Chapter 3
Time-Domain Finite Element Methods for Metamaterials

In this chapter, we present several fully discrete mixed finite element methods for solving Maxwell's equations in metamaterials described by the Drude model and the Lorentz model. In Sects. 3.1 and 3.2, we respectively discuss the constructions of divergence and curl conforming finite elements, and the corresponding interpolation error estimates. These two sections are quite important, since we will use both the divergence and curl conforming finite elements for solving Maxwell's equations in the rest of the book. The material for Sects. 3.1 and 3.2 is quite classic, and we mainly follow the book by Monk (Finite element methods for Maxwell's equations. Oxford Science Publications, New York, 2003). After introducing the basic theory of divergence and curl conforming finite elements, we focus our discussion on developing some finite element methods for solving the time-dependent Maxwell's equations when metamaterials are involved. More specifically, in Sect. 3.3, we discuss the well posedness of the Drude model. Then in Sects. 3.4 and 3.5, we present detailed stability and error analysis for the Crank-Nicolson scheme and the leap-frog scheme, respectively. Finally, we extend our discussion on the well posedness, scheme development and analysis to the Lorentz model and the Drude-Lorentz model in Sects. 3.6 and 3.7, respectively.

3.1 Divergence Conforming Elements

3.1.1 Finite Element on Hexahedra and Rectangles

If a vector function has a continuous normal derivative, then such a finite element is usually called *divergence conforming*. More specifically, similar to the H^1 conforming finite elements discussed in Chap. 2, we can prove the following result.

J. Li and Y. Huang, *Time-Domain Finite Element Methods for Maxwell's Equations in Metamaterials*, Springer Series in Computational Mathematics 43,
DOI 10.1007/978-3-642-33789-5_3, © Springer-Verlag Berlin Heidelberg 2013

Lemma 3.1. *Let K_1 and K_2 be two non-overlapping Lipschitz domains having a common interface Λ such that $\overline{K_1} \cap \overline{K_2} = \Lambda$. Assume that $\mathbf{u}_1 \in H(div; K_1)$ and $\mathbf{u}_2 \in H(div; K_2)$, and $\mathbf{u} \in (L^2(K_1 \cup K_2 \cup \Lambda))^d$ be defined by*

$$\mathbf{u} = \begin{cases} \mathbf{u}_1 & on\ K_1, \\ \mathbf{u}_2 & on\ K_2. \end{cases}$$

Then $\mathbf{u}_1 \cdot \mathbf{n} = \mathbf{u}_2 \cdot \mathbf{n}$ on Λ implies that $\mathbf{u} \in H(div; K_1 \cup K_2 \cup \Lambda)$, where $\mathbf{n}$ is the unit normal vector to Λ.

Proof. Suppose that we have a function $\mathbf{u} \in (L^2(K_1 \cup K_2 \cup \Lambda))^d$ defined by $\mathbf{u}|_{K_i} = \mathbf{u}_i, i = 1, 2$, and $\mathbf{u}_1 \cdot \mathbf{n} = \mathbf{u}_2 \cdot \mathbf{n}$ on Λ. To prove that $\mathbf{u} \in H(\text{div}; K_1 \cup K_2 \cup \Lambda)$, we only need to show that $\nabla \cdot \mathbf{u} \in L^2(K_1 \cup K_2 \cup \Lambda)$. For any function $\phi \in C_0^\infty(K_1 \cup K_2 \cup \Lambda)$, using integration by parts, we have

$$\int_{K_1 \cup K_2 \cup \Lambda} \mathbf{u} \cdot \nabla \phi d\mathbf{x}$$
$$= -\int_{K_1} \nabla \cdot (\mathbf{u}|_{K_1}) \phi d\mathbf{x} - \int_{K_2} \nabla \cdot (\mathbf{u}|_{K_2}) \phi d\mathbf{x} + \int_{\Lambda} (\mathbf{u}_1 \cdot \mathbf{n}_1 + \mathbf{u}_2 \cdot \mathbf{n}_2) \phi ds,$$

where $\mathbf{n}_1$ and $\mathbf{n}_2$ denote the unit outward normals to ∂K_1 and ∂K_2, respectively.

Denote a function v such that $v|_{K_l} = \nabla \cdot (\mathbf{u}|_{K_l}), l = 1, 2$. Using the assumption that $\mathbf{u}_1 \cdot \mathbf{n} = \mathbf{u}_2 \cdot \mathbf{n}$ on Λ, we see that the boundary integral term vanishes. Hence, we have

$$\int_{K_1 \cup K_2 \cup \Lambda} \mathbf{u} \cdot \nabla \phi d\mathbf{x} = -\int_{K_1 \cup K_2 \cup \Lambda} v \phi d\mathbf{x},$$

which shows that $\nabla \cdot \mathbf{u} \in L^2(K_1 \cup K_2 \cup \Lambda)$ by the definition of weak derivative. This concludes our proof. □

Now let us consider a divergence conforming element on a reference hexahedron.

Definition 3.1. For any integer $k \geq 1$, the Nédélec divergence conforming element is defined by the triple:

$$\hat{K} = (0, 1)^3,$$
$$P_{\hat{K}} = Q_{k,k-1,k-1} \times Q_{k-1,k,k-1} \times Q_{k-1,k-1,k},$$
$$\Sigma_{\hat{K}} = M_{\hat{f}}(\hat{\mathbf{u}}) \cup M_{\hat{K}}(\hat{\mathbf{u}}),$$

where $M_{\hat{f}}(\hat{\mathbf{u}})$ is the set of degrees of freedom given on all faces $\hat{f}_i$ of $\hat{K}$, each with the outward normal $\mathbf{n}_i$:

$$M_{\hat{f}}(\hat{\mathbf{u}}) = \{\int_{\hat{f}_i} \hat{\mathbf{u}} \cdot \hat{\mathbf{n}}_i q dA,\ \forall\ q \in Q_{k-1,k-1}(\hat{f}_i), i = 1, \cdots, 6\} \tag{3.1}$$

and $M_{\hat{K}}(\hat{\mathbf{u}})$ is the set of degrees of freedom given on the element $\hat{K}$:

$$M_{\hat{K}}(\hat{\mathbf{u}}) = \{\int_{\hat{K}} \hat{\mathbf{u}}\cdot\hat{\mathbf{q}} dV, \ \forall\, \hat{\mathbf{q}} \in Q_{k-2,k-1,k-1}\times Q_{k-1,k-2,k-1}\times Q_{k-1,k-1,k-2}\}. \quad (3.2)$$

First we want to prove that the Nédélec element defined in Definition 3.1 is indeed unisolvent.

Theorem 3.1. *The degrees of freedom (3.1) and (3.2) uniquely determine a vector function* $\hat{\mathbf{u}} \in Q_{k,k-1,k-1} \times Q_{k-1,k,k-1} \times Q_{k-1,k-1,k}$ *on* $\hat{K} = (0,1)^3$.

Proof. Our proof follows [217].

(i) First we show that if all the face degrees of freedom in (3.1) on a face (say $\hat{f}_i$) are zero, then $\hat{\mathbf{u}}\cdot\hat{\mathbf{n}}_i = 0$ on this face. Considering that all faces are parallel to the coordinate axes, on any face $\hat{f}_i$, $\hat{\mathbf{u}}\cdot\hat{\mathbf{n}}_i \in Q_{k-1,k-1}$. Hence choosing $\hat{\mathbf{q}} = \hat{\mathbf{u}}\cdot\hat{\mathbf{n}}_i$ in (3.1) immediately leads to $\hat{\mathbf{u}}\cdot\hat{\mathbf{n}}_i = 0$.

(ii) Now let us consider the unisolvence. Note that the dimension of P_K is $3k^2(k+1)$, which equals the total number of degrees of freedom in $\Sigma_{\hat{K}}$. Hence we just need to prove that vanishing all degrees of freedom for $\hat{\mathbf{u}} \in P_{\hat{K}}$ yields $\hat{\mathbf{u}} = 0$. From Part (i), we know that $\hat{\mathbf{u}}\cdot\hat{\mathbf{n}}_i = 0$ on all faces, which implies that $\hat{\mathbf{u}}$ can be written as

$$\hat{\mathbf{u}} = (\hat{x}_1(1-\hat{x}_1)\hat{r}_1, \hat{x}_2(1-\hat{x}_2)\hat{r}_2, \hat{x}_3(1-\hat{x}_3)\hat{r}_3)^T,$$

where $\hat{r}_1 \in Q_{k-2,k-1,k-1}$, $\hat{r}_2 \in Q_{k-1,k-2,k-1}$, and $\hat{r}_3 \in Q_{k-1,k-1,k-2}$. Choosing $\hat{\mathbf{q}} = \hat{\mathbf{r}} \equiv (\hat{r}_1, \hat{r}_2, \hat{r}_3)^T$ in (3.2) shows that $\hat{\mathbf{r}} = \mathbf{0}$, which completes the proof. □

By trace theorem [2], we have $\hat{\mathbf{u}}|_{\hat{f}} \in (H^{\delta}(\hat{f}))^3 \subset (L^2(\hat{f}))^3$. Hence the degrees of freedom (3.1) and (3.2) are well defined for any $\hat{\mathbf{u}} \in (H^{\frac{1}{2}+\delta}(\hat{f}))^3, \delta > 0$.

After obtaining the basis function on the reference hexahedron $\hat{K}$, we can derive the basis function on a general element K by mapping. To make the degrees of freedom (3.1) and (3.2) invariant, we need the following special transformation

$$\mathbf{u} \circ F_K = \frac{1}{\det(B_K)} B_k \hat{\mathbf{u}}, \quad (3.3)$$

where F_K is the affine mapping defined in (2.18). For technical reasons, we assume that B_K is a diagonal matrix, i.e., the mapped element K has all edges parallel to the coordinate axes. The unit outward normal vector $\mathbf{n}$ to ∂K is obtained by the transformation [217, Eq. (5.21)]:

$$\mathbf{n} \circ F_K = \frac{1}{|B_K^{-T}\hat{\mathbf{n}}|} B_k^{-T}\hat{\mathbf{n}}, \quad (3.4)$$

where $\hat{\mathbf{n}}$ is the unit outward normal vector to $\partial\hat{K}$.

Lemma 3.2. *Suppose that* $det(B_K) > 0$ *and the function* $\mathbf{u}$ *and the normal* $\mathbf{n}$ *on* K *are obtained by the transformations (3.3) and (3.4), respectively. Then the degrees of freedom of* $\mathbf{u}$ *on* K *given by*

$$M_f(\mathbf{u}) = \{\int_{f_i} \mathbf{u} \cdot \mathbf{n}_i q dA, \ \forall\, q \in Q_{k-1,k-1}(f_i), \ i = 1, \cdots, 6\}, \tag{3.5}$$

$$M_K(\mathbf{u}) = \{\int_K \mathbf{u} \cdot \mathbf{q} dV, \ \forall\, \mathbf{q} \circ F_K = B_K^{-T} \hat{\mathbf{q}}, \ \textit{where}$$
$$\hat{\mathbf{q}} \in Q_{k-2,k-1,k-1} \times Q_{k-1,k-2,k-1} \times Q_{k-1,k-1,k-2}\}, \tag{3.6}$$

are identical to the degrees of freedom for $\hat{\mathbf{u}}$ *on* $\hat{K}$ *given in (3.1) and (3.2).*

Proof. (i) By the transformations (3.3) and (3.4), we have

$$\int_f \mathbf{u} \cdot \mathbf{n} q dA = \int_{\hat{f}} \frac{1}{\det(B_K)|B_K^{-T}\hat{\mathbf{n}}|} \hat{\mathbf{u}} \cdot \hat{\mathbf{n}} \hat{q} \frac{\text{area}(f)}{\text{area}(\hat{f})} d\hat{A} = \int_{\hat{f}} \hat{\mathbf{u}} \cdot \hat{\mathbf{n}} \hat{q} d\hat{A}, \tag{3.7}$$

where in the last step we used the fact that

$$\text{area}(f) = \det(B_K)|B_K^{-T}\hat{\mathbf{n}}|\text{area}(\hat{f}).$$

Equation (3.7) shows that the degrees of freedom $M_f(\mathbf{u})$ is invariant.

(ii) The invariance of $M_K(\mathbf{u})$ is easy to see by noting that

$$\begin{aligned}\int_K \mathbf{u} \cdot \mathbf{q} dV &= \int_{\hat{K}} \frac{1}{\det(B_K)} B_K \hat{\mathbf{u}} \cdot B_K^{-T} \hat{\mathbf{q}} \det(B_K)\, d\hat{V} \\ &= \int_{\hat{K}} \hat{\mathbf{u}} B_K^T \cdot B_K^{-T} \hat{\mathbf{q}} d\hat{V} = \int_{\hat{K}} \hat{\mathbf{u}} \cdot \hat{\mathbf{q}} d\hat{V}.\end{aligned}$$

□

Now suppose that we have a regular family of meshes of Ω denoted by T_h, and we form a global set of degrees of freedom by assembling the degrees of freedom from each element K in T_h, i.e.,

$$\Sigma = \cup_{K \in T_h} \Sigma_K.$$

If all neighboring elements match the whole common face (i.e., the face degrees of freedom match), then $\mathbf{u} \cdot \mathbf{n}$ is continuous by the proof of Part (i) in Theorem 3.1. Hence the finite element space W_h obtained by mapping the reference element in Definition 3.1 through transformation (3.3) is divergence conforming, i.e., W_h is a subset of $H(\text{div}, \Omega)$. Therefore, we can write W_h explicitly as

$$W_h = \{\mathbf{u}_h \in H(\text{div}, \Omega) : \mathbf{u}_h|_K \in \\ Q_{k,k-1,k-1} \times Q_{k-1,k,k-1} \times Q_{k-1,k-1,k}, \ \forall\, K \in T_h\}. \tag{3.8}$$

Similarly, we can define divergence conforming elements on rectangles.

Definition 3.2. For any integer $k \geq 1$, a divergence conforming element can be defined by the triple:

$$\hat{K} = (0,1)^2,$$
$$P_{\hat{K}} = Q_{k,k-1} \times Q_{k-1,k},$$
$$\Sigma_{\hat{K}} = M_{\hat{f}}(\hat{\mathbf{u}}) \cup M_{\hat{K}}(\hat{\mathbf{u}}),$$

where $M_{\hat{f}}(\hat{\mathbf{u}})$ is the set of degrees of freedom given on all faces $\hat{f}_i$ of $\hat{K}$, each with the outward normal $\mathbf{n}_i$:

$$M_{\hat{f}}(\hat{\mathbf{u}}) = \{\int_{\hat{f}_i} \hat{\mathbf{u}} \cdot \hat{\mathbf{n}}_i \, q dA, \ \forall\, q \in P_{k-1}(\hat{f}_i), i = 1, \cdots, 4\} \tag{3.9}$$

and $M_{\hat{K}}(\hat{\mathbf{u}})$ is the set of degrees of freedom given on the element $\hat{K}$:

$$M_{\hat{K}}(\hat{\mathbf{u}}) = \{\int_{\hat{K}} \hat{\mathbf{u}} \cdot \hat{\mathbf{q}} dV, \ \forall\, \hat{\mathbf{q}} \in Q_{k-2,k-1} \times Q_{k-1,k-2}\}. \tag{3.10}$$

Using the same technique as in the proof of Theorem 3.1, we can easily prove that the divergence element defined in Definition 3.2 is unisolvent.

Theorem 3.2. *The degrees of freedom (3.9) and (3.10) uniquely determine a vector function* $\hat{\mathbf{u}} \in Q_{k,k-1} \times Q_{k-1,k}$ *on* $\hat{K} = (0,1)^2$.

Below we present two often used divergence conforming elements: one for a cubic element; and one for a rectangular element.

Example 3.1. Choosing $k = 1$ in Definition 3.1, we know that $\mathbf{u}_{\hat{K}} \in Q_{1,0,0} \times Q_{0,1,0} \times Q_{0,0,1}$. Hence we can represent $\mathbf{u}_{\hat{K}}$ as follows:

$$\mathbf{u}_{\hat{K}} = (a_1 + b_1\hat{x}_1, a_2 + b_2\hat{x}_2, a_3 + b_3\hat{x}_3)^T,$$

where the constants can be determined by the six face degrees of freedom of (3.1).

If we label the six faces in the following order: front, right, back, left, bottom and top, then the outward normals are:

$$\mathbf{n}_1 = (0,-1,0)', \ \ \mathbf{n}_2 = (1,0,0)', \ \ \mathbf{n}_3 = (0,1,0)',$$
$$\mathbf{n}_4 = (-1,0,0)', \ \ \mathbf{n}_5 = (0,0,-1)', \ \ \mathbf{n}_6 = (0,0,1)',$$

substituting which into (3.1) gives all the coefficients:

$$a_1 = -\int_{left} \mathbf{u}\cdot\mathbf{n}_4 dA, \quad b_1 = \int_{right} \mathbf{u}\cdot\mathbf{n}_2 dA + \int_{left} \mathbf{u}\cdot\mathbf{n}_4 dA,$$

$$a_2 = -\int_{front} \mathbf{u}\cdot\mathbf{n}_1 dA, \quad b_2 = \int_{front} \mathbf{u}\cdot\mathbf{n}_1 dA + \int_{back} \mathbf{u}\cdot\mathbf{n}_3 dA,$$

$$a_3 = -\int_{bottom} \mathbf{u}\cdot\mathbf{n}_5 dA, \quad b_3 = \int_{bottom} \mathbf{u}\cdot\mathbf{n}_5 dA + \int_{top} \mathbf{u}\cdot\mathbf{n}_6 dA.$$

Hence we can write $\mathbf{u}_{\hat{K}}$ as

$$\begin{aligned}
\mathbf{u}_{\hat{K}}(\mathbf{x}) &= \begin{pmatrix} (\hat{x}_1 - 1)\int_{left} \mathbf{u}\cdot\mathbf{n}_4 dA + \hat{x}_1 \int_{right} \mathbf{u}\cdot\mathbf{n}_2 dA \\ (\hat{x}_2 - 1)\int_{front} \mathbf{u}\cdot\mathbf{n}_1 dA + \hat{x}_2 \int_{back} \mathbf{u}\cdot\mathbf{n}_3 dA \\ (\hat{x}_3 - 1)\int_{bottom} \mathbf{u}\cdot\mathbf{n}_5 dA + \hat{x}_3 \int_{top} \mathbf{u}\cdot\mathbf{n}_6 dA \end{pmatrix} \\
&= \text{`}(\int_{left} \mathbf{u}\cdot\mathbf{n}_4 dA)\mathbf{N}_{left}(\mathbf{x}) + (\int_{left} \mathbf{u}\cdot\mathbf{n}_2 dA)\mathbf{N}_{right}(\mathbf{x}) \\
&\quad + (\int_{left} \mathbf{u}\cdot\mathbf{n}_1 dA)\mathbf{N}_{front}(\mathbf{x}) + (\int_{left} \mathbf{u}\cdot\mathbf{n}_3 dA)\mathbf{N}_{back}(\mathbf{x}) \\
&\quad + (\int_{left} \mathbf{u}\cdot\mathbf{n}_5 dA)\mathbf{N}_{bottom}(\mathbf{x}) + (\int_{left} \mathbf{u}\cdot\mathbf{n}_6 dA)\mathbf{N}_{top}(\mathbf{x}),
\end{aligned}$$

where the basis functions $\mathbf{N}_{**}$ are as follows:

$$\mathbf{N}_{left} = \begin{pmatrix} \hat{x}_1 - 1 \\ 0 \\ 0 \end{pmatrix}, \ \mathbf{N}_{right} = \begin{pmatrix} \hat{x}_1 \\ 0 \\ 0 \end{pmatrix}, \ \mathbf{N}_{front} = \begin{pmatrix} 0 \\ \hat{x}_2 - 1 \\ 0 \end{pmatrix},$$

$$\mathbf{N}_{back} = \begin{pmatrix} 0 \\ \hat{x}_2 \\ 0 \end{pmatrix}, \ \mathbf{N}_{bottom} = \begin{pmatrix} 0 \\ 0 \\ \hat{x}_3 - 1 \end{pmatrix}, \ \mathbf{N}_{top} = \begin{pmatrix} 0 \\ 0 \\ \hat{x}_3 \end{pmatrix}.$$

Example 3.2. For rectangular elements, we can similarly define the divergence conforming finite element space W_h:

$$W_h = \{\mathbf{u}_h \in H(\text{div}, \Omega) : \ \mathbf{u}_h|_K \in Q_{k,k-1} \times Q_{k-1,k}, \ \forall \ K \in T_h\}. \tag{3.11}$$

For example, consider a rectangle $K = [x_c - h_x, x_c + h_x] \times [y_c - h_y, y_c + h_y]$. Then a function $\Pi_K^d \mathbf{u} \in W_h$ with $k = 1$ can be expressed as

$$\Pi_K^d \mathbf{u}(x, y) = \sum_{j=1}^{4} (\int_{l_j} \mathbf{u}\cdot\mathbf{n}_j dl)\mathbf{N}_j(x, y), \tag{3.12}$$

where l_j denote the four edges of the element K, which start from the bottom edge and are oriented counterclockwise. By satisfying the interpolation condition $\int_{l_j}(\Pi_K^d\mathbf{u}-\mathbf{u})\cdot\mathbf{n}_j\,dl=0$, we can obtain the face element basis functions $\mathbf{N}_j$ as follows:

$$\mathbf{N}_1=\begin{pmatrix}0\\ \dfrac{y-(y_c+h_y)}{4h_xh_y}\end{pmatrix},\quad \mathbf{N}_2=\begin{pmatrix}\dfrac{x-(x_c-h_x)}{4h_xh_y}\\ 0\end{pmatrix},$$

$$\mathbf{N}_3=\begin{pmatrix}0\\ \dfrac{y-(y_c-h_y)}{4h_xh_y}\end{pmatrix},\quad \mathbf{N}_4=\begin{pmatrix}\dfrac{x-(x_c+h_x)}{4h_xh_y}\\ 0\end{pmatrix}.$$

It is easy to check that the basis functions $\mathbf{N}_i$ satisfy the conditions:

$$\int_{l_j}\mathbf{N}_i\cdot\mathbf{n}_j\ dl=\delta_{ij},\quad i,j=1,\cdots,4.$$

3.1.2 Interpolation Error Estimates

From the unisolvence and divergence conforming property proved in last section, we see that with sufficient regularity, there exists a well-defined $H(div;\Omega)$ interpolation operator on K denoted as Π_K^d. For example, if we assume that $\mathbf{u}\in(H^{1/2+\delta}(K))^3,\delta>0$, then there is a unique function

$$\Pi_K^d\mathbf{u}\in Q_{k,k-1,k-1}\times Q_{k-1,k,k-1}\times Q_{k-1,k-1,k}$$

such that

$$M_f(\mathbf{u}-\Pi_K^d\mathbf{u})=0\ \text{ and }\ M_K(\mathbf{u}-\Pi_K^d\mathbf{u})=0,$$

where M_f and M_K are the sets of degrees of freedom in (3.5) and (3.6), respectively. More specifically, this is equivalent to requiring that: For all faces $f_i, i=1,\cdots,6$,

$$\int_{f_i}(\mathbf{u}-\Pi_K^d\mathbf{u})\cdot\mathbf{n}_i q\,dA=0,\ \ \forall\, q\in Q_{k-1,k-1}(f_i),\tag{3.13}$$

and

$$\int_K(\mathbf{u}-\Pi_K^d\mathbf{u})\cdot\mathbf{q}\,dV=0,\ \forall\,\mathbf{q}\circ F_K=B_K^{-T}\hat{\mathbf{q}},$$
$$\hat{\mathbf{q}}\in Q_{k-2,k-1,k-1}\times Q_{k-1,k-2,k-1}\times Q_{k-1,k-1,k-2}.\tag{3.14}$$

Before we prove the interpolation error estimate, we need to prove the following lemma, which shows that the interpolant on a general element K and the interpolation on the reference element $\hat{K}$ are closely related.

Lemma 3.3. *Suppose that* $\mathbf{u}$ *is sufficiently smooth such that* $\Pi_K^d \mathbf{u}$ *is well defined. Then under transformation (3.3), we have*

$$\widehat{\Pi_K^d \mathbf{u}} = \Pi_{\hat{K}}^d \hat{\mathbf{u}}.$$

Proof. By the definition of operator Π_K^d, we know that

$$M_f(\mathbf{u} - \Pi_K^d \mathbf{u}) = M_K(\mathbf{u} - \Pi_K^d \mathbf{u}) = 0,$$

which, along with the invariance of the degrees of freedom by transformation (3.3), leads to

$$M_{\hat{f}}(\mathbf{u} - \widehat{\Pi_K^d \mathbf{u}}) = M_{\hat{K}}(\mathbf{u} - \widehat{\Pi_K^d \mathbf{u}}) = 0.$$

Then by the unisolvence of the degrees of freedom, we obtain

$$\Pi_{\hat{K}}^d(\mathbf{u} - \widehat{\Pi_K^d \mathbf{u}}) = \Pi_{\hat{K}}^d(\hat{\mathbf{u}} - \widehat{\Pi_K^d \mathbf{u}}) = 0. \tag{3.15}$$

By the unisolvence again, we have $\Pi_{\hat{K}}^d(\widehat{\Pi_K^d \mathbf{u}}) = \widehat{\Pi_K^d \mathbf{u}}$, which together with (3.15) yields $\Pi_{\hat{K}}^d \hat{\mathbf{u}} = \widehat{\Pi_K^d \mathbf{u}}$. □

From the local interpolation operator Π_K^d, we can define a global interpolation operator

$$\Pi_h^d : (H^{\frac{1}{2}+\delta}(\Omega))^3 \to W_h, \forall\ \delta > 0,$$

element-wisely by

$$(\Pi_h^d \mathbf{u})|_K = \Pi_K^d(\mathbf{u}|_K) \quad \text{for each } K \in T_h.$$

The following theorem gives an error estimate for this interpolant.

Theorem 3.3. *Assume that* $0 < \delta < \frac{1}{2}$ *and* T_h *is a regular family of hexahedral meshes on* Ω *with faces aligning with the coordinate axes. Then if* $\mathbf{u} \in (H^s(\Omega))^3, \frac{1}{2} + \delta \le s \le k$, *there is a constant* $C > 0$ *independent of* h *and* $\mathbf{u}$ *such that*

$$||\mathbf{u} - \Pi_h^d \mathbf{u}||_{(L^2(\Omega))^3} \le Ch^s ||\mathbf{u}||_{(H^s(\Omega))^3}, \quad \frac{1}{2} + \delta \le s \le k. \tag{3.16}$$

Proof. For simplicity, here we only prove the result for integer $s = k \ge 1$. Proofs for more general cases can be found in other references (e.g., [5] for $\frac{1}{2} + \delta \le s < 1$).

As usual, we start with a local estimate on one element K. Using (3.3) and Lemma 2.5, we have

$$||\mathbf{u} - \Pi_K^d \mathbf{u}||^2_{(L^2(K))^3} = \int_K |\mathbf{u} - \Pi_K^d \mathbf{u}|^2 dV$$

$$= \int_{\hat{K}} |B_K(\hat{\mathbf{u}} - \widehat{\Pi_K^d \mathbf{u}})|^2 \frac{1}{|\det(B_K)|} d\hat{V}$$
$$\leq \frac{||B_K||^2}{|\det(B_K)|} ||\hat{\mathbf{u}} - \widehat{\Pi_K^d \mathbf{u}}||^2_{(L^2(\hat{K}))^3} \leq \frac{Ch_K^2}{|\det(B_K)|} ||\hat{\mathbf{u}} - \widehat{\Pi_K^d \mathbf{u}}||^2_{(L^2(\hat{K}))^3}.$$

By Lemma 3.3, the fact that

$$(I - \Pi_{\hat{K}}^d)\hat{\mathbf{p}} = 0 \qquad \forall\ \hat{\mathbf{p}} \in (Q_{k-1,k-1,k-1})^3,$$

and the Sobolev Embedding Theorem 2.1, we have

$$||\hat{\mathbf{u}} - \widehat{\Pi_K^d \mathbf{u}}||_{(L^2(\hat{K}))^3} = ||\hat{\mathbf{u}} - \Pi_{\hat{K}}^d \hat{\mathbf{u}}||_{(L^2(\hat{K}))^3}$$
$$= ||(I - \Pi_{\hat{K}}^d)(\hat{\mathbf{u}} + \hat{\mathbf{p}})||_{(L^2(\hat{K}))^3} \leq C||\hat{\mathbf{u}} + \hat{\mathbf{p}}||_{(H^k(\hat{K}))^3}.$$

Using the vector form of Lemma 2.6 to the previous inequality, we have

$$||\hat{\mathbf{u}} - \Pi_{\hat{K}}^d \hat{\mathbf{u}}||_{(L^2(\hat{K}))^3} \leq C \inf_{\hat{\mathbf{p}} \in (Q_{k-1,k-1,k-1})^3} ||\hat{\mathbf{u}} + \hat{\mathbf{p}}||_{(H^k(\hat{K}))^3} \leq C|\hat{\mathbf{u}}|_{(H^k(\hat{K}))^3}.$$

Combining the above estimates gives us

$$||\mathbf{u} - \Pi_K^d \mathbf{u}||_{(L^2(K))^3} \leq \frac{Ch_K}{|\det(B_K)|^{1/2}} |\hat{\mathbf{u}}|_{(H^k(\hat{K}))^3}. \tag{3.17}$$

Using (3.3), we have

$$|\hat{\mathbf{u}}|_{(H^k(\hat{K}))^3} = (\int_{\hat{K}} |\hat{\partial}_{\hat{x}}^k \hat{\mathbf{u}}|^2 d\hat{V})^{1/2}$$
$$= (\int_K |\partial_x^k (B_K^{-1}(det(B_K)\mathbf{u}))B_K^k|^2 \frac{dV}{|\det(B_K)|})^{1/2}$$
$$\leq |\det(B_K)|^{1/2} ||B_K^{-1}|| \cdot ||B_K||^k |\mathbf{u}|_{(H^k(K))^3}.$$

Substituting the previous estimate into (3.17) and using Lemma 2.5, we obtain

$$||\mathbf{u} - \Pi_K^d \mathbf{u}||_{(L^2(K))^3} \leq \frac{Ch_K}{|\det(B_K)|^{1/2}} \frac{|\det(B_K)|^{1/2}}{\rho_K} h_K^k |\mathbf{u}|_{(H^k(K))^3}.$$

Finally, substituting the previous estimate into the identity

$$||\mathbf{u} - \Pi_K^d \mathbf{u}||^2_{(L^2(\Omega))^3} = \sum_{K \in T_h} ||\mathbf{u} - \Pi_K^d \mathbf{u}||^2_{(L^2(K))^3}$$

and using the regularity of the mesh, we complete the proof. □

3.1.3 Finite Elements on Tetrahedra and Triangles

In this section, we will introduce some divergence conforming elements on tetrahedra and triangles. For tetrahedra, we define the space

$$D_k = (P_{k-1})^3 \oplus \tilde{P}_{k-1}\mathbf{x}. \tag{3.18}$$

Recall that $\tilde{P}_{k-1}$ represents the space of homogeneous polynomial of degree $k-1$.

It is easy to check that the dimension of D_k is

$$\begin{aligned} dim(D_k) &= 3 * dim(P_{k-1}) + dim(P_{k-1}) - dim(P_{k-2}) \\ &= 4 \cdot \frac{(k+2)(k+1)k}{3!} - \frac{(k+1)k(k-1)}{3!} = \frac{1}{2}(k+1)k(k+3). \end{aligned}$$

One interesting property about D_k is that $\nabla \cdot D_k \in P_{k-1}$.

Lemma 3.4. *Let D_k be the space defined by (3.18). Then $\nabla \cdot D_k \in P_{k-1}$.*

Proof. We can express any $\mathbf{u} \in D_k$ as $\mathbf{u}(\mathbf{x}) = \mathbf{p}(\mathbf{x}) + q(\mathbf{x})\mathbf{x}$, where $\mathbf{p}(\mathbf{x}) \in (P_{k-1})^3$ and $q(\mathbf{x}) \in \tilde{P}_{k-1}$. Hence

$$\begin{aligned} \nabla \cdot (q(\mathbf{x})\mathbf{x}) &= \partial_{x_1}(qx_1) + \partial_{x_2}(qx_2) + \partial_{x_3}(qx_3) \\ &= \nabla q \cdot \mathbf{x} + 3q = (k-1)q + 3q = (k+2)q, \end{aligned}$$

where we used the fact that $\nabla q \cdot \mathbf{x} = (k-1)q$ for any $q \in \tilde{P}_{k-1}$. Thus $\nabla \cdot (q(\mathbf{x})\mathbf{x}) \in P_{k-1}$, which along with the fact that $\nabla \cdot (P_{k-1})^3 \in P_{k-2}$ concludes the proof. □

Now let us construct a divergence conforming element on a reference tetrahedron $\hat{K}$, which has four vertices as $(0,0,0), (1,0,0), (0,1,0), (0,0,1)$. We assume that the four faces are labelled as that the outward unit normals of the first three faces (i.e., left, front and bottom) are

$$\mathbf{n}_1 = (-1,0,0)', \quad \mathbf{n}_2 = (0,-1,0)', \quad \mathbf{n}_3 = (0,0,-1)'.$$

Definition 3.3. For any integer $k \geq 1$, the divergence conforming element is defined by the triple:

$$\begin{aligned} &\hat{K} = \text{the reference tetrahedron}, \\ &P_{\hat{K}} = D_k, \\ &\Sigma_{\hat{K}} = M_{\hat{f}}(\hat{\mathbf{u}}) \cup M_{\hat{K}}(\hat{\mathbf{u}}), \end{aligned}$$

where $M_{\hat{f}}(\hat{\mathbf{u}})$ and $M_{\hat{K}}(\hat{\mathbf{u}})$ are the degrees of freedom defined as follows:

$$M_{\hat{f}}(\hat{\mathbf{u}}) = \{\int_{\hat{f}_i} \hat{\mathbf{u}} \cdot \hat{\mathbf{n}}_i \hat{q} d\hat{A}, \ \forall \hat{q} \in P_{k-1}(\hat{f}_i), i = 1, \cdots, 4\}, \tag{3.19}$$

$$M_{\hat{K}}(\hat{\mathbf{u}}) = \{\int_{\hat{K}} \hat{\mathbf{u}} \cdot \hat{\mathbf{q}} d\hat{V}, \ \forall \hat{\mathbf{q}} \in (P_{k-2})^3\}. \tag{3.20}$$

Below we show that this element is indeed unisolvent.

Theorem 3.4. *The degrees of freedom (3.19) and (3.20) uniquely define a vector function $\hat{\mathbf{u}} \in D_k$ on the reference tetrahedron $\hat{K}$.*

Proof. Note that the total number of degrees of freedom in Definition 3.3 is

$$\begin{aligned} 4 * dim(P_{k-1}(f)) + 3 * dim(P_{k-2}) &= 4 \cdot \frac{1}{2}(k+1)k + 3 \cdot \frac{1}{3!}(k+1)k(k-1) \\ &= \frac{1}{2}(k+1)k(k+3), \end{aligned}$$

which is the same as $dim(D_k)$. Hence to prove the unisolvence, we only need to show that vanishing all degrees of freedom for $\hat{\mathbf{u}} \in D_k$ gives $\hat{\mathbf{u}} = 0$. For simplicity, we drop the hat sign in the rest proof.

(i) First we prove that if all degrees of freedom (3.19) on a face vanish, then $\mathbf{u}\cdot\mathbf{n} = 0$ on that face. Since $\mathbf{u} \in D_k$, we can write $\mathbf{u} = \mathbf{p} + q\mathbf{x}$ for some $\mathbf{p} \in (P_{k-1})^3$ and $q \in \tilde{P}_{k-1}$. Assume that face f contains a point $\mathbf{a}$, then for any $\mathbf{x} \in f$, we have $(\mathbf{x} - \mathbf{a}) \cdot \mathbf{n} = 0$, where $\mathbf{n}$ is the unit outward normal to f. Hence, we obtain

$$\mathbf{u} \cdot \mathbf{n} = \mathbf{p} \cdot \mathbf{n} + q\mathbf{x} \cdot \mathbf{n} = \mathbf{p} \cdot \mathbf{n} + q\mathbf{a} \cdot \mathbf{n} \in P_{k-1}.$$

Therefore, choosing $q = \mathbf{u} \cdot \mathbf{n}$ in (3.19) leads to $\mathbf{u} \cdot \mathbf{n} = 0$.

(ii) From (i), we have $\mathbf{u} \cdot \mathbf{n} = 0$ on ∂K. For any $\phi \in P_{k-1}$, using integration by parts and the assumption of vanishing degrees of freedom (3.20), we obtain

$$\int_K \nabla \cdot \mathbf{u} \phi dV = \int_{\partial K} \mathbf{u} \cdot \mathbf{n} \phi dA - \int_K \mathbf{u} \cdot \nabla \phi dV = 0.$$

Choosing $\phi = \nabla \cdot \mathbf{u} \in P_{k-1}$ (by Lemma 3.4) yields $\nabla \cdot \mathbf{u} = 0$.

On the other hand, from proof of Lemma 3.4, for any $\mathbf{u} = \mathbf{p} + q\mathbf{x}$ with some $\mathbf{p} \in (P_{k-1})^3$ and $q \in \tilde{P}_{k-1}$, we have $\nabla \cdot \mathbf{u} = \nabla \cdot \mathbf{p} + (k+2)q$, which leads to $q = -\nabla \cdot \mathbf{p}/(k+2) \in P_{k-2}$. Hence $q = 0$, which yields

$$\mathbf{u} = \mathbf{p} \in (P_{k-1})^3. \tag{3.21}$$

If $k = 1$, (3.21) along with the condition $\mathbf{u} \cdot \mathbf{n} = 0$ on ∂K immediately implies that $\mathbf{u} = 0$.

