

The Use of Huygens' Equivalence Principle for Solving 3-D Volume Integral Equation of Scattering

Cai-Cheng Lu and Weng Cho Chew, *Fellow, IEEE*

Abstract—A three-dimensional (3-D) version of the nested equivalent principle algorithm (NEPAL) is presented. In 3-D, a scatterer is first decomposed into N subscatterers. Then, spherical wave functions are used to represent the scattered field of the subscatterers. Subscatterers are divided into different levels of groups in a nested manner. For example, each group consists of eight subgroups, and each subgroup contains eight sub-subgroups, and so on. For each subgroup, the scattering solution is first solved and the number of subscatterers of the subgroup is then reduced by replacing the interior subscatterers with boundary subscatterers using Huygens' equivalence principle. As a result, when the subgroups are combined to form a higher level group, the group will have a smaller number of subscatterers. This process is repeated for each level, and in the last level, the number of subscatterers is proportional to that of boundary size of the scatterers. This algorithm has a computational complexity of $O(N^2)$ in three dimensions for all excitations and has the advantage of solving large scattering problems for multiple excitations. This is in contrast to Gaussian elimination which has a computational complexity of $O(N^3)$.

I. INTRODUCTION

THE computation of electromagnetic scattering of three-dimensional (3-D) objects finds applications in many areas. Hence, it has been earnestly studied by many engineers, scientists, and mathematicians alike. As a result, many algorithms have been developed for solving 3-D scattering problems. Among these algorithms, finite difference time domain methods are popular due to their lower complexity and ease of implementations. Many of these methods are for one excitation only, however, and the solution process needs to be repeated for multiple excitations. Moreover, absorbing boundary conditions are required for those algorithms. On the other hand, the solution to integral equations automatically satisfies the radiation condition. As a result, an integral equation solver with computational complexity comparable to differential equation solvers is a viable alternative solution technique to these scattering problems.

This paper presents an integral equation solver using the nested equivalence principle algorithm (NEPAL) which has been successfully applied to two-dimensional (2-D) problems [1]. It is known that in solving an integral equation, one can

first replace the volume scatterer by small subscatterers where the size of a subscatterer is much smaller than a wavelength [2]. The unknown function to be sought is expanded in terms of basis functions which usually have their supports on the subscatterers. By matching the field on the subscatterers, a set of linear equations are formed. The number of unknowns is proportional to the number of subscatterers in this case. Physically, each subscatterer can be considered a scattering center as in [3]–[7].

The interaction of a subscatterer with the other subscatterers can be described by interaction matrices [8]. If there are N subscatterers, then there will be N^2 interaction matrices since each subscatterer will interact with all the other subscatterers including itself. The N^2 interaction matrices can be found with N^3 operations as in [8] and [9]. The idea of NEPAL is to reduce the number of scattering centers, and hence to reduce the CPU time required for the solution, and essentially to find the inverse of the integral operator with computational complexity less than $O(N^3)$.

Similar algorithms for inversion of a matrix using the nested dissection method for the finite element method can be found in [10]. It is shown in [1] that the computational complexity of NEPAL is asymptotically $N^{1.5}$ for 2-D problems. In this paper, we will first formulate NEPAL for 3-D problems which has a computational complexity of N^2 . In addition, we will use some of the symbols differently. For example, i is used either as an index to a variable or as the imaginary number $\sqrt{-1}$. The meaning is deduced from the context.

II. FORMULATION OF THE PROBLEM

A 3-D volume can be discretized into a set of small cubic boxes called subscatterers. Each subscatterer has a volume Δ^3 . The scattering property of each subscatterer can be represented by a multipole with order proportional to the electrical size of the subscatterer. The scattered field due to the i th subscatterer in isolation, located at $\mathbf{r}_i$, can be written as [8]

$$\begin{aligned} E_i^{\text{sca}}(\mathbf{r}) &= \sum_{nm} \{ \mathbf{M}_{nm}(\mathbf{r} - \mathbf{r}_i) [\mathbf{b}_i^M]_{nm} \\ &\quad + \mathbf{N}_{nm}(\mathbf{r} - \mathbf{r}_i) [\mathbf{b}_i^E]_{nm} \} \\ &= \bar{\psi}^t(\mathbf{r} - \mathbf{r}_i) \cdot \mathbf{b}_i \end{aligned} \quad (1)$$

where $\mathbf{b}_i = [\mathbf{b}_i^M, \mathbf{b}_i^E]^t$ is the multipole coefficients and superscripts M and E stand for TE and TM components, respectively. Here, $\bar{\psi}^t(\mathbf{r}) = [\mathbf{M}, \mathbf{N}]$ is the vector spherical

Manuscript received July 7, 1994; revised November 30, 1994. This work was supported in part by Office of Naval Research Grant N00014-89-J1286, the Army Research Office Contract DAAL03-91-G-0339, The National Science Foundation Grant NSF ECS 92-24466, NASA Grant NASA NAG 2-871, and the National Center for Supercomputing Applications (NCSA) at the University of Illinois, Urbana-Champaign.

The authors are with the Department of Electrical and Computer Engineering, University of Illinois, Urbana, IL 61801 USA.

IEEE Log Number 9410725.

wave function, and $\mathbf{M}$ and $\mathbf{N}$ are given by [6] and [11]

$$\begin{aligned}\mathbf{M}_{nm}(\mathbf{r}) &= \hat{\theta} \frac{im}{\sin \theta} z_n(kr) Y_{nm}(\theta, \phi) \\ &\quad - \hat{\phi} z_n(kr) \frac{\partial Y_{nm}(\theta, \phi)}{\partial \theta}, \\ \mathbf{N}_{nm}(\mathbf{r}) &= \hat{r} \frac{n(n+1)}{kr} z_n(kr) Y_{nm}(\theta, \phi) \\ &\quad + \hat{\theta} \frac{1}{kr} \frac{\partial}{\partial r} [r z_n(kr)] \frac{\partial Y_{nm}(\theta, \phi)}{\partial \theta} \\ &\quad + \hat{\phi} \frac{im}{kr \sin \theta} \frac{\partial}{\partial r} [r z_n(kr)] Y_{nm}(\theta, \phi) \\ &\quad n = 1, 2, \dots, m = -n, \dots, n.\end{aligned}$$

In the above, $z_n(kr)$ is a spherical Bessel function or spherical Hankel function of the first kind depending on whether $\bar{\psi}^t(\mathbf{r})$ is an incoming wave or an outgoing wave, respectively. $\Re_g \bar{\psi}$ denotes the regular part of $\bar{\psi}$, which is obtained from $\bar{\psi}$ by setting the Neumann function to zero. In the above, Y_{nm} is a spherical harmonic, and it is defined by [12] and [13]

$$\begin{aligned}Y_{nm}(\theta, \phi) &= (-1)^m \sqrt{\frac{2n+1}{4\pi} \frac{(n-m)!}{(n+m)!}} P_n^m(\cos \theta) e^{im\phi}, \\ &\quad m \geq 0, \\ Y_{nm}(\theta, \phi) &= (-1)^{|m|} Y_{n|m|}(\theta, \phi), \quad m < 0\end{aligned}$$

where, P_n^m is the associated Legendre function.

