JOURNAL OF COMPUTATIONAL PHYSICS 82, 172–192 (1989)

# Exact Non-reflecting Boundary Conditions

JOSEPH B. KELLER

*Departments of Mathematics and Mechanical Engineering, Stanford University, Stanford, California 94305*

AND

DAN GIVOLI

*Department of Mechanical Engineering, Stanford University, Stanford, California 94305*

Received December 4, 1987; revised June 29, 1988

An exact non-reflecting boundary condition is devised for use in solving the reduced wave equation in an infinite domain. The domain is made finite by the introduction of an artificial boundary on which this exact condition is imposed. In the finite domain a finite element method is employed. Although the boundary condition is non-local, that does not affect the efficiency of the computational scheme. Numerical examples are presented which show that the use of the exact non-local boundary condition yields results which are much more accurate than those obtained with various approximate local conditions. The method can also be used to solve problems in large finite domains by reducing them to smaller domains, and it can be adapted to other differential equations. © 1989 Academic Press, Inc.

## 1. INTRODUCTION

To solve the reduced wave equation numerically in an unbounded domain, it is usual to introduce an artificial boundary $\mathscr{B}$ to make the computational domain finite. Then some boundary condition must be imposed on $\mathscr{B}$. We shall show how to obtain an exact boundary condition there and how to combine it with the finite element method in the computational domain. This leads to an efficient method which eliminates the defects, such as spurious reflections from $\mathscr{B}$, which arise in other procedures.

We shall also show how to solve a problem in a large finite domain by cutting out a large regular subdomain, thus introducing an artificial boundary $\mathscr{B}$. This leaves a smaller computational domain bounded in part by $\mathscr{B}$. On $\mathscr{B}$ we shall again obtain an exact boundary condition and use it with the finite element method in the computational domain.

Naturally there has been a great deal of previous work on these subjects. In order to see how the present work is related to it and to what extent the present work



0021-9991/89 $3.00

is novel, we shall describe some of the previous methods. The simplest and most usual boundary condition is

$$u_\nu(\mathbf{x}) = iku(\mathbf{x}), \qquad \mathbf{x} \text{ on } \mathscr{B}. \tag{1}$$

Here $u(\mathbf{x})$ is the unknown scattered field, $u_\nu$ is its outward normal derivative, and $k$ is the wave number. This condition is of the same form as the Sommerfeld radiation condition, which is exactly correct when imposed at infinity but only approximately correct when imposed at a finite boundary $\mathscr{B}$. As a consequence the use of (1) leads to the spurious reflection of waves from $\mathscr{B}$. The magnitude of the reflected wave increases the more the direction of the wave $u$ deviates from the direction normal to $\mathscr{B}$, so it increases as $\mathscr{B}$ is moved closer to the scattering region.

In order to diminish the spurious reflection, various authors have devised improved local boundary conditions. Engquist and Majda [1] did so by expressing $u_\nu$ exactly as a pseudodifferential operator applied to $u$ on $\mathscr{B}$, and then approximating this operator by the local differential operators given in (65) and (66) below. Bayliss and Turkel [2] used the asymptotic expansion of $u$ far from the scatterer to obtain similar approximate local boundary conditions. Feng [3] obtained an exact non-local condition involving an integral over $\mathscr{B}$ of $u$ multiplied by a Green's function, and then he approximated it by various local conditions, such as (67) below. However, all these local boundary conditions still lead to spurious reflection.

Gustafsson and Kreiss [4] considered a hyperbolic system of equations for two functions of $x$, $y$, and $t$ in a waveguide. They obtained an exact non-local boundary condition involving the Fourier coefficients of the solution and discussed its use in a finite difference method. Hagstrom and Keller [5] formed an exact boundary condition in a cylindrical domain in terms of the eigenfunctions and eigenvalues of a problem in the cross section of the cylinder. They also proved the existence of an exact boundary condition for certain nonlinear problems and gave an asymptotic expansion for it. Ting and Miksis [6] obtained an exact boundary condition by expressing $u$ on $\mathscr{B}$ as an integral of $u$ and $u_\nu$ over a surface interior to $\mathscr{B}$, using the free space Green's function.

The work most closely related to ours is contained in Fix and Marin [7], MacCamy and Marin [8], Marin [9, 10], and Goldstein [11]. In all of them an exact boundary condition is imposed on $\mathscr{B}$ and the finite element method is employed in the computational domain. In [7] this boundary condition is found analytically for a waveguide and some numerical results are given. In [8] 2-dimensional exterior domains are considered. The boundary condition involves the solution of an integral equation on $\mathscr{B}$ for which numerical methods of solution are given. The convergence of the finite element method with this boundary condition is proved. A summary of these results and a numerical example are presented in [9]. Goldstein [11] presents a survey of previous work and gives a convergence proof for the finite element method with an exact boundary condition on $\mathscr{B}$ in the case of 2-dimensional exterior problems and waveguides.

In contrast with the method of MacCamy and Marin [8] and Marin [9], we

choose $\mathscr{B}$ to be a circle in 2-dimensional exterior problems and a sphere in 3-dimensional ones. As a consequence we can express the exact boundary condition explicitly in terms of known functions rather than in terms of the solution of an integral equation which is obtained numerically. This results in a simpler and more accurate method, which is similar to that used by Fix and Marin [7] for waveguides. It enables us to discuss the bandedness of the finite element stiffness matrix, the separability of the integrals in the boundary condition, and the best choice of the radius of $\mathscr{B}$ in terms of the computational accuracy and cost.

We shall present the solutions of some test problems obtained by our method. We shall compare them with exact solutions and with numerical solutions involving the Sommerfeld condition (1), the Engquist–Majda conditions (65) and (66) and the Feng condition (67).

We shall also obtain the exact boundary condition explicitly for Laplace's equation in both exterior and interior domains. Then we shall present a simple convergence proof of the finite element method with this boundary condition.

At first sight it seems that the non-locality of the exact boundary condition might spoil the banded structure of the finite element matrix, and the complexity of this condition might require a great deal of computation. However, neither of these difficulties occurs. In fact our results using the non-local conditions are more accurate than those obtained by using approximate local conditions, and they require about the same computational work.

In Section 2 we present the finite element formulation for the reduced wave equation in an infinite domain. The exact boundary condition is formulated in Section 3, where we also show how the method can be applied to large finite domains. The convergence of the scheme and some computational aspects are discussed in Sections 4 and 5. In Section 6 we present some numerical results and compare them to those obtained by using the approximate local boundary conditions proposed by other authors.

## 2. Finite Element Formulation

We consider the following problem for a solution $u$ of the inhomogeneous reduced wave equation in an infinite domain $\mathscr{R}$ bounded internally by the surface $\Gamma = \Gamma_g \cup \Gamma_h$ of an obstacle:

$$\nabla^2 u + k^2 u + f = 0 \qquad \text{in } \mathscr{R} \tag{2}$$

$$u = g \qquad \text{on } \Gamma_g \tag{3}$$

$$u_\nu = h \qquad \text{on } \Gamma_h \tag{4}$$

$$\lim_{r \to \infty} r^{(d-1)/2}(u_r - iku) = 0. \tag{5}$$

Here $k$, $f$, $g$, and $h$ are given functions, $d$ is the spatial dimension, and (5) is the

Sommerfeld radiation condition. We assume that outside a sphere (or circle) $\mathscr{B}$ of radius $R$, $f=0$ and $k$ is constant.

