Self-Adjoint Variational Formulation of Problems Having Non-Self-Adjoint Operators

GARBO JENG AND ALVIN WEXLER, MEMBER, IEEE

Abstract—**A systematic approach is given for deriving a variational formulation, previously stated by Stakgold, of non-self-adjoint operators from the standard quadratic functional for self-adjoint operators given by Mikhlin. If the same set of basis functions is used to approximate the solution of the operator equation and its adjoint equation, the resulting system of equations is identical to that derived from the Galerkin method. By using two differing sets of basis functions, one obtains a system of equations which corresponds to that derived from the moment method in general. As a particular and important example, the integral equation for the interface problem between differing media is considered. Compared to the method used by McDonald, Friedman, and Wexler, the present formulation involves no danger of finding a false solution, results in a simpler set of equations, requires fewer integrations, and is seen—in the case of integral equations—to correspond to the Galerkin method. It is also shown that for wave propagation through a lossy medium, which involves the solution of the non-self-adjoint complex Helmholtz equation, the resulting system of linear equations takes the same form as those for the real self-adjoint case but for the addition of complex arithmetic.**

I. Introduction

IN A RECENT paper [1] by McDonald, Friedman, and Wexler, the integral equation formulation of electrostatic field problems was solved using a variational method. The interface problem resulted in integral equations whose kernels are nonsymmetric. Hence, the quadratic functional [2, pp. 74–95]

$$F = \langle K\sigma,\sigma\rangle - \langle \sigma,g\rangle - \langle g,\sigma\rangle \tag{1}$$

whose stationary point corresponds to the solution of

$$K\sigma = g \tag{2}$$

provided K is self-adjoint (that is, real symmetric kernel for the case of an integral operator), is not directly applicable.

In order to apply (1) to the non-self-adjoint interface problem, McDonald *et al.* [1] proceeded by defining a modified self-adjoint operator K' using the following relation

$$K'\sigma = \langle K\sigma,G\rangle \tag{3}$$

where G is the free space Green's function of the problem concerned. It was argued that $K'\sigma = 0$ implies $K\sigma = 0$ (for the homogeneous case of $g = 0$), and hence, a solution of $K'\sigma = 0$ implies a solution of $K\sigma = 0$. However, $K\sigma = 0$ is a sufficient but not a necessary condition for $K'\sigma$ to vanish. For example, it is obvious that certain choices of the approximating function to σ cause the inner product $\langle K\sigma,G\rangle$ to vanish due to orthogonality of $K\sigma$ and G. When this is the case, a solution to $K'\sigma = 0$ need not form an approximate solution to $K\sigma = 0$.

Here an adjoint formulation, derived from the quadratic functional for self-adjoint operators, is used so that the danger of finding a false solution does not exist. Additionally, the extra surface integration, as introduced from the definition of a new operator using relation (3), is not required.

II. Extension of the Quadratic Functional for Non-Self-Adjoint Operators

Consider the inhomogeneous equation

$$K\sigma = g. \tag{4}$$

Assume that there is an adjoint operator K^a to the operator K with the property

$$\langle K\sigma,\tau\rangle = \langle \sigma,K^a\tau\rangle \tag{5}$$

where the pair of brackets is defined by

$$\langle u,v\rangle = \int uv^*\, d\Omega \tag{6}$$

and the superscript * denotes the complex-conjugate. An operator K is said to be self-adjoint if

$$\langle K\sigma,\tau\rangle = \langle \sigma,K\tau\rangle. \tag{7}$$

Equation (7) follows directly from (5) if K is self-adjoint, that is, if $K = K^a$.

