New Higher-Order Basis Functions for Curvilinear Finite Elements

Dissertation
zur Erlangung des Grades
des Doktors der Ingenieurwissenschaften
der Naturwissenschaftlich-Technischen Fakultät
der Universität des Saarlandes

von

László Levente Tóth

Saarbrücken

2022

Tag des Kolloquiums:	27. Juni 2023
Dekan:	Prof. Dr. Ludger Santen
Berichterstatter:	Prof. Dr. techn. Romanus Dyczij-Edlinger Prof. Dr. Sergej Rjasanow Prof. Dr. Ralf Hiptmair
Vorsitz:	Prof. Dr. Dr. h.c. Heiko Rieger
Akad. Beisitzer:	Dr.-Ing. Amine Othmane

Acknowledgement

I thank my wife for supporting me and having the patience to endure my long-lasting work. I am grateful to my grandparents and parents for their continuous encouragement and teachings, furthermore, for all the knowledge I received that made me curious from a young age. I also thank my sister for her support and cheerful entertainment.

I thank my colleagues, Petra Braun, Peter Reiter, Rolf Baltes, Adrián-Amor Martín, and Lukas David Schuck, who helped me with solutions to various problems and always provided a fun time.

Lastly, I have to say thanks to my supervisor, Prof. Dr. Romanus Dyczij-Edlinger, for welcoming me to his chair and providing the opportunity to pursue a Ph.D. while giving me the possibility to have the experience of working and teaching at the university. I am also grateful to him for his care in general, for being precise, for providing advice with strong theoretical and technical knowledge, and for trusting me to work independently on my research.

Abstract

The first contribution is a fast calculation method for tetrahedral finite element matrices which is applicable to curvilinear geometries and inhomogeneous material properties. The element matrices are obtained at a low computational cost via scaled additions of universal matrices. The proposed technique is more efficient than competing approaches and provides well-defined lower and upper bounds for the required number of matrices.

In the case of tetrahedral $\boldsymbol{H}(\mathrm{div})$ elements, a new set of basis functions is proposed for the mixed-order Nédélec space. The specialty of the functions is a high level of orthogonality which applies to arbitrary straight-sided tetrahedra. The resulting condition numbers, compared to competing bases, are significantly lower.

The remaining contributions concern hexahedral elements, where a new, mixed-order serendipity element is proposed for $\boldsymbol{H}(\mathrm{curl})$-conforming functions. It allows the construction of a single set of hierarchical basis functions that can also be used to span various other finite element spaces. Therefore, it is possible to use different finite element spaces within the same mesh while maintaining conformity. In the curvilinear case, a special yet versatile way of mesh refinement is proposed along with serendipity basis functions for the interpolation of the geometry. The main advantage of the proposed methods is the resulting algebraic rate of convergence in $\boldsymbol{H}(\mathrm{curl})$-norm with the least possible number of unknowns.

Kurzfassung

Der erste Beitrag ist eine schnelle Berechnungsmethode von Finite-Elemente-Matrizen für Tetraeder, die auf krummlinige Geometrien und inhomogene Materialeigenschaften anwendbar ist. Die Elementmatrizen werden mit geringem Rechenaufwand durch skalierte Addition vorgefertigter Matrizen erstellt. Die vorgeschlagene Methode ist effizienter als vergleichbare Ansätze und liefert wohldefinierte obere und untere Schranken für die Anzahl der benötigten Matrizen.

Für $\boldsymbol{H}(\mathrm{div})$-konforme Elemente auf Tetraedern werden neue Ansatzfunktionen für den Nédélec-Raum gemischter Ordnung vorgestellt. Die Besonderheit dieser Funktionen ist ein hohes Maß an Orthogonalität für beliebige geradlinige Tetraeder. Im Vergleich zu anderen Ansatzfunktionen sind die resultierenden Konditionszahlen deutlich kleiner.

Die übrigen Beiträge betreffen Hexaeder, für die ein neues Serentipity-Element gemischter Ordnung vorgestellt wird. Es ermöglicht die Konstruktion hierarchischer Ansatzfunktionen, die auch zum Aufspannen anderer Finite-Elemente-Räume angewandt werden kann. Daher ist es möglich, verschiedene Finite-Elemente-Räume auf dem gleichen Netz zu verwenden und dabei Konformität zu bewahren. Für den krummlinigen Fall wird eine spezielle aber vielseitige Methode zur Netzverfeinerung mit Serentipity-Ansatzfunktionen zur Interpolation der Geometrie vorgestellt. Der Hauptvorteil der vorgestellten Methoden ist die algebraische Konvergenz in der Norm des $\boldsymbol{H}(\mathrm{rot})$ mit der kleinstmöglichen Anzahl an Unbekannten.

Contents

Chapter 1

Introduction

Commonly, a physical model of an engineering problem contains not only straight-sided geometric shapes but often also complicated curvilinear domains. The finite element method is a robust numerical analysis tool for investigating such complicated structures. To harvest the method's full potential, such as the algebraic rate of convergence and adaptability, it must be implemented appropriately. The starting point of my work was to implement different higher-order tetrahedral elements that can handle curvilinear geometries. Generally, these topics were considered relatively old, since most developments on curvilinear finite elements ended in the last century. The main results are summarized in the isoparametric finite elements [Cia02, Chapter 4.3], which require the same order of polynomial interpolation of the curvilinear geometry as the order of finite element basis functions. Hence, the same basis functions can be used to represent the geometry and the fields in the case of H^1-conforming functions [SF73, Chapter 3.3]. On the other hand, there is a new emerging method, namely the isogeometric method. Similarly to the isoparametric element, it uses spline-based basis functions for the representation of the geometry and for the unknown fields [BSV10]. However, it requires the availability of an exact, spline-based description of the geometry and the parametrization of the volumetric domain [RS12]. The disadvantage is that this information might not be available or may be hard to provide for most practical examples with complex geometry. For these practical reasons, I favored the flexibility of finite element meshes and continued with traditional curvilinear finite elements. Here, the essential aspects are the accurate, reliable, and efficient calculation of the different finite element matrices for curvilinear domains and inhomogeneous material properties.

The most common way to deal with the calculation of finite element matrices is to use numerical integration, where the computational cost scales highly with the order of finite element basis functions. A solution for this issue can be the use of master matrices or the so-called universal matrices [Sil78], that are a set of geometry- and material-independent (precalculated) element matrices. For straight-sided elements

with homogeneous materials, the element matrices are calculated by the scaled matrix addition of universal matrices. However, in the case of curvilinear elements or inhomogeneous materials, the geometry- and material-dependent terms not only consist of scalar constants but also depend on the coordinates. Therefore, traditional universal matrices are not applicable, whereas there is no problem with using numerical integration. This issue motivated the work of Chapter 4, constructing hierarchical universal matrices applicable to curvilinear geometries and inhomogeneous materials. The proposed method proved to be more efficient and reliable than competing approaches.

In previous works [Ing06], Ingelström proposed basis functions for the polynomial H^1-conforming space and for the $\boldsymbol{H}(\mathrm{curl})$-conforming mixed-order Nédélec space [Néd80], that are utilized in various finite element formulations and computer programs. Their common property is the orthogonality with respect to the interpolation operator, which yields a high level of sparsity and linear independence. Moreover, the $\boldsymbol{H}(\mathrm{curl})$-conforming basis reuses the gradient of the higher-order basis functions from the H^1-conforming basis as the null space functions of the curl operator. These properties can be highly useful for p adaptive methods [NW04] and multi-level preconditioners [ZC02]. This work thus motivated the construction of the remaining set, an $\boldsymbol{H}(\mathrm{div})$-conforming basis in Chapter 5 with the corresponding interpolation operator orthogonality. The previously developed H^1 and $\boldsymbol{H}(\mathrm{curl})$ basis and the newly developed $\boldsymbol{H}(\mathrm{div})$ bases fulfill the commutation properties of the continuous and discrete de Rham complexes (short: de Rham commutation), not just on the level of finite element spaces but also on the level of higher-order basis functions. They also show similar orthogonality properties with respect to the relevant interpolation operator. The finite element matrices resulting from the present basis functions are highly sparse and produce significantly lower condition numbers than competing bases.

In practical applications, the most common meshes are tetrahedral since this shape is the most versatile for general geometries. However, structured domains, such as an extruded surface or a curvilinear block, can be meshed by prismatic or hexahedral elements very efficiently. For this reason, the best meshes are considered to be composite meshes of different element types. As it turned out, the commonly used finite element spaces for hexahedra, such as the tensor product polynomial space for H^1-conforming functions or the mixed-order Nédélec space, produce a significantly larger number of unknowns than their tetrahedral counterpart. These observations motivated me to rigorously investigate the known hexahedral elements and look for better alternatives.

Most works on finite element error estimation, such as in [Cia02, Chapter 4.3], [ABF02], [FGM11], [ABB15], [Ape98], and [SF73, Chapter 3.3], address the convergence and its requirements in terms of the mesh parameter or element size h and the order of finite element basis functions p. In the case of isoparametric tetrahedral

elements, it is known that the interpolation of the geometry is required to be of the same order as the finite element basis functions [Cia02, Chapter 4.3]. Based on these analyses, several different finite elements have been developed for hexahedra. A large group is the serendipity elements [AA11], [AA14], [GK19], [GKS19], whose dimensions are smaller than the other tensor product spaces, although they yield the same $\mathcal{O}(h^p)$ rate of convergence for parallelepiped meshes with elements of affine geometry mappings. However, for elements with non-affine geometry mapping (not parallelepiped meshes), these elements are considered insufficient due to the degrading rate of convergence [ABB15], [DG19], [GK19].

Another large group is the tensor product elements, such as in [BF12, Chapter 2.], [AP02], [Néd80], [Néd86], [ABB15], [FGM11], [BD13b] and [BD13a], that yield the same $\mathcal{O}(h^p)$ rate of convergence for both affine geometry mappings and arbitrary straight-sided elements. Given our interest in electromagnetics, the recently developed lowest order ($p = 1$) space of Falk et al. [FGM11] has great importance since their space provides the mentioned convergence in $\boldsymbol{H}(\text{curl})$-norm, whereas the mixed-order space of Nédélec provides it only in $\boldsymbol{L}^2$-norm. The corresponding higher-order finite element was reported by Bergot et al. [BD13b]. However, no proof of the space minimality and no error estimation have been published. The case of the curvilinear hexahedron appears even more problematic because it was proven in [ABB15] that none of the elements above provide $\mathcal{O}(h^p)$ convergence rate in the $\boldsymbol{H}(\text{curl})$-norm for general curvilinear meshes. This suggests that even richer spaces might be necessary. These discouraging results motivated my research for alternatives that feature the least possible number of unknowns while providing optimal rates of convergence [B2].

The new aspects, compared to other works, lie in an alternative definition of the mesh parameter and refinement and their relation to the interpolated geometry mappings. In order to lay down the foundations, rigorous finite element error analyses are given in Chapter 3. The resulting error estimations provide additional ideas for maintaining the ideal convergence rates. As an outcome, general convergence requirements are given in the proposed framework for curvilinear finite elements with interpolated geometry mappings, considering different ways of mesh refinements.

In Chapter 6, definitions are provided for some of the best-known hexahedral elements that fit into the framework of the present general error estimates. Furthermore, a new finite element space, the "mixed-order serendipity element", is introduced for $\boldsymbol{H}(\text{curl})$-conforming functions. The main advantage is in attaining the least possible number of unknowns while providing $\mathcal{O}(h^p)$ rate of convergence in $\boldsymbol{H}(\text{curl})$-norm for affinely refined meshes. Moreover, it allows the construction of hierarchical basis functions that are compatible with the mixed-order Nédélec space and the tensor product space of Falk et al. and Bergot et al. A single set of basis functions is provided in the thesis that can span these spaces, that is orthogonal with respect to the interpolation operator, and has a separate higher-order subset

of gradient functions. The existence of such a basis allowed the development of two new methods, the mixed-order mixed-space element for straight-sided meshes and the iso-serendipity element for curvilinear meshes. To my knowledge, these are the first methods for non-parallelepiped hexahedral meshes that apply serendipity spaces with minimal dimensions while providing an $\mathcal{O}(h^p)$ convergence rate in $\boldsymbol{H}(\text{curl})$-norm.

Chapter 2

Mathematical and Physical Foundations

2.1 Maxwell's Equations

Physics, as a field, describes the behavior of nature in terms of mathematical equations. Considering the restrictions of classical physics, electromagnetic fields are described by Maxwell's equations. The differential representation of the equations in the time domain reads as follows:

$$\nabla \times \boldsymbol{\mathcal{E}}(\boldsymbol{r},t) = -\frac{\partial}{\partial t}\boldsymbol{\mathcal{B}}(\boldsymbol{r},t), \tag{2.1a}$$

$$\nabla \times \boldsymbol{\mathcal{H}}(\boldsymbol{r},t) = \frac{\partial}{\partial t}\boldsymbol{\mathcal{D}}(\boldsymbol{r},t) + \boldsymbol{\mathcal{J}}(\boldsymbol{r},t), \tag{2.1b}$$

$$\nabla \cdot \boldsymbol{\mathcal{D}}(\boldsymbol{r},t) = \varrho(\boldsymbol{r},t), \tag{2.1c}$$

$$\nabla \cdot \boldsymbol{\mathcal{B}}(\boldsymbol{r},t) = 0. \tag{2.1d}$$

Here, $\boldsymbol{\mathcal{E}}$ is the electric field, $\boldsymbol{\mathcal{H}}$ is the magnetic field, $\boldsymbol{\mathcal{D}}$ is the electric displacement field, $\boldsymbol{\mathcal{B}}$ is the magnetic flux density, $\boldsymbol{\mathcal{J}}$ is the electric current density, and ϱ is the electric charge density. Each quantity may depend on the coordinate vector $\boldsymbol{r}$ and time variable t. The connection to classical mechanics is given by the Lorentz force

$$\boldsymbol{F} = q\left(\boldsymbol{\mathcal{E}} + \boldsymbol{v} \times \boldsymbol{\mathcal{B}}\right), \tag{2.2}$$

where q is the point charge and $\boldsymbol{v}$ is the velocity. The remaining fields, $\boldsymbol{\mathcal{H}}$, $\boldsymbol{\mathcal{D}}$, and $\boldsymbol{\mathcal{J}}$ are related to $\boldsymbol{\mathcal{E}}$ and $\boldsymbol{\mathcal{B}}$ through the constitutive relations

$$\boldsymbol{\mathcal{D}} = \varepsilon\boldsymbol{\mathcal{E}}, \tag{2.3a}$$

$$\boldsymbol{\mathcal{B}} = \mu\boldsymbol{\mathcal{H}}, \tag{2.3b}$$

$$\boldsymbol{\mathcal{J}} = \sigma\boldsymbol{\mathcal{E}}. \tag{2.3c}$$

Here, $\boldsymbol{\varepsilon}$ is the electric permittivity, $\boldsymbol{\mu}$ is the magnetic permeability, and $\boldsymbol{\sigma}$ is the conductivity. In the general case, these quantities are operators that describe the electrical properties of the materials. However, in many cases, it is sufficient to consider these operators to be linear and time-invariant. In those cases, the constitutive relations become

$$\boldsymbol{\mathcal{D}} = \varepsilon_0 \boldsymbol{\varepsilon_r} \boldsymbol{\mathcal{E}}, \tag{2.4a}$$
$$\boldsymbol{\mathcal{B}} = \mu_0 \boldsymbol{\mu_r} \boldsymbol{\mathcal{H}}, \tag{2.4b}$$
$$\boldsymbol{\mathcal{J}} = \boldsymbol{\sigma} \boldsymbol{\mathcal{E}}, \tag{2.4c}$$

where ε_0 is the permittivity of the vacuum, $\boldsymbol{\varepsilon_r}$ is the relative permittivity tensor, μ_0 is the permeability of the vacuum, $\boldsymbol{\mu_r}$ is the relative permeability tensor, and $\boldsymbol{\sigma}$ is the conductivity tensor. These tensorial quantities describe the effect of the materials relative to the case of vacuum.

The excitation in electromagnetics is often a periodic field in terms of time. Thus, any field $\boldsymbol{\mathcal{Y}}$ can be written as

$$\boldsymbol{\mathcal{Y}} = \Re\left\{\boldsymbol{Y} e^{j\omega t}\right\}, \tag{2.5}$$

where $\boldsymbol{Y}$ is the complex amplitude, ω is the angular frequency, and $j = \sqrt{-1}$ is the imaginary unit. The operator $\Re$ takes the real part of a complex quantity. The advantage of the time-harmonic representation is that the time derivative becomes a multiplication:

$$\frac{\partial}{\partial t} \mapsto j\omega. \tag{2.6}$$

Therefore, Maxwell's equations in the frequency domain read

$$\nabla \times \boldsymbol{E}(\boldsymbol{r}, \omega) = -\mathrm{j}\omega \boldsymbol{B}(\boldsymbol{r}, \omega), \tag{2.7a}$$
$$\nabla \times \boldsymbol{H}(\boldsymbol{r}, \omega) = \mathrm{j}\omega \boldsymbol{D}(\boldsymbol{r}, \omega) + \boldsymbol{J}(\boldsymbol{r}, \omega), \tag{2.7b}$$
$$\nabla \cdot \boldsymbol{D}(\boldsymbol{r}, \omega) = \rho(\boldsymbol{r}, \omega), \tag{2.7c}$$
$$\nabla \cdot \boldsymbol{B}(\boldsymbol{r}, \omega) = 0. \tag{2.7d}$$

Here, the quantities $\boldsymbol{E}$, $\boldsymbol{H}$, $\boldsymbol{D}$, $\boldsymbol{B}$, and ρ are the corresponding complex amplitudes of the fields $\boldsymbol{\mathcal{E}}$, $\boldsymbol{\mathcal{H}}$, $\boldsymbol{\mathcal{D}}$, $\boldsymbol{\mathcal{B}}$, and ϱ in the frequency domain. The form of the constitutive relations remains the same, since only time-invariant material properties were considered:

$$\boldsymbol{D} = \varepsilon_0 \boldsymbol{\varepsilon_r} \boldsymbol{E}, \tag{2.8a}$$
$$\boldsymbol{B} = \mu_0 \boldsymbol{\mu_r} \boldsymbol{H}, \tag{2.8b}$$
$$\boldsymbol{J} = \boldsymbol{\sigma} \boldsymbol{E}. \tag{2.8c}$$

2.1.1 Interface Conditions

In electromagnetics, the interface conditions define the behavior of fields on the interfaces of domains Ω_1 and Ω_2 with different material properties. The unit-normal vector of the interface is denoted by $\boldsymbol{n}_{12}$ and the electric field of Ω_1 and Ω_2 is denoted by $\boldsymbol{E}_1$ and $\boldsymbol{E}_2$ on the common interface. Then, the interface condition for the electric field reads

$$\boldsymbol{n}_{12} \times (\boldsymbol{E}_2 - \boldsymbol{E}_1) = 0. \tag{2.9}$$

This implies that the tangential component is continuous, whereas the normal component may jump on the interface. A similar relationship holds for magnetic fields $\boldsymbol{H}_1$ and $\boldsymbol{H}_2$, whose tangential discontinuity is determined by the surface current $\boldsymbol{K}$ as

$$\boldsymbol{n}_{12} \times (\boldsymbol{H}_2 - \boldsymbol{H}_1) = \boldsymbol{K}. \tag{2.10}$$

Therefore, the normal component of the magnetic field may be discontinuous on the interface, whereas both $\boldsymbol{E}$ and $\boldsymbol{H}$ are tangentially continuous in the absence of surface currents.

The interface condition for the magnetic flux densities $\boldsymbol{B}_1$ and $\boldsymbol{B}_2$ reads

$$\boldsymbol{n}_{12} \cdot (\boldsymbol{B}_2 - \boldsymbol{B}_1) = 0, \tag{2.11}$$

which means that these fields are required to have normal continuity, whereas the tangential component may be discontinuous on the interface. At last, the electric displacement fields $\boldsymbol{D}_1$ and $\boldsymbol{D}_2$ are considered. The discontinuity of the normal component is determined by

$$\boldsymbol{n}_{12} \cdot (\boldsymbol{D}_2 - \boldsymbol{D}_1) = \sigma_q, \tag{2.12}$$

where σ_q is the surface charge density. Similar to the previous case, the tangential components may be discontinuous, whereas both $\boldsymbol{B}$ and $\boldsymbol{D}$ have continuous normal components on the interface in the absence of surface charges.

2.2 Function Spaces

When numerical techniques such as the finite element method (FEM) are used to solve Maxwell's equations, the computational domain $\Omega = \cup_e \Omega_e$ is subdivided into disjointed elements Ω_e. In the present discussion, the domain is considered to be $\Omega \subset \mathbb{R}^3$ and the coordinate vector is denoted by $\boldsymbol{r}$. On each element, basis functions are used for the approximation of the ideal, exact solution. In most cases, however, the finite element approximation converges to the exact solution but never reaches

it. Hence, it is essential to choose basis functions and formalism in such a way that they preserve some important properties of the exact solution, such as the continuity requirements and interface conditions. The right choice of function spaces can help to ensure that these requirements are met.

The spaces of scalar- and vector-valued p-integrable functions are denoted by $L^p(\Omega)$ and $\boldsymbol{L}^p(\Omega)$, respectively. Their formal definitions read

$$L^p(\Omega) := \{f \in \mathbb{C} \mid \|f\|^p_{L^p(\Omega)} < \infty \text{ for } p \in \mathbb{Z}^+\}, \tag{2.13a}$$

$$\boldsymbol{L}^p(\Omega) := \{\boldsymbol{f} \in \mathbb{C}^n \mid \|\boldsymbol{f}\|^p_{\boldsymbol{L}^p(\Omega)} < \infty \text{ for } n, p \in \mathbb{Z}^+\}, \tag{2.13b}$$

where the corresponding norms are

$$\|f\|_{L^p(\Omega)} := \left(\int_\Omega |f|^p \, \mathrm{d}\boldsymbol{r} \right)^{1/p}, \tag{2.14a}$$

$$\|\boldsymbol{f}\|_{\boldsymbol{L}^p(\Omega)} := \left(\sum_{i=n}^{n} \| [\boldsymbol{f}]_i \|^p_{L^p(\Omega)} \, \mathrm{d}\boldsymbol{r} \right)^{1/p}. \tag{2.14b}$$

In the limiting case of $p = \infty$, an alternative definition is provided, which applies not only for scalar- or vector-valued functions but also to matrix-valued functions,

$$L^\infty(\Omega) := \{\boldsymbol{F}(\boldsymbol{r}) \in \mathbb{C}^m \times \mathbb{C}^n \mid \|\boldsymbol{F}(\boldsymbol{r})\|_{L^\infty(\Omega)} < \infty \text{ for } m, n \in \mathbb{Z}^+\}, \tag{2.15}$$

as the corresponding norm is defined as the largest essential supremum of all components, as follows:

$$\|\boldsymbol{F}(\boldsymbol{r})\|_{L^\infty(\Omega)} := \max_{i,j} \left(\operatorname*{ess.\,sup}_{\boldsymbol{r} \in \Omega} | [\boldsymbol{F}(\boldsymbol{r})]_{ij} | \right). \tag{2.16}$$

Next, some additional definitions are provided for function spaces that impose some requirements on the function derivatives. Let us denote the weak or distributional derivative D^α (noted in other references as $\partial^\alpha \Phi$ [Cia02, Chapter 1], [M$^+$03, Chapter 3]) as of order $p = |\alpha|$. The $\alpha = (\alpha_1, \alpha_2, \alpha_3)$ denote a multi-index of non-negative integers $\alpha_i \in \mathbb{Z}^+$ with the absolute value of $|\alpha| = \alpha_1 + \alpha_2 + \alpha_3$. In the case of an infinitely differentiable scalar-valued function $\Phi \in C^\infty(\Omega)$ or a vector-valued function $\Phi \in \boldsymbol{C}^\infty(\Omega)$, the following definitions [M$^+$03, Chapter 3] are equivalents:

$$D^\alpha \Phi = \partial^\alpha \Phi = \frac{\partial^{|\alpha|} \Phi}{\partial \boldsymbol{r}^\alpha} = \frac{\partial^{|\alpha|} \Phi}{\partial r_1^{\alpha_1} \partial r_2^{\alpha_2} \partial r_3^{\alpha_3}}. \tag{2.17}$$

However, if the function is integrable but not differentiable over the entire domain, a general definition based on distributions is required [Tar07, Chapter 4]. For any distribution Φ one defines the distribution $D^\alpha \Phi$ by the formula

$$(D^\alpha \Phi, \Theta)_{\boldsymbol{L}^2(\Omega)} = (-1)^{|\alpha|} \, (\Phi, D^\alpha \Theta)_{\boldsymbol{L}^2(\Omega)}. \tag{2.18}$$

Here, Θ is an $|\alpha|$ times differentiable smooth function (in $\Theta \in C^{|\alpha|}(\Omega)$ or $\Theta \in \boldsymbol{C}^{|\alpha|}(\Omega)$) with compact support over the domain Ω. Such a general definition of the derivative is required for the definition of the finite element approximation spaces. Based on the notations from [Cia02, Chapter 1], the scalar- and vector-valued Hilbert spaces are denoted by $H^m(\Omega)$ and $\boldsymbol{H}^m(\Omega)$, respectively. Their formal definitions read

$$H^m(\Omega) := \left\{ f \in L^2(\Omega) \,|\, \partial^\alpha f \in L^2(\Omega) \text{ for } |\alpha_i| \leq m \right\}, \tag{2.19a}$$

$$\boldsymbol{H}^m(\Omega) := \left\{ \boldsymbol{f} \in \boldsymbol{L}^2(\Omega) \,|\, \partial^\alpha \boldsymbol{f} \in \boldsymbol{L}^2(\Omega) \text{ for } |\alpha_i| \leq m \right\}, \tag{2.19b}$$

with corresponding norms

$$\|f\|_{H^m(\Omega)} := \sqrt{\sum_{|\alpha| \leq m} \|\partial^\alpha f\|^2_{L^2(\Omega)}}, \tag{2.20a}$$

$$\|\boldsymbol{f}\|_{\boldsymbol{H}^m(\Omega)} := \sqrt{\sum_{|\alpha| \leq m} \|\partial^\alpha \boldsymbol{f}\|^2_{\boldsymbol{L}^2(\Omega)}}, \tag{2.20b}$$

and semi-norms

$$|f|_{H^m(\Omega)} := \sqrt{\sum_{|\alpha| = m} \|\partial^\alpha f\|^2_{L^2(\Omega)}}, \tag{2.21a}$$

$$|\boldsymbol{f}|_{\boldsymbol{H}^m(\Omega)} := \sqrt{\sum_{|\alpha| = m} \|\partial^\alpha \boldsymbol{f}\|^2_{\boldsymbol{L}^2(\Omega)}}. \tag{2.21b}$$

For later usage in error estimation, the generalized space of $\boldsymbol{H}^m(\Omega)$, the Sobolev space [Cia02, Section 3.1] is denoted by $W^{m,p}(\Omega)$ for a domain Ω, any integer $m \geq 0$, and any number p satisfying $1 \leq p \leq \infty$. The space $W^{m,p}(\Omega)$ consists of those functions $\boldsymbol{f} \in \boldsymbol{L}^p(\Omega)$ for which all distributional derivatives of a multi-index α belong to the space $\boldsymbol{L}^p(\Omega)$ if $|\alpha| \leq m$. The corresponding norm is

$$\begin{aligned} \|\boldsymbol{f}\|_{m,p,\Omega} &:= \left(\sum_{|\alpha| \leq m} \|\partial^\alpha \boldsymbol{f}\|^p_{\boldsymbol{L}^p(\Omega)} \right)^{1/p} && \text{if } 1 \leq p < \infty, \\ \|\boldsymbol{f}\|_{m,\infty,\Omega} &:= \max_{|\alpha| \leq m} \|\partial^\alpha \boldsymbol{f}\|_{L^\infty(\Omega)} && \text{if } p = \infty. \end{aligned} \tag{2.22}$$

and the semi-norm is

$$\begin{aligned} |\boldsymbol{f}|_{m,p,\Omega} &:= \left(\sum_{|\alpha| = m} \|\partial^\alpha \boldsymbol{f}\|^p_{\boldsymbol{L}^p(\Omega)} \right)^{1/p} && \text{if } 1 \leq p < \infty, \\ |\boldsymbol{f}|_{m,\infty,\Omega} &:= \max_{|\alpha| = m} \|\partial^\alpha \boldsymbol{f}\|_{L^\infty(\Omega)} && \text{if } p = \infty. \end{aligned} \tag{2.23}$$

Note that for the case $p = 2$, $W^{m,2}(\Omega) = \boldsymbol{H}^m(\Omega)$ and the corresponding norms are equivalent.

Finally, the definitions of the less general Hilbert spaces, which are of great importance in electromagnetics, are given:

$$\boldsymbol{H}(\mathrm{curl}, \Omega) := \left\{ \boldsymbol{w} \in \boldsymbol{L}^2(\Omega) \,|\, \nabla \times \boldsymbol{w} \in \boldsymbol{L}^2(\Omega) \right\}, \tag{2.24a}$$

$$\boldsymbol{H}(\mathrm{div}, \Omega) := \left\{ \boldsymbol{d} \in \boldsymbol{L}^2(\Omega) \,|\, \nabla \cdot \boldsymbol{d} \in L^2(\Omega) \right\}. \tag{2.24b}$$

The corresponding norms are

$$\|\boldsymbol{w}\|_{\boldsymbol{H}(\mathrm{curl},\Omega)} := \sqrt{\|\boldsymbol{w}\|^2_{\boldsymbol{L}^2(\Omega)} + \|\nabla \times \boldsymbol{w}\|^2_{\boldsymbol{L}^2(\Omega)}}, \tag{2.25a}$$

$$\|\boldsymbol{d}\|_{\boldsymbol{H}(\mathrm{div},\Omega)} := \sqrt{\|\boldsymbol{d}\|^2_{\boldsymbol{L}^2(\Omega)} + \|\nabla \cdot \boldsymbol{d}\|^2_{L^2(\Omega)}}. \tag{2.25b}$$

For a simply connected domain $\Omega \subset \mathbb{R}^3$, the de Rham complex [BBF13, Chapter 2.1.4] reads

$$H^1(\Omega) \xrightarrow{\nabla} \boldsymbol{H}(\mathrm{curl}, \Omega) \xrightarrow{\nabla\times} \boldsymbol{H}(\mathrm{div}, \Omega) \xrightarrow{\nabla\cdot} L^2(\Omega). \tag{2.26}$$

This sequence describes how the function spaces of $H^1(\Omega)$, $\boldsymbol{H}(\mathrm{curl}, \Omega)$, $H(\mathrm{div}, \Omega)$, and $L^2(\Omega)$ are connected by the differential operators ∇, $\nabla\times$ and $\nabla\cdot$ [Bos98, Chapter 5.1]. The image of the scalar function $v \in H^1(\Omega)$ becomes

$$\nabla v \in \boldsymbol{H}(\mathrm{curl}, \Omega),$$

which is in the null space of the $\nabla\times$ operator,

$$\nabla \times \nabla v = 0.$$

Similarly, the image of a vector function $\boldsymbol{w} \in \boldsymbol{H}(\mathrm{curl}, \Omega)$ becomes

$$\nabla \times \boldsymbol{w} \in \boldsymbol{H}(\mathrm{div}, \Omega),$$

which is in the null space of the $\nabla\cdot$ operator,

$$\nabla \cdot \nabla \times \boldsymbol{w} = 0.$$

Moreover, the image of a function $\boldsymbol{d} \in \boldsymbol{H}(\mathrm{div}, \Omega)$ belongs to the space of

$$\nabla \cdot \boldsymbol{d} \in L^2(\Omega).$$

Applying these principles to Maxwell's equations (2.1) revealed [Bos98, Chapter 5.1] that the natural way of approximating the electric field is to use functions $\boldsymbol{w} \in \boldsymbol{H}(\mathrm{curl}, \Omega)$ and for the magnetic flux density $\boldsymbol{d} \in \boldsymbol{H}(\mathrm{div}, \Omega)$. Similar concepts hold for the other vector fields and for the related potentials. Let us denote the

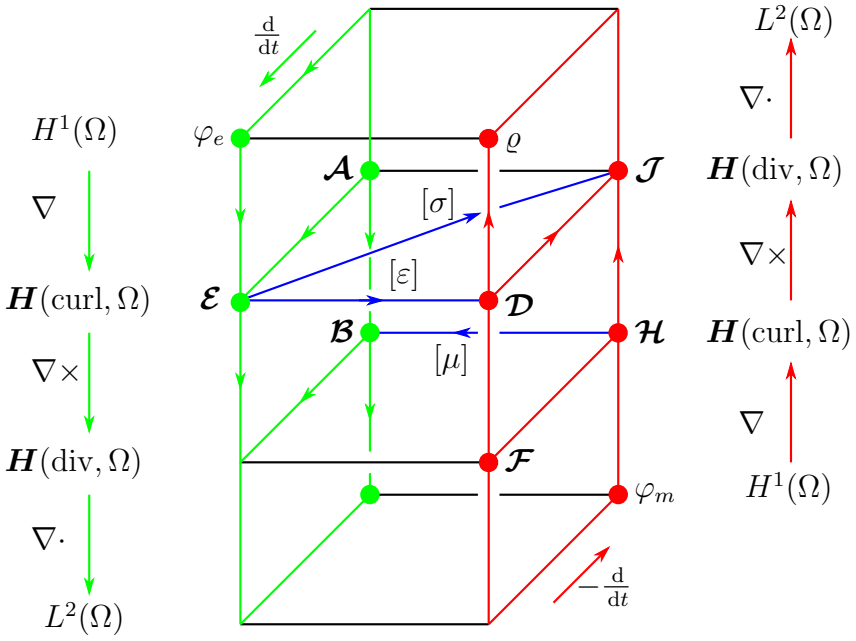


Figure 2.1: The structure of Maxwell's equations [Bal18].

electric and magnetic scalar potentials by φ_e and φ_m and the electric and magnetic vector potentials by $\boldsymbol{\mathcal{F}}$ and $\boldsymbol{\mathcal{A}}$, respectively. Their formal definitions read

$$\nabla \times \boldsymbol{\mathcal{A}} = \boldsymbol{\mathcal{B}}, \tag{2.27a}$$

$$-\nabla\varphi_e = \boldsymbol{\mathcal{E}} + \frac{\partial}{\partial t}\boldsymbol{\mathcal{A}}, \tag{2.27b}$$

$$\nabla \times \boldsymbol{\mathcal{F}} = -\boldsymbol{\mathcal{D}}, \tag{2.27c}$$

$$-\nabla\varphi_m = \boldsymbol{\mathcal{H}} + \frac{\partial}{\partial t}\boldsymbol{\mathcal{F}}. \tag{2.27d}$$

The entire structure of Maxwell's equations is illustrated in Figure 2.1 [Bal18], which is the so-called "Maxwell's house" [Bos98]. The de Rham complex of the function spaces is shown along the vertical direction, and the structure is composed of two distinct sequences. The left green side and the right red side represent the primal and dual complexes that are connected by the material operators. This structure should also be preserved in the discrete setting when the different fields are approximated via the finite element method.

2.3 Electric Field Formulation

In this section, the focus is on the frequency-domain solution of Maxwell's equations for the case $\omega > 0$, where the time-harmonic electric and magnetic fields are coupled. With these assumptions, Maxwell's equations read

$$\nabla \times \boldsymbol{E} = -\mathrm{j}\omega \boldsymbol{B}, \tag{2.28a}$$
$$\nabla \times \boldsymbol{H} = \mathrm{j}\omega \boldsymbol{D} + \boldsymbol{J}, \tag{2.28b}$$
$$\nabla \cdot \boldsymbol{D} = 0, \tag{2.28c}$$
$$\nabla \cdot \boldsymbol{B} = 0. \tag{2.28d}$$

The speed of light in vacuum is denoted by c_0, and it is related to the permittivity and permeability by

$$c_0 = \frac{1}{\sqrt{\mu_0 \varepsilon_0}}. \tag{2.29}$$

The free-space wavenumber k_0 for an angular frequency ω reads

$$k_0 = \frac{\omega}{c_0}. \tag{2.30}$$

Hence, the definition for the characteristic impedance of free space η_0 is given as

$$\eta_0 = \sqrt{\frac{\mu_0}{\varepsilon_0}} = \frac{\omega \mu_0}{k_0}. \tag{2.31}$$

Using the constitutive relations and substituting (2.28a) into (2.28b), the curl-curl wave equation of the electric field is obtained:

$$\nabla \times \boldsymbol{\mu}_r^{-1} \nabla \times \boldsymbol{E} + \mathrm{j} k_0 \eta_0 \boldsymbol{\sigma} \boldsymbol{E} - k_0^2 \boldsymbol{\varepsilon}_r \boldsymbol{E} = \boldsymbol{0}. \tag{2.32}$$

In the case of a computational domain Ω with a finite volume, boundary conditions need to be imposed on $\partial\Omega$. In the present setting, the boundary consists of perfect electrical condors (PEC) on Γ_E and multiple different port boundaries on Γ_i indexed by the integer i:

$$\partial\Omega = \Gamma_E \cup \Gamma_1 \cup \Gamma_2 \cdots \cup \Gamma_i \cdots . \tag{2.33}$$

All of the considered port boundaries and their neighborhoods are restricted to have homogeneous isotropic material properties, and only the dominant TE mode is assumed on each Γ_i. In this case, the associated modal impedance Z_i [Har01, Chapter 4] on Γ_i and the (imaginary) propagation constant $j\beta_i$ are related by

$$Z_i = \frac{\omega \mu_0 \mu_r}{\beta_i}. \tag{2.34}$$

The model is excited via the incident electric field $\boldsymbol{E}_k^{\mathrm{inc}}$ of the dominant mode on a given port boundary Γ_k of index k.

The corresponding boundary value problem [Ing06, ZC06] reads

$$\nabla \times \boldsymbol{\mu}_r^{-1} \nabla \times \boldsymbol{E} + \mathrm{j} k_0 \eta_0 \boldsymbol{\sigma} \boldsymbol{E} - k_0^2 \boldsymbol{\varepsilon}_r \boldsymbol{E} = \boldsymbol{0}, \qquad \text{in } \Omega, \tag{2.35a}$$

$$\boldsymbol{n} \times (\boldsymbol{E} \times \boldsymbol{n}) = \boldsymbol{0}, \qquad \text{on } \Gamma_E, \tag{2.35b}$$

$$(\boldsymbol{\mu}_r^{-1} \nabla \times \boldsymbol{E}) \times \boldsymbol{n} + \frac{j k_0 \eta_0}{Z_i} \boldsymbol{n} \times \boldsymbol{E} \times \boldsymbol{n} = \delta_{ik} \frac{2 j k_0 \eta_0}{Z_k} \boldsymbol{n} \times \boldsymbol{E}_k^{\mathrm{inc}} \times \boldsymbol{n} \qquad \text{on } \Gamma_i. \tag{2.35c}$$

Here, $\boldsymbol{n}$ is the outward oriented unit-length normal vector on the boundaries and δ_{ik} is the Kronecker delta function. Let the subspace $\boldsymbol{H}_E(\mathrm{curl}, \Omega)$ be defined by

$$\boldsymbol{H}_E(\mathrm{curl}, \Omega) = \left\{ \boldsymbol{w} \in \boldsymbol{H}(\mathrm{curl}, \Omega) \,|\, \boldsymbol{n} \times (\boldsymbol{w} \times \boldsymbol{n}) = \boldsymbol{0} \text{ on } \Gamma_E \right\}. \tag{2.36}$$

Then, the corresponding weak formulation, obtained via the Petrov-Galerkin method, [Ing06, ZC06] reads

$$\int_\Omega \nabla \times \boldsymbol{w} \cdot \boldsymbol{\mu}_r^{-1} \nabla \times \boldsymbol{E} \, \mathrm{d}\boldsymbol{r} + j k_0 \eta_0 \int_\Omega \boldsymbol{w} \cdot \boldsymbol{\sigma} \boldsymbol{E} \, \mathrm{d}\boldsymbol{r} - k_0^2 \int_\Omega \boldsymbol{w} \cdot \boldsymbol{\varepsilon}_r \boldsymbol{E} \, \mathrm{d}\boldsymbol{r} +$$

$$\sum_i \frac{j k_0 \eta_0}{Z_i} \int_{\Gamma_i} (\boldsymbol{w} \times \boldsymbol{n}) \cdot (\boldsymbol{E} \times \boldsymbol{n}) \, \mathrm{d}\boldsymbol{r} = \frac{2 j k_0 \eta_0}{Z_k} \int_{\Gamma_k} (\boldsymbol{w} \times \boldsymbol{n}) \cdot \left(\boldsymbol{E}_k^{\mathrm{inc}} \times \boldsymbol{n} \right) \mathrm{d}\boldsymbol{r}, \tag{2.37}$$

where the test function $\boldsymbol{w}$ and electric field are in the subspace $\boldsymbol{w}, \boldsymbol{E} \in \boldsymbol{H}_E(\mathrm{curl}, \Omega)$. Using appropriate normalizations for the modal electric fields $\boldsymbol{E}_i^{\mathrm{inc}}$ on the different ports,

$$\frac{1}{Z_i} \int_{\Gamma_i} \left(\boldsymbol{E}_i^{\mathrm{inc}} \times \boldsymbol{n} \right) \cdot \left(\boldsymbol{E}_i^{\mathrm{inc}} \times \boldsymbol{n} \right) \mathrm{d}\boldsymbol{r} = 1, \tag{2.38}$$

the scattering parameters are calculated [ZC06, Chapter 6.3.1] as

$$S_{ik} = \frac{1}{Z_i} \int_{\Gamma_i} (\boldsymbol{E} \times \boldsymbol{n}) \cdot \left(\boldsymbol{E}_i^{\mathrm{inc}} \times \boldsymbol{n} \right) \mathrm{d}\boldsymbol{r} - \delta_{ik}. \tag{2.39}$$

There are other formulations that can handle higher-order modes on the ports [Bal18, Chapter 4] or yield more stability for low frequencies [Joc13]. However, the main contributions of this thesis are related to the spaces of basis functions, and this formulation is sufficient for this purpose without loss of generality.

2.4 The Finite Element Method

So far, only functions that belong to an infinite-dimensional function space have been considered. In order to obtain a numerical solution, discretization is required with a finite-dimensional approximation space. The finite element method subdivides the computational domain Ω into a finite number of elements Ω_e with a given topology. Each Ω_e is a simply connected compact sub-domain of Ω with Lipschitz continuous boundaries. Hence, each element consists of nodes N, edges E, faces F, and volume V. The critical parameter that characterizes an element is the mesh parameter or mesh size $h = \mathrm{diam}(\Omega_e)$ [Cia02, Chapter 3.1]. Each element is associated with an approximation space, which is often a polynomial space. Further on, the polynomial space of order p is denoted by $\mathbb{P}_p$, which is spanned by all monomials of order $i \leq p$. Moreover, the space of homogeneous polynomials is denoted by $\widetilde{\mathbb{P}}_p$, which is spanned by the monomials of exact order $i = p$.

The best-known mathematical definition of a finite element is given by Ciarlet [Cia02, Chapter 2.3]. In order to fit into the present framework, only some notations were changed.

Definition 2.1 (Finite Element)**.** *The* Finite Element *is a triplet* $\{\Omega_e, P, \Sigma\}$*, where*

1. Ω_e *is the domain of the finite element,*
2. P *is the space of real-valued functions defined over the set* Ω_e*,*
3. Σ *is the finite set of linearly independent linear forms* M_a*, defined over the space* P*.*

The linear forms M_a are also called moments or degrees of freedoms of an element. Here, a is an index for the different linearly independent moments. For an arbitrary $v \in P$, there exists a unique real scalar α_a which satisfies

$$M_a(v) = \alpha_a. \tag{2.40}$$

Consequently, there exists a corresponding set of functions $v_a \in P$, which produces a linearly independent matrix $\boldsymbol{A}$ as

$$[\boldsymbol{A}]_{ab} = M_b(v_a). \tag{2.41}$$

This set can be used as a basis for the representation of an arbitrary $v \in P$. Particularly, if $\boldsymbol{A}$ is an identity matrix, then

$$v = \sum_a \alpha_a v_a. \tag{2.42}$$

Hence, one may review the bases $\{M_a\}$ and $\{v_a\}$ as dual bases in the algebraic sense [Cia02, Chapter 2.3]. This also reflects that the number of moments must be the same as the finite element basis functions. Moreover, the approximation space is determined by $\operatorname{span}\{v_a\} = P$. Thus, in Definition 2.1, the second point highly depends on the third point. Therefore, a similar definition as Definition 2.1 with the same structure will always require a unisolvent finite element and an invertible matrix $\boldsymbol{A}$.

In the case of $\Omega \subset \mathbb{R}^3$, the important spaces for conforming finite elements are $H^1(\Omega)$, $\boldsymbol{H}(\operatorname{curl}, \Omega)$, $\boldsymbol{H}(\operatorname{div}, \Omega)$ and $L^2(\Omega)$. The corresponding finite-dimensional approximation spaces of order p are denoted by

$$\mathcal{V}_p \subset H^1(\Omega), \tag{2.43a}$$
$$\mathcal{W}_p \subset \boldsymbol{H}(\operatorname{curl}, \Omega), \tag{2.43b}$$
$$\mathcal{D}_p \subset \boldsymbol{H}(\operatorname{div}, \Omega), \tag{2.43c}$$
$$\mathcal{P}_p \subset L^2(\Omega). \tag{2.43d}$$

In the case of $\mathcal{V}_p$, the index p indicates that the polynomial space of order $\mathbb{P}_p$ is included in the approximation space. Generally, however, p indicates that the asymptotic rate of convergence is $\mathcal{O}(h^p)$ in the corresponding natural norm. In order to preserve the structure of Maxwell's equations, as shown in Figure 2.1, and the conformity between the elements, the discrete approximation spaces must also form a de Rham complex. This discrete complex of the pth-order approximation spaces reads

$$\mathcal{V}_p \xrightarrow{\nabla} \mathcal{W}_p \xrightarrow{\nabla\times} \mathcal{D}_p \xrightarrow{\nabla\cdot} \mathcal{P}_p. \tag{2.44}$$

Historically, this concept was first utilized for first-order finite elements on simplicial domains, where it is referred to as the Whitney complex [Bos98, Chapter 5.2.2].

The Finite Element Equation System

In the following, finite element discretization is applied to the weak formulation of (2.37). Let us assume that the unknown electric field is $\boldsymbol{E} \in \mathcal{W}_p \subset \boldsymbol{H}(\operatorname{curl}, \Omega)$ on each element. Then, the general expression of the electric field on a single element can be written as the linear combination of $\boldsymbol{w}_b \in \mathcal{W}_p$ with some constant coefficients x_b^e:

$$\boldsymbol{E} = \sum_b x_b^e \boldsymbol{w}_b, \qquad \text{on } \Omega_e. \tag{2.45}$$

The coefficients x_b^e can be interpreted as the moments or degrees of freedoms of the finite element that are unknown. Furthermore, this expressed electric field must satisfy the PEC boundary condition in (2.37) and $\boldsymbol{E} \in \boldsymbol{H}_E(\operatorname{curl}, \Omega)$. This requirement

gives a constraint for the coefficients x_b^e and basis functions $\boldsymbol{w}_b$. When satisfied, the contribution of each element can be assembled globally. Then, the electric field can be written in terms of all basis functions $\boldsymbol{w}_b$ from all elements as

$$\boldsymbol{E} = \sum_b x_b \boldsymbol{w}_b, \qquad \text{on } \Omega, \tag{2.46}$$

where x_b are the corresponding global scaling coefficients. The next step is to choose a sufficient, finite-dimensional set of testing functions. This is done by following the Petrov-Galerkin method, which uses the same functions for approximating the solution as for the testing functions. Then, an arbitrary testing function $\boldsymbol{w} \in \boldsymbol{H}_E(\mathrm{curl}, \Omega)$ is also expressible as the linear combination of $\boldsymbol{w}_a \in \mathcal{W}_p$ basis functions with the appropriate constant coefficients y_a as

$$\boldsymbol{w} = \sum_a y_a \boldsymbol{w}_a, \qquad \text{on } \Omega. \tag{2.47}$$

With the previously discussed discretization, the weak formulation results in a finite element equation system

$$\left(\boldsymbol{S}^{\mu} + jk_0\eta_0 \boldsymbol{T}^{\sigma} - k_0^2 \boldsymbol{T}^{\varepsilon} + \sum_i \frac{jk_0\eta_0}{Z_i} \boldsymbol{Z}^i \right) \boldsymbol{x} = \frac{2jk_0\eta_0}{Z_k} \boldsymbol{b}_{E_k}, \tag{2.48}$$

where the vectors and matrices are calculated as

$$[\boldsymbol{S}^{\mu}]_{ab} = \int_{\Omega} \nabla \times \boldsymbol{w}_a \cdot \boldsymbol{\mu}_r^{-1} \nabla \times \boldsymbol{w}_b \, \mathrm{d}\boldsymbol{r}, \tag{2.49a}$$

$$[\boldsymbol{T}^{\sigma}]_{ab} = \int_{\Omega} \boldsymbol{w}_a \cdot \boldsymbol{\sigma} \boldsymbol{w}_b \, \mathrm{d}\boldsymbol{r}, \tag{2.49b}$$

$$[\boldsymbol{T}^{\varepsilon}]_{ab} = \int_{\Omega} \boldsymbol{w}_a \cdot \boldsymbol{\varepsilon}_r \boldsymbol{w}_b \, \mathrm{d}\boldsymbol{r}, \tag{2.49c}$$

$$\left[\boldsymbol{Z}^i\right]_{ab} = \int_{\Gamma_i} (\boldsymbol{w}_a \times \boldsymbol{n}) \cdot (\boldsymbol{w}_b \times \boldsymbol{n}) \, \mathrm{d}\boldsymbol{r}, \tag{2.49d}$$

$$[\boldsymbol{x}]_b = x_b, \tag{2.49e}$$

$$[\boldsymbol{b}_{E_k}]_b = \int_{\Gamma_k} (\boldsymbol{w}_b \times \boldsymbol{n}) \cdot \left(\boldsymbol{E}_k^{\mathrm{inc}} \times \boldsymbol{n}\right) \, \mathrm{d}\boldsymbol{r}. \tag{2.49f}$$

Here, $\boldsymbol{S}^{\mu}$ is the stiffness matrix, $\boldsymbol{T}^{\sigma}$ and $\boldsymbol{T}^{\varepsilon}$ are mass matrices of different material properties, $\boldsymbol{Z}^i$ is the so-called impedance matrix, $\boldsymbol{b}_{E_k}$ is the excitation vector, and $\boldsymbol{x}$ is the unknown solution vector. In practice, these integrals are evaluated element by element, resulting in element matrices. Then, the global matrices are obtained by the assembly of the element-wise contributions. For the sake of completeness, the scattering parameters are calculated as

$$S_{ik} = \frac{1}{Z_i} \boldsymbol{x} \cdot \boldsymbol{b}_{E_i} - \delta_{ik}. \tag{2.50}$$

In the following, two subcases of the previously described system are considered. These formulations are used to test various basis functions.

The first is the lossless driven case, where the conductivity is zero $\boldsymbol{\sigma} = \mathbf{0}$. Hence, the corresponding matrix vanishes, $\boldsymbol{M}^{\sigma} = \mathbf{0}$, and the finite element equation system becomes

$$\left(\boldsymbol{S}^{\mu} - k_0^2 \boldsymbol{T}^{\varepsilon} + \sum_i \frac{j k_0 \eta_0}{Z_i} \boldsymbol{Z}^i \right) \boldsymbol{x} = \frac{2 j k_0 \eta_0}{Z_k} \boldsymbol{b}_{E_k}. \tag{2.51}$$

In the second case, apart from non-conductive materials, no excitations are considered. Hence, the corresponding matrices $\boldsymbol{M}^{\sigma} = \boldsymbol{Z}^i = \mathbf{0}$ and the excitation vector $\boldsymbol{b}_{E_k} = \mathbf{0}$ become zero. Thus, the finite element equation system becomes a generalized algebraic eigenvalue problem,

$$\left(\boldsymbol{S}^{\mu} - k_0^2 \boldsymbol{T}^{\varepsilon} \right) \boldsymbol{x} = \mathbf{0}, \tag{2.52}$$

where k_0^2 is the eigenvalue, and $\boldsymbol{x}$ is its associated eigenvector. This problem corresponds to the resonant mode and frequency calculation of a cavity resonator.

2.5 Geometry Mapping Invariance

In a general finite element mesh, different types of elements with different geometry deformations may occur. A practical way of handling all elements is to work on a so-called local or reference element and map it to each different global element. This procedure is efficient from a computational point of view, as it allows the construction of precalculated matrices and the uniform application of finite element basis functions on various elements. However, the conformity and de Rham complex of the approximation spaces must also be preserved on the mapped elements. In order to investigate this issue, formal and general definitions of the moments used for finite elements are given. The interpolation operators are then induced by the moments, which are the theoretical tools of the finite element approximation process. In a last step, the local definitions of the finite elements are given with the corresponding mappings. It is proven in this section that when moments are used in a given general form, neither the conformity nor the de Rham complex of the approximation spaces is affected. Moreover, the moments are invariant of the geometry mapping. In later chapters, several different finite element spaces are introduced as particular cases of the general theory, inheriting all properties from the generic case. Note that the overall consequences are already known [Cia02, Chapter 4.3] without general proofs. Similar theories exist for the first-order elements [BBF13, Chapter 2], [FGM11], as well as for higher-order elements with straight-sided geometries [ABDG98], [M$^+$03, Chapter 8.2].

2.5.1 The Interpolation Operator

The interpolation operator π is a useful theoretical tool that allows the finite element analysis of fields v that do not belong to the approximation space $v \notin P$ (in Definition 2.1). It is a projection operator to the approximation space, allowing the projection (interpolant) to be expressed as the linear combination of basis functions $v_a \in P$,

$$\pi v = \sum_a M_a(v) v_a = \sum_a \alpha_a v_a. \tag{2.53}$$

In the case of $v \in P$, the interpolant is equivalent to the original function

$$\pi v = v. \tag{2.54}$$

Historically, pointwise interpolation operators [Sch64], that match an original function and its interpolant in some interpolation locations, were first developed. Later, different interpolation operators were introduced involving the function integrals [M$^+$03, Chapter 5].

The unit tangential vector is denoted by $\boldsymbol{t}$ on an edge or on a face, and the unit normal vector on a face is denoted by $\boldsymbol{n}$. The interior domain of an edge, a face, and volume is denoted by E, F, and V, respectively.

The moments of a function $v \in H^1(\Omega)$ are denoted by $M_p^{(\cdot)}(v)$. These moments are associated to the node $M_p^{vn}(v)$, edge $M_p^{ve}(v)$, face $M_p^{vf}(v)$, or to the volume $M_p^{vv}(v)$. Their formal definitions read

$$M_p^{vn}(v) = v\Big|_N, \tag{2.55a}$$

$$M_p^{ve}(v) = \int_E v\,(\boldsymbol{\omega}\cdot\boldsymbol{t})\,\mathrm{d}\boldsymbol{r}, \qquad \forall\,(\boldsymbol{\omega}\cdot\boldsymbol{t}) \in \mathcal{Q}_p^{ve} \subset C^p(E), \tag{2.55b}$$

$$M_p^{vf}(v) = \int_F v\,(\boldsymbol{b}\cdot\boldsymbol{n})\,\mathrm{d}\boldsymbol{r}, \qquad \forall\,(\boldsymbol{b}\cdot\boldsymbol{n}) \in \mathcal{Q}_p^{vf} \subset C^p(F), \tag{2.55c}$$

$$M_p^{vv}(v) = \int_V v\,q\,\mathrm{d}\boldsymbol{r}, \qquad \forall\,q \in \mathcal{Q}_p^{vv} \subset C^p(V), \tag{2.55d}$$

where the spaces of edge-, face-, and volume-associated test functions are $\mathcal{Q}_p^{ve}$, $\mathcal{Q}_p^{vf}$, and $\mathcal{Q}_p^{vv}$, respectively. Here, the first superscript indicates the space, and the second denotes the geometric object. The test functions are constructed from the appropriate coordinate vectors and functions from the approximation spaces $\boldsymbol{\omega} \in \mathcal{W}_p \subset \boldsymbol{H}(\mathrm{curl}, \Omega_e)$, $\boldsymbol{b} \in \mathcal{D}_p \subset \boldsymbol{H}(\mathrm{div}, \Omega_e)$, and $q \in \mathcal{P}_p \subset L^2(\Omega)$. The space of test functions and the approximation space $\mathcal{V}_p$ are assumed to induce a unisolvent $H^1(\Omega)$-conforming finite element. Then, the definition of the elementwise interpolation operator of order p reads

$$M_p^{(\cdot)}\left(v - \pi_p^v v\right) = 0, \qquad \forall\,(\cdot) \in \{vn, ve, vf, vv\}. \tag{2.56}$$

The global interpolation operator $\pi_{p,h}^v$ of order p and mesh parameter h is defined piecewise on each element $(\pi_{p,h}^v v)|_{\Omega_e} = \pi_p^v(v|_{\Omega_e})$. For all other function spaces, the global interpolation operator can be obtained in the same way.

Next, moments are defined for approximating functions $\boldsymbol{w} \in \boldsymbol{H}(\mathrm{curl}, \Omega)$ via the approximation space $\mathcal{W}_p$. The moments are assigned to the edge $M_p^{we}(\boldsymbol{w})$, face $M_p^{wf}(\boldsymbol{w})$, and volume $M_p^{wv}(\boldsymbol{w})$ as

$$M_p^{we}(\boldsymbol{w}) = \int_E \boldsymbol{w}\cdot(u\,\boldsymbol{t})\,\mathrm{d}\boldsymbol{r}, \qquad \forall\,(u\,\boldsymbol{t}) \in \mathcal{Q}_p^{we} \subset \boldsymbol{C}^p, \tag{2.57a}$$

$$M_p^{wf}(\boldsymbol{w}) = \int_F \boldsymbol{w}\cdot(\boldsymbol{\omega}\times\boldsymbol{n})\,\mathrm{d}\boldsymbol{r}, \qquad \forall\,(\boldsymbol{\omega}\times\boldsymbol{n}) \in \mathcal{Q}_p^{wf} \subset \boldsymbol{C}^p(F), \tag{2.57b}$$

$$M_p^{wv}(\boldsymbol{w}) = \int_V \boldsymbol{w}\cdot\boldsymbol{b}\,\mathrm{d}\boldsymbol{r}, \qquad \forall\,\boldsymbol{b} \in \mathcal{Q}_p^{wv} \subset \boldsymbol{C}^p(V). \tag{2.57c}$$

Here, the spaces of edge-, face-, and volume-associated test functions are $\mathcal{Q}_p^{we}$, $\mathcal{Q}_p^{wf}$, and $\mathcal{Q}_p^{wv}$, respectively. The test functions are constructed from the coordinate vectors and functions of the approximation spaces $u \in \mathcal{V}_p \subset H^1(\Omega_e)$, $\boldsymbol{\omega} \in \mathcal{W}_p \subset \boldsymbol{H}(\mathrm{curl}, \Omega_e)$, and $\boldsymbol{b} \in \mathcal{D}_p \subset \boldsymbol{H}(\mathrm{div}, \Omega_e)$. Again, it is assumed that the space of test functions and the approximation space $\mathcal{W}_p$ induce a unisolvent finite element. Then, the corresponding elementwise interpolation operator π_p^w of order p is defined as

$$M_p^{(\cdot)}\left(\boldsymbol{w} - \pi_p^w \boldsymbol{w}\right) = 0, \qquad \forall\,(\cdot) \in \{we, wf, wv\}. \tag{2.58}$$

Finally, the moments of function $\boldsymbol{d} \in \boldsymbol{H}(\mathrm{div}, \Omega)$ are considered. They are assigned to the faces $M_p^{df}(\boldsymbol{d})$ and to the volume $M_p^{dv}(\boldsymbol{d})$ as

$$M_p^{df}(\boldsymbol{d}) = \int_F \boldsymbol{d} \cdot (u\,\boldsymbol{n})\,\mathrm{d}\boldsymbol{r}, \qquad \forall\,(u\,\boldsymbol{n}) \in \mathcal{Q}_p^{df} \subset \boldsymbol{C}^p(F), \tag{2.59a}$$

$$M_p^{dv}(\boldsymbol{d}) = \int_V \boldsymbol{d} \cdot \boldsymbol{\omega}\,\mathrm{d}\boldsymbol{r}, \qquad \forall\,\boldsymbol{\omega} \in \mathcal{Q}_p^{dv} \subset \boldsymbol{C}^p(V). \tag{2.59b}$$

Here, the space of face- and volume-associated test functions are $\mathcal{Q}_p^{df}$ and $\mathcal{Q}_p^{dv}$, respectively. The test functions consist of coordinate vectors and functions of the approximation spaces $u \in \mathcal{V}_p^f \subset H^1(\Omega_e)$ and $\boldsymbol{\omega} \in \mathcal{W}_p \subset \boldsymbol{H}(\mathrm{curl}, \Omega_e)$. Assuming that the moments and the approximation space $\mathcal{D}_p$ induce a unisolvent finite element, the corresponding elementwise interpolation operator π_p^d is defined as

$$M_p^{(\cdot)}\left(\boldsymbol{d} - \pi_p^d \boldsymbol{d}\right) = 0, \qquad \forall\,(\cdot) \in \{df, dv\}. \tag{2.60}$$

2.5.2 Function Mappings

In order to use the same basis functions on each element, it is necessary to define a so-called reference or local element. Then, the functions of each global element are obtained as the appropriate map of the functions of the reference element. Let $\widehat{\Omega}_e \subset \mathbb{R}^3$ denote a reference tetrahedron, cube, or extruded triangle (prism) with a coordinate vector $\widehat{\boldsymbol{r}} = [\widehat{r}_1, \widehat{r}_2, \widehat{r}_3]^T$. The corresponding geometry mapping $\boldsymbol{g} : \widehat{\boldsymbol{r}} \to \boldsymbol{r}$ is required to be continuously differentiable, invertible, and bijective over $\widehat{\Omega}_e$. Thus, the inverse mapping $\boldsymbol{g}^{-1} : \boldsymbol{r} \to \widehat{\boldsymbol{r}}$ exists and is bijective, too. Further on, each quantity that belongs to the reference element is denoted by $\widehat{(\cdot)}$. Consequently, the operator $\widehat{\nabla} = [\frac{\partial}{\partial \widehat{r}_1}, \frac{\partial}{\partial \widehat{r}_2}, \frac{\partial}{\partial \widehat{r}_3}]^T$ denotes the gradient, $\widehat{\nabla}\times$ the curl, and $\widehat{\nabla}\cdot$ the divergence operator on $\widehat{\Omega}_e$.

As detailed in [RT77], [Néd80], [Hip99], [BBF13, Section 2.1.3], the natural way of transforming functions $\widehat{v} \in H^1(\widehat{\Omega}_e)$, $\widehat{\boldsymbol{w}} \in \boldsymbol{H}(\mathrm{curl}, \widehat{\Omega}_e)$, $\widehat{\boldsymbol{d}} \in \boldsymbol{H}(\mathrm{div}, \widehat{\Omega}_e)$, and $\widehat{q} \in L^2(\widehat{\Omega}_e)$ is via the Piola transformations

$$v = P_0\left\{\widehat{v}\right\} = \widehat{v} \circ \boldsymbol{g}^{-1}, \tag{2.61a}$$

$$\boldsymbol{w} = P_1\left\{\widehat{\boldsymbol{w}}\right\} = \left(\boldsymbol{J}^{-1}\widehat{\boldsymbol{w}}\right) \circ \boldsymbol{g}^{-1}, \tag{2.61b}$$

$$\boldsymbol{d} = P_2\left\{\widehat{\boldsymbol{d}}\right\} = \left(\det(\boldsymbol{J})^{-1}\boldsymbol{J}^T\widehat{\boldsymbol{d}}\right) \circ \boldsymbol{g}^{-1}, \tag{2.61c}$$

$$q = P_3\left\{\widehat{q}\right\} = \left(\det(\boldsymbol{J})^{-1}\widehat{q}\right) \circ \boldsymbol{g}^{-1}, \tag{2.61d}$$

where the Jacobian of mapping $\boldsymbol{g}$ is

$$\boldsymbol{J} := \left[\widehat{\nabla}\,\boldsymbol{r}^T\right] = \left[\widehat{\nabla}\,\boldsymbol{g}\,(\widehat{\boldsymbol{r}})^T\right]. \tag{2.62}$$

Since $\boldsymbol{J}^{-1}$, $\det(\boldsymbol{J})^{-1}\boldsymbol{J}^T$, and $\det(\boldsymbol{J})^{-1}$ are invertible on $\widehat{\Omega}_e$, the corresponding inverse transformations are

$$\widehat{v} = P_0^{-1}\left\{v\right\} = v \circ \boldsymbol{g}, \tag{2.63a}$$

$$\widehat{\boldsymbol{w}} = P_1^{-1}\left\{\boldsymbol{w}\right\} = \boldsymbol{J}\,(\boldsymbol{w} \circ \boldsymbol{g}), \tag{2.63b}$$

$$\widehat{\boldsymbol{d}} = P_2^{-1}\left\{\boldsymbol{d}\right\} = \det(\boldsymbol{J})\boldsymbol{J}^{-T}\,(\boldsymbol{d} \circ \boldsymbol{g}), \tag{2.63c}$$

$$\widehat{q} = P_3^{-1}\left\{q\right\} = \det(\boldsymbol{J})\,(q \circ \boldsymbol{g}). \tag{2.63d}$$

Using Piola transformations, the de Rham complex is satisfied because they preserve the corresponding differential operators,

$$\nabla v = P_1\left\{\widehat{\nabla}\widehat{v}\right\}, \tag{2.64a}$$

$$\nabla \times \boldsymbol{w} = P_2\left\{\widehat{\nabla} \times \widehat{\boldsymbol{w}}\right\}, \tag{2.64b}$$

$$\nabla \cdot \boldsymbol{d} = P_3\left\{\widehat{\nabla} \cdot \widehat{\boldsymbol{d}}\right\}, \tag{2.64c}$$

independently of the geometry mapping. Hence, the null spaces are also preserved:

$$\nabla \times \nabla v = P_2\left\{\widehat{\nabla} \times \widehat{\nabla}\widehat{v}\right\} = P_2\left\{\boldsymbol{0}\right\} = \boldsymbol{0}, \tag{2.65a}$$

$$\nabla \cdot \nabla \times \boldsymbol{w} = P_3\left\{\widehat{\nabla} \cdot \widehat{\nabla} \times \widehat{\boldsymbol{w}}\right\} = P_3\left\{0\right\} = 0. \tag{2.65b}$$

Consequently, it is permissible to construct finite elements and perform differential operations on the reference domain and map them to the global elements [M$^+$03, Chapter 3.9]. The functions of the global approximation spaces originate from the reference element:

$$u = P_0\left\{\widehat{u}\right\} \in \mathcal{V}_p, \qquad \forall \widehat{u} \in \widehat{\mathcal{V}}_p, \tag{2.66a}$$

$$\boldsymbol{\omega} = P_1\left\{\widehat{\boldsymbol{\omega}}\right\} \in \mathcal{W}_p, \qquad \forall \widehat{\boldsymbol{\omega}} \in \widehat{\mathcal{W}}_p, \tag{2.66b}$$

$$\boldsymbol{d} = P_2\left\{\widehat{\boldsymbol{d}}\right\} \in \mathcal{D}_p, \qquad \forall \widehat{\boldsymbol{d}} \in \widehat{\mathcal{D}}_p, \tag{2.66c}$$

$$q = P_3\left\{\widehat{q}\right\} \in \mathcal{P}_p, \qquad \forall \widehat{q} \in \widehat{\mathcal{P}}_p. \tag{2.66d}$$

2.5.3 Invariance of the Moments

In this subsection, it will be shown that the previously presented moments remain unchanged for all bijective and continuously differentiable geometry mappings, including the curvilinear case.

An edge E and a face F of an element can always be parametrized by the local coordinates. The orientation and the shape of a local element can be chosen such that a given edge E is parametrized by the variable $\widehat{r}_1$ over the interval of $[0,1]$. Thus, integration over the edge in the reference element requires the edge element

$$\left| \frac{\partial \boldsymbol{r}}{\partial \widehat{r}_1} \right| = \left| \boldsymbol{J}^T \widehat{\boldsymbol{t}} \right|, \tag{2.67}$$

where $\widehat{\boldsymbol{t}} = [1,0,0]^T$ is the unit tangential vector of the straight edge on the reference element. Hence, the unit length tangential vector of a global edge, in general, can be written as

$$\boldsymbol{t} = \left(\frac{\boldsymbol{J}^T \widehat{\boldsymbol{t}}}{\left| \boldsymbol{J}^T \widehat{\boldsymbol{t}} \right|} \right) \circ \boldsymbol{g}^{-1}. \tag{2.68}$$

Furthermore, the local parametrization of the element can be chosen such that a given face F is parametrized by the local coordinates $\widehat{r}_1$ and $\widehat{r}_2$ within the interval of $[0,1]^2$. The related surface element of this parameterization reads

$$\left| \frac{\partial \boldsymbol{r}}{\partial \widehat{r}_1} \times \frac{\partial \boldsymbol{r}}{\partial \widehat{r}_2} \right|. \tag{2.69}$$

Using vector identities and the Piola transforms, the surface element can be rewritten as

$$\left| \frac{\partial \boldsymbol{r}}{\partial \widehat{r}_1} \times \frac{\partial \boldsymbol{r}}{\partial \widehat{r}_2} \right| = \left| \det(\boldsymbol{J}) \boldsymbol{J}^{-1} \left(\frac{\partial \widehat{\boldsymbol{r}}}{\partial \widehat{r}_1} \times \frac{\partial \widehat{\boldsymbol{r}}}{\partial \widehat{r}_2} \right) \right| = |\det(\boldsymbol{J})| \left| \boldsymbol{J}^{-1} \widehat{\boldsymbol{n}} \right|, \tag{2.70}$$

where $\widehat{\boldsymbol{n}} = [1,0,0]^T$ is the unit normal vector on the planar face of the reference element. Hence, the natural expression of a normal vector on a face F, in general, is

$$\boldsymbol{n} = \left(\frac{\boldsymbol{J}^{-1} \widehat{\boldsymbol{n}}}{\left| \boldsymbol{J}^{-1} \widehat{\boldsymbol{n}} \right|} \right) \circ \boldsymbol{g}^{-1}. \tag{2.71}$$

As bijective geometry mappings have been assumed, the Jacobian is always invertible over the reference element. Thus, choosing a right-handed orientation for the local coordinates the volume element becomes

$$|\det(\boldsymbol{J})| = \det(\boldsymbol{J}) > 0, \qquad \text{on } \widehat{\Omega}_e. \tag{2.72}$$

Next, the invariance of the discussed moments is shown. It is possible to evaluate the integrals of (2.55) over the reference element by changing the variables to the local coordinates. Hence, these moments are also associated with the nodes $\widehat{N}$, edges $\widehat{E}$, faces $\widehat{F}$, and volume $\widehat{V}$ of the reference element. The functions of the reference element (2.61), thus the functions of the approximation spaces (2.66), are transformed via the Piola transformations. One can show that these mapped integrals result in local moments $\widehat{M}_p^{vn}(\widehat{v})$, $\widehat{M}_p^{ve}(\widehat{v})$, $\widehat{M}_p^{vf}(\widehat{v})$, and $\widehat{M}_p^{vv}(\widehat{v})$ that are independent of the geometry mapping:

$$M_p^{vn}(v) = \widehat{v}\Big|_{\widehat{N}} = \widehat{M}_p^{vn}(\widehat{v}), \tag{2.73a}$$

$$M_p^{ve}(v) = \int_{\widehat{E}} \widehat{v} \left(\boldsymbol{J}^{-1}\widehat{\boldsymbol{\omega}} \cdot \frac{\boldsymbol{J}^T\widehat{\boldsymbol{t}}}{\left|\boldsymbol{J}^T\widehat{\boldsymbol{t}}\right|} \right) \left|\boldsymbol{J}^T\widehat{\boldsymbol{t}}\right| \,\mathrm{d}\widehat{\boldsymbol{r}} = \int_{\widehat{E}} \widehat{v}\,(\widehat{\boldsymbol{\omega}} \cdot \widehat{\boldsymbol{t}})\,\mathrm{d}\widehat{\boldsymbol{r}} = \widehat{M}_p^{ve}(\widehat{v}), \tag{2.73b}$$

$$\begin{aligned} M_p^{vf}(v) &= \int_{\widehat{F}} \widehat{v} \left(\det(\boldsymbol{J})^{-1}\boldsymbol{J}^T\widehat{\boldsymbol{b}} \cdot \frac{\boldsymbol{J}^{-1}\widehat{\boldsymbol{n}}}{\left|\boldsymbol{J}^{-1}\widehat{\boldsymbol{n}}\right|} \right) |\det(\boldsymbol{J})| \left|\boldsymbol{J}^{-1}\widehat{\boldsymbol{n}}\right| \,\mathrm{d}\widehat{\boldsymbol{r}}, \\ &= \int_{\widehat{F}} \widehat{v}\,(\widehat{\boldsymbol{b}} \cdot \widehat{\boldsymbol{n}})\,\mathrm{d}\widehat{\boldsymbol{r}} = \widehat{M}_p^{vf}(\widehat{v}), \end{aligned} \tag{2.73c}$$

$$M_p^{vv}(v) = \int_{\widehat{V}} \widehat{v}\,(\det(\boldsymbol{J})^{-1}\widehat{q})\,|\det(\boldsymbol{J})|\,\mathrm{d}\widehat{\boldsymbol{r}} = \int_{\widehat{V}} \widehat{v}\,\widehat{q}\,\mathrm{d}\widehat{\boldsymbol{r}} = \widehat{M}_p^{vv}(\widehat{v}). \tag{2.73d}$$

Here, $\widehat{\boldsymbol{t}}$ is the unit length tangential vector of the straight edge $\widehat{E}$, and $\widehat{\boldsymbol{n}}$ is the unit length normal vector of a face $\widehat{F}$. Similarly to the case of the global element, the functions $\widehat{u}$, $\widehat{\boldsymbol{\omega}}$, $\widehat{\boldsymbol{b}}$, and $\widehat{q}$ belong to the local approximation spaces of $\widehat{\mathcal{V}}_p$, $\widehat{\mathcal{W}}_p$, $\widehat{\mathcal{D}}_p$, and $\widehat{\mathcal{P}}_p$, respectively.

Using a similar approach as in the $H^1(\widehat{\Omega}_e)$ case, it can be shown that the $\boldsymbol{H}(\mathrm{curl}, \widehat{\Omega}_e)$ moments associated with the edge $\widehat{M}_p^{we}(\widehat{\boldsymbol{w}})$, face $\widehat{M}_p^{wf}(\widehat{\boldsymbol{w}})$, and volume $\widehat{M}_p^{wv}(\widehat{\boldsymbol{w}})$ over the reference domain are identical to the moments of the global element:

$$M_p^{we}(\boldsymbol{w}) = \int_{\widehat{E}} \boldsymbol{J}^{-1}\widehat{\boldsymbol{w}} \cdot \left(\widehat{u}\,\frac{\boldsymbol{J}^T\widehat{\boldsymbol{t}}}{\left|\boldsymbol{J}^T\widehat{\boldsymbol{t}}\right|} \right) \left|\boldsymbol{J}^T\widehat{\boldsymbol{t}}\right| \,\mathrm{d}\widehat{\boldsymbol{r}} = \int_{\widehat{E}} \widehat{\boldsymbol{w}} \cdot (\widehat{u}\,\widehat{\boldsymbol{t}})\,\mathrm{d}\widehat{\boldsymbol{r}} = \widehat{M}_p^{we}(\widehat{\boldsymbol{w}}), \tag{2.74a}$$

$$\begin{aligned} M_p^{wf}(\boldsymbol{w}) &= \int_{\widehat{F}} \boldsymbol{J}^{-1}\widehat{\boldsymbol{w}} \cdot \left(\boldsymbol{J}^{-1}\widehat{\boldsymbol{\omega}} \times \frac{\boldsymbol{J}^{-1}\widehat{\boldsymbol{n}}}{\left|\boldsymbol{J}^{-1}\widehat{\boldsymbol{n}}\right|} \right) |\det(\boldsymbol{J})| \left|\boldsymbol{J}^{-1}\widehat{\boldsymbol{n}}\right| \,\mathrm{d}\widehat{\boldsymbol{r}}, \\ &= \int_{\widehat{F}} \widehat{\boldsymbol{w}} \cdot (\widehat{\boldsymbol{\omega}} \times \widehat{\boldsymbol{n}})\,\mathrm{d}\widehat{\boldsymbol{r}} = \widehat{M}_p^{wf}(\widehat{\boldsymbol{w}}), \end{aligned} \tag{2.74b}$$

$$M_p^{wv}(\boldsymbol{w}) = \int_{\widehat{V}} \boldsymbol{J}^{-1}\widehat{\boldsymbol{w}} \cdot (\det(\boldsymbol{J})^{-1}\boldsymbol{J}^T\widehat{\boldsymbol{b}})\,|\det(\boldsymbol{J})|\;\mathrm{d}\widehat{\boldsymbol{r}} = \int_{\widehat{V}} \widehat{\boldsymbol{w}} \cdot \widehat{\boldsymbol{b}}\,\mathrm{d}\widehat{\boldsymbol{r}} = \widehat{M}_p^{wv}(\widehat{\boldsymbol{w}}). \tag{2.74c}$$

Moreover, the $\boldsymbol{H}(\mathrm{div}, \widehat{\Omega}_e)$ moments on the reference element also become independent of the geometry mapping. These local moments are associated with the face $\widehat{M}_p^{df}(\widehat{\boldsymbol{d}})$ and the volume $\widehat{M}_p^{dv}(\widehat{\boldsymbol{d}})$ as

$$\begin{aligned} M_p^{df}(\boldsymbol{d}) &= \int_{\widehat{F}} \det(\boldsymbol{J})^{-1}\boldsymbol{J}^T\widehat{\boldsymbol{d}} \cdot \left(\widehat{u}\,\frac{\boldsymbol{J}^{-1}\widehat{\boldsymbol{n}}}{\left|\boldsymbol{J}^{-1}\widehat{\boldsymbol{n}}\right|}\right) |\det(\boldsymbol{J})|\left|\boldsymbol{J}^{-1}\widehat{\boldsymbol{n}}\right|\;\mathrm{d}\widehat{\boldsymbol{r}}, \\ &= \int_{\widehat{F}} \widehat{\boldsymbol{d}} \cdot (\widehat{u}\,\widehat{\boldsymbol{n}})\,\mathrm{d}\widehat{\boldsymbol{r}} = \widehat{M}_p^{df}(\widehat{\boldsymbol{d}}), \end{aligned} \tag{2.75a}$$

$$\begin{aligned} M_p^{dv}(\boldsymbol{d}) &= \int_{\widehat{V}} \det(\boldsymbol{J})^{-1}\boldsymbol{J}^T\widehat{\boldsymbol{d}} \cdot (\boldsymbol{J}^{-1}\widehat{\boldsymbol{\omega}})\,|\det(\boldsymbol{J})|\;\mathrm{d}\widehat{\boldsymbol{r}} \\ &= \int_{\widehat{V}} \widehat{\boldsymbol{d}} \cdot \widehat{\boldsymbol{\omega}}\,\mathrm{d}\widehat{\boldsymbol{r}} = \widehat{M}_p^{dv}(\widehat{\boldsymbol{d}}). \end{aligned} \tag{2.75b}$$

The corresponding interpolation operators over the reference element are defined as

$$\widehat{M}_p^{(\cdot)}\left(\widehat{v} - \widehat{\pi}_p^v\widehat{v}\right) = 0, \qquad \forall\,(\cdot) \in \{vn, ve, vf, vv\}, \tag{2.76a}$$

$$\widehat{M}_p^{(\cdot)}\left(\widehat{\boldsymbol{w}} - \widehat{\pi}_p^w\widehat{\boldsymbol{w}}\right) = 0, \qquad \forall\,(\cdot) \in \{we, wf, wv\}, \tag{2.76b}$$

$$\widehat{M}_p^{(\cdot)}\left(\widehat{\boldsymbol{d}} - \widehat{\pi}_p^d\widehat{\boldsymbol{d}}\right) = 0, \qquad \forall\,(\cdot) \in \{df, dv\}. \tag{2.76c}$$

Since the moments are independent of the geometry mapping, the relations between the local and global interpolants are given by the Piola transformations:

$$\pi_p^v v = P_0\left\{\widehat{\pi}_p^v\widehat{v}\right\}, \tag{2.77a}$$

$$\pi_p^w \boldsymbol{w} = P_1\left\{\widehat{\pi}_p^w\widehat{\boldsymbol{w}}\right\}, \tag{2.77b}$$

$$\pi_p^d \boldsymbol{d} = P_2\left\{\widehat{\pi}_p^d\widehat{\boldsymbol{d}}\right\}. \tag{2.77c}$$

2.5.4 Commutation Properties of the Approximation Spaces

Let the approximation spaces be defined via the interpolation operators as

$$\mathcal{V}_p = \left\{ u = \pi_p^v v \,|\, \forall v \in H^1(\Omega_e) \right\}, \tag{2.78a}$$
$$\mathcal{W}_p = \left\{ \boldsymbol{\omega} = \pi_p^w \boldsymbol{w} \,|\, \forall \boldsymbol{w} \in \boldsymbol{H}(\mathrm{curl}, \Omega_e) \right\}, \tag{2.78b}$$
$$\mathcal{D}_p = \left\{ \boldsymbol{b} = \pi_p^d \boldsymbol{d} \,|\, \forall \boldsymbol{d} \in \boldsymbol{H}(\mathrm{div}, \Omega_e) \right\}. \tag{2.78c}$$

In this section, it is shown that the following commutative diagram holds,

$$\begin{array}{ccccc} H^1(\Omega_e) & \xrightarrow{\nabla} & \boldsymbol{H}(\mathrm{curl}, \Omega_e) & \xrightarrow{\nabla\times} & \boldsymbol{H}(\mathrm{div}, \Omega_e) \\ \downarrow{\scriptstyle \pi_p^v} & & \downarrow{\scriptstyle \pi_p^w} & & \downarrow{\scriptstyle \pi_p^d} \\ \mathcal{V}_p & \xrightarrow{\nabla} & \mathcal{W}_p & \xrightarrow{\nabla\times} & \widehat{\mathcal{D}}_p \end{array} \tag{2.79}$$

provided that $\mathcal{W}_p$ and $\mathcal{D}_p$, respectively, includes the irrotational and divergence-free subspaces

$$\nabla \mathcal{V}_p \subset \mathcal{W}_p, \tag{2.80a}$$
$$\nabla \times \mathcal{W}_p \subset \mathcal{D}_p. \tag{2.80b}$$

These commutative relations are required by the approximation estimates in later chapters. Sufficient conditions for the validity of the commutative diagram are

$$\pi_p^w \nabla v = \nabla \pi_p^v v, \tag{2.81a}$$
$$\pi_p^d \nabla \times \boldsymbol{w} = \nabla \times \pi_p^w \boldsymbol{w}. \tag{2.81b}$$

These requirements are difficult to grant directly because the approximation spaces of the different real elements may differ in the curvilinear case due to the varying element shapes. However, the presented general moments are independent of the geometry mapping, resulting in geometry-independent interpolation operators (2.76) as well as approximation spaces over the reference element:

$$\widehat{\mathcal{V}}_p = \left\{ \widehat{u} = \widehat{\pi}_p^v \widehat{v} \,|\, \forall \widehat{v} \in H^1(\widehat{\Omega}_e) \right\}, \tag{2.82a}$$
$$\widehat{\mathcal{W}}_p = \left\{ \widehat{\boldsymbol{\omega}} = \widehat{\pi}_p^w \widehat{\boldsymbol{w}} \,|\, \forall \widehat{\boldsymbol{w}} \in \boldsymbol{H}(\mathrm{curl}, \widehat{\Omega}_e) \right\}, \tag{2.82b}$$
$$\widehat{\mathcal{D}}_p = \left\{ \widehat{\boldsymbol{b}} = \widehat{\pi}_p^d \widehat{\boldsymbol{d}} \,|\, \forall \widehat{\boldsymbol{d}} \in \boldsymbol{H}(\mathrm{div}, \widehat{\Omega}_e) \right\}. \tag{2.82c}$$

With the introduction of the local approximation spaces, the commutative diagram becomes

$$\begin{array}{ccccc}
H^1(\Omega_e) & \xrightarrow{\nabla} & \boldsymbol{H}(\mathrm{curl},\Omega_e) & \xrightarrow{\nabla\times} & \boldsymbol{H}(\mathrm{div},\Omega_e) \\
\downarrow{\scriptstyle P_0^{-1}} & & \downarrow{\scriptstyle P_1^{-1}} & & \downarrow{\scriptstyle P_2^{-1}} \\
H^1(\widehat{\Omega}_e) & \xrightarrow{\widehat{\nabla}} & \boldsymbol{H}(\mathrm{curl},\widehat{\Omega}_e) & \xrightarrow{\widehat{\nabla}\times} & \boldsymbol{H}(\mathrm{div},\widehat{\Omega}_e) \\
\downarrow{\scriptstyle \widehat{\pi}_p^v} & & \downarrow{\scriptstyle \widehat{\pi}_p^w} & & \downarrow{\scriptstyle \widehat{\pi}_p^d} \\
\widehat{\mathcal{V}}_p & \xrightarrow{\widehat{\nabla}} & \widehat{\mathcal{W}}_p & \xrightarrow{\widehat{\nabla}\times} & \widehat{\mathcal{D}}_p \\
\downarrow{\scriptstyle P_0} & & \downarrow{\scriptstyle P_1} & & \downarrow{\scriptstyle P_2} \\
\mathcal{V}_p & \xrightarrow{\nabla} & \mathcal{W}_p & \xrightarrow{\nabla\times} & \mathcal{D}_p
\end{array} \tag{2.83}$$

Therefore, the finite element approximation consists of three theoretical steps:

1. The global fields are mapped to the reference element.
2. The finite element approximation is a projection of the local fields to the local approximation spaces.
3. The local approximations are transformed back to the global element.

Consequently, if the local approximation spaces over the reference element include the appropriate subspaces

$$\widehat{\nabla}\widehat{\mathcal{V}}_p \subset \widehat{\mathcal{W}}_p, \tag{2.84a}$$

$$\widehat{\nabla}\times\widehat{\mathcal{W}}_p \subset \widehat{\mathcal{D}}_p, \tag{2.84b}$$

and the interpolants of the functions $\widehat{v} \in \widehat{\mathcal{V}}_p$ and $\widehat{\boldsymbol{w}} \in \widehat{\mathcal{W}}_p$ satisfy

$$\widehat{\pi}_p^w\widehat{\nabla}\widehat{v} = \widehat{\nabla}\widehat{\pi}_p^v\widehat{v}, \tag{2.85a}$$

$$\widehat{\pi}_p^d\widehat{\nabla}\times\widehat{\boldsymbol{w}} = \widehat{\nabla}\times\widehat{\pi}_p^w\widehat{\boldsymbol{w}}, \tag{2.85b}$$

then both local and global approximation spaces form a de Rham complex.

Note that the last approximation space $\mathcal{P}_p$ of the entire sequence was not considered. This is the appropriate space for approximating functions in $L^2(\Omega)$, which may be discontinuous between the elements. Therefore, no additional functions are needed to maintain continuity between the elements. The analogy between the continuous (2.26) and discrete (2.44) de Rham complex implies that the image of the previous space must be included. Therefore, the approximation spaces on the reference element must satisfy

$$\widehat{\nabla}\cdot\widehat{\mathcal{D}}_p \subset \widehat{\mathcal{P}}_p. \tag{2.86}$$

As both $\widehat{\mathcal{D}}_p$ and $\widehat{\mathcal{P}}_p$ are polynomial spaces, the functions of $\widehat{\nabla}\cdot\widehat{\mathcal{D}}_{p+1}$ are also suitable for $\widehat{\mathcal{P}}_p$.

The Gradient Subspace

This section provides proof for (2.85a), assuming that the gradient subspace is included in the approximation space (2.84a). First, the interpolant $\widehat{\pi}_p^w$ of equation (2.85a) is taken as

$$\widehat{\pi}_p^w \left(\widehat{\pi}_p^w \widehat{\nabla} \widehat{v} \right) = \widehat{\pi}_p^w \left(\widehat{\nabla} \widehat{\pi}_p^v \widehat{v} \right). \tag{2.87}$$

Using the fact that the interpolant of an interpolant becomes $\widehat{\pi}_p^w \widehat{\pi}_p^w(\cdot) = \widehat{\pi}_p^w(\cdot)$, it is sufficient to show that

$$\widehat{\pi}_p^w \left(\widehat{\nabla} \left(\widehat{v} - \widehat{\pi}_p^v \widehat{v} \right) \right) = \mathbf{0}, \qquad \forall \widehat{v}. \tag{2.88}$$

Thus, the $\boldsymbol{H}(\mathrm{curl})$-conforming moments of an arbitrary H^1 interpolation error function $\widehat{v}_{err} = \widehat{v} - \widehat{\pi}_p^v \widehat{v}$ must vanish

$$\widehat{M}_p^{(\cdot)} \left(\widehat{\nabla} \widehat{v}_{err} \right) = 0, \qquad \forall (\cdot) \in \{we, wf, wv\}. \tag{2.89}$$

First, the edge-associated moments are considered:

$$\widehat{M}_p^{we}(\widehat{\nabla} \widehat{v}_{err}) = \int_{\widehat{E}} \widehat{\nabla} \widehat{v}_{err} \cdot (\widehat{u}\, \widehat{\boldsymbol{t}}) \, \mathrm{d}\widehat{\boldsymbol{r}}. \tag{2.90}$$

One can rewrite this integral with the use of the vector identity

$$\widehat{\nabla} \widehat{v}_{err} \cdot \left(\widehat{u}\, \widehat{\boldsymbol{t}} \right) = \widehat{\nabla} \left(\widehat{v}_{err}\, \widehat{u} \right) \widehat{\boldsymbol{t}} - \widehat{v}_{err}\, \widehat{\nabla} \widehat{u} \cdot \widehat{\boldsymbol{t}}, \tag{2.91}$$

and applying the gradient theorem as

$$\widehat{M}_p^{we}(\widehat{\nabla} \widehat{v}_{err}) = \left[\widehat{v}_{err}\, \widehat{u} \right]_{\widehat{N}_1}^{\widehat{N}_2} - \int_{\widehat{E}} \widehat{v}_{err} \left(\widehat{\nabla} \widehat{u} \cdot \widehat{\boldsymbol{t}} \right) \, \mathrm{d}\widehat{\boldsymbol{r}}. \tag{2.92}$$

Here, the first term is equivalent to the node-associated moments (2.73a), that are multiplied by a scalar $\widehat{u}$. Moreover, the second term is equivalent to a moment over the edge (2.73b) since $\widehat{\boldsymbol{\omega}} = \widehat{\nabla} \widehat{u} \in \widehat{\nabla} \widehat{\mathcal{V}}_p \subset \widehat{\mathcal{W}}_p$. Therefore, both terms become zero due to the definition of the corresponding interpolation operator (2.76a) for an error function.

Next, let us continue with the face-associated moments:

$$\widehat{M}_p^{wf}(\widehat{\nabla} \widehat{v}_{err}) = \int_{\widehat{F}} \widehat{\nabla} \widehat{v}_{err} \cdot (\widehat{\boldsymbol{\omega}} \times \widehat{\boldsymbol{n}}) \, \mathrm{d}\widehat{\boldsymbol{r}}. \tag{2.93}$$

One can apply the vector identity

$$\widehat{\nabla}\widehat{v}_{err} \cdot (\widehat{\boldsymbol{\omega}} \times \widehat{\boldsymbol{n}}) = \widehat{\nabla} \times (\widehat{v}_{err}\, \widehat{\boldsymbol{\omega}}) - \widehat{v}_{err} \widehat{\nabla} \times \widehat{\boldsymbol{\omega}} \cdot \widehat{\boldsymbol{n}}, \tag{2.94}$$

and Stokes' theorem to rewrite the moment as a line integral over a closed curve of the edges and a face integral:

$$\widehat{M}_p^{wf}(\widehat{\nabla}\widehat{v}_{err}) = \oint_{\widehat{E}} \widehat{v}_{err} \left(\widehat{\boldsymbol{\omega}} \cdot \widehat{\boldsymbol{t}}\right) \mathrm{d}\widehat{\boldsymbol{r}} - \int_{\widehat{F}} \widehat{v}_{err} \left(\widehat{\nabla} \times \widehat{\boldsymbol{\omega}} \cdot \widehat{\boldsymbol{n}}\right) \mathrm{d}\widehat{\boldsymbol{r}}. \tag{2.95}$$

Here, each line integral is equivalent to the edge-associated moments (2.74a). Furthermore, the surface integral is equivalent to the face-associated moments (2.73c) since $\widehat{\boldsymbol{b}} = \widehat{\nabla} \times \widehat{\boldsymbol{\omega}} \in \widehat{\nabla} \times \widehat{\mathcal{W}}_p \subset \widehat{\mathcal{D}}_p$. Thus, by the definition of the interpolation operator (2.76a), both terms become zero for the error function.

Finally, the case of the volume-associated moments is considered:

$$\widehat{M}_p^{wv}(\widehat{\nabla}\widehat{v}_{err}) = \int_{\widehat{V}} \widehat{\nabla}\widehat{v}_{err} \cdot \widehat{\boldsymbol{b}}\, \mathrm{d}\widehat{\boldsymbol{r}}. \tag{2.96}$$

One can apply the vector identity

$$\widehat{\nabla}\widehat{v}_{err} \cdot \widehat{\boldsymbol{b}} = \widehat{\nabla} \cdot \left(\widehat{v}_{err}\, \widehat{\boldsymbol{b}}\right) - \widehat{v}_{err} \widehat{\nabla} \cdot \widehat{\boldsymbol{b}}, \tag{2.97}$$

and the divergence theorem to rewrite the moment as surface integrals and a volume integral:

$$\widehat{M}_p^{wv}(\widehat{\nabla}\widehat{v}_{err}) = \oint_{\widehat{F}} \widehat{v}_{err} \left(\widehat{\boldsymbol{b}} \cdot \widehat{\boldsymbol{n}}\right) \mathrm{d}\widehat{\boldsymbol{r}} - \int_{\widehat{V}} \widehat{v}_{err} \left(\widehat{\nabla} \cdot \widehat{\boldsymbol{b}}\right) \mathrm{d}\widehat{\boldsymbol{r}}. \tag{2.98}$$

In the first term, each face integral is equivalent to a face-associated moment (2.73c). Moreover, the second term is equivalent to a volume-associated moment (2.73d) with the test function $\widehat{q} = \widehat{\nabla} \cdot \widehat{\boldsymbol{b}} \in \widehat{\nabla} \cdot \widehat{\mathcal{D}}_p \subset \widehat{\mathcal{P}}_p$. Hence, both terms become zero due to the definition of the interpolation operator (2.76a) for the error function.

The Solenoidal Subspace

In this section, proof is given for (2.85b), assuming the inclusion of the appropriate discrete subspace (2.84). First, the operator $\widehat{\pi}_p^d$ is applied to (2.85b) as

$$\widehat{\pi}_p^d \left(\widehat{\pi}_p^d \widehat{\nabla} \times \widehat{\boldsymbol{w}}\right) = \widehat{\pi}_p^d \left(\widehat{\nabla} \times \widehat{\pi}_p^w \widehat{\boldsymbol{w}}\right), \tag{2.99}$$

where the term $\widehat{\pi}_p^d \widehat{\pi}_p^d(\cdot)$ can be simplified to a single interpolation $\widehat{\pi}_p^d(\cdot)$. Hence, it is sufficient to show that the relation

$$\widehat{\pi}_p^d \left(\widehat{\nabla} \times \left(\widehat{\boldsymbol{w}} - \widehat{\pi}_p^w \widehat{\boldsymbol{w}} \right) \right) = \mathbf{0} \tag{2.100}$$

holds for all $\widehat{\boldsymbol{w}}$ functions. Therefore, all $\boldsymbol{H}(\mathrm{div})$-conforming finite element moments must vanish for an error function $\widehat{\boldsymbol{w}}_{err} = \widehat{\boldsymbol{w}} - \widehat{\pi}_p^w \widehat{\boldsymbol{w}}$:

$$\widehat{M}_p^{(\cdot)} \left(\widehat{\nabla} \times \widehat{\boldsymbol{w}}_{err} \right) = 0, \qquad \forall (\cdot) \in \{df, dv\}. \tag{2.101}$$

First, the face-associated moments are considered:

$$\widehat{M}_p^{df} (\widehat{\nabla} \times \widehat{\boldsymbol{w}}_{err}) = \int_{\widehat{F}} \widehat{\nabla} \times \widehat{\boldsymbol{w}}_{err} \cdot (\widehat{u}\, \widehat{\boldsymbol{n}})\, \mathrm{d}\widehat{\boldsymbol{r}}. \tag{2.102}$$

Applying the vector identity

$$\widehat{\nabla} \times \widehat{\boldsymbol{w}}_{err} \cdot (\widehat{u}\, \widehat{\boldsymbol{n}}) = \widehat{\nabla} \times (\widehat{u}\, \widehat{\boldsymbol{w}}_{err}) \cdot \widehat{\boldsymbol{n}} + \widehat{\boldsymbol{w}}_{err} \cdot \left(\widehat{\nabla} \widehat{u} \times \widehat{\boldsymbol{n}} \right), \tag{2.103}$$

as well as Stokes' theorem, the face-associated moments become

$$\widehat{M}_p^{df} (\widehat{\nabla} \times \widehat{\boldsymbol{w}}_{err}) = \oint_{\widehat{E}} \widehat{\boldsymbol{w}}_{err} \cdot \left(\widehat{u}\, \widehat{\boldsymbol{t}} \right) \mathrm{d}\widehat{\boldsymbol{r}} + \int_{\widehat{F}} \widehat{\boldsymbol{w}}_{err} \cdot \left(\widehat{\nabla} \widehat{u} \times \widehat{\boldsymbol{n}} \right) \mathrm{d}\widehat{\boldsymbol{r}}. \tag{2.104}$$

Here, the first term is a line integral over all the edges of a face. Each edge integral is equivalent to the edge-associated moments (2.74a). The second term is a face integral, which is identical to the face-associated moments (2.74b) with $\widehat{\boldsymbol{\omega}} = \widehat{\nabla} \widehat{u} \in \widehat{\nabla} \widehat{\mathcal{V}}_p \subset \widehat{\mathcal{W}}_p$. Using the definition of the interpolation operator (2.76b), both integrals become zero for an error function.

Finally, the volume-associated moments are investigated:

$$\widehat{M}_p^{dv} (\widehat{\nabla} \times \widehat{\boldsymbol{w}}_{err}) = \int_{\widehat{V}} \widehat{\nabla} \times \widehat{\boldsymbol{w}}_{err} \cdot \widehat{\boldsymbol{\omega}}\, \mathrm{d}\widehat{\boldsymbol{r}}. \tag{2.105}$$

One can apply the vector identity

$$\widehat{\nabla} \times \widehat{\boldsymbol{w}}_{err} \cdot \widehat{\boldsymbol{\omega}} = \widehat{\nabla} \cdot (\widehat{\boldsymbol{w}}_{err} \times \widehat{\boldsymbol{\omega}}) + \widehat{\boldsymbol{w}}_{err} \cdot \widehat{\nabla} \times \widehat{\boldsymbol{\omega}}, \tag{2.106}$$

as well as the divergence theorem to rewrite the moment as

$$\widehat{M}_p^{dv} (\widehat{\nabla} \times \widehat{\boldsymbol{w}}_{err}) = \oint_{\widehat{F}} \widehat{\boldsymbol{w}}_{err} \cdot (\widehat{\boldsymbol{\omega}} \times \widehat{\boldsymbol{n}})\, \mathrm{d}\widehat{\boldsymbol{r}} + \int_{\widehat{V}} \widehat{\boldsymbol{w}}_{err} \cdot \widehat{\nabla} \times \widehat{\boldsymbol{\omega}}\, \mathrm{d}\widehat{\boldsymbol{r}}. \tag{2.107}$$

Here, the first boundary integral term is equivalent to the sums of face-associated moments (2.74b). Furthermore, the second volume integral is equivalent to the volume-associated moments (2.74c), with $\widehat{\boldsymbol{b}} = \widehat{\nabla} \times \widehat{\boldsymbol{\omega}} \in \widehat{\nabla} \widehat{\mathcal{W}}_p \subset \widehat{\mathcal{D}}_p$. Hence, both terms become zero due to the definition of the interpolation operator (2.76b) for an error function.

2.6 Tetrahedral Elements

The most used meshes for practical applications consist of simplexes due to their versatile usage and ease of generation for both structured and unstructured geometries. In three dimensions, the best-known tetrahedral finite element is the polynomial H^1-conforming element [BBF13, Section 2.2] for scalar fields. For the approximation of $\boldsymbol{H}(\mathrm{curl})$-conforming vector fields, the complete- [Néd86] and mixed-order Nédélec elements [Néd80] are used. For the approximation of $\boldsymbol{H}(\mathrm{div})$-conforming vector fields, the Raviart-Thomas elements (identical to the mixed-order face elements of Nédélec [Néd80]), the Brezzi-Douglas-Marini elements (identical to the complete-order face elements of Nédélec [Néd86]), and the Brezzi-Douglas-Fortin-Marini elements are used [BBF13, Section 2.3, 2.6]. These finite elements differ in their continuity requirements and in their approximation properties. The present work only considers the smallest spaces, the H^1-conforming scalar polynomial space and the mixed-order Nédélec space, that provide algebraic convergence rates in their corresponding natural norms. In this section, general definitions are given for these spaces, and explicit basis functions of Ingelström [Ing06] are provided for the H^1- and $\boldsymbol{H}(\mathrm{curl})$-conforming cases.

Let us start with the definition of the reference tetrahedron. In terms of local coordinates $\widehat{\boldsymbol{r}}$, the four nodes of the element are $\widehat{\boldsymbol{r}}_1 = [1,0,0]^T$, $\widehat{\boldsymbol{r}}_2 = [0,1,0]^T$, $\widehat{\boldsymbol{r}}_3 = [0,0,1]^T$, and $\widehat{\boldsymbol{r}}_4 = [0,0,0]$. Each of the six straight edges is bounded by two nodes and each flat triangular face is bounded by three edges. The boundary of the element consists of four flat triangular faces. Thus, each node is associated to an index i. Similarly, all edges, faces, and volume can be associated with the indices of the bounding nodes ij, ijk, and $ijkl$, respectively. The reference element of the present configuration is shown in Figure 2.2, where the arrows indicate the orientation of the edges. For all of the discussed tetrahedral elements, the local and global ordering of the nodes and the orientation of the edges and faces are determined by the node indices in ascending order.

In order to construct basis functions with compact support and symmetry, it is common to introduce barycentric coordinates $\boldsymbol{\varphi} = [\varphi_1, \varphi_2, \varphi_3, \varphi_4]^T$ [Ska08], [Cox61, Section 13.7] which are related to the reference element coordinates $\widehat{\boldsymbol{r}} = [\widehat{r}_1, \widehat{r}_2, \widehat{r}_3]^T$ by

$$\begin{bmatrix} \widehat{r}_1 \\ \widehat{r}_2 \\ \widehat{r}_3 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} \varphi_1 \\ \varphi_2 \\ \varphi_3 \\ \varphi_4 \end{bmatrix} . \tag{2.108}$$

Each φ_i gives the value of one on the associated node $\widehat{\boldsymbol{r}}_i$ and zero on all other nodes.

