Available online at www.sciencedirect.com
SCIENCE DIRECT®
C. R. Physique 6 (2005) 604–617
COMPTES RENDUS PHYSIQUE
http://france.elsevier.com/direct/COMREN/

Interaction of electromagnetic fields with the environment/Interaction du champ électromagnétique avec l'environnement

Fast algorithms for layered media

Weng Cho Chew [a,*], Bin Hu [b], Y.C. Pan [b], L.J. Jiang [c]

[a] *Center for Computational Electromagnetics and Electromagnetics Laboratory, University of Illinois, Urbana-Champaign, IL 61801, USA*
[b] *Intel Corporation, Hillsboro, OR 97214, USA*
[c] *IBM, Yorktown Heights, NY 10598, USA*

Available online 30 August 2005

Abstract

This paper gives an overview of computational electromagnetics for simulating complex structures with a large number of unknowns using integral equation method. First, it discusses the pros and cons of differential equation solvers versus integral equation solvers. Then it gives an overview of different fast algorithms for solving integral equations. It also discusses the difficulty of solving electromagnetic problems in the "twilight zone" between static and full-wave electromagnetics. Finally, it discusses various methods for fast solutions of scattering and coupling problems involving layered media both for statics and dynamic problems, and presents some numerical results. ***To cite this article: W.C. Chew et al., C. R. Physique 6 (2005).***

Résumé

Algorithmes rapides pour des milieux en couches. Cet article présente une vue d'ensemble de l'électromagnétisme en vue du calcul pour la simulation de structures complexes avec un grand nombre d'inconnues à l'aide de méthodes intégrales. Il discute d'abord les avantages et désavantages des codes basés sur la résolution des équations différentielles comparés à ceux sur les équations intégrales. Il donne ensuite une vue d'ensemble des différents algorithmes rapides pour résoudre les équations intégrales. Il discute aussi la difficulté de résoudre les problèmes électromagnétiques dans la « zone d'ombre » entre l'électromagnétisme statique et hautes fréquences. Enfin, il discute les diverses méthodes pour des résolutions rapides de problèmes de diffraction et de couplage impliquant des milieux en couches à la fois pour l'électrostatique et l'électrodynamique, et présente quelques résultats numériques. ***Pour citer cet article : W.C. Chew et al., C. R. Physique 6 (2005).***

Keywords: Computational electromagnetics; Integral equations solvers; Layered media

Mots-clés : Électromagnétisme numérique ; Équations intégrales ; Milieux en couches

1. Introduction

Computational electromagnetics is becoming increasing important due to its potential to replace many laboratory experiments. This essentially stems from the predictive power of the laws of electromagnetism, or Maxwell's equations. Maxwell's equations are valid over a wide range of length-scales, and also valid over different frequencies or time scales. Consequently, when Maxwell's equations can be solved accurately, the result is often used as a reliable method to predict the performance of a design or the outcome of an experiment or a system.

* Corresponding author.
E-mail address: w-chew@uiuc.edu (W.C. Chew).

1631-0705/$ – see front matter
doi:10.1016/j.crhy.2005.06.004

Due to the predictive capability of Maxwell's equations, mathematical modeling has been very important from very early time [1]. Mathematical modeling was used by early civilizations to design buildings. For instance, Pythagoras theorem was known both in the Far East as well as the West from very early times [2]. One can say that pre-calculus, which was used by Kepler to predict celestial events in the 1600s, marked a new beginning for the use of mathematical modeling for prediction. In the late 1600s and early 1700s, Leibniz and Newton invented calculus that even further empowered us to describe natural phenomena [3]. Navier–Stokes equations were developed in the 1800s to describe the physics of fluid with great success [4 and references therein]. In the early 1800s, telegraphy grew to be quite important, and submarine cables were even drawn between America and Europe before the laws of electromagnetism were fully understood [5]. However, it was quite clear that the telegraphy equations predicted wave phenomenon, but the laws of electromagnetism did not then. So it was almost de rigueur that Maxwell's displacement current term be added in 1864 to complete the laws of electromagnetism [6].

After electromagnetic theory was fully formulated, early mathematical modeling was done with closed form solutions involving simple shapes such as spheres and cylinders [7]. However, such closed form analytic solutions for mathematical modeling were insufficient for many queries of science and engineering. As a consequence, approximate solutions were sought to expand the scope of problems that were solvable with analytic methods. These were the asymptotic methods, such as the high-frequency method where a small parameter is assumed [8 and references therein]. The solutions are often ansatz-based, and the errors in the solutions are often not controllable.

The advent of digital computer in the 1940s [9] was mainly driven by our need to solve the fluid equations which are nonlinear, and the only viable solutions are by numerical methods. Thus, numerical methods were developed early for computational fluid dynamics [10 and references therein].

However, Maxwell's equations are pervasively used in many areas of engineering and science, because many phenomena, including light phenomena, are electromagnetic in nature. Hence, there has been an urgent need to solve Maxwell's equations for increasingly more difficult problems, and this often calls for numerical solutions via computational electromagnetics. In the early days, both differential equation solvers and integral equation solvers found their ways into computational electromagnetics [11,12]. Examples of differential equation solvers are the finite difference time domain method [13], and the finite element method [14–16], while examples of integral equation solvers are the method of moments, and the Nyström method [17–19]. Differential equation solvers and integral equation solvers have their pros and cons as follows [1]:

Differential equation solvers:

- Con: Field is the unknown to be sought, and it permeates all of space.
- Con: Number of unknowns scales volumetrically as $(kD)^{3+\alpha}$.
- Con: Grid dispersion error requires higher grid density for larger problems.
- Pro: Simplicity for implementation.
- Pro: Sparse matrix system.

Integral equation solvers:

- Pro: Current is the unknown that often resides only on surfaces.
- Pro: Unknown N scales as$(kD)^2$.
- Pro: Green's function is used as a propagator that reduces propagation error.
- Con: Difficult to implement (has to deal with singular integrals).
- Con: Dense matrix system.

In the above, k is the wavenumber and D is the diameter of the simulation region. In differential equation solvers, due to the grid dispersion error, the grid density often increases with problem size, and hence, the reason for $N \sim (kD)^{3+\alpha}$ in the above [20]. The effect of grid dispersion error can be mitigated with high accuracy, higher-order numerical methods [21]. The exponent α depends on the order of the numerical used. The higher the order of the numerical method (which generally means higher accuracy), the smaller α is.

Integral equation solvers find its great advantage in surface integral equations, where the number of unknowns is generally much smaller than those in differential equation solvers. What is holding back integral equation solvers has been the dense matrix system which is both expensive to solve and store. However, fast solvers can alleviate this problem [1].

2. Overview of fast algorithms

Differential equation solvers are, in general, quite efficient for the number of operations needed to solve the number of unknowns. Because of the sparse nature of the matrix system, an N unknown problem requires $O(N)$ operations for a matrix-

vector product. Therefore, when the matrix-vector product is used in an iterative solver or a time-marching scheme, the cost per iteration or per time step with respect to the unknown count is quite efficient. The drawback is that N can be exorbitantly large.

On the other hand, integral equation solvers come with dense matrix systems. Dense matrix systems can be solved efficiently if they are of Toeplitz nature. This occurs when an FFT-based method is used to solve the problem [22–24]. The matrix storage can then be reduced to $O(N)$ from $O(N^2)$. Moreover, a matrix-vector product can be performed in $O(N \log N)$ instead of $O(N^2)$ operations. Early forms of FFT-base algorithms used rectilinear grids, limiting the versatility of the geometry that can be modeled. However, the introduction of the pre-corrected FFT method and the adaptive integral method lifted this limitation [25,26]. The breakaway from rectilinear grid can also be achieved using interpolation and anterpolation [27].

The FFT-based method is only optimally efficient when applied to problems where the domain in which the FFT has to be performed is approximately the same as the domain where the unknowns reside. This happens in problems where the unknowns are volumetrically distributed or where the unknowns are distributed on the surface of a flat plate. When the unknowns are distributed on the surface of a round or fat scatterer, an FFT cannot be performed efficiently, due to the zero padding that is necessary in order to use the FFT algorithm. Hence, for a fat scatterer where the surface area is proportional to D^2 and the volume is proportional to D^3 where D is the diameter of the scatterer, the cost of performing an FFT is of $O(N_s^{3/2} \log N_s)$ where N_s is the number of surface unknowns. The use of sparse-data FFT does not help either, as fast convolution cannot be performed by such a method [28].

