MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

①

APPLICATIONS OF THE VARIATIONAL INTEGRAL IN

ITERATIVE NUMERICAL SOLUTIONS TO THE

STATIONARY HEAT EQUATIONS

THESIS

Mark L. MacDonald
Second Lieutenant, USAF

AFIT/GNE/ENP/84M-11

DTIC
ELECTE
MAY 1 5 1984
S D
B

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY
AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

84 05 14 117

APPLICATIONS OF THE VARIATIONAL INTEGRAL IN

ITERATIVE NUMERICAL SOLUTIONS TO THE

STATIONARY HEAT EQUATIONS

THESIS

Mark L. MacDonald
Second Lieutenant, USAF

AFIT/GNE/ENP/84M-11

DTIC
SELECTED
MAY 1 5 1984

B

APPLICATIONS OF THE VARIATIONAL INTEGRAL

IN ITERATIVE NUMERICAL SOLUTIONS

TO THE

STATIONARY HEAT EQUATIONS

THESIS

Presented to the Faculty of the School of Engineering

of the Air Force Institute of Technology

Air University

in Partial Fulfillment of the

Requirements for the Degree of

Master of Science in Nuclear Engineering

Mark L. MacDonald, B.S.

Second Lieutenant, USAF

March 1984

Acknowledgments

I would like to express my appreciation to Dr. Bernard Kaplan of the Air Force Insitute of Technology for his guidance and assistance. I would also like to thank Dr. N. Pagano of the Air Force Materials Laboratory for sponsoring this research project.

Mark L. MacDonald

Contents

List of Figures

List of Tables

Abstract

Three stationary heat equations were solved using the finite-difference method. The resulting set of algebraic equations were solved using the Gauss-Seidel iterative technique. The temperatures at the nodal points were substituted into a numerical approximation of the variational integral. The variational integral approximation was used to determine when to stop the iterative process. The variational integral stopping criterion was compared to a stopping criterion that uses the displacement between iterations to approximate the error between the iterative solution and the exact solution. The variational integral was found to be less effective as a stopping criterion than the error estimate.

The variational integral was examined as a method of determining whether the finite-difference technique or the finite-element technique gave a more accurate solution. It was found that the variational integral failed, in some cases, to predict the more accurate method.

I Introduction

Background

In practice, the number of boundary value problems that can be solved analytically is quite limited. Consequently, numerical techniques are used to approximate the partial differential equations associated with the boundary value problems. These numerical techniques often lead to large sets of simultaneous equations. These ec   .iuns are usually sparse (they contain few non-zero elements in the coeffi-cient matrix).

Iterative techniques are often used to solve large sets of simultaneous equations. These techniques have two major advantages over methods that solve the simultaneous equations directly. First, iterative methods avoid calculations with the many zero elements in the coefficient matrix while direct methods perform calculations on every element of the coefficient matrix. Secondly, iterative methods are not as susceptible to the accumulation of round-off error.

An important difficulty with the iterative techniques is determining when a sufficient number of iterations has been completed to ensure an accurate solution. This thesis addresses the problem of when to stop iterating.

Many boundary value problems can be expressed as variational integrals to be minimized. Accordingly, one possible test for convergence of the numerical solution to the boundary value problem is whether it minimizes the associated variational integral.

1

Problem

The primary purpose of this thesis is to determine the usefulness of the variational integral as a test for terminating the iterative process.

The secondary purpose is to determine the usefulness of the variational integral as a means of determining which numerical method gives a more accurate solution.

Scope

This thesis is limited to the study of stationary (time-independent) heat equations in one and two spacial dimensions. The numerical approximations to the boundary value problems consist of the finite-difference and finite-element techniques. The iterative scheme used is the Gauss-Seidel method (method of successive displacements).

Approach and Presentation

Section II introduces the theory necessary to understand the variational integral, the finite-difference method, and the finite-element method. The Gauss-Seidel iterative technique is discussed. The concept of norms is introduced and is applied in examining stopping criteria.

Section III introduces the three stationary heat equations considered in this thesis. The variational integral formulations of the problems are given and the numerical approximation techniques are applied.

Section IV presents the results of the study. First, errors caused by approximating the variational integral are examined. Next, the utility of the variational integral as a stopping criterion is

2

discussed. Finally, the usefulness of the variational integral as a means of determining which numerical method gives a more accurate solution is analyzed.

Section V draws conclusions and gives recommendations for future studies.

<u>Variational Integral</u>

Many problems in the physical sciences can be solved in two ways. First, they can be written as differential equations subject to given boundary conditions. Secondly, they can be written as variational integrals to be extremized (maximized or minimized). The two formulations are equivalent because the function that will extremize the variational integral will also satisfy the differential equation. This equivalence ia demonstrated by the fact that the variational integral is extremized only when an equation, called the Euler-Lagrange equation, is satisfied. This Euler-Lagrange equation is precisely the same as the differential equation formulation of the problem (Ref 8:67). In general, only elliptic differential equations can be written in variational form (Refs 8:117; 5:164) but some exceptions have been found (Ref 7:252-256).

In variational calculus, one looks for the function $t(x,y)$ with continuous second partial derivatives, that extremizes $I(t)$ on region R of the x-y plane:

$$I(t) = \iint_R F(x,y,t,t_x,t_y)\,dxdy \qquad (2.1)$$

where

$$t = t(x,y) \qquad (2.2)$$

$$t_x = \frac{\partial}{\partial x}t(x,y) \qquad (2.3)$$

$$t_y = \frac{\partial}{\partial y}t(x,y) \qquad (2.4)$$

To find $t(x,y)$, examine the set of functions

$$\tilde{t}(x,y) = t(x,y) + \varepsilon\eta(x,y) \qquad (2.5)$$

where ε takes on some value close to zero and $\eta(x,y)$ is an arbitrary function with continuous second partial derivatives. At this point, an important assumption is made. It is assumed that t is given on the boundary as a constant or as a function of x and y. As a consequence, $\tilde{t}$ is not subject to variation on the boundary and $\eta(x,y)$ must equal zero on the boundary.

Substituting Eq (2.5) into Eq (2.1) gives

$$J(\varepsilon) \equiv I(t+\varepsilon\eta) = \iint_R F(x,y,t+\varepsilon\eta,t_x+\varepsilon\eta_x,t_y+\varepsilon\eta_y)dxdy \qquad (2.6)$$

A necessary condition for I to have an extremum at $\tilde{t} = t$ is that $J(\varepsilon)$ must have an extremum at $\varepsilon = 0$. So

$$\frac{\partial}{\partial\varepsilon}J(\varepsilon)\Big|_{\varepsilon=0} = 0 \qquad (2.7)$$

Differentiating Eq (2.6) with respect to ε gives

$$\frac{\partial}{\partial\varepsilon}J(\varepsilon) = \iint_R \left(\frac{\partial F}{\partial\tilde{t}} \cdot \frac{\partial\tilde{t}}{\partial\varepsilon} + \frac{\partial F}{\partial\tilde{t}_x} \cdot \frac{\partial\tilde{t}_x}{\partial\varepsilon} + \frac{\partial F}{\partial\tilde{t}_y} \cdot \frac{\partial\tilde{t}_y}{\partial\varepsilon} \right) dxdy \qquad (2.8)$$

or

$$\frac{\partial}{\partial\varepsilon}J(\varepsilon) = \iint_R \left(\eta\frac{\partial F}{\partial\tilde{t}} + \eta_x\frac{\partial F}{\partial\tilde{t}_x} + \eta_y\frac{\partial F}{\partial\tilde{t}_y} \right) dxdy \qquad (2.9)$$

By the chain rule (Ref 1:92)

$$\frac{\partial}{\partial x}\left(\eta\frac{\partial F}{\partial\tilde{t}_x} \right) = \frac{\partial F}{\partial\tilde{t}_x}\eta_x + \eta\frac{\partial}{\partial x}\left(\frac{\partial F}{\partial\tilde{t}_x} \right) \qquad (2.10)$$

5

therefore

$$\eta_x \frac{\partial F}{\partial \tilde{t}_x} = \frac{\partial}{\partial x}\left(\eta \frac{\partial F}{\partial \tilde{t}_x}\right) - \eta \frac{\partial}{\partial x}\left(\frac{\partial F}{\partial \tilde{t}_x}\right) \tag{2.11}$$

and

$$\eta_y \frac{\partial F}{\partial \tilde{t}_y} = \frac{\partial}{\partial y}\left(\eta \frac{\partial F}{\partial \tilde{t}_y}\right) - \eta \frac{\partial}{\partial y}\left(\frac{\partial F}{\partial \tilde{t}_y}\right) \tag{2.12}$$

Substituting Eqs (2.11) and (2.12) into Eq (2.9) gives

$$\frac{\partial}{\partial \varepsilon}J(\varepsilon) = \iint_R \eta \left[\frac{\partial F}{\partial \tilde{t}} - \frac{\partial}{\partial x}\left(\frac{\partial F}{\partial \tilde{t}_x}\right) - \frac{\partial}{\partial y}\left(\frac{\partial F}{\partial \tilde{t}_y}\right)\right] dxdy$$

$$+ \iint_R \left[\frac{\partial}{\partial x}\left(\eta \frac{\partial F}{\partial \tilde{t}_x}\right) + \frac{\partial}{\partial y}\left(\eta \frac{\partial F}{\partial \tilde{t}_y}\right)\right] dxdy \tag{2.13}$$

By applying Green's Theorem (Ref 1:197), the second integral in Eq (2.13) can be transformed into a line integral

$$\frac{\partial}{\partial \varepsilon}J(\varepsilon) = \iint_R \eta \left[\frac{\partial F}{\partial \tilde{t}} - \frac{\partial}{\partial x}\left(\frac{\partial F}{\partial \tilde{t}_x}\right) - \frac{\partial}{\partial y}\left(\frac{\partial F}{\partial \tilde{t}_y}\right)\right] dxdy$$

$$+ \oint_L \left[\left(\eta \frac{\partial F}{\partial \tilde{t}_x}\right) dy - \left(\eta \frac{\partial F}{\partial \tilde{t}_y}\right) dx\right] \tag{2.14}$$

But since $\eta(x,y) = 0$ ⌐ ⌐e boundary, C

$$\frac{\partial}{\partial \varepsilon}J(\varepsilon) = \iint_R \eta \left[\frac{\partial F}{\partial \tilde{t}} - \frac{\partial}{\partial x}\left(\frac{\partial F}{\partial \tilde{t}_x}\right) - \frac{\partial}{\partial y}\left(\frac{\partial F}{\partial \tilde{t}_y}\right)\right] dxdy \tag{2.15}$$

6

Recall that

$$\frac{\partial}{\partial \epsilon} J(\epsilon) \Big|_{\epsilon=0} = 0 \qquad (2.7)$$

if I is to have an extremum at $\tilde{t} = t$. Therefore

$$\iint_{R} \eta \left[\frac{\partial F}{\partial t} - \frac{\partial}{\partial x} \left(\frac{\partial F}{\partial t_x} \right) - \frac{\partial}{\partial y} \left(\frac{\partial F}{\partial t_y} \right) \right] dxdy = 0 \qquad (2.16)$$

Since η is an arbitrary function of x and y then

$$\frac{\partial F}{\partial t} - \frac{\partial}{\partial x} \left(\frac{\partial F}{\partial t_x} \right) - \frac{\partial}{\partial y} \left(\frac{\partial F}{\partial t_y} \right) = 0 \qquad (2.17)$$

Eq (2.17) is the Euler-Lagrange equation and it must be satisfied for t to extremize I(t). In the derivation that lead to Eq (2.17), it was assumed that the function, t, was specified on the boundaries (Dirichlet boundary conditions). Other boundary conditions could produce a change in the integral that is to be extremized. A method of determining the variational integral for other boundary conditions is given in Mikhlin (Ref 12:116-121). For all the problems considered in this thesis, the Euler-Lagrange equation given as Eq (2.17) will give us a function, F that extremizes I(t) as given in Eq (2.1).

Finite-Element Method

The finite element method involves extremizing the variational formulation of the problem. The solution region is divided into interconnected subdomains — the finite elements. The elements in one-dimensional problems are line segments. The most versatile elements in higher dimensions are triangles for two-dimensional problems and tetrahedrons for three-dimensional problems (Ref 8:86). Next, iterpolation functions are chosen to represent the temperatures

7

over each element. These interpolation functions are often poly-nomials. Suppose the function being integrated under the variational integral contains derivatives up to the (r+1)th order. Then, the interpolation functions must have rth order continuous derivatives at the element interfaces and (r+1)th order continuous derivatives within each element (Ref 8:124). When these criteria are met, the variational integral for the entire solution region can be written as the sum of the variational integrals of each element (Ref 8:78).

An interpolation function for the temperatures on an element is written in terms of the nodal values associated with that element. The variational integral is evaluated over the element and the integral is differentiated with respect to the nodes and set equal to zero in order to extremize the variational integral on that element. The equations for each element are assembled to give the equations governing the temperature over the entire region. In this manner, a set of equations is created which can be solved simultaneously to find the temperatures at the element nodes. This set of simultaneous equations can be written in matrix form

$$\underline{A}\underline{t} = \underline{b} \tag{2.18}$$

where

$\underline{A}$ = a coefficient matrix

$\underline{t}$ = vector of unknown nodal temperatures

$\underline{b}$ = vector of known constants

Rather than go into the details of the process here, an example is given in Appendix A.

8

Finite-Difference Method

A differential equation can be converted into finite-difference form by approximating the derivatives in the equation by differences. The elementary definition of a derivative is (Ref 3:54)

$$\frac{dt}{dx} = \lim_{\Delta x \to 0} \frac{t(x + \Delta x) - t(x)}{\Delta x} \tag{2.19}$$

The derivative is approximated by a divided difference

$$\frac{dt}{dx} \approx \frac{t(x + \Delta x) - t(x)}{\Delta x} \tag{2.20}$$

In one-dimensional problems, the solution region is broken into line segments and a nodal point and its neighboring nodal points are used to approximate the derivative at that nodal point. In this manner, a linear difference equation is created for each nodal point. The equations for the individual nodal points can be assembled to give a set of algebraic equations to be solved simultaneously. Linear combinations of Taylor series expansions give several difference formulas with their associated truncation error (Ref 3:55-59):

First forward difference

$$\frac{dt}{dx} = \frac{t(x + h) - t(x)}{h} + O(h) \tag{2.21}$$

First central difference

$$\frac{dt}{dx} = \frac{t(x + h) - t(x - h)}{2h} + O(h^2) \tag{2.22}$$

First backward difference

$$\frac{dt}{dx} = \frac{t(x) - t(x - h)}{h} + O(h) \tag{2.23}$$

Second forward difference

$$\frac{d^2t}{dx^2} = \frac{t(x + 2h) - 2t(x + h) + t(x)}{h^2} + O(h) \tag{2.24}$$

Second central difference

$$\frac{d^2t}{dx^2} = \frac{t(x + h) - 2t(x) + t(x - h)}{h^2} + O(h^2) \tag{2.25}$$

Second backward difference

$$\frac{d^2t}{dx^2} = \frac{t(x) - 2t(x - h) + t(x - 2h)}{h^2} + O(h) \tag{2.26}$$

where h is the distance between nodal points. Truncation error is a result of ignoring higher order terms in the Taylor series. The finite-difference method is also subject to something called round-off error. Round-off error is the result of the fact that a computer can store a limited number of significant figures. The total error due to both truncation and round-off error will be referred to as the discretization error. It is the difference between the analytical solution and the finite-difference solution at a particular nodal point.

For two-dimensional problems, a rectangular mesh is imposed over the solution region in order to find approximate temperatures at the nodal point of the mesh.

10
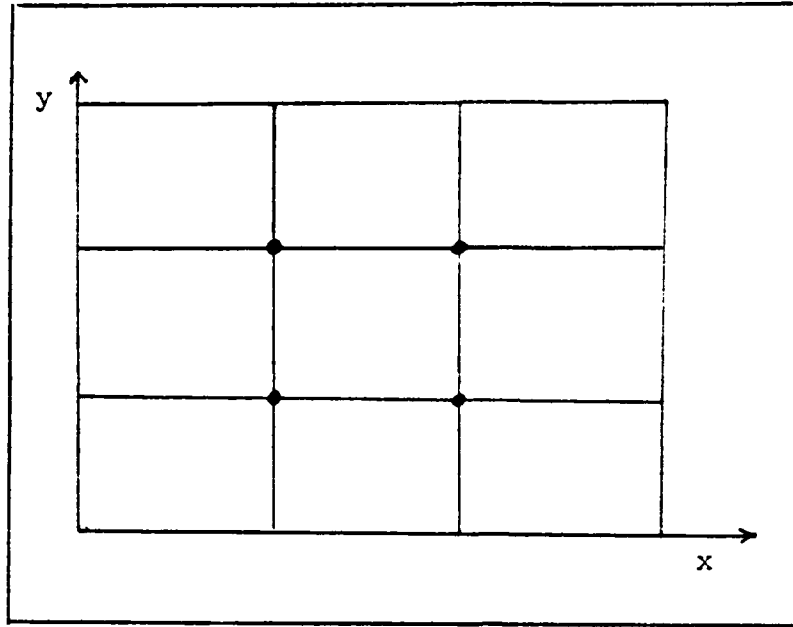
Figure 2.1. Mesh for a Rectangular Region

Solution of Simultaneous Equations

Both the finite-element and finite-difference methods lead to a
set of simultaneous liner algebraic equations. There are two methods
of solving for the unknowns in these equations — direct methods and
indirect methods. Direct methods lead to an exact answer, excluding
round-off error, after a finite number of arithmetic operations.
Direct methods tend to be time consuming since each element in the
coefficient matrix must be operated upon. In general, N^3 operations
are required where N is the number of equations (Refs 3:111; 5:209).
Some examples of exact methods are Gaussian Elimination and the Thomas
method (Ref 3:9-12; 44-48).

Indirect iterative methods require an infinite number of opera-
tions to solve the equations exactly, but an adequate solution can

11

usually be achieved in a finite number of steps. These methods usually require less than N^3 operations, especially if there are many zeros in the coefficient matrix as is the case in the finite-element and finite-difference equations. Furthermore, cumulative round-off error does not grow as in the direct methods (Ref 3:111).

The matrix equation

$$\underline{At} = \underline{b} \tag{2.18}$$

can be re-written

$$\underline{t} = \underline{Gt} + \underline{c} \tag{2.27}$$

An iterative solution can be written in the form

$$\underline{t}^{(n+1)} = \underline{Gt}^{(n)} + \underline{c} \tag{2.28}$$

where

$\underline{t}^{(n)}$ = nth iterated vector

$\underline{G}$ = iteration matrix

$\underline{c}$ = known constant vector

The three most commonly used iteration methods are the Jacobi, Gauss-Seidel, and successive over-relaxation methods. They are derived by factoring the coefficient matrix into strictly upper triangular, diagonol and lower triangular matricies denoted by $\underline{U}$, $\underline{D}$, and $\underline{L}$ respectively (Ref 3:46-47):

$$(\underline{U} + \underline{D} + \underline{L})\underline{t} = \underline{b} \tag{2.29}$$

The Gauss-Seidel Method (also known as the method of successive displacements) is used in this thesis. It expresses the (n+1)th

12

iterative vector in terms of both the nth iterative vector elements and the (n+1)th iterative vector elements as soon as they become available:

$$(\underline{L} + \underline{D})\underline{t}^{(n+1)} = -\underline{U}\underline{t}^{(n)} + \underline{b} \tag{2.30}$$

or written in a form similar to Eq (2.27)

$$\underline{t}^{(n+1)} = -(\underline{L} + \underline{D})^{-1}\underline{U}\underline{t}^{(n)} + (\underline{L} + \underline{D})^{-1}\underline{b} \tag{2.31}$$

Translated into a computational algorithm:

$$t_i^{(n+1)} = -\frac{1}{a_{ii}}\left[\sum_{j=1}^{i-1}a_{ij}t_j^{(n+1)} + \sum_{j=i+1}^{n}a_{ij}t_j^{(n)} - b_i\right] \tag{2.32}$$

where a_{ij} is the element in the ith row and the jth column of the coefficient matrix. One begins the process by choosing some arbitrary initial trial vector, $\underline{t}^{(0)}$.

Norms

In the following sections, we will be concerned with measuring the difference between exact temperatures derived analytically and the temperatures found from the iterative process. The error vector is defined as

$$\underline{e}^{(n)} = \underline{t} - \underline{t}^{(n)} \tag{2.33}$$

where

$\underline{e}^{(n)}$ = the error vector for the nth iteration

$\underline{t}$ = vector of exact nodal temperatures

$\underline{t}^{(n)}$ = vector of nodal temperatures from the nth iteration

The displacement vector (also known as the "residual vector") is defined as

13

$$\underline{d}^{(n)} = \underline{t}^{(n+1)} - \underline{t}^{(n)} \qquad (2.34)$$

where

$\underline{t}^{(n+1)}$ = vector of nodal temperatures from the (n+1)th iteration

$\underline{t}^{(n)}$ = vector of temperatures from the nth iteration

It is often more convenient to measure the size of these vectors as scalar quantities. A norm, symbolized by $||\cdot||$, will allow one to do just that. A vector n-norm can be defined by

$$||\underline{t}||_n = (\sum_i |t_i|^n)^{1/n} \qquad (2.35)$$

where

$\underline{t}$ = vector

t_i = ith element of vector t

The norm used in this thesis is

$$||\underline{t}||_1 = \sum_i |t_i| \qquad (2.36)$$

Stopping Criteria

Standard Method. One of the fundamental difficulties with iterative methods is determining when to stop iterating. Some common criteria are (Ref 4:227)

$$\left|\left|\underline{t}^{(n)} - \underline{t}^{(n-1)}\right|\right| < \varepsilon \qquad (2.37)$$

or

$$\left|\left|\frac{\underline{t}^{(n)} - \underline{t}^{(n-1)}}{\underline{t}^{(n)}}\right|\right| < \varepsilon \qquad (2.38)$$

for some prescribed ε or

14

$$n > N \qquad (2.39)$$

for some given N.

The problem with the above criteria is that

$$\lVert \underline{t}^{(n)} - \underline{t}^{(n-1)} \rVert < \varepsilon \qquad (2.37)$$

does not imply that

$$\lVert t^{(n)} - t \rVert < \varepsilon \qquad (2.40)$$

Figure 2.2 illustrates this fact (Ref 9:5). Thus, for problems where convergence is slow, $\underline{t}^{(n)}$ may be far from the exact temperature even though it would have eventually been reached. However, choosing an iteration limit that is to stringent might lead to a considerable waste of computer time.

