

# Survey of meshless and generalized finite element methods: A unified approach

Ivo Babuška\*

*Institute for Computational Engineering and Sciences,  
ACE 6.412, University of Texas at Austin,  
Austin, TX 78712, USA*

Uday Banerjee†

*Department of Mathematics, 215 Carnegie,  
Syracuse University, Syracuse, NY 13244, USA  
E-mail: banerjee@syr.edu  
http://bhaskara.syr.edu*

John E. Osborn†

*Department of Mathematics, University of Maryland,  
College Park, MD 20742, USA  
E-mail: jeo@math.umd.edu  
http://www.math.umd.edu/~jeo*

In the past few years meshless methods for numerically solving partial differential equations have come into the focus of interest, especially in the engineering community. This class of methods was essentially stimulated by difficulties related to mesh generation. Mesh generation is delicate in many situations, for instance, when the domain has complicated geometry; when the mesh changes with time, as in crack propagation, and remeshing is required at each time step; when a Lagrangian formulation is employed, especially with nonlinear PDEs. In addition, the need for flexibility in the selection of approximating functions (*e.g.*, the flexibility to use non-polynomial approximating functions), has played a significant role in the development of meshless methods. There are many recent papers, and two books, on meshless methods; most of them are of an engineering character, without any mathematical analysis.

In this paper we address meshless methods and the closely related generalized finite element methods for solving linear elliptic equations, using variational principles. We give a unified mathematical theory with proofs, briefly address implementational aspects, present illustrative numerical examples, and provide a list of references to the current literature.

\* Supported by NSF grant DMS-98-02367 and ONR grant N00014-99-1-0724.

† Partially supported by the TICAM Visiting Faculty Research Fellowship Program.

The aim of the paper is to provide a survey of a part of this new field, with emphasis on mathematics. We present proofs of essential theorems because we feel these proofs are essential for the understanding of the mathematical aspects of meshless methods, which has approximation theory as a major ingredient. As always, any new field is stimulated by and related to older ideas. This will be visible in our paper.

## CONTENTS

1	Introduction	2
2	The model problem	8
3	Approximation by local functions in $\mathbb{R}^n$ : the $h$ -version analysis	9
4	Construction and selection of particle shape functions	39
5	Superconvergence of the gradient of the solution in $L_2$	57
6	The generalized finite element method	86
7	Solutions of elliptic boundary value problems	98
8	Implementational aspects of meshless methods	113
9	Examples	116
10	Future challenges	119
	References	120

## 1. Introduction

### 1.1. *A brief historical review of the numerical solution of partial differential equations*

The numerical solution of partial differential equations has been of central importance for many years. Significant progress has been made in this area, especially in the last 30 years; this progress is directly related to the developments in computer technology. Methods such as, for example, the finite element method, are used in many applications.

Although significant progress has been made, numerical methods for the solution of differential equations are still often based on heuristic ideas, and verified by numerical experiments. Mathematical analysis is often shallow, and fails to address fully important issues that arise in the application of the methods to important problems in engineering and science.

There are three classical families of numerical methods for solving PDEs:

- (1) finite difference methods;
- (2) finite volume methods;
- (3) finite element methods.

These three families have two common, basic features:

- (a) they employ a mesh;
- (b) they use local approximation by polynomials.

We discuss each of these features in turn.

Mesh generation is often very expensive – especially in human effort. There are several reasons for this effort.

- The domain of the problem can have very complex geometry.
- The domain of the problem may change with time, which requires remeshing at each time step, as for example in the problem of crack propagation or when Lagrangian coordinates are used.
- Adaptive procedures require changes of mesh during computation.

Although great progress has been made in the theory and practice of mesh generation, the construction of the mesh is still a very delicate component of the numerical solution of differential equations. For this reason there is an interest in the development of methods that eliminate or reduce the need for a mesh.

Although polynomials have outstanding approximation properties, there are situations in which they are not effective. We mention problems whose solutions are not smooth, in the sense that they may not have several bounded derivatives. For such problems, there are sometimes other effective approximating functions, which we will refer to as *special*. The classical methods are not flexible in this regard: they do not use these special non-polynomial approximating functions. There is thus an interest in developing and analysing methods that can flexibly use these special approximation functions.

This created the need to develop methods that address both of these issues: the elimination, completely or partially, of the need for meshes; and the effective use of special (nonpolynomial) approximating functions. The inspiration for such methods came mainly from two sources.

The first of these sources is the class of classical particle methods that arise in physical simulation in connection with the Boltzmann equation or with fluid dynamics. Particle methods attempt to describe the motion of the atoms or their averages (or density) in Lagrangian coordinates (see Gingold and Monaghan (1977, 1982), Monaghan (1982, 1988), Nanbu (1980) and Neunzert and Struckmeier (1995), for example).

The other source is the idea of interpolation in the context of general *variational methods* (of Galerkin type). These methods select approximations from a finite-dimensional space, called the *trial space*, and, under certain general conditions, it is known that the error in approximation is no larger than a constant times the error in best approximation by functions in the trial space. Thus the quality of the method is determined by the approximation property of the trial space. It is thus natural to try to find a trial space that has good approximation properties. This property relates directly to interpolation by the approximating functions. For functions in one dimension this is a classical issue in numerical analysis, and, from around 1950, was studied in higher dimensions and for arbitrary distributions of points. It was recognized that the construction of trial spaces could be based on the idea of interpolation.

### 1.2. Meshless methods

Let us now make the discussion of variational methods more precise. We consider an elliptic PDE, which has the variational or weak form

$$u \in H_1, \quad B(u, v) = \mathcal{F}(v), \quad \text{for all } v \in H_2, \quad (1.1)$$

where  $H_1, H_2$  are two Hilbert spaces,  $B(u, v)$  is a bounded bilinear form on  $H_1 \times H_2$ , and  $\mathcal{F}(v)$  is a continuous linear functional on  $H_2$ . Under certain general conditions (the inf-sup or BB condition; see Babuška (1971), Babuška and Aziz (1972)), the solution  $u$  is characterized by (1.1). We are interested in approximating  $u$ . To that end, we assume we have two finite-dimensional spaces  $M_1 \subset H_1, M_2 \subset H_2$  that satisfy the discrete inf-sup condition (see Babuška and Aziz (1972)). The approximate solution  $u_{M_1}$  is characterized by

$$u_{M_1} \in M_1, \quad B(u_{M_1}, v) = \mathcal{F}(v), \quad \text{for all } v \in M_2. \quad (1.2)$$

As a consequence of the fact that  $M_1$  and  $M_2$  satisfy the discrete inf-sup condition, we know that the approximation  $u_{M_1}$  is quasi-optimal, that is,

$$\|u - u_{M_1}\|_{H_1} \leq C \inf_{\chi \in M_1} \|u - \chi\|_{H_1}. \quad (1.3)$$

We note that there are delicate problems related to the discrete inf-sup condition for problems that are not coercive, or where the spaces  $M_1$  and  $M_2$  are different, *e.g.*, in the mixed method or non-self-adjoint problems. Atluri and Shen (2002) use different spaces (without mathematical analysis of the discrete inf-sup condition).

Thus the quality of the approximation, that is, the error  $\|u - u_{M_1}\|_{H_1}$ , is mainly determined by the approximation properties of the trial space  $M_1$ , that is, by

$$E_1 = \inf_{\chi \in M_1} \|u - \chi\|_{H_1}.$$

It is therefore natural to select the trial space  $M_1$  so that  $E_1$  is small. To do this effectively we should use whatever information is available for the solution  $u$ . Note that, with a general variational method, as we have formulated it, there is no mention of a mesh. Of course, we may use a mesh to construct a good trial space; that, in fact, is exactly what is done with a finite element method. For example, the trial space might be the space of piecewise linear functions over a mesh.

Meshless methods, however, either avoid the use of a mesh, or use a mesh only minimally, for example, only for the numerical integration. The Petrov–Galerkin method given by (1.2) is a meshless method if the construction of  $M_1$  and  $M_2$  either does not require a mesh or requires a mesh only minimally. Thus, in designing meshless methods within the framework of variational methods, we have two general goals.

- (1) The construction of trial spaces  $M_1$  that effectively approximate the solution, and the construction of test spaces  $M_2$  ensuring the inf-sup (stability) condition.

If the solution has special features, *e.g.*, if it is not smooth, we should have the flexibility to use special approximating functions.

- (2) The minimizing of the need for a mesh.

In meshless methods, there is sometimes a mesh in the background, used for numerical integration, but we may not need a mesh generator.

We note that there are meshless methods that are not of the type given by (1.2), for instance, methods based on collocation, but the construction of approximating space follows the guidelines of the construction of  $M_1$ , as mentioned before.

The approximating (trial) spaces can be the spans of specific approximating functions (shape functions), with either global or local supports. Polynomials and non-compactly supported radial basis functions are examples of approximating functions that are defined over the entire domain of interest. See Mikhlin (1971) for a discussion of the use of polynomials and Buhmann (2000) and Powell (1992) for a discussion of the use of radial basis functions. Another type of approximating function is related to interpolation and data fitting procedures. For a survey of various approaches we refer to Atluri and Shen (2002), Dierckx (1995), Franke (1978, 1979), Gordon and Wixon (1978), Lancaster and Salkauskas (1981, 1986), McLain (1974), and Shepard (1968). Typical finite element approximating functions and spline functions have local supports. Babuška, Caloz and Osborn (1994) identified and analysed shape functions that are effective for the approximation of solutions of elliptic equations with rough coefficient; the idea in this paper was extended and developed in Babuška and Melenk (1997). The approximating functions used in Babuška *et al.* (1994) and Babuška and Melenk

(1997) can be characterized as solutions of particular homogeneous differential equations. In one dimension,  $L$ -splines – a generalization of splines that satisfy a differential equation – are used as approximating functions: see, *e.g.*, Varga (1971). Principles for the selection of shape function were addressed in Babuška, Banerjee and Osborn (2001, 2002*b*).

We note that in the engineering literature many names are used for methods that differ only in their implementation or in the shape functions employed: see, *e.g.*, De and Bathe (2001) and Sukumar, Moes, Moran and Belytschko (2000), among others. For a survey of results on meshless methods we refer to Babuška, Banerjee and Osborn (2002*a*), Belytschko, Krongauz, Organ, Fleming and Krysl (1996), Duarte (1995), Griebel and Schweitzer, eds (2002*a*), Li and Liu (2002), Liu (2002) and Schweitzer (200*x*).

One of the major problems of meshless methods is the imposition of boundary conditions, especially Dirichlet boundary conditions. It is well known that, if the underlying problem is a Dirichlet BVP, the essential boundary condition is addressed with a method such as the penalty method or the Lagrange multiplier method. On the other hand, the boundary condition of a Neumann problem is natural, and does not need to be explicitly imposed in the variational formulation. In both situations, a simple uniform mesh on a rectangle containing the domain can be used; the mesh need not conform to the boundary and a mesh generator is not needed. These ideas are classical and have been extensively analysed (for example, see Babuška and Aziz (1972)). This way of imposing boundary conditions can be used in the context of meshless methods, and this approach was also mentioned in Li and Liu (2002). The boundary of the domain does come into play in the construction of the stiffness matrix, but a mesh generator is not needed. This approach was generalized and used together with the ideas in Babuška *et al.* (1994), Babuška and Melenk (1997) and Strouboulis, Babuška and Copps (2001*a*) in solving problems with very complex geometries: see, *e.g.*, Strouboulis, Copps and Babuška (2001*b*).

We finally mention a meshless method – the *generalized finite element method* (GFEM) – which attempts to achieve simultaneously the two goals of variationally formulated meshless methods. With this method we begin with a *partition of unity*. Construction of a partition of unity is a relatively simple task. It can be done by various means. One is to use a simple mesh, for example a uniform mesh, and use the associated hat functions as the partition of unity. We could also use ideas from various interpolation procedures, for instance, the Shepard method. It is essential that the construction can, but is not required to, utilize the geometry of the domain. The partition of unity on the domain is obtained by restriction. The partition of unity functions typically have compact supports with small diameters.

Then we multiply the partition of unity functions by functions that are defined separately and independently on the supports of the partition of

unity functions. In this way we create shape functions that belong to  $H^1(\Omega)$ , and can be used in the variational method. We thereby obtain a large flexibility in the construction of shape functions, and the associated trial spaces. This flexibility can be used to construct approximations that utilize the available information, the character of a singularity, or a boundary layer, *e.g.*, on the approximated function (solution). Hence the method achieves the goals mentioned above.

We do face three serious difficulties in the implementation of the GFEM. First there is the problem of numerical integrations when the areas over which we integrate are not simple triangles, simplices, *etc.*, as with the usual FEM. We note, however, that the process is completely parallelizable. A second difficulty is the treatment of essential boundary conditions. The third issue concerns the system of linear equations. It may be singular, and thus certain classical methods, such as multigrid, may not be applicable. These difficulties can be, and have been, overcome in some implementations, so it is clear that the GFEM shows a definite advantage over the classical FEM in certain situations. We mention problems with complex geometry, crack propagation, and analysis of multi-site local damage.

Of course, any new method should be compared with previously developed methods, and the class of problems for which the new method is superior should be identified. Theoretical and practical experience (see Babuška *et al.* (2002*a*), Li and Liu (2002) and Strouboulis *et al.* (2001*b*)) is progressing in this direction. Meshless methods in various forms, *e.g.*, within the framework of collocation or variational methods, are now the subject of many papers and (engineering) books, which mainly focus on practical aspects without serious theoretical analysis.

This paper focuses on ideas and theoretical results. Some are adjustments of old ideas and results. Some results are based on papers that are submitted or in the final stage of preparation. Although we focus on the theory, we have attempted to address theoretical issues that illuminate practical issues. We will show that the results presented here are natural generalizations of the classical FEM, which is a special case of some of the methods presented here. This paper addresses only problems related to linear PDEs.

Various relevant and typical references are provided. The reference list is not comprehensive, but together with the citations in the references provide, in our opinion, a very reasonable description of the current state of the art in meshless methods.

### 1.3. *The scope of this paper*

The short Section 2 defines the model problem, a linear elliptic boundary value problem. Section 3.1 presents approximation results when the particles are uniformly distributed. The presented results were obtained

using the Fourier transform. Section 3.2 presents an alternative proof of the approximation results that can be generalized to the case of non-uniformly distributed particles. Section 3.3 discusses approximation for arbitrarily distributed particles. Section 4 discusses the construction of shape functions, and presents some results on interpolation and on the asymptotic form of the error. Section 5 addresses the question of superconvergence. Section 6 discusses the generalized finite element method. Section 7 discusses the application of the approximation results developed in Section 3, and discusses the treatment of Dirichlet boundary conditions. Section 8 explains some implementational aspects. Section 9 reports some numerical examples obtained by the GFEM, when the domain is very complex. Finally, Section 10 presents additional results and challenges.

## 2. The model problem

For concreteness and simplicity we will address the weak solution of the model problem

$$-\Delta u + u = f(x), \quad \text{on } \Omega \subset \mathbb{R}^n \quad (2.1)$$

and

$$\frac{\partial u}{\partial n} = 0, \quad \text{on } \partial\Omega, \quad (2.2)$$

or

$$u = 0, \quad \text{on } \partial\Omega, \quad (2.3)$$

where  $f \in L_2(\Omega)$  is given. We will assume that  $\Omega$  is a Lipschitz domain; additional assumptions on  $\partial\Omega$  will be given as needed.

The weak solution  $u_0 \in H^1(\Omega)$  ( $H_0^1(\Omega)$ , respectively) satisfies

$$B(u_0, v) = \mathcal{F}(v), \quad \text{for all } v \in H^1(\Omega) \quad (v \in H_0^1(\Omega), \text{ respectively}), \quad (2.4)$$

where

$$B(u, v) \equiv \int_{\Omega} (\nabla u \cdot \nabla v + uv) \, dx \quad \text{and} \quad \mathcal{F}(v) \equiv \int_{\Omega} f v \, dx. \quad (2.5)$$

The energy norm of  $u_0$  is defined by

$$\|u_0\|_E \equiv B(u_0, u_0)^{1/2} = \|u_0\|_{H^1(\Omega)}. \quad (2.6)$$

We will write  $H$  instead of  $H^1(\Omega)$  or  $H_0^1(\Omega)$  if no misunderstanding can occur.

Let  $S \subset H$  be a finite-dimensional subspace, called the approximation space. Then the Galerkin approximation,  $u_S \in S$ , to  $u_0$  is determined by

$$\tilde{B}(u_S, v) = \mathcal{F}(v), \quad \text{for all } v \in S, \quad (2.7)$$



where  $\tilde{B}$  is either  $B$  or a perturbation of  $B$ . If  $\tilde{B} = B$ , it is immediate that

$$\|u_0 - u_S\|_{H^1(\Omega)} = \inf_{\chi \in S} \|u_0 - \chi\|_{H^1(\Omega)}. \tag{2.8}$$

Hence, the main problem is the approximation of  $u_0$  by functions in  $S$ .

**Remark 1.** The *finite element method* (FEM) is the Galerkin method where  $S$  is the span of functions with small supports. For the history of the FEM, see Babuška (1994) and the references therein.

**Remark 2.** The classical Ritz method uses spaces of polynomials on the entire domain  $\Omega$  for the approximation spaces: see, *e.g.*, Mikhlin (1971).

As mentioned above, the finite element method uses basis functions with small supports, for example, ‘hill’ functions. The theory of approximation with general hill functions with translation-invariant supports was developed in Babuška (1970) using the Fourier transform. The results in Babuška (1970) were applied to the numerical solution of PDEs in Babuška (1971). A very similar theory, also based on the Fourier transform, was later developed in Strang (1971) and Strang and Fix (1973); see also Li and Liu (1996). Later, hill functions were, in another context, called *particle functions* (see Gingold and Monaghan (1977)). In the 1990s, hill functions began to be used in the framework of *meshless methods*. For a broad survey of meshless methods see Li and Liu (2002). A survey of the approximation properties of radial hill functions, previously referred to as radial basis functions, is given in Buhmann (2000).

In this paper we will survey basic meshless approximation results and their use in the framework of Galerkin methods.

### 3. Approximation by local functions in $\mathbb{R}^n$ : the $h$ -version analysis

As mentioned in Section 2, we are interested in the approximation of functions by particle shape functions. We first consider uniformly distributed particles, and then general (non-uniformly distributed) particles.

#### 3.1. Uniformly distributed particles and associated particle shape functions

Let  $\mathbb{Z}$  be the integer lattice, and, for  $j = (j_1, \dots, j_n) \in \mathbb{Z}^n$  and  $0 < h \leq 1$ , let

$$x_j^h = (j_1 h, \dots, j_n h) = h j.$$

The points  $x_j^h$  are called *uniformly distributed particles*. When considering such families of particles, we often construct associated shape functions, called particle shape functions, as follows. Let  $\phi \in H^q(\mathbb{R}^n)$ , for some  $0 \leq q$ ,

be a function with compact support; let  $\eta \equiv \text{supp } \phi$ , and suppose

$$\eta \subset B_\rho = \{x \in \mathbb{R}^n : \|x\|^2 = x_1^2 + \cdots + x_n^2 < \rho\}.$$

We assume that  $0 \in \mathring{\eta}$  ( $\mathring{\eta}$  is the interior of  $\eta$ ). Then define

$$\phi_j^h(x) = \phi_j^h(x_1, \dots, x_n) = \phi\left(\frac{x-jh}{h}\right) = \phi\left(\frac{x_1-j_1h}{h}, \dots, \frac{x_n-j_nh}{h}\right), \quad (3.1)$$

for  $j \in \mathbb{Z}^n$  and  $0 < h \leq 1$ . Clearly,

$$\eta_j^h \equiv \text{supp } \phi_j^h = \left\{x : \frac{x-jh}{h} \in \eta\right\} \subset B_{\rho h}^j = \{x : \|x-x_j^h\| < \rho h\},$$

and  $x_j^h \in \mathring{\eta}_j^h$ . Particles and particle shape functions defined in this way will be called *translation-invariant*, since they satisfy

$$x_{j+l}^h = x_j^h + x_l^h \text{ and } \phi_{j+l}^h(x) = \phi_j^h(x-x_l^h).$$

They are a special case of general (non-uniformly distributed) particles, which will be introduced in Section 3.3. We refer to  $h$  as the size of the particle, and the function  $\phi$  is called the *basic shape function*. In this section we will be interested in the approximation properties of

$$V_h^{k,q} \equiv \left\{v = v(x) = \sum_{j \in \mathbb{Z}^n} w_j^h \phi_j^h(x) : w_j^h \in \mathbb{R}\right\}, \quad (3.2)$$

which is the linear span of the associated shape functions, as  $h \rightarrow 0$ . The parameter  $k$  in  $V_h^{k,q}$  is related to a property of the  $\{\phi_j^h(x)\}$ , which will be discussed later. We will refer to  $V_h^{k,q}$  as the particle space in  $\mathbb{R}^n$ . The  $\{w_j^h\}$  are called *weights*. Specifically, given  $u \in H^{k+1}(\mathbb{R}^n)$ , we are interested in estimating

$$\inf_{\chi \in V_h^{k,q}} \|u - \chi\|_{H^s(\mathbb{R}^n)}, \quad (3.3)$$

for  $0 \leq s \leq \min\{q, k+1\}$ . We are especially interested in the maximum  $\mu$  such that

$$\inf_{\chi \in V_h^{k,q}} \|u - \chi\|_{H^s(\mathbb{R}^n)} \leq C(k, q) h^\mu \|u\|_{H^{k+1}(\mathbb{R}^n)}, \quad (3.4)$$

for  $0 \leq s \leq \min\{q, k+1\}$ , where the constant  $C = C(k, q)$  depends on  $k, q$ , but is independent of  $h$  ( $C$  also depends on  $\phi$ ).

Because we are assuming the particles are uniformly distributed, and hence the particles and shape functions are translation-invariant, estimates of the form (3.4) can be obtained via the Fourier transform. This was done in Babuška (1970), Strang (1971) and Strang and Fix (1973). We will cite one of the results in the last paper.

Let

$$\hat{\phi}(\xi) = \hat{\phi}(\xi_1, \dots, \xi_n) \equiv \int_{\mathbb{R}^n} \phi(x) e^{-ix \cdot \xi} dx$$

denote the Fourier transform of  $\phi(x)$ . We note that  $\hat{\phi}(\xi) \in C^\infty(\mathbb{R}^n)$  since  $\phi(x)$  has compact support. We use the usual multi-index notation:  $\alpha = (\alpha_1, \dots, \alpha_n)$ , with  $\alpha_i \geq 0$ ;  $|\alpha| = \alpha_1 + \dots + \alpha_n$ ;  $x^\alpha = x_1^{\alpha_1} \dots x_n^{\alpha_n}$ ; and

$$D^\alpha \hat{\phi} = \frac{\partial^{|\alpha|} \hat{\phi}}{\partial \xi_1^{\alpha_1} \dots \partial \xi_n^{\alpha_n}}.$$

**Theorem 3.1. (Strang and Fix 1973)** Suppose  $\phi \in H^q(\mathbb{R}^n)$  has compact support, where the smoothness index  $q \geq 0$  is an integer. Then the following three conditions are equivalent:

(1) 
$$\hat{\phi}(0) \neq 0 \tag{3.5}$$

and

$$D^\alpha \hat{\phi}(2\pi j) = 0, \quad \text{for } 0 \neq j \in \mathbb{Z}^n \text{ and } |\alpha| \leq k, \tag{3.6}$$

where  $k$  is a nonnegative integer.

(2) For  $|\alpha| \leq k$ ,

$$\sum_{j \in \mathbb{Z}^n} j^\alpha \phi(x - j) = \lambda x^\alpha + q^\alpha(x), \quad \text{for all } x \in \mathbb{R}^n, \tag{3.7}$$

where  $\lambda \neq 0$  and  $q^\alpha(x)$  is a polynomial with degree  $< |\alpha|$ .

The equality in (3.7) is for almost all  $x \in \mathbb{R}^n$ . The function of the right-hand side of (3.7) is, of course, continuous. If the function on the left-hand side is continuous, which will be the case if  $q > n/2$ , then (3.7) will hold for all  $x \in \mathbb{R}^n$ .

(3) For each  $u \in H^{k+1}(\mathbb{R}^n)$  there are weights  $w_j^h \in \mathbb{R}$ , for  $j \in \mathbb{Z}^n$  and  $0 < h$ , such that

$$\begin{aligned} & \left\| u - \sum_{j \in \mathbb{Z}^n} w_j^h \phi_j^h \right\|_{H^s(\mathbb{R}^n)} \\ & \leq Ch^{k+1-s} \|u\|_{H^{k+1}(\mathbb{R}^n)}, \quad \text{for } 0 \leq s \leq \min\{q, k+1\}, \end{aligned} \tag{3.8}$$

and

$$h^n \sum_{j \in \mathbb{Z}^n} (w_j^h)^2 \leq K^2 \|u\|_{H^0(\mathbb{R}^n)}^2. \tag{3.9}$$

Here  $C$  and  $K$  may depend on  $q, k$ , and  $s$ , but are independent of  $u$  and  $h$ . The exponent  $k + 1 - s$  is the best possible if  $k$  is the largest integer for which (3.7) holds.

If (3.7) holds, the basic shape function  $\phi$  is called *quasi-reproducing of order  $k$* . If (3.7) holds with  $\lambda = 1$  and  $q^\alpha(x) = 0$ ,  $\phi$  is called *reproducing of order  $k$* . If  $\phi$  is quasi-reproducing of order  $k$  (respectively, reproducing of order  $k$ ), then the corresponding particle shape functions  $\phi_i^h$  are also called quasi-reproducing of order  $k$  (respectively, reproducing of order  $k$ ). The parameter  $k$  in  $V_h^{k,q}$ , defined in (3.2), is the quasi-reproducing order of the basic shape function  $\phi$ .

**Remark 3.** If one were to define the notion of quasi-reproducing basic shape function  $\phi$  of order  $k$  by the formula

$$\sum_{j \in \mathbb{Z}^n} j^\alpha \phi(x - j) = \lambda_\alpha x^\alpha + q^\alpha(x), \quad \text{for all } x \in \mathbb{R}^n, \quad \text{for } |\alpha| \leq k, \quad (3.10)$$

where  $\lambda_\alpha \neq 0$ , it might appear that this would lead to a larger class of functions. This, however, is not the case; it is easily shown that if  $\phi$  satisfies (3.10), then  $\lambda_\alpha = \lambda$ , for  $|\alpha| \leq k$ .

In one dimension we can prove more.

**Theorem 3.2. (Strang and Fix 1973)** Suppose  $\phi \in H^q(\mathbb{R})$  (in one dimension) has compact support and satisfies condition (1) of Theorem 3.1, that is, it satisfies (3.5) and (3.6). Then

$$\hat{\phi}(\xi) = Z(\xi) \left( \frac{\sin(\xi/2)}{\xi/2} \right)^{k+1}, \quad (3.11)$$

where  $Z(\xi)$  is an entire function.

*Proof.* Because  $\phi$  has compact support,  $\hat{\phi}(\xi)$  is an entire function and, because of (3.5) and (3.6),  $\hat{\phi}(0) \neq 0$ , and  $\hat{\phi}(\xi)$  has zeros of at least order  $k$  at  $2\pi j$ ,  $0 \neq j \in \mathbb{Z}$ . Let

$$\hat{\sigma}_k(\xi) = \left( \frac{\sin(\xi/2)}{\xi/2} \right)^{k+1}. \quad (3.12)$$

The function  $\hat{\sigma}_k(\xi)$  is entire with only zeros of order  $k + 1$  at  $2\pi j$ , for  $0 \neq j \in \mathbb{Z}$ . Hence

$$Z(\xi) = \hat{\phi}(\xi) / \hat{\sigma}_k(\xi)$$

is entire, and

$$\hat{\phi}(\xi) = Z(\xi) \left( \frac{\sin(\xi/2)}{\xi/2} \right)^{k+1}, \quad (3.13)$$

as desired.  $\square$

**Theorem 3.3. (Babuška 1970)** Suppose  $\phi \in H^q(\mathbb{R})$  (in one dimension) has compact support and satisfies condition (1) of Theorem 3.1, that is,

it satisfies (3.5) and (3.6). Then, for any  $\epsilon > 0$ ,

$$\text{supp } \phi \not\subset \left[ -\frac{(k+1)}{2} + \epsilon, \frac{(k+1)}{2} - \epsilon \right]. \tag{3.14}$$

*Proof.* Suppose, on the contrary, that

$$\text{supp } \phi \subset [-(k+1)/2 + \epsilon, (k+1)/2 - \epsilon], \quad \text{for some } \epsilon > 0. \tag{3.15}$$

We will show that this assumption leads to a contradiction.

The function  $\hat{\phi}(\xi)$  is entire and, with  $\xi = \xi_1 + i\xi_2$ , (3.15) implies

$$|\hat{\phi}(\xi)| \leq C e^{(\frac{k+1}{2} - \epsilon)|\xi_2|}. \tag{3.16}$$

This estimate follows directly from the definition of the Fourier transform and assumption (3.15). Using elementary properties of the sine function, we find that

$$\left| \left( \frac{\sin(\xi/2)}{\xi/2} \right)^{k+1} \right| \geq C \frac{e^{\frac{k+1}{2}|\xi_2|}}{|\xi_2|^{k+1}}, \quad \text{for } |\xi_2| \text{ large.} \tag{3.17}$$

Using (3.5), (3.6), (3.16), and (3.17), we have

$$|Z(\xi)| = \left| \frac{\hat{\phi}(\xi)}{\left(\frac{\sin(\xi/2)}{\xi/2}\right)^{k+1}} \right| \leq C_0 + C_{k+1}|\xi|^{k+1}, \quad \text{for all } \xi \in \mathcal{C}, \tag{3.18}$$

where  $Z(\xi)$  is as in (3.11). Since  $Z(\xi)$  is entire, estimate (3.18) implies (via a generalization of Liouville’s theorem for entire functions) that  $Z(\xi)$  is a polynomial of degree  $\leq k + 1$ . Next, we use (3.11) and (3.16) to get

$$\left| Z(\xi) \left( \frac{\sin(\xi/2)}{\xi/2} \right)^{k+1} \right| = |\hat{\phi}(\xi)| \leq C e^{(\frac{k+1}{2} - \epsilon)|\xi_2|}. \tag{3.19}$$

Combining this estimate with the lower bound in (3.17), we have

$$|Z(\xi)| \leq C |\xi|^{k+1} e^{-\epsilon|\xi_2|}, \quad \text{for } |\xi_2| \text{ large.} \tag{3.20}$$

This implies  $Z(\xi) = 0$ . Thus, (3.11) implies  $\hat{\phi}(\xi) = 0$ , which contradicts (3.5). Thus (3.15) is false, which proves (3.14).  $\square$

The case of uniformly distributed particles is very special, but we have considered it, and cited Theorem 3.1 from Strang and Fix (1973) because the result provides necessary and sufficient conditions on the basic shape function  $\phi$  for the validity of the approximability result (3.8) and (3.9), leading to the optimal value for  $\mu$  in (3.4).

