JOURNAL OF COMPUTATIONAL PHYSICS 17, 181–191 (1975)

Engineering Applications of ADI Methods to Piecewise Linear Multidimensional Heat Transfer*

JIMMIE H. SMITH

Los Alamos Scientific Laboratory, Los Alamos, New Mexico 87544

Received March 14, 1974; revised October 2, 1974

Transient multidimensional forms of the heat equation are developed with alternating-direction-implicit (ADI) methods. Nonlinearities which stem from boundary conditions and variable properties are approximated in a piecewise linear manner. The resulting ADI forms of the heat equation are successfully demonstrated on a diverse pair of applications consisting of an air-to-ground missile design and an orbiting hollow cube exposed to short-duration radiation. In both applications, the ADI equations prove to be an order of magnitude more economical of computer time than the traditional Crank–Nicolson method. Moreover, the ADI equations prove to be less sensitive to approximation of nonlinearities than the DuFort–Frankel explicit method.

I. INTRODUCTION

In the present paper, alternating-direction-implicit (ADI) methods are applied to the transient multidimensional heat equation with variable properties. With analytical methods, representation of variable properties requires modification of the governing equation. However, with numerical methods, the use of the constant-property equation in a piecewise linear manner will yield approximate solutions when reasonably small time steps are used and nonlinearities are not too severe. In most practical applications, these constraints are satisfied. Two actual problem solutions, an air-to-ground missile flight and an orbiting hollow cube, are presented to illustrate the engineering approximations which facilitate this otherwise formidable task. Comparisons to better known methods help demonstrate the efficiency of ADI methods.

An earlier extensive survey of difference methods [1] found that for three-dimensional conduction with radiation boundary conditions, the ADI methods of Douglas [2] and Douglas and Gunn [3] were an order of magnitude faster than the widely used Crank–Nicolson method [4]. Another ADI method by Brian [5] was considered for this work but offered no advantage over that of Douglas.

Explicit differencing, while simple in concept, usually requires an order of

* This work was performed under the auspices of the US Atomic Energy Commission.

magnitude more small time steps than implicit methods in order to maintain stability. An exceptional scheme which is both explicit and unconditionally stable was developed by DuFort and Frankel [6] as a perturbation of the famous Richardson method which is unconditionally unstable as discussed in Ames [7, p. 327]. The DuFort–Frankel method is a very effective technique in practice [8] when nonlinearities from boundary conditions and/or variable properties are not dominant. Other unconditionally stable explicit methods such as those of Saul'yev [9] and Barakat and Clark [10] are also available.

The decision to use an implicit method, the best approach to which is ADI [1], was dictated by the severe nonlinearities in the applications of interest. That is, severe nonlinearities necessitate approximations to which ADI techniques are relatively insensitive. To further establish the efficacy of ADI methods, consider statements in the literature. Isaacson and Keller [11] survey difference methods for partial differential equations and conclude that "... the analytical methods developed for the treatment of partial differential equations are, in general, not suited for the efficient numerical evaluation of solutions." Further along, these authors [11] conclude, "We have thus shown that the alternating direction method is more efficient than any of the other iterative schemes, even when parameters that are not necessarily optimal are employed." In another text, Smith [12] says, "The most efficient method at present for rectangular regions is one proposed by Peaceman and Rachford [13] in 1955." Here it should be noted that the generalized ADI methods of Douglas [2], which we will use, are equivalent to that of Peaceman and Rachford in two dimensions and that of Crank–Nicolson in one dimension.

With the foregoing justifications, we proceed to develop the ADI equations for the fully three-dimensional case as well as another set for the two-dimensional axisymmetric case. Following this, the equations are applied with approximations which simplify the solution while reasonable engineering accuracy, as demonstrated by experiment, is maintained.

