CONSTRAINED VARIATIONAL SOLUTION OF FIELD PROBLEMS

A Thesis

Presented to

The Faculty of Graduate Studies

The University of Manitoba

In Partial Fulfillment

of the requirements for the degree

Doctor of Philosophy

by

Bruce Henry McDonald

February, 1975.

CONSTRAINED VARIATIONAL SOLUTION OF FIELD PROBLEMS

by

BRUCE H. McDONALD

A dissertation submitted to the Faculty of Graduate Studies of the University of Manitoba in partial fulfillment of the requirements of the degree of

DOCTOR OF PHILOSOPHY

ABSTRACT

Scalar field problems are solved variationally, with a guarantee of energy convergence. Static and time-harmonic fields are obtained directly by solution of a partial differential equation using finite-element segmentation of the physical region which may be inhomogeneous or anisotropic. Fields in free-space regions are obtained by seeking layer source distributions in solution to a Fredholm integral equation of the first kind, using only the simple free-space Green's functions with finite-element segmentation of the source layer. Unbounded field problems where there are local inhomogeneities are solved by combining these field and source methods: the inhomogeneities are enclosed in picture-frames and field is sought; and the remaining free-space region is represented by a layer source distribution on the picture-frame boundaries. Field and source finite elements are mutually constrained to solve the picture-frame problem. The Green's function singularities are treated by removing them from general integrals and dealing with them analytically. This addition-subtraction technique is also found useful in handling special singular source distributions which are introduced to improve convergence. The Green's function behaves regularly in one picture-frame method presented: the integral equation is used to relate potential at the picture-frame boundaries to that on a contour within the picture-frames, and the separation between the boundaries and the contour permits numerical methods to be applied directly. The methods are general purpose in nature, and there is no restriction on the shape of inhomogeneous objects.

ACKNOWLEDGEMENTS

The author wishes to thank Professor A. Wexler, advisor and friend, for his tuition, guidance and encouragement over the duration of the project.

Thanks also go to Dr. M. Friedman for his suggestions for handling singular integrands and for the several examples he presented. He and D.J. Richards are to be commended for making the program MANFEP so easy to use; this facilitated development of the mutual constraint picture-frame technique. The author is grateful to Dr. Bill Mathers of AECL Pinawa, for many stimulating discussions about source distribution techniques, and also to M.C. Décreton for the examples he did using singular functions to improve convergence of the finite-element method for field problems. Conversations with G.I. Costache about complex operators were very helpful, and he and Miss G. Jeng are to be thanked for their programming assistance in the development of the integral equation variational procedures.

The author acknowledges gratefully the financial support of the National Research Council of Canada.

Finally, many thanks to Mrs. Aileen Winstanley who typed the manuscript with care and accuracy, and with such speed that Greek symbols appeared in her nightmares.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER I

INTRODUCTION

In his quest to understand the world about him, man has devised theories, based often on empirical evidence, which have helped him to explain natural phenomena. With the rapid growth of scientific knowledge in the past century, it was only to be expected that different theories would arise to explain the same phenomenon and that each theory would have its own advocates today.

Hammond [1] has identified this situation in electromagnetism: "... The history of electromagnetic investigation is the history of the interplay of two fundamentally different modes of thought. The first of these, the method of electromagnetic fields which ascribes the action to a continuum, is associated with such thinkers as Gilbert, Faraday and Maxwell. The second, the method of electromagnetic sources, concentrates attention on the forces between electric and magnetic bodies and is associated with Franklin, Cavendish and Ampere .... field problems are conveniently handled by differential equations and sources by integral equations .... the rigid distinction between field and source methods can be discarded only when it is seen that there is a freedom of choice ...".

The purpose of the work reported here has been to study the use of both methods, separately and simultaneously, in solving static and time-harmonic scalar problems in a region of space Ω. The static problem

requires solution of Poisson's equation

$$-\nabla\cdot(\varepsilon\nabla\phi) = \rho \tag{1.1}$$

and the time-harmonic or propagation problem requires solution of the deterministic Helmholtz equation

$$-\nabla\cdot(\varepsilon\nabla\phi) - W\phi = \rho \quad . \tag{1.2}$$

Here, ϕ is the unknown scalar potential to be determined and ρ is a given source distribution. The permittivity ε is a scalar function of position unless the material in Ω is anisotropic (properties dependent upon direction) in which case ε is a tensor. The propagation function W is a known scalar function which will be specified shortly in terms of the material properties of Ω and the frequency of propagation.

While equations (1.1) or (1.2) prescribe the field behaviour at any point in Ω , it is usually necessary, for complete problem definition, to have ϕ satisfy certain conditions on S, the boundary of Ω . The Dirichlet boundary condition is

$$\phi(\underline{r}) = g(\underline{r}) \; ; \; \underline{r} \text{ on } S_1 \quad . \tag{1.3a}$$

The Neumam boundary condition is

$$\hat{n}\cdot(\varepsilon(\underline{r})\nabla\phi(\underline{r})) = h(\underline{r}) \; ; \; \underline{r} \text{ on } S_2 \tag{1.3b}$$

and the mixed boundary condition is

$$\hat{n}\cdot(\varepsilon(\underline{r})\,\nabla\phi(\underline{r})) + \alpha(\underline{r})\phi(\underline{r}) = h(\underline{r}) \; ; \; \underline{r} \text{ on } S_3 \quad . \tag{1.3c}$$

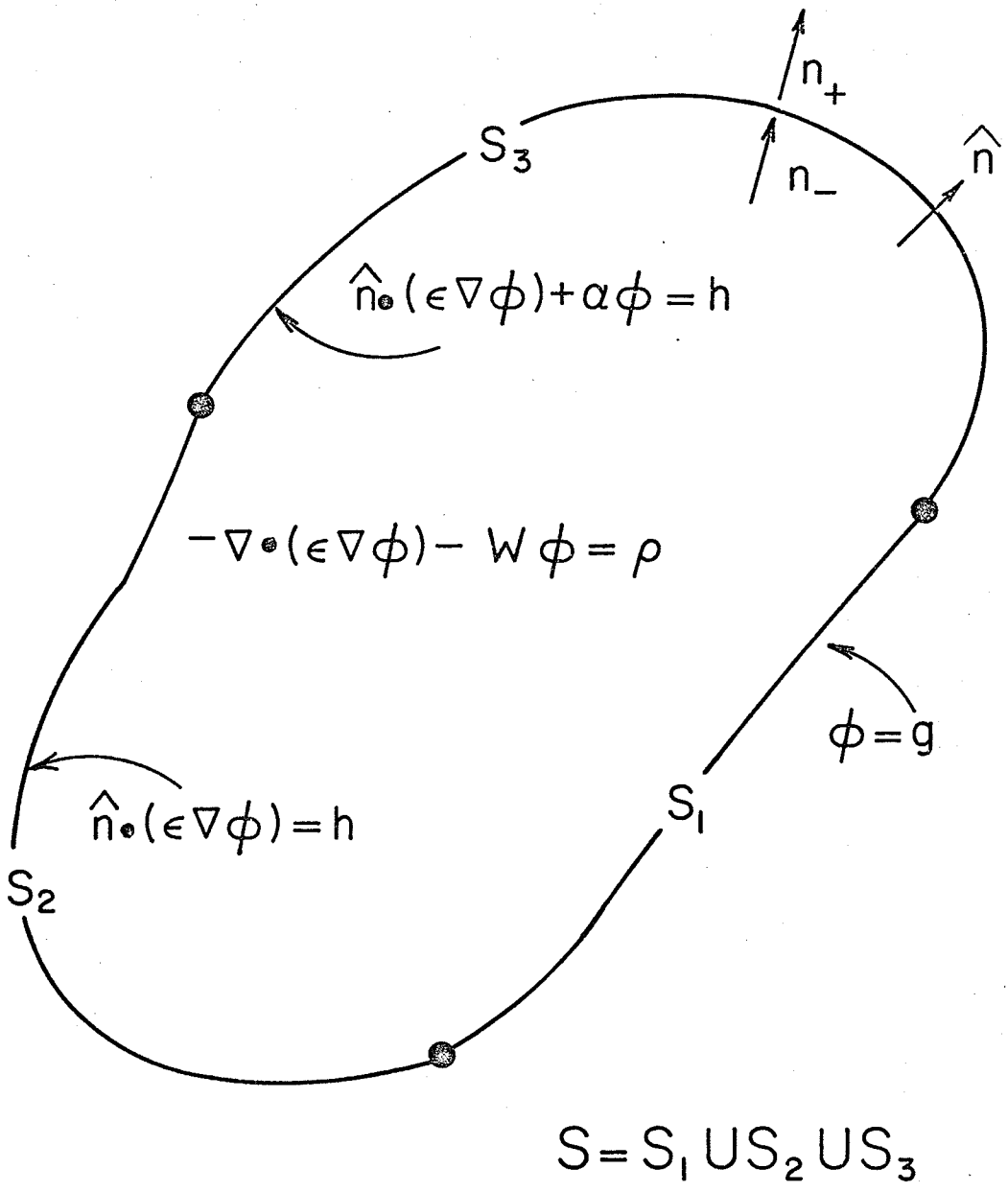


Fig. 1.1 The Boundary Value Problem

The unit vector $\hat{n}$ is taken pointing out of Ω and the boundary of Ω is $S = S_1+S_2+S_3$, as shown in Fig. 1.1. Part or all of S may be at infinity, in which case Ω is an unbounded region.

If the region Ω is anisotropic (i.e. with tensor ε) there is no difficulty in the static case; the scalar potential is adequate. Propagation in such material normally requires a vector formulation if electromagnetic phenomena are to be handled correctly [2, p. 104-109]; the scalar potential ϕ is inadequate. As the work here deals with scalar potential, only isotropic material is considered in the case of propagation. Nevertheless the mathematical treatment is general, with W taken to be a known scalar function.

When the region Ω is isotropic, the propagation function W in (1.2) is simply

$$W = \omega^2\mu\varepsilon^2 \tag{1.4}$$

where ω is the frequency (radians/second) and μ is the permeability of the material in Ω. If the region is homogeneous free space (i.e. $\varepsilon=\varepsilon_o = 8.85 \times 10^{-12}$ farads/meter and $\mu=\mu_o = 4\pi \times 10^{-7}$ henrys/meter), the Helmholtz equation (1.2) assumes the more familiar form

$$- \nabla^2\phi - k^2\phi = \frac{\rho}{\varepsilon_o} \tag{1.5}$$

where the propagation constant k is defined by

$$k^2 = \omega^2\mu_o\varepsilon_o \quad . \tag{1.6}$$

Throughout this work the permeability is taken to be μ_o. (As a matter of interest, note that if the material were anisotropic, with tensor ε, and one attempted to use (1.4), a tensor propogation function W would result. In this case (1.2) would not make sense unless the potential were a vector function.)

Equation (1.5) has an analytical solution in free space which is given by:

$$\phi(\underline{r}) = \frac{1}{\varepsilon_o} \int_{\Omega} \rho(\underline{r}') \, G(\underline{r}|\underline{r}') \, d\Omega' \quad . \tag{1.7}$$

The Green's function for this problem is [3, p. 55]:

$$G(\underline{r}|\underline{r}') = \frac{e^{jk|\underline{r}-\underline{r}'|}}{4\pi|\underline{r}-\underline{r}'|} \quad . \tag{1.8}$$

This is an"action-at-a-distance" function with an integrable singularity at $\underline{r}'=\underline{r}$. It is the solution of the PDE [2]

$$-\nabla^2 G(\underline{r}|\underline{r}') - k^2 G(\underline{r}|\underline{r}') = \delta(\underline{r}-\underline{r}') \tag{1.9}$$

where δ is the Dirac delta function [4, p. 340] and where G satisfies the radiation condition at infinity [5]. If k=0, Poisson's equation is solved with vanishing potential at infinity.

An addition and a subtraction allow (1.8) to be written

$$G(\underline{r}|\underline{r}') = \frac{1}{4\pi|\underline{r}-\underline{r}'|} + \frac{e^{jk|\underline{r}-\underline{r}'|}-1}{4\pi|\underline{r}-\underline{r}'|} \quad . \tag{1.10}$$

The second term of this expression is not singular at $\underline{r}=\underline{r}'$, as

$$\lim_{d\to 0} \frac{e^{jkd}-1}{4\pi d} = \lim_{d\to 0} \frac{1}{4\pi} \{ jk + \frac{(jk)^2 d}{2!} + + \} = \frac{jk}{4\pi} \tag{1.11}$$

and it vanishes at k=0. The first term of (1.10) is the Green's function for Poisson's equation, and contains the integrable singularity. It is important to note that consideration of the singularity for the static case suffices for the propagation case as well.

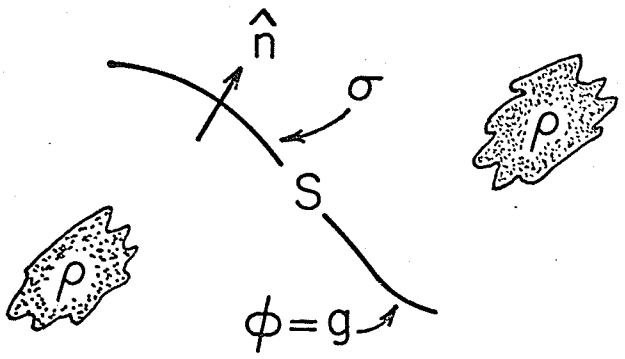
Suppose now that a Dirichlet condition is specified on some surface S in free space. If it should happen that (1.7) generates a potential satisfying $\phi=g$ on S, the problem is solved. If not, it is convenient to place a simple layer of source σ on S, as in Fig. 1.2 and adjust it until the Dirichlet condition is satisfied. Then, throughout free space,

$$\phi(\underline{r}) = \frac{1}{\varepsilon_o} \int_S \sigma(\underline{r}') \, G(\underline{r}|\underline{r}')ds' + \frac{1}{\varepsilon_o} \int_\Omega \rho(\underline{r}') \, G(\underline{r}|\underline{r}')d\Omega' \, . \tag{1.12}$$

To find σ, one must solve a Fredholm integral equation (IE) of the first kind [3, pp.48-49]:

$$\frac{1}{\varepsilon_o} \int_S \sigma(\underline{r}') \, G(\underline{r}|\underline{r}')ds' = g(\underline{r}) - \frac{1}{\varepsilon_o} \int_\Omega \rho(\underline{r}') \, G(\underline{r}|\underline{r}')d\Omega' \; ; \; \underline{r} \text{ on S} \, . \tag{1.13}$$

The right-hand side of (1.13) is completely known, and the unknown source distribution σ appears within an integral. This problem has a unique solution if S is piecewise continuous [6, pp.277-315; 7, pp.180-186] and k is not an eigenvalue of equation (1.13) [5]. One benefit of this approach is that ϕ generated by (1.12) satisfies the PDE (1.5) exactly at any point not on S.



$$\phi = \frac{1}{\epsilon_o}\int_S \sigma G\, ds + \frac{1}{\epsilon_o}\int_\Omega \rho G\, d\Omega$$

Fig. 1.2. The Integral Equation Dirichlet Problem.

The PDE and IE methods may now be compared. The first method requires finding ϕ in Ω which satisfies not only boundary conditions at S but also the PDE in Ω. Inhomogeneities are handled fairly easily: one substitutes (1.2) for (1.5) as the PDE to be satisfied. Approximate solutions often require a number of unknowns proportional to the size of Ω [8, pp.11-14], so that large regions may be less easily accomodated. On the other hand, the IE method requires finding σ on S to satisfy the Dirichlet condition at S, with the result that ϕ may be obtained anywhere in the infinite region Ω by an integration. If the region is not homogeneous the simple free-space Green's function (1.8) is not by itself applicable; dyadic Green's functions [9, pp.1769-1772;

10, p.46-66], or other modified Green's functions [11-12] are often required, complicating the IE procedure. Futhermore, computation is made difficult by the presence of the Green's function singularity; it must be taken into account in any computation [13, p.422-436]. If the region Ω is infinitely large and inhomogeneous, neither method is easily applied: the PDE approach suffers due to the size of the region, and the IE method, due to the presence of the inhomogeneities.

Consider the case where the region is infinitely large with localized inhomogeneities, as shown in Fig. 1.3. Transformation and other analytic techniques [14, pp.72-120] are not easily applied. One may enclose the inhomogeneities in small regions Ω_i called picture-frames, and in each of these the PDE method is applicable, given appropriate conditions on the Ω_i boundaries. The remaining region $\Omega_e = \Omega - \Sigma\Omega_i$ is homogeneous, and the IE method is applicable

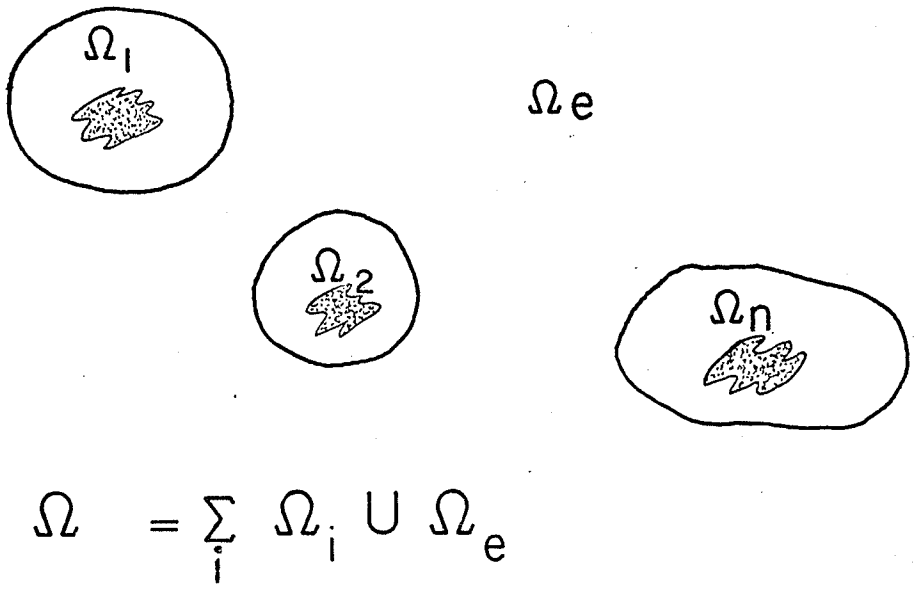


Fig. 1.3. Picture-frame Segmentation of Ω .

with a suitable Dirichlet condition on the Ω_i boundaries. Forcing ϕ to be smooth and continuous at the Ω_i boundaries permits the two methods to be used together, simultaneously, in production of the solution over the entire infinite region. There have been several recent reports of such techniques [15-21], and here, in Chapter III a method is presented whereby the PDE picture-frame solutions, constrained by an IE expression for Ω_e, are obtained variationally using finite elements [22].

In Chapter IV methods are described which permit IE problems to be solved variationally [23-24] and in Chapter VI, variational methods for both the PDE and the IE portions of the picture-frame problem are mutually constrained to produce the solution. In Chapter V a somewhat different approach is presented: when the inhomogeneity is uniform (for example a homogeneous dielectric with relative permittivity ε_r), purely IE variational methods may be mutually constrained to produce the solution [24].

In each example the approximate solution is obtained by solving a system of linear equations:

$$[A]\ \underline{x} = \underline{b} \quad . \tag{1.14}$$

The variational methods described here are used to construct the matrix $[A]$ and the vector $\underline{b}$. While less emphasis has been placed on techniques for solving (1.14) than on techniques for constructing it, certain useful approaches for handling systems of equations are described in Chapter VII.

The presentation begins by describing the application of variational methods to the PDE, using finite-element techniques.

1.1 References

[1] P. Hammond, "Sources and fields - some thoughts on teaching the principles of electromagnetism." Int. Jnl. Elect. Eng. Educ., Vol. 7, pp.65-76, 1969.

[2] L. Cairo and T. Kahan, Variational Techniques in Electromagnetism. London: Blackie, 1965.

[3] I. Stakgold, Boundary Value Problems of Mathematical Physics, Vol. II. New York: Macmillan, 1968.

[4] C.R. Wylie, Advanced Engineering Mathematics. New York: McGraw-Hill, 1960.

[5] A.J. Burton and G.F. Miller, "The application of integral equation methods to the numerical solution of some exterior boundary value problems." Proc. of the Royal Society, Series A, Vol. 323, No. 1553, pp. 201-210, 8 June 1971.

[6] O.D. Kellogg, Foundations of Potential Theory. Berlin: Springer-Verlag, 1929.

[7] N.I. Muskhelishvili, Singular Integral Equations. Groningen: Noordhoff, 1946.

[8] G.E. Forsythe and W.R. Wasow, Finite Difference Methods for Partial Differential Equations. New York: John Wiley and Sons, 1960.

[9] P. Morse and H. Feshback, Methods of Theoretical Physics. New York: McGraw-Hill, 1953.

[10] C.T. Tai, Dyadic Green's Functions in Electromagnetic Theory. Scranton: INTEXT Educational Publishers, 1971.

[11] A.F. dos Santos and V.E. Vieira, "The Green's function for Poisson's equation in a two dielectric region." IEEE Trans., MTT-20, p.777, Nov. 1972.

[12] P. Benedek and P. Silvester, "Capacitance of parallel rectangular plates separated by a dielectric sheet." IEEE Trans. MTT-20, pp.504-510, Aug. 1972.

[13] F.S. Acton, Numerical Methods That Work. New York: Harper and Row, 1970.

[14] W.R. Smythe, Static and Dynamic Electricity. New York: McGraw-Hill, 1968.

[15] I. Cermak, "Solution of unbounded field problems by boundary relaxation." Ph.D. Thesis, Dept. of Elect. Eng., McGill University, Montreal, March, 1969.

[16] I. Cermak and P. Silvester, "Solution of two-dimensional field problems by boundary relaxation." Proc. IEE, Vol. 115, pp.1341-1348, Sept. 1968.

[17] I. Cermak and P. Silvester, "Analysis of Coaxial line discontinuities by boundary relaxation." IEEE Trans. MTT-17, pp.489-495, August 1969.

[18] C.G. Williams and G.K. Cambrell, "Efficient numerical solution of unbounded field problems." Elect. Letts., Vol. 8, pp.247-248, May 1972.

[19] F. Sandy and J. Sage, "Use of finite difference approximations to partial differential equations for problems having boundaries at infinity." IEEE Trans. MTT-19, pp.484-486, May 1971.

[20] D. Greenspan and P. Werner, "A numerical method for the exterior Dirichlet problem for the reduced wave equation." Arch. Ration. Mech. Anal., Vo. 23, pp.288-316, 1966.

[21] P. Silvester and M.S. Hsieh, "Finite element solution of 2-dimensional exterior field problems." Proc. IEE., Vol. 118, pp. 1743-1747, Dec. 1971.

[22] B.H. McDonald and A. Wexler, "Finite-element solution of unbounded field problems." IEEE Trans. MTT-20, pp.841-847, Dec. 1972.

[23] B.H. McDonald, M. Friedman, M. Décreton and A. Wexler, "Integral finite-element approach for solving the Laplace equation." Elect. Lett., Vol. 9, pp.242-244, May 1973.

[24] B.H. McDonald, M. Friedman and A. Wexler, "Variational solution of integral equations." IEEE Trans. MTT-22, pp.237-248, March 1974.

CHAPTER II

BASIC VARIATIONAL METHODS

In this chapter functional minimization is presented as a means of obtaining energy convergent approximate solutions to PDE problems. A proof of convergence for certain applications of Galerkin's Method, given by Mikhlin [1] permits energy convergent solutions to be obtained for propagation problems by seeking functions which make the functional stationary.

The variational approach involves use of the Fundamental Lemma of the calculus of variations [2, p.39] which may be stated as follows: if $\eta(\underline{r})$ is an arbitrary piecewise continuous function in R, the only function u which satisfies

$$I = \int_R \eta(\underline{r})\, u(\underline{r})\, dR = 0 \tag{2.1}$$

is the function $u=0$. If $u>0$ in some elementary region R', it is only necessary to set $\eta=1$ in R' and $\eta=0$ in R-R' to have

$$I = \int_{R'} u(\underline{r})\, dR > 0 \quad . \tag{2.2}$$

Therefore contradictions result unless

$$u = 0 \quad . \tag{2.3}$$

2.1 Linear Operators

Consider the linear operator equation

$$Lu = f \tag{2.4}$$

where the function u, in the domain of the operator (D_L), is sought to make (2.4) true at all points in a region R. The PDE (1.2) and the IE (1.13) of Chapter I are both linear operator equations. In the case of a PDE it is assumed for the moment that u satisfies any given boundary conditions.

An inner product over R may be defined by [3]

$$\langle u,v \rangle = \int_R w(\underline{r})\ u(\underline{r})\ v^*(\underline{r})\ dR \tag{2.5}$$

where w is a weighting function (here taken to be unity) and where, for real variables, the complex conjugate v^* equals v. If u and v are vectors, two inner products may be defined; the scalar form:

$$\langle \underline{u}^T, \underline{v} \rangle = \sum_{i=1}^{N} \langle u_i, v_i \rangle \tag{2.6}$$

and the dyadic form:

$$\langle \underline{u}, \underline{v}^T \rangle = \begin{bmatrix} \langle u_1, v_1 \rangle & \cdots & \langle u_1, v_N \rangle \\ \vdots & & \vdots \\ \langle u_N, v_1 \rangle & \cdots & \langle u_N, v_N \rangle \end{bmatrix} . \tag{2.7}$$

The vector conventions used are described in Appendix A.

The operator L is self-adjoint [1, p.26] if

$$\langle Lu, v \rangle = \langle u, Lv \rangle \tag{2.8}$$

and it is positive-definite [1, p.31] if

$$\langle Lu,u\rangle \begin{cases} > 0 & u\neq 0 \\ = 0 & u=0 \end{cases} \quad . \tag{2.9}$$

The norm of u over R [1, p.15] is given by

$$|| u || = \langle u,u\rangle^{\frac{1}{2}} \tag{2.10}$$

and the energy norm of u over R [1, p.52], by

$$\| u \| = \langle Lu,u\rangle^{\frac{1}{2}} \quad . \tag{2.11}$$

It is clear that the energy norm is real-valued only if L is a positive-definite operator. The operator is positive-bounded-below [1, p.36] if

$$\| u \| \geq \gamma || u|| \quad ; \quad \gamma>0 \tag{2.12}$$

where γ is a real-valued constant.

One useful result of positive-definiteness is that the solution to (2.4), if it exists, is unique [1, p.74] . Suppose that there are 2 solutions, u_1 and u_2. Let $v = u_1-u_2$ and consider

$$Lu_1 - Lu_2 = Lv = 0 \quad ; \tag{2.13}$$

one may then write

$$\langle Lv, v\rangle = 0 \quad . \tag{2.14}$$

But L is positive-definite and therefore $v=0$. Hence

$$u_1 = u_2 \tag{2.15}$$

and the solution is unique. It follows as well that the only solution to $Lv = 0$ is the trivial solution, $v=0$.

The Laplacian operator, $-\nabla^2$, is self-adjoint and positive-definite with homogeneous Dirichlet boundary conditions. Using Green's first identity [1, p.12] and real-valued ϕ:

$$\begin{aligned} \langle -\nabla^2\phi \ \phi \rangle &= \int_\Omega \nabla\phi \cdot \nabla\phi \ d\Omega - \int_S \phi \frac{\partial\phi}{\partial n} ds \\ &= \int_\Omega \nabla\phi \cdot \nabla\phi \ d\Omega \quad . \end{aligned} \tag{2.16}$$

This is proportional to the electrostatic energy U, and for this reason, $\langle Lu,u \rangle$ is called the energy of u in L over R. Mikhlin has established that the Laplacian is positive-bounded-below for finite regions with homogeneous mixed or Dirichlet boundary conditions [1, pp.138-151].

2.2 The Quadratic Functional and Energy Convergence

The Minimal Functional Theorem [1, p.318] states that if L is a positive-definite, self-adjoint operator, the function u which minimizes the functional

$$F(u) = \langle Lu,u\rangle - \langle u,f\rangle - \langle f,u\rangle \tag{2.17}$$

also satisfies the equation Lu=f. Consider the first variation of F(u):

$$\delta F(u) = \langle L\delta u,u\rangle + \langle Lu,\delta u\rangle - \langle \delta u,f\rangle - \langle f,\delta u\rangle \quad . \tag{2.18}$$

The self-adjointness of L allows this to be written as

$$\delta F(u) = \langle \delta u,Lu\rangle + \langle Lu,\delta u\rangle - \langle \delta u,f\rangle - \langle f,\delta u\rangle \quad . \tag{2.19}$$

The stationary point is obtained by setting the first variation to zero:

$$\begin{aligned}\delta F(u) &= \langle \delta u,Lu-f\rangle + \langle Lu-f,\delta u\rangle \\ &= 2\ \mathrm{Re}\colon \langle Lu-f,\delta u\rangle = 0 \quad .\end{aligned} \tag{2.20}$$

Since δu is arbitrary, Lu=f by the Fundamental Lemma. (See the discussion at (2.1).) The second variation of F at the solution point is

$$\delta^2 F(u) = 2\langle L\delta u,\delta u\rangle \quad . \tag{2.21}$$

Since L is positive-definite, this is non-negative (2.9), and the stationary point of F is a minimum. Note that self-adjointness is required to obtain a stationary point, and that positive-definiteness is required to show that this point is a minimum.

Suppose that the solution is to be approximated by a linear combination of real-valued functions with possibly complex coefficients:

$$u_N(\underline{r}) = \sum_{i=1}^{N} c_i v_i(\underline{r}) = \underline{c}^T \underline{v} = \underline{v}^T \underline{c} \quad . \tag{2.22}$$

The error of this approximation is defined by

$$e_N = u - u_N \quad . \tag{2.23}$$

The set of functions $\{v_i\}$ is said to be complete [1, p.65] if

$$\lim_{N\to\infty} ||e_N|| = 0 \tag{2.24}$$

and it is energy complete [1, p.68] if

$$\lim_{N\to\infty} \|e_N\| = 0 \quad . \tag{2.25}$$

If L is positive-bounded-below as well, from equation (2.12)

$$\lim_{N\to\infty} \|e_N\| \geq \lim_{N\to\infty} \gamma ||e_N|| \neq 0 \quad ; \quad \gamma > 0 \tag{2.26}$$

and u_N converges to u in the mean (i.e. in the rms sense). Convergence is said to be in the energy of u_N in L, and the sequence of approximations, u_N, is called a minimizing sequence [1, p.83].

Completeness in energy may be used to derive the quadratic functional (2.17) as follows: from (2.11) and (2.25)

$$\lim_{N\to\infty} \| e_N \| = \lim_{N\to\infty} \langle Le_N, e_N \rangle^{\frac{1}{2}} = 0 \quad . \tag{2.27}$$

Then, using (2.23) and (2.4)

$$\begin{aligned} \langle Le_N, e_N \rangle &= \langle L(u-u_N),\ u-u_N \rangle \\ &= \langle f,u \rangle - \langle Lu_N, u \rangle - \langle f, u_N \rangle + \langle Lu_N, u_N \rangle \geq 0 \ . \end{aligned} \tag{2.28}$$

The inequality follows from the definition of a positive-definite operator (2.9). Now, if L is self-adjoint

$$\langle Lu_N, u \rangle = \langle u_N, Lu \rangle = \langle u_N, f \rangle \tag{2.29}$$

and

$$\langle Le_N, e_N \rangle = \langle Lu_N, u_N \rangle - \langle u_N, f \rangle - \langle f, u_N \rangle + \langle f, u \rangle \geq 0 \quad . \tag{2.30}$$

This may be written as

$$F(u_N) = \langle Lu_N, u_N \rangle - \langle f, u_N \rangle - \langle u_N, f \rangle \geq - \langle f, u \rangle \tag{2.31}$$

and is seen, in comparison with (2.17), to be a statement of the Minimal Functional Theorem. The minimum value of F is $-\langle f,u\rangle$, achieved in the limit as $N\to\infty$. That is

$$\lim_{N\to\infty} F(u_N) = F(u) = -\langle f,u\rangle \quad . \tag{2.32}$$

From this derivation it is evident that minimization of the functional is equivalent to minimization of the energy norm of the error e_N.

2.3 The Rayleigh-Ritz and Galerkin Methods

The Rayleigh-Ritz system of equations [1,p.85] arises by taking $\partial F/\partial c_i = 0$ for each parameter c_i, i=1...N. In vector notation (see Appendix A), the functional (2.17) becomes

$$F(\underline{c}) = \underline{c}^{*T} \int_R L\underline{v}\, \underline{v}^T dR\, \underline{c} - \underline{c}^{*T} \int_R \underline{v}\, f\, dR - \underline{c}^T \int_R \underline{v}\, f^* dR \qquad (2.33)$$

and the variation yields the matrix equation

$$\int_R L\underline{v}\, \underline{v}^T\, dR\, \underline{c} = \int_R \underline{v}\, f\, dR \quad . \qquad (2.34)$$

Since the operator is self-adjoint, the matrix is symmetric and (2.34) may be written

$$\int_R \underline{v}\, L\underline{v}^T dR\, \underline{c} = \int_R \underline{v}\, f\, dR \qquad (2.35)$$

as well. Solving this system of equations produces those coefficients c_i, i=1...N, which give the functional its least value. The value of the functional with this solution is

$$F(\underline{c}) = -\, \underline{c}^{*T} \int_R L\underline{v}\, \underline{v}^T\, dR\, \underline{c} = -\, \langle Lu_N, u_N \rangle \quad . \qquad (2.36)$$

Since L is positive-definite, $\langle Lu_N, u_N \rangle$ is positive in the non-trivial case, and the value of the functional is negative. Minimization of the functional in this case clearly implies maximization of its magnitude $|F(\underline{c})|$, which is the same as maximization of $\langle Lu_N, u_N \rangle$.

In computation, the matrix equation (2.35) is produced directly, knowing $\underline{v}$, L and f for the problem at hand. It is clearly possible to construct the matrix equation when L is neither self-adjoint nor positive-definite. One may lose convergence guarantees in such cases, however. Galerkin's method [1, p.448] involves direct computation of the matrix equation as well. It is a special weighted residual method [4]. Let a residual R_N be defined by

$$\begin{aligned} R_N &= Lu_N - f \\ &= L\underline{v}^T\underline{c} - f \end{aligned} \tag{2.37}$$

and force this residual to be orthogonal to each function v_i over R. This results in

$$\int_R \underline{v}(L\underline{v}^T\underline{c} - f)\, dR = 0 \tag{2.38}$$

or

$$\int_R \underline{v}L\underline{v}^T dR\, \underline{c} = \int_R \underline{v}\, f\, dR \quad . \tag{2.39}$$

This equation is the same as the Rayleigh-Ritz matrix equation (2.35) which was obtained by functional minimization. When L is self-adjoint and positive-definite the Galerkin and Rayleigh-Ritz methods coincide [1, pp.489-490]. But note that the matrix equation (2.38) may be constructed without regard for special properties of L: Often very little may be said about convergence [4, pp.352-398].

Suppose that the equation Lu=f has a unique solution, and that L may be split into two operators:

$$L = L_o + L' \tag{2.40}$$

where L_o is self-adjoint and positive-bounded-below, and where L' is linear and continuous. If the solution is in the domain of L_o, Galerkin's method converges in the energy of u_N in L_o. This is a theorem proved by Mikhlin [1, p.313; pp.469-475], and the result is that L need be neither self-adjoint nor positive-definite, but must possess a component with such properties. The Galerkin matrix equation may be written as

$$\int_R \underline{v} L_o \underline{v}^T dR \; \underline{c} + \int_R \underline{v} L' \underline{v}^T dR \; \underline{c} = \int_R \underline{v} \, f \, dR \quad . \tag{2.41}$$

This may be premultiplied by $\underline{c}^{*T}$. Reference to equations (2.35) and (2.37) permits the result to be expressed as

$$\langle L_o u_N, u_N \rangle + \int_R L' u_N u_N^* \, dR = \int_R u_N^* f \, dR \quad . \tag{2.42}$$

The first term on the left side is clearly the energy in L_o. Since L_o is positive-bounded-below, equation (2.26) applies, provided that u_N is in D_{Lo}. The result is a guarantee of convergence in the mean to the true solution, the same guarantee as obtained when L itself is self-adjoint and positive-bounded-below.

Galerkin's method is a special method of weighted residuals. In general, one forces orthogonality of R_N to a weighting vector $\underline{w}$ [5, pp.230-288]:

$$\int_R \underline{w} \, R_N \, dR = o \tag{2.43}$$

or, using (2.37)

$$\int_R \underline{w} L \underline{v}^T \, dR \; \underline{c} = \int_R \underline{w} \, f \, dR = 0 \quad . \tag{2.44}$$

Galerkin's method results when $\underline{w} = \underline{v}$. When Dirac delta functions are used:

$$w_j = \delta(x-x_j) \; ; \; j=1 \; \ldots \; N \tag{2.45}$$

the collocation (or point-matching) method results:

$$\sum_{i=1}^{N} v_i(x_j)c_i = f(x_j) \; ; \; j=1 \; \ldots \; N \quad . \tag{2.46}$$

When $\underline{w}$ is a vector of the form

$$\underline{w}^T = (1 \; x \; x^2 \; \ldots \; x^N) \tag{2.47}$$

the moment method results. If it happens that $\underline{v}$ is also a vector of the same elements ($\underline{w}=\underline{v}$) the Galerkin and moment methods coincide. There is some confusion in the literature concerning momenclature. What are called weighted residual methods here, after Finlayson [4], are called moment methods by Harrington [6]. Several papers (for example [7-8]) have appeared which describe collocation as a moment method. It is the Galerkin method which is to be recommended for the problems discussed here; for all of the operators dealt with here contain a positive-definite, self-adjoint component, the Laplacian, and energy convergence is guaranteed.

2.4 The Helmholtz Functional and an Example

Consider the propagation problem shown in Fig. 2.1. The PDE (1.2) is to be satisfied at all points in Ω, and ε may be a scalar or tensor (the latter in the case of anisotropic material). The propagation function W is taken to be a scalar function of position. A mixed boundary condition (1.3c) is to be satisfied on S_1 and a Dirichlet condition (1.3a) on the remainder of S. The problem of finding ϕ may be solved by seeking the stationary point of the functional

$$F(\phi) = \int_{\Omega} (\nabla\phi^* \cdot (\varepsilon\nabla\phi) - W\,\phi\phi^* - \phi^*\rho - \phi\rho^*)\, d\,\Omega \qquad (2.48)$$

$$+ \int_{S_1} (\alpha\phi\phi^* - h\phi^* - h^*\phi)\, ds$$

where for the mixed condition on S_1, α is a real number. The Dirichlet condition on $S-S_1$ is enforced by restricting the trial functions to satisfy it.

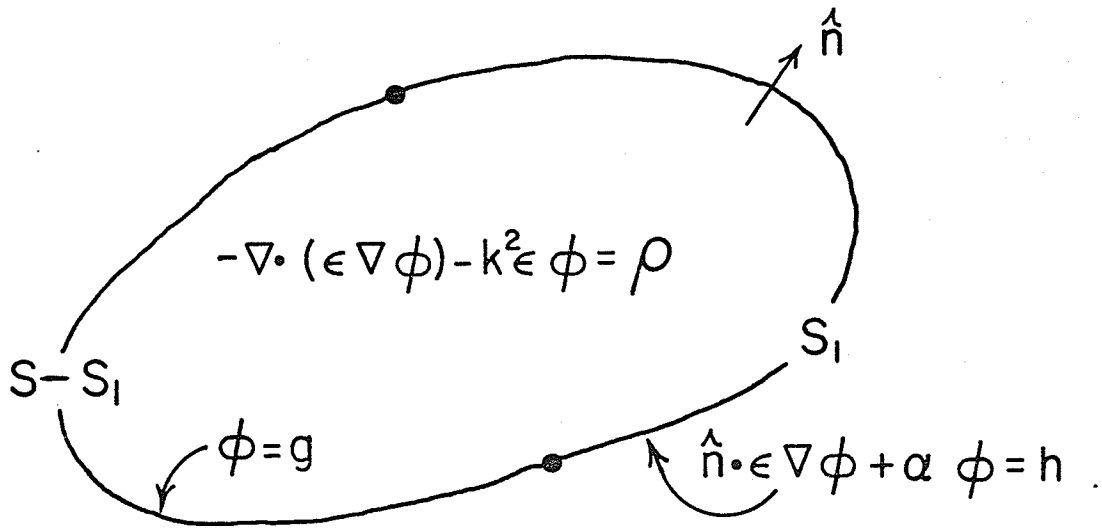


Fig. 2.1 The Helmholtz Problem.

Consider the first variation of (2.48).

$$\delta F(\phi) = \int_{\Omega} (\nabla\delta\phi^{*}\cdot(\varepsilon\nabla\phi) + \nabla\phi^{*}\cdot(\varepsilon\nabla\delta\phi) - W\delta\phi^{*}\phi - W\phi^{*}\delta\phi - \delta\phi^{*}\rho - \delta\phi\rho^{*})d\Omega + \int_{S_1} (\alpha\delta\phi\phi^{*} + \alpha\phi\delta\phi^{*} - h\delta\phi^{*} - h^{*}\delta\phi)ds \quad . \tag{2.50}$$

Green's first identity [1,p.12] may be applied in the following fashion

$$\int_{\Omega} \nabla\delta\phi^{*}\cdot(\varepsilon\nabla\phi)d\Omega = \int_{\Omega} \delta\phi^{*}(-\nabla\cdot(\varepsilon\nabla\phi))d\Omega + \int_{S_1} \delta\phi^{*}(\hat{n}\cdot(\varepsilon\nabla\phi))ds \tag{2.51}$$

to yield

$$\begin{aligned}\delta F(\phi) = {} & \int_{\Omega} \delta\phi^{*}(-\nabla\cdot(\varepsilon\nabla\phi) - W\phi - \rho)d\Omega + \int_{\Omega} \delta\phi(-\nabla\cdot(\varepsilon\nabla\phi^{*}) - W\phi^{*} - \rho^{*})d\Omega \\ & + \int_{S_1} \delta\phi^{*}(\hat{n}\cdot(\varepsilon\nabla\phi) + \alpha\phi - h)ds + \int_{S_1} \delta\phi(\hat{n}\cdot(\varepsilon\nabla\phi^{*}) + \alpha\phi^{*} - h^{*})ds \qquad (2.52) \\ = {} & 2\,\mathrm{Re}\left\{\int_{\Omega} \delta\phi^{*}(-\nabla\cdot(\varepsilon\nabla\phi) - W\phi - \rho)d\Omega + \int_{S_1} \delta\phi^{*}(\hat{n}\cdot(\varepsilon\nabla\phi) + \alpha\phi - h)ds\right\} .\end{aligned}$$

Setting $\delta F(\phi)$ to zero reveals that at the stationary point

$$-\nabla\cdot(\varepsilon\nabla\phi) - W\phi = \rho \; ; \; \underline{r} \text{ in } \Omega \tag{2.53a}$$

and

$$\hat{n}\cdot(\varepsilon\nabla\phi) + \alpha\phi = h \; ; \; \underline{r} \text{ on } S_1 \quad . \tag{2.53b}$$

The Dirichlet condition is satisfied by restriction of the trial functions on $S-S_1$. This forces $\delta\phi=0$ on that portion of the boundary and explains why the boundary integral in (2.51) is defined only over S_1. Note that if $\alpha=h=0$ the integral over S_1 in the functional need not

be computed (it vanishes). In this case the homogeneous Neumann boundary condition is produced naturally on S_1:

$$\hat{n}\cdot(\varepsilon\nabla\phi) = 0 \; ; \; \underline{r} \text{ on } S_1, \; \alpha=h=0 \quad . \tag{2.53c}$$

The second variation of (2.48) reveals whether or not the stationary point is a minimum:

$$\delta^2 F(\phi) = 2\int_\Omega (\nabla\delta\phi^*\cdot(\varepsilon\nabla\delta\phi) - W\delta\phi^*\delta\phi)\,d\Omega + 2\int_{S_1} \alpha\,\delta\phi^*\delta\phi\,ds \quad . \tag{2.54}$$

For $\alpha\geq 0$ the integral over S_1 is non-negative. For $W=0$ (the static problem) the integral over Ω is positive for ε positive (or positive-definite if ε is a tensor). For the propagation case there may exist values of W for which the integral over Ω is not necessarily positive. For example, if ε is constant in Ω, equation (1.5) applies, and the integral over Ω could be negative for a value of the propagation constant $k>\lambda$ where λ is the smallest system eigenvalue. Nevertheless one expects energy convergence in the Laplacian (the self-adjoint and positive-bounded-below component) by Mikhlin's theorem. That is,

$$I = \int_\Omega \nabla e_N^*\cdot(\varepsilon\nabla e_N)\,d\Omega \tag{2.55}$$

should be minimized and should vanish as $N\to\infty$, irrespective of the value of k. Furthermore, as the Laplacian is positive-bounded-below [1,pp.138-151], one should expect from (2.26) that

$$\int_\Omega \nabla e_N^*\cdot(\varepsilon\nabla e_N)\,d\Omega \geq \gamma \int_\Omega \varepsilon e_N^* e_N \,d\Omega \; ; \; \gamma > 0 \quad . \tag{2.56}$$

This points out the usefulness of Mikhlin's theorem to the propagation problem: the convergence guarantee is independent of frequency. Of some importance to analysts however, is that the rate of convergence is not assured and may well deteriorate as frequency is increased.

A simple example will serve to illustrate the application of the Rayleigh-Ritz method to such propagation problems. Consider the one-dimensional problem shown in Fig. 2.2:

$$\frac{-d^2\phi(x)}{dx^2} - k^2\phi(x) = \rho \; ; \; 0 < x < 1 \tag{2.57a}$$

with boundary conditions

$$\phi(0) = 0 \; , \; \phi(1) = 1 \quad . \tag{2.57b}$$

The propagation constant k and the constant source ρ may be varied to facilitate the study of different cases. The Laplacian, $-d^2/dx^2$, is positive-definite (2.16) but depending upon the value of k^2 the

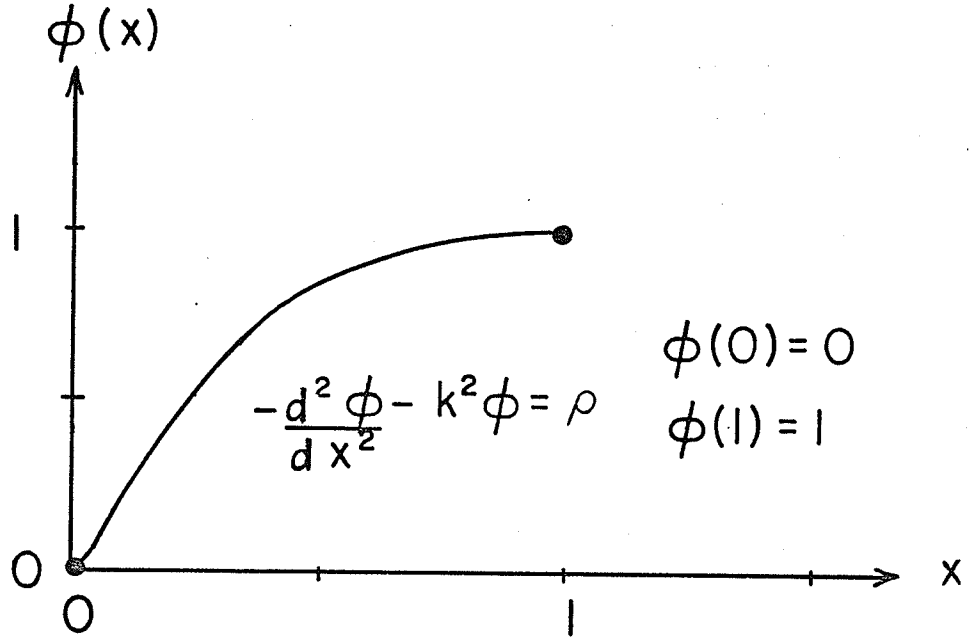


Fig. 2.2 The One-Dimensional PDE Problem:Configuration.

Helmholtz operator may not be.

The true solution to this problem is

$$\phi(\rho, k, x) = \frac{\rho}{k^2}\left(\frac{\sin k(1-x) + \sin k\,x}{\sin k}\right) + \frac{\sin kx}{\sin k} \qquad (2.58a)$$

with the special case

$$\phi(\rho, o, x) = \frac{\rho}{2}\,(x-x^2) + x \quad . \qquad (2.58b)$$

The eigenvalues are $\pi/2$, $3\pi/2$ .... $(2\eta-1)\pi/2$.., and the solution is not defined for $k=\eta\pi$, for which values the denominator in (2.58a) vanishes.

Suppose that ϕ is to be approximated by

$$\phi_N(x) = \sum_{i=1}^{N} c_i\, x^i = \underline{f}^T\underline{c}\ ; \quad \sum_{i=1}^{N} c_i = 1 \quad . \qquad (2.59)$$

This trial function satisfies the Dirichlet boundary conditions. One may, at first, ignore the condition at x=1 and write the functional (2.48) in the form of (2.33):

$$F(\underline{c}) = \underline{c}^T \int_o^1 \frac{d\underline{f}}{dx}\frac{d\underline{f}^T}{dx}\,dx\ \underline{c} \quad -\underline{c}^T k^2 \int_o^1 \underline{f}\underline{f}^T\,dx\ \underline{c} - 2\underline{c}^T\rho \int_o^1 \underline{f}\,dx \quad . \qquad (2.60)$$

The first term on the right side is the energy of the approximation in the Laplacian, U_L. The matrix equation is obtained by taking $\partial F/\partial c_i = o$, $i=1 \ldots N$:

$$\left(\int_o^1 \frac{d\underline{f}}{dx}\frac{d\underline{f}^T}{dx}\,dx - k^2 \int_o^1 \underline{f}\,\underline{f}^T dx\right)\underline{c} = \rho \int_o^1 \underline{f}\,dx \qquad (2.61a)$$

or $$\left[A - k^2 B\right]\underline{c} = \rho\,\underline{b} \quad . \qquad (2.61b)$$

The matrix and vector entries are, quite simply

$$A_{ij} = ij/(i+j-1) \tag{2.62a}$$

$$B_{ij} = 1/(i+j+1) \tag{2.62b}$$

$$b_i = 1/(i+1) \quad . \tag{2.62c}$$

Note that the matrix equation (2.61b) may be constructed directly using (2.62). In practice one does not start with (2.60), but rather with (2.61a). For this example (2.61a) is also the Galerkin matrix equation (see Mikhlin [1, pp.489-490]). Thus the example is one also of the application of Galerkin's method.

To this point no attempt has been made to enforce the Dirichlet condition at x=1. Solving (2.61) as it stands will give a solution satisfying the natural condition, $\partial\phi/\partial x = o$, at x=1. The Dirichlet condition at x=1 is now imposed by using a Lagrange constraint [2, pp.12-17]; the matrix equation (2.61b) is augmented by one equation:

$$\begin{bmatrix} A - k^2 B & \underline{\ell} \\ \underline{\ell}^T & 0 \end{bmatrix} \begin{pmatrix} \underline{c} \\ \lambda \end{pmatrix} = \begin{pmatrix} \rho\, \underline{b} \\ 1 \end{pmatrix} \quad . \tag{2.63}$$

The vector $\underline{\ell}$ is a vector of ones:

$$\underline{\ell}^T = (1\ 1\ 1\ \ldots\ 1) \tag{2.64a}$$

whence the additional equation is

$$\underline{\ell}^T \underline{c} = \sum_{i=1}^{N} c_i = 1 \quad . \tag{2.64b}$$

The method of computation is quite straightforward. For given values of ρ, k, and N the matrix equation (2.61) is constructed and solved by Gaussian elimination and back-substitution. (See Chapter VII.) The vector of coefficients $\underline{c}$ may then be used to compute ϕ_N at any

point from (2.59). The error e_N of the approximation is easily obtained at any point as the difference between ϕ_N and the analytical solution at that point (2.58), and the error norm $||e_N||$ (2.10) is readily computed by numerical integration. Similarly, the approximate and exact solutions are differentiated analytically to give de_N/dx at any point, and the energy norm of the error in the Laplacian operator $\|e_N\|$ (2.55) is computed numerically.

The value of the functional is computed from

$$F(\underline{c}) = \underline{c}^T \left[A - k^2 B\right] \underline{c} - 2 \underline{c}^T \underline{b} \rho \qquad (2.65)$$

and the energy in the Helmholtz operator, U_H is

$$U_H(\underline{c}) = \underline{c}^T \left[A - k^2 B\right] \underline{c} \qquad (2.66)$$

When the Helmholtz operator is positive-definite, U_H is positive and the coefficient matrix in (2.61b) is positive-definite. Conversely if U_H is negative, the coefficient matrix in (2.61b) is not positive-definite, and the functional (2.60) is not necessarily minimized. Note that the Lagrange multiplier λ is not used in these computations; it is simply an extra variable included to permit the constraint equation (2.64b) to be added to the system.

Several examples are presented. The homogeneous case (ρ=0) and an inhomogeneous case (ρ=2) were each tested with k at 0.5, 2.0, and 4.0. For each value of ρ and k the polynomial order N was varied from 1 up to a maximum of ten.* The graphs of $||e_N||$ and $\|e_N\|$ (energy in the Laplacian), shown in Fig. 2.3, illustrate two points. First of all, it is seen that the maximum value of N used varies from 5 to 8, depending upon the case. Higher values of N produced no further improvement, indicating round-off error (due probably to ill-conditioning; see Chapter VII). The second point is that in all

* Double precision arithmetic was used in this example.

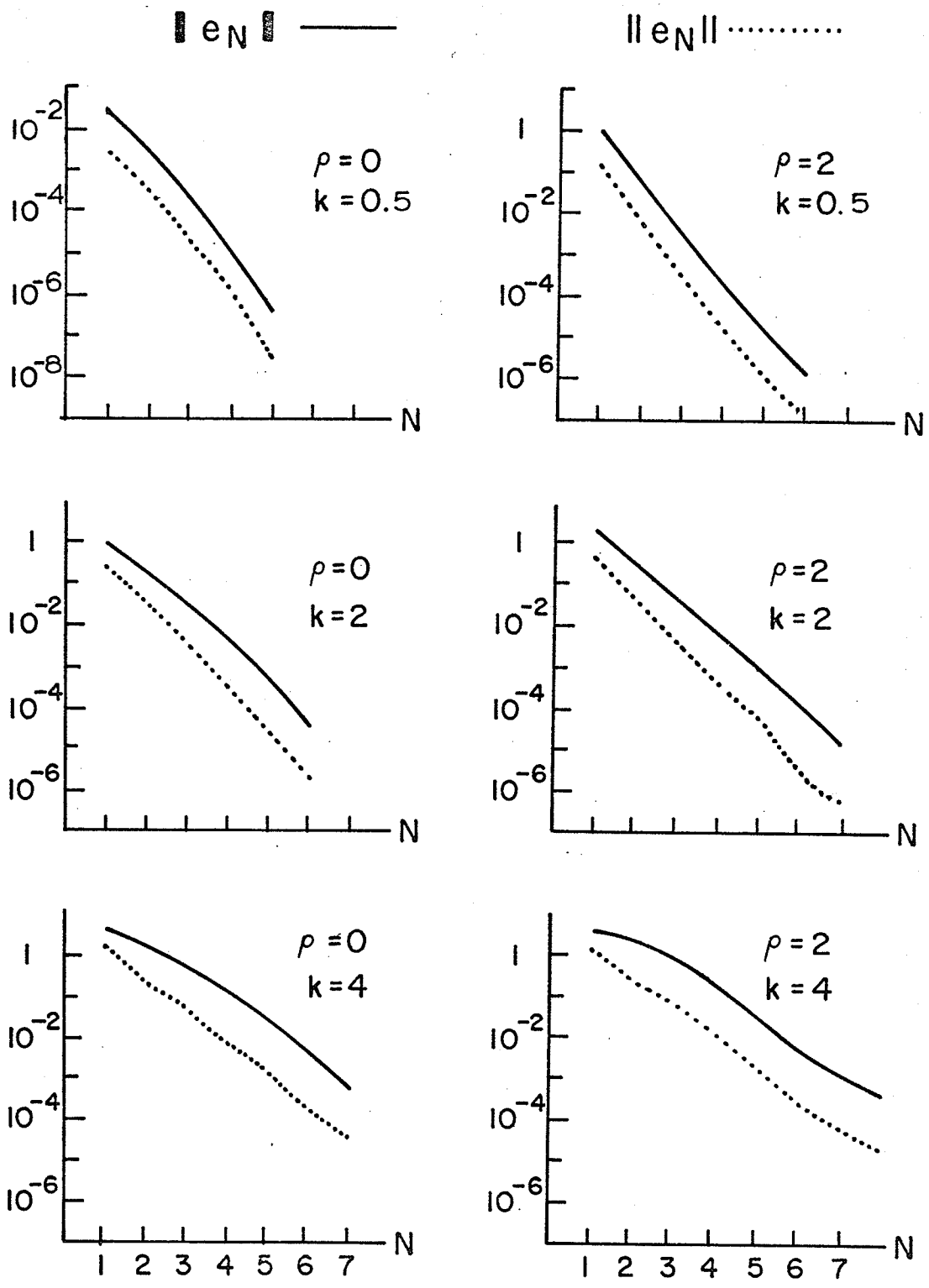


Fig. 2.3. The One-Dimensional PDE Problem: Energy Convergence Properties.

TABLE 2.1 THE ONE-DIMENSIONAL PDE PROBLEM, $\rho=0$

Prop. constant k	Number of terms N	Laplacian Energy U_L	Functional Value F	Error norm $\|\|e_N\|\| \times 10^{-5}$
0.0	1	1.000	1.000	0.000
0.5	5	1.002	0.915	0.006
1.0	5	1.027	0.642	0.033
1.5	5	1.183	0.106	0.066
2.0	6	1.961	-0.915	0.198
2.5	6	7.052	-3.347	0.286
3.0	6	215.440	-21.046	0.982
3.5	7	54.449	9.346	0.224
4.0	7	15.695	3.455	2.630
4.5	7	11.081	0.970	6.492
5.0	8	12.854	-1.479	15.415

TABLE 2.2 THE ONE-DIMENSIONAL PDE PROBLEM $\rho=2.0$

Prop. constant k	Number of terms N	Laplacian Energy U_L	Functional Value F	Error norm $\|\|e_N\|\| \times 10^{-5}$
0.0	2	1.333	-1.333	0.001
0.5	6	1.396	-1.469	0.010
1.0	6	1.645	-1.914	0.008
1.5	7	2.366	-2.808	0.038
2.0	7	4.763	-4.587	0.110
2.5	7	17.149	-9.063	0.261
3.0	7	448.210	-43.581	0.961
3.5	8	94.732	17.009	0.141
4.0	8	23.420	6.163	2.350
4.5	8	14.470	2.378	6.102
5.0	8	14.860	-0.064	13.915

cases $||e_N||$ and $\|e_N\|$ exhibit monotonic reduction as the polynomial order N is increased, with $\|e_N\| \geq \frac{\pi}{2}||e_N||$, as expected from Mikhlin's theorem.* It is noteworthy that the rate of convergence varies from case to case, in general growing worse as the propagation constant k is increased.

The summary data presented in Tables 2.1 and 2.2 are also interesting. For each value of k from 0 to 5, N was increased to give the smallest value of $||e_N||$, and for this solution the values of the functional F, and the value of the energy in the Laplacian U_L are tabulated. There are three points to be made. First, the minimum error $||e_N||$ increases, essentially monotonically, with k, due probably to the fact that higher order polynomials are required to approximate sin k x. Secondly, the values at k=3 (close to π, where the solution is not defined) indicate that although the solution is not as smooth as elsewhere (indicated by the higher values of U_L), the accuracy is reasonably good.

The third point to be made requires more discussion. Consider Table 2.1. There is no source term (ρ=0), and from (2.65) and (2.66), the value of the functional is equal to the energy in the Helmholtz operator, U_H. With k=2, U_H is negative. This means, quite simply, that the functional (2.60) is not necessarily minimized in this case. Accuracy, however, is not seriously affected.

This simple exercise has demonstrated how one may use the Rayleigh-Ritz method (or equivalently, the Galerkin method) for both static and propagation problems. Perhaps the most important conclusion to be drawn is that convergence curves are similar in all the cases presented, as seen by inspection of Fig. 2.3.

---

* The Laplacian is positive-bounded-below and $\pi/2$ is the smallest system eigenvalue. See Mikhlin [1,p.481].

The error norm $||e_N||$ decreases monotonically as N is increased, and is bounded above, for each value of N, by $\gamma \|e_N\|$, $\gamma>0$ (energy in the Laplacian). This shows the practical importance of Mikhlin's theorem: if an operator L contains a self-adjoint positive-bounded-below component L_0, the Galerkin solution converges in the mean as $N\to\infty$, and error is bounded by the energy norm of the error (energy in L_0) regardless of other properties of the linear, continuous operator L.

This exercise presents a case where the theorem applies, as the Dirichlet boundary conditions ensure that the Laplacian is positive-bounded-below. (See Mikhlin [1,p.119-120].) However, if $\int\phi\partial\phi/\partial n\, ds$ on the boundary does not vanish, positive-definiteness may not be demonstrable. Consider the radiation problem in free space, where potential behaves as e^{jkr}/r (1.8). The boundary integral at infinity may be formulated as

$$\lim_{r\to\infty} (jke^{2jkr} - e^{2jkr}/r).$$

For the static problem (k=0) both terms vanish, and the Laplacian is positive-definite. For the time-harmonic problem ($k\neq0$) the first term is not necessarily zero; it behaves sinusoidally as $r\to\infty$. The boundary integral may not vanish and the Laplacian may not be positive-definite. In other words, the solution of Lu=f may not be in the domain of L_0. In such cases the theorem does not apply.

