TENSOR PRODUCT ANALYSIS OF PARTIAL DIFFERENCE EQUATIONS

BY R. E. LYNCH, J. R. RICE AND D. H. THOMAS

Communicated by A. S. Householder, January 6, 1963

1. **Introduction.** This note presents a few of the results which we have obtained by applying a classical and fundamental idea to the analysis of certain partial difference equations. The idea is that certain multidimensional problems can be solved by solving a few one-dimensional problems—it is the basis of the classical method of separation of variables of mathematical physics. In the case of partial difference equations, this idea leads to *tensor product* analysis of the matrices involved.

With this approach we accomplish the following: (i) Explicit *exact* solutions of problems consisting of separable partial difference equations and boundary conditions are obtained, (ii) A stable algorithm is devised with which these exact solutions can be evaluated with less work than approximate solutions can be computed by overrelaxation techniques, (iii) A simple, direct analysis of certain alternating direction implicit (ADI) methods is carried out and, as a result, a simple explanation of the power of this method is given, (iv) A necessary and sufficient condition is found for commutativity of certain matrices which occur in ADI schemes.

2. **Tensor products applied to elliptic and parabolic boundary value problems.** The *tensor product* (Kronecker product, direct product) of two matrices $A = \{a_{ij}\}$ and $B = \{b_{kl}\}$, denoted by $A \otimes B$, can be written as a matrix in block partition form:

$$A \otimes B = \begin{pmatrix} a_{11}B & \cdots & a_{1n}B \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ a_{m1}B & \cdots & a_{mn}B \end{pmatrix}.$$

A detailed account of properties of tensor products is given in [8]. Some of the elementary properties are:

$$(A+B) \otimes C = A \otimes C + B \otimes C, \qquad A \otimes (B+C) = A \otimes B + A \otimes C,$$
$$(A \otimes B)(C \otimes D) = AC \otimes BD, \qquad (A \otimes B)^{-1} = A^{-1} \otimes B^{-1}.$$

For brevity, we do not indicate explicitly the sizes of the matrices involved; we assume throughout that the sizes of matrices and vectors are compatible with the indicated operations.

Consider a second order linear elliptic partial differential equation in two independent variables x and y

$$(2.1) \qquad\qquad \mathcal{L}[v] = f,$$

in which the differential operator $\mathcal{L}$ has the special form

$$(2.2) \qquad\qquad \mathcal{L} = \mathcal{L}_x + \mathcal{L}_y,$$

where $\mathcal{L}_x(\mathcal{L}_y)$ is an operator with derivatives only with respect to $x(y)$ and coefficients which depend only on $x(y)$. For example, $\mathcal{L}_x$ might be

$$\mathcal{L}_x = - a_2(x)\, \frac{\partial^2}{\partial x^2} + a_1(x)\, \frac{\partial}{\partial x} + a_0(x), \qquad a_2 > 0,$$

or it might be in self-adjoint form

$$\mathcal{L}_x = - a_2(x)\, \frac{\partial}{\partial x}\left(a_1(x)\, \frac{\partial}{\partial x}\right) + a_0(x), \qquad a_2 a_1 > 0.$$

We treat problems consisting of (2.1) in the unit square

$$R = \left\{ (x, y)\,\middle|\, 0 < x < 1,\ 0 < y < 1\right\}$$

and any of a number of conditions specified on the boundary of R such as Dirichlet, mixed, or periodic boundary conditions.

A partial difference equation is derived from (2.1) by the usual procedure of placing mesh points in R and replacing derivatives in $\mathcal{L}$ with divided differences. The set of mesh points are the intersections of a set of lines parallel to the x-axis with lines parallel to the y-axis; the spacing between lines need not be uniform. The result is a system of linear equations which can be written as

$$(2.3) \qquad\qquad Lu = g$$

where L is a matrix and u is the approximation to the solution v of (2.1) at mesh points. The vector g involves values of both the right side of (2.1) and boundary values. The exact form of L depends on the operator $\mathcal{L}$, the difference approximation, and the boundary conditions; but, for five point approximations and $\mathcal{L}$ of the form (2.2), it is readily verified that (2.3) is of the form

$$(2.4) \qquad\qquad (I \otimes A + B \otimes I)u = g$$

where I denotes identity matrices of appropriate sizes.