If $k \geq 2$, then (3.21) and the fact $\mathbf{u} \cdot \mathbf{n} = 0$ on ∂K imply that

$$\mathbf{u} = (x_1 r_1, x_2 r_2, x_3 r_3)^T, \text{ for some } \mathbf{r} = (r_1, r_2, r_3)^T \in (P_{k-2})^3.$$

Choosing $\mathbf{q} = \mathbf{r}$ in the vanishing degrees of freedom (3.20) shows that $\mathbf{r} = 0$, hence $\mathbf{u} = 0$, which concludes our proof. □

The divergence conforming element on a general tetrahedron K can be obtained by mapping the finite element on the reference tetrahedron $\hat{K}$ given by Definition 3.3 through the transformation (3.3).

Lemma 3.5. *Suppose that $det(B_K) > 0$ and the function $\mathbf{u}$ and the normals $\mathbf{n}$ on K are obtained by the transformations (3.3) and (3.4). Then the degrees of freedom of $\mathbf{u}$ on K given by*

$$M_f(\mathbf{u}) = \{\int_{f_i} \mathbf{u} \cdot \mathbf{n}_i q dA, \ \forall\, q \in P_{k-1}(f_i), i = 1, \cdots, 4\}, \tag{3.22}$$

$$M_K(\mathbf{u}) = \{\int_K \mathbf{u} \cdot \mathbf{q} dV, \ \forall\, \mathbf{q} \circ F_K = B_K^{-T} \hat{\mathbf{q}}, \ \hat{\mathbf{q}} \in (P_{k-2})^3\} \tag{3.23}$$

are identical to the degrees of freedom for $\hat{\mathbf{u}}$ on $\hat{K}$ given in (3.19) and (3.20).

Given a regular family of tetrahedral meshes of Ω denoted as T_h, we can define a finite element space W_h using the degrees of freedom (3.22) and (3.23). From Part (i) in the proof of Theorem 3.4, we know that if two neighboring elements share the same face degrees of freedom, then $\mathbf{u} \cdot \mathbf{n}$ is continuous across the neighboring common face. Hence the finite element space W_h is globally divergence conforming, i.e., W_h is a subset of $H(\mathrm{div}, \Omega)$. Therefore, we can write W_h explicitly as

$$W_h = \{\mathbf{u}_h \in H(\mathrm{div}, \Omega) : \ \mathbf{u}_h|_K \in D_k, \ \forall\, K \in T_h\}. \tag{3.24}$$

By the same technique used for hexahedral element, we can define a global interpolation operator $\Pi_h^d : (H^{1/2+\delta}(\Omega))^3 \to W_h, \delta > 0$. The same interpolation error estimate as Theorem 3.3 holds true, and the proof is exactly the same as that carried out for Theorem 3.3. Details can consult Monk's book [217].

Below we show an example for the divergence conforming element defined in Definition 3.3 on the reference tetrahedron when $k = 1$.

Example 3.3. When $k = 1$, any function $\hat{\mathbf{u}}$ in the divergence conforming element can be expressed as

$$\hat{\mathbf{u}} = \mathbf{a} + b\hat{\mathbf{x}},$$

where the coefficients $\mathbf{a} = (a_1, a_2, a_3)^T$ and b can be determined by the four face degrees of freedom $\int_{f_i} \hat{\mathbf{u}} \cdot \hat{\mathbf{n}}_i dA, i = 1, \cdots, 4$. Note that for our reference tetrahedron, the unit outward normals are given by

$$\hat{\mathbf{n}}_1 = (-1,0,0)',\ \hat{\mathbf{n}}_2 = (0,-1,0)',\ \hat{\mathbf{n}}_3 = (0,0,-1)',\ \hat{\mathbf{n}}_4 = \frac{1}{\sqrt{3}}(1,1,1)'.$$

The condition $\int_{f_1} \hat{\mathbf{u}} \cdot \hat{\mathbf{n}}_1 dA$ gives us

$$\int_{f_1} \hat{\mathbf{u}} \cdot \hat{\mathbf{n}}_1 dA = -\int_{x_1=0,x_2,x_3\geq 0,x_2+x_3\leq 1} (a_1 + bx_1)dA = -a_1 \cdot \frac{1}{2},$$

which leads to $a_1 = -2\int_{f_1} \hat{\mathbf{u}} \cdot \hat{\mathbf{n}}_1 dA$.

By the same arguments, we obtain

$$a_2 = -2\int_{f_2} \hat{\mathbf{u}} \cdot \hat{\mathbf{n}}_2 dA,\ \ a_3 = -2\int_{f_3} \hat{\mathbf{u}} \cdot \hat{\mathbf{n}}_3 dA.$$

Note that face f_4 can be expressed by $x_1 + x_2 + x_3 = 1$, and has area $area(f_4) = \frac{\sqrt{3}}{2}$. Hence we have

$$\begin{aligned}\int_{f_4} \hat{\mathbf{u}} \cdot \hat{\mathbf{n}}_4 dA &= \int_{f_4} \frac{1}{\sqrt{3}}(a_1 + bx_1 + a_2 + bx_2 + a_3 + bx_3)dA \\ &= \int_{f_4} \frac{1}{\sqrt{3}}(a_1 + a_2 + a_3 + b)dA = \frac{1}{2}(a_1 + a_2 + a_3 + b),\end{aligned}$$

which leads to

$$b = 2\left(\int_{f_4} \hat{\mathbf{u}} \cdot \hat{\mathbf{n}}_4 dA + \int_{f_1} \hat{\mathbf{u}} \cdot \hat{\mathbf{n}}_1 dA + \int_{f_2} \hat{\mathbf{u}} \cdot \hat{\mathbf{n}}_2 dA + \int_{f_3} \hat{\mathbf{u}} \cdot \hat{\mathbf{n}}_3 dA\right).$$

Hence, the interpolation function on the reference element $\hat{K}$ can be written as

$$\Pi_{\hat{K}}^d \hat{\mathbf{u}} = \sum_{j=1}^{4}\left(\int_{f_j} \hat{\mathbf{u}} \cdot \mathbf{n}_j dA\right)\hat{\mathbf{N}}_j(\hat{\mathbf{x}}),$$

where the basis function $\hat{\mathbf{N}}_j$ are:

$$\hat{\mathbf{N}}_i = 2(\hat{\mathbf{x}} - \mathbf{e}_i),\ i = 1,2,3,\ \ \hat{\mathbf{N}}_4 = 2\hat{\mathbf{x}},$$

where $\mathbf{e}_i$ is the opposite vertex of each f_i, i.e., $\mathbf{e}_1 = (1,0,0)'$, $\mathbf{e}_2 = (0,1,0)'$, and $\mathbf{e}_3 = (0,0,1)'$. Moreover, it is easy to check that the basis functions satisfy the conditions $\int_{f_i} \hat{\mathbf{N}}_j \cdot \hat{\mathbf{n}}_i dA = \delta_{ij},\ \ i,j = 1,2,3,4.$

Example 3.4. Note that for triangular elements, the divergence conforming element space can still be defined using (3.24). The only difference is that we have to change three to two in the definition of D_k given by (3.18). Below we construct the lowest-order divergence conforming triangular element.

First let us consider a reference triangle $\hat{K}$, which is formed by vertices $\hat{A}_i, i = 1, 2, 3$, where

$$\hat{A}_1 = (0, 0),\ \hat{A}_2 = (1, 0),\ \hat{A}_3 = (0, 1).$$

The unit outward normal vectors are defined as follows:

$$\hat{\mathbf{n}}_1 = (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})',\ \ \hat{\mathbf{n}}_2 = (-1, 0)',\ \ \hat{\mathbf{n}}_3 = (0, -1)'.$$

The interpolation on the reference element $\hat{K}$ can be written as

$$\Pi^d_{\hat{K}}\mathbf{E} = \sum_{j=1}^{3}(\int_{l_j} \mathbf{E} \cdot \hat{\mathbf{n}}_j dl)\hat{\mathbf{N}}_j(\hat{x}, \hat{y}),$$

where the basis function $\hat{\mathbf{N}}_j = (a_1 + b\hat{x}, a_2 + b\hat{y})'$ satisfies the conditions

$$\int_{l_i} \hat{\mathbf{N}}_j \cdot \hat{\mathbf{n}}_i dl = \delta_{ij},\ \ i, j = 1, 2, 3.$$

It can be shown that the basis functions $\hat{\mathbf{N}}_j$ are:

$$\hat{\mathbf{N}}_1 = \begin{pmatrix} \hat{x} \\ \hat{y} \end{pmatrix},\ \ \hat{\mathbf{N}}_2 = \begin{pmatrix} -1 + \hat{x} \\ \hat{y} \end{pmatrix},\ \ \hat{\mathbf{N}}_3 = \begin{pmatrix} \hat{x} \\ -1 + \hat{y} \end{pmatrix}.$$

Then for a general triangle K with vertices $A_i = (x_i, y_i),\ i = 1, 2, 3$, we can use the affine mapping $F_K : \hat{\mathbf{x}} \rightarrow \mathbf{x}$ defined by

$$x = x_1 + (x_2 - x_1)\hat{x} + (x_3 - x_1)\hat{y},$$
$$y = y_1 + (y_2 - y_1)\hat{x} + (y_3 - y_1)\hat{y},$$

to map the reference element $\hat{K}$ to the element K.

Let $|K|$ be the area of K. After some lengthy algebra, we can find the inverse mapping F_K^{-1} of F_K as follows:

$$\hat{x} = 2|K|[(y_3 - y_1)(x - x_1) - (x_3 - x_1)(y - y_1)],$$
$$\hat{y} = 2|K|[-(y_2 - y_1)(x - x_1) + (x_2 - x_1)(y - y_1)],$$

from which we can obtain the basis function on K defined as:

$$\mathbf{N}_i(x, y) = \hat{\mathbf{N}}_i \circ F_K^{-1},\quad i = 1, 2, 3.$$

3.2 Curl Conforming Elements

3.2.1 Finite Element on Hexahedra and Rectangles

If a vector function has a continuous tangential component, then such a finite element is usually called *curl conforming*. Similar to the H^1 conforming finite elements, we can prove the following result.

Lemma 3.6. *Let K_1 and K_2 be two non-overlapping Lipschitz domains having a common interface Λ such that $\overline{K_1} \cap \overline{K_2} = \Lambda$. Assume that $\mathbf{u}_1 \in H(curl; K_1)$ and $\mathbf{u}_2 \in H(curl; K_2)$, and $\mathbf{u} \in (L^2(K_1 \cup K_2 \cup \Lambda))^3$ be defined by*

$$\mathbf{u} = \begin{cases} \mathbf{u}_1 & on\ K_1, \\ \mathbf{u}_2 & on\ K_2. \end{cases}$$

Then $\mathbf{u}_1 \times \mathbf{n} = \mathbf{u}_2 \times \mathbf{n}$ on Λ implies that $\mathbf{u} \in H(curl; K_1 \cup K_2 \cup \Lambda)$, where $\mathbf{n}$ is the unit normal vector to Λ.

Proof. The proof can be carried out in exactly the same way as that given for Lemma 3.1 by using the following identity: For any function $\boldsymbol{\phi} \in (C_0^\infty(K_1 \cup K_2 \cup \Lambda))^3$,

$$\begin{aligned} &\int_{K_1 \cup K_2 \cup \Lambda} \mathbf{u} \cdot \nabla \times \boldsymbol{\phi} d\mathbf{x} \\ &= \int_{K_1} \nabla \times \mathbf{u}_1 \cdot \boldsymbol{\phi} d\mathbf{x} + \int_{K_2} \nabla \times \mathbf{u}_2 \cdot \boldsymbol{\phi} d\mathbf{x} + \int_{\Lambda} (\mathbf{u}_1 \times \mathbf{n}_1 + \mathbf{u}_2 \times \mathbf{n}_2) \cdot \boldsymbol{\phi} ds, \end{aligned}$$

where $\mathbf{n}_i$ is the unit outward normal to ∂K_i, and $\mathbf{u}_i = \mathbf{u}|_{K_i}, i = 1, 2$. □

Let us consider the curl conforming elements on a reference hexahedron.

Definition 3.4. For any integer $k \geq 1$, the Nédélec curl conforming element is defined by the triple:

$$\begin{aligned} \hat{K} &= (0,1)^3, \\ P_{\hat{K}} &= Q_{k-1,k,k} \times Q_{k,k-1,k} \times Q_{k,k,k-1}, \\ \Sigma_{\hat{K}} &= M_{\hat{e}}(\hat{\mathbf{u}}) \cup M_{\hat{f}}(\hat{\mathbf{u}}) \cup M_{\hat{K}}(\hat{\mathbf{u}}), \end{aligned}$$

where $M_{\hat{e}}(\hat{\mathbf{u}})$ is the set of degrees of freedom (DOFs) given on all edges $\hat{e}_i$ of $\hat{K}$, each with the unit tangential vector $\hat{\boldsymbol{\tau}}_i$ in the direction of $\hat{e}_i$:

$$M_{\hat{e}}(\hat{\mathbf{u}}) = \{\int_{\hat{e}_i} \hat{\mathbf{u}} \cdot \hat{\tau}_i \hat{q} d\hat{s},\ \forall\, \hat{q} \in P_{k-1}(\hat{e}_i),\ i = 1, \cdots, 12\}, \tag{3.25}$$

$M_{\hat{f}}(\hat{\mathbf{u}})$ is the set of degrees of freedom given on all faces $\hat{f}_i$ of $\hat{K}$, each with the unit outward normal vector $\mathbf{n}_i$:

$$M_{\hat{f}}(\hat{\mathbf{u}}) = \{\int_{\hat{f}_i} \hat{\mathbf{u}} \times \hat{\mathbf{n}}_i \cdot \hat{\mathbf{q}} d\hat{A}, \ \forall\, \hat{\mathbf{q}} \in Q_{k-2,k-1}(\hat{f}_i) \times Q_{k-1,k-2}(\hat{f}_i), i = 1, \cdots, 6\}, \tag{3.26}$$

and $M_{\hat{K}}(\hat{\mathbf{u}})$ is the set of degrees of freedom given on the element $\hat{K}$:

$$M_{\hat{K}}(\hat{\mathbf{u}}) = \{\int_{\hat{K}} \hat{\mathbf{u}} \cdot \hat{\mathbf{q}} d\hat{V}, \ \forall\, \hat{\mathbf{q}} \in Q_{k-1,k-2,k-2} \times Q_{k-2,k-1,k-2} \times Q_{k-2,k-2,k-1}\}. \tag{3.27}$$

Hence we have a total of $12k$ edge DOFs, $6 \cdot 2(k-1)k$ face DOFs, and $3 \cdot k(k-1)^2$ element DOFs. It is easy to see that

$$\dim(P_{\hat{K}}) = 12k + 6 \cdot 2(k-1)k + 3 \cdot k(k-1)^2 = 3k(k+1)^2.$$

First we want to prove that the element defined in Definition 3.4 is indeed unisolvent.

Theorem 3.5. *The degrees of freedom (3.25)–(3.27) uniquely determine a vector function* $\mathbf{u} \in Q_{k-1,k,k} \times Q_{k,k-1,k} \times Q_{k,k,k-1}$ *on* $\hat{K} = (0,1)^3$.

Proof. (i) First we show that if all the face degrees of freedom (3.25) and (3.26) on a face (say $\hat{f}_i$) are zero, then $\hat{\mathbf{u}} \times \hat{\mathbf{n}}_i = \mathbf{0}$ on this face. Without loss of generality, let us consider face $\hat{x}_1 = 0$. On this face, noting that $\hat{\mathbf{u}} \times \hat{\mathbf{n}}_1 = -\hat{u}_3 \mathbf{j} + \hat{u}_2 \mathbf{k}$, hence the tangential components of $\hat{\mathbf{u}}$ on $\hat{x}_1 = 0$ are:

$$\hat{u}_3 \in Q_{k,k-1}(\hat{x}_2, \hat{x}_3), \quad \hat{u}_2 \in Q_{k-1,k}(\hat{x}_2, \hat{x}_3).$$

Thus on every edge of face $\hat{x}_1 = 0$, we have $\hat{\mathbf{u}} \cdot \hat{\tau} \in P_{k-1}$. Then choosing $q = \hat{\mathbf{u}} \cdot \hat{\tau}$ in (3.25) leads to $\hat{\mathbf{u}} \cdot \hat{\tau} = 0$ on each edge of this face.

Furthermore, because $\hat{\mathbf{u}} \cdot \hat{\tau} = 0$ on each edge of face $\hat{x}_1 = 0$, we know that the tangential components of $\hat{\tau}$ on this face can be written as:

$$\hat{u}_2 = \hat{x}_3(1-\hat{x}_3)\hat{v}_2, \quad \hat{v}_2 \in Q_{k-1,k-2}(\hat{x}_2, \hat{x}_3),$$
$$\hat{u}_3 = \hat{x}_2(1-\hat{x}_2)\hat{v}_3, \quad \hat{v}_3 \in Q_{k-2,k-1}(\hat{x}_2, \hat{x}_3).$$

Hence on $\hat{x}_1 = 0$, choosing $\hat{\mathbf{q}} = (-\hat{v}_3, \hat{v}_2)$ in (3.26) shows that $\hat{v}_2 = \hat{v}_3 = 0$ on this face, i.e., $\hat{\mathbf{u}} \times \hat{\mathbf{n}} = \mathbf{0}$ on face $\hat{x}_1 = 0$, which proves the curl conformity.

(ii) Now let us consider the unisolvence. Note that the dimension of $P_{\hat{K}}$ is $3k(k+1)^2$, which is same as the total number of degrees of freedom in $\Sigma_{\hat{K}}$. Hence we just need to prove that vanishing all degrees of freedom for $\hat{\mathbf{u}} \in P_{\hat{K}}$ yields $\hat{\mathbf{u}} = 0$. From Part (i), we know that $\hat{\mathbf{u}} \times \hat{\mathbf{n}}_i = \mathbf{0}$ on all faces, which implies that $\hat{\mathbf{u}}$ can be written as

$$\hat{\mathbf{u}} = (\hat{x}_2(1-\hat{x}_2)\hat{x}_3(1-\hat{x}_3)\hat{r}_1, \hat{x}_1(1-\hat{x}_1)\hat{x}_3(1-\hat{x}_3)\hat{r}_2, \hat{x}_1(1-\hat{x}_1)\hat{x}_2(1-\hat{x}_2)\hat{r}_3)^T,$$

where $\hat{r}_1 \in Q_{k-1,k-2,k-2}$, $\hat{r}_2 \in Q_{k-2,k-1,k-2}$, and $\hat{r}_3 \in Q_{k-2,k-2,k-1}$. Choosing $\hat{\mathbf{q}} = \hat{\mathbf{r}} \equiv (\hat{r}_1, \hat{r}_2, \hat{r}_3)^T$ in (3.27) shows that $\hat{\mathbf{r}} = 0$, which completes the proof. □

After obtaining the basis function on the reference element $\hat{K}$, we can derive the basis function on a general element K through mapping. To make the degrees of freedom (3.25)–(3.27) invariant, we need the following special transformation

$$\mathbf{u} \circ F_K = B_k^{-T}\hat{\mathbf{u}}, \tag{3.28}$$

where F_K is the affine mapping defined in (2.18). For technical reasons, we assume that B_K is a diagonal matrix, hence the mapped element K has all edges parallel to the coordinate axes. The unit outward normal vector $\mathbf{n}$ to K is obtained by the transformation (3.4), and the unit tangential vector $\boldsymbol{\tau}$ along edge e of K is given by:

$$\boldsymbol{\tau} = B_K\hat{\boldsymbol{\tau}}/|B_K\hat{\boldsymbol{\tau}}|, \tag{3.29}$$

where $\hat{\boldsymbol{\tau}}$ is a unit tangential vector along edge $\hat{e}$ of $\hat{K}$. Note that (3.29) can be seen as follows: a tangent vector $\hat{\boldsymbol{\tau}} = \hat{\mathbf{x}}_1 - \hat{\mathbf{x}}_2$ is transformed into

$$\mathbf{x}_1 - \mathbf{x}_2 = B_K(\hat{\mathbf{x}}_1 - \hat{\mathbf{x}}_2) = B_K\hat{\boldsymbol{\tau}},$$

normalizing which leads to (3.29).

Lemma 3.7. *Suppose that $\hat{\mathbf{u}} \in H(curl; \hat{K})$, and $\mathbf{u}$ is mapped from $\hat{\mathbf{u}}$ by (3.28). Then $\mathbf{u} \in H(curl; K)$ and*

$$\nabla \times \mathbf{u} = \frac{1}{det(B_K)} B_K \hat{\nabla} \times \hat{\mathbf{u}}. \tag{3.30}$$

Proof. From (3.28), we have

$$\hat{u}_i = b_{1i}u_1 + b_{2i}u_2 + b_{3i}u_3, \quad i = 1, 2, 3.$$

From mapping (2.18), we have

$$\frac{\partial}{\partial\hat{x}_1} = \frac{\partial}{\partial x_1}\frac{\partial x_1}{\partial\hat{x}_1} + \frac{\partial}{\partial x_2}\frac{\partial x_2}{\partial\hat{x}_1} + \frac{\partial}{\partial x_3}\frac{\partial x_3}{\partial\hat{x}_1} = b_{11}\frac{\partial}{\partial x_1} + b_{21}\frac{\partial}{\partial x_2} + b_{31}\frac{\partial}{\partial x_3},$$
$$\frac{\partial}{\partial\hat{x}_2} = \frac{\partial}{\partial x_1}\frac{\partial x_1}{\partial\hat{x}_2} + \frac{\partial}{\partial x_2}\frac{\partial x_2}{\partial\hat{x}_2} + \frac{\partial}{\partial x_3}\frac{\partial x_3}{\partial\hat{x}_2} = b_{12}\frac{\partial}{\partial x_1} + b_{22}\frac{\partial}{\partial x_2} + b_{32}\frac{\partial}{\partial x_3},$$
$$\frac{\partial}{\partial\hat{x}_3} = \frac{\partial}{\partial x_1}\frac{\partial x_1}{\partial\hat{x}_3} + \frac{\partial}{\partial x_2}\frac{\partial x_2}{\partial\hat{x}_3} + \frac{\partial}{\partial x_3}\frac{\partial x_3}{\partial\hat{x}_3} = b_{13}\frac{\partial}{\partial x_1} + b_{23}\frac{\partial}{\partial x_2} + b_{33}\frac{\partial}{\partial x_3},$$

using which we obtain the first component of $\hat{\nabla} \times \hat{\mathbf{u}}$ as

$$
\begin{aligned}
& (\hat{\nabla} \times \hat{\mathbf{u}})_1 = \frac{\partial \hat{u}_3}{\partial \hat{x}_2} - \frac{\partial \hat{u}_2}{\partial \hat{x}_3} \\
= & \frac{\partial}{\partial \hat{x}_2}(b_{13}u_1 + b_{23}u_2 + b_{33}u_3) - \frac{\partial}{\partial \hat{x}_3}(b_{12}u_1 + b_{22}u_2 + b_{32}u_3) \\
= & b_{13}(b_{12}\frac{\partial u_1}{\partial x_1} + b_{22}\frac{\partial u_1}{\partial x_2} + b_{32}\frac{\partial u_1}{\partial x_3}) + b_{23}(b_{12}\frac{\partial u_2}{\partial x_1} + b_{22}\frac{\partial u_2}{\partial x_2} + b_{32}\frac{\partial u_2}{\partial x_3}) + \cdots \\
= & (-b_{32}b_{23} + b_{33}b_{22})(\frac{\partial u_3}{\partial x_2} - \frac{\partial u_2}{\partial x_3}) + (-b_{12}b_{33} + b_{13}b_{32})(\frac{\partial u_1}{\partial x_3} - \frac{\partial u_3}{\partial x_1}) \\
& +(b_{23}b_{12} - b_{22}b_{13})(\frac{\partial u_2}{\partial x_1} - \frac{\partial u_1}{\partial x_2}) \\
= & det(B_K) \cdot (B_K^{-1}\nabla \times \mathbf{u})_1,
\end{aligned}
$$

where in the last step we used the fact that: The inverse of matrix A can be written as $A^{-1} = \frac{1}{det(A)}C^T$, where C is the matrix of cofactors, i.e., each element c_{ij} of C is the cofactor corresponding to element a_{ij} of A.

By the same technique, we can prove that

$$
\begin{aligned}
(\hat{\nabla} \times \hat{\mathbf{u}})_2 &= \frac{\partial \hat{u}_1}{\partial \hat{x}_3} - \frac{\partial \hat{u}_3}{\partial \hat{x}_1} = det(B_K) \cdot (B_K^{-1}\nabla \times \mathbf{u})_2, \\
(\hat{\nabla} \times \hat{\mathbf{u}})_3 &= \frac{\partial \hat{u}_2}{\partial \hat{x}_1} - \frac{\partial \hat{u}_1}{\partial \hat{x}_2} = det(B_K) \cdot (B_K^{-1}\nabla \times \mathbf{u})_3,
\end{aligned}
$$

which concludes our proof. □

A more general result

$$
(\nabla \times \mathbf{u}) \circ F_K = \frac{1}{det(dF_K)} dF_K \hat{\nabla} \times \hat{\mathbf{u}} \tag{3.31}
$$

holds true [217, Corollary 3.58], where the mapping $F_K : \hat{K} \to K$ is assumed to be continuously differentiable, invertible and surjective, i.e., F_K is not restricted to an affine mapping. Here $dF_K = dF_K(\hat{x})/d\hat{x}$ is the jacobian of the mapping. It is easy to see that for the affine mapping $F_K(\hat{x}) = B_K\hat{x} + b_K$, the jacobian $dF_K = B_K$, and (3.31) reduces to (3.30).

Lemma 3.8. *Suppose that $det(B_K) > 0$, and the function $\mathbf{u}$ and the tangential vector $\boldsymbol{\tau}$ are obtained by the transformations (3.28) and (3.29), respectively. Then the degrees of freedom of $\mathbf{u}$ on K given by*

$$
M_e(\mathbf{u}) = \{\int_{e_i} \mathbf{u} \cdot \boldsymbol{\tau}_i q ds, \ \forall\, q \in P_{k-1}(e_i), \ i = 1, \cdots, 12\},
$$

$$M_f(\mathbf{u}) = \{\int_{f_i} \mathbf{u} \times \mathbf{n}_i \cdot \mathbf{q} dA,$$
$$\forall\, \mathbf{q} \circ F_K = B_K^{-T} \hat{\mathbf{q}},\ \ \hat{\mathbf{q}} \in Q_{k-2,k-1}(\hat{f}_i) \times Q_{k-1,k-2}(\hat{f}_i),\ i = 1, \cdots, 6\},$$
$$M_K(\mathbf{u}) = \{\int_K \mathbf{u} \cdot \mathbf{q} dV,\ \forall\, \mathbf{q} \circ F_K = \frac{1}{det(B_K)} B_K \hat{\mathbf{q}},$$
$$\hat{\mathbf{q}} \in Q_{k-1,k-2,k-2} \times Q_{k-2,k-1,k-2} \times Q_{k-2,k-2,k-1}\},$$

are identical to the degrees of freedom for $\hat{\mathbf{u}}$ *on* $\hat{K}$ *given in (3.25)–(3.27).*

Proof. (i) By the transformations (3.28) and (3.29), we have

$$\int_e \mathbf{u} \cdot \boldsymbol{\tau} q ds = \int_{\hat{e}} B_K^{-T} \hat{\mathbf{u}} \cdot \frac{1}{|B_K \hat{\boldsymbol{\tau}}|} B_K \hat{\boldsymbol{\tau}} \cdot \hat{q} \cdot \frac{ds}{d\hat{s}} d\hat{s} = \int_{\hat{e}} \hat{\mathbf{u}} \cdot \hat{\boldsymbol{\tau}} \hat{q} d\hat{s},$$

which shows that the degrees of freedom $M_e(\mathbf{u})$ are invariant.

(ii) By Green's formula and (3.30), we have

$$\int_{\partial K} \mathbf{n} \times \mathbf{u} \cdot \mathbf{q} dA = \int_K (\nabla \times \mathbf{u} \cdot \mathbf{q} - \mathbf{u} \cdot \nabla \times \mathbf{q}) dV$$
$$= \int_{\hat{K}} [\frac{1}{det(B_K)} B_K \hat{\nabla} \times \hat{\mathbf{u}} \cdot B_K^{-T} \hat{\mathbf{q}} - B_K^{-T} \hat{\mathbf{u}} \cdot \frac{1}{det(B_K)} B_K \hat{\nabla} \times \hat{\mathbf{q}}] det(B_K) d\hat{V}$$
$$= \int_{\hat{K}} (\hat{\nabla} \times \hat{\mathbf{u}} \cdot \hat{\mathbf{q}} - \hat{\mathbf{u}} \cdot \hat{\nabla} \times \hat{\mathbf{q}}) d\hat{V} = \int_{\partial \hat{K}} \hat{\mathbf{n}} \times \hat{\mathbf{u}} \cdot \hat{\mathbf{q}} d\hat{A},$$

which shows that the degrees of freedom $M_f(\mathbf{u})$ are invariant.

(iii) The invariance of $M_K(\mathbf{u})$ is easy to see by noting that

$$\int_K \mathbf{u} \cdot \mathbf{q} dV = \int_{\hat{K}} B_K^{-T} \hat{\mathbf{u}} \cdot \frac{1}{det(B_K)} B_K \hat{\mathbf{q}} \cdot det(B_K) d\hat{V} = \int_{\hat{K}} \hat{\mathbf{u}} \cdot \hat{\mathbf{q}} d\hat{V}.$$

□

Now suppose that we have a regular family of hexahedral meshes of Ω, denoted as T_h. We can define a curl conforming finite element space V_h on the mesh T_h by assembling the degrees of freedom from each element K in T_h, i.e.,

$$\Sigma = \cup_{K \in T_h} \left(M_e(\mathbf{u}) \cup M_f(\mathbf{u}) \cup M_K(\mathbf{u}) \right).$$

More specifically, we can write V_h explicitly as

$$V_h = \{\mathbf{u}_h \in H(\mathrm{curl}; \Omega) :\ \mathbf{u}_h|_K \in$$
$$Q_{k-1,k,k} \times Q_{k,k-1,k} \times Q_{k,k,k-1},\ \forall\, K \in T_h\}. \tag{3.32}$$

The curl conforming finite element space (3.32) can be extended similarly to rectangular elements, in which case V_h becomes:

$$V_h = \{\mathbf{u}_h \in H(\text{curl}; \Omega) : \quad \mathbf{u}_h|_K \in Q_{k-1,k} \times Q_{k,k-1}, \ \forall \ K \in T_h\}. \tag{3.33}$$

On a reference rectangle $\hat{K} = (0,1)^2$, the set of DOFs for the curl conforming element is formed by edge DOFs:

$$M_{\hat{e}}(\hat{\mathbf{u}}) = \{\int_{\hat{e}_i} \hat{\mathbf{u}} \cdot \hat{\boldsymbol{\tau}}_i \hat{q} d\hat{s}, \ \forall \ \hat{q} \in P_{k-1}(\hat{e}_i), i = 1, \cdots, 4\},$$

and element DOFs:

$$M_{\hat{K}}(\hat{\mathbf{u}}) = \{\int_{\hat{K}} \hat{\mathbf{u}} \cdot \hat{\mathbf{q}} d\hat{V}, \ \forall \ \hat{\mathbf{q}} \in Q_{k-2,k-1} \times Q_{k-1,k-2}\}.$$

Hence we have a total of $4k$ edge DOFs, and $2 \cdot (k-1)k$ element DOFs, whose summation equals

$$4k + 2 \cdot (k-1)k = 2k(k+1) = \dim(Q_{k-1,k} \times Q_{k,k-1}).$$

The DOFs on general rectangles can be defined similarly as shown in Lemma 3.8. More specifically, we only need the following DOFs:

$$M_e(\mathbf{u}) = \{\int_{e_i} \mathbf{u} \cdot \boldsymbol{\tau}_i q ds, \ \forall \ q \in P_{k-1}(e_i), \ i = 1, \cdots, 4\},$$
$$M_K(\mathbf{u}) = \{\int_K \mathbf{u} \cdot \mathbf{q} dV, \ \forall \ \mathbf{q} \circ F_K = \frac{1}{det(B_K)} B_K \hat{\mathbf{q}}, \ \hat{\mathbf{q}} \in Q_{k-2,k-1} \times Q_{k-1,k-2}\}.$$

Below we present some exemplary curl conforming finite elements.

Example 3.5. Consider a cube $K = (x_c - h_x, x_c + h_x) \times (y_c - h_y, y_c + h_y) \times (z_c - h_z, z_c + h_z)$. The lowest-order curl conforming finite element (i.e., $k = 1$ in Definition 3.4) has $\mathbf{u}_{\hat{K}} \in Q_{0,1,1} \times Q_{1,0,1} \times Q_{1,1,0}$. Hence we can represent $\mathbf{u}_{\hat{K}}$ as follows:

$$\mathbf{u}_{\hat{K}} = ((a_1 + b_1 y)(c_1 + d_1 z), (a_2 + b_2 x)(c_2 + d_2 z), (a_3 + b_3 x)(c_3 + d_3 y))^T,$$

where the constants can be determined by the 12 edge degrees of freedom of (3.25).

The 12 edges are labeled as follows:

$$l_1 : (x_c - h_x, y_c - h_y, z_c - h_z) \to (x_c + h_x, y_c - h_y, z_c - h_z),$$
$$l_2 : (x_c + h_x, y_c - h_y, z_c - h_z) \to (x_c + h_x, y_c + h_y, z_c - h_z),$$
$$l_3 : (x_c - h_x, y_c + h_y, z_c - h_z) \to (x_c + h_x, y_c + h_y, z_c - h_z),$$

$$
\begin{aligned}
&l_4 : (x_c - h_x, y_c - h_y, z_c - h_z) \rightarrow (x_c - h_x, y_c + h_y, z_c - h_z),\\
&l_5 : (x_c - h_x, y_c - h_y, z_c + h_z) \rightarrow (x_c + h_x, y_c - h_y, z_c + h_z),\\
&l_6 : (x_c + h_x, y_c - h_y, z_c + h_z) \rightarrow (x_c + h_x, y_c + h_y, z_c + h_z),\\
&l_7 : (x_c - h_x, y_c + h_y, z_c + h_z) \rightarrow (x_c + h_x, y_c + h_y, z_c + h_z),\\
&l_8 : (x_c - h_x, y_c - h_y, z_c + h_z) \rightarrow (x_c - h_x, y_c + h_y, z_c + h_z),\\
&l_9 : (x_c + h_x, y_c - h_y, z_c - h_z) \rightarrow (x_c + h_x, y_c - h_y, z_c + h_z),\\
&l_{10} : (x_c + h_x, y_c + h_y, z_c - h_z) \rightarrow (x_c + h_x, y_c + h_y, z_c + h_z),\\
&l_{11} : (x_c - h_x, y_c + h_y, z_c - h_z) \rightarrow (x_c - h_x, y_c + h_y, z_c + h_z),\\
&l_{12} : (x_c - h_x, y_c - h_y, z_c - h_z) \rightarrow (x_c - h_x, y_c - h_y, z_c + h_z).
\end{aligned}
$$

For any $\mathbf{E} \in H(\text{curl}; K)$, its curl interpolation $\Pi_K^c \mathbf{E}$ satisfying

$$\int_{l_i} (\mathbf{E} - \Pi_K^c \mathbf{E}) \cdot \boldsymbol{\tau}_i dl = 0, \quad i = 1, \cdots, 12, \tag{3.34}$$

where $\boldsymbol{\tau}_i$ is the corresponding unit tangential vector along each edge l_i.

Using (3.34) and after some algebraic calculations, we obtain

$$\Pi_K^c \mathbf{E}(x, y, z) = \sum_{j=1}^{12} \Big(\int_{l_j} \mathbf{E} \cdot \boldsymbol{\tau}_j dl \Big) \mathbf{N}_j(x, y, z),$$

where the basis functions $\mathbf{N}_j$ are given as follows:

$$\mathbf{N}_1 = \begin{pmatrix} \frac{(y_c + h_y - y)(z_c + h_z - z)}{|K|} \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{N}_2 = \begin{pmatrix} 0 \\ \frac{(x - x_c + h_x)(z_c + h_z - z)}{|K|} \\ 0 \end{pmatrix},$$

$$\mathbf{N}_3 = \begin{pmatrix} \frac{(y_c - h_y - y)(z_c + h_z - z)}{|K|} \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{N}_4 = \begin{pmatrix} 0 \\ \frac{(x - x_c - h_x)(z_c + h_z - z)}{|K|} \\ 0 \end{pmatrix},$$

where $|K| = 8h_x h_y h_z$ denotes the volume of K.