Expressing the incident field in the same manner

$$\mathbf{E}^{\text{inc}}(\mathbf{r}) = \Re_f \bar{\psi}^t(\mathbf{r} - \mathbf{r}_i) \cdot \bar{\alpha}_{is} \cdot \mathbf{a}_s \quad (2)$$

we can relate the unknown coefficients $\mathbf{b}_i$ to $\mathbf{a}_s$ via a T -matrix by matching the field boundary of the subscatterer as in [3], [4], and [11]

$$\mathbf{b}_i = \bar{T}_i \cdot \bar{\alpha}_{is} \mathbf{a}_s. \quad (3)$$

In the above, $\bar{\alpha}_{is}$ is a translation matrix [6], [13], and [14] that changes the coordinates of the multipole expressions, $\bar{T}_i$ is the T -matrix for the i th subscatterer in isolation, and $\mathbf{a}_s$ is the incident wave amplitude.

When Δ is small, $\bar{T}_i$ is diagonal and can be approximated by that of an equivolume sphere,¹ and can be found using the Mie series [6]. When N_A subscatterers are present, the coefficients $\mathbf{b}_i$ will be unknown which satisfy the following brute force equation [3], [5], and [7]

$$\mathbf{b}_i = \bar{T}_i \cdot \left[\bar{\alpha}_{is} \cdot \mathbf{a}_s + \sum_{\substack{j=1 \\ j \neq i}}^{N_A} \bar{\alpha}_{ij} \cdot \mathbf{b}_j \right], \quad i = 1, 2, \dots, N_A. \quad (4)$$

Since the solution to the above equations has to be linearly proportional to $\bar{\alpha}_{js} \cdot \mathbf{a}_s$, it can be written as

$$\mathbf{b}_i = \sum_{j=1}^{N_A} \bar{I}_{ij(N_A)} \cdot \bar{\alpha}_{js} \cdot \mathbf{a}_s, \quad i = 1, 2, \dots, N_A. \quad (5)$$

¹This approximation was first used by Richmond [2].

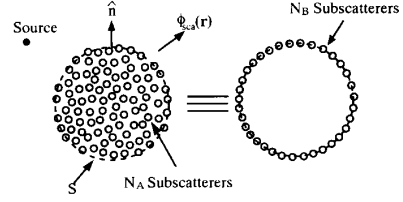


Fig. 1. The replacement of volume scattering centers by surface scattering centers using Huygens' principle can greatly reduce the number of scattering centers needed to represent the scatterer.

Hence, the total scattered field is given by

$$\mathbf{E}^{\text{sca}}(\mathbf{r}) = \sum_{i=1}^{N_A} \bar{\psi}^t(\mathbf{r} - \mathbf{r}_i) \cdot \sum_{j=1}^{N_A} \bar{I}_{ij(N_A)} \cdot \bar{\alpha}_{js} \cdot \mathbf{a}_s. \quad (6)$$

In the above, $\bar{I}_{ij(N_A)}$, $i, j = 1, 2, \dots, N_A$ are called interaction matrices; they describe the interactions between N_A subscatterers. They also describe a scattering process whereby the wave enters through the j th scattering center and leaves through the i th scattering center. They can be found in a variety of ways, for example, by Gaussian elimination or by a recursive algorithm as in [8] and [9].

III. EQUIVALENCE OF INTERACTION MATRICES

As shown in the above section, the scattered field can be determined by the N_A^2 interaction matrices for a group of N_A subscatterers. In this section, we will show by the Huygens' equivalence principle that there is another set of interaction matrices which will generate the same scattered field. To this end, we first give a definition of equivalent interaction matrices.

Definition: Two sets of interaction matrices $\bar{I}_{ij(N_A)}$ and $\bar{I}_{ij(N_B)}$ are said to be equivalent if they generate the same scattered field via (6) for the same incident field.

In the following, we will give a mathematical description for the equivalence of the interior subscatterers by those on the surface using Huygens' equivalence principle [13] (see Fig. 1).

First, we assume that S is the boundary of a volume region V , and V contains sources which generate field $\mathbf{E}$ and $\mathbf{H}$ on the boundary. Next, we let $\mathbf{r}'$ be a point outside V , as shown in Fig. 2. Huygens' equivalence principle states that the field at $\mathbf{r}'$ due to sources inside V can be represented via equivalent sources on the boundary

$$\begin{aligned}\mathbf{E}(\mathbf{r}') &= - \oint_S dS [\hat{n} \times \mathbf{E}(\mathbf{r}) \cdot \nabla \times \bar{\mathbf{G}}_e(\mathbf{r}, \mathbf{r}') \\ &\quad + i\omega\mu\hat{n} \times \mathbf{H}(\mathbf{r}) \cdot \bar{\mathbf{G}}_e(\mathbf{r}, \mathbf{r}')] \quad (7)\end{aligned}$$

where

$$\bar{\mathbf{G}}_e(\mathbf{r}, \mathbf{r}') = \left(\bar{\mathbf{I}} + \frac{\nabla \nabla}{k^2} \right) \frac{e^{ik|\mathbf{r} - \mathbf{r}'|}}{4\pi|\mathbf{r} - \mathbf{r}'|} \quad (7a)$$

is the free space dyadic Green's function. Equation (7) has a double role. First, it provides an indirect approach to compute the field at points outside V . Second, it tells us how to construct the equivalent sources: the equivalent sources are simply the

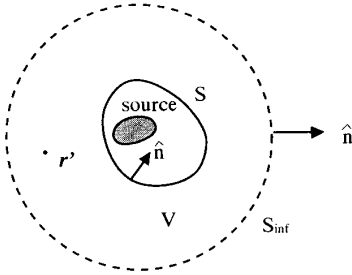


Fig. 2. The Huygens' equivalence principle.

tangential components of the fields, which are electric and magnetic surface currents.

Apart from the above two points, we also observed that: 1) The equivalence principle does not specify the type and the number of the original sources inside V ; they could be any kind of sources: induced sources, other equivalent sources, a point source, several point sources, or distributed sources. 2) The equivalent sources are uniquely determined by the shape of the surface and the density of the tangential field components. The original sources could be volumetric, and hence, reside in a 3-D space, while the equivalence sources reside on a surface which is on a space of dimensionality one less than the 3-D space. Consequently, this allows us to replace a large number of sources with a relatively small number of sources.

