the amplifier noise figure was 25 dB. Noise figures in the range 15 to 20 dB have been reported [7] for coaxial amplifiers.

The amplifier results reported here were obtained by using a 10-dB directional coupler to monitor the output. The load VSWR at the amplifier input was less than 1.2:1 over the X band.

Discussion

The similarity between the effects of bias voltage and heat-sink temperature indicate a reduction in carrier mobility due to thermal effects as part of the stabilization mechanism of supercritically doped Gunn-diode amplifiers. Such diodes proved to be easily stabilized in waveguide circuits that could provide low enough values of load resistance. For the circuit shown in Fig. 1 the calculated load resistance decreased from 140 Ω at 8 GHz to 42 Ω at 12.4 GHz.

The demonstration of stable operation of Gunn-diode amplifiers in waveguide circuits opens the possibility of solid-state amplifiers at millimeter wavelengths.

Acknowledgment

The authors wish to thank the former Monsanto Microwave Group for generously providing the Gunn diodes.

References

[1] H. W. Thim and M. R. Barber, "Microwave amplification in a GaAs bulk semiconductor," *IEEE Trans. Electron Devices (Special Issue on Semiconductor Bulk-Effect and Transit-Time Devices)*, vol. ED-13, pp. 110–114, Jan. 1966.
[2] H. W. Thim, "Linear negative conductance amplification with Gunn oscillators," *Proc. IEEE* (Lett.), vol. 55, pp. 446–447, Mar. 1967.
[3] T. E. Walsh, B. S. Perlman, and R. E. Enstrom, "Stabilized supercritical transferred electron amplifiers," *IEEE J. Solid-State Circuits (Special Issue on Optoelectronic Circuits and Solid-State Microwave Circuits)*, vol. SC-4, pp. 374–376, Dec. 1969.
[4] J. Magarshack and A. Mircea, "Wideband CW amplification in X-band with Gunn diodes," in *Dig. IEEE Int. Solid-State Circuits Conf.* (Philadelphia, Pa.), Feb. 1970.
[5] H. Pollman and R. W. H. Englemann, "On supercritical reflection-type amplification and the stability criterion in bulk GaAs devices," in *Proc. MOGA—70 Conf.* (Amsterdam, The Netherlands), Sept. 7–11, 1970, pp. 16-24–16-28.
[6] J. Magarshack and A. Mircea, "Stabilization and wideband amplification using overcritically doped transferred electron diodes," in *Proc. MOGA—70 Conf.* (Amsterdam, The Netherlands), Sept. 7–11, 1970, pp. 16-19–16-23.
[7] B. S. Perlman, "CW microwave amplification from circuit stabilized epitaxial GaAs transferred electron devices," in *Dig. IEEE Int. Solid-State Circuits Conf.* (Philadelphia, Pa.), Feb. 1970.
[8] B. S. Perlman, C. L. Upadhyayula, and W. W. Siekanowicz, "Microwave properties and applications of negative conductance transferred electron devices," *Proc. IEEE (Special Issue on Microwave Semiconductors)*, vol. 59, pp. 1229–1237, Aug. 1971.
[9] D. E. McCumber and A. G. Chynoweth, "Theory of negative-conductance amplification and of Gunn instabilities in 'two-valley' semiconductors," *IEEE Trans. Electron Devices (Special Issue on Semiconductor Bulk-Effect and Transit-Time Devices)*, vol. ED-13, pp. 4–21, Jan. 1966.
[10] G. S. Hobson, *Electron. Lett.*, vol. 2, p. 207, 1966.
[11] L. Young, "Tables for cascaded homogeneous quarter-wave transformers," *IRE Trans. Microwave Theory Tech.*, vol MTT-7, pp. 233–237, Apr. 1959.
[12] J. M. Roe and F. J. Rosenbaum, "Characterization of packaged microwave diodes in reduced-height waveguide," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-18, pp. 638–642, Sept. 1970.
[13] H. W. Thim, "Noise reduction in bulk negative-resistance amplifiers," *Electron. Lett.*, vol. 7, pp. 106–108, Feb. 25, 1971.
[14] M. E. Hines and C. Buntschuh, "Broad-band power amplification with Gunn-effect diodes," *IEEE J. Solid-State Circuits (Special Issue on Optoelectronic Circuits and Solid-State Microwave Circuits)*, vol. SC-4, pp. 370–374, Dec. 1969.