The next step is the definition of the finite element approximation spaces. The most used space for approximating $H^1(\Omega)$-conforming functions is the polynomial

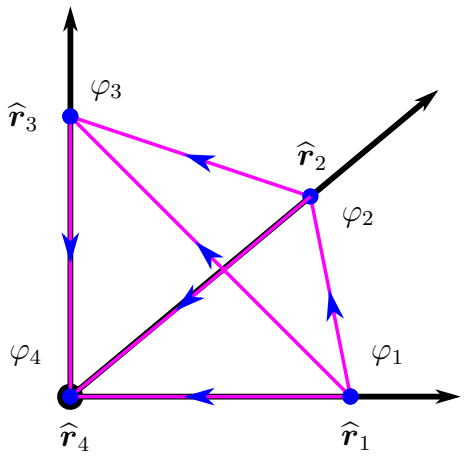


Figure 2.2: The reference tetrahedron.

space $\mathbb{P}_p$ of order p [M$^+$03, Chapter 5.6]. For the approximation of $\boldsymbol{H}(\mathrm{curl}, \Omega)$- and $\boldsymbol{H}(\mathrm{div}, \Omega)$-conforming functions, Nédélec introduced the incomplete- or mixed-order polynomial spaces [Néd80] of

$$R_p = (\mathbb{P}_{p-1})^3 \oplus \mathcal{S}_p, \tag{2.109a}$$
$$D_p = (\mathbb{P}_{p-1})^3 \oplus \mathcal{S}_p^{\perp}, \tag{2.109b}$$

where the highest order subspaces $\mathcal{S}_p$ and $\mathcal{S}_p^{\perp}$ are defined as

$$\mathcal{S}_p = \left\{ \boldsymbol{u} \in (\widetilde{\mathbb{P}}_p)^3 \middle| \boldsymbol{u} \cdot \widehat{\boldsymbol{r}} = 0 \right\}, \tag{2.110a}$$
$$\mathcal{S}_p^{\perp} = \left\{ \boldsymbol{u} \in (\widetilde{\mathbb{P}}_p)^3 \middle| \boldsymbol{u} \times \widehat{\boldsymbol{r}} = 0 \right\}. \tag{2.110b}$$

These spaces are chosen as local approximation spaces:

$$\widehat{\mathcal{V}}_p = \mathbb{P}_p, \qquad \widehat{\mathcal{W}}_p = R_p, \qquad \widehat{\mathcal{D}}_p = D_p. \tag{2.111}$$

The third step in the definition of a finite element is the determination of the moments. Let the $H^1(\widehat{\Omega}_e)$-conforming moments be defined as

$$\widehat{M}_p^{vn}(\widehat{v}) = \widehat{v}\Big|_{\widehat{N}}, \tag{2.112a}$$
$$\widehat{M}_p^{ve}(\widehat{v}) = \int_{\widehat{E}} \widehat{v}\,\widehat{q}\,\mathrm{d}\widehat{\boldsymbol{r}}, \qquad \forall\, \widehat{q} \in \mathbb{P}_{p-2}(\widehat{E}), \tag{2.112b}$$
$$\widehat{M}_p^{vf}(\widehat{v}) = \int_{\widehat{F}} \widehat{v}\,\widehat{q}\,\mathrm{d}\widehat{\boldsymbol{r}}, \qquad \forall\, \widehat{q} \in \mathbb{P}_{p-3}(\widehat{F}), \tag{2.112c}$$
$$\widehat{M}_p^{vv}(\widehat{v}) = \int_{\widehat{V}} \widehat{v}\,\widehat{q}\,\mathrm{d}\widehat{\boldsymbol{r}}, \qquad \forall\, \widehat{q} \in \mathbb{P}_{p-4}(\widehat{V}). \tag{2.112d}$$

Let the $\boldsymbol{H}(\mathrm{curl}, \widehat{\Omega}_e)$-conforming moments be defined as

$$\widehat{M}_p^{we}(\widehat{\boldsymbol{w}}) = \int_{\widehat{E}} \widehat{\boldsymbol{w}} \cdot (\widehat{q}\,\widehat{\boldsymbol{t}})\,\mathrm{d}\widehat{\boldsymbol{r}}, \qquad \forall\, \widehat{q} \in \mathbb{P}_{p-1}(\widehat{E}), \tag{2.113a}$$

$$\widehat{M}_p^{wf}(\widehat{\boldsymbol{w}}) = \int_{\widehat{F}} \widehat{\boldsymbol{w}} \cdot (\widehat{\boldsymbol{q}} \times \widehat{\boldsymbol{n}})\,\mathrm{d}\widehat{\boldsymbol{r}}, \qquad \forall\, \widehat{\boldsymbol{q}} \in (\mathbb{P}_{p-2}(\widehat{F}))^3, \tag{2.113b}$$

$$\widehat{M}_p^{wv}(\widehat{\boldsymbol{w}}) = \int_{\widehat{V}} \widehat{\boldsymbol{w}} \cdot \widehat{\boldsymbol{q}}\,\mathrm{d}\widehat{\boldsymbol{r}}, \qquad \forall\, \widehat{\boldsymbol{q}} \in (\mathbb{P}_{p-3}(\widehat{V}))^3. \tag{2.113c}$$

Let the $\boldsymbol{H}(\mathrm{div}, \widehat{\Omega}_e)$-conforming moments be defined as

$$\widehat{M}_p^{df}(\widehat{\boldsymbol{d}}) = \int_{\widehat{F}} \widehat{\boldsymbol{d}} \cdot (\widehat{q}\,\widehat{\boldsymbol{n}})\,\mathrm{d}\widehat{\boldsymbol{r}}, \qquad \forall\, \widehat{q} \in \mathbb{P}_{p-1}(\widehat{F}), \tag{2.114a}$$

$$\widehat{M}_p^{dv}(\widehat{\boldsymbol{d}}) = \int_{\widehat{V}} \widehat{\boldsymbol{d}} \cdot \widehat{\boldsymbol{q}}\,\mathrm{d}\widehat{\boldsymbol{r}}, \qquad \forall\, \widehat{\boldsymbol{q}} \in (\mathbb{P}_{p-2}(\widehat{V}))^3. \tag{2.114b}$$

In the case of straight-sided elements, these moments are equivalent to cases where the same moments and approximation spaces are assumed on the global elements [M$^+$03, Chapter 5], [Néd80]. For such cases, the induced finite element is unisolvent and conforming [M$^+$03, Chapter 5], [Néd80]. However, Piola transformations would involve rational functions for curvilinear geometry mappings. Hence, each element with different curvilinear geometry mapping would produce a different approximation space on the same reference element. This would make the general theoretical analysis impossible. Thus, only the locally defined moments are appropriate. It can be seen that (2.112), (2.113), and (2.114) are equivalent to the general moments (2.73), (2.74), and (2.75) with the approximation spaces $\widehat{\mathcal{V}}_p$, $\widehat{\mathcal{W}}_p$, and $\widehat{\mathcal{D}}_p$. These details are usually not observed because the basis functions are defined on the reference element and mapped to the curvilinear elements via Piola transformations. Hence, even if the moments are defined incorrectly on the global element, the definitions of the basis functions correspond to the locally defined interpolation operator and the correct geometry-invariant moments. In the case of smooth fields, the asymptotic convergence rates [M$^+$03, Chapter 5], [Néd80] for straight-sided meshes are

$$\left\|v - \pi_p^v v\right\|_{L^2(\Omega_e)} \leq Ch^{p+1}, \tag{2.115a}$$

$$\left\|v - \pi_p^v v\right\|_{H^1(\Omega_e)} \leq Ch^{p}, \tag{2.115b}$$

$$\left\|\boldsymbol{w} - \pi_p^w \boldsymbol{w}\right\|_{\boldsymbol{L}^2(\Omega_e)} \leq Ch^{p}, \tag{2.115c}$$

$$\left\|\boldsymbol{w} - \pi_p^w \boldsymbol{w}\right\|_{\boldsymbol{H}(\mathrm{curl},\Omega_e)} \leq Ch^{p}, \tag{2.115d}$$

$$\left\|\boldsymbol{d} - \pi_p^d \boldsymbol{d}\right\|_{\boldsymbol{L}^2(\Omega_e)} \leq Ch^{p}, \tag{2.115e}$$

$$\left\|\boldsymbol{d} - \pi_p^d \boldsymbol{d}\right\|_{\boldsymbol{H}(\mathrm{div},\Omega_e)} \leq Ch^{p}. \tag{2.115f}$$

Table 2.1: Basis functions for the approximation space $\widetilde{\mathcal{V}}_p$ from [Ing06].

$\widetilde{\mathcal{V}}_p$	Association	Basis function
1	node i	$\varphi_i,$
2	edge ij	$\varphi_i\varphi_j,$
3	edge ij	$\varphi_i\varphi_j\left(\varphi_i-\varphi_j\right),$
	face ijk	$\varphi_i\varphi_j\varphi_k,$
4	edge ij	$\varphi_i\varphi_j\left(\varphi_i^2-3\varphi_i\varphi_j+\varphi_j^2\right),$
	face ijk	$\varphi_i\varphi_j\varphi_k\left(\varphi_j-\varphi_k\right),$
	face ijk	$\varphi_i\varphi_j\varphi_k\left(\varphi_k-\varphi_i\right),$
	volume $ijkl$	$\varphi_i\varphi_j\varphi_k\varphi_l,$

Now that the finite element approximation spaces are defined, the focus is on the utilized basis functions. While they span the same space, specific bases differ in their numerical properties, such as sparsity, orthogonality, or condition numbers. The best-known basis functions are probably the interpolatory basis functions, where each function is associated with a single interpolation location, and the coefficients are associated with the field values. This is a great advantage because the unknowns of the finite element equation system have a direct interpretation. However, one disadvantage is that different orders of basis functions cannot be used for neighboring elements in the same mesh. Moreover, if the order of the elements needs to be increased, a new set of basis functions and interpolation locations are required.

These properties motivated the development of hierarchical basis functions. Generally, they are no longer associated with any interpolation location. On the other hand, degree elevation is possible by adding a set of incremental-order functions to the set of lower-order functions. This property can be utilized for *hp*-adaptive techniques [BS93], [Ban96], and [AP02] and for multilevel preconditioners [Ing06]. Let the incremental subspaces for $\mathcal{V}_p$ be denoted by $\widetilde{\mathcal{V}}_p$, for $\mathcal{W}_p$ by $\widetilde{\mathcal{W}}_p$, and for $\mathcal{D}_p$ by $\widetilde{\mathcal{D}}_p$. Then, the decompositions of the approximation spaces read

$$\mathcal{V}_p = \widetilde{\mathcal{V}}_1 \oplus ... \oplus \widetilde{\mathcal{V}}_p, \tag{2.116a}$$

$$\mathcal{W}_p = \widetilde{\mathcal{W}}_1 \oplus ... \oplus \widetilde{\mathcal{W}}_p, \tag{2.116b}$$

$$\mathcal{D}_p = \widetilde{\mathcal{D}}_1 \oplus ... \oplus \widetilde{\mathcal{D}}_p. \tag{2.116c}$$

In the following, the bases of Ingelström [Ing06] are preferred for the approximation spaces of $\mathcal{V}_p$ (Table 2.1) and $\mathcal{W}_p$ (Table 2.2). An important property of these bases is that each higher order ($p > 1$) hierarchical space $\widetilde{\mathcal{W}}_p$ is subdivided into the gradient subspace $\nabla\widetilde{\mathcal{V}}_p$ and a rotational subspace $\widetilde{\mathcal{A}}_p$ with non-zero curl. This decomposition

reads

$$\widetilde{\mathcal{W}}_p = \widetilde{\mathcal{A}}_p \oplus \nabla\widetilde{\mathcal{V}}_p, \qquad p > 1. \tag{2.117}$$

The advantage is that the functions of $\widetilde{\mathcal{V}}_p$ can be reused as gradient subspace functions in $\nabla\widetilde{\mathcal{V}}_p \subset \widetilde{\mathcal{W}}_p$, which is in the null space of the curl operator. Furthermore, it is possible to define an even finer subspace structure based on the association type of the moments. Those functions of $\widetilde{\mathcal{V}}_p$, which are associated with the nodes, belong to the subspace $\widetilde{\mathcal{V}}_p^n$, the edge-associated functions to $\widetilde{\mathcal{V}}_p^e$, the face-associated functions to $\widetilde{\mathcal{V}}_p^f$, and the volume functions to $\widetilde{\mathcal{V}}_p^v$. Hence, the incremental space $\widetilde{\mathcal{V}}_p$ is the sum of the different subspaces:

$$\widetilde{\mathcal{V}}_p = \widetilde{\mathcal{V}}_p^n \oplus \widetilde{\mathcal{V}}_p^e \oplus \widetilde{\mathcal{V}}_p^f \oplus \widetilde{\mathcal{V}}_p^v. \tag{2.118}$$

Similarly, the edge-, face-, and volume-associated functions of $\widetilde{\mathcal{A}}_p$ belong to the subspaces of $\widetilde{\mathcal{A}}_p^e$, $\widetilde{\mathcal{A}}_p^f$, and $\widetilde{\mathcal{A}}_p^v$, respectively. Hence, the incremental space $\widetilde{\mathcal{A}}_p$ is the sum of these subspaces:

$$\widetilde{\mathcal{A}}_p = \widetilde{\mathcal{A}}_p^e \oplus \widetilde{\mathcal{A}}_p^f \oplus \widetilde{\mathcal{A}}_p^v. \tag{2.119}$$

Another important property of Ingelström's basis is the orthogonality with respect to the interpolation operator:

$$\pi_p^v v_q = 0, \qquad \forall v_q \in \mathcal{V}_q \text{ and } q > p, \tag{2.120a}$$

$$\pi_p^w \boldsymbol{w}_q = \boldsymbol{0}, \qquad \forall \boldsymbol{w}_q \in \mathcal{W}_q \text{ and } q > p. \tag{2.120b}$$

This means that the lower-order interpolation of a higher-order interpolation is equivalent to a single, lower-order interpolation:

$$\pi_p^w(\pi_{p+1}^w \boldsymbol{w}) = \sum_{q=1}^{p} \sum_i c_q^i \boldsymbol{w}_q^i = \pi_p^w \boldsymbol{w}. \tag{2.121}$$

Therefore, any higher-order interpolation can be obtained by the superposition of the lower-order interpolation and the higher-order incremental term:

$$\pi_{p+1}^w \boldsymbol{w} = \sum_{q=1}^{p} \sum_i c_q^i \boldsymbol{w}_q^i + \sum_i c_{p+1}^i \boldsymbol{w}_{p+1}^i = \pi_p^w \boldsymbol{w} + \sum_i c_{p+1}^i \boldsymbol{w}_{p+1}^i. \tag{2.122}$$

There is no need to compute all the coefficients because the lower-order terms remain unchanged during degree elevation. This enables a hierarchical interpolation scheme via the previously described interpolation operators π_p^w, π_p^v with the corresponding basis functions. Moreover, the bases not only provide a high level of sparsity but are also useful [Ing06] for hierarchical preconditioners [ZC02] and solvers [PDEL99]. One contribution of the thesis is a family of hierarchical $\boldsymbol{H}(\mathrm{div})$ bases for the mixed-order Nédélec space with the same properties as Ingelström's $\boldsymbol{H}(\mathrm{curl})$ basis. Hence, with the present work, basis functions that are orthogonal with respect to the relevant interpolation operators are available for all spaces of the discrete de Rham complex.

Table 2.2: Basis functions for the approximation space $\widetilde{\mathcal{A}}_p$ from [Ing06].

$\widetilde{\mathcal{A}}_p$	Association	Basis function
1	edge ij	$\varphi_i \nabla \varphi_j - \varphi_j \nabla \varphi_i,$
2	face ijk	$3\varphi_j\varphi_k \nabla \varphi_i - \nabla (\varphi_i \varphi_j \varphi_k),$
	face ijk	$3\varphi_k\varphi_i \nabla \varphi_j - \nabla (\varphi_i \varphi_j \varphi_k),$
3	face ijk	$4\varphi_j\varphi_k (\varphi_j - \varphi_k) \nabla \varphi_i - \nabla (\varphi_i \varphi_j \varphi_k (\varphi_j - \varphi_k)),$
	face ijk	$4\varphi_k\varphi_i (\varphi_k - \varphi_i) \nabla \varphi_j - \nabla (\varphi_i \varphi_j \varphi_k (\varphi_k - \varphi_i)),$
	face ijk	$4\varphi_i\varphi_j (\varphi_i - \varphi_j) \nabla \varphi_k - \nabla (\varphi_i \varphi_j \varphi_k (\varphi_i - \varphi_j)),$
	volume $ijkl$	$4\varphi_j\varphi_k\varphi_l \nabla \varphi_i - \nabla (\varphi_i \varphi_j \varphi_k \varphi_l),$
	volume $ijkl$	$4\varphi_k\varphi_l\varphi_i \nabla \varphi_j - \nabla (\varphi_i \varphi_j \varphi_k \varphi_l),$
	volume $ijkl$	$4\varphi_l\varphi_i\varphi_j \nabla \varphi_k - \nabla (\varphi_i \varphi_j \varphi_k \varphi_l),$
4	face ijk	$5\varphi_j\varphi_k \left(\varphi_j^2 - 3\varphi_j\varphi_k + \varphi_k^2\right) \nabla \varphi_i -$ $\nabla \left(\varphi_i \varphi_j \varphi_k \left(\varphi_j^2 - 3\varphi_j\varphi_k + \varphi_k^2\right)\right),$
	face ijk	$5\varphi_k\varphi_i (\varphi_k^2 - 3\varphi_k\varphi_i + \varphi_i^2) \nabla \varphi_j -$ $\nabla (\varphi_i \varphi_j \varphi_k (\varphi_k^2 - 3\varphi_k\varphi_i + \varphi_i^2)),$
	face ijk	$5\varphi_i\varphi_j \left(\varphi_i^2 - 3\varphi_i\varphi_j + \varphi_j^2\right) \nabla \varphi_k -$ $\nabla \left(\varphi_i \varphi_j \varphi_k \left(\varphi_i^2 - 3\varphi_i\varphi_j + \varphi_j^2\right)\right),$
	face ijk	$(6\varphi_i - \varphi_j - \varphi_k) (\varphi_j - \varphi_k) \varphi_j\varphi_k \nabla \varphi_i +$ $(6\varphi_j - \varphi_k - \varphi_i) (\varphi_k - \varphi_i) \varphi_k\varphi_i \nabla \varphi_j +$ $(6\varphi_k - \varphi_i - \varphi_j) (\varphi_i - \varphi_j) \varphi_i\varphi_j \nabla \varphi_k,$
	volume $ijkl$	$5\varphi_j\varphi_k\varphi_l (\varphi_j - \varphi_k) \nabla \varphi_i - \nabla (\varphi_i \varphi_j \varphi_k \varphi_l (\varphi_j - \varphi_k)),$
	volume $ijkl$	$5\varphi_j\varphi_k\varphi_l (\varphi_k - \varphi_l) \nabla \varphi_i - \nabla (\varphi_i \varphi_j \varphi_k \varphi_l (\varphi_k - \varphi_l)),$
	volume $ijkl$	$5\varphi_k\varphi_l\varphi_i (\varphi_k - \varphi_l) \nabla \varphi_j - \nabla (\varphi_i \varphi_j \varphi_k \varphi_l (\varphi_k - \varphi_l)),$
	volume $ijkl$	$5\varphi_k\varphi_l\varphi_i (\varphi_l - \varphi_i) \nabla \varphi_j - \nabla (\varphi_i \varphi_j \varphi_k \varphi_l (\varphi_l - \varphi_i)),$
	volume $ijkl$	$5\varphi_l\varphi_i\varphi_j (\varphi_l - \varphi_i) \nabla \varphi_k - \nabla (\varphi_i \varphi_j \varphi_k \varphi_l (\varphi_l - \varphi_i)),$
	volume $ijkl$	$5\varphi_l\varphi_i\varphi_j (\varphi_i - \varphi_j) \nabla \varphi_k - \nabla (\varphi_i \varphi_j \varphi_k \varphi_l (\varphi_i - \varphi_j)),$
	volume $ijkl$	$5\varphi_i\varphi_j\varphi_k (\varphi_i - \varphi_j) \nabla \varphi_l - \nabla (\varphi_i \varphi_j \varphi_k \varphi_l (\varphi_i - \varphi_j)),$
	volume $ijkl$	$5\varphi_i\varphi_j\varphi_k (\varphi_j - \varphi_k) \nabla \varphi_l - \nabla (\varphi_i \varphi_j \varphi_k \varphi_l (\varphi_j - \varphi_k)),$

Chapter 3

Requirements for Algebraic Asymptotic Convergence Rates

A significant part of this thesis concerns new finite element spaces and basis functions that provide algebraic asymptotic convergence rates in terms of the finite element order and in the base of the mesh parameter. In order to lay the foundations for specific error estimations and convergence requirements, this chapter provides the general theoretical foundations and framework. The main, original contributions are the alternative definition of the mesh parameter, the interpretation of the mesh refinement, and the associated finite element error estimations. The results are given in a general way and apply to all elements whose finite element moments are subcases of the defined general moments (2.55), (2.57), and (2.59), and for which the approximation space of the reference element is polynomial. Therefore, the results apply to typical hexahedral, tetrahedral, and triangular-based prism elements.

3.1 Errors Related to Geometry Representation

The previously defined general moments (2.55), (2.57), and (2.59) require the geometry mapping $\boldsymbol{g}(\widehat{\boldsymbol{r}}) \in \boldsymbol{C}^p(\widehat{\Omega})$ to be p-times differentiable. Hence, its Taylor expansion of order p exists in every point $\widehat{\boldsymbol{r}}_i \in \widehat{\Omega}_e$ of the reference element. Therefore, it is possible to approximate or interpolate the geometry mappings by polynomials. In general, geometry mappings can take up various forms and may significantly differ element by element. In some cases, the geometry mappings might not be available as functions in closed forms but rather as sets of points on the boundaries. Hence, it is practical to interpolate the geometry mappings by polynomials $\boldsymbol{g}_k(\widehat{\boldsymbol{r}})$ of order k and use the same basis functions on each element to describe their shape. Then, an exact geometry mapping $\boldsymbol{g}(\widehat{\boldsymbol{r}})$ can be rewritten as

$$\boldsymbol{g}\left(\widehat{\boldsymbol{r}}\right) = \boldsymbol{g}_k\left(\widehat{\boldsymbol{r}}\right) + \boldsymbol{R}_k\left(\widehat{\boldsymbol{r}}\right), \tag{3.1}$$

where the remainder term $\boldsymbol{R}_k\left(\widehat{\boldsymbol{r}}\right)$ is the interpolation error. In the case of polynomial geometry mappings $\boldsymbol{g} \in (\mathbb{P}_k)^3$, the interpolation is exact (i.e., $\boldsymbol{g}\left(\widehat{\boldsymbol{r}}\right) = \boldsymbol{g}_k\left(\widehat{\boldsymbol{r}}\right)$ and $\boldsymbol{R}_k = \boldsymbol{0}$). Now that each element is equipped with an interpolated geometry mapping, the neighboring elements are required to map their adjacent boundaries to the same interface to preserve the elementary continuity of the mesh. Thus, the mappings of the neighboring elements are required to be in the same space on their common nodes $\widehat{N}_i$, edges $\widehat{E}_i$, and faces $\widehat{F}_i$. Here, subscript i is an index for the different subdomains. Therefore, the error is not just required to be zero for the interpolation of kth-order polynomials on the reference element, but also for all nodes, edges, and faces simultaneously for a given $\boldsymbol{g}$:

$$\boldsymbol{R}_k(\widehat{N}_i) = \boldsymbol{0}, \quad \forall i, \tag{3.2a}$$

$$\boldsymbol{g}(\widehat{E}_i) \in (\mathbb{P}_k(\widehat{E}_i))^3 \Longrightarrow \boldsymbol{R}_k(\widehat{E}_i) = \boldsymbol{0}, \quad \forall i, \tag{3.2b}$$

$$\boldsymbol{g}(\widehat{F}_i) \in (\mathbb{P}_k(\widehat{F}_i))^3 \Longrightarrow \boldsymbol{R}_k(\widehat{F}_i) = \boldsymbol{0}, \quad \forall i, \tag{3.2c}$$

$$\boldsymbol{g} \in (\mathbb{P}_k)^3 \Longrightarrow \boldsymbol{R}_k = \boldsymbol{0}. \tag{3.2d}$$

This induces a set of interpolation locations $\{\widehat{\boldsymbol{r}}_i\}$ for the reference element which belong to the nodes, edges, faces, and volume, as well as a corresponding set of global coordinates $\{\boldsymbol{r}_i\}$, where the error of the interpolated mapping is required to be zero:

$$\boldsymbol{g}_k\left(\widehat{\boldsymbol{r}}_i\right) = \boldsymbol{g}\left(\widehat{\boldsymbol{r}}_i\right) = \boldsymbol{r}_i. \tag{3.3}$$

In practice, an H^1-conforming finite element basis is used for the geometry interpolation because it satisfies (3.2), providing C^0 continuity.

3.1.1 The Geometry Interpolation Error

The investigation begins with the interpolation of a function along an edge $\widehat{E}_i$, which requires $k+1$ interpolation locations for kth-order polynomials. It is assumed that the interpolation locations are uniformly distributed. If an edge $\widehat{E}_i$ of the reference element is parametrized by the variable $\varphi_i \in [0,1]$, then the pointwise interpolation error of $\boldsymbol{g}(\varphi_i) \in \boldsymbol{C}^{k+1}$ over the edge [Run03] reads

$$|\boldsymbol{R}_k(\widehat{\boldsymbol{r}})| \leq \frac{1}{4(k+1)} \left(\frac{1}{k}\right)^{k+1} \sup_{\varphi_i \in [0,1]} \left| \frac{\partial^{k+1} \boldsymbol{g}(\varphi_i)}{\partial \varphi_i^{k+1}} \right|, \quad \forall \widehat{\boldsymbol{r}} \in \widehat{E}_i. \tag{3.4}$$

As the parametrization interval is independent of the actual geometry, the term

$$C_k = \frac{1}{4(k+1)} \left(\frac{1}{k}\right)^{k+1} \tag{3.5}$$

is also independent of the geometry mapping $\boldsymbol{g}$. Indeed, (3.5) depends only on the order of interpolation k and tends to zero when $k \to \infty$. By choosing the interpolation locations in a certain non-uniform pattern, it is possible to decrease the constant C_k for a given order k [BS07, Section 10.4]. However, the distribution of the interpolation locations does not change the fact that C_k is a positive constant with a finite upper bound C and

$$|\boldsymbol{g}(\widehat{\boldsymbol{r}}) - \boldsymbol{g}_k(\widehat{\boldsymbol{r}})| = |\boldsymbol{R}_k(\widehat{\boldsymbol{r}})| \leq C \sup_{\varphi_i \in [0,1]} \left| \frac{\partial^{k+1} \boldsymbol{g}(\varphi_i)}{\partial \varphi_i^{k+1}} \right|, \quad \forall \widehat{\boldsymbol{r}} \in \widehat{E}_i. \tag{3.6}$$

As the geometry interpolation space is also a polynomial space that is complete up to order k over the faces and the volume, similar error estimations can be obtained in a more general form for the other geometry partitions. Choosing the reference element to be $\widehat{\Omega}_e \subseteq [0,1]^3$, the application of the Bramble-Hilbert lemma [BS07,Section 4.3] yields the following estimate:

$$|\boldsymbol{g} - \boldsymbol{g}_k|_{s,q,(\cdot)} = |\boldsymbol{R}_k|_{s,q,(\cdot)} \leq C d^{k+1-s} \, |\boldsymbol{g}|_{k+1,q,(\cdot)}, \qquad \forall (\cdot) \in (\widehat{E}_i, \widehat{F}_i, \widehat{\Omega}_e), \tag{3.7}$$

where $C = C((\cdot), k)$ is a positive $\boldsymbol{g}$-independent constant and $d \leq 1$ is the diameter of the domain $(\cdot)$. Note that the term of $d^{k+1-s} \leq 1$ tends to zero when $k \to \infty$.

In the following, two subcases of (3.7) are given that are used to establish finite element error estimates in later sections. Using a given polynomial order k, the L^∞-norm estimate of the interpolation error ((3.7) with $s = 0$ and $q = \infty$) reads

$$\|\boldsymbol{R}_k\|_{L^\infty(\cdot)} \leq C \, |\boldsymbol{g}|_{k+1,\infty,(\cdot)}, \qquad \forall (\cdot) \in (\widehat{E}_i, \widehat{F}_i, \widehat{\Omega}_e), \tag{3.8}$$

where C is a positive $\boldsymbol{g}$ independent constant. The second subcase is the L^∞-norm estimate of the derivative of the interpolation error with respect to a multi-index $|\alpha| \leq k+1$ ((3.7) with $s \leq k+1$ and $q = \infty$):

$$\left\| \widehat{D}^\alpha \boldsymbol{R}_k \right\|_{L^\infty(\cdot)} \leq C \, |\boldsymbol{g}|_{k+1,\infty,(\cdot)}, \qquad \forall (\cdot) \in (\widehat{E}_i, \widehat{F}_i, \widehat{\Omega}_e). \tag{3.9}$$

3.1.2 Mesh Parameter and Mesh Refinement

The common way to uniformly refine a straight-sided mesh of a straight-sided geometry is to equally subdivide the edges and also take the midpoints of the faces and the volume if necessary. However, generating a completely new mesh with a halved element size could provide a similar result. In both cases, the mesh parameters are identical, but only the first approach maintains a connection between the initial and refined meshes. In the case of curvilinear elements and interpolated geometry mappings, these processes are less straightforward. In most cases, the mesh parameter is considered to be associated with the size of the underlying straight-sided mesh [Cia02, ABB15]. The elements with higher-order interpolation nodes are considered [Cia02] to be perturbations of the straight-sided elements, leaving the mesh parameter unchanged. Therefore, two different refinement methods are being used, similarly to the straight-sided case. First, either the underlying straight-sided mesh is refined homogeneously or an entirely new straight-sided mesh is generated with half the size of the previous mesh. Then, the higher-order interpolation nodes are inserted for the elements next to the curvilinear boundaries and are displaced to the exact curvilinear boundaries.

The present framework takes a different approach. First, the existence of an initial macro mesh, which represents the geometry without any error, is assumed. Hence, each macro element of this initial mesh has a perfect geometry mapping $\boldsymbol{r} = \widetilde{\boldsymbol{g}}_0(\widehat{\boldsymbol{r}}_0)$ that maps the reference domain of the macro element $\widehat{\Omega}_e$ to the global domain of the macro element Ω_M:

$$\widetilde{\boldsymbol{g}}_0(\widehat{\boldsymbol{r}}_0) : \widehat{\boldsymbol{r}}_0 \in \widehat{\Omega}_e \longrightarrow \boldsymbol{r} \in \Omega_M. \tag{3.10}$$

Then, the mesh refinement is considered to be an arbitrary reparametrization procedure of each initial reference element. After refining the initial mesh n-times, the mapping,

$$\boldsymbol{g}(\widehat{\boldsymbol{r}}) : \widehat{\boldsymbol{r}} \in \widehat{\Omega}_e \longrightarrow \boldsymbol{r} \in \Omega_e, \tag{3.11}$$

of an arbitrary element is obtained as

$$\boldsymbol{g}(\widehat{\boldsymbol{r}}) = \widetilde{\boldsymbol{g}}_0(\widehat{\boldsymbol{r}}_0) \circ \widetilde{\boldsymbol{g}}_1(\widehat{\boldsymbol{r}}_1) \circ \widetilde{\boldsymbol{g}}_2(\widehat{\boldsymbol{r}}_2) \cdots \circ \widetilde{\boldsymbol{g}}_n(\widehat{\boldsymbol{r}}_n) = \widetilde{\boldsymbol{g}}_0(\widehat{\boldsymbol{r}}_0) \circ \boldsymbol{f}(\widehat{\boldsymbol{r}}_n = \widehat{\boldsymbol{r}}), \tag{3.12}$$

where a single function composition with a mapping

$$\widetilde{\boldsymbol{g}}_n(\widehat{\boldsymbol{r}}_n) : \widehat{\boldsymbol{r}}_n \in \widehat{\Omega}_e \longrightarrow \widehat{\boldsymbol{r}}_{n-1} \in \widehat{\Omega}_{n-1} \subset \widehat{\Omega}_e, \quad \text{for} \quad n > 0, \tag{3.13}$$

corresponds to one refinement. Here, $\widehat{\Omega}_{n-1}$ is a domain of an element that is obtained by the refinement of $\widehat{\Omega}_e$ in the $(n-1)$th step. The function

$$\boldsymbol{f}(\widehat{\boldsymbol{r}}_n = \widehat{\boldsymbol{r}}) : \widehat{\boldsymbol{r}}_n = \widehat{\boldsymbol{r}} \in \widehat{\Omega}_e \longrightarrow \widehat{\boldsymbol{r}}_0 \in \widehat{\Omega}_0 \subset \widehat{\Omega}_e, \quad \text{for} \quad n > 0, \tag{3.14}$$

is the function composition of all refinements. The previously described mappings, $\widetilde{\boldsymbol{g}}_0$, $\widetilde{\boldsymbol{g}}_i$, and $\boldsymbol{f}$, are required to be sufficiently smooth: to be $\boldsymbol{C}^1$ diffeomorphisms (bijective with $\boldsymbol{C}^1$ inverse functions) and

$$\|\widetilde{\boldsymbol{g}}_0\|_{1,\infty,\widehat{\Omega}_e} \leq C, \qquad \|\widetilde{\boldsymbol{g}}_i\|_{1,\infty,\widehat{\Omega}_e} \leq C, \qquad \|\boldsymbol{f}\|_{1,\infty,\widehat{\Omega}_e} \leq C, \tag{3.15}$$

where C is a positive constant. Therefore, any geometry mapping $\boldsymbol{g}$ of a regular element inherits these properties and the corresponding inverse-mapping $\boldsymbol{g}^{-1}$ also satisfy

$$\|\boldsymbol{g}\|_{1,\infty,\widehat{\Omega}_e} \leq C, \qquad \left\|\boldsymbol{g}^{-1}\right\|_{1,\infty,\Omega_e} \leq C. \tag{3.16}$$

The proposed interpretation is illustrated in Figure 3.1. A consequence is that the geometry boundaries and the element boundaries of the initial mesh are undisturbed by the refinement, and the geometry error of the geometry interpolation is only introduced after the refinements. Note that not all meshes and geometry mappings

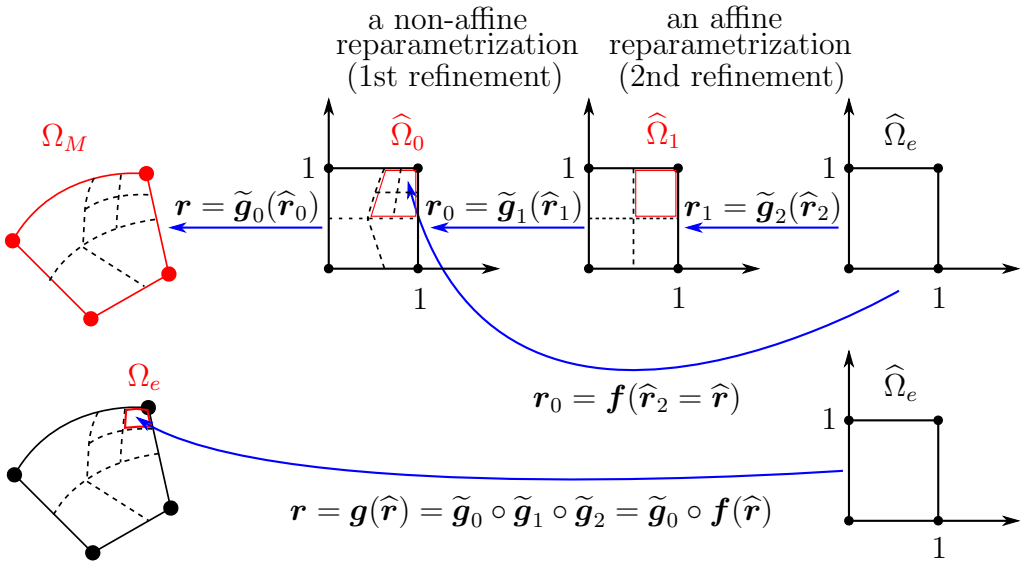


Figure 3.1: Illustration of an element's geometry mapping in the proposed alternative interpretation.

of the elements are obtained in the previously described way. However, it is always possible to artificially construct a mesh refinement history in this interpretation by providing an initial mesh with initial geometry mappings $\widetilde{\boldsymbol{g}}_0$ and creating appropriate reparametrization functions $\boldsymbol{f}$ that map and refine the initial mesh into the available refined mesh. The artificially constructed mesh refinement histories are not unique and hard to obtain. However, the following theories require only some characterizations of the overall refinement functions $\boldsymbol{f}$ and not the initial mesh or geometry description.

Now that the refinement is clarified, the mesh parameter h can be defined. Instead of defining it via the usual approach as giving a geometrical interpretation, it is introduced as a certain quantity, in which base, the algebraic convergence might occur. Then, a geometrical interpretation will be provided for some special cases. Provided that the largest extent of the reference element in the direction of the local coordinates is normalized to one (i.e., $\widehat{\Omega}_e \subseteq [0,1]^3$), the definition of the mesh parameter reads

$$h := \max_{|\alpha|\geq 1} \left\{ \left\| \widehat{D}^{\alpha} \boldsymbol{f} \right\|_{L^{\infty}(\widehat{\Omega}_e)} \right\}. \tag{3.17}$$

Then, the global mesh parameter is given as the largest h of all the elements. Note that the usual definition of the mesh parameter

$$h := \max_{|\alpha|=1} \left\{ \left\| \widehat{D}^{\alpha} \boldsymbol{f} \right\|_{L^{\infty}(\widehat{\Omega}_e)} \right\}, \tag{3.18}$$

only involves the first-order derivative of the mapping $\boldsymbol{f}$, and the higher-order derivatives are irrelevant. However, the two definitions, (3.17) and (3.18), become equivalent in the asymptotic case for $h \to 0$ if $\boldsymbol{f}$ is asymptotically affine.

Next, estimations are given for the derivatives of geometry mappings in order to determine the effect of mesh refinement on the geometry interpolation error (3.8). Only the $L^{\infty}(\widehat{\Omega}_e)$-norm is considered because this result is required later to establish finite element approximation estimates. However, similar estimates can be obtained for the $L^{\infty}(\widehat{E})$ and $L^{\infty}(\widehat{F})$ norms.

Estimation of the Geometry Mapping Error for Affine Refinement

Using the rule for the derivative of function composition [Leo17], the derivative of the geometry mapping with respect to the local coordinates reads

$$\widehat{D}^{\alpha}\boldsymbol{g} = \widehat{D}^{\alpha}\left(\widetilde{\boldsymbol{g}}_0 \circ \boldsymbol{f}\right) = \sum_{\beta,\gamma,l} C_{\beta,\gamma,l}\left((\widehat{D}^{\beta}\widetilde{\boldsymbol{g}}_0) \circ \boldsymbol{f}\right)\prod_{i=1}^{|\beta|}\widehat{D}^{\gamma_i}\boldsymbol{f}_{l_i}. \tag{3.19}$$

This is a combinatorial form, where the sums and multiplications of the functions are required for all combinations of β, γ, and l. Here, α and β are multi-indices with the properties of

$$\alpha = (\alpha_1, \alpha_2, \alpha_3) \in (\mathbb{Z}^+)^3 \text{ and } 1 \leq |\alpha| \leq k+1, \tag{3.20a}$$

$$\beta = (\beta_1, \beta_2, \beta_3) \in (\mathbb{Z}^+)^3 \text{ and } 1 \leq |\beta| \leq |\alpha|\,. \tag{3.20b}$$

Moreover, γ is a set of multi-indices for all values of β satisfying

$$\gamma = (\gamma_1, \cdots, \gamma_{|\beta|}), \quad \gamma_i \in (\mathbb{Z}^+)^3 \text{ and } \sum_{i=1}^{|\beta|}\gamma_i = \alpha, \tag{3.21}$$

and l is a set of corresponding indices for vector components of $\boldsymbol{f}$ with the values of

$$l = (l_1, \cdots, l_{|\beta|}) \text{ and } l_i \in \{1, 2, 3\}. \tag{3.22}$$

The scaling term $C_{\beta,\gamma,l}$ is a geometry-independent constant for all combinations of the indices.

This section considers the case of affine refinement, where each element is obtained by a series of affine refinement steps. The function of a single reparametrization in the nth refinement step reads

$$\widetilde{\boldsymbol{g}}_n(\widehat{\boldsymbol{r}}_n) = \boldsymbol{A}_n\widehat{\boldsymbol{r}}_n + \boldsymbol{b}_n \in (\mathbb{P}_1)^3\,, \tag{3.23}$$

where $\boldsymbol{A}_n$ is a constant matrix and $\boldsymbol{b}_n$ is a constant vector. For a uniform affine refinement step, the size of each domain is uniformly halved, and $\|\boldsymbol{A}_n\|_{\infty} = 1/2$. Since the function composition of n affine and uniform refinement step is also an affine mapping, the bound for an arbitrary derivative of $\boldsymbol{f}$ is

$$\left\|\widehat{D}^{\gamma_i}\boldsymbol{f}_{l_i}\right\|_{L^{\infty}(\widehat{\Omega}_e)} = \left\|\boldsymbol{A}_1\boldsymbol{A}_2\cdots\boldsymbol{A}_n\frac{\partial^{|\gamma_i|}\widehat{\boldsymbol{r}}_{l_i}}{\partial\widehat{\boldsymbol{r}}^{\gamma_i}}\right\|_{L^{\infty}(\widehat{\Omega}_e)} \leq \begin{cases} 2^{-n} = h & \text{for } |\gamma_i| = 1 \\ 0 & \text{for } |\gamma_i| > 1 \end{cases}. \tag{3.24}$$

Hence, the mesh parameter (3.17) after n refinement steps becomes $h = 1/2^n$. The $L^{\infty}(\widehat{\Omega}_e)$-norm estimate can be obtained for (3.19) by applying the generalized Hölder and triangle inequalities:

$$\left\|\widehat{D}^{\alpha}\boldsymbol{g}\right\|_{L^{\infty}(\widehat{\Omega}_e)} = \sum_{\beta,\gamma,l} C_{\beta,\gamma,l}\left\|(\widehat{D}^{\beta}\widetilde{\boldsymbol{g}}_0) \circ \boldsymbol{f}\right\|_{L^{\infty}(\widehat{\Omega}_e)}\prod_{i=1}^{|\beta|}\left\|\widehat{D}^{\gamma_i}\boldsymbol{f}_{l_i}\right\|_{L^{\infty}(\widehat{\Omega}_e)}. \tag{3.25}$$

As all the higher-order derivatives of $\boldsymbol{f}$ vanish in (3.24), only the terms with $|\beta| = |\alpha|$ remain in (3.25) due to (3.20) and (3.21). Thus the substitution of $|\alpha| = k+1$ and (3.24) for (3.25) yields

$$\left\| \widehat{D}^{\alpha} \boldsymbol{g} \right\|_{L^{\infty}(\widehat{\Omega}_e)} \leq \sum_{|\beta|=k+1} C_{\beta} \left\| (\widehat{D}^{\beta} \widetilde{\boldsymbol{g}}_0) \circ \boldsymbol{f} \right\|_{L^{\infty}(\widehat{\Omega}_e)} \prod_{i=1}^{|\beta|} h \leq C h^{k+1} \left| \widetilde{\boldsymbol{g}}_0 \right|_{k+1,\infty,\widehat{\Omega}_e}, \quad (3.26)$$

where C_{β} and C are constants independent of the geometry. This result can be applied to estimate the geometry interpolation error of the affinely refined initial mesh where the geometry mappings $\boldsymbol{g}$ are asymptotically affine. The application of (3.26) for the error term (3.8) leads to

$$\left\| \boldsymbol{R}_k \right\|_{L^{\infty}(\widehat{\Omega}_e)} \leq C h^{k+1} \left| \widetilde{\boldsymbol{g}}_0 \right|_{k+1,\infty,\widehat{\Omega}_e}, \quad (3.27)$$

where C is a geometry-independent constant. Thus, the geometry error depends exponentially on the order of interpolation k in the base of mesh parameter h, which becomes half of its previous size in each refinement step for uniform refinements. The last factor $|\widetilde{\boldsymbol{g}}_0|_{k+1,\infty,\widehat{\Omega}_e}$ depends on the smoothness of the initial geometry mapping. Note that if the geometry and the initial mapping are described by polynomials of order smaller than $k+1$, the interpolation error becomes zero since $|\widetilde{\boldsymbol{g}}_0|_{k+1,\infty,\widehat{\Omega}_e} = 0$.

Estimation of the Geometry Mapping Error for Polynomial Refinement

The second type of refinement that we will consider is a higher-order polynomial refinement with an overall refinement mapping $\boldsymbol{f} \in (\mathbb{P}_r)^3$. Applying the definition of the mesh parameter (3.17), the bound for the derivative of $\boldsymbol{f}$ with a multi-index γ_i becomes

$$\left\| \widehat{D}^{\gamma_i} \boldsymbol{f}_{l_i} \right\|_{L^{\infty}(\widehat{\Omega}_e)} \leq \begin{cases} h & \text{for } |\gamma_i| \leq r \\ 0 & \text{for } |\gamma_i| > r \end{cases}. \quad (3.28)$$

Hence, the worst-case scenario in the estimation of the multiplicative term of (3.19) happens when $|\gamma_i| = r$ for all indices i. For any higher-order derivative, at least a single term must become zero, which leads to a zero multiplicative term, i.e.,

$$\sum_{i=1}^{|\beta|} |\gamma_i| = |\alpha| = k+1 \; > \; (|\beta|\, r), \qquad \Longrightarrow \qquad \left\| \prod_{i=1}^{|\beta|} \widehat{D}^{\gamma_i} \boldsymbol{f}_{l_i} \right\|_{L^{\infty}(\widehat{\Omega}_e)} = 0. \quad (3.29)$$

Thus, the value of the multi-index β is bounded from below by $|\beta| \geq \lceil (k+1)/r \rceil$, and the estimation of the multiplicative term becomes

$$\left\| \prod_{i=1}^{|\beta|} \widehat{D}^{\gamma_i} \boldsymbol{f}_{l_i} \right\|_{L^{\infty}(\widehat{\Omega}_e)} \leq C h^{|\beta|} \leq C h^{\lceil (k+1)/r \rceil}, \quad (3.30)$$

where C is a geometry-independent constant and $\lceil\cdot\rceil$ denotes the rounding up operator for the closest integer (ceil function). Hence, estimate can be obtained for the geometry error (3.8) with the substitution of (3.25) and (3.30) with $|\alpha| = k+1$:

$$\begin{aligned}\|\boldsymbol{R}_k\|_{L^\infty(\widehat{\Omega}_e)} &\leq \sum_{k+1\geq|\beta|\geq\lceil(k+1)/r\rceil} C_\beta \left\|(\widehat{D}^\beta \widetilde{\boldsymbol{g}}_0)\circ\boldsymbol{f}\right\|_{L^\infty(\widehat{\Omega}_e)} h^{\lceil(k+1)/r\rceil} \\ &\leq C h^{\lceil(k+1)/r\rceil}\,\|\widetilde{\boldsymbol{g}}_0\|_{k+1,\infty,\widehat{\Omega}_e}\,. \end{aligned} \tag{3.31}$$

Here, C_β and C are geometry-independent constants, and the factor of $\|\widetilde{\boldsymbol{g}}_0\|_{k+1,\infty,\widehat{\Omega}_e}$ is determined by the smoothness of the initial geometry mapping $\widetilde{\boldsymbol{g}}_0$. It is easy to conclude that for higher order ($r > 1$) refinements, the convergence of the interpolation error becomes worse compared to the case of the affine refinement ($r = 1$).

Estimation of the Geometry Mapping Error for General Refinement

The last considered refinement case is when the overall refinement mapping $\boldsymbol{f}$ is a general, infinitely differentiable function over $\widehat{\Omega}_e$. The resulting estimate is equivalent to the previous case of the rth-order polynomial refinements, for the limit of $r \to \infty$. Therefore, the interpolation error of the geometry only linearly depends on the mesh parameter:

$$\|\boldsymbol{R}_k\|_{L^\infty(\widehat{\Omega}_e)} \leq C h\,\|\widetilde{\boldsymbol{g}}_0\|_{k+1,\infty,\widehat{\Omega}_e}\,. \tag{3.32}$$

Summary of the Geometry Mapping Error Estimates

To summarize the results, the mesh refinement is considered to be a reparametrization of an initial mesh, one which exactly represents the geometry. The interpolated geometry mappings are determined by the interpolation locations on these refined elements. In general, the mesh parameter h is only related to the overall refinement function (3.17) and not to the actual geometry. However, for affine refinements ($r = 1$), the multi-index in the definition of the mesh parameter (3.17) is $|\alpha| = 1$, and the geometrical interpretation of h becomes the largest side-length (in the direction of the reference element axis) of the element in the initial normalized domain $\widehat{\Omega}_0$. This geometrical interpretation can be lost for non-affine refinements ($r > 1$) when the higher-order derivatives ($|\alpha| > 1$) are also present in the definition (3.17).

The obtained asymptotic convergence rates of the geometry error for interpolated geometry mappings (assuming smooth initial geometry mappings $\|\widetilde{\boldsymbol{g}}_0\|_{k+1,\infty,\widehat{\Omega}_e}$ with

a finite upper bound) are

$$\|\boldsymbol{R}_k\|_{L^\infty(\widehat{\Omega}_e)}, |\boldsymbol{g}|_{k+1,\infty,\widehat{\Omega}_e} \leq \begin{cases} Ch^{k+1} & \text{for affine refinement,} \\ Ch^{\lceil (k+1)/r \rceil} & \text{for polynomial refinement of order } r, \\ Ch & \text{for general non-affine refinement,} \end{cases} \tag{3.33}$$

where C is a constant, independent of h. Hence, the affine refinement is the most advantageous among the investigated methods. Note that the constant C in the front depends on the smoothness of the initial geometry mapping. Therefore, it is reasonable to choose or construct this initial mapping as smoothly as possible to reduce errors. Moreover, the smoothness also influences where the asymptotic region of convergence begins for the geometry interpolation error.

Error Estimation of the Geometry Mapping Derivatives

Here, an estimate is given for the derivative of the geometry interpolation error, which is required by the finite element error estimations. As the form of the geometry error estimate (3.8) is the same as for its derivative (3.9), the same estimation applies as for (3.33), i.e.,

$$\left\| \widehat{D}^\alpha \boldsymbol{R}_k \right\|_{L^\infty(\widehat{\Omega}_e)} \leq Ch^{\lceil (k+1)/r \rceil}, \tag{3.34}$$

where C is a constant, independent of h.

Since the determinant of the Jacobian is the sum of triple products of the geometry mapping derivatives, the piecewise application of

$$|\boldsymbol{g}|_{k,\infty,\widehat{\Omega}_e} \leq Ch^{\lceil k/r \rceil} \tag{3.35}$$

leads to the estimate

$$|\det \boldsymbol{J}|_{k,\infty,\widehat{\Omega}_e} \leq Ch^{3+\lceil k/r \rceil}. \tag{3.36}$$

Error Estimation of the Inverse Geometry Mapping

The estimates for inverse geometry mappings are less straightforward in the general case. However, considering only first-order derivatives, the connection between the derivative geometry mapping and the derivative of the inverse geometry mapping is given by the invertible Jacobians (inverse function theorem):

$$\widehat{\nabla} \boldsymbol{g}^T = \boldsymbol{J} = \widehat{\boldsymbol{J}}^{-1} \circ \boldsymbol{g} = \left[\nabla \left(\boldsymbol{g}^{-1} \right)^T \right]^{-1} \circ \boldsymbol{g}. \tag{3.37}$$

Moreover, it can be easily shown for a β derivative of a function $v \in W^{|\beta|,\infty}(\Omega_e)$ that the infinite-norm is equivalent in any domain because Ω_e and $\widehat{\Omega}_e$ are interchangeable due to the vanishing effect of the volume element:

$$\left|\left(D^{\beta} v\right) \circ \boldsymbol{g}\right|_{0,\infty,\widehat{\Omega}_e} = \left|D^{\beta} v\right|_{0,\infty,\Omega_e} \leq |v|_{|\beta|,\infty,\Omega_e} . \tag{3.38}$$

Therefore, an estimate for the derivative of $\boldsymbol{g}^{-1} \in W^{1,\infty}(\Omega_e)$ can be obtained by inverting the right-hand side in (3.35), due to (3.37) and (3.38), as

$$\left|\left[\boldsymbol{g}^{-1}\right]_i\right|_{1,\infty,\Omega_e} \leq C h^{-1}, \tag{3.39}$$

where C is another bounded positive constant, independent of h. For the same reason, the estimate for the inverse Jacobian determinant becomes

$$\left|\det \widehat{\boldsymbol{J}}\right|_{0,\infty,\Omega_e} = \left|\det(\boldsymbol{J}^{-1})\right|_{0,\infty,\widehat{\Omega}_e} \leq C h^{-3}. \tag{3.40}$$

The next step of this investigation shows that the higher-order derivatives of the inverse geometry mapping converge at the same rate as for the first derivative (3.39), assuming affine refinements only ($r = 1$). First, the derivative rule of function compositions (3.19) is applied for the derivative of the identity mapping

$$\begin{aligned} \mathbf{0} = \widehat{D}^{\alpha} \widehat{\boldsymbol{r}} = \widehat{D}^{\alpha} \left(\boldsymbol{g}^{-1} \circ \boldsymbol{g}\right), \qquad & \text{for } |\alpha| > 1, \\ \mathbf{0} \overset{(3.19)}{=} \sum_{\beta,\gamma,l} C_{\alpha,\beta,\gamma,l} \left(\left(D^{\beta} \boldsymbol{g}^{-1}\right) \circ \boldsymbol{g}\right) \prod_{i=1}^{|\beta|} \widehat{D}^{\gamma_i} [\boldsymbol{g}]_{l_i}. \end{aligned} \tag{3.41}$$

Then, it is always possible to take such a linear combination of (3.41) with constant coefficients $\widetilde{C}_{\alpha_n}$ for the different multi-indices $|\alpha_n| = |\alpha|$, which satisfies

$$\begin{aligned} \mathbf{0} = \sum_{n}^{|\alpha_n|=|\alpha|} \widetilde{C}_{\alpha_n} \widehat{D}^{\alpha_n} \left(\boldsymbol{g}^{-1} \circ \boldsymbol{g}\right) = \left(D^{\alpha} \boldsymbol{g}^{-1}\right) \circ \boldsymbol{g} \sum_{\gamma,l} \widetilde{C}_{\alpha,\gamma,l} \prod_{i=1}^{|\alpha|} \widehat{D}^{\gamma_i} [\boldsymbol{g}]_{l_i} \\ + \sum_{\beta,\gamma,l}^{|\beta|\leq|\alpha|-1} \widetilde{C}_{\alpha,\beta,\gamma,l} \left(D^{\beta} \boldsymbol{g}^{-1}\right) \circ \boldsymbol{g} \prod_{i=1}^{|\beta|} \widehat{D}^{\gamma_i} [\boldsymbol{g}]_{l_i}, \qquad \text{for } |\alpha| > 1, \end{aligned} \tag{3.42}$$

where $\widetilde{C}_{\alpha,\gamma,l}$ and $\widetilde{C}_{\alpha,\beta,\gamma,l}$ are some constant coefficients. Thus, an α-derivative of the inverse geometry mapping can be expressed by rearranging (3.42):

$$\left(D^{\alpha} \boldsymbol{g}^{-1}\right) \circ \boldsymbol{g} = \frac{\sum_{\beta,\gamma,l}^{|\beta|\leq|\alpha|-1} \widetilde{C}_{\alpha,\beta,\gamma,l} \left(D^{\beta} \boldsymbol{g}^{-1}\right) \circ \boldsymbol{g} \prod_{i=1}^{|\beta|} \widehat{D}^{\gamma_i} [\boldsymbol{g}]_{l_i}}{-\sum_{\gamma,l} \widetilde{C}_{\alpha,\gamma,l} \prod_{i=1}^{|\alpha|} \widehat{D}^{\gamma_i} [\boldsymbol{g}]_{l_i}}. \tag{3.43}$$

At this point, an estimate can be obtained for the norm of each term in (3.43) using (3.33) and (3.38):

$$\left|\left(D^{\beta}\boldsymbol{g}^{-1}\right)\circ\boldsymbol{g}\right|_{0,\infty,\widehat{\Omega}_e} \overset{(3.38)}{\leq} \left|D^{\beta}\boldsymbol{g}^{-1}\right|_{0,\infty,\Omega_e}, \tag{3.44}$$

$$\left|\left(\prod_{i=1}^{|\alpha|}\widehat{D}^{\gamma_i}\left[\boldsymbol{g}\right]_{l_i}\right)^{-1}\right|_{0,\infty,\widehat{\Omega}_e} \leq \prod_{i=1}^{|\alpha|}\left|\left(\widehat{D}^{\gamma_i}\left[\boldsymbol{g}\right]_{l_i}\right)^{-1}\right|_{0,\infty,\widehat{\Omega}_e}, \qquad \text{for } |\gamma_i|=1,$$

$$\overset{(3.33)}{\leq} C\prod_{i=1}^{|\alpha|}h^{-1} \leq Ch^{-|\alpha|}. \tag{3.45}$$

Moreover,

$$\left|\prod_{i=1}^{|\beta|}\widehat{D}^{\gamma_i}\left[\boldsymbol{g}\right]_{l_i}\right|_{0,\infty,\widehat{\Omega}_e} \leq \prod_{i=1}^{|\beta|}\left|\widehat{D}^{\gamma_i}\left[\boldsymbol{g}\right]_{l_i}\right|_{0,\infty,\widehat{\Omega}_e}, \qquad \text{for } \sum_{i=1}^{|\beta|}|\gamma_i| = |\alpha|,$$

$$\overset{(3.33)}{\leq} C\prod_{i=1}^{|\beta|}h^{\lceil|\gamma_i|/r\rceil} \leq Ch^{\lceil|\alpha|/r\rceil}. \tag{3.46}$$

Thus, using the separate estimates of each term (3.44), (3.45), and (3.46), the estimate of the derivative inverse geometry mapping for affine refinements ($r = 1$) becomes

$$\left|D^{\alpha}\boldsymbol{g}^{-1}\right|_{0,\infty,\Omega_e} \leq C\sum_{\beta}^{|\beta|\leq|\alpha|-1}\left|D^{\beta}\boldsymbol{g}^{-1}\right|_{0,\infty,\Omega_e}h^{|\alpha|}h^{-|\alpha|} = C\sum_{\beta}^{|\beta|\leq|\alpha|-1}\left|D^{\beta}\boldsymbol{g}^{-1}\right|_{0,\infty,\Omega_e}. \tag{3.47}$$

Using mathematical induction form the order $|\alpha| = 2$, increasing $|\alpha|$ one-by-one, the general estimate of the inverse geometry mapping remains the same as for first-order derivatives (3.39):

$$\left|\boldsymbol{g}^{-1}\right|_{k,\infty,\Omega_e} \leq Ch^{-1}, \qquad \text{for } k \geq 1. \tag{3.48}$$

For the same reason, the estimate for the inverse Jacobian determinant is

$$\left|\det\widehat{\boldsymbol{J}}\right|_{k,\infty,\Omega_e} = \left|\det(\boldsymbol{J}^{-1})\right|_{k,\infty,\widehat{\Omega}_e} \leq Ch^{-3}. \tag{3.49}$$

3.1.3 Function Mappings of Interpolated Geometries

This section shows how the interpolated geometry mappings affect the mapping of functions in different function spaces. According to diagram (2.83), the finite element interpolation is always applied on the reference element introducing the local field error. Then, these local fields are mapped to the global element where the geometry error is introduced when using interpolated geometry mappings. Therefore, the starting point should always be the reference element where the fields are assumed to be free of geometry errors.

Let the domain of the interpolated element be denoted by Ω_e^k that is described by the interpolated geometry mapping $\boldsymbol{g}_k$ of order k. Thus, the exact geometry mapping $\boldsymbol{g}$ and its interpolant $\boldsymbol{g}_k$ are mappings from $\widehat{\Omega}_e$ to Ω_e and Ω_e^k, respectively:

$$\boldsymbol{g}: \quad \widehat{\Omega}_e \to \Omega_e, \tag{3.50a}$$

$$\boldsymbol{g}_k: \quad \widehat{\Omega}_e \to \Omega_e^k. \tag{3.50b}$$

Note that if $\boldsymbol{g} \neq \boldsymbol{g}_k$, then $\Omega_e \neq \Omega_e^k$ but $\Omega_e \cap \Omega_e^k$ is sufficiently large. The largest extent of the deviation of Ω_e^k from Ω_e is given by (3.33), i.e.,

$$\|\boldsymbol{g} - \boldsymbol{g}_k\|_{L^\infty(\widehat{\Omega}_e)} \leq Ch^{\lceil (k+1)/r \rceil}, \tag{3.51}$$

for polynomial refinements of order r. Therefore, the volume of the difference domains $\Omega_e \setminus \Omega_e^k$ and $\Omega_e^k \setminus \Omega_e$ must converge with at least the same rate as (3.51):

$$\text{vol.}(\Omega_e^k \setminus \Omega_e) \leq Ch^{\lceil (k+1)/r \rceil}, \tag{3.52a}$$

$$\text{vol.}(\Omega_e \setminus \Omega_e^k) \leq Ch^{\lceil (k+1)/r \rceil}. \tag{3.52b}$$

Hence, in the context of asymptotic convergence rates in a given norm over an element, the contribution of any sufficiently smooth field with a finite norm is negligible in the difference regions $\Omega_e \setminus \Omega_e^k$ and $\Omega_e^k \setminus \Omega_e$ since they show at least the same convergence rates as the geometry interpolation error. As an outcome, it is sufficient to investigate the effects of the functions mappings in the intersection region

$$\widetilde{\Omega}_e^k = \Omega_e \cap \Omega_e^k. \tag{3.53}$$

Let us denote the functions of the reference element by $\widehat{v} \in H^1(\widehat{\Omega}_e)$, $\widehat{\boldsymbol{w}} \in \boldsymbol{H}(\text{curl}, \widehat{\Omega}_e)$, $\widehat{\boldsymbol{d}} \in \boldsymbol{H}(\text{div}, \widehat{\Omega}_e)$, and $\widehat{q} \in L^2(\widehat{\Omega}_e)$, that are assumed to be free of geometry errors. Moreover, let us denote the corresponding functions of the interpolated global element by $v \in H_1(\Omega_e^k)$, $\boldsymbol{w} \in \boldsymbol{H}(\text{curl}, \Omega_e^k)$, $\boldsymbol{d} \in \boldsymbol{H}(\text{div}, \Omega_e^k)$, and $q \in L^2(\Omega_e^k)$, that are determined by Piola transformations of the local functions. For interpolated geometry mappings, the exact transformations (2.61) need to be replaced by kth-order Piola transformations, which utilize the interpolated geometry mapping $\boldsymbol{g}_k$ of order k and its Jacobian $\boldsymbol{J}_k = \left[\nabla \boldsymbol{g}_k^T\right]$:

$$v = P_{0,k}\left\{\widehat{v}\right\} = \widehat{v} \circ \boldsymbol{g}_k^{-1}, \tag{3.54a}$$
$$\boldsymbol{w} = P_{1,k}\left\{\widehat{\boldsymbol{w}}\right\} = \left(\boldsymbol{J}_k^{-1}\widehat{\boldsymbol{w}}\right) \circ \boldsymbol{g}_k^{-1}, \tag{3.54b}$$
$$\boldsymbol{d} = P_{2,k}\left\{\widehat{\boldsymbol{d}}\right\} = \left(\det(\boldsymbol{J}_k)^{-1}\boldsymbol{J}_k^T\widehat{\boldsymbol{d}}\right) \circ \boldsymbol{g}_k^{-1}, \tag{3.54c}$$
$$q = P_{3,k}\left\{\widehat{q}\right\} = \left(\det(\boldsymbol{J}_k)^{-1}\widehat{q}\right) \circ \boldsymbol{g}_k^{-1}. \tag{3.54d}$$

Assuming that the interpolated geometry mappings are bijective over the elements, the inverse Piola transformations of order k are obtained as:

$$\widehat{v} = P_{0,k}^{-1}\left\{v\right\} = v \circ \boldsymbol{g}_k, \tag{3.55a}$$
$$\widehat{\boldsymbol{w}} = P_{1,k}^{-1}\left\{\boldsymbol{w}\right\} = \boldsymbol{J}_k(\boldsymbol{w} \circ \boldsymbol{g}_k), \tag{3.55b}$$
$$\widehat{\boldsymbol{d}} = P_{2,k}^{-1}\left\{\boldsymbol{d}\right\} = \det(\boldsymbol{J}_k)\boldsymbol{J}_k^{-T}(\boldsymbol{d} \circ \boldsymbol{g}_k), \tag{3.55c}$$
$$\widehat{q} = P_{3,k}^{-1}\left\{q\right\} = \det(\boldsymbol{J}_k)(q \circ \boldsymbol{g}_k). \tag{3.55d}$$

Therefore, the finite element approximation of the solution may be exact on the reference element, but the transformed fields on the global element are affected by the geometry interpolation error. The main issue with this error is the possibility of dominating the finite element approximation error or even degrading the convergence rate. In the following sections, the sufficient order of geometry interpolation is investigated that leaves the finite element convergence unaffected.

3.1.4 Effect of the Interpolated Geometry on the H^1 Interpolation Operator

In order to show the effect of the interpolated geometry mapping, we introduce the geometric distortion operator Π_k^v of order k, which acts on an arbitrary function $v \in H^1(\Omega_e)$ as

$$\Pi_k^v v = P_{0,k}\left\{\widehat{v}\right\} = P_{0,k}\left\{P_0^{-1}\left\{v\right\}\right\} = v \circ \boldsymbol{g} \circ \boldsymbol{g}_k^{-1}. \tag{3.56}$$

This geometric distortion is equivalent to the map of the corresponding exact local function $\widehat{v}$ from the reference element to the global element by the interpolation based mapping $P_{0,k}$. Hence, this geometric distortion operator fits into the presented finite element process, where the approximated fields of the reference element are mapped to the real element via an interpolated geometry mapping

$$\Pi_k^v \pi_p^v v = \Pi_k^v v_p = P_{0,k}\left\{\widehat{\pi}_p^v \widehat{v}\right\}. \tag{3.57}$$

In practice, $\Pi_k^v \pi_p^v$ is the only available interpolation operator because the exact geometry mapping is usually not available, and thus π_p^v is unknown.

Next, an estimate for the geometric distortion error in $L^2(\widetilde{\Omega}_e^k)$-norm is derived for functions $v_p = \pi_p^v v$ over the intersection (3.53) of domains Ω_e and Ω_e^k. First, the geometry mapping is separated into the sum of the interpolated mapping and the error term as $\boldsymbol{g} = \boldsymbol{g}_k + \boldsymbol{R}_k$. Second, v_p is assumed to be $\max\{p,k\}$ times differentiable over the element, which allows the representation

$$v_p(\boldsymbol{r}) = a + \boldsymbol{b}\cdot\boldsymbol{r} + \widetilde{v}_p^2(\boldsymbol{r}), \qquad \text{for } \boldsymbol{r} \in \widetilde{\Omega}_e^k, \tag{3.58}$$

with a constant scalar a, constant vector $\boldsymbol{b}$, and a second order remainder term $\widetilde{v}_p^2$. Then, the geometric distortion of v_p can be written as

$$\begin{aligned}
\Pi_k^v v_p &= v_p \circ (\boldsymbol{g}_k + \boldsymbol{R}_k) \circ \boldsymbol{g}_k^{-1} = v_p \circ (\boldsymbol{r} + \boldsymbol{R}_k \circ \boldsymbol{g}_k^{-1}),\\
&= (a + \boldsymbol{b}\cdot\boldsymbol{r} + \widetilde{v}_p^2) \circ (\boldsymbol{r} + \boldsymbol{R}_k \circ \boldsymbol{g}_k^{-1}),\\
&= a + \boldsymbol{b}\cdot\boldsymbol{r} + \boldsymbol{b}\cdot\boldsymbol{R}_k \circ \boldsymbol{g}_k^{-1} + \widetilde{v}_p^2 \circ (\boldsymbol{r} + \boldsymbol{R}_k \circ \boldsymbol{g}_k^{-1}).
\end{aligned} \tag{3.59}$$

Thus, using (3.59), the initial estimation of the distortion error reads

$$\begin{aligned}
\|v_p - \Pi_k^v v_p\|_{L^2(\widetilde{\Omega}_e^k)} &= \left\|\boldsymbol{b}\cdot\boldsymbol{R}_k \circ \boldsymbol{g}_k^{-1} + \widetilde{v}_p^2 - \widetilde{v}_p^2 \circ (\boldsymbol{r} + \boldsymbol{R}_k \circ \boldsymbol{g}_k^{-1})\right\|_{L^2(\widetilde{\Omega}_e^k)},\\
\|v_p - \Pi_k^v v_p\|_{L^2(\widetilde{\Omega}_e^k)} &\le \left\|\boldsymbol{b}\cdot\boldsymbol{R}_k \circ \boldsymbol{g}_k^{-1}\right\|_{L^2(\widetilde{\Omega}_e^k)} + \left\|\widetilde{v}_p^2 - \widetilde{v}_p^2 \circ (\boldsymbol{r} + \boldsymbol{R}_k \circ \boldsymbol{g}_k^{-1})\right\|_{L^2(\widetilde{\Omega}_e^k)},\\
\|v_p - \Pi_k^v v_p\|_{L^2(\widetilde{\Omega}_e^k)} &\le C\,\|v_p\|_{H^p(\widetilde{\Omega}_e^k)} \left\|\boldsymbol{R}_k \circ \boldsymbol{g}_k^{-1}\right\|_{L^\infty(\widetilde{\Omega}_e^k)}\\
&\quad + \left\|\widetilde{v}_p^2 - \widetilde{v}_p^2 \circ (\boldsymbol{r} + \boldsymbol{R}_k \circ \boldsymbol{g}_k^{-1})\right\|_{L^2(\widetilde{\Omega}_e^k)},
\end{aligned} \tag{3.60}$$

where C is a positive constant independent of k. Here, it can be shown that the remainder term converges exponentially to zero:

$$\left\|\boldsymbol{R}_k \circ \boldsymbol{g}_k^{-1}\right\|_{L^\infty(\widetilde{\Omega}_e^k)} \le \left\|\boldsymbol{R}_k \circ \boldsymbol{g}_k^{-1}\right\|_{L^\infty(\Omega_e^k)} \overset{(3.38)}{=} \|\boldsymbol{R}_k\|_{L^\infty(\widehat{\Omega}_e)} \overset{(3.33)}{\le} C h^{\lceil (k+1)/r \rceil}. \tag{3.61}$$

Moreover, the higher-order term with $\widetilde{v}_p^2$ in (3.60) should converge with at least the same rate as the first-order term

$$\left\|\widetilde{v}_p^2 - \widetilde{v}_p^2 \circ (\boldsymbol{r} + \boldsymbol{R}_k \circ \boldsymbol{g}_k^{-1})\right\|_{L^2(\widetilde{\Omega}_e^k)} \leq C \left\|v_p\right\|_{H^p(\widetilde{\Omega}_e^k)} \left\|\boldsymbol{R}_k \circ \boldsymbol{g}_k^{-1}\right\|_{L^\infty(\widetilde{\Omega}_e^k)}, \tag{3.62}$$

due to the higher or equal order power of the error term (3.61). Thus, the geometric distortion error of an H^1-conforming function is

$$\left\|v_p - \Pi_k^v v_p\right\|_{L^2(\widetilde{\Omega}_e^k)} \leq C h^{\lceil (k+1)/r \rceil} \left\|v_p\right\|_{H^p(\widetilde{\Omega}_e^k)}, \tag{3.63}$$

where C is some constant, independent of h. Hence, if the interpolation order of the geometry is determined as $k = r(p-1)$, (3.63) leads to an $\mathcal{O}(h^p)$ rate of asymptotic convergence, i.e.,

$$\left\|v_p - \Pi_k^v v_p\right\|_{L^2(\widetilde{\Omega}_e^k)} \leq C h^p \left\|v_p\right\|_{H^p(\widetilde{\Omega}_e^k)}. \tag{3.64}$$

For the case of affine refinements ($r = 1$), the required interpolation order $k = p-1$ is minimal. However, for non-affine refinements, a higher interpolation order is needed to achieve the same algebraic rate of convergence. In conclusion, for $k = r(p-1)$ order of geometry interpolations, it is permissible to replace the ideal finite element interpolation operator π_p^v by the practically available interpolation operator $\Pi_k^v \pi_p^v$ because the rate of convergence,

$$\left\|\pi_p^v v - \Pi_k^v \pi_p^v v\right\|_{L^2(\widetilde{\Omega}_e^k)} \leq C h^p \left\|\pi_p^v v\right\|_{H^p(\widetilde{\Omega}_e^k)} \leq C h^p \left\|v\right\|_{H^p(\widetilde{\Omega}_e^k)}, \tag{3.65}$$

is identical to the expected finite element convergence.

3.1.5 Effect of the Interpolated Geometry on the $\boldsymbol{H}(\mathrm{curl})$ Interpolation Operator

This section proceeds by considering a geometric distortion operator Π_k^w for $\boldsymbol{w} \in \boldsymbol{H}(\mathrm{curl}, \Omega_e)$ functions, similarly to the previous case. The geometrical distortion of a function $\boldsymbol{w}$ is required to be equivalent to the corresponding function $\widehat{\boldsymbol{w}}$ on the reference element, which is mapped to the real element via the Piola transformation of order k:

$$\begin{aligned}\Pi_k^w \boldsymbol{w} &= P_{1,k}\left\{\widehat{\boldsymbol{w}}\right\} = P_{1,k}\left\{P_1^{-1}\left\{\boldsymbol{w}\right\}\right\},\\ &= \left(\boldsymbol{J}_k^{-1}\boldsymbol{J}\right) \circ \boldsymbol{g}_k^{-1}\left(\boldsymbol{w} \circ \boldsymbol{g} \circ \boldsymbol{g}_k^{-1}\right),\\ &= \left(\left[\widehat{\nabla}\boldsymbol{g}_k^T\right]^{-1}\left[\widehat{\nabla}\boldsymbol{g}^T\right]\right) \circ \boldsymbol{g}_k^{-1}\left(\boldsymbol{w} \circ \boldsymbol{g} \circ \boldsymbol{g}_k^{-1}\right). \end{aligned} \tag{3.66}$$

Therefore, this operator is suitable to express the practically available interpolation operator $\Pi_k^w \pi_p^w$, which is an interpolation over the reference element and a mapping to the real element via an interpolation-based function mapping:

$$\Pi_k^w \pi_p^w \boldsymbol{w} = \Pi_k^w \boldsymbol{w}_p = P_{1,k}\left\{\widehat{\pi}_p^w \widehat{\boldsymbol{w}}\right\}. \tag{3.67}$$

The first step of the distortion error estimation for functions $\boldsymbol{w}_p = \pi_p^w \boldsymbol{w}$ is to rewrite the geometry mapping as the sum of the interpolated geometry and an error term, which results in a separable Jacobian $\boldsymbol{J} = \left[\widehat{\nabla}\boldsymbol{g}^T\right] = \left[\widehat{\nabla}\boldsymbol{g}_k^T\right] + \left[\widehat{\nabla}\boldsymbol{R}_k^T\right]$. Hence, the geometric distortion (3.66) can be rewritten as

$$\Pi_k^w \boldsymbol{w}_p = \left(\boldsymbol{I} + \left(\left[\widehat{\nabla}\boldsymbol{g}_k^T\right]^{-1} \circ \boldsymbol{g}_k^{-1}\right)\left(\left[\widehat{\nabla}\boldsymbol{R}_k^T\right] \circ \boldsymbol{g}_k^{-1}\right)\right)\left(\boldsymbol{w}_p \circ \boldsymbol{g} \circ \boldsymbol{g}_k^{-1}\right), \tag{3.68}$$

where $\boldsymbol{I}$ is the identity matrix. The initial $\boldsymbol{L}^2(\widetilde{\Omega}_e^k)$-norm error estimation can be obtained by utilizing the triangle inequality, as

$$\begin{aligned}\left\|\boldsymbol{w}_p - \Pi_k^w \boldsymbol{w}_p\right\|_{\boldsymbol{L}^2(\widetilde{\Omega}_e^k)} \leq{}& \left\|\boldsymbol{w}_p - \boldsymbol{w}_p \circ \boldsymbol{g} \circ \boldsymbol{g}_k^{-1}\right\|_{\boldsymbol{L}^2(\widetilde{\Omega}_e^k)} + \\ &\left\|\left(\left[\widehat{\nabla}\boldsymbol{g}_k^T\right]^{-1} \circ \boldsymbol{g}_k^{-1}\right)\left(\left[\widehat{\nabla}\boldsymbol{R}_k^T\right] \circ \boldsymbol{g}_k^{-1}\right)\left(\boldsymbol{w}_p \circ \boldsymbol{g} \circ \boldsymbol{g}_k^{-1}\right)\right\|_{\boldsymbol{L}^2(\widetilde{\Omega}_e^k)}. \end{aligned} \tag{3.69}$$

The first term of (3.69) can be further simplified using the estimate

$$\left\|\boldsymbol{w}_p - \boldsymbol{w}_p \circ \boldsymbol{g} \circ \boldsymbol{g}_k^{-1}\right\|_{\boldsymbol{L}^2(\widetilde{\Omega}_e^k)} \leq C h^{\lceil (k+1)/r \rceil} \left\|\boldsymbol{w}_p\right\|_{\boldsymbol{H}^p(\widetilde{\Omega}_e^k)}, \tag{3.70}$$

which is obtained the same way as (3.63) for (3.60).

The simplification of the second term in (3.69) requires further attention since it consists of three multiplicative factors.

Using the generalized Hölder's inequality, it is possible to obtain an estimate with the multiplicative factor of estimates for each term:

$$\left\|\boldsymbol{w}_p \circ \boldsymbol{g} \circ \boldsymbol{g}_k^{-1}\right\|_{\boldsymbol{L}^2(\widetilde{\Omega}_e^k)} \leq C \left\|\boldsymbol{w}_p\right\|_{\boldsymbol{L}^2(\widetilde{\Omega}_e^k)}, \tag{3.71a}$$

$$\left\|\left[\widehat{\nabla}\boldsymbol{g}_k^T\right]^{-1} \circ \boldsymbol{g}_k^{-1}\right\|_{L^\infty(\widetilde{\Omega}_e^k)} \overset{(3.37)}{=} \left\|\nabla(\boldsymbol{g}_k^{-1})^T\right\|_{L^\infty(\widetilde{\Omega}_e^k)} \overset{(3.48)}{\leq} Ch^{-1}, \tag{3.71b}$$

$$\left\|\left[\widehat{\nabla}\boldsymbol{R}_k^T\right] \circ \boldsymbol{g}_k^{-1}\right\|_{L^\infty(\widetilde{\Omega}_e^k)} \overset{(3.38)}{\leq} \left\|\widehat{\nabla}\boldsymbol{R}_k^T\right\|_{L^\infty(\widehat{\Omega}_e)} \overset{(3.34)}{\leq} Ch^{\lceil (k+1)/r \rceil}. \tag{3.71c}$$

Hence, the estimate of the second term in (3.69) becomes

$$\left\|\left[\widehat{\nabla}\boldsymbol{g}_k^T\right]^{-1} \circ \boldsymbol{g}_k^{-1}\left[\widehat{\nabla}\boldsymbol{R}_k^T\right] \circ \boldsymbol{g}_k^{-1}\boldsymbol{w}_p \circ \boldsymbol{g} \circ \boldsymbol{g}_k^{-1}\right\|_{\boldsymbol{L}^2(\widetilde{\Omega}_e^k)} \leq Ch^{\lceil (k+1)/r \rceil - 1} \left\|\boldsymbol{w}_p\right\|_{\boldsymbol{L}^2(\widetilde{\Omega}_e^k)}. \tag{3.72}$$

The final error estimate of the geometric distortion error is obtained by the substitution of (3.70) and (3.72) into (3.69),

$$\left\|\boldsymbol{w}_p - \Pi_k^w \boldsymbol{w}_p\right\|_{\boldsymbol{L}^2(\widetilde{\Omega}_e^k)} \leq Ch^{\lceil (k+1)/r \rceil - 1} \left\|\boldsymbol{w}_p\right\|_{\boldsymbol{H}^p(\widetilde{\Omega}_e^k)}, \tag{3.73}$$

where C is some constant, independent of h. Choosing the order of geometry interpolation $k = rp$, the asymptotic convergence rate of the distortion error is $\mathcal{O}(h^p)$ due to the estimate of

$$\left\|\boldsymbol{w}_p - \Pi_k^w \boldsymbol{w}_p\right\|_{\boldsymbol{L}^2(\widetilde{\Omega}_e^k)} \leq Ch^p \left\|\boldsymbol{w}_p\right\|_{\boldsymbol{H}^p(\widetilde{\Omega}_e^k)}. \tag{3.74}$$

For affine refinements ($r = 1$), the classical isoparametric result is achieved, which requires the same order of geometry and field representation $k = p$. The consequence is the permissible replacement of the ideal interpolation operator π_p^w by the practically available interpolation operator $\Pi_k^w \pi_p^w$ without affecting the expected convergence rate:

$$\left\|\pi_p^w \boldsymbol{w} - \Pi_k^w \pi_p^w \boldsymbol{w}\right\|_{\boldsymbol{L}^2(\widetilde{\Omega}_e^k)} \leq Ch^p \left\|\pi_p^w \boldsymbol{w}\right\|_{\boldsymbol{H}^p(\widetilde{\Omega}_e^k)} \leq Ch^p \left\|\boldsymbol{w}\right\|_{\boldsymbol{H}^p(\widetilde{\Omega}_e^k)}. \tag{3.75}$$

3.1.6 Effect of the Interpolated Geometry on the $\boldsymbol{H}(\mathrm{div})$ Interpolation Operator

The definition of the geometric distortion operator Π_k^d for functions $\boldsymbol{d} \in \boldsymbol{H}(\mathrm{div}, \Omega_e)$ is analogous to that of previous cases. The distortion of a function is required to be the same as the corresponding exact local function $\widehat{\boldsymbol{d}}$, which is mapped to the global element by a Piola transformation of order k, as

$$\begin{aligned}
\Pi_k^d \boldsymbol{d} &= P_{2,k}\left\{\widehat{\boldsymbol{d}}\right\} = P_{2,k}\left\{P_2^{-1}\left\{\boldsymbol{d}\right\}\right\}, \\
&= \left(\left[\det\left(\boldsymbol{J}_k\right)^{-1}\boldsymbol{J}_k^T \det\left(\boldsymbol{J}\right)\boldsymbol{J}^{-T}\right]\boldsymbol{d}\circ\boldsymbol{g}\right)\circ\boldsymbol{g}_k^{-1}, \\
&= \left[\mathrm{adj}\left(\boldsymbol{J}_k^{-1}\boldsymbol{J}\right)^T\circ\boldsymbol{g}_k^{-1}\right]\boldsymbol{d}\circ\boldsymbol{g}\circ\boldsymbol{g}_k^{-1}, \\
&= \left[\mathrm{adj}\left(\boldsymbol{I}+\left[\widehat{\nabla}\boldsymbol{g}_k^T\right]^{-1}\left[\widehat{\nabla}\boldsymbol{R}_k^T\right]\right)^T\circ\boldsymbol{g}_k^{-1}\right]\boldsymbol{d}\circ\boldsymbol{g}\circ\boldsymbol{g}_k^{-1}, \\
&= \boldsymbol{d}\circ\boldsymbol{g}\circ\boldsymbol{g}_k^{-1} - \left[\boldsymbol{I}-\mathrm{adj}\left(\boldsymbol{I}+\left[\widehat{\nabla}\boldsymbol{g}_k^T\right]^{-1}\circ\boldsymbol{g}_k^{-1}\left[\widehat{\nabla}\boldsymbol{R}_k^T\right]\circ\boldsymbol{g}_k^{-1}\right)^T\right]\boldsymbol{d}\circ\boldsymbol{g}\circ\boldsymbol{g}_k^{-1}.
\end{aligned} \tag{3.76}$$

Therefore, it is suitable to express the practically available interpolation operator $\Pi_k^d\pi_p^d$, which is an interpolation operator over the reference domain and a corresponding interpolation-based Piola transformation:

$$\Pi_k^d\pi_p^d\boldsymbol{d} = \Pi_k^d\boldsymbol{d}_p = P_{2,k}\left\{\widehat{\pi}_p^d\widehat{\boldsymbol{d}}\right\}. \tag{3.77}$$

Further simplification of (3.76) is not possible in a closed form. The initial estimate for the geometric distortion error for functions $\boldsymbol{d}_p = \pi_p^d\boldsymbol{d}$ is obtained by the application of the triangle and the generalized Hölder's inequality, by taking the norm of each term:

$$\begin{aligned}
\left\|\boldsymbol{d}_p - \Pi_k^d\boldsymbol{d}_p\right\|_{\boldsymbol{L}^2(\tilde{\Omega}_e^k)} \leq{}& \left\|\boldsymbol{d}_p - \boldsymbol{d}_p\circ\boldsymbol{g}\circ\boldsymbol{g}_k^{-1}\right\|_{\boldsymbol{L}^2(\tilde{\Omega}_e^k)} + \\
&\left\|\boldsymbol{I}-\mathrm{adj}\left(\boldsymbol{I}+\left[\widehat{\nabla}\boldsymbol{g}_k^T\right]^{-1}\circ\boldsymbol{g}_k^{-1}\left[\widehat{\nabla}\boldsymbol{R}_k^T\right]\circ\boldsymbol{g}_k^{-1}\right)^T\right\|_{L^\infty(\tilde{\Omega}_e^k)}\left\|\boldsymbol{d}_p\circ\boldsymbol{g}\circ\boldsymbol{g}_k^{-1}\right\|_{\boldsymbol{L}^2(\tilde{\Omega}_e^k)}.
\end{aligned} \tag{3.78}$$

Next, it is possible to use that the term $\left[\widehat{\nabla}\boldsymbol{g}_k^T\right]^{-1}\circ\boldsymbol{g}_k^{-1}\left[\widehat{\nabla}\boldsymbol{R}_k^T\right]\circ\boldsymbol{g}_k^{-1}$ is a matrix function over $\mathbb{R}^3\oplus\mathbb{R}^3$ and all components are expected to converge to zero, at least, with the order of $\mathcal{O}(h)$; see (3.71). Otherwise, there is no convergence for $k+1 \leq r$ due to the estimate (3.72) with $\boldsymbol{w}_p = \boldsymbol{d}_p$. For such a matrix, with arbitrary coefficients of a, b, c, d, e, f, g, h, and i, the following identity applies:

$$\boldsymbol{I}-\mathrm{adj}\left(\boldsymbol{I}+\begin{bmatrix}a & b & c\\ d & e & f\\ g & h & i\end{bmatrix}\right) = \begin{bmatrix}-e-i & b & c\\ d & -a-i & f\\ g & h & -a-e\end{bmatrix} + \begin{bmatrix}fh-ei & bi-ch & ce-bf\\ di-fg & cg-ai & af-cd\\ eg-dh & ah-bg & bd-ae\end{bmatrix}, \tag{3.79}$$

where the convergence rate of the second term is always higher than or equal to the first term. Thus, the second term can be neglected in the following estimations. Moreover, using the fact that

$$\left\| \begin{bmatrix} -e-i & b & c \\ d & -a-i & f \\ g & h & -a-e \end{bmatrix} \right\|_{L^\infty} < C \left\| \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \right\|_{L^\infty}, \tag{3.80}$$

for a positive constant C, it is permissible to get an estimate in the form of

$$\left\| \boldsymbol{I} - \operatorname{adj}\left(\boldsymbol{I} + \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \right) \right\|_{L^\infty} \leq C \left\| \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}^T \right\|_{L^\infty}. \tag{3.81}$$

Therefore, the equivalent estimation for the second term of (3.78) reads

$$\left\| \boldsymbol{I} - \operatorname{adj}\left(\boldsymbol{I} + \left[\widehat{\nabla} \boldsymbol{g}_k^T\right]^{-1} \circ \boldsymbol{g}_k^{-1} \left[\widehat{\nabla} \boldsymbol{R}_k^T\right] \circ \boldsymbol{g}_k^{-1} \right)^T \right\|_{L^\infty(\widetilde{\Omega}_e^k)} \leq C \left\| \left[\widehat{\nabla} \boldsymbol{g}_k^T\right]^{-1} \circ \boldsymbol{g}_k^{-1} \left[\widehat{\nabla} \boldsymbol{R}_k^T\right] \circ \boldsymbol{g}_k^{-1} \right\|_{L^\infty(\widetilde{\Omega}_e^k)}. \tag{3.82}$$

The application of (3.82) for (3.78) results in

$$\begin{aligned} \left\| \boldsymbol{d}_p - \Pi_k^d \boldsymbol{d}_p \right\|_{\boldsymbol{L}^2(\widetilde{\Omega}_e^k)} \leq{}& \left\| \boldsymbol{d}_p - \boldsymbol{d}_p \circ \boldsymbol{g} \circ \boldsymbol{g}_k^{-1} \right\|_{\boldsymbol{L}^2(\widetilde{\Omega}_e^k)} + \\ & C \left\| \left[\widehat{\nabla} \boldsymbol{g}_k^T\right]^{-1} \circ \boldsymbol{g}_k^{-1} \left[\widehat{\nabla} \boldsymbol{R}_k^T\right] \circ \boldsymbol{g}_k^{-1} \right\|_{\boldsymbol{L}^\infty(\widetilde{\Omega}_e^k)} \left\| \boldsymbol{d}_p \circ \boldsymbol{g} \circ \boldsymbol{g}_k^{-1} \right\|_{\boldsymbol{L}^2(\widetilde{\Omega}_e^k)}. \end{aligned} \tag{3.83}$$

Notice that this form is equivalent to that of (3.69), with the replacement of $\boldsymbol{w}_p$ by $\boldsymbol{d}_p$. Thus, the same estimate applies as (3.73),

$$\left\| \boldsymbol{d}_p - \Pi_k^d \boldsymbol{d}_p \right\|_{\boldsymbol{L}^2(\widetilde{\Omega}_e^k)} \leq C h^{\lceil (k+1)/r \rceil - 1} \left\| \boldsymbol{d}_p \right\|_{\boldsymbol{H}^p(\widetilde{\Omega}_e^k)}, \tag{3.84}$$

where C is some constant, independent of h. Choosing the interpolation order of the geometry $k = pr$, the asymptotic convergence rate of the distortion error becomes $\mathcal{O}(h^p)$ due to the estimate

$$\left\| \boldsymbol{d}_p - \Pi_k^d \boldsymbol{d}_p \right\|_{\boldsymbol{L}^2(\widetilde{\Omega}_e^k)} \leq C h^p \left\| \boldsymbol{d}_p \right\|_{\boldsymbol{H}^p(\widetilde{\Omega}_e^k)}. \tag{3.85}$$

Similar to the previous cases, by setting $r = 1$, the theoretical result of the isoparametric elements is obtained. If the convergence rate of the distortion error is $\mathcal{O}(h^p)$, it is permissible to replace the ideal interpolation operator by the practically available interpolation operator as it does not affect the expected convergence rate:

$$\left\| \pi_p^d \boldsymbol{d} - \Pi_k^d \pi_p^d \boldsymbol{d} \right\|_{\boldsymbol{L}^2(\widetilde{\Omega}_e^k)} \leq C h^p \left\| \pi_p^d \boldsymbol{d} \right\|_{\boldsymbol{H}^p(\widetilde{\Omega}_e^k)} \leq C h^p \left\| \boldsymbol{d} \right\|_{\boldsymbol{H}^p(\widetilde{\Omega}_e^k)}. \tag{3.86}$$

3.1.7 Asymptotic Limit of the Finite Element Metrics

The effects of the function mappings on the interpolation operators have been analyzed in the previous sections. However, their effect on finite element formulations is also an important issue. These formulations are typically based on matrices and vectors, which are obtained via some inner product integrals. These integrals are evaluated over the reference element, where the effect of the geometry mapping and the material properties, or generally speaking some metric tensors of the global element, are contained by a single metric tensor. As the geometry mappings of the curvilinear elements are not affine transformations, whether they are exact or interpolated mappings, the corresponding function mappings result in position-dependent metrics whose components are typically rational functions. Therefore, exact integration is not possible in general, and some numerical integration schemes are necessary. Most of these methods are based on the interpolation or approximation of the integrand via some known functions, commonly polynomials, that have a known exact integral. Therefore, it is crucial to know the required order of polynomials or the required number of sampling points. In the following, the necessary representation of the local metric terms is investigated which leaves the expected asymptotic rate of convergence unaffected.

Let us consider the first Strang lemma [Cia02], which states that the finite element error is bounded, up to a constant factor, by the sum of the approximation error and the consistency errors. Let $I(\boldsymbol{a}, \boldsymbol{b})$ denote an elliptic bilinear form with a metric tensor $\boldsymbol{\chi}$ for functions $\boldsymbol{a}, \boldsymbol{b} \in \mathcal{I}_p \in \{\mathcal{V}_p, \mathcal{W}_p, \mathcal{D}_p, \mathcal{P}_p\}$. Furthermore, let $I_K(\boldsymbol{a}, \boldsymbol{b})$ denote the corresponding approximated bilinear form, which is obtained element by element via the approximation of the metric tensor over the reference element by polynomials of order K. Then, the consistency error of the bilinear form over a single element is defined [Cia02] as

$$\sup_{\boldsymbol{a} \in \mathcal{I}_p} \frac{|I(\boldsymbol{a}, \boldsymbol{b}) - I_K(\boldsymbol{a}, \boldsymbol{b})|}{\|\boldsymbol{a}\|_{0,1,\Omega}}. \tag{3.87}$$

Since the initial bilinear form is defined over the global domain and the considered metric approximation is performed element by element over the reference element, the functions and the integration domain need to be mapped to the reference element. The mapping of $\boldsymbol{a}$ and $\boldsymbol{b}$ between the reference and the real element is given as

$$\boldsymbol{a} = (\boldsymbol{M}\,\widehat{\boldsymbol{a}}) \circ \boldsymbol{g}^{-1}, \quad \text{and} \quad \boldsymbol{b} = (\boldsymbol{M}\,\widehat{\boldsymbol{b}}) \circ \boldsymbol{g}^{-1}, \tag{3.88}$$

where

$$\boldsymbol{M} \in \left\{1, \boldsymbol{J}^{-1}, (\det \boldsymbol{J})^{-1} \boldsymbol{J}^T, (\det \boldsymbol{J})^{-1}\right\} \tag{3.89}$$

is an invertible matrix over $\widehat{\Omega}_e$, whose components $[\boldsymbol{M}]_{ij} \in W^{K+1,\infty}(\widehat{\Omega}_e)$ are only dependent of the Jacobian. Thus, the corresponding inverse transformations are

$$\widehat{\boldsymbol{a}} = \boldsymbol{M}^{-1}\,\boldsymbol{a} \circ \boldsymbol{g}, \quad \text{and} \quad \widehat{\boldsymbol{b}} = \boldsymbol{M}^{-1}\,\boldsymbol{b} \circ \boldsymbol{g}. \tag{3.90}$$

Using the previously introduced notations, the change of the integration domain reads

$$\int_{\Omega_e} \boldsymbol{a} \cdot \boldsymbol{\chi} \boldsymbol{b} \, \mathrm{d}\boldsymbol{r} = \int_{\widehat{\Omega}_e} (\boldsymbol{M}\widehat{\boldsymbol{a}}) \cdot \boldsymbol{\chi} \circ \boldsymbol{g}(\boldsymbol{M}\widehat{\boldsymbol{b}}) \det(\boldsymbol{J}) \, \mathrm{d}\widehat{\boldsymbol{r}} = \int_{\widehat{\Omega}_e} \widehat{\boldsymbol{a}} \cdot \boldsymbol{\Lambda} \widehat{\boldsymbol{b}} \, \mathrm{d}\widehat{\boldsymbol{r}}, \tag{3.91}$$

where the local metric tensor is expressed as

$$\boldsymbol{\Lambda} = \boldsymbol{M}^T (\boldsymbol{\chi} \circ \boldsymbol{g}) \boldsymbol{M} \det(\boldsymbol{J}). \tag{3.92}$$

For the simplification of the notations, only a single, arbitrary component of the metric $\Lambda = [\boldsymbol{\Lambda}]_{ij}$ and the corresponding field components $\widehat{a}(\widehat{\boldsymbol{r}}) = [\widehat{\boldsymbol{a}}]_i$ and $\widehat{b}(\widehat{\boldsymbol{r}}) = [\widehat{\boldsymbol{b}}]_j$ are considered because the derivation is independent of i and j. Therefore, the general integral over the reference element yields

$$I = \int_{\widehat{\Omega}_e} \widehat{a}(\widehat{\boldsymbol{r}}) \Lambda(\widehat{\boldsymbol{r}}) \widehat{b}(\widehat{\boldsymbol{r}}) \, \mathrm{d}\widehat{\boldsymbol{r}}. \tag{3.93}$$

Next, an approximated metric term $\Lambda_K(\widehat{\boldsymbol{r}})$ of order K is considered, as well as the corresponding integral scheme

$$I_K = \int_{\widehat{\Omega}_e} \widehat{a}(\widehat{\boldsymbol{r}}) \Lambda_K(\widehat{\boldsymbol{r}}) \widehat{b}(\widehat{\boldsymbol{r}}) \, \mathrm{d}\widehat{\boldsymbol{r}}, \qquad \text{with } \Lambda_K \in \mathbb{P}_K. \tag{3.94}$$

The induced truncation error ϵ_K of the approximated metric can be written as

$$\epsilon_K = |I - I_K| = \left| \int_{\widehat{\Omega}_e} \widehat{a}\, \widehat{b}\, (\Lambda - \Lambda_K) \, \mathrm{d}\widehat{\boldsymbol{r}} \right|. \tag{3.95}$$

The first estimate is obtained by interchanging the integral and the absolute value operator and then applying the Cauchy-Schwarz inequality:

$$\epsilon_K \leq \left| (\widehat{a}\, \widehat{b}) \, (\Lambda - \Lambda_K) \right|_{0,1,\widehat{\Omega}_e}, \tag{3.96a}$$

$$\epsilon_K \leq \left| \widehat{a}\, \widehat{b} \right|_{0,2,\widehat{\Omega}_e} |\Lambda - \Lambda_K|_{0,2,\widehat{\Omega}_e}. \tag{3.96b}$$

Notice that the first term depends only on the local fields, whereas the second term is the metric error in $L^2(\widehat{\Omega}_e)$-norm. In the following, the result of (3.96b) is not used and an alternative estimate is obtained from (3.96a) via Hölder's inequality,

$$\epsilon_K \leq \left| \widehat{a}\, \widehat{b} \right|_{0,1,\widehat{\Omega}_e} |\Lambda - \Lambda_K|_{0,\infty,\widehat{\Omega}_e}, \tag{3.97}$$

which eases the derivations and helps to obtain the final result in a simpler form while providing the same theoretical conclusion.

The General Metric Truncation Error

The next step of the derivation is the application of the Bramble-Hilbert lemma [BS07, Section 4.3] to the metric truncation error, and applying a function composition with the identity mapping $\widehat{\boldsymbol{r}} = \boldsymbol{g}^{-1} \circ \boldsymbol{g}$:

$$\begin{aligned} |\Lambda - \Lambda_K|_{0,\infty,\widehat{\Omega}_e} &\leq C \sum_{|\alpha|=K+1} \left| \widehat{D}^{\alpha} \Lambda \right|_{0,\infty,\widehat{\Omega}_e}, \\ &\leq C \sum_{|\alpha|=K+1} \left| \widehat{D}^{\alpha} \left[(\Lambda \circ \boldsymbol{g}^{-1}) \circ \boldsymbol{g} \right] \right|_{0,\infty,\widehat{\Omega}_e}. \end{aligned} \tag{3.98}$$

Then, the general formula [Leo17, Theorem 11.54] is considered for the derivative function compositions of $\Lambda \circ \boldsymbol{g}^{-1}$ with $\boldsymbol{g}$, $|\alpha| = K+1$ times differentiable functions:

$$\widehat{D}^{\alpha} \left((\Lambda \circ \boldsymbol{g}^{-1}) \circ \boldsymbol{g} \right) = \sum_{\beta,\gamma,l} C_{\alpha,\beta,\gamma,l} \left(D^{\beta} (\Lambda \circ \boldsymbol{g}^{-1}) \right) \circ \boldsymbol{g} \prod_{i=1}^{|\beta|} \widehat{D}^{\gamma_i} \left[\boldsymbol{g} \right]_{l_i}. \tag{3.99}$$

Here, the multi-index β and γ_i are required to be $1 \leq |\beta| \leq |\alpha|$, $|\gamma_i| > 0$, and $\sum_{i=1}^{|\beta|} \gamma_i = \alpha$. Moreover, l is a set of indices $l = (l_1, \cdots, l_{|\beta|})$ for the vector components of $\boldsymbol{g}$ with the values of $l_i \in \{1, 2, 3\}$. The constants $C_{\alpha,\beta,\gamma,l}$ correspond to each multi-index and are independent of $(\Lambda \circ \boldsymbol{g}^{-1})$ and $\boldsymbol{g}$.

