

# On the Solution of Block-Tridiagonal Systems Arising from Certain Finite-Difference Equations\*

By J. M. Varah

**Abstract.** We consider the solution of the linear systems arising from certain implicit finite-difference approximations to systems of linear differential equations. In particular, we consider those schemes which lead to matrices of block-tridiagonal form. There are two common methods for solving such equations: using a block-tridiagonal factorization (blocksolve), or treating the matrix as a band matrix (bandsolve).

First, we discuss conditions for ensuring the numerical stability of the block-tridiagonal factorization for general matrices of this form. Then, we compare the two methods for general block-tridiagonal matrices (including matrices arising from the Crank-Nicolson scheme for systems of parabolic equations) and for a more specialized block-tridiagonal matrix which arises from schemes of H. B. Keller for systems of two-point boundary value problems and parabolic equations.

**1. Introduction.** The approximation of a time-dependent system of linear partial differential equations in one space variable by an implicit finite-difference scheme leads to a linear algebraic system with a matrix of a definite band structure. This system must be solved at each time-step and thus it is worthwhile to solve this system as efficiently as possible. When only one equation is involved, this is fairly straightforward: For example, approximating the heat equation

$$u_t = u_{xx}, \quad 0 \leq x \leq 1, \quad 0 \leq t \leq T,$$

$$u(x, 0) = f(x), \quad u(0, t) = g_1(t), \quad u(1, t) = g_2(t)$$

by the Crank-Nicolson scheme

$$\left( I - \frac{k}{2} D_+ D_- \right) v_i^{m+1} = \left( I + \frac{k}{2} D_+ D_- \right) v_i^m$$

leads to a tridiagonal matrix equation, for which the direct solution, using Gaussian elimination, is well known (see e.g. [3, p. 56]). Here

$$v_i^m = v(j\Delta x, m\Delta t),$$

$$k = \Delta t = \text{time-step},$$

$$h = \Delta x = \text{space-step } (h = 1/(n - 1)),$$

$D_+, D_-$  are the usual forward and backward spatial difference operators.

However, when the problem in question is a *system* of differential equations, or if it is put in system form for solution, the structure of the accompanying matrix

---

Received February 14, 1972.

AMS 1970 subject classifications. Primary 65N20; Secondary 65F05.

Key words and phrases. Solution of difference equations, band matrices, block-tridiagonal matrices.

\* Supported by National Research Council of Canada grant #A8240.

is more complicated. In particular, let us assume that the unknown grid functions  $\mathbf{v}_i^m = ((v_1)_i^m, (v_2)_i^m, \dots, (v_p)_i^m)$  are numbered for the matrix equation so that components are consecutive; then the matrix has a block-banded structure with  $p \times p$  blocks instead of single elements.

*Note.* This ordering may not in fact be optimal in the sense of number of operations for solution (see George [9] for some very nice new results on optimal orderings). However, the ordering given here is certainly natural, easy to program, and leads to a bandwidth as small as possible.

For many of the basic schemes used for solving such systems, the matrix is in fact block-tridiagonal, i.e.,

$$(1) \quad A = \begin{pmatrix} B_1 & C_1 & & & \\ A_2 & B_2 & \cdot & & \\ & \cdot & \cdot & \cdot & \\ & & \cdot & \cdot & C_{n-1} \\ & & & \cdot & \\ & & & & A_n & B_n \end{pmatrix}$$

with each  $A_i, B_i, C_i$  square. For example, with the parabolic system (here  $\mathbf{u} = (u_1, \dots, u_p)$ )

$$(2) \quad \begin{aligned} \mathbf{u}_t &= P(x, t)\mathbf{u}_{xx}, & 0 \leq x \leq 1, 0 \leq t \leq T, \\ \mathbf{u}(x, 0) &= \mathbf{f}(x), \\ \left. \begin{aligned} Q_1\mathbf{u}_x + Q_0\mathbf{u} &= g^I(t) \\ Q_2\mathbf{u} &= g^{II}(t) \end{aligned} \right\} & \text{at } x = 0, \\ \left. \begin{aligned} R_1\mathbf{u}_x + R_0\mathbf{u} &= g^{III}(t) \\ R_2\mathbf{u} &= g^{IV}(t) \end{aligned} \right\} & \text{at } x = 1, \end{aligned}$$

