Cyclic Reduction

Scott Cohen *

November 10, 1994

Abstract. The method of Block Cyclic Reduction (BCR) is described in the context of solving Poisson's equation with Dirichlet boundary conditions. The numerical instability of the original BCR algorithm is shown. A stable variant to BCR, credited to Oscar Buneman, is described and a proof of stabilty is given. Finally, the Sweet device for parallelization is presented.

1 Introduction

Differential equations are in the realm of continuous mathematics. Solution functions are defined on continuous sets such as an interval or a rectangle. In order to use a computer to help "solve" a differential equation, we must first discretize the problem. The idea is to program a computer to compute approximations to the solution function on some discrete, finite subset of the original domain. Such a subset is called a mesh. Once a mesh has been chosen, we can discretize the differential equation by replacing derivatives at mesh points by approximations involving nearby mesh points. This "finite difference" method of discretization results in a linear system whose unknowns are the approximations to the solution function on the mesh.

One way to solve a linear system is to use Gaussian elimination. The basic idea behind Gaussian elimination is to transform the system

$$\begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

into an equivalent upper triangular system

$$\begin{pmatrix} a'_{1,1} & a'_{1,2} & \cdots & a'_{1,n} \\ & a'_{2,2} & \cdots & a'_{2,n} \\ & & \ddots & \vdots \\ & & & a'_{n,n} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b'_1 \\ b'_2 \\ \vdots \\ b'_n \end{pmatrix}$$

*Thesis Advisor: Professor Gene Golub, Computer Science Department, Stanford University which can be easily solved. This transformation is iterative in the sense that the rth step clears the subdiagonal entries in the rth column of the coefficient matrix. Each step brings the system one column closer to the equivalent upper triangular system.

The coefficient matrix of a discretized problem is usually very sparse and well-structured (e.g. symmetric and tridiagonal). This is because a finite difference derivative approximation at a given point involves only a regular pattern of a few neighboring mesh points. Cyclic reduction is a recursive, as opposed to iterative, algorithm which takes advantage of the structure of such a coefficient matrix. The first step of cyclic reduction (sometimes called "odd-even" reduction) is to reduce the original linear system to one of the same form, but approximately half the size as the original. This strategy is possible because of some special form of the coefficient matrix. The solution to the reduced system allows one to solve the original system. We can repeatedly reduce the system size until we arrive at a 1×1 system (i.e. a single equation). The solution to the simplest system is used to solve the previously generated system, whose solution is used to solve the previously generated system, and so on, until all the unknowns in the original system have been found.

2 The Model Problem

Consider Poisson's equation on a rectangle, R, with Dirichlet boundary conditions

$$(1) \qquad \begin{aligned} \tfrac{\partial^2 u}{\partial x^2}(x,y) + \tfrac{\partial^2 u}{\partial y^2}(x,y) &= f(x,y) \qquad \text{for } (x,y) \in R, \\ u(x,y) &= g(x,y) \qquad \text{for } (x,y) \in \partial R, \end{aligned}$$

where f and g are given, $R = \{(x,y) : a < x < b,\ c < y < d\}$, and ∂R denotes the boundary of R. Introduce a grid (x_i, y_j) on $\bar{R} = R \cup \partial R$ with equal horizontal and vertical step length $h = \frac{b-a}{m+1} = \frac{d-c}{n+1}$ so that

$$\begin{aligned} x_i &= a + ih \qquad i = 0, \ldots, m+1, \\ y_j &= c + jh \qquad j = 0, \ldots, n+1. \end{aligned}$$

(For simplicity, we assume that is possible to divide $\bar{R}$ into squares of side h.) Using the central difference approximations

$$\begin{aligned} \frac{\partial^2 u}{\partial x^2}(x_i, y_j) &\approx \frac{u(x_{i-1}, y_j) - 2u(x_i, y_j) + u(x_{i+1}, y_j)}{h^2}, \\ \frac{\partial^2 u}{\partial y^2}(x_i, y_j) &\approx \frac{u(x_i, y_{j-1}) - 2u(x_i, y_j) + u(x_i, y_{j+1})}{h^2} \end{aligned}$$

the discretized form of (1) is

$$(2) \qquad \begin{aligned} u_{i,j-1} + (u_{i-1,j} - 4u_{i,j} + u_{i+1,j}) + u_{i,j+1} &= h^2 f_{i,j} \qquad \begin{matrix} i = 1, \ldots, m, \\ j = 1, \ldots, n \end{matrix} \\ u_{i,j} &= g_{i,j} \qquad (x_i, y_j) \in \partial R \end{aligned}$$

where $u_{i,j} \approx u(x_i, y_j)$, $f_{i,j} = f(x_i, y_j)$, and $g_{i,j} = g(x_i, y_j)$. We can rewrite (2) as

$$(3) \qquad \begin{aligned} \boldsymbol{u}_{j-1} + A\boldsymbol{u}_j + \boldsymbol{u}_{j+1} &= h^2 \boldsymbol{f}_j \qquad j = 1, \ldots, n, \\ \boldsymbol{u}_j &= \boldsymbol{g}_j \qquad j = 0, n+1 \end{aligned}$$

where A is the $m \times m$ tridiagonal, symmetric matrix

$$A = \begin{pmatrix} -4 & 1 & & & \\ 1 & -4 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & -4 & 1 \\ & & & 1 & -4 \end{pmatrix} \tag{4}$$

and

$$\boldsymbol{u}_j = \begin{pmatrix} u_{1,j} \\ u_{2,j} \\ \vdots \\ u_{m-1,j} \\ u_{m,j} \end{pmatrix}, \ \boldsymbol{f}_j = \begin{pmatrix} f_{1,j} \\ f_{2,j} \\ \vdots \\ f_{m-1,j} \\ f_{m,j} \end{pmatrix}, \ \boldsymbol{g}_j = \begin{pmatrix} g_{1,j} \\ g_{2,j} \\ \vdots \\ g_{m-1,j} \\ g_{m,j} \end{pmatrix}.$$

Finally, (3) can be written as an $n \times n$ symmetric, tridiagonal block system

$$\begin{bmatrix} A & I & & & \\ I & A & I & & \\ & \ddots & \ddots & \ddots & \\ & & I & A & I \\ & & & I & A \end{bmatrix} \begin{bmatrix} \boldsymbol{u}_1 \\ \boldsymbol{u}_2 \\ \vdots \\ \boldsymbol{u}_{n-1} \\ \boldsymbol{u}_n \end{bmatrix} = \begin{bmatrix} \boldsymbol{b}_1 \\ \boldsymbol{b}_2 \\ \vdots \\ \boldsymbol{b}_{n-1} \\ \boldsymbol{b}_n \end{bmatrix} \tag{5}$$

where

$$\begin{bmatrix} \boldsymbol{b}_1 \\ \boldsymbol{b}_2 \\ \vdots \\ \boldsymbol{b}_{n-1} \\ \boldsymbol{b}_n \end{bmatrix} = \begin{bmatrix} h^2 \boldsymbol{f}_1 - \boldsymbol{g}_0 \\ h^2 \boldsymbol{f}_2 \\ \vdots \\ h^2 \boldsymbol{f}_{n-1} \\ h^2 \boldsymbol{f}_n - \boldsymbol{g}_{n+1} \end{bmatrix}.$$

The goal is to solve (5) for $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n$; i.e., we want to solve for the $m \times n$ unknown approximations $u_{i,j}$, $i = 1, \ldots, m$, $j = 1, \ldots, n$ to the solution u of (1). The method of block cyclic reduction was originally developed by Gene Golub and R. W. Hockney in the mid 1960's to numerically solve Poisson's equation with periodic boundary conditions (see [6]). The system that arose in Hockney's analysis of the "48×48 Plasma Problem"is the same as (5), except that the block coefficient matrix also has I in the $(1, m)$ and $(m, 1)$ entries.

3 Some Preliminary Results

Before we describe the method of Block Cyclic Reduction, let us study the matrix A in (4) that arises in the discretization of our model problem. In particular, we shall find the eigenvalues of A. More generally, we seek the eigenvalues of the $m \times m$ tridiagonal, symmetric matrix T,

$$T = \begin{pmatrix} a & b & & & \\ b & a & b & & \\ & \ddots & \ddots & \ddots & \\ & & b & a & b \\ & & & b & a \end{pmatrix}. \tag{6}$$

We will use the following Lemmas.

Lemma 1 *The $m \times m$ matrix*

$$B = \begin{pmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ & & & -1 & 2 \end{pmatrix}, \tag{7}$$

has eigenvectors

$$\boldsymbol{v}_j = (\sin jh, \sin 2jh, \ldots, \sin mkh)^T, \tag{8}$$

wtih corresponding eigenvalues

$$\lambda_j = 2 - 2\cos jh, \tag{9}$$

where $j = 1, \ldots, m$ and $h = \pi/(m+1)$.