Estimates for the norm of each multiplicative term are developed separately because this process is equivalent to the application of Hölder's inequality for infinity norms. Considering that $\sum_{i=1}^{|\beta|} |\gamma_i| = |\alpha|$, the multiplicative term can be bounded by

$$\prod_{i=1}^{|\beta|} \left| \widehat{D}^{\gamma_i} \left[\boldsymbol{g} \right]_{l_i} \right|_{0,\infty,\widehat{\Omega}_e} \overset{(3.35)}{\leq} C \prod_{i=1}^{|\beta|} h^{\lceil |\gamma_i|/r \rceil} \leq C h^{\lceil |\alpha|/r \rceil}. \tag{3.100}$$

Applying the domain independence of the infinite norm (3.38) and (3.100) to the piecewise norm of each term in (3.99), the estimate of a derivative function composition becomes

$$\left| \widehat{D}^{\alpha} \left((\Lambda \circ \boldsymbol{g}^{-1}) \circ \boldsymbol{g} \right) \right|_{0,\infty,\widehat{\Omega}_e} \leq C \sum_{\beta} \left| (\Lambda \circ \boldsymbol{g}^{-1}) \right|_{|\beta|,\infty,\Omega_e} h^{\lceil |\alpha|/r \rceil}, \tag{3.101a}$$

$$\leq C h^{\lceil |\alpha|/r \rceil} \left\| (\Lambda \circ \boldsymbol{g}^{-1}) \right\|_{|\alpha|,\infty,\Omega_e}. \tag{3.101b}$$

Then, the general metric-independent result is obtained via the substitution of (3.101b) into (3.98) and then into (3.97):

$$\epsilon_K \leq C h^{\lceil (K+1)/r \rceil} \left| \widehat{a}\, \widehat{b} \right|_{0,1,\widehat{\Omega}_e} \left\| \Lambda \circ \boldsymbol{g}^{-1} \right\|_{K+1,\infty,\Omega_e}. \tag{3.102}$$

This estimation can be further detailed by considering the underlying structure of the metric and the function mappings. One can obtain an estimation for the first term of (3.102) by using Hölder's inequality and (3.90) as

$$\begin{aligned}
\left|\widehat{a}\,\widehat{b}\right|_{0,1,\widehat{\Omega}_e} &\leq C\,|\widehat{a}|_{0,2,\widehat{\Omega}_e}\left|\widehat{b}\right|_{0,2,\widehat{\Omega}_e}\,,\\
&\leq C\,|\widehat{\boldsymbol{a}}|_{0,2,\widehat{\Omega}_e}\left|\widehat{\boldsymbol{b}}\right|_{0,2,\widehat{\Omega}_e}\,,\\
&\leq C\left|\boldsymbol{M}^{-1}\boldsymbol{a}\circ\boldsymbol{g}\right|_{0,2,\widehat{\Omega}_e}\left|\boldsymbol{M}^{-1}\boldsymbol{b}\circ\boldsymbol{g}\right|_{0,2,\widehat{\Omega}_e}\,,\\
&\leq C\left|\boldsymbol{M}^{-1}\right|^2_{0,\infty,\widehat{\Omega}_e}|\boldsymbol{a}\circ\boldsymbol{g}|_{0,2,\widehat{\Omega}_e}|\boldsymbol{b}\circ\boldsymbol{g}|_{0,2,\widehat{\Omega}_e}\,,\\
&\leq C\left|\boldsymbol{M}^{-1}\right|^2_{0,\infty,\widehat{\Omega}_e}\left|(\det\boldsymbol{J})^{-1}\right|^{1/2}_{0,\infty,\widehat{\Omega}_e}|\boldsymbol{a}|_{0,2,\Omega_e}\left|(\det\boldsymbol{J})^{-1}\right|^{1/2}_{0,\infty,\widehat{\Omega}_e}|\boldsymbol{b}|_{0,2,\Omega_e}\,.
\end{aligned} \tag{3.103}$$

Similarly, an estimate can be obtained for the second term of (3.102) using Hölder's inequality and (3.92) as

$$\begin{aligned}
\left\|\Lambda\circ\boldsymbol{g}^{-1}\right\|_{K+1,\infty,\Omega_e} &\leq C\left\|\boldsymbol{\Lambda}\circ\boldsymbol{g}^{-1}\right\|_{K+1,\infty,\Omega_e}\,,\\
&\leq C\,\|\boldsymbol{M}\|^2_{K+1,\infty,\widehat{\Omega}_e}\,\|\det\boldsymbol{J}\|_{K+1,\infty,\widehat{\Omega}_e}\,\|\boldsymbol{\chi}\|_{K+1,\infty,\Omega_e}\,.
\end{aligned} \tag{3.104}$$

At this point, it is known that the norm of $\boldsymbol{M}$ and $\det\boldsymbol{J}$ is dependent on the first derivative of only $\boldsymbol{g}$ and $\boldsymbol{g}^{-1}$. Moreover, the norm of any higher-order derivative of $\boldsymbol{g}$ converges with a higher rate (3.35), whereas the norm of $\boldsymbol{g}^{-1}$ remains unchanged (3.48) for any higher-order derivative when affine refinement is used. Therefore,

$$\|\boldsymbol{M}\|_{K+1,\infty,\widehat{\Omega}_e} \leq C\,|\boldsymbol{M}|_{0,\infty,\widehat{\Omega}_e}\,, \tag{3.105a}$$

$$\|\det\boldsymbol{J}\|_{K+1,\infty,\widehat{\Omega}_e} \leq C\,|\det\boldsymbol{J}|_{0,\infty,\widehat{\Omega}_e}\,, \tag{3.105b}$$

and (3.104) yields

$$\left\|\Lambda\circ\boldsymbol{g}^{-1}\right\|_{K+1,\infty,\Omega_e} \leq C\,|\boldsymbol{M}|^2_{0,\infty,\widehat{\Omega}_e}\,|\det\boldsymbol{J}|_{0,\infty,\widehat{\Omega}_e}\,\|\boldsymbol{\chi}\|_{K+1,\infty,\Omega_e}\,. \tag{3.106}$$

The general result can be assembled by substituting in the estimates (3.103) and (3.106) into (3.102) with $r = 1$, yielding

$$\epsilon_K \leq Ch^{K+1}\,\|\boldsymbol{\chi}\|_{K+1,\infty,\Omega_e}\,|\boldsymbol{a}|_{0,2,\Omega_e}\,|\boldsymbol{b}|_{0,2,\Omega_e}\,. \tag{3.107}$$

The General Asymptotic Limit

The main result is that algebraic convergence can be achieved independently of the actual choice of the function mapping $\boldsymbol{M}$ (3.89) for affinely refined meshes. Using a testing function $\boldsymbol{a}$ that is independent of K and h, with a finite norm $|\boldsymbol{a}|_{0,2,\Omega_e} \leq C$, the general result (3.107) simplifies to

$$\epsilon_K \leq Ch^{K+1} \left\|\boldsymbol{\chi}\right\|_{K+1,\infty,\Omega_e} |\boldsymbol{b}|_{0,2,\Omega_e} . \tag{3.108}$$

It can be seen that the error depends on the smoothness of the global metric tensor with $\left\|\boldsymbol{\chi}\right\|_{K+1,\infty,\Omega_e}$ and also on the $\boldsymbol{L}^2$-norm of the unknown field with $|\boldsymbol{b}|_{0,2,\Omega_e}$. Note that the geometry mapping dependent terms in (3.103) and (3.106) cancel out in an asymptotic sense but contribute to the constant factor C in (3.107) and (3.108). Therefore, the smoothness of the geometry mappings contributes to the metric error and affects the region where the asymptotic rate of convergence starts.

Consequently, if the asymptotic convergence rate of the finite element is $\mathcal{O}(h^p)$ and the mesh is affinely refined, then the sufficient asymptotic limit for the metric expansion order is

$$K = p - 1 \tag{3.109}$$

since the convergence rate of the metric truncation error is the same:

$$\epsilon_K \leq Ch^{(p-1)+1} \left\|\boldsymbol{\chi}\right\|_{K+1,\infty,\Omega_e} |\boldsymbol{b}|_{0,2,\Omega_e} = Ch^p \left\|\boldsymbol{\chi}\right\|_{K+1,\infty,\Omega_e} |\boldsymbol{b}|_{0,2,\Omega_e} . \tag{3.110}$$

3.2 Errors Related to the Field Representation

In this section, general error estimates are provided for finite elements. The main focus is on the determination of appropriate requirements, that result in the convergence rate of $\mathcal{O}(h^p)$ for the fields $v \in H^1(\Omega_e)$, $\boldsymbol{w} \in \boldsymbol{H}(\mathrm{curl}, \Omega_e)$, and $\boldsymbol{d} \in \boldsymbol{H}(\mathrm{div}, \Omega_e)$ in $\boldsymbol{L}^2$-norm. Therefore, the finite element order and mesh parameter independent quantities, such as the quality of the initial mesh or the optimization of the interpolation locations, are not considered for minimizing the error.

3.2.1 H^1 Approximation Estimates

This section deals with the error estimation for a scalar function $v \in H^1(\Omega_e)$. In the proposed interpretation, the approximation is performed over the reference element and then mapped to the global element; see (2.83). Thus, the resulting local error can be mapped to the global domain by the appropriate Piola transformation

$$\left\| v - \pi_p^v v \right\|_{L^2(\Omega_e)} = \left\| P_0 \left\{ \widehat{v} - \widehat{\pi}_p^v \widehat{v} \right\} \right\|_{L^2(\Omega_e)} . \tag{3.111}$$

The first estimate is obtained by changing the global L^2-norm error to an error over the reference element

$$\left\| v - \pi_p^v v \right\|_{L^2(\Omega_e)} \leq C\, h^{3/2} \left\| \widehat{v} - \widehat{\pi}_p^v \widehat{v} \right\|_{L^2(\widehat{\Omega}_e)} , \tag{3.112}$$

where C is a positive constant and $h^{3/2}$ arises from the effect of the L^2-norm (square root) on the volume element

$$\sqrt{\int_{\widehat{\Omega}_e} \det\left(\boldsymbol{J}\right) \, \mathrm{d}\widehat{\boldsymbol{r}}} \leq C\sqrt{h^3}. \tag{3.113}$$

The Bramble-Hilbert lemma can be applied to the local error if the approximation space contains a polynomial space

$$\mathbb{P}_{p-1} \subseteq \widehat{\mathcal{V}}_p. \tag{3.114}$$

Then, the obtained local function is transformed to the global element via the inverse Piola transformation

$$\begin{aligned} \left\| v - \pi_p^v v \right\|_{L^2(\Omega_e)} &\leq C\, h^{3/2} \left\| \widehat{D}^\alpha \widehat{v} \right\|_{L^2(\widehat{\Omega}_e)} , \\ &\leq C\, h^{3/2} \left\| \widehat{D}^\alpha \left[P_0^{-1} \left\{ v \right\} \right] \right\|_{L^2(\widehat{\Omega}_e)} , \quad \text{with } |\alpha| = p. \end{aligned} \tag{3.115}$$

Next, the general formula for the derivative function compositions (3.19) is applied to calculate the pth order derivative

$$\widehat{D}^{\alpha}[P_0^{-1}\{v\}] = \widehat{D}^{\alpha}(v \circ \boldsymbol{g}) = \sum_{\beta,\gamma,l} C_{\beta,\gamma,l} \left((D^{\beta} v) \circ \boldsymbol{g}\right) \prod_{i=1}^{|\beta|} \widehat{D}^{\gamma_i} \boldsymbol{g}_{l_i}. \tag{3.116}$$

Here, the multi-indices β and γ_i are required to be $1 \leq |\beta| \leq p$, $|\gamma_i| > 0$ and $\sum_{i=1}^{|\beta|} \gamma_i = \alpha$. Moreover, l is a set of indices $l = (l_1, \cdots, l_{|\beta|})$ for the vector components of $\boldsymbol{g}$ with the values of $l_i \in \{1, 2, 3\}$. The constants $C_{\beta,\gamma,l}$ correspond to each multi-index and are independent of the field and the mesh parameter. Thus, using the triangle and generalized Hölder inequalities, the upper bound of the function composition reads

$$\begin{aligned} \left\| \widehat{D}^{\alpha}(v \circ \boldsymbol{g}) \right\|_{L^2(\widehat{\Omega}_e)} &\leq \sum_{\beta,\gamma,l} C_{\beta,\gamma,l} \left\| (D^{\beta} v) \circ \boldsymbol{g} \right\|_{L^2(\widehat{\Omega}_e)} \prod_{i=1}^{|\beta|} \left\| \widehat{D}^{\gamma_i} \boldsymbol{g}_{l_i} \right\|_{L^{\infty}(\widehat{\Omega}_e)}, \\ &\overset{(3.113)}{\leq} h^{-3/2} \sum_{\beta,\gamma,l} C_{\beta,\gamma,l} \left\| D^{\beta} v \right\|_{L^2(\Omega_e)} \prod_{i=1}^{|\beta|} \left\| \widehat{D}^{\gamma_i} \boldsymbol{g}_{l_i} \right\|_{L^{\infty}(\widehat{\Omega}_e)}. \end{aligned} \tag{3.117}$$

Next, the estimate of (3.33) is applied to the derivative of the geometry mapping

$$\left\| \widehat{D}^{\gamma_i} \boldsymbol{g}_{l_i} \right\|_{L^{\infty}(\widehat{\Omega}_e)} \leq |\boldsymbol{g}|_{|\gamma_i|,\infty,\widehat{\Omega}_e} \leq C h^{\lceil |\gamma_i|/r \rceil}. \tag{3.118}$$

Considering that $\sum_{i=1}^{|\beta|} |\gamma_i| = |\alpha| = p$, the multiplicative term can be bounded by

$$\prod_{i=1}^{|\beta|} \left\| \widehat{D}^{\gamma_i} \boldsymbol{g}_{l_i} \right\|_{L^{\infty}(\widehat{\Omega}_e)} \leq C_{\beta} \prod_{i=1}^{|\beta|} h^{\lceil |\gamma_i|/r \rceil} \leq C_{\beta} h^{\lceil p/r \rceil}, \tag{3.119}$$

where C_{β} is a constant. Applying (3.119) to (3.117), the estimation of a derivative function composition becomes

$$\begin{aligned} \left\| \widehat{D}^{\alpha}(v \circ \boldsymbol{g}) \right\|_{L^2(\widehat{\Omega}_e)} &\leq \sum_{\beta} C_{\beta} \left\| D^{\beta} v \right\|_{L^2(\Omega_e)} h^{-3/2 + \lceil p/r \rceil}, \\ &\leq C h^{-3/2 + \lceil p/r \rceil} \left\| v \right\|_{H^{|\alpha|}(\Omega_e)}. \end{aligned} \tag{3.120}$$

Then, the finite element error estimation is obtained via the substitution of (3.120) into (3.115), as

$$\left\| v - \pi_p^v v \right\|_{L^2(\Omega_e)} \leq C h^{\lceil p/r \rceil} \left\| v \right\|_{H^p(\Omega_e)}, \tag{3.121}$$

where C is a constant, independent of h and v. It can be seen that the desired rate is met only for affine refinements ($r = 1$) with the requirement of (3.114). However, there is a way to compensate for the effect of the higher-order refinements. The usual

procedure [FGM11] is to apply the Bramble-Hilbert lemma on the global element for (3.111). However, the difficulty with this approach is that the interpolant $\pi_p^v v$ is not a polynomial for a general curvilinear element, and the Bramble-Hilbert lemma is not directly applicable.

Up to this point, only the π_p^v interpolation operator has been considered, which produces no geometry error because the local interpolant is mapped to the global element via exact geometry mapping. However, in Section 3.1.4, it was shown that it is permissible to replace π_p^v by the practically available interpolation operator $\Pi_k^v \pi_p^v$, which maps the local interpolant to the global domain via an interpolated geometry mapping of order $k = r(p-1)$. Now that the geometry mapping is a polynomial, it is possible to apply the Bramble-Hilbert lemma on the global element by requiring that the global approximation space $\Pi_k^v \pi_p^v v \in \mathcal{V}_p$ contains a polynomial space

$$\mathbb{P}_{p-1} \subseteq \mathcal{V}_p \tag{3.122}$$

for arbitrary $P_{0,k}$ mappings. Thus, the estimate over the global element becomes

$$\left\| v - \Pi_k^v \pi_p^v v \right\|_{L^2(\widetilde{\Omega}_e^k)} \leq C h^p \left\| v \right\|_{H^p(\widetilde{\Omega}_e^k)}, \tag{3.123}$$

where C is a constant, independent of h and v.

3.2.2 $\boldsymbol{H}(\mathrm{curl})$ Approximation Estimates

The $\boldsymbol{L}^2(\Omega_e)$-norm error estimation for a $\boldsymbol{w} \in \boldsymbol{H}(\mathrm{curl}, \Omega_e)$ function is carried out similarly as for a $v \in H^1(\Omega_e)$ function in the previous subsection. The process starts with the error over the global element and maps it to the local element via the Piola transformation

$$\left\|\boldsymbol{w} - \pi_p^w \boldsymbol{w}\right\|_{\boldsymbol{L}^2(\Omega_e)} = \left\|P_1\left\{\widehat{\boldsymbol{w}} - \widehat{\pi}_p^w \widehat{\boldsymbol{w}}\right\}\right\|_{\boldsymbol{L}^2(\Omega_e)}. \tag{3.124}$$

Then, the first estimate is obtained by changing the domain to the reference element, as

$$\left\|\boldsymbol{w} - \pi_p^w \boldsymbol{w}\right\|_{\boldsymbol{L}^2(\Omega_e)} \leq C\, h^{1/2} \left\|\widehat{\boldsymbol{w}} - \widehat{\pi}_p^w \widehat{\boldsymbol{w}}\right\|_{\boldsymbol{L}^2(\widehat{\Omega}_e)}, \tag{3.125}$$

where $h^{1/2}$ arises from the effect of the norm on the volume element (3.113) and the Jacobian of the Piola transformation

$$\left\|\boldsymbol{J}^{-1}\right\|_{L^\infty(\widehat{\Omega}_e)} \leq Ch^{-1}. \tag{3.126}$$

Requiring the approximation space to contain a polynomial space

$$(\mathbb{P}_{p-1})^3 \subseteq \widehat{\mathcal{W}}_p, \tag{3.127}$$

the Bramble-Hilbert lemma can be used on the local element

$$\left\|\boldsymbol{w} - \pi_p^w \boldsymbol{w}\right\|_{\boldsymbol{L}^2(\Omega_e)} \leq C\, h^{1/2} \left\|\widehat{D}^{\tilde{\alpha}} \widehat{\boldsymbol{w}}\right\|_{\boldsymbol{L}^2(\widehat{\Omega}_e)}, \quad \text{with } |\tilde{\alpha}| = p. \tag{3.128}$$

Then, using the inverse Piola transformation, the estimate is rewritten in terms of the initial global function

$$\begin{aligned}\left\|\boldsymbol{w} - \pi_p^w \boldsymbol{w}\right\|_{\boldsymbol{L}^2(\Omega_e)} &\leq C\, h^{1/2} \left\|\widehat{D}^{\tilde{\alpha}} \left[P_1^{-1}\left\{\boldsymbol{w}\right\}\right]\right\|_{\boldsymbol{L}^2(\widehat{\Omega}_e)}, \\ &\leq C\, h^{1/2} \left\|\widehat{D}^{\tilde{\alpha}} \left[\boldsymbol{J}(\boldsymbol{w} \circ \boldsymbol{g})\right]\right\|_{\boldsymbol{L}^2(\widehat{\Omega}_e)}.\end{aligned} \tag{3.129}$$

Next, the general Leibniz rule is applied and the piecewise norm of each term (triangle and generalized Hölder inequalities) is taken to obtain the estimate

$$\left\|\boldsymbol{w} - \pi_p^w \boldsymbol{w}\right\|_{\boldsymbol{L}^2(\Omega_e)} \leq h^{1/2} \sum_{0 \leq |\alpha| \leq p} C_\alpha \left\|\widehat{D}^{\tilde{\alpha}-\alpha} \boldsymbol{J}\right\|_{L^\infty(\widehat{\Omega}_e)} \left\|\widehat{D}^{\alpha}(\boldsymbol{w} \circ \boldsymbol{g})\right\|_{\boldsymbol{L}^2(\widehat{\Omega}_e)}, \tag{3.130}$$

where C_α is a constant for all multi-index α. The first term with the derivative of the Jacobian can be estimated by the upper bound of the derivative geometry mapping (3.33) as

$$\left\|\widehat{D}^{\tilde{\alpha}-\alpha} \boldsymbol{J}\right\|_{L^\infty(\widehat{\Omega}_e)} \leq C \left\|\widehat{D}^{\tilde{\alpha}-\alpha+1} \boldsymbol{g}\right\|_{L^\infty(\widehat{\Omega}_e)} \leq Ch^{\lceil (p-|\alpha|+1)/r \rceil}. \tag{3.131}$$

The second term can be estimated the same way as (3.120), i.e.,

$$\left\|\widehat{D}^{\alpha}(\boldsymbol{w}\circ\boldsymbol{g})\right\|_{\boldsymbol{L}^2(\widehat{\Omega}_e)} \leq Ch^{-3/2+\lceil|\alpha|/r\rceil}\left\|\boldsymbol{w}\right\|_{\boldsymbol{H}^{|\alpha|}(\Omega_e)}. \tag{3.132}$$

Therefore, the finite element error estimation becomes

$$\left\|\boldsymbol{w}-\pi_p^w\boldsymbol{w}\right\|_{\boldsymbol{L}^2(\Omega_e)} \leq \sum_{0\leq|\alpha|\leq p} C_\alpha h^{\lceil(p-|\alpha|+1)/r\rceil+\lceil|\alpha|/r\rceil-1}\left\|\boldsymbol{w}\right\|_{\boldsymbol{H}^{|\alpha|}(\Omega_e)}. \tag{3.133}$$

It can be seen that the achieved convergence rate for $r=1$ (affine refinements) is

$$\left\|\boldsymbol{w}-\pi_p^w\boldsymbol{w}\right\|_{\boldsymbol{L}^2(\Omega_e)} \leq \sum_{0\leq|\alpha|\leq p} C_\alpha h^p\left\|\boldsymbol{w}\right\|_{\boldsymbol{H}^{|\alpha|}(\Omega_e)} \leq Ch^p\left\|\boldsymbol{w}\right\|_{\boldsymbol{H}^{p}(\Omega_e)}, \tag{3.134}$$

where C is a constant, independent of h and $\boldsymbol{w}$. However, the rate of convergence does not reach $\mathcal{O}(h^p)$ for non-affine refinements.

Similar to the case of H^1-conforming estimates, it is possible to recover the convergence for non-affine refinements by some additional requirements on the global element. In Section 3.1.5, it was shown that it is permissible to replace π_p^w by the practically available interpolation operator $\Pi_k^w\pi_p^w$, which maps the local interpolant to the global domain via an interpolated geometry mapping of order $k=rp$. Moreover, requiring the global approximation space $\Pi_k^w\pi_p^w\boldsymbol{w}\in\mathcal{W}_p$ to contain a polynomial space

$$(\mathbb{P}_{p-1})^3\subseteq\mathcal{W}_p \tag{3.135}$$

for arbitrary $P_{1,k}$ mappings, the Bramble-Hilbert lemma can be applied directly on the global element. Thus, the obtained finite element error estimate reads

$$\left\|\boldsymbol{w}-\Pi_k^w\pi_p^w\boldsymbol{w}\right\|_{\boldsymbol{L}^2(\widetilde{\Omega}_e^k)} \leq Ch^p\left\|\boldsymbol{w}\right\|_{\boldsymbol{H}^{p}(\widetilde{\Omega}_e^k)}, \tag{3.136}$$

where C is a constant, independent of h and $\boldsymbol{w}$.

3.2.3 $\boldsymbol{H}(\mathrm{div})$ Approximation Estimates

The interpolation error of a function $\boldsymbol{d} \in \boldsymbol{H}(\mathrm{div}, \Omega_e)$ reads

$$\left\|\boldsymbol{d} - \pi_p^d \boldsymbol{d}\right\|_{\boldsymbol{L}^2(\Omega_e)} = \left\| P_2 \left\{ \widehat{\boldsymbol{d}} - \widehat{\pi}_p^d \widehat{\boldsymbol{d}} \right\} \right\|_{\boldsymbol{L}^2(\Omega_e)}, \tag{3.137}$$

where P_2 is the appropriate Piola transformation. The first estimate is obtained by changing the norm's domain:

$$\left\|\boldsymbol{d} - \pi_p^d \boldsymbol{d}\right\|_{\boldsymbol{L}^2(\Omega_e)} \leq C\, h^{-1/2} \left\| \widehat{\boldsymbol{d}} - \widehat{\pi}_p^d \widehat{\boldsymbol{d}} \right\|_{\boldsymbol{L}^2(\widehat{\Omega}_e)}, \tag{3.138}$$

where $h^{-1/2}$ arises from the norm's effect on the volume element (3.113) and the multiplicative term of the Piola transformation

$$\left\|\det(\boldsymbol{J})^{-1} \boldsymbol{J}^T\right\|_{L^\infty(\widehat{\Omega}_e)} \leq C h^{-2}. \tag{3.139}$$

If the local approximation space contains a polynomial space

$$\left(\mathbb{P}_{p-1}\right)^3 \subseteq \widehat{\mathcal{D}}_p, \tag{3.140}$$

then the Bramble-Hilbert lemma can be applied to obtain the estimate

$$\left\|\boldsymbol{d} - \pi_p^d \boldsymbol{d}\right\|_{\boldsymbol{L}^2(\Omega_e)} \leq C\, h^{-1/2} \left\| \widehat{D}^{\tilde{\alpha}} \widehat{\boldsymbol{d}} \right\|_{\boldsymbol{L}^2(\widehat{\Omega}_e)}, \quad \text{with } |\widetilde{\alpha}| = p. \tag{3.141}$$

Then, using the inverse Piola transformation, the estimate is rewritten in terms of the global function

$$\begin{aligned} \left\|\boldsymbol{d} - \pi_p^d \boldsymbol{d}\right\|_{\boldsymbol{L}^2(\Omega_e)} &\leq C\, h^{-1/2} \left\| \widehat{D}^{\tilde{\alpha}} \left[P_2^{-1} \{\boldsymbol{d}\} \right] \right\|_{\boldsymbol{L}^2(\widehat{\Omega}_e)}, \\ &\leq C\, h^{-1/2} \left\| \widehat{D}^{\tilde{\alpha}} \left[\det(\boldsymbol{J}) \boldsymbol{J}^{-T} (\boldsymbol{d} \circ \boldsymbol{g}) \right] \right\|_{\boldsymbol{L}^2(\widehat{\Omega}_e)}. \end{aligned} \tag{3.142}$$

This form is further expanded by utilizing the Leibniz rule and applying triangle and Hölder inequalities:

$$\left\|\boldsymbol{d} - \pi_p^d \boldsymbol{d}\right\|_{\boldsymbol{L}^2(\Omega_e)} \leq h^{-1/2} \sum_{0 \leq |\alpha| \leq p} C_\alpha \left\| \widehat{D}^{\tilde{\alpha}-\alpha} \left[\det(\boldsymbol{J}) \boldsymbol{J}^{-T} \right] \right\|_{L^\infty(\widehat{\Omega}_e)} \left\| \widehat{D}^{\alpha} \left[\boldsymbol{d} \circ \boldsymbol{g} \right] \right\|_{\boldsymbol{L}^2(\widehat{\Omega}_e)}, \tag{3.143}$$

where C_α is a constant for all multi-index α. The estimate of the first matrix term,

$$\left\| \widehat{D}^{\tilde{\alpha}-\alpha} \left[\det(\boldsymbol{J}) \boldsymbol{J}^{-T} \right] \right\|_{L^\infty(\widehat{\Omega}_e)} = \left\| \widehat{D}^{\tilde{\alpha}-\alpha} \left[\mathrm{adj}(\boldsymbol{J}) \right] \right\|_{L^\infty(\widehat{\Omega}_e)} \leq C h^{1+\lceil (p-|\alpha|+1)/r \rceil}, \tag{3.144}$$

is based on the application of (3.33) for the derivative adjugate matrix components, which contains only pairwise multiplications of the Jacobian components. Thus,

applying the estimate of function compositions (3.132) with $\boldsymbol{w} = \boldsymbol{d}$ and (3.144) for (3.143), the detailed finite element error estimate becomes

$$\left\|\boldsymbol{d} - \pi_p^d \boldsymbol{d}\right\|_{\boldsymbol{L}^2(\Omega_e)} \leq \sum_{0 \leq |\alpha| \leq p} C_\alpha h^{\lceil (p-|\alpha|+1)/r \rceil + \lceil |\alpha|/r \rceil - 1} \left\|\boldsymbol{d}\right\|_{\boldsymbol{H}^{|\alpha|}(\Omega_e)} . \tag{3.145}$$

It can be seen that the achieved convergence rate for $r = 1$ (affine refinements) is

$$\left\|\boldsymbol{d} - \pi_p^d \boldsymbol{d}\right\|_{\boldsymbol{L}^2(\Omega_e)} \leq \sum_{0 \leq |\alpha| \leq p} C_\alpha h^p \left\|\boldsymbol{d}\right\|_{\boldsymbol{H}^{|\alpha|}(\Omega_e)} \leq C h^p \left\|\boldsymbol{d}\right\|_{\boldsymbol{H}^p(\Omega_e)} , \tag{3.146}$$

where C is a constant, independent of h and $\boldsymbol{d}$. Furthermore, the rate of convergence does not reach $\mathcal{O}(h^p)$ for non-affine refinements. This result is identical to the case of the $\boldsymbol{H}(\mathrm{curl})$-conforming estimate (3.134). Hence, the same (3.136) procedure can be used to improve the convergence rate for non-affine refinements ($r > 1$). In Section 3.1.6, it was shown that it is permissible to replace π_p^d by the practically available interpolation operator $\Pi_k^d \pi_p^d$, which maps the local interpolant to the global domain via an interpolated geometry mapping of order $k = rp$. Then, requiring the global approximation space $\Pi_k^d \pi_p^d \boldsymbol{d} \in \mathcal{D}_p$ to contain a complete order polynomial space

$$(\mathbb{P}_{p-1})^3 \subseteq \mathcal{D}_p \tag{3.147}$$

for arbitrary geometry mappings $P_{2,k}$, the Bramble-Hilbert lemma can be applied directly on the global element

$$\left\|\boldsymbol{d} - \Pi_k^d \pi_p^d \boldsymbol{d}\right\|_{\boldsymbol{L}^2(\widetilde{\Omega}_e^k)} \leq C h^p \left\|\boldsymbol{d}\right\|_{\boldsymbol{H}^p(\widetilde{\Omega}_e^k)} . \tag{3.148}$$

Here, C is a constant, independent of h and $\boldsymbol{d}$.

3.3 Summary of Convergence Requirements

This section provides a summary for the approximation spaces and the geometry mappings, which produce the $\mathcal{O}(h^p)$ rate of convergence. To do so, it is necessary to impose some conditions.

- In the geometry representation:
  - The exact mapping and its interpolant are bijective on each element.
  - The mesh parameter h is defined as (3.17).
  - The polynomial order r of the refinement is defined as (3.12).
  - The order of geometry interpolation k is defined as (3.1).
- In the field representation:
  - The finite element moments are chosen according to (2.55), (2.57), and (2.59), being associated with nodes, edges, faces, and volume.
  - The local approximation spaces on the reference element $\widehat{\Omega}_e$ include the subspaces of $\widehat{\nabla}\widehat{\mathcal{V}}_p \subset \widehat{\mathcal{W}}_p$ and $\widehat{\nabla} \times \widehat{\mathcal{W}}_p \subset \widehat{\mathcal{D}}_p$.
  - The approximation spaces $\mathcal{V}_p$, $\mathcal{W}_p$, and $\mathcal{D}_p$ of each global element Ω_e are such that the approximated fields of the computational domain Ω are $H^1(\Omega)$-, $\boldsymbol{H}(\mathrm{curl}, \Omega)$-, and $\boldsymbol{H}(\mathrm{div}, \Omega)$-conforming, respectively.

As a general consideration, the usage of the exact geometry and affine refinement ($r = 1$) is preferable. Thus, there is no geometry error, and the necessary approximation space is much smaller. However, such geometry mappings are unavailable for some applications. In those cases, interpolated geometry mappings can be used. The convergence rate of the geometry error and the convergence rate of the field error are matched by choosing an appropriate interpolation order. Hence, the overall convergence rate is not affected. Moreover, if the exact geometry is a polynomial of order $\tilde{k}$, then the interpolation order can be limited by $k \leq \tilde{k}$. Due to the generality of the proofs, the obtained necessary conditions apply for hexahedral, tetrahedral, and triangular-based prism elements and are identical to well-known results of the conforming tetrahedral isoparametric elements [Cia02, M$^+$03, SF73]. However, the application and the interpretation differ due to the different interpretations of the mesh parameter and mesh refinement.

The disadvantage of the non-affine refinement ($r > 1$) is that the local finite element approximation spaces need to be dependent on the geometry mappings. Hence, much larger spaces are required than for affine refinements. However, it can be advantageous for the *hp*-adaptive techniques that the refined mesh is not restricted by the previous, coarser mesh. Another advantage is the capability to improve the

geometry representation and mesh quality in each refinement step without any precautions. This is not possible for affine refinements when a mesh with interpolated geometry mappings is refined because the geometry error of the initial mesh carries over to the refined elements, too. The issue of the affine and non-affine refinements is not relevant for tetrahedral meshes. The reason is that the straight-sided geometries are described by affine mappings for tetrahedral elements, and the midpoints of the edges determine all the refined elements. The consequence is that the higher-order terms of a polynomial geometry mapping have a higher rate of convergence, and the natural way of refinement is affine refinement. However, an affine geometry mapping does not necessarily describe a straight-sided hexahedron or prismatic element. In those cases, the refinement involves not just the midpoints of the edges but also the midpoint of the faces. Hence, the elements of a refined mesh do not necessarily converge to elements with affine geometry mappings. This can be described as a non-affine refinement.

Note that the requirements obtained for the approximation spaces are determined for the elementwise error. However, for the same global error, the conformity of the elements is also required. Therefore, depending on the element type, the geometry mapping, and the refinement, a larger space might be necessary. The incremental monomials of this larger space can be used to maintain the conformity between the elements but do not contribute to the elementwise convergence rate.

3.3.1 H^1 Convergence Requirements

Let us start with the necessary requirements for approximating a function $v \in H^1(\Omega_e)$ with the asymptotic convergence rate

$$\left\| v - \pi_p^v v \right\|_{L^2(\Omega_e)} = \mathcal{O}(h^p). \tag{3.149}$$

Here, the required order of geometry interpolation is $k = r(p-1)$ due to (3.64). For affine refinements ($r = 1$), the required geometry order becomes significantly less, and the appropriate approximation space (3.114) of the reference element is

$$\mathbb{P}_{p-1} \subseteq \widehat{\mathcal{V}}_p. \tag{3.150}$$

However, for non-affine refinements ($r > 1$), the situation is more complicated. In such cases, the global approximation space $\Pi_k^v \pi_p^v v \in \mathcal{V}_p$ is required to contain the same polynomial space (3.122), i.e.,

$$\mathbb{P}_{p-1} \subseteq \mathcal{V}_p. \tag{3.151}$$

Because the global approximation of the fields is equivalent to the local approximation with the corresponding Piola transformation

$$\Pi_k^v \pi_p^v v = P_{0,k} \left\{ \widehat{\pi}_p^v \widehat{v} \right\} = (\widehat{\pi}_p^v \widehat{v}) \circ \boldsymbol{g}_k^{-1} \in \mathcal{V}_p, \tag{3.152}$$

the necessary condition in terms of the reference element reads

$$\mathbb{P}_{p-1} \circ \boldsymbol{g}_k \subseteq \widehat{\mathcal{V}}_p, \quad \forall \boldsymbol{g}_k. \tag{3.153}$$

Next, the necessary requirements are summarized for the $H^1(\Omega_e)$-norm convergence

$$\left\| v - \pi_p^v v \right\|_{H^1(\Omega_e)} = \mathcal{O}(h^p). \tag{3.154}$$

Hence, in addition to the previous case, the convergence

$$\left\| \nabla v - \nabla \pi_p^v v \right\|_{\boldsymbol{L}^2(\Omega_e)} = \mathcal{O}(h^p) \tag{3.155}$$

of the gradient functions $\nabla v \in \boldsymbol{H}(\mathrm{curl}, \Omega)$ is also required. Using the commutation of the interpolation operators

$$\left\| \nabla v - \nabla \pi_p^v v \right\|_{\boldsymbol{L}^2(\Omega_e)} = \left\| \nabla v - \pi_p^w \nabla v \right\|_{\boldsymbol{L}^2(\Omega_e)} = \mathcal{O}(h^p), \tag{3.156}$$

the requirement of the $\nabla v \in \boldsymbol{H}(\mathrm{curl}, \Omega_e)$ functions can be used. Thus, the required order of geometry interpolation is $k = rp$ due to (3.74). In the case of affine refinements ($r = 1$), the local approximation space of the gradient functions $\widehat{\pi}_p^w \widehat{\nabla} \widehat{v} = \widehat{\nabla} \widehat{\pi}_p^v \widehat{v} \in \widehat{\nabla} \widehat{\mathcal{V}}_p$ is required to contain the polynomial space (3.127) of $\nabla \mathbb{P}_p \subseteq \widehat{\nabla} \widehat{\mathcal{V}}_p$. Hence, the approximation space must contain the complete order polynomial space

$$\mathbb{P}_p \subseteq \widehat{\mathcal{V}}_p. \tag{3.157}$$

For higher-order refinements ($r > 1$), the same condition (3.135) applies for the approximation space of the global element

$$\Pi_k^w \pi_p^w \nabla v = P_{1,k} \left\{ \widehat{\nabla} \widehat{\pi}_p^v \widehat{v} \right\} = \nabla \left(\left(\widehat{\pi}_p^v \widehat{v} \right) \circ \boldsymbol{g}_k^{-1} \right) \in \nabla \mathcal{V}_p, \tag{3.158a}$$

$$\nabla \mathbb{P}_p \subseteq \nabla \mathcal{V}_p. \tag{3.158b}$$

On the reference element, this requirement reads

$$\mathbb{P}_p \circ \boldsymbol{g}_k \subseteq \widehat{\mathcal{V}}_p, \quad \forall \boldsymbol{g}_k, \tag{3.159}$$

which is dependent on the interpolated geometry mapping. Table 3.1 contains these results in a structured way. Similar results have also been presented in the literature [ABF02, ABB15].

Table 3.1: Necessary conditions for the elementwise $\mathcal{O}(h^p)$ convergence rates of H^1-conforming functions.

	$\left\|v - \pi_p^v v\right\|_{L^2(\Omega_e)} = \mathcal{O}(h^p)$	$\left\|\nabla\left(v - \pi_p^v v\right)\right\|_{\boldsymbol{L}^2(\Omega_e)} = \mathcal{O}(h^p)$
$r = 1$	$\mathbb{P}_{p-1} \subseteq \widehat{\mathcal{V}}_p$	$\mathbb{P}_p \subseteq \widehat{\mathcal{V}}_p$
$r > 1$	$\mathbb{P}_{p-1} \circ \boldsymbol{g}_k \subseteq \widehat{\mathcal{V}}_p$	$\mathbb{P}_p \circ \boldsymbol{g}_k \subseteq \widehat{\mathcal{V}}_p$
	$k = r(p-1)$	$k = rp$

3.3.2 $\boldsymbol{H}(\mathrm{curl})$ Convergence Requirements

First, the necessary conditions are summarized for the convergence rate of

$$\left\|\boldsymbol{w} - \pi_p^w \boldsymbol{w}\right\|_{\boldsymbol{L}^2(\Omega_e)} = \mathcal{O}(h^p). \tag{3.160}$$

It is permissible to use interpolated geometry mappings of order $k = rp$ without affecting the convergence of the $\boldsymbol{w} \in \boldsymbol{H}(\mathrm{curl}, \Omega_e)$ functions (see (3.74)). Moreover, the approximation space on the reference element (3.127) is required to be

$$(\mathbb{P}_{p-1})^3 \subseteq \widehat{\mathcal{W}}_p. \tag{3.161}$$

For non-affine refinements ($r > 1$), however, the inclusion of the same polynomial space is required (3.135) on the global element, i.e.,

$$(\mathbb{P}_{p-1})^3 \subseteq \mathcal{W}_p. \tag{3.162}$$

Applying the relationship between the global and local interpolation

$$\Pi_k^w \pi_p^w \boldsymbol{w} = P_{1,k}\left\{\widehat{\pi}_p^w \widehat{\boldsymbol{w}}\right\} = \left(\boldsymbol{J}_k^{-1} \widehat{\pi}_p^w \widehat{\boldsymbol{w}}\right) \circ \boldsymbol{g}_k^{-1} \in \mathcal{W}_p, \tag{3.163}$$

the necessary approximation space of the reference element is determined as

$$\boldsymbol{J}_k (\mathbb{P}_{p-1})^3 \circ \boldsymbol{g}_k \subseteq \widehat{\mathcal{W}}_p, \quad \forall \boldsymbol{g}_k. \tag{3.164}$$

Next, the necessary conditions are summarized for the convergence rate

$$\left\|\boldsymbol{w} - \pi_p^w \boldsymbol{w}\right\|_{\boldsymbol{H}(\mathrm{curl},\Omega_e)} = \mathcal{O}(h^p). \tag{3.165}$$

Hence, in addition to the previous case, it is also required that

$$\left\|\nabla \times \boldsymbol{w} - \nabla \times \pi_p^w \boldsymbol{w}\right\|_{\boldsymbol{L}^2(\Omega_e)} = \mathcal{O}(h^p). \tag{3.166}$$

Here, the commutation of the interpolation operators is used to express the curl of an interpolant as an interpolant of a function $\nabla \times \boldsymbol{w} \in \boldsymbol{H}(\mathrm{div}, \Omega_e)$, i.e.,

$$\widehat{\pi}_p^d \widehat{\nabla} \times \widehat{\boldsymbol{w}} = \widehat{\nabla} \times \widehat{\pi}_p^w \widehat{\boldsymbol{w}} \in \widehat{\nabla} \times \widehat{\mathcal{W}}_p \subset \widehat{\mathcal{D}}_p. \tag{3.167}$$

Hence, for affine refinements, the necessary condition of the convergence

$$\widehat{\nabla} \times (\mathbb{P}_p)^3 \subseteq \widehat{\nabla} \times \widehat{\mathcal{W}}_p \tag{3.168}$$

is inherited from the criteria of $\boldsymbol{H}(\mathrm{div}, \Omega_e)$-conforming functions (3.140). In the case of higher-order refinements ($r > 1$), the same condition applies over the global element

$$\nabla \times (\mathbb{P}_p)^3 \subseteq \nabla \times \mathcal{W}_p, \tag{3.169}$$

which is also inherited from the criteria of $\boldsymbol{H}(\mathrm{div}, \Omega_e)$-conforming functions (3.147). Using the appropriate Piola transformation

$$\begin{aligned} \Pi_k^w \pi_p^d \nabla \times \boldsymbol{w} &= P_{2,k} \left\{ \widehat{\nabla} \times \widehat{\pi}_p^w \widehat{\boldsymbol{w}} \right\}, \\ &= \left[\det(\boldsymbol{J}_k)^{-1} \boldsymbol{J}_k^T \widehat{\nabla} \times \widehat{\pi}_p^w \widehat{\boldsymbol{w}} \right] \circ \boldsymbol{g}^{-1} \in \nabla \times \mathcal{W}_p, \end{aligned} \tag{3.170}$$

the necessary condition for the reference element is obtained as

$$\det(\boldsymbol{J}_k) \boldsymbol{J}_k^{-T} \left[\nabla \times (\mathbb{P}_p)^3 \right] \circ \boldsymbol{g}_k \subseteq \widehat{\nabla} \times \widehat{\mathcal{W}}_p, \quad \forall \boldsymbol{g}_k. \tag{3.171}$$

All of these results are summarized in Table 3.2.

Table 3.2: Necessary conditions for the elementwise $\mathcal{O}(h^p)$ convergence rates of $\boldsymbol{H}(\mathrm{curl})$-conforming functions.

	$\left\| \boldsymbol{w} - \pi_p^w \boldsymbol{w} \right\|_{\boldsymbol{L}^2(\Omega_e)} = \mathcal{O}(h^p)$	$\left\| \nabla \times \left(\boldsymbol{w} - \pi_p^w \boldsymbol{w} \right) \right\|_{\boldsymbol{L}^2(\Omega_e)} = \mathcal{O}(h^p)$
$r = 1$	$(\mathbb{P}_{p-1})^3 \subseteq \widehat{\mathcal{W}}_p$	$\widehat{\nabla} \times (\mathbb{P}_p)^3 \subseteq \widehat{\nabla} \times \widehat{\mathcal{W}}_p$
$r > 1$	$\boldsymbol{J}_k (\mathbb{P}_{p-1})^3 \circ \boldsymbol{g}_k \subseteq \widehat{\mathcal{W}}_p$	$\det(\boldsymbol{J}_k) \boldsymbol{J}_k^{-T} \left[\nabla \times (\mathbb{P}_p)^3 \right] \circ \boldsymbol{g}_k \subseteq \widehat{\nabla} \times \widehat{\mathcal{W}}_p$
	$k = rp$	$k = rp$

3.3.3 $\boldsymbol{H}(\mathrm{div})$ Convergence Requirements

In case of the $\boldsymbol{H}(\mathrm{div}, \Omega_e)$-conforming functions, only the conditions of the

$$\left\|\boldsymbol{d} - \pi_p^d \boldsymbol{d}\right\|_{\boldsymbol{L}^2(\Omega_e)} = \mathcal{O}(h^p), \tag{3.172}$$

convergence are considered. Here, it is allowed to use interpolated geometry mappings of order $k = rp$ without affecting the rate of convergence (3.84). For affine refinements, it is necessary to include a polynomial space (3.140) in the approximation space of the reference element

$$(\mathbb{P}_{p-1})^3 \subseteq \widehat{\mathcal{D}}_p. \tag{3.173}$$

For higher-order refinements ($r > 1$), the same polynomial space is required for the global approximation space (3.147). Thus, using the appropriate mapping

$$\Pi_k^d \pi_p^d \boldsymbol{d} = P_{2,k}\left\{\widehat{\pi}_p^d \widehat{\boldsymbol{d}}\right\} \in (\mathbb{P}_{p-1})^3 \subseteq \mathcal{D}_p, \tag{3.174}$$

the necessary condition for the approximation space of the reference element is expressed as

$$\det(\boldsymbol{J}_k)\boldsymbol{J}_k^{-T}(\mathbb{P}_{p-1})^3 \circ \boldsymbol{g}_k \subseteq \widehat{\mathcal{D}}_p, \quad \forall \boldsymbol{g}_k. \tag{3.175}$$

Table 3.3 presents all of these results.

Table 3.3: Necessary conditions for the elementwise $\mathcal{O}(h^p)$ convergence rate of $\boldsymbol{H}(\mathrm{div})$-conforming functions.

	$\left\|\boldsymbol{d} - \pi_p^d \boldsymbol{d}\right\|_{\boldsymbol{L}^2(\Omega_e)} = \mathcal{O}(h^p)$
$r = 1$	$(\mathbb{P}_{p-1})^3 \subseteq \widehat{\mathcal{D}}_p$
$r > 1$	$\det(\boldsymbol{J}_k)\boldsymbol{J}_k^{-T}(\mathbb{P}_{p-1})^3 \circ \boldsymbol{g}_k \subseteq \widehat{\mathcal{D}}_p$
	$k = rp$

Chapter 4

Hierarchical Universal Matrices

The most used elements for three-dimensional finite element codes are tetrahedra, hexahedra, triangular prisms, and pyramids. Other polyhedral shapes [RW14, LKI17, CXX+22] are less important because they can be easily decomposed to the previously listed elements. In each case, the approximation error is controllable by changing the overall mesh parameter h by applying mesh refinement and/or increasing the polynomial degree of the basis functions p. When basis functions of higher order are applied to curvilinear geometries, the geometries of the elements must also be made curvilinear. Otherwise, the error of the geometry representation would dominate the local finite element approximation error, and the convergence rate would decrease to the case of $p = 1$. [SF73, Chapter 3.3] and [Cia02, Chapter 4.3] indicate that isoparametric modeling suffices to prevent this pitfall from happening in the H^1-conforming case. The finite element representation of the boundary and the element geometry mapping must be made curvilinear, and the same order interpolation of the geometry mapping is sufficient for the asymptotic rate of convergence as the order of finite element basis functions. Similar results were obtained from the analysis in Chapter 3 for H^1, $\boldsymbol{H}(\mathrm{curl})$, and $\boldsymbol{H}(\mathrm{div})$ elements, where I provide an alternative definition for the mesh parameter and focus on the effect of mesh refinement instead of the quality of the geometry mapping.

To demonstrate the issue with the geometry representation, let us consider the boundary value problem

$$\nabla \times \boldsymbol{\mu}_r^{-1} \nabla \times \boldsymbol{E} - k_0^2 \boldsymbol{\varepsilon}_r \boldsymbol{E} = \boldsymbol{0}, \qquad \text{in } \Omega, \tag{4.1a}$$

$$\boldsymbol{n} \times (\boldsymbol{E} \times \boldsymbol{n}) = \boldsymbol{0}, \qquad \text{in } \Gamma. \tag{4.1b}$$

These material properties are position-dependent tensors in the general, inhomogeneous, and anisotropic case. Following Section 2.4, the weak formulation of the

problem with a test functions $\boldsymbol{w}$ becomes

$$\int_{\Omega} \nabla \times \boldsymbol{w} \cdot \boldsymbol{\mu}_r^{-1} \nabla \times \boldsymbol{E} \, \mathrm{d}\boldsymbol{r} - k_0^2 \int_{\Omega} \boldsymbol{w} \cdot \boldsymbol{\varepsilon}_r \boldsymbol{E} \, \mathrm{d}\boldsymbol{r} = 0, \tag{4.2}$$

and the corresponding finite element model is a generalized eigenvalue problem

$$\left(\boldsymbol{S} - k_0^2 \boldsymbol{T}\right) \boldsymbol{x} = \boldsymbol{0} \tag{4.3}$$

with the eigenvalue k_0^2 and the associated eigenvector $\boldsymbol{x}$. The so-called stiffness matrix $\boldsymbol{S}$ and the mass matrix $\boldsymbol{T}$ are calculated as

$$[\boldsymbol{S}]_{ab} = \int_{\Omega} \nabla \times \boldsymbol{w}_a \cdot \boldsymbol{\mu}_r^{-1} \nabla \times \boldsymbol{w}_b \mathrm{d}\boldsymbol{r}, \tag{4.4a}$$

$$[\boldsymbol{T}]_{ab} = \int_{\Omega} \boldsymbol{w}_a \cdot \boldsymbol{\varepsilon}_r \boldsymbol{w}_b \mathrm{d}\boldsymbol{r}. \tag{4.4b}$$

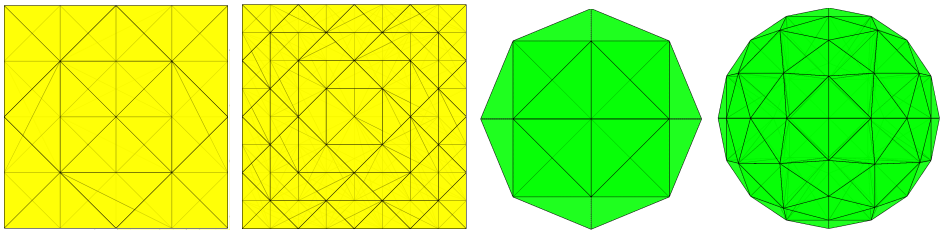


Figure 4.1: Straight-sided tetrahedral meshes for a cube and a sphere.

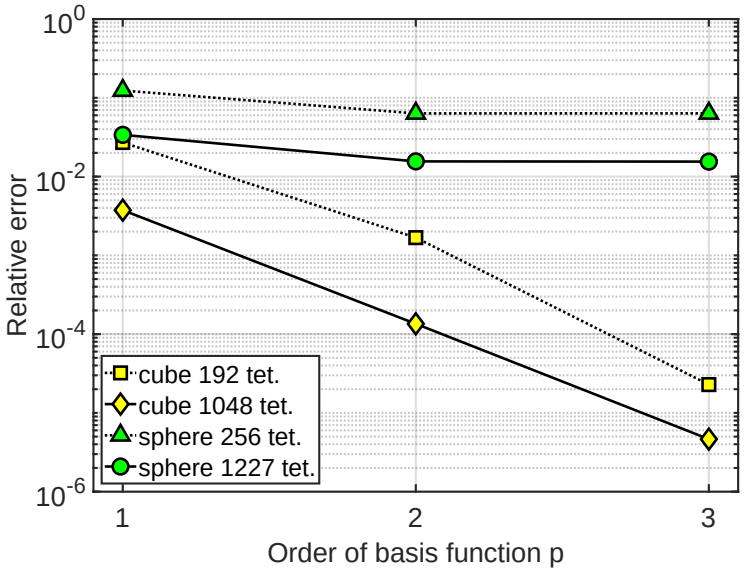


Figure 4.2: Relative error of the smallest nonzero eigenvalues.

In the following example, an empty cubic and a spherical cavity resonator are considered with perfect electrical conducting boundaries. Uniform, straight-sided tetrahedral discretization is applied with the basis functions of order p from Table 2.2 and Table 2.1. The smallest nonzero eigenvalue is calculated for both cavities with different numbers of elements and different orders ($p = 1, 2, 3$) of basis functions. The relative error is shown in Figure 4.2 in terms of basis order p for the meshes of Figure 4.1. In the case of the cube, the affine geometry mapping perfectly represents the flat boundary of the computational domain. With the coarser mesh, one can see a deviation from the algebraic rate due to the element-wise highly changing field in the pre-asymptotic region. Nevertheless, the asymptotic convergence rate is algebraic on the finer mesh because no geometrical error is committed, and the fields are sufficiently smooth on each element. In the case of spheres, however, the geometry representation is of only first order since straight-sided elements describe the curvilinear boundaries. Therefore, the higher-order finite element basis functions do not improve the convergence.

4.1 Universal Matrices

Numerical integration techniques are widely used for the calculation of finite element matrices. These methods can be applied to various shapes, basis functions, and material properties. However, higher-order elements yield significantly increased calculation costs. To remedy this, universal matrices (UMs) can be constructed for the case of straight-sided elements [Sil78]. The advantage is that the element matrices are obtained at a low computational cost as weighted sums of universal matrices [Sil78]. Since the number of universal matrices is independent of the basis order, this approach is more efficient than numerical integration.

Let us denote a component of an element stiffness and mass matrix by $[\boldsymbol{S}_e]_{ab}$ and $[\boldsymbol{T}_e]_{ab}$, respectively. The indices a and b correspond to the basis functions $\boldsymbol{w}_a$ and $\boldsymbol{w}_b$, which have support over the element Ω_e. Using the invariance of the finite element moments, it is permissible to perform the differential operations on the reference domain and map each function to the global domain by the appropriate mapping. Hence, the integrals of arbitrary elements can be mapped to the reference element as

$$
\begin{aligned}
[\boldsymbol{T}_e]_{ab} &= \int_{\Omega_e} \boldsymbol{w}_a \cdot \boldsymbol{\varepsilon}_r \boldsymbol{w}_b \mathrm{d}\boldsymbol{r} = \int_{\widehat{\Omega}_e} \left(\boldsymbol{J}^{-1}\widehat{\boldsymbol{w}}_a\right) \cdot (\boldsymbol{\varepsilon}_r \circ \boldsymbol{g}) \left(\boldsymbol{J}^{-1}\widehat{\boldsymbol{w}}_b\right) \det(\boldsymbol{J}) \mathrm{d}\widehat{\boldsymbol{r}}, \\
&= \int_{\widehat{\Omega}_e} \widehat{\boldsymbol{w}}_a \cdot \widehat{\boldsymbol{\Lambda}}^1 \widehat{\boldsymbol{w}}_b \mathrm{d}\widehat{\boldsymbol{r}}, \qquad (4.5a) \\
[\boldsymbol{S}_e]_{ab} &= \int_{\Omega_e} \nabla \times \boldsymbol{w}_a \cdot \boldsymbol{\mu}_r^{-1} \nabla \times \boldsymbol{w}_b \mathrm{d}\boldsymbol{r},
\end{aligned}
$$

$$[\boldsymbol{S}_e]_{ab} = \int_{\widehat{\Omega}_e} \left(\det(\boldsymbol{J})^{-1}\boldsymbol{J}^T\widehat{\nabla}\times\widehat{\boldsymbol{w}}_a\right)\cdot\left(\boldsymbol{\mu}_r^{-1}\circ\boldsymbol{g}\right)\left(\det(\boldsymbol{J})^{-1}\boldsymbol{J}^T\widehat{\nabla}\times\widehat{\boldsymbol{w}}_b\right)\det(\boldsymbol{J})\mathrm{d}\widehat{\boldsymbol{r}},$$
$$= \int_{\widehat{\Omega}_e} \widehat{\nabla}\times\widehat{\boldsymbol{w}}_a\cdot\widehat{\boldsymbol{\Lambda}}^2\widehat{\nabla}\times\widehat{\boldsymbol{w}}_b\mathrm{d}\widehat{\boldsymbol{r}}, \tag{4.5b}$$

where the transformed local metric functions $\widehat{\boldsymbol{\Lambda}}^{(\cdot)}$ consist of the material properties and the Jacobian

$$\widehat{\boldsymbol{\Lambda}}^1 = \det(\boldsymbol{J})\,\boldsymbol{J}^{-T}\left(\boldsymbol{\varepsilon}_r\circ\boldsymbol{g}\right)\boldsymbol{J}^{-1}, \tag{4.6a}$$
$$\widehat{\boldsymbol{\Lambda}}^2 = \det(\boldsymbol{J})^{-1}\,\boldsymbol{J}\left(\boldsymbol{\mu}_r^{-1}\circ\boldsymbol{g}\right)\boldsymbol{J}^T. \tag{4.6b}$$

In the case of straight-sided elements and elementwise constant material properties, the local metric terms simplify to constant matrices. Therefore, the integrals of (4.5a) and (4.5b) can be calculated exactly using polynomial basis functions on the reference domain. Taking advantage of the linearity of the integration, all constant coefficients $\left[\widehat{\boldsymbol{\Lambda}}^{(\cdot)}\right]_{ij}$ with the indices $i,j\in\{1,2,3\}$ can be pulled out of the integral

$$[\boldsymbol{T}_e]_{ab} = \int_{\widehat{\Omega}_e}\sum_{ij}[\widehat{\boldsymbol{w}}_a]_i\left[\widehat{\boldsymbol{\Lambda}}^1\right]_{ij}[\widehat{\boldsymbol{w}}_b]_j\,\mathrm{d}\widehat{\boldsymbol{r}} \overset{\widehat{\boldsymbol{\Lambda}}^1=\text{const.}}{=} \sum_{ij}\left[\widehat{\boldsymbol{\Lambda}}^1\right]_{ij}\left[\boldsymbol{T}^0_{ij}\right]_{ab}, \tag{4.7}$$

$$[\boldsymbol{S}_e]_{ab} = \int_{\widehat{\Omega}_e}\sum_{ij}\left[\widehat{\nabla}\times\widehat{\boldsymbol{w}}_a\right]_i\left[\widehat{\boldsymbol{\Lambda}}^2\right]_{ij}\left[\widehat{\nabla}\times\widehat{\boldsymbol{w}}_b\right]_j\mathrm{d}\widehat{\boldsymbol{r}} \overset{\widehat{\boldsymbol{\Lambda}}^2=\text{const.}}{=} \sum_{ij}\left[\widehat{\boldsymbol{\Lambda}}^2\right]_{ij}\left[\boldsymbol{S}^0_{ij}\right]_{ab}, \tag{4.8}$$

obtaining universal matrices

$$\left[\boldsymbol{T}^0_{ij}\right]_{ab} = \int_{\widehat{\Omega}_e}[\widehat{\boldsymbol{w}}_a]_i\,[\widehat{\boldsymbol{w}}_b]_j\,\mathrm{d}\widehat{\boldsymbol{r}}, \tag{4.9}$$
$$\left[\boldsymbol{S}^0_{ij}\right]_{ab} = \int_{\widehat{\Omega}_e}\left[\widehat{\nabla}\times\widehat{\boldsymbol{w}}_a\right]_i\left[\widehat{\nabla}\times\widehat{\boldsymbol{w}}_b\right]_j\mathrm{d}\widehat{\boldsymbol{r}}. \tag{4.10}$$

Since these matrices are independent of the geometry and material, they can be used for all elements in the mesh. Thus, the universal matrices $\boldsymbol{T}^0_{ij}$ and $\boldsymbol{S}^0_{ij}$ are only calculated once. The technique, in this form, does not apply to curvilinear elements since the metric term would become a rational tensor function, and analytical integration would not be possible. In such cases, one may apply numerical integration over the reference domain. However, these methods [IIN09, MMIN12] tend to be computationally expensive because the element matrices need to be instantiated at each quadrature point [Coo02], and both matrix size and the number of quadrature points increase rapidly with the finite element order [Bor13].

A different approach is to employ polynomial interpolation for the metric parts [VW99, BWPL12]. This enables analytic integration in terms of some interpolation basis functions, and the calculation of element matrices can be obtained by UMs as in the straight-sided case. An efficient scheme for H^1-conforming elements is presented in [VW99]. It uses hierarchical Newton-Lagrange interpolation with nonsymmetric locations of the interpolation points. The advantage is the reusable lower-order interpolation of the metric term, which remains unchanged when the order is increased. The corresponding UMs also inherit this hierarchical property. On the other hand, universal matrices for $\boldsymbol{H}(\mathrm{curl})$-conforming basis functions are presented in [BWPL12]. This approach is based on Lagrange interpolation, which is not hierarchical but features symmetric locations of the interpolation points. Here, the interpolation process is faster and better conditioned than using hierarchical interpolation. However, a different set of universal matrices is required for the different order of metric interpolations because it is not hierarchical. The results from the lower-order interpolations cannot be reused.

In the following sections, a different method for obtaining $\boldsymbol{H}(\mathrm{curl})$ UMs is presented which is applicable to curvilinear elements with inhomogeneous material properties. The approximation of the metric terms is proposed using hierarchical polynomials. Hence, the resulting UMs inherit the hierarchical property. The approximation process involves numerical integration rules with symmetric quadrature locations. If the order of the approximation polynomial is higher than a limit specified by the order of basis functions, the corresponding UMs become zero due to the hierarchical and orthogonality property of the metric. It will be shown that the calculation of FE matrices is exact using a finite number of UMs. Moreover, it was shown in Section 3.1.7 that the metric expansion may be terminated at a much earlier point when the finite element convergence is in the asymptotic region. In most cases, there is no need to obtain a metric approximation up to the representation limit because the much lower asymptotic limit can be used.

4.2 Hierarchical Metric Expansion

Let $\mathbb{P}_K$ denote the space of polynomials up to degree K on $\widehat{\Omega}$, and $\{b_k^l\}$ be a basis, where k indicates the degree and l a specific polynomial of degree k. Furthermore, $L(k)$ stands for the number of homogeneous polynomials of degree k. Provided that $\boldsymbol{\Lambda}^{(\cdot)}$ are square-integrable on $\widehat{\Omega}_e$, there exists a unique best polynomial approximation $\boldsymbol{\Lambda}_K^{(\cdot)}$ in the L^2-norm [Dav75, Section 7.5],

$$\widehat{\boldsymbol{\Lambda}}_K^{(\cdot)}(\widehat{\boldsymbol{r}}) = \sum_{k=0}^{K} \sum_{l=1}^{L(k)} \widehat{\boldsymbol{\Lambda}}_{kl}^{(\cdot)} b_k^l(\widehat{\boldsymbol{r}}), \qquad \widehat{\boldsymbol{r}} \in \widehat{\Omega}_e, \tag{4.11}$$

with matrix-valued coefficients $\widehat{\mathbf{\Lambda}}_{kl}^{(\cdot)}$. The proposed choice of basis functions $b_k^l(\widehat{\boldsymbol{r}})$ are hierarchical scalar polynomials of order k in the space of incremental order polynomials $\widetilde{\mathcal{P}}_k$. The space of polynomials up to order K is the sum of the incremental spaces

$$\mathbb{P}_K = \widetilde{\mathcal{P}}_1 \oplus ... \oplus \widetilde{\mathcal{P}}_k \oplus ... \oplus \widetilde{\mathcal{P}}_K. \tag{4.12}$$

The main property of these functions is the pairwise orthonormality on the reference domain:

$$\int_{\widehat{\Omega}_e} b_k^l \, b_{k'}^{l'} \mathrm{d}\widehat{\boldsymbol{r}} = \begin{cases} 1 & \text{for } (k,l) = (k',l'), \\ 0 & \text{else.} \end{cases} \tag{4.13}$$

Such an orthonormal basis can be obtained by a Gram-Schmidt process or simply normalizing the classical orthogonal polynomials for the simplex [DX14, Proposition 2.3.8]. In addition, I have developed a special tetrahedron symmetric set of hierarchical orthonormal polynomials in [A4] that are aesthetically appealing and expected to be more robust regarding the round-off errors. The method of construction and the first few order functions are given in Appendix A. The constant matrix coefficients $\widehat{\mathbf{\Lambda}}_{kl}^{(\cdot)}$ are determined such that the L^2-norm error is minimized, i.e.,

$$\int_{\widehat{\Omega}} \left| \left[\widehat{\mathbf{\Lambda}}^{(\cdot)} - \widehat{\mathbf{\Lambda}}_K^{(\cdot)} \right]_{ij} \right|^2 \mathrm{d}\widehat{\boldsymbol{r}} \stackrel{!}{=} \min, \qquad \forall (i,j). \tag{4.14}$$

Since $\mathbb{P}(\widehat{\Omega})$ is dense in $L^2(\widehat{\Omega})$, the approximation is guaranteed to converge for $K \to \infty$ [Sch69, CQ82]. By the projection theorem [Dav75, Section 12.5], the coefficients $\widehat{\mathbf{\Lambda}}_{kl}^{(\cdot)}$ are obtained from the orthogonality condition

$$\int_{\widehat{\Omega}} \left(\widehat{\mathbf{\Lambda}}^{(\cdot)} - \widehat{\mathbf{\Lambda}}_K^{(\cdot)} \right) b_{k'}^{l'} \mathrm{d}\widehat{\boldsymbol{r}} = 0, \qquad \forall b_{k'}^{l'} \in \mathbb{P}_K. \tag{4.15}$$

Substituting (4.11) for $\widehat{\mathbf{\Lambda}}_K^{(\cdot)}(\widehat{\boldsymbol{r}})$ yields

$$\int_{\widehat{\Omega}_e} \widehat{\mathbf{\Lambda}}^{(\cdot)}(\widehat{\boldsymbol{r}}) b_{k'}^{l'} \mathrm{d}\widehat{\boldsymbol{r}} = \sum_{k=0}^{K} \sum_{l=1}^{L(k)} \widehat{\mathbf{\Lambda}}_{kl}^{(\cdot)} \int_{\widehat{\Omega}_e} b_k^l \, b_{k'}^{l'} \mathrm{d}\widehat{\boldsymbol{r}} = \widehat{\mathbf{\Lambda}}_{k'l'}^{(\cdot)} \int_{\widehat{\Omega}_e} b_{k'}^{l'} \, b_{k'}^{l'} \mathrm{d}\widehat{\boldsymbol{r}} = \widehat{\mathbf{\Lambda}}_{k'l'}^{(\cdot)}. \tag{4.16}$$

Hence, the matrix coefficients are determined by evaluating a scalar product. In practice, this is achieved via numerical integration, which is the sole source of error in this method. Nevertheless, efficient symmetric quadrature schemes can be used, which are at least capable of integrating polynomials of order $2K$ because $(b_{k'}^{l'} \, b_{k'}^{l'}) \in \mathbb{P}_{2K}$.

4.2.1 Physical Interpretation of Metric Expansion

One can also investigate how these techniques effect the material properties. Rearranging (4.6a), the electric permittivity can be expressed in terms of the metric expansion

$$\boldsymbol{\varepsilon}_r \circ \boldsymbol{g} = \det(\boldsymbol{J})^{-1} \sum_{k=0}^{K} \sum_{l=1}^{L(k)} \boldsymbol{J}^T \widehat{\boldsymbol{\Lambda}}^1_{kl} \boldsymbol{J} b^l_k. \tag{4.17}$$

Here, the first scaling term $\det(\boldsymbol{J})^{-1}$ is the inverse of the volume element. Therefore, the remaining term has a volume distributional interpretation. The second, tensorial terms

$$\boldsymbol{J}^T \widehat{\boldsymbol{\Lambda}}^1_{kl} \boldsymbol{J} = \sum_{i=1}^{3} \sum_{j=1}^{3} \left[\widehat{\boldsymbol{\Lambda}}^1_{kl}\right]_{ij} \frac{\partial \boldsymbol{r}}{\partial \widehat{r}_i} \otimes \frac{\partial \boldsymbol{r}}{\partial \widehat{r}_j}, \tag{4.18}$$

are equivalent to the sum of scaled contravariant dyadics. Hence, the proposed metric expansion is equivalent to the contravariant approximation of the electric permittivity in the form

$$\boldsymbol{\varepsilon}_r \circ \boldsymbol{g} = \det\left(\widehat{\nabla} \boldsymbol{r}^T\right)^{-1} \sum_{i=1}^{3} \sum_{j=1}^{3} \frac{\partial \boldsymbol{r}}{\partial \widehat{r}_i} \otimes \frac{\partial \boldsymbol{r}}{\partial \widehat{r}_j} \sum_{k=0}^{K} \sum_{l=1}^{L(k)} \left[\widehat{\boldsymbol{\Lambda}}^1_{kl}\right]_{ij} b^l_k. \tag{4.19}$$

For the magnetic susceptibility $\boldsymbol{\mu}_r^{-1}$, the metric expansion results in

$$\boldsymbol{\mu}_r^{-1} = \det(\boldsymbol{J}) \circ \boldsymbol{g}^{-1} \sum_{k=0}^{K} \sum_{l=1}^{L(k)} \left(\boldsymbol{J}^{-1} \widehat{\boldsymbol{\Lambda}}^2_{kl} \boldsymbol{J}^{-T}\right) \circ \boldsymbol{g}^{-1} \left(b^l_k \circ \boldsymbol{g}^{-1}\right). \tag{4.20}$$

Here, the first scaling term is the mapped volume element, which is equal to $\det(\boldsymbol{J}) \circ \boldsymbol{g}^{-1} = \det(\nabla \widehat{\boldsymbol{r}}^T)^{-1}$ due to the inverse function theorem [HH15]. Moreover, the application of the inverse function theorem on the tensorial term leads to

$$\left(\boldsymbol{J}^{-1} \widehat{\boldsymbol{\Lambda}}^2_{kl} \boldsymbol{J}^{-T}\right) \circ \boldsymbol{g}^{-1} = \sum_{i=1}^{3} \sum_{j=1}^{3} \left[\widehat{\boldsymbol{\Lambda}}^2_{kl}\right]_{ij} \nabla \widehat{r}_i \otimes \nabla \widehat{r}_j, \tag{4.21}$$

which is the sum of covariant dyadics scaled by the metric coefficients. Thus, the metric expansion is equivalent to the covariant approximation of the magnetic susceptibility in the form of

$$\boldsymbol{\mu}_r^{-1} = \det\left(\nabla \widehat{\boldsymbol{r}}^T\right)^{-1} \sum_{i=1}^{3} \sum_{j=1}^{3} \nabla \widehat{r}_i \otimes \nabla \widehat{r}_j \sum_{k=0}^{K} \sum_{l=1}^{L(k)} \left[\widehat{\boldsymbol{\Lambda}}^2_{kl}\right]_{ij} b^l_k \circ \boldsymbol{g}^{-1}. \tag{4.22}$$

4.3 Hierarchical Universal Matrices

By applying the metric expansion to the stiffness matrix, one can obtain a sum of integrals with polynomial basis functions

$$[\boldsymbol{S}_e]_{ab} = \sum_{k=0}^{K} \sum_{l=1}^{L(k)} \int_{\widehat{\Omega}_e} \widehat{\nabla} \times \widehat{\boldsymbol{w}}_a \cdot \widehat{\boldsymbol{\Lambda}}_{kl}^2 b_k^l \widehat{\nabla} \times \widehat{\boldsymbol{w}}_b \mathrm{d}\widehat{\boldsymbol{r}}. \tag{4.23}$$

From this point on, analytic integration can be used because the integrand is a polynomial. Moreover, the geometry and material information is only present in the constant matrix coefficients, which allows the construction of universal matrices

$$[\boldsymbol{S}_{ijl}^k]_{ab} = \int_{\widehat{\Omega}_e} \left[\widehat{\nabla} \times \widehat{\boldsymbol{w}}_a\right]_i b_k^l \left[\widehat{\nabla} \times \widehat{\boldsymbol{w}}_b\right]_j \mathrm{d}\widehat{\boldsymbol{r}}. \tag{4.24}$$

If the polynomial order of the tetrahedral FE basis is denoted by p, the product of the curl of basis functions satisfies

$$\left[\widehat{\nabla} \times \widehat{\boldsymbol{w}}_a\right]_i \left[\widehat{\nabla} \times \widehat{\boldsymbol{w}}_b\right]_j \in \mathbb{P}_{2p-2}. \tag{4.25}$$

Hence, (4.25) can be expressed in terms of a hierarchical basis $\left\{b_{k'}^{l'}\right\}$ and constant scalar coefficients $\alpha_{k'l'}^{abij}$:

$$\left[\widehat{\nabla} \times \widehat{\boldsymbol{w}}_a\right]_i \left[\widehat{\nabla} \times \widehat{\boldsymbol{w}}_b\right]_j = \sum_{k'=0}^{2p-2} \sum_{l'=1}^{L(k')} \alpha_{k'l'}^{abij}\, b_{k'}^{l'}. \tag{4.26}$$

The calculation of the universal matrices then reads

$$[\boldsymbol{S}_{ijl}^k]_{ab} = \sum_{k'=0}^{2p-2} \sum_{l'=1}^{L(k')} \alpha_{k'l'}^{abij} \int_{\widehat{\Omega}_e} b_k^l\, b_{k'}^{l'} \mathrm{d}\widehat{\boldsymbol{r}}, \tag{4.27}$$

which implies that the higher $k > 2p-2$ order matrices $\boldsymbol{S}_{ijl}^k$ become zero due to the orthogonal property. This leads to the following metric representation limit for the stiffness matrix:

$$k \leq K = 2p - 2. \tag{4.28}$$

Therefore, a finite order metric expansion and a corresponding finite number of universal matrices are required for the calculation of the stiffness matrix, namely

$$\boldsymbol{S}_e = \sum_{k=0}^{2p-2} \sum_{l=1}^{L(k)} \sum_{ij} \left[\widehat{\boldsymbol{\Lambda}}_{kl}^2\right]_{ij} \boldsymbol{S}_{ijl}^k, \tag{4.29}$$

since the contribution of the higher-order terms is zero. To calculate the element matrices, scaled matrix additions of $\boldsymbol{S}_{ijl}^{k}$ are required. Another advantage of the method is that the calculation of a metric coefficient is independent of the higher- or lower-order terms due to the orthogonal property (4.13). Hence, the lower-order coefficients remain the same when the expansion order is increased. Moreover, the universal matrices inherit this hierarchical property, and the higher-order matrices can be added as a correction during degree elevation because the lower-order matrix terms remain the same.

One can apply the same approach for the mass matrix where also metric expansion of order K and finite element basis functions of order p are used

$$[\boldsymbol{T}_e]_{ab} = \sum_{k=0}^{K} \sum_{l=1}^{L(k)} \int_{\widehat{\Omega}_e} \widehat{\boldsymbol{w}}_a \cdot \widehat{\boldsymbol{\Lambda}}_{kl}^{1} b_k^l \widehat{\boldsymbol{w}}_b \mathrm{d}\widehat{\boldsymbol{r}}. \tag{4.30}$$

Here, the definition of the material and geometry-independent universal matrices reads

$$[\boldsymbol{T}_{ijl}^{k}]_{ab} = \int_{\widehat{\Omega}_e} [\widehat{\boldsymbol{w}}_a]_i \, b_k^l \, [\widehat{\boldsymbol{w}}_b]_j \, \mathrm{d}\widehat{\boldsymbol{r}}. \tag{4.31}$$

Since the product of two arbitrary tetrahedral basis functions is

$$[\widehat{\boldsymbol{w}}_a]_i \, [\widehat{\boldsymbol{w}}_b]_j \in \mathbb{P}_{2p}, \tag{4.32}$$

it can be expanded in terms of a hierarchical basis $\left\{b_{k'}^{l'}\right\}$ and constant scalar coefficients $\beta_{k'l'}^{abij}$:

$$[\widehat{\boldsymbol{w}}_a]_i \, [\widehat{\boldsymbol{w}}_b]_j = \sum_{k'=0}^{2p} \sum_{l'=1}^{L(k')} \beta_{k'l'}^{abij} \, b_{k'}^{l'}. \tag{4.33}$$

Hence, the universal matrices read

$$[\boldsymbol{T}_{ijl}^{k}]_{ab} = \sum_{k'=0}^{2p} \sum_{l'=1}^{L(k')} \beta_{k'l'}^{abij} \int_{\widehat{\Omega}_e} b_k^l \, b_{k'}^{l'} \mathrm{d}\widehat{\boldsymbol{r}}. \tag{4.34}$$

This representation implies that the terms of $k > 2p$ order are zero due to orthogonality. Thus, the resulting metric representation limit is

$$k \leq K = 2p, \tag{4.35}$$

which also limits the number of required universal matrices for the calculation of the mass matrix

$$\boldsymbol{T}_e = \sum_{k=0}^{2p} \sum_{l=1}^{L(k)} \sum_{ij} \left[\widehat{\boldsymbol{\Lambda}}_{kl}^{1}\right]_{ij} \boldsymbol{T}_{ijl}^{k}. \tag{4.36}$$

Therefore, the mass matrix is calculated by metric coefficient scaled matrix additions of $\boldsymbol{T}^k_{ijl}$ universal mass matrices. These matrices are also hierarchical in the order of the metric expansion k, which can be increased without recalculating the lower-order terms.

4.3.1 Appropriate Order of Metric Expansion

Now it is known that the polynomial expansions of the metric terms and the corresponding universal matrices may be terminated at the metric representation limits

$$K = \begin{cases} 2p & \text{for } \boldsymbol{T}_e, \\ 2p-2 & \text{for } \boldsymbol{S}_e. \end{cases} \tag{4.37}$$

These limits are independent of the geometry representation and the quality of the field representation. Up to this point, the method can be considered as a numerical integration scheme.

Now, it is natural to ask whether it is possible to use a lower-order metric expansion than the representation limit. In order to discuss this issue, it is necessary to consider the effect of mesh refinement and geometry error. Using affine refinements or some equivalently smooth geometry mappings $\boldsymbol{g}$ that show an algebraic (asymptotic) convergence rate

$$|\boldsymbol{g}|_{k,\infty,\widehat{\Omega}_e} \leq Ch^k \tag{4.38}$$

with a constant C independent of k and h, the results of Section 3.1.7 can be directly applied. Since the initial weak formulation of the eigenvalue problem (4.2) is a specific case of the considered general product (3.91), the corresponding general estimate (3.108) can be applied for the metric truncation error

$$\epsilon_K \leq Ch^{K+1}\left(\|\boldsymbol{E}\|_{\boldsymbol{L}^2(\Omega_e)} \|\boldsymbol{\varepsilon}_r\|_{K+1,\infty,\Omega_e} + \|\nabla\times\boldsymbol{E}\|_{\boldsymbol{L}^2(\Omega_e)} \left\|\boldsymbol{\mu}_r^{-1}\right\|_{K+1,\infty,\Omega_e} \right), \tag{4.39}$$

$$\epsilon_K \leq Ch^{K+1} \|\boldsymbol{E}\|_{\boldsymbol{H}(\mathrm{curl},\Omega_e)} \left(\|\boldsymbol{\varepsilon}_r\|_{K+1,\infty,\Omega_e} + \left\|\boldsymbol{\mu}_r^{-1}\right\|_{K+1,\infty,\Omega_e} \right), \tag{4.40}$$

where C is a positive constant independent of K and h while $\boldsymbol{E}$ is the field approximated by the finite element. Thus, if the asymptotic convergence rate of the field error in $\boldsymbol{H}(\mathrm{curl})$-norm is $\mathcal{O}(h^p)$, the sufficient asymptotic limit for the metric expansion order is

$$K = p - 1 \tag{4.41}$$

because the overall convergence rate is not affected. However, if the geometry mapping does not satisfy the requirement (4.38) or the finite element convergence is not in the asymptotic region, then the metric expansion order may need to be increased until the metric representation limits (4.37).

4.3.2 Application to Hierarchical Basis Functions

The previously introduced metric expansion and universal matrix schemes are applicable to various element types and polynomial approximation spaces. So far, polynomial finite element basis functions that could be either hierarchical or interpolatory have been considered. However, one can further increase the efficiency of the method by using hierarchical finite element basis functions and hierarchical approximation spaces, such as $\widetilde{\mathcal{W}}_i$, whose direct sum gives the entire approximation space $\mathcal{W}_p$:

$$\mathcal{W}_p = \widetilde{\mathcal{W}}_1 \oplus ... \oplus \widetilde{\mathcal{W}}_i \oplus ... \oplus ...\widetilde{\mathcal{W}}_p, \qquad 1 < i < p. \tag{4.42}$$

Hierarchical FE bases such as [Web99], [SLC01], and [Ing06] are very attractive for applications like p adaptive methods [NW04] or multilevel preconditioners [ZC02]. The present method ought to work with various hierarchical bases that, at most, belong to $(\mathbb{P}_p)^3$. My implementation employs the basis functions of Ingelström [Ing06], which span the Nédélec space of incomplete order [Néd80] (see Section 2.6). Let us denote the corresponding hierarchical basis functions by $\boldsymbol{w}_m^a \in \widetilde{\mathcal{W}}_m \subset (\mathbb{P}_m)^3$ and $\boldsymbol{w}_n^b \in \widetilde{\mathcal{W}}_n \in (\mathbb{P}_n)^3$, where a and b are indices for the basis functions within the same incremental space and $n, m \leq p$. The application of hierarchical basis functions leads to a fine structure for the stiffness and mass element matrices

$$\boldsymbol{S}_e = \left[\begin{array}{c|c|c} \boldsymbol{S}_{11} & \boldsymbol{S}_{12} & \dots \\ \hline \boldsymbol{S}_{21} & \boldsymbol{S}_{22} & \dots \\ \hline \vdots & \vdots & \ddots \end{array}\right], \qquad \boldsymbol{T}_e = \left[\begin{array}{c|c|c} \boldsymbol{T}_{11} & \boldsymbol{T}_{12} & \dots \\ \hline \boldsymbol{T}_{21} & \boldsymbol{T}_{22} & \dots \\ \hline \vdots & \vdots & \ddots \end{array}\right], \tag{4.43}$$

consisting of submatrices representing the couplings between the hierarchical basis functions $\boldsymbol{w}_m^a$ of order m and $\boldsymbol{w}_n^b$ of order n. The calculations of these submatrices over the local domain read

$$[\boldsymbol{S}_{mn}]_{ab} = \int_{\widehat{\Omega}_e} \widehat{\nabla} \times \widehat{\boldsymbol{w}}_m^a \cdot \widehat{\boldsymbol{\Lambda}}^2 \widehat{\nabla} \times \widehat{\boldsymbol{w}}_n^b \mathrm{d}\widehat{\boldsymbol{r}}, \tag{4.44a}$$

$$[\boldsymbol{T}_{mn}]_{ab} = \int_{\widehat{\Omega}_e} \widehat{\boldsymbol{w}}_m^a \cdot \widehat{\boldsymbol{\Lambda}}^1 \widehat{\boldsymbol{w}}_n^b \mathrm{d}\widehat{\boldsymbol{r}}. \tag{4.44b}$$

The corresponding universal matrices

$$[\boldsymbol{S}_{ijl}^{mnk}]_{ab} = \int_{\widehat{\Omega}_e} \left[\widehat{\nabla} \times \widehat{\boldsymbol{w}}_m^a\right]_i b_k^l \left[\widehat{\nabla} \times \widehat{\boldsymbol{w}}_n^b\right]_j \mathrm{d}\widehat{\boldsymbol{r}}, \tag{4.45a}$$

$$[\boldsymbol{T}_{ijl}^{mnk}]_{ab} = \int_{\widehat{\Omega}_e} [\widehat{\boldsymbol{w}}_m^a]_i \, b_k^l \left[\widehat{\boldsymbol{w}}_n^b\right]_j \mathrm{d}\widehat{\boldsymbol{r}}, \tag{4.45b}$$

are not only hierarchical in the order of metric expansion but also in the finite element basis functions order m and n. The main advantage of hierarchical FE basis functions is a hierarchical representation limit that separately applies to each submatrix:

$$k \leq K = m + n - 2 \qquad \text{for } \boldsymbol{S}_{mn}, \tag{4.46a}$$

$$k \leq K = m + n \qquad \text{for } \boldsymbol{T}_{mn}. \tag{4.46b}$$

The higher-order universal matrices are zero and do not contribute to the FE solution. Hence, the maximum number of scaled matrix additions is much lower for the lower-order block matrices;

$$\boldsymbol{S}_{mn} = \sum_{k=0}^{m+n-2} \sum_{l=1}^{L(k)} \sum_{ij} \left[\widehat{\boldsymbol{\Lambda}}^2_{kl}\right]_{ij} \boldsymbol{S}^{mnk}_{ijl}, \tag{4.47a}$$

$$\boldsymbol{T}_{mn} = \sum_{k=0}^{m+n} \sum_{l=1}^{L(k)} \sum_{ij} \left[\widehat{\boldsymbol{\Lambda}}^1_{kl}\right]_{ij} \boldsymbol{T}^{mnk}_{ijl}. \tag{4.47b}$$

Another advantage of the proposed method is the higher efficiency in the case of increasing either the order of finite element basis functions p or the metric expansion K. In those cases, only the contributions of the higher, incremental order universal matrices are required. The lower-order matrices and metric coefficients remain the same. Note that the asymptotic limit remains the same as (4.41) for all submatrices. Therefore, the required number of universal matrices for the lower-order submatrices might be dominated by the representation limit (4.46), and the asymptotic limit becomes irrelevant.

4.3.3 Complexity

The following analysis is for integrating the element matrices (4.5) for a hierarchical FE basis of order p, assuming that the metric terms are polynomials of order K. Unlike in [VW99] and [BWPL12], symmetries in the basis functions and matrices are not exploited because they would lower the costs of all methods equally.

Let $N(p)$ denote the number of polynomials up to order p in three dimensions,

$$N(p) = (1 + p)(2 + p)(3 + p)/6, \tag{4.48}$$

and let $G(p)$ be the number of quadrature points for integrating polynomials up to order p over the reference element. Table 4.1 presents data for $G(p)$ based on the quadrature scheme of [ZCL09].

Table 4.1: Number of quadrature points for integrating polynomials of order p over a tetrahedral domain [ZCL09].

p	1	2	3	4	5	6	7	8	9	10	11	12	13	14
$G(p)$	1	4	8	14	14	24	36	46	61	81	109	140	171	236

Proposed Method

The calculation consists of two steps. The first one involves the calculation of a few metric expansion coefficients. The second step is the calculation of the element matrices by scaled matrix additions of universal matrices $\boldsymbol{S}_{ijl}^{kmn}$ and $\boldsymbol{T}_{ijl}^{kmn}$. First, the focus is on the second step. The number of scaled additions $a_{\boldsymbol{S}}^{App.}$ for the stiffness matrix and $a_{\boldsymbol{T}}^{App.}$ for the mass matrix are determined as the number of metric expansion coefficients times the number of entries in the universal matrices (4.47). Provided that the material property tensors are symmetric, the metric tensors are 3×3 matrices with only six independent entries. Hence, the number of scaled matrix additions is $6N(K)$, where the order of metric expansion K is bounded by the representation limit (4.46). Moreover, the number of FE basis functions in each incremental subspace is $\dim \widetilde{\mathcal{W}}_p = (p+2)(3p+1)/2$ [Ing06], which determines the number of entries in each hierarchical UM $\boldsymbol{S}_{ijl}^{kmn}$ and $\boldsymbol{T}_{ijl}^{kmn}$ as $(\dim \widetilde{\mathcal{W}}_m)(\dim \widetilde{\mathcal{W}}_n)$. Thus, the numbers of scaled additions are obtained as

$$a_{\boldsymbol{S}}^{App.} = \sum_{m=1}^{p} \sum_{n=1}^{p} (\dim \widetilde{\mathcal{W}}_m)(\dim \widetilde{\mathcal{W}}_n) 6N(\min\{K, m+n-2\}), \tag{4.49a}$$

$$a_{\boldsymbol{T}}^{App.} = \sum_{m=1}^{p} \sum_{n=1}^{p} (\dim \widetilde{\mathcal{W}}_m)(\dim \widetilde{\mathcal{W}}_n) 6N(\min\{K, m+n\}). \tag{4.49b}$$

Next, the computational cost of obtaining $6N(K)$ different metric expansion coefficients is considered. The expansion is obtained via the application of a numerical integration scheme for a scalar product (4.16), which is capable of integrating polynomials of order $2K$. Hence, the evaluation of the quadrature scheme [ZCL09] requires $G(2K)$ number of scaled additions, where K is bounded by the representation limit (4.46). Thus, the numbers of scaled additions $a_{\widehat{\boldsymbol{\Lambda}}^1}$ for $\widehat{\boldsymbol{\Lambda}}^1$ and $a_{\widehat{\boldsymbol{\Lambda}}^2}$ for $\widehat{\boldsymbol{\Lambda}}^2$ are given as

$$a_{\widehat{\boldsymbol{\Lambda}}^1} = \begin{cases} 6N(K)\, G(2K) & \text{if } K \leq 2p, \\ 6N(2p)\, G(4p) & \text{else,} \end{cases} \tag{4.50a}$$

$$a_{\widehat{\boldsymbol{\Lambda}}^2} = \begin{cases} 6N(K)\, G(2K) & \text{if } K \leq 2p-2, \\ 6N(2p-2)\, G(4p-4) & \text{else.} \end{cases} \tag{4.50b}$$

The obtained relevant values are shown in Table 4.2, Table 4.3, and Table 4.4. It can be seen that $a_{\boldsymbol{T}}^{App.}$ and $a_{\boldsymbol{S}}^{App.}$ become dominant when the FE basis order p is increased and the cost of the metric expansion is negligible. This is because the number of entries in an element matrix is $(\dim \mathcal{W}_p)^2 \sim p^6$ [Néd80], whereas the number of required quadrature points is $G(4p) \sim p^3$ [SH12].

Complexity of Interpolation Based UM Approaches

Next, let us consider another UM based approach [BWPL12] as a comparison. It is based on the interpolation of the geometry and material property dependent terms on uniformly distributed locations over the reference element. This results in a straightforward interpolation scheme; however, unlike with the proposed method, no approximation or hierarchical property is involved. Moreover, the sufficient order of the interpolation is unknown. In any case, a polynomial representation of the metric term allows the construction of UMs in the form of (4.31), and (4.24), and provides the same evaluation scheme as (4.29) and (4.36) with some interpolation values of the metric $\widehat{\boldsymbol{\Lambda}}_{kl}^{(\cdot)}$. Hence, the number of scaled additions is given as the number of interpolation coefficients $6N(K)$ times the number of entries in the universal matrices. Since the number of FE basis functions is $\dim \mathcal{W}_p = p(p+2)(p+3)/2$ [Néd80], the number of entries in each UM is $(\dim \mathcal{W}_p)^2$. Hence, the numbers of scaled additions for the stiffness $a_{\boldsymbol{S}}^{\text{Int.}}$ and mass matrix $a_{\boldsymbol{T}}^{\text{Int.}}$ are given as

$$a_{\boldsymbol{S}}^{\text{Int.}} = (\dim \mathcal{W}_p)^2 \; 6N(K), \tag{4.51a}$$

$$a_{\boldsymbol{T}}^{\text{Int.}} = (\dim \mathcal{W}_p)^2 \; 6N(K). \tag{4.51b}$$

Based on this analysis, the proposed method seems more efficient for higher-order metric terms due to the effect of the metric representation limit.

Complexity of Numerical Integration Based Approaches

Finally, let us consider the case where numerical integration is applied for curvilinear elements [Bor13]. Here, the computation cost is the evaluation of a quadrature scheme capable of integrating polynomials of order $2(p-1)+K$ and $2p+K$ for the stiffness and mass matrices, respectively [BWPL12]. Since the matrices have $(\dim \mathcal{W}_p)^2$ entries, the required numbers of scaled additions $a_{\boldsymbol{S}}^{\text{Num.}}$ and $a_{\boldsymbol{T}}^{\text{Num.}}$ are

$$a_{\boldsymbol{S}}^{\text{Num.}} = (\dim \mathcal{W}_p)^2 \; 6G(2p+K-2), \tag{4.52a}$$

$$a_{\boldsymbol{T}}^{\text{Num.}} = (\dim \mathcal{W}_p)^2 \; 6G(2p+K). \tag{4.52b}$$

Some numerical values for these relevant computation costs are shown in Table 4.2 and Table 4.3. The advantage of the proposed method is clear, especially considering that only this method provides clear directions for the necessary and sufficient order for K.

Table 4.2: Number of scaled additions for the stiffness matrix.

		Order of metric expansion K						
	p	0	1	2	3	4	5	6
$a_{\boldsymbol{S}}^{\text{Num.}}$	1	216	216	864	1728	3024	3024	5184
	2	9600	19200	33600	33600	57600	86400	110400
	3	170100	170100	291600	437400	558900	741150	984150
	4	1016064	1524096	1947456	2582496	3429216	4614624	5927040
$a_{\boldsymbol{S}}^{\text{Int.}}$	1	216	864	2160	4320	7560	12096	18144
	2	2400	9600	24000	48000	84000	134400	201600
	3	12150	48600	121500	243000	425250	680400	1020600
	4	42336	169344	423360	846720	1481760	2370816	3556224
$a_{\boldsymbol{S}}^{\text{App.}}$	1	216	216	216	216	216	216	216
	2	2400	8952	16008	16008	16008	16008	16008
	3	12150	47952	113508	193008	249258	249258	249258
	4	42336	168696	415368	796728	1263648	1700994	1956522

Table 4.3: Number of scaled additions for the mass matrix.

		Order of metric expansion K						
	p	0	1	2	3	4	5	6
$a_{\boldsymbol{T}}^{\text{Num.}}$	1	864	1728	3024	3024	5184	7776	9936
	2	33600	33600	57600	86400	110400	146400	194400
	3	291600	437400	558900	741150	984150	1324350	1701000
	4	1947456	2582496	3429216	4614624	5927040	7239456	9991296
$a_{\boldsymbol{T}}^{\text{Int.}}$	1	216	864	2160	4320	7560	12096	18144
	2	2400	9600	24000	48000	84000	134400	201600
	3	12150	48600	121500	243000	425250	680400	1020600
	4	42336	169344	423360	846720	1481760	2370816	3556224
$a_{\boldsymbol{T}}^{\text{App.}}$	1	216	864	2160	2160	2160	2160	2160
	2	2400	9600	24000	45840	63480	63480	63480
	3	12150	48600	121500	240840	404730	571680	676680
	4	42336	169344	423360	844560	1461240	2262096	3133680

Table 4.4: Number of scaled additions for the metric expansion.

	FE order	Order of metric expansion K						
	p	0	1	2	3	4	5	6
$a_{\widehat{\mathbf{\Lambda}}^1}$	1	6	96	840	840	840	840	840
	2	6	96	840	2880	9660	9660	9660
	3	6	96	840	2880	9660	27216	70560
	4	6	96	840	2880	9660	27216	70560
$a_{\widehat{\mathbf{\Lambda}}^2}$	1	6	6	6	6	6	6	6
	2	6	96	840	840	840	840	840
	3	6	96	840	2880	9660	9660	9660
	4	6	96	840	2880	9660	27216	70560

4.3.4 Numerical Examples

Curvilinear Tetrahedron with Inhomogeneous Material Properties

In the first numerical example, the representation limit of the metric expansion is demonstrated using a single element. Moreover, it is verified that the metric expansion does not affect the null space and range space of the $\nabla\times$ operator even for curvilinear geometries and inhomogeneous material properties. The initial shape of the element is taken as a straight-sided tetrahedron with unit length edges that is deformed by a second-order polynomial mapping on a face while leaving the remaining faces planar (see Figure 4.3). The deformation is obtained as the displacement of the edge midpoints on the face to the surface of a unit sphere whose center $\boldsymbol{r}_0$ is located at the opposite node of the face. The material properties are

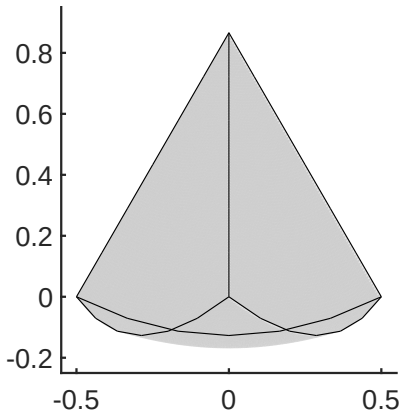


Figure 4.3: Tetrahedron with a second-order curvilinear face.

$$\boldsymbol{\mu}_r = \boldsymbol{I}, \tag{4.53a}$$
$$\boldsymbol{\varepsilon}_r = (11 + 8\tanh(|\boldsymbol{r} - \boldsymbol{r}_0| - 0.2))\,\boldsymbol{I}, \tag{4.53b}$$

where $\boldsymbol{I}$ denotes an identity matrix. These properties correspond to a certain ceramic paste [VW99]. From the physical perspective, the boundary is considered to be a perfect magnetic conductor. From the algebraic perspective, this means that the generalized eigenvalue problem

$$(\boldsymbol{S}_e - \lambda \boldsymbol{T}_e)\,\boldsymbol{x} = \boldsymbol{0}, \tag{4.54}$$

arises from the element matrices of a single element. The list of calculated eigenvalues is provided in Table 4.5, Table 4.6, and Table 4.7 for FE basis function orders $p = 1$, $p = 2$, and $p = 3$, respectively. The zero eigenvalues are at the level of numerical noise, and the dimension of the null space is not affected by the metric expansion order.

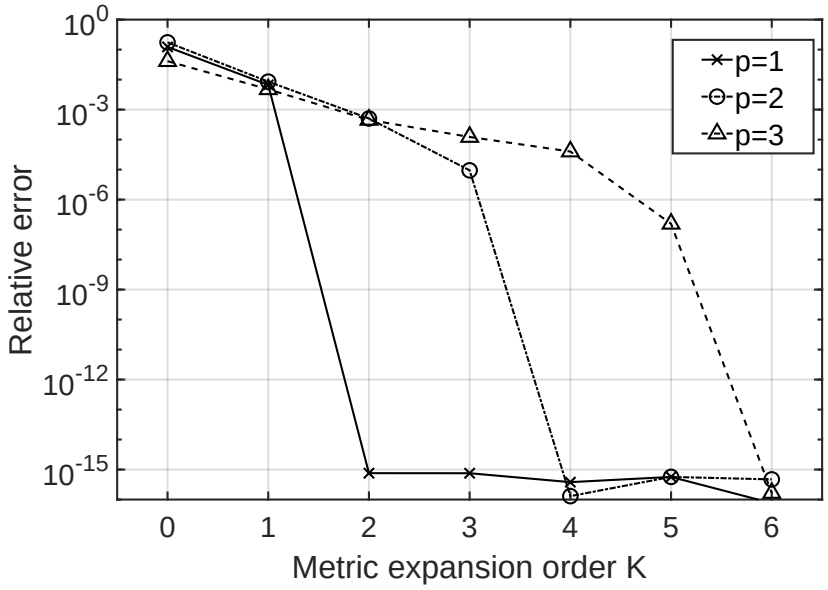


Figure 4.4: Relative error of the smallest nonzero eigenvalue in terms of the metric expansion order K. The solutions with the highest expansion order $K = 6$ are chosen as the reference values.

Figure 4.4 shows the relative error of the smallest nonzero eigenvalue in terms of the metric expansion order K, relative to the case $K = 6$, for the different p-order finite elements. The effect of the metric representation limit ($K = 2p$ for the mass matrix) is clearly visible since the error decreases to the level of numerical noise. Note that the asymptotic limit ($K = p - 1$) is out of question at this point because the true solution error is far from the asymptotic region of convergence.

Table 4.5: Tetrahedron eigenvalues using FE basis functions of order $p = 1$.

Order of metric expansion K						
0	1	2	3	4	5	6
-9.1e-16	-5.0e-16	-2.5e-16	-2.5e-16	-2.5e-16	-2.5e-16	-2.5e-16
-3.2e-16	-4.4e-17	8.6e-17	8.6e-17	8.6e-17	8.6e-17	8.6e-17
5.1e-16	1.3e-16	2.6e-15	2.6e-15	2.6e-15	2.6e-15	2.6e-15
3.54669	4.07217	4.04512	4.04512	4.04512	4.04512	4.04512
3.54669	4.07217	4.05563	4.05563	4.05563	4.05563	4.05563
4.57506	4.87751	4.86753	4.86753	4.86753	4.86753	4.86753

Table 4.6: Tetrahedron eigenvalues using FE basis functions of order $p = 2$.

Order of metric expansion K						
0	1	2	3	4	5	6
-9.2e-16	-1.7e-15	-2.2e-15	-2.8e-15	-8.5e-16	-8.5e-16	-8.5e-16
-6.8e-16	-8.2e-16	-1.5e-15	-9.3e-16	-5.7e-16	-5.7e-16	-5.7e-16
-1.4e-16	-3.4e-16	-6.0e-16	-3.8e-16	-2.5e-16	-2.5e-16	-2.5e-16
2.2e-17	-7.0e-17	-5.1e-16	-3.1e-16	-1.2e-16	-1.2e-16	-1.2e-16
2.7e-16	1.1e-16	-3.4e-16	8.6e-17	4.9e-17	4.9e-17	4.9e-17
6.2e-16	5.4e-16	-1.3e-16	2.3e-16	3.8e-16	3.8e-16	3.8e-16
7.8e-16	9.0e-16	2.2e-16	2.5e-16	8.3e-16	8.3e-16	8.3e-16
1.5e-15	1.3e-15	3.6e-16	7.2e-16	1.2e-15	1.2e-15	1.2e-15
2.0e-15	5.4e-15	6.4e-16	1.7e-15	2.4e-15	2.4e-15	2.4e-15
3.20914	3.92166	3.89031	3.88838	3.88834	3.88834	3.88834
3.20914	3.92166	3.92328	3.92086	3.92082	3.92082	3.92082
3.77856	4.26043	4.25261	4.25511	4.25501	4.25501	4.25501
7.70363	10.0889	9.94368	9.92834	9.92529	9.92529	9.92529
7.70363	10.1512	10.0227	9.99505	9.99258	9.99258	9.99258
9.06713	10.1512	10.3550	10.3244	10.3224	10.3224	10.3224
10.1377	11.2556	11.1198	11.1294	11.1278	11.1278	11.1278
10.1377	11.2556	11.1959	11.2148	11.2118	11.2118	11.2118
13.0817	11.3184	11.2961	11.3127	11.3093	11.3093	11.3093
13.0817	11.3184	11.5029	11.5306	11.5260	11.5260	11.5260
13.7726	15.8516	15.8954	15.8920	15.8856	15.8856	15.8856

Table 4.7: Tetrahedron eigenvalues using FE basis functions of order $p = 3$.