In order to solve the problem numerically we define a computational domain $\Omega$ which is bounded internally by $\Gamma$ and externally by the artificial boundary $\mathscr{B}$. (See Fig. 1.) The domain outside $\mathscr{B}$, $\mathscr{R}-\Omega$, we denote $D$.

On $\mathscr{B}$ we impose the boundary condition $u_v=-Mu$. Here $M$ is an operator called the *Dirichlet to Neumann* (*DtN*) map, because it relates the Dirichlet datum $u$ to the Neumann datum $u_v$ on $\mathscr{B}$. This is an *exact* boundary condition and will be discussed in detail in the next section. We then replace (2)–(5) by the equivalent problem

$$\nabla^2 u + k^2 u + f = 0 \qquad \text{in } \Omega \tag{6}$$

$$u = g \qquad \text{on } \Gamma_g \tag{7}$$

$$u_v = h \qquad \text{on } \Gamma_h \tag{8}$$

$$u_v = -Mu \qquad \text{on } \mathscr{B}. \tag{9}$$

To obtain a weak form of this problem we introduce the two function spaces

$$\mathscr{S} = \{u \mid u \in H^1(\Omega) \text{ and } u = g \text{ on } \Gamma_g\} \tag{10}$$

and

$$\mathscr{S}_0 = \{w \mid w \in H^1(\Omega) \text{ and } w = 0 \text{ on } \Gamma_g\}, \tag{11}$$

where $H^1$ is the Sobolev space defined by

$$H^1(\Omega) = \left\{\psi(\mathbf{x}) \mid \psi \in L_2(\Omega), \frac{\partial \psi}{\partial x_i} \in L_2(\Omega)\right\}. \tag{12}$$

Then (6)–(9) is equivalent to the following variational problem:

Find $u \in \mathscr{S}$ such that for all $w \in \mathscr{S}_0$,

$$a(w, u) + b(w, u) = (w, f) + (w, h)_\Gamma, \tag{13}$$

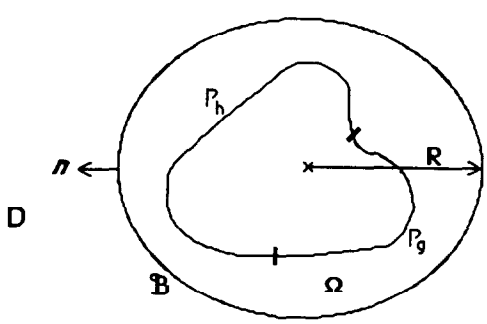


FIG. 1. Geometry of a typical reduced wave equation problem.

where

$$a(w, u) = \int_{\Omega} (\nabla w \cdot \nabla u - k^2 wu)\, d\Omega \tag{14}$$

$$b(w, u) = \int_{\mathscr{B}} wMu\, d\mathscr{B} \tag{15}$$

$$(w, f) = \int_{\Omega} wf\, d\Omega \tag{16}$$

$$(w, h)_{\Gamma} = \int_{\Gamma_h} wh\, d\Gamma. \tag{17}$$

In order to solve this variational problem by the Galerkin method, we introduce finite dimensional subspaces of $\mathscr{S}$ and $\mathscr{S}_0$, denoted respectively $\mathscr{S}^h$ and $\mathscr{S}_0^h$. Then we consider the following approximation to (13):

Find $u^h \in \mathscr{S}^h$ such that for all $w^h \in \mathscr{S}_0^h$,

$$a(w^h, u^h) + b(w^h, u^h) = (w^h, f) + (w^h, h)_{\Gamma}. \tag{18}$$

To construct the finite element spaces $\mathscr{S}^h$ and $\mathscr{S}_0^h$, we discretize $\Omega$ into a finite number of element domains. Each element is associated with a finite number of nodes on its boundaries and possibly in its interior. We define $\eta_g$ to be the set of all nodes on the boundary $\Gamma_g$ where $u$ is prescribed, and $\eta$ to be the set of all the other nodes, at which the value of $u$ is unknown. (See Fig. 2.) We introduce a shape (or basis) function $N_A$ for each node $A \in \eta$, and a shape function $N_A^g$ for each node $A \in \eta_g$. Then the functions $u^h$ and $w^h$ are assumed to have the form

$$u^h(\mathbf{x}) = \sum_{A \in \eta} d_A N_A(\mathbf{x}) + \sum_{A \in \eta_g} g_A N_A^g(\mathbf{x}); \qquad g_A = g(\mathbf{x}_A) \text{ for } A \in \eta_g \tag{19}$$

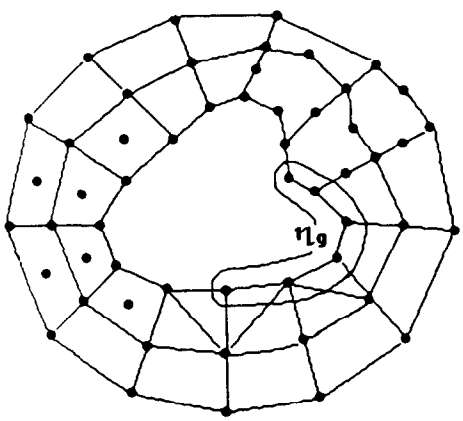


FIG. 2. Typical discretization of the computational domain $\Omega$. Also shown is the set $\eta_g$ of all nodes on $\Gamma_g$.

and

$$w^h(\mathbf{x}) = \sum_{A \in \eta} c_A N_A(\mathbf{x}). \tag{20}$$

Here $d_A$ and $c_A$ are constants and $\mathbf{x}_A$ is the physical point at node number $A$.

We now substitute (19) and (20) into (18). Since $w^h$ and therefore $c_A$ are arbitrary, the expression multiplying $c_A$ must vanish for each $A$. This results in the finite element matrix formulation of the problem:

$$\mathbf{Kd} = \mathbf{F} \tag{21}$$

$$\mathbf{K} = \mathbf{K}^a + \mathbf{K}^b \tag{22}$$

$$\mathbf{K}^a = [K^a_{AB}], \qquad \mathbf{K}^b = [K^b_{AB}], \qquad \mathbf{d} = \{d_B\}, \qquad \mathbf{F} = \{F_A\} \tag{23}$$

$$K^a_{AB} = a(N_A, N_B), \qquad K^b_{AB} = b(N_A, N_B) \tag{24}$$

$$F_A = (N_A, f) + (N_A, h)_\Gamma - \sum_{B \in \eta_g} g_B[a(N_A, N^g_B) + b(N_A, N^g_B)]. \tag{25}$$

Here $A$ and $B$ on the left-hand side of (24) and (25) are the positions in the matrix or vector corresponding to nodes $A$ and $B$. The number of equations in the linear algebraic system (21) is equal to the number of nodes in $\eta$. Its solution $\mathbf{d}$ determines the finite element solution $u^h$ via (19).