Note that this arises when seeking ϕ throughout the infinite region. It is far better to seek a source distribution on a surface within the region in this case, and a method will be described in Section 6.1.

2.5 Finite Elements

In the example of the previous section a single trial function was defined over the whole region. It is often more convenient to define trial functions over subregions and to connect these representations together for the complete solution.

To begin with, consider two sub-regions, or finite elements, Ω_1 and Ω_2, as shown in Fig. 2.4, with $\hat{n}$ drawn from Ω_1 to Ω_2. For simplicity, a homogeneous Dirichlet condition is specified on S. The functional (2.48) may be computed over each element separately:

$$F(\phi) = \int_{\Omega_1} (\nabla\phi^* \cdot (\varepsilon_1 \nabla\phi) - W_1\phi^*\phi - \rho^*\phi - \phi^*\rho)\, d\Omega$$

$$+ \int_{\Omega_2} (\nabla\phi^* \cdot (\varepsilon_2 \nabla\phi) - W_2\phi^*\phi - \rho^*\phi - \phi^*\rho)\, d\Omega \quad . \tag{2.66}$$

Suppose that ϕ is continuous along S_I, the inter-element boundary. The first variation of F, then, is

$$\delta F(\phi) = 2\ \mathrm{Re}: \int_{\Omega_1} \delta\phi^*(-\nabla\cdot(\varepsilon_1\nabla\phi) - W_1\phi - \rho)\, d\Omega$$

$$+ \int_{\Omega_2} \delta\phi^*(-\nabla\cdot(\varepsilon_2\nabla\phi) - W_2\phi - \rho)\, d\Omega \tag{2.67}$$

$$+ \int_{S_I} \delta\phi^*(\hat{n}\cdot(\varepsilon_1\nabla\phi) - \hat{n}\cdot(\varepsilon_2\nabla\phi))\, ds \quad .$$

At the stationary point the PDE is satisfied in each element and the interface condition is satisfied naturally. When ε_1 and ε_2 are scalars, the more common interface condition results:

$$\varepsilon_1 \frac{\partial\phi}{\partial n}_- = \varepsilon_2 \frac{\partial\phi}{\partial n}_+ \tag{2.68}$$

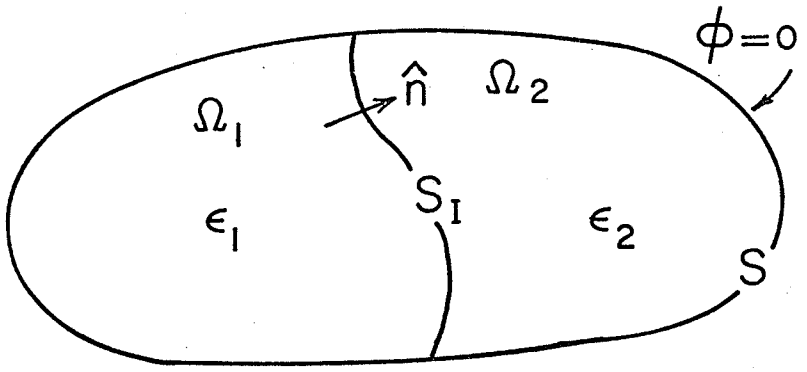


Fig. 2.4 Two Finite Elements

where n- is the normal approaching S_I, and n+ is the normal leaving S_I; both in the direction $\hat{n}$.

Each finite element may be treated independently from the others, and they are connected for problem solution by enforcing continuity at the inter-element boundaries. Any number of elements may be used.

It is convenient to have the trial (or approximating) function over an element as an interpolate of the potential values of certain points (nodes) in the element. This facilitates the enforcement of continuity between elements, and also the imposition of Dirichlet boundary conditions. The trial function over the element, then, is

$$\phi(\underline{r}) = \sum_{i=1}^{N} \underline{f}_i(\underline{r})\phi_i = \underline{f}^T\underline{\phi} \quad . \tag{2.69}$$

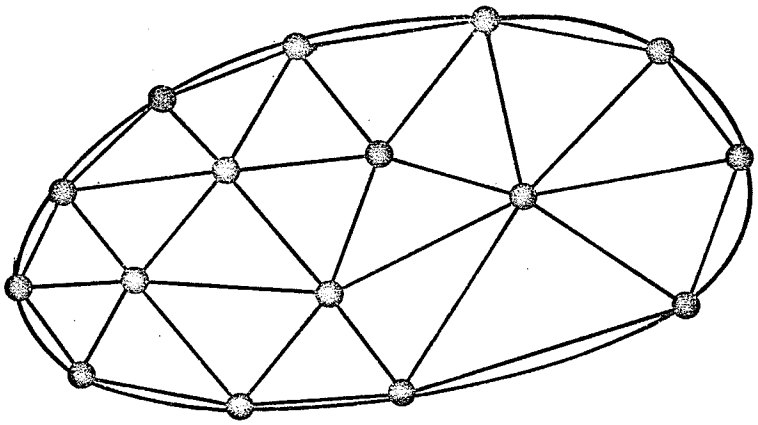


Fig. 2.5 Division of Ω into Triangular Elements

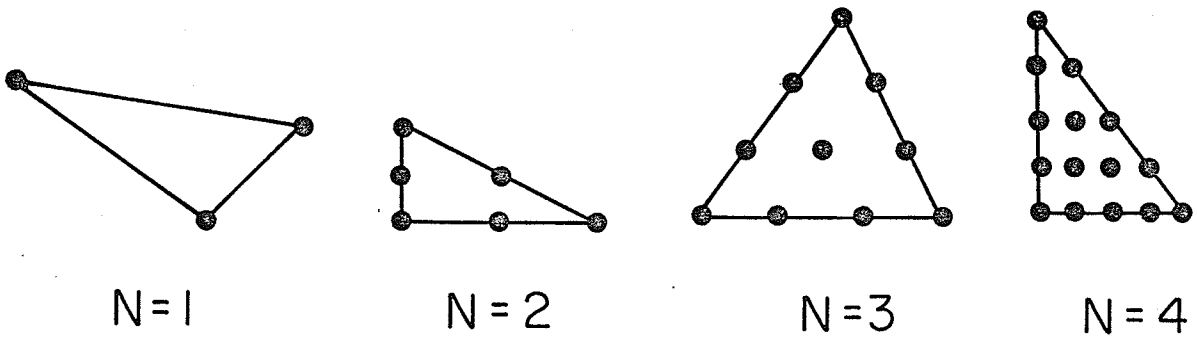


Fig. 2.6 Automatic Node Placement in Triangular Elements

Suppose that Ω is a region in the plane and is divided into triangular elements, as shown in Fig. 2.5. If a complete polynomial of degree M is used for approximation, there will be N=(M+1)(M+2)/2 nodes [9], which may be placed automatically in each triangular element as shown in Fig. 2.6. The M+1 nodes on each triangle side are equally spaced. This permits inter-element boundary continuity to be enforced easily: boundary nodes occupying the same physical location are assigned the same potential. A Dirichlet condition is readily enforced: boundary nodes are set to the specified values.

The interpolatory functions may be determined as follows. Let $\underline{g}$ be a vector of polynomials:

$$\underline{g}^T = (1 \; x \; y \; x^2 \; xy \; y^2 \; .... \; xy^{M-1} \; y^M) \quad . \tag{2.70}$$

Then

$$\phi(\underline{r}) = \underline{g}^T \, \underline{c} \tag{2.71}$$

where $\underline{c}$ is the vector of coefficients. Equation (2.71) may be evaluated at each node point in the element. The result is a vector of node potentials $\underline{\phi}$:

$$\underline{\phi} = \begin{bmatrix} \underline{g}_1^T \\ \vdots \\ \underline{g}_N^T \end{bmatrix} \underline{c} = [G] \, \underline{c} \quad . \tag{2.72}$$

Equations (2.71) and (2.72) may be combined to yield

$$\phi(\underline{r}) = \underline{g}^T \left[G^{-1}\right] \underline{\phi} \quad . \tag{2.73}$$

The vector interpolatory function $\underline{f}$ (2.69) is then given by

$$\underline{f}^T = \underline{g}^T \left[G^{-1}\right] \quad . \tag{2.74}$$

The matrix equation for the element may now be constructed:

$$\left[G^{-1}\right]^T \int_{\Omega_i} (\nabla \underline{g} \cdot \varepsilon \nabla \underline{g}^T - W \underline{g}\, \underline{g}^T)\, dx\, dy \quad \left[G^{-1}\right] \underline{\phi} = \left[G^{-1}\right]^T \int_{\Omega_i} \underline{g}\, \rho \quad dx\, dy . \tag{2.75}$$

The convention is to use real-valued interpolatory functions, whence the matrix computations may be performed using real arithmetic for real-valued ε, W. The integrations may be carried out analytically in the element for scalar, polynomial ε, W. Richards and Wexler [10] describe the algorithm used, with an extension to allow sides of the triangular elements to be curved. Wexler [11] describes an extension to anisotropic material for the static case (W=0).

The algorithm is inefficient due to the necessity of inverting the [G] matrix at equation (2.73) and due to the integrations in (2.75) which must be computed in the coordinates of each triangular element. A better method involves the use of Lagrange interpolation polynomials and isoparametric elements. Nevertheless, the individual elements, so treated, may be combined by enforcing continuity of ϕ between elements. There is a trade-off between order of approximation N and number of elements. The same number of unknown node potentials- in total -can result by using many low-order elements or fewer high-order elements. The former approach results in a less dense matrix. The matrix entry A_{ij} will vanish only when nodes i and j are not defined within the same element. Thus a single high-order element for the problem would produce a totally dense matrix. At the other extreme, use of first-order (N=1) elements to model the problem would result in a larger number of zero entries and a matrix structure similar to that of a finite-difference matrix. This will be expanded upon in Chapter VII.

2.6 Isoparametric Finite Elements

This technique, described in Zienkiewicz [12], permits all integrations to be performed numerically in one simple element (the local element), and the matrix inversion in equation (2.73) is eliminated.

Consider the local element shown in Fig. 2.7, and let the N node points (determined by M, the degree of the approximation [9]) be placed automatically, as shown previously in Fig. 2.6. Potential in the local element is then given by

$$\phi(\zeta,\eta) = \sum_{i=1}^{N} \alpha_i(\zeta,\eta)\, \phi_i = \underline{\alpha}^T \underline{\phi} \quad ; \tag{2.76}$$

when (ζ,η) is node point j all the α_i vanish, except for α_j, which is unity.

Consider now the finite element in the original (or global) coordinates shown in Fig. 2.8. The node points in the global element are described by two vectors, $\underline{x}$ and $\underline{y}$, the x and y coordinates of the nodes. Equation (2.76) may be applied as follows:

$$x(\zeta,\eta) = \underline{\alpha}^T \underline{x} \tag{2.77a}$$

$$y(\zeta,\eta) = \underline{\alpha}^T \underline{y} \tag{2.77b}$$

When (ζ,η) is node point j in the local element, (x,y) determined from (2.77), is node point j in the global element. Equations (2.77) thus represent the isoparametric transformation from local to global coordinates. The global node points may be placed arbitrarily, but it is convenient to place them in a regular triangle, as in the previous section. If a triangle is to have a curved side, the node points on

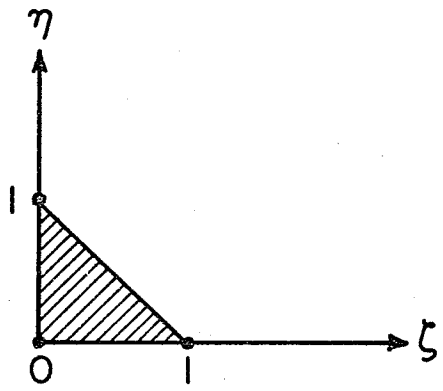


Fig. 2.7 The Local Element

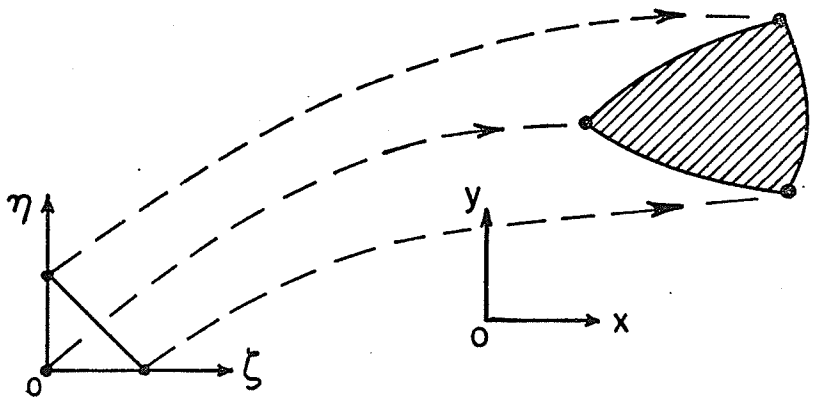


Fig. 2.8 The Global Element and the Mapping

that side may be shifted to the desired locations, as described in [13].

If any point (ζ,η) is specified in the local element, the corresponding global point (x,y) is known immediately from (2.77). The inverse transformation is not so simple. If the global element is a regular triangle, defined by its vertex coordinates, the point (ζ,η) corresponding to (x,y) may be found geometrically. (See Appendix E). If the global element is not regular it is convenient to use a Newton-Raphson method [14] to find the (ζ,η) corresponding to (x,y). A suitable starting point for the Newton-Raphson procedure is that point (ζ_o,η_o) determined by presuming the global element to be a regular triangle and applying the geometric procedure. This approach was suggested by M. Friedman and is documented in [13].

To compute the integrals required, one needs to know the gradient of potential in the local coordinates. This is accomplished by use of the Jacobian, here a 2 x 2 matrix whose elements are

$$\begin{aligned} J_{11} &= \frac{\partial x(\zeta,\eta)}{\partial \zeta} = \frac{\partial \underline{\alpha}^T}{\partial \zeta}\,\underline{x} \\ J_{12} &= \frac{\partial y(\zeta,\eta)}{\partial \zeta} = \frac{\partial \underline{\alpha}^T}{\partial \zeta}\,\underline{y} \\ J_{21} &= \frac{\partial x(\zeta,\eta)}{\partial \eta} = \frac{\partial \underline{\alpha}^T}{\partial \eta}\,\underline{x} \\ J_{22} &= \frac{\partial y(\zeta,\eta)}{\partial \eta} = \frac{\partial \underline{\alpha}^T}{\partial \eta}\,\underline{y} \quad . \end{aligned} \tag{2.78}$$

Then

$$\begin{aligned} \begin{pmatrix} \frac{\partial \phi(\zeta,\eta)}{\partial x} \\ \frac{\partial \phi(\zeta,\eta)}{\partial y} \end{pmatrix} &= \left[J(\zeta,\eta)\right]^{-1} \begin{pmatrix} \frac{\partial \phi(\zeta,\eta)}{\partial \zeta} \\ \frac{\partial \phi(\zeta,\eta)}{\partial \eta} \end{pmatrix} \\ &= \left[J\right]^{-1}\left[\nabla \underline{\alpha}\right]^T \underline{\phi} \end{aligned} \tag{2.79}$$

where

$$[\nabla\underline{\alpha}]^T = \begin{bmatrix} \dfrac{\partial \underline{\alpha}^T(\zeta,\eta)}{\partial\zeta} \\ \dfrac{\partial \underline{\alpha}^T(\zeta,\eta)}{\partial\eta} \end{bmatrix} . \tag{2.80}$$

The integration element is also transformed by $[J]$:

$$dx\ dy = \det[J]\ d\zeta\ d\eta . \tag{2.81}$$

The integration may now be computed entirely in the local coordinates, over the local element, by the use of Gaussian quadrature (described in Appendix D). At each integration point (ζ,η) values of ε, W, and ρ are obtained, using (2.77) to get the corresponding global point.* The element matrix equation, in the form of (2.34), is

$$\iint_{\substack{\text{local}\\ \text{element}}} ([\nabla\underline{\alpha}^T]^T\ [J^{-1}]^T \cdot \varepsilon [J^{-1}]\ [\nabla\underline{\alpha}^T] - W\ \underline{\alpha}\ \underline{\alpha}^T)\ \det\ [J] d\zeta d\eta\ \underline{\phi}$$
$$= \iint_{\substack{\text{local}\\ \text{element}}} \underline{\alpha}\ \rho\ \det\ [J] d\zeta d\eta \quad . \tag{2.82}$$

As the shape functions $\underline{\alpha}$ are real-valued, the matrix and vector contributions involve only real arithmetic for real-valued ε, W, and ρ.

Boundary integrals are computed in a similar fashion with special care being taken to obtain the elementary length ds in the local element; it is necessary to know which side of the local elememt is involved:

$$ds = \sqrt{\left(\frac{\partial x}{\partial\zeta}\right)^2 + \left(\frac{\partial y}{\partial\zeta}\right)^2}\, d\zeta \tag{2.83a}$$

$$ds = \sqrt{\left(\frac{\partial x}{\partial\eta}\right)^2 + \left(\frac{\partial y}{\partial\eta}\right)^2}\, d\eta \tag{2.83b}$$

or for the hypoteneuse

*W, ε, and ρ are required at points in the global element for computation of entries to the matrix equation, as in (2.75).

$$ds = \sqrt{\left(\frac{\partial x}{\partial \zeta} - \frac{\partial x}{\partial \eta}\right)^2 + \left(\frac{\partial y}{\partial \zeta} - \frac{\partial y}{\partial \eta}\right)^2}\ d\zeta \quad . \tag{2.83c}$$

The partial derivatives are all elements of the Jacobian (2.78).

The use of isoparametric elements increases the speed at which finite-element contributions to the system matrix are computed. Two-dimensional elements are sufficient for the problems considered here. Zienkiewicz [12] describes elements which may be used for three-dimensional problems.

2.7 Concluding Remarks

The most important result in this chapter is that Mikhlin's theorem permits energy convergent solutions of the deterministic Helmholtz equation to be obtained using finite elements in the Rayleigh-Ritz procedure. Energy convergence is highly desirable, for it guarantees that a more careful approximation (more terms in u_N) will result in a better approximation. The rate of convergence depends upon several things; in particular upon the trial functions used. It can rarely be predicted in advance.

The polynomic and isoparametric finite elements are general purpose elements, perfectly adequate for representation of scalar potential. The recent book edited by Whiteman [15] provides an excellent basis for general studies in finite-element methods, and among the contributors are Zlámal [16-17] and Oden [18]. Application of finite elements to the solution of Helmholtz problems has also been reported [19-21]. It has been demonstrated by Décreton [22] that convergence is improved when the general purpose elements are augmented by special functions to handle known non-polynomic behaviour.

The finite-element approach is flexible and convenient. The entire region may be taken as a single element, or it may be divided into several elements. One has freedom to choose the shape and size of the element for best representation of part of a region, and the isoparametric finite-element procedure provides an economical way of computing the integrals required. The analyst need be concerned only about one element at a time, for the elements are easily put together to produce the complete, continuous solution.

Mikhlin's theorem on the convergence of Galerkin's method is not to be underestimated. As the Laplacian, $-\nabla^2$, appears in the deterministic Helmholtz equation, convergence in energy is guaranteed, even if one cannot show the functional to be minimized at the solution point. One need seek only the stationary point. In this context, a paper by Hazel and Wexler [23] is very interesting. They use a functional which produces the Dirichlet condition naturally, after a suggestion in Hildebrand [24, p.200], but are not able to show that the functional is minimized. Nevertheless, the results they give show uniform convergence properties - to be expected by the theorem. They also show, however, that faster convergence to an exact solution is obtained by restricting the trial functions to satisfy the Dirichlet conditions, that is, $||e_N||$ converges toward zero more rapidly. For this reason Dirichlet problems here are handled by restricting the trial functions [13].

The tools used to solve the deterministic Helmholtz equation in a bounded region with Dirichlet or mixed boundary conditions have been presented. One divides the region into convenient finite elements and constructs the linear system matrix equation by including the contributions from each element in turn. Provided that the propagation function is not an eigenvalue of the system equation (in which case the matrix is singular and one seeks an eigensolution), convergence in the energy of the approximation in the Laplacian is obtained.

2.8 References

[1] S.G. Mikhlin, Variational Methods in Mathematical Physics. New York: MacMillan, 1964.

[2] M.J. Forray, Variational Calculus in Science and Engineering. New York: McGraw-Hill, 1968.

[3] A. Wexler, "Computation of electromagnetic fields" (invited paper). IEEE Trans. MTT-17, pp. 416-439, August, 1969.

[4] B.A. Finlayson, The Method of Weighted Residuals and Variational Principles. New York: Academic Press, 1972.

[5] S.H. Crandall, Engineering Analysis. New York: McGraw-Hill, 1956.

[6] R.F. Harrington, Field Computation by Moment Methods. New York: Macmillan, 1968.

[7] V.K. Thong, "Solutions for some waveguide discontinuities by the method of moments." IEEE Trans. MTT-20, pp. 416-418, June, 1972.

[8] Y. Chow and S. Wu, "A moment method with mixed linear basis functions for scatterings by waveguide junctions." IEEE Trans. MTT-21, pp. 333-340, May, 1973.

[9] P. Silvester, "High-order polynomial triangular finite-elements for potential problems." Int. Jnl. Eng. Sci., Vol. 7, pp. 849-861, 1969.

[10] D.J. Richards and A. Wexler, "Finite-element solutions within curved boundaries." IEEE Trans. MTT-20, pp. 650-657, Oct. 1972.

[11] A. Wexler, "Finite-element field analysis of an inhomogeneous, anisotropic, reluctance machine rotor." IEEE Trans. PAS-92, pp. 145-149, Jan. 1973.

[12] O.C. Zienkiewicz, The Finite-Element Method in Engineering Science. London: McGraw-Hill, 1971.

[13] M. Friedman, D.J. Richards and A. Wexler, University of Manitoba Finite-Element Program (MANFEP). Dept. of Elect. Eng., University of Manitoba, Winnipeg., Nov. 1973.

[14] E. Isaacson and H.B. Keller, Analysis of Numerical Methods. New York: J. Wiley, 1966.

[15] J.R. Whiteman (editor) The Mathematics of Finite Elements and Applications. London: Academic Press, 1973.

[16] M. Zlámal, "The finite-element method in domains with curved boundaries," Int. Jnl. for Num. Meth. Eng., Vol. 5, pp.367-373, 1973.

[17] M. Zlámal, "On the finite-element method." Num. Math., Vol. 12, pp.394-409, 1968.

[18] J.T. Oden, "A general theory of finite elements I and II." Int. Jnl. Num. Meth. Eng., Vol. 1, pp. 205-221, pp.247-260, 1969.

[19] P.L. Arlett, A.K. Bahrani and O.C. Zienkiewicz, "Applications of finite-elements to the solution of Helmholtz's equation." Proc. IEE, Vol. 115, pp.1762-1766, Dec. 1968.

[20] P. Silvester and A. Konrad, "Axisymmetric triangular finite-elements for the scalar Helmholtz equation." Int. Jnl. for Num. Meth. Eng., Vol. 5, pp. 481-497, 1973.

[21] P. Daly and J.D. Helps, "Exact finite-element solutions to the Helmholtz equation." Int. Jnl. for Num. Meth. Eng., Vol. 6, pp. 529-542, 1973.

[22] M.C. Décreton, "Treatment of singularities in the finite-element method." M.Sc. Thesis, University of Manitoba, Winnipeg, 1972.

[23] T.G. Hazel and A. Wexler, "Variational formulation of the Dirichlet boundary condition." IEEE Trans. MTT-20, pp.385-390. June, 1972.

[24] F.R. Hildebrand, Methods of Applied Mathematics. Englewood Cliffs, N.J.: Prentice-Hall, 1965.

CHAPTER III

A PARTIAL DIFFERENTIAL EQUATION WITH AN INTEGRAL CONSTRAINT

The techniques described in Chapter II permit solution of the Helmholtz and Poisson equations in bounded regions. But what can one do if the region is unbounded with localized regions of source and inhomogeneity? It is simply not feasible to divide the infinite region into finite elements.* One could construct an approximate bounded region problem by placing a boundary S sufficiently far from all sources that potential is approximately zero - and then proceed with the finite-element approach assuming a homogeneous Dirichlet condition on S. However, if the sources are separated by large distances the resulting bounded region may be inconveniently large. Another approach is desirable.

Suppose that each localized region of source or inhomogeneity is separated from the infinite region by placement of a boundary S_i around it. This results in small inhomogeneous regions called picture-frames, each of which may be treated easily by the PDE finite-element technique. The question then arises as to what boundary condition(s) to impose on S_i, the picture-frame boundaries.

Consider the infinite region Ω_e external to all the picture-frames. This region, by definition, is homogeneous and source-free. By equation (1.12) it is feasible to construct an integral equation to determine the source σ on the S_i if the potential is known on S_i. Solution of the IE problem would then produce the proper potential

* At least not by the methods outlined in Chapter II.

throughout Ω_e. Suppose that the normal derivative of potential on S_i can be related to σ on S_i. Then the IE (1.12) will involve only functions of potential (along with the Green's functions). One may now think of the IE as a function of potential on the S_i which ensures satisfaction of the prescribed problem in Ω_e. This function may be used as a boundary condition on the PDE picture-frame solutions, and will ensure that potential satisfies the prescribed equation in Ω_e. The IE expression may be thought of as a constraint upon the PDE picture-frame solutions. It is convenient in computation to use the Lagrange constraint method introduced in Section 2.4 as a means of enforcing the IE constraint.

This presentation deals with problems in the plane; that is, those problems having uniform cross-section. To be specific, analysis proceeds in the x-y plane at z=0, and it is assumed that the solution is the same for any z. See. Fig. 3.1.

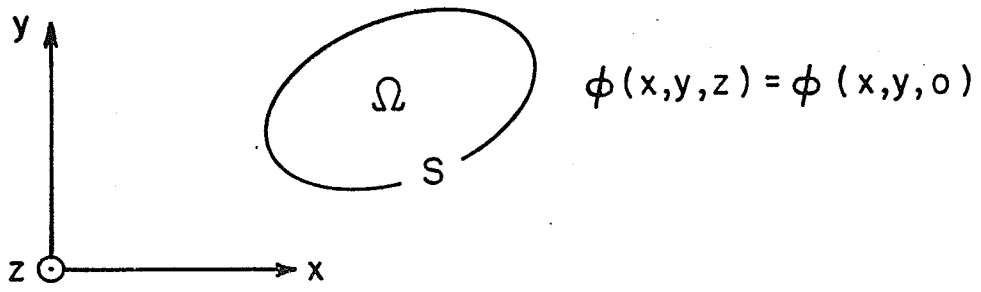


Fig. 3.1 Uniform Cross-Section Configuration

3.1 The PDE Matrix Equation

The finite-element method described in Chapter II is applied to each picture-frame region using the Helmholtz functional (2.48), with an assumed homogeneous Neumann condition on the S_i, the picture-frame boundaries. This results in the matrix equation

$$[S]\ \underline{\phi} = \underline{b} \quad . \tag{3.1}$$

The entries to $[S]$ and $\underline{b}$ are computed in the global coordinates by the methods described in Section 2.5. The vector $\underline{\phi}$ includes all nodes in all picture-frames. Solution of (3.1) produces a potential satisfying the Helmholtz equation in each picture-frame. One, of course, would expect difficulty if the propagation constant k coincided with an eigenvalue of the system (3.1), for $[S]$ would be singular. The homogeneous Neumann condition on the S_i is produced naturally, but (except fortuitously) potential in the region Ω_e external to the picture-frames does not satisfy the Helmholtz equation. It is the IE constraint which will provide the proper solution.

Silvester and Hsieh [1] and Williams and Cambrell [2] have also used finite-element methods to construct PDE solutions within picture-frame regions, to be inter-related by an IE constraint for solution of the unbounded static problem. Cermak and Silvester [3-5] and Sandy and Sage [6] have used finite-difference methods for construction of the PDE solutions within picture-frame regions, with relaxation techniques employed to obtain solution for the static case.

3.2 The Integral Constraint

Here the integral expression used to constrain the PDE picture-frame solutions is introduced. Consider the following two equations in a homogeneous bounded region Ω:

$$\nabla^2\phi(\underline{r}') + k^2\phi(\underline{r}') = 0$$

$$-\nabla^2 G(\underline{r}|\underline{r}') - k^2 G(\underline{r}|\underline{r}') = \delta(\underline{r}-\underline{r}') \quad . \tag{3.2}$$

Multiplying the first by G, the second by ϕ and adding them gives

$$G(\underline{r}|\underline{r}')\ \nabla^2\phi(\underline{r}') - \phi(\underline{r}')\nabla^2 G(\underline{r}|\underline{r}') = \phi(\underline{r}')\delta(\underline{r}-\underline{r}') \quad . \tag{3.3}$$

Integrating (3.3) over Ω and applying Green's theorem [7:p.12] with $\hat{n}$ pointing out of Ω, results in

$$\phi(\underline{r}) = \int_S \left(G(\underline{r}|\underline{r}')\ \frac{\partial\phi(\underline{r}')}{\partial n'} - \phi(\underline{r}')\ \frac{\partial G(\underline{r}|\underline{r}')}{\partial n'}\right) ds' \ ;\ \underline{r} \text{ in } \Omega \tag{3.4}$$

where $\underline{r}$ is not on S. If $\underline{r}$ is outside Ω, potential computed using (3.4) vanishes.

By use of (3.4) potential anywhere in Ω may be computed from an integration over the boundary, involving $\phi, \partial\phi/\partial n$ and $G, \partial G/\partial n$ on S.

For problems in the plane, the static free-space Green's function is given by [8, p.50]

$$G(\underline{r}|\underline{r}') = -\frac{1}{2\pi}\ \ln|\underline{r}-\underline{r}'| \tag{3.5a}$$

$$\frac{\partial G(\underline{r}|\underline{r}')}{\partial n} = \frac{-\cos\theta}{2\pi|\underline{r}-\underline{r}'|} \tag{3.5b}$$

and for the time-harmonic case ($k\neq 0$), by [8, p.55]

$$G(\underline{r}|\underline{r}') = -\frac{j}{4} H_0 (k|\underline{r}-\underline{r}'|) \tag{3.6a}$$

$$\frac{\partial G(\underline{r}|\underline{r}')}{\partial n} = \frac{jk}{4} H_1 (k|\underline{r}-\underline{r}'|) \cos\theta \ . \tag{3.6b}$$

The angle θ lies between the line from $\underline{r}'$ to $\underline{r}$ and $\hat{n}$, as shown in Fig.3.2.

Equation (3.4) may be thought of as potential produced by the superposition of two source distributions on S; a simple layer $\sigma = \varepsilon_o \partial\phi/\partial n$ which produces a continuous potential governed by [8, pp.48-49]

$$\phi(\underline{r}) = \frac{1}{\varepsilon_o} \int_S \sigma(\underline{r}') \, G(\underline{r}|\underline{r}') ds' \tag{3.7}$$

and a dipole layer of strength $\beta = -\varepsilon_o \phi$ (see Fig. 3.3), producing a potential which is discontinuous at S [8, pp.113-119]:

$$\phi(\underline{r}) = \frac{1}{\varepsilon_o} \int_S \beta(\underline{r}') \frac{\partial G(\underline{r}|\underline{r}')}{\partial n'} ds' \ ; \ \underline{r} \text{ not on S} \tag{3.8a}$$

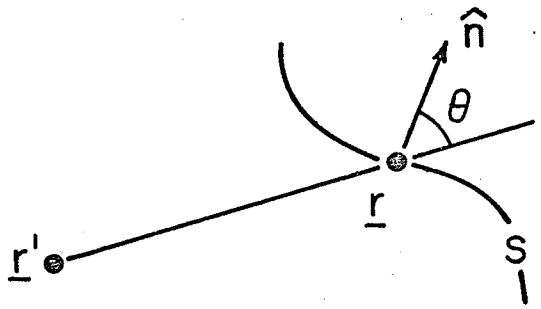
$$\phi_-(\underline{r}) = \frac{-\beta(\underline{r}')}{2\varepsilon_o} + \frac{1}{\varepsilon_o} \int_S \beta(\underline{r}') \frac{\partial G(\underline{r}|\underline{r}')}{\partial n'} ds' \ ; \ \underline{r} \text{ on S} \tag{3.8b}$$

$$\phi_+(\underline{r}) = \frac{\beta(\underline{r}')}{2\varepsilon_o} + \frac{1}{\varepsilon_o} \int_S \beta(\underline{r}') \frac{\partial G(\underline{r}|\underline{r}')}{\partial n'} ds' \ ; \ \underline{r} \text{ on S} \ . \tag{3.8c}$$

Equation (3.8b) may be used to show that when $\underline{r}$ is on S,

$$\phi(\underline{r}) = 2 \int_S \left(G(\underline{r}|\underline{r}') \frac{\partial\phi(\underline{r}')}{\partial n'} - \phi(\underline{r}') \frac{\partial G(\underline{r}|\underline{r}')}{\partial n'}\right) ds' ; \underline{r} \text{ on S} . \tag{3.9}$$

The basic equations to be used for construction of the integral constraint are now available. Equation (3.9) relates ϕ and $\partial\phi/\partial n$ on S in such a way that ϕ satisfies the homogeneous Helmholtz equation everywhere in the region Ω. Once ϕ and $\partial\phi/\partial n$ are known, equation (3.4)

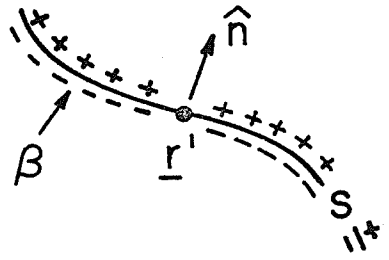


$$d = |\underline{r} - \underline{r}'|$$

$$G = -\frac{1}{2\pi} \ln(d)$$

$$\frac{\partial G}{\partial n} = \frac{-1}{2\pi d} \cos\theta$$

Fig. 3.2 Static 2-D Green's Function: Geometry



$$\phi(\underline{r}) = \frac{1}{\epsilon_0} \int_S \beta(\underline{r}') \frac{\partial G(\underline{r}|\underline{r}')}{\partial n'} ds'$$

Fig. 3.3 A Dipole Layer Source Distribution

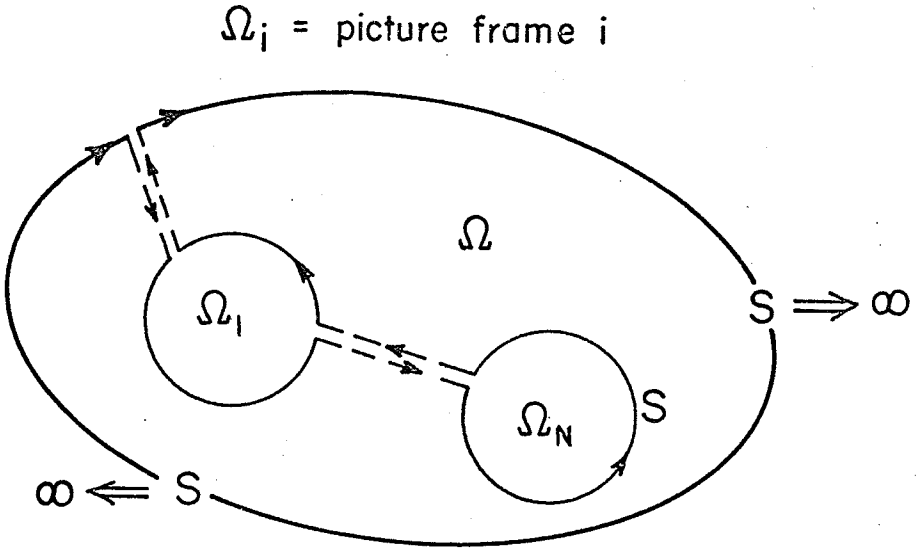


Fig. 3.4 The Region Outside the Picture-Frames

may be used to compute ϕ at any point not on S. It is now necessary to relate the region Ω to the exterior region Ω_e of the picture-frame problem, and S to the picture-frame boundaries, the S_i.

Consider Fig. 3.4, where S has been drawn with cuts between regions Ω_i, i=1 .. N, and with a cut to a portion which will be taken to infinity. The regions Ω_i may be associated with the picture-frames, and when the outer portion of S is taken to infinity, Ω will be associated with Ω_e, the infinite region exterior to the picture-frames. Integration of (3.9) along S so defined reveals the required integral constraint. There is no net contribution to the integral of (3.9) from the cuts [9, p.45], as the contribution resulting from the integral in one direction along the cut is exactly cancelled by the contribution from the integral in the opposite direction. Since G

satisfies the radiation condition at infinity [10], the contribution from the portion of S at infinity vanishes [8, pp.174-176]. One is left with contributions solely from the portions of S enclosing the Ω_i.

Therefore one may construct the integral constraint simply by integrating (3.9) over the picture-frame boundaries, the S_i, with $\hat{n}$ pointing into each picture-frame. Williams and Cambrell [2] have used (3.9) as the IE expression for the static case in this manner, and Burton and Miller [10] assert that the relation is valid for any value of the propagation constant k. Thus time-harmonic problems may be considered. As well, Greenspan and Werner [11] have examined the Helmholtz problem with the conclusion that the picture-frame method, using (3.9) as the constraint, produces a unique solution.

Silvester and Hsieh [1] use a simple layer distribution alone for the constraint expression, in the form

$$\phi(\underline{r}) = \int_S \frac{\partial \phi(\underline{r}')}{\partial n'} \; G(\underline{r}|\underline{r}') \; ds' \tag{3.10}$$

and use a Galerkin procedure to produce a matrix constraint equation, relating ϕ to $\partial\phi/\partial n$ on S. From equation (3.7), the normal derivatives of the simple layer distribution are [8, p.119]

$$\frac{\partial \phi(\underline{r})}{\partial n_-} = \frac{\sigma(\underline{r})}{2\varepsilon_o} + \frac{1}{\varepsilon_o} \int_S \sigma(\underline{r}') \; \frac{\partial G(\underline{r}|\underline{r}')}{\partial n} \; ds' \tag{3.11a}$$

and

$$\frac{\partial \phi(\underline{r}')}{\partial n_+} = -\frac{\sigma(\underline{r})}{2\varepsilon_o} + \frac{1}{\varepsilon_o} \int_S \sigma(\underline{r}') \; \frac{\partial G(\underline{r}|\underline{r}')}{\partial n} \; ds' \tag{3.11b}$$

or

$$\frac{\partial\phi(\underline{r})}{\partial n_-} - \frac{\partial\phi(\underline{r})}{\partial n_+} = \frac{\sigma(\underline{r})}{\varepsilon_o} \quad . \tag{3.11c}$$

The relation used by Silvester and Hsieh [1] is incomplete unless $\partial\phi/\partial n_+ = 0$. Williams and Cambrell [2] also note this and show by example that (3.9) produces better results, using a point-matching technique to produce the matrix equation relating ϕ to $\partial\phi/\partial n$ on S.

Cermak and Silvester [3-5] define an auxiliary contour S_c, lying within the picture-frame, as shown in Fig. 3.5. They define a single layer distribution σ on S_c, identified by the coefficients σ_i, i=1 ... M, at the M finite-difference nodes on S_c. They then compute potential at these M points on S_c:

$$\underline{\phi}_{S_c} = \left[M_{S_c}\right] \underline{\sigma} \tag{3.12}$$

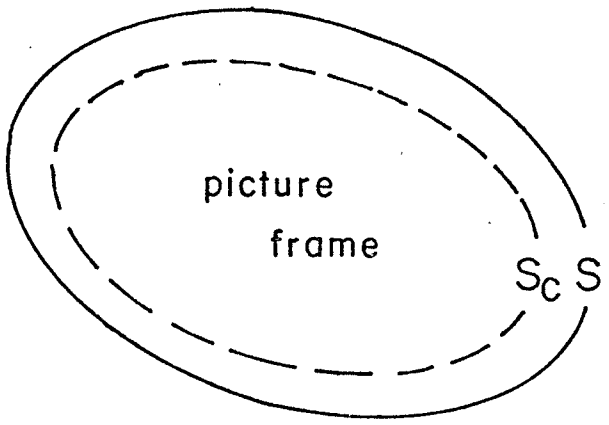


Fig. 3.5 The Auxiliary Contour S_c in the Picture-Frames

and also at M suitably chosen nodes on S:

$$\underline{\phi}_S = [M_S]\ \underline{\sigma} \quad . \tag{3.13}$$

By eliminating $\underline{\sigma}$, they obtain

$$\underline{\phi}_S = [M_S]\ [M_{S_c}]^{-1}\ \underline{\phi}_{S_c} \tag{3.14}$$

and this "shift matrix" is used as the constraint on ϕ in the picture-frame(s). The size of the matrices required in their finite-difference technique can demand that relaxation be used to obtain the shift matrix. Since they use a single layer alone, resonances could occur at picture-frame eigenvalues [10] in the time-harmonic case.

An alternative approach is to use this auxiliary contour with equation (3.4):

$$\phi(\underline{r}) = \int_{S_c} \left(G(\underline{r}|\underline{r}')\ \frac{\partial \phi(\underline{r}')}{\partial n'} - \phi(\underline{r}')\ \frac{\partial G(\underline{r}|\underline{r}')}{\partial n'} \right) ds' ;\ \underline{r} \text{ on } S \tag{3.15}$$

where S_c has a portion within each picture-frame, as shown in Fig. 3.6. The potential and its derivatives may all be obtained in terms of the PDE approximation, so that there are no new variables to be introduced. Furthermore, if S_c and S are sufficiently far apart - at least 0.25 - as shown in Fig. 3.7, the integrands are regular and numerical integration may be used directly. This means, essentially, that Green's function singularities may be ignored.

To facilitate computation, S_c is divided into equal length line segments. Each segment is defined by its end-point coordinates, and with $\hat{n}$ pointing to the left, into the picture-frame, as shown in Fig. 3.8:

$$\hat{n} = \frac{(y_a - y_b)\ \hat{i} - (x_a - x_b)\ \hat{j}}{\sqrt{(x_a - x_b)^2 + (y_a - y_b)^2}} = n_x \hat{i} + n_y \hat{j} \tag{3.16}$$

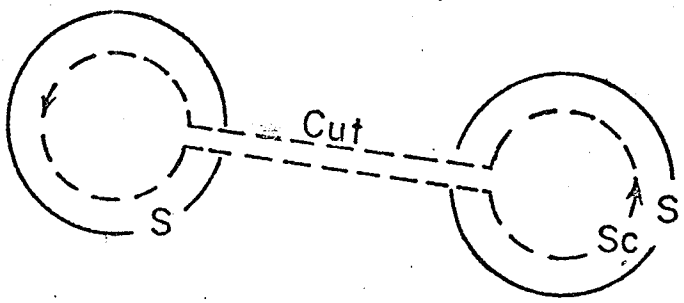


Fig. 3.6 The Auxiliary Contour and Multiple Picture-Frames

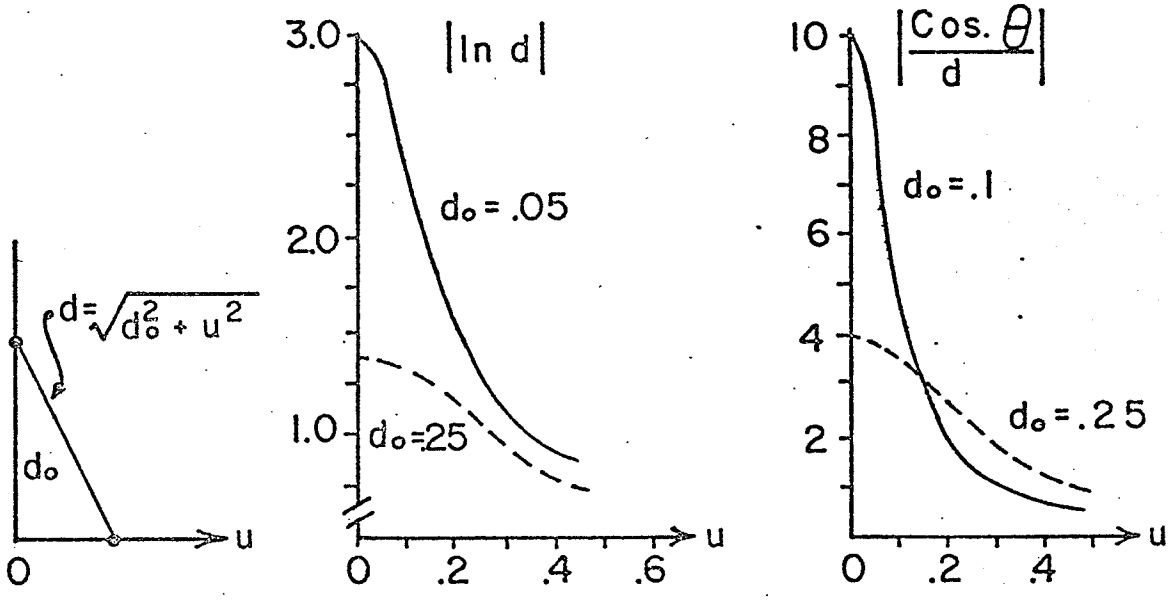


Fig. 3.7 The Behaviour of ln(d), cos θ/d with Separation d_o Between S_c and the Picture-Frame Boundary ($d=|\underline{r}-\underline{r}'|$; when u=0, $\underline{r}'$ is point on S_c closest to $\underline{r}$ on S; u is distance on a linear S_c from that point).

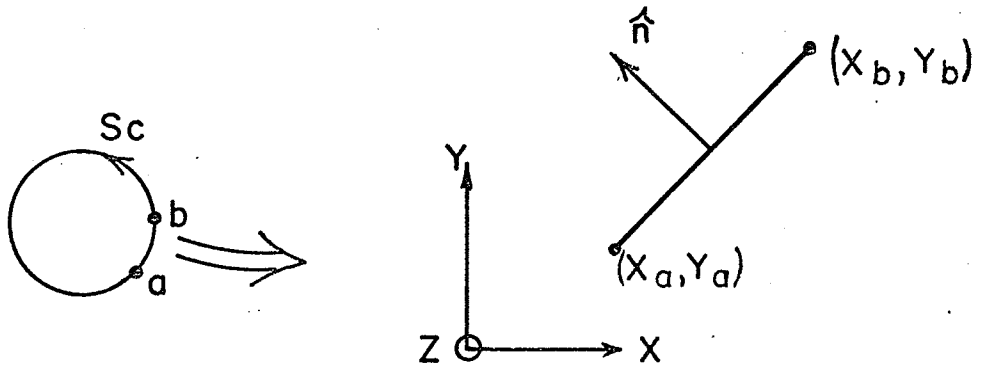


Fig. 3.8 Orientation of a Line Segment of S_c

Simple trapezoidal integration [12, p.197] is used over the line segment. The line from $\underline{r}$ to integration point $\underline{r}'$ has the direction

$$\hat{m} = \frac{(x'-x)\,\hat{i} + (y'-y)\,\hat{j}}{|\underline{r}-\underline{r}'|} = m_x\,\hat{i} + m_y\,\hat{j} \quad . \tag{3.17}$$

The cosine required for $\partial G/\partial n$ is simply obtained as

$$\cos\theta = n_x\, m_x + n_y\, m_y \quad . \tag{3.18}$$

All contributions for the Green's functions (3.5, 3.6) may now be evaluated.

Potential is given anywhere in a PDE finite-element by

$$\phi(\underline{r}') = \underline{f}^T\,\underline{\phi} \tag{3.19}$$

where $\underline{f}$ and $\underline{\phi}$ are the finite-element functions and coefficients, respectively, as described in Chapter II. Then

$$\frac{\partial \phi(\underline{r}')}{\partial n'} = \left(\frac{\partial \underline{f}^T}{\partial x'} n_x + \frac{\partial \underline{f}^T}{\partial y'} n_y\right) \underline{\phi} \tag{3.20}$$

Now, for any observation point $\underline{r}$ on S (or in the external region) the integration over S_c (3.15) may be computed in the form:

$$\phi(\underline{r}) = \underline{g}(\underline{r})^T \underline{\phi} \tag{3.21a}$$

where

$$\underline{g}(\underline{r})^T = \int_{S_c} \left(G(\underline{r}|\underline{r}') \left(\frac{\partial \underline{f}^T}{\partial x'} n_x + \frac{\partial \underline{f}^T}{\partial y'} n_y\right) - \underline{f}^T p(\underline{r}|\underline{r}') (n_x m_x + n_y m_y) \right) ds' \tag{3.21b}$$

and in which p, for the static case (3.5b) is given by

$$p(\underline{r}|\underline{r}') = \frac{-1}{2\pi|\underline{r}-\underline{r}'|} \tag{3.21c}$$

or, for the time-harmonic case (3.6b) by

$$p(\underline{r}|\underline{r}') = \frac{jk}{4} H_1 (k|\underline{r}-\underline{r}'|) \quad . \tag{3.21d}$$

In this expression (3.21) $\underline{\phi}$ is the vector of all node potentials in all the finite-elements. Note that if the point $\underline{r}$ is not in the region external to S_c, $\phi(\underline{r})$ in (3.21a) should vanish. (See the discussion at (3.4).) This is a useful check on the trapezoidal integration procedure, and the integration step may be adjusted to improve accuracy.

The integral constraint is constructed by evaluating (3.21a) at all the node points on the picture-frame boundary. If $\underline{\phi}_B$ is the vector of these node potentials, the following matrix equation results

$$\underline{\phi}_B = [M] \underline{\phi} \tag{3.22}$$

If $\underline{\phi}_I$ is the vector of internal node potentials (that is, $\underline{\phi}^T=(\underline{\phi}_I{}^T \underline{\phi}_B{}^T)$), (3.22) may be rewritten in the form

$$\underline{\phi}_B - [N]\,\underline{\phi}_I = 0 \tag{3.23}$$

The matrix $[N]$ is not square unless the number of nodes on S equals the number of those not on S. The Rayleigh-Ritz matrix equation for PDE solution (3.1) may be written as

$$\begin{bmatrix} S_{II} & S_{IB} \\ S_{BI} & S_{BB} \end{bmatrix} \begin{pmatrix} \underline{\phi}_I \\ \underline{\phi}_B \end{pmatrix} = \begin{pmatrix} \underline{b}_I \\ \underline{b}_B \end{pmatrix} \tag{3.24}$$

and (3.24) is to be solved subject to (3.23). This is accomplished by using a Lagrange constraint [13, pp.12-17]

$$\begin{bmatrix} S_{II} & S_{IB} & -N^{*T} \\ S_{BI} & S_{BB} & I \\ -N & I & 0 \end{bmatrix} \begin{pmatrix} \underline{\phi}_I \\ \underline{\phi}_B \\ \underline{\lambda} \end{pmatrix} = \begin{pmatrix} \underline{b}_I \\ \underline{b}_B \\ \underline{0} \end{pmatrix} . \tag{3.25}$$

Equation (3.25) may be solved directly, or $\underline{\phi}_B$ and $\underline{\lambda}$ may be eliminated algebraically to yield

$$\left[S_{II} + S_{IB}N + N^{*T} S_{BI} + N^{*T} S_{BB} N\right]\ \underline{\phi}_I = \underline{b}_I + \left[N^{*T}\right] \underline{b}_B \tag{3.26}$$

Here only the internal node potentials are sought, with the boundary node potentials recoverable from (3.23) after solution of (3.26). It should be remembered that the Green's function is complex for $k \neq 0$ (See equation (3.6)), so that the matrix $[N]$ will, in general, be

complex-valued.

There are no free sources in any of the examples, so that the PDE is always homogeneous.

3.3 TheStrip Capacitor: Electrostatic Solution

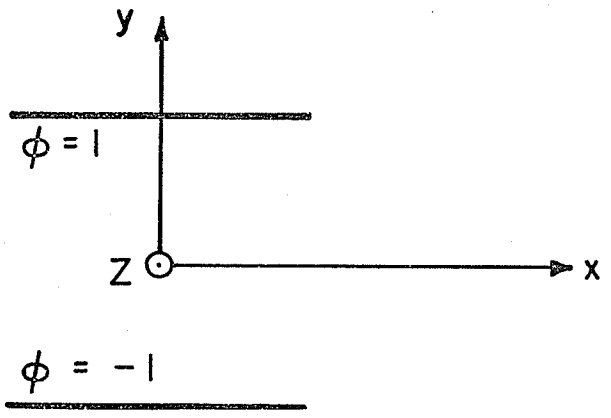
Consider the geometry shown in Fig. 3.9a, representing the cross-section of an infinitely-extending parallel-strip capacitor. The symmetries of the problem allow solution to be sought only in the positive quadrant, and the reduced picture-frame with the contour S_c are shown in Fig. 3.9b. With the polynomial of degree N=4 there are 65 internal node potentials to be obtained. The contour S_c is constructed with 10 line segments, and 4 point trapezoidal integration is used to compute terms in (3.15)*.

The symmetry of the geometry about the x=0 plane is handled by imposing the homogeneous Neumann condition at x=0. This condition is natural to the PDE formulation, and need not be enforced. The anti-symmetry about the y=0 plane is handled by enforcing the homogeneous Dirichlet condition at y=0. The integration over S_c is done in four parts, one for each quadrant, by the technique described in Appendix B.

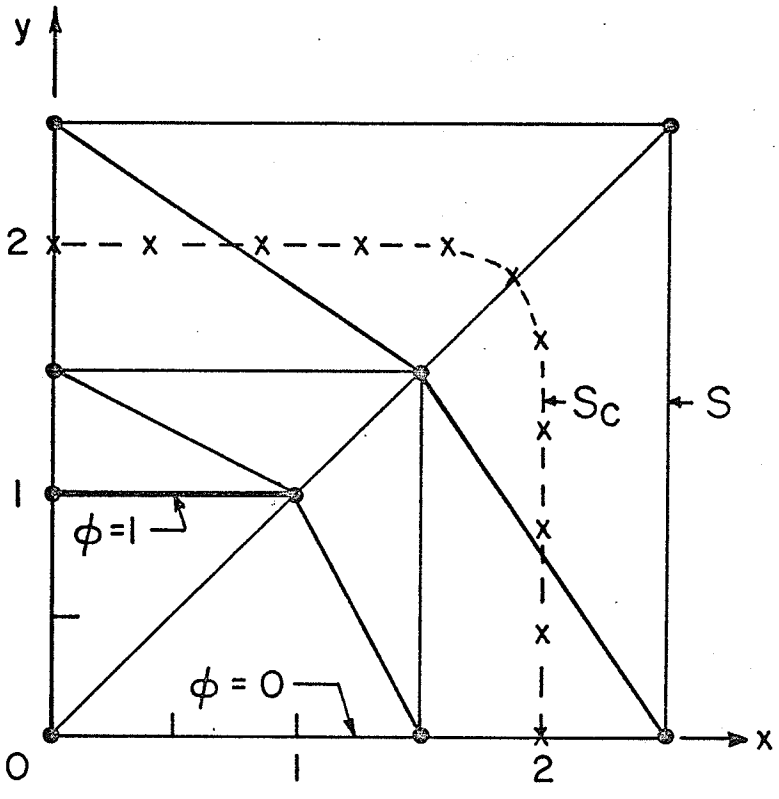
For the Dirichlet potential ϕ=1 specified on the strip, the resulting equipotentials are shown in Fig. 310. For comparison, an "exact" equipotential plot, accurate to within one per cent, is also shown. It may be observed that the greatest error occurs in the vicinity of the strip edge. This error is due to the known singular behaviour of the derivatives not being well modelled by the polynomial approximations. (This problem will be discussed further in Chapter IV.) Décreton [14] has shown that the introduction of special functions to handle such behaviour can greatly improve the approximation.

The problem was investigated early by Palmer [15] who used Schwartz-Christoffel techniques to obtain an analytic solution, and is a special

* The contour S_c is chosen to ensure separation from S ($\geq$0.25, as in Fig 3.7). The "rounding" of S_c near (2,2) is not significant; the contour may be constructed of any straight line segments.



(a)



(b)

Fig. 3.9 The Strip Capacitor
a) Full View of Capacitor in Cross-Section
b) The Reduced Picture-Frame

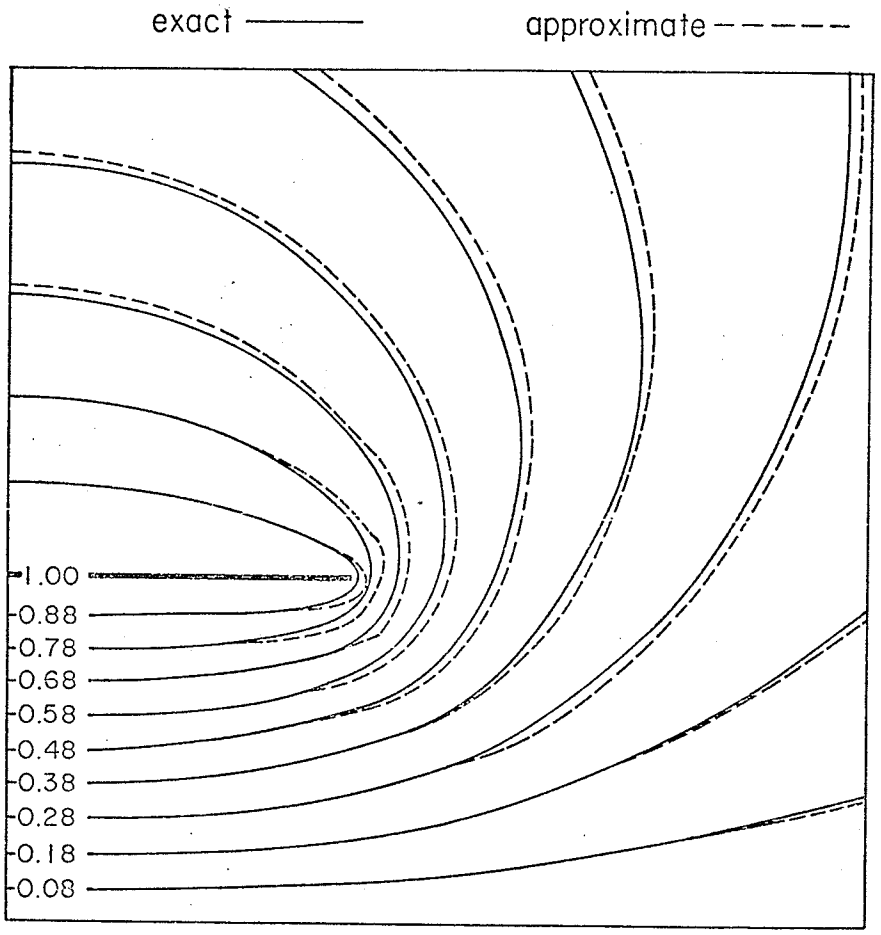


Fig. 3.10 The Strip Capacitor: Static Equipotentials within the Reduced Picture-Frame

case of problems investigated by Wheeler [16] and Bryant and Weiss [17]. The capacitance of this strip capacitor is 18.7246 pf/m*. This value was computed by use of Palmer's algorithm. Capacitance may also be computed from the electrostatic energy U [18, p.36]:

$$C = \frac{2U}{V^2} = \frac{U}{2} \tag{3.27}$$

where V is the potential difference between the strips. The electrostatic energy is also given by the integral [18, p.109]

$$U = \frac{\varepsilon_o}{2} \int_{\text{all space}} \nabla\phi \cdot \nabla\phi \, d\Omega \quad . \tag{3.28}$$

Applying Green's first identity to the region external to S, with $\hat{n}$ pointing out of Ω, and using the symmetries, allow this to be written as

$$U = 2\varepsilon_o \int_{\Omega} \nabla\phi \cdot \nabla\phi \, d\Omega + 2\varepsilon_o \int_S \phi \frac{\partial\phi}{\partial n} \, ds \tag{3.29}$$

Substitution of (3.29) into (3.27) yields the appropriate equation for capacitance:

$$C = \varepsilon_o \int_{\Omega} \nabla\phi \cdot \nabla\phi \, d\Omega + \varepsilon_o \int_S \phi \frac{\partial\phi}{\partial n} \, ds \tag{3.30}$$

where Ω and S are the reduced picture-frame and its external boundary, respectively. The first integral in (3.30) is computed by integrating the polynomial approximation exactly, and the second, by 21 point trapezoidal integration. The result is 18.9 pf/m, which is in error by about one per cent.

---

* The reference for this was given as [17] in [19]. This is incorrect, as the value was taken from tables in a preprint of [17] which did not appear in the published work.

3.4 The Capacitor with a Dielectric Slab: Electrostatic Solution

Suppose that a dielectric slab (ε_r=9) is placed between the strips as shown in Fig. 3.11a. The symmetries allow the solution to be obtained in the positive quadrant and the reduced picture-frame is shown in Fig. 3.11b. Here, the controur S_c is composed of 13 segments, and again, 4 point trapezoidal integration is employed on each segment.

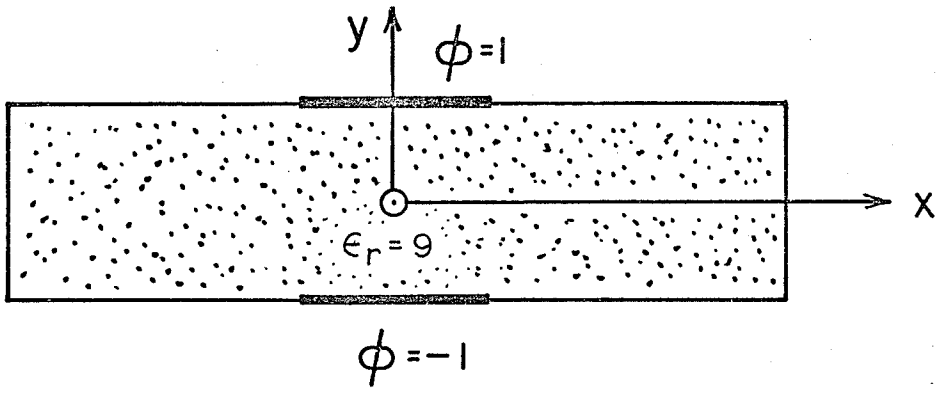
The problem is similar to certain open microstrip problems in which the dielectric is taken to be infinitely wide, which permits use of integral transform techniques [20] and integral equation techniques, such as that of Bryant and Weiss [17]. For a slab of finite width those methods are not directly applicable, although the latter technique could be suitably modified. This will be pursued further in Chapter V.