the Crank-Nicolson scheme

$$\left( I - \frac{k}{2} D_+ P\left(x_i - \frac{h}{2}, t_{m+1}\right) D_- \right) \mathbf{v}_i^{m+1} = \left( I + \frac{k}{2} D_+ P\left(x_i - \frac{h}{2}, t_m\right) D_- \right) \mathbf{v}_i^m$$

with any kind of discrete boundary conditions can be expressed in the form (1) with each block of order  $p$ .

Recently, Keller has proposed new schemes for a first-order system of two-point boundary value problems [4], and for parabolic systems [5]; both of these can be expressed as block-tridiagonal systems (see Section 3). Other examples of block-tridiagonal schemes can be found in Richtmyer and Morton [6].

There are two common methods for solving the system (1); treating it as a band matrix, or using its block-tridiagonal structure. Either way, for an accurate solution, we need to ensure that the intermediate quantities do not become too large (we call this *numerical stability*). In the case of a band matrix, this can be done by partial pivoting; however, this increases the amount of work necessary, so it is useful to know when pivoting is not required. Sufficient conditions for this (namely  $A$  diagonally dominant or  $A$  symmetric positive definite) were first given by Wilkinson [8] and are also discussed in Wendroff [7].

In Section 2, we discuss the same problem for a general block-tridiagonal system, and, in Section 3, compare the work involved in the band and block methods for the particular schemes mentioned above.

**2. Solution of General Block-Tridiagonal Systems.** The direct solution of (1) by block-Gaussian elimination without pivoting is well known (see for example Isaacson and Keller [3, p. 59]); the block-triangular decomposition can be expressed as

$$(3) \quad \begin{pmatrix} B_1 & C_1 & & & & \\ A_2 & B_2 & C_2 & & & \\ & \cdot & \cdot & \cdot & & \\ & & \cdot & \cdot & \cdot & \\ & & & \cdot & \cdot & C_{n-1} \\ & & & & A_n & B_n \end{pmatrix} = \begin{pmatrix} I & & & & & \\ L_2 & I & & & & \\ & L_3 & I & & & \\ & & \cdot & \cdot & \cdot & \\ & & & \cdot & \cdot & \\ & & & & L_n & I \end{pmatrix} \begin{pmatrix} U_1 & C_1 & & & & \\ & U_2 & C_2 & & & \\ & & \cdot & \cdot & \cdot & \\ & & & \cdot & \cdot & \\ & & & & \cdot & C_{n-1} \\ & & & & & U_n \end{pmatrix}$$

or  $A = LU$ . The recurrence for the  $L_i, U_i$  is

$$(4) \quad \left. \begin{aligned} U_1 &= B_1 \\ L_i &= A_i U_{i-1}^{-1} \\ U_i &= B_i - L_i C_{i-1} \end{aligned} \right\} \quad i = 2, \dots, n.$$

For numerical stability, we need a bound on the pivotal growth, i.e.,  $\max(\|L_i\|, \|U_i\|) \leq K$  for some norm. In the scalar case, this is guaranteed if  $A$  is diagonally dominant; we wish to show that a similar result holds here.

*Definition 1.* The matrix  $A$  of (1) is *block diagonally dominant* with respect to the matrix norm  $\|\cdot\|$  if

$$(5) \quad \|B_i^{-1}\| (\|A_i\| + \|C_i\|) \leq 1, \quad i = 1, \dots, n.$$

This concept has been discussed often in the literature; see, for example, Feingold and Varga [1].

A necessary condition for this, of course, is that the diagonal blocks  $B_i$  be non-singular, and we assume this in what follows. We also assume  $C_i \neq 0$ , for, if some  $C_i = 0$ , the system would be decoupled.