Finite-Element Solutions within Curved Boundaries

DAVID J. RICHARDS AND ALVIN WEXLER

Abstract—**The paper shows that a curved boundary need not be approximated by a small number of finite-element sides, resulting in a coarse polygonal approximation to the shape of the region and consequent inaccuracies, but may be defined as accurately as desired. An algorithm and associated mathematics are presented for locating the stationary point of a functional by the Rayleigh-Ritz method with a two-variable power series as a trial function. As a particular example, the functional employed is one that is made stationary by the solution of Poisson's equation under mixed, Dirichlet, or Neumann boundary conditions. The technique is based on the fact that the three boundary conditions are natural ones. Results are presented for a problem involving curved boundaries under mixed and Neumann conditions and for the capacitance calculations of a pair of noncoaxial cylinders having specified potentials. Comparisons are made with the finite-difference method. It is concluded that the finite-element method is, in nearly all aspects, superior to finite differences—particularly when curved boundary modeling errors are reduced. It is expected that the method described will be equally useful for, and quite simple to adapt to, the solution of the Helmholtz equation in an enclosed region.**

Manuscript received September 22, 1971; revised May 10, 1972. This work was conducted with the support of Atomic Energy of Canada Ltd., Pinawa, Man., Canada.

The authors are with the Department of Electrical Engineering, University of Manitoba, Winnipeg 19, Man., Canada.

I. Introduction

As a particular example of the application of the finite-element technique to regions that are bounded by curves, this paper treats the Poisson equation

$$-\nabla^2\phi = p(x, y) \tag{1}$$

under the following boundary conditions:

$$\left.\frac{\partial\phi}{\partial n}\right|_s + \sigma(s)\phi(s) = h(s) \tag{2}$$

and

$$\phi(s) = g(s) \tag{3}$$

within a homogeneous, isotropic region. The mixed boundary condition (2) holds over a portion of the boundary C_1. (When $\sigma = 0$, this becomes the Neumann boundary condition.) The Dirichlet boundary condition (3) holds over a portion of the boundary denoted C_2. C denotes the entire boundary.

As indicated in [1], the appropriate functional is

$$F(u) = \int_R\int [\,|\nabla u|^2 - 2up(x, y)]\,dxdy + \int_{C_1} [\sigma(s)u^2 - 2h(s)u]\,ds - 2\int_{C_2} [u - g(s)]\frac{\partial u}{\partial n}\,ds \tag{4}$$

for a two-dimensional region as a specific example. u is a trial function approximation to ϕ. As shown in [1], (2) and (3) are boundary conditions that appear naturally at the stationary point of (4).

II. Rayleigh–Ritz Procedure

Because boundary conditions (2) and (3) are natural, we can employ as a trial function a linear combination of functions, none of which need satisfy the stated boundary conditions individually.

A convenient trial function is the polynomial

$$u(x, y) = \sum_{i=1}^{M} c_i x^m y^n \tag{5}$$

where x and y are the independent variables and m and n are integers. The c_i, known as the variational parameters, are coefficients to be determined in an attempt to find the stationary point of (4). This is accomplished by substituting (5) into (4), performing the indicated operations, then differentiating F with respect to each c_i and setting each resulting equation to zero. We are left with a set of M simultaneous linear equations in M unknowns and may solve for all the unknown c_i coefficients. If a sufficient number of terms are employed, $u(x, y)$ should be a reasonable approximation to $\phi(x, y)$.

The order of the polynomial (5) is N. Thus

$$0 \leq m + n \leq N \tag{6}$$

and for all possible mn combinations, there must be

$$M = \tfrac{1}{2}(N + 1)(N + 2) \tag{7}$$

terms.

A simple, linear ordering scheme for the terms of the complete polynomial (5) is obtained by taking

$$i = \tfrac{1}{2}(m + n)(m + n + 1) + n + 1 \tag{8}$$

thus giving an expanded version of (5) that is

$$u(x, y) = c_1 + c_2x + c_3y + c_4x^2 + c_5xy + \cdots. \tag{9}$$

This is the system employed in the program associated with this paper. For convenience, we can write

$$u(x, y) = \sum_{i=1}^{M} c_i f_i \tag{10}$$

where

$$f_i = x^m y^n \tag{11}$$

represents the variable terms of the polynomial and the subscript is set by (8).

Once the coefficients are computed, the field approximation $u(x, y)$ is completely specified. For reasons of convenience, pointed out later, it is often preferred to require that the polynomial collocate with the potential $u_j(x_j, y_j)$ at M particular node points specified by x_jy_j coordinate pairs. (The subscript j indicates that u_j is not a continuous function of position.) To accomplish this, the polynomial (10) is successively expressed at each node point $j = 1, 2, \cdots, M$, i.e.,

$$u_j(x_j, y_j) = \sum_{i=1}^{M} c_i f_{ji} \tag{12}$$

where

$$f_{ji} = x_j^m y_j^n. \tag{13}$$

Thus we have M equations in M unknowns giving

$$\mathbf{u} = B\mathbf{c} \tag{14}$$

where

$$B = [f_{ji}]. \tag{15}$$

Therefore, solving for $\mathbf{c}$, the vector of coefficients of (10),

$$\mathbf{c} = A\mathbf{u} \tag{16}$$

where the square matrix

$$A = [a_{ij}] = B^{-1} \tag{17}$$

is either computed by inversion or assembled by the Lagrange interpolation technique, one example being [2]. Note that difficulties arise in the inversion of B as N

increases. This is due to the fact that B is of Vandermonde type [3, p. 157], [4, pp. 323–324], and consequently becomes increasingly ill-conditioned with N.