A simple technique has been found that will allow one to use the displacement vector, $\lVert \underline{d}^{(n)} \rVert$ to estimate the error, $\lVert \underline{e}^{(n)} \rVert$ (Ref 14). Recall that the matrix equation can be written as

$$\underline{t} = \underline{G}\underline{t} + \underline{c} \qquad (2.27)$$

And an iterative solution can be written as

$$\underline{t}^{(n+1)} = \underline{G}\underline{t}^{(n)} + \underline{c} \qquad (2.28)$$

By subtract Eq (2.28) from Eq (2.27), the error vector can be written

$$\underline{e}^{(n+1)} = \underline{G}\underline{e}^{(n)} \qquad (2.41)$$

The following discussion holds for all real symmetric and some real nonsymmetric matricies (Ref 3:27). Assume that the initial error

15
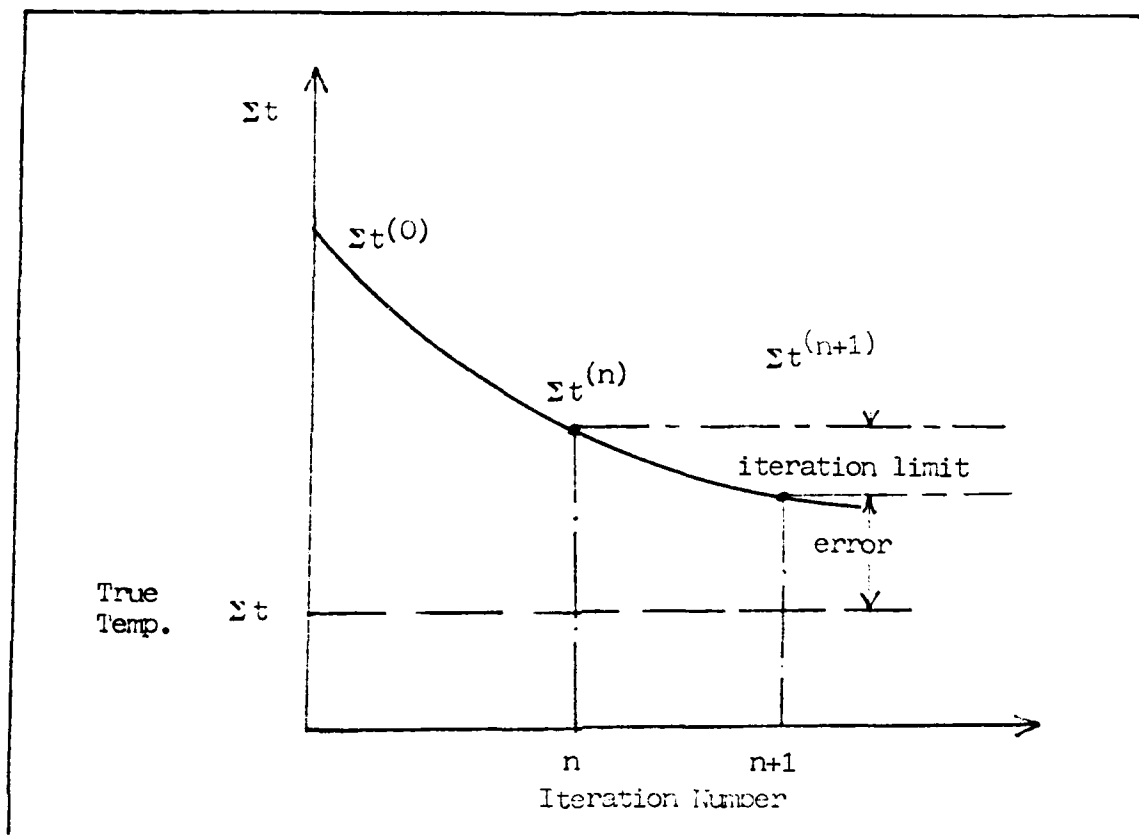
Figure 2.2. Iteration Limit.

vector can be expanded in terms of eigenvectors $\underline{V}_i$ of the iteration matrix $\underline{G}$ (Ref 14:236)

$$\underline{e}^{(0)} = \sum_{i=1}^{m} C_i \underline{V}_i \qquad (2.42)$$

where C_i are scalars and m is the dimension of the square coefficient matrix. From Eq (2.41)

$$\underline{e}^{(n)} = \underline{G}^n \underline{e}^{(0)} \qquad (2.43)$$

where $\underline{G}^n$ is $\underline{G}$ raised to the nth power.

$$\underline{e}^{(n)} = \underline{G}^n \sum_{i=1}^{m} C_i \underline{V}_i \qquad (2.44)$$

$$\underline{e}^{(n)} = \sum_{i=1}^{m} C_i \underline{G}^n \underline{V}_i \qquad (2.45)$$

But $\underline{GV}_i = \lambda_i \underline{V}_i$ by the definition of an eigenvalue where λ_i is an eigenvalue corresponding to $\underline{V}_i$. Then

$$\underline{e}^{(n)} = \sum_{i=1}^{m} C_i \lambda_i^n \underline{V}_i \qquad (2.46)$$

$$\underline{e}^{(n)} = \lambda_1^n \left\{ C_1 \underline{V}_1 + \left(\frac{\lambda_2}{\lambda_1} \right)^n C_2 \underline{V}_2 + \ldots + \left(\frac{\lambda_m}{\lambda_1} \right)^n C_m \underline{V}_m \right\} \qquad (2.47)$$

If $|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \ldots \geq |\lambda_m|$

then for large sufficiently large n

$$\underline{e}^{(n)} \approx \lambda_1^n C_1 \underline{V}_1 \qquad (2.48)$$

$$\underline{e}^{(n+1)} \approx \lambda_1^{n+1} C_1 \underline{V}_1 \qquad (2.49)$$

so

$$\underline{e}^{(n+1)} \approx \lambda_1 \underline{e}^{(n)} \qquad (2.50)$$

In order for the error vector to vanish

$$|\lambda_1| < 1 \qquad (2.51)$$

$|\lambda_1|$ is called the spectral radius. Eq (2.50) shows that if λ_1 is small, convergence occurs rapidly. In general, λ_1 increases for a larger number of unknowns (Ref 3:116; 16:87). Thus, a problem takes more iterations to solve as one increases the number of nodal points used in the finite-difference and finite-element methods.

17

From Eq (2.33)

$$\underline{t} = \underline{t}^{(n)} + \underline{e}^{(n)} \qquad (2.52)$$

and from Eq (2.50)

$$\underline{t} \approx \underline{t}^{(n+1)} + \lambda_1 \underline{e}^{(n)} \qquad (2.53)$$

eliminating $\underline{e}^{(n)}$ between Eq (2.52) and Eq (2.53) gives

$$\underline{t} \approx \frac{\underline{t}^{(n+1)} - \lambda_1 \underline{t}^{(n)}}{1 - \lambda_1} \qquad (2.54)$$

or

$$\underline{t} \approx \underline{t}^{(n)} + \frac{\underline{t}^{(n+1)} - \underline{t}^{(n)}}{1 - \lambda_1} \qquad (2.55)$$

From Eq (2.55), it can be seen that, if λ_1 is close to one, then small differences between successive iterations does not imply that the iterative solution is close to the exact solution. But, given the definition of the error vector and the displacement vector, the error vector can be approximated by

$$\underline{e}^{(n)} \approx \frac{\underline{d}^{(n)}}{1 - \lambda_1} \qquad (2.56)$$

or

$$||\underline{e}^{(n)}|| \approx \frac{||\underline{d}^{(n)}||}{|1 - \lambda_1|} \qquad (2.57)$$

18

For most problems λ_1 will not be known but can be approximated in the following manner (Ref 14:241). For sufficiently large n

$$\underline{e}^{(n)} \approx \lambda_1 \underline{e}^{(n-1)} \tag{2.58}$$

$$\underline{e}^{(n+1)} - \underline{e}^{(n)} \approx \lambda_1 (\underline{e}^{(n)} - \underline{e}^{(n-1)}) \tag{2.59}$$

$$\underline{t}^{(n+1)} - \underline{t}^{(n)} \approx \lambda_1 (\underline{t}^{(n)} - \underline{t}^{(n-1)}) \tag{2.60}$$

$$\underline{d}^{(n)} \approx \lambda_1 \underline{d}^{(n-1)} \tag{2.61}$$

or

$$|\lambda_1| \approx \frac{||\underline{d}^{(n)}||}{||\underline{d}^{(n-1)}||} \tag{2.62}$$

Therefore

$$||\underline{e}^{(n)}|| \approx \frac{||\underline{d}^{(n)}||}{\left(1 - \frac{||\underline{d}^{(n)}||}{||\underline{d}^{(n-1)}||}\right)} \tag{2.63}$$

One now has a simple means of estimating the error between the true solution and the iterative solution at the nth iteration and, therefore, a criterion for stopping the iterative process. The variational integral will be judged against this criterion. Young provides a more refined version of this error approximation stopping criterion (Ref 17:226-227). Warren gives a more detailed analysis of this and related error estimates (Ref 16).

Variational Integral. Since the variational integral is extremized for the exact solution to the boundary value problem, it is

19

thought that the nodal temperatures from the iterative method can be substituted into a numerical approximation of the variational integral. As the nodal temperatures approach the exact values at the nodal points, the variational integral should be extremized. Thus, an error criterion might be

$$|I^{(n)} - I^{(n-1)}| < \varepsilon \tag{2.64}$$

for some prescribed ε. This criterion might not have the difficulty that the standard criteria had for slowly converging problems as illustrated in Figure 2.3.
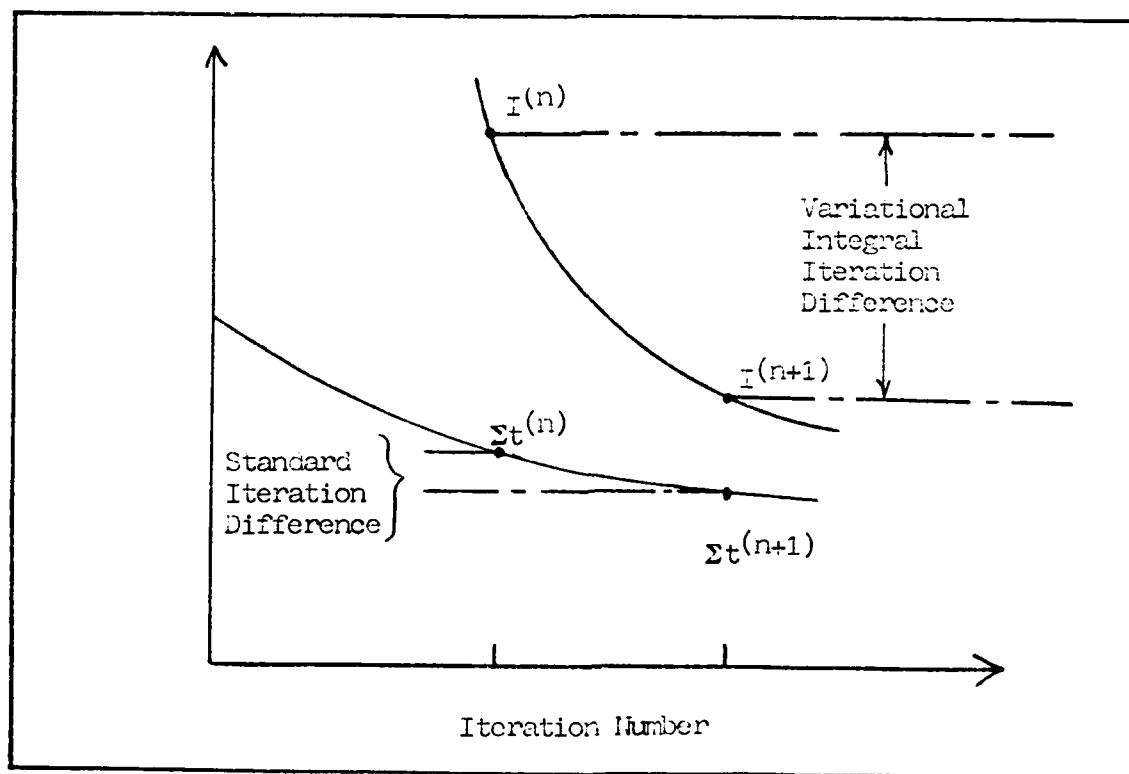
Figure 2.3. Hypothetical Iteration Limit Comparison

It is also thought that the variational integral can be used as a stopping criteria for a special case of problems where the temperatures from the iterative technique are either mostly above or mostly below the exact values. Suppose the initial trial solution for nodal temperatures is above the exact values but the iterative method converges to temperatures below the exact values; then, at some iteration, the iterative solution will pass through the exact solution and the variational integral will be extremized. In practice, one will not know the exact values but an initial trial solution can be chosen that is known to be above the true solution and then an initial trial solution can be chosen that is known to be below the true solution. One of these trial solutions will pass through the exact solution.

III Procedure

In this section, the three heat equations used in this thesis and the computer codes used to evaluate the problems are examined.

One-Dimensional Heat Equation

The problem to be examined is shown in Figure 3.1. The equation to be solved is

$$\frac{d^2t}{dx^2} - m^2t = 0 \tag{3.1}$$

where m is a constant. The boundary conditions are

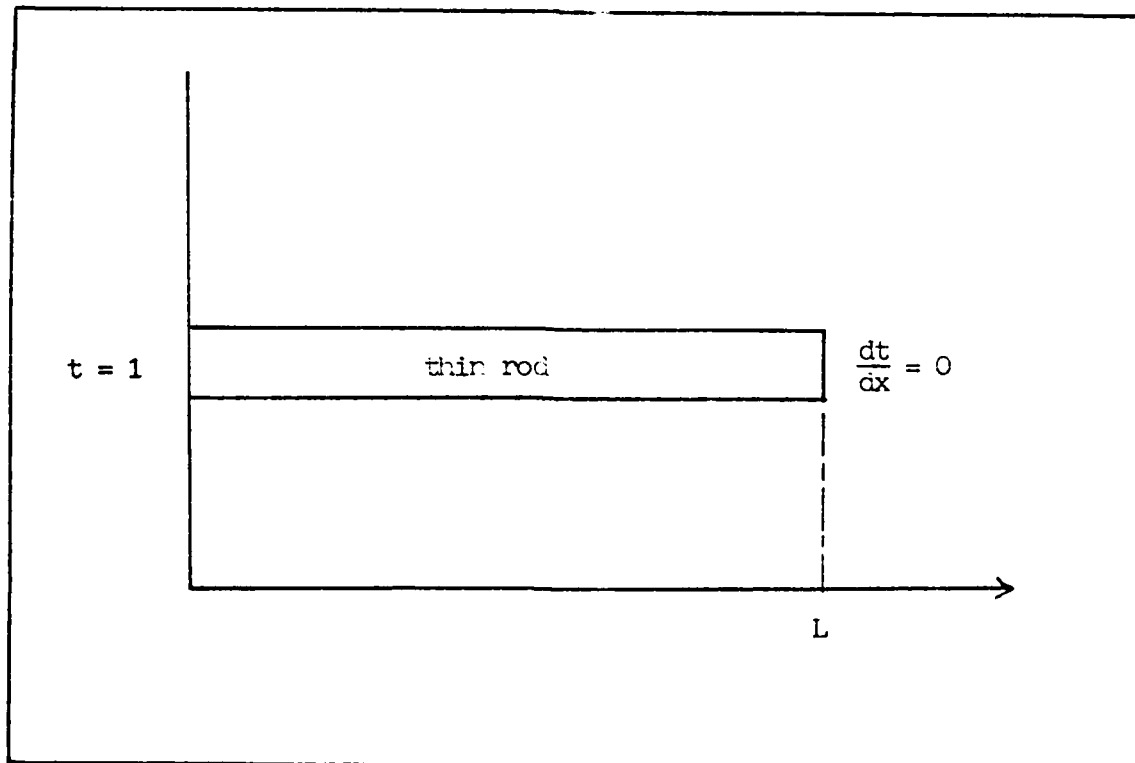$$t(0) = 1 \tag{3.2}$$

$$\frac{d}{dx}t(L) = 0 \tag{3.3}$$

Figure 3.1. Heat Conduction in a Thin Rod.

The equation is easily solved (Ref 11:21):

$$t = \frac{\cosh[m(L - x)]}{\cos[(mL)]} \tag{3.4}$$

Variational Integral. The variational integral for the one-dimensional heat equation can be developed by comparing the governing differential equation (Eq (3.1)) with the Euler-Lagrange equation in one dimension:

$$\frac{\partial F}{\partial t} - \frac{d}{dx}\left(\frac{\partial F}{\partial t_x}\right) = 0 \tag{3.5}$$

The result of the comparison is that

$$F = \frac{1}{2}m^2t^2 + \frac{1}{2}\left(\frac{dt}{dx}\right)^2 \tag{3.6}$$

Thus, the integral to be extremized can be written

$$I(t) = \frac{1}{2}\int_o^L \left[\left(\frac{dt}{dx}\right)^2 + m^2t^2\right]dx \tag{3.7}$$

In order to use the variational integral when only the nodal temperatures are known, the derivative in Eq (3.7) is approximated by a forward difference and the integral is approximated by a sum (Ref 10:235):

$$I(t) \approx \frac{1}{2}\sum_{i=0}^N \left[\left(\frac{t_{i+1} - t_i}{x_{i+1} - x_i}\right)^2 + m^2t_i^2\right](x_{i+1} - x_i) \tag{3.8}$$

One can use the boundary condition given in Eq (3.3) and the forward difference approximation given by Eq (2.21) to show that for $i = N$,

23

$$t_{N+1} = t_N \qquad (3.9)$$

For the purpose of validating the computer code, the variational integral was solved analytically in Appendix A:

$$I(t) = \tanh(mL) \qquad (3.10)$$

Finite-Difference Formulation. The thin rod shown in Figure 3.1 is divided with a number of equally spaced nodes as shown in Figure 3.2.
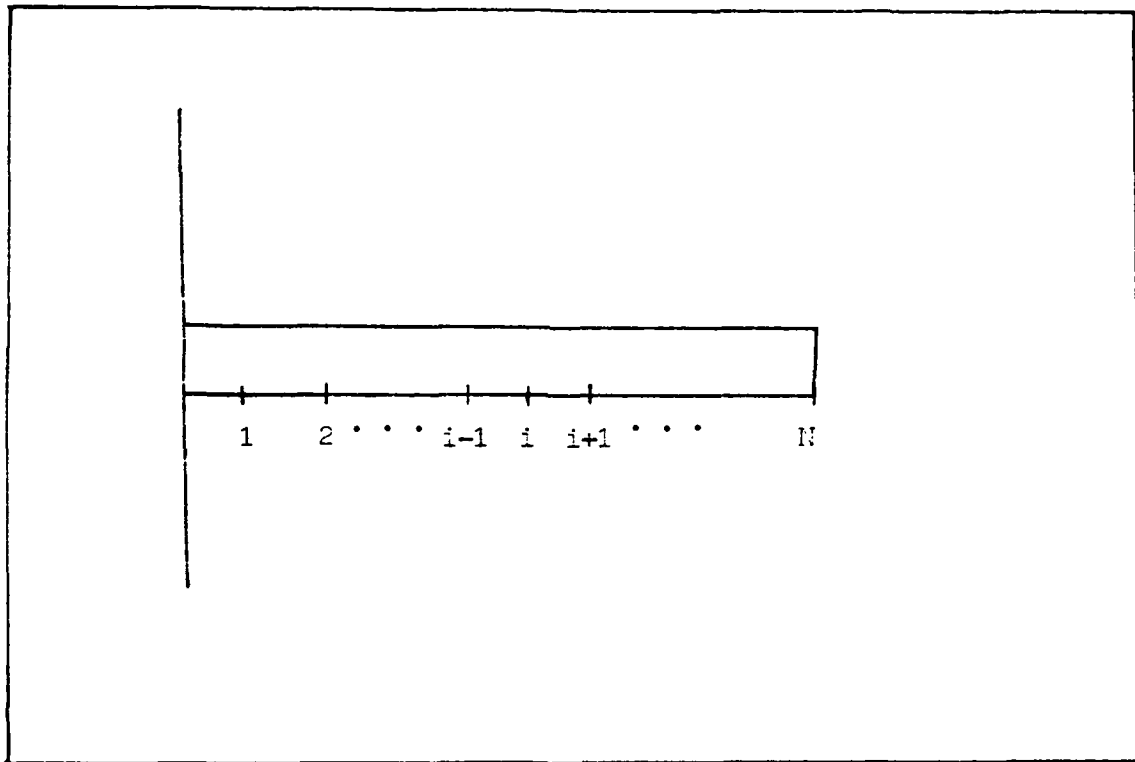
Figure 3.2. Nodal Divisions for Thin Rod.

The differential equation is approximated using Eq (2.25):

$$\frac{t_{i+1} - 2t_i + t_{i-1}}{h^2} + m^2 t_i = 0 \qquad (3.11)$$

24

where

$$h = x_{i+1} - x_i \qquad (3.12)$$

Eq (3.11) can also be written as

$$-t_{i-1} + Dt_i - t_{i+1} = 0 \qquad (3.13)$$

where

$$D = 2 + (mhL)^2 \qquad (3.14)$$

Using the boundary condition given as Eq (3.3) and the central difference approximation given by Eq (2.22),

$$t_{N+1} = t_{N-1} \qquad (3.15)$$

thus Eq (3.13) for $i = N$ is given by

$$-2t_{N-1} + Dt_N = 0 \qquad (3.16)$$

and using the boundary condition given as Eq (3.2)

$$Dt_1 - t_2 = 1 \qquad (3.17)$$

Therefore, for a problem with four unknown nodal temperatures, the simultaneous equations can be written in matrix form as shown in Figure 3.3.

$$\begin{bmatrix} D & -1 & 0 & 0 \\ -1 & D & -1 & 0 \\ 0 & -1 & D & -1 \\ 0 & 0 & -2 & D \end{bmatrix} \begin{bmatrix} t_1 \\ t_2 \\ t_3 \\ t_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Figure 3.3. Matrix Equation for Finite-Difference Approximation, Four Nodes.

25

In order to validate the computer code, $m = 2$ and $L = 1$ so that the temperatures could be compared with those given in the literature (Ref 11:243).

Finite-Element Formulation. The finite-element method is more involved than the finite-difference method and the derivation of the matrix equation is given in Appendix B. The matrix formulation of the problem turns out to be identical to the finite-difference formulation (Figure 3.3) except that

$$D = \frac{12 + 4(mhL)^2}{6 - (mhL)^2} \tag{3.18}$$

Two-Dimensional Poisson's Equation

The problem to be examined is a square plate with uniform energy generation across the plate. The equation to be solved is

$$\frac{\partial^2 t}{\partial x^2} + \frac{\partial^2 t}{\partial y^2} + \frac{g}{k} = 0 \tag{3.19}$$

where g and k are constants. The boundary conditions are

$$\frac{\partial t}{\partial x}(0,y) = 0 \tag{3.20}$$

$$\frac{\partial t}{\partial y}(x,0) = 0 \tag{3.21}$$

$$t(L,y) = 0 \tag{3.22}$$

$$t(x,L) = 0 \tag{3.23}$$

26

The analytic solution is rather involved and is developed in Appendix C. For the validation purposes, g, k, and L were set equal to one so that the numerical solutions could be compared with the literature (Ref 11:260,384).