While a fast matrix-vector product involving an integral equation operator cannot be performed rapidly using FFTs when the data are sparsely distributed in space, fast algorithms existed in the astronomy community whereby Coulombic interactions between cluster of stars can be rapidly evaluated [29,30]. This is equivalent to evaluating the action of an integral operator on a cluster of static sources. The algorithm is later systematized by Greengard and Rokhlin [31]. Meanwhile, other fast matrix-vector product algorithm for static field existed using wavelet decomposition [32 and references therein].

However, fast matrix-vector product algorithms for electrodynamic problems have been elusive. This is due to the oscillatory nature of electrodynamic fields, or Helmholtz-type field [33]. Oscillatory field sends information long range, whereas a Laplacian field or a static field does not send information long range. In fact, a static field becomes increasingly smooth the further it is from the source, implying the loss of information or bandwidth. Hence, simple compression scheme by wavelet algorithm works for static fields, but not for oscillatory fields. Moreover, the task of transmitting information from object A to object B using multipole translation becomes increasingly expensive for a highly oscillatory field and large objects.

The translation of the field from point to point is intimately related to the translation operator which forms the translation group [34]. This group is Abelian, and hence, a diagonal representation of such a group exists when plane waves are used as the basis rather than multipoles when seeking the matrix representation of such a translator. Under the plane wave basis, the translators are diagonal [1,33]. Strictly speaking, the 'multipole algorithm' is a misnomer when applied to electrodynamic problems.

Therefore, a suitable representation of the field from a group of sources is in terms of plane waves radiating from them in all direction: a function that represents the strength of the plane wave radiation field in different directions is called the radiation field pattern of the source. Hence, the information of the radiation field from a group of sources is captured in its radiation pattern. When the observation point is outside the evanescent field of the sources, the outgoing plane waves can be converted into incoming plane waves by the use of an outgoing wave to incoming wave (O2I) translator. Such a translator is only efficient for electrodynamic problems if it is diagonal, because the information content of the field grows with the size of the group sources.

A suggestion for such an algorithm was first given by Rokhlin [35]. However, unlike what was suggested, a simple nesting of such an algorithm into a multilevel (tree) algorithm does not yield an $O(N \log N)$ algorithm for a matrix-vector product for dynamic problems. In a multilevel algorithm, the group sizes of the sources are different at different levels [1]. Therefore, the information content of the field enriches as the group sizes become large. While the radiation pattern for larger group sizes are sampled at a higher sampling rate, those for smaller group sizes can be sampled at a lower sampling rate. Hence, the vector arrays that are used to store the information contents of the radiation patterns for different group sizes are different—with smaller arrays near the leafy branches of the tree algorithm and larger arrays near the root of the tree. Since the groups near the leafy branches are numerous, and the groups near the root are scarce, the use of different array sizes greatly reduces the storage and workload at the near-leafy level of the algorithm. However, in order to enable the use of different array sizes at different levels, the radiation patterns have to be interpolated and anterpolated (transpose interpolated) between different levels. This idea is akin to the idea of multilevel multigrid where such operations are used to reduce workloads at coarser levels [36]. As a consequence, a matrix-vector product can be effected in $O(N \log N)$ operations for dynamic problems [37,38].

3. The twilight zone

Another challenge region in fast algorithm is the 'twilight zone' where a field is not truly static nor truly dynamic, viz., the domain of the quasistatic field [39–41]. When the operating frequencies of integrated circuits (IC) were low, most electrical phenomena in a computer chip can be described using circuit theory. However, the increasing operating frequency of IC chips calls for powerful and efficient full-wave electromagnetic simulation tools in this 'twilight zone'. In this regime, static solvers are inadequate, and the non-zero frequency effect of the field cannot be ignored. Consequently, broadband fast solvers from static to microwave range are needed that will capture the wave physics as well as the circuit physics inherent in the electrical design of computer chips. Due to the increasing complexity of computer chip design, it is imperative that these fast solvers handle problems with millions of unknowns, representing orders of magnitude in size capability over traditional solvers. While this has been possible in the electrodynamic regime, or the electrostatic regime, it has been elusive in the 'twilight zone' between static and radio frequencies [42].

To understand this, we look at Maxwell's equation more carefully in this regime:

$$\nabla \times \mathbf{E} = 0, \qquad \nabla \cdot \varepsilon \mathbf{E} = \rho = \lim_{\omega \to 0} \nabla \cdot \mathbf{J}/\mathrm{i}\omega \tag{1}$$

$$\nabla \times \mathbf{H} = \mathbf{J}, \qquad \nabla \cdot \mu \mathbf{H} = 0 \tag{2}$$

This physics in this regime can be gleaned from Eqs. (1) and (2) when the frequency is zero. The electric field is decoupled from the magnetic field. The current associated with (2) is divergence free, while that with (1) is non-divergence free. Hence, the current decomposes itself into a curl free (irrotational) component and a divergence free (solenoidal) component. The electric field represents the world of capacitors, while the magnetic field represents the world of inductors. Though uncoupled at static, they are weakly coupled in the quasi-static regime. Consequently, we have

$$\mathbf{J} = \mathbf{J}_{\mathrm{sol}} + \mathbf{J}_{\mathrm{irr}} \tag{3}$$

This decomposition is known as the Helmholtz decomposition [43]. Moreover, by comparing (1.1) and (1.2), we see that

$$\mathbf{J}_{\mathrm{irr}} \sim \mathrm{O}(\omega), \quad \mathbf{J}_{\mathrm{sol}} \gg \mathbf{J}_{\mathrm{irr}}, \quad \omega \to 0 \tag{4}$$

Any numerical code that fails to capture the aforementioned physics correctly will break down. This is because the irrotational current that models the capacitors is much smaller than the solenoidal current, but both currents are equally important. Since this instability is inherently due to the disparate amplitude of the irrotational current and the solenoidal current, the use of second kind integral equations does not remedy this problem [44–46]. To remedy this instability, one needs to decompose the current into irrotational and solenoidal currents, partition the matrix, and perform frequency normalization in order to stabilize the integral equation [47,39].

Alternatively, one can look at the vector wave equation

$$\nabla \times \nabla \times \mathbf{E} - k^2 \mathbf{E} = \mathrm{i}k\mathbf{J}/\eta \tag{5}$$

When $k \to 0$, a perturbation expansion in small k, $\mathbf{E} \sim \mathbf{E}^{(0)} + k\mathbf{E}^{(1)} + \cdots$ and a similar one for $\mathbf{J}$ can be substituted into the above, and the leading terms satisfy the equations

$$\nabla \times \nabla \times \mathbf{E}^{(0)} = 0, \qquad \nabla \times \nabla \times \mathbf{E}^{(1)} = (\mathrm{i}k/\eta)\mathbf{J}^{(0)} \tag{6}$$

which have no unique solution, since the curl operator has a null space. In other words, the zero frequency point is a singular perturbation point [48].

Another potential hazard of the 'twilight zone' is the breakdown of acceleration techniques. The factorizations used in the fast multipole method and the multilevel fast multipole algorithm designed for dynamic problems breakdown when $k \to 0$. In other words, the error becomes larger for lower frequencies [49,50]. The factorization for oscillatory Green's function is based on plane wave radiation to the far field, assuming that evanescent waves are unimportant. When the frequency goes to zero, there is no plane wave to speak of regarding the field emitted by a source, and evanescent field dominates the total field. This morphing of the physics is essentially the reason for the breakdown of the factorization. Hence, when evanescent waves dominate, it is more expedient to represent the field in terms of multipole expansions, as has been done in many electromagnetic scattering problems [51]. Fortunately, multipole fields in the quasi-static regime have short range interaction, and hence, do not increase the complexity of the algorithm. In fact, in the limit when $k \to 0$, for surface distributed sources in 3D, a matrix vector product can be effected in $\mathrm{O}(N)$ operations, with similar complexity for memory requirement.

Yet another hazard exists for the acceleration technique in the 'twilight zone' using multipole expansions. The factors of the factorized oscillatory Green's function, $\exp(\mathrm{i}kr)/(4\pi r)$, does not transition smoothly to the factors of the factorized static Green's function, $1/(4\pi r)$. One can see this in the 2D expression for the translational addition theorem:

$$H_0^{(1)}\big(k|\boldsymbol{\rho} - \boldsymbol{\rho}'|\big) = \sum_{n=0}^{\infty} H_n^{(1)}(k\rho) J_n(k\rho') \mathrm{e}^{\mathrm{i}n(\phi - \phi')} \tag{7}$$

when $\rho > \rho'$. In the above, when $k \to 0$, the left-hand side becomes infinitely large, while on the right-hand side, the Hankel function $H_n^{(1)}(k\rho) \sim \mathrm{O}((k\rho)^{-n})$ (except for the $n = 0$ term), and the Bessel function $J_n(k\rho') \sim \mathrm{O}((k\rho')^n)$. So when the right-hand side is expressed as the inner product between two vectors, the elements of the vectors can overflow and underflow, giving rise to numerical instability. This instability exists also when the Green's function is factorized into a product of a vector times a matrix times a vector.

