - 1$, we produce two eigenvalues for each value of j .*

Corollary 7.5. *Of all the eigenvalues of $C^{-1}A$, the one that is greatest in absolute value corresponds to $j = m$.*

Proof. The eigenvalues δ are given by the quadratic formula

$$\delta = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}, \quad (7.12)$$

where a , b , and c are given in (7.10). To prove the corollary, it is clear that it suffices to show that $j = m$ maximizes a (i.e., makes the denominator in (7.12) a negative number of smallest possible absolute value); minimizes b (i.e., makes $-b$ the largest (positive) number possible); and maximizes $b^2 - 4ac$ (i.e., makes $b^2 - 4ac$ the largest (positive) number possible). Recall that $j = m$ corresponds to $\zeta = -1$.

We saw in the proof of Lemma 7.2 that a^* is linear in ζ . Since $\frac{\partial a^*}{\partial \zeta} = 1 - 4\xi^2 > 0$, we see that a^* is minimized at $\zeta = -1$. Thus a is maximized at $\zeta = -1$.

We also saw in the proof of Lemma 7.2 that b^* is linear in ζ . Since $\frac{\partial b^*}{\partial \zeta} = 4\xi^2 > 0$, we see that b^* is minimized at $\zeta = -1$. Thus b is minimized at $\zeta = -1$.

Since $b < 0$ is minimized at $\zeta = -1$, we see that b^2 is maximized at $\zeta = -1$. To maximize δ in (7.12), we thus want to subtract from b^2 the non-negative quantity $4ac$ of minimum magnitude. This is achieved when $c = 0$, which, as we saw above, occurs when $\zeta = -1$. ■

Corollary 7.6. *For the one-dimensional problem, the eigenvalue of $C^{-1}A$ of greatest magnitude is $\delta = \frac{-24}{h^2(1-4\xi^2)}$.*

Proof. It is clear that, with respect to (7.12), we have

$$\delta_+ < \delta_- \leq 0.$$

When $\zeta = -1$, which corresponds to the eigenvalue of greatest magnitude, we obtain $\delta_+ = \frac{-24}{h^2(1-4\xi^2)}$. ■

We make one last comment before tackling the two dimensional problem. It is clear that once we have obtained the value of an eigenvalue δ , we may substitute this value into either (7.6) or (7.7) to find the value of the corresponding parameter c_2 , thus allowing us to compute the eigenvector corresponding to the eigenvalue δ .

B. Two-dimensional formulation

We now discuss the analogous two dimensional problem. That is, we want to solve the eigenproblem

$$C^{-1}A\mathbf{x} = \delta\mathbf{x}. \quad (7.13)$$

In this case, matrix A represents the differential operator $\frac{\partial^2(\bullet)}{\partial x^2} + \frac{\partial^2(\bullet)}{\partial y^2}$ while matrix C represents the operator (\bullet) . As in the one-dimensional case, the eigenproblem (7.13) may be written

$$A\mathbf{x} - \delta C\mathbf{x} = \mathbf{0}. \quad (7.14)$$

This is merely the collocation discretization of the (continuous) PDE

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} - \delta u = 0 \quad (7.15)$$

with homogeneous Dirichlet boundary conditions on the boundary of $[0, 1] \times [0, 1]$. The solution of (7.15) consists of the eigenfunctions

$$u = \sin j\pi x \sin \ell\pi y \quad (7.16)$$

and the eigenvalues

$$\delta = -(j\pi)^2 - (\ell\pi)^2, \quad (7.17)$$

$j, \ell = 1, 2, 3, \dots$. Let us interpret the result (7.17) as follows. The eigenvalues given by (7.17) are precisely the sum of the eigenvalues of two one-dimensional problems (7.3).

The corresponding result for the discrete two-dimensional eigenproblem (7.14) is completely analogous to that for the continuous two-dimensional eigenproblem (7.15). We first consider to have two sets of identical eigenvalues to the one-dimensional problem: Let $\Delta_j = \{\delta_1^j, \delta_2^j, \dots, \delta_{2m}^j\}$ be the eigenvalues of the one-dimensional problem (7.2) and let $\Delta_\ell = \{\delta_1^\ell, \delta_2^\ell, \dots, \delta_{2m}^\ell\}$ be such that $\delta_k^j = \delta_k^\ell$ for $k = 1, 2, \dots, 2m$. Write δ in (7.14) as $\delta = \delta_j + \delta_\ell$, where $\delta_j \in \Delta_j$ and $\delta_\ell \in \Delta_\ell$.