Variational Integral. The variational integral is developed by comparing the differential equation with the Euler-Lagrange equation in two dimensions (Eq 2.17)). The integral is given by

$$I(t) = \frac{1}{2}\iint_R \left[\left(\frac{\partial t}{\partial x}\right)^2 + \left(\frac{\partial t}{\partial y}\right)^2 - \frac{2gt}{K}\right]dxdy \qquad (3.25)$$

The integral is approximated by sums over both the x and y directions. The derivatives are approximated by backwards differences in order to take advantage of the derivative boundary conditions in the same way that the forward difference was used to take advantage of the derivative boundary condition in the one-dimensional problem (pg. 23):

$$I(t) \approx \frac{1}{2}\sum_{i=0}^{N}\sum_{j=0}^{M}\left[\left(\frac{t_{i,j} - t_{i-1,j}}{\Delta x}\right)^2\right.$$

$$\left. + \left(\frac{t_{i,j} - t_{i,j-1}}{\Delta y}\right)^2 - \frac{2g}{K}t_{i,j}\right]\Delta x\Delta y \qquad (3.26)$$

In addition, the variational integral was solved analytically for computer code validation. The derivation is very similar to the derivation for the one-dimensional problem given in Appendix A.

Finite-Difference Formulation. The square plate is overlayed by a system of square meshes each of dimension h by h.

27

The heat equation Eq (3.19)) can be approximated using Eq (2.25):

$$\frac{1}{h^2}(t_{i+1,j} - 2t_{i,j} + t_{i-1,j}) + \frac{1}{h^2}(t_{i,j+1} - 2t_{i,j} + t_{i,j-1}) + \frac{g}{K} = 0 \quad (3.27)$$

where

$$x = ih \qquad (3.28)$$

$$y = jh \qquad (3.29)$$

The subscripts must be combined before the function, t, can be represented by a vector. This is easily done by incrementing the subscript j through its range of values while, for each j, incrementing i through its range of values. A matrix equation is found for the two-dimensional problem in a manner very similar to that of the one-dimensional problem discussed on page 24.

Finite-Element Formulation. Triangular elements are used in the two-dimensional problem. If one uses the same nodal points as used in the finite-difference scheme, then the elements can be oriented in two different fashions. The orientation in Figure 3.4 will be referred to as "Case One" and the orientation in Figure 3.5 will be referred to as "Case Two." The two arrangements lead to different matrix equations and, consequently, different solutions.

Two-Dimensional Laplace's Equation

The problem to be considered is a square plate

$$\frac{\partial^2 t}{\partial x^2} + \frac{\partial^2 t}{\partial y^2} = 0 \qquad (3.30)$$

28

subject to the boundary conditions

$$t(x,0) = 0 \qquad (3.31)$$

$$t(0,y) = 0 \qquad (3.32)$$

$$t(x,L) = 0 \qquad (3.33)$$
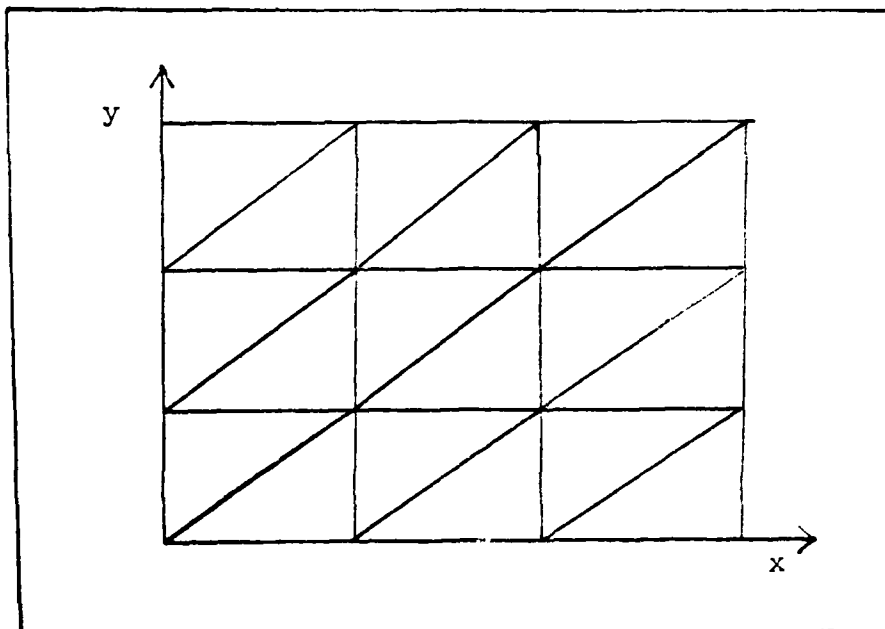
$$t(L,y) = 100 \qquad (3.34)$$

Figure 3.4. Finite-Element Arrangement (Case One).

29
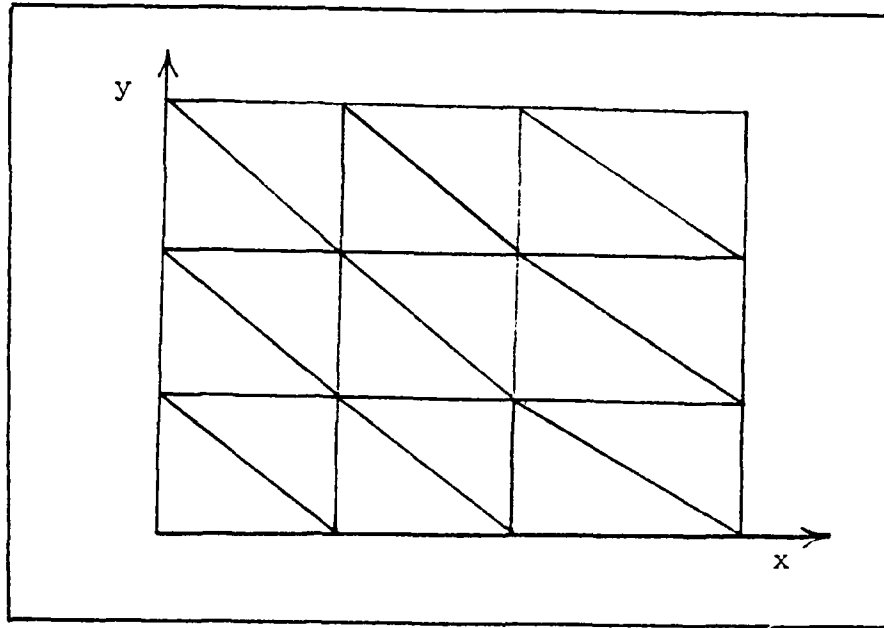
Figure 3.5. Finite-Element Arrangement (Case Two).

The problem is a rather standard boundary value problem. Therefore, the analytic solution will not be derived in this thesis. The solution is given by (Ref 6:107):

$$t(x,y) = \sum_{n=1}^{\infty} \frac{200\left(1 - \cos(n\pi)\right) \sinh\left(\frac{n\pi x}{L}\right) \sin\left(\frac{n\pi y}{L}\right)}{n\pi \sinh(n\pi)} \qquad (3.35)$$

Variational Integral. The variational integral is developed by comparing the differential equation with the Euler-Lagrange equation. The integral is given by

$$I(t) = \frac{1}{2} \iint_R \left[\left(\frac{\partial t}{\partial x}\right)^2 + \left(\frac{\partial t}{\partial y}\right)^2 \right] dxdy \qquad (3.36)$$

30

The integral is approximated by sums over both the x and y directions. The derivations are approximated by central differences.

$$I(t) \approx \frac{1}{2} \sum_{i=0}^{N} \sum_{j=0}^{M} \left[\left(\frac{t_{i+1,j} - t_{i-1,j}}{2h} \right)^2 + \left(\frac{t_{i,j+1} - t_{i,j-1}}{2h} \right)^2 \right] h^2 \quad (3.37)$$

Finite-Difference Formulation. The square plate is overlayed by a system of square meshes and the problem is solved in the same way as the Poisson problem. Due to time constraints, the problem was not solved using finite-element techniques.

Computer Codes

All codes were written in Fortran-77 for a DEC Vax 11/780 computer. The computer programs were quite similar for all three heat equations.

For each of the boundary value problems, a program was developed that calculated the exact temperatures at the nodes and then varied these temperatures by a given percentage. These temperatures were substituted into the variational integral approximation to determine what effect errors in the temperatures had on the value of the variational integral.

Another code was written for each problem that was used in the analysis of the variational integral as a stopping criterion and as an accuracy criterion. The matrix equations from the finite-difference and finite-element methods were solved iteratively using the Gauss-Seidel method. The exact solution at the nodal points was calculated using the analytical solution. After each iteration, the difference

31

between the exact and iterative temperatures was used to calculate the percent difference between the exact and approximate temperatures

$$||\underline{\tilde{e}}^{(n)}|| = \left|\left|\frac{\underline{t} - \underline{t}^{(n)}}{\underline{t}}\right|\right|$$ (3.38)

The difference between successive iterations was used to calculate the percent change in temperature

$$||\underline{\tilde{d}}^{(n)}|| = \left|\left|\frac{\underline{t}^{(n+1)} - \underline{t}^{(n)}}{\underline{t}^{(n)}}\right|\right|$$ (3.39)

The reason for these new relationships is that when one actually uses a stopping criteria, the percent error in a solution is usually more useful than the absolute error. This also allows a comparison between the standard stopping criterion and the variational integral criterion by eliminating the magnitudes of the temperatures and variational integrals. Note that Eqs (3.38) and (3.39) are very similar to the error vector and displacement vector previously defined. In the following sections the terms "error norm" and "displacement norm" will refer to Eqs (3.38) and (3.39). The relationship

$$||\underline{\tilde{e}}^{(n)}|| = \frac{||\underline{\tilde{d}}^{(n)}||}{\left(1 - \frac{||\underline{\tilde{d}}^{(n)}||}{||\underline{\tilde{d}}^{(n-1)}||}\right)}$$ (3.40)

is assumed to hold although it cannot be rigorously developed as was Eq (2.63).

The variational integral was computed after each iteration by substituting the temperatures from the iterative process into the variational integral approximation. To compare the variational integral as a stopping criterion to the stopping criterion defined in Eq (3.40), the percent change in the variational integral is defined as

$$|\Delta I^{(n)}| \equiv \frac{\left|I^{(n)} - I^{(n-1)}\right|}{|I^{(n-1)}|}$$ (3.41)

IV Numerical Results

Variational Integral Minimum

Recall that to calculate the variational integral based on the nodal temperatures, the integral had to be approximated using differences for the derivatives and sums for the integrals. These approximations lead to the concern that the variational integral approximation might not be extremized at the exact nodal temperatures.

This was found to be the case for all three problems considered. The exact temperature for each node was varied by a given percentage of the exact temperature and the variational integral was calculated. Figures 4.1, 4.2, and 4.3 illustrate that the integral for the one-dimensional heat equation had a minimum for some given values of temperature and that this minimum occurred closer to the exact temperatures as the number of nodes increased. Figures 4.4 and 4.5 illustrate that, for the two-dimensional problems, the variational integrals are minimized and these minima occur at temperatures other than the exact values. Table I summarizes the results for all three problems. For the purpose of this thesis, one variational integral approximation will be said to be more accurate than another if the minimum in the variational integral occurs for temperatures closer to the exact values. Thus, Table I shows that the variational integral approximation becomes more accurate as the number of nodes increase. In practice, one will not know the exact temperatures and will have no way of determining the accuracy of an integral approximation as was done in Table I.

Figure 4.1. Variational Integral Minimum, 1-D Problem, Four Nodes
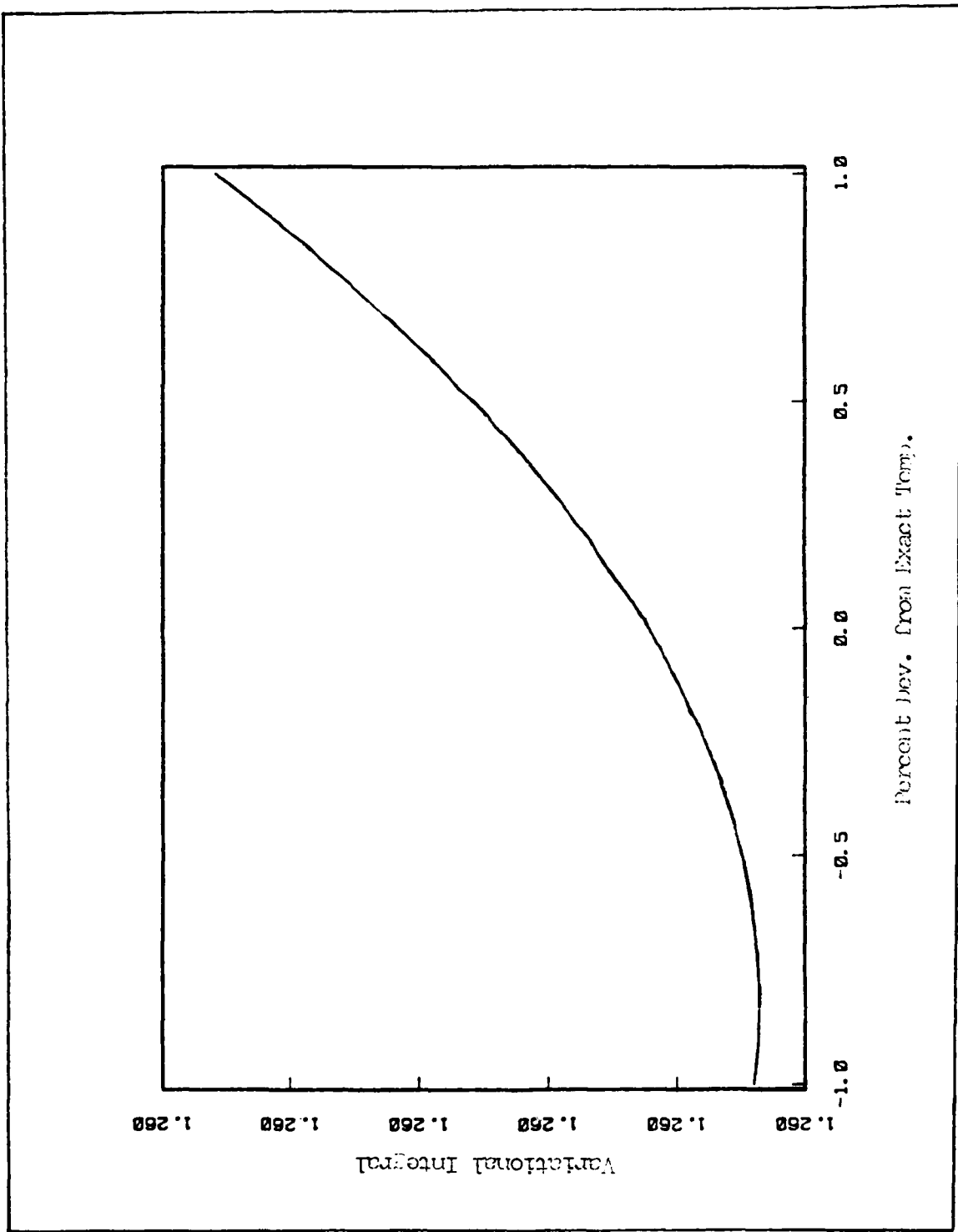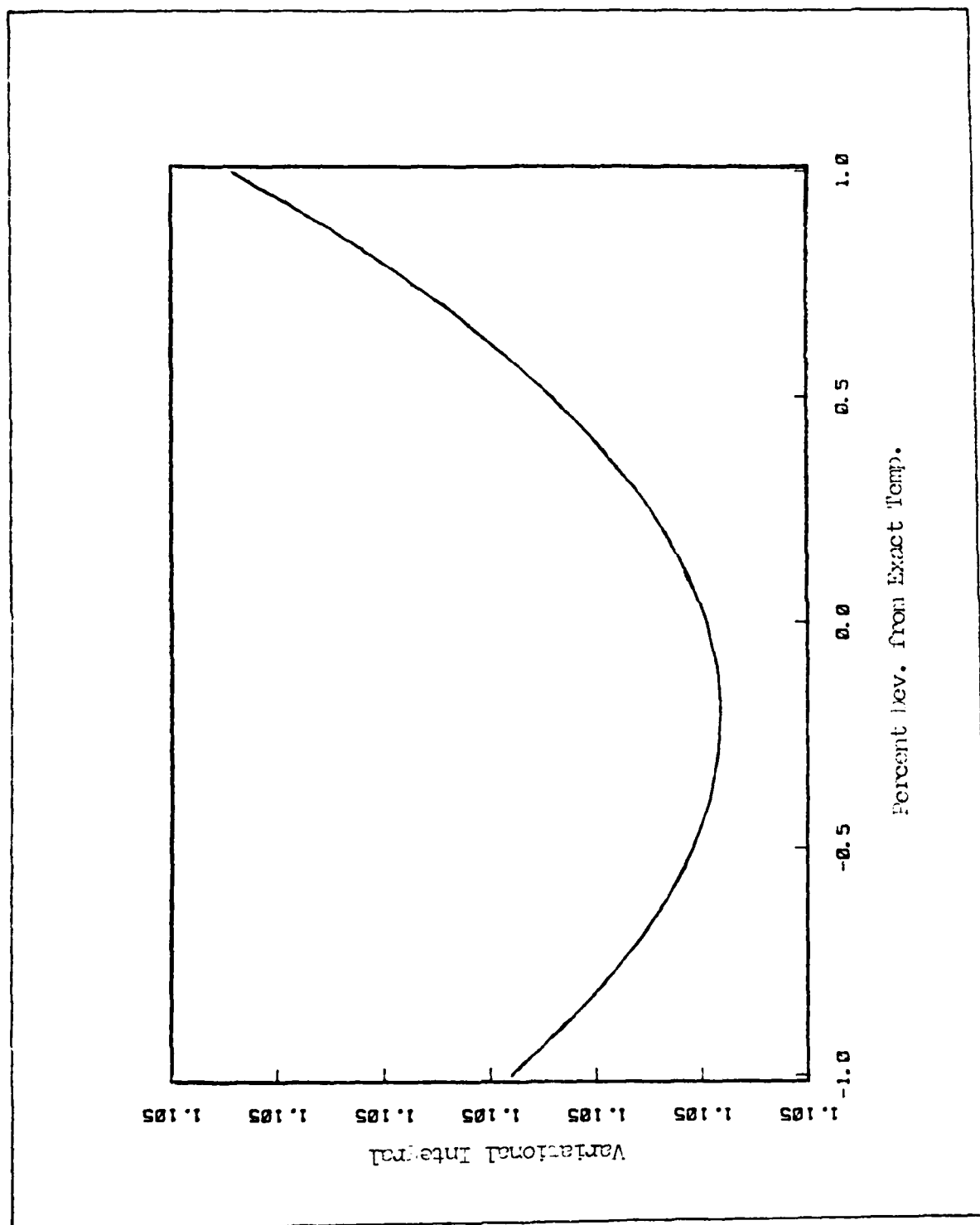
Percent Dev. from Exact Temp.

Figure 4.2. Variational Integral Minimum, 1-D Problem, Eight Nodes
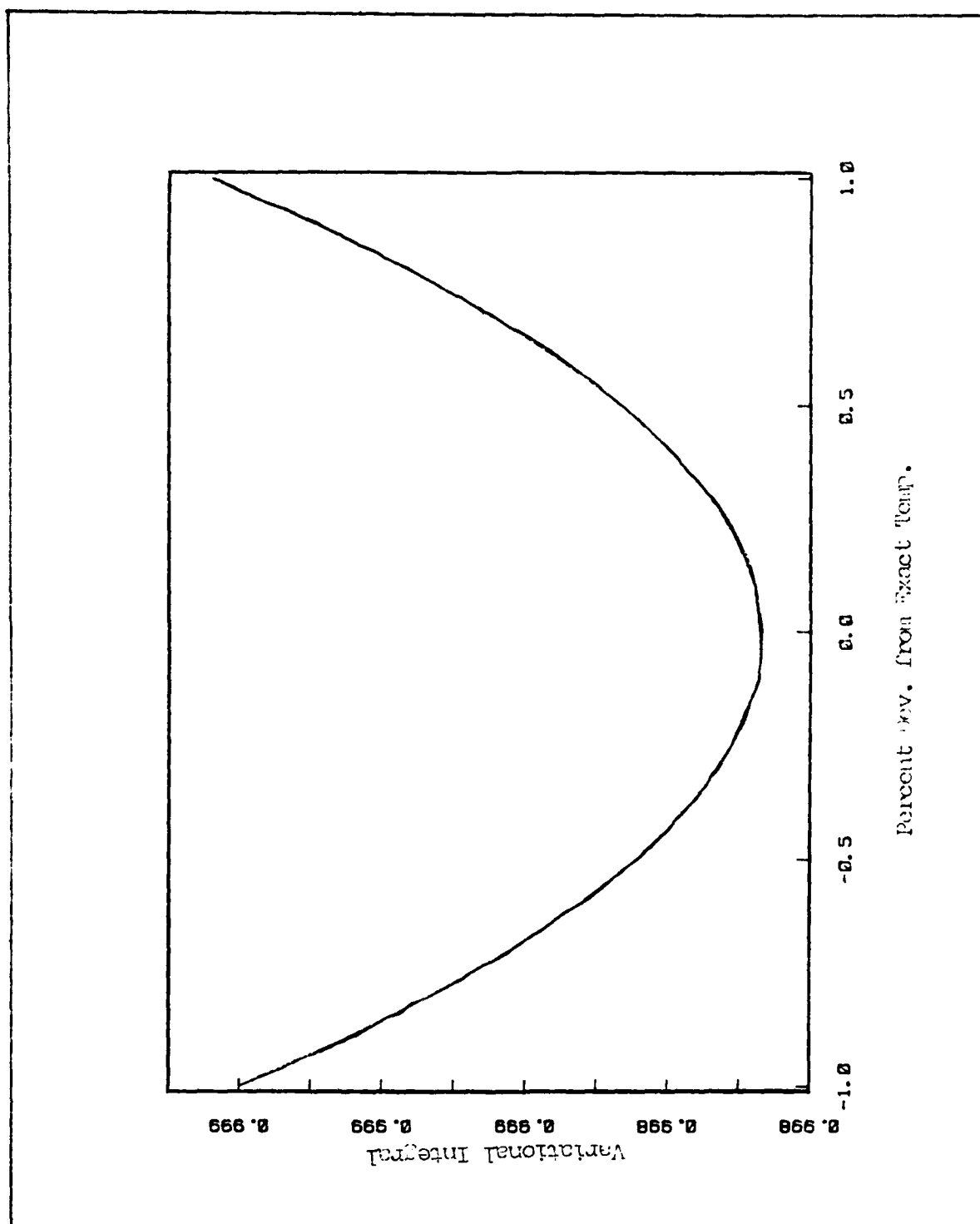
36

