# Numerical analysis of the potential formulation of the volume integral equation for electromagnetic scattering

**Johannes Markkanen**

[1]Department of Physics, University of Helsinki

**Key Points:**

- Discretization of the potential volume integral equation for light scattering is presented
- Accuracy and stability of the method is studied
- The spectrum of the discretized system coinsides with the theoretically predicted spectrum

Corresponding author: Johannes Markkanen, `johannes.markkanen@helsinki.fi`

**Abstract**

We present a discretization of the potential volume integral equation (PVIE) for electromagnetic scattering by dielectric objects. The equations are written for the vector and scalar potentials combined with the Lorenz gauge condition. The advantage of the potential formulation over the more conventional formulations, i.e., the field-, flux-, or current-formulations, is that the potentials are continuous across material interfaces enabling the use of fully continuous basis functions. The discretization with the fully continuous basis functions leads to a better conditioned system matrix compared to those of the other formulations discretized with the fully or partially discontinuous functions. Thus, it speeds up the iterative solution. The method is accelerated with the pre-corrected fast Fourier transform algorithm.

# 1 Introduction

Electromagnetic wave scattering by strongly inhomogenous objects is a common problem arising from many different fields in science and engineering such as remote sensing, astronomy, electromagnetic imaging, bioelectromagnetics, and metamaterials. The volume-integral-equation (VIE) methods are typically applied for such problems. Strong material contrasts including negative parameters, however, give rise to a major numerical challenge because of an ill-conditioned matrix when the system is solved iteratively.

The VIE formulations can be written for the fields [*Sun and Chew*, 2009], flux densities [*Schaubert et al.*, 1984], polarization currents [*Yurkin et al.*, 2010; *Markkanen et al.*, 2012a], and potentials [*De Doncker*, 2001; *Chang and Lomakin*, 2011]. The choice of the formulation and the unknown allows us to construct several different VIE formulations for dielectric scattering problems [*Botha*, 2006; *Sun and Chew*, 2009; *Markkanen et al.*, 2012b]. The choice also decides the type of basis and testing functions that should be used in order to obtain a converging numerical solution. The basis and testing functions should span the proper finite-dimensional function spaces [*van Beurden and van Eijndhoven*, 2007, 2008].

The use of partially or fully discontinuous basis functions to discretize the VIEs may introduce spurious solutions associated with the element boundaries due to the spreading of the spectrum of the discretized system [*Markkanen and Ylä-Oijala*, 2016]. Particularly, this may happen when the material parameter is negative. In the PVIE, however, the unknown functions are fully continuous scalar and vector potentials which yields to the spectrum without the artificial continuous part for objects with smooth boundaries as demonstrated by *Markkanen and Ylä-Oijala* [2016].

In [*De Doncker*, 2001; *DeDoncker*, 2003], the PVIE was discretized with curvilinear cubes in combination with the nodal Lagrange basis functions and the collocation method. Later, *Chang and Lomakin* [2011] replaced the scalar potential equation in the PVIE by the discretized version of the Lorenz gauge condition in order to remove the singular surface integral equation. Both approaches, however, have shown some serious accuracy problems when the domain is discretized with linear tetrahedral elements [*Adil*, 2013; *Markkanen and Ylä-Oijala*, 2016]. Finally, *Markkanen* [2016] showed that the accuracy problems in the DeDoncker's formulation arise from the properties of the discrete basis functions in the tetrahedral mesh, i.e., the volume charge can be non-zero in an element with a constant permittivity. The problem was solved by adding a term to the scalar potential equation which was assumed to vanish in [*De Doncker*, 2001; *DeDoncker*, 2003].

In this paper, we study the properties of the PVIE formulation by *Markkanen* [2016]. We show that the collocation method leads to a robust discretization scheme that allows for an efficient solution for sufficiently smooth dielectric scattering problems. The eigenvalue distribution of the system matrix corresponds almost exactly the theoretically predicted one. Albeit, the optimal convergence may not be achieved and spurious solutions

may appear for objects with negative permittivities and sharp corners. This is because the testing functions do not span the proper function space, i.e., the dual space of the range space of the operator. These findings suggest that the PVIE is an attractive formulation for the analysis of complex media but further studies are required to find the proper set of testing functions. The rest of the paper is organized as follows. In Section 2, the PVIE is derived and its low frequency spectral properties are investigated in Section 3. The discretization scheme and the precorrected-FFT (pFFT) acceleration technique are presented in Sections 4 and 5, respectively. Numerical results are shown and discussed in Section 6, and the paper is concluded in Section 7.

## 2 Potential formulation

Consider a time-harmonic electromagnetic scattering by a three-dimensional bounded object $D$ in a homogeneous background medium with the constant permittivity $\epsilon_0$ and permeability $\mu_0$. The time factor $e^{-i\omega t}$, where $\omega$ is the angular frequency, is assumed and suppressed. We assume that the object $D$ is isotropic with the position $\mathbf{r}$ dependent relative permittivity function $\epsilon_r(\mathbf{r})$. The relative permability is restricted to unity, i.e., $\mu_r(\mathbf{r}) = 1$.

The electric $\mathbf{E}$ and magnetic $\mathbf{H}$ fields can be represented in terms of the vector $\mathbf{A}$ and scalar $V$ potentials as

$$\mathbf{E}(\mathbf{r}) = i\omega\mathbf{A}(\mathbf{r}) - \nabla V(\mathbf{r}), \ \ \mathbf{H}(\mathbf{r}) = \frac{1}{\mu_0}\nabla \times \mathbf{A}(\mathbf{r}). \tag{1}$$

To define the potentials unambiguously, we apply the Lorenz gauge condition, namely

$$i\omega\nabla \cdot \mathbf{A}(\mathbf{r}) = -k^2 V(\mathbf{r}). \tag{2}$$

With the Lorenz gauge imposed, the vector and scalar potentials satisfy the vector and scalar Helmholtz equations, i.e.,

$$\nabla^2\mathbf{A}(\mathbf{r}) + k^2\mathbf{A}(\mathbf{r}) = -\mu_0\mathbf{J}(\mathbf{r}), \tag{3}$$

$$\nabla^2 V(\mathbf{r}) + k^2 V(\mathbf{r}) = -\frac{\rho(\mathbf{r})}{\epsilon_0}. \tag{4}$$

The solutions for the vector and scalar Helmholtz equations in terms of the free-space Green's function $G$ read as

$$\mathbf{A}(\mathbf{r}) = \mu_0 \int_D G(\mathbf{r},\mathbf{r}')\mathbf{J}(\mathbf{r}')\,d\mathbf{r}' \tag{5}$$

and

$$V(\mathbf{r}) = \frac{1}{\epsilon_0} \int_D G(\mathbf{r},\mathbf{r}')\rho(\mathbf{r}')\,d\mathbf{r}', \tag{6}$$

where $\mathbf{J}$ and $\rho$ are the source current and charge densities, respectively.