By expanding (16) we are able to express the unknown coefficients c_i in terms of the unknown node potentials u_j, i.e.,

$$c_i = \sum_{j=1}^{M} a_{ij} u_j. \tag{18}$$

The variational parameters c_i, in (5), may now be replaced by the new Rayleigh–Ritz parameters u_j.

The source function $p(x, y)$ is defined in terms of potentials p_j, one value at each of the node points $j=1, 2, \cdots, M$, and an order N interpolating polynomial gives $p(x, y)$ as a continuous function. The polynomial coefficients are given by (18), with the a_{ij} terms as previously found and with u_j replaced by p_j.

Functional Stationary Point

To approximate the stationary point of the functional (4) we differentiate F with respect to each variational parameter u_k and set each resulting equation to zero. Rearranging, we obtain

$$\begin{aligned}
&\iint_R \left[\frac{\partial u}{\partial x}\frac{\partial}{\partial x}\left(\frac{\partial u}{\partial u_k}\right) + \frac{\partial u}{\partial y}\frac{\partial}{\partial y}\left(\frac{\partial u}{\partial u_k}\right)\right] dx dy \\
&\quad + \int_{C_1} \sigma u \frac{\partial u}{\partial u_k}\, ds - \int_{C_2} \left[\frac{\partial u}{\partial u_k}\frac{\partial u}{\partial n} + u\frac{\partial}{\partial n}\left(\frac{\partial u}{\partial u_k}\right)\right] ds \\
&= \iint_R \frac{\partial u}{\partial u_k}\, p\, dx dy + \int_{C_1} h \frac{\partial u}{\partial u_k}\, ds \\
&\quad - \int_{C_2} g \frac{\partial}{\partial n}\left(\frac{\partial u}{\partial u_k}\right) ds.
\end{aligned} \tag{19}$$

Equation (19) is arranged so that all unknown u_k appear linearly on the left-hand side and the right-hand side is constant. It is easy to see that this is so. From (10) and (18) we have

$$u = \sum_{j=1}^{M}\left(\sum_{i=1}^{M} a_{ij} f_i\right) u_j. \tag{20}$$

u_j appears linearly in (20), and in $\partial u/\partial x$ and $\partial u/\partial y$ as well. Note that

$$\frac{\partial u}{\partial u_k} = \sum_{i=1}^{M} a_{ik} f_i \tag{21}$$

where a_{ik} is a known constant, $f_i = x^m y^n$ with specified m and n, and u_k does not appear. Also

$$\frac{\partial u}{\partial n} = \frac{\partial u}{\partial x}\frac{\partial x}{\partial n} + \frac{\partial u}{\partial y}\frac{\partial y}{\partial n} \tag{22}$$

where $\partial x/\partial n$ and $\partial y/\partial n$ are direction cosines, has u_k appearing linearly. All x and y variables vanish upon integration and σ, p, h, and g are known functions of position. Therefore, when (19) is repeated for all unknown u_k with $k=1, 2, \cdots, M$, we obtain the matrix equation

$$Su = b \tag{23}$$

where S turns out to be symmetric. If no natural Dirichlet boundary conditions are required, S is positive definite as well. Note that the usual method of specifying a Dirichlet condition along a triangle side is to force the $N+1$ nodes along that side to the Dirichlet value. In this case again S is positive definite.

III. Subsectional Bases

It is often advisable not to assume one polynomial over the entire region of interest, but rather to subdivide the region and consider a polynomial over each resulting smaller region. Thus the finite-element method is a particular implementation of the variational method with subsectional bases. There are at least three practical reasons for the use of a finite-element scheme: 1) it permits the equivalent of finite-difference mesh refinement in regions of rapid field variation; 2) it produces a matrix S that has null matrix blocks, and thus permits storage economies to be effected and block-iterative schemes to be employed in large problems; and 3) it allows complicated boundary shapes to be modeled in a piecewise-linear (i.e., polygonal) sense when elements with straight sides are employed. The last reason is not a very convincing one, as many triangle sides are often required to model a curved boundary accurately. This paper describes an improved scheme for representing curved boundaries accurately.

Triangular Subdivision

A convenient element shape for two-dimensional regions is the triangle, and is consequently the one most frequently employed.

Consider that each triangular subdivision has a piecewise-plane trial function $u(x, y)$ expressed over it. This corresponds to the first three terms of (9). Equivalently, the trial function may be expressed in terms of the potential at each of three noncollinear node points, conveniently located at the vertices of the triangle [5], [6]. Consider also that the potential at node points, at which the vertices of two or more triangles meet, is common to each of the triangles and that sides common to two adjacent triangles are of equal length. Qualitatively, if one thinks of the exact potential solution as a surface over the x–y plane, the finite-element method produces a piecewise-planar approximation to the solution resembling triangular facets cut on a diamond.

