# PRECONDITIONERS FOR THE DISCRETIZED TIME-HARMONIC MAXWELL EQUATIONS IN MIXED FORM*

CHEN GREIF† AND DOMINIK SCHÖTZAU‡

**Key words.** time-harmonic Maxwell equations, finite element methods, saddle point linear systems, preconditioners, Krylov subspace solvers

**AMS subject classifications.** 65F10, 65N30

**Abstract.** We introduce a new preconditioning technique for iteratively solving linear systems arising from finite element discretizations of the mixed formulation of the time-harmonic Maxwell equations. The preconditioners are motivated by spectral equivalence properties of the discrete operators, but are augmentation-free and Schur complement-free. We provide a complete spectral analysis, and show that the eigenvalues of the preconditioned matrix are strongly clustered. The analytical observations are accompanied by numerical results that demonstrate the scalability of the proposed approach.

**1. Introduction.** We introduce new preconditioners for linear systems arising from finite element discretization of the mixed formulation of the time-harmonic Maxwell equations in lossless media with perfectly conducting boundaries [4, 5, 15, 20]. The following model problem with constant coefficients is considered: find the vector field $u$ and the multiplier $p$ such that

$$\begin{aligned} \nabla \times \nabla \times u - k^2 u + \nabla p &= f && \text{in } \Omega, \\ \nabla \cdot u &= 0 && \text{in } \Omega, \\ u \times n &= 0 && \text{on } \partial\Omega, \\ p &= 0 && \text{on } \partial\Omega. \end{aligned} \tag{1.1}$$

Here $\Omega \subset \mathbb{R}^3$ is a simply connected polyhedron domain with a connected boundary $\partial\Omega$, and $n$ denotes the outward unit normal on $\partial\Omega$. The datum $f$ is a given generic source (not necessarily divergence-free), and the wave number satisfies $k^2 = \omega^2 \epsilon\mu$, where $\omega \geq 0$ is the temporal frequency, and $\epsilon$ and $\mu$ are positive permittivity and permeability parameters. We assume that $k^2$ is not a Maxwell eigenvalue and that

$$k^2 \ll 1.$$

The introduction of the scalar variable $p$ guarantees the stability and well-posedness of the equations as $k$ tends to 0, including the limit case $k = 0$; see the discussion in [5, Section 3].

Finite element discretization using Nédélec elements of the first kind [19] for the approximation of the vector field and standard nodal elements for the multiplier yields a saddle point linear system of the form

$$\begin{pmatrix} A - k^2 M & B^T \\ B & 0 \end{pmatrix} \begin{pmatrix} u \\ p \end{pmatrix} = \begin{pmatrix} g \\ 0 \end{pmatrix}, \tag{1.2}$$

*This work was supported in part by the Natural Sciences and Engineering Research Council of Canada.

†Department of Computer Science, University of British Columbia, Vancouver, B.C., V6T 1Z4, Canada, greif@cs.ubc.ca

‡Department of Mathematics, University of British Columbia, Vancouver, B.C., V6T 1Z2, Canada, schoetzau@math.ubc.ca

where now $u \in \mathbb{R}^n$ and $p \in \mathbb{R}^m$ are finite arrays representing the finite element approximations, and $g \in \mathbb{R}^n$ is the load vector associated with $f$. The matrix $A \in \mathbb{R}^{n\times n}$ is symmetric positive semidefinite with nullity $m$, and corresponds to the discrete curl-curl operator; $B \in \mathbb{R}^{m\times n}$ is a discrete divergence operator with full row rank, and $M \in \mathbb{R}^{n\times n}$ is the vector mass matrix.

It is possible to decouple (1.2) into two separate problems, using the discrete Helmholtz decomposition [18, Section 7.2.1]. For $p$ we obtain a standard Poisson equation, for which many efficient solution methods exist. Then, once $p$ is available, a way of dealing with the high nullity of the discrete curl-curl operator in the resulting equation for $u$ is by applying a procedure of augmentation: the matrix $A$ is replaced by $A_W = A + B^T W^{-1} B$, where $W \in \mathbb{R}^{m\times m}$ is a weight matrix, chosen so that $A_W$ is symmetric positive definite. This does not change the solution, due to the divergence-free condition $Bu = 0$. Popular choices for $W$ that have been considered in the literature are scaled identity matrices or lumped mass matrices; see [14, pp. 319–320] and references therein. A similar approach in the context of finite volume methods has been proposed in [12].

We note that if $k \neq 0$, a direct approach based on solving $(A - k^2 M)u = g$ automatically enforces $Bu = 0$, provided the right hand side is divergence-free. A multigrid technique for this case has been proposed in [8]. A regularization technique is introduced in [21] to deal with the case $k = 0$, whereby $A$ is replaced by $A + \sigma M$, where $\sigma$ is a regularization parameter. Algebraic multigrid is shown to converge even for small $\sigma$. The solution is divergence-free for divergence-free data, but it changes with the parameter.

Leaving the saddle point system intact is a viable approach that works naturally for the limiting case $k = 0$, which is our main interest in this paper. The linear system does not have to be modified or regularized even if its (1,1) block is singular. The block structure of the saddle point matrix lends itself to effective preconditioners. Indeed, there has been great progress in the last few years in solution methods for saddle point systems, and there is a number of available robust solution methods [2].

An Uzawa-type algorithm for the saddle point system, coupled with a domain decomposition approach, has been proposed in [17]. The original system is transformed into a new system by augmentation with the scalar Laplacian as a weight matrix and it is shown that the condition number of the resulting matrix grows logarithmically with respect to the ratio between the subdomain diameter and the mesh size. The method incorporates augmentation and is parameter-dependent. Its convergence properties rely on extreme eigenvalues of the augmented Schur complement, which may be difficult to evaluate.

In this paper we introduce a new block diagonal preconditioning technique for the iterative solution of the saddle point linear system. It is motivated by spectral equivalence properties similar to those in [17]. However, we avoid augmenting the original system and obtain a preconditioned matrix that is completely parameter-free, and does not rely on the (augmented) Schur complement even in the convergence analysis. We show several equivalence properties of the matrices, and present spectral bounds based on the stability constants of the differential operators.