To derive the potential formulation, we apply the volume-equivalence principle, i.e., the scatterer is removed and replaced by the equivalent current and charge densities defined as

$$\mathbf{J}_{\text{eq}}(\mathbf{r}) = -i\omega\epsilon_0(\epsilon_r(\mathbf{r}) - 1)\mathbf{E}(\mathbf{r}) \tag{7}$$

$$\rho_{\text{eq}}(\mathbf{r}) = \frac{1}{i\omega}\nabla \cdot \mathbf{J}_{\text{eq}}(\mathbf{r}). \tag{8}$$

The equivalent current and charge densities are the sources for the scattered potentials given by

$$\mathbf{A}^s(\mathbf{r}) = \mu_0 \int_D G(\mathbf{r},\mathbf{r}')\mathbf{J}_{\text{eq}}(\mathbf{r}')\,d\mathbf{r}' \tag{9}$$

$$V^s(\mathbf{r}) = \frac{1}{\epsilon_0} \int_D G(\mathbf{r},\mathbf{r}')\rho_{\text{eq}}(\mathbf{r}')\,d\mathbf{r}' \tag{10}$$

and the total potentials can be represented as a sum of the incident and scattered potentials as

$$\mathbf{A}(\mathbf{r}) = \mathbf{A}^{\text{inc}}(\mathbf{r}) + \mathbf{A}^{\text{s}}(\mathbf{r}). \tag{11}$$

$$V(\mathbf{r}) = V^{\text{inc}}(\mathbf{r}) + V^{\text{s}}(\mathbf{r}). \tag{12}$$

Substituting (9) into (11) and using the definition of the equivalent current, the integral equation for the vector potential reads as

$$i\omega\mathbf{A}^{\text{inc}}(\mathbf{r}) = i\omega\mathbf{A}(\mathbf{r}) - k^2 \int_D (\epsilon_r(\mathbf{r}') - 1)(i\omega\mathbf{A}(\mathbf{r}') - \nabla'V(\mathbf{r}'))G(\mathbf{r},\mathbf{r}')\,\mathrm{d}\mathbf{r}'. \tag{13}$$

Taking the divergence of the above equation and using the Lorenz gauge (2), the equation for the scalar potential yields

$$V^{\text{inc}}(\mathbf{r}) = V(\mathbf{r}) + \nabla \cdot \int_D (\epsilon_r(\mathbf{r}') - 1)(i\omega\mathbf{A}(\mathbf{r}') - \nabla'V(\mathbf{r}'))G(\mathbf{r},\mathbf{r}')\,\mathrm{d}\mathbf{r}'. \tag{14}$$

Next, we define the normalized vector potential $\tilde{\mathbf{A}} = i\omega\mathbf{A}$ to improve numerical balance of the unknowns. By moving the divergence operator inside the integral and using $\nabla G = -\nabla'G$, yields

$$\begin{cases} \tilde{\mathbf{A}}^{\text{inc}}(\mathbf{r}) & = & \tilde{\mathbf{A}}(\mathbf{r}) - k^2 \int_D (\epsilon_r(\mathbf{r}') - 1)(\tilde{\mathbf{A}}(\mathbf{r}') - \nabla'V(\mathbf{r}'))G(\mathbf{r},\mathbf{r}')\,\mathrm{d}\mathbf{r}' \\ V^{\text{inc}}(\mathbf{r}) & = & V(\mathbf{r}) + \int_D \nabla' \cdot [(\epsilon_r(\mathbf{r}') - 1)(\tilde{\mathbf{A}}(\mathbf{r}') - \nabla'V(\mathbf{r}'))]G(\mathbf{r},\mathbf{r}')\,\mathrm{d}\mathbf{r}' \\ & & - \int_{\partial D} \mathbf{n}(\mathbf{r}') \cdot [(\epsilon_r(\mathbf{r}') - 1)(\tilde{\mathbf{A}}(\mathbf{r}') - \nabla'V(\mathbf{r}'))G(\mathbf{r},\mathbf{r}')]\,\mathrm{d}\mathbf{r}'. \end{cases} \tag{15}$$

The surface and volume integrals in the scalar potential equation correspond the contributions of the surface and volume charges, respectively. It is worth noting that in the discretized potential formulations [*De Doncker*, 2001; *DeDoncker*, 2003; *Markkanen and Ylä-Oijala*, 2016], the volume integral in the scalar potential equation was assumed to be zero inside each element and dropped out. This assumption, however, leads to a diverging numerical solution if the basis functions do not inherently satisfy the zero divergence condition [*Markkanen*, 2016].

## 3  Essential spectrum

In this section, we analyze the spectral properties of the PVIE. The spectrum of a bounded operator $T$ is a set of complex-valued numbers $\lambda_s$ such that $\lambda_s I - T$ is not invertible, where $I$ is the identity operator. The spectrum of a continuous operator is related to the eigenvelues of the discterized operator, i.e., the system matrix. Thus, understanding the spectrum of the VIE operator is of interest not only in order to understand the resonances associated with the given structure but also from the preconditioning point of view. Moreover, the knowledge of the spectrum can be used to estimate the number of iterations required to solve a given problem (See, e.g., [*Yurkin*, 2016] and references therein).

The spectrum of the operator can be decomposed into two parts. The essential spectrum is a closed subset of the spectrum which is stable under compact perturbations, and the remaining part contains isolated eigenvalues and possible continuous part. We can find the essential spectrum of the PVIE by removing all compact operators. First we write the operator as a matrix form:

$$T = \begin{pmatrix} \mathbf{I} - k^2\mathcal{V} & \mathcal{U} \\ \nabla \cdot \mathcal{V} & I - \nabla \cdot \mathcal{U} \end{pmatrix} \tag{16}$$

where

$$\mathcal{V}(\mathbf{A})(\mathbf{r}) = \int_D (\epsilon_r(\mathbf{r}') - 1)G(\mathbf{r}, \mathbf{r}')\mathbf{A}(\mathbf{r}')\, d\mathbf{r}', \tag{17}$$

and

$$\mathcal{U}(V)(\mathbf{r}) = \int_D (\epsilon_r(\mathbf{r}') - 1)G(\mathbf{r}, \mathbf{r}')\nabla' V(\mathbf{r}')\, d\mathbf{r}'. \tag{18}$$

The operators $\mathcal{V}$ and $\mathcal{U}$ are bounded from $H^1$ to $H^2$, where $H^s$ refers to the standard Sobolev spaces with $s \in \mathbb{R}$, hence they are compact in $H^1$ by the embedding theorem.