II. The ADI Forms of the Heat Equation

One-dimensional Crank–Nicolson [4] differencing of the transient heat equation yields tridiagonal matrices which are economical to solve by a variant of Gaussian elimination as suggested, for example, in Varga [14]. Moreover, tridiagonal systems which stem from finite-difference approximations to parabolic partial differential equations tend to be diagonally dominant and well conditioned as discussed in Forsythe and Moler [15, p. 117]. For multidimensional problems, alternating-direction-implicit or ADI methods of solution are generically referred to. These methods entail cyclic application of one-dimensional Crank–Nicolson differencing to one spatial coordinate after another.

Prior to expressing the transient three-dimensional heat or diffusion equation ($\partial T/\partial t = \alpha\nabla^2 T + \dot{q}'''/\rho c$) in ADI form, certain ground rules are formulated. The nonlinearity caused by temperature-dependent properties will be handled in an approximate manner as, for example, in [16]. This approximate method of evaluating properties at the last known temperature avoids the complexity of solving a system of nonlinear algebraic equations at each time step. In other words, piecewise linearization is used to approximate the actual property curves by a sequence of connected plateaus. The effect of this approximation on consistency and convergence cannot be rigorously demonstrated because of the nonlinearities.

To facilitate writing the ADI formulations, a second-order spatial differencing operator, $\delta_p{}^2$, where p denotes a spatial coordinate, is defined. For example, if $p = x$, then

$$\delta_x{}^2 T^{n+1}_{i,j,k} = (T^{n+1}_{i+1,j,k} - 2T^{n+1}_{i,j,k} + T^{n+1}_{i-1,j,k})/(\Delta x)^2.$$

Note that this differencing operator acting on temperature T does not modify superscripts. The subscripts (i, j, k) indicate position (x, y, z) where $x = i\,\Delta x$, $y = j\,\Delta y$, and $z = k\,\Delta z$. Right-hand superscripts such as $(n + 1)$ denote the time plane, $t = (n + 1)\,\Delta t$. Left-hand superscripts such as ${}^*T^{n+1}$ or ${}^{**}T^{n+1}$ denote first and second iterates, respectively. The third and final iterate (for three dimensions) will be written merely as T^{n+1} in preparation for using it to start the next cycle. Finally, let thermal diffusivity, $\alpha = k/(\rho c)$, where ρ = density, k = thermal conductivity, and c = specific heat capacity. With these conventions, the three-dimensional ADI form of the homogeneous isotropic heat equation may be written as:

$$\begin{aligned} {}^*T^{n+1}_{i,j,k} - T^n_{i,j,k} = {} & \alpha\,\Delta t[\delta_x{}^2({}^*T^{n+1}_{i,j,k} + T^n_{i,j,k})/2 \\ & + \delta_y{}^2 T^n_{i,j,k} + \delta_z{}^2 T^n_{i,j,k}] + (\dot{q}''')^n\,\Delta t/(\rho c), \end{aligned} \tag{1}$$

$$\begin{aligned} {}^{**}T^{n+1}_{i,j,k} - T^n_{i,j,k} = {} & \alpha\,\Delta t[\delta_x{}^2({}^*T^{n+1}_{i,j,k} + T^n_{i,j,k})/2 \\ & + \delta_y{}^2({}^{**}T^{n+1}_{i,j,k} + T^n_{i,j,k})/2 + \delta_z{}^2 T^n_{i,j,k}] + (\dot{q}''')^n\,\Delta t/(\rho c), \end{aligned} \tag{2}$$

$$\begin{aligned} T^{n+1}_{i,j,k} - T^n_{i,j,k} = {} & \alpha\,\Delta t[\delta_x{}^2({}^*T^{n+1}_{i,j,k} + T^n_{i,j,k})/2 + \delta_y{}^2({}^{**}T^{n+1}_{i,j,k} + T^n_{i,j,k})/2 \\ & + \delta_z{}^2(T^{n+1}_{i,j,k} + T^n_{i,j,k})/2] + (\dot{q}''')^n\,\Delta t/(\rho c). \end{aligned} \tag{3}$$

To avoid the homogeneous isotropic limitation, one would include spatial variation of properties with $\delta^2(\alpha T)$ rather than $\delta^2 T$. All T's are known for the nth time-plane and must be found for the $(n + 1)$st time-plane in the following manner. Equation (1) is applied successively to each of n nodes, and an $n \times n$ tridiagonal system is solved for ${}^*T^{n+1}_{i,j,k}$, the first estimate of $T^{n+1}_{i,j,k}$ at each node. Similarly Eq. (2) is applied to each node to find ${}^{**}T^{n+1}_{i,j,k}$, the second estimate of $T^{n+1}_{i,j,k}$.