The finite element method requires that the shape functions $N_A$ and $N^g_A$ be continuous and of *local support*. It is standard to define $N_A$ and $N^g_A$ to have the value 1 at node $A$, to equal zero at every other node, and to vanish outside a local patch of elements which share node $A$. We note that the last term in (25), $b(N_A, N^g_B)$, vanishes because each $N^g_B$ is identically zero on the boundary $\mathscr{B}$. Therefore (25) becomes

$$F_A = (N_A, f) + (N_A, h)_\Gamma - \sum_{B \in \eta_g} g_B\, a(N_A, N^g_B). \tag{26}$$

This equation and (22) show that the effect of the DtN boundary condition on the standard finite element scheme is just the inclusion of $\mathbf{K}^b$ in the matrix $\mathbf{K}$ in (21). Furthermore, $K^b_{AB}$ is nonzero only if *both* nodes $A$ and $B$ lie on the boundary $\mathscr{B}$.

Usually the locality of the shape functions makes the matrix $\mathbf{K}$ *banded*, because there is no interaction between two nodes not belonging to the same element. Here, however, $K^b_{AB}$ is nonzero for *any* pair of nodes $A$ and $B$ on $\mathscr{B}$, due to the non-local character of $b(N_A, N_B)$ in (15). This seems to spoil the bandedness of $\mathbf{K}$, but in fact it does not. This is because finite element schemes usually employ the "skyline" method, in which only the portion of $\mathbf{K}$ from the diagonal to the upper "skyline" of the matrix (i.e., up to the last nonzero entry in each column) is stored and operated on. The best simple method of numbering the nodes is that illustrated in Fig. 3. It consists in numbering circumferentially, and shifting one node "back-

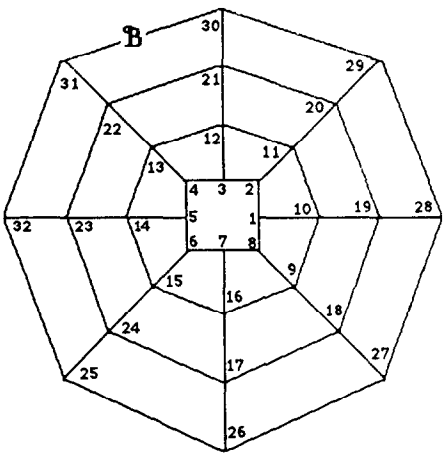


FIG. 3. Optimal node numbering of a typical mesh.

wards" when advancing outward from one line to the next. With this numbering, the pairs $(A, B)$ of nodes on the boundary $\mathscr{B}$ are all *below* the "skyline" of the matrix.

## 3. The DtN Boundary Condition

The way to find the DtN boundary condition on $\mathscr{B}$ is to solve the Dirichlet problem in the external domain $D$. We choose $\mathscr{B}$ to be a circle or sphere of radius $R$ and then this problem is

$$\nabla^2 u + k^2 u = 0 \qquad \text{in } D \tag{27}$$

$$u = u(R, \theta) \qquad \text{on } \mathscr{B} \tag{28}$$

$$\lim_{r \to \infty} r^{(d-1)/2}(u_r - iku) = 0. \tag{29}$$

In two dimensions the solution of (27)–(29) is

$$u(r, \theta) = \frac{1}{\pi} \sum_{n=0}^{\infty}{}' \int_0^{2\pi} \frac{H_n^{(1)}(kr)}{H_n^{(1)}(kR)} \cos n(\theta - \theta')\, u(R, \theta')\, d\theta'. \tag{30}$$

Here $H_n^{(1)}$ is the Hankel function of the first kind. The prime after the sum indicates that a factor of $\frac{1}{2}$ multiples the term with $n = 0$. We now differentiate (30) with respect to $r$ and set $r = R$ to obtain

$$u_v(R, \theta) = -\sum_{n=0}^{\infty}{}' \int_0^{2\pi} m_n(\theta - \theta')\, u(R, \theta')\, d\theta' \tag{31}$$

$$m_n(\theta - \theta') = -\frac{k}{\pi} \frac{H_n^{(1)\prime}(kR)}{H_n^{(1)}(kR)} \cos n(\theta - \theta'). \tag{32}$$

In (32) the $m_n(\theta-\theta')$ are the DtN kernels, the symmetry of which is responsible for the symmetry of $\mathbf{K}^b$ in (22). When $u(R,\theta)$ in (30) is replaced by the finite element solution $u^h(\mathbf{x})$ evaluated at $\mathbf{x}$ on $\mathscr{B}$, (30) yields the approximate value of $u(r,\theta)$ at points outside $\mathscr{B}$.

In three dimensions the solution of (27)–(29) is

$$u(r,\theta,\phi)=\frac{\sqrt{R/r}}{2\pi R^2}\sum_{n=0}^{\infty}{\sum_{j=0}^{n}}'\frac{H^{(1)}_{n+1/2}(kr)}{H^{(1)}_{n+1/2}(kR)}\frac{(2n+1)(n-j)!}{(n+j)!}$$

$$\times\int_{\mathscr{B}}P_n^j(\cos\phi)\,P_n^j(\cos\phi')\cos j(\theta-\theta')\,u(R,\theta',\phi')\,d\mathscr{B}'. \tag{33}$$

Here $d\mathscr{B}'=R^2\sin\phi'\,d\theta'\,d\phi'$, and $P_n^j$ is the associated Legendre function of the first kind. From (33) we obtain the DtN boundary condition

$$u_{,\nu}(R,\theta,\phi)=-\sum_{n=0}^{\infty}\int_{\mathscr{B}}m_n(\theta,\phi,\theta',\phi')\,u(R,\theta',\phi')\,d\mathscr{B}', \tag{34}$$

where the DtN kernel is

$$m_n(\theta,\phi,\theta',\phi')={\sum_{j=0}^{n}}'\beta_{jn}P_n^j(\cos\phi)\,P_n^j(\cos\phi')\cos j(\theta-\theta') \tag{35}$$

$$\beta_{jn}=-\frac{(2n+1)(n-j)!\;\gamma_n}{2\pi R^2(n+j)!} \tag{36}$$

$$\gamma_n=\frac{(\partial/\partial R)[R^{-1/2}H^{(1)}_{n+1/2}(kR)]}{R^{-1/2}H^{(1)}_{n+1/2}(kR)}. \tag{37}$$

We shall now show how to use the method we have presented to solve problems with large finite domains. To do so we consider a finite domain $\mathscr{R}$ with an irregularly shaped boundary $\partial\mathscr{R}$. We suppose that $f$ is zero and that $k$ is constant, except in a region near $\partial\mathscr{R}$. To obtain a small computational domain $\Omega$ we cut out of $\mathscr{R}$ a spherical or circular domain $D$ within which $f=0$ and $k$ is constant. In $D$ the solution $u$ can be obtained analytically in terms of its values on $\mathscr{B}=\partial D$. Thus the computational domain $\Omega=\mathscr{R}-D$ is bounded internally by $\mathscr{B}$ and externally by $\partial R$, so the bigger $D$ is the smaller $\Omega$ is.