If A and B are nondefective, then there are nonsingular matrices P and Q whose columns are eigenvectors of A and B, respectively, such that

$$P^{-1}AP = \Lambda(A), \qquad Q^{-1}BQ = \Lambda(B),$$

where $\Lambda(A)$ denotes a diagonal matrix of eigenvalues of A. For the difference approximations and the boundary conditions commonly used, A and B are nondefective. It can easily be checked, by direct substitution and simplification, that the exact solution of (2.4) is

$$(2.5) \qquad u = \{Q \otimes P[I \otimes \Lambda(A) + \Lambda(B) \otimes I]^{-1}Q^{-1} \otimes P^{-1}\}g.$$

This is, of course, the analog of the familiar solution of partial differential equations in terms of a Green's function.

For some classical equations, such as the Laplace, Poisson, Helmholtz, biharmonic, etc., in rectangular Cartesian coordinates, the matrices P and Q are known explicitly for mesh points equally spaced in the coordinate directions. In other cases, when P and Q must be computed, the amount of computation is small compared with the total amount necessary for evaluating the solution. This is due to the fact that A and B are band matrices. For self-adjoint problems, P and Q are real orthogonal matrices so that $P^{-1} = P^T$, $Q^{-1} = Q^T$. In nonsymmetric problems, one does not invert P and Q, rather one determines the eigenvectors of the matrices A^T and B^T, a simpler problem than the inversion.

Although the matrix on the right side of (2.5) is the inverse of L, it is not efficient to compute L^{-1}. However, the solution of the boundary value problem can be computed efficiently by evaluating

$$(2.6) \qquad u_{ij} = \sum_\alpha q_{i\alpha} \sum_\beta p_{j\beta}[\lambda_\beta(A) + \lambda_\alpha(B)]^{-1} \sum_k \bar{q}_{\alpha k} \sum_l \bar{p}_{\beta l}g_{kl}$$

in which $\{p_{j\beta}\} = P$, $\{q_{i\alpha}\} = Q$, $\{\bar{p}_{\beta l}\} = P^{-1}$, $\{\bar{q}_{\alpha k}\} = Q^{-1}$, and $\lambda_\beta(A)$, $\lambda_\alpha(B)$ are eigenvalues of A, B, respectively. The values of u can be found from (2.6) by computing, successively, three intermediate matrices, each of the same size as u. The evaluation of (2.6) is computationally stable.

We now compare the work required to solve the Laplace equation with a grid of n points in each direction by (i) optimum successive overrelaxation (SOR), (ii) direct solution by tensor products (TP), (iii) the Peaceman-Rachford alternating direction implicit method (ADI). The SOR and ADI methods are iterative and we assume that the iteration is terminated when the reduction in error is consistent with the discretization error of the difference approximation. Moreover, we consider the standard five point approximation to the Laplacian so that the iteration is terminated when the initial error is reduced by a factor of n^{-2}. The asymptotic order (as n tends to

infinity) of the total number of arithmetic operations (additions, multiplications, divisions) for each method is[1]

Method:	SOR	TP	ADI
Operations:	$14 \, n^3 \log n$	$4n^3$	$40 \, n^2 \log^2 n$

The ratio of additions to multiplications is approximately the same for each method. The superiority of the ADI method over the TP method is to be expected because it is an efficient approximation method which is based on the tensor product nature of the problem (see §3).

Thus, it is clear that problems of the form (2.1)–(2.2) can be solved by first solving two one-dimensional problems. The method is easily extended to higher dimensions. The following are examples of the tensor product formulation of common partial difference equations. As above, the region is a rectangle.

Nine point approximation to the Laplace (Poisson) equation:

$$[6I \otimes A + 6B \otimes I + B \otimes A]u = g.$$

Five point approximation to the Helmholtz equation:

$$[I \otimes A + B \otimes I + \sigma I \otimes I]u = g, \qquad \sigma = \text{constant}.$$

Thirteen point approximation to the biharmonic equation:

$$[I \otimes A + B \otimes I]^2 u = [I \otimes A^2 + 2B \otimes A + B^2 \otimes I]u = g.$$

The Laplace equation in spherical coordinates (3 dimensions):

$$\{[C \otimes B_1 + I \otimes B_2] \otimes I + I \otimes I \otimes A\}u = g.$$

This tensor product method can also be used to solve parabolic finite difference equations. A typical scheme is the (implicit) Crank-Nicholson method for the heat equation which results in a partial difference equation of the form

$$[I \otimes A + B \otimes I][u(t + \Delta t) + u(t)] = 2\sigma[u(t + \Delta t) - u(t)]$$

where σ depends on the grid spacing.