We now write the four (because, for two-dimensional problems, we have four collocation points per rectangular finite element) equations analogous to (7.6) and (7.7):

$$\begin{bmatrix} M_{11} \\ M_{12} \\ M_{13} \\ -M_{14} \\ M_{21} \\ M_{22} \\ M_{23} \\ -M_{24} \\ M_{31} \\ M_{32} \\ M_{33} \\ -M_{34} \\ -M_{41} \\ -M_{42} \\ -M_{43} \\ M_{44} \end{bmatrix}^T \mathbf{v} = \begin{bmatrix} M_{13} \\ M_{14} \\ M_{11} \\ -M_{12} \\ M_{23} \\ M_{24} \\ M_{21} \\ -M_{22} \\ M_{33} \\ M_{34} \\ M_{31} \\ -M_{32} \\ -M_{43} \\ -M_{44} \\ -M_{41} \\ M_{42} \end{bmatrix}^T \mathbf{v} = \begin{bmatrix} M_{31} \\ M_{32} \\ M_{33} \\ -M_{34} \\ M_{41} \\ M_{42} \\ M_{43} \\ -M_{44} \\ M_{11} \\ M_{12} \\ M_{13} \\ -M_{14} \\ -M_{21} \\ -M_{22} \\ -M_{23} \\ M_{24} \end{bmatrix}^T \mathbf{v} = \begin{bmatrix} M_{33} \\ M_{34} \\ M_{31} \\ -M_{32} \\ M_{43} \\ M_{44} \\ M_{41} \\ -M_{42} \\ M_{13} \\ M_{14} \\ M_{11} \\ -M_{12} \\ -M_{23} \\ -M_{24} \\ -M_{21} \\ M_{22} \end{bmatrix}^T \mathbf{v} = 0 \quad (7.18)$$

where

$$\begin{aligned}
M_{11} &= a_{1,1} - \delta c_{1,1} = \left(\frac{1}{2} + \zeta\right) [-\alpha - \delta_j \left(\frac{1}{2} + \zeta\right)] + \left(\frac{1}{2} + \zeta\right) [-\alpha - \delta_\ell \left(\frac{1}{2} + \zeta\right)] \\
M_{12} &= a_{1,2} - \delta c_{1,2} = \left(\frac{1}{2} + \zeta\right) [-\beta - \delta_j \chi] + \chi [-\alpha - \delta_\ell \left(\frac{1}{2} + \zeta\right)] \\
M_{13} &= a_{1,3} - \delta c_{1,3} = \left(\frac{1}{2} + \zeta\right) [\alpha - \delta_j \left(\frac{1}{2} - \zeta\right)] + \left(\frac{1}{2} - \zeta\right) [-\alpha - \delta_\ell \left(\frac{1}{2} + \zeta\right)] \\
M_{14} &= a_{1,4} - \delta c_{1,4} = \left(\frac{1}{2} + \zeta\right) [\gamma - \delta_j \psi] + \psi [-\alpha - \delta_\ell \left(\frac{1}{2} + \zeta\right)] \\
M_{21} &= a_{2,1} - \delta c_{2,1} = \chi [-\alpha - \delta_j \left(\frac{1}{2} + \zeta\right)] + \left(\frac{1}{2} + \zeta\right) [-\beta - \delta_\ell \chi] \\
M_{22} &= a_{2,2} - \delta c_{2,2} = \chi [-\beta - \delta_j \chi] + \chi [-\beta - \delta_\ell \chi] \\
M_{23} &= a_{2,3} - \delta c_{2,3} = \chi [\alpha - \delta_j \left(\frac{1}{2} - \zeta\right)] + \left(\frac{1}{2} - \zeta\right) [-\beta - \delta_\ell \chi] \\
M_{24} &= a_{2,4} - \delta c_{2,4} = \chi [\gamma - \delta_j \psi] + \psi [-\beta - \delta_\ell \chi] \\
M_{31} &= a_{3,1} - \delta c_{3,1} = \left(\frac{1}{2} - \zeta\right) [-\alpha - \delta_j \left(\frac{1}{2} + \zeta\right)] + \left(\frac{1}{2} + \zeta\right) [\alpha - \delta_\ell \left(\frac{1}{2} - \zeta\right)] \\
M_{32} &= a_{3,2} - \delta c_{3,2} = \left(\frac{1}{2} - \zeta\right) [-\beta - \delta_j \chi] + \chi [\alpha - \delta_\ell \left(\frac{1}{2} - \zeta\right)] \\
M_{33} &= a_{3,3} - \delta c_{3,3} = \left(\frac{1}{2} - \zeta\right) [\alpha - \delta_j \left(\frac{1}{2} - \zeta\right)] + \left(\frac{1}{2} - \zeta\right) [\alpha - \delta_\ell \left(\frac{1}{2} - \zeta\right)] \\
M_{34} &= a_{3,4} - \delta c_{3,4} = \left(\frac{1}{2} - \zeta\right) [\gamma - \delta_j \psi] + \psi [\alpha - \delta_\ell \left(\frac{1}{2} - \zeta\right)] \\
M_{41} &= a_{4,1} - \delta c_{4,1} = \psi [-\alpha - \delta_j \left(\frac{1}{2} + \zeta\right)] + \left(\frac{1}{2} + \zeta\right) [\gamma - \delta_\ell \psi] \\
M_{42} &= a_{4,2} - \delta c_{4,2} = \psi [-\beta - \delta_j \chi] + \chi [\gamma - \delta_\ell \psi] \\
M_{43} &= a_{4,3} - \delta c_{4,3} = \psi [\alpha - \delta_j \left(\frac{1}{2} - \zeta\right)] + \left(\frac{1}{2} - \zeta\right) [\gamma - \delta_\ell \psi] \\
M_{44} &= a_{4,4} - \delta c_{4,4} = \psi [\gamma - \delta_j \psi] + \psi [\gamma - \delta_\ell \psi]
\end{aligned}$$