**THEOREM 2.1.** *If  $A$  is block-diagonally dominant, then the block-triangular decomposition (3) is numerically stable and, in fact,*

$$\|L_i\| \leq \|A_i\|/\|C_{i-1}\|, \quad \|U_i\| \leq \|B_i\| + \|A_i\|.$$

*Proof.* From (4), it is clear we need only a bound for  $\|U_i^{-1}\|$ . We show by induction that  $\|U_i^{-1}\| \leq 1/\|C_i\|$ . This is clear for  $i = 1$  from (4) and (5). Assume for  $i - 1$ ; then

$$\begin{aligned} U_i &= B_i - A_i U_{i-1}^{-1} C_{i-1} \\ &= B_i(I - B_i^{-1} A_i U_{i-1}^{-1} C_{i-1}) = B_i(I - R). \end{aligned}$$

Now  $\|R\| \leq 1 - \|B_i^{-1}\| \|C_i\|$  from (5) and the induction hypothesis. Thus,

$$\|U_i^{-1}\| \leq \|B_i^{-1}\| / (\|B_i^{-1}\| \|C_i\|) = 1/\|C_i\|.$$

Now the above bounds on  $\|L_i\|, \|U_i\|$  follow easily from (4). Q.E.D.

If we assume  $A$  is block-diagonally dominant by *columns* rather than rows, a similar analysis gives  $\|L_i\| \leq 1, \|U_i\| \leq \|B_i\| + \|C_i\|$ , which is the obvious generalization of Theorem 5, p. 56 of [3].

We can, in fact, be more general than this by including block-diagonal scaling via  $DAE$ ,

$$D = \begin{bmatrix} D_1 & & & \\ & \ddots & & \\ & & \ddots & \\ & & & D_n \end{bmatrix}, \quad E = \begin{bmatrix} E_1 & & & \\ & \ddots & & \\ & & \ddots & \\ & & & E_n \end{bmatrix}.$$

If  $A = LU$  is the block-triangular factorization (3), then the corresponding factorization for  $DAE$  is  $(DL)(UE)$ , and we can apply Theorem 2.1 to this matrix. Thus, we have (as in the scalar case)

**COROLLARY.** *The block-triangular decomposition (3) is numerically stable if  $DAE$  is block-diagonally dominant for some block-diagonal  $D, E$  with  $\|D\|, \|D^{-1}\|, \|E\|, \|E^{-1}\|$  bounded.*

One fruitful choice of  $D, E$  is  $D_i = B_i^{-1}, E_i = e_i I$ . This leads to the following:

**THEOREM 2.2.** *Let  $A$  be as in (1) and define  $\alpha_i = (\|B_i^{-1}C_i\| \|B_{i+1}^{-1}A_{i+1}\|)^{1/2}$  (assume  $\alpha_i \neq 0$ ). Then the block-tridiagonal factorization (3) is numerically stable if the matrix*

$$S = \begin{bmatrix} 1 & \alpha_1 & & & \\ \alpha_1 & 1 & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & \alpha_{n-1} \\ & & & & \alpha_{n-1} & 1 \end{bmatrix}$$

*is positive semidefinite.*

Before proving this, we should remark that this theorem includes the result of Richtmyer and Morton [6, p. 275 ff.] as a special case, namely with  $\alpha_i \equiv \alpha$  and the condition  $\alpha \leq \frac{1}{2}$ .

*Proof.* With  $D_i = B_i^{-1}, E_i = e_i I$ ,  $DAE$  is block-diagonally dominant if the following set of inequalities holds:

$$\begin{aligned} e_2 \|B_1^{-1}C_1\| &\leq e_1, \\ e_{i-1} \|B_i^{-1}A_i\| + e_{i+1} \|B_i^{-1}C_i\| &\leq e_i, \quad i = 2, \dots, n-1, \\ e_{n-1} \|B_n^{-1}A_n\| &\leq e_n. \end{aligned}$$