**Remark 4.** Condition (2) of Theorem 3.1 implies that

$$\sum_{j \in \mathbb{Z}^n} \phi(x - j) = b. \tag{3.21}$$

Hence the functions  $\phi(x - j)/b, j \in \mathbb{Z}^n$ , form a *partition of unity*. Thus sets  $\eta_j^b$  form an open cover of  $\mathbb{R}^n$ .

**Remark 5.** Condition (2), *i.e.*, (3.7), of Theorem 3.1 is the definition of the notion of quasi-reproducing of order  $k$ , and from Theorem 3.1 we see that this notion is equivalent to condition (1), *i.e.*, to (3.5) and (3.6). It is of interest to have a condition similar to (3.5) and (3.6) that is equivalent to the related notion of reproducing of order  $k$ , which can be stated as

$$\sum_{j \in \mathbb{Z}^n} p(j)\phi(x - j) = p(x), \quad \text{for all polynomials } p(x) \text{ of degree } \leq k.$$

It can be shown that  $\phi$  is reproducing of order  $k$  if and only if

- (a)  $\hat{\phi}(0) = 1$ ,
- (b)  $D^\alpha \hat{\phi}(0) = 0$ , for  $1 \leq |\alpha| \leq k$ ,
- (c)  $D^\alpha \hat{\phi}(2\pi j) = 0$ , for  $0 \neq j \in \mathbb{Z}^n$  and  $|\alpha| \leq k$ .

The proof follows the argument of the proof of Theorem 3.1 in Strang and Fix (1973).

**Remark 6.** The B-spline of order  $k$ , denoted by  $\sigma_k(x)$ , is the  $(k + 1)$ -fold convolution of the characteristic function of the cube  $(-1/2, 1/2)^n$ . The support of  $\sigma_k(x)$  is the closed cube  $[-(k + 1)/2, (k + 1)/2]^n$ , and the Fourier transform of  $\sigma_k(x)$  is given by

$$\hat{\sigma}_k(\xi) = \prod_{i=1}^n \left( \frac{\sin \xi_i/2}{\xi_i/2} \right)^{k+1}.$$

Here  $\hat{\sigma}_k(\xi)$  satisfies (3.5) and (3.6), and thus  $\sigma_k(x)$  is quasi-reproducing of order  $k$ . We note, however, that  $\sigma_k(x)$  is not reproducing of order  $k$  for  $k > 1$ , since  $\hat{\sigma}_k(\xi)$  does not satisfy condition (b) in Remark 2 for  $k > 1$ .

In one dimension ( $n = 1$ ), we can say more. If  $\phi(x)$  satisfies (3.5) and (3.6), then from Theorem 3.2,  $\hat{\phi}(\xi)$  is the product of  $\hat{\sigma}_k(\xi)$  and a suitable entire function  $Z(\xi)$ . Using the Paley–Wiener theorem (see Rudin (1991)), it can be shown that  $Z(\xi)$  is the Fourier transform of a distribution  $\psi$  with compact support:  $Z(\xi) = \hat{\psi}(\xi)$ . We can thus express (3.11) as

$$\hat{\phi}(\xi) = \hat{\psi}(\xi)\hat{\sigma}_k(\xi).$$

Thus any  $\phi(x)$  that satisfies (3.5) and (3.6), which may or may not be piecewise polynomial, can be constructed via the convolution of the B-spline with a distribution with compact support. If  $n > 1$ , no such divisor  $\hat{\phi}/\hat{\sigma}_k$  exists in general.

**Remark 7.** Theorem 3.3 has an especially simple interpretation for  $\phi$  satisfying (3.5) and (3.6), and whose support is a symmetric interval about 0;

namely,  $\text{supp } \phi \supseteq [-\frac{k+1}{2}, \frac{k+1}{2}]$ , and hence grows with  $k$ . As mentioned above, the support of  $\sigma_k(x)$  is  $[-\frac{k+1}{2}, \frac{k+1}{2}]$ , and hence  $\sigma_k(x)$  has minimal support.

**Remark 8.** The size of the support  $\eta$  of  $\phi$ , *i.e.*, the diameter of the smallest closed ball containing  $\eta$ , plays an important role when particle shape functions are used in a Galerkin method to approximate the solution of a variationally posed boundary value problem. The Galerkin method leads to a linear algebraic system, where the (stiffness) matrix is banded, and it is well known that such linear systems can be solved efficiently by the elimination method when the bandwidth is small. The size of the bandwidth is directly proportional to the size of  $\eta$ . Thus it is desirable to use a  $\phi$  whose support is as small as possible, of course, without sacrificing the accuracy of the computed solution of the differential equation. In one dimension, as mentioned above, the basic shape function  $\sigma_k(x)$  has the minimal support  $[-(k + 1)/2, (k + 1)/2]$ , which increases with  $k$ .

**Remark 9.** Suppose  $\phi(x) \in H^q(\mathbb{R}^n)$  has compact support and is quasi-reproducing of order  $k$ . For  $s \leq q$ , let

$$\phi^*(x) = \left( 1 + \sum_{|\beta| \leq s} \alpha_\beta D^\beta \right) \phi(x).$$

Then  $\phi^*(x) \in H^{q-s}(\mathbb{R}^n)$ ,  $\phi^*$  is quasi-reproducing of order  $k$ , and  $\text{supp } \phi^* = \text{supp } \phi$ .

**Remark 10.** Suppose  $\phi \in H^q(\mathbb{R}^n)$  is quasi-reproducing of order  $k$ . It is possible to construct another shape function  $\phi^*$ , in terms of  $\phi$ , that is reproducing of order  $k$ , is in  $H^q(\mathbb{R}^n)$ , but will have a larger support than  $\phi$ . For example,  $\phi^*$  can be constructed as a linear combination of translates of  $\phi$ .

**Remark 11.** For the sake of simplicity, suppose  $k = 2$  or  $3$  in this remark. Consider the function  $\phi(x), x \in \mathbb{R}$ , whose Fourier transform is given by

$$\hat{\phi}_k(\xi) = \hat{\sigma}_k(\xi) \left( 1 - \frac{\hat{\sigma}_k''(0)}{2} \xi^2 \right), \tag{3.22}$$

where  $\sigma_k(x)$  is the B-spline of order  $k$  (recall  $\sigma_k(x)$  is quasi-reproducing of order  $k$  and its support is  $[-(k + 1)/2, (k + 1)/2]$ ). Here  $\hat{\phi}_k(\xi)$  satisfies conditions (a), (b), (c) of Remark 2, and hence  $\phi_k(x)$  is reproducing of order  $k$ . From (3.22) it is clear that

$$\phi_k(x) = \left[ \sigma_k * \left( \delta_0 - \frac{\hat{\sigma}_k''(0)}{2} \delta_0'' \right) \right](x),$$

where  $\delta_0$  is the Dirac distribution, and hence  $\text{supp } \phi_k = \text{supp } \sigma_k$ . From the

expression for  $\hat{\sigma}_k(\xi)$ , we easily see that  $\sigma_k \in H^{k+1/2-\epsilon}(\mathbb{R})$ . Using the expression for  $\hat{\phi}_k(\xi)$  (i.e., (3.22)), we get  $\phi_k \in H^{k-3/2-\epsilon}(\mathbb{R})$ , but  $\phi_k \notin H^{k-1}(\mathbb{R})$ . Thus, for  $k = 2$ , we have  $\phi_2 \in H^{1/2-\epsilon}(\mathbb{R})$ , but  $\phi_2 \notin H^1(\mathbb{R})$ . And for  $k = 3$ , we see that  $\phi_3 \in H^{3/2-\epsilon}(\mathbb{R})$ , and hence  $\phi_3 \in H^1(\mathbb{R})$ . For approximating solutions of second-order differential equations, it is important that the shape functions are in  $H^1$ . This dichotomy between the cases  $k = 2$  (even) and  $k = 3$  (odd), in fact, holds for all  $k$ . Thus, for  $k$  even, there is no  $\phi_k \in H^1(\mathbb{R})$  that is reproducing of order  $k$  with support  $[-\frac{k+1}{2}, \frac{k+1}{2}]$ . This latter result can be seen as follows. From Remark 3 we know that  $\phi_k$  is the convolution of  $\sigma_k$  and a distribution with compact support. If we require  $\text{supp } \phi_k = \text{supp } \sigma_k$ , then this distribution must be supported at the origin, and hence its Fourier transform must be a polynomial. We have examined the case when the polynomial is  $1 - \frac{\hat{\sigma}_k''(0)}{2}\xi^2$ ; the general situation is similar.

**Remark 12.** The weights in (3.8) depend on  $u$ , but they are not unique. We note that the functions  $\phi_j^h$  may be linearly dependent. Taking  $q = k + 1$  allows application of Theorem 3.1 for all  $s \leq k + 1$ . Taking  $q > k + 1$ , i.e., assuming extra smoothness on the particle shape functions, does not change the estimate. The approximability of the classical finite element shape functions (the hat functions) can be analysed with Theorem 3.1 with  $q = k = 1$ .

**Remark 13.** The space  $V_h^{k,q}$  is a  $S^{t,k^*}$ -regular system (this notion will be introduced in Section 3.2), with  $k^* = q$  and  $t = k + 1$ .  $S^{t,k^*}$ -regular systems are analysed in Babuška and Aziz (1972). They have many important properties, some of which will be used in the following sections.

### 3.2. Alternative proof for uniformly distributed particles and particle shape functions

In this section we first give an alternative proof that condition (2) of Theorem 3.1 implies estimate (3.8), again for uniformly distributed particles and associated shape functions. This alternative proof does not use the Fourier transform, and it can be naturally generalized to non-uniformly distributed particles.

We review our notation before stating the theorem. Recall that

$$x_j^h = jh, \quad \text{for } j \in \mathbb{Z}^n \text{ and } 0 < h,$$

are the particles, and  $\phi \in H^q(\mathbb{R}^n)$ , with  $q \geq 0$ , is the basic shape function. Also  $\eta = \text{supp } \phi \subset B_\rho$ , and  $0 \in \overset{\circ}{\eta}$ . Then the particle shape functions,  $\phi_j^h(x)$ , are defined by

$$\phi_j^h(x) = \phi\left(\frac{x - jh}{h}\right);$$



it is immediate that

$$\eta_j^h = \text{supp } \phi_j^h \subset B_{\rho h}^j,$$

and  $x_j^h \in \bar{\eta}_j^h$ .

**Theorem 3.4.** Suppose  $\phi \in H^q(\mathbb{R}^n)$ , with smoothness index  $q \geq 0$ , has compact support  $\eta \subset B_\rho$ , and suppose the  $\phi_j^h(x)$  are defined in (3.1). Suppose  $k = 0, 1, 2, \dots$  and suppose, for  $|\alpha| \leq k$ ,

$$\sum_{j \in \mathbb{Z}^n} j^\alpha \phi(x - j) = \lambda x^\alpha + q^\alpha(x); \tag{3.23}$$

here  $\lambda \neq 0$ , and  $q^\alpha(x)$  is a polynomial of degree  $< |\alpha|$ , i.e., suppose  $\phi$  is quasi-reproducing of order  $k$ . Suppose  $u$  satisfies

$$\sum_{j \in \mathbb{Z}^n} \|u\|_{H^{r_j+1}(B_{\bar{\rho}h}^j)}^2 < \infty, \quad \text{where } 0 \leq r_j \leq k, \tag{3.24}$$

where  $\bar{\rho} \geq 1$  is sufficiently large and independent of  $h$ . Then there exist weights  $w_l^h$  such that

$$\left\| u - \sum_{l \in \mathbb{Z}^n} w_l^h \phi_l^h \right\|_{H^s(\mathbb{R}^n)}^2 \leq C \sum_{j \in \mathbb{Z}^n} h^{2(r_j+1-s)} \|u\|_{H^{r_j+1}(B_{\bar{\rho}h}^j)}^2, \tag{3.25}$$

for  $0 \leq s \leq \min\{q, r_j + 1\}$ ,

where  $C$  is independent of  $u$  and  $h$ . If  $u \in H^{k'+1}(\mathbb{R}^n)$ , where  $0 \leq k' \leq k$ , then

$$\left\| u - \sum_{l \in \mathbb{Z}^n} w_l^h \phi_l^h \right\|_{H^s(\mathbb{R}^n)} \leq C h^{k'+1-s} \|u\|_{H^{k'+1}(\mathbb{R}^n)}, \quad \text{for } 0 \leq s \leq \min\{q, k' + 1\}. \tag{3.26}$$

*Proof.* The proof is in several steps.

1. Suppose  $\phi$  satisfies (3.23), and write  $q^\alpha(x) = \sum_{|\gamma| \leq |\alpha|-1} d_{\gamma\alpha} x^\gamma$ . Then

$$\begin{aligned} \sum_{j \in \mathbb{Z}^n} (x_j^h)^\alpha \phi_j^h(x) &= \sum_{j \in \mathbb{Z}^n} (jh)^\alpha \phi_j^h(x) \\ &= h^{|\alpha|} \sum_{j \in \mathbb{Z}^n} j^\alpha \phi\left(\frac{x}{h} - j\right) \\ &= h^{|\alpha|} \left\{ \lambda \left(\frac{x}{h}\right)^\alpha + q^\alpha\left(\frac{x}{h}\right) \right\} \end{aligned}$$

$$\begin{aligned}
&= \lambda x^\alpha + h^{|\alpha|} \sum_{|\gamma| \leq |\alpha| - 1} d_{\gamma\alpha} \left(\frac{x}{h}\right)^\gamma \\
&= \lambda x^\alpha + \sum_{|\gamma| \leq |\alpha| - 1} h^{|\alpha| - |\gamma|} d_{\gamma\alpha} x^\gamma, \quad \text{for } |\alpha| \leq k. \quad (3.27)
\end{aligned}$$

Equations (3.23) and (3.27) are in fact equivalent: (3.27) could be viewed as a scaled version of (3.23). For any  $p \in \mathcal{P}^k$ , there is a uniquely determined  $w = w_{p,h} \in \mathcal{P}^k$  satisfying

$$p(x) = \sum_{j \in \mathbb{Z}^n} w_{p,h}(x_j^h) \phi_j^h(x), \quad \text{for all } x \in \mathbb{R}^n. \quad (3.28)$$

We first prove the existence of  $w_{p,h}$ , and begin by considering the monic polynomials  $p_\alpha = x^\alpha$ . Suppose  $|\alpha| = 0$ . Then from (3.27) we have

$$1 = \sum_{j \in \mathbb{Z}^n} \frac{1}{\lambda} \phi_j^h(x) = \sum_{j \in \mathbb{Z}^n} w_{\{1\},h}(x_j^h) \phi_j^h(x),$$

where  $w_{\{1\},h}(x) = 1/\lambda$ . Next suppose  $|\alpha| = 1$ . Using (3.27) again we have

$$x^\alpha = \sum_{j \in \mathbb{Z}^n} \frac{1}{\lambda} (x_j^h)^\alpha \phi_j^h(x) - \frac{h d_{0\alpha}}{\lambda} = \sum_{j \in \mathbb{Z}^n} w_{\{x^\alpha\},h}(x_j^h) \phi_j^h(x)$$

where

$$w_{\{x^\alpha\},h}(x) = \frac{x^\alpha}{\lambda} - \frac{h d_{0\alpha}}{\lambda^2}.$$

Proceeding in this way, by induction, we get  $w_{\{x^\alpha\},h}(x)$  for  $|\alpha| \leq k$ , where  $w_{\{x^\alpha\},h}(x)$  is of the form

$$w_{\{x^\alpha\},h}(x) = e_{\alpha\alpha} x^\alpha + \sum_{|\beta| \leq |\alpha| - 1} e_{\alpha\beta} h^{|\alpha| - |\beta|} x^\beta, \quad (3.29)$$

where  $e_{\alpha\alpha} = \lambda^{-1}$  and  $e_{\alpha\beta}$  are expressions in  $d_{\gamma\alpha}$ ,  $|\gamma| < |\alpha|$ . For  $p(x) = \sum_{|\alpha| \leq k} c_\alpha x^\alpha$ , we let  $w_{p,h}(x) = \sum_{|\alpha| \leq k} c_\alpha w_{\{x^\alpha\},h}(x)$ . It is immediate that

$$p(x) = \sum_{j \in \mathbb{Z}^n} w_{p,h}(x_j^h) \phi_j^h(x),$$

which establishes the existence of  $w_{p,h}(x)$ . We can show that

$$w_{p,h}(x) = \sum_{|\beta| \leq k} \left[ \sum_{|\beta| + 1 \leq |\alpha| \leq k, \alpha = \beta} c_\alpha d_{\alpha\beta} h^{|\alpha| - |\beta|} \right] x^\beta. \quad (3.30)$$

To prove uniqueness, suppose  $w_{p,h}(x) = 0$ . We will show that  $p(x) = \sum_{|\alpha| \leq k} c_\alpha x^\alpha = 0$ . Since  $w_{p,h}(x) = 0$ , it is clear from (3.30) that the coefficient of  $x^\beta$  is zero for  $|\beta| \leq k$ , from which we can deduce that  $c_\alpha = 0$ ,  $|\alpha| \leq k$ ,

and thus  $p(x) = 0$ . It will be convenient to write  $w_{p,h}(x) = \mathcal{A}^h p$ . Then  $\mathcal{A}^h : \mathcal{P}^k \rightarrow \mathcal{P}^k$  is a bijection satisfying

$$p(x) = \sum_{j \in \mathbb{Z}^n} (\mathcal{A}^h p)(x_j^h) \phi_j^h(x), \quad \text{for all } x \in \mathbb{R}^n, \quad \text{for any } p \in \mathcal{P}^k. \quad (3.31)$$

We define  $\mathcal{A} = \mathcal{A}^h$  when  $h = 1$ . We note that  $\mathcal{A}$  satisfies (3.31) with  $h = 1$ . We also have

$$[(\mathcal{A}^h)^{-1} w](x) = \sum_{j \in \mathbb{Z}^n} w(x_j^h) \phi_j^h, \quad \text{for all } x \in \mathbb{R}^n, \quad \text{for any } w \in \mathcal{P}^k. \quad (3.32)$$

It is also clear from the construction that  $\mathcal{A}^h : \mathcal{P}^i \rightarrow \mathcal{P}^i$ , for  $i \leq k$ .

2. Define the cells  $\omega_j$  and  $\omega_j^h$ :

$$\omega_j = \left\{ x : \|x - j\|_\infty \equiv \max_{i=1,\dots,n} |x_i - j_i| < \rho \right\}$$

and

$$\omega_j^h = \left\{ x : \|x - x_j^h\|_\infty \equiv \max_{i=1,\dots,n} |x_i - x_{j_i}^h| < \rho h \right\}.$$

The families  $\{\omega_j\}_{j \in \mathbb{Z}^n}$  and  $\{\omega_j^h\}_{j \in \mathbb{Z}^n}$  are open covers of  $\mathbb{R}^n$  provided  $\rho > 1/2$ . Let

$$A_j^h = \{l \in \mathbb{Z}^n : \eta_l^h \cap \omega_j^h \neq \emptyset\},$$

and define

$$\Omega_j^h = \cup_{l \in A_j^h} \omega_l^h.$$

It is immediate that one can select  $M$  and  $\bar{\rho}$  such that

$$\text{card } A_j^h \leq M \quad (3.33)$$

and

$$\Omega_j^h \subset B_{\bar{\rho}h}^j. \quad (3.34)$$

The constants  $M$  and  $\bar{\rho}$  are independent of  $j$  and  $h$ , but do depend on  $\phi$ ; specifically on  $\rho$ .

For any  $l \in \mathbb{Z}^n$ , since  $u \in H^{r_l+1}(B_{\bar{\rho}h}^l)$ , it is well known (Bramble and Hilbert 1970, Bramble and Hilbert 1971, Ciarlet 1980) that there is a polynomial  $p^{l,h} = p_k^{l,h}$  of degree  $\leq k$  such that

$$\|u - p^{l,h}\|_{H^s(B_{\bar{\rho}h}^l)} \leq Ch^{r_l+1-s} \|u\|_{H^{r_l+1}(B_{\bar{\rho}h}^l)}, \quad \text{for } 0 \leq s \leq r_l + 1 \leq k + 1, \quad (3.35)$$

where  $C$  is independent of  $u$ ,  $h$ , and  $l$ , but does depend on  $k$  ( $p^{l,h}$  can, in fact, be chosen such that its degree  $\leq r_l$ ). Define the weights

$$w_l^h = (\mathcal{A}^h p^{l,h})(x_l^h). \quad (3.36)$$

Let  $j$  be fixed. We will work with the polynomial  $p^{j,h}$ , which satisfies (3.35) with  $l = j$ , as well as the polynomial  $p^{l,h}$ . Using (3.36), we find

$$\begin{aligned} & \left\| u - \sum_{l \in \mathbb{Z}^n} w_l^h \phi_l^h \right\|_{H^s(\omega_j^h)} \\ & \leq \left\| u - \sum_{l \in A_j^h} w_l^h \phi_l^h \right\|_{H^s(\omega_j^h)} \\ & \leq \left\| u - \sum_{l \in A_j^h} (\mathcal{A}^h p^{j,h})(x_l^h) \phi_l^h \right\|_{H^s(\omega_j^h)} \\ & \quad + \sum_{l \in A_j^h} |(\mathcal{A}^h p^{j,h})(x_l^h) - (\mathcal{A}^h p^{l,h})(x_l^h)| \|\phi_l^h\|_{H^s(\omega_j^h)}. \end{aligned} \tag{3.37}$$

We now estimate the two terms on the right side of (3.37).

**3.** From (3.31) and the definition of  $A_j^h$ , we have

$$p(x) = \sum_{l \in \mathbb{Z}^n} (\mathcal{A}^h p)(x_l^h) \phi_l^h(x) = \sum_{l \in A_j^h} (\mathcal{A}^h p)(x_l^h) \phi_l^h(x), \quad \text{for } x \in \omega_j^h,$$

for any  $p \in \mathcal{P}^k$ . Using this formula and (3.35) with  $l = j$ , we obtain the estimate

$$\begin{aligned} \left\| u - \sum_{l \in A_j^h} (\mathcal{A}^h p^{j,h})(x_l^h) \phi_l^h \right\|_{H^s(\omega_j^h)} &= \|u - p^{j,h}\|_{H^s(\omega_j^h)} \\ &\leq Ch^{r_j+1-s} \|u\|_{H^{r_j+1}(B_{\bar{\rho}h}^j)}, \end{aligned} \tag{3.38}$$

for the first term of (3.37).

A scaling argument shows that

$$\|\phi_l^h\|_{H^s(\omega_j^h)} \leq h^{-s+n/2} \|\phi\|_{H^s(\mathbb{R}^n)}.$$

Thus

$$\begin{aligned} & \sum_{l \in A_j^h} |(\mathcal{A}^h p^{j,h})(x_l^h) - (\mathcal{A}^h p^{l,h})(x_l^h)| \|\phi_l^h\|_{H^s(\omega_j^h)} \\ & \leq Ch^{-s+n/2} \sum_{l \in A_j^h} |(\mathcal{A}^h p^{j,h})(x_l^h) - (\mathcal{A}^h p^{l,h})(x_l^h)|. \end{aligned} \tag{3.39}$$

It remains to estimate the right-hand side of this inequality.

For  $l \in A_j^h$ ,  $\omega_l^h \subset \Omega_j^h$ , and hence, using (3.34),  $\omega_l^h \subset B_{\bar{\rho}h}^j$ . Also  $\omega_l^h \subset B_{\bar{\rho}h}^l$ .

Thus, using (3.35) with  $s = 0$ , we have

$$\begin{aligned} \|p^{j,h} - p^{l,h}\|_{H^0(\omega_l^h)} &\leq \|p^{j,h} - u\|_{H^0(\omega_l^h)} + \|u - p^{l,h}\|_{H^0(\omega_l^h)} \\ &\leq \|p^{j,h} - u\|_{H^0(B_{\rho h}^j)} + \|u - p^{l,h}\|_{H^0(B_{\rho h}^l)} \\ &\leq Ch^{r_j+1}\|u\|_{H^{r_j+1}(B_{\rho h}^j)} \\ &\quad + Ch^{r_l+1}\|u\|_{H^{r_l+1}(B_{\rho h}^l)}. \end{aligned} \tag{3.40}$$

It is easily shown that there is a constant  $C$  such that

$$\|w\|_{L^\infty(\omega_l^h)} \leq Ch^{-n/2}\|w\|_{H^0(\omega_l^h)}, \quad \text{for any } w \in \mathcal{P}^k; \tag{3.41}$$

$C$  is independent of  $w$ ,  $h$ , and  $l$ . Applying (3.41) to  $w = \mathcal{A}^h p^{j,h} - \mathcal{A}^h p^{l,h}$ , we have

$$\begin{aligned} |(\mathcal{A}^h p^{j,h})(x_l^h) - (\mathcal{A}^h p^{l,h})(x_l^h)| \\ \leq \|\mathcal{A}^h p^{j,h} - \mathcal{A}^h p^{l,h}\|_{L^\infty(\omega_l^h)} \\ \leq Ch^{-n/2}\|\mathcal{A}^h p^{j,h} - \mathcal{A}^h p^{l,h}\|_{H^0(\omega_l^h)}. \end{aligned} \tag{3.42}$$

For any  $p \in \mathcal{P}^k$ , we write  $p(x) = \tilde{p}\left(\frac{x-x_l^h}{h}\right)$ , where  $\tilde{p} \in \mathcal{P}^k$ . Using (3.31) with  $h = 1$  (recall that  $\mathcal{A} = \mathcal{A}^h$  for  $h = 1$ ), we see that

$$\tilde{p}(x) = \sum_{i \in \mathbb{Z}^n} (\mathcal{A}\tilde{p})(i)\phi(x - i),$$

and therefore

$$\begin{aligned} p(x) &= \tilde{p}\left(\frac{x - x_l^h}{h}\right) \\ &= \sum_{i \in \mathbb{Z}^n} (\mathcal{A}\tilde{p})(i)\phi\left(\frac{x - x_{i+l}^h}{h}\right) \\ &= \sum_{i \in \mathbb{Z}^n} (\mathcal{A}\tilde{p})(i)\phi_{i+l}^h(x) \\ &= \sum_{i \in \mathbb{Z}^n} (\mathcal{A}\tilde{p})(i - l)\phi_i^h(x) \\ &= \sum_{i \in \mathbb{Z}^n} (\mathcal{A}\tilde{p})\left(\frac{x_i^h - x_l^h}{h}\right)\phi_i^h(x). \end{aligned}$$

Comparing the above expression with (3.31) and using the uniqueness of the representation (3.28), we obtain

$$(\mathcal{A}^h p)(x) = (\mathcal{A}\tilde{p})\left(\frac{x - x_l^h}{h}\right). \tag{3.43}$$

We further note that  $\|\mathcal{A}\tilde{p}\|_{H^0(\omega_0)}$  is a norm on  $\tilde{p}$ , and since all norms are equivalent on a finite-dimensional space, we have

$$\|\mathcal{A}\tilde{p}\|_{H^0(\omega_0)} \leq C\|\tilde{p}\|_{H^0(\omega_0)}. \quad (3.44)$$

Therefore, from (3.43) and (3.44), and using the transformation  $y = (x - x_l^h)/h$ , we have

$$\begin{aligned} \|\mathcal{A}^h p\|_{H^0(\omega_l^h)}^2 &= \int_{\omega_l^h} |(\mathcal{A}^h p(x))^2| dx = \int_{\omega_0} |(\mathcal{A}\tilde{p}(y))^2| dy \\ &\leq C \int_{\omega_0} |\tilde{p}(y)|^2 dy \\ &= C \int_{\omega_l^h} |\tilde{p}((x - x_l^h)/h)|^2 dx \\ &= C \int_{\omega_l^h} |p(x)|^2 dx \\ &= C\|p\|_{H^0(\omega_l^h)}^2, \quad \text{for } p \in \mathcal{P}^k, \end{aligned} \quad (3.45)$$

with  $C$  independent of  $p$ ,  $l$ , and  $h$ . Combining (3.42), (3.45) with  $p^{l,h} - p^{j,h}$  and (3.40) yields

$$\begin{aligned} &|(\mathcal{A}^h p^{j,h})(x_l^h) - (\mathcal{A}^h p^{l,h})(x_l^h)| \\ &\leq Ch^{-n/2}(h^{r_j+1}\|u\|_{H^{r_j+1}(B_{\rho h}^j)} + h^{r_l+1}\|u\|_{H^{r_l+1}(B_{\rho h}^l)}), \end{aligned}$$

and hence, using (3.33), we have the estimate

$$\begin{aligned} &\sum_{l \in A_j^h} |(\mathcal{A}^h p^{j,h})(x_l^h) - (\mathcal{A}^h p^{l,h})(x_l^h)| \\ &\leq Ch^{-n/2} \{ Mh^{r_j+1}\|u\|_{H^{r_j+1}(B_{\rho h}^j)} \\ &\quad + \sum_{l \in A_j^h} h^{r_l+1}\|u\|_{H^{r_l+1}(B_{\rho h}^l)} \} \end{aligned} \quad (3.46)$$

for the right side of (3.37). Now we combine (3.37), (3.38), (3.39), and (3.46) to obtain

$$\left\| u - \sum_{l \in \mathbb{Z}^n} w_l^h \phi_l^h \right\|_{H^s(\omega_j^h)} \leq C \sum_{l \in A_j^h} h^{r_l+1-s} \|u\|_{H^{r_l+1}(B_{\rho h}^l)}. \quad (3.47)$$

4. Finally, we estimate  $\|u - \sum_{l \in \mathbb{Z}^n} w_l^h \phi_l^h\|_{H^s(\mathbb{R}^n)}$ . Using (3.47), which is valid for all  $j \in \mathbb{Z}^n$ , and (3.33) we obtain

$$\begin{aligned} \left\| u - \sum_{l \in \mathbb{Z}^n} w_l^h \phi_l^h \right\|_{H^s(\mathbb{R}^n)}^2 &\leq \sum_{j \in \mathbb{Z}^n} \left\| u - \sum_{l \in \mathbb{Z}^n} w_l^h \phi_l^h \right\|_{H^s(\omega_j^h)}^2 \\ &\leq C \sum_{j \in \mathbb{Z}^n} \sum_{l \in \mathcal{A}_j^h} h^{2(r_l+1-s)} \|u\|_{H^{r_l+1}(B_{\bar{\rho}h}^l)}^2 \\ &\leq C \sum_{j \in \mathbb{Z}^n} h^{2(r_j+1-s)} \|u\|_{H^{r_j+1}(B_{\bar{\rho}h}^j)}^2, \end{aligned} \tag{3.48}$$

where  $C$  (which depends on  $M$ ) is independent of  $u$  and  $h$ . This proves (3.25).