The method is improved by defining a polynomial solution approximation, in the two independent space variables, over each triangular interval. Although, due to storage limitations on the total number of free variables, the number of elements must be reduced, the

result is usually a significant increase in accuracy. A good example of this approach, applied to the solution of the Helmholtz equation in waveguides, is furnished by [7].

We will enforce continuity of potential across adjacent triangle sides. Each polynomial is defined in terms of either coefficients or node potentials, the number of such parameters being given by (7). An order N polynomial is uniquely defined along a straight line by requiring it to collocate with potentials at $N+1$ node points. Thus, if $N+1$ nodes are common to two adjacent triangles, we are insured that the potential shall be continuous, although the normal derivative need not be. However, it can be shown that the interface condition—i.e., continuity of flux crossing the boundary between adjacent elements—is a natural one and will be more precisely satisfied as N increases. Therefore, because of this convenience, nodal potentials are employed rather than polynomial coefficients as variational parameters in the finite-element method.

It is usual to place a node at each vertex. As $N+1$ are required along each side, a total of $3N$ nodes are distributed over the boundary. The remaining $M-3N = (N-1)(N-2)/2$ nodes are inserted within the triangle interior.

When the region is not subdivided, (23) results by performing the operations described in (19). For a subdivided region, the functional (4) has a contribution from each subdivision, i.e.,

$$F = \sum F_r. \tag{24}$$

F_r is the contribution from one element region r. Then

$$\frac{\partial F}{\partial u_k} = \sum \frac{\partial F_r}{\partial u_k} = 0. \tag{25}$$

Thus S and $\boldsymbol{b}$ are assembled by accumulating contributions of each subregion.

Curved Boundaries

Most previous implementations of the finite-element method introduced considerable modeling errors in attempting to approximate curved boundaries by triangle sides. The errors occur because there are severe limitations on the number of triangles that one can expect to accommodate. However, there is no need to model a curved boundary with triangle sides. This is because one need not restrict the polynomial trial function to exist only within the triangle, i.e., the polynomial is defined over all space by node potentials only on the triangle. And, conversely, the polynomial trial function need not be used to represent the field over the entire element. It could be restricted to part of an element, although some node points lie outside of that part.

In the work reported in this paper, two vertices of each triangle adjacent to a curved boundary are located on the boundary. The boundary may curve into the triangle or away from it. If the former, that part of each surface integral of (19) is performed not over the entire triangle but only over that part of it in the interior of the region. If the boundary curves away from the triangle side, then the integration is performed over the extra adjacent area, using the polyomial defined within the triangle. This is entirely proper, as the polynomial exists over all space and not just within the region containing the defining node potentials u_j. Likewise, each contour integral follows the boundary exactly.

To actually accomplish this end the boundary could be defined in a piecewise-polynomial sense. Thus the resulting integrations all involve polynomials and may be fairly easily accomplished. In this work, for simplicity, the boundary was defined in a piecewise-linear (polygonal) fashion—thus requiring more computation than would otherwise be needed, as a large number of linear pieces needed to be employed.

The above polygonal approach should not be confused with that of approximating the boundary with triangle sides. The latter approach is limited as to the number of linear segments available, whereas the former permits one to use virtually as many as desired in order to approximate the boundary with a precision consistent with the overall accuracy expected.

IV. Integration Schemes

Following (22), it was pointed out that the integrands of (19) are power series containing terms of the form $x^m y^n$. We shall see that if the boundary is represented by piecewise polynomials of the same type, the resulting integrations are conveniently performed after certain simple polynomial multiplications are carried out. Such operations can be efficiently programmed.

Contour Integrals

Between any two points a and b the contour integral consists of a sequence of terms of the form $\int_a^b x^m y^n\, ds$. Note that the argument can include $\sigma(s)$ and $h(s)$ as specified polynomial functions of position.

The equation of the line connecting a to b is

$$y = \alpha + \beta x \tag{26}$$

where α and β are easily found. Since

$$ds = \sqrt{1+\beta^2}\, dx \tag{27}$$

we get

$$\begin{aligned}\int_a^b x^m y^n\, ds &= \sqrt{1+\beta^2} \int_{x_a}^{x_b} x^m(\alpha+\beta x)^n\, dx \\ &= \sqrt{1+\beta^2} \left\{ \frac{x^{m+1}(\alpha+\beta x)^n}{m+n+1} \Bigg|_{x_a}^{x_b} \right. \\ &\quad \left. + \frac{\alpha n}{m+n+1} \int_{x_a}^{x_b} x^m(\alpha+\beta x)^{n-1}\, dx \right\}\end{aligned} \tag{28}$$

which gives an analytic solution by a recursive definition. This is repeated until $n-1=0$ and the final integration remaining involves only x^m.