Order of metric expansion k						
0	1	2	3	4	5	6
-1.1e-14	-5.4e-15	-5.1e-15	-4.1e-15	-4.2e-15	-4.5e-15	-2.8e-15
-4.4e-15	-4.5e-15	-3.8e-15	-3.4e-15	-2.8e-15	-3.4e-15	-2.6e-15
-3.6e-15	-4.1e-15	-2.9e-15	-2.2e-15	-2.4e-15	-2.8e-15	-1.4e-15
-2.0e-15	-3.0e-15	-2.5e-15	-1.4e-15	-2.2e-15	-1.8e-15	-1.1e-15
-1.6e-15	-2.5e-15	-1.3e-15	-1.2e-15	-1.4e-15	-1.1e-15	-7.4e-16
-1.5e-15	-1.6e-15	-8.7e-16	-3.9e-16	-8.2e-16	-6.8e-16	-3.7e-16
-7.7e-16	-1.2e-15	-5.4e-16	-3.2e-16	-2.8e-16	-2.8e-16	-1.4e-16
-3.7e-16	-7.2e-16	-2.2e-16	3.3e-17	-2.5e-16	-4.2e-17	-3.7e-17
-2.2e-16	-5.5e-16	-3.0e-17	6.1e-17	-2.3e-16	4.4e-17	5.7e-18
2.6e-18	-3.4e-16	4.2e-17	1.1e-16	-8.5e-17	1.0e-16	5.9e-17
6.9e-17	-6.2e-17	1.6e-16	3.1e-16	8.1e-17	2.3e-16	6.8e-17
4.8e-16	8.9e-17	4.0e-16	7.0e-16	1.8e-16	5.6e-16	1.52e-16
8.7e-16	3.8e-16	4.7e-16	7.5e-16	5.0e-16	6.3e-16	3.16e-16
1.2e-15	4.8e-16	6.6e-16	1.0e-15	6.7e-16	9.5e-16	8.19e-16
1.7e-15	1.5e-15	1.2e-15	1.3e-15	1.5e-15	1.2e-15	1.61e-15
2.1e-15	1.9e-15	1.7e-15	1.9e-15	1.9e-15	2.0e-15	2.31e-15
2.8e-15	2.9e-15	2.0e-15	3.4e-15	2.8e-15	3.0e-15	3.76e-15
3.4e-15	5.0e-15	3.4e-15	3.6e-15	4.2e-15	3.7e-15	5.09e-15
5.2e-15	1.0e-14	3.8e-15	4.3e-15	6.9e-15	4.6e-15	6.87e-15
3.85914	4.00591	4.02690	4.02457	4.02491	4.02507	4.02507
3.85914	4.00591	4.03875	4.03927	4.03924	4.03940	4.03939
4.63812	4.41168	4.43079	4.42736	4.42748	4.42751	4.42751
7.69534	9.37035	9.30923	9.29498	9.28879	9.28847	9.28846
7.69534	10.1040	10.0715	10.0759	10.0753	10.0751	10.0751
7.99637	10.1040	10.1402	10.1490	10.1474	10.1474	10.1474
9.74354	10.2679	10.2203	10.2319	10.2307	10.2308	10.2308
9.74354	10.2679	10.3403	10.3345	10.3332	10.3327	10.3326
9.92180	11.0233	10.8660	10.8725	10.8648	10.8651	10.8651
10.0845	11.0233	11.0747	11.0845	11.0773	11.0775	11.0775
10.0845	11.7428	11.8666	11.8715	11.8693	11.8699	11.8698
15.1656	18.8665	18.5171	18.4602	18.4568	18.4575	18.4593
15.1656	18.8665	18.9623	18.9079	18.9000	18.9007	18.9023
15.2941	19.4149	19.4308	19.4237	19.4100	19.4110	19.4093
17.5062	19.9842	19.8279	19.8369	19.8351	19.8348	19.8349
17.5062	19.9842	20.0182	20.0396	20.0369	20.0371	20.0373
20.9928	21.2803	21.2784	21.2788	21.2689	21.2682	21.2664
20.9928	21.2803	21.5936	21.6249	21.6184	21.6171	21.6152
22.8766	23.3017	23.1930	23.2567	23.2497	23.2490	23.2495
25.1567	23.4416	23.9201	23.9452	23.9406	23.9412	23.9410
25.1567	25.0627	24.7550	24.7641	24.7562	24.7574	24.7577
26.2930	25.1496	25.0506	25.0263	25.0130	25.0150	25.0153
27.0736	25.1496	25.4559	25.3960	25.3818	25.3850	25.3855
27.0736	28.2898	28.0542	28.0351	28.0290	28.0308	28.0305
29.1468	28.2898	28.5514	28.5315	28.5248	28.5267	28.5264
36.0214	40.1133	39.9746	39.9499	39.9218	39.9121	39.9077

Spherical Resonator

The main goal of the second example is to demonstrate the method on a curvilinear domain, which has a smooth but nonpolynomial solution. Hence, the exact solution cannot be reached using a finite number of basis functions, and the expected asymptotic convergence rate for the dominant eigenvalue is $\mathcal{O}(h^{2p})$. This is therefore ideal for the demonstration of convergence in terms of the metric expansion order K and finite element basis order p. For these reasons, the eigenvalue problem of a spherical cavity resonator has been chosen with perfect electrical conducting boundaries and material properties of the vacuum. The model is represented by a mesh of 2048 elements and elementwise third-order interpolated geometry mappings (see Figure 4.5). In Figure 4.6, the relative error of the smallest nonzero eigenvalue

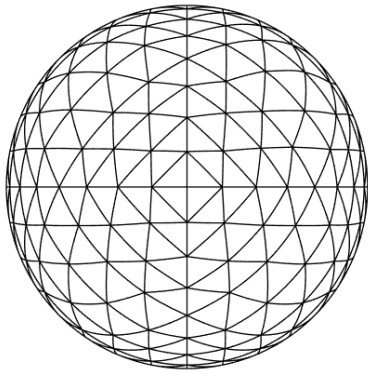


Figure 4.5: Boundary mesh of the unit sphere with third-order elements.

is shown with respect to the analytic solution. One can observe the convergence of the eigenvalues by increasing the metric expansion order K for the different finite element basis order p. In Figure 4.7, also the relative error of the first nonzero eigenvalue is shown with respect to the analytic solution, but now versus the order p of the finite element basis functions. The upper line corresponds to the solution obtained by a straight-sided interpolation of the curvilinear geometry. As expected, this solution does not converge when increasing the finite element basis order. All the other curves correspond to the cases where third-order polynomial interpolation is used for the curvilinear geometry and metric expansion of order K is applied. The calculated convergence rates for the $K = 6$ cases are $\mathcal{O}(h^{2.19p})$ and $\mathcal{O}(h^{2.05p})$ for orders $p = 2$ and $p = 3$, respectively. As expected, the solution does not change at all after the representation limit ($K = 2p$ for the mass matrix). However, the termination of the expansion is permissible at a much earlier point since the convergence reached its asymptotic region. As predicted by the asymptotic limit in Section 4.3.1, Figure 4.6 clearly shows that the expansion order $K = p - 1$ is sufficient for the correct convergence rate.

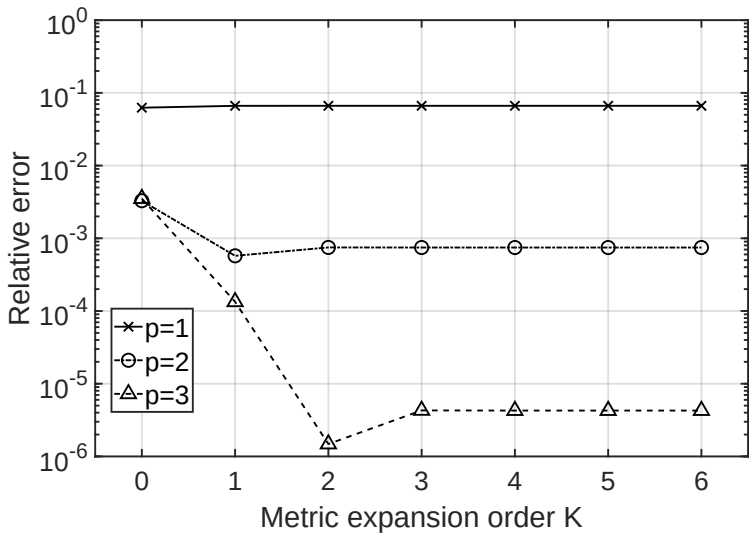


Figure 4.6: Relative error of the smallest nonzero eigenvalue in terms of the metric expansion order K. Parameter: FE basis order p.

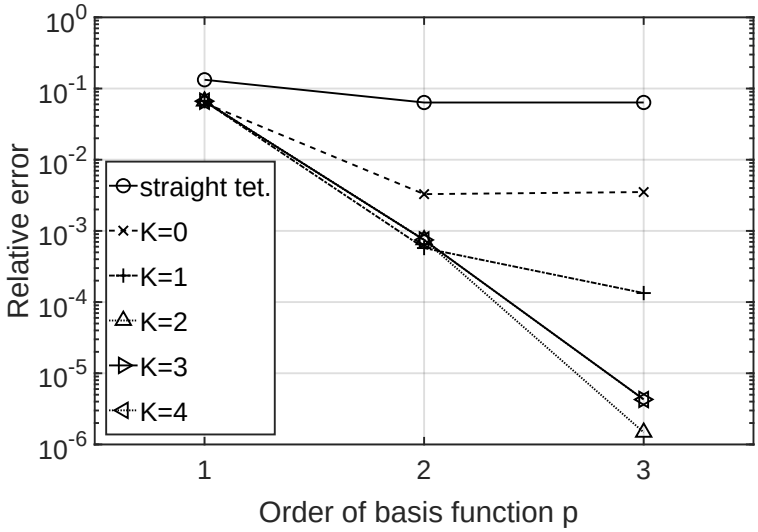


Figure 4.7: Relative error of the smallest nonzero eigenvalue in terms of the FE basis order p. Geometry representation: straight-sided and polynomial third-order tetrahedral mesh with metric expansion. Parameter: metric expansion of order K.

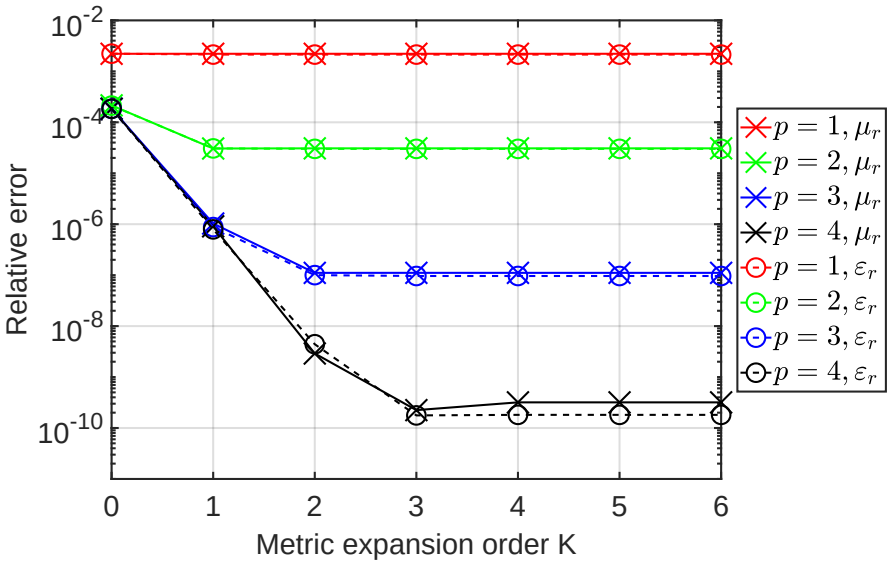


Figure 4.8: Relative error of the dominant eigenvalue in terms of the metric order K for the inhomogeneous μ_r and ε_r configurations. Parameter: FE basis order p.

Rectangular Resonator with Inhomogeneous Materials

The following example demonstrates the method for straight elements and inhomogeneous materials. The model is a rectangular cavity with linear dimensions of $1\text{m} \times 2\text{m} \times 3\text{m}$ bounded by perfect electrical conductors. The utilized mesh consists of 1152 straight-sided elements. Two complementary sets of material properties are considered. In the first case, only the permittivity is inhomogeneous

$$\boldsymbol{\varepsilon}_r = (11 + 8\tanh(|\boldsymbol{r}| - 0.2)), \ \boldsymbol{\mu}_r = 1. \tag{4.55}$$

In the second configuration, the material properties are interchanged

$$\boldsymbol{\mu}_r = (11 + 8\tanh(|\boldsymbol{r}| - 0.2)), \ \boldsymbol{\varepsilon}_r = 1, \tag{4.56}$$

resulting in a smooth inhomogeneous permeability distribution. In both cases, the relative error of the dominant eigenvalue is calculated with respect to the solution obtained via $p = 4$ orders of FE basis functions with $K = 6$ order metric expansion on an additionally homogeneously refined mesh. The obtained relative errors are shown in Figure 4.8. The asymptotic limit $K = p - 1$ is sufficient for both configurations because the higher-order expansion orders do not improve on the solution.

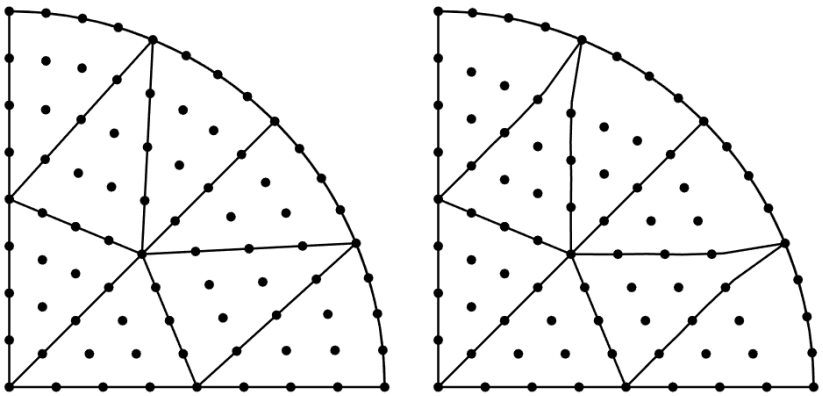


Figure 4.9: Side view of the mesh for a cylindrical resonator section with regular (left) and non-regular (right) fourth-order elements.

Cylindrical Resonator

The following example concerns the calculation of a cylindrical resonator of radius 1 m with perfect electric conductor boundaries and material properties of the vacuum. Due to the sole interest in the dominant eigenvalue, the symmetries of the dominant mode are utilized, modeling only the quarter geometry with a perfect electrical and a perfect magnetic conducting boundaries at cylinder cross-sections. The applied two different meshes are shown in Figure 4.9. The elements are curvilinear and interpolate the true geometry by fourth-order basis functions. The two different mesh versions only differ in the locations of the higher-order nodes with a regular and a slightly non-regular distribution, which affects the smoothness of the geometry mappings. Figure 4.10 shows the relative error of the smallest nonzero eigenvalue with respect to the analytic solution. The different curves correspond to solutions of different FE orders p and to the regular and non-regular meshes of Figure 4.9. One can observe convergence by increasing the metric expansion order. As expected, the accuracy of the solution is not improved after reaching a certain metric expansion order K determined by the representation limit ($K = 2p - 2$ for the stiffness matrix). Moreover, a rapid convergence can be seen in the case of the regular mesh where the geometry mappings are smoother. As it is predicted by the theorems of Section 4.3.1, the expansion can be terminated at a much earlier stage if the geometry representation and the finite element convergence have reached the asymptotic region. The proposed asymptotic limit $K = p - 1$ is in good agreement with the results of Figure 4.10 for the case of the regular mesh.

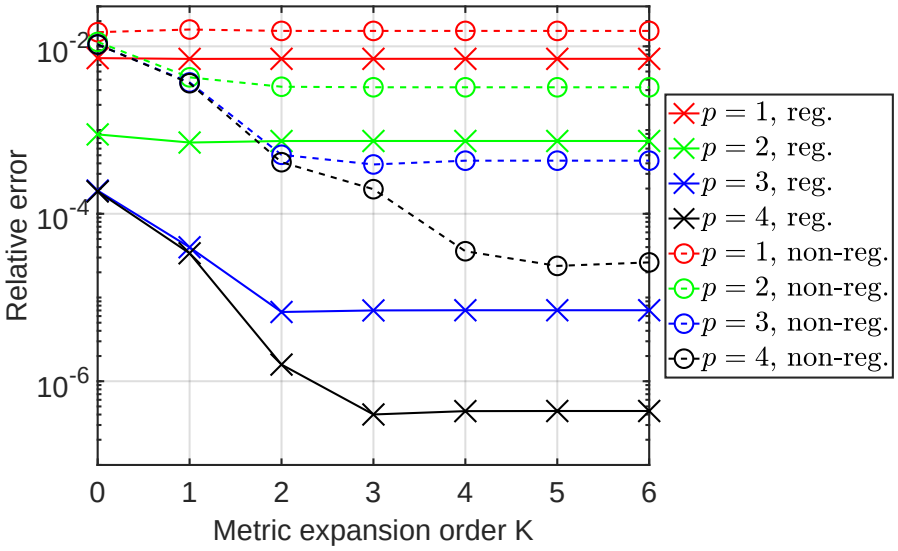


Figure 4.10: Relative error of the dominant eigenvalue in terms of the metric expansion order K. Parameter: FE basis order p, applied to the regularized (reg.) and non-regularized (non-reg.) higher-order mesh.

Cylindrical Cavity with Corrugated Post

The purpose of the last example is to qualitatively demonstrate the method on a geometry that can be used for real-life applications. Therefore, a cylindrical cavity, which includes a conducting post in the middle, is chosen. Such structures are used to design tunable waveguide filters [CMPR98] and resonators [LM07]. The initial geometry is an empty metallic cylinder with a radius of 30 mm and height of 86 mm, which is loaded with a post of varying diameter (see Figure 4.11). The three different sections have radii of 10 mm, 5 mm, and 10 mm and a height of 25 mm, 15 mm, and 25 mm, respectively. All metallic surfaces are perfect electric conductors, and the interior domain is vacuum. The geometry and the measurement results are taken from [LM07], as well as the reference results using Mode Matching (MM) and the Finite Difference Time Domain (FDTD) methods. In addition, Ansys HFSS Version 12.0 is used to obtain reference finite element solutions utilizing curvilinear elements with $p = 3$ order finite element basis functions. In the proposed implementation, curvilinear elements are applied with basis functions of order $p = 3$ as well as metric expansion order $K = 6$. The number of unknowns is 19248 in the proposed method and 22374 in the case of the HFSS solution. The resonance frequencies are provided in Table 4.8. Despite the relatively low number of unknowns, the proposed method shows a good agreement with the reference solutions and measurements. Note that this comparison only produces qualitative information. This is due to the fact that the approximation of the solution also depends on the mesh quality and the representation of the curvilinear geometry.

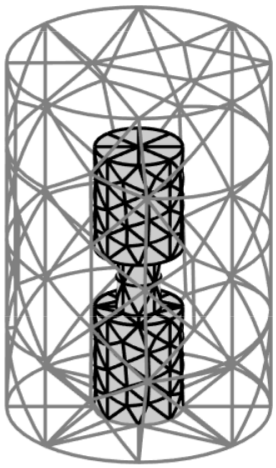


Figure 4.11: Mesh of the cylindrical cavity with a corrugated post.

Table 4.8: Resonance frequency comparison of the resonator in Fig. 4.11.

f_r [GHz]				
MM [LM07]	FDTD [LM07]	HFSS FEM	Proposed	Measurements [LM07]
2.976	2.969	2.968	2.977	2.970
4.169	4.149	4.155	4.166	4.153
5.010	5.008	5.005	5.010	4.990
5.840	5.838	5.834	5.839	5.820
6.862	6.864	6.861	6.862	6.830
7.007	7.001	7.000	7.003	7.004
7.483	7.499	7.485	7.478	7.475

4.4 Universal Matrices for Sensitivity Analysis

Sensitivity analysis plays an important role in the design process of microwave devices. It tells how the solutions are affected by small changes in a given variable or design parameter. A typical application is to incorporate the sensitivity values into the optimization process or to investigate the effect of minor deviations in the parameters, such as in the case of manufacturing errors. Generally, the first step is to obtain the derivatives of the FE equation system with respect to the desired design variables. All quantities of the equation system can be sorted into two groups. The first group consists of the known matrices and excitation vectors, and the second group consists of the unknown solution vectors. To express and calculate the derivatives of the solution vectors, the calculations of the FE matrices and excitation vectors, as well as their derivatives, are necessary. In this section, the previously described eigenvalue problem is considered for sensitivity analysis. This section shows how the hierarchical universal matrices can be used for the calculation of the derivative matrices, which are required for the calculation of eigenvalue derivatives. The determination of the derivative eigenvectors also requires the same matrices [Nel76]. However, the solution process differs for eigenvectors with eigenvalues of different multiplicity [Dai89]. Moreover, even the multiplicity of the derivative of the eigenvalues effects the solution process [ATMM07]. Hence, only eigenvalue sensitivities of a single multiplicity are considered to avoid technical complications.

Let us begin with the algebraic eigenvalue problem of (4.3) with a given eigenpair $(\boldsymbol{x}_i, \lambda_i)$ of eigenvector $\boldsymbol{x}_i$ and eigenvalue λ_i. Provided that $\boldsymbol{\mu}_r, \boldsymbol{\varepsilon}_r \in \mathbb{R}^+$ are real symmetric matrices, $\boldsymbol{S}$ and $\boldsymbol{T}$ become positive definite. Furthermore, an arbitrarily chosen design parameter is denoted by $\tau \in \mathbb{R}$, which smoothly parametrizes both the eigenpair $(\boldsymbol{x}_i, \lambda_i)$ and the finite element matrices. First, the derivative of (4.3) is calculated with respect to τ and then left multiplied by $\boldsymbol{x}_i^T$:

$$\left[\boldsymbol{x}_i^T \frac{\partial}{\partial \tau}\right] (\boldsymbol{S} - \lambda_i \boldsymbol{T})\, \boldsymbol{x}_i = 0, \tag{4.57}$$

$$\boldsymbol{x}_i^T \left(\frac{\partial}{\partial \tau}\boldsymbol{S} - \left(\frac{\partial}{\partial \tau}\lambda_i\right)\boldsymbol{T} - \lambda_i \frac{\partial}{\partial \tau}\boldsymbol{T}\right)\boldsymbol{x}_i + \boldsymbol{x}_i^T (\boldsymbol{S} - \lambda_i \boldsymbol{T}) \frac{\partial \boldsymbol{x}_i}{\partial \tau} = 0. \tag{4.58}$$

Here, the term

$$\boldsymbol{x}_i^T (\boldsymbol{S} - \lambda_i \boldsymbol{T}) = \boldsymbol{0}^T \tag{4.59}$$

becomes zero due to the definition of the eigenvector. Therefore, with the help of the normalization

$$\boldsymbol{x}_i^T \boldsymbol{T} \boldsymbol{x}_i = \boldsymbol{I} \tag{4.60}$$

and (4.59), (4.58) can be simplified, and the derivative of the eigenvalue is expressed as

$$\frac{\partial}{\partial \tau}\lambda_i = \boldsymbol{x}_i^T \left(\frac{\partial}{\partial \tau}\boldsymbol{S} - \lambda_i \frac{\partial}{\partial \tau}\boldsymbol{T}\right)\boldsymbol{x}_i. \tag{4.61}$$

Thus, the sensitivity analysis requires the considered eigenpair $(\lambda_i, \boldsymbol{x}_i)$ as well as the matrix derivatives $\frac{\partial}{\partial\tau}\boldsymbol{S}$ and $\frac{\partial}{\partial\tau}\boldsymbol{T}$. Let us consider the calculation of the hierarchical matrices $\boldsymbol{S}_{mn}$ (4.44a) and $\boldsymbol{T}_{mn}$ (4.44b), which are block matrices of the element matrices. Since the integrals are performed over the reference element, the basis functions $\widehat{\boldsymbol{w}}_m^a \in (\mathbb{P}_m)^3$ and $\widehat{\boldsymbol{w}}_n^b \in (\mathbb{P}_n)^3$ with $m+n \leq 2p$ are independent of the geometry and the material properties thus also of τ. Hence, the τ-derivatives of the matrices (4.44) read

$$\frac{\partial}{\partial\tau}[\boldsymbol{S}_{mn}]_{ab} = \int_{\widehat{\Omega}_e} \widehat{\nabla}\times\widehat{\boldsymbol{w}}_m^a \cdot \left(\frac{\partial}{\partial\tau}\widehat{\boldsymbol{\Lambda}}^2\right)\widehat{\nabla}\times\widehat{\boldsymbol{w}}_n^b \mathrm{d}\widehat{\boldsymbol{r}}, \tag{4.62a}$$

$$\frac{\partial}{\partial\tau}[\boldsymbol{T}_{mn}]_{ab} = \int_{\widehat{\Omega}_e} \widehat{\boldsymbol{w}}_m^a \cdot \left(\frac{\partial}{\partial\tau}\widehat{\boldsymbol{\Lambda}}^1\right)\widehat{\boldsymbol{w}}_n^b \mathrm{d}\widehat{\boldsymbol{r}}, \tag{4.62b}$$

where only the metric term can be dependent of τ. The structure of (4.62) implies that $\frac{\partial}{\partial\tau}\boldsymbol{S}$ and $\frac{\partial}{\partial\tau}\boldsymbol{T}$ may also be constructed from UMs. The procedure becomes the same as for the original metric expansion but with a different metric term. The polynomial expansion of $\frac{\partial}{\partial\tau}\boldsymbol{\Lambda}^{(\cdot)}$ in terms of b_k^l basis functions reads

$$\frac{\partial}{\partial\tau}\widehat{\boldsymbol{\Lambda}}^{(\cdot)}(\tau,\widehat{\boldsymbol{r}}) = \sum_{k=0}^{K}\sum_{l=1}^{L(k)}\left(\frac{\partial}{\partial\tau}\widehat{\boldsymbol{\Lambda}}_{kl}^{(\cdot)}(\tau)\right) b_k^l(\widehat{\boldsymbol{r}}), \tag{4.63}$$

where the derivative coefficients are calculated via the scalar product

$$\frac{\partial}{\partial\tau}\widehat{\boldsymbol{\Lambda}}_{kl}^{(\cdot)} = \int_{\widehat{\Omega}_e}\left(\frac{\partial}{\partial\tau}\widehat{\boldsymbol{\Lambda}}^{(\cdot)}\right) b_k^l(\widehat{\boldsymbol{r}})\mathrm{d}\widehat{\boldsymbol{r}}. \tag{4.64}$$

The substitution of the series expansion (4.63) for (4.62) leads to a polynomial representation, which allows for analytical integration. Because the only difference from the regular matrix calculation is the different metric coefficients, the same universal matrices and calculation procedure can be used with the derivative metric coefficients

$$\frac{\partial}{\partial\tau}\boldsymbol{S}_{mn} = \sum_{k=0}^{m+n-2}\sum_{l=1}^{L(k)}\sum_{i,j}\left[\frac{\partial}{\partial\tau}\widehat{\boldsymbol{\Lambda}}_{kl}^2\right]_{ij}\boldsymbol{S}_{ijl}^{mnk}, \tag{4.65a}$$

$$\frac{\partial}{\partial\tau}\boldsymbol{T}_{mn} = \sum_{k=0}^{m+n}\sum_{l=1}^{L(k)}\sum_{i,j}\left[\frac{\partial}{\partial\tau}\widehat{\boldsymbol{\Lambda}}_{kl}^1\right]_{ij}\boldsymbol{T}_{ijl}^{mnk}. \tag{4.65b}$$

Due to the same universal matrices, the computation cost of this evaluation is the same as in the regular case (see Section 4.3.3). Moreover, the same asymptotic (lower) limit, $K = p-1$, and representation (upper) limit, $K = m+n-2$ for $\frac{\partial}{\partial\tau}\boldsymbol{S}_{mn}$ and $K = m+n$ for $\frac{\partial}{\partial\tau}\boldsymbol{T}_{mn}$, applies to the derivative metric expansion.

4.4.1 Calculation of Metric Derivatives

The remaining question is how to obtain the exact derivatives of the metric terms, which are the subject of the expansion (4.63) and required in (4.64). Let the geometry be described by H^1 interpolatory FE basis functions $L_i(\widehat{\boldsymbol{r}}) \in \widehat{\mathcal{V}}_p = \mathbb{P}_p$, using the parameter-dependent locations of the element nodes $\boldsymbol{r}_i(\tau)$ as interpolation points. Denoting the matrix of node coordinates by $\boldsymbol{R}(\tau)$ and the vector of basis functions by $\boldsymbol{L}(\widehat{\boldsymbol{r}})$, the geometry mapping of an element $\boldsymbol{g}_p \in (\widehat{\mathcal{V}}_p)^3$ takes the form

$$\boldsymbol{r} = \boldsymbol{g}_p(\tau, \widehat{\boldsymbol{r}}) = \sum_i \boldsymbol{r}_i(\tau) L_i(\widehat{\boldsymbol{r}}) = \boldsymbol{R}(\tau)\boldsymbol{L}(\widehat{\boldsymbol{r}}). \tag{4.66}$$

Thus, the derivative of the Jacobian reads

$$\frac{\partial}{\partial\tau}\boldsymbol{J} = \frac{\partial}{\partial\tau}\widehat{\nabla}\boldsymbol{g}^T = \left[\widehat{\nabla}\boldsymbol{L}^T\right]\left[\frac{\partial}{\partial\tau}\boldsymbol{R}^T\right]. \tag{4.67}$$

Utilizing Jacobi's rule and the product rule, the derivatives of the metric tensors $\widehat{\boldsymbol{\Lambda}}^1$ and $\widehat{\boldsymbol{\Lambda}}^2$ are obtained by

$$\begin{aligned}
\frac{\partial}{\partial\tau}\widehat{\boldsymbol{\Lambda}}^1 = {}& \widehat{\boldsymbol{\Lambda}}^1 \operatorname{tr}\left[\boldsymbol{J}^{-1}\frac{\partial}{\partial\tau}\boldsymbol{J}\right] - \det(\boldsymbol{J})\left(\boldsymbol{J}^{-T}\frac{\partial \boldsymbol{J}^T}{\partial\tau}\boldsymbol{J}^{-T}\left(\boldsymbol{\varepsilon}_r \circ \boldsymbol{g}\right)\boldsymbol{J}^{-1}\right) \\
& - \det(\boldsymbol{J})\left(\boldsymbol{J}^{-T}\left(\boldsymbol{\varepsilon}_r \circ \boldsymbol{g}\right)\boldsymbol{J}^{-1}\frac{\partial \boldsymbol{J}}{\partial\tau}\boldsymbol{J}^{-1}\right) \\
& + \det(\boldsymbol{J})\,\boldsymbol{J}^{-T}\frac{\partial}{\partial\tau}\left(\boldsymbol{\varepsilon}_r \circ \boldsymbol{g}\right)\boldsymbol{J}^{-1}, \qquad (4.68a) \\
\frac{\partial}{\partial\tau}\widehat{\boldsymbol{\Lambda}}^2 = {}& \det(\boldsymbol{J})^{-1}\left(\frac{\partial \boldsymbol{J}}{\partial\tau}\left(\boldsymbol{\mu}_r^{-1} \circ \boldsymbol{g}\right)\boldsymbol{J}^T + \boldsymbol{J}\left(\boldsymbol{\mu}_r^{-1} \circ \boldsymbol{g}\right)\frac{\partial \boldsymbol{J}^T}{\partial\tau}\right) \\
& - \widehat{\boldsymbol{\Lambda}}^2 \operatorname{tr}\left[\boldsymbol{J}^{-1}\frac{\partial}{\partial\tau}\boldsymbol{J}\right] + \det(\boldsymbol{J})^{-1}\,\boldsymbol{J}\frac{\partial}{\partial\tau}\left(\boldsymbol{\mu}_r^{-1} \circ \boldsymbol{g}\right)\boldsymbol{J}^T. \qquad (4.68b)
\end{aligned}$$

The advantage of the previous form (4.68) with (4.67) is that the metric derivatives can be formulated for arbitrary parametrizations. As an input, only the derivatives of the nodal interpolation locations are required in (4.67) with respect to the parameter. Note that all involved terms are 3×3 matrices. Therefore, the computational cost is relatively low. In order to speed up the evaluation of $\boldsymbol{J}$ or $\frac{\partial}{\partial\tau}\boldsymbol{J}$ in some quadrature locations, a precalculated set of $\boldsymbol{L}(\widehat{\boldsymbol{r}})$ and $\widehat{\nabla}\boldsymbol{L}^T$ can be stored on the typical numerical integration locations.

4.4.2 Numerical Example

The following example features an empty spherical resonator placed at the center of the coordinate system and terminated by perfect electric conductor boundaries. The nominal radius is 1 m, and the nominal coordinate vector of the domain is denoted by $\widetilde{\boldsymbol{r}}$. The coordinate vector $\boldsymbol{r}$ of the parameterized sphere smoothly depends on a parameter τ and the nominal coordinate vector

$$\boldsymbol{r}(\tau) = (1+\tau)\widetilde{\boldsymbol{r}}, \qquad |\widetilde{\boldsymbol{r}}| \leq 1. \tag{4.69}$$

Thus, the interpolation nodes determining the nominal geometry depend also on the parameter via (4.69). This can be considered as a proportional radial contraction or expansion of the sphere. The goal is to compute the sensitivity of the smallest nonzero eigenvalue with respect to τ. To avoid technicalities with higher-multiplicity eigenvalues, only one-quarter of the structure is modeled. Thereby, perfect magnetic conductor boundary conditions are applied on the two cross-sections due to the symmetry of the mode. The structure is discretized into 512 tetrahedra using elementwise third-order polynomial interpolation for the curvilinear boundaries. The reference values are obtained from analytical calculations [Har01].

Figure 4.12 shows the relative error in the sensitivity of the dominant eigenvalue at the $\tau = 0$ location of the parameter space. It can be seen that straight tetrahedral elements are not sufficient and it is necessary to use curvilinear elements and an appropriate approximation order for the metric tensors and their derivatives. Figure 4.13 showcases the same data as Figure 4.12 but in terms of the metric expansion order instead of the finite element order. As expected, increasing the metric expansion beyond a finite element specific limit does not change the solution and the error remains unchanged. From the properties of the universal matrices it is known that the metric representation ($K = 2p - 2$ for the stiffness matrix) provides this sharp upper limit.

For smooth fields, the asymptotic convergence rate of the eigenvalue sensitivity $\frac{\partial}{\partial\tau}\lambda_i$ is expected to be $\mathcal{O}(h^{2p})$. The rates obtained for metric expansion order $K = 6$ are of order $\mathcal{O}(h^{2.04p})$ and $\mathcal{O}(h^{2.07p})$ when the $p = 1$ order error is compared to the $p = 2$ and $p = 3$ cases. Since the geometry representation and the finite element convergence have reached their asymptotic region, the asymptotic limit $K = p - 1$ of the metric expansion is sufficient. Figure 4.13 clearly demonstrates that above these limits, the error may change slightly but should not affect the convergence rates.

Note how similarly the proposed method works for the ordinary eigenvalue problem and eigenvalue sensitivity calculations. The obtained convergence curves show an excellent agreement in tendency with Figure 4.6 and Figure 4.7, where only the convergence of the eigenvalue is investigated.

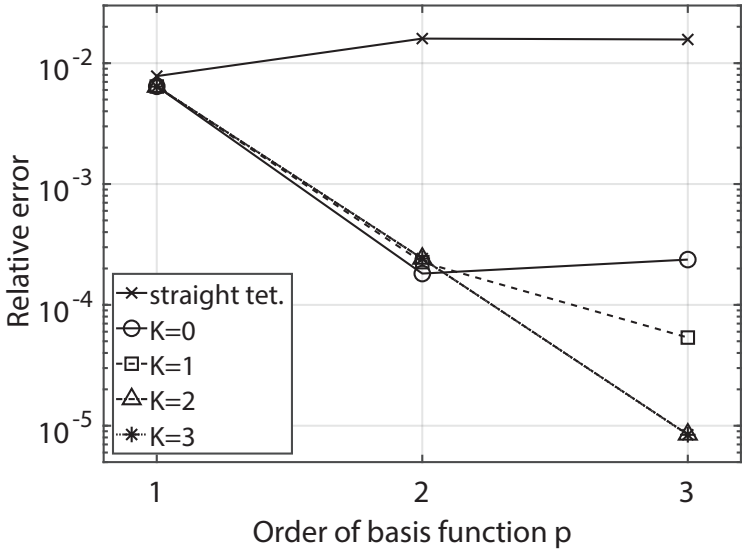


Figure 4.12: Relative sensitivity error of the smallest nonzero eigenvalue in terms of the FE basis order p. Geometry representation: straight-sided and polynomial third-order tetrahedral mesh with metric expansion. Parameter: metric expansion of order K.

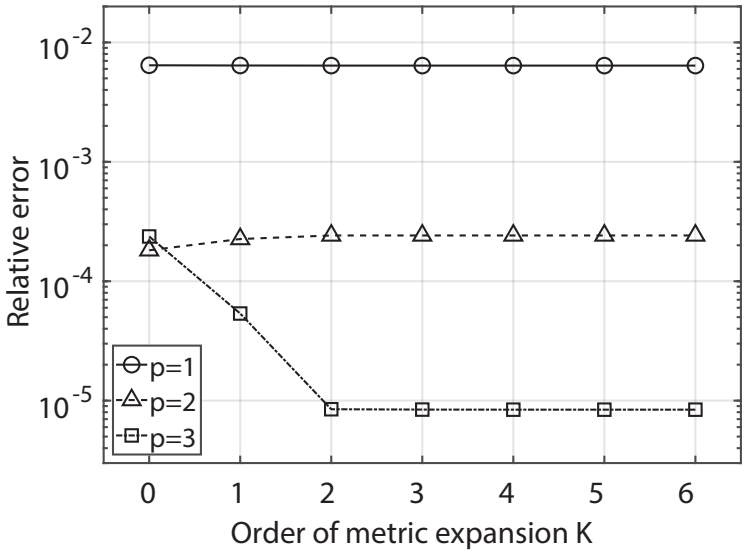


Figure 4.13: Relative sensitivity error of the smallest nonzero eigenvalue in terms of the metric expansion order K. Parameter: order of FE basis functions p.

4.5 Summary

This chapter presents a new calculation scheme for curvilinear finite element matrices. The method is based on a polynomial expansion of the material- and geometry-dependent metric terms of the finite element matrices, allowing the construction of universal matrices. For this purpose, a new set of hierarchical and pairwise orthonormal polynomials are developed that reflect the topological symmetries of arbitrary tetrahedra. As an outcome, the element matrices can be calculated efficiently by scaled matrix additions of the precalculated universal matrices. Based on the complexity analysis, the proposed method is more efficient than competing approaches. The unique feature is a proven upper representation limit for the required order of expansion and universal matrices. Hence, a finite number of universal matrices is sufficient for the exact calculation of the element matrices. Moreover, if the finite element convergence is in the asymptotic region, it is permissible to terminate the metric expansion and the number of universal matrices at a well-defined earlier point than the representation limit. These lower and upper limits are unique features of the proposed method and provide reliable and efficient calculations.

Chapter 5

A New Set of $\boldsymbol{H}$(div)-Conforming Basis Functions for Tetrahedra

$\boldsymbol{H}$(div)-conforming elements have a wide range of applications. In electromagnetics, they are commonly used for integral equation based methods [Bot06, CZZ$^+$15, PYA18] and finite elements [Néd80]. There are many applications in mechanics, fluid dynamics [Néd80, Néd86], and in acoustic field modeling [Gev88], [Jen07], [KK15]. In these fields, the $\boldsymbol{H}$(div)-conforming elements are commonly used by mixed finite element methods [BF12, LAH09, WXY12, CDF$^+$16].

There are two main types of bases in the literature. The first type is the interpolatory basis, where the coefficients represent the field values at specified interpolation locations. The second type is the hierarchical basis, where the order can be increased by including higher incremental order functions and allows heterogeneous basis function orders within the same mesh. However, the constant coefficients of the basis functions are no longer the field values. In the present work, hierarchical bases are preferred due to their efficient application in *hp*-adaptive schemes [AP02] and multilevel solvers [AFW97, Hip97]. Moreover, the proposed metric expansion and universal matrix schemes (Chapter 4) are directly applicable to polynomial basis functions over the reference tetrahedron.

The two best known $\boldsymbol{H}$(div)-conforming tetrahedral elements in the electromagnetic community are the complete-order [Néd86] and mixed-order Nédélec elements [Néd80]. These elements are also referred to as Brezzi-Douglas-Marini and Raviart-Thomas elements [BF12]. In [SWG84], a mixed first-order basis was given for tetrahedra, which is equivalent to the first-order Raviart-Thomas element [RT77]. Later on, Coyle and Ainsworth [AC03] as well as Zaglmayr [Zag06] developed higher-order hierarchical bases for the complete-order case with explicitly given functions. Both bases are constructed from orthogonal polynomials to maintain linear independence and reduce the condition number of the mass matrix. However, the orthogonality

of the specified functions only holds on the reference domain, and it is not guaranteed for arbitrary mesh deformations. In [Bot07], Botha presented hierarchical bases for both complete- and mixed-order cases. The common property of the bases of Botha and Zaglmayr is that the functions are partitioned according to the null space of the divergence operator. Similar functions have been presented for the $\boldsymbol{H}$(curl)-conforming case in [Web99], [SLC01], and [Ing06]. These partitionings are advantageous for some solvers and preconditioners, such as [AFW97], [Hip97], and [HX07], because they require the decomposition of the basis with respect to the kernel of the divergence operator. Moreover, a higher level of sparsity can be obtained for the stiffness matrices. The disadvantage of Botha's bases is the high condition number of the Gram matrix. Graglia and Peterson followed an alternative approach and proposed a basis [GP12] for the mixed-order case, which does not span exactly the same space as Nédélec. Their functions cannot be partitioned according to the null space; however, they provide significantly better condition numbers compared to Botha and Zaglmayr. This property is obtained by orthogonalizing the functions on a regular tetrahedron and using special weighting factors.

The present work aims at developing a polynomial basis that produces sparse element matrices with low condition numbers for arbitrary straight-sided tetrahedra and provides a fast yet accurate calculation procedure for the finite element matrices. The proposed basis is hierarchical, spans the mixed-order space of Nédélec, and the higher-order functions are partitioned according to the null space of the divergence operator. The orthogonality of the basis with respect to the interpolation operator results in a hierarchical interpolation scheme and leads to general orthogonality properties over arbitrary straight-sided tetrahedra. The condition numbers are reduced by an additional orthogonalization step over a given, regular tetrahedron. As a result, the element matrices are sparse, and the higher-order region of the stiffness matrix contains nonzero elements only at the diagonal. A short comparison of properties for some available, explicitly given, and hierarchical $\boldsymbol{H}$(div) bases is provided in Table 5.1.

Table 5.1: A short comparison of some hierarchical $\boldsymbol{H}$(div) bases for tetrahedra.

	Zaglmayr [Zag06]	Botha [Bot07]	Graglia-Peterson [GP12]	Proposed
Order	complete	both	mixed	mixed
Subspace division	✓	✓	×	✓
Orthogonality	fixed shape	×	fixed shape	arbitrary shape

5.1 The Proposed Family of Basis

The present approach is inspired by the $\boldsymbol{H}(\mathrm{curl})$ basis of Ingelström [Ing06]. The aim is to achieve similar orthogonal properties with respect to the $\boldsymbol{H}(\mathrm{div})$ interpolation operator

$$\pi_p^d \boldsymbol{d}_q = \boldsymbol{0}, \qquad \forall \boldsymbol{d}_q \in \mathcal{D}_q \text{ and } q > p, \tag{5.1}$$

as for the $\boldsymbol{H}(\mathrm{curl})$ case in (2.120). Hence, the following hierarchical decomposition is considered:

$$\mathcal{D}_p = \widetilde{\mathcal{D}}_1 \oplus ... \oplus \widetilde{\mathcal{D}}_p, \tag{5.2}$$

where $\widetilde{\mathcal{D}}_p$ denotes the incremental space of order p. The goal is to reuse the range space functions in $\mathcal{A}_p$ for the construction of null space functions in the $\boldsymbol{H}(\mathrm{div})$-conforming case as $\nabla \times \mathcal{A}_p \subset \mathcal{D}_p$. Therefore, each incremental space is decomposed as

$$\widetilde{\mathcal{D}}_1 = \widetilde{\mathcal{F}}_1, \tag{5.3a}$$

$$\widetilde{\mathcal{D}}_p = \widetilde{\mathcal{F}}_p \oplus \nabla \times \widetilde{\mathcal{A}}_p, \qquad \text{for } p > 1. \tag{5.3b}$$

The first-order case deserves special treatment because $\widetilde{\mathcal{F}}_1$ already includes $\nabla \times \widetilde{\mathcal{A}}_1$. This is inevitable since there is no other way to maintain the continuity of normal components at the boundaries [Bot07], similarly to the Whitney functions in the $\boldsymbol{H}(\mathrm{curl}, \Omega)$-conforming case [Web99]. However, there are hierarchical subsets of functions in the higher-order case that form a basis for the higher-order null space $\nabla \times \widetilde{\mathcal{A}}_p$ of the divergence operator. Hence, except for the first-order case, it is sufficient to have basis functions for $\widetilde{\mathcal{V}}_p$, $\widetilde{\mathcal{A}}_p$, and $\widetilde{\mathcal{F}}_p$ with the same hierarchical and orthogonal properties. With these functions, not only the sequence of higher-order approximation spaces becomes complete,

$$\widetilde{\mathcal{V}}_p \xrightarrow{\nabla} \widetilde{\mathcal{A}}_p \oplus \nabla\widetilde{\mathcal{V}}_p \xrightarrow{\nabla\times} \widetilde{\mathcal{F}}_p \oplus \nabla \times \widetilde{\mathcal{A}}_p \xrightarrow{\nabla\cdot} \nabla \cdot \widetilde{\mathcal{F}}_p \quad \text{for } p > 1, \tag{5.4}$$

but the actual basis functions themselves show a sequence for the de Rham complex. Moreover, in each space, the corresponding basis functions show the same hierarchical and orthogonality properties with respect to the interpolation operator and provide the same hierarchical interpolation scheme as (2.122). Based on the two different types of finite element moments (2.114), namely face and volume, the subspaces can be furthermore divided into

$$\widetilde{\mathcal{F}}_1 = \widetilde{\mathcal{F}}_1^f, \tag{5.5a}$$

$$\widetilde{\mathcal{F}}_p = \widetilde{\mathcal{F}}_p^v, \qquad \text{for } p > 1, \tag{5.5b}$$

$$\nabla \times \widetilde{\mathcal{A}}_p = \nabla \times \widetilde{\mathcal{A}}_p^f \oplus \nabla \times \widetilde{\mathcal{A}}_p^v, \qquad \text{for } p > 1. \tag{5.5c}$$

Here, $\widetilde{\mathcal{F}}_1^f$ is the first-order face-associated function space, and $\widetilde{\mathcal{F}}_p^v$ is the higher-order volume-associated space with the dimensions of $\dim\{\widetilde{\mathcal{F}}_1^f\} = 4$ and $\dim\{\widetilde{\mathcal{F}}_p^v\} = p(p+1)/2$, respectively [Bot07]. Moreover, $\nabla \times \widetilde{\mathcal{A}}_p^f$ is the face- and $\nabla \times \widetilde{\mathcal{A}}_p^v$ the volume-associated divergence-free subspace with the dimensions of $\dim\{\nabla \times \widetilde{\mathcal{A}}_p^f\} = 4p$ and $\dim\{\nabla \times \widetilde{\mathcal{A}}_p^v\} = p(p-2)$, respectively [Ing06]. The proposed set of basis functions for the space $\widetilde{\mathcal{F}}_p^v$ is shown in Table 5.2 in terms of barycentric coordinates (see Section 2.6). From (5.1), the resulting orthogonality properties for the inner products of functions $\boldsymbol{d}_p \in \mathcal{D}_p$ and $\boldsymbol{d}_q \in \mathcal{D}_q$ are

$$\int_{\Omega_e} \nabla \cdot \boldsymbol{d}_p \nabla \cdot \boldsymbol{d}_q \,\mathrm{d}\boldsymbol{r} = 0, \qquad \mid p - q \mid > 0, \tag{5.6a}$$

$$\int_{\Omega_e} \boldsymbol{d}_p \cdot \boldsymbol{d}_q \,\mathrm{d}\boldsymbol{r} = 0, \qquad \mid p - q \mid > 2, \tag{5.6b}$$

that apply for arbitrary straight-sided elements Ω_e. The orthogonality relations with respect to a scalar polynomial function $f_p \in \mathbb{P}_p$ of order p are

$$\int_{\Omega_e} f_p \nabla \cdot \boldsymbol{d}_q \,\mathrm{d}\boldsymbol{r} = 0, \qquad q > p + 1, \tag{5.7a}$$

$$\int_{\Omega_e} \nabla f_p \cdot \boldsymbol{d}_q \,\mathrm{d}\boldsymbol{r} = 0, \qquad q > p + 1. \tag{5.7b}$$

The proofs of these relationships are provided in Appendix B.2. Moreover, there is a way to achieve an even higher level of orthogonality. It is possible to take any linear combination of the functions in the same subspace because it is also orthogonal with respect to the interpolation operator. First, the functions $\boldsymbol{f}_p \in \widetilde{\mathcal{F}}_p^v$ are ordered into matrices $\boldsymbol{F}_p^v = [\boldsymbol{f}_p^1, \boldsymbol{f}_p^2 \dots]$. To avoid misunderstandings, the ordering of the functions is the same as shown in Table 5.2. It is then permissible to choose alternative bases as

$$\boldsymbol{F}_{p,\mathrm{ort}}^v = [\boldsymbol{f}_{p,\mathrm{ort}}^1, \boldsymbol{f}_{p,\mathrm{ort}}^2 \dots], \tag{5.8}$$

that are obtained by linear combinations

$$\boldsymbol{F}_{p,\mathrm{ort}}^v = \boldsymbol{F}_p^v \boldsymbol{C}_p^{Fv}, \tag{5.9}$$

with invertible matrices $\boldsymbol{C}_p^{Fv}$. These matrices are chosen such that the corresponding stiffness matrices over the regular equilateral tetrahedron Ω_{reg} with unit-length edges are

$$\int_{\Omega_{\mathrm{reg}}} (\nabla \cdot \boldsymbol{F}_{p,\mathrm{ort}}^v)^T (\nabla \cdot \boldsymbol{F}_{p,\mathrm{ort}}^v) \,\mathrm{d}\boldsymbol{r} = \frac{3}{\sqrt{2}} \boldsymbol{I}, \qquad p > 1, \tag{5.10}$$

where $\boldsymbol{I}$ is an identity matrix. Because the remaining functions are null space functions, the higher-order stiffness matrix contains only elements at the diagonal for any other straight-sided elements. Such $\boldsymbol{C}_p^{Fv}$ matrices are given for the orders $p = 2$, $p = 3$, and $p = 4$ in Appendix B.3, Tables B.1, B.2, and B.3.

A similar procedure can be performed on the null space functions associated with the face or volume (see (2.119)). In this case, no additional orthogonality can be achieved on arbitrarily deformed elements. However, it is possible to lower the condition number of the mass matrix by orthogonalizing the functions of the same subspaces over a fixed domain. The chosen domain is the regular equilateral tetrahedron Ω_{reg} with unit-length edges. First, the basis functions in Table 2.2 that are associated to the volume $\boldsymbol{w}_p^v \in \widetilde{\mathcal{A}}_p^v$ are ordered into matrices $\boldsymbol{A}_p^v$. Second, the functions in Table 2.2 that are associated with a given face and belong to a specific incremental order $\boldsymbol{w}_p^f \in \widetilde{\mathcal{A}}_p^f$ are ordered into matrices $\boldsymbol{A}_p^f$. Then, the sets of $\boldsymbol{H}(\text{div})$ null space basis functions that are associated to a given face or volume are stored in the matrices

$$\nabla \times \boldsymbol{A}_p^{(\cdot)} = [\nabla \times \boldsymbol{w}_p^{(\cdot),1}, \nabla \times \boldsymbol{w}_p^{(\cdot),2}, \dots], \qquad \text{for } (\cdot) \in \{f, v\}. \tag{5.11}$$

In all cases, the ordering of functions within the matrices is the same as in Table 2.2. Furthermore, corresponding alternative sets of basis functions can also be stored in matrices

$$\nabla \times \boldsymbol{A}_{p,\text{ort}}^{(\cdot)} = [\nabla \times \boldsymbol{w}_{p,\text{ort}}^{(\cdot),1}, \nabla \times \boldsymbol{w}_{p,\text{ort}}^{(\cdot),2} \dots], \qquad \text{for } (\cdot) \in \{f, v\}, \tag{5.12}$$

that are obtained as linear combinations with coefficient matrices $\boldsymbol{C}_p^{Af}$ and $\boldsymbol{C}_p^{Av}$:

$$\nabla \times \boldsymbol{A}_{p,\text{ort}}^f = \nabla \times \boldsymbol{A}_p^f \boldsymbol{C}_p^{Af}, \tag{5.13a}$$

$$\nabla \times \boldsymbol{A}_{p,\text{ort}}^v = \nabla \times \boldsymbol{A}_p^v \boldsymbol{C}_p^{Av}. \tag{5.13b}$$

In the present case, the additional freedom of coefficient matrices is used to achieve the desired orthogonality

$$\int_{\Omega_{\text{reg}}} (\nabla \times \boldsymbol{A}_{p,\text{ort}}^{(\cdot)})^T (\nabla \times \boldsymbol{A}_{p,\text{ort}}^{(\cdot)}) \, \mathrm{d}\boldsymbol{r} = \boldsymbol{D}, \qquad p > 1, \tag{5.14}$$

where $\boldsymbol{D}$ is a diagonal matrix. Note that the orthogonalization matrices $\boldsymbol{C}_p^{Af}$ are the same for the same sets of functions associated with different faces due to symmetry. Therefore, the calculation of $\boldsymbol{C}_p^{Af}$ matrices is only required for functions associated with a given face and can also be used for the other face associated functions.

The details of the orthogonalization procedure are given in Appendix B.3. In the case of face-associated functions, the coefficient matrices $\boldsymbol{C}_p^{Af}$ for the orders $p = 2$, $p = 3$, and $p = 4$ are provided in Tables B.4, B.5, and B.7, respectively. In case of the volume-associated functions, the coefficient matrices $\boldsymbol{C}_p^{Av}$ for the orders $p = 3$ and $p = 4$ are provided in Tables B.6 and B.8, respectively.

Table 5.2: Basis functions for the approximation space $\widetilde{\mathcal{F}}_p$.

$\widetilde{\mathcal{F}}_p$	Association	Basis functions
1	face ijk	$\boldsymbol{f}_1^{ijk} = \varphi_i \nabla\varphi_j \times \nabla\varphi_k + \varphi_j \nabla\varphi_k \times \nabla\varphi_i + \varphi_k \nabla\varphi_i \times \nabla\varphi_j$
2	volume $ijkl$	$\boldsymbol{f}_2^{ijkl} = 3\varphi_i \boldsymbol{f}_1^{jkl}$, $\boldsymbol{f}_2^{jikl}$, $\boldsymbol{f}_2^{kijl}$
3	volume $ijkl$	$\boldsymbol{f}_2^{ijkl}(2\varphi_j - \varphi_i) - \boldsymbol{f}_2^{jikl}(2\varphi_i - \varphi_j)$, $\boldsymbol{f}_2^{jikl}(2\varphi_k - \varphi_j) - \boldsymbol{f}_2^{kijl}(2\varphi_j - \varphi_k)$, $\boldsymbol{f}_2^{kijl}(2\varphi_l - \varphi_k) - \boldsymbol{f}_2^{lijk}(2\varphi_k - \varphi_l)$, $\boldsymbol{f}_2^{lijk}(2\varphi_i - \varphi_l) - \boldsymbol{f}_2^{ijkl}(2\varphi_l - \varphi_i)$, $\boldsymbol{f}_2^{ijkl}(2\varphi_k - \varphi_i) + \boldsymbol{f}_2^{kijl}(2\varphi_i - \varphi_k)$, $\boldsymbol{f}_2^{jikl}(2\varphi_l - \varphi_j) + \boldsymbol{f}_2^{lijk}(2\varphi_j - \varphi_l)$
4	volume $ijkl$	$\boldsymbol{f}_2^{ijkl}(6\varphi_i\varphi_j - 3\varphi_j^2 - \varphi_i^2) + \boldsymbol{f}_2^{jikl}(6\varphi_i\varphi_j - 3\varphi_i^2 - \varphi_j^2)$, $\boldsymbol{f}_2^{jikl}(6\varphi_j\varphi_k - 3\varphi_k^2 - \varphi_j^2) + \boldsymbol{f}_2^{kijl}(6\varphi_j\varphi_k - 3\varphi_j^2 - \varphi_k^2)$, $\boldsymbol{f}_2^{kijl}(6\varphi_k\varphi_l - 3\varphi_l^2 - \varphi_k^2) + \boldsymbol{f}_2^{lijk}(6\varphi_k\varphi_l - 3\varphi_k^2 - \varphi_l^2)$, $\boldsymbol{f}_2^{lijk}(6\varphi_l\varphi_i - 3\varphi_i^2 - \varphi_l^2) + \boldsymbol{f}_2^{ijkl}(6\varphi_l\varphi_i - 3\varphi_l^2 - \varphi_i^2)$, $\boldsymbol{f}_2^{ijkl}(6\varphi_i\varphi_k - 3\varphi_k^2 - \varphi_i^2) - \boldsymbol{f}_2^{kijl}(6\varphi_i\varphi_k - 3\varphi_i^2 - \varphi_k^2)$, $\boldsymbol{f}_2^{jikl}(6\varphi_j\varphi_l - 3\varphi_l^2 - \varphi_j^2) - \boldsymbol{f}_2^{lijk}(6\varphi_j\varphi_l - 3\varphi_j^2 - \varphi_l^2)$, $\varphi_i \boldsymbol{f}_1^{jkl}(28\varphi_i^2 - 21\varphi_i + 3)$, $\varphi_j \boldsymbol{f}_1^{ikl}(28\varphi_j^2 - 21\varphi_j + 3)$, $\varphi_k \boldsymbol{f}_1^{ijl}(28\varphi_k^2 - 21\varphi_k + 3)$, $\varphi_l \boldsymbol{f}_1^{ijk}(28\varphi_l^2 - 21\varphi_l + 3)$

5.2 Application of Basis Functions

As an example, the acoustic vector wave equation [Jen07] is chosen over a domain Ω with homogeneous Dirichlet boundary conditions on the boundary Γ_D and homogeneous Neumann boundary conditions on Γ_N:

$$\rho\frac{\partial^2 \boldsymbol{u}}{\partial t^2} - \nabla\lambda_L\nabla\cdot\boldsymbol{u} = \boldsymbol{f}, \qquad \text{on } \Omega, \tag{5.15a}$$

$$\boldsymbol{n}\cdot\boldsymbol{u} = 0, \qquad \text{on } \Gamma_D, \tag{5.15b}$$

$$\nabla\cdot\boldsymbol{u} = 0, \qquad \text{on } \Gamma_N. \tag{5.15c}$$

Here, $\boldsymbol{u}$ is the displacement vector, ρ is the material density, λ_L is the Lamé coefficient, and $\boldsymbol{f}$ is the excitation vector. The notation $\boldsymbol{n}$ stands for the unit normal vector on Γ_D. In the case of homogeneous anisotropic material properties and time-harmonic fields, the corresponding eigenvalue problem of (5.15a) reads

$$\lambda\rho\boldsymbol{u} - \nabla\lambda_L\nabla\cdot\boldsymbol{u} = \boldsymbol{0}, \tag{5.16}$$

where $\boldsymbol{u}$ is the eigenfunction and λ is the associated eigenvalue. Using the Galerkin approach, the weak formulation of the problem becomes [Bot07]

$$\lambda\int_\Omega \boldsymbol{u}\cdot(\rho\boldsymbol{d})\,\mathrm{d}\boldsymbol{r} - \int_\Omega \nabla\cdot\boldsymbol{u}\,(\lambda_L\nabla\cdot\boldsymbol{d})\,\mathrm{d}\boldsymbol{r} = 0, \tag{5.17}$$

where $\boldsymbol{u}$ and the test functions $\boldsymbol{d}$ are in the space of $\boldsymbol{H}(\mathrm{div},\Omega)$ and satisfy the essential boundary condition (5.15b). The finite element discretization with basis functions $\boldsymbol{d}_a \in \mathcal{D}_p$ and $\boldsymbol{d}_b \in \mathcal{D}_p$ leads to a generalized algebraic eigenvalue problem

$$(\boldsymbol{S} - \lambda\boldsymbol{T})\boldsymbol{x} = \boldsymbol{0} \tag{5.18}$$

with an eigenvalue λ and the associated eigenvector $\boldsymbol{x}$. The mass matrix $\boldsymbol{T}$ and the stiffness matrix $\boldsymbol{S}$ are obtained by the assembly of the corresponding element matrices

$$[\boldsymbol{T}_e]_{ab} = \int_{\Omega_e} \boldsymbol{d}_a\cdot(\rho\boldsymbol{d}_b)\,\mathrm{d}\boldsymbol{r}, \tag{5.19a}$$

$$[\boldsymbol{S}_e]_{ab} = \int_{\Omega_e} \nabla\cdot\boldsymbol{d}_a\,(\lambda_L\nabla\cdot\boldsymbol{d}_b)\,\mathrm{d}\boldsymbol{r}. \tag{5.19b}$$

Here, one can make use of the hierarchical property of the proposed functions. As in the case of $\boldsymbol{H}(\mathrm{curl},\Omega)$ matrices, hierarchical block matrices $\boldsymbol{T}_{mn}$ and $\boldsymbol{S}_{mn}$ can be constructed using basis functions $\boldsymbol{d}_m^a \in \widetilde{\mathcal{D}}_m$ of order $m \le p$ and $\boldsymbol{d}_n^b \in \widetilde{\mathcal{D}}_n$ of order $n \le p$,

$$\boldsymbol{S}_e = \left[\begin{array}{c|c|c} \boldsymbol{S}_{11} & \boldsymbol{S}_{12} & \cdots \\ \hline \boldsymbol{S}_{21} & \boldsymbol{S}_{22} & \cdots \\ \hline \vdots & \vdots & \ddots \end{array}\right], \qquad \boldsymbol{T}_e = \left[\begin{array}{c|c|c} \boldsymbol{T}_{11} & \boldsymbol{T}_{12} & \cdots \\ \hline \boldsymbol{T}_{21} & \boldsymbol{T}_{22} & \cdots \\ \hline \vdots & \vdots & \ddots \end{array}\right]. \tag{5.20}$$

Moreover, applying the appropriate function mappings (2.61c) and (2.61d), the calculation can be performed over the reference element $\widehat{\Omega}_e$:

$$[\boldsymbol{T}_{mn}]_{ab} = \int_{\Omega_e} \boldsymbol{d}_m^a \cdot (\rho \boldsymbol{d}_n^b)\,\mathrm{d}\boldsymbol{r} = \int_{\widehat{\Omega}_e} \widehat{\boldsymbol{d}}_m^a \cdot \widehat{\boldsymbol{\Lambda}}^2 \widehat{\boldsymbol{d}}_n^b\,\mathrm{d}\widehat{\boldsymbol{r}}, \tag{5.21a,b}$$

$$[\boldsymbol{S}_{mn}]_{ab} = \int_{\Omega_e} \nabla \cdot \boldsymbol{d}_m^a (\lambda_L \nabla \cdot \boldsymbol{d}_n^b)\,\mathrm{d}\boldsymbol{r} = \int_{\widehat{\Omega}_e} \widehat{\nabla} \cdot \widehat{\boldsymbol{d}}_m^a\, \widehat{\Lambda}^3\, \widehat{\nabla} \cdot \widehat{\boldsymbol{d}}_n^b \mathrm{d}\widehat{\boldsymbol{r}}. \tag{5.21c}$$

Here, the metric terms $\widehat{\boldsymbol{\Lambda}}^2$ and $\widehat{\Lambda}^3$ include the effect of the function mappings (2.61c) and (2.61d) as well as the volume element (2.72):

$$\widehat{\boldsymbol{\Lambda}}^2 = \det(\boldsymbol{J})^{-1} \boldsymbol{J} \rho \boldsymbol{J}^T, \tag{5.22a}$$

$$\widehat{\Lambda}^3 = \det(\boldsymbol{J})^{-1} \lambda_L. \tag{5.22b}$$

These metrics become constant matrices for straight-sided elements and elementwise inhomogeneous materials, and the integration can be performed analytically. However, in the case of curvilinear geometries, the metric is a rational tensor function, and exact evaluation is not possible.

5.2.1 Hierarchical Universal Matrices

For the efficient calculation of the finite element matrices, hierarchical metric expansions (Section 4.2) and hierarchical universal matrices (Section 4.3) are applied. In the present case, the Kth-order metric expansions read

$$\widehat{\boldsymbol{\Lambda}}^2(\widehat{\boldsymbol{r}}) = \sum_{k=0}^{K} \sum_{l=1}^{L(k)} \widehat{\boldsymbol{\Lambda}}_{kl}^2 b_k^l(\widehat{\boldsymbol{r}}), \qquad \widehat{\boldsymbol{r}} \in \widehat{\Omega}_e, \tag{5.23}$$

$$\widehat{\Lambda}^3(\widehat{\boldsymbol{r}}) = \sum_{k=0}^{K} \sum_{l=1}^{L(k)} \widehat{\Lambda}_{kl}^3 b_k^l(\widehat{\boldsymbol{r}}), \qquad \widehat{\boldsymbol{r}} \in \widehat{\Omega}_e, \tag{5.24}$$

where $L(k)$ is the number of homogeneous polynomials of order k and l is the index of the hierarchical orthonormal basis function b_k^l within the incremental order. Moreover, $\widehat{\boldsymbol{\Lambda}}_{kl}^2$ denotes the constant 3×3 matrix coefficients and $\widehat{\Lambda}_{kl}^3$ denotes the constant scalar coefficients of the metric expansion. The determination of these coefficients is done by evaluating the scalar products

$$\int_{\widehat{\Omega}_e} \widehat{\boldsymbol{\Lambda}}^2(\widehat{\boldsymbol{r}}) b_k^l \mathrm{d}\widehat{\boldsymbol{r}} = \widehat{\boldsymbol{\Lambda}}_{kl}^2, \tag{5.25a}$$

$$\int_{\widehat{\Omega}_e} \widehat{\Lambda}^3(\widehat{\boldsymbol{r}}) b_k^l \mathrm{d}\widehat{\boldsymbol{r}} = \widehat{\Lambda}_{kl}^3. \tag{5.25b}$$

These integrals are evaluated numerically by a quadrature rule capable of integrating polynomials of order $2K$ at minimum. Because the polynomial orders of basis functions within the mass and stiffness matrices are identical to the $\boldsymbol{H}(\mathrm{curl})$ case in Section 4.3.2, the same theories apply to the present case. Therefore, the $\boldsymbol{H}(\mathrm{div}, \Omega)$ hierarchical universal matrices are obtained by

$$[\boldsymbol{S}_l^{mnk}]_{ab} = \int_{\widehat{\Omega}_e} \widehat{\nabla} \cdot \widehat{\boldsymbol{d}}_m^a b_k^l \widehat{\nabla} \cdot \widehat{\boldsymbol{d}}_n^b \mathrm{d}\widehat{\boldsymbol{r}}, \tag{5.26a}$$

$$[\boldsymbol{T}_{ijl}^{mnk}]_{ab} = \int_{\widehat{\Omega}_e} \left[\widehat{\boldsymbol{d}}_m^a\right]_i b_k^l \left[\widehat{\boldsymbol{d}}_n^b\right]_j \mathrm{d}\widehat{\boldsymbol{r}}, \tag{5.26b}$$

where $i, j \in \{1, 2, 3\}$ denote the indices for the different vector components. Due to the same metric representation limit (4.46), a finite number of universal matrices is sufficient for the exact calculation. Hence, the calculation of the hierarchic element matrices reads

$$\boldsymbol{S}_{mn} = \sum_{k=0}^{m+n-2} \sum_{l=1}^{L(k)} \widehat{\Lambda}_{kl}^3 \boldsymbol{S}_l^{mnk}, \tag{5.27a}$$

$$\boldsymbol{T}_{mn} = \sum_{k=0}^{m+n} \sum_{l=1}^{L(k)} \sum_{ij} \left[\widehat{\boldsymbol{\Lambda}}_{kl}^2\right]_{ij} \boldsymbol{T}_{ijl}^{mnk}. \tag{5.27b}$$

Moreover, similarly to the $\boldsymbol{H}(\mathrm{curl}, \Omega)$ case in Section 4.3.1, all integrals schemes are specific subcases of the investigated general product (3.91). Therefore, in the asymptotic region of the geometry error convergence, the general estimate of the metric truncation error (3.108) applies. The application of (3.108) for the initial problem (5.17) leads to the metric truncation error

$$\epsilon_K \leq C h^{K+1} \left(\|\boldsymbol{u}\|_{\boldsymbol{L}^2(\Omega_e)} \|\rho\|_{K+1,\infty,\Omega_e} + \|\nabla \cdot \boldsymbol{u}\|_{L^2(\Omega_e)} \|\lambda_L\|_{K+1,\infty,\Omega_e} \right), \tag{5.28a}$$

$$\epsilon_K \leq C h^{K+1} \|\boldsymbol{u}\|_{\boldsymbol{H}(\mathrm{div},\Omega_e)} \left(\|\rho\|_{K+1,\infty,\Omega_e} + \|\lambda_L\|_{K+1,\infty,\Omega_e} \right), \tag{5.28b}$$

where C is a positive constant independent of K and h and $\boldsymbol{u}$ is the exact field approximated by the finite element. Thus, if the expected convergence rate of the finite element in $\boldsymbol{H}(\mathrm{div})$-norm is $\mathcal{O}(h^p)$, the sufficient asymptotic limit for the metric expansion order is

$$K = p - 1. \tag{5.29}$$

However, if the convergence of the geometry error or the solution is not in the asymptotic region, the metric expansion order may need to be increased until the metric representation limits (4.37) (see Section 4.3.1 and Section 3.1.7).

5.2.2 Numerical Examples

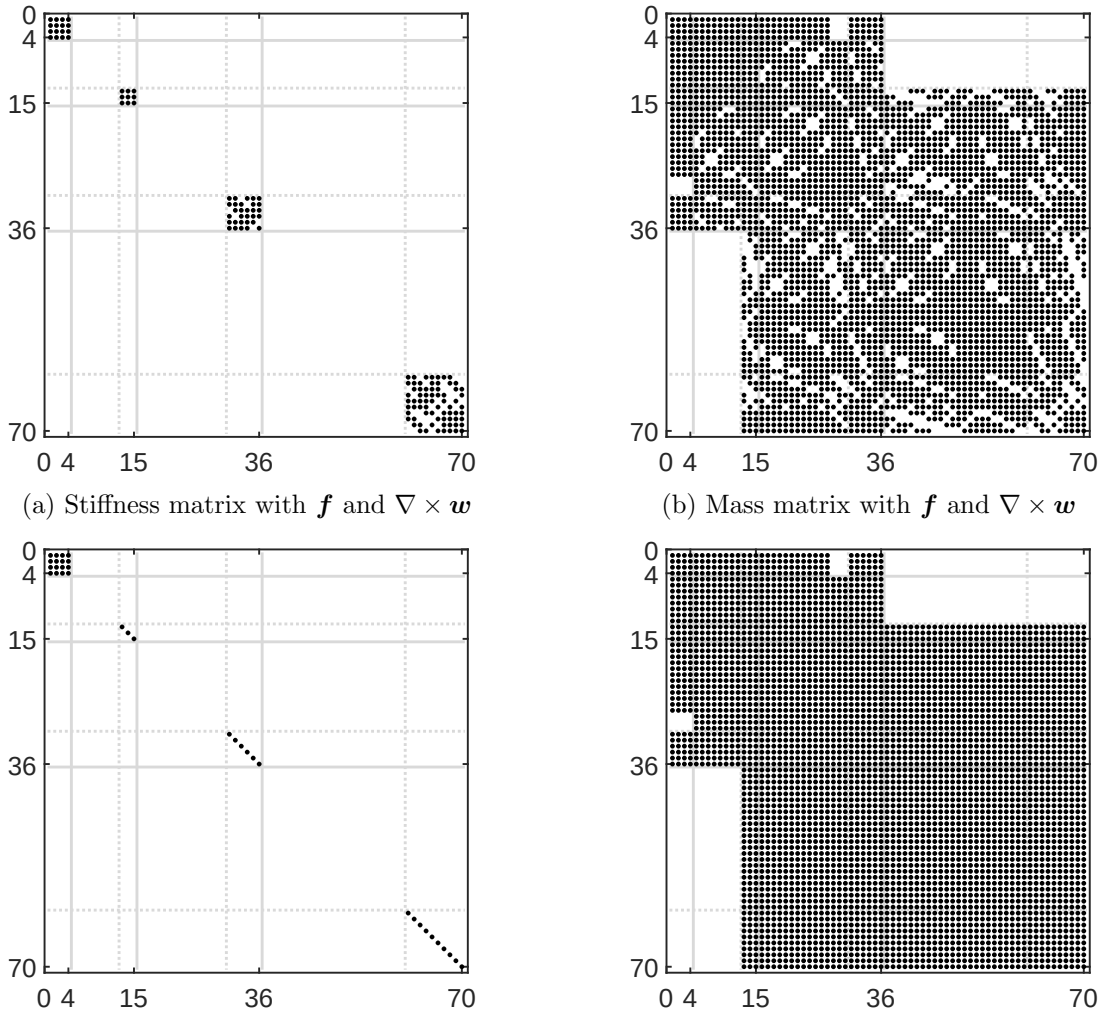
In this section, some numerical results are presented for the validation of the proposed bases. The proposed functions lead to four cases based on the combinations of subspaces with different orthogonality properties:

- Range space functions $\boldsymbol{f}$ (Table 5.2) and null space functions $\nabla \times \boldsymbol{w}$ (Table 2.2).
- Range space functions $\boldsymbol{f}$ (Table 5.2) and orthogonalized null space functions $\nabla \times \boldsymbol{w}_{ort}$ (orthogonalized face- and volume-associated functions (5.12)).
- Orthogonalized range space functions $\boldsymbol{f}_{\text{ort}}$ (orthogonalized volume-associated functions (5.8)) and null space functions $\nabla \times \boldsymbol{w}$ (Table 2.2).
- Orthogonalized range space functions $\boldsymbol{f}_{\text{ort}}$ (orthogonalized volume-associated functions (5.8)) and orthogonalized null space functions $\nabla \times \boldsymbol{w}_{\text{ort}}$ (orthogonalized face- and volume-associated functions (5.12)).

Since these different basis functions span the same space, they only differ in their numerical properties. Therefore, a comparison is given for the numerical properties with respect to other basis functions from the literature.

Straight-Sided Tetrahedron

First, a single straight-sided tetrahedron is chosen to show the advantage of the presented bases. For arbitrary affine geometry mappings, the sparsity pattern of the mass and stiffness matrix is shown in Figure 5.1 for the cases when the functions of both subspaces are either orthogonalized ($\boldsymbol{f}_{\text{ort}}, \nabla \times \boldsymbol{w}_{\text{ort}}$) or not ($\boldsymbol{f}, \nabla \times \boldsymbol{w}$). Here, the stiffness matrix becomes block diagonal or even diagonal in the orthogonalized case for the higher-order functions. On the other hand, the mass matrix is less sparse in the orthogonalized case, but the condition number is significantly lower. In order to obtain a fair comparison with other bases, the calculation of the condition numbers is performed for the element mass matrices after diagonal scaling. In Figure 5.2, these condition numbers are shown in terms of the basis function order p for all combinations of the bases. The geometry is chosen to be a regular tetrahedron with unit-length edges for which the Graglia-Peterson basis is considered state-of-the-art in terms of the condition number. It can be seen that the fully orthogonalized version of the proposed bases performs the best in the higher-order case. Moreover, the growth of the condition number is not significantly changed for the different deformations of the regular tetrahedron. This can be seen in Figure 5.3 for a reference tetrahedron with nodes of (0,0,0), (1,0,0), (0,1,0), and (0,0,1) and in Figure 5.4 for a test tetrahedron with nodes of (0,0,0), (1,2,3), (-1,2,-1), and (-3,0,2). One can also compare the condition numbers to the basis of Zaglmayr and Botha.



(a) Stiffness matrix with $\boldsymbol{f}$ and $\nabla \times \boldsymbol{w}$

(b) Mass matrix with $\boldsymbol{f}$ and $\nabla \times \boldsymbol{w}$

(c) Stiffness matrix with $\boldsymbol{f}_{\text{ort}}$ and $\nabla \times \boldsymbol{w}_{\text{ort}}$

(d) Mass matrix with $\boldsymbol{f}_{\text{ort}}$ and $\nabla \times \boldsymbol{w}_{\text{ort}}$

Figure 5.1: Sparsity patterns of the fourth-order $\boldsymbol{H}(\text{div})$ bases.

It is clear from Table 5.3 that the presented orthogonalized basis not just possesses the advantageous properties of the previous bases but also has the strongest linear independence. As the order of basis functions becomes higher, this basis can maintain a reasonable condition number on arbitrary geometries.

Moreover, Beuchler, Pillwein, and Zaglmayr suggest the usage of static condensation with higher-order basis functions, where the condition number of the volume-associated functions is of great importance [BPZ12]. They developed a new, high-order basis [BPZ12] using weighted Jacobi polynomials and used tensor product construction on the reference tetrahedron. Their functions produce low condition numbers and high sparsity for the mass matrix of the volume-associated functions $\boldsymbol{T}_v$. Moreover, their stiffness matrix for the higher-order volume-associated functions $\boldsymbol{S}_v$ is purely diagonal, which is similar to the present case. In [BPZ12], it is

suggested that the practically achieved condition number of the matrix $(\boldsymbol{T}_v + \boldsymbol{S}_v)$ should be considered with diagonal preconditioning. To compare the present work, the test tetrahedron with the nodes of (0,0,0), (0.315,0.632,0.158), (1.5,0,0), and (0,0,1) is used as suggested in [BPZ12, BPZ10]. The resulting condition numbers for different p orders of basis functions are shown in Table 5.4. It can be seen that proposed orthogonalized basis performs the best for the investigated orders.

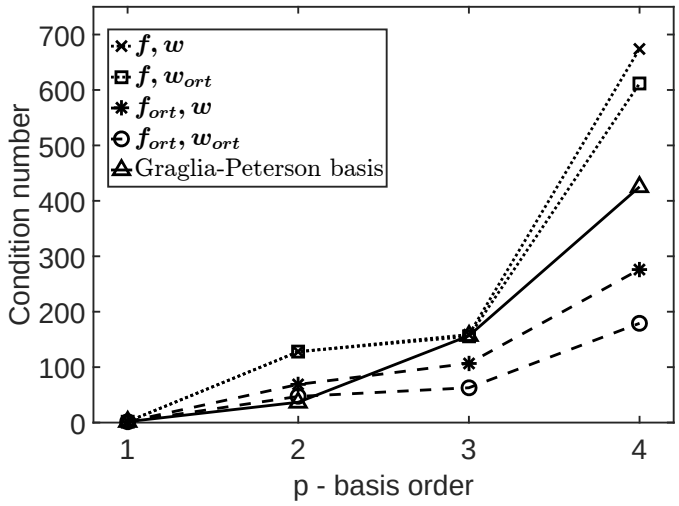


Figure 5.2: Condition number of the normalized mass matrix for the regular tetrahedron with unit length edges.

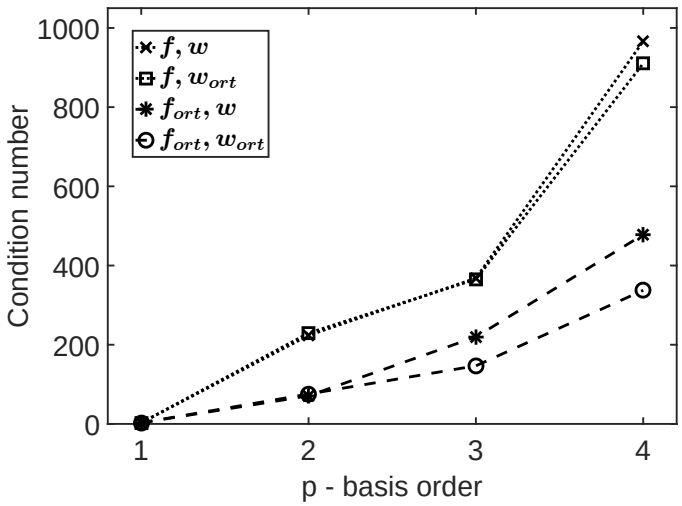


Figure 5.3: Condition number of the normalized mass matrix for a reference tetrahedron with nodes of (0,0,0), (1,0,0), (0,1,0), and (0,0,1).

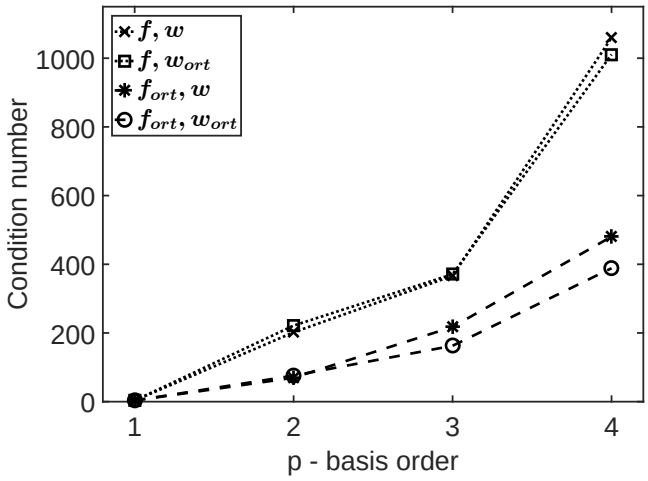


Figure 5.4: Condition number of the normalized mass matrix for a straight-sided test tetrahedron with nodes of (0,0,0), (1,2,3), (-1,2,-1), and (-3,0,2).

Table 5.3: Condition number of mass matrices after diagonal scaling. The reference values are taken from [GP12].

Regular tetrahedron with unit length edges							
p	Zaglmayr basis [Zag06]	Botha basis [Bot07]	Graglia-Peterson basis [GP12]	Proposed basis			
				$\boldsymbol{f}, \boldsymbol{w}$	$\boldsymbol{f}, \boldsymbol{w}_{\text{ort}}$	$\boldsymbol{f}_{\text{ort}}, \boldsymbol{w}$	$\boldsymbol{f}_{\text{ort}}, \boldsymbol{w}_{\text{ort}}$
1	1.667	1.667	1.667	1.667	1.667	1.667	1.667
2	47.2	143.8	36.71	127.8	127.9	68.87	47.19
3	188.1	303.1	156.6	158.7	155.4	106.8	62.82
4	n/a	n/a	424.3	674.1	611.6	276.2	179.0
Reference tetrahedron							
p	Zaglmayr basis [Zag06]	Botha basis [Bot07]	Graglia-Peterson basis [GP12]	Proposed basis			
				$\boldsymbol{f}, \boldsymbol{w}$	$\boldsymbol{f}, \boldsymbol{w}_{\text{ort}}$	$\boldsymbol{f}_{\text{ort}}, \boldsymbol{w}$	$\boldsymbol{f}_{\text{ort}}, \boldsymbol{w}_{\text{ort}}$
1	2.816	2.816	2.816	2.816	2.816	2.816	2.816
2	50.82	252.2	71.48	223.5	229.4	70.88	75.54
3	309.6	480.4	410.8	368.1	365.2	219.7	146.7
4	n/a	n/a	n/a	966.2	910.3	477.8	337.4

Table 5.4: Practically achieved condition numbers of $(\boldsymbol{T}_v + \boldsymbol{S}_v)$ with only volume associated basis functions. The reference values are taken from [BPZ12].

Test tetrahedron					
p	Beuchler-Pillwein-Zaglmayr basis [BPZ12], [BPZ10]	Proposed basis			
		$\boldsymbol{f}, \boldsymbol{w}$	$\boldsymbol{f}, \boldsymbol{w}_{\text{ort}}$	$\boldsymbol{f}_{\text{ort}}, \boldsymbol{w}$	$\boldsymbol{f}_{\text{ort}}, \boldsymbol{w}_{\text{ort}}$
2	6.5	5.9	5.9	1.5	1.5
3	17	9.6	8.2	7.3	4.7
4	37	82	82	21	7.5

Rectangular Resonator

In order to demonstrate the standard convergence behavior of the mixed-order Nédélec element with the proposed basis, the eigenvalue problem of (5.18) is chosen for a rectangular resonator with the dimensions of 3 m $\times$ 2 m $\times$ 2 m. All boundaries are set to homogeneous Dirichlet boundary conditions, and the material density $\rho = 1$ and Lamé coefficient $\lambda_L = 1$ are unit valued over the whole domain. The calculations are performed on a uniformly refined mesh where the size of the elements is the same. The calculated first nonzero eigenvalues are shown in Figure 5.5.

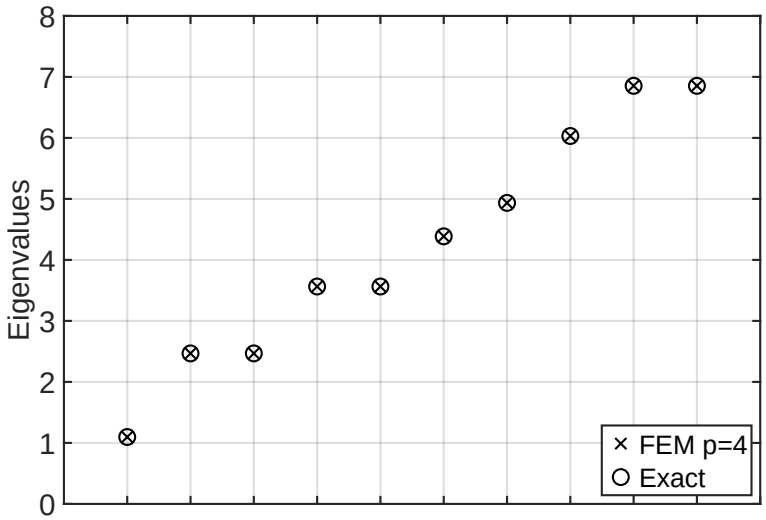


Figure 5.5: The first nonzero eigenvalues of the rectangular resonator.

Further on, only the dominant mode is considered which has the smallest eigenvalue. Figure 5.6 shows the relative error for different p orders of basis functions in terms of the number of refinements. It can be seen that an algebraic rate of convergence is achieved. In the case of smooth fields and sufficient geometry representation, the asymptotic rate of convergence of the Nédélec [Néd80] elements is $\mathcal{O}(h^{2p})$ for such an eigenvalue [Bot06], where h is the element size and p is the finite element order. From the determined steepness of the $p = 1, 2, 3,$ and 4 order convergence curves, the obtained convergence rates are $\mathcal{O}(h^{2.91})$, $\mathcal{O}(h^{3.97})$, $\mathcal{O}(h^{6.32})$, and $\mathcal{O}(h^{8.14})$, respectively. Thus, the performance of the proposed basis shows good agreement with the theoretical value.

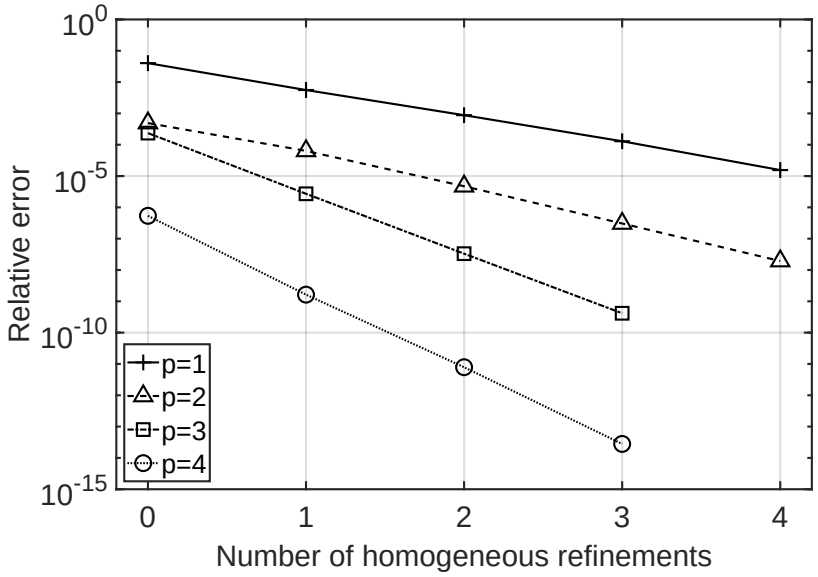


Figure 5.6: Relative error of the smallest nonzero eigenvalue in terms of the number of homogeneous refinements.

Spherical Resonator

In the case of curvilinear elements, the exact integration of the finite element matrices is not possible. However, it is possible to use numerical integration or to apply the proposed method based on hierarchical universal matrices. In this example, the basis functions and the corresponding universal matrices are demonstrated on the eigenvalue problem of a spherical resonator with a unit radius. Both the material density $\rho = 1$ and the Lamé coefficient $\lambda_L = 1$ are unit valued over the whole domain. For the dominant mode, the symmetry of the solution is used to reduce the computational domain to one eighth of the sphere. Hence, a smaller number of unknowns is required, and the multiplicity of the smallest eigenvalue is reduced to one. Thus, homogeneous Dirichlet boundary condition is applied to the curvilinear outer boundary and to two of the cross-section planes. The single remaining cross-section plane is a homogeneous Neumann boundary. Two different discretizations are used for the domain: a coarser mesh with 678 tetrahedra and a finer mesh with 43392 tetrahedra. The finer mesh is obtained by two-times uniform refinement of the coarser mesh, and both meshes interpolate the sphere by fourth-order basis functions. For this choice, it is known that the error in geometry representation will be at least in the same order of magnitude or less than in the FE solution for $p \leq 4$. These utilized meshes are shown in Figure 5.7. The first few nonzero eigenvalues are given in Figure 5.8, where the proposed universal matrices are applied up to order $K = 6$ and FE basis functions of order $p = 4$ on the fine mesh. Next, only the dominant mode and the corresponding smallest nonzero eigenvalue are considered. Figure 5.9 and Figure 5.10 show the convergence of this eigenvalue in terms of the metric ex-

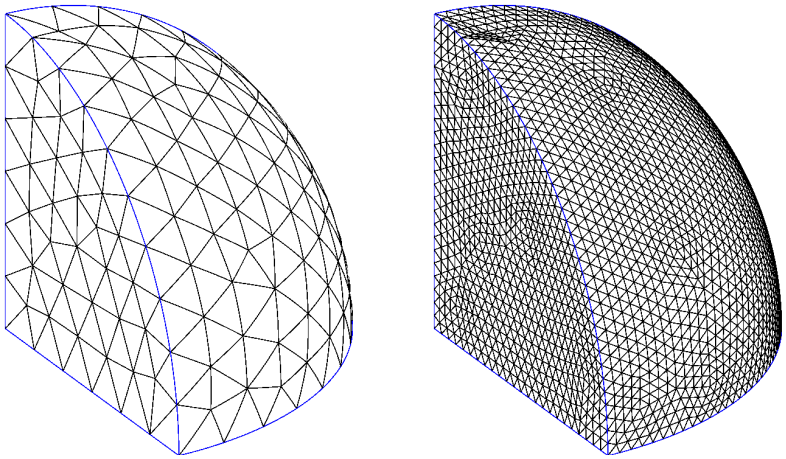


Figure 5.7: Coarse (left) and fine (right) mesh of the sphere section.

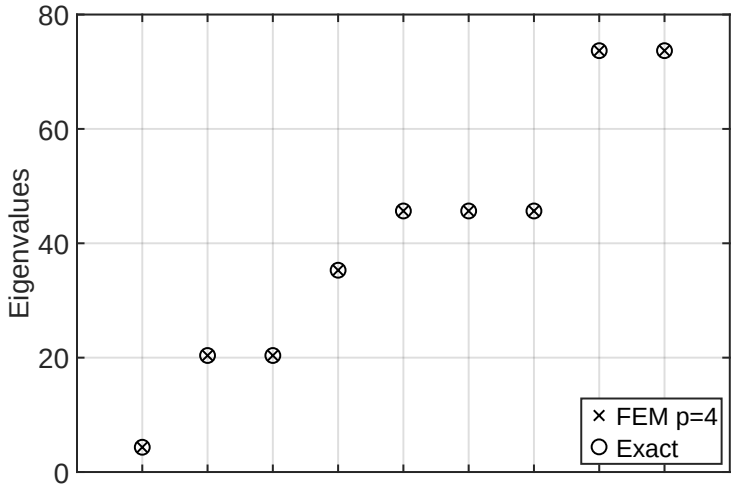


Figure 5.8: The first nonzero eigenvalues using metric expansion order $K = 6$ and FE basis functions of order $p = 4$ on the fine mesh.

pansion order K for the different p order basis functions. As expected, higher-order basis yields better solutions but also requires higher-order metric expansion in order to reach its ideal performance. The solutions obtained for the coarser and finer mesh fit the theoretical predictions, displaying the two limiting cases. In any scenario, the sufficient order of metric expansion is given by the representation limit ($K \leq 2p-2$ for the stiffness matrix), which is $K = 0, 2, 4, 6$ for the stiffness matrix orders of $p = 1, 2, 3, 4$, respectively. This can be observed in Figure 5.9, where the FE convergence has not fully reached its asymptotic region. However, Figure 5.10 shows the solution for the finer mesh, where asymptotic region of convergence is reached. In this case, it is possible to terminate the expansion at an earlier point determined

by the asymptotic limit ($K = p - 1$), which is $K = 0, 1, 2, 3$ for FE basis functions of order $p = 1, 2, 3, 4$, respectively. The present numerical experiments show a good agreement with these theoretical limits.

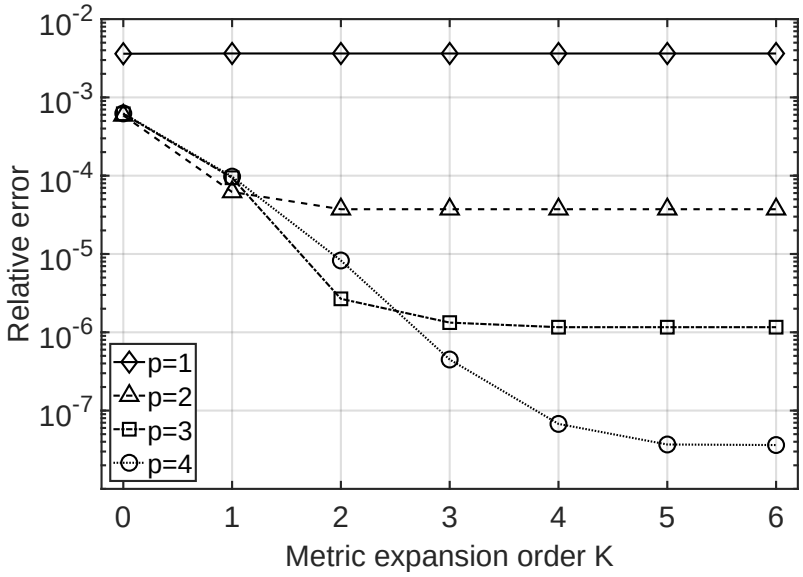


Figure 5.9: Convergence of the smallest nonzero eigenvalue by using FE basis functions of order p and metric expansions of order K for the coarse mesh.

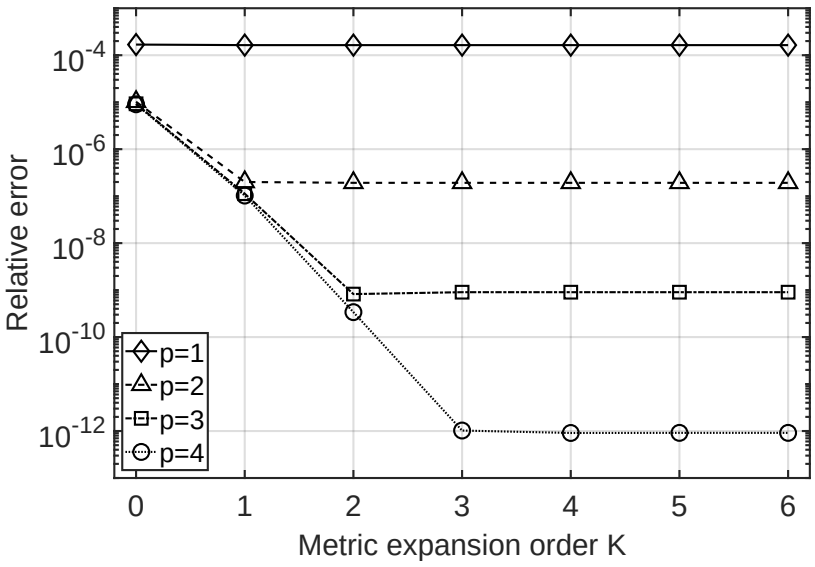


Figure 5.10: Convergence of the smallest nonzero eigenvalue by using FE basis functions of order p and metric expansions of order K for the fine mesh.

5.3 Summary

In this chapter, a new set of $\boldsymbol{H}$(div) basis functions has been presented, which completes the sequence of the discrete de Rham complex with the previously available H^1 and $\boldsymbol{H}$(curl) bases [Ing06]. The common properties of these bases are the orthogonality with respect to the interpolation operators and distinct subsets of functions for the higher-order null space functions. This leads to element matrices with advantageous numerical properties, such as a high level of sparsity and orthogonality properties, for arbitrary straight-sided elements. A special feature of the presented orthogonalized basis is that the higher-order sections of the generated stiffness matrices are purely diagonal for arbitrary straight-sided elements. Moreover, the condition number of the resulting mass matrix is significantly lower than other competing basis functions from the literature. The element matrix calculation of the basis functions is accelerated with hierarchical universal matrices. The proposed functions and methods are demonstrated via eigenvalue problems.

Chapter 6

Hexahedral Finite Elements

Historically, the first hexahedral elements were defined as tensor product approximation spaces [Cia02, Chapter 2.2], [SF73, Section 1.9]. The advantage is the straightforward construction from the one-dimensional case [SF73, Section 1.9]. However, the higher-order spaces became significantly larger than the same-order spaces for tetrahedra. To overcome this disadvantage, serendipity spaces were developed with fewer moments, typically leaving out face- or volume-associated functions [ZK00]. Thus, the spaces were reduced, becoming more efficient from the computational perspective. It has been proven that including a complete-order polynomial space and maintaining the conformity are sufficient for the theoretical rate of convergence [AA11], [AA14] of H^1-conforming elements with affine geometry mappings. Similar results have been obtained for the $\boldsymbol{H}(\mathrm{curl})$- and $\boldsymbol{H}(\mathrm{div})$-conforming spaces as well [AA14], [GK19], [GKS19]. The drawback of such spaces is the restriction to affine geometry mappings (i.e., parallelepiped elements) [Ape98], [ABF02], [ABB15], which limits their application. For general hexahedral meshes, the traditional tensor product [BF12, Chapter 2], [AP02], [Néd80], [Néd86] or even larger and more complicated spaces [ABB15], [FGM11], [BD13b], [BD13a] are being used to maintain the rate of convergence at the price of a greatly increased number of unknowns. In most cases, error estimations are available; however, their application is limited to straight-sided geometries. Moreover, up to the author's knowledge, the effect of different mesh refinements has not been considered yet.

In Chapter 3, convergence requirements of curvilinear elements are discussed from the point of view of geometry interpolation and mesh refinement. These new aspects provide additional possibilities to maintain the theoretical rate of convergence for highly deformed meshes. The first contributions of this chapter are alternative definitions of the different finite elements, which are appropriate for the theories of Chapter 3. Note that these definitions are equivalent to the known finite elements in the literature [Néd80], [Néd86], [FGM11], [BD13b], [GK19], [GKS19] defined for straight-sided elements. However, the present definitions of the finite element mo-

ments are always given for the reference element, yielding different moments for the corresponding mapped curvilinear elements. As a general consequence, all the considered spaces are capable of $\mathcal{O}(h^p)$ convergence rates for curvilinear elements when only affine mesh refinement is applied. In addition to these known finite element spaces, a new serendipity function space, the mixed- or incomplete-order serendipity space, is proposed for $\boldsymbol{H}(\mathrm{curl})$-conforming functions. The resulting convergence rate of the $\boldsymbol{H}(\mathrm{curl})$-norm error is $\mathcal{O}(h^p)$ for curvilinear meshes when affine refinement is applied. Compared to other competing elements, the advantages are the fewest number of finite element moments, the capability to produce hierarchical basis functions, and the compatibility with the other mixed- or incomplete-order tensor product spaces.

The second contribution of this chapter is a new set of basis functions for the defined $\boldsymbol{H}(\mathrm{curl})$-conforming elements. The higher-order functions are partitioned according to the null space of the curl operator and are orthogonal with respect to the corresponding interpolation operator. Hence, the resulting finite element matrices are highly sparse.

The third contribution is a new method, namely the mixed-order mixed-space finite element, which adaptively applies different finite element spaces. Based on the geometry mapping and mesh refinement, the smallest approximation space can be used on each element that is appropriate for the $\mathcal{O}(h^p)$ rate of convergence. Hence, a significantly smaller number of unknowns is required than for the uniform application of a larger space.

Finally, the fourth contribution is another new method, the iso-serendipity element. It applies the defined serendipity spaces and its basis functions for both the unknown fields and for the interpolation of the curvilinear geometry. The method only allows a special way of mesh refinement, which can improve the geometry representation in each step. Up to the author's knowledge, the iso-serendipity is the first method to utilize a serendipity space for curvilinear meshes while providing an $\mathcal{O}(h^p)$ rate of convergence in $\boldsymbol{H}(\mathrm{curl})$-norm.

6.1 Geometry Description of Hexahedra

Following Definition 2.1 of a general finite element, the domain of the element needs to be declared first. For a general hexahedron, the reference element is a cube with unit length edges over the space of $[0,1]^3 \subset \mathbb{R}^3$. Each global element is produced by the map $\boldsymbol{g}$ of the reference element to the global domain. This mapping is assumed to be smooth and bijective; otherwise, the finite element can lose the conformity or subspace preserving properties. Section 3.1 shows that an appropriate order geometry interpolant $\boldsymbol{g}_k$ may be used instead of the exact mapping because it does not affect the convergence rate. This interpolant is also assumed to be bijective over the reference element. Hence, the conformity and the subspace preserving property also apply to interpolated geometries. First, linear functions of the local coordinates $(\hat{r}_1, \hat{r}_2, \hat{r}_3)$ are introduced that have a support over the reference domain:

$$\begin{aligned} x_0 &= 1 - \hat{r}_1, & y_0 &= 1 - \hat{r}_2, & z_0 &= 1 - \hat{r}_3, \\ x_1 &= \hat{r}_1, & y_1 &= \hat{r}_2, & z_1 &= \hat{r}_3. \end{aligned} \tag{6.1}$$

The values of x_0, y_0, and z_0 are one at the beginning of the corresponding intervals $\hat{r}_i \in [0,1]$ and zero at the end. The values of x_1, y_1, and z_1 are one at the end of the corresponding intervals and zero at the beginning. Hence, the monomials of x_i, y_j, and z_k with the indices $i, j, k \in \{0,1\}$ are well suited for constructing conforming polynomials. With these notations, a first-order geometry mapping reads

$$\boldsymbol{r} = \boldsymbol{g}_1(\hat{\boldsymbol{r}}) = \sum \boldsymbol{r}_{ijk} x_i(\hat{r}_1) y_j(\hat{r}_2) z_k(\hat{r}_3), \tag{6.2}$$

where $\boldsymbol{r}_{ijk}$ is the coordinate vector of the corresponding node. Furthermore, the tensor product space of polynomials is denoted by $\mathbb{Q}_{a,b,c}$ with the order of a in $\hat{r}_1$ variable, b in $\hat{r}_2$ variable, and c in $\hat{r}_3$ variable. The monomials of the geometry mapping then belong to the space of $x_i y_j z_k \in \mathbb{Q}_{1,1,1}$. If there are only four linearly independent node coordinates, the geometry is a parallelepiped, and the geometry mapping is a first-order polynomial $\boldsymbol{g}_1 \in (\mathbb{P}_1)^3$. However, in the general case (i.e., all eight nodes are linearly independent), the geometry mapping is in the space of $\boldsymbol{g}_1 \in (\mathbb{Q}_{1,1,1})^3$. In the following sections, the first, second, and third components of the mapping are denoted by g_x, g_y, and g_z, respectively. One can write these components as polynomials in the form of

$$\begin{aligned} \boldsymbol{g}_1(\hat{\boldsymbol{r}}) &= \begin{bmatrix} g_x(\hat{r}_1, \hat{r}_2, \hat{r}_3) \\ g_y(\hat{r}_1, \hat{r}_2, \hat{r}_3) \\ g_z(\hat{r}_1, \hat{r}_2, \hat{r}_3) \end{bmatrix}, \\ &= \begin{bmatrix} a_1 + b_1\hat{r}_1 + c_1\hat{r}_2 + d_1\hat{r}_3 + e_1\hat{r}_1\hat{r}_2 + f_1\hat{r}_2\hat{r}_3 + g_1\hat{r}_3\hat{r}_1 + h_1\hat{r}_1\hat{r}_2\hat{r}_3 \\ a_2 + b_2\hat{r}_1 + c_2\hat{r}_2 + d_2\hat{r}_3 + e_2\hat{r}_1\hat{r}_2 + f_2\hat{r}_2\hat{r}_3 + g_2\hat{r}_3\hat{r}_1 + h_2\hat{r}_1\hat{r}_2\hat{r}_3 \\ a_3 + b_3\hat{r}_1 + c_3\hat{r}_2 + d_3\hat{r}_3 + e_3\hat{r}_1\hat{r}_2 + f_3\hat{r}_2\hat{r}_3 + g_3\hat{r}_3\hat{r}_1 + h_3\hat{r}_1\hat{r}_2\hat{r}_3 \end{bmatrix}, \end{aligned} \tag{6.3}$$

where $a_i, b_i, \ldots$ are constants dependent on the node coordinates. The reference element of the present configuration is shown in Figure 6.1, where the arrows indicate

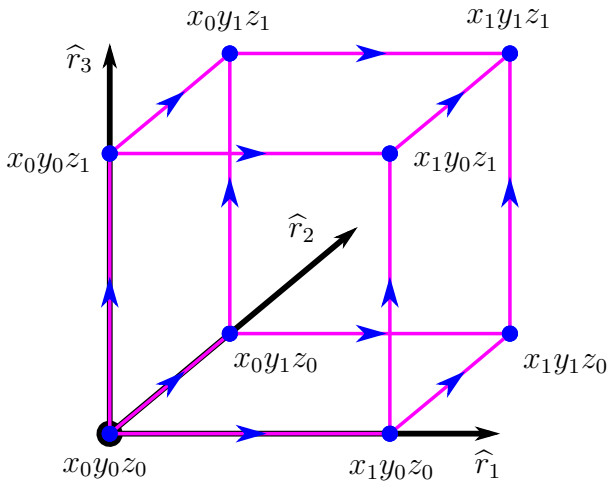


Figure 6.1: The reference element of a general hexahedron.

the orientation of the edges. Each edge is associated with a monomial of x_iy_j, x_iz_k, or y_jz_k, which has a unit value only on the associated edge. Moreover, each face is associated with x_i, y_j, or z_k, which has a unit value only on the associated face. For all of the considered hexahedral elements, the local ordering and orientation of the nodes, edges, and faces are determined by the direction of the reference axis and the order of coordinate components.