Suppose  $u \in H^{k'+1}(\mathbb{R}^n)$ , where  $0 \leq k' \leq k$ . Then taking  $r_j = k'$  in (3.48), and using the fact that the overlap in  $\{B_{\bar{\rho}h}^j\}_{j \in \mathbb{Z}^n}$  is bounded independently of  $h$ , we get

$$\begin{aligned} \left\| u - \sum_{l \in \mathbb{Z}^n} w_l^h \phi_l^h \right\|_{H^s(\mathbb{R}^n)} &\leq Ch^{k'+1-s} \left( \sum_{j \in \mathbb{Z}^n} \|u\|_{H^{k'+1}(B_{\bar{\rho}h}^j)}^2 \right)^{1/2} \\ &\leq Ch^{k'+1-s} \|u\|_{H^{k'+1}(\mathbb{R}^n)}, \end{aligned} \tag{3.49}$$

where  $C$  is independent of  $u$  and  $h$ , which is (3.26). □

**Remark 14.** Estimates (3.25) and (3.26) have been established provided  $\bar{\rho}$  is sufficiently large; specifically, provided (3.33) holds. As pointed out in connection with (3.34),  $\bar{\rho}$  depends on  $\rho$ . Note that the constants  $C$  in (3.25) and (3.26) depend on  $\bar{\rho}$ .

So far in this section, we have considered functions  $u$  defined on  $\mathbb{R}^n$ , and have presented a result on the approximation of  $u$  by particle shape functions. We now consider functions  $u$  defined on a bounded domain  $\Omega$  in  $\mathbb{R}^n$  with Lipschitz-continuous boundary. We will show that  $V_h^{k,q}$ , defined in (3.2), when restricted to  $\Omega$ , provides accurate approximation to  $u$ .

We first recall the well-known extension result (Stein 1970) that there is a bounded extension operator  $E : L_2(\Omega) \rightarrow L_2(\mathbb{R}^n)$ , *i.e.*, an operator  $E$  satisfying  $Eu|_{\Omega} = u$  for all  $u \in L_2(\Omega)$ , such that if  $u \in H^m(\Omega)$  then  $Eu \in H^m(\mathbb{R}^n)$  and

$$\|Eu\|_{H^m(\mathbb{R}^n)} \leq C_m \|u\|_{H^m(\Omega)}, \quad \text{for all } u \in H^m(\Omega), \quad m = 0, 1, \dots \tag{3.50}$$

Here  $C_m$  is independent of  $u$  but depends on  $m$ .

We define the subset  $\mathbb{Z}_{\Omega}^n$  of  $\mathbb{Z}^n$ , which will be used in the next result, by

$$\mathbb{Z}_{\Omega}^n = \{j \in \mathbb{Z}^n : \eta_j^h \cap \Omega \neq \emptyset\}, \tag{3.51}$$

where  $\eta_j^h = \text{supp } \phi_j^h$ .

**Theorem 3.5.** Suppose  $\phi \in H^q(\mathbb{R}^n)$ , with smoothness index  $q \geq 0$ , has compact support  $\eta \subset B_\rho$ , and is quasi-reproducing of order  $k$ . Suppose  $u \in H^{k'+1}(\Omega)$ , where  $0 \leq k' \leq k$ . Then there are weights  $w_j^h$  such that

$$\left\| u - \sum_{l \in \mathbb{Z}_\Omega^n} w_l^h \phi_l^h \right\|_{H^s(\Omega)} \leq Ch^{k'+1-s} \|u\|_{H^{k'+1}(\Omega)}, \quad 0 \leq s \leq \min(q, k' + 1), \quad (3.52)$$

where  $C$  is independent of  $u$  and  $h$ .

*Proof.* Suppose  $u \in H^{k'+1}(\Omega)$ , and let  $\bar{u} = Eu$ , where  $E$  is the extension operator mentioned above. Applying (3.26) of Theorem 3.4 to  $\bar{u}$ , there are weights  $w_l^h$  such that

$$\left\| \bar{u} - \sum_{l \in \mathbb{Z}^n} w_l^h \phi_l^h \right\|_{H^s(\mathbb{R}^n)} \leq Ch^{k'+1-s} \|\bar{u}\|_{H^{k'+1}(\mathbb{R}^n)}.$$

Therefore, from (3.50) with  $m = k' + 1$ , we have

$$\begin{aligned} \left\| u - \sum_{l \in \mathbb{Z}^n} w_l^h \phi_l^h \right\|_{H^s(\Omega)} &= \left\| \bar{u} - \sum_{l \in \mathbb{Z}^n} w_l^h \phi_l^h \right\|_{H^s(\Omega)} \\ &\leq \left\| \bar{u} - \sum_{l \in \mathbb{Z}^n} w_l^h \phi_l^h \right\|_{H^s(\mathbb{R}^n)} \\ &\leq Ch^{k'+1-s} \|\bar{u}\|_{H^{k'+1}(\mathbb{R}^n)} \\ &\leq Ch^{k'+1-s} \|u\|_{H^{k'+1}(\Omega)}. \end{aligned} \quad (3.53)$$

From the definition of  $\mathbb{Z}_\Omega^h$  in (3.51), it is clear that

$$\left\| u - \sum_{l \in \mathbb{Z}^n} w_l^h \phi_l^h \right\|_{H^s(\Omega)} = \left\| u - \sum_{l \in \mathbb{Z}_\Omega^n} w_l^h \phi_l^h \right\|_{H^s(\Omega)},$$

and therefore from (3.53) we get the desired result.  $\square$

By examining the approximation of  $u$ , obtained in Theorem 3.5, namely

$$\sum_{l \in \mathbb{Z}_\Omega^n} w_l^h \phi_l^h|_\Omega,$$

we see that the sum involves only those  $l$ s for which

$$\text{supp } \phi_l^h \cap \Omega \neq \emptyset,$$

that is, only those particles  $x_l^h$  such that  $\text{dist}(x_l^h, \Omega) < \rho h$ . So the approximation involves particle shape functions corresponding to the particles



inside  $\Omega$ , as well as some particles lying outside  $\Omega$ . We will denote the span of these shape functions by

$$V_{\Omega,h}^{k,q} = \text{span}\{\phi_j^h|_{\Omega} : \text{supp } \phi_j^h \cap \Omega \neq \emptyset\}. \tag{3.54}$$

Thus the functions in  $V_{\Omega,h}^{k,q}$  are the functions in  $V_h^{k,q}$  restricted to  $\Omega$ .

*(t, k\*)-regular systems*

We now introduce  $(t, k^*)$ -regular systems of functions (cf. Babuška and Aziz (1972)). Let  $\Omega \subseteq \mathbb{R}^n$ , and suppose  $S_h(\Omega)$ ,  $0 < h \leq 1$ , is a one-parameter family of linear spaces of functions on  $\Omega$ . For  $0 \leq k^* \leq t$ ,  $S_h(\Omega)$  will be called a  $(t, k^*)$ -regular system and will be denoted by  $S_h^{t,k^*}(\Omega)$  if

(1)

$$(1) S_h^{t,k^*}(\Omega) \subset H^{k^*}(\Omega), \quad \text{for } 0 < h \leq 1; \tag{3.55}$$

(2) For every  $u \in H^l(\Omega)$ , with  $0 \leq l$ , there is a function  $g_h \in S_h^{t,k^*}$  such that

$$\|u - g_h\|_{H^s(\Omega)} \leq Ch^\mu \|u\|_{H^l(\Omega)}, \quad \text{for } 0 \leq s \leq \min\{l, k^*\}, \tag{3.56}$$

where  $\mu = \min\{t - s, l - s\}$ . The constant  $C$  is independent of  $u$  and  $h$ .

We now introduce two additional notions.

(LA) A  $(t, k^*)$ -regular system  $S_h^{t,k^*}(\Omega)$  is said to satisfy a *local assumption* if, for  $u \in H^l(\Omega)$ , with support  $S$ , the function  $g_h \in S_h^{t,k^*}(\Omega)$  in (3.56) can be chosen so that the support  $S_h$  of  $g_h$  has the property that

$$S_h \subset S^{\lambda h} \equiv \{x \in \Omega : d(x, S) \leq \lambda h\},$$

where  $d(x, S)$  is the distance from  $x$  to  $S$ , and  $\lambda$  is independent of  $h$ .

(IA) We say that  $S_h^{t,k^*}(\Omega)$  satisfies an *inverse assumption* (cf. Babuška and Aziz (1972)) if there is an  $0 \leq \epsilon \leq k^*$  such that

$$\|g\|_{H^{k^*}(\Omega)} \leq Ch^{-(k^*-r)} \|g\|_{H^r(\Omega)},$$

for all  $k^* - \epsilon \leq r \leq k^*$  and all  $g \in S_h^{t,k^*}(\Omega)$ ,

where  $C$  is independent of  $h$  and  $g$  (it may depend on  $k^*$  and  $\epsilon$ ).

A  $(t, k^*)$ -regular system is referred to as a  $(t, k)$ -regular system in classical literature (Babuška and Aziz 1972). We have used  $k^*$  instead of  $k$  in this paper for notational clarity.

The approximation space  $V_{\Omega,h}^{k,q}$ , defined in (3.54), is a  $(t, k^*)$ -regular system: more precisely, we have the following result.

**Theorem 3.6.** Suppose  $0 \leq q < k + 1$ , and suppose  $\phi \in H^q(\mathbb{R}^n)$  has compact support and is quasi-reproducing of order  $k$ . Then  $V_{\Omega,h}^{k,q}$  is a  $(k + 1, q)$ -regular system.

*Proof.* We show that  $V_{\Omega,h}^{k,q}$  is a  $(t, k^*)$ -regular system with  $t = k + 1$  and  $k^* = q$ . Since  $\phi \in H^q(\mathbb{R}^n)$ , it is clear that  $V_{\Omega,h}^{k,q} \subset H^q(\Omega)$  and thus (3.55) is immediate with  $k^* = q$ . Next we show that (3.56) is satisfied. Suppose  $u \in H^l(\mathbb{R}^n)$  with  $l \geq 0$ . If  $l = 0$ , (3.56) is trivial. So, suppose  $1 \leq l$ . Applying Theorem 3.5, specifically (3.52) with  $k' = \min(k + 1, l) - 1$ , we get

$$\left\| u - \sum_{j \in \mathbb{Z}_\Omega^n} w_j^h \phi_j^h \right\|_{H^s(\Omega)} \leq Ch^{\min(l-s, k+1-s)} \|u\|_{H^{\min(l, k+1)}(\Omega)} \leq Ch^\mu \|u\|_{H^l(\mathbb{R}^n)},$$

for  $0 \leq s \leq \min\{q, \min\{l, k + 1\}\} = \min\{l, q\}$  (since  $q < k + 1$ ), where  $\mu = \min\{k + 1 - s, l - s\}$ . This is (3.56), with  $g_h = \sum_{l \in \mathbb{Z}_\Omega^n} w_l^h \phi_l^h|_\Omega$ ,  $t = k + 1$ , and  $k^* = q$ . □

We now show that  $V_{\Omega,h}^{k,q}$  satisfies the local assumption, LA.

**Theorem 3.7.** Suppose  $\phi \in H^q(\mathbb{R}^n)$ , where  $0 \leq q \leq k + 1$ , has compact support and is quasi-reproducing of order  $k$ . Then  $V_{\Omega,h}^{k,q}$  satisfies the local assumption, LA.

*Proof.* Suppose  $u \in H^l(\Omega)$  such that  $\text{supp } u = S \subset \Omega$ . Consider the approximation of  $u$ , obtained in Theorem 3.5, namely

$$g_h = \sum_{j \in \mathbb{Z}_\Omega^n} w_j^h \phi_j^h. \tag{3.57}$$

A careful study of the proofs of Theorems 3.5 and 3.4, and considering the zero extension of  $u$  outside  $\Omega$ , reveals that, for  $j \in \mathbb{Z}_\Omega^n$ ,

$$w_j^h = 0, \quad \text{if and only if } B_{\bar{\rho}h}^j \cap S = \emptyset.$$

Now, for  $j \in \mathbb{Z}_\Omega^n$  such that  $w_j^h \neq 0$ , we know that  $\eta_j^h = \text{supp } \phi_j^h \subset B_{\bar{\rho}h}^j$ . Therefore,  $S_h = \text{supp } g_h = \{x \in \mathbb{R}^n : d(x, S) \leq (\bar{\rho} + \rho)h\}$ , and so we can take  $\lambda = (\bar{\rho} + \rho)$  in the definition of LA. For small  $h$ , we have  $S_h \subset \Omega$ . Hence  $V_{\Omega,h}^{k,q}$  satisfies the local assumption, LA. □

**Remark 15.** The particle space  $V_h^{k,q}$  is  $(k + 1, q)$ -regular and satisfies the local assumption, LA, for  $\Omega = \mathbb{R}^n$ .

We note that the particle spaces  $V_{\Omega,h}^{k,q}$  and  $V_h^{k,q}$  will also satisfy the inverse assumption, IA, if additional conditions are imposed on the shape functions  $\{\phi_j^h\}$ . We will formulate these conditions in Section 3.3 in the context of non-uniformly distributed particles; the corresponding conditions on the shape functions associated with uniformly distributed particles can then be obtained as a special case.

3.3. Approximation by particle shape functions associated with arbitrary (non-uniformly distributed) particles in  $\mathbb{R}^n$ : the  $h$ -version

In this section we will generalize the major part of Theorem 3.4.

Suppose  $\{X^\nu\}_{\nu \in N}$  is a family of countable subsets of  $\mathbb{R}^n$ ; the family is indexed by the parameter  $\nu$ , which varies over the index set  $N$ . The points in  $X^\nu$  are called *particles*, and will be denoted by  $\underline{x}$ , to distinguish them from general points in  $\mathbb{R}^n$ . If it is necessary to emphasize that  $\underline{x} \in X^\nu$ , we will write  $\underline{x} = \underline{x}^\nu$ . With each  $\underline{x}^\nu \in X^\nu$  we associate

- $h_{\underline{x}^\nu}^\nu = h_{\underline{x}}^\nu$ , a positive number;
- $\omega_{\underline{x}^\nu}^\nu = \omega_{\underline{x}}^\nu$ , a bounded domain in  $\mathbb{R}^n$ ;
- $\phi_{\underline{x}^\nu}^\nu = \phi_{\underline{x}}^\nu$ , a function in  $H^q(\mathbb{R}^n)$ , with  $q \geq 0$  and with  $\eta_{\underline{x}^\nu}^\nu = \eta_{\underline{x}}^\nu \equiv \text{supp } \phi_{\underline{x}^\nu}^\nu$  assumed compact.

The numbers  $h_{\underline{x}^\nu}^\nu = h_{\underline{x}}^\nu$  will be referred to as the *sizes* of the particles  $\underline{x}$ , and the functions  $\phi_{\underline{x}^\nu}^\nu$  are called the *particle shape functions*. For a given  $\nu \in N$ , let

$$\mathcal{M}^\nu = \{X^\nu, \{h_{\underline{x}}^\nu, \omega_{\underline{x}}^\nu, \phi_{\underline{x}}^\nu\}_{\underline{x} \in X^\nu}\}.$$

$\mathcal{M}^\nu$  will be referred to as a *particle-shape function system* – and  $\{\mathcal{M}^\nu\}_{\nu \in N}$  as a family of particle-shape function systems. This nomenclature is similar to that used in the FEM when we speak of a triangulation and a family of triangulations.

Regarding the particle-shape function system, we make several assumptions.

A1. For each  $\nu$ ,

$$\bigcup_{\underline{x} \in X^\nu} \omega_{\underline{x}}^\nu = \mathbb{R}^n,$$

*i.e.*, for each  $\nu$ ,  $\{\omega_{\underline{x}}^\nu\}_{\underline{x} \in X^\nu}$  is an open cover of  $\mathbb{R}^n$ .

A2. For  $\underline{x} \in X^\nu$ , let

$$S_{\underline{x}}^\nu \equiv \{y \in X^\nu : \omega_{\underline{x}}^\nu \cap \omega_y^\nu \neq \emptyset\}.$$

There is a constant  $\kappa < \infty$ , which may depend on  $\{\mathcal{M}^\nu\}_{\nu \in N}$ , but neither on  $\nu$  nor on  $\underline{x} \in X^\nu$ , such that

$$\text{card } S_{\underline{x}}^\nu \leq \kappa, \quad \text{for all } \underline{x} \in X^\nu \text{ and all } \nu \in N.$$

A3. For all  $\underline{x} \in X^\nu$ , and  $\nu \in N$ ,  $\underline{x} \in \hat{\eta}_{\underline{x}}^\nu$  and  $\hat{\eta}_{\underline{x}}^\nu \subset \omega_{\underline{x}}^\nu$ .

A4. For  $\underline{x} \in X^\nu$ , let

$$\Omega_{\underline{x}}^\nu = \bigcup_{\underline{y} \in Q_{\underline{x}}^\nu} \omega_{\underline{y}}^\nu,$$

where

$$Q_{\underline{x}}^\nu \equiv \{\underline{y} \in X^\nu : \eta_{\underline{y}}^\nu \cap \omega_{\underline{x}}^\nu \neq \emptyset\}.$$

There is a  $0 < \bar{\rho} < \infty$ , which may depend on  $\{\mathcal{M}^\nu\}_{\nu \in N}$  but is independent of  $\underline{x}$  and  $\nu$ , such that

$$\Omega_{\underline{x}}^\nu \subset B_{\bar{\rho}h_{\underline{x}}^\nu}^{\underline{x}}, \quad \text{for all } \underline{x} \in X^\nu \text{ and } \nu \in N,$$

where  $B_{\bar{\rho}h_{\underline{x}}^\nu}^{\underline{x}}$  is the ball of radius  $\bar{\rho}h_{\underline{x}}^\nu$  centred at  $\underline{x}$ , namely,

$$B_{\bar{\rho}h_{\underline{x}}^\nu}^{\underline{x}} = \{x \in \mathbb{R}^n : \|x - \underline{x}\| \leq \bar{\rho}h_{\underline{x}}^\nu\}.$$

A5. For each  $\underline{x} \in X^\nu$ , there is a one-to-one mapping  $\mathcal{A}_{\underline{x}}^\nu : \mathcal{P}^k \rightarrow \mathcal{P}^k$  such that

$$\sum_{\underline{y} \in Q_{\underline{x}}^\nu} (\mathcal{A}_{\underline{x}}^\nu p)(\underline{y}) \phi_{\underline{y}}^\nu(x) = p(x), \quad \text{for } x \in \omega_{\underline{x}}^\nu, \text{ and any } p \in \mathcal{P}^k, \quad (3.58)$$

and

$$\|\mathcal{A}_{\underline{x}}^\nu p\|_{L^2(\omega_{\underline{y}}^\nu)} \leq C \|p\|_{L^2(\omega_{\underline{y}}^\nu)}, \quad \text{for all } p \in \mathcal{P}^k, \underline{y} \in Q_{\underline{x}}^\nu, \text{ and } \underline{x} \in X^\nu.$$

A6. For any  $0 \leq s \leq q$ ,

$$\|\phi_{\underline{y}}^\nu\|_{H^s(\omega_{\underline{x}}^\nu)} \leq C (h_{\underline{y}}^\nu)^{-s+n/2}, \quad \text{for all } \underline{y} \in Q_{\underline{x}}^\nu.$$

The constant  $C$  may depend on  $\{\mathcal{M}^\nu\}_{\nu \in N}$ , but is independent of  $\underline{y}, \underline{x}$ , and  $\nu$ .

A7. There is a constant  $C$  such that

$$\|w\|_{L^\infty(\omega_{\underline{y}}^\nu)} \leq C (h_{\underline{y}}^\nu)^{-n/2} \|w\|_{L^2(\omega_{\underline{y}}^\nu)}, \quad \text{for any } w \in \mathcal{P}^k,$$

where  $C$  is independent of  $\underline{y}$  and  $\nu$ .

**Remark 16.** From the definitions of  $Q_{\underline{x}}^\nu$  and  $S_{\underline{x}}^\nu$ , and assumption A3, it is clear that  $Q_{\underline{x}}^\nu \subset S_{\underline{x}}^\nu$ . Hence from assumption A2, it is immediate that

$$\text{card } Q_{\underline{x}}^\nu \leq \kappa. \quad (3.59)$$

We could, of course, have stated (3.59) as an assumption, but have chosen to state  $\text{card } S_{\underline{x}}^\nu \leq \kappa$  as an assumption because, generally,  $S_{\underline{x}}^\nu$  is easier to work with than  $Q_{\underline{x}}^\nu$ . We also note that assumptions A1–A4 imply that, for any  $x \in \mathbb{R}^n$ ,

$$\text{card } \{\underline{x} \in X^\nu : x \in \overset{\circ}{\eta}_{\underline{x}}^\nu\} \leq \kappa, \quad \text{for all } \nu \in N. \quad (3.60)$$

**Remark 17.** We note that the left-hand side of (3.58) in A5 is defined for all  $x \in \mathbb{R}^n$ , but the equality is required to hold only for  $x \in \omega_{\underline{x}}^\nu$ .

**Remark 18.** Note that assumptions A1–A7 can be thought of as assumptions on  $\mathcal{M}^\nu$ , for each  $\nu \in N$ ; they are assumptions on  $\{\mathcal{M}^\nu\}_{\nu \in N}$  in that they are assumptions on  $\mathcal{M}^\nu$  for each  $\nu$  and that the constants in the assumptions do not depend on  $\nu$ .

**Remark 19.** Assumption A5 effectively defines the notion of quasi-reproducing shape functions  $\phi_{\underline{x}}^\nu$  of order  $k$ . Note that the condition is local: the operator  $\mathcal{A}_{\underline{x}}^\nu$  depends on  $\underline{x}$ , the sum is taken only over  $\underline{y} \in Q_{\underline{x}}^\nu$ , and the equation holds only for  $x \in \omega_{\underline{x}}^\nu$ . The shape functions  $\phi_{\underline{x}}^\nu$  are said to be reproducing of order  $k$  if

$$\sum_{\underline{y} \in X^\nu} p(\underline{y}) \phi_{\underline{y}}^\nu(x) = p(x), \quad \text{for } x \in \mathbb{R}^n \text{ and any } p \in \mathcal{P}^k.$$

If the shape functions  $\phi_{\underline{x}}^\nu$  are reproducing of order  $k$ , then it is immediate that they satisfy A5 with  $\mathcal{A}_{\underline{x}}^\nu$  equal to the identity mapping for each  $\underline{x}$ .

**Remark 20.** Assumption A5 implies

$$\bigcup_{\underline{x} \in X^\nu} \hat{\eta}_{\underline{x}}^\nu = \mathbb{R}^n, \quad \text{for each } \nu.$$

**Remark 21.** Consider uniformly distributed particles,  $x_j^h$ , and associated particle shape functions,  $\phi_j^h$ , as defined in Section 3.2. Then with  $\underline{x}^\nu = x_j^h$ ,  $h_{\underline{x}}^\nu = h$ ,  $\phi_{\underline{x}}^\nu = \phi_j^h$ , and  $\omega_{\underline{x}}^\nu = \omega_j^h$ , as defined in the proof of Theorem 3.4, the associated particle-shape function system satisfies assumptions A1–A7. Note that  $\mathcal{A}_{\underline{x}}^\nu = \mathcal{A}_{x_j^h}^h = \mathcal{A}^h$  satisfies A5.

Suppose  $\{\mathcal{M}^\nu\}_{\nu \in N}$  is a family of particle-shape function systems, satisfying A1–A7. Define

$$\mathbb{V}_{\nu}^{k,q} = \text{span} \{ \phi_{\underline{x}}^\nu : \underline{x} \in X^\nu \}, \quad \text{for each } \nu \in N. \tag{3.61}$$

The next theorem gives an approximation error estimate when a function  $u$ , defined on  $\mathbb{R}^n$ , is approximated by an appropriate function in  $\mathbb{V}_{\nu}^{k,q}$ ,  $\nu \in N$ .

**Theorem 3.8.** Suppose the family  $\{\mathcal{M}^\nu\}_{\nu \in N}$  of particle-shape function systems satisfies A1–A7, and  $h_{\underline{x}}^\nu \leq 1$  for  $\underline{x} \in X^\nu$ ,  $\nu \in N$ . Suppose

$$\sum_{\underline{x} \in X^\nu} \|u\|_{H^{r_{\underline{x}}^\nu+1}(B_{\bar{\rho}h_{\underline{x}}^\nu}^{\underline{x}})}^2 < \infty, \quad \text{where } r_{\underline{x}}^\nu \leq k, \quad \text{for all } \underline{x} \in X^\nu \text{ and } \nu \in N, \tag{3.62}$$

where  $\bar{\rho}$  is introduced in A4. Further, suppose that operators  $\mathcal{A}_{\underline{x}}^\nu$ , introduced in A5, satisfy

$$\|(\mathcal{A}_{\underline{x}}^\nu - \mathcal{A}_{\underline{y}}^\nu)p\|_{H^0(\omega_{\underline{y}}^\nu)} \leq C(h_{\underline{x}}^\nu)^{r_{\underline{x}}^\nu+1} \|p\|_{H^0(\omega_{\underline{y}}^\nu)}, \quad \text{for all } p \in \mathcal{P}^k, \quad \underline{y} \in Q_{\underline{x}}^\nu, \tag{3.63}$$

for all  $\underline{x} \in X^\nu$ ,  $\nu \in N$ , where  $C$  is independent of  $\underline{x}$  and  $\nu$ . Then there are weights  $w_{\underline{y}}^\nu \in \mathbb{R}$ , for  $\underline{y} \in X^\nu$  and for all  $\nu \in N$ , such that

$$\begin{aligned} & \left\| u - \sum_{\underline{y} \in X^\nu} w_{\underline{y}}^\nu \phi_{\underline{y}}^\nu \right\|_{H^s(\mathbb{R}^n)} \\ & \leq C \left( \sum_{\underline{y} \in X^\nu} (h_{\underline{y}}^\nu)^{2(r_{\underline{y}}^\nu + 1 - s)} \|u\|_{H^{r_{\underline{y}}^\nu + 1}(B_{\rho h_{\underline{y}}^\nu}^{\underline{y}})} \right)^{1/2}, \end{aligned} \tag{3.64}$$

for  $0 \leq s \leq \inf\{q, r_{\underline{y}}^\nu + 1 : \underline{y} \in X^\nu, \nu \in N\}$ . The constant  $C$  depends on the constants in Assumptions A1–A7 and on (3.63), but neither on  $u$  nor on  $\nu$ .

**Note.** If  $\{\phi_{\underline{x}}^\nu\}$  is reproducing of order  $k$ , then (3.63) is trivially satisfied (cf. Remark 19). Since shape functions that are reproducing of order  $k$  are mainly used in practice, we have not included (3.63) in the set of basic assumptions (A1–A7).

*Proof.* The proof of this result is analogous to the proof of Theorem 3.4.

**1.** The sets  $\omega_{\underline{x}}^\nu$  play the role of the sets  $\omega_j^h$  in the proof of Theorem 3.4. The sets  $Q_{\underline{x}}^\nu$ ,  $\Omega_{\underline{x}}^\nu$ ,  $B_{\rho h_{\underline{x}}^\nu}^{\underline{x}}$ , and the mapping  $\mathcal{A}_{\underline{x}}^\nu$  play the roles of the sets  $A_j^h$ ,  $\Omega_j^h$ ,  $B_{\rho h}^j$ , and the mapping  $\mathcal{A}_j^h$ , respectively, in the proof of Theorem 3.4. Assumptions A1–A7 state the properties of these sets and the mappings we will need in this proof.

