SOLVING THE 3D HARMONIC MAXWELL EQUATIONS WITH FINITE ELEMENTS, LAGRANGE MULTIPLIERS AND ITERATIVE METHODS

Laurent Crouzet
Gérard Meurant

CEA
Centre d'Etudes de Limeil-Valenton,
94195 Villeneuve St Georges cedex, France

Abstract. In this paper we consider the problem of the scattering of a plane wave by a three dimensional perfectly conducting bounded object. We embed the object in a sphere on the boundary of which we apply an absorbing boundary condition. The harmonic Maxwell equations are discretized with conforming P_1 finite elements in $[H^1(\Omega)]^3$ on tetrahedrons. The perfectly conducting boundary condition on the object is handled via Lagrange multipliers. This amounts to solving a sparse symmetric indefinite linear system with complex coefficients. Iterative methods like GMRES and Bi–CGSTAB are described for solving this problem with suitable efficient preconditioners. Numerical experiments for several model problems are given.

1. Description of the problem

We consider the scattering of a plane wave by a perfectly conducting three dimensional bounded object Ω. Our final goal is to compute the diffracted fields, that is solving Maxwell equations outside Ω. The general Maxwell equations are,

$$\begin{cases} \vec{curl}\vec{E} = -\dfrac{\partial \vec{B}}{\partial t} \\ \vec{curl}\vec{H} = \vec{J} + \dfrac{\partial \vec{D}}{\partial t} \\ div\vec{D} = \rho \\ div\vec{B} = 0 \end{cases}$$

where $\vec{E}$ is the electric field, $\vec{H}$ the magnetic field, $\vec{D}$ the electric induction, $\vec{B}$ the magnetic induction, $\vec{J}$ the current density and ρ the charge density. Moreover, we have the charge conservation equation:

$$div\vec{J} = -\frac{\partial \rho}{\partial t}.$$

1991 *Mathematics Subject Classification.* 49J40, 65F10, 65L60.
Key words and phrases. Maxwell equations, finite elements, iterative methods.

We also have the constitutive relations,

$$\begin{cases} \vec{D} = \epsilon\ \vec{E} \\ \vec{B} = \mu\ \vec{H} \\ \vec{J} = \sigma\ \vec{E} \end{cases}$$

where ϵ is the electric permittivity, μ the magnetic permeability and σ the electric conductivity.

We are looking for time harmonic solutions whose time dependency is $e^{-i\omega t}$. Then,

$$\begin{cases} \vec{curl}\vec{E} - i\omega\mu\vec{H} = 0 \\ \vec{curl}\vec{H} + i\omega\epsilon\vec{E} = \vec{J} \\ div\vec{J} - i\omega\rho = 0 \\ \vec{J} = \sigma\vec{E} \end{cases}$$

In the media outside the object Ω (usually air) there are no charges or currents. Therefore,

$$\vec{J} = 0,\ \rho = 0.$$

The object Ω is supposed to be perfectly conducting, so

$$\vec{E} = \vec{H} = 0 \text{ in } \Omega.$$

The interface condition (that is the boundary condition on the boundary Γ_1 of Ω) is

$$\vec{E} \times \vec{n} = 0,$$
$$\vec{H}.\vec{n} = 0.$$

The equations for the diffracted fields outside Ω are

$$\begin{cases} \vec{curl}\vec{E}^d - i\omega\mu\vec{H}^d = 0 \quad \text{outside } \Omega \\ \vec{curl}\vec{H}^d + i\omega\epsilon\vec{E}^d = 0 \quad \text{outside } \Omega \\ \vec{E}^d \times \vec{n} = -\vec{E}^{inc} \times \vec{n} \quad \text{on } \Gamma_1 \\ \vec{H}^d.\vec{n} = -\vec{H}^{inc}.\vec{n} \quad \text{on } \Gamma_1 \end{cases}$$

To have a unique solution we must also impose that we have a radiation condition at infinity, the so called Silver–Müller condition

$$\lim_{r\to\infty} r(\sqrt{\mu}\vec{H}^d \times \frac{\vec{r}}{r} - \sqrt{\epsilon}\vec{E}^d) = 0,$$