The source of this 'problem' arises from the fact that Helmholtz equation has a 'yardstick' which is the wavelength. This yardstick remains for all wavelengths, and the Green's function and its factorization can be written in terms of a dimensionless parameter $k\rho$. When k is identically zero, the corresponding Laplace's equation is scale invariant, this yardstick disappears and the solution and its factorization cannot be written in terms of this dimensionless parameter anymore. This instability can be alleviated by defining normalized Hankel and Bessel functions [52,53].

In modeling multi-scale problems using the acceleration method, the geometry in the leafy levels may be in the quasistatic regime, while the geometry on the whole may be in the wave regime. In this manner, it is necessary to factorize the Green's function using a mixed-form fast multipole algorithm, where the source fields are represented in terms of multipoles near the leafy levels, and in terms of plane waves above the leafy levels [54]. A transformer is needed to bridge between the dichotomous nature of the field representation.

$$\begin{aligned}\big[\alpha_{LL'}(\mathbf{r}_{ji})\big]_{L\times L'} = {} & \big[\beta_{LL_1}(\mathbf{r}_{jJ_1})\big]_{L\times L_1}\cdot\big[\beta_{L_1L_2}(\mathbf{r}_{J_1J_2})\big]_{L_1\times L_2}\cdot\big[\beta_{L_2L_3}(\mathbf{r}_{J_2J_3})\big]_{L_2\times L_3}\cdot[D]^{\dagger}_{S_4\times L_3}\\ & \cdot\operatorname{diag}\big[\mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\mathbf{r}_{J_3J_4}}\big]_{S_4\times S_4}\cdot[I]^{T}_{S_5\times S_4}\cdot\operatorname{diag}\big[\mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\mathbf{r}_{J_4J_5}}\big]_{S_5\times S_5}\\ & \cdot\operatorname{diag}\big[\widetilde{T}(\Omega_{S_5},\mathbf{r}_{J_5I_5})\omega_{S_5}\big]_{S_5\times S_5}\cdot\operatorname{diag}\big[\mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\mathbf{r}_{I_5I_4}}\big]_{S_5\times S_5}\cdot[I]_{S_5\times S_4}\cdot\operatorname{diag}\big[\mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\mathbf{r}_{I_4I_3}}\big]_{S_4\times S_4}\\ & \cdot[D]_{S_4\times L_3}\cdot\big[\beta_{L_3L_2}(\mathbf{r}_{I_3I_2})\big]_{L_3\times L_2}\cdot\big[\beta_{L_2L_1}(\mathbf{r}_{I_2I_1})\big]_{L_2\times L_1}\cdot\big[\beta_{L_1L'}(\mathbf{r}_{I_1i})\big]_{L_1\times L'}\end{aligned}\tag{8}$$

The above shows the factorization of a translator into factors that use multipole as well as plane waves for representations. The diagonal factors follow from plane wave basis, while the non-diagonal factors follow from the multipole basis. D is a transformer matrix that transforms back and forth between the two bases. I is an interpolation matrix between different levels.

4. Fast algorithms for layered media

While much advances have been made in designing fast algorithms for the homogeneous medium Green's function, designing them for a layered medium Green's function is a challenge [55–59]. A prime reason is the complex nature of the layered medium Green's function: a layered medium Green's function is expressible in terms Sommerfeld integrals, which usually do not have closed forms [60]. Fortunately, these integrals can be written in terms of plane wave expansions, making them particularly suitable for their factorization. As mentioned before, plane wave basis diagonalizes the translator. A layered medium Green's function can be written as

$$g(\mathbf{r}_j,\mathbf{r}_i) = \int \mathrm{d}\hat{k}\, W(\hat{k})\mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\mathbf{r}_{ji}}\tag{9}$$

where the integral summation is over plane waves as well as inhomogeneous waves and $\mathbf{r}_{ji} = \mathbf{r}_j - \mathbf{r}_i$. For 2D problem, the above is a single integral while it is a double integral for 3D problems. Hence, the factor $\mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\mathbf{r}_{ji}}$ involves both kinds of waves. By letting $\mathbf{r}_{ji} = \mathbf{r}_{jC_J} + \mathbf{r}_{C_JC_I} + \mathbf{r}_{C_Ii}$ as shown in Fig. 1, the above integral can be factorized easily:

$$g(\mathbf{r}_j,\mathbf{r}_i) = \int \mathrm{d}\hat{k}\, W(\hat{k})\mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\mathbf{r}_{jC_J}}\cdot\mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\mathbf{r}_{C_JC_I}}\cdot\mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\mathbf{r}_{C_Ii}}\tag{10}$$

Here, $\mathbf{r}_{C_I}, \mathbf{r}_{C_J}$ are the centers of group I and group J respectively.

In the above, $W(\hat{k})\mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\mathbf{r}_{ji}}$ is the integrand, and we assume that $W(\hat{k})$ is slowing varying, while $\mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\mathbf{r}_{ji}}$ is rapidly varying because $\mathbf{r}_{ji}$ is large. Hence, there is a steepest descent path (SDP) determined by $\mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\mathbf{r}_{ji}}$ as well as a steepest ascent path (SAP) as shown in Fig. 2. A common integration path should avoid crossing the SAP of the pair-wise interaction between the source and the observation point, as the integrand becomes exponentially large on the SAP, a potential source of numerical instability.

Fig. 3 shows a viable numerical integration path that consists of the sum of Paths I, II, and III. By choosing the length of Path II, the path need not cross any of the SAPs for all the pair-wise interactions for all the sources between the two groups. Moreover, the integrands of all the pair-wise interactions will decay along I and III. Consequently, after the use of a quadrature rule in (10), we have (see Fig. 4)

$$g(\mathbf{r}_j,\mathbf{r}_i) = \sum_q \underbrace{\mathrm{e}^{\mathrm{i}\mathbf{k}(\Omega_q)\cdot\mathbf{r}_{jC_J}}}_{\beta_{jC_J}(\Omega_q)}\underbrace{w_q\mathrm{e}^{\mathrm{i}\mathbf{k}(\Omega_q)\cdot\mathbf{r}_{C_JC_I}}}_{\alpha_{C_JC_I}(\Omega_q)}\underbrace{\mathrm{e}^{\mathrm{i}\mathbf{k}(\Omega_q)\cdot\mathbf{r}_{C_Ii}}}_{\beta_{C_Ii}(\Omega_q)}\tag{11}$$

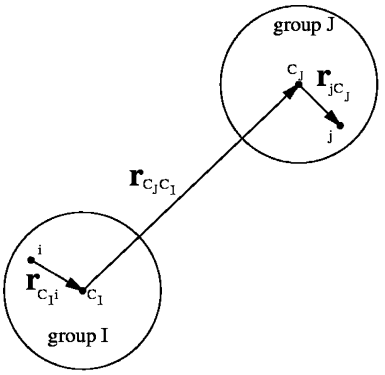


Fig. 1. Translation of a field from point i to point j can be decomposed into a product of three translations.

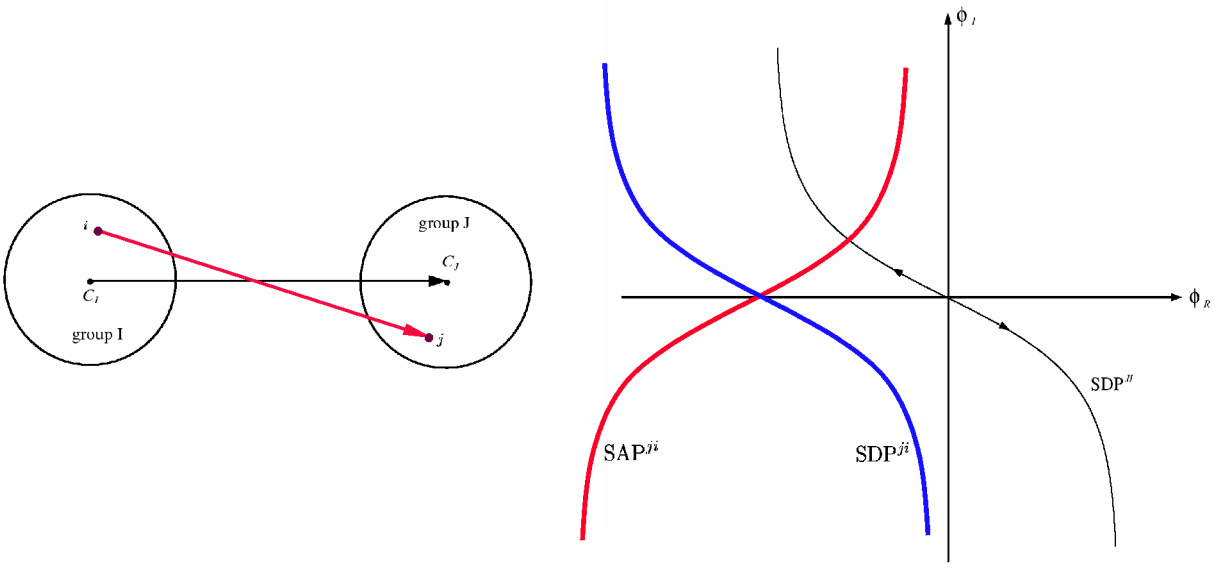


Fig. 2. Different source and observation points will have a different steepest descent path (SDP) and a steepest ascent path (SAP).