According to Mikhlin [2, pp. 74–95], if K is self-adjoint, the quadratic functional

$$F = \langle K\sigma,\sigma\rangle - \langle \sigma,g\rangle - \langle g,\sigma\rangle \quad \text{(for complex } \sigma \text{ and/or } g) \tag{8}$$

or

$$F = \langle K\sigma,\sigma\rangle - 2\langle \sigma,g\rangle \quad \text{(for the real case)} \tag{9}$$

is stationary at a solution of

$$K\sigma = g. \tag{10}$$

To show this, let σ_0 be a stationary point of the functional (8) and consider a small perturbation $\varepsilon\eta$ from the stationary point, that is,

$$F(\varepsilon) = F(\sigma_0 + \varepsilon\eta) \tag{11}$$

Manuscript received December 4, 1975; revised August 25, 1977. This work was supported by the National Research Council of Canada.

The authors are with the University of Manitoba, Department of Electrical Engineering, Winnipeg, Man., Canada R3T 2N2.

0018-9480/78/0200-0091$00.75

where η is an arbitrary function in the domain of definition of the operator K and ε is a real scalar multiplier. Expanding (11) using (8), we have

$$F(\varepsilon) = \langle K\sigma_0,\sigma_0\rangle + \varepsilon\langle K\eta,\sigma_0\rangle + \varepsilon\langle K\sigma_0,\eta\rangle + \varepsilon^2\langle K\eta,\eta\rangle - \varepsilon\langle\eta,g\rangle - \varepsilon\langle g,\eta\rangle - \langle\sigma_0,g\rangle - \langle g,\sigma_0\rangle.$$

The first derivative of the functional (8) at $\varepsilon = 0$ is given by

$$\left.\frac{dF}{d\varepsilon}\right|_{\varepsilon=0} = \langle K\eta,\sigma_0\rangle + \langle K\sigma_0,\eta\rangle - \langle\eta,g\rangle - \langle g,\eta\rangle \quad (12)$$

which must vanish at the stationary point. Hence,

$$\langle K\eta,\sigma_0\rangle + \langle K\sigma_0,\eta\rangle - \langle\eta,g\rangle - \langle g,\eta\rangle = 0. \quad (13)$$

The self-adjointness of K allows (13) to be written as

$$\langle\eta,\ K\sigma_0 - g\rangle + \langle K\sigma_0 - g,\ \eta\rangle = 0. \quad (14)$$

Since the mapping defined by (6) satisfies the following properties

1) $\langle u,v\rangle = \langle v,u\rangle^*$

2) $\langle a_1u_1 + a_2u_2,\ v\rangle = a_1\langle u_1,v\rangle + a_2\langle u_2,v\rangle$

3) $\langle u,u\rangle \geq 0$

4) if $\langle u,u\rangle = 0$, then $u \equiv 0$ (15)

of an inner product, (14) yields, for arbitrary η, $K\sigma_0 - g = 0$.

Careful study of the steps leading to this result reveals that the requirement of an inner product is unnecessarily restrictive insofar as the proof of stationarity is concerned. To prove stationarity, without requiring a minimum of the functional, the following conditions are sufficient:

1) The operator K satisfies the relation (7).

2) $\langle a_1u_1 + a_2u_2,\ v\rangle = a_1\langle u_1,v\rangle + a_2\langle u_2,v\rangle$

3) $\langle u,\ a_1v_1 + a_2v_2\rangle = a_1^*\langle u,v_1\rangle + a_2^*\langle u,v_2\rangle$ (16)

4) if $\langle u,v\rangle + \langle v,u\rangle = 0$

for arbitrary v, then $u \equiv 0$.

Obviously, an inner product satisfies the above properties 2)–4) and is a more restrictive mapping. We shall now derive a functional for non-self-adjoint operators based on the functional (8).

First, forming a matrix operator

$$\mathcal{K} = \begin{bmatrix} K & 0 \\ 0 & K^a \end{bmatrix}.$$

The domain and the range of the operator $\mathcal{K}$ are formed by the elements

$$\tilde{\sigma} = \begin{bmatrix} \sigma \\ \tau \end{bmatrix} \quad \text{and} \quad \tilde{g} = \begin{bmatrix} g \\ h \end{bmatrix}$$

respectively, where τ and h are in the domain and the range, respectively, of the operator K^a.