Finally, Eq. (3) is applied to each node for the third and final estimate of $T^{n+1}_{i,j,k}$ at each node.

Inspection shows that one-dimensional Crank–Nicolson differencing, which is applied only to the x-coordinate in Eq. (1), propagates through the y and z coordinates in Eqs. (2) and (3). It might seem wise to replace the first estimate $*T^{n+1}_{i,j,k}$ in Eq. (3) by the second estimate $**T^{n+1}_{i,j,k}$ which is known before Eq. (3) is used. This logical fallacy destroys the unconditional stability as shown in Richtmyer and Morton [17]. Although Eqs. (1–3) are effective in making the ADI method transparent, a more efficient set of equations from a computational viewpoint is obtained by retaining Eq. (1), replacing Eq. (2) by the difference between (1) and (2) which we will call (4), and replacing Eq. (3) by the difference between (2) and (3) which we will call (5):

$$**T^{n+1}_{i,j,k} - *T^{n+1}_{i,j,k} = \alpha\,\Delta t\,\delta_y{}^2(**T^{n+1}_{i,j,k} - T^{n}_{i,j,k})/2, \tag{4}$$

$$T^{n+1}_{i,j,k} - **T^{n+1}_{i,j,k} = \alpha\,\Delta t\,\delta_z{}^2(T^{n+1}_{i,j,k} - T^{n}_{i,j,k})/2. \tag{5}$$

The final trio of three-dimensional equations consists of Eqs. (1), (4), and (5). Their use is demonstrated in Section IV.

For a two-dimensional axisymmetric geometry, another coordinate system yields simpler equations. A curvilinear body-contour coordinate s and an orthogonal coordinate r will be used as shown in Section III. The orthogonal curvilinear ADI forms of the homogeneous isotropic heat equation are:

$$*T^{n+1}_{s,r} - T^{n}_{s,r} = \alpha\,\Delta t[\delta_s{}^2(*T^{n+1}_{s,r} + T^{n}_{s,r})/2 + \delta_r{}^2T^{n}_{s,r}] + (\dot{q}''')^n\,\Delta t/(\rho c), \tag{6}$$

$$T^{n+1}_{s,r} - *T^{n+1}_{s,r} = \alpha\,\Delta t\,\delta_r{}^2(T^{n+1}_{s,r} - T^{n}_{s,r})/2. \tag{7}$$

Equations (6) and (7) are applied to an air-to-ground missile flight in Section III. Absolute measurements of s and r are not needed for a nodal solution because only the incremental values Δs and Δr are used. Figure 1 illustrates the incremental values for a coarse nodal struture. For accuracy, the example of Section III

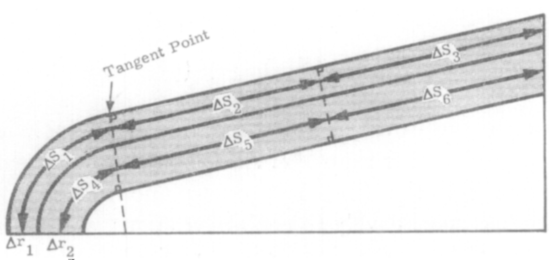

FIG. 1. Orthogonal curvilinear coordinate increments.

required a finer mesh. This choice of coordinate system essentially unfolds the layered heat shield into a planar layered structure. The usual axisymmetric coordinates would have done nicely for the near-conical (actually an ogive) portion of the vehicle but not for the hemispherical nose cap.

The preceding equations each yield tridiagonal systems. Banded matrices which occur in practice, especially those that arise from finite-difference approximations to parabolic differential equations, are usually diagonally dominant and well conditioned.