With N=3 there are 71 node potentials to be determined for the present problem. The resulting equipotentials are shown in Fig. 3.12. The strip edge singularity effect is clearly observed in the perturbed contours in the finite-elements shown by the dashed lines.

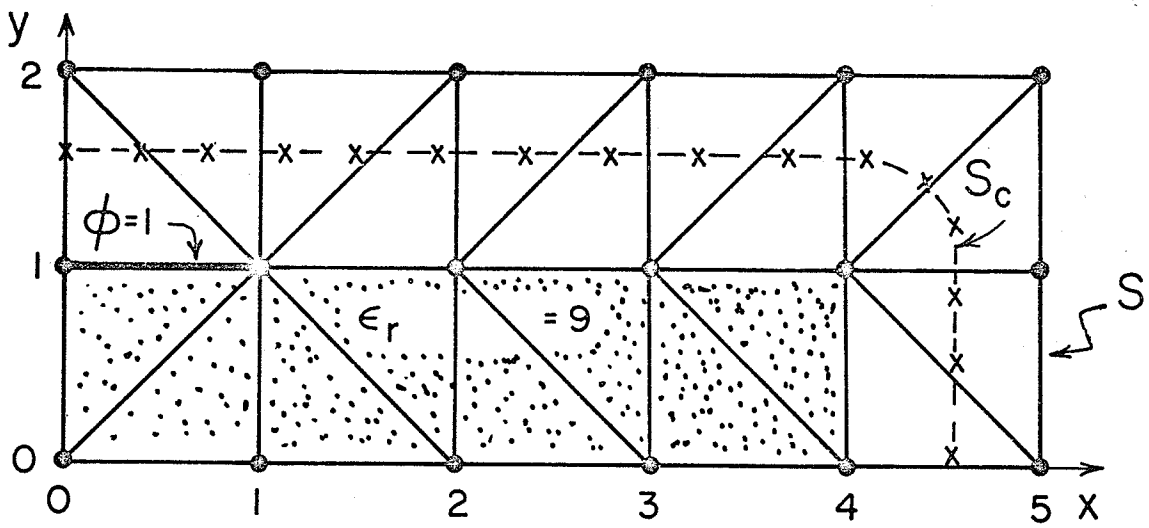
The effective dielectric constant is obtained as the ratio of capacitance with the dielectric slab to capacitance without the slab. Using field energies to compute capacitance, as in the previous example, the effective dielectric constant is computed to be 6.38. Bryant and Weiss [17] report a value of 6.47* for an infinitely wide slab (ε_r=9).

---

* This reference, as given in [19], is not completely correct. The value referred to appeared in a preprint of [17], not in the final publication.



(a)



(b)

Fig. 3.11 The Strip Capacitor with a Dielectric Slab (ε_r=9)
a) Full View of Problem in Cross-Section
b) The Reduced Picture-Frame

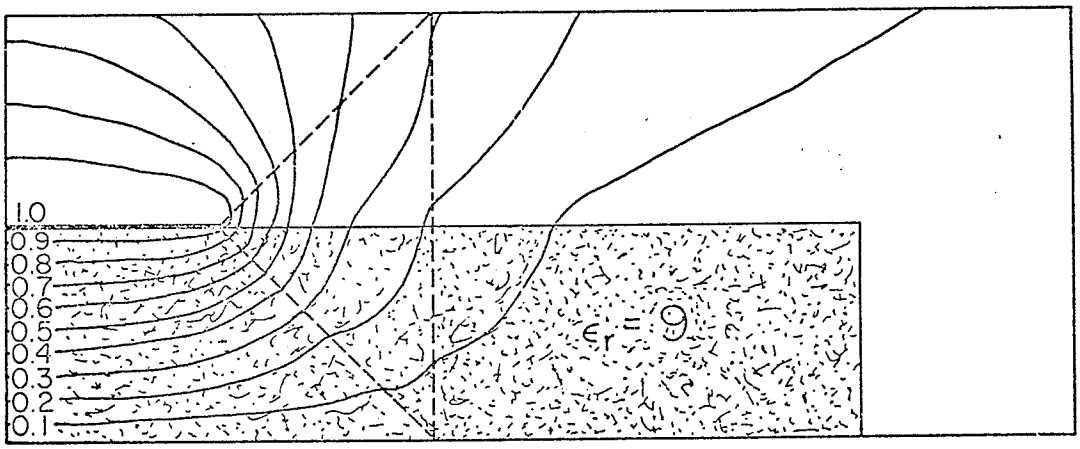


Fig. 3.12 The Strip Capacitor with a Dielectric Slab (ε_r=9): Static Equipotentials

One would expect the value to be higher for the latter case, due to the much larger dielectric. However, as seen in Fig. 3.12, the field in the dielectric becomes small as distance from the conductor is increased, and this implies but a small contribution to energy from the field in the dielectric at large distances from the conductor.

3.5 The Capacitor: Time-Harmonic Solution

Here the electric vector potential described by Harrington [21, pp.129-131] is used to obtain solution to the capacitor problem of Section 3.3. There are several algorithmic modifications required. Equipotentials here are of H_z and are tangent to the direction of the electric field. In the static case equipotentials are orthogonal to the electrostatic equipotentials obtained previously. The symmetries must be interchanged in constructing solutions using the reduced picture-frame shown in Fig. 3.13. The triangular finite-elements are placed as in the example of Section 3.3, but here S_c is placed somewhat differently; again, however, as ten equal length line segments. The source is conceived as a line dipole connected between the strips at x=0, which produces the Neumann boundary condition shown in Fig. 3.13.

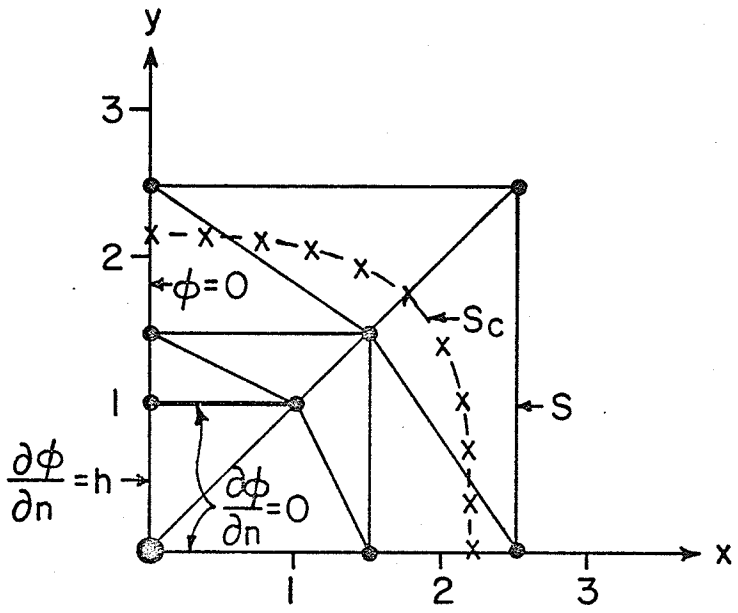


Fig. 3.13 The Reduced Picture-Frame for the Strip Capacitor with the Electric Vector Potential

The results were obtained with N=2 (55 node potentials) and the static equipotentials of H_z are shown in Fig. 3.14. These are seen to be orthogonal to the electrostatic equipotentials obtained previously (illustrated in Fig. 3.10). The equipotentials of the real part of H_z, obtained with the propagation constant k=4 (representing a frequency of 19 GHz with dimensions in centimeters), are shown in Fig. 3.15. These may be compared with those of a dipole source radiating into free space, as shown by Harrington [21, p.277].

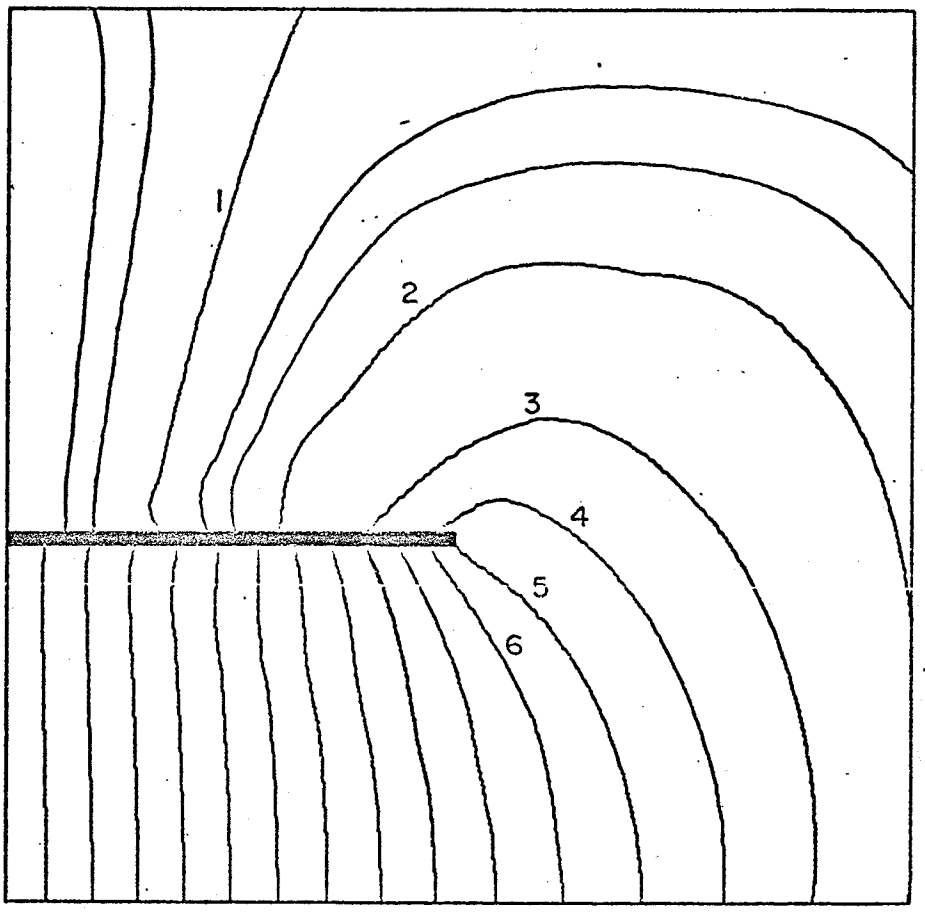


Fig. 3.14 The Strip Capacitor; Electric Vector Potential: Contours of H_z, Static Case

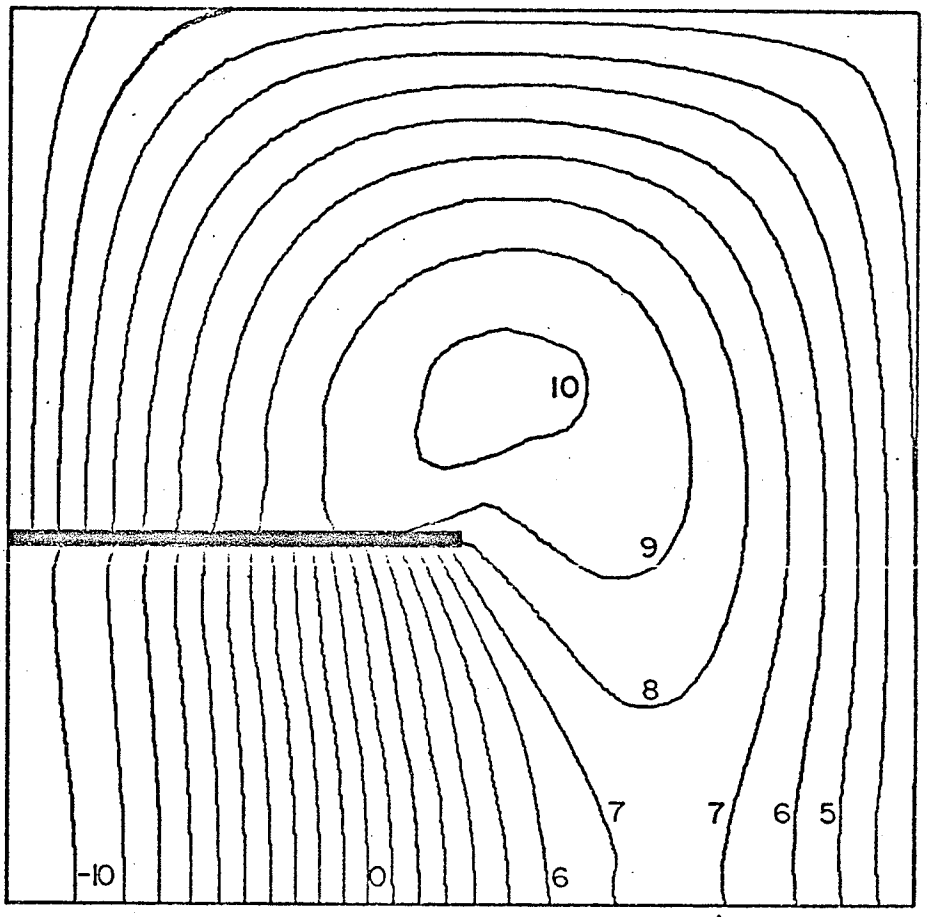


Fig. 3.15 The Strip Capacitor; Electric Vector Potentials: Contours of H_z at 19 GHz.

3.6 An Antenna-Type Problem

Consider the geometry shown in Fig. 3.16. This is a cross-section of a thick "antenna" infinite in the z direction with two dielectric obstacles. The dielectric constant is chosen as ε_r=100, to ensure that the dielectric perturbs the free-space solution. The electric vector potential is used, and the symmetries may be used to construct the reduced picture-frames shown in Fig. 3.17. There are two separate picture-frames in Fig. 3.17, and the triangular elements and the contours S_c are placed as shown. The source is a line dipole connected between the antenna plates at x=0, resulting in the Neumann boundary condition shown.

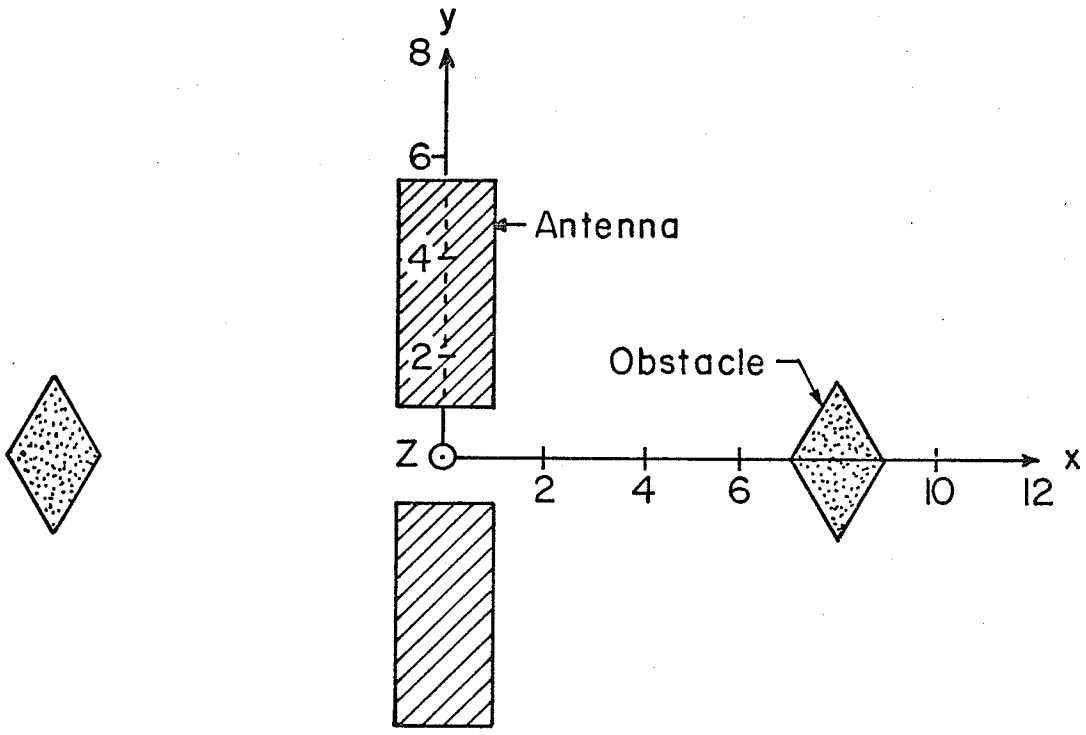


Fig. 3.16 An Antenna-Type Problem: Full View

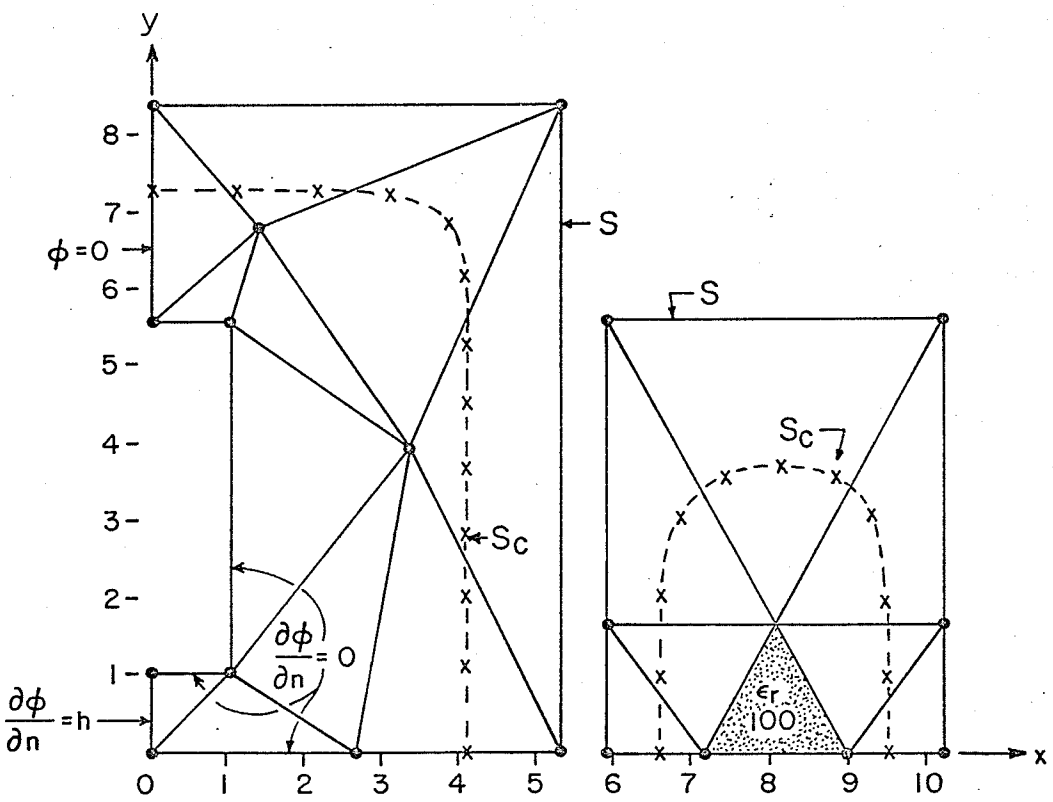


Fig. 3.17 The Antenna-Type Problem: Reduced Picture-Frames

With N=2 there are 57 node potentials to be determined.* The frequency was chosen to be 5 GHz, resulting in a propagation function $W=1.1\varepsilon_r$ (with dimensions in centimeters). The equipotentials of the real part of H_z in the two picture-frames are shown in Fig. 3.18. The wobble of the contours, quite noticeable in the larger picture-frame results from using too low an order of approximation. There is continuity of potential between the picture-frames, as shown particularly by the contour numbered 4. Field behaviour is seen to be correct near the dielectric obstacle.

It must be noted that this solution was obtained without z-direction propagation. If such propagation were allowed, a single vector potential would not be sufficient to represent the electromagnetic field due to the coupling of electric and magnetic fields at the air-dielectric interface [18, p.303]. In such cases full vector potential formulation may be required, as described by English [22], although in certain cases [23-24] coupled E_z, H_z potentials may be sufficient. Investigation of such problems would be a logical extension of this work.

---

* The selection N=2 was made for convenience, and in no way limits the method. The packaged algorithms catered for a maximum of 75 unknowns, and this value was felt to be sufficiently large for demonstration purposes.

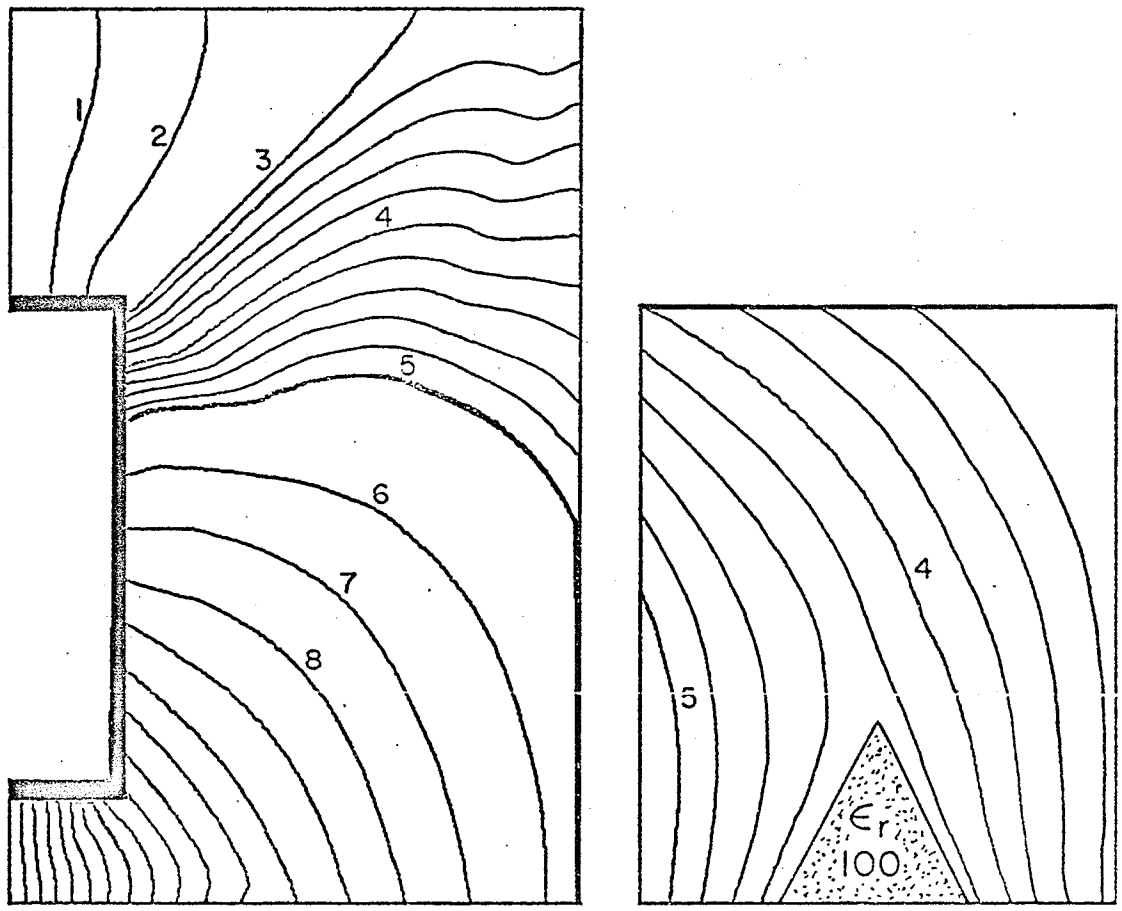


Fig. 3.18 The Antenna-Type Problem: Contours of H_z at 5 GHz

3.7 Conclusions

Unbounded field problems are conveniently solved by constraining PDE variational expressions in picture-frames to satisfy an IE condition. Potential is obtained directly, within the picture-frames, and may be generated at any point outside the picture-frames by an integration. The potential in the exterior region satisfies-exactly-the homogeneous PDE.

Picture-frames may be placed wherever they are required, and the distance between picture-frames is not important. Computing effort is conserved: PDE solutions are computed only within picture-frames, and the IE constraint is computed using only the simplest, free-space Green's functions. Explicit treatment of the Green's function singularity is avoided by the separation of S and S_c; the singularity is literally sidestepped. This device permits the use of quite simple algorithms.

There is some trade-off, however. In order to sidestep the singularities a homogeneous overlap region between S and S_c is required, and this makes the picture-frames possibly larger than necessary, thereby increasing computation for the PDE portion. It is also necessary to specify the contour S_c separately, but this is not difficult.

A significant contribution is the extension to the Helmholtz problem. With the proper exterior IE expression, energy convergence within the picture-frames is guaranteed by Mikhlin's theorem.

The use of finite-elements for representation of ϕ permits one to deal with much smaller (although denser) matrices than does the use of finite differences. Expensive iteration schemes often associated with finite-difference methods are avoided, although certain direct methods useful in finite-difference work are applicable (see Chapter VII). For the examples presented here, simple direct methods yield fast, accurate solution of the matrix equations.

3.8 References

[1] P. Silvester and M.S. Hsieh, "Finite-element solution of 2-dimensional exterior field problems." Proc. IEE, Vol. 118, pp.1743-1747, Dec. 1971.

[2] C.G. Williams and G.K. Cambrell, "Efficient numerical solution of unbounded field problems." Elect. Letts., Vol. 8, pp.247-248, May 1972.

[3] I. Cermak, "Solution of unbounded field problems by boundary relaxation." Ph.D. Thesis, Dept. of Elect. Eng., McGill University, Montreal, March 1969.

[4] I. Cermak and P. Silvester, "Solution of two-dimensional field problems by boundary relaxation." Proc. IEE, Vol. 115, pp.1341-1348, September 1968.

[5] I. Cermak and P. Silvester, "Analysis of coaxial line discontinuities by boundary relaxation." IEEE Trans. MTT-17, pp.489-495, August 1969.

[6] F. Sandy and J. Sage, "Use of finite difference approximations to partial differential equations for problems having boundaries at infinity." IEEE Trans. MTT-19, pp.484-486, May 1971

[7] S.G. Mikhlin, Variational Methods in Mathematical Physics. New York: Macmillan, 1964.

[8] I. Stakgold, Boundary Value Problems of Mathematical Physics, Vol. II. New York: Macmillan, 1968.

[9] R.E. Collin, Field Theory of Guided Waves. New York: McGraw-Hill, 1960.

[10] A.J. Burton and G.F. Miller, "The application of integral equation methods to the numerical solution of some exterior boundary value problems." Proc. of the Royal Society, Series A, Vol. 323, No. 1553, pp. 201-210, 8 June 1971.

[11] D. Greenspan and P. Werner, "A numerical method for the exterior Dirichlet problem for the reduced wave equation." Arch. Ration. Mech. Anal., Vol. 23, pp.288-316, 1966.

[12] G.E. Forsythe and W.R. Wasow, Finite Difference Methods for Partial Differential Equations. New York: John Wiley and Sons, 1960.

[13] M.J. Forray, Variational Calculus in Science and Engineering. New York: McGraw-Hill, 1968.

[14] M.C. Décreton, "Treatment of singularities in the finite-element method." M.Sc. Thesis, University of Manitoba, Winnipeg, 1972.

[15] H.B. Palmer, "The capacitance of the parallel plate capacitor by the Schwartz-Christoffel transformation." AIEE Trans., Vol. 56, pp. 363-366, 1937.

[16] H.A. Wheeler, "Transmission line properties of parallel strips separated by a dielectric sheet," IEEE Trans. MTT-13, pp.172-185, March 1965.

[17] T.G. Bryant and J.A. Weiss, "Parameters of microstrip transmission lines and of coupled pairs of microstrip lines." IEEE Trans. MTT-16, pp. 1021-1027, December 1968.

[18] J.R. Reitz and F.J. Milford, Foundations of Electromagnetic Theory. Reading, Mass.: Addison-Wesley, 1960.

[19] B.H. McDonald and A. Wexler, "Finite-element solution of unbounded field problems." IEEE Trans. MTT-20, pp.841-847, December 1972.

[20] E. Yamashita and R. Mittra, "Variational Method for the Analysis of Microstrip lines." IEEE Trans. MTT-16, pp.251-256, April 1968.

[21] R.F. Harrington, Time Harmonic Electromagnetic Fields. New York: McGraw-Hill, 1961.

[22] W.J. English, "A computer-implemented vector variational solution of loaded rectangular waveguides." SIAM J. App'l Math., Vol. 21, pp. 461-468, November 1971.

[23] Z.J. Csendes and P. Silvester, "Numerical solution of dielectric loaded waveguides: I-finite-element analysis." IEEE Trans., MTT-18, pp.1124-1131, December 1970.

[24] Z.J. Csendes and P. Silvester, "Numerical solution of dielectric loaded waveguides: II-modal approximation techniques." IEEE Trans. MTT-19, pp.504-509, June 1971.

CHAPTER IV

VARIATIONAL SOLUTION OF THE FREDHOLM INTEGRAL EQUATION

In this chapter variational techniques are described for obtaining energy convergent approximate solutions to the static problem

$$\frac{1}{\varepsilon_o} \int_S \sigma(\underline{r}')\, G(\underline{r}|\underline{r}')\, ds' = g(\underline{r})\,;\ \underline{r} \text{ on } S \tag{4.1}$$

where G is the free-space Green's function. For three-dimensional problems it is (1.8)

$$G(\underline{r}|\underline{r}') = \frac{1}{4\pi|\underline{r}-\underline{r}'|} \tag{4.2a}$$

and for two-dimensional problems it is (3.5a)

$$G(\underline{r}|\underline{r}') = \frac{-1}{2\pi} \ln|\underline{r}-\underline{r}'| \quad . \tag{4.2b}$$

In operator notation, (4.1) may be written as

$$K\sigma = g \tag{4.3}$$

and of special interest are the properties of the integral operator K. One may observe from (4.1) and (4.2) that the Green's function singularity cannot be avoided - it must be treated explicitly, as $\underline{r}'$ and $\underline{r}$ in (4.1) both lie on S.

There is also the problem of singular charge distributions to be considered. Suppose that S is a rectangular conducting plate lying in the y=0 plane with edges at $z=\pm b$ and $x=\pm a$, as shown in Fig. 4.1a. The Dirichlet problem (4.1) is

$$\frac{1}{4\pi\varepsilon_o}\int_{-b}^{b}\int_{-a}^{a}\frac{\sigma(x,z)\,dx\,dz}{\sqrt{(x-x')^2+(z-z')^2}} = g \qquad (4.4)$$

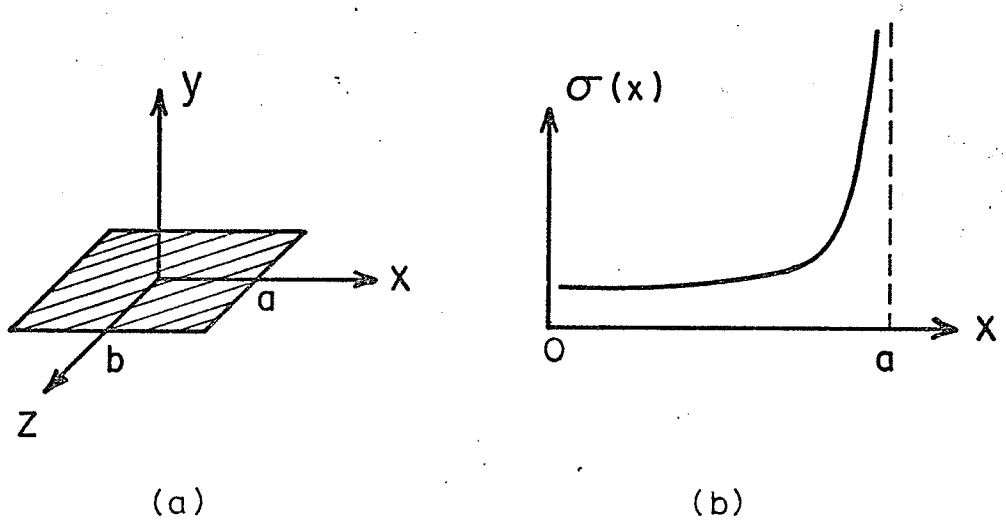
Consider the partial derivative with respect to x' :

$$\frac{1}{4\pi\varepsilon_o}\int_{-b}^{b}\int_{-a}^{a}\frac{\sigma(x,z)\,(x-x')}{((x-x')^2+(z-z')^2)^{3/2}}\,dx\,dz = 0$$

or

$$\frac{1}{4\pi\varepsilon_o}\int_{-b}^{b}\int_{-a}^{x'}\frac{\sigma(x,z)(x'-x)dxdz}{((x-x')^2+(z-z')^2)^{3/2}}$$

$$= \frac{1}{4\pi\varepsilon_o}\int_{-b}^{b}\int_{x'}^{a}\frac{\sigma(x,z)(x-x')\,dxdz}{((x-x')^2+(z-z')^2)^{3/2}} \qquad (4.5)$$



(a) (b)

Fig. 4.1 Source Distribution on a Rectangular Conductor
a) A View of the Configuration
b) The Singular Source Distribution

This equation states that the x component of the electrostatic force must balance at any point $-a<x'<a$ on the plate. Consider what happens as $x'\to a$. The integral on the right-hand side of (4.5) is computed over a vanishing area, but the value of the integral must be finite for any non-zero value of $x'-a$. Therefore the integrand must grow infinitely large as $x'\to a$. (See Fig. 4.1b) This is the only way of producing a finite, non-zero integral to satisfy (4.5). Jones [1, pp.566-569] summarizes much of the work done prior to 1966 in investigating this singularity problem and the well known result is that in the absence of corners

$$\sigma(x) \propto \frac{1}{\sqrt{x}} \tag{4.6}$$

where x is the distance from the edge.

4.1 The Integral Operator

The operator K is self-adjoint due to the symmetry of the Green's function (4.2): the inner product (see Section 2.1) is taken over S, and

$$\begin{aligned} \langle Ku,v\rangle &= \frac{1}{\varepsilon_o}\int_S v(\underline{r})\int_S u(\underline{r}')\ G(\underline{r}|\underline{r}')ds'ds \\ &= \frac{1}{\varepsilon_o}\int_S u(\underline{r})\int_S v(\underline{r}')\ G(\underline{r}|\underline{r}')\ ds'\ ds \\ &= \langle u,Kv\rangle \ . \end{aligned} \tag{4.7}$$

S is a curve in the plane or a surface in space for problems in two or three dimensions, respectively.

The operator is also positive-definite if ϕ vanishes at infinity [2, p.158]. Using equations (3.11)

$$\langle K\sigma,\sigma\rangle = \frac{1}{\varepsilon_o}\int_S \sigma(\underline{r})\int_S \sigma(\underline{r}')\ G(\underline{r}|\underline{r}')\ ds'\ ds = \int_S \sigma(\underline{r})\ \phi(\underline{r})\ ds$$

$$= \varepsilon_o \int_S \left(\frac{\partial\phi}{\partial n_-} - \frac{\partial\phi}{\partial n_+}\right)\phi\ ds \ . \tag{4.8}$$

Applying Green's identity, (with $\phi\to 0$ at infinity)yields

$$\langle K\sigma,\sigma\rangle = \int_S \sigma\ \phi ds = \varepsilon_o \int_\Omega \nabla\phi\cdot\nabla\phi\ \ d\Omega \tag{4.9}$$

where here Ω is all space: the infinite region. K is positive-definite, and the energy in K is proportional to electrostatic energy, as in the PDE case.

It is now necessary to examine the behaviour of ϕ at infinity. If S is of finite extent, ϕ vanishes at infinity if the Green's function so vanishes. Let $\underline{r}'$ be a point on S. For three-dimensional problems

$$\lim_{\underline{r}\to\infty} \quad G(\underline{r}|\underline{r}') \;=\; \frac{1}{4\pi|\underline{r}'-\infty|} \;=\; 0 \tag{4.10}$$

and the operator is positive-definite. For two-dimensional problems, however

$$\lim_{\underline{r}\to\infty} \quad G(\underline{r}|\underline{r}') \;=\; -\frac{1}{2\pi} \ln|\underline{r}'-\infty| \;=\; -\infty \quad . \tag{4.11}$$

Therefore potential can become infinite as $\underline{r}\to\infty$, and another approach is necessary.

Suppose that G is strictly positive for any two points $\underline{r},\underline{r}'$ on S; that is

$$G(\underline{r}|\underline{r}') \geq \gamma > 0 \; ; \quad \underline{r},\underline{r}' \text{ on } S \quad . \tag{4.12}$$

Then

$$\begin{aligned} \langle K\sigma,\sigma\rangle \;&=\; \frac{1}{\varepsilon_o} \int_S \sigma(\underline{r}) \int_S \sigma(\underline{r}')\, G(\underline{r}|\underline{r}')\, ds'\, ds \\ &\geq \frac{\gamma}{\varepsilon_o} \left(\int_S \sigma(\underline{r})\, ds \right)^2 \geq 0 \end{aligned} \tag{4.13}$$

and the operator is positive-definite.

Consider a modified Green's function for two-dimensional problems:

$$G(\underline{r}|\underline{r}') \;=\; -\frac{1}{2\pi} \ln\left(\frac{|\underline{r}'-\underline{r}|}{d}\right) \tag{4.14}$$

where, for any two points $\underline{r},\underline{r}'$ on S, $d > |\underline{r}-\underline{r}'|$. The argument of the logarithm is less than unity, and G is strictly positive on S.

Equation (4.13) holds and positive-definiteness is established. Since K is a self-adjoint positive-definite operator, or can be made so for a particular two-dimensional problem by the inclusion of a suitable constant, one may seek an energy convergent solution by minimizing the functional

$$F(\sigma) = \langle K\sigma,\sigma\rangle - 2\langle\sigma,g\rangle$$

$$= \frac{1}{\varepsilon_o}\int_S \sigma(\underline{r}) \int_S \sigma(\underline{r}')G(\underline{r}|\underline{r}')\, ds'\, ds$$

$$- 2\int_S \sigma(\underline{r})\, g(\underline{r})\, ds \quad . \qquad (4.15)$$

4.2 Previous Work

There have been many variational procedures reported (of which [3-5] are recent examples) for waveguide problems which have functionals of the form

$$F(\sigma) = \frac{\int_S \sigma(\underline{r}) \int_S \sigma(\underline{r}') \, G(\underline{r}|\underline{r}') ds' \, ds}{\left(\int_S \sigma(\underline{r}) \, ds \right)^2} \quad . \tag{4.16}$$

These procedures date back to Levine and Schwinger [6] and the general procedure may be found in Collin [7, p.154]. In this formulation G is not the free-space Green's function; rather it is a Green's function satisfying the boundary conditions that is constructed as a sum of the waveguide eigenmodes.

Integral transform techniques such as that of Yamashita and Mittra [8] are limited by the problem configuration: S must usually be an infinite straight line or plane.

Some very early work in solving the IE of interest here (4.1) was reported by Crout [9] and Crout and Hildebrand [10] who used polynomial trial functions for σ and a point-matching technique. Harrington [11] describes how the moment method may be used to solve IE problems. Spielman and Harrington [12] describe an IE moment method for finding eigenvalues; unfortunately their approach creates a nonlinear problem, as the eigenvalues occur within a transcendental Green's function. Silvester and Hsieh [13] describe a Galerkin procedure for solution of (4.1) and Benedek and Silvester [14] have investigated its application to the parallel strip capacitor. Pontoppidan [15] reports an IE solution by point-matching for static problems in two dimensions.

The bibliography produced by Noble [16] is invaluable. Green's functions are well covered by Roach [17], and Bergman [18] discusses IE operators. Kellogg [19], as well, is a basic reference for potential theory, and Mikhlin's two books on integral equations [20-21] are extremely useful.

4.3 Finite Elements

The functional (4.15) is minimized in a fashion analogous to that used for the PDE functional (2.48) earlier. Here S is divided into finite-elements which will be treated independently. The procedure is more complicated here than for the PDE case, however. Suppose that there are two elements, S_1 and S_2 in Fig. 4.2, each with a constant source σ_i. The functional, then, is

$$F(\sigma) = \frac{\sigma_1^2}{\varepsilon_o} \int_{S_1} \int_{S_1} G(\underline{r}|\underline{r}')ds'ds + \frac{2\sigma_1\sigma_2}{\varepsilon_o} \int_{S_1} \int_{S_2} G(\underline{r}|\underline{r}')ds'\, ds$$

$$+ \frac{\sigma_2^2}{\varepsilon_o} \int_{S_2} \int_{S_2} G(\underline{r}|\underline{r}')ds'\, ds - 2\,\sigma_1 \int_{S_1} g(\underline{r})ds - 2\sigma_2 \int_{S_2} g(\underline{r})ds \,. \tag{4.17}$$

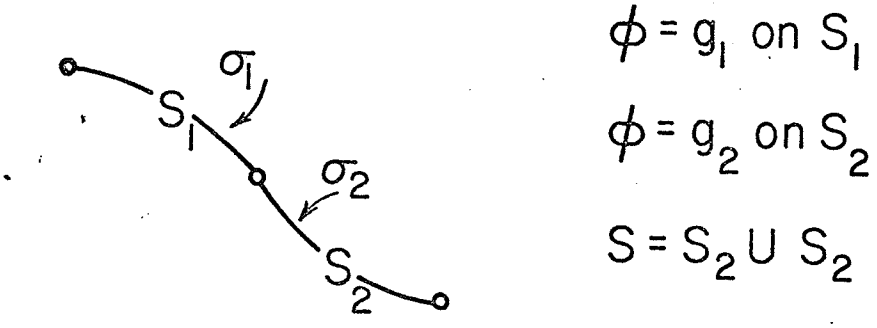


Fig. 4.2 IE Finite Elements

The Rayleigh-Ritz matrix equation is obtained by setting $\partial F/\partial \sigma_i = 0$; i=1, 2. It is

$$\frac{1}{\varepsilon_o}\begin{bmatrix} \int_{S_1}\int_{S_1} G(\underline{r}|\underline{r}')ds'ds & \int_{S_1}\int_{S_2} G(\underline{r}|\underline{r}')ds'ds \\ \int_{S_2}\int_{S_1} G(\underline{r}|\underline{r}')ds'ds & \int_{S_2}\int_{S_2} G(\underline{r}|\underline{r}')ds'ds \end{bmatrix}\begin{pmatrix} \sigma_1 \\ \sigma_2 \end{pmatrix} = \begin{pmatrix} \int_{S_1} g(\underline{r})ds \\ \int_{S_2} g(\underline{r})ds \end{pmatrix} \quad (4.18)$$

The matrix is symmetric and dense. There is contribution from every source to every other source. This is characteristic of such IE formulations.

One may ask whether or not it is possible to use a finite-element approach here and treat each element individually, in turn, as was done for PDE problems in Chapter II. Is it possible to treat element S_1 independently from S_2? The answer is yes, provided that a simple device is used: in dealing with element S_1 one simply assumes that the source on element S_2 is known. This "known" source σ_2 produces a potential at S_1 which is subtracted from the specified Dirichlet condition, resulting in a new Dirichlet problem to solve. In principle, then, one treats each element S_i separately and includes contributions from other elements in the form of an augmented Dirichlet condition.

Applying this idea to the present example, consider the first element S_1 and assume σ_2 to be completely known. Then the equation to be solved is

$$\frac{\sigma_1}{\varepsilon_o}\int_{S_1} G(\underline{r}|\underline{r}')ds' = g_1(\underline{r}) \quad (4.19)$$

where the augmented Dirichlet condition

$$g_1(\underline{r}) = g(\underline{r}) - \frac{\sigma_2}{\varepsilon_o} \int_{S_2} G(\underline{r}|\underline{r}')ds' \tag{4.20}$$

is assumed to be completely known. The variational procedure may be applied to (4.19) to yield

$$\frac{\sigma_1}{\varepsilon_o} \int_{S_1} \int_{S_1} G(\underline{r}|\underline{r}')ds'ds = \int_{S_1} g_1(\underline{r})ds \quad . \tag{4.21}$$

The second element may be treated similarly, and the following matrix equation results:

$$\frac{1}{\varepsilon_o} \begin{bmatrix} \int_{S_1} \int_{S_1} G(\underline{r}|\underline{r}')ds'ds & 0 \\ 0 & \int_{S_2} \int_{S_2} G(\underline{r}|\underline{r}')ds'ds \end{bmatrix} \begin{pmatrix} \sigma_1 \\ \sigma_2 \end{pmatrix} = \begin{pmatrix} \int_{S_1} g_1(\underline{r})ds \\ \int_{S_2} g_2(\underline{r})ds \end{pmatrix} . \tag{4.22}$$

If g_1 and g_2 were actually known, this matrix equation could produce the solution. However, g_1 and g_2 contain the unknowns. Using equation (4.20) for S_1 and performing a similar operation for S_2 results in

$$\begin{pmatrix} \int_{S_1} g_1(\underline{r})ds \\ \int_{S_2} g_2(\underline{r})ds \end{pmatrix} = \begin{pmatrix} \int_{S_1} g(\underline{r})ds \\ \int_{S_2} g(\underline{r})ds \end{pmatrix} - \frac{1}{\varepsilon_o} \begin{bmatrix} 0 & \int_{S_1}\int_{S_2} G(\underline{r}|\underline{r}')ds'ds \\ \int_{S_2} \int_{S_1} G(\underline{r}|\underline{r}')ds'ds & 0 \end{bmatrix} \begin{pmatrix} \sigma_1 \\ \sigma_2 \end{pmatrix} . \tag{4.23}$$

This may be substituted into (4.22), and equation (4.18) immediately results. One may think of these equations as having come from the operator equation

$$(K_o + K')\ \sigma = g \tag{4.24}$$

where K_o is the operator over each element, taken individually, and where K' is the operator which interconnects or mutually constrains the elements. K_o produces the matrix in (4.22); K' the matrix is (4.23).

One may pause for a moment to consider what this means in comparison with the PDE formulation presented in Chapter II. In the latter, it is necessary to have continuity of potential between any two elements and in practise one assigns a single node potential to node points at the same physical location. Suppose that this were not the case, that instead, two coincident node points on an inter-element boundary were each assigned a separate potential, one for each element. It would then be necessary to force these two potentials to assume the same value (for continuity) and Lagrange constraints, for example, could be used to accomplish this. The PDE finite-elements would then be mutually constrained to enforce continuity.

Returning now to the IE problem, two simple elements are presented for use in the subsequent examples: the line element for problems in the plane, and the rectangular element for three dimensional field problems. On each element the source is taken to be constant. This is denoted pulse function approximation, and the approximate solution is complete when the pulse heights, σ_i, i=1 .. N are found. These elements are depicted in Fig. 4.3.

Consider the line element first. Suppose the normal direction is to the left of the element. The direction cosines of the normal may be obtained from the end point coordinates of the line, just as in (3.16). Let the (u,v) coordinate system be defined as shown in Fig. 4.4. The (u,v) coordinates of the observer point $\underline{r}_p$ are given by

$$\begin{aligned} u_p &= (y_a - y_p)n_x + (x_p - x_a)n_y \\ v_p &= (x_p - x_a)n_x + (y_p - y_a)n_y \end{aligned} \qquad (4.25)$$

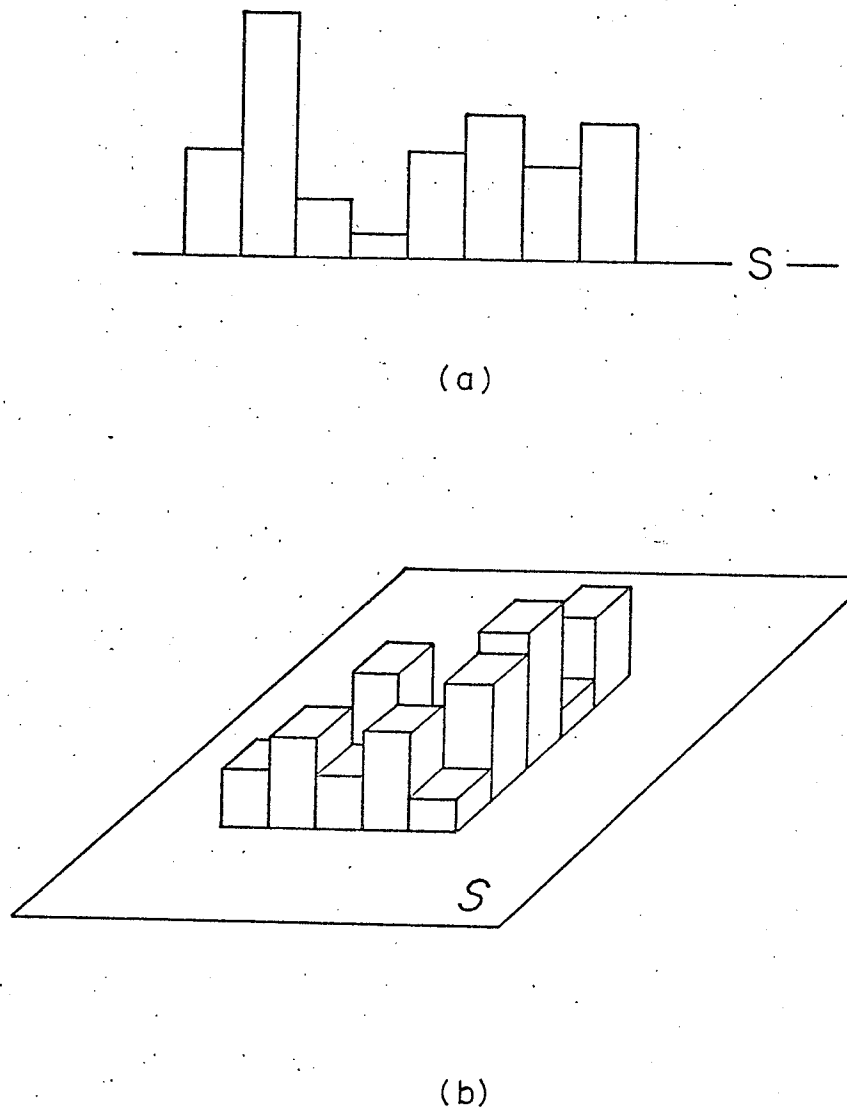


Fig. 4.3 Simple Pulse Function Elements
a) Line Elements for 2-D Problems
b) Rectangular Elements for 3-D Problems

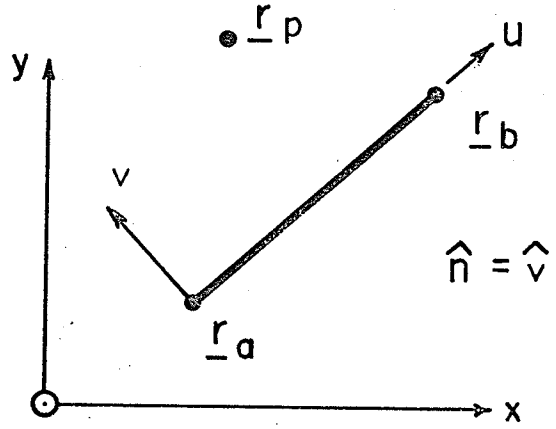


Fig. 4.4 A Local Coordinate System for Line Elements

Note that v_p is the perpendicular distance from $\underline{r}_p$ to the line element. If $v_p \neq 0$, $\underline{r}_p$ is not on the element. Also, if $u_p < 0$ or if $u_p > u_b$, the point $\underline{r}_p$ is not on the element, even when $v_p = 0$. It will be found useful later to test whether or not an observation point is on an element, and, if it is not, to know how close it lies to a particular element.

The integral of the Green's function over the element is of the form:

$$I = \int_{\underline{r}_a}^{\underline{r}_b} -\frac{1}{2\pi} \ln|\underline{r}-\underline{r}_p| \, ds \quad . \tag{4.26}$$

In the (u,v) coordinate system this integral becomes

$$I = -\frac{1}{2\pi} \int_0^{u_b} \ln \sqrt{(v_p)^2 + (u-u_p)^2} \, du \quad . \tag{4.27}$$

Making the substitution $t=u-u_p$, this integral becomes

$$I = \frac{-1}{2\pi}\int_{-u_p}^{u_b-u_p} \ln\sqrt{v_p^2+t^2}\,dt = \frac{-1}{2\pi}\int_{-u_p}^{0} \ln\sqrt{v_p^2+t^2}\,dt - \frac{1}{2\pi}\int_{0}^{u_b-u_p} \ln\sqrt{v_p^2+t^2}\,dt \tag{4.28}$$

which may be integrated analytically by the formula

$$\int_a^b \ln\sqrt{v_p^2+t^2}\,dt = t\{\ln\sqrt{v_p^2+t^2}-1\} + v_p\tan^{-1}\left(\frac{t}{v_p}\right)\Bigg|_a^b . \tag{4.29}$$

The result is

$$I = \frac{-1}{2\pi}\left\{ u_p\{\ln\sqrt{u_p^2+v_p^2}-1\} + v_p\tan^{-1}\left(\frac{u_p}{v_p}\right) + (u_b-u_p)\{\ln\sqrt{(u_b-u_p)^2+v_p^2}-1\} + v_p\tan^{-1}\left(\frac{u_b-u_p}{v_p}\right)\right\} \tag{4.30}$$

and u_b is simply the length of the line element. Thus the integral of G is obtained analytically for any observation point $\underline{r}_p$. Note that if $u_p=0$ or $u_p=u_b$, one of the integrals in (4.28) vanishes and need not be computed.

Consider now the rectangular element shown in Fig. 4.5. The coordinate system (u,v,w) is defined as shown. The normal direction is in the +w direction, and the direction cosines of $\hat{n}$ are given by

$$n_x = \{(y_b-y_a)(z_c-z_b) - (y_c-y_b)(z_b-z_a)\}/\Delta \tag{4.31a}$$

$$n_y = \{(z_b-z_a)(x_c-x_a) - (x_b-x_a)(z_c-z_b)\}/\Delta \tag{4.31b}$$

$$n_z = \{(x_b-x_a)(y_c-y_b) - (x_c-x_a)(y_b-y_a)\}/\Delta \tag{4.31c}$$

where Δ is the element area $\Delta=u_b v_d$. The coordinates of the observation point $\underline{r}_p$ in the (u,v,w) system are given by

$$u_p = \{(x_b-x_a)(x_p-x_a) + (y_b-y_a)(y_p-y_a) + (z_b-z_a)(z_p-z_a)\}/u_b \quad (4.32a)$$

$$v_p = \{(x_d-x_a)(x_p-x_a) + (y_d-y_a)(y_p-y_a) + (z_d-z_a)(z_p-z_a)\}/v_d \quad (4.32b)$$

$$w_p = (x_p-x_a)n_x + (y_p-y_a)n_y + (z_p-z_a)n_z \quad (4.32c)$$

As in the two dimensional case, the (u,v,w) coordinates of $\underline{r}_p$ may be used to determine whether or not $\underline{r}_p$ is on the element, and, if it is not, how close it is.

The integral of the Green's function over the rectangular element is of the form

$$I = \frac{1}{4\pi} \int_S \frac{ds}{|\underline{r}-\underline{r}_p|} \quad . \quad (4.33)$$

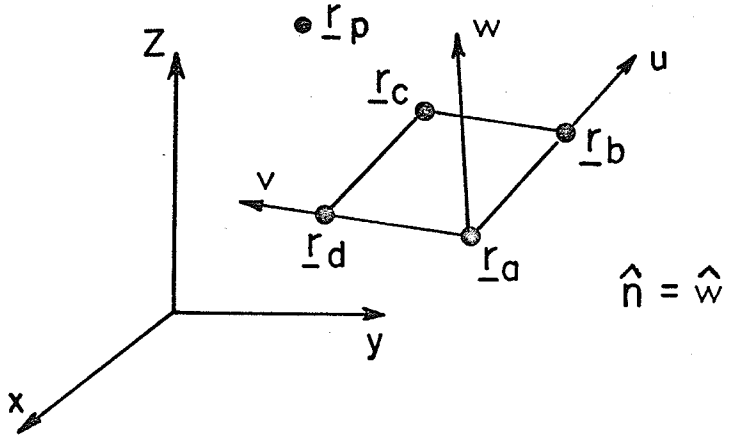


Fig. 4.5 A Local Coordinate System for Rectangular Elements

In the (u,v,w) coordinate system this integral becomes

$$I = \frac{1}{4\pi} \int_0^{v_d} \int_0^{u_b} \frac{du\ dv}{\sqrt{w_p^2+(u-u_p)^2+(v-v_p)^2}} \tag{4.34}$$

If $w_p \neq 0$ the addition-subtraction technique described in Appendix C, and suggested for use here by M. Friedman, is applied to give

$$I = \frac{1}{4\pi} \int_0^{v_d} \int_0^{u_b} \left(\frac{1}{\sqrt{w_p^2+(u-u_p)^2+(v-v_p)^2}} - \frac{1}{\sqrt{(u-u_p)^2+(v-v_p)^2}} \right) du\ dv$$

$$+ \frac{1}{4\pi} \int_0^{v_d} \int_0^{u_b} \frac{du\ dv}{\sqrt{(u-u_p)^2+(v-v_p)^2}} \ . \tag{4.35}$$

The first integrand is regular. This permits one to use a numerical integration technique, such as Gaussian quadrature.

The second integral in (4.35) may be obtained analytically. Let $t=u-u_p$ and $t'=v-v_p$. Then the required integral is of the form

$$I_2 = \int_{-v_p}^{v_d-v_p} \int_{-u_p}^{u_b-u_p} \frac{dt\ dt'}{\sqrt{t^2+t'^2}} \tag{4.36}$$

or

$$I_2 = \int_0^{-v_p} \int_0^{-u_p} \frac{dt\ dt'}{\sqrt{t^2+t'^2}} + \int_0^{v_d-v_p} \int_0^{u_b-u_p} \frac{dt\ dt'}{\sqrt{t^2+t'^2}}$$

$$- \int_0^{v_d-v_p} \int_0^{-u_p} \frac{dt\ dt'}{\sqrt{t^2+t'^2}} - \int_0^{-v_p} \int_0^{u_b-u_p} \frac{dt\ dt'}{\sqrt{t^2+t'^2}} \ . \tag{4.37}$$

Each of these is evaluated as

$$I(T,T') = \int_0^T \int_0^{T'} \frac{dt\, dt'}{\sqrt{t^2+t'^2}} = T' \ln\left(\frac{T}{T'} + \sqrt{1+\left(\frac{T}{T'}\right)^2}\right)$$

$$+ T \ln\left(\frac{T'}{T} + \sqrt{1+\left(\frac{T'}{T}\right)^2}\right); T,\ T' > 0 \quad . \tag{4.38}$$

Here T, T' are the absolute values of the limits in the four integrals in (4.37). The sign of the integral is positive if the limits are of the same sign, and the integral is negative if the signs of the limits are different. If either T or T' is zero, the integral (4.38) is simply not computed, for it is over a zero area in the (u,v) plane.

Note that a rectangular element always lies in a plane.

4.4 The Strip Capacitor

The two dimensional example problem of Section 3.3 is revisited, this time to seek an IE solution. Because of the symmetries, the source need be sought only in the positive quadrant, on the line y=1, in the region $0 \leq x \leq 1$. This region is conveniently divided into N line elements, each of length 1/N, as in Fig. 4.6, and pulse functions may be used to approximate the source. The inclusion of special functions to handle the known singular behaviour of the source distribution is also investigated.

4.4.1 Pulse Function Approximation

Pulse functions may be placed on the N line elements to approximate the source:

$$\sigma(x) = \sigma_i \quad \text{for } x \text{ on } S_i \, . \tag{4.39}$$

The symmetries are conveniently handled by using a mirror image Green's function, defined as G=P+Q, where [22, p.430]

$$P(\underline{r}|\underline{r}') = -\frac{1}{2\pi} \ln \left(\frac{(x+x')^2 + (y-y')^2}{\left[(x-x')^2 + (y+y')^2\right]\left[(x+x')^2 + (y+y')^2\right]} \right)^{1/2} \tag{4.40a}$$

and

$$Q(\underline{r}|\underline{r}') = -\frac{1}{2\pi} \ln \sqrt{\{(x-x')^2 + (y-y')^2\}} \, . \tag{4.40b}$$

The image symmetry at y=0 ensures that potential vanishes at infinity, and the operator is therefore positive-definite. Potential is given by

$$\phi(\underline{r}') = \frac{1}{\varepsilon_o} \int_0^1 \sigma(x) \, G(x',y'|x,1) \, dx \tag{4.41}$$

$$= \sum_{i=1}^{N} \frac{\sigma_i}{\varepsilon_o} \int_{S_i} G(x',y'|x,1) \, dx \, .$$

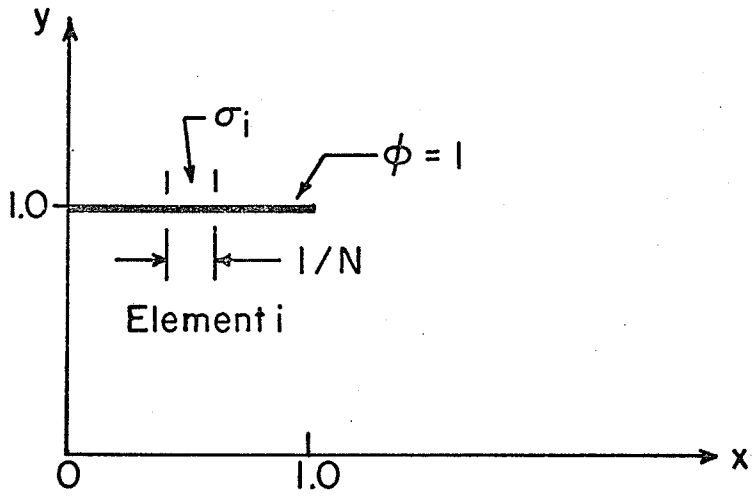


Fig. 4.6 Finite Elements for Strip Capacitor; IE Solution

The function P is treated as a regular function and Gaussian quadrature is used to compute

$$I = \int_{S_i} P(x',y'|x,1)\, dx \simeq \sum_{j=1}^{M} c_j\, P(x',y'|x_j,1) \qquad (4.42)$$

where c_j and x_j; j=1 .. M, are the M Gauss coefficients and integration points, respectively. See Appendix D.