If

$$\tilde{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \quad \text{and} \quad \tilde{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$$

are two elements in the domain of definition of $\mathcal{K}$, we define

$$\langle\tilde{u},\tilde{v}\rangle = \int W\tilde{u}\cdot\tilde{v}^*\,d\Omega \quad (17)$$

where W is a weighting factor defined by

$$W = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}.$$

For example,

$$\langle\mathcal{K}\tilde{u},\tilde{u}\rangle = \int \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} K & 0 \\ 0 & K^a \end{bmatrix}\begin{bmatrix} u_1 \\ u_2 \end{bmatrix}\cdot\begin{bmatrix} u_1 \\ u_2 \end{bmatrix}^* d\Omega$$

$$= \int Ku_1u_2^*\,d\Omega$$

$$= \langle Ku_1,u_2\rangle. \quad (18)$$

Using (17), it is easily shown that

$$\langle\mathcal{K}\tilde{u},\tilde{v}\rangle = \langle\tilde{u},\mathcal{K}\tilde{v}\rangle.$$

Thus (7) is satisfied. Moreover, the mapping defined by (17) satisfies the properties 2)–4) of (16) as required for stationarity. However, as is obvious from (18), even if $\mathcal{K} = I$, an identity operator, one cannot conclude that $\langle I\tilde{u},\tilde{u}\rangle \geq 0$. The definition given by (17) is thus not applicable for cases where positive definiteness is to be proven. However, it is sufficient for the purpose of proving stationarity. Thus the functional

$$F = \langle\mathcal{K}\tilde{\sigma},\tilde{\sigma}\rangle - \langle\tilde{\sigma},\tilde{g}\rangle - \langle\tilde{g},\tilde{\sigma}\rangle \quad (19)$$

is stationary at a solution of

$$\mathcal{K}\tilde{\sigma} = \tilde{g}.$$

Using (17), the functional (19) is

$$F = \langle K\sigma,\tau\rangle - \langle g,\tau\rangle - \langle\sigma,h\rangle. \quad (20)$$

The above functional, which is here derived in a direct fashion, is identical to that stated in [3, p. 357]. The stationary points of F, with respect to τ and σ, give the solution to

$$K\sigma = g$$

and

$$K^a\tau = h$$

respectively. Note that by taking the approximating functions σ and τ to be

$$\sigma = \boldsymbol{\sigma}^T\boldsymbol{\alpha} = \boldsymbol{\alpha}^T\boldsymbol{\sigma} \quad (21)$$

and

$$\tau = \boldsymbol{\tau}^T\boldsymbol{\alpha} = \boldsymbol{\alpha}^T\boldsymbol{\tau} \quad (22)$$

where $\boldsymbol{\alpha}$ is a vector of interpolating functions and $\boldsymbol{\sigma}$ and $\boldsymbol{\tau}$ contain the variational parameters, the functional given by (20) produces

$$F = \boldsymbol{\sigma}^T\langle K\boldsymbol{\alpha},\boldsymbol{\alpha}^T\rangle\boldsymbol{\tau} - \langle g,\boldsymbol{\alpha}^T\rangle\boldsymbol{\tau} - \boldsymbol{\sigma}^T\langle\boldsymbol{\alpha},h\rangle. \quad (23)$$

Upon taking the variation with respect to τ, (23) results in a system of equations

$$\langle K\boldsymbol{\alpha},\boldsymbol{\alpha}^T\rangle\boldsymbol{\sigma} = \langle g,\boldsymbol{\alpha}\rangle \tag{24}$$

which is algorithmically identical to that obtained from Galerkin's method. It may also be noted that by taking the set of approximating functions to be

$$\sigma = \boldsymbol{\sigma}^T\boldsymbol{\alpha} = \boldsymbol{\alpha}^T\boldsymbol{\sigma} \tag{25}$$

and

$$\tau = \boldsymbol{\tau}^T\boldsymbol{\beta} = \boldsymbol{\beta}^T\boldsymbol{\tau} \tag{26}$$

the system of equations (24) is

$$\langle \boldsymbol{\beta},K\boldsymbol{\alpha}^T\rangle\boldsymbol{\sigma} = \langle \boldsymbol{\beta},g\rangle \tag{27}$$

which is the system of equations resulting from the moment method in general.

