J. W. White

Removal of Numerical Instability in the Solution of Nonlinear Heat Exchange Equations

Abstract: Current relaxation methods used by most large-scale thermal-computation programs for solving steady-state temperature distributions are subject to numerical instability when radiation is present in the system under study. The instability is usually due to the linearization of the numerical formulation of the governing equations and manifests itself as an oscillatory or divergent solution. A new method of solution is presented, which takes advantage of the diagonal dominance of the coefficient matrices of the linear and nonlinear transfer modes and uses a Gauss-Seidel system relaxation method together with a Newton-Raphson root-evaluation technique. System solution appears similar to that for a diagonally dominant linear system for all magnitudes of radiation. Test cases show the method to be monotonically convergent over the entire range of the radiation parameter considered, while previous methods were found to fail at certain magnitudes of the nonlinear term.

Nomenclature

$\mathbf{A}$	Conduction coefficient matrix
A_k	Surface area of node k
a_{ij}	Element of matrix $\mathbf{A}$
$\mathbf{B}$	Radiation coefficient matrix
b_{ij}	Element of matrix $\mathbf{B}$
$\mathbf{C}$	Heat generation matrix
Δx	Spatial increment
$\mathscr{F}_{i\rightarrow j}$	Overall radiation factor between nodes i and j
$f(T_k)$	Summation of all heat fluxes into node k
K	Thermal conductivity
$\mathbf{q}_r$	Radiant heat flux
Q'''	Volume rate of heat generation
Qgen	Heat generation
R_{ni}cond	Conduction resistance
R_{ni}rad	Radiation resistance
T_{mi}	Temperature of adjacent node point
T_k	Temperature of central node point
T_∞	Radiation sink temperature
σ	Stefan-Boltzmann constant

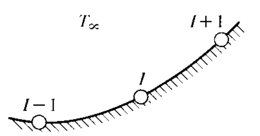


Figure 1 Node point distribution for conduction and radiation.

Introduction

In the high temperature applications of modern technology, thermal radiation has become an important mode of heat transfer. Whether it be heat rejection from a spacecraft surface to deep space or from a hot electronic component on a printed circuit card to a container structure, radiation is quite often an important effect which should be taken into account during design. When radiation is present, the thermal energy equation appears as

$$\text{div}\,(K\,\text{grad}\,T) - \text{div}\,\mathbf{q}_r + Q''' = 0, \tag{1}$$

where $\mathbf{q}_r$ is the radiative flux vector.

A standard procedure for obtaining a numerical solution to Eq. (1) is to approximate the gradient terms by finite differences and the nonlinear radiation terms by small areas of uniform temperature. The resulting system of algebraic equations is then solved for the temperature values. Without the radiation terms, the solution is quite simple; with them, computational difficulties may arise if the magnitude of radiation is comparable to that of conduction.

Consider a general case with radiation and conduction. Assuming that the mesh or node point I is in thermal contact with two adjacent nodes and with deep space (see Fig. 1) and, for simplicity here, assuming that the overall radiation coefficients $\mathscr{F}_{i\rightarrow j}$ are all equal, we can reduce the finite difference form of Eq. (1) to

$$T_l + \frac{3\sigma\mathscr{F}_{i\to j}}{K}\Delta x T_l^4 - \left(\frac{T_{l-1}}{2} + \frac{T_{l+1}}{2} + \frac{\sigma\mathscr{F}_{i\to j}}{K}\Delta x T_{l-1}^4 + \frac{\sigma\mathscr{F}_{i\to j}}{K}\Delta x T_{l+1}^4\right) - \frac{\sigma\mathscr{F}_{i\to j}}{K}\Delta x T_\infty^4 - \frac{Q_l'''}{2K}(\Delta x)^2 = 0. \quad (2)$$

Now, instead of needing to solve a system of linear algebraic equations as is the case for conduction only, the problem is one involving nonlinear algebraic equations. It is this nonlinearity that presents the difficulty in obtaining a solution to Eq. (1) by relaxation methods. One of the most frequently used methods [1–3] for solving systems of nonlinear equations in thermal analysis consists of a linearization of the nonlinear terms with a Gauss-Seidel [4] solution of the resulting system. The finite difference form of Eq. (1) may be written as

$$\sum_i \left(\frac{KA}{\Delta x}\right)_{k,mi} (T_{mi} - T_k) + \sum_i \sigma\mathscr{F}_{k\to mi} A_k (T_{mi}^4 - T_k^4) + Q_k\text{gen} = 0, \quad (3)$$

where the subscript k corresponds to the node point being solved for and mi refers to an adjacent node point.