With straightforward algebraic manipulations (and dropping the index d for simplicity), we finally have the problem we want to solve,

$$\begin{cases} \vec{curl}\ \vec{curl}\vec{E} - k^2\vec{E} = 0 \\ div(\vec{E}) = 0 \\ \vec{E} \times n = g\ on\ \Gamma_1 \\ \lim_{r\to\infty} r(\vec{curl}\vec{E} \times \frac{\vec{r}}{r} - ik\vec{E}) = 0 \end{cases}$$

2. The variational formulations.

There are many ways to solve the problem described in the first section. All the techniques transform the problem in the unbounded region outside Ω to a problem on a bounded set. A common way to do so is to embed the object Ω into a sphere of boundary Γ_2, to consider only the bounded open set Ω_{ext} between Γ_1 and Γ_2 and to apply an artificial boundary condition that mimics the Silver–Müller condition on Γ_2 ([3],[4]). Conditions of these types are called absorbing boundary conditions and are devised in order to have no reflected waves coming back from Γ_2 to perturb the solution.

To obtain a variational formulation, we use the Green's formula:

$$\int_\Omega F.(\vec{curl}V)\,dx = \int_\Omega (\vec{curl}F).V\,dx + \int_\Gamma (F\times n).V\,d\Gamma$$

Then, we have

$$\int_\Omega (\vec{curl}\ \vec{curl}E - k^2E)\bar{E}\ dx = 0$$

$$\Leftrightarrow \int_\Omega \vec{curl}E.\vec{curl}\bar{E} - k^2E.\bar{E}\ dx + \int_\Gamma (\bar{E}\times n).(\vec{curl}E)\ d\Gamma = 0$$

We introduce the following spaces,

$$\begin{aligned}
H(\Omega,curl) &= \{q \mid q\in L^2(\Omega)^3,\quad \vec{curl}\ q\in L^2(\Omega)^3\}\\
H(\Omega,div) &= \{q \mid q\in L^2(\Omega)^3,\quad div\ q\in L^2(\Omega)\}\\
X_g &= \{q \mid q\in H(\Omega,curl),\quad q\times n = g \text{ on } \Gamma_1\}\\
Y_g &= \{q \mid q\in H(\Omega,curl)\cap H(\Omega,div),\quad q\times n = g \text{ on } \Gamma_1\}\\
Z_g &= \{q \mid q\in H^1(\Omega)^3,\quad q\times n = g \text{ on } \Gamma_1\}
\end{aligned}$$

Because $\bar{E}\times n = 0$ on Γ_1, we have

$$\int_\Omega (\vec{curl}E.\vec{curl}\bar{E} - k^2E.\bar{E})\ dx + \int_{\Gamma_2} (\bar{E}\times n).(\vec{curl}E)\ d\Gamma_2 = 0$$

$$\Leftrightarrow \int_\Omega (\vec{curl}E.\vec{curl}\bar{E} - k^2E.\bar{E})\ dx + \int_{\Gamma_2} (n\times \vec{curl}E).\bar{E}\ d\Gamma_2 = 0$$

So, classically, the first variational formulation is

$$\begin{cases} \text{Find } E\in X_g \text{ such that :}\\ \displaystyle\int_\Omega (\vec{curl}E.\vec{curl}\bar{E} - k^2E.\bar{E})\ dx + \int_{\Gamma_2} (n\times \vec{curl}E).\bar{E}\ d\Gamma_2 = 0,\quad \forall \bar{E}\in X_0. \end{cases}$$

This formulation can be used to construct a method of approximation based on finite elements which are conforming in $H(\Omega,curl)$. If we want to use classical finite elements, we need to introduce another variational formulation, see [1], [6].

First, to ensure the divergence free constraint, following [5], we add a divergence term to the variational formulation and we consider

$$\left\{\begin{array}{l}\text{Find } E \in Y_g \text{ such that :}\\ \displaystyle\int_\Omega (\vec{curl}E.\vec{curl}\bar{E} - k^2 E.\bar{E} \ + t\ divE.div\bar{E})\ dx\\ \displaystyle + \ \int_{\Gamma_2} (n \times \vec{curl}E).\bar{E}\ d\Gamma_2 = 0, \quad \forall \bar{E} \in Y_0.\end{array}\right.$$

If E has enough regularity, we can replace Y by Z. As it is not easy to handle the condition on $E \times n$ in the functional spaces, we introduce a Lagrange multiplier λ, $\lambda \in H^{-\frac{1}{2}}(\Gamma)^3$. Then, we obtain the final variational formulation