Each iteration of our scheme requires solving a linear system one of whose associated matrices is $A + \gamma M$, where $\gamma > 0$ is given. For such systems solution techniques with linear complexity are available; see [1, 13, 16, 21] and references therein. We show that the spectral distribution of the preconditioned matrices is favorable for Krylov subspace solvers in terms of clustering of eigenvalues. We also derive explicit expressions for the eigenvectors in terms of the null vectors of the discrete operators $A$ and $B$, which makes the convergence analysis complete.

Our numerical results indicate that the proposed technique scales extremely well with the mesh size, both on uniformly and locally refined meshes. In this paper we only focus on the performance of the outer solver, and do not consider computational issues related to how to solve the inner iterations associated with (implicit) inversion of the preconditioner.

The remainder of the paper is structured as follows. In Section 2 we present the mixed finite element formulation, make some necessary definitions, and discuss the algebraic properties of the discrete operators. In Sections 3 and 4 we discuss spectral equivalence and augmentation. In Section 5 we introduce and analyze the proposed preconditioners. In Section 6 we provide numerical examples that confirm the analysis and demonstrate the scalability of our approach. Finally, in Section 7 we draw some conclusions.

**2. Mixed finite element formulation.** In this section we provide details on the finite element formulation leading to the saddle point system (1.2).

**2.1. Discretization.** To discretize (1.1), we consider conforming and shape-regular partitions $\mathcal{T}_h$ of $\Omega$ into tetrahedra $\{K\}$. We denote the diameter of the tetrahedron $K$ by $h_K$ for all $K \in \mathcal{T}_h$ and define $h = \max_{K\in\mathcal{T}_h} h_K$. Let $\mathcal{P}_\ell(K)$ be the space of polynomials of degree $\ell$ on $K$ and let $\mathcal{N}_\ell(K)$ be the space of Nédélec vector polynomials of the first kind [18, 19]. The index $\ell$ is chosen so that $\mathcal{P}_{\ell-1}(K)^3 \subset \mathcal{N}_\ell(K) \subset \mathcal{P}_\ell(K)^3$. For $\ell \geq 1$, the finite element spaces for the approximation of the electric field and the multiplier are taken as

$$
\begin{aligned}
V_h &= \{v_h \in H_0(\mathrm{curl}) \,|\, v_{h|K} \in \mathcal{N}_\ell(K), K \in \mathcal{T}_h\},\\
Q_h &= \{q_h \in H_0^1(\Omega) \,|\, q_{h|K} \in \mathcal{P}_\ell(K), K \in \mathcal{T}_h\}.
\end{aligned}
$$

Here we use the Sobolev space

$$
H_0(\mathrm{curl}) = \left\{\, v \in L^2(\Omega)^3 \;:\; \nabla\times v \in L^2(\Omega)^3\,,\; v\times n = 0 \text{ on } \partial\Omega \,\right\}.
$$

We consider the following finite element formulation: find $(u_h, p_h) \in V_h \times Q_h$ such that

$$
\begin{aligned}
\int_\Omega (\nabla\times u_h)\cdot(\nabla\times v_h)\,dx - k^2\int_\Omega u_h\cdot v_h\,dx + \int_\Omega v_h\cdot\nabla p_h\,dx &= \int_\Omega f\cdot v_h\,dx,\\
\int_\Omega u_h\cdot\nabla q_h\,dx &= 0
\end{aligned}
\tag{2.1}
$$

for all $(v_h, q_h) \in V_h \times Q_h$.

To transform (2.1) into matrix form, let $\langle\psi_j\rangle_{j=1}^n$ and $\langle\phi_i\rangle_{i=1}^m$ be standard finite element bases for the spaces $V_h$ and $Q_h$ respectively:

$$
V_h = \mathrm{span}\langle\psi_j\rangle_{j=1}^n, \qquad Q_h = \mathrm{span}\langle\phi_i\rangle_{i=1}^m. \tag{2.2}
$$

Define

$$
\begin{aligned}
A_{i,j} &= \int_\Omega (\nabla\times\psi_j)\cdot(\nabla\times\psi_i)\,dx, && 1\leq i,j\leq n,\\
M_{i,j} &= \int_\Omega \psi_j\cdot\psi_i\,dx, && 1\leq i,j\leq n,\\
B_{i,j} &= \int_\Omega \psi_j\cdot\nabla\phi_i\,dx, && 1\leq i\leq m,\ 1\leq j\leq n,
\end{aligned}
$$

and let $A \in \mathbb{R}^{n\times n}$, $M \in \mathbb{R}^{n\times n}$, and $B \in \mathbb{R}^{m\times n}$ be the corresponding matrices. Let us also define the scalar Laplace matrix on $Q_h$ as $L = (L_{i,j})_{i,j=1}^m \in \mathbb{R}^{m\times m}$, where

$$L_{i,j} = \int_\Omega \nabla\phi_j \cdot \nabla\phi_i \, dx. \tag{2.3}$$

We further introduce the load vector $g \in \mathbb{R}^n$ by setting

$$g_i = \int_\Omega f \cdot \psi_i \, dx, \qquad 1 \le i \le n,$$

where $f$ is the source term in (1.1). We identify finite element functions $u_h \in V_h$ or $p_h \in Q_h$ with their coefficient vectors $u = (u_1, \ldots, u_n)^T \in \mathbb{R}^n$ and $p = (p_1, \ldots, p_m)^T \in \mathbb{R}^m$, with respect to the bases (2.2). The finite element solution of (2.1) is computed by solving the saddle point linear system (1.2).

**2.2. Properties of the discrete operators.** Let us now present a few key properties of the operators, using the well-known discrete Helmholtz decomposition for Nédélec elements. To that end, note that $\nabla Q_h \subset V_h$, and let us introduce the matrix $C \in \mathbb{R}^{n\times m}$ by setting

$$\nabla\phi_j = \sum_{i=1}^n C_{i,j}\psi_i, \qquad j = 1, \ldots, m.$$

For a function $q \in Q_h$ given by $q_h = \sum_{j=1}^m q_j\phi_j$, we then have

$$\nabla q_h = \sum_{i=1}^n \sum_{j=1}^m C_{i,j} q_j \psi_i,$$

so that $u = Cq$ is the coefficient vector of $u_h = \nabla q_h$ in the basis $\langle\psi_i\rangle_{i=1}^n$.