The finite element formulation is exactly the same as that given before, but the DtN map is different. To obtain it we solve the problem (27) and (28). Then we find that (30)–(37) hold with the Hankel function $H_n^{(1)}$ replaced everywhere by the Bessel function $J_n$. Thus (32) becomes

$$m_n(\theta-\theta')=-\frac{k}{\pi}\frac{J_n'(kR)}{J_n(kR)}\cos n(\theta-\theta') \tag{38}$$

and (37) becomes

$$\gamma_n = \frac{(\partial/\partial R)[R^{-1/2} J_{n+1/2}(kR)]}{R^{-1/2} J_{n+1/2}(kR)}. \tag{39}$$

Since this DtN boundary condition is real, the whole problem is real.

All the preceding considerations apply to the Laplace or Poisson equation to which (2) reduces when $k = 0$. Then (5) is replaced by the simpler condition that $u$ tends to zero at infinity. By letting $k$ tend to zero in (30)–(39) we can obtain the corresponding solutions and DtN boundary conditions for Laplace's equation in $D$. Alternatively we can solve Laplace's equation directly. In either way we find that outside a circle of radius $R$, $u$ is given by

$$u(r, \theta) = \frac{1}{\pi} \sum_{n=1}^{\infty} (R/r)^n \int_0^{2\pi} \cos n(\theta - \theta')\, u(R, \theta')\, d\theta'. \tag{40}$$

Then we find that the boundary condition (31) holds with $m_n(\theta - \theta')$ given by

$$m_n(\theta - \theta') = \frac{n}{\pi R} \cos n(\theta - \theta'). \tag{41}$$

Exactly the slame expression (41) applies when $D$ is the interior of the circle. An alternative closed form solution for $u_\nu$ in terms of $u$ can be given, but it is less convenient computationally.

Suppose that $k = 0$ and that the Neumann condition is given on all of $\Gamma$, so that $\Gamma = \Gamma_h$. Then the solution of the original problem (2)–(4) does not vanish at infinity unless $f$ and $h$ satisfy the condition

$$\int_\Omega f\, d\Omega + \int_{\Gamma_h} h\, d\Gamma = 0. \tag{42}$$

When this condition holds, the problem (6)–(9) has a solution but it is unique only up to an additive constant. This solution can be made unique by requiring that

$$\int_{\mathscr{B}} u(R, \theta)\, d\theta = 0. \tag{43}$$

Instead of imposing (43) we can include the constant $m_0(\theta - \theta') = 1/\pi R$ in (41) to get the modified DtN boundary condition

$$u_\nu(R, \theta) = -\frac{1}{\pi R} \int_0^{2\pi} u(R, \theta')\, d\theta' - \sum_{n=1}^{\infty} \frac{n}{\pi R} \int_0^{2\pi} \cos n(\theta - \theta')\, u(R, \theta')\, d\theta'. \tag{44}$$

The additional term will not contribute anything to $u_\nu$ if (43) is satisfied, but it will render the solution of (6)–(9) unique. Furthermore, it is easy to show that this unique solution satisfies (43).

## 4. Convergence

We shall now examine the convergence of the numerical procedure for Laplace's equation, which is the special case of the reduced wave equation with $k=0$. The proof in this case is much simpler than those for $k\neq 0$ given by MacCamy and Marin [8] and Goldstein [11].

For $k=0$ *with a local condition instead of the DtN boundary condition* the proof is standard and can be found in Strang and Fix [12]. It is composed of three ingredients. The first is the *best approximation property*: the finite element solution $u^h$ is the best among all members of $\mathcal{S}^h$ in that it gives minimal error in the "energy norm" $a(v, v)^{1/2}$. This can be stated as

$$a(u-u^h, u-u^h) \leqslant a(u-\tilde{u}, u-\tilde{u}), \qquad \forall \tilde{u} \in \mathcal{S}^h. \tag{45}$$

The second ingredient is the following approximation theorem: If $u \in H^{p+1}$ then $\exists u_I \in \mathcal{S}^h$ such that

$$\|u-u_I\|_1 \leqslant \hat{c} h^p, \tag{46}$$

where $p$ is the highest degree of complete polynomial in $\mathcal{S}^h$, $\hat{c}$ is a constant depending only on $u$, $h$ is the mesh parameter, and $\|\cdot\|_1$ denotes the norm in $H^1$. The last ingredient is the equivalence of the norm $\|\cdot\|_1$ with the "energy norm" $a(\cdot, \cdot)^{1/2}$:

$$c_1 \|v\|_1 \leqslant a(v, v)^{1/2} \leqslant c_2 \|v\|_1 \qquad \forall v \in \mathcal{S}_0. \tag{47}$$

Combining (45), (46), and (47) yields

$$\|u-u^h\|_1 \leqslant c h^p \tag{48}$$

which is the standard error estimate.

When the DtN operator $b(\cdot, \cdot)$ is included in the left side of (13) the "energy norm" is no longer $a(v, v)^{1/2}$, but rather $[a(v, v)+b(v, v)]^{1/2}$. The best approximation property becomes

$$\begin{aligned} &a(u-u^h, u-u^h)+b(u-u^h, u-u^h) \\ &\leqslant a(u-\tilde{u}, u-\tilde{u})+b(u-\tilde{u}, u-\tilde{u}) \qquad \forall \tilde{u} \in \mathcal{S}^h \end{aligned} \tag{49}$$

and to recover the usual error estimate (48) it remains to show that

$$c_1 \|v\|_1 \leqslant [a(v, v)+b(v, v)]^{1/2} \leqslant c_2 \|v\|_1 \qquad \forall v \in \mathcal{S}_0. \tag{50}$$

Now, since (47) is known to hold, the proof will be complete if we show that

$$0 \leqslant b(v, v) \leqslant \|v\|_1^2 \qquad \forall v \in \mathcal{S}_0. \tag{51}$$

We shall suppose first that $D$ is infinite. The left inequality of (51), namely the positivity of $b(v, v)$, can be shown directly from the form of the DtN kernels (41)

or in the following more general way. Any $v$ defined on $\mathscr{B}$ can be extended to all of $D$ such that it satisfies $\nabla^2 v = 0$ there and vanishes at infinity. Thus, the DtN boundary condition is satisfied on $\mathscr{B}$ and we have

$$b(v, v) = \int_{\mathscr{B}} vMv \, dB = -\int_{\mathscr{B}} vv_\nu \, d\mathscr{B}$$

$$= \int_D \nabla \cdot (v \nabla v) \, d\mathbf{x}$$

$$= \int_D \nabla v \cdot \nabla v \, d\mathbf{x} \geqslant 0. \tag{52}$$

To prove the right-hand inequality in (51) we assume that $\Gamma = \Gamma_g$. Then we denote by $w_0$ the function that minimizes the Dirichlet integral over $\Omega$ among all functions in $\mathscr{S}_0$ which are equal to $v$ on $\mathscr{B}$. Thus

$$\|v\|_1^2 = \|v\|_0^2 + \|\nabla v\|_0^2 \geqslant \|\nabla v\|_0^2 \geqslant \|\nabla w_0\|_0^2. \tag{53}$$