If the slope of the straight line becomes too steep, then by writing

$$x = \gamma + \delta y \tag{29}$$

we integrate

$$\int_a^b x^m y^n \, ds = \sqrt{1+\delta^2} \int_{y_a}^{y_b} (\gamma + \delta y)^m y^n \, dy \tag{30}$$

instead.

Note that all contour integrals are performed in a counterclockwise direction about any region.

The unit normal $\hat{n}$ to a point on the boundary is easily defined in terms of an angle θ measured counterclockwise from the positive x direction. Again, considering a small linear boundary segment,

$$\hat{n} = \hat{\imath} \cos\theta + \hat{\jmath} \sin\theta. \tag{31}$$

To obtain the normal derivative of a polynomial u, say, it is only required to multiply each term of u by the appropriate direction cosine:

$$\frac{\partial u}{\partial n} = \frac{\partial u}{\partial x} \cos\theta + \frac{\partial u}{\partial y} \sin\theta. \tag{32}$$

Equation (19) also requires $(\partial/\partial n)(\partial u/\partial u_k)$ and this is obtained in the same fashion.

Surface Integrals

Surface integrals involve terms of the form

$$-\int_{x_a}^{x_b} \int_0^y x^m y^n \, dy dx$$

where the limit y on the inner integral is a function of x. This gives the integral over the region enclosed by the straight lines a–b, $y=0$, $x=x_a$, and $x=x_b$. Therefore,

$$\begin{aligned} -\int_{x_a}^{x_b} \int_0^y x^m y^n \, dy dx &= -\int_{x_a}^{x_b} \int_0^{\alpha+\beta x} x^m y^n \, dy dx \\ &= \frac{-1}{n+1} \int_{x_a}^{x_b} x^m (\alpha + \beta x)^{n+1} \, dx \end{aligned} \tag{33}$$

which may be integrated as in (28). If the slope is steep, a change of variable of integration (from x to y) is made as in (30). Note that for a closed region the directions of integration must be counterclockwise around the boundary.

V. Results

In order to obtain an insight into the accuracy of the results produced by the program, a number of tests were carried out. Problems having analytical solutions (found by separation of variables or conformal transformations) were solved numerically. An error measure

$$E = \sqrt{\frac{\sum_{i=1}^{m} (\phi_i - u_i)^2}{\sum_{i=1}^{m} \phi_i^2}} \tag{34}$$

was evaluated over a discrete-point set. ϕ_i is the exact solution and u_i the computed value at node i. The number of points was approximately $m=400$. Equation (34) is actually the ratio of the Euclidian error norm to the Euclidian function norm (i.e., a fractional error norm) with delta-function weighting factors.

Another useful measure of numerical solution quality is the computation of the flux imbalance F. The fractional flux imbalance is

$$F = \frac{\int_s \left.\frac{\partial u}{\partial n}\right|_s ds}{\int_{s_i} \left.\frac{\partial u}{\partial n}\right|_s ds} \tag{35}$$

where the integral on the denominator is performed along the boundary across which flux enters the region, and the numerator consists of an integral about the entire boundary. A homogeneous region and $p(x, y)=0$ is assumed.

A residual measure R was also defined to indicate, particularly in the absence of an exact solution, the fractional deviation of $\nabla^2 u$ from zero. Thus

$$R = \sqrt{\frac{\sum_{i=1}^{m} (\nabla^2 u_i)^2}{\sum_{i=1}^{m} u_i^2}} \tag{36}$$

is computed by evaluating $\nabla^2 u(x, y)$ at each of the m points specified. Thus we have a fractional residual norm.

The smaller the value of E the more accurate the answer. Usually, F and R decrease with increasing accuracy. Note that we cannot usually obtain E, as the solution is rarely known exactly. However, F and R can be computed from the approximate solution. If we can qualitatively relate F and R to E, then an assessment of the accuracy of a given solution can be made.

Among the types of problems that were solved, for which exact solutions were known, were a rectangular region, a sector of an annulus, and a re-entrant corner. Boundary conditions tested on these shapes were homogeneous and nonhomogeneous Dirichlet, mixed, and Neumann. Problems solved generally included at least two of the above conditions.