We shall denote by $\langle\cdot,\cdot\rangle$ the standard Euclidean inner product in $\mathbb{R}^n$ or $\mathbb{R}^m$, and by null$(\cdot)$ the null space of a matrix. For a given positive (semi)definite matrix $W$ and a vector $x$, we define the (semi)norm

$$|x|_W = \sqrt{\langle Wx, x\rangle}.$$

PROPOSITION 2.1. *The following relations hold:*

*(i)* $\mathbb{R}^n = \mathrm{null}(A) \oplus \mathrm{null}(B)$.
*(ii) For any* $u \in \mathrm{null}(A)$ *there is a unique* $q \in \mathbb{R}^m$ *such that* $u = Cq$.
*(iii)* $\langle Mu, Cq\rangle = \langle Bu, q\rangle$ *for* $u \in \mathbb{R}^n$ *and* $q \in \mathbb{R}^m$.
*(iv)* $\langle MCp, Cq\rangle = \langle Lp, q\rangle$ *for* $p, q \in \mathbb{R}^m$.
*(v) Let* $u \in \mathrm{null}(A)$ *with* $u = Cp$. *Then* $|u|_M = |p|_L$.

*Proof.* The first two relations follow from the discrete Helmholtz decomposition [18, Section 7.2.1]. If $u_h$ and $\nabla q_h$ are the finite element functions associated with the vectors $u$ and $Cq$, then we have

$$\langle Mu, Cq\rangle = \int_\Omega u_h \cdot \nabla q_h \, dx = \langle Bu, q\rangle,$$

which shows (ii). Relation (iv) follows similarly, and (v) follows from (iv). ∎

Let us further show a few more properties of $C$, and connections to the other matrices we have.

PROPOSITION 2.2. *The following relations hold:*

*(i)* $AC = 0$.
*(ii)* $BC = L$.
*(iii)* $MC = B^T$.
*(iv)* *If the datum $f$ is divergence-free, then $C^T g = 0$.*

*Proof.* The first assertion is obvious since the null space of $A$ is equal to the range of $C$, by Proposition 2.1. The defining properties of $B$, $C$ and $L$ yield, for $1 \le i, j \le m$,

$$(BC)_{i,j} = \sum_{k=1}^{n} B_{i,k} C_{k,j} = \int_\Omega \left( \sum_{k=1}^{n} C_{k,j} \psi_k \right) \cdot \nabla \phi_i \, dx = \int_\Omega \nabla \phi_j \cdot \nabla \phi_i \, dx = L_{i,j}.$$

This shows identity (ii). The third one follows similarly. Finally, to see (iv), note that for $1 \le j \le m$, using integration by parts and the divergence-free condition, we obtain

$$(C^T g)_j = \sum_{i=1}^{n} C_{i,j} g_i = \int_\Omega f \cdot \nabla \phi_j \, dx = -\int_\Omega (\nabla \cdot f) \, \phi_j \, dx = 0.$$

This completes the proof. ◻

An orthogonality property with respect to the inner product $\langle M\cdot, \cdot \rangle$ is obtained as follows. Let $u_A \in \text{null}(A)$ and $u_B \in \text{null}(B)$. Setting $u_A = Cq$, we have

$$\langle M u_A, u_B \rangle = \langle M u_B, Cq \rangle = \langle B u_B, q \rangle = 0, \tag{2.4}$$

by relation (iii) in Proposition 2.1. Consequently, we also have the following result.

PROPOSITION 2.3. *Let $u = u_A + u_B$ with $u_A \in \text{null}(A)$ and $u_B \in \text{null}(B)$. Then we have $|u|_M^2 = |u_A|_M^2 + |u_B|_M^2$.*

Let us now present stability properties of the matrices $A$ and $B$. First, by the Cauchy-Schwarz inequality, we obviously have

$$|\langle Au, v \rangle| \le |u|_A |v|_A, \qquad u, v \in \mathbb{R}^n.$$

A similar continuity property holds for $B$:

$$|\langle Bv, q \rangle| \le |v|_M |q|_L, \qquad v \in \mathbb{R}^n, \; q \in \mathbb{R}^m. \tag{2.5}$$

Secondly, the matrix $A$ is positive definite on null$(B)$ and

$$\langle Au, u \rangle \ge \alpha \left( |u|_A^2 + |u|_M^2 \right), \qquad u \in \text{null}(B), \tag{2.6}$$

with a stability constant $\alpha$ which is independent of the mesh size and only depends on the shape regularity of the mesh and the approximation order $\ell$ [14, Theorem 4.7]. Note that, since $\langle Au, u \rangle = |u|_A^2$, we must have $0 < \alpha < 1$ and then also

$$|u|_A^2 \ge \bar{\alpha} |u|_M^2, \qquad u \in \text{null}(B), \tag{2.7}$$

with

$$\bar{\alpha} = \frac{\alpha}{1 - \alpha}. \tag{2.8}$$

Finally, the matrix $B$ satisfies the discrete inf-sup condition

$$\inf_{0 \ne q \in \mathbb{R}^m} \sup_{0 \ne v \in \text{null}(A)} \frac{\langle Bv, q \rangle}{|v|_M |q|_L} \ge \beta > 0, \tag{2.9}$$

with an inf-sup constant $\beta > 0$ that only depends on the domain $\Omega$; see [18, p. 179] or [14, p. 319].

The above stated properties and the theory of mixed finite element methods [3] ensure that (2.1) is well-posed and the saddle point system is uniquely solvable (provided that the mesh size is sufficiently small). Moreover, it can been shown that asymptotically the method is optimally convergent in the mesh size; see [18, Chapter 7].

**3. Spectral equivalence properties.** Consider the augmented matrix

$$A_L = A + B^T L^{-1} B, \tag{3.1}$$

where $L$ is the scalar Laplacian defined in (2.3). The spectral equivalence properties derived below motivate the preconditioners presented in Section 5.

Applying the discrete Helmholtz decomposition in Proposition 2.1, we have:

LEMMA 3.1. *Let $u = u_A + u_B$ with $u_A \in \mathrm{null}(A)$ and $u_B \in \mathrm{null}(B)$. Then*

$$|Bu|_{L^{-1}} = |u_A|_M.$$

*Proof.* From Proposition 2.1, we have $u_A = Cp$ for a vector $p \in \mathbb{R}^m$. Using the identity $BC = L$ in Proposition 2.2, we obtain

$$|Bu|_{L^{-1}} = \langle L^{-1}Bu, Bu\rangle = \langle L^{-1}Bu_A, Bu_A\rangle = \langle L^{-1}BCp, BCp\rangle = \langle Lp, p\rangle = |p|_L.$$

Since $|p|_L = |u_A|_M$, the result follows. □

As an immediate consequence of Lemma 3.1, we conclude that $B^T L^{-1} B$ and $M$ are spectrally equivalent on the null space of $A$.