Other basis functions can be obtained similarly. For example,

$$\mathbf{N}_5 = \begin{pmatrix} \frac{(y_c + h_y - y)(z - z_c + h_z)}{|K|} \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{N}_9 = \begin{pmatrix} 0 \\ 0 \\ \frac{(x - x_c + h_x)(y_c + h_y - y)}{|K|} \end{pmatrix}.$$

It is easy to check that the basis functions $\mathbf{N}_j$ satisfy the property

$$\int_{l_i} \mathbf{N}_j \cdot \boldsymbol{\tau}_i dl = \delta_{ij}, \;\; i, j = 1, \cdots 12.$$

Example 3.6. Consider a rectangle $K = [x_c - h_x, x_c + h_x] \times [y_c - h_y, y_c + h_y]$. For the lowest-order edge element $Q_{0,1} \times Q_{1,0}$, the interpolation $\Pi_K^c \mathbf{u}$ of any $\mathbf{u} \in H(\text{curl}; K)$ can be written as

$$\Pi_K^c \mathbf{u}(x, y) = \sum_{j=1}^{4} (\int_{l_j} \mathbf{u} \cdot \boldsymbol{\tau}_j dl) \mathbf{N}_j(x, y), \tag{3.35}$$

where l_j denote the four edges of the element, which start from the bottom and are oriented counterclockwise. Furthermore, $|l_j|$ and $\boldsymbol{\tau}_j$ represent the length of edge l_j and the unit tangent vector along l_j, respectively. The edge element basis functions $\mathbf{N}_j$ are as follows:

$$\mathbf{N}_1 = \begin{pmatrix} \frac{(y_c + h_y) - y}{4h_x h_y} \\ 0 \end{pmatrix}, \;\; \mathbf{N}_2 = \begin{pmatrix} 0 \\ \frac{x - (x_c - h_x)}{4h_x h_y} \end{pmatrix},$$

$$\mathbf{N}_3 = \begin{pmatrix} \frac{(y_c - h_y) - y}{4h_x h_y} \\ 0 \end{pmatrix}, \;\; \mathbf{N}_4 = \begin{pmatrix} 0 \\ \frac{x - (x_c + h_x)}{4h_x h_y} \end{pmatrix}.$$

Example 3.7. Consider a rectangle $K = [x_c - h_x, x_c + h_x] \times [y_c - h_y, y_c + h_y]$. For the second-order edge element $Q_{1,2} \times Q_{2,1}$, the interpolation $\Pi_K^c \mathbf{u}$ of any $\mathbf{u} \in H(\text{curl}; K)$ can be obtained by satisfying

$$\int_{l_i} (\mathbf{u} - \Pi_K^c \mathbf{u}) \cdot \boldsymbol{\tau}_i q dl = 0, \;\forall\, q \in P_1(l_i), \; i = 1, \cdots, 4,$$

$$\int_K (\mathbf{u} - \Pi_K^c \mathbf{u}) \cdot \mathbf{q} dxdy = 0, \;\forall\, \mathbf{q} \in Q_{0,1} \times Q_{1,0}.$$

Let us denote the unit tangent vectors $\boldsymbol{\tau}_j$ along l_j be:

$$\boldsymbol{\tau}_1 = (1, 0)', \;\; \boldsymbol{\tau}_2 = (0, 1)', \;\; \boldsymbol{\tau}_3 = (-1, 0)', \;\; \boldsymbol{\tau}_4 = (0, -1)'.$$

After lengthy calculations, we can write the interpolation $\Pi_K^c \mathbf{u}$ as

$$\Pi_K^c \mathbf{u}(x, y) = \sum_{j=1}^{12} c_j \mathbf{N}_j(x, y), \tag{3.36}$$

where the DOFs c_j are

$$c_j = \int_{l_j} \mathbf{u}(x, y) \cdot \boldsymbol{\tau}_j dl, \quad j = 1, \cdots, 4,$$
$$c_5 = \int_{l_1} (x - x_c)\mathbf{u}(x, y) \cdot \boldsymbol{\tau}_1 dl, \quad c_6 = \int_{l_2} (y - y_c)\mathbf{u}(x, y) \cdot \boldsymbol{\tau}_2 dl,$$
$$c_7 = \int_{l_3} (x - x_c)\mathbf{u}(x, y) \cdot \boldsymbol{\tau}_3 dl, \quad c_8 = \int_{l_4} (y - y_c)\mathbf{u}(x, y) \cdot \boldsymbol{\tau}_4 dl,$$
$$c_9 = \int_K \mathbf{u}(x, y) \cdot (1, 0)' dxdy, \quad c_{10} = \int_K \mathbf{u}(x, y) \cdot (0, 1)' dxdy,$$
$$c_{11} = \int_K \mathbf{u}(x, y) \cdot (x - x_c, 0)' dxdy, \quad c_{12} = \int_K \mathbf{u}(x, y) \cdot (0, y - y_c)' dxdy,$$

and the basis functions $\mathbf{N}_j$ can be expressed as:

$$\mathbf{N}_1 = \begin{pmatrix} \frac{[3(y-y_c)+h_y][(y-y_c)-h_y]}{8h_x h_y^2} \\ 0 \end{pmatrix}, \quad \mathbf{N}_2 = \begin{pmatrix} 0 \\ \frac{[3(x-x_c)-h_x][(x-x_c)+h_x]}{8h_x^2 h_y} \end{pmatrix},$$

$$\mathbf{N}_3 = \begin{pmatrix} \frac{-[3(y-y_c)-h_y][(y-y_c)+h_y]}{8h_x h_y^2} \\ 0 \end{pmatrix}, \quad \mathbf{N}_4 = \begin{pmatrix} 0 \\ \frac{-[3(x-x_c)+h_x][(x-x_c)-h_x]}{8h_x^2 h_y} \end{pmatrix},$$

$$\mathbf{N}_5 = \begin{pmatrix} \frac{(x-x_c)[3(y-y_c)-3h_y][3(y-y_c)+h_y]}{8h_x^3 h_y^2} \\ 0 \end{pmatrix}, \quad \mathbf{N}_6 = \begin{pmatrix} 0 \\ \frac{(y-y_c)[3(x-x_c)+3h_x][3(x-x_c)-h_x]}{8h_x^2 h_y^3} \end{pmatrix},$$

$$\mathbf{N}_7 = \begin{pmatrix} \frac{-(x-x_c)[3(y-y_c)+3h_y][3(y-y_c)-h_y]}{8h_x^3 h_y^2} \\ 0 \end{pmatrix}, \mathbf{N}_8 = \begin{pmatrix} 0 \\ \frac{-(y-y_c)[3(x-x_c)-3h_x][3(x-x_c)+h_x]}{8h_x^2 h_y^3} \end{pmatrix},$$

$$\mathbf{N}_9 = \begin{pmatrix} \frac{-3[(y-y_c)-h_y][(y-y_c)+h_y]}{8h_x h_y^3} \\ 0 \end{pmatrix}, \quad \mathbf{N}_{10} = \begin{pmatrix} 0 \\ \frac{-3[(x-x_c)-h_x][(x-x_c)+h_x]}{8h_x^3 h_y} \end{pmatrix},$$

$$\mathbf{N}_{11} = \begin{pmatrix} \frac{-9(x-x_c)[(y-y_c)-h_y][(y-y_c)+h_y]}{8h_x^3 h_y^3} \\ 0 \end{pmatrix}, \quad \mathbf{N}_{12} = \begin{pmatrix} 0 \\ \frac{-9(y-y_c)[(x-x_c)+h_x][(x-x_c)-h_x]}{8h_x^3 h_y^3} \end{pmatrix}.$$

3.2.2 Interpolation Error Estimates

With sufficient regularity, there exists a well-defined $H(curl)$ interpolation operator on K denoted as Π_K^c. For example, if we assume that $\mathbf{u}, \nabla \times \mathbf{u} \in (H^{1/2+\delta}(K))^3$, $\delta > 0$, then there is a unique function

$$\Pi_K^c \mathbf{u} \in Q_{k-1,k,k} \times Q_{k,k-1,k} \times Q_{k,k,k-1}$$

such that

$$M_e(\mathbf{u} - \Pi_K^c \mathbf{u}) = 0, \quad M_f(\mathbf{u} - \Pi_K^c \mathbf{u}) = 0 \text{ and } M_K(\mathbf{u} - \Pi_K^c \mathbf{u}) = 0,$$

where M_e, M_f and M_K are the sets of degrees of freedom stated in Lemma 3.8.

Similar to the proof carried out for the $H(div)$ interpolation operator, we can easily prove the following lemma, which shows that the interpolant $\Pi_K^c \mathbf{u}$ on a general element K and the interpolation $\Pi_{\hat{K}}^c \hat{\mathbf{u}}$ on the reference element $\hat{K}$ are closely related.

Lemma 3.9. *Suppose that* $\mathbf{u}$ *is sufficiently smooth such that* $\Pi_K^c \mathbf{u}$ *is well defined. Then under transformation (3.28), we have*

$$\widehat{\Pi_K^c \mathbf{u}} = \Pi_{\hat{K}}^c \hat{\mathbf{u}}.$$

From the local interpolation operator Π_K^c, we can define a global interpolation operator

$$\Pi_h^c : (H^{\frac{1}{2}+\delta}(\Omega))^3 \to W_h, \quad \forall\, \delta > 0,$$

element-wisely by

$$(\Pi_h^c \mathbf{u})|_K = \Pi_K^c(\mathbf{u}|_K) \quad \text{for each } K \in T_h.$$

The following theorem shows that there is a close connection between the curl interpolation and divergence interpolation.

Theorem 3.6. *For the space* W_h *given by (3.8) and* V_h *given by (3.32), we have*

$$\nabla \times V_h \subset W_h.$$

Furthermore, if we assume that $\mathbf{u}$ *is smooth enough such that* $\Pi_h^c \mathbf{u}$ *and* $\Pi_h^d \nabla \times \mathbf{u}$ *are well defined, then we have*

$$\nabla \times \Pi_h^c \mathbf{u} = \Pi_h^d \nabla \times \mathbf{u}. \tag{3.37}$$

Proof. For any $\mathbf{u}_h \in V_h$, it is easy to see that

$$\nabla \times \mathbf{u}_h|_K \in Q_{k,k-1,k-1} \times Q_{k-1,k,k-1} \times Q_{k-1,k-1,k},$$

which leads to $\nabla \times V_h \subset W_h$.

Without loss of generality, we just prove that (3.37) for a reference element K (for simplicity, we drop the hat notation). Noting that $\nabla \times \Pi_h^c \mathbf{u} - \Pi_h^d \nabla \times \mathbf{u} \in W_h \subset H(\text{div}; \Omega)$, hence proof of (3.37) is equivalent to prove that the degrees of freedom given in (3.1) and (3.2) vanish for $\nabla \times \Pi_h^c \mathbf{u} - \Pi_h^d \nabla \times \mathbf{u}$.

(i) Consider a face f of K with face normal $\mathbf{n}$, and let $q \in Q_{k-1,k-1}(f)$. Using (3.5) and integration by parts, we have

$$\int_f (\nabla \times \Pi_K^c \mathbf{u} - \Pi_K^d \nabla \times \mathbf{u}) \cdot \mathbf{n} q dA = \int_f (\nabla \times \Pi_K^c \mathbf{u} - \nabla \times \mathbf{u}) \cdot \mathbf{n} q dA$$

$$= -\int_f \nabla_f \cdot (\mathbf{n} \times (\Pi_K^c \mathbf{u} - \mathbf{u})) q dA$$

$$= \int_f \mathbf{n} \times (\Pi_K^c \mathbf{u} - \mathbf{u}) \cdot \nabla_f q dA - \int_{\partial f} \mathbf{n}_{\partial f} \cdot (\mathbf{n} \times (\Pi_K^c \mathbf{u} - \mathbf{u})) q ds, \tag{3.38}$$

where $\mathbf{n}_{\partial f}$ is the unit outward normal to ∂f on the plane f. Note that in the second equality we used an identity [217, (3.52)], and ∇_f denotes the surface gradient. The first term in (3.38) actually becomes zero by noting that $\nabla_f q \in Q_{k-2,k-1}(f) \times Q_{k-1,k-2}(f)$. Furthermore, the second term in (3.38) can be rewritten as

$$\int_{\partial f} \mathbf{n}_{\partial f} \cdot (\mathbf{n} \times (\Pi_K^c \mathbf{u} - \mathbf{u})) q ds = \int_{\partial f} (\mathbf{n}_{\partial f} \times \mathbf{n}) \cdot (\Pi_K^c \mathbf{u} - \mathbf{u}) q ds,$$

which vanishes since $q \in P_{k-1}(e)$ on each edge of f.

(ii) Let $\mathbf{q} \in Q_{k-2,k-1,k-1} \times Q_{k-1,k-2,k-1} \times Q_{k-1,k-1,k-2}$. Using (3.6) and integration by parts, we have

$$\int_K (\nabla \times \Pi_K^c \mathbf{u} - \Pi_K^d \nabla \times \mathbf{u}) \cdot \mathbf{q} dV = \int_K (\nabla \times \Pi_K^c \mathbf{u} - \nabla \times \mathbf{u}) \cdot \mathbf{q} dV$$

$$= \int_K (\Pi_K^c \mathbf{u} - \mathbf{u}) \cdot \nabla \times \mathbf{q} dV + \int_{\partial K} (\mathbf{n} \times (\Pi_K^c \mathbf{u} - \mathbf{u})) \cdot \mathbf{q} dA.$$

The right hand side vanishes by using (3.27) and (3.26), and this concludes the proof.

□

Lemma 3.10. *Suppose that* $\mathbf{v}$ *and* $\hat{\mathbf{v}}$ *are related by the transformation (3.28). Then for any* $s \geq 0$, *we have*

$$|\hat{\mathbf{v}}|_{(H^s(\hat{K}))^3} \leq C|det(B_K)|^{-1/2} ||B_K||^{s+1} |\mathbf{v}|_{(H^s(\hat{K}))^3},$$

$$|\hat{\nabla} \times \hat{\mathbf{v}}|_{(H^s(\hat{K}))^3} \leq C|det(B_K)|^{1/2} ||B_K||^{s-1} |\nabla \times \mathbf{v}|_{(H^s(\hat{K}))^3}.$$

Proof. From $\hat{\mathbf{v}} = B_K^T \mathbf{v} \circ F_K$, we have

$$\frac{\partial^\alpha \hat{\mathbf{v}}}{\partial \hat{\mathbf{x}}^\alpha} = B_K^T \frac{\partial^\alpha}{\partial \hat{\mathbf{x}}^\alpha} (\mathbf{v} \circ F_K) = B_K^T (B_K)^\alpha \frac{\partial^\alpha \mathbf{v}}{\partial \mathbf{x}^\alpha},$$

which leads to

$$||\frac{\partial^\alpha \hat{\mathbf{v}}}{\partial \hat{\mathbf{x}}^\alpha}||_{(L^2(\hat{K}))^3} = (\int_K |B_K^T (B_K)^\alpha \frac{\partial^\alpha \mathbf{v}}{\partial \mathbf{x}^\alpha}|^2 \cdot \frac{1}{det(B_K)} dV)^{1/2}$$
$$\leq C |det(B_K)|^{-1/2} ||B_K||^{|\alpha|+1} ||\frac{\partial^\alpha \mathbf{v}}{\partial \mathbf{x}^\alpha}||_{(L^2(K))^3},$$

and summing all multi-indices $|\alpha|_1 = s$ completes the proof of the first part.

Using the fact that $\hat{\nabla} \times \hat{\mathbf{v}} = det(B_K) B_K^{-1} \nabla \times \mathbf{v}$, we can prove the second part similarly by noting that

$$||\frac{\partial^\alpha}{\partial \hat{\mathbf{x}}^\alpha}(\hat{\nabla} \times \hat{\mathbf{v}})||_{(L^2(\hat{K}))^3}$$
$$= (\int_K |det(B_K) B_K^{-1} (B_K)^\alpha \frac{\partial^\alpha}{\partial \mathbf{x}^\alpha}(\nabla \times \mathbf{v})|^2 \cdot \frac{1}{det(B_K)} dV)^{1/2}$$
$$\leq C |det(B_K)|^{1/2} ||B_K||^{|\alpha|-1} ||\frac{\partial^\alpha}{\partial \mathbf{x}^\alpha}(\nabla \times \mathbf{v})||_{(L^2(K))^3}.$$

□

Now we can prove the error estimate for Π_h^c interpolation operator.

Theorem 3.7. *Assume that* $0 < \delta < \frac{1}{2}$ *and* T_h *is a regular family of hexahedral meshes on* Ω *with faces aligning with the coordinate axes. If* $\mathbf{u}, \nabla \times \mathbf{u} \in (H^s(\Omega))^3$, $\frac{1}{2} + \delta \leq s \leq k$, *then there is a constant* $C > 0$ *independent of* h *and* $\mathbf{u}$ *such that*

$$||\mathbf{u} - \Pi_h^c \mathbf{u}||_{(L^2(\Omega))^3} + ||\nabla \times (\mathbf{u} - \Pi_h^c \mathbf{u})||_{(L^2(\Omega))^3}$$
$$\leq Ch^s (||\mathbf{u}||_{(H^s(\Omega))^3} + ||\nabla \times \mathbf{u}||_{(H^s(\Omega))^3}), \quad \frac{1}{2} + \delta \leq s \leq k. \tag{3.39}$$

Proof. For simplicity, here we only prove the result for integer $s = k \geq 1$.

As usual, we start with a local estimate on one element K. By (3.28), we have

$$||\mathbf{u} - \Pi_K^c \mathbf{u}||_{(L^2(K))^3} = (\int_K |\mathbf{u} - \Pi_K^c \mathbf{u}|^2 dV)^{1/2}$$
$$= (\int_{\hat{K}} |B_K^{-T} (\hat{\mathbf{u}} - \widehat{\Pi_K^c \mathbf{u}})|^2 |\det(B_K)| d\hat{V})^{1/2}$$
$$\leq |\det(B_K)|^{1/2} ||B_K^{-1}|| \, ||\hat{\mathbf{u}} - \widehat{\Pi_K^c \mathbf{u}}||_{(L^2(\hat{K}))^3}. \tag{3.40}$$

By Lemma 3.9 and the fact that

$$(I - \Pi_{\hat{K}}^c)\hat{\mathbf{p}} = 0 \quad \forall\, \hat{\mathbf{p}} \in (Q_{k-1,k-1,k-1})^3,$$

we have

$$||\hat{\mathbf{u}} - \widehat{\Pi_K^c \mathbf{u}}||_{(L^2(\hat{K}))^3} = ||\hat{\mathbf{u}} - \Pi_{\hat{K}}^c \hat{\mathbf{u}}||_{(L^2(\hat{K}))^3} = ||(I - \Pi_{\hat{K}}^c)(\hat{\mathbf{u}} + \hat{\mathbf{p}})||_{(L^2(\hat{K}))^3}$$
$$\leq C(||\hat{\mathbf{u}} + \hat{\mathbf{p}}||_{(H^k(\hat{K}))^3} + ||\hat{\nabla} \times (\hat{\mathbf{u}} + \hat{\mathbf{p}})||_{(H^k(\hat{K}))^3}). \tag{3.41}$$

Using the fact that [217, (5.12)]: If $\mathbf{v}, \nabla \times \mathbf{v} \in (H^s(K))^3$ for $0 \leq s \leq k$, then

$$\inf_{\phi \in Q^3_{k-1,k-1,k-1}} (||\mathbf{v} + \phi||_{(H^s(K))^3} + ||\nabla \times (\mathbf{v} + \phi)||_{(H^s(K))^3})$$
$$\leq C(|\mathbf{v}|_{(H^s(K))^3} + |\nabla \times \mathbf{v}|_{(H^s(K))^3} + |\nabla \times \mathbf{v}|_{(H^{[s]}(K))^3}),$$

where $[s]$ is the integer part of s, we obtain

$$||\hat{\mathbf{u}} - \Pi_{\hat{K}}^c \hat{\mathbf{u}}||_{(L^2(\hat{K}))^3} \leq C(|\hat{\mathbf{u}}|_{(H^k(K))^3} + |\hat{\nabla} \times \hat{\mathbf{u}}|_{(H^k(K))^3}). \tag{3.42}$$

Substituting (3.41) and (3.42) into (3.40) and using Lemma 3.10, we obtain

$$\begin{aligned}
||\mathbf{u} - \Pi_K^c \mathbf{u}||_{(L^2(K))^3} &\leq |det(B_K)|^{1/2} |B_K^{-1}| \cdot C(|\hat{\mathbf{u}}|_{(H^k(K))^3} + |\hat{\nabla} \times \hat{\mathbf{u}}|_{(H^k(K))^3}) \\
&\leq C|det(B_K)|^{1/2} |B_K^{-1}| \cdot (|B_K|^{k+1} |det(B_K)|^{-1/2} |\mathbf{u}|_{(H^k(K))^3} \\
&\quad + |B_K|^{k-1} |det(B_K)|^{1/2} |\nabla \times \mathbf{u}|_{(H^k(K))^3}) \\
&\leq C h_K^k (|\mathbf{u}|_{(H^k(K))^3} + |\nabla \times \mathbf{u}|_{(H^k(K))^3}),
\end{aligned}$$

which completes the proof for the L_2 error estimate.

Using (3.37) and Theorem 3.3, we can prove the curl estimate:

$$||\nabla \times (\mathbf{u} - \Pi_K^c \mathbf{u})||_{(L^2(K))^3} = ||(I - \Pi_K^d) \nabla \times \mathbf{u}||_{(L^2(K))^3} \leq C h^k ||\nabla \times \mathbf{u}||_{(H^k(K))^3}.$$

□

3.2.3 Finite Elements on Tetrahedra and Triangles

Before we construct a curl conforming finite element on a tetrahedron, we need to define a subspace of homogeneous vector polynomials of degree k denoted by

$$\mathscr{S}_k = \{\mathbf{p} \in (\tilde{P}_k)^3 : \ \mathbf{x} \cdot \mathbf{p} = 0\}. \tag{3.43}$$

Note that $\mathbf{x} \cdot \mathbf{p} \in \tilde{P}_{k+1}$, hence the dimension of $\mathscr{S}_k$ can be calculated as follows:

$$\begin{aligned}
dim(\mathscr{S}_k) &= 3dim(\tilde{P}_k) - dim(\tilde{P}_{k+1}) \\
&= 3(dim(P_k) - dim(P_{k-1})) - (dim(P_{k+1}) - dim(P_k))
\end{aligned}$$

$$= 3(\frac{(k+3)(k+2)(k+1)}{3!} - \frac{(k+2)(k+1)k}{3!})$$
$$-(\frac{(k+4)(k+3)(k+2)}{3!} - \frac{(k+3)(k+2)(k+1)}{3!}) = (k+2)k.$$

We need another important polynomial space

$$C_k = (P_{k-1})^3 \oplus \mathscr{S}_k. \tag{3.44}$$

It is easy to check that the dimension of C_K is

$$dim(C_k) = 3dim(P_{k-1}) + dim(\mathscr{S}_k) = \frac{1}{2}(k+3)(k+2)k.$$

Now we can define the curl conforming element on the reference tetrahedron $\hat{K}$ with four vertices: $(0,0,0)$, $(1,0,0)$, $(0,1,0)$ and $(0,0,1)$.

Definition 3.5. For any integer $k \geq 1$, the Nédélec curl conforming element is defined by the triple:

$$\hat{K} \text{ is the reference tetrahedron,}$$
$$P_{\hat{K}} = C_k,$$
$$\Sigma_{\hat{K}} = M_{\hat{e}}(\hat{\mathbf{u}}) \cup M_{\hat{f}}(\hat{\mathbf{u}}) \cup M_{\hat{K}}(\hat{\mathbf{u}}),$$

where $M_{\hat{e}}(\hat{\mathbf{u}})$, $M_{\hat{f}}(\hat{\mathbf{u}})$ and $M_{\hat{K}}(\hat{\mathbf{u}})$ are the sets of degrees of freedom given on edges of $\hat{K}$, faces of $\hat{K}$, and $\hat{K}$ itself:

$$M_{\hat{e}}(\hat{\mathbf{u}}) = \{\int_{\hat{e}_i} \hat{\mathbf{u}} \cdot \hat{\tau}_i \hat{q} d\hat{s},\ \forall\, \hat{q} \in P_{k-1}(\hat{e}_i),\ i = 1,\cdots,6\}, \tag{3.45}$$

$$M_{\hat{f}}(\hat{\mathbf{u}}) = \{\frac{1}{area(\hat{f}_i)} \int_{\hat{f}_i} \hat{\mathbf{u}} \cdot \hat{\mathbf{q}} d\hat{A},\ \forall\, \hat{\mathbf{q}} \in (P_{k-2}(\hat{f}_i))^3$$
$$\text{and } \hat{\mathbf{q}} \cdot \hat{\mathbf{n}}_i = 0,\ i = 1,\cdots,4\}, \tag{3.46}$$

$$M_{\hat{K}}(\hat{\mathbf{u}}) = \{\int_{\hat{K}} \hat{\mathbf{u}} \cdot \hat{\mathbf{q}} d\hat{V},\ \forall\, \hat{\mathbf{q}} \in (P_{k-3}(\hat{K}))^3\}. \tag{3.47}$$

Note that the face degrees of freedom defined by (3.46) look different from the original ones given by Nédélec [222], they are actually equivalent as remarked in Monk [217, p. 129]. Note that any $\hat{\mathbf{q}} \in (P_{k-2}(\hat{f}))^3$ satisfying $\hat{\mathbf{q}} \cdot \hat{\mathbf{n}} = 0$ can be written as $\hat{\mathbf{q}} = (\hat{\mathbf{n}} \times \hat{\mathbf{q}}) \times \hat{\mathbf{n}}$, from which we have

$$\int_{\hat{f}} \hat{\mathbf{u}} \cdot \hat{\mathbf{q}} d\hat{A} = \int_{\hat{f}} \hat{\mathbf{u}} \times \hat{\mathbf{n}} \cdot \hat{\mathbf{q}} \times \hat{\mathbf{n}} d\hat{A},$$

which is equivalent to

$$\int_{\hat{f}} \hat{\mathbf{u}} \times \hat{\mathbf{n}} \cdot \hat{\mathbf{r}} d\hat{A}, \quad \hat{\mathbf{r}} \in (P_{k-2}(\hat{f}))^2,$$

since $\hat{\mathbf{q}} \times \hat{\mathbf{n}} \in (P_{k-2}(\hat{f}))^2$.

Lemma 3.11. *Suppose that $det(B_K) > 0$ and the function $\mathbf{u}$ and the tangential vector $\boldsymbol{\tau}$ are obtained by the transformations (3.28) and (3.29). Then the degrees of freedom of $\mathbf{u}$ on K given by*

$$M_e(\mathbf{u}) = \{\int_{e_i} \mathbf{u} \cdot \boldsymbol{\tau}_i q ds, \ \forall\, q \in P_{k-1}(e_i), \ i = 1, \cdots, 6\},$$

$$M_f(\mathbf{u}) = \{\int_{f_i} \mathbf{u} \cdot \mathbf{q} dA, \quad \forall\, \mathbf{q} \circ F_K = B_K \hat{\mathbf{q}},$$

$$\hat{\mathbf{q}} \in (P_{k-2}(\hat{f}_i))^3, \ \hat{\mathbf{q}} \cdot \hat{\mathbf{n}}_i = 0, \ i = 1, \cdots, 4\},$$

$$M_K(\mathbf{u}) = \{\int_K \mathbf{u} \cdot \mathbf{q} dV, \ \forall\, \mathbf{q} \circ F_K = \frac{1}{det(B_K)} B_K \hat{\mathbf{q}}, \ \hat{\mathbf{q}} \in (P_{k-3}(\hat{K}))^3\},$$

are identical to the degrees of freedom for $\hat{\mathbf{u}}$ on $\hat{K}$ given in Definition 3.5.

The proof of this lemma is very similar to that given for Lemma 3.8. Details can be found in [217, Lemma 5.34]. Similarly, the finite element given in Lemma 3.11 is curl conforming and unisolvent. Readers interested in the detailed proof can consult [217, pp. 133–134].

Furthermore, we can construct the global curl conforming finite element space on a tetrahedral mesh T_h of Ω by

$$V_h = \{\mathbf{u} \in H(curl; \Omega) : \ \ \mathbf{u}|_K \in C_k \ \text{ for all } K \in T_h\}. \tag{3.48}$$

If $\mathbf{u}$ is smooth enough, then on any element $K \in T_h$ we can define the element-wise interpolant $\Pi_K^c \mathbf{u} \in C_k$ satisfying

$$M_e(\mathbf{u} - \Pi_K^c \mathbf{u}) = M_f(\mathbf{u} - \Pi_K^c \mathbf{u}) = M_K(\mathbf{u} - \Pi_K^c \mathbf{u}) = 0. \tag{3.49}$$

Hence we can define the global interpolant $\Pi_h^c \mathbf{u} \in V_h$ element by element:

$$(\Pi_h^c \mathbf{u})|_K = \Pi_K^c(\mathbf{u}|_K) \quad \forall\, K \in T_h.$$

Furthermore, we can prove that the global curl interpolant $\Pi_h^c \mathbf{u}$ and the global divergence interpolant $\Pi_h^d \mathbf{u}$ defined in Sect. 3.1 satisfy the relation [217, Lemma 5.40]:

$$\nabla \times \Pi_h^c \mathbf{u} = \Pi_h^d (\nabla \times \mathbf{u}).$$

Also the same interpolation error estimate stated in Theorem 3.7 holds true. Detailed proof can be found in [217, Theorem 5.41].

The above construction can be extended to triangular elements, in which case (3.43) and (3.44) become as:

$$C_k = (P_{k-1})^2 \oplus \mathscr{S}_k, \quad \mathscr{S}_k = \{\mathbf{p} \in (\tilde{P}_k)^2 : \ \mathbf{x} \cdot \mathbf{p} = 0\}. \tag{3.50}$$

It can be seen that on triangles,

$$\dim(C_k) = 2\dim(P_{k-1}) + \dim(\mathscr{S}_k) = 2 \cdot \frac{(k+1)k}{2} + k = k(k+2).$$

Similar to Definition 3.5, the curl conforming element on a reference triangle $\hat{K}$ can be formed by the following edge and element DOFs:

$$M_{\hat{e}}(\hat{\mathbf{u}}) = \{\int_{\hat{e}_i} \hat{\mathbf{u}} \cdot \hat{\tau}_i \hat{q} d\hat{s}, \ \forall \hat{q} \in P_{k-1}(\hat{e}_i), \ i = 1, 2, 3\}, \tag{3.51}$$

$$M_{\hat{K}}(\hat{\mathbf{u}}) = \{\int_{\hat{K}} \hat{\mathbf{u}} \cdot \hat{\mathbf{q}} d\hat{V}, \ \forall \hat{\mathbf{q}} \in (P_{k-2}(\hat{K}))^2\}. \tag{3.52}$$

It is easy to see that the total edge DOFs are $3k$, and the total element DOFs are $2 \cdot \frac{k(k-1)}{2}$, whose summation is

$$3k + 2 \cdot \frac{k(k-1)}{2} = k(k+2) = \dim(C_k).$$

Moreover, from (3.50) we easily write the spaces C_1 and C_2 on $\hat{K}$ as:

$$C_1 = \langle \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} \hat{y} \\ -\hat{x} \end{pmatrix} \rangle,$$

$$C_2 = (P_1(\hat{K}))^2 \oplus \langle \begin{pmatrix} \hat{x}\hat{y} \\ -\hat{x}^2 \end{pmatrix}, \begin{pmatrix} \hat{y}^2 \\ -\hat{x}\hat{y} \end{pmatrix} \rangle.$$

The Nédélec curl conforming element on general triangles can be obtained through transformations (3.28) and (3.29) and the degrees of freedom given by

$$M_e(\mathbf{u}) = \{\int_{e_i} \mathbf{u} \cdot \boldsymbol{\tau}_i q ds, \ \forall q \in P_{k-1}(e_i), \ i = 1, 2, 3\},$$

$$M_K(\mathbf{u}) = \{\int_K \mathbf{u} \cdot \mathbf{q} dV, \ \forall \mathbf{q} \circ F_K = \frac{1}{det(B_K)} B_K \hat{\mathbf{q}}, \ \hat{\mathbf{q}} \in (P_{k-2}(\hat{K}))^2\}.$$

Below we present two lowest-order curl conforming elements: one for tetrahedra, and another one for triangles.

Example 3.8. For $k = 1$ in (3.44), Nédélec [222] shows that

$$C_1 = \{\mathbf{u}(\mathbf{x}) = \mathbf{a} + \mathbf{b} \times \mathbf{x}, \quad \text{where } \mathbf{a}, \mathbf{b} \in R^3\},$$

where $\mathbf{a}$ and $\mathbf{b}$ are uniquely determined by the edge degrees of freedom $\int_e \mathbf{u} \times \boldsymbol{\tau}\, ds$ of K. Here K is assumed to be a general non-degenerate tetrahedron formed by vertices A_i, where $i = 1, 2, 3, 4$. From (3.44) and (3.43), we can write the space C_1 as:

$$C_1 = (P_0(\hat{K}))^3 \oplus \langle \begin{pmatrix} 0 \\ -x_3 \\ x_2 \end{pmatrix}, \begin{pmatrix} -x_3 \\ 0 \\ x_1 \end{pmatrix}, \begin{pmatrix} -x_2 \\ x_1 \\ 0 \end{pmatrix} \rangle.$$

To obtain a better form for the basis functions of C_1, we need to use the barycentric coordinate function λ_i corresponding to node A_i. More specifically, if we denote $\lambda_i = \alpha_{i0} + \alpha_{i1} x + \alpha_{i2} y + \alpha_{i3} z$, then $\lambda_i(A_j)$ satisfies

$$\lambda_i(A_j) = \delta_{i,j}, \quad i, j = 1, \cdots, 4, \tag{3.53}$$

which has a unique solution for each λ_i. For example, when $i = 1$, (3.53) can be written as follows:

$$\begin{pmatrix} 1 & x_1 & y_1 & z_1 \\ 1 & x_2 & y_2 & z_2 \\ 1 & x_3 & y_3 & z_3 \\ 1 & x_4 & y_4 & z_4 \end{pmatrix} \begin{pmatrix} \alpha_{10} \\ \alpha_{11} \\ \alpha_{12} \\ \alpha_{13} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix},$$

whose coefficient matrix has determinant as six times of the volume of K, and hence the system has a unique solution. With barycentric coordinate function λ_i, it can be shown that the basis function of C_1 with unit integral on an edge formed by vertices A_i and A_j is given by

$$\phi_{i,j} = \lambda_i \nabla \lambda_j - \lambda_j \nabla \lambda_i, \quad i, j = 1, \cdots, 4. \tag{3.54}$$

Note that elements such as C_1 depend on the edge degrees of freedom and are often called *edge elements*. C_1 is also called *Whitney element*, since Whitney [294] introduced this right framework in which to develop a finite element discretization of electromagnetic theory.

Example 3.9. Similarly, we can construct the lowest-order curl conforming element on a general triangle K formed by vertices $A_i, i = 1, 2, 3$. It can be shown that the basis function of C_1 on an edge formed by vertices A_i and A_j is given by

$$\phi_{i,j} = \lambda_i \nabla \lambda_j - \lambda_j \nabla \lambda_i, \quad i, j = 1, \cdots, 3.$$

3.3 Mathematical Analysis of the Drude Model

From Chap. 1, the governing equations used for modeling wave propagation in metamaterials with the Drude model can be written as:

$$\epsilon_0 \frac{\partial \mathbf{E}}{\partial t} = \nabla \times \mathbf{H} - \mathbf{J}, \tag{3.55}$$

$$\mu_0 \frac{\partial \mathbf{H}}{\partial t} = -\nabla \times \mathbf{E} - \mathbf{K}, \tag{3.56}$$

$$\frac{1}{\epsilon_0 \omega_{pe}^2} \frac{\partial \mathbf{J}}{\partial t} + \frac{\Gamma_e}{\epsilon_0 \omega_{pe}^2} \mathbf{J} = \mathbf{E}, \tag{3.57}$$

$$\frac{1}{\mu_0 \omega_{pm}^2} \frac{\partial \mathbf{K}}{\partial t} + \frac{\Gamma_m}{\mu_0 \omega_{pm}^2} \mathbf{K} = \mathbf{H}. \tag{3.58}$$

For simplicity, we assume that the modeling domain is $\Omega \times (0, T)$, where Ω is a bounded Lipschitz polyhedral domain in $\mathscr{R}^3$ with connected boundary $\partial\Omega$. Furthermore, we assume that the boundary of Ω is perfect conducting so that

$$\mathbf{n} \times \mathbf{E} = \mathbf{0} \quad \text{on } \partial\Omega, \tag{3.59}$$

where $\mathbf{n}$ is the unit outward normal to $\partial\Omega$. Also, the initial conditions for (3.55)–(3.58) are assumed to be as follows:

$$\mathbf{E}(\mathbf{x}, 0) = \mathbf{E}_0(\mathbf{x}), \quad \mathbf{H}(\mathbf{x}, 0) = \mathbf{H}_0(\mathbf{x}), \tag{3.60}$$

$$\mathbf{J}(\mathbf{x}, 0) = \mathbf{J}_0(\mathbf{x}), \quad \mathbf{K}(\mathbf{x}, 0) = \mathbf{K}_0(\mathbf{x}), \tag{3.61}$$

where $\mathbf{E}_0(\mathbf{x}), \mathbf{H}_0(\mathbf{x}), \mathbf{J}_0(\mathbf{x})$ and $\mathbf{K}_0(\mathbf{x})$ are some given functions.

First, we can show that the model problem (3.55)–(3.61) is stable.