**2.** For any  $\underline{y} \in X^\nu$ , since  $u \in H^{r_{\underline{y}}^\nu + 1}(B_{\rho h_{\underline{y}}^\nu}^{\underline{y}})$ , it is well known that there is a polynomial  $p^{\underline{y}, \nu} = p_k^{\underline{y}, \nu}$  of degree  $\leq k$ , such that

$$\|u - p_k^{\underline{y}, \nu}\|_{H^s(B_{\rho h_{\underline{y}}^\nu}^{\underline{y}})} \leq C (h_{\underline{y}}^\nu)^{r_{\underline{y}}^\nu + 1 - s} \|u\|_{H^{r_{\underline{y}}^\nu + 1}(B_{\rho h_{\underline{y}}^\nu}^{\underline{y}})}, \tag{3.65}$$

for  $0 \leq s \leq r_{\underline{y}}^\nu + 1 \leq k + 1$ , where  $C$  is independent of  $u$ ,  $\nu$  and  $\underline{y}$ , but does depend on  $k$  ( $p_k^{\underline{y}, \nu}$  can, in fact, be selected so that its degree  $\leq r_{\underline{y}}^\nu$ ). Define the weights

$$w_{\underline{y}}^\nu = (\mathcal{A}_{\underline{y}}^\nu p^{\underline{y}, \nu})(\underline{y}), \tag{3.66}$$

where  $\mathcal{A}_{\underline{y}}^\nu$  is the operator introduced in assumption A5.

Let  $\underline{x} \in X^\nu$  be fixed. We will work with the polynomial  $p^{\underline{x}, \nu}$ , which satisfies (3.65) with  $\underline{y} = \underline{x}$ , as well as the polynomial  $p^{\underline{y}, \nu}$ . Using (3.66)

we find

$$\begin{aligned}
 & \left\| u - \sum_{\underline{y} \in X^\nu} w_{\underline{y}}^\nu \phi_{\underline{y}}^\nu \right\|_{H^s(\omega_{\underline{x}}^\nu)} \\
 & \leq \left\| u - \sum_{\underline{y} \in Q_{\underline{x}}^\nu} w_{\underline{y}}^\nu \phi_{\underline{y}}^\nu \right\|_{H^s(\omega_{\underline{x}}^\nu)} \\
 & \leq \left\| u - \sum_{\underline{y} \in Q_{\underline{x}}^\nu} (\mathcal{A}_{\underline{x}}^\nu p^{x,\nu})(\underline{y}) \phi_{\underline{y}}^\nu \right\|_{H^s(\omega_{\underline{x}}^\nu)} \\
 & \quad + \sum_{\underline{y} \in Q_{\underline{x}}^\nu} |(\mathcal{A}_{\underline{x}}^\nu p^{x,\nu})(\underline{y}) - (\mathcal{A}_{\underline{y}}^\nu p^{y,\nu})(\underline{y})| \|\phi_{\underline{y}}^\nu\|_{H^s(\omega_{\underline{x}}^\nu)}. \tag{3.67}
 \end{aligned}$$

We now estimate the two terms on the right-hand side of (3.67).

**3.** From assumption A5, we know that

$$\sum_{\underline{y} \in Q_{\underline{x}}^\nu} (\mathcal{A}_{\underline{x}}^\nu p)(\underline{y}) \phi_{\underline{y}}^\nu(x) = p(x), \quad \text{for } x \in \omega_{\underline{x}}^\nu, \text{ and any } p \in \mathcal{P}^k.$$

Using this formula and (3.65) with  $\underline{y} = \underline{x}$ , we obtain the estimate

$$\begin{aligned}
 \left\| u - \sum_{\underline{y} \in Q_{\underline{x}}^\nu} (\mathcal{A}_{\underline{x}}^\nu p^{x,\nu})(\underline{y}) \phi_{\underline{y}}^\nu \right\|_{H^s(\omega_{\underline{x}}^\nu)} & \leq \|u - p^{x,\nu}\|_{H^s(\omega_{\underline{x}}^\nu)} \\
 & \leq C(h_{\underline{x}}^\nu)^{r_{\underline{x}}^\nu+1-s} \|u\|_{H^{r_{\underline{x}}^\nu+1}(B_{\frac{\rho}{2}h_{\underline{x}}^\nu}^{\underline{x}})} \tag{3.68}
 \end{aligned}$$

for the first term.

Using assumption A6, we have

$$\|\phi_{\underline{y}}^\nu\|_{H^s(\omega_{\underline{x}}^\nu)} \leq C(h_{\underline{y}}^\nu)^{-s+n/2}, \quad \text{for all } \underline{y} \in Q_{\underline{x}}^\nu.$$

Thus

$$\begin{aligned}
 & \sum_{\underline{y} \in Q_{\underline{x}}^\nu} |(\mathcal{A}_{\underline{x}}^\nu p^{x,\nu})(\underline{y}) - (\mathcal{A}_{\underline{y}}^\nu p^{y,\nu})(\underline{y})| \|\phi_{\underline{y}}^\nu\|_{H^s(\omega_{\underline{x}}^\nu)} \\
 & \leq \sum_{\underline{y} \in Q_{\underline{x}}^\nu} (h_{\underline{y}}^\nu)^{-s+n/2} |(\mathcal{A}_{\underline{x}}^\nu p^{x,\nu})(\underline{y}) - (\mathcal{A}_{\underline{y}}^\nu p^{y,\nu})(\underline{y})|. \tag{3.69}
 \end{aligned}$$

It remains to estimate the right-hand side of this inequality.

For  $\underline{y} \in Q_{\underline{x}}^\nu$ , we have  $\omega_{\underline{y}}^\nu \subset \Omega_{\underline{x}}^\nu$ , and hence  $\omega_{\underline{y}}^\nu \subset B_{\frac{\rho}{2}h_{\underline{x}}^\nu}^{\underline{x}}$ , using assumption A4.

Also  $\omega_{\underline{y}}^\nu \subset B_{\rho h_{\underline{y}}^\nu}^{\underline{y}}$ . Thus, using (3.65) with  $s = 0$ , we have

$$\begin{aligned} & \|p^{\underline{x},\nu} - p^{\underline{y},\nu}\|_{H^0(\omega_{\underline{y}}^\nu)} \\ & \leq \|p^{\underline{x},\nu} - u\|_{H^0(\omega_{\underline{y}}^\nu)} + \|u - p^{\underline{y},\nu}\|_{H^0(\omega_{\underline{y}}^\nu)} \\ & \leq (h_{\underline{x}}^\nu)^{r_{\underline{x}}^\nu+1} \|u\|_{H^{r_{\underline{x}}^\nu+1}(B_{\rho h_{\underline{x}}^\nu}^{\underline{x}})} + (h_{\underline{y}}^\nu)^{r_{\underline{y}}^\nu+1} \|u\|_{H^{r_{\underline{y}}^\nu+1}(B_{\rho h_{\underline{y}}^\nu}^{\underline{y}})}. \end{aligned} \tag{3.70}$$

Now, using assumption A7, we have

$$\begin{aligned} & |(\mathcal{A}_{\underline{x}}^\nu p^{\underline{x},\nu})(\underline{y}) - (\mathcal{A}_{\underline{y}}^\nu p^{\underline{y},\nu})(\underline{y})| \\ & \leq |[(\mathcal{A}_{\underline{x}}^\nu - \mathcal{A}_{\underline{y}}^\nu)p^{\underline{x},\nu}](\underline{y})| + |[\mathcal{A}_{\underline{y}}^\nu(p^{\underline{x},\nu} - p^{\underline{y},\nu})](\underline{y})| \\ & \leq \|(\mathcal{A}_{\underline{x}}^\nu - \mathcal{A}_{\underline{y}}^\nu)p^{\underline{x},\nu}\|_{L^\infty(\omega_{\underline{y}}^\nu)} + \|\mathcal{A}_{\underline{y}}^\nu(p^{\underline{x},\nu} - p^{\underline{y},\nu})\|_{L^\infty(\omega_{\underline{y}}^\nu)} \\ & \leq C(h_{\underline{y}}^\nu)^{-n/2} \{ \|(\mathcal{A}_{\underline{x}}^\nu - \mathcal{A}_{\underline{y}}^\nu)p^{\underline{x},\nu}\|_{H^0(\omega_{\underline{y}}^\nu)} \\ & \quad + \|\mathcal{A}_{\underline{y}}^\nu(p^{\underline{x},\nu} - p^{\underline{y},\nu})\|_{H^0(\omega_{\underline{y}}^\nu)} \}. \end{aligned} \tag{3.71}$$

Also, using assumption A5 and (3.70), we obtain

$$\begin{aligned} & \|\mathcal{A}_{\underline{y}}^\nu(p^{\underline{x},\nu} - p^{\underline{y},\nu})\|_{H^0(\omega_{\underline{y}}^\nu)} \\ & \leq C \|p^{\underline{x},\nu} - p^{\underline{y},\nu}\|_{H^0(\omega_{\underline{y}}^\nu)} \\ & \leq C \{ (h_{\underline{x}}^\nu)^{r_{\underline{x}}^\nu+1} \|u\|_{H^{r_{\underline{x}}^\nu+1}(B_{\rho h_{\underline{x}}^\nu}^{\underline{x}})} + (h_{\underline{y}}^\nu)^{r_{\underline{y}}^\nu+1} \|u\|_{H^{r_{\underline{y}}^\nu+1}(B_{\rho h_{\underline{y}}^\nu}^{\underline{y}})} \}. \end{aligned} \tag{3.72}$$

Moreover, from (3.63), we have

$$\|(\mathcal{A}_{\underline{x}}^\nu - \mathcal{A}_{\underline{y}}^\nu)p^{\underline{x},\nu}\|_{H^0(\omega_{\underline{y}}^\nu)} \leq (h_{\underline{x}}^\nu)^{r_{\underline{x}}^\nu+1} \|p^{\underline{x},\nu}\|_{H^0(\omega_{\underline{y}}^\nu)}, \tag{3.73}$$

and from (3.65), with  $\underline{y} = \underline{x}$ , and recalling that  $h_{\underline{x}}^\nu \leq 1$ , we get

$$\begin{aligned} \|p^{\underline{x},\nu}\|_{H^0(\omega_{\underline{y}}^\nu)} & \leq \|p^{\underline{x},\nu} - u\|_{H^0(\omega_{\underline{y}}^\nu)} + \|u\|_{H^0(\omega_{\underline{y}}^\nu)} \\ & \leq C(h_{\underline{x}}^\nu)^{r_{\underline{x}}^\nu+1} \|u\|_{H^{r_{\underline{x}}^\nu+1}(B_{\rho h_{\underline{x}}^\nu}^{\underline{x}})} + \|u\|_{H^0(\omega_{\underline{y}}^\nu)} \\ & \leq C \|u\|_{H^{r_{\underline{x}}^\nu+1}(B_{\rho h_{\underline{x}}^\nu}^{\underline{x}})}. \end{aligned} \tag{3.74}$$

Combining (3.71)–(3.74), we have

$$\begin{aligned} & |(\mathcal{A}_{\underline{x}}^\nu p^{\underline{x},\nu})(\underline{y}) - (\mathcal{A}_{\underline{y}}^\nu p^{\underline{y},\nu})(\underline{y})| \\ & \leq C(h_{\underline{y}}^\nu)^{-n/2} \{ (h_{\underline{x}}^\nu)^{r_{\underline{x}}^\nu+1} \|u\|_{H^{r_{\underline{x}}^\nu+1}(B_{\rho h_{\underline{x}}^\nu}^{\underline{x}})} + (h_{\underline{y}}^\nu)^{r_{\underline{y}}^\nu+1} \|u\|_{H^{r_{\underline{y}}^\nu+1}(B_{\rho h_{\underline{y}}^\nu}^{\underline{y}})} \}. \end{aligned} \tag{3.75}$$

Then we combine (3.69), (3.75), (3.59), and assumption A2 (see (3.59)),



to obtain

$$\begin{aligned} & \sum_{\underline{y} \in Q_{\underline{x}}^\nu} |(\mathcal{A}_{\underline{x}}^\nu p^{\underline{x},\nu})(\underline{y}) - (\mathcal{A}_{\underline{y}}^\nu p^{\underline{y},\nu})(\underline{y})| \|\phi_{\underline{y}}^\nu\|_{H^s(\omega_{\underline{x}})} \\ & \leq C \left\{ \kappa (h_{\underline{x}}^\nu)^{r_{\underline{x}}^\nu+1} \|u\|_{H^{r_{\underline{x}}^\nu+1}(B_{\rho h_{\underline{x}}^\nu}^{\underline{x}})} + \sum_{\underline{y} \in Q_{\underline{x}}^\nu} (h_{\underline{y}}^\nu)^{r_{\underline{y}}^\nu+1} \|u\|_{H^{r_{\underline{y}}^\nu+1}(B_{\rho h_{\underline{y}}^\nu}^{\underline{y}})} \right\} \\ & \leq C \sum_{\underline{y} \in Q_{\underline{x}}^\nu} (h_{\underline{y}}^\nu)^{r_{\underline{y}}^\nu+1} \|u\|_{H^{r_{\underline{y}}^\nu+1}(B_{\rho h_{\underline{y}}^\nu}^{\underline{y}})}, \end{aligned} \tag{3.76}$$

which is an estimate for the second term in (3.67). Thus, from (3.67), (3.68), and (3.76), we have

$$\left\| u - \sum_{\underline{y} \in Q_{\underline{x}}^\nu} w_{\underline{y}}^\nu \phi_{\underline{y}}^\nu \right\|_{H^s(\omega_{\underline{x}})} \leq C \sum_{\underline{y} \in Q_{\underline{x}}^\nu} (h_{\underline{y}}^\nu)^{r_{\underline{y}}^\nu+1} \|u\|_{H^{r_{\underline{y}}^\nu+1}(B_{\rho h_{\underline{y}}^\nu}^{\underline{y}})}. \tag{3.77}$$

4. It remains to estimate  $\|u - \sum_{\underline{y} \in X^\nu} w_{\underline{y}}^\nu \phi_{\underline{y}}^\nu\|_{H^s(\mathbb{R}^n)}$ . Using (3.77), which is valid for all  $\underline{x} \in X^\nu$ , and assumptions A1, A2, A4, we obtain

$$\begin{aligned} \left\| u - \sum_{\underline{y} \in X^\nu} w_{\underline{y}}^\nu \phi_{\underline{y}}^\nu \right\|_{H^s(\mathbb{R}^n)}^2 & \leq \sum_{\underline{x} \in X^\nu} \left\| u - \sum_{\underline{y} \in Q_{\underline{x}}^\nu} w_{\underline{y}}^\nu \phi_{\underline{y}}^\nu \right\|_{H^s(\omega_{\underline{x}})}^2 \\ & \leq C \sum_{\underline{x} \in X^\nu} \sum_{\underline{y} \in Q_{\underline{x}}^\nu} (h_{\underline{y}}^\nu)^{2(r_{\underline{y}}^\nu+1-s)} \|u\|_{H^{r_{\underline{y}}^\nu+1}(B_{\rho h_{\underline{y}}^\nu}^{\underline{y}})}^2 \\ & \leq C \sum_{\underline{y} \in X^\nu} (h_{\underline{y}}^\nu)^{2(r_{\underline{y}}^\nu+1-s)} \|u\|_{H^{r_{\underline{y}}^\nu+1}(B_{\rho h_{\underline{y}}^\nu}^{\underline{y}})}^2, \end{aligned} \tag{3.78}$$

which is (3.64). □

It will be useful to state estimate (3.64) in Theorem 3.8 in certain alternative ways. Given a family of particle-shape function systems  $\{\mathcal{M}^\nu\}_{\nu \in \mathbb{N}}$  satisfying A1–A7, define

$$h^\nu = \sup_{\underline{x} \in X^\nu} h_{\underline{x}}^\nu, \quad \text{for each } \nu. \tag{3.79}$$

With this definition, from (3.64) we have

$$\left\| u - \sum_{\underline{y} \in X^\nu} w_{\underline{y}}^\nu \phi_{\underline{y}}^\nu \right\|_{H^s(\mathbb{R}^n)} \leq C \left( \sum_{\underline{y} \in X^\nu} (h_{\underline{y}}^\nu)^{2(r_{\underline{y}}^\nu+1-s)} \|u\|_{H^{r_{\underline{y}}^\nu+1}(B_{\rho h_{\underline{y}}^\nu}^{\underline{y}})} \right)^{1/2}. \tag{3.80}$$

Now, if  $r_{\underline{x}}^\nu = k'$ , where  $0 \leq k' \leq k$ , for all  $\underline{y} \in X^\nu$  and  $\nu$ , then (3.80) leads to the following result.

**Theorem 3.9.** Suppose the family  $\{\mathcal{M}^\nu\}_{\nu \in N}$  of particle-shape function systems satisfies A1–A7, (3.63), and in addition, suppose  $h^\nu \leq 1$ , for all  $\nu$ . Suppose  $\|u\|_{H^{k'+1}(\mathbb{R}^n)} < \infty$ , where  $0 \leq k' \leq k$ . Then there are weights  $w_{\underline{y}}^\nu \in \mathbb{R}$  such that

$$\left\| u - \sum_{\underline{y} \in X^\nu} w_{\underline{y}}^\nu \phi_{\underline{y}}^\nu \right\|_{H^s(\mathbb{R}^n)} \leq C(h^\nu)^{k'+1-s} \|u\|_{H^{k'+1}(\mathbb{R}^n)}, \tag{3.81}$$

for  $0 \leq s \leq \min(q, k' + 1)$ , where  $C$  is independent of  $u$  and  $\nu$ .

We note that if  $h^{\nu_1} < h^{\nu_2}$ ,  $\nu_1, \nu_2 \in N$ , then we would view  $\mathcal{M}^{\nu_1}$  as a ‘refinement’ of  $\mathcal{M}^{\nu_2}$ .

There is yet another way to state the estimate (3.81). Let  $0 < h \leq 1$ , and suppose  $\{\mathcal{M}^\nu\}_{\nu \in N}$ , a family of particle-shape function systems satisfying A1–A7, (3.63), and in addition,

$$h^\nu = \sup_{\underline{x}^\nu \in X^\nu} h_{\underline{x}^\nu}^\nu \leq h, \quad \text{for each } \nu. \tag{3.82}$$

We can now think of  $\nu = \nu(h)$  as determined by  $h$ , although, of course, many particle-shape function systems satisfy (3.82). We can, in fact, think of having a one-to-one correspondence between  $\nu$  and  $h$ . Thus we can regard  $h$  as the parameter and write a family of particle-shape function systems as

$$\{\mathcal{M}^h\}_{0 < h \leq 1} = \{X^h, \{h_{\underline{x}}^h, \omega_{\underline{x}}^h, \phi_{\underline{x}}^h\}_{\underline{x} \in X^h}\}_{0 < h \leq 1}$$

instead of  $\{\mathcal{M}^\nu\}_{\nu \in N}$ . With this understanding that  $\nu = \nu(h)$ , the estimate (3.81) can be written as

$$\left\| u - \sum_{\underline{y} \in X^h} w_{\underline{y}}^h \phi_{\underline{y}}^h \right\|_{H^s(\mathbb{R}^n)} \leq Ch^{k+1-s} \|u\|_{H^{k+1}(\mathbb{R}^n)}. \tag{3.83}$$

We are naturally interested in having  $h \downarrow 0$ , and hence in considering  $\nu(h)$ s for which  $h^{\nu(h)} \downarrow 0$ . More specifically, we will often consider a sequence  $h_m \downarrow 0$ , and the corresponding sequence of particle systems  $\mathcal{M}^{\nu_1}, \mathcal{M}^{\nu_2}, \dots$ , where  $\nu_l = \nu(h_l)$ .

We remark that the estimate (3.64) is stronger than (3.81) and (3.83), in that (3.64) uses  $h_{\underline{x}^\nu}^\nu$  instead of the larger  $h^\nu$ , and (3.64) allows a more general regularity assumption on the function  $u$ . The viewpoint outlined in this paragraph is similar to the usual view of meshes in the FEM.

For a given family  $\{\mathcal{M}^\nu\}_{\nu \in N}$  of particle-shape function systems, we defined the space  $\mathbb{V}_\nu^{k,q}$  in (3.61). With  $h$ ,  $0 < h \leq 1$ , as the parameter, *i.e.*, for a given family  $\mathcal{M}^h$ ,  $0 < h \leq 1$ , we will use the space

$$\mathbb{V}_h^{k,q} \equiv \mathbb{V}_{\nu(h)}^{k,q} = \text{span} \{ \phi_{\underline{x}}^h : \underline{x} \in X^h \}. \tag{3.84}$$

So far, we have discussed the approximation of a function  $u$  defined on  $\mathbb{R}^n$ , by particle shape functions. We now consider  $u$  defined on  $\Omega$ , where  $\Omega$  is a bounded domain, with Lipschitz-continuous boundary, in  $\mathbb{R}^n$ . We now show that functions in  $\mathbb{V}_{\Omega,h}^{k,q}$ , defined by

$$\mathbb{V}_{\Omega,h}^{k,q} = \text{span}\{\phi_{\underline{x}}^h|_{\Omega} : \phi_{\underline{x}}^h \in \mathbb{V}_h^{k,q}, \text{ for } \underline{x} \in A_{\Omega}^h\}, \tag{3.85}$$

where

$$A_{\Omega}^h = \{\underline{x} \in X^h : \hat{\eta}_{\underline{x}}^h \cap \Omega \neq \emptyset\},$$

provide accurate approximation of functions  $u$  defined on  $\Omega$ .

**Theorem 3.10.** Suppose  $\mathcal{M}^h$ ,  $0 < h \leq 1$ , is a family of particle-shape function systems satisfying A1–A7 and (3.63). Let  $\Omega \subset \mathbb{R}^n$  be a bounded domain with Lipschitz-continuous boundary, and suppose  $u \in H^{k'+1}(\Omega)$ , where  $0 \leq k' \leq k$ . Then there are weights  $w_{\underline{y}}^h \in \mathbb{R}$  such that

$$\left\| u - \sum_{\underline{y} \in A_{\Omega}^h} w_{\underline{y}}^h \phi_{\underline{y}}^h \right\|_{H^s(\Omega)} \leq Ch^{k'+1-s} \|u\|_{H^{k'+1}(\Omega)}, \tag{3.86}$$

for  $0 \leq s \leq \min(q, k' + 1)$ , where the constant  $C$  is independent of  $u$  and  $h$ .

The proof of this theorem is based on using (3.83) on the extension  $\bar{u} = Eu$ , and is similar to the proof of Theorem 3.5. We omit the proof of this theorem. We note that the approximation  $\sum_{\underline{y} \in A_{\Omega}^h} w_{\underline{y}}^h \phi_{\underline{y}}^h$ , obtained in Theorem 3.10, is such that

$$\sum_{\underline{y} \in A_{\Omega}^h} w_{\underline{y}}^h \phi_{\underline{y}}^h|_{\Omega} \in \mathbb{V}_{\Omega,h}^{k,q}.$$

In Section 3.2, we reviewed the notion of  $(t, k^*)$ -regular system  $S_h(\Omega)$ . In the next theorem, we show that  $\mathbb{V}_{\Omega,h}^{k,q}$  is a  $(k + 1, q)$ -system.

**Theorem 3.11.** Suppose  $\mathcal{M}^h$ ,  $0 < h \leq 1$ , is a family of particle-shape function systems satisfying A1–A7 and (3.63). Let  $\Omega \subset \mathbb{R}^n$  be a bounded domain with Lipschitz-continuous boundary. Then  $\mathbb{V}_{\Omega,h}^{k,q}$  is a  $(k+1, q)$ -regular system, where  $k$  is the order of quasi-reproducing shape functions in  $\mathcal{M}^h$ .

The proof of this theorem is similar to the proof of Theorem 3.6, and will be omitted.

**Remark 22.** The space  $\mathbb{V}_{\Omega,h}^{k,q}$  satisfies the local assumption, LA.

*Quasi-uniform particle-shape function system*

We will call a family of particle-shape function systems  $\{\mathcal{M}^h\}_{0 < h \leq 1}$  quasi-uniform if there is a  $\beta$ ,  $1 < \beta < \infty$ , such that

$$\beta^{-1} \leq \frac{h_{\underline{x}}^h}{h_{\underline{y}}^h} \leq \beta, \quad \text{for all } \underline{x}, \underline{y} \in X^h \text{ and } 0 < h \leq 1. \quad (3.87)$$

We note that (3.87) is equivalent to

$$\beta^{-1} \leq \frac{h}{h_{\underline{y}}^h} \leq \beta, \quad \text{for all } \underline{y} \in X^h \text{ and } 0 < h \leq 1, \quad (3.88)$$

where  $h$  is defined by (3.82).

**Remark 23.** We can also define *uniform* particle-shape function system by imposing the condition

$$h_{\underline{x}}^h = h_{\underline{y}}^h, \quad \text{for all } \underline{x}, \underline{y} \in X^h \text{ and } 0 < h \leq 1.$$

We note that the system with uniformly distributed particles and the associated shape functions as defined in Section 3.1 is uniform. But uniform particle-shape function systems may have particles that are not uniformly distributed.

Consider a family of particle-shape function systems  $\{\mathcal{M}^h\}_{0 < h \leq 1}$  satisfying assumptions A1–A7. Let  $\Omega \subset \mathbb{R}^n$  be a bounded domain, and suppose  $\mathcal{M}^h$  satisfies the following additional assumptions:

- $\mathcal{M}^h$  is quasi-uniform, *i.e.*, (3.88) holds;
- for all  $x \in A_{\Omega}^h$ , there is a  $\beta > 0$  such that, for  $0 \leq s \leq q$ ,

$$\beta^{-1} h^{\frac{n}{2}-s} \leq \|\phi_{\underline{y}}^h\|_{H^s(\omega_{\underline{x}}^h \cap \Omega)} \leq \beta h^{\frac{n}{2}-s}, \quad \text{for all } \underline{y} \in Q_{\underline{x}}^h, \quad (3.89)$$

where  $Q_{\underline{x}}^h = \{\underline{y} \in X^h : \eta_{\underline{y}}^h \cap \omega_{\underline{x}}^h \neq \emptyset\}$  (*cf.* A4);

- for all  $w_{\underline{y}} \in \mathbb{R}$ , for  $\underline{y} \in Q_{\underline{x}}^h$ , and  $\underline{x} \in A_{\Omega}^h$ , there exists  $C > 0$ , independent of  $\underline{x}$ , such that

$$h^{-s} \left[ \sum_{\underline{y} \in Q_{\underline{x}}^h} |w_{\underline{y}}|^2 h^n \right]^{1/2} \leq C \left\| \sum_{\underline{y} \in Q_{\underline{x}}^h} w_{\underline{y}} \phi_{\underline{y}}^h \right\|_{H^s(\omega_{\underline{x}}^h \cap \Omega)}, \quad \text{for } 0 \leq s \leq q. \quad (3.90)$$

Then the particle space  $\mathbb{V}_{\Omega, h}^{k, q}$  satisfies the inverse assumption IA, introduced in Section 3.2. To see this, consider  $\underline{x} \in A_{\Omega}^h$ . Then, using (3.89) and (3.90),

we have

$$\begin{aligned}
 \left\| \sum_{\underline{y} \in Q_{\underline{x}}^h} w_{\underline{y}} \phi_{\underline{y}}^h \right\|_{H^q(\omega_{\underline{x}}^h \cap \Omega)} &\leq \sum_{\underline{y} \in Q_{\underline{x}}^h} |w_{\underline{y}}| \|\phi_{\underline{y}}^h\|_{H^q(\omega_{\underline{x}}^h \cap \Omega)} \\
 &\leq Ch^{\frac{n}{2}-q} \left( \sum_{\underline{y} \in Q_{\underline{x}}^h} |w_{\underline{y}}|^2 \right)^{1/2} \\
 &= Ch^{s-q} h^{-s} \left( \sum_{\underline{y} \in Q_{\underline{x}}^h} |w_{\underline{y}}|^2 h^n \right)^{1/2} \\
 &\leq Ch^{s-q} \left\| \sum_{\underline{y} \in Q_{\underline{x}}^h} w_{\underline{y}} \phi_{\underline{y}}^h \right\|_{H^s(\omega_{\underline{x}}^h \cap \Omega)}, \tag{3.91}
 \end{aligned}$$

where  $C$  depends on  $\kappa$  (cf. A2). Thus

$$\begin{aligned}
 \left\| \sum_{\underline{y} \in A_{\Omega}^h} w_{\underline{y}} \phi_{\underline{y}}^h \right\|_{H^q(\Omega)}^2 &\leq \sum_{\underline{x} \in A_{\Omega}^h} \left\| \sum_{\underline{y} \in Q_{\underline{x}}^h} w_{\underline{y}} \phi_{\underline{y}}^h \right\|_{H^q(\omega_{\underline{x}}^h \cap \Omega)}^2 \\
 &\leq Ch^{2(s-q)} \sum_{\underline{x} \in A_{\Omega}^h} \left\| \sum_{\underline{y} \in Q_{\underline{x}}^h} w_{\underline{y}} \phi_{\underline{y}}^h \right\|_{H^s(\omega_{\underline{x}}^h \cap \Omega)}^2 \\
 &\leq Ch^{2(s-q)} \left\| \sum_{\underline{y} \in A_{\Omega}^h} w_{\underline{y}} \phi_{\underline{y}}^h \right\|_{H^s(\omega_{\underline{x}}^h \cap \Omega)}^2.
 \end{aligned}$$

Since any element  $g$  of  $\mathbb{V}_{\Omega,h}^{k,q}$  is of the form  $\sum_{\underline{y} \in A_{\Omega}^h} w_{\underline{y}} \phi_{\underline{y}}^h|_{\Omega}$ , we have shown that  $\mathbb{V}_{\Omega,h}^{k,q}$  satisfies the inverse assumption, IA. We summarize the above discussion in the following theorem.

**Theorem 3.12.** Suppose  $\mathcal{M}^h$ ,  $0 < h \leq 1$ , is a family of quasi-uniform particle-shape function systems satisfying A1–A7, (3.89), and (3.90). Let  $\Omega \subset \mathbb{R}^n$  be a bounded domain with Lipschitz-continuous boundary. Then  $\mathbb{V}_{\Omega,h}^{k,q}$  satisfies the inverse assumption, IA.