With  $\delta_i = \|B_i^{-1}C_i\|, \gamma_i = \|B_i^{-1}A_i\|$ , this holds if there is a positive vector  $\mathbf{e}$  so that

$$P_{\mathbf{e}} \equiv \begin{bmatrix} 1 & -\delta_1 & & & & \\ -\gamma_2 & 1 & -\delta_2 & & & \\ & \cdot & \cdot & \cdot & & \\ & & \cdot & \cdot & \cdot & \\ & & & \cdot & -\delta_{n-1} & \\ & & & -\gamma_n & 1 & \end{bmatrix} \mathbf{e} \geq 0.$$

But this holds whenever it holds for the symmetrized matrix

$$DPD^{-1} = \begin{pmatrix} 1 & -\alpha_1 & & & & \\ -\alpha_1 & 1 & -\alpha_2 & & & \\ & \cdot & \cdot & \cdot & & \\ & & \cdot & \cdot & \cdot & \\ & & & \cdot & \cdot & -\alpha_{n-1} \\ -\alpha_{n-1} & & & & & 1 \end{pmatrix} \equiv I - R.$$

Now if  $(I - R)$  is positive definite, we know from its  $M$ -matrix properties (see, for example, Householder [2, p. 58]) that there is a positive vector  $\mathbf{e}$  so that  $(I - R)\mathbf{e} > 0$ . And if it is singular but positive semidefinite, we take for  $\mathbf{e}$  the null vector of  $(I - R)$ ; it must have all positive components as it is the Perron vector of  $R$  ( $R$  is irreducible because no  $\alpha_i = 0$ ). So a sufficient condition for numerical stability is  $(I - R)$  positive semidefinite. But since  $R$ 's eigenvalues occur in  $(+, -)$  pairs, this holds if and only if  $I + R = S$  is positive semidefinite. Q.E.D.

**3. Solution of Systems: Blocksolve vs. Bandsolve.** We consider two kinds of systems: the general block-tridiagonal matrix (1) and a specialized form arising from the solution of first-order systems. The general form (1) arises for example in the solution of (2) via the Crank-Nicolson scheme. For a constant coefficient problem (i.e.,  $P = \text{constant}$ ), we have  $A_i = C_i = -\lambda P/2$ ,  $B_i = I + \lambda P$  where  $\lambda = k/h^2$ , and since  $P$  is positive definite, we can apply Theorem 2.2 to show the numerical stability of the block-tridiagonal factorization. For variable coefficients, we will certainly have this stability for  $h$  small enough. Treating the system as a band matrix, we may have to row pivot the first  $p$  equations (these are the boundary conditions) but this is not necessary with the rest because of the positive definiteness.

So we compare operations of bandsolve and blocksolve for general systems (1) without pivoting (the blocks are  $p \times p$  and there are  $n$  such blocks in each row and column).

(a) *Bandsolve:*

(i) *LU decomposition:* Each set of  $p$  equations requires

$$(2p)(2p - 1) + (2p - 1)(2p - 2) + \dots + (p + 1)p = p(7p^2 - 1)/3 \text{ multiplications.}$$

(ii) *Backsolution:* Both forward and backward substitutions require

$$(p + 1) + \dots + 2p = p(3p + 1)/2 \text{ operations per } p \text{ equations.}$$

This gives a total of  $np^2(7p/3 + 3)$  multiplications (carrying only  $O(np^2)$  terms).

(b) *Blocksolve:* (See (4).)

(i) *Block LU decomposition:* For  $i = 2, \dots, n$ , we have

$$\left. \begin{array}{l} LU \text{ decomposition of } U_{i-1}: p^3/3 \text{ operations} \\ \text{solution of } L_i(U_{i-1}) = A_i: p^3 \text{ operations} \\ \text{formation of } U_i = B_i - L_i C_{i-1}: p^3 \text{ operations} \end{array} \right\} \text{total } \frac{7np^3}{3}.$$

(ii) Backsolution:

$$\begin{aligned} Ly &= f \text{ takes } np^2, \\ Ux &= y \text{ takes } 2np^2 \text{ (using the saved } LU \text{ decomposition of each } U_i). \end{aligned}$$

This gives a total of  $np^2(7p/3 + 3)$ , the same as for bandsolve.