With all problems, E decreased as the order of the polynomial increased to $N=5$. In all cases also F and R decreased with increasing polynomial order. For example, in solving the field within the sector of an annulus having mixed boundary conditions, the percentage flux imbalance (i.e., $F\times 100$) dropped from 3.6 to 0.38

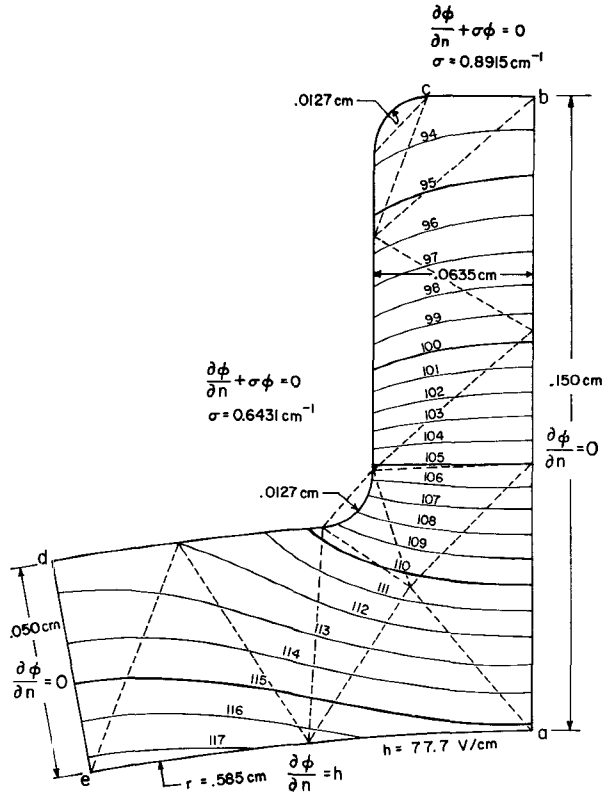

Fig. 1. Equipotential contours V within a region having curved boundaries, given by the solution of Laplace's equation with Neumann and mixed boundary conditions. Dashed lines indicate the finite-element triangle arrangement employed. For clarity, dashed lines are omitted along boundary segments having large radii. $N=2$.

percent, and the percentage error norm (i.e., $E\times100$) dropped from 0.17 to 0.034 percent, while polynomial order increased from three to four using two triangles. Thus E and F may be considered very roughly proportional. These figures are representative of all tests run.

The field in the immediate vicinity of an internal corner is not found accurately, as the power series is incomplete in such a case. However, the field reasserts its correct shape a small distance away. The above-mentioned results also apply in the case of internal corners.

Example with Neumann and Mixed Boundary Conditions

Fig. 1 shows a fairly complicated region bounded by two radial lines (d–e and a–b), one line parallel to a–b, three coaxial curves of fairly large radius, and two arcs having a significantly smaller radius. Neumann and mixed boundary conditions are imposed as indicated. The curve b–d was described by 60 linear segments and a–e by 40 of them so that the shapes plotted on $8\frac{1}{2}$-by-11-in paper looked smooth and not polygonal.

The problem is not just an academic abstraction but is a practical heat-flow problem translated to electrical terms. It represents an asymmetric portion of a section of finned nuclear fuel sheath with a uranium heat source in the center of a pipe that is immersed in coolant. The flow of heat into the coolant is impeded by the effect of the boundary layer, i.e., convective heat dissipation.

TABLE I
FINITE DIFFERENCE–FINITE ELEMENT COMPARISON

	IBM 360/65 Computer		
	Number of Equations	CPU Time	Percentage of Flux Imbalance
Finite difference	155	3.0 min[a]	0.70
Finite element	42[b]	0.1 min	0.06

[a] Optimum acceleration factor used.
[b] Polynomial order $N=2$ and number of triangles equal to 13 as shown in Fig. 1.

This is analogous to a resistive boundary in which the current crossing it is proportional to the voltage difference.

Although the finite-element method is generally considered to be more efficient than the finite-difference method, it could be argued that the latter is sometimes more accurate within curved boundaries. Because of the large number of finite-difference nodes that may be accommodated, by using successive overrelaxation, the boundary may be fairly precisely described. Using finite elements, and direct-solution methods, fewer nodes are employed with only a small number adjacent to the boundary, thus resulting in larger modeling errors due to the coarse polygonal approximation. However, this argument cannot be sustained when the accurate boundary-definition method, described in this paper, is used in the finite-element scheme.

The problem defined in Fig. 1 was also solved by finite differences. The Laplacian operator in polar coordinates was discretized. Uniform angular and radial intervals were used, except near boundaries where nonuniform operators were employed when needed. To speed up the finite-difference program as much as possible, the equations resulting from nonuniform operators were stored and retrieved as required rather than regenerated at each iteration. Table I shows two representative computer runs comparing finite differences and finite elements. A brief glance at the results indicates that the finite-element program produces results that are very much more accurate than finite differences in much less time. The accuracy is improved by perhaps a factor of 10 while the time is reduced by a factor of 30 with roughly equal storage requirements.

As a check, the fractional residual norm R was calculated using the five-point discretization of $\nabla^2 u$ over the difference and variational solutions. R was found to be about 25 times larger in the former than in the latter results. This served to indicate that the difference equations were not as accurately solved as they might have been. Some, but not very much, accuracy could still be gained by continuing the successive overrelaxation (SOR) process at the expense of still more computing time.

