

Maximally Orthogonal High-Order Basis Functions have a Well-Conditioned Gram Matrix

Ignace Bogaert, *Member, IEEE*, and Francesco P. Andriulli, *Senior Member, IEEE*

Abstract—Recently, a novel high-order finite element space for wires, quadrilaterals and hexahedrons was presented [1]. Numerical results have shown a very favorable behavior of the condition number of the Gram matrix of this finite element space as a function of the polynomial degree. In this paper, this high-order finite element space is recognized to be expressible in terms of Jacobi polynomials, which can be easily computed using a three-term recurrence. In addition, the condition number of the Gram matrix of the one-dimensional finite element space is rigorously analyzed for the general case of a piecewise smooth (possibly curved) geometry. An explicit upper bound for the condition number in terms of the mesh quality is proved. This bound implies that the one-dimensional finite element space is stable for arbitrarily high polynomial degree. Numerical results corroborate the theoretical results and show that the basis can be used to perform *hp*-refinement, leading to an accurate handling of both large smooth regions and corners.

Index Terms—High-order basis functions, polynomials, conditioning

I. INTRODUCTION

HIGH-order finite elements have attracted much attention [2]–[7] because of their potential to yield much better accuracy for the same number of unknowns or, similarly, a smaller number of unknowns for the same accuracy. Unfortunately, high-order subdomain finite element spaces have historically suffered from significant linear dependence between the various basis functions. This in turn leads to ill-conditioned matrices resulting from the discretization of partial differential equations or integral equations. While the ill-conditioning due to the pseudo-differential nature of the equation (be it partial differential or integral) can usually be mitigated using multigrid, multiresolution or Calderón preconditioners [8]–[11], this is usually not as easily done for the ill-conditioning due to linear dependence between the different basis functions. It is therefore of great practical importance to find high-order subdomain finite element spaces that show sufficient linear independence between the various finite elements.

A way to quantify the amount of linear dependence of a finite element space is to look at the condition number of the Gram matrix associated with the space. If the finite elements are denoted as $b_n, \forall n \in [1, B]$, the elements of the Gram matrix are defined as

$$[G]_{n,m} = \int_{\Omega} b_n(\mathbf{r})b_m(\mathbf{r})d\mathbf{r}, \forall n, m \in \{1, B\}, \quad (1)$$

Ignace Bogaert is with the Department of Information Technology, Ghent University, Sint-Pietersnieuwstraat 41, B-9000 Gent, Belgium, e-mail: Ignace.Bogaert@intec.UGent.be. The work of I. Bogaert is supported by a postdoctoral grant from the Research Foundation Flanders (FWO Vlaanderen).

Francesco Andriulli is with the Microwave Department of Telecom Bretagne - Institut Mines-Telecom, Brest, France.

where the integration is over the computational domain Ω . For a boundary element method, Ω can be the boundary of another, higher-dimensional, domain or even a one-dimensional wire embedded in three dimensional space. It is clear that if two finite elements are linearly dependent, the smallest singular value of the Gram matrix will be zero, leading to an infinite condition number. In this sense the condition number of the Gram matrix is related to the amount of linear dependence in the finite element space. However, there is also a more direct link. For example, when discretizing an integral equation of the second kind such as the magnetic field integral equation [12], [13], the Gram matrix arises as an integral part of the final matrix equation. Clearly, acquiring explicit bounds for the condition number of the Gram matrix has immediate use for these types of equations.

For a long time, no div-conforming high-order subdomain finite element spaces were known that exhibit bounded condition numbers for their Gram matrix. The div-conformity of the finite elements requires that the finite elements have continuous normal components globally across the mesh. In practice, this requirement turned out to be hard to satisfy while keeping the condition number of the Gram matrix under control, both for hierarchical [5], [7], [14], [15] and non-hierarchical bases [2].

However, recently, novel div-conforming high-order finite element spaces were presented for generalized wires, quadrilaterals and hexahedrons [1]. Preliminary results were also presented in [16]. By means of numerical experiments, these finite element spaces have been shown to be almost orthogonal up to degree 8. One of the most significant observations from these numerical results is that the orthogonality seems to *improve* with higher polynomial degree, which lends credibility to the idea that this finite element space might have a bounded condition number, irrespective of the polynomial degree. This is clearly a property that should be regarded as highly valuable, especially because the space also allows full *hp*-refinement [17]–[19], just like hierarchical finite element spaces.

The finite element spaces from [1] are based on a set of two node functions $Q_0(s), Q_1(s)$ and $N-1$ segment functions $Q_n(s), \forall n \in [2, N]$ (where N is the polynomial degree of the finite element space) that is maximally orthogonalized. Other sets of functions have been proposed [7], [20], but these have been shown to not lead to well-conditioned Gram matrices by means of numerical experiments.

In this paper, the fact that the functions $Q_n(s)$ do lead to a well-conditioned Gram matrix will be proved in a mathematically rigorous way for the specific case of the one-dimensional finite elements (i.e. the finite element space for generalized

wires). First, explicit formulas will be developed for the segment and node functions. For the segment functions, it will be shown that the Gram-Schmidt procedure proposed in [1] leads to Jacobi polynomials, such that a three-term recurrence can be used instead of the Gram-Schmidt orthogonalization. For the node functions, a representation in terms of derived Legendre polynomials is proved. These explicit representations are useful for all three finite element spaces proposed in [1]. Subsequently, these explicit formulas are used to derive an upper bound for the condition number of the Gram matrix of the generalized wire finite element space. It is shown that this upper bound indeed decreases for increasing polynomial degree, thereby proving the stability of the finite element space up to arbitrary degree. Numerical results subsequently confirm the theoretical predictions for polynomial degrees as high as 3036.

Throughout the rest of this paper, the polynomial degree of the finite element space will be denoted by N .

II. EXPLICIT EXPRESSIONS FOR THE SEGMENT AND NODE FUNCTIONS

Here, the segment and node functions proposed in [1] will be cast into analytical formulas. These formulas have advantages both from an implementation point of view and for the theoretical understanding of the stability of the finite element space. Among other things, it will be proved that the orthogonality factor defined in [1] monotonically improves with rising polynomial degree N .

A. Segment Functions

In [1], the segment function $Q_n(s)$, with $n \in [2, N]$, is constructed such that it is:

- a polynomial of degree n ,
- L_2 -orthogonal to $Q_m(s), \forall m \in [2, n-1]$,
- zero on the boundary, i.e. for $s = \pm 1$.