With use of the electrical analogy, thermal resistances are defined as

$$R_{ni}\text{cond} = \frac{\Delta x}{KA}\ .$$

and

$$R_{ni}\text{rad} = [\sigma\mathscr{F}_{k\to mi} A_k (T_{mi}^2 + T_k^2)(T_{mi} + T_k)]^{-1}.$$

Equation (3) then takes the form

$$\left(\sum_i \frac{1}{R_{ni}\text{cond}} + \sum_i \frac{1}{R_{ni}\text{rad}}\right) T_k = \sum_i \frac{T_{mi}}{R_{ni}\text{cond}} + \sum_i \frac{T_{mi}}{R_{ni}\text{rad}} + Q_k\text{gen}. \quad (4)$$

The coefficient of the left side of Eq. (4) is clearly dominant over those of the right side, a condition that is necessary to insure convergence in a Gauss-Seidel method of solution. A problem arises, however, in that the radiation resistances R_{ni} rad are functions of the temperature T_k. Since the radiation resistance is proportional to the inverse cube of the temperature, small errors in calculated values of T_k cause a larger error in the corresponding radiation resistance. This tends to cause an oscillatory system response that may or may not be damped out, depending on the relative magnitudes of the radiation and conduction terms.

Another way [2] of handling the general case with radiation and conduction is to classify each node as either conduction or radiation dominant. The conduction dominant solution is similar to that of the linearized method whereas the radiation dominant solution determines a new value of temperature from only the radiation term. In the latter case the conduction term is included but is based on the last calculated value. As would be expected, the dominant-mode method works in cases with extreme ratios of conduction to radiation. There are usually regions with conduction and radiation of the same order of magnitude, however, where this method fails.

The work described in this paper was undertaken to find an approximate method of solution for Eq. (1) that completely removes the numerical instabilities associated with previous methods of solution. The requirements seemed to suggest that the full nonlinear terms would have to be kept and dealt with. One approach might have been to use a generalized Newton's method of solution [5]; however, the diagonal dominance of the coefficient matrices of both the conduction and radiation terms indicated that a Gauss-Seidel method should work, which was the case. It was felt that a tangible reward for success in this effort would be a great decrease in the amount of computation time necessary for solution, while intangible gains would include increased user confidence in final results.

The rest of this paper includes a presentation of the new technique, the several levels of tolerance required for system convergence, and the initial data requirements. Results of several test cases are presented, which compare the new method of solution with previous ones.

Solution technique

With the motivation that the full nonlinear terms should be retained, the radiation resistances are redefined as

$$R_{ni}\text{rad} = (\sigma\mathscr{F}_{k\to mi} A_k)^{-1}.$$

Equation (3) is then recast as

$$\sum_i \frac{T_{mi} - T_k}{R_{ni}\text{cond}} + \sum_i \frac{T_{mi}^4 - T_k^4}{R_{ni}\text{rad}} + Q_k\text{gen} = 0. \quad (5)$$

Rearrangement gives

$$\sum_i \frac{1}{R_{ni}\text{cond}} T_k + \sum_i \frac{1}{R_{ni}\text{rad}} T_k^4 - \left(\sum_i \frac{T_{mi}}{R_{ni}\text{cond}} + \sum_i \frac{T_{mi}^4}{R_{ni}\text{rad}} + Q_k\text{gen}\right) = 0. \quad (6)$$

Equation (6) may be written in matrix form as

$$\mathbf{AT} + \mathbf{BT}^4 = \mathbf{C}, \quad (7)$$

where **A** and **B** are coefficient matrices of the conduction and radiation terms, respectively, and **C** is a vector composed of the components of heat generation. From examination of Eq. (6), the **A** and **B** matrices are diagonally dominant, satisfying, respectively,

$$|a_{ii}| \geq \sum_{j \neq i} |a_{ij}|;$$

$$|b_{ii}| \geq \sum_{j \neq i} |b_{ij}|,$$

with the strict inequality holding for at least one mesh point.