The function Q is singular, but it may be integrated analytically in the local coordinates as described at (4.26), with $\underline{r}'$ as the observation point.

The Rayleigh-Ritz matrix equation is constructed exactly as described at equation (4.18) for an N=2 example. Here the NxN system is given by

$$\begin{bmatrix} \vdots & \\ \cdots A_{ij} & \cdots \\ \vdots & \end{bmatrix} \begin{pmatrix} \vdots \\ \sigma_i \\ \vdots \end{pmatrix} = \begin{pmatrix} \vdots \\ 1/N \\ \vdots \end{pmatrix} . \qquad (4.43)$$

Since the Dirichlet condition is unity the right-hand side entries are simply the element lengths. The matrix entries are

$$A_{ij} = \int_{S_i} \int_{S_j} P(\underline{r}|\underline{r}')ds'ds + Q(\underline{r}|\underline{r}')\, ds'\, ds$$

$$= \int_{S_i} f(\underline{r})\, ds \qquad (4.44)$$

where

$$f(\underline{r}) = \int_{S_j} P(\underline{r}'|\underline{r})\, ds' + \int_{S_j} Q(\underline{r}|\underline{r}')\, ds' \quad . \qquad (4.45)$$

The first integral in (4.45) is computed numerically (as at (4.42)) and the second, analytically. Finally (4.44) is computed numerically, for $f(\underline{r})$ is regular. Order 3 Gaussian quadrature is used for all numerical integrations.

Capacitance may be determined by the ratio of charge on one strip to the potential difference. Since half the charge on one strip is determined ($0 \leq x \leq 1$) and the potential difference is 2,

$$C = \frac{1}{N} \sum_{i=1}^{N} \sigma_i \quad . \qquad (4.46)$$

Results are given in Table 4.1. The choice of (40/9,10) as the point to display potential is somewhat arbitrary: this potential exhibited greater variation with N than others in an early test and was retained for comparison. Both capacitance and potential converge, and the correct capacitance, 18.72 pf/M is attained with 50 pulse functions. This solution was used to produce the "exact" equipotentials given in Section 3.3, Fig. 3.10. Note that the

capacitance (and hence potential energy) is maximized by this procedure. This is to be expected, as the IE is not homogeneous. See the comment at equation (2.36).

TABLE 4.1 THE STRIP CAPACITOR;
IE SOLUTION: PULSES

Number of Pulses	Capacitance pf/M	Potential ϕ (40/9, 10)
1	17.72	0.1063
2	18.20	0.1092
5	18.51	0.1110
10	18.62	0.1117
20	18.68	0.1121
30	18.70	0.1122
40	18.71	0.1123
50	18.72	0.1124

4.4.2 Approximation with Singular Trial Functions

It is known that the source distribution is singular at the strip edges [1, p.569]. It behaves as

$$\sigma_s(x) \sim \frac{1}{\sqrt{1+x}\,\sqrt{1-x}} = \frac{1}{\sqrt{1-x^2}} \quad . \tag{4.47}$$

Since the problem is solved in the positive quadrant, $x \geq 0$, it is possible to use the simpler function

$$\sigma_s(x) = \frac{1}{\sqrt{1-x}} \quad . \tag{4.48}$$

This function, with its own variational parameter σ_o, is introduced in an attempt to improve convergence. The approximation now is

$$\sigma(x) = \sigma_i + \frac{\sigma_o}{\sqrt{1-x}}, \quad x \text{ on } S_i \quad . \tag{4.49}$$

One equation is added to the matrix equation (4.43), and is obtained by setting $\partial F/\partial \sigma_o = 0$. It is

$$\sum_{i=1}^{N} \frac{\sigma_i}{\varepsilon_o} \int_0^1 \frac{1}{\sqrt{1-x}} \int_{S_i} G(x,1|x',1)dx'dx + \frac{\sigma_o}{\varepsilon_o} \int_0^1 \int_0^1 \frac{G(x,1|x',1)dx\, dx'}{\sqrt{1-x}\ \sqrt{1-x'}}$$

$$= \int_0^1 \frac{dx}{\sqrt{1-x}} \quad . \tag{4.50}$$

With subscript $_o$ denoting this equation

$$b_o = \int_0^1 \frac{dx}{\sqrt{1-x}} = 2 \tag{4.51a}$$

$$A_{oo} = \frac{1}{\varepsilon_o} \int_0^1 \frac{1}{\sqrt{1-x}} \int_0^1 \frac{G(x,1|x',1)\ dx'}{\sqrt{1-x'}}\, dx \tag{4.51b}$$

and

$$A_{oi} = A_{io} = \frac{1}{\varepsilon_o} \int_0^1 \frac{1}{\sqrt{1-x}} \int_{S_i} G(x,1|x',1)dx'\, dx \quad . \tag{4.51c}$$

The matrix entries are symmetric (4.51c) since the functional is symmetric in i and j. (Recall the derivation from (4.17).)

The addition-subtraction technique (Appendix C) is used to evaluate (4.51b) and (4.51c). Consider the off-diagonal entry first.

Equation (4.51c) is written as

$$A_{oi} = \frac{1}{\varepsilon_o} \int_0^1 \frac{1}{\sqrt{1-x}} \{f(x)-f(1)\}dx + \frac{f(1)}{\varepsilon_o} \int_0^1 \frac{dx}{\sqrt{1-x}} \tag{4.52}$$

where

$$f(x) = \int_{S_i} G(x,1|x',1)dx' \quad . \tag{4.53}$$

Equation (4.53) has already been considered, at (4.45). The last integral of (4.52) is given by (4.51a) and the integrand of the first integral of (4.52) is regular, permitting numerical integration.

The diagonal entry is handled similarly:

$$A_{oo} = \frac{1}{\varepsilon_o} \int_0^1 \frac{1}{\sqrt{1-x}} \{f(x)-f(1)\}dx + \frac{f(1)}{\varepsilon_o} \int_0^1 \frac{dx}{\sqrt{1-x}} \tag{4.54}$$

where

$$f(x) = \int_0^1 \frac{G(x,1|x',1)\, dx'}{\sqrt{1-x'}} \quad . \tag{4.55}$$

The addition-subtraction technique is applied to (4.55) to obtain f(x). Recalling the definitions of P and Q (4.40), f(x) is written as

$$f(x) = \int_0^1 \left\{\frac{P(x,1|x',1)-P(x,1|1,1)}{\sqrt{1-x'}}\right\} dx' + P(x,1|1,1) \int_0^1 \frac{dx}{\sqrt{1-x}}$$

$$- \frac{1}{2\pi} \int_0^1 \frac{\ln|x-x'|dx'}{\sqrt{1-x'}} \quad . \tag{4.56}$$

P is regular, permitting numerical evaluation of the first integral. The second integral has already been evaluated, at (4.51a) and the last may be obtained analytically:

$$-\frac{1}{2\pi}\int_0^1 \frac{\ln|x-x'|\,dx'}{\sqrt{1-x'}} = \frac{1}{\pi}\left\{2-\ln x + \sqrt{1-x}\,\ln\left(\frac{1-\sqrt{1-x}}{1+\sqrt{1-x}}\right)\right\} . \qquad (4.57)$$

The function f(x) is now available, and (4.54) is integrated in the same way as (4.52). The matrix equation may now be completed.

Capacitance is still given by the total charge on the strip $(0\leq x\leq 1)$

$$C = \frac{1}{N}\sum_{i=1}^{N}\sigma_i + 2\,\sigma_o \qquad (4.58)$$

unless N=0, in which case there is but one equation to be solved:

$$A_{oo}\;\sigma_o = b_o \qquad (4.59)$$

and here $C = 2\sigma_o$.

The results obtained using order 3 Gaussian quadrature for all numerical integrations are listed in Table 4.2. Convergence is much

TABLE 4.2 THE STRIP CAPACITOR; IE SOLUTION: PULSES AND SINGULAR FUNCTION

Number of Pulses	Number of unknowns	Capacitance pf/M	Potential $\phi(40/9,10)$
0	1	18.57	0.1115
1	2	18.72	0.1123
4	5	18.72	0.1124

faster with the special singular function: a total of 5 unknowns produces the accuracy previously only attained using 50 unknowns as given in Table 4.1. Although energy convergence is guaranteed, the rate of convergence depends very much on the choice of trial functions. An improvement with a special function to handle known edge singularities was observed for PDE formulations by Décreton [23], who introduced such functions to handle derivative ($\partial\phi/\partial n$) singularities at edges.

Another choice of singular trial function was suggested by M. Friedman:

$$\sigma(x) = \sum_{i=1}^{N} \frac{c_i x^{i-1}}{\sqrt{1-x}} = \underline{f}^T \underline{c};\ 0 \leq x \leq 1 \quad . \tag{4.60}$$

This function is defined over the entire range, so that in effect there is but one finite-element. Note that each term in the expansion is singular - the solution is constrained to this singular function. A similar trial function (with $\sqrt{1-x^2}$ in the denominator - see (4.47)) has been used by Silvester and Benedek [24] in investigation of certain open microstrip problems. They use a "quasi moment method" with Legendre polynomial weighting functions and compute all integrals numerically with specially modified Gaussian quadrature formulae.

Here, the Rayleigh-Ritz matrix equation is

$$[A]\ \underline{c} = \underline{b} \quad . \tag{4.61}$$

The entries to $\underline{b}$ may be computed recursively by the formula

$$\int_0^1 \frac{x^i dx}{\sqrt{1-x}} = \frac{2^i}{2i+1} \int_0^1 \frac{x^{i-1} dx}{\sqrt{1-x}} \quad . \tag{4.62}$$

Therefore

$$b_1 = 2$$

$$b_i = \frac{2^i}{2i+1}\, b_{i-1} \quad ; \quad i>1 \quad . \tag{4.63}$$

The matrix contributions are given by

$$A_{ij} = \frac{1}{\varepsilon_o} \int_0^1 \frac{x^{i-1}}{\sqrt{1-x}} \int_0^1 \frac{x'^{j-1}}{\sqrt{1-x'}}\, G(x,1|x',1)dx'\, dx \tag{4.64}$$

and the addition-subtraction technique is applied:

$$A_{ij} = \frac{1}{\varepsilon_o} \int_0^1 \frac{x^{i-1}}{\sqrt{1-x}} \{f(x)-f(1)\}dx + \frac{f(1)}{\varepsilon_o} \int_0^1 \frac{x^{i-1}}{\sqrt{1-x}}\, dx \quad . \tag{4.65}$$

The first integrand is regular, and the second integral has just been investigated. The function f(x) is computed using the addition-subtraction technique as follows:

$$f(x) = \int_0^1 \frac{x'^{j-1} G(x,1|x',1)dx'}{\sqrt{1-x'}} = \int_0^1 \frac{x'^{j-1}}{\sqrt{1-x'}} \{P(x,1|x',1)-P(x,1|1,1)\}dx'$$

$$+ P(x,1|1,1) \int_0^1 \frac{x'^{j-1}dx'}{\sqrt{1-x'}} - \frac{1}{2\pi} \int_0^1 \ln|x-x'| \left\{ \frac{x'^{j-1}}{\sqrt{1-x'}} - \frac{x^{j-1}}{\sqrt{1-x}} \right\} dx'$$

$$- \frac{x^{j-1}}{2\pi\sqrt{1-x}} \int_0^1 \ln|x-x'|dx' - \frac{1}{2\pi} \int_0^1 \frac{\ln|x-x'|\, dx'}{\sqrt{1-x'}} \quad . \tag{4.66}$$

The second, fourth and fifth integrals have already been obtained analytically, at (4.51), (4.26) and (4.57) respectively. The remaining two have regular integrands and may be obtained numerically. Although each may be integrated directly over (0,1) it is perhaps advisable to

perform the numerical integration in two steps; over (0,x) and (x,1). This would relieve difficulties caused by a possible derivative singularity at x. (See Appendix C). The example is presented using order 10 Gaussian quadrature directly over (0,1), and the results are listed in Table 4.3. Convergence is essentially the same as in the previous example, with 4 figure accuracy being obtained with 5 unknowns.

TABLE 4.3 THE STRIP CAPACITOR; IE SOLUTION
SINGULAR POLYNOMIAL FUNCTION

Order of Polynomial	Number of unknowns	Capacitance pf/M	Potential $\phi(40/9,10)$
0	1	18.57	0.1115
1	2	18.72	0.1124
2	3	18.72	0.1124

It is computationally easier to include the singular function separately, with its own variational parameter. It may also be better in cases where the singularity is not as well known; forcing every trial function to behave incorrectly may not be wise. If an inappropriate function is introduced with a separate parameter, the variational method may simply ignore it by setting the parameter to zero if it is inappropriate.

4.5 A Three-Dimensional Example

Consider the two T shaped conductors shown in Fig. 4.7a with a potential difference of 2 volts. With the coordinate system shown, the symmetries may be used to reduce the problem to that of finding a source distribution on the portion shown in Fig. 4.7b, at z=1. The image Green's function is also constructed as G=P+Q where

$$P(\underline{r}|\underline{r}') = \frac{1}{4\pi}\left\{\frac{1}{\sqrt{(x-x')^2+(y+y')^2+(z-z')^2}} - \frac{1}{\sqrt{(x-x)^2+(y-y')^2+(z+z')^2}}\right.$$

$$\left. - \frac{1}{\sqrt{(x-x')^2+(y+y')^2+(z+z')^2}}\right\} \qquad (4.67a)$$

and

$$Q(\underline{r}|\underline{r}') = \frac{1}{4\pi\sqrt{(x-x')^2+(y-y')^2+(z-z')^2}} \quad . \qquad (4.67b)$$

Since the problem is three-dimensional the operator is positive-definite (4.10).

The function P is regular and is integrated over a rectangular region by double application of simple Gaussian quadrature (Appendix D):

$$\int_0^a \int_0^b f(x,y)dy\,dx \simeq \sum_{i=1}^{M}\sum_{j=1}^{M} c_i(a)c_j(b)\, f(x_i(a), y_j(b)) \qquad (4.68)$$

The integral of Q over the rectangular element is obtained by the method described at (4.33).

The problem is investigated first with pulse function approximation and then with a special singular trial function to handle the edge and corner singularities of the plate source distribution.

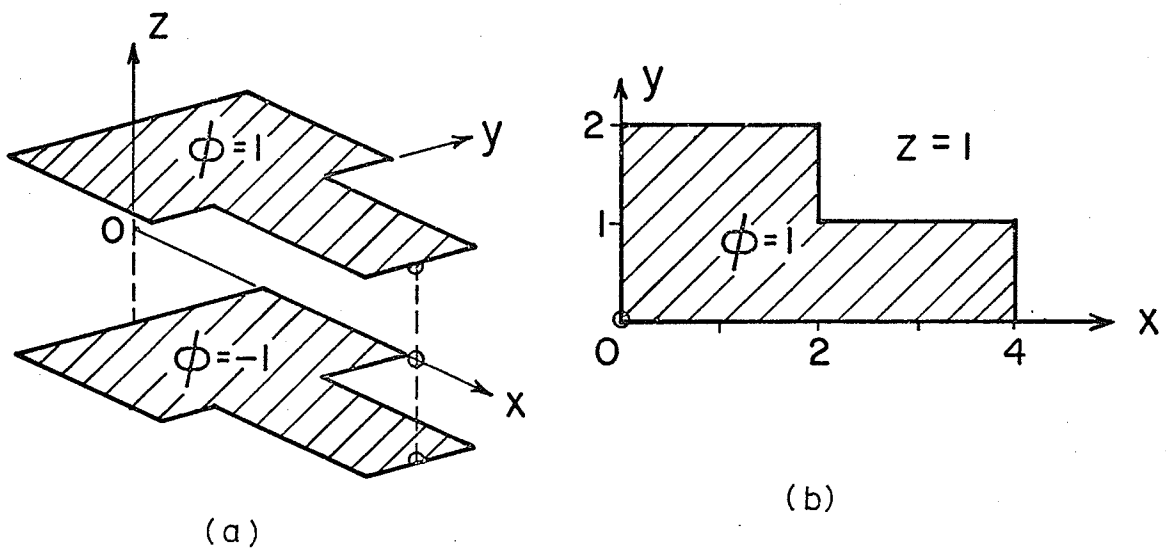


Fig. 4.7 A 3-D Example.
a) Full View of Both Plates
b) Reduced Geometry; View from above.

4.5.1 Pulse Function Approximation

The plate region is divided into square elements of side 1/N. Therefore there are $6N^2$ elements, and with pulse function approximation, $6N^2$ pulse heights σ_i to be found. Capacitance is given by

$$C = \frac{1}{N^2} \sum_{i=1}^{6N^2} \sigma_i \quad . \tag{4.69}$$

The Rayleigh-Ritz matrix equation is constructed exactly as in the simple example of Section 4.3, at equation (4.18). Here, it is

$$[\,A\,]\ \underline{\sigma}\ =\ \underline{b} \tag{4.70}$$

With ϕ=1 on the plate, the vector entries are all given by $1/N^2$, the element area, and the matrix entries are given by

$$A_{ij}\ =\ \frac{1}{\varepsilon_o}\int_{S_i}\int_{S_j} G(\underline{r}|\underline{r}')\ ds'\ ds\ . \tag{4.71}$$

These integrals are computed numerically:

$$A_{ij}\ =\ \frac{1}{\varepsilon_o}\int_{S_i} f(\underline{r})ds \tag{4.72}$$

where

$$f(\underline{r})\ =\ \int_{S_j} P(\underline{r}|\underline{r}')\ ds'\ +\ \int_{S_j} Q(\underline{r}|\underline{r}')ds'\ . \tag{4.73}$$

Equation (4.72) is integrated numerically, and f, in (4.73) is obtained by methods described above, in Section 4.3.

The results obtained, using order 3 Gaussian quadrature for all numerical integration, are listed in Table 4.4. Neither capacitance

TABLE 4.4 THE 3-D PROBLEM; IE SOLUTION: PULSES

N	Number of Pulses	Capacitance pf	Potential ϕ $(2,^3,2)$
1	6	119.4	0.1968
2	24	125.6	0.2064
3	54	127.8	0.2102
4	96	129.2	0.2127

nor field has converged with N=4 (96 unknowns) although energy convergence is monotonic, and from below as expected. Field convergence is also monotonic, as illustrated by potential at the point (2,3,2). (This point was chosen somewhat arbitrarily; it was a point of greatest change in potential with N in an early test and was retained for comparison purposes.)

Tests with N=5 (150 unknowns) resulted in a vanishing matrix determinant. This was originally thought to be due to ill-conditioning since such a symptom is common [25, pp.174-177], but some later thoughts suggest that the problem may have been due to the equation solving algorithm used (see Chapter VII) and the numerical values of the matrix [A]. If [A] were the NxN unit matrix, multiplication of [A] by γ would result in a determinant equal to γ^N. With the computer zero (IBM 360/65) occurring at $\sim 10^{-80}$, and the computer infinity at $\sim 10^{80}$, the permissable range of γ with N=80 is $0.1<\gamma<10.0$.

Pulse function representation for source in IE solution of a three-dimensional problem has resulted in a slowly converging minimizing sequence. It is in such problems as this that accurate representation of known singular behaviour may be expected to assist greatly in obtaining the solution. It would also be possible to concentrate the elements near the singular corners, and perhaps this merits investigation.

4.5.2 Approximation with a Singular Trial Function

The introduction of a special singular function to handle edge singularities in the strip capacitor problem resulted in greatly improved convergence properties (Section 4.4). If an accurate singu-

larity function can be obtained for the T shaped conductor, similar improvements might be expected here. Consider the following function, defined over the conductor in three parts as shown in Fig. 4.8.

$$\sigma_s(x,y) = \begin{cases} \{x(2-x)(2-y)\}^{-\alpha} & \text{on } S_a \quad (4.74a) \\ \{x(3-x-y)\}^{-\alpha} & \text{on } S_b \quad (4.74b) \\ \{(4-x)(1-y)\}^{-\alpha} & \text{on } S_c \quad . \quad (4.74c) \end{cases}$$

This function is continuous over the conductor. (This is not necessary, but it maintains consistency with the examples in Section 4.4)

Jones [1,pp.566-569], in summarizing previous work in analysing edge and corner singularities, suggests that the edge condition alone is sufficient and that corner singularities may be neglected. Therefore choosing α=0.5 in (4.74) creates a singular function σ_s which behaves

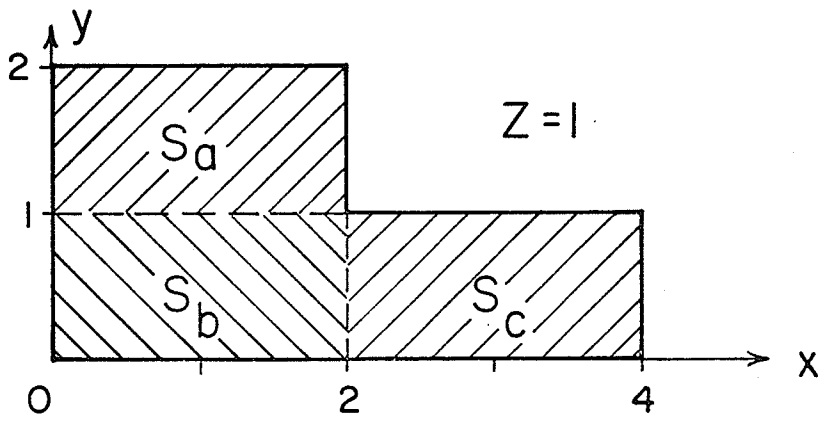


Fig. 4.8 The 3-D Example; 3 part segmentation of Plate for Singular Source Function

near the edges in the same fashion as the function used in Section 4.4 (at equation (4.48)). This function is introduced into the functional with its own variational parameter

$$\sigma(\underline{r}) = \sigma_i + \sigma_o \, \sigma_s(\underline{r}) \; ; \text{ in } S_i \tag{4.75}$$

and an additional equation results, similar to that at (4.50), which is added to the matrix equation (4.70). The additional equation is

$$\frac{1}{\varepsilon_o} \sum_{i=1}^{6N^2} \sigma_i \int_S \sigma_s(\underline{r}) \int_{S_i} G(\underline{r}|\underline{r}')ds' \, ds + \frac{\sigma_o}{\varepsilon_o} \int_S \sigma_s(\underline{r}) \int_S \sigma_s(\underline{r}') \, G(\underline{r}|\underline{r}')ds'ds$$

$$= \int_S \sigma_s(\underline{r})ds \quad . \tag{4.76}$$

The methods used for computing these integrals will be given in the next subsection.

The results, presented in Table 4.5, indicate that while convergence is faster with σ_s, it is not significantly faster (Table 4.4). The last column of Table 4.5 provides the answer; the variational method is choosing to ignore σ_s as more pulses become available to represent the source distribution. Thus the assumed edge and corner singularity function is not very good.

TABLE 4.5 THE 3-D PROBLEM; IE SOLUTION:
PULSES AND SINGULAR FUNCTION, α=0.5

N	Number of pulses	Capacitance pf	Potential $\phi(2,3,2)$	% capacitance from σ_s
0	0	122.8	0.2008	100.0
1	6	124.0	0.2048	77.9
2	24	126.9	0.2095	51.0
3	54	128.8	0.2129	34.0
4	96	129.9	0.2138	7.8

Braunbek [26] (to whom Jones also refers) suggests that the singularity near a corner is a function of the angle of a corner; that is

$$\sigma(\underline{r}) \sim \eta^{-\alpha} \qquad (4.77)$$

where η is distance from the corner, and α depends upon the angle.* For a 90° corner (of which the whole T shaped conductor has 6), he estimates a value $\alpha=0.3$, and for a 270° corner (of which the whole conductor has 2) he estimates a value $\alpha=0$; there is no singularity as the corner is approached from points on the conductor. Perhaps it is possible to select a different value of α in (4.74) which will compensate for the corner effect. But what value of α should be used?

Returning briefly to the pure edge singularity, consider the functions

$$\sigma_1(x) = (1-x)^{-\alpha} \qquad (4.78)$$

and

$$\sigma_2(x) = (1-x^2)^{-\alpha} \quad . \qquad (4.79)$$

With $\alpha=0.5$, these are the singular functions (4.48) and (4.47) which were prescribed for the parallel strip problem in Section 4.42. The addition-subtraction algorithm may be applied to compute

$$\phi(x) = \frac{1}{\varepsilon_o} \int_0^1 \sigma_i(x')\, G(x,1|x',1)ds' \; ; \; i=1,2 \qquad (4.80)$$

for various values of α. If σ_i may be thought of as the first term in a series expansion of σ, one could argue that adjusting α so that potential is most nearly constant in (4.78) might permit faster convergence of the series. Fig. 4.9 displays the results of the experiment. The functions $(1-x)^{-.46}$ and $(1-x^2)^{-.54}$ succeed in minimizing the standard deviation from a constant potential. The values of α are within ten percent of the correct value of 0.5.

* His solution is a "rather poor approximation"; α(angle) "may lie between... two curves", and between these two the "difference...is considerable".

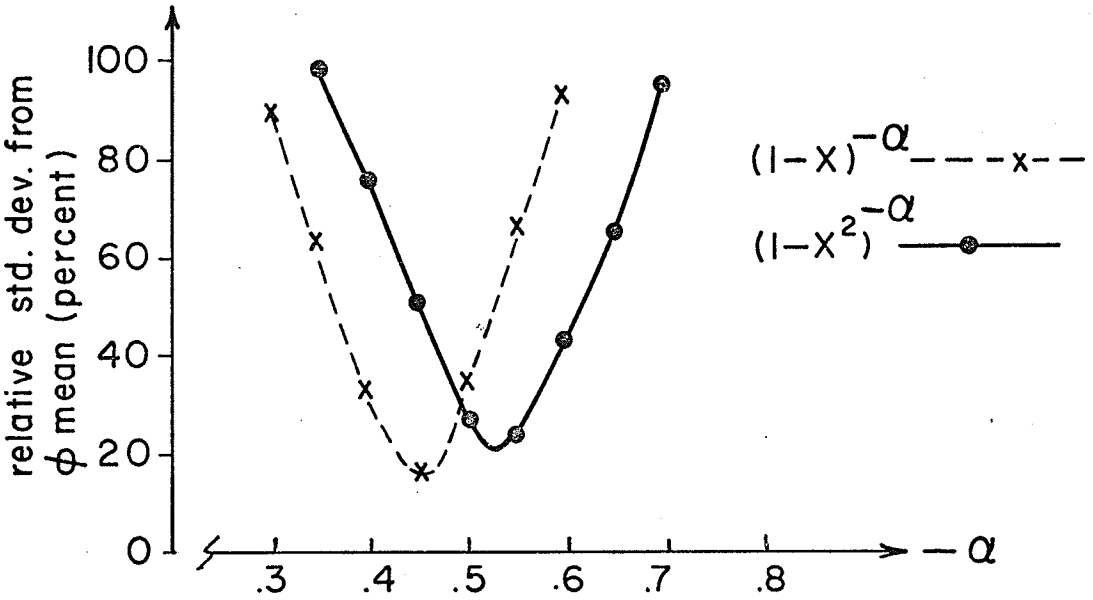


Fig. 4.9 Variation from ϕ(mean) as a Function of α; Edge Singularity Functions Alone

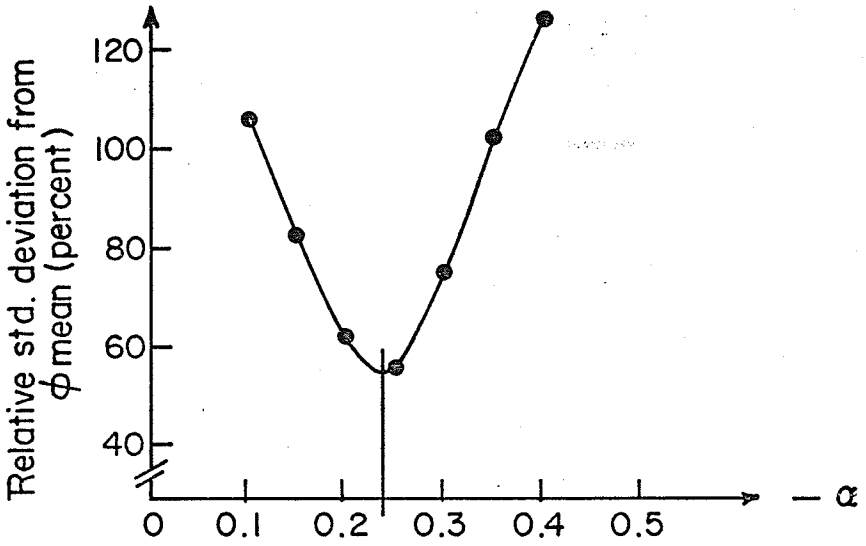


Fig. 4.10 Variation from ϕ(mean) as a Function of α; Edge and Corner Singularity Function

Potential due to σ_s in the three-dimensional problem is given by

$$\phi(\underline{r}) = \frac{1}{\varepsilon_o} \int_S \sigma_s(\underline{r}') \, G(\underline{r}|\underline{r}') \, ds' \, . \qquad (4.81)$$

The same experiment was carried out here, with deviation in ϕ from ϕ (mean) evaluated at 60 points on S. It may be observed in Fig. 4.10 that α=0.24 succeeds in producing the most nearly constant potential over S.

The variation procedure was repeated, with the addition of (4.76) to the matrix equation and with α=0.25. The results listed in Table 4.6 are somewhat better than those obtained with α=0.5 (in Table 4.5) but convergence is still quite slow, and the last column in Table 4.6 indicates that the variational method is again choosing to ignore σ_s. The assumed edge and corner singularity is still not very good.

TABLE 4.6 THE 3-D PROBLEM; IE SOLUTION: PULSES AND SINGULAR FUNCTION, α=0.25

N	Number of pulses	Capacitance pf	Potential $\phi(2,3,2)$	% capacitance from σ_s
0	0	124.6	0.2060	100.0
1	6	126.2	0.2078	78.2
2	24	128.0	0.2108	51.3
3	54	129.5	0.2131	34.1
4	96	130.3	0.2142	7.8

Another approach is presented by Silvester and Benedek [24] for corners in microstrip lines: they first of all determine the source distribution for a semi-infinite strip with all trial functions constrained to behave as $(1-x^2)^{-1/2}$ and with $\sigma(y)$ constant to the end of the strip. They then superimpose two such strips at right angles, to represent a 90^o corner, as shown in Fig. 4.11, and seek solution to

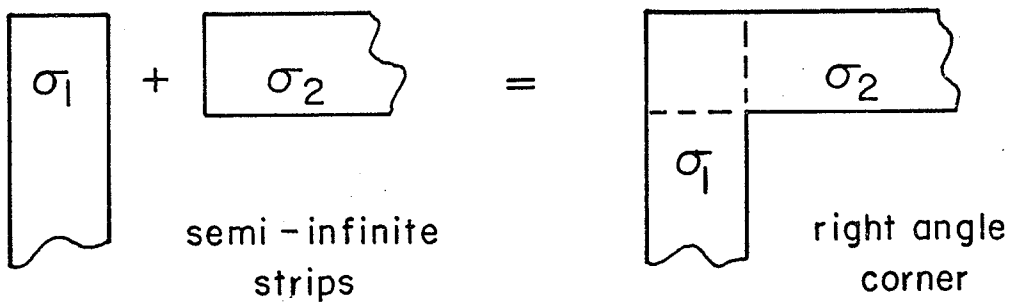


Fig. 4.11 Superposition of Semi-infinite Strips
to Represent a Corner

the Dirichlet problem. The result is a potential which does not (can not - due to the singular function behaviour) exactly satisfy the Dirichlet condition. This "error" is called a potential residual. They seek another source distribution on the strips to produce the potential residual. This additional charge produces an excess capacitance, which they are able to compare to measured data. Experiments such as Silvester and Benedek report may be useful in determining the nature of source distribution at corners. One result of the situation reported here is that work needs to be done in analysing source behaviour near edges and corners.

4.5.3 Integration of the Singular Functions

A summary is presented here of the techniques used to compute the integrals required for equations (4.76) and (4.81). First of all, (4.81) may be written as

$$\phi(\underline{r}) = \frac{1}{\varepsilon_o} \int_S \sigma_s(\underline{r}') \, P(\underline{r}|\underline{r}')ds' + \frac{1}{\varepsilon_o} \int_S \sigma_s(r) \, Q(\underline{r}|\underline{r}')ds' \, . \qquad (4.82)$$

Since $P(\underline{r}|\underline{r}')$ is regular, the first integral of (4.82) is included in integrals of the form

$$I = \int_S \sigma_s(\underline{r}) \, f(\underline{r})ds \qquad (4.83)$$

where f is a regular function. Note that when

$$f(\underline{r}) = \int_{S_i} G(\underline{r}|\underline{r}')ds' \qquad (4.84)$$

the first integral of (4.76) is included, and that when f=1 the right side of (4.76) is included.

The addition-subtraction technique is applied over each of the three regions, S_a, S_b, and S_c (see Fig. 4.8) and the contributions are added. Discussion of the integration over S_c is sufficient to demonstrate the technique; the other two regions are handled in a similar fashion:

$$I = \int_{S_c} \sigma_s(\underline{r}) f(\underline{r})ds = \int_0^1 \frac{1}{(1-y)^\alpha} \int_2^4 \frac{f(x,y)}{(4-x)^\alpha} \, dxdy$$

$$= \int_0^1 \frac{1}{(1-y)^\alpha} \{g(y)-g(1)\}dy + g(1) \int_0^1 \frac{dy}{(1-y)^\alpha} \qquad (4.85)$$

where

$$g(y) = \int_2^4 \frac{f(x,y)dx}{(4-x)^{\alpha}} = \int_2^4 \frac{1}{(4-x)^{\alpha}} \{f(x,y)-f(4,y)\}dx + f(4,y) \int_2^4 \frac{dx}{(4-x)^{\alpha}} . \tag{4.86}$$

Gaussian quadrature of order 6 is used for the numerical integrations, and the analytic integrals are simply obtained.

The second integral of (4.76) is somewhat more complicated. First of all, it may be written in the form

$$\int_S \sigma_s(\underline{r}) \int_S \sigma_s(\underline{r}') \, G(\underline{r}|\underline{r}')ds'ds = \int_S \sigma_s(\underline{r}) \, f(\underline{r}) \, ds \tag{4.87}$$

where

$$f(\underline{r}) = \int_S \sigma_s(\underline{r}') \, G(\underline{r}|\underline{r}') \, ds' . \tag{4.88}$$

The function $f(\underline{r})$ is regular, so that the single integral in (4.87) may be computed as (4.83). One may identify (4.88) with (4.81) and (4.82). Therefore only the integral of $\sigma_s Q$ is required, and once again, it is sufficient to consider the integration over S_c.

If $\underline{r}$ does not lie in the plane of S, the perpendicular projection point $\underline{r}''$ is found and

$$f(\underline{r}) = \int_{S_c} \sigma_s(\underline{r}') \, \{G(\underline{r}|\underline{r}') - Q(\underline{r}''|\underline{r}')\}ds' + \int_{S_c} \sigma_s(\underline{r}') \, Q(\underline{r}'|\underline{r}'')ds' . \tag{4.89}$$

The first integral is computed as (4.83). If the projection point $\underline{r}''$ does not lie in S_c, Q is treated as being regular. If $\underline{r}''$ does lie within S_c, the integral is transformed as follows:

$$I = \int_0^1 \frac{1}{(1-y')^{\alpha}} \int_0^4 \frac{1}{(4-x')^{\alpha}\sqrt{(x''-x')^2+(y''-y')^2}}\, dx'dy'$$

$$= \int_{-y''}^{1-y''} \frac{1}{(1-y''-v)^{\alpha}} \int_{2-x''}^{4-x''} \frac{1}{(4-x''-u)^{\alpha}\sqrt{u^2+v^2}}\, du\, dv \quad . \tag{4.90}$$

This transformation simply shifts the origin to (x",y"), the singular point of Q. Now, (4.90) may be evaulated as the appropriate sum of 4 integrals of the form

$$I_1 = \int_0^k \int_0^b f(u,v)\, du\, dv = \int_0^k \int_0^{\frac{bv}{k}} f\, du\, dv + \int_0^b \int_0^{\frac{ku}{v}} f\, dv\, du \tag{4.91}$$

similarly to (4.37). In this expression

$$f(u,v) = \frac{1}{(1-y''-v)^{\alpha}(4-x''-u)^{\alpha}\sqrt{u^2+v^2}} \quad . \tag{4.92}$$

Consider the first integral of (4.91). The addition-subtraction technique is used to write

$$I_2 = \int_0^k \int_0^{\frac{bv}{k}} f\, du\, dv = \int_0^k \frac{1}{(1-y''-v)^{\alpha}}\{g(v)-g(1-y'')\}dv + g(1-y'')\int_0^k \frac{dv}{(1-y''-v)^{\alpha}} \tag{4.93}$$

and g is obtained as

$$g(v) = \int_0^{\frac{bv}{k}} \frac{1}{(4-x''-u)^{\alpha}}\left\{\frac{1}{\sqrt{u^2+v^2}} - \frac{1}{\sqrt{(4-x'')^2+v^2}}\right\}du + \frac{1}{\sqrt{(4-x'')^2+v^2}} \int_0^{\frac{bv}{k}} \frac{du}{(4-x''-u)^{\alpha}} \quad . \tag{4.94}$$

The first integrals of (4.93) and (4.94) are computed using order 6 Gaussian quadrature, and the analytic integrals are readily obtainable. This procedure is applicable when $\underline{r}''$ lies on the edge of S_c, that is when $x''=4$ or $y''=1$. In this case however, integrals at (4.91) which are over zero area are not computed. The algorithms for integration over S_a and S_b are much the same.

4.6 Axially Symmetric Problems

Consider the three dimensional problem shown in Fig. 4.12. The (η,ψ,z) cylindrical coordinate system is used. If ϕ and σ are independent of the angle ψ the problem is axially (or axi-) symmetric. It is possible to eliminate the angle ψ from the IE. For example, the Dirichlet equation

$$\frac{1}{\varepsilon_o}\int_S \sigma(\underline{r}')\, G(\underline{r}|\underline{r}')ds' = g(\underline{r}) \tag{4.95}$$

where S is a surface in 3-D space, becomes

$$\frac{1}{\varepsilon_o}\int_S \eta'\sigma(\eta',z')\int_0^{2\pi} \frac{d\psi\quad ds'}{4\pi\sqrt{\eta^2+\eta'^2+(z-z')^2-2\eta\eta'\cos\psi}} = g(\eta,z)\ . \tag{4.96}$$

Here S is a curve in the (η,z) plane. The integral over ψ may be

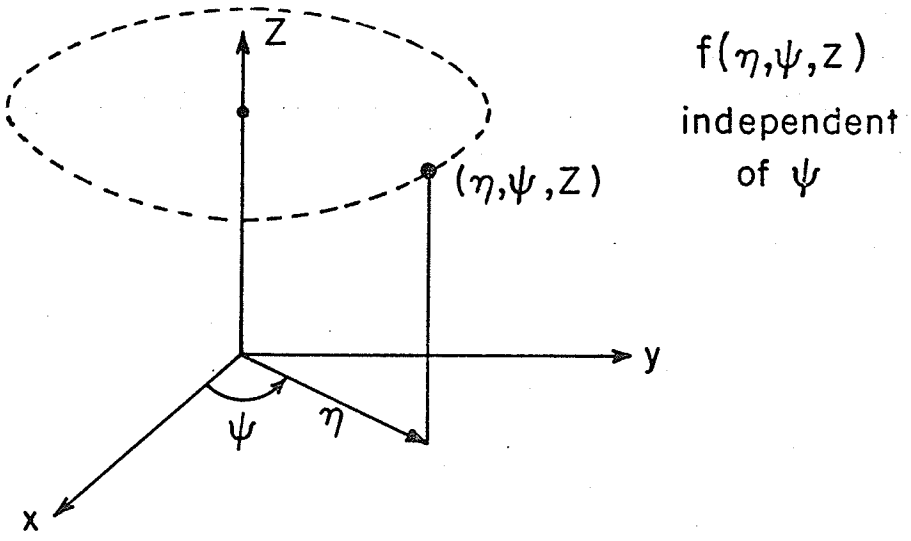


Fig. 4.12 The Axisymmetric Geometry

identified as a complete elliptic integral of the first kind K, [27,p.30], whence (4.96) becomes

$$\frac{1}{\varepsilon_o}\int_S \eta'\sigma(\eta',z')\ \frac{K(\xi)}{\pi\chi}\,ds' = g(\eta,z) \quad . \tag{4.97}$$

The argument of K is given by

$$\xi^2 = 4\eta\eta'/\chi^2 \tag{4.98}$$

and

$$\chi^2 = (\eta+\eta')^2+(z-z')^2 \quad . \tag{4.99}$$

The functional (4.15) involves a further integral over the original surface in three-dimensional space. It becomes

$$F(\sigma) = \frac{2\pi}{\varepsilon_o}\int_S \eta\sigma(\eta,z)\int_S \eta'\sigma(\eta',z')\ \frac{K(\xi)}{\pi\chi}\,ds'ds - 4\pi\int_S \eta\ g(\eta,z)ds \tag{4.100}$$

where, it is stressed, S is a curve in the (η,z) plane, $\eta \geq 0$.

The Green's function singularity now occurs as $\xi\to 1$. The small argument form for K [28, p.11] may be used as follows:

$$\lim_{\xi\to 1} K(\xi) = \ln 4 - \ln\sqrt{1-\xi^2} \quad . \tag{4.101}$$

The distance between the two points $\underline{r}$ and $\underline{r}'$ satisfies

$$d^2 = (\eta-\eta')^2 +(z-z')^2 = (\eta+\eta')^2+(z-z')^2-4\eta\eta'$$

$$= \chi^2(1-\xi)^2 \quad . \tag{4.102}$$

Therefore

$$\ln 4 - \ln\sqrt{1-\xi^2} = \ln 4\chi - \ln|\underline{r}-\underline{r}'| \quad . \tag{4.103}$$

Note as well, from (4.88), that when $|\underline{r}-\underline{r}'|\to 0$, $\chi=2\eta'=2\eta$. Consider the limit

$$\lim_{|\underline{r}-\underline{r}'|\to 0} \frac{\eta' K(\xi)}{\pi\chi} = \frac{\eta'}{2\pi\eta'} (\ln 8\eta' - \ln|\underline{r}-\underline{r}'|) = \frac{-1}{2\pi} \ln|\underline{r}-\underline{r}'| + \frac{\ln 8\eta'}{2\pi} \tag{4.104}$$

Provided that $\eta' \neq 0$, the Green's function here behaves like the two-dimensional logarithmic Green's function as $|\underline{r}-\underline{r}'| \to 0$.

Consider now the integration of (4.86) where σ is a regular function

$$g(\eta,z) = \int_S \left[\sigma(\eta',z') - \sigma(\eta,z)\right] \frac{\eta' K(\xi)}{\pi\chi} \, ds' + \sigma(\eta,z) \int_S \frac{\eta' K(\xi)}{\pi\chi} \, ds' . \tag{4.105}$$

The first integrand is regular, and the second integral is obtained by using the addition-subtraction technique:

$$\int_S \frac{\eta' K(\xi)}{\pi\chi} \, ds' = \int_S \left(\frac{\eta' K(\xi)}{\pi\chi} + \frac{1}{2\pi} \ln|\underline{r}-\underline{r}'|\right) ds' - \frac{1}{2\pi} \int_S \ln|\underline{r}-\underline{r}'| \, ds' . \tag{4.106}$$

The last integral has been integrated analytically, at (4.26) and the first integrand is regular, with the value $(\ln 8\eta')/2\pi$ (4.104) when $\underline{r}=\underline{r}'$.

As a first example, consider the sphere, with radius 0.5 and Dirichlet condition $\phi=2$, shown in Fig. 4.13a. The analytic solution is $\phi=1/r$; $r \geq 0.5$. The solution may be obtained using pulse functions over N (straight) line elements as shown in Fig. 4.13b, with the image Green's function (G=P+Q, defined using the elliptic integrals (4.97)) used to cater for the symmetry at z=0. The capacitance of the configuration, with $\varepsilon_o=1$, is π. Order 3 Gaussian quadrature was used for the numerical integrations, and the results are presented in Table 4.7. Convergence is smooth and monotonic, with much of the residual error due probably to using a polygon to model a circle.

As a second example, consider the doughnut shown in Fig. 4.14a. The cross-section of the doughnut is a sphere of radius 1 centered at

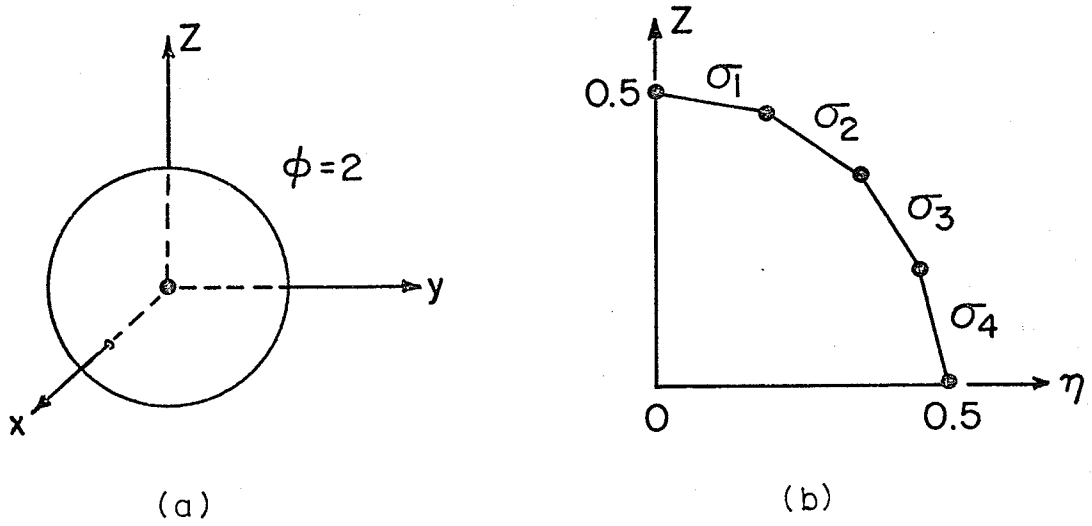


Fig. 4.13 The Sphere, Radius 0.5
a) 3-D Geometry
b) IE Finite Elements, N=4

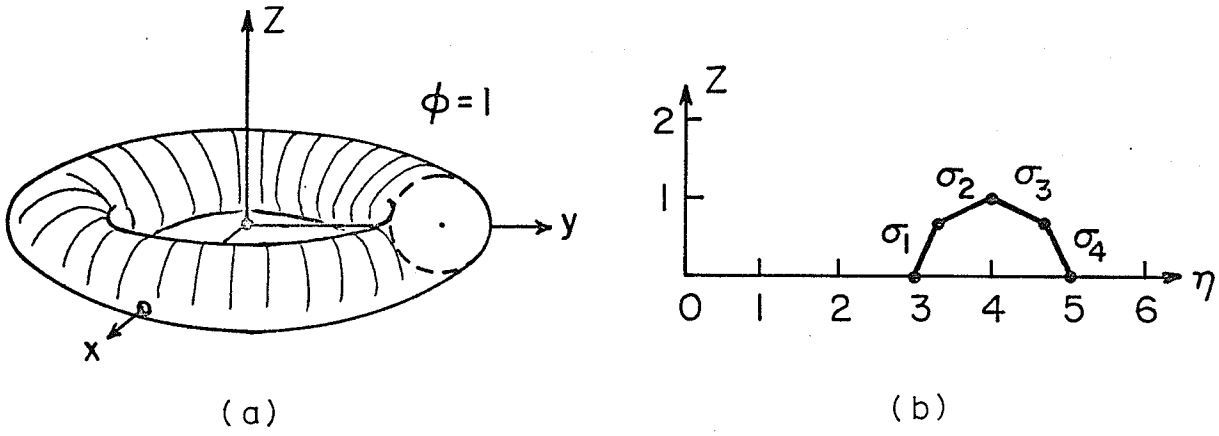


Fig. 4.14 The Doughnut Problem
a) 3-D Geometry
b) IE Finite Elements, N=4

TABLE 4.7 THE SPHERE; AXISYMMETRIC IE SOLUTION

Number of pulses N	Capacitance (norm)	Potential ϕ at r		
		r=1	r=2	r=3
5	3.13644	0.994250	0.497124	0.331416
10	3.14004	0.998540	0.499270	0.332846
20	3.14114	0.999640	0.499820	0.333213
30	3.14137	0.999844	0.499922	0.333281
40	3.14146	0.999900	0.499955	0.333303
50	3.14151	0.999926	0.499965	0.333310
exact	3.14159	1.0	0.5	0.333333
|% error| (N=50)	0.0026	0.0074	0.0070	0.0069

(4,0), as shown in Fig. 4.14b. Symmetry about the z=0 plane permits a solution to be sought only in the positive quadrant, and N line elements with pulses are placed in the positive quadrant of the (η,z) plane as shown. The results displayed in Table 4.8 were obtained using order 3 Gaussian quadrature, and, once again, monotonic convergence is demonstrated. Equipotentials will be shown later, in Fig. 6.7 .

TABLE 4.8 THE DOUGHNUT; AXISYMMETRIC IE SOLUTION

Number of pulses	Potential ϕ at (r,0)				
	r=0	r=2	r=3	r=6	r=8
15	0.862430	0.913497	0.999161	0.734389	0.505740
20	0.862621	0.913676	0.999399	0.734686	0.505943
25	0.862709	0.913758	0.999534	0.734821	0.506092
30	0.862756	0.913802	0.999620	0.734894	0.506171
35	0.862783	0.913825	0.999709	0.734932	0.506227
40	0.862792	0.913840	0.999772	0.734964	0.506264

4.7 Conclusions

Fredholm integral equations of the first kind with positive kernels may be solved using the Rayleigh-Ritz procedure with finite elements. For electrostatic problems the kernel - the Green's function - is positive-definite (or for two-dimensional problems, can be made so by including a suitable constant if the surface S is of finite extent). Convergence is in electrostatic energy.

The singularity in G is conveniently handled by the addition-subtraction technique, which produces a regular integrand, permitting numerical integration, and a singular integrand which is integrated analytically. The procedure is general purpose in nature and is useful in handling singularities in σ, the source distribution, as well.

The surface S (a curve in the plane for two-dimensional problems) is arbitrary; it need not be a special surface. This is an advantage over procedures where S is restricted to be an infinite plane, for example. The finite-elements used are linear elements (plane-rectangular or straight-line segments) and this introduces some modelling error for arbitrary surfaces. But the technology of finite-elements is available, and is applicable here. A useful extension of the work reported would be to employ curved elements, and the isoparametric technique of Section 2.6 might be the logical approach to use. The greatest difficulty would probably be in computing the analytic integrals in the local element, for the Green's function involve distances which will not, in general, be mapped linearly between the local and global elements.

The use of special singular functions to handle known source behaviour near edges and corners results in improved convergence properties

at the expense of more involved computation. This expense may be reduced by introducing the special functions independently, with separate variational parameters. Singular behaviour near corners is not as well understood as it is near an edge. Some useful work might well be done in determining what special function(s) should be introduced for corner singularities. The results obtained here, and those reported by Silvester and Benedek [22] indicate that special treatment is required.

The concept of mutual constraint has been introduced as an alternative way of constructing the Rayleigh-Ritz matrix equation. One treats each element separately, assuming all external effects to be known - they are taken to the right hand side of the equation. After applying the variational procedure to the element, the actually unknown contributions from the other elements are returned to the left side of the matrix equation to permit simultaneous solution. The PDE finite-element procedure described in Chapter II and used in Chapter III may be thought of similarly: the PDE elements are treated independently and then are mutually constrained to produce continuous potential across the inter-element boundaries. Those examples, and the IE examples presented in this chapter have one thing in common: each finite-element in the system possesses the same operator, and the mutual constraint equation coincides with the Rayleigh-Ritz matrix equation for the whole problem.

One advantage of the PDE element technique is that elements may be placed where desired for best representation. Perhaps it would be useful to use the IE elements in the same fashion, concentrating them, for example, near corners, to compensate for singular behaviour. The use here of identical square pulse function elements for all source representation (except for special singular functions) is by no means optimum.

4.8 References

[1] D.S. Jones, The Theory of Electromagnetism. Oxford: Pergamon, 1964.

[2] F.G. Tricomi, Integral Equations. New York: Interscience, 1957.

[3] D. Gish, "Characteristic Impedance and phase velocity of a dielectric supported air strip transmission line with sidewalls." IEEE Trans. MTT-18, pp.131-148, March 1970.

[4] J.L. Allen and M.F. Estes, "Broadside coupled strips in a layered dielectric medium." IEEE Trans. MTT-20,pp.662-668, October 1972.

[5] M. Maeda, "An analysis of gap in microstrip transmission lines." IEEE Trans. MTT-20, pp.390-396, June 1972.

[6] H. Levine and J. Schwinger, "On the theory of diffraction by an aperture in an infinite plane screen." Phys. Rev., Vol. 74, pp. 958-974, 1948.

[7] R.E. Collin, Field Theory of Guided Waves, New York: McGraw-Hill,1960.

[8] E. Yamashita and R. Mittra, "Variational Method for the Analysis of microstrip lines." IEEE Trans. MTT-16, pp.251-256, April 1968.

[9] P.D. Crout, "An application of polynomial approximation to the solution of integral equations arising in physical systems." Jnl. Math. and Phys., Vol. 19, pp.34-92, 1940.

[10] P.D. Crout and F.B. Hildebrand, "A least squares procedure for solving integral equations by polynomial approximation." Jnl. Math and Phys., Vol. 20, pp.310-335, 1941.

[11] R.F. Harrington, Field Computation by Moment Methods. New York: Macmillan, 1968.

[12] B.E. Spielman and R.F. Harrington, "Waveguides of arbitrary cross-section by solution of a nonlinear integral eigenvalue problem." IEEE Trans. MTT-20, pp.578-585, September 1972.

[13] P. Silvester and M.S. Hsieh, "Projective solution of integral equations arising in electric and magnetic field problems." Int. Jnl. Comput. Phys., Vol. 8, pp.73-82, 1971.

[14] P. Silvester and P. Benedek, "Electrostatics of the microstrip-revisited." IEEE Trans. MTT-20, pp.756-758, November 1972.

[15] K. Pontoppidan, "Equivalent source solutions for electrostatic and magnetostatic problems." IEEE Int. Conv. Digest, New York, pp.620-621, March 1971.

[16] B. Noble, A Bibliography on: Methods for Solving Integral Equations. University of Wisconsin, Mathematics Research Centre, MRC 1176, September 1971.

[17] G.F. Roach, Green's Functions - Introductory Theory with Applications. London: Van Nostrand Reinhold, 1970.

[18] S. Bergman, Integral Operators in the Theory of Partial Differential Equations. New York: Springer-Verlag, 1969

[19] O.D. Kellogg, Foundations of Potential Theory. Berlin: Springer-Verlag, 1929.

[20] S.G. Mikhlin, Integral Equations (and their applications to certain problems in mechanics, mathematical physics and technology). New York: Macmillan, 1964.

[21] S.G. Mikhlin, Linear Integral Equations. Delhi; Hindustan Publ. Corp., 1960.

[22] F.S. Acton, Numerical Methods That Work. New York: Harper and Row, 1970.

[23] M. C. Décreton, "Treatment of singularities in the finite-element method." M.Sc Thesis, University of Manitoba, Winnipeg, 1972.

[24] P. Silvester and P. Benedek, "Microstrip discontinuity capacitances for right angle bends, T junctions and crossings." IEEE Trans. MTT-21, pp.341-346, May 1973.

[25] J.N. Franklin, Matrix Theory. Englewood Cliffs, New Jersey: Prentice-Hall, 1968.

[26] W. Braunbek, "On the diffraction field near a plane screen corner." IEEE Trans.AP-4, pp.219-223, April 1956.

[27] P. Silvester, Modern Electromagnetic Fields. Englewood Cliffs, N.J., Prentice-Hall, 1968.

[28] P.F. Byrd and M.D. Friedman, Handbook of Elliptical Integrals for Engineers. Berlin: Springer, 1954.

CHAPTER V

INTEGRAL VARIATIONAL SOLUTION OF INTERFACE PROBLEMS

Bryant and Weiss [1] in 1969 presented a point-matching technique for solving the air-dielectric interface problem in open microstrip based on an IE formulation using the static free-space Green's function. In this chapter the IE variational with finite-elements is applied to similar problems. Whereas the earlier method was presented for an infinite planar interface, the method described here is applicable to arbitrary surfaces, provided that these surfaces may be adequately represented by rectangular finite-elements (or line elements in two-dimensional problems). This method involves the mutual constraint of two different variational forms, and the constrained form is not self-adjoint. Nevertheless it contains a positive-definite, self-adjoint component, and the theorem given by Mikhlin (Section 2.3) guarantees energy convergence.

Benedek and Silvester [2] employ a Galerkin method to solve the IE problem for an infinity extending interface but use a special interface-cum-mirror image Green's function which is expressed as an infinite series. A similar approach is described by Farrar and Adams [3] who use a point-matching technique. The nature of the special Green's function appears to limit these methods to infinite, planar interfaces. An interesting approach is described by Itoh, Mittra, and Ward [4] for this problem: they apply Galerkin's method in the Fourier Transform domain and in this case the special Green's function is deterministic, rather than being expressed as an infinite series.

There are several recent reports available [5-8] wherein variational techniques dating from Levine and Schwinger [9] are used to determine properties of shielded microstrip. There, the Green's functions are constructed from the waveguide eigenmodes. Moment methods have been reported as well for solution of such problems by Mittra, Itoh and Li [10] and Minor and Bole [11]. Thong [12] and Chow and Wu [13] use moment methods with dyadic Green's functions. These methods are really not applicable here; the goal here is to solve unbounded problems using the simple free-space Green's functions.

5.1 The Interface Functional

Consider the problem shown in Fig. 5.1. The region Ω_1 is a dielectric (ε_r) and Ω_2 is free-space and extends to infinity. The surface S_I separating Ω_1 and Ω_2 is the interface. Let $\hat{n}$ be drawn from Ω_1 to Ω_2, suppose that a free source h resides on S_I, and that potential vanishes at infinity.

The PDE functional for the static problem is (from (2.48) with W=0):

$$F(\phi) = \varepsilon_r\varepsilon_o \int_{\Omega_1} \nabla\phi\cdot\nabla\phi \, d\Omega + \varepsilon_o \int_{\Omega_2} \nabla\phi\cdot\nabla\phi \, d\Omega - 2 \int_{S_I} h\,\phi \, ds \quad . \qquad (5.1)$$

The first variation is

$$\delta F(\phi) = 2 \int_{\Omega_1} \delta\phi(-\varepsilon_r\varepsilon_o\nabla^2\phi)d\Omega + 2\int_{\Omega_2} \delta\phi(-\varepsilon_o\nabla^2\phi)_{d\Omega} + 2\varepsilon_o\int_{S_I} \delta\phi\left(\varepsilon_r\frac{\partial\phi}{\partial n_-} - \frac{\partial\phi}{\partial n_+} - \frac{h}{\varepsilon_o}\right)ds \quad . \qquad (5.2)$$

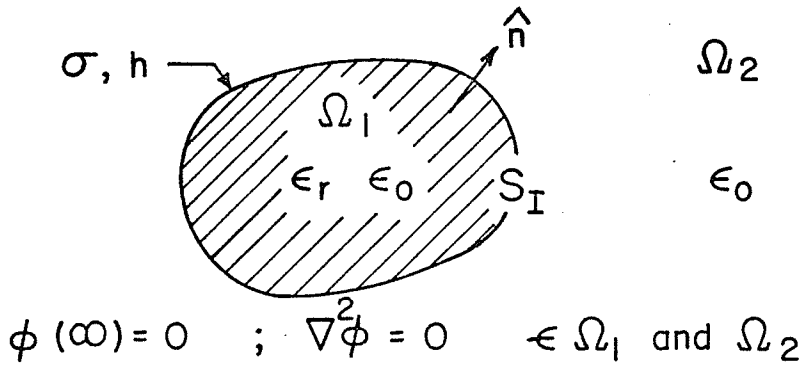


Fig. 5.1 The Interface Problem

At the minimum the Laplacian is satisfied throughout and on the interface S_I

$$\varepsilon_r \frac{\partial\phi}{\partial n_-} - \frac{\partial\phi}{\partial n_+} = \frac{h}{\varepsilon_o} \quad . \tag{5.3}$$

If there is no free-source on S_I (h=0) the normal interface condition prevails (2.68) but in the absence of any other source, and with ϕ=0 at infinity, ϕ=0 throughout.