III. Application to the Interface Problem

From the condition of continuity of flux at the interface, it has been shown [1] that the equation governing the charge at the interface between dielectrics having relative permittivities ε_1 and ε_2, respectively, is given by

$$\frac{\varepsilon_1+\varepsilon_2}{2\varepsilon_0\varepsilon_2}\sigma(s) + \frac{\varepsilon_1-\varepsilon_2}{\varepsilon_0\varepsilon_2}\int_s \sigma(s')\frac{\partial G}{\partial n}(s\,|\,s')\,ds' = 0 \tag{28}$$

where $\sigma(s)$ is an equivalent single-layer charge distribution. This is a Fredholm integral equation of the second kind. From the definition of the adjoint operator given by (5), the adjoint of (28) is obtained by replacing the Kernel $(\partial G/\partial n)(s\,|\,s')$ by $(\partial G/\partial n')(s'\,|\,s)$. In general, the adjoint operator of a complex integral operator is one with the kernel replaced by its complex-conjugate transpose. For a general curve, it is easily shown that

$$\frac{\partial G}{\partial n}(s\,|\,s') \neq \frac{\partial G}{\partial n'}(s'\,|\,s). \tag{29}$$

Hence, the integral operator of (28) is not self-adjoint. Using (20), the functional for the operator equation (28) is

$$F = \int_s \tau(s)\left[\frac{\varepsilon_1+\varepsilon_2}{2\varepsilon_0\varepsilon_2}\sigma(s) + \frac{\varepsilon_1-\varepsilon_2}{2\varepsilon_0\varepsilon_2}\int \sigma(s')\cdot\frac{\partial G}{\partial n}(s\,|\,s')\,ds'\right]ds. \tag{30}$$

As an example, the 3-D microstrip problem presented in [1] is solved by the Ritz method using functional (30). A T-shaped thin conductor at unit potential is placed on top of a dielectric slab, with relative permittivity ε_r, lying on an infinite ground conductor in the $z = 0$ plane. On the conductor plate S_P, the potential ϕ is constant, say $\phi(s) = g$. At the dielectric interface S_I, (28) holds. Therefore, we have

$$\int_S G(s\,|\,s')\sigma(s')\,ds' = g, \qquad s \in S_P \tag{31}$$

$$\frac{\varepsilon_r+1}{2\varepsilon_0}\sigma(s) + \frac{\varepsilon_r-1}{\varepsilon_0}\int_S K(s\,|\,s')\sigma(s')\,ds' = 0, \qquad s \in S_I \tag{32}$$

where

$$S = S_I \cup S_P$$

$$G(s\,|\,s') = \frac{1}{4\pi|s-s'|}$$

and

$$K(s\,|\,s') = \frac{\partial G}{\partial n}(s\,|\,s').$$

For the purpose of comparing results with those reported in [1], the unknown functions σ and τ are approximated by constants over square equal-area subregions. An improved technique, using triangular finite elements with higher order approximations and applicable to problems having arbitrary geometries in space, has recently been developed for the integral operators considered. This is discussed elsewhere [6], [7].