The next step is to specify the type of refinement. When the mesh exactly represents the geometry without errors, the volume of the initial mesh ought to be preserved after each uniform refinement step. Hence, there are two natural choices for the overall refinement functions $\boldsymbol{f}_r \in (\mathbb{P}_r)^3$ (see section 3.1.2) for straight-sided geometries. The first choice is an affine refinement

$$\boldsymbol{f}_r \in (\mathbb{P}_1)^3, \quad \text{with } r = 1, \tag{6.4}$$

while the second choice is a first-order tensor product (general piecewise linear) refinement

$$\boldsymbol{f}_r \in (\mathbb{Q}_{1,1,1})^3 \subset (\mathbb{P}_3)^3, \quad \text{with } r = 3. \tag{6.5}$$

The difference between these refinements is the additional freedom in choosing the interior nodes on the faces and on the volume. Hence, the second choice provides freedom for choosing the boundaries of the elements in the interior domain, allowing inhomogeneous mesh refinements. When interpolated geometry mappings are used for the curvilinear geometries, non-affine refinements, such as (6.5), can be used to reduce the geometry error in each refinement step. However, the affine refinement of the mesh with exact mappings is already sufficient for the theoretical convergence rate if the exact geometry description is available.

In the general curvilinear case, non-affine refinements have two main disadvantages:

1. The required order of geometry interpolation $k = rp$ is higher for non-affine refinements ($r > 1$); see Section 3.3. However, for polynomial geometries of order $\widetilde{k} < rp$, $k = \widetilde{k}$ is also sufficient.

2. The approximation space becomes dependent on the geometry mapping and needs to be larger; see Section 3.3.

Hence, the even higher order ($r > 3$) overall refinement functions are less advantageous because they increase the approximation space without providing useful advantages (to the author's knowledge). Thus, only the simplest scenario is considered for non-affine refinements:

- The geometry is straight-sided.
- The mesh represents the geometry without any error ($k = \widetilde{k} = 1$).
- The refinement function is a first-order tensor product polynomial ($r = 3$).

6.2 Serendipity Spaces for Affine Refinements

In the following sections, serendipity spaces are proposed for the approximation of H^1- and $\boldsymbol{H}$(curl)-conforming functions. The definitions are given for the reference element and the moments are specific subcases of the general moments (2.73) and (2.74). Therefore, the various subspaces are preserved for arbitrary curvilinear elements (see Section 2.5.4). Compared to the tensor product spaces, the advantage is the smaller dimension, while the resulting convergence rate $\mathcal{O}(h^p)$ is the same for the affinely refined meshes (see Section 3.3). Another advantage is that the edge-, face-, and volume-associated spaces are subspaces of the corresponding edge-, face-, and volume-associated tensor product spaces. Hence, the mixed usage of serendipity and tensor product spaces is possible within one mesh. This is advantageous in cases where the majority of the elements are located in the interior of the mesh and do not have any common face with any physical boundary. Therefore, affine refinement and serendipity elements can be used on the dominant interior part of the mesh. The resulting number of unknowns is significantly lower than for the uniform application of the tensor product spaces, whereas the convergence rates are the same.

6.2.1 Serendipity H^1-Conforming Element

First, the definition of the moments is provided, which is the subset of the general moments (2.73). Hence, the same form of definition and notations are used.

The space of test functions associated with the edge $\widehat{\mathcal{Q}}_p^{ve}$, face $\widehat{\mathcal{Q}}_p^{vf}$, and volume $\widehat{\mathcal{Q}}_p^{vv}$ are given as

- $\widehat{\mathcal{Q}}_p^{ve} = \{\widehat{q} \in \mathbb{Q}_{a-2} \mid a = p\} = \mathbb{P}_{p-2}(\widehat{E})$,
- $\widehat{\mathcal{Q}}_p^{vf} = \{\widehat{q} \in \mathbb{Q}_{a-2,b-2} \mid a + b = p\} = \mathbb{P}_{p-4}(\widehat{F})$,
- $\widehat{\mathcal{Q}}_p^{vv} = \{\widehat{q} \in \mathbb{Q}_{a-2,b-2,c-2} \mid a + b + c = p\} = \mathbb{P}_{p-6}(\widehat{V})$.

Here, a, b, and c as well as all subscripts are required to be non-negative integers. The superscripts ve, vf, and vv indicate the H^1-conforming edge-, face-, and volume-associated space, respectively. The corresponding moments are

$$\widehat{M}_p^{vn}(\widehat{v}) = \widehat{v}\Big|_{\widehat{N}}, \tag{6.6a}$$

$$\widehat{M}_p^{ve}(\widehat{v}) = \int_{\widehat{E}} \widehat{v}\widehat{q}\,\mathrm{d}\widehat{\boldsymbol{r}}, \qquad \forall \widehat{q} \in \widehat{\mathcal{Q}}_p^{ve}, \tag{6.6b}$$

$$\widehat{M}_p^{vf}(\widehat{v}) = \int_{\widehat{F}} \widehat{v}\widehat{q}\,\mathrm{d}\widehat{\boldsymbol{r}}, \qquad \forall \widehat{q} \in \widehat{\mathcal{Q}}_p^{vf}, \tag{6.6c}$$

$$\widehat{M}_p^{vv}(\widehat{v}) = \int_{\widehat{V}} \widehat{v}\widehat{q}\,\mathrm{d}\widehat{\boldsymbol{r}}, \qquad \forall \widehat{q} \in \widehat{\mathcal{Q}}_p^{vv}. \tag{6.6d}$$

Hence, the number of node-associated moments $\widehat{M}_p^{vn}$ is 8, independently of the order. The number of edge-associated moments $\widehat{M}_p^{ve}$ is

$$12 \cdot \dim(\widehat{\mathcal{Q}}_p^{ve}) = 12(p-1), \qquad p > 0. \tag{6.7}$$

The number of face-associated moments $\widehat{M}_p^{vf}$ is

$$6 \cdot \dim(\widehat{\mathcal{Q}}_p^{vf}) = 6 \sum_{a=2}^{p-2} (p-a-1) = 3(p-3)(p-2), \qquad p > 1. \tag{6.8}$$

The number of volume-associated moments $\widehat{M}_p^{vv}$ is

$$\dim(\widehat{\mathcal{Q}}_p^{vv}) = \sum_{a=2}^{p-4} \sum_{b=2}^{p-a-2} (p-a-b-1) = \frac{(p-5)(p-4)(p-3)}{6}, \qquad p > 2. \tag{6.9}$$

These moments induce an H^1-conforming hexahedral finite element over a space of $\widehat{\mathcal{V}}_{S,p}$ whose dimension is the same as the number of moments. This approximation space is defined by its conforming basis functions, which are associated with nodes, edges, faces, and volume. The first-order node-associated functions $\widehat{v}_{ijk} \in \widehat{\mathcal{V}}_{S,1}^n$ must vanish on all the other nodes (and on the faces that are not adjacent) since they are

induced by the node-associated moments that define a scalar value for each node and are required to be linearly independent. These scalar node-associated values of the moments are smoothly extended to the entire volume using $x_i y_j z_k$ functions, which have a unit value on the corresponding node and are zero on the other nodes. Hence, the general form of the node-associated functions is

$$\widehat{v}_{ijk} = x_i y_j z_k. \tag{6.10}$$

The higher-order edge-associated functions $\widehat{v}_{jk}, \widehat{v}_{ik}, \widehat{v}_{ij} \in \widehat{\mathcal{V}}^e_{S,p}$ must provide compact support on the associated edge and vanish on all other edges. Therefore, the edge-associated functions of $\widehat{\mathcal{V}}^e_{S,p}$ are extensions of the functions of $\widehat{\mathcal{Q}}^{ve}_p$ to the entire volume as

$$\widehat{v}_{jk} = x_0 x_1 y_j z_k \, f_{jk}(x_1), \qquad \text{for } f_{jk}(x_1) \in \widehat{\mathcal{Q}}^{ve}_p \text{ on edge } y_j z_k, \tag{6.11a}$$

$$\widehat{v}_{ik} = y_0 y_1 x_i z_k \, f_{ik}(y_1), \qquad \text{for } f_{ik}(y_1) \in \widehat{\mathcal{Q}}^{ve}_p \text{ on edge } x_i z_k, \tag{6.11b}$$

$$\widehat{v}_{ij} = z_0 z_1 x_i y_j \, f_{ij}(z_1), \qquad \text{for } f_{ij}(z_1) \in \widehat{\mathcal{Q}}^{ve}_p \text{ on edge } x_i y_j. \tag{6.11c}$$

The face-associated functions $\widehat{v}_k, \widehat{v}_j, \widehat{v}_i \in \widehat{\mathcal{V}}^f_{S,p}$ are required to provide compact support on the associated face and vanish on all other faces. Hence, their general form is obtained as the natural extension of the face-associated test functions:

$$\widehat{v}_k = x_0 x_1 y_0 y_1 z_k \, f_k(x_1, y_1), \qquad \text{for } f_k(x_1, y_1) \in \widehat{\mathcal{Q}}^{vf}_p \text{ on face } z_k, \tag{6.12a}$$

$$\widehat{v}_j = x_0 x_1 y_j z_0 z_1 \, f_j(x_1, z_1), \qquad \text{for } f_j(x_1, z_1) \in \widehat{\mathcal{Q}}^{vf}_p \text{ on face } y_j, \tag{6.12b}$$

$$\widehat{v}_i = x_i y_0 y_1 z_0 z_1 \, f_i(y_1, z_1), \qquad \text{for } f_i(y_1, z_1) \in \widehat{\mathcal{Q}}^{vf}_p \text{ on face } x_i. \tag{6.12c}$$

Figure 6.1 illustrates the associated geometry entities of the node-type functions, and Figure 6.2 illustrates the edge- and face-type functions. Finally, the volume-associated functions $\widehat{v} \in \widehat{\mathcal{V}}^v_{S,p}$ are considered, which must vanish on all boundaries. Hence, their general form is

$$\widehat{v} = x_0 x_1 y_0 y_1 z_0 z_1 \, f(x_1, y_1, z_1), \qquad f(x_1, y_1, z_1) \in \widehat{\mathcal{Q}}^{vv}_p. \tag{6.13}$$

It is easy to conclude that the dimension of the edge-, face- and volume-associated space of test functions is the same as the corresponding approximation spaces

$$\dim(\widehat{\mathcal{V}}^e_{S,p}) = 12 \cdot \dim(\widehat{\mathcal{Q}}^{ve}_p), \tag{6.14a}$$

$$\dim(\widehat{\mathcal{V}}^f_{S,p}) = 6 \cdot \dim(\widehat{\mathcal{Q}}^{vf}_p), \tag{6.14b}$$

$$\dim(\widehat{\mathcal{V}}^v_{S,p}) = 1 \cdot \dim(\widehat{\mathcal{Q}}^{vv}_p). \tag{6.14c}$$

The resulting finite element space $\widehat{\mathcal{V}}_{S,p}$ of order p is given as the direct sum of the node-, edge-, face-, and volume-associated subspaces

$$\widehat{\mathcal{V}}_{S,p} = \widehat{\mathcal{V}}^n_{S,1} \oplus \widehat{\mathcal{V}}^e_{S,p} \oplus \widehat{\mathcal{V}}^f_{S,p} \oplus \widehat{\mathcal{V}}^v_{S,p}, \tag{6.15}$$

which yields the following important property:

$$\mathbb{P}_p \subset \widehat{\mathcal{V}}_{S,p} \subset \mathbb{Q}_{p,p,p}. \tag{6.16}$$

Using the invariance of the moments, the global approximation space $\mathcal{V}_{S,p}$ is obtained by the Piola transformation of the local space on each element. In the case of exact geometry mappings and affine refinements, the resulting convergence rates are

$$\left\|v - \pi^v_{p,h} v\right\|_{L^2(\Omega)} \leq C h^{p+1}, \tag{6.17a}$$

$$\left\|v - \pi^v_{p,h} v\right\|_{H^1(\Omega)} \leq C h^{p}, \tag{6.17b}$$

because $\widehat{\mathcal{V}}_{S,p}$ has a complete p order subspace (see Table 3.1).

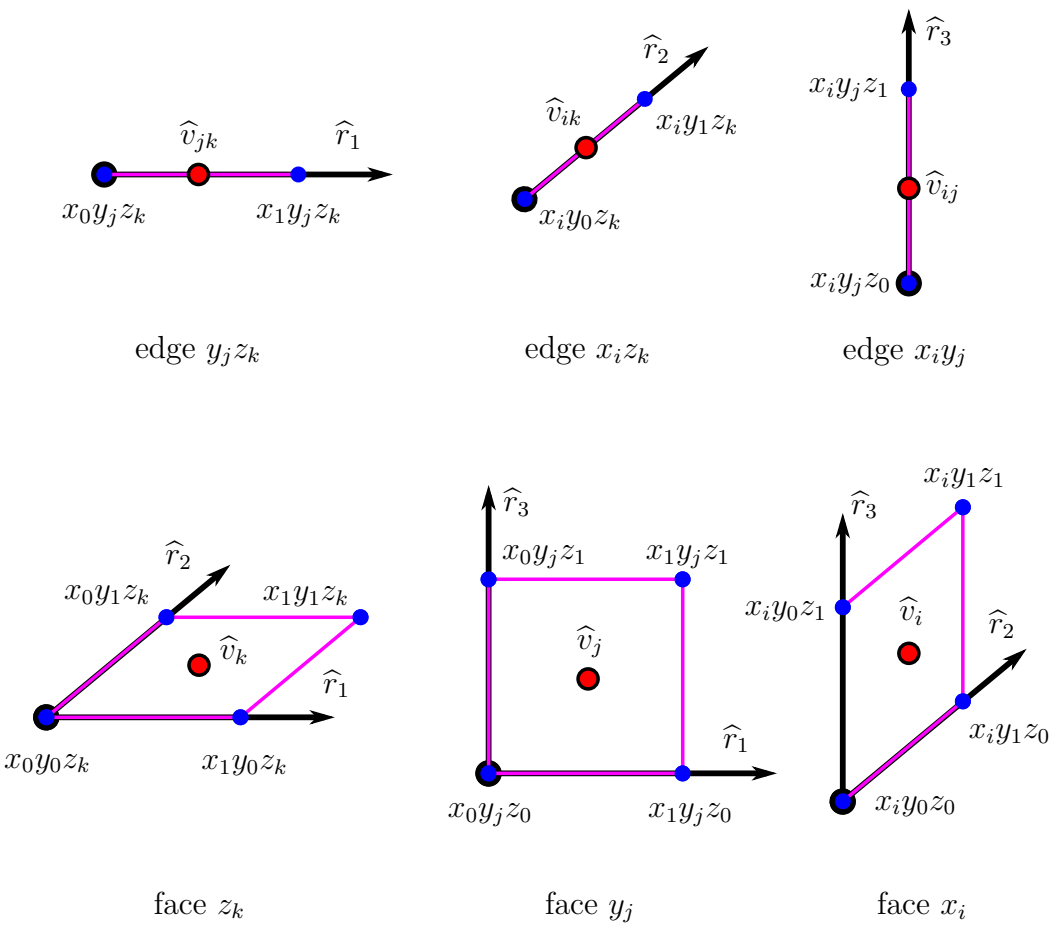


Figure 6.2: The associated edges and faces of the edge- and face-type functions.

6.2.2 Serendipity $\boldsymbol{H}(\mathrm{curl})$-Conforming Element

This section defines the complete-order serendipity element whose local approximation space $\widehat{\mathcal{W}}_{S,p}$ contains $(\mathbb{P}_p)^3$ as a subspace. The space of test functions associated with the edge $\widehat{\mathcal{Q}}_p^{we}$, face $\widehat{\mathcal{Q}}_p^{wf}$, and volume $\widehat{\mathcal{Q}}_p^{wv}$ are defined as

- $\widehat{\mathcal{Q}}_p^{we} = \{\boldsymbol{q} \in \widehat{\boldsymbol{e}}_1 \mathbb{Q}_a \mid a = p\} = \widehat{\boldsymbol{e}}_1 \mathbb{P}_p(\widehat{E})$,
- $\widehat{\mathcal{Q}}_p^{wf} = \{\boldsymbol{q} \in (\widehat{\boldsymbol{e}}_1 \mathbb{Q}_{a,b-2},\, \widehat{\boldsymbol{e}}_2 \mathbb{Q}_{a-2,b}) \mid a + b = p\} = (\mathbb{P}_{p-2}(\widehat{F}))^2$,
- $\widehat{\mathcal{Q}}_p^{wv} = \{\boldsymbol{q} \in (\widehat{\boldsymbol{e}}_1 \mathbb{Q}_{a,b-2,c-2},\, \widehat{\boldsymbol{e}}_2 \mathbb{Q}_{a-2,b,c-2},\, \widehat{\boldsymbol{e}}_3 \mathbb{Q}_{a-2,b-2,c}) \mid a + b + c = p\}$,
  $= (\mathbb{P}_{p-4}(\widehat{V}))^3$.

Here, a, b, and c as well as each subscript are required to be non-negative integers, and $\widehat{\boldsymbol{e}}_1$, $\widehat{\boldsymbol{e}}_2$, and $\widehat{\boldsymbol{e}}_3$ are the standard unit vectors. The superscripts we, wf, and wv indicate the $\boldsymbol{H}(\mathrm{curl})$-conforming edge-, face-, and volume-associated space, respectively. The corresponding moments associated with the edge $\widehat{M}_p^{we}$, face $\widehat{M}_p^{wf}$, and volume $\widehat{M}_p^{wv}$ are defined as

$$\widehat{M}_p^{we}(\widehat{\boldsymbol{w}}) = \int_{\widehat{E}} \widehat{\boldsymbol{w}} \cdot \widehat{\boldsymbol{q}} \,\mathrm{d}\widehat{\boldsymbol{r}}, \qquad \forall \widehat{\boldsymbol{q}} \in \widehat{\mathcal{Q}}_p^{we}, \tag{6.18a}$$

$$\widehat{M}_p^{wf}(\widehat{\boldsymbol{w}}) = \int_{\widehat{F}} \widehat{\boldsymbol{w}} \cdot \widehat{\boldsymbol{q}} \,\mathrm{d}\widehat{\boldsymbol{r}}, \qquad \forall \widehat{\boldsymbol{q}} \in \widehat{\mathcal{Q}}_p^{wf}, \tag{6.18b}$$

$$\widehat{M}_p^{wv}(\widehat{\boldsymbol{w}}) = \int_{\widehat{V}} \widehat{\boldsymbol{w}} \cdot \widehat{\boldsymbol{q}} \,\mathrm{d}\widehat{\boldsymbol{r}}, \qquad \forall \widehat{\boldsymbol{q}} \in \widehat{\mathcal{Q}}_p^{wv}, \tag{6.18c}$$

with the dimensions of

$$12 \cdot \dim(\widehat{\mathcal{Q}}_p^{we}) = 12(p+1), \tag{6.19a}$$

$$6 \cdot \dim(\widehat{\mathcal{Q}}_p^{wf}) = 6 \cdot 2 \sum_{a=0}^{p-2} (p-a-1) = 6(p-1)p, \tag{6.19b}$$

$$1 \cdot \dim(\widehat{\mathcal{Q}}_p^{wv}) = 1 \cdot 3 \sum_{a=0}^{p-4} \sum_{b=2}^{p-a-2} (p-a-b-1) = \frac{(p-1)(p-2)(p-3)}{2}, \qquad p > 0. \tag{6.19c}$$

These moments induce an $\boldsymbol{H}(\mathrm{curl})$-conforming hexahedral finite element over a space $\widehat{\mathcal{W}}_{S,p}$. The edge-associated functions $\widehat{\boldsymbol{w}}_{jk}, \widehat{\boldsymbol{w}}_{ik}, \widehat{\boldsymbol{w}}_{ij} \in \widehat{\mathcal{W}}_{S,p}^{e}$ of order p must have a zero tangential component on all other edges. Here, the subscripts indicate the associated edges as described by Figure 6.3. Therefore, any edge-associated function may be

written as a test function of the edge, which is extended to the entire volume with the appropriate multiplicative factor of local coordinates:

$$\widehat{\boldsymbol{w}}_{jk} = y_j z_k \, f_{jk}(x_1)\widehat{\boldsymbol{e}}_1, \qquad \text{for } f_{jk}(x_1)\widehat{\boldsymbol{e}}_1 \in \widehat{\mathcal{Q}}_p^{we} \text{ on edge } y_j z_k, \tag{6.20a}$$

$$\widehat{\boldsymbol{w}}_{ik} = x_i z_k \, f_{ik}(y_1)\widehat{\boldsymbol{e}}_2, \qquad \text{for } f_{ik}(y_1)\widehat{\boldsymbol{e}}_2 \in \widehat{\mathcal{Q}}_p^{we} \text{ on edge } x_i z_k, \tag{6.20b}$$

$$\widehat{\boldsymbol{w}}_{ij} = x_i y_j \, f_{ij}(z_1)\widehat{\boldsymbol{e}}_3, \qquad \text{for } f_{ij}(z_1)\widehat{\boldsymbol{e}}_3 \in \widehat{\mathcal{Q}}_p^{we} \text{ on edge } x_i y_j. \tag{6.20c}$$

The face-associated functions of order p are denoted by $\widehat{\boldsymbol{w}}_k^x, \widehat{\boldsymbol{w}}_k^y, \widehat{\boldsymbol{w}}_j^x, \widehat{\boldsymbol{w}}_j^z, \widehat{\boldsymbol{w}}_i^y, \widehat{\boldsymbol{w}}_i^z \in \widehat{\mathcal{W}}_{S,p}^f$. Here, the subscripts indicate the associated faces, and the superscripts indicate the direction as described by Figure 6.3. These functions are required to have tangential compact support over the associated face and a zero tangential component on all other faces. Hence, the general forms of such functions are

$$\widehat{\boldsymbol{w}}_k^x = y_0 y_1 z_k \, f_k^x(x_1, y_1)\widehat{\boldsymbol{e}}_1, \qquad \text{for } f_k^x(x_1, y_1)\widehat{\boldsymbol{e}}_1 \in \widehat{\mathcal{Q}}_p^{wf} \text{ on face } z_k, \tag{6.21a}$$

$$\widehat{\boldsymbol{w}}_k^y = x_0 x_1 z_k \, f_k^y(x_1, y_1)\widehat{\boldsymbol{e}}_2, \qquad \text{for } f_k^y(x_1, y_1)\widehat{\boldsymbol{e}}_2 \in \widehat{\mathcal{Q}}_p^{wf} \text{ on face } z_k, \tag{6.21b}$$

$$\widehat{\boldsymbol{w}}_j^x = z_0 z_1 y_j \, f_j^x(x_1, z_1)\widehat{\boldsymbol{e}}_1, \qquad \text{for } f_j^x(x_1, z_1)\widehat{\boldsymbol{e}}_1 \in \widehat{\mathcal{Q}}_p^{wf} \text{ on face } y_j, \tag{6.21c}$$

$$\widehat{\boldsymbol{w}}_j^z = x_0 x_1 y_j \, f_j^z(x_1, z_1)\widehat{\boldsymbol{e}}_3, \qquad \text{for } f_j^z(x_1, z_1)\widehat{\boldsymbol{e}}_3 \in \widehat{\mathcal{Q}}_p^{wf} \text{ on face } y_j, \tag{6.21d}$$

$$\widehat{\boldsymbol{w}}_i^y = z_0 z_1 x_i \, f_i^y(y_1, z_1)\widehat{\boldsymbol{e}}_2, \qquad \text{for } f_i^y(y_1, z_1)\widehat{\boldsymbol{e}}_2 \in \widehat{\mathcal{Q}}_p^{wf} \text{ on face } x_i, \tag{6.21e}$$

$$\widehat{\boldsymbol{w}}_i^z = y_0 y_1 x_i \, f_i^z(y_1, z_1)\widehat{\boldsymbol{e}}_3, \qquad \text{for } f_i^z(y_1, z_1)\widehat{\boldsymbol{e}}_3 \in \widehat{\mathcal{Q}}_p^{wf} \text{ on face } x_i. \tag{6.21f}$$

The volume-associated functions of order p are denoted by $\widehat{\boldsymbol{w}}^x, \widehat{\boldsymbol{w}}^y, \widehat{\boldsymbol{w}}^z \in \widehat{\mathcal{W}}_{S,p}^v$. These functions must provide tangential compact support over the volume of the element and thus have a zero tangential component on all faces. Their general forms are

$$\widehat{\boldsymbol{w}}^x = y_0 y_1 z_0 z_1 \, f^x(x_1, y_1, z_1)\widehat{\boldsymbol{e}}_1, \qquad f_x(x_1, y_1, z_1)\widehat{\boldsymbol{e}}_1 \in \widehat{\mathcal{Q}}_p^{wv}, \tag{6.22a}$$

$$\widehat{\boldsymbol{w}}^y = z_0 z_1 x_0 x_1 \, f^y(x_1, y_1, z_1)\widehat{\boldsymbol{e}}_2, \qquad f_y(x_1, y_1, z_1)\widehat{\boldsymbol{e}}_2 \in \widehat{\mathcal{Q}}_p^{wv}, \tag{6.22b}$$

$$\widehat{\boldsymbol{w}}^z = x_0 x_1 y_0 y_1 \, f^z(x_1, y_1, z_1)\widehat{\boldsymbol{e}}_3, \qquad f_z(x_1, y_1, z_1)\widehat{\boldsymbol{e}}_3 \in \widehat{\mathcal{Q}}_p^{wv}. \tag{6.22c}$$

Therefore, the dimension of the edge-, face- and volume-associated approximation space is the same as the space of test functions

$$\dim(\widehat{\mathcal{W}}_{S,p}^{we}) = 12 \cdot \dim(\widehat{\mathcal{Q}}_p^{we}), \tag{6.23a}$$

$$\dim(\widehat{\mathcal{W}}_{S,p}^{wf}) = 6 \cdot \dim(\widehat{\mathcal{Q}}_p^{wf}), \tag{6.23b}$$

$$\dim(\widehat{\mathcal{W}}_{S,p}^{wv}) = 1 \cdot \dim(\widehat{\mathcal{Q}}_p^{wv}). \tag{6.23c}$$

The local approximation space $\widehat{\mathcal{W}}_{S,p}$ is the direct sum of the edge-, face-, and volume-associated subspaces:

$$\widehat{\mathcal{W}}_{S,p} = \widehat{\mathcal{W}}_{S,p}^e \oplus \widehat{\mathcal{W}}_{S,p}^f \oplus \widehat{\mathcal{W}}_{S,p}^v. \tag{6.24}$$

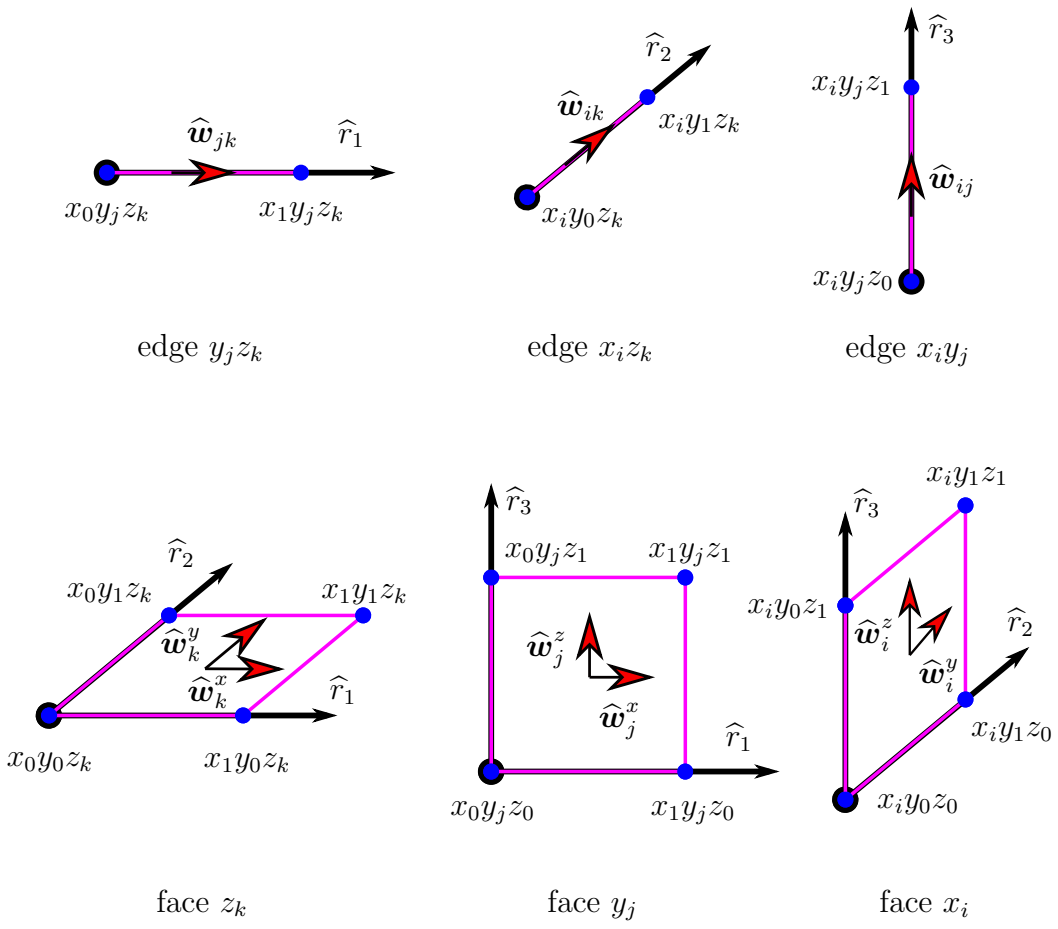


Figure 6.3: Associated edges and faces of the edge- and face-type functions.

The relations between these spaces and the tensor product spaces are

$$\widehat{\mathcal{W}}^e_{S,p} \subset (\mathbb{Q}_{p,p,p})^3, \tag{6.25a}$$
$$\widehat{\mathcal{W}}^f_{S,p} \subset (\mathbb{Q}_{p,p,p})^3, \tag{6.25b}$$
$$\widehat{\mathcal{W}}^v_{S,p} \subset (\mathbb{P}_p)^3 \subset (\mathbb{Q}_{p,p,p})^3, \tag{6.25c}$$
$$(\mathbb{P}_p)^3 \subset \widehat{\mathcal{W}}_{S,p} \subset (\mathbb{Q}_{p,p,p})^3. \tag{6.25d}$$

The global approximation space $\mathcal{W}_{S,p}$ of order p is obtained by the transformation of the local space to the global space via Piola transformation. The resulting asymptotic rate of convergence is

$$\left\|\boldsymbol{w} - \pi^w_{p,h}\boldsymbol{w}\right\|_{\boldsymbol{L}^2(\Omega)} \leq Ch^{p+1}, \tag{6.26a}$$
$$\left\|\boldsymbol{w} - \pi^w_{p,h}\boldsymbol{w}\right\|_{\boldsymbol{H}(\mathrm{curl},\Omega)} \leq Ch^{p}, \tag{6.26b}$$

for affine refinements and exact geometry mappings; see Table 3.2.

6.2.3 Mixed-Order $\boldsymbol{H}(\mathrm{curl})$-Conforming Serendipity Element

This section introduces the proposed new finite element of the thesis, namely the mixed-order (or incomplete order) serendipity element, for the approximation of $\boldsymbol{H}(\mathrm{curl})$-conforming functions. The main advantage is the fewer dimension compared to the complete-order serendipity space $\widehat{\mathcal{W}}_{S,p}$, while they both yield the same rate of convergence in $\boldsymbol{H}(\mathrm{curl},\Omega)$-norm. The term "mixed-order" or "incomplete-order" refers to the fact that the new approximation space $\widehat{\mathcal{W}}_{SI,p}$ includes a polynomial space $(\mathbb{P}_{p-1})^3 \subset \widehat{\mathcal{W}}_{SI,p}$, which yields the convergence rate of

$$\left\|\boldsymbol{w} - \pi^{w}_{p,h}\boldsymbol{w}\right\|_{\boldsymbol{L}^2(\Omega)} \leq Ch^p. \tag{6.27}$$

However, it is incomplete for the pth order space $(\mathbb{P}_p)^3 \not\subset \widehat{\mathcal{W}}_{SI,p}$ since the rate of convergence

$$\left\|\nabla\times\boldsymbol{w} - \nabla\times\pi^{w}_{p,h}\boldsymbol{w}\right\|_{\boldsymbol{L}^2(\Omega)} \leq Ch^p \tag{6.28}$$

only requires $\widehat{\nabla}\times(\mathbb{P}_p)^3 \subset \widehat{\nabla}\times\widehat{\mathcal{W}}_{SI,p}$ (see Table 3.2). Therefore, it is permissible to subtract the highest incremental order gradient space $\widehat{\nabla}(\widehat{\mathcal{V}}_{S,p+1}\setminus\widehat{\mathcal{V}}_{S,p})$ from $\widehat{\mathcal{W}}_{S,p}$ because it belongs to the null space of the $\widehat{\nabla}\times$-operator, which is not required for the $\boldsymbol{L}^2$-norm convergence. Such induction of the mixed-order space is necessary for the generation of hierarchical basis functions, which can be decomposed into a set of null space and complementary space functions. Hence, the subscript SI is an abbreviation for the serendipity space, which is an incomplete pth-order polynomial space. In the case of tetrahedral elements, the procedure is straightforward. The highest incremental order space is the space of homogeneous polynomials $\mathbb{P}_{p+1}\setminus\mathbb{P}_p = \widetilde{\mathbb{P}}_{p+1}$, whose gradient is a subspace of the complete-order polynomial space $\widehat{\nabla}\widetilde{\mathbb{P}}_{p+1}\subset(\mathbb{P}_p)^3$. Hence, the mixed-order space for tetrahedral elements is identical to the rotational space of Nédélec (2.109a):

$$R_p = (\mathbb{P}_p)^3 \setminus \widehat{\nabla}\widetilde{\mathbb{P}}_{p+1}. \tag{6.29}$$

However, the procedure becomes more complicated for hexahedral elements where the highest incremental order gradient space is not a subspace of the complete-order space, i.e.,

$$\widehat{\nabla}(\widehat{\mathcal{V}}_{S,p+1}\setminus\widehat{\mathcal{V}}_{S,p}) \not\subset \widehat{\mathcal{W}}_{S,p}. \tag{6.30}$$

The proposed solution eliminates those test functions of the complete-order finite element moments, which correspond to the incremental order gradient space. Therefore, the dimension of the finite element is reduced without affecting the dimension of the range space.

Let the derivation begin with an edge-associated function y_jz_k (6.11a) which yields a single incremental order function

$$\widetilde{v}_{jk} = x_0x_1y_jz_kx_1^{p-1} \in \widehat{\mathcal{V}}^e_{S,p+1}\setminus\widehat{\mathcal{V}}^e_{S,p} \qquad \text{on edge } y_jz_k. \tag{6.31}$$

It can be seen that the gradient function

$$\widehat{\nabla}\widetilde{v}_{jk} = y_j z_k \frac{\partial x_0 x_1^p}{\partial \widehat{r}_1}\widehat{\boldsymbol{e}}_1 + z_k x_0 x_1^p \frac{\partial y_j}{\partial \widehat{r}_2}\widehat{\boldsymbol{e}}_2 + y_j x_0 x_1^p \frac{\partial z_k}{\partial \widehat{r}_3}\widehat{\boldsymbol{e}}_3 \tag{6.32}$$

contains the higher-order term of $x_0 x_1^p = (1-\widehat{r}_1)\widehat{r}_1^p \in \mathbb{Q}_{p+1,0,0}$ in the second and third components. Here, $\widehat{\boldsymbol{e}}_1$ is the tangential vector, whereas $\widehat{\boldsymbol{e}}_2$ and $\widehat{\boldsymbol{e}}_3$ are the normal vectors on the edge $y_j z_k$. The consequence is that $\widehat{\nabla}\widetilde{v}_{jk} \notin (\mathbb{Q}_{p,p,p})^3$ and $\widehat{\nabla}\widetilde{v}_{jk} \notin \widehat{\mathcal{W}}_{S,p}$. However, these higher-order terms do not contribute to the finite element moments because they are only present in the normal components, whereas the moments are assigned to the tangential component. These other components can be viewed as necessary byproducts of the polynomial basis construction (6.11), ensuring the conformity of the H^1-conforming basis functions. Therefore, these non-tangential components do not contribute to the rate of convergence and are not required for the conformity of $\boldsymbol{H}(\mathrm{curl})$-conforming functions. From these considerations, the corresponding test function $\widetilde{\boldsymbol{q}}^e_{p+1}$ of $\widehat{\nabla}\widetilde{v}_{jk}$ is obtained as a tangential trace and the inverse of the test function extension (6.20a), i.e.,

$$\begin{aligned}\widetilde{\boldsymbol{q}}^e_{p+1} &= \frac{1}{y_j z_k}(\widehat{\boldsymbol{e}}_1 \cdot \widehat{\nabla}\widetilde{v}_{jk})\widehat{\boldsymbol{e}}_1 = \frac{\partial x_0 x_1^p}{\partial \widehat{r}_1}\widehat{\boldsymbol{e}}_1, \\ &= \frac{\partial (1-\widehat{r}_1)\widehat{r}_1^p}{\partial \widehat{r}_1}\widehat{\boldsymbol{e}}_1 = (p\widehat{r}_1^{p-1} - (p+1)\widehat{r}_1^p)\widehat{\boldsymbol{e}}_1, \qquad \text{on edge } y_j z_k. \end{aligned} \tag{6.33}$$

The results for the other edges are analogous. Therefore, the corresponding space of test functions $\mathcal{G}^e_p$ for the incremental order gradient space $\widehat{\nabla}(\widehat{\mathcal{V}}^e_{S,p+1} \setminus \widehat{\mathcal{V}}^e_{S,p})$, which is associated with the edges, is defined as

$$\mathcal{G}^e_p = \{\widetilde{\boldsymbol{q}}^e_{p+1}\}. \tag{6.34}$$

The next step is to determine the space of test functions $\widehat{\mathcal{Q}}^{we}_p$ for the edge associated moments of the mixed-order serendipity elements. The sufficient oversized trial space is

$$\widehat{\boldsymbol{e}}_1 \mathbb{P}_p(\widehat{E}) = \{\widehat{\boldsymbol{e}}_1 \widehat{r}_1^p\} \oplus \widehat{\boldsymbol{e}}_1 \mathbb{P}_{p-1}(\widehat{E}), \qquad \text{for } \widehat{\mathcal{Q}}^{we}_p, \tag{6.35}$$

which is identical to the case of complete-order serendipity element. Based on the estimation (6.26), it is clear that the one order less space $\widehat{\boldsymbol{e}}_1 \mathbb{P}_{p-1}(\widehat{E}) \subseteq \widehat{\mathcal{Q}}^{we}_p$ is a necessary condition of the convergence. Moreover, it is known that the moments of the space $\mathcal{G}^e_p$ do not contribute to the convergence. Combining these results, the space of edge-associated test functions is defined as

$$\widehat{\mathcal{Q}}^{we}_p = \Big(\{\widehat{\boldsymbol{e}}_1 \widehat{r}_1^p\} \setminus (\mathcal{G}^e_p \oplus \widehat{\boldsymbol{e}}_1 \mathbb{P}_{p-1}(\widehat{E}))\Big) \oplus \widehat{\boldsymbol{e}}_1 \mathbb{P}_{p-1}(\widehat{E}) = \widehat{\boldsymbol{e}}_1 \mathbb{P}_{p-1}(\widehat{E}). \tag{6.36}$$

Let us continue with a z_k face-associated function (6.12a), which yields the following incremental order function:

$$\widehat{\nabla}\widetilde{v}_k = \widehat{\nabla}(x_0 x_1 y_0 y_1 z_k x_1^a y_1^b) \in \widehat{\nabla}(\widehat{\mathcal{V}}^f_{S,p+1} \setminus \widehat{\mathcal{V}}^f_{S,p}),\ \forall\,(a+b) = p-3,\ \text{ on face } z_k. \tag{6.37}$$

Similarly to the edge-associated functions, the corresponding test functions of the face-associated moments can be obtained as a tangential trace and the inverse of the test function extensions (6.21a) and (6.21b), i.e.,

$$\begin{aligned}\widetilde{\boldsymbol{q}}^{f,ab}_{p+1} &= \begin{bmatrix} \frac{1}{y_0 y_1 z_k} & 0 & 0 \\ 0 & \frac{1}{x_0 x_1 z_k} & 0 \end{bmatrix} \widehat{\nabla}\widetilde{v}_k = \frac{\partial x_0 x_1^{a+1} y_1^b}{\partial \widehat{r}_1}\widehat{\boldsymbol{e}}_1 + \frac{\partial y_0 y_1^{b+1} x_1^a}{\partial \widehat{r}_2}\widehat{\boldsymbol{e}}_2, \\ &= \widehat{r}_2^b \frac{\partial (1-\widehat{r}_1)\widehat{r}_1^{a+1}}{\partial \widehat{r}_1}\widehat{\boldsymbol{e}}_1 + \widehat{r}_1^a \frac{\partial (1-\widehat{r}_2)\widehat{r}_2^{b+1}}{\partial \widehat{r}_2}\widehat{\boldsymbol{e}}_2, \\ &= \begin{bmatrix} \widehat{r}_2^b((a+1)\widehat{r}_1^a - (a+2)\widehat{r}_1^{a+1}) \\ \\ \widehat{r}_1^a((b+1)\widehat{r}_2^b - (b+2)\widehat{r}_2^{b+1}) \end{bmatrix}, \quad \forall\,(a+b) = p-3 \quad \text{on face } z_k. \end{aligned} \tag{6.38}$$

The substitution of the condition $(a+b) = p-3$ into the three different cases $a > b \geq 0$, $b > a \geq 0$, and $a = b$ leads to the relations of $(p-3) \geq a > (p-3)/2$, $(p-3) \geq b > (p-3)/2$, and $a = b = (p-3)/2$, respectively. Note that the last case is only present for the odd values of p, due to the integers a and b. The space of test functions $\mathcal{G}_p^f$ for the incremental order gradient functions is defined as the direct sum of the three different cases

$$\begin{aligned}\mathcal{G}_p^f =& \mathcal{G}_p^{a>b} \oplus \mathcal{G}_p^{a<b} \oplus \mathcal{G}_p^{a=b}, \\ =& \{\widetilde{\boldsymbol{q}}^{f,ab}_{p+1} \mid \quad \forall\,(p-3) \geq a > (p-3)/2,\, b = p-3-a\} \oplus \\ & \{\widetilde{\boldsymbol{q}}^{f,ab}_{p+1} \mid \quad \forall\,(p-3) \geq b > (p-3)/2,\, b = p-3-a\} \oplus \\ & \{\widetilde{\boldsymbol{q}}^{f,ab}_{p+1} \mid \quad a = b = (p-3)/2 \quad \text{for odd } p\}. \end{aligned} \tag{6.39}$$

The next step is the determination of the test function space $\widehat{\mathcal{Q}}_p^{wf}$ for the face-associated moments. The procedure is the same as for the edge-associated space, where the unnecessary terms are removed from the case of the complete-order space. The sufficient oversized trial space reads

$$(\mathbb{P}_{p-2}(\widehat{F}))^2 = (\mathbb{P}_{p-3}(\widehat{F}))^2 \oplus \mathcal{P}_p^{a\leq b} \oplus \mathcal{P}_p^{a>b}, \qquad \text{for } \widehat{\mathcal{Q}}_p^{wf}, \tag{6.40}$$

where the highest order subspaces $\mathcal{P}_p^{a\leq b}$ and $\mathcal{P}_p^{a>b}$ are defined as

$$\mathcal{P}_p^{a\leq b} = \{\widehat{\boldsymbol{e}}_i \widehat{r}_i^a \widehat{r}_j^b \mid \forall (a+b) = p-2 \text{ and } a \leq b\}, \tag{6.41a}$$

$$\begin{aligned}\mathcal{P}_p^{a>b} =& \{\widehat{\boldsymbol{e}}_i \widehat{r}_i^a \widehat{r}_j^b \mid \forall (a+b) = p-2 \text{ and } a > b\}, \\ =& \{\widehat{\boldsymbol{e}}_i \widehat{r}_i^{a+1} \widehat{r}_j^b \mid \forall (a+b) = p-3 \text{ and } a \geq b\}, \\ =& \{\widehat{\boldsymbol{e}}_i \widehat{r}_i^{a+1} \widehat{r}_j^b \mid p-3 \geq a \geq (p-3)/2,\, b = p-3-a\}, \\ =& \{\widehat{\boldsymbol{e}}_1 \widehat{r}_1^{a+1} \widehat{r}_2^b \mid p-3 \geq a \geq (p-3)/2,\, b = p-3-a\} \\ & \oplus \{\widehat{\boldsymbol{e}}_2 \widehat{r}_1^a \widehat{r}_2^{b+1} \mid p-3 \geq b \geq (p-3)/2,\, b = p-3-a\}. \end{aligned} \tag{6.41b}$$

It is known from the error estimation (6.26) of the complete-order serendipity element that $(\mathbb{P}_{p-3}(\widehat{F}))^2$ is a necessary condition of the convergence. Moreover, we have

also considered that the curl operator requires the derivative of a vector components with respect to the other variables, i.e., $\widehat{\boldsymbol{e}}_i\widehat{r}_j$ for $i \neq j$. Hence, any $\widehat{\boldsymbol{e}}_i\widehat{r}_i^a\widehat{r}_j^b$ monomial, for arbitrary indices $i, j \in \{1, 2\}$ and $a \leq b$, corresponds to a higher-order range space than its counterpart $\widehat{\boldsymbol{e}}_i\widehat{r}_i^a\widehat{r}_j^b$ for $a > b$. Based on these considerations, it is reasonable to preserve the subspace $\mathcal{P}_p^{a\leq b}$ and only eliminate functions from $\mathcal{P}_p^{a>b}$ via $\mathcal{G}_p^f$. This constraint enforces the inclusion of the serendipity space functions into the lowest possible order tensor product space, since always the monomials with the highest powers are eliminated. Thus, the compatibility of the functions within the different function spaces can be ensured. Consequently, the definition for the space of face-associated test functions is

$$\widehat{\mathcal{Q}}_p^{wf} = (\mathbb{P}_{p-3}(\widehat{F}))^2 \oplus \mathcal{P}_p^{a\leq b} \oplus \left(\mathcal{P}_p^{a>b} \setminus (\mathcal{G}_p^f \oplus \mathcal{P}_p^{a\leq b} \oplus (\mathbb{P}_{p-3}(\widehat{F}))^2)\right). \tag{6.42}$$

For even p orders, all functions of $\mathcal{P}_p^{a>b}$ can be eliminated by the functions of $\mathcal{G}_p^f = \mathcal{G}_p^{a>b} \oplus \mathcal{G}_p^{a<b}$, and the remaining space becomes $\mathcal{P}_p^{a\leq b} \oplus (\mathbb{P}_{p-3}(\widehat{F}))^2$. However, there is only one function in $\mathcal{G}_p^{a=b}$ for odd p orders which cannot eliminate both the corresponding functions $\widehat{\boldsymbol{e}}_1\widehat{r}_1^{\frac{p-3}{2}+1}\widehat{r}_2^{\frac{p-3}{2}} \in \mathcal{P}_p^{a>b}$ and $\widehat{\boldsymbol{e}}_2\widehat{r}_1^{\frac{p-3}{2}}\widehat{r}_2^{\frac{p-3}{2}+1} \in \mathcal{P}_p^{a>b}$. Hence, an additional function, and the corresponding space

$$\mathcal{R}_p^{a=b} = \left\{ \widehat{\boldsymbol{e}}_1\widehat{r}_1^{a+1}\widehat{r}_2^a - \widehat{\boldsymbol{e}}_2\widehat{r}_1^a\widehat{r}_2^{a+1} \;\middle|\; \text{for } a = \left\lfloor \frac{p-3}{2} \right\rfloor = \frac{p-3}{2} \right\}, \tag{6.43}$$

needs to be preserved for the odd orders. Therefore, the final expression for the space of face-associated test functions is

$$\widehat{\mathcal{Q}}_p^{wf} = (\mathbb{P}_{p-3}(\widehat{F}))^2 \oplus \mathcal{P}_p^{a\leq b} \oplus \mathcal{R}_p^{a=b}. \tag{6.44}$$

The results for the other faces are analogous.

Finally, a volume-associated function (6.13) is considered, which yields the following incremental order function:

$$\widehat{\nabla}\widetilde{v} = \widehat{\nabla}(x_0x_1y_0y_1z_0z_1x_1^ay_1^bz_1^c) \in \widehat{\nabla}(\widehat{\mathcal{V}}_{S,p+1}^v \setminus \widehat{\mathcal{V}}_{S,p}^v), \quad \forall\,(a+b+c) = p-5. \tag{6.45}$$

The corresponding test function of the volume-associated moments can be obtained as the inverse of the test function extensions (6.22), i.e.,

$$\begin{aligned}
\widetilde{\boldsymbol{q}}_{p+1}^{v,abc} &= \begin{bmatrix} \dfrac{1}{y_0y_1z_0z_1} & 0 & 0 \\ 0 & \dfrac{1}{x_0x_1z_0z_1} & 0 \\ 0 & 0 & \dfrac{1}{x_0x_1y_0y_1} \end{bmatrix} \widehat{\nabla}\widetilde{v} = \begin{bmatrix} \widehat{r}_3^c\widehat{r}_2^b((a+1)\widehat{r}_1^a - (a+2)\widehat{r}_1^{a+1}) \\ \widehat{r}_3^c\widehat{r}_1^a((b+1)\widehat{r}_2^b - (b+2)\widehat{r}_2^{b+1}) \\ \widehat{r}_1^a\widehat{r}_2^b((c+1)\widehat{r}_3^c - (c+2)\widehat{r}_3^{c+1}) \end{bmatrix}, \\
&= \widehat{r}_1^a\widehat{r}_2^b\widehat{r}_3^c \begin{bmatrix} a+1 \\ b+1 \\ c+1 \end{bmatrix} - \begin{bmatrix} (a+2)\widehat{r}_1^{a+1}\widehat{r}_2^b\widehat{r}_3^c \\ (b+2)\widehat{r}_1^a\widehat{r}_2^{b+1}\widehat{r}_3^c \\ (c+2)\widehat{r}_1^a\widehat{r}_2^b\widehat{r}_3^{c+1} \end{bmatrix}, \quad \forall\,(a+b+c) = p-5.
\end{aligned} \tag{6.46}$$

Hence, the space of test functions $\mathcal{G}_p^v$ for the incremental order gradient function is defined as the direct sum of the different subspaces, depending on the relations of a, b, and c:

$$\begin{aligned}
\mathcal{G}_p^v =& \mathcal{G}_p^{a>b,c} \oplus \mathcal{G}_p^{b>a,c} \oplus \mathcal{G}_p^{c>a,b} \oplus \mathcal{G}_p^{a=b>c} \oplus \mathcal{G}_p^{a=c>b} \oplus \mathcal{G}_p^{b=c>a} \oplus \mathcal{G}_p^{a=b=c}, \\
=& \{\widetilde{\boldsymbol{q}}_{p+1}^{v,abc} \mid \quad \forall\,(a+b+c) = p-5 \text{ and } a > b,c\} \oplus \\
& \{\widetilde{\boldsymbol{q}}_{p+1}^{v,abc} \mid \quad \forall\,(a+b+c) = p-5 \text{ and } b > a,c\} \oplus \\
& \{\widetilde{\boldsymbol{q}}_{p+1}^{v,abc} \mid \quad \forall\,(a+b+c) = p-5 \text{ and } c > a,b\} \oplus \\
& \{\widetilde{\boldsymbol{q}}_{p+1}^{v,abc} \mid \quad \forall\,(a+b+c) = p-5 \text{ and } a = b > c\} \oplus \\
& \{\widetilde{\boldsymbol{q}}_{p+1}^{v,abc} \mid \quad \forall\,(a+b+c) = p-5 \text{ and } a = c > b\} \oplus \\
& \{\widetilde{\boldsymbol{q}}_{p+1}^{v,abc} \mid \quad \forall\,(a+b+c) = p-5 \text{ and } b = c > a\} \oplus \\
& \{\widetilde{\boldsymbol{q}}_{p+1}^{v,abc} \mid \quad \forall\,(a+b+c) = p-5 \text{ and } a = b = c\}.
\end{aligned} \tag{6.47}$$

The next step is the determination of the test function space $\widehat{\mathcal{Q}}_p^{wv}$ for the volume-associated moments. The procedure is the same as for the face-associated space, where the unnecessary terms are removed from the complete-order space. The sufficient oversized trial space reads

$$(\mathbb{P}_{p-4}(\widehat{V}))^3 = (\mathbb{P}_{p-5}(\widehat{V}))^3 \oplus \mathcal{P}_p^{a\leq b,c} \oplus \bar{\mathcal{P}}_p^{a\leq b,c} \qquad \text{for } \widehat{\mathcal{Q}}_p^{wv}, \tag{6.48}$$

where the highest-order subspaces $\mathcal{P}_p^{a\leq b,c}$ and $\bar{\mathcal{P}}_p^{a\leq b,c}$ are defined as

$$\mathcal{P}_p^{a\leq b,c} = \{\widehat{\boldsymbol{e}}_i \widehat{r}_i^a \widehat{r}_j^b \widehat{r}_k^c \mid \forall (a+b+c) = p-4 \text{ and } a \leq b \text{ or } a \leq c\}, \tag{6.49a}$$

$$\begin{aligned}
\bar{\mathcal{P}}_p^{a\leq b,c} =& \{\widehat{\boldsymbol{e}}_i \widehat{r}_i^a \widehat{r}_j^b \widehat{r}_k^c \mid \forall (a+b+c) = p-4 \text{ and } a > b,c\}, \\
=& \{\widehat{\boldsymbol{e}}_i \widehat{r}_i^{a+1} \widehat{r}_j^b \widehat{r}_k^c \mid \forall (a+b+c) = p-5 \text{ and } a \geq b,c\}.
\end{aligned} \tag{6.49b}$$

The one order less subspace $(\mathbb{P}_{p-5}(\widehat{V}))^3$ is a necessary condition of the convergence due to the error estimation (6.26) of the complete-order serendipity element. Moreover, using a similar argumentation as for the face-associated space, it is also considered that the curl operator requires the derivative of vector components with respect to the other variables, i.e., $\widehat{\boldsymbol{e}}_i \widehat{r}_j$ for $i \neq j$. Hence, any $\widehat{\boldsymbol{e}}_i \widehat{r}_i^a \widehat{r}_j^b \widehat{r}_k^c$ monomial, for arbitrary indices $i, j, k \in \{1, 2, 3\}$ and $a \leq b$ or $\leq c$, corresponds to a higher-order range space than its counterparts $\widehat{\boldsymbol{e}}_i \widehat{r}_i^a \widehat{r}_j^b \widehat{r}_k^c$ for $a > b, c$. Based on these considerations, the subspace $\mathcal{P}_p^{a\leq b,c}$ is preserved, and functions are only eliminated from $\bar{\mathcal{P}}_p^{a\leq b,c}$ via $\mathcal{G}_p^v$. Therefore, the definition for the space of volume-associated test functions is

$$\widehat{\mathcal{Q}}_p^{wv} = (\mathbb{P}_{p-5}(\widehat{V}))^3 \oplus \mathcal{P}_p^{a\leq b,c} \oplus \left(\bar{\mathcal{P}}_p^{a\leq b,c} \setminus (\mathcal{G}_p^v \oplus \mathcal{P}_p^{a\leq b,c} \oplus (\mathbb{P}_{p-5}(\widehat{V}))^3)\right). \tag{6.50}$$

In order to simplify this expression, a further subdivision of $\bar{\mathcal{P}}_p^{a\leq b,c}$ is required for the distinct cases

$$\begin{aligned}\bar{\mathcal{P}}_p^{a\leq b,c} =& \mathcal{P}_p^{a>b,c} \oplus \mathcal{P}_p^{a=b>c} \oplus \mathcal{P}_p^{a=c>b} \oplus \mathcal{P}_p^{a=b=c},\\ =& \{\widehat{\boldsymbol{e}}_i \widehat{r}_i^{a+1} \widehat{r}_j^b \widehat{r}_k^c \mid \forall (a+b+c) = p-5 \text{ and } a > b, c\}\\ &\oplus \{\widehat{\boldsymbol{e}}_i \widehat{r}_i^{a+1} \widehat{r}_j^b \widehat{r}_k^c \mid \forall (a+b+c) = p-5 \text{ and } a = b > c\}\\ &\oplus \{\widehat{\boldsymbol{e}}_i \widehat{r}_i^{a+1} \widehat{r}_j^b \widehat{r}_k^c \mid \forall (a+b+c) = p-5 \text{ and } a = c > b\}\\ &\oplus \{\widehat{\boldsymbol{e}}_i \widehat{r}_i^{a+1} \widehat{r}_j^b \widehat{r}_k^c \mid \forall (a+b+c) = p-5 \text{ and } a = b = c\}. \end{aligned} \tag{6.51}$$

It can be seen that the subspace $\mathcal{P}_p^{a>b,c}$ is completely eliminated by $\mathcal{G}_p^{a>b,c}$, $\mathcal{G}_p^{b>a,c}$, and $\mathcal{G}_p^{c>a,b}$ in (6.50). However, the remaining subspaces cannot be eliminated entirely. For example, $\mathcal{P}_p^{a=b=c}$ contains three functions, due to the different vector components, whereas $\mathcal{G}_p^{a=b=c}$ is only of a single dimension. Hence, (6.50) requires preserving the following two functions, and the corresponding space

$$\begin{aligned}\mathcal{R}_p^{a=b=c} = \Bigg\{ &\left(\widehat{\boldsymbol{e}}_1 \widehat{r}_1^{a+1} \widehat{r}_2^a \widehat{r}_3^a - \widehat{\boldsymbol{e}}_2 \widehat{r}_1^a \widehat{r}_2^{a+1} \widehat{r}_3^a\right), \left(\widehat{\boldsymbol{e}}_1 \widehat{r}_1^{a+1} \widehat{r}_2^a \widehat{r}_3^a - \widehat{\boldsymbol{e}}_3 \widehat{r}_1^a \widehat{r}_2^a \widehat{r}_3^{a+1}\right) \Bigg|\\ &\text{for } a = \left\lfloor \frac{p-5}{3} \right\rfloor = \frac{p-5}{3} \Bigg\}. \end{aligned} \tag{6.52}$$

Finally, the functions of $\mathcal{G}_p^{a=b>c} \oplus \mathcal{G}_p^{a=c>b} \oplus \mathcal{G}_p^{b=c>a}$ can be used to eliminate either $\mathcal{P}_p^{a=b>c}$ or $\mathcal{P}_p^{a=c>b}$ in (6.50), since their dimensions are identical and their highest-order monomials are the same. However, the difference between the two functions is picked to obtain symmetrical functions instead of choosing one over the other. Thus, resulting the remaining subspace is

$$\begin{aligned}\mathcal{R}_p^{a=b,a=c} = \Bigg\{ &\widehat{\boldsymbol{e}}_i \widehat{r}_i^{a+1} \widehat{r}_j^a \widehat{r}_k^b - \widehat{\boldsymbol{e}}_i \widehat{r}_i^{a+1} \widehat{r}_j^b \widehat{r}_k^a \Bigg|\\ &\text{for } \frac{p-5}{2} \geq a > \frac{p-5}{3},\ b = p-5-2a \Bigg\}. \end{aligned} \tag{6.53}$$

Definition of the Finite Element Moments

In order to summarize the previously obtained results, a short definition is given for the finite element moments of the mixed-order serendipity element. The spaces of test functions associated with the edge $\widehat{\mathcal{Q}}_p^{we}$, face $\widehat{\mathcal{Q}}_p^{wf}$, and volume $\widehat{\mathcal{Q}}_p^{wv}$ are

$$\widehat{\mathcal{Q}}_p^{we} = \mathbb{P}_{p-1}(\widehat{E})\widehat{\boldsymbol{e}}_1, \tag{6.54a}$$

$$\widehat{\mathcal{Q}}_p^{wf} = (\mathbb{P}_{p-3}(\widehat{F}))^2 \oplus \mathcal{P}_p^{a\leq b} \oplus \mathcal{R}_p^{a=b}, \tag{6.54b}$$

$$\widehat{\mathcal{Q}}_p^{wv} = (\mathbb{P}_{p-5}(\widehat{V}))^3 \oplus \mathcal{P}_p^{a\leq b,c} \oplus \mathcal{R}_p^{a=b,a=c} \oplus \mathcal{R}_p^{a=b=c}, \tag{6.54c}$$

where the subspaces are defined as

$$\begin{aligned}
\mathcal{P}_p^{a\leq b} =& \{\widehat{\boldsymbol{e}}_i \widehat{r}_i^a \widehat{r}_j^b \mid \forall (a+b) = p-2 \text{ and } a \leq b\},\\
\mathcal{P}_p^{a\leq b,c} =& \{\widehat{\boldsymbol{e}}_i \widehat{r}_i^a \widehat{r}_j^b \widehat{r}_k^c \mid \forall (a+b+c) = p-4 \text{ and } a \leq b \text{ or } a \leq c\},\\
\mathcal{R}_p^{a=b} =& \left\{ \widehat{\boldsymbol{e}}_1 \widehat{r}_1^{a+1} \widehat{r}_2^a - \widehat{\boldsymbol{e}}_2 \widehat{r}_1^a \widehat{r}_2^{a+1} \;\middle|\; \text{for } a = \left\lfloor \frac{p-3}{2} \right\rfloor = \frac{p-3}{2} \right\},\\
\mathcal{R}_p^{a=b,a=c} =& \left\{ \widehat{\boldsymbol{e}}_i \widehat{r}_i^{a+1} \widehat{r}_j^a \widehat{r}_k^b - \widehat{\boldsymbol{e}}_i \widehat{r}_i^{a+1} \widehat{r}_j^b \widehat{r}_k^a \;\middle|\; \text{for } \frac{p-5}{2} \geq a > \frac{p-5}{3},\, b = p-5-2a \right\},\\
\mathcal{R}_p^{a=b=c} =& \left\{ \left(\widehat{\boldsymbol{e}}_1 \widehat{r}_1^{a+1} \widehat{r}_2^a \widehat{r}_3^a - \widehat{\boldsymbol{e}}_2 \widehat{r}_1^a \widehat{r}_2^{a+1} \widehat{r}_3^a \right),\, \left(\widehat{\boldsymbol{e}}_1 \widehat{r}_1^{a+1} \widehat{r}_2^a \widehat{r}_3^a - \widehat{\boldsymbol{e}}_3 \widehat{r}_1^a \widehat{r}_2^a \widehat{r}_3^{a+1} \right) \;\middle|\; \right.\\
& \qquad \left. \text{for } a = \left\lfloor \frac{p-5}{3} \right\rfloor = \frac{p-5}{3} \right\},
\end{aligned}$$

with a, b, and c non-negative integers. The corresponding moments, associated with the edges $\widehat{M}_p^{we}$, faces $\widehat{M}_p^{wf}$, and volume $\widehat{M}_p^{wv}$ are defined identically to the complete-order serendipity element, i.e.,

$$\widehat{M}_p^{we}(\widehat{\boldsymbol{w}}) = \int_{\widehat{E}} \widehat{\boldsymbol{w}} \cdot \widehat{\boldsymbol{q}} \, \mathrm{d}\widehat{\boldsymbol{r}}, \qquad \forall \widehat{\boldsymbol{q}} \in \widehat{\mathcal{Q}}_p^{we}, \tag{6.56a}$$

$$\widehat{M}_p^{wf}(\widehat{\boldsymbol{w}}) = \int_{\widehat{F}} \widehat{\boldsymbol{w}} \cdot \widehat{\boldsymbol{q}} \, \mathrm{d}\widehat{\boldsymbol{r}}, \qquad \forall \widehat{\boldsymbol{q}} \in \widehat{\mathcal{Q}}_p^{wf}, \tag{6.56b}$$

$$\widehat{M}_p^{wv}(\widehat{\boldsymbol{w}}) = \int_{\widehat{V}} \widehat{\boldsymbol{w}} \cdot \widehat{\boldsymbol{q}} \, \mathrm{d}\widehat{\boldsymbol{r}}, \qquad \forall \widehat{\boldsymbol{q}} \in \widehat{\mathcal{Q}}_p^{wv}. \tag{6.56c}$$

Definition of the Approximation Space

After the declaration of the domain and the finite element moments, only the definition of the approximation space remains. The natural way to obtain approximation spaces is via smooth extensions of test functions, associated with the edge (6.20), face (6.21), and volume (6.22), to the entire domain while providing the necessary compact support. However, these extensions are not capable of generating functions for $\widehat{\nabla}\widehat{\mathcal{V}}_{S,p}$ since there are higher-order polynomial components normal to the associated edge (6.32) or face (6.37). This would prevent constructing higher-order basis functions with subsets of purely null space functions. To overcome this problem, only the range space functions are obtained as the natural extensions of the test functions, whereas the null space functions are already available from (6.11), (6.12), and (6.13) via $\widehat{\nabla}\widehat{\mathcal{V}}_{S,p}$.

In the first-order case, no separation is possible since the scalar functions are only associated with the nodes (6.10), whereas the vector functions are associated to the edges with different conformity requirements. Hence, the first-order functions $\widehat{\boldsymbol{w}}_{jk}, \widehat{\boldsymbol{w}}_{ik}, \widehat{\boldsymbol{w}}_{ij} \in \widehat{\mathcal{W}}^{e}_{SI,1}$ must be the same as for the other finite element spaces (6.20):

$$\widehat{\boldsymbol{w}}_{jk} = (y_j z_k \widehat{\boldsymbol{e}}_1), \qquad \text{on edge } y_j z_k, \tag{6.57a}$$

$$\widehat{\boldsymbol{w}}_{ik} = (x_i z_k \widehat{\boldsymbol{e}}_2), \qquad \text{on edge } x_i z_k, \tag{6.57b}$$

$$\widehat{\boldsymbol{w}}_{ij} = (x_i y_j \widehat{\boldsymbol{e}}_3), \qquad \text{on edge } x_i y_j. \tag{6.57c}$$

Here, the subscripts indicate the associated edges as shown in Figure 6.3. Considering that the highest-order test functions of the moments contribute only to the range space, they can be used to generate range space basis functions. The difference between the test function space of the mixed-order element (6.54) and the one-order less test function space of the complete-order element (6.18) yields $\widehat{\mathcal{Q}}^{we}_{p} \setminus \widehat{\boldsymbol{e}}_1 \mathbb{P}_{p-1}(\widehat{E}) = \emptyset$. Hence, there are no edge-associated range space functions for $p > 1$, and the definition of the edge-associated approximation space $\widehat{\mathcal{W}}^{e}_{SI,p}$ reads

$$\widehat{\mathcal{W}}^{e}_{SI,p} = \widehat{\mathcal{W}}^{e}_{SI,1} \oplus \widehat{\nabla}\widehat{\mathcal{V}}^{e}_{S,p}. \tag{6.58}$$

However, both null and range space functions are present in the higher-order case for the face- and volume-associated functions. The face-associated functions of order p are denoted by $\widehat{\boldsymbol{w}}_k, \widehat{\boldsymbol{w}}_j, \widehat{\boldsymbol{w}}_i \in \widetilde{\mathcal{A}}^{f}_{SI,p}$, that contribute to the range space of the curl operator. Here, the subscripts are indices for the associated faces as shown in Figure 6.3. These functions are required to have tangential compact support over the associated face and zero tangential component on all other faces. The general form of such a functions are

$$\widehat{\boldsymbol{w}}_k = \begin{bmatrix} y_0 y_1 z_k f^x_k(x_1, y_1) \\ x_0 x_1 z_k f^y_k(x_1, y_1) \\ 0 \end{bmatrix}, \quad \text{for } \begin{bmatrix} f^x_k(x_1, y_1) \\ f^y_k(x_1, y_1) \end{bmatrix} \in (\widehat{\mathcal{Q}}^{wf}_{p} \setminus (\mathbb{P}_{p-3}(\widehat{F}))^2) \text{ on face } z_k, \tag{6.59a}$$

$$\widehat{\boldsymbol{w}}_j = \begin{bmatrix} z_0 z_1 y_j f_j^x(x_1, z_1) \\ 0 \\ x_0 x_1 y_j f_j^z(x_1, z_1) \end{bmatrix}, \quad \text{for} \quad \begin{bmatrix} f_j^x(x_1, z_1) \\ f_j^z(x_1, z_1) \end{bmatrix} \in (\widehat{\mathcal{Q}}_p^{wf} \setminus (\mathbb{P}_{p-3}(\widehat{F}))^2) \text{ on face } y_j, \tag{6.59b}$$

$$\widehat{\boldsymbol{w}}_i = \begin{bmatrix} 0 \\ z_0 z_1 x_i f_i^y(y_1, z_1) \\ y_0 y_1 x_i f_i^z(y_1, z_1) \end{bmatrix}, \quad \text{for} \quad \begin{bmatrix} f_i^y(y_1, z_1) \\ f_i^z(y_1, z_1) \end{bmatrix} \in (\widehat{\mathcal{Q}}_p^{wf} \setminus (\mathbb{P}_{p-3}(\widehat{F}))^2) \text{ on face } x_i. \tag{6.59c}$$

Therefore, the face-associated approximation space $\widehat{\mathcal{W}}_{SI,p}^f$ is given as a direct sum of the incremental order spaces $\widetilde{\mathcal{A}}_{SI,p}^f$ and the null space $\widehat{\nabla}\widehat{\mathcal{V}}_{S,p}^f$, i.e.,

$$\widehat{\mathcal{W}}_{SI,p}^f = \widetilde{\mathcal{A}}_{SI,2}^f \oplus \ldots \oplus \widetilde{\mathcal{A}}_{SI,p}^f \oplus \widehat{\nabla}\widehat{\mathcal{V}}_{S,p}^f. \tag{6.60}$$

The volume-associated range space functions of incremental order p are denoted by $\widehat{\boldsymbol{w}} \in \widetilde{\mathcal{A}}_{SI,p}^v$. These functions must provide tangential compact support for the element, having zero tangential components on all faces. Their general form is

$$\widehat{\boldsymbol{w}} = \begin{bmatrix} y_0 y_1 z_0 z_1 f^x(x_1, y_1, z_1) \\ x_0 x_1 z_0 z_1 f^y(x_1, y_1, z_1) \\ x_0 x_1 y_0 y_1 f^z(x_1, y_1, z_1) \end{bmatrix}, \quad \begin{bmatrix} f^x(x_1, y_1, z_1) \\ f^y(x_1, y_1, z_1) \\ f^z(x_1, y_1, z_1) \end{bmatrix} \in (\widehat{\mathcal{Q}}_p^{wv} \setminus (\mathbb{P}_{p-5}(\widehat{V}))^3). \tag{6.61a}$$

Therefore, the volume-associated approximation space $\widehat{\mathcal{W}}_{SI,p}^v$ is given as a direct sum of the incremental order spaces $\widetilde{\mathcal{A}}_{SI,p}^v$ and the null space $\widehat{\nabla}\widehat{\mathcal{V}}_{S,p}^v$, i.e.,

$$\widehat{\mathcal{W}}_{SI,p}^v = \widetilde{\mathcal{A}}_{SI,4}^v \oplus \ldots \oplus \widetilde{\mathcal{A}}_{SI,p}^v \oplus \widehat{\nabla}\widehat{\mathcal{V}}_{S,p}^v. \tag{6.62}$$

Using the fact that the finite element moments induce the approximation space, the dimensions of the subspaces can be determined as

$$\dim(\widehat{\mathcal{W}}_{SI,p}^e) = \dim(\widehat{\mathcal{W}}_{S,p}^e) + \dim(\widehat{\mathcal{V}}_{S,p}^e) - \dim(\widehat{\mathcal{V}}_{S,p+1}^e) = 12p, \tag{6.63a}$$

$$\begin{aligned} \dim(\widehat{\mathcal{W}}_{SI,p}^f) &= \dim(\widehat{\mathcal{W}}_{S,p}^f) + \dim(\widehat{\mathcal{V}}_{S,p}^f) - \dim(\widehat{\mathcal{V}}_{S,p+1}^f), \\ &= 6(p^2 - 2p + 2), \qquad p > 1, \end{aligned} \tag{6.63b}$$

$$\begin{aligned} \dim(\widehat{\mathcal{W}}_{SI,p}^v) &= \dim(\widehat{\mathcal{W}}_{S,p}^v) + \dim(\widehat{\mathcal{V}}_{S,p}^v) - \dim(\widehat{\mathcal{V}}_{S,p+1}^v), \\ &= \frac{(p-3)(p^2 - 4p + 6)}{2}, \qquad p > 2. \end{aligned} \tag{6.63c}$$

Thus, the finite element approximation space $\widehat{\mathcal{W}}_{SI,p}$ of order p is defined as the direct sum of the edge-, face-, and volume-associated subspaces:

$$\widehat{\mathcal{W}}_{SI,p} = \widehat{\mathcal{W}}_{SI,p}^e \oplus \widehat{\mathcal{W}}_{SI,p}^f \oplus \widehat{\mathcal{W}}_{SI,p}^v. \tag{6.64}$$

6.2.4 Comparison to Other Serendipity Spaces

The first dimension independent definition of the scalar serendipity element was given by Arnold and Awanou [AA11]. They showed that these spaces provide the same convergence rate as the same order tensor product space for straight-sided geometries that are discretized by parallelepiped elements with affine geometry mappings. After them, several researchers have revisited the topic [RGB14], [AA14], [GKS19], [GK19]. In all cases, including the presented space $\widehat{\mathcal{V}}_{S,p}$, the approximation space turned out to be the same for hexahedra. Further, similar results have also been obtained for $\boldsymbol{H}(\mathrm{curl})$- and $\boldsymbol{H}(\mathrm{div})$-conforming elements. Arnold and Awanou published a general definition for serendipity spaces [AA14] based on differential forms. Their spaces for $\boldsymbol{H}(\mathrm{curl})$-conforming hexahedra turned out to be the same as the complete-order serendipity space $\widehat{\mathcal{W}}_{S,p}$ in this work. Exact basis functions were not provided; however, some nodal association locations were given for an interpolatory basis. Based on the analysis of Arnold, Boffi, and Bonizzoni [ABB15], the theoretical convergence rate is only achieved for serendipity spaces when each element is obtained via affine geometry mappings (parallelotope meshes). The main advantage of the presented new definitions over the existing ones is the explicit expression of the moments and the approximation spaces on the reference element, allowing the direct application of the results of Chapter 3. Hence, the ideal rate of convergence is guaranteed for affine refinements of curvilinear meshes and not just for parallelotope meshes with affine geometry mappings.

Recently, Gillette and Tyler published papers [GK19, GKS19] on the so-called trimmed serendipity spaces for $\boldsymbol{H}(\mathrm{curl})$-conforming hexahedra. Their theory is based on the application of differential forms for the globally defined finite element moments. They showed that the exclusion of some higher-order test functions of the finite element moments that contribute to the highest-order null space is permissible. Furthermore, they provided a semi-hierarchical basis for hexahedra, which is hierarchical except for the highest-order case [GKS19]. This is considered as a "feature of serendipity spaces, not a bug"[GKS19]. They suspected [GKS19] that it is not possible to achieve a fully hierarchical basis with the trimmed serendipity space. Their space has the same dimension as the proposed mixed-order serendipity space $\widehat{\mathcal{W}}_{SI,p}$. However, the actual finite element moments and the approximation space are different due to the difference in the highest-order incremental space. The main reason for this is that the approximation spaces are usually determined via smooth extensions of the test functions to the entire element domain. Oftentimes, these smooth extensions are multiplications with functions of the local coordinates, which provides sufficient compact support and conformity requirements. However, different conformity requirements yield different smooth extensions for the hexahedral spaces $\widehat{\mathcal{W}}_{S,p}$ and $\widehat{\nabla}\widehat{\mathcal{V}}_{S,p+1}$, resulting in $\widehat{\nabla}(\widehat{\mathcal{V}}_{S,p+1} \setminus \widehat{\mathcal{V}}_{S,p}) \not\subset \widehat{\mathcal{W}}_{S,p}$. Since the trimmed serendipity space of Gillette and Tyler is a subspace of the complete-order serendipity space $\widehat{\mathcal{W}}_{S,p}$, the highest-order functions must contribute to both the highest-order range space and null space. Therefore, the trimmed serendipity space

is not suitable for the generation of truly hierarchical basis functions. In contrast, the present analysis utilizes the de Rham complex of the approximation spaces to obtain the finite element moments instead of using the appropriate differential forms. The test functions of the moments that correspond to the highest-order gradient space $\widehat{\nabla}(\widehat{\mathcal{V}}_{S,p+1} \setminus \widehat{\mathcal{V}}_{S,p})$ are determined. After the elimination of these moments from the complete-order serendipity space $\widehat{\mathcal{W}}_{S,p}$, the highest-order functions only belong to the range space and are independent of the highest-order null space functions. Subsequently, the mixed-order (or incomplete-order) approximation space $\widehat{\mathcal{W}}_{SI,p}$ is defined as the direct sum of incremental order range spaces and a lower-order null space. Hence, the proposed mixed-order serendipity space is defined to facilitate the construction of a hierarchical basis with subsets of purely range space and null space functions for $p > 1$.

6.3 Tensor Product Spaces

Tensor product spaces are widely used for their structured way of implementation, allowing systematic procedures for the finite element assembly [ALM+09, AMGC19]. Another advantage is that some important properties from the lower dimensions can be easily extended to higher dimensions, such as the approximation properties. Generally, these spaces yield more accurate solutions and are applicable to general straight-sided geometries without affecting the convergence rate of the $\boldsymbol{L}^2$-norm error [ABB15]. In the framework presented in Chapter 3, this is equivalent to the elements, that are obtained via the piecewise linear (non-affine) refinements of straight-sided meshes. Each refinement ought to preserve the node locations of the initial elements; however, the newly inserted node locations on the faces and on the volume can be placed freely. Hence, piecewise linear refinement can reduce the geometry error in each refinement step if interpolated geometry mappings are used. This is the main advantage over affine refinement, where the nodes of the refined elements are determined by the mappings of the parent elements, denying any correction. Non-affine refinements always degrade the rate of convergence for serendipity elements. However, the main disadvantage of tensor product elements compared to serendipity elements is their much larger size, which increases the number of unknowns in the finite element equation system. In the following subsections, alternative definitions are provided for the best-known tensor product spaces for H^1- and $\boldsymbol{H}(\mathrm{curl})$-conforming hexahedral elements. The form and the notations of the finite element moments are identical to the general moments (2.73) and (2.74). Hence, the results of Chapter 3 are directly applicable.

6.3.1 H^1-Conforming Tensor Product Element for Straight-Sided Geometries

The space of test functions associated with the edge $\widehat{\mathcal{Q}}_p^{ve}$, face $\widehat{\mathcal{Q}}_p^{vf}$, and volume $\widehat{\mathcal{Q}}_p^{vv}$ are given as

- $\widehat{\mathcal{Q}}_p^{ve} = \mathbb{Q}_{p-2}$,
- $\widehat{\mathcal{Q}}_p^{vf} = \mathbb{Q}_{p-2,p-2}$,
- $\widehat{\mathcal{Q}}_p^{vv} = \mathbb{Q}_{p-2,p-2,p-2}$,

with the dimensions of

$$\dim(\widehat{\mathcal{Q}}_p^{ve}) = 12(p-1), \qquad p > 0, \tag{6.65a}$$

$$\dim(\widehat{\mathcal{Q}}_p^{vf}) = 6(p-1)^2, \qquad p > 0, \tag{6.65b}$$

$$\dim(\widehat{\mathcal{Q}}_p^{vv}) = (p-1)^3, \qquad p > 0. \tag{6.65c}$$

The finite element moments are defined on the reference element and are associated with the nodes $\widehat{M}_p^{vn}$, edges $\widehat{M}_p^{ve}$, faces $\widehat{M}_p^{vf}$, and volume $\widehat{M}_p^{vv}$ of the unit-cube:

$$\widehat{M}_p^{vn}(\widehat{v}) = \widehat{v}\Big|_{\widehat{N}}, \tag{6.66a}$$

$$\widehat{M}_p^{ve}(\widehat{v}) = \int_{\widehat{E}} \widehat{v}\widehat{q}\,\mathrm{d}\widehat{\boldsymbol{r}}, \qquad \forall \widehat{q} \in \widehat{\mathcal{Q}}_p^{ve}, \tag{6.66b}$$

$$\widehat{M}_p^{vf}(\widehat{v}) = \int_{\widehat{F}} \widehat{v}\widehat{q}\,\mathrm{d}\widehat{\boldsymbol{r}}, \qquad \forall \widehat{q} \in \widehat{\mathcal{Q}}_p^{vf}, \tag{6.66c}$$

$$\widehat{M}_p^{vv}(\widehat{v}) = \int_{\widehat{V}} \widehat{v}\widehat{q}\,\mathrm{d}\widehat{\boldsymbol{r}}, \qquad \forall \widehat{q} \in \widehat{\mathcal{Q}}_p^{vv}. \tag{6.66d}$$

Hence, the number of node-associated moments $\widehat{M}_p^{vn}$ and the dimension of the node-associated approximation space $\widehat{\mathcal{V}}_{T,p}^n$ is 8, independent of the order. The resulting finite element space of order p is denoted by $\widehat{\mathcal{V}}_{T,p}$, and it is the direct sum of the node-, edge-, face-, and volume-associated subspaces:

$$\widehat{\mathcal{V}}_{T,p} = \widehat{\mathcal{V}}_{T,1}^n \oplus \widehat{\mathcal{V}}_{T,p}^e \oplus \widehat{\mathcal{V}}_{T,p}^f \oplus \widehat{\mathcal{V}}_{T,p}^v. \tag{6.67}$$

As shown for the serendipity spaces in Section 6.2.1, the smooth extensions of the test functions to the entire volume with (6.10), (6.11), (6.12), and (6.13) lead to the tensor product space of $\widehat{\mathcal{V}}_{T,p} = \mathbb{Q}_{p,p,p}$. This can also be verified by comparing the dimensions of the moments and the approximation space:

$$\dim(\widehat{\mathcal{V}}_{T,p}) = (p+1)^3 = 8 + 12(p-1) + 6(p-1)^2 + (p-1)^3. \tag{6.68}$$

One can see that the serendipity space $\widehat{\mathcal{V}}_{S,p}$ of order p is a subspace $\widehat{\mathcal{V}}_{S,p} \subset \widehat{\mathcal{V}}_{T,p}$, which yields the convergence rate of (6.17). Moreover, the one order higher serendipity space is not a subspace $\widehat{\mathcal{V}}_{S,p+1} \not\subset \widehat{\mathcal{V}}_{T,p}$. Therefore, both the tensor product and the serendipity space yield the same convergence rate,

$$\left\|v - \pi_{p,h}^v v\right\|_{L^2(\Omega)} \leq Ch^{p+1}, \tag{6.69a}$$

$$\left\|v - \pi_{p,h}^v v\right\|_{H^1(\Omega)} \leq Ch^{p}, \tag{6.69b}$$

for affine refinements of the exact geometry mapping. Furthermore, in the case of straight-sided elements and piecewise linear (non-affine) refinements, the sufficient condition for the same rate of convergence is $\mathbb{P}_p \circ \boldsymbol{g}_1 = \widehat{\mathcal{V}}_{T,p}$ for $\forall \boldsymbol{g}_1$; see Table 3.1. Since $\mathbb{P}_p \circ \boldsymbol{g}_1 = \mathbb{Q}_{p,p,p}$, the rate of convergence is not affected for arbitrary straight-sided elements. This finite element space [M$^+$03, Chapter 6] and its convergence property [ABF02] are the best-known results of tensor product finite elements due to their wide range of applications.

6.3.2 $\boldsymbol{H}(\mathrm{curl})$-Conforming Tensor Product Element for Straight-Sided Geometries

The space of test functions associated with the edge $\widehat{\mathcal{Q}}_p^{we}$, face $\widehat{\mathcal{Q}}_p^{wf}$, and volume $\widehat{\mathcal{Q}}_p^{wv}$ are defined as

- $\widehat{\mathcal{Q}}_p^{we} = \widehat{\boldsymbol{e}}_1 \mathbb{Q}_{p-1}$,
- $\widehat{\mathcal{Q}}_p^{wf} = \widehat{\boldsymbol{e}}_1 \mathbb{Q}_{p-1,p-1} \oplus \widehat{\boldsymbol{e}}_2 \mathbb{Q}_{p-1,p-1}$,
- $\widehat{\mathcal{Q}}_p^{wv} = \widehat{\boldsymbol{e}}_1 \mathbb{Q}_{p-1,p-1,p-1} \oplus \widehat{\boldsymbol{e}}_2 \mathbb{Q}_{p-1,p-1,p-1} \oplus \widehat{\boldsymbol{e}}_3 \mathbb{Q}_{p-1,p-1,p-1}$.

The dimensions of these spaces are

$$12 \cdot \dim(\widehat{\mathcal{Q}}_p^{we}) = 12p, \tag{6.70a}$$

$$6 \cdot \dim(\widehat{\mathcal{Q}}_p^{wf}) = 12p^2 \tag{6.70b}$$

$$1 \cdot \dim(\widehat{\mathcal{Q}}_p^{wv}) = 3p^3. \tag{6.70c}$$

The corresponding moments associated with the edge $\widehat{M}_p^{we}$, face $\widehat{M}_p^{wf}$, and volume $\widehat{M}_p^{wv}$ are

$$\widehat{M}_p^{we}(\widehat{\boldsymbol{w}}) = \int_{\widehat{E}} \widehat{\boldsymbol{w}} \cdot \widehat{\boldsymbol{q}} \, \mathrm{d}\widehat{\boldsymbol{r}}, \qquad \forall \widehat{\boldsymbol{q}} \in \widehat{\mathcal{Q}}_p^{we}, \tag{6.71a}$$

$$\widehat{M}_p^{wf}(\widehat{\boldsymbol{w}}) = \int_{\widehat{F}} \widehat{\boldsymbol{w}} \cdot \widehat{\boldsymbol{q}} \, \mathrm{d}\widehat{\boldsymbol{r}}, \qquad \forall \widehat{\boldsymbol{q}} \in \widehat{\mathcal{Q}}_p^{wf}, \tag{6.71b}$$

$$\widehat{M}_p^{wv}(\widehat{\boldsymbol{w}}) = \int_{\widehat{V}} \widehat{\boldsymbol{w}} \cdot \widehat{\boldsymbol{q}} \, \mathrm{d}\widehat{\boldsymbol{r}}, \qquad \forall \widehat{\boldsymbol{q}} \in \widehat{\mathcal{Q}}_p^{wv}, \tag{6.71c}$$

which induce the corresponding approximation spaces the same way as for the serendipity elements. The edge-, face-, and volume-associated approximation spaces are denoted by $\widehat{\mathcal{W}}_{T,p}^e$, $\widehat{\mathcal{W}}_{T,p}^f$, and $\widehat{\mathcal{W}}_{T,p}^v$, respectively. These spaces are defined by the smooth extensions of the edge- (6.20), face- (6.21), and volume-associated (6.22) test functions to the entire domain. Their direct sum is the finite element approximation space $\widehat{\mathcal{W}}_{T,p}$, which is the tensor product space of

$$\widehat{\mathcal{W}}_{T,p} = \widehat{\mathcal{W}}_{T,p}^e \oplus \widehat{\mathcal{W}}_{T,p}^f \oplus \widehat{\mathcal{W}}_{T,p}^v = \begin{bmatrix} \mathbb{Q}_{p-1,p+1,p+1} \\ \mathbb{Q}_{p+1,p-1,p+1} \\ \mathbb{Q}_{p+1,p+1,p-1} \end{bmatrix}. \tag{6.72}$$

This can also be verified by comparing the dimension of the space to the dimension of the moments:

$$\dim(\widehat{\mathcal{W}}_{T,p}) = 3p(p+2)^2 = 12p + 12p^2 + 3p^3. \tag{6.73}$$

Convergence Properties

The first-order $\widehat{\mathcal{W}}_{T,1}$ approximation space was reported by Falk et al. [FGM11], whereas their definition of moments was different. Falk et al. rely on the construction of the $\boldsymbol{H}(\mathrm{div})$-conforming space, which is used for the separation of the curl operator null space. The authors prove that in the case of a general straight-sided hexahedron, the sufficient condition of the first-order convergence is

$$\widehat{\mathcal{W}}_{T,1} = \begin{bmatrix} \mathbb{Q}_{0,2,2} \\ \mathbb{Q}_{2,0,2} \\ \mathbb{Q}_{2,2,0} \end{bmatrix} \implies \begin{cases} \left\| \boldsymbol{w} - \pi^w_{p,h} \boldsymbol{w} \right\|_{\boldsymbol{L}^2(\Omega)} & = \mathcal{O}(h), \\ \left\| \boldsymbol{w} - \pi^w_{p,h} \boldsymbol{w} \right\|_{\boldsymbol{H}(\mathrm{curl},\Omega)} & = \mathcal{O}(h). \end{cases} \tag{6.74}$$

In a later work, Bergot and Duruflé claim [BD13b] that the optimal, sufficient, and necessary condition for the higher-order approximation space is

$$\widehat{\mathcal{W}}_{T,p} = \begin{bmatrix} \mathbb{Q}_{p-1,p+1,p+1} \\ \mathbb{Q}_{p+1,p-1,p+1} \\ \mathbb{Q}_{p+1,p+1,p-1} \end{bmatrix} \implies \begin{cases} \left\| \boldsymbol{w} - \pi^w_{p,h} \boldsymbol{w} \right\|_{\boldsymbol{L}^2(\Omega)} & = \mathcal{O}(h^p), \\ \left\| \boldsymbol{w} - \pi^w_{p,h} \boldsymbol{w} \right\|_{\boldsymbol{H}(\mathrm{curl},\Omega)} & = \mathcal{O}(h^p). \end{cases} \tag{6.75}$$

Their reasoning is that the global approximation space $\mathcal{W}_{T,p}$ must include the rotational space R_p of Nédélec (2.109a). Bergot et al. claim that the optimal $\widehat{\mathcal{W}}_{T,p}$ space is obtained via the Piola transformation of such a global space R_p to the reference element, which yields the desired rate of convergence (6.75). However, they provide no definition of the degrees of freedom and no proofs of error estimations for hexahedra. These elements were validated by numerical experiments where the convergence rates seem correct. However, no details of the model geometries are given, and the applied mesh [BD13b] seems to be a straight-sided mesh of a sphere inside a cube. In that case, achieving a higher convergence rate than $\mathcal{O}(h^1)$ is theoretically impossible; see Section 3.3 or [Cia02, Section 4.2] or [SF73, Section 3.2]. Furthermore, even if curvilinear meshes are used, achieving the correct convergence rates (6.75) is theoretically not possible for general curvilinear meshes; see Section 3.3 or [ABB15]. In the following, however, proofs are given for the minimality of such a $\widehat{\mathcal{W}}_{T,p}$ space for the restricted case of straight-sided geometries.

The starting point is the $\boldsymbol{L}^2(\Omega)$-norm estimate. According to Table 3.2, the sufficient condition for the convergence in case of exact, piecewise linear $\boldsymbol{g}_1$ geometry mappings is

$$\forall \boldsymbol{g}_1, \quad \boldsymbol{J}_1 (\mathbb{P}_{p-1})^3 \circ \boldsymbol{g}_1 \subset \widehat{\mathcal{W}}_{T,p} \implies \left\| \boldsymbol{w} - \pi^w_p \boldsymbol{w} \right\|_{\boldsymbol{L}^2(\Omega_e)} = \mathcal{O}(h^p). \tag{6.76}$$

Thus, the transformation involves the geometry mapping and its Jacobian. This is why the $\boldsymbol{H}(\mathrm{curl})$-conforming function space is more intricate than the H^1-conforming space. In the rest of the section, the condition of $\forall \boldsymbol{g}_1$ applies to each equation; however, it is left out to ease the readability. Using the definition of $\boldsymbol{g}_1$ from equation

(6.3), the general form of the Jacobian becomes

$$\begin{aligned}\boldsymbol{J}_1^T &= \left[\widehat{\nabla} g_x, \widehat{\nabla} g_y, \widehat{\nabla} g_z\right]^T, \\ &= \begin{bmatrix} b_1{+}e_1\hat{r}_2{+}g_1\hat{r}_3{+}h_1\hat{r}_2\hat{r}_3,\ c_1{+}e_1\hat{r}_1{+}f_1\hat{r}_3{+}h_1\hat{r}_1\hat{r}_3,\ d_1{+}f_1\hat{r}_2{+}g_1\hat{r}_1{+}h_1\hat{r}_1\hat{r}_2 \\ b_2{+}e_2\hat{r}_2{+}g_2\hat{r}_3{+}h_2\hat{r}_2\hat{r}_3,\ c_2{+}e_2\hat{r}_1{+}f_2\hat{r}_3{+}h_2\hat{r}_1\hat{r}_3,\ d_2{+}f_2\hat{r}_2{+}g_2\hat{r}_1{+}h_2\hat{r}_1\hat{r}_2 \\ b_3{+}e_3\hat{r}_2{+}g_3\hat{r}_3{+}h_3\hat{r}_2\hat{r}_3,\ c_3{+}e_3\hat{r}_1{+}f_3\hat{r}_3{+}h_3\hat{r}_1\hat{r}_3,\ d_3{+}f_3\hat{r}_2{+}g_3\hat{r}_1{+}h_3\hat{r}_1\hat{r}_2 \end{bmatrix}. \end{aligned} \tag{6.77}$$

It can be seen that each column vector of the Jacobian is in the space of

$$\boldsymbol{J}_1\widehat{\boldsymbol{e}}_i \in \begin{bmatrix} \mathbb{Q}_{0,1,1} \\ \mathbb{Q}_{1,0,1} \\ \mathbb{Q}_{1,1,0} \end{bmatrix}. \tag{6.78}$$

Moreover, the function composition of a $(p-1)$-order polynomial and the geometry mapping satisfies

$$(\mathbb{P}_{p-1})^3 \circ \boldsymbol{g}_1 \subset (\mathbb{Q}_{p-1,p-1,p-1})^3. \tag{6.79}$$

Combining (6.78) and (6.79), the sufficient space for the $\boldsymbol{L}^2(\Omega_e)$-norm convergence becomes

$$\boldsymbol{J}_1(\mathbb{P}_{p-1})^3 \circ \boldsymbol{g}_1 \subset \begin{bmatrix} \mathbb{Q}_{p-1,p,p} \\ \mathbb{Q}_{p,p-1,p} \\ \mathbb{Q}_{p,p,p-1} \end{bmatrix}. \tag{6.80}$$

Since this space is a subspace of $\widehat{\mathcal{W}}_{T,p}$, the resulting error estimate reads

$$\begin{bmatrix} \mathbb{Q}_{p-1,p,p} \\ \mathbb{Q}_{p,p-1,p} \\ \mathbb{Q}_{p,p,p-1} \end{bmatrix} \subset \widehat{\mathcal{W}}_{T,p} \implies \left\|\boldsymbol{w} - \pi_p^w \boldsymbol{w}\right\|_{\boldsymbol{L}^2(\Omega_e)} = \mathcal{O}(h^p), \tag{6.81}$$

for straight-sided elements.

Next, the investigation continues with the $\boldsymbol{H}(\text{curl})$-norm estimate. Two different proofs are provided for the determination of the minimal approximation space. The first proof is based on the estimation of functions $\nabla \times \boldsymbol{w} \in \boldsymbol{H}(\text{div})$. Then, the sufficient space $\widehat{\mathcal{W}}_{T,p}$ is determined by using the commutation of the interpolation operator, as well as the mappings of the differential operators. The second proof is based on the estimation of functions $\boldsymbol{d} \in \boldsymbol{H}(\text{div})$, where the sufficient $\widehat{\mathcal{D}}_{T,p}$ space is determined. Then, all those subspaces that are not in the form of $\widehat{\nabla} \times \widehat{\boldsymbol{w}} \subset \widehat{\nabla} \times \widehat{\mathcal{W}}_{T,p}$ are eliminated. Hence, there is no need to use the commutation properties or to assume that the dimension of the subspaces is unaffected. The provision of these proofs from both sides guarantees the minimality of $\widehat{\mathcal{W}}_{T,p}$.