**Remark 24.** We can show that the particle space  $\mathbb{V}_h^{k,q}$  also satisfies the inverse assumption IA, if  $\mathcal{M}^h$ , satisfying A1–A7, also satisfies (3.89) and (3.90) with  $\Omega = \mathbb{R}^n$ .

So far we have addressed the approximation properties of the spaces  $\mathbb{V}_h^{k,q}$  and  $\mathbb{V}_{\Omega,h}^{k,q}$ , *i.e.*, approximation by the functions of the form

$$\sum_{y \in X^h} w_y^h \phi_y^h \quad \text{and} \quad \sum_{y \in A_\Omega^h} w_y^h \phi_y^h.$$

We will now present the approximation properties of the space

$$\mathbb{W}_{\Omega,h}^{k',q} = \left\{ v|_\Omega : v = \sum_{\underline{x} \in A_\Omega^h} \phi_{\underline{x}}^h \psi_{\underline{x}}^h, \psi_{\underline{x}}^h \in \mathcal{P}^{k'}(\hat{\eta}_{\underline{x}}^h) \right\},$$

where  $\{\phi_{\underline{x}}^h\}$  form a partition of unity. The space  $\mathbb{W}_{\Omega,h}^{k',q}$  was used in Taylor, Zienkiewicz and Onate (1998), and is a special case of the space  $V^\nu$  considered in Section 6. The approximation properties of the space  $\mathbb{W}_{\Omega,h}^{k',q}$  are similar to the approximation properties of  $\mathbb{V}_{\Omega,h}^{k,q}$ .

**Theorem 3.13.** Suppose  $\mathcal{M}^h, 0 < h \leq 1$ , is a family of particle-shape function system that satisfy A1–A7 with  $k = 0$  and  $\mathcal{A}_{\underline{x}}^h = I$ , which implies that the shape functions  $\{\phi_{\underline{x}}^h\}$  form a partition of unity. Let  $\Omega \in \mathbb{R}^n$  be a bounded domain with Lipschitz-continuous boundary and suppose  $u \in H^{k'+1}(\Omega)$ , where  $0 \leq k'$ . Then there exist  $\psi_{\underline{x}}^h \in \mathcal{P}^{k'}(\hat{\eta}_{\underline{x}}^h)$  such that

$$\left\| u - \sum_{\underline{x} \in A_\Omega^h} \phi_{\underline{x}}^h \psi_{\underline{x}}^h \right\|_{H^1(\Omega)} \leq Ch^{k'} \|u\|_{H^{k'+1}(\Omega)},$$

where  $C$  is independent of  $u$  and  $h$ .

The proof of this result can be obtained directly from Theorems 6.1–6.3, and we will comment on the proof in Section 6.

**Remark 25.** In Theorem 3.13, we assumed that  $\{\phi_{\underline{x}}^h\}$  are reproducing of order  $k = 0$  (*i.e.*, they form a partition of unity) because, as indicated in Section 4, such shape functions are easier to construct. The general case where  $\{\phi_{\underline{x}}^h\}$  are assumed to be quasi-reproducing of order  $k$  will be addressed in a forthcoming paper. In this situation, for  $0 \leq s \leq q$ , we expect the error estimate to be

$$\left\| u - \sum_{\underline{x} \in A_\Omega^h} \phi_{\underline{x}}^h \psi_{\underline{x}}^h \right\|_{H^s(\Omega)} \leq Ch^{k'+k+1-s} \|u\|_{H^{k'+k+1}(\Omega)}.$$

**Remark 26.** Note that in the situation addressed in Theorem 3.13, we associate with each particle  $\underline{x} \in X^h$ , multiple shape functions

$$\phi_{\underline{x}}^h \psi_{\underline{x},j}^h, j = 1, 2, \dots, N,$$

where  $\{\psi_{x,j}^h\}_{j=1}^N$  is a basis for  $\mathcal{P}^{k'}(\tilde{\eta}_x^h)$ . This is in contrast to the situation discussed earlier in this section, where with each particle we associated only the single shape function  $\phi_x^h$ . In a forthcoming paper, we will address, in a more general context, the use of multiple shape functions associated with a single particle.

#### 4. Construction and selection of particle shape functions

In Section 3, we presented an abstract description of particle-shape function systems with respect to uniform as well as non-uniform distribution of particles. We showed that if these particle-shape function systems satisfy certain properties (assumptions A1–A7 and (3.63)), they will have good approximation properties. In this section we will present an example of a particular particle-shape function system, where the shape functions are reproducing of order  $k$ , and show that under certain conditions they satisfy assumptions A1–A7, and hence have good approximation properties. We note that (3.63) is trivially satisfied. This example will also show that a wide variety of particle shape functions can be constructed. Therefore it is important to address the issue of selecting an appropriate class of shape functions that would yield efficient approximation of the solution of a particular problem, or a class of problems. We also present an interpolation result that will indicate a procedure for choosing a class of shape functions, among a given collection of such classes. Such shape functions will yield the smallest value of the usual Sobolev norm interpolation error, when the interpolated function is included in a higher-order Sobolev space.

##### 4.1. An example of a class of particle shape functions

Several particle shape functions have been developed over the past decade. SPH shape functions (Gingold and Monaghan 1977) were introduced in the context of fluid dynamics, whereas Shepard functions (Shepard 1968) and MLS shape functions (Lancaster and Salkauskas 1981) were introduced in the context of data fitting with respect to irregularly distributed particles in higher dimensions. In the 1990s, RKP (reproducing kernel particle) shape functions were introduced (Liu, Jun and Zhang 1995) in the context of approximation of solutions of partial differential equations. In this paper, we describe the construction of RKP shape functions for non-uniform as well as uniform distribution of particles, and relate them to the abstract setting given in Section 3. Specifically, we will show that the resulting particle-shape function system satisfies assumptions A1–A7.

##### *Non-uniformly distributed particles*

For  $\nu \in N$ ,  $N$  an index set, let  $X^\nu = \{x_i^\nu\}_{i \in \mathbb{Z}}$  where  $x_i^\nu \in \mathbb{R}^n$ . With each  $x_i^\nu \in X^\nu$  we associate a positive number  $h_i^\nu$ . We consider a fixed  $\nu$  and often

suppress the superscript  $\nu$ , for example, we write  $x_i$  and  $h_i$  instead of  $x_i^\nu$  and  $h_i^\nu$ , respectively. We will comment on  $\nu$  when appropriate.

Let  $w(x) \geq 0$  be a continuous function with compact support, specifically,

$$\eta \equiv \text{supp } w(x) = \overline{B}_R(0), \quad R > 0. \quad (4.1)$$

The function  $w(x)$  is called a weight function (or window function).

The commonly used weight functions in one dimension are as follows.

(a) Gaussian:

$$w(x) = \begin{cases} \frac{e^{\delta(x/R)^2} - e^\delta}{1 - e^\delta}, & |x| \leq R \\ 0, & |x| \geq R, \end{cases} \quad (4.2)$$

where  $\delta > 0$ .

(b) cubic spline:

$$w(x) = \begin{cases} \frac{2}{3} - 4(x/R)^2 + 4(|x|/R)^3, & |x| \leq R/2 \\ \frac{4}{3} - 4(|x|/R) + 4(x/R)^2 - \frac{4}{3}(|x|/R)^3, & R/2 \leq |x| \leq R \\ 0, & |x| > R. \end{cases} \quad (4.3)$$

(c) conical:

$$w(x) = \begin{cases} [1 - (x/R)^2]^l, & |x| \leq R \\ 0, & |x| > R, \end{cases} \quad (4.4)$$

where  $l = 1, 2, \dots$

We note that one may consider nonsymmetric versions of some of these weight functions, as was done in Armentano and Duran (2001).

In  $\mathbb{R}^n$ ,  $w(x)$  can be constructed from a one-dimensional weight function  $w(x)$  (symmetric) as  $w(x) = w(\|x\|)$ , where  $\|x\|$  is the Euclidean length of  $x$ . Further,  $w(x)$  can also be constructed via  $w(x) = \prod_{j=1}^n w_j(x)$ , where  $x = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$ . Consequently,  $\eta$  will be an  $n$ -cube. However, we will assume  $\eta$  given by (4.1) in this section.

For each  $j$ , we define

$$w_j(x) = w\left(\frac{x - x_j}{h_j}\right). \quad (4.5)$$

Clearly,

$$\eta_j \equiv \text{supp } w_j(x) = \overline{B}_{Rh_j}(x_j). \quad (4.6)$$

Let

$$Q_i = \{x_j : \dot{\eta}_i \cap \dot{\eta}_j \neq \emptyset\}, \quad (4.7)$$



and assume that

$$\cup_{j \in \mathbb{Z}} \overset{\circ}{\eta}_j = \mathbb{R}^n, \tag{4.8}$$

$$\text{card } Q_i \leq \kappa, \quad \text{for all } i \in \mathbb{Z}, \tag{4.9}$$

where  $\kappa$  is independent of  $i$  and  $\nu$ .

For a given integer  $k, k \geq 0$ , the RKP shape function  $\phi_j(x)$ , associated with the particle  $x_j$ , is defined by

$$\phi_j(x) = w_j(x) \sum_{|\alpha| \leq k} (x - x_j)^\alpha b_\alpha(x), \tag{4.10}$$

where  $b_\alpha(x)$  are chosen so that

$$\sum_{j \in \mathbb{Z}} p(x_j) \phi_j(x) = p(x), \quad \text{for } x \in \mathbb{R}^n \text{ and } p \in \mathcal{P}^k(\mathbb{R}^n), \tag{4.11}$$

so that  $\{\phi_j(x)\}_{j \in \mathbb{Z}}$  are reproducing of order  $k$ . This gives rise to a linear system in  $b_\alpha(x)$ , namely

$$\sum_{|\alpha| \leq k} m_{\alpha+\beta}(x) b_\alpha(x) = \delta_{|\beta|,0}, \quad \text{for } |\beta| \leq k, \tag{4.12}$$

where  $\delta_{|\beta|,|\alpha|}$  is the Kronecker delta, and

$$m_\alpha(x) = \sum_{j \in \mathbb{Z}} w_j(x) (x - x_j)^\alpha.$$

It is clear from (4.6) and (4.10) that

$$\text{supp } \phi_j(x) = \text{supp } w_j(x) = \eta_j. \tag{4.13}$$

We now briefly describe the derivation of (4.12). For a fixed  $y \in \mathbb{R}^n$ , consider

$$p_\beta(x) = (y - x)^\beta, \quad 0 \leq |\beta| \leq k.$$

Using  $p(x) = p_\beta(x)$  in (4.11), we get

$$\sum_{j \in \mathbb{Z}} (y - x_j)^\beta \phi_j(x) = (y - x)^\beta,$$

and letting  $y = x$  in the above equality, we have

$$\sum_{j \in \mathbb{Z}} (x - x_j)^\beta \phi_j(x) = \delta_{|\beta|,0}, \quad 0 \leq |\beta| \leq k. \tag{4.14}$$

Thus (4.11) implies (4.14); one can also show that (4.14) implies (4.11).

Now using (4.10) in (4.14), we get

$$\begin{aligned}
 \delta_{|\beta|,0} &= \sum_{j \in \mathbb{Z}} (x - x_j)^\beta \phi_j(x) \\
 &= \sum_{j \in \mathbb{Z}} (x - x_j)^\beta w_j(x) \sum_{|\alpha| \leq k} (x - x_j)^\alpha b_\alpha(x) \\
 &= \sum_{|\alpha| \leq k} b_\alpha(x) \sum_{j \in \mathbb{Z}} w_j(x) (x - x_j)^{\alpha+\beta} \\
 &= \sum_{|\alpha| \leq k} m_{\alpha+\beta}(x) b_\alpha(x), \tag{4.15}
 \end{aligned}$$

which is (4.12).

We now consider the unique solvability of (4.11). For  $k = 0$ , the linear system (4.12) is  $m_0(x)b_0(x) = [\sum_{i \in \mathbb{Z}} w_i(x)]b_0(x) = 1$ . Assuming  $\sum_{i \in \mathbb{Z}} w_i(x) \neq 0$ ,  $x \in \mathbb{R}^n$ , we have  $b_0(x) = 1/m_0(x)$ . Therefore from (4.10), we have

$$\phi_j(x) = \frac{w_j(x)}{\sum_{i \in \mathbb{Z}} w_i(x)}, \quad j \in \mathbb{Z}.$$

This expression for  $\{\phi_j(x)\}$  gives another verification that they form a partition of unity. These shape functions, introduced in Shepard (1968), are called Shepard functions.

The unique solvability of (4.12), for  $k \geq 1$ , depends on the weight functions and on the distribution of the particles  $\{x_i\}$  in  $\mathbb{R}^n$ . The required distribution of particles is in turn related to the interpolation problem in  $\mathbb{R}^n$ . It was shown in Han and Meng (2001) that a necessary condition for unique solvability of (4.12) is that, for  $x \in \mathbb{R}^n$ ,

$$\text{card } A(x) \geq \dim \mathcal{P}^k, \tag{4.16}$$

where

$$A(x) = \{x_l : x \in \mathring{\eta}_l\}. \tag{4.17}$$

For  $k = 1$ , Han and Meng (2001) showed that the linear system (4.12) is nonsingular if the following conditions are satisfied.

- (a) There are constants  $C_1, C_2 > 0$ , independent of  $\nu$ , and  $h > 0$ , such that

$$C_1 \leq \frac{h_i}{h} \leq C_2, \quad \text{for all } i \in \mathbb{Z}. \tag{4.18}$$

- (b) There are constants  $\tilde{C}_1, \tilde{C}_2 > 0$ , independent of  $\nu$ , such that, for any  $x \in \mathbb{R}^n$ , there are  $(n + 1)$  particles  $x_{i_l} \in A(x), l = 0, \dots, n$ , such that

$$\min_{0 \leq l \leq n} w\left(\frac{x - x_{i_l}}{h}\right) \geq \tilde{C}_1 > 0 \tag{4.19}$$

and

$$\text{Volume } K(x_{i_0}, x_{i_1}, \dots, x_{i_n}) \geq \tilde{C}_2 h^n, \tag{4.20}$$

where  $K(x_{i_0}, x_{i_1}, \dots, x_{i_n})$  is the simplex with vertices  $x_{i_0}, x_{i_1}, \dots, x_{i_n}$ .

We will now cast RKP shape functions, discussed above, in the framework of a particle-shape function system, introduced in Section 3.3. We started with a collection of particles  $X^\nu = \{x_j^\nu\}_{j \in \mathbb{Z}}$ , where  $x_j^\nu \in \mathbb{R}^n$ , and positive numbers  $h_j^\nu$ . Corresponding to each particle  $x_j^\nu \in X^\nu$ , we associated, in (4.10), the RKP shape function,  $\phi_j^\nu = \phi_j$  with compact support  $\eta_j^\nu = \eta_j = \overline{B}_{Rh_j}(x_j)$ , where the parameter  $R$  was related to the weight function  $w(x)$ . It was shown in Han and Meng (2001) that if  $w(x) \in C^q(\mathbb{R}^n)$ , then  $\phi_j \in C^q(\mathbb{R}^n)$ , and thus  $\phi_j \in H^q(\mathbb{R}^n)$ ; here we assume  $q = 1$ . We recall that the conditions (4.8), (4.8), (4.16), (4.18)–(4.20) were required for the construction of shape functions,  $\phi_j, j \in \mathbb{Z}$ . We let  $\omega_j^\nu = \omega_j \equiv \dot{\eta}_j$ ; certainly each  $\omega_j^\nu$  is a bounded domain. We now show that the family of particle-shape function systems  $\{\mathcal{M}^\nu\}_{\nu \in N}$ , where

$$\mathcal{M}^\nu = \{X^\nu, \{h_i^\nu, \omega_i^\nu, \phi_i^\nu\}\},$$

with these choices for  $\phi_i^\nu$  and  $\omega_i^\nu$ , satisfies assumptions A1–A7 in Section 3.3. We will continue to use the notation introduced earlier in this section, and suppress  $\nu$ ; the statements of A1–A7 using this notation should be clear.

- Since  $\omega_i = \dot{\eta}_i$  for  $i \in \mathbb{Z}$ , assumption A1 follows from (4.8). We also see that the sets  $S_x^\nu \equiv S_i$  and  $Q_x^\nu \equiv Q_i$ , introduced in assumptions A2 and A4 are the same. Thus A2 follows from (4.9).
- Assumption A3 is immediate from the definition  $\omega_i$ .
- Since  $\omega_j = \dot{\eta}_j$ , the set  $\Omega_x^\nu$ , introduced in assumption A4, is given by  $\Omega_x^\nu = \Omega_i = \cup_{x_j \in Q_i} \dot{\eta}_j$ . Since each  $\eta_j$  is a ball of radius  $Rh_j$ , it is easily seen, using (4.18), that assumption A4 is satisfied with  $\bar{\rho} = 3RC_2/C_1$ .
- RKP shape functions,  $\{\phi_j\}$ , considered here, are reproducing of order  $k = 1$ , *i.e.*, they satisfy (4.11) with  $k = 1$ . Thus A5 is satisfied with  $\mathcal{A}_x^\nu = \mathcal{A}_i = I$  (identity), for all  $i \in \mathbb{Z}$ , with  $k = 1$ .
- Han and Meng (2001) showed that, if the weight function  $w(x) \in C^q$ , then

$$\|\phi_i\|_{W^{s,\infty}(\dot{\eta}_i)} \leq C(h_i)^{-s}, \quad \text{for } 0 \leq s \leq q \text{ and } i \in \mathbb{Z}.$$

Thus using a scaling argument and this estimate, we obtain

$$\|\phi_i\|_{H^s(\dot{\eta}_i)} \leq C(h_i)^{-s+n/2}, \quad \text{for } 0 \leq s \leq q \text{ and } i \in \mathbb{Z}, \tag{4.21}$$

where  $h$  and  $h_i$  satisfy (4.18). Recall that we assumed  $q = 1$ . Now let  $x_j \in Q_i$ . Then

$$\|\phi_j\|_{H^s(\dot{\eta}_i)} = \|\phi_j\|_{H^s(\dot{\eta}_i \cap \dot{\eta}_j)} \leq \|\phi_j\|_{H^s(\dot{\eta}_j)},$$

and combining this with (4.21), we get, for  $0 \leq s \leq 1$ ,

$$\|\phi_j\|_{H^s(\tilde{\eta}_i)} \leq C(h_i)^{-s+n/2}, \quad \text{for all } x_j \in Q_i \text{ and } i \in \mathbb{Z},$$

which is assumption A6 with  $q = 1$ .

- A scaling argument shows that assumption A7 is satisfied.

We remark that (3.60), together with condition (b) (following (4.18)), establishes a lower bound of  $\kappa$ , namely  $(n+1) \leq \kappa$ .

We have thus shown that assumptions A1–A7, with  $k = 1$  and  $q = 1$ , are satisfied by RKP particle-shape function systems provided (4.8), (4.9), (4.16), (4.18)–(4.20) are satisfied. Thus we can apply Theorem 3.8 to obtain an approximation error estimate for RKP particle-shape function systems. Note that the condition (3.63) in Theorem 3.8 is trivially satisfied with  $\mathcal{A}_{\underline{x}}^\nu = I$  for all  $\underline{x} \in X^\nu$ . We remark that an interpolation error estimate, under the assumptions (4.8), (4.9), (4.16), (4.18)–(4.20), was also obtained in Han and Meng (2001).

We note that A1–A7 only guarantee good approximability of the shape functions; they do not provide a recipe to construct particle shape functions that are quasi-reproducing or reproducing of order  $k$ . In fact the availability of such particle shape functions is assumed in A5. Further assumptions may be needed to construct such shape functions; for example (4.16), (4.18)–(4.20) were needed to construct RKP particle shape functions. Therefore, there should be enough restrictions on the particle distributions and the supports of shape functions for it to be possible to construct these shape functions satisfying A1–A7, thereby ensuring good approximation properties.

#### *Uniformly distributed particles*

We consider the uniformly distributed particles  $x_j^h = jh$ ,  $j \in \mathbb{Z}^n$  as in Section 3.2. This is a special case of the non-uniformly distributed particles considered in the first part of this section. For each  $x_j^h$ , we define  $w_j^h(x) = w(\frac{x-x_j^h}{h})$ , where  $w(x) \geq 0$  is a continuous weight function with compact support  $\eta = \overline{B}_R(0)$ . Clearly,  $\eta_j^h \equiv \text{supp } w_j^h(x) = \overline{B}_{Rh}(x_j^h)$ . It can be easily shown that if  $R = 3\sqrt{n}/2$  (in fact, we need only  $R > \sqrt{n}$ ), then (4.8), (4.9) with  $\kappa = (4R+1)^n$ , (4.18) with  $C_1 = C_2 = 1$ , and (4.20) with  $\tilde{C}_2 = 1/2$  are satisfied. If  $w(x) = w(r)$ , with  $r = \|x\|$ , is monotonically decreasing in  $r$ , then it can also be shown easily that (4.19) is satisfied with  $\tilde{C}_1 = w(\sqrt{n})$ . Therefore, RKP shape functions  $\phi_i^h(x)$ , associated with  $x_i^h$ , for all  $i \in \mathbb{Z}^n$ , can be constructed using the procedure described in (4.10), (4.11) and (4.12) for  $k = 1$ , namely

$$\phi_j^h(x) = w_j^h(x) \sum_{|\alpha| \leq k} (x - x_j^h)^\alpha b_\alpha^h(x), \quad (4.22)$$

where  $\{b_\alpha^h(x)\}_{|\alpha|\leq k}$  is the solution of

$$\sum_{|\alpha|\leq k} m_{\alpha+\beta}^h(x)b_\alpha^h(x) = \delta_{|\beta|,0}, \quad |\beta| \leq k, \tag{4.23}$$

with  $k = 1$ , and

$$m_\alpha^h(x) = \sum_{j \in \mathbb{Z}^n} w_j^h(x)(x - x_j^h)^\alpha. \tag{4.24}$$

The shape functions  $\{\phi_j^h\}$  satisfy

$$\sum_{j \in \mathbb{Z}^n} p(x_j^h)\phi_j^h(x) = p(x), \quad \text{for all } x \in \mathbb{R}^n \text{ and } p \in \mathcal{P}^k(\mathbb{R}^n). \tag{4.25}$$

As with the non-uniformly distributed particles, we consider the family of particle-shape function systems

$$\mathcal{M}^h = \{X^h, \{h_{\underline{x}}, \omega_{\underline{x}}^h, \phi_{\underline{x}}^h\}_{\underline{x} \in X^h}\}, \quad 0 < h \leq 1,$$

for RKP shape functions with respect to uniformly distributed particles, by letting  $X^h = \{\underline{x} = x_j^h : j \in \mathbb{Z}^n\}$  and using  $h_{\underline{x}}^h = h$ ,  $\omega_{\underline{x}}^h = \eta_j^h$  and  $\phi_{\underline{x}}^h = \phi_j^h$ . Note that here we used the parameter  $h$  instead of  $\nu$ . We have shown above that conditions (4.8), (4.9), (4.18)–(4.20) are satisfied, with  $w(x) = w(r)$ , a monotonically decreasing weight function in  $r$ , and  $R = 3\sqrt{n}/2$ . Therefore, based on the discussion on RKP particle-shape function systems for non-uniformly distributed particles, it clear that  $\{\mathcal{M}^h\}_{0 < h \leq 1}$  satisfies assumptions A1–A7 with  $k = 1$ , ensuring good approximation properties of the RKP shape functions.

We recall that in Section 3.1, the particle shape function  $\phi_i^h(x)$  was defined in (3.1) by scaling and translating the basic shape function  $\phi(x)$  for uniformly distributed particles, *i.e.*, they were translation-invariant. We will show that the RKP shape functions  $\{\phi_i^h\}_{i \in \mathbb{Z}^n}$ , constructed via (4.22) and (4.23), also satisfy (3.1) with  $\phi(x) = \phi_0^1(x)$  (*i.e.*, with  $i = 0$  and  $h = 1$ ). We assume that the linear system (4.23) is invertible for  $k \geq 1$ .

From (4.22) and (4.23) with  $i = 0$  and  $h = 1$ , we have

$$\phi(x) = w(x) \sum_{|\alpha|\leq k} x^\alpha b_\alpha^1(x), \tag{4.26}$$

where  $b_\alpha^1(x)$  are the solutions of

$$\sum_{|\alpha|\leq k} m_{\alpha+\beta}^1(x)b_\alpha^1(x) = \delta_{|\beta|,0}, \quad \text{for } |\beta| \leq k, \tag{4.27}$$

and

$$m_\alpha^1(x) = \sum_{j \in \mathbb{Z}^n} w(x - j)(x - j)^\alpha. \tag{4.28}$$

We replace  $x$  by  $\frac{x-x_i^h}{h}$  in (4.27) and (4.28) to get

$$\sum_{|\alpha| \leq k} m_{\alpha+\beta}^1 \left( \frac{x-x_i^h}{h} \right) b_\alpha \left( \frac{x-x_i^h}{h} \right) = \delta_{|\beta|,0}, \quad \text{for } |\beta| \leq k, \quad (4.29)$$

where

$$\begin{aligned} m_\alpha^1 \left( \frac{x-x_i^h}{h} \right) &= \sum_{j \in \mathbb{Z}^n} w \left( \frac{x-x_i^h}{h} - j \right) \left( \frac{x-x_i^h}{h} - j \right)^\alpha \\ &= \sum_{j \in \mathbb{Z}^n} w \left( \frac{x-x_{i+j}^h}{h} \right) \left( \frac{x-x_{i+j}^h}{h} \right)^\alpha \\ &= \frac{1}{h^{|\alpha|}} \sum_{j \in \mathbb{Z}^n} w_j^h(x) (x-x_j^h)^\alpha \\ &= \frac{1}{h^{|\alpha|}} m_\alpha^h(x). \end{aligned} \quad (4.30)$$

Using (4.30) in (4.29), we get

$$\sum_{|\alpha| \leq k} \frac{1}{h^{|\alpha+\beta|}} m_{\alpha+\beta}^h(x) b_\alpha \left( \frac{x-x_i^h}{h} \right) = \delta_{|\beta|,0},$$

and therefore

$$\sum_{|\alpha| \leq k} m_{\alpha+\beta}^h(x) \frac{1}{h^{|\alpha|}} b_\alpha \left( \frac{x-x_i^h}{h} \right) = h^{|\beta|} \delta_{|\beta|,0} = \delta_{|\beta|,0}, \quad \text{for all } |\beta| \leq k. \quad (4.31)$$

Since the  $\{b_\alpha^h(x)\}$  is the unique solution for (4.23), it is clear from (4.31) that

$$b_\alpha^h(x) = \frac{1}{h^{|\alpha|}} b_\alpha \left( \frac{x-x_i^h}{h} \right),$$

and thus from (4.26) we have

$$\begin{aligned} \phi \left( \frac{x-x_i^h}{h} \right) &= w \left( \frac{x-x_i^h}{h} \right) \sum_{|\alpha| \leq k} \left( \frac{x-x_i^h}{h} \right)^\alpha b_\alpha \left( \frac{x-x_i^h}{h} \right) \\ &= w_i^h(x) \sum_{|\alpha| \leq k} (x-x_i^h)^\alpha \frac{1}{h^{|\alpha|}} b_\alpha \left( \frac{x-x_i^h}{h} \right) \\ &= w_i^h(x) \sum_{|\alpha| \leq k} (x-x_i^h)^\alpha b_\alpha^h(x) \\ &= \phi_i^h(x). \end{aligned}$$

Thus, for uniformly distributed particles, RKP shape functions satisfy (3.1), *i.e.*, they are translation-invariant.

**Remark 27.** To approximate functions defined on a bounded domain  $\Omega$ , we use the restrictions of RKP shape functions on  $\Omega$ , as described in Section 3.3 (*cf.* (3.85) and Theorem 3.10). We note that the RKP shape functions corresponding to the particles near the boundary of  $\Omega$ , as defined here, are different from the RKP shape functions defined in Han and Meng (2001) and Liu *et al.* (1995). But they are the same for particles inside  $\Omega$ , sufficiently away from the boundary  $\partial\Omega$ . They are also the same when  $\Omega = \mathbb{R}^n$ .

**Remark 28.** RKP shape functions are not available analytically in simple forms. Their values at a point  $x \in \mathbb{R}^n$  are computed via (4.10), which involves the solution of the linear system (4.12) for each  $x$ . Thus computation of RKP shape functions, which are reproducing of order  $k$ , is not easy, especially for higher  $k$ . Moreover, as a consequence of (4.16), it is necessary that the supports of these shape functions be large so that the linear system (4.12) is invertible. We note that it is much easier to construct the shape functions for the space  $\mathbb{W}_{\Omega,h}^{k,q}$ , introduced near the end of Section 3.3.

#### 4.2. Interpolation and selection

In this section, we will address the interpolation of a function in terms of particle shape functions, and will propose a procedure to select a shape function that will yield efficient approximation. We consider uniformly distributed particles  $\{x_j^h\}$  in  $\mathbb{R}^n$ , and the associated particle shape functions  $\{\phi_j^h\}$ , defined in (3.1), where  $\phi \in H^q(\mathbb{R}^n)$  with  $q \geq 1$  has compact support;  $\text{supp } \phi \subset B_R(0)$ . We have seen that  $\{\phi_j^h\}$  is translation-invariant,  $\text{supp } \phi_j^h \in B_{Rh}(x_j^h)$ , and in addition

$$\|\phi_j^h\|_{H^1(\mathbb{R}^n)} \leq h^{n/2-1} \|\phi\|_{H^1(\mathbb{R}^n)}. \tag{4.32}$$

We assume that  $\{\phi_j^h\}$  is reproducing of order  $k$ , *i.e.*, (4.25) holds.

Let  $\Omega$  be a bounded domain in  $\mathbb{R}^n$ . We will consider a smooth function  $u(x)$  defined in  $\Omega$  and study the error  $u - \tilde{\mathcal{I}}_h u$ , where  $\tilde{\mathcal{I}}_h u$  is the ‘interpolant’ of  $u$  in terms of  $\phi_j^h$ . The results in this subsection are from Babuška, Banerjee and Osborn (200x), and we refer to this paper for some of the details that we do not present here.