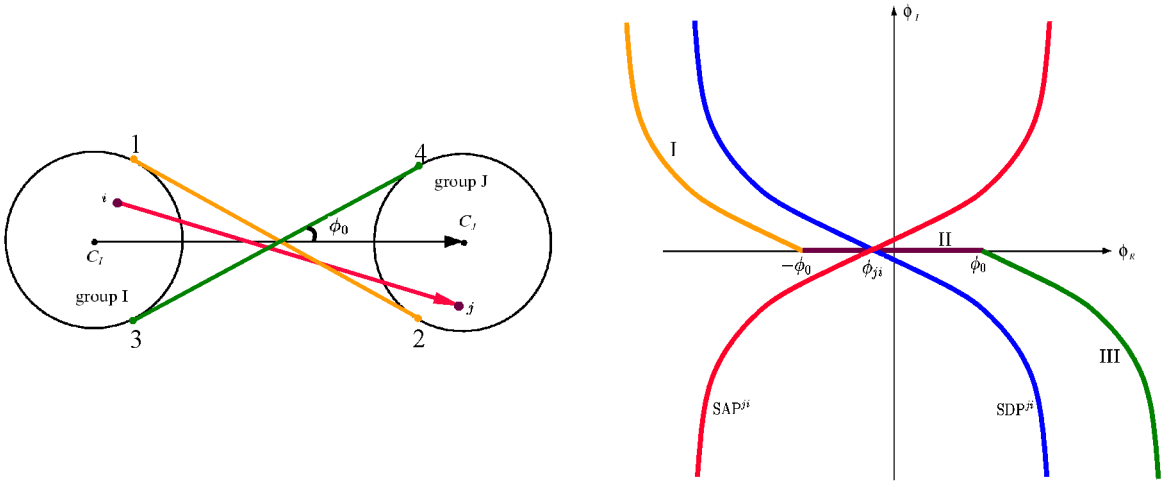


Fig. 3. A viable integration path that consist of path I + II + III for the Sommerfeld integration that crosses no SAPs.

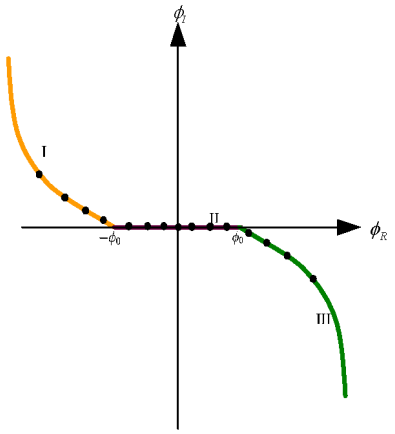


Fig. 4. Sample quadrature points on the integration path.

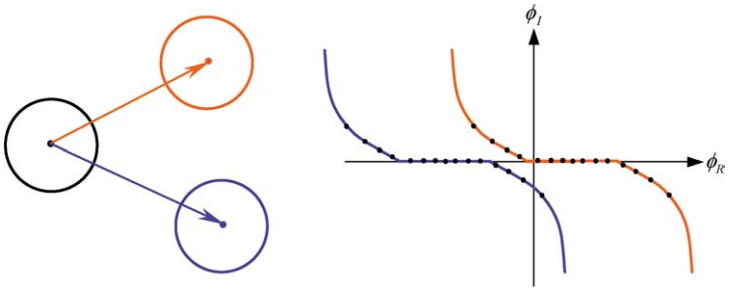


Fig. 5. Different sample quadrature points on the complex planes are needed for computing different group interactions.

Fig. 5 shows that for different group interactions, the integration paths translate horizontally.

In the multilevel algorithm, only the radiation patterns of sources and the aggregated radiation patterns of sources are stored and reused. The same applies to the receiving patterns during the disaggregation procedure. Therefore, in order to reduce memory storage, it desirable to store only the radiation and receiving patterns on the real axis. Therefore, we express the values of the patterns on the complex plane in terms of the value on the real axis as follows:

$$\beta(\Omega_q) = \sum_s I(\Omega_q, \Omega_s)\beta(\Omega_s) \tag{12}$$

where $I(\Omega_q, \Omega_s)$ is an interpolator that can be developed from an interpolation function. Using the above in Eq. (11), we have

$$g(\mathbf{r}_j, \mathbf{r}_i) = \sum_s \beta_{jC_J}(\Omega_s) \cdot \beta_{C_I i}(\Omega_s) \cdot \left(\sum_q \mathbf{I}(\Omega_q, \Omega_s) \cdot \alpha_{C_J C_I}(\Omega_q) \right) = \sum_s \beta_{jC_J}(\Omega_s) \cdot \mathcal{T}_{C_J C_I}(\Omega_s) \cdot \beta_{C_I i}(\Omega_s) \tag{13}$$

Hence, points on the real axis can be reused for different group interaction. The above represents the factorization of the Green's function using the fast inhomogeneous plane wave algorithm (FIPWA). Because only one summation is involved, the translator in the above is diagonal. Error analysis of such a factorization has also been performed in [61,62].

When a point source is placed on top of a layered medium in a 3D space, the Green's function can be decomposed into a direct term and the reflected wave term. The direct term can be treated with the conventional fast method, while the reflected wave term can be treated with the aforementioned method.

$$G(\mathbf{r}_j, \mathbf{r}_i) = G^d(\mathbf{r}_j, \mathbf{r}_i) + G^r(\mathbf{r}_j, \mathbf{r}_i) \tag{14}$$

where

$$\begin{aligned} G^d(\mathbf{r}_j, \mathbf{r}_i) &= \int_{\text{SIP}} \mathrm{d}k_\rho \, \frac{k_\rho}{k_z} H_0\big(k_\rho |\boldsymbol{\rho}_j - \boldsymbol{\rho}_i|\big) \mathrm{e}^{\mathrm{i}k_z |z_j - z_i|} \\ G^r(\mathbf{r}_j, \mathbf{r}_i) &= \int_{\text{SIP}} \mathrm{d}k_\rho \, W(k_\rho) H_0\big(k_\rho |\boldsymbol{\rho}_j - \boldsymbol{\rho}_i|\big) \mathrm{e}^{\mathrm{i}k_z |z_j - z_{iI}|} \end{aligned} \tag{15}$$

where SIP stands for the Sommerfeld integration path, $W(k_\rho)$ is a function that contains the effect of the wave reflection from the layered medium. Furthermore, $\mathbf{r}_{iI} = \boldsymbol{\rho}_i + \hat{z} z_{iI}$ represents the image location of the source point $\mathbf{r}_i$.

The contribution from the reflected wave term can be thought of as coming from an effective image source located at the optical image location, but it sends field to the field point via a complex Green's function given by the second term of (15). This is an image theorem but not in the conventional sense (see Fig. 6).

The interaction of the image source with the field point can be computed similar to the previous discussion, but the SDP path will be decided by the image source point and field point. Moreover, contributions from branch point and poles are possible when the integration path is deformed from SIP to the SDP (see Fig. 7, [60]). Consequently, the reflected wave term can be further decomposed into three contributions, viz.,

$$G^r(\mathbf{r}_j, \mathbf{r}_i) = G^{r\text{SDP}}(\mathbf{r}_j, \mathbf{r}_i) + G^{rP}(\mathbf{r}_j, \mathbf{r}_i) + G^{rB}(\mathbf{r}_j, \mathbf{r}_i) \tag{16}$$

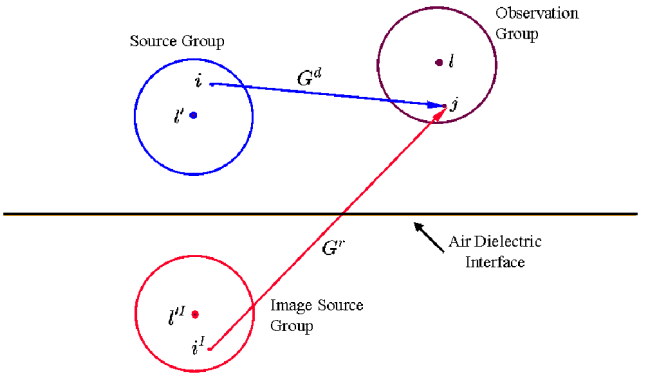


Fig. 6. Effective image theorem for FIPWA.