Now suppose that potential everywhere is to be produced by a source σ on S_I:

$$\phi(\underline{r}) = \frac{1}{\varepsilon_o} \int_{S_I} \sigma(\underline{r}')\, G(\underline{r}|\underline{r}')\, ds' \quad . \tag{5.4}$$

Substituting this directly into the first term of (5.1) yields

$$\bar{I} = \varepsilon_r\varepsilon_o \int_{\Omega_1} \nabla\phi\cdot\nabla\phi \; d\Omega$$

$$= \frac{\varepsilon_r\varepsilon_o}{\varepsilon_o^2} \int_{\Omega_1} \left\{ \int_{S_I} \sigma(\underline{r}')\nabla G(\underline{r}|\underline{r}')ds' \right\}\cdot\left\{ \int_{S_I} \sigma(\underline{r}'')\nabla\, G(\underline{r}|\underline{r}'')ds'' \right\} d\Omega \tag{5.5a}$$

$$= \frac{\varepsilon_r}{\varepsilon_o} \int_{S_I} \sigma(\underline{r}') \int_{S_I} \sigma(\underline{r}'') \left\{ \int_{\Omega_1} \nabla G(\underline{r}|\underline{r}')\cdot\nabla G(\underline{r}|\underline{r}'')\; d\Omega \right\} ds''\, ds'. \tag{5.5b}$$

Green's first identity [14, p.12] may be used to show that the integral over Ω_1 in (5.5b) is equal to

$$-\int_{\Omega_1} G(\underline{r}|\underline{r}')\nabla^2 G(\underline{r}|\underline{r}'')d\Omega \; + \lim_{\underline{r}',\underline{r}''\to S_I} \int_{S_I} G(\underline{r}|\underline{r}')\, \frac{\partial G(\underline{r}|\underline{r}'')}{\partial n}\, ds \quad .$$

For the Laplacian $\nabla^2 G = -\delta$ (see equation (1.9)) and the limits for the second integral involve a contribution as at eqation (3.8b). Therefore

this expression is equal to

$$\frac{1}{2}\, G(\underline{r}'|\underline{r}'') + \int_{S_I} G(\underline{r}|\underline{r}')\, \frac{\partial G(\underline{r}|\underline{r}'')}{\partial n}\, ds \quad .$$

It is symmetric in $\underline{r}', \underline{r}''$, and this is easily shown by exchanging superscripts before applying Green's first identity. Therefore the integral in (5.5b) becomes

$$I = \frac{\varepsilon_r}{\varepsilon_o} \int_{S_I} \int_{S_I} \sigma(\underline{r}')\sigma(\underline{r}'') \left\{ \frac{G(\underline{r}'|\underline{r}'')}{2} + \int_{S_I} G(\underline{r}|\underline{r}')\, \frac{\partial G(\underline{r}|\underline{r}'')ds}{\partial n} \right\} ds''ds' \,. \tag{5.6}$$

Substitution of (5.4) into the second term of the functional (5.1) produces the same result, except that the sign of the integral of $G\partial G/\partial n$ changes, as $\hat{n}$ points into Ω_2. The functional (5.1) may now be written as

$$F(\sigma) = \frac{\varepsilon_r+1}{2\varepsilon_o} \int_{S_I} \int_{S_I} \sigma(\underline{r}')\sigma(\underline{r}'')G(\underline{r}'|\underline{r}'')ds''ds' + \frac{\varepsilon_r-1}{\varepsilon_o} \int_{S_I} \int_{S_I} \sigma(\underline{r}')\sigma(\underline{r}'')$$

$$\times \int_{S_I} G(\underline{r}\,|\underline{r}')\, \frac{\partial G(\underline{r}|\underline{r}'')}{\partial n}\, ds\, ds''ds' - \frac{2}{\varepsilon_o} \int_{S_I} \sigma(\underline{r}') \int_{S_I} h(\underline{r})G(\underline{r}|\underline{r}')ds\, ds' \,. \tag{5.7}$$

This functional may be minimized to find σ. The first variation is

$$\delta F(\sigma) = 2 \int_{S_I} \delta\sigma(\underline{r}') \int_{S_I} \frac{G(\underline{r}'|\underline{r}'')}{\varepsilon_o}\, f(\underline{r}'')\, ds''ds' \tag{5.8}$$

where

$$f(\underline{r}'') = \frac{\varepsilon_r+1}{2}\, \sigma(\underline{r}'') + (\varepsilon_r-1) \int_{S_I} \sigma(\underline{r})\, \frac{\partial G(\underline{r}|\underline{r}'')}{\partial n''}\, ds \tag{5.9}$$

and a change in superscript has been employed to enable G to be factored out. At the stationary point, from (5.8) it is seen that $\int fG=0$. But this is a positive-definite integral operator, so that $f=0$ is the only solution. (See Section 2.1). Using equations (3.11) the result, from (5.9) with $f=0$ is

$$\varepsilon_r \frac{\partial\phi}{\partial n_-} - \frac{\partial\phi}{\partial n_+} = \frac{h}{\varepsilon_o} \text{ on } S_I \quad . \tag{5.10}$$

Therefore the field produced by σ which makes the IE functional (5.3) stationary is precisely the field ϕ which makes the PDE functional (5.1) stationary - at the unique solution point of the problem.

The second variation of (5.7) may be written as

$$\delta^2 F(\sigma) = \frac{2}{\varepsilon_o} \int_{S_I} \int_{S_I} \delta\sigma(\underline{r}')G(\underline{r}'|\underline{r}'')\{\frac{\varepsilon_r+1}{2}\,\delta\sigma(\underline{r}'')+(\varepsilon_r-1)\int_{S_I}\delta\sigma(\underline{r})\frac{\partial G(\underline{r}|\underline{r}'')}{\partial n''}ds\}ds'ds'' . \tag{5.11}$$

Using (3.11) this is the same as

$$\delta^2 F(\sigma) = 2\varepsilon_o \int_{S_I} \delta\phi(\underline{r}'')\{\varepsilon_r \frac{\partial\delta\phi\ (\underline{r}'')}{\partial n_-} - \frac{\partial\delta\phi\ (\underline{r}'')}{\partial n_+}\}\ ds'' \tag{5.12}$$

which, using Green's first identity and the fact that ϕ vanishes at infinity, is the same as

$$\delta^2 F(\sigma) = \varepsilon_o\varepsilon_r \int_{\Omega_1} \nabla\delta\phi\cdot\nabla\delta\phi d\Omega + \varepsilon_o \int_{\Omega_2} \nabla\delta\phi\cdot\nabla\delta\phi\ \ d\Omega\ \geq 0 \quad . \tag{5.13}$$

The operator is positive-definite.

One may note that the source h on S_I generates a potential which may be viewed as a Dirichlet condition on S_I :

$$g(\underline{r}') = \frac{1}{\varepsilon_o} \int_{S_I} h(\underline{r})\ G(\underline{r}|\underline{r}')\ ds \quad . \tag{5.14}$$

Now, if $\varepsilon_r=1$ in (5.7), and (5.14) is used,

$$F(\sigma) = \frac{1}{\varepsilon_o} \int_{S_I} \int_{S_I} \sigma(\underline{r}')\sigma(\underline{r}'')\ G(\underline{r}'|\underline{r}'')ds''ds' - 2 \int_{S_I} \sigma(\underline{r}')g(\underline{r}')ds' . \tag{5.17}$$

This is identical to the IE functional for the Dirichlet problem (4.15). It is interesting that the IE and PDE functionals may be derived one from the other, but not surprising when one realizes that the IE and PDE are but two ways of looking at the same problem.

The computation is complicated by the presence of triple integrals, but no new singularities are introduced, and algorithms for computation will be presented in due course.

In this presentation there has been no restriction on the shape of S_I; it must be continuous and closed, of course, if it is to be the boundary of a dielectric object with finite dimensions. It is also necessary that ε_r be constant throughout the dielectric material. Otherwise ε_r cannot be removed from the integral over Ω_1; that is, equation (5.5) cannot be written as shown, and the derivation following (5.5) is invalid.

5.2 The Interface in the Presence of a Conductor

Suppose that a surface S is placed in the free space region, with a Dirichlet condition $\phi=g$. This could, for example, be a conductor (and a perfect conductor if g is constant). Suppose that sources σ and σ_I are placed on S and S_I, respectively, as shown in Fig. 5.2. Potential everywhere is to be given by σ and σ_I:

$$\phi(\underline{r}) = \frac{1}{\varepsilon_o}\int_S \sigma(\underline{r}')G(\underline{r}|\underline{r}')ds' + \frac{1}{\varepsilon_o}\int_{S_I} \sigma_I(\underline{r}')G(\underline{r}|\underline{r}')ds' \quad . \tag{5.18}$$

If σ_I were known, the Dirichlet problem could be solved by applying the IE functional directly:

$$F(\sigma) = \frac{1}{\varepsilon_o}\int_S\int_S \sigma(\underline{r})\sigma(\underline{r}')G(\underline{r}|\underline{r}')ds'ds - 2\int_S \sigma(\underline{r})g_s(\underline{r})ds \tag{5.19}$$

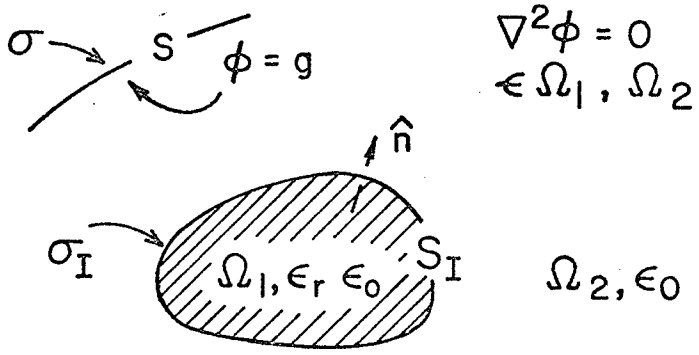


Fig. 5.2 The Interface S_I with a Conductor S.

where

$$g_s(\underline{r}) = g(\underline{r}) - \int_{S_I} \sigma_I(\underline{r}')G(\underline{r}|\underline{r}')ds' \; ; \; \underline{r} \text{ on } S \tag{5.20}$$

and is completely known.

At the minimum of (5.19)

$$\frac{1}{\varepsilon_o}\int_S \sigma(\underline{r}')G(\underline{r}|\underline{r}')ds' = g_s(\underline{r}) \; ; \; \underline{r} \text{ on } S \; . \tag{5.21}$$

Equation (5.20) may be substituted now into (5.21) and the two charge distributions, σ and σ_I, produce a potential satisfying the Dirichlet condition on S.

Now consider σ to be known. This charge produces a potential, say ϕ_s, whose normal derivatives at S_I (for S and S_I not coincident) are, from (3.11),

$$\frac{\partial\phi_s(\underline{r})}{\partial n_-} = \frac{1}{\varepsilon_o}\int_S \sigma(\underline{r}'')\frac{\partial G(\underline{r}|\underline{r}'')}{\partial n}ds'' \; ; \; \underline{r} \text{ on } S_I \tag{5.22a}$$

and

$$\frac{\partial\phi_s(\underline{r})}{\partial n_+} = \frac{1}{\varepsilon_o}\int_S \sigma(\underline{r}'')\frac{\partial G(\underline{r}|\underline{r}'')}{\partial n}ds'' \; ; \; \underline{r} \text{ on } S_I \; . \tag{5.22b}$$

Consider the source h_s defined by

$$\frac{-h_s(\underline{r})}{\varepsilon_o} = \varepsilon_r \frac{\partial\phi_s}{\partial n_-} - \frac{\partial\phi_s}{\partial n_+} = \frac{(\varepsilon_r - 1)}{\varepsilon_o}\int_S \sigma(\underline{r}'')\frac{\partial G(\underline{r}|\underline{r}'')}{\partial n}ds''; \; \underline{r} \text{ on } S_I \; . \tag{5.23}$$

If this source is used in the IE interface functional (5.7), the result will be a complete potential satisfying the interface condition. Consider that the resulting σ_I will produce a potential ϕ_I satisfying (5.10); that is,

$$\varepsilon_r \frac{\partial\phi_I}{\partial n_-} - \frac{\partial\phi_I}{\partial n_+} = \frac{h_s(\underline{r})}{\varepsilon_o} = -\varepsilon_r \frac{\partial\phi_s}{\partial n_-} + \frac{\partial\phi_s}{\partial n_+} \; . \tag{5.24}$$

But the total potential ϕ is the sum of ϕ_s and ϕ_I. Therefore

$$\varepsilon_r \frac{\partial \phi}{\partial n_-} - \frac{\partial \phi}{\partial n_+} = 0 \; ; \; \underline{r} \text{ on } S_I \tag{5.25}$$

and the interface condition, with no free charge on S_I, is satisfied. Thus a known source σ on S creates an effective source h_s on the interface which may be used in the IE interface functional over S_I alone to determine σ_I.

Suppose, for simplicity, that S and S_I each consist of one finite-element, and let σ and σ_I be constants. The Rayleigh-Ritz equation for σ on S, from (5.19), is

$$\sigma \frac{1}{\varepsilon_o} \int_S \int_S G(\underline{r}|\underline{r}')ds'ds = \int_S g_s(\underline{r})ds \tag{5.26a}$$

or

$$A\sigma = b_1' \; . \tag{5.26b}$$

The Rayleigh-Ritz equation for σ_I on S_I, from (5.7) is

$$\sigma_I \left\{ \frac{\varepsilon_r+1}{2\varepsilon_o} \int_{S_I} \int_{S_I} G(\underline{r}'|\underline{r}'')ds''ds' + \frac{\varepsilon_r-1}{\varepsilon_o} \int_{S_I} \int_{S_I} \int_{S_I} G(\underline{r}|\underline{r}') \frac{\partial G(\underline{r}|\underline{r}'')}{\partial n} dsds''ds' \right\}$$

$$= \frac{1}{\varepsilon_o} \int_{S_I} \int_{S_I} h_s(\underline{r})G(\underline{r}|\underline{r}')ds'ds \tag{5.27a}$$

or

$$D\,\sigma_I = b_2' \; . \tag{5.27b}$$

Were g_s and h_s known, the individual equations would satisfy the matrix equation

$$\begin{bmatrix} A & 0 \\ 0 & D \end{bmatrix} \begin{pmatrix} \sigma \\ \sigma_I \end{pmatrix} = \begin{pmatrix} b_1' \\ b_2' \end{pmatrix} \; . \tag{5.28}$$

However, (5.20) and (5.23) must be used to construct the mutual constraint matrix equation:

$$\begin{pmatrix} b_1' \\ b_2' \end{pmatrix} = \begin{pmatrix} b_1 \\ 0 \end{pmatrix} - \begin{bmatrix} 0 & A' \\ D' & 0 \end{bmatrix} \begin{pmatrix} \sigma \\ \sigma_I \end{pmatrix} \tag{5.29}$$

which, when substituted into (5.28) yields

$$\frac{1}{\varepsilon_o} \left[\begin{array}{c|c} \displaystyle\iint_S \iint_S G(\underline{r}|\underline{r}')ds'ds & \displaystyle\iint_S \iint_{S_I} G(\underline{r}|\underline{r}')ds'ds \\ \hline (\varepsilon_r-1)\displaystyle\iint_S \iint_{S_I} \int_{S_I} G(\underline{r}|\underline{r}')\frac{\partial G(\underline{r}|\underline{r}'')}{\partial n}dsds'ds'' & \dfrac{\varepsilon_r+1}{2}\displaystyle\iint_{S_I}\iint_{S_I} G(\underline{r}'|\underline{r}'')ds''ds' + \\ & (\varepsilon_r-1)\displaystyle\iint_{S_I}\iint_{S_I}\int_{S_I} G(\underline{r}|\underline{r}')\frac{\partial G(\underline{r}|\underline{r}'')}{\partial n}dsds'ds'' \end{array}\right]$$

$$\times \begin{pmatrix} \sigma \\ \\ \sigma_I \end{pmatrix} = \begin{pmatrix} \displaystyle\int_S g(\underline{r})ds \\ \\ 0 \end{pmatrix} . \tag{5.30}$$

The coefficient matrix in (5.30) is not generally symmetric (for example when $\varepsilon_r=1$). Therefore the "operator" is not self-adjoint, but a positive-definite, self-adjoint component exists (from the matrix in (5.28)). Energy convergence is guaranteed, provided that the solution is unique.

If the solution is not unique there must be charges σ and σ_I which produce a potential vanishing at S, at infinity and which satisfies the interface condition. Positive-definiteness of the Laplacian is sufficient to show that this potential vanishes everywhere. It is necessary to show,

for non-uniqueness, that there exist non-zero σ and σ_I such that

$$\int_S \sigma(\underline{r}')G(\underline{r}|\underline{r}')ds' + \int_{S_I} \sigma_I(\underline{r}')G(\underline{r}|\underline{r}') = 0 \ ; \text{ any } \underline{r} \ . \tag{5.31}$$

One may define a surface S', by means of cuts, which includes S and S_I, as shown in Fig. 5.3. Then

$$\int_{S'} \sigma'(\underline{r}')G(\underline{r}|\underline{r}')ds'=0, \ \underline{r} \text{ on } S' \ . \tag{5.32}$$

Positive-definiteness of the integral operator demands that σ' on S', (including σ and σ_I) vanishes. Therefore the solution to the mutually constrained problem is unique.

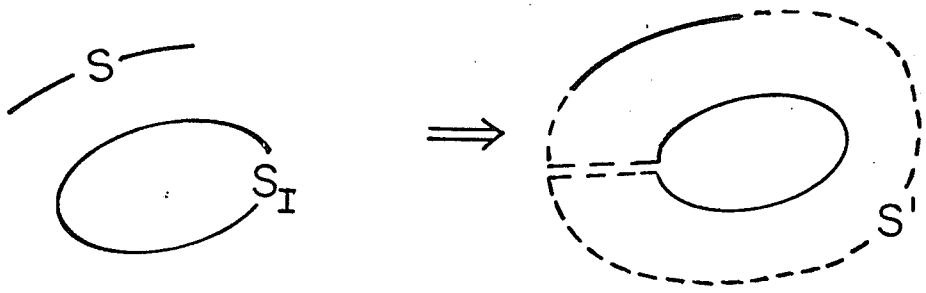


Fig. 5.3 A Single Surface S' Including S and S_I.

5.3 Finite Element Considerations

To facilitate computation of the integrals for the Dirichlet functional contributions to the matrix equation it is convenient to use finite elements for S_I which are the same as for S. Therefore line elements are used for two-dimensional problems, and rectangular elements for problems in three dimensions. No singular function representation is used for σ_I, and solutions are obtained using pulse functions alone. The algorithms in Section 4.3 are entirely sufficient for computing the constraint contributions from the elements on S_I for two- and three-dimensional problems, respectively. To be specific, suppose there are N elements on S and N_I elements on S_I. The matrix equation components from the Dirichlet functional [A] and [A'] in (5.28) and (5.29) are given by

$$A_{ij} = \frac{1}{\varepsilon_o} \int_{S_i} \int_{S_j} G(\underline{r}|\underline{r}')\, ds'\, ds; \quad i=1 \,..\, N \;;\; j=1 \,..\, N \tag{5.33a}$$

and

$$A'_{ij} = \frac{1}{\varepsilon_o} \int_{S_i} \int_{S_{Ij}} G(\underline{r}|\underline{r}')ds'\, ds \;;\; i=1 \,..\, N \;;\; j=1 \,..\, N_I \;. \tag{5.33b}$$

The components from the interface functional, [D] and [D'] in (5.28) and (5.29), are given by

$$D_{ij} = \frac{\varepsilon_r+1}{2\varepsilon_o} \int_{S_{Ii}} \int_{S_{Ij}} G(\underline{r}|\underline{r}')ds'ds$$

$$+ \frac{\varepsilon_r-1}{\varepsilon_o} \int_{S_{Ii}} \int_{S_{Ij}} \sum_{k=1}^{N_I} \int_{S_{Ik}} G(\underline{r}|\underline{r}')\frac{\partial G(\underline{r}|\underline{r}'')}{\partial n}\, ds\, ds''ds';$$

$$i=1 \;\; N_I \;;\; j=1 \,..\, N_I \tag{5.34a}$$

and

$$D'_{ij} = \frac{\varepsilon_r - 1}{\varepsilon_o} \int_{S_{Ii}} \int_{S_j} \sum_{k=1}^{N_I} \int_{S_{Ik}} G(\underline{r}|\underline{r}') \frac{\partial G(\underline{r}|\underline{r}'')}{\partial n} ds\, ds''ds; \quad i=1..N_I; \quad j=1..N \,. \tag{5.34b}$$

The two integrals in (5.33) and the first in (5.34a) are all handled in the same way, and are computed by the methods described in Section 4.3. The integrals over S_{Ik} in (5.34a) and (5.34b) are obtained as follows

Let
$$I_k = \int_{S_{I_k}} G(\underline{r}|\underline{r}') \frac{\partial G(\underline{r}|\underline{r}'')}{\partial n} ds \,. \tag{5.35}$$

Then the second integral in (5.34a) and the single integral (5.34b) are of the form

$$I = \int_{S_i} \int_{S_j} f(\underline{r}', \underline{r}'') ds'ds \tag{5.36a}$$

where

$$f(\underline{r}', \underline{r}'') = \sum_{k=1}^{N} I_k \tag{5.36b}$$

Numerical integration techniques are readily applied in integrating (5.36).

The integral (5. 35) for a line element (two-dimensional problem) may be transformed to (u,v) coordinates, as shown in Fig. 5.4. It is computed, then, as two of the form

$$I = v'' \int_0^a \frac{\ln \sqrt{u^2+v'^2}}{v''^2+(u''-u)^2} du \,. \tag{5.37}$$

If $v''=0$ the integral vanishes. If $v''\neq 0$, the addition-subtraction technique is applied:

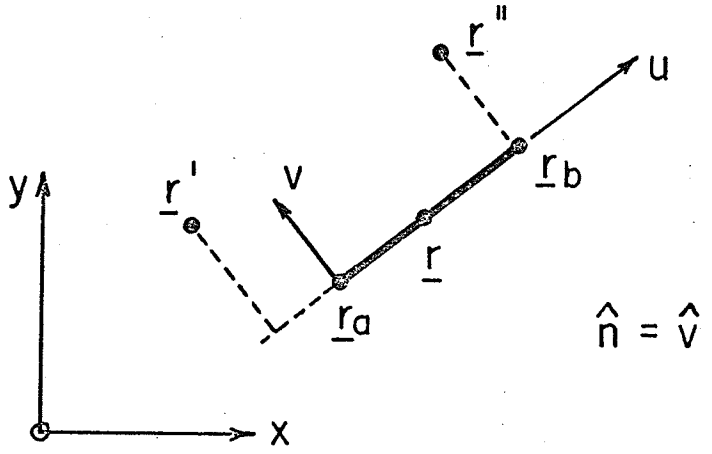


Fig. 5.4 Local Coordinates for the 2-D Interface Problem

$$I = v'' \int_0^a \left\{ \frac{\ln\sqrt{u^2+v'^2}}{v''^2+(u''-u)^2} - \frac{\ln\sqrt{u^2+v'^2}}{v''^2+u''^2} \right\} du + \frac{v''}{v''^2+u''^2} \int_0^a \ln\sqrt{u^2+v'^2}\, du \quad . \tag{5.38}$$

The second integral has been treated previously, at (4.27). The first integrand is regular, permitting numerical integration to be used. To reduce computational effort, it is convenient to treat the integrand as regular if $v' \geq 0.25$. Also, if the projected point u' lies off the line segment, it is convenient to compute (5.37) as one integral, numerically, and directly over the line element, for the integrand is regular.

For the three-dimensional problem the methods described in Section 4.2 are used to determine whether or not $\underline{r}''$ is in the plane of S_{I_k}.

If so, no computation is done, as the integral vanishes. Otherwise, if the projection of $\underline{r}'$ to S_{I_k} does not lie within the element, numerical integration is used directly. If the projection of point $\underline{r}'$ does lie in S_{I_k}, the integral is written as four of the form

$$I = \int_0^a \int_0^b \frac{w''\, du\, dv}{\sqrt{w'^2+u^2+v^2}\ \{w''^2+(u''-u)^2+(v''-v)^2\}^{3/2}} \tag{5.39}$$

where u and v are measured from the projection of $\underline{r}'$. See Fig. 5.5. In these local coordinates

$$G(\underline{r}|\underline{r}') = \{w'^2+u^2+v^2\}^{-1/2} \tag{5.40a}$$

and

$$\frac{G(\underline{r}|\underline{r}'')}{\partial n} = \frac{\cos\theta}{\{w''^2+(u''-u)^2+(v''-v)^2\}\ 1/2} = \frac{\cos\theta}{d} \tag{5.40b}$$

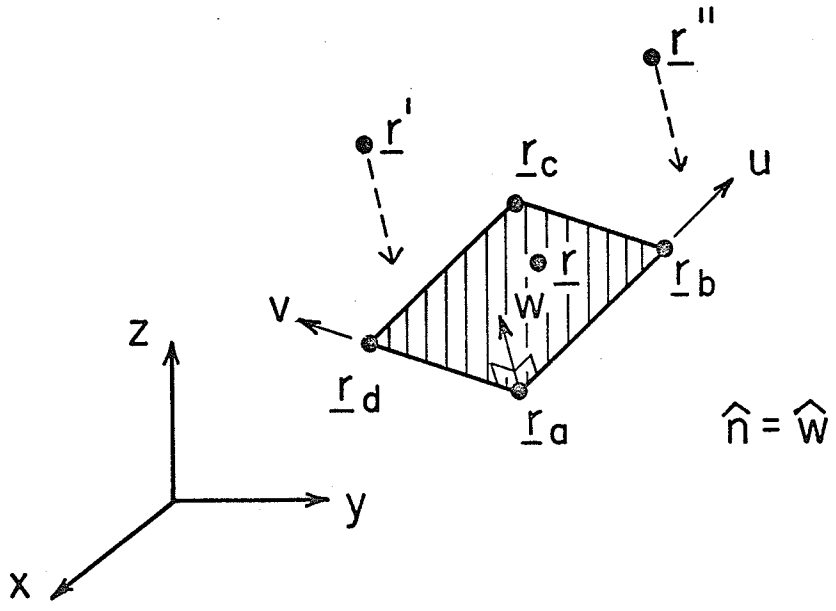


Fig. 5.5 Local Coordinates for the 3-D Interface Problem

where

$$\cos\theta = \frac{w''}{d} \quad . \tag{5.40c}$$

Equations (5.40) were used to construct (5.39).

The addition-subtraction technique is applied to (5.39):

$$I = \int_0^a \int_0^b \frac{w''}{\sqrt{w'^2+u^2+v^2}} \left\{ \frac{1}{\{w''^2+(u''-u)^2+(v''-v)^2\}^{3/2}} - \frac{1}{\{w''^2+u''^2+v''^2\}^{3/2}} \right\} du\ dv$$

$$+ \frac{w''}{\{w''^2+u''^2+v''^2\}^{3/2}} \int_0^a \int_0^b \frac{du\ dv}{\sqrt{w'^2+u^2+v^2}} \quad . \tag{5.41}$$

The first integrand is regular, permitting numerical integration, and the second integral has been investigated previously, at (4.38).

The algorithms are now complete for linear pulse function elements, and the complete matrix equation may be constructed.

5.4 A Two-Dimensional Example

Consider the square parallel-strip capacitor with the gap filled by a dielectric (ε_r) as shown in Fig. 5.6a. The mirror image symmetries are used to reduce the problem to that of finding σ and σ_I in the positive quadrant as shown in Fig. 5.6b. Image Green's functions are constructed during integration by the procedure described in Appendix B.

Note that S and S_I coincide on $0 \leq x \leq 1$, $y=1$. This appears to be in conflict with the derivation (see equation (5.22)) but that is not really the case. All integrations involving $\partial G(\underline{r}|\underline{r}'')/\partial n$, where $\underline{r},\underline{r}''$ are both on $0 \leq x \leq 1$, $y=1$, vanish. This allows one to obtain a net charge $\sigma+\sigma_I$ on S, and for convenience, this charge is simply denoted σ. Note that if the entire interface S_I lies in a plane, all integrals involving $\partial G/\partial n$ vanish. This is the special case investigated by Bryant and Weiss [1].

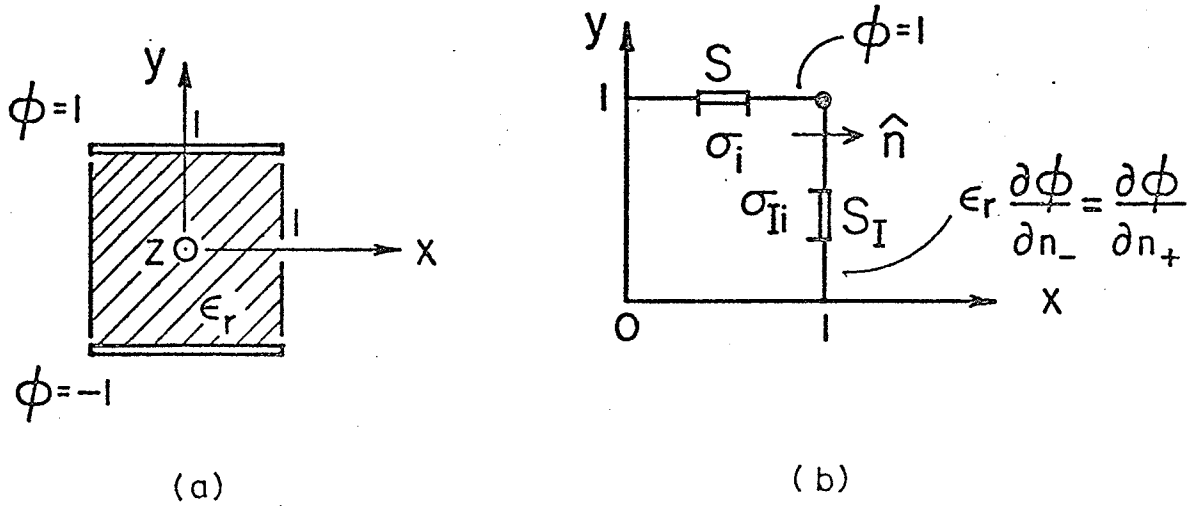


Fig. 5.6 Strip Capacitor with Dielectric Slab: Full View
a) Full View of Problem in Cross-Section
b) Reduced Geometry and Element Placement

Line elements, each of length 1/N are placed on S and S_I, resulting in a (2N)x(2N) matrix equation for all pulse heights. All numerical integrations are performed using order 3 Gaussian quadrature.

The interface condition (5.25) may be rewritten as

$$\frac{\partial \phi}{\partial n_-} = \frac{1}{\varepsilon_r} \frac{\partial \phi}{\partial n_+} \quad . \tag{5.42}$$

Then

$$\lim_{\varepsilon_r \to \infty} \frac{\partial \phi}{\partial n_-} = 0 \quad . \tag{5.43}$$

In this case, for this problem the solution ($\varepsilon_r \to \infty$) within the dielectric should satisfy

$$\phi(x,y) = y \quad ; \quad 0 \leq x \leq 1, \; 0 \leq y \leq 1 \quad . \tag{5.44}$$

In order to prevent scaling problems, the interface functional is divided by ε_r before computation, whence, in (5.30) the coefficient $(\varepsilon_r+1)/2\varepsilon_o$ becomes $(1+1/\varepsilon_r)/2\varepsilon_o$, and $(\varepsilon_r-1)/\varepsilon_o$ becomes $(1-1/\varepsilon_r)/\varepsilon_o$. In effect the equations are normalized to ε_r.

The results presented in Table 5.1 were obtained with N=10. As ε_r is made large, $\phi(1,y) \sim y$, within 0.01 percent at $\varepsilon_r = 10^6$. This verifies the algorithm for this problem.

TABLE 5.1 THE 2-D INTERFACE PROBLEM;
STRIBS WITH SLAB

y	Potential at (y,1) $\varepsilon_r=1$	$\varepsilon_r=10$	$\varepsilon_r=10^3$	$\varepsilon_r=10^6$	|% diff.| $\phi$to y $(\varepsilon_r=10^6)$
0.05	0.037925	0.048108	0.050100	0.050006	0.012
0.15	0.113960	0.114360	0.150299	0.150018	0.012
0.25	0.190573	0.240728	0.250498	0.250029	0.012
0.35	0.268209	0.337299	0.350698	0.350041	0.012
0.45	0.347432	0.434184	0.450898	0.450052	0.012
0.55	0.429015	0.531535	0.551098	0.550063	0.012
0.65	0.514139	0.629579	0.651298	0.650074	0.011
0.75	0.604837	0.728689	0.751477	0.750085	0.011
0.85	0.705232	0.829472	0.851446	0.850096	0.011
0.95	0.823645	0.928782	0.945868	0.949910	0.010

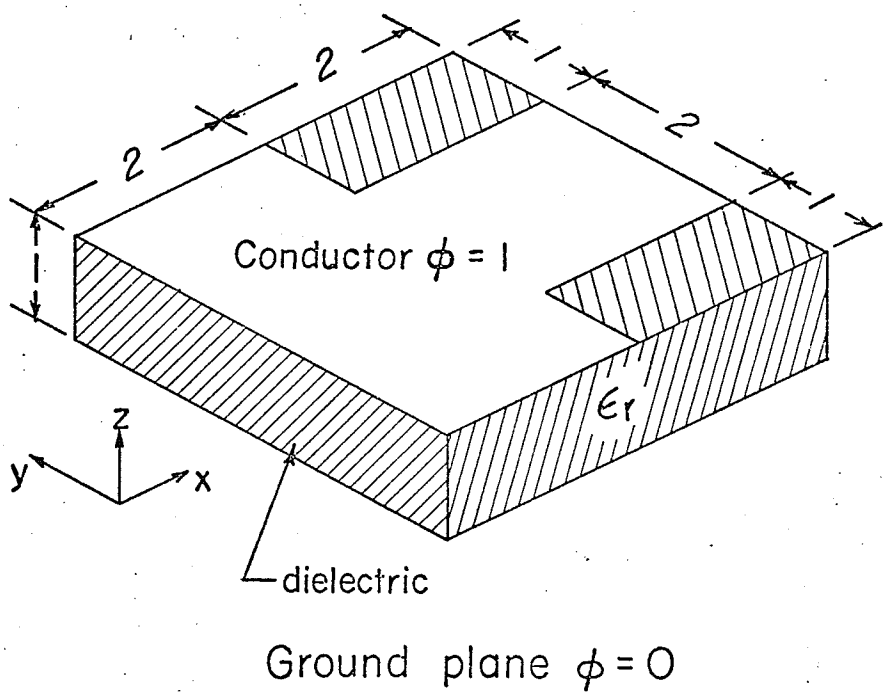


Fig. 5.7 3-D Interface Problem - The T-shaped Conductor on a Dielectric (ε_r=10): Full View

5.5 A Three-Dimensional Example

The T shaped conductor of Section 4.5 is placed on a dielectric slab (ε_r=10) as shown in Fig. 5.7. This problem is illustrative of a microstrip discontinuity (with quasi-static analysis) or of a portion of an integrated circuit. The potential specified on the conductor is ϕ=1, and the symmetries are used as in Section 4.6 to reduce the problem as depicted in Fig. 5.8.

Square pulse function elements of side 1/N are placed on S (where the net charge σ is sought) and on S_I. There are $6N^2$ elements on the conductor, as before, with $2N^2$ elements on S_I at z=1, x=0, and x=4, and $4N^2$ at y=2. Therefore there are $16N^2$ pulse heights in total to be found. Order 3 Gaussian quadrature is used for all numerical integrations involving the pulse elements. The matrix equation for pulse functions

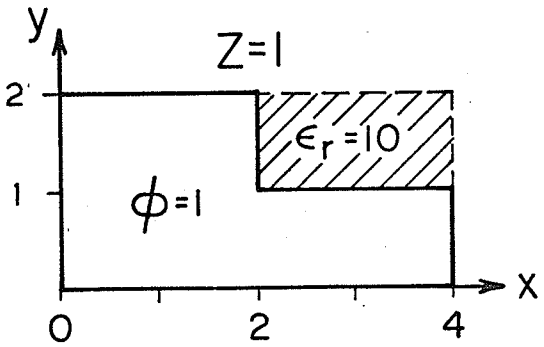


Fig. 5.8 3-D Interface Problem: Reduced Geometry Viewed From Above.

is constructed as in the example of Section 5.4, but using the formulas for three dimensional analysis presented in Section 5.3.

The special function σ_s (4.74) is introduced to assist in modelling the singularity, with the choice α=0.25, as in Section 4.5. This adds one entry to the right hand vector, and a row and column to the matrix, which now may be represented as

$$\begin{bmatrix} A & \underline{\ell} & A' \\ \underline{\ell}^T & a_o & \underline{p}^T \\ D' & \underline{q} & D \end{bmatrix} \begin{pmatrix} \underline{\sigma} \\ \sigma_o \\ \underline{\sigma}_I \end{pmatrix} = \begin{pmatrix} \underline{b} \\ b_o \\ \underline{0} \end{pmatrix} . \qquad (5.45)$$

The vector entry $\underline{\ell}$, the diagonal entry a_o, and the vector entry b_o are identical to the additions made in Section 4.5, at equation (4.76). The elements of the vector $\underline{p}$ are given by

$$p_i = \int_S \sigma_s(\underline{r}) \int_{S_{Ii}} G(\underline{r}|\underline{r}')ds' \, ds \; ; \; i=1 \ldots N_I \qquad (5.46)$$

and those of $\underline{q}$ are given by

$$q_i = \frac{\varepsilon_r - 1}{\varepsilon_o} \int_S \sigma_s(\underline{r}'') \int_{S_{Ii}} \sum_{k=1}^{N_I} \int_{S_{Ik}} G(\underline{r}|\underline{r}') \frac{G(\underline{r}|\underline{r}'')}{\partial n} ds \, ds'ds'' \; i=1 \ldots N_I . \qquad (5.47)$$

These integrals are readily computed by methods presented previously. All numerical integrations involving σ_s, that is over S_a, S_b and S_c as defined in Fig. 4.8, are performed using order 6 Gaussian quadrature.* Equipotentials are shown in Figs. 5.9-5.12 at several cross-sections, for the case N=2. For the case N=3 there are 145 unknowns, and the matrix determinant vanished, as in Section 4.5.

* Order 3 quadrature was adequate for the small pulse regions but order 6 was needed in the larger regions for σ_s (S_a, S_b, and S_c).

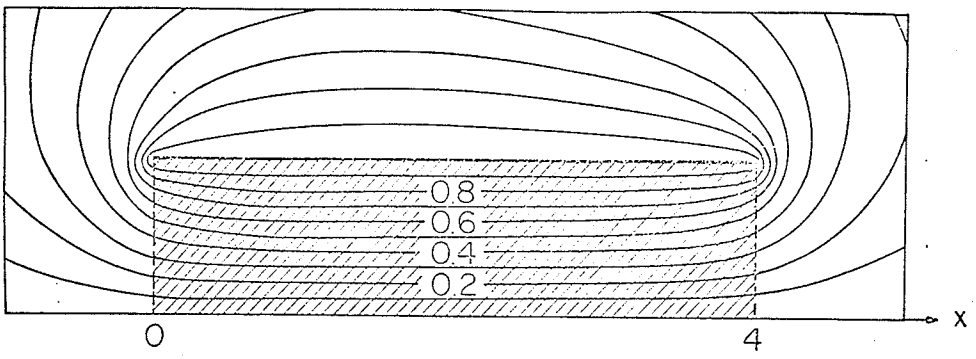


Fig. 5.9 3-D Interface Problem:
Equipotentials at y=0.5

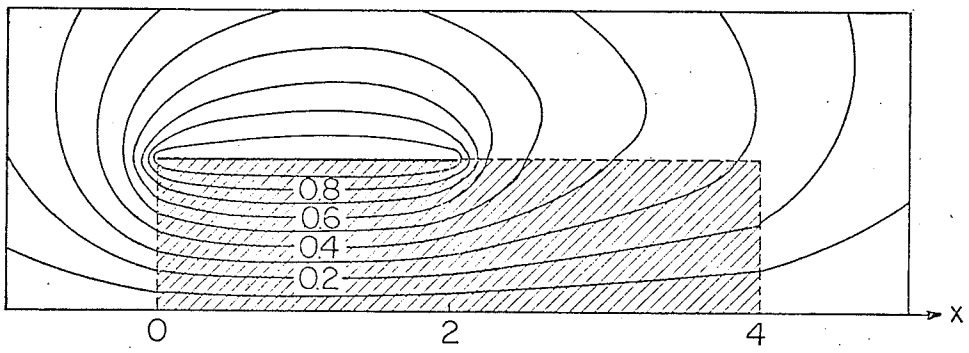


Fig. 5.10 3-D Interface Problem:
Equipotentials at y=1.5

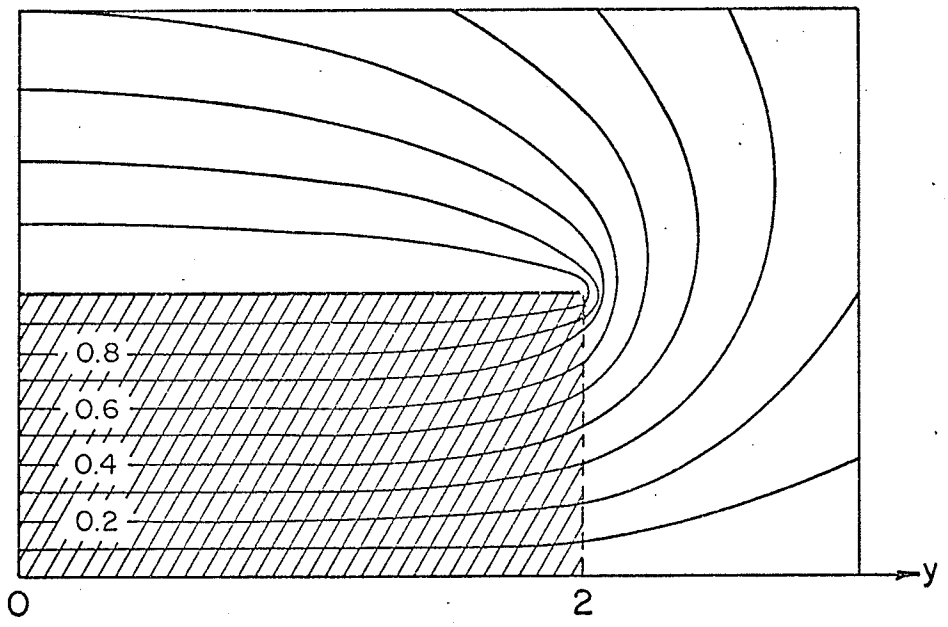


Fig. 5.11 3-D Interface Problem:
Equipotentials at x=1.0

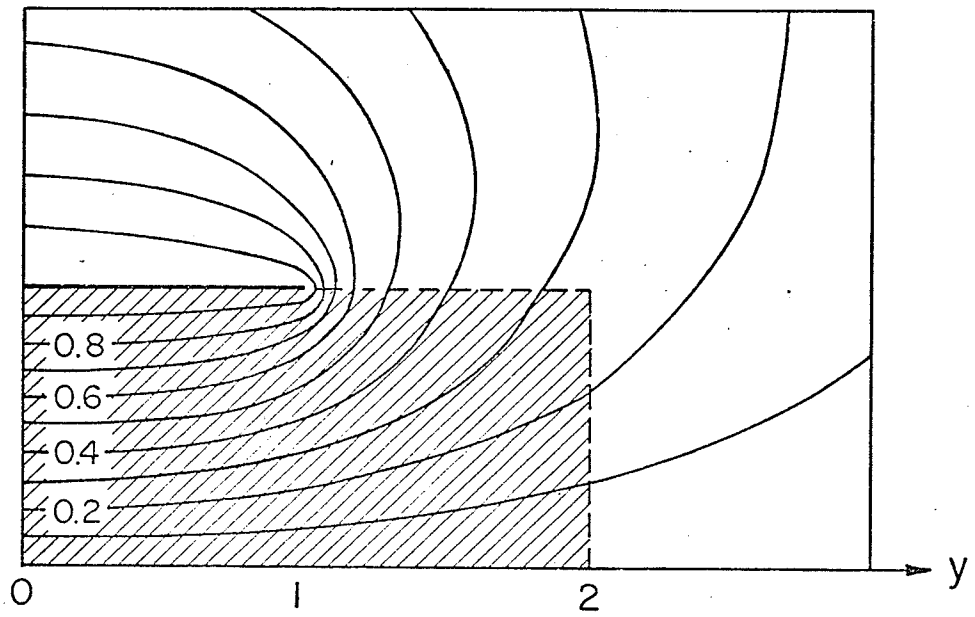


Fig. 5.12 3-D Interface Problem: Equipotentials
at x=3.0

5.6 Conclusions

Interface problems may be solved by mutual constraint of IE variational expressions in the static case. Energy convergence is guaranteed when the operator in each finite element is positive-definite and self-adjoint.

The interface problem demonstrates the duality of the PDE and IE functionals; one may be derived from the other. This makes relevant Hammond's statement: there is a freedom of choice between field and source methods. One may solve variationally for either potential or source, and in either case, the interface problem is solved.

The method of mutual constraint described here is general purpose in nature. There are no special restrictions on the shapes of S or S_I, but there can be modelling errors when line elements or rectangular elements are used as described here. The technology of finite-elements is available and when the way is found of performing the required analytic integration over curved elements, isoparametric finite-elements may be used to more adequately model arbitrary shapes.

It is necessary that the dielectric be homogeneous. From another point of view (see Stratton [15, p.187]) the charge σ_I is an effective polarization charge which replaces the dielectric. If the dielectric is inhomogeneous, a polarization charge $\underline{P}$ may be introduced throughout the dielectric, to satisfy the equation $\nabla \cdot \underline{P} = 0$. The method described in this chapter could be extended to handle this case, but there would be the unknown polarization $\underline{P}$ to determine within the dielectric region, with an appropriate increase in the number of unknowns. It is better to use picture-frame methods in such cases, and mutually constrain PDE and IE formulations.

One advantage of the present approach over many is that only the free-space Green's function is used; there is no need to construct special Green's functions (such as the two-dielectric Green's function given in [16]). However, to apply the addition-subtraction technique it is necessary to be able to integrate G and $G\partial G/\partial n$ analytically over each finite-element, and this has been done for the linear elements used here.

One may compare this technique to the picture-frame method of Chapter III. The dielectric may loosely be thought of as a picture-frame, and the interface condition at S_I as the external region constraint. Since the dielectric is homogeneous, the static picture-frame solution for ϕ may be replaced by a solution for σ_I on the boundary of the dielectric S_I. Thus there is a relationship between the picture-frame method and this IE method. When the region is piecewise homogeneous this method is better because fewer unknowns are involved, which can result in reduced computer storage and time requirements. On the other hand, when the dielectric is not homogeneous or is anisotropic the method described in this chapter is not directly applicable and one is advised to use the picture-frame method.

5.7 References

[1] T.G. Bryant and J.A. Weiss, "Parameters of microstrip transmission lines and of coupled pairs of microstrip lines." IEEE Trans. MTT-16, pp.1021-1027, December 1968.

[2] P. Benedek and P. Silvester, "Capacitance of parallel rectangular plates separated by a dielectric sheet." IEEE Trans. MTT-20, pp.504-510, August, 1972.

[3] A. Farrar and A.T. Adams, "Matrix methods for microstrip three dimensional problems." IEEE Trans., MTT-20, pp.497-504, August 1972.

[4] T. Itoh, R. Mittra and R. Ward, "A method for computing edge capacitance of finite and semi-finite lines." IEEE Trans.MTT-20, pp.847-849, December 1972.

[5] D. Gish, "Characteristic Impedance and phase velocity of a dielectric supported air strip transmission line with sidewalls." IEEE Trans. MTT-18, pp.131-148, March 1970.

[6] J.L. Allen and M.F. Estes, "Broadside coupled strips in a layered dielectric medium." IEEE Trans. MTT-20, pp.662-668, October 1972.

[7] M. Maeda, "An analysis of gap in microstrip transmission lines." IEEE Trans. MTT-20, pp.390-396, June 1972.

[8] T.E. Rozzi, "The variational treatment of thick interacting inductive crises." IEEE Trans. MTT-21, pp.82-88, February 1973.

[9] H. Levine and J. Schwinger, "On the theory of diffraction by an aperture in an infinite plane screen." Phys. Rev. Vol. 74, pp.958-974, 1948.

[10] R. Mittra, T. Itoh and T. Li, "Analytical and numerical studies of the relative convergence phenomena arising in the solution of an integral equation by the moment method." IEEE Trans. MTT-20 pp. 96-104, February 1972.

[11] J. Minor and D. Bole, "Modes in the shielded microstrip on a ferrite substrate transversely magnetized in the plane of the substrate." IEEE Trans. MTT-19, pp. 570-576, July 1971.

[12] V.K. Thong, "Solutions for some waveguide discontinuities by the method of moments." IEEE Trans. MTT-20, pp.416-418, June 1972.

[13] Y. Chow and S. Wu, "A moment method with mixed linear basis functions for scatterings by waveguide junctions." IEEE Trans. MTT-21, pp. 333-340, May 1973.

[14] S.G. Mikhlin, Variational Methods in Mathematical Physics, New York: Macmillan, 1964.

[15] J.A. Stratton, Electromagnetic Theory. New York: McGraw-Hill, 1941.

[16] A.F. dos Santos and V.R. Vieira, "The Green's function for Poisson's equation in a two dielectric region." IEEE Trans. MTT-20, p.777, November 1972.

CHAPTER VI

MUTUALLY CONSTRAINED IE AND PDE VARIATIONAL METHODS

The technique described in Chapter III permits simultaneous solution of PDE and IE expressions by a Lagrange constraint method. While the PDE expression is variational, the IE expression is constructed by a point-matching technique. The methods of Chapters IV and V allow IE problems to be solved variationally, and allow different IE variational expressions to be mutually constrained for problem solution. In this chapter the IE portion of the picture-frame problem is constructed variationally and the PDE and IE expressions are solved by mutual constraint. (The PDE portion is variational, as before.)

The static IE functional (4.15) must be extended to the time-harmonic case. As the Green's function (1.8) is complex the integral operator is complex, and adjoint expressions are required to construct the functional. The PDE functional (2.48) is readily extended to the axi-symmetric problem. It becomes possible to study four cases of potential in the plane: static and time-harmonic problems having either uniform cross-section or axial symmetry. The technique is particularly applicable to problems over infinitely extending regions with isolated inhomogeneities, as in Chapter III.

6.1 A Functional for the Complex Integral Operator

In Section 4.1 the static integral operator was shown to be self-adjoint and positive-definite (or easily made so for finite S by adding a constant for two-dimensional problems). The quadratic functional (4.15) is directly applicable. For Helmholtz problems (the time-harmonic case) the operator is not self-adjoint:

$$\langle Ku,v\rangle = \int_S v^*(\underline{r}) \int_S u(\underline{r}') \frac{e^{jk|\underline{r}-\underline{r}'|}}{4\pi\varepsilon_o|\underline{r}-\underline{r}'|}\, ds'ds \qquad (6.1a)$$

and

$$\langle u,Kv\rangle = \int_S v^*(\underline{r}) \int_S u(\underline{r}') \frac{e^{-jk|\underline{r}-\underline{r}'|}}{4\pi\varepsilon_o|\underline{r}-\underline{r}'|}\, ds'ds \quad ; \qquad (6.1b)$$

therefore

$$\langle Ku,v\rangle \neq \langle u,Kv\rangle \qquad (6.1c)$$

and the quadratic functional may not be used directly.

Certain properties of complex numbers are useful: if u=a+jb, then

$$\left.\begin{aligned} u\,u &= (a+jb)(a+jb) = (a^2-b^2) + j(2ab) \\ u\,u\,u &= (a^3-3ab^2) + j(3a^2b-b^3) \\ u^*u^* &= (a^2-b^2) - j(2ab) \\ u^*u^*u^* &= (a^3-3ab^2) - j(3a^2b-b^3) \\ u\,u+u^*u^* &= 2(a^2-b^2) = 2\ \mathrm{Re}\colon u\,u \end{aligned}\right\} \qquad (6.2)$$

and

$$u\,u\,u+u^*u^*u^* = 2(a^3-3ab^2) = 2\ \mathrm{Re}\colon u\,u\,u \quad .$$

Consider the functional

$$F(\sigma) = \frac{1}{\varepsilon_o} \int_S \int_S \{\sigma^*(\underline{r})\sigma^*(\underline{r}')G^*(\underline{r}|\underline{r}') + \sigma(\underline{r})\sigma(\underline{r}')G(\underline{r}|\underline{r}')\}\, ds'ds$$

$$- 2\int_S \{\sigma^*(\underline{r})g^*(\underline{r}) + \sigma(\underline{r})g(\underline{r})\}ds \quad . \tag{6.3}$$

It is easily shown, using (6.2) that F is a real number, and, in fact

$$F(\sigma) = 2\,\mathrm{Re}: \quad \frac{1}{\varepsilon_o} \int_S \int_S \sigma(\underline{r})\sigma(\underline{r}')G(\underline{r}|\underline{r}')ds'ds - 2\int_S \sigma(\underline{r})g(\underline{r})ds \, . \tag{6.4}$$

Consider the first variation of (6.3):

$$\delta F(\sigma) = 2\int_S \delta\sigma^*(\underline{r})\left\{ \int_S \frac{\sigma^*(\underline{r}')}{\varepsilon_o} G(\underline{r}|\underline{r}')ds' - g^*(\underline{r}) \right\} ds$$

$$+ 2\int_S \delta\sigma(\underline{r})\left\{ \int_S \frac{\sigma(\underline{r}')}{\varepsilon_o} G(\underline{r}|\underline{r}')ds' - g(\underline{r}) \right\} ds$$

$$= 4\,\mathrm{Re}: \int_S \delta\sigma(\underline{r})\left\{ \int_S \frac{\sigma(\underline{r}')}{\varepsilon_o} G(\underline{r}|\underline{r}')\, ds' - g(\underline{r}) \right\} ds \quad . \tag{6.5}$$

At the stationary point of F the Helmholtz equation is satisfied. The Rayleigh-Ritz (or Galerkin) equation may be written directly, just as in the static case. That is, if

$$\sigma(\underline{r}) = \underline{f}^T \underline{\sigma} \tag{6.6}$$

the matrix equation is

$$\int_S \int_S \underline{f}(\underline{r})\ \underline{f}(\underline{r}')^T G(\underline{r}|\underline{r}')ds'\ ds\ \ \underline{\sigma} \ = \int_S \underline{f}(\underline{r})g(\underline{r})ds \ . \qquad (6.7)$$

The matrix is not Hermitian, but this equation arises at the stationary point of the quadratic functional (6.3). It is convenient to use real-valued functions $\underline{f}$ with complex coefficients $\underline{\sigma}$ as in Chapter III.

It is interesting to see how (6.3) may be derived using a matrix operator. Let

$$\mathcal{K} \ = \ \begin{bmatrix} \int_S G\ ds' & 0 \\ 0 & \int_S G^* ds' \end{bmatrix} \qquad (6.8a)$$

$$\underset{\sim}{\sigma} \ = \ \begin{pmatrix} \sigma \\ \sigma^* \end{pmatrix} \qquad (6.8b)$$

$$\underset{\sim}{g} \ = \ \begin{pmatrix} g \\ g^* \end{pmatrix} \qquad (6.8c)$$

and consider the functional

$$F(\sigma) = \langle \mathcal{K}\ \underset{\sim}{\sigma}, \underset{\sim}{\sigma} \rangle - \langle \underset{\sim}{g}, \underset{\sim}{\sigma} \rangle \ - \langle \underset{\sim}{\sigma}, \underset{\sim}{g} \rangle \quad . \qquad (6.9)$$

The inner product is defined with a weighting matrix

$$w \ = \ \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad . \qquad (6.10)$$

Then

$$\langle \mathcal{K}\underset{\sim}{u}, \underset{\sim}{v} \rangle \ = \ \int_S (v^* v) \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \int_S G\ ds' & 0 \\ 0 & \int_S G^* ds \end{bmatrix} \begin{pmatrix} u \\ u^* \end{pmatrix} ds$$

$$= \int_S \int_S (v^* u^* G^* + vuG)\ ds'ds \qquad (6.11)$$

and

$$\langle \underset{\sim}{u}, \mathcal{K}\underset{\sim}{v} \rangle = \int_S \left[\int_S G^* v^* ds' \quad \int_S G v ds' \right] \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{pmatrix} u \\ u^* \end{pmatrix} ds$$

$$= \int_S \int_S (v^* u^* G^* + vu\, G)\, ds' ds \quad . \tag{6.12}$$

Therefore

$$\langle \mathcal{K}\underset{\sim}{u}, \underset{\sim}{v} \rangle = \langle \underset{\sim}{u}, \mathcal{K}\underset{\sim}{v} \rangle \tag{6.13}$$

and the operator $\mathcal{K}$ is self-adjoint. Quite clearly

$$\langle \mathcal{K}\underset{\sim}{\sigma}, \underset{\sim}{\sigma} \rangle = \int_S \int_S (\sigma^* \sigma^* G + \sigma\sigma G)\, ds' ds \tag{6.14}$$

and

$$\langle \underset{\sim}{g}, \underset{\sim}{\sigma} \rangle = \int_S (\sigma^* \sigma) \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} g \\ g^* \end{bmatrix} ds = \int_S (\sigma^* g^* + \sigma g) ds \quad ; \tag{6.15}$$

also

$$\langle \underset{\sim}{\sigma}, \underset{\sim}{g} \rangle = \int_S (\sigma^* g^* + \sigma g) ds \quad . \tag{6.16}$$

It is apparent that the two functionals (6.3) and (6.9) are identical.

It is not possible to demonstrate positive-definiteness, for the second variation of (6.3) is

$$\delta^2 F(\sigma) = 4\ \mathrm{Re}: \frac{1}{\varepsilon_o} \int_S \int_S \delta\sigma(\underline{r})\, \delta\sigma(\underline{r}')\, G(\underline{r}|\underline{r}')\, ds'\, ds \ . \tag{6.17}$$

Consider $\delta\sigma = a+jb$ and $G=M+jN$. The integrand of (6.17) is then

$$\mathrm{Re}: \delta\sigma\ \delta\sigma\ G = (a^2 - b^2)M - 2ab\ N \tag{6.18}$$

and this number may be negative. Nevertheless the operator contains the Laplacian as a component, and by Mikhlin's theorem, energy convergence in the Laplacian is guaranteed, provided that the problem has a unique solution. Energy convergence is in

$$U(\sigma) = \int_S \int_S \sigma(\underline{r})^* \, \sigma(\underline{r}') \, \frac{1}{4\pi|\underline{r}-\underline{r}'|} \, ds' \, ds \quad . \tag{6.19}$$

This is clearly real, positive, and finite (for finite S and integrable σ).

Incidentally, the preceeding discussion suggests that the IE variational techniques presented in Chapters IV and V for static problems could be extended in a straightforward manner to the propagation (time-harmonic) case. Perhaps this might be a useful extension as well.

6.2 Integral Solution of the Exterior Problem

As demonstrated in Section 3.2, the potential for the exterior region Ω_e (outside the picture-frames) satisfies

$$\phi(\underline{r}) = 2\int_S \left(G(\underline{r}|\underline{r}') \frac{\partial \phi(\underline{r}')}{\partial n'} - \phi(\underline{r}') \frac{\partial G(\underline{r}|\underline{r}')}{\partial n'} \right) ds' \; ; \; \underline{r} \text{ on } S \qquad (6.20)$$

where S is the cumulative boundary of all picture-frames (i.e. the boundary of Ω_e) and $\hat{n}$ points out of Ω_e. If a Dirichlet condition $\phi=g$ is specified on S, one can construct the following IE:

$$\int_S \frac{\sigma(\underline{r}')}{\varepsilon_o} G(\underline{r}|\underline{r}')ds' = \frac{g(\underline{r})}{2} + \int_S g(\underline{r}') \frac{\partial G(\underline{r}|\underline{r}')}{\partial n'} ds' \qquad (6.21)$$

where $\sigma = \varepsilon_o \partial\phi/\partial n$. This is a Fredholm IE of the first kind and has been considered for the static case in Chapter IV. Burton and Miller [1] assert that (6.20) is satisfied uniquely for any propagation constant k, including picture-frame eigenvalues. By use of equation (6.7) a matrix equation may be written directly. If $\sigma(\underline{r})=\underline{f}^T\underline{\sigma}$ (with $\underline{f}$ real-valued) the equation is

$$\frac{1}{\varepsilon_o} \int_S \int_S \underline{f}\, \underline{f}^T G \, ds' \, ds \, \underline{\sigma} = \int_S \underline{f} \left(\frac{g}{2} + \int_S g \frac{\partial G}{\partial n'} ds' \right) ds \quad . \qquad (6.22)$$

When S is composed of line elements the methods of Chapter IV may be used to compute the integrations required in (6.22). If $\underline{f}$ is a regular function on each element S_i:

$$\int_S \underline{f}(\underline{r})\, p(\underline{r})\, ds = \sum_{i=1}^{N} \int_{S_i} \underline{f}(\underline{r})\, p(\underline{r})\, ds \quad . \qquad (6.23)$$

Furthermore, if p is regular, the integration over each element may be computed numerically.

The integral on the left of (6.22) is included when p is a vector function

$$\underline{p}(\underline{r})^T = \sum_{i=1}^{N} \int_{S_i} \underline{f}(\underline{r}')^T \, G(\underline{r}|\underline{r}') \, ds' \tag{6.24}$$

and what is required is the integral of $\underline{f}^T G$ over each element.

The right hand side of (6.22) is included when

$$p(\underline{r}) = \frac{g(\underline{r})}{2} + \sum_{i=1}^{N} \int_{S_i} g(\underline{r}') \, \frac{\partial G(\underline{r}|\underline{r}')}{\partial n'} \, ds' \quad . \tag{6.25}$$

The integrals which include singularities are thus isolated :

$$I_1 = \int_{S_i} \underline{f}(\underline{r}')^T \, G(\underline{r}|\underline{r}') \, ds' \tag{6.26a}$$

and

$$I_2 = \int_{S_i} \underline{f}(\underline{r}')^T \, \frac{\partial G(\underline{r}|\underline{r}')}{\partial n'} \, ds' \tag{6.26b}$$

where, simply for convenience, g on S has been approximated by the same representation as σ; that is $g(\underline{r}') = \underline{f}^T \underline{g}$. For the Dirichlet problem it is simply necessary to post-multiply (6.26b) by $\underline{g}$, the given Dirichlet node potentials on S_i. It will be seen that for the picture-frame problem $\underline{g}$ will not be known, and the representation here is convenient.