Find $\{E, \lambda\} \ \in H^1(\Omega)^3 \times H^{-\frac{1}{2}}(\Gamma)^3$ such that :

$$\begin{array}{rl} & \displaystyle\int_\Omega (\vec{curl}E.\vec{curl}\bar{E} - k^2 E.\bar{E} \ + t\ divE.div\bar{E})\, dx\\ + & \displaystyle\int_{\Gamma_2} (n \times \vec{curl}E).\bar{E}\ d\Gamma_2 + \int_{\Gamma_1} (\lambda \times n)(\bar{E} \times n)\, d\Gamma_1 = 0 \quad \forall \bar{E} \in H^1(\Omega)^3\\ & \displaystyle\int_{\Gamma_1} (\nu \times n)(E \times n - g)\, d\Gamma_1 = 0, \quad \forall \nu \in H^{-\frac{1}{2}}(\Gamma)^3\\ & \displaystyle\int_{\Gamma_1} (\lambda.n)\theta\, d\Gamma_1 = 0, \quad \forall \theta \in H^{\frac{1}{2}}(\Gamma)\end{array}$$

The last equation is needed to obtain the uniqueness of λ. It means that λ is tangent to Γ_1.

On Γ_2 we use a first order absorbing boundary condition that is introduced in the integral on Γ_2.

The domain between Γ_1 and Γ_2, Ω_{ext} is discretized with a classical triangulation $\mathcal{T}_h$ using tetrahedrons and we define the following finite dimensional spaces,

$$Z_h = \{F_h \in C^0(\Omega)^3,\ \forall K_h \in \mathcal{T}_h,\ F_{h|K_h} \in P_1(K_h)^3\},$$

$$M_h = \{\lambda \in C^0(\Gamma_1)^3,\ \forall K_h^{\Gamma_1} \in \mathcal{T}_h^{\Gamma_1},\ \lambda_{|K_h^{\Gamma_1}} \in P_1(K_h^{\Gamma_1})^3\}.$$

$\mathcal{T}_h^{\Gamma_1}$ is the trace of the triangulation on Γ_1. It can be shown that the discrete version of the last equation of the variational formulation is automatically verified by using a standard quadrature formula.

In the next Section, to describe the iterative methods we will use a simpler model problem that is the Helmholtz equation on the same domain as before with a first order absorbing boundary condition on Γ_2,

$$\left\{\begin{array}{rl} -\Delta u - k^2 u = 0 & \text{in } \Omega\\ u = g_1 & \text{on } \Gamma_1\\ \displaystyle\frac{\partial u}{\partial n} - ik\, u = g_2 & \text{on } \Gamma_2,\end{array}\right.$$

The variational formulation is then

$$\begin{cases} \text{Find } \{u,\lambda\} \in H^1(\Omega)\times H^{-1/2}(\Gamma_1) \text{ such that,} \\ \displaystyle\int_\Omega (\nabla u \nabla v - k^2 uv)\,dx - i\int_{\Gamma_2} k\,uv\,d\Gamma_2 = \int_{\Gamma_2} g_2 v\,d\Gamma_2 + \int_{\Gamma_1} \lambda v\,d\Gamma_1 \quad \forall v \in H^1(\Omega), \\ \displaystyle\int_{\Gamma_1} \mu(u-g_1)\,d\Gamma_1 = 0 \quad \forall \mu \in H^{-1/2}(\Gamma_1). \end{cases}$$

and we use classical P_1 approximations. The mesh we consider has 11664 tetrahedrons and 2440 vertices as well as 1944 triangles on the boundaries. We choose a value of $k = 0.5$ and we obtain an l_2 relative error of 10^{-4}.

3. The iterative methods for the Helmholtz equation

After discretization we obtain a linear system

$$\mathcal{A}\begin{pmatrix} u \\ \lambda \end{pmatrix} = \begin{pmatrix} A & B^t \\ B & 0 \end{pmatrix}\begin{pmatrix} u \\ \lambda \end{pmatrix} = \begin{pmatrix} b \\ c \end{pmatrix}$$

Here the matrix A has complex entries and its Hermitian part may be indefinite. However A is symmetric.