By removing the compact operators, the essential spectrum can be found by solving

$$det \begin{vmatrix} \mathbf{I} - \lambda_1\mathbf{I} & 0 \\ \nabla \cdot \mathcal{V} & (I - \nabla \cdot \mathcal{U}) - \lambda_2 I \end{vmatrix} = (\mathbf{I} - \lambda_1\mathbf{I})((I - \nabla \cdot \mathcal{U}) - \lambda_2 I) = 0. \tag{19}$$

Therefore, the essential spectrum is the union of the essential spectra of the diagonal operators in (16). It is clear that $\lambda_1 = 1$ but finding $\lambda_2$ values is more complicated. The operator $\nabla \cdot \mathcal{U}$ appears in acoustic scattering and it has been analyzed in [*Costabel*, 2015]. The essential spectrum corresponds to that of the double layer potential operator, which can be written in the principal value sense as

$$\sum_i \left[ \frac{\Omega_{r,i}(\mathbf{r})}{4\pi}(\epsilon_{r,i}(\mathbf{r}) - 1)V(\mathbf{r}) + p.v. \int_{\partial D_i} (\epsilon_{r,i}(\mathbf{r}') - 1)\partial_{n'}G(\mathbf{r}, \mathbf{r}')V(\mathbf{r}')\, d\mathbf{r}' \right]. \tag{20}$$

Here, we have divided the domain $D$ into the subdomains $D_i$ with constant permittivity $\epsilon_{r,i}$. The solid angle $\Omega_{r,i}$ is the opening angle that the observation point $\mathbf{r}$ sees the domain $D_i$, and $\partial_n$ corresponds the normal derivative. The principal value part is compact on smooth surfaces and bounded on non-smooth surfaces. Thus, the spectrum spreads out around the accumulation points related to the non-smooth parts. By adding the contribution of the identity operator, the accumulation points of the essential spectrum of the PVIE reads as

$$\sigma_e(\epsilon_{r,i}) = \{1, 1 + \frac{\Omega_{r,i}}{4\pi}(\epsilon_{r,i} - 1)\}. \tag{21}$$

For homogeneous objects with smooth surfaces, only three accumulation points exist, hence $\sigma_e(\epsilon_r) = \{1, \frac{1}{2}(\epsilon_r + 1), \epsilon_r\}$. For smooth surfaces $\Omega_{r,i} = 2\pi$ and the accumulation point is at $\frac{1}{2}(\epsilon_r + 1)$. The accumulation point at $\epsilon_r$ appears when the observation point is inside a homogeneous domain because $\Omega_{r,i} = 4\pi$. Finally, we note that the essential spectrum of the PVIE is the same for other VIE formulations [*Costabel et al.*, 2012; *Markkanen*, 2014].

## 4 Discretization

The vector and scalar potential equations define bounded mappings given by

$$\begin{aligned} \mathbf{H}^1(\mathbb{R}^3)^3 \times H^1(\mathbb{R}^3)^1 &\to \mathbf{H}^1(\mathbb{R}^3)^3 \\ \mathbf{H}^1(\mathbb{R}^3)^3 \times H^1(\mathbb{R}^3)^1 &\to H^1(\mathbb{R}^3)^1, \end{aligned} \tag{22}$$

where $\mathbf{H}^1(\mathbb{R}^3)^3$ is the space of vector functions whose components are in $H^1(\mathbb{R}^3)^1$, and $H^1(\mathbb{R}^3)^1$ is the space of continuous scalar functions. Since the material parameter $(\epsilon_r - 1)$ vanishes outside the domain $D$, we can consider the PVIE on $D$. This is easy to verify since if the potentials solves the PVIE on $D$ they can be extended outside on $D$ by the same equation.

We discretize the unknowns by dividing the domain into tetrahedral elements and expand the scalar potential with fully continuous basis functions except on the boundary

$\partial D$. The scalar potential is expanded by linear nodal basis functions $N_n$ which span the finite-dimensional space which is a subspace of $H^1(D)^1$ as

$$V \approx \sum_n c_n N_n \tag{23}$$

where $c_n$ are the unknown coefficients. On $\partial D$, nodal basis functions defined on the boundary nodes are used. Furthermore, each component of the vector potential $\mathbf{A}$ is expanded by the same functions as

$$
\begin{aligned}
A_x &\approx \sum_n x_n N_n \\
A_y &\approx \sum_n y_n N_n \\
A_z &\approx \sum_n z_n N_n.
\end{aligned}
\tag{24}
$$

This discretization of the unknowns results in $4N_{nodes}$ scalar equations in $H^1(D)$ where $N_{nodes}$ is the number of nodes in the tetrahedral mesh. We apply the projection method with the symmetric $L_2$ product to discretize the equations. Thus, to guarantee that the solution converges in a proper norm the test functions should span the dual space of $H^1(D)$. We discretize the equations with the collocation method and the Galerkin method. In the collocation method, the test functions are distributions and it is typically applied for problems with sufficiently smooth material parameter profiles. In the Galerkin method, the test and basis functions are identical. Consequently, the test space is a small subspace of the dual space of the range of the PVIE, and the convergence in the norm of the solution cannot be guaranteed. Interestingly, in the case of acoustic surface integral equations in $H^{1/2}$, *Ylä-Oijala et al.* [2015] did not find any significant differences in accuracy when the equations were discretized with $H^{-1/2}$ and $H^{1/2}$ conforming test functions.

Next, we write the elements of the resulting system matrix discretized by the Galerkin and the collocation methods. By applying Galerkin's technique where testing and basis functions are identical, the elements of the system matrix can be written as

$$
Z_{mn} = \begin{pmatrix}
A_x A_x & 0 & 0 & A_x V \\
0 & A_y A_y & 0 & A_y V \\
0 & 0 & A_z A_z & A_z V \\
V A_x & V A_y & V A_z & V V
\end{pmatrix}
\tag{25}
$$

where

$$
\begin{aligned}
(A_i A_j)_{mn} &= \hat{\mathbf{e}}_i \cdot \hat{\mathbf{e}}_j \int_{V_m} N_m N_n \, dV - \hat{\mathbf{e}}_i \cdot \hat{\mathbf{e}}_j k^2 \int_{V_m} \int_{V_n} (\epsilon'_n - 1) N_m N'_n G \, dV' dV \\
(A_i V)_{mn} &= \hat{\mathbf{e}}_i \cdot k^2 \int_{V_m} \int_{V_n} (\epsilon'_n - 1) N_m \nabla' N'_n G \, dV' dV \\
(V A_j)_{mn} &= -\int_{V_m} \int_{\partial V_n} \mathbf{n}' \cdot \hat{\mathbf{e}}_j (\epsilon'_n - 1) N_m N'_n G \, dS' dV \\
&\quad + \int_{V_m} \int_{V_n} (\epsilon'_n - 1) \hat{\mathbf{e}}_j \cdot N_m \nabla' N'_n G dV' dV \\
(V V)_{mn} &= \int_{V_m} N_m N_m \, dV + \int_{V_m} \int_{\partial V_n} \mathbf{n}' \cdot (\epsilon'_n - 1) N_m \nabla' N'_n G \, dS' dV,
\end{aligned}
\tag{26}
$$

in which $i$ and $j$ obtain values $x$, $y$, and $z$. The collocation scheme is obtained by replacing the test functions $N_m$ with the delta functions $\delta_{mn}$ ($\delta_{mn} = 1$ when $m = n$ and $\delta_{mn} = 0$ when $m \neq n$). All singular integrals are evaluated with the singularity extraction method *Järvenpää et al.* [2003].

## 5 Acceleration with the FFT

The adaptive integral method (AIM) or precorrected-FFT algorithm (pFFT) [*Bleszynski et al.*, 1996; *Phillips and White*, 1997] separates the computations of the near and far field interactions. The near field interactions are directly computed by using (26) and the far field interactions are approximated with the interpolation and the FFT.