Proof. We need to check that $B\boldsymbol{v}_j = \lambda_j \boldsymbol{v}_j$. This can be done using the trigonometric formulae $\sin(x \pm y) = \sin x \cos y \pm \cos x \sin y$. The kth entry of $B\boldsymbol{v}_j$ is simply

$$(B\boldsymbol{v}_j)^k = -\sin(k-1)jh + 2\sin kjh - \sin(k+1)jh.$$

But the sum and difference identities for sin imply that

$$\begin{aligned} \sin(k-1)jh &= \sin kjh \cos jh - \cos kjh \sin jh, \\ \sin(k+1)jh &= \sin kjh \cos jh + \cos kjh \sin jh. \end{aligned}$$

Some simple alebgraic manipulation yields

$$(B\boldsymbol{v}_j)^k = (2 - 2\cos jh)\sin kjh = (\lambda_j \boldsymbol{v}_j)^k.$$

Q.E.D.

Lemma 2 *Suppose the $m \times m$ matrix M has eigenvectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m$, with corresponding eigenvalues $\lambda_1, \ldots, \lambda_m$. Then for any constants α and β, the matrix $\alpha M + \beta I$ has eigenvalues $\alpha\lambda_j + \beta$ with corresponding eigenvectors $\boldsymbol{v}_j$.*

Proof. The proof is straightforward:

$$\begin{aligned} (\alpha M + \beta I)\boldsymbol{v}_j &= \alpha M\boldsymbol{v}_j + \beta I\boldsymbol{v}_j \\ &= \alpha(\lambda_j \boldsymbol{v}_j) + \beta \boldsymbol{v}_j \\ (\alpha M + \beta I)\boldsymbol{v}_j &= (\alpha\lambda_j + \beta)\boldsymbol{v}_j. \end{aligned}$$

Q.E.D.

Our main result of the section follows from the previous lemmas.

Theorem 1 *The $m \times m$ tridiagonal, symmetric matrix*

$$T = \begin{pmatrix} a & b & & & \\ b & a & b & & \\ & \ddots & \ddots & \ddots & \\ & & b & a & b \\ & & & b & a \end{pmatrix},$$

has eigenvalues $\lambda_j = a + 2b\cos(j\pi/(m+1))$, for $j = 1, \ldots, m$.

Proof. Write $T = -bB + (a+2b)I$, with B as in lemma 1. By the previous two lemmas, the eigenvalues of T are (with $\alpha = -b$, $\beta = a + 2b$)

$$\begin{aligned} \lambda_j &= -b[2 - 2\cos(j\pi/(m+1))] + (a+2b) \\ \lambda_j &= a + 2b\cos(j\pi/(m+1)), \end{aligned}$$

Q.E.D.

Applying the previous theorem with $a = -4$ and $b = 1$, we see that the eigenvalues of the matrix A are

$$\lambda_j = -2(2 + \cos(j\pi/m+1)), \quad j = 1, \ldots, m. \tag{10}$$

From (10) we see that $\lambda_j < 0$. This implies that the matrix A is negative definite ($\boldsymbol{v}^T A \boldsymbol{v} < 0$ $\forall \boldsymbol{v} \neq \mathbf{o}$). Since A has m distinct eigenvalues $\lambda_1, \ldots, \lambda_m$, it must have m linearly independent eigenvectors. It follows that A is invertible. Since the invertible matrix A is both symmetric and negative definite, any linear system $A\boldsymbol{v} = \boldsymbol{w}$ may be solved numerically using Gaussian elimination without partial pivoting (cf [3], pp. 207-208). A somewhat easier way to arrive at this same conclusion is to notice that A is diagonally dominant (The absolute value of each diagonal element of A is greater than or equal to the sum of the absolute values of the other entries in its row, with strict inequality for at least one row.) Alternatively, a linear system $A\boldsymbol{v} = \boldsymbol{w}$ may be solved using scalar cyclic reduction as described by Bondeli and Gander in [1]. Equation (10) also shows that $2 < |\lambda_j| < 6 \ \forall j$. This fact will be used later during our discussion of the accuracy of the basic Block Cyclic Reduction algorithm which we now describe.

4 Block Cyclic Reduction

We shall illustrate the method of Block Cyclic Reduction (BCR) in the context of (5). The method proceeds in two stages: reduction and backsubstitution. During each step of the reduction stage, we eliminate approximately half the unknowns in the system. After $O(\log_2 n)$ reductions we are left with a 1×1 block system. After solving this system, the previously eliminated unknowns are computed by backsubstitution.

4.1 The Reduction Stage of BCR

Assume for simplicity that the system size n is of the form $n = 2^{k+1} - 1$. (The case for general n is described by Roland Sweet in [8].) Organize the n block equations of (5) into groups of the three:

$$\begin{array}{rccccccccc} (11) \quad & u_{j-2} & + & Au_{j-1} & + & u_j & & & = & b_{j-1} \\ & & & u_{j-1} & + & Au_j & + & u_{j+1} & = & b_j \\ & & & & & u_j & + & Au_{j+1} + u_{j+2} & = & b_{j+1} \end{array}$$

for each $j = 2, 4, \ldots, n-3, n-1$ (where we set $\boldsymbol{u}_0 = \boldsymbol{u}_{n+1} = \mathbf{o}$). If we multiply the middle equation of (11) by $-A$ and then add the three equations, we get

$$\boldsymbol{u}_{j-2} + (2I - A^2)\boldsymbol{u}_j + \boldsymbol{u}_{j+1} = \boldsymbol{b}_j^{(1)} \quad j = 2, 4, \ldots, n-3, n-1, \tag{12}$$

where $\boldsymbol{b}_j^{(1)} = \boldsymbol{b}_{j-1} - A\boldsymbol{b}_j + \boldsymbol{b}_{j+1}$. Note that (12) only involves $\boldsymbol{u}_j$'s with an even index. The "reduced" linear system corresponding to (5) is approximately half the size of (5) and is given by

$$\begin{bmatrix} A^{(1)} & I & & & \\ I & A^{(1)} & I & & \\ & \ddots & \ddots & \ddots & \\ & & I & A^{(1)} & I \\ & & & I & A^{(1)} \end{bmatrix} \begin{bmatrix} \boldsymbol{u}_2 \\ \boldsymbol{u}_4 \\ \vdots \\ \boldsymbol{u}_{n-3} \\ \boldsymbol{u}_{n-1} \end{bmatrix} = \begin{bmatrix} \boldsymbol{b}_2^{(1)} \\ \boldsymbol{b}_4^{(1)} \\ \vdots \\ \boldsymbol{b}_{n-3}^{(1)} \\ \boldsymbol{b}_{n-1}^{(1)} \end{bmatrix}, \tag{13}$$

where $A^{(1)} = 2I - A^2$. Note that (13) is an $\frac{n-1}{2} \times \frac{n-1}{2} = (2^k - 1) \times (2^k - 1)$ linear system of exactly the same form as (5). Therefore, we can apply the same reduction process to (13). Since the reduction step preserves the block symmetric tridiagonal form, we can apply the reduction step repeatedly. After k reductions ($n = 2^{k+1} - 1$) we will be left with a 1×1 block system

$$A^{(k)}\boldsymbol{u}_{2^k} = \boldsymbol{b}_{2^k}^{(k)}. \tag{14}$$

In general, during the rth reduction step we

- eliminate $\boldsymbol{u}_{1\cdot 2^{r-1}},\ \boldsymbol{u}_{3\cdot 2^{r-1}},\ \boldsymbol{u}_{5\cdot 2^{r-1}},\ \ldots,\ \boldsymbol{u}_{(2^{k+2-r}-1)\cdot 2^{r-1}}$
- retain $\boldsymbol{u}_{1\cdot 2^r},\ \boldsymbol{u}_{2\cdot 2^r},\ \boldsymbol{u}_{3\cdot 2^r},\ \ldots,\ \boldsymbol{u}_{(2^{k+1-r}-1)\cdot 2^r}$

The operator and RHS of the $(2^{k+1-r} - 1) \times (2^{k+1-r} - 1)$ reduced system created by the rth reduction step are

$$\begin{bmatrix} A^{(r)} & I & & & \\ I & A^{(r)} & I & & \\ & \ddots & \ddots & \ddots & \\ & & I & A^{(r)} & I \\ & & & I & A^{(r)} \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} \boldsymbol{b}_{1\cdot 2^r}^{(r)} \\ \boldsymbol{b}_{2\cdot 2^r}^{(r)} \\ \vdots \\ \boldsymbol{b}_{(2^{k+1-r}-2)\cdot 2^r}^{(r)} \\ \boldsymbol{b}_{(2^{k+1-r}-1)\cdot 2^r}^{(r)} \end{bmatrix},$$

respectively, where the $A^{(r)}$'s and $\boldsymbol{b}_j^{(r)}$'s are defined by the recurrence relations

$$A^{(0)} = A, \quad A^{(r+1)} = 2I - \left(A^{(r)}\right)^2 \tag{15}$$

$$\boldsymbol{b}_j^{(0)} = \boldsymbol{b}_j, \quad \boldsymbol{b}_j^{(r+1)} = \boldsymbol{b}_{j-2^r}^{(r)} - A^{(r)}\boldsymbol{b}_j^{(r)} + \boldsymbol{b}_{j+2^r}^{(r)} \tag{16}$$

Solving (14) is discussed later. Once (14) has been solved, the reduction phase of BCR is complete and the backsubstitution phase begins.