COROLLARY 3.2. *For any $u$ in the null space of $A$ the following relation holds:*

$$\langle B^T L^{-1} Bu, u\rangle = \langle Mu, u\rangle.$$

THEOREM 3.3. *The matrices $A_L$ and $A + M$ are spectrally equivalent:*

$$\alpha \leq \frac{\langle A_L u, u\rangle}{\langle (A+M)u, u\rangle} \leq 1,$$

*for any $u \in \mathbb{R}^n$, where $0 < \alpha < 1$ is the coercivity constant in (2.6).*

By noticing that

$$\langle (A+M)u, u\rangle = |u|_A^2 + |u|_M^2,$$

the proof of Theorem 3.3 is readily obtained from the bounds in the subsequent lemma. We note that a similar result can be found in [17, Theorem 3.1].

LEMMA 3.4. *The following relations hold:*

*(i)* $|\langle A_L u, v\rangle| \leq \left(|u|_A^2 + |u|_M^2\right)^{1/2} \left(|v|_A^2 + |v|_M^2\right)^{1/2}$ *for* $u, v \in \mathbb{R}^n$.

*(ii)* $\langle A_L u, u\rangle \geq \alpha \left(|u|_A^2 + |u|_M^2\right)$ *for* $u \in \mathbb{R}^n$.

*In (ii)* $0 < \alpha < 1$ *is the coercivity constant given in (2.6).*

*Proof.* By Proposition 2.1, we may decompose $u$ and $v$ into $u = u_A + u_B$ and $v = v_A + v_B$ with $u_A$, $v_A \in \mathrm{null}(A)$ and $u_B$, $v_B \in \mathrm{null}(B)$. Furthermore, there are vectors $p$ and $q$ in $\mathbb{R}^m$ such that $u_A = Cp$ and $v_A = Cq$.

Let us show the first assertion. By the Cauchy-Schwarz inequality,

$$|\langle Au, v\rangle| \leq |u_B|_A |v_B|_A = |u|_A |v|_A.$$

Similarly, the Cauchy-Schwarz inequality, Lemma 3.1 and the orthogonality in Proposition 2.3 yield

$$|\langle B^T L^{-1} Bu, v\rangle| = |\langle L^{-1} Bu, Bv\rangle| \leq |Bu|_{L^{-1}} |Bv|_{L^{-1}} = |u_A|_M |v_A|_M \leq |u|_M |v|_M.$$

The first assertion readily follows from summing the last two inequalities and applying again the Cauchy-Schwarz inequality. To show the result in (ii), note that the stability property (2.6) of the matrix $A$ yields:

$$\langle Au, u\rangle = \langle Au_B, u_B\rangle \geq \alpha \left(|u_B|_A^2 + |u_B|_M^2\right).$$

From Lemma 3.1,

$$\langle L^{-1} Bu, Bu\rangle = |Bu|_{L^{-1}}^2 = |u_A|_M^2,$$

and hence

$$\langle A_L u, u\rangle \geq \alpha \left(|u_A|_M^2 + |u_B|_A^2 + |u_B|_M^2\right).$$

By the orthogonality relation in Proposition 2.3 we have $|u|_M^2 = |u_A|_M^2 + |u_B|_M^2$, from which relation (ii) follows. □

We end this section by pointing out a connection between $L$ and the Schur complement associated with $A_L$, $S = BA_L^{-1}B^T$. The matrices $S$ and $L$ are spectrally equivalent; we have

$$\alpha\beta^2 \leq \frac{\langle Sp, p\rangle}{\langle Lp, p\rangle} \leq \alpha^{-1},$$

for any $p \in \mathbb{R}^m$. Here, $\alpha$ and $\beta$ are the coercivity and inf-sup constants from (2.6) and (2.9), respectively. We provide a full proof in [10] and also refer the reader to [17, Theorem 3.3]. It is a consequence of Lemma 3.4, the inf-sup condition in (2.9), and standard arguments for mixed finite element methods [3]. As a consequence, the preconditioners we propose in Section 5 are closely related to block preconditioners that rely on forming approximations of the Schur complement. Such techniques have been successfully used in a variety of applications, notably for the discretized Stokes and Navier-Stokes equations [6, 7].

**4. Augmentation with the scalar Laplacian.** We now turn our attention to the linear system and consider augmentation with the Laplacian as a starting point. We will assume that $A - k^2M$ is nonsingular; this can always be achieved by choosing the mesh size sufficiently small [18, Corollary 7.3].

Consider the matrix of (1.2):

$$\mathcal{K} = \begin{pmatrix} A - k^2M & B^T \\ B & 0 \end{pmatrix}, \tag{4.1}$$

and define the symmetric positive definite block diagonal matrix

$$\mathcal{K}_L = \begin{pmatrix} A_L - k^2M & 0 \\ 0 & L \end{pmatrix}. \tag{4.2}$$

We stress that the original system is not augmented and that $\mathcal{K}_L$ will not be the preconditioner that we eventually use; it is only used to lay the theoretical basis and motivation for the preconditioning approach that we propose in Section 5.

THEOREM 4.1. *The matrix $\mathcal{K}_L^{-1}\mathcal{K}$ has two distinct eigenvalues, given by*

$$\mu_+ = 1\ ; \qquad \mu_- = -\frac{1}{1-k^2}\ ,$$

*with algebraic multiplicities $n$ and $m$ respectively.*

*Proof.* Since $\mathcal{K}_L$ is symmetric positive definite and $\mathcal{K}$ is symmetric, $\mathcal{K}_L^{-1}\mathcal{K}$ has a complete set of linearly independent eigenvectors that span $\mathbb{R}^{n+m}$. The corresponding eigenvalue problem is

$$\begin{pmatrix} A-k^2M & B^T \\ B & 0 \end{pmatrix}\begin{pmatrix} v \\ q \end{pmatrix} = \mu \begin{pmatrix} A-k^2M+B^TL^{-1}B & 0 \\ 0 & L \end{pmatrix}\begin{pmatrix} v \\ q \end{pmatrix}.$$