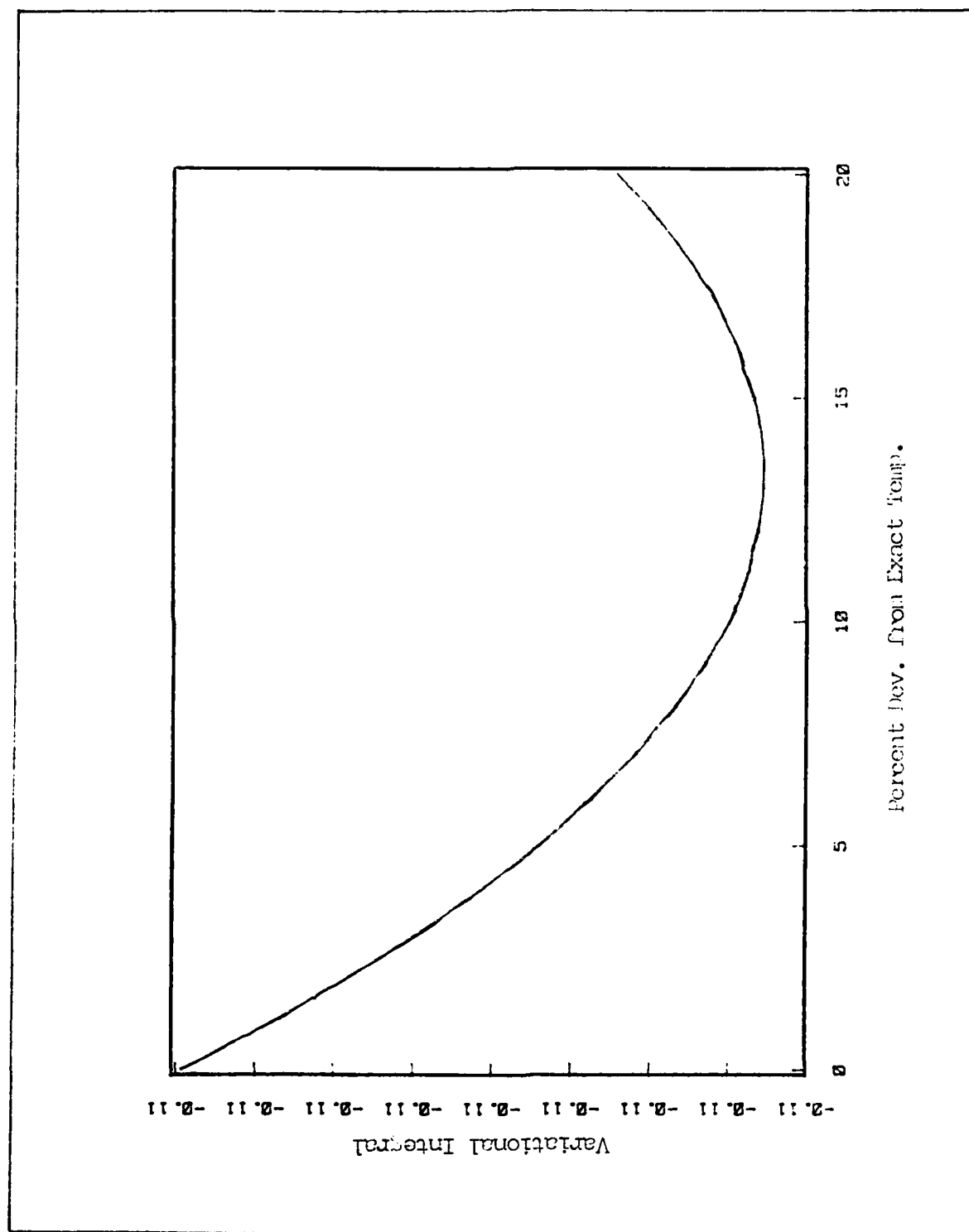Figure 4.3. Variational Integral Minimum, 1-D Problem, Thirty-two Nodes

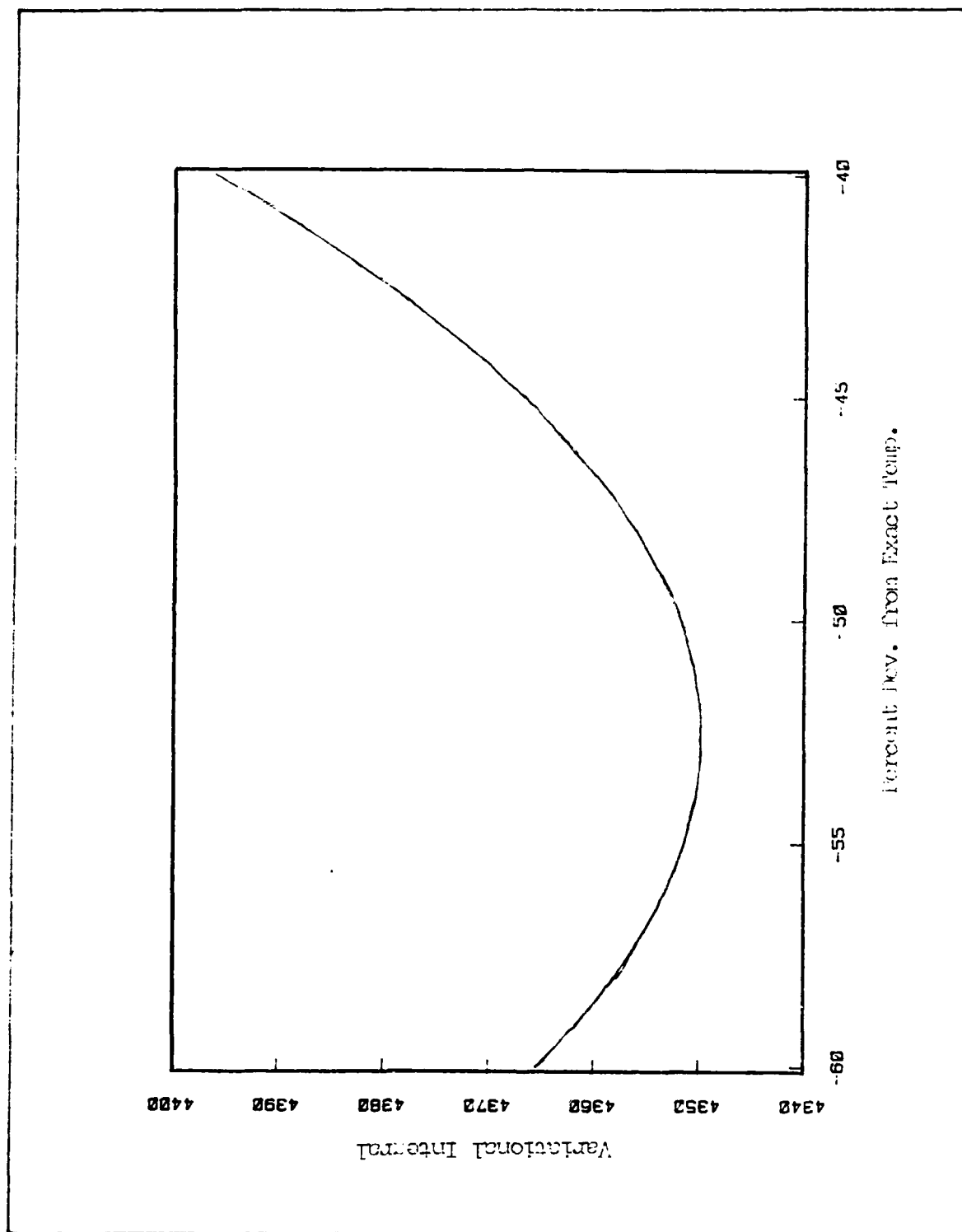Figure 4.4. Variational Integral Minimum, Poisson Problem, Four Nodes

Figure 4.5. Variational Integral Minimum, Laplace Problem, Four Poles

The exact value of the variational integral is calculated using the analytical solution as shown in Appendix A. It can be seen, in Table I, that the value of the variational integral approximation does not approach the value of the exact variational integral in the Poisson problem even though the accuracy of the integral approximation improves for an increasing number of nodes (using the definition of accuracy given above). Thus, relating the value of the integral approximation to the exact value is of little use. Furthermore, one will not be able to calculate the exact variational integral in "real-world" problems since the analytic solution will not be known.

Table I

Variational Integral Approximation

Problem	Exact Integral[1]	Nodes	Integral Approximation[2]	Percent Deviation[3]
One-Dimen.	0.964	4	1.260	−0.80
		8	1.105	−0.21
		16	1.033	−0.05
		32	0.993	−0.01
Two-Dimen. Poisson	−0.1406	4x4	−0.1070	+13
		6x6	−0.0947	+9
		8x8	−0.0884	+7
Two-Dimen. Laplace	1.04×10^4	4x4	4.35×10^3	−53
		6x6	5.92×10^3	−45
		8x8	7.21×10^3	−39

1. Value of variational integral using analytic solution.

2. Value of variational integral approximation at its minimum.

3. Percent deviation from exact nodal temperatures where the minimum in the variational integral occurs.

Stopping Criteria

Recall from Section II, that for a problem with slow convergence, the difference between temperatures for successive iterations could be small even though the difference between the iterated value and the exact value was still quite large. It was thought that the variational integral might give an alternative evaluation for termination since the integral is extremized as the iterated nodal temperatures approach the exact nodal temperatures. Figure 4.6 shows that as the sum of the iterated temperatures approaches the sum of the exact temperatures, the variational integral converges to a constant value.

Figure 4.7 compares a common stopping criterion, $||\tilde{d}^{(n)}||$ with the variational integral stopping criterion, $|\Delta I^{(n)}|$. It can be seen from Figure 4.7 that, past a certain iteration, $||\tilde{d}^{(n)}||$ and $|\Delta I^{(n)}|$ decrease at the same rate.

One can easily explain the linear decrease in $||\tilde{d}^{(n)}||$ (keeping in mind that the y-axis in Figure 4.7 is a logarithmic scale). The slope of $||\tilde{d}^{(n)}||$ is given by

$$m = \frac{\log||\tilde{d}^{(n)}|| - \log||\tilde{d}^{(n-1)}||}{n - (n-1)} \qquad (4.1)$$

Recall that for large n

$$|\lambda_1| \approx \frac{||\tilde{d}^{(n)}||}{||\tilde{d}^{(n-1)}||} \qquad (4.2)$$

or

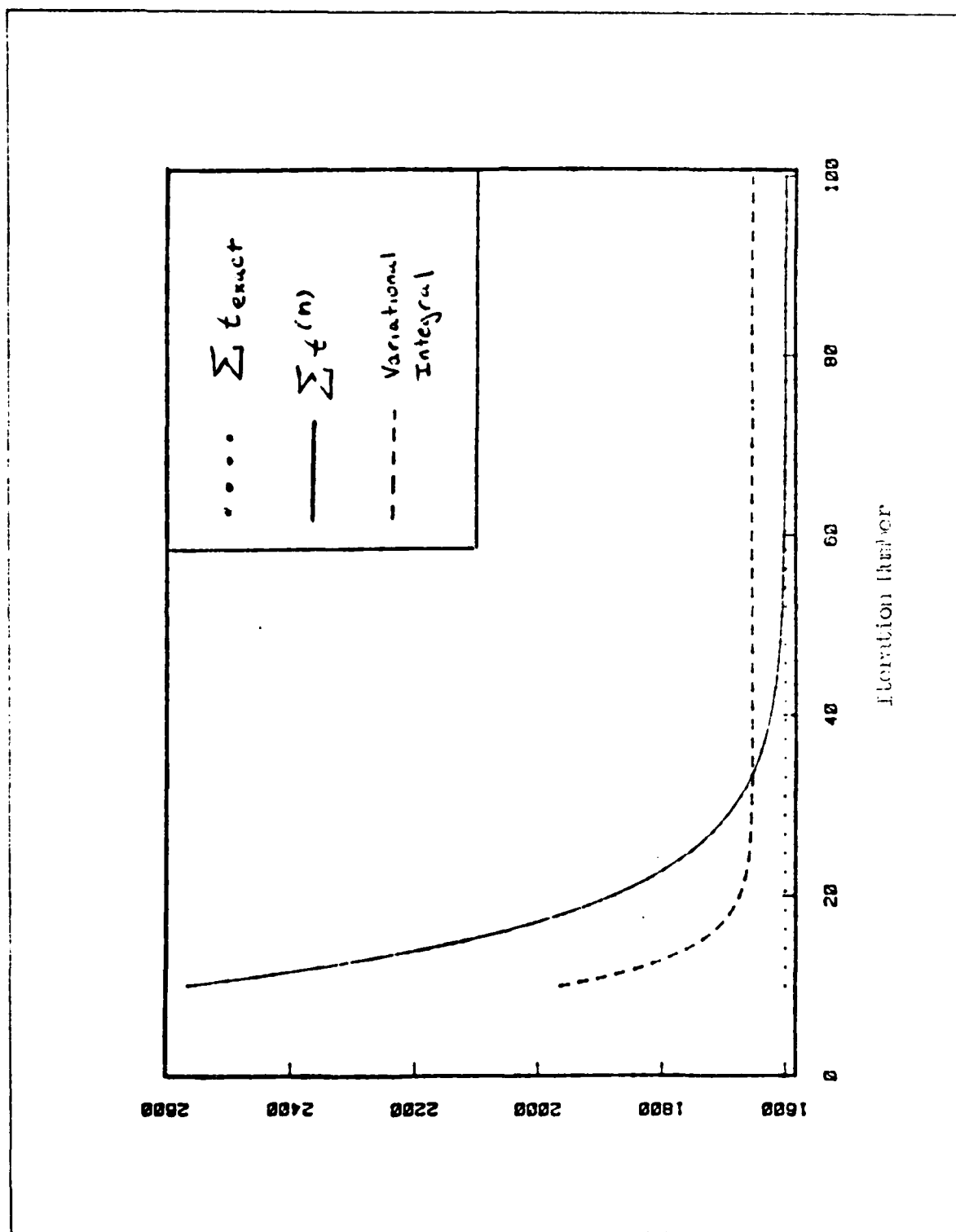$$\log|\lambda_1| \approx \log||\tilde{d}^{(n)}|| - \log||d^{(n-1)}|| \qquad (4.3)$$

41

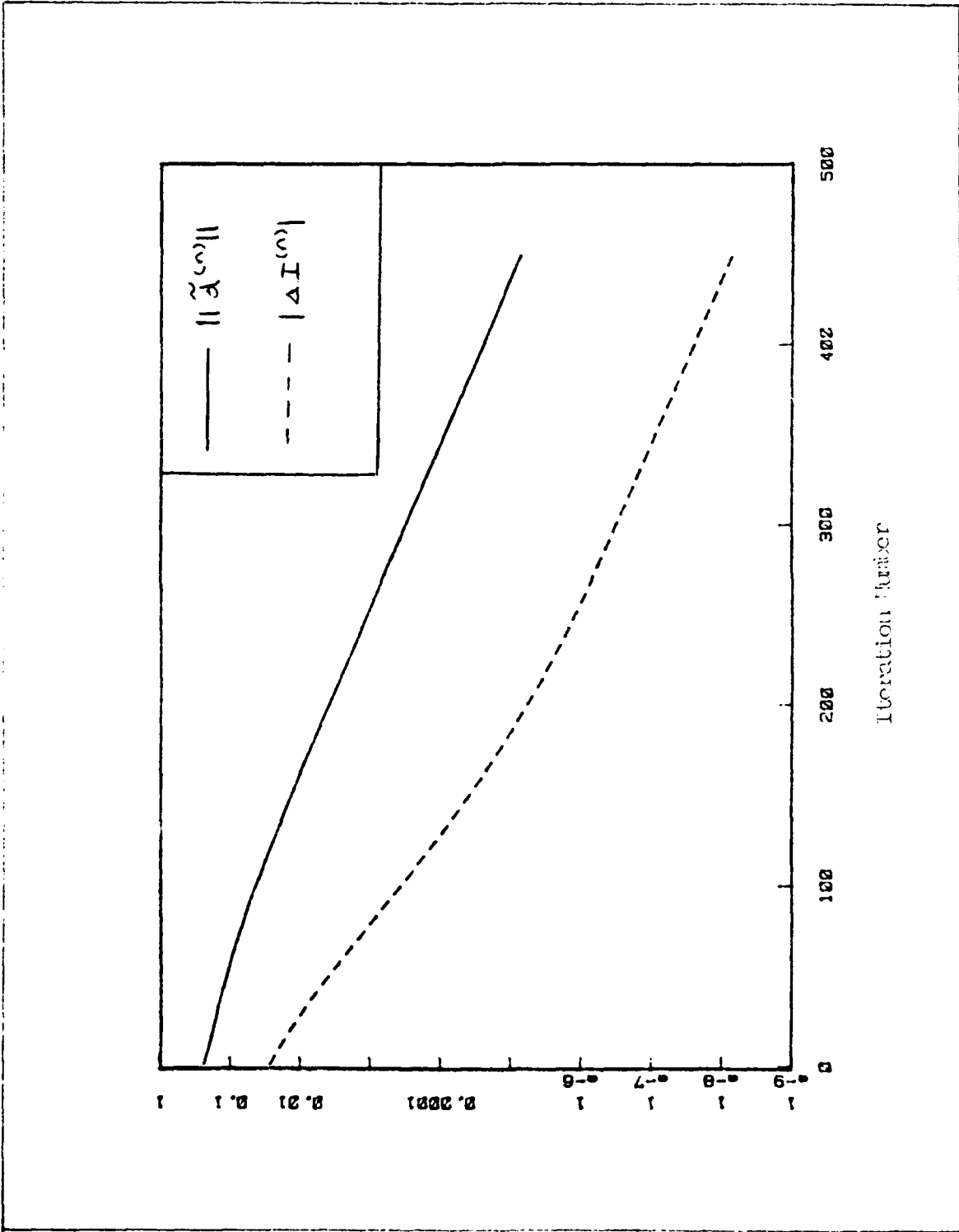Figure 4.8. Temperature and Variational Integral Convergence

Figure 4.7. Stopping Criterion Comparison

By comparison of Eq. (4.1) and (4.3), it can be seen that

$$m = \log |\lambda_1| \qquad (4.4)$$

More importantly, Figure 4.7 shows that for sufficiently large n

$$|\lambda_1| \approx \frac{|\Delta I^{(n)}|}{|\Delta I^{(n-1)}|} \qquad (4.5)$$

Eq. (4.5) was found to hold for all three heat problems and all nodal densities considered. Thus, the variational integral does not provide a better stopping criterion because $|\Delta I^{(n)}|$ decreases at the same rate as $||\tilde{d}^{(n)}||$.

Furthermore, the standard method can provide an estimate of the error while the variational integral cannot. Figure 4.8 shows the error norm and the error norm estimate. The error estimate approximates the error once the spectral radius approximation has converged to its final value. Figures 4.9 through 4.11 show that the spectral radius approximation rapidly reaches a given value and that this value increases as the number of nodes increase. Returning to Figure 4.8, it can be seen that the estimate of the error norm continues to decrease even though the error norm has stopped decreasing. This is due to the fact that the error norm estimate depends only on temperatures from successive iterations. The error norm decreases until, at some iteration, the discretization error prevents any further decrease.

In Section II, it was predicted that the variational integral could be useful when the iterative temperatures passed through the exact values. Figure 4.12 shows the sum of the nodal temperatures from
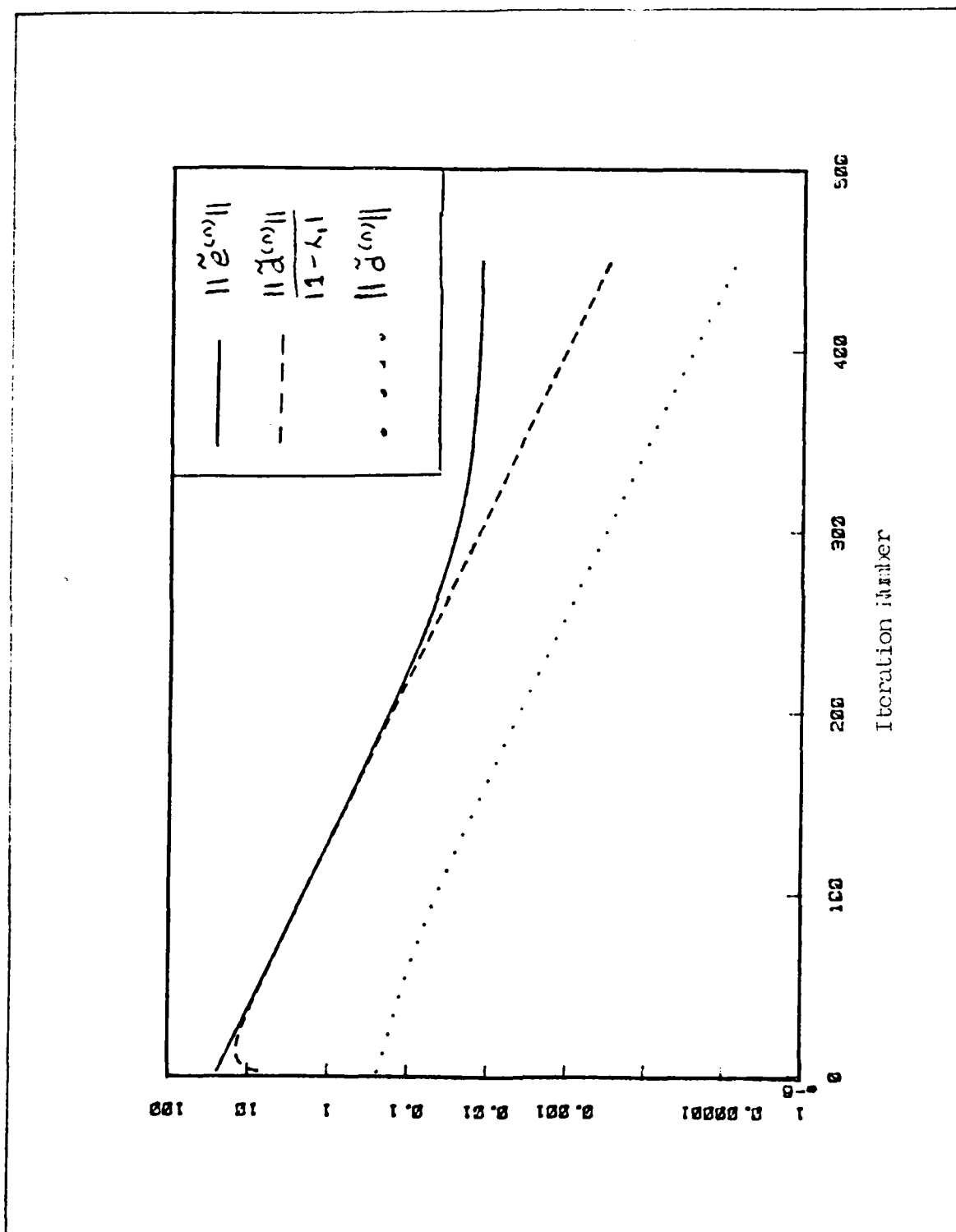
44