The mathematical representation of a source can be different. For example, (2) represents an incident field by its multipole expansion about the source point, while (6) represents the induced sources by a set of matrices $\bar{I}_{ij(N_A)}$, $i, j = 1, 2, \dots, N_A$. Similarly, we can represent equivalent sources in the same manner, i.e., via another set of interaction matrices $\bar{I}_{ij(N_B)}$, $i, j = 1, 2, \dots, N_B$. We shall now construct the relation between the two sets of matrices using (7).

A. The Outside Equivalence Problem

For waves leaving the scatterer, we can replace the sources of the outgoing waves with Huygens' equivalent sources. First, we divide the surface S into N_B small surfaces or patches, where each patch is of area ΔS . As a result, (7) can be replaced by a discrete summation using integration quadrature rule yielding

$$\begin{aligned} \mathbf{E}(\mathbf{r}') = & - \sum_{i=1}^{N_B} \Delta S w_i [\hat{n} \times \mathbf{E}(\mathbf{r}_i) \cdot \nabla \times \bar{\mathbf{G}}_e(\mathbf{r}_i, \mathbf{r}') \\ & + i\omega\mu\hat{n} \times \mathbf{H}(\mathbf{r}_i) \cdot \bar{\mathbf{G}}_e(\mathbf{r}_i, \mathbf{r}')] \end{aligned} \quad (8)$$

where the weight w_i depends on the quadrature rule used [15]. At the i th patch, we let $\mathbf{r} = \mathbf{r}'' + \mathbf{r}_i$. Here, $\mathbf{r}''$ is the local coordinate for the i th patch. Then

$$\mathbf{r} - \mathbf{r}' = \mathbf{r}'' + \mathbf{r}_i - \mathbf{r}' = \mathbf{r}'' - (\mathbf{r}' - \mathbf{r}_i). \quad (9)$$

Under the condition that $|\mathbf{r}' - \mathbf{r}_i| > |\mathbf{r}''|$, the dyadic Green's function $\bar{\mathbf{G}}_e(\mathbf{r}, \mathbf{r}')$ can be expanded as $(\mathbf{r}'_i = \mathbf{r}' - \mathbf{r}_i)$, [13, p.

409])

$$\begin{aligned} \bar{\mathbf{G}}_e(\mathbf{r}'', \mathbf{r}'_i) = & \sum_{nm} \frac{ik(-)^m}{n(n+1)} [\Re_g \mathbf{M}_{n,-m}(\mathbf{r}'') \mathbf{M}_{nm}(\mathbf{r}'_i) \\ & + \Re_g \mathbf{N}_{n,-m}(\mathbf{r}'') \mathbf{N}_{nm}(\mathbf{r}'_i)]. \end{aligned} \quad (10)$$

Then it is easy to show that

$$\bar{\mathbf{G}}_e(\mathbf{r}_i, \mathbf{r}') = \bar{\mathbf{G}}_e(\mathbf{r}'' = 0, \mathbf{r}'_i). \quad (11)$$

Consequently, (8) can be written more compactly as

$$\mathbf{E}(\mathbf{r}') = \sum_i \sum_{nm} \{ \mathbf{M}_{nm}(\mathbf{r}'_i) \cdot [\mathbf{b}_i^M]_{nm} + \mathbf{N}_{nm}(\mathbf{r}'_i) [\mathbf{b}_i^E]_{nm} \} \quad (12)$$

where

$$[\mathbf{b}_i^M]_{nm} = \frac{-ik^2(-)^m}{n(n+1)} \Delta S w_i [\hat{n}_i \times \mathbf{E}(\mathbf{r}_i) \cdot \Re_g \mathbf{N}_{n,-m}(0)], \quad (13a)$$

$$[\mathbf{b}_i^E]_{nm} = \frac{-ik^2(-)^m}{n(n+1)} \Delta S w_i [i\eta \hat{n}_i \times \mathbf{H}(\mathbf{r}_i) \cdot \Re_g \mathbf{N}_{n,-m}(0)] \quad (13b)$$

$\hat{n}_i$ is the average normal of the i th patch, and $\eta = \sqrt{\mu_0/\epsilon_0}$ is the free space impedance.

In the above derivation, the identities

$$\nabla \times \mathbf{M} = k\mathbf{N}, \quad \nabla \times \mathbf{N} = k\mathbf{M}$$

have been used. Equation (12) gives the source-field relation for the equivalent sources on the surface S .

We assume that the source-field relation for the interior sources is of the form as (1) with multipole amplitude given by $\mathbf{a}_j$, $j = 1, 2, \dots, N_A$. Then

$$\begin{aligned} \mathbf{E}(\mathbf{r}) = & \sum_{j,\nu\mu} \{ \mathbf{M}_{\nu\mu}(\mathbf{r} - \mathbf{r}_j) \cdot [\mathbf{a}_j^M]_{\nu\mu} + \mathbf{N}_{\nu\mu}(\mathbf{r} - \mathbf{r}_j) \\ & \cdot [\mathbf{a}_j^E]_{\nu\mu} \}, \end{aligned} \quad (14a)$$

$$\begin{aligned} \mathbf{H}(\mathbf{r}) = & \frac{1}{i\eta} \sum_{j,\nu\mu} \{ \mathbf{N}_{\nu\mu}(\mathbf{r} - \mathbf{r}_j) \cdot [\mathbf{a}_j^M]_{\nu\mu} + \mathbf{M}_{\nu\mu}(\mathbf{r} - \mathbf{r}_j) \\ & \cdot [\mathbf{a}_j^E]_{\nu\mu} \}. \end{aligned} \quad (14b)$$

We let $\mathbf{r} = \mathbf{r}_i$ in (14a) and (14b), and substitute them into (13a), we have

$$\begin{aligned} [\mathbf{b}_i^M]_{nm} = & \frac{-ik^2(-)^m}{n(n+1)} \Delta S w_i \Re_g \mathbf{N}_{n,-m}(0) \\ & \cdot \sum_{j,\nu\mu} (\hat{n}_i \times \mathbf{M}_{\nu\mu}(\mathbf{r}_{ji}) \cdot [\mathbf{a}_j^M]_{\nu\mu} + \hat{n}_i \\ & \cdot \mathbf{N}_{\nu\mu}(\mathbf{r}_{ji}) \cdot [\mathbf{a}_j^E]_{\nu\mu}) \end{aligned}$$

where $\mathbf{r}_{ji} = \mathbf{r}_i - \mathbf{r}_j$. Alternatively, we can write the above as

$$[\mathbf{b}_i^M]_{nm} = \sum_{j,\nu\mu} \{ [\bar{\mathbf{c}}_{ij}^{MM}]_{nm,\nu\mu} \cdot [\mathbf{a}_j^M]_{\nu\mu} + [\bar{\mathbf{c}}_{ij}^{ME}]_{nm,\nu\mu} [\mathbf{a}_j^E]_{\nu\mu} \}. \quad (15a)$$