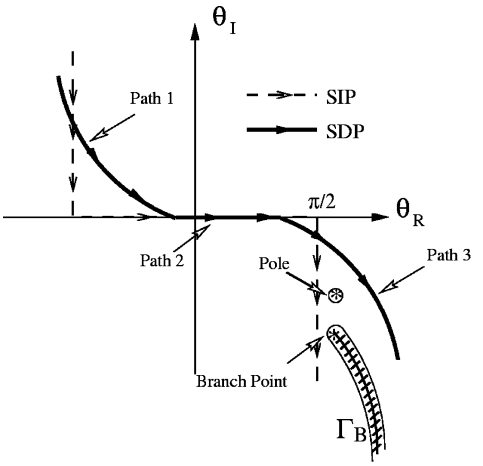


Fig. 7. The integration path along SDP, plus possible pole and branch point contributions when the integration path is deformed from SIP to SDP.

In the reflected wave term, the Hankel function can be factorized and diagonalized by 2D FIPWA or the conventional FMM (fast multipole method). Consequently, each of the terms in the above can be factorized and diagonalized.

Consequently, the layered medium Green's function can be decomposed into the sum of contributions from the direct wave, the reflected wave that consists of the SDP, the pole, and the branch-point contributions.

$$G(\mathbf{r}_j,\mathbf{r}_i)=\sum_s \beta_{jl}(\Omega_s)\cdot\mathcal{T}_{ll'}(\Omega_s)\cdot\beta_{l'i}(\Omega_s)+\sum_s \beta_{jl}(\Omega_s)\cdot\mathcal{T}^{r\,\mathrm{SDP}}_{ll'I}(\Omega_s)\cdot\beta_{l'Ii^I}(\Omega_s)$$
$$+\sum_P \beta_{jl}(\Omega_P)\cdot\mathcal{T}^{r\,P}_{ll'I}(\Omega_P)\cdot\beta_{l'Ii^I}(\Omega_P)+\sum_n \beta_{jl}(\Omega_n)\cdot\mathcal{T}^{r\,B}_{ll'I}(\Omega_n)\cdot\beta_{l'Ii^I}(\Omega_n) \tag{17}$$

The algorithm for a buried object in a half space proceeds in a similar manner. The complication arises when the object is embedded in between layered media. The expression will involve four image terms corresponding to the reflected fields rather than just one in (14).

5. Objects embedded in layered media

The case of when the object is embedded in a layered medium is a more difficult case. In this case, the object will have four effective image terms as oppose to one in the aforementioned discussion [60]. Hence, the book-keeping is rather involved. Even though the general dynamic case has not been fully worked out, we have developed fast algorithms for this case for static fields [63]. The translator will similarly contain five terms: the self term plus four effective image terms:

$$\alpha^{jk}_{nm}(\mathbf{r}_D)=\tilde{\Lambda}_O+\tilde{\Lambda}_{UU}+\tilde{\Lambda}_{UL}+\tilde{\Lambda}_{LU}+\tilde{\Lambda}_{LL} \tag{18}$$

According to 'optical' image theorem, there should be infinitely many image terms. However, the images can be summed into four main groups above.

When inter-layer translations are involved, there are only four terms to consider. These cases are shown in Fig. 8. A typical translator can be expressed in terms of Sommerfeld-type integrals as:

$$\tilde{\Lambda}_L(\mathbf{r},\mathbf{r}';n,m,j,k)=\frac{i^{3j+n-|m-k|}}{\pi}A^{jk}_{nm}\mathrm{e}^{\mathrm{i}(m-k)\phi}+\left[I_{L1}+(-1)^{j+n-m+k}I_{L2}\right]$$
$$I_{L1}=\int_0^\infty \mathrm{d}\lambda\,\lambda^{j+n}K_{m-k}(\lambda\rho)\mathrm{e}^{\mathrm{i}\lambda Z}\Omega(-\lambda)\qquad I_{L2}=\int_0^\infty \mathrm{d}\lambda\,\lambda^{j+n}K_{m-k}(\lambda\rho)\mathrm{e}^{-\mathrm{i}\lambda Z}\Omega(\lambda) \tag{19}$$

where Ω is a function that contains information on the layered medium [64]. Alternatively, they can be expressed as an infinite sum of images, which is not efficient for computation compared to the above. On the other hand, the discrete complex image method (DCIM) can be used to expedite the evaluation of the above expression, greatly expediting their evaluations [65].

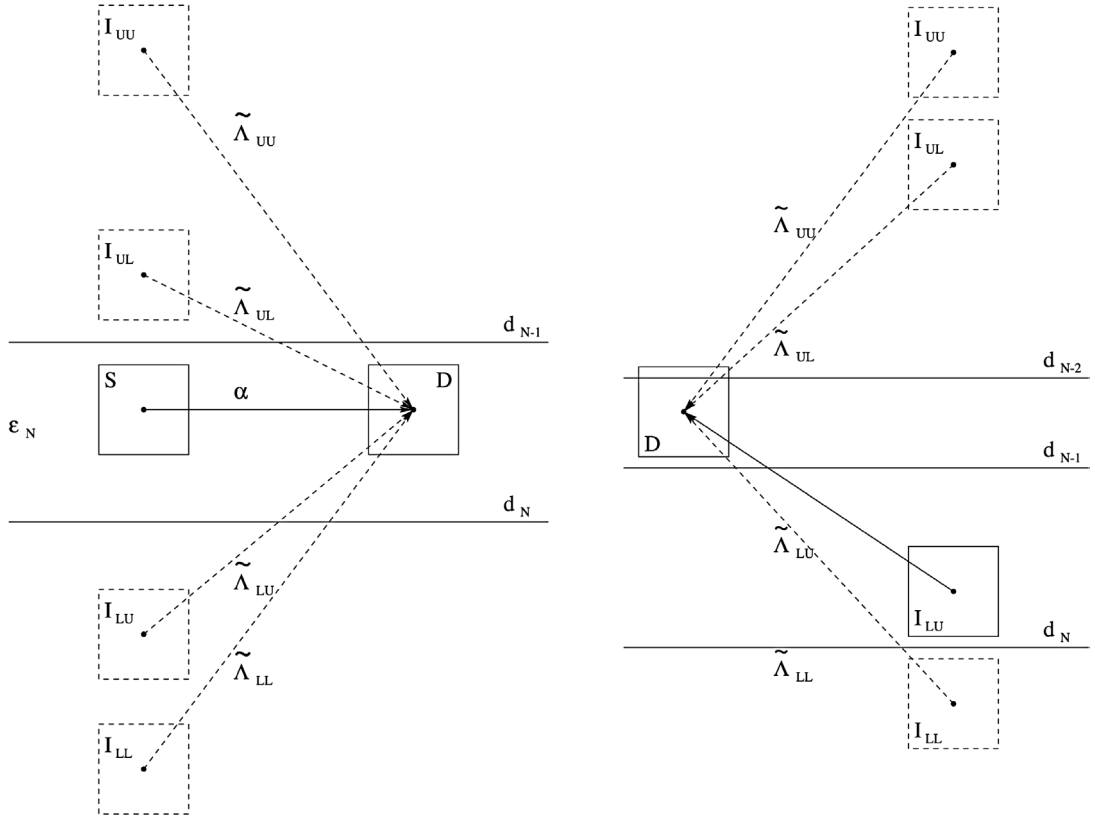


Fig. 8. Schematic for intra-layer translation (left) and inter-layer translation (right).