Figure 4.6. Error Norm Approximation
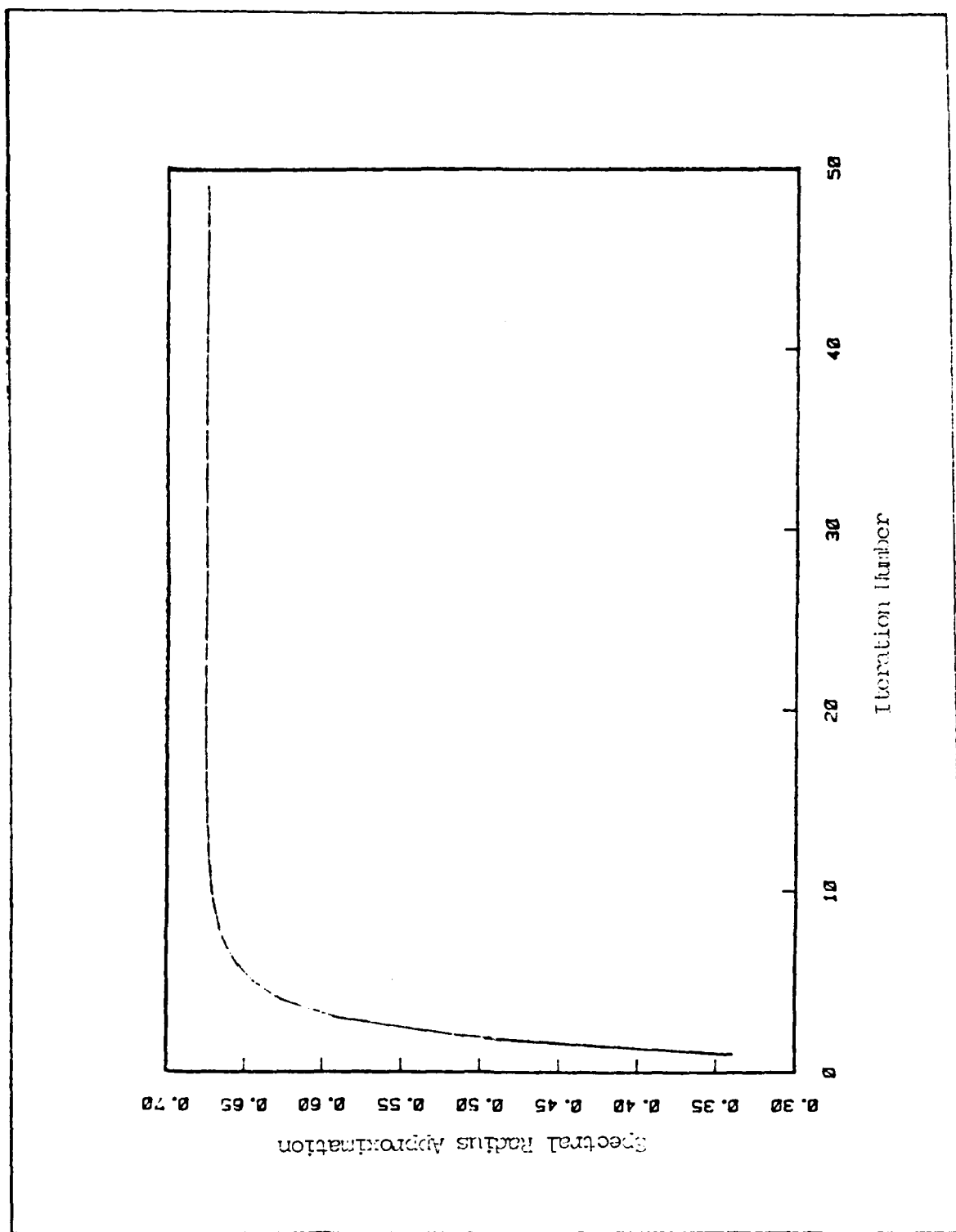
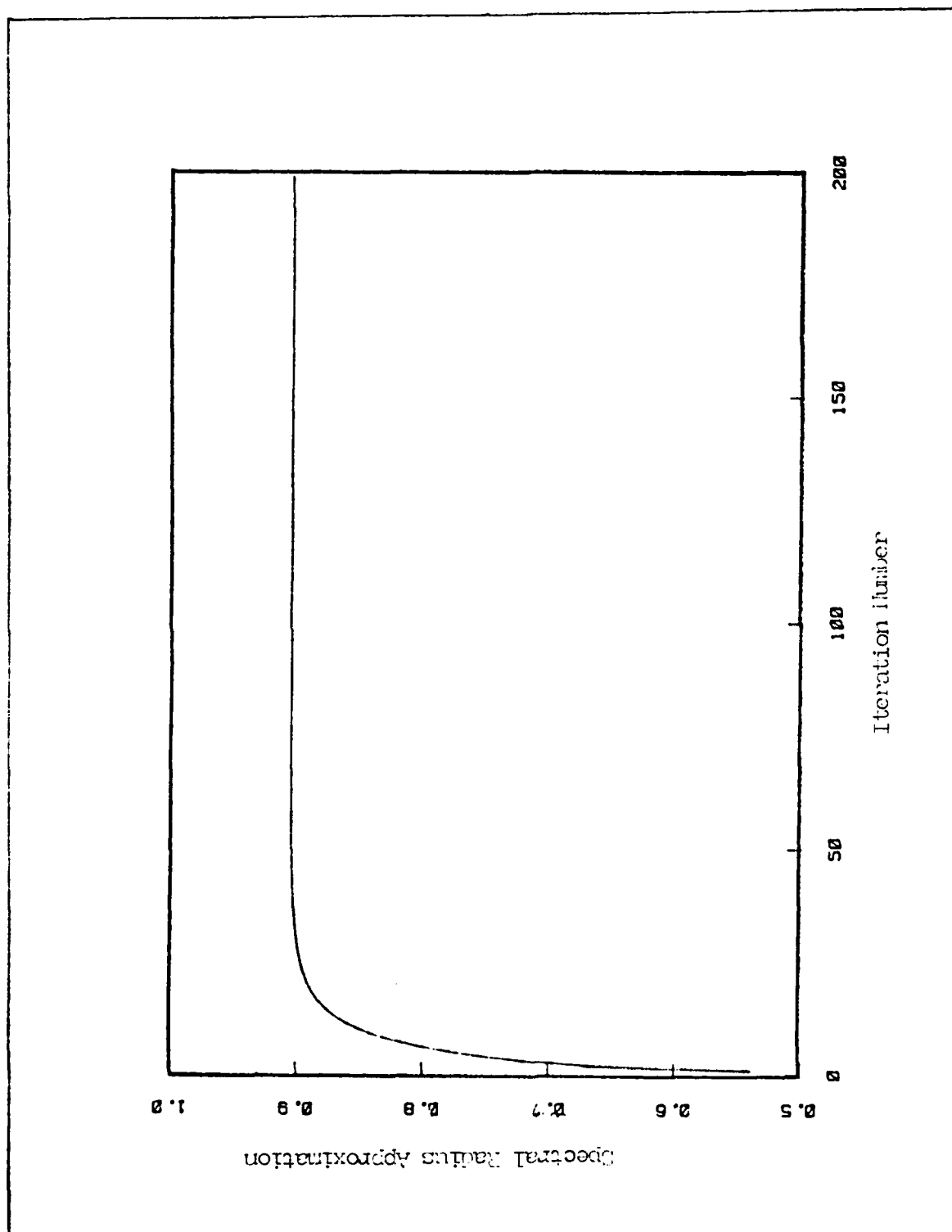Figure 4.9. Spectral Radius Approximation, 1-D Problem, Four Holes

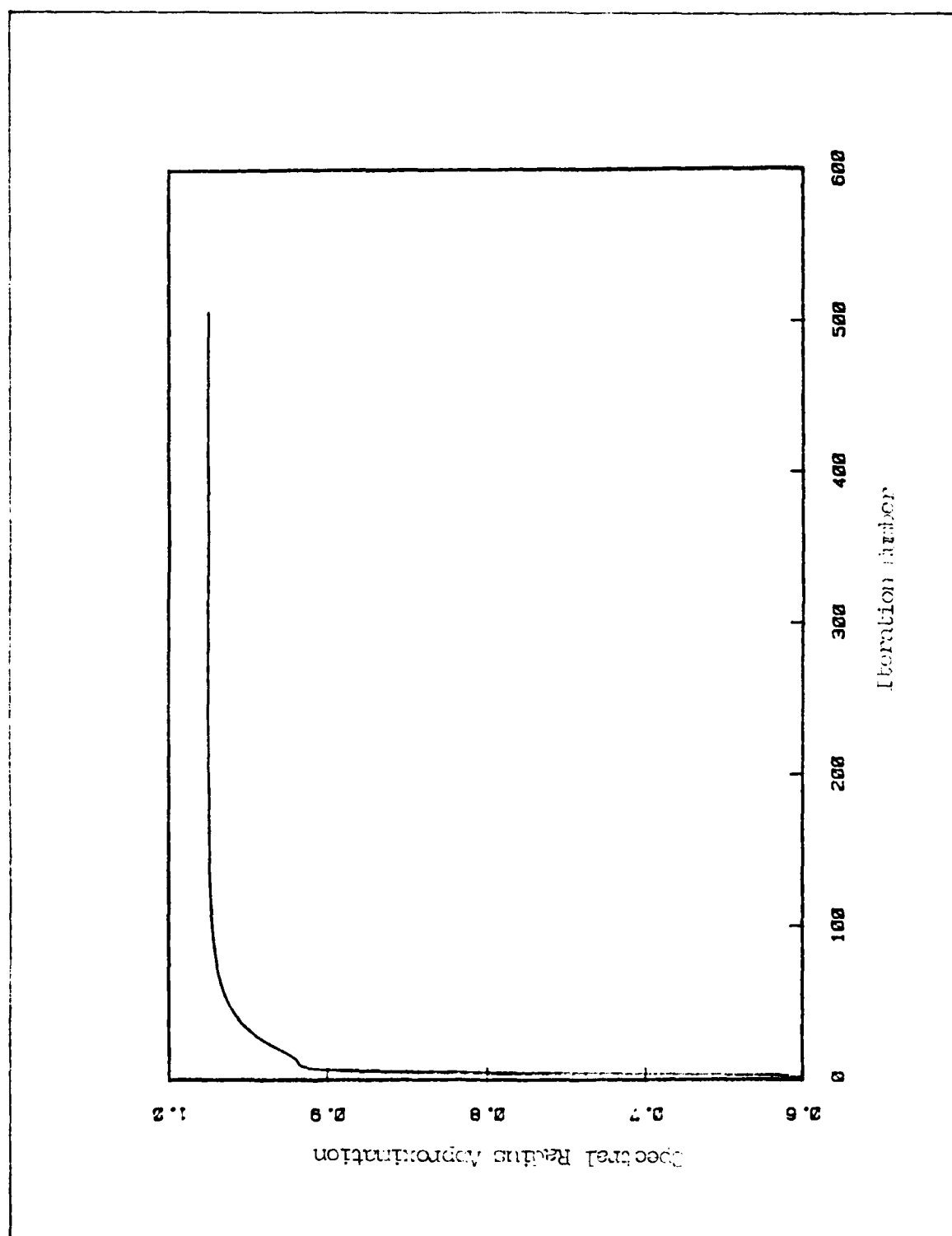Figure 4.10. Spectral Radius Approximation, 1-D Problem, Eight Nodes

Figure 4.11. Spectral Radius Approximation, 1-D Problem, Sixteen Nodes
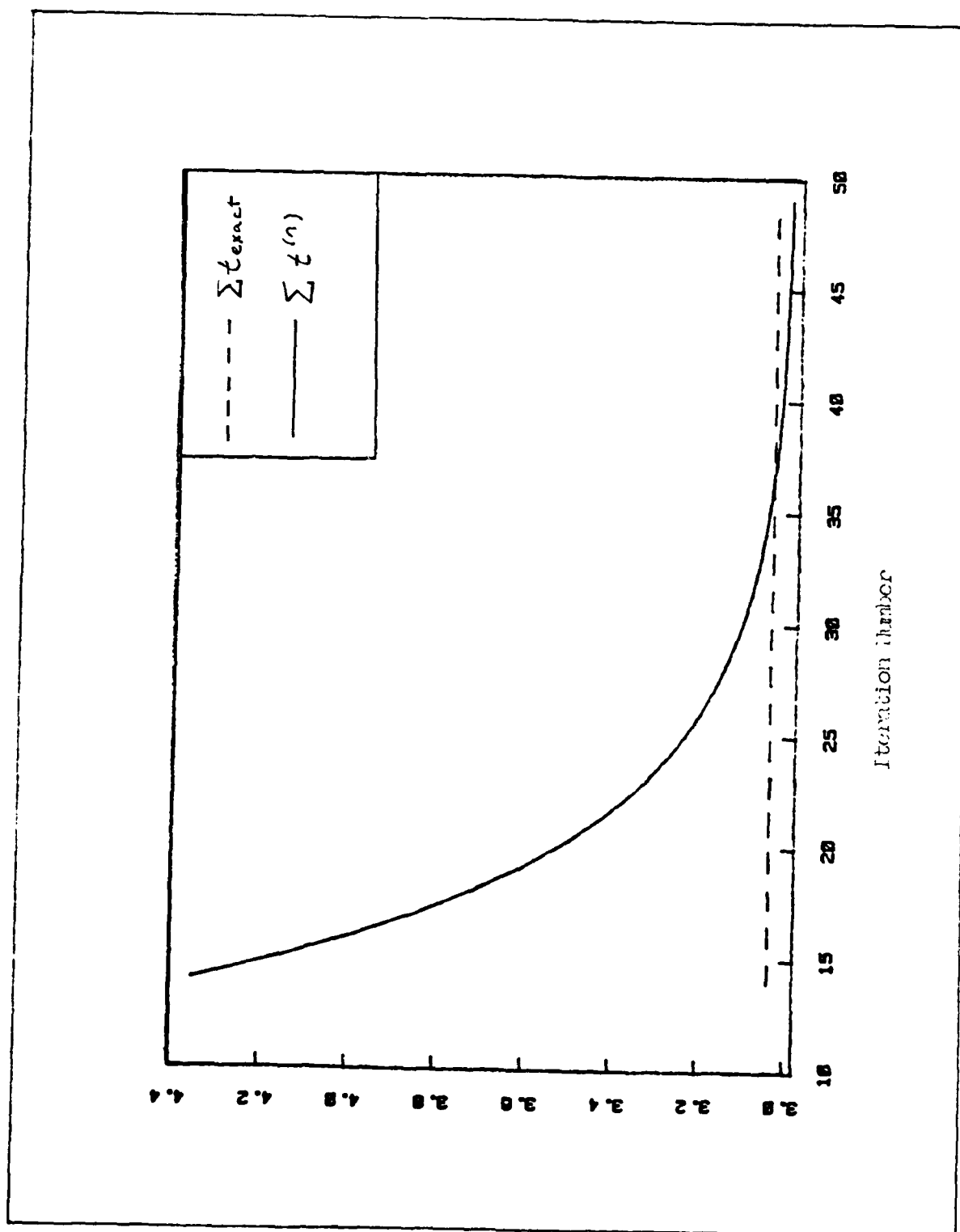
Figure 4.12. Total Temperature data

the iterative method passing through the sum of the exact nodal temper-
atures after thirty-seven iterations. Thus, one would expect the
variational integral to be extremized at the thirty-seventh iteration.
Figure 4.13 shows that the variational integral was actually minimized
after nineteen iterations. This is understandable in light of the
previous discussion in which it was found that the variational integral
approximation is minimized for temperatures other than the exact
values. The results from Figures 4.12 and 4.13 can be represented more
concisely using $|\Delta I^{(n)}|$ and $||\tilde{e}^{(n)}||$. Figure 4.14 shows that the
variational integral stopping criterion, $|\Delta I^{(n)}|$ has a "cusp" or local
minimum where the variational integral is a minimum (nineteenth itera-
tion) and the error norm, $||\tilde{e}^{(n)}||$ has a cusp where the iterative
solution passes through the exact solution (thirty-seventh iteration).
In Figure 4.14, the initial trial solution was greater than the exact
solution. The finite-difference equations converged to temperatures
below the exact values. Figure 4.15 shows that when the initial trial
solution was below the exact solution, the iterative temperatures did
not pass through the exact values and no cusps were observed for either
$|\Delta I^{(n)}|$ or $||\tilde{e}^{(n)}||$. Table II compares the iteration number where the
integral approximation is minimized to the iteration number where the
error norm is minimized. Appendix D shows that the results of the
stopping criteria study hold for all three problems and all nodal
densities considered.
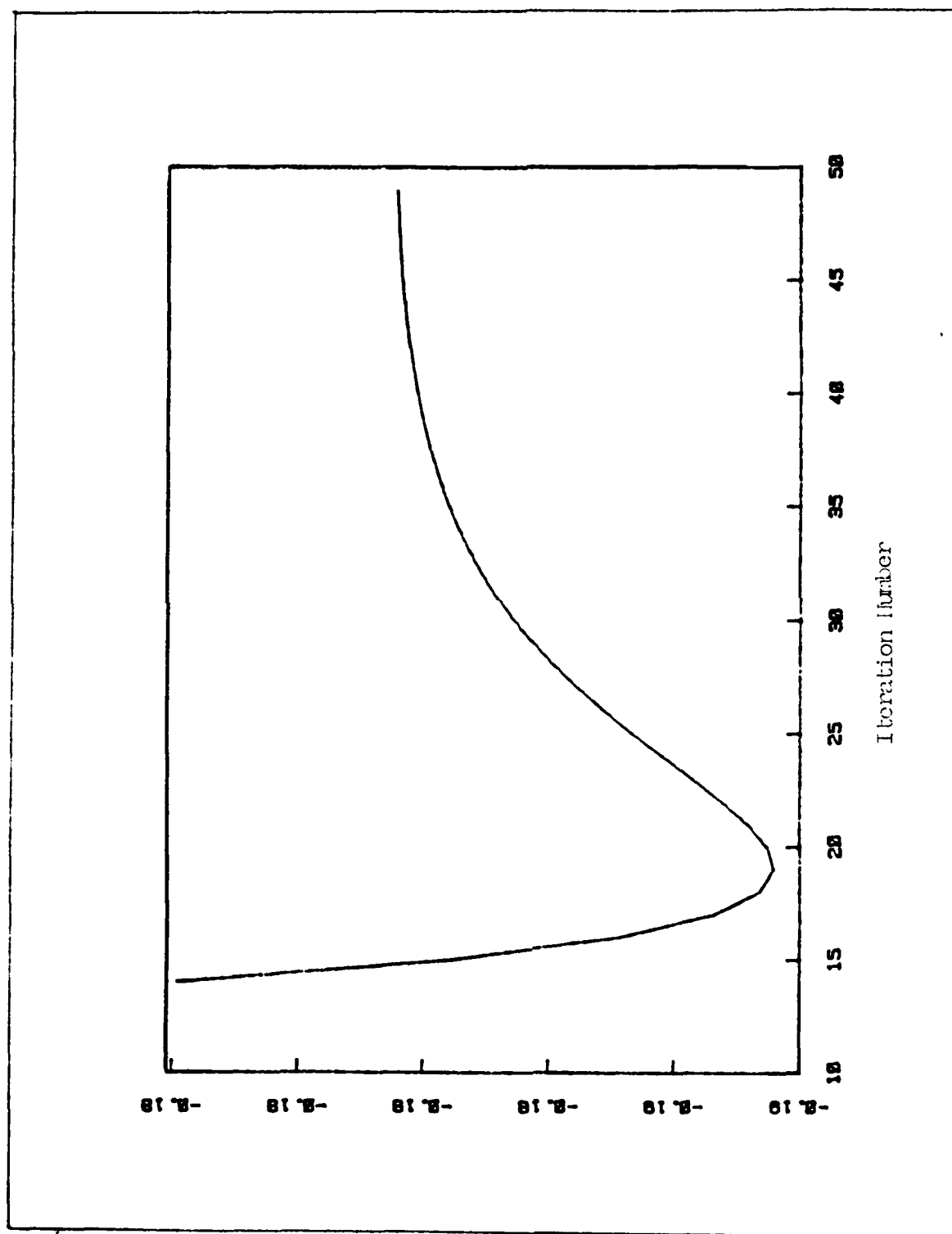
Figure 4.13. Variational Integral
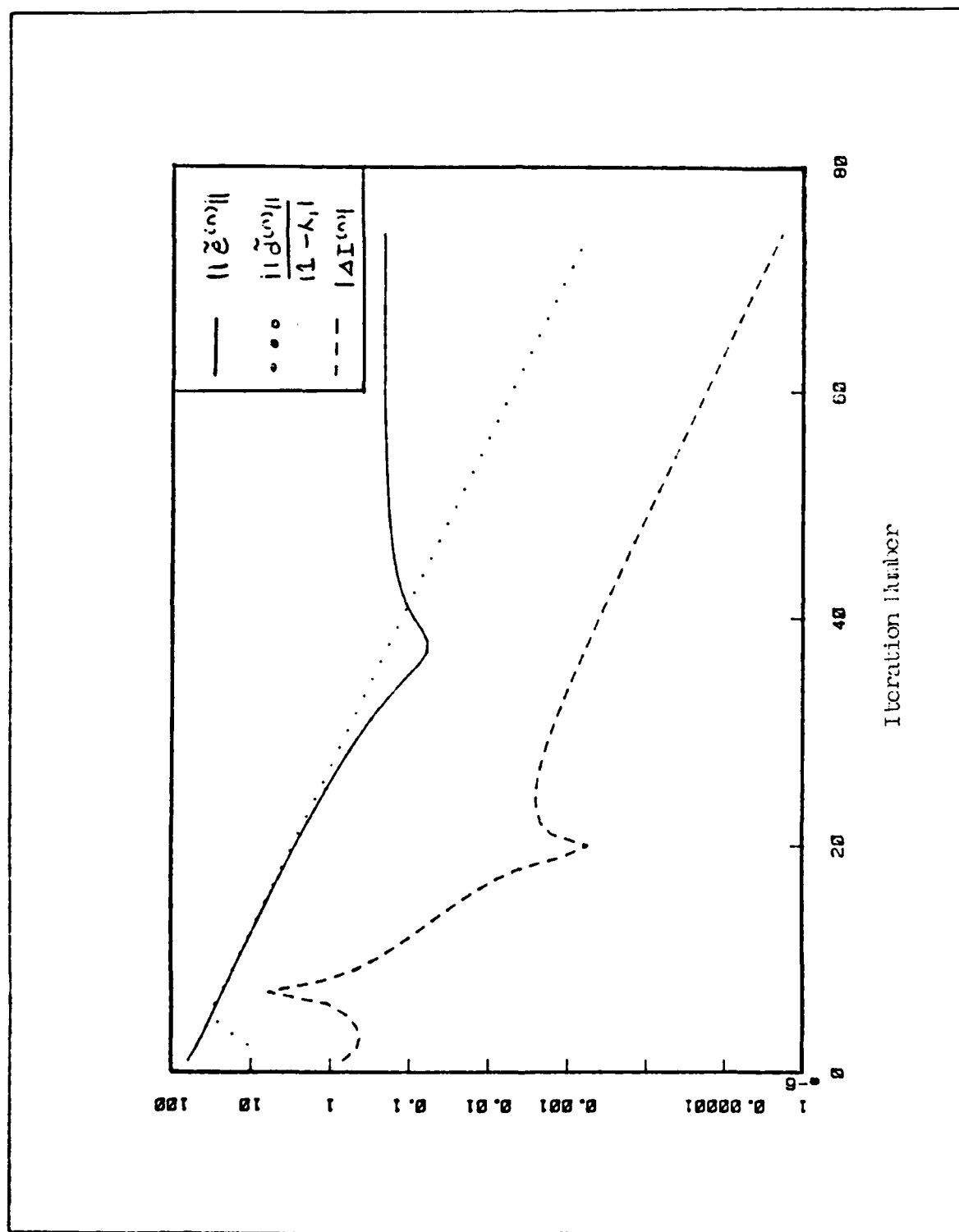
Figure 4.14. Stopping Criteria, Initial Trial Solution Above the Exact Solution
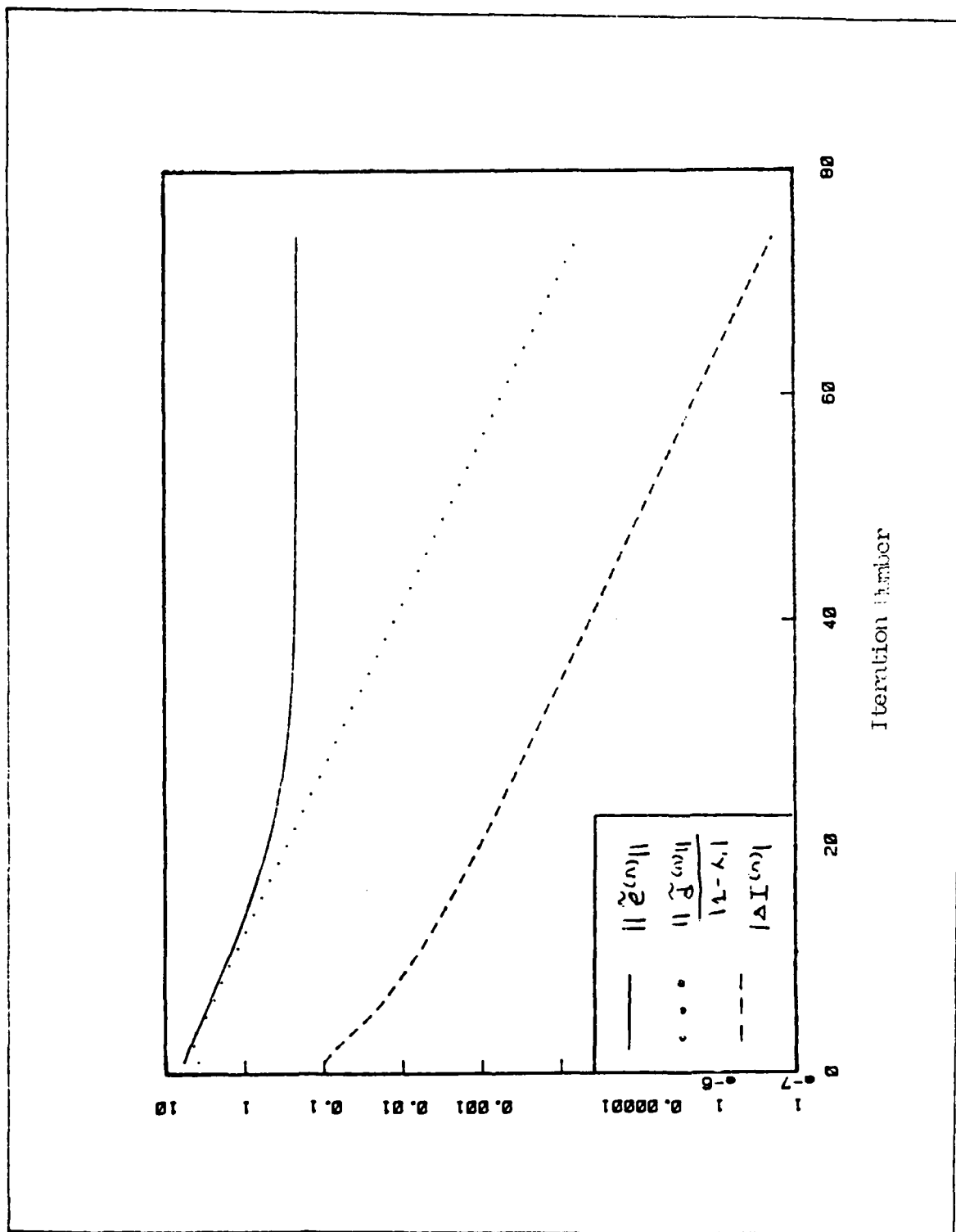
Figure 4.15. Stopping Criteria, Initial Trial Solution Below Exact Solution