First, the tetrahedral mesh is enclosed by a uniform auxiliary grid and the auxiliary grid is divided into equi-sized cubes with the edge length $\delta_{\text{cube}}$ such that each cube contains $p^3$ grid points and few basis functions. In this paper we use order $p = 3$, i.e., each cube has 27 interpolation nodes. Second, we compute and store 8 projection matrices. The projection matrices are mappings from the primary basis functions in the tetrahedral mesh into the auxiliary sources in the uniform grid. The projections are defined by requiring the basis functions in a cube and the auxiliary sources related to the same cube to radiate the same fields at given test points $\mathbf{r}_t$. The test points can be selected, e.g., on the surface of a sphere or cube with radius/edgelength $R$ whose center coincides with the center of the cube. We have used a cube $R = 4\delta_{\text{cube}}$ with 98 evenly distributed points $\mathbf{r_t}$ on the surface of the cube. The number of test points $\mathbf{r}_t$ should be larger than the number of interpolation nodes $p^3$.

The mapping from the grid sources to the test potentials can be defined as

$$P_{t,l} = G(\mathbf{r}_t, \mathbf{r}_l) \tag{27}$$

where $\mathbf{r}_l \ l = 1, 2, ..., p^3$ are the positions of the grid points in a cube. Mappings from the basis functions and their derivatives to the test potentials are given by

$$
\begin{aligned}
P_{t,n}^A &= -\hat{\mathbf{e}}_i \cdot \hat{\mathbf{e}}_j k^2 \int_{V_n} (\epsilon_n - 1) N_n G(\mathbf{r}_t, \mathbf{r}) \, d\mathbf{r} \\
P_{t,n}^{dV} &= \hat{\mathbf{e}}_i \cdot k^2 \int_{V_n} (\epsilon_n - 1) \nabla N_n G(\mathbf{r}_t, \mathbf{r}) \, d\mathbf{r} \\
P_{t,n}^{dA} &= -\int_{\partial V_n} (\epsilon_n - 1) \mathbf{n} \cdot \hat{\mathbf{e}}_j N_n G(\mathbf{r}_t, \mathbf{r}) \, d\mathbf{r} \\
&+ \int_{V_n} (\epsilon_n - 1) \hat{\mathbf{e}}_j \cdot \nabla N_n G(\mathbf{r}_t, \mathbf{r}) \, d\mathbf{r} \\
P_{t,n}^V &= \int_{\partial V_n} (\epsilon_n - 1) \mathbf{n} \cdot \hat{\mathbf{e}}_j \nabla N_n G(\mathbf{r}_t, \mathbf{r}) \, d\mathbf{r}
\end{aligned}
\tag{28}
$$

Finally, we can define mappings from the basis functions to the grid sources by requiring that they both radiate the same fields at test points, namely

$$
\begin{aligned}
\Pi_{l,n}^A &= \sum_k P_{l,t}^+ P_{t,n}^A \\
\Pi_{l,n}^{dV} &= \sum_k P_{l,t}^+ P_{t,n}^{dV} \\
\Pi_{l,n}^{dA} &= \sum_k P_{l,t}^+ P_{t,n}^{dA} \\
\Pi_{l,n}^V &= \sum_k P_{l,t}^+ P_{t,n}^V
\end{aligned}
\tag{29}
$$

in which $P_{l,t}^+$ denotes the pseudo inverse of $P_{t,l}$.

The interpolation operator is defined as a mapping from the grid sources to the testing functions, namely

$$\Lambda_{m,l} = \left( \sum_k P_{l,t}^+ I_{t,m} \right)^T \tag{30}$$

where $()^T$ denotes the transpose. For Galerkin's method

$$I_{t,m} = \int_{V_m} N_m G(\mathbf{r}_t, \mathbf{r}) \, d\mathbf{r}, \tag{31}$$

and for the collocation

$$I_{t,m} = G(\mathbf{r}_t, \mathbf{r}_m) \tag{32}$$

where $\mathbf{r}_m$ is the collocation point.

A matrix-vector multiplication for the far-field part can be executed with the FFT for each components separately as

$$
\begin{aligned}
ax_m &= \Lambda_{m,l} \mathcal{F}^{-1} \left( \mathcal{G} \mathcal{F} \left( \Pi_{l,n}^{A,xx} x_n + \Pi_{l,n}^{dV,x} c_n \right) \right) \\
ay_m &= \Lambda_{m,l} \mathcal{F}^{-1} \left( \mathcal{G} \mathcal{F} \left( \Pi_{l,n}^{A,yy} y_n + \Pi_{l,n}^{dV,y} c_n \right) \right) \\
az_m &= \Lambda_{m,l} \mathcal{F}^{-1} \left( \mathcal{G} \mathcal{F} \left( \Pi_{l,n}^{A,zz} z_n + \Pi_{l,n}^{dV,z} c_n \right) \right) \\
ac_m &= \Lambda_{m,l} \mathcal{F}^{-1} \left( \mathcal{G} \mathcal{F} \left( \Pi_{l,n}^{dA,x} x_n + \Pi_{l,n}^{dA,y} y_n + \Pi_{l,n}^{dA,z} z_n + \Pi_{l,n}^{V} c_n \right) \right),
\end{aligned}
\tag{33}
$$

in which $\mathcal{F}$ denotes the FFT, $\mathcal{F}^{-1}$ the inverse FFT, and $\mathcal{G}$ is the zero-padded Fourier transformed Green's function matrix related to the uniform grid.

Therefore we have computed the far-field interactions by 4 3D-FFTs and IFFTs in addition to 14 sparse matrix-vector multiplications. It worth noting here that accelerating, e.g., the JVIE formulation [*Markkanen et al.*, 2012a] with the pFFT requires 6 3D FFTs and IFFTs instead of 4. The discretized JVIE is written for 3 current components and their normal or rotated tangential traces resulting in 6 FFTs and IFFTs. Finally, since the interpolations are not accurate in the near-zone, the interactions between basis function in the near-zone (distance between basis functions $< l_{nz}$) are corrected by computing them directly from (26).

## 6 Numerical results

In this section, we present various numerical examples to validate the method. The algorithm has been implemented with modern Fortran language and the FFTW-3.3.6 library is used to compute the FFTs. Further, the code is parallelized with the openMPI. In all computations, the near-zone distance is chosen to be $l_{nz} = 3.5\delta_{box}$ and $\delta_{box}$ is selected to be a mean edge length of the mesh. The order $p = 3$, i.e., each cube contains $p^3$ grid points. For numerical integration, the third order Gaussian quadrature for tetrahedral elements is applied with the singularity subtraction method in which one term is subtracted from the kernel and integrated analytically. The generalized minimal residual method GMRES(50) is used to solve the linear systems with tolerance $10^{-5}$ ($L_2$-norm).