6. Numerical results

In this section, we show some numerical results from layered medium simulations. Fig. 9 shows the simulation result of the scattering of a buried land mine using multilevel FIPWA consisting of three levels. The land mine is simulated as a PEC with

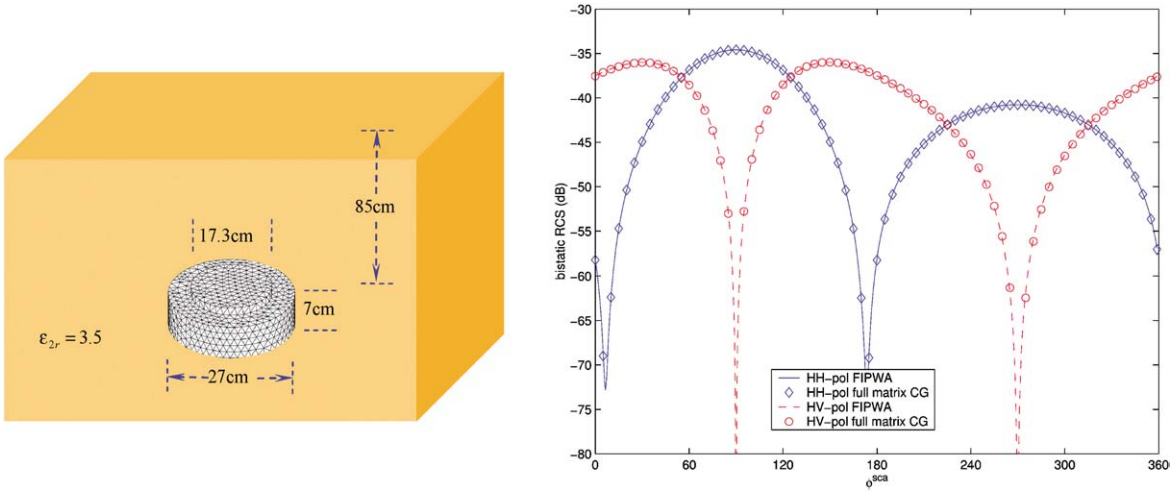


Fig. 9. The simulation of a buried land mine using three-level ML-FIPWA with $(\theta^{\text{inc}}, \phi^{\text{inc}}) = (60°, -90°)$. The frequency is 600 MHz with 2646 unknowns.

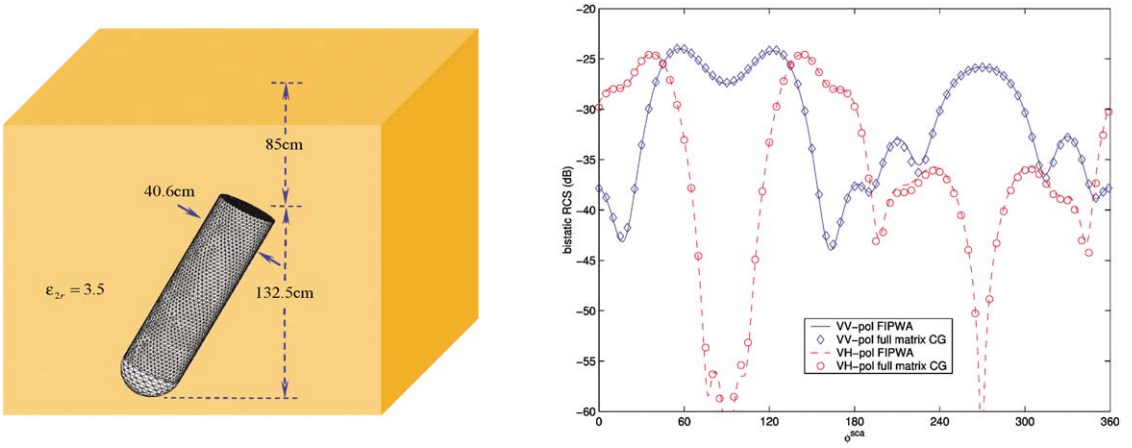


Fig. 10. The simulation of a buried unexploded ordnance using five-level ML-FIPWA with$(\theta^{inc}, \phi^{inc}) = (60°, -90°)$. The frequency is 600 MHz with 8532 unknowns.

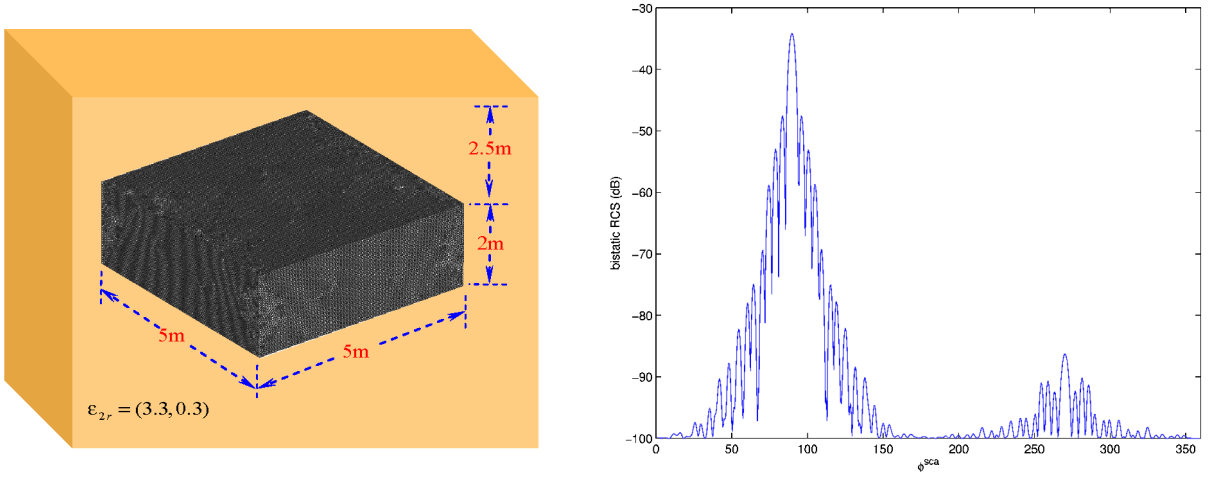


Fig. 11. Simulation of a buried underground bunker using ML-FIPWA. The incident angles are the same as before, and the frequency is 900 MHz. The number of unknowns used is 1 074 588 and eight levels are used in ML-FIPWA.

dimensions shown, buried in a lossless half-space with a relative permittivity of 3.5. The direct calculations are done by solving the MOM matrix equation without using acceleration method. Good agreement with direct calculations is observed.

Fig. 10 shows the simulation of scattering from a buried unexploded ordnance in the same lossless ground, using 8532 unknowns, and five levels of ML-FIPWA. Again the agreement with direct calculations is excellent.

Fig. 11 shows the bistatic RCS (VV-pol) scattering solution of a buried bunker in a lossy ground with $\varepsilon_r = (3.3, 0.3)$, requiring 1 074 588 unknowns for its simulation at 900 MHz. The incident angles are $(\theta^{inc}, \phi^{inc}) = (60°, -90°)$, the same as before. Because of the large number of unknowns, eight levels of ML-FIPWA are required. The calculation took 18 hours of CPU time on one SGI R10000/195-MHz processor in the Origin 2000 array at NCSA, using 3.94 GB of memory. If conventional matrix-based algorithm is used, 9.3 TB of memory would be required, and 11 years of CPU would be needed to solve this problem. The data shows strong specular scattering, with smaller backscattering. The backscattering would have been absent if the buried bunker is absent.

Fig. 12 shows the comparison of the memory requirements and CPU usage for a matrix-vector product of ML-FIPWA for the buried object case and the layered medium case. The additional resource for the buried object case is marginally more than the free-space case in both the memory requirements and CPU usage. The comparison simulation was done on a 600 MHz DEC Alpha 21 164 workstation. The reason that only marginally more resources are needed is because there is hardly any near-field interaction between the image sources and the field point.

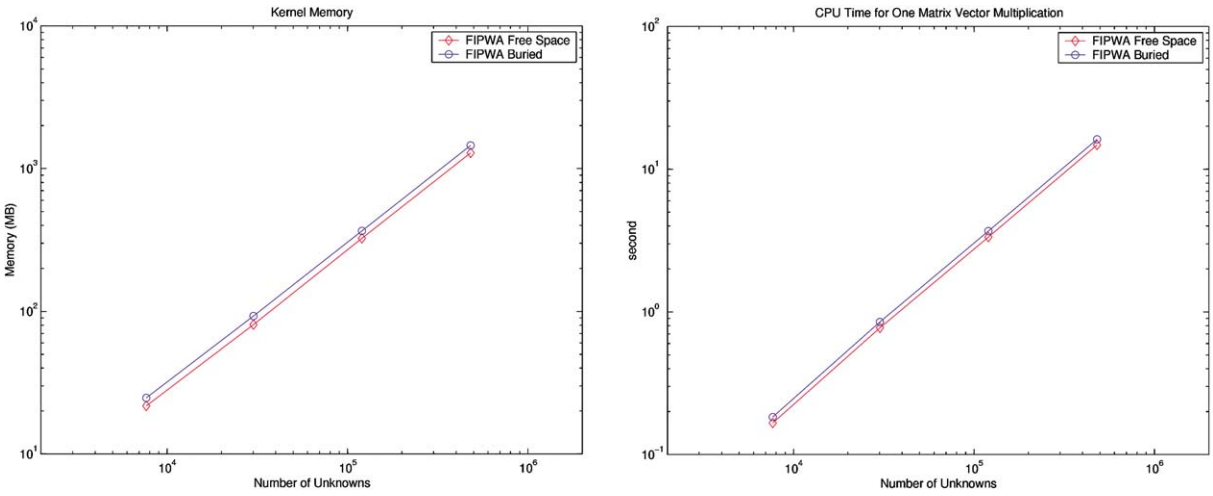


Fig. 12. The comparison of the memory requirements and CPU usage with the free-space case. The memory requirement for the buried object case is 13% more than the free-space case, and the CPU usage is about 10% more.