Similarly, we have

$$[\mathbf{b}_i^E]_{nm} = \sum_{j,\nu\mu} \{ [\bar{\mathbf{c}}_{ij}^{EM}]_{nm,\nu\mu} \cdot [\mathbf{a}_j^M]_{\nu\mu} + [\bar{\mathbf{c}}_{ij}^{EE}]_{nm,\nu\mu} \cdot [\mathbf{a}_j^E]_{\nu\mu} \}. \quad (15b)$$

In the above

$$[\bar{\mathbf{c}}_{ij}^{MM}]_{nm,\nu\mu} = \frac{-ik^2(-)^m}{n(n+1)} \Delta S w_i \Re_g \mathbf{N}_{n,-m}(0) \cdot \hat{n}_i \times \mathbf{M}_{\nu\mu}(\mathbf{r}_{ji}), \quad (16a)$$

$$[\bar{\mathbf{c}}_{ij}^{ME}]_{nm,\nu\mu} = \frac{-ik^2(-)^m}{n(n+1)} \Delta S w_i \Re_g \mathbf{N}_{n,-m}(0) \cdot \hat{n}_i \times \mathbf{N}_{\nu\mu}(\mathbf{r}_{ji}), \quad (16b)$$

$$[\bar{\mathbf{c}}_{ij}^{EM}]_{nm,\nu\mu} = [\bar{\mathbf{c}}_{ij}^{ME}]_{nm,\nu\mu}, \quad (16c)$$

$$[\bar{\mathbf{c}}_{ij}^{EE}]_{nm,\nu\mu} = [\bar{\mathbf{c}}_{ij}^{MM}]_{nm,\nu\mu}. \quad (16d)$$

Denoting

$$\bar{\mathbf{h}}_{ij}^{(o)} = \bar{\mathbf{c}}_{ij} = \begin{bmatrix} \bar{\mathbf{c}}_{ij}^{MM} & \bar{\mathbf{c}}_{ij}^{ME} \\ \bar{\mathbf{c}}_{ij}^{EM} & \bar{\mathbf{c}}_{ij}^{EE} \end{bmatrix}$$

we have the relation between the equivalent sources and the original sources

$$\mathbf{b}_i = \sum_j \bar{\mathbf{h}}_{ij}^{(o)} \cdot \mathbf{a}_j$$

where $\mathbf{b}_i = [\mathbf{b}_i^M, \mathbf{b}_i^E]^t$ and $\mathbf{a}_j = [\mathbf{a}_j^M, \mathbf{a}_j^E]^t$. If there is only one interior source located at $\mathbf{r}_j$ inside V , then the field at $\mathbf{r}'$ corresponding to this source is given by (14a) with $N_A = 1$, or

$$\mathbf{E}(\mathbf{r}') = \bar{\psi}^t(\mathbf{r}' - \mathbf{r}_j) \cdot \mathbf{a}_j.$$

The above is the direct source-field relation. On the other hand, we have an indirect source-field relation in (12) via the Huygens' equivalence principle, or

$$\mathbf{E}(\mathbf{r}') = \sum_{i=1}^{N_B} \bar{\psi}^t(\mathbf{r}' - \mathbf{r}_i) \cdot \bar{\mathbf{h}}_{ij}^{(o)} \cdot \mathbf{a}_j.$$

Equating the two representations, we have for an arbitrary $\mathbf{a}_j$

$$\bar{\psi}^t(\mathbf{r} - \mathbf{r}_j) = \sum_{i=1}^{N_B} \bar{\psi}^t(\mathbf{r} - \mathbf{r}_i) \cdot \bar{\mathbf{h}}_{ij}^{(o)}. \quad (17) \quad \text{or}$$

Equation (17) is the first equation which will be used to derive the equivalence between interaction matrices of interior subscatterers and boundary subscatterers.

B. The Inside Equivalence Problem

In a similar manner, the sources generating incident waves impinging on a scatterer can be replaced by Huygens' equivalent sources. Now we consider the sources outside of V and compute the field at $\mathbf{r}'$ inside V due to outside sources. Using similar steps as in deriving $\bar{\mathbf{h}}_{ii'}^{(o)}$, we can write the field $\mathbf{E}(\mathbf{r}')$ in the same form as (12), (13a), and (13b), except that $\mathbf{r}'$ is inside V in this case.

Suppose that the source is located at $\mathbf{r}_s$ with the multipole amplitude $\mathbf{a}_s$, then the fields at $\mathbf{r}$ due to this source are given by

$$\mathbf{E}(\mathbf{r}) = \sum_{\nu\mu} \{ \mathbf{M}_{\nu\mu}(\mathbf{r} - \mathbf{r}_s) \cdot [\mathbf{a}_s^M]_{\nu\mu} + \mathbf{N}_{\nu\mu}(\mathbf{r} - \mathbf{r}_s) \cdot [\mathbf{a}_s^E]_{\nu\mu} \}, \quad (18a)$$

$$\mathbf{H}(\mathbf{r}) = \frac{1}{i\eta} \sum_{\nu\mu} \{ \mathbf{N}_{\nu\mu}(\mathbf{r} - \mathbf{r}_s) \cdot [\mathbf{a}_s^M]_{\nu\mu} + \mathbf{M}_{\nu\mu}(\mathbf{r} - \mathbf{r}_s) \cdot [\mathbf{a}_s^E]_{\nu\mu} \}. \quad (18b)$$

Equation (18) gives an outgoing wave centered at $\mathbf{r}_s$, and the field point $\mathbf{r}$ is not specified yet. Now, we let $\mathbf{r}$ be a point in the vicinity of $\mathbf{r}_i \in S$, with $i = 1, 2, \dots, N_B$. Observe that at $\mathbf{r}$, the field can also be considered incoming waves focusing at $\mathbf{r}_i$. This is more rigorously given by the vector wave translation equation from the vector addition theorem as in [12] and [14]

$$\begin{aligned} \mathbf{M}(\mathbf{r} - \mathbf{r}_s) &= \Re_g \mathbf{M}(\mathbf{r} - \mathbf{r}_i) \cdot \bar{\mathbf{A}}_{is} + \Re_g \mathbf{N}(\mathbf{r} - \mathbf{r}_i) \cdot \bar{\mathbf{B}}_{is}, \\ \mathbf{N}(\mathbf{r} - \mathbf{r}_s) &= \Re_g \mathbf{M}(\mathbf{r} - \mathbf{r}_i) \cdot \bar{\mathbf{B}}_{is} + \Re_g \mathbf{N}(\mathbf{r} - \mathbf{r}_i) \cdot \bar{\mathbf{A}}_{is}. \end{aligned}$$