First, we study the convergence of the FFT-accelerated PVIE solution with respect to the element size defined as a mean edge length of tetrahedral mesh. The error norms are defined as follows:

The total solution error reads

$$err_{tot} = \frac{|\sum_v \mathbf{E}_v \cdot \mathbf{E}_v^* - \sum_v \mathbf{E}_v^{\mathrm{mie}} \cdot \mathbf{E}_v^{*,\mathrm{mie}}|}{\sum_v \mathbf{E}_v^{\mathrm{mie}} \cdot \mathbf{E}_v^{*,\mathrm{mie}}} \tag{34}$$

which is related to the total energy error of the system. Here, $()^*$ denotes the complex conjugate. The total energy error, however, does not tell anything about the error distribution, hence we are also interested in $L_1$-norm of the error expressed as

$$err_{L_1} = \frac{\sum_v |(\mathbf{E}_v \cdot \mathbf{E}_v^* - \mathbf{E}_v^{\mathrm{mie}} \cdot \mathbf{E}_v^{*,\mathrm{mie}})|}{\sum_v \mathbf{E}_v^{\mathrm{mie}} \cdot \mathbf{E}_v^{*,\mathrm{mie}}}. \tag{35}$$

A measure for the worst case error is provide by the $L_\infty$-norm

$$err_{L_\infty} = \frac{\max_v |(\mathbf{E}_v \cdot \mathbf{E}_v^* - \mathbf{E}_v^{\mathrm{mie}} \cdot \mathbf{E}_v^{*,\mathrm{mie}})|}{\max_v (\mathbf{E}_v^{\mathrm{mie}} \cdot \mathbf{E}_v^{*,\mathrm{mie}})}. \tag{36}$$

We study the errors integrated over small equi-sized cubical elements rather than point-wise errors. To compute these integrals, the scattering object is enclosed by a $60 \times 60 \times 60$ cubical array. The electric fields $\mathbf{E}_v$, and $\mathbf{E}_v^{\text{mie}}$ indicate the integrated field values over a cube $v$ in the array, computed by the PVIE and the analytical Mie solution, respectively. Eight point integration rule is applied for each cube.

Fig. 1 shows the errors as functions of element size in the free space wavelengths for a homogeneous sphere of size $kr = 2.037$ and $\epsilon_r = 5.0 + i1$ computed with the collocation method and the Galerkin method. The number of tetrahedral elements used in the computations were 376,1120, 4357,15695,53441, and 139010. In this case both testing procedures lead to almost identical $L_1$ and $L_\infty$ error convergence behavior with respect to the element size. The integrated error with the collocation method converges slightly faster. We note that the collocation method is much faster since the calculation of the pre-correction matrix is cheaper with the collocation than that of with the Galerkin method approximately by a factor of 10. Moreover, the number of iterations required to solve the system was 27 for the collocation and 53 for the Galerkin method. Time per iteration is the same for both methods.
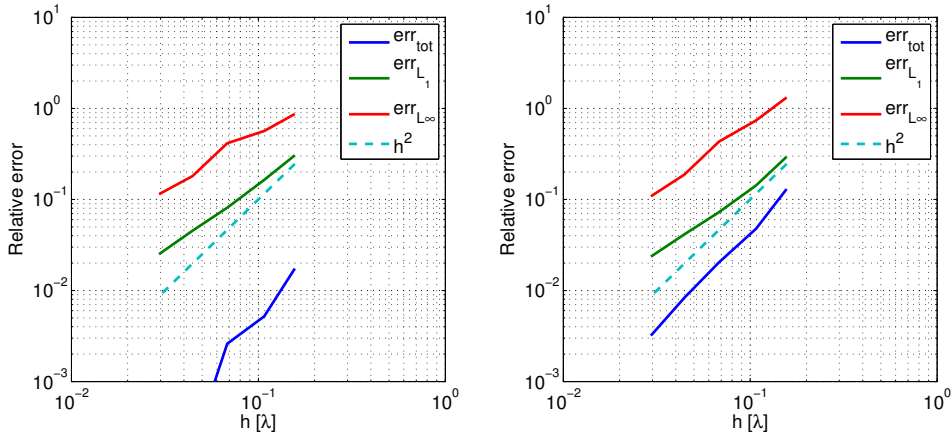


**Figure 1.** Relative errors calculated by the PVIE with the collocation method (left) and the Galerkin method (right) for a sphere $kr = 2.037$ and $\epsilon_r = 5.0 + i1$.

Next example is a larger sphere. The size parameter $kr = 10.186$ and the permittivity $\epsilon_r = 2.9 + i0.1$. Fig. 2 displays relative errors for the collocation and the Galerkin methods, respectively. The same meshes are used in the computations as in Fig. 1 with one addition mesh containing 380000 elements. As opposed to the smaller sphere, the convergence rate of the collocation method is slower than that of the Galerkin method with large elements but the asymptotic convergence rate seems to be the same.

Based on the above results and several other examples we have studied, the collocation method is favorable over Galerkin's for problems involving small details or inhomogeneities. Galerkin's method works better for large homogeneous or weakly inhomogenous objects. In the rest of the paper we only consider the collocation method.

Let us consider scattering by a $20 \times 20 \times 20$ cubical lattice of dielectric spheres. The length of the unit cell edge is 3 in arbitrary units. The permittivity of each sphere $\epsilon_r = 3 + i$ and the radius $r = 1$ (arb. units). Each sphere is discretized with 300 tetrahedral elements giving in total 2.4M elements and 2.816M unknowns. The mean edgelength equals $0.07\lambda$. The incident planewave propagates along $z$-axis with the wavenumber $k = 1$ (arb. units)$^{-1}$. Fig. 3 shows the computed S11 element of the Mueller matrix as a function
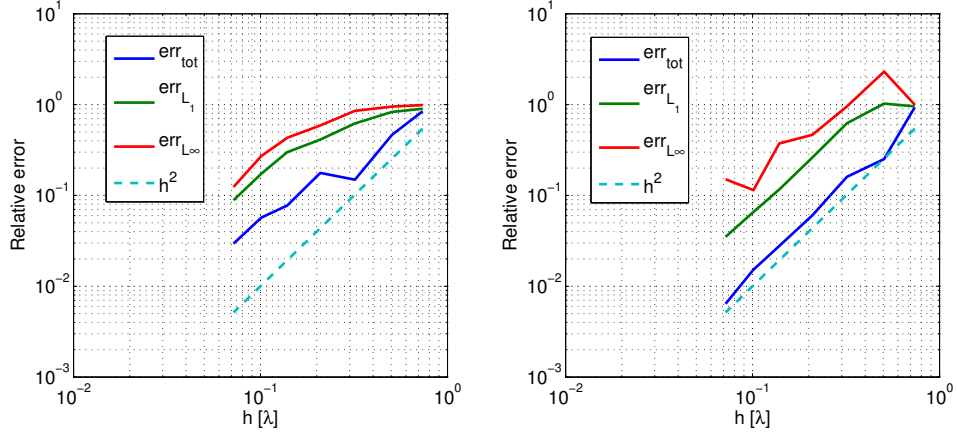
**Figure 2.** Same as in Fig. 1 but for a sphere $kr = 10.186$ and $\epsilon_r = 2.9 + i0.1$.

of scattering angle in the $yz$–plane. The reference results is computed semi-analytically with the fast superposition T-matrix method (FaSTMM) [*Markkanen and Yuffa*, 2017]. Excellent agreement is observed also for the other Mueller matrix elements.
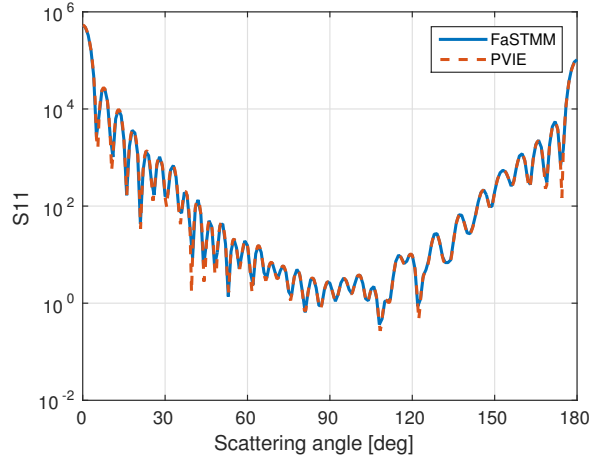


**Figure 3.** Computed Mueller matrix element S11 of a cubical lattice of spheres with the PVIE and the semi-analytical superposition T-matrix method. The scattering plane is the $yz$–plane and the forward scattering direction is at $0°$.