There is no acceleration factor that one need be concerned with in the finite-element scheme, the equations being solved directly. The finite-difference results were obtained by using the optimum acceleration factor determined experimentally in advance. Generally, the

factor will not be known and so the 3.0 min central processing unit (CPU) time stated should be considered a lower bound for this problem. The execution time for the finite-element program depends primarily upon the polynomial order of the approximate trial field and increases roughly in proportion to N^3 to N^5 for a polynomial of order N.

Along the radial lines and the inner surface, boundary condition expressions were written using central difference derivative expressions of order h^2. Everywhere else, boundary node potentials were determined by linear interpolation between two adjacent interior nodes and by employing the stated derivative boundary condition. This is a first-order approach and was employed in preference to a more complicated higher order approach. The flux imbalance was not always found to be a monotonic function of mesh interval spacing. For example, upon going from 155 to 200 equations, the heat-flux imbalance *increased* from 0.7 (one of the better figures) to 3.3 percent, and the temperature plots suffered as well. On the other hand, the finite-element results always behaved monotonically and settled down quickly with increasing polynomial order. When the case $N=3$ was run (78 equations), for the problem of Fig. 1, the potential contours could not be distinquished from those presented in this paper.

It has been found that the accuracy of the finite-difference method suffers greatly by having boundaries not coincident with mesh lines. In order to improve the accuracy it is necessary to have a fine grid everywhere, thus causing an unnecessary increase in interior nodes and resulting in increased computational and storage requirements. As an alternative to this, the mesh can gradually be refined in regions adjacent to the boundary. However, this introduces great programming difficulties and dubious accuracies. Previous finite-element schemes have introduced considerable modeling errors by approximating curved boundaries by triangle sides. Here we have been able to describe the boundary to any required accuracy without the expense of introducing additional unknowns.

One should be wary of nonconvex regions, even if the internal curves are far from being corners. Our experience has shown that such bends should be enclosed within one or more triangles of fairly small extent, as in Fig. 1. This has the effect of preventing "contamination" of the rest of the solution. An injudicious rearrangement of triangles can cause the error to increase by an order of magnitude.

Off-Center Cylinders

The capacitance calculation of the pair of noncoaxial cylinders, of Fig. 2, serves as an example of curved boundaries under Dirichlet boundary conditions. Due to symmetry, only half of the cable was considered with an element arrangement as indicated. The capacitance

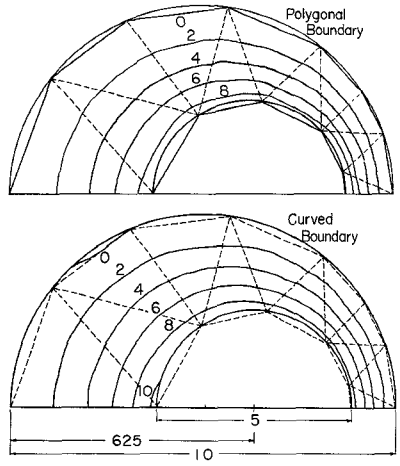

Fig. 2. Off-center cylinders modeled by triangle sides (top) and by small linear segments (bottom) with 90 segments per curve. Equipotential contours are indicated with 1 V between conducting surfaces. $N=3$.

in farads/meter, is given by Smythe [9, pp. 76–78]:

$$C = 2\pi\epsilon\left[\cosh^{-1}\left(\frac{R_1^2 + R_2^2 - D^2}{2R_1R_2}\right)\right]^{-1}. \tag{37}$$

R_1 and R_2 are the radii of the cylinders and D is the distance between their centers.

There are two ways of calculating the capacitance of the system. By computing the stored energy, and noting that the conducting surfaces have a potential difference of 1 V between them,

$$C = \epsilon \iint |\nabla\phi|^2 \, dxdy \tag{38}$$

where the integration is performed over the entire cross section (i.e., twice the region shown in Fig. 2). Alternatively, the capacitance is equal to the charge stored when $V=1$, i.e.,

$$C = \epsilon \int \frac{\partial\phi}{\partial n} \, ds \tag{39}$$

where the integral is performed over the outer or inner conductor.

The usual way of modeling such curved boundaries is by means of triangle sides, resulting in the rather coarse approximation shown at the top of Fig. 2. The Dirichlet condition is enforced. Having introduced this amount of modeling error, one can only expect to deduce accurately the solution of a structure that is of no interest. In the bottom part of the figure, the actual problem is much more faithfully represented by using 90 linear segments for each curve with the boundary condition made natural.