These matrix properties indicate that a Gauss-Seidel method of solution will work, solving not for the single variable unknown T_k but for a polynomial in T_k given by Eq. (6):

$$\left(\sum_i \frac{1}{R_{ni}\text{cond}}\right) T_k + \left(\sum_i \frac{1}{R_{ni}\text{rad}}\right) T_k^4 = \sum_i \frac{T_{mi}}{R_{ni}\text{cond}} + \sum_i \frac{T_{mi}^4}{R_{ni}\text{rad}} + Q_k\text{gen}. \quad (6)$$

The right side of Eq. (6) is known from input or previous calculations and is treated as constant. Thus the conduction and radiation resistances are all now constant, and Eq. (6) can now be given more simply as

$$\alpha T_k + \beta T_k^4 - \zeta = 0, \quad (8)$$

where

$$\alpha = \sum_i \frac{1}{R_{ni}\text{cond}},$$

$$\beta = \sum_i \frac{1}{R_{ni}\text{rad}},$$

$$\zeta = \sum_i \frac{T_{mi}}{R_{ni}\text{cond}} + \sum_i \frac{T_{mi}^4}{R_{ni}\text{rad}} + Q_k\text{gen}.$$

It remains now to obtain the proper root of the polynomial expression for T_k. Define a heat input function as

$$f(T_k) = \alpha T_k + \beta T_k^{\,4} - \zeta = 0\,. \quad (9)$$

Equation (9) is sketched in Fig. 2. Of the four roots of Eq. (9), two are real and two are a complex conjugate pair. We are interested in the one real positive root, which is the only possible interpretation for absolute temperature.

The function $f(T_k)$ is always concave upward as is seen by taking the second derivative

$$f''(T_k) = 12\,\beta(T_k)^2.$$

This type of function yields nicely to root evaluation by a Newton-Raphson technique. Expanding $f(T_k)$ in a Taylor series around the most recently calculated value and truncating higher order terms, we obtain

$$f(T_{k,i+1}) \approx f(T_{k,i}) + f'(T_{k,i})(T_{k,i+1} - T_{k,i}),$$

where i corresponds to the ith iteration for T_k.

Requiring that $f(T_{k,i+1})$ equal zero (as it must for solution) gives, after rearrangement,

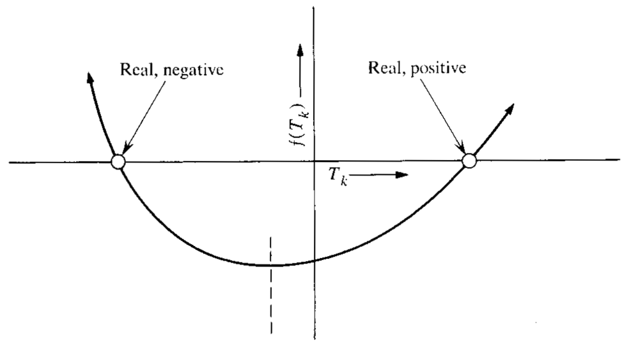


Figure 2 Graphical form of heat input function.

$$T_{k,i+1} = T_{k,i} - \frac{f(T_{k,i})}{f'(T_{k,i})}. \quad (10)$$

Equation (10) is evaluated repetitively until the difference $(T_{k,i+1} - T_{k,i})$ is less than some prescribed value of tolerance. Once this condition is satisfied, the next equation is solved in a similar manner by using the latest information available for the evaluation of the "constant" ζ. The net result is that the nonlinear system of equations approximating Eq. (1) is solved with a solution response similar to that for a linear system. At each mesh point the root of each equation is evaluated several times to satisfy tolerance requirements. This procedure requires more time than does the solution of a linear system, but overall the solution is in general more rapid than other methods we have considered, and is much more stable.

- *Initial conditions*

To define an appropriate range of initial temperature values in starting the relaxation process, we again consider Eq. (10). No problems are encountered as long as $f'(T_k)$ is nonzero, since it appears in the denominator. Three cases warrant consideration: 1) no conduction ($\alpha = 0$); 2) no radiation ($\beta = 0$); 3) both conduction and radiation present ($\alpha \neq 0, \beta \neq 0$).