Applying (23) and the Ritz procedure, the system of equations for solving σ is

$$\sum_{j=1}^{n}\sigma_j\int_{\Delta_k}\int_{\Delta_j} G(s\,|\,s')\,ds'\,ds - \int_{\Delta_k} g\,ds = 0, \qquad k = 1, 2, \cdots, m \tag{33}$$

$$\frac{\varepsilon_r+1}{2\varepsilon_0}\sigma_k\Delta_k + \frac{\varepsilon_r-1}{\varepsilon_0}\sum_{j=1}^{n}\sigma_j\int_{\Delta_k}\int_{\Delta_j} K(s\,|\,s')\,ds'\,ds = 0, \qquad k = m+1, \cdots, n \tag{34}$$

where Δ_k denotes the kth subregion, m = the number of subregions on the plate, and n = total number of subregions ($n > m$). Note that equations (33) are identical to equations (55a) of [1] neglecting for the moment the term involving the singular function f_s, as this is not our concern here. This is to be expected as the operator is self-adjoint here and the functional (20) reduces identically to functional (8). Equation (34), however, takes a much simpler form than the corresponding equation (55c) in [1]. It should be noted that the above equations could have been obtained from Galerkin's methods directly, using unit pulse weighting functions. It is also to be remarked that although the physical energy is minimized, for this case, one is unfortunately unable to show that the functional is minimized due to the introduction of the augmented operator $\mathscr{K}$ as mentioned in Section II.

Equations (33) and (34) are solved simultaneously for the unknown charge distribution σ on the conductor and the dielectric surfaces. Once σ is known, the potential ϕ, at any point in space, can be evaluated from the integral

$$\phi(s) = \frac{1}{\varepsilon_0}\int_S G(s\,|\,s')\sigma(s')\,ds'. \tag{35}$$

IV. Application to Wave Propagation in a Lossy Medium

For time-harmonic wave propagation problem, the Helmholtz equation

$$-(\nabla^2 + k^2)\phi = f \tag{36}$$

holds. Equation (36) in operator notation is

$$L\phi = f \tag{37}$$

where

$$L = -\nabla^2 - k^2$$

and

$$k^2 = \omega^2 \mu\varepsilon. \tag{38}$$

If the dielectric medium is lossy, the permittivity ε is a complex quantity. In such cases, from (6) and (7), it can be shown that the operator L is non-self-adjoint.

From the functional (20) for non-self-adjoint operators, we have

$$F = \langle L\phi,\phi^a\rangle - \langle f,\phi^a\rangle - \langle \phi,f^a\rangle. \tag{39}$$

Assuming homogeneous Neumann or Dirichlet boundary conditions for ϕ and ϕ^a, the explicit form of the functional (39) is

$$F = \iint \nabla\phi \cdot \nabla\phi^{a*}\, d\Omega - \iint k^2\phi\phi^{a*}\, d\Omega - \iint \phi^{a*} f\, d\Omega - \iint f^{a*}\phi\, d\Omega. \tag{40}$$

Letting

$$\phi = \boldsymbol{\phi}^T\boldsymbol{\alpha} = \boldsymbol{\alpha}^T\boldsymbol{\phi}$$

and

$$\phi^a = \boldsymbol{\phi}^{aT}\boldsymbol{\alpha} = \boldsymbol{\alpha}^T\boldsymbol{\phi}^a \tag{41}$$

and substituting into (40), we have

$$F = \boldsymbol{\phi}^T \iint \nabla\boldsymbol{\alpha} \cdot \nabla\boldsymbol{\alpha}^T\, d\Omega\boldsymbol{\phi}^{a*} - \boldsymbol{\phi}^T \iint k^2\boldsymbol{\alpha}\boldsymbol{\alpha}^T\, d\Omega\boldsymbol{\phi}^{a*} - \iint \boldsymbol{\alpha}^T f\, d\Omega\boldsymbol{\phi}^{a*} - \iint f^{a*}\boldsymbol{\alpha}^T\, d\Omega\boldsymbol{\phi} \tag{42}$$

or, replacing the integrations by matrices,

$$F = \boldsymbol{\phi}^T R\boldsymbol{\phi}^{a*} - \boldsymbol{b}^T\boldsymbol{\phi}^{a*} - \boldsymbol{b}^{aT}\boldsymbol{\phi} \tag{43}$$