Therefore, the above can be rewritten as

$$\mathbf{E}(\mathbf{r}) = \sum_{\nu\mu} \{ \Re_g \mathbf{M}_{\nu\mu}(\mathbf{r} - \mathbf{r}_i) \cdot [\mathbf{a}_{is}^M]_{\nu\mu} + \Re_g \mathbf{N}_{\nu\mu}(\mathbf{r} - \mathbf{r}_i) \cdot [\mathbf{a}_{is}^E]_{\nu\mu} \}, \quad (18c)$$

$$\mathbf{H}(\mathbf{r}) = \frac{1}{i\eta} \sum_{\nu\mu} \{ \Re_g \mathbf{N}_{\nu\mu}(\mathbf{r} - \mathbf{r}_i) \cdot [\mathbf{a}_{is}^M]_{\nu\mu} + \Re_g \mathbf{M}_{\nu\mu}(\mathbf{r} - \mathbf{r}_i) \cdot [\mathbf{a}_{is}^E]_{\nu\mu} \} \quad (18d)$$

where $\mathbf{a}_{is}$ is related to $\mathbf{a}_s$ by

$$\mathbf{a}_{is} = \begin{bmatrix} \bar{\mathbf{A}}_{is} & \bar{\mathbf{B}}_{is} \\ \bar{\mathbf{B}}_{is} & \bar{\mathbf{A}}_{is} \end{bmatrix} \cdot \mathbf{a}_s = \bar{\alpha}_{is} \cdot \mathbf{a}_s.$$

Substitution of (18c) and (18d) into (13a), we have

$$\begin{aligned} [\mathbf{b}_i^M]_{nm} &= \frac{-ik^2(-)^m}{n(n+1)} \Delta S w_i \Re_g \mathbf{N}_{n,-m}(0) \\ &\cdot \hat{n}_i \times \sum_{\nu\mu} (\Re_g \mathbf{M}_{\nu\mu}(0) \cdot [\mathbf{a}_{is}^M]_{\nu\mu} \\ &+ \Re_g \mathbf{N}_{\nu\mu}(0) \cdot [\mathbf{a}_{is}^E]_{\nu\mu}) \end{aligned}$$

$$[\mathbf{b}_i^E]_{nm} = \sum_{\nu\mu} \{ [\bar{\mathbf{c}}_i^{MM}]_{nm,\nu\mu} \cdot [\mathbf{a}_{is}^M]_{\nu\mu} + [\bar{\mathbf{c}}_i^{ME}]_{nm,\nu\mu} \cdot [\mathbf{a}_{is}^E]_{\nu\mu} \}. \quad (19a)$$

Similarly, we have

$$[\mathbf{b}_i^E]_{nm} = \sum_{\nu\mu} \{ [\bar{\mathbf{c}}_i^{EM}]_{nm,\nu\mu} \cdot [\mathbf{a}_{is}^M]_{\nu\mu} + [\bar{\mathbf{c}}_i^{EE}]_{nm,\nu\mu} \cdot [\mathbf{a}_{is}^E]_{\nu\mu} \}. \quad (19b)$$

Hence

$$\mathbf{b}_i = \bar{\mathbf{c}}_i \cdot \mathbf{a}_s = \begin{bmatrix} \bar{\mathbf{c}}_i^{MM} & \bar{\mathbf{c}}_i^{ME} \\ \bar{\mathbf{c}}_i^{ME} & \bar{\mathbf{c}}_i^{EE} \end{bmatrix} \cdot \mathbf{a}_{is} \quad (19c)$$

where

$$[c_i^{MM}]_{nm,\nu\mu} = \frac{-ik^2(-)^m}{n(n+1)} \Delta Sw_i \Re_g \mathbf{N}_{n,-m}(0) \cdot \hat{n}_i \times \Re_g \mathbf{M}_{\nu\mu}(0), \quad (20a)$$

$$[c_i^{ME}]_{nm,\nu\mu} = \frac{-ik^2(-)^m}{n(n+1)} \Delta Sw_i \Re_g \mathbf{N}_{n,-m}(0) \cdot \hat{n}_i \times \Re_g \mathbf{N}_{\nu\mu}(0), \quad (20b)$$

$$[c_i^{EM}]_{nm,\nu\mu} = [c_i^{ME}]_{nm,\nu\mu}, \quad (20c)$$

$$[c_i^{EE}]_{nm,\nu\mu} = [c_i^{MM}]_{nm,\nu\mu}. \quad (20d)$$

In the above

$$\Re_g \mathbf{M}_{\nu\mu}(0) = 0$$

and

$$\begin{aligned} \Re_g \mathbf{N}_{\nu\mu}(0) &= 0, \quad \nu \neq 1, \\ \Re_g \mathbf{N}_{1,\mp 1}(0) &= \frac{2}{3} \sqrt{\frac{3}{8\pi}} (\pm \hat{x} - i\hat{y}), \\ \Re_g \mathbf{N}_{1,0}(0) &= \frac{2}{3} \sqrt{\frac{3}{4\pi}} \hat{z}. \end{aligned}$$

Now, assume that the source is located outside V at $\mathbf{r}_s$, and $\mathbf{r}_j$ is an interior point, the incident field in the vicinity of $\mathbf{r}_j$ due to source at $\mathbf{r}_s$ can be written as

$$\mathbf{E}^{\text{inc}}(\mathbf{r}) = \Re_g \bar{\psi}^t(\mathbf{r} - \mathbf{r}_j) \cdot \bar{\alpha}_{js} \cdot \mathbf{a}_s. \quad (21)$$

Applying the equivalence principle, this field can be thought of as coming from the equivalent sources on the boundary

$$\mathbf{E}^{\text{inc}}(\mathbf{r}) = \sum_{i=1}^{N_B} \bar{\psi}^t(\mathbf{r} - \mathbf{r}_i) \cdot \mathbf{b}_i \quad (22)$$

where $\mathbf{b}_i$ is given by (19c).

Using translation formula to translate $\bar{\psi}^t(\mathbf{r} - \mathbf{r}_i)$, the outgoing wave originated at $\mathbf{r}_i$ to an incoming wave centered at $\mathbf{r}_j$, we have

$$\bar{\psi}^t(\mathbf{r} - \mathbf{r}_i) = \Re_g \bar{\psi}^t(\mathbf{r} - \mathbf{r}_j) \cdot \bar{\alpha}_{ji}. \quad (23)$$

Substituting $\bar{\psi}^t(\mathbf{r} - \mathbf{r}_i)$ in (21) by (23), and equating the resultant equation and (22) for arbitrary point $\mathbf{r}$ near $\mathbf{r}_j$, we obtain

$$\bar{\alpha}_{js} \cdot \mathbf{a}_s = \sum_{i=1}^{N_B} \bar{\alpha}_{ji} \cdot \bar{\mathbf{c}}_i \cdot \bar{\alpha}_{is} \cdot \mathbf{a}_s = \sum_{i=1}^{N_B} \bar{\mathbf{h}}_{ji}^{(i)} \cdot \bar{\alpha}_{is} \cdot \mathbf{a}_s \quad (24)$$

with

$$\bar{\mathbf{h}}_{ji}^{(i)} = \bar{\alpha}_{ji} \cdot \bar{\mathbf{c}}_i. \quad (25)$$

Equation (24) is the second equation which will be used for our equivalence problem.