The used construction method is a Gram-Schmidt orthogonalization procedure, which starts from a collection of $N-1$ polynomials $S_k(s) = P_k(s) - P_{k-2}(s), \forall k \in [2, N]$. The polynomial $P_k(s)$ is the Legendre polynomial [21] of degree k , which guarantees that $S_k(\pm 1) = 0$. In general, the result of Gram-Schmidt orthogonalization depends on the order (i.e. which one comes first) of the input functions. However, here, the condition that $Q_n(s)$ is a polynomial of degree n (and not higher) fixes the input order. Hence the functions $Q_n(s)$ are uniquely determined, up to normalization factors, by the three conditions (2).

It will now be shown that the Gram-Schmidt procedure can be replaced by a three-term recurrence on a suitable set of Jacobi polynomials [21]. We claim that the segment functions can be defined as follows:

$$Q_n(s) = K_n(1-s^2)P_{n-2}^{2,2}(s), \quad \forall n \in [2, N]. \quad (3)$$

with

$$K_n = \frac{1}{4} \sqrt{\frac{(2n+1)(n+1)(n+2)}{2n(n-1)}}. \quad (4)$$

The functions $Q_n(s)$ are depicted in Figure 1 for $n \in [2, 4]$. Due to the uniqueness of the segment functions, it needs only be shown that these segment functions satisfy the three conditions (2) to prove that they constitute the same set of functions (up to normalization factors).

The first condition is trivially satisfied because the Jacobi polynomial $P_{n-2}^{2,2}(s)$ is of degree $n-2$. The factor $(1-s^2)$ in front of the the Jacobi polynomial then drives up the degree of $Q_n(s)$ to n . This factor is also directly responsible for the satisfaction of the third condition in (2). That leaves the second condition to be proved. Due to the orthogonality properties of the Jacobi polynomials [21], the following holds

$$\int_{-1}^1 Q_n(s)Q_m(s)ds = \delta_{n,m}, \quad \forall n, m \geq 2 \quad (5)$$

i.e. the segment functions form an orthonormal set. Clearly, definition (3) is advantageous from an implementation point of view because the Jacobi polynomials satisfy the recurrence relation

$$2n(n+4)P_n^{2,2}(s) = (2n+4)(2n+3)sP_{n-1}^{2,2}(s) - (n+1)(2n+4)P_{n-2}^{2,2}(s). \quad (6)$$

This allows for a rapid and numerically stable evaluation of the segment functions $Q_n(s)$ for arbitrarily high degree.

It should be noted that these segment functions (3) have also been discussed in [22] and [4]. There, however, the accompanying node functions were not chosen maximally orthogonal to the segment functions, leading to a growing linear dependence between the various finite elements as the degree is increased.

B. Node Functions

The node functions $Q_0(s)$ and $Q_1(s)$ are polynomial functions of degree N that satisfy the following conditions

- $Q_0(1) = 0 = Q_1(-1), \quad Q_0(-1) \neq 0 \neq Q_1(1)$,
- $Q_0(s), Q_1(s)$ are L_2 -orthogonal to $Q_m(s), \forall m \in [2, N]$.

(7)

To find these functions, one possibility is to project away the segment functions $Q_m(s), \forall m \in [2, N]$ from initial functions $1-s$ and $1+s$ [1]. However, as will be shown, it is also possible to relate the node functions to the derivative of Legendre polynomials. To show this, consider the following two polynomials:

$$u_0(s) = \frac{d}{ds}P_N(s), \quad (8)$$

$$u_1(s) = \frac{d}{ds}P_{N+1}(s). \quad (9)$$

These two polynomials are orthogonal to the segment functions because $\forall j \in \{0, 1\}$:

$$\begin{aligned}
 \int_{-1}^1 u_j(s) Q_m(s) ds &= \\
 &= K_m \int_{-1}^1 P_{m-2}^{2,2}(s) (1-s^2) \frac{d}{ds} P_{N+j}(s) ds, \\
 &= \frac{2K_m}{n+2} \int_{-1}^1 \frac{d}{ds} P_{m-1}^{1,1}(s) (1-s^2) \frac{d}{ds} P_{N+j}(s) ds, \\
 &= -\frac{2K_m}{n+2} \int_{-1}^1 P_{m-1}^{1,1}(s) \frac{d}{ds} \left[(1-s^2) \frac{d}{ds} P_{N+j}(s) \right] ds, \\
 &= (N+j)(N+j+1) \frac{2K_m}{n+2} \int_{-1}^1 P_{m-1}^{1,1}(s) P_{N+j}(s) ds, \\
 &= 0, \forall m \leq N. \tag{10}
 \end{aligned}$$

This result implies that the functions $u_j(s)$ span the space of node functions. Then the only task left is to find the linear combinations that satisfy the first of conditions (7):

$$Q_0^N(s) = (-1)^N \frac{Nu_1(s) - (N+2)u_0(s)}{\sqrt{2N(N+1)^2(N+2)}}, \tag{11}$$

$$Q_1^N(s) = \frac{Nu_1(s) + (N+2)u_0(s)}{\sqrt{2N(N+1)^2(N+2)}}. \tag{12}$$

The factor in the denominators has been chosen such that

$$\int_{-1}^1 Q_j^N(s) Q_j^N(s) ds = 1, \forall j \in \{0, 1\}. \tag{13}$$

Because the functions defined in (11) and (12) depend on the maximal degree N , the superscript has been added in these definitions. For uniformity of notation, this superscript may be used for the segment functions as well throughout this paper. Figure 1 depicts these two node functions for $N = 4$. As can be seen, these functions have most of their weight clustered at the edges of the interval $[-1, 1]$. This clustering increases for increasing degree N , due to the orthogonality with the segment functions.

For the purpose of stitching together two node functions on adjacent segments to form a continuous function, it is useful to know the value of the node functions in the points $s = \pm 1$. Now, since

$$u_j(\pm 1) = (\pm 1)^{N+j+1} \frac{(N+j)(N+j+1)}{2}, \tag{14}$$

the node functions can be evaluated on the boundary, which yields

$$Q_0^N(-1) = Q_1^N(1) = \sqrt{\frac{N(N+2)}{2}}. \tag{15}$$

C. Orthogonality

Now that expressions have been derived for the segment and node functions, it becomes possible to determine the orthogonality properties of these functions on one segment. To this end, consider the $N+1$ by $N+1$ Gram matrix of the segment and node functions:

$$[\mathbb{G}_N]_{n,m} = \int_{-1}^1 Q_n^N(s) Q_m^N(s) ds, \forall n, m \in \{0, N\}. \tag{16}$$

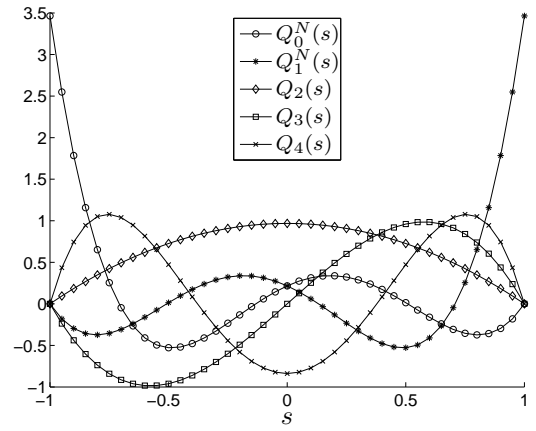


Fig. 1. The first few segment and node functions for $N = 4$.