From the nonsingularity of $\mathcal{K}_L^{-1}\mathcal{K}$ it follows that $\mu \neq 0$. Substituting $q = \frac{1}{\mu}L^{-1}Bv$, we obtain for the first block row

$$\mu(A-k^2M)v + B^TL^{-1}Bv = \mu^2(A-k^2M+B^TL^{-1}B)v. \tag{4.3}$$

By inspection it is straightforward to see that any vector $v \in \mathbb{R}^n$ satisfies (4.3) with $\mu = 1$, and thus the latter is an eigenvalue of $\mathcal{K}_L^{-1}\mathcal{K}$, with eigenvectors of the form $(v, L^{-1}Bv)$, where $v \neq 0$. We claim that the eigenvalue $\mu = 1$ has algebraic multiplicity $n$. (That is, there is no other eigenvector associated with $\mu = 1$ in addition to the above set.) This can be concluded by using standard arguments to show that if the set of vectors $\{(v_i, L^{-1}Bv_i)\}_{i=1}^{n+r}$, with $r \geq 0$, are linearly independent then necessarily $\{v_i\}_{i=1}^{n+r}$ are also linearly independent, and the latter cannot be so unless $r = 0$.

It is possible according to Proposition 2.1 to form $v = v_A + v_B$, where $v_A \in \text{null}(A)$ and $v_B \in \text{null}(B)$. We now show that if an eigenvector $(v, \frac{1}{\mu}L^{-1}Bv) = (v_A + v_B, \frac{1}{\mu}L^{-1}Bv_A)$ has a nonzero $v_B$ component, then its associated eigenvalue must necessarily be $\mu = 1$. Noting that by (2.4)

$$\langle M(v_A+v_B), v_B\rangle = |v_B|_M^2,$$

after taking inner products of (4.3) with $v_B$ and dividing by $\mu$ we get

$$(\mu - 1)\left(|v_B|_A^2 - k^2|v_B|_M^2\right) = 0.$$

Since the symmetric matrix $A - k^2M$ is nonsingular, it follows that for $v_B \neq 0$ we must have $|v_B|_A^2 - k^2|v_B|_M^2 = \langle (A-k^2M)v_B, v_B\rangle \neq 0$, and hence $\mu = 1$.

Next, we argue that at least $2m$ of the vectors $v$ must have a nonzero $v_A$ component. Let us prove this by showing that assuming otherwise leads to a contradiction. Suppose the eigenvectors are given by $(v, \frac{1}{\mu}L^{-1}Bv)$ for a set of $n+m$ choices of $v$. If our argument does not hold, then more than $n-m$ eigenvectors satisfy $v = v_B$, and must be of the form $(v_B, 0)$. But since the null space of $B$ is of rank $n-m$, there cannot be more than this number of linearly independent vectors $(v_B, 0)$.

Since at least $2m$ of the eigenvectors satisfy $v_A \neq 0$, and since the multiplicity of $\mu = 1$ is $n$, it follows that at least $m$ of the eigenvectors associated with $\mu = 1$ satisfy $v_A \neq 0$. Thus, consider $m$ such vectors, $v = v_A + v_B$ with $v_A \neq 0$. Then (4.3) reads

$$\mu\left(Av_B - k^2M(v_A+v_B)\right) + B^TL^{-1}Bv_A = \mu^2\left(Av_B - k^2M(v_A+v_B) + B^TL^{-1}Bv_A\right).$$

Taking inner products with the vectors $v_A$ and noting that by (2.4)

$$\langle M(v_A+v_B), v_A\rangle = |v_A|_M^2,$$

and by Corollary 3.2 we have

$$\langle B^T L^{-1} B v_A, v_A\rangle = \langle M v_A, v_A\rangle,$$

it follows that

$$-(\mu^2-\mu)k^2|v_A|_M^2 + (\mu^2-1)|v_A|_M^2 = 0.$$

Hence we have

$$(1-k^2)\mu^2 + k^2\mu - 1 = 0, \tag{4.4}$$

from which it follows that $\mu_+ = 1$ and $\mu_- = -\frac{1}{1-k^2}$. We have thus shown that $\mu_-$ is the only possible eigenvalue that is not equal to 1, and its algebraic multiplicity must be equal to $m$. This completes the proof. ☐

The proof of Theorem 4.1 in fact shows that the eigenspace of $\mathcal{K}_L^{-1}\mathcal{K}$ can be expressed in terms of the null vectors of $A$ and $B$, as follows.

COROLLARY 4.2. *Let $\{v_i\}_{i=1}^m$ be a basis for the null space of $A$ and $\{z_i\}_{i=1}^{n-m}$ a basis for the null space of $B$. Then $\{(v_i, L^{-1}Bv_i)\}_{i=1}^m$ and $\{(z_i,0)\}_{i=1}^{n-m}$ are $n$ linearly independent eigenvectors associated with the eigenvalue $\mu_+$. The vectors $\{(v_i, -(1-k^2)L^{-1}Bv_i)\}_{i=1}^m$ are $m$ linearly independent eigenvectors associated with the eigenvalue $\mu_-$. Grouped together, those eigenvectors form a complete eigenspace that spans $\mathbb{R}^{n+m}$.*

From Theorem 4.1 it follows that if MINRES were to be used for solving (1.2), with $\mathcal{K}_L$ as a preconditioner, then convergence would require merely two iterations, if roundoff errors are ignored. However, forming $A_L$ may be too computationally costly. We mention also that the results in Theorem 5.2 can be extended to general algebraic settings, i.e. not necessarily to the Maxwell operator, as we show in [11].

**5. The proposed augmentation-free preconditioners.** Define

$$\mathcal{P}_M = A + \gamma M, \tag{5.1}$$

where $\gamma = 1-k^2$. For the saddle point system (1.2) we consider the preconditioner

$$\mathcal{P}_{M,L} = \begin{pmatrix} \mathcal{P}_M & 0 \\ 0 & L \end{pmatrix}. \tag{5.2}$$

Throughout, we will assume that preconditioned MINRES for the saddle point system is used. A crucial factor in the speed of convergence of this method is the distribution of the eigenvalues; strong clustering yields fast convergence [9, Section 3.1]. The choice of $\mathcal{P}_M$ and $\mathcal{P}_{M,L}$ is motivated by the spectral equivalence results given in Theorem 3.3 and the eigenvalue distribution observed in Theorem 4.1, which allow us to observe that $\mathcal{P}_M \approx A_L - k^2 M$ and $\mathcal{P}_{M,L} \approx \mathcal{K}_L$. Thus, the overall computational cost of the solution procedure will depend on the ability to efficiently solve linear systems whose associated matrices are $A+\gamma M$ and $L$ (or approximations thereof). For solving the former we refer the reader to [1, 13, 16, 21].