and \mathbf{v} (analogous to (7.8)) contains the linear combination of relevant entries of the discretized eigenfunctions (7.16):

$$\mathbf{v} = \sum_{k=1}^8 d_k \mathbf{w}_k,$$

where

$$\mathbf{w}_1 = \begin{bmatrix} 0 \\ (m-j)\pi \sin L \\ \sin J \sin L \\ -(m-j)\pi \cos J \sin L \\ \hline 0 \\ -(m-j)(m-\ell)\pi^2 \cos L \\ -(m-\ell)\pi \cos L \sin J \\ (m-j)(m-\ell)\pi^2 \cos J \cos L \\ \hline 0 \\ 0 \\ 0 \\ 0 \\ \hline 0 \\ (m-j)(m-\ell)\pi^2 \\ (m-\ell)\pi \sin J \\ -(m-j)(m-\ell)\pi^2 \cos J \end{bmatrix} \quad \mathbf{w}_2 = \begin{bmatrix} 0 \\ -(m-j)\pi \sin L \\ -\sin J \sin L \\ (m-j)\pi \cos J \sin L \\ \hline 0 \\ -(m-j)(m+\ell)\pi^2 \cos L \\ -(m+\ell)\pi \cos L \sin J \\ (m-j)(m+\ell)\pi^2 \cos J \cos L \\ \hline 0 \\ 0 \\ 0 \\ 0 \\ \hline 0 \\ (m-j)(m+\ell)\pi^2 \\ (m+\ell)\pi \sin J \\ -(m-j)(m+\ell)\pi^2 \cos J \end{bmatrix}$$

$$\mathbf{w}_3 = \left[\begin{array}{c} 0 \\ (m-\ell)\pi \sin J \\ \sin J \sin L \\ -(m-\ell)\pi \cos L \sin J \\ \hline 0 \\ -(m-j)(m-\ell)\pi^2 \cos J \\ -(m-j)\pi \cos J \sin L \\ (m-j)(m-\ell)\pi^2 \cos J \cos L \\ \hline 0 \\ 0 \\ 0 \\ 0 \\ \hline 0 \\ (m-j)(m-\ell)\pi^2 \\ (m-j)\pi \sin L \\ -(m-j)(m-\ell)\pi^2 \cos L \end{array} \right]$$