*Note.* In a private communication, Gene Golub has shown that, for a symmetric positive definite block-tridiagonal matrix  $A$ , using the Cholesky decomposition once can reduce the operation count for blocksolve to  $5np^3/3 + O(np^2)$ .

Now consider the more specialized matrix form:

(6)

$$\begin{array}{ccccc}
 & & & & \\
 & & & & \\
 \boxed{F_0} & | & & & \\
 & | & & & \\
 \boxed{F_1} & | & \boxed{G_1} & | & \text{--- ---} \\
 & | & | & | & \\
 \text{--- ---} & | & \boxed{F_2} & | & \boxed{G_2} \\
 & | & | & | & \\
 & | & & & \boxed{G_0} \\
 & | & & & \\
 & | & & & \\
 & | & & & 
 \end{array}$$

Here all blocks are  $p \times p$  except  $F_0$  ( $q \times p$ ) and  $G_0$  ( $p - q \times p$ ). When split up so all blocks are  $p \times p$  (as indicated by the dotted lines), it becomes a block-tridiagonal system.

This form arises in solving the general first-order system

$$u'(x) = K(x)u(x) + f(x), \quad 0 \leq x \leq 1,$$

with separated boundary conditions

$$F_0u(0) = \alpha, \quad G_0u(1) = \beta$$

by the midpoint rule

(7)

$$\frac{v_i - v_{i-1}}{h_i} = K(x_{i-1/2})\left(\frac{v_i + v_{i-1}}{2}\right) + f(x_{i-1/2}).$$

See Keller [4] for details. Here  $F_0$  has  $q$  rows (assumed linearly independent) and  $G_0$  has  $p - q$  rows.

This matrix also comes up in solving the parabolic system (2) by Keller's box scheme [5]. Since the problem is converted to a first-order system (with unknowns  $u, u_x$ ), the matrix blocks are of order  $2p$  now, and  $q = p$ .

We again wish to compare the solution of (6) by band and block methods. First, we examine the stability or pivoting problem for the midpoint rule and box scheme. For the midpoint rule, from (7), we have

$$F_i = -I - h_i K(x_{i-1/2})/2, \quad G_i = I - h_i K(x_{i-1/2})/2.$$

Thus, for  $\max, h_i = h$  small enough, these are nonsingular. However, as a band

matrix there is no guarantee that the diagonal pivots will be large, and pivoting may be necessary; for example, the first column of  $F_0$  may be all zero necessitating an initial row interchange using a row of  $F_1$ . This, of course, increases the bandwidth and the work involved. We include below a calculation of the work required for bandsolve using pivoting. Similarly, for the block-tridiagonal decomposition, the diagonal blocks  $B_i$  may be singular. Thus,  $B_i$  will be singular if the  $q$  rows of  $F_0$  and the first  $(p - q)$  rows of  $F_1$  are dependent. But if so, we can interchange rows of  $F_1$  (and  $G_1$ ) so  $B_i$  is nonsingular; moreover, this preserves the matrix form (6). We can apply a similar “pivoting” for each such diagonal block as we proceed, and it seems clear that this will be numerically stable for  $h$  small enough. Indeed, if we assume the boundary matrix  $F_0$  is scaled so  $F_0 = (H_0 I)$ , then for  $h$  small enough

$$U_1 = B_1 = \left( \begin{array}{c|c} H_0 & I \\ \hline I & 0 \end{array} \right) + O(h);$$

and since

$$A_2 = \left( \begin{array}{c|c} 0 & I \\ \hline 0 & 0 \end{array} \right) + O(h),$$

we have

$$U_2 = \left( \begin{array}{c|c} H_0 & I \\ \hline I & 0 \end{array} \right) + O(h),$$

so the diagonal blocks  $U_i$  stay well conditioned.