Suppose that $\underline{r}''$ is the point on S_i closest to $\underline{r}$; that is $|\underline{r}''-\underline{r}|$ is a minimum distance. Suppose further that G=P+Q where Q is the singular part. Then the addition-subtraction technique is applied to (6.26) to yield

$$I_1 = \int_{S_i} \left\{ \underline{f}(\underline{r}')^T G(\underline{r}|\underline{r}') - \underline{f}(\underline{r}'')^T Q(\underline{r}|\underline{r}') \right\} ds + \underline{f}(\underline{r})^T \int_{S_i} Q(\underline{r}|\underline{r}')ds' \tag{6.27a}$$

and

$$I_2 = \int_{S_i} \{ \underline{f}(\underline{r}')^T \frac{\partial G(\underline{r}|\underline{r}')}{\partial n'} - \underline{f}(\underline{r}'')^T \frac{\partial Q(\underline{r}|\underline{r}')}{\partial n'} \} ds' + \underline{f}(\underline{r}'')^T \int_{S_i} \frac{\partial Q(\underline{r}|\underline{r}')}{\partial n'} ds' . \tag{6.27b}$$

The first integrands in I_1, I_2 are regular, and numerical quadrature may be applied directly. (Note: if $\underline{r}''$ lies on S_i, and is not an end point, the numerical integration may be done in two parts, to ease derivative singularities as indicated in Appendix C). If $\underline{r}$ lies sufficiently far from S_i, equations (6.26) may be integrated numerically. The coordinate transformations described in Section 4.3 allow the minimum distance between $\underline{r}'$ and S_i to be easily established, and as shown in Fig. 3.7 a minimum distance of 0.25 is sufficient to ensure well behaved integrands.

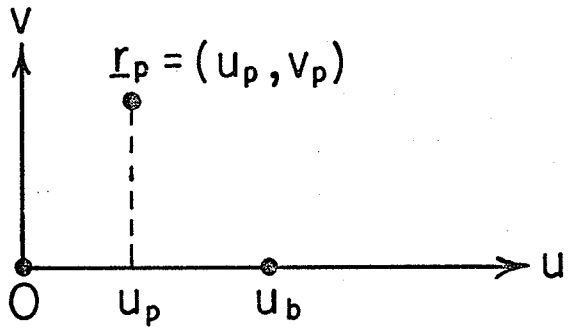
The (u,v) coordinate system of Section 4.3 may be used to compute the integrations in (6.27). The observation point here is $\underline{r}$, and in the local coordinate is (u_p, v_p), as shown in Fig. 6.1. If $u_p<0$, the point on the u axis closest to the observation point is u=0, and if $u_p>u_b$ it is $u=u_b$. Otherwise it is $u=u_p$. To be general, this point is denoted as $\underline{r}''$, and lies on the element; it is the point on the element closest to the observation point $\underline{r}$.

The singular function Q is given by

$$Q(\underline{r}|\underline{r}') = -\frac{1}{2\pi} \ln \sqrt{v_p^2+(u-u_p)^2} \tag{6.28}$$

and its derivative is given by

$$\frac{\partial Q(\underline{r}|\underline{r}')}{\partial n'} = +\frac{1}{2\pi} \frac{v_p}{(v_p^2+(u-u_p)^2)} \tag{6.29}$$



$$u_p \leq 0 \rightarrow u'' = 0$$
$$u_p \geq u_b \rightarrow u'' = u_b$$
$$0 < u_p < u_b \rightarrow u'' = u_p$$

Fig. 6.1 Local Coordinates for Singular Point u".

where $\hat{n}$ is taken in the +w direction, to the left of the line from a to b, as before. If $v_p=0$ the normal derivative vanishes and (6.29) is simply not computed.

The coordinates of $\underline{r}'$ corresponding to an integration point on the u axis are needed, and these are given by

$$x' = x_a + \frac{u}{u_b} (x_b-x_a) \tag{6.30a}$$

$$y' = y_a + \frac{u}{u_b} (y_b-y_a) \quad . \tag{6.30b}$$

The numerically computed integral for I_1 (6.27a) is

$$I_{1N}=\int_0^{u_b} \left\{ \underline{f}^T(\underline{r}')\ G(\underline{r}|\underline{r}')+ \frac{1}{2\pi}\ \underline{f}^T(\underline{r}'')\ \ln\ \sqrt{v_p^{\,2}+(u-u_p)^2} \right\} du \tag{6.31a}$$

and for I_2 (6.27b) it is

$$I_{2N} = \int_0^{u_b} \left\{ \underline{f}^T(\underline{r}')\ \frac{\partial G(\underline{r}|\underline{r}')}{\partial n'} - \frac{1}{2\pi}\ \underline{f}^T(\underline{r}'')\ \frac{v_p}{(v_p^{\,2}+(u-u_p)^2)} \right\} du \ ;\ v_p\neq 0 \ . \tag{6.31b}$$

The analytic integral for Q has already been given (4.27), and for $\partial Q/\partial n$ (6.29) it is computed by substituting $t=u-u_p$:

$$\frac{1}{2\pi}\int_0^{u_b} \frac{v_p\ du}{(v_p^{\,2}+(u-u_p)^2)} = \frac{1}{2\pi}\int_{-u_p}^{u_b-u_p} \frac{v_p\ dt}{v_p^{\,2}+t^2} = \frac{1}{2\pi}\left\{ \tan^{-1}\frac{u_p}{v_p} + \tan^{-1}\frac{u_b-u_p}{v_p} \right\}$$

$$;\ v_p \neq 0 \tag{6.32}$$

It is now necessary to specify G, $\partial G/\partial n$ for each case to be studied, and to ensure that Q contains the singularity. Consider the problem having uniform cross-section. In this case G=Q and no problems arise as the

numerical integrations (6.31) are readily performed. For the time-harmonic case, the Green's function is specified at (3.6) and involves Hankel functions. The small argument formula for the Hankel function [2, p.139] may be used to show that P=G-Q is regular at $\underline{r}=\underline{r}'$:

$$\lim_{\underline{r}\to\underline{r}'} P(\underline{r}|\underline{r}') = \lim_{\underline{r}\to\underline{r}'}\left(\frac{-j\, H_o(k|\underline{r}-\underline{r}'|)}{4} + \frac{1}{2\pi}\ln|\underline{r}-\underline{r}'|\right) = \frac{1}{2\pi}\left(\ln\frac{k}{2}+\gamma\right)-\frac{j}{4} \tag{6.33}$$

where Euler's constant is $\gamma=0.5772157$. For $k>0$ this limit is finite, and the above algorithm may be applied directly. If $\underline{r}=\underline{r}'$, the integrand of (6.31a) is replaced by

$$\underline{f}(\underline{r}')\left\{\frac{1}{2\pi}\left(\ln\frac{k}{2}+\gamma\right)-\frac{j}{4}\right\} \quad .$$

For the integral given at (6.31b) the $\underline{r}$ and $\underline{r}'$ will coincide only at the numerical integration points on S_i. The integrand vanishes (i.e. $\cos\theta = -v_p/\sqrt{v_p^2+(u-u_p)^2} = 0$) for any observation point on S_i, including $\underline{r}=\underline{r}'$.

For axially symmetric problems the discussion in Section 4.6 is pertinent. The function $p(\underline{r})$ in (6.24) and (6.25) is replaced by $2\pi\eta\, p(\underline{r})$, and the functions G, $\partial G/\partial n$ in (6.31) must be specified.

First of all, from (1.10) it is evident that the singularity is contained in the Laplacian for three-dimensional problems, including those with axial symmetry. The function to be used for G in (6.31) is given by

$$G(\underline{r}|\underline{r}') = \eta'\int_0^{2\pi}\frac{d\,\psi}{4\pi|\underline{r}-\underline{r}'|} + 2\eta'\int_0^{\pi}\frac{e^{jk|\underline{r}-\underline{r}'|}-1}{4\pi|\underline{r}-\underline{r}'|}\,d\,\psi \quad . \tag{6.34}$$

The second integrand here is regular; the integral may be computed to six figure accuracy using order 8 Gaussian quadrature. It vanishes when $\underline{r}=\underline{r}'$. The first integral has already been obtained, at (4.97). The first term in (6.34) is

$$\eta' \, K(\xi)/\pi\chi$$

where ξ and χ are defined at (4.98), (4.99), and are readily computed for the numerical integration points in (6.31a). It has been demonstrated at (4.104) that Q contains the singularity. At $\underline{r}=\underline{r}'$, the integrand of (4.31a) is replaced by

$$\underline{f}(\underline{r}')^T \, \frac{\ln 8\eta'}{2\pi} \, .$$

The normal derivative is more complicated. It may be formulated as follows for (6.31b) :

$$\frac{\partial G(\underline{r}|\underline{r}')}{\partial n'} = \eta' \int_0^{2\pi} \frac{\partial}{\partial n'} \left\{ \frac{e^{jk|\underline{r}-\underline{r}'|}}{4\pi|\underline{r}-\underline{r}'|} \right\} d\psi = \eta' \int_0^{2\pi} \frac{-\cos\theta \, d\psi}{4\pi|\underline{r}-\underline{r}'|}$$

$$+ 2\eta' \int_0^{\pi} \frac{\{e^{jk|\underline{r}-\underline{r}'|}(jk|\underline{r}-\underline{r}'|-1) + 1\} \cos\theta \, d\psi}{4\pi|\underline{r}-\underline{r}'|^2} \, . \qquad (6.35)$$

The second integrand is regular: as

$$\lim_{d\to 0} \frac{1}{4\pi d^2} (e^{jkd}(jkd-1) + 1) = \frac{1}{4\pi d^2} \{ (1+jkd - \frac{k^2d^2}{2})(jkd-1) + 1\}$$

$$= \frac{-k^2}{8\pi} \, . \qquad (6.36)$$

Numerical quadrature may be employed directly to evaluate the second integral. The first integral is given by Silvester [3, pp.154-156] and

may be put in the form

$$\eta' \int_0^{2\pi} \frac{-\cos\theta \, d\psi}{4\pi|\underline{r}-\underline{r}'|} = -\eta' \frac{E(\xi)\cos\theta}{\pi\chi|\underline{r}-\underline{r}'|} + n'_x \eta' \left\{ \frac{E(\xi) - K(\xi)}{2\pi \eta \chi} \right\} \qquad (6.37)$$

where E is the complete elliptic integral of the second kind and n'_x is the direction cosine of the normal to S in the η direction at the source point $\underline{r}'$, as in Fig. 6.2. Both terms in (6.37) are singular, so that (6.31b) is not applicable without modification; the following form is used instead:

$$I_{2N} = \int_0^{u_b} \{\underline{f}^T(\underline{r}') \frac{\partial G(\underline{r}|\underline{r}')}{\partial n'} - \frac{1}{2\pi} \underline{f}^T(\underline{r}'') \frac{v_p}{\{v_p^2+(u-u_p)^2\}}$$

$$- \frac{n'_x}{4\pi\eta'} \underline{f}^T(\underline{r}'') \ln \sqrt{v_p^2+(u-u_p)^2} \}du \qquad (6.38)$$

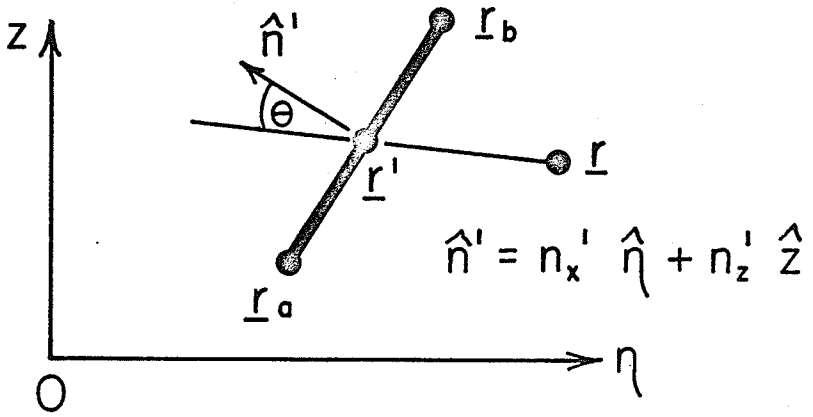


Fig. 6.2 Axisymmetric Geometry for Normal Derivative Determination

and the analytic integrals to be added are given by

$$I = \underline{f}^T(\underline{r}'')\left\{\int_{S_i} \frac{\partial Q(\underline{r}|\underline{r}')}{\partial n'}\, ds' - \frac{n_x'}{2\eta'}\int_{S_i} Q(\underline{r}|\underline{r}')ds'\right\}; \qquad (6.39)$$

the integrals in (6.39) have already been obtained.

At the singular point $\underline{r}\to\underline{r}'$; $\xi=1$. In this case, as in Section 4.6, $\chi=2\eta'$, and $K(\xi)=\ln 8\eta' - \ln\sqrt{v_p^2+(u-u_p)^2}$. Therefore only (6.37) contributes to$\partial G/\partial n$. The integrand of (6.38) becomes

$$\underline{f}^T(\underline{r}')\left\{\frac{-\eta'}{2\pi\eta'}\cdot\frac{-v_p}{\{v_p^2+(u-u_p)^2\}} + \frac{n_x'\,\eta'}{4\pi\eta'\,\eta'}(1-\ln 8\eta' + \ln\sqrt{v_p^2+(u-u_p)^2})\right.$$

$$\left. - \frac{1}{2\pi}\cdot\frac{v_p}{\{v_p^2+(u-u_p)^2\}} - \frac{n_x'}{4\pi\eta'}\ln\sqrt{u_p^2+(u-u_p)^2}\right\}$$

$$= \underline{f}^T(\underline{r}')\ \frac{n_x'}{4\pi\eta'}(1-\ln 8\eta') \quad . \qquad (6.40)$$

This is regular and is used as the integrand when $\underline{r}=\underline{r}'$ at the numerical integration points.

The integrals required to complete the matrix equation (6.22) may now be computed for each case to be considered.

The choice of line elements for the S_i has been made to facilitate computation of the analytic integrals required. It also allows the normal direction to S_i to be computed but once. If curved elements are required, here is one way of computing the integrals of Q and $\partial Q/\partial n$; suppose the curved element is represented by the equation w(u), such that w=0 if the element is linear, as shown in Fig. 6.3. The integral of Q may be obtained as two of the form

$$\int_0^a \ln \sqrt{u^2+(w(u)-v_p)^2}\, du = \int_0^a \ln \sqrt{\frac{u^2+(w(u)-v_p)^2}{u^2+(w(o)-v_p)^2}}\, du + \int_0^a \ln \sqrt{u^2+(w(o)-v_p)^2}\, du \,. \tag{6.41}$$

The first integrand is regular, permitting numerical integration, and the second integral may be handled by the methods presented previously, at (4.27). The integral of $\partial Q/\partial n$ may be obtained as two of the form

$$\int_0^a \frac{(v_p-w(u) + u \frac{\partial w(u)}{\partial n})\, du}{\{u^2+(v_p-w(u))^2\}\{1+(\frac{\partial w(u)}{\partial n})^2\}^{1/2}} =$$

$$\int_0^a \left\{ \frac{(v_p-w(u)+u \frac{\partial w(u)}{\partial u})}{(u^2+(v_p-w(u))^2)(1+(\frac{\partial w(u)}{\partial u})^2)^{1/2}} - \frac{v_p - w(o)}{(u^2+(v_p-w(o))^2)\ (1+(\frac{\partial w(o)}{\partial u})^2)^{1/2}} \right\} du$$

$$+ \frac{1}{(1+(\frac{\partial w(o)}{\partial u})^2)^{1/2}} \int_0^a \frac{(v_p-w(o))\, du}{(u^2+(v_p-w(o))^2)} \,. \tag{6.42}$$

The first integrand is regular, and the second may be obtained analytically.

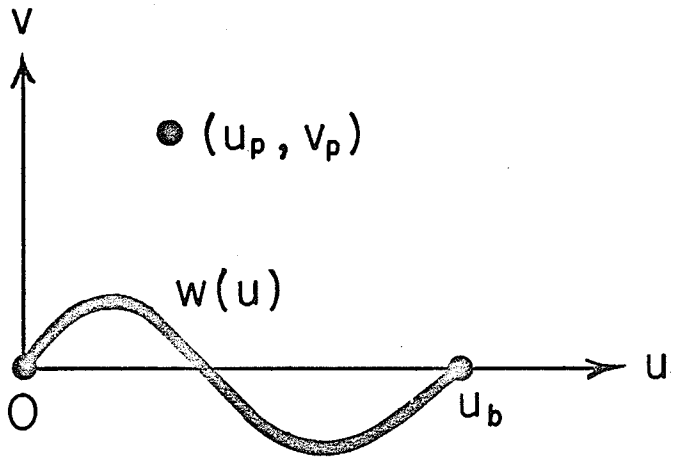


Fig. 6.3 Possible Local Coordinates for a Curved IE Element

6.3 The Mutual Constraint

The PDE functional form is examined first. Consider picture-frame Ω_i with an inhomogeneous Neumann boundary condition defined by

$$\hat{n} \cdot (\varepsilon \nabla \phi(\underline{r})) = -\sigma(\underline{r}) \; ; \; \underline{r} \text{ on } S_i \; . \tag{6.43}$$

Here $\hat{n}$ points out of the picture-frame and the negative sign on σ is used, because in the previous section the normal derivative was taken pointing into the picture-frame. The PDE functional (2.48) in the case of uniform cross-section, then, is

$$F(\phi) = \int_{\Omega_i} \{\nabla\phi^* \cdot (\varepsilon\nabla\phi) - W\phi^*\phi - p\phi^* - p^*\phi\} d\Omega + \int_{S_i} \{\sigma^*\phi + \phi^*\sigma\} \, ds \tag{6.44}$$

and node potentials are fixed to any Dirichlet conditions specified within Ω_i. It is easily verified that at the stationary point the Helmholtz equation (1.2) is satisfied in Ω_i and (6.43) is satisfied on S_i. The region Ω_i and the curve S_i are in the (x,y) plane. As indicated in Chapter I, the propagation function W is scalar, whereas ε may be either scalar or tensor depending upon the properties of the material in Ω_i.

For the axi-symmetric case it is simply necessary to remember that the problem is three-dimensional with no dependence on the angle ψ in (η,ψ,z) coordinates. The volume element is $\eta d\Omega$ and the surface element is ηds, where Ω_i and S_i are a region and a curve, respectively, in the (η,z) plane. The functional becomes:

$$\begin{aligned} F(\phi) &= \int_0^{2\pi} \int_{\Omega_i} \eta \{\nabla\phi^* \cdot (\varepsilon\nabla\phi) - W\phi^*\phi - p\phi^* - p^*\phi\} d\Omega d\psi + \int_0^{2\pi} \int_{S_i} \eta\{\sigma^*\phi + \phi^*\sigma\} ds \, d\psi \\ &= 2\pi \int_{\Omega_i} \eta\{\nabla\phi^* \cdot (\varepsilon\nabla\phi) - W\phi^*\phi - p\phi^* - p^*\phi\} d\Omega + 2\pi \int_{S_i} \eta\{\sigma^*\phi + \phi^*\sigma\} ds \; . \end{aligned} \tag{6.45}$$

A boundary relaxation procedure [4] may now be prescribed. If σ is known on S_i, the PDE functional may be used to find ϕ in Ω_i and on S_i. This potential on S_i may be used as a Dirichlet condition, and the IE functional may be used to find σ on S_i. This σ may, in turn, be used with the PDE functional to find ϕ - the iterative process is continued until σ and ϕ are found which satisfy both the PDE and IE problems. Potential is clearly continuous; the use of ϕ on S_i as a Dirichlet condition for the IE problem ensures this. Also, at solution,

$$\sigma = - \varepsilon_o \frac{\partial \phi}{\partial n_+} \tag{6.46}$$

where $\hat{n}$ points into Ω_e. (The IE problem was constructed with $\hat{n}$ pointing out of Ω_e, and $\sigma=\varepsilon_o \partial\phi/\partial n$; it is simply necessary to change the sign to obtain (6.46).) Then using (6.43), at the solution

$$\hat{n}\cdot(\varepsilon \nabla\phi(\underline{r})) = \varepsilon_o \frac{\partial \phi(\underline{r})}{\partial n_+} \quad ; \ \underline{r} \text{ on } S_i \ . \tag{6.47}$$

If ε is scalar, this becomes

$$\varepsilon_r \varepsilon_o \frac{\partial \phi(\underline{r})}{\partial n_-} = \varepsilon_o \frac{\partial \phi(\underline{r})}{\partial n_+} \tag{6.48}$$

and is simply the interface condition (2.65). With this formulation the picture-frame boundary S_i may be placed directly on an interface, and the proper condition results naturally. This is in contrast to the method in Chapter III, where a homogeneous region was imposed between S_i and a contour within the picture-frame S_{c_i}.

Convergence of boundary relaxation procedures depends upon positive-definite matrix properties [4], and for time-harmonic problems one could expect numerical problems to arise if $k=\lambda_i$, a picture-frame eigenvalue.

The problem may also be formulated as one of mutual constraint. As in Chapter III, the vector of node potentials $\underline{\phi}$ is introduced, and the

matrix equation is constructed using isoparametric finite-elements. Suppose, for simplicity, that σ on the S_i is represented by the same interpolatory function as ϕ on S_i. The resulting PDE matrix equation, for all Ω_i, is

$$[\,A\,]\,\underline{\phi} \;=\; -\,[A']\,\underline{\sigma} + \underline{b} \quad . \tag{6.49}$$

The matrix $[A]$ is symmetric and square. The matrix $[A']$ is constructed from the integrals over S_i in (6.44) or (6.45), and is not square; it has a number of rows equal to the total number of nodes in all the Ω_i, and a number of columns equal to the number of nodes on all the S_i. The vector $\underline{b}$ arises from the integrations of p (source) over Ω_i, and also from imposed Dirichlet node potentials in the Ω_i. If the Neumann condition were known, $\underline{\sigma}$ would be available, and the right side of (6.49) would be completely known:

$$\begin{aligned} [\,A\,]\;\underline{\phi} \;&=\; -\,[A']\;\underline{\sigma} + \underline{b} \\ &=\; \underline{b}' + \underline{b} \\ &=\; \underline{b}'' \quad . \end{aligned} \tag{6.50}$$

Solution of (6.50) would solve the Neumann problem in each and every Ω_i. (In the boundary relaxation procedure, (6.50) is used to find ϕ using the σ obtained from the previous IE solution for Ω_e. $[A]$ would be singular if $k=\lambda_i$, a picture-frame eigenvalue).

Turning now to the IE problem, one may write (6.22) using the same node points on $S=\Sigma\, S_i$, and the same interpolatory functions for $\underline{\sigma}$, $\underline{g}$ on S (see equation (6.26)):

$$[\,D\,]\,\underline{\sigma} \;=\; -\,[\,D''\,]\;\underline{g} \quad . \tag{6.51}$$

The matrices $[D]$ and $[D'']$ are both square, but neither is Hermitian in

the time-harmonic case (complex integral operator). The nature of the PDE isoparametric formulation is such that potential on a straight side triangle boundary is an interpolate only of the node potentials on that side; the other node potentials in the element make no contribution [5]. Therefore, without loss of generality, one may rewrite (6.51) as

$$[D]\ \underline{\sigma} = -[D']\ \underline{\phi} \tag{6.52}$$

where $\underline{\phi}$ is the vector of all node potentials in all Ω_i. The matrix $[D']$ is not square. It still has rows equal to the number of nodes on $S = \Sigma\ S_i$, but it now has columns equal to the total number of nodes, with zero entries being added. If $\underline{\phi}$ is known, so is $\underline{g}$, the vector of node potentials on S. Therefore a known vector $\underline{b}$ may be constructed:

$$\underline{b} = -[D'']\ \underline{g} \equiv -[D']\ \underline{\phi} \tag{6.53}$$

and (6.51) becomes

$$[D]\ \underline{\sigma} = \underline{b}\ . \tag{6.54}$$

This equation is the solution to the Dirichlet problem in Ω_e, and in the boundary relaxation procedure is used to find $\underline{\sigma}$ employing $\underline{\phi}$ from the previous PDE solution.

Equations (6.49) and (6.52) may be solved simultaneously by mutual constraint:

$$\begin{bmatrix} A & A' \\ D' & D \end{bmatrix} \begin{pmatrix} \underline{\phi} \\ \underline{\sigma} \end{pmatrix} = \begin{pmatrix} \underline{b} \\ \underline{0} \end{pmatrix} \tag{6.55}$$

provided that, the propagation constant k is not an eigenvalue of the coefficient matrix. In this regard, the coefficient matrix is not Hermitian, and the eigenvalues, in general, will be complex numbers. The problem will arise only if by chance the matrix has a real eigen-

value λ_i and $k=\lambda_i$. Quite clearly there is some work to be done to determine under what conditions (6.53) has a unique solution, but no difficulties have been encountered in any of the test examples.

Every finite-element - each PDE isoparametric element in the Ω_i for potential representation and each IE line element on the S_i for source representation - possesses an operator with a self-adjoint, positive-definite component, the Laplacian. Convergence is guaranteed in energy; over the picture-frames:

$$U_1 = \sum_i \int_{\Omega_i} \nabla\phi^* \cdot (\varepsilon \nabla\phi) d\Omega \tag{6.56}$$

and over the picture-frame boundaries:

$$U_2 = \sum_{ij} \frac{1}{\varepsilon_o} \int_{S_i} \int_{S_j} \sigma(\underline{r})^* \sigma(\underline{r}') G_L(\underline{r}|\underline{r}') ds' ds \tag{6.57}$$

where G_L is the Laplacian Green's function for the problem.

Of course it is not necessary to use the same interpolatory functions to represent σ and ϕ on the S_i; one may choose almost any integrable functions in linear combination to represent σ. It is convenient, however, to use the same representation, in that only one set of functions need be dealt with in computation.

It is interesting to consider the methods presented by Silvester and Hsieh [6] and Williams and Cambrell [7] within the present context. Each involves a PDE matrix equation (for the static problem) using finite-elements, similar to (6.49). Each also involves an IE matrix equation similar to (6.53) using, in [6] a projective Galerkin method with the incomplete Green's function (see equation (3.10)), and in [7] , a point-matching technique with the complete Green's function. Each then involves an inversion of the [D] matrix:

$$\underline{\sigma} = - [D]^{-1} [D'] \underline{\phi} \ . \tag{6.58}$$

Substitution of (6.58) into (6.49), gives an equation for $\underline{\phi}$:

$$[A - A'D^{-1}D'] \underline{\phi} = \underline{b} \ . \tag{6.59}$$

This reduces the number of unknowns, but at the expense of a matrix inversion and two matrix multiplications. The point of similarity between these two methods and the present one is that each involves the construction of two matrix equations relating $\underline{\sigma}$ and $\underline{\phi}$.

6.4 Twin Cylinders and the Doughnut

Consider the two infinitely long cylinders shown in Fig. 6.4a. With symmetry about the y=0 plane and antisymmetry about the x=0 plane, solution is required only in the positive quadrant and a picture-frame with 16 triangular elements is placed as shown in Fig. 6.4b. Four of the elements have curved sides (arcs of a circle) and on these sides the Dirichlet condition $\phi=1$ is enforced. The symmetries are handled in computation by allowing the natural PDE condition $\partial\phi/\partial n$ to apply at x=0 in the picture-frame, and the IE integrals are handled by the method described in Appendix B. Each of the PDE elements is handled by use of the isoparametric technique of Section 2.6.

With N=2 there are 43 node potentials to be found in the picture-frame, and 9 node sources to be found on the picture-frame boundary S.

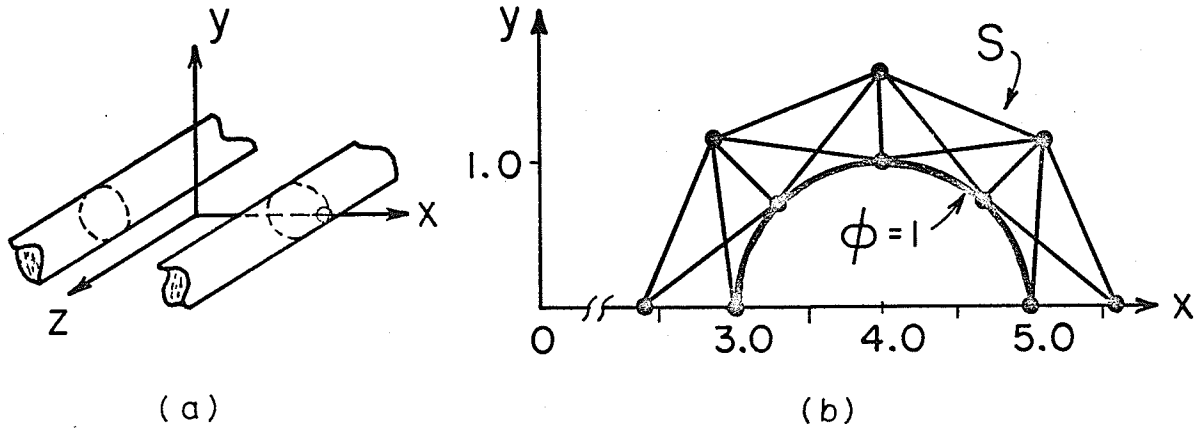


Fig. 6.4 Twin Cylinders.
a) 3-D Full View
b) Reduced Picture-Frame

The matrix equation (6.15) therefore involves 52 unknowns. Order 4 Gaussian quadrature was employed for all numerical integration in the IE portion. The resulting static equipotentials are shown in Fig. 6.5, and these are seen to be continuous at the picture-frame boundary.

This problem has an analytical solution [8, p.178] and for points on the x axis is given by

$$\phi_e(x) = \frac{1}{\ln(4-\sqrt{15})} \ln\left|\frac{x-\sqrt{15}}{x+\sqrt{15}}\right| ; \quad \begin{matrix} 0\leq x\leq 3 \\ x\geq 5 \end{matrix} . \qquad (6.60)$$

The experimental results are compared to the exact solution in Table 6.1. The relative error norm $||\phi-\phi_e|| \; / ||\phi_e||$ is obtained using trapezoidal integration at 41 points on the x axis $0\leq x\leq 3$; $5\leq x\leq 8$ and is less than 0.25 percent.

The problem may also be solved by boundary relaxation. If $\underline{\phi}^{(n)}$, $\underline{\sigma}^{(n)}$ are the node potentials and sources respectively at the n^{th}

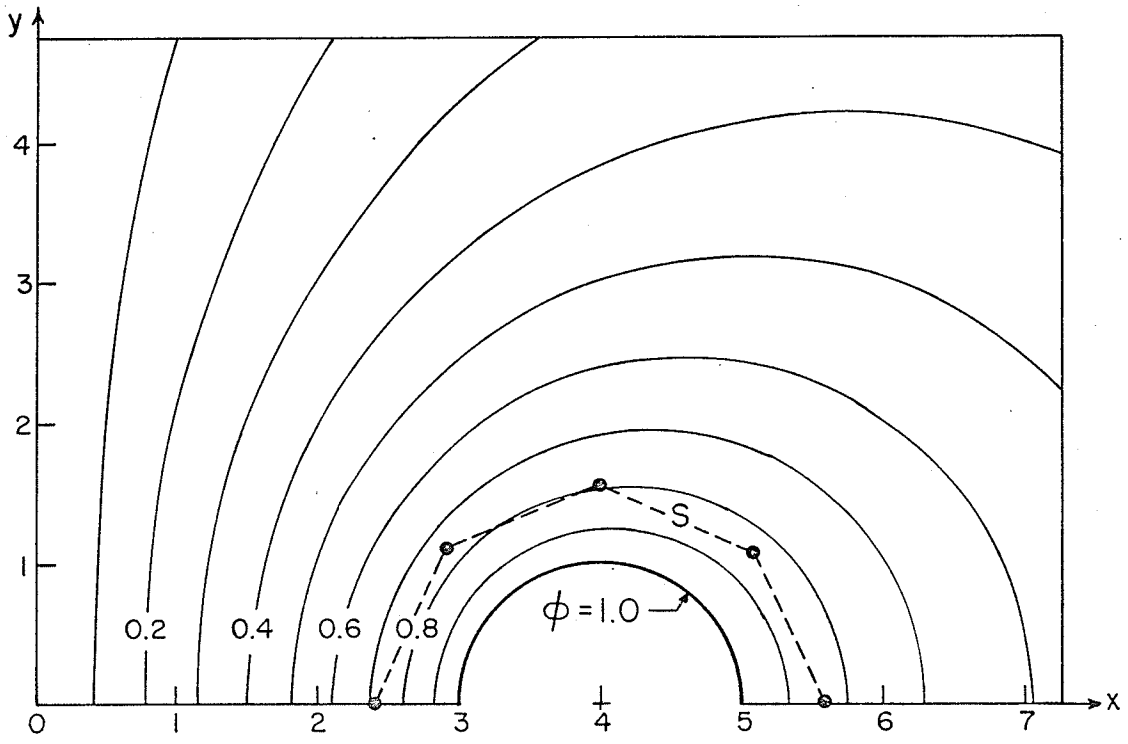


Fig. 6.5 Twin Cylinders: Static Equipotentials

TABLE 6.1 THE TWIN CYLINDERS:
PDE-IE STATIC SOLUTION, N=2

x	Potential ϕ at (x,0) computed	exact	|% error| in potential
0.0	0.0	0.0	0.0
1.0	0.256610	0.256056	0.216
2.0	0.555673	0.533848	0.329
3.0	1.0	1.0	0.0
5.0	1.0	1.0	0.0
6.0	0.744748	0.743944	0.108
7.0	0.604414	0.603945	0.078
8.0	0.512522	0.512112	0.080
% relative error norm	0.222		

iteration, the updated values for iteration n+1 are given by

$$\underline{\phi}^{(n+1)} = (1-\omega)\underline{\phi}^{(n)} + \omega[A^{-1}]\{\underline{b} - [A']\ \underline{\sigma}^{(n)}\} \tag{6.61a}$$

and

$$\underline{\sigma}^{(n+1)} = (1-\omega)\underline{\sigma}^{(n)} - \omega[D^{-1}]\ [D']\ \underline{\phi}^{(n+1)} \tag{6.61b}$$

where ω is a relaxation factor [4, p.239] which may be adjusted to optimize convergence. The error at iteration n is obtained as a relative error norm with respect to the direct solution of the matrix equation,

$$e^{(n)} = \left\{ \sum_{i=1}^{N} |\phi_i{}^{(n)} - \phi_i| + \sum_{i=1}^{N} |\sigma_i{}^{(n)} - \sigma_i| \right\} / \left\{ \sum_{i=1}^{N} |\phi_i| + \sum_{i=1}^{N} |\sigma_i| \right\}. \tag{6.62}$$

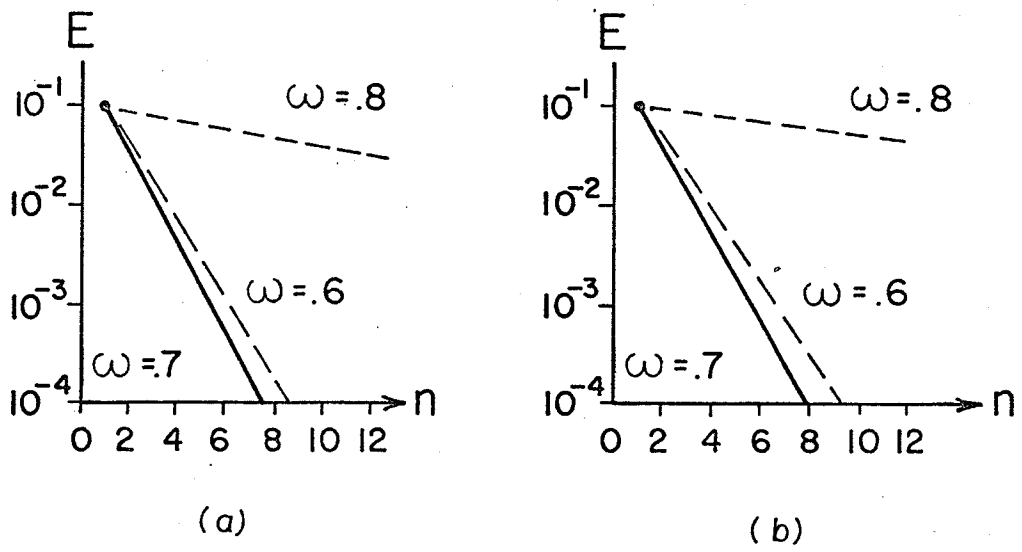


Fig. 6.6 Convergence of Boundary Relaxation for the Twin Cylinder Problem. a) Static case (k=0). b) Propdgation case (k=0.5). (E is the relative error norm described on page 185.)

Fig. 6.6a shows certain of the results obtained. It is interesting that under-relaxation is required for convergence (ω_{opt}=0.7); this was also observed by Sandy and Sage [9], although Cermak [10] was able to over-relax, probably because he relaxed each node point in the picture-frame in turn, rather than in the block fashion employed here. (As a check, when each individual node potential and source is updated in turn using latest available values, ω_{opt}=1,15 for the twin cylinder problem.)

The time-harmonic problem may be investigated by repeating the procedure with propdgation constant k=0.5 (corresponding to a frequency of $\sim$ 2.5 Ghz with dimensions in centimeters). The problem was solved directly and then by boundary relaxation. The real part of the solution

is used in (6.62) and the results are shown in Fig. 6.6b. Again ω_{opt}=0.7, and in fact, the curves are practically indistinguishable from those in Fig. 6.6a; identical convergence properties are obtained with boundary relaxation for the cases k=0, k=.5. This is not to say that the solutions are equally accurate, however; it simply relates to methods of solving the matrix equation (6.15).

The doughnut problem of Section 4.6 may be investigated here by using the same picture-frame (Fig. 6.4b) with symmetry about the y=0 plane, and the axi-symmetric functionals (6.45)(4.100). With N=2 there are, again, 52 unknowns, and order 4 Gaussian quadrature was used for all numerical integration in the IE portion. The resulting static equipotentials are shown in Fig. 6.7 and results are listed in Table 6.2, along with the best results obtained using purely IE functionals from Table 4.8. It is observed that the difference in the two solutions on

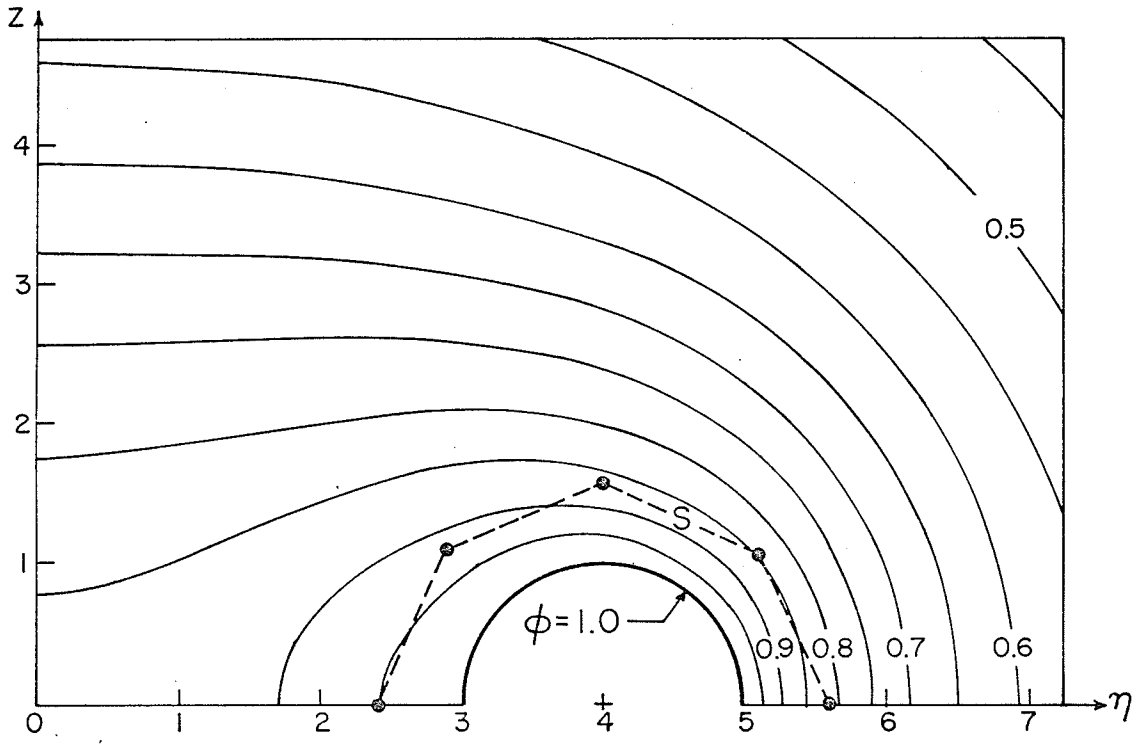


Fig. 6.7 The Doughnut: Static Equipotentials

TABLE 6.2 THE DOUGHNUT;
PDE-IE STATIC SOLUTION, N=2

x	Potential ϕ at (x,0) computed	pure IE solution (Table 5.5)	|difference| x 10^{-4}	|% variation| between approx solutions
0.0	0.863095	0.862792	3.03	0.035
1.0	0.874695	0.874401	2.94	0.034
2.0	0.914058	0.913840	2.18	0.024
3.0	1.0	0.999772	2.28	0.023
5.0	1.0	0.999009	9.91	0.099
6.0	0.736289	0.734964	13.25	0.180
7.0	0.596958	0.596449	5.09	0.085
8.0	0.506490	0.506264	2.26	0.045

the x axis is less than 0.2 percent; the present method works well for axi-symmetric problems.

The matrix equation for this problem may be solved by boundary relaxation, in a fashion identical to that for the twin cylinder problem. Equations (6.61) are used, and (6.62) is applied to the real parts of the solution variables, just as before. Convergence curves are shown in Figs. 6.8 . Again, the two sets of curves are virtually indistinguishable, with under-relaxation again required, and ω_{opt}=0.6.

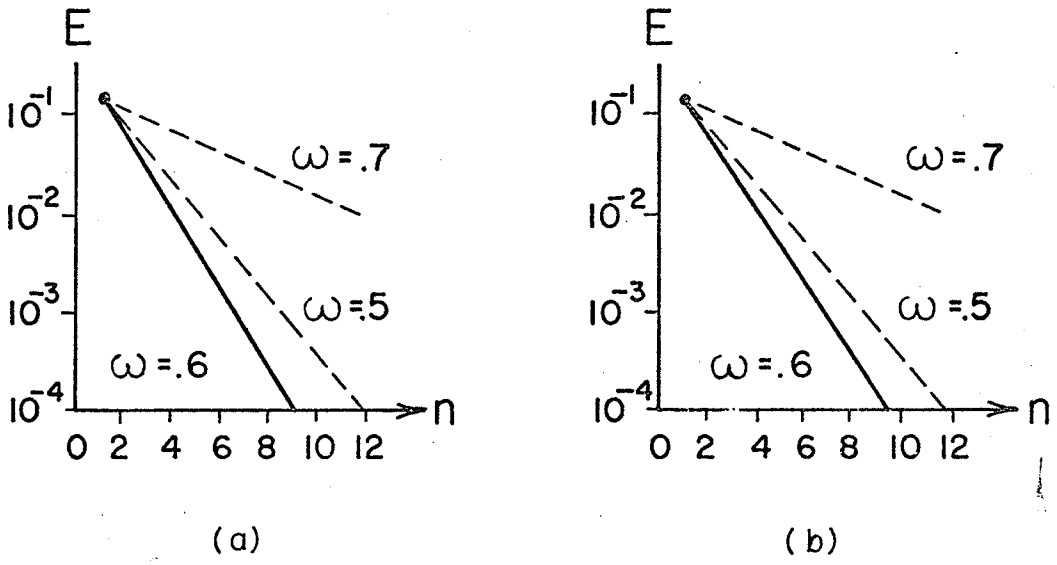


Fig. 6.8 Convergence of Boundary Relaxation for the Doughnut Problem.
a) Static case (k=0).
b) Propdgation case (k=0.5).
(E is the relative error norm described on page 185.)

6.5 Accuracy Tests with a Circular Cross-Section

Consider the picture-frame shown in Fig. 6.9. If one uses the PDE and IE uniform cross-section functionals (6.44), (6.3) with symmetry about the x=0, y=0 planes, the configuration represents an infinitely long cylinder, with radius 0.5, centered on the z axis. If one uses the axi-symmetric functionals (6.45), (4.100) with symmetry about the y=0 plane, a sphere centered at the origin is represented. The analytic solutions are all known.

The picture-frame has 12 triangular elements, two with curved sides on which a Dirichlet potential is to be specified. With N=2 there are 35 node potentials and 9 node sources (44 unknowns) to be obtained, and with N=3, there are 70 node potentials and 13 node sources (83 unknowns). Isoparametric techniques are used for the PDE elements, and numerical integrations for the IE elements are performed using order 4 Gaussian quadrature, with the symmetries included by the method given in Appendix B.

The sphere is investigated first. The solution is to satisfy

$$\phi(r) = \frac{e^{jkr}}{r} \qquad r \geq 0.5 \tag{6.63}$$

where r is the distance from the origin, and if (x,y) coordinates are used

$$r = \sqrt{x^2 + y^2} \quad . \tag{6.64}$$

The Dirichlet condition to be imposed is simply

$$\phi(0.5) = 2e^{j0.5k} \quad . \tag{6.65}$$

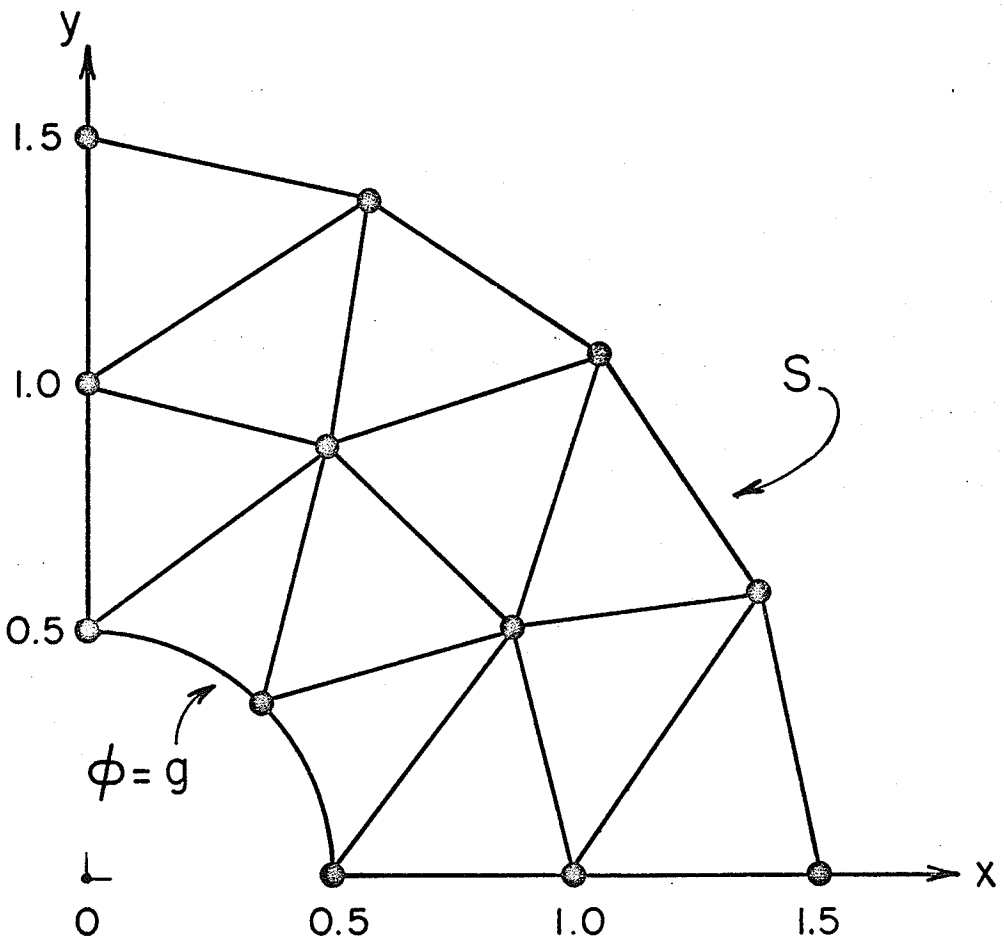


Fig. 6.9 Reduced Picture-Frame for Accuracy Tests Showing Finite-Element Placement.

The results obtained with k=0 (the static case) and with k=1.0 (representing a frequency of $\sim$ 5 GHz with dimensions in centimeters) are listed in Tables 6.3 and 6.4, respectively. The results are clearly very good, with relative error norms with N=4 for both ϕ over Ω, and σ over S, being less than 0.1 percent (real parts are used). The source σ is observed to be slightly less accurate than ϕ; that σ is integrated to obtain potential tends to smooth out local variations.

TABLE 6.3 THE SPHERE; PDE-IE SOLUTION, k=0

r	Potential ϕ at r N=2	N=3	exact	|% error| in ϕ, N=3
0.50	2.0	2.0	2.0	0.0
0.67	1.503677	1.501679	1.5	0.112
1.00	1.002521	1.001329	1.0	0.133
1.33	0.750023	0.750000	0.75	0.0
1.67	0.600597	0.600268	0.60	0.045
2.00	0.500578	0.500209	0.50	0.042
2.33	0.429011	0.428750	0.428571	0.042
2.67	0.375389	0.375156	0.375000	0.042
3.00	0.333684	0.333472	0.333333	0.042
% relative error norm	0.1561	0.0735		

r	source σ at r N=2	N=3	exact	|% error| in σ,N=3
1.50	0.682623	0.683264	0.683054	0.31
% relative error norm	0.1863	0.0914		

TABLE 6.4 THE SPHERE; PDE-IE SOLUTION, k=1

r	Re: Potential ϕ at r N=2	N=3	exact	|% error| in ϕ, N=3
0.50	1.755165	1.755165	1.755165	0.0
0.75	0.971063	0.973960	0.975585	0.167
1.00	0.541103	0.540629	0.540302	0.061
1.25	0.251871	0.252151	0.252258	0.042
1.50	0.047806	0.047400	0.047158	0.512
1.75	-0.101563	-0.101769	-0.101855	0.084
2.00	-0.208011	-0.208056	-0.208073	0.008
2.25	-0.279241	-0.279217	-0.279188	0.010
2.50	-0.320683	-0.320526	-0.320458	0.021
2.75	-0.336408	-0.336207	-0.336110	0.029
3.00	-0.330219	-0.330113	-0.329998	0.035
% relative error norm	0.2131	0.0775		

r	Re: source σ at r N=2	N=3	exact	|% error| in σ, N=3
1.5	0.435004	0.435585	0.435905	0.074
% relative error norm	0.2615	0.0983		

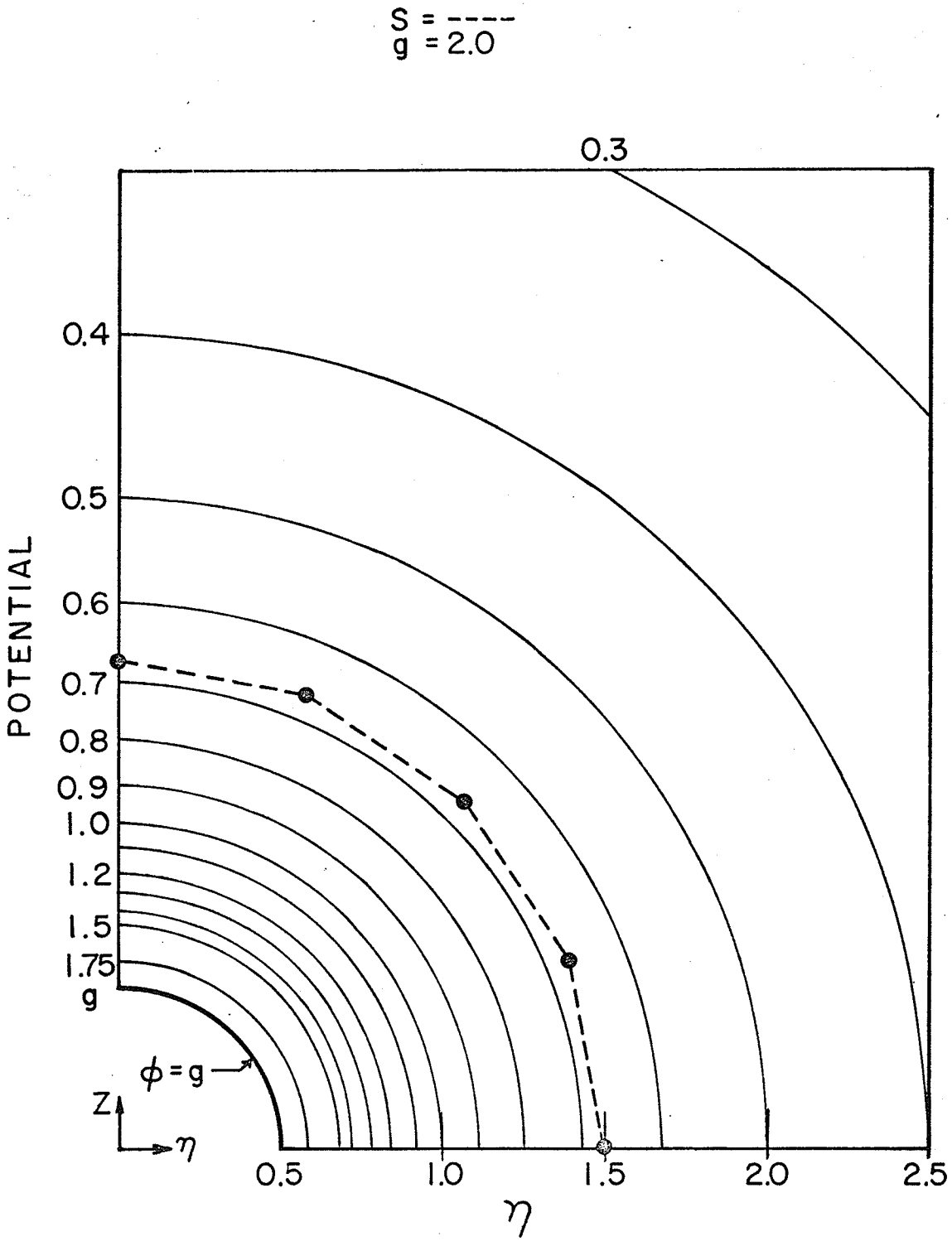


Fig. 6.10 The Sphere: Equipotentials, k=0

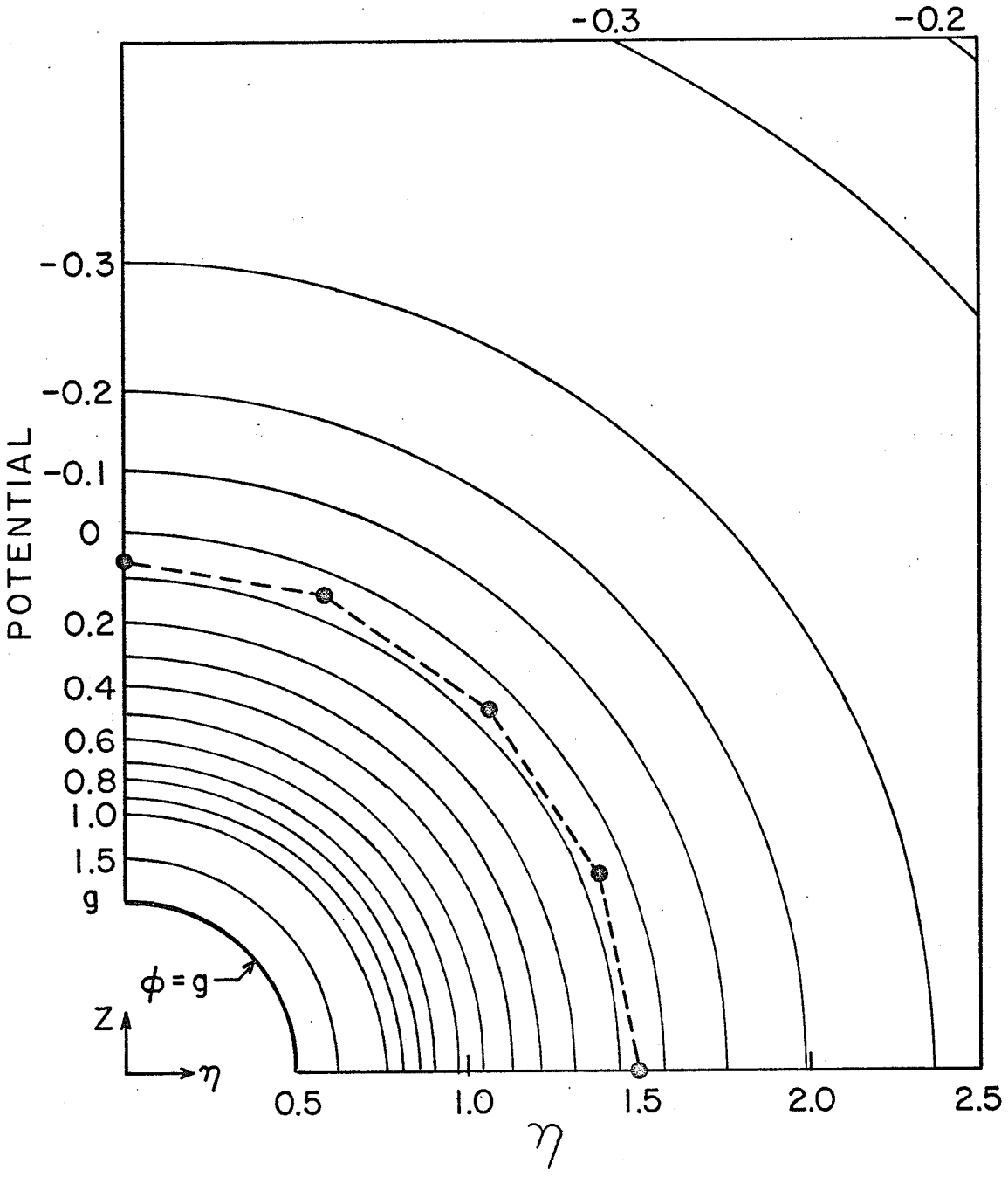


Fig. 6.11 The Sphere: Equipotentials (real part), k=1.0.

Equipotentials for the two cases are shown in Figs. 6.10 and 6.11. They are "exact", for errors of less than 0.1 percent cannot be observed.

Turning to the cylinder problem, the solution for the static case is to satisfy

$$\phi(r) = \ln(r) \qquad r \geq 0.5 \tag{6.66}$$

and the Dirichlet condition is

$$\phi(0.5) = \ln(0.5) .$$

The results in Table 6.5 were obtained with N=2 and 3. The solution is very good, with error norms of less than 0.05 percent. Again σ is less accurate than ϕ. Equipotentials are shown in Fig. 6.12.

For this problem the magnitude of the integral of $\partial\phi/\partial n$ over S is constant for a circular arc S',

$$I = \int_{S'} \frac{\partial\phi}{\partial n}\, ds = \int_0^{\frac{\pi}{2}} \frac{r d\psi}{r} = \frac{\pi}{2} = 1.57080 . \tag{6.67}$$

With ε_o=1, the integral of σ over the picture-frame boundary in the positive quadrant, obtained by order 4 Gaussian quadrature is 1.57042 (an error of -0.026 percent). The integral of $\partial\phi/\partial n$ over S, obtained by differentiating the PDE finite-element functions and using order 4 Gaussian quadrature, is 1.57231 (an error of +0.1 percent). Therefore one may deduce that σ is more accurate than $\partial\phi/\partial n$ - which is not surprising, for if σ and ϕ are both polynomial on S, $\partial\phi/\partial n$ is a polynomial of lower order. Alternatively, potential is obtained from σ by an integration involving G, and may be expected to be a higher order approximation. Therefore larger error norms for σ thanϕshould not be expected to degrade the solution for ϕ.