Next, we study materials with negative permittivities. Fig. 4 illustrates the extinction cross sections (left) and the number of iterations (right) of a sphere discretized with 376 and 15695 linear tetrahedral elements as functions of the real part of the permittivity. The imaginary part of the permittivity equals 0.01, and the size parameter equals 0.01.

For the ideal sphere, there is a static resonance at $\epsilon_r = -2$ which is visible as a spike in the extinction cross section. We observe that the PVIE finds the static resonance if the discretization is dense enough. With coarser mesh the position of the spike has shifted because of the discretized geometry considerably deviates from the ideal sphere. The solid angles $\Omega_{r,i}$ averaged over surface nodes equal 4.76 and 5.76 for the coarse and
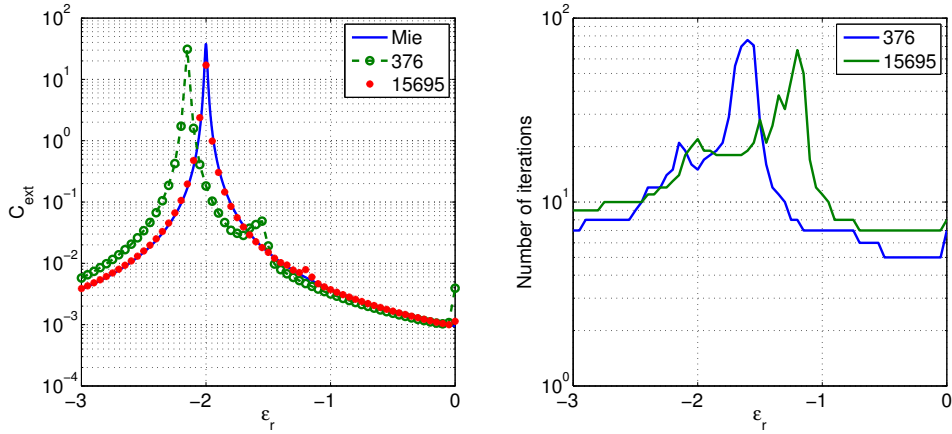
**Figure 4.** Computed extinction cross sections and the number of iterations as functions of the real part of permittivity. The sphere of size $kr$=0.01 is discretized with 376 and 15695 linear tetrahedral elements. The solid angle $\Omega_{r,i}$ avaraged over the surface nodes equals 4.76 (376) and 5.76 (15695) corresponding the unstable points around $\epsilon_r = -1.64$ and $\epsilon_r = -1.18$, respectively.

dense mesh, respectively. For the ideal sphere the solid angle should be $2\pi$. From the spectral analysis, we know that for the ideal sphere, the PVIE is not solvable at $\epsilon_r = -1$. Since the discretized PVIE is forced to be valid at the node points (collocation scheme), which are not smooth, the unstable points depend on the mesh. Using (21) and the averaged solid angles, we expect the unstable points to appear around $\epsilon_r = -1.64$ and $\epsilon_r = -1.18$. This is clearly supported by the convergence rate of the iterative solution. The eigenvalues for the same meshes in the case of $\epsilon_r = -2 + i0.01$ are plotted in Fig. 5. We can see the accumulation points around the expected values as well as the isolated eigenvalues that correspond the static resonance around zero.
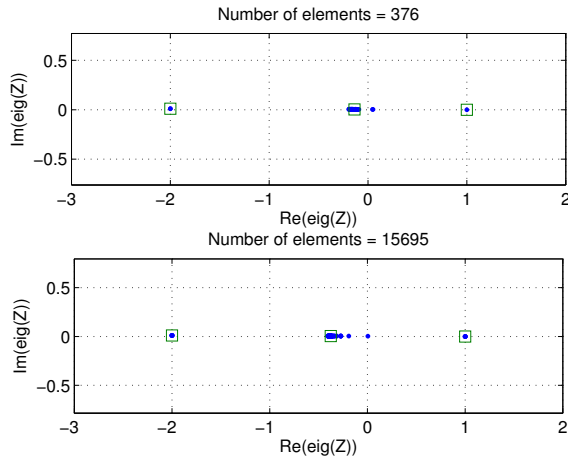


**Figure 5.** Computed eigenvalues of the system matrix for a dielectric ($\epsilon_r = -2 + i0.01$) sphere of size 0.01 (dots). The sphere is discretized with 376 (top) and 15695 (bottom) tetrahedral elements. The expected accumulation points of eigenvalues are computed from (21) using solid angles $\Omega_{r,i}$ averaged over all surface nodes (squares). The isolated eigenvalues associated to the static resonance are also visible around zero.

**Table 1.**  For a cube, the equation $\sigma_e(\Omega_r, \epsilon_r) = 0$ is satisfied with the arguments:

|         | $\Omega_r$ | $\epsilon_r$ |
| ------- | ---------- | ------------ |
| Volume  | $4\pi$     | 0            |
| Surface | $2\pi$     | -1           |
| Edge    | $\pi$      | -3           |
| Corner  | $\pi/2$    | -7           |

Next, we study the unstable solutions associated with the essential spectrum of a cube. We expect to find the unstable solutions when the essential spectrum (21) has an accumulation point at 0, i.e., $\sigma_e(\Omega_r, \epsilon_r) = 0$. The observation point can locate inside of the cube ($\Omega_r = 4\pi$), on the flat surface ($\Omega_r = 2\pi$), on the edge ($\Omega_r = \pi$), or on the corner ($\Omega_r = \pi/2$). Thus, we solve the equation $\sigma_e(\Omega_r, \epsilon_r) = 0$ for the permittivity in the cases of the above mentioned solid angles. The permittivities are shown in table 1.

The permittivities in table 1 correspond to the unstable solutions, hence the system matrix is not invertible for these values. In Fig. 6, the extinction cross section and the number of iterations are plotted for a small cube ($kl = 0.0162$) as a function of permittivity from 0 to -8 containing the unstable values. We have added a small imaginary part $\text{Im}\epsilon_r = 0.01$ to avoid theoretical instabilities. The cube is discretized with 4639 and 32307 elements.