Here $\|\cdot\|_0$ denotes the norm in $L_2$. Now we extend $w_0$ to be identically zero inside [illegible] integral over $Q$ among all functions in $\mathscr{S}_1$ which are equal to $v$ on $\mathscr{B}$. The function $w_0$, extended to be identically zero in the region enclosed by $\Gamma$, is one of these functions, and its Dirichlet integrals over $Q$ and $\Omega$ are equal. Thus $\|\nabla w_0\|_0^2 \geqslant \|\nabla W\|_0^2$. Upon combining this inequality with (53) we get $\|v\|_1^2 \geqslant \|\nabla W\|_0^2$. But $W$ is the solution of $\nabla^2 W = 0$ in $Q$ with $W = v$ on $\mathscr{B}$. Then by proceeding as in (52) and noting that the kernels $m_n(\theta - \theta')$ are the slame for the interior of $\mathscr{B}$ and for the exterior, we get

$$\|v\|_1^2 \geqslant \|\nabla W\|_0^2 = \int_{\mathscr{B}} vMv \, d\mathscr{B} = b(v, v). \tag{54}$$

This is the desired inequality (51).

When $D$ is finite the proof still applies, with $Q$ denoting the domain exterior to $\mathscr{B}$.

Thus the DtN finite element formulation for Laplace's equation converges with the standard finite element rate of convergence. When the sum in (31) is truncated after $N$ terms, there is an additional error which decreases to zero as $N$ tends to infinity.

## 5. Computational Aspects

We have already observed that the DtN boundary condition does not interfere with the symmetry and bandedness of the finite element matrix. Now we are going

to derive explicit expressions for the DtN term $\mathbf{K}^b$. To do so we note first that the DtN kernels $m_n(\mathbf{x}, \mathbf{x}')$ are separable. For example, in the 2-dimensional case we have from (32)

$$m_n(\theta - \theta') = -\frac{k}{\pi}\frac{H_n^{(1)\prime}(kR)}{H_n^{(1)}(kR)}(\cos n\theta \cos n\theta' + \sin n\theta \sin n\theta'). \tag{55}$$

For all the cases considered here (i.e., the reduced wave equation and Laplace's equation in two and three dimensions) we can write

$$Mu(\mathbf{x}) = \sum_{j=0}^{\infty} a_j F_j(\mathbf{x}) \int_{\mathscr{B}} F_j(\mathbf{x}')\, u(\mathbf{x}')\, d\mathbf{x}', \qquad \mathbf{x} \text{ on } \mathscr{B}. \tag{56}$$

Now from (24) and (15) we have

$$\begin{aligned} K_{AB}^b = b(N_A, N_B) &= \int_{\mathscr{B}} N_A M N_B \, d\mathscr{B} \\ &= \sum_{j=0}^{\infty} a_j \left( \int_{\mathscr{B}} N_A(\mathbf{x})\, F_j(\mathbf{x})\, d\mathscr{B} \right) \left( \int_{\mathscr{B}} N_B(\mathbf{x})\, F_j(\mathbf{x})\, d\mathscr{B} \right). \end{aligned} \tag{57}$$

Therefore, in a $d$-dimensional problem, only integrals of dimension $d-1$ have to be computed to form $\mathbf{K}^b$. Moreover the $F_j$ are simple trigonometric functions (or in the 3-dimensional case polynomials in trigonometric functions) which suggests evaluating the integrals explicitly. Altogether we have to evaluate $n_{\mathscr{B}}$ integrals for each $j$, where $n_{\mathscr{B}}$ is the number of nodes on $\mathscr{B}$. The integration associated with node $A$ has to be performed only over the sides of the elements connected to the node, because the shape function $N_A(\mathbf{x})$ is identically zero elsewhere.

The storage associated with the DtN term is therefore $O(n_{\mathscr{B}})$. The addition of this term to the finite element matrix must be performed on the "global level," not on the "element level," since assembling small element matrices into the global matrix is not possible here.

The coefficients $a_j$ in (56) are the expressions involving the Bessel functions in (32), (37), (38), and (39). *The cost of computing them is marginal* because there are typically only a few of them (depending on the number of harmonics that are taken into account), and they have to be evaluated only once for each choice of $kR$. In fact, they can be tabulated once for an interval of values of $kR$ and used as needed.

The main additional time-consuming step in using the DtN boundary condition is the evaluation of the integrals in (57) involving trigonometric expressions. This time can be reduced by evaluating the integrals explicitly, as was mentioned before. But in any case, *this computational effort is insignificant for a "large" problem*, where the dominant cost is that associated with the large number of elements in the mesh rather than with the nodes on part of the boundary.

We now consider how to choose the two parameters $R$, the radius of the boundary $\mathscr{B}$, and $N$, the number of terms to be used in the DtN map. When the DtN map

is given by a *closed-form* expression the choice of $R$ does not affect the accuracy, but in the more common case increasing $R$, as well as increasing $N$, will improve accuracy.

First of all, in the case where the original domain is finite and $k \neq 0$, certain values of the radius $R$ are not permitted. These are the values that make the denominator of one of the coefficients in (38) or (39) vanish. In fact to avoid large round-off errors one should not use values of $R$ which make this denominator very small. In other words, $R$ should be chosen such that $kR$ is not very close to a root of $J_n$ in two dimensions or to a root of $J_{n+1/2}$ in three dimensions. This consideration applies to the values of $n$ ranging from 0 to $N$. No such restriction exists in the case where the original domain is infinite or where $k = 0$.

For fixed $R$ and $h$ there is a certain $N$ which is optimal, in the sense that no additional accuracy can be gained by taking $N$ larger. This value $N_{\text{opt}}$ occurs when the finite element discretization error equals the series truncation error. In the case of Laplace's equation it is possible to estimate $N_{\text{opt}}$ very crudely in the following manner.

The rate of convergence of the finite element method in the $L_2$ norm can be shown to be $p+1$, which is higher by 1 than the $H_1$ convergence rate (see (48)). Empirically the pointwise rate of convergence is usually the same as that of the $L_2$ norm. Thus we write the finite element error in the form

$$e_{\text{f.e.}} = ch^{p+1}. \tag{58}$$

The truncation error due to neglecting terms after the $N$th term can be estimated by

$$e_t^N = \sum_{n=N+1}^{\infty} u^{(n)}(R, \theta) \simeq u^{(N)}(R, \theta), \tag{59}$$

where $u^{(n)}$ is the $n$th term in the expansion of $u$. Now $u^{(N)}$ can be written as $T_N R^{-N}$, where $T_N$ is a function of $\theta$ but does not depend upon $R$. Therefore we have

$$e_t^N = T_N R^{-N}. \tag{60}$$

Equating (58) and (60) and taking the logarithm of both sides results in

$$\log c + (p+1)\log h = \log T_N - N \log R. \tag{61}$$

Assuming that $c$ and $T_N$ are of order one with respect to $R$ and $h$, we can neglect $\log c$ and $\log T_N$ in comparison with the other terms, and obtain

$$N_{\text{opt}} = -(p+1)\frac{\log h}{\log R} \tag{62}$$

which is the desired estimate.