In [1], a numerical study was done on a so-called orthogonality factor O_N , which is a quantity derived from the Gram matrix:

$$O_N = \sum_{n=0}^N \sum_{m=0, m \neq n}^N \left| [\mathbb{G}_N]_{n,m} \right|. \tag{17}$$

It is clear that a low orthogonality factor is desirable, because the Gram matrix of the segment and node functions becomes more diagonally dominant, resulting in a better condition number and a better linear independence. A remarkable result from the numerical experiment in [1] was that the orthogonality factor *decreases* as the degree is increased. This unique property will now be proved using the analytical expressions derived in the above.

First, it is clear that the only nonzero off-diagonal elements of the Gram matrix are

$$[\mathbb{G}_N]_{0,1} = [\mathbb{G}_N]_{1,0} = \int_{-1}^1 Q_0^N(s) Q_1^N(s) ds. \tag{18}$$

Expressed in terms of the functions $u_0(s)$ and $u_1(s)$, this integral becomes:

$$\int_{-1}^1 Q_0^N(s) Q_1^N(s) ds = C_N \int_{-1}^1 N^2 u_1^2(s) - (N+2)^2 u_0^2(s) ds, \tag{19}$$

with

$$C_N = \frac{(-1)^N}{2N(N+1)^2(N+2)}. \tag{20}$$

Using the result

$$\int_{-1}^1 \left[\frac{d}{ds} P_n(s) \right]^2 ds = n(n+1), \tag{21}$$

the following is immediately obtained

$$\int_{-1}^1 Q_0^N(s) Q_1^N(s) ds = -\frac{(-1)^N}{N+1}. \tag{22}$$

As a consequence, the orthogonality factor becomes

$$O_N = \frac{2}{N+1}. \tag{23}$$

This theoretical result matches perfectly with the numerically obtained results in Figure 5 in [1]. It also proves that the orthogonality factor will keep on decreasing for arbitrarily high degree N , ultimately converging to zero.

As an alternative to the orthogonality factor, one can also compute the condition number κ_N of the Gram matrix. The proof is elementary and we give only the result:

$$\kappa_N = 1 + \frac{2}{N}. \quad (24)$$

This shows once more that the segment and node functions $Q_n(s)$ constitute an almost orthogonal set for arbitrarily high polynomial degrees. A useful additional result is that

$$\frac{N}{N+1} \|\mathbf{v}\|^2 \leq \mathbf{v}^T \cdot \mathbf{G}_N \cdot \mathbf{v} \leq \frac{N+2}{N+1} \|\mathbf{v}\|^2, \quad \forall \mathbf{v} \in \mathbb{R}^{N+1}. \quad (25)$$

III. A STABLE FINITE ELEMENT SPACE

In the previous section, the segment and node functions $Q_n(s)$ were shown to constitute an almost-orthonormal set on the range $[-1, 1]$. Of course, in practice, these functions are not used on the range $[-1, 1]$, but are rather used to construct a finite element space that is div-conforming. Imposing global normal continuity and the presence of coordinate stretching due to geometry representation will deteriorate the orthogonality properties of the basis. This notwithstanding, it will be shown that the condition number of the normally continuous finite element space can be bounded from above using a simple normalization strategy, provided that some regularity conditions are satisfied for the mesh. This normalization behaves similarly to symmetric diagonal preconditioning, which led to the good numerical results in [1].

To avoid overcomplicating the problem, we will focus on a one-dimensional finite element space, which can be used in for example three-dimensional wire scattering or in two-dimensional scattering for the transverse-electric case (magnetic field parallel to the invariant axis). The analysis of the condition number will be performed for a general (possibly) curved geometry, such that the most general conclusions can be drawn.

A. Finite Element Space

Consider a connected, closed scatterer that consists of P segments σ_p , $p \in [1, P]$, such that each segment is parameterized by s , which is restricted to the range $[-1, 1]$. The point on σ_p corresponding to a specific value of s is $\mathbf{r}_p(s)$. Associated with this representation is the [Jacobian](#)

$$J_p(s) = \|\mathbf{a}_p(s)\|, \quad (26)$$

with the tangential vector $\mathbf{a}_p(s)$ defined as

$$\mathbf{a}_p(s) = \frac{d}{ds} \mathbf{r}_p(s). \quad (27)$$

Using the Jacobian (26), the segment length l_p can be computed

$$l_p = \int_{-1}^1 J_p(s) ds. \quad (28)$$

Without loss of generality, it will be assumed that two segments have consecutive indices p and $p+1$ if they touch in a point. To make the scatterer closed, the index $P+1$ should be regarded as equivalent 1. In addition, it will be assumed that they are ordered such that $\mathbf{r}_p(1) = \mathbf{r}_{p+1}(-1)$.

In the discretization, each segment gets assigned a degree N_p , which signifies the polynomial degree of the expansion on this segment. Note that different segments are allowed to have different degrees, such that full hp -refinement becomes possible [17]–[19]. The finite elements \mathbf{g}_{mn} will be understood to be defined on the entire scatterer, being zero on all but one or two segments. To simplify the definition of these necessarily piecewise polynomial functions, we introduce the restriction operator $|_p$, such that

$$\mathbf{g}_{mn}|_p(s) = \text{the part of } \mathbf{g}_{mn} \text{ on segment } \sigma_p. \quad (29)$$

The finite element space consists of two parts, corresponding to the segment and node functions introduced in the above. The segment finite elements, $\forall n \in [2, N_m]$, are defined as

$$\mathbf{g}_{mn}|_p(s) = \begin{cases} p = m : & c_m Q_n(s) \frac{\mathbf{a}_p(s)}{J_p(s)} \\ p \neq m : & 0 \end{cases}. \quad (30)$$

From this, it becomes clear that m is the 'segment-index' and n is the 'degree-index' of \mathbf{g}_{mn} . Since the segment functions are zero on the boundary of the segment, these finite elements satisfy the normal continuity constraint. Also, the finite elements contain a constant factor c_m that will be used for the normalization purposes mentioned earlier. The factor c_m will be defined further on, in equation (35).