THEOREM 5.1. *The matrix*

$$\mathcal{P}_M^{-1}(A_L - k^2 M)$$

*has an eigenvalue $\mu = 1$ of algebraic multiplicity $m$. The rest of the eigenvalues are bounded as follows:*

$$\frac{\bar{\alpha} - k^2}{\bar{\alpha} + 1 - k^2} < \mu < 1, \tag{5.3}$$

*with $\bar{\alpha}$ defined in (2.8).*

*Proof.* The corresponding eigenvalue problem is

$$(A - k^2 M + B^T L^{-1} B)v = \mu \left(A + (1-k^2)M\right) v.$$

Suppose $v = v_A + v_B$, where $v_A \in \text{null}(A)$ and $v_B \in \text{null}(B)$. We then have

$$Av_B - k^2 M(v_A + v_B) + B^T L^{-1} B v_A = \mu \left(Av_B + (1-k^2)M(v_A + v_B)\right).$$

By linear independence considerations, there are at least $m$ vectors $v$ that satisfy $v_A \neq 0$. For $m$ such vectors, taking inner products with $v_A$ and noting that by Corollary 3.2

$$\langle B^T L^{-1} B v_A, v_A\rangle = |v_A|_M^2$$

and that by (2.4) we have

$$\langle M(v_A + v_B), v_A\rangle = \langle M v_A, v_A\rangle = |v_A|_M^2,$$

we get

$$\mu(1-k^2)|v_A|_M^2 = (1-k^2)|v_A|_M^2.$$

It follows that $\mu = 1$ is an eigenvalue of multiplicity $m$.

For the rest of the eigenvectors we must have $v_B \neq 0$, and now taking inner products with $v_B$ and noting that

$$\langle B^T L^{-1} B v_A, v_B\rangle = \langle L^{-1} B v_A, B v_B\rangle = 0,$$

and that by (2.4) we have

$$\langle M(v_A + v_B), v_B\rangle = \langle M v_B, v_B\rangle = |v_B|_M^2,$$

it follows that

$$(1-\mu)|v_B|_A^2 = \left((1-k^2)\mu + k^2\right)|v_B|_M^2. \tag{5.4}$$

It is impossible to have $\mu = 1$, since in this case (5.4) collapses into $|v_B|_M = 0$, which cannot hold for $v_B \neq 0$. We cannot have $\mu > 1$ either, since that would imply that in (5.4) the left hand side is negative but the right hand side is positive. (Recall that we assume $k \ll 1$.) We conclude that we must have $\mu < 1$.

From (2.7) we recall that for any $u \in \text{null}(B)$, $|u|_A^2 \geq \bar{\alpha}|u|_M^2$ with $\bar{\alpha} = \frac{\alpha}{1-\alpha} > 0$. Applying this to (5.4) we conclude $(1-k^2)\mu + k^2 \geq \bar{\alpha}(1-\mu)$, and since $1-k^2+\bar{\alpha} > 0$ we obtain (5.3). Since $\mu$ can be either equal to 1 or satisfy (5.3), but not simultaneously both, the algebraic multiplicities follow. ☐

THEOREM 5.2. *Let $\mathcal{K}$ be the saddle point matrix (4.1). Then $\mu_+ = 1$ and $\mu_- = -\frac{1}{1-k^2}$ are eigenvalues of the preconditioned matrix $\mathcal{P}_{M,L}^{-1}\mathcal{K}$, each with algebraic multiplicity $m$. The rest of the eigenvalues satisfy the bound (5.3).*

*Proof.* The eigenvalue problem for $\mathcal{P}_{M,L}^{-1}\mathcal{K}$ is

$$\begin{pmatrix} A-k^2M & B^T \\ B & 0 \end{pmatrix}\begin{pmatrix} v \\ q \end{pmatrix} = \mu \begin{pmatrix} A+(1-k^2)M & 0 \\ 0 & L \end{pmatrix}\begin{pmatrix} v \\ q \end{pmatrix}.$$

Setting $q = \frac{1}{\mu}L^{-1}Bv$ and multiplying the resulting equation for $v$ by $\mu$, we have

$$\left[(\mu^2-\mu)A + ((1-k^2)\mu^2 + k^2\mu)M\right]v = B^TL^{-1}Bv.$$

The rest of the proof follows by taking the same steps taken in the proof of Theorem 5.1. We get $m$ equations of the form

$$(1-k^2)\mu^2 + k^2\mu - 1 = 0,$$

from which $\mu_+$ and $\mu_-$ are obtained. Note that this quadratic equation is identical to equation (4.4) for the eigenvalues of $\mathcal{K}_L^{-1}\mathcal{K}$, cf. Theorem 4.1. This reinforces that $\mathcal{P}_{M,L}$ is an effective sparse approximation of $\mathcal{K}_L$. Obtaining the bound (5.3) is done in a way identical to the last part of the proof of Theorem 5.1. □

For $k = 0$ the result of Theorem 5.2 simplifies as follows.

COROLLARY 5.3. *For the preconditioned matrix $\mathcal{P}_{M,L}^{-1}\mathcal{K}$ with $k = 0$, there are eigenvalues $\mu_\pm = \pm 1$, with algebraic multiplicity $m$ each. The rest of the eigenvalues satisfy $\alpha < \mu < 1$.*

## 6. Numerical experiments.

Our numerical experiments were performed using MATLAB; for generating the meshes we used the PDE toolbox. We implemented the two-dimensional version of the time-harmonic Maxwell equations. The lowest order elements were used, i.e. $\ell = 1$. The solutions of the preconditioned systems in each iteration were computed exactly.