The key to solution is Gaussian elimination wherein the number of unknowns in each equation is reduced until a simple explicit solution is possible. Such a decomposition requires nm^2 multiplications for a tridiagonal matrix while $n^3/3$ are required for a full matrix [15]. For a tridiagonal matrix, bandwidth $= 3 = 2m + 1$ so $m = 1$ and $n =$ order. Thus nm^2 reduces to n-multiplications as compared with $n^3/3$ for a full matrix of order n. This major reduction justifies the statement that tridiagonal matrices are economical to solve.

Presentation of finite-difference equations cannot be considered complete without statements concerning consistency, stability, and convergence as discussed in detail in O'Brien, Hyman, and Kaplan [18] or Carnahan, Luther, and Wilkes [19]. Consistency implies that truncation errors vanish as spatial and temporal increments approach zero. That is, a consistent difference equation approximates the partial differential equation which it represents. Stability merely means that the amplification of errors is bounded. A theorem attributed to Lax guarantees convergence when both stability and consistency are satisfied.

Unconditional stability may be demonstrated for either set of ADI equations (1, 4, 5) or (6, 7) by means of the weakened von Neumann necessary condition on the amplification factor, $\xi(t + \Delta t)/\xi(t)$:

$$| \xi(t + \Delta t)/\xi(t)| \leqslant 1 + \mathcal{O}(\Delta t). \tag{8}$$

The term $\xi(t)$ represents temporal error only. The term $\mathcal{O}(\Delta t)$, which represents terms of the order of Δt, is the "weakening" referred to above which allows for legitimate exponential growth. When physics rules out solutions of the form $e^{\Delta t}$, the more stringent condition, $| \xi(t + \Delta t)/\xi(t)| \leqslant 1$, may be used. To find the amplification factor with the von Neumann method, one assumes that the temporal and spatial error contributions may be separated and that their product satisfies the difference equations. $T = \xi(t) \exp(i \sum \alpha_j x_j)$ is substituted into the difference equations to form the product of their amplification factors.

In an unconditionally stable scheme, such as the present ADI method, accuracy alone dictates the size of time increment. Accuracy is complicated by the fact that both truncation and round-off errors combine. Truncation error is established by considering the difference between the truncated difference expressions already presented and the full Taylor expansions for each derivative. This consideration shows the truncation error to be second-order in space and time plus the error associated with the source term when properties are evaluated at the midtime plane $(n + \frac{1}{2})$. Unfortunately this evaluation requires time-consuming iterative or predictor–corrector methods as suggested by Douglas [2]. For expediency, one often sacrifices the second-order accuracy in time and evaluates properties explicitly at the nth time-plane. In practice, this accuracy is generally as good as the imprecise knowledge of material properties anyway. Estimating the error in the source term is not generally feasible because of wide variation from case to case. Section III, Air-to-Ground Missile Application illustrates the complexity possible in the source term. In this application, a previously written boundary layer cold wall heating code and a hot wall correction were needed to evaluate the source term at each point prior to solving the diffusion equation.

In addition to truncation error, the finite word length on any digital computer causes what is termed "round-off error." In using Gaussian elimination to solve the tridiagonal systems given by the ADI equations, round-off error can accumulate to undesirable levels. Opposing effects occur when the time increment is reduced. Smaller time increments yield a more accurate approximate to a partial derivative. However, reducing the time increment greatly increases the number of operations required with the Gaussian elimination method in order to reach a given final time. Thus the round-off error increases greatly and may overwhelm the expected gain in accuracy from a reduced time increment.

In the two example problems considered here, let n be the number of nodes chosen and the order of the tridiagonal systems solved. For the extremely large problem in which an $n \times n$ tridiagonal system requires large amounts of core and solution time, a further approximation may be desirable. Beginning at the boundary of an arbitrary body and working inward, the ADI equations yield four unknowns and three independent equations for each trio of nodes. By evaluating the fourth unknown at the nth time-frame instead of the correct $(n + 1)$st time, one solves only 3×3 tridiagonal matrices for the complete problem. The accuracy of this approximation is controlled by the choice of time increment down to the point of domination by round-off error as discussed previously. This crude approximation should only be used with caution since it may destroy unconditional stability.