With (17) and (24), we can easily relate the equivalent interaction matrices $\bar{\mathbf{I}}_{ij(N_B)}$ with the original ones $\bar{\mathbf{I}}_{ij(N_A)}$.

To this end, we start with (6), and replace $\bar{\psi}^t(\mathbf{r} - \mathbf{r}_i)$ in it by (17), and replace $\bar{\alpha}_{is} \cdot \mathbf{a}_s$ by (24), to obtain

$$\begin{aligned} \mathbf{E}^{\text{sca}}(\mathbf{r}) &= \sum_{m=1}^{N_B} \bar{\psi}^t(\mathbf{r} - \mathbf{r}_m) \cdot \sum_{i=1}^{N_A} \bar{\mathbf{h}}_{mi}^{(o)} \cdot \sum_{j=1}^{N_A} \bar{\mathbf{I}}_{ij(N_A)} \\ &\quad \cdot \sum_{m'=1}^{N_B} \bar{\mathbf{h}}_{jm'}^{(i)} \cdot \bar{\alpha}_{m's} \cdot \mathbf{a}_s. \end{aligned} \quad (26)$$

The above can be rewritten as

$$\mathbf{E}^{\text{sca}}(\mathbf{r}) = \sum_{m=1}^{N_B} \bar{\psi}^t(\mathbf{r} - \mathbf{r}_m) \cdot \sum_{m'=1}^{N_B} \bar{\mathbf{I}}_{mm'(N_B)} \cdot \bar{\alpha}_{m's} \cdot \mathbf{a}_s \quad (27)$$

where

$$\bar{\mathbf{I}}_{mm'(N_B)} = \sum_{i=1}^{N_A} \bar{\mathbf{h}}_{mi}^{(o)} \cdot \sum_{j=1}^{N_A} \bar{\mathbf{I}}_{ij(N_A)} \cdot \bar{\mathbf{h}}_{jm'}^{(i)}. \quad (28)$$

Equation (28) specifies the equivalence relation between the original interaction matrices and the equivalent ones. This is the key equation for NEPAL. We will explain how this equation is used to reduce the number of interaction matrices.

For a volume scatterer, the total subscatterers are divided into boundary subscatterers and interior subscatterers. If we define $\bar{\mathbf{h}}_{mi}^{(o)} = \bar{\mathbf{h}}_{mi}^{(i)} = \bar{\mathbf{I}}$ for $i = m$, then the right hand side of (28) may also include interaction matrices of boundary subscatterers.

Let N_A be the total number of subscatterers, and N_B , the number of boundary subscatterers. Then the number of interaction matrices is reduced from N_A^2 to N_B^2 . This could be significant if $N_B \ll N_A$ as shown in Fig. 1. Furthermore, there is no violation of addition theorem in using this equivalence principle.

IV. THE NESTED EQUIVALENCE PRINCIPLE ALGORITHM (NEPAL)

As mentioned in the introduction of this paper, the idea of NEPAL is to reduce the number of scattering centers, or the number of interaction matrices, as shown by (28). In this section, we will describe the steps involved in this algorithm.

To begin, we define a nested group systems which consists of different levels of groups, i.e., a level l group contains eight members, each member itself is a group of level $l-1$, where, $l = L, L = l, \dots, 2, 1$. In the scattering problem, we consider all the subscatterers in the original volume scatterer as a level L group. This group is first divided into eight level $L-1$ groups, each $L-1$ group is then divided into eight level $L-2$ groups, and so on. The process continues until the lowermost level of groups is reached. It can be seen that if the number of total subscatterers is $N = 8^L$, then the number of level l group is $N/8^l$.

We can start to solve the problem at any level. A brute force approach is to start at level L . In the approach of this paper, we start from level 2 (one can see from the following discussions that starting from level 1 results in exactly the same algorithm as starting from level 2 since there are no

interior subscatterers in level 1 groups). In each level, we find the scattering solution for each of the groups in this level, and then use Huygens' equivalence principle (28) to replace the interior subscatterers of a group by subscatterers on the boundary of the group, so that the scattering properties of the group are represented by the subscatterers on the boundary. As a result, the interior subscatterers in each group are "removed" and the number of subscatterers in each group is reduced.

For example, in the starting level (level 2), we first solve for the scattering solution for each group of 64 subscatterers. Then (28) is used to remove the interior subscatterers. In this level, there are only eight interior subscatterers for each group. After this step, each group contains 56 subscatterers with a known solution. One is to be reminded, however, that the solutions are for each of the groups only, i.e., the solutions account only for the interactions within each group in isolation. Moreover, we need to repeat this process in this level for $N/64$ times since there are $N/8^2$ groups in this level.

Next, we go to the next higher level—level 3. In this level, each level 3 group contains 448 subscatterers, as it is made up of eight level 2 groups and each level 2 group now contains 56 subscatterers. The solution to the 448 subscatterers together is not known. Hence, we first solve for the scattering solution for each level 3 group with 448 subscatterers. Then their interior subscatterers (152 for each group in this level) are removed via (28) and 296 subscatterers remain in each group of this level with known solution. Again, we need to repeat such process for $N/8^3$ times for all the $N/8^3$ groups in this level.

We see that at each level, upon removal of the interior subscatterers, a group contains fewer subscatterers than what it has originally. The process is continued until the highest level (level L) is reached, where there is only one group which is made up of eight level $L - 1$ groups. Again, the solution must be sought for the group; however, since the interior subscatterers for the eight level $L - 1$ groups are removed, this group contains much less subscatterers than N , the total number of subscatterers for the original problem. As a result, much less operations are expended to find the final solution.

It is seen that the work at each level contains two parts: one to find the scattering solution, the other to remove the interior subscatterers. It is important that the removal of the interior subscatterers of a group does not change the scattering property of the group. The solution of interaction matrices can be found in the same way as described in [1].

Using similar analysis as presented for the 2-D case [1], the computational complexity can be shown to be CN^2 for N small spheres, where C is a constant. In the above, we describe the algorithm using base eight, i.e., a level l group contains eight level $l - 1$ groups for $l = L, L - 1, \dots, 1$. We can also use other bases, however, such as base four and base two. Though the computational complexity, CN^2 , does not depend on the base used, practical implementation shows that base two algorithm is the most efficient one in the sense that it has a smaller constant C (roughly 900). The memory requirement is of $O(N^{4/3})$ in 3-D and of $O(N \log N)$ in 2-D.