To solve this system we consider and compare two different iterative methods: GMRES [8] and Bi–CGSTAB [9] generalized to complex data (see [7]) and used with suitable preconditioners. We will see that solving for the preconditioner will introduce another level of inner iterations. We briefly recall GMRES and Bi–CGSTAB.

To solve $\tilde{\mathcal{A}}x = \tilde{f}$, the GMRES algorithm is defined as follows.

GMRES(nd)

x_0 given, $r_0 = \tilde{f} - \tilde{\mathcal{A}}x_0, v_1 = \frac{r_0}{\|r_0\|}$. For j=1 to nd, do

$$h_{i,j} = (\tilde{\mathcal{A}}v_j, v_i) \qquad i = 1,..,j$$

$$\hat{v}_{j+1} = \tilde{\mathcal{A}}v_j - \sum_{i=1}^{j} h_{i,j}v_i$$

$$h_{j+1,j} = \|\hat{v}_{j+1}\|$$

$$v_{j+1} = \frac{\hat{v}_{j+1}}{h_{j+1,j}}$$

$$V_k = (v_1, v_2, \dots, v_k)$$

$$\bar{H}_k = \begin{pmatrix} h_{1,1} & h_{1,2} & \dots & h_{1,n} \\ h_{2,1} & \ddots & & \\ 0 & \ddots & \ddots & \\ \vdots & \ddots & h_{k,k-1} & h_{k,k} \\ 0 & \dots & 0 & h_{k+1,k} \end{pmatrix}$$

The approximate solution is computed as :

$$x_{nd} = x_0 + V_{nd}\,y_{nd} \text{ where } y_{nd} \text{ minimizes } \|d_{nd} - \bar{H}_{nd}y\|$$

$$d_k = (\|r_0\| \ 0 \ \ldots \ 0)^t$$

Restart :

$$\text{Compute} \qquad r_{nd} = \tilde{f} - \tilde{\mathcal{A}} x_{nd}, \text{ stop if } \frac{\|r_{nd}\|}{\|r_0\|} < \epsilon$$

$$\text{otherwise do} \qquad x_0 = x_{nd}, \ \ v_1 = \frac{r_{nd}}{\|r_{nd}\|}, \text{ and restart.}$$

The Bi–CGSTAB algorithm is defined as

Bi-CGSTAB

x_0 given, $r_0 = \tilde{f} - \tilde{\mathcal{A}} x_0, p_0 = r_0, \hat{\omega}_0 = \beta = \alpha_0 = 1, v_0 = q_0 = 0$ For n=1, ... , do

$$\hat{\beta} = (p_n, r_{n-1})$$

$$\omega_n = (\hat{\beta}/\beta)(\hat{\omega}_{n-1}, \alpha_{n-1})$$

$$\beta = \hat{\beta}$$

$$q_n = r_{n-1} + \omega_n (q_{n-1} - \alpha_{n-1} v_{n-1})$$

$$v_n = \tilde{\mathcal{A}} q_n$$

$$\hat{\omega}_n = \hat{\beta}_n / (p, v_n)$$

$$s = r_{n-1} - \hat{\omega}_n v_n$$

$$t = \tilde{\mathcal{A}} s$$

$$\alpha_n = (t, s)/(t, t)$$

$$x_n = x_{n-1} + \hat{\omega}_n q_n + \alpha_n s$$

$$r_n = s - \alpha_n t$$

We define a preconditioned matrix $\tilde{\mathcal{A}} = M^{-1}\mathcal{A}$ and $\tilde{f} = M^{-1} f$. Following [2], the preconditioner M is chosen as

$$M = \begin{pmatrix} A & B^t \\ B & -\epsilon_{mat} I \end{pmatrix}.$$

At each iteration we have to solve systems like

$$Mz = r.$$

This is done also with an (unpreconditioned) iterative method (either GMRES or Bi–CGSTAB) directly on the system with M or by eliminating one variable and considering the matrix $A + \frac{1}{\epsilon_{mat}} B^t B$. However, ϵ_{mat} should be small enough in order for M to be as close as possible to $\mathcal{A}$ but not too small as in that case the condition number of $A + \frac{1}{\epsilon_{mat}} B^t B$ deteriorates.