Table II

Integral and Error Norm Minima

Problem[1]	Nodes	Integral[2]	$\|\tilde{e}\|^2$
One-Dimen.	4	6	9
	6	15	26
	8	30	53
	10	51	90
	16	150	280
Two-Dimen. Poisson	4x4	20	38
	6x6	43	96
	8x8	74	185
Two-Dimen. Laplace	4x4	7	32
	6x6	18	66
	8x8	35	111

1. Using the finite-difference method

2. Iteration number where the minimum occurs

Accuracy Criterion

Since the variational integral is minimized for the exact solution, it was thought that the variational integral approximation could be used to determine which numerical method would give the best approximation to the true solution. Two methods were considered—the finite difference method and the finite-element method. Due to time constraints, only the one-dimensional problem and the Poisson problem were considered. Table III shows the results for the one-dimensional problem.

Table III

Integral and Error Norm for 1-D Problem

Nodes	Method[1]	Integral	$||\tilde{e}||$
3	FD	1.3720	0.0759
	FE	1.3695	0.0812
4	FD	1.2606	0.0549
	FE	1.2594	0.0569
5	FD	1.1967	0.0429
	FE	1.1960	0.0436
6	FD	1.1553	0.0351
	FE	1.1550	0.0352
7	FD	1.1264	0.0298
	FE	1.1262	0.0295
8	FD	1.1051	0.0259
	FE	1.1049	0.0253

1. FD: Finite-Difference Method
   FE: Finite-Element Method

Note that for nodal densities of less than six, the temperatures from the finite-element method minimize the variational integral even though the finite-difference method minimizes the error norm. Nodal densities greater than eight were considered but they gave the same results as densities six through eight and, consequently, were not included in the table. The discrepancy between the integral and error norm for nodal densities less than six can be explained. Table IV shows that the temperatures for the finite-element method are less than the exact temperatures while the temperatures from the finite-difference method are greater than the exact temperatures.

55

Table IV

Nodal Temperatures for 1-D Problem (Four Nodes)

Node	Exact	Finite-Diff.	Finite-Elem.
1	0.6252	0.6289	0.6215
2	0.4102	0.4150	0.4051
3	0.2998	0.3049	0.2943
4	0.2658	0.2710	0.2604

Figure 4.16 is based on Figures 4.1 through 4.5 and shows how the variational integral could fail to predict which method is most accurate. For nodal densities greater than or equal to six, the finite-element method gives more accurate temperatures and, therefore, the variational integral approximation correctly predicts the more accurate method.

The results for the Poisson problem are summarized in Table V. It was found that one finite-element case gave the most accurate temperatures while the other finite-element case gave the least accurate answers. For all nodal densities considered, the variational integral correctly predicted which method gave the most accurate solution. The difficulty that occurred in the one-dimensional problem did not occur for the Poisson problem because the majority of temperatures for all three methods gave temperatures below the exact values.
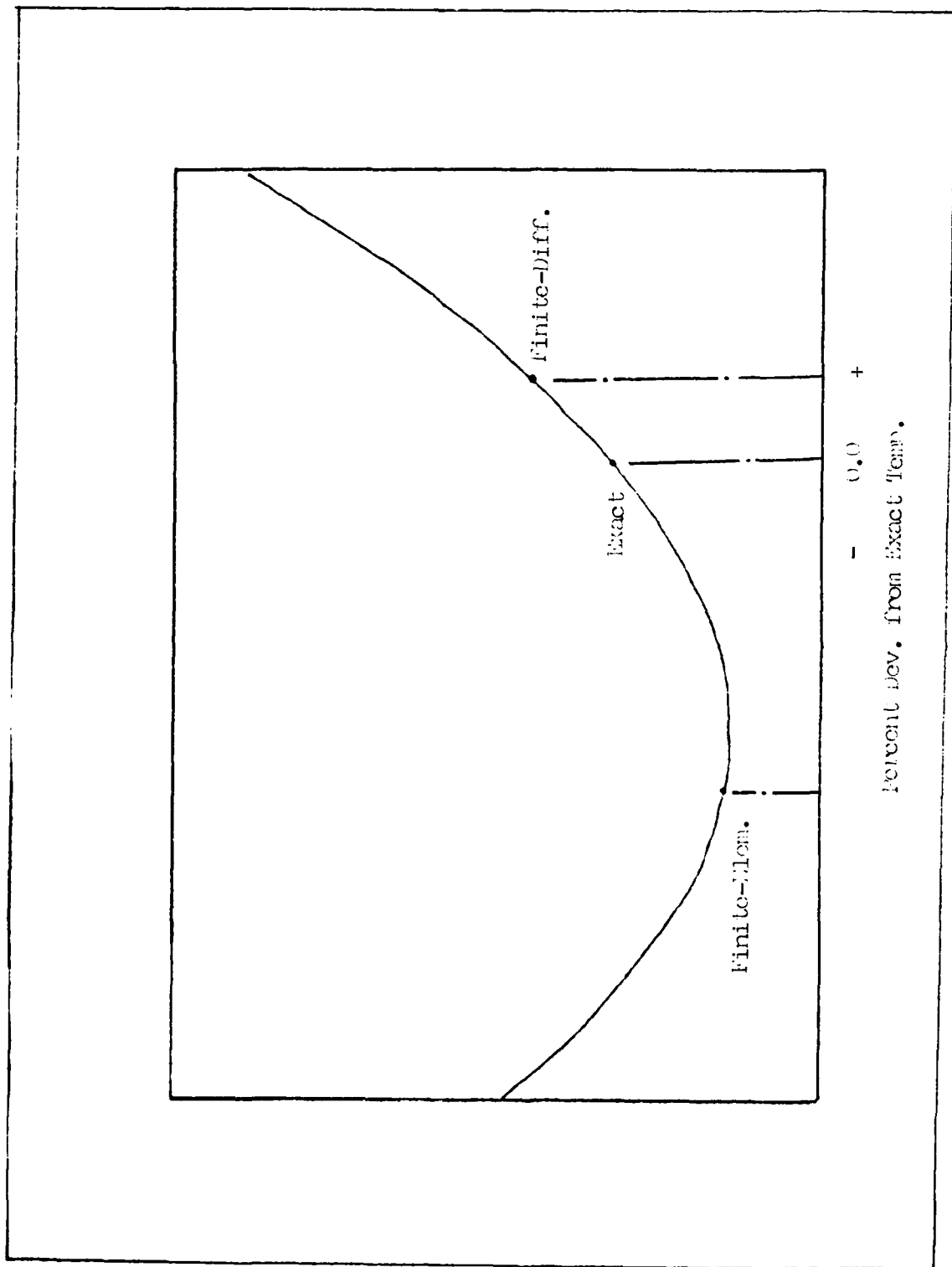
Figure 4.16. Schematic Illustration of Accuracy Criterion Failure

57

Table V

Integral and Error Norm for Poisson Problem

Nodes	Method[1]	Integral	$\|\|\tilde{e}\|\|$
2×2	FE-2	-0.1632	0.4147
	FD	-0.1680	0.1891
	FE-1	-0.1710	0.0844
4×4	FE-2	-0.1806	0.3966
	FD	-0.1815	0.2143
	FE-1	-0.1821	0.1024
8×8	FE-2	-0.1940	0.3961
	FD	-0.1942	0.2366
	FE-1	-0.1944	0.1223

1. FD: Finite-Difference Method
   FE-1: Finite-Element Method (Case One)
   FE-2: Finite-Element Method (Case Two)

Conclusions and Recommendations

Conclusions

The variational integral approximation for each of the three boundary value problems was minimized for temperatures other than the exact values. As the number of nodes was increased, the approximations were minimized at temperatures closer to the exact values.

As a stopping criterion, the variational integral was found to be less effective than the error estimate criterion. For a sufficiently large number of iterations, the variational integral stopping criterion decreased at the same rate as the error norm estimate criterion. But the error norm estimate provides an estimate of the actual error norm for a limited number of iterations while no such error estimate is available from the variational integral stopping criterion. Furthermore, the error norm estimate is easier to calculate.

For the special case where the temperatures from the iterative methods passed through the exact temperatures, it was found that the error norm was minimized at a different iteration than was the variational integral approximation. This was expected since the variational integral approximations were not minimized for the exact temperatures. Thus, the utility of the variational integral as a stopping criterion for this special case depends on the accuracy of the variational integral approximation.

The variational integral approximation did not correctly predict which method gave the most accurate solution in some of the cases studied. It is believed that this failure is due to the variational

integral approximation being minimized at temperatures other than the exact values.

Recommendations

Many of the difficulties of using the variational integral as a stopping criterion or an accuracy criterion are believed to be due to the fact that the variational integral approximation is minimized at temperatures other than the exact values. Thus, attention should be focused on reducing or, at least, predicting this error. Better methods of approximating the derivatives and integrals under the variational integral may exist.

For the accuracy criterion, consideration should be given to using linear interpolation functions to transform the nodal temperatures into a piecewise continuous function. These functions can be substituted into the variational integral and the integral can be solved analytically. Other techniques that solve the boundary value problem over the entire solution region such as the Rayleigh-Ritz and Galerkin methods can also be substituted into the variational integral to determine which method gives the most accurate solution.

Bibliography

1. Amazigo, John C. and Lester A. Rubenfeld. _Advanced Calculus and Its Applications to the Engineering and Physical Sciences_. New York: John Wiley and Sons, Inc., 1980.

2. Churchill, Ruel V. and James Ward Brown. _Fourier Series and Boundary Value Problems (Third Edition)_. New York: McGraw-Hill Book Company, 1978.

3. Clark, Melville Jr. and Kent F. Hansen. _Numerical Methods of Reactor Analysis_. New York: Academic Press, Inc., 1964.

4. Conte, S. D. and Carl De Boor. _Elementary Numerical Analysis (Third Edition)_. New York: McGraw-Hill Book Company, 1980.

5. Forsythe, George E. and Wolfgang R. Wasow. _Finite-Difference Methods for Partial Differential Equations_. New York: John Wiley and Sons, Inc., 1960.

6. Gallof, Sanford. _A Numerical Investigation and Utilization of Green's Functions_. MS thesis. Wright-Patterson AFB, Ohio: Air Force Institute of Technology, June 1968.

7. Gurtin, M. E. "Variational Principles for Linear Initial Value Problems," _Quarterly Journal of Applied Mathematics, 22_: 252–256 (1964).

8. Huebner, H. Kenneth. _The Finite Element Method for Engineers_. New York: John Wiley and Sons, Inc., 1975.

9. Kaplan, B. _Investigation of the "Iteration Limit" in the Generalized Transient Heat Transfer Program (THT)_. DC 59-11-198. Aircraft Nuclear Propulsion Department, General Electric, November 1959.

10. Martin, Harold C. and Graham F. Carey. _Introduction to Finite Element Analysis_. New York: Mcgraw-Hill Book Company, 1973.

11. Myers, Glen E. _Analytical Methods in Conduction Heat Transfer_. New York: McGraw-Hill Book Company, 1971.

12. Mikhlin, S. G. _Variational Methods in Mathematical Physics_. New York: The Macmillan Company, 1964.

13. Schechter, R. S. _The Variational Method in Engineering_. New York: McGraw-Hill Book Company, 1967.

14. Smith, G. D. _Numerical Solution of Partial Differential Equations (Second Edition)_. Oxford: Clarendon Press, 1978.

15. Tyn Myint-U. _Partial Differential Equations of Mathematical Physics_. New York: Elsevier North Holland, Inc., 1980.

16. Warren, Robert A. _Convergence and Error Criteria of Iterative Numerical Solutions to the Transient Heat Conduction Equation._ MS thesis. Wright-Patterson AFB, Ohio: Air Force Institute of Technology, March 1982.

17. Young, David M. and Louis A. Hageman. _Applied Iterative Methods_. New York: Academic Press, Inc., 1981.

Appendix A

Exact Variational Integral

The variational integral for the one-dimensional heat equation is

$$I(t) = \frac{1}{2} \int_0^L \left[\left(\frac{dt}{dx} \right)^2 + m^2 t^2 \right] dx \qquad (A.1)$$

Integrating the first term by parts allows Eq (A.1) to be written

$$I(t) = \frac{1}{2} \left[t \frac{dt}{dx} \right]_{x=0}^{L}$$

$$- \frac{1}{2} \int_{x=0}^L \left[t \frac{d}{dx} \left(\frac{dt}{dx} \right) - m^2 t^2 \right] dx \qquad (A.2)$$

Since the heat equation is given by

$$\frac{d^2 t}{dx^2} - m^2 t = 0 \qquad (A.3)$$

Eq (A.2) reduces to

$$I(t) = \frac{1}{2} \left[t \frac{dt}{dx} \right]_{x=0}^{L} \qquad (A.4)$$

Using the derivative boundary condition

$$\frac{d}{dx} t(L) = 0 \qquad (A.5)$$

the integral can be written as

$$I(t) = - \frac{1}{2} \left[t \frac{dt}{dx} \right] \Big|_{x=0} \qquad (A.6)$$

63

Substituting the exact temperature at x equal to zero gives

$$I(t) = \tanh(mL) \tag{A.7}$$

Finite-Element Derivation for the 1-D Heat Equation (Ref 11:332-339)

Use linear interpolation functions on the elements for the one-dimensional heat equation as shown in Figure B.1.
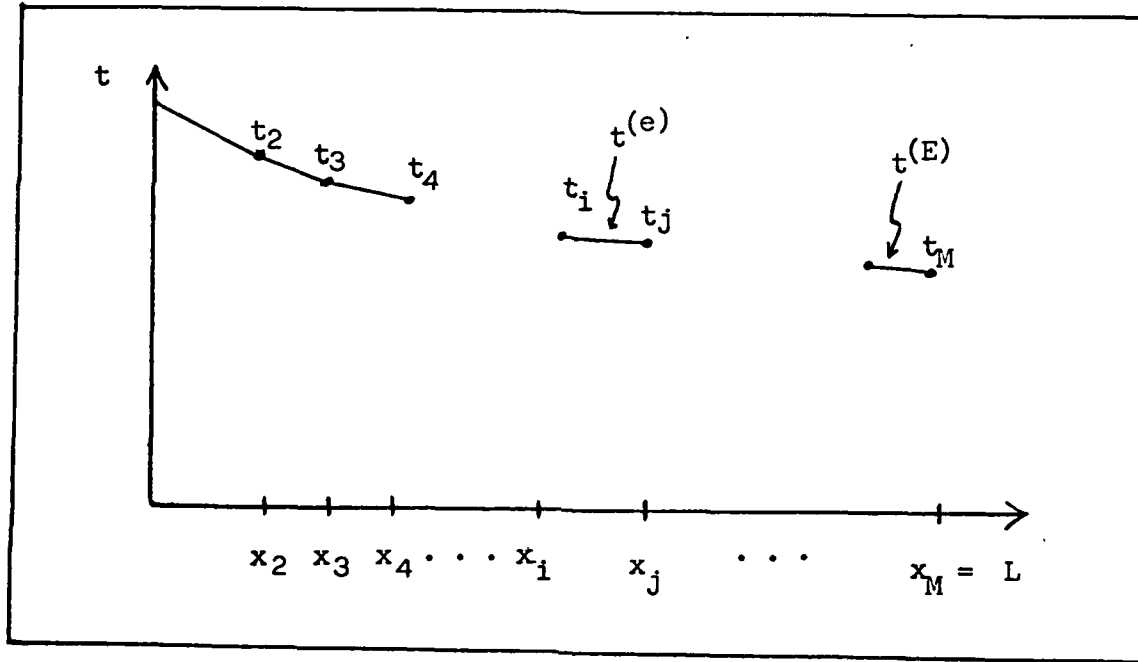

Figure B.1. Finite-Element Temperature Interpolations

Therefore, the temperature within element e is given by $t^{(e)} = C_1^{(e)} + C_2^{(e)}x$. The constants may be determined by evaluating at x_i and x_j:

$$C_1^{(e)} = \frac{x_j t_i - x_i t_j}{x_j - x_i} \qquad \text{(B.1)}$$

$$C_2^{(e)} = \frac{t_j - t_i}{x_j - x_i} \qquad \text{(B.2)}$$

Consequently

$$t^{(e)} = \frac{x_j t_i - x_i t_j}{x_j - x_i} + \frac{t_j - t_i}{x_j - x_i} \qquad \text{(B.3)}$$

and

$$\frac{dt}{dx} = \frac{t_j - t_i}{x_j - x_i} \qquad \text{(B.4)}$$

The integral for a given element can be written

$$I^{(e)} = \frac{1}{2} \int_{x_i}^{x_j} \left(\frac{t_j - t_i}{x_j - x_i} \right)^2 dx$$

$$+ \frac{1}{2} \int_{x_i}^{x_j} m^2 \left(\frac{x_j t_i - x_i t_j}{x_j - x_i} + \frac{t_j - t_i}{x_j - x_i} x \right)^2 dx \qquad \text{(B.5)}$$

After $I^{(e)}$ is integrated, it will be a function of t_i and t_j only. To minimize $I^{(e)}$, it must be differentiated with respect to both t_i and t_j. One can arbitrarily do the differentiation first and then the integration:

$$\frac{\partial I^{(e)}}{\partial t_i} = \int_{x_i}^{x_j} \frac{t_i - t_j}{(x_j - x_i)^2} + \int_{x_i}^{x_j} \frac{m^2}{(x_j - x_i)^2} \left[(x_j^2 t_i - x_i x_j t_j) \right.$$

$$+ (x_i t_j - 2 x_j t_i + x_i t_j) x + (t_i - t_j) x^2 \right] dx \qquad \text{(B.6)}$$

The integration is carried out to give

$$\frac{\partial I^{(e)}}{\partial t_i} = \frac{t_i - t_j}{x_j - x_i}$$

$$+ \frac{m^2}{(x_j - x_i)^2} \left[(\tfrac{1}{3} t_i + \tfrac{1}{6} t_j)(x_j^3 - 3 x_j^2 x_i + 3 x_j x_i^2 - x_i^3) \right] \qquad \text{(B.7)}$$

66

which can be simplified:

$$\frac{\partial I^{(e)}}{\partial t_i} = \frac{t_i - t_j}{x_j - x_i} + \frac{(x_j - x_i)}{6} (2t_i + t_j) \qquad (B.8)$$

Similarly

$$\frac{\partial I^{(e)}}{\partial t_j} = \frac{t_j - t_i}{x_j - x_i} + \frac{(x_j - x_i)}{6} (t_i + 2t_j) \qquad (B.9)$$

Assembling the integral of the elements, we find

$$I(t_1, t_2, \ldots t_m, \ldots t_M) = \sum_{e=1}^{E} I^{(e)} \qquad (B.10)$$

To find the minimum of I, differentiate with respect to each m and set each derivative equal to zero. Since there are M nodes, there will be M equations. Each equations will of the form

$$\frac{\partial I^{(1)}}{\partial t_m} + \frac{\partial I^{(2)}}{\partial t_m} + \ldots + \frac{\partial I^{(a)}}{\partial t_m} + \frac{\partial I^{(b)}}{\partial t_m} + \ldots + \frac{\partial I^{(E)}}{\partial t_m} = 0 \qquad (B.11)$$

The only two elements that will contain t_m in their integrals are the elements immediately adjacent to node m. If the elements denoted (a) and (b) are the elements on either side of node m, then Eq (B.11) reduces to

$$\frac{\partial I^{(a)}}{\partial t_m} + \frac{\partial I^{(b)}}{\partial t_m} = 0 \qquad (B.12)$$

For element (a), use Eq (B.9) and let i = m - 1 and j = m

$$\frac{\partial I^{(a)}}{\partial t_m} = \frac{t_m - t_{m-1}}{\Delta x} + \frac{m^2 \Delta x}{6} (t_{m-1} + 2t_m) \qquad (B.13)$$

67

For element (b), use Eq (B.8) and let $i = n$ and $j = m + 1$

$$\frac{\partial I^{(b)}}{\partial t_m} = \frac{t_m - t_{m+1}}{\Delta x} + \frac{m^2 \Delta x}{6} (2t_m + t_{m+1}) \qquad (B.14)$$

where

$$\Delta x = x_j - x_i \qquad (B.15)$$

Substituting Eqs (B.14) and (B.13) into (B.12) gives

$$- t_{m-1} + Dt_m - t_{m+1} = 0 \qquad (B.16)$$

where

$$D = \frac{12 + 4m^2 (\Delta x)^2}{6 - m^2 (\Delta x)^2} \qquad (B.17)$$

Eq (B.16) holds for all interior nodes. For the first node, the boundary condition can be used:

$$t_0 = 1 \qquad (B.18)$$

For node M, the temperature appears only in element (E). Setting $i = M - 1$ and $j = M$ and using Eq (B.9)

$$\frac{\partial I^{(E)}}{\partial t_m} = \frac{t_M - t_{M-1}}{\Delta x} + \frac{m^2 \Delta x}{6} (t_{M-1} + 2t_M) \qquad (B.19)$$

or

$$-2t_{M-1} + Dt_M = 0 \qquad (B.20)$$

where D is as in Eq (B.17). One now has the M linear algebraic equations to solve simultaneously.