Table II indicates, as one would expect, that the capacitance is more accurately determined when the

TABLE II

Capacitance Calculation Errors for a Pair of Off-Center Cylinders (Fig. 2)

$C_{exact} = 92.229$ pF/m

	N	$\epsilon_0 \int\int \lvert\nabla\phi\rvert^2 \, dxdy$	$\epsilon_0 \int \frac{\partial\phi}{\partial n} ds$
Polygonal boundary (with forced Dirichlet)	2	−1.4 percent	−9.0 percent
	3	−2.4 percent	3.4 percent
Curved boundary (with natural Dirichlet)	2	13.2 percent	0.9 percent
	3	1.3 percent	0.003 percent

boundaries are more realistically represented. The polygonal shape could, fortuitously, give an accurate result for a low polynomial order. But when $N=3$ in our example, the curved boundary results are uniformly better. Note that, for the curved boundary, the charge-determination method gives better results than the energy method. It is not obvious why this should be so. However, there is no reason to expect that the energy integral should give better results when the Dirichlet condition is natural. From [1] it is clear that the energy (and hence the capacitance) is not a stationary function of the trial field $u(x, y)$.

VI. Conclusions

The variational approach has a great many advantages over the traditional finite-difference method. Mainly, the advantages relate to accuracy in both modeling the problem and in computing the answer, and in economy of computation. Another advantage is that the solution is presented in terms of a continuous polynomial (or is easily converted to one by an algorithm already programmed) rather than as a set of potentials at discrete points. Thus interpolation is no problem at all. Oftentimes, the field itself is not the required answer, but some overall function of the field is wanted, e.g., electrical energy stored (a volume integral involving the gradient of potential squared), overall electrical resistance (requiring an integral involving the normal derivative of potential over the boundary), etc. These integrals are conveniently performed using the polnomial field approximation and integration routines within the program.

It has often been stated that finite differences are easier to program than finite elements. This is true only for the simplest problems or when one is content with both long computation times and low accuracies. For instance, to improve the finite-difference scheme to adequately cope with arbitrary boundaries is much more difficult to accomplish than was the program described in this paper.

In regions where the field varies rapidly, one may wish to use smaller triangles in order to maintain overall accuracy. This is the finite-element equivalent of mesh grading but, in marked contrast, it is very easily accomplished.

Because of its efficiency, the finite-element method can be used within an optimization program. In the first stages of optimization, some accuracy can be sacrificed for speed. Reducing the polynomial order from two to one would permit the majority of analyses to be made within 2 s each. Thus there is no great problem in using the finite-element program as an automatic design tool. Reference [10] is an excellent review paper on the subject of optimization.

Although the Helmholtz equation has not been explicitly studied here, there is every reason to expect that accurate definition of boundary shape (in the fashion described in this paper) will yield results improved over those otherwise obtained.

The reader's attention is directed to the isoparametric element [8], [5, pp. 129–153], in which a curved boundary is defined by a few nodes positioned along it and an interpolating polynomial. The isoparametric element is subject to nonunique mapping of coordinate transformations [5, p. 134]. The element described in this paper is not subject to such problems, but may involve more computation. As pointed out in [1, pp. 389–390], making the Dirichlet boundary condition natural destroys positive-definiteness and may reduce the accuracy. On the other hand, because the isoparametric element permits Dirichlet boundary conditions to be enforced along a curve, positive-definiteness is maintained. A detailed comparison of these two alternate approaches would be useful.

Acknowledgment

The authors wish to thank Dr. W. G. Mathers, of Atomic Energy of Canada Ltd., for a large number of very stimulating dicussions on this topic.

References

[1] T. G. Hazel and A. Wexler, "Variational formulation of the Dirichlet boundary condition," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-20, pp. 385–390, June 1972.
[2] P. Silvester, "High-order polynomial triangular finite elements for potential problems," *Int. J. Eng. Sci.*, vol. 7, pp. 849–861, 1969.
[3] I. S. Berezin and N. P. Zhidkov, *Computing Methods*, vol. 1. Oxford, England: Pergamon, 1965.
[4] L. G. Kelly, *Handbook of Numerical Methods and Applications*. Reading, Mass: Addison-Wesley, 1967.
[5] O. C. Zienkiewicz and Y. K. Cheung, *The Finite Element Method in Engineering and Science*. New York: McGraw-Hill, 1971.
[6] O. C. Zienkiewicz and Y. K. Cheung, "Finite elements in the solution of field problems," *The Engineer*, pp. 507–510, Sept. 24, 1965.
[7] P. Silvester, "A general high-order finite-element waveguide analysis program," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-17, pp. 204–210, Apr. 1969.
[8] I. Ergatoudis, B. M. Irons, and O. C. Zienkiewicz, "Curved, isoparametric, 'quadrilateral' elements for finite element analysis," *Int. J. Solids Struct.*, vol. 4, pp. 31–42, 1968.
[9] W. R. Smythe, *Static and Dynamic Electricity*. New York: McGraw-Hill, 1968.
[10] J. W. Bandler, "Optimization methods for computer-aided design," *IEEE Trans. Microwave Theory Tech.* (*Special Issue on Computer-Oriented Microwave Practices*), vol. MTT-17, pp. 533–552, Aug. 1969.