Lemma 3.12. *The solution* $(\mathbf{E}, \mathbf{H}, \mathbf{J}, \mathbf{K})$ *of the problem (3.55)–(3.61) satisfies the following stability estimate:*

$$\begin{aligned} &\epsilon_0 ||\mathbf{E}(t)||_0^2 + \mu_0 ||\mathbf{H}(t)||_0^2 + \frac{1}{\mu_0 \omega_{pm}^2} ||\mathbf{K}(t)||_0^2 + \frac{1}{\epsilon_0 \omega_{pe}^2} ||\mathbf{J}(t)||_0^2 \\ &\le \epsilon_0 ||\mathbf{E}(0)||_0^2 + \mu_0 ||\mathbf{H}(0)||_0^2 + \frac{1}{\mu_0 \omega_{pm}^2} ||\mathbf{K}(0)||_0^2 + \frac{1}{\epsilon_0 \omega_{pe}^2} ||\mathbf{J}(0)||_0^2. \end{aligned} \tag{3.62}$$

Proof. Multiplying Eq. (3.55)–(3.58) by $\mathbf{E}, \mathbf{H}, \mathbf{J}, \mathbf{K}$ and integrating over the domain Ω, respectively, adding the resultants together, and using the identity

$$\int_\Omega \nabla \times \mathbf{H} \cdot \mathbf{E} d\mathbf{x} = \int_\Omega \mathbf{H} \cdot \nabla \times \mathbf{E} d\mathbf{x} - \int_{\partial\Omega} \mathbf{H} \cdot \mathbf{n} \times \mathbf{E} ds$$

and the boundary condition (3.59), we obtain

$$\frac{1}{2}\frac{d}{dt}[\epsilon_0||\mathbf{E}(t)||_0^2 + \mu_0||\mathbf{H}(t)||_0^2 + \frac{1}{\mu_0\omega_{pm}^2}||\mathbf{K}(t)||_0^2 + \frac{1}{\epsilon_0\omega_{pe}^2}||\mathbf{J}(t)||_0^2]$$
$$+\frac{\Gamma_m}{\mu_0\omega_{pm}^2}||\mathbf{K}(t)||_0^2 + \frac{\Gamma_e}{\epsilon_0\omega_{pe}^2}||\mathbf{J}(t)||_0^2 = 0,$$

which easily leads to the stability estimate (3.62). □

Now we want to show that the model problem (3.55)–(3.61) exists a unique solution.

Theorem 3.8. *There exists a unique solution* $\mathbf{E}\in H_0(curl;\Omega)$ *and* $\mathbf{H}\in H(curl;\Omega)$ *for the system (3.55)–(3.61).*

Proof. Let us denote the Laplace transform of a function $f(t)$ defined for $t \geq 0$ by $\hat{f}(s) = \int_0^\infty e^{-st} f(t)dt$. Taking the Laplace transform of (3.55)–(3.58), we have

$$\epsilon_0(s\hat{\mathbf{E}} - \mathbf{E}_0) = \nabla\times\hat{\mathbf{H}} - \hat{\mathbf{J}}, \tag{3.63}$$

$$\mu_0(s\hat{\mathbf{H}} - \mathbf{H}_0) = -\nabla\times\hat{\mathbf{E}} - \hat{\mathbf{K}}, \tag{3.64}$$

$$(s+\Gamma_e)\hat{\mathbf{J}} = \mathbf{J}_0 + \epsilon_0\omega_{pe}^2\hat{\mathbf{E}}, \tag{3.65}$$

$$(s+\Gamma_m)\hat{\mathbf{K}} = \mathbf{K}_0 + \mu_0\omega_{pm}^2\hat{\mathbf{H}}. \tag{3.66}$$

Combining (3.63) with (3.65), we obtain

$$\epsilon_0[s(s+\Gamma_e) + \omega_{pe}^2]\hat{\mathbf{E}} = (s+\Gamma_e)\nabla\times\hat{\mathbf{H}} + \epsilon_0(s+\Gamma_e)\mathbf{E}_0 - \mathbf{J}_0. \tag{3.67}$$

Similarly, combining (3.64) with (3.66), we obtain

$$\mu_0[s(s+\Gamma_m) + \omega_{pm}^2]\hat{\mathbf{H}} = \mu_0(s+\Gamma_m)\mathbf{H}_0 - (s+\Gamma_m)\nabla\times\hat{\mathbf{E}} - \mathbf{K}_0,$$

whose curl gives

$$\mu_0[s(s+\Gamma_m) + \omega_{pm}^2]\nabla\times\hat{\mathbf{H}}$$
$$= \mu_0(s+\Gamma_m)\nabla\times\mathbf{H}_0 - (s+\Gamma_m)\nabla\times\nabla\times\hat{\mathbf{E}} - \nabla\times\mathbf{K}_0. \tag{3.68}$$

Adding the result of (3.67) multiplied by $\mu_0[s(s+\Gamma_m) + \omega_{pm}^2]$ to the result of (3.68) multiplied by $(s+\Gamma_e)$, we have

$$\epsilon_0\mu_0[s(s+\Gamma_e) + 2\omega_{pe}^2][s(s+\Gamma_m) + \omega_{pm}^2]\hat{\mathbf{E}} + (s+\Gamma_m)(s+\Gamma_e)\nabla\times\nabla\times\hat{\mathbf{E}}$$
$$= \mu_0[s(s+\Gamma_m) + \omega_{pm}^2][\epsilon_0(s+\Gamma_e)\mathbf{E}_0 - \mathbf{J}_0]$$
$$+(s+\Gamma_e)[\mu_0(s+\Gamma_m)\nabla\times\mathbf{H}_0 - \nabla\times\mathbf{K}_0]. \tag{3.69}$$

A weak formulation of (3.69) is: Find $\hat{\mathbf{E}} \in H_0(curl;\Omega)$ such that

$$\begin{aligned}
&\epsilon_0\mu_0[s(s+\Gamma_e)+2\omega_{pe}^2][s(s+\Gamma_m)+\omega_{pm}^2](\hat{\mathbf{E}},\boldsymbol{\phi})\\
&+(s+\Gamma_m)(s+\Gamma_e)(\nabla\times\hat{\mathbf{E}},\nabla\times\boldsymbol{\phi})\\
=\ &\mu_0[s(s+\Gamma_m)+\omega_{pm}^2](\epsilon_0(s+\Gamma_e)\mathbf{E}_0-\mathbf{J}_0,\boldsymbol{\phi})\\
&+(s+\Gamma_e)(\mu_0(s+\Gamma_m)\nabla\times\mathbf{H}_0-\nabla\times\mathbf{K}_0,\boldsymbol{\phi}) \quad \forall\,\boldsymbol{\phi}\in H_0(curl;\Omega),
\end{aligned}$$

which has a unique solution by the Lax-Milgram lemma. The inverse Laplace transform of the function $\hat{\mathbf{E}}$ is the solution $\mathbf{E}$ of (3.55)–(3.61), and the uniqueness of $\mathbf{E}$ follows from the uniqueness of the Laplace transform.

Existence and uniqueness of solution $\mathbf{H}$ can be proved similarly. □

Finally, we can prove that the electric and magnetic fields also satisfy the Gauss' law if the initial fields are divergence free. More specifically, we have

Lemma 3.13. *Assume that the initial conditions are divergence free, i.e.,*

$$\nabla\cdot(\epsilon_0\mathbf{E}_0)=0,\ \ \nabla\cdot(\mu_0\mathbf{H}_0)=0,\ \ \nabla\cdot\mathbf{J}_0=0,\ \ \nabla\cdot\mathbf{K}_0=0. \tag{3.70}$$

Then for any time $t>0$, we have

$$\nabla\cdot(\epsilon_0\mathbf{E}(t))=0,\ \ \nabla\cdot(\mu_0\mathbf{H}(t))=0,\ \ \nabla\cdot\mathbf{J}(t)=0,\ \ \nabla\cdot\mathbf{K}(t)=0.$$

Proof. Taking the divergence of (3.55), we have

$$\frac{\partial}{\partial t}(\nabla\cdot(\epsilon_0\mathbf{E}))=-\nabla\cdot\mathbf{J}. \tag{3.71}$$

Then taking the divergence of (3.57), we have

$$\frac{\partial}{\partial t}(\nabla\cdot\mathbf{J})+\Gamma_e\nabla\cdot\mathbf{J}=\omega_{pe}^2\nabla\cdot(\epsilon_0\mathbf{E}). \tag{3.72}$$

Substituting (3.71) into (3.72), we obtain a second-order constant coefficient ordinary differential equation

$$\left(\frac{\partial^2}{\partial t^2}+\Gamma_e\frac{\partial}{\partial t}+\omega_{pe}^2\right)\nabla\cdot(\epsilon_0\mathbf{E})=0, \tag{3.73}$$

which has initial conditions (from (3.70) and (3.71))

$$\nabla\cdot(\epsilon_0\mathbf{E})(0)=\frac{\partial}{\partial t}(\nabla\cdot(\epsilon_0\mathbf{E}))(0)=0. \tag{3.74}$$

From the basic theory of ordinary differential equation, we know that the problem (3.73) and (3.74) only has zero solution, i.e.,

$$\nabla \cdot (\epsilon_0 \mathbf{E}(t)) = 0,$$

substituting which into (3.71) leads to $\nabla \cdot \mathbf{J}(t) = 0$.

By symmetry, we can prove $\nabla \cdot (\mu_0 \mathbf{H}(t)) = 0$ and $\nabla \cdot \mathbf{K}(t) = 0$. □

3.4 The Crank-Nicolson Scheme for the Drude Model

3.4.1 The Raviart-Thomas-Nédélec Finite Elements

To design a finite element method, we assume that Ω is partitioned by a family of regular tetrahedral (or cubic) meshes T_h with maximum mesh size h. Depending upon the regularity of the solution of the problem, we can use proper order divergence and curl conforming (often called as Raviart-Thomas-Nédélec) tetrahedral elements discussed in Sects. 3.1 and 3.2: For any $l \geq 1$,

$$\mathbf{U}_h = \{\mathbf{u}_h \in H(div; \Omega) : \ \mathbf{u}_h|_K \in (p_{l-1})^3 \oplus \tilde{p}_{l-1}\mathbf{x}, \ \ \forall\ K \in T_h\}, \tag{3.75}$$

$$\mathbf{V}_h = \{\mathbf{v}_h \in H(\text{curl}; \Omega) : \ \mathbf{v}_h|_K \in (p_{l-1})^3 \oplus S_l, \ \ \forall\ K \in T_h\}, \tag{3.76}$$

where the space

$$S_l = \{\mathbf{p} \in (\tilde{p}_l)^3, \quad \mathbf{x} \cdot \mathbf{p} = 0\},$$

or Raviart-Thomas-Nédélec cubic elements:

$$\mathbf{U}_h = \{\mathbf{u}_h \in H(div; \Omega) : \mathbf{u}_h|_K \in Q_{l,l-1,l-1} \times Q_{l-1,l,l-1} \times Q_{l-1,l-1,l}, \forall K \in T_h\},$$

$$\mathbf{V}_h = \{\mathbf{v}_h \in H(\text{curl}; \Omega) : \ \mathbf{v}_h|_K \in Q_{l-1,l,l} \times Q_{l,l-1,l} \times Q_{l,l,l-1}, \ \forall\ K \in T_h\}.$$

Recall that $\tilde{p}_k$ denotes the space of homogeneous polynomials of degree k, and $Q_{i,j,k}$ denotes the space of polynomials whose degrees are less than or equal to i, j, k in variables x, y, z, respectively. To impose the boundary condition $\mathbf{n} \times \mathbf{E} = \mathbf{0}$ on the boundary $\partial\Omega$ of Ω, we introduce a subspace of $\mathbf{V}_h$:

$$\mathbf{V}_h^0 = \{\mathbf{v} \in \mathbf{V}_h : \mathbf{v} \times \mathbf{n} = \mathbf{0} \ \text{ on } \partial\Omega\}.$$

In the error analysis below, we will often use the following fact that

$$\nabla \times \mathbf{V}_h \subset \mathbf{U}_h. \tag{3.77}$$

Also we will need two operators. The first one is the standard $L^2(\Omega)$-projection operator: For any $\mathbf{H} \in (L^2(\Omega))^d$, $P_h\mathbf{H} \in \mathbf{U}_h$ satisfies

$$(P_h\mathbf{H} - \mathbf{H}, \boldsymbol{\psi}_h) = 0, \quad \forall\ \psi_h \in \mathbf{U}_h.$$

Another one is the standard Nédélec interpolation operator Π_h^c mapped from $H(\text{curl};\Omega)$ to $\mathbf{V}_h$. To simplify the notation, we will just use Π_h for Π_h^c in the rest of this chapter.

Recall that we have the following interpolation error estimate: For any $\mathbf{E} \in H^l(\text{curl};\Omega),\ 1 \le l$, we have

$$||\mathbf{E} - \Pi_h\mathbf{E}||_0 + ||\nabla \times (\mathbf{E} - \Pi_h\mathbf{E})||_0 \le Ch^l||\mathbf{E}||_{l,\text{curl}}, \tag{3.78}$$

and the projection error estimate:

$$||\mathbf{H} - P_h\mathbf{H}||_0 \le Ch^l||\mathbf{H}||_l, \quad \forall\ \mathbf{H} \in (H^l(\Omega))^d,\ 0 \le l. \tag{3.79}$$

To define a fully discrete scheme, we divide the time interval $[0, T]$ into M uniform subintervals by points $t_k = k\tau$, where $\tau = \frac{T}{M}$ and $k = 0, 1, \cdots, M$.

3.4.2 *The Scheme and Its Stability Analysis*

Now we can formulate a Crank-Nicolson mixed finite element scheme for (3.55)–(3.58): for $k = 1, 2, \cdots, M$, find $\mathbf{E}_h^k \in \mathbf{V}_h^0, \mathbf{J}_h^k \in \mathbf{V}_h, \mathbf{H}_h^k, \mathbf{K}_h^k \in \mathbf{U}_h$ such that

$$\epsilon_0(\delta_\tau\mathbf{E}_h^k, \boldsymbol{\phi}_h) - (\overline{\mathbf{H}}_h^k, \nabla \times \boldsymbol{\phi}_h) + (\overline{\mathbf{J}}_h^k, \boldsymbol{\phi}_h) = 0, \tag{3.80}$$

$$\mu_0(\delta_\tau\mathbf{H}_h^k, \boldsymbol{\psi}_h) + (\nabla \times \overline{\mathbf{E}}_h^k, \boldsymbol{\psi}_h) + (\overline{\mathbf{K}}_h^k, \boldsymbol{\psi}_h) = 0, \tag{3.81}$$

$$\frac{1}{\epsilon_0\omega_{pe}^2}(\delta_\tau\mathbf{J}_h^k, \tilde{\boldsymbol{\phi}}_h) + \frac{\Gamma_e}{\epsilon_0\omega_{pe}^2}(\overline{\mathbf{J}}_h^k, \tilde{\boldsymbol{\phi}}_h) = (\overline{\mathbf{E}}_h^k, \tilde{\boldsymbol{\phi}}_h), \tag{3.82}$$

$$\frac{1}{\mu_0\omega_{pm}^2}(\delta_\tau\mathbf{K}_h^k, \tilde{\boldsymbol{\psi}}_h) + \frac{\Gamma_m}{\mu_0\omega_{pm}^2}(\overline{\mathbf{K}}_h^k, \tilde{\boldsymbol{\psi}}_h) = (\overline{\mathbf{H}}_h^k, \tilde{\boldsymbol{\psi}}_h), \tag{3.83}$$

for any $\boldsymbol{\phi}_h \in \mathbf{V}_h^0$, $\boldsymbol{\psi}_h \in \mathbf{U}_h$, $\tilde{\boldsymbol{\phi}}_h \in \mathbf{V}_h$, $\tilde{\boldsymbol{\psi}}_h \in \mathbf{U}_h$, subject to the initial conditions

$$\mathbf{E}_h^0(\mathbf{x}) = \Pi_h\mathbf{E}_0(\mathbf{x}), \quad \mathbf{H}_h^0(\mathbf{x}) = P_h\mathbf{H}_0(\mathbf{x}),$$
$$\mathbf{J}_h^0(\mathbf{x}) = \Pi_h\mathbf{J}_0(\mathbf{x}), \quad \mathbf{K}_h^0(\mathbf{x}) = P_h\mathbf{K}_0(\mathbf{x}).$$

In (3.80)–(3.83), we use the central difference and average operators at time lever $k + \frac{1}{2}$:

$$\delta_\tau u^k = (u^k - u^{k-1})/\tau, \quad \overline{u}^k = (u^k + u^{k-1})/2,$$

where $u^k = u(k\tau)$.

First, let us look at the scheme (3.80)–(3.83) carefully. It can be seen that (3.82) and (3.83) are equivalent to

$$\mathbf{J}_h^k = \frac{\epsilon_0\omega_{pe}^2}{2\tau^{-1}+\Gamma_e}(\mathbf{E}_h^k+\mathbf{E}_h^{k-1}) + \frac{2\tau^{-1}-\Gamma_e}{2\tau^{-1}+\Gamma_e}\mathbf{J}_h^{k-1}, \tag{3.84}$$

$$\mathbf{K}_h^k = \frac{\mu_0\omega_{pm}^2}{2\tau^{-1}+\Gamma_m}(\mathbf{H}_h^k+\mathbf{H}_h^{k-1}) + \frac{2\tau^{-1}-\Gamma_m}{2\tau^{-1}+\Gamma_m}\mathbf{K}_h^{k-1}. \tag{3.85}$$

Then substituting (3.84) and (3.85) into (3.80) and (3.81), respectively, we obtain

$$\begin{aligned}(\frac{2\epsilon_0}{\tau} + \frac{\epsilon_0\omega_{pe}^2}{2\tau^{-1}+\Gamma_e})(\mathbf{E}_h^k,\boldsymbol{\phi}_h) - (\mathbf{H}_h^k,\nabla\times\boldsymbol{\phi}_h) &= -\frac{4\tau^{-1}}{2\tau^{-1}+\Gamma_e}(\mathbf{J}_h^{k-1},\boldsymbol{\phi}_h)\\ +(\frac{2\epsilon_0}{\tau} - \frac{\epsilon_0\omega_{pe}^2}{2\tau^{-1}+\Gamma_e})(\mathbf{E}_h^{k-1},\boldsymbol{\phi}_h) &+ (\mathbf{H}_h^{k-1},\nabla\times\boldsymbol{\phi}_h),\end{aligned} \tag{3.86}$$

$$\begin{aligned}(\frac{2\mu_0}{\tau} + \frac{\mu_0\omega_{pm}^2}{2\tau^{-1}+\Gamma_m})(\mathbf{H}_h^k,\boldsymbol{\psi}_h) + (\nabla\times\mathbf{E}_h^k,\boldsymbol{\psi}_h) &= -\frac{4\tau^{-1}}{2\tau^{-1}+\Gamma_m}(\mathbf{K}_h^{k-1},\boldsymbol{\psi}_h)\\ +(\frac{2\mu_0}{\tau} - \frac{\mu_0\omega_{pm}^2}{2\tau^{-1}+\Gamma_m})(\mathbf{H}_h^{k-1},\boldsymbol{\psi}_h) &- (\nabla\times\mathbf{E}_h^{k-1},\boldsymbol{\psi}_h).\end{aligned} \tag{3.87}$$

Hence, to solve the system (3.80)–(3.83) at each time step, we just need to solve the smaller system (3.86) and (3.87) for $\mathbf{E}_h^k$ and $\mathbf{H}_h^k$, then update $\mathbf{J}_h^k$ and $\mathbf{K}_h^k$ using (3.84) and (3.85).

We want to assure that the system (3.86) and (3.87) is invertible.

Lemma 3.14. *At each time step, the system (3.86) and (3.87) is uniquely solvable.*

Proof. Note that the coefficient matrix for the system (3.86) and (3.87) with the vector solution $(\mathbf{E}_h^k,\mathbf{H}_h^k)'$ can be written as

$$Q \equiv \begin{pmatrix} A & -B \\ B' & D \end{pmatrix},$$

where the matrices $A = (\frac{2\epsilon_0}{\tau}+\frac{\epsilon_0\omega_{pe}^2}{2\tau^{-1}+\Gamma_e})(\boldsymbol{\phi}_h,\boldsymbol{\phi}_h)$ and $D = (\frac{2\mu_0}{\tau}+\frac{\mu_0\omega_{pm}^2}{2\tau^{-1}+\Gamma_m})(\boldsymbol{\psi}_h,\boldsymbol{\psi}_h)$ are symmetric positive definite, and the matrix $B = (\boldsymbol{\psi}_h,\nabla\times\boldsymbol{\phi}_h)$. Here $\boldsymbol{\phi}_h$ and $\boldsymbol{\psi}_h$ are arbitrary functions from $\mathbf{V}_h^0$ and $\mathbf{U}_h$, respectively.

It is easy to check that the determinant of Q equals $\det(A)\det(D+B'A^{-1}B)$, which is obviously non-zero. Hence, Q is non-singular, which concludes the proof. □

Finally, we want to show that the scheme (3.80)–(3.83) is unconditionally stable and has a discrete stability similar to the continuous case stated in Lemma 3.12.

Lemma 3.15. *For the solution of (3.80)–(3.83), we have*

$$\epsilon_0||\mathbf{E}_h^k||_0^2 + \mu_0||\mathbf{H}_h^k||_0^2 + \frac{1}{\epsilon_0\omega_{pe}^2}||\mathbf{J}_h^k||_0^2 + \frac{1}{\mu_0\omega_{pm}^2}||\mathbf{K}_h^k||_0^2$$
$$\leq \epsilon_0||\mathbf{E}_h^0||_0^2 + \mu_0||\mathbf{H}_h^0||_0^2 + \frac{1}{\epsilon_0\omega_{pe}^2}||\mathbf{J}_h^0||_0^2 + \frac{1}{\mu_0\omega_{pm}^2}||\mathbf{K}_h^0||_0^2.$$

Proof. Choosing $\boldsymbol{\phi}_h = \tau(\mathbf{E}_h^k + \mathbf{E}_h^{k-1}), \boldsymbol{\psi}_h = \tau(\mathbf{H}_h^k + \mathbf{H}_h^{k-1}), \tilde{\boldsymbol{\phi}}_h = \tau(\mathbf{J}_h^k + \mathbf{J}_h^{k-1})$, $\tilde{\boldsymbol{\psi}}_h = \tau(\mathbf{K}_h^k + \mathbf{K}_h^{k-1})$ in (3.80)–(3.83), respectively, and adding the resultants together, we obtain

$$\epsilon_0(||\mathbf{E}_h^k||_0^2 - \mathbf{E}_h^{k-1}||_0^2) + \mu_0(||\mathbf{H}_h^k||_0^2 - ||\mathbf{H}_h^{k-1}||_0^2)$$
$$+\frac{1}{\epsilon_0\omega_{pe}^2}(||\mathbf{J}_h^k||_0^2 - ||\mathbf{J}_h^{k-1}||_0^2) + \frac{\tau\Gamma_e}{2\epsilon_0\omega_{pe}^2}||\mathbf{J}_h^k + \mathbf{J}_h^{k-1}||_0^2$$
$$+\frac{1}{\mu_0\omega_{pm}^2}(||\mathbf{K}_h^k||_0^2 - ||\mathbf{K}_h^{k-1}||_0^2) + \frac{\tau\Gamma_m}{2\mu_0\omega_{pm}^2}||\mathbf{K}_h^k + \mathbf{K}_h^{k-1}||_0^2 = 0,$$

from which it is easy to obtain the following unconditional stability

$$\epsilon_0||\mathbf{E}_h^k||_0^2 + \mu_0||\mathbf{H}_h^k||_0^2 + \frac{1}{\epsilon_0\omega_{pe}^2}||\mathbf{J}_h^k||_0^2 + \frac{1}{\mu_0\omega_{pm}^2}||\mathbf{K}_h^k||_0^2$$
$$\leq \epsilon_0||\mathbf{E}_h^{k-1}||_0^2 + \mu_0||\mathbf{H}_h^{k-1}||_0^2 + \frac{1}{\epsilon_0\omega_{pe}^2}||\mathbf{J}_h^{k-1}||_0^2 + \frac{1}{\mu_0\omega_{pm}^2}||\mathbf{K}_h^{k-1}||_0^2$$
$$\leq \cdots \leq \epsilon_0||\mathbf{E}_h^0||_0^2 + \mu_0||\mathbf{H}_h^0||_0^2 + \frac{1}{\epsilon_0\omega_{pe}^2}||\mathbf{J}_h^0||_0^2 + \frac{1}{\mu_0\omega_{pm}^2}||\mathbf{K}_h^0||_0^2.$$

□

3.4.3 The Optimal Error Estimate

In this section, we shall prove that the Crank-Nicolson scheme (3.80)–(3.83) is optimally convergent. To prove that, we need the following estimates.

Lemma 3.16. *Denote* $\overline{\mathbf{u}}^k = \frac{1}{2}(\mathbf{u}^k + \mathbf{u}^{k-1})$. *Then we have*

$$(i)\ \ \|\delta_\tau \mathbf{u}^k\|_0^2 = \|\frac{\mathbf{u}^k - \mathbf{u}^{k-1}}{\tau}\|_0^2 \leq \frac{1}{\tau}\int_{t_{k-1}}^{t_k} \|\mathbf{u}_t(t)\|_0^2 dt \quad \forall\, \mathbf{u} \in H^1(0,T;(L^2(\Omega))^3),$$

$$(ii)\ \ \|\overline{\mathbf{u}}^k - \frac{1}{\tau}\int_{t_{k-1}}^{t_k} \mathbf{u}(t)dt\|_0^2 \leq \frac{\tau^3}{4}\int_{t_{k-1}}^{t_k} \|\mathbf{u}_{tt}(t)\|_0^2 dt \quad \forall \mathbf{u} \in H^2(0,T;(L^2(\Omega))^3).$$

Proof. (i) The proof follows by squaring the identity

$$\delta_\tau \mathbf{u}^k = \frac{1}{\tau}\int_{t_{k-1}}^{t_k} \mathbf{u}_t(t)dt$$

and using the Cauchy-Schwarz inequality.

(ii) Squaring both sides of the integral identity

$$\frac{1}{2}(\mathbf{u}^k + \mathbf{u}^{k-1}) - \frac{1}{\tau}\int_{t_{k-1}}^{t_k} \mathbf{u}(t)dt = \frac{1}{2\tau}\int_{t_{k-1}}^{t_k} (t - t_{k-1})(t_k - t)\mathbf{u}_{tt}(t)dt, \quad (3.88)$$

we can obtain

$$|\overline{\mathbf{u}}^k - \frac{1}{\tau}\int_{t_{k-1}}^{t_k} \mathbf{u}(t)dt|^2 \le \frac{1}{4\tau^2}(\int_{t_{k-1}}^{t_k} (t - t_{k-1})^2(t_k - t)^2 dt)(\int_{t_{k-1}}^{t_k} |\mathbf{u}_{tt}(t)|^2 dt)$$
$$\le \frac{1}{4}\tau^3 \int_{t_{k-1}}^{t_k} |\mathbf{u}_{tt}(t)|^2 dt,$$

integrating which over Ω concludes the proof.

□

Below we will often use the so-called discrete Gronwall inequality ([243, p. 14], [114]).

Theorem 3.9. *Let $f(t)$ and $g(t)$ be nonnegative functions defined on $t_j = j\tau$, $j = 0, 1, \cdots, M$, and $g(t)$ be non-decreasing. If*

$$(t_k) \le g(t_k) + r\tau \sum_{j=0}^{k-1} f(t_j),$$

where r is a positive constant, then we have

$$f(t_k) \le g(t_k)\exp(kr\tau).$$

Now we can prove the following optimal error estimate for the scheme (3.80)–(3.83).

Theorem 3.10. *Let $(\mathbf{E}^n, \mathbf{H}^n)$ and $(\mathbf{E}_h^n, \mathbf{H}_h^n)$ be the analytic and finite element solutions at time $t = t_n$, respectively. Under the regularity assumptions*

$$\mathbf{H}, \mathbf{K} \in (L^2(0, T; (H^l(\Omega))^3))^3,$$
$$\mathbf{E}, \mathbf{J} \in L^\infty(0, T; H^l(curl; \Omega)), \quad \mathbf{E}_t, \mathbf{J}_t \in (L^2(0, T; H^l(curl; \Omega)))^3,$$
$$\mathbf{H}_{tt}, \mathbf{E}_{tt}, \mathbf{K}_{tt}, \mathbf{J}_{tt}, \nabla \times \mathbf{H}_{tt}, \nabla \times \mathbf{E}_{tt} \in (L^2(0, T; (L^2(\Omega))^3))^3,$$

there exists a constant $C = C(T, \epsilon_0, \mu_0, \omega_{pe}, \omega_{pm}, \Gamma_e, \Gamma_m, \mathbf{E}, \mathbf{H}, \mathbf{K}, \mathbf{L})$, independent of both time step τ and mesh size h, such that

$$\max_{1\le n\le M}(||\mathbf{E}^n-\mathbf{E}_h^n||_0+||\mathbf{H}^n-\mathbf{H}_h^n||_0+||\mathbf{J}^n-\mathbf{J}_h^n||_0+||\mathbf{K}^n-\mathbf{K}_h^n||_0)\le C(\tau^2+h^l),$$

where $l\ge 1$ is the order of basis functions in spaces $\mathbf{U}_h$ and $\mathbf{V}_h$.

Proof. Multiplying (3.55)–(3.58) by $\frac{1}{\tau}\boldsymbol{\phi}_h\in\mathbf{V}_h^0$, $\frac{1}{\tau}\boldsymbol{\psi}_h\in\mathbf{U}_h$, $\frac{1}{\tau}\tilde{\boldsymbol{\phi}}_h\in\mathbf{V}_h$, $\frac{1}{\tau}\tilde{\boldsymbol{\psi}}_h\in\mathbf{U}_h$, respectively, integrating the resultants in time over $I^k=[t_{k-1},t_k]$ and in space over Ω, then using the Stokes' formula

$$\int_\Omega\nabla\times\mathbf{E}\cdot\boldsymbol{\psi}=\int_{\partial\Omega}n\times\mathbf{E}\cdot\boldsymbol{\psi}+\int_\Omega\mathbf{E}\cdot\nabla\times\boldsymbol{\psi},\tag{3.89}$$

we obtain

$$\epsilon_0(\delta_\tau\mathbf{E}^k,\boldsymbol{\phi}_h)-(\frac{1}{\tau}\int_{I^k}\mathbf{H}(s)ds,\nabla\times\boldsymbol{\phi}_h)+(\frac{1}{\tau}\int_{I^k}\mathbf{J}(s)ds,\boldsymbol{\phi}_h)=0,\tag{3.90}$$

$$\mu_0(\delta_\tau\mathbf{H}^k,\boldsymbol{\psi}_h)+(\nabla\times\frac{1}{\tau}\int_{I^k}\mathbf{E}(s)ds,\boldsymbol{\psi}_h)+(\frac{1}{\tau}\int_{I^k}\mathbf{K}(s)ds,\boldsymbol{\psi}_h)=0,\tag{3.91}$$

$$\frac{1}{\epsilon_0\omega_{pe}^2}(\delta_\tau\mathbf{J}^k,\tilde{\boldsymbol{\phi}}_h)+\frac{\Gamma_e}{\epsilon_0\omega_{pe}^2}(\frac{1}{\tau}\int_{I^k}\mathbf{J}(s)ds,\tilde{\boldsymbol{\phi}}_h)=(\frac{1}{\tau}\int_{I^k}\mathbf{E}(s)ds,\tilde{\boldsymbol{\phi}}_h),\tag{3.92}$$

$$\frac{1}{\mu_0\omega_{pm}^2}(\delta_\tau\mathbf{K}^k,\tilde{\boldsymbol{\psi}}_h)+\frac{\Gamma_m}{\mu_0\omega_{pm}^2}(\frac{1}{\tau}\int_{I^k}\mathbf{K}(s)ds,\tilde{\boldsymbol{\psi}}_h)=(\frac{1}{\tau}\int_{I^k}\mathbf{H}(s)ds,\tilde{\boldsymbol{\psi}}_h).\tag{3.93}$$

Denote $\xi_h^k=\Pi_h\mathbf{E}^k-\mathbf{E}_h^k$, $\eta_h^k=P_h\mathbf{H}^k-\mathbf{H}_h^k$, $\tilde{\xi}_h^k=\Pi_h\mathbf{J}^k-\mathbf{J}_h^k$, $\tilde{\eta}_h^k=P_h\mathbf{K}^k-\mathbf{K}_h^k$. Subtracting (3.80)–(3.83) from (3.90)–(3.93), respectively, we obtain the error equations

$$\begin{aligned}
(i)\quad &\epsilon_0(\delta_\tau\xi_h^k,\boldsymbol{\phi}_h)-(\overline{\eta}_h^k,\nabla\times\boldsymbol{\phi}_h)=\epsilon_0(\delta_\tau(\Pi_h\mathbf{E}^k-\mathbf{E}^k),\boldsymbol{\phi}_h)\\
&-(P_h\overline{\mathbf{H}}^k-\frac{1}{\tau}\int_{I^k}\mathbf{H}(s)ds,\nabla\times\boldsymbol{\phi}_h)-(\frac{1}{\tau}\int_{I^k}\mathbf{J}(s)ds-\overline{\mathbf{J}}_h^k,\boldsymbol{\phi}_h),\\
(ii)\quad &\mu_0(\delta_\tau\eta_h^k,\boldsymbol{\psi}_h)+(\nabla\times\overline{\xi}_h^k,\boldsymbol{\psi}_h)=\mu_0(\delta_\tau(P_h\mathbf{H}^k-\mathbf{H}^k),\boldsymbol{\psi}_h)\\
&+(\nabla\times(\Pi_h\overline{\mathbf{E}}^k-\frac{1}{\tau}\int_{I^k}\mathbf{E}(s)ds),\boldsymbol{\psi}_h)-(\frac{1}{\tau}\int_{I^k}\mathbf{K}(s)ds-\overline{\mathbf{K}}_h^k,\boldsymbol{\psi}_h),\\
(iii)\quad &\frac{1}{\epsilon_0\omega_{pe}^2}(\delta_\tau\tilde{\xi}_h^k,\tilde{\boldsymbol{\phi}}_h)+\frac{\Gamma_e}{\epsilon_0\omega_{pe}^2}(\overline{\tilde{\xi}}_h^k,\tilde{\boldsymbol{\phi}}_h)=\frac{1}{\epsilon_0\omega_{pe}^2}(\delta_\tau(\Pi_h\mathbf{J}^k-\mathbf{J}^k),\tilde{\boldsymbol{\phi}}_h)\\
&+\frac{\Gamma_e}{\epsilon_0\omega_{pe}^2}(\Pi_h\overline{\mathbf{J}}^k-\frac{1}{\tau}\int_{I^k}\mathbf{J}(s)ds,\tilde{\boldsymbol{\phi}}_h)+(\frac{1}{\tau}\int_{I^k}\mathbf{E}(s)ds-\overline{\mathbf{E}}_h^k,\tilde{\boldsymbol{\phi}}_h),\\
(iv)\quad &\frac{1}{\mu_0\omega_{pm}^2}(\delta_\tau\tilde{\eta}_h^k,\tilde{\boldsymbol{\psi}}_h)+\frac{\Gamma_m}{\mu_0\omega_{pm}^2}(\overline{\tilde{\eta}}_h^k,\tilde{\boldsymbol{\psi}}_h)=\frac{1}{\mu_0\omega_{pm}^2}(\delta_\tau(P_h\mathbf{K}^k-\mathbf{K}^k),\tilde{\boldsymbol{\psi}}_h)
\end{aligned}$$

$$+\frac{\Gamma_m}{\mu_0\omega_{pm}^2}(P_h\overline{\mathbf{K}}^k-\frac{1}{\tau}\int_{I^k}\mathbf{K}(s)ds,\tilde{\boldsymbol{\psi}}_h)+(\frac{1}{\tau}\int_{I^k}\mathbf{H}(s)ds-\overline{\mathbf{H}}_h^k,\tilde{\boldsymbol{\psi}}_h).$$

Choosing $\phi_h=\tau(\xi_h^k+\xi_h^{k-1})$, $\psi_h=\tau(\eta_h^k+\eta_h^{k-1})$, $\tilde{\phi_h}=\tau(\tilde{\xi}_h^k+\tilde{\xi}_h^{k-1})$, $\tilde{\psi_h}=\tau(\tilde{\eta}_h^k+\tilde{\eta}_h^{k-1})$ in the above error equations, adding the resultants together, and using the property of operator P_h, we obtain

$$\begin{aligned}
&\epsilon_0(||\xi_h^k||_0^2-||\xi_h^{k-1}||_0^2)+\mu_0(||\eta_h^k||_0^2-||\eta_h^{k-1}||_0^2)\\
&+\frac{1}{\epsilon_0\omega_{pe}^2}(||\tilde{\xi}_h^k||_0^2-||\tilde{\xi}_h^{k-1}||_0^2)+\frac{2\tau\Gamma_e}{\epsilon_0\omega_{pe}^2}||\overline{\tilde{\xi}}_h^{k+1}||_0^2\\
&+\frac{1}{\mu_0\omega_{pm}^2}(||\tilde{\eta}_h^k||_0^2-||\tilde{\eta}_h^{k-1}||_0^2)+\frac{2\tau\Gamma_m}{\mu_0\omega_{pm}^2}||\overline{\tilde{\eta}}_h^{k+1}||_0^2\\
=\;&2\tau\epsilon_0(\delta_\tau(\Pi_h\mathbf{E}^k-\mathbf{E}^k),\overline{\xi}_h^k)-2\tau(\overline{\mathbf{H}}^k-\frac{1}{\tau}\int_{I^k}\mathbf{H}(s)ds,\nabla\times\overline{\xi}_h^k)\\
&-2\tau(\frac{1}{\tau}\int_{I^k}\mathbf{J}(s)ds-\overline{\mathbf{J}}_h^k,\overline{\xi}_h^k)+2\tau(\nabla\times(\Pi_h\overline{\mathbf{E}}^k-\frac{1}{\tau}\int_{I^k}\mathbf{E}(s)ds),\overline{\eta}_h^k)\\
&-2\tau(\frac{1}{\tau}\int_{I^k}\mathbf{K}(s)ds-\overline{\mathbf{K}}_h^k,\overline{\eta}_h^k)+\frac{2\tau}{\epsilon_0\omega_{pe}^2}(\delta_\tau(\Pi_h\mathbf{J}^k-\mathbf{J}^k),\overline{\tilde{\xi}}_h^k)\\
&+\frac{2\tau\Gamma_e}{\epsilon_0\omega_{pe}^2}(\Pi_h\overline{\mathbf{J}}^k-\frac{1}{\tau}\int_{I^k}\mathbf{J}(s)ds,\overline{\tilde{\xi}}_h^k)+2\tau(\frac{1}{\tau}\int_{I^k}\mathbf{E}(s)ds-\overline{\mathbf{E}}_h^k,\overline{\tilde{\xi}}_h^k)\\
&+\frac{2\tau\Gamma_m}{\mu_0\omega_{pm}^2}(\overline{\mathbf{K}}^k-\frac{1}{\tau}\int_{I^k}\mathbf{K}(s)ds,\overline{\tilde{\eta}}_h^k)+2\tau(\frac{1}{\tau}\int_{I^k}\mathbf{H}(s)ds-\overline{\mathbf{H}}_h^k,\overline{\tilde{\eta}}_h^k)\\
=\;&\sum_{i=1}^{10}(Err)_i.
\end{aligned}\tag{3.94}$$

Since this is our first error analysis of numerical schemes for solving Maxwell's equations in metamaterials, below we provide detailed estimates of each $(Err)_i$ in (3.94).