$$\mathbf{w}_4 = \left[\begin{array}{c} 0 \\ -(m-\ell)\pi \sin J \\ -\sin J \sin L \\ (m-\ell)\pi \cos L \sin J \\ \hline 0 \\ -(m+j)(m-\ell)\pi^2 \cos J \\ -(m+j)\pi \cos J \sin L \\ (m+j)(m-\ell)\pi^2 \cos J \cos L \\ \hline 0 \\ 0 \\ 0 \\ 0 \\ \hline 0 \\ (m+j)(m-\ell)\pi^2 \\ (m+j)\pi \sin L \\ -(m+j)(m-\ell)\pi^2 \cos L \end{array} \right]$$

$$\mathbf{w}_5 = \left[\begin{array}{c} 0 \\ (m+\ell)\pi \sin J \\ -\sin J \sin L \\ -(m+\ell)\pi \cos L \sin J \\ \hline 0 \\ -(m-j)(m+\ell)\pi^2 \cos J \\ (m-j)\pi \cos J \sin L \\ (m-j)(m+\ell)\pi^2 \cos J \cos L \\ \hline 0 \\ 0 \\ 0 \\ 0 \\ \hline 0 \\ (m-j)(m+\ell)\pi^2 \\ -(m-j)\pi \sin L \\ -(m-j)(m+\ell)\pi^2 \cos L \end{array} \right]$$

$$\mathbf{w}_6 = \left[\begin{array}{c} 0 \\ -(m+\ell)\pi \sin J \\ \sin J \sin L \\ (m+\ell)\pi \cos L \sin J \\ \hline 0 \\ -(m+j)(m+\ell)\pi^2 \cos J \\ (m+j)\pi \cos J \sin L \\ (m+j)(m+\ell)\pi^2 \cos J \cos L \\ \hline 0 \\ 0 \\ 0 \\ 0 \\ \hline 0 \\ (m+j)(m+\ell)\pi^2 \\ -(m+j)\pi \sin L \\ -(m+j)(m+\ell)\pi^2 \cos L \end{array} \right]$$

$$\mathbf{w}_7 = \left[\begin{array}{c} 0 \\ (m+j)\pi \sin L \\ -\sin J \sin L \\ -(m+j)\pi \cos J \sin L \\ \hline 0 \\ -(m+j)(m-\ell)\pi^2 \cos L \\ (m-\ell)\pi \cos L \sin J \\ (m+j)(m-\ell)\pi^2 \cos J \cos L \\ \hline 0 \\ 0 \\ 0 \\ 0 \\ \hline 0 \\ (m+j)(m-\ell)\pi^2 \\ -(m-\ell)\pi \sin J \\ -(m+j)(m-\ell)\pi^2 \cos J \end{array} \right] \quad \mathbf{w}_8 = \left[\begin{array}{c} 0 \\ -(m+j)\pi \sin L \\ \sin J \sin L \\ (m+j)\pi \cos J \sin L \\ \hline 0 \\ -(m+j)(m+\ell)\pi^2 \cos L \\ (m+\ell)\pi \cos L \sin J \\ (m+j)(m+\ell)\pi^2 \cos J \cos L \\ \hline 0 \\ 0 \\ 0 \\ 0 \\ \hline 0 \\ (m+j)(m+\ell)\pi^2 \\ -(m+\ell)\pi \sin J \\ -(m+j)(m+\ell)\pi^2 \cos J \end{array} \right].$$

Here $J = j\pi h$ and $L = \ell\pi h$. We have assumed here (without loss of generality and by analogy to our assumption in the one-dimensional analysis) that $k = 0$, providing $x_k = y_k = 0$ and $x_{k+1} = x_1 = y_{k+1} = y_1 = h$. The values d_k , $k = 1, 2, \dots, 8$, that produce the eigenvector \mathbf{x} in (7.14) (or, equivalently, give the vector \mathbf{v} that satisfies (7.18)), are seen to be

$$\begin{aligned} d_3 &= (m+j)(m+\ell), & d_1 &= \frac{\ell}{m} (1 - c_2^\ell) - c_2^\ell - d_3 \\ d_5 &= -(m+j)(m-\ell), & d_2 &= \frac{\ell}{m} (1 - c_2^\ell) - 1 - d_5 \\ d_4 &= -(m-j)(m+\ell), & d_7 &= \frac{\ell}{m} c_2^j (1 - c_2^\ell) - c_2^j c_2^\ell - d_4 \\ d_6 &= (m-j)(m-\ell), & d_8 &= \frac{\ell}{m} c_2^j (1 - c_2^\ell) - c_2^j - d_6. \end{aligned}$$