We now define the ‘interpolant’  $\tilde{\mathcal{I}}_h u$  of a function  $u$ . For any  $x \in \mathbb{R}^n$ , let

$$A_x^h = \{k \in \mathbb{Z}^n : x \in \mathring{\eta}_k^h\}.$$

Here  $A_x^h$  is called the influence set for the point  $x$ . Then  $(\tilde{\mathcal{I}}_h u)(x)$  is defined as

$$(\tilde{\mathcal{I}}_h u)(x) = \sum_{j \in A_x^h} u(x_j^h) \phi_j^h(x). \tag{4.33}$$

In (4.33), we, of course, assume that  $u(x_j^h)$  is defined for all  $j \in A_x^h$ . If  $p \in \mathcal{P}^k$ , then from (4.25) we have

$$\sum_{j \in A_x^h} p(x_j^h) \phi_j^h(x) = \sum_{j \in \mathbb{Z}^n} p(x_j^h) \phi_j^h(x) = p(x), \quad \text{for all } x \in \mathbb{R}^n, \tag{4.34}$$

*i.e.*,  $\tilde{\mathcal{I}}_h p = p$ . Now let  $u \in H^s(\Omega)$  with  $s > n/2$ . For some  $x \in \Omega$ , the particles  $x_j^h$  for  $j \in A_x^h$  may be outside  $\Omega$ , and  $u(x_j^h)$  may not be defined. To define  $\tilde{\mathcal{I}}_h u(x)$  for  $u \in H^s(\Omega)$  and for all  $x \in \Omega$ , we need an extension  $\bar{u}$  of  $u$  in a ball  $B_{R_0}$  containing  $\Omega$  such that  $\text{dist}(\partial\Omega, \partial B_{R_0}) > \rho h$ , and  $\bar{u} \in H^s(B_{R_0})$ . Then,

$$(\tilde{\mathcal{I}}_h u)(x) \equiv (\tilde{\mathcal{I}}_h \bar{u})(x) = \sum_{j \in A_x^h} \bar{u}(x_j^h) \phi_j^h(x), \quad \text{for } x \in \Omega, \tag{4.35}$$

is well defined. For an extension  $\bar{u}$ , we may use  $\bar{u} = Eu$ , where  $Eu$  was defined in (3.50). Thus,  $(\tilde{\mathcal{I}}_h u)(x)$  for  $x \in \Omega$  will depend on few values of  $\bar{u}(x_j^h)$ , where the particle  $x_j^h$  is just outside  $\Omega$ . We remark that  $\tilde{\mathcal{I}}_h u$  is not an interpolant of  $u$  in the usual sense, since, generally,  $\phi_j^h(x_i^h) \neq \delta_{ij}$ , and hence  $(\tilde{\mathcal{I}}_h u)(x_j^h) \neq u(x_j^h)$ .

We define the function

$$\xi_\alpha^h(x) = x^\alpha - \sum_{i \in A_x^h} (x_i^h)^\alpha \phi_i^h(x), \quad |\alpha| = k + 1, \quad \text{for } x \in \mathbb{R}^n. \tag{4.36}$$

We will also use

$$\xi_\alpha(x) \equiv \xi_\alpha^1(x) = x^\alpha - \sum_{i \in A_x^1} i^\alpha \phi(x - i), \quad |\alpha| = k + 1, \quad \text{for } x \in \mathbb{R}^n, \tag{4.37}$$

where  $A_x^1$  is  $A_x^h$  with  $h = 1$ . These functions will play an important role in the analysis presented in this subsection, as well as in Section 5. We note that  $\xi_\alpha(x)$  is the error between the polynomial  $x^\alpha$ , with  $|\alpha| = k + 1$ , and its interpolant when  $h = 1$ . In one dimension, we will write these functions as  $\xi_{k+1}^h(x)$  and  $\xi_{k+1}(x)$  respectively.

We begin with certain results about these functions. We first present some notation that will be used in these results. Let  $I_j^h$  be the *cell* centred at  $x_j^h$ , defined by

$$I_j^h = \{x : \|x - x_j^h\|_\infty \equiv \max_{i=1, \dots, n} |x_i - x_{j_i}| \leq h/2\}.$$



For each  $I_j^h$ , we define

$$A_j^h = \{k \in \mathbb{Z}^n : \eta_k^h \cap I_j^h \neq \emptyset\},$$

and

$$B_j^h = \{\cup_{k \in A_j^h} B_{Rh}(x_k^h)\} \cup I_j^h.$$

We note that the cardinality of  $A_j^h$  is finite, and is bounded independently of  $j$  and  $h$ . Further, there exists  $\bar{R} > 0$ , independent of  $j$  and  $h$ , such that  $B_j^h \subset \tilde{B}_j^h \equiv B_{\bar{R}h}(x_j^h)$  and  $\cup_{j \in \mathbb{Z}^n} \tilde{B}_j^h = \mathbb{R}^n$ .

**Lemma 4.1.**  $\xi_\alpha^h(x)$ , with  $|\alpha| = k + 1$ , is periodic, *i.e.*,

$$\xi_\alpha^h(x + x_j^h) = \xi_\alpha^h(x), \quad \text{for any } x_j^h. \tag{4.38}$$

*Proof.* We first note that

$$(x + x_j^h)^\alpha = x^\alpha + p(x; x_j^h), \tag{4.39}$$

where  $p(x; x_j^h)$  is a polynomial in  $x$  of degree  $\leq k$  with coefficients that depend on  $x_j^h$ . Now using (4.39), with  $x = x_i^h$ , and the fact that the  $\{\phi_i^h\}$  is translation-invariant and reproducing of order  $k$ , we get

$$\begin{aligned} \sum_{i \in \mathbb{Z}^n} (x_i^h)^\alpha \phi_i^h(x + x_j^h) &= \sum_{i \in \mathbb{Z}^n} (x_i^h)^\alpha \phi_{i-j}^h(x) \\ &= \sum_{i \in \mathbb{Z}^n} (x_{i+j}^h)^\alpha \phi_i^h(x) \\ &= \sum_{i \in \mathbb{Z}^n} (x_i^h + x_j^h)^\alpha \phi_i^h(x) \\ &= \sum_{i \in \mathbb{Z}^n} (x_i^h)^\alpha \phi_i^h(x) + \sum_{i \in \mathbb{Z}^n} p(x_i^h; x_j^h) \phi_i^h(x) \\ &= \sum_{i \in \mathbb{Z}^n} (x_i^h)^\alpha \phi_i^h(x) + p(x, x_j^h). \end{aligned} \tag{4.40}$$

From (4.36), (4.39) and (4.40), we get

$$\xi_\alpha^h(x + x_j^h) = x^\alpha - \sum_{i \in \mathbb{Z}^n} (x_i^h)^\alpha \phi_i^h(x) = \xi_\alpha^h(x),$$

which is the desired result. □

**Lemma 4.2.** Let  $\alpha = \alpha(i)$ ,  $i = 1, \dots, M_k$  be an enumeration of the multi-indices  $\alpha$  with  $|\alpha(i)| = k + 1$ . Let  $I_j^h$  be the cell centred at the particle  $x_j^h$ . Then, for  $d_\alpha \in \mathbb{R}$ , we have

$$\left\| \sum_{|\alpha|=k+1} \frac{1}{\alpha!} d_\alpha \xi_\alpha^h(x) \right\|_{H^1(I_j^h)}^2 = h^{2k+n} \mathbb{V}^T (A + h^2 B) \mathbb{V}, \tag{4.41}$$

where  $\mathbb{V} = [d_{\alpha(1)}, d_{\alpha(2)}, \dots, d_{\alpha(M_k)}]^T$  and  $A, B$  are  $M_k \times M_k$  matrices given by

$$A_{lm} = \int_I \frac{1}{\alpha(l)! \alpha(m)!} \nabla \xi_{\alpha(l)} \cdot \nabla \xi_{\alpha(m)} \, dx, \quad (4.42)$$

$$B_{lm} = \int_I \frac{1}{\alpha(l)! \alpha(m)!} \xi_{\alpha(l)} \xi_{\alpha(m)} \, dx, \quad (4.43)$$

respectively, and  $I = [-1/2, 1/2]^n$ .

**Note.** The matrices  $A$  and  $B$  are independent of  $I_j^h$ .

*Proof.* A simple scaling argument, used with (3.1), shows that

$$\xi_{\alpha} \left( \frac{x}{h} \right) = h^{-(k+1)} \xi_{\alpha}^h(x).$$

Now, using the periodicity of  $\xi_{\alpha}^h(x)$ , a standard scaling argument, and this identity, we have

$$\begin{aligned} \left\| \sum_{|\alpha|=k+1} \frac{1}{\alpha!} d_{\alpha} \nabla \xi_{\alpha}^h(x) \right\|_{H^0(I_j^h)}^2 &= \left\| \sum_{|\alpha|=k+1} \frac{1}{\alpha!} d_{\alpha} \nabla \xi_{\alpha}^h(x) \right\|_{H^0(I_0^h)}^2 \\ &= h^{2(k+1)} \left\| \sum_{|\alpha|=k+1} \frac{1}{\alpha!} d_{\alpha} \nabla \left[ \xi_{\alpha} \left( \frac{x}{h} \right) \right] \right\|_{H^0(I_0^h)}^2 \\ &= h^{2(k+1)} h^{n-2} \left\| \sum_{|\alpha|=k+1} \frac{1}{\alpha!} d_{\alpha} \nabla \xi_{\alpha}(y) \right\|_{H^0(I)}^2 \\ &= h^{2k+n} \mathbb{V}^T A \mathbb{V}. \end{aligned} \quad (4.44)$$

Using a similar argument, we have

$$\left\| \sum_{|\alpha|=k+1} \frac{1}{\alpha!} d_{\alpha} \xi_{\alpha}^h(x) \right\|_{H^0(I_j^h)}^2 = h^{2k+2+n} \mathbb{V}^T B \mathbb{V}.$$

Combining this identity with (4.44), we get

$$\left\| \sum_{|\alpha|=k+1} \frac{1}{\alpha!} d_{\alpha} \xi_{\alpha}^h(x) \right\|_{H^1(I_j^h)}^2 = h^{2k+n} \mathbb{V}^T (A + h^2 B) \mathbb{V},$$

which is the desired result.  $\square$

**Lemma 4.3.** Let  $I_j^h$  be the cell centred at  $x_j^h$ , and consider the corresponding set  $\tilde{B}_j^h$ . Suppose  $u \in H^{k+2+q}(\tilde{B}_j^h)$  with  $q > \frac{n}{2}$  when  $n \geq 2$ , and  $q = 0$  when  $n = 1$ . Then,

(a) for any  $\delta > 0$ ,

$$\begin{aligned} & \|u - \tilde{\mathcal{I}}_h u\|_{H^1(I_j^h)}^2 \\ & \leq (1 + \delta^2) \left\| \sum_{|\alpha|=k+1} \frac{1}{\alpha!} (D^\alpha u)(x_j^h) \xi_\alpha^h(x) \right\|_{H^1(I_j^h)}^2 \\ & \quad + \left(1 + \frac{1}{\delta^2}\right) Ch^{2k+2} \sum_{|\alpha|=k+2} \|D^\alpha u\|_{H^q(\tilde{B}_j^h)}^2. \end{aligned} \tag{4.45}$$

and

(b) for any  $\delta > 0$ ,

$$\begin{aligned} & \left\| \sum_{|\alpha|=k+1} \frac{1}{\alpha!} (D^\alpha u)(x_j^h) \xi_\alpha^h(x) \right\|_{H^1(I_j^h)}^2 \\ & \leq (1 + \delta^2) \|u - \tilde{\mathcal{I}}_h u\|_{H^1(I_j^h)}^2 \\ & \quad + \left(1 + \frac{1}{\delta^2}\right) Ch^{2k+2} \sum_{|\alpha|=k+2} \|D^\alpha u\|_{H^q(\tilde{B}_j^h)}^2. \end{aligned} \tag{4.46}$$

The proof of this result is based on Taylor’s theorem, a bound on the remainder in Taylor’s theorem, and a bound on the interpolant of the same remainder. We do not include the proof here, and refer to Babuška *et al.* (200x).

We will now study the interpolation error  $u - \tilde{\mathcal{I}}_h u$ , where  $u$  is a smooth function in  $\Omega$ . An interpolation error estimate for RKP shape functions, namely  $\|u - \tilde{\mathcal{I}}_h u\|_{W^{l,q}(\Omega)} \approx O(h^{k+1-l})$  for certain values of  $l$ , was proved in Han and Meng (2001); the same estimate for  $q = 2$  was proved in Liu, Li and Belytschko (1997). A similar order of convergence in the  $H^{1,\infty}$  norm was also obtained for MLS shape functions in Armentano (2002) and Armentano and Duran (2001). We note that the definitions of  $\tilde{\mathcal{I}}_h u$  for the RKP shape functions and MLS shape functions, presented in these papers, are slightly different from our definition as given in (4.35). From the proof of the our next result, we will obtain an estimate of  $\|u - \tilde{\mathcal{I}}_h u\|_{H^1(\Omega)}$  where the shape functions are reproducing of order  $k$ . Moreover, this theorem gives some information on the size of  $h^k \|u - \tilde{\mathcal{I}}_h u\|_{H^1(\Omega)}$ , which facilitates the selection of ‘good’ shape functions, which will be discussed later.

We now present the main result of this section. We define certain sets

which will be used in this result:

$$\begin{aligned} \bar{\mathcal{A}}^h &= \{k \in \mathbb{Z}^n : \Omega \cap \dot{I}_k^h \neq \emptyset\}, & \bar{\Omega}_h &= \cup_{j \in \bar{\mathcal{A}}^h} I_j^h, \\ \underline{\mathcal{A}}^h &= \{k \in \mathbb{Z}^n : I_k^h \subset \Omega\}, & \underline{\Omega}_h &= \cup_{j \in \underline{\mathcal{A}}^h} I_j^h, \\ \underline{\mathcal{B}}^h &= \{\cup_{j \in \underline{\mathcal{A}}^h} \tilde{B}_j^h\} \cup \Omega, & \bar{\mathcal{B}}^h &= \cup_{j \in \bar{\mathcal{A}}^h} \tilde{B}_j^h. \end{aligned}$$

It is clear that  $\underline{\Omega}_h \subset \Omega \subset \bar{\Omega}_h$ , and  $|\Omega - \underline{\Omega}_h| \rightarrow 0$ ,  $|\bar{\Omega}_h - \Omega| \rightarrow 0$  as  $h \rightarrow 0$ . Also  $\Omega \subset \underline{\mathcal{B}}^h \subset \bar{\mathcal{B}}^h$ , and  $|\underline{\mathcal{B}}^h - \Omega| \rightarrow 0$ ,  $|\bar{\mathcal{B}}^h - \Omega| \rightarrow 0$  as  $h \rightarrow 0$ .

**Theorem 4.4. (Babuška *et al.* (200x))** Let  $\bar{\lambda}$  be the largest eigenvalue of the matrix  $A$  given in (4.42). Suppose  $q > \frac{n}{2}$  when  $n \geq 2$ , and  $q = 0$  when  $n = 1$ . Then we have

$$\sup_{u \in H^{k+2+q}(\Omega)} \lim_{h \rightarrow 0} \frac{\|u - \tilde{\mathcal{I}}_h u\|_{H^1(\Omega)}^2}{h^{2k} Q_h(u)} = \bar{\lambda}, \tag{4.47}$$

where

$$Q_h(u) = |u|_{H^{k+1}(\Omega)}^2 + h \sum_{|\alpha|=k+2} \|D^\alpha u\|_{H^q(\Omega)}^2. \tag{4.48}$$

**Note.** In (4.47), we consider  $u \in H^{k+2+q}(\Omega)$  such that  $u \notin \mathcal{P}^k$ .

*Proof.* We will first prove that, for  $u \in H^{k+2+q}(\Omega)$ ,

$$\lim_{h \rightarrow 0} \frac{\|u - \tilde{\mathcal{I}}_h u\|_{H^1(\Omega)}^2}{h^{2k} Q_h(u)} = \frac{\int_{\Omega} V^T(x) A V(x) dx}{|u|_{H^{k+1}(\Omega)}^2}, \tag{4.49}$$

where

$$V^T(x) = [D^{\alpha(1)}u(x), D^{\alpha(2)}u(x), \dots, D^{\alpha(M_k)}u(x)],$$

and  $\alpha(i)$ ,  $1 \leq i \leq M_k$ , are the multi-indices with  $|\alpha(i)| = k + 1$ .

Let  $u \in H^{k+2+q}(\Omega)$ , and suppose  $\bar{u}$  is an extension of  $u$ , as discussed before. Since,  $\Omega \subset \bar{\Omega}_h$ , we have

$$\|u - \tilde{\mathcal{I}}_h u\|_{H^1(\Omega)}^2 \leq \|\bar{u} - \tilde{\mathcal{I}}_h \bar{u}\|_{H^1(\bar{\Omega}_h)}^2 = \sum_{j \in \bar{\mathcal{A}}^h} \|\bar{u} - \tilde{\mathcal{I}}_h \bar{u}\|_{H^1(I_j^h)}^2.$$

Therefore, using (4.45), (4.41), and recalling that  $\bar{\mathcal{B}}^h = \cup_{j \in \bar{\mathcal{A}}^h} \tilde{B}_j^h$ , we get for any  $\delta > 0$ ,

$$\begin{aligned} \|u - \tilde{\mathcal{I}}_h u\|_{H^1(\Omega)}^2 &\leq (1 + \delta^2) \sum_{j \in \bar{\mathcal{A}}^h} \left\| \sum_{|\alpha|=k+1} \frac{1}{\alpha!} (D^\alpha \bar{u})(x_j^h) \xi_\alpha^h(x) \right\|_{H^1(I_j^h)}^2 \\ &\quad + \left(1 + \frac{1}{\delta^2}\right) C h^{2k+2} \sum_{j \in \bar{\mathcal{A}}^h} \sum_{|\alpha|=k+2} \|D^\alpha \bar{u}\|_{H^q(\tilde{B}_j^h)}^2 \end{aligned}$$

$$\begin{aligned} &\leq (1 + \delta^2)h^{2k} \sum_{j \in \bar{A}^h} h^n V_j^T (A + h^2 B) V_j \\ &\quad + \left(1 + \frac{1}{\delta^2}\right) Ch^{2k+2} \sum_{|\alpha|=k+2} \|D^\alpha \bar{u}\|_{H^q(\bar{B}^h)}^2, \end{aligned} \tag{4.50}$$

where

$$V_j^T = [D^{\alpha(1)} \bar{u}(x_j^h), D^{\alpha(2)} \bar{u}(x_j^h), \dots, D^{\alpha(M_k)} \bar{u}(x_j^h)].$$

Therefore, dividing (4.50) by  $h^{2k} Q_h(u)$ , where  $Q_h(u)$  is defined in (4.48), we get

$$\begin{aligned} \frac{\|u - \tilde{\mathcal{I}}_h u\|_{H^1(\Omega)}^2}{h^{2k} Q_h(u)} &\leq (1 + \delta^2) \frac{\sum_{j \in \bar{A}^h} h^n V_j^T (A + h^2 B) V_j}{Q_h(u)} \\ &\quad + \left(1 + \frac{1}{\delta^2}\right) Ch^2 \frac{\sum_{|\alpha|=k+2} \|D^\alpha \bar{u}\|_{H^q(\bar{B}^h)}^2}{Q_h(u)}. \end{aligned} \tag{4.51}$$

A typical term of the quadratic form  $V_j^T (A + h^2 B) V_j$  is

$$D^{\alpha(i)} \bar{u}(x_j^h) (A_{il} + h^2 B_{il}) D^{\alpha(l)} \bar{u}(x_j^h).$$

Since

$$\lim_{h \rightarrow 0} \sum_{j \in \bar{A}^h} h^n D^{\alpha(i)} \bar{u}(x_j^h) A_{il} D^{\alpha(l)} \bar{u}(x_j^h) = \int_{\Omega} D^{\alpha(i)} u(x) A_{il} D^{\alpha(l)} u(x) \, dx$$

and

$$\lim_{h \rightarrow 0} h^2 \sum_{j \in \bar{A}^h} h^n D^{\alpha(i)} \bar{u}(x_j^h) B_{il} D^{\alpha(l)} \bar{u}(x_j^h) = 0,$$

we have

$$\lim_{h \rightarrow 0} \sum_{j \in \bar{A}^h} h^n V_j^T (A + h^2 B) V_j = \int_{\Omega} V^T(x) A V(x) \, dx. \tag{4.52}$$

Since  $|\bar{B}^h - \Omega| \rightarrow 0$  as  $h \rightarrow 0$ , we have

$$\lim_{h \rightarrow 0} \sum_{|\alpha|=k+2} \|D^\alpha \bar{u}\|_{H^q(\bar{B}^h)}^2 = \sum_{|\alpha|=k+2} \|D^\alpha u\|_{H^q(\Omega)}^2. \tag{4.53}$$

Also  $\lim_{h \rightarrow 0} Q_h(u) = |u|_{H^{k+1}(\Omega)}$ . Thus, for any  $\delta > 0$ , using (4.52) and (4.53) in (4.51), we get

$$\limsup_{h \rightarrow 0} \frac{\|u - \tilde{\mathcal{I}}_h u\|_{H^1(\Omega)}^2}{h^{2k} Q_h(u)} \leq (1 + \delta^2) \frac{\int_{\Omega} V^T(x) A V(x) \, dx}{|u|_{H^{k+1}(\Omega)}^2},$$

and, since  $\delta > 0$  is arbitrary, we have

$$\limsup_{h \rightarrow 0} \frac{\|u - \tilde{\mathcal{I}}_h u\|_{H^1(\Omega)}^2}{h^{2k} Q_h(u)} \leq \frac{\int_{\Omega} V^T(x)AV(x) \, dx}{|u|_{H^{k+1}(\Omega)}^2}. \tag{4.54}$$

Following the argument leading to (4.54), but using  $\underline{A}^h$ ,  $\underline{B}^h$ , and (4.46) instead of  $\bar{A}^h$ ,  $\bar{B}^h$ , and (4.45), respectively, we can also show that

$$\frac{\int_{\Omega} V^T(x)AV(x) \, dx}{|u|_{H^{k+1}(\Omega)}^2} \leq \liminf_{h \rightarrow 0} \frac{\|u - \tilde{\mathcal{I}}_h u\|_{H^1(\Omega)}^2}{h^{2k} Q_h(u)}. \tag{4.55}$$

Combining (4.54) and (4.55), we see that

$$\lim_{h \rightarrow 0} \frac{\|u - \tilde{\mathcal{I}}_h u\|_{H^1(\Omega)}^2}{h^{2k} Q_h(u)}$$

exists, and

$$\lim_{h \rightarrow 0} \frac{\|u - \tilde{\mathcal{I}}_h u\|_{H^1(\Omega)}^2}{h^{2k} Q_h(u)} = \frac{\int_{\Omega} V^T(x)AV(x) \, dx}{|u|_{H^{k+1}(\Omega)}^2},$$

which is (4.49).

Since  $\bar{\lambda}$  is the largest eigenvalue of the matrix  $A$ , from the usual variational characterization of eigenvalues we have

$$\int_{\Omega} V^T(x)AV(x) \, dx \leq \bar{\lambda} \int_{\Omega} \sum_{i=1}^{M_k} |D^{\alpha(i)}u(x)|^2 \, dx = \bar{\lambda}|u|_{H^{k+1}(\Omega)}^2.$$

Thus, from (4.49) we get

$$\lim_{h \rightarrow 0} \frac{\|u - \tilde{\mathcal{I}}_h u\|_{H^1(\Omega)}^2}{h^{2k} Q_h(u)} \leq \bar{\lambda}, \quad \text{for any } u \in H^{k+2+q}(\Omega).$$

Hence

$$\sup_{u \in H^{p+2+q}(\Omega)} \lim_{h \rightarrow 0} \frac{\|u - \tilde{\mathcal{I}}_h u\|_{H^1(\Omega)}^2}{h^{2k} Q_h(u)} \leq \bar{\lambda}. \tag{4.56}$$

Let  $\bar{v} = [v_1, v_2, \dots, v_{M_k}]^T$  be an eigenvector of  $A$  corresponding to  $\bar{\lambda}$ . Then it is easily seen that there is a  $u \in \mathcal{P}^{k+1}$  such that the vector  $V(x) = \bar{v}$ . For this particular  $u$ , we have

$$\frac{\int_{\Omega} V^T(x)AV(x) \, dx}{|u|_{H^{k+1}(\Omega)}^2} = \bar{\lambda}.$$

Hence, from (4.56) we conclude that

$$\sup_{u \in H^{k+2+q}(\Omega)} \lim_{h \rightarrow 0} \frac{\|u - \tilde{\mathcal{I}}_h u\|_{H^1(\Omega)}^2}{h^{2k} Q_h(u)} = \bar{\lambda},$$

which is the desired result. □

**Remark 29.** We know from (4.35) that the interpolant of a smooth function depends on its extension to  $\mathbb{R}^n$ . But it is clear from the proof of Theorem 4.4 that (4.47) is valid for any extension satisfying (3.50).

**Remark 30.** We note that the same result holds for the  $H^1$ -seminorm of the interpolation error, *i.e.*, for  $q > \frac{n}{2}$  when  $n \geq 2$ , and  $q = 0$  when  $n = 1$ , we have

$$\sup_{u \in H^{k+2+q}(\Omega)} \lim_{h \rightarrow 0} \frac{|u - \tilde{\mathcal{I}}_h u|_{H^1(\Omega)}^2}{h^{2k} [ \|u\|_{H^{k+1}(\Omega)}^2 + h \sum_{|\alpha|=k+2} \|D^\alpha u\|_{H^q(\Omega)}^2 ]} = \bar{\lambda}.$$

**Remark 31.** From (4.51) in the proof of Theorem 4.4, we can obtain an interpolation error estimate,

$$\|u - \tilde{\mathcal{I}}_h u\|_{H^1(\Omega)} \leq Ch^k \|u\|_{H^{k+2+q}(\Omega)},$$

where  $C$  may depend on  $\Omega$ , but is independent of  $u$  and  $h$ . We note, however, that this is not the optimal error estimate. For an outline of the proof, see Babuška *et al.* (200x).

We have seen in Remark 31 that, if the particle shape functions are reproducing of order  $k$ , then for a smooth function  $u$ ,

$$\|u - \tilde{\mathcal{I}}_h u\|_{H^1(\Omega)} \approx O(h^k),$$

where  $\tilde{\mathcal{I}}_h u$  is the interpolation of  $u$  as defined in (4.35). There are many classes of shape functions that have these properties. We have seen in Section 4.1 that translation-invariant RKP shape functions depend on the weight function  $w(x)$ , and different choices of  $w(x)$  will generate different classes of such shape functions.

We will assess the approximability of a family  $\{\phi_j^h\}$  of shape functions by the size of  $\bar{\lambda}$ , the largest eigenvalue of the matrix  $A$  defined in (4.42). We note that  $\bar{\lambda}$  is computable, and depends only on the basic shape function  $\phi(x)$ . We emphasize that  $\bar{\lambda}$  does not depend on  $u$  or on  $h$ . From (4.47), we know that

$$\frac{\|u - \tilde{\mathcal{I}}_h u\|_{H^1(\Omega)}}{h^p \sqrt{Q_h(u)}} \lesssim \sqrt{\bar{\lambda}}, \quad \text{for small } h.$$

Thus we see that  $\bar{\lambda}$  is a useful measure of the approximability of the family  $\{\phi_j^h\}$ , determined from the basic shape function  $\phi(x)$ .

We will illustrate our selection scheme in one dimension, and will rank the shape functions according to their approximability. In one dimension,

$$\bar{\lambda} = \left( \frac{|\xi_{k+1}|_{H^1(0,1)}}{(k+1)!} \right)^2.$$

In the rest of this paper we will suppress  $H^1(0,1)$  in  $|\xi_{k+1}|_{H^1(0,1)}$ , and instead write  $|\xi_{k+1}|_1$ .

We considered four different classes of RKP shape functions, reproducing of order 1, corresponding to four different weight functions  $w(x)$ . These weight functions were given by (4.2) with  $\delta = 2$ , (4.3), and (4.4) with  $l = 2, 4$ . We then computed  $|\xi_{k+1}|_1$  for each of these four classes of shape functions for  $R = 1.7$ ; we obtained

$$|\xi_{k+1}|_1 = \begin{cases} 0.237, & \text{for } w(x) \text{ in (4.4), } l = 2, \\ 0.203, & \text{for } w(x) \text{ in (4.2), } \delta = 2, \\ 0.095, & \text{for } w(x) \text{ in (4.3),} \\ 0.029, & \text{for } w(x) \text{ in (4.4), } l = 4. \end{cases}$$

We choose the RKP shape functions corresponding to  $w(x)$  given in (4.4) with  $l = 4$ , since these shape functions yield the smallest value of  $|\xi_{k+1}|_1$ . We note that the value of  $|\xi_{k+1}|_1$  depends strongly on  $R$ , and the shape function corresponding to  $w(x)$  given in (4.4) with  $l = 4$  may not be our choice for other values of  $R$ . We refer to Babuška *et al.* (200x) for further discussion on this issue.

To validate our criterion for selection of the shape functions, we have considered the function  $u(x) = x^4$  on the interval  $\Omega = (0,1)$  and computed the error  $|u - \tilde{\mathcal{I}}_h u|_{H^1(\Omega)}$ .  $\tilde{\mathcal{I}}_h u$  is the interpolant of  $u$  with respect to the four classes of RKP shape functions described in the last paragraph, with  $h = 1/n$ ,  $n = 40, 50, \dots, 100$ . We note that the definition of  $\tilde{\mathcal{I}}_h u$  requires the values of  $u(x)$  in a small neighbourhood of  $\Omega$ , and we have extended  $u = x^4$  outside  $\Omega$  by itself.