4.2 The Backsubstitution Stage of BCR

The system (5) is equivalent to the reduced system (13) along with the "eliminated" system

$$\tag{17} \begin{bmatrix} A & & & & \\ & A & & & \\ & & \ddots & & \\ & & & A & \\ & & & & A \end{bmatrix} \begin{bmatrix} \boldsymbol{u}_1 \\ \boldsymbol{u}_3 \\ \vdots \\ \boldsymbol{u}_{n-2} \\ \boldsymbol{u}_n \end{bmatrix} \begin{bmatrix} \boldsymbol{b}_1 - \boldsymbol{u}_2 \\ \boldsymbol{b}_3 - \boldsymbol{u}_2 - \boldsymbol{u}_4 \\ \vdots \\ \boldsymbol{b}_{n-2} - \boldsymbol{u}_{n-3} - \boldsymbol{u}_{n-1} \\ \boldsymbol{b}_n - \boldsymbol{u}_{n-1} \end{bmatrix}$$

The diagonal block system (17) follows by simply moving all even indexed $\boldsymbol{u}_i$'s from the LHS to the RHS of the odd-indexed block matrix equations (3). Each reduction step produces both a reduced system and a corresponding eliminated system to be solved during backsubstitution. In general, the rth reduction step produces the eliminated system which has the diagonal block coefficient matrix with $A^{(r-1)}$ on the diagonal and RHS

$$\tag{18} \begin{bmatrix} \boldsymbol{b}^{(r-1)}_{1\cdot 2^{r-1}} - \boldsymbol{u}_{1\cdot 2^r} \\ \boldsymbol{b}^{(r-1)}_{3\cdot 2^{r-1}} - \boldsymbol{u}_{1\cdot 2^r} - \boldsymbol{u}_{2\cdot 2^r} \\ \vdots \\ \boldsymbol{b}^{(r-1)}_{(2^{k+2-r}-3)\cdot 2^{r-1}} - \boldsymbol{u}_{(2^{k+1-r}-2)\cdot 2^r} - \boldsymbol{u}_{(2^{k+1-r}-1)\cdot 2^r} \\ \boldsymbol{b}^{(r-1)}_{(2^{k+2-r}-1)\cdot 2^{r-1}} - \boldsymbol{u}_{(2^{k+1-r}-1)\cdot 2^r} \end{bmatrix}.$$

After computing $\boldsymbol{u}_{2^k}$ from (14), backsubstitution begins (step 1) by solving the eliminated system produced by the kth reduction step

$$\begin{bmatrix} A^{(k-1)} & \\ & A^{(k-1)} \end{bmatrix} \begin{bmatrix} \boldsymbol{u}_{2^k - 2^{k-1}} \\ \boldsymbol{u}_{2^k + 2^{k-1}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{b}^{(k-1)}_{2^k - 2^{k-1}} - \boldsymbol{u}_{2^k} \\ \boldsymbol{b}^{(k-1)}_{2^k + 2^{k-1}} - \boldsymbol{u}_{2^k} \end{bmatrix}.$$

The vectors $\boldsymbol{u}_{2^k - 2^{k-1}}$, $\boldsymbol{u}_{2^k}$, and $\boldsymbol{u}_{2^k + 2^{k-1}}$ can then be used to solve the eliminated system produced by the $(k-1)st$ reduction step. Continuing in this way, all the $\boldsymbol{u}_i$ can be computed in k backsubstitution steps. The last step performed is to solve (17). During the rth backsubstitution step we solve for the unknowns $\boldsymbol{u}_j$ with

$$j = 2^{k-r} + l \cdot 2^{k-r+1} \quad l = 0, 1, 2, \ldots, 2^r - 1.$$

The backsubstitution stage requires the solution of systems of the form $A^{(r)}\boldsymbol{v} = \boldsymbol{w}$, where the RHS $\boldsymbol{w}$ involves the $\boldsymbol{b}_j^{(r)}$'s (see (18)). In fact, (14) is also of this form.

4.2.1 Solving the Linear Systems in BCR

We shall prove the following factor decomposition for $A^{(r)}$:

Theorem 2 *For $r = 1, 2, \ldots, k$,*

$$A^{(r)} = -\prod_{l=1}^{2^r} (A + 2\cos(\theta_l^{(r)}) I),$$

where

$$\theta_l^{(r)} = \frac{2l-1}{2^{r+1}}\pi.$$

Proof. From (15) we see that $A^{(r)}$ is a polynomial of degree 2^r in A. So for each $r = 0, 1, \ldots, k$, $\exists$ a polynomial $p_{2^r}(a)$ of degree 2^r such that

$$A^{(r)} = p_{2^r}(A),$$

where the polynomials $p_{2^r}(a)$ must satisfy

$$\begin{array}{rcl} p_{2^0}(a) & \equiv & a \\ p_{2^{r+1}}(a) & = & 2 - (p_{2^r}(a))^2. \end{array} \tag{19}$$

We will use the following lemma.

Lemma 3 *For* $r = 0, 1, 2, \ldots, k$,

$$p_{2^r}(-2\cos\theta) = -2\cos(2^r\theta).$$

Proof of Lemma. The proof proceeds by induction on r. The formula is true for $r = 0$ since $p_{2^0}(-2\cos\theta) = -2\cos\theta = -2cos(2^0\theta)$. Now suppose $p_{2^r}(-2\cos\theta) = -2\cos(2^r\theta)$ for some r such that $0 \le r \le k-1$. Then

$$\begin{array}{rcl} p_{2^{r+1}}(-2\cos\theta) & = & 2 - (p_{2^r}(-2\cos\theta))^2 \\ & = & 2 - (-2\cos(2^r\theta))^2 \\ & = & 2(1 - 2\cos^2(2^r\theta)) \\ & = & -2\cos(2 \cdot 2^r\theta) \\ p_{2^{r+1}}(-2\cos\theta) & = & -2\cos(2^{r+1}\theta). \end{array}$$

This completes the proof of the lemma.

By the lemma, $p_{2^r}(a)$ has 2^r distinct real zeros at

$$a_l = -2\cos(\frac{2l-1}{2^{r+1}}\pi) = -2\cos\theta_l^{(r)}, \quad l = 1, 2, 3, \ldots, 2^r.$$

In the non-trival cases $r \ge 1$, the coefficient of A^{2^r} in $p_{2^r}(A)$ is clearly -1. Therefore, for $r = 1, 2, \ldots, k$, we have

$$p_{2^r}(a) = -\prod_{l=1}^{2^r}(a + 2\cos\theta_l^{(r)})$$

and

$$A^{(r)} = p_{2^r}(A) = -\prod_{l=1}^{2^r}(A + 2\cos(\theta_l^{(r)})I).$$

Q.E.D.

Now define $A_l^{(r)} \equiv A + 2\cos(\theta_l^{(r)})I$ so that

$$A^{(r)} = -\prod_{l=1}^{2^r} A_l^{(r)}. \tag{20}$$

To solve $A^{(r)}\boldsymbol{v} = \boldsymbol{w}$, we solve

$$
\begin{aligned}
A_1^{(r)}\boldsymbol{v}_1 &= -\boldsymbol{w} && \Longrightarrow \boldsymbol{v}_1, \\
A_2^{(r)}\boldsymbol{v}_2 &= \boldsymbol{v}_1 && \Longrightarrow \boldsymbol{v}_2, \\
&\vdots \\
A_{2^r}^{(r)}\boldsymbol{v}_{2^r} &= \boldsymbol{v}_{2^r-1} && \Longrightarrow \boldsymbol{v}_{2^r} \equiv \boldsymbol{v}.
\end{aligned}
$$

The $m \times m$ invertible matrices $A_l^{(r)}$ are all diagonally dominant, so each of the 2^r above linear systems can be solved by Gaussian elimination without pivoting. Alternatively, these linear systems may be solved using scalar cyclic reduction because each $A_l^{(r)}$ is symmetric and tridiagonal.

4.2.2 Performing the Matrix-Vector Multiplications in BCR

To compute $A^{(r)}\boldsymbol{v}$, we could perform the 2^r matrix multiplications

$$
\begin{aligned}
-A_1^{(r)}\boldsymbol{v} &= \boldsymbol{y}_1 && \Longrightarrow \boldsymbol{y}_1, \\
A_2^{(r)}\boldsymbol{y}_1 &= \boldsymbol{y}_2 && \Longrightarrow \boldsymbol{y}_2, \\
&\vdots \\
A_{2^r}^{(r)}\boldsymbol{y}_{2^r-1} &= \boldsymbol{y}_{2^r} && \Longrightarrow \boldsymbol{y}_{2^r} \equiv A^{(r)}\boldsymbol{v}.
\end{aligned}
$$

It is also possible to take advantage of the recursive nature of the polynomials $p_{2^r}(a)$ in the proof of theorem 2.