III. Air-to-Ground Missile (AGM) Application

Pyrolysis and the resulting carbonaceous char formation in the phenolic refrasil nosetip of an air-to-ground missile highly attenuate electromagnetic radiation at certain wavelengths. Thus, in order to conduct test flights requiring telemetry, it was necessary to insert a noncharring antenna window in the nosetip. The analysis consists of determining the specific thickness of antenna window (slip cast fused silica) and insulating foam which guarantees that the underlying phenolic refrasil will not rise above some predetermined temperature. This predetermined temperature comes from thermogravimetric and differential thermal analyses as discussed in Smith [20].

A constraint on total thickness was dictated by the outer body contour in conjunction with the minimum central volume for the antenna proper. Another constraint is that a certain minimum thickness of phenolic refrasil must be present to serve as structural support for the antenna components.

To carry out this task in an ideal manner would necessitate a simultaneous solution of the boundary-layer convective heating equations in conjuction with the transient diffusion equation. Instead, the surface recession was minimal so that an engineering approximation allowed uncoupling and separate solutions of the two sets of equations. First, the trajectory and boundary-layer heating calculations were performed for an assumed "cold wall" or reference temperature at the outher surface of the missile. These "cold wall heating rates" were then modified with a "hot wall correction" and combined with fourth-power radiation to space to form the source term in Eqs. (6) and (7).

With the hot-wall correction consisting of a ratio of enthalpy differences, hot-wall heating rates are computed as follows.

$$\dot{q}_{hw} = \dot{q}_{cw}(h_r - h_{hw})/(h_s - h_{cw}) \tag{9}$$

($h \sim$ enthalpy; $\dot{q} \sim$ heat flux; subscripts: $r \sim$ recovery; $s \sim$ stagnation; $cw \sim$ cold wall; $hw \sim$ hot wall).

Complete details of this application of the ADI equations (6, 7) are given in [20]. However, for the present purposes, a summary will suffice. Figure 1 illustrates a coarse mesh of orthogonal curvilinear coordinate increment for use with Eqs. (6) and (7). The actual solution required a much finer mesh for accuracy. The AGM geometry is comprised of a near-conical ogive and a hemispherical nose cap.

Solutions to Eqs. (6) and (7) yield the temperature profiles plotted in Fig. 2. To confirm their accuracy, an independent solution was obtained with an existing thermal analyzer [21] with Crank–Nicolson differencing. The agreement between the two methods was good enough to obviate any discussion of the differences. The importance of the application dictated experimental verification in radiant-heat and arc-jet tests. Redundant thermocouples were placed symmetrically to check the lower curve of Fig. 2. One thermocouple agreed almost perfectly, while the other showed almost 50°C higher at the end of flight. Fortunately, observation through a viewing port detected a flame near the high thermocouple. This provided a strong justification for discarding an extraneous data point. Furthermore, the agreement among two independent numerical methods and one thermocouple should be sufficiently miraculous for anyone.

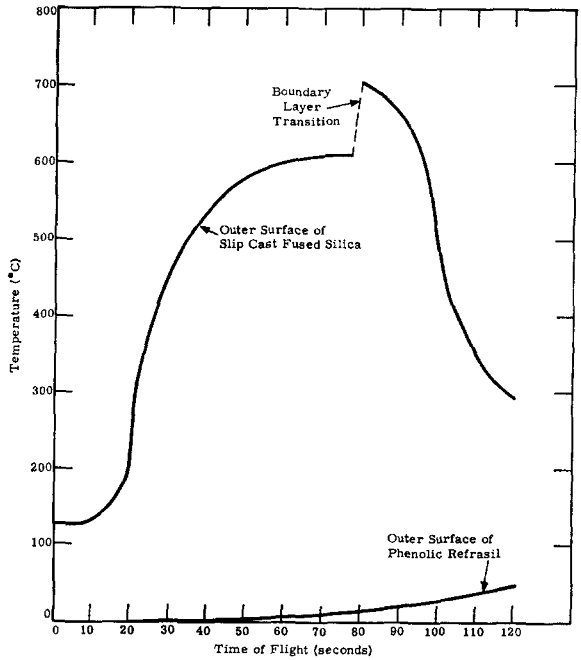

FIG. 2. Temperature at antenna window of AGM.