For all of the cases considered, the positive real root of Eq. (9) will be approached with the Newton-Raphson method if all the initialized temperature values are greater than 0°R. This presents no limitation because the user can usually provide fairly good representative values of temperature based on experience and knowledge of the particular application.

- *Tolerances*

The temperature value of each node is iteratively calculated to within some tolerance, say DELTA 1 on each system pass. The system of equations is then said to converge when each nodal value changes by less than

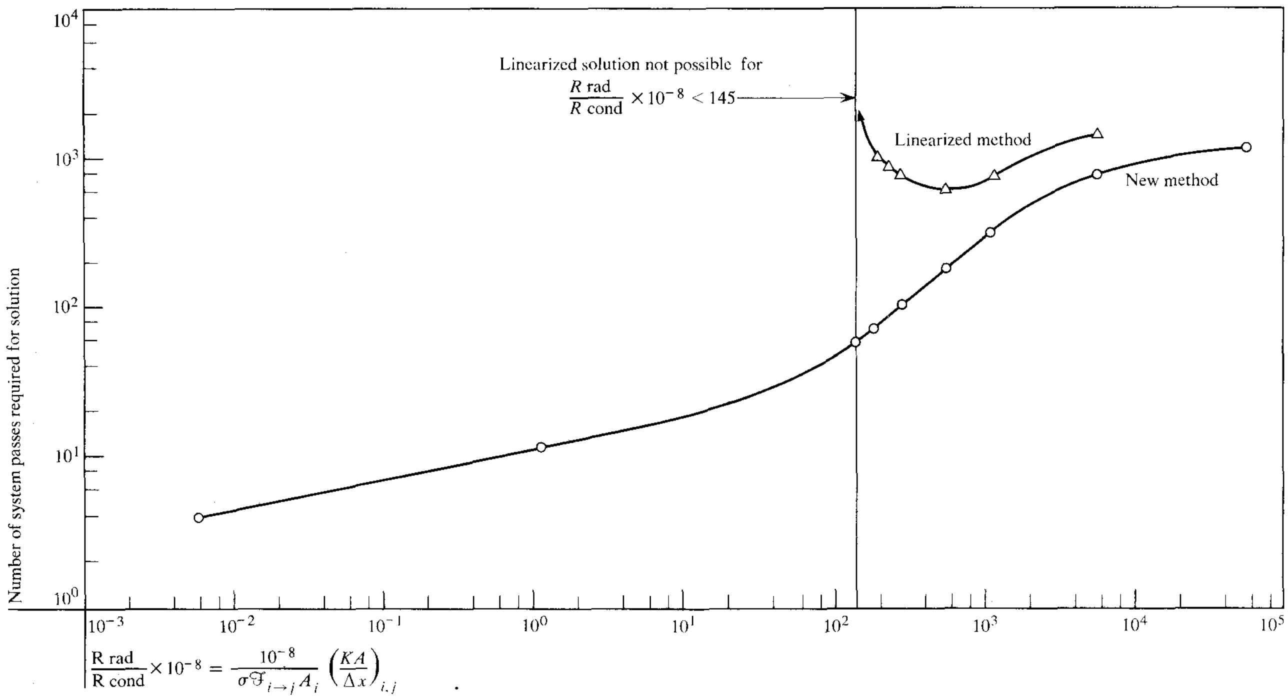


Figure 3 Comparison of new and linearized methods for radiation fin problem. ($T_\infty = 0$°R; $Q_i$gen = 100 Btu/hr; required tolerance on successive passes $T_{i,n+1} - T_{i,n} \leq 0.001$ °R.)

DELTA 2 (another tolerance) on successive system passes. Values of root evaluation tolerance DELTA 1 should be significantly smaller than the final required system tolerance DELTA 2. Test cases indicate that a good empirical rule is to require a root evaluation tolerance about ten times smaller than the system tolerance. It has been found that computation time can be minimized by using several corresponding pairs of values for root evaluation and system tolerances.

Initially, guessed values of temperature for all node points must be provided as input to the program. This results in the "constant" ζ being calculated rather crudely in the early stages of the solution. However, as the relaxation process continues the value of ζ becomes more accurate. What this implies is that initially the value of DELTA 1 should not be as small as later in the solution. It is not useful to calculate roots of the polynomial very accurately until all temperature values have relaxed to a fairly good degree of accuracy.