Theorem 3 *Let* $\{\boldsymbol{\eta}_s\}_{s=0}^{2^r}$ *be defined by the recurrence relation*

$$
\begin{aligned}
&\boldsymbol{\eta}_0 = -2\boldsymbol{v},\ \boldsymbol{\eta}_1 = A\boldsymbol{v}, \\
&\boldsymbol{\eta}_s = -A\boldsymbol{\eta}_{s-1} - \boldsymbol{\eta}_{s-2}, \quad s = 2, 3, \ldots, 2^r.
\end{aligned}
\tag{21}
$$

Then $\boldsymbol{\eta}_{2^r} = A^{(r)}\boldsymbol{v}$.

Proof. We shall use some elementary trigonometric results.

Lemma 4 *The following identities for* $\cos(s\theta)$ *and* $\cosh(s\theta)$ *hold:*

$$
\begin{aligned}
\cos(s\theta) &= 2\cos\theta\cos((s-1)\theta) - \cos((s-2)\theta) \\
\cosh(s\theta) &= 2\cosh\theta\cosh((s-1)\theta) - \cosh((s-2)\theta)
\end{aligned}
\tag{22}
$$

Proof of Lemma. Using the sum and product formulae

$$
\begin{aligned}
\cos(x+y) &= \cos x\cos y - \sin x\sin y, \\
\cosh(x+y) &= \cosh x\cosh y + \sinh x\sinh y, \\
\sin x\sin y &= (\cos(x-y) - \cos(x+y))/2, \\
\sinh x\sinh y &= (\cosh(x+y) - \cosh(x-y))/2,
\end{aligned}
$$

we have

$$\begin{aligned}
\cos s\theta &= \cos(\theta + (s-1)\theta) \\
&= \cos\theta\cos((s-1)\theta) - \sin((s-1)\theta)\sin\theta \\
\cos s\theta &= \cos\theta\cos((s-1)\theta) - (\frac{\cos((s-2)\theta) - \cos s\theta}{2})
\end{aligned}$$

Solving the last equation for $\cos s\theta$ gives the desired result. The result for cosh is proven similarly.

For $s = 0, 1, \ldots, 2^r$ let

$$\tilde{p}_s(a) = \begin{cases} -2\cos s\theta & \text{if } |a| \le 2,\ a = -2\cos\theta \\ -2\cosh s\theta & \text{if } |a| > 2,\ a = -2\cosh\theta \end{cases}.$$

Then using lemma 4 we can quickly verify that for $s \ge 2$,

$$\tilde{p}_s(a) = -a\tilde{p}_{s-1}(a) - \tilde{p}_{s-2}(a) \quad \forall a.$$

It's also easy to check that $\tilde{p}_0(a) \equiv -2$ and $\tilde{p}_1(a) \equiv a$. Since $\tilde{p}_s(a) = p_s(a)\ \forall s = 1, 2, 4, \ldots, 2^r$, it follows that

$$A^r\boldsymbol{v} = p_{2^r}(A)\boldsymbol{v} = \tilde{p}_{2^r}(A)\boldsymbol{v}.$$

We can therefore compute $A^{(r)}\boldsymbol{v}$ via the sequence (21).

Q.E.D.

4.3 The Accuracy of BCR

Suppose we write a computer program to compute $A^{(r)}\boldsymbol{v}$ using the sequence (21) as suggested in Section 4.2.2. In the computation of each $\boldsymbol{\eta}_s$, there will be some roundoff error $\boldsymbol{\delta}_{s-1}$. The actual sequence computed is thus given by

$$\text{(23)} \qquad \begin{aligned} &\tilde{\boldsymbol{\eta}}_0 = -2\boldsymbol{v},\ \tilde{\boldsymbol{\eta}}_1 = A\boldsymbol{v} + \boldsymbol{\delta}_0 \\ &\tilde{\boldsymbol{\eta}}_s = -A\tilde{\boldsymbol{\eta}}_{s-1} - \tilde{\boldsymbol{\eta}}_{s-2} + \boldsymbol{\delta}_{s-1} \quad s = 2, 3, \ldots 2^r. \end{aligned}$$

Define the error sequence $\{\boldsymbol{\varepsilon}_s\}_{s=0}^{2^r}$ by $\boldsymbol{\varepsilon}_s = \tilde{\boldsymbol{\eta}}_s - \boldsymbol{\eta}_s$, so that we have the error recurrence

$$\text{(24)} \qquad \begin{aligned} &\boldsymbol{\varepsilon}_0 = \mathbf{o},\ \boldsymbol{\varepsilon}_1 = \boldsymbol{\delta}_0 \\ &\boldsymbol{\varepsilon}_s = -A\boldsymbol{\varepsilon}_{s-1} - \boldsymbol{\varepsilon}_{s-2} + \boldsymbol{\delta}_{s-1}, \quad s = 2, \ldots, 2^r. \end{aligned}$$

Since A is real and symmetric, $\exists$ a real, orthogonal matrix Q such that

$$Q^T A Q = \Lambda = \mathrm{diag}(\lambda_i)_{i=1}^m,$$

where $\lambda_1, \lambda_2, \ldots, \lambda_m$ are the eigenvalues of the matrix A. Multiplying each equation in (24) by Q^T, we get

$$\text{(25)} \qquad \begin{aligned} &\boldsymbol{\xi}_0 = \mathbf{o},\ \boldsymbol{\xi}_1 = \boldsymbol{\tau}_0 \\ &\boldsymbol{\xi}_s = -\Lambda\boldsymbol{\xi}_{s-1} - \boldsymbol{\xi}_{s-2} + \boldsymbol{\tau}_{s-1}, \quad s = 2, \ldots, 2^r, \end{aligned}$$

where $\boldsymbol{\xi}_l = Q^T \boldsymbol{\varepsilon}_l$ and $\boldsymbol{\tau}_l = Q^T \boldsymbol{\delta}_l$. If we denote the ith component of a vector $\boldsymbol{w}$ by $(\boldsymbol{w})^i$, then for each $j = 1, \ldots, m$, we can rewrite (25) as a scalar recurrence

$$(\boldsymbol{\xi}_0)^j = 0,\ (\boldsymbol{\xi}_1)^j = (\boldsymbol{\tau}_0)^j \\ (\boldsymbol{\xi}_s)^j = -\lambda_j (\boldsymbol{\xi}_{s-1})^j - (\boldsymbol{\xi}_{s-2})^j + (\boldsymbol{\tau}_{s-1})^j, \quad s = 2, \ldots, 2^r, \tag{26}$$

In order to avoid subscript and superscript confusion, let us work with the scalar recurrence

$$\alpha_0 = 0,\ \alpha_1 = \beta_0 \\ \alpha_s = -\lambda \alpha_{s-1} - \alpha_{s-2} + \beta_{s-1}. \tag{27}$$

Now define the recurrence relation

$$\gamma_0 = 0,\ \gamma_1 = 1 \\ \gamma_s = -\lambda \gamma_{s-1} - \gamma_{s-2}. \tag{28}$$

Then a simple induction shows that the solution to (27) is given by

$$\alpha_s = \sum_{i=0}^{s-1} \gamma_{s-i} \beta_i. \tag{29}$$

We now seek a closed form for γ_s.
Case 1. $|\lambda| < 2$.
Let $\lambda = -2\cos\phi$. Then the characteristic equation

$$x^2 - 2(\cos\phi)x + 1 = 0$$

has distinct roots

$$x_\pm = e^{\pm i\phi}.$$

The closed form is thus given by

$$\gamma_s = C e^{is\phi} + D e^{-is\phi},$$

where C and D are constants determinined by the initial conditions γ_0 and γ_1. Using these initial conditions, we find that $D = -C$ and $C = 1/(e^{i\phi} - e^{-i\phi})$. Therefore,

$$\begin{aligned} \gamma_s &= \frac{e^{is\phi} - e^{-is\phi}}{e^{i\phi} - e^{-i\phi}} \\ \gamma_s &= \frac{\sin(s\phi)}{\sin\phi} \end{aligned} \tag{30}$$

Case 2. $|\lambda| > 2$.
Let $\lambda = -2\cosh\phi$. Then characteristic equation

$$x^2 - 2(\cosh\phi)x + 1 = 0$$

has distinct roots

$$x_\pm = \cosh\phi \pm \sinh\phi.$$

Using the formulae

$$\cosh\phi = \tfrac{e^{\phi}+e^{-\phi}}{2}, \qquad \text{(31)}$$
$$\sinh\phi = \tfrac{e^{\phi}-e^{-\phi}}{2},$$

we have

$$x_{\pm} = e^{\pm\phi}.$$

Using these initial conditions γ_0 and γ_1 along with the hyperbolic trigonometry formulae (31), we find that the closed form solution to (28) is

$$\gamma_s = \frac{\sinh(s\phi)}{\sin\phi}. \qquad \text{(32)}$$

Case 3. $\lambda = -2$.
The characteristic equation

$$x^2 - 2x + 1 = 0$$

has a double root at $x = 1$. The closed form is thus given by

$$\gamma_s = C(1)^s + Ds(1)^s = C + Ds.$$

From the initial conditions, it follows easily that

$$\gamma_s \equiv s.$$

Case 4. $\lambda = 2$.
The characteristic equation

$$x^2 + 2x + 1 = 0$$

has a double root at $x = -1$. The closed form is thus given by

$$\gamma_s = C(-1)^s + Ds(-1)^s.$$

From the initial conditions, we find that

$$\gamma_s = s(-1)^{s+1}.$$

To summarize, the closed form solution to (28) is given by

$$\gamma_s = \begin{cases} \frac{\sin(s\phi)}{\sin\phi}, & \text{if } |\lambda| < 2,\ \lambda = -2\cos\phi \\ \frac{\sinh(s\phi)}{\sinh\phi}, & \text{if } |\lambda| > 2,\ \lambda = -2\cosh\phi \\ s & \text{if } \lambda = -2 \\ s(-1)^{s+1} & \text{if } \lambda = 2 \end{cases}.$$