### 6.1. A smooth domain with a quasi-uniform grid.

In this example the domain is the unit square. Uniformly refined meshes were constructed. The refined mesh is obtained from the original one by dividing each triangle into four congruent ones. The number of elements and matrix sizes are given in Table 6.1.

| Grid | Nel | $n+m$ |
|---|---|---|
| G1 | 64 | 113 |
| G2 | 256 | 481 |
| G3 | 1024 | 1985 |
| G4 | 4096 | 8065 |
| G5 | 16384 | 32513 |
| G6 | 65536 | 130561 |
| G7 | 262144 | 523265 |

TABLE 6.1
*Number of elements (Nel) and the size of the linear systems $(n+m)$ for seven grids used in Example 6.1.*

First, we set the right hand side function so that the exact solution is given by

$$u(x,y) = \begin{pmatrix} u_1(x,y) \\ u_2(x,y) \end{pmatrix} = \begin{pmatrix} 1-y^2 \\ 1-x^2 \end{pmatrix}$$

and $p \equiv 0$. The datum $f$ in this case is divergence-free. We ran MINRES with the preconditioner $\mathcal{P}_{M,L}$. The counts of the outer iterations are given in Table 6.2. The inner iterations were solved by the conjugate gradient method, preconditioned with incomplete Cholesky factorization, using a tight convergence tolerance. As expected, the outer solver scales extremely well with hardly any sensitivity to the mesh size and the wave number.

| Grid | $k=0$ | $k=\frac{1}{8}$ | $k=\frac{1}{4}$ | $k=\frac{1}{2}$ |
|---|---|---|---|---|
| G1 | 5 | 5 | 5 | 5 |
| G2 | 5 | 5 | 5 | 5 |
| G3 | 5 | 5 | 5 | 5 |
| G4 | 6 | 6 | 5 | 6 |
| G5 | 6 | 6 | 6 | 6 |
| G6 | 6 | 6 | 6 | 6 |
| G7 | 6 | 6 | 6 | 6 |

TABLE 6.2

*Iteration counts for Example 6.1 with a divergence-free right hand side, for various meshes and values of $k$, using MINRES for solving the saddle point system with the preconditioner $\mathcal{P}_{M,L}$. The outer iteration was stopped once the initial relative residual was reduced by a factor of $10^{-10}$.*

| Grid | $k=0$ | $k=\frac{1}{8}$ | $k=\frac{1}{4}$ | $k=\frac{1}{2}$ |
|---|---|---|---|---|
| G1 | 5 | 5 | 5 | 5 |
| G2 | 6 | 6 | 6 | 6 |
| G3 | 6 | 6 | 6 | 6 |
| G4 | 6 | 6 | 6 | 7 |
| G5 | 7 | 7 | 7 | 7 |
| G6 | 7 | 7 | 7 | 7 |
| G7 | 7 | 7 | 7 | 7 |

TABLE 6.3

*Iteration counts for Example 6.1 with a right hand side that is not divergence-free, for various meshes and values of $k$, using MINRES for solving the saddle point system with the preconditioner $\mathcal{P}_{M,L}$. The outer iteration was stopped once the initial relative residual was reduced by a factor of $10^{-10}$.*

We also ran the saddle point solver on an example with a right hand side function that was not divergence-free. We took the same $u$ as above, and $p=(1-x^2)(1-y^2)$. The iteration counts are given in Table 6.3. As before, the solver scales very well. Figure 6.1 depicts the eigenvalues of the preconditioned matrix $\mathcal{P}_{M,L}^{-1}\mathcal{K}$ for grid G2 with $k=\frac{1}{4}$. This linear system has 481 degrees of freedom, with $n=368$ and $m=113$. As is expected from Theorem 5.2, the $m$ negative eigenvalues of the matrix are equal to $-\frac{1}{1-k^2}=-\frac{16}{15}=-1.0666\ldots$, and for the positive ones, $m$ of them are equal to 1 and the remaining $n-m$ eigenvalues are bounded away from 0 and below 1. In our computations we observed strong clustering beyond what can be concluded from Theorem 5.2. Three of the positive eigenvalues are between 0.7 and 0.9, with the smallest equal to 0.706..., and four additional ones are between 0.9 and 0.95. The remaining 361 eigenvalues are all between 0.95 and 1, with 113 of them identically equal to 1, again as is known by the same theorem. This clustering effect explains the fast convergence of the preconditioned iterative solver.

**6.2. An L-shaped domain with locally refined grids.** In this example we consider an L-shaped domain, as depicted in Figure 6.2. The meshes were locally

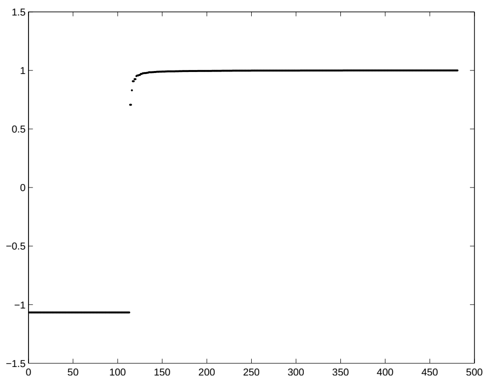


FIG. 6.1. *Plot of the eigenvalues of the preconditioned matrix $\mathcal{P}_{M,L}^{-1}\mathcal{K}$, for $k = \frac{1}{4}$, for grid G2 in Example 6.1.*

refined at the nonconvex corner at the origin; the number of elements and sizes are given in Table 6.4. Four of the five grids that were used are depicted in Figure 6.2. We set up the problem so that the right hand side function is equal to 1 throughout the domain. As in the previous example, we applied MINRES, preconditioned by $\mathcal{P}_{M,L}$, to the saddle point system. Table 6.5 demonstrates the scalability of the solvers: the outer iteration counts do not seem to be sensitive to changes in the mesh size.

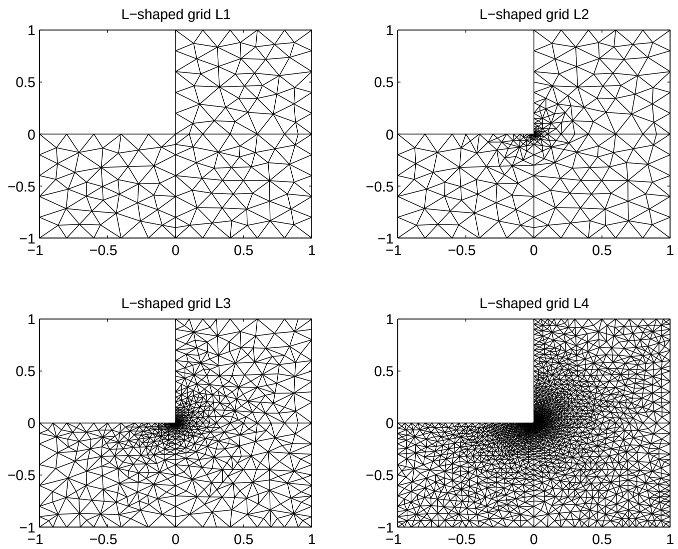


FIG. 6.2. *Grids L1 through L4 for Example 6.2.*

**7. Conclusions.** A new augmentation-free and Schur complement-free block diagonal preconditioner has been introduced for the discretized mixed formulation of the time-harmonic Maxwell equations. We present a complete spectral analysis. The outer iteration counts are hardly sensitive to changes in the mesh size or in small values of the wave number.