We can see that the extinction cross section computed with the different meshes deviate from each other. This is not surprising since the fields inside the cube can be strongly localized or even singular near sharp corners and wedges. Thus, a fine mesh, especially near corners and wedges, is needed to obtain an accurate solution providing that the solution exists and is unique. We see several spikes in the extinction cross section that are associated with the accumulation points and isolated eigenvalues. The spikes at $\epsilon_r = -1, -3$, and $-7$ may be artificial since the integral equation is not solvable for these values, and the finite numerical accuracy of the method may spoil the solution near these values. Specifically, the use of the collocation method may be responsible of these spikes since the PVIE may not be satisfied in the proper function space. It is worth noting that the resonance at $\epsilon_r = 0$ is not excited by the planewave incidence, and as shown in [*Li and Chew*, 2006] the conjugate gradient type iterative solvers converge if the right-hand-side vector is not in the null-space of the matrix. The other spikes are the static resonances of a cube. *Fuchs* [1975] showed that there are six major static resonances between $\epsilon_r = -0.42$, and $\epsilon_r = -3.68$. However, as pointed out by *Helsing and Perfekt* [2013], some of them may be artifacts due to insufficient resolution or unintended rounding of the corners. Be that as it may, there are no such things as infinitely sharp corners or wedges described by the bulk permittivity in the physical world. Thus, rounding corners and wedges is a commonly used approach in the numerical analysis of non-smooth plasmonic particles. See, e.g., [*Wallén et al.*, 2010] and references therein.

Finally, it is of interest to look at the energy distribution of the solution near the accumulation points of the essential spectrum. In Fig. 7 we plot the absorbed power for the small cubes ($kl = 0.00162$) with permittivitites $\epsilon_r = 0 + i0.01$, (top left) $\epsilon_r = -1 + i0.01$ (top right), $\epsilon_r = -3 + i0.01$ (bottom left), and $\epsilon_r = -7 + i0.01$ (bottom right). The cube is discretized with 400k elements. We can see that the absorbed power is confined in the geometrical entities (volume, surfaces, edges, corners) which give rise to the zero eigenvalues with the given permittivity. From the mathematical point of view, the PVIE is not a Fredholm of index zero when the permittivity is 0, $-1, -3$, or $-7$. This means that the inverse of the system matrix is not bounded, and consequently, the solution is not unique or the energy is infinite. Fig. 7 shows the energy density is indeed very high at surfaces,
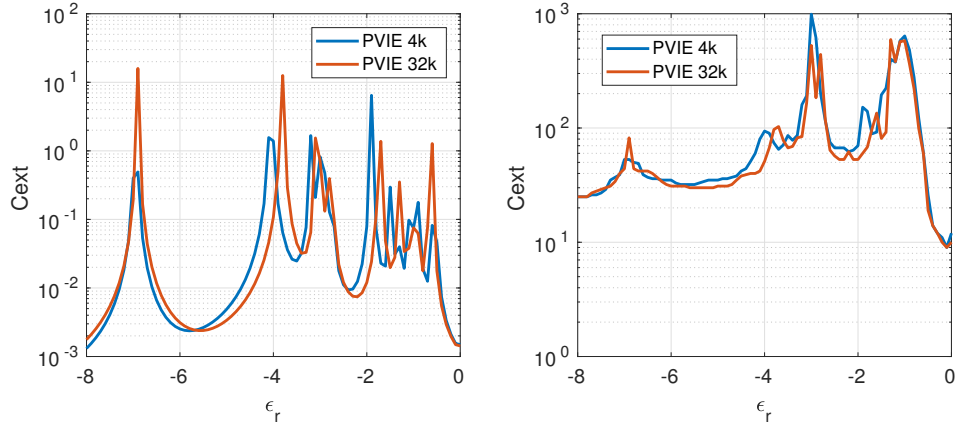
**Figure 6.** The extinction cross section (left) and the number of iterations (right) as functions of permittivity computed by the PVIE. The scatterer is a cube of size $kl = 0.0162$. The cube is discterized with 4639 (4k) and 32307 (32k) tetrahedral elements.

wedges, and corners depending on the permittivity which implies that the energy may not be bounded in these cases. This suggests that the solution obtained by the collocation method may diverge near the accumulation points since the equation may not be satisfied in $H^1$. Thus, a set of testing functions in the dual space of the range, i.e., $H^{-1}$, should be used to discretize the PVIE in order to remove the artificial solutions.

## 7 Conclusions

We have studied the potential volume-integral-equation method for the electromagnetic scattering analysis by dielectric objects. The formulation has two equations and two unknowns: the vector and scalar potentials. The PVIE allows us to use fully continuous basis functions which is favorable to the conditioning of the resulting system matrix.

We applied the collocation and Galerkin methods to discretize the system. For objects with smooth surfaces both schemes lead roughly to the same convergence rate of accuracy whereas the collocation method is more stable. The eigenvalues of the system matrix discretized with the collocation method coincide almost exactly with the theoretically predicted values. This is in contrast to the standard volume-integral-equation formulations where the eigenvalue spectra contain some artificial parts. Further, the eigenvalues are directly associated with the test functions which allows for building a simple scaling-based preconditioner without complicated Helmholtz decomposition as in [*Markkanen*, 2014].

The numerical experiments and theoretical considerations, however, imply that neither of the testing schemes (collocation or Galerkin), is the most optimal. They can give rise to spurious solution especially for non-smooth objects with negative permittivity. Thus, constructing a proper set of testing function is an interesting topic for future research.
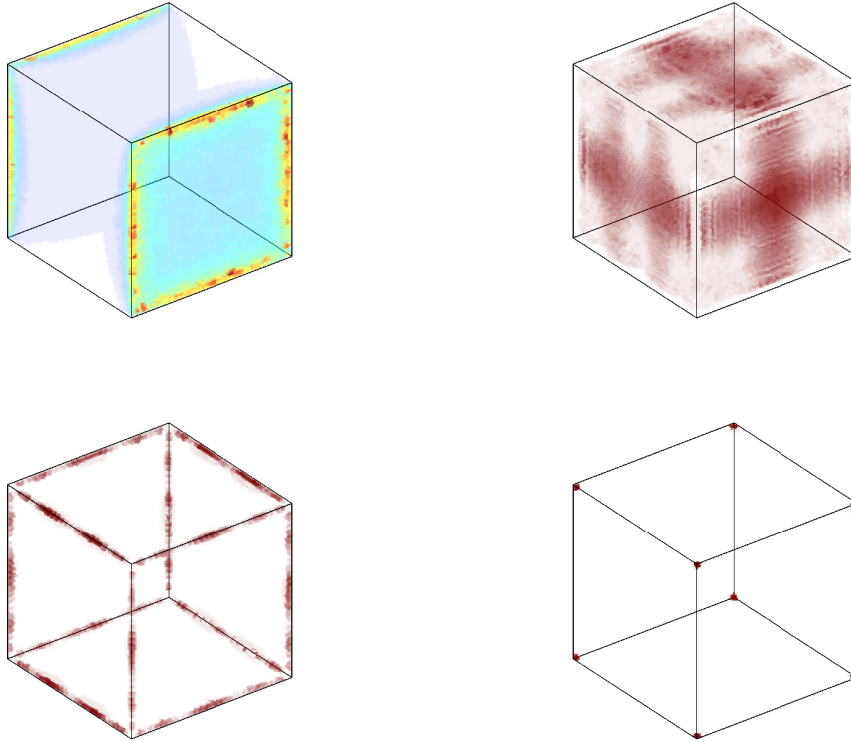
**Figure 7.** Absorbed power in a cube with $\epsilon_r = 0 + i0.01$ (top left), $\epsilon_r = -1 + i0.01$ (top right), $\epsilon_r = -3 + i0.01$ (bottom left), and $\epsilon_r = -7 + i0.01$ (bottom right).

the Fortran implementation of the method presented in the paper will be available upon request.