We summarize the results in Table 4.1. We note that columns 2–5 correspond to different classes of RKP shape functions constructed using different weight functions  $w(x)$ ; for example, the column headed ‘Cubic spline’ refers to the cubic spline weight function given by (4.3). It is clear that the error  $|u - \tilde{\mathcal{I}}_h u|_{H^1(\Omega)}$  can be ranked according to the size of  $|\xi_2|_1$  for the four choices of  $w(x)$  considered here with  $R = 1.7$ ; the error and  $|\xi_2|_1$  are both minimal when  $w(x)$  is the conical weight function with  $l = 4$ .

This selection scheme is based on (4.47), and we know from Remark 29 that (4.47) is valid for any extension. We refer to Babuška *et al.* (200x) for an experimental illustration of this fact. We remark that this selection scheme is also valid for the projection error, which will be indicated by our results in the next section.



Table 4.1. The  $H^1$ -seminorm of the error,  $|u - \tilde{I}_h u|_{H^1(\Omega)}$ , where  $\tilde{I}_h u$  is the interpolant of  $u(x) = x^4$  using RKP shape functions that are reproducing of order 1, corresponding to different weight functions  $w(x)$ . The radius of support of  $\omega(x)$  is  $R = 1.7$ .

$n$	$ u - \tilde{I}_h u _{H^1(\Omega)}$			
	Conical: $l = 2$	Gauss: $\delta = 2$	Cubic spline	Conical: $l = 4$
40	1.607e-2	1.376e-2	6.435e-3	2.283e-3
50	1.281e-2	1.096e-2	5.130e-3	1.730e-3
60	1.066e-2	9.112e-3	4.267e-3	1.396e-3
70	9.126e-3	7.800e-3	3.653e-3	1.172e-3
80	7.980e-3	6.819e-3	3.194e-3	1.012e-3
90	7.090e-3	6.058e-3	2.838e-3	8.908e-4
100	6.379e-3	5.449e-3	2.553e-3	7.962e-4

### 5. Superconvergence of the gradient of the solution in $L_2$

Superconvergence is an important feature of finite element methods, which allows an accurate approximation of the derivatives of the solution of the underlying BVP. In this section, we will discuss the idea of superconvergence when particle shape functions are used to approximate the solution of a BVP. We will consider uniformly distributed particles and the associated particle shape functions, which were developed in Sections 3.1 and 3.2. For uniformly distributed particles, a careful analysis in one dimension can be easily generalized to higher dimensions. Thus, in this section, we present the results in one dimension, thereby avoiding some technical details arising in the higher-dimensional analysis.

We will use the notation introduced in Section 3.1, but restricted to one dimension, *i.e.*, for  $h > 0$ , we consider  $x_j^h = jh, j \in \mathbb{Z}$ , and the corresponding shape function  $\phi_j^h$  defined in (3.1). We assume that the shape functions are reproducing of order  $k$ . We use the following notation:

$$I_j^h = (x_j^h, x_{j+1}^h), \quad A_j^h = \{m \in \mathbb{Z} : \eta_m^h \cap I_j^h \neq \emptyset\};$$

$$I_j = (j, j + 1), \quad A_j \equiv A_j^1 \quad (\text{with } h = 1).$$

We assume that

$$\text{card}(A_j) \leq \kappa,$$

or equivalently,

$$\text{card}(A_j^h) \leq \kappa,$$

where  $\kappa$  is independent of  $j$  and  $h$ . We assume that the basic shape function  $\phi(x)$  is such that, for any  $v(x) = \sum_{i \in \mathbb{Z}} c_i \phi_i(x)$  for  $x \in I_0$ , there exist positive constants  $C_1, C_2$ , independent of  $v$ , but possibly depending on  $\kappa$ , such that

$$C_1 \sum_{j \in A_0} c_j^2 \leq \int_{I_0} v^2 \, dx \leq C_2 \sum_{j \in A_0} c_j^2. \tag{5.1}$$

This implies that the functions  $\{\phi_i(x)\}_{i \in A_0}$  are linearly independent in  $I_0$ , that is,

$$\sum_{j \in A_0} c_j \phi_j(x) = 0, \quad x \in I_0 \text{ implies } c_j = 0, j \in A_0.$$

Throughout this section, we use  $C, C_1, C_2$  as generic constants, which will have different values in different places.

Consider  $\Omega = (-c, d) \subset \mathbb{R}$ . Let  $u_0 \in H^1(\Omega)$  be the solution of the Neumann problem

$$B(u_0, v) = \mathcal{F}(v), \quad \text{for all } v \in H^1(\Omega), \tag{5.2}$$

where

$$B(u, v) = \int_{\Omega} (u'v' + uv) \, dx \quad \text{and} \quad \mathcal{F}(v) = \int_{\Omega} f v \, dx$$

as in (2.4) and (2.5). We will often use the notation  $B^F(u, v)$  to denote the above bilinear form, where the  $\Omega$  is replaced by another domain  $F$ .

Let  $u_h \in V_{\Omega,h}^{k,q}$  be the solution of

$$B(u_h, v) = \mathcal{F}(v), \quad \text{for all } v \in V_{\Omega,h}^{k,q}, \tag{5.3}$$

where  $V_{\Omega,h}^{k,q}$  was defined in (3.54). It is clear from (5.2) and (5.3) that

$$B(u_0 - u_h, v) = 0, \quad \text{for all } v \in V_{\Omega,h}^{k,q}, \tag{5.4}$$

and we easily have

$$\|u_h\|_{H^1(\Omega)} \leq \|u_0\|_{H^1(\Omega)}. \tag{5.5}$$

Recall that the functions in  $V_{\Omega,h}^{k,q}$  are restrictions of the functions in  $S_h \equiv V_h^{k,q}$  on  $\Omega$  (cf. (3.2) and (3.54)). Thus (5.4) is true when  $V_{\Omega,h}^{k,q}$  is replaced by  $S_h$ .

We assume that, for any  $\rho > 0$ ,

$$\|u_0 - u_h\|_{L_2(B_\rho(0))} \leq Ch^{k+1} \|u_0\|_{H^{k+1}(\Omega)} \rho^{\frac{1}{2}}, \tag{5.6}$$

and there are positive constants  $C_1, C_2$ , independent of  $u, h$ , and  $\rho$ , such that

$$C_1 h^k \rho^{\frac{1}{2}} \leq \frac{\|u'_0 - u'_h\|_{L_2(B_\rho(0))}}{\|u_0\|_{H^{k+1}(\Omega)}} \leq C_2 h^k \rho^{\frac{1}{2}}, \tag{5.7}$$

where  $B_\rho(0) = \{x : |x| < \rho\}$  and  $B_\rho(0) \subset \Omega$ . We will write  $B_\rho \equiv B_\rho(0)$  throughout this section.

The main goal of this section is to investigate the error  $u'(x) - u'_h(x)$  in a neighbourhood of  $x = 0$ , *i.e.*, for  $x \in B_H \subset \subset \Omega$  and  $H = h^\gamma$ ,  $\gamma < 1$ , where  $\gamma$  will be chosen later. We will prove the following result.

**Theorem 5.1.** Suppose  $u_0$  and  $u_h$  satisfy (5.6) and (5.7), and let  $e_h = u_0 - u_h$ . Moreover, assume that  $u_0 \in W_\infty^{k+2}(B_{2H})$ . Then, for sufficiently small  $h$ , there exists  $\epsilon^* > 0$  such that

$$\frac{\|e'_h - T(u_0)\xi_{k+1}^h\|_{L_2(B_H)}}{\|e'_h\|_{L_2(B_H)}} \leq Ch^{\epsilon^*},$$

where

$$T(u_0) = \frac{u_0^{(k+1)}(0)}{(k+1)!} \quad \text{and} \quad \xi_{k+1}^h(x) = h^{k+1}\xi_{k+1}\left(\frac{x}{h}\right);$$

$\xi_{k+1}$  is defined in (4.37).

**Remark 32.** Theorem 5.1 is a superconvergence result. It shows that

$$\|e'_h - T(u_0)\xi_{k+1}^h\|_{L_2(B_H)} \ll \|e'_h\|_{L_2(B_H)}.$$

This allows us, for example, to analyse the effectiveness of an error estimator, as was done in Babuška, Strouboulis, Upadhyay and Gangaraj (1996) and Babuška and Strouboulis (2001).

Since the all the results in this paper have been presented in terms of  $L_2$ -based norms (*i.e.*, in terms of the usual Sobolev norms), we also present this result in terms of  $L_2$ -based norms. Superconvergence in  $L_\infty$  will be addressed in a forthcoming paper. Assuming superconvergence in  $L_\infty$ , the superconvergence points and superconvergence recoveries in the case of particle shape functions can be obtained as in Babuška and Strouboulis (2001). At the end of this section, we will see an example where the superconvergence points are distributed in a different way to that of the classical FEM.

**Remark 33.** The essential aspects of superconvergence analysis in the classical FEM are interior estimates, developed in Nitsche and Schatz (1974), Schatz and Wahlbin (1995) and Wahlbin (1995). This analysis strongly utilizes the polynomial character of the shape functions. Here, in the case of particle shape functions, we had to develop another approach to the analysis of superconvergence, which is based on weighted Sobolev spaces. The main idea of the proof of our superconvergence result is to show that *locally* the approximation error is asymptotically the same as the error in the interpolation of a polynomial of degree  $k + 1$  by particle shape functions. The analysis is technical; we present the main idea of this analysis in this section.

**Remark 34.** Assumptions (5.6) and (5.7) are directly related to the control of pollution, as in the FEM. The assumption  $u_0 \in W_\infty^{k+1}(B_\rho)$  is analogous to the assumption in the FEM (see Babuška *et al.* (1996) and Babuška and Strouboulis (2001)).

To prove Theorem 5.1, we will first develop certain ideas and establish several technical results. To that end, for given parameters  $H = h^\gamma$ , with  $\gamma < 1$ , and  $\alpha \geq 1$ , we define the function  $g(x)$  by

$$g(x) = \begin{cases} 1, & -H \leq x \leq H, \\ e^{-\alpha(x-H)}, & x > H, \\ e^{\alpha(H+x)}, & x < -H, \end{cases} \quad (5.8)$$

where  $\alpha$  is such that  $\alpha h < 1$ , and will be chosen later. We note that a proper choice of  $\gamma$  and  $\alpha$  is crucial for the analysis presented in this section. Often, we will use  $g \equiv g(x)$ ,  $g_i \equiv g(x_i^h)$  and  $g_{i+\frac{1}{2}} \equiv g(x_i^h + h/2)$ .

#### *Generalized interpolant and certain norm estimates*

We first introduce the idea of a generalized interpolant of a function  $u$ , which is different to the  $\tilde{I}_h u$  defined in Section 4.2. Let  $\tilde{I}_0 \equiv I_{-1} \cup I_0 \cup \{0\} = (-1, 1)$  and  $\tilde{A}_0 \equiv A_{-1} \cup A_0$ . Then, from (5.1), it is clear that there are positive constants  $C_1, C_2$ , independent of  $v = \sum_{i \in \mathbb{Z}} c_i \phi_i(x)$ , but possibly depending on  $\kappa$ , such that

$$C_1 \sum_{j \in \tilde{A}_0} c_j^2 \leq \int_{\tilde{I}_0} v^2 dx \leq C_2 \sum_{j \in \tilde{A}_0} c_j^2, \quad (5.9)$$

which implies that  $\{\phi_i(x)\}_{i \in \tilde{A}_0}$  are also linearly independent in  $\tilde{I}_0$ . We define  $\psi_0(x) = \sum_{i \in \tilde{A}_0} a_i \phi_i(x)$  with  $\text{supp } \psi_0 = \overline{\tilde{I}_0}$  (closure of  $\tilde{I}_0$ ), such that

$$\begin{aligned} \int_{\tilde{I}_0} \psi_0(x) \phi_0(x) dx &= 1, \\ \int_{\tilde{I}_0} \psi_0(x) \phi_j(x) dx &= 0, \quad \text{for all } j \in \tilde{A}_0, \quad j \neq 0. \end{aligned} \quad (5.10)$$

Using (5.9), we can show that

$$\|\psi_0\|_{L_2(\tilde{I}_0)} \leq C. \quad (5.11)$$

We also note that, since  $\{\phi_i(x)\}_{i \in \tilde{A}_0}$  form a partition unity on  $\tilde{I}_0$ , from (5.10) we have

$$\int_{\tilde{I}_0} \psi_0(x) dx = \int_{\tilde{I}_0} \psi_0(x) \sum_{i \in \tilde{A}_0} \phi_i(x) dx = \int_{\tilde{I}_0} \psi_0(x) \phi_0(x) dx = 1. \quad (5.12)$$

Let  $\psi_i^h(x) = \psi_0(\frac{x}{h} - i)$ . Then  $\text{supp } \psi_i^h = \overline{\tilde{I}_i^h}$ , where  $\tilde{I}_i^h = (x_{i-1}^h, x_{i+1}^h)$ . Note

that  $\cup_{i \in \mathbb{Z}} \tilde{I}_i^h = \mathbb{R}$ . Now, for a given  $v \in L_2(\mathbb{R})$ , we define the *generalized interpolant* of  $v$  as

$$\tilde{\mathcal{I}}_h^* v(x) = \sum_{i \in \mathbb{Z}} \Psi_i^h(v) \phi_i^h(x), \tag{5.13}$$

where

$$\Psi_i^h(v) = \frac{1}{h} \int_{\tilde{I}_i^h} \psi_i^h(x) v(x) \, dx. \tag{5.14}$$

We note that  $\tilde{\mathcal{I}}_h^* v(x)$  depends on the  $v(y)$  for  $y \in \cup_{i \in A^h(x)} \tilde{I}_i^h$ , where  $A^h(x) = \{l \in \mathbb{Z} : x \in \eta_l^h\}$ . We also define

$$\tilde{A}_i^h = \{m \in \mathbb{Z} : \eta_m^h \cap \tilde{I}_i^h \neq \emptyset\}.$$

**Lemma 5.2.** Suppose  $v(x) = \sum_{i \in \mathbb{Z}} c_i^h \phi_i^h(x)$ . Then

$$c_i^h = \Psi_i^h(v), \quad \text{and} \tag{5.15}$$

$$\tilde{\mathcal{I}}_h^* v(x) = v(x). \tag{5.16}$$

*Proof.* From (5.10) and the definition of  $\Psi_i^h(v)$  in (5.14),  $i \in \mathbb{Z}$ , we have

$$\begin{aligned} \Psi_i^h(v) &= \frac{1}{h} \int_{\tilde{I}_i^h} \psi_i^h(x) v(x) \, dx \\ &= \frac{1}{h} \int_{\tilde{I}_i^h} \psi_0\left(\frac{x}{h} - i\right) \sum_{j \in \tilde{A}_i^h} c_j^h \phi\left(\frac{x}{h} - j\right) \, dx \\ &= \int_{\tilde{I}_0} \psi_0(y) \sum_{j \in \tilde{A}_i^h} c_j^h \phi_{j-i}(y) \, dy \\ &= c_i^h, \end{aligned}$$

which is (5.15). Now using (5.15) in (5.13), we get (5.16). □

**Remark 35.** We note that if  $v$  is a local linear combination of the  $\{\phi_i^h\}$ , *i.e.*, in a bounded open interval, then  $\tilde{\mathcal{I}}_h^* v = v$  only in the interior of that open interval. More precisely,  $\tilde{\mathcal{I}}_h^* v = v$  in an interval  $I$  if  $v$  is a linear combination of the  $\{\phi_i\}$  in  $\cup_{x \in I} \cup_{i \in A^h(x)} \tilde{I}_i^h$ .

We will use the following result later.

**Lemma 5.3.** Let  $\Omega$  be a bounded interval, and suppose  $u \in L_2(\Omega)$ . Then

$$\|\tilde{\mathcal{I}}_h^* E u\|_{H^1(\mathbb{R})} \leq C h^{-1} \|u\|_{L_2(\Omega)},$$

where  $E$  is the extension operator satisfying (3.50).

*Proof.* We first note that the extension  $Eu$  of  $u$  satisfies  $\|Eu\|_{L_2(\mathbb{R})} \leq C\|u\|_{L_2(\Omega)}$ . Now, from (5.13) and (4.32),

$$\begin{aligned} \|\tilde{\mathcal{I}}_h^* Eu\|_{H^1(\tilde{I}_i^h)}^2 &\leq C \sum_{j \in \tilde{A}_i^h} |\Psi_j^h(Eu)|^2 \|\phi_j^h\|_{H^1(\tilde{\eta}_j)}^2 \\ &\leq Ch^{-1} \sum_{j \in \tilde{A}_i^h} |\Psi_j^h(Eu)|^2, \end{aligned} \tag{5.17}$$

where  $C$  depends on  $\kappa$ ; and using the Schwartz inequality on (5.14) with  $v = Eu$ , and a scaling argument, we get

$$\begin{aligned} |\Psi_j^h(Eu)|^2 &\leq \frac{1}{h^2} \left( \int_{\tilde{I}_j^h} \psi_j^h(x) Eu(x) \, dx \right)^2 \\ &\leq \frac{1}{h^2} \left[ \int_{\tilde{I}_j^h} (\psi_j^h)^2 \, dx \right] \left[ \int_{\tilde{I}_j^h} (Eu)^2 \, dx \right] \\ &\leq \frac{1}{h} \|\psi_0\|_{L_2(\tilde{I}_0)}^2 \left[ \int_{\tilde{I}_j^h} (Eu)^2 \, dx \right]. \end{aligned} \tag{5.18}$$

Thus, from (5.17), (5.18), and the fact that  $\|Eu\|_{L_2(\mathbb{R})} \leq C\|u\|_{L_2(\Omega)}$ , we have

$$\|\tilde{\mathcal{I}}_h^* Eu\|_{H^1(\mathbb{R})}^2 \leq Ch^{-2} \|u\|_{L_2(\Omega)}^2,$$

which is the desired result. □

**Remark 36.** We can also show that

$$\|\tilde{\mathcal{I}}_h^* Eu\|_{L_2(\mathbb{R})} \leq C\|u\|_{L_2(\Omega)},$$

using the same arguments as in the proof of Lemma 5.3.

Consider the function  $v(x) = \sum_{i \in A_j^h} c_i^h \phi_i^h(x)$  on  $I_j^h$ . Then, using scaling, translation, and (5.1), we have

$$C_1 h \sum_{j \in A_j^h} (c_i^h)^2 \leq \int_{I_j^h} v^2 \, dx \leq C_2 h \sum_{j \in A_j^h} (c_i^h)^2, \tag{5.19}$$

where  $C_1, C_2$  are positive constants, independent of  $h$  and  $j$ , but possibly depending on  $\kappa$ . Using (5.19), we can show that if  $v(x) = \sum_{i \in \mathbb{Z}} c_i^h \phi_i^h(x) = 0$  in  $L_2$ , then  $c_i^h = 0$ , for all  $i \in \mathbb{Z}$ , i.e.,  $\{\phi_i^h\}$  are linearly independent.

We will now prove certain lower bounds for  $\int_{I_j^h} g v^2 \, dx$  and  $\int_{I_j^h} g v'^2 \, dx$ , where  $g(x)$  has been defined before. We first prove the following inequality.

**Lemma 5.4.** Let  $i_0, i_1$  be integers such that  $i_0 < i_1$ , and suppose  $\{c_i\}_{i=i_0}^{i_1}$  are real numbers. Then there exists a positive constant  $C$ , depending only on  $i_1 - i_0$ , such that, for any  $k, i_0 \leq k \leq i_1$ , we have

$$\sum_{i=i_0}^{i_1} g_{i+\frac{1}{2}}(c_i - c_k)^2 \leq C \sum_{i=i_0}^{i_1-1} g_{i+\frac{1}{2}}(c_{i+1} - c_i)^2. \tag{5.20}$$

*Proof.* Suppose the integers  $i_0, i_1$  are such that  $H < i_0 h < i_1 h$ , where  $H = h^\gamma, \gamma < 1$ . Then

$$\sum_{i=i_0}^{i_1} g_{i+\frac{1}{2}}(c_i - c_k)^2 = \sum_{i=i_0}^{k-1} g_{i+\frac{1}{2}}(c_i - c_k)^2 + \sum_{i=k+1}^{i_1} g_{i+\frac{1}{2}}(c_i - c_k)^2. \tag{5.21}$$

We first note that

$$\begin{aligned} & \sum_{i=k+1}^{i_1} g_{i+\frac{1}{2}}(c_i - c_k)^2 \\ & \leq C \sum_{i=k+1}^{i_1} \sum_{j=k}^{i-1} g_{i+\frac{1}{2}}(c_{j+1} - c_j)^2 \\ & = C \sum_{i=k+1}^{i_1} \sum_{j=k}^{i-1} \left(1 + \frac{g_{i+\frac{1}{2}} - g_{j+\frac{1}{2}}}{g_{j+\frac{1}{2}}}\right) g_{j+\frac{1}{2}}(c_{j+1} - c_j)^2. \end{aligned} \tag{5.22}$$

But from the definition of  $g(x)$  in (5.8), we have

$$\left(1 + \frac{g_{i+\frac{1}{2}} - g_{j+\frac{1}{2}}}{g_{j+\frac{1}{2}}}\right) \leq e^{-\alpha(i-j)h} \leq e^{\alpha(i_1-i_0)h} \leq C, \tag{5.23}$$

and using this in (5.22), we get

$$\begin{aligned} \sum_{i=k+1}^{i_1} g_{i+\frac{1}{2}}(c_i - c_k)^2 & \leq C \sum_{i=k+1}^{i_1} \sum_{j=k}^{i-1} g_{j+\frac{1}{2}}(c_{j+1} - c_j)^2 \\ & \leq C \sum_{j=k}^{i_1-1} g_{j+\frac{1}{2}}(c_{j+1} - c_j)^2, \end{aligned} \tag{5.24}$$

where  $C$  depends on  $(i_1 - i_0)$ .

Using similar arguments we can show that

$$\sum_{i=i_0}^{k-1} g_{i+\frac{1}{2}}(c_k - c_i)^2 \leq C(k - 1 - i_0) \sum_{j=i_0}^{k-1} g_{j+\frac{1}{2}}(c_{j+1} - c_j)^2, \tag{5.25}$$

where  $C$  depends on  $(i_1 - i_0)$ . Therefore, combining (5.21), (5.24), and

(5.25), we have

$$\sum_{i=i_0}^{i_1} g_{i+\frac{1}{2}}(c_i - c_k)^2 \leq C \sum_{i=i_0}^{i_1} g_{j+\frac{1}{2}}(c_{j+1} - c_j)^2, \quad (5.26)$$

where  $C$  depends on  $(i_i - i_0)$ . Using similar arguments, we can prove (5.26) for all integers  $i_0, i_1$  such that  $i_0 < i_1$ .  $\square$

**Lemma 5.5.** Suppose  $v(x) = \sum_{i \in \mathbb{Z}} c_i^h \phi_i^h(x)$ . Then:

- (a) there are positive constants  $C_1, C_2$ , independent of  $v, h$  and  $j$ , but possibly depending on  $\kappa$ , such that

$$C_1 h \sum_{i \in A_j^h} g_i(c_i^h)^2 \leq \int_{I_j^h} g v^2 dx \leq C_2 h \sum_{i \in A_j^h} g_i(c_i^h)^2; \quad (5.27)$$

- (b) there is a positive constant  $C$ , independent of  $v$  and  $h$ , such that

$$\frac{1}{h} \sum_{i \in \mathbb{Z}^n} g_{i+\frac{1}{2}}(c_{i+1}^h - c_i^h)^2 \leq C \int_{\mathbb{R}} g v'^2 dx. \quad (5.28)$$

*Proof.* (a) Consider  $j \in \mathbb{Z}$  and the corresponding  $A_j^h$  such that, for  $i \in A_j^h$ ,  $H < x_i^h$ . Let  $g_M = \max_{i \in A_j^h} g(x_i^h)$  and  $g_m = \min_{i \in A_j^h} g(x_i^h)$ . Then, it is easy to check that  $\frac{g_M}{g_m} \leq C$ , where  $C$  depends  $\kappa$ . Now, using (5.19), we have

$$\begin{aligned} h \sum_{i \in A_j^h} g_i(c_i^h)^2 &\leq g_M h \sum_{i \in A_j^h} (c_i^h)^2 \\ &\leq \frac{g_M}{C_1 g_m} \int_{I_j^h} g v^2 dx \\ &\leq C \int_{I_j^h} g v^2 dx. \end{aligned} \quad (5.29)$$

Using a similar argument, we get

$$\int_{I_j^h} g v^2 dx \leq C \sum_{i \in A_j^h} g_i(c_i^h)^2.$$

Combining the above with (5.29) gives the required result. Using similar arguments, we can prove (5.27) for any  $j \in \mathbb{Z}$ .  $\square$

- (b) Let  $u = \sum_{i \in \mathbb{Z}} c_i \phi_i(x)$ . Then, from (5.14) and (5.15) with  $h = 1$ , we have

$$c_i = \Psi_i^1(u) = \int_{i-1}^{i+1} \psi_i^1(x) u(x) dx, \quad (5.30)$$



and therefore

$$\begin{aligned} c_{i+1} - c_i &= \int_i^{i+2} \psi_{i+1}^1(x)u(x) \, dx - \int_{i-1}^{i+1} \psi_i^1(x)u(x) \, dx \\ &= \int_{i-1}^{i+2} (\psi_{i+1}^1(x) - \psi_i^1(x))u(x) \, dx. \end{aligned} \tag{5.31}$$

Let  $F(x) = \int_{i-1}^x [\psi_{i+1}^1(t) - \psi_i^1(t)] \, dt$ . Using translation and (5.12), it is easily seen that

$$\int_{i-1}^{i+1} \psi_i^1(t) \, dt = \int_i^{i+2} \psi_{i+1}^1(t) \, dt = 1,$$

and therefore  $F(i - 1) = F(i + 2) = 0$ . Also, using the Schwartz inequality and (5.11), we can show that

$$\int_{i-1}^{i+2} F^2 \, dx \leq C.$$

Now, using the above bound, integrating (5.31) by parts, and using the Schwartz inequality, we get

$$(c_{i+1} - c_i)^2 = \left( \int_{i-1}^{i+2} Fu' \, dx \right)^2 \leq C \int_{i-1}^{i+2} u'^2 \, dx. \tag{5.32}$$

Let  $v = \sum_{i \in \mathbb{Z}} c_i^h \phi_i^h(x)$ . Then, by a standard scaling argument, we have

$$\int_{x_{i-1}^h}^{x_{i+2}^h} (v'(x))^2 \, dx = \frac{1}{h} \int_{i-1}^{i+1} (u'(y))^2 \, dy, \tag{5.33}$$

where  $u(y) = \sum_{i \in \mathbb{Z}} c_i^h \phi_i(y)$ . Therefore, from (5.32) and (5.33), we have

$$\frac{1}{h} (c_{i+1}^h - c_i^h)^2 \leq C \int_{x_{i-1}^h}^{x_{i+2}^h} v'^2 \, dx. \tag{5.34}$$

From the definition of  $g(x)$ , we can show that

$$\left( 1 + \frac{g_{i+1/2} - g(x)}{g(x)} \right) \leq C, \quad \text{for } x \in (x_{i-1}^h, x_{i+2}^h).$$

Therefore,

$$\begin{aligned} \frac{1}{h} g_{i+1/2} (c_{i+1}^h - c_i^h)^2 &\leq C \int_{x_{i-1}^h}^{x_{i+2}^h} gv'^2 \left( 1 + \frac{g_{i+1/2} - g}{g} \right) \, dx \\ &\leq C \int_{x_{i-1}^h}^{x_{i+2}^h} gv'^2 \, dx, \end{aligned}$$

and hence

$$\frac{1}{h} \sum_{i \in \mathbb{Z}} g_{i+\frac{1}{2}} (c_{i+1}^h - c_i^h)^2 \leq C \sum_{i \in \mathbb{Z}} \int_{x_{i-1}^h}^{x_{i+2}^h} g v'^2 dx \leq C \int_{\mathbb{R}} g v'^2 dx,$$

which is the required result.  $\square$

**Remark 37.** We note that it is possible to show that

$$\int_{\mathbb{R}} g v'^2 dx \leq C \frac{1}{h} \sum_{i \in \mathbb{Z}^n} g_{i+\frac{1}{2}} (c_{i+1}^h - c_i^h)^2,$$

and together with (5.28) we see that  $\frac{1}{h} \sum_{i \in \mathbb{Z}^n} g_{i+\frac{1}{2}} (c_{i+1}^h - c_i^h)^2$  is equivalent to  $|v|_{H^1(\mathbb{R})}^2$ . The proof of this fact is easier than the proof of (5.28), and we do not provide the proof here.

*A perturbed bilinear form  $B_{\Theta}(u, v)$  and related results*

For a given  $\Theta \geq 1$ , we now consider the bilinear form

$$B_{\Theta}^{\mathbb{R}}(u, v) \equiv B^{\mathbb{R}}(u, v) + \Theta D^{\mathbb{R}}(u, v),$$

where

$$B^{\mathbb{R}}(u, v) = \int_{\mathbb{R}} (u'v' + uv) dx \quad \text{and} \quad D^{\mathbb{R}}(u, v) = \int_{\mathbb{R}} uv dx.$$

We will write  $B_{\Theta}(u, v) \equiv B_{\Theta}^{\mathbb{R}}(u, v)$ , but will use  $B_{\Theta}^F(u, v)$  when the domain of integration is  $F$  instead of  $\mathbb{R}$ . Also, we will use  $D^F(u, v)$ , where  $\mathbb{R}$  is replaced by a domain  $F$  in the definition of  $D^{\mathbb{R}}(u, v)$ .