Now consider the box scheme (see Keller [5]). For simplicity, we will only consider it applied to the heat equation  $u_t = u_{xx}$  with  $\alpha_0 u + \alpha_1 u_x$  prescribed at  $x = 0$ . This can be written as a first-order system:  $u_x = v$ ,  $v_x = u_t$ , and then differenced as

$$(8) \quad D_{-x}u_i^n = \frac{1}{2}(v_i^n + v_{i-1}^n), \quad D_{-x}(v_i^n + v_{i-1}^n) = D_{-t}(u_i^n + u_{i-1}^n).$$

This leads to an implicit equation for  $(u_i, v_i)$  on the  $n$ th time line with matrix

$$B = \left( \begin{array}{cc|cc|c} \alpha_0 & \alpha_1 & 0 & 0 & \\ \hline 1 & h/2 & -1 & h/2 & 0 \\ \hline h/k & 1 & h/k & -1 & \cdot \\ \hline 0 & 0 & 1 & h/2 & \cdot \\ \hline & & & & \cdot \\ \hline 0 & & & & \cdot \\ \hline & & & & \cdot \end{array} \right).$$

Clearly, we cannot guarantee the stability of bandsolve without pivoting, and, similarly, the block-tridiagonal factorization may be unstable ( $\alpha_1 = 0$  implies  $\|U_1^{-1}\| = O(1/h)$ , and, in fact, then each  $\|U_i\| = O(1/h)$ ,  $i = 2, 3, \dots, n$ ). However, if we

interchange the equations in (8), we obtain the matrix

$$B' = \begin{pmatrix} \alpha_0 & \alpha_1 & 0 & 0 & \mathbf{0} \\ h/k & 1 & h/k & -1 & \\ \hline 1 & h/2 & -1 & h/2 & \cdot \\ 0 & 0 & h/k & 1 & \cdot \\ \hline \mathbf{0} & & \cdot & \cdot & \cdot \end{pmatrix}.$$

Treating this as a band matrix, it is easy to see the bandsolve decomposition is numerically stable with a slight restriction on  $h$  and  $k$ . Indeed, because of the sign pattern, if our  $LU$  decomposition proceeds so that the third pivot  $B'_{33}$  is  $< 0$  independent of  $h$ , then from that point on the pivots will slowly increase in magnitude (by  $O(h)$  on each step), giving a stable decomposition. But this third pivot is, omitting  $O(h)$  terms,  $(-2 + 1/(1 - h\alpha_1/k\alpha_0))$ . If  $\alpha_0\alpha_1 \leq 0$ , we have  $-2 \leq B'_{33} \leq -1$  with no restriction on  $h$  and  $k$ ; if  $\alpha_0\alpha_1 > 0$ , we have  $-3 \leq B'_{33} \leq -2$  for  $k/h \leq \alpha_1/2\alpha_0$  and  $-1 \leq B'_{33} \leq -\frac{1}{2}$  for  $k/h \geq 3\alpha_1/\alpha_0$ .

Treating  $B'$  as a block-tridiagonal matrix, for numerical stability we again need to ensure  $k/h \neq \alpha_1/\alpha_0$ , so  $U_i$  is nonsingular. In fact if we carry through the algorithm, we see that

$$U_i = \begin{pmatrix} h/ke_i - 1 & -1/e_i \\ h/k & 1 \end{pmatrix} + O(h), \quad i = 2, 3, \dots,$$

where  $e_2 = \alpha_0/\alpha_1 - h/k$ ,  $e_{i+1} = e_i - 2(h/k)e_i^2$ ,  $i = 2, 3, \dots$ . For stability, we need only to keep the  $e_i$  away from zero. Note that  $e_{i+1} < e_i$  and, thus, we have a stable decomposition for all  $h$  and  $k$  if  $\alpha_0\alpha_1 \leq 0$ . If  $\alpha_0\alpha_1 > 0$ , it is stable if  $e_2 < 0$  (for example, if  $k/h \leq \frac{1}{2}\alpha_1/\alpha_0$ ); if  $\alpha_1/\alpha_0 \ll 1$  so this would be unduly restrictive, we have stability if we choose  $k/h$  so that  $e_2 > 0$ ,  $e_3 < 0$ .

Although we have only discussed the box scheme for the heat equation, it is clear that similar considerations could be applied to the box scheme for more general parabolic equations and systems.