TABLE 6.5 THE CYLINDER; PDE-IE SOLUTION, k=0

r	Potential ϕ at r N=2	N=3	exact	|% error| in ϕ, N=3
0.50	-0,693147	-0.693147	-0.693147	0.0
0.67	-0.405862	-0.405594	-0.405465	0.032
1.00	-0.001217	-0.000504	0.0	-
1.33	0.288316	0.287903	0.287681	0.077
1.67	0.511123	0.510943	0.510826	0.023
2.00	0.693589	0.693325	0.693147	0.026
2.33	0.847837	0.847516	0.847298	0.026
2.67	0.981458	0.981082	0.980830	0.026
3.00	1.099312	1.098902	1.098610	0.026
% relative error norm	0.0776	0.0312		

r	source σ at r N=2	N=3	exact	|% error| in σ, N=3
1.5	-0.653317	-0.653639	-0.653857	0.032
% relative error norm	0.0916	0.0421		

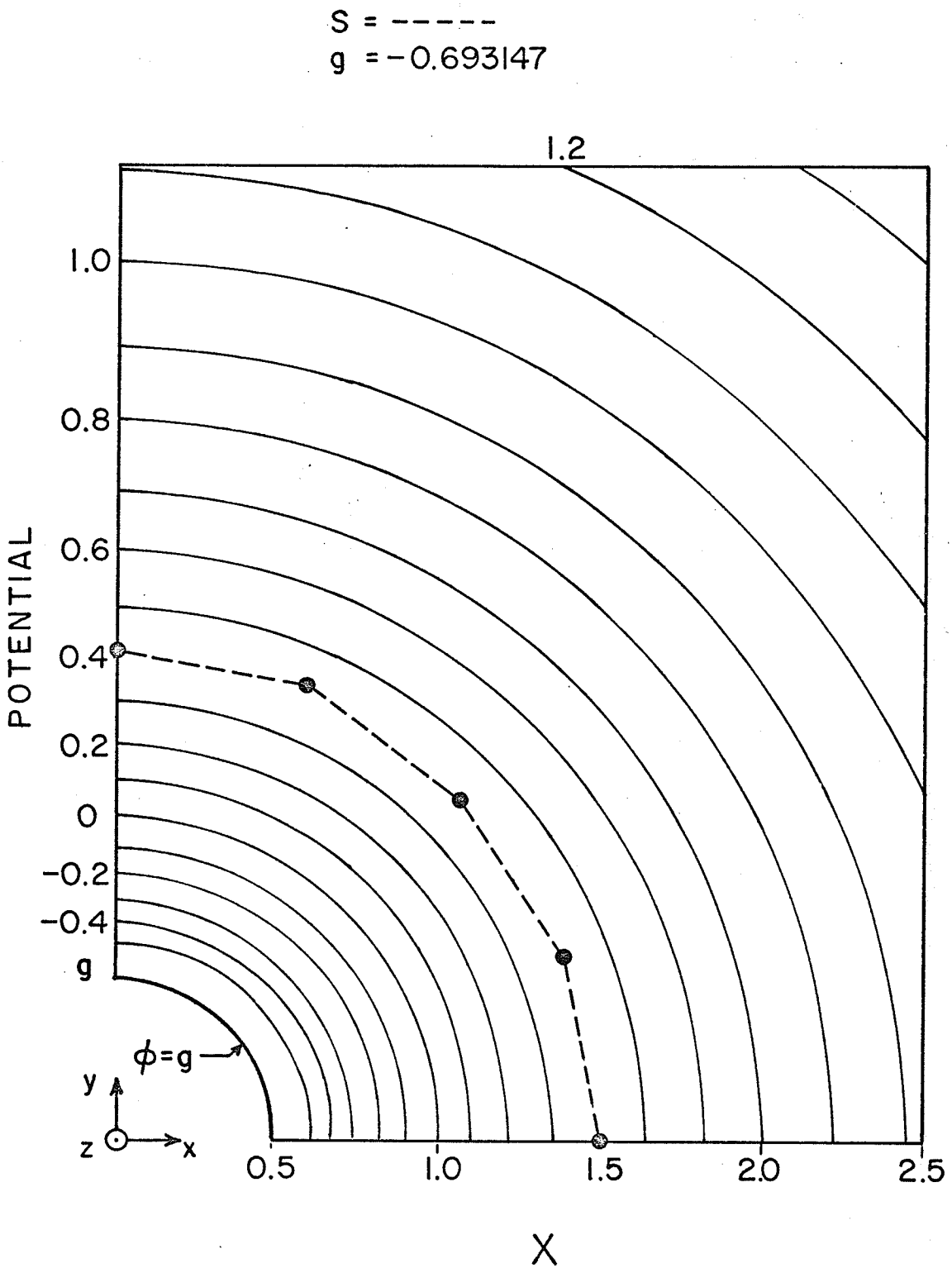


Fig. 6.12 The Cylinder: Equipotentials, k=0

For the time-harmonic cylinder problem, the solution is to satisfy

$$\phi(r) = - H_o (kr) \quad r \geq 0.5 \tag{6.68}$$

and the Dirichlet condition is

$$\phi(0.5) = - H_o (0.5k). \tag{6.69}$$

Two examples are given: k=1 (representing a frequency of $\sim$5 GHz with dimensions in centimeters), and k=2 (a frequency of $\sim$ 10 GHz). The results (real parts of ϕ and σ) are listed in Tables 6.6 and 6.7, and the equipotentials are shown in Fig. 6.13 for the case k=1. For this case the relative error norms are both less than 0.05 percent (N=3) indicating a very good solution. The results with k=2 are obtained with N=2 and show somewhat more error (0.3 percent relative error norm for ϕ). Nevertheless the results are still good, and such errors are virtually zero for engineering purposes.

The examples presented in this section show that highly accurate solutions can be achieved with the method, but there is no assurance that such accuracy will always be achieved with so few unknowns in other problems. There is the guarantee of energy convergence, however, so that including more unknowns (more finite-elements, higher order approximation) will provide better results. The results of Chapter IV show that the number of unknowns required for a specified accuracy depends to a large extent on the ability of the trial functions to model the true solution.

TABLE 6.6 THE CYLINDER; PDE-IE SOLUTION, k=1.0

	Re: potential ϕ at r			|% error| in ϕ, N=3
r	N=2	N=3	exact	
0.50	0.444519	0.444519	0.444519	0.0
0.75	0.138400	0.137416	0.137173	0.177
1.00	-0.087871	-0.088113	-0.088257	0.163
1.25	-0.258098	-0.258187	-0.258217	0.012
1.50	-0.383188	-0.382527	-0.382449	0.002
1.75	-0.465446	-0.465483	-0.465493	0.002
2.00	-0.510405	-0.510370	-0.510376	0.001
2.25	-0.520191	-0.520107	-0.520065	0.008
2.50	-0.498279	-0.498151	-0.498070	0.016
2.75	-0.448929	-0.448971	-0.448659	0.070
3.00	-0.377161	-0.376894	-0.376858	0.009
% relative error norm	0.0890	0.0329		

	Re: source σ at r			|% error| in σ, N=3
r	N=2	N=3	exact	
1.5	0.403761	0.404618	0.404386	0.057
% relative error norm	0.1007	0.0446		

TABLE 6.7 THE CYLINDER; PDE-IE SOLUTION
WITH N=2, k=2.0

r	Re: Potential ϕ at r (computed)	Re: Potential ϕ at r (exact)	|% error| in ϕ
0.50	-0.088257	-0.088257	0.0
0.75	-0.382361	-0382449	0.023
1.00	-0.509513	-0.510376	0.169
1.25	-0.496622	-0.498070	0.291
1.50	-0.378818	-0.376850	0.522
1.75	-0.19)244	-0.189022	0.646
2.00	0.016634	0.016941	1.810
2.25	0.193894	0.194705	0.416
2.50	0.308144	0.308518	0.121
2.75	0.339582	0.339481	0.630
3.00	0.288702	0.288195	0.176
% relative error norm	0.303		

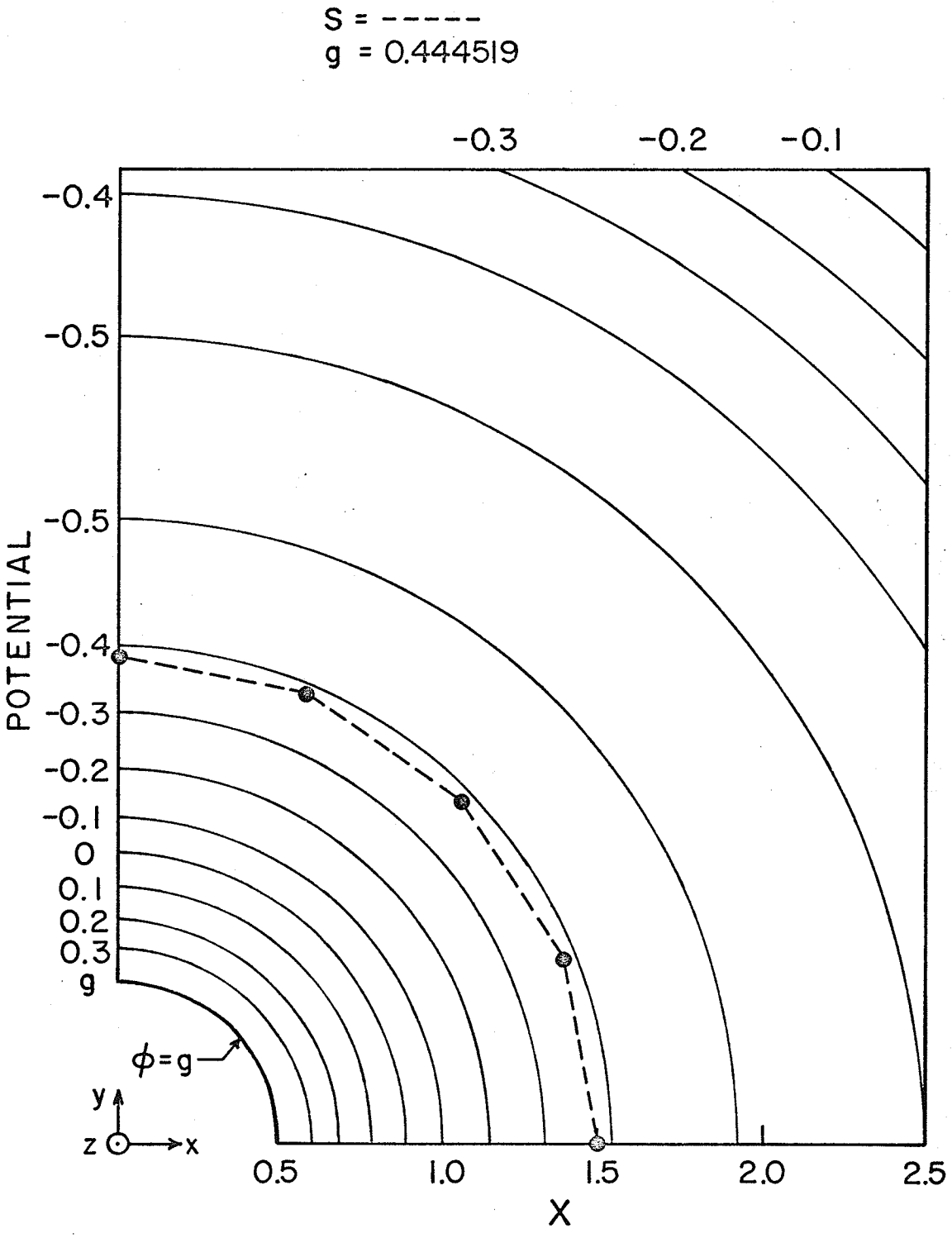


Fig. 6.13 The Cylinder: Equipotentials (real part), k=1.0.

6.6 The Strip Capacitor

The example problem of Section 3.3 is discussed again to show that the IE portion of this formulation may be used alone to solve the static exterior problem. Here the symmetries are employed as before to reduce the problem to the positive quadrant, and in contrast to the earlier examples, the strip is given a small thickness, so that S excises a finite region from the infinite plane. Five straight-line elements for S are specified as shown in Fig. 6.14, and continuous σ is sought, represented by a polynomial interpolate of node sources, in the same fashion as in the previous sections. No special singular function is used for the edge condition.

If N is the order of the polynomial approximation, there are

$$M = 7 + 6(N-1) = 6N+1 \tag{6.70}$$

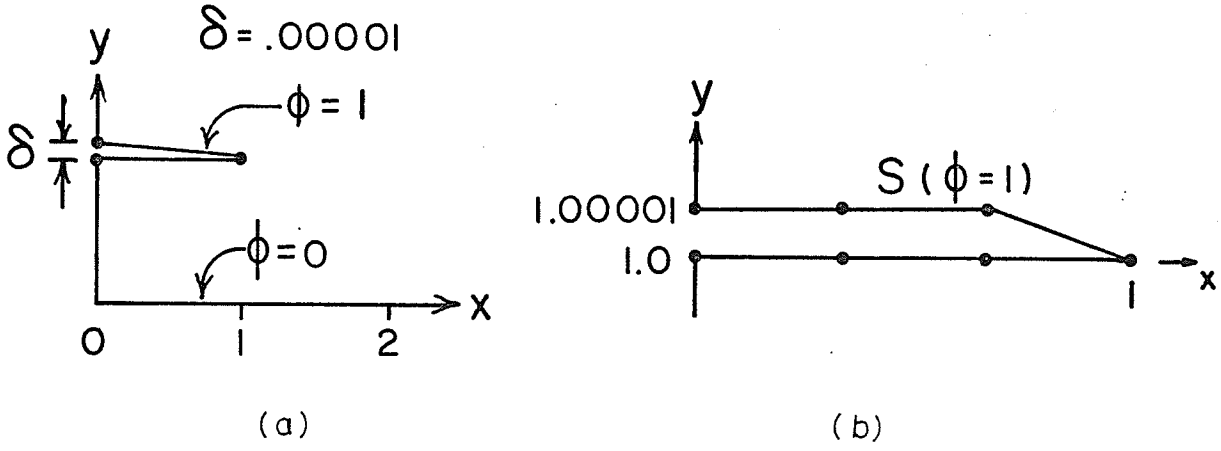


Fig. 6.14 The Thick Strip Capacitor
a) The Reduced Geometry
b) Finite-Element Placement on S.

node sources to be obtained. The matrix equation (6.51) is constructed with ϕ=1 on S, and the symmetries are handled by the method in Appendix B. Order 4 Gaussian quadrature was used for all numerical integrations. The results obtained are listed in Table 6.8, and the solution is almost 4 figures accurate with N=4 (25 unknowns).

TABLE 6.8 THE STRIP CAPACITOR; PDE-IE STATIC SOLUTION

Polynomial order	Number of nodes	Capacitance pf/M	Potential ϕ(40/9,10)
1	7	18.63	0.1118
2	13	18.68	0.1121
3	19	18.70	0.1122
4	25	18.71	0.1123

In Fig. 6.15 convergence properties for this approximation are compared to those with pulse function approximation, and to those with special singular functions, which were obtained in Section 4.4. Continuous polynomial representation is better than pulses alone, but the greatest gains are made using singular functions, as described in Chapter IV.

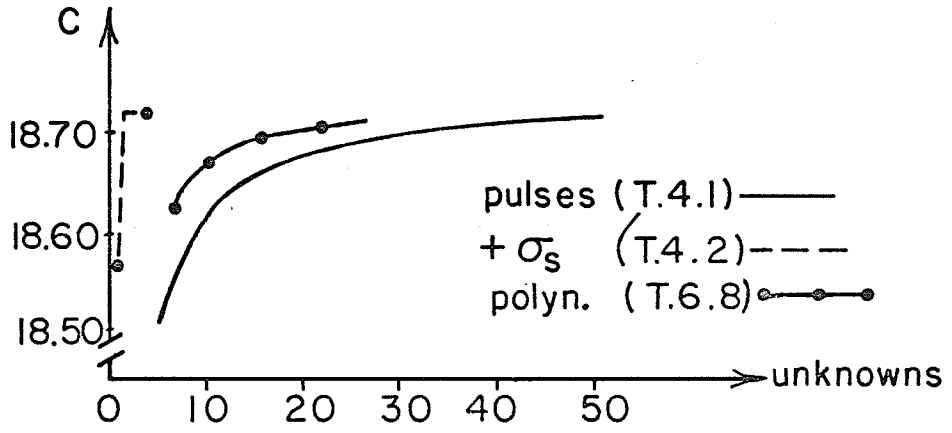


Fig. 6.15 Comparison of Convergence Properties for Different Solutions of the Strip Capacitor Problem.

6.7 A Pin Insulator

Consider the simple pin insulator shown in Fig. 6.16. The problem is axi-symmetric, and the reduced picture-frame shown in Fig. 6.17 is sufficient, with 24 triangular elements. Notice that the insulator has only 3 straight element sides. Since S has been restricted to consist only of line elements, it is necessary to enclose the insulator within S as shown. Had curved elements been allowed, only 9 elements would be required, all within the insulator. This may be compared to the technique of Chapter III where an overlap region is required; the 24 elements are always necessary if the method of Chapter III is applied.

The dielectric constant of the insulator is $\varepsilon_r=2.1$ [4], although it is easily taken as $\varepsilon_r(\underline{r})$ by the PDE formulation*. In fact, the PDE method has been used by Wexler [11] for anisotropic material, and the particular algorithm used here (MANFEP [5]) includes the feature.

The Dirichlet conditions $\phi=0$, $\phi=1$ are enforced as shown, and with N=2, there are 63 node potentials and 13 sources (76 unknowns) to be found. The static equipotentials are shown in Fig. 6.18; continuous and smooth ϕ results throughout.

A contaminated insulator is represanted by permitting conducting material to reside on the surface of the insulator. Thus a portion of the interface is an equipotential. This is easily enforced by using Lagrange constraints; node potentials or the contaminated surface are all forced to assume the same values. For simplicity, here the contaminant touches the conductor, and the Dirichlet condition $\phi=1$ is extended along the insulator surface. The static equipotentials resulting from two tests are shown in Figs. 6.19 and 6.20. The effect of the contaminant is clearly observed in the shift in the location of

* That is, the region may be inhomogeneous or anisotropic.

potential contours towards the ground. This significantly increases the possibility of electrical breakdown.

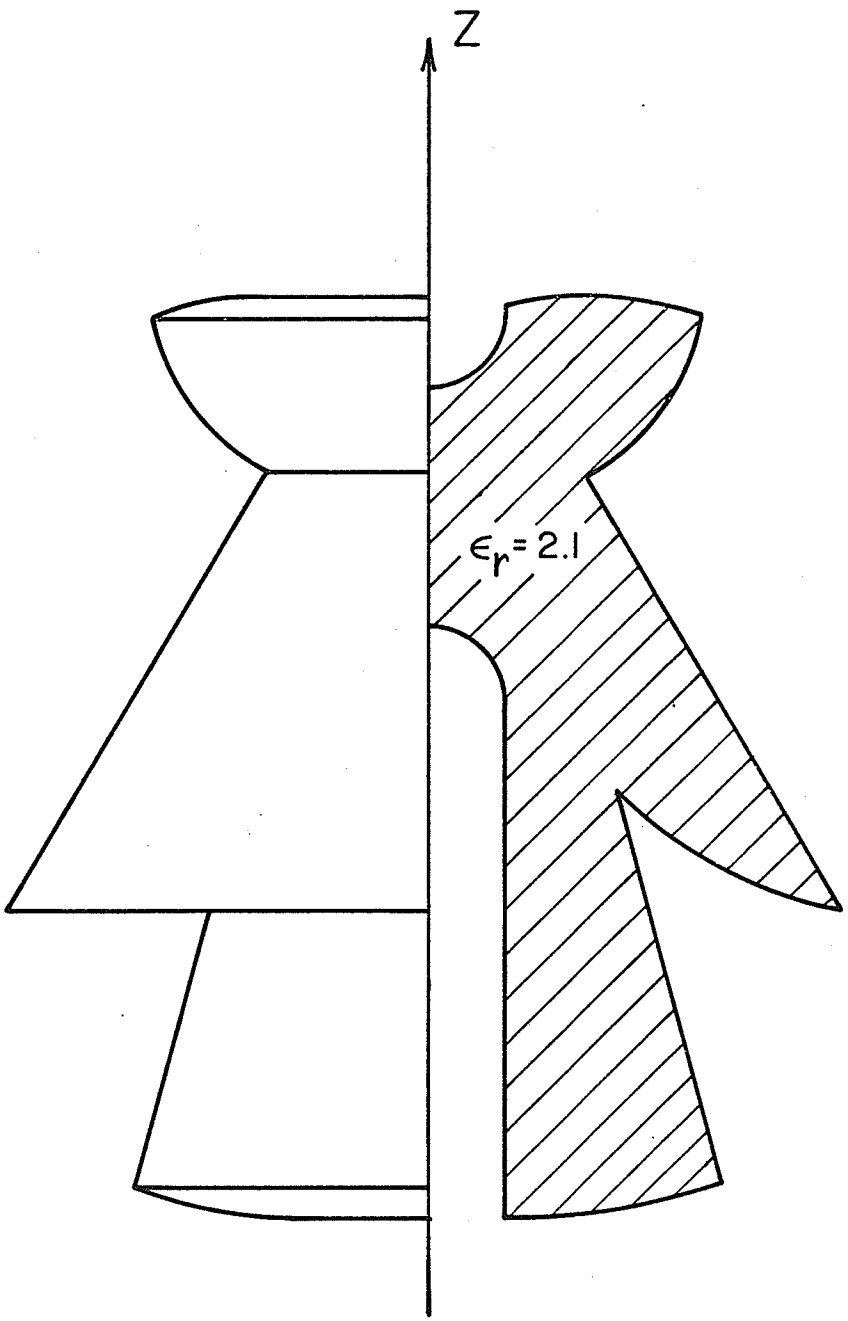


Fig. 6.16 A Typical Pin Insulator

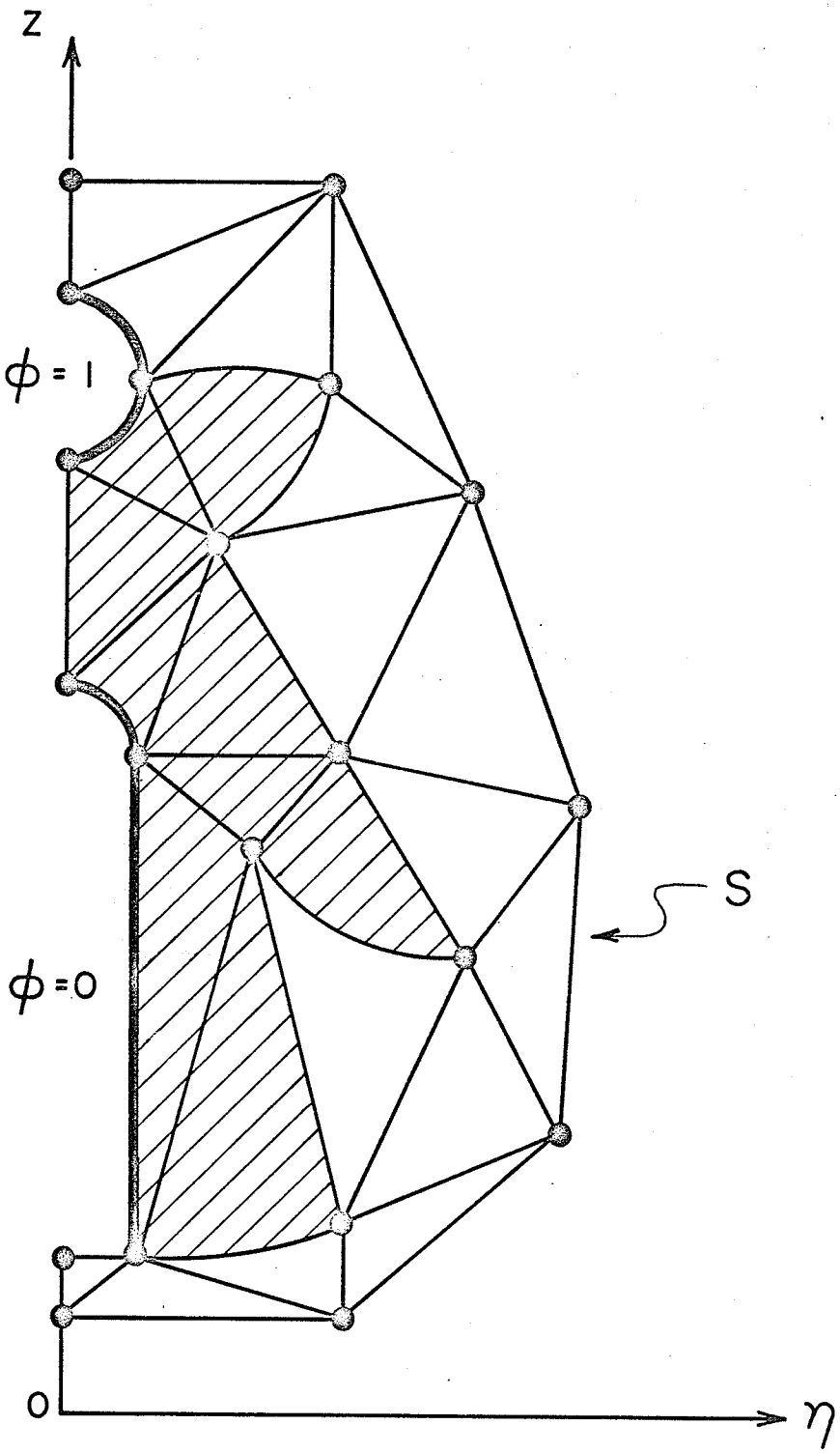


Fig. 6.17 Finite Elements for the Pin Insulator Problem.

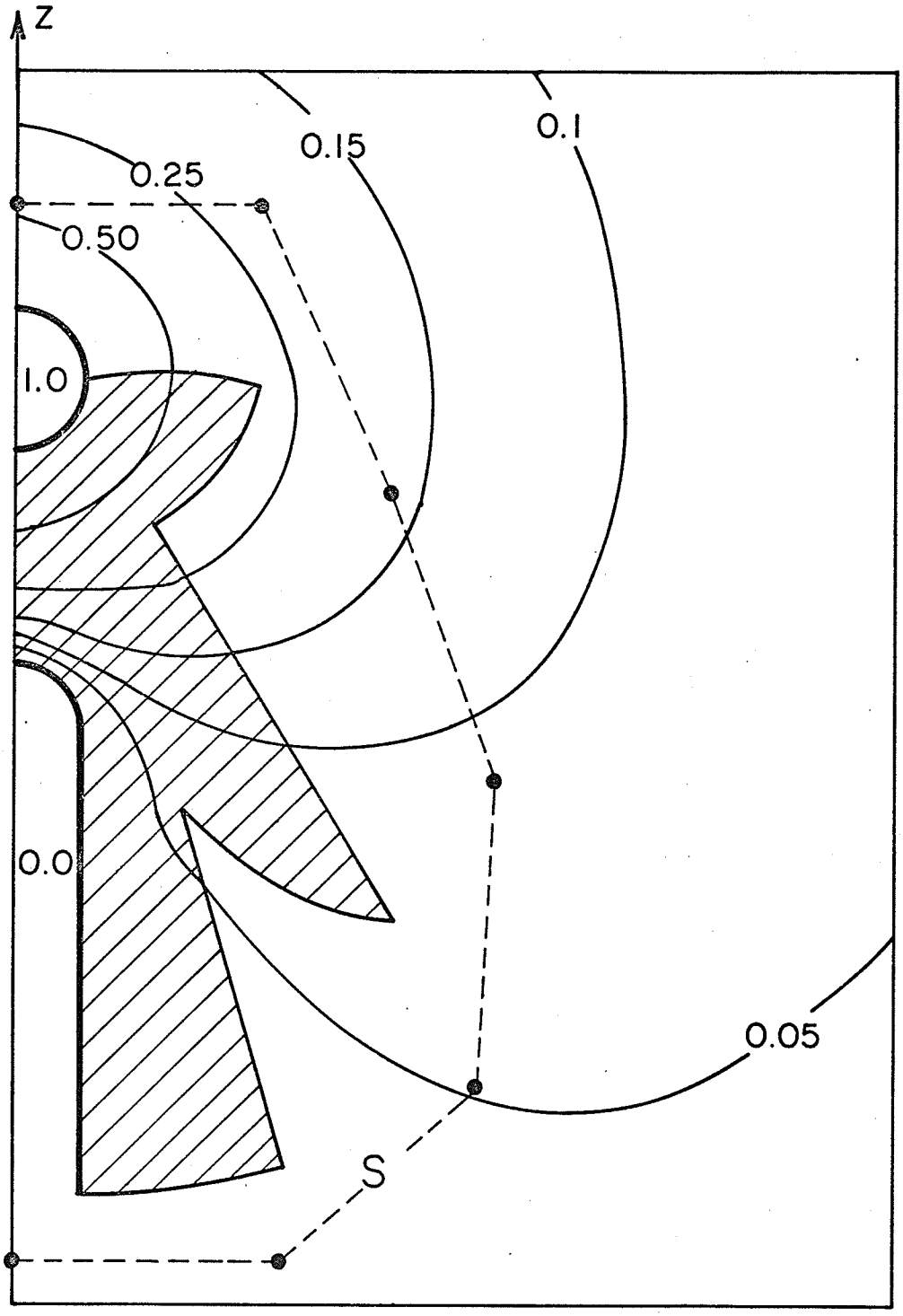


Fig. 6.18 Equipotentials for the Pin Insulator Problem.

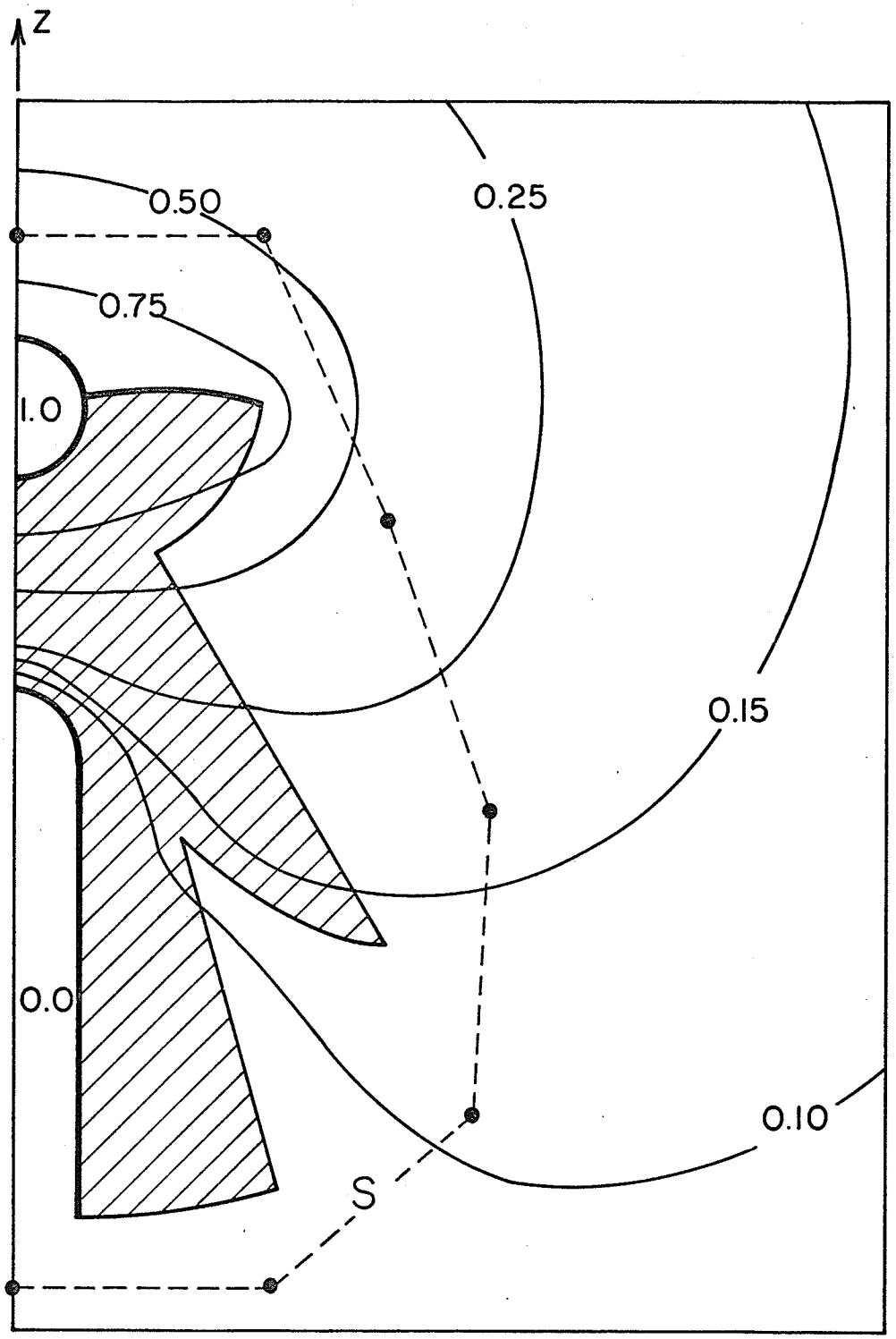


Fig. 6.19 Equipotentials for the Contaminated Pin Insulator, Case I.

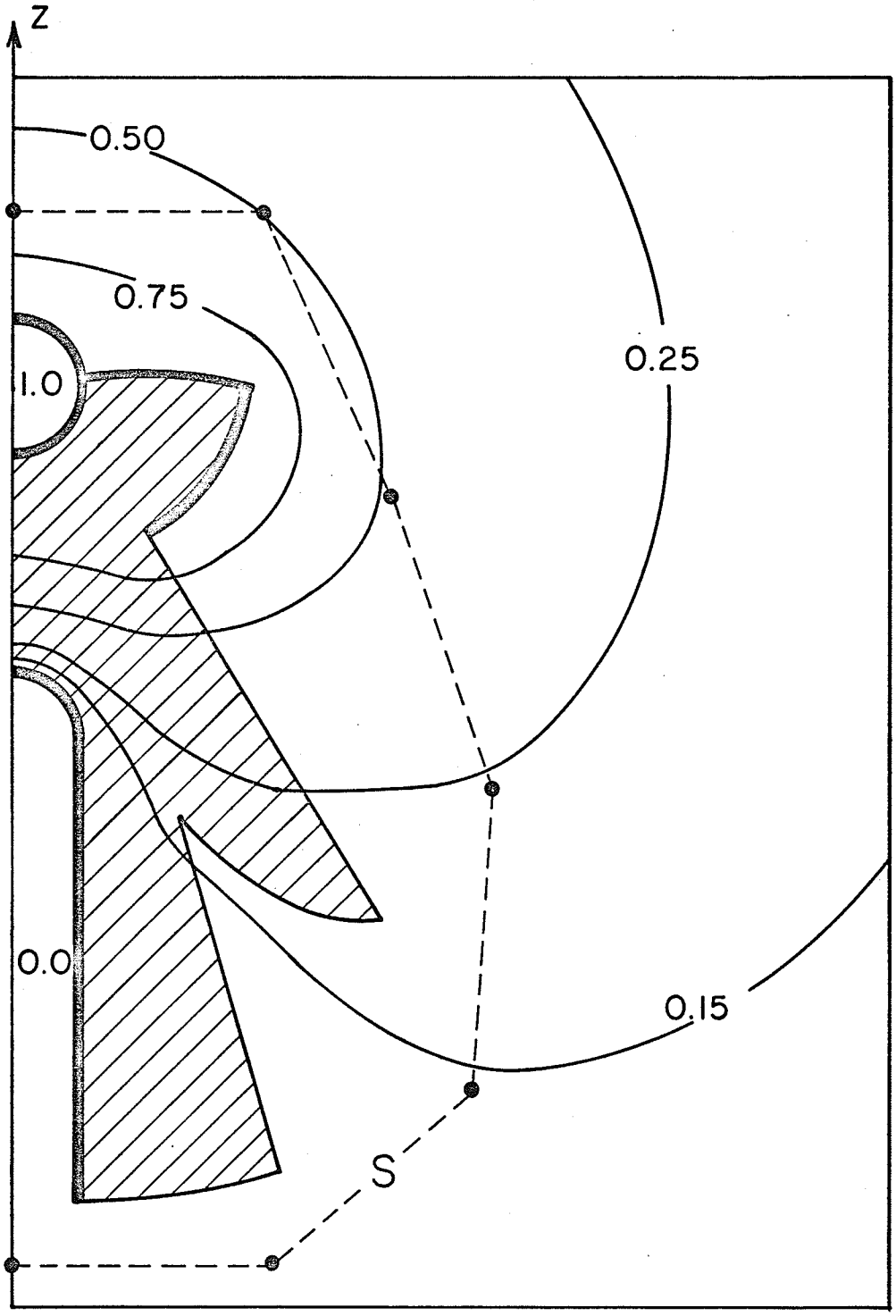


Fig. 6.20 Equipotentials for the Contaminated Pin Insulator, Case II.

6.8 Conclusions

Mutual constraint of PDE and IE variational expressions for unbounded field problems results in energy convergent approximate solutions. The method has been shown capable of high accuracy in static and time-harmonic problems having either uniform cross-section or axial symmetry. The accuracy of the method rests upon the ability of the PDE approximation to represent ϕ in the picture-frames, and upon good representation of $\partial\phi/\partial n$ by σ.

The complex integral operator for time-harmonic problems is not self-adjoint, but the adjoint operator equation may be introduced into the functional to allow the Rayleigh-Ritz matrix equation to be obtained in the usual fashion [12]. Only the simplest free-space Green's functions need be used in the formulation, and even these require sophisticated algorithms to be handled properly. The addition-subtraction technique has been found adequate for treating the singular integrands, and appropriate algorithms have been presented when the finite-elements on S are line segments. An important extension would be to include curved elements.

It is of interest to note that the IE portion of the method is applicable to bounded regions. This is apparent from the original formulation, at equation (3.4). In [13] an example constructed by M. Friedman shows how the method is applicable to such regions. It is possible, therefore, to solve bounded field problems by the mutual constraint method of this chapter. Consider the problem depicted in Fig. 6.21. The large homogeneous region Ω_1 need not have PDE finite elements placed for ϕ; it is sufficient to seek a source $\sigma=\varepsilon_o\partial\phi/\partial n$ on

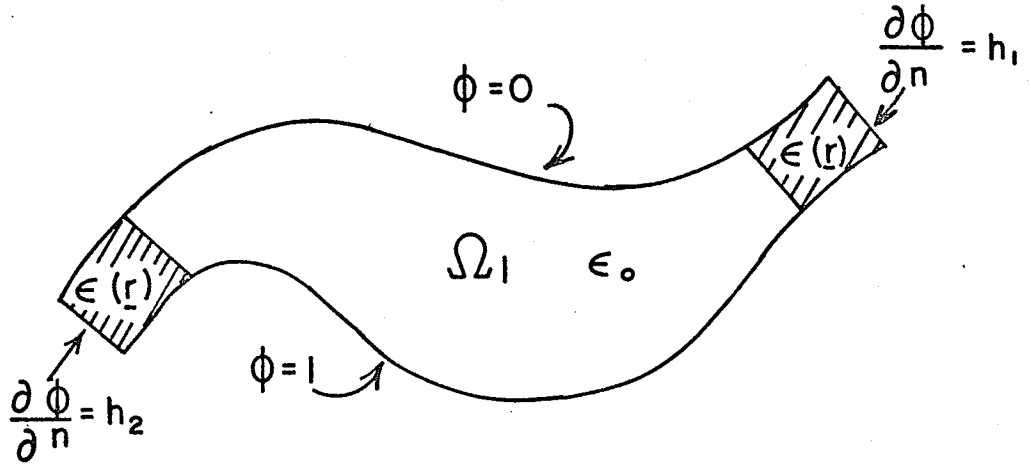


Fig. 6.21 A Bounded Problem.

the boundary of Ω_1. PDE elements may be placed in the two inhomogeneous regions, and mutual constraint may be used to achieve the overall solution.

In addition, one may compare this method to that of Chapter III. The earlier method employed a point-matching technique for the IE portion, whereas this one seeks the solution variationally, with demonstrable energy convergence properties. The earlier method allowed the IE portion to be handled without explicit treatment of Green's function singularities, and this is an algorithmic advantage over the methods used here, where the formulation demands explicit treatment. The earlier method demanded a homogeneous region between the picture-frame boundary and any inhomogeneous material within the picture-frame: this one does not. However, if

dielectric boundaries cannot be modelled by line elements, an homogeneous region results anyway, as in the insulator problem of Section 6.7. The trade-off is essentially between the energy convergence properties of this method and the simpler algorithms of the earlier one. The existence of efficient accurate routines to handle Green's function singularities tend to make this method preferable.

One may also compare the computing requirements of the two methods. Although this method requires more storage for the algorithms than that of Chapter III (because the algorithms are more complicated), the bulk storage required for data (the matrices) is the same. Approximately twenty percent more computation time is required for constructing the variational IE matrix contributions here than for constructing the IE constraint matrix contributions of Chapter III. However, when the computation time for constructing the PDE matrix contributions and for solving the resulting matrix equation (using Gauss elimination and back-substitution) is included, the method described here is only about five percent slower, for the same degree of accuracy.

This picture-frame method may also be compared to the purely IE formulation for interface problems in Chapter V. When the picture-frame region is homogeneous (for example, a homogeneous dielectric) the picture-frame may be replaced by a source distribution on S; and the IE formulation of Chapter V may be used. The advantage is in the reduction in number of unknowns which can also result in a considerable reduction in computer storage and time requirements. On the other hand, when the picture-frame region is not homogeneous or is anisotropic, the method of Chapter V is not applicable and a picture-frame method is recommended.

6.9 References

[1] A.J. Burton and G.F. Miller, "The application of integral equation methods to the numerical solution of some exterior boundary value problems." Proc. of the Royal Society, Series A, Vol. 323, No. 1553, pp. 201-210, 8 June 1971.

[2] E. Jahnke, F. Emde and F. Losch, Tables of Higher Functions. New York: McGraw-Hill, 1960.

[3] P. Silvester, Modern Electromagnetic Fields. Englewood Cliffs, N.J.: Prentice-Hall, 1968.

[4] G.E. Forsythe and W.R. Wasow, Finite Difference Methods for Partial Differential Equations. New York: John Wiley and Sons, 1960.

[5] M. Friedman, D.J. Richards and A. Wexler, University of Manitoba Finite-Element Program (MANFEP). Dept of Elect. Eng., University of Manitoba, Winnipeg, November 1973.

[6] P. Silvester and M.S. Hsieh, "Finite element solution of 2-dimensional exterior field problems." Proc. IEE., Vol. 118, pp. 1743-1747, December 1971.

[7] C.G. Williams and G.K. Cambrell, "Efficient numerical solution of unbounded field problems." Elect. Letts., Vol. 8, pp.247-248, May 1972.

[8] J.R. Reitz and F.J. Milford, Foundations of Electromagnetic Theory. Reading, Mass: Addison-Wesley, 1960.

[9] F. Sandy and J. Sage, "Use of finite difference approximations to partial differential equations for problems having boundaries at infinity." IEEE Trans. MTT-19, pp.484-486, May 1971.

[10] I. Cermak, "Solution of unbounded field problems by boundary relaxation." Ph.D. Thesis., Department of Elect. Eng., McGill University, Montreal, March 1969.

[11] A. Wexler, "Finite-element field analysis of an inhomogeneous, anisotropic, reluctance machine rotor." IEEE Trans. PAS-92, pp.145-149, January 1973.

[12] S.G. Mikhlin, Variational Methods in Mathematical Physics. New York: Macmillan, 1964.

[13] B.H. McDonald, M. Friedman, M. Decreton and A. Wexler, "Integral finite-element approach for solving the Laplace equation." Elect. Lett., Vol. 9, pp. 242-244, May 1973.

CHAPTER VII

SOLUTION OF MATRIX EQUATIONS

Very little has been said to this point about the actual solution of the matrix equations arising in the formulations of the preceeding chapters:

$$[A]\ \underline{x} = \underline{b} \tag{7.1}$$

where $[A]$ is an NxN square matrix, apart from the fact that Gaussian elimination with back-substitution is used. (In the 3 dimensional IE examples it was mentioned that the matrix determinant vanished - or at least reached the computer zero - when over 100 equations were involved.) Instead, the emphasis has been largely on the construction of (7.1) such that $\underline{x}$ is meaningful in terms of the physical problem under consideration.

The term "ill-conditioned" has been used several times in the description of certain matrices. While there are several possible definitions of a matrix "condition number" [1, pp.174-177], it may be conveniently defined as the square root of the ratio of the largest eigenvalue of $[A^TA]$ to the smallest. The larger the condition number, the more ill-conditioned the matrix $[A]$.

Now consider (7.1). If $[A]$ or $\underline{b}$ is perturbed slightly (for example by round-off error in computation) the solution $\underline{x}$ will also be perturbed, but hopefully not by too much. The upper bound on the relative perturbation of $\underline{x}$ is proportional to the condition number. This means that if $[A]$ is ill-conditioned, a small amount of round-off error in computation can cause a very large perturbation in $\underline{x}$ from the true solution, and indeed, can make the computed $\underline{x}$ meaningless. Although a symptom of ill-

conditioning is a vanishing matrix determinant, it is not necessarily true that a small determinant implies a large condition number. For example, if $[A]$ is multiplied by a very small number $\gamma>0$ the determinant is multiplied by γ^N, but the condition number, as a ratio of eigenvalues, is unchanged.

In this chapter the structure of the matrix equations is examined more closely, and alternative strategies are described which might be used for more efficient solution. This will be particularly important when large-scale problems are attempted.

Much of the material presented here is taken from the book edited by Reid [2], which, in many respects, represents the state of the art in 1970.

7.1 The Matrix Equations

The matrix equations presented in this work have arisen from several distinct algorithmic procedures, and it is useful to review briefly the matrix structures at this point.

The PDE finite-element procedures in Chapters III and VI produce matrices which may be sparse. If [A] is the matrix produced, entry A_{ij} will be non zero only if nodes i and j are present in the same finite-element. If there is but one element the matrix will be dense. On the other hand if linear right triangles are used as shown in Fig. 7.1, each node will exist in at most 8 triangles, and therefore there will be at most 9 non zero entries in each row of [A], no matter how many elements are used. This particular case produces the finite-difference matrix [3] if the right triangles are isosceles and boundary conditions are all

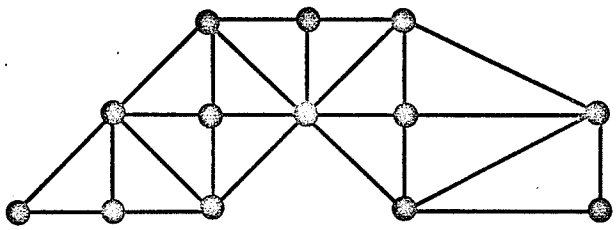


Fig. 7.1 Linear Right Triangular Finite Elements.

Dirichlet. Treating PDE finite-element matrices as being dense can be wasteful.

The matrices produced by IE procedures in Chapters III - VI are dense by comparison. This is because the Green's functions link all elements together via an integration. In certain cases (Chapters IV, V, and equation (6.51)) the matrices are completely dense. In other cases not all the unknowns are required to represent σ in the integral of σG. This occurs in the IE constraint equation of Chapter III; only those node potentials in elements through which S_c passes are involved. It also occurs in Chapter VI (at equation (6.52)); only node potentials (and sources) on S are involved. Therefore in problems where only IE operations are involved the matrices must be treated as being dense, whereas in the PDE-IE picture-frame problems, depending upon the PDE representation and contour placement, the matrices need not be treated as being completely dense.

The picture-frame method in Chapter III involves a Lagrange constraint, and (3.21) may be written in the general form

$$\begin{bmatrix} A & M^T \\ M & 0 \end{bmatrix} \begin{pmatrix} \underline{\phi} \\ \underline{\lambda} \end{pmatrix} = \begin{pmatrix} \underline{b} \\ \underline{0} \end{pmatrix} . \tag{7.2}$$

The PDE matrix $[A]$ is square and sparse, depending upon the finite-element representation, and the non-square IE constraint matrix $[M]$ may be sparse depending upon the PDE finite elements and the placement of S_c. $[M]$ will also be complex-valued for values of the propagation constant $k>0$.

When the boundary node points and the Lagrange multiplier $\underline{\lambda}$ are eliminated, as at (3.26) the matrix is smaller but much more dense, due to the matrix multiplications and additions involved.

Equation (7.2) may be split into

$$\left(\begin{bmatrix} A & 0 \\ 0 & I \end{bmatrix} + \begin{bmatrix} 0 & M^T \\ M^T & -I \end{bmatrix} \right) \begin{pmatrix} \underline{\phi} \\ \underline{\lambda} \end{pmatrix} = \begin{pmatrix} \underline{b} \\ \underline{0} \end{pmatrix} \tag{7.3}$$

where the first matrix is essentially the PDE finite-element matrix and the second matrix may be thought of as the IE constraint matrix.

The IE variational procedures in Chapters IV and V produce completely dense matrices in equations of the form

$$[A] \quad \underline{\sigma} = \underline{b} \tag{7.4}$$

which may be split into

$$\{ [A_o] + [A'] \} \underline{\sigma} = \underline{b} \tag{7.5}$$

where $[A_o]$ is the block diagonal matrix formed by treating each IE finite element in isolation, and where $[A']$ is the mutual constraint matrix. Equation (7.5) is in the same form as (7.3).

A third general form arises in the PDE-IE mutual constraint method of Chapter VI, and as at (6.55) is

$$\begin{bmatrix} A & A' \\ D' & D \end{bmatrix} \begin{pmatrix} \underline{\phi} \\ \underline{\sigma} \end{pmatrix} = \begin{pmatrix} \underline{b} \\ \underline{0} \end{pmatrix} . \tag{7.6}$$

The PDE finite-element matrix $[A]$ may be sparse, but the IE finite-element matrix $[D]$ is completely dense. The mutual constraint entries $[A']$ and $[D']$ are sparse, for the number of nodes on S is often much less than the total number of node points. Equation (7.6) may also be split into two parts

$$\left(\begin{bmatrix} A & 0 \\ 0 & D \end{bmatrix} + \begin{bmatrix} 0 & A' \\ D' & 0 \end{bmatrix} \right) \begin{pmatrix} \underline{\phi} \\ \underline{\sigma} \end{pmatrix} = \begin{pmatrix} \underline{b} \\ \underline{0} \end{pmatrix} \tag{7.7}$$

where the first matrix arises by treating the PDE and IE problems separately, and the second is the mutual constraint matrix. Equation (7.7) is in the same form as (7.5) and (7.3).

One may also separate the static and time-harmonic parts of the equations. (See (1.2) and (1.10)). It is possible, therefore, to represent all the matrix equations as

$$\left[A_o + A' \right] \quad \underline{x} = \underline{b} \tag{7.8}$$

where $[A_o]$ is real-valued, symmetric positive-definite, and sparse, and where $[A']$ may be complex-valued, non-Hermitian, not of a definite form, and relatively dense. It is necessary that $[A] = [A_o + A']$ be non-singular in order to solve (7.8).

7.2 Gaussian Elimination and Back-Substitution

This technique, as applied in all the examples, treats the coefficient matrix in (7.1) as being completely dense. Row interchange, or partial pivoting, is used "to give smaller bounds for the perturbations due to rounding error" (Walsh, in [2, p.43]). Gaussian elimination produces an upper triangular matrix from which the solution is obtained by back-substitution.

Consider a simple 3x3 example (without pivoting), representative of a larger NxN system:

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} . \qquad (7.9)$$

The first step is to eliminate a_{21} and a_{31}. This is done by forming a_{i1}/a_{11}, i=2 .. N and subtracting the first equation times this number from the i^{th} equation:

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22}-\frac{a_{21}}{a_{11}}a_{12} & a_{23}-\frac{a_{21}}{a_{11}}a_{13} \\ 0 & a_{32}-\frac{a_{31}}{a_{11}}a_{12} & a_{33}-\frac{a_{31}}{a_{11}}a_{13} \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2-\frac{a_{21}}{a_{11}}b_1 \\ b_3-\frac{a_{31}}{a_{11}}b_1 \end{pmatrix} . \qquad (7.10)$$

The next step is to eliminate the entry $a_{32}-a_{31}a_{12}/a_{11}$. This is done by forming $(a_{i2}-\frac{a_{i1}}{a_{11}}a_{12})/(a_{22}-\frac{a_{21}}{a_{11}}a_{12})$, i=3 .. N, and subtracting the second equation times this number from the $i^{\underline{th}}$ equation :

$$\begin{bmatrix} a_{11} & & a_{13} \\ 0 & a_{22}-\frac{a_{21}}{a_{11}}a_{12} & a_{23}-\frac{a_{21}}{a_{11}}a_{13} \\ 0 & 0 & (a_{33}-\frac{a_{31}}{a_{11}}a_{13})-(a_{32}-\frac{a_{31}}{a_{11}}a_{12})(a_{23}-\frac{a_{21}}{a_{11}}a_{13})/(a_{22}-\frac{a_{21}}{a_{11}}a_{12}) \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

$$= \begin{pmatrix} b_1 \\ b_2-\frac{a_{21}}{a_{11}}b_1 \\ (b_3-\frac{a_{31}}{a_{11}}b_1) - (a_{32}-\frac{a_{31}}{a_{11}}a_{12})(b_2-\frac{a_{21}}{a_{11}}b_1)/(a_{22}-\frac{a_{21}}{a_{11}}a_{12} \end{pmatrix} . \qquad (7.11)$$

The procedure continues until the [A] matrix is in upper diagonal form:

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22}' & a_{23}' \\ 0 & 0 & a_{33}' \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2' \\ b_3' \end{pmatrix} . \qquad (7.12)$$

Consider the products of the diagonals

$$\prod_{i=1}^{N} a_{ii} = \det [A] . \qquad (7.13)$$

If N=1, immediately a_{11} is the determinant. With N=2

$$a_{11}a_{22}' = a_{11}\left(a_{22} - \frac{a_{21}\,a_{12}}{a_{11}}\right) = a_{11}\,a_{22} - a_{21}\,a_{12} \qquad (7.14)$$

and with N=3

$$a_{11}a_{22}'a_{33}' = a_{11}\left(a_{22}-\frac{a_{21}a_{12}}{a_{11}}\right)\left(a_{33}-\frac{a_{31}a_{13}}{a_{11}}\right)$$

$$- a_{11}\left(a_{32}-\frac{a_{31}a_{12}}{a_{11}}\right)\left(a_{23}-\frac{a_{21}a_{13}}{a_{11}}\right)$$

$$= a_{11}\,a_{22}\,a_{33} + a_{12}\,a_{23}\,a_{31} + a_{13}\,a_{21}\,a_{32}$$

$$- a_{31}\,a_{22}\,a_{13} - a_{32}\,a_{23}\,a_{11} - a_{33}\,a_{21}\,a_{12} \qquad (7.15)$$

which, again, is the determinant.

If the matrix is sparse and may be put into a banded form:

$$[A] = \begin{bmatrix} \ddots & & 0 \\ & \ddots & \\ 0 & & \ddots \end{bmatrix} \qquad (7.16)$$

the method may be applied just over the band (Walsh, in [2, p.43]).

7.3 Symmetric Choleski Factorization

In the special case where the coefficient matrix is symmetric the Choleski method [4, pp.76-78; pp.95-97] is faster than Gaussian elimination. This form occurs in the static PDE-IE Lagrange constraint method of Chapter III, and in the integral variational methods of Chapter IV.

For a symmetric matrix one may write

$$[A] = [L]\,[L]^T \tag{7.17}$$

where $[L]$ is a lower triangular matrix. Consider the simple 2x2 case:

$$[A] = \begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix} \tag{7.18a}$$

and

$$[L] = \begin{bmatrix} l_{11} & 0 \\ l_{12} & l_{22} \end{bmatrix} . \tag{7.18b}$$

Then, from (7.17)

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix} = \begin{bmatrix} l_{11}^2 & l_{11}l_{21} \\ l_{11}\,l_{21} & l_{11}^2 + l_{22}^2 \end{bmatrix} \tag{7.19}$$

and the elements of $[L]$ may be obtained:

$$l_{11} = \sqrt{a_{11}} \tag{7.20a}$$

$$l_{12} = a_{12}/l_{11} \tag{7.20b}$$

$$l_{22} = \sqrt{a_{22} - l_{11}^2} \quad . \tag{7.20c}$$

This may be extended readily to an NxN matrix.

The appearance of the square root in l_{ii} may indicate the necessity of complex arithmetic; however if $[A]$ is positive-definite this is not necessary [4, p.78].

The equation (7.1) becomes

$$[LL^T]\ \underline{x} = \underline{b} \tag{7.21}$$

or

$$[L]\ \underline{y} = \underline{b} \tag{7.22}$$

where

$$[L]^T\ \underline{x} = \underline{y}\ . \tag{7.23}$$

For the simple example (7.23) takes the form

$$\begin{bmatrix} l_{11} & 0 \\ l_{12} & l_{22} \end{bmatrix} \begin{Bmatrix} y_1 \\ y_2 \end{Bmatrix} = \begin{Bmatrix} b_1 \\ b_2 \end{Bmatrix} \tag{7.24}$$

and substitution is required to find the y_i. In the general case

$$y_i = (b_i - \sum_{k=1}^{i-1} l_{ik}\, y_k)/l_{ii}\ ;\ i=1,\ 2\ ..\ N \tag{7.25}$$

and the x_i are obtained in a similar fashion

$$x_i = (y_i - \sum_{k=i+1}^{N} l_{ki}\, x_k)/l_{ii}\ ;\ i=1,\ 2\ ..\ N\ . \tag{7.26}$$

No additional computer storage is required, as the matrix $[L]$ may be written directly over $[A]$. Also, the vector $\underline{y}$ may be written directly over $\underline{b}$, and $\underline{x}$ directly over $\underline{y}$. This method is to be preferred to Gaussian elimination for solution of the dense, symmetric matrix equations produced in solution of the IE problems in Chapter IV.

7.4 A Bifactorization Method

An algorithm for direct solution of (7.1) based on a bifactorization method is presented by Zollenkopf in [2, pp.75-96]; the method is computationally convenient when $[A]$, the coefficient matrix, has symmetrical zero entries. The method is based on the equation

$$[L^{(N)} \ldots L^{(2)} L^{(1)} A R^{(1)} R^{(2)} \ldots R^{(N)}] = I \qquad (7.27)$$

where the $[L]$ and $[R]$ are left and right hand factor matrices, respectively, and N is the number of equations in the matrix system (7.1). Then

$$[A^{-1}] = [R^{(1)} R^{(2)} \ldots R^{(N)} L^{(N)} \ldots L^{(2)} L^{(1)}] \qquad (7.28)$$

and the solution of (7.1) is simply

$$\underline{x} = [A^{-1}] \underline{b} \ . \qquad (7.29)$$

To produce the factor matrices the following sequence of intermediate matrices is produced:

$$\begin{aligned} [A^{(0)}] &= [A] \\ [A^{(j)}] &= [L^{(j)} A^{(j-1)} R^{(j)}] \\ [A^{(N)}] &= [I] \ . \end{aligned} \qquad (7.30)$$

The reduced matrix $[A^{(j)}]$ is produced as:

$$A_{jj}^{(j)} = 1 \ ; \ A_{ij}^{(j)} = A_{jk}^{(j)} = 0 \ ;$$

$$A_{ik}^{(j)} = A_{ik}^{(j-1)} - \frac{A_{ij}^{(j-1)} A_{jk}^{(j-1)}}{A_{jj}^{(j-1)}} \qquad (7.31)$$

where j is the pivotal index j=1 ... N, and i,k = j+1 ... N. Clearly $[A^{(N)}] = [I]$. The left hand factor matrices are very sparse and differ from I only in column j

$$L_{jj}^{(j)} = 1/A_{jj}^{(j-1)}$$

$$L_{ij}^{(j)} = -A_{ij}^{(j-1)}/A_{jj}^{(j-1)} \quad ; \; i = J+1 \ldots N \; . \tag{7.32}$$

Similarly, the right hand factor matrices differ from I only in row j

$$R_{jj}^{(j)} = 1$$

$$R_{jk}^{(j)} = -A_{jk}^{(j-1)}/A_{jj}^{(j-1)} \quad k = j+1 \,..\, N \; . \tag{7.33}$$

It is clear that very straightforward algorithms may be used to compute these factor matrices.

Consider a simple 2x2 example :

$$[A^{(0)}] = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \; . \tag{7.34}$$

Then

$$[A^{(1)}] = \begin{bmatrix} 1 & 0 \\ 0 & \dfrac{\Delta}{a_{11}} \end{bmatrix} \tag{7.35}$$

$$[A^{(2)}] = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{7.36}$$

where $\Delta = a_{11}a_{22} - a_{21}a_{12}$ is the matrix determinant. The left hand factor matrices are then

$$[L^{(1)}] = \begin{bmatrix} \frac{1}{a_{11}} & 0 \\ \frac{-a_{21}}{a_{11}} & 1 \end{bmatrix} \tag{7.37}$$

and

$$[L^{(2)}] = \begin{bmatrix} 1 & 0 \\ 0 & \frac{a_{11}}{\Delta} \end{bmatrix} . \tag{7.38}$$

The right hand factor matrices are

$$[R^{(1)}] = \begin{bmatrix} 1 & \frac{-a_{12}}{a_{11}} \\ 0 & 1 \end{bmatrix} \tag{7.39}$$

and

$$[R^{(2)}] = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} . \tag{7.40}$$

Then from (7.28)

$$[A^{-1}] = \begin{bmatrix} 1 & \frac{-a_{12}}{a_{11}} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \frac{a_{11}}{\Delta} \end{bmatrix} \begin{bmatrix} \frac{1}{a_{11}} & 0 \\ \frac{-a_{21}}{a_{11}} & 1 \end{bmatrix}$$

$$= \frac{1}{\Delta} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix} \tag{7.41}$$

and this is correct.

Zollenkopf's algorithm exploits the benefits of sparsity by using a packed matrix storage scheme in which only the non-zero terms are retained (thus requiring indexing algorithms) and "requires only a small fraction of the memory space used for storage of the full matrix and its inverse" [2, p.85].

7.5 Iterative Methods

The procedures outlined above are direct methods, in that the solution to (7.1) is produced in one step. There are also methods which take an assumed solution for $\underline{x}$ and create a better one; repeated application of the algorithm will hopefully generate the true solution to (7.1) after a finite number of steps, or iterations.