Putting all the preceeding results in this section together,

$$(\boldsymbol{\xi}_s)^j = \sum_{i=0}^{s-1} (\gamma_{s-i})^j (\boldsymbol{\tau}_i)^j, \qquad \text{(33)}$$

where

$$(34) \qquad (\gamma_s)^j = \begin{cases} \frac{\sin(s\phi_j)}{\sin\phi_j}, & \text{if } |\lambda_j| < 2,\ \lambda_j = -2\cos\phi_j \\ \frac{\sinh(s\phi_j)}{\sinh\phi_j}, & \text{if } |\lambda_j| > 2,\ \lambda_j = -2\cosh\phi_j \\ s & \text{if } \lambda_j = -2 \\ s(-1)^{s+1} & \text{if } \lambda_j = 2 \end{cases}.$$

Note that for $|\lambda_j| < 2$, the sequence $\{(\gamma_s)^j\}_s$ remains bounded as $s \uparrow \infty$:

$$|(\gamma_s)^j| = \left|\frac{\sin s\phi_j}{\sin\phi_j}\right| \le \frac{1}{|\sin\phi_j|},$$

where $\lambda_j = -2\cos\phi_j$. For $|\lambda_j| = 2$, the sequence $\{(\gamma_s)^j\}_s$ is clearly unbounded as $s \uparrow \infty$. This is also true when $|\lambda_j| > 2$. In this case, $\{(\gamma_s)^j\}_s$ is a positive, strictly increasing sequence growing exponentially in s at the rate

$$\frac{d(\gamma_s)^j}{ds} = \frac{\phi_j}{e^{\phi_j} - e^{-\phi_j}}\left(e^{s\phi_j} + e^{-s\phi_j}\right).$$

Thus for $|\lambda_j| \ge 2$, the error component $(\boldsymbol{\varepsilon}_{2^r})^j = (Q\boldsymbol{\xi}_{2^r})^j$ may be unacceptably large with respect to the desired value $(\boldsymbol{\eta}_{2^r})^j \equiv (A^{(r)}\boldsymbol{v})^j$. As we saw in section 3, each of the eigenvalues λ_j, $j = 1, \ldots, m$ of A obtained in discretizing our model problem satisfies $|\lambda_j| > 2$. So computing $A^{(r)}\boldsymbol{v}$ via the recurrence relation (21) is a numerically unstable method for our model problem. The underlying difficulty in computing $A^{(r)}\boldsymbol{v}$ is that the matrix $A^{(r)}$ becomes very ill-conditioned as r increases. A more precise statement of this difficulty may be found in the landmark paper by Buzbee, Golub, and Nielson ([2], p. 648).

5 Buneman's Algorithm - A Stable Variant of BCR

5.1 Description of the Algorithm

Consider the jth equation of the first reduced system:

$$\boldsymbol{u}_{j-2} + (2I - A^2)\boldsymbol{u}_j + \boldsymbol{u}_{j+2} = \boldsymbol{b}_j^{(1)}.$$

The RHS is given by

$$\begin{aligned} \boldsymbol{b}_j^{(1)} &= \boldsymbol{b}_{j-1} + \boldsymbol{b}_{j+1} - A\boldsymbol{b}_j \\ &= A^{(1)}A^{-1}\boldsymbol{b}_j + \boldsymbol{b}_{j-1} + \boldsymbol{b}_{j+1} - 2A^{-1}\boldsymbol{b}_j. \end{aligned}$$

The last equality follows from the fact that

$$A^{(1)}A^{-1} = (2I - A^2)A^{-1} = 2A^{-1} - A.$$

If we let $\boldsymbol{p}_j^{(1)} = A^{-1}\boldsymbol{b}_j$ and $\boldsymbol{q}_j^{(1)} = \boldsymbol{b}_{j-1} + \boldsymbol{b}_{j+1} - 2A^{-1}\boldsymbol{b}_j$, then we can write

$$\boldsymbol{b}_j^{(1)} = A^{(1)}\boldsymbol{p}_j^{(1)} + \boldsymbol{q}_j^{(1)}.$$

In general, we have

$$
\begin{aligned}
\boldsymbol{b}_j^{(r+1)} &= \boldsymbol{b}_{j-2^r}^{(r)} + \boldsymbol{b}_{j+2^r}^{(r)} - A^{(r)}\boldsymbol{b}_j^{(r)} \\
&= A^{(r+1)}(A^{(r)})^{-1}\boldsymbol{b}_j^{(r)} + \boldsymbol{b}_{j-2^r}^{(r)} + \boldsymbol{b}_{j+2^r}^{(r)} - 2(A^{(r)})^{-1}\boldsymbol{b}_j^{(r)} \\
\boldsymbol{b}_j^{(r+1)} &= A^{(r+1)}\boldsymbol{p}_j^{(r+1)} + \boldsymbol{q}_j^{(r+1)},
\end{aligned} \tag{35}
$$

where

$$
\begin{aligned}
\boldsymbol{p}_j^{(r+1)} &= (A^{(r)})^{-1}\boldsymbol{b}_j^{(r)} \\
\boldsymbol{q}_j^{(r+1)} &= \boldsymbol{b}_{j-2^r}^{(r)} + \boldsymbol{b}_{j+2^r}^{(r)} - 2(A^{(r)})^{-1}\boldsymbol{b}_j^{(r)}.
\end{aligned} \tag{36}
$$

The reorganization of the computation of the $\boldsymbol{b}_j^{(r)}$'s is due to Oscar Buneman (cf [2]). We can get a recurrence relation for the $\boldsymbol{p}_j^{(r)}$'s and $\boldsymbol{q}_j^{(r)}$'s by substituting (36) into (35) and using the identity $(A^{(r)})^2 = 2I - A^{(r+1)}$:

$$
\begin{aligned}
\boldsymbol{b}_j^{(r+1)} &= \boldsymbol{b}_{j-2^r}^{(r)} + \boldsymbol{b}_{j+2^r}^{(r)} - A^{(r)}\boldsymbol{b}_j^{(r)} \\
&= A^{(r)}\boldsymbol{p}_{j-2^r}^{(r)} + \boldsymbol{q}_{j-2^r}^{(r)} + A^{(r)}\boldsymbol{p}_{j+2^r}^{(r)} + \boldsymbol{q}_{j+2^r}^{(r)} - \\
&\quad A^{(r)}(A^{(r)}\boldsymbol{p}_j^{(r)} + \boldsymbol{q}_j^{(r)}) \\
A^{(r+1)}\boldsymbol{p}_j^{(r+1)} + \boldsymbol{q}_j^{(r+1)} &= A^{(r+1)}(\boldsymbol{p}_j^{(r)} - (A^{(r)})^{-1}(\boldsymbol{p}_{j-2^r}^{(r)} + \boldsymbol{p}_{j+2^r}^{(r)} - \boldsymbol{q}_j^{(r)})) + \\
&\quad \boldsymbol{q}_{j-2^r}^{(r)} + \boldsymbol{q}_{j+2^r}^{(r)} - \\
&\quad 2(\boldsymbol{p}_j^{(r)} - (A^{(r)})^{-1}(\boldsymbol{p}_{j-2^r}^{(r)} + \boldsymbol{p}_{j+2^r}^{(r)} - \boldsymbol{q}_j^{(r)})).
\end{aligned}
$$

Comparing terms, we get

$$
\boldsymbol{p}_j^{(r+1)} = \boldsymbol{p}_j^{(r)} - (A^{(r)})^{-1}(\boldsymbol{p}_{j-2^r}^{(r)} + \boldsymbol{p}_{j+2^r}^{(r)} - \boldsymbol{q}_j^{(r)}), \tag{37}
$$

$$
\boldsymbol{q}_j^{(r+1)} = \boldsymbol{q}_{j-2^r}^{(r)} + \boldsymbol{q}_{j+2^r}^{(r)} - 2\boldsymbol{p}_j^{(r+1)}. \tag{38}
$$

We compute the $\boldsymbol{p}_j^{(r)}$'s and $\boldsymbol{q}_j^{(r)}$'s as follows:

1. For $j = 1, 2, \ldots, 2^{k+1} - 1$, initialize $\boldsymbol{p}_j^{(0)} = \mathbf{o}$, $\boldsymbol{q}_j^{(0)} = \boldsymbol{b}_j$.
2. For $r = 1, 2, \ldots, k$ (reduction step r)
   For $j = 1 \cdot 2^r, 2 \cdot 2^r, \ldots, 2^{k+1-r} \cdot 2^r \equiv 2^{k+1}$

   - Solve $A^{(r-1)}\boldsymbol{v} = \boldsymbol{p}_{j-2^{r-1}}^{(r-1)} + \boldsymbol{p}_{j+2^{r-1}}^{(r-1)} - \boldsymbol{q}_j^{(r-1)}$ for $\boldsymbol{v}$ using the method of section 4.2.1.
   - Compute $\boldsymbol{p}_j^{(r)} = \boldsymbol{p}_j^{(r-1)} - \boldsymbol{v}$.
   - Compute $\boldsymbol{q}_j^{(r)} = \boldsymbol{q}_{j-2^{r-1}}^{(r-1)} + \boldsymbol{q}_{j+2^{r-1}}^{(r-1)} - 2\boldsymbol{p}_j^{(r)}$.