where

$$R = S - k^2 T = \iint (\nabla\boldsymbol{\alpha} \cdot \nabla\boldsymbol{\alpha}^T - k^2\boldsymbol{\alpha}\boldsymbol{\alpha}^T)\, d\Omega \tag{44}$$

$$\boldsymbol{b} = \iint \boldsymbol{\alpha} f\, d\Omega \tag{45}$$

and

$$\boldsymbol{b}^a = \iint f^{a*}\boldsymbol{\alpha}\, d\Omega. \tag{46}$$

Taking derivatives with respect to each of the variational parameters $\boldsymbol{\phi}^a$ in turn, we get

$$R\boldsymbol{\phi} = \boldsymbol{b}. \tag{54}$$

Note that for the non-self-adjoint complex operator case shown here, the system of equations takes the same form as those for the real self-adjoint case. The generation of the matrix R, can, therefore, be obtained as before, except with the use of complex arithmetic.

We have assumed homogeneous Neumann or Dirichlet boundary conditions in this study. The extension to the mixed and inhomogeneous boundary conditions is straightforward and follows the approach outlined by Mikhlin [2, pp. 116–121]. It is a fairly straightforward matter to extend the functional to continuously inhomogeneous and orthotropic media as outlined in [4] or, preferably, [5]. Again, it is only required that complex arithmetic be employed. The formulation is otherwise unchanged.

V. Conclusion

A generalized functional was derived from the quadratic functional for self-adjoint operators and was shown to be applicable to integral equations with nonsymmetric kernels and for non-self-adjoint partial differential operators.

For the interface problem, using the integral equation formulation, comparing (34) with (55c) of [1], it is apparent that with the present formulation the amount of computation involved is reduced. Moreover, as pointed out in the Introduction, the solution process is not based upon the argument that $\langle K\sigma,G\rangle = 0$ implies $K\sigma = 0$.

For the non-self-adjoint Helmholtz problem, it is seen that the discretization is obtained in the same fashion as required for the real self-adjoint case except for the use of complex arithmetic.

At present, finite-element algorithms for the solution of Fredholm integral equations of the first and second kind are being developed (e.g., [6] and [7]). The algorithms permit problems involving arbitrary curved surfaces in space to be handled by an extension of the isoparametric finite-element approach.

Acknowledgment

The authors wish to acknowledge Prof. H. Lakser and other colleagues at the University of Manitoba and one of the reviewers for some very useful discussions and critical comments. Dr. G. I. Costache is gratefully acknowledged for contributions that led to the results of Section IV.

References

[1] B. H. McDonald, M. Friedman, and A. Wexler, "Variational solution of integral equations," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-22, pp. 237–238, Mar. 1974.

[2] S. G. Mikhlin, *Variational Methods in Mathematical Physics*. New York: Macmillan, 1974.

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

[4] A Wexler, "Finite-element analysis of an inhomogeneous, anisotropic, reluctance machine," *IEEE Trans. Power App. Syst.*, vol. PAS-92, pp. 145–149, Jan./Feb. 1973.

[5] ——, "Isoparametric finite elements for continuously inhomogeneous and orthotropic media," in *Proc Int. Conf. Finite Elements for Water Resources*, (July 12–16, 1976), W. G. Gray, G. F. Pinder, and C. A. Brebbia, Eds. London: Pentech Press, 1977, pp. 2.3–2.24.

[6] G. Jeng and A. Wexler, "Finite-element solution of boundary integral equations," in *Proc. Int. Symp. Large Engineering Systems*, (August 9–12, 1976), A. Wexler, Ed. New York: Pergamon Press, 1977, pp. 112–121.

[7] ——, "Isoparametric, finite-element, variational solution of integral equations for three-dimensional fields," *Int. J. Num. Meth. Eng.*, vol. 11, pp. 1455–1471, 1977.