Derivation of an Analytical Solution for Poisson's Equation

The equation to be solved is

$$\frac{\partial^2 t}{\partial x^2} + \frac{\partial^2 t}{\partial y^2} = - \frac{g}{k} \tag{C.1}$$

$$t(x,L) = 0 \tag{C.2}$$

$$t(L,y) = 0 \tag{C.3}$$

$$\frac{\partial t}{\partial x}(0,y) = 0 \tag{C.4}$$

$$\frac{\partial t}{\partial y}(x,0) = 0 \tag{C.5}$$

Assume that the solution can be written as $t(x,y) = v(x,y) + w(x,y)$ where v is the particular solution of Poisson's Equation and w is the solution of the associated homogeneous equation (Ref 16:239):

$$\frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} = - \frac{g}{k} \tag{C.6}$$

$$\frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 w}{\partial y^2} = 0 \tag{C.7}$$

with

$$w(x,L) = - v(x,L) \tag{C.8}$$

$$w(L,y) = - v(L,y) \tag{C.9}$$

$$\frac{\partial}{\partial x}w(0,y) = - \frac{\partial}{\partial x}v(0,y) \tag{C.10}$$

69

$$\frac{\partial}{\partial y}w(x,0) = -\frac{\partial}{\partial y}v(x,0) \tag{C.11}$$

Assume v is of the form

$$v(x,y) = A + Bx + Cy + Dx^2 + Exy + Fy^2 \tag{C.12}$$

Substituting Eq (C.12) into Eq (C.6) gives

$$2D + 2F = -\frac{g}{k} \tag{C.13}$$

Let

$$F = -\frac{g}{k} \tag{C.14}$$

and

$$D = 0 \tag{C.15}$$

The other coefficients are arbitrary so let

$$v(x,y) = \frac{gL^2}{2k} - \frac{g}{2k}y^2 \tag{C.16}$$

so that v reduces to zero on the sides $y = 0$ and $y = L$.

Now solve for w using separation of variables:

$$w(x,y) = X(x)Y(y) \tag{C.17}$$

This gives

$$X''(x) - \lambda X(x) = 0 \tag{C.18}$$

$$X(L) = \left(\frac{gL^2}{2k} - \frac{gy^2}{2k} \right) \frac{1}{Y(y)} \tag{C.19}$$

$$X'(0) = 0$$

70

and

$$Y''(y) + \lambda Y(y) = 0 \qquad \text{(C.21)}$$

$$Y(L) = 0 \qquad \text{(C.22)}$$

$$Y'(L) = 0 \qquad \text{(C.23)}$$

Try

$$Y = A\cos\left(\sqrt{\lambda}\, y\right) + B\sin\left(\sqrt{\lambda}\, y\right) \qquad \text{(C.24)}$$

Apply boundary conditions:

$$Y'(0) = 0 \qquad \text{(C.25)}$$

therefore

$$B = 0 \qquad \text{(C.26)}$$

and

$$Y(L) = 0 \qquad \text{(C.27)}$$

consequently

$$\lambda = \left((2n + 1)\frac{\pi}{2L}\right)^2 \qquad n = 0, 1, 2.. \qquad \text{(C.28)}$$

and

$$Y(y) = \cos\left[(2n + 1)\frac{\pi y}{2L}\right] \qquad \text{(C.29)}$$

From Eq (C.15)

$$X''(x) - \left[(2n + 1)\frac{\pi}{2L}\right]^2 X(x) = 0 \qquad \text{(C.30)}$$

71

Try

$$X(x) = Ce^{rx} + De^{-rx} \tag{C.31}$$

where

$$r = [(2n + 1)\frac{\pi}{2L}] \tag{C.32}$$

Applying boundary condition

$$X'(0) = 0 \tag{C.33}$$

Therefore

$$X(x) = C[e^{rx} + e^{-rx}] \tag{C.34}$$

or

$$X(x) = \cosh[(2n + 1)\frac{\pi x}{2L}] \tag{C.35}$$

By superposition (Ref 15:7)

$$w(x,y) = \sum_{n=0}^{\infty} C_n \cos\left[(2n + 1)\frac{\pi y}{2L}\right] \cosh\left[(2n + 1)\frac{\pi x}{2L}\right] \tag{C.36}$$

Applying the last boundary condition (Eq (C.19)):

$$\frac{gL^2}{2k} - \frac{gy^2}{2k} = \sum_{n=0}^{\infty} C_n \cos\left[(2n + 1)\frac{\pi y}{2L}\right] \cosh\left[(2n + 1)\frac{\pi}{2}\right] \tag{C.37}$$

Multiplying both sides of Eq (C.37) by the term $\cos[(2m + 1)\frac{\pi y}{2L}]$
and then integrating with respect to y:

$$\int_0^L \frac{gL^2}{2k} \cos\left[(2m + 1)\frac{\pi y}{2L}\right] - \int_0^L \frac{gy^2}{2k} \cos\left[(2m + 1)\frac{\pi y}{2L}\right] dy$$

$$= \int_0^L \sum_{n=0}^{\infty} C_n \cos\left[(2m+1)\frac{\pi y}{2L}\right] \cos\left[(2n+1)\frac{\pi y}{2L}\right] \cosh\left[(2n+1)\frac{\pi}{2}\right] dy \tag{C.38}$$

72

For $m \neq n$, the right-hand side equals zero. Therefore

$$C_n = \frac{gL^2}{2k} \frac{(-1)^{n+1} \, 32}{(2n + 1)^3 \pi^3 \cosh[(2n + 1)\frac{\pi}{2}]} \qquad (C.39)$$

and

$$w(x,y) = \frac{gL^2}{2k} \frac{32}{\pi} \sum_{n=0}^{\infty} \frac{(-1)^{n+1} \cos[(2n+1)\frac{\pi y}{2L}] \cosh[(2n+1)\frac{\pi x}{2L}]}{(2n+1)^3 \cosh[(2n+1)\frac{\pi}{2}]} \qquad (C.40)$$

Since

$$t(x,y) = v(x,y) + w(x,y) \qquad (C.41)$$

then

$$t(x,y) = \frac{gL^2}{2k} - \frac{gy^2}{2k}$$

$$+ \frac{gL^2}{2k} \frac{32}{\pi^3} \sum_{n=0}^{\infty} \frac{(-1)^{n+1} \cos[(2n+1)\frac{\pi y}{2L}] \cosh[(2n+1)\frac{\pi x}{2L}]}{(2n+1)^3 \cosh[(2n+1)\frac{\pi}{2}]} \qquad (C.42)$$

Appendix D

Stopping Criteria

The following figures show the error norm, the error norm esti-
mate, and the variational integral stopping criterion for all three
heat equations solved by the finite-difference method. The figures
show that for a sufficiently large number of iterations, the error norm
approximation and the variational integral stopping criterion decrease
at the same rate.

The figures also show that when the iterative solution passes
through the exact solution, the variational integral achieves a local
minimum at a different iteration than does the error norm. In the case
of Laplace's problem, the error norm achieves a local minimum but it is
not pronounced enough to be visible on the graphs.
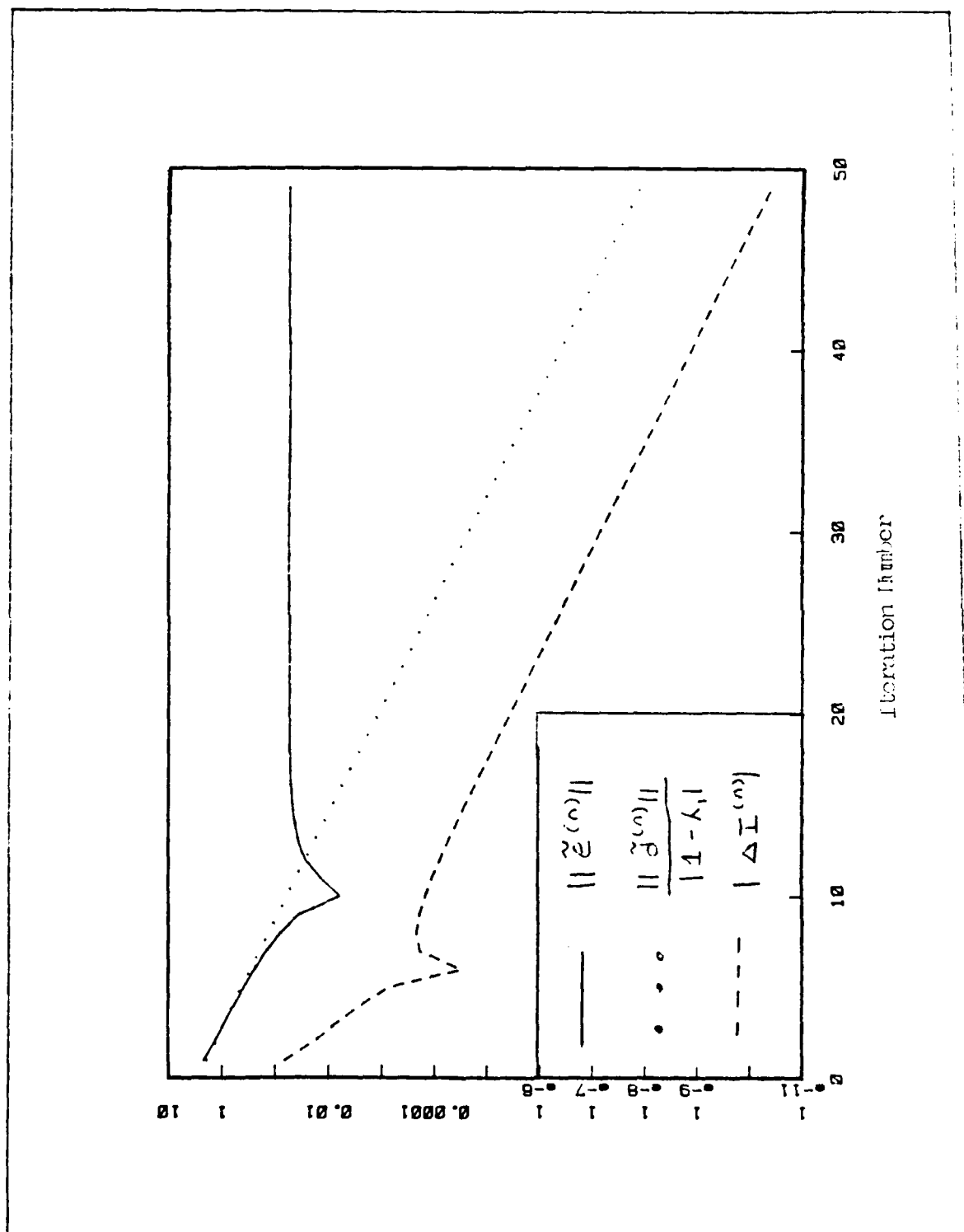
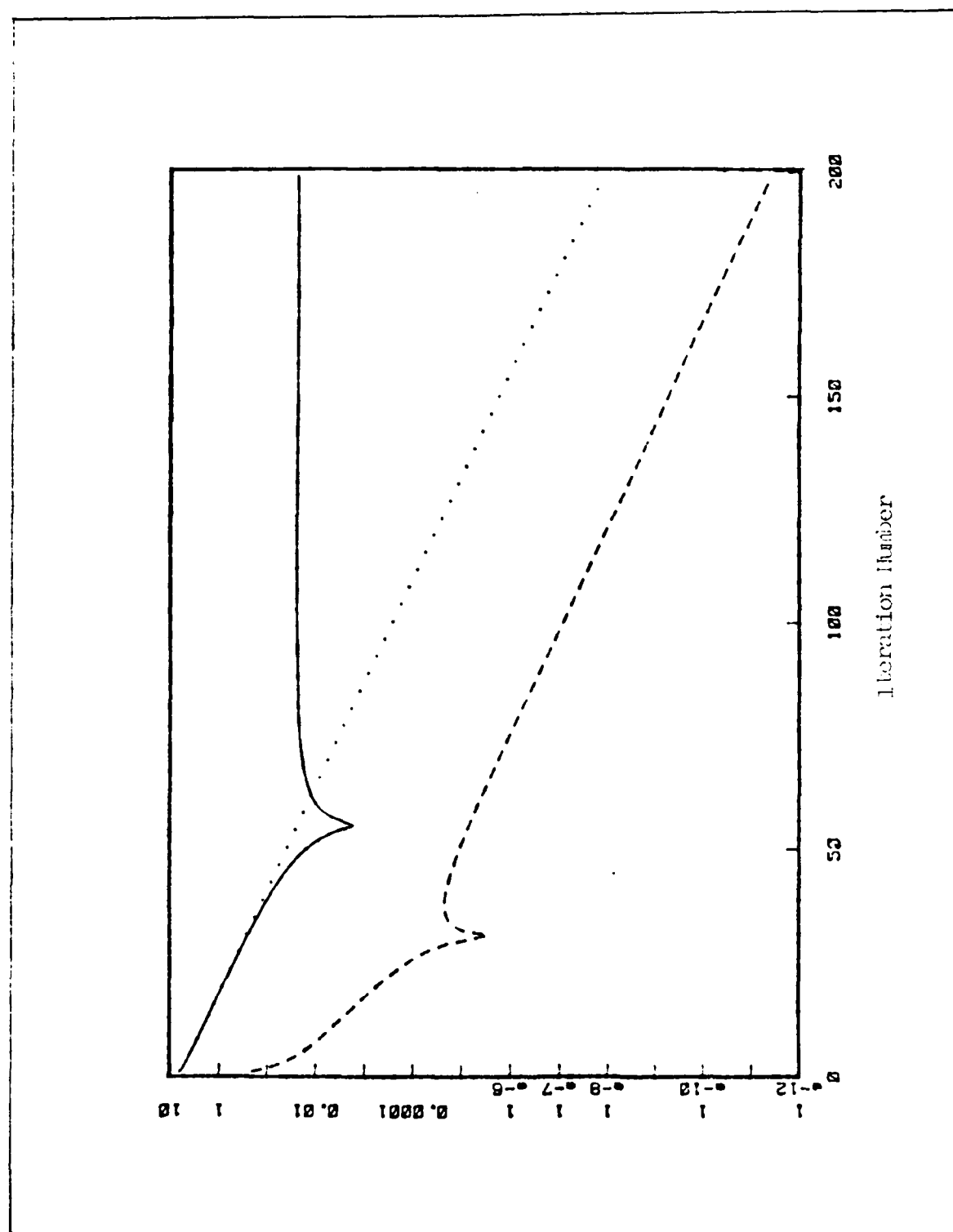Figure D.1. Stopping Criteria, 1-D, Four Nodes
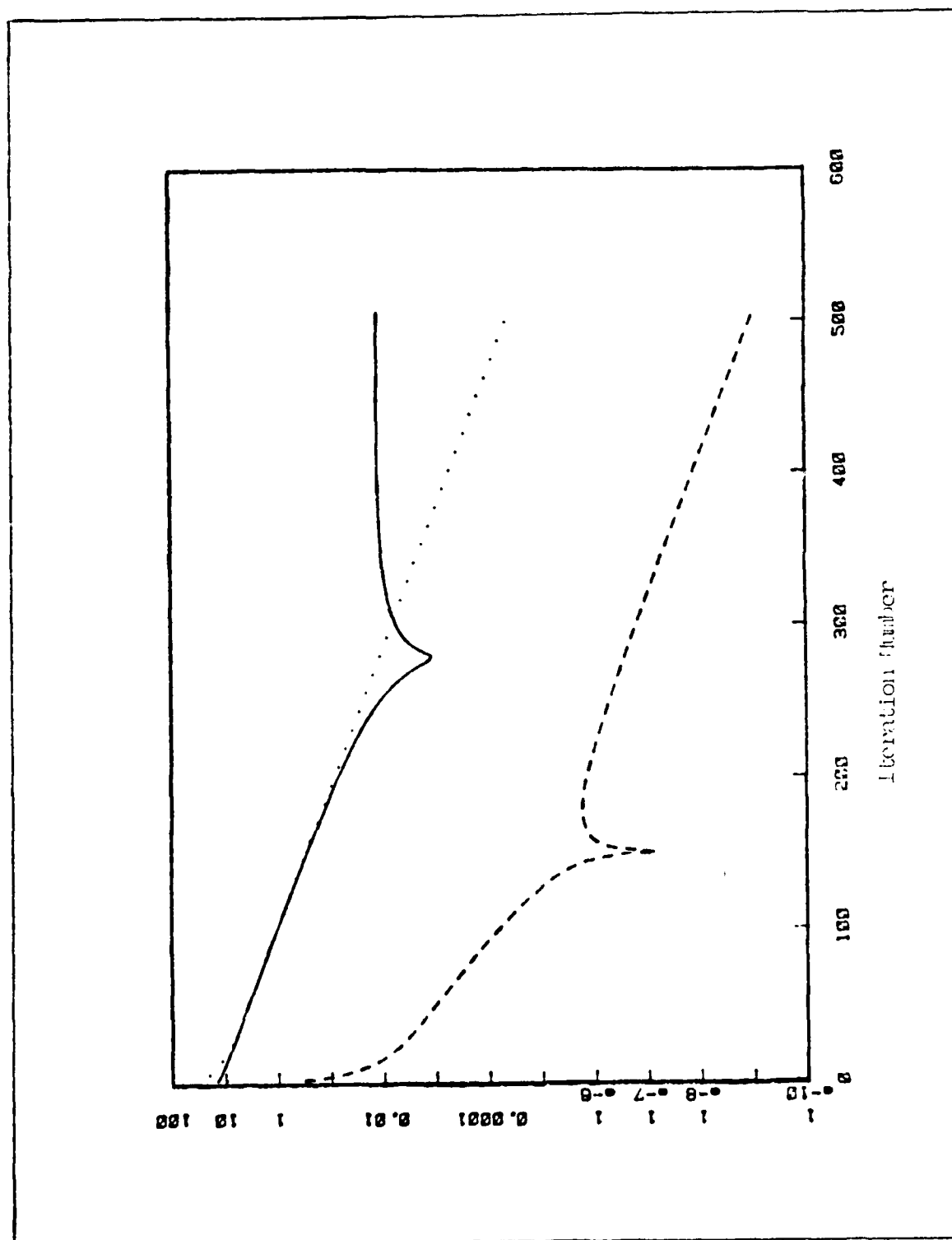
Figure 5.2. Stopping Criteria, 1-D, 3 9 Nodes

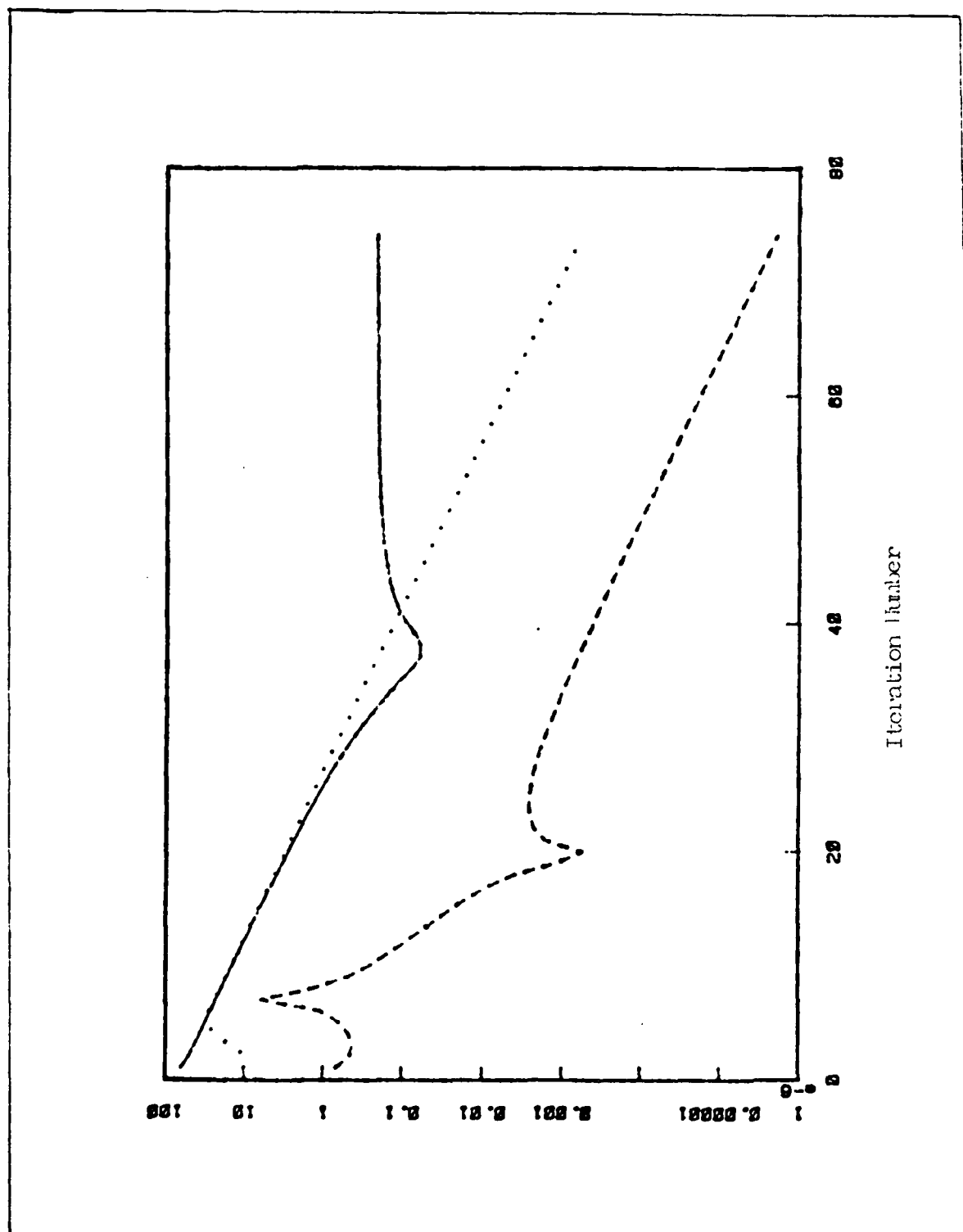Figure D.3. Stopping Criteria, 1-D, Sixteen Nodes