To do the backsubstitution steps, note that

$$
\begin{aligned}
\boldsymbol{u}_{j-2^r} + A^{(r)}\boldsymbol{u}_j + \boldsymbol{u}_{j+2^r} &= A^{(r)}\boldsymbol{p}_j^{(r)} + \boldsymbol{q}_j^{(r)} \\
A^{(r)}(\boldsymbol{u}_j - \boldsymbol{p}_j^{(r)}) &= \boldsymbol{q}_j^{(r)} - \boldsymbol{u}_{j-2^r} - \boldsymbol{u}_{j+2^r}.
\end{aligned}
$$

So to compute $\boldsymbol{u}_j$ from $\boldsymbol{u}_{j-2^r}$ and $\boldsymbol{u}_{j+2^r}$, first solve

$$A^{(r)}\boldsymbol{v} = \boldsymbol{q}_j^{(r)} - \boldsymbol{u}_{j-2^r} - \boldsymbol{u}_{j+2^r}$$

for $\boldsymbol{v}$, and then compute

$$\boldsymbol{u}_j = \boldsymbol{p}_j^{(r)} + \boldsymbol{v}.$$

5.2 Proof of Stability

Our first step in showing the stability of the Buneman algorithm is to write out the $\boldsymbol{p}_j^{(r)}$'s and $\boldsymbol{q}_j^{(r)}$'s in terms of the unknown approximations $\boldsymbol{u}_j$.

Theorem 4 *For the Buneman algorithm, the following relations hold:*

$$\begin{aligned} \boldsymbol{p}_j^{(r)} &= \boldsymbol{u}_j + \boldsymbol{z}_j^{(r)}, \\ \boldsymbol{q}_j^{(r)} &= \boldsymbol{u}_{j-2^r} + \boldsymbol{u}_{j+2^r} - A^{(r)}\boldsymbol{z}_j^{(r)}, \end{aligned}$$

where

(40) $$\boldsymbol{z}_j^{(r)} = (-1)^{r+1}S^{(r)}\left(\sum_{k=1}^{2^{r-1}}\left(\boldsymbol{u}_{j-(2k-1)} + \boldsymbol{u}_{j+(2k-1)}\right)\right)$$

and

(41) $$S^{(r)} = (A^{(r-1)}A^{(r-2)}\ldots A^{(0)})^{-1}.$$

Proof. The proof is by induction on r. First let us check the base case $r = 1$.

$$\begin{aligned} \boldsymbol{p}_j^{(1)} &= \boldsymbol{p}_j^{(0)} - (A^{(0)})^{-1}(\boldsymbol{p}_{j-1}^{(0)} + \boldsymbol{p}_{j+1}^{(0)} - \boldsymbol{q}_j^{(0)}) \\ &= A^{-1}\boldsymbol{q}_j^{(0)} = A^{-1}\boldsymbol{b}_j \\ &= A^{-1}(\boldsymbol{u}_{j-1} + A\boldsymbol{u}_j + \boldsymbol{u}_{j+1}) \\ &= \boldsymbol{u}_j + A^{-1}(\boldsymbol{u}_{j-1} + \boldsymbol{u}_{j+1}) \\ \boldsymbol{p}_j^{(1)} &= \boldsymbol{u}_j + \boldsymbol{z}_j^{(1)} \end{aligned}$$

We also have

$$\begin{aligned} \boldsymbol{q}_j^{(1)} &= \boldsymbol{q}_{j-1}^{(0)} + \boldsymbol{q}_{j+1}^{(0)} - 2\boldsymbol{p}_j^{(1)} \\ &= \boldsymbol{b}_{j-1} + \boldsymbol{b}_{j+1} - 2(\boldsymbol{u}_j + \boldsymbol{z}_j^{(1)}) \\ &= (\boldsymbol{u}_{j-2} + A\boldsymbol{u}_{j-1} + \boldsymbol{u}_j) + (\boldsymbol{u}_j + A\boldsymbol{u}_{j+1} + \boldsymbol{u}_{j+2}) \\ &\quad -2(\boldsymbol{u}_j + \boldsymbol{z}_j^{(1)}) \\ &= \boldsymbol{u}_{j-2} + \boldsymbol{u}_{j+2} + A^2A^{-1}(\boldsymbol{u}_{j-1} + \boldsymbol{u}_{j+1}) \\ &\quad -2\boldsymbol{z}_j^{(1)} \\ &= \boldsymbol{u}_{j-2} + \boldsymbol{u}_{j+2} + (A^2 - 2I)\boldsymbol{z}_j^{(1)} \\ \boldsymbol{q}_j^{(1)} &= \boldsymbol{u}_{j-2} + \boldsymbol{u}_{j+2} - A^{(1)}\boldsymbol{z}_j^{(1)} \end{aligned}$$

Now we perform the inductive step and show that the theorem statement holds for the $r+1$ case assuming the r case as the inductive hypothesis. Again, first we do the $\boldsymbol{p}_j^{(r)}$ case followed by the $\boldsymbol{q}_j^{(r)}$ case. We start in both cases with the basic recursive definitions for $\boldsymbol{p}_j^{(r+1)}$ and $\boldsymbol{q}_j^{(r+1)}$:

$$\begin{aligned}
\boldsymbol{p}_j^{(r+1)} &= \boldsymbol{p}_j^{(r)} - (A^{(r)})^{-1}(\boldsymbol{p}_{j-2^r}^{(r)} + \boldsymbol{p}_{j+2^r}^{(r)} - \boldsymbol{q}_j^{(r)}), \\
&= (\boldsymbol{u}_j + \boldsymbol{z}_j^{(r)}) - (A^{(r)})^{-1}(\boldsymbol{u}_{j-2^r} + \boldsymbol{z}_{j-2^r}^{(r)} \\
&\quad + \boldsymbol{u}_{j+2^r} + \boldsymbol{z}_{j+2^r}^{(r)} \\
&\quad - (\boldsymbol{u}_{j-2^r} + \boldsymbol{u}_{j+2^r} - A^{(r)}\boldsymbol{z}_j^{(r)})) \\
&= \boldsymbol{u}_j - (A^{(r)})^{-1}(\boldsymbol{z}_{j-2^r}^{(r)} + \boldsymbol{z}_{j+2^r}^{(r)}) \\
\boldsymbol{p}_j^{(r+1)} &= \boldsymbol{u}_j + \boldsymbol{z}_j^{(r+1)}.
\end{aligned}$$

The last step follows from the fact that

$$\boldsymbol{z}_j^{(r+1)} = -(A^{(r)})^{-1}(\boldsymbol{z}_{j-2^r}^{(r)} + \boldsymbol{z}_{j+2^r}^{(r)}). \tag{42}$$

Equation (42) from the definitions of $\boldsymbol{z}_j^{(r)}$ and $S^{(r)}$ given in equations (40) and (41), respectively.

For the $\boldsymbol{q}_j^{(r)}$ case, we have

$$\begin{aligned}
\boldsymbol{q}_j^{(r+1)} &= \boldsymbol{q}_{j-2^r}^{(r)} + \boldsymbol{q}_{j+2^r}^{(r)} - 2\boldsymbol{p}_j^{(r+1)} \\
&= (\boldsymbol{u}_{j-2^r-2^r} + \boldsymbol{u}_{j-2^r+2^r} - A^{(r)}\boldsymbol{z}_{j-2^r}^{(r)}) \\
&\quad + (\boldsymbol{u}_{j+2^r-2^r} + \boldsymbol{u}_{j+2^r+2^r} - A^{(r)}\boldsymbol{z}_{j+2^r}^{(r)}) \\
&\quad - 2(\boldsymbol{u}_j + \boldsymbol{z}_j^{(r+1)}) \\
&= \boldsymbol{u}_{j-2^{r+1}} + \boldsymbol{u}_{j+2^{r+1}} \\
&\quad (A^{(r)})^2(A^{(r)})^{-1}(\boldsymbol{z}_{j-2^r}^{(r)} + \boldsymbol{z}_{j+2^r}^{(r)}) - 2\boldsymbol{z}_j^{(r+1)} \\
&= \boldsymbol{u}_{j-2^{r+1}} + \boldsymbol{u}_{j+2^{r+1}} - (2I - (A^{(r)})^2))\boldsymbol{z}_j^{(r+1)} \\
\boldsymbol{q}_j^{(r+1)} &= \boldsymbol{u}_{j-2^{r+1}} + \boldsymbol{u}_{j+2^{r+1}} - A^{(r+1)}\boldsymbol{z}_j^{(r+1)}
\end{aligned}$$

Q.E.D.