The nodal finite elements are more cumbersome, requiring the stitching together of two node functions on subsequent segments σ_m and σ_{m+1} . The nodal finite element \mathbf{g}_{m1} on this pair of segments will be defined as follows:

$$\mathbf{g}_{m1}|_p(s) = \begin{cases} p = m : & d_m \frac{Q_0^{N_p}(s)}{\sqrt{N_p(N_p+2)}} \frac{\mathbf{a}_p(s)}{J_p(s)} \\ p = m+1 : & d_m \frac{Q_0^{N_p}(s)}{\sqrt{N_p(N_p+2)}} \frac{\mathbf{a}_p(s)}{J_p(s)} \\ p \neq m \text{ or } m+1 : & 0 \end{cases}. \quad (31)$$

This definition guarantees div-conformity because, using equation (15), it can be checked that

$$\frac{\mathbf{a}_m(1)}{J_m(1)} \cdot \mathbf{g}_{m1}|_m(1) = \frac{\mathbf{a}_{m+1}(-1)}{J_{m+1}(-1)} \cdot \mathbf{g}_{m1}|_{m+1}(-1) = \frac{d_m}{\sqrt{2}}. \quad (32)$$

B. Normalization and Mesh Regularity Assumptions

For the purpose of analyzing the condition number of this finite element space's Gram matrix, the normalization factors c_m and d_m need to be defined. To this end, we define

$$J_p^{\min} = \min_{s \in [-1, 1]} J_p(s), \quad (33)$$

$$J_p^{\max} = \max_{s \in [-1, 1]} J_p(s). \quad (34)$$

Using these extremal values of the Jacobian on each segment, the normalization factors will be defined as

$$c_p = \frac{1}{\sqrt{J_p^{\min}}}, \quad (35)$$

$$d_p = \sqrt{\max \left[\frac{N_p(N_p + 2)}{2J_p^{\min}}, \frac{N_{p+1}(N_{p+1} + 2)}{2J_{p+1}^{\min}} \right]}. \quad (36)$$

In addition, two regularity assumptions will be made. The first is an assumption on the geometrical refinement rate, and is expressed as follows:

$$\frac{\max [J_p^{\max}, J_{p+1}^{\max}]}{\min [J_p^{\min}, J_{p+1}^{\min}]} \leq \gamma < \infty, \forall p \in [1, P]. \quad (37)$$

A simple interpretation of this inequality can be found for meshes consisting of straight segments, for which $r_p(s)$ is a linear function of s . In this case it holds that

$$J_p^{\max} = J_p^{\min} = 2l_p, \quad (38)$$

such that criterion (37) simplifies to

$$\frac{\max [l_p, l_{p+1}]}{\min [l_p, l_{p+1}]} \leq \gamma < \infty, \forall p \in [1, P]. \quad (39)$$

This criterion is simply a bound on how quickly the segment length is allowed to change. This is a common measure for mesh quality and it is therefore not surprising that a generalization is found to be necessary for the case with curved segments. A simple but useful corollary of (37) is

$$\frac{J_{p+i}^{\max}}{J_{p+j}^{\min}} \leq \gamma, \forall p \in [1, P], \forall i, j \in \{0, 1\}. \quad (40)$$

The second assumption has to do with how quickly the segment degree N_p is allowed to change. It will be formalized as

$$\frac{\max [N_p(N_p + 2), N_{p+1}(N_{p+1} + 2)]}{\min [N_p(N_p + 2), N_{p+1}(N_{p+1} + 2)]} \leq \beta < \infty, \forall p \in [1, P]. \quad (41)$$

A constraint on the rate of change for N_p may seem a bit surprising at first, but it becomes more logical if one thinks of increasing the degree as increasing the spatial bandwidth of the finite element space. Indeed, refining the mesh very similarly increases this spatial bandwidth. Therefore, assumptions (37) and (41) can be seen as bounds on how quickly the spatial bandwidth of the finite element space is allowed to change.

C. The Gram Matrix

Now that the finite element space has been defined, we can move on to investigate the condition number of the Gram matrix of this finite element space. Due to the presence of multiple coordinate systems, the Gram matrix can be defined in multiple ways. However, in this work, the Gram matrix induced by the L_2 norm on the boundary of the scatterer will be used because of its physical interpretation and its relevance for the conditioning of the magnetic field integral equation.

Written in terms of integrations over the P segments, the elements of the Gram matrix are given by

$$[\mathbf{G}_{FE}]_{mn, m'n'} = \sum_{p=1}^P \int_{-1}^1 g_{mn}|_p(s) g_{m'n'}|_p(s) J_p(s) ds. \quad (42)$$

Here, mn and $m'n'$ are composite indices for the rows and columns of \mathbf{G}_{FE} respectively, running over both the segment and degree indices.

Using the Gram matrix, the L_2 norm of any linear combination of finite elements h

$$h = \sum_{mn} [\mathbf{b}]_{mn} g_{mn}, \quad (43)$$

can be written as

$$\|h\|_{L_2}^2 = \mathbf{b}^T \cdot \mathbf{G}_{FE} \cdot \mathbf{b}. \quad (44)$$

To prove an upper bound for the condition number of the Gram matrix, a bound of the following type will be derived:

$$\alpha_- \|\mathbf{b}\|^2 \leq \mathbf{b}^T \cdot \mathbf{G}_{FE} \cdot \mathbf{b} \leq \alpha_+ \|\mathbf{b}\|^2. \quad (45)$$

The norm $\|\mathbf{b}\|^2$ is the euclidian norm of the coefficient vector. Of course, the above inequality implies that the smallest singular value of \mathbf{G}_{FE} is larger than α_- , and that the largest singular value is smaller than α_+ . This in turn leads to α_+/α_- as an upper bound on the condition number.