As an example, let DELTA $1 = 0.1$. Since several values of tolerance are used, this quantity should be indexed. Thus DELTA 1 (1) = 0.1. Select DELTA 2 (1) = 1.0 [notice that DELTA 1 (1) is ten times smaller than DELTA 2 (1)]. The system solution will proceed until the value DELTA 2 (1) is satisfied by all nodes. At that point, both DELTA 1 and DELTA 2 are updated. Next select DELTA 1 (2) = 0.0001 and DELTA 2 (2) = 0.001. These tolerances are then used until final system solution is achieved. From test cases, it appears that little or no further saving in computation time is achieved by using more than two levels of tolerance.

Results of test cases [6]

First, consider a radiation fin that conducts heat along its longitudinal axis and radiates heat to its surroundings at a uniform temperature T_∞. Fifty node points were used to thermally describe the fin, each of which had an assigned quantity of heat generation. The ratio of radiation to conduction resistance was varied over several orders of magnitude to show both the utility and the limitations of the linearized method, as well as the superiority of the new method of solution.

Results are shown in Fig. 3. Notice that the linearized method works satisfactorily for large values of Rrad/Rcond (conduction dominant). For values of Rrad/Rcond less than about 145×10^8, however, no solution was possible for this example with the linearized method. The new method of solution, on the other hand, was completely stable over all ranges of input data. For this particular example the number of iterations for system solution continues to decrease as the conduction resistance is increased. This is because the increased conduction resistance tends to isolate each node from adjacent nodes and allows the final temperature at a node to be fixed by the thermal radiation environment.

Second, an example is included for a three-node, conduction-radiation situation (see Fig. 4), which illustrates the damped oscillation characteristics of the linearized method. Each node point radiates to space (T_∞). Node 2 generates heat and nodes 1 and 3 are connected to node 2 by conduction. For this example the ratio of conduction to radiation heat transfer is about 1.3 for node 2. The numerical results are plotted in Fig. 5. Our method of solution satisfied the convergence requirements in 18 system passes while the linearized method required 82 passes to satisfy the same tolerance. Note the monotonic convergence of the new method, which is also a characteristic of the linear system.

Finally, an example is presented to illustrate an attempt to classify each node point as either conduction or radiation dominant. The problem involves 1000 node points, with each node conducting and radiating to each of its adjacent nodes in three dimensions. Each node has an assigned quantity of heat generation. Results are presented in Fig. 6 and, as would be expected, the dominant-mode method works for extreme ratios of radiation-to-conduction resistance. The significance of the results is that the regions of dominant-mode solution do not overlap. There exists a region ($3 < R\text{rad}/R\text{cond} \times 10^{-8} < 27$) where neither of the dominant-mode methods of solution will solve the problem. The new method, however, provides stable convergence in this region. In fact, we expect that solution is possible with the new method over the semi-infinite region $0 \leq R\text{rad}/R\text{cond} < \infty$.

Summary

Numerical instability associated with the steady-state solution of nonlinear thermal systems can be eliminated by use of a two-level iteration procedure. Because of the diagonal dominance of the coefficient matrices, it is intuitive that this method insures stability and convergence. Test cases substantiate this conclusion, but it remains to be proved in a rigorous manner. Due to the monotonic behavior of the system response, a solution can be obtained in general much more quickly than with previous methods. This characteristic reduces computation time and increases the user's confidence that the final results are in fact a real solution to his problem. The method can be applied to many types of steady-state heat transfer analyses.

Acknowledgment

The programming required to implement the new method of solution within an existing thermal analysis program was done by W. A. Hall, Jr., whose help is gratefully acknowledged.

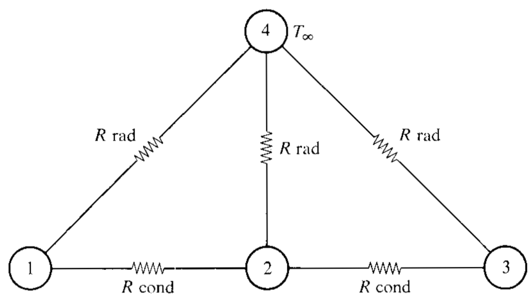


Figure 4 Nodal arrangement for three-point problem.