Numerical experiments show that formula (62) is a quite reasonable although crude estimate. For example, with bilinear quadrilateral elements ($p=1$), $R=2$, and uniform radial spacing of $h=0.2$, (62) yields 4.6 whereas numerical experiments yield $N_{\text{opt}}=3$. With $R=2.5$ and $h=0.75$, (62) gives 0.62 versus the experimental $N_{\text{opt}}=1$.

A reasonable order of choosing the computational parameters is as follows. First the boundary $\Gamma$ is discretized, according to the shape of $\Gamma$ and the accuracy desired of the solution. This determines $h$ and the circumferential discretization of the domain. Now there are two extreme approaches to choosing $R$ and $N$. The first is to take the smallest $R$ possible which bounds the "irregular" region (i.e., put $\mathscr{B}$ as close as possible to $\Gamma$ in Fig. 1). Substituting this $R$ into formula (62), yields the corresponding $N_{\text{opt}}$. The other extreme is to take $N=1$ in (62) and find the corresponding radius $R$. In the first approach the computational domain, and hence the finite element cost, is minimized, whereas in the second the DtN cost is minimized. An intermediate choice is probably preferable. $R$ and $h$ determine the shape of the whole mesh since all the elements should have aspect ratios close to one.

## 6. Numerical Experiments

A number of approximate local boundary conditions have been suggested for use in numerical schemes for solving the 2-dimensional problem. Those which are compatible with the finite element method have the form

$$-u_\nu = C_1 u + C_2 \frac{\partial^2 u}{\partial \theta^2} \qquad \text{on } \mathscr{B}. \tag{63}$$

They yield exactly the same finite element formulation as in Section 2 with

$$b(w,u) = C_1 \int_{\mathscr{B}} wu \, d\mathscr{B} - C_2 \int_{\mathscr{B}} \frac{\partial w}{\partial \theta}\frac{\partial u}{\partial \theta} \, d\mathscr{B}. \tag{64}$$

The simplest approximate boundary condition is the Sommerfeld radiation condition (5) applied on $\mathscr{B}$, which corresponds to $C_1=-ik$, $C_2=0$ in (63). To demonstrate its deficiency we compare the result of using it with that of using the DtN boundary condition in the problem of a point source in an infinite plane. Figure 4 shows the mesh, composed of linear triangles and bilinear quadrilateral elements. The source is located at a node which is three nodes from the boundary. We use $k=1$, $R=1$, and 4 terms in the DtN kernel. Figures 5 and 6 are contour plots of the imaginary part of the solution for the DtN and the Sommerfeld boundary conditions, respectively. The crudeness of the mesh is responsible for the unsmooth "circles" in the DtN case. The figures show that the Sommerfeld condition yields a severe spurious reflection from the artificial boundary but the DtN condition does not.

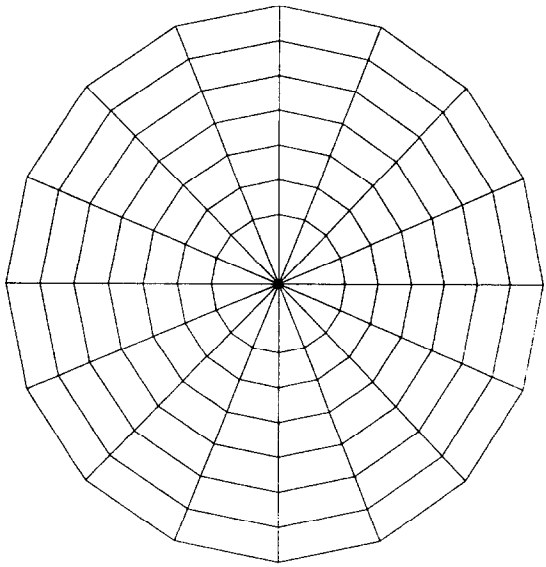

FIG. 4. Mesh for the point source problem.

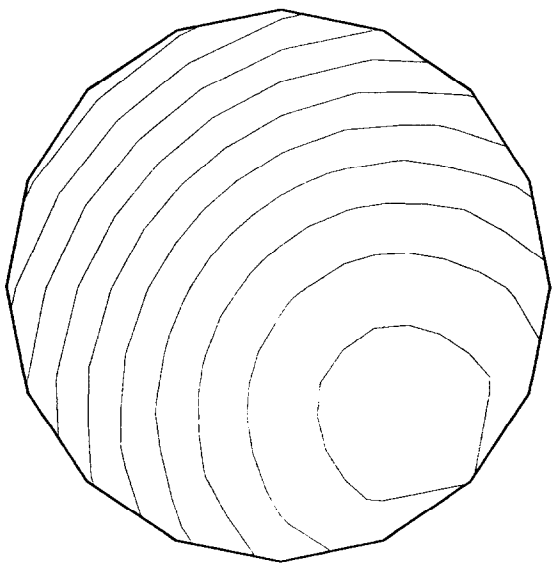

FIG. 5. The point source problem: contour plot of the imaginary part of the solution with the DtN boundary condition.

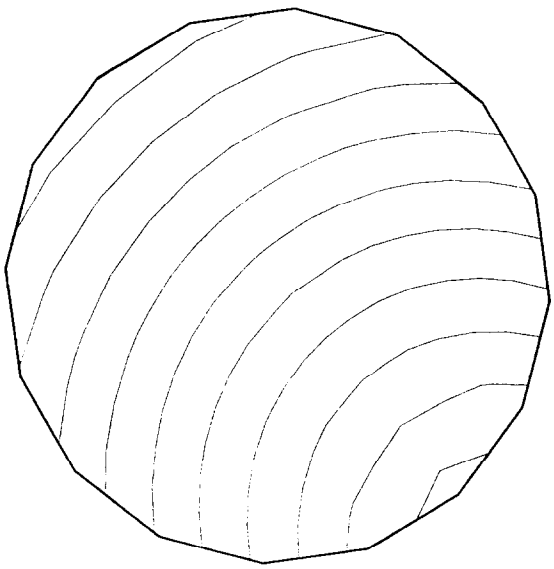

FIG. 6. The point source problem: contour plot of the imaginary part of the solution with the Sommerfeld boundary condition.

In Table I we compare the results using the two boundary conditions quantitatively. We list the values of the imaginary part of the exact solution at various points on the artificial boundary together with the values obtained by using the

obtained with the DtN boundary condition are better by far. It took 8.3 s to run the scheme using the Sommerfeld condition, and only 0.4 s more to run the one using the DtN condition.