To derive an inequality such as (45), we start from a corollary of (42):

$$\sum_{p=1}^P J_p^{\min} \int_{-1}^1 [h|_p(s)]^2 ds \leq \|h\|_{L_2}^2 \leq \sum_{p=1}^P J_p^{\max} \int_{-1}^1 [h|_p(s)]^2 ds. \quad (46)$$

Let us now derive bounds for the integrals appearing in both the right and left hand sides. Inserting the definition of the finite elements into $h|_p(s)$ yields

$$\begin{aligned} h|_p(s) &= \sum_{mn} [\mathbf{b}]_{mn} g_{mn}|_p(s), \\ &= \sum_{n=2}^N [\mathbf{b}]_{pn} c_p Q_n(s) \frac{\mathbf{a}_p(s)}{J_p(s)} \\ &\quad + [\mathbf{b}]_{p1} d_p \frac{Q_1^{N_p}(s)}{\sqrt{N_p(N_p + 2)}} \frac{\mathbf{a}_p(s)}{J_p(s)} \\ &\quad + [\mathbf{b}]_{(p-1)1} d_{p-1} \frac{Q_0^{N_p}(s)}{\sqrt{N_p(N_p + 2)}} \frac{\mathbf{a}_p(s)}{J_p(s)}. \end{aligned} \quad (47)$$

Using this expression and (25), it becomes possible to derive the following bounds

$$\frac{N_p}{N_p + 1} D_p \leq \int_{-1}^1 [h|_p(s)]^2 ds \leq \frac{N_p + 2}{N_p + 1} D_p, \quad (48)$$

with

$$D_p = c_p^2 \sum_{n=2}^N [\mathbf{b}]_{pn}^2 + \frac{[\mathbf{b}]_{p1}^2 d_p^2}{N_p(N_p + 2)} + \frac{[\mathbf{b}]_{(p-1)1}^2 d_{p-1}^2}{N_p(N_p + 2)} \quad (49)$$

The two inequalities in (48) will now be used to find a lower bound for the left hand side of (46) and an upper bound for the right hand side of (46), thereby providing suitable values for α_{\pm} .

1) *Determining α_-* : Using (35) and (36), the following can be shown:

$$J_p^{\min} D_p = J_p^{\min} \left[c_p^2 \sum_{n=2}^N [\mathbf{b}]_{pn}^2 + \frac{[\mathbf{b}]_{p1}^2 d_p^2}{N_p(N_p+2)} + \frac{[\mathbf{b}]_{(p-1)1}^2 d_{p-1}^2}{N_p(N_p+2)} \right] \quad (50)$$

$$\geq \sum_{n=2}^N [\mathbf{b}]_{pn}^2 + \frac{1}{2} [\mathbf{b}]_{p1}^2 + \frac{1}{2} [\mathbf{b}]_{(p-1)1}^2. \quad (51)$$

The latter is proved by noting that

$$\frac{J_p^{\min} d_p^2}{N_p(N_p+2)} = \frac{1}{2} \max \left[1, \frac{J_p^{\min} N_{p+1}(N_{p+1}+2)}{J_{p+1}^{\min} N_p(N_p+2)} \right] \geq \frac{1}{2}. \quad (52)$$

Similarly:

$$\frac{J_p^{\min} d_{p-1}^2}{N_p(N_p+2)} \geq \frac{1}{2}. \quad (53)$$

Inequality (51) can now be leveraged to show that

$$\sum_{p=1}^P J_p^{\min} \frac{N_p}{N_p+1} D_p \geq \frac{N^{\min}}{N^{\min}+1} \|\mathbf{b}\|^2, \quad (54)$$

with

$$N^{\min} = \min_p [N_p]. \quad (55)$$

This result proves the leftmost inequality in (45) with

$$\alpha_- = \frac{N^{\min}}{N^{\min}+1}. \quad (56)$$

2) *Determining α_+* : On the other hand

$$J_p^{\max} D_p = J_p^{\max} \left[c_p^2 \sum_{n=2}^N [\mathbf{b}]_{pn}^2 + \frac{[\mathbf{b}]_{p1}^2 d_p^2}{N_p(N_p+2)} + \frac{[\mathbf{b}]_{(p-1)1}^2 d_{p-1}^2}{N_p(N_p+2)} \right], \quad (57)$$

$$\leq \gamma \sum_{n=2}^N [\mathbf{b}]_{pn}^2 + \frac{\gamma\beta}{2} [\mathbf{b}]_{p1}^2 + \frac{\gamma\beta}{2} [\mathbf{b}]_{(p-1)1}^2. \quad (58)$$

The reasoning behind this inequality is

$$J_p^{\max} \frac{d_p^2}{N_p(N_p+2)} = \frac{1}{2} \max \left[\frac{J_p^{\max}}{J_p^{\min}}, \frac{J_p^{\max} N_{p+1}(N_{p+1}+2)}{J_{p+1}^{\min} N_p(N_p+2)} \right], \quad (59)$$

$$\leq \frac{\gamma}{2} \max \left[1, \frac{N_{p+1}(N_{p+1}+2)}{N_p(N_p+2)} \right], \quad (60)$$

$$\leq \frac{\gamma}{2} \max [1, \beta], \quad (61)$$

$$\leq \frac{\gamma\beta}{2}. \quad (62)$$

In a similar way, the following can be derived

$$J_p^{\max} \frac{d_{p-1}^2}{N_p(N_p+2)} \leq \frac{\gamma\beta}{2}. \quad (63)$$

Inequality (58) allows us to proceed with finding an upper bound for the right hand side of (46):

$$\begin{aligned} & \sum_{p=1}^P J_p^{\max} \int_{-1}^1 [h|_p(s)]^2 ds \\ & \leq \sum_{p=1}^P J_p^{\max} \frac{N_p+2}{N_p+1} D_p, \\ & \leq \gamma \sum_{p=1}^P \frac{N_p+2}{N_p+1} \left[\sum_{n=2}^N [\mathbf{b}]_{pn}^2 + \frac{\beta}{2} [\mathbf{b}]_{p1}^2 + \frac{\beta}{2} [\mathbf{b}]_{(p-1)1}^2 \right], \\ & \leq \gamma \frac{N^{\min}+2}{N^{\min}+1} \sum_{p=1}^P \left[\sum_{n=2}^N [\mathbf{b}]_{pn}^2 + \frac{\beta}{2} [\mathbf{b}]_{p1}^2 + \frac{\beta}{2} [\mathbf{b}]_{(p-1)1}^2 \right], \\ & \leq \gamma\beta \frac{N^{\min}+2}{N^{\min}+1} \|\mathbf{b}\|^2. \end{aligned} \quad (64)$$

This means that

$$\alpha_+ = \gamma\beta \frac{N^{\min}+2}{N^{\min}+1}, \quad (65)$$

leads to a provably correct rightmost inequality in (45).

3) *Bounding the condition number*: Now we are in a position to construct the bound on the condition number of the Gram matrix of the finite element space:

$$\kappa [G_{FE}] \leq \frac{\alpha_+}{\alpha_-} = \gamma\beta \left[1 + \frac{2}{N^{\min}} \right]. \quad (66)$$

From this, it is clear that a higher minimal degree leads to a (slightly) better bound on the condition number. However, more importantly, it unequivocally proves that the condition number of the Gram matrix for the finite element space is bounded by a constant determined only by the rate of change of the mesh density and the selected polynomial degree. Therefore, it can be concluded that finite element spaces of arbitrary order can be used without compromising the condition number of the Gram matrix.