The next two theorems will allow us to estimate the two quantities

- $\left\| \boldsymbol{p}_j^{(r)} - \boldsymbol{u}_j \right\|_2$
- $\left\| \boldsymbol{q}_j^{(r)} - (\boldsymbol{u}_{j-2^r} + \boldsymbol{u}_{j+2^r}) \right\|_2$.

Theorem 5 $\left\| S^{(r)} \right\|_2 < e^{(-2^r-1)\theta_1}$, *where*

$$\theta_i = \cosh^{-1}(-\lambda_i/2)$$

and $\{\lambda_i\}_{i=1}^m$ *are the eigenvalues of* A.

Proof. Working from the definition of $S^{(r)}$,

$$\begin{aligned}\left\|S^{(r)}\right\|_2 &= \left\|\prod_{j=0}^{r-1}(A^{(j)})^{-1}\right\|_2 \\ &\leq \prod_{j=0}^{r-1}\left\|(A^{(j)})^{-1}\right\|_2 \\ &\leq \prod_{j=0}^{r-1}\max_{\{\lambda_i\}}\left(\frac{1}{|p_{2^j}(\lambda_i)|}\right),\end{aligned}$$

where the p_{2^j} are as defined in section 4.2.1. Using a proof similar to the proof of Lemma 3, we can show that

$$p_{2^j}(\lambda_i) = -2\cosh(2^j\theta_i).$$

It follows that

$$\left\|S^{(r)}\right\|_2 \leq \frac{1}{2^r}\prod_{j=0}^{r-1}\max_{\{\theta_i\}}\left(\frac{1}{\cosh(2^j\theta_i)}\right). \tag{43}$$

But we know that $\lambda_i = -2(2-\cos(i\pi/m+1))$. So

$$\theta_i = \cosh^{-1}(2-\cos(i\pi/m+1))$$

and

$$1 < \theta_1 < \theta_2 < \cdots < \theta_m.$$

Therefore,

$$\max_{\{\theta_i\}}\left(\frac{1}{\cosh(2^j\theta_i)}\right) = \frac{1}{\cosh(2^j\theta_1)}.$$

Substituting the previous result into (43), we get

$$\left\|S^{(r)}\right\|_2 \leq \frac{1}{2^r}\prod_{j=0}^{r-1}\left(\frac{1}{\cosh(2^j\theta_1)}\right).$$

Writing out $\cosh(2^j\theta_1)$ in terms of exponentials we find that

$$\begin{aligned}\left\|S^{(r)}\right\|_2 &\leq \prod_{j=0}^{r-1}\frac{1}{e^{2^j\theta_1}+e^{-2^j\theta_1}} \\ &\leq \prod_{j=0}^{r-1}\frac{e^{-2^j\theta_1}}{1+e^{-2^{j+1}\theta_1}} \\ &< \prod_{j=0}^{r-1}e^{-2^j\theta_1} \\ \left\|S^{(r)}\right\|_2 &< e^{-(2^r-1)\theta_1}\end{aligned}$$

Q.E.D.

Theorem 6 $\left\|A^{(r)}S^{(r)}\right\|_2 < 2e^{\theta_m}$.

Proof. Using essentially the same reasoning as in the proof of theorem 5 we can bound $\left\|A^{(r)}S^{(r)}\right\|_2$ by

$$\left\|A^{(r)}S^{(r)}\right\|_2 \leq \frac{1}{2^{r-1}} \max_{\{\theta_i\}} \left\{ \left(\prod_{j=0}^{r-1} \frac{1}{\cosh(2^j\theta_i)} \right) \cdot \cosh(2^r\theta_i) \right\} \tag{44}$$

Now we want to show that the maximum in (44) occurs at θ_m. Define the functions $\Psi_l(\theta)$ by

$$\Psi_l(\theta) = \left(\prod_{j=0}^{l-1} \frac{1}{\cosh(2^j\theta)} \right) \cdot \cosh(2^l\theta).$$

We claim that Ψ_l is an increasing function on $\theta \in (0, \infty)$ for each $l \geq 1$. This can be shown by induction on l. For the base case $l = 1$, we have

$$\begin{aligned} \Psi_1(\theta) &= \frac{\cosh 2\theta}{\cosh \theta} \\ &= \frac{e^{2\theta} + e^{-2\theta}}{e^{\theta} + e^{-\theta}} \end{aligned}$$

Then some simple calculus and algebraic manipulation yields

$$\Psi_1'(\theta) = \frac{(e^{3\theta} - e^{-3\theta}) + 3(e^{\theta} - e^{-\theta})}{(e^{\theta} + e^{-\theta})^2}.$$

Hence,

$$\Psi_1'(\theta) > 0 \quad \forall \theta > 0.$$

So Ψ_1 is increasing on $\theta \in (0, \infty)$. Given the induction hypothesis that Ψ_l is increasing on $\theta \in (0, \infty)$ for a fixed $l \geq 1$, we must show that Ψ_{l+1} is also increasing on this interval. Using the trigonometric identity for $\cosh(2x)$,

$$\cosh(2^{l+1}\theta) = \cosh(2 \cdot 2^l\theta) = \cosh^2(2^l\theta) + \sinh^2(2^l\theta).$$

Thus we can write

$$\Psi_{l+1}(\theta) = \left(1 + \frac{\sinh^2(2^r\theta)}{\cosh^2(2^r\theta)}\right) \Psi_l(\theta). \tag{45}$$

It's an easy matter to check that

$$\frac{d}{dx}\frac{\sinh^2 x}{\cosh^2 x} = \frac{2\cosh x \sinh x}{\cosh^4 x} > 0 \quad \forall x > 0.$$

This means that the first factor in the RHS of (45) is increasing on $\theta \in (0, \infty)$. By the inductive hypothesis, the second factor is also increasing on this interval. Therefore Ψ_{l+1} is increasing on $\theta \in (0, \infty)$ and the inductive proof is complete. Since

$$1 < \theta_1 < \theta_2 < \cdots < \theta_m,$$

we can conclude that

$$\max_{\{\theta_i\}} \left\{ \left(\prod_{j=0}^{r-1} \frac{1}{\cosh(2^j\theta_i)} \right) \cdot \cosh(2^r\theta_i) \right\} = \left(\prod_{j=0}^{r-1} \frac{1}{\cosh(2^j\theta_m)} \right) \cdot \cosh(2^r\theta_m) \tag{46}$$

Substituting (46) into (44):

$$\begin{aligned} \left\| A^{(r)} S^{(r)} \right\|_2 &\le \frac{1}{2^{r-1}} \left(\prod_{j=0}^{r-1} \frac{1}{\cosh(2^j\theta_m)} \right) \cdot \cosh(2^r\theta_m) \\ &\le \left(\prod_{j=0}^{r-1} \frac{1}{e^{2^j\theta_m} + e^{-2^j\theta_m}} \right) \left(e^{2^r\theta_m} + e^{-2^r\theta_m} \right) \\ &< e^{-(2^r-1)\theta_m} \left(e^{2^r\theta_m} + e^{-2^r\theta_m} \right) \\ &< e^{\theta_m}(1 + e^{-2^{r+1}\theta_m}) \\ \left\| A^{(r)} S^{(r)} \right\|_2 &< 2e^{\theta_m} \end{aligned}$$

Q.E.D.

Using theorems 4 and 5, and the triangle inequality, we see that

$$\left\| \boldsymbol{p}_j^{(r)} - \boldsymbol{u}_j \right\|_2 \le \mu e^{-(2^r-1)\theta_1}, \tag{47}$$

where

$$\mu = \sum_{j=1}^{n} \|\boldsymbol{u}_j\|_2 .$$

Therefore $\boldsymbol{p}_j^{(r)}$ is a good approximation to $\boldsymbol{u}_j$ for large values of r. Using theorems 4 and 6, we also see that

$$\left\| \boldsymbol{q}_j^{(r)} - (\boldsymbol{u}_{j-2^r} + \boldsymbol{u}_{j+2^r}) \right\|_2 < 2\mu e^{\theta_m}. \tag{48}$$

So $\left\| \boldsymbol{q}_j^{(r)} \right\|_2$ remains bounded during calculation. Therefore, Buneman's algorithm gives numerically stable results when used to solve our model problem.