Applying the results of Table 3.2, the convergence requirement of functions $\nabla \times \boldsymbol{w} \in \boldsymbol{H}(\text{div})$ for straight-sided elements and piecewise linear refinements is

$$\det(\boldsymbol{J}_1)\boldsymbol{J}_1^{-T}\left(\nabla \times (\mathbb{P}_p)^3\right) \circ \boldsymbol{g}_1 \subset \widehat{\nabla} \times \widehat{\mathcal{W}}_{T,p} \implies \left\|\nabla \times (\boldsymbol{w} - \pi_p^w \boldsymbol{w})\right\|_{\boldsymbol{L}^2(\Omega_e)} = \mathcal{O}(h^p). \tag{6.82}$$

Let us consider a decomposition [DVBMR17] for the complete-order polynomial space

$$(\mathbb{P}_p)^3 = (\mathbb{P}_{p-1})^3 \oplus \boldsymbol{r} \times (\widetilde{\mathbb{P}}_{p-1})^3 \oplus \nabla\widetilde{\mathbb{P}}_{p+1}, \tag{6.83}$$

where $\boldsymbol{r}$ denotes the coordinate vector and $\widetilde{\mathbb{P}}_{p-1}$ is the space of homogeneous polynomials of order $p-1$. Here, the highest-order gradient subspace $\nabla\widetilde{\mathbb{P}}_{p+1}$ is negligible. since it is in the null space of the curl operator. Thus, the sufficient space is simplified to

$$\det(\boldsymbol{J}_1)\boldsymbol{J}_1^{-T}\left(\nabla \times \left((\mathbb{P}_{p-1})^3 \oplus \boldsymbol{r} \times (\widetilde{\mathbb{P}}_{p-1})^3\right)\right) \circ \boldsymbol{g}_1 \subset \widehat{\nabla} \times \widehat{\mathcal{W}}_{T,p}. \tag{6.84}$$

Further simplification is possible through applying the mapping of the curl operator (2.64) to the local domain:

$$\det(\boldsymbol{J}_1)\boldsymbol{J}_1^{-T}\left(\nabla \times \left((\mathbb{P}_{p-1})^3 \oplus \boldsymbol{r} \times (\widetilde{\mathbb{P}}_{p-1})^3\right)\right) \circ \boldsymbol{g}_1 = \widehat{\nabla} \times \left(\boldsymbol{J}_1\left((\mathbb{P}_{p-1})^3 \oplus \boldsymbol{r} \times (\widetilde{\mathbb{P}}_{p-1})^3\right) \circ \boldsymbol{g}_1\right). \tag{6.85}$$

Hence, the sufficient condition becomes

$$\boldsymbol{J}_1\left((\mathbb{P}_{p-1})^3 \oplus \boldsymbol{r} \times (\widetilde{\mathbb{P}}_{p-1})^3\right) \circ \boldsymbol{g}_1 \subset \widehat{\mathcal{W}}_{T,p} \implies \left\|\nabla \times (\boldsymbol{w} - \pi_p^w \boldsymbol{w})\right\|_{\boldsymbol{L}^2(\Omega_e)} = \mathcal{O}(h^p). \tag{6.86}$$

Here, the polynomial space $(\mathbb{P}_{p-1})^3 \oplus \boldsymbol{r} \times (\widetilde{\mathbb{P}}_{p-1})^3 = R_p$ becomes the rotational space of Nédélec (2.109a). Therefore, it is equivalent to the statement of Bergot and Duruflé [BD13a]. The first term

$$\boldsymbol{J}_1(\mathbb{P}_{p-1})^3 \circ \boldsymbol{g}_1 \subset \begin{bmatrix} \mathbb{Q}_{p-1,p,p} \\ \mathbb{Q}_{p,p-1,p} \\ \mathbb{Q}_{p,p,p-1} \end{bmatrix} \subset \widehat{\mathcal{W}}_{T,p} \tag{6.87}$$

is already included in the space $\widehat{\mathcal{W}}_{T,p}$, since it is required for the $\boldsymbol{L}^2(\Omega_e)$-norm convergence. Hence, the remaining requirement is to include the space of

$$\boldsymbol{J}_1\left(\boldsymbol{r} \times (\widetilde{\mathbb{P}}_{p-1})^3\right) \circ \boldsymbol{g}_1 = \left(\boldsymbol{J}_1\left(\boldsymbol{g}_1 \times (\mathbb{Q}_{p-1,p-1,p-1})^3\right)\right) \subset \widehat{\mathcal{W}}_{T,p}. \tag{6.88}$$

All functions of this space are of the form

$$\boldsymbol{J}_1\left(\boldsymbol{g}_1 \times \boldsymbol{w}\right) = (g_z\widehat{\nabla}g_y - g_y\widehat{\nabla}g_z)[\boldsymbol{w}]_1 + (g_x\widehat{\nabla}g_z - g_z\widehat{\nabla}g_x)[\boldsymbol{w}]_2 + (g_y\widehat{\nabla}g_x - g_x\widehat{\nabla}g_y)[\boldsymbol{w}]_3, \quad (6.89)$$

where $[\boldsymbol{w}]_i \in \mathbb{Q}_{p-1,p-1,p-1}$ and $g_x, g_y, g_z \in \mathbb{Q}_{1,1,1}$ are polynomials. Next, the investigation is only continued for the last term of (6.89) and only for the first vector component

$$[(g_y\widehat{\nabla}g_x - g_x\widehat{\nabla}g_y)]_1 = g_y\frac{\partial g_x}{\partial \widehat{r}_1} - g_x\frac{\partial g_y}{\partial \widehat{r}_1}, \quad (6.90)$$

since the other components and terms would yield similar results due to symmetry. Consider a decomposition for the geometry mapping based on the variable $\widehat{r}_1$:

$$g_x = a_x + b_x\widehat{r}_1, \qquad a_x, b_x \in \mathbb{Q}_{0,1,1}, \quad (6.91a)$$
$$g_y = a_y + b_y\widehat{r}_1, \qquad a_y, b_y \in \mathbb{Q}_{0,1,1}. \quad (6.91b)$$

Thus, the term of (6.90) yields

$$g_y\frac{\partial g_x}{\partial \widehat{r}_1} - g_x\frac{\partial g_y}{\partial \widehat{r}_1} = (a_y + b_y\widehat{r}_1)b_x - (a_x + b_x\widehat{r}_1)b_y = a_yb_x - a_xb_y \in \mathbb{Q}_{0,2,2}. \quad (6.92)$$

Hence, each term of (6.89) belongs to the space of

$$(g_z\widehat{\nabla}g_y - g_y\widehat{\nabla}g_z),\ (g_x\widehat{\nabla}g_z - g_z\widehat{\nabla}g_x),\ (g_y\widehat{\nabla}g_x - g_x\widehat{\nabla}g_y) \in \begin{bmatrix}\mathbb{Q}_{0,2,2}\\ \mathbb{Q}_{2,0,2}\\ \mathbb{Q}_{2,2,0}\end{bmatrix}. \quad (6.93)$$

Moreover, the effect of the multiplication with a component of $\boldsymbol{w} \in (\mathbb{Q}_{p-1,p-1,p-1})^3$ polynomial in (6.89) yields

$$\boldsymbol{J}_1\left(\boldsymbol{g}_1 \times \boldsymbol{w}\right) \in \begin{bmatrix}\mathbb{Q}_{p-1,p+1,p+1}\\ \mathbb{Q}_{p+1,p-1,p+1}\\ \mathbb{Q}_{p+1,p+1,p-1}\end{bmatrix}, \quad (6.94)$$

which is the sufficient approximation space

$$\begin{bmatrix}\mathbb{Q}_{p-1,p+1,p+1}\\ \mathbb{Q}_{p+1,p-1,p+1}\\ \mathbb{Q}_{p+1,p+1,p-1}\end{bmatrix} = \widehat{\mathcal{W}}_{T,p}. \quad (6.95)$$

The second approach for the derivation of the sufficient space $\widehat{\mathcal{W}}_{T,p}$ uses the approximation theory of functions $\boldsymbol{d} \in \boldsymbol{H}(\mathrm{div})$. This method is suggested by Falk in [FGM11], where the analysis is carried out for first-order elements with exact,

straight-sided geometry mapping $\boldsymbol{g}_1$. In the following, this theory of Falk is extended to the higher-order case, starting with the requirement (3.175) for the $\boldsymbol{L}^2(\Omega_e)$-norm convergence:

$$\det(\boldsymbol{J}_1)\boldsymbol{J}_1^{-T}(\mathbb{P}_{p-1})^3 \circ \boldsymbol{g}_1 \subset \widehat{\mathcal{D}}_{T,p} \implies \left\|\boldsymbol{d} - \pi_p^d \boldsymbol{d}\right\|_{\boldsymbol{L}^2(\Omega_e)} = \mathcal{O}(h^p). \tag{6.96}$$

Following the proof of Falk, the definition of the adjugate matrix can be used to show that the multiplicative term of $\det(\boldsymbol{J}_1)\boldsymbol{J}_1^{-T}$ is a polynomial. Applying the general form of the Jacobian (6.77), the first column of the matrix $\det(\boldsymbol{J}_1)\boldsymbol{J}_1^{-T}$ becomes

$$\begin{aligned}\det(\boldsymbol{J}_1)\boldsymbol{J}_1^{-T}\widehat{\boldsymbol{e}}_1 &= \operatorname{adj}(\boldsymbol{J}_1^T)\widehat{\boldsymbol{e}}_1 \\ &= \begin{bmatrix} A_1^1+(D_3^1-C_2^1)\hat{r}_1+C_1^1\hat{r}_2+D_1^1\hat{r}_3-(E_2^1+G_2^1)\hat{r}_1\hat{r}_2+(E_3^1-G_3^1)\hat{r}_1\hat{r}_3+G_1^1\hat{r}_1^2+H_3^1\hat{r}_1^2\hat{r}_2-H_2^1\hat{r}_1^2\hat{r}_3 \\ A_2^1+B_2^1\hat{r}_1+(B_1^1-D_3^1)\hat{r}_2+D_2^1\hat{r}_3+(E_1^1-G_1^1)\hat{r}_2\hat{r}_1-(E_3^1+G_3^1)\hat{r}_2\hat{r}_3+G_2^1\hat{r}_2^2-H_3^1\hat{r}_1\hat{r}_2^2+H_1^1\hat{r}_2^2\hat{r}_3 \\ A_3^1+B_3^1\hat{r}_1+C_3^1\hat{r}_2+(C_2^1-B_1^1)\hat{r}_3-(E_1^1+G_1^1)\hat{r}_3\hat{r}_1+(E_2^1-G_2^1)\hat{r}_3\hat{r}_2+G_3^1\hat{r}_3^2+H_2^1\hat{r}_1\hat{r}_3^2-H_1^1\hat{r}_2\hat{r}_3^2 \end{bmatrix},\end{aligned}$$

where $\widehat{\boldsymbol{e}}_1 = (1,0,0)^T$ is a unit vector. The constant coefficients are determined as

$$\begin{aligned}
&A_1^1 = c_2d_3 - d_2c_3, && A_2^1 = d_2b_3 - b_2d_3, && A_3^1 = b_2c_3 - c_2b_3,\\
&B_1^1 = f_2b_3 - b_2f_3, && B_2^1 = g_2b_3 - b_2g_3, && B_3^1 = b_2e_3 - e_2b_3,\\
&C_1^1 = c_2f_3 - f_2c_3, && C_2^1 = g_2c_3 - c_2g_3, && C_3^1 = e_2c_3 - c_2e_3,\\
&D_1^1 = f_2d_3 - d_2f_3, && D_2^1 = d_2g_3 - g_2d_3, && D_3^1 = e_2d_3 - e_3d_2,\\
&E_1^1 = h_2b_3 - h_3b_2, && E_2^1 = h_2c_3 - h_3c_2, && E_3^1 = h_2d_3 - h_3d_2,\\
&G_1^1 = e_2g_3 - g_2e_3, && G_2^1 = f_2e_3 - e_2f_3, && G_3^1 = g_2f_3 - f_2g_3,\\
&H_1^1 = f_2h_3 - h_2f_3, && H_2^1 = h_3g_2 - h_2g_3, && H_3^1 = e_2h_3 - h_2e_3.
\end{aligned}$$

Similar results can be obtained for all other column vectors of $\operatorname{adj}(\boldsymbol{J}_1^T)$ by replacing the constant coefficients $A_1^1, A_2^1, A_3^1, B_1^1 \ldots$ by $A_1^i, A_2^i, A_3^i, B_1^i \ldots$. Thus, all column vectors belong to the space of

$$\det(\boldsymbol{J}_1)\boldsymbol{J}_1^{-T}\widehat{\boldsymbol{e}}_i \subset \begin{bmatrix} \mathbb{Q}_{2,1,0} \cup \mathbb{Q}_{2,0,1} \\ \mathbb{Q}_{0,2,1} \cup \mathbb{Q}_{1,2,0} \\ \mathbb{Q}_{1,0,2} \cup \mathbb{Q}_{0,1,2} \end{bmatrix}. \tag{6.97}$$

Then, the sufficient approximation space is obtained by the combination of such an adjugate matrix and the composition of polynomials (6.79) as

$$\det(\boldsymbol{J}_1)\boldsymbol{J}_1^{-T}(\mathbb{P}_{p-1})^3 \circ \boldsymbol{g}_1 \in \widehat{\mathcal{D}}_{T,p} = \begin{bmatrix} \mathbb{Q}_{p+1,p,p-1} \cup \mathbb{Q}_{p+1,p-1,p} \\ \mathbb{Q}_{p-1,p+1,p} \cup \mathbb{Q}_{p,p+1,p-1} \\ \mathbb{Q}_{p,p-1,p+1} \cup \mathbb{Q}_{p-1,p,p+1} \end{bmatrix}. \tag{6.98}$$

The next step of the proof is to look for the appropriate curl subspace $\widehat{\nabla} \times \widehat{\mathcal{W}}_{T,p} \subset \widehat{\mathcal{D}}_{T,p}$, where $\dim(\widehat{\mathcal{W}}_{T,p})$ is minimal. One can obtain a lower and upper bound for the space $\widehat{\mathcal{W}}_{T,p}$ with the $\boldsymbol{L}^2(\Omega_e)$-norm convergent space (6.81) of order p and $p+1$;

$$\begin{bmatrix} \mathbb{Q}_{p-1,p,p} \\ \mathbb{Q}_{p,p-1,p} \\ \mathbb{Q}_{p,p,p-1} \end{bmatrix} \subseteq \widehat{\mathcal{W}}_{T,p} \subseteq \begin{bmatrix} \mathbb{Q}_{p,p+1,p+1} \\ \mathbb{Q}_{p+1,p,p+1} \\ \mathbb{Q}_{p+1,p+1,p} \end{bmatrix}. \tag{6.99}$$

Thus, the incremental trial spaces that need to be investigated are

$$\begin{bmatrix}\mathbb{Q}_{p,p+1,p+1}\\ \mathbb{Q}_{p+1,p,p+1}\\ \mathbb{Q}_{p+1,p+1,p}\end{bmatrix} \setminus \begin{bmatrix}\mathbb{Q}_{p-1,p,p}\\ \mathbb{Q}_{p,p-1,p}\\ \mathbb{Q}_{p,p,p-1}\end{bmatrix} = \begin{bmatrix}\hat{r}_2^{p+1}\hat{r}_3^{p+1}\mathbb{Q}_{p-1,0,0}\\ \hat{r}_1^{p+1}\hat{r}_3^{p+1}\mathbb{Q}_{0,p-1,0}\\ \hat{r}_1^{p+1}\hat{r}_2^{p+1}\mathbb{Q}_{0,0,p-1}\end{bmatrix} \oplus \begin{bmatrix}\hat{r}_2^{p+1}\mathbb{Q}_{p-1,0,p}\\ \hat{r}_1^{p+1}\mathbb{Q}_{0,p-1,p}\\ \hat{r}_1^{p+1}\mathbb{Q}_{0,p,p-1}\end{bmatrix} \oplus \begin{bmatrix}\hat{r}_3^{p+1}\mathbb{Q}_{p-1,p,0}\\ \hat{r}_3^{p+1}\mathbb{Q}_{p,p-1,0}\\ \hat{r}_2^{p+1}\mathbb{Q}_{p,0,p-1}\end{bmatrix} \oplus \begin{bmatrix}\hat{r}_1^{p}\mathbb{Q}_{0,p+1,p+1}\\ \hat{r}_2^{p}\mathbb{Q}_{p+1,0,p+1}\\ \hat{r}_3^{p}\mathbb{Q}_{p+1,p+1,0}\end{bmatrix}. \tag{6.100}$$

In the following, the proof is determined only for the $\widehat{\boldsymbol{e}}_1$-component of the space, due to the structured nature of the geometry. Then, an arbitrary monomial of the first term of the decomposition reads $\hat{r}_1^i\hat{r}_2^{p+1}\hat{r}_3^{p+1} \in \hat{r}_2^{p+1}\hat{r}_3^{p+1}\mathbb{Q}_{p-1,0,0}$ for all indices $i \leq p-1$. It can be shown that the curl of this monomial belongs to the space of

$$\widehat{\nabla} \times (\widehat{\boldsymbol{e}}_1\hat{r}_1^i\hat{r}_2^{p+1}\hat{r}_3^{p+1}) = \begin{bmatrix}0\\ \hat{r}_1^i\hat{r}_2^{p+1}\hat{r}_3^{p}(p+1)\\ -\hat{r}_1^i\hat{r}_2^{p}\hat{r}_3^{p+1}(p+1)\end{bmatrix} \in \widehat{\mathcal{D}}_{T,p}. \tag{6.101}$$

The same procedure can be repeated for the monomials of the second $\hat{r}_1^i\hat{r}_2^{p+1}\hat{r}_3^k \in \hat{r}_2^{p+1}\mathbb{Q}_{p-1,0,p}$ and third $\hat{r}_1^i\hat{r}_2^k\hat{r}_3^{p+1} \in \hat{r}_3^{p+1}\mathbb{Q}_{p-1,p,0}$ terms of the decomposition. For all indices of $i \leq p-1$ and $k \leq p$, the curl of the monomials belong to the space of

$$\widehat{\nabla} \times (\widehat{\boldsymbol{e}}_1\hat{r}_1^i\hat{r}_2^{p+1}\hat{r}_3^{k}) = \begin{bmatrix}0\\ \hat{r}_1^i\hat{r}_2^{p+1}\hat{r}_3^{k-1}(k)\\ -\hat{r}_1^i\hat{r}_2^{p}\hat{r}_3^{k}(p+1)\end{bmatrix} \in \widehat{\mathcal{D}}_{T,p}, \tag{6.102}$$

$$\widehat{\nabla} \times (\widehat{\boldsymbol{e}}_1\hat{r}_1^i\hat{r}_2^{k}\hat{r}_3^{p+1}) = \begin{bmatrix}0\\ \hat{r}_1^i\hat{r}_2^{k}\hat{r}_3^{p}(p+1)\\ -\hat{r}_1^i\hat{r}_2^{k-1}\hat{r}_3^{p+1}(k)\end{bmatrix} \in \widehat{\mathcal{D}}_{T,p}. \tag{6.103}$$

Hence, the condition of

$$\begin{bmatrix}\mathbb{Q}_{p-1,p+1,p+1}\\ \mathbb{Q}_{p+1,p-1,p+1}\\ \mathbb{Q}_{p+1,p+1,p-1}\end{bmatrix} \subseteq \widehat{\mathcal{W}}_{T,p}, \tag{6.104}$$

is necessary for the convergence. In the remaining part, it is shown that this is also a sufficient condition. A further decomposition can be obtained for $\widehat{\boldsymbol{e}}_1\hat{r}_1^p\mathbb{Q}_{0,p+1,p+1}$, which is the last remaining term of (6.100):

$$\hat{r}_1^p\mathbb{Q}_{0,p+1,p+1} = \hat{r}_1^p\hat{r}_2^{p+1}\mathbb{Q}_{0,0,p+1} \oplus \hat{r}_1^p\hat{r}_3^{p+1}\mathbb{Q}_{0,p,0} \oplus \hat{r}_1^p\mathbb{Q}_{0,p,p}. \tag{6.105}$$

Here, any monomial of the first term can be written as $\hat{r}_1^p\hat{r}_2^{p+1}\hat{r}_3^i \in \hat{r}_1^p\hat{r}_2^{p+1}\mathbb{Q}_{0,0,p+1}$ for all indices $i \leq p+1$. Moreover, any monomial of the second term can be written as

$\hat{r}_1^p \hat{r}_2^k \hat{r}_3^{p+1} \in \hat{r}_1^p \hat{r}_3^{p+1} \mathbb{Q}_{0,p,0}$ for all indices $k \leq p$. Since the curl of these monomials do not belong to the space of $\widehat{\mathcal{D}}_{T,p}$

$$\widehat{\nabla} \times (\widehat{\boldsymbol{e}}_1 \hat{r}_1^p \hat{r}_2^{p+1} \hat{r}_3^i) = \begin{bmatrix} 0 \\ \hat{r}_1^p \hat{r}_2^{p+1} \hat{r}_3^{i-1}(i) \\ -\hat{r}_1^p \hat{r}_2^p \hat{r}_3^i (p+1) \end{bmatrix} \notin \widehat{\mathcal{D}}_{T,p}, \tag{6.106}$$

$$\widehat{\nabla} \times (\widehat{\boldsymbol{e}}_1 \hat{r}_1^p \hat{r}_2^k \hat{r}_3^{p+1}) = \begin{bmatrix} 0 \\ \hat{r}_1^p \hat{r}_2^k \hat{r}_3^p (p+1) \\ -\hat{r}_1^p \hat{r}_2^{k-1} \hat{r}_3^{p+1}(k) \end{bmatrix} \notin \widehat{\mathcal{D}}_{T,p}, \tag{6.107}$$

they are unnecessary for the space of $\widehat{\mathcal{W}}_{T,p}$ as well. Finally, the remaining space of $\widehat{\boldsymbol{e}}_1 \hat{r}_1^p \mathbb{Q}_{0,p,p}$ needs to be eliminated from $\widehat{\mathcal{W}}_{T,p}$. For this, one may use an arbitrary linear combination of the highest-order gradient functions in $\nabla(\widehat{\mathcal{V}}_{T,p+1} \setminus \widehat{\mathcal{V}}_{T,p})$, since they do not contribute to the curl-subspace, i.e.,

$$\widehat{\nabla} \times \widehat{\mathcal{W}}_{T,p} = \widehat{\nabla} \times \left(\widehat{\mathcal{W}}_{T,p} \setminus \widehat{\nabla}(\widehat{\mathcal{V}}_{T,p+1} \setminus \widehat{\mathcal{V}}_{T,p}) \right). \tag{6.108}$$

Hence, the highest-order gradient functions of

$$\nabla(\hat{r}_1^{p+1} \mathbb{Q}_{0,p,p}) = \begin{bmatrix} \hat{r}_1^p \mathbb{Q}_{0,p,p} \\ \hat{r}_1^{p+1} \frac{\partial}{\partial \hat{r}_2} \mathbb{Q}_{0,p,p} \\ \hat{r}_1^{p+1} \frac{\partial}{\partial \hat{r}_3} \mathbb{Q}_{0,p,p} \end{bmatrix} \subset \nabla(\widehat{\mathcal{V}}_{T,p+1} \setminus \widehat{\mathcal{V}}_{T,p}), \tag{6.109}$$

can be used to eliminate $\widehat{\boldsymbol{e}}_1 \hat{r}_1^p \mathbb{Q}_{0,p,p}$, since the second $\widehat{\boldsymbol{e}}_1 \hat{r}_1^{p+1} \frac{\partial}{\partial \hat{r}_2} \mathbb{Q}_{0,p,p}$ and third components $\widehat{\boldsymbol{e}}_1 \hat{r}_1^{p+1} \frac{\partial}{\partial \hat{r}_3} \mathbb{Q}_{0,p,p}$ are already included in $\widehat{\mathcal{W}}_{T,p}$. One can obtain analogous results for the other vector components of (6.100), since the symmetry of the reference hexahedron corresponds to a symmetric approximation space. Therefore, the sufficient and necessary condition becomes

$$\begin{bmatrix} \mathbb{Q}_{p-1,p+1,p+1} \\ \mathbb{Q}_{p+1,p-1,p+1} \\ \mathbb{Q}_{p+1,p+1,p-1} \end{bmatrix} = \widehat{\mathcal{W}}_{T,p} \implies \left\| \boldsymbol{w} - \pi_p^w \boldsymbol{w} \right\|_{\boldsymbol{H}(\mathrm{curl},\Omega_e)} = \mathcal{O}(h^p), \tag{6.110}$$

for general straight-sided elements.

6.3.3 Complete-Order Nédélec Element

Nédélec proposed two different approximation spaces for hexahedra. The so-called Nédélec element of the second kind [Néd86] is also referred to as the complete-order Nédélec element, since it contains a complete p-order polynomial subspace. Furthermore, the so-called Nédélec elements of the first kind [Néd80] are also known as mixed-order or incomplete-order elements. Here, the dimension of the range space is identical to the complete p-order case, whereas the dimension of the null space is identical to the case of $(p-1)$-order. The mixed-order Nédélec space is obtained via removal of the appropriate functions from the complete-order Nédélec space. Hence, the main difference between these spaces is their dimension and the resulting convergence rates. In this subsection, a definition is given for the complete-order Nédélec element. As always, the reference domain is a cube with unit length edges. The spaces of test functions associated with the edge $\widehat{\mathcal{Q}}_p^{we}$, face $\widehat{\mathcal{Q}}_p^{wf}$, and volume $\widehat{\mathcal{Q}}_p^{wv}$ are defined as

- $\widehat{\mathcal{Q}}_p^{we} = \widehat{\boldsymbol{e}}_1 \mathbb{Q}_p$,
- $\widehat{\mathcal{Q}}_p^{wf} = \widehat{\boldsymbol{e}}_1 \mathbb{Q}_{p,p-2} \oplus \widehat{\boldsymbol{e}}_2 \mathbb{Q}_{p-2,p}$,
- $\widehat{\mathcal{Q}}_p^{wv} = \widehat{\boldsymbol{e}}_1 \mathbb{Q}_{p,p-2,p-2} \oplus \widehat{\boldsymbol{e}}_2 \mathbb{Q}_{p-2,p,p-2} \oplus \widehat{\boldsymbol{e}}_3 \mathbb{Q}_{p-2,p-2,p}$,

with the dimensions of

$$12 \cdot \dim(\widehat{\mathcal{Q}}_p^{we}) = 12(p+1), \tag{6.111a}$$

$$6 \cdot \dim(\widehat{\mathcal{Q}}_p^{wf}) = 12(p+1)(p-1), \tag{6.111b}$$

$$1 \cdot \dim(\widehat{\mathcal{Q}}_p^{wv}) = 3(p+1)(p-1)^2. \tag{6.111c}$$

The corresponding moments are associated with the edge $\widehat{M}_p^{we}$, face $\widehat{M}_p^{wf}$, and volume $\widehat{M}_p^{wv}$ as

$$\widehat{M}_p^{we}(\widehat{\boldsymbol{w}}) = \int_{\widehat{E}} \widehat{\boldsymbol{w}} \cdot \widehat{\boldsymbol{q}} \, \mathrm{d}\widehat{\boldsymbol{r}}, \qquad \forall \widehat{\boldsymbol{q}} \in \widehat{\mathcal{Q}}_p^{we}, \tag{6.112a}$$

$$\widehat{M}_p^{wf}(\widehat{\boldsymbol{w}}) = \int_{\widehat{F}} \widehat{\boldsymbol{w}} \cdot \widehat{\boldsymbol{q}} \, \mathrm{d}\widehat{\boldsymbol{r}}, \qquad \forall \widehat{\boldsymbol{q}} \in \widehat{\mathcal{Q}}_p^{wf}, \tag{6.112b}$$

$$\widehat{M}_p^{wv}(\widehat{\boldsymbol{w}}) = \int_{\widehat{V}} \widehat{\boldsymbol{w}} \cdot \widehat{\boldsymbol{q}} \, \mathrm{d}\widehat{\boldsymbol{r}}, \qquad \forall \widehat{\boldsymbol{q}} \in \widehat{\mathcal{Q}}_p^{wv}. \tag{6.112c}$$

These moments induce subspaces associated to the edges $\widehat{\mathcal{W}}_{N,p}^{e}$, faces $\widehat{\mathcal{W}}_{N,p}^{f}$, and volume $\widehat{\mathcal{W}}_{N,p}^{v}$ via the smooth extensions of the edge- (6.20), face- (6.21), and volume-

associated (6.22) test functions to the entire domain. The direct sum of the subspaces provides the approximation space

$$\widehat{\mathcal{W}}_{N,p} = \widehat{\mathcal{W}}^{e}_{N,p} \oplus \widehat{\mathcal{W}}^{f}_{N,p} \oplus \widehat{\mathcal{W}}^{v}_{N,p} = (\mathbb{Q}_{p,p,p})^3 \tag{6.113}$$

with the dimension of

$$\dim(\widehat{\mathcal{W}}_{N,p}) = 3(p+1)^3. \tag{6.114}$$

The relation to the complete-order serendipity spaces is

$$\widehat{\mathcal{W}}_{S,p} \subset \widehat{\mathcal{W}}_{N,p}. \tag{6.115}$$

Hence, the same (6.26) error estimation applies for affine refinements of the exact geometry mappings (this includes the parallelotope meshes), i.e.,

$$\left\|\boldsymbol{w} - \pi^{w}_{p,h}\boldsymbol{w}\right\|_{\boldsymbol{L}^2(\Omega)} \leq Ch^{p+1}, \tag{6.116a}$$

$$\left\|\boldsymbol{w} - \pi^{w}_{p,h}\boldsymbol{w}\right\|_{\boldsymbol{H}(\mathrm{curl},\Omega)} \leq Ch^{p}. \tag{6.116b}$$

However, this rate changes for general straight-sided geometry mappings and non-affine refinements. One can see that the sufficient condition (6.81) of the $\boldsymbol{L}^2$-norm convergence is met, whereas the $\boldsymbol{H}(\mathrm{curl})$-norm condition (6.110) is not. Hence, the rate of convergence suffers a degrade of one order compared to (6.116b) in $\boldsymbol{H}(\mathrm{curl})$-norm, i.e.,

$$\left\|\boldsymbol{w} - \pi^{w}_{p,h}\boldsymbol{w}\right\|_{\boldsymbol{L}^2(\Omega)} \leq Ch^{p+1}, \tag{6.117a}$$

$$\left\|\boldsymbol{w} - \pi^{w}_{p,h}\boldsymbol{w}\right\|_{\boldsymbol{H}(\mathrm{curl},\Omega)} \leq Ch^{p-1}. \tag{6.117b}$$

In the case of curvilinear elements and non-affine refinements, the rate of convergence would degrade even more for both norms due to the general requirements of Table 3.2.

6.3.4 Mixed-Order Nédélec Element

In this section, a definition is given for the mixed-order Nédélec element. The reference domain is a cube with unit length edges. The spaces of test functions are defined as

- $\widehat{\mathcal{Q}}_p^{we} = \widehat{\boldsymbol{e}}_1 \mathbb{Q}_{p-1}$,
- $\widehat{\mathcal{Q}}_p^{wf} = \widehat{\boldsymbol{e}}_1 \mathbb{Q}_{p-1,p-2} \oplus \widehat{\boldsymbol{e}}_2 \mathbb{Q}_{p-2,p-1}$,
- $\widehat{\mathcal{Q}}_p^{wv} = \widehat{\boldsymbol{e}}_1 \mathbb{Q}_{p-1,p-2,p-2} \oplus \widehat{\boldsymbol{e}}_2 \mathbb{Q}_{p-2,p-1,p-2} \oplus \widehat{\boldsymbol{e}}_3 \mathbb{Q}_{p-2,p-2,p-1}$,

with the dimensions of

$$12 \cdot \dim(\widehat{\mathcal{Q}}_p^{we}) = 12(p), \tag{6.118a}$$

$$6 \cdot \dim(\widehat{\mathcal{Q}}_p^{wf}) = 12p(p-1), \tag{6.118b}$$

$$1 \cdot \dim(\widehat{\mathcal{Q}}_p^{wv}) = 3p(p-1)^2. \tag{6.118c}$$

The corresponding moments are given as

$$\widehat{M}_p^{we}(\widehat{\boldsymbol{w}}) = \int_{\widehat{E}} \widehat{\boldsymbol{w}} \cdot \widehat{\boldsymbol{q}} \, \mathrm{d}\widehat{\boldsymbol{r}}, \qquad \forall \widehat{\boldsymbol{q}} \in \widehat{\mathcal{Q}}_p^{we}, \tag{6.119a}$$

$$\widehat{M}_p^{wf}(\widehat{\boldsymbol{w}}) = \int_{\widehat{F}} \widehat{\boldsymbol{w}} \cdot \widehat{\boldsymbol{q}} \, \mathrm{d}\widehat{\boldsymbol{r}}, \qquad \forall \widehat{\boldsymbol{q}} \in \widehat{\mathcal{Q}}_p^{wf}, \tag{6.119b}$$

$$\widehat{M}_p^{wv}(\widehat{\boldsymbol{w}}) = \int_{\widehat{V}} \widehat{\boldsymbol{w}} \cdot \widehat{\boldsymbol{q}} \, \mathrm{d}\widehat{\boldsymbol{r}}, \qquad \forall \widehat{\boldsymbol{q}} \in \widehat{\mathcal{Q}}_p^{wv}. \tag{6.119c}$$

Similar to the complete-order case, these moments induce subspaces associated with the edges $\widehat{\mathcal{W}}_{NI,p}^e$, faces $\widehat{\mathcal{W}}_{NI,p}^f$, and volume $\widehat{\mathcal{W}}_{NI,p}^v$, whose direct sum gives the approximation space

$$\widehat{\mathcal{W}}_{NI,p} = \widehat{\mathcal{W}}_{NI,p}^e \oplus \widehat{\mathcal{W}}_{NI,p}^f \oplus \widehat{\mathcal{W}}_{NI,p}^v = \begin{bmatrix} \mathbb{Q}_{p-1,p,p} \\ \mathbb{Q}_{p,p-1,p} \\ \mathbb{Q}_{p,p,p-1} \end{bmatrix} \tag{6.120}$$

with the dimension of

$$\dim(\widehat{\mathcal{W}}_{NI,p}) = 3p(p+1)^2. \tag{6.121}$$

The important relations to the other finite element approximation spaces are

$$\widehat{\mathcal{W}}_{NI,p} \subset \widehat{\mathcal{W}}_{N,p}, \tag{6.122a}$$

$$\widehat{\mathcal{W}}_{SI,p} \subset \widehat{\mathcal{W}}_{NI,p} \subset \widehat{\mathcal{W}}_{T,p}. \tag{6.122b}$$

Therefore, the mixed-order Nédélec element yields the same rate of convergence for the affine refinement of the exact geometry mappings as the mixed-order serendipity element $\widehat{\mathcal{W}}_{SI,p}$ (6.27), (6.28), i.e.,

$$\left\|\boldsymbol{w} - \pi^{w}_{p,h}\boldsymbol{w}\right\|_{\boldsymbol{L}^2(\Omega)} \leq Ch^p, \tag{6.123a}$$

$$\left\|\boldsymbol{w} - \pi^{w}_{p,h}\boldsymbol{w}\right\|_{\boldsymbol{H}(\mathrm{curl},\Omega)} \leq Ch^p. \tag{6.123b}$$

For general straight-sided elements (and non-affine refinements), the $\boldsymbol{L}^2$-norm convergence rate is unaffected due to (6.81). However, the convergence rate of the $\boldsymbol{H}(\mathrm{curl})$-norm degrades by one order, i.e.,

$$\left\|\boldsymbol{w} - \pi^{w}_{p,h}\boldsymbol{w}\right\|_{\boldsymbol{L}^2(\Omega)} \leq Ch^p, \tag{6.124a}$$

$$\left\|\boldsymbol{w} - \pi^{w}_{p,h}\boldsymbol{w}\right\|_{\boldsymbol{H}(\mathrm{curl},\Omega)} \leq Ch^{p-1}, \tag{6.124b}$$

since the condition (6.110) is not met. Similar results were shown by several authors [FGM11], [ABB15], [BD13b]. The advantage over the complete-order element is the smaller space, which produces fewer number of unknowns for the same convergence rate in $\boldsymbol{H}(\mathrm{curl})$-norm.

6.4 Hierarchical Basis Functions

Hierarchical basis functions are widely used for their flexibility in degree elevation, since there is no need to use a completely different basis for the different orders. Only the higher, incremental order functions need to be added to the lower-order basis, resulting in faster assembly times and minimizing the storage requirements. The second advantage of hierarchical basis functions is that they allow the usage of different order basis functions within one mesh without affecting conformity. For an element with higher-order neighbors, only the incremental order functions that belong to the common interfaces need to be included. In the following, a set of hierarchical basis functions is provided, capable of spanning most of the finite element spaces discussed in the previous sections.

The basis functions are defined on the reference element in terms of $x_i(\widehat{\boldsymbol{r}})$, $y_j(\widehat{\boldsymbol{r}})$, and $z_k(\widehat{\boldsymbol{r}})$, that are given in equation (6.1). On the global element, these are functions of the global coordinates $x_i(\boldsymbol{r}) = x_i(\widehat{\boldsymbol{r}}) \circ \boldsymbol{g}^{-1}$, $y_j(\boldsymbol{r}) = y_j(\widehat{\boldsymbol{r}}) \circ \boldsymbol{g}^{-1}$, and $z_k(\boldsymbol{r}) = z_k(\widehat{\boldsymbol{r}}) \circ \boldsymbol{g}^{-1}$, due to the inverse geometry mapping $\widehat{\boldsymbol{r}} = \boldsymbol{g}^{-1}(\boldsymbol{r})$. The construction of basis functions is based on Jacobi polynomials due to their orthogonality properties. Let $P_q^{ij}(x_1)$ denote the Jacobi polynomial of order q with the indices i and j. This is a classical orthogonal polynomial over the interval of $[-1, 1]$. For present purposes, the domain of orthogonality is shifted to $[0, 1]$, which gives rise to the following definition of the shifted Jacobi polynomial $\mathcal{P}_q^{ij}$:

$$\mathcal{P}_q^{ij}(x_1) = P_q^{ij}(2x_1 - 1). \tag{6.125}$$

Hence, the orthogonality relation for arbitrary functions $v_p(x_1) \in \mathbb{P}_p$ and $f_q(x_1) \in \mathcal{P}_q^{ij}(x_1)$ yields

$$\int_0^1 v_p(x_1)(1 - x_1)^i x_1^j f_q(x_1)\,\mathrm{d}x_1 = \int_0^1 v_p(x_1) x_0^i x_1^j f_q(x_1)\,\mathrm{d}x_1 = 0, \quad q > p, \tag{6.126}$$

with respect to the weighting function of $x_0^i x_1^j$. The next step is to provide the basis functions in a structured way. Since the finite element moments are determined the same way on all nodes, edges, and faces, it is reasonable to provide the node-associated functions for one general node, the edge-associated functions for one general edge, and the face-associated functions for one general face. Let the general variables be denoted by a_i, b_j, and c_k, which are combinations of x_i, y_j, and z_k variables. A general node of a hexahedral element is associated with the function $a_i b_j c_k$, since it has a unit value on the associated node. Similarly, a generic edge is parametrized by the variable c_1 and associated with the function $a_i b_j$, since it has a unit value over the edge. A generic face is parametrized by the variables a_1 and b_1 and associated with a function c_k, since it has a unit value over the face.

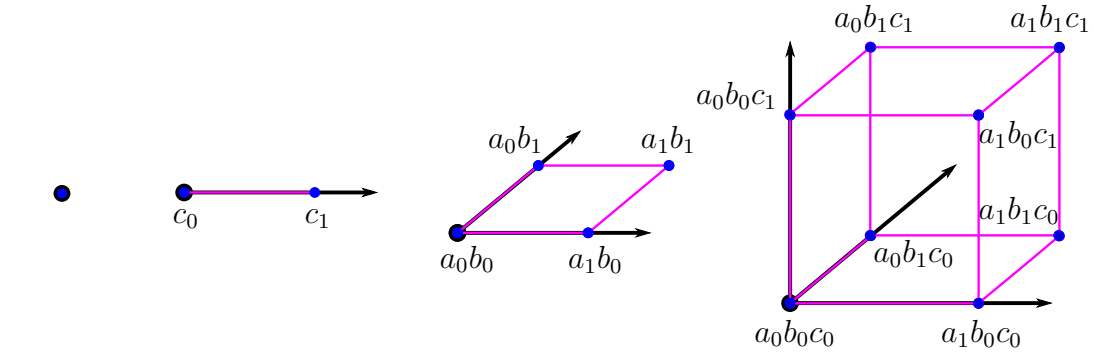


Figure 6.4: Associated functions of a general node, edge, and face.

Hence, the possible combinations are

$$
\begin{aligned}
(\text{ node } a_ib_jc_k) &= (\text{ node } x_iy_jz_k),\\
(\text{ edge } a_ib_j,\ \text{variable: } c_1) &\in \left\{ \begin{matrix} (\text{ edge } y_jz_k,\ \text{variable: } x_1)\\ (\text{ edge } x_iz_k,\ \text{variable: } y_1)\\ (\text{ edge } x_iy_j,\ \text{variable: } z_1) \end{matrix} \right\},\\
(\text{ face } c_k,\ \text{variables: } a_1, b_1) &\in \left\{ \begin{matrix} (\text{ face } z_k,\ \text{variables: } x_1, y_1)\\ (\text{ face } y_j,\ \text{variables: } x_1, z_1)\\ (\text{ face } x_i,\ \text{variables: } y_1, z_1) \end{matrix} \right\},\\
(\text{ volume,\ variables: } a_i, b_j, c_k) &= (\text{ volume,\ variables: } x_i, y_j, z_k).
\end{aligned}
$$

The graphical representation of these general partitions are shown in Figure 6.4. It can be seen that these general nodes, edges, faces, and volume are identical to the previously specified association types in Figure 6.1, Figure 6.2, or Figure 6.3. However, these notations allow a more compact representation of bases with large dimensions. With these notations, the scalar basis functions are shown in Table 6.1 and their dimensions in Table 6.2. The H^1-conforming approximation spaces of $\mathcal{V}_{S,p}$ and $\mathcal{V}_{T,p}$ contain hierarchical subspaces

$$\mathcal{V}_{(\cdot),p} = \widetilde{\mathcal{V}}_{(\cdot),1} \oplus \cdots \oplus \widetilde{\mathcal{V}}_{(\cdot),p}, \qquad (\cdot) \in \{S, T\}, \tag{6.128a}$$

that are associated with the nodes, edges, faces, and volume of the element

$$\widetilde{\mathcal{V}}_{(\cdot)} = \widetilde{\mathcal{V}}^n_{(\cdot)} \oplus \widetilde{\mathcal{V}}^e_{(\cdot)} \oplus \widetilde{\mathcal{V}}^f_{(\cdot)} \oplus \widetilde{\mathcal{V}}^v_{(\cdot)}, \qquad (\cdot) \in \{S, T\}. \tag{6.129}$$

The basis functions of the $\boldsymbol{H}(\text{curl})$-conforming spaces $\mathcal{A}_{S,p}$, $\mathcal{A}_{N,p}$, and $\mathcal{A}_{T,p}$ are shown in Table 6.4 and their dimensions in Table 6.3. Each space is hierarchical in terms of the polynomial order

$$\mathcal{A}_{(\cdot),p} = \widetilde{\mathcal{A}}_{(\cdot),1} \oplus \cdots \oplus \widetilde{\mathcal{A}}_{(\cdot),p} \qquad (\cdot) \in \{S, N, T\}, \tag{6.130}$$

as well as containing subspaces, which are associated with the edges, faces, and volume:

$$\widetilde{\mathcal{A}}_{(\cdot)} = \widetilde{\mathcal{A}}^e_{(\cdot)} \oplus \widetilde{\mathcal{A}}^f_{(\cdot)} \oplus \widetilde{\mathcal{A}}^v_{(\cdot)}, \qquad (\cdot) \in \{S, N, T\}. \tag{6.131}$$

Table 6.1: H^1-conforming basis functions.

$\widetilde{\mathcal{V}}_{T,p}$	$\widetilde{\mathcal{V}}_{S,p}$	Assoc.	Num.	Basis functions
1	1	node $a_ib_jc_k$	1	$v_{ijk} = a_ib_jc_k,$
2	2	edge a_ib_j	1	$v^{\gamma}_{ij} = c_0c_1a_ib_j\mathcal{P}^{1,1}_{\gamma}(c_1), \vert\gamma = 0$
2	4	face c_k	1	$v^{\alpha\beta}_{k} = a_0a_1b_0b_1c_k\mathcal{P}^{1,1}_{\alpha}(a_1)\mathcal{P}^{1,1}_{\beta}(b_1), \vert\alpha,\beta = 0$
2	6	volume	1	$v^{\alpha\beta\gamma} = a_0a_1b_0b_1c_0c_1\mathcal{P}^{1,1}_{\alpha}(a_1)\mathcal{P}^{1,1}_{\beta}(b_1)\mathcal{P}^{1,1}_{\gamma}(c_1), \vert\alpha,\beta,\gamma = 0$
3	3	edge a_ib_j	1	$v^{1}_{ij},$
3	$\alpha+\beta+4$	face c_k	3	$v^{\alpha\beta}_{k}, \quad 0 \le \alpha,\beta \le 1$ while $\alpha = 1$ or $\beta = 1$
3	$\alpha+\beta+6$	volume	7	$v^{\alpha\beta\gamma}, \quad 0 \le \alpha,\beta,\gamma \le 1$ while $\alpha = 1$ or $\beta = 1$ or $\gamma = 1$
4	4	edge a_ib_j	1	$v^{2}_{ij},$
4	$\alpha+\beta+4$	face c_k	5	$v^{\alpha\beta}_{k}, \quad 0 \le \alpha,\beta \le 2$ while $\alpha = 2$ or $\beta = 2$
4	$\alpha+\beta+6$	volume	19	$v^{\alpha\beta\gamma}, \quad 0 \le \alpha,\beta,\gamma \le 2$ while $\alpha = 2$ or $\beta = 2$ or $\gamma = 2$

Table 6.2: Dimensions of the H^1-conforming bases.

p	$\mathcal{V}^n_{T,p}$	$\mathcal{V}^e_{T,p}$	$\mathcal{V}^f_{T,p}$	$\mathcal{V}^v_{T,p}$	$\mathcal{V}_{T,p}$	$\mathcal{V}^n_{S,p}$	$\mathcal{V}^e_{S,p}$	$\mathcal{V}^f_{S,p}$	$\mathcal{V}^v_{S,p}$	$\mathcal{V}_{S,p}$
1	8	0	0	0	8	8	0	0	0	8
2	8	12	6	1	27	8	12	0	0	20
3	8	24	24	8	64	8	24	0	0	32
4	8	36	54	27	125	8	36	6	0	50

Table 6.3: Dimensions of the range space $\boldsymbol{H}$(curl)-conforming bases.

p	$\mathcal{A}^e_{T,p}$	$\mathcal{A}^f_{T,p}$,	$\mathcal{A}^v_{T,p}$	$\mathcal{A}^e_{N,p}$	$\mathcal{A}^f_{N,p}$	$\mathcal{A}^v_{N,p}$	$\mathcal{A}^e_{S,p}$	$\mathcal{A}^f_{S,p}$	$\mathcal{A}^v_{S,p}$
1	12	12	3	12	0	0	12	0	0
2	12	42	23	12	18	5	12	12	0
3	12	84	73	12	48	28	12	30	0
4	12	138	165	12	90	81	12	54	3

Table 6.4: Range space $\boldsymbol{H}(\mathrm{curl})$-conforming basis functions.

$\widetilde{\mathcal{A}}_{T,p}$	$\widetilde{\mathcal{A}}_{N,p}$	$\widetilde{\mathcal{A}}_{S,p}$	Assoc.	Num.	Basis functions
1	1	1	edge a_ib_j	1	$a_ib_j\nabla c_1,$
1	2	2	face c_k	1	$\boldsymbol{w}_k^{a,\alpha\beta} = b_0b_1c_k\mathcal{P}_\alpha^{0,0}(a_1)\mathcal{P}_\beta^{1,1}(b_1)\nabla a_1, \mid \alpha,\beta = 0$
1	2	2	face c_k	1	$\boldsymbol{w}_k^{b,\alpha\beta} = a_0a_1c_k\mathcal{P}_\alpha^{1,1}(a_1)\mathcal{P}_\beta^{0,0}(b_1)\nabla b_1, \mid \alpha,\beta = 0$
1	2	4	volume	1	$\boldsymbol{w}^{a,\alpha\beta\gamma} = b_0b_1c_0c_1\mathcal{P}_\alpha^{0,0}(a_1)\mathcal{P}_\beta^{1,1}(b_1)\mathcal{P}_\gamma^{1,1}(c_1)\nabla a_1, \mid \alpha,\beta,\gamma = 0$
1	2	4	volume	1	$\boldsymbol{w}^{b,\alpha\beta\gamma} = a_0a_1c_0c_1\mathcal{P}_\alpha^{1,1}(a_1)\mathcal{P}_\beta^{0,0}(b_1)\mathcal{P}_\gamma^{1,1}(c_1)\nabla b_1, \mid \alpha,\beta,\gamma = 0$
1	2	4	volume	1	$\boldsymbol{w}^{c,\alpha\beta\gamma} = a_0a_1b_0b_1\mathcal{P}_\alpha^{1,1}(a_1)\mathcal{P}_\beta^{1,1}(b_1)\mathcal{P}_\gamma^{0,0}(c_1)\nabla c_1, \mid \alpha,\beta,\gamma = 0$
2	2	3	face c_k	1	$\boldsymbol{w}_k^{a,10} - \boldsymbol{w}_k^{b,01},$
2	2	5	volume	1	$\boldsymbol{w}^{a,100} - \boldsymbol{w}^{b,010},$
2	2	5	volume	1	$\boldsymbol{w}^{a,100} - \boldsymbol{w}^{c,001},$
2	3	$\alpha+\beta+2$	face c_k	2	$\boldsymbol{w}_k^{a,\alpha\beta}, \quad 0 \le \alpha \le 1,\ \beta = 1$
2	3	$\alpha+\beta+2$	face c_k	2	$\boldsymbol{w}_k^{b,\alpha\beta}, \quad 0 \le \beta \le 1,\ \alpha = 1$
2	3	$\alpha+\beta+4$	volume	6	$\boldsymbol{w}^{a,\alpha\beta\gamma}, \quad 0 \le \alpha,\beta,\gamma \le 1$ while $\beta = 1$ or $\gamma = 1$
2	3	$\alpha+\beta+4$	volume	6	$\boldsymbol{w}^{b,\alpha\beta\gamma}, \quad 0 \le \alpha,\beta,\gamma \le 1$ while $\alpha = 1$ or $\gamma = 1$
2	3	$\alpha+\beta+4$	volume	6	$\boldsymbol{w}^{c,\alpha\beta\gamma}, \quad 0 \le \alpha,\beta,\gamma \le 1$ while $\alpha = 1$ or $\beta = 1$
3	3	5	face c_k	1	$\boldsymbol{w}_k^{a,21} - \boldsymbol{w}_k^{b,12},$
3	3	7	volume	1	$\boldsymbol{w}^{a,210} - \boldsymbol{w}^{a,201},$
3	3	7	volume	1	$\boldsymbol{w}^{b,021} - \boldsymbol{w}^{b,120},$
3	3	7	volume	1	$\boldsymbol{w}^{a,102} - \boldsymbol{w}^{a,012},$
3	3	8	volume	1	$\boldsymbol{w}^{a,211} - \boldsymbol{w}^{b,121},$
3	3	8	volume	1	$\boldsymbol{w}^{a,211} - \boldsymbol{w}^{c,112},$
3	4	$\alpha+\beta+2$	face c_k	3	$\boldsymbol{w}_k^{a,\alpha\beta}, \quad 0 \le \alpha \le 2,\ \beta = 2$
3	4	$\alpha+\beta+2$	face c_k	3	$\boldsymbol{w}_k^{b,\alpha\beta}, \quad 0 \le \beta \le 2,\ \alpha = 2$
3	4	$\alpha+\beta+4$	volume	15	$\boldsymbol{w}^{a,\alpha\beta\gamma}, \quad 0 \le \alpha,\beta,\gamma \le 2$ while $\beta = 2$ or $\gamma = 2$
3	4	$\alpha+\beta+4$	volume	15	$\boldsymbol{w}^{b,\alpha\beta\gamma}, \quad 0 \le \alpha,\beta,\gamma \le 2$ while $\alpha = 2$ or $\gamma = 2$
3	4	$\alpha+\beta+4$	volume	15	$\boldsymbol{w}^{c,\alpha\beta\gamma}, \quad 0 \le \alpha,\beta,\gamma \le 2$ while $\alpha = 2$ or $\beta = 2$
4	4	7	face c_k	1	$\boldsymbol{w}_k^{a,32} - \boldsymbol{w}_k^{b,23},$
4	4	9	volume	1	$\boldsymbol{w}^{a,320} - \boldsymbol{w}^{a,302},$
4	4	9	volume	1	$\boldsymbol{w}^{b,230} - \boldsymbol{w}^{b,032},$
4	4	9	volume	1	$\boldsymbol{w}^{c,023} - \boldsymbol{w}^{c,203},$
4	4	10	volume	1	$\boldsymbol{w}^{a,321} - \boldsymbol{w}^{a,312},$
4	4	10	volume	1	$\boldsymbol{w}^{b,231} - \boldsymbol{w}^{b,132},$
4	4	10	volume	1	$\boldsymbol{w}^{c,123} - \boldsymbol{w}^{c,213},$
4	4	11	volume	1	$\boldsymbol{w}^{a,322} - \boldsymbol{w}^{b,232},$
4	4	11	volume	1	$\boldsymbol{w}^{a,322} - \boldsymbol{w}^{c,223},$
4	5	$\alpha+\beta+2$	face c_k	4	$\boldsymbol{w}_k^{a,\alpha\beta}, \quad 0 \le \alpha \le 3,\ \beta = 3$
4	5	$\alpha+\beta+2$	face c_k	4	$\boldsymbol{w}_k^{b,\alpha\beta}, \quad 0 \le \beta \le 3,\ \alpha = 3$
4	5	$\alpha+\beta+4$	volume	28	$\boldsymbol{w}^{a,\alpha\beta\gamma}, \quad 0 \le \alpha,\beta,\gamma \le 3$ while $\beta = 3$ or $\gamma = 3$
4	5	$\alpha+\beta+4$	volume	28	$\boldsymbol{w}^{b,\alpha\beta\gamma}, \quad 0 \le \alpha,\beta,\gamma \le 3$ while $\alpha = 3$ or $\gamma = 3$
4	5	$\alpha+\beta+4$	volume	28	$\boldsymbol{w}^{c,\alpha\beta\gamma}, \quad 0 \le \alpha,\beta,\gamma \le 3$ while $\alpha = 3$ or $\beta = 3$

This single set of basis functions for each subspace enables most previously described finite element spaces for $\boldsymbol{H}(\mathrm{curl})$-conforming functions. The mixed-order serendipity space reads

$$\mathcal{W}_{SI,p} = \widetilde{\mathcal{W}}_{SI,1} \oplus \cdots \oplus \widetilde{\mathcal{W}}_{SI,p}, \tag{6.132a}$$
$$\widetilde{\mathcal{W}}_{SI,1} = \widetilde{\mathcal{A}}_{S,1}, \tag{6.132b}$$
$$\widetilde{\mathcal{W}}_{SI,p} = \widetilde{\mathcal{A}}_{S,p} \oplus \nabla\widetilde{\mathcal{V}}_{S,p}, \quad \text{for } p > 1. \tag{6.132c}$$

The mixed-order Nédélec space reads

$$\mathcal{W}_{NI,p} = \widetilde{\mathcal{W}}_{NI,1} \oplus \cdots \oplus \widetilde{\mathcal{W}}_{NI,p}, \tag{6.133a}$$
$$\widetilde{\mathcal{W}}_{NI,1} = \widetilde{\mathcal{A}}_{N,1}, \tag{6.133b}$$
$$\widetilde{\mathcal{W}}_{NI,p} = \widetilde{\mathcal{A}}_{N,p} \oplus \nabla\widetilde{\mathcal{V}}_{T,p}, \quad \text{for } p > 1. \tag{6.133c}$$

The tensor product space for general straight-sided geometries reads

$$\mathcal{W}_{SI,p} = \widetilde{\mathcal{W}}_{T,1} \oplus \cdots \oplus \widetilde{\mathcal{W}}_{T,p}, \tag{6.134a}$$
$$\widetilde{\mathcal{W}}_{T,1} = \widetilde{\mathcal{A}}_{T,1}, \tag{6.134b}$$
$$\widetilde{\mathcal{W}}_{T,p} = \widetilde{\mathcal{A}}_{T,p} \oplus \nabla\widetilde{\mathcal{V}}_{T,p}, \quad \text{for } p > 1. \tag{6.134c}$$

Table 6.5: Dimensions of $\boldsymbol{H}(\mathrm{curl})$-conforming bases.

p	$\mathcal{W}^e_{T,p}$	$\mathcal{W}^f_{T,p}$,	$\mathcal{W}^v_{T,p}$	$\mathcal{W}_{T,p}$	$\mathcal{W}^e_{NI,p}$	$\mathcal{W}^f_{NI,p}$	$\mathcal{W}^v_{NI,p}$	$\mathcal{W}_{NI,p}$	$\mathcal{W}^e_{SI,p}$	$\mathcal{W}^f_{SI,p}$	$\mathcal{W}^v_{SI,p}$	$\mathcal{W}_{SI,p}$
1	12	12	3	27	12	0	0	12	12	0	0	12
2	24	48	24	96	24	24	6	54	24	12	0	36
3	36	108	81	225	36	72	36	144	36	30	0	66
4	48	192	192	432	48	144	108	300	48	60	3	111

The main orthogonality property of the bases is provided by the Jacobi polynomials. These properties can be formalized in a compact form using the corresponding p-order interpolation operator of the approximation space, which acts on a higher, q-order basis function as

$$\pi^v_p v_q = 0, \qquad v_q \in \mathcal{V}_{(\cdot),q} \text{ for } q > p, \tag{6.135a}$$
$$\pi^w_p \nabla v_q = \mathbf{0}, \qquad \nabla v_q \in \mathcal{W}_{(\cdot),q} \text{ for } q > p. \tag{6.135b}$$

Here, the second relation follows from the first one using the commutation property of the interpolation operators.

Similar orthogonality properties apply to the vector functions

$$\pi_p^w \boldsymbol{w}_q = \boldsymbol{0}, \qquad \boldsymbol{w}_q \in \mathcal{W}_{(\cdot),q} \text{ for } q > p, \tag{6.136a}$$

$$\pi_p^d \nabla \times \boldsymbol{w}_q = \boldsymbol{0}, \qquad \nabla \times \boldsymbol{w}_q \in \mathcal{D}_{(\cdot),q} \text{ for } q > p, \tag{6.136b}$$

$$\pi_p^d \nabla \times \boldsymbol{w}_q = \boldsymbol{0}, \qquad \boldsymbol{w}_q \in \nabla \mathcal{V}_{(\cdot),q}. \tag{6.136c}$$

Here, (6.136b) follows from (6.136a) through the commutation property of the interpolation operators. The third relation is the result of the separated set of basis functions for the null space of the curl operator. The consequences of the orthogonality and the distinguished null space functions are the advantageous numerical properties of the finite element matrices. For the demonstration, consider the normalized element mass matrices

$$[\boldsymbol{T}]_{ab} = \frac{1}{\|v_a\|_{L^2(\Omega_e)} \|v_b\|_{L^2(\Omega_e)}} \int_{\Omega_e} v_a v_b \, \mathrm{d}\boldsymbol{r}, \qquad \text{for } v_a, v_b \in \mathcal{V}_{(\cdot)}, \tag{6.137a}$$

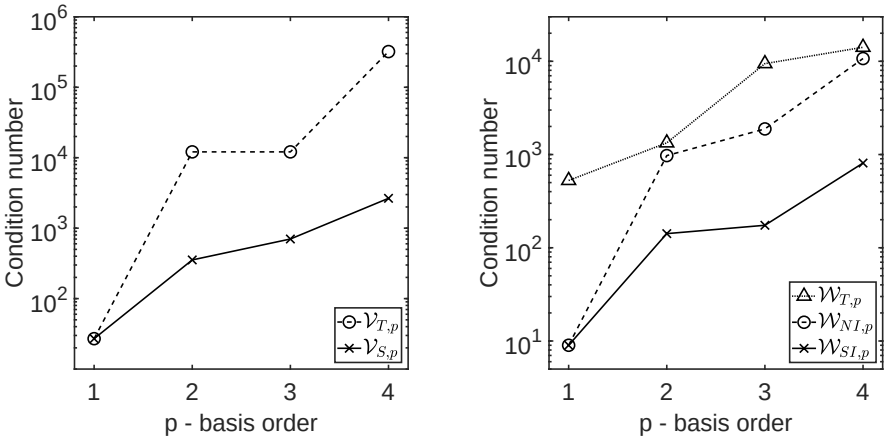
$$[\boldsymbol{T}]_{ab} = \frac{1}{\|\boldsymbol{w}_a\|_{\boldsymbol{L}^2(\Omega_e)} \|\boldsymbol{w}_b\|_{\boldsymbol{L}^2(\Omega_e)}} \int_{\Omega_e} \boldsymbol{w}_a \cdot \boldsymbol{w}_b \, \mathrm{d}\boldsymbol{r}, \qquad \text{for } \boldsymbol{w}_a, \boldsymbol{w}_b \in \mathcal{W}_{(\cdot)}, \tag{6.137b}$$

and the corresponding stiffness matrices

$$[\boldsymbol{S}]_{ab} = \frac{1}{\|v_a\|_{L^2(\Omega_e)} \|v_b\|_{L^2(\Omega_e)}} \int_{\Omega_e} \nabla v_a \cdot \nabla v_b \, \mathrm{d}\boldsymbol{r}, \qquad \text{for } v_a, v_b \in \mathcal{V}_{(\cdot)}, \tag{6.138a}$$

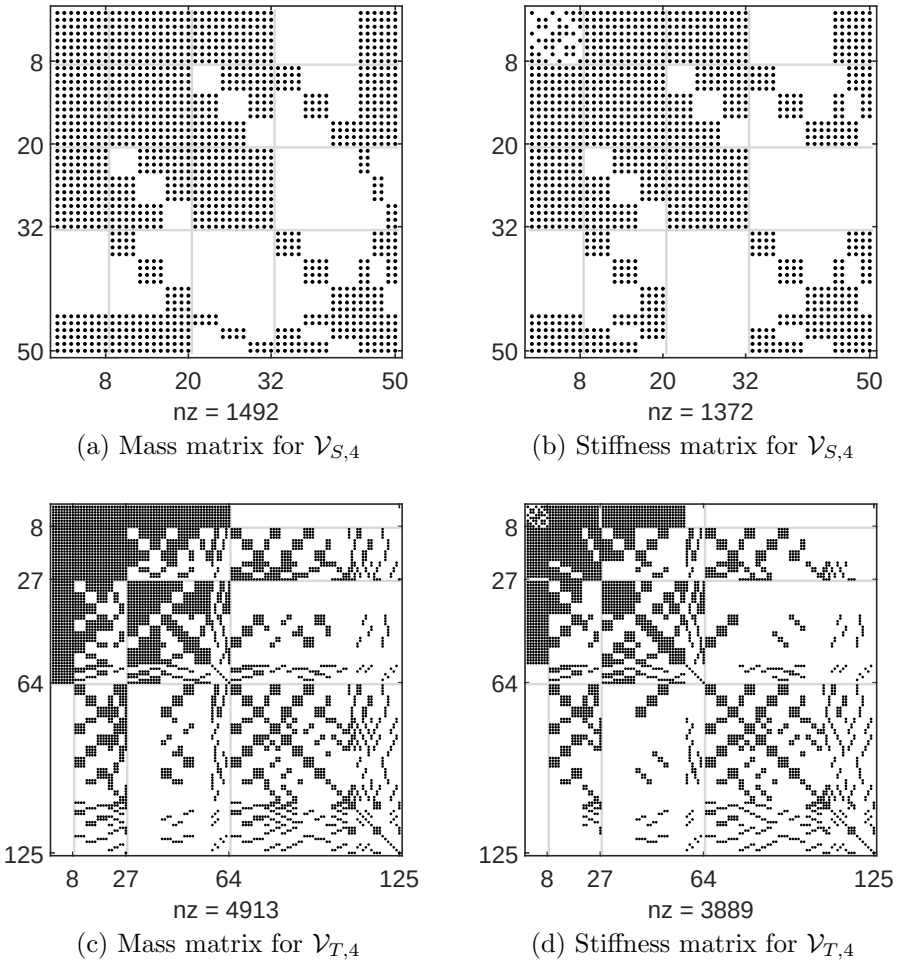
$$[\boldsymbol{S}]_{ab} = \frac{1}{\|\boldsymbol{w}_a\|_{\boldsymbol{L}^2(\Omega_e)} \|\boldsymbol{w}_b\|_{\boldsymbol{L}^2(\Omega_e)}} \int_{\Omega_e} \nabla \times \boldsymbol{w}_a \cdot \nabla \times \boldsymbol{w}_b \, \mathrm{d}\boldsymbol{r}, \quad \text{for } \boldsymbol{w}_a, \boldsymbol{w}_b \in \mathcal{W}_{(\cdot)}. \tag{6.138b}$$

The shape of the element is set to be a cube with unit length edges $\Omega_e = [0, 1]^3$. The calculated sparsity patterns are shown in Figure 6.6 and Figure 6.7. The zero blocks of the stiffness matrices in Figure 6.7 are provided by the separable higher-order subset of gradient functions, whose curl is zero independent of the geometry mapping. The condition numbers of the normalized mass matrices are shown in Figure 6.5 for the different finite element spaces, in terms of the basis order p. Evidently, the serendipity spaces are more advantageous in terms of these properties.



(a) H^1 bases (b) $\boldsymbol{H}(\text{curl})$ bases

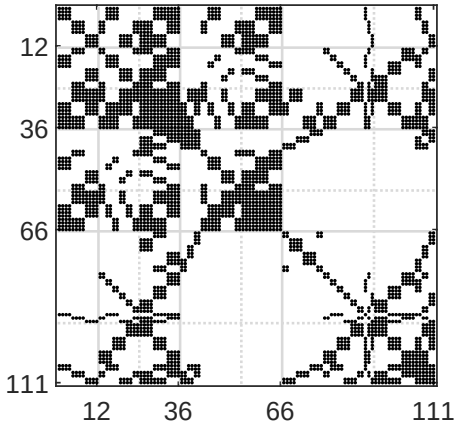
Figure 6.5: Condition number of mass matrices.



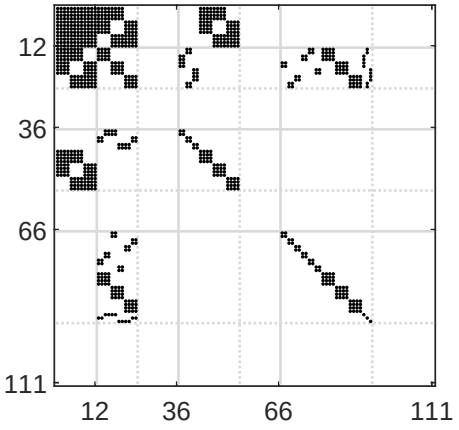
(a) Mass matrix for $\mathcal{V}_{S,4}$ (b) Stiffness matrix for $\mathcal{V}_{S,4}$

(c) Mass matrix for $\mathcal{V}_{T,4}$ (d) Stiffness matrix for $\mathcal{V}_{T,4}$

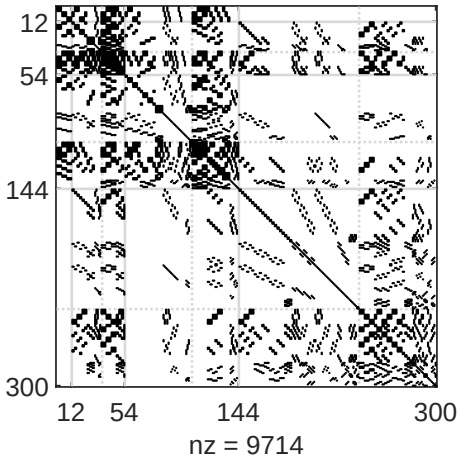
Figure 6.6: Sparsity patterns of the H^1 bases.



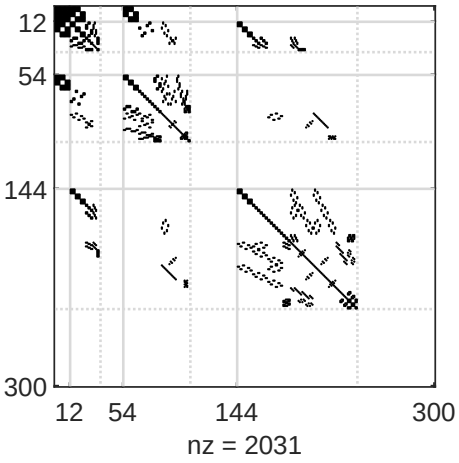
(a) Mass matrix for $\mathcal{W}_{SI,4}$



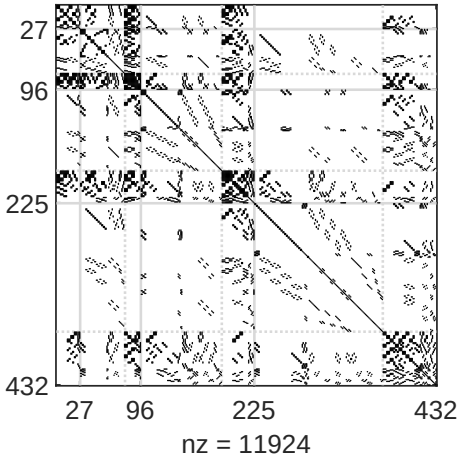
(b) Stiffness matrix for $\mathcal{W}_{SI,4}$



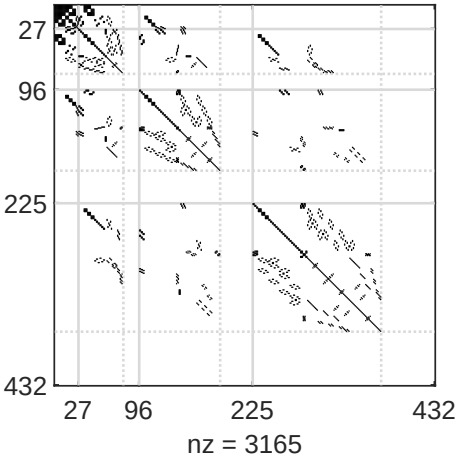
(c) Mass matrix for $\mathcal{W}_{NI,4}$



(d) Stiffness matrix for $\mathcal{W}_{NI,4}$



(e) Mass matrix for $\mathcal{W}_{T,4}$



(f) Stiffness matrix for $\mathcal{W}_{T,4}$

Figure 6.7: Sparsity patterns of the $\boldsymbol{H}(\mathrm{curl})$ bases.

6.4.1 Comparison to Other Basis Functions

The main advantage of the proposed single set of hierarchical basis functions is its capability to span various finite element spaces: $\mathcal{V}_{S,p}$, $\mathcal{V}_{T,p}$, $\mathcal{W}_{SI,p}$, $\mathcal{W}_{NI,p}$, and $\mathcal{W}_{T,p}$. Hence, it is possible to use different orders of elements and different finite element spaces within one mesh. To the author's knowledge, no such basis functions have existed so far. Therefore, only a short comparison is given for the basis functions of the specific finite element spaces.

In the case of the complete-order serendipity space $\mathcal{W}_{S,p}$, Arnold and Awanou [AA14] provided sets of nodes that are associated with the finite element moments. Therefore, explicit basis functions were not provided; however, one can interpret the locations as an interpolatory basis. The disadvantage of the space $\mathcal{W}_{S,p}$ is that it requires a larger number of unknowns for the same rate of convergence in the $\boldsymbol{H}$(curl)-norm than the mixed-order serendipity space $\mathcal{W}_{SI,p}$. Moreover, their interpolatory functions are not compatible with other finite element spaces and cannot be partitioned according to the null space.

Gillette and Kloefkorn proposed basis functions [GK19] for their trimmed serendipity space with the same dimension and convergence properties as the proposed basis for $\mathcal{W}_{SI,p}$. However, the finite element moments and the approximation space are different due to the difference in the highest-order functions. The definitions of the space $\mathcal{W}_{SI,p}$ and finite element moments are designed to construct hierarchical basis functions, which can be partitioned according to the null space. Unlike the basis of Gillette and Kloefkorn, the proposed basis is orthogonal with respect to the interpolation operator and compatible with other finite element spaces.

Next, let us consider the scalar serendipity space $\mathcal{V}_{S,p}$ for H^1-conforming functions. Historically, the 8-point serendipity basis functions have been presented first for quadrilateral (face) elements [SB91, Chapter 6] [ZK00]. These are interpolatory basis functions that span the proposed face-associated serendipity space $\mathcal{V}^f_{S,p}$ for $p = 2$. Later on, a generalized framework was given by Arnold and Awanou based on differential forms [AA11]. They showed interpolation locations for a higher-order hexahedra, which can be interpreted as an interpolatory basis. However, explicit basis functions were not provided. Gillette and Kloefkorn also proposed basis functions [GK19] for $\mathcal{V}_{S,p}$ that are hierarchical. Since all the above mentioned bases span the same space, their approximation property is the same. The main difference between these bases and the proposed hierarchical basis is the orthogonality with respect to the interpolation operator.

Next, let us consider the tensor product space $\mathcal{W}_{T,p}$ for $\boldsymbol{H}$(curl)-conforming functions. The first-order space and the corresponding basis functions were reported in [FGM11]. After that, Bergot and Duruflé presented hierarchical basis functions with the same orthogonality property as the proposed functions. However, no subset of

null space functions is available for $p > 1$. The advantages of the present basis over that of Bergot and Duruflé are the much sparser stiffness matrix, the straightforward way for the separation of the higher-order gradient subspace, and the compatibility with other finite element spaces $\mathcal{W}_{SI,p}$, $\mathcal{W}_{NI,p}$, and $\mathcal{W}_{T,p}$.

Finally, the best-known spaces are considered: the tensor product scalar space $\mathcal{V}_{T,p}$ and the mixed-order Nédélec space $\mathcal{W}_{NI,p}$. Many different bases have been developed for these spaces with similar numerical properties. Recently, Martin et al. [C5] proposed a hierarchical basis for $\mathcal{W}_{NI,p}$ which, similar to the proposed basis, reuse a scalar basis for the gradient subspace $\nabla\mathcal{V}_{T,p} \subset \mathcal{W}_{NI,p}$ and yield similar orthogonality properties. The difference between the basis in [C5] from the proposed functions is the lack of compatibility with the other finite element spaces $\mathcal{W}_{SI,p}$ and $\mathcal{W}_{T,p}$.

6.5 Mixed-Order Mixed-Space Finite Elements

At this point, it is possible to implement various finite element spaces using basis functions from the same set. Thus, the only open problem is the choice of the finite element space. The proposed solution is to use the smallest appropriate space, which results in an $\mathcal{O}(h^p)$ rate of convergence in the natural norms. The present discussion is restricted to computational domains with straight-sided boundaries, allowing to mesh the geometry with straight-sided elements without introducing any geometry error. In practice, most meshes are such that the major number of the elements are parallelepiped, described by affine geometry mappings, or where affine refinement is acceptable. The domain of all such elements is denoted by Ω_A. Over all the elements of Ω_A, the serendipity spaces $\mathcal{V}_{S,p}$ and $\mathcal{W}_{SI,p}$ are sufficient for the correct, asymptotic rate of convergence. The other tensor product spaces, $\mathcal{V}_{T,p}$ and $\mathcal{W}_{T,p}$, might provide better approximations but with the same convergence rate and higher numbers of unknowns. On the other hand, tensor product spaces are advantageous since they can provide the theoretical convergence rate for arbitrary straight-sided meshes without restrictions on the refinement. If the elements are not parallelepipeds and the way of refinement is unknown, the tensor product spaces of $\mathcal{V}_{T,p}$ and $\mathcal{W}_{T,p}$ should be used to preserve the desired asymptotic convergence rates. However, in most practical applications, such elements are only present in small numbers, typically close to the boundaries. Those domains with non-affine refinements are denoted by $\Omega_{\bar{A}}$.

These arguments motivate the mixed-space finite elements, utilizing different finite element spaces within one mesh, since the heterogeneous application of different spaces is much more efficient than using tensor product spaces on all elements. These finite elements are also mixed-order elements in the $\boldsymbol{H}(\mathrm{curl})$-conforming case, because the highest-order gradient subspace is not included. In order to preserve the conformity between elements of Ω_A and $\Omega_{\bar{A}}$, the functions, belonging to the common interfaces $\Omega_A \cap \Omega_{\bar{A}} = \partial\Omega_{A\bar{A}}$ need to be included in Ω_A. The compatibility of the proposed basis functions with both the tensor product space and the serendipity space results in incremental spaces:

$$\mathcal{V}_{T,p} = \mathcal{V}_{S,p} \oplus \widetilde{\mathcal{V}}^{e}_{S\to T,p} \oplus \widetilde{\mathcal{V}}^{f}_{S\to T,p} \oplus \widetilde{\mathcal{V}}^{v}_{S\to T,p}. \tag{6.139}$$

Here, $\widetilde{\mathcal{V}}^{e}_{S\to T,p}$ is the incremental edge-associated, $\widetilde{\mathcal{V}}^{f}_{S\to T,p}$ is the face-associated, and $\widetilde{\mathcal{V}}^{v}_{S\to T,p}$ is the volume-associated subspace. Using this decomposition, the definition of the H^1-conforming mixed-space reads

$$\mathcal{V}_{\mathrm{adapt},p} \begin{cases} \mathcal{V}_{T,p} & \text{on } \Omega_{\bar{A}}, \\ \mathcal{V}_{S,p} \oplus \widetilde{\mathcal{V}}^{e}_{S\to T,p}\left(\partial\Omega_{A\bar{A}}\right) \oplus \widetilde{\mathcal{V}}^{f}_{S\to T,p}\left(\partial\Omega_{A\bar{A}}\right) & \text{on } \Omega_A. \end{cases} \tag{6.140}$$

The notation $(\partial\Omega_{A\bar{A}})$ after the incremental spaces means that only those functions need to be included which are associated with the boundary $\partial\Omega_{A\bar{A}}$. The advantage

is that the asymptotic rate of convergence,

$$\left\|v - \pi^v_{p,h} v\right\|_{H^1(\Omega)} \leq Ch^p, \tag{6.141}$$

remains unchanged, whereas $\dim(\mathcal{V}_{\text{adapt},p}) \leq \dim(\mathcal{V}_{T,p})$. A similar definition is possible for the $\boldsymbol{H}(\text{curl})$-conforming functions. There, the decomposition of the tensor product space reads

$$\mathcal{W}_{T,p} = \mathcal{W}_{SI,p} \oplus \widetilde{\mathcal{W}}^e_{SI\to T,p} \oplus \widetilde{\mathcal{W}}^f_{SI\to T,p} \oplus \widetilde{\mathcal{W}}^v_{SI\to T,p}, \tag{6.142}$$

where $\widetilde{\mathcal{W}}^e_{SI\to T,p}$ is the incremental edge-associated, $\widetilde{\mathcal{W}}^f_{SI\to T,p}$ is the face-associated, and $\widetilde{\mathcal{W}}^v_{SI\to T,p}$ is the volume-associated subspace. Thus, the $\boldsymbol{H}(\text{curl})$-conforming mixed-order mixed-space reads

$$\mathcal{W}_{\text{adapt},p} \begin{cases} \mathcal{W}_{T,p} & \text{on } \Omega_{\bar{A}}, \\ \mathcal{W}_{SI,p} \oplus \widetilde{\mathcal{W}}^e_{SI\to T,p}\left(\partial\Omega_{A\bar{A}}\right) \oplus \widetilde{\mathcal{W}}^f_{SI\to T,p}\left(\partial\Omega_{A\bar{A}}\right) & \text{on } \Omega_A. \end{cases} \tag{6.143}$$

Note that the adaptive gradient subspace is always included in the adaptive $\boldsymbol{H}(\text{curl})$-conforming approximation space,

$$\nabla\mathcal{V}_{\text{adapt},p} \subset \mathcal{W}_{\text{adapt},p}, \tag{6.144}$$

by construction. Hence, the discrete de Rham complex, as well as all previously established error estimates, apply:

$$\left\|\boldsymbol{w} - \pi^w_{p,h}\boldsymbol{w}\right\|_{\boldsymbol{L}^2(\Omega)} \leq Ch^p, \tag{6.145a}$$

$$\left\|\boldsymbol{w} - \pi^w_{p,h}\boldsymbol{w}\right\|_{\boldsymbol{H}(\text{curl},\Omega)} \leq Ch^p. \tag{6.145b}$$

Definition 6.1 (Mixed-order mixed-space $\boldsymbol{H}(\text{curl})$-conforming element)**.**
The mixed-order mixed-space $\boldsymbol{H}(\text{curl})$-conforming element *is defined by the following three points:*

- *The reference cube $\widehat{\Omega}_e$ is mapped to the global element Ω_e via the geometry mapping $\boldsymbol{g}_1 \in (\mathcal{V}_{S,1})^3 = (\mathbb{Q}_{111})^3$.*
- *The space of approximation is $\mathcal{W}_{adapt,p}$.*
- *The finite element moments are the hybrid combination of the moments of $\mathcal{W}_{SI,p}$ and $\mathcal{W}_{T,p}$.*

6.5.1 Numerical Examples

The main goal of the present numerical examples is to compare the finite element solutions obtained for the following approximation spaces:

- Tensor product space $\mathcal{W}_{T,p}$ of order p,
- Mixed-order Nédélec space $\mathcal{W}_{NI,p}$ of order p,
- Mixed-order serendipity space $\mathcal{W}_{SI,p}$ of order p,
- Mixed-order mixed-space (adaptive) $\mathcal{W}_{\text{adapt},p}$ of order p.

Each space is represented by basis functions from the same set (Table 6.1, Table 6.4). Even though the calculations are only performed with the proposed basis functions, any other set would yield the same solution that spans the same space. The only differences concern the numerical properties, such as the condition number, orthogonality, or sparsity.

The following numerical examples concern the calculation of the electric field $\boldsymbol{E}$ in a rectangular waveguide with the dimensions of 1 m × 2 m × 4 m. The boundary value problem (2.35) leads to the finite element discretization of (2.51) in a lossless media. The numerical model consists of PEC boundaries on the sides of the waveguide and wave ports on the two ends. In all cases, the model is excited by the electric field of the dominant TE_{01} mode [Poz11, Chapter 3.3] at the frequency of 140MHz. The focus is on the asymptotic rate of convergence in terms of the mesh parameter h, using $\boldsymbol{L}^2$- and $\boldsymbol{H}(\mathrm{curl})$-norms. The desired ideal rates in all cases are considered to be

$$\left\|\boldsymbol{E} - \pi^w_{p,h}\boldsymbol{E}\right\|_{\boldsymbol{L}^2(\Omega)} \leq Ch^p, \tag{6.146a}$$

$$\left\|\boldsymbol{E} - \pi^w_{p,h}\boldsymbol{E}\right\|_{\boldsymbol{H}(\mathrm{curl},\Omega)} \leq Ch^p. \tag{6.146b}$$

In some cases, the $\boldsymbol{H}(\mathrm{curl})$-norm error might be dominated by the $\boldsymbol{L}^2$-norm in the early asymptotic stage of the convergence. Therefore, the error of the electric field $\boldsymbol{E}$ and $\nabla \times \boldsymbol{E}$ are separately investigated in the $\boldsymbol{L}^2$-norm, relative to the analytical solution. Hence, the equivalent rates of (6.146) revealing the correctness in a more clearer way are

$$\left\|\boldsymbol{E} - \pi^w_{p,h}\boldsymbol{E}\right\|_{\boldsymbol{L}^2(\Omega)} \leq Ch^p, \tag{6.147a}$$

$$\left\|\nabla \times \boldsymbol{E} - \nabla \times \pi^w_{p,h}\boldsymbol{E}\right\|_{\boldsymbol{L}^2(\Omega)} \leq Ch^p. \tag{6.147b}$$

In all the following examples, these errors are calculated and normalized with reference to the analytical solutions and the convergence rates are compared to the ideal rates. The main argument for the watertight nature of these tests involves the

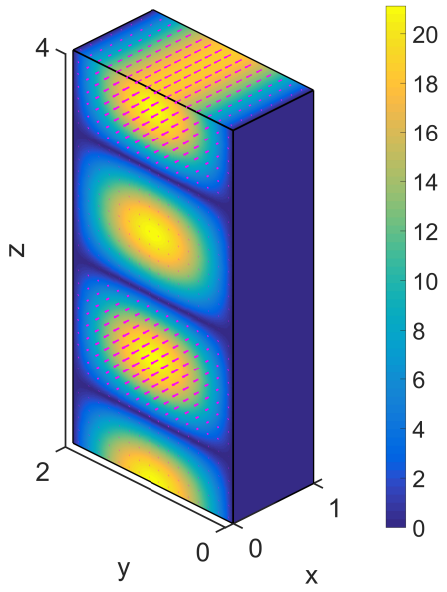


Figure 6.8: Magnitude of the calculated electric field strength (TE_{01} mode). All values are given in coherent SI units.

applications of several different meshes with highly distorted elements. Hence, even if the underlying analytical solution is the same, each element of the different meshes must approximate a different local finite element solution. Moreover, different ways of mesh refinement are applied to each initial mesh, which allows to demonstrate the effectiveness of the mixed-order mixed-space elements. The first method is the complete new remeshing of the domain with elements whose shape does not depends on the previous mesh. Indeed, meshing simple block-like domains with uniform cubic elements of affine geometry mappings is easy. In the general case, however, variously deformed hexahedrons are expected with non-affine geometry mappings. Thus, this remeshing process is equivalent to the case where the elements are obtained via non-affine refinements, allowing arbitrary node placement for the new elements. The second method is the affine refinement, which obtains any new element of the refined mesh through the affine reparametrization of the corresponding parent element. Details on the mesh refinement can be found in Section 3.1.2.

Uniform Mesh with Non-Affine Refinement

The first mesh consists of elements with the same shape, but with different translations, orientations, rotations, and scales. Each face of the elements is planar; however, the geometry mappings are non-affine. In each refinement, a completely new mesh is generated, consisting of the scaled-down versions of the initial element shapes. Hence, the smoothness of the geometry mappings are uniform on each ele-

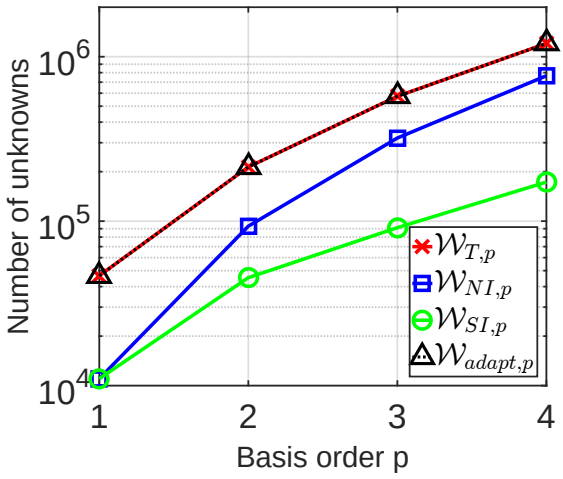


Figure 6.9: Number of unknowns for the finite element spaces $\mathcal{W}_{SI,p}$, $\mathcal{W}_{NI,p}$, $\mathcal{W}_{T,p}$, and $\mathcal{W}_{\text{adapt},p}$, with respect to the basis function order p. All values are obtained for the mesh with $h = 2^{-3}$.

ment and also on each refined mesh. Since the initial elements' geometry mapping are non-affine, such a sequence of meshes is equivalent to the non-affine mesh refinement of the initial mesh. The side view of the mesh sequence is shown in Figure 6.10. In Figure 6.11, the PEC boundary faces are colored red and the ports of the model are colored yellow.

The calculated electric field strength can be seen in Figure 6.8, which shows the reflection-free propagation of the dominant mode. Figure 6.12 shows the convergence of the relative errors in terms of the mesh parameter h, with respect to the analytical solution. For each convergence curve, the corresponding reference rate ($\mathcal{O}(h^p)$) is plotted with red color, starting backward from the finest calculation step, in order to neglect the preasymptotic region. It can be seen that the serendipity space $\mathcal{W}_{SI,p}$ does not meet the required rate in any case. The mixed-order Nédélec space $\mathcal{W}_{NI,p}$ performs better. The convergence rate of the electric field $\boldsymbol{E}$ meets the expectations in the higher-order case ($p > 1$). However, the convergence rate of $\nabla \times \boldsymbol{E}$ becomes one order less $\mathcal{O}(h^{p-1})$. Thus, in the first-order case ($p = 1$), neither $\nabla \times \boldsymbol{E}$ nor $\boldsymbol{E}$ shows any convergence for the space $\mathcal{W}_{NI,1}$. In order to achieve the same rate of convergence in both norms, either $\mathcal{W}_{T,p}$ or $\mathcal{W}_{\text{adapt},p}$ should be used, which happen to be the same for a uniform, non-affinely refined mesh. The number of unknowns in the finite element equation systems is shown in Figure 6.9 for the mesh of $h = 2^{-3}$, in terms of the basis order p. It can be seen that the serendipity space $\mathcal{W}_{SI,p}$ has the lowest number of unknowns; however, it produces low rates of convergence. The mixed-order Nédélec space $\mathcal{W}_{NI,p}$ produces a compromised solution, with acceptable (but not ideal) convergence and a higher number of unknowns. Finally, spaces $\mathcal{W}_{T,p}$ and $\mathcal{W}_{\text{adapt},p}$ produce the highest number of unknowns but achieve the expected ideal convergence rates.

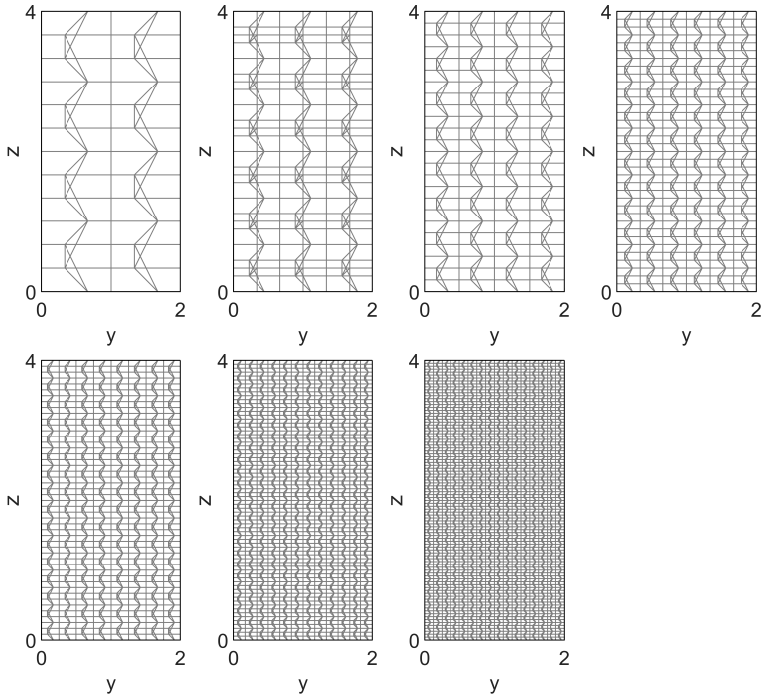


Figure 6.10: Side view of the uniform mesh sequence during non-affine refinement. The illustration includes the interior edges of the mesh.

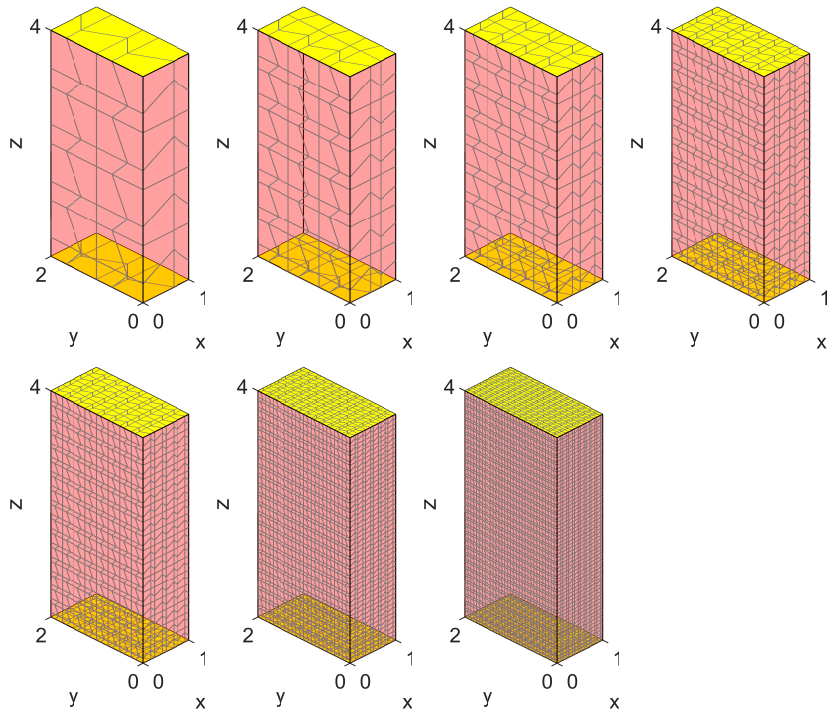


Figure 6.11: Boundaries of the uniform mesh sequence during non-affine refinement. The red faces are PEC and the yellow faces are wave ports.

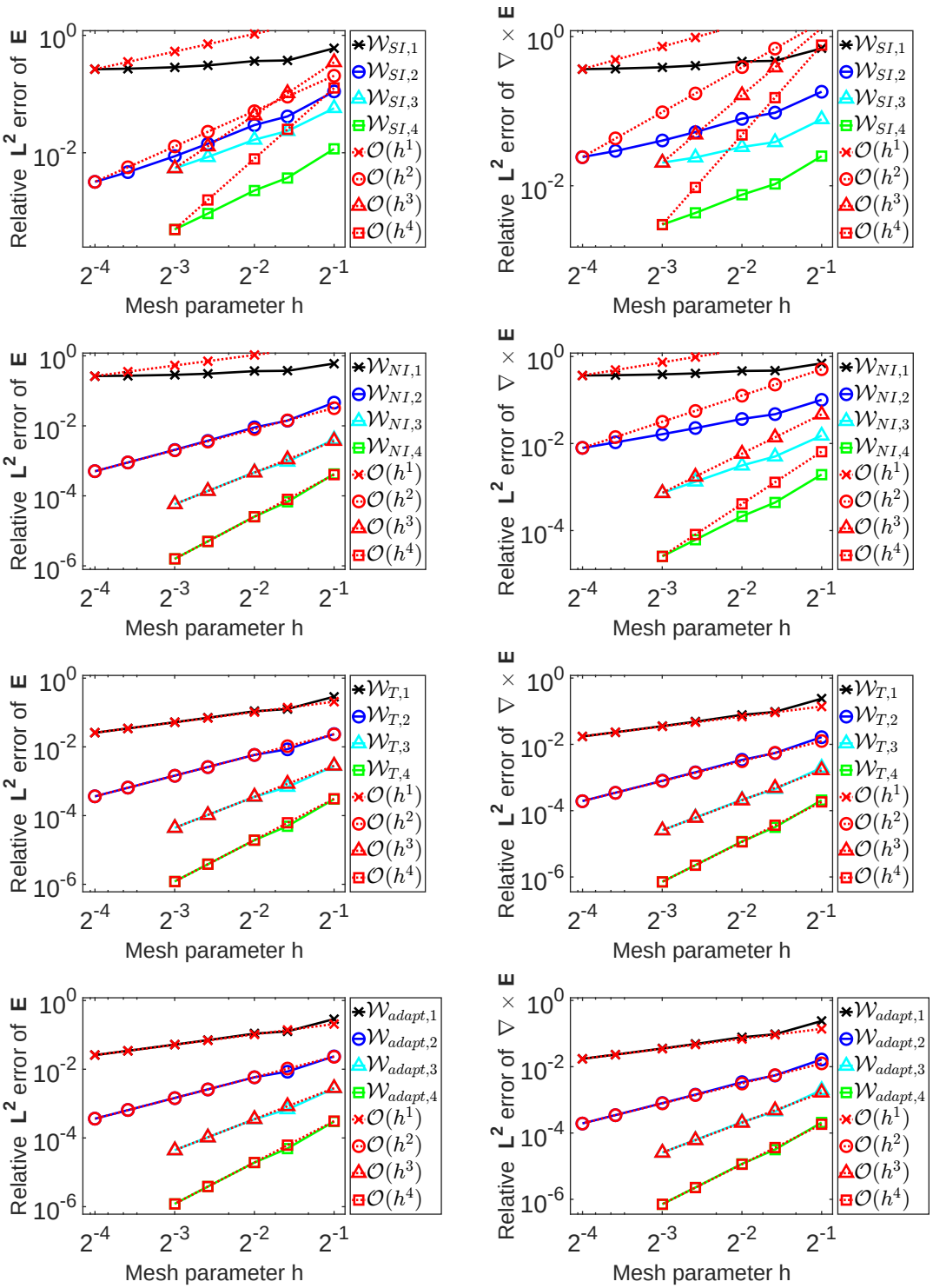


Figure 6.12: Relative error of the electric field $\boldsymbol{E}$ and its curl $\nabla \times \boldsymbol{E}$, in terms of the mesh parameter h. The calculations are obtained for the finite element spaces $\mathcal{W}_{SI,p}$, $\mathcal{W}_{NI,p}$, $\mathcal{W}_{T,p}$, and $\mathcal{W}_{\text{adapt},p}$, which are compared to the ideal rates $\mathcal{O}(h^p)$.

Half Affinely and Half Non-Affinely Refined Uniform Mesh

In practice, it is unusual for an entire mesh to consist of hexahedral elements with a non-regular shape and for all of them to be refined in a non-affine way. To get a compromised solution, a mesh whose upper half utilizes affine refinement is considered, whereas the bottom half uses non-affine refinement. The side view of the mesh sequence is shown in Figure 6.14. In Figure 6.15, the PEC boundary faces are colored red and the wave ports of the model are marked with yellow. The calculated convergence curves are shown in Figure 6.16, in terms of the mesh parameter h. Similarly to the previous example, the serendipity space $\mathcal{W}_{SI,p}$ shows poor performance, and the mixed-order Nédélec space $\mathcal{W}_{NI,p}$ provides $\mathcal{O}(h^{p-1})$ rate for $\nabla \times \boldsymbol{E}$. Moreover, both the tensor product space $\mathcal{W}_{T,p}$ and the mixed-order mixed-space $\mathcal{W}_{\text{adapt},p}$ finite elements show ideal convergence rates. The main difference from the previous example is that the serendipity space $\mathcal{W}_{SI,p}$ is sufficient for the $\mathcal{O}(h^p)$ convergence over the affinely refined (upper) half of the mesh. Therefore, the mixed-order mixed-space elements with $\mathcal{W}_{\text{adapt},p}$ require much fewer unknowns than the tensor product space $\mathcal{W}_{T,p}$ while providing the same convergence rates. These achieved numbers of unknowns, as shown in Figure 6.13, are comparatively the same for the Nédélec space $\mathcal{W}_{NI,4}$ and the mixed-order mixed-space $\mathcal{W}_{\text{adapt},4}$, while only the second one provides the correct rates.

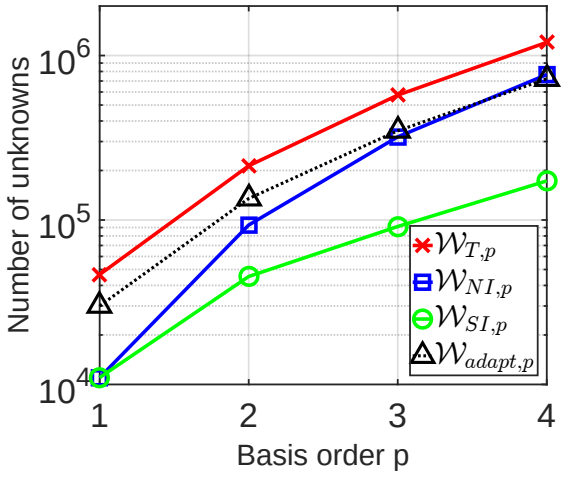


Figure 6.13: Number of unknowns for the finite element spaces $\mathcal{W}_{SI,p}$, $\mathcal{W}_{NI,p}$, $\mathcal{W}_{T,p}$, and $\mathcal{W}_{\text{adapt},p}$, with respect to the basis function order p. All values are obtained for the mesh with $h = 2^{-3}$.

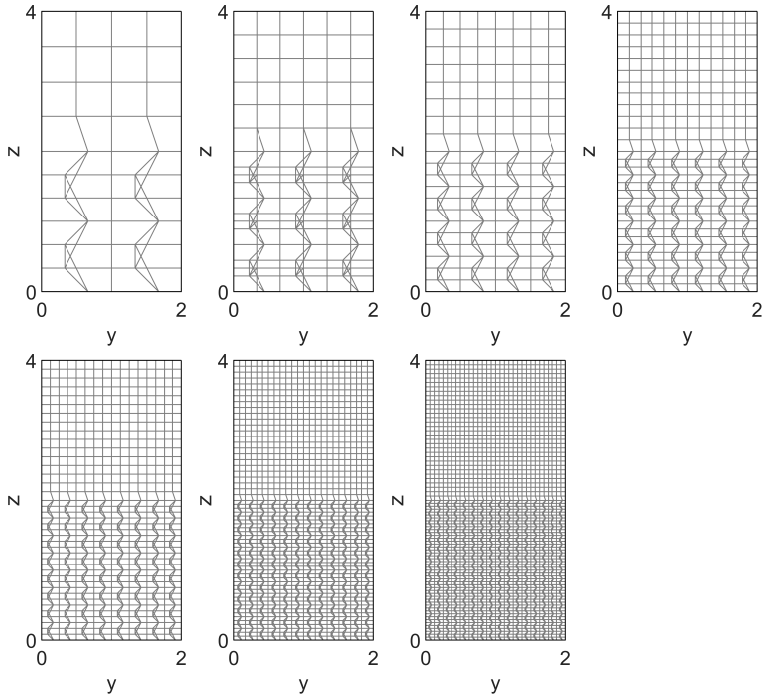


Figure 6.14: Side view of the half affinely and half non-affinely refined mesh sequence. The illustration includes the interior edges of the mesh.

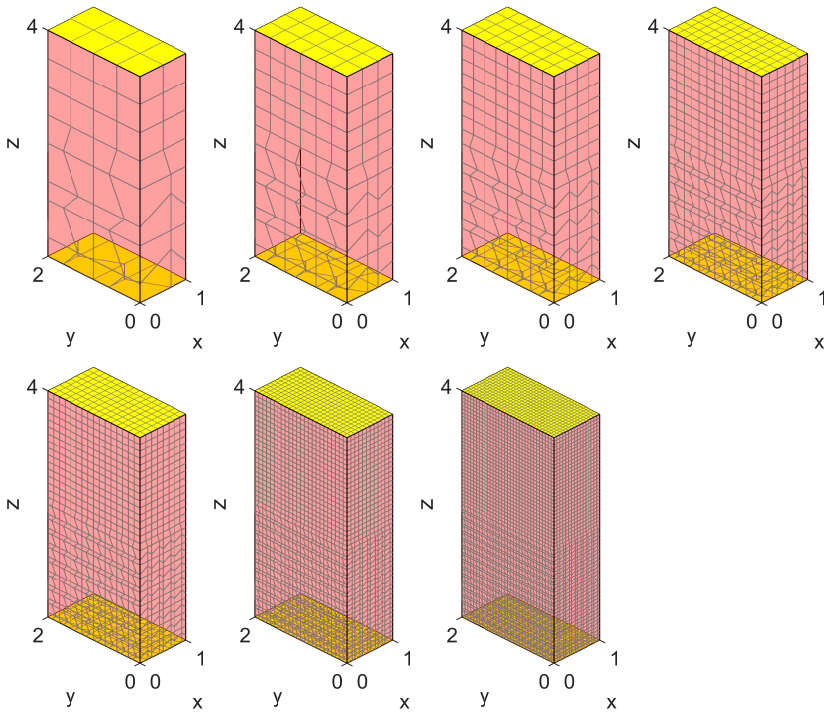


Figure 6.15: Boundaries of the half affinely and half non-affinely refined mesh sequence. The red faces are PEC and the yellow faces are wave ports.

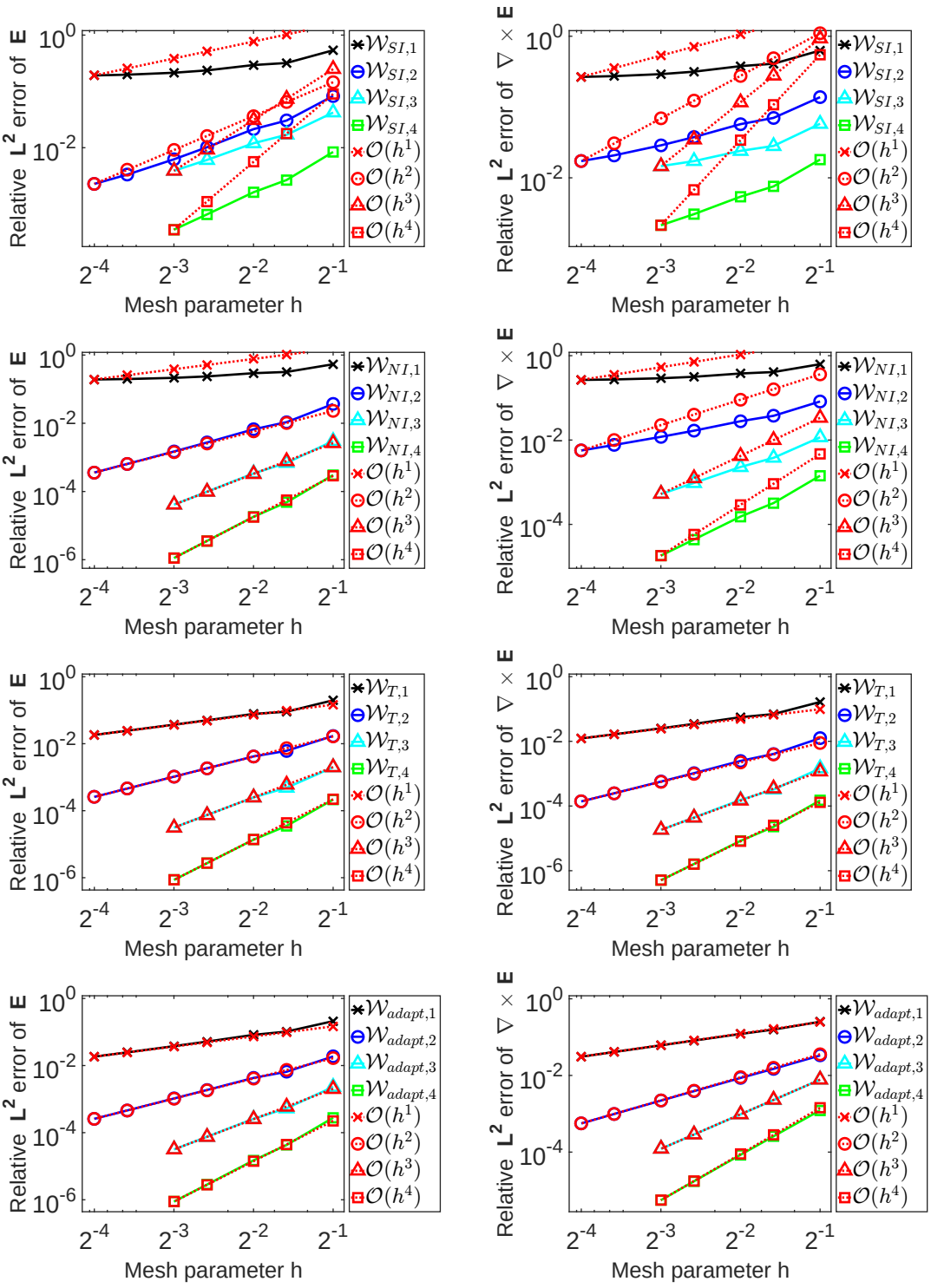


Figure 6.16: Relative error of the electric field $\boldsymbol{E}$ and its curl $\nabla \times \boldsymbol{E}$, in terms of the mesh parameter h. The calculations are obtained for the finite element spaces $\mathcal{W}_{SI,p}$, $\mathcal{W}_{NI,p}$, $\mathcal{W}_{T,p}$, and $\mathcal{W}_{\text{adapt},p}$, which are compared to the ideal rates $\mathcal{O}(h^p)$.