IV. NUMERICAL RESULTS

Numerical results will now be presented to support the theoretical derivations in the above. To this end, the condition number of the Gram matrix and the two dimensional magnetic field integral equation for the transverse-electric case will be investigated.

First, the condition number of the Gram matrix will be compared to the theoretically derived bounds. Then the full magnetic field integral equation will be solved to show that the finite element space indeed leads to correct solutions.

A. Gram Matrix

To test the condition number of the Gram matrix, an inhomogeneously curved geometry should be used. In this work, the lemniscate of Gerono [23] was used, with parameter representation

$$x(\phi) = \cos \phi, \quad y(\phi) = \frac{1}{2} \sin 2\phi, \quad \text{with } \phi \in \left[-\frac{\pi}{2}, \frac{\pi}{2} \right]. \quad (67)$$

This closed curve is depicted in Figure 2, along with the segmentation that was used. Clearly, the presence of the corner

requires the mesh to be refined in that region. Indeed, even though no singularity occurs in the current, the charge can be singular, which means that the current can have an infinite derivative. The m th segment is defined as the part of the lemniscate that lies between ϕ_m and ϕ_{m+1} with

$$\phi_m = \begin{cases} m = 0 : & -\frac{\pi}{2} \\ 0 < m \leq M : & [2^{m-M} - 1] \frac{\pi}{2} \\ M < m \leq 2M : & -\phi_{2M-m} \end{cases} \quad (68)$$

In total, there are $2M$ segments, where the parameter M determines the amount of refinement that is done around the corner. It is clear that, using such a refinement strategy, the length of subsequent segments will not differ by substantially more than a factor two, which is why (37) will be satisfied with a bounded γ .

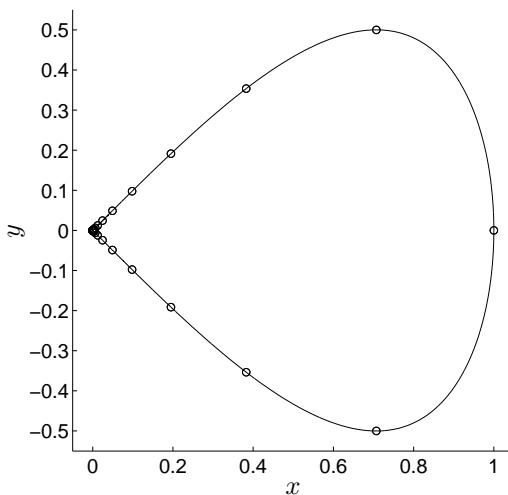


Fig. 2. The used lemniscate, refined near the corner by means of the division strategy given in (68) with $M = 25$.

For the numerical test, the degree assigned to each segment will be given by

$$N_p = 3 + \left\lceil \frac{5kl_p}{2\pi} \right\rceil. \quad (69)$$

Here, k is a number that governs how much spatial bandwidth the finite element space need to accommodate. The constants appearing in (69) are chosen somewhat arbitrarily, but the general form is constructed such that electrically long segments get assigned high polynomial degrees while short segments get assigned a fixed degree that is deemed sufficient for the 'low-frequency' region near the corner. This is a choice that corresponds to what one would normally choose while simulating a scattering problem at a wave number around k . Also, (69) guarantees that the rate of change of the degree is limited, based on the fact that the rate of change of the segment length is limited by (37). Therefore, (41) will be satisfied with a bounded β .

Table I shows the numerical results obtained using the choices (68) and (69) for various combinations of the refinement parameter M and the wave number k . The values for γ and β have been computed for each case, such that the bound (66) can be evaluated. It is clearly seen that the

M	k	γ	β	N^{\min}	N^{\max}	Bound (66)	$\kappa [G_{FE}]$
1	20	2.138	1	28	28	2.291	2.130
5	20	2.828	2.263	6	14	8.533	1.762
25	20	2.828	2.263	4	14	9.600	2.093
1	100	2.138	1	125	125	2.172	2.138
5	100	2.828	2.881	14	55	9.314	1.664
25	100	2.828	2.881	4	55	12.225	2.091
1	500	2.138	1	610	610	2.145	2.138
5	500	2.828	3.154	57	259	9.235	1.684
25	500	2.828	3.287	4	259	13.947	2.090
1	2500	2.138	1	3036	3036	2.139	2.138
5	2500	2.828	3.328	272	1279	9.481	1.695
25	2500	2.828	3.784	4	1279	16.056	2.086

TABLE I

THE BOUND ON THE CONDITION NUMBER AND THE ACTUAL COMPUTED CONDITION NUMBER FOR VARIOUS COMBINATIONS OF M AND k .

condition number of the Gram matrix is smaller than the corresponding bound for all cases, even for extreme refinement of the geometry (the longest segment is over three million times longer than the shortest one) and for large variations of the polynomial degree (4 on the shortest segments versus 1279 on the longest ones). It is also evident that basis functions of extremely high degree can be used without compromising the condition number of the Gram matrix, see for example the case where the degree of the polynomials used is 3036.

B. Magnetic Field Integral Equation

Finally, it will be shown that the proposed finite element space can be used to accurately solve scattering problems. The two dimensional magnetic field integral equation for the transverse-electric case is chosen to avoid ill-conditioning due to the pseudo-differential nature of the operator that is being discretized. Indeed, if the electric field integral equation were to be used, it would still lead to ill-conditioned matrices due to the dense-mesh breakdown. However, the crux is that many techniques exist for removing the ill-conditioning originating from the equation. Examples are multigrid, multiresolution or operator preconditioning [8], [9], [11]. Though these techniques are very effective, they cannot remove the ill-conditioning that comes from the linear dependence between finite elements if the Gram matrix bears no resemblance to any physical operator relevant to the problem being solved. The proposed finite element space removes this issue by having a well-conditioned Gram matrix.