Once the solution is found, the computation of the scattered field is of a subdominant cost. This can be easily seen from (27), in which a_n is the incident wave expansion coefficient

and is arbitrary, so that once $\bar{I}_{mm'}(N_B)$ are computed, they can be used to calculate the scattered field for any incident wave at the cost of $O(N_B^2)$ operations, where, $N_B^2 \sim O(N^{4/3})$ is negligible compared to the total computational cost $O(N^2)$.

V. PARALLELIZATION OF THE CODE ON CM-5

The Connection Machine CM-5 of Thinking Machine Corp. at the National Center for Supercomputing Application at the University of Illinois is a massively parallel processor (MPP) machine with 512 processors. A version of NEPAL was first written on a SUN SPARC workstation which was a serial machine. The version of the code did not run efficiently on a massively parallel machine. Much of the code has to be rewritten for an MPP machine.

There are two major processes in implementing NEPAL: they are the interaction matrix algorithm (IMA) process and the equivalence principle process. The IMA consumed most of the CPU time. When NEPAL was first implemented on the CM-5, the code called an external function in the library of CM-5. This function inverted a matrix by Gauss-Jordan method (the name of the function was "gen_gj_invert"). When this function was called in the IMA part, the CPU time for matrix fill was dominant, because the matrix inverse was greatly expedited. This was not the case on a serial machine. Hence, the matrix fill had to be expedited, making it a small portion of the total CPU. An estimated throughput of about 1 GFLOPS on a 64 processor partition of the CM-5 was obtained using this approach.

VI. NUMERICAL RESULTS

Using this algorithm, we have developed a program to compute the scattering solution of a rectangular cubic dielectric scatterer. In our numerical simulation, six spherical modes (three for TE and three for TM) are used to expand the scattered field of a subscatterer. Fig. 4(a) and 4(b) show the scattered field (magnitude and phase, respectively) of a cubic dielectric scatterer with side length $a = b = c = 1.6\lambda$ and $\epsilon_r = 2.6$. The incident wave is coming from $\theta = 180$ degrees and $\phi = 0$ degrees with θ polarization and a frequency of 300 MHz. The scattered field is observed at $\theta = 0$ degrees to $\theta = 180$ degrees and $\phi = 0$ degrees. Only the E_θ component is plotted. The results agree well with that of the brute-force solution using Gaussian elimination.

Fig. 5 shows RCS of a dielectric rod of radius 0.4λ and of length 0.8λ , with $\epsilon_r = 2.05$. The discretization $\Delta = 0.1\lambda$, while the wave is incident from the axis of the cylinder, and the RCS is plotted as a function of the off-axis angle. The result compares well with the brute-force solution using Gaussian elimination.

As for the computational complexity, Fig. 6 shows the comparison of the CPU time (on the CM-5) of this method (NEPAL) with the brute force solution. It is seen that when N , the number of unknowns, is small, NEPAL is not as efficient as brute force. When N is large, however, NEPAL becomes more efficient than brute force. The crossover occurs at about $N = 1800$. It is also seen that the slope of the CPU time

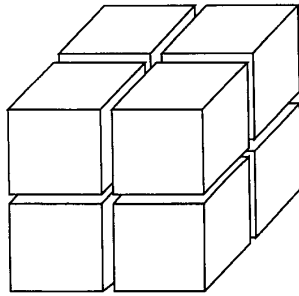
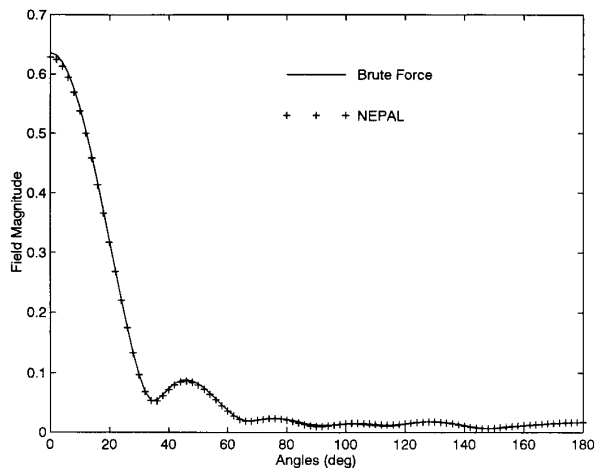
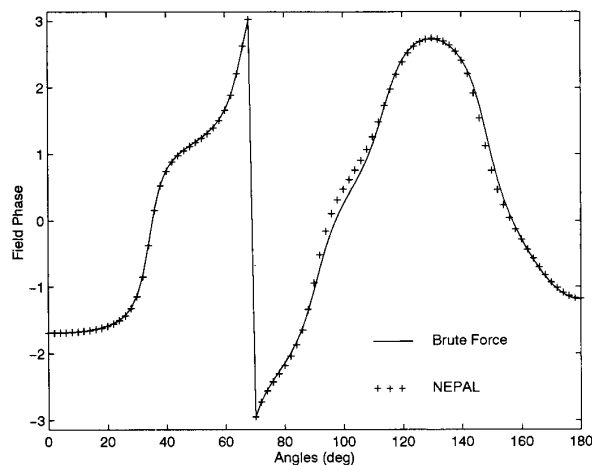


Fig. 3. At the lowest level, each cubic box representing a group contains 64 subscatterers. The scattering solution of each group can be found, and then grouped by groups of two for the next level.



(a)



(b)

Fig. 4. (a) Scattered field (magnitude) of a cubic dielectric scatterer for plane wave incidence at $\theta = 180$ degrees. Side length $= 1.6\lambda$, $\epsilon_r = 2.6$, $f = 300$ (MHz). Field points are on the circle of $r = 10\lambda$. The cube is divided into 512 small subscatterers. (b) Scattered field (phase) of a dielectric cube with the same parameters as in (a).

curve for NEPAL is decreasing and approaches the slope of an N^2 curve.

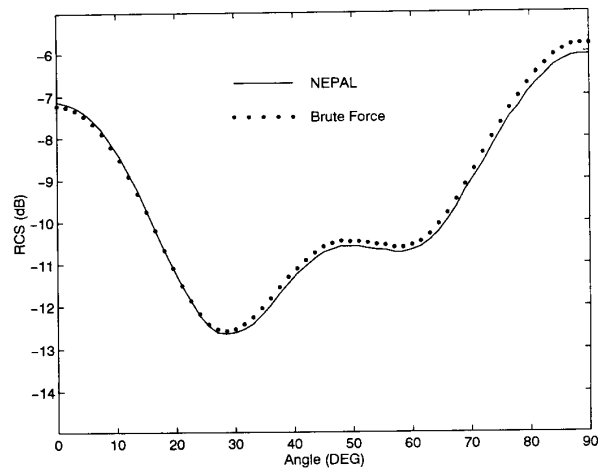


Fig. 5. The RCS of a circular dielectric rod of finite length whose radius is 0.4λ , length is 0.8λ , and with $\epsilon_r = 2.05$. The discretization $\Delta = 0.1\lambda$, the incident wave is from the axis, and the RCS is plotted as a function of off-axis angle. The number of unknowns is 3036.