Here c_2^j is the value of c_2 corresponding to the eigenvalue δ_j of the one-dimensional problem while c_2^ℓ is the value of c_2 corresponding to the eigenvalue δ_ℓ of the one-dimensional

problem. With these choices of d_k , $k = 1, 2, \dots, 8$, the vector \mathbf{v} becomes

$$\mathbf{v} = \begin{bmatrix} 0 \\ \left[v_2^j \right] \left[(1 - c_2^\ell) \sin L \right] \\ \left[(1 - c_2^j) \sin J \right] \left[(1 - c_2^\ell) \sin L \right] \\ \left[-v_2^j \cos J \right] \left[(1 - c_2^\ell) \sin L \right] \\ \hline 0 \\ \left[v_2^j \right] \left[v_2^\ell \cos L \right] \\ \left[(1 - c_2^j) \sin J \right] \left[v_2^\ell \cos L \right] \\ \left[-v_2^j \cos J \right] \left[v_2^\ell \cos L \right] \\ \hline 0 \\ 0 \\ 0 \\ 0 \\ \hline 0 \\ \left[v_2^j \right] \left[-v_2^\ell \right] \\ \left[(1 - c_2^j) \sin J \right] \left[-v_2^\ell \right] \\ \left[-v_2^j \cos J \right] \left[-v_2^\ell \right] \end{bmatrix}, \quad (7.19)$$

where

$$\begin{aligned} v_2^j &= \pi \left[(m - j) + c_2^j (m + j) \right] \\ v_2^\ell &= \pi \left[(m - \ell) + c_2^\ell (m + \ell) \right]. \end{aligned}$$

(Note the analogy to (7.8) (for the one-dimensional problem).)

It is interesting to observe, given the definition of \mathbf{v} in (7.19), that the left sides of the equations (7.18) can be represented as linear combinations of

$$\begin{aligned} z_1^j &= [-\beta - \delta_j \chi] v_2 + \left[\alpha - \delta_j \left(\frac{1}{2} - \zeta \right) \right] v_3 + [-\gamma + \delta_j \psi] v_4 = 0 \\ z_2^j &= [\gamma - \delta_j \psi] v_2 + \left[-\alpha - \delta_j \left(\frac{1}{2} + \zeta \right) \right] v_3 + [\beta + \delta_j \chi] v_4 = 0 \\ z_1^\ell &= [-\beta - \delta_\ell \chi] v_2 + \left[\alpha - \delta_\ell \left(\frac{1}{2} - \zeta \right) \right] v_3 + [-\gamma + \delta_\ell \psi] v_4 = 0 \\ z_2^\ell &= [\gamma - \delta_\ell \psi] v_2 + \left[-\alpha - \delta_\ell \left(\frac{1}{2} + \zeta \right) \right] v_3 + [\beta + \delta_\ell \chi] v_4 = 0, \end{aligned}$$

which are simply the left sides of (7.6) and (7.7) under the simplifications which result in (7.8) for eigenvalues δ_j and δ_ℓ . Note that there are four separate equations in (7.18) of the form $\mathbf{m}^T \mathbf{v} = 0$; denote them as (a), (b), (c), (d) (from left to right). With this

notation, we find:

$$\begin{aligned}
[\text{left side of (7.18(a))}] &= \left[\left(\frac{1}{2} + \zeta \right) (1 - c_2^\ell) \sin L + \chi v_2^\ell \cos L + \psi v_2^\ell \right] z_1^j \\
&\quad + \left[\left(\frac{1}{2} - \zeta \right) (1 - c_2^j) \sin J + \psi v_2^j \cos J + \chi v_2^j \right] z_2^\ell \\
[\text{left side of (7.18(b))}] &= \left[\left(\frac{1}{2} + \zeta \right) (1 - c_2^\ell) \sin L + \chi v_2^\ell \cos L + \psi v_2^\ell \right] z_2^j \\
&\quad + \left[\left(\frac{1}{2} + \zeta \right) (1 - c_2^j) \sin J + \chi v_2^j \cos J + \psi v_2^j \right] z_2^\ell \\
[\text{left side of (7.18(c))}] &= \left[\left(\frac{1}{2} - \zeta \right) (1 - c_2^\ell) \sin L + \psi v_2^\ell \cos L + \chi v_2^\ell \right] z_1^j \\
&\quad + \left[\left(\frac{1}{2} - \zeta \right) (1 - c_2^j) \sin J + \psi v_2^j \cos J + \chi v_2^j \right] z_1^\ell \\
[\text{left side of (7.18(d))}] &= \left[\left(\frac{1}{2} - \zeta \right) (1 - c_2^\ell) \sin L + \psi v_2^\ell \cos L + \chi v_2^\ell \right] z_2^j + \\
&\quad \left[\left(\frac{1}{2} + \zeta \right) (1 - c_2^j) \sin J + \chi v_2^j \cos J + \psi v_2^j \right] z_1^\ell
\end{aligned}$$

Now that the eigenvalues and eigenvectors of the two-dimensional problem are fully described in terms of their one-dimensional counterparts, we may state our results:

Theorem 7.7. *Let $\Delta_j = \{\text{eigenvalues of the one-dimensional problem given in Theorem 7.4}\}$. We know, from Theorem 7.4, that Δ_j contains $2m$ elements. Let $\Delta_\ell = \Delta_j$. Then the $4m^2$ eigenvalues δ of (7.14), which are all negative, are given by all possible sums of the form*

$$\delta = \delta_j + \delta_\ell$$

where $\delta_j \in \Delta_j$ and $\delta_\ell \in \Delta_\ell$.

Using Corollary 7.6 and Theorem 7.7, we obtain

Corollary 7.8. *For the two-dimensional problem, the eigenvalue of $C^{-1}A$ of greatest magnitude is*

$$\delta = \frac{-48}{h^2(1 - 4\xi^2)}.$$

This result was used above in Equation (5.6).

REFERENCES

1. Brill, S. H. and Pinder, G. F. (in press). Eigenvalue Analysis of a Block Red-Black Gauss-Seidel Preconditioner Applied to the Hermite Collocation Discretization of Poisson's Equation. *Numerical Methods for Partial Differential Equations*.
2. van der Vorst, H. A. (1992). Bi-CGSTAB: A Fast and Smoothly Converging Variant of Bi-CG for the Solution of Nonsymmetric Linear Systems. *SIAM J. Sci. Stat. Comput.*, 13:631-644.
3. Dyksen, W. R. and Rice, J. R. (1986). The Importance of Scaling for the Hermite Bicubic Collocation Equations. *SIAM J. Sci. Stat. Comput.*, 7:707-719.
4. Lai, Y.-L., Hadjidimos, A., Houstis, E. N., and Rice, J. R. (1995). On the Iterative Solution of Hermite Collocation Equations. *SIAM J. Matrix Anal. Appl.*, 16:254-277.
5. Papatheodorou, T. S. (1983). Block AOR Iteration for Nonsymmetric Matrices. *Mathematics of Computation*, 41:511-525.

6. Golub, G. H. and van Loan, C. F. (1996). *Matrix Computations*, Third ed., The Johns Hopkins University Press, Baltimore, Maryland.
7. Prenter, P. M. and Russell, R. D. (1976). Orthogonal Collocation for Elliptic Partial Differential Equations. *SIAM J. Numer. Anal.*, 13:923–939.
8. Brill, S. H. (1998). Documentation for Collocation Software. RCGRD Publication #98-01, Research Center for Groundwater Remediation Design, University of Vermont, Burlington, Vermont.
9. Murtaugh, B. A. and Saunders, M. A. (1987). MINOS 5.1 User's Guide. Technical Report SOL 83-20R, Department of Operations Research, Stanford University, Stanford, California.
10. Johnson, L. W., Riess, R. D., and Arnold, J. T. (1993). *Introduction to Linear Algebra*, Third ed., Addison-Wesley Publishing Company, Reading, Massachusetts.
11. Bialecki, B., Fairweather, G., and Bennett, K. R. (1992). Fast Direct Solvers for Piecewise Hermite Bicubic Orthogonal Spline Collocation Equations. *SIAM J. Numer. Anal.*, 29:156–173.