For the scattering by a two-dimensional perfectly conducting scatterer Ω , the magnetic field integral equation for the transverse-electric case can be written as

$$\frac{1}{2}H_z(\mathbf{r}) + \int_{\partial\Omega} \hat{\mathbf{n}}(\mathbf{r}') \cdot \nabla G(\mathbf{r} - \mathbf{r}') H_z(\mathbf{r}') d\mathbf{r}' = H_z^{\text{inc}}(\mathbf{r}), \quad (70)$$

with $\hat{\mathbf{n}}(\mathbf{r}')$ the unit normal vector to the boundary $\partial\Omega$ and the Green's function

$$G(\mathbf{r}) = -\frac{j}{4} H_0^{(2)}(k \|\mathbf{r}\|), \quad (71)$$

where k is the wave number. The integral in (70) should be interpreted as *on* the boundary of the scatterer, such that the Dirac delta distribution in the integrand is not seen. Rather,

this part is the first term of (70). Also, the z -directed magnetic field is used to represent the tangential current on the scatterer: $\mathbf{J}(\mathbf{r}) = \hat{\mathbf{n}}(\mathbf{r}) \times [\hat{\mathbf{e}}_z H_z(\mathbf{r})]$. This current will be discretized by means of div-conforming finite elements, such that the solution has a continuous normal component. This avoids point charges, which would be non-physical. The test functions are chosen identical to the basis functions, such that the first term of (70) gives rise to the Gram matrix of the basis functions.

As a first scattering problem, a cylinder with unit radius is chosen because of the availability of an analytical solution. This allows the condition number of the *continuous* MFIE to be computed. Indeed, for a cylinder with radius a , the eigenvalues of the operator in the left hand side of (70) can be shown to be

$$\lambda_n = \frac{j\pi}{4ka} (H_{n-1}^{(2)}(ka) - H_{n+1}^{(2)}(ka)) J_n(ka). \quad (72)$$

Since the MFIE operator on a cylinder is diagonalized by unitary operators (i.e. the periodic Fourier transform on the boundary), the singular values are $|\lambda_n|$. Therefore, the condition number of the continuous MFIE is given by

$$\kappa_{\text{MFIE}}^{\text{Cont.}} = \frac{\sup_n |\lambda_n|}{\inf_n |\lambda_n|}. \quad (73)$$

It is worthwhile to point out that, for large n , $|\lambda_n|$ converges to $\frac{1}{2}$. This means that a moderate n is usually sufficient to compute $\kappa_{\text{MFIE}}^{\text{Cont.}}$.

For the numerical testing, a similar refinement as for the Lemniscate was used. The m th segment is defined as the part on the parameter curve

$$x(\phi) = \cos(2\phi + \pi), \quad y(\phi) = \sin(2\phi + \pi), \quad (74)$$

with ϕ between ϕ_m and ϕ_{m+1} , which have been defined in (68) and in which $M = 5$ was taken. This describes a cylinder and the resulting segmentation is shown in Figure 3. Of course, the refinement is not strictly necessary for a smooth geometry such as the cylinder, but it allows us to show that the robustness of the condition number does not rely on the uniformity of the mesh. The segment degree was

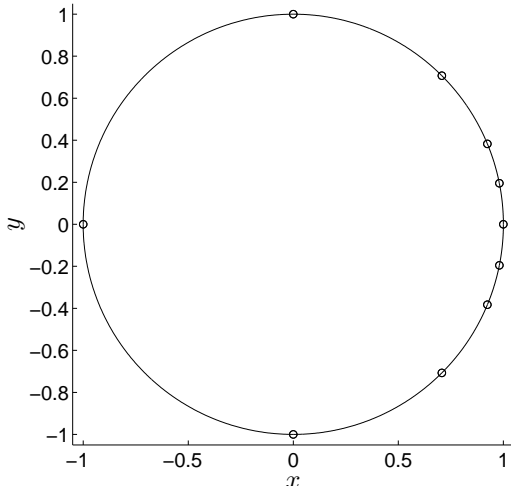


Fig. 3. The used cylinder, with parameter equation (74), refined near $\phi = \pm \frac{\pi}{2}$ by means of the division strategy given in (68) with $M = 5$.

k	Bound (66)	κ_{GFE}	$\kappa_{\text{MFIE}}^{\text{Cont.}}$	$\kappa_{\text{MFIE}}^{\text{Discr.}}$	L_2 -Error
1	4.38	1.97	2.36	2.57	0.13×10^{-3}
4	6.19	1.59	6.09	6.12	0.16×10^{-3}
16	7.43	1.38	30.5	30.6	0.31×10^{-4}
64	8.16	1.47	16.3	16.3	0.55×10^{-6}
256	8.11	1.49	5672	5673	0.58×10^{-9}

TABLE II
THE CONDITION NUMBER OF THE GRAM MATRIX, THE CONTINUOUS MFIE AND THE DISCRETIZED MFIE FOR VARIOUS VALUES OF k .

again set by means of equation (69). Table II shows the upper bound (66), the condition number of the Gram matrix, the condition number of the continuous MFIE and the condition number of the discretized MFIE (denoted by $\kappa_{\text{MFIE}}^{\text{Discr.}}$) for various wave numbers k . The achieved relative error (on the current) in the L_2 norm was also computed. As can be seen, the bound (66) is satisfied for all cases. In addition, the condition number of the discretized equation very closely matches the condition number of the continuous equation, especially when the condition number of the continuous equation is very high. Finally, it is also interesting to see that the relative L_2 -error decreases dramatically as the wave number k is increased. This is due to the fact that the constants in equation (69) are suboptimal for large k , i.e. they lead to segment degrees that are too high for the used wave number. This in turn leads to highly accurate results, demonstrating the approximation power of the finite element space for smooth solutions.

As a final numerical example, scattering of a plane wave impinging on the Lemniscate is modeled for $k = 25\text{m}^{-1}$, $M = 25$. The incoming plane wave propagates along the x -axis:

$$H_z^{\text{inc}}(\mathbf{r}) = e^{-jkx}. \quad (75)$$

The z -component of the total magnetic field is shown in figure 4, for points on the line $\mathbf{r} = (x, 0)$, $x \in [-1, 2]$. As can be seen, the total magnetic field is of the order 10^{-8} inside the structure. This shows that the high-order finite element space is capable of very high accuracy, even when the structure has sharp corners, if an appropriate refinement strategy is used. In addition, this refinement is compatible with a low condition number, in this case $\kappa_{\text{GFE}} = 2.092$ and $\kappa_{\text{MFIE}}^{\text{Discr.}} = 24.29$.

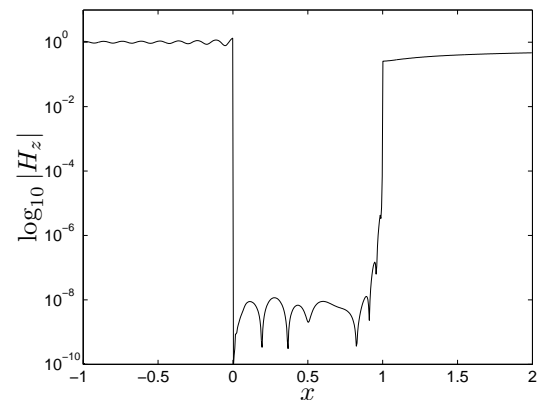


Fig. 4. The total magnetic field on the line $y = 0$, which cuts through the Lemniscate along the symmetry axis.