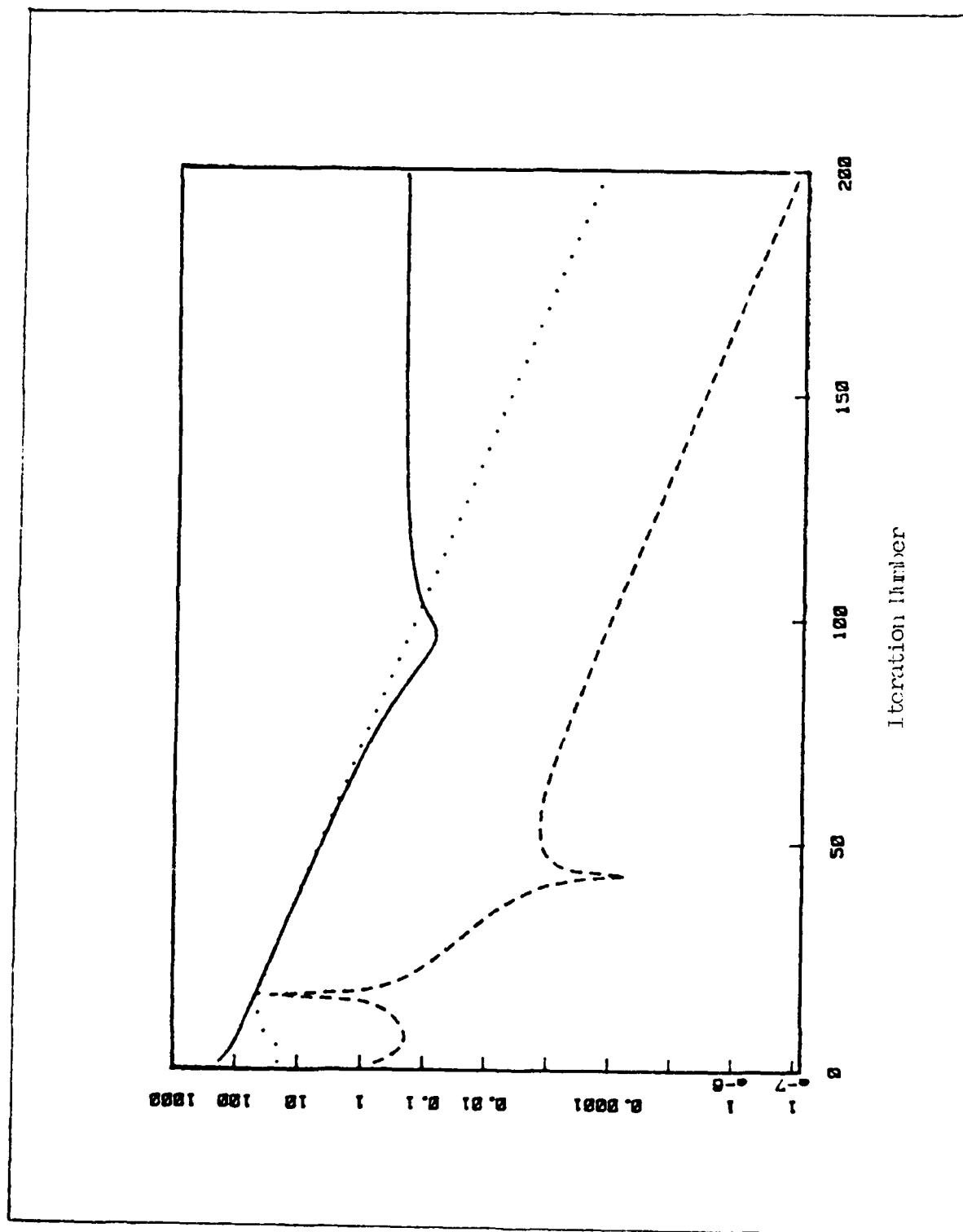Figure D.4. Stopping Criteria, Poisson Problem, 4 x 4 Nodes

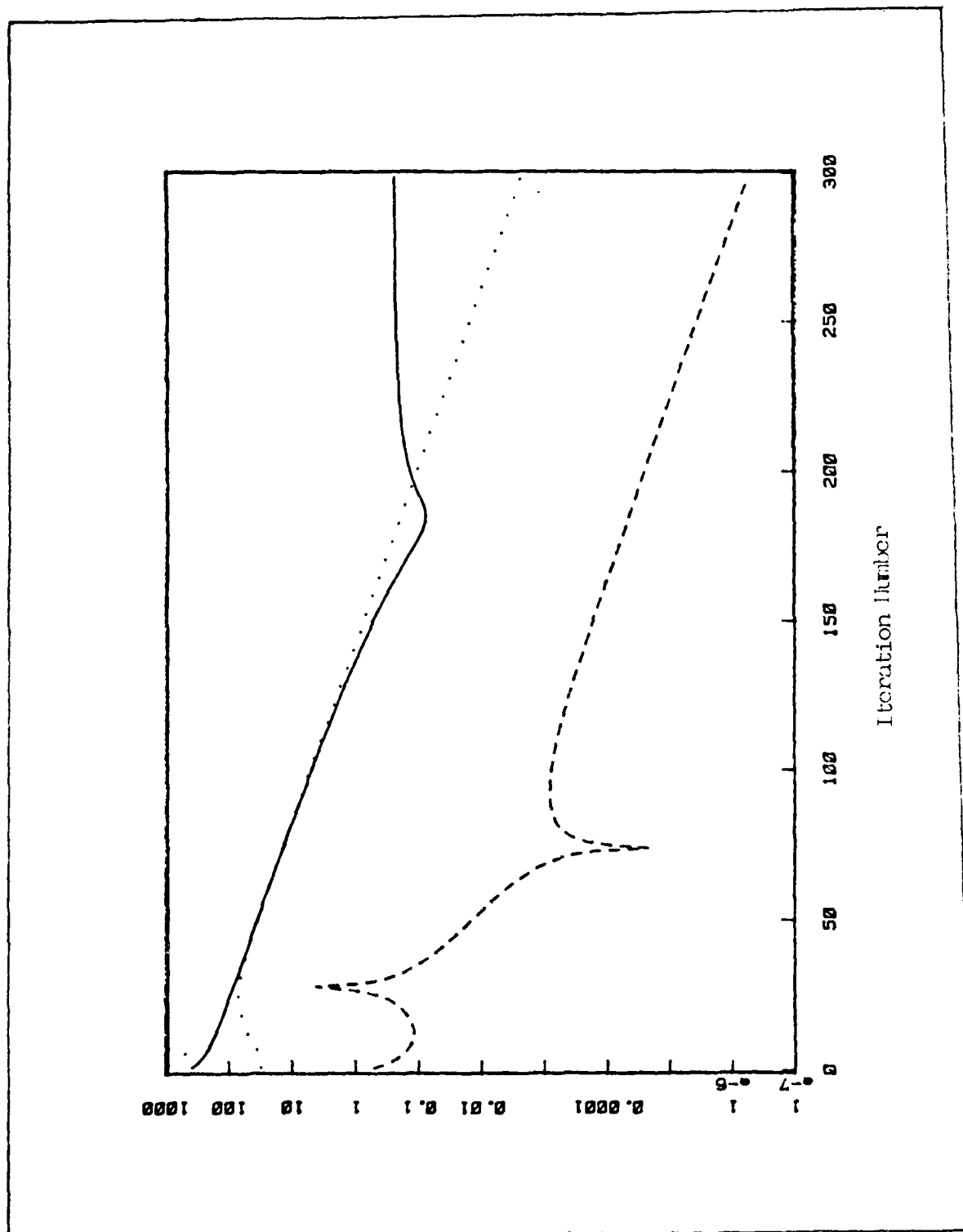Figure D.5. Stopping Criteria, Poisson Problem, 6 x 6 Nodes

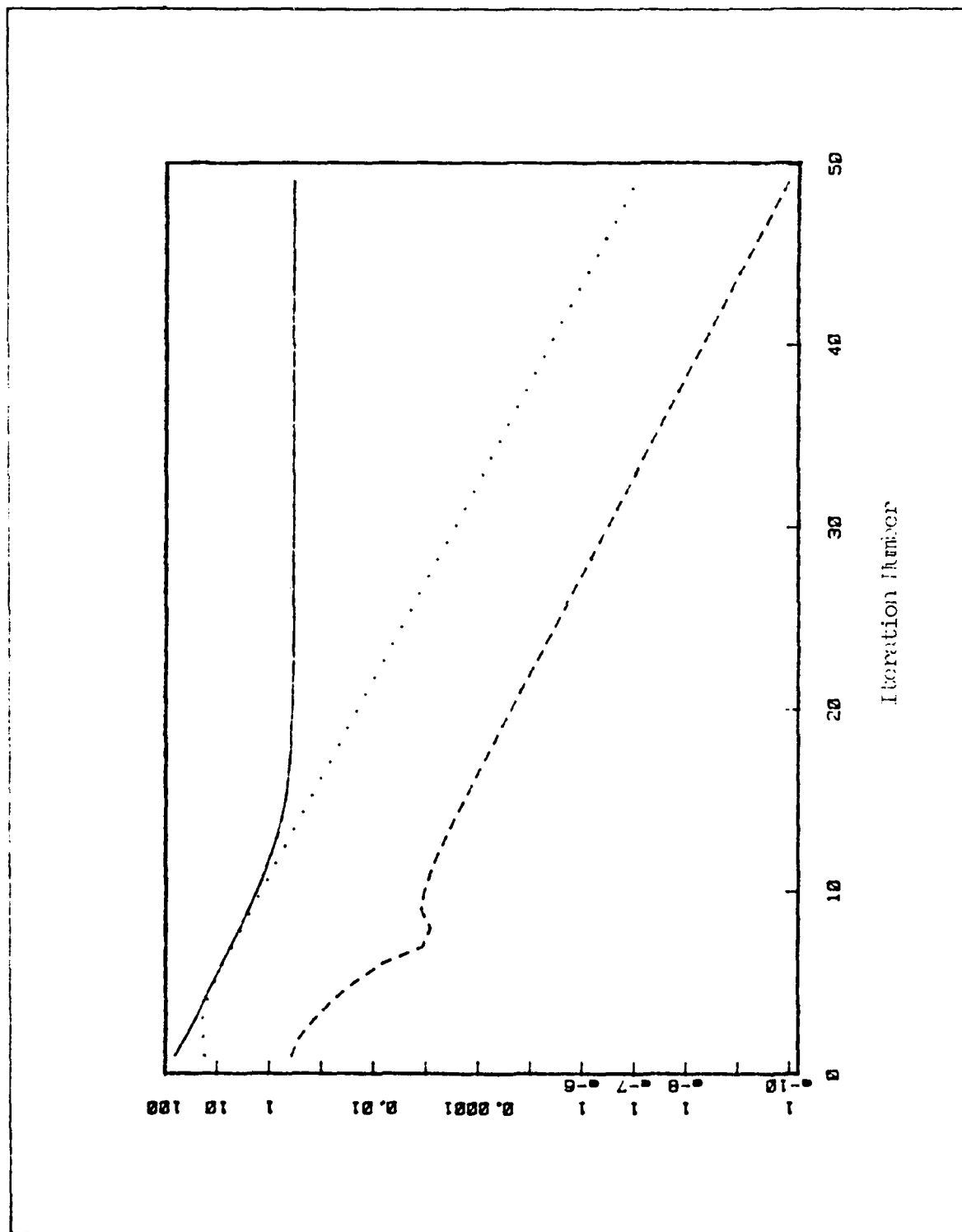Figure 3.6. Stopping Criteria, Poisson Problem, 8 x 8 Nodes

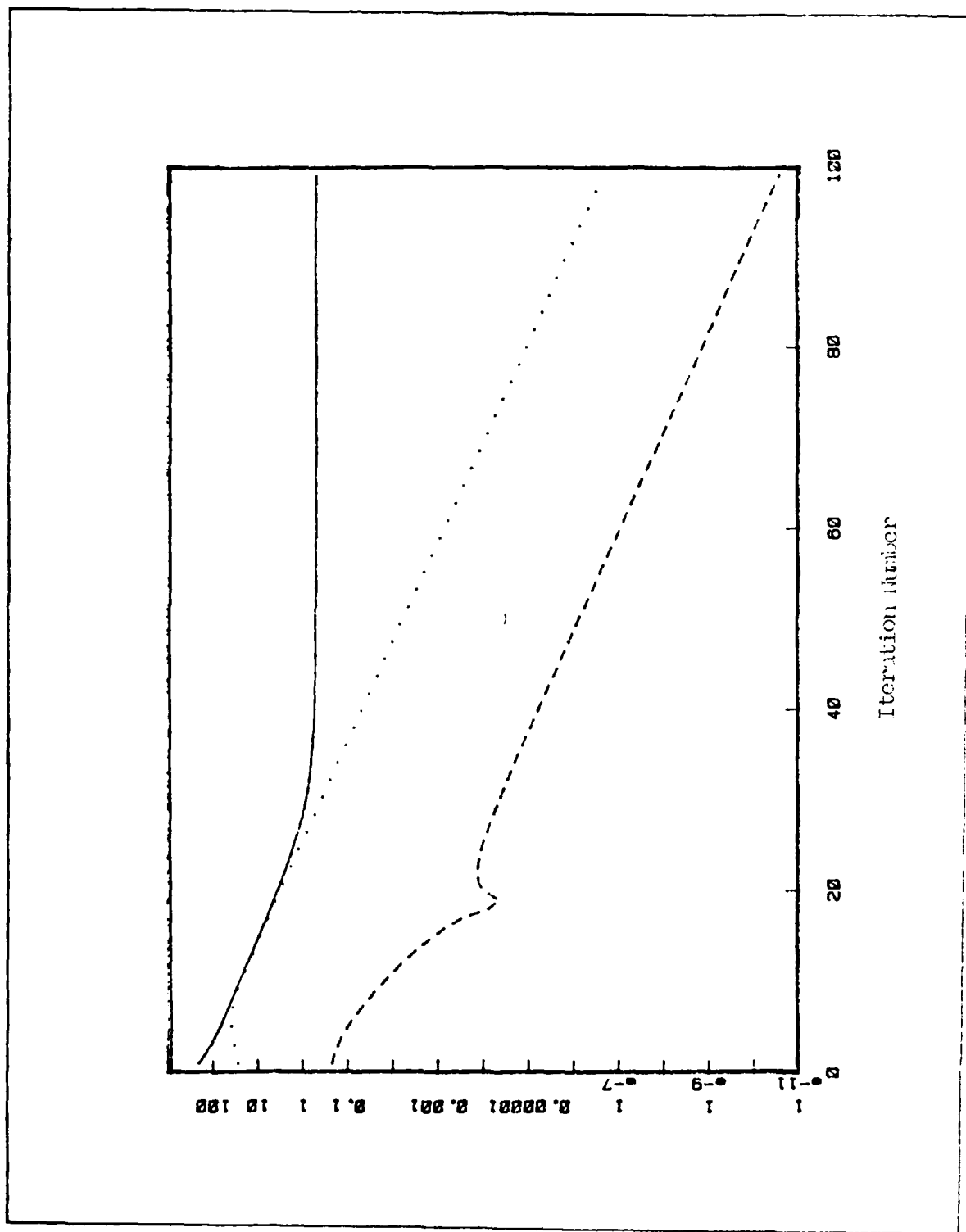Figure D.7. Stopping Criteria, Laplace Problem, 4 x 4 Nodes

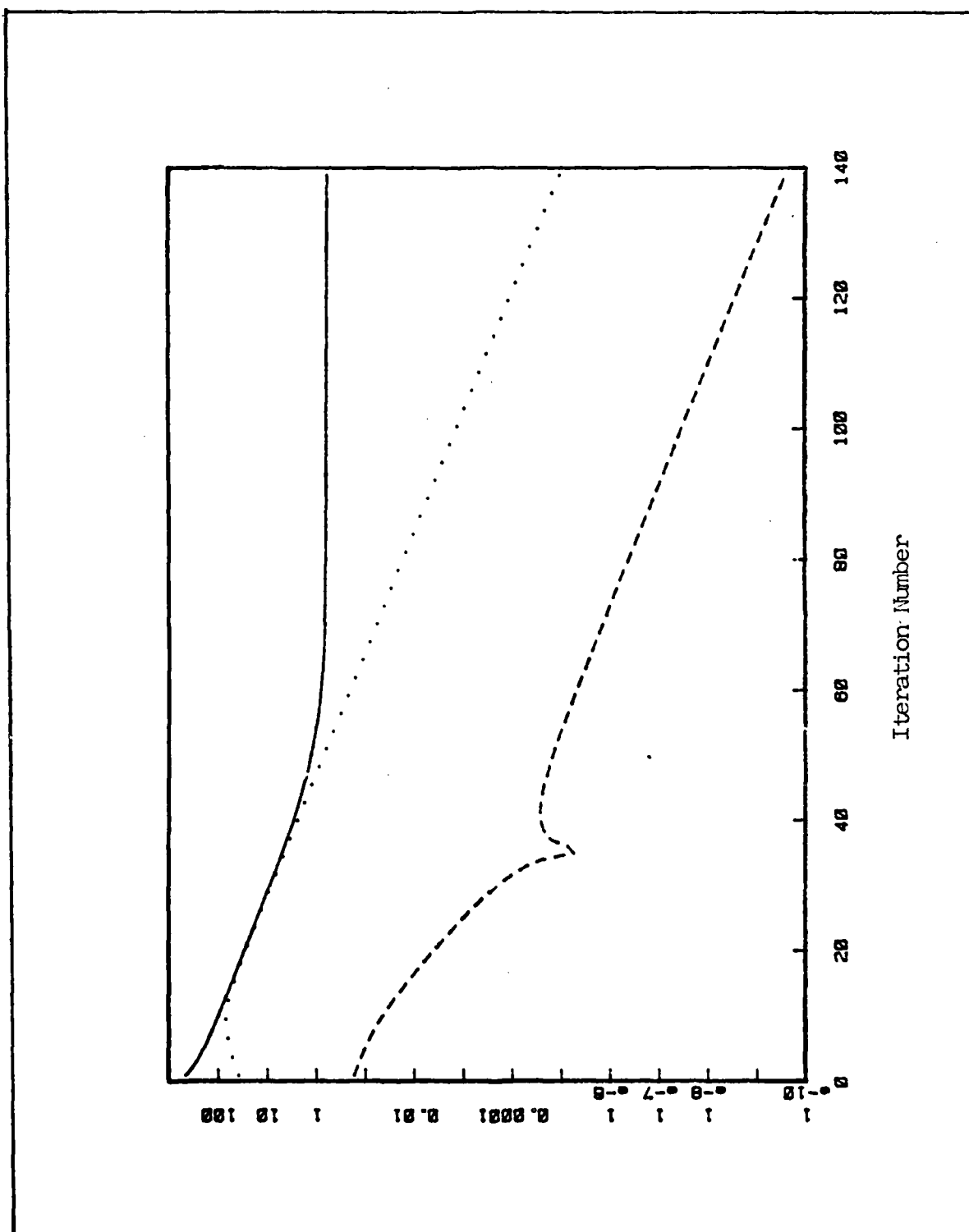Figure 5.3. Stopping Criteria, Laplace Problem, 6 x 6 Nodes

Figure D.9. Stopping Criteria, Laplace Problem, 8 x 8 Nodes

<u>Vita</u>

Mark L. MacDonald was born 19 July 1961 in Augsburg, Germany. After graduating from high school in Leavenworth, Kansas, he enrolled in Iowa State University. In May 1982, he received the degree of Bachelor of Science with Distinction majoring in Physics. Upon graduation, he received a commission in the USAF through the ROTC program. His first assignment was to the School of Engineering, Air Force Institute of Technology.

Permanent Address: 12614 Whispering Hills Avenue

Creve Coeur, MO 63141

REPORT DOCUMENTATION PAGE

1. REPORT SECURITY CLASSIFICATION	1b. RESTRICTIVE MARKINGS
UNCLASSIFIED	

2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT
	Approved for public release;
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	distribution unlimited.

4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S)
AFIT/GNE/ENP/84M-11	

6a. NAME OF PERFORMING ORGANIZATION	6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION
School of Engineering	AFIT/ENP	

6c. ADDRESS (City, State and ZIP Code)	7b. ADDRESS (City, State and ZIP Code)
Air Force Institute of Technology Wright-Patterson AFB, Ohio 45433	

8a. NAME OF FUNDING/SPONSORING ORGANIZATION	8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER

8c. ADDRESS (City, State and ZIP Code)	10. SOURCE OF FUNDING NOS.			
	PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT NO.

11. TITLE (Include Security Classification)
See Box 19

12. PERSONAL AUTHOR(S)
Mark L. MacDonald, B.S., 2d Lt, USAF

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Yr., Mo., Day)	15. PAGE COUNT
MS Thesis	FROM _____ TO _____	1984 March	93

16. SUPPLEMENTARY NOTATION

Approved for public release IAW AFR 190-17.
RYAN E. WOLAVER 7 May 84
Dean for Research and Professional Development
Air Force Institute of Technology (ATC)
Wright-Patterson AFB OH 45433

17.	COSATI CODES		18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB. GR.	Error Criteria, Numerical Convergence, Heat Conduction, Variational Integral
20	13		

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

Title: APPLICATIONS OF THE VARIATIONAL INTEGRAL IN ITERATIVE NUMERICAL SOLUTIONS TO THE STATIONARY HEAT EQUATIONS

Thesis Chairman: Dr. Bernard Kaplan

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT	21. ABSTRACT SECURITY CLASSIFICATION
UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☐ DTIC USERS ☐	UNCLASSIFIED

22a. NAME OF RESPONSIBLE INDIVIDUAL	22b. TELEPHONE NUMBER (Include Area Code)	22c. OFFICE SYMBOL
Dr. Bernard Kaplan	513-255-5533	AFIT/ENP

DD FORM 1473, 83 APR EDITION OF 1 JAN 73 IS OBSOLETE.

(Box 19)

Three stationary heat equations were solved using the finite-difference method. The resulting set of algebraic equations were solved using the Gauss-Seidel iterative technique. The temperatures at the nodal points were substituted into a numerical approximation of the variational integral. The variational integral approximation was used to determine when to stop the iterative process. The variational integral stopping criterion was compared to a stopping criterion that uses the displacement between iterations to approximate the error between the iterative solution and the exact solution. The variational integral was found to be less effective as a stopping criterion than the error estimate.

The variational integral was examined as a method of determining whether the finite-difference technique or the finite-element technique gave a more accurate solution. It was found that the variational integral failed, in some cases, to predict the more accurate method.