---

[1] For computing machines with $32K$ core storage, n can be as large as (about) 150. For $n > 150$, tapes must be used. For such problems, the TP method has the additional advantage that *fewer* tape operations are required for it than for ADI. Currently, 10,000 words can be transferred to or from tapes in a time interval equivalent to about 10^6 arithmetic operations. Consequently, for $150 \leq n \leq 1000$, the TP method is more economical than ADI for solving the Laplace equation.

In addition to elliptic and parabolic partial difference equations, tensor product analysis can be applied to other problems such as difference-differential equations, integral equations in more than one variable, integro-differential equations, etc.

3. Tensor product analysis of alternating direction implicit methods. Consider the differential equation (2.1) and partial difference equation (2.3) derived from it. Even if the operator $\mathcal{L}$ is not of the form (2.2), one can write (2.3) in the form

$$(3.1) \qquad\qquad (H + V)u = g$$

where H and V are the matrices derived from differentiation with respect to x and y, respectively. We assume that any cross derivative term has been eliminated. If $\mathcal{L}$ is of the form (2.2) then $H = I \otimes A$, $V = B \otimes I$ as in (2.4).

It is well known [2] that commutativity of H and V plays an important role in the analysis of ADI methods. It is clear that H and V commute if L is of the form (2.4). It is also known [3] that (2.4) is not necessary for commutativity. We now give a necessary and sufficient condition for commutativity for the special case of five point approximations. We say that H and V are of *standard block tridiagonal form* if H is block diagonal where each diagonal block has all nonzero entries on the tridiagonal and if V is block tridiagonal where each of the tridiagonal blocks is a nonsingular diagonal matrix.

THEOREM. *If H and V are of standard block tridiagonal form, then a necessary and sufficient condition for $HV = VH$ is that there exist a nonsingular diagonal matrix D and tridiagonal matrices A and B such that*

$$(3.2) \qquad H = D^{-1}[I \otimes A]D, \qquad V = D^{-1}[B \otimes I]D.$$

This theorem states, in essence, that the most general differential equation which yields the equation (3.1) by means of a five point approximation and with $HV = VH$ is

$$(3.3) \qquad\qquad \frac{1}{d(x, y)} \, \mathcal{L}[d(x, y)v] = f$$

where $\mathcal{L}$ is of the form (2.2).

Note that for *any* matrices A and B which have complete sets of orthonormal eignvectors e_i and f_j, corresponding to eigenvalues λ_i and σ_j, respectively, the vectors $f_j \otimes e_i$ form a complete set of orthonormal eigenvectors for both $I \otimes A$ and $B \otimes I$. Moreover, the vectors $D^{-1}f_j \otimes e_i$

form a complete set of eigenvectors of $D^{-1}[I \otimes A + B \otimes I]D$ orthogonal with respect to D^2.

The following ADI scheme is given in [10]

$$(3.4) \quad \begin{aligned} (H + \rho_m I)u^{(m+1/2)} &= g - (V - \rho_m I)u^{(m)} \\ (V + \rho_m I)u^{(m+1)} &= (V - \omega\rho_m I)u^{(m)} + (1 + \omega)\rho_m u^{(m+1/2)} \end{aligned}$$

with fixed ω and iteration parameters ρ_m. The values $\omega = 1$, 0 correspond to the Peaceman-Rachford scheme [9] and the Douglas-Rachford scheme [6], respectively. If H and V commute, then by the theorem, they are of the form (3.2). Hence the error $\epsilon^{(m)} = u^{(m)} - u$ at the mth iterate satisfies

$$D^{-1}[I \otimes (A + \rho_m I)]D\epsilon^{(m+1/2)} = - D^{-1}[(B - \rho_m I) \otimes I]D\epsilon^{(m)},$$

$$\begin{aligned} D^{-1}[(B + \rho_m I) \otimes I]D\epsilon^{(m+1)} &= D^{-1}[(B - \omega\rho_m I) \otimes I]D\epsilon^{(m)} \\ &\quad + (1 + \omega)\rho_m \epsilon^{(m+1/2)}. \end{aligned}$$

The error can be expanded in terms of the eigenvectors as

$$\epsilon^{(m)} = \sum \alpha_{ij}^{(m)} D^{-1} f_j \otimes e_i.$$

A simple computation shows

$$\alpha_{ij}^{(m+1)} = \alpha_{ij}^{(m)} \left[\frac{(\lambda_i - \rho_m)(\sigma_j - \rho_m)}{(\lambda_i + \rho_m)(\sigma_j + \rho_m)} + (1 - \omega) \frac{\rho_m(\lambda_i + \sigma_j)}{(\lambda_i + \rho_m)(\sigma_j + \rho_m)} \right].$$