V. CONCLUSION

A recently proposed high-order finite element space has been rigorously analyzed. An alternative expression for the finite elements in terms of known orthogonal polynomials has been derived, leading to a simplified computation. **In addition, an explicit bound for the condition number of the Gram matrix of this finite element space has been derived for the one-dimensional case. This shows that the mesh refinement rate and degree variation rate are the two main factors determining the condition number.** To the best knowledge of the authors, this is the first div-conforming finite element space that is usable in hp -refinement and has a provably bounded condition number. Numerical results confirm the theoretical predictions, even for extreme polynomial degrees (3036) and mesh refinements.

ACKNOWLEDGMENT

The work of I. Bogaert was supported by a postdoctoral grant from the Research Foundation–Flanders (FWO–Vlaanderen).

REFERENCES

- [1] M. Kostic and B. Kolundzija, "Maximally Orthogonalized Higher Order Bases Over Generalized Wires, Quadrilaterals, and Hexahedra," *IEEE Transactions on Antennas and Propagation*, vol. 61, no. 6, pp. 3135–3148, 2013.
- [2] R. Graglia, D. Wilton, and A. Peterson, "Higher Order Interpolatory Vector Bases for Computational Electromagnetics," *Antennas and Propagation, IEEE Transactions on*, vol. 45, no. 3, pp. 329–342, 1997.
- [3] D. Sun, J. Lee, and Z. Cendes, "Construction of Nearly Orthogonal Nedelec Bases for Rapid Convergence with Multilevel Preconditioned Solvers," *SIAM Journal on Scientific Computing*, vol. 23, no. 4, pp. 1053–1076, 2001.
- [4] G. E. Karniadakis and S. J. Sherwin, *Spectral/hp Element Methods for CFD*. New York: Oxford University Press, 1999.
- [5] J. Schöberl and S. Zaglmayr, "High Order Nédélec Elements with Local Complete Sequence Properties," *International Journal for Computation and Mathematics in Electrical and Electronic Engineering*, vol. 24, no. 2, pp. 374–384, 2005.
- [6] M. Djordjevic and B. M. Notaro, "Higher-Order Hierarchical Basis Functions with Improved Orthogonality Properties for Moment-Method Modeling of Metallic and Dielectric Microwave Structures," *Microwave and Optical Technology Letters*, vol. 37, no. 2, pp. 83–88, 2003.
- [7] E. Jorgensen, J. Volakis, P. Meincke, and O. Breinbjerg, "Higher Order Hierarchical Legendre Basis Functions for Electromagnetic Modeling," *Antennas and Propagation, IEEE Transactions on*, vol. 52, no. 11, pp. 2985–2995, 2004.
- [8] Yu Zhu and Andreas C. Cangellaris, *Multigrid Finite Element Methods for Electromagnetic Field Modeling*. Wiley-IEEE Press, 2006.
- [9] F. P. Andriulli, A. Tabacco, and G. Vecchi, "A Multiresolution Approach to the Electric Field Integral Equation in Antenna Problems," *SIAM Journal on Scientific Computing*, vol. 29, no. 1, pp. 1–21, 2007.
- [10] F. P. Andriulli, K. Cools, H. Bagci, F. Olyslager, A. Buffa, S. Christiansen, and E. Michielssen, "A Multiplicative Calderon Preconditioner for the Electric Field Integral Equation," *IEEE Transactions on Antennas and Propagation*, vol. 56, no. 8, pp. 2398–2412, 2008.
- [11] R. Hiptmair, "Operator Preconditioning," *Computers and Mathematics with Applications*, vol. 52, no. 5, pp. 699 – 706, 2006.
- [12] F. Murray, "Conductors in an Electromagnetic Field," *Ant. J. Math.*, vol. 53, pp. 275–288, 1931.
- [13] A. Maue, "On the Formulation of a General Scattering Problem by Means of an Integral Equation," *Z. Phys.*, vol. 126(7/9), pp. 601–618, 1949.
- [14] M. Ainsworth and J. Coyle, "Hierarchic hp -edge Element Families for Maxwell's Equations on Hybrid Quadrilateral/Triangular Meshes," *Computer Methods in Applied Mechanics and Engineering*, vol. 190, no. 49-50, pp. 6709–6733, 2001.
- [15] R. Graglia, A. Peterson, and F. Andriulli, "Curl-Conforming Hierarchical Vector Bases for Triangles and Tetrahedra," *IEEE Transactions on Antennas and Propagation*, vol. 59, no. 3, pp. 950–959, 2011.
- [16] D. Sumic and B. Kolundzija, "Efficient Iterative Solution of Surface Integral Equations Based on Maximally Orthogonalized Higher Order Basis Functions," in *Antennas and Propagation Society International Symposium, 2005 IEEE*, vol. 4A, 2005, pp. 288–291 vol. 4A.
- [17] I. Babuska and M. Suri, "The p and $h-p$ Versions of the Finite Element Method, Basic Principles and Properties," *SIAM Review*, vol. 36, no. 4, p. 578632, 1994.
- [18] L. Demkowicz, P. Monk, L. Vardapetyan, and W. Rachowicz, "de Rham Diagram for hp -Finite Element Spaces," *Computers and Mathematics with Applications*, vol. 39, pp. 29–38, 2000.
- [19] L. Demkowicz, W. Rachowicz, and P. Devloo, "A Fully Automatic hp -Adaptivity," *Journal of Scientific Computing*, vol. 17, no. 1-4, pp. 117–142, 2002.
- [20] B. M. Kolundzija, "Electromagnetic Modeling of Composite Metallic and Dielectric Structures," *IEEE Trans. Microw. Theory Tech.*, vol. 47, p. 10211032, 1999.
- [21] M. Abramowitz and I.A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables*, ser. Advanced Mathematics. New York: Dover Publications, Inc., 1965.
- [22] J. P. Webb and R. Abouchacra, "Hierarchical Triangular Elements Using Orthogonal Polynomials," *International Journal for Numerical Methods in Engineering*, vol. 38, pp. 245–257, 1995.
- [23] Wikipedia, "Lemniscate of Geronon — Wikipedia. The Free Encyclopedia," accessed 7-Januari-2014. [Online]. Available: http://en.wikipedia.org/wiki/Lemniscate_of_Geronon