Next we consider the problem of a circular boundary in an infinite plane with nonuniform boundary values. Figure 7 shows the mesh, where the radius of the boundary is $a = 0.5$, $\mathscr{B}$ is taken at $R = 1, k = 1$, and the elements are bilinear. The boundary value is $\cos j\theta$, where $j$ ranges from 0 (uniform boundary value) to 5.

TABLE I

The Point Source Problem: The Imaginary Part of the Solution at Various Points on the Artificial Boundary

| $\theta$ | Exact | DtN | Sommerfeld ($S_0$) | DtN error | $S_0$ error |
|---|---|---|---|---|---|
| −45° | 0.2413 | 0.2419 | 0.3308 | 0.2% | 37.1% |
| 0° | 0.2193 | 0.2195 | 0.2320 | 0.1% | 5.8% |
| 45° | 0.1704 | 0.1696 | 0.1136 | 0.5% | 33.3% |
| 90° | 0.1268 | 0.1256 | 0.0468 | 0.9% | 63.1% |
| 135° | 0.1103 | 0.1089 | 0.0261 | 1.3% | 76.3% |

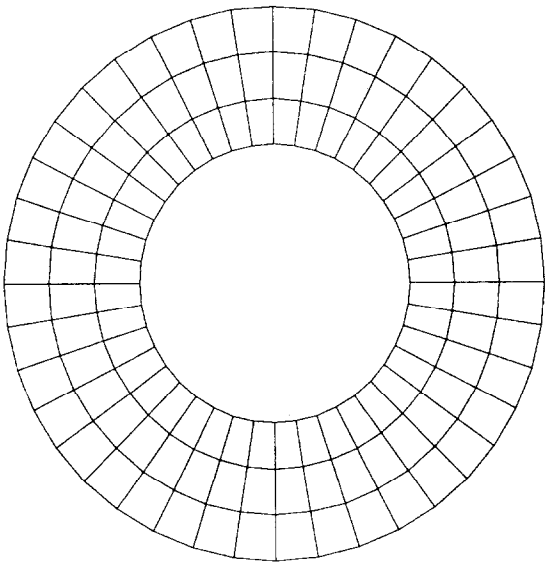

FIG. 7. Mesh for the non-uniform circular boundary problem.

Beside the Sommerfeld condition, three other approximate boundary conditions are examined:

$$E_1: \quad -u_\nu = \left(-ik + \frac{1}{2R}\right) u; \tag{65}$$

$$E_2: \quad -u_\nu = \left(-ik + \frac{1}{2R}\right) u - \left(\frac{i}{2kR^2} + \frac{1}{2k^2R^3}\right) \frac{\partial^2 u}{\partial \theta^2}; \tag{66}$$

$$A_3: \quad -u_\nu = \left(-ik + \frac{1}{2R} - \frac{i}{8kR^2} + \frac{1}{8k^2R^3}\right) u - \left(\frac{i}{2kR^2} + \frac{1}{2k^2R^3}\right) \frac{\partial^2 u}{\partial \theta^2}. \tag{67}$$

$E_1$ and $E_2$ were proposed by Engquist and Majda [1] and $A_3$ by Feng [3]. Figure 8 illustrates the real part of the solution along the boundary $\mathscr{B}$ for the case $j = 2$. The exact solution is given by (30) where the integration is over the inner boundary, and $R$ is replaced by $a$. The DtN solution is hardly distinguishable from the exact solution, while the solutions using $E_1$, $E_2$, and $A_3$ are off by about 25% at the peaks. The solution based on the Sommerfeld condition is the least accurate. The scheme using $A_3$ took 10 s to run; that using the DtN—half a second more, including the evaluation of four coefficients in the DtN kernel (32).

We have computed the exact solution and those based on the DtN and the $A_3$ boundary conditions for $0 \leqslant j \leqslant 5$. In Table II we list their values on the boundary $\mathscr{B}$ at $\theta = 0$, together with their relative errors. The results show that both finite element solutions deteriorate when $j$ becomes large. This occurs because the

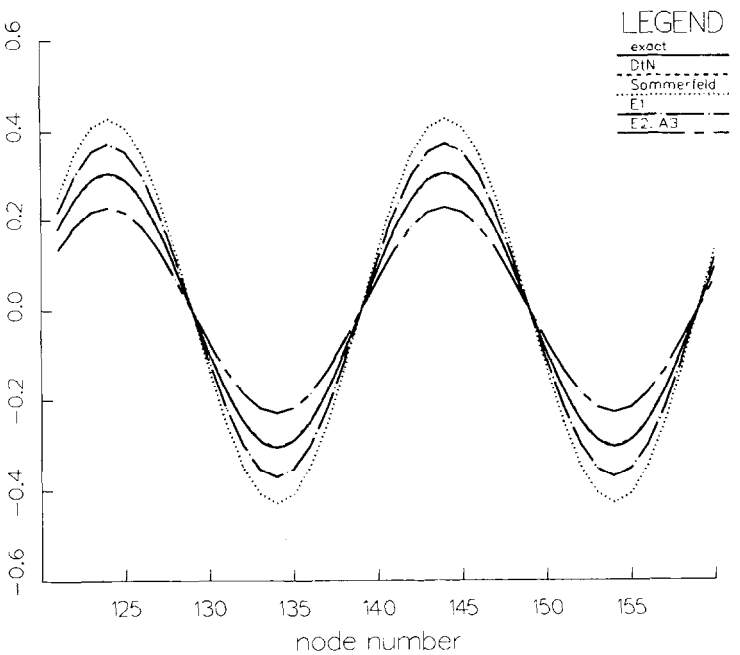


FIG. 8. The non-uniform circular boundary problem: comparison between various solutions for the boundary value $\cos 2\theta$. The real part of the solution along the boundary $\mathscr{B}$ is displayed.

elements are too large to accommodate so large a variation and because only four terms are used in the DtN kernel. However, for $j \leqslant 3$ the relative error of the $A_3$ scheme is larger than that of the DtN scheme by an order of magnitude. The DtN error remains small for all $j \leqslant 3$ while the $A_3$ error increases rapidly with $j$. This is related to the fact that all the approximate boundary conditions are most accurate for waves which hit the artificial boundary normally, and they become less accurate as the incidence becomes more oblique.