Using the Cauchy-Schwarz inequality, the arithmetic-geometric mean inequality

$$(a,b)\le\delta||a||_0^2+\frac{1}{4\delta}||b||_0^2\quad\forall\,\delta>0,\tag{3.95}$$

Lemma 3.16, and the interpolation estimate (3.78), we have

$$\begin{aligned}
Err_1&\le\tau\epsilon_0(2\delta_1||\overline{\xi}_h^k||_0^2+\frac{1}{2\delta_1}||\delta_\tau(\Pi_h\mathbf{E}^k-\mathbf{E}^k)||_0^2)\\
&\le\tau\epsilon_0[\delta_1(||\xi_h^k||_0^2+||\xi_h^{k-1}||_0^2)+\frac{1}{2\delta_1\tau}\int_{I^k}||\partial_t(\Pi_h\mathbf{E}^k-\mathbf{E}^k)||_0^2dt]
\end{aligned}$$

$$\leq \tau\epsilon_0[\delta_1(||\xi_h^k||_0^2 + ||\xi_h^{k-1}||_0^2) + \frac{1}{2\delta_1\tau}\int_{I^k} Ch^{2l}||\mathbf{E}_t||_{l,\mathrm{curl}}^2 dt].$$

Similarly, we can obtain

$$Err_2 \leq \tau[2\delta_2||\overline{\xi}_h^k||_0^2 + \frac{1}{2\delta_2}||\nabla \times (\overline{\mathbf{H}}^k - \frac{1}{\tau}\int_{I^k}\mathbf{H}(s)ds||_0^2$$
$$\leq \tau[\delta_2(||\xi_h^k||_0^2 + ||\xi_h^{k-1}||_0^2) + \frac{\tau^3}{8\delta_2}\int_{I^k}||\nabla \times \mathbf{H}_{tt}(s)||_0^2 ds].$$

$$Err_3 = -2\tau(\frac{1}{\tau}\int_{I^k}\mathbf{J}(s)ds - \overline{\mathbf{J}}^k + \overline{\mathbf{J}}^k - \Pi_h\overline{\mathbf{J}}^k + \overline{\tilde{\xi}}_h^k, \overline{\xi}_h^k)$$
$$\leq -2\tau(\overline{\tilde{\xi}}_h^k, \overline{\xi}_h^k) + \tau[2\delta_3||\overline{\xi}_h^k||_0^2 + \frac{1}{\delta_3}(||\frac{1}{\tau}\int_{I^k}\mathbf{J}(s)ds - \overline{\mathbf{J}}^k||_0^2 + ||\overline{\mathbf{J}}^k - \Pi_h\overline{\mathbf{J}}^k||_0^2)]$$
$$\leq -2\tau(\overline{\tilde{\xi}}_h^k, \overline{\xi}_h^k) + \tau[\delta_3(||\xi_h^k||_0^2 + ||\xi_h^{k-1}||_0^2)$$
$$+ \frac{\tau^3}{4\delta_3}\int_{I^k}||\mathbf{J}_{tt}(s)||_0^2 ds + Ch^{2l}||\mathbf{J}||^2_{L^\infty(0,T;H^l(\mathrm{curl};\Omega))}].$$

By the same arguments, we have

$$Err_4 = 2\tau(\nabla \times (\Pi_h\overline{\mathbf{E}}^k - \overline{\mathbf{E}}^k) + \nabla \times (\overline{\mathbf{E}}^k - \frac{1}{\tau}\int_{I^k}\mathbf{E}(s)ds), \overline{\eta}_h^k)$$
$$\leq \tau[2\delta_4||\overline{\eta}_h^k||_0^2 + \frac{1}{\delta_4}(||\nabla \times (\Pi_h\overline{\mathbf{E}}^k - \overline{\mathbf{E}}^k)||_0^2 + ||\nabla \times (\overline{\mathbf{E}}^k - \frac{1}{\tau}\int_{I^k}\mathbf{E}(s)ds)||_0^2)]$$
$$\leq \tau\delta_4(||\eta_h^k||_0^2 + ||\eta_h^{k-1}||_0^2) + \frac{\tau}{\delta_4}(Ch^{2l}||\mathbf{E}||^2_{L^\infty(0,T;H^l(\mathrm{curl};\Omega))} + \frac{\tau^3}{4}\int_{I^k}||\nabla \times \mathbf{E}_{tt}(s)||_0^2 ds),$$

and

$$Err_5 = -2\tau(\frac{1}{\tau}\int_{I^k}\mathbf{K}(s)ds - \overline{\mathbf{K}}^k + \overline{\mathbf{K}}^k - P_h\overline{\mathbf{K}}^k + \overline{\tilde{\eta}}_h^k, \overline{\eta}_h^k)$$
$$\leq -2\tau(\overline{\tilde{\eta}}_h^k, \overline{\eta}_h^k) + \tau\delta_5(||\eta_h^k||_0^2 + ||\eta_h^{k-1}||_0^2) + \frac{\tau^4}{8\delta_5}\int_{I^k}||\mathbf{K}_{tt}||_0^2 ds.$$

Similar to Err_1, we have

$$Err_6 = \frac{2\tau}{\epsilon_0\omega_{pe}^2}(\delta_\tau(\Pi_h\mathbf{J}^k - \mathbf{J}^k), \overline{\tilde{\xi}}_h^k)$$
$$\leq \frac{\tau}{\epsilon_0\omega_{pe}^2}[\delta_6(||\tilde{\xi}_h^k||_0^2 + ||\tilde{\xi}_h^{k-1}||_0^2) + \frac{1}{2\delta_6\tau}\int_{I^k} Ch^{2l}||\mathbf{J}_t||_{l,\mathrm{curl}}^2 dt].$$

Furthermore, we have

$$
\begin{aligned}
Err_7 &= \frac{2\tau\Gamma_e}{\epsilon_0\omega_{pe}^2}(\Pi_h\overline{\mathbf{J}}^k - \overline{\mathbf{J}}^k + \overline{\mathbf{J}}^k - \frac{1}{\tau}\int_{I^k}\mathbf{J}(s)ds, \overline{\tilde{\xi}}_h^k) \\
&\leq \frac{\tau\Gamma_e}{\epsilon_0\omega_{pe}^2}[\delta_7(||\tilde{\xi}_h^k||_0^2 + ||\tilde{\xi}_h^{k-1}||_0^2) + \frac{1}{\delta_7}(\frac{\tau^3}{4}\int_{I^k}||\mathbf{J}_{tt}||_0^2 ds + Ch^{2l}||\mathbf{J}||^2_{L^\infty(0,T;H^l(\mathrm{curl};\Omega))})],
\end{aligned}
$$

$$
\begin{aligned}
Err_8 &= 2\tau(\frac{1}{\tau}\int_{I^k}\mathbf{E}(s)ds - \overline{\mathbf{E}}^k + \overline{\mathbf{E}}^k - \Pi_h\overline{\mathbf{E}}^k + \overline{\xi}_h^k, \overline{\tilde{\xi}}_h^k) \\
&\leq 2\tau(\overline{\xi}_h^k, \overline{\tilde{\xi}}_h^k) + \tau[2\delta_8||\overline{\tilde{\xi}}_h^k||_0^2 + \frac{1}{\delta_8}(||\frac{1}{\tau}\int_{I^k}\mathbf{E}(s)ds - \overline{\mathbf{E}}^k||_0^2 + ||\overline{\mathbf{E}}^k - \Pi_h\overline{\mathbf{E}}^k||_0^2)] \\
&\leq 2\tau(\overline{\xi}_h^k, \overline{\tilde{\xi}}_h^k) + \tau\delta_8(||\tilde{\xi}_h^k||_0^2 + ||\tilde{\xi}_h^{k-1}||_0^2) \\
&\quad + \frac{\tau}{\delta_8}(\frac{\tau^3}{4}\int_{I^k}||\mathbf{E}_{tt}||_0^2 ds + Ch^{2l}||\mathbf{E}||^2_{L^\infty(0,T;H^l(\mathrm{curl};\Omega))}), \\
Err_9 &\leq \frac{\tau\Gamma_m}{\mu_0\omega_{pm}^2}[\delta_9(||\tilde{\eta}_h^k||_0^2 + ||\tilde{\eta}_h^{k-1}||_0^2) + \frac{\tau^3}{8\delta_9}\int_{I^k}||\mathbf{K}_{tt}||_0^2 ds],
\end{aligned}
$$

and

$$
\begin{aligned}
Err_{10} &= 2\tau(\frac{1}{\tau}\int_{I^k}\mathbf{H}(s)ds - \overline{\mathbf{H}}^k + \overline{\mathbf{H}}^k - P_h\overline{\mathbf{H}}^k + \overline{\eta}_h^k, \overline{\tilde{\eta}}_h^k) \\
&\leq 2\tau(\overline{\eta}_h^k, \overline{\tilde{\eta}}_h^k) + \tau[\delta_{10}(||\tilde{\eta}_h^k||_0^2 + ||\tilde{\eta}_h^{k-1}||_0^2) + \frac{\tau^3}{8\delta_{10}}\int_{I^k}||\mathbf{H}_{tt}||_0^2 ds].
\end{aligned}
$$

Substituting the estimates of Err_i into (3.94), and summing up the results from $k = 1$ to n ($n \leq M - 1$), and using the facts $n\tau \leq T$ and $\xi_h^0 = \eta_h^0 = \tilde{\xi}_h^0 = \tilde{\eta}_h^0 = 0$, we can obtain (details see [191])

$$
\begin{aligned}
&\epsilon_0||\xi_h^n||_0^2 + \mu_0||\eta_h^n||_0^2 + \frac{1}{\epsilon_0\omega_{pe}^2}||\tilde{\xi}_h^n||_0^2 + \frac{1}{\mu_0\omega_{pm}^2}||\tilde{\eta}_h^n||_0^2 \\
&\leq C\tau\sum_{k=1}^{n-1}(||\xi_h^k||_0^2 + ||\eta_h^k||_0^2 + ||\tilde{\xi}_h^k||_0^2 + ||\tilde{\eta}_h^k||_0^2) + C(h^{2l} + \tau^4),
\end{aligned}
$$

which, along with the discrete Gronwall inequality, the triangle inequality, the estimates (3.78) and (3.79), completes the proof. □

3.5 The Leap-Frog Scheme for the Drude Model

3.5.1 The Leap-Frog Scheme

The Crank-Nicolson scheme discussed in last section is implicit, hence we have to solve a linear system at each time step, which is quite computationally intensive. Using a similar idea to the famous Yee scheme [299], we can construct an explicit leap-frog finite element scheme [183]: Given initial approximations $\mathbf{E}_h^0, \mathbf{K}_h^0, \mathbf{H}_h^{\frac{1}{2}}, \mathbf{J}_h^{\frac{1}{2}}$, for $k = 1, 2, \cdots$, find $\mathbf{E}_h^k \in \mathbf{V}_h^0, \mathbf{J}_h^{k+\frac{1}{2}} \in \mathbf{V}_h, \mathbf{H}_h^{k+\frac{1}{2}}, \mathbf{K}_h^k \in \mathbf{U}_h$ such that

$$\epsilon_0(\frac{\mathbf{E}_h^k - \mathbf{E}_h^{k-1}}{\tau}, \boldsymbol{\phi}_h) - (\mathbf{H}_h^{k-\frac{1}{2}}, \nabla \times \boldsymbol{\phi}_h) + (\mathbf{J}_h^{k-\frac{1}{2}}, \boldsymbol{\phi}_h) = 0, \tag{3.96}$$

$$\mu_0(\frac{\mathbf{H}_h^{k+\frac{1}{2}} - \mathbf{H}_h^{k-\frac{1}{2}}}{\tau}, \boldsymbol{\psi}_h) + (\nabla \times \mathbf{E}_h^k, \boldsymbol{\psi}_h) + (\mathbf{K}_h^k, \boldsymbol{\psi}_h) = 0, \tag{3.97}$$

$$\frac{1}{\epsilon_0\omega_{pe}^2}(\frac{\mathbf{J}_h^{k+\frac{1}{2}} - \mathbf{J}_h^{k-\frac{1}{2}}}{\tau}, \tilde{\boldsymbol{\phi}}_h) + \frac{\Gamma_e}{\epsilon_0\omega_{pe}^2}(\frac{\mathbf{J}_h^{k+\frac{1}{2}} + \mathbf{J}_h^{k-\frac{1}{2}}}{2}, \tilde{\boldsymbol{\phi}}_h) = (\mathbf{E}_h^k, \tilde{\boldsymbol{\phi}}_h), \tag{3.98}$$

$$\frac{1}{\mu_0\omega_{pm}^2}(\frac{\mathbf{K}_h^k - \mathbf{K}_h^{k-1}}{\tau}, \tilde{\boldsymbol{\psi}}_h) + \frac{\Gamma_m}{\mu_0\omega_{pm}^2}(\frac{\mathbf{K}_h^k + \mathbf{K}_h^{k-1}}{2}, \tilde{\boldsymbol{\psi}}_h) = (\mathbf{H}_h^{k-\frac{1}{2}}, \tilde{\boldsymbol{\psi}}_h), \tag{3.99}$$

for any $\boldsymbol{\phi}_h \in \mathbf{V}_h^0$, $\boldsymbol{\psi}_h \in \mathbf{U}_h$, $\tilde{\boldsymbol{\phi}}_h \in \mathbf{V}_h$, $\tilde{\boldsymbol{\psi}}_h \in \mathbf{U}_h$.

Note that (3.98) and (3.99) can be simplified to

$$\mathbf{J}_h^{k+\frac{1}{2}} = \frac{2\epsilon_0\omega_{pe}^2}{2\tau^{-1} + \Gamma_e}\mathbf{E}_h^k + \frac{2\tau^{-1} - \Gamma_e}{2\tau^{-1} + \Gamma_e}\mathbf{J}_h^{k-\frac{1}{2}}, \tag{3.100}$$

$$\mathbf{K}_h^k = \frac{2\mu_0\omega_{pm}^2}{2\tau^{-1} + \Gamma_m}\mathbf{H}_h^{k-\frac{1}{2}} + \frac{2\tau^{-1} - \Gamma_m}{2\tau^{-1} + \Gamma_m}\mathbf{K}_h^{k-1}. \tag{3.101}$$

respectively.

In practice, the above leap-frog scheme can be implemented as follows: at each time step, we first solve (3.96) for $\mathbf{E}_h^k$ and update $\mathbf{K}_h^k$ using (3.101) in parallel, then solve (3.97) for $\mathbf{H}_h^{k+\frac{1}{2}}$ and update $\mathbf{J}_h^{k+\frac{1}{2}}$ by (3.100) in parallel. Compared to the Crank-Nicolson scheme presented in the last section, the leap-frog scheme is more efficient for solving large-scale problems, since no large global coefficient matrix has to be stored and inversed. Of course, we still have to inverse two mass matrices: one for (3.96) and one for (3.98). For the lowest-order cubic (or rectangular) edge element, we can even use mass-lumping technique [217, p. 352] for the mass matrix in (3.96) to speed up the computation, in which case, the mass matrix becomes a diagonal matrix. Of course, being an explicit scheme, the leap-frog scheme has a time step constraint as we will show in next section.

3.5.2 The Stability Analysis

In this section, we shall prove that the leap-frog scheme (3.98) and (3.99) is conditionally stable and has a discrete stability similar to the continuous stability obtained in Lemma 3.12.

Lemma 3.17. *Denote* $\gamma_e = |\frac{2\tau^{-1}-\Gamma_e}{2\tau^{-1}+\Gamma_e}|$, $\gamma_m = |\frac{2\tau^{-1}-\Gamma_m}{2\tau^{-1}+\Gamma_m}|$. *For the recursively defined* $\mathbf{J}_h^{k+\frac{1}{2}}$ *and* $\mathbf{K}_h^k$, *we have*

$$(i)\ \ ||\mathbf{J}_h^{k+\frac{1}{2}}||_0^2 \le 2[\epsilon_0^2\omega_{pe}^4\tau T\sum_{l=1}^{k}||\mathbf{E}_h^l||_0^2 + \gamma_e^{2k}||\mathbf{J}_h^{\frac{1}{2}}||_0^2], \tag{3.102}$$

$$(ii)\ \ ||\mathbf{K}_h^k||_0^2 \le 2[\mu_0^2\omega_{pm}^4\tau T\sum_{l=0}^{k-1}||\mathbf{H}_h^{l+\frac{1}{2}}||_0^2 + \gamma_m^{2k}||\mathbf{K}_h^0||_0^2]. \tag{3.103}$$

Proof. From (3.100) and the triangle inequality, we have

$$\begin{aligned}
||\mathbf{J}_h^{k+\frac{1}{2}}||_0 &\le \epsilon_0\omega_{pe}^2\tau||\mathbf{E}_h^k||_0 + \gamma_e||\mathbf{J}_h^{k-\frac{1}{2}}||_0 \\
&\le \epsilon_0\omega_{pe}^2\tau||\mathbf{E}_h^k||_0 + \gamma_e(\epsilon_0\omega_{pe}^2\tau||\mathbf{E}_h^{k-1}||_0 + \gamma_e||\mathbf{J}_h^{k-\frac{3}{2}}||_0) \\
&\le \cdots \\
&\le \epsilon_0\omega_{pe}^2\tau(||\mathbf{E}_h^k||_0 + \gamma_e||\mathbf{E}_h^{k-1}||_0 + \cdots + \gamma_e^{k-1}||\mathbf{E}_h^1||_0) + \gamma_e^k||\mathbf{J}_h^{\frac{1}{2}}||_0 \\
&\le \epsilon_0\omega_{pe}^2\tau\sum_{l=1}^{k}||\mathbf{E}_h^l||_0 + \gamma_e^k||\mathbf{J}_h^{\frac{1}{2}}||_0,
\end{aligned} \tag{3.104}$$

where we used the fact $\gamma_e < 1$ in the last step.

Squaring both sides of (3.104), we further have

$$\begin{aligned}
||\mathbf{J}_h^{k+\frac{1}{2}}||_0^2 &\le 2[\epsilon_0^2\omega_{pe}^4\tau^2(\sum_{l=1}^{k}||\mathbf{E}_h^l||_0)^2 + \gamma_e^{2k}||\mathbf{J}_h^{\frac{1}{2}}||_0^2] \\
&\le 2[\epsilon_0^2\omega_{pe}^4\tau^2(\sum_{l=1}^{k}1^2)(\sum_{l=1}^{k}||\mathbf{E}_h^l||_0^2) + \gamma_e^{2k}||\mathbf{J}_h^{\frac{1}{2}}||_0^2] \\
&\le 2[\epsilon_0^2\omega_{pe}^4\tau T\sum_{l=1}^{k}||\mathbf{E}_h^l||_0^2 + \gamma_e^{2k}||\mathbf{J}_h^{\frac{1}{2}}||_0^2],
\end{aligned} \tag{3.105}$$

where we used the fact $k\tau \le T$ in the last step.

Similarly, from (3.101), we have

$$||\mathbf{K}_h^k||_0 \le \mu_0\omega_{pm}^2\tau||\mathbf{H}_h^{k-\frac{1}{2}}||_0 + \gamma_m||\mathbf{K}_h^{k-1}||_0,$$

using which and repeating the above procedure, we can obtain

$$||\mathbf{K}_h^k||_0^2 \le 2[\mu_0^2\omega_{pm}^4\tau T\sum_{l=0}^{k-1}||\mathbf{H}_h^{l+\frac{1}{2}}||_0^2 + \gamma_m^{2k}||\mathbf{K}_h^0||_0^2], \tag{3.106}$$

which completes the proof. □

Theorem 3.11. *Let $C_v = 1/\sqrt{\mu_0\epsilon_0}$ denote the speed of light in vacuum, and C_{inv} denote the constant from the standard inverse estimate*

$$||\nabla\times\psi_h||_0 \le C_{inv}h^{-1}||\psi_h||_0, \quad \psi_h\in\mathbf{V}_h. \tag{3.107}$$

Then under the assumption that the time step

$$\tau = \min(\frac{h}{2C_{inv}C_v}, 1), \tag{3.108}$$

the solutions of (3.96)–(3.99) satisfy the following:

$$\epsilon_0||\mathbf{E}_h^n||_0^2 + \mu_0||\mathbf{H}_h^{n+\frac{1}{2}}||_0^2 + \frac{1}{\mu_0\omega_{pm}^2}||\mathbf{K}_h^n||_0^2 + \frac{1}{\epsilon_0\omega_{pe}^2}||\mathbf{J}_h^{n+\frac{1}{2}}||_0^2$$
$$\le C\left(||\mathbf{E}_h^0||_0^2 + ||\mathbf{H}_h^{\frac{1}{2}}||_0^2 + ||\mathbf{J}_h^{\frac{1}{2}}||_0^2 + ||\mathbf{K}_h^0||_0^2\right), \quad \forall\, n\ge 1,$$

where the constant $C > 1$ is independent of h and τ.

Proof. Choosing $\boldsymbol{\phi}_h = \tau(\mathbf{E}_h^k + \mathbf{E}_h^{k-1})$ in (3.96), $\boldsymbol{\psi}_h = \tau(\mathbf{H}_h^{k+\frac{1}{2}} + \mathbf{H}_h^{k-\frac{1}{2}})$ in (3.97), adding (3.96) and (3.97) together, then using the following identity

$$-(\mathbf{H}_h^{k-\frac{1}{2}}, \nabla\times(\mathbf{E}_h^k + \mathbf{E}_h^{k-1})) + (\nabla\times\mathbf{E}_h^k, \mathbf{H}_h^{k+\frac{1}{2}} + \mathbf{H}_h^{k-\frac{1}{2}})$$
$$= -(\mathbf{H}_h^{k-\frac{1}{2}}, \nabla\times\mathbf{E}_h^{k-1}) + (\nabla\times\mathbf{E}_h^k, \mathbf{H}_h^{k+\frac{1}{2}}), \tag{3.109}$$

and summing the resultants from $k=1$ to $k=n$, we obtain

$$\begin{aligned}
&\epsilon_0(||\mathbf{E}_h^n||_0^2 - ||\mathbf{E}_h^0||_0^2) + \mu_0(||\mathbf{H}_h^{n+\frac{1}{2}}||_0^2 - ||\mathbf{H}_h^{\frac{1}{2}}||_0^2)\\
=\ &\tau[(\mathbf{H}_h^{\frac{1}{2}}, \nabla\times\mathbf{E}_h^0) - (\nabla\times\mathbf{E}_h^n, \mathbf{H}_h^{n+\frac{1}{2}})]\\
&-\tau\sum_{k=1}^n(\mathbf{J}_h^{k-\frac{1}{2}}, \mathbf{E}_h^k + \mathbf{E}_h^{k-1}) - \tau\sum_{k=1}^n(\mathbf{K}_h^k, \mathbf{H}_h^{k+\frac{1}{2}} + \mathbf{H}_h^{k-\frac{1}{2}}).
\end{aligned} \tag{3.110}$$

By the Cauchy-Schwartz inequality and the inverse estimate (3.107), we have

$$\begin{aligned}\tau(\nabla\times\mathbf{E}_h^n,\mathbf{H}_h^{n+\frac{1}{2}}) &\leq \tau\cdot C_{inv}h^{-1}||\mathbf{E}_h^n||_0||\mathbf{H}_h^{n+\frac{1}{2}}||_0\\ &= \tau\cdot C_{inv}h^{-1}\cdot C_v\sqrt{\epsilon_0}||\mathbf{E}_h^n||_0\cdot\sqrt{\mu_0}||\mathbf{H}_h^{n+\frac{1}{2}}||_0\\ &\leq \delta_1\epsilon_0||\mathbf{E}_h^n||_0^2+\frac{1}{4\delta_1}(\frac{C_{inv}C_v\tau}{h})^2\mu_0||\mathbf{H}_h^{n+\frac{1}{2}}||_0^2, \quad (3.111)\end{aligned}$$

and

$$\begin{aligned}\tau\sum_{k=1}^{n}(\mathbf{J}_h^{k-\frac{1}{2}},\mathbf{E}_h^k+\mathbf{E}_h^{k-1}) &\leq \tau\sum_{k=1}^{n}||\mathbf{J}_h^{k-\frac{1}{2}}||_0(||\mathbf{E}_h^k||_0+||\mathbf{E}_h^{k-1}||_0)\\ &\leq \tau\sum_{k=1}^{n}[\delta_2||\mathbf{E}_h^k||_0^2+\delta_3||\mathbf{E}_h^{k-1}||_0^2+(\frac{1}{4\delta_2}+\frac{1}{4\delta_3})||\mathbf{J}_h^{k-\frac{1}{2}}||_0^2].\end{aligned}$$

Furthermore, from (3.105) and the fact that $\gamma_e < 1$, we have

$$\begin{aligned}\tau\sum_{k=1}^{n}||\mathbf{J}_h^{k-\frac{1}{2}}||_0^2 &\leq 2\tau\sum_{k=1}^{n}[\epsilon_0^2\omega_{pe}^4\tau T\sum_{l=1}^{k-1}||\mathbf{E}_h^l||_0^2+\gamma_e^{2(k-1)}||\mathbf{J}_h^{\frac{1}{2}}||_0^2]\\ &\leq 2[\epsilon_0^2\omega_{pe}^4\tau T^2\sum_{l=1}^{n-1}||\mathbf{E}_h^l||_0^2+T||\mathbf{J}_h^{\frac{1}{2}}||_0^2].\end{aligned}$$

Similarly, we have

$$\begin{aligned}&\tau\sum_{k=1}^{n}(\mathbf{K}_h^k,\mathbf{H}_h^{k+\frac{1}{2}}+\mathbf{H}_h^{k-\frac{1}{2}})\\ &\leq \tau\sum_{k=1}^{n}[\delta_4||\mathbf{H}_h^{k+\frac{1}{2}}||_0^2+\delta_5||\mathbf{H}_h^{k-\frac{1}{2}}||_0^2+(\frac{1}{4\delta_4}+\frac{1}{4\delta_5})||\mathbf{K}_h^k||_0^2],\end{aligned}$$

and

$$\begin{aligned}\tau\sum_{k=1}^{n}||\mathbf{K}_h^k||_0^2 &\leq 2\tau\sum_{k=1}^{n}[\mu_0^2\omega_{pm}^4\tau T\sum_{l=0}^{k-1}||\mathbf{H}_h^{l+\frac{1}{2}}||_0^2+\gamma_m^{2k}||\mathbf{K}_h^0||_0^2]\\ &\leq 2[\mu_0^2\omega_{pm}^4\tau T^2\sum_{l=0}^{n-1}||\mathbf{H}_h^{l+\frac{1}{2}}||_0^2+T||\mathbf{K}_h^0||_0^2].\end{aligned}$$

Substituting the above estimates and the following (let $n=0$ in (3.111))

$$\tau(\mathbf{H}_h^{\frac{1}{2}}, \nabla\times\mathbf{E}_h^0) \le \delta_1\epsilon_0||\mathbf{E}_h^0||_0^2 + \frac{1}{4\delta_1}(\frac{C_{inv}C_v\tau}{h})^2\mu_0||\mathbf{H}_h^{\frac{1}{2}}||_0^2$$

into (3.110), we obtain

$$\begin{aligned}
&\epsilon_0(||\mathbf{E}_h^n||_0^2 - ||\mathbf{E}_h^0||_0^2) + \mu_0(||\mathbf{H}_h^{n+\frac{1}{2}}||_0^2 - ||\mathbf{H}_h^{\frac{1}{2}}||_0^2)\\
\le\ &\frac{\tau}{2}||\nabla\times\mathbf{H}_h^{\frac{1}{2}}||_0^2 + (\frac{\tau}{2}+\tau\delta_3)||\mathbf{E}_h^0||_0^2\\
&+(\delta_1+\frac{\tau\delta_2}{\epsilon_0})\epsilon_0||\mathbf{E}_h^n||_0^2 + [\frac{1}{4\delta_1}(\frac{C_{inv}C_v\tau}{h})^2 + \frac{\tau\delta_4}{\mu_0}]\mu_0||\mathbf{H}_h^{n+\frac{1}{2}}||_0^2\\
&+\tau[(\frac{1}{2\delta_2}+\frac{1}{2\delta_3})\epsilon_0^2\omega_{pe}^4T^2+\delta_2+\delta_3]\sum_{l=1}^{n-1}||\mathbf{E}_h^l||_0^2 + (\frac{1}{2\delta_2}+\frac{1}{2\delta_3})T||\mathbf{J}_h^{\frac{1}{2}}||_0^2\\
&+\tau[(\frac{1}{2\delta_4}+\frac{1}{2\delta_5})\mu_0^2\omega_{pm}^4T^2+\delta_4+\delta_5]\sum_{l=0}^{n-1}||\mathbf{H}_h^{l+\frac{1}{2}}||_0^2 + (\frac{1}{2\delta_4}+\frac{1}{2\delta_5})T||\mathbf{K}_h^0||_0^2.
\end{aligned}$$

By choosing δ_i small enough and $\tau = O(h)$ such that $||\mathbf{E}_h^n||_0^2$ and $||\mathbf{H}_h^{n+\frac{1}{2}}||_0^2$ can be controlled by the left-hand side (e.g., $\delta_1=\frac{1}{4}$, $\delta_2=\frac{1}{4}\epsilon_0$, $\delta_4=\frac{1}{4}\mu_0$, $\tau=\min(\frac{h}{2C_{inv}C_v},1)$, $\delta_3=\epsilon_0, \delta_5=\mu_0$), and using the discrete Gronwall inequality, we have

$$\epsilon_0||\mathbf{E}_h^n||_0^2 + \mu_0||\mathbf{H}_h^{n+\frac{1}{2}}||_0^2 \le C[||\mathbf{E}_h^0||_0^2 + ||\mathbf{H}_h^{\frac{1}{2}}||_0^2 + ||\mathbf{J}_h^{\frac{1}{2}}||_0^2 + ||\mathbf{K}_h^0||_0^2],$$

which, along with (3.105) and (3.106), concludes the proof. □

Remark 3.1. Note that when h is small enough, the time step constraint (3.108) reduces to the standard CFL condition $\tau = O(h)$, which is often imposed on explicit schemes used to solve the first-order hyperbolic systems.

A tight and accurate estimate of the constant C_{inv} in (3.107) is quite challenging, since it depends on the element shape and the order of the basis function. Below we just show a tight estimate of C_{inv} for the lowest rectangular edge element.

Lemma 3.18. *Consider a domain Ω is triangulated by a mesh T_h formed by m rectangles $K_i=[x_c^i-h_x, x_c^i+h_x]\times[y_c^i-h_y, y_c^i+h_y]$, $i=1,\cdots,m$. Let $h=\max\{h_x,h_y\}$. Then we have*

$$C_{inv} \ge \max\{\frac{\sqrt{3}}{2}\frac{h}{h_x}, \frac{\sqrt{3}}{2}\frac{h}{h_y}\}. \tag{3.112}$$

Proof. Recall that the lowest edge element basis functions are (cf. Example 3.6):

$$N_1^i(x,y) = \begin{pmatrix} \frac{(y_c^i+h_y)-y}{4h_xh_y} \\ 0 \end{pmatrix}, \quad N_2^i(x,y) = \begin{pmatrix} 0 \\ \frac{x-(x_c^i-h_x)}{4h_xh_y} \end{pmatrix},$$
$$N_3^i(x,y) = \begin{pmatrix} \frac{(y_c^i-h_y)-y}{4h_xh_y} \\ 0 \end{pmatrix}, \quad N_4^i(x,y) = \begin{pmatrix} 0 \\ \frac{x-(x_c^i+h_x)}{4h_xh_y} \end{pmatrix},$$

where $N_j^i, j = 1,2,3,4$, start from the bottom edge and orient counterclockwisely.

It is easy to check that the 2-D curl of N_j^i satisfies

$$\int_{K_i} |\nabla \times N_j^i|^2 dxdy = \int_{K_i} |\frac{1}{4h_xh_y}|^2 dxdy = \frac{1}{4h_xh_y},$$

and N_j^i satisfies

$$\int_{K_i} |N_1^i|^2 dxdy = \int_{K_i} (\frac{y_c^i + h_y - y}{4h_xh_y})^2 dxdy$$
$$= \frac{2h_x}{(4h_xh_y)^2} \cdot \frac{-1}{3}(y_c^i + h_y - y)^3 |_{y=y_c^i-h_y}^{y_c^i+h_y} = \frac{h_y}{3h_x},$$
$$\int_{K_i} |N_3^i|^2 dxdy = \frac{h_y}{3h_x}, \quad \int_{K_i} |N_2^i|^2 dxdy = |N_4^i|^2 dxdy = \frac{h_x}{3h_y},$$

from which we can see that

$$\frac{||\nabla \times N_j^i||_{0,K_i}^2}{||N_j^i||_{0,K_i}^2} = \frac{3}{4h_y^2}, \quad j = 1,3. \tag{3.113}$$

Similarly, we have

$$\frac{||\nabla \times N_j^i||_{0,K_i}^2}{||N_j^i||_{0,K_i}^2} = \frac{3}{4h_x^2}, \quad j = 2,4,$$

applying which to (3.107) we complete the proof. □

By Lemma 3.18, we should try to use shape regular meshes and avoid anisotropic meshes in practice computation, since the anisotropic mesh may have a very large C_{inv} and lead to a very small time step according to (3.108).

3.5.3 The Optimal Error Estimate

To carry out the error analysis for the scheme (3.96)–(3.99), we need some preliminary estimates.

Lemma 3.19. *Denote $u^j = u(\cdot, j\tau)$. For any $u \in H^2(0,T;L^2(\Omega))$, we have*

$$(i)\ \ ||u^k - \frac{1}{\tau}\int_{t_{k-\frac{1}{2}}}^{t_{k+\frac{1}{2}}} u(s)ds||_0^2 \le \frac{\tau^3}{4}\int_{t_{k-\frac{1}{2}}}^{t_{k+\frac{1}{2}}} ||u_{tt}(s)||_0^2 ds,$$

$$(ii)\ \ ||u^{k-\frac{1}{2}} - \frac{1}{\tau}\int_{t_{k-1}}^{t_k} u(s)ds||_0^2 \le \frac{\tau^3}{4}\int_{t_{k-1}}^{t_k} ||u_{tt}(s)||_0^2 ds,$$

$$(iii)\ \ ||\frac{1}{2}(u^{k-1} + u^k) - \frac{1}{\tau}\int_{t_{k-1}}^{t_k} u(s)ds||_0^2 \le \frac{\tau^3}{4}\int_{t_{k-1}}^{t_k} ||u_{tt}(s)||_0^2 ds,$$

$$(iv)\ \ ||\frac{1}{2}(u^{k-\frac{1}{2}} + u^{k+\frac{1}{2}}) - \frac{1}{\tau}\int_{t_{k-\frac{1}{2}}}^{t_{k+\frac{1}{2}}} u(s)ds||_0^2 \le \frac{\tau^3}{4}\int_{t_{k-\frac{1}{2}}}^{t_{k+\frac{1}{2}}} ||u_{tt}(s)||_0^2 ds.$$

Furthermore, for any $u \in H^1(0,T;L^2(\Omega))$, we have

$$(v)\ \ ||\delta_\tau u^{k+\frac{1}{2}}||_0^2 = ||\frac{u^{k+\frac{1}{2}} - u^{k-\frac{1}{2}}}{\tau}||_0^2 \le \frac{1}{\tau}\int_{t_{k-\frac{1}{2}}}^{t_{k+\frac{1}{2}}} ||u_t(t)||_0^2 dt.$$

Proof. (i) Using the following integral identity

$$u(s) = u(t_k) + (s - t_k)u_t(t_k) + \int_s^{t_k} (r - s)u_{tt}(r)dr$$

we obtain

$$|u^k - \frac{1}{\tau}\int_{t_{k-\frac{1}{2}}}^{t_{k+\frac{1}{2}}} u(s)ds|^2 = |-\frac{1}{\tau}\int_{t_{k-\frac{1}{2}}}^{t_{k+\frac{1}{2}}} [\int_s^{t_k} (r-s)u_{tt}(r)dr]ds|^2$$

$$\le \frac{1}{\tau^2}(\int_{t_{k-\frac{1}{2}}}^{t_{k+\frac{1}{2}}} [\int_s^{t_k} (r-s)u_{tt}(r)dr]^2 ds)(\int_{t_{k-\frac{1}{2}}}^{t_{k+\frac{1}{2}}} 1^2 ds)$$

$$\le \frac{1}{\tau}\int_{t_{k-\frac{1}{2}}}^{t_{k+\frac{1}{2}}} (\int_s^{t_k} (r-s)^2 dr)(\int_s^{t_k} |u_{tt}(r)|^2 dr)ds \le \frac{1}{4}\tau^3 \int_{t_{k-\frac{1}{2}}}^{t_{k+\frac{1}{2}}} |u_{tt}(r)|^2 dr,$$

integrating which over Ω concludes the proof.