Now we consider the operation counts for the matrix (6) using both bandsolve and blocksolve. For this, we make use of the zero patterns, but make no assumptions about the nonzero elements for the sake of generality. Thus, the operation times could be cut somewhat for special nonzero elements; for example, in the box scheme one might take special account of the ones appearing (this is done by Keller [5]).

(a) *Bandsolve for (6) (no pivoting).*

(i) *LU decomposition:* Each set of  $p$  rows takes

$$\begin{aligned} & p(p + q - 1) + (p - 1)(p + q - 2) + \dots + (p - q + 1)p \\ & \quad + (2p - q)(p - 1) + (2p - q - 1)(p - 2) + \dots + (p + 1)q \\ & = np(5p^2 + 3pq - 3q^2 - 3p + 6q - 2)/6 \quad \text{for all } np \text{ rows.} \end{aligned}$$



(ii) Backsolution:

$Ly = f$  takes for each set of  $p$  rows (except first)

$$[(p + 1) + \cdots + (p + q)] + [(q + 1) + \cdots + p],$$

$Ux = y$  takes for each set of  $p$  rows (except last)

$$[p + (p - 1) + \cdots + (p - q + 1)] + [(2p - q) + \cdots + (p + 1)]$$

giving a total of  $np(2p + 1)$  multiplications.

(b) *Bandsolve for (6) (with pivoting)*. Now  $L$  will have the same form, but  $U$  could look like the upper triangular part of a general block-tridiagonal matrix.

(i) *LU decomposition*: Each set of  $p$  rows takes

$$2p(p + q - 1) + (2p - 1)(p + q - 2) + \cdots + (p + 1)q \\ = np(5p^2 + 9pq - 3p + 3q - 2)/6 \quad \text{for all } np \text{ rows.}$$

(ii) Backsolution:

$Ly = f$  same as (a), i.e.

$$n[(p + 1) + \cdots + (p + q) + (q + 1) + \cdots + p] \\ = np(p + 2q + 1)/2,$$

$Ux = y$  same as for general block-tridiagonal, i.e.  $np(3p + 1)/2$

giving a total of  $np(2p + q + 1)$  multiplications for the backsolution.

(c) *Blocksolve for (6)*.

(i) Block *LU decomposition*: For  $i = 2, \dots, n$ , we have

$$\left. \begin{array}{l} LU \text{ decomposition of } U_{i-1}: p^3/3 \text{ operations} \\ \text{solution of } L_i(U_{i-1}) = A_i: qp^2 \text{ operations} \\ \text{formation of } U_i = B_i - L_i C_{i-1}: pq(p - q) \text{ opns.} \end{array} \right\} \begin{array}{l} \text{total} \\ np(p^2/3 + 2pq - q^2). \end{array}$$

(Note that  $L_i$  is zero in its last  $(p - q)$  rows and  $C_i$  is zero in its first  $q$  rows.)

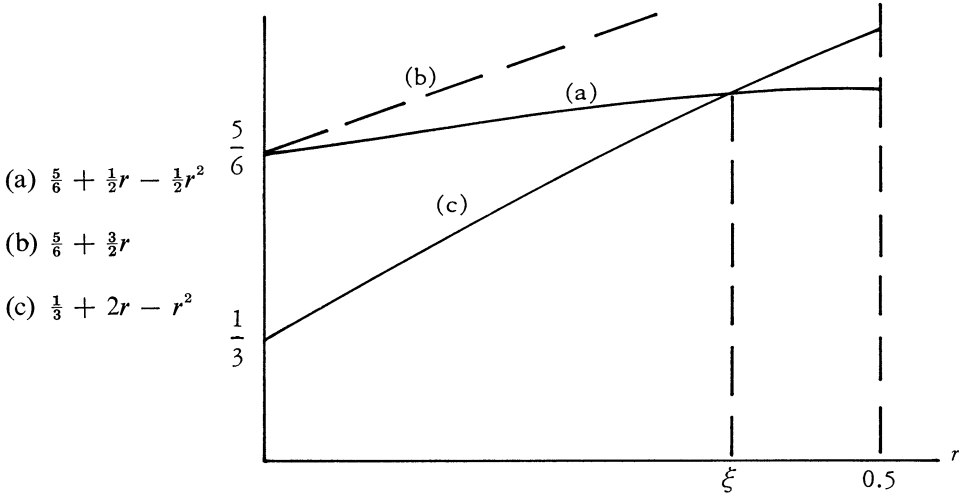
(ii) Backsolution:

$$Ly = f \text{ takes } npq \text{ operations,} \\ Ux = y \text{ takes } np(2p - q) \text{ operations.}$$