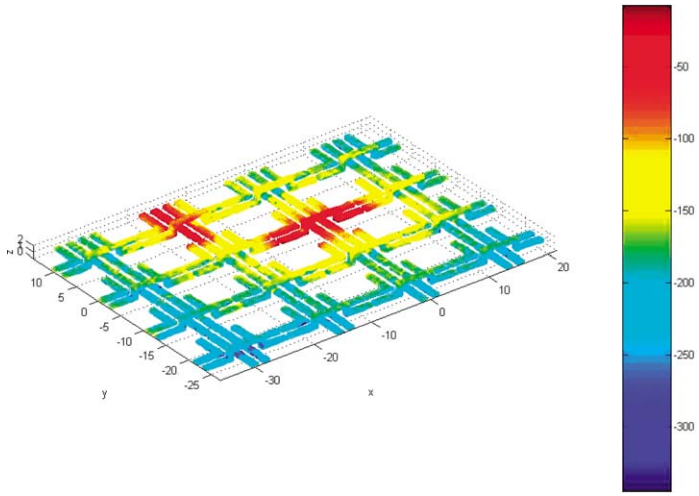


Fig. 13. The use of stratified medium fast multipole algorithm allows a problem of 1 204 352 to be solvable.

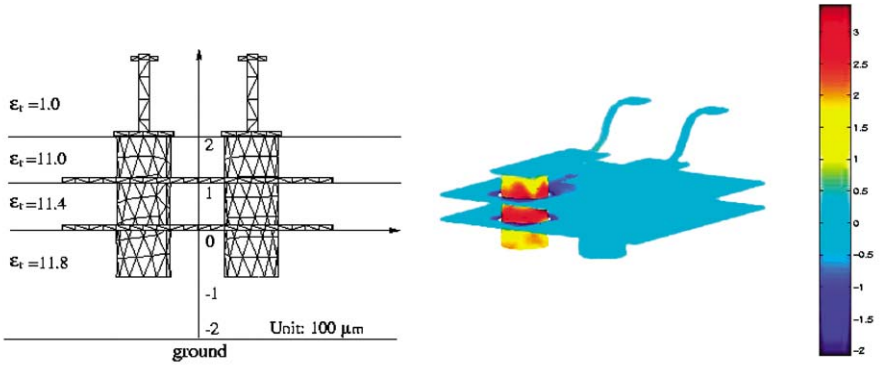
Fig. 13 shows the use of the stratified medium fast multipole algorithm (FMA) to solve a problem with over 1 million unknowns using about 1.41 GB of memory on an old SGI Power Challenger. The computation time was 991.1 s per iteration. This is only marginally more expensive than the free space case that took 966.8 s per iteration using 1.39 GB of memory.

Fig. 14 shows the use of DCIM to expedite the setup times of both the FMA part as well as the MOM part of the code. Speedup factors of more than 40 are possible, without severe degradation of the result.

7. Conclusions

We have given an overview of fast algorithm research in a complex environment. This topic is becoming exceedingly important because many real world applications require the simulation of electromagnetic field in such environment. A difficult regime is designing fast algorithms and solvers for the 'twilight' zone where wave theory meets static field theory. This area is very important for the simulation of electrical signals inside computer circuits and computer chips, as well as nano-structures.

The solution of a scattering and coupling problem involving layered media involves another level of complexity and bookkeeping because of the complex nature of the Green's function involved. We have shown that general fast algorithms can be developed for Laplace's equation for layered media. For full-wave problems, we have developed fast algorithms for layered



	FMA Setup CPU (sec)	MoM Setup CPU (sec)
Old Method	15104.2	1366.49
New Method	342.712	31.496
Speed Up	44.07	43.39

Fig. 14. The use of discrete complex image method (DCIM) can greatly expedite both the setup time for FMA and MOM.

media when the object is above the layered media or below the layered media, but not for objects embedded in a layered medium. Much book-keeping is expected of such work, as it quickly becomes very complex.

References

[1] W.C. Chew, J. Jin, E. Michielssen, J.M. Song (Eds.), Fast and Efficient Algorithms in Computational Electromagnetics, Artech House, Norwood, MA, July 2001.
[2] J. Needham, Science and Civilization in China, Cambridge Univ. Press, Cambridge, UK, 2005.
[3] R. Cooke, The History of Mathematics: A Brief Course, Wiley Interscience, New York, 1997.
[4] A.R. Choudhuri, The Physics of Fluids and Plasma, Cambridge Univ. Press, Cambridge, UK, 1998.
[5] G. Cookson, The Cable: The Wire That Changed the World, Tempus Publishing, Ltd., 2003.
[6] J.C. Maxwell, A Treatise of Electricity and Magnetism, 2 vols., Clarendon Press, Oxford, 1873;
Also, see in: P.M. Harman (Ed.), The Scientific Letters and Papers of James Clerk Maxwell, vol. II, 1862–1873, Cambridge Univ. Press, Cambridge, UK, 1995.
[7] J.J. Bowman, T.B.A. Senior, P.L.E. Uslenghi, Electromagnetic and Acoustic Scattering by Simple Shapes, North-Holland, Amsterdam, 1969, and references therein.
[8] R.C. Hansen (Ed.), Geometric Theory of Diffraction, IEEE Press, Piscataway, NJ, 1981.
[9] P.E. Ceruzzi, A History of Modern Computing, MIT Press, Cambridge, MA, 1998.
[10] J.C. Tannehill, D.A. Anderson, R.H. Pletcher, Computational Fluid Mechanics and Heat Transfer, first ed., Hemisphere Publishing, Washington, DC/New York, 1984; second ed., Taylor & Francis, Philadelphia, 1997.
[11] K.S. Yee, Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media, IEEE Trans. Ant. Propagat. 14 (1966) 302–307.
[12] R.F. Harrington, Field Computation by Moment Method, Krieger Publ., Malabar, FL, 1982.
[13] A. Taflove, Computational Electrodynamics: The Finite-Difference Time-Domain Method, Artech House, Norwood, MA, 1995.
[14] P.P. Silvester, R.L. Ferrari, Finite Elements for Electrical Engineers, second ed., Cambridge Univ. Press, Cambridge, UK, 1990.
[15] J.M. Jin, The Finite Element Method in Electromagnetics, John Wiley & Sons, New York, 1993.
[16] J.L. Volakis, A. Chatterjee, L.C. Kempel, Finite Element Method for Electromagnetics, IEEE Press, New York, 1998.
[17] S.M. Rao, D.R. Wilton, A.W. Glisson, Electromagnetic scattering by surfaces of arbitrary shape, IEEE Trans. Antennas Propagat. 30 (3) (1982) 409–418.
[18] E.J. Nyström, Uber die praktische auflosung von linearen integralgleichungen mit anwendungen auf randwertaufgaben der potentialtheorie, Commentationes Physico-Mathematicae 4 (15) (1928).
[19] L.S. Canino, J.J. Ottusch, M.A. Stalzer, J.L. Visher, S. Wandzura, Numerical solution of the Helmholtz equation in 2D and 3D using a higher order Nyström discretization, J. Comput. Phys. 146 (1998) 627–663.
[20] R. Lee, A.C. Cangellaris, A study of discretization error in the finite element approximation of wave solution, IEEE Trans. Ant. Propagat. 40 (5) (1992) 542–549.
[21] J.M. Jin, J. Liu, Z. Lou, C.S. Liang, A fully high-order finite element simulation of scattering by deep cavities, IEEE Trans. Antennas Propagat. 51 (9) (September 2003) 2420–2429.