We have limited the discussion in this paper to exploring the convergence of the outer iteration, relying on the assumption that robust solution techniques exist for

| Grid | Nel | $n+m$ |
|---|---|---|
| L1 | 258 | 451 |
| L2 | 458 | 813 |
| L3 | 1403 | 2608 |
| L4 | 5164 | 9927 |
| L5 | 19339 | 37882 |

TABLE 6.4

*Number of elements (Nel) and the size of the linear systems* $(n+m)$ *for five grids used in Example 6.2.*

| Grid | $k=0$ | $k=\frac{1}{8}$ | $k=\frac{1}{4}$ | $k=\frac{1}{2}$ |
|---|---|---|---|---|
| L1 | 5 | 5 | 5 | 5 |
| L2 | 5 | 5 | 5 | 5 |
| L3 | 5 | 5 | 5 | 5 |
| L4 | 5 | 5 | 5 | 5 |
| L5 | 4 | 4 | 4 | 4 |

TABLE 6.5

*Iteration counts for Example 6.2 with various meshes and values of* $k$*, using MINRES for solving the saddle point system with the preconditioner* $\mathcal{P}_{M,L}$*. The outer iteration was stopped once the initial relative residual was reduced by a factor of* $10^{-10}$.

solving a system whose associated matrix is $A+\gamma M$. Future research will focus on exploring further computational aspects of our solution technique. We will explore using efficient inner solvers. Finally, we will explore whether similar preconditioners can be applied to problems in three dimensions and problems with variable coefficients.

## REFERENCES


[1] D.N. Arnold, R.S. Falk, and R. Winther. Multigrid in $H(\mathrm{div})$ and $H(\mathrm{curl})$. *Numer. Math.*, 85:197–217, 2000.

[2] M. Benzi, G.H. Golub, and J. Liesen. Numerical solution of saddle point problems. *Acta Numerica*, 14:1–137, 2005.

[3] F. Brezzi and M. Fortin. Mixed and hybrid finite element methods. In *Springer Series in Computational Mathematics*, volume 15. Springer–Verlag, New York, 1991.

[4] Z. Chen, Q. Du, and J. Zou. Finite element methods with matching and nonmatching meshes for Maxwell equations with discontinuous coefficients. *SIAM J. Numer. Anal.*, 37:1542–1570, 2000.

[5] L. Demkowicz and L. Vardapetyan. Modeling of electromagnetic absorption/scattering problems using *hp*–adaptive finite elements. *Comput. Methods Appl. Mech. Engrg.*, 152:103–124, 1998.

[6] H.C. Elman, D.J. Silvester, and A.J. Wathen. Iterative methods for problems in computational fluid dynamics. In R.H. Chan, T.F. Chan, and G.H. Golub, editors, *Iterative Methods in Scientific Computing*, pages 271–327. Springer-Verlag, Singapore, 1997.

[7] H.C. Elman, D.J. Silvester, and A.J. Wathen. *Finite Elements and Fast Iterative Solvers.* Oxford University Press, 2005.

[8] J. Gopalakrishnan, J.E. Pasciak, and L.F. Demkowicz. Analysis of a multigrid algorithm for time harmonic Maxwell equations. *SIAM J. Numer. Anal.*, 42(1):90–108, 2004.

[9] A. Greenbaum. *Iterative Methods for Solving Linear Systems.* SIAM, 1997.

[10] C. Greif and D. Schötzau. Preconditioners for the discretized time-harmonic Maxwell equations in mixed form. Technical Report TR-2006-01, Department of Computer Science, The University of British Columbia.

[11] C. Greif and D. Schötzau. Preconditioners for saddle point linear systems with highly singular (1,1) blocks. *ETNA, Special Volume on Saddle Point Problems*, 22:114–121, 2006.

[12] E. Haber, U.M. Ascher, D. Aruliah, and D. Oldenburg. Fast simulation of 3D electromagnetic problems using potentials. *J. Comput. Phys.*, 163:150–171, 2000.

[13] R. Hiptmair. Multigrid method for Maxwell's equations. *SIAM J. Numer. Anal.*, 36:204–225, 1998.

[14] R. Hiptmair. Finite elements in computational electromagnetism. *Acta Numerica*, 11:237–339, 2002.

[15] P. Houston, I. Perugia, A. Schneebeli, and D. Schötzau. Mixed discontinuous Galerkin approximation of the Maxwell operator: the indefinite case. *Modél. Math. Anal. Numér.*, (39):727–754, 2005.

[16] J.J. Hu, R.S. Tuminaro, P.B. Bochev, C.J. Garasi, and A.C. Robinson. Towards an $h$-independent algebraic multigrid method for Maxwell's equations. *SIAM J. Sci. Comput.*, to appear.

[17] Q. Hu and J. Zou. Substructuring preconditioners for saddle-point problems arising from Maxwell's equations in three dimensions. *Math. Comp.*, 73(245):35–61, 2004.

[18] P. Monk. *Finite element methods for Maxwell's equations.* Oxford University Press, New York, 2003.

[19] J.C. Nédélec. Mixed finite elements in $\mathbb{R}^3$. *Numer. Math.*, 35:315–341, 1980.

[20] I. Perugia, D. Schötzau, and P. Monk. Stabilized interior penalty methods for the time-harmonic Maxwell equations. *Comput. Methods Appl. Mech. Engrg.*, 191:4675–4697, 2002.

[21] S. Reitzinger and J. Schöberl. An algebraic multigrid method for finite element discretizations with edge elements. *Numer. Linear Algebra Appl.*, 9:223–238, 2002.