6 Identifying Parallelism in Buneman's Algorithm

The discussion in this section closely follows Sweet's discussion ([9], pp.763-764). We want to determine the computations in the Buneman variant of BCR that may be performed in parallel. During the rth reduction step we compute the vector pairs $(\boldsymbol{p}_j^{(r)}, \boldsymbol{q}_j^{(r)})$ for $j = 1 \cdot \Delta, 2 \cdot \Delta, \ldots, (2^{k+1-r} - 1) \cdot \Delta$, where $\Delta = 2^r$. These computations are done according to the instructions given in the section 5.1. Note that $\boldsymbol{p}_j^{(r)}$ depends only on $\boldsymbol{p}_j^{(r-1)}$, $\boldsymbol{p}_{j-2^{r-1}}^{(r-1)}$, $\boldsymbol{p}_{j+2^{r-1}}^{(r-1)}$, and $\boldsymbol{q}_j^{(r-1)}$ which were computed during the $(r-1)st$ reduction step. Similarly, $\boldsymbol{q}_j^{(r)}$ depends only on $\boldsymbol{p}_j^{(r)}$ and the vectors $\boldsymbol{q}_{j-2^{r-1}}^{(r-1)}$, $\boldsymbol{q}_{j+2^{r-1}}^{(r-1)}$ which were also computed during the previous reduction step. Thus, during the rth reduction step we can compute each of the $2^{k+1-r} - 1$ vector pairs $(\boldsymbol{p}_j^{(r)}, \boldsymbol{q}_j^{(r)})$ in parallel. Since the number of vector pairs to compute is roughly halved with each successive reduction, the reduction stage benefits less and less from parallelization as the reduction stage proceeds. Furthermore, at reduction step r we must solve 2^{r-1} tridiagonal linear systems as outlined in section 4.2.1 in order to compute each pair $(\boldsymbol{p}_j^{(r)}, \boldsymbol{q}_j^{(r)})$. Thus, in addition to the collapse in parallelism, the amount of work to done to compute a vector pair doubles with each successive reduction step.

In contrast, the backsubstitution phase begins very serial and becomes more and more parallel as backsubstitution proceeds. During the rth backsubstitution step, we solve 2^{k-r} tridiagonal systems in order to compute each of the unknowns $\boldsymbol{u}_{2^{k-r}+l\cdot 2^{k-r+1}}$ for $l = 0, 1, \ldots, 2^r - 1$. Since all eliminated systems are block diagonal, we can compute these $\boldsymbol{u}_j$'s in parallel using 2^r processors. Thus the parallelism of the backsubstitution phase doubles, while the work to compute a particular $\boldsymbol{u}_j$ is halved with each successive backsubstitution step.

7 The Sweet Device for Parallelization

The collapse of parallelism during the reduction stage could be avoided if we could solve a system such as $A^{(r)}\boldsymbol{v} = \boldsymbol{w}$ in a parallel manner instead of the serial manner suggested in section 4.2.1. We want to compute

$$\boldsymbol{v} = (A^{(r)})^{-1}\boldsymbol{w}.$$

We know that $A^{(r)}$ is a polynomial of degree 2^r in A and that it is factored as in (20). The brilliant idea of Roland Sweet in [9] to obtain parallelism is to use the partial fraction expansion of $(A^{(r)})^{-1}$. We need the following lemma from complex analysis:

Lemma 5 *Let $p(x)$ and $q(x)$ be relatively prime polynomials with $\deg q < \deg p = n$. Suppose the roots, $\sigma_1, \sigma_2, \ldots, \sigma_n$ of $p(x)$ are distinct. Then*

$$\frac{q(x)}{p(x)} = \sum_{l=1}^{n} \frac{c_l}{x - \sigma_l},$$

where

$$c_l = \frac{q(\sigma_l)}{p'(\sigma_l)}.$$

Applying the lemma with $q(x) \equiv 1$ and $p(x) = p_{2^r}(x)$, the solution of $A^{(r)}\boldsymbol{v} = \boldsymbol{w}$ can be written as

$$\boldsymbol{v} = \sum_{l=1}^{2^r} c_l^{(r)} (A + 2\theta_l^{(r)} I)^{-1} \boldsymbol{w},$$

where

$$c_l^{(r)} = \frac{1}{p'_{2^r}(-2\theta_l^{(r)})}.$$

If we let $\boldsymbol{v}_l^{(r)} = c_l^{(r)}(A + 2\theta_l^{(r)})^{-1}\boldsymbol{w}$, then

$$\boldsymbol{v} = \sum_{l=1}^{2^r} \boldsymbol{v}_l^{(r)}.$$

The vectors $\boldsymbol{v}_l^{(r)}$, $l = 1, \ldots, 2^r$ may be computed in parallel using 2^r processors, where the lth processor solves the tridiagonal linear system

$$(A + 2\theta_l^{(r)} I)\boldsymbol{v}_l^{(r)} = \boldsymbol{w}.$$

We recover $\boldsymbol{v}$ by summing the output of the 2^r processors. During step r of the reduction phase in Buneman's algorithm, there are $2^{k+1-r} - 1$ independent systems to solve involving $A^{(r-1)}$. We can thus perform the rth reduction step using

$$(2^{k+1-r} - 1)(2^{r-1}) = 2^k - 2^{r-1}$$

parallel processes. During step r of the backsubstitution phase, we must solve 2^r independent systems involving $A^{(k-r)}$. Thus we can perform the rth backsubstitution step of Buneman's algorithm using

$$2^r \cdot 2^{k-r} = 2^k$$

parallel processes.

8 Conclusion

The method of cyclic reduction can be applied to Poisson's equation with boundary conditions other than Dirichlet. For example, we may specify the normal derivative $\partial u/\partial \eta$ on ∂R. This is called Neumann boundary conditions. We may also have periodic boundary conditions $u(a, y) = u(b, y), u(x, c) = u(x, d)$. When Poisson's equation with these two types of boundary conditions is discretized, we get coefficient matrices (M_N for the Neumann case and M_P for the periodic case):

$$M_N = \begin{bmatrix} A_N & 2I & & & \\ I & A_N & I & & \\ & \ddots & \ddots & \ddots & \\ & & I & A_N & I \\ & & & 2I & A_N \end{bmatrix}$$

and

$$M_P = \begin{bmatrix} A_P & I & & & I \\ I & A_P & I & & \\ & \ddots & \ddots & \ddots & \\ & & I & A_P & I \\ I & & & I & A_P \end{bmatrix},$$

where

$$A_N = \begin{pmatrix} -4 & 2 & & & \\ 1 & -4 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & -4 & 1 \\ & & & 2 & -4 \end{pmatrix}$$

and

$$A_P = \begin{pmatrix} -4 & 1 & & & 1 \\ 1 & -4 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & -4 & 1 \\ 1 & & & 1 & -4 \end{pmatrix}.$$

Systems with coefficient matrix M_N or M_P may be reduced to systems half the original size using the same ideas as in section 4.1. An excellent source for the details of solving Poisson's equation with Dirichlet, Neumann, or periodic boundary conditions is [2].

The Poisson problem (i.e., finding a function whose laplacian is equal to a given function in the interior of a region with certain boundary conditions) is not the only problem for which cyclic reduction may be applied. The method also applies when we discretize (using finite difference approximations for the derivatives) elliptic partial differential equations of the form

$$a(x)\frac{\partial^2 u}{\partial x^2} + b(x)\frac{\partial u}{\partial x} + c(x)u + \frac{\partial^2 u}{\partial y^2} = f. \tag{49}$$

For example, with Dirichlet boundary conditions, the discretization of (49) using central difference approximations for the derivatives produces a linear system with coefficient matrix of the form

$$\begin{bmatrix} A & -I & & & \\ -I & A & -I & & \\ & \ddots & \ddots & \ddots & \\ & & -I & A & -I \\ & & & -I & A \end{bmatrix}. \tag{50}$$

If we multiply this system with coefficient matrix (50) by -1, then we are get an equivalent system of exactly the same form as (5) with $-A$ in place of A. Thus, with Dirichlet boundary conditions, the underlying linear system for the problems (5) and (49) are the same. This is also true when Neumann or periodic boundary conditions have been specified.

In this paper we dealt only with the case when the size of the coefficient matrix is of the form $n = 2^{k+1} - 1$. In [8], Roland Sweet provides extensions to the Buneman variant of BCR for equations of the form (49) with Dirichlet, Neumann, or periodic boundary conditions. For each of these boundary conditions, the extended algorithm presented by Sweet reduces to the Buneman algorithm in the case $n = 2^{k+1} - 1$.

References

[1] S. Bondeli and W. Gander. Cyclic reduction for special tridiagonal systems. Unpublished.

[2] B. L. Buzbee, G. H. Golub, and C. W. Nielson. On direct methods for solving poisson's equations. *SIAM J.Numer. Anal.*, 7(4):627–656, December 1970.

[3] Lars Elden and Linde Wittmeyer-Koch. *Numerical Analysis: An Introduction*. Academic Press, inc., 1990.

[4] Oliver Ernst, Margot Gerritsen, and Hamid Samandari. Parallel cyclic reduction on the cm-2. Final Report from CS315B at Stanford.

[5] G. H. Golub and J. H. Ortega. *Scientific Computing and Differential Equations*. Academic Press, inc., 1992.

[6] R. W. Hockney. A fast direct solution of poisson's equation using fourier analysis. *Journal of the Association for Computing Machinery*, 12(1):95–113, January 1965.

[7] J. Stoer and R. Burlirsch. *Introduction to Numerical Analysis*. Springer-Verlag New York, inc., 1980.

[8] Roland Sweet. A cyclic reduction algorithm for solving block tridiagonal systems of arbitrary dimension. *SIAM J. Numer. Anal.*, 14(4):706–720, September 1977.

[9] Roland Sweet. A parallel and vector variant of the cyclic reduction algorithm. *SIAM J. Sci. Stat. Comput.*, 9(4):761–765, July 1988.