Uniform Mesh with a Non-Affine Element Layer

Taking a further step from the previous example, consider a mesh that mainly consists of regular cubical elements but has a non-affinely refined boundary layer of non-regular elements. This mesh sequence is shown in Figure 6.18 and Figure 6.19. Usually, the practical applications only require the elements to be non-regular and to be refined in a non-affine way if they are located next to a domain boundary, which is necessary for the correct representation of the geometry. However, it is allowed to set the dominant number of elements on the interior to be affine elements (parallelepipeds) and utilize affine refinement. In all those elements, the serendipity space $\mathcal{W}_{SI,p}$ is sufficient, which leads to a much lower number of unknowns for the mixed-order mixed-space $\mathcal{W}_{\text{adapt},p}$ finite element than for $\mathcal{W}_{NI,p}$ or $\mathcal{W}_{T,p}$. The comparison of the number of unknowns is shown in Figure 6.17 and the achieved convergences rates in Figure 6.20. It can be seen that the serendipity space $\mathcal{W}_{SI,p}$ performs better than in the previews examples, but still far from the ideal behavior. The Nédélec space $\mathcal{W}_{NI,p}$ shows almost perfect convergence; however, it cannot reach exactly the $\mathcal{O}(h^p)$ rate for $\nabla \times \boldsymbol{E}$. Furthermore, both the mixed-order mixed-space $\mathcal{W}_{\text{adapt},p}$ and the tensor product space $\mathcal{W}_{T,p}$ comply with the ideal convergence rates. The main advantage of the proposed method and space $\mathcal{W}_{\text{adapt},p}$ is that it provides the ideal rates while requiring the least number of unknowns.

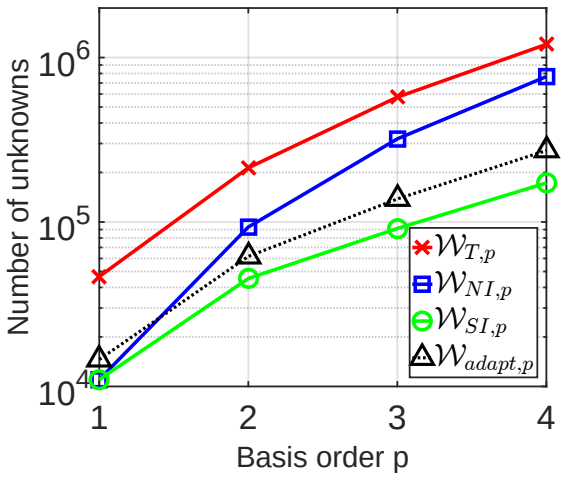


Figure 6.17: Number of unknowns for the finite element spaces $\mathcal{W}_{SI,p}$, $\mathcal{W}_{NI,p}$, $\mathcal{W}_{T,p}$, and $\mathcal{W}_{\text{adapt},p}$, with respect to the basis function order p. All values are obtained for the mesh with $h = 2^{-3}$.

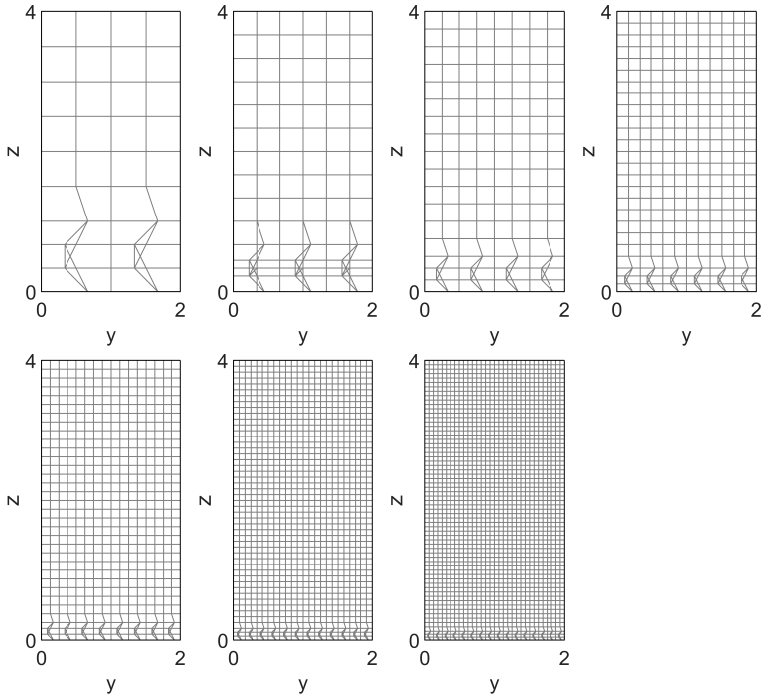


Figure 6.18: Side view of the dominantly cubic mesh sequence with a non-affinely refined mesh layer. The illustration includes the interior edges of the mesh.

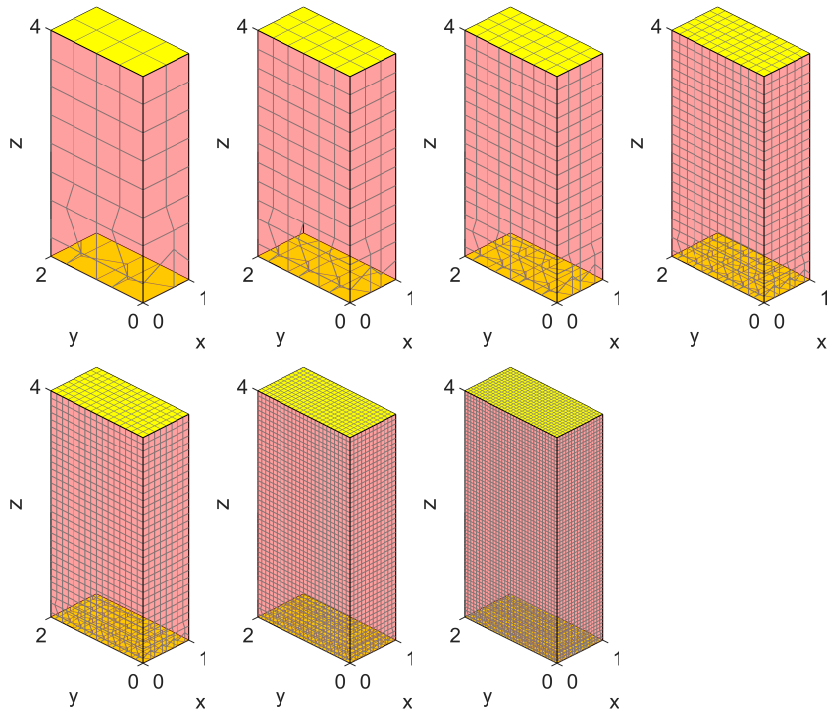


Figure 6.19: Boundaries of the dominantly cubic mesh sequence with a non-affinely refined mesh layer. The red faces are PEC and the yellow faces are wave ports.

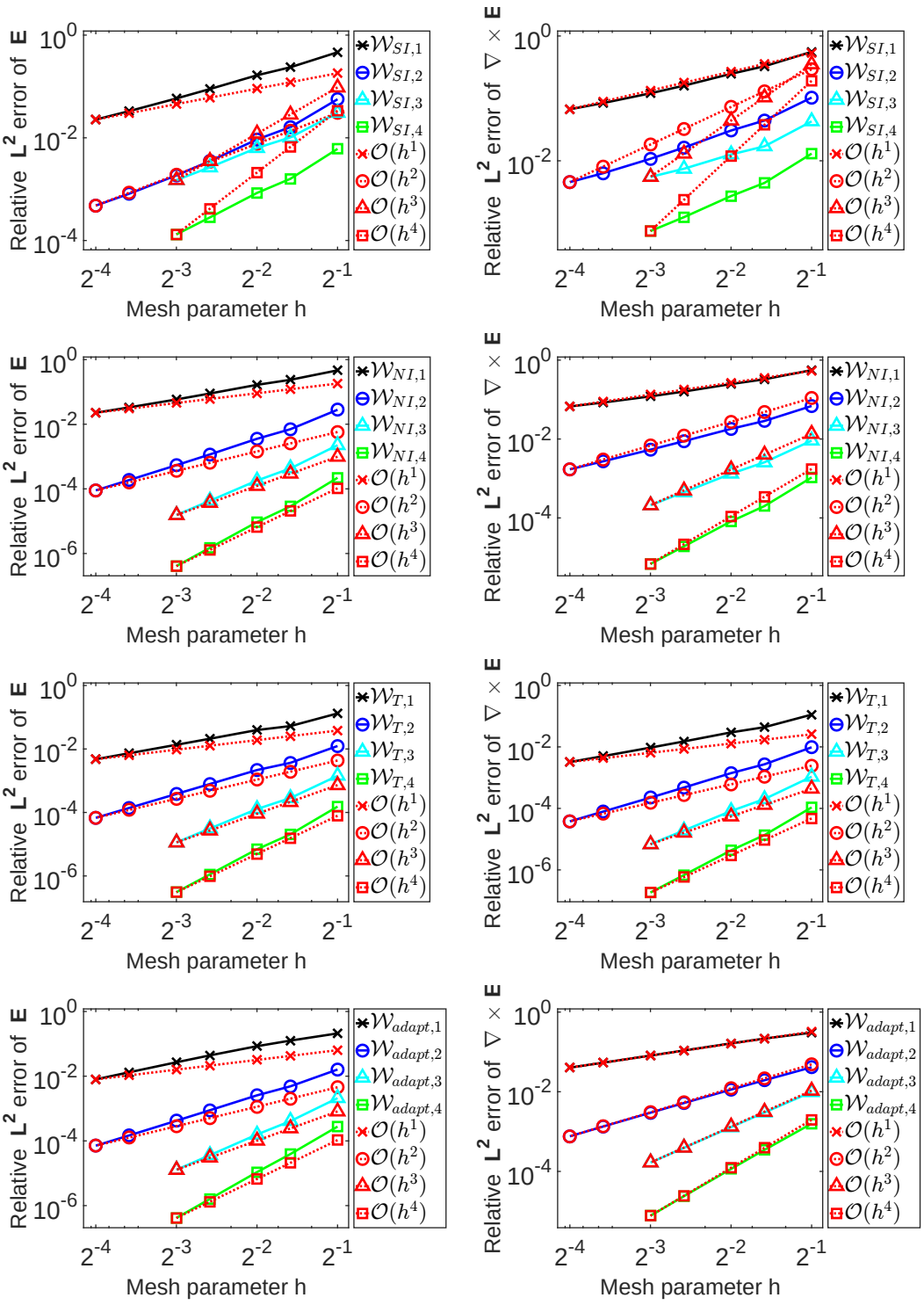


Figure 6.20: Relative error of the electric field $\boldsymbol{E}$ and its curl $\nabla \times \boldsymbol{E}$, in terms of the mesh parameter h. The calculations are obtained for the finite element spaces $\mathcal{W}_{SI,p}$, $\mathcal{W}_{NI,p}$, $\mathcal{W}_{T,p}$, and $\mathcal{W}_{\text{adapt},p}$, which are compared to the ideal rates $\mathcal{O}(h^p)$.

Non-Uniform Mesh with Affine Refinement

As the last example, consider an initial cubic mesh whose nodes are randomly displaced with the maximum deviation of 1/6 element side length in each direction. Therefore, the initial mesh consists of highly distorted elements with non-flat faces in the interior of the domain. However, each element is refined in an affine way by the affine reparametrization of the corresponding parent element. The applied mesh sequence is shown in Figure 6.22 and Figure 6.23. Due to the uniformly applied affine refinement, the mixed-order mixed-space $\mathcal{W}_{\text{adapt},p}$ becomes identical to the serendipity space $\mathcal{W}_{SI,p}$. Therefore, the necessary number of unknowns becomes minimal in this case (see Figure 6.21). The achieved convergence curves are shown in Figure 6.24. As expected, all finite element spaces produce the expected or even better convergence rates. The main difference between them is the higher accuracy on the given mesh, which comes at the price of a much higher number of unknowns. However, the space $\mathcal{W}_{\text{adapt},p}$ can afford a higher number of refinement than $\mathcal{W}_{T,p}$ or $\mathcal{W}_{NI,p}$ (if $p > 1$) for a given number of unknowns, which generally yields higher accuracy.

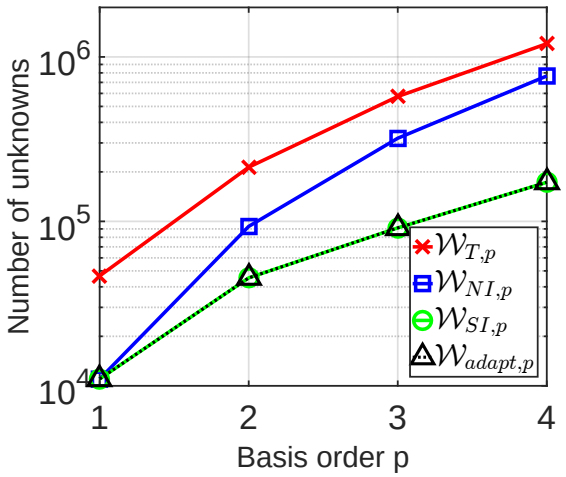


Figure 6.21: Number of unknowns for the finite element spaces $\mathcal{W}_{SI,p}$, $\mathcal{W}_{NI,p}$, $\mathcal{W}_{T,p}$, and $\mathcal{W}_{\text{adapt},p}$, with respect to the basis function order p. All values are obtained for the mesh with $h = 2^{-3}$.

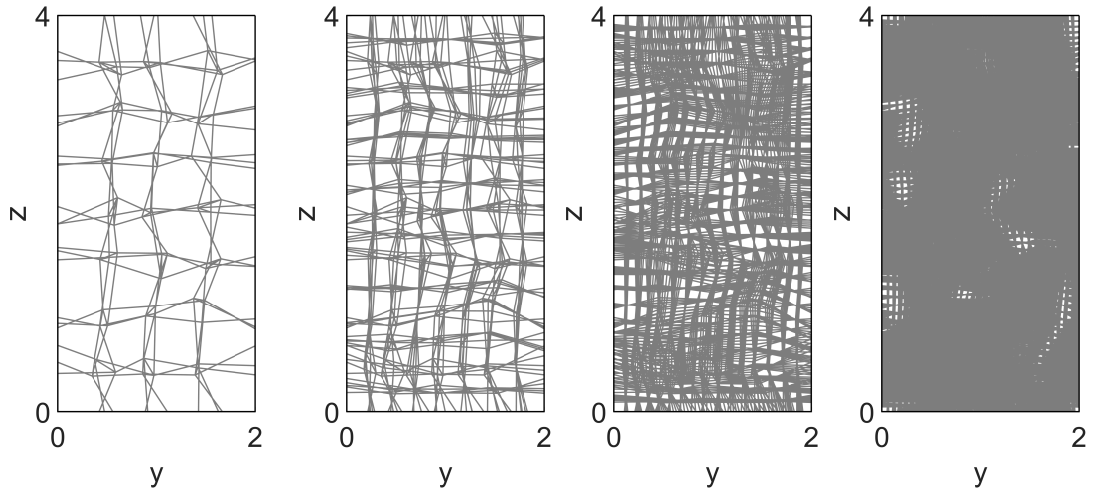


Figure 6.22: Side view of the non-uniform mesh sequence with affine refinement. The illustration includes the interior edges of the mesh.

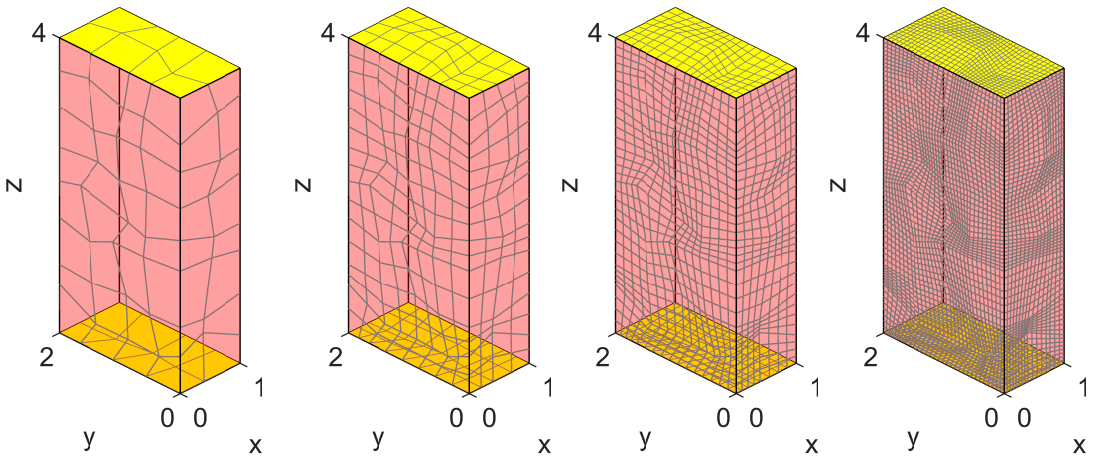


Figure 6.23: Boundaries of the non-uniform mesh sequence with affine refinement. The red faces are PEC and the yellow faces are wave ports.

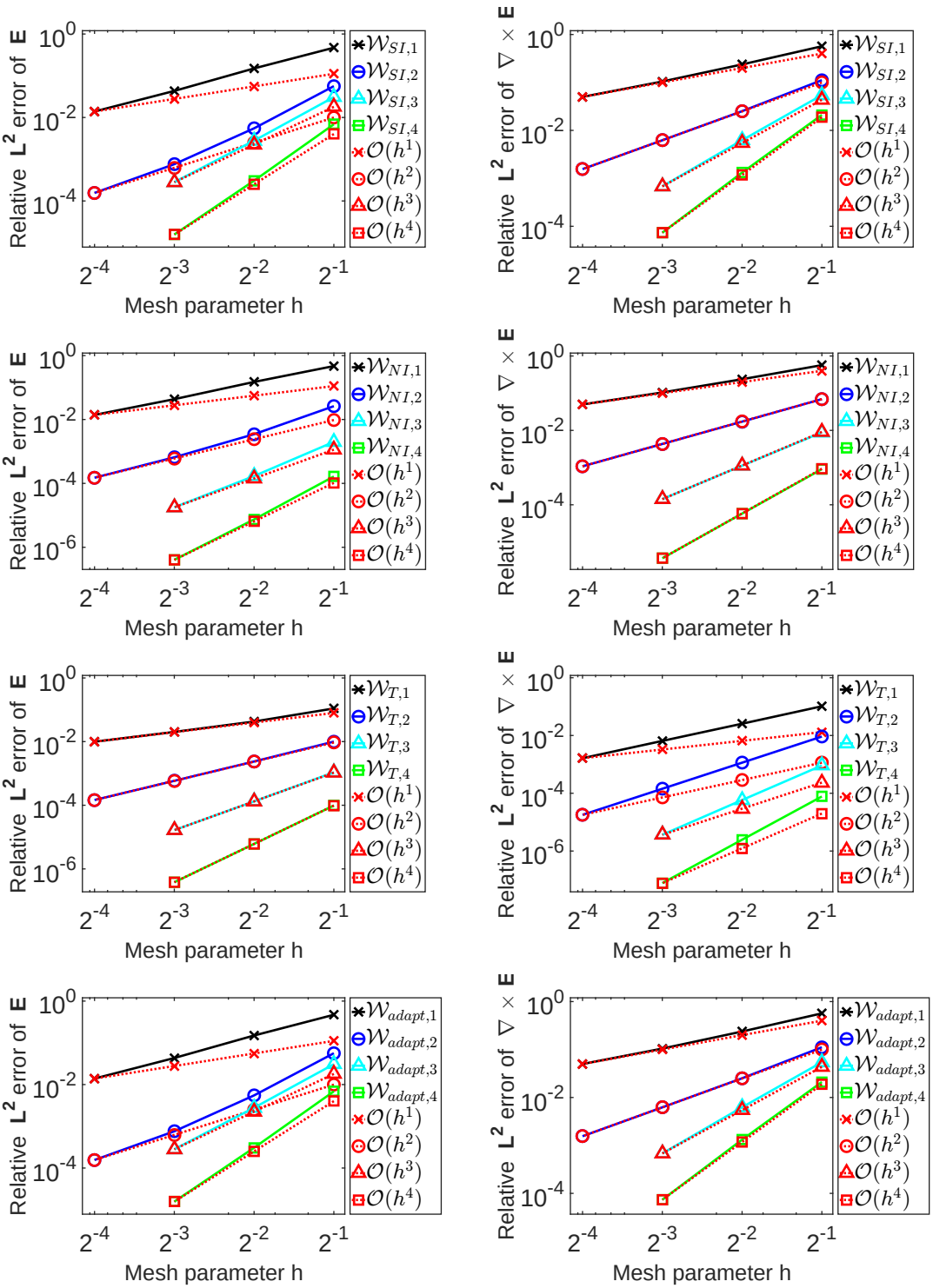


Figure 6.24: Relative error of the electric field $\boldsymbol{E}$ and its curl $\nabla \times \boldsymbol{E}$, in terms of the mesh parameter h. The calculations are obtained for the finite element spaces $\mathcal{W}_{SI,p}$, $\mathcal{W}_{NI,p}$, $\mathcal{W}_{T,p}$, and $\mathcal{W}_{\text{adapt},p}$, which are compared to the ideal rates of $\mathcal{O}(h^p)$.

6.6 Iso-Serendipity Finite Elements

The previous section only considered geometries that can be meshed by straight-sided elements without introducing any geometry error. General domains with curvilinear elements have not been discussed yet. The main problem with curvilinear geometries is that exact initial meshes are usually not available without geometry errors or are hard to construct. In most practical applications, only some interpolation locations are provided by the mesh generators or solid modellers and not the exact mappings of the initial elements. In other cases, exact mappings of the curvilinear boundaries might be available but not for the mapping of the volume. Therefore, the affine refinement of these unavailable mappings is almost impossible. The problem now is that the ideal convergence rate of the investigated finite elements is only guaranteed in the curvilinear case for the affine refinement of the exact meshes, where there are no geometry errors. To resolve this issue, a special way of refinement, namely the quasi-affine refinement, is proposed for meshes with interpolated geometry mappings. Quasi-affine refinement is based on the affine refinement of the interpolated geometry mappings and on a correction step for the boundary nodes. The main advantage is that interpolation nodes are always available, and the procedure is equivalent to the affine refinement of an explicitly unknown but exact initial mesh. There is no need for the exact geometry mappings since they are iteratively constructed in each refinement step.

Consider an element Ω_n with an exact geometry mapping, obtained as the n-times affine refinement of the initial element Ω_0. According to Section 3.1.2, the geometry mapping $\boldsymbol{g}(\widehat{\boldsymbol{r}}_n)$ of the element Ω_n is a function composition of the mapping of the initial element $\widetilde{\boldsymbol{g}}_0(\widehat{\boldsymbol{r}}_0)$ and some affine refinement function $\widehat{\boldsymbol{r}}_0 = \boldsymbol{f}_n(\widehat{\boldsymbol{r}}_n)$ with a unique inverse $\widehat{\boldsymbol{r}}_n = \boldsymbol{f}_n^{-1}(\widehat{\boldsymbol{r}}_0)$,

$$\boldsymbol{g}(\widehat{\boldsymbol{r}}_n) = \widetilde{\boldsymbol{g}}_0(\widehat{\boldsymbol{r}}_0) \circ \boldsymbol{f}_n(\widehat{\boldsymbol{r}}_n). \tag{6.148}$$

Then, the pth-order interpolant of this mapping is denoted by $\boldsymbol{g}_p(\widehat{\boldsymbol{r}}_n)$, which introduces a geometry error $\boldsymbol{R}_p(\widehat{\boldsymbol{r}}_n)$ for the refined element as

$$\boldsymbol{r} = \widetilde{\boldsymbol{g}}_0(\widehat{\boldsymbol{r}}_0) \circ \boldsymbol{f}_n(\widehat{\boldsymbol{r}}_n) = \boldsymbol{g}_p(\widehat{\boldsymbol{r}}_n) + \boldsymbol{R}_p(\widehat{\boldsymbol{r}}_n). \tag{6.149}$$

However, this error vanishes on the associated interpolation locations of the global $\boldsymbol{r}_{n,i}$ and reference element $\widehat{\boldsymbol{r}}_{n,i}$ by the definition of interpolation

$$\boldsymbol{r}_{n,i} = \widetilde{\boldsymbol{g}}_0(\widehat{\boldsymbol{r}}_0) \circ \boldsymbol{f}_n(\widehat{\boldsymbol{r}}_{n,i}) = \boldsymbol{g}_p(\widehat{\boldsymbol{r}}_{n,i}). \tag{6.150}$$

According to Section 3.1 and Section 3.3, if the interpolated geometry mapping is H^1-conforming and capable of interpolating polynomials of order p, then the finite element convergence rates are not affected. Thus, an appropriate choice of such a geometry interpolating basis is the H^1-conforming serendipity basis $v_i \in \mathcal{V}_{S,p}$. Therefore, the interpolated mapping can be written as

$$\boldsymbol{g}_p = \sum_i \boldsymbol{c}_i v_i, \qquad \text{for } v_i \in \mathcal{V}_{S,p}, \tag{6.151}$$

where the vector coefficients $\boldsymbol{c}_i$ are determined through interpolation. In the present setting, the higher-order locations are uniformly distributed over the edges, faces, and volumes of the global and reference elements. A different selection could provide a better solution, but without improving the asymptotic convergence rate.

So far, the initial mapping $\widetilde{\boldsymbol{g}}_0(\widehat{\boldsymbol{r}}_0)$ is undetermined, whereas the type of refinement and the interpolation locations are known. Thus, the only requirement for maintaining the affine refinement of $\widetilde{\boldsymbol{g}}_0(\widehat{\boldsymbol{r}}_0)$ is the backward compatibility with the previous interpolations. Since the interpolated geometry mapping is uniquely determined by the interpolation locations $\widehat{\boldsymbol{r}}_{m,i}$ for $m < n$, the compatibility of the actual and the parent element is guaranteed if the points of previous interpolation locations map to the same global point after each refinement:

$$\boldsymbol{r}_{m,i} = \left(\boldsymbol{g}_p\left(\widehat{\boldsymbol{r}}_n\right) \circ \boldsymbol{f}_n^{-1}\left(\widehat{\boldsymbol{r}}_0\right)\right) \circ \boldsymbol{f}_m\left(\widehat{\boldsymbol{r}}_{m,i}\right), \qquad \forall \boldsymbol{r}_{m,i} \in \Omega_n \text{ with } m < n. \tag{6.152}$$

A straightforward way to ensure this is to use all the old interpolation locations of the parent elements $\boldsymbol{r}_{m,i} \in \Omega_n$ with $m < n$ for the new interpolation locations on the refined element. The remaining new boundary interpolation locations must be placed at the exact curvilinear surface to eliminate the pointwise geometry error. The rest of the new interpolation locations without constraints are located in the interior of the domain and can be placed freely. Naturally, placing these free nodes to yield smoother geometry mappings is advantageous since the smoothness is a significant source of the finite element error (see Section 3.1). The previously discussed points result in the quasi-affine refinement.

Definition 6.2 (Quasi-affine refinement)**.**
Quasi-affine refinement *is a mesh refinement procedure that consists of three steps:*

1. *Affine refinement of the mesh where the elements are equipped with interpolated geometry mappings* $\boldsymbol{g}_p \in (\mathcal{V}_{S,p})^3$.
2. *Displacement of the new interpolation nodes on the curvilinear boundaries to the true (exact) curvilinear boundary.*
3. *All face- and volume-associated interpolation locations that do not coincide with previous interpolation locations can be displaced as long as they do not affect the geometry representation. These locations can be regularized such that the polynomial orders of the mappings are minimized.*

This way, the initial mappings free of the geometry error are iteratively constructed in each refinement step. In the case of $p = 2$ order elements, this process is illustrated in the bottom half of Figure 6.25, where the second-order continuity of the parent element is preserved. Moreover, the upper half of Figure 6.25 illustrates a counterexample, taking the affine refinement of the corresponding straight-sided element and then displacing the boundary points. Therefore, this is a non-affine refinement

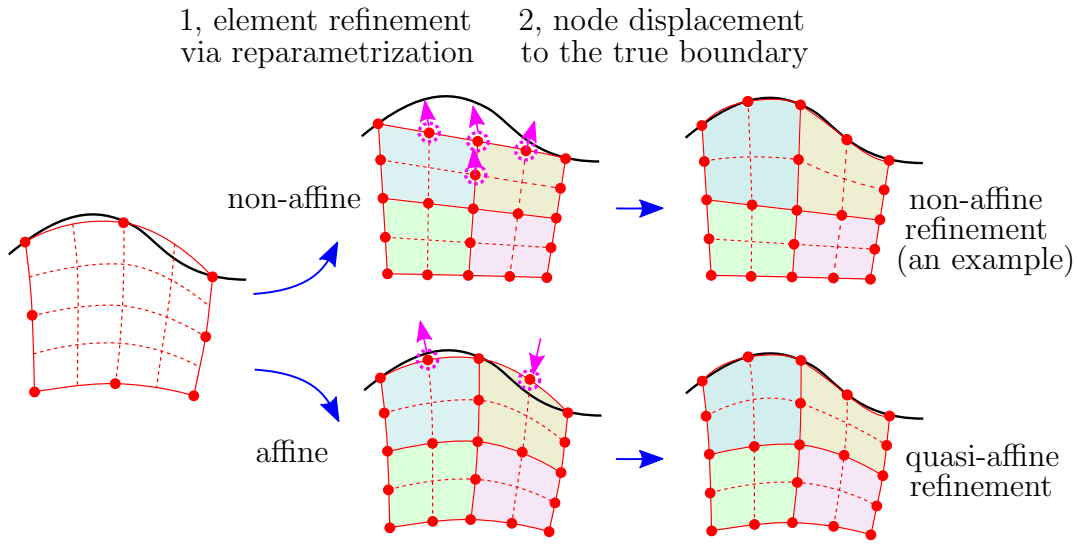


Figure 6.25: Illustration of the quasi-affine refinement and a counterexample for a second-order element.

that only maintains first-order continuity. A common approach for obtaining a finer mesh for curvilinear geometries is the generation of a completely new straight-sided mesh with half the size of the previous mesh. Then, higher-order nodes are inserted and displaced to the true curvilinear boundaries. The problem with this method is that each mesh introduces a different representation of the boundary, independent of the previous mesh. Hence, all constants in the convergence proofs might differ. Moreover, the convergence rate is affected by the differences in mesh continuity between consecutive meshes. An average mesh generator can provide such smooth meshes for simple structured domains that result in solutions close to the ideal theoretical limit. However, for complicated and unstructured domains, this is an unlikely outcome.

Some technical aspects of the quasi-affine refinement only arise in the higher-order case. For even numbers of geometry interpolation orders, such as $3, 5, ...$, the interpolation locations are uniformly distributed on the edges and do not contain the midpoint. Hence, compatibility with the previous interpolation locations and conformity of the mapping would yield nonuniform interpolation locations. Moreover, the third point of Definition 6.2 concerns the regularization of the face- and volume-associated interpolation locations, which do not affect the boundary of the element and the previous interpolation locations. In the $p > 3$ order case, some interpolation locations must be located on the face or within the element's interior volume. However, for all flat faces with curvilinear edges, the face-associated functions and interpolation locations do not contribute to the geometry representation. Similarly, none of the volume-associated interpolation functions and interpolation locations contributes to the boundary representation. Therefore, these interpolation locations can be displaced such that the polynomial order and the smoothness of

the mapping is decreased. The advantage of smoothness is the better approximation since the geometry mappings yield a smaller constant in (3.33). Moreover, the geometry error convergence and, consequently, the finite element error convergence reach their asymptotic region faster.

Overall, quasi-affine refinement with interpolated geometry mappings $\boldsymbol{g}_p \in (\mathcal{V}_{S,p})^3$ resolves the problems with affine refinement when exact geometry mappings are unavailable. Hence, any finite element is capable of achieving an $\mathcal{O}(h^p)$ convergence rate for curvilinear geometries if it is appropriate for parallelepipeds. Thus, an appropriate choice of finite element is the scalar serendipity space $\mathcal{V}_{S,p}$ for H^1-conforming functions and the mixed-order (or incomplete) serendipity space $\mathcal{W}_{SI,p}$ for $\boldsymbol{H}(\mathrm{curl})$-conforming functions, since they possess the least dimensions. Due to this feature, the method is called "iso-serendipity", since the same orders of serendipity spaces are used for both the geometry and field representations. The method can be viewed as the serendipity version of the isoparametric elements with some constraints on the mesh refinement procedure.

Definition 6.3 (Iso-serendipity H^1-conforming element)**.**
The iso-serendipity H^1-conforming element *is defined by the following four points:*

- *The reference element* $\widehat{\Omega}_e = [0,1]^3$ *is mapped to the global element by an interpolated geometry mapping* $\boldsymbol{g}_p \in (\mathcal{V}_{S,p})^3$*.*
- *The finite element moments are defined according to Section 6.2.1.*
- *The approximation space is the scalar* H^1*-conforming serendipity space* $\mathcal{V}_{S,p}$*.*
- *The only allowed way of refinement is* quasi-affine refinement*.*

Definition 6.4 (Iso-serendipity $\boldsymbol{H}(\mathrm{curl})$-conforming element)**.**
The iso-serendipity $\boldsymbol{H}(\mathrm{curl})$-conforming element *is defined by the following four points:*

- *The reference element* $\widehat{\Omega}_e = [0,1]^3$ *is mapped to the global element by an interpolated geometry mapping* $\boldsymbol{g}_p \in (\mathcal{V}_{S,p})^3$*.*
- *The finite element moments are defined according to Section 6.2.3.*
- *The approximation space is the* $\boldsymbol{H}(\mathrm{curl})$*-conforming mixed-order serendipity space* $\mathcal{W}_{SI,p}$*.*
- *The only allowed way of refinement is* quasi-affine refinement*.*

The main advantages of the iso-serendipity elements are the small number of unknowns and the resulting asymptotic convergence rates for smooth fields:

$$\left\| v - \pi^v_{p,h} v \right\|_{H^1(\Omega)} \leq C h^p, \tag{6.153a}$$

$$\left\|\boldsymbol{w} - \pi^{w}_{p,h}\boldsymbol{w}\right\|_{\boldsymbol{H}(\mathrm{curl},\Omega)} \leq Ch^{p}. \tag{6.153b}$$

At the same time, tensor product finite elements, such as in [Néd80, FGM11, BD13b], may be used instead, as they might provide better solutions. However, the convergence rates are expected to be the same, whereas the required number of unknowns is significantly higher for tensor product spaces.

6.6.1 Numerical Examples

In the following, the iso-serendipity finite element method is demonstrated with several numerical examples. In most cases, the model is chosen to be a coaxial waveguide whose electric field $\boldsymbol{E}$ is calculated. There are multiple reasons for this choice. First, the boundary of the domain is required to be curvilinear, which cannot be interpolated exactly by a finite number of polynomial basis functions; thus, a perfect geometry representation is not possible. Second, explicit analytical solutions are expected for both eigenvalue and excitation problems that cannot be represented exactly by a finite number of basis functions. This solution must be p times differentiable so that the desired convergence rates can be achieved:

$$\left\|\boldsymbol{E} - \pi^{w}_{p,h}\boldsymbol{E}\right\|_{\boldsymbol{L}^2(\Omega)} \leq Ch^{p}, \tag{6.154a}$$

$$\left\|\nabla \times \boldsymbol{E} - \nabla \times \pi^{w}_{p,h}\boldsymbol{E}\right\|_{\boldsymbol{L}^2(\Omega)} \leq Ch^{p}. \tag{6.154b}$$

In each calculation, these relative errors are obtained with respect to the exact analytical solution. The third reason is purely practical. If the hexahedral mesh is topologically not equivalent to a block structure with missing elements, it might be impossible to define only one reference element with a given orientation, which is only determined by the indices of the hexahedra nodes. In order to avoid the programming technicalities of the element orientations based on the mesh topology, my decision was to choose a structured mesh. However, note that the iso-serendipity method is not limited to structured meshes.

In each numerical example, the following three finite element spaces are compared using basis function from the same set (Table 6.4, Table 6.1):

- Mixed-order serendipity space $\mathcal{W}_{SI,p}$ (iso-serendipity method),
- Mixed-order Nédélec space $\mathcal{W}_{NI,p}$,
- Tensor product space $\mathcal{W}_{T,p}$.

The geometry mapping $\boldsymbol{g}_p$ of each element is expressed in terms of the serendipity basis functions $v_i \in \mathcal{V}_{S,p}$ as

$$\boldsymbol{g}_p = \sum_i \boldsymbol{c}_i v_i, \tag{6.155}$$

where the vector coefficients $\boldsymbol{c}_i$ are determined through interpolation.

As a remark, no separate numerical example is given for the demonstration of the H^1-conforming iso-serendipity elements and the corresponding serendipity space $\mathcal{V}_{S,p}$. However, the approximation space $\mathcal{W}_{SI,p}$ of the $\boldsymbol{H}(\mathrm{curl})$-conforming iso-serendipity element is such that it includes the gradient serendipity space $\nabla\mathcal{V}_{S,p} \subset \mathcal{W}_{SI,p}$. Furthermore, the very same basis functions are used for $v \in \mathcal{V}_{S,p}$ as for the gradient subspace $\boldsymbol{w} = \nabla v \in \nabla\mathcal{V}_{S,p} \subset \mathcal{W}_{SI,p}$ in the higher-order ($p > 1$) case. Therefore, the commutation of the interpolation operators yields

$$\left\|\boldsymbol{w} - \pi^w_{p,h}\boldsymbol{w}\right\|_{\boldsymbol{L}^2(\Omega)} \overset{\boldsymbol{w}=\nabla v}{=} \left\|\nabla v - \pi^w_{p,h}\nabla v\right\|_{\boldsymbol{L}^2(\Omega)} = \left\|\nabla(v - \pi^v_{p,h}v)\right\|_{\boldsymbol{L}^2(\Omega)}, \tag{6.156}$$

which guarantees the H^1-conforming convergence rate

$$\left\|v - \pi^v_{p,h}v\right\|_{H^1(\Omega)} \leq Ch^p, \qquad \text{for } \pi^v_p v \in \mathcal{V}_{S,p}, \tag{6.157}$$

due to Table 3.1 if the convergence rate of

$$\left\|\boldsymbol{w} - \pi^w_{p,h}\boldsymbol{w}\right\|_{\boldsymbol{L}^2(\Omega)} \leq Ch^p, \qquad \text{for } \pi^w_p \boldsymbol{w} \in \mathcal{W}_{SI,p}, \tag{6.158}$$

is satisfied for $\boldsymbol{H}(\mathrm{curl})$-conforming functions.

Coaxial Cavity with a Quasi-Affinely Refined Mesh

The purpose of the first numerical example is to demonstrate quasi-affine refinement on a simple finite element model using different orders of interpolated geometry mappings. The geometry consists of a single coaxial domain with the material properties of the vacuum. The considered boundary value problem (2.35) of the electric field leads to the algebraic eigenvalue problem (2.52) for lossless domains. The dominant mode of the coaxial line is the TEM mode, which is comparatively simple to approximate due to the angular independence of the electric field [Poz11]. It is easy to provide a curvilinear mesh where the direction of the electric field is perfectly aligned with the mesh edges, and the approximation of the solution is not challenging. In order to obtain a higher-order and less trivial physical mode as the dominant mode of the numerical model, only half of the coaxial cavity is considered, and the cross-section plane is replaced by a PEC. Thus, each boundary became a PEC, allowing us to calculate only the antisymmetric resonant field patterns of the entire coaxial cavity. The second advantage of this configuration is the single multiplicity of the smallest nonzero eigenvalue and the uniqueness of the dominant mode. The investigation of higher-order eigenvalues is not favorable since they are less smooth and the preasymptotic region of the convergence is significantly larger. Moreover, eigenvalues with higher multiplicity are also not preferred since they might result in reduced convergence rates.

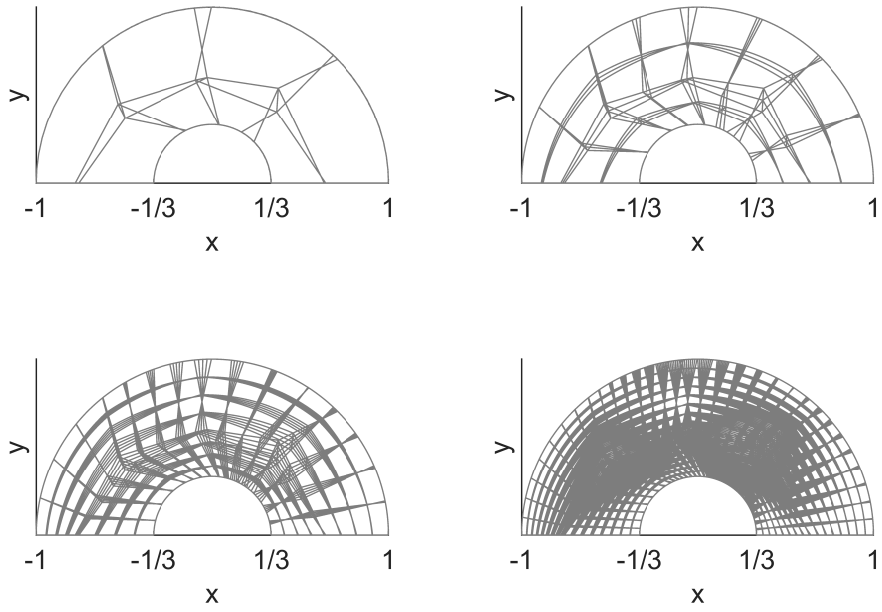


Figure 6.26: Side view of the mesh for the coaxial cavity during quasi-affine refinement. The illustration includes the interior edges of the mesh.

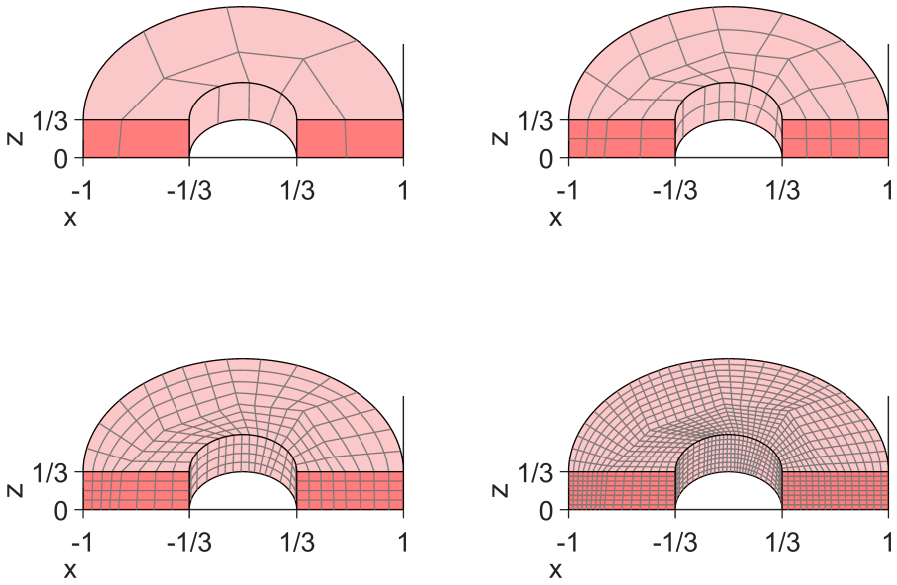


Figure 6.27: Boundaries of the mesh for the coaxial cavity during quasi-affine refinement.

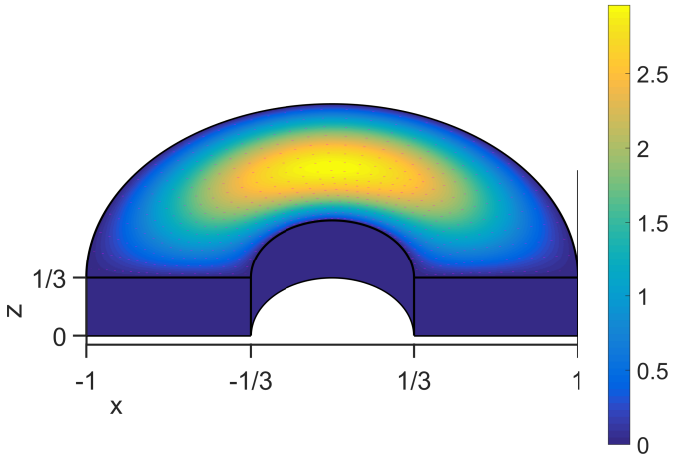


Figure 6.28: Magnitude of the calculated electric field strength for the dominant mode. All values are given in coherent SI units.

Two different mesh sequences are used in the calculations. In both cases, the vertices of the elements are the same; however, the interpolated geometry order is different. In the first case, the geometry mappings are chosen to be second-order $\boldsymbol{g}_2 \in (\mathcal{V}_{S,2})^3$. In the second case, the geometry mappings are chosen to be fourth-order $\boldsymbol{g}_4 \in (\mathcal{V}_{S,4})^3$. Since the two cases cannot be distinguished visually, only the meshes of the $\boldsymbol{g}_4$ mappings are shown in Figure 6.26 and Figure 6.27. In both cases, the quasi-affine refinement is applied and the smallest non-zero eigenvalue is calculated. The corresponding dominant mode is shown in Figure 6.28.

In the first case, the relative error of the calculated eigenvalue is shown in Figure 6.29 with reference to the analytical solution, in terms of the mesh parameter h. For each finite element space, the ideal rate of convergence $\mathcal{O}(h^{2p})$ is plotted backward from the finest step, in order to neglect the preasymptotic region. It can be seen that the serendipity space $\mathcal{W}_{SI,p}$ and the Nédélec space $\mathcal{W}_{NI,p}$ show the desired rates for orders $p = 1$ and $p = 2$. However, in the higher order ($p > 2$) case, the rate of convergence is unchanged. This is the expected behavior for using only second-order geometry mappings $\boldsymbol{g}_2$. The result of the tensor product space $\mathcal{W}_{T,p}$ is slightly inconsistent. This can be explained by the fact that $\mathcal{W}_{T,p}$ contains much higher-order monomials than p, whereas the geometry mapping only belongs to the serendipity space $\boldsymbol{g}_2 \in (\mathcal{V}_{S,2})^3$. This strong difference between the field and geometry representations distorts the results.

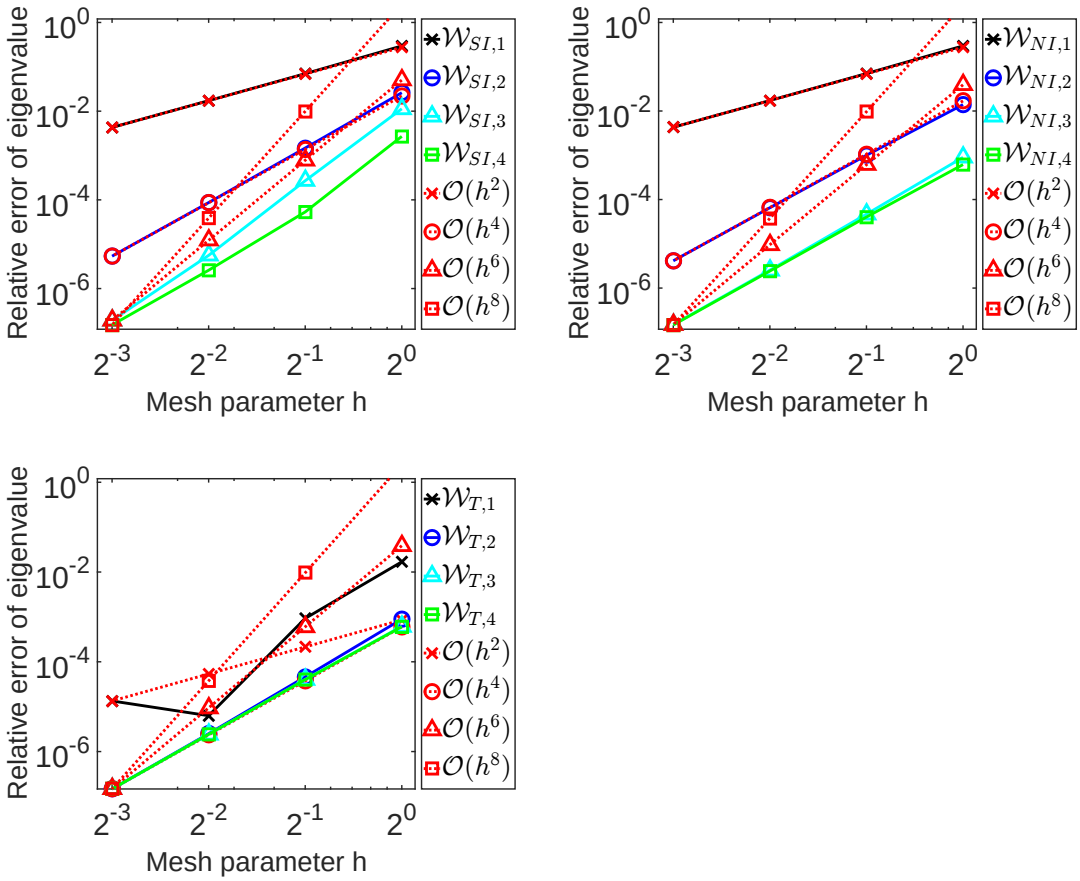


Figure 6.29: Relative error of the smallest nonzero eigenvalue in terms of the mesh parameter h. The calculations are obtained for the second-order mesh with spaces $\mathcal{W}_{SI,p}$, $\mathcal{W}_{NI,p}$, and $\mathcal{W}_{T,p}$, which are compared to the ideal rates of $\mathcal{O}(h^{2p})$.

In the second case, fourth-order geometry mappings $\boldsymbol{g}_4 \in (\mathcal{V}_{S,4})^3$ are applied. The relative error of the calculated (smallest non-zero) eigenvalue is shown in Figure 6.30 in terms of the mesh parameter h. The serendipity space $\mathcal{W}_{SI,p}$ clearly provides the expected convergence rates for the orders $p = 1, 2, 3, 4$. However, both $\mathcal{W}_{NI,p}$ and $\mathcal{W}_{T,p}$ show inconsistent results. Again, this can be explained by the enlarged range of the preasymptotic region due to the large difference between the geometry and field representations. The numbers of unknowns in the finite element equation systems are shown in Figure 6.31 for different p-orders of basis functions. According to the results, the serendipity space requires the lowest number of unknowns and the lowest order in the geometry representation while providing the ideal convergence rates.

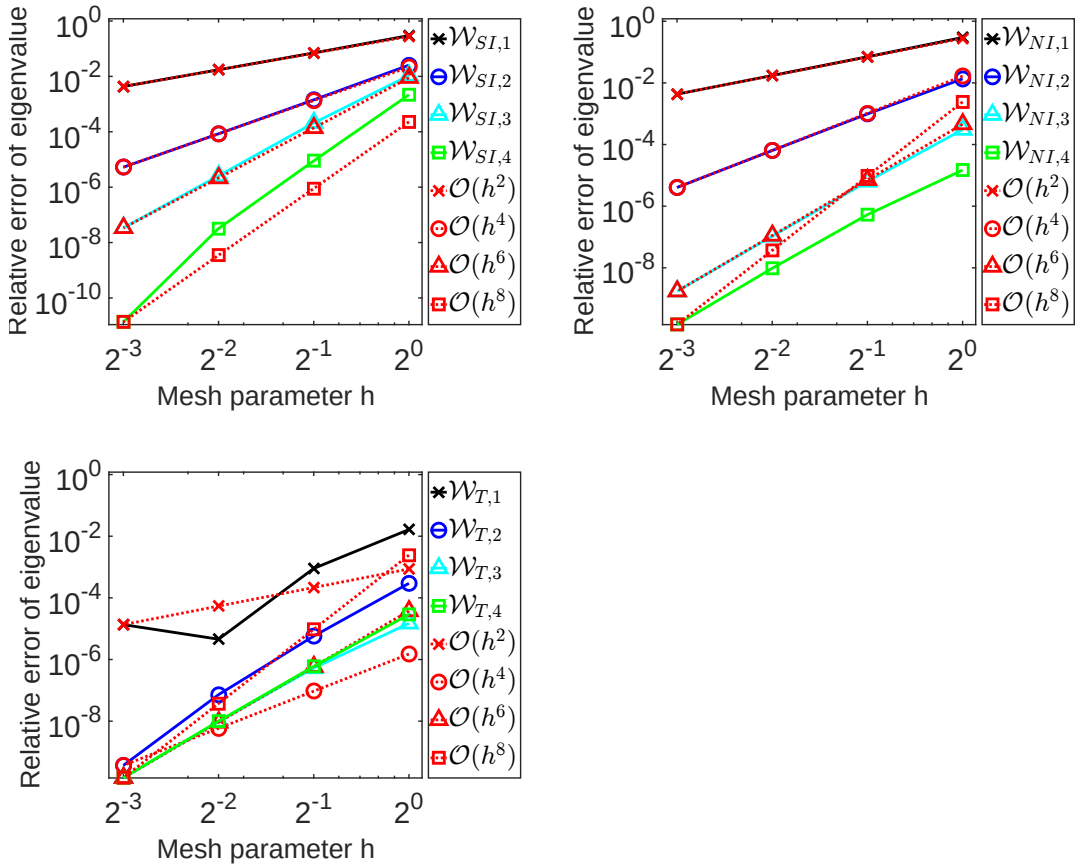


Figure 6.30: Relative error of the smallest nonzero eigenvalue in terms of the mesh parameter h. The calculations are obtained for the fourth-order mesh with spaces $\mathcal{W}_{SI,p}$, $\mathcal{W}_{NI,p}$, and $\mathcal{W}_{T,p}$, which are compared to the ideal rates of $\mathcal{O}(h^{2p})$.

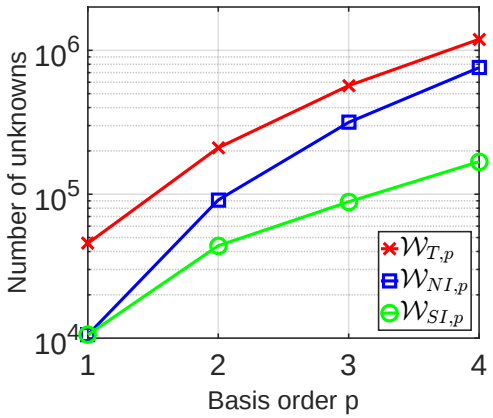


Figure 6.31: Number of unknowns for the finite element spaces $\mathcal{W}_{SI,p}$, $\mathcal{W}_{NI,p}$, and $\mathcal{W}_{T,p}$, with respect to the basis function order p. All values are obtained for the mesh with $h = 2^{-3}$.

Coaxial Cavity with a Non-Affinely Refined Mesh

The purpose of the second example is to demonstrate the effect of the non-affine refinement compared with the quasi-affine refinement, as shown in Figure 6.25. In order to make a fair comparison, the same calculation is performed on the same fourth-order initial mesh as in the previous example. The only difference is that the applied non-affine refinement (Figure 6.25) only preserves the first-order continuity of the elements instead of the required fourth-order continuity. This non-affinely refined mesh sequence is shown in Figure 6.33 and Figure 6.34. The relative error of the first nonzero eigenvalue is shown in Figure 6.32 in terms of the mesh parameter h. As expected, only the first-order $\mathcal{W}_{SI,1}$, $\mathcal{W}_{NI,1}$, and $\mathcal{W}_{T,1}$ result in the correct rate of convergence $\mathcal{O}(h^2)$. For all the higher-order cases, the convergence rates are not sufficient and are inconsistent.

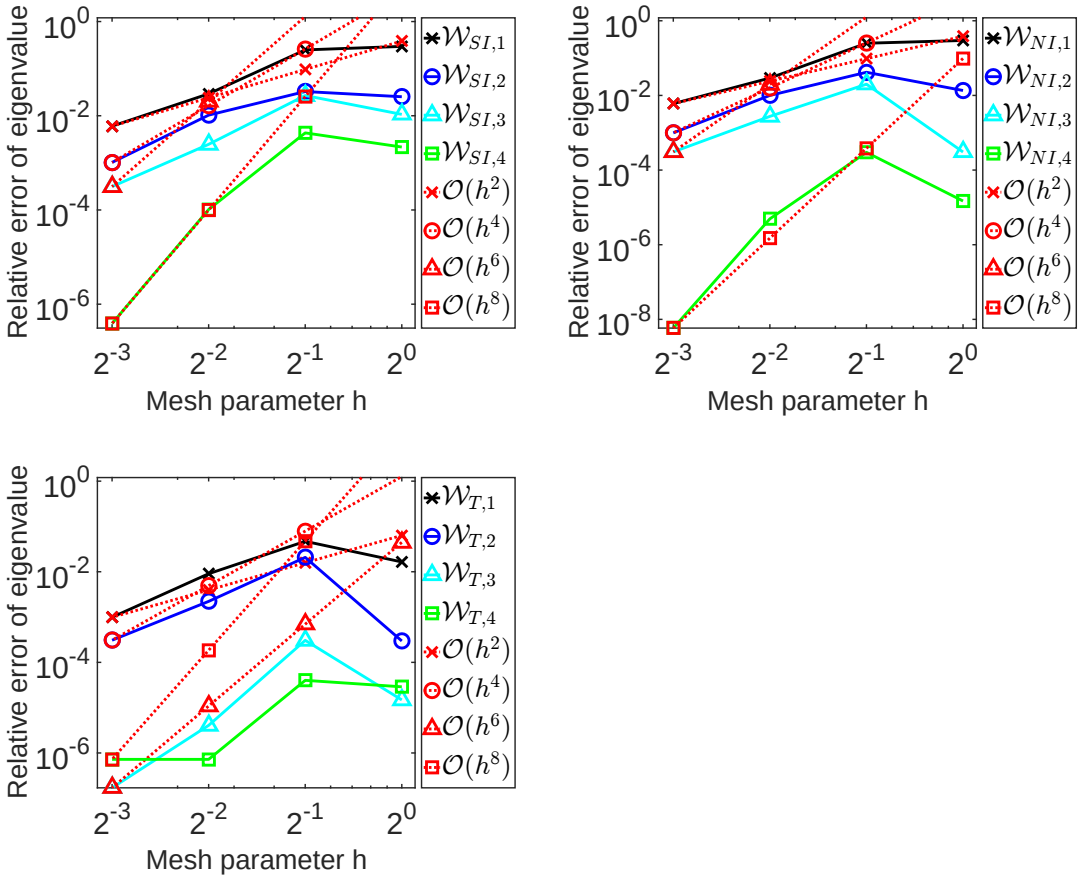


Figure 6.32: Relative error of the smallest nonzero eigenvalue in terms of the mesh parameter h. The calculations are obtained for the fourth-order mesh with spaces $\mathcal{W}_{SI,p}$, $\mathcal{W}_{NI,p}$, and $\mathcal{W}_{T,p}$, which are compared to the ideal rates of $\mathcal{O}(h^{2p})$.

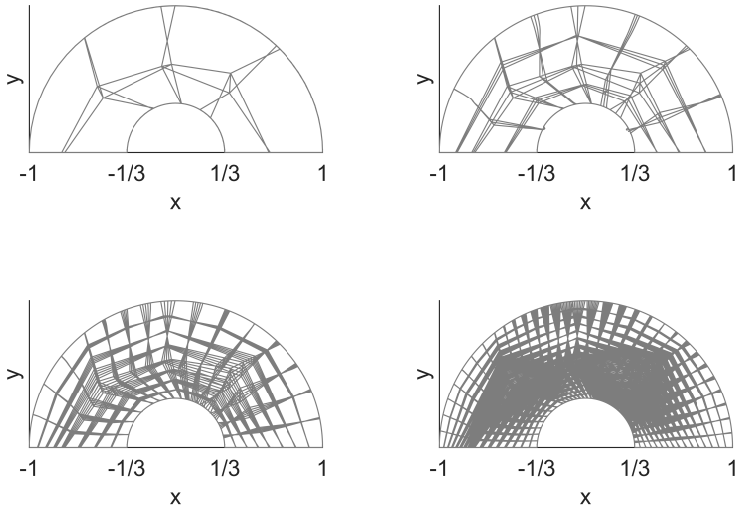


Figure 6.33: Side view of the mesh for the coaxial cavity during non-affine refinement. The illustration includes the interior edges of the mesh.

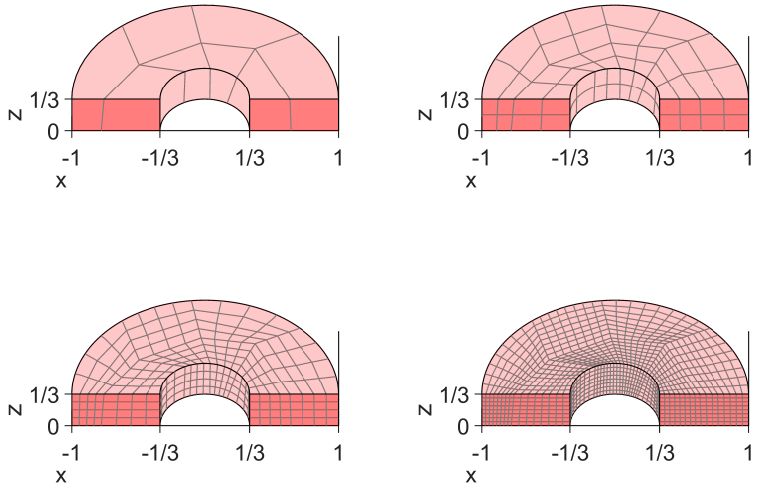


Figure 6.34: Boundaries of the mesh for the coaxial cavity during non-affine refinement.

Coaxial Waveguide with a Quasi-Affinely Refined Regular Mesh

The following numerical example is an excitation problem with the finite element formulation (2.51). The inner and outer conductors of the coaxial waveguide are PEC, whereas the two ends are wave ports. The dominant mode of the coaxial line is the TEM mode, which has a comparatively simple field pattern [Poz11]. It is possible

to find a structured mesh with edges in the direction of the electric field. Therefore, the next, higher-order TE_{11} mode is considered. With this excitation, the electric field changes its direction smoothly from radial to tangential, and the hexahedral mesh cannot be aligned everywhere. The symmetry of the TE_{11} mode allows for the simulation of only a quarter of the coaxial waveguide, placing appropriate PEC and PMC boundary conditions. Hence, the model is excited by the electric field of the TE_{11} mode [Mar51, Section 2-4.] on the quarter of the coaxial cross-section. The sequence of the fourth-order quasi-affinely refined mesh is shown in Figure 6.36 and Figure 6.37.

Since the space $\mathcal{W}_{T,4}$ is the largest in dimension, it is expected to provide the best solution. The corresponding electric field and its error are shown in Figure 6.38 for the calculation on the finest mesh. The error pattern makes clear that the dominant error occurs in the regions close to the curvilinear boundary. However, the error pattern is smooth and similar to the solution. The calculated convergence curves for the electric field $\boldsymbol{E}$ and $\nabla \times \boldsymbol{E}$ are shown in Figure 6.39. In accordance with the theoretical predictions, the required order of convergence $\mathcal{O}(h^p)$ is reached in the $\boldsymbol{H}(\mathrm{curl})$- and $\boldsymbol{L}^2$-norm for all elements. The only notable difference between the investigated elements is the higher convergence rate of $\nabla \times \boldsymbol{E}$ for the tensor product space $\mathcal{W}_{T,p}$ than the stated ideal rate (6.154b). This is the expected behavior since $\mathcal{W}_{T,p}$ is a larger space tailored for general straight-sided elements and non-affine refinements. In the case of using affine or quasi-affine refinements, $\mathcal{W}_{T,p}$ satisfies the higher-order convergence requirement for $\nabla \times \boldsymbol{E}$ but not for $\boldsymbol{E}$ (see Table 3.2). In terms of the $\boldsymbol{H}(\mathrm{curl})$-norm, the presented elements only differ in the accuracy of the solutions on the initial mesh. Naturally, the higher accuracy comes at the price of a much higher number of unknowns; see Figure 6.35. However, space $\mathcal{W}_{SI,p}$ can afford a higher number of refinements that $\mathcal{W}_{T,p}$ or $\mathcal{W}_{NI,p}$ (if $p > 1$) for a given number of unknowns, which generally yields higher accuracy.

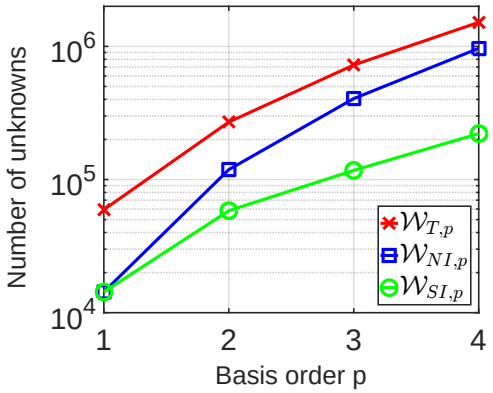


Figure 6.35: Number of unknowns for the finite element spaces $\mathcal{W}_{SI,p}$, $\mathcal{W}_{NI,p}$, and $\mathcal{W}_{T,p}$, with respect to the basis function order p. All values are obtained for the mesh with $h = 2^{-3}$.

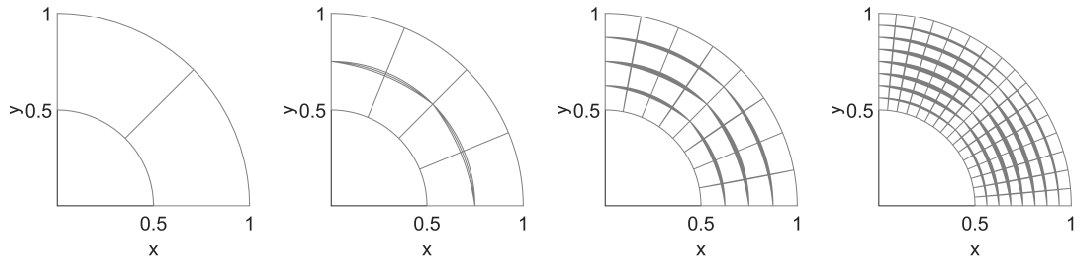


Figure 6.36: Side view of the mesh for the coaxial waveguide during quasi-affine refinement. The illustration includes the interior edges of the mesh.

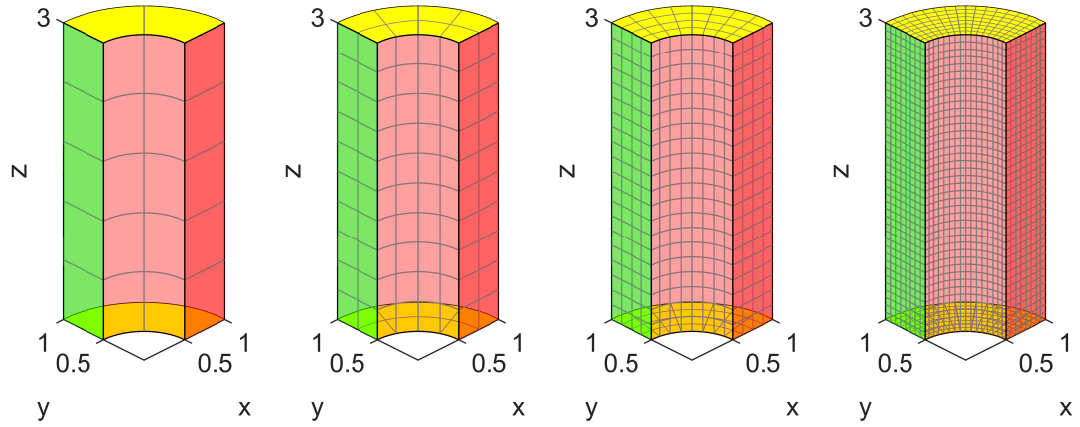


Figure 6.37: Boundaries of the coaxial mesh during quasi-affine refinement. The red faces are PEC, the green faces are PMC, and the yellow faces are wave ports.

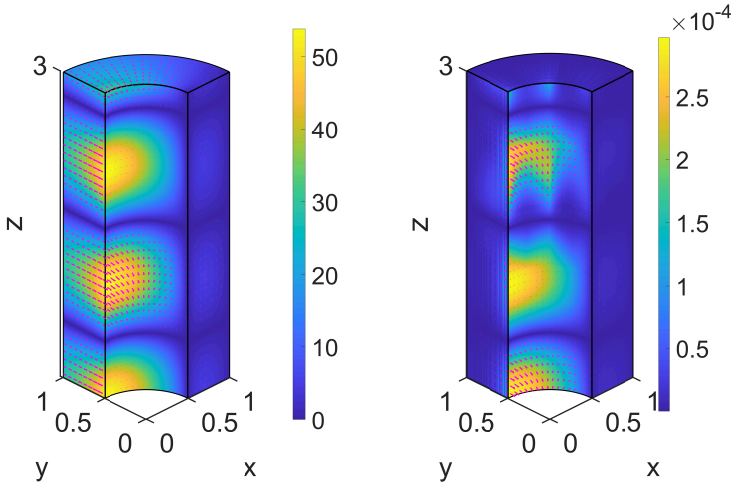


Figure 6.38: Magnitude of the calculated electric field strength (left) and its error with respect to the analytical solution (right). The results are obtained for the space $\mathcal{W}_{T,4}$ on the mesh with $h = 2^{-3}$. All values are given in coherent SI units.

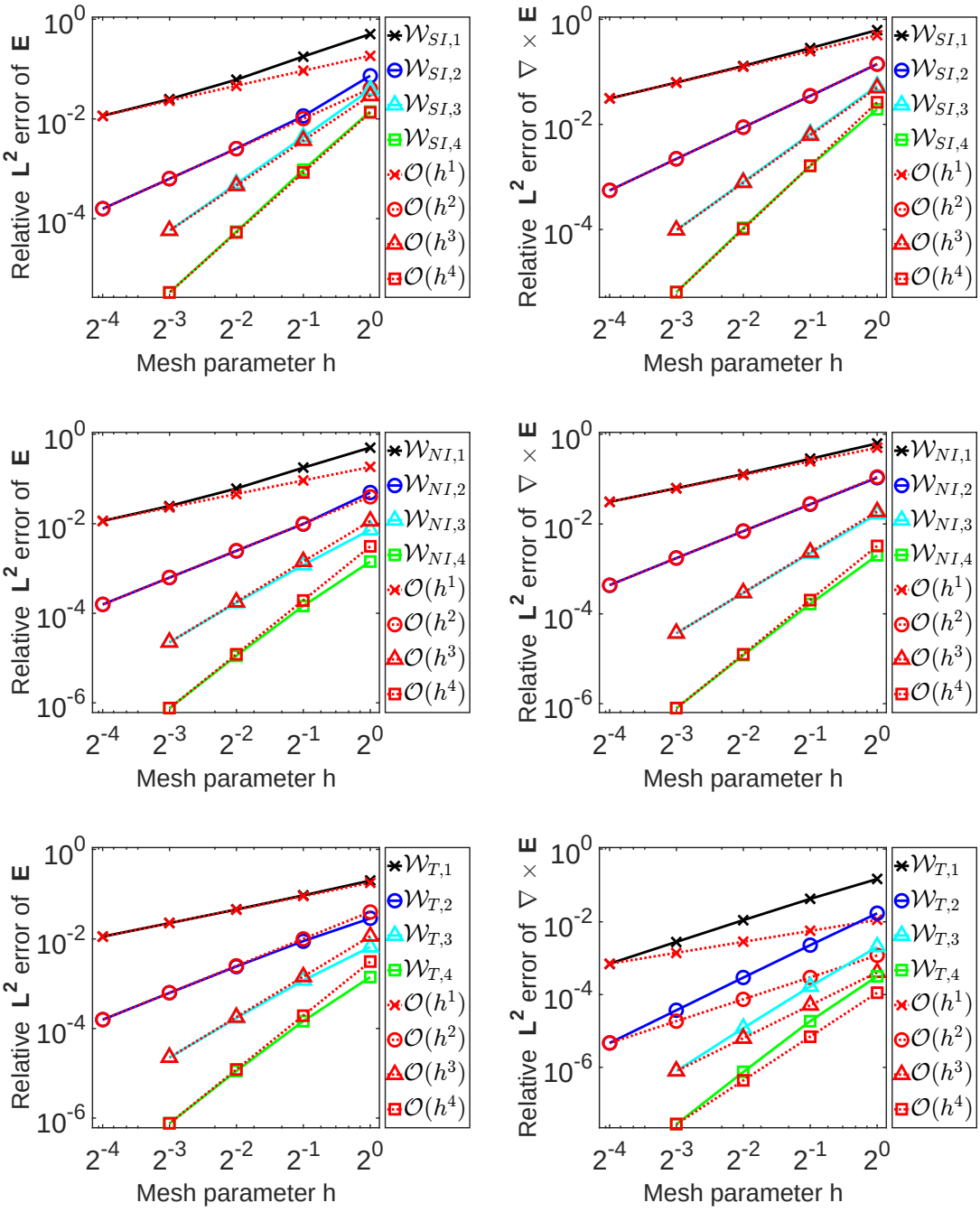


Figure 6.39: Relative error of the electric field $\boldsymbol{E}$ and its curl $\nabla \times \boldsymbol{E}$ in terms of the mesh parameter h. The calculations are obtained for the spaces $\mathcal{W}_{SI,p}$, $\mathcal{W}_{NI,p}$, and $\mathcal{W}_{T,p}$, which are compared to the ideal rates of $\mathcal{O}(h^p)$.

Coaxial Waveguide with a Non-Affinely Refined Regular Mesh

This numerical example shows how the convergence rate is affected by a change in the refinement method from quasi-affine refinement (as in the previous section)

to non-affine refinement. The applied sequence of fourth-order meshes is shown in Figure 6.41 and Figure 6.42, where the initial mesh is the same as in the previous section. The electric field and its error are shown in Figure 6.38 from the calculation on the finest mesh with the space of $\mathcal{W}_{T,4}$. It can be seen that the numerical error is much higher than in the previous case. Furthermore, the error pattern is determined by the pattern of the mesh and does not resemble the solution. The obtained convergence curves are shown in Figure 6.44. For all first-order finite element spaces $\mathcal{W}_{SI,1}$, $\mathcal{W}_{NI,1}$, and $\mathcal{W}_{T,1}$, the correct convergence rates $\mathcal{O}(h^1)$ are achieved for both $\boldsymbol{E}$ and $\nabla \times \boldsymbol{E}$. However, in all higher-order cases, the desired rates are not met. This behavior is expected since the utilized non-affine refinement (illustrated in Figure 6.25) only preserves the first-order continuity of the initial mapping. Hence, in terms of convergence rates, all the investigated spaces are equally insufficient.

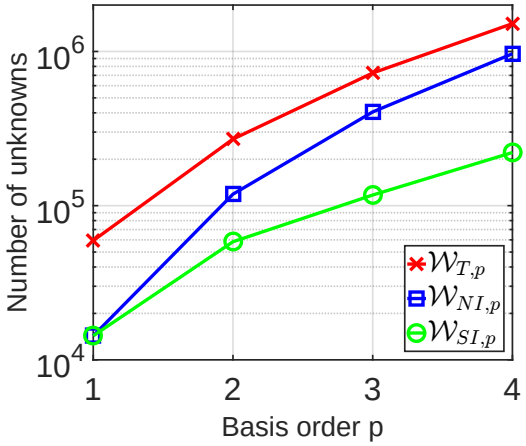


Figure 6.40: Number of unknowns for the finite element spaces $\mathcal{W}_{SI,p}$, $\mathcal{W}_{NI,p}$, and $\mathcal{W}_{T,p}$, with respect to the basis function order p. All values are obtained for the mesh with $h = 2^{-3}$.

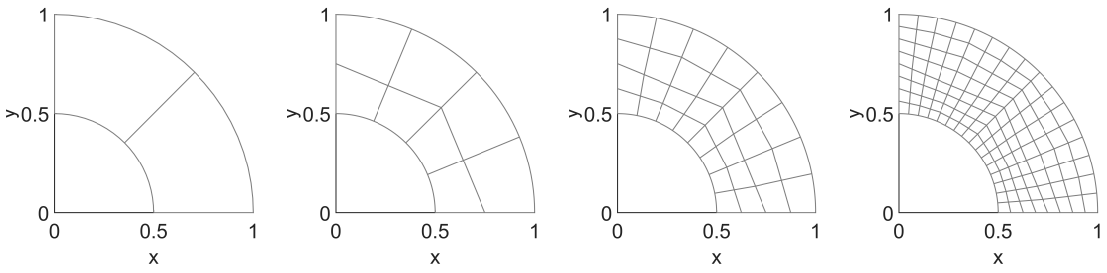


Figure 6.41: Side view of the mesh for the coaxial waveguide during non-affine refinement. The illustration includes the interior edges of the mesh.

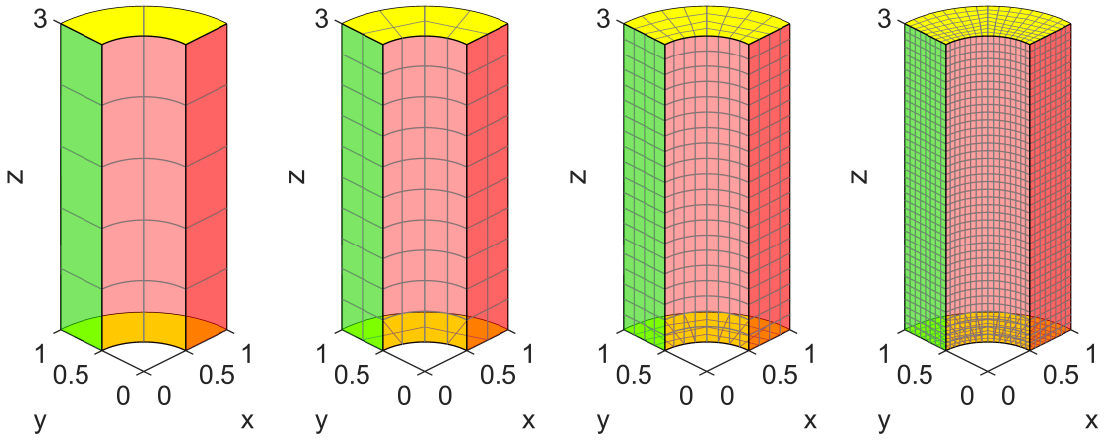


Figure 6.42: Boundaries of the coaxial waveguide mesh during non-affine refinement. The red faces are PEC, the green faces are PMC, and the yellow faces are wave ports.

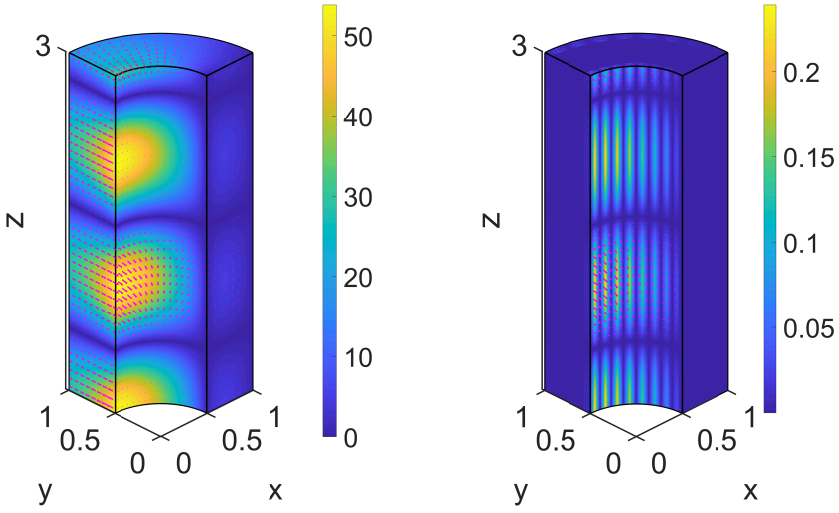


Figure 6.43: Magnitude of the calculated electric field strength (left) and its error with respect to the analytical solution (right). The calculation is obtained for the space $\mathcal{W}_{T,4}$ on the mesh with $h = 2^{-3}$. All values are given in coherent SI units.

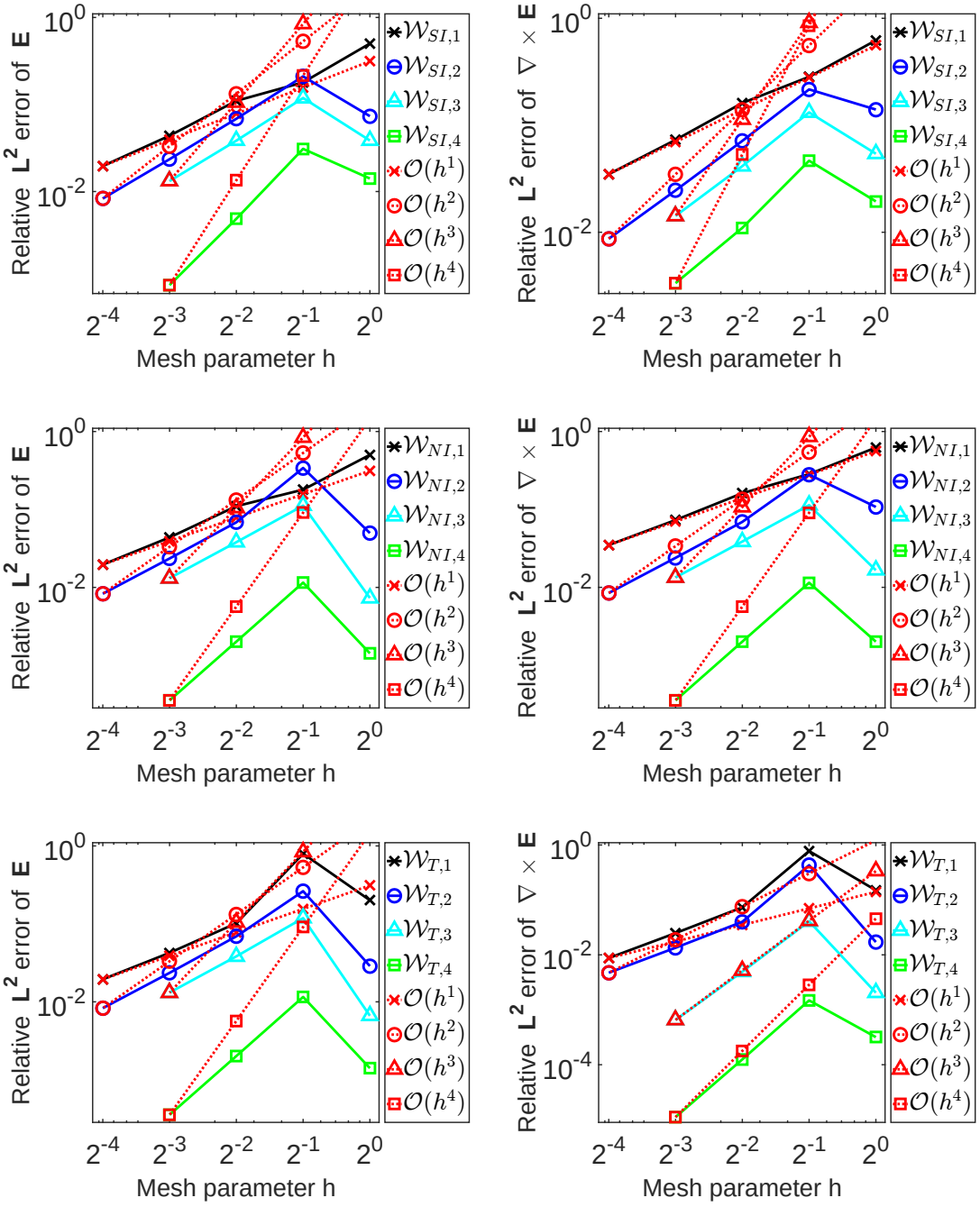


Figure 6.44: Relative error of the electric field $\boldsymbol{E}$ and its curl $\nabla \times \boldsymbol{E}$ in terms of the mesh parameter h. The calculations are obtained for the spaces $\mathcal{W}_{SI,p}$, $\mathcal{W}_{NI,p}$, and $\mathcal{W}_{T,p}$, which are compared to the ideal rates of $\mathcal{O}(h^p)$.

Coaxial Waveguide with a Quasi-Affinely Refined Non-Regular Mesh

To demonstrate the robustness of quasi-affine refinement, a similar coaxial waveguide problem is solved as in the previous examples but with a much worst initial mesh. All inner faces within the mesh are also curvilinear and not parallel to any other boundary. Moreover, all edges on the interior of the boundaries and on the interior of the domain are not aligned with any edge of the geometrical domain. The applied sequence of fourth-order meshes is shown in Figure 6.46 and Figure 6.47. Due to the quasi-affine refinement, the nodes of the refined elements are always displaced to the true boundary. Therefore, none of these elements is a parallelepiped. Even in an asymptotic sense, the boundary elements do not converge to parallelotope elements due to the repeated correction step. However, the continuity of the fourth-order geometry mappings of the initial elements is preserved by the quasi-affine refinement.

The calculated electric field and its error can be seen in Figure 6.48 for the finite element space $\mathcal{W}_{T,4}$ on the finest mesh. This solution is supposed to provide the lowest error since it is the largest space. It appears that the main source of error is the deformed elements on the curvilinear boundary. The calculated convergence rates are shown in Figure 6.49. In accordance with the theoretical predictions, the required order of convergence $\mathcal{O}(h^p)$ is reached in the $\boldsymbol{H}(\mathrm{curl})$- and $\boldsymbol{L}^2$-norm for all elements. Among them, the serendipity space $\mathcal{W}_{SI,4}$ has the lowest number of unknowns; see Figure 6.45. The other spaces result in higher accuracy at the price of a much higher number of unknowns. However, the space $\mathcal{W}_{SI,p}$ can afford a higher number of refinements that $\mathcal{W}_{T,p}$ or $\mathcal{W}_{NI,p}$ (if $p > 1$) for a given number of unknowns, which generally yields higher accuracy. Therefore, the iso-serendipity element is a promising new method.

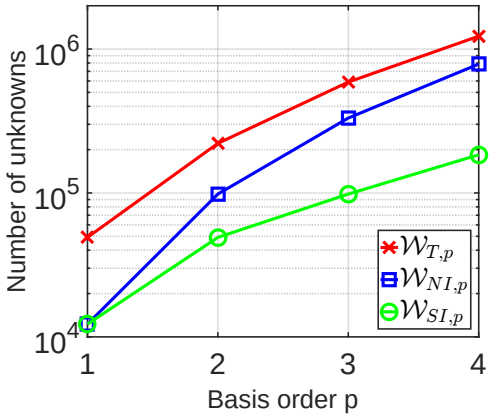


Figure 6.45: Number of unknowns for the finite element spaces $\mathcal{W}_{SI,p}$, $\mathcal{W}_{NI,p}$, and $\mathcal{W}_{T,p}$, with respect to the basis function order p. All values are obtained for the mesh with $h = 2^{-3}$.

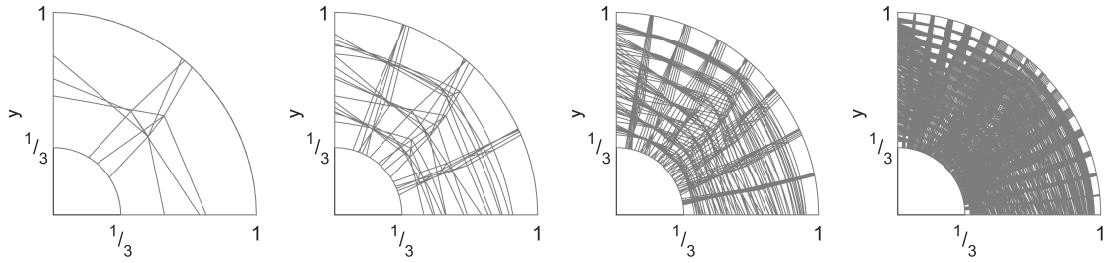


Figure 6.46: Side view of the non-regular mesh for the coaxial waveguide during quasi-affine refinement. The illustration includes the interior edges of the mesh.

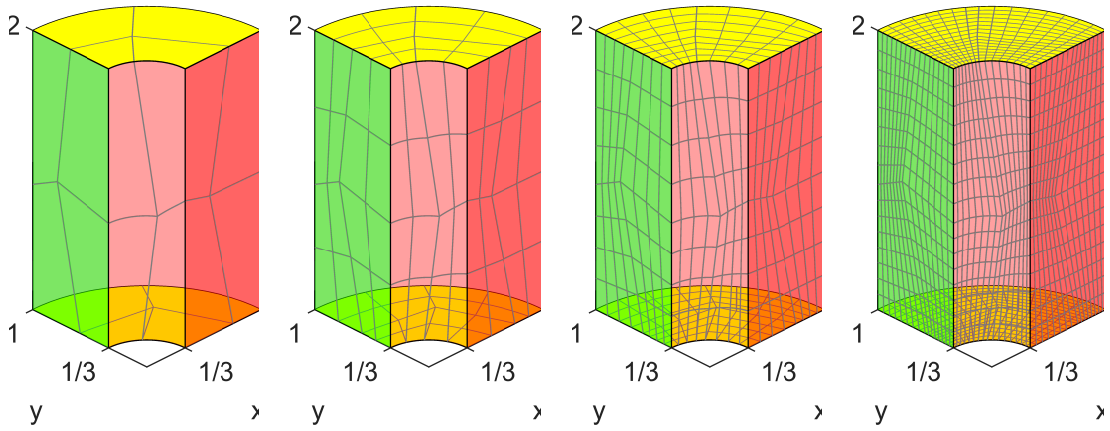


Figure 6.47: Boundaries of the non-regular mesh during quasi-affine refinement. The red faces are PEC, the green faces are PMC, and the yellow faces are the ports.

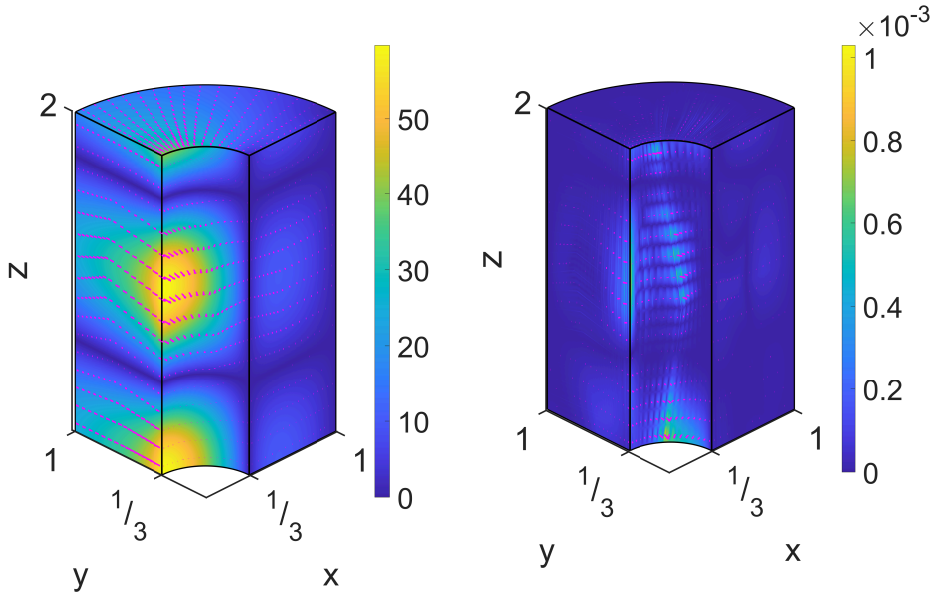


Figure 6.48: Magnitude of the calculated electric field strength (left) and its error with respect to the analytical solution (right). The calculation is obtained for the space $\mathcal{W}_{T,4}$ on the mesh with $h = 2^{-3}$. All values are given in coherent SI units.

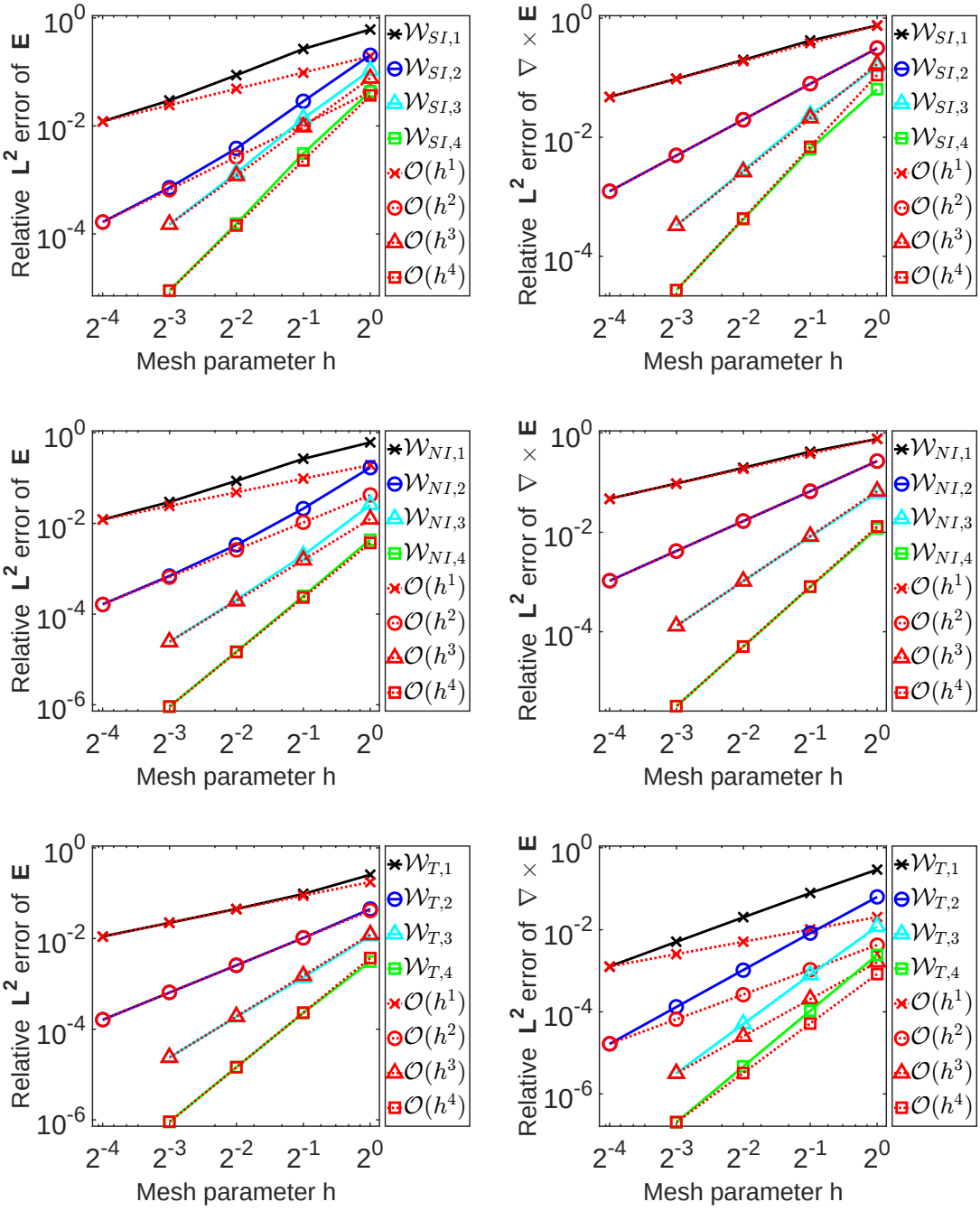


Figure 6.49: Relative error of the electric field $\boldsymbol{E}$ and its curl $\nabla \times \boldsymbol{E}$ in terms of the mesh parameter h. The calculations are obtained for the spaces $\mathcal{W}_{SI,p}$, $\mathcal{W}_{NI,p}$, and $\mathcal{W}_{T,p}$, which are compared to the ideal rates of $\mathcal{O}(h^p)$.

Scattering Parameter Calculation of a Rectangular Waveguide with a Conducting Post

The aim of this last example is to demonstrate the iso-serendipity method for models closer to realistic applications, where the geometry is partially curvilinear. In most practical applications, microwave devices are characterized based on a scattering parameter description, and the fields have only secondary importance. Therefore, the convergence of the scattering parameters is of primary importance.

For the model geometry, a rectangular waveguide section with the dimensions 10 mm $\times$ 39 mm $\times$ 50 mm is chosen, in which a round conducting post of radius 2 mm is embedded in the middle perpendicular to the H-plane. Such structures are widely used for constructing microwave filters [LM04, AKIH08], since they exhibit inductive behavior [PKCR13] in a lumped-element sense. From the modeling perspective, all boundaries and the cylindrical post are considered to be a perfect electric conductor, and the material properties in the interior are of the vacuum. The structure is terminated by two wave ports at the ends, of which the bottom one is excited by the incident electric field of the dominant TE_{01} mode [Poz11] at 5.5 GHz. Due to the symmetry of the geometry and the excitation, it suffices to model the half domain and impose a perfect magnetic boundary condition on the symmetry plane. The initial mesh and the magnitude of the electric field strength on a refined mesh are given in Figure 6.50.

According to the iso-serendipity method, fourth-order geometry mappings $\boldsymbol{g}_4 \in (\mathcal{V}_{S,4})^3$ and affine refinement are applied for the mesh and the finite element basis functions of the mixed-order serendipity elements $\mathcal{W}_{SI,4}$. Therefore, minimal number of unknowns are required for the ideal algebraic convergence rates. In contrast to the previous examples, the $L^2(\Omega)$-norm error of the reflection coefficient S_{11} is calculated in each refinement step, and its error is determined with reference to the solution for the largest (and most accurate) FE space, $\mathcal{W}_{T,4}$, on the finest possible mesh that I was able to compute. Note that the finite element convergence rate in energy is twice the convergence rate of the natural $\boldsymbol{H}(\mathrm{curl}, \Omega)$ intensity norm. Since the scattering parameters are related to the energy and are stationary quantities, the ideal theoretical rate is $\mathcal{O}(h^{2p})$. The obtained rates are shown in Figure 6.51. It can be clearly seen that the method tends to the ideal convergence rate for the higher refinement steps. The deviations from the ideal rates with the coarser meshes are due to the enlarged preasymptotic region since the smoothness of the geometry mappings affects the constant scaling of the convergence. Due to the improvement of geometry representation at each refinement step (quasi-affine refinement), the geometry mappings of the elements change next to the boundaries. Thus, the constant factor of the convergence rate changes in the preasymptotic region of convergence. Note that this effect can be significantly reduced by generating a good initial mesh where the curvilinear geometry mappings of the initial elements are already relatively smooth; see the coaxial waveguide examples.

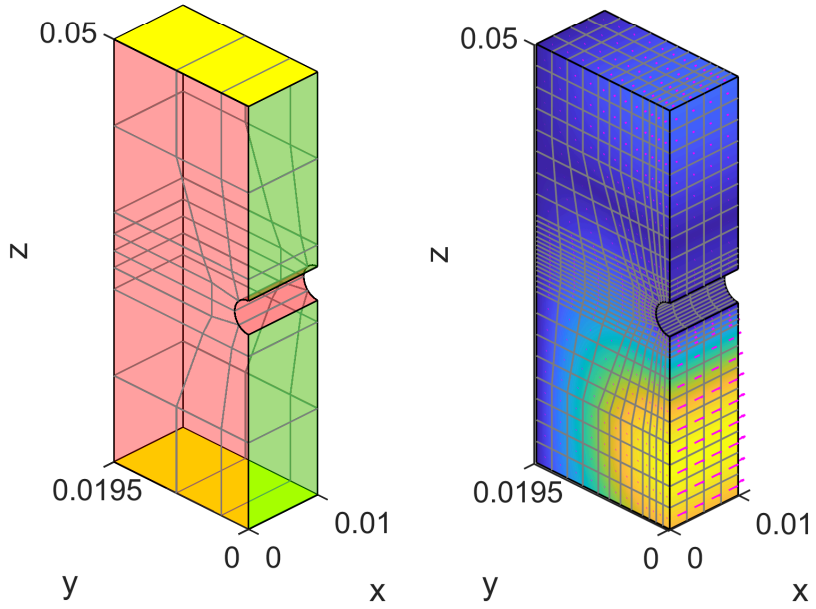


Figure 6.50: The initial mesh (left) and the magnitude of the calculated electric field strength on a refined mesh (right). The red faces are PEC, the green faces are PMC, and the yellow faces are wave ports.

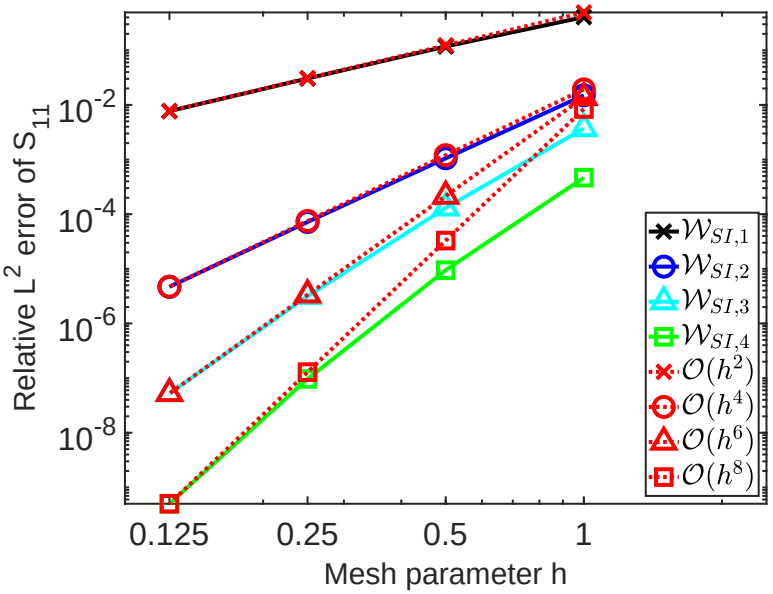


Figure 6.51: Relative L^2-norm error of the S_{11} parameter in terms of the mesh parameter h, compared to the corresponding ideal rates $\mathcal{O}(h^{2p})$.

6.7 Summary

This chapter has given an overview and provided alternative definitions for some of the commonly used H^1 and $\boldsymbol{H}(\mathrm{curl})$ hexahedral finite elements to which the convergence requirements specified in Chapter 3 are directly applicable. Moreover, a new serendipity element, the so-called mixed-order serendipity element, is presented for $\boldsymbol{H}(\mathrm{curl})$-conforming functions. Its main feature is the least number of unknowns while providing the asymptotic convergence rate $\mathcal{O}(h^p)$ in the $\boldsymbol{H}(\mathrm{curl})$-norm for the affine refinement of the mesh. Moreover, hierarchical basis functions are provided that have a separate subset of higher-order null space functions and are compatible with other tensor product elements. These properties have led to a new method, the mixed-order mixed-space finite element, which uses different finite element spaces within one mesh without affecting conformity. As a result, it is possible to use the smallest necessary space on each element, based on the type of refinement and the geometry mapping, and achieve the ideal rate of convergence globally.