The ADI solutions required 30 times less (8 sec vs 245 sec) CPU time on a CDC 6600 than conventional Crank–Nicolson differencing. Because of the impossibility of accounting for input–output and peripheral processing, this factor of 30 implies unwarranted significance. Henceforth, this gain will be spoken of as an order of magnitude.

IV. ORBITING HOLLOW CUBE APPLICATION

An orbiting hollow cube at a uniform initial temperature is exposed on two sides and one end to short-duration radiant energy deposition. Complete details of this application are given in [1]. The three-dimensional ADI forms of the heat equations (1, 4, 5) provide the long-term transient solution. For comparison purposes, two cases are considered: (1) equilibration by conduction through cube walls only; and (2) conduction through walls, radiation to space, and radiation across the interior of the cube.

Figure 3 shows temperature profiles which agree with intuitive notions. Radiation with conduction accelerates the equilibration process. Also the asymptotic steady-

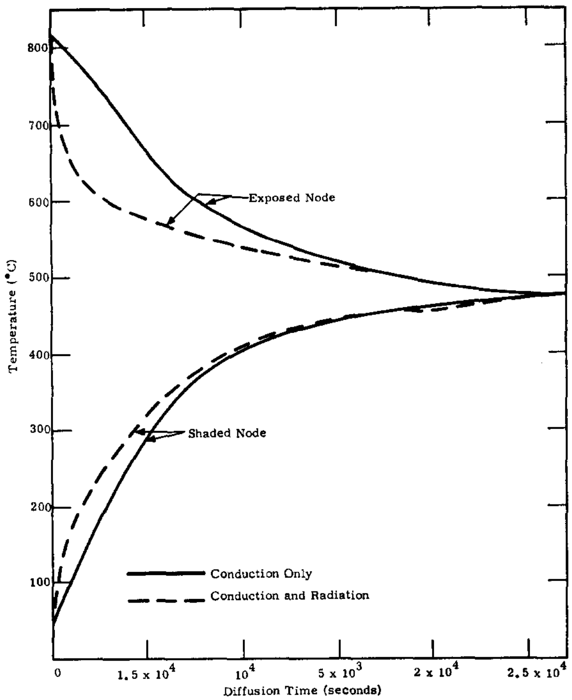

FIG. 3. Orbiting hollow cube exposed asymmetrically to radiant deposition (———, Conduction only; - - -, Conduction and radiation).

state solution is the same with or without radiation. To further confirm these results, four curves congruent to those in Fig. 3 were generated independently with Crank–Nicolson methods.

Both with and without radiation, the ADI solutions required 12 times less time (128 min vs 11 and 85 min vs 7) on the UNIVAC 1108 than with the conventional Crank–Nicolson differencing. Again this is better expressed as an order-of-magnitude improvement.

V. RESULTS AND RECOMMENDATIONS

Many current conduction codes [22] use differencing techniques which are obsolete in the sense that they require unnecessarily large amounts of computer time for a specified accuracy. The approach in the present paper has been to select well-known implicit and explicit schemes as fiducial methods to illuminate characteristics of the ADI methods.

An explicit technique is preferred for simplicity, but stability constraints often force the use of small time steps. This, in turn, increases computer cost as well as round-off error. A rare explicit scheme by DuFort and Frankel [6] is unconditionally stable. Yet it too has drawbacks when compared with ADI methods. For one thing, the DuFort–Frankel method requires two planes of initial values. Providing the second plane of starter values usually requires the troublesome use for one time step of another technique [8]. Furthermore, explicit techniques appear to be more sensitive to approximation of nonlinearities than ADI methods. In a simplistic intuitive manner, one might say that implicit methods dampen the sensitivity to boundary conditions by simultaneously considering multiple points.