[22] D.T. Borup, O.P. Gandhi, Fast-Fourier transform method for calculation of SAR distributions in finely discretized inhomogeneous models of biological bodies, IEEE Trans. Microwave Theory Tech. 32 (4) (1984) 355–360.
[23] C.Y. Shen, K.J. Glover, M.I. Sancer, A.D. Varvatsis, The discrete Fourier transform method of solving differential-integral equations in scattering theory, IEEE Trans. Ant. Propagat. 37 (8) (1989) 1032–1041.
[24] H. Gan, W.C. Chew, A discrete BCG-FFT algorithm for solving 3D inhomogeneous scatterer problems, J. Electromag. Waves Appl. 9 (10) (1995) 1339–1357.
[25] J.R. Phillips, J.K. White, Efficient capacitance computation of 3D structures using generalized pre-corrected FFT methods, in: Proceedings of the 3rd Topical Meeting on Electric Performance of Electronic Packaging, November 2–4, Monterey, CA, 1994.
[26] E. Bleszynski, M. Bleszynski, T. Jaroszewicz, A fast integral-equation solver for electromagnetic scattering problems, IEEE APS Int. Symp. Dig. (1994) 416–419.
[27] W.C. Chew, G.L. Wang, Anterpolation precorrected FFT for integral equations, in: PIERS 2003, Hawaii, October 13, 2003.
[28] A. Aydiner, W.C. Chew, J.M. Song, T.J. Cui, A Sparse Data Fast Fourier Transform (SDFFT), IEEE Antennas Propagat. 51 (11) (November 2003) 3161–3170.
[29] A.W. Appel, An efficient program for many-body simulation, SIAM J. Sci. Stat. Comput. 6 (1) (1985) 85–103.
[30] J. Barnes, P. Hut, A hierarchical O($N \log N$) force calculation algorithm, Nature 324 (1986) 446–449.
[31] L. Greengard, V. Rokhlin, A fast algorithm for particle simulations, J. Comput. Phys. 73 (1987) 325–348.
[32] R.L. Wagner, W.C. Chew, A study of wavelets for the solution of electromagnetic integral equations, IEEE Trans. Ant. Propagat. 43 (8) (1995) 802–810.
[33] W.C. Chew, Computational electromagnetics—the physics of smooth versus oscillatory fields, Philos. Trans. Royal Soc. London Ser. A, Math., Phys. Eng. Sci. Theme Issue Short Wave Scattering 362 (1816) (March 15, 2004) 579–602.
[34] W.K. Tung, Group Theory in Physics, World Scientific Publ., Philadelphia, PA, 1985.
[35] V. Rokhlin, Rapid solution of integral equations of scattering theory in two dimensions, J. Comput. Phys. 86 (1990) 414–439.
[36] A. Brandt, Multilevel computations of integral transforms and particle interactions with oscillatory kernels, Comp. Phys. Comm. 65 (1991) 24–38.
[37] C.C. Lu, W.C. Chew, A multilevel algorithm for solving boundary-value scattering, Micro. Opt. Tech. Lett. 7 (10) (July 1994) 466–470.
[38] J.M. Song, W.C. Chew, Multilevel fast-multipole algorithm for solving combined field integral equations of electromagnetic scattering, Micro. Opt. Tech. Lett. 10 (1) (September 1995) 14–19.
[39] J.S. Zhao, W.C. Chew, Integral equation solution of Maxwell's equations from zero frequency to microwave frequencies, IEEE Trans. Antennas Propagat., James R. Wait Memorial Special Issue 48 (10) (October 2000) 1635–1645.
[40] L.J. Jiang, W.C. Chew, Low frequency inhomogeneous plane wave algorithm—LF-FIPWA, Micro. Opt. Tech. Lett. 40 (2) (January 2004) 117–122.
[41] Y. Chu, W.C. Chew, S. Chen, J. Zhao, A surface integral equation method for low-frequency scattering from a composite object, IEEE Trans. Antennas Propagat. 51 (10) (October 2003) 2837–2844.
[42] Y.H. Chu, W.C. Chew, Large-scale computation for electrically small structures using surface integral equation method, Micro. Opt. Tech. Lett., 2005, in press.
[43] R.E. Collin, Foundation for Microwave Engineering, second ed., McGraw–Hill, New York, 1992.
[44] H. Contopanagos, B. Dembart, M. Epton, J.J. Ottusch, V. Rokhlin, J.L. Visher, S.M. Wandzura, Well-conditioned boundary integral equations for three-dimensional electromagnetic scattering, IEEE Trans. Antennas Propagat. 50 (December 2002) 1824–1830.
[45] Y. Zhang, T.J. Cui, W.C. Chew, J.S. Zhao, Magnetic field integral equation at very low frequencies, IEEE Trans. Antennas Propagat. 51 (8) (August 2003) 1864–1871.
[46] R.J. Adams, Physical and analytical properties of a stabilized electric field integral equation, IEEE Trans. Antennas Propagat. 52 (2) (2004) 362–372.
[47] D.R. Wilton, A.W. Glisson, On improving the electric field integral equation at low frequencies, in: URSI Radio Science Meeting Digest, Los Angeles, CA, June 1981, p. 24.
[48] D. Colton, R. Kress, Integral Equation methods in Scattering Theory, Wiley, New York, 1983.
[49] S. Ohnuki, W.C. Chew, Truncation error analysis of multipole expansion, SIAM J. Scientific Computing 25 (4) (2003) 1293–1306.
[50] M.L. Hastriter, S. Ohnuki, W.C. Chew, Error control of the translation operator in 3-D MLFMA, Micro. Opt. Tech. Lett. 37 (May 2003) 184–188.
[51] J.S. Zhao, W.C. Chew, Three dimensional multilevel fast multipole algorithm from static to electrodynamic, Micro. Opt. Tech. Lett. 26 (1) (July 2000) 43–48.
[52] W.C. Chew, J. Friedrich, R. Geiger, A multiple scattering solution for the effective permittivity of a sphere mixture, IEEE Trans. Geoscience and Remote Sensing 28 (2) (March 1990) 207–214.
[53] J.S. Zhao, W.C. Chew, Applying LF-MLFMA to solve complex PEC structures, Micro. Opt. Tech. Lett. 28 (3) (February 5, 2001) 155–160.
[54] L.J. Jiang, W.C. Chew, A mixed-form fast multipole algorithm, IEEE Trans. Antennas Propagat. (2005), in press.
[55] B. Hu, W.C. Chew, E. Michielssen, J. Zhao, An improved fast steepest descent algorithm for the fast analysis of two-dimensional scattering problems, Radio Science 34 (4) (July–August 1999) 759–772.
[56] B. Hu, W.C. Chew, Fast inhomogeneous plane wave algorithm for electromagnetic solutions in layered medium structures—2D case, Radio Science 35 (1) (2000) 31–43.
[57] B. Hu, W.C. Chew, Fast inhomogeneous plane wave algorithm for scattering from objects above the multi-layered medium, IEEE Trans. Geosci. Remote Sensing 39 (5) (May 2001) 1028–1038.
[58] B. Hu, W.C. Chew, S. Velamparambil, Fast inhomogeneous plane wave algorithm (FIPWA) for analysis of electromagnetic scattering, Radio Science 36 (6) (December 2001) 1327–1340.

[59] B. Hu, Fast inhomogeneous plane wave algorithm for electromagnetic scattering problems, Ph.D. Thesis, Dept. Elec. Comp. Engrg., U. Illinois, January 2001.
[60] W.C. Chew, Waves and Fields in Inhomogeneous Media, Van Nostrand Reinhold, New York, 1990. Reprinted by IEEE Press, 1995.
[61] S. Ohnuki, W.C. Chew, Error analysis of the fast inhomogeneous plane wave algorithm for 2-D free space cases, Micro. Opt. Tech. Lett. 38 (4) (2003) 300–304.
[62] H.S. Park, S. Ohnuki, B. Hu, H.T. Kim, W.C. Chew, Error control of the fast inhomogeneous plane wave algorithm for 3-D free space cases, Research Report: CCEM No. 03-05, Ctr Comput. Electromag. and Electromag. Laboratory, U. Illinois, May 20, 2005.
[63] Y.C. Pan, W.C. Chew, A fast multipole method for embedded structure in a stratified medium, Progress in Electromagnetics Research PIER 44 (2004) 1–38.
[64] Y.C. Pan, Development of fast multipole method for stratified medium, Ph.D. Thesis, Dept. Elec. Comp. Engrg., U. Illinois, May 2002.
[65] L.J. Jiang, W.C. Chew, Y.C. Pan, Capacitance extraction in the multilayer medium using DCIM and SMFMA, JEMWA, submitted for publication.