To demonstrate how the DtN method works on a problem with a finite domain,

TABLE II

The Non-uniform Circular Boundary Problem:
Results Are Given at $\theta = 0$ on the Boundary $\mathscr{B}$

| $j$ | Exact | DtN | $A_3$ | DtN error | $A_3$ error |
|---|---|---|---|---|---|
| 0 | 0.630 | 0.632 | 0.589 | 0.35% | 6.4% |
| 1 | 0.565 | 0.568 | 0.503 | 0.55% | 10.9% |
| 2 | 0.303 | 0.305 | 0.225 | 0.36% | 25.7% |
| 3 | 0.138 | 0.135 | 0.080 | 2.67% | 42.5% |
| 4 | 0.067 | 0.059 | 0.027 | 10.8% | 58.8% |
| 5 | 0.033 | 0.024 | 0.009 | 26.1% | 73.1% |

we consider a 2-dimensional problem in the interior of a circle. The circle is of radius $a = 1$, $u$ is prescribed as zero on its boundary, and $k = 1$. In addition, on a circular arc of 45° with radius $r = \frac{5}{6}$ we prescribe $u = 1$. We introduce the artificial boundary at $R = 0.5$ and again use the mesh in Fig. 7. Note that the DtN boundary condition (31), (38) is applied along the *inner* boundary. The line where $u = 1$ is represented by an arc of six nodes, at a distance of one layer of elements from the outer boundary. Figure 9 is a contour plot of the solution. The values vary from 0 on the outer boundary to 1 on the arc.

We close with an application of the DtN boundary condition (31), (41) to Laplace's equation outside a circular boundary with the boundary value $\cos 2\theta$. We use exactly the same setup as before, only now $k = 0$. We will use two approximate boundary conditions on the artificial boundary, either $u = 0$ or $u_\nu = 0$, since the exact solution and all of its derivatives tend to zero at infinity. In Fig. 10 a comparison is made between the exact solution and those obtained using the DtN and the two approximate boundary conditions. The values shown lie on the circle $r = \frac{2}{3}$, which is at a distance of one layer of elements from the inner boundary. The superiority of the DtN boundary condition over the two approximate ones is apparent again. At points closer to the artificial boundary the solution using the DtN condition is equally good, while the approximate solutions are even less accurate.

In order to use the approximate boundary condition, we might attempt to put the artificial boundary at a larger distance from the source, where the approximate boundary condition is more accurate. But then if we do not want to increase the cost by adding more elements to the mesh, we must "stretch" the mesh radially,

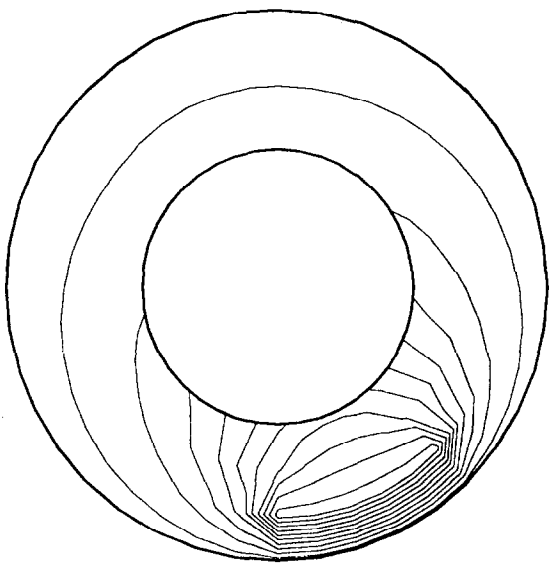

FIG. 9. The 45° circular arc problem: contour plot of the solution with the DtN boundary condition.

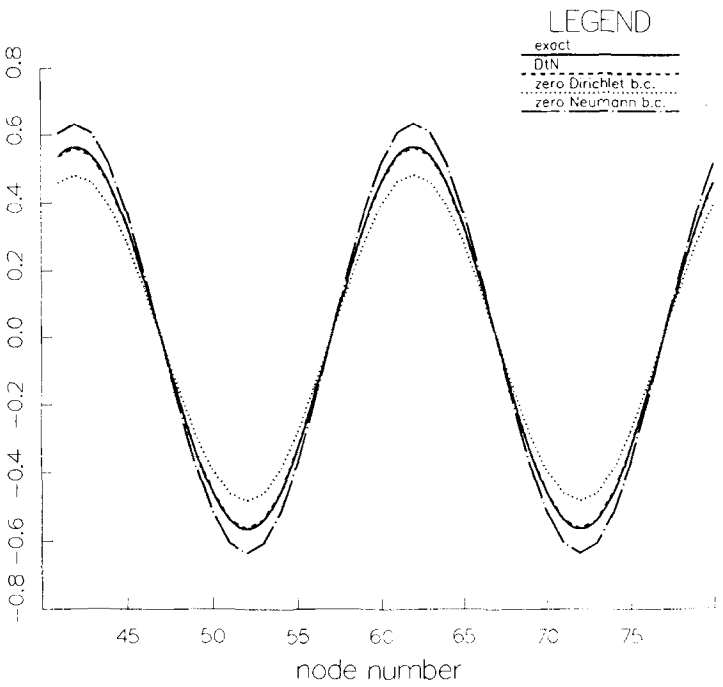


FIG. 10. The Laplace problem: comparison between various solutions for $\cos 2\theta$ distribution on the inner boundary. The solution along the circle $r = \frac{2}{3}$ is displayed.

resulting in slender elements. Numerical tests show that the ill-shaped elements have the effect of deteriorating the results even more. Thus no matter where the artificial boundary is placed, the solution using the DtN condition is more accurate than those using approximate boundary conditions.

## 7. SUMMARY

We have shown how to obtain explicitly the *exact* Dirichlet-to-Neumann boundary condition on an artificial boundary for the reduced wave equation and for Laplace's equation in infinite domains and in large domains. When applied in conjunction with the finite element method it results in an efficient and accurate scheme. The non-locality of the boundary condition has no effect on the banded structure of the finite element matrix, and the cost due to the complexity of the boundary condition is marginal.

## ACKNOWLEDGMENTS

This work was supported in part by ONR, AFOSR, and NSF.

## REFERENCES

1. B. ENGQUIST AND A. MAJDA, *Math. Comput.* **31**, 629 (1977).
2. A. BAYLISS AND E. TURKEL, *Commun. Pure Appl. Math.* **33**, 707 (1980).
3. K. FENG, in *Proceedings, International Congress of Mathematicians, Warsaw, Poland, 1983*, p. 1439.
4. B. GUSTAFSSON AND H. O. KREISS, *J. Comput. Phys.* **33**, 333 (1979).
5. T. HAGSTROM AND H. B. KELLER, *SIAM J. Math. Anal.* **17**, 322 (1986).
6. L. TING AND M. J. MIKSIS, *J. Acoust. Soc. Amer.* **80**, 1825 (1986).
7. G. J. FIX AND S. P. MARIN, *J. Comput. Phys.* **28**, 253 (1978).
8. R. C. MACCAMY AND S. P. MARIN, *Int. J. Math. Math. Sci.* **3**, 311 (1980).
9. S. P. MARIN, *IEEE Trans. Antennas Propag.* **AP-30**, 1045 (1982).
10. S. P. MARIN, Thesis, Carnegie-Mellon University, 1978 (unpublished).
11. C. I. GOLDSTEIN, *Math. Comput.* **39**, 309 (1982).
12. G. STRANG AND G. J. FIX, *An Analysis of the Finite Element Method* (Prentice-Hall, Englewood Cliffs, NJ, 1973).