In conclusion, for at least two applications, ADI methods have been shown to be an order of magnitude faster than the widely used Crank–Nicolson method with far less sensitivity to boundary approximations than the DuFort–Frankel method. These boundary approximations are almost mandatory in that they preclude solving simultaneous nonlinear algebraic equations at each time step. Since ADI methods are not completely understood by anyone, it is conceivable that, in some cases, the approximations and linearizations could lead to an undamped error growth. In that event, one would of necessity resort to some other method. However, this failure has not yet happened in a practical application so confidence is growing. It would seem that ADI methods could be recommended with just one reservation involving existing thermal analyzers. The reservation stems from the necessity of tagging conductors with spatial dimensions in order to cyclically difference one spatial coordinate after another. For this reason, adding ADI methods to existing thermal analyzers may be undesirably troublesome. Existing thermal analyzers are not easily modified to retain knowledge of the spatial direction of each conductor in an orthogonal grid system.

References

1. Jimmie H. Smith, "Survey of Three-Dimensional Finite Difference Forms of Heat Equation," Sandia Laboratories Report, SC-M-70-83, March 1970.
2. Jim Douglas, Jr., *Numer. Math.* **4** (1962), 41–63.
3. J. Douglas and J. Gunn, *Numer. Math.* **6** (1964), 428.
4. J. Crank and P. Nicolson, *Proc. Cambridge Philos. Soc.* **43** (1947), 50–67.
5. P. L. T. Brian, *A.I.Ch.E.J.* **7** (1961), 367.
6. E. C. DuFort and S. P. Frankel, *Mathematical Tables and Other Aids to Computation* **7** (1953), 135–152.
7. William F. Ames, "Nonlinear Partial Differential Equations in Engineering," p. 328, Academic Press, New York and London, 1965.
8. Jimmie H. Smith, "Improved Thermal Designs for a Cylindrical Electrical Igniter," accepted for presentation at the Fifth International Heat Transfer Conference in Tokyo, Japan, September 2–7, 1974.

9. V. K. SAUL'YEV, "Integration of Equations of Parabolic Type by the Method of Nets," MacMillan, New York, 1964.
10. H. Z. BARAKAT AND J. A. CLARK, *J. Heat Transfer* **88** (1966), 421–427.
11. EUGENE ISAACSON AND HERBERT BISHOP KELLER, "Analysis of Numerical Methods," Wiley, New York, 1966.
12. G. D. SMITH, "Numerical Solution of Partial Differential Equations," Oxford University Press, New York and London, 1965.
13. D. W. PEACEMAN AND H. H. RACHFORD, JR., *J. Soc. Ind. Appl. Math.* **3** (1955), 28.
14. 14. R. S. VARGA, "Matrix Iterative Analysis," Prentice–Hall, Englewood Cliffs, NJ, 1962.
15. GEORGE FORSYTHE AND CLEVE B. MOLER, "Computer Solution of Linear Algebraic Systems," Prentice–Hall, Englewood Cliffs, NJ, 1967.
16. C. BONACINA, G. COMINI, A. FASANO AND M. PRIMICERIO, *Intern. J. Heat Mass Transfer.* **16** (1973), 1825–1832.
17. ROBERT D. RICHTMYER AND K. W. MORTON, "Difference Methods for Initial-Value Problems," Interscience, New York, 1967.
18. G. O'BRIEN, M. HYMAN, AND S. KAPLAN, *J. Math. Phys.* **29** (1951), 233–251.
19. BRICE CARNAHAN, H. A. LUTHER, JAMES O. WILKES, "Applied Numerical Methods," Wiley, New York, 1969.
20. JIMMIE H. SMITH, *J. Spacecraft Rockets* **11** (1973), 198–200.
21. D. R. LEWIS, J. D. GASKI AND L. R. THOMPSON, "Chrysler Improved Numerical Differencing Analyzer for 3rd Generation Computers," Chrysler Corporation Space Division, TN-AP-67-187, October 1967 (The latest version of this code, known as SINDA, is available from NASA-MSC).
22. W. A. SHUKER, *Nuclear Safety* **12** (1971), 569–582.