We can summarize these operation counts, keeping only terms  $O(npq)$ :

$$\begin{array}{ll} \text{(a) bandsolve (no pivoting)} & - np(\frac{5}{6}p^2 + \frac{1}{2}pq - \frac{1}{2}q^2 + \frac{3}{2}p + q), \\ \text{(b) bandsolve (with pivoting)} & - np(\frac{5}{6}p^2 + \frac{3}{2}pq + \frac{3}{2}p + \frac{3}{2}q), \\ \text{(c) blocksolve} & - np(\frac{1}{3}p^2 + 2pq - q^2 + 2p). \end{array}$$

It is clear that (b) is always slower than (c), but the interesting comparison is (a) vs. (c): (c) is better for  $r = q/p$  near zero (e.g.,  $q = 1$ ), but (a) is better for  $r$  near  $\frac{1}{2}$ . (We need only consider  $0 < r \leq \frac{1}{2}$  since for  $r > \frac{1}{2}$  it is more efficient to turn the matrix around.) To get a graphical comparison, we can drop the terms in  $npq$ ,  $np^2$ , and divide by  $np^3$ ; then each count is only a function of  $r$ :



Using just these terms, we see that (a) is better for  $r < \xi$ , (c) better for  $r > \xi$ ,  $\xi \cong 0.38$ , with the difference at  $r = \frac{1}{2}$  amounting to about 12%. Actually, the graph is biased in favor of (c) especially for  $r$  small, because the term in  $np^2$  has been dropped. For an additional comparison, we give the operation counts from the table for  $p = 8$ :

count	$q$			
	1	2	3	4
(a)	70	73	76	77
(b)	78	92	105	119
(c)	52	65	76	85

Department of Computer Science  
 University of British Columbia  
 Vancouver 8, Canada

1. D. G. FEINGOLD & R. S. VARGA, "Block diagonally dominant matrices and generalizations of the Gerschgorin circle theorem," *Pacific J. Math.*, v. 12, 1962, pp. 1241-1250. MR 27 #1458.
2. A. S. HOUSEHOLDER, *The Theory of Matrices in Numerical Analysis*, Blaisdell, Waltham, Mass., 1964. MR 30 #5475.
3. E. A. ISAACSON & H. B. KELLER, *Analysis of Numerical Methods*, Wiley, New York, 1966. MR 34 #924.
4. H. B. KELLER, "Accurate difference methods for linear ordinary differential systems subject to linear constraints," *SIAM J. Numer. Anal.*, v. 6, 1969, pp. 8-30.
5. H. B. KELLER, "A new difference scheme for parabolic problems," *Numerical Solution of Partial Differential Equations II (SYNSPADE 1970)* (Proc. Sympos., Univ. of Maryland, College Park, Md., 1970), Academic Press, New York, 1971, pp. 327-350. MR 43 #2866.
6. R. D. RICHTMYER & K. W. MORTON, *Difference Methods for Initial-Value Problems*, 2nd ed., Interscience Tracts in Pure and Appl. Math., no. 4, Interscience, New York, 1967. MR 36 #3515.
7. B. WENDROFF, *Theoretical Numerical Analysis*, Academic Press, New York, 1966. MR 33 #5080.
8. J. H. WILKINSON, "Error analysis of direct methods of matrix inversion," *J. Assoc. Comput. Mach.*, v. 8, 1961, pp. 281-330. MR 31 #874.
9. J. A. GEORGE, "An optimal ordering of a regular finite element mesh," *SIAM J. Numer. Anal.* (To appear.)