Now, we have to consider 4 combinations of the outer and inner iterative methods. Another parameter to consider is the stopping criteria we choose for the inner iterations. This is governed by a parameter ϵ_{resol}. The numerical results for the model problem follow.

ϵ_{mat}	10^{-3}			10^{-4}		
ϵ_{resol}	10^{-3}	10^{-5}	10^{-7}	10^{-3}	10^{-5}	10^{-7}
$\#it_{out}$	-	-	-	11^*	10	11
$\#it_{in}$	(NC)	(NC)	(NC)	115	132	119
ϵ_{mat}	10^{-5}			10^{-6}		
ϵ_{resol}	10^{-3}	10^{-5}	10^{-7}	10^{-3}	10^{-5}	10^{-7}
$\#it_{out}$	10^*	8	7	-	-	4
$\#it_{in}$	34	48	64	(NC)	(NC)	350

$Bi-CGSTAB_{out}$ and $Bi-CGSTAB_{in}$

$\epsilon_{mat} = 10^{-5}$ and $\epsilon_{resol} = 10^{-5}$ seem a good choice.

ϵ_{mat}	10^{-3}			10^{-4}		
ϵ_{resol}	10^{-3}	10^{-5}	10^{-7}	10^{-3}	10^{-5}	10^{-7}
$\#it_{out}$	-	-	-	11^*	10	-
$\#it_{in}$	(NC)	(NC)	(NC)	5×30	5×35	(NC)
ϵ_{mat}	10^{-5}			10^{-6}		
ϵ_{resol}	10^{-3}	10^{-5}	10^{-7}	10^{-3}	10^{-5}	10^{-7}
$\#it_{out}$	7^*	7	-	18^*	7	-
$\#it_{in}$	5×7	5×9	(NC)	5×15	5×35	(NC)

$Bi-CGSTAB_{out}$ and $GMRES(5)_{in}$

Again $\epsilon_{mat} = 10^{-5}$ and $\epsilon_{resol} = 10^{-5}$ seem a good choice.

ϵ_{mat}	10^{-3}			10^{-4}		
ϵ_{resol}	10^{-3}	10^{-5}	10^{-7}	10^{-3}	10^{-5}	10^{-7}
$\#it_{out}$	-	-	-	2×9	2×11	2×8
$\#it_{in}$	(NC)	(NC)	(NC)	118	143	198
ϵ_{mat}	10^{-5}			10^{-6}		
ϵ_{resol}	10^{-3}	10^{-5}	10^{-7}	10^{-3}	10^{-5}	10^{-7}
$\#it_{out}$	2×7	2×4	2×7	-	-	2×6
$\#it_{in}$	33	86	283	(NC)	(NC)	284

$GMRES(2)_{out}$ and $Bi-CGSTAB_{in}$

$\epsilon_{mat} = 10^{-5}$ and $\epsilon_{resol} = 10^{-3}$ are a good choice.

ϵ_{mat}	10^{-3}			10^{-4}		
ϵ_{resol}	10^{-3}	10^{-5}	10^{-7}	10^{-3}	10^{-5}	10^{-7}
$\#it_{out}$	2×21	-	-	2×7	2×7	-
$\#it_{in}$	2×48	(NC)	(NC)	2×37	2×52	(NC)
ϵ_{mat}	10^{-5}			10^{-6}		
ϵ_{resol}	10^{-3}	10^{-5}	10^{-7}	10^{-3}	10^{-5}	10^{-7}
$\#it_{out}$	2×13	2×5	-	2×12	(NC)	-
$\#it_{in}$	2×11	2×32	(NC)	2×21	-	(NC)

$GMRES(2)_{out}$ and $GMRES(2)_{in}$

Once again, $\epsilon_{mat} = 10^{-5}$ and $\epsilon_{resol} = 10^{-5}$ give satisfactory results.

These computations where for a value of k smaller than the maximum that can be reached with the given mesh but other runs were done with different values of k up to the maximum value and the numbers of inner iterations were relatively insensitive to the value of k, only the number of outer iterations is depending on k. From the experiments it seems that the number of outer iterations is $O(\sqrt{k})$.

From the point of view of computing time, the smallest one is given by $Bi-CGSTAB_{out}$ and $GMRES_{in}$.