Note that for the special choice $\omega = 1$ (the Peaceman-Rachford scheme), $\alpha_{ij}^{(m+1)}$ can be made zero for *all* i by the choice $\rho_m = \sigma_j$ for some j. This explains the power of the Peaceman-Rachford scheme—a large number of the components of the error vector can be annihilated *simultaneously*. Furthermore, this is done without increasing any other components of the error vector if the λ_i, σ_j and ρ_m are positive, as is usually the case. In contrast, the SOR method can annihilate at most *one* component of the error; the Douglas-Rachford scheme *cannot* annihilate any error component. Therefore, the Peaceman-Rachford scheme can be exact (except for roundoff) in a number of iterates equal to the number of grid points taken in either the x or y direction. Nevertheless, it is normally more efficient to use a set of "optimum parameters" [7]. This is due to the fact that satisfactory (but inexact) results can normally be obtained with a smaller total number of iterations [3].

Tensor product analysis has been applied to schemes proposed by Douglas [5] for parabolic equations, Baker and Oliphant [1], Conte

and Dames [4] and others. The results will be reported in the paper which describes the foregoing results in detail.

This analysis has also been applied to problems in higher dimensions and allows one to easily generate several ADI schemes in higher dimensions.

Added in proof. Formula (2.5) is implicitly given by studies of the separation of variables in an abstract setting, e.g. [11; 12]. Let $A_0 = \{-\delta_{i,i-1} + 2\delta_{ii} - \delta_{i,i+1}\}$. Egerváry [13] gives (2.5) for $L = (A_0 \otimes I + I \otimes A_0)$, but does not give the efficient computational method (2.6). Heller [14] has treated ADI methods for $L = (A_0 \otimes I + I \otimes A_0)^p$.

REFERENCES

1. G. A. Baker, Jr. and T. A. Oliphant, *An implicit, numerical method for solving the two-dimensional heat equation*, Quart. Appl. Math. 17 (1960), 361–373.

2. G. Birkhoff and R. S. Varga, *Implicit alternating direction methods*, Trans. Amer. Math. Soc. 92 (1959), 13–24.

3. G. Birkhoff, R. S. Varga and D. M. Young, *Alternating direction implicit methods*, Advances in Computers, Vol. III, pp. 189–273, Academic Press, New York, 1962.

4. S. D. Conte and R. T. Dames, *An alternating direction method for solving the biharmonic equation*, Math. Comp. 12 (1958), 198–205.

5. Jim Douglas, *Alternating direction methods for three space variables*, Numer. Math. 4 (1962), 41–63.

6. Jim Douglas, Jr. and H. H. Rachford, Jr., *On the numerical solution of heat conduction problems in two or three space variables*, Trans. Amer. Math. Soc. 82 (1956), 421–439.

7. C. de Boor, and J. R. Rice, *Chebyshev approximation by $\alpha\Pi(x-r_i)/(x+s_i)$ and application to ADI iteration*, J. Soc. Indust. Appl. Math. 11 (1963), 159–169.

8. P. R. Halmos, *Finite-dimensional vector spaces*, Chapter II, Van Nostrand, Princeton, N. J., 1958.

9. D. W. Peaceman and H. H. Rachford, Jr., *The numerical solution of parabolic and elliptic differential equations*, J. Soc. Indust. Appl. Math. 3 (1955), 28–41.

10. R. S. Varga, *Matrix iterative analysis*, Prentice-Hall, Englewood Cliffs, N. J., 1962.

11. H. O. Cordes, *Separation der Variablen in Hilbertschen Räumen*, Math. Ann. 125 (1953), 401–434.

12. B. Friedman, *An abstract formulation of the method of separation of variables*, Proceedings of the conference on differential equations (dedicated to A. Weinstein), pp. 209–226, Univ. of Maryland Book Store, 1956.

13. E. Egerváry, *Über eine Methode zur numerischen Lösung der Poissonschen Differenzengleichung für beliebige Gebiete*, Acta Math. Acad. Sci. Hungar. 11 (1960), 341–361.

14. J. Heller, *Simultaneously, successive, and alternating direction iteration schemes*, J. Soc. Indust. Appl. Math. 8 (1960), 150–173.

GENERAL MOTORS CORPORATION, WARREN, MICHIGAN