For curvilinear geometries where only pointwise (interpolatory) information is available on the boundary, another new method is proposed, namely the iso-serendipity element. This element uses the functions of the defined serendipity spaces for both geometry interpolation and field representation. It is more restrictive than the mixed-order mixed-space element since it only allows the developed quasi-affine refinement of the mesh. Its advantage over affine refinement is that while the geometry representation improves with each refinement step, only some interpolation locations are required on the curvilinear boundary instead of the exact mapping. Consequently, the resulting number of unknowns is minimal, and the asymptotic convergence rate is $\mathcal{O}(h^p)$ in the $\boldsymbol{H}(\mathrm{curl})$-norm. Therefore, the number of unknowns with the iso-serendipity element is significantly less than with other known tensor product finite elements with the same convergence rates.

Chapter 7

Summary

The thesis concerns new higher-order finite element basis functions and their applications for curvilinear finite elements. In order to lay down the foundations for the different contributions, rigorous finite element error analyses are provided in Chapter 3, which verify and motivate the different methods. The new aspects lie in the alternative definition of the mesh parameter and mesh refinement and in their relation to the interpolated geometry mappings. The resulting different finite element error estimations provide new possibilities for maintaining the ideal convergence rates. As an outcome, general convergence requirements are provided in the presented framework for curvilinear finite elements with interpolated geometry mappings, considering different ways of mesh refinements.

The second contribution of the thesis, in Chapter 4, is an efficient calculation scheme for curvilinear finite element matrices. The method is based on a polynomial expansion of the material- and geometry-dependent metric terms of the finite element matrices, allowing the construction of universal matrices. For this purpose, a new set of hierarchical and pairwise orthonormal polynomials was developed, reflecting the topological symmetries of arbitrary tetrahedra. As an outcome, the element matrices can be calculated efficiently by the scaled matrix additions of the precalculated universal matrices. Based on a complexity analysis, this procedure is more efficient than competing methods. The unique feature is a proven upper representation limit for the required order of expansion and universal matrices. Hence, a finite number of universal matrices is sufficient for the exact calculation of the element matrices. Moreover, the expansion can be terminated at a much earlier point if the finite element convergence is in the asymptotic region. These well-defined lower and upper limits are unique features of the method and yield efficient and reliable calculations.

For $\boldsymbol{H}(\mathrm{div})$-conforming tetrahedral elements, a new set of basis functions is given in Chapter 5, which completes the sequence of the discrete de Rham complex with

the previously available H^1- and $\boldsymbol{H}(\mathrm{curl})$-conforming bases [Ing06]. The common properties of these bases are the orthogonality with respect to the interpolation operator and a subset of functions for the higher-order null space. These features lead to element matrices with advantageous numerical properties, such as a high level of sparsity and preservation of orthogonality for arbitrary straight-sided elements. The higher-order sections of the generated stiffness matrices are purely diagonal for arbitrary straight-sided elements. Moreover, the condition numbers of the resulting mass matrices are significantly lower than for competing basis functions from the literature.

Pertaining to the field of hexahedral finite elements, several original contributions are presented in Chapter 6. It is generally believed that serendipity spaces can only produce an $\mathcal{O}(h^p)$ rate of convergence in the $\boldsymbol{H}(\mathrm{curl})$-norm for parallelepiped meshes [ABB15], [DG19], [GK19]. Several different proofs in the literature reflect this result: [ABF02], [FGM11], [Ape98], [ABB15]. In order to overcome this problem, error estimations are provided for a general element, considering the effect of geometry interpolation and different mesh refinements. Based on this new viewpoint, modified definitions and error estimations are proposed for the best-known hexahedral finite elements and introduced a new mixed-order serendipity element for $\boldsymbol{H}(\mathrm{curl})$-conforming functions. The main advantage of these definitions is the resulting $\mathcal{O}(h^p)$ rate of convergence in the $\boldsymbol{H}(\mathrm{curl})$-norm, which applies to affinely refined curvilinear meshes. The proposed mixed-order serendipity element provides the least number of unknowns while allowing the construction of hierarchical basis functions. The presented basis functions are hierarchical, have a subset of null space functions, and are compatible with other tensor product finite elements. These properties led to the development of the mixed-order mixed-space finite element, using different finite element spaces within one mesh without affecting conformity. The smallest finite element space is chosen for each element which is sufficient for the theoretical convergence rate. This can extensively lower the number of unknowns for meshes where most elements are paralleled or affinely refined.

For curvilinear geometries, another new method is proposed: the iso-serendipity element. This element uses the functions of the defined serendipity spaces for both geometry interpolation and field representation. It is more restrictive than the mixed-order mixed-space element since it only allows a special way of mesh refinement, the quasi-affine refinement. Its advantage over affine refinement is that while the geometry representation improves with each refinement step, only some interpolation locations on the curvilinear boundary are required instead of the exact mapping. It is shown that quasi-affine refinement is equivalent to the affine refinement of an exact yet unknown mesh. Therefore, the proposed iso-serendipity element is guaranteed to converge with the ideal rate while requiring a minimal number of unknowns. To the author's best knowledge, the iso-serendipity element is the first method for curvilinear meshes which applies a serendipity space and provides $\mathcal{O}(h^p)$ rate of convergence in the $\boldsymbol{H}(\mathrm{curl})$-norm.

Appendix A

Hierarchical Orthonormal Basis for Tetrahedra

This section deals with the construction of polynomial functions used for the metric expansion of finite element matrices in Chapter 4 and Section 5.2.1.

The most straightforward way of generating an orthonormal basis is via the Gram-Schmidt method, applied to the monomials of linearly independent coordinates. This approach is applicable to arbitrary domains and produces basis functions that are hierarchical, greatly simplifying the degree elevation of orthogonal expansions. On the contrary, the resulting basis does not inherit any symmetry property from the underlying domain since the subsequent functions within the same order consist of increasing numbers of monomials. Symmetric functions are generally considered more advantageous since their evaluation and approximation may be performed more accurately and efficiently [JGBY16, BDO21, ZCL09, Coo97].

There also exist methods that start from tensor product domains or from the n-dimensional ball and apply reparameterization techniques for mapping it to the simplex [DX14, Koo75, GS11, OTV20, AAG20]. They produce hierarchical, pairwise orthogonal polynomials but do not reflect the natural symmetries of the simplex. Other classes of orthogonal polynomials are the monomial orthogonal bases and the corresponding biorthogonal bases [DX14, Sec. 5.2]. These are symmetry-preserving on the simplex but not hierarchical. Also, they do not provide orthogonality between the same-order functions.

Lately, several different approaches have been presented, combining some explicitly available orthogonal polynomials and a further orthogonalization and symmetrization step [HS98, FGS03, RA05, Sau07]. However, none of the previously mentioned bases has the three essential properties simultaneously: pairwise orthonormality, degree-ordered hierarchical functions, and symmetry over arbitrary tetrahedron.

These features are generally useful for adaptive approximation problems and particularly advantageous for the metric expansion method.

A.1 Basis Properties

The goal is to produce hierarchical and L^2 orthonormal (4.13) (Legendre-type) basis functions $b_q^l \in \widetilde{\mathcal{P}}_q$ that span the polynomial space $\mathbb{P}_p$ over the reference element $\widehat{\Omega}_e$. Here, $l \in \{1, \cdots L(q)\}$ and $\widetilde{\mathcal{P}}_q$ is an incremental space of degree q such that

$$\mathbb{P}_p(\widehat{\Omega}_e) = \widetilde{\mathcal{P}}_0 \oplus \cdots \oplus \widetilde{\mathcal{P}}_q \oplus \cdots \oplus \widetilde{\mathcal{P}}_p, \tag{A.1a}$$

$$\dim \mathbb{P}_p = (p+1)(p+2)(p+3)/6, \tag{A.1b}$$

$$\dim \widetilde{\mathcal{P}}_q = L(q) = (q+1)(q+2)/2. \tag{A.1c}$$

The reference element is defined in Section 2.6. The domain is parameterized by the linearly independent Cartesian coordinates $\widehat{\boldsymbol{r}} = [\widehat{r}_1, \widehat{r}_2, \widehat{r}_3]^T$ and determined by the coordinates of the vertices $\widehat{\boldsymbol{n}}_1 = [1,0,0]^T$, $\widehat{\boldsymbol{n}}_2 = [0,1,0]^T$, $\widehat{\boldsymbol{n}}_3 = [0,0,1]^T$ and $\widehat{\boldsymbol{n}}_4 = [0,0,0]^T$. Any vector of this reference element can be described in terms of barycentric coordinates $(\varphi_1, \ldots, \varphi_4)$, where $0 \leq \varphi_i \leq 1$. According to [Ska08], [Cox61, Section 13.7], the relationship to the linearly independent coordinates is given by

$$\begin{bmatrix} \widehat{r}_1 \\ \widehat{r}_2 \\ \widehat{r}_3 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} \varphi_1 \\ \varphi_2 \\ \varphi_3 \\ \varphi_4 \end{bmatrix}, \tag{A.2}$$

yielding the linear combination of the node coordinates

$$\widehat{\boldsymbol{r}} = \sum_i \widehat{\boldsymbol{n}}_i \varphi_i. \tag{A.3}$$

Their advantage is the uniform parameterization of the element since each barycentric coordinate is associated to a given node and gives a zero value on the other nodes. Moreover, the permutation of the node coordinates is equivalent to the permutation of the barycentric coordinates; see (A.3). Therefore, they are appropriate for describing symmetries over tetrahedral domains since the symmetric functions of barycentric coordinates are symmetric with respect to the geometry (defined by the nodes). Therefore, the aim is to obtain an ordered set of $b_q^l \in \widetilde{\mathcal{P}}_q$ functions as a vector

$$\widetilde{\boldsymbol{b}}_q = [\underbrace{b_q^1 \ldots b_q^i}_{\widetilde{\boldsymbol{b}}_{q,1}} \ldots \underbrace{b_q^k \ldots b_q^l}_{\widetilde{\boldsymbol{b}}_{q,s}} \ldots], \tag{A.4}$$

where the $\widetilde{\boldsymbol{b}}_{q,s}$ sub-vectors contain the sth set of tetrahedron symmetric functions.

A.2 Basis Symmetry

Further on, let us denote the node set of a tetrahedron by N, the edge set by E, the face set by F, and the volume by V. Moreover, let us denote a symmetry group of functions over a general tetrahedron by $S_{[\text{type}]}$, which consists of such a symmetrical functions $v(\varphi_i, \varphi_j, \varphi_k, \varphi_l)$ of the barycentric coordinates (A.2) that remain unchanged for some permutations of its arguments. The naturally arising symmetry groups for the different permutations are

$$\begin{aligned} S_V &= \{v \,|\, v(\varphi_i, \varphi_j, \varphi_k, \varphi_l) = v(\varphi_1, \varphi_2, \varphi_3, \varphi_4), \\ &\qquad \forall (i,j,k,l) = \text{Perm.}(1,2,3,4)\}, && \text{(A.5a)} \\ S_{NF} &= \{v \,|\, v(\varphi_i, \varphi_j, \varphi_k, \varphi_l) = v(\varphi_i, \varphi_b, \varphi_c, \varphi_d), \\ &\qquad \forall (b,c,d) = \text{Perm.}(j,k,l)\}, && \text{(A.5b)} \\ S_{EE} &= \{v \,|\, v(\varphi_i, \varphi_j, \varphi_k, \varphi_l) = v(\varphi_a, \varphi_b, \varphi_c, \varphi_d), \\ &\qquad \forall (a,b) = \text{Perm.}(i,j) \text{ and } (c,d) = \text{Perm.}(k,l)\}, && \text{(A.5c)} \\ S_E &= \{v \,|\, v(\varphi_i, \varphi_j, \varphi_k, \varphi_l) = v(\varphi_j, \varphi_i, \varphi_k, \varphi_l)\}\,. && \text{(A.5d)} \end{aligned}$$

Here, the functions of S_V show a complete volumetric symmetry, S_{NF} a node-type or face-type symmetry, and S_{EE} and S_E edge-type symmetries. For example, the uniform exponents of the barycentric coordinates are $\{\varphi_i^m\} \in S_{NF}$, $\{\varphi_i^m \varphi_j^m\} \in S_{EE}$ or $\{\varphi_k^m \varphi_l^n\} \in S_E$. Unfortunately, not all symmetries can be maintained for hierarchical bases where the functions are degree ordered. Since the set of barycentric coordinates already spans the first-order polynomial space and forms a node-type symmetry group S_{NF}, some adjustments are required to allow a separate representation for the constant term. This attempt results in breaking the natural symmetries. However, it is possible to maintain a structure by forming additional groups of antisymmetric functions

$$\begin{aligned} S_{\overline{V}} &= \{v \in S_{EE} \,|\, v(\varphi_i, \varphi_j, \varphi_k, \varphi_l) = -v(\varphi_k, \varphi_l, \varphi_i, \varphi_j)\}\,, && \text{(A.6a)} \\ S_{\overline{EE}} &= \{v \in S_E \,|\, v(\varphi_i, \varphi_j, \varphi_k, \varphi_l) = -v(\varphi_i, \varphi_j, \varphi_l, \varphi_k)\}\,, && \text{(A.6b)} \\ S_{\overline{E}} &= \{v \,|\, v(\varphi_i, \varphi_j, \varphi_k, \varphi_l) = -v(\varphi_j, \varphi_i, \varphi_k, \varphi_l)\}\,, && \text{(A.6c)} \end{aligned}$$

and a sequential, zig-zag symmetry along two complementary paths

$$S_Z = \{v \,|\, v(\varphi_i, \varphi_j, \varphi_k, \varphi_l) = v(\varphi_l, \varphi_k, \varphi_j, \varphi_i), \text{ for } (ijkl) \in \{(1234), (2413)\}\}\,. \tag{A.7}$$

The functions of $S_{\overline{V}}$ show volume-type and the functions of $S_{\overline{EE}}$ and $S_{\overline{E}}$ show edge-type antisymmetries. The dimensions of the defined groups are given in Table A.1.

Table A.1: Dimensions of the different symmetry and antisymmetry groups.

	S_V	S_{NF}	S_{EE}	S_E	$S_{\overline{V}}$	$S_{\overline{EE}}$	$S_{\overline{E}}$	S_Z	$\{\text{Perm.}(1,2,3,4)\}$
dim	1	4	6	12	3	6	12	2	24

A.3 The Initial Symmetric Basis

The first step of the basis construction is the generation of such incremental order sets $\widetilde{\mathcal{V}}_q$ of qth-order functions that span the completer-order polynomial space

$$\mathbb{P}_p(\widehat{\Omega}_e) = \widetilde{\mathcal{V}}_0 \oplus \cdots \oplus \widetilde{\mathcal{V}}_q \oplus \cdots \oplus \widetilde{\mathcal{V}}_p, \tag{A.8a}$$

$$\dim \widetilde{\mathcal{V}}_q = L(q) = (q+1)(q+2)/2, \tag{A.8b}$$

and possess symmetries on a general tetrahedron. With the notations introduced in the previous sections, the general definition of the incremental sets of basis functions are

$$\widetilde{\mathcal{V}}_0 = \{1\}, \qquad \widetilde{\mathcal{V}}_0 \in S_V, \tag{A.9a}$$

$$\widetilde{\mathcal{V}}_1 = \{(\varphi_3+\varphi_4)-(\varphi_1+\varphi_2),\ (\varphi_2+\varphi_4)-(\varphi_1+\varphi_3),\ (\varphi_1+\varphi_4)-(\varphi_2+\varphi_3)\}, \qquad \widetilde{\mathcal{V}}_1 \in S_{\overline{V}}, \tag{A.9b}$$

$$\widetilde{\mathcal{V}}_q = \left(\bigoplus_{ij} \widetilde{\mathcal{V}}_q^{ij}\right) \oplus \left(\bigoplus_{ijk} \widetilde{\mathcal{V}}_q^{ijk}\right) \oplus \widetilde{\mathcal{V}}_q^{Z} \oplus \widetilde{\mathcal{V}}_q^{1234} \quad \text{for } q \geq 2, \tag{A.9c}$$

where the higher-order functions are partitioned into different subsets $\widetilde{\mathcal{V}}_q^{ij}$, $\widetilde{\mathcal{V}}_q^{ijk}$, $\widetilde{\mathcal{V}}_q^{Z}$, and $\widetilde{\mathcal{V}}_q^{1234}$. In the beginning, let us set $\widetilde{\mathcal{V}}_q^{Z} = \emptyset$ as an empty set. The subset of edge-associated functions $\widetilde{\mathcal{V}}_q^{ij}$, with indices $ij \in (12, 13, 14, 23, 24, 34)$, is given as

$$\widetilde{\mathcal{V}}_2^{ij} = \{\varphi_i\varphi_j\}, \qquad \widetilde{\mathcal{V}}_2^{ij} \in S_{EE}, \tag{A.10a}$$

$$\widetilde{\mathcal{V}}_3^{ij} = \{\varphi_i\varphi_j(\varphi_i-\varphi_j)\}, \qquad \widetilde{\mathcal{V}}_3^{ij} \in S_{\overline{EE}}, \tag{A.10b}$$

$$\widetilde{\mathcal{V}}_q^{ij} = \varphi_i\varphi_j\widetilde{\mathcal{V}}_{q-2}^{ij} \quad \text{for } q \geq 4, \qquad \widetilde{\mathcal{V}}_q^{ij} \in S_{(\cdot)}. \tag{A.10c}$$

The subset of face-associated functions $\widetilde{\mathcal{V}}_q^{ijk}$, with indices $ijk \in (123, 124, 134, 234)$, is given as

$$\widetilde{\mathcal{V}}_q^{ijk} = \emptyset \quad \text{for } q \leq 2, \tag{A.11a}$$

$$\widetilde{\mathcal{V}}_3^{ijk} = \{\varphi_i\varphi_j\varphi_k\}, \qquad \widetilde{\mathcal{V}}_3^{ijk} \in S_{NF}, \tag{A.11b}$$

$$\widetilde{\mathcal{V}}_4^{ijk} = \{\varphi_i\varphi_j\varphi_k(\varphi_k-\varphi_i),\ \varphi_i\varphi_j\varphi_k(\varphi_k-\varphi_j)\}, \tag{A.11c}$$

$$\widetilde{\mathcal{V}}_q^{ijk} = \underbrace{\varphi_i\varphi_j\varphi_k(\widetilde{\mathcal{V}}_{q-3}^{ij} \oplus \widetilde{\mathcal{V}}_{q-3}^{ik} \oplus \widetilde{\mathcal{V}}_{q-3}^{jk})}_{\in S_E \text{ or } S_{\overline{E}}} \oplus \underbrace{\varphi_i\varphi_j\varphi_k\widetilde{\mathcal{V}}_{q-3}^{ijk}}_{\in S_E,\ S_{\overline{E}}, \text{ or } S_{NF}} \quad \text{for } q \geq 5. \tag{A.11d}$$

The subset of volume-associated functions $\widetilde{\mathcal{V}}_q^{1234}$ is given as

$$\widetilde{\mathcal{V}}_q^{1234} = \emptyset \quad \text{for } q \leq 3, \tag{A.12a}$$

$$\widetilde{\mathcal{V}}_q^{1234} = \varphi_1\varphi_2\varphi_3\varphi_4\widetilde{\mathcal{V}}_{q-4} \quad \text{for } q \geq 4, \qquad \widetilde{\mathcal{V}}_q^{1234} \in S_{(\cdot)}. \tag{A.12b}$$

The dimensions of these subsets are

$$\dim \widetilde{\mathcal{V}}_q^{ij} = 1, \qquad \text{for } q \geq 2, \tag{A.13a}$$

$$\dim \widetilde{\mathcal{V}}_q^{ijk} = k - 2, \qquad \text{for } q \geq 2, \tag{A.13b}$$

$$\dim \widetilde{\mathcal{V}}_q^{1234} = \dim \widetilde{\mathcal{V}}_{q-4} = L(q-4) = (q-3)(q-2)/2, \qquad \text{for } q \geq 2. \tag{A.13c}$$

It can be shown by using (A.8) and (A.13) that the generated functions form a valid basis since the proposed set span a polynomial space with the correct dimension

$$\begin{aligned} \dim \widetilde{\mathcal{V}}_q &= 6 \dim \widetilde{\mathcal{V}}_q^{ij} + 4 \dim \widetilde{\mathcal{V}}_q^{ijk} + \dim \widetilde{\mathcal{V}}_q^{1234}, \\ &= 6 + 4(q-2) + (q-3)(q-2)/2 = (q+1)(q+2)/2 = \text{(A.8b)}. \end{aligned} \tag{A.14}$$

A.3.1 The Problem of Face-Type Symmetries

The only remaining problem of the previous initial basis is that the functions are grouped into symmetry or antisymmetry groups except for $\widetilde{\mathcal{V}}_4^{ijk}$ (A.11c) since their dimension does not match any group

$$4 \text{ dim } \widetilde{\mathcal{V}}_4^{ijk} = 8 \neq \text{dim } S_{(\cdot)}, \qquad \forall(\cdot) \in \left\{V, NF, EE, E, \overline{V}, \overline{EE}, \overline{E}, Z\right\}. \tag{A.15}$$

Due to geometrical reasons, any complete symmetry group associated with triangular faces would require 1 or 3 functions for each face. Moreover, any complete face-associated antisymmetry group would require 3 functions per face. In the present case, this is not possible since the three available function $\left\{\varphi_i^2\varphi_j\varphi_k, \varphi_i\varphi_j^2\varphi_k, \varphi_i\varphi_j\varphi_k^2\right\}$ for $\widetilde{\mathcal{V}}_4^{ijk}$ would be linearly dependent on the face with the lower-order function $\varphi_i\varphi_j\varphi_k$. To resolve the problem, we decided to replace the face-type sets $\widetilde{\mathcal{V}}_4^{ijk}$ with a correction set $\widetilde{\mathcal{V}}_4^{Z}$, consisting of such linear combinations of the previous functions that result in edge-type $S_{\overline{EE}}$ and zig-zag type S_Z symmetries. Therefore, all functions belong to a complete group, and the dimension of the basis is correct

$$4 \text{ dim } \widetilde{\mathcal{V}}_4^{ijk} = 8 = \text{dim } \widetilde{\mathcal{V}}_4^{Z} = \text{dim } S_{\overline{EE}} + \text{dim } S_Z = 6 + 2. \tag{A.16}$$

Since the basis generation (A.11d) recursively multiplies the lower-order functions with the face monomial $\varphi_i\varphi_j\varphi_k$, the same corrections are required in the higher-order cases

$$\left\{\begin{array}{c} \bigoplus\limits_{ijk} \varphi_i^n\varphi_j^n\varphi_k^n\widetilde{\mathcal{V}}_4^{ijk} \\ \widetilde{\mathcal{V}}_q^{Z} = \emptyset, \end{array}\right\} \Longrightarrow \left\{\begin{array}{c} \bigoplus\limits_{ijk} \varphi_i^n\varphi_j^n\varphi_k^n\widetilde{\mathcal{V}}_4^{ijk} = \emptyset \\ \widetilde{\mathcal{V}}_q^{Z} \neq \emptyset \end{array}\right\}, \quad \text{for } n \in \mathbb{N}_0 \text{ and } q = 3n + 4. \tag{A.17}$$

The correction set is defined as

$$\widetilde{\mathcal{V}}_q^Z = \underbrace{\bigoplus_{ij}\{(\varphi_k\varphi_l)^{n+1}(\varphi_i^{n+2}-\varphi_j^{n+2})\}}_{\in S_{\overline{EE}}} \oplus \underbrace{\{v_n^Z(\varphi_1,\varphi_2,\varphi_3,\varphi_4), v_n^Z(\varphi_2,\varphi_4,\varphi_1,\varphi_3)\}}_{\in S_Z},$$
$$\text{for } \frac{q-4}{3} = n \in \mathbb{N}_0, \tag{A.18}$$

and the zig-zag functions v_n^Z are given as

$$\begin{aligned} v_n^Z(\varphi_1,\varphi_2,\varphi_3,\varphi_4) =& (\varphi_1\varphi_2)^{n+1}(\varphi_3^{n+2}+\varphi_4^{n+2}) - \\ & (\varphi_2\varphi_3)^{n+1}(\varphi_4^{n+2}+\varphi_1^{n+2}) + (\varphi_3\varphi_4)^{n+1}(\varphi_1^{n+2}+\varphi_2^{n+2}). \end{aligned} \tag{A.19}$$

As an outcome, the structure of the initial basis (A.9c) remains the same, and the higher-order functions are generated recursively with (A.10c), (A.11d), and (A.12b). The only differences are $\widetilde{\mathcal{V}}_4^{ijk} = \emptyset$ and there is an additional correction set $\widetilde{\mathcal{V}}_q^Z$ for $q = 4, 7, 10, \cdots, 4+3n$.

A.3.2 Recursively Inherited Symmetries

In the higher-order case, all incremental order subsets are generated recursively by multiplying the already available lower-order functions with either the corresponding edge monomial $\varphi_i\varphi_j$, face monomial $\varphi_i\varphi_j\varphi_k$, or volume monomial $\varphi_i\varphi_j\varphi_k\varphi_l$ (except for $\widetilde{\mathcal{V}}_q^Z$). Since these monomials are symmetric, all the recursively obtained higher-order functions inherit the same symmetry or antisymmetry of the lower-order functions. For example, all edge-associated functions are generated by (A.10c), yielding alternating symmetric and antisymmetric functions

$$\widetilde{\mathcal{V}}_2^{ij} = \{\varphi_i\varphi_j\}, \qquad \widetilde{\mathcal{V}}_2^{ij} \in S_{EE}, \tag{A.20a}$$
$$\widetilde{\mathcal{V}}_3^{ij} = \{\varphi_i\varphi_j(\varphi_i-\varphi_j)\}, \qquad \widetilde{\mathcal{V}}_3^{ij} \in S_{\overline{EE}}, \tag{A.20b}$$
$$\widetilde{\mathcal{V}}_4^{ij} = \varphi_i\varphi_j\widetilde{\mathcal{V}}_2^{ij} = \varphi_i\varphi_j\{\varphi_i\varphi_j\} = \{\varphi_i^2\varphi_j^2\}, \qquad \widetilde{\mathcal{V}}_4^{ij} \in S_{EE}, \tag{A.20c}$$
$$\widetilde{\mathcal{V}}_5^{ij} = \varphi_i\varphi_j\widetilde{\mathcal{V}}_3^{ij} = \varphi_i\varphi_j\{\varphi_i\varphi_j(\varphi_i-\varphi_j)\} = \{\varphi_i^2\varphi_j^2(\varphi_i-\varphi_j)\}, \quad \widetilde{\mathcal{V}}_5^{ij} \in S_{\overline{EE}}. \tag{A.20d}$$

The face-associated functions are generated by (A.11d), reusing both the lower-order edge- and face-associated functions. Therefore, the higher-order functions are partitioned into multiple different symmetry groups

$$\widetilde{\mathcal{V}}_3^{ijk} = \{\varphi_i\varphi_j\varphi_k\}, \qquad \widetilde{\mathcal{V}}_3^{ijk} \in S_{NF}, \tag{A.21a}$$
$$\widetilde{\mathcal{V}}_4^{ijk} = \emptyset, \tag{A.21b}$$
$$\widetilde{\mathcal{V}}_5^{ijk} = \varphi_i\varphi_j\varphi_k(\widetilde{\mathcal{V}}_2^{ij} \oplus \widetilde{\mathcal{V}}_2^{ik} \oplus \widetilde{\mathcal{V}}_2^{jk}) \ \oplus \ \varphi_i\varphi_j\varphi_k\widetilde{\mathcal{V}}_2^{ijk}, \tag{A.21c}$$
$$= \underbrace{\varphi_i\varphi_j\varphi_k(\{\varphi_i\varphi_j\} \oplus \{\varphi_i\varphi_k\} \oplus \{\varphi_j\varphi_k\})}_{\in S_E} \oplus \emptyset, \tag{A.21d}$$

$$= \{\varphi_i^2\varphi_j^2\varphi_k,\ \varphi_i^2\varphi_j\varphi_k^2,\ \varphi_i\varphi_j^2\varphi_k^2\}, \qquad \widetilde{\mathcal{V}}_5^{ijk} \in S_E, \tag{A.21e}$$

$$\widetilde{\mathcal{V}}_6^{ijk} = \varphi_i\varphi_j\varphi_k(\widetilde{\mathcal{V}}_3^{ij} \oplus \widetilde{\mathcal{V}}_3^{ik} \oplus \widetilde{\mathcal{V}}_3^{jk}) \ \oplus \ \varphi_i\varphi_j\varphi_k\widetilde{\mathcal{V}}_3^{ijk}, \tag{A.21f}$$

$$= \underbrace{\{\varphi_i^2\varphi_j^2\varphi_k(\varphi_i-\varphi_j),\ \varphi_i^2\varphi_j\varphi_k^2(\varphi_i-\varphi_k),\ \varphi_i\varphi_j^2\varphi_k^2(\varphi_j-\varphi_k)\}}_{\in S_{\overline{E}}} \oplus \underbrace{\{\varphi_i^2\varphi_j^2\varphi_k^2\}}_{\in S_{NF}}. \tag{A.21g}$$

The volume-associated functions are generated by (A.12b), reusing all the functions of the $(q-4)$th-order incremental space (A.9c). Thus, all the previously defined symmetry groups will appear

$$\widetilde{\mathcal{V}}_4^{1234} = \varphi_1\varphi_2\varphi_3\varphi_4\widetilde{\mathcal{V}}_0 = \varphi_1\varphi_2\varphi_3\varphi_4\{1\} = \{\varphi_1\varphi_2\varphi_3\varphi_4\}, \qquad \widetilde{\mathcal{V}}_4^{1234} \in S_V, \tag{A.22a}$$

$$\widetilde{\mathcal{V}}_5^{1234} = \varphi_1\varphi_2\varphi_3\varphi_4\widetilde{\mathcal{V}}_1, \qquad \widetilde{\mathcal{V}}_5^{1234} \in S_{\overline{V}}, \tag{A.22b}$$

$$\widetilde{\mathcal{V}}_6^{1234} = \varphi_1\varphi_2\varphi_3\varphi_4\widetilde{\mathcal{V}}_2 = \varphi_1\varphi_2\varphi_3\varphi_4\left(\bigoplus_{ij}\widetilde{\mathcal{V}}_2^{ij}\right), \qquad \widetilde{\mathcal{V}}_6^{1234} \in S_{EE}. \tag{A.22c}$$

A.4 Symmetric Orthonormalization

A.4.1 Orthogonalization Between Different Orders

Now that an initial basis is available with subsets of symmetric functions, the remaining task is an orthonormalization procedure that transforms the functions of each symmetry group symmetrically. The first step is to place the functions of $\widetilde{\mathcal{V}}_0, ..., \widetilde{\mathcal{V}}_q$ in ascending order into a row vector $\boldsymbol{v}_q$, and apply an ordering within each order according to (A.9c). Let us denote the row vector of the functions in the highest order incremental space $\widetilde{\mathcal{V}}_q$ by $\widetilde{\boldsymbol{v}}_q$, which gives $\boldsymbol{v}_q = [\boldsymbol{v}_{q-1}, \widetilde{\boldsymbol{v}}_q]$. Thus, the orthogonalization of $\widetilde{\boldsymbol{v}}_q$ with respect to the lower-order functions $\boldsymbol{v}_{q-1}$ can be uniquely obtained in the form of

$$\widetilde{\boldsymbol{v}}_q^{\text{low}} = \widetilde{\boldsymbol{v}}_q + \boldsymbol{v}_{q-1}\boldsymbol{A}_q, \tag{A.23}$$

where the orthogonalization matrix $\boldsymbol{A}_q$ is determined by the orthogonality requirement

$$\int_{\widehat{\Omega}_e} \boldsymbol{v}_{q-1}^T\widetilde{\boldsymbol{v}}_q^{\text{low}}\mathrm{d}\widehat{\boldsymbol{r}} = \int_{\widehat{\Omega}_e} \boldsymbol{v}_{q-1}^T\widetilde{\boldsymbol{v}}_q\,\mathrm{d}\widehat{\boldsymbol{r}} + \int_{\widehat{\Omega}_e} \boldsymbol{v}_{q-1}^T\boldsymbol{v}_{q-1}\,\mathrm{d}\widehat{\boldsymbol{r}}\boldsymbol{A}_q = \boldsymbol{0}, \tag{A.24a}$$

$$\boldsymbol{A}_q = -\left[\int_{\widehat{\Omega}_e} \boldsymbol{v}_{q-1}^T\boldsymbol{v}_{q-1}\,\mathrm{d}\widehat{\boldsymbol{r}}\right]^{-1}\int_{\widehat{\Omega}_e} \boldsymbol{v}_{q-1}^T\widetilde{\boldsymbol{v}}_q\,\mathrm{d}\widehat{\boldsymbol{r}}. \tag{A.24b}$$

This step can be interpreted as a blockwise Gram-Schmidt method. The following way obtained functions are shown in Table A.2 up to order 3 (additional scaling is applied to avoid rational coefficients). The obtained functions of $\widetilde{\boldsymbol{v}}_q^{\text{low}}$ are hierarchical, symmetrical, and only orthogonal with respect to the lower-order functions.

Table A.2: Symmetric basis functions that are orthogonal with respect to lower-order functions.

	q	l	$\left[\widetilde{\boldsymbol{v}}_q^{\text{low}}\right]_l$
S_V	0	1	$\sqrt{6}$,
$S_{\overline{V}}$	1	1	$\sqrt{30}(1-2(\varphi_1+\varphi_2))$,
	1	2	$\sqrt{30}(1-2(\varphi_2+\varphi_3))$,
	1	3	$\sqrt{30}(1-2(\varphi_3+\varphi_1))$,
S_{EE}	2	1	$1-5(\varphi_1+\varphi_2)+30\varphi_1\varphi_2$,
	2	2	$1-5(\varphi_1+\varphi_3)+30\varphi_1\varphi_3$,
	2	3	$1-5(\varphi_1+\varphi_4)+30\varphi_1\varphi_4$,
	2	4	$1-5(\varphi_2+\varphi_3)+30\varphi_2\varphi_3$,
	2	5	$1-5(\varphi_2+\varphi_4)+30\varphi_2\varphi_4$,
	2	6	$1-5(\varphi_3+\varphi_4)+30\varphi_3\varphi_4$,
$S_{\overline{EE}}$	3	1	$(2-7(\varphi_1+\varphi_2)+56\varphi_1\varphi_2)(\varphi_1-\varphi_2)$,
	3	2	$(2-7(\varphi_1+\varphi_3)+56\varphi_1\varphi_3)(\varphi_1-\varphi_3)$,
	3	3	$(2-7(\varphi_1+\varphi_4)+56\varphi_1\varphi_4)(\varphi_1-\varphi_4)$,
	3	4	$(2-7(\varphi_2+\varphi_3)+56\varphi_2\varphi_3)(\varphi_2-\varphi_3)$,
	3	5	$(2-7(\varphi_2+\varphi_4)+56\varphi_2\varphi_4)(\varphi_2-\varphi_4)$,
	3	6	$(2-7(\varphi_3+\varphi_4)+56\varphi_3\varphi_4)(\varphi_3-\varphi_4)$,
S_{NF}	3	7	$-1+6(\varphi_1+\varphi_2+\varphi_3)-42(\varphi_1\varphi_2+\varphi_1\varphi_3+\varphi_2\varphi_3)+336\varphi_1\varphi_2\varphi_3$,
	3	8	$-1+6(\varphi_1+\varphi_2+\varphi_4)-42(\varphi_1\varphi_2+\varphi_1\varphi_4+\varphi_2\varphi_4)+336\varphi_1\varphi_2\varphi_4$,
	3	9	$-1+6(\varphi_1+\varphi_3+\varphi_4)-42(\varphi_1\varphi_3+\varphi_1\varphi_4+\varphi_3\varphi_4)+336\varphi_1\varphi_3\varphi_4$,
	3	10	$-1+6(\varphi_2+\varphi_3+\varphi_4)-42(\varphi_2\varphi_3+\varphi_2\varphi_4+\varphi_3\varphi_4)+336\varphi_2\varphi_3\varphi_4$,

A.4.2 Orthogonalization Between Same-Order Functions

The remaining task is a symmetrical orthogonalization within the same order functions via $\boldsymbol{B}_q$ coefficient matrices

$$\widetilde{\boldsymbol{b}}_q = \widetilde{\boldsymbol{v}}_q^{\text{low}} \boldsymbol{B}_q. \tag{A.25}$$

Orthogonalization Between Different Groups

In order to obtain orthogonality between the functions of different symmetry groups within $\widetilde{\boldsymbol{v}}_q^{\text{low}}$, the same procedure can be continued as between the different order functions. Let us denote the row vector of the functions in the first $(s-1)$th symmetry group by $\boldsymbol{v}_{q,s-1}^{\text{low}}$ and the sth group by $\widetilde{\boldsymbol{v}}_{q,s}^{\text{low}}$. Thus, a similar decomposition is possible as in the initial case (A.23),

$$\widetilde{\boldsymbol{v}}_q^{\text{low}} = [\underbrace{\widetilde{\boldsymbol{v}}_{q,1}^{\text{low}} \cdots \widetilde{\boldsymbol{v}}_{q,s-1}^{\text{low}}}_{\boldsymbol{v}_{q,s-1}^{\text{low}}} \widetilde{\boldsymbol{v}}_{q,s}^{\text{low}}, ...] \tag{A.26}$$

and the orthogonalization can be uniquely obtained in the form of

$$\widetilde{\boldsymbol{v}}_{q,s}^{\text{ort}} = \widetilde{\boldsymbol{v}}_{q,s}^{\text{low}} + \boldsymbol{v}_{q,s-1}^{\text{low}} \boldsymbol{A}_{q,s}, \tag{A.27}$$

where the orthogonalization matrix $\boldsymbol{A}_{q,s}$ is determined by

$$\boldsymbol{A}_{q,s} = -\left[\int_{\widehat{\Omega}_e} \left(\boldsymbol{v}_{q,s-1}^{\text{low}}\right)^T \boldsymbol{v}_{q,s-1}^{\text{low}} \, \mathrm{d}\widehat{\boldsymbol{r}}\right]^{-1} \int_{\widehat{\Omega}_e} \left(\boldsymbol{v}_{q,s-1}^{\text{low}}\right)^T \widetilde{\boldsymbol{v}}_{q,s}^{\text{low}} \, \mathrm{d}\widehat{\boldsymbol{r}}. \tag{A.28}$$

Orthogonalization Within Groups

The single remaining problem is the symmetric orthonormalization of functions within each $\widetilde{\boldsymbol{v}}_{q,s}^{\text{ort}}$ symmetry group. Let us denote a vector of orthonormal functions by $\widetilde{\boldsymbol{b}}_{q,s}$, which is obtained by an orthonormalization matrix $\boldsymbol{B}_{q,s}$ as

$$\widetilde{\boldsymbol{b}}_{q,s} = \widetilde{\boldsymbol{v}}_{q,s}^{\text{ort}} \boldsymbol{B}_{q,s}. \tag{A.29}$$

The orthogonality requirement between these functions yields

$$\int_{\widehat{\Omega}_e} \widetilde{\boldsymbol{b}}_{q,s}^T \widetilde{\boldsymbol{b}}_{q,s} \mathrm{d}\widehat{\boldsymbol{r}} = \boldsymbol{B}_{q,s}^T \int_{\widehat{\Omega}_e} \left(\widetilde{\boldsymbol{v}}_{q,s}^{\text{ort}}\right)^T \left(\widetilde{\boldsymbol{v}}_{q,s}^{\text{ort}}\right) \mathrm{d}\widehat{\boldsymbol{r}} \, \boldsymbol{B}_{q,s} = \boldsymbol{B}_{q,s}^T \boldsymbol{G}_{q,s} \boldsymbol{B}_{q,s} = \boldsymbol{I}, \tag{A.30}$$

where $\boldsymbol{I}$ is the identity matrix and $\boldsymbol{G}_{q,s}$ is the Gram matrix. Since $\boldsymbol{G}_{q,s}$ is an invertible symmetric square matrix, it can be orthogonalized

$$\boldsymbol{G}_{q,s} = \boldsymbol{V}_{q,s}\boldsymbol{D}_{q,s}\boldsymbol{V}_{q,s}^{T}, \tag{A.31}$$

with eigenvectors contained in $\boldsymbol{V}_{q,s}$ and the corresponding eigenvalues at the diagonal of matrix $\boldsymbol{D}_{q,s}$. Due to the positive definiteness of the Gramm matrix, all eigenvalues are real and positive. It can be seen that with the symmetric matrix

$$\boldsymbol{B}_{q,s} = \boldsymbol{G}_{q,s}^{-1/2} = \boldsymbol{V}_{q,s}\boldsymbol{D}_{q,s}^{-1/2}\boldsymbol{V}_{q,s}^{T}, \tag{A.32}$$

the orthogonality of (A.30) is satisfied since $\boldsymbol{V}_{q,s}$ is orthogonal

$$\boldsymbol{B}_{q,s}^{T}\boldsymbol{G}_{q,s}\boldsymbol{B}_{q,s} = \left(\boldsymbol{V}_{q,s}\boldsymbol{D}_{q,s}^{-1/2}\boldsymbol{V}_{q,s}^{T}\right)\boldsymbol{V}_{q,s}\boldsymbol{D}_{q,s}\boldsymbol{V}_{q,s}^{T}\left(\boldsymbol{V}_{q,s}\boldsymbol{D}_{q,s}^{-1/2}\boldsymbol{V}_{q,s}^{T}\right) = \boldsymbol{I}. \tag{A.33}$$

This symmetric orthogonalization within the same symmetry group turned out to be the same as Löwdin's symmetric orthogonalization for orbitals [Löw50]. It has been shown that $\boldsymbol{B}_{q,s}$ is unique for a given $\widetilde{\boldsymbol{v}}_{q,s}^{\mathrm{ort}}$ [AEG80]. Moreover, the generated set of functions remains unchanged for the permutation or orthogonal transformation of the vector $\widetilde{\boldsymbol{v}}_{q,s}^{\mathrm{ort}}$ [AEG80]. Thus, the associated rotations of the initial symmetric functions result in the same orthogonalization [May02, May04, RSS08]. Consequently, the generated set $\widetilde{\boldsymbol{b}}_{q,s}$ will inherit the symmetry of the starting set $\widetilde{\boldsymbol{v}}_{q,s}^{\mathrm{ort}}$ [AEG80]. An additional feature of Löwdin's orthogonalization is that it minimizes the pairwise L^2-norm error of the initial and orthogonal functions

$$\sum_i \left\| \left[\widetilde{\boldsymbol{b}}_{q,s} - \widetilde{\boldsymbol{v}}_{q,s}^{\mathrm{ort}}\right]_i \right\|_{L^2}^2 = \text{ min.} \tag{A.34}$$

Therefore, the functions of $\widetilde{\boldsymbol{b}}_{q,s}$ are highly similar to the functions in $\widetilde{\boldsymbol{v}}_{q,s}^{\mathrm{ort}}$.

The functions of $\widetilde{\boldsymbol{v}}_0^{\mathrm{low}} = \widetilde{\boldsymbol{b}}_0 \in S_V$ and $\widetilde{\boldsymbol{v}}_1^{\mathrm{low}} = \widetilde{\boldsymbol{b}}_1 \in S_{\overline{V}}$ in Table A.2 are already orthonormal, so no further orthogonalization is required. The functions in $\widetilde{\boldsymbol{v}}_2^{\mathrm{low}}$ belong to a single symmetry group, S_{EE}, so the corresponding vector of orthonormal functions $\widetilde{\boldsymbol{b}}_2$ is obtained via

$$\widetilde{\boldsymbol{b}}_2 = \widetilde{\boldsymbol{b}}_{2,1} = \widetilde{\boldsymbol{v}}_{2,1}^{\mathrm{ort}}\boldsymbol{B}_{2,1} = \widetilde{\boldsymbol{v}}_2^{\mathrm{low}}\boldsymbol{B}_2, \tag{A.35a}$$

$$\boldsymbol{B}_2 = \begin{bmatrix} a & b & b & b & b & -c \\ b & a & b & b & -c & b \\ b & b & a & -c & b & b \\ b & b & -c & a & b & b \\ b & -c & b & b & a & b \\ -c & b & b & b & b & a \end{bmatrix}, \text{ where } \begin{array}{l} a = 2.86155787682788, \\ b = 0.229204245589015, \\ c = 0.0367174724100048. \end{array} \tag{A.35b}$$

The functions in $\widetilde{\boldsymbol{v}}_3^{\mathrm{low}}$ belong to two symmetry groups, so the corresponding vector of orthonormal functions $\widetilde{\boldsymbol{b}}_{3,1} \in S_{\overline{EE}}$ and $\widetilde{\boldsymbol{b}}_{3,2} \in S_{NF}$ are obtained via

$$\widetilde{\boldsymbol{b}}_3 = \left[\widetilde{\boldsymbol{b}}_{3,1}, \widetilde{\boldsymbol{b}}_{3,2}\right] = \widetilde{\boldsymbol{v}}_3^{\mathrm{low}} \begin{bmatrix} \boldsymbol{B}_{3,1} & \boldsymbol{A}_{3,2}\boldsymbol{B}_{3,2} \\ \boldsymbol{0} & \boldsymbol{B}_{3,2} \end{bmatrix} = \widetilde{\boldsymbol{v}}_3^{\mathrm{low}}\boldsymbol{B}_3, \tag{A.36a}$$

$$\boldsymbol{B}_3 = \begin{bmatrix} a & b & b & -b & -b & 0 & 0 & 0 & -d & d \\ b & a & b & b & 0 & -b & 0 & -d & 0 & d \\ b & b & a & 0 & b & b & -d & 0 & 0 & d \\ -b & b & 0 & a & b & -b & 0 & -d & d & 0 \\ -b & 0 & b & b & a & b & -d & 0 & d & 0 \\ 0 & -b & b & -b & b & a & -d & d & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & c & e & e & e \\ 0 & 0 & 0 & 0 & 0 & 0 & e & c & e & e \\ 0 & 0 & 0 & 0 & 0 & 0 & e & e & c & e \\ 0 & 0 & 0 & 0 & 0 & 0 & e & e & e & c \end{bmatrix}, \quad \begin{array}{l} a = 3.99973289167987, \\ b = 0.0982693727007718, \\ c = 2.15382222237972, \\ d = 0.132170133159420, \\ e = 0.127213513935276. \end{array} \tag{A.36b}$$

These results are also summarized in Table A.3. In the higher-order case, all functions and matrices are obtained by repeating the described orthogonalization steps.

Table A.3: Symmetric, hierarchical, pairwise orthonormal basis functions.

	q	$\widetilde{\boldsymbol{b}}_q$
S_V	0	$\widetilde{\boldsymbol{v}}_0^{\text{low}}$
$S_{\overline{V}}$	1	$\widetilde{\boldsymbol{v}}_1^{\text{low}}$
S_{EE}	2	$\widetilde{\boldsymbol{v}}_2^{\text{low}}\boldsymbol{B}_2$
$S_{\overline{EE}}$, S_{NF}	3	$\widetilde{\boldsymbol{v}}_3^{\text{low}}\boldsymbol{B}_3$

Appendix B

Construction of $\boldsymbol{H}$(div) Bases for Tetrahedra

This section presents the construction of the proposed $\boldsymbol{H}$(div)-conforming bases. In the first subsection, only the higher-order volume-associated functions that are orthogonal only with respect to the interpolation operator (Table 5.2) are discussed. Then, proofs are provided for the verification of the orthogonality properties (5.6a), (5.6b), (5.7a), and (5.7b). Moreover, a symmetric orthogonalization process is presented over a given fixed domain. This method is used to achieve the desired orthogonality properties (5.10) and (5.14) for the range space of the higher $p > 1$ order functions in $\nabla \cdot \widetilde{\mathcal{F}}_p$ and $\nabla \times \widetilde{\mathcal{A}}_p$, respectively. As a result, separate matrices are obtained for functions of Table 5.2 and Table 2.2, which describe their appropriate linear combinations.

B.1 Construction of Range Space Functions

The first order $\boldsymbol{H}(\mathrm{div})$-conforming basis functions are the so-called Raviart-Thomas functions $\boldsymbol{f}_1^{ijk} \in \mathcal{D}_1$ that are associated to each triangular face of an element with the node indices ijk. They have zero normal components on all boundaries except the associated face. In the case of straight-sided elements, their general form reads [M$^+$03, Section 5.4] $\boldsymbol{f}_1^{ijk} = \alpha^{ijk}\boldsymbol{r} + \boldsymbol{\beta}^{ijk}$, where α^{ijk} is a constant scalar, $\boldsymbol{r}$ is the coordinate vector, and $\boldsymbol{\beta}^{ijk}$ is a constant vector. Hence, it is sufficient to construct the basis functions $\boldsymbol{f}_p \in \mathcal{D}_p$ of order p as

$$\boldsymbol{f}_p = \sum_{ijk} f_{p-1}^{ijk} \left(\alpha^{ijk}\boldsymbol{r} + \boldsymbol{\beta}^{ijk}\right), \qquad f_{p-1}^{ijk} \in \mathbb{P}_{p-1}(\Omega_e), \tag{B.1}$$

since there is no contribution to the highest-order null space. In order to obtain basis functions with the appropriate mappings for curvilinear elements, it is reasonable to rewrite this expression in terms of the barycentric coordinates φ_i and work on the reference element $\widehat{\Omega}_e$. Therefore, the general expression for the basis functions [ZC06, Section 2.5.3] reads

$$\widehat{\boldsymbol{f}}_1^{ijk} = \varphi_i \widehat{\nabla}\varphi_j \times \widehat{\nabla}\varphi_k + \varphi_j \widehat{\nabla}\varphi_k \times \widehat{\nabla}\varphi_i + \varphi_k \widehat{\nabla}\varphi_i \times \widehat{\nabla}\varphi_j, \tag{B.2a}$$

$$\widehat{\boldsymbol{f}}_p = \sum_{ijk} \widehat{f}_{p-1}^{ijk} \widehat{\boldsymbol{f}}_1^{ijk}, \qquad \widehat{f}_{p-1}^{ijk} \in \mathbb{P}_{p-1}(\widehat{\Omega}_e). \tag{B.2b}$$

Since the higher-order range space functions are all volume-associated and the finite element moments are independent of the domain, the orthogonality with respect to the interpolation operator (5.1) is satisfied by the following conditions:

$$\widehat{\boldsymbol{f}}_p \cdot \widehat{\boldsymbol{n}} = 0, \qquad p > 1, \tag{B.3a}$$

$$\int_{\widehat{V}} \widehat{\boldsymbol{f}}_p \cdot \widehat{\boldsymbol{q}} \,\mathrm{d}\widehat{\boldsymbol{r}} = 0, \qquad \forall\, \widehat{\boldsymbol{q}} \in (\mathbb{P}_{p-3}(\widehat{V}))^3. \tag{B.3b}$$

The first requirement (B.3a) is easy to satisfy since any $\varphi_l \widehat{\boldsymbol{f}}_1^{ijk}$ function has a zero normal component on the faces. Hence, it can be multiplied by an arbitrary linear combination of polynomials $f_{p-2}^{ijk,a} \in \mathbb{P}_{p-2}(\widehat{\Omega}_e)$ of order $p-2$ and corresponding constant coefficients $c_p^{ijk,a}$. Here a is an index for the different polynomials that span $\mathbb{P}_{p-2}$. These linear combinations can be formulated as a matrix-vector product of the coefficients $c_p^{ijk,a}$, ordered into a column vector $\boldsymbol{c}_p$, and a matrix $\boldsymbol{F}_p$ of linearly independent vector functions $\widehat{f}_{p-2}^{ijk,a} \varphi_l \widehat{\boldsymbol{f}}_1^{ijk}$, i.e.,

$$\widehat{\boldsymbol{f}}_p = \sum_{ijk} \sum_a \widehat{f}_{p-2}^{ijk,a} \varphi_l \widehat{\boldsymbol{f}}_1^{ijk} c^{ijk,a} = \boldsymbol{F}_p \boldsymbol{c}_p, \qquad \mathrm{span}\{f_{p-2}^{ijk,a}\} = \mathbb{P}_{p-2}(\widehat{\Omega}_e). \tag{B.4}$$

The second property (B.3b) requires an orthogonalization with respect to the test functions. Here, we considered the test functions to be in the incomplete order space

$\widehat{\boldsymbol{q}} \in \widehat{\mathcal{W}}_{p-2} = R_{p-2}$ of Nédélec (2.109a), since it includes the desired polynomial space $\mathbb{P}_{p-3} \subset \widehat{\mathcal{W}}_{p-2}$. The advantage is that this larger space corresponds to the volume-associated moments of the complete-order Nédélec space [Néd86]. Hence, the same orthogonality property $\pi_p^d \boldsymbol{f}_q = 0, q > p$ is obtained for both the complete [Néd86] and incomplete [Néd80] order interpolation operators of Nédélec. This way, all the available polynomials are utilized to obtain the orthogonality property with respect to the (complete-order) interpolation operator. Since the $\boldsymbol{H}$(curl)-conforming approximation space on the reference tetrahedra is $\widehat{\mathcal{W}}_p$, it is possible to express an arbitrary test function $\widehat{\boldsymbol{q}}$ in terms of basis functions $\widehat{\boldsymbol{w}}_{p-2}^b \in \widehat{\mathcal{W}}_{p-2}$, i.e.,

$$\widehat{\boldsymbol{q}} = \sum_b \alpha_{p-2}^b \widehat{\boldsymbol{w}}_{p-2}^b. \tag{B.5}$$

Here, b is an index between the basis functions and α_{p-2}^b is the corresponding constant coefficient. By ordering the basis functions $\widehat{\boldsymbol{w}}_{p-2}^b$ into a matrix $\boldsymbol{W}_{p-2}$ and the corresponding coefficients α_{p-2}^b into a vector $\boldsymbol{\alpha}_{p-2}$, any test function can be expressed as a matrix-vector product

$$\widehat{\boldsymbol{q}} = \sum_b \alpha_{p-2}^b \widehat{\boldsymbol{w}}_{p-2}^b = \boldsymbol{W}_{p-2}\, \boldsymbol{\alpha}_{p-2}. \tag{B.6}$$

Then, the construction of the $\boldsymbol{H}$(div) basis functions reduces to the calculation of an appropriate set of linear independent $\boldsymbol{c}_p$ coefficient vectors which satisfies (B.3b) for arbitrary $\boldsymbol{\alpha}_{p-2}$ vectors. Let us order all coefficient vectors into a matrix $\boldsymbol{C}_p = [\boldsymbol{c}_p^1, ..., \boldsymbol{c}_p^i, ...]$. Then, the criterion (B.3b) requires this coefficient matrix to be in the null space of the following product

$$\text{Null}\left(\int_{\widehat{\Omega}_e} \boldsymbol{W}_{p-2}^T \boldsymbol{F}_p \, \mathrm{d}v \right) = \boldsymbol{C}_p. \tag{B.7}$$

Note that there is no unique solution since any linear combination of a specific $\boldsymbol{C}_p$ satisfies the same condition. Hence, a manually selected appropriate $\boldsymbol{C}_p$ is chosen, which provides tetrahedron symmetric sets of basis functions; see Table 5.2.

B.2 Orthogonality on Straight-Sided Elements

This subsection provides proofs for the orthogonality of the basis functions over an arbitrary straight-sided tetrahedron. In Section B.1, the orthogonality of the range space functions is discussed with respect to the interpolation operator. However, the null space functions $\nabla \times \boldsymbol{w}_p \in \nabla \times \mathcal{W}_p$ have not been considered. The reason for this is that the commutation of the interpolation operators $\pi_p^d \nabla \times \boldsymbol{w}_q = \nabla \times \pi_p^w \boldsymbol{w}_q$ leads to the case of $\boldsymbol{H}(\mathrm{curl})$ interpolation, where the orthogonality is already satisfied (2.120). Then, the properties of (5.6b) and (5.7b) are direct consequences of the orthogonality with respect to the interpolation operator. Therefore, proofs are only provided for (5.6a) and (5.7a).

First, the orthogonality of the stiffness matrix is discussed. The general form of the integrals is

$$\int_{\Omega_e} \nabla \cdot \boldsymbol{d}_p \nabla \cdot \boldsymbol{d}_q \,\mathrm{d}\boldsymbol{r}, \tag{B.8}$$

where $\boldsymbol{d}_p \in \widetilde{\mathcal{D}}_p$ and $\boldsymbol{d}_q \in \widetilde{\mathcal{D}}_q$ are the hierarchical basis functions. Here, the integrand can be transformed via the vector identity

$$\nabla \cdot \boldsymbol{d}_p \nabla \cdot \boldsymbol{d}_q = \nabla \cdot (\boldsymbol{d}_q \nabla \cdot \boldsymbol{d}_p) - \boldsymbol{d}_q \cdot \nabla \nabla \cdot \boldsymbol{d}_p. \tag{B.9}$$

Moreover, by applying the divergence theorem, the integral becomes

$$\int_{\Omega_e} \nabla \cdot \boldsymbol{d}_p \nabla \cdot \boldsymbol{d}_q \,\mathrm{d}\boldsymbol{r} = \oint_F \boldsymbol{d}_q \cdot \boldsymbol{n} \,(\nabla \cdot \boldsymbol{d}_p)\, \mathrm{d}\boldsymbol{r} - \int_{\Omega_e} \boldsymbol{d}_q \cdot \nabla \nabla \cdot \boldsymbol{d}_p \,\mathrm{d}\boldsymbol{r}. \tag{B.10}$$

Here, $\boldsymbol{n}$ denotes the normal vector of the different planar faces F for an arbitrary straight-sided tetrahedron. The first surface integral vanishes since all the higher-order $p > 1$ range space functions have a zero normal component on the faces. Moreover, $\nabla \nabla \cdot \boldsymbol{d}_p \in (\mathbb{P}_{p-2})^3$ is a polynomial of order $p-2$. Hence, this term also vanishes for the cases $|p-q| > 0$ due to (B.3b). The consequence is that each hierarchical basis function is orthogonal with respect to the lower-order functions in the stiffness matrix (5.6a). In the case of mixed inner products (5.7a), the same proof holds with the replacement of the $\nabla \cdot \boldsymbol{d}_p$ by a polynomial of order $f_p \in \mathbb{P}_p$. In that case, the resulting orthogonality reads

$$\int_{\Omega_e} f_p \nabla \cdot \boldsymbol{d}_q \,\mathrm{d}\boldsymbol{r} = \oint_F \boldsymbol{d}_q \cdot \boldsymbol{n} f_p \,\mathrm{d}\boldsymbol{r} - \int_{\Omega_e} \boldsymbol{d}_q \cdot \nabla f_p \,\mathrm{d}\boldsymbol{r} = 0, \qquad q > p+1. \tag{B.11}$$

B.3 Symmetric Orthogonalization

The presented hierarchical bases in Table 2.2 and Table 5.2 fulfill the orthogonality requirements with respect to the interpolation operators. Moreover, each function is either associated with the face or the volume, as well as belonging to the range or null space of the divergence operator. However, the bases are not unique since any linear combination of functions $\nabla \cdot \boldsymbol{f}_p^i \in \nabla \cdot \widetilde{\mathcal{F}}_p^v$ or $\nabla \times \boldsymbol{w}_p^i \in \nabla \times \widetilde{\mathcal{A}}_p^f$ or $\nabla \times \boldsymbol{w}_p^i \in \nabla \times \widetilde{\mathcal{A}}_p^v$ in the corresponding subspaces fulfills the same requirements. This additional freedom can be used to orthogonalize the functions within the same subspace for a given shape. My choice is the regular equilateral tetrahedron Ω_{reg} with unit length edges. Conceptually, this is the same orthogonalization procedure as in Section A.4, but for vector-valued functions and over a regular element. Therefore, the initial function subsets with symmetries are transformed in a symmetric way, unlike with the Gram-Schmidt process.

Let $\boldsymbol{d}_p^i$ denote a function which needs to be orthogonalized, and i is the index of the function in the corresponding subspace. Every $\boldsymbol{d}_p^i$ function is linearly independent of the other. Hence, for all nonsingular coefficient matrices $\boldsymbol{C}_p^{(\cdot)}$, the linear combination of

$$[\boldsymbol{d}_{p,\text{ort}}^1, \boldsymbol{d}_{p,\text{ort}}^2, ...] = [\boldsymbol{d}_p^1, \boldsymbol{d}_p^2, ...]\boldsymbol{C}_p^{(\cdot)}, \tag{B.12}$$

also provides an appropriate basis. Here, $(\cdot)$ denotes Fv for functions in $\nabla \cdot \widetilde{\mathcal{F}}_p^v$, Af for functions in $\nabla \times \widetilde{\mathcal{A}}_p^f$, and Av in case of $\nabla \times \widetilde{\mathcal{A}}_p^v$.

The goal is to construct these coefficient matrices in such a way that the symmetry of the original $\boldsymbol{d}_p^i$ functions is preserved. Let the matrix $\boldsymbol{S}_p$ contain the first n functions as $\boldsymbol{S}_p = [\boldsymbol{d}_p^1, ..., \boldsymbol{d}_p^n]$, which are symmetric with respect to each other on a tetrahedral domain and let $\boldsymbol{R}_p$ contain the remaining functions as $\boldsymbol{R}_p = [\boldsymbol{d}_p^{n+1}, \boldsymbol{d}_p^{n+2}, ...]$. In order to orthogonalize $\boldsymbol{R}_p$ with respect to $\boldsymbol{S}_p$, one may use their linear combination as $\widetilde{\boldsymbol{R}}_p = \boldsymbol{R}_p + \boldsymbol{S}_p\boldsymbol{A}_p$ with the constant coefficient matrix $\boldsymbol{A}_p$. The orthogonality over a regular tetrahedron Ω_{reg} reads

$$\int_{\Omega_{\text{reg}}} \boldsymbol{S}_p^T \tilde{\boldsymbol{R}}_p \, \mathrm{d}\boldsymbol{r} = \int_{\Omega_{\text{reg}}} \boldsymbol{S}_p^T \left(\boldsymbol{R}_p + \boldsymbol{S}_p\boldsymbol{A}_p\right) \mathrm{d}\boldsymbol{r} = \boldsymbol{0}. \tag{B.13}$$

After evaluating the integrals, the linear equation system can be solved for the coefficient matrix $\boldsymbol{A}_p$, which reads

$$\boldsymbol{A}_p = -\left[\int_{\Omega_{\text{reg}}} \boldsymbol{S}_p^T \boldsymbol{S}_p \, \mathrm{d}\boldsymbol{r}\right]^{-1} \int_{\Omega_{\text{reg}}} \boldsymbol{S}_p^T \boldsymbol{R}_p \, \mathrm{d}\boldsymbol{r}. \tag{B.14}$$

Since $\boldsymbol{S}_p$ is orthogonal with respect to the remaining functions in $\widetilde{\boldsymbol{R}}_p$, any linear combination of these functions is orthogonal as well. This relation can be written as $\boldsymbol{S}_p\boldsymbol{B}_p$, where the constant matrix $\boldsymbol{B}_p$ describes such an arbitrary linear combination. Moreover, the orthogonality of the functions in $\boldsymbol{S}_p\boldsymbol{B}_p$ with respect to each other requires

$$\int_{\Omega_{\mathrm{reg}}} (\boldsymbol{S}_p\boldsymbol{B}_p)^T \boldsymbol{S}_p\boldsymbol{B}_p \,\mathrm{d}\boldsymbol{r} = \boldsymbol{B}_p^T \int_{\Omega_{\mathrm{reg}}} \boldsymbol{S}_p^T\boldsymbol{S}_p \,\mathrm{d}\boldsymbol{r}\ \boldsymbol{B}_p = \boldsymbol{D}_p, \tag{B.15}$$

where $\boldsymbol{D}_p$ is a diagonal matrix. Since the second integral of (B.15) results in a symmetric positive definite matrix, it is orthogonally diagonalizable as

$$\int_{\Omega_{\mathrm{reg}}} \boldsymbol{S}_p^T\boldsymbol{S}_p \,\mathrm{d}\boldsymbol{r}\ = \boldsymbol{V}_p\boldsymbol{\Lambda}_p\boldsymbol{V}_p^T. \tag{B.16}$$

Here, the matrix $\boldsymbol{V}_p$ contains the eigenvectors as column vectors and $\boldsymbol{\Lambda}_p$ is a diagonal matrix containing the eigenvalues. One can show by substitution that with the symmetric choice of

$$\boldsymbol{B}_p = \boldsymbol{V}_p\boldsymbol{\Lambda}_p^{-\frac{1}{2}}\boldsymbol{V}_p^T\boldsymbol{D}_p^{\frac{1}{2}}, \tag{B.17}$$

the orthogonality (B.15) is satisfied. Hence, the first n orthogonalized functions are calculated as

$$[\boldsymbol{d}_{p,\mathrm{ort}}^1, ..., \boldsymbol{d}_{p,\mathrm{ort}}^n] = \boldsymbol{S}_p\boldsymbol{B}_p. \tag{B.18}$$

If the matrix of the remaining functions $\widetilde{\boldsymbol{R}}_p$ is an empty matrix, then $\boldsymbol{C}_p^{(\cdot)} = \boldsymbol{B}_p$. Otherwise, the same process needs to be repeated until reaching that point. Note that in the orthogonalization procedure for basis functions of order $p = 4$, the matrix of the remaining functions $\widetilde{\boldsymbol{R}}_p$ consists of a single set of symmetric functions at most. Thus, the orthogonalization procedure needs to be repeated only once. In the second round, $\boldsymbol{B}_{p,2}$ denotes the unknown matrix which determines the appropriate linear combination of the remaining functions as $\widetilde{\boldsymbol{R}}_p\boldsymbol{B}_{p,2}$. The new functions of $\widetilde{\boldsymbol{R}}_p\boldsymbol{B}_{p,2}$ are required to be piecewise orthogonal. Therefore, the product of

$$\int_{\Omega_{\mathrm{reg}}} \left(\widetilde{\boldsymbol{R}}_p\boldsymbol{B}_{p,2}\right)^T \widetilde{\boldsymbol{R}}_p\boldsymbol{B}_{p,2} \,\mathrm{d}\boldsymbol{r} = \boldsymbol{B}_{p,2}^T \int_{\Omega_{\mathrm{reg}}} \widetilde{\boldsymbol{R}}_p^T\widetilde{\boldsymbol{R}}_p \,\mathrm{d}\boldsymbol{r}\ \boldsymbol{B}_{p,2} = \boldsymbol{D}_{p,2}, \tag{B.19}$$

should result in a diagonal matrix $\boldsymbol{D}_{p,2}$. Then, the orthogonal diagonalization is required with the matrix of eigenvectors $\boldsymbol{V}_{p,2}$ and the diagonal matrix of eigenvalues $\boldsymbol{\Lambda}_{p,2}$, i.e.,

$$\int_{\Omega_{\mathrm{reg}}} \widetilde{\boldsymbol{R}}_p^T\widetilde{\boldsymbol{R}}_p \,\mathrm{d}\boldsymbol{r}\ = \boldsymbol{V}_{p,2}\boldsymbol{\Lambda}_{p,2}\boldsymbol{V}_{p,2}^T. \tag{B.20}$$

Therefore, the calculation of $\boldsymbol{B}_{p,2}$ with a diagonal, normalization matrix $\boldsymbol{D}_{p,2}$ reads

$$\boldsymbol{B}_{p,2} = \boldsymbol{V}_{p,2}\boldsymbol{\Lambda}_{p,2}^{-\frac{1}{2}}\boldsymbol{V}_{p,2}^{T}\boldsymbol{D}_{p,2}^{\frac{1}{2}}, \tag{B.21}$$

and the full orthogonalization matrix becomes

$$\boldsymbol{C}_{p}^{(\cdot)} = \left[\begin{array}{c|c} \boldsymbol{B}_{p} & \boldsymbol{A}_{p}\boldsymbol{B}_{p,2} \\ \hline \boldsymbol{0} & \boldsymbol{B}_{p,2} \end{array}\right]. \tag{B.22}$$

Note that the orthogonality of functions $\nabla \cdot \boldsymbol{f}_{p}^{1}$ not just applies to a regular tetrahedron Ω_{reg} but also to arbitrary straight-sided tetrahedron since it is a scalar function.

Table B.1: Orthogonalization matrix (5.9) for functions in $\widetilde{\mathcal{F}}_{2}^{v}$ from Table 5.2.

$$\boldsymbol{C}_{2}^{Fv} = \begin{bmatrix} a & -b & b \\ -b & a & -b \\ b & -b & a \end{bmatrix}, \quad \begin{aligned} a &= 1.490711984999860 \\ b &= 0.372677996249965 \end{aligned}$$

Table B.2: Orthogonalization matrix (5.9) for functions in $\widetilde{\mathcal{F}}_{3}^{v}$ from Table 5.2.

$$\boldsymbol{C}_{3}^{Fv} = \begin{bmatrix} a & b & c & b & -b & b \\ b & a & b & c & b & -b \\ c & b & a & b & -b & b \\ b & c & b & a & b & -b \\ -b & b & -b & b & a & -c \\ b & -b & b & -b & -c & a \end{bmatrix}, \quad \begin{aligned} a &= 1.227406393014804 \\ b &= 0.088770422604629 \\ c &= 0.044190436394881 \end{aligned}$$

Table B.3: Orthogonalization matrix (5.9) for functions in $\widetilde{\mathcal{F}}_{4}^{v}$ from Table 5.2.

$$\boldsymbol{C}_{4}^{Fv} = \begin{bmatrix} a & -b & 0 & -b & -b & -b & c & c & 0 & 0 \\ -b & a & -b & 0 & b & -b & 0 & c & c & 0 \\ 0 & -b & a & -b & b & b & 0 & 0 & c & c \\ -b & 0 & -b & a & -b & b & c & 0 & 0 & c \\ -b & b & b & -b & a & 0 & c & 0 & -c & 0 \\ -b & -b & b & b & 0 & a & 0 & c & 0 & -c \\ 0 & 0 & 0 & 0 & 0 & 0 & d & e & -e & e \\ 0 & 0 & 0 & 0 & 0 & 0 & e & d & e & -e \\ 0 & 0 & 0 & 0 & 0 & 0 & -e & e & d & e \\ 0 & 0 & 0 & 0 & 0 & 0 & e & -e & e & d \end{bmatrix}, \quad \begin{aligned} a &= 1.533295567153832 \\ b &= 0.038428702322497 \\ c &= 2.568915451961508 \\ d &= 5.059024700653003 \\ e &= 0.996276007541980 \end{aligned}$$

Table B.4: Orthogonalization matrix (5.13a) for functions in $\nabla\times\widetilde{\mathcal{A}}_2^f$ from Table 2.2.

$$\boldsymbol{C}_2^{Af}=\begin{bmatrix} a & b\\ b & a\end{bmatrix},\quad \begin{array}{l}\text{a=0.544237547893580}\\ \text{b=0.145828011448779}\end{array}$$

Table B.5: Orthogonalization matrix (5.13a) for functions in $\nabla\times\widetilde{\mathcal{A}}_3^f$ from Table 2.2.

$$\boldsymbol{C}_3^{Af}=\begin{bmatrix} a & -b & -b\\ -b & a & -b\\ -b & -b & a\end{bmatrix},\quad \begin{array}{l}a=0.446430946524381\\ b=0.042073263934821\end{array}$$

Table B.6: Orthogonalization matrix (5.13b) for functions in $\nabla\times\widetilde{\mathcal{A}}_3^v$ from Table 2.2.

$$\boldsymbol{C}_3^{Av}=\begin{bmatrix} a & b & b\\ b & a & b\\ b & b & a\end{bmatrix},\quad \begin{array}{l}a=1.080123449734652\\ b=0.2700308624336632\end{array}$$

Table B.7: Orthogonalization matrix (5.13a) for functions in $\nabla\times\widetilde{\mathcal{A}}_4^f$ from Table 2.2.

$$\boldsymbol{C}_4^{Af}=\begin{bmatrix} a & b & b & 0\\ b & a & b & 0\\ b & b & a & 0\\ 0 & 0 & 0 & c\end{bmatrix},\quad \begin{array}{l}a=0.415182195396832\\ b=0.014959527169861\\ c=0.694685625986152\end{array}$$

Table B.8: Orthogonalization matrix (5.13b) for functions in $\nabla\times\widetilde{\mathcal{A}}_4^v$ from Table 2.2.

$$\boldsymbol{C}_4^{Av}=\begin{bmatrix} a & b & -g & -f & -e & -c & -g & d\\ b & a & d & -g & -c & -e & -f & -g\\ -g & d & a & b & -g & -f & -e & -c\\ -f & -g & b & a & d & -g & -c & -e\\ -e & -c & -g & d & a & b & -g & -f\\ -c & -e & -f & -g & b & a & d & -g\\ -g & -f & -e & -c & -g & d & a & b\\ d & -g & -c & -e & -f & -g & b & a\end{bmatrix},\quad \begin{array}{l}a=1.577430193626924\\ b=0.429924197921561\\ c=0.216403501134851\\ d=0.211046239810806\\ e=0.142510605811004\\ f=0.076210515325377\\ g=0.067417862242715\end{array}$$

List of Publications

Peer-Reviewed Journal Articles

[A1] L. L. Tóth, R. Dyczij-Edlinger, "Hierarchical universal matrices for sensitivity analysis by curvilinear finite elements," In: *Applied Computational Electromagnetics Society Journal* (ACES) 34 (2019), Feb., Nr. 2, S. 206–210

[A2] L. L. Tóth, A. A. Martin, R. Dyczij-Edlinger, "Hierarchical universal matrices for curvilinear tetrahedral H(curl) finite elements with inhomogeneous material properties," In: *IEEE Transactions on Antennas and Propagation*, Accepted (2023).

[A3] A. A. Martin, L. C. Garcia, L. L. Tóth, O. Floch, R. Dyczij-Edlinger, "A Rigorous verification process of the domain decomposition method in a finite element method code for electromagnetics," In: *IEEE Transactions on Antennas and Propagation*, Accepted (2023).

[A4] L. L. Tóth, R. Dyczij-Edlinger, "Structurally symmetric, hierarchical, pairwise orthonormal Legendre-type polynomials on tetrahedra," In: *Journal of Computational and Applied Mathematics*, In preparation.

[A5] L. L. Tóth, A. A. Martin, R. Dyczij-Edlinger, "A new set of hierarchical H(div)-conforming basis functions for tetrahedral meshes," In: *International Journal for Numerical Methods in Engineering*, In preparation.

[A6] L. L. Tóth, A. A. Martin, R. Dyczij-Edlinger, "Mixed-order mixed-space H(curl)-conforming elements for hexahedra," In: *International Journal for Numerical Methods in Engineering*, In preparation.

[A7] L. L. Tóth, A. A. Martin, R. Dyczij-Edlinger, "The iso-serendipity finite element method," In: *International Journal for Numerical Methods in Engineering*, In preparation.

Peer-Reviewed Conference Articles

[B1] L. L. Tóth, R. Dyczij-Edlinger, "Hierarchical universal matrices for sensitivity analysis by curvilinear finite elements," In: *2018 International Applied Computational Electromagnetics Society Symposium, IEEE Xplore*, (2018), S. 1–2

[B2] L. L. Tóth, A. A. Martin, R. Dyczij-Edlinger, "The mixed-order serendipity finite element for H(curl)-conforming hexahedra," In: *2021 Kleinheubach Conference, IEEE Xplore*, (2021), S. 1–4

[B3] L. D. Schuck, L. L. Tóth, R. Dyczij-Edlinger, "Time domain state space solver for microwave systems based on Laguerre polynomials," In: *The 10th Vienna International Conference on Mathematical Modelling, IFAC-PapersOnLine* 55 (2022), Nr. 20, S. 25–30

[B4] L. L. Tóth, A. A. Martin, R. Dyczij-Edlinger, "Convergence study of H(curl) serendipity basis functions for hexahedral finite-elements," In: *24th International Microwave and Radar Conference, IEEE Xplore*, (2022), S. 1–3

[B5] L. L. Tóth, L. D. Schuck, R. Dyczij-Edlinger, "Interpolation-operator orthogonal, hierarchical H1-conforming basis functions for tetrahedral finite elements," In: *24th International Conference on Electromagnetics in Advanced Applications*, (2023), Accepted.

Conference Presentations

[C1] L. L. Tóth, R. Dyczij-Edlinger, "Universal finite-element matrices for curvilinear tetrahedra with continuously varying material properties,"
Pres. KH2017-Di-B1-01, *Kleinheubacher Tagung 2017*, Miltenberg, Germany: September 25–27, 2017

[C2] L. L. Tóth, R. Dyczij-Edlinger, "Hierarchical universal matrices for sensitivity analysis by curvilinear finite elements," *2018 International Applied Computational Electromagnetics Society (ACES) Symposium*, Denver, Colorado USA: March 24–29, 2018

[C3] L. L. Tóth, R. Dyczij-Edlinger, "Hierarchic divergence-conforming basis functions for tetrahedral finite-elements," Pres. KH2018-Di-B1-05, *Kleinheubacher Tagung 2018*, Miltenberg, Germany: September 24–26, 2018

[C4] L. L. Tóth, A. A. Martin, R. Dyczij-Edlinger, "Hierarchical H(div) basis functions and universal matrices for curvilinear finite elements," *Kinetics and BEM on the Saar 2019*, Saarbrücken, Germany: June 4–7, 2019

[C5] A. A. Martin, L. L. Tóth, R. Dyczij-Edlinger, "H(curl)-conforming hierarchical basis functions on prisms and hexahedra," Pres. KH2019-Di-B1-01, *Kleinheubacher Tagung 2019*, Miltenberg, Germany: September 23–25, 2019

[C6] L. L. Tóth, R. Dyczij-Edlinger, "The mixed-order serendipity finite element for H(curl)-conforming hexahedra," Pres. KH2021-Mi-B1-07, *Kleinheubacher Tagung 2021*, Miltenberg, Germany: September 28–30, 2021

[C7] L. L. Tóth, R. Dyczij-Edlinger, "Fast calculation of curvilinear finite element matrices via hierarchical universal matrices," *URSI-B Workshop "Felder und Wellen"*, TUM Akademiezentrum Raitenhaslach, Germany: June 29–July 1, 2022

[C8] L. L. Tóth, A. A. Martin, R. Dyczij-Edlinger, "Convergence study of H(curl) serendipity basis functions for hexahedral finite-elements," *24th International Microwave and Radar Conference* (MIKON-2022), Gdansk, Poland: September 11–15, 2022

List of Figures

List of Tables

Bibliography

[AA11] Arnold, Douglas N. ; Awanou, Gerard: The serendipity family of finite elements. In: *Foundations of Computational Mathematics* 11 (2011), Nr. 3, S. 337–344

[AA14] Arnold, Douglas ; Awanou, Gerard: Finite element differential forms on cubical meshes. In: *Mathematics of Computation* 83 (2014), Nr. 288, S. 1551–1570

[AAG20] Aktaş, Rabia ; Area, Iván ; Güldoğan, Esra: A new family of orthogonal polynomials in three variables. In: *Journal of Inequalities and Applications* 2020 (2020), Nr. 1, S. 1–27

[ABB15] Arnold, Douglas N. ; Boffi, Daniele ; Bonizzoni, Francesca: Finite element differential forms on curvilinear cubic meshes and their approximation properties. In: *Numerische Mathematik* 129 (2015), Nr. 1, S. 1–20

[ABDG98] Amrouche, Cherif ; Bernardi, Christine ; Dauge, Monique ; Girault, Vivette: Vector potentials in three-dimensional non-smooth domains. In: *Mathematical Methods in the Applied Sciences* 21 (1998), Nr. 9, S. 823–864

[ABF02] Arnold, Douglas ; Boffi, Daniele ; Falk, Richard: Approximation by quadrilateral finite elements. In: *Mathematics of Computation* 71 (2002), Nr. 239, S. 909–922

[AC03] Ainsworth, Mark ; Coyle, Joe: Hierarchic finite element bases on unstructured tetrahedral meshes. In: *International Journal for Numerical Methods in Engineering* 58 (2003), Nr. 14, S. 2103–2130

[AEG80] Aiken, John G. ; Erdos, John A. ; Goldstein, Jerome A.: On löwdin orthogonalization. In: *International Journal of Quantum Chemistry* 18 (1980), Nr. 4, S. 1101–1108

[AFW97] Arnold, Douglas ; Falk, Richard ; Winther, Ragnar: Preconditioning in H(div) and applications. In: *Mathematics of Computation* 66 (1997), Nr. 219, S. 957–984

[AKIH08] Azuma, Y. ; Kaneko, Takuya ; Ishikawa, Kou ; Horii, Yasushi: Ultra-wideband (UWB) bandpass filter composed of periodically-installed metallic posts in a rectangular waveguide. In: *2008 Asia-Pacific Microwave Conference* (2008), S. 1–4

[ALM+09] Alnæs, Martin S. ; Logg, Anders ; Mardal, Kent-Andre ; Skavhaug, Ola ; Langtangen, Hans P.: Unified framework for finite element assembly. In: *International Journal of Computational Science and Engineering* 4 (2009), Nr. 4, S. 231–244

[AMGC19] Amor-Martin, Adrian ; Garcia-Castillo, Luis E.: Construction of higher-order curl-conforming finite elements and its assembly. In: *International Journal of RF and Microwave Computer-Aided Engineering* 29 (2019), Nr. 8, S. e21753

[AP02] Ainsworth, M. ; Pinchedez, K.: hp-approximation theory for BDFM and RT finite elements on quadrilaterals. In: *SIAM Journal on Numerical Analysis* 40 (2002), Nr. 6, S. 2047–2068

[Ape98] Apel, Th.: Anisotropic interpolation error estimates for isoparametric quadrilateral finite elements. In: *Computing* 60 (1998), Nr. 2, S. 157–174

[ATMM07] Aa, Nico van d. ; Ter Morsche, H ; Mattheij, R: Computation of eigenvalue and eigenvector derivatives for a general complex-valued eigensystem. In: *The Electronic Journal of Linear Algebra* 16 (2007)

[Bal18] Baltes, Rolf-Björn: *Verfahren der Ordnungsreduktion zur Charakterisierung Passiver Mikrowellenstrukturen im Frequenz- und Zeitbereich auf Basis der Finite-Elemente-Methode*. Saarländische Universitäts- und Landesbibliothek, 2018

[Ban96] Bank, Randolph E.: Hierarchical bases and the finite element method. In: *Acta Numerica* 5 (1996), S. 1–43

[BBF13] Boffi, Daniele ; Brezzi, Franco ; Fortin, Michel: *Mixed Finite Element Methods and Applications*. Bd. 44. Springer, 2013

[BD13a] Bergot, Morgane ; Duruflé, Marc: Approximation of H(div) with high-order optimal finite elements for pyramids, prisms and hexahedra. In: *Communications in Computational Physics* 14 (2013), Nr. 5, S. 1372–1414

[BD13b] Bergot, Morgane ; Duruflé, Marc: High-order optimal edge elements for pyramids, prisms and hexahedra. In: *Journal of Computational Physics* 232 (2013), Nr. 1, S. 189–213

[BDO21] Bachmayr, Markus ; Dusson, Geneviève ; Ortner, Christoph: Polynomial approximation of symmetric functions. In: *arXiv preprint arXiv:2109.14771* (2021)

[BF12] Brezzi, Franco ; Fortin, Michel: *Mixed and Hybrid Finite Element Methods*. Bd. 15. Springer Science & Business Media, 2012

[Bor13] Borzi, Giuseppe: Universal matrices for edge elements. In: *IEEE Transactions on Microwave Theory and Techniques* 61 (2013), Nr. 6, S. 2275–2281

[Bos98] Bossavit, Alain: *Computational Electromagnetism: Variational Formulations, Complementarity, Edge Elements*. Academic Press, 1998

[Bot06] Botha, Matthys M.: Solving the volume integral equations of electromagnetic scattering. In: *Journal of Computational Physics* 218 (2006), Oktober, Nr. 1, S. 141–158

[Bot07] Botha, Matthys M.: Fully hierarchical divergence-conforming basis functions on tetrahedral cells, with applications. In: *International Journal for Numerical Methods in Engineering* 71 (2007), Nr. 2, S. 127–148

[BPZ10] Beuchler, Sven ; Pillwein, Veronika ; Zaglmayr, Sabine: Sparsity optimized high order finite element functions for H(div) on simplices. (2010)

[BPZ12] Beuchler, Sven ; Pillwein, Veronika ; Zaglmayr, Sabine: Sparsity optimized high order finite element functions for H(div) on simplices. In: *Numerische Mathematik* 122 (2012), Oct, Nr. 2, S. 197–225

[BS93] Bank, R. ; Smith, R.: A posteriori error estimates based on hierarchical bases. In: *SIAM Journal on Numerical Analysis* 30 (1993), Nr. 4, S. 921–935

[BS07] Brenner, Susanne ; Scott, Ridgway: *The Mathematical Theory of Finite Element Methods*. Bd. 15. Springer Science & Business Media, 2007

[BSV10] Buffa, Annalisa ; Sangalli, Giancarlo ; Vázquez, Rafael: Isogeometric analysis in electromagnetics: B-splines approximation. In: *Computer Methods in Applied Mechanics and Engineering* 199 (2010), Nr. 17-20, S. 1143–1152

[BWPL12] Beig, D. Ansari O. ; Wang, J. ; Peng, Z. ; Lee, J. F.: A universal array approach for finite elements with continuously inhomogeneous material properties and curved boundaries. In: *IEEE Transactions on Antennas and Propagation* 60 (2012), Oct, Nr. 10, S. 4745–4756

[CDF+16] Castro, Douglas A. ; Devloo, Philippe R. ; Farias, Agnaldo M. ; Gomes, Sônia M. ; Siqueira, Denise de ; Durán, Omar: Three dimensional hierarchical mixed finite element approximations with enhanced primal variable accuracy. In: *Computer Methods in Applied Mechanics and Engineering* 306 (2016), S. 479 – 502

[Cia02] Ciarlet, P.G.: *The Finite Element Method for Elliptic Problems*. Society for Industrial and Applied Mathematics, 2002 (Classics in Applied Mathematics)

[CMPR98] Coccioli, R. ; Morini, A. ; Pelosi, G. ; Rozzi, T.: Design of tolerance-corrected filters employing half-cylinder posts. In: *IEEE Transactions on Microwave Theory and Techniques* 46 (1998), Jan, Nr. 1, S. 116–118

[Coo97] Cools, Ronald: Constructing cubature formulae: the science behind the art. In: *Acta numerica* 6 (1997), S. 1–54

[Coo02] Cools, Ronald: Advances in multidimensional integration. In: *Journal of Computational and Applied Mathematics* 149 (2002), Nr. 1, S. 1 – 12. – Scientific and Engineering Computations for the 21st Century - Methodologies and Applications Proceedings of the 15th Toyota Conference

[Cox61] Coxeter, Harold Scott M.: *Introduction to Geometry*. New York, London, 1961

[CQ82] Canuto, Claudio ; Quarteroni, Alfio: Approximation results for orthogonal polynomials in Sobolev spaces. In: *Mathematics of Computation* 38 (1982), Nr. 157, S. 67–86

[CXX+22] Cao, Juan ; Xiao, Yi ; Xiao, Yanyang ; Chen, Zhonggui ; Xue, Fei ; Wei, Xiaodong ; Zhang, Yongjie J.: Quadratic serendipity element shape functions on general planar polygons. In: *Computer Methods in Applied Mechanics and Engineering* 392 (2022), S. 114703

[CZZ+15] Cai, Q. ; Zhao, Y. ; Zheng, Y. ; Jia, M. ; Zhao, Z. ; Nie, Z.: Volume integral equation with higher order hierarchical basis functions for analysis of dielectric electromagnetic scattering. In: *IEEE Transactions on Antennas and Propagation* 63 (2015), Nov, Nr. 11, S. 4964–4975

[Dai89] Dailey, R L.: Eigenvector derivatives with repeated eigenvalues. In: *AIAA journal* 27 (1989), Nr. 4, S. 486–491

[Dav75] Davis, Philip J.: *Interpolation and Approximation*. Courier Corporation, 1975

[DG19] Duczek, S ; Gravenkamp, H: Critical assessment of different mass lumping schemes for higher order serendipity finite elements. In: *Computer Methods in Applied Mechanics and Engineering* 350 (2019), S. 836–897

[DVBMR17] Da Veiga, Lourenco B. ; Brezzi, Franco ; Marini, Luisa D. ; Russo, Alessandro: Serendipity face and edge VEM spaces. In: *Rendiconti Lincei* 28 (2017), Nr. 1, S. 143–180

[DX14] Dunkl, Charles F. ; Xu, Yuan: *Orthogonal Polynomials of Several Variables*. Cambridge University Press, 2014 (155)

[FGM11] Falk, Richard S. ; Gatto, Paolo ; Monk, Peter: Hexahedral H(div) and H(curl) finite elements. In: *ESAIM: Mathematical Modelling and Numerical Analysis* 45 (2011), Nr. 1, S. 115–143

[FGS03] Farouki, Rida T. ; Goodman, Tim N. ; Sauer, Thomas: Construction of orthogonal bases for polynomials in Bernstein form on triangular and simplex domains. In: *Computer Aided Geometric Design* 20 (2003), Nr. 4, S. 209–230

[Gev88] Geveci, Tunc: On the application of mixed finite element methods to the wave equations. In: *ESAIM: Mathematical Modelling and Numerical Analysis* 22 (1988), Nr. 2, S. 243–250

[GK19] Gillette, Andrew ; Kloefkorn, Tyler: Trimmed serendipity finite element differential forms. In: *Mathematics of Computation* 88 (2019), Nr. 316, S. 583–606

[GKS19] Gillette, Andrew ; Kloefkorn, Tyler ; Sanders, Victoria: Computational serendipity and tensor product finite element differential forms. In: *The SMAI Journal of Computational Mathematics* 5 (2019), S. 1–21

[GP12] Graglia, R. D. ; Peterson, A. F.: Hierarchical divergence-conforming Nédélec elements for volumetric cells. In: *IEEE Transactions on Antennas and Propagation* 60 (2012), Nov, Nr. 11, S. 5215–5227

[GS11] Griffiths, Robert C. ; Spano, Dario: Multivariate Jacobi and Laguerre polynomials, infinite-dimensional extensions, and their probabilistic connections with multivariate Hahn and Meixner polynomials. In: *Bernoulli* 17 (2011), Nr. 3, S. 1095–1125

[Har01] Harrington, R.F.: *Time-Harmonic Electromagnetic Fields*. Wiley, 2001 (IEEE Press Series on Electromagnetic Wave Theory)

[HH15] *Kapitel* 2.9. In: HUBBARD, J. H. ; HUBBARD, B. B.: *Vector Calculus, Linear Algebra, and Differential Forms: A Unified Approach (5th edition)*. Matrix Editions, 2015

[Hip97] HIPTMAIR, Ralf: Multigrid method for H(div) in three dimensions. In: *Electron. Trans. Numer. Anal* 6 (1997), Nr. 1, S. 133–152

[Hip99] HIPTMAIR, Ralf: Canonical construction of finite elements. In: *Mathematics of Computation* 68 (1999), Nr. 228, S. 1325–1346

[HS98] HUI, Gary Man-Kwong ; SWANN, Howard: On orthogonal polynomial bases for triangles and tetrahedra invariant under the symmetric group. In: *Contemporary Mathematics* 218 (1998), S. 438–446

[HX07] HIPTMAIR, Ralf ; XU, Jinchao: Nodal auxiliary space preconditioning in H(curl) and H(div) spaces. In: *SIAM Journal on Numerical Analysis* 45 (2007), Nr. 6, S. 2483–2509

[IIN09] ILIC, Milan M. ; ILIC, Andjelija Z. ; NOTAROŠ, Branislav M.: Continuously inhomogeneous higher order finite elements for 3-D electromagnetic analysis. In: *IEEE Transactions on Antennas and Propagation* 57 (2009), Nr. 9, S. 2798–2803

[Ing06] INGELSTRÖM, P.: A new set of H(curl)-conforming hierarchical basis functions for tetrahedral meshes. In: *IEEE Transactions on Microwave Theory and Techniques* 54 (2006), Jan, Nr. 1, S. 106–114

[Jen07] JENKINS, Eleanor W.: Numerical solution of the acoustic wave equation using Raviart-Thomas elements. In: *Journal of Computational and Applied Mathematics* 206 (2007), Nr. 1, S. 420 – 431

[JGBY16] JIANG, Hao ; GRAILLAT, Stef ; BARRIO, Roberto ; YANG, Canqun: Accurate, validated and fast evaluation of elementary symmetric functions and its application. In: *Applied Mathematics and Computation* 273 (2016), S. 1160–1178

[Joc13] JOCHUM, Martin T.: *Niederfrequenzstabile Potenzialformulierungen zur Finite-Elemente-Simulation elektromagnetischer Felder im Frequenzbereich*. Saarländische Universitäts-und Landesbibliothek, 2013

[KK15] KIRBY, Robert C. ; KIEU, Thinh T.: Symplectic-mixed finite element approximation of linear acoustic wave equations. In: *Numerische Mathematik* 130 (2015), Jun, Nr. 2, S. 257–291

[Koo75] KOORNWINDER, Tom: Two-variable analogues of the classical orthogonal polynomials. In: *Theory and Application of Special Functions*. Elsevier, 1975, S. 435–495

[LAH09] Li, Jichun ; Arbogast, Todd ; Huang, Yunqing: Mixed methods using standard conforming finite elements. In: *Computer Methods in Applied Mechanics and Engineering* 198 (2009), Nr. 5, S. 680 – 692

[Leo17] Leoni, G.: *A First Course in Sobolev Spaces*. American Mathematical Society, 2017 (Graduate studies in mathematics)

[LKI17] Lee, Chan ; Kim, Hobeom ; Im, Seyoung: Polyhedral elements by means of node/edge-based smoothed finite element method. In: *International Journal for Numerical Methods in Engineering* 110 (2017), Nr. 11, S. 1069–1100

[LM04] Lech, R. ; Mazur, J.: Propagation in rectangular waveguides periodically loaded with cylindrical posts. In: *IEEE Microwave and Wireless Components Letters* 14 (2004), Nr. 4, S. 177–179

[LM07] Lech, R. ; Mazur, J.: Analysis of circular cavity with cylindrical objects. In: *IEEE Transactions on Microwave Theory and Techniques* 55 (2007), Oct, Nr. 10, S. 2115–2123

[Löw50] Löwdin, Per-Olov: On the non-orthogonality problem connected with the use of atomic wave functions in the theory of molecules and crystals. In: *The Journal of Chemical Physics* 18 (1950), Nr. 3, S. 365–375

[M+03] Monk, Peter u. a.: *Finite Element Methods for Maxwell's Equations*. Oxford University Press, 2003

[Mar51] Marcuvitz, Nathan: *Waveguide Handbook*. McGraw-Hill, 1951 (IEEE electromagnetic waves series 21)

[May02] Mayer, István: On Löwdin's method of symmetric orthogonalization. In: *International Journal of Quantum Chemistry* 90 (2002), Nr. 1, S. 63–65

[May04] Mayer, I: Löwdin population analysis is not rotationally invariant. In: *Chemical physics letters* 393 (2004), Nr. 1-3, S. 209–212

[MMIN12] Manić, Ana B. ; Manić, Sanja B. ; Ilić, Milan M. ; Notaroš, Branislav M.: Large anisotropic inhomogeneous higher order hierarchical generalized hexahedral finite elements for 3-D electromagnetic modeling of scattering and waveguide structures. In: *Microwave and Optical Technology Letters* 54 (2012), Nr. 7, S. 1644–1649

[Néd80] Nédelec, J. C.: Mixed finite elements in R3. In: *Numerische Mathematik* 35 (1980), Sep, Nr. 3, S. 315–341

[Néd86] Nédélec, J. C.: A new family of mixed finite elements in R3. In: *Numerische Mathematik* 50 (1986), Jan, Nr. 1, S. 57–81

[Nel76] Nelson, R. B.: Simplified calculation of eigenvector derivatives. In: *AIAA Journal* 14 (1976), Nr. 9, S. 1201–1205

[NW04] Nair, D ; Webb, JP: P-adaptive computation of the scattering parameters of 3-D microwave devices. In: *IEEE Transactions on Magnetics* 40 (2004), Nr. 2, S. 1428–1431

[OTV20] Olver, Sheehan ; Townsend, Alex ; Vasil, Geoffrey M.: Recurrence relations for a family of orthogonal polynomials on a triangle. In: *Spectral and High Order Methods for Partial Differential Equations ICOSAHOM 2018*. Springer, Cham, 2020, S. 79–92

[PDEL99] Peng, Guanghua ; Dyczij-Edlinger, R. ; Lee, Jin-Fa: Hierarchical methods for solving matrix equations from TVFEMs for microwave components. In: *IEEE Transactions on Magnetics* 35 (1999), Nr. 3, S. 1474–1477

[PKCR13] Patel, Amit ; Kosta, Yogeshwarprasad ; Chhasatia, Neetirajsinh ; Raval, Falguni: Design and fabrication of microwave waveguide resonator: with improved characteristic response. In: *European Journal of Scientific Research* 102 (2013), Nr. 2, S. 163–174

[Poz11] Pozar, D.M.: *Microwave Engineering, 4th Edition*. Wiley, 2011

[PYA18] Pfeiffer, R. A. ; Young, J. C. ; Adams, R. J.: Divergence-conforming constrained basis functions for hexahedral volume elements. In: *IEEE Transactions on Antennas and Propagation* 66 (2018), Jan, Nr. 1, S. 501–504

[RA05] Rababah, A ; Alqudah, M: Jacobi-weighted orthogonal polynomials on triangular domains. In: *Journal of Applied Mathematics* 2005 (2005), Nr. 3, S. 205–217

[RGB14] Rand, Alexander ; Gillette, Andrew ; Bajaj, Chandrajit: Quadratic serendipity finite elements on polygons using generalized barycentric coordinates. In: *Mathematics of Computation* 83 (2014), Nr. 290, S. 2691–2716

[RS12] Ratnani, Ahmed ; Sonnendrücker, Eric: An arbitrary high-order spline finite element solver for the time domain Maxwell equations. In: *Journal of Scientific Computing* 51 (2012), Nr. 1, S. 87–106

[RSS08] Rokob, András T ; Szabados, Ágnes ; Surján, Peter R.: A Note on the Symmetry Properties of Löwdin's Orthogonalization Schemes. In: *Collection of Czechoslovak Chemical Communications* 73 (2008), Nr. 6, S. 937–944

[RT77] Raviart, Pierre-Arnaud ; Thomas, Jean-Marie: A mixed finite element method for 2-nd order elliptic problems. In: *Mathematical Aspects of Finite Element Methods*. Springer, 1977, S. 292–315

[Run03] Runborg, Olof: Notes on polynomial interpolation. In: *Kungliga Tekniska Högskolan, 2d1250, Tillämpade numeriska metoder II* (2003). https://www-users.cselabs.umn.edu/classes/Spring-2020/csci5302/Notes/Classnotes/interp.pdf

[RW14] Rjasanow, Sergej ; Weisser, Steffen: FEM with Trefftz trial functions on polyhedral elements. In: *Journal of Computational and Applied Mathematics* 263 (2014), S. 202–217

[Sau07] Sauer, Tomas: Jacobi polynomials in Bernstein form. In: *Journal of Computational and Applied Mathematics* 199 (2007), Nr. 1, S. 149–158

[SB91] Szabó, Barna ; Babuška, Ivo: *Finite Element Analysis*. John Wiley & Sons, 1991

[Sch64] Schoenberg, IJ: On best approximations of linear operators. In: *Nederl. Akad. Wetensch. Proc. Ser. A* Bd. 67, 1964, S. 155–163

[Sch69] Schultz, Martin H.: L2-multivariate approximation theory. In: *SIAM Journal on Numerical Analysis* 6 (1969), Nr. 2, S. 184–209

[SF73] Strang, Gilbert ; Fix, George J.: *An Analysis of the Finite Element Method*. Bd. 212. Prentice-hall Englewood Cliffs, NJ, 1973

[SH12] Shunn, L. ; Ham, F.: Symmetric quadrature rules for tetrahedra based on a cubic close-packed lattice arrangement. In: *Journal of Computational and Applied Mathematics* 236 (2012), Nr. 17, S. 4348–4364

[Sil78] Silvester, P: Construction of triangular finite element universal matrices. In: *International Journal for Numerical Methods in Engineering* 12 (1978), Nr. 2, S. 237–244

[Ska08] Skala, Vaclav: Barycentric coordinates computation in homogeneous coordinates. In: *Computers and Graphics* 32 (2008), Nr. 1, S. 120–127

[SLC01] Sun, Din-Kow ; Lee, Jin-Fa ; Cendes, Zoltan: Construction of nearly orthogonal Nedelec bases for rapid convergence with multilevel preconditioned solvers. In: *SIAM Journal on Scientific Computing* 23 (2001), Nr. 4, S. 1053–1076

[SWG84] Schaubert, D. ; Wilton, D. ; Glisson, A.: A tetrahedral modeling method for electromagnetic scattering by arbitrarily shaped inhomogeneous dielectric bodies. In: *IEEE Transactions on Antennas and Propagation* 32 (1984), January, Nr. 1, S. 77–85

[Tar07] Tartar, Luc: *An Introduction to Sobolev Spaces and Interpolation Spaces*. Bd. 3. Springer Science & Business Media, 2007

[VW99] Villeneuve, D. ; Webb, J. P.: Hierarchical universal matrices for triangular finite elements with varying material properties and curved boundaries. In: *International Journal for Numerical Methods in Engineering* 44 (1999), Nr. 2, S. 215–228

[Web99] Webb, J. P.: Hierarchal vector basis functions of arbitrary order for triangular and tetrahedral finite elements. In: *IEEE Transactions on Antennas and Propagation* 47 (1999), Aug, Nr. 8, S. 1244–1253

[WXY12] Wheeler, Mary ; Xue, Guangri ; Yotov, Ivan: A multipoint flux mixed finite element method on distorted quadrilaterals and hexahedra. In: *Numerische Mathematik* 121 (2012), May, Nr. 1, S. 165–204

[Zag06] Zaglmayr, Sabine: *High Order Finite Element Methods for Electromagnetic Field Computation*. A-4040 Linz, Altenbergerstraße 69, Johannes Kepler Universität, Linz, Diss., 2006

[ZC02] Zhu, Yu ; Cangellaris, Andreas C.: Hierarchical multilevel potential preconditioner for fast finite-element analysis of microwave devices. In: *IEEE Transactions on Microwave Theory and Techniques* 50 (2002), Nr. 8, S. 1984–1989

[ZC06] Zhu, Yu ; Cangellaris, Andreas C.: *Multigrid Finite Element Methods for Electromagnetic Field Modeling*. Bd. 28. John Wiley & Sons, 2006

[ZCL09] Zhang, Linbo ; Cui, Tao ; Liu, Hui: A set of symmetric quadrature rules on triangles and tetrahedra. In: *Journal of Computational Mathematics* (2009), S. 89–96

[ZK00] Zhang, Jing ; Kikuchi, Fumio: Interpolation error estimates of a modified 8-node serendipity finite element. In: *Numerische Mathematik* 85 (2000), Nr. 3, S. 503–524