The general form of a stationary iterative process is obtained by splitting $[A]$ into the sum of two matrices $[B] + [C]$, where $[B]$ is non singular, and writing (Walsh, in [2, p.52])

$$[B] \quad \underline{x}^{(\ell+1)} = - [C] \quad \underline{x}^{(\ell)} + \underline{b} \quad . \tag{7.42}$$

Convergence ($\ell \to \infty$) is obtained if and only if the eigenvalues of $[-B^{-1} C]$ are less than unity in modulus [2, p.52]. Consider the first step, with a starting value $\underline{x}^{(0)}$:

$$\underline{x}^{(1)} = [-B^{-1} C] \quad \underline{x}^{(0)} \quad + \quad [B^{-1}] \quad \underline{b} \quad . \tag{7.43}$$

The next step is

$$\begin{aligned} \underline{x}^{(2)} &= [-B^{-1}C] \quad \underline{x}^{(1)} \quad + \quad [B^{-1}] \quad \underline{b} \\ &= [-B^{-1}C]^2 \quad \underline{x}^{(0)} \quad + \quad [I-B^{-1}C] \quad [B^{-1}] \quad \underline{b} \end{aligned} \tag{7.44}$$

and step ℓ is given by

$$\underline{x}^{(\ell)} = [-B^{-1}C]^{\ell} \quad \underline{x}^{(0)} + \sum_{i=0}^{\ell} \quad [-B^{-1}C]^{i} \quad [B^{-1}] \quad \underline{b} \quad . \tag{7.45}$$

Now, the inverse of $[I+ B^{-1}C]$ is given by

$$[I+B^{-1}C]^{-1} = \sum_{i=0}^{\infty} \quad [-B^{-1}C]^{i} \quad . \tag{7.46}$$

Therefore, in the limit $\ell \to \infty$

$$\underline{x}^{(\infty)} = \left[-B^{-1}C\right]^{\infty} \underline{x}^{(0)} + \left[I+B^{-1}C\right]^{-1} \left[B^{-1}\right] \underline{b}$$

$$= \left[-B^{-1}C\right]^{\infty} \underline{x}^{(0)} + \left[B(I+B^{-1}C)\right]^{-1} \underline{b}$$

$$= \left[-B^{-1}C\right]^{\infty} \underline{x}^{(0)} + \left[A^{-1}\right] \underline{b}$$

$$= \left[-B^{-1}C\right]^{\infty} \underline{x}^{(0)} + \underline{x} \; . \qquad (7.47)$$

If $\underline{x}^{(\infty)}$ is to be the correct solution $\underline{x}$, $\left[-B^{-1}C\right]^{\infty} \equiv [0]$, and this is guaranteed by the earlier condition on the eigenvalues. Although exact convergence is obtained only in the limit, solutions of specified accuracy may often be obtained with only a few iterations.

If $[A]$ is of general sparse form and can be written as

$$[A] = [\, L + D + U \,] \qquad (7.48)$$

where $[L]$ and $[U]$ are strictly lower and upper triangular, respectively, and where $[D]$ is diagonal and non-singular, the relaxation process

$$[D + \omega L] \; \underline{x}^{(\ell+1)} = \{(1-\omega)[D] + \omega \left[L^T\right]\} \; \underline{x}^{(\ell)} + \omega \, \underline{b} \qquad (7.49)$$

converges for $0 < \omega < 2$ if $[A]$ is also positive-definite and symmetric. This result has led to the type of iterative schemes used by Cermak [5] and Sandy and Sage [6] for the solution of the finite-difference equations in the picture-frame problems.

The boundary relaxation technique applied here in Section 6.4 is a variant obtained by splitting the matrix into $[A_o] + [A']$, where $[A_o]$ is formed by treating the PDE-IE equations separately and $[A']$ is the mutual constraint matrix. One may then write

$$[A_o] \; \underline{x}^{(\ell+1)} = - [A'] \; \underline{x}^{(\ell)} + \underline{b} \qquad (7.50)$$

and solve by relaxation:

$$[A_o]\ \underline{x}^{(\ell+1)} = (1-\omega)[A_o]\ \underline{x}^{(\ell)} + \omega\ \{-[A']\ \underline{x}^{(\ell)} + \underline{b}\} \qquad (7.51a)$$

or

$$\underline{x}^{(\ell+1)} = (1-\omega)\ \underline{x}^{(\ell)} + \omega\{\ -[A_o^{-1}A']\ \underline{x}^{(\ell)} + [A_o^{-1}]\ \underline{b}\ \}\ . \qquad (7.51b)$$

When a new value $x_i^{(\ell+1)}$ is obtained it may be used immediately in computation in place of $x_i^{(\ell)}$. The scheme in Section 6.4 involves updating $\underline{x}$ on a block basis: it is block relaxation.

In order to implement the algorithm it is necessary to invert $[A_o]$; or at least to solve the system

$$[A_o]\ \underline{x}^{(\ell+1)} = \underline{b}'\ . \qquad (7.52)$$

When $[A_o]$ is block diagonal it is possible to solve (7.44) one block at a time, and this is the technique actually used in Section 6.4.

From the discussion in Section 7.1 it is apparent that all the matrix equations used in this work may be split so that $[A_o]$ is symmetric, sparse and positive-definite. Techniques such as that described by Zollenkopf may be used to determine $[A_o^{-1}]$ - and this is needed only once. Such an approach may be desirable for large systems. It is necessary to ensure that the convergent solution is correct, and investigation is required into the properties of $[-A_o^{-1}A']$ for each case; the eigenvalues must be, in modulus, less than unity.

A basic reference for iterative methods is the text book by Varga [7] and Mohberg and Reynolds [8] describe an application to the Helmholtz equation in lossy material.

7.6 Diakoptical Analysis

Diakoptics is a method whereby large, complicated systems can be torn into a set of smaller parts. Each part is solved, conditional upon the solution of the others and the parts are then reconnected. Much of the general theory is presented in [9] by Kron, the principal author of this approach.

Suppose that the solution of a PDE problem is described by

$$[A] \ \underline{\phi} = \underline{b} \tag{7.53}$$

where there are M node potentials in several finite-elements. Let $\underline{\phi}_N$ be the vector of node potentials in certain elements - the part of the system under consideration, and let $\underline{\phi}_K$ be the vector of node potentials in elements to be removed or "torn out" of the region. Then (7.53) may be written

$$\begin{bmatrix} A_{NN} & A_{NK} \\ A_{KN} & A_{KK} \end{bmatrix} \begin{pmatrix} \underline{\phi}_N \\ \underline{\phi}_K \end{pmatrix} = \begin{pmatrix} \underline{b}_N \\ \underline{b}_K \end{pmatrix} . \tag{7.54}$$

The idea is to remove the K nodes and solve for $\underline{\phi}_N$ alone. Clearly

$$[A_{NN}] \ \underline{\phi}_N = \underline{b}_N - [A_{NK}] \ \underline{\phi}_K \tag{7.55}$$

and this would be the method of mutual constraint: but the vector $\underline{\phi}_K$ would be presumed to be known.

In diakoptics, however, (7.55) is written as

$$[A_{NN}] \ \underline{\phi}_N = \underline{b}_N - \tilde{\underline{b}}_N \tag{7.56}$$

where $\tilde{b}_N$ is the "correction vector" resulting from the removal of the elements containing the K nodes. Now it is clear from the discussion of PDE finite-element methods in Chapter II that one need only enforce continuity between the elements of interest and those removed; only nodes in elements at the "tears" need be considered as contributing to $\tilde{b}_N$. Let the vector of these nodes be $\underline{\phi}_T$ and consider the equation

$$[A_{TT}] \quad \underline{\phi}_T \quad = \quad \underline{b}_T \ . \tag{7.57}$$

This is an equation relating the nodes on the "tears" in such a way that ϕ satisfies the PDE in the region removed. (In principle, it is not unlike the IE constraint equation of Chapter III.)

Now define a connection matrix $[C_{TN}]$ which selects the entries of $\underline{\phi}_T$ from $\underline{\phi}_N$:

$$\underline{\phi}_T \quad = \quad [C_{TN}]^T \quad \underline{\phi}_N \ . \tag{7.58}$$

It is clear that

$$\tilde{\underline{b}}_N = [C_{NT}] \ \underline{b}_T \ . \tag{7.59}$$

Substituting (7.58) into (7.57) yields

$$[A_{TT}] \ [C_{TN}]^T \ \underline{\phi}_N \ = \ \underline{b}_T \ . \tag{7.60}$$

Using (7.59), (7.56) becomes

$$[A_{NN}] \ \underline{\phi}_N \ = \ \underline{b}_N \ - \ [C_{NT}] \ \underline{b}_T$$

or

$$\underline{\phi}_N \ = \ [A_{NN}]^{-1} \ \{ \ b_N \ - \ [C_{NT}] \ \ \underline{b}_T \ \} \ . \tag{7.61}$$

Substituting (7.61) into (7.60) and rearranging yields

$$[C_{TN}]^T \ [A_{NN}]^{-1} \ \ \underline{b}_N \ = \ \{ [A_{TT}]^{-1} \ + \ [C_{TN}]^T \ [A_{NN}]^{-1} \ [C_{NT}] \} \ \underline{b}_T \ . \tag{7.62}$$

From (7.62) one may obtain $\underline{b}_T$ as a function of $\underline{b}_N$:

$$\underline{b}_T = [\tilde{A}_{TT}] \; [C_{TN}]^T \; [A_{NN}]^{-1} \; \underline{b}_N \tag{7.63}$$

where

$$[\tilde{A}_{TT}]^{-1} = [A_{TT}]^{-1} + [C_{TN}]^T \; [A_{NN}]^{-1} \; [C_{NT}] \; . \tag{7.64}$$

Substituting (7.63) into (7.61) gives the final equation of solution:

$$\underline{\phi}_N = [A_{NN}]^{-1} \; \underline{b}_N - [A_{NN}]^{-1} [C_{NT}] [\tilde{A}_{TT}] [C_{TN}]^T \; [A_{NN}]^{-1} \; \underline{b}_N \; . \tag{7.65}$$

Note that this equation is independent of $\underline{\phi}_K$; the connection with the removed nodes is through $[A_{TT}]$.

One may begin at the right-hand side of (7.65) to evaluate $\underline{\phi}_N$. The product $[A_{NN}]^{-1} \; \underline{b}_N$ can be solved by a Choleski factorization scheme, or by Zollenkopf's bifactorization method. The matrix $[\tilde{A}_{TT}]$ is found from (7.64) and cannot be computed by a sparse matrix technique. Therefore T should be a fairly small number if the scheme is to be effective.

The power of this method rests in the fact that the entire system need not be recomputed if a part is changed. (The "constraint" equation (7.57) is employed via (7.64) and (7.65) to link the changed element(s) to the rest of the system.) It is particularly effective for PDE formulations, but not so useful for IE applications. In the latter it is not possible to "tear" a charge distribution and proceed considering only the charge at the tear; all of the charge contributes to potential at every point through an integration. Therefore T is a large number for IE applications.

7.7 Conclusions

Although Gaussian elimination and back-substitution has been adequate to solve the matrix equations for most of the examples presented, better, more economical methods are to be preferred for larger systems. The algorithm developed by Zollenkopf is useful for sparse matrices with symmetrically placed zeros, and such matrices arise in the picture-frame methods of Chapter III and VI. The requirement of matrix positive-definiteness "for reasons of round-off error" [2, p.75] is clearly violated for certain time-harmonic problems, and investigation into the type and magnitude of such errors for the problems considered here might be useful if Zollenkopf's algorithm is selected.

The IE matrix equations in Chapter IV and V possess dense coefficient matrices. Symmetric Choleski factorization is recommended for the symmetric matrix equations in Chapter IV, as it is faster than Gaussian elimination. The non-symmetric, dense matrix produced by the IE interface problem method in Chapter V remains somewhat difficult to handle. A.J. Burton, in private communication [10] , has indicated that he is about to publish an efficient, direct method for solving large, dense matrix equations and this might prove useful.

Construction of an iterative process, where the matrix equation to be solved at each step has a sparse, symmetric, positive-definite coefficient matrix, is possible for all the problems considered, but there is a question of convergence to be explored. The experiences reported in Section 6.4 indicate that for at least some Helmholtz problems one can be certain of convergence properties similar to those for the static case. Zollenkopf's algorithm can be used to invert the coefficient matrix (which is the same at each iteration) and this will facilitate computation.

Although diakoptical analysis shows much promise for PDE applications it may not be as effective for IE problems. Nevertheless, investigation of its possible uses for IE problems could be fruitful.

7.8 References

[1] J.N. Franklin, Matrix Theory, Englewood Cliffs, New Jersey: Prentice-Hall, 1968.

[2] J.K. Reid (editor), Large Sparse Sets of Linear Equations. London: Academic Press, 1971.

[3] J.G.A. Cross and A.C. Walker, "The finite-difference and localized Ritz methods." Int. Jnl. for Num. Meth. Eng., Vol. 3, pp.155-160, 1971.

[4] C.E. Froberg, Introduction to Numerical Analysis. Reading, Mass: Addison-Wesley, 1965.

[5] I. Cermak, "Solution of unbounded field problems by boundary relaxation." Ph.D Thesis. Dept. of Elect. Eng., McGill University, Montreal, March 1969.

[6] F. Sandy and J. Sage, "Use of finite difference approximations to partial differential equations for problems having boundaries at infinity." IEEE Trans., MTT-19, pp.484-486, May 1971.

[7] R.S. Varga, Matrix Iterative Analysis, Englewood Cliffs, N.J.: Prentice-Hall, 1965.

[8] J. Molberg and D. Reynolds, "Iterative solutions of the scalar Helmholtz equation in lossy regions." IEEE Trans. MTT-17, pp.460-464, August 1969.

[9] G. Kron, Diakoptics, London: MacDonald, 1963.

[10] A.J. Burton, private communication, May 1974.

CHAPTER VIII

GENERAL CONCLUSIONS

Field problems may be solved directly by seeking the solution of a PDE, or indirectly be seeking a source distribution as the solution of an IE. In either case it is convenient to subdivide the region which the solution is to span into finite elements and treat the elements one at a time. The solution over each element may be obtained variationally by application of the Rayleigh-Ritz method. It is necessary to connect the element solutions together if one is to obtain the solution to the entire problem, and this procedure may be thought of as mutually constraining the finite-elements.

For each finite-element, whether a PDE element over which the field is sought or an IE element over which a source distribution is sought, it has been possible to show that for the static and propagation problems dealt with here the operator has a positive-definite component. Application of the Rayleigh-Ritz procedure therefore yields a minimizing sequence. Convergence in energy to the unique problem solution is assured by the theorem given in Mikhlin [1]. This is very important as it guarantees that as one refines the approximation one improves the approximation: rms error is reduced. Inhomogeneous anisotropic material has been catered for in consideration of the PDE element (Chapter II), and the complex integral operator occurring in IE propagation problems has been considered as well (Chapter VI). Lossy material, where ε is complex-valued, should be easily included in the PDE formulation.

PDE finite-element methods are well documented [2-3] and the isoparametric finite-element method is recommended for obtaining the PDE

matrix equation for each element. PDE elements are mutually constrained easily by demanding field continuity at inter-element boundaries. The general purpose polynomial trial functions used in the application of the method here are adequate for modelling most fields. However it is possible to improve the representation by including special functions to cater for known non-polynomic field behaviour. This has been demonstrated by Décreton [4] who was interested in modelling field behaviour near edges and corners.

IE finite-elements are not handled so simply, due primarily to the presence of the singular Green's function. In order to make computation convenient only the simple free-space Green's functions have been employed, and the elements have been linear: lines in the plane and rectangles in space. The addition-subtraction technique (Appendix C) has been usefully employed to automate the necessary integrations over each IE finite-element. Pulse functions (Chapters IV and V) and polynomials (Chapter VI) have both been used to approximate the source distribution on each element, and the introduction of special trial functions to handle edge and corner singularities (Chapter IV) has been shown to improve convergence. The improvement was dramatic in the case of an edge singularity, but it was not as significant in the three-dimensional example presented. This is probably because the assumed edge and corner singularity function was not very good; there is work to be done in analysing field behaviour near corners.

The use of free-space Green's functions is algorithmically convenient, as it has been observed that finding the special problem-dependent Green's function can be as difficult as solving the problem itself [5]. This advantage is offset by a very important restriction: the space must be homogeneous, isotropic and lossless. The permittivity must be a constant

throughout. This was no disadvantage in the free-space examples of Chapter IV, and the technique presented in Chapter V permits regions where ε is piecewise constant to be handled by use of the simple Green's functions. For example, a dielectric obstacle is replaced by a derivative constraint (an equivalent interface source distribution) on its boundary. This permits a source distribution to be sought on the interface by IE element techniques.

The finite-elements used in the specific examples presented here are of three types. The PDE isoparametric element is used to represent the field subject to a Dirichlet or Neumann or a mixed condition on the element boundary. The IE "Dirichlet" element is used for representation of the source distribution where a Dirichlet condition is specified, and the IE "interface" element is used for representation of the equivalent interface source distribution which results when a homogeneous dielectric obstacle is removed. The algorithms for constructing the element matrix equation for each element have been presented. In addition, algorithms for mutual constraint of elements of the same kind have been presented, and in this case, when the elements are all of the same kind, the resulting matrix equation for the entire problem is the same as the Rayleigh-Ritz matrix equation for the entire problem.

It has also been convenient to mutually constrain elements of different kinds for particular problems. In the case of an air-dielectric interface in the presence of a charged conductor it has been useful to mutually constrain the IE Dirichlet and interface elements (Chapter V). Here the final matrix equation has not, in general, a symmetric matrix. The matrix equation produced by mutual constraint coincides not with the Rayleigh-Ritz but rather with the Galerkin matrix equation for the problem. Nevertheless, a positive-definite operator component exists, and energy

convergence is guaranteed by Mikhlin's theorem. It is important to note that the shapes of the conductors and the dielectric obstacles are not restricted by the method. The interface need not be in a plane, for example. The method is quite general purpose, and although it has been described for static problems, extension to the time-harmonic problem should be straightforward.

In the case of inhomogeneous or anisotropic or lossy material the method breaks down, as the simple free-space Green's function is inadequate: "action-at-a-distance" is most easily prescribed through free space. Here the picture-frame method (Chapters III and VI) is readily applied if the regions of inhomogeneity can be excised from the larger free-space region. The excised regions, the picture-frames, are most easily represented by the PDE isoparametric element technique, and the free-space region is readily represented in terms of an equivalent source distribution on the picture-frame boundaries. It is possible to mutually-constrain the PDE and IE finite elements, and a method for accomplishing this has been presented (Chapter VI). The procedure results in the field over the picture-frames and a source distribution on the picture-frame boundaries which may be used to obtain the potential at any point outside the picture-frames by an integration. Energy convergence for the mutually constrained solution is guaranteed, as the operator in each element has a positive-definite component.

Note that this picture-frame method requires explicit treatment of the Green's function singularities. The IE elements on the picture-frame boundaries are treated in the same fashion as the IE elements in the Dirichlet problems of Chapter IV. The Green's function singularities may be ignored in computation by the very simple device of defining an auxiliary contour within each picture-frame over which the IE expression is to be integrated (Chapter III). A matrix constraint equation is readily

obtained by point-matching: potentials at nodes on the picture-frame boundaries are determined in terms of potential on the auxilary contour by integration of the IE expression. The separation of the boundaries from the contour permits the Green's function to be treated as a regular function. The PDE element solutions within the picture-frames are obtained subject to the IE constraint. This technique allows simpler algorithms to be used, but it has two disadvantages in comparison with the mutual constraint technique presented in Chapter VI: a homogeneous region is required in each picture-frame between the auxiliary contour and the boundary; and energy convergence may be guaranteed only for the picture-frame solutions, as the exterior solution is obtained by point-matching.

An especially interesting situation results when the inhomogeneity is a dielectric, with constant relative permittivity ε_r. The methods of Chapters III, V and VI are all applicable. One may replace the dielectric by an equivalent interface source distribution and seek the solution by purely IE methods. One may also use a picture-frame method. To employ the mutual constraint method presented in Chapter VI one could place the picture-frame boundary at the air-dielectric interface. To use the simpler method of Chapter III a larger picture-frame is required, to allow the existence of a free-space region between the auxilary contour and the picture-frame boundary. The purely IE method has the fewest unknowns (only on the interface) but the matrix is completely dense. The picture-frame methods on the other hand require more unknowns (throughout the picture-frame), but the matrices may be relatively sparse. The algorithms available for solving matrix equations (Chapter VII) tend to favor sparse large matrices: dense large matrices are not as easily handled. This suggests that picture-frame methods may be preferable

when the geometry is very complicated; the relatively larger number of unknowns could be more than balanced by ease (and cost) with which the matrix equation may be solved.

The methods described here may be extended to the vector problems which result when Maxwell's equations [6, p.303] are not satisfied by the solution of simple scalar potential problems [6, p.315]. This occurs, for example, in propdgation problems with dielectric interfaces. English [7] and Daly [8] have described variational techniques for solution of the vector eigenproblem for bounded regions, and Harrington [9] has described moment methods for solution of the integral equations whic h involve vector current (rather than charge) distributions. As well, Csendes and Silvester [10-11] and Daly [12] have described coupled vector potential techniques by which loaded waveguides with uniform cross-section may be solved variationally. (Here there are but two unknowns at each point in space, rather than the six involved with a full vector representation.) Basic variational approaches for the vector problem may be found in Morse and Feshbach [13] and in Cairo and Kahan [14]. A systematic approach is indicated, as the vector problem is more complicated; but methods similar to those presented here are feasible.

The solution of problems is fundamental to the practice of engineering. The modular nature of a mutually constrained finite-element method permits the choice of the more convenient representation-field or source - for each part of a given problem. The general purpose nature of the procedures described here allows problems with arbitrary configuration to be handled, as there are no geometric restrictions on conductors or interfaces. If a region is large and homogeneous it is unreasonable to model the entire region when a source distribution may be sought on its boundary. If a region is small and inhomogeneous or anisotropic it is unreasonable to

wrestle with special Green's functions when a field solution is more easily obtained. It is sensible to use the method which is more convenient. Mutual constraint permits both methods to be applied together in solution of complex engineering problems.

8.1 References

[1] S.G. Mikhlin, Variational Methods in Mathematical Physics. New York: Macmillan, 1964.

[2] O.C. Zienkiewicz, The Finite Element Method in Engineering Science. London: McGraw-Hill, 1971.

[3] M. Friedman, D.J. Richards and A. Wexler, University of Manitoba Finite-Element Program (MANFEP). Dept. of Elect. Eng., University of Manitoba, Winnipeg, November 1973.

[4] M.C. Décreton. "Treatment of singularities in the finite-element method." M.Sc Thesis, University of Manitoba, Winnipeg, 1972.

[5] A. Wexler, "Computation of electromagnetic fields" (invited paper). IEEE Trans. MTT-17, pp.416-439, August, 1969.

[6] J.R. Reitz and F.J. Milford, Foundations of Electromagnetic Theory. Reading, Mass: Addison-Wesley, 1960.

[7] W.J. English, "A computer-implemented vector variational solution of loaded rectangular waveguides." SIAM J. App'l Math., Vol. 21, pp. 461-468, November, 1971.

[8] P. Daly, "Hybrid mode analysis of microstrip by finite-element methods." IEEE Trans., MTT-19, pp.19-25, January 1971.

[9] R.F. Harrington, Field Computation by Moment Methods. New York: Macmillan, 1968.

[10] Z.J. Csendes and P. Silvester, "Numerical solution of dielectric loaded waveguides: I-finite-element analysis." IEEE Trans., MTT-18, pp.1124-1131, December 1970.

[11] Z.J. Csendes and P. Silvester, "Numerical solution of dielectric loaded waveguides: II-modal approximation techniques." IEEE Trans. MTT-19, pp. 504-509, June 1971.

[12] P. Daly, "Finite-element coupling matrices." Elect. Lett., Vol. 5, pp.613-615, November 1969.

[13] P. Morse and H. Feshbach, Methods of Theoretical Physics. New York: McGraw-Hill, 1953.

[14] L. Cairo and T. Kahan, Variational Techniques in Electromagnetism. London: Blackie, 1965.

BIBLIOGRAPHY

F.S. Acton, Numerical Methods That Work. New York: Harper and Row, 1970.

J.L. Allen and M.F. Estes, "Broadside coupled stripe in a layered dielectric medium." IEEE Trans. MTT-20, pp.662-668, October 1972.

P.L. Arlett, A.K. Bahrani and O.C. Zienkiewicz, "Applications of finite-elements to the solution of Helmholtz's equation." Proc. IEE, Vol. 115, pp.1762-1766, December 1968.

P. Benedek and P. Silvester, "Capacitance of parallel rectangular plates separated by a dielectric sheet." IEEE Trans. MTT-20, pp.504-510, August 1972.

S. Bergman, Integral Operators in the Theory of Partial Differential Equations. New York: Springer-Verlag, 1969.

W. Braunbek, "On the diffraction field near a plane screen corner." IEEE Trans. AP-4, pp.219-223, April 1956.

T.G. Bryant and J.A. Weiss, "Parameters of microstrip transmission lines and of coupled pairs of microstrip lines." IEEE Trans. MTT-16, pp.1021-1027, December 1968.

A.J. Burton and G.F. Miller, "The application of integral equation methods to the numerical solution of some exterior boundary value problems." Proc. of the Royal Society, Series A, Vol. 323, No. 1553, pp.201-210, 8 June 1971.

P.F. Byrd and M.D. Friedman, Handbook of Elliptical Integrals for Engineers. Berlin: Springer, 1954.

L. Cairo and T. Kahan, Variational Techniques in Electromagnetism. London: Blackie, 1965.

B. Carnahan, H.A. Luther and J.O. Wilkes, Applied Numerical Methods. New York: Wiley, 1969.

I. Cermak, "Solution of unbounded field problems by boundary relaxation." Ph.D Thesis. Dept. of Elect. Eng., McGill University, Montreal, March 1969.

I. Cermak and P. Silvester, "Solution of two-dimensional field problems by boundary relaxation." Proc. IEE., Vol. 115, pp.1341-1348, September, 1968.

I. Cermak and P. Silvester, "Analysis of Coaxial line discontinuities by boundary relaxation." IEEE Trans. MTT-17, pp.489-495, August 1969.

Y. Chow and S. Wu, "A moment method with mixed linear basis functions for scatterings by waveguide junctions." IEEE Trans. MTT-21, pp.333-340, May 1973.

R.E. Collin, Field Theory of Guided Waves. New York: McGraw-Hill, 1960.

S.H. Crandall, Engineering Analysis. New York: McGraw-Hill, 1956.

J.G.A. Croll and A.C. Walker, "The finite-difference and localized Ritz methods." Int. Jnl. for Num. Meth. Eng. Vol. 3, pp.155-160 1971.

P.D. Crout, "An application of polynomial approximation to the solution of integral equations arising in physical systems." Jnl. Math. and Phys., Vol. 19, pp.34-92, 1940.

P.D. Crout and F.B. Hildebrand, "A least squares procedure for solving integral equations by polynomial approximation." Jnl. Math. and Phys., Vol. 20, pp.310-335, 1941.

Z.J. Csendes and P. Silvester, "Numerical solution of dielectric loaded waveguides: I-finite-element analysis." IEEE Trans., MTT-18, pp.1124-1131, December 1970.

Z.J. Csendes and P. Silvester, "Numerical solution of dielectric loaded waveguides: II-modal approximation techniques." IEEE Trans. MTT-19, pp.504-509, June 1971.

P. Daly, "Finite-element coupling matrices." Elect. Lett., Vol. 5, pp.613-615, November 1969.

P. Daly, "Hybrid mode analysis of microstrip by finite-element methods." IEEE Trans., MTT-19, pp.19-25, January 1971.

P. Daly and J.D. Helps, "Exact finite-element solutions to the Helmholtz equation." Int. Jnl. for Num. Meth. Eng., Vol. 6,

M. C. Décreton. "Treatment of singularities in the finite-element method." M.Sc. Thesis, University of Manitoba, Winnipeg, 1972.

A.F. dos Santos and V.R. Vieira, "The Green's function for Poisson's equation in a two dielectric region." IEEE Trans., MTT-20, p. 777, November 1972.

W.J. English, "A computer-implemented vector variational solution of loaded rectangular wave guides." SIAM J. App'l Math., Vol. 21, pp.461-468, November 1971.

A. Farrar and A.T. Adams, "Matrix methods for microstrip three dimensional problems." IEEE Trans., MTT-20, pp.497-504, August 1972.

B.A. Finlayson, The Method of Weighted Residuals and Variational Principles. New York: Academic Press, 1972.

M.J. Forray, Variational Calculus in Science and Engineering. New York: McGraw-Hill, 1968.

G.E. Forsythe and W.R. Wasow, Finite Difference Methods for Partial Differential Equations. New York: John Wiley and Sons, 1960.

J.N. Franklin, Matrix Theory. Englewood Cliffs, New Jersey: Prentice-Hall, 1968.

M. Friedman, D.J. Richards and A. Wexler, University of Manitoba Finite-Element Program (MANFEP). Dept. of Elect. Eng., University of Manitoba, Winnipeg, November 1973.

C.E. Froberg, Introduction to Numerical Analysis. Reading, Mass: Addison-Wesley, 1965.

D. Gish, "Characteristic Impedance and phase velocity of a dielectric supported air strip transmission line with sidewalls." IEEE Trans. MTT-18, pp.131-143, March 1970.

D. Greenspan and P. Werner, "A numerical method for the exterior Dirichlet problem for the reduced wave equations." Arch. Ration. Mech. Anal., Vol. 23, pp.288-316, 1966.

P. Hammond, "Sources and fields - some thoughts on teaching the principles of electromagnetism." Int. Jnl. Elect. Eng., Vol. 7, pp. 65-76, 1969.

R.F. Harrington, Field Computation by Moment Methods, New York: Macmillan, 1968.

R.F. Harrington, Time Harmonic Electromagnetic Fields. New York: McGraw-Hill, 1961.

T.G. Hazel and A. Wexler, "Variational formulation of the Dirichlet boundary condition." IEEE Trans., MTT-20, pp.385-390, June 1972.

F.R. Hildebrand, Methods of Applied Mathematics. Englewood Cliffs, N.J.: Prentice-Hall, 1965.

E. Isaacson and H.B. Keller, Analysis of Numerical Methods. New York: J. Wiley, 1966.

T. Itoh, R. Mittra and R. Ward, "A method for computing edge capacitance of finite and semi-infinite lines." IEEE Trans. MTT-20, pp.847-849, December 1972.

E. Johnke, F. Emde and F. Losch, Tables of Higher Functions. New York: McGraw-Hill, 1960.

D.S. Jones, The Theory of Electromagnetism. Oxford: Pergamon, 1964.

L.V. Kantorovich and V.I. Krylov, Approximate Methods of Higher Analysis. New York: J. Wiley, 1955.

O.D. Kellogg, Foundations of Potential Theory. Berlin: Springer-Verlag, 1929.

F. Kikuchi, "A finite-element method for non-self-adjoint problems." Int. Jnl. for Num. Meth. Eng., Vol 6, pp.39-54, 1973.

G. Kron, Diakoptics. London : MacDonald, 1963.

H. Levine and J. Schwinger, "On the theory of diffraction by an aperture in an infinite plane screen." Phys. Rev. Vol. 74, pp.958-974, 1948.

M. Maeda, "An analysis of gap in microstrip transmission lines." IEEE Trans. MTT-20, pp.390-396, June 1972.

B.H. McDonald, M. Friedman, M. Décreton and A. Wexler, "Integral finite-element approach for solving the Laplace equation." Elect. Lett., Vol. 9, pp.242-244, May 1973.

B.H. McDonald, M. Friedman and A. Wexler, "Variational solution of integral equations." IEEE Trans. MTT-22, pp.237-248, March 1974.

B.H. McDonald and A. Wexler, "Finite-element solution of unbounded field problems." IEEE Trans. MTT-20, pp.841-847, December 1972.

S.G. Mikhlin, Variational Methods in Mathematical Physics. New York: Macmillan, 1964.

S.G. Mikhlin, Integral Equations (and their applications to certain problems in mechanics, mathematical physics and technology). New York: Macmillan, 1964.

S.G. Mikhlin, Linear Integral Equations. Delhi, Hindustan Publ. Corp., 1960.

J. Minor and D. Bole, "Modes in the shielded microstrip on a finite substrate transversely magnetized in the plane of the substrate." IEEE Trans. MTT-19, pp.570-576, July 1971.

R. Mittra, T. Itoh and T. Li, "Analytical and numerical studies of the relative convergence phenomena arising in the solution of an integral equation by the moment method." IEEE Trans. MTT-20, pp.96-104, February 1972.

J. Molberg and D. Reynolds, "Iterative solutions of the scalar Helmholtz equation in lossy regions." IEEE Trans. MTT-17, pp.460-464, August 1969.

P, Morse and H. Feshbach., Methods of Theoretical Physics. New York: McGraw-Hill, 1953.

N.I. Muskhelishvili, Singular Integral Equations. Groningen: Noordhoff, 1946.

B. Noble, A Bibliography on: Methods for Solving Integral Equations. University of Wisconsin, Mathematics Research Centre, MRC 1176, September, 1971.

J.T. Oden, "A general theory of finite-elements I and II." Int. Jnl. Num. Meth. Eng., Vol. 1, pp.205-221, pp.247-260, 1969.

H.B. Palmer, "The capacitance of the parallel plate capacitor by the Schwartz-Christoffel transformation." AIEE Trans., Vol. 56, pp.363-366, 1937.

K. Pontoppidan, "Equivalent source solutions for electrostatic and magnetostatic problems." IEEE Int. Conv. Digest, New York, pp.620-621, March 1971.

J.K. Reid, (editor), Large Sparse Sets of Linear Equations. London: Academic Press, 1971.

J.R. Reitz and F.J. Milford, Foundations of Electromagnetic Theory. Reading, Mass: Addison-Wesley, 1960.

D.J. Richards and A. Wexler, "Finite-element solutions within curved boundaries." IEEE Trans. MTT-20, pp.650-657, October 1972.

G.F. Roach, Green's Functions - Introductory Theory with Applications. London: Van Nostrand Reinhold, 1970.

S.M. Robinson and A.H. Stroud, "The approximate solution of an integral equation using high-order Gaussian quadrature formulas." Maths. Comp. Vol. 15, pp.286-288, 1961.

T.E. Rozzi, "The variational treatment of thick interacting inductive irises" IEEE Trans. MTT-21, pp.82-88, February 1973.

F. Sandy and J. Sage, "Use of finite difference approximations to partial differential equations for problems having boundaries at infinity." IEEE Trans. MTT-19, pp. 484-486, May 1971.

P. Silvester, Modern Electromagnetic Fields. Englewood Cliffs, N.J. : Prentice-Hall, 1968.

P. Silvester, "High-order polynomial triangular finite-elements for potential problems." Int. Jnl. Eng. Sci. Vol. 7, pp.849-861, 1969

P. Silvester and P. Benedek, "Electrostatics of the microstrip-revisited." IEEE Trans. MTT-20, pp.756-758, November 1972.

P. Silvester and P. Benedek, "Microstrip discontinuity capacitance for right angle bends, T junctions and crossings. IEEE Trans. MTT-21, pp.341-346, May 1973.

P. Silvester and M.S. Hsieh, "Finite element solution of 2-dimensional exterior field problems." Proc. IEE, Vol. 118, pp.1743-1747, December 1971.

P. Silvester and M.S. Hsieh, "Projective solution of integral equations arising in electric and magnetic field problems." Int. Jnl. Comput Phys. Vol. 8, pp.73-82, 1971.

P. Silvester and A. Konrad, "Axisymmetric triangular finite-elements for the scalar Helmholtz equation." Int. Jnl. for Num. Meth. Eng., Vol. 5, pp.481-497, 1973.

W.R. Smythe, Static and Dynamic Electricity. New York: McGraw-Hill, 1968.

B.E. Spielman and R.F. Harrington, "Waveguides of arbitrary cross-section by solution of a nonlinear integral eigenvalue problem." IEEE Trans. MTT-20, pp.578-585, September 1972.

I. Stakgold, Boundary Value Problems of Mathematical Physics, Vol II. New York: Macmillan, 1968.

J.A. Stratton, Electromagnetic Theory. New York: McGraw-Hill,1941

A.H. Stroud, Approximate Calculation of Multiple Integrals. Englewood Cliffs, N.J.: Prentice-Hall, 1971.

A.H. Stroud and D. Secrest, Gaussian Quadrature Formulas. Englewood Cliffs, N.J. : Prentice-Hall, 1966.

C.T. Tai, Dyadic Green's Functions in Electromagnetic Theory. Scranton: INTEXT Educational Publishers, 1971.

V.K. Thong, "Solutions for some waveguide discontinuities by the method of moments." IEEE Trans. MTT-20, pp.416-418,June 1972.

F.G. Tricomi, Integral Equations. New York: Interscience, 1957.

R.S. Varga, Matrix Iterative Analysis. Englewood Cliffs, N.J.: Prentice-Hall, 1965.

A. Wexler, "Computation of electromagnetic fields" (invited paper). IEEE Trans. MTT-17, pp.416-439, August 1969.

A. Wexler, "Finite-element field analysis of an inhomogeneous, anisotropic, reluctance machine rotor." IEEE Trans. PAS-92, pp.145-149, January 1973.

H.A. Wheeler, "Transmission line properties of parallel strips separated by a dielectric sheet." IEEE Trans. MTT-13, pp.172-185, March 1965.

J.R. Whiteman (editor), The Mathematics of Finites Elements and Applications. London: Academic Press, 1973.

C.G. Williams and G.K. Cambrell, "Efficient numerical solution of unbounded field problems." Elect. Letts. Vol. 8, pp.247-248 May 1972.

C.R. Wylie, Advanced Engineering Mathematics. New York: McGraw-Hill, 1960.

E. Yamashita and R. Mittra, "Variational Method for the analysis of microstrip lines." IEEE Trans. MTT-16, pp.251-256, April 1968.

O.C. Zienkiewicz, The Finite Element Method in Engineering Science. London: McGraw-Hill, 1971.

M. Zlámal, "The finite-element method in domains with curved boundaries." Int. Jnl. for Num. Meth. Eng., Vol. 5, pp.367-373, 1973.

M. Zlámal, "On the finite-element method." Num. Math. Vol.12, pp.394-409, 1968.

APPENDIX A

MATRIX VECTOR NOTATION

The vector $\underline{r}$ is used generally as a position vector. In two dimensional space $\underline{r}^T=(x\ y)$ and in three, $\underline{r}^T=(x\ y\ z)$. Therefore $u(\underline{r})=u(x,y)$. When a set of N such functions is used, these functions may be collected in a vector;

$$\underline{u}(\underline{r}) = \begin{pmatrix} u_1(x,y) \\ \vdots \\ u_N(x,y) \end{pmatrix} \tag{A-1}$$

for example in two-dimensional space.

When a linear combination of these N functions is required the vector dot product may be used

$$v(\underline{r}) = \sum_{i=1}^{N} c_i\, u_i\,(x,y) = \underline{c} \cdot \underline{u} \tag{A-2}$$

where $\underline{c}$ is the vector of coefficients. Alternatively, regular matrix multiplication may be used if the first vector is transposed:

$$v(\underline{r}) = \underline{c}^T\underline{u} = (c_1 \ldots c_N) \begin{pmatrix} u_1 \\ \vdots \\ u_N \end{pmatrix} . \tag{A-3}$$

The dyadic product of two vectors is a matrix, and arises when the second vector is transposed:

$$\underline{v}\,\underline{u}^T = \begin{pmatrix} v_1 \\ \cdot \\ \cdot \\ u_N \end{pmatrix} (u_1 \ldots u_N) = \begin{bmatrix} v_1u_1 & \cdots & v_1u_N \\ \cdot & & \cdot \\ \cdot & & \cdot \\ v_Nu_1 & \cdots & v_Nu_N \end{bmatrix} . \tag{A-4}$$

The following scalar matrix equation is quadratic in $\underline{x}$:

$$F(\underline{x}) = \underline{x}^T [A] \underline{x} + \underline{x}^T \underline{b} + c \quad . \tag{A-5}$$

The derivative with respect to the vector $\underline{x}$ is

$$\underline{f}(\underline{x}) = [A+A^T] \underline{x} + \underline{b} \tag{A-6}$$

and is a vector. It may be obtained by taking $\partial F/\partial x_i$, which creates the row of $\underline{f}$, or it may be done by differentiating directly:

$$\frac{\partial F(\underline{x})}{\partial \underline{x}} = (x^T [A])^T + [A] \underline{x} + \underline{b}$$

$$= [A+A^T] \underline{x} + \underline{b} \quad . \tag{A-7}$$

The proof is obtained by expanding (A-5)

$$F(\underline{x}) = x_1 (A_{11} x_1 + + A_{1N} x_N + b_1) + + x_N (A_{N1} x_1 + + A_{NN} x_N + b_N) + c \tag{A-8}$$

and taking derivatives :

$$\frac{\partial F(\underline{x})}{\partial \underline{x}} = \begin{bmatrix} \frac{\partial F(\underline{x})}{\partial x_1} \\ \vdots \\ \frac{\partial F(\underline{x})}{\partial x_N} \end{bmatrix} = \begin{bmatrix} (A_{11}x_1 + + A_{1N}x_N + b_1) + (x_1 A_{11} + + x_N A_{N1}) \\ \vdots \\ (A_{N1}x_1 + + A_{NN}x_N + b_N) + (x_1 A_{1N} + + x_N A_{NN}) \end{bmatrix} , \tag{A-9}$$

which is the same as

$$\frac{\partial F(\underline{x})}{\partial \underline{x}} = \begin{bmatrix} A_{11} + A_{11} & \cdots & A_{1N} + A_{N1} \\ & & \\ A_{N1} + A_{1N} & \cdots & A_{NN} + A_{NN} \end{bmatrix} \underline{x} + \underline{b}$$

$$= [A + A^T] \underline{x} + \underline{b} \quad . \tag{A-10}$$

Note that the resulting matrix is always symmetric. The second derivative of F is a matrix

$$\frac{\partial^2 F(\underline{x})}{\partial \underline{x}^2} = \left[A + A^T\right] \tag{A-11}$$

which arises by taking

$$\frac{\partial^2 F(\underline{x})}{\partial \underline{x}^2} = \begin{bmatrix} \frac{\partial}{\partial x_1} \left\{ \frac{\partial F(\underline{x})}{\partial \underline{x}} \right\} \\ \vdots \\ \frac{\partial}{\partial x_N} \left\{ \frac{\partial F(\underline{x})}{\partial \underline{x}} \right\} \end{bmatrix} = \left[A + A^T\right] . \tag{A-12}$$

APPENDIX B

IMAGE SYMMETRIES

Consider the problem of integrating

$$\phi(\underline{r}) = \int_S \sigma(\underline{r}')\, G(\underline{r}|\underline{r}')\, ds' \tag{B-1}$$

with the image symmetry shown in Fig. B.1. For every point $\underline{r}$ there is an image point $\underline{r}_I$. The integral may be computed in two parts:

$$\phi(\underline{r}) = \int_{S_1} \sigma(\underline{r}')\, G(\underline{r}|\underline{r}')ds' + \int_{S_2} \sigma(\underline{r}_I')\, G(\underline{r}|\underline{r}_I')\, ds_I' . \tag{B-2}$$

Now, by virtue of the symmetries, the image contribution (Fig. B.1c) may be included as follows:

$$\phi(\underline{r}) = \int_{S_1} \sigma(\underline{r}')\, \{ G(\underline{r}|\underline{r}') + G(\underline{r}|\underline{r}_I')\}\, ds' . \tag{B-3}$$

The image Green's function is

$$G(\underline{r}|\underline{r}') + G(\underline{r}|\underline{r}_I') .$$

This method is used in Chapter IV.

An alternative method is to compute

$$\phi_1(\underline{r}) = \int_{S_1} \sigma(\underline{r}')\, G(\underline{r}|\underline{r}')\, ds' \tag{B-4a}$$

and, as shown in Fig. B.1d:

$$\phi_2(\underline{r}_I) = \int_{S_1} \sigma(\underline{r}')\, G(\underline{r}_I|\underline{r}')\, ds' . \tag{B-4b}$$

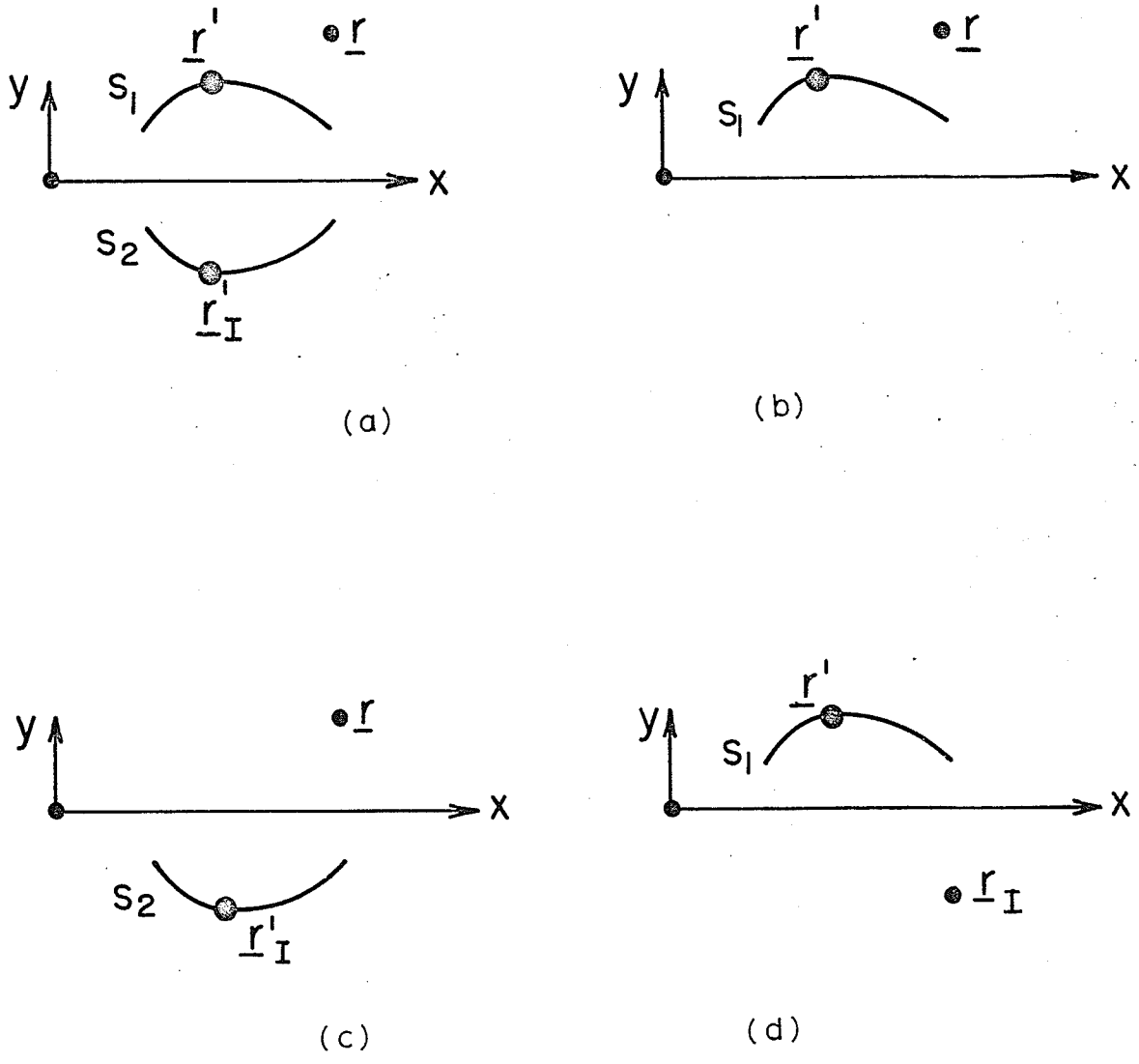


Fig. B.1 Mirror Image Symmetries. a) Full Problem. b) Major Contribution. c) Image Contribution (source point mapped). d) Image Contribution (observation point mapped).

Then, since G is symmetric

$$\phi(\underline{r}) = \phi_1(\underline{r}) + \phi_2(\underline{r}_I) \quad . \tag{B-5}$$

The form at (B-3) arises by mapping the source point across the plane of symmetry, and at (B-5) by mapping the observation point.

Consider now the integral

$$\phi(\underline{r}) = \int_S \left\{ \frac{\partial \phi(\underline{r}')}{\partial n'} G(\underline{r}|\underline{r}') - \phi(\underline{r}') \frac{\partial G(\underline{r}|\underline{r}')}{\partial n'} \right\} ds' . \tag{B-6}$$

If the source point $\underline{r}'$ is mapped, so must be the normal derivatives: it is much more convenient to map the observation point:

$$\phi_1(\underline{r}) = \int_{S_1} \left\{ \frac{\partial \phi(\underline{r}')}{\partial n'} G(\underline{r}|\underline{r}') - \phi(\underline{r}') \frac{\partial G(\underline{r}|\underline{r}')}{\partial n'} \right\} ds' \tag{B-7a}$$

and

$$\phi_2(\underline{r}_I) = \int_{S_1} \left\{ \frac{\partial \phi(\underline{r}')}{\partial n'} G(\underline{r}_I|\underline{r}') - \phi(\underline{r}') \frac{\partial G(\underline{r}_I|\underline{r}')}{\partial n'} \right\} ds' . \tag{B-7b}$$

Then, using (B-5)

$$\phi(\underline{r}) = \int_{S_1} \left\{ \frac{\partial \phi(\underline{r}')}{\partial n'} \left\{ G(\underline{r}|\underline{r}') + G(\underline{r}_I|\underline{r}') \right\} \right.$$

$$\left. - \phi(\underline{r}') \left\{ \frac{\partial G(\underline{r}|\underline{r}')}{\partial n'} + \frac{\partial G(\underline{r}_I|\underline{r}')}{\partial n'} \right\} \right\} ds' . \tag{B-8}$$

This is the method used for the mirror symmetries in the problems of Chapters III, V, and VI. In positive symmetry contributions are added, as above. For negative symmetry (antisymmetry), contributions are subtracted.

APPENDIX C

THE ADDITION-SUBTRACTION INTEGRATION TECHNIQUE

The method was suggested for use here by M. Friedman. It is described by Isaacson and Keller [1, pp.347-349] and a form was given by Robinson and Stroud [2] based on a suggestion in Kantorovich and Krylov [3, p.101].

Consider the following integral

$$I = \int_a^b f(x)\, g(x)\, dx \qquad \text{(C-1)}$$

where g(x) is a regular function and where f(x) has an integrable singularity at x=c. Suppose that there exists a simply integrable function h(x) such that

$$R(x) = f(x) - h(x) \qquad \text{(C-2)}$$

is regular in $a \leq x \leq b$.

Then I may be integrated as

$$I = \int_a^b (f(x)g(x) - h(x)g(c)\ )\, dx + g(c) \int_a^b h(x)\, dx$$

$$= \int_a^b R(x)g(x)\, dx + \int_a^b h(x)\{g(x)-g(c)\}dx + g(c) \int_a^b h(x)\, dx \qquad \text{(C-3)}$$

The first integrand is regular, and Gauss quadrature may be used. The last integral may be obtained analytically, by assumption.

The integral in question is

$$I_2 = \int_a^b h(x)\{g(x) - g(c)\}\, dx = \int_a^b p(x)dx \qquad \text{(C-4)}$$

and Fig. C.1 indicates that p(x) is regular. A Taylor series expansion of g about x=c:

$$g(x) = g(c) + \sum_{i=1}^{\infty} (x-c)^i \frac{g^{(i)}(c)}{i!} \qquad \text{(C-5)}$$

allows one to write

$$p(c) = \lim_{x \to c} (x-c)\, h(x)\, g'(c) = 0$$

for the functions used here.

Unfortunately there could be an infinite derivative at x=c:

$$p'(x) = h'(x)\,\{g(x)-g(c)\} + h(x)\{g'(x)\} \qquad \text{(C-7)}$$

and using (C-5)

$$p'(c) = \lim_{x \to c} \{(x-c)\, h'(x)g'(c) + h(x)g'(x)\} \sim h(c)g'(c)\,. \qquad \text{(C-8)}$$

Unless g'(c)=0, p' is singular at x=c. Fig. C.2 shows in more detail what happens at x=c. In such circumstances it is better to perform the numerical quadrature in two parts:

$$I_2 = \int_a^c p(x)dx + \int_c^b p(x)dx \qquad \text{(C-9)}$$

so that the derivative singularity does not occur within the integration range.

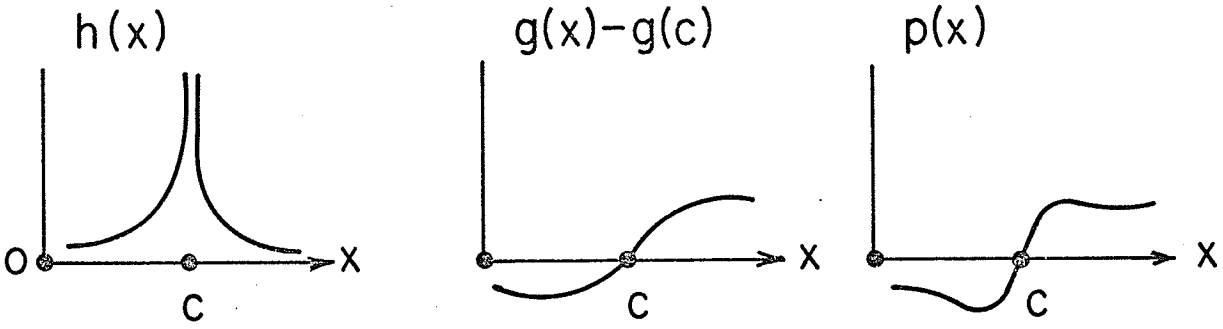


Fig. C.1 Constructing the Regular Function p(x)

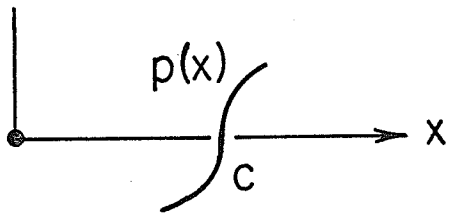


Fig. C.2 Close-up of behaviour near x=c.

References

[1] E. Isaacson and H.B. Keller, Analysis of Numerical Methods. New York: J. Wiley, 1966.

[2] S.M. Robinson and A.H. Stroud, "The approximate solution of an integral equation using high-order Gaussian quadrature formulas." Maths. Comp., Vol. 15, pp. 286-288, 1961.

[3] L.V. Kantorovich and V.I. Krylov, Approximate Methods of Higher Analysis. New York: J. Wiley, 1955.

APPENDIX D

GAUSSIAN QUADRATURE

The integral of f(x) weighted by $\omega(x)>0$ is sought as follows [1]

$$\int_a^b \omega(x) f(x) dx \simeq \sum_{i=1}^{N} a_i f(x_i) \qquad \text{(D-1)}$$

where a_i is a Gauss coefficient and x_i is a Gauss integration point. N is the order of the quadrature.

Suppose that the weights may be integrated analytically:

$$c_i = \int_a^b \omega(x) x^i \, dx \quad ; \quad i=0, 1, 2 \ldots N \, . \qquad \text{(D-2)}$$

Consider the orthogonal polynomial defined by

$$P_N(x) = x^N + \sum_{i=0}^{N-1} k_i x^i \; ; \qquad \text{(D.3a)}$$

$$\int_a^b \omega(x) x^j P_N(x) dx = 0 \; ; \; j=0, 1 \ldots N-1 \quad , \qquad \text{(D.3b)}$$

and this may be written as

$$\sum_{i=0}^{N-1} k_i \int_a^b \omega(x) x^{i+j} dx = - \int_a^b \omega(x) x^{j+N} dx; \; j=0 \ldots N-1 \qquad \text{(D-4)}$$

This defines a matrix equation for the k_i. Using (D-2)

$$\begin{bmatrix} \ddots & \vdots \\ \cdots & c_{i+j} \end{bmatrix} \begin{pmatrix} \vdots \\ k_i \end{pmatrix} = - \begin{pmatrix} \vdots \\ c_{i+N} \end{pmatrix} \, . \qquad \text{(D-5)}$$

The N roots of $P_N(x)$ are now found and used as Gauss integration points. The coefficients a_i are obtained by forcing exact integration for polynomials to order N-1:

$$\int_a^b \omega(x)x^j\,dx = c_j = \sum_{i=1}^{N} a_i\, x_i^{\,j}\;;j=0\ldots N-1 \qquad \text{(D-6)}$$

or

$$\begin{bmatrix} \cdots & x_i^{\,j} & \cdots \end{bmatrix} \begin{pmatrix} \vdots \\ a_i \\ \vdots \end{pmatrix} = \begin{pmatrix} \vdots \\ c_j \\ \vdots \end{pmatrix} \quad . \qquad \text{(D-7)}$$

By definition of the orthogonal polynomial (D-3b)

$$\int_a^b \omega(x)\, x^j P_N(x)\,dx \;\simeq\; \sum_{i=1}^{N} a_i x^j P_N(x_i) = 0 \quad . \qquad \text{(D-8)}$$

The right-hand is zero because P_N vanishes at its roots. The left side is zero for $j \leq N-1$ (equation (D-3b)). The maximum degree of the polynomial is N+N-1. Therefore any polynomial of order $j \leq 2N-1$ is integrated exactly by this selection of a_i and x_i.

Consider a simple example :

$$\int_0^1 f(x)dx \;\sim\; a_1 f(x_1) + a_2 f(x_2) \quad . \qquad \text{(D-9)}$$

Here $\omega=1$. Immediately

$$c_i = \int_0^1 x^i\,dx = 1/i+1 \quad ; \qquad \text{(D-10)}$$

the polynomial is

$$P_2(x) = x^2 + k_0 + k_1 x \quad . \qquad \text{(D-11)}$$

Then

$$\begin{bmatrix} 1 & 1/2 \\ 1/2 & 1/3 \end{bmatrix} \begin{bmatrix} k_0 \\ k_1 \end{bmatrix} = - \begin{bmatrix} 1/3 \\ 1/4 \end{bmatrix} \tag{D-12}$$

and

$$P_2(x) = x^2 + 1/6 - x = \left(x - \frac{\sqrt{3}+1}{2\sqrt{3}}\right) \cdot \left(x - \frac{\sqrt{3}-1}{2\sqrt{3}}\right) , \tag{D-13}$$

and the roots are obtained. The Gauss coefficients are obtained from

$$\begin{bmatrix} \frac{1}{\sqrt{3}+1} & \frac{1}{\sqrt{3}-1} \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = \begin{bmatrix} 1 \\ \sqrt{3} \end{bmatrix} \tag{D-14}$$

and $a_0=a_1=1/2$. Therefore

$$\int_0^1 f(x)dx \simeq \frac{1}{2}\left[f\left(\frac{\sqrt{3}-1}{2\sqrt{3}}\right) + f\left(\frac{\sqrt{3}+1}{2\sqrt{3}}\right)\right] \tag{D-15}$$

This formula is exact for x^0, x^1, x^2, x^3 and produces a value of 7/36 for $f(x)=x^4$.

Gauss formulae are tabulated in [2] for several $\omega(x)$ and are usually over the range (-1,1), where Legendre polynomials may be used as the $P_N(x)$.

If $$\int_{-1}^{1} \omega(x)\, f(x)dx \simeq \sum_{i=1}^{N} a_i\, f(x_i)$$

Then

$$\int_p^q \omega(x)f(x)dx = \frac{q-p}{2} \sum_{i=1}^{N} a_i\, f\left[\frac{p+q}{2} + x_i\left(\frac{q-p}{2}\right)\right] . \tag{D-16}$$

References

[1] A.H. Stroud, Approximate Calculation of Multiple Integrals. Englewood Cliffs, N.J.: Prentice-Hall, 1971.

[2] A.H. Stroud and D. Secrest, Gaussian Quadrature Formulas. Englewood Cliffs, N.J.: Prentice-Hall, 1966.

APPENDIX E

LOCATING A POINT IN A REGULAR TRIANGLE

Consider the regular triangle and the local element shown in Fig. E.1. The problem is to determine if a point (x,y) lies in the global element. Consider mapping the triangle into the local element by the transformation

$$\zeta = \frac{(1\ x\ y)}{D} \begin{pmatrix} x_c y_a - x_a y_c \\ y_c - y_a \\ x_a - x_c \end{pmatrix} \tag{E-1}$$

$$\eta = \frac{(1\ x\ y)}{D} \begin{pmatrix} x_a y_b - x_b y_a \\ y_a - y_b \\ x_b - x_a \end{pmatrix} \tag{E-2}$$

where D is the area of the triangle:

$$D = x_b y_c + x_a y_b + x_c y_a - x_b y_a - x_c y_b - x_a y_c \ . \tag{E-3}$$

If the point (ζ,η) is in the local element, (x,y) is in the triangle. If the following conditions are all true, (ζ,η) is in the local element:

$$\begin{aligned} \zeta\ (1-\zeta) &> 0 \\ \eta\ (1-\eta) &> 0 \\ \zeta+\eta &< 1 \ . \end{aligned} \tag{E-4}$$

This algorithm was suggested by M. Friedman.

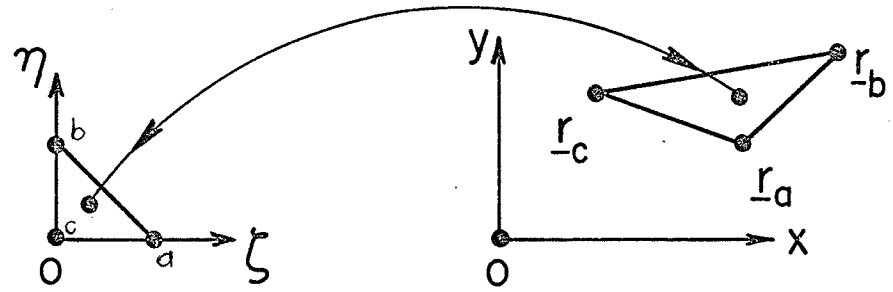


Fig. E.1 A Regular Triangle Mapped to the Local Element