## References

Adil, M. (2013), Volume integral equations for the study of electromagnetic scattering by bi-anisotropic objects, Master's thesis, School of Electrical Engineering, Aalto University.

Bleszynski, E., M. Bleszynski, and T. Jaroszewicz (1996), AIM: Adaptive integral method for solving large-scale electromagnetic scattering and radiation problems, *Radio Science*, *31*, 1225–1252.

Botha, M. (2006), Solving the volume integral equations of electromagnetic scattering, *Journal of Computational Physics*, *218*(1), 141–158, doi:10.1016/j.jcp.2006.02.004.

Chang, R., and V. Lomakin (2011), Potential-based volume integral equations, *IEEE International Symposium on Antennas and Propagation (APSURSI)*, pp. 2712–2715.

Costabel, M. (2015), On the spectrum of volume integral operators in acoustic scattering, in *Integral Methods in Science and Engineering: Theoretical and Computational Advances*, edited by C. Constanda and A. Kirsch, pp. 119–127, Springer International Publishing, Cham.

Costabel, M., E. Darrigrand, and H. Sakly (2012), The essential spectrum of the volume integral operator in electromagnetic scattering by a homogeneous body, *Comptes Rendus Mathematique*, *350*(3), 193–197, doi:10.1016/j.crma.2012.01.017.

De Doncker, P. (2001), A potential integral equations method for electromagnetic scattering by penetrable bodies, *IEEE Transactions on Antennas and Propagation*, *49*(7), 1037–1042, doi:10.1109/8.933483.

DeDoncker, P. (2003), A volume/surface potential formulation of the method of moments applied to electromagnetic scattering, *Engineering analysis with boundary elements*, *27*(4), 325–331, doi:10.1016/S0955-7997(02)00120-0.

Fuchs, R. (1975), Theory of the optical properties of ionic crystal cubes, *Physical Review B*, *11*, 1732–1740, doi:10.1103/PhysRevB.11.1732.

Helsing, J., and K.-M. Perfekt (2013), On the polarizability and capacitance of the cube, *Applied and Computational Harmonic Analysis*, *34*(3), 445–468, doi: http://dx.doi.org/10.1016/j.acha.2012.07.006.

Järvenpää, S., M. Taskinen, and P. Ylä-Oijala (2003), Singularity extraction technique for integral equation methods with higher order basis functions on plane triangles and tetrahedra, *International Journal for Numerical Methods in Engineering*, *58*(8), 1149–1165.

Li, M.-K., and W. Chew (2006), Applying divergence-free condition in solving the volume integral equation, *Progress In Electromagnetics Research, PIER*, *57*, 311–333.

Markkanen, J. (2014), Discrete Helmholtz decomposition for electric current volume integral equation formulation, *IEEE Transactions on Antennas and Propagation*, *62*(12), 6282–6289, doi:10.1109/TAP.2014.2364614.

Markkanen, J. (2016), Volume potential-integral-equation formulation for electromagnetic scattering by dielectric objects, in *2016 URSI International Symposium on Electromagnetic Theory (EMTS)*, pp. 468–471, doi:10.1109/URSI-EMTS.2016.7571428.

Markkanen, J., and P. Ylä-Oijala (2016), Numerical comparison of spectral properties of volume-integral-equation formulations, *Journal of Quantitative Spectroscopy and Radiative Transfer*, *178*, 269–275, doi:http://dx.doi.org/10.1016/j.jqsrt.2016.01.020.

Markkanen, J., and A. Yuffa (2017), Fast superposition T-matrix solution for clusters with arbitrarily-shaped constituent particles, *Journal of Quantitative Spectroscopy and Radiative Transfer*, *189*, 181–188, doi:http://dx.doi.org/10.1016/j.jqsrt.2016.11.004.

Markkanen, J., P. Ylä-Oijala, and A. Sihvola (2012a), Discretization of volume integral equation formulations for extremely anisotropic materials, *IEEE Transactions on Antennas and Propagation*, *60*(11), 5195–5202, doi:10.1109/TAP.2012.2207675.

Markkanen, J., C.-C. Lu, X. Cao, and P. Yla-Oijala (2012b), Analysis of volume integral equation formulations for scattering by high-contrast penetrable objects, *IEEE Transactions on Antennas and Propagation*, *60*(5), 2367–2374, doi:10.1109/TAP.2012.2189704.

Phillips, J. R., and J. K. White (1997), A precorrected-FFT method for electrostatic analysis of complicated 3-D structures, *IEEE Transactions on CAD of Integrated Circuits and Systems*, *16*(10), 1059–1072, doi:10.1109/43.662670.

Schaubert, D., D. Wilton, and A. Glisson (1984), A tetrahedral modeling method for electromagnetic scattering by arbitrarily inhomogeneous dielectric bodies, *IEEE Transactions on Antennas and Propagation*, *32*(1), 77–85, doi:10.1109/TAP.1984.1143193.

Sun, L., and W. Chew (2009), A novel formulation of the volume integral equation for electromagnetic scattering, *Waves in Random and Complex Media*, *19*(1), 162–180, doi: 10.1080/17455030802545658.

van Beurden, M., and S. van Eijndhoven (2007), Gaps in present discretization schemes for domain integral equations, *International Conference on Electromagnetics in Advanced Applications, ICEAA 2007, Torino*.

van Beurden, M., and S. van Eijndhoven (2008), Well-posedness of domain integral equations for a dielectric object in homogeneous background, *Journal of Engineering Mathematics*, *62*, 289–302, doi:10.1007/s10665-008-9218-2.

Wallén, H., H. Kettunen, and A. Sihvola (2008), Surface modes of negative-parameter interfaces and the importance of rounding sharp corners, *Metamaterials*, *2*, pp. 113–121, doi:10.1016/j.metmat.2008.07.005.

Ylä-Oijala, P., S. Kiminki, and S. Järvenpää (2015), Conforming boundary element methods in acoustics, *Engineering Analysis with Boundary Elements*, *50*, 447–458, doi:

http://dx.doi.org/10.1016/j.enganabound.2014.10.002.

Yurkin, M. (2016), Performance of iterative solvers in the discrete dipole approximation, in *2016 URSI International Symposium on Electromagnetic Theory (EMTS)*, pp. 488–491, doi:10.1109/URSI-EMTS.2016.7571433.

Yurkin, M. A., M. Min, and A. G. Hoekstra (2010), Application of the discrete dipole approximation to very large refractive indices: Filtered coupled dipoles revived, *Physical Review E*, *82*(3), 036,703, doi:10.1103/PhysRevE.82.036703.