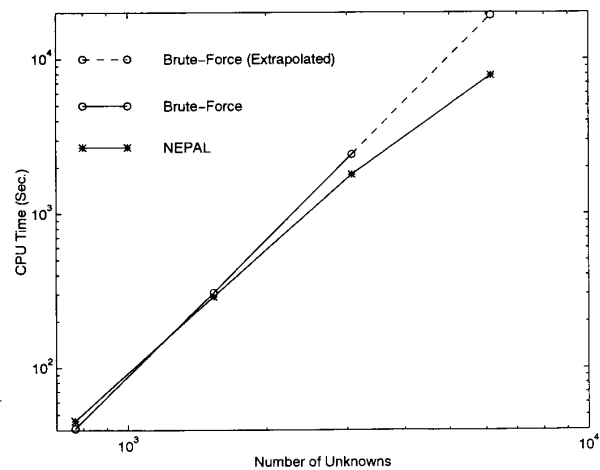


Fig. 6. Comparison of CPU time of NEPAL and Gaussian elimination.

VII. CONCLUSION

We have presented in this paper the extension of the nested equivalence principle algorithm (NEPAL) to three dimensions. The algorithm is based on Huygens' equivalence principle by nesting small algorithms within a larger one. Therefore, the key element is to divide the computation into several stages and reduce the number of unknowns at each stage. This represents an efficient algorithm for directly solving the integral equation of scattering with reduced computational complexity of $O(N^2)$. In other words, this method seeks the inverse of the integral operator in less than $O(N^3)$ operations. Hence, it can be used to compute the scattering solution of large objects for many incident waves.

REFERENCES

- [1] W. C. Chew and C. C. Lu, "The use of Huygens' equivalence principle for solving the volume integral equation of scattering," *IEEE Trans. Antennas Propagat.*, vol. 41, no. 7, pp. 897-904, July 1993.

- [2] J. H. Richmond, "Scattering by a dielectric cylinder of arbitrary cross section shape," *IEEE Trans. Antennas Propagat.*, vol. AP-14, no. 4, pp. 460-464, July 1966.
- [3] B. Peterson and S. Ström, "T-matrix for electromagnetic scattering from an arbitrary number of scatterers and representation of $E(3)$," *Phys. Rev.*, vol. D8, pp. 3661-3678, 1973.
- [4] W. C. Chew, J. Friedrich, and R. Geiger, "A multiple scattering solution for the effective permittivity of a sphere mixture," *IEEE Trans. Geosci. Remote Sensing*, vol. 28, no. 2, p. 207, 1990.
- [5] L. Tsang and C. E. Mandt, "Monte-Carlo simulations of the extinction rate of dense media with randomly distributed dielectric spheres based on solution of Maxwell's equations," *Optics Lett.*, vol. 17, no. 5, pp. 314-316, Mar. 1992.
- [6] Y. M. Wang and W. C. Chew, "A recursive T-matrix approach for the solution of electromagnetic scattering by many spheres," *IEEE Trans. Antennas Propagat.*, vol. 41, no. 12, pp. 1633-1639, Dec. 1993.
- [7] G. P. Otto and W. C. Chew, "Inverse scattering of Hz waves using local shape function imaging: A T-matrix formulation," *Int. J. Imaging Syst. Technol.*, vol. 5, pp. 22-27, 1994.
- [8] L. Gürel and W. C. Chew, "Recursive algorithms for calculating the scattering from N strips or patches," *IEEE Trans. Antennas Propagat.*, vol. 38, no. 4, pp. 507-515, Apr. 1990.
- [9] C. Lanczos, *Applied Analysis*. Englewood Cliffs, NJ: Prentice-Hall, 1956, p. 41.
- [10] A. George, "Nested dissection of a regular finite element mesh," *SIAM J. Numer. Anal.*, vol. 10, pp. 345-363, 1973.
- [11] C. F. Bohren and D. R. Huffman, *Absorption and Scattering of Light by Small Particles*, ch. 4. New York: Wiley, 1983.
- [12] J. D. Jackson, *Classical Electrodynamics*, 2nd ed. New York: Wiley, 1975.
- [13] W. C. Chew, *Waves and Fields in Inhomogeneous Media*. New York: Van Nostrand, 1990.
- [14] W. C. Chew and Y. M. Wang, "Efficient ways to compute the vector addition theorem," *J. Electromag. Waves Appl.*, vol. 7, no. 5, pp. 651-665, 1993.
- [15] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions*. New York: Dover, 1965.



Weng Cho Chew (S-'79-M'80-SM'86-F'93) received the B.S. degree in 1976, both the M.S. and engineer's degree in 1978, and the Ph.D. degree in 1980, all in electrical engineering from the Massachusetts Institute of Technology, Cambridge.

From 1981 to 1985, he was with Schlumberger-Doll Research in Ridgefield, CT, where he was a Program Leader and a Department Manager. From 1985 to 1990, he was a Associate Professor with the University of Illinois. He is currently a Professor there and teaches graduate courses in Waves and

Fields in Inhomogeneous Media, and Theory of Microwave and Optical Waveguides, and supervises a graduate research program. His current research interests are in wave propagation, scattering, inverse scattering, and fast algorithms related to scattering, inhomogeneous media for geophysical subsurface sensing, and nondestructive testing applications. Previously, he has also analyzed electrochemical effects and dielectric properties of composite materials, microwave, and optical waveguides, and microstrip antennas.

Dr. Chew is often listed in the *List of Excellent Instructors* on campus. He has authored a book, *Waves and Fields in Inhomogeneous Media*, published over 110 scientific journal articles, and presented over 130 conference papers. He is a member of Eta Kappa Nu, Tau Beta Pi, URSI Commission B and F, and an active member in the Society of Exploration Geophysics. He was an NSF Presidential Young Investigator for 1986. He was also an AdCom member and is presently an Associate Editor with the International Journal of Image Systems and Technology, and has been a Guest Editor of Radio Science and International Journal of Imaging Systems and Technology. From 1989 to 1993, he was the Associate Director of the Advanced Construction Technology Center at the University of Illinois.



Cai-Cheng Lu received the B.S. and M.S. degrees, both in electrical engineering, from Beijing University of Aeronautics and Astronautics, China, in 1983 and 1986, respectively. Currently, he is pursuing the Ph.D. degree at the University of Illinois at Urbana-Champaign.

From 1986 to 1990, he was with the Department of Electronic Engineering, University of Aeronautics and Astronautics, China. From 1991 to 1992, he was a Visiting Scholar at the Department of Electrical and Computer Engineering, University of Illinois

at Urbana-Champaign. His current research interests are in electromagnetic wave scattering and inverse scattering, antennas, and wave propagation.