Another possibility is to obtain a complete or incomplete factorization $\mathcal{L}\mathcal{D}^{-1}\mathcal{L}^t$ of $A + \frac{1}{\epsilon_{mat}} B^t B$ and to set

$$M = \begin{pmatrix} \mathcal{L} & -\frac{1}{\epsilon_{mat}} B^t \\ 0 & I \end{pmatrix} \begin{pmatrix} \mathcal{D}^{-1} & 0 \\ 0 & -\epsilon_{mat} I \end{pmatrix} \begin{pmatrix} \mathcal{L}^t & 0 \\ -\frac{1}{\epsilon_{mat}} B & I \end{pmatrix}$$

If we use a complete decomposition, the solve for M is "exact". Providing there is enough storage available this gives the second best computing time.

4. Numerical results for Maxwell equations

To be able to compare to an exact solution, we use firstly a boundary condition

$$E \times n = g, \quad \text{on } \Gamma_2$$

in place of the absorbing boundary condition.

This problem is much harder to solve than the Helmholtz equation. To obtain convergence, we must precondition the inner iterative method. We use an incomplete factorization of $A + \frac{1}{\epsilon_{mat}} B^t B$ where we keep some fill in according to the size of the fills.

As an example, we use the combination $Bi-CGSTAB_{out}$ and $Bi-CGSTAB_{in}$. We note that we need to use smaller values of $\epsilon_{mat} = 10^{-11}$ than for the Helmholtz equation and $\epsilon_{resol} = 10^{-2}$.

We give the number of outer and inner iterations as a function of the percentage of eliminated entries. In all cases, we have a relative error of the order of 10^{-4}.

% elim	0.9	4.8	15.7	37.3	64.7	82.1	84.5	85.4
# it out	5	5	5	5	5	5	9	9
# it in	1	1	1	1	1	1	3^*	67^*

We see that we can drop a large part of the fill in (up to 80%) without affecting the number of outer and inner iterations which are quite small. In fact, we do not even need the inner iterations in that particular case.

Therefore, it seems we have an efficient method for solving the Maxwell equations with the given boundary condition. Right now, we are doing extensive numerical tests to verify these preliminary results and we are introducing in the code the absorbing boundary conditions. The numerical results will be described in a forthcoming report.

5. Conclusion

In this paper we have presented preliminary numerical results using nested iterative methods for solving the Helmholtz and Maxwell equations. For the later, we use classical conforming finite elements, the boundary condition on the perfectly conducting body being handled via a Lagrange multiplier.

The iterative methods we used are GMRES and Bi–CGSTAB generalized to complex matrices. For both problems, we obtain convergence in a small number of outer and inner iterations. This shows that this kind of approach allows to efficiently solve these difficult problems.

References

[1] F.Assous, P.Degond, E.Heintze, P.A.Raviart, J.Segre, *On a finite-element method for solving the three-dimensional Maxwell equations*, to appear in J. Comp. Physics (1993).

[2] O. Axelsson, *Preconditioning of indefinite problems by regularization*, SIAM J. Numer. Anal. **16 No 1** (1979), 58–69.

[3] A.Bayliss, E.Turkel, *Radiation boundary conditions for wave-like equations*, Comm. Pure Appl. Math. **33** (1980).

[4] B.Engquist, A.Majda, *Absorbing boundary conditions for the numerical simulation of waves*, Math. Comp. **31, No 139** (1977), 629–651.

[5] R.Glowinski, J.Périaux, *Solving the Maxwell équations by Lagrange multipliers*, Proceedings of the sixth international symposium on domain decomposition methods, 1992, to appear.

[6] E.Heintze, *Résolution des équations de Maxwell tridimensionnelles instationnaires par une méthode d'éléments finis conformes*, Thèse de Doctorat de l'Université de Paris 6 /yr 1992.

[7] P. Joly, G. Meurant, *Complex conjugate gradient methods*, Numerical Algorithms **4** (1993), 379–406.

[8] Y. Saad, M.H. Schultz, *Gmres,a generalized minimal residual algorithm for solving non symmetric linear systems*, SIAM J. Sci. Stat. Comput. **7** (1986), 856–869.

[9] H. Van der Vorst, *Bi-CGSTAB: a fast and smoothly converging variant of Bi-CG for the solution of nonsymmetric linear systems*, SIAM J. Sci. Stat. Comput. **13, No 2** (1992), 631–644.