Figure 5 System response with damped oscillation. ($\mathscr{F}_{i\to j}A_i = 1.0\,\text{ft}^2$; $R\text{cond} = 0.2\,°\text{R/Btu/hr}$; $Q\text{gen}$, node 2 = 1000 Btu/hr.)

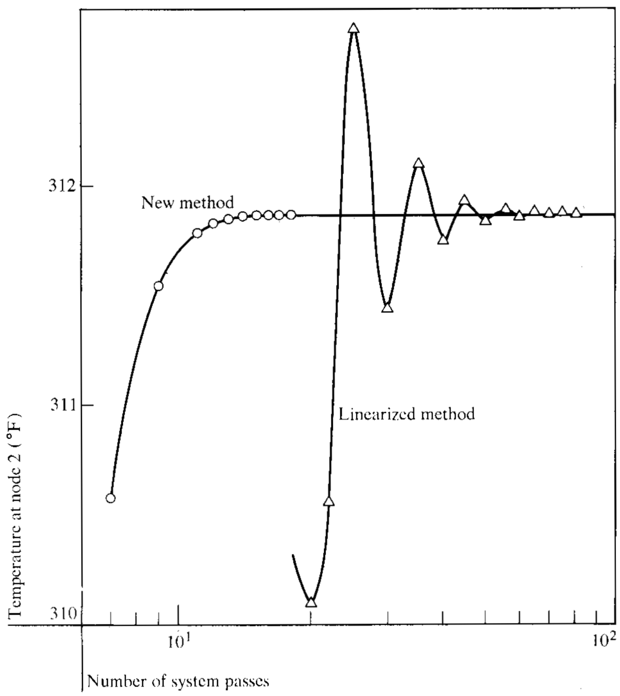
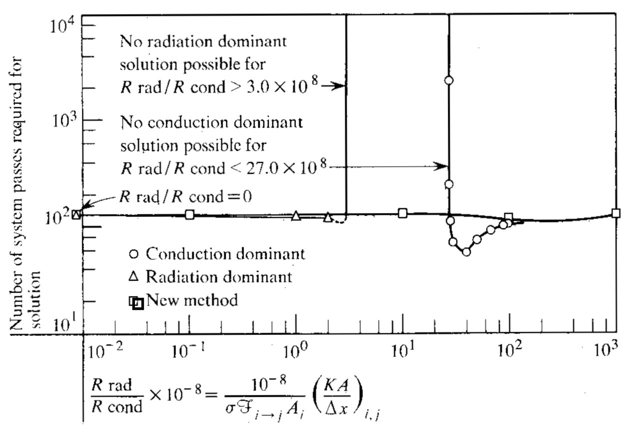


Figure 6 Comparison of new and dominant-mode methods of solution. (1000-node parallelepiped; $Q_i\text{gen} = 1000$ Btu/hr.)



References and notes

1. C. C. Logan, R. B. David and L. H. Michel, "Thermal Analyzer," *Report L. R. 17708*, Lockheed California Company, Burbank, California, March 13, 1964.
2. K. W. Lallier and B. R. Pagnani, "Numerical Analysis for Thermal Applications," *Report 63-825-862*, IBM Corporation, Space Guidance Center, Owego, N.Y., May 1963.
3. J. D. Gaski and D. R. Lewis, "Chrysler Improved Numerical Differencing Analyzer," *TN-AP-66-15*, Chrysler Corporation Space Division, New Orleans, La., April 30, 1966.
4. M. G. Salvadori, M. L. Baron, *Numerical Methods in Engineering*, second edition, Prentice-Hall, Inc., Englewood Cliffs, N.J. 1961, p. 33.
5. G. E. Forsythe, C. B. Moler, *Computer Solution of Linear Algebraic Systems*, Prentice-Hall, Inc., Englewood Cliffs, N.J. 1967, p. 132.
6. The test cases described in this paper were programmed in FORTRAN IV on an experimental basis and executed on an IBM System 360/75.

Received July 16,1971

The author is located at the IBM Data Processing Division Branch Office, P.O. Box 80, Winston-Salem, North Carolina 27106. The work reported in this paper was done at the IBM Systems Development Division Laboratory, San Jose, California 95114.