(ii) The proof is all the same as (i) except we use the following identity

$$u(s) = u(t_{k-\frac{1}{2}}) + (s - t_{k-\frac{1}{2}})u_t(t_{k-\frac{1}{2}}) + \int_{t_{k-\frac{1}{2}}}^{s} (s - r)u_{tt}(r)dr.$$

(iii) The proof is given in Lemma 3.16.

(iv) The proof is based on the following identity

$$\frac{1}{2}(u^{k-\frac{1}{2}}+u^{k+\frac{1}{2}})-\frac{1}{\tau}\int_{t_{k-\frac{1}{2}}}^{t_{k+\frac{1}{2}}}u(s)ds=\frac{1}{2\tau}\int_{t_{k-\frac{1}{2}}}^{t_{k+\frac{1}{2}}}(s-t_{k-\frac{1}{2}})(t_{k+\frac{1}{2}}-s)u_{ss}(s)ds.$$

(v) The proof is easily obtained by using $\frac{u^{k+\frac{1}{2}}-u^{k-\frac{1}{2}}}{\tau}=\int_{t_{k-\frac{1}{2}}}^{t_{k+\frac{1}{2}}}u_t(t)dt$.

□

Suppose that the solution is smooth enough, then we can prove the following optimal error estimate for the leap-frog scheme (3.96)–(3.99).

Theorem 3.12. *Let* $(\mathbf{E}^n,\mathbf{H}^n)$ *and* $(\mathbf{E}_h^n,\mathbf{H}_h^n)$ *be the analytic and finite element solutions at time* $t=t_n$*, respectively. Under the regularity assumptions*

$$\begin{aligned}
&\mathbf{H},\mathbf{K}\in L^\infty(0,T;(H^l(\Omega))^3),\\
&\mathbf{E},\mathbf{J},\mathbf{E}_t,\mathbf{J}_t\in L^\infty(0,T;H^l(curl;\Omega)),\\
&\mathbf{E}_{tt},\mathbf{H}_{tt},\mathbf{J}_{tt},\mathbf{K}_{tt},\nabla\times\mathbf{E}_{tt},\nabla\times\mathbf{H}_{tt}\in L^2(0,T;(L^2(\Omega))^3),
\end{aligned}$$

there exists a constant $C=C(T,\epsilon_0,\mu_0,\omega_{pe},\omega_{pm},\Gamma_e,\Gamma_m,\mathbf{E},\mathbf{H},\mathbf{J},\mathbf{K})$*, independent of both time step* τ *and mesh size* h*, such that*

$$\begin{aligned}
&\max_{1\le n}(||\mathbf{E}^n-\mathbf{E}_h^n||_0+||\mathbf{H}^{n+\frac{1}{2}}-\mathbf{H}_h^{n+\frac{1}{2}}||_0+||\mathbf{J}^{n+\frac{1}{2}}-\mathbf{J}_h^{n+\frac{1}{2}}||_0+||\mathbf{K}^n-\mathbf{K}_h^n||_0)\\
&\le C(\tau^2+h^l)+C\Big(||\mathbf{E}^0-\mathbf{E}_h^0||_0+||\mathbf{H}^{\frac{1}{2}}-\mathbf{H}_h^{\frac{1}{2}}||_0+||\mathbf{J}^{\frac{1}{2}}-\mathbf{J}_h^{\frac{1}{2}}||_0+||\mathbf{K}^0-\mathbf{K}_h^0||_0\Big).
\end{aligned}$$

where $l\ge 1$ *is the order of basis functions defined in spaces* $\mathbf{U}_h$ *and* $\mathbf{V}_h$*.*

Proof. Integrating the governing equations (3.55) and (3.58) from t_{k-1} to t_k, and (3.56) and (3.57) from $t_{k-\frac{1}{2}}$ to $t_{k+\frac{1}{2}}$, then multiplying the respective resultants by $\frac{\boldsymbol{\phi}_h}{\tau},\frac{\boldsymbol{\psi}_h}{\tau},\frac{\tilde{\boldsymbol{\phi}}_h}{\tau},\frac{\tilde{\boldsymbol{\psi}}_h}{\tau}$ and integrating over Ω, we have

$$\epsilon_0(\frac{\mathbf{E}^k-\mathbf{E}^{k-1}}{\tau},\boldsymbol{\phi}_h)-(\frac{1}{\tau}\int_{t_{k-1}}^{t_k}\mathbf{H}(s)ds,\nabla\times\boldsymbol{\phi}_h)+(\frac{1}{\tau}\int_{t_{k-1}}^{t_k}\mathbf{J}(s)ds,\boldsymbol{\phi}_h)=0, \tag{3.114}$$

$$\mu_0(\frac{\mathbf{H}^{k+\frac{1}{2}}-\mathbf{H}^{k-\frac{1}{2}}}{\tau},\boldsymbol{\psi}_h)+(\frac{1}{\tau}\int_{t_{k-\frac{1}{2}}}^{t_{k+\frac{1}{2}}}\nabla\times\mathbf{E}(s)ds,\boldsymbol{\psi}_h)+(\frac{1}{\tau}\int_{t_{k-\frac{1}{2}}}^{t_{k+\frac{1}{2}}}\mathbf{K}(s)ds,\boldsymbol{\psi}_h)=0, \tag{3.115}$$

$$\frac{1}{\epsilon_0\omega_{pe}^2}(\frac{\mathbf{J}^{k+\frac{1}{2}}-\mathbf{J}^{k-\frac{1}{2}}}{\tau},\tilde{\boldsymbol{\phi}}_h)+\frac{\Gamma_e}{\epsilon_0\omega_{pe}^2}(\frac{1}{\tau}\int_{t_{k-\frac{1}{2}}}^{t_{k+\frac{1}{2}}}\mathbf{J}(s)ds,\tilde{\boldsymbol{\phi}}_h)=(\frac{1}{\tau}\int_{t_{k-\frac{1}{2}}}^{t_{k+\frac{1}{2}}}\mathbf{E}(s)ds,\tilde{\boldsymbol{\phi}}_h), \tag{3.116}$$

$$\frac{1}{\mu_0\omega_{pm}^2}(\frac{\mathbf{K}^k-\mathbf{K}^{k-1}}{\tau},\tilde{\boldsymbol{\psi}}_h)+\frac{\Gamma_m}{\mu_0\omega_{pm}^2}(\frac{1}{\tau}\int_{t_{k-1}}^{t_k}\mathbf{K}(s)ds,\tilde{\boldsymbol{\psi}}_h)=(\frac{1}{\tau}\int_{t_{k-1}}^{t_k}\mathbf{H}(s)ds,\tilde{\boldsymbol{\psi}}_h). \tag{3.117}$$

Denote

$$\xi_h^k = \Pi_h \mathbf{E}^k - \mathbf{E}_h^k, \quad \tilde{\xi}_h^{k-\frac{1}{2}} = \Pi_h \mathbf{J}^{k-\frac{1}{2}} - \mathbf{J}_h^{k-\frac{1}{2}},$$
$$\eta_h^{k-\frac{1}{2}} = P_h \mathbf{H}^{k-\frac{1}{2}} - \mathbf{H}_h^{k-\frac{1}{2}}, \quad \tilde{\eta}_h^k = P_h \mathbf{K}^k - \mathbf{K}_h^k.$$

Subtracting (3.96)–(3.99) from (3.114)–(3.117), respectively, we obtain

$$\begin{aligned}
&\epsilon_0(\frac{\xi_h^k - \xi_h^{k-1}}{\tau}, \boldsymbol{\phi}_h) - (\eta_h^{k-\frac{1}{2}}, \nabla \times \boldsymbol{\phi}_h) \\
&= \epsilon_0(\delta_\tau(\Pi_h \mathbf{E}^k - \mathbf{E}^k), \boldsymbol{\phi}_h) - (P_h \mathbf{H}^{k-\frac{1}{2}} - \frac{1}{\tau}\int_{t_{k-1}}^{t_k} \mathbf{H}(s)ds, \nabla \times \boldsymbol{\phi}_h) \\
&\quad +(-\tilde{\xi}_h^{k-\frac{1}{2}} + \Pi_h \mathbf{J}^{k-\frac{1}{2}} - \frac{1}{\tau}\int_{t_{k-1}}^{t_k} \mathbf{J}(s)ds, \boldsymbol{\phi}_h),
\end{aligned} \tag{3.118}$$

$$\begin{aligned}
&\mu_0(\frac{\eta_h^{k+\frac{1}{2}} - \eta_h^{k-\frac{1}{2}}}{\tau}, \boldsymbol{\psi}_h) + (\nabla \times \xi_h^k, \boldsymbol{\psi}_h) \\
&= \mu_0(\delta_\tau(P_h \mathbf{H}^{k+\frac{1}{2}} - \mathbf{H}^{k+\frac{1}{2}}), \boldsymbol{\psi}_h) + (\nabla \times (\Pi_h \mathbf{E}^k - \frac{1}{\tau}\int_{t_{k-\frac{1}{2}}}^{t_{k+\frac{1}{2}}} \mathbf{E}(s)ds), \boldsymbol{\psi}_h) \\
&\quad +(-\tilde{\eta}_h^k + P_h \mathbf{K}^k - \frac{1}{\tau}\int_{t_{k-\frac{1}{2}}}^{t_{k+\frac{1}{2}}} \mathbf{K}(s)ds, \boldsymbol{\psi}_h),
\end{aligned} \tag{3.119}$$

$$\begin{aligned}
&\frac{1}{\epsilon_0\omega_{pe}^2}(\frac{\tilde{\xi}_h^{k+\frac{1}{2}} - \tilde{\xi}_h^{k-\frac{1}{2}}}{\tau}, \tilde{\boldsymbol{\phi}}_h) + \frac{\Gamma_e}{\epsilon_0\omega_{pe}^2}(\frac{1}{2}(\tilde{\xi}_h^{k+\frac{1}{2}} + \tilde{\xi}_h^{k-\frac{1}{2}}), \tilde{\boldsymbol{\phi}}_h) \\
&= \frac{1}{\epsilon_0\omega_{pe}^2}(\delta_\tau(\Pi_h \mathbf{J}^{k+\frac{1}{2}} - \mathbf{J}^{k+\frac{1}{2}}), \tilde{\boldsymbol{\phi}}_h) + \frac{\Gamma_e}{\epsilon_0\omega_{pe}^2}(\frac{1}{2}(\Pi_h \mathbf{J}^{k+\frac{1}{2}} + \Pi_h \mathbf{J}^{k-\frac{1}{2}}) \\
&\quad -\frac{1}{\tau}\int_{t_{k-\frac{1}{2}}}^{t_{k+\frac{1}{2}}} \mathbf{J}(s)ds, \tilde{\boldsymbol{\phi}}_h) + (\frac{1}{\tau}\int_{t_{k-\frac{1}{2}}}^{t_{k+\frac{1}{2}}} \mathbf{E}(s)ds - \Pi_h \mathbf{E}^k + \xi_h^k, \tilde{\boldsymbol{\phi}}_h),
\end{aligned} \tag{3.120}$$

and

$$\begin{aligned}
&\frac{1}{\mu_0\omega_{pm}^2}(\frac{\tilde{\eta}_h^k - \tilde{\eta}_h^{k-1}}{\tau}, \tilde{\boldsymbol{\psi}}_h) + \frac{\Gamma_m}{\mu_0\omega_{pm}^2}(\frac{1}{2}(\tilde{\eta}_h^k + \tilde{\eta}_h^{k-1}), \tilde{\boldsymbol{\psi}}_h) \\
&= \frac{1}{\mu_0\omega_{pm}^2}(\delta_\tau(P_h \mathbf{K}^k - \mathbf{K}^k), \tilde{\boldsymbol{\psi}}_h) + \frac{\Gamma_m}{\mu_0\omega_{pm}^2}(\frac{1}{2}(P_h \mathbf{K}^k + P_h \mathbf{K}^{k-1}) \\
&\quad -\frac{1}{\tau}\int_{t_{k-1}}^{t_k} \mathbf{K}(s)ds, \tilde{\boldsymbol{\psi}}_h) + (\frac{1}{\tau}\int_{t_{k-1}}^{t_k} \mathbf{H}(s)ds - P_h \mathbf{H}^{k-\frac{1}{2}} + \eta_h^{k-\frac{1}{2}}, \tilde{\boldsymbol{\psi}}_h).
\end{aligned} \tag{3.121}$$

Choosing

$$\boldsymbol{\phi}_h = \xi_h^k + \xi_h^{k-1}, \ \ \boldsymbol{\psi}_h = \eta_h^{k+\frac{1}{2}} + \eta_h^{k-\frac{1}{2}}, \ \ \tilde{\boldsymbol{\phi}}_h = \tilde{\xi}_h^{k+\frac{1}{2}} + \tilde{\xi}_h^{k-\frac{1}{2}}, \ \ \tilde{\boldsymbol{\psi}}_h = \tilde{\eta}_h^k + \tilde{\eta}_h^{k-1},$$

in (3.118)–(3.121), respectively, then using the following identities:

$$\begin{aligned} &(\nabla \times \xi_h^k, \eta_h^{k+\frac{1}{2}} + \eta_h^{k-\frac{1}{2}}) - (\eta_h^{k-\frac{1}{2}}, \nabla \times (\xi_h^k + \xi_h^{k-1})) \\ &= (\nabla \times \xi_h^k, \eta_h^{k+\frac{1}{2}}) - (\nabla \times \xi_h^{k-1}, \eta_h^{k-\frac{1}{2}}), \end{aligned}$$

and

$$\begin{aligned} &-(\tilde{\xi}_h^{k-\frac{1}{2}}, \xi_h^k + \xi_h^{k-1}) - (\tilde{\eta}_h^k, \eta_h^{k+\frac{1}{2}} + \eta_h^{k-\frac{1}{2}}) \\ &+(\xi_h^k, \tilde{\xi}_h^{k+\frac{1}{2}} + \tilde{\xi}_h^{k-\frac{1}{2}}) + (\eta_h^{k-\frac{1}{2}}, \tilde{\eta}_h^k + \tilde{\eta}_h^{k-1}) \\ = &-(\tilde{\xi}_h^{k-\frac{1}{2}}, \xi_h^{k-1}) + (\tilde{\xi}_h^{k+\frac{1}{2}}, \xi_h^k) - (\tilde{\eta}_h^k, \eta_h^{k+\frac{1}{2}}) + (\tilde{\eta}_h^{k-1}, \eta_h^{k-\frac{1}{2}}), \end{aligned}$$

and summing up the resultants from $k = 1$ to $k = n$, we obtain

$$\begin{aligned} &\epsilon_0(||\xi_h^n||_0^2 - ||\xi_h^0||_0^2) + \mu_0(||\eta_h^{n+\frac{1}{2}}||_0^2 - ||\eta_h^{\frac{1}{2}}||_0^2) \\ &\quad + \frac{1}{\epsilon_0\omega_{pe}^2}(||\tilde{\xi}_h^{n+\frac{1}{2}}||_0^2 - ||\tilde{\xi}_h^{\frac{1}{2}}||_0^2) + \frac{1}{\mu_0\omega_{pm}^2}(||\tilde{\eta}_h^n||_0^2 - ||\tilde{\eta}_h^0||_0^2) \\ \le\ &\tau\epsilon_0 \sum_{k=1}^n (\delta_\tau(\Pi_h \mathbf{E}^k - \mathbf{E}^k), \xi_h^k + \xi_h^{k-1}) \\ &-\tau \sum_{k=1}^n (P_h \mathbf{H}^{k-\frac{1}{2}} - \frac{1}{\tau}\int_{t_{k-1}}^{t_k} \mathbf{H}(s)ds, \nabla \times (\xi_h^k + \xi_h^{k-1})) \\ &+\tau \sum_{k=1}^n (\Pi_h \mathbf{J}^{k-\frac{1}{2}} - \frac{1}{\tau}\int_{t_{k-1}}^{t_k} \mathbf{J}(s)ds, \xi_h^k + \xi_h^{k-1}) \\ &+\tau\mu_0 \sum_{k=1}^n (\delta_\tau(P_h \mathbf{H}^{k+\frac{1}{2}} - \mathbf{H}^{k+\frac{1}{2}}), \eta_h^{k+\frac{1}{2}} + \eta_h^{k-\frac{1}{2}}) \\ &+\tau \sum_{k=1}^n (\nabla \times (\Pi_h \mathbf{E}^k - \frac{1}{\tau}\int_{t_{k-\frac{1}{2}}}^{t_{k+\frac{1}{2}}} \mathbf{E}(s)ds), \eta_h^{k+\frac{1}{2}} + \eta_h^{k-\frac{1}{2}}) \\ &+\tau \sum_{k=1}^n (P_h \mathbf{K}^k - \frac{1}{\tau}\int_{t_{k-\frac{1}{2}}}^{t_{k+\frac{1}{2}}} \mathbf{K}(s)ds, \eta_h^{k+\frac{1}{2}} + \eta_h^{k-\frac{1}{2}}) \end{aligned}$$

$$
\begin{aligned}
&+\frac{\tau}{\epsilon_0\omega_{pe}^2}\sum_{k=1}^{n}(\delta_\tau(\Pi_h\mathbf{J}^{k+\frac{1}{2}}-\mathbf{J}^{k+\frac{1}{2}}),\tilde{\xi}_h^{k+\frac{1}{2}}+\tilde{\xi}_h^{k-\frac{1}{2}})\\
&+\frac{\tau\Gamma_e}{\epsilon_0\omega_{pe}^2}\sum_{k=1}^{n}(\frac{1}{2}(\Pi_h\mathbf{J}^{k+\frac{1}{2}}+\Pi_h\mathbf{J}^{k-\frac{1}{2}})-\frac{1}{\tau}\int_{t_{k-\frac{1}{2}}}^{t_{k+\frac{1}{2}}}\mathbf{J}(s)ds,\tilde{\xi}_h^{k+\frac{1}{2}}+\tilde{\xi}_h^{k-\frac{1}{2}})\\
&+\tau\sum_{k=1}^{n}(\frac{1}{\tau}\int_{t_{k-\frac{1}{2}}}^{t_{k+\frac{1}{2}}}\mathbf{E}(s)ds-\Pi_h\mathbf{E}^k,\tilde{\xi}_h^{k+\frac{1}{2}}+\tilde{\xi}_h^{k-\frac{1}{2}})\\
&+\frac{\tau}{\mu_0\omega_{pm}^2}\sum_{k=1}^{n}(\delta_\tau(P_h\mathbf{K}^k-\mathbf{K}^k),\tilde{\eta}_h^k+\tilde{\eta}_h^{k-1})\\
&+\frac{\tau\Gamma_m}{\mu_0\omega_{pm}^2}\sum_{k=1}^{n}(\frac{1}{2}(P_h\mathbf{K}^k+P_h\mathbf{K}^{k-1})-\frac{1}{\tau}\int_{t_{k-1}}^{t_k}\mathbf{K}(s)ds,\tilde{\eta}_h^k+\tilde{\eta}_h^{k-1})\\
&+\tau\sum_{k=1}^{n}(\frac{1}{\tau}\int_{t_{k-1}}^{t_k}\mathbf{H}(s)ds-P_h\mathbf{H}^{k-\frac{1}{2}},\tilde{\eta}_h^k+\tilde{\eta}_h^{k-1})+\tau(\tilde{\xi}_h^{n+\frac{1}{2}},\xi_h^n)-\tau(\tilde{\xi}_h^{\frac{1}{2}},\xi_h^0)\\
&-\tau(\tilde{\eta}_h^n,\eta_h^{n+\frac{1}{2}})+\tau(\tilde{\eta}_h^0,\eta_h^{\frac{1}{2}})+\tau(\nabla\times\xi_h^0,\eta_h^{\frac{1}{2}})-\tau(\nabla\times\xi_h^n,\eta_h^{n+\frac{1}{2}})\\
=&\sum_{i=1}^{18}Err_i. \qquad (3.122)
\end{aligned}
$$

The proof is completed by using Lemma 3.19 and careful estimating all Err_i. Readers can consult the proof of Theorem 3.10 or the original paper [183]. □

Remark 3.2. We like to remark that a similar leap-frog scheme can be developed for Maxwell's equations in free space by dropping the constitutive equations (3.57) and (3.58), and treating **J** and **K** as fixed sources in (3.55) and (3.56). Following the same proof as carried out above, we can show that the stability and error estimate in the free space become as:

$$
\epsilon_0||\mathbf{E}_h^n||_0^2+\mu_0||\mathbf{H}_h^{n+\frac{1}{2}}||_0^2\le C[||\mathbf{E}_h^0||_0^2+||\mathbf{H}_h^{\frac{1}{2}}||_0^2],
$$

and

$$
\begin{aligned}
&\max_{1\le n}(||\mathbf{E}^n-\mathbf{E}_h^n||_0+||\mathbf{H}^{n+\frac{1}{2}}-\mathbf{H}_h^{n+\frac{1}{2}}||_0)\\
&\le C(\tau^2+h^l)+C(||\mathbf{E}^0-\mathbf{E}_h^0||_0+||\mathbf{H}^{\frac{1}{2}}-\mathbf{H}_h^{\frac{1}{2}}||_0).
\end{aligned}
$$

3.6 Extensions to the Lorentz Model

3.6.1 The Well-Posedness of the Lorentz Model

Another popular model for describing wave propagation in metamaterials is the Lorentz model discussed in Chap. 1. Recall that the time-domain Lorentz model is described by the following governing equations:

$$\epsilon_0 \frac{\partial \mathbf{E}}{\partial t} + \frac{\partial \mathbf{P}}{\partial t} - \nabla \times \mathbf{H} = 0, \tag{3.123}$$

$$\mu_0 \frac{\partial \mathbf{H}}{\partial t} + \frac{\partial \mathbf{M}}{\partial t} + \nabla \times \mathbf{E} = 0, \tag{3.124}$$

$$\frac{1}{\epsilon_0 \omega_{pe}^2} \frac{\partial^2 \mathbf{P}}{\partial t^2} + \frac{\Gamma_e}{\epsilon_0 \omega_{pe}^2} \frac{\partial \mathbf{P}}{\partial t} + \frac{\omega_{e0}^2}{\epsilon_0 \omega_{pe}^2} \mathbf{P} - \mathbf{E} = 0, \tag{3.125}$$

$$\frac{1}{\mu_0 \omega_{pm}^2} \frac{\partial^2 \mathbf{M}}{\partial t^2} + \frac{\Gamma_m}{\mu_0 \omega_{pm}^2} \frac{\partial \mathbf{M}}{\partial t} + \frac{\omega_{m0}^2}{\mu_0 \omega_{pm}^2} \mathbf{M} - \mathbf{H} = 0. \tag{3.126}$$

Note that the governing equations written in this way are convenient to obtain the stability and error analysis elegantly. Hopefully, readers may appreciate this as we move forward.

For simplicity, we assume that the model problem (3.123)–(3.126) is complemented by a perfect conducting boundary condition (3.59), and initial conditions

$$\mathbf{E}(\mathbf{x}, 0) = \mathbf{E}_0(\mathbf{x}), \quad \mathbf{H}(\mathbf{x}, 0) = \mathbf{H}_0(\mathbf{x}), \quad \mathbf{P}(\mathbf{x}, 0) = \mathbf{P}_0(\mathbf{x}), \tag{3.127}$$

$$\mathbf{M}(\mathbf{x}, 0) = \mathbf{M}_0(\mathbf{x}), \ \frac{\partial \mathbf{P}}{\partial t}(\mathbf{x}, 0) = \mathbf{P}_1(\mathbf{x}), \ \frac{\partial \mathbf{M}}{\partial t}(\mathbf{x}, 0) = \mathbf{M}_1(\mathbf{x}), \tag{3.128}$$

where $\mathbf{E}_0, \mathbf{H}_0, \mathbf{P}_0, \mathbf{M}_0, \mathbf{P}_1$, and $\mathbf{M}_1$ are some given functions.

Denote $\mathbf{V}^*$ as the dual space of $\mathbf{V} = H_0(\text{curl}; \Omega)$. Then a weak formulation of (3.123)–(3.126) can be formed as: Find the solution $\mathbf{E} \in H^1(0, T; \mathbf{V}^*) \cap (L^2(0, T; \mathbf{V}))^3, \mathbf{H} \in H^1(0, T; (L^2(\Omega))^3), \mathbf{P} \in H^2(0, T; \mathbf{V}^*), \mathbf{M} \in H^2(0, T; (L^2(\Omega))^3)$ such that

$$\epsilon_0 (\frac{\partial \mathbf{E}}{\partial t}, \boldsymbol{\phi}) + (\frac{\partial \mathbf{P}}{\partial t}, \boldsymbol{\phi}) - (\mathbf{H}, \nabla \times \boldsymbol{\phi}) = 0, \tag{3.129}$$

$$\mu_0 (\frac{\partial \mathbf{H}}{\partial t}, \boldsymbol{\psi}) + (\frac{\partial \mathbf{M}}{\partial t}, \boldsymbol{\psi}) + (\nabla \times \mathbf{E}, \boldsymbol{\psi}) = 0, \tag{3.130}$$

$$\frac{1}{\epsilon_0 \omega_{pe}^2} [(\frac{\partial^2 \mathbf{P}}{\partial t^2}, \tilde{\boldsymbol{\phi}}) + \Gamma_e (\frac{\partial \mathbf{P}}{\partial t}, \tilde{\boldsymbol{\phi}}) + \omega_{e0}^2 (\mathbf{P}, \tilde{\boldsymbol{\phi}})] - (\mathbf{E}, \tilde{\boldsymbol{\phi}}) = 0, \tag{3.131}$$

$$\frac{1}{\mu_0 \omega_{pm}^2} [(\frac{\partial^2 \mathbf{M}}{\partial t^2}, \tilde{\boldsymbol{\psi}}) + \Gamma_m (\frac{\partial \mathbf{M}}{\partial t}, \tilde{\boldsymbol{\psi}}) + \omega_{m0}^2 (\mathbf{M}, \tilde{\boldsymbol{\psi}})] - (\mathbf{H}, \tilde{\boldsymbol{\psi}}) = 0, \tag{3.132}$$

hold true for any $\boldsymbol{\phi} \in H_0(\text{curl};\Omega), \boldsymbol{\psi} \in (L^2(\Omega))^3, \tilde{\boldsymbol{\phi}} \in H(\text{curl};\Omega)$, and $\tilde{\boldsymbol{\psi}} \in (L^2(\Omega))^3$.

First, we have the following stability for the Lorentz model (3.129)–(3.132).

Lemma 3.20.

$$\begin{aligned}
&\epsilon_0||\mathbf{E}(t)||_0^2 + \mu_0||\mathbf{H}(t)||_0^2 + \frac{1}{\epsilon_0\omega_{pe}^2}||\frac{\partial \mathbf{P}}{\partial t}(t)||_0^2 \\
&\quad + \frac{1}{\mu_0\omega_{pm}^2}||\frac{\partial \mathbf{M}}{\partial t}(t)||_0^2 + \frac{\omega_{e0}^2}{\epsilon_0\omega_{pe}^2}||\mathbf{P}(t)||_0^2 + \frac{\omega_{m0}^2}{\mu_0\omega_{pm}^2}||\mathbf{M}(t)||_0^2 \\
&\le \epsilon_0||\mathbf{E}(0)||_0^2 + \mu_0||\mathbf{H}(0)||_0^2 + \frac{1}{\epsilon_0\omega_{pe}^2}||\frac{\partial \mathbf{P}}{\partial t}(0)||_0^2 \\
&\quad + \frac{1}{\mu_0\omega_{pm}^2}||\frac{\partial \mathbf{M}}{\partial t}(0)||_0^2 + \frac{\omega_{e0}^2}{\epsilon_0\omega_{pe}^2}||\mathbf{P}(0)||_0^2 + \frac{\omega_{m0}^2}{\mu_0\omega_{pm}^2}||\mathbf{M}(0)||_0^2.
\end{aligned}$$

Proof. Choosing $\boldsymbol{\phi} = \mathbf{E}, \boldsymbol{\psi} = \mathbf{H}, \tilde{\boldsymbol{\phi}} = \frac{\partial \mathbf{P}}{\partial t}, \tilde{\boldsymbol{\psi}} = \frac{\partial \mathbf{M}}{\partial t}$ in (3.129)–(3.132) respectively, then summing up the resultants, we have

$$\begin{aligned}
&\frac{1}{2}\frac{d}{dt}[\epsilon_0||\mathbf{E}||_0^2 + \frac{1}{\epsilon_0\omega_{pe}^2}||\frac{\partial \mathbf{P}}{\partial t}||_0^2 + \frac{\omega_{e0}^2}{\epsilon_0\omega_{pe}^2}||\mathbf{P}||_0^2] + \frac{\Gamma_e}{\epsilon_0\omega_{pe}^2}||\frac{\partial \mathbf{P}}{\partial t}||_0^2 \\
&+\frac{1}{2}\frac{d}{dt}[\mu_0||\mathbf{H}||_0^2 + \frac{1}{\mu_0\omega_{pm}^2}||\frac{\partial \mathbf{M}}{\partial t}||_0^2 + \frac{\omega_{m0}^2}{\mu_0\omega_{pm}^2}||\mathbf{M}||_0^2] + \frac{\Gamma_m}{\mu_0\omega_{pm}^2}||\frac{\partial \mathbf{M}}{\partial t}||_0^2 = 0,
\end{aligned}$$

integrating which from 0 to t completes the proof. □

The existence and uniqueness of the solution for the Lorentz model (3.123)–(3.128) can be proved by following the same technique developed earlier in Theorem 3.8 for the Drude model.

Theorem 3.13. *The model (3.123)–(3.128) exists a unique solution* $\mathbf{E} \in H_0(\text{curl};\Omega)$ *and* $\mathbf{H} \in H(\text{curl};\Omega)$.

Proof. Taking the Laplace transform of (3.123) and (3.124), we obtain

$$\epsilon_0(s\hat{\mathbf{E}} - \mathbf{E}_0) + s\hat{\mathbf{P}} - \mathbf{P}_0 - \nabla \times \hat{\mathbf{H}} = 0 \tag{3.133}$$

$$\mu_0(s\hat{\mathbf{H}} - \mathbf{H}_0) + s\hat{\mathbf{M}} - \mathbf{M}_0 + \nabla \times \hat{\mathbf{E}} = 0. \tag{3.134}$$

Taking the Laplace transform of (3.125) and (3.126), we have

$$\hat{\mathbf{P}} = [(s+\Gamma_e)\mathbf{P}_0 + \mathbf{P}'(0) + \hat{\mathbf{E}}]/(s^2 + \Gamma_e s + \omega_{e0}^2) \tag{3.135}$$

$$\hat{\mathbf{M}} = [(s+\Gamma_m)\mathbf{M}_0 + \mathbf{M}'(0) + \hat{\mathbf{H}}]/(s^2 + \Gamma_m s + \omega_{m0}^2). \tag{3.136}$$

Substituting (3.135) into (3.133), we have

$$(\epsilon_0 s + \frac{s}{s^2 + \Gamma_e s + \omega_{e0}^2})\hat{\mathbf{E}} - \nabla \times \hat{\mathbf{H}}$$
$$= \epsilon_0 \mathbf{E}_0 + \mathbf{P}_0 - \frac{[(s + \Gamma_e)\mathbf{P}_0 + \mathbf{P}'(0)]s}{s^2 + \Gamma_e s + \omega_{e0}^2} \equiv \mathbf{f}_0, \tag{3.137}$$
$$(\mu_0 s + \frac{s}{s^2 + \Gamma_m s + \omega_{m0}^2})\hat{\mathbf{H}} + \nabla \times \hat{\mathbf{E}}$$
$$= \mu_0 \mathbf{H}_0 + \mathbf{M}_0 - \frac{[(s + \Gamma_m)\mathbf{M}_0 + \mathbf{M}'(0)]s}{s^2 + \Gamma_m s + \omega_{m0}^2} \equiv \mathbf{g}_0. \tag{3.138}$$

Eliminating $\hat{\mathbf{H}}$ from (3.137) and (3.138), we obtain

$$(\epsilon_0 s + \frac{s}{s^2 + \Gamma_e s + \omega_{e0}^2})(\mu_0 s + \frac{s}{s^2 + \Gamma_m s + \omega_{m0}^2})\hat{\mathbf{E}} + \nabla \times \nabla \times \hat{\mathbf{E}}$$
$$= \nabla \times \mathbf{g}_0 + (\mu_0 s + \frac{s}{s^2 + \Gamma_m s + \omega_{m0}^2})\mathbf{f}_0, \tag{3.139}$$

whose weak formulation can be written as: Find $\hat{\mathbf{E}} \in H_0(\text{curl}; \Omega)$ such that

$$(\epsilon_0 s + \frac{s}{s^2 + \Gamma_e s + \omega_{e0}^2})(\mu_0 s + \frac{s}{s^2 + \Gamma_m s + \omega_{m0}^2})(\hat{\mathbf{E}}, \boldsymbol{\phi}) + (\nabla \times \hat{\mathbf{E}}, \nabla \times \boldsymbol{\phi})$$
$$= (\nabla \times \mathbf{g}_0 + (\mu_0 s + \frac{s}{s^2 + \Gamma_m s + \omega_{m0}^2})\mathbf{f}_0, \boldsymbol{\phi}), \quad \forall\, \boldsymbol{\phi} \in H_0(\text{curl}; \Omega). \tag{3.140}$$

By the Lax-Milgram lemma, the Eq. (3.140) exists a unique solution $\hat{\mathbf{E}} \in H_0(\text{curl}; \Omega)$ for any $s > 0$. The existence and uniqueness of $\hat{\mathbf{H}} \in H(\text{curl}; \Omega)$ can be assured by using the same argument, in which case we just eliminate $\hat{\mathbf{E}}$ from (3.137) and (3.138). The solutions $\mathbf{E}$ and $\mathbf{H}$ are the inverse Laplace transforms of $\hat{\mathbf{E}}$ and $\hat{\mathbf{H}}$, respectively. □

3.6.2 The Crank-Nicolson Scheme and Error Analysis

3.6.2.1 The Scheme and Stability Analysis

We first consider a Crank-Nicolson type scheme: For $k = 1, 2, \cdots, N-1$, find $\mathbf{E}_h^{k+1} \in \mathbf{V}_h^0, \mathbf{P}_h^{k+1} \in \mathbf{V}_h, \mathbf{H}_h^{k+1}, \mathbf{M}_h^{k+1} \in \mathbf{U}_h$ such that

$$\epsilon_0(\delta_{2\tau}\mathbf{E}_h^k, \boldsymbol{\phi}_h) + (\delta_{2\tau}\mathbf{P}_h^k, \boldsymbol{\phi}_h) - (\overline{\mathbf{H}}_h^k, \nabla \times \boldsymbol{\phi}_h) = 0, \tag{3.141}$$

$$\mu_0(\delta_{2\tau}\mathbf{H}_h^k, \boldsymbol{\psi}_h) + (\delta_{2\tau}\mathbf{M}_h^k, \boldsymbol{\psi}_h) + (\nabla \times \overline{\mathbf{E}}_h^k, \boldsymbol{\psi}_h) = 0, \tag{3.142}$$

$$\frac{1}{\epsilon_0\omega_{pe}^2}[(\frac{\delta_\tau\mathbf{P}_h^{k+1} - \delta_\tau\mathbf{P}_h^k}{\tau}, \tilde{\boldsymbol{\phi}}_h) + \Gamma_e(\delta_{2\tau}\mathbf{P}_h^k, \tilde{\boldsymbol{\phi}}_h) + \omega_{e0}^2(\overline{\mathbf{P}}_h^k, \tilde{\boldsymbol{\phi}}_h)] = (\overline{\mathbf{E}}_h^k, \tilde{\boldsymbol{\phi}}_h), \tag{3.143}$$

$$\frac{1}{\mu_0\omega_{pm}^2}[(\frac{\delta_\tau\mathbf{M}_h^{k+1} - \delta_\tau\mathbf{M}_h^k}{\tau}, \tilde{\boldsymbol{\psi}}_h) + \Gamma_m(\delta_{2\tau}\mathbf{M}_h^k, \tilde{\boldsymbol{\psi}}_h) + \omega_{m0}^2(\overline{\mathbf{M}}_h^k, \tilde{\boldsymbol{\psi}}_h)] = (\overline{\mathbf{H}}_h^k, \tilde{\boldsymbol{\psi}}_h), \tag{3.144}$$

hold true for any $\boldsymbol{\phi}_h \in \mathbf{V}_h^0, \boldsymbol{\psi}_h \in \mathbf{U}_h, \tilde{\boldsymbol{\phi}}_h \in \mathbf{V}_h$ and $\tilde{\boldsymbol{\psi}}_h \in \mathbf{U}_h$. Here we denote

$$\delta_\tau u^k = (u^k - u^{k-1})/\tau, \quad \delta_{2\tau} u^k = (u^{k+1} - u^{k-1})/2\tau, \quad \overline{u}^k = (u^{k+1} + u^{k-1})/2.$$

To implement this scheme, we use the following initial approximations:

$$\mathbf{E}_h^0(\mathbf{x}) = \Pi_h\mathbf{E}_0(\mathbf{x}), \quad \mathbf{H}_h^0(\mathbf{x}) = P_h\mathbf{H}_0(\mathbf{x}), \tag{3.145}$$

$$\mathbf{P}_h^0(\mathbf{x}) = \Pi_h\mathbf{P}_0(\mathbf{x}), \quad \mathbf{M}_h^0(\mathbf{x}) = P_h\mathbf{M}_0(\mathbf{x}), \tag{3.146}$$

$$\begin{aligned}\mathbf{E}_h^1(\mathbf{x}) &= \Pi_h\mathbf{E}(\mathbf{x}, \tau) \approx \Pi_h(\mathbf{E}(\mathbf{x}, 0) + \tau\mathbf{E}_t(\mathbf{x}, 0)) \\ &= \Pi_h[\mathbf{E}_0(\mathbf{x}) + \tau\epsilon_0^{-1}(\nabla \times \mathbf{H}_0(\mathbf{x}) - \mathbf{P}_1(\mathbf{x}))],\end{aligned} \tag{3.147}$$

$$\mathbf{H}_h^1(\mathbf{x}) = P_h\mathbf{H}(\mathbf{x}, \tau) \approx P_h[\mathbf{H}_0(\mathbf{x}) - \tau\mu_0^{-1}(\nabla \times \mathbf{E}_0(\mathbf{x}) + \mathbf{M}_1(\mathbf{x}))], \tag{3.148}$$

$$\begin{aligned}\mathbf{P}_h^1(\mathbf{x}) &= \Pi_h\mathbf{P}(\mathbf{x}, \tau) \approx \Pi_h(\mathbf{P}(\mathbf{x}, 0) + \tau\mathbf{P}_t(\mathbf{x}, 0) + \frac{\tau^2}{2}\mathbf{P}_{tt}(\mathbf{x}, 0)) \\ &= \Pi_h[\mathbf{P}_0(\mathbf{x}) + \tau\mathbf{P}_1(\mathbf{x}) + \frac{\tau^2}{2}(\epsilon_0\omega_{pe}^2\mathbf{E}_0 - \omega_{e0}^2\mathbf{P}_0 - \Gamma_e\mathbf{P}_1)],\end{aligned} \tag{3.149}$$

$$\mathbf{M}_h^1(\mathbf{x}) = P_h[\mathbf{M}_0(\mathbf{x}) + \tau\mathbf{M}_1(\mathbf{x}) + \frac{\tau^2}{2}(\mu_0\omega_{pm}^2\mathbf{H}_0 - \omega_{m0}^2\mathbf{M}_0 - \Gamma_m\mathbf{M}_1)]. \tag{3.150}$$

For this scheme, we have the following stability:

Lemma 3.21. *For any $k \geq 2$, we have*

$$\begin{aligned}&\epsilon_0||\mathbf{E}_h^k||_0^2 + \mu_0||\mathbf{H}_h^k||_0^2 + \frac{2}{\epsilon_0\omega_{pe}^2}||\delta_\tau\mathbf{P}_h^k||_0^2 \\ &+ \frac{2}{\mu_0\omega_{pm}^2}||\delta_\tau\mathbf{M}_h^k||_0^2 + \frac{\omega_{e0}^2}{\epsilon_0\omega_{pe}^2}||\mathbf{P}_h^k||_0^2 + \frac{\omega_{m0}^2}{\mu_0\omega_{pm}^2}||\mathbf{M}_h^k||_0^2 \\ &\leq \epsilon_0||\mathbf{E}_h^1||_0^2 + \mu_0||\mathbf{H}_h^1||_0^2 + \frac{\omega_{e0}^2}{\epsilon_0\omega_{pe}^2}||\mathbf{P}_h^1||_0^2 + \frac{\omega_{m0}^2}{\mu_0\omega_{pm}^2}||\mathbf{M}_h^1||_0^2 \\ &+ \epsilon_0||\mathbf{E}_h^0||_0^2 + \mu_0||\mathbf{H}_h^0||_0^2 + \frac{2}{\epsilon_0\omega_{pe}^2}||\delta_\tau\mathbf{P}_h^1||_0^2\end{aligned}$$

$$+\frac{2}{\mu_0\omega_{pm}^2}||\delta_\tau \mathbf{M}_h^1||_0^2+\frac{\omega_{e0}^2}{\epsilon_0\omega_{pe}^2}||\mathbf{P}_h^0||_0^2+\frac{\omega_{m0}^2}{\mu_0\omega_{pm}^2}||\mathbf{M}_h^0||_0^2.$$

Proof. Choosing $\boldsymbol{\phi}_h=\tau(\mathbf{E}_h^{k+1}+\mathbf{E}_h^{k-1})$, $\boldsymbol{\psi}_h=\tau(\mathbf{H}_h^{k+1}+\mathbf{H}_h^{k-1})$, $\tilde{\boldsymbol{\phi}}_h=\tau(\delta_\tau\mathbf{P}_h^{k+1}+\delta_\tau\mathbf{P}_h^{k})$, $\tilde{\boldsymbol{\psi}}_h=\tau(\delta_\tau\mathbf{M}_h^{k+1}+\delta_\tau\mathbf{M}_h^{k})$ in (3.141)–(3.144), respectively, and adding the resultants together, we have

$$\begin{aligned}
&\frac{\epsilon_0}{2}(||\mathbf{E}_h^{k+1}||_0^2-||\mathbf{E}_h^{k-1}||_0^2)+\frac{\mu_0}{2}(||\mathbf{H}_h^{k+1}||_0^2-||\mathbf{H}_h^{k-1}||_0^2)\\
&+\frac{1}{\epsilon_0\omega_{pe}^2}(||\delta_\tau\mathbf{P}_h^{k+1}||_0^2-||\delta_\tau\mathbf{P}_h^{k}||_0^2)+\frac{1}{\mu_0\omega_{pm}^2}(||\delta_\tau\mathbf{M}_h^{k+1}||_0^2-||\delta_\tau\mathbf{M}_h^{k}||_0^2)\\
&+\frac{\omega_{e0}^2}{2\epsilon_0\omega_{pe}^2}(||\mathbf{P}_h^{k+1}||_0^2-||\mathbf{P}_h^{k-1}||_0^2)+\frac{\omega_{m0}^2}{2\mu_0\omega_{pm}^2}(||\mathbf{M}_h^{k+1}||_0^2-||\mathbf{M}_h^{k-1}||_0^2)\le 0.
\end{aligned}$$

Summing up the above estimate from $k=1$ to $k=n-1$, and using the identity

$$\sum_{k=1}^{n}(a_{k+1}^2-a_{k-1}^2)=a_{n+1}^2+a_n^2-a_1^2-a_0^2,$$

we can easily see that the proof completes. □

Let us look at the scheme carefully. From (3.143), we have

$$\mathbf{P}_h^{k+1}=a_1(\mathbf{E}_h^{k+1}+\mathbf{E}_h^{k-1})+a_2\mathbf{P}_h^{k}-a_3\mathbf{P}_h^{k-1}, \tag{3.151}$$

where

$$a_1=\frac{\epsilon_0\omega_{pe}^2\tau^2}{2+\tau\Gamma_e+\tau^2\omega_{e0}^2},\ a_2=\frac{4}{2+\tau\Gamma_e+\tau^2\omega_{e0}^2},\ a_3=\frac{2-\tau\Gamma_e+\tau^2\omega_{e0}^2}{2+\tau\Gamma_e+\tau^2\omega_{e0}^2}.$$

Similarly, from (3.144), we have

$$\mathbf{M}_h^{k+1}=\tilde{a}_1(\mathbf{H}_h^{k+1}+\mathbf{H}_h^{k-1})+\tilde{a}_2\mathbf{M}_h^{k}-\tilde{a}_3\mathbf{M}_h^{k-1}, \tag{3.152}$$

where

$$\tilde{a}_1=\frac{\mu_0\omega_{pm}^2\tau^2}{2+\tau\Gamma_m+\tau^2\omega_{m0}^2},\ \tilde{a}_2=\frac{4}{2+\tau\Gamma_m+\tau^2\omega_{m0}^2},\ \tilde{a}_3=\frac{2-\tau\Gamma_m+\tau^2\omega_{m0}^2}{2+\tau\Gamma_m+\tau^2\omega_{m0}^2}.$$

Substituting (3.151) and (3.152) into (3.141) and (3.142), respectively, we obtain

$$\begin{aligned}
&(\epsilon_0+a_1)(\mathbf{E}_h^{k+1},\boldsymbol{\phi}_h)-\tau(\mathbf{H}_h^{k+1},\nabla\times\boldsymbol{\phi}_h)=(\epsilon_0-a_1)(\mathbf{E}_h^{k-1},\boldsymbol{\phi}_h)\\
&-a_2(\mathbf{P}_h^{k},\boldsymbol{\phi}_h)+(1+a_3)(\mathbf{P}_h^{k-1},\boldsymbol{\phi}_h)+\tau(\mathbf{H}_h^{k-1},\nabla\times\boldsymbol{\phi}_h),
\end{aligned} \tag{3.153}$$

and

$$(\mu_0 + \tilde{a}_1)(\mathbf{H}_h^{k+1}, \boldsymbol{\psi}_h) + \tau(\nabla \times \mathbf{E}_h^{k+1}, \boldsymbol{\psi}_h) = (\mu_0 - \tilde{a}_1)(\mathbf{H}_h^{k-1}, \boldsymbol{\psi}_h)$$
$$-\tilde{a}_2(\mathbf{M}_h^k, \boldsymbol{\psi}_h) + (1 + \tilde{a}_3)(\mathbf{M}_h^{k-1}, \boldsymbol{\psi}_h) - \tau(\nabla \times \mathbf{E}_h^{k-1}, \boldsymbol{\psi}_h), \qquad (3.154)$$

whose system's coefficient matrix can be written as $\begin{pmatrix} A & -B' \\ B & D \end{pmatrix}$, where matrices $A = (\epsilon_0 + a_1)(\boldsymbol{\phi}_h, \boldsymbol{\phi}_h)$ and $D = (\mu_0 + \tilde{a}_1)(\boldsymbol{\psi}_h, \boldsymbol{\psi}_h)$ are symmetric positive definite, and matrix $B = \tau(\nabla \times \boldsymbol{\phi}_h, \boldsymbol{\psi}_h)$. Here B' denotes the transpose of B. It is easy to see that the determinant of the coefficient matrix is

$$\begin{vmatrix} A & -B' \\ B & D \end{vmatrix} = |A| \cdot \begin{vmatrix} I & -A^{-1}B' \\ B & D \end{vmatrix} = |A| \cdot \begin{vmatrix} I & -A^{-1}B' \\ 0 & D + BA^{-1}B' \end{vmatrix} = |A| \cdot |D + BA^{-1}B'|,$$

which is non-zero. Hence in practical implementation of the scheme (3.141)–(3.144), at each time step we can solve (3.153) and (3.154) for $(\mathbf{E}_h^{k+1}, \mathbf{H}_h^{k+1})$, then update $\mathbf{P}_h^{k+1}$ and $\mathbf{M}_h^{k+1}$ using (3.151) and (3.152), respectively.

3.6.2.2 The Optimal Error Estimate

Before proving the optimal error estimate for the Crank-Nicolson scheme (3.141)–(3.144), we need some estimates.

Lemma 3.22 ([184, Lemma 4.3]). *For any $p(x,t) \in H^3(0,T; L^2(\Omega))$ and $k \geq 1$, we have*

$$||(\frac{p^{k+1} + p^k}{2} - p^{k+\frac{1}{2}}) - (\frac{p^k + p^{k-1}}{2} - p^{k-\frac{1}{2}})||_0^2 \leq \frac{\tau^5}{8} \int_{t_{k-1}}^{t_{k+1}} ||p_{s^3}(s)||_0^2 ds.$$

Proof. By Taylor expansions, it is easy to see the following identity is true:

$$\frac{p(t) + p(t-\tau)}{2} - p(t - \frac{\tau}{2}) = \frac{1}{2}[\int_{t-\frac{\tau}{2}}^{t} (t-s)p_{s^2}(s)ds + \int_{t-\tau}^{t-\frac{\tau}{2}} (s-t+\tau)p_{s^2}(s)ds],$$

applying which to p_t we obtain

$$\begin{aligned}
&||(\frac{p^{k+1} + p^k}{2} - p^{k+\frac{1}{2}}) - (\frac{p^k + p^{k-1}}{2} - p^{k-\frac{1}{2}})||_0^2 \\
&= ||\int_{t_k}^{t_{k+1}} \frac{d}{dt}(\frac{p(t) + p(t-\tau)}{2} - p(t - \frac{\tau}{2}))dt||_0^2 \\
&= ||\frac{1}{2}\int_{t_k}^{t_{k+1}} [\int_{t-\frac{\tau}{2}}^{t} (t-s)p_{s^3}(s)ds + \int_{t-\tau}^{t-\frac{\tau}{2}} (s-t+\tau)p_{s^3}(s)ds]dt||_0^2
\end{aligned}$$

$$\leq \frac{\tau}{4}\int_{t_k}^{t_{k+1}} 2[||\int_{t-\frac{\tau}{2}}^{t}(t-s)p_{s^3}(s)ds||_0^2 + ||\int_{t-\tau}^{t-\frac{\tau}{2}}(s-t+\tau)p_{s^3}(s)ds||_0^2]dt$$

$$\leq \frac{\tau}{2}\int_{t_k}^{t_{k+1}} [\frac{\tau}{2}||\int_{t-\frac{\tau}{2}}^{t}(t-s)^2||p_{s^3}(s)||_0^2 ds + \frac{\tau}{2}\int_{t-\tau}^{t-\frac{\tau}{2}}(s-t+\tau)^2||p_{s^3}(s)||_0^2 ds]dt$$

$$\leq \frac{\tau^3}{2}\cdot\frac{\tau^2}{4}\int_{t_{k-1}}^{t_{k+1}}||p_{s^3}(s)||_0^2 ds,$$

from which the proof completes. □

By similar arguments, the following two estimates can be proved (cf. [184]).

Lemma 3.23. *For any $p(x,t) \in H^4(0,T;L^2(\Omega))$ and $k \geq 1$, we have*

$$||(\delta_\tau p^{k+1} - p_t^{k+\frac{1}{2}}) - (\delta_\tau p^k - p_t^{k-\frac{1}{2}})||_0^2 \leq \frac{\tau^5}{32}\int_{t_{k-1}}^{t_{k+1}}||p_{s^4}(s)||_0^2 ds.$$

Lemma 3.24. *For any $p(x,t) \in H^2(0,T;L^2(\Omega))$ and $k \geq 1$, we have*

$$||\frac{p^{k+1}+p^{k-1}}{2} - p^k||_0^2 \leq \tau^3\int_{t_{k-1}}^{t_{k+1}}||p_{s^2}(s)||_0^2 ds.$$

With the above estimates and proper regularity assumptions, we can prove the following optimal error estimate.

Theorem 3.14. *Let $(\mathbf{E}^m, \mathbf{H}^m, \mathbf{P}^m, \mathbf{M}^m)$ and $(\mathbf{E}_h^m, \mathbf{H}_h^m, \mathbf{P}_h^m, \mathbf{M}_h^m)$ be the analytic and numerical solutions of (3.129)–(3.132) and (3.141)–(3.144) at time t_m, respectively. Under the regularity assumptions*

$$\mathbf{E} \in H^1(0,T;H^l(curl;\Omega)) \cap L^\infty(0,T;H^l(curl;\Omega)) \cap H^2(0,T;H(curl;\Omega)),$$
$$\mathbf{P} \in H^2(0,T;H^l(curl;\Omega)) \cap L^\infty(0,T;H^l(curl;\Omega)) \cap H^4(0,T;(L^2(\Omega))^3),$$
$$\mathbf{H} \in H^2(0,T;H(curl;\Omega)) \cap L^\infty(0,T;(H^l(\Omega))^3),$$
$$\mathbf{M} \in H^4(0,T;(L^2(\Omega))^3) \cap L^\infty(0,T;(H^l(\Omega))^3),$$

there exists a constant $C > 0$ independent of mesh size h and time step τ, such that

$$\epsilon_0||\mathbf{E}^n - \mathbf{E}_h^n||_0^2 + \frac{\omega_{e0}^2}{\epsilon_0\omega_{pe}^2}||\mathbf{P}^n - \mathbf{P}_h^n||_0^2 + \mu_0||\mathbf{H}^n - \mathbf{H}_h^n||_0^2 + \frac{\omega_{m0}^2}{\mu_0\omega_{pm}^2}||\mathbf{M}^n - \mathbf{M}_h^n||_0^2$$
$$\leq C(h^{2l}+\tau^4) + C(||\xi_h^0||_0^2 + ||\tilde{\xi}_h^0||_0^2 + ||\eta_h^0||_0^2 + ||\tilde{\eta}_h^0||_0^2),$$
$$+C[||\xi_h^1||_0^2 + ||\delta_\tau\tilde{\xi}_h^1||_0^2 + ||\tilde{\xi}_h^1||_0^2 + ||\eta_h^1||_0^2 + ||\delta_\tau\tilde{\eta}_h^1||_0^2 + ||\tilde{\eta}_h^1||_0^2],$$

where $\xi_h^k = \Pi_h \mathbf{E}^k - \mathbf{E}_h^k$, $\eta_h^k = P_h \mathbf{H}^k - \mathbf{H}_h^k$, $\tilde{\xi}_h^k = \Pi_h \mathbf{P}^k - \mathbf{P}_h^k$, $\tilde{\eta}_h^k = P_h \mathbf{M}^k - \mathbf{M}_h^k$, *and* $l \geq 1$ *is the order of the basis functions in spaces* $\mathbf{U}_h$ *and* $\mathbf{V}_h$.

Remark 3.3. Note that the initial approximations (3.145)–(3.150) yield the initial errors

$$||\xi_h^0||_0 = ||\tilde{\xi}_h^0||_0 = ||\eta_h^0||_0 = ||\tilde{\eta}_h^0||_0 = 0,$$
$$||\xi_h^1||_0 = ||\delta_\tau \tilde{\xi}_h^1||_0 = ||\tilde{\xi}_h^1||_0 = ||\eta_h^1||_0 = ||\delta_\tau \tilde{\eta}_h^1||_0 = ||\tilde{\eta}_h^1||_0 = O(h^l + \tau^2),$$

from which we have the optimal error estimate

$$||\mathbf{E} - \mathbf{E}_h^n||_0 + ||\mathbf{P} - \mathbf{P}_h^n||_0 + ||\mathbf{H} - \mathbf{H}_h^n||_0 + ||\mathbf{M} - \mathbf{M}_h^n||_0 \leq C(h^l + \tau^2).$$

Proof. Integrating (3.129) and (3.130) in time from $t = t_{k-1}$ to t_{k+1} and dividing all by 2τ, and integrating (3.131) and (3.132) in time from $t = t_{k-\frac{1}{2}}$ to $t = t_{k+\frac{1}{2}}$ and then dividing by τ, we have

$$\epsilon_0(\delta_{2\tau}\mathbf{E}^k, \boldsymbol{\phi}) + (\delta_{2\tau}\mathbf{P}^k, \boldsymbol{\phi}) - (\frac{1}{2\tau}\int_{t_{k-1}}^{t_{k+1}} \mathbf{H}(s)ds, \nabla \times \boldsymbol{\phi}) = 0, \tag{3.155}$$

$$\mu_0(\delta_{2\tau}\mathbf{H}^k, \boldsymbol{\psi}) + (\delta_{2\tau}\mathbf{M}^k, \boldsymbol{\psi}) + (\frac{1}{2\tau}\int_{t_{k-1}}^{t_{k+1}} \nabla \times \mathbf{E}(s)ds, \boldsymbol{\psi}) = 0, \tag{3.156}$$

$$\begin{aligned}
&\frac{1}{\tau\epsilon_0\omega_{pe}^2}(\frac{\partial \mathbf{P}^{k+\frac{1}{2}}}{\partial t} - \frac{\partial \mathbf{P}^{k-\frac{1}{2}}}{\partial t}, \tilde{\boldsymbol{\phi}}) + \frac{\Gamma_e}{\epsilon_0\omega_{pe}^2}(\frac{\mathbf{P}^{k+\frac{1}{2}} - \mathbf{P}^{k-\frac{1}{2}}}{\tau}, \tilde{\boldsymbol{\phi}}) \\
&\quad + \frac{\omega_{e0}^2}{\epsilon_0\omega_{pe}^2}(\frac{1}{\tau}\int_{t_{k-\frac{1}{2}}}^{t_{k+\frac{1}{2}}} \mathbf{P}(s)ds, \tilde{\boldsymbol{\phi}}) - (\frac{1}{\tau}\int_{t_{k-\frac{1}{2}}}^{t_{k+\frac{1}{2}}} \mathbf{E}(s)ds, \tilde{\boldsymbol{\phi}}) = 0,
\end{aligned} \tag{3.157}$$

$$\begin{aligned}
&\frac{1}{\tau\mu_0\omega_{pm}^2}(\frac{\partial \mathbf{M}^{k+\frac{1}{2}}}{\partial t} - \frac{\partial \mathbf{M}^{k-\frac{1}{2}}}{\partial t}, \tilde{\boldsymbol{\psi}}) + \frac{\Gamma_m}{\mu_0\omega_{pm}^2}(\frac{\mathbf{M}^{k+\frac{1}{2}} - \mathbf{M}^{k-\frac{1}{2}}}{\tau}, \tilde{\boldsymbol{\psi}}) \\
&\quad + \frac{\omega_{m0}^2}{\mu_0\omega_{pm}^2}(\frac{1}{\tau}\int_{t_{k-\frac{1}{2}}}^{t_{k+\frac{1}{2}}} \mathbf{M}(s)ds, \tilde{\boldsymbol{\psi}}) - (\frac{1}{\tau}\int_{t_{k-\frac{1}{2}}}^{t_{k+\frac{1}{2}}} \mathbf{H}(s)ds, \tilde{\boldsymbol{\psi}}) = 0.
\end{aligned} \tag{3.158}$$

Denote $\xi_h^k = \Pi_h \mathbf{E}^k - \mathbf{E}_h^k$, $\eta_h^k = P_h \mathbf{H}^k - \mathbf{H}_h^k$, $\tilde{\xi}_h^k = \Pi_h \mathbf{P}^k - \mathbf{P}_h^k$, $\tilde{\eta}_h^k = P_h \mathbf{M}^k - \mathbf{M}_h^k$. Subtracting (3.141)–(3.144) from (3.155)–(3.158) with $\boldsymbol{\phi} = \boldsymbol{\phi}_h$, $\boldsymbol{\psi} = \boldsymbol{\psi}_h$, $\tilde{\boldsymbol{\phi}} = \tilde{\boldsymbol{\phi}}_h$, and $\tilde{\boldsymbol{\psi}} = \tilde{\boldsymbol{\psi}}_h$, using the property of operator P_h, we obtain the error equations

$$\begin{aligned}
(i) \quad &\epsilon_0(\frac{\xi_h^{k+1} - \xi_h^{k-1}}{2\tau}, \boldsymbol{\phi}_h) + (\frac{\tilde{\xi}_h^{k+1} - \tilde{\xi}_h^{k-1}}{2\tau}, \boldsymbol{\phi}_h) - \frac{1}{2}(\eta_h^{k+1} + \eta_h^{k-1}, \nabla \times \boldsymbol{\phi}_h) \\
&= \epsilon_0(\frac{(\Pi_h \mathbf{E}^{k+1} - \Pi_h \mathbf{E}^{k-1}) - (\mathbf{E}^{k+1} - \mathbf{E}^{k-1})}{2\tau}, \boldsymbol{\phi}_h)
\end{aligned}$$

$$+\left(\frac{(\Pi_h\mathbf{P}^{k+1}-\Pi_h\mathbf{P}^{k-1})-(\mathbf{P}^{k+1}-\mathbf{P}^{k-1})}{2\tau},\boldsymbol{\phi}_h\right)$$
$$-\left(\frac{1}{2}(\mathbf{H}^{k+1}+\mathbf{H}^{k-1})-\frac{1}{2\tau}\int_{t_{k-1}}^{t_{k+1}}\mathbf{H}(s)ds,\nabla\times\boldsymbol{\phi}_h\right), \quad (3.159)$$

$$(ii)\quad \mu_0\left(\frac{\eta_h^{k+1}-\eta_h^{k-1}}{2\tau},\boldsymbol{\psi}_h\right)+\left(\frac{\tilde{\eta}_h^{k+1}-\tilde{\eta}_h^{k-1}}{2\tau},\boldsymbol{\psi}_h\right)+\frac{1}{2}(\nabla\times(\xi_h^{k+1}+\xi_h^{k-1}),\boldsymbol{\psi}_h)$$
$$=\left(\frac{1}{2}(\nabla\times\Pi_h\mathbf{E}^{k+1}+\nabla\times\Pi_h\mathbf{E}^{k-1})-\frac{1}{2\tau}\int_{t_{k-1}}^{t_{k+1}}\nabla\times\mathbf{E}(s)ds,\boldsymbol{\psi}_h\right), \quad (3.160)$$

$$(iii)\quad \frac{1}{\tau\epsilon_0\omega_{pe}^2}(\delta_\tau\tilde{\xi}_h^{k+1}-\delta_\tau\tilde{\xi}_h^k,\tilde{\boldsymbol{\phi}}_h)+\frac{\Gamma_e}{\epsilon_0\omega_{pe}^2}\left(\frac{\tilde{\xi}_h^{k+1}-\tilde{\xi}_h^{k-1}}{2\tau},\tilde{\boldsymbol{\phi}}_h\right)$$
$$+\frac{\omega_{e0}^2}{\epsilon_0\omega_{pe}^2}\left(\frac{\tilde{\xi}_h^{k+1}+\tilde{\xi}_h^{k-1}}{2},\tilde{\boldsymbol{\phi}}_h\right)-\left(\frac{\xi_h^{k+1}+\xi_h^{k-1}}{2},\tilde{\boldsymbol{\phi}}_h\right)$$
$$=\frac{1}{\tau\epsilon_0\omega_{pe}^2}((\delta_\tau\Pi_h\mathbf{P}^{k+1}-\delta_\tau\Pi_h\mathbf{P}^k)-(\mathbf{P}_t^{k+\frac{1}{2}}-\mathbf{P}_t^{k-\frac{1}{2}}),\tilde{\boldsymbol{\phi}}_h)$$
$$+\frac{\Gamma_e}{\epsilon_0\omega_{pe}^2}\left(\frac{\Pi_h\mathbf{P}^{k+1}-\Pi_h\mathbf{P}^{k-1}}{2\tau}-\frac{\mathbf{P}^{k+\frac{1}{2}}-\mathbf{P}^{k-\frac{1}{2}}}{\tau},\tilde{\boldsymbol{\phi}}_h\right)$$
$$+\frac{\omega_{e0}^2}{\epsilon_0\omega_{pe}^2}\left(\frac{\Pi_h\mathbf{P}^{k+1}+\Pi_h\mathbf{P}^{k-1}}{2}-\frac{1}{\tau}\int_{t_{k-\frac{1}{2}}}^{t_{k+\frac{1}{2}}}\mathbf{P}(s)ds,\tilde{\boldsymbol{\phi}}_h\right)$$
$$-\left(\frac{\Pi_h\mathbf{E}^{k+1}+\Pi_h\mathbf{E}^{k-1}}{2}-\frac{1}{\tau}\int_{t_{k-\frac{1}{2}}}^{t_{k+\frac{1}{2}}}\mathbf{E}(s)ds,\tilde{\boldsymbol{\phi}}_h\right), \quad (3.161)$$

$$(iv)\quad \frac{1}{\tau\mu_0\omega_{pm}^2}(\delta_\tau\tilde{\eta}_h^{k+1}-\delta_\tau\tilde{\eta}_h^k,\tilde{\boldsymbol{\psi}}_h)+\frac{\Gamma_m}{\mu_0\omega_{pm}^2}\left(\frac{\tilde{\eta}_h^{k+1}-\tilde{\eta}_h^{k-1}}{2\tau},\tilde{\boldsymbol{\psi}}_h\right)$$
$$+\frac{\omega_{m0}^2}{\mu_0\omega_{pm}^2}\left(\frac{\tilde{\eta}_h^{k+1}+\tilde{\eta}_h^{k-1}}{2},\tilde{\boldsymbol{\psi}}_h\right)-\left(\frac{\eta_h^{k+1}+\eta_h^{k-1}}{2},\tilde{\boldsymbol{\psi}}_h\right)$$
$$=\frac{1}{\tau\mu_0\omega_{pm}^2}((\delta_\tau\mathbf{M}^{k+1}-\delta_\tau\mathbf{M}^k)-(\mathbf{M}_t^{k+\frac{1}{2}}-\mathbf{M}_t^{k-\frac{1}{2}}),\tilde{\boldsymbol{\psi}}_h)$$
$$+\frac{\Gamma_m}{\mu_0\omega_{pm}^2}\left(\frac{\mathbf{M}^{k+1}-\mathbf{M}^{k-1}}{2\tau}-\frac{\mathbf{M}^{k+\frac{1}{2}}-\mathbf{M}^{k-\frac{1}{2}}}{\tau},\tilde{\boldsymbol{\psi}}_h\right)$$

$$+\frac{\omega_{m0}^2}{\epsilon_0\omega_{pe}^2}(\frac{\mathbf{M}^{k+1}+\mathbf{M}^{k-1}}{2}-\frac{1}{\tau}\int_{t_{k-\frac{1}{2}}}^{t_{k+\frac{1}{2}}}\mathbf{M}(s)ds,\tilde{\boldsymbol{\psi}}_h)$$

$$-(\frac{\mathbf{H}^{k+1}+\mathbf{H}^{k-1}}{2}-\frac{1}{\tau}\int_{t_{k-\frac{1}{2}}}^{t_{k+\frac{1}{2}}}\mathbf{H}(s)ds,\tilde{\boldsymbol{\psi}}_h). \tag{3.162}$$

Choosing $\phi_h=\tau(\xi_h^{k+1}+\xi_h^{k-1}), \psi_h=\tau(\eta_h^{k+1}+\eta_h^{k-1}), \tilde{\phi}_h=\tau(\delta_\tau\tilde{\xi}_h^{k+1}+\delta_\tau\tilde{\xi}_h^k), \tilde{\psi}_h=\tau(\delta_\tau\tilde{\eta}_h^{k+1}+\delta_\tau\tilde{\eta}_h^k)$ in the above error equations, and adding the resultants together, we obtain

$$\frac{\epsilon_0}{2}(||\xi_h^{k+1}||_0^2-||\xi_h^{k-1}||_0^2)+\frac{1}{\epsilon_0\omega_{pe}^2}(||\delta_\tau\tilde{\xi}_h^{k+1}||_0^2-||\delta_\tau\tilde{\xi}_h^k||_0^2)$$
$$+\frac{\omega_{e0}^2}{2\epsilon_0\omega_{pe}^2}(||\tilde{\xi}_h^{k+1}||_0^2-||\tilde{\xi}_h^{k-1}||_0^2)+\frac{\mu_0}{2}(||\eta_h^{k+1}||_0^2-||\eta_h^{k-1}||_0^2)$$
$$+\frac{1}{\mu_0\omega_{pm}^2}(||\delta_\tau\tilde{\eta}_h^{k+1}||_0^2-||\delta_\tau\tilde{\eta}_h^k||_0^2)+\frac{\omega_{m0}^2}{2\mu_0\omega_{pm}^2}(||\tilde{\eta}_h^{k+1}||_0^2-||\tilde{\eta}_h^{k-1}||_0^2)$$
$$\leq\sum_{i=1}^{12}Err_i, \tag{3.163}$$

where Err_i are those right hand side terms from (3.159) to (3.162).

The proof can be done by carefully estimating all Err_i. Details can be seen in the original paper [184]. □

3.6.3 Some Other Schemes

By introducing the induced electric and magnetic currents $\mathbf{J}=\frac{\partial\mathbf{P}}{\partial t}$ and $\mathbf{K}=\frac{\partial\mathbf{M}}{\partial t}$, respectively, we can rewrite the Lorentz model (3.123)–(3.126) as

$$\epsilon_0\frac{\partial\mathbf{E}}{\partial t}=\nabla\times\mathbf{H}-\mathbf{J} \tag{3.164}$$

$$\mu_0\frac{\partial\mathbf{H}}{\partial t}=-\nabla\times\mathbf{E}-\mathbf{K} \tag{3.165}$$

$$\frac{1}{\epsilon_0\omega_{pe}^2}\frac{\partial\mathbf{J}}{\partial t}+\frac{\Gamma_e}{\epsilon_0\omega_{pe}^2}\mathbf{J}=\mathbf{E}-\frac{\omega_{e0}^2}{\epsilon_0\omega_{pe}^2}\mathbf{P} \tag{3.166}$$

$$\frac{\omega_{e0}^2}{\epsilon_0\omega_{pe}^2}\frac{\partial\mathbf{P}}{\partial t}=\frac{\omega_{e0}^2}{\epsilon_0\omega_{pe}^2}\mathbf{J} \tag{3.167}$$

$$\frac{1}{\mu_0\omega_{pm}^2}\frac{\partial\mathbf{K}}{\partial t}+\frac{\Gamma_m}{\mu_0\omega_{pm}^2}\mathbf{K}=\mathbf{H}-\frac{\omega_{m0}^2}{\mu_0\omega_{pm}^2}\mathbf{M} \tag{3.168}$$

$$\frac{\omega_{m0}^2}{\mu_0\omega_{pm}^2}\frac{\partial \mathbf{M}}{\partial t} = \frac{\omega_{m0}^2}{\mu_0\omega_{pm}^2}\mathbf{K}. \tag{3.169}$$

Multiplying the above equations by $\mathbf{E}, \mathbf{H}, \mathbf{J}, \mathbf{P}, \mathbf{K}$ and $\mathbf{M}$, respectively, then integrating over Ω and summing up the resultants, we can easily obtain the following stability.

Lemma 3.25.

$$\begin{aligned}
&\epsilon_0||\mathbf{E}(t)||_0^2 + \mu_0||\mathbf{H}(t)||_0^2 + \frac{1}{\epsilon_0\omega_{pe}^2}||\mathbf{J}(t)||_0^2 \\
&\quad + \frac{1}{\mu_0\omega_{pm}^2}||\mathbf{K}(t)||_0^2 + \frac{\omega_{e0}^2}{\epsilon_0\omega_{pe}^2}||\mathbf{P}(t)||_0^2 + \frac{\omega_{m0}^2}{\mu_0\omega_{pm}^2}||\mathbf{M}(t)||_0^2 \\
&\le \epsilon_0||\mathbf{E}(0)||_0^2 + \mu_0||\mathbf{H}(0)||_0^2 + \frac{1}{\epsilon_0\omega_{pe}^2}||\mathbf{J}(0)||_0^2 \\
&\quad + \frac{1}{\mu_0\omega_{pm}^2}||\mathbf{K}(0)||_0^2 + \frac{\omega_{e0}^2}{\epsilon_0\omega_{pe}^2}||\mathbf{P}(0)||_0^2 + \frac{\omega_{m0}^2}{\mu_0\omega_{pm}^2}||\mathbf{M}(0)||_0^2.
\end{aligned}$$

We can construct a Crank-Nicolson scheme for solving the system (3.164)–(3.169): For $k = 1, 2, \cdots, N$, find $\mathbf{E}_h^k \in \mathbf{V}_h^0, \mathbf{J}_h^k, \mathbf{P}_h^k \in \mathbf{V}_h, \mathbf{H}_h^k, \mathbf{K}_h^k, \mathbf{M}_h^k \in \mathbf{U}_h$ such that

$$\epsilon_0(\delta_\tau \mathbf{E}_h^k, \boldsymbol{\phi}_h) - (\overline{\mathbf{H}}_h^k, \nabla\times\boldsymbol{\phi}_h) + (\overline{\mathbf{J}}_h^k, \boldsymbol{\phi}_h) = 0, \tag{3.170}$$

$$\mu_0(\delta_\tau \mathbf{H}_h^k, \boldsymbol{\psi}_h) + (\nabla\times\overline{\mathbf{E}}_h^k, \boldsymbol{\psi}_h) + (\overline{\mathbf{K}}_h^k, \boldsymbol{\psi}_h) = 0, \tag{3.171}$$

$$\frac{1}{\epsilon_0\omega_{pe}^2}[(\delta_\tau \mathbf{J}_h^k, \boldsymbol{\phi}_{1h}) + \Gamma_e(\overline{\mathbf{J}}_h^k, \boldsymbol{\phi}_{1h}) + \omega_{e0}^2(\overline{\mathbf{P}}_h^k, \boldsymbol{\phi}_{1h})] = (\overline{\mathbf{E}}_h^k, \boldsymbol{\phi}_{1h}), \tag{3.172}$$

$$\frac{\omega_{e0}^2}{\epsilon_0\omega_{pe}^2}(\delta_\tau \mathbf{P}_h^k, \boldsymbol{\phi}_{2h}) - \frac{\omega_{e0}^2}{\epsilon_0\omega_{pe}^2}(\overline{\mathbf{J}}_h^k, \boldsymbol{\phi}_{2h}) = 0, \tag{3.173}$$

$$\frac{1}{\mu_0\omega_{pm}^2}[(\delta_\tau \mathbf{K}_h^k, \boldsymbol{\psi}_{1h}) + \Gamma_m(\overline{\mathbf{K}}_h^k, \boldsymbol{\psi}_{1h}) + \omega_{m0}^2(\overline{\mathbf{M}}_h^k, \boldsymbol{\psi}_{1h})] = (\overline{\mathbf{H}}_h^k, \boldsymbol{\psi}_{1h}), \tag{3.174}$$

$$\frac{\omega_{m0}^2}{\mu_0\omega_{pm}^2}(\delta_\tau \mathbf{M}_h^k, \boldsymbol{\psi}_{2h}) - \frac{\omega_{m0}^2}{\mu_0\omega_{pm}^2}(\overline{\mathbf{K}}_h^k, \boldsymbol{\psi}_{2h}) = 0, \tag{3.175}$$

hold true for any $\boldsymbol{\phi}_h \in \mathbf{V}_h^0, \boldsymbol{\psi}_h, \boldsymbol{\psi}_{1h}, \boldsymbol{\psi}_{2h} \in \mathbf{U}_h, \boldsymbol{\phi}_{1h}, \boldsymbol{\phi}_{2h} \in \mathbf{V}_h$, and are subject to the initial approximations

$$\begin{aligned}
&\mathbf{E}_h^0(\mathbf{x}) = \Pi_h\mathbf{E}_0(\mathbf{x}), \ \ \mathbf{J}_h^0(\mathbf{x}) = \Pi_h\mathbf{J}_0(\mathbf{x}), \ \ \mathbf{P}_h^0(\mathbf{x}) = \Pi_h\mathbf{P}_0(\mathbf{x}), \\
&\mathbf{H}_h^0(\mathbf{x}) = P_h\mathbf{H}_0(\mathbf{x}), \ \ \mathbf{K}_h^0(\mathbf{x}) = P_h\mathbf{K}_0(\mathbf{x}), \ \ \mathbf{M}_h^0(\mathbf{x}) = P_h\mathbf{M}_0(\mathbf{x}).
\end{aligned}$$

Choosing $\boldsymbol{\phi}_h = \overline{\mathbf{E}}_h^k, \boldsymbol{\psi}_h = \overline{\mathbf{H}}_h^k, \boldsymbol{\phi}_{1h} = \overline{\mathbf{J}}_h^k, \boldsymbol{\phi}_{2h} = \overline{\mathbf{P}}_h^k, \boldsymbol{\psi}_{1h} = \overline{\mathbf{K}}_h^k, \boldsymbol{\psi}_{2h} = \overline{\mathbf{M}}_h^k$ in (3.170)–(3.175), respectively, and adding the resultants together, we can obtain the following discrete stability in exactly the same form as in the continuous case.

Lemma 3.26. *For any $k \geq 1$, we have*

$$\begin{aligned}
&\epsilon_0||\mathbf{E}_h^k||_0^2 + \mu_0||\mathbf{H}_h^k||_0^2 + \frac{1}{\epsilon_0\omega_{pe}^2}||\mathbf{J}_h^k||_0^2 \\
&\quad + \frac{1}{\mu_0\omega_{pm}^2}||\mathbf{K}_h^k||_0^2 + \frac{\omega_{e0}^2}{\epsilon_0\omega_{pe}^2}||\mathbf{P}_h^k||_0^2 + \frac{\omega_{m0}^2}{\mu_0\omega_{pm}^2}||\mathbf{M}_h^k||_0^2 \\
&\leq \epsilon_0||\mathbf{E}_h^0||_0^2 + \mu_0||\mathbf{H}_h^0||_0^2 + \frac{1}{\epsilon_0\omega_{pe}^2}||\mathbf{J}_h^0||_0^2 \\
&\quad + \frac{1}{\mu_0\omega_{pm}^2}||\mathbf{K}_h^0||_0^2 + \frac{\omega_{e0}^2}{\epsilon_0\omega_{pe}^2}||\mathbf{P}_h^0||_0^2 + \frac{\omega_{m0}^2}{\mu_0\omega_{pm}^2}||\mathbf{M}_h^0||_0^2.
\end{aligned}$$

We want to show that practical implementation of (3.170)–(3.175) is actually not that scary. Solving (3.173), we obtain

$$\mathbf{P}_h^k = \mathbf{P}_h^{k-1} + \frac{\tau}{2}(\mathbf{J}_h^k + \mathbf{J}_h^{k-1}). \tag{3.176}$$

Substituting (3.176) into (3.172) and simplifying the result, we have

$$\beta\mathbf{J}_h^k = (1 - \frac{\tau\Gamma_e}{2} - \frac{\tau^2\omega_{e0}^2}{4})\mathbf{J}_h^{k-1} - \tau\omega_{e0}^2\mathbf{P}_h^{k-1} + \frac{\tau\epsilon_0\omega_{pe}^2}{2}(\mathbf{E}_h^k + \mathbf{E}_h^{k-1}), \tag{3.177}$$

where we denote $\beta = 1 + \frac{\tau\Gamma_e}{2} + \frac{\tau^2\omega_{e0}^2}{4}$.

Then substituting (3.177) into (3.170) and simplifying the result, we have

$$\begin{aligned}
&\epsilon_0(1 + \frac{\tau^2\omega_{pe}^2}{4\beta})(\mathbf{E}_h^k, \boldsymbol{\phi}_h) - \frac{\tau}{2}(\mathbf{H}_h^k, \nabla\times\boldsymbol{\phi}_h) = \epsilon_0(1 - \frac{\tau^2\omega_{pe}^2}{4\beta})(\mathbf{E}_h^{k-1}, \boldsymbol{\phi}_h) \\
&+ \frac{\tau}{2}(\mathbf{H}_h^{k-1}, \nabla\times\boldsymbol{\phi}_h) - \frac{1}{\beta}(\tau\mathbf{J}_h^{k-1} - \frac{\tau^2\omega_{e0}^2}{2}\mathbf{P}_h^{k-1}, \boldsymbol{\phi}_h).
\end{aligned} \tag{3.178}$$

Similarly, from (3.173) to (3.175), we can obtain

$$\mathbf{M}_h^k = \mathbf{M}_h^{k-1} + \frac{\tau}{2}(\mathbf{K}_h^k + \mathbf{K}_h^{k-1}), \tag{3.179}$$

$$\tilde{\beta}\mathbf{K}_h^k = (1 - \frac{\tau\Gamma_m}{2} - \frac{\tau^2\omega_{m0}^2}{4})\mathbf{K}_h^{k-1} - \tau\omega_{m0}^2\mathbf{M}_h^{k-1} + \frac{\tau\mu_0\omega_{me}^2}{2}(\mathbf{H}_h^k + \mathbf{H}_h^{k-1}), \tag{3.180}$$

$$\mu_0(1+\frac{\tau^2\omega_{me}^2}{4\tilde{\beta}})(\mathbf{H}_h^k,\boldsymbol{\psi}_h)+\frac{\tau}{2}(\nabla\times\mathbf{E}_h^k,\boldsymbol{\psi}_h)=\mu_0(1-\frac{\tau^2\omega_{me}^2}{4\tilde{\beta}})(\mathbf{H}_h^{k-1},\boldsymbol{\psi}_h)$$
$$-\frac{\tau}{2}(\nabla\times\mathbf{E}_h^{k-1},\boldsymbol{\psi}_h)-\frac{1}{\tilde{\beta}}(\tau\mathbf{K}_h^{k-1}-\frac{\tau^2\omega_{m0}^2}{2}\mathbf{M}_h^{k-1},\boldsymbol{\psi}_h). \tag{3.181}$$

Hence, at each time step, we first solve a system formed by (3.178) and (3.181) for $\mathbf{E}_h^k$ and $\mathbf{H}_h^k$; then use (3.177) and (3.180) to update $\mathbf{J}_h^k$ and $\mathbf{K}_h^k$; finally, use (3.176) and (3.179) to update $\mathbf{P}_h^k$ and $\mathbf{M}_h^k$.

With proper regularity assumption, we can similarly prove the following optimal error estimate:

$$||\mathbf{E}^n-\mathbf{E}_h^n||_0+||\mathbf{P}^n-\mathbf{P}_h^n||_0+||\mathbf{H}^n-\mathbf{H}_h^n||_0+||\mathbf{M}^n-\mathbf{M}_h^n||_0$$
$$+||\mathbf{J}^n-\mathbf{J}_h^n||_0+||\mathbf{K}^n-\mathbf{K}_h^n||_0\le C(h^l+\tau^2). \tag{3.182}$$

Finally, we like to mention that leap-frog type schemes can be constructed for solving the system (3.164)–(3.169). For example, one leap-frog scheme is given as following: For $k\ge 0$, find $\mathbf{J}_h^{k+\frac{1}{2}},\mathbf{P}_h^{k+1}\in\mathbf{V}_h,\mathbf{E}_h^k\in\mathbf{V}_h^0,\mathbf{K}_h^{k+\frac{1}{2}},\mathbf{M}_h^{k+1},\mathbf{H}_h^{k+\frac{3}{2}}\in\mathbf{U}_h$ such that

$$\frac{1}{\epsilon_0\omega_{pe}^2}\frac{\mathbf{J}_h^{k+\frac{1}{2}}-\mathbf{J}_h^{k-\frac{1}{2}}}{\tau}+\frac{\Gamma_e}{2\epsilon_0\omega_{pe}^2}(\mathbf{J}_h^{k+\frac{1}{2}}+\mathbf{J}_h^{k-\frac{1}{2}})=\mathbf{E}_h^k-\frac{\omega_{e0}^2}{\epsilon_0\omega_{pe}^2}\mathbf{P}_h^k,$$
$$\frac{\omega_{e0}^2}{\epsilon_0\omega_{pe}^2}\frac{\mathbf{P}_h^{k+1}-\mathbf{P}_h^k}{\tau}=\frac{\omega_{e0}^2}{\epsilon_0\omega_{pe}^2}\mathbf{J}_h^{k+\frac{1}{2}},$$
$$\epsilon_0(\frac{\mathbf{E}_h^{k+1}-\mathbf{E}_h^k}{\tau},\boldsymbol{\phi}_h)-(\mathbf{H}_h^{k+\frac{1}{2}},\nabla\times\boldsymbol{\phi}_h)+(\mathbf{J}_h^{k+\frac{1}{2}},\boldsymbol{\phi}_h)=0,$$
$$\frac{1}{\mu_0\omega_{pm}^2}\frac{\mathbf{K}_h^{k+\frac{1}{2}}-\mathbf{K}_h^{k-\frac{1}{2}}}{\tau}+\frac{\Gamma_m}{2\mu_0\omega_{pm}^2}(\mathbf{K}_h^{k+\frac{1}{2}}+\mathbf{K}_h^{k-\frac{1}{2}})$$
$$=\frac{1}{2}(\mathbf{H}_h^{k+\frac{1}{2}}+\mathbf{H}_h^{k-\frac{1}{2}})-\frac{\omega_{m0}^2}{\mu_0\omega_{pm}^2}\mathbf{M}_h^k,$$
$$\frac{\omega_{m0}^2}{\mu_0\omega_{pm}^2}\frac{\mathbf{M}_h^{k+1}-\mathbf{M}_h^k}{\tau}=\frac{\omega_{m0}^2}{\mu_0\omega_{pm}^2}\mathbf{K}_h^{k+\frac{1}{2}},$$
$$\mu_0(\frac{\mathbf{H}_h^{k+\frac{3}{2}}-\mathbf{H}_h^{k+\frac{1}{2}}}{\tau},\boldsymbol{\psi}_h)+(\nabla\times\mathbf{E}_h^{k+1},\boldsymbol{\psi}_h)+(\mathbf{K}_h^{k+\frac{1}{2}},\boldsymbol{\psi}_h)=0,$$

hold true for any $\boldsymbol{\phi}_h\in\mathbf{V}_h^0$ and $\boldsymbol{\psi}_h\in\mathbf{U}_h$. Readers are encouraged to carry out the stability and error analysis by following the technique developed in Sect. 3.5 for the Drude model.

3.7 Extensions to the Drude-Lorentz Model

3.7.1 The Well-Posedness

Recall from Chap. 1 that the governing equations for the Drude-Lorentz model are:

$$\epsilon_0 \frac{\partial \mathbf{E}}{\partial t} = \nabla \times \mathbf{H} - \mathbf{J}, \tag{3.183}$$

$$\mu_0 \frac{\partial \mathbf{H}}{\partial t} = -\nabla \times \mathbf{E} - \mathbf{K}, \tag{3.184}$$

$$\frac{1}{\mu_0 \omega_0^2 F} \frac{\partial \mathbf{K}}{\partial t} + \frac{\gamma}{\mu_0 \omega_0^2 F} \mathbf{K} + \frac{1}{\mu_0 F} \mathbf{M} = \mathbf{H}, \tag{3.185}$$

$$\frac{1}{\mu_0 F} \frac{\partial \mathbf{M}}{\partial t} = \frac{1}{\mu_0 F} \mathbf{K}, \tag{3.186}$$

$$\frac{1}{\epsilon_0 \omega_p^2} \frac{\partial \mathbf{J}}{\partial t} + \frac{\nu}{\epsilon_0 \omega_p^2} \mathbf{J} = \mathbf{E}. \tag{3.187}$$

To complete the problem, we assume that the perfect conducting boundary condition (3.59) is imposed, and the initial conditions

$$\mathbf{E}(\mathbf{x}, 0) = \mathbf{E}_0(\mathbf{x}), \quad \mathbf{H}(\mathbf{x}, 0) = \mathbf{H}_0(\mathbf{x}), \tag{3.188}$$

$$\mathbf{K}(\mathbf{x}, 0) = \mathbf{K}_0(\mathbf{x}), \quad \mathbf{M}(\mathbf{x}, 0) = \mathbf{M}_0(\mathbf{x}), \quad \mathbf{J}(\mathbf{x}, 0) = \mathbf{J}_0(\mathbf{x}), \tag{3.189}$$

hold true. Here $\mathbf{E}_0, \mathbf{H}_0, \mathbf{K}_0, \mathbf{M}_0$ and $\mathbf{J}_0$ are some given functions.

First, we have the following stability for our model problem (3.183)–(3.187).

Lemma 3.27. *For the solution* $(\mathbf{E}, \mathbf{H}, \mathbf{K}, \mathbf{M}, \mathbf{J})$ *of problem (3.183)–(3.187) subject to boundary condition (3.59) and initial conditions (3.188) and (3.189), the following stability holds true:*

$$\begin{aligned} &\epsilon_0 ||\mathbf{E}(t)||_0^2 + \mu_0 ||\mathbf{H}(t)||_0^2 + \frac{1}{\mu_0 \omega_0^2 F} ||\mathbf{K}(t)||_0^2 + \frac{1}{\mu_0 F} ||\mathbf{M}(t)||_0^2 + \frac{1}{\epsilon_0 \omega_p^2} ||\mathbf{J}(t)||_0^2 \\ &\leq \epsilon_0 ||\mathbf{E}_0||_0^2 + \mu_0 ||\mathbf{H}_0||_0^2 + \frac{1}{\mu_0 \omega_0^2 F} ||\mathbf{K}_0||_0^2 + \frac{1}{\mu_0 F} ||\mathbf{M}_0||_0^2 + \frac{1}{\epsilon_0 \omega_p^2} ||\mathbf{J}_0||_0^2. \end{aligned} \tag{3.190}$$

Proof. Note that the problem (3.183)–(3.187) can be rewritten as

$$\frac{\partial}{\partial t} \mathscr{A} \mathbf{u}(t) = (\mathscr{B} + \mathscr{C}) \mathbf{u}(t), \tag{3.191}$$

where vector $\mathbf{u}(t) = (\mathbf{E}, \mathbf{H}, \mathbf{K}, \mathbf{M}, \mathbf{J})'$, matrices

$$\mathscr{A} = \mathrm{diag}(\epsilon_0 I_3, \mu_0 I_3, \frac{1}{\mu_0\omega_0^2 F} I_3, \frac{1}{\mu_0 F} I_3, \frac{1}{\epsilon_0\omega_p^2} I_3),$$

$$\mathscr{C} = \mathrm{diag}(0_3, 0_3, -\frac{\gamma}{\mu_0\omega_0^2 F} I_3, 0_3, -\frac{\nu}{\epsilon_0\omega_p^2} I_3),$$

and

$$\mathscr{B} = \begin{pmatrix} 0_3 & \nabla\times & 0_3 & 0_3 & -I_3 \\ -\nabla\times & 0_3 & -I_3 & 0_3 & 0_3 \\ 0_3 & I_3 & 0_3 & -\frac{1}{\mu_0 F} I_3 & 0_3 \\ 0_3 & 0_3 & \frac{1}{\mu_0 F} I_3 & 0_3 & 0_3 \\ I_3 & 0_3 & 0_3 & 0_3 & 0_3 \end{pmatrix}.$$

Here I_3 denotes a 3×3 identity matrix, and 0_3 denotes a 3×3 zero matrix.

To prove the stability, left-multiplying (3.191) by $\mathbf{u}'$, then integrating over Ω, and using the property $\mathbf{u}'\mathscr{B}\mathbf{u} = 0$, we obtain

$$\frac{d}{dt}(\mathbf{u}'\mathscr{A}\mathbf{u}) = \mathbf{u}'\mathscr{C}\mathbf{u} \le 0,$$

integrating which with respect to t leads to the stability (3.190). □

Remark 3.4. The stability (3.190) can be proved directly as we did previously for the Drude model and Lorentz model. Multiplying (3.183)–(3.187) by $\mathbf{E}, \mathbf{H}, \mathbf{K}, \mathbf{M}, \mathbf{J}$ and integrating over Ω, then adding the resultants together, we obtain

$$\begin{aligned} &\frac{1}{2}\frac{d}{dt}[\epsilon_0||\mathbf{E}(t)||_0^2 + \mu_0||\mathbf{H}(t)||_0^2 + \frac{1}{\mu_0\omega_0^2 F}||\mathbf{K}(t)||_0^2 + \frac{1}{\mu_0 F}||\mathbf{M}(t)||_0^2 \\ &+ \frac{1}{\epsilon_0\omega_p^2}||\mathbf{J}(t)||_0^2] \\ &+ \frac{\nu}{\epsilon_0\omega_p^2}||\mathbf{J}(t)||_0^2 + \frac{\gamma}{\mu_0\omega_0^2 F}||\mathbf{K}(t)||_0^2 = 0, \end{aligned}$$

integrating which from 0 to t leads to (3.190). Rewriting the last three governing equations (3.185)–(3.187) in this way leads to this elegant proof.

Now, let us prove the existence for the model problem (3.183)–(3.187).

Theorem 3.15. *The problem (3.183)–(3.187) has a unique solution* $(\mathbf{E}, \mathbf{H})$ *in* $H(curl; \Omega) \oplus H(curl; \Omega)$.

Proof. Though the technique developed in Theorems 3.8 and 3.13 can still be used here, we apply a different technique developed in [122].

From ordinary differential equation theory, it is easy to see that the solutions of (1.29) and (1.30) with zero initial conditions can be expressed as

$$\mathbf{P}(\mathbf{x},t) = \frac{\epsilon_0\omega_p^2}{\nu}\int_0^t (1-e^{-\nu(t-s)})\mathbf{E}(\mathbf{x},s)ds, \tag{3.192}$$

and

$$\mathbf{M}(\mathbf{x},t) = \mu_0 F\omega_0^2\int_0^t g(t-s)\mathbf{H}(\mathbf{x},s)ds, \tag{3.193}$$

respectively. Here the kernel $g(t) = \frac{1}{\alpha}e^{-\frac{\gamma}{2}t}\sin(\alpha t)$, where $\alpha = \sqrt{\omega_0^2 - (\frac{\gamma}{2})^2}$.

Using the definition $\mathbf{J} = \frac{\partial \mathbf{P}}{\partial t}$ and $\mathbf{K} = \frac{\partial \mathbf{M}}{\partial t}$ introduced in Chap. 1, then substituting (3.192) and (3.193) into (3.183) and (3.184), respectively, we can rewrite (3.183) and (3.184) as:

$$\frac{d}{dt}(A\mathscr{E} + K * \mathscr{E}) = L\mathscr{E} + \mathscr{F}, \tag{3.194}$$

where we denote $\mathscr{E} = (\mathbf{E},\mathbf{H})'$, $*$ for the convolution product, $\mathscr{F}$ for source terms obtained by transforming a problem with non-zero initial conditions to a problem with zero initial conditions. Moreover

$$A = \begin{pmatrix} \epsilon_0 I_3 & 0_3 \\ 0_3 & \mu_0 I_3 \end{pmatrix}, \quad K = \begin{pmatrix} \epsilon_1 I_3 & 0_3 \\ 0_3 & \mu_1 I_3 \end{pmatrix}, \quad L = \begin{pmatrix} 0_3 & \nabla\times \\ -\nabla\times & 0_3 \end{pmatrix},$$

where $\epsilon_1 = \frac{\epsilon_0\omega_p^2}{\nu}(u(t) - e^{-\nu t})$, $\mu_1 = \mu_0 F\omega_0^2 g(t)$, and $u(t)$ denotes the unit step function.

Note that problem (3.194) is a special case of Problem I of [122], whose existence and uniqueness is guaranteed by Theorem 3.1 of [122]. □

3.7.2 Two Numerical Schemes

In this section, we present two fully-discrete schemes developed in [195] for solving the problem (3.183)–(3.187).

First, let us start with a Crank-Nicolson type scheme: For $k = 1, 2, \cdots, N$, find $\mathbf{E}_h^k \in \mathbf{V}_h^0, \mathbf{J}_h^k \in \mathbf{V}_h, \mathbf{H}_h^k, \mathbf{K}_h^k, \mathbf{M}_h^k \in \mathbf{U}_h$ such that

$$\epsilon_0(\delta_\tau \mathbf{E}_h^k, \boldsymbol{\phi}_h) - (\overline{\mathbf{H}}_h^k, \nabla\times\boldsymbol{\phi}_h) + (\overline{\mathbf{J}}_h^k, \boldsymbol{\phi}_h) = 0, \tag{3.195}$$

$$\mu_0(\delta_\tau \mathbf{H}_h^k, \boldsymbol{\psi}_h) + (\nabla\times\overline{\mathbf{E}}_h^k, \boldsymbol{\psi}_h) + (\overline{\mathbf{K}}_h^k, \boldsymbol{\psi}_h) = 0, \tag{3.196}$$

$$\frac{1}{\mu_0\omega_0^2 F}(\delta_\tau \mathbf{K}_h^k, \tilde{\boldsymbol{\psi}}_{1h}) + \frac{\gamma}{\mu_0\omega_0^2 F}(\overline{\mathbf{K}}_h^k, \tilde{\boldsymbol{\psi}}_{1h}) + \frac{1}{\mu_0 F}(\overline{\mathbf{M}}_h^k, \tilde{\boldsymbol{\psi}}_{1h}) = (\overline{\mathbf{H}}_h^k, \tilde{\boldsymbol{\psi}}_{1h}),$$

$$\frac{1}{\mu_0 F}(\delta_\tau \mathbf{M}_h^k, \tilde{\boldsymbol{\psi}}_{2h}) = \frac{1}{\mu_0 F}(\overline{\mathbf{K}}_h^k, \tilde{\boldsymbol{\psi}}_{2h}), \tag{3.197}$$

$$\frac{1}{\epsilon_0\omega_p^2}(\delta_\tau \mathbf{J}_h^k, \tilde{\boldsymbol{\phi}}_h) + \frac{\nu}{\epsilon_0\omega_p^2}(\overline{\mathbf{J}}_h^k, \tilde{\boldsymbol{\phi}}_h) = (\overline{\mathbf{E}}_h^k, \tilde{\boldsymbol{\phi}}_h), \tag{3.198}$$

hold true for any $\boldsymbol{\phi}_h \in \mathbf{V}_h^0, \boldsymbol{\psi}_h, \tilde{\boldsymbol{\psi}}_{1h}, \tilde{\boldsymbol{\psi}}_{2h} \in \mathbf{U}_h, \tilde{\boldsymbol{\phi}}_h \in \mathbf{V}_h$, and are subject to the initial approximations

$$\mathbf{E}_h^0(\mathbf{x}) = \Pi_h \mathbf{E}_0(\mathbf{x}), \;\; \mathbf{H}_h^0(\mathbf{x}) = P_h \mathbf{H}_0(\mathbf{x}),$$
$$\mathbf{K}_h^0(\mathbf{x}) = P_h \mathbf{K}_0(\mathbf{x}), \; \mathbf{M}_h^0(\mathbf{x}) = P_h \mathbf{M}_0(\mathbf{x}), \; \mathbf{J}_h^0(\mathbf{x}) = \Pi_h \mathbf{J}_0(\mathbf{x}).$$

Below we show that the scheme (3.195)–(3.198) satisfies a discrete stability, which has exactly the same form as the continuous case proved in Lemma 3.27.

Lemma 3.28. *For any $k \geq 1$, we have*

$$\begin{aligned}
&\epsilon_0 ||\mathbf{E}_h^k||_0^2 + \mu_0 ||\mathbf{H}_h^k||_0^2 + \frac{1}{\epsilon_0 \omega_p^2} ||\mathbf{J}_h^k||_0^2 + \frac{1}{\mu_0 \omega_0^2 F} ||\mathbf{K}_h^k||_0^2 + \frac{1}{\mu_0 F} ||\mathbf{M}_h^k||_0^2 \\
\leq\, &\epsilon_0 ||\mathbf{E}_h^0||_0^2 + \mu_0 ||\mathbf{H}_h^0||_0^2 + \frac{1}{\epsilon_0 \omega_p^2} ||\mathbf{J}_h^0||_0^2 + \frac{1}{\mu_0 \omega_0^2 F} ||\mathbf{K}_h^0||_0^2 + \frac{1}{\mu_0 F} ||\mathbf{M}_h^0||_0^2.
\end{aligned}$$

Proof. Choosing $\boldsymbol{\phi}_h = \tau \overline{\mathbf{E}}_h^k, \boldsymbol{\psi}_h = \tau \overline{\mathbf{H}}_h^k, \tilde{\boldsymbol{\psi}}_{1h} = \tau \overline{\mathbf{K}}_h^k, \tilde{\boldsymbol{\psi}}_{2h} = \tau \overline{\mathbf{M}}_h^k, \tilde{\boldsymbol{\phi}}_h = \tau \overline{\mathbf{J}}_h^k$ in (3.195)–(3.198), respectively, then adding the resultants together, we have

$$\begin{aligned}
&\frac{\epsilon_0}{2}(||\mathbf{E}_h^k||_0^2 - ||\mathbf{E}_h^{k-1}||_0^2) + \frac{\mu_0}{2}(||\mathbf{H}_h^k||_0^2 - ||\mathbf{H}_h^{k-1}||_0^2) + \frac{1}{2\epsilon_0 \omega_p^2}(||\mathbf{J}_h^k||_0^2 - ||\mathbf{J}_h^{k-1}||_0^2) \\
&+ \frac{\tau \nu}{\epsilon_0 \omega_p^2} ||\overline{\mathbf{J}}_h^k||_0^2 + \frac{1}{2\mu_0 \omega_0^2 F}(||\mathbf{K}_h^k||_0^2 - ||\mathbf{K}_h^{k-1}||_0^2) \\
&+ \frac{\tau \gamma}{\mu_0 \omega_0^2 F} ||\overline{\mathbf{K}}_h^k||_0^2 + \frac{1}{2\mu_0 F}(||\mathbf{M}_h^k||_0^2 - ||\mathbf{M}_h^{k-1}||_0^2) = 0,
\end{aligned}$$

which easily concludes the proof. □

For the Crank-Nicolson scheme (3.195)–(3.198), the following optimal error estimate can be proved similarly to Theorem 3.10. Details can be found in the original paper [195].

Theorem 3.16. *Let $(\mathbf{E}^m, \mathbf{H}^m, \mathbf{K}^m, \mathbf{M}^m, \mathbf{J}^m)$ and $(\mathbf{E}_h^m, \mathbf{H}_h^m, \mathbf{K}_h^m, \mathbf{M}_h^m, \mathbf{J}_h^m)$ be the analytic and numerical solutions of (3.183)–(3.187) and (3.195)–(3.198) at time t_m, respectively. Under proper regularity assumptions, there exists a constant $C > 0$ independent of mesh size h and time step τ, such that*

$$\begin{aligned}
&\sqrt{\epsilon_0} ||\mathbf{E}^n - \mathbf{E}_h^n||_0 + \sqrt{\mu_0} ||\mathbf{H}^n - \mathbf{H}_h^n||_0 + \frac{1}{\sqrt{\epsilon_0 \omega_p^2}} ||\mathbf{J}^n - \mathbf{J}_h^n||_0 \\
&+ \frac{1}{\sqrt{\mu_0 \omega_0^2 F}} ||\mathbf{K}^n - \mathbf{K}_h^n||_0 + \frac{1}{\sqrt{\mu_0 F}} ||\mathbf{M}^n - \mathbf{M}_h^n||_0 \leq C(h^l + \tau^2),
\end{aligned}$$

where $l \geq 1$ is the order of the basis functions in spaces $\mathbf{U}_h$ and $\mathbf{V}_h$.

Note that the Crank-Nicolson scheme (3.195)–(3.198) has a non-symmetric linear system of as many as 15 unknown functions (five 3-D unknown variables), which results a very large-scale system even for linear edge elements. Hence directly solving the coupled system is quite challenging. In this aspect the leap-frog scheme developed below (cf. (3.199)–(3.202)) is more practical and appealing, since one unknown variable is solved at each step. Of course, the leap-frog scheme has to obey the CFL time step constraint.

A leap-frog type scheme for solving (3.183)–(3.187) is proposed in [195]: For $k = 1, 2, \cdots$, find $\mathbf{E}_h^k \in \mathbf{V}_h^0, \mathbf{J}_h^{k+\frac{1}{2}} \in \mathbf{V}_h, \mathbf{H}_h^{k+\frac{1}{2}}, \mathbf{K}_h^k, \mathbf{M}_h^{k+\frac{1}{2}} \in \mathbf{U}_h$ such that

$$\epsilon_0(\frac{\mathbf{E}_h^k - \mathbf{E}_h^{k-1}}{\tau}, \boldsymbol{\phi}_h) - (\mathbf{H}_h^{k-\frac{1}{2}}, \nabla \times \boldsymbol{\phi}_h) + (\mathbf{J}_h^{k-\frac{1}{2}}, \boldsymbol{\phi}_h) = 0, \tag{3.199}$$

$$\mu_0(\frac{\mathbf{H}_h^{k+\frac{1}{2}} - \mathbf{H}_h^{k-\frac{1}{2}}}{\tau}, \boldsymbol{\psi}_h) + (\nabla \times \mathbf{E}_h^k, \boldsymbol{\psi}_h) + (\mathbf{K}_h^k, \boldsymbol{\psi}_h) = 0, \tag{3.200}$$

$$\begin{aligned}&\frac{1}{\mu_0\omega_0^2 F}(\frac{\mathbf{K}_h^k - \mathbf{K}_h^{k-1}}{\tau}, \tilde{\boldsymbol{\psi}}_{1h}) + \frac{\gamma}{\mu_0\omega_0^2 F}(\frac{\mathbf{K}_h^k + \mathbf{K}_h^{k-1}}{2}, \tilde{\boldsymbol{\psi}}_{1h}) + \frac{1}{\mu_0 F}(\mathbf{M}_h^{k-\frac{1}{2}}, \tilde{\boldsymbol{\psi}}_{1h})\\ &\quad = (\mathbf{H}_h^{k-\frac{1}{2}}, \tilde{\boldsymbol{\psi}}_{1h}),\end{aligned}$$

$$\frac{1}{\mu_0 F}(\frac{\mathbf{M}_h^{k+\frac{1}{2}} - \mathbf{M}_h^{k-\frac{1}{2}}}{\tau}, \tilde{\boldsymbol{\psi}}_{2h}) = \frac{1}{\mu_0 F}(\mathbf{K}_h^k, \tilde{\boldsymbol{\psi}}_{2h}), \tag{3.201}$$

$$\frac{1}{\epsilon_0\omega_p^2}(\frac{\mathbf{J}_h^{k+\frac{1}{2}} - \mathbf{J}_h^{k-\frac{1}{2}}}{\tau}, \tilde{\boldsymbol{\phi}}_h) + \frac{\nu}{\epsilon_0\omega_p^2}(\frac{\mathbf{J}_h^{k+\frac{1}{2}} + \mathbf{J}_h^{k-\frac{1}{2}}}{2}, \tilde{\boldsymbol{\phi}}_h) = (\mathbf{E}_h^k, \tilde{\boldsymbol{\phi}}_h), \tag{3.202}$$

hold true for any $\boldsymbol{\phi}_h \in \mathbf{V}_h^0, \boldsymbol{\psi}_h, \tilde{\boldsymbol{\psi}}_{1h}, \tilde{\boldsymbol{\psi}}_{2h} \in \mathbf{U}_h, \tilde{\boldsymbol{\phi}}_h \in \mathbf{V}_h$, and are subject to the initial approximations

$$\mathbf{E}_h^0(\mathbf{x}) = \Pi_h \mathbf{E}_0(\mathbf{x}), \quad \mathbf{K}_h^0(\mathbf{x}) = P_h \mathbf{K}_0(\mathbf{x}), \tag{3.203}$$

$$\mathbf{H}_h^{\frac{1}{2}}(\mathbf{x}) = P_h[\mathbf{H}_0(\mathbf{x}) - \frac{\tau}{2}\mu_0^{-1}(\nabla \times \mathbf{E}_0(\mathbf{x}) + \mathbf{K}_0(\mathbf{x}))],$$

$$\mathbf{M}_h^{\frac{1}{2}}(\mathbf{x}) = P_h[\mathbf{M}_0(\mathbf{x}) + \frac{\tau}{2}\mathbf{K}_0(\mathbf{x})],$$

$$\mathbf{J}_h^{\frac{1}{2}}(\mathbf{x}) = \Pi_h[\mathbf{J}_0(\mathbf{x}) + \frac{\tau}{2}(\epsilon_0\omega_p^2 \mathbf{E}_0(\mathbf{x}) - \nu \mathbf{J}_0(\mathbf{x}))]. \tag{3.204}$$

The following discrete stability for the leap-frog scheme (3.199)–(3.202) can be proved similarly to Theorem 3.11.

Theorem 3.17. *Under the time step constraint*

$$\tau = \min\{\frac{1}{2\omega_0\sqrt{F}}, \frac{1}{2\omega_0}, \frac{1}{2\omega_p}, \frac{h}{2C_v C_{inv}}\}, \tag{3.205}$$

where C_v and C_{inv} are defined in Theorem 3.11. Then for any $k \geq 1$, we have

$$\epsilon_0||\mathbf{E}_h^k||_0^2 + \mu_0||\mathbf{H}_h^{k+\frac{1}{2}}||_0^2 + \frac{1}{\epsilon_0\omega_p^2}||\mathbf{J}_h^{k+\frac{1}{2}}||_0^2 + \frac{1}{\mu_0\omega_0^2 F}||\mathbf{K}_h^k||_0^2 + \frac{1}{\mu_0 F}||\mathbf{M}_h^{k+\frac{1}{2}}||_0^2$$
$$\leq C[\epsilon_0||\mathbf{E}_h^0||_0^2 + \mu_0||\mathbf{H}_h^{\frac{1}{2}}||_0^2 + \frac{1}{\epsilon_0\omega_p^2}||\mathbf{J}_h^{\frac{1}{2}}||_0^2 + \frac{1}{\mu_0\omega_0^2 F}||\mathbf{K}_h^0||_0^2 + \frac{1}{\mu_0 F}||\mathbf{M}_h^{\frac{1}{2}}||_0^2],$$

where $C > 1$ is independent of h and τ.

Similarly to Theorem 3.12, the following optimal error estimate can be proved.

Theorem 3.18. *Let $(\mathbf{E}^m, \mathbf{H}^{m+\frac{1}{2}}, \mathbf{K}^m, \mathbf{M}^{m+\frac{1}{2}}, \mathbf{J}^{m+\frac{1}{2}})$ and $(\mathbf{E}_h^m, \mathbf{H}_h^{m+\frac{1}{2}}, \mathbf{K}_h^m, \mathbf{M}_h^{m+\frac{1}{2}}, \mathbf{J}_h^{m+\frac{1}{2}})$ be the analytic and numerical solutions of (3.183)–(3.187) and (3.199)–(3.202), respectively. Under proper regularity assumptions, there exists a constant $C > 0$ independent of h and τ such that*

$$\sqrt{\epsilon_0}||\mathbf{E}^n - \mathbf{E}_h^n||_0 + \sqrt{\mu_0}||\mathbf{H}^{n+\frac{1}{2}} - \mathbf{H}_h^{n+\frac{1}{2}}||_0 + \frac{1}{\sqrt{\epsilon_0\omega_p^2}}||\mathbf{J}^{n+\frac{1}{2}} - \mathbf{J}_h^{n+\frac{1}{2}}||_0$$
$$+\frac{1}{\sqrt{\mu_0\omega_0^2 F}}||\mathbf{K}^n - \mathbf{K}_h^n||_0 + \frac{1}{\sqrt{\mu_0 F}}||\mathbf{M}^{n+\frac{1}{2}} - \mathbf{M}_h^{n+\frac{1}{2}}||_0 \leq C(h^l + \tau^2),$$

where $l \geq 1$ is the order of the basis functions in spaces $\mathbf{U}_h$ and $\mathbf{V}_h$.

3.8 Bibliographical Remarks

In this chapter we presented some basic time-domain finite element (FETD) methods developed for Maxwell's equations in metamaterials. Some early works on FETD can be found in papers [79, 178] and references cited therein. Since 2000, in addition to our own work on FETD (e.g., [189, 190, 193, 194]), there has been a growing interest in developing FETD methods for Maxwell's equations in dispersive media [25, 160, 207, 251, 258, 290]. A nice list of literature on FETD methods for general complex media (including metamaterials) can be found in the review paper by Teixeira [277], which provides over 300 papers (though many are on FDTD methods) published by 2007. Another very recent and excellent review on FETD was written by Chen and Monk [68], which provides some numerical analysis on use of edge elements and certain A-stable schemes.

For more advanced finite element theory for Maxwell's equations, interested readers should consult some more theoretical papers such as [11, 13, 40, 41, 75, 145, 222] and the classic book by Monk [217].