Let  $H_{g, \Theta}^1$  and  $H_{g^{-1}, \Theta}^1$  be Hilbert spaces defined as

$$H_{g, \Theta}^1 = \left\{ u : \|u\|_{1, g, \Theta}^2 \equiv \int_{\mathbb{R}} g u'^2 dx + (1 + \Theta) \int_{\mathbb{R}} g u^2 dx < \infty \right\},$$

$$H_{g^{-1}, \Theta}^1 = \left\{ u : \|u\|_{1, g^{-1}, \Theta}^2 \equiv \int_{\mathbb{R}} g^{-1} u'^2 dx + (1 + \Theta) \int_{\mathbb{R}} g^{-1} u^2 dx < \infty \right\}.$$

We will choose  $\Theta$  later. The choice of  $\Theta$ , along with the choices of  $\gamma$  and  $\alpha$ , mentioned before, is important for the main result of this section. We assume that  $\alpha^2/\bar{\Theta} < 1$  where  $\bar{\Theta} = 1 + \Theta$ .

We will often suppress  $\Theta$  in  $\|u\|_{1, g, \Theta}$  and  $\|u\|_{1, g^{-1}, \Theta}$  and instead write  $\|u\|_{1, g}$  and  $\|u\|_{1, g^{-1}}$  respectively. We will also use the fact that  $|g'/g| \leq \alpha$ , which is obvious from the definition of  $g(x)$ .

**Remark 38.** The space  $H_{g, \Theta}^1$  is directed towards obtaining interior estimates of  $e'_h$ , i.e.,  $e'_h$  is locally characterized through the use of the space  $H_{g, \Theta}$ .

We now consider  $B_\Theta(\cdot, \cdot) : H_{g,\Theta}^1 \times H_{g^{-1},\Theta}^1 \rightarrow \mathbb{R}$ .

**Lemma 5.6.** The bilinear form  $B_\Theta(\cdot, \cdot)$  is bounded on  $H_{g,\Theta}^1 \times H_{g^{-1},\Theta}^1$ , that is,

$$B_\Theta(u, v) \leq C \|u\|_{1,g} \|v\|_{1,g^{-1}}, \quad \text{for all } u \in H_{g,\Theta}^1, \quad v \in H_{g^{-1},\Theta}^1.$$

*Proof.* Let  $u \in H_{g,\Theta}^1$  and  $v \in H_{g^{-1},\Theta}^1$ . Then

$$\begin{aligned} B_\Theta(u, v) &= \int_{\mathbb{R}} [u'v' + (1 + \Theta)uv] \, dx \\ &= \int_{\mathbb{R}} [g^{1/2}u'g^{-1/2}v' + (1 + \Theta)^{1/2}g^{1/2}u(1 + \Theta)^{1/2}g^{-1/2}v] \, dx \\ &\leq C \left[ \int_{\mathbb{R}} (gu'^2 + (1 + \Theta)gu^2) \, dx \right]^{1/2} \left[ \int_{\mathbb{R}} (g^{-1}v'^2 + (1 + \Theta)g^{-1}v^2) \, dx \right]^{1/2} \\ &= C \|u\|_{1,g} \|v\|_{1,g^{-1}}. \quad \square \end{aligned}$$

**Lemma 5.7.** Suppose  $\alpha^2/\bar{\Theta} < 1$ . Then there is a constant  $C > 0$ , which depends on  $\alpha^2/\bar{\Theta}$ , such that

$$\inf_{u \in H_{g,\Theta}^1} \sup_{v \in H_{g^{-1},\Theta}^1} \frac{B_\Theta(u, v)}{\|u\|_{1,g} \|v\|_{1,g^{-1}}} \geq C > 0.$$

*Proof.* Suppose  $u \in H_{g,\Theta}^1$ . We consider  $v = gu$ . Now,

$$\begin{aligned} B_\Theta(u, v) &= \int_{\mathbb{R}} [u'v' + \bar{\Theta}uv] \, dx \\ &= \int_{\mathbb{R}} [u'(gu' + g'u) + \bar{\Theta}gu^2] \, dx \\ &= \int_{\mathbb{R}} [gu'^2 + \bar{\Theta}gu^2] \, dx + \int_{\mathbb{R}} uu'g' \, dx. \end{aligned} \tag{5.35}$$

Now, for  $\epsilon > 0$ ,

$$\begin{aligned} \left| \int_{\mathbb{R}} uu'g' \, dx \right| &= \left| \int_{\mathbb{R}} guu' \left( \frac{g'}{g} \right) \, dx \right| \\ &\leq \alpha \int_{\mathbb{R}} |g^{1/2}ug^{1/2}u'| \, dx \\ &\leq \alpha \left[ \epsilon \int_{\mathbb{R}} gu'^2 \, dx + \frac{1}{\epsilon} \int_{\mathbb{R}} gu^2 \, dx \right], \end{aligned}$$

and therefore, from (5.35), we get

$$\begin{aligned} B_{\Theta}(u, v) &\geq \int_{\mathbb{R}} (gu'^2 + \bar{\Theta}gu^2) dx - \alpha \left[ \epsilon \int_{\mathbb{R}} gu'^2 dx + \frac{1}{\epsilon} \int_{\mathbb{R}} gu^2 dx \right] \\ &= (1 - \alpha\epsilon) \int_{\mathbb{R}} gu'^2 dx + \left(1 - \frac{\alpha}{\epsilon\bar{\Theta}}\right) \int_{\mathbb{R}} \bar{\Theta}gu^2 dx. \end{aligned} \quad (5.36)$$

We choose  $\epsilon$  such that  $\alpha\epsilon < 1$  and  $\alpha/\epsilon\bar{\Theta} < 1$ , and therefore, from (5.36), we have

$$B_{\Theta}(u, v) \geq C_1 \|u\|_{1,g}^2, \quad (5.37)$$

where

$$C_1 = \min \left[ (1 - \alpha\epsilon), \left(1 - \frac{\alpha}{\epsilon\bar{\Theta}}\right) \right] > 0. \quad (5.38)$$

We next show that  $\|v\|_{1,g^{-1}} \leq C_2 \|u\|_{1,g}$ . First note that

$$\begin{aligned} \int_{\mathbb{R}} g^{-1}v'^2 dx &= \int_{\mathbb{R}} g^{-1}(gu' + g'u)^2 dx \\ &= \int_{\mathbb{R}} gu'^2 dx + \int_{\mathbb{R}} g^{-1}g'^2u^2 dx + 2 \int_{\mathbb{R}} g'uu' dx. \end{aligned} \quad (5.39)$$

Now,

$$\int_{\mathbb{R}} g^{-1}g'^2u^2 dx = \int_{\mathbb{R}} g \left(\frac{g'}{g}\right)^2 u^2 dx \leq \alpha^2 \int_{\mathbb{R}} gu^2 dx, \quad (5.40)$$

and

$$\begin{aligned} 2 \int_{\mathbb{R}} g'uu' dx &= 2 \int_{\mathbb{R}} g \left(\frac{g'}{g}\right) uu' dx \leq 2 \int_{\mathbb{R}} |\alpha g^{1/2} u g^{1/2} u'| dx \\ &\leq \int_{\mathbb{R}} (gu'^2 + \alpha^2 gu^2) dx. \end{aligned} \quad (5.41)$$

Therefore, using (5.40) and (5.41) in (5.39), we get

$$\int_{\mathbb{R}} g^{-1}v'^2 dx \leq 2 \int_{\mathbb{R}} gu'^2 dx + \frac{2\alpha^2}{\bar{\Theta}} \int_{\mathbb{R}} \bar{\Theta}gu^2 dx. \quad (5.42)$$

Thus, combining

$$\bar{\Theta} \int_{\mathbb{R}} g^{-1}v^2 dx = \bar{\Theta} \int_{\mathbb{R}} g^{-1}g^2u^2 dx = \bar{\Theta} \int_{\mathbb{R}} gu^2 dx$$

with (5.42), we get

$$\|v\|_{1,g^{-1}}^2 \leq 2 \int_{\mathbb{R}} gu'^2 dx + \left(1 + \frac{2\alpha^2}{\bar{\Theta}}\right) \int_{\mathbb{R}} \bar{\Theta}gu^2 dx. \quad (5.43)$$

Since  $\alpha^2/\bar{\Theta} < 1$ , from (5.43) we have

$$\|v\|_{1,g^{-1}}^2 \leq 3 \|u\|_{1,g}^2. \quad (5.44)$$

Thus  $v \in H^1_{g^{-1}, \Theta}$ , and combining (5.37) and (5.44), we get

$$\inf_{u \in H^1_{g, \Theta}} \sup_{v \in H^1_{g^{-1}, \Theta}} \frac{B_{\Theta}(u, v)}{\|u\|_{1, g} \|v\|_{1, g^{-1}}} \geq C > 0,$$

where

$$C = \frac{\min[(1 - \alpha\epsilon), (1 - \frac{\alpha}{\epsilon\bar{\Theta}})]}{\sqrt{3}}. \quad \square$$

We now prove the inf-sup condition on  $S_h \times S_h$ . In the proof, we will use the function  $d_i(x), x \in I_k^h$  and  $i \in A_k^h$ , to denote the following similar functions:

$$\frac{g_i - g(x)}{\sqrt{g_i g(x)}}, \quad \frac{g_{i+\frac{1}{2}} - g(x)}{\sqrt{g_{i+\frac{1}{2}} g(x)}}, \quad \frac{g_{i+\frac{1}{2}} - g_{l_k+\frac{1}{2}}}{\sqrt{g_{l_k+\frac{1}{2}} g(x)}},$$

where  $l_k \in A_k^h$ . It is easily seen from the definition of  $g(x)$  that

$$|d_i(x)| \leq C\alpha h. \tag{5.45}$$

**Lemma 5.8.** Suppose  $\alpha^2/\bar{\Theta} < C_1$  and  $\alpha h < C_2$ , where  $C_1, C_2$  are sufficiently small. Then there is a constant  $C > 0$ , independent of  $u, v$ , and  $h$ , but possibly depending on  $\kappa$  and  $\alpha^2/\bar{\Theta}$ , such that, for sufficiently small  $h$ ,

$$\inf_{u \in S_h} \sup_{v \in S_h} \frac{B_{\Theta}(u, v)}{\|u\|_{1, g} \|v\|_{1, g^{-1}}} \geq C > 0. \tag{5.46}$$

*Proof.* Let  $u = \sum_{i \in \mathbb{Z}} c_i^h \phi_i^h$  in  $S_h$  such that  $\|u\|_{1, g} \leq \infty$ . Then, for  $x \in I_k^h$ , we have  $u = \sum_{i \in A_k^h} c_i^h \phi_i^h$ . Since  $\sum_{i \in A_k^h} \phi_i^{h'}(x) = 0$  for  $x \in I_k^h$ , we have

$$u'(x) = \sum_{i \in A_k^h} c_i^h \phi_i^{h'}(x) = \sum_{i \in A_k^h} (c_i^h - c_{l_k}^h) \phi_i^{h'}(x), \quad x \in I_k^h,$$

where  $l_k \in A_k^h$  is a fixed integer for given  $k$ .

We now choose  $v = \sum_{i \in \mathbb{Z}} c_i^h g_{i+\frac{1}{2}} \phi_i^h$  in  $S_h$  and, as before, for  $x \in I_k^h$ ,

$$\begin{aligned} v'(x) &= \sum_{i \in A_k^h} c_i^h g_{i+\frac{1}{2}} \phi_i^{h'}(x) \\ &= \sum_{i \in A_k^h} (c_i^h g_{i+\frac{1}{2}} - c_{l_k}^h g_{l_k+\frac{1}{2}}) \phi_i^{h'}(x) \\ &= \sum_{i \in A_k^h} (c_i^h - c_{l_k}^h) g_{i+\frac{1}{2}} \phi_i^{h'}(x) + c_{l_k}^h \sum_{i \in A_k^h} (g_{i+\frac{1}{2}} - g_{l_k+\frac{1}{2}}) \phi_i^{h'}(x). \end{aligned}$$

Now,

$$\int_{\mathbb{R}} u'v' dx = \int_{\mathbb{R}} gu'^2 dx + \int_{\mathbb{R}} u'(v' - gu') dx. \tag{5.47}$$

For  $\epsilon > 0$ , we have

$$\begin{aligned} \left| \int_{\mathbb{R}} u'(v' - gu') \, dx \right| &= \left| \int_{\mathbb{R}} g^{1/2} u' \frac{(v' - gu')}{g^{1/2}} \, dx \right| \\ &\leq \epsilon \int_{\mathbb{R}} gu'^2 \, dx + \frac{1}{\epsilon} \int_{\mathbb{R}} \frac{(v' - gu')^2}{g} \, dx. \end{aligned} \quad (5.48)$$

Now, from the definition of  $v'$  and  $u'$ ,

$$\begin{aligned} \int_{I_k^h} \frac{1}{g} (v' - gu')^2 \, dx &= \int_{I_k^h} \left[ \sum_{i \in A_k^h} (c_i^h - c_{l_k}^h) \frac{g_{i+\frac{1}{2}} - g}{g^{1/2}} \phi_i^{h'} \right. \\ &\quad \left. + c_{l_k}^h \sum_{i \in A_k^h} \frac{g_{i+\frac{1}{2}} - g_{l_k+\frac{1}{2}}}{g^{1/2}} \phi_i^{h'} \right]^2 \, dx \\ &\leq C \int_{I_k^h} \left[ \sum_{i \in A_k^h} (c_i^h - c_{l_k}^h) \frac{g_{i+\frac{1}{2}} - g}{g^{1/2}} \phi_i^{h'} \right]^2 \, dx \\ &\quad + C \int_{I_k^h} (c_{l_k}^h)^2 \left[ \sum_{i \in A_k^h} \frac{g_{i+\frac{1}{2}} - g_{l_k+\frac{1}{2}}}{g^{1/2}} \phi_i^{h'} \right]^2 \, dx. \end{aligned} \quad (5.49)$$

The first term of the right-hand side of the above inequality, employing (5.45) and (5.20), gives

$$\begin{aligned} &\int_{I_k^h} \left[ \sum_{i \in A_k^h} (c_i^h - c_{l_k}^h) \frac{g_{i+\frac{1}{2}} - g}{g^{1/2}} \phi_i^{h'} \right]^2 \, dx \\ &= \int_{I_k^h} \left[ \sum_{i \in A_k^h} (c_i^h - c_{l_k}^h) (g_{i+\frac{1}{2}})^{1/2} \frac{g_{i+\frac{1}{2}} - g}{(g_{i+\frac{1}{2}})^{1/2} g^{1/2}} \phi_i^{h'} \right]^2 \, dx \\ &\leq C \int_{I_k^h} \sum_{i \in A_k^h} (c_i^h - c_{l_k}^h)^2 g_{i+\frac{1}{2}} d_i^2 (\phi_i^{h'})^2 \, dx \\ &\leq C \alpha^2 h^2 \sum_{i \in A_k^h} (c_i^h - c_{l_k}^h)^2 g_{i+\frac{1}{2}} \int_{I_k^h} (\phi_i^{h'})^2 \, dx \\ &\leq C \alpha^2 h^2 \frac{1}{h} \sum_{i \in A_k^h} (c_i^h - c_{l_k}^h)^2 g_{i+\frac{1}{2}} \\ &\leq C \alpha^2 h^2 \frac{1}{h} \sum_{i, (i+1) \in A_k^h} (c_{i+1}^h - c_i^h)^2 g_{i+\frac{1}{2}}, \end{aligned} \quad (5.50)$$

where  $C$  is independent of  $\alpha$ ,  $h$ , but depends on  $\kappa$ .

The second term of the right-hand side of (5.49), employing (5.45), gives

$$\begin{aligned}
 & (c_{l_k}^h)^2 \int_{I_k^h} \left[ \sum_{i \in A_k^h} \frac{g_{i+\frac{1}{2}} - g_{l_k+\frac{1}{2}}}{g^{1/2}} \phi_i^{h'} \right]^2 dx \\
 &= (c_{l_k}^h)^2 g_{l_k+\frac{1}{2}} \int_{I_k^h} \left[ \sum_{i \in A_k^h} \frac{g_{i+\frac{1}{2}} - g_{l_k+\frac{1}{2}}}{(g_{l_k+\frac{1}{2}})^{1/2} g^{1/2}} \phi_i^{h'} \right]^2 dx \\
 &\leq C (c_{l_k}^h)^2 g_{l_k+\frac{1}{2}} \sum_{i \in A_k^h} \int_{I_k^h} d_i^2 (\phi_i^{h'})^2 dx \\
 &\leq C \alpha^2 h (c_{l_k}^h)^2 g_{l_k}, \tag{5.51}
 \end{aligned}$$

where  $C$  depends on  $\kappa$ , but is independent of  $\alpha$ ,  $h$ . Therefore, from (5.49), (5.50), and (5.51) we have

$$\int_{I_k^h} \frac{1}{g} (v' - gu')^2 dx \leq C \alpha^2 h^2 \frac{1}{h} \sum_{i, (i+1) \in A_k^h} (c_{i+1}^h - c_i^h)^2 g_{i+\frac{1}{2}} + C \alpha^2 h (c_{l_k}^h)^2 g_{l_k}.$$

Now summing the above inequality over  $k \in \mathbb{Z}$ , and using (5.27) and (5.28), we get

$$\begin{aligned}
 \int_{\mathbb{R}} \frac{1}{g} (v' - gu')^2 dx &= \sum_{k \in \mathbb{Z}} \int_{I_k^h} \frac{1}{g} (v' - gu')^2 dx \\
 &\leq C \alpha^2 h^2 \frac{1}{h} \sum_{k \in \mathbb{Z}} \sum_{i, (i+1) \in A_k^h} (c_{i+1}^h - c_i^h)^2 g_{i+\frac{1}{2}} \\
 &\quad + C \alpha^2 h \sum_{k \in \mathbb{Z}} \sum_{i \in A_k^h} (c_i^h)^2 g_i \\
 &\leq C \alpha^2 h^2 \frac{1}{h} \sum_{i \in \mathbb{Z}} (c_{i+1}^h - c_i^h)^2 g_{i+\frac{1}{2}} \\
 &\quad + C \alpha^2 \sum_{k \in \mathbb{Z}} \int_{I_k^h} gu^2 dx \\
 &\leq C \alpha^2 h^2 \int_{\mathbb{R}} gu'^2 dx + C \alpha^2 \int_{\mathbb{R}} gu^2 dx. \tag{5.52}
 \end{aligned}$$

Then, from (5.48) and (5.52) we have

$$\begin{aligned} \int_{\mathbb{R}} u'(v' - gu') \, dx &\leq \epsilon \int_{\mathbb{R}} gu'^2 \, dx \\ &\quad + \frac{1}{\epsilon} \left[ C\alpha^2 h^2 \int_{\mathbb{R}} gu'^2 \, dx + C\alpha^2 \int_{\mathbb{R}} gu^2 \, dx \right] \\ &= \left( \epsilon + \frac{C\alpha^2 h^2}{\epsilon} \right) \int_{\mathbb{R}} gu'^2 \, dx + \frac{C\alpha^2}{\epsilon \bar{\Theta}} \int_{\mathbb{R}} \bar{\Theta} gu^2 \, dx. \end{aligned} \quad (5.53)$$

We next consider

$$\bar{\Theta} \int_{\mathbb{R}} uv \, dx = \bar{\Theta} \int_{\mathbb{R}} gu^2 \, dx + \bar{\Theta} \int_{\mathbb{R}} u(v - gu) \, dx. \quad (5.54)$$

For  $\epsilon_1 > 0$ , we have

$$\begin{aligned} \left| \int_{\mathbb{R}} u(v - gu) \, dx \right| &= \left| \int_{\mathbb{R}} g^{1/2} u \frac{v - gu}{g^{1/2}} \, dx \right| \\ &\leq \epsilon_1 \int_{\mathbb{R}} gu^2 \, dx + \frac{1}{\epsilon_1} \int_{\mathbb{R}} \frac{(v - gu)^2}{g} \, dx. \end{aligned} \quad (5.55)$$

Now,

$$\begin{aligned} \int_{I_k^h} \frac{(v - gu)^2}{g} \, dx &= \int_{I_k^h} \frac{1}{g} \left[ \sum_{i \in A_k^h} c_i^h (g_i - g) \phi_i^h \right]^2 \, dx \\ &= \int_{I_k^h} \left[ \sum_{i \in A_k^h} c_i^h g_i^{1/2} \frac{(g_i - g)}{g_i^{1/2} g^{1/2}} \phi_i^h \right]^2 \, dx \\ &\leq C \int_{I_k^h} \sum_{i \in A_k^h} (c_i^h)^2 g_i d_i^2 \phi_i^{h^2} \, dx \\ &\leq C\alpha^2 h^2 h \sum_{i \in A_k^h} (c_i^h)^2 g_i. \end{aligned}$$

Therefore, using (5.27), we get

$$\begin{aligned} \int_{\mathbb{R}} \frac{(v - gu)^2}{g} \, dx &= \sum_{k \in \mathbb{Z}} \int_{I_k^h} \frac{(v - gu)^2}{g} \, dx \\ &\leq \sum_{k \in \mathbb{Z}} C\alpha^2 h^2 h \sum_{i \in A_k^h} (c_i^h)^2 g_i \\ &\leq C\alpha^2 h^2 \int_{\mathbb{R}} gu^2 \, dx. \end{aligned} \quad (5.56)$$



Thus, from (5.55), (5.56), we have

$$\bar{\Theta} \left| \int_{\mathbb{R}} u(v - gu) \, dx \right| \leq \left( \epsilon_1 + \frac{C\alpha^2 h^2}{\epsilon_1} \right) \int_{\mathbb{R}} \bar{\Theta} gu^2 \, dx, \tag{5.57}$$

and combining (5.47), (5.53), (5.54), and (5.57), we get

$$\begin{aligned} |B_{\Theta}(u, v)| &\geq \int_{\mathbb{R}} gu'^2 \, dx + \bar{\Theta} \int_{\mathbb{R}} gu^2 \, dx \\ &\quad - \left| \int_{\mathbb{R}} u'(v' - gu') \, dx \right| - \bar{\Theta} \left| \int_{\mathbb{R}} u(v - gu) \, dx \right| \\ &\geq \left( 1 - \epsilon - \frac{C\alpha^2 h^2}{\epsilon} \right) \int_{\mathbb{R}} gu'^2 \, dx \\ &\quad + \left( 1 - \epsilon_1 - \frac{C\alpha^2 h^2}{\epsilon_1} - \frac{C\alpha^2}{\epsilon \bar{\Theta}} \right) \int_{\mathbb{R}} \bar{\Theta} gu^2 \, dx. \end{aligned}$$

Now we can choose  $\epsilon$  and  $\epsilon_1$ , for sufficiently small  $h$ , such that

$$|B_{\Theta}(u, v)| \geq C_1 \|u\|_{1,g}^2, \tag{5.58}$$

where  $C_1 > 0$ , since  $\alpha^2/\bar{\Theta} \ll 1$ ,  $\alpha h \ll 1$  by assumption.

We now show that  $\|v\|_{1,g^{-1}} \leq C\|u\|_{1,g}$ . From the definition of  $v'$ , we have

$$\begin{aligned} \int_{I_k^h} g^{-1} v'^2 \, dx &= \int_{I_k^h} g^{-1} \left[ \sum_{i \in A_k^h} (c_i^h - c_{l_k}^h) g_{i+\frac{1}{2}} \phi_i^{h'} \right. \\ &\quad \left. + c_{l_k} \sum_{i \in A_k^h} (g_{i+\frac{1}{2}} - g_{l_k+\frac{1}{2}}) \phi_i^{h'} \right]^2 \, dx \\ &\leq \int_{I_k^h} \left[ \sum_{i \in A_k^h} (c_i^h - c_{l_k}^h) \frac{g_{i+\frac{1}{2}}}{g^{1/2}} \phi_i^{h'} \right]^2 \, dx \\ &\quad + c_{l_k}^2 \int_{I_k^h} \left[ \sum_{i \in A_k^h} \frac{(g_{i+\frac{1}{2}} - g_{l_k+\frac{1}{2}})}{g^{1/2}} \phi_i^{h'} \right]^2 \, dx. \end{aligned} \tag{5.59}$$

Now,

$$\begin{aligned} &\int_{I_k^h} \left[ \sum_{i \in A_k^h} (c_i^h - c_{l_k}^h) \frac{g_{i+\frac{1}{2}}}{g^{1/2}} \phi_i^{h'} \right]^2 \, dx \\ &= \int_{I_k^h} \left[ \sum_{i \in A_k^h} (c_i^h - c_{l_k}^h) g^{1/2} \phi_i^{h'} + \sum_{i \in A_k^h} (c_i^h - c_{l_k}^h) g_{i+\frac{1}{2}}^{1/2} \left( \frac{g_{i+\frac{1}{2}} - g}{g_{i+\frac{1}{2}}^{1/2} g^{1/2}} \right) \phi_i^{h'} \right]^2 \, dx \end{aligned}$$

$$\begin{aligned}
&\leq C \int_{I_k^h} g \left[ \sum_{i \in A_k^h} (c_i^h - c_{l_k}^h) \phi_i^{h'} \right]^2 dx + C \int_{I_k^h} \sum_{i \in A_k^h} (c_i^h - c_{l_k}^h)^2 g_{i+\frac{1}{2}} d_i^2 (\phi_i^{h'})^2 dx \\
&\leq C \int_{I_k^h} g u'^2 dx + C \int_{I_k^h} \sum_{i \in A_k^h} (c_i^h - c_{l_k}^h)^2 g_{i+\frac{1}{2}} d_i^2 (\phi_i^{h'})^2 dx. \tag{5.60}
\end{aligned}$$

Also using (5.45) and (5.20), we have

$$\begin{aligned}
\sum_{i \in A_k^h} (c_i^h - c_{l_k}^h)^2 g_{i+\frac{1}{2}} \int_{I_k^h} d_i^2 (\phi_i^{h'})^2 dx &\leq C \alpha^2 h^2 \frac{1}{h} \sum_{i \in A_k^h} (c_i^h - c_{l_k}^h)^2 g_{i+\frac{1}{2}} \tag{5.61} \\
&\leq C \alpha^2 h^2 \frac{1}{h} \sum_{i, i+1 \in A_k^h} (c_{i+1}^h - c_i^h)^2 g_{i+\frac{1}{2}}.
\end{aligned}$$

Therefore, using (5.60), (5.61) and (5.51) in (5.59), we get

$$\begin{aligned}
&\int_{I_k^h} g^{-1} v'^2 dx \\
&\leq C \int_{I_k^h} g u'^2 dx + C \alpha^2 h (c_{l_k}^h)^2 g_{l_k} + C \alpha^2 h^2 \frac{1}{h} \sum_{i, i+1 \in A_k^h} (c_{i+1}^h - c_i^h)^2 g_{i+\frac{1}{2}} \\
&\leq C \int_{I_k^h} g u'^2 dx + C \alpha^2 h \sum_{i \in A_k^h} (c_i^h)^2 g_i + C \alpha^2 h^2 \frac{1}{h} \sum_{i, i+1 \in A_k^h} (c_{i+1}^h - c_i^h)^2 g_{i+\frac{1}{2}} \\
&\leq C \int_{I_k^h} g u'^2 dx + C \alpha^2 \int_{I_k^h} g u^2 dx + C \alpha^2 h^2 \frac{1}{h} \sum_{i, i+1 \in A_k^h} (c_{i+1}^h - c_i^h)^2 g_{i+\frac{1}{2}}.
\end{aligned}$$

Now, summing the above inequality for all  $k$  and using (5.20), we get

$$\int_{\mathbb{R}} g^{-1} v'^2 dx \leq C(1 + \alpha^2 h^2) \int_{\mathbb{R}} g u'^2 dx + C \alpha^2 \int_{\mathbb{R}} g u^2 dx. \tag{5.62}$$

Again,

$$\int_{\mathbb{R}} g^{-1} v^2 dx = \int_{\mathbb{R}} g^{-1} \left( \sum_{i \in \mathbb{Z}} c_i^h g_i \phi_i^h \right)^2 dx = \int_{\mathbb{R}} \left( \sum_{i \in \mathbb{Z}} c_i^h \frac{g_i}{g^{1/2}} \phi_i^h \right)^2 dx. \tag{5.63}$$

Now using (5.45), we get

$$\begin{aligned}
&\int_{I_k^h} \left( \sum_{i \in A_k^h} c_i^h \frac{g_i}{g^{1/2}} \phi_i^h \right)^2 dx \\
&= \int_{I_k^h} \left[ \sum_{i \in A_k^h} c_i^h g^{1/2} \phi_i^h + \sum_{i \in A_k^h} c_i^h \frac{g_i - g}{g^{1/2}} \phi_i^h \right]^2 dx
\end{aligned}$$

$$\begin{aligned}
 &\leq C \int_{I_k^h} \left( \sum_{i \in A_k^h} c_i^h g^{1/2} \phi_i^h \right)^2 dx + C \int_{I_k^h} \sum_{i \in A_k^h} (c_i^h)^2 g_i \left( \frac{g_i - g}{g^{1/2} g_i^{1/2}} \right)^2 \phi_i^{h^2} dx \\
 &\leq C \int_{I_k^h} g u^2 dx + C \sum_{i \in A_k^h} (c_i^h)^2 g_i \int_{I_k^h} d_i^2 \phi_i^{h^2} dx \\
 &\leq C \int_{I_k^h} g u^2 dx + C \alpha^2 h^2 h \sum_{i \in A_k^h} (c_i^h)^2 g_i \\
 &\leq C \int_{I_k^h} g u^2 dx + C \alpha^2 h^2 C \int_{I_k^h} g u^2 dx \\
 &\leq C(1 + \alpha^2 h^2) \int_{I_k^h} g u^2 dx,
 \end{aligned}$$

and therefore, from (5.63) and the above inequality,

$$\begin{aligned}
 \int_{\mathbb{R}} g^{-1} v^2 dx &\leq C \sum_{k \in \mathbb{Z}} \int_{I_k^h} \left( \sum_{i \in A_k^h} c_i^h \frac{g_i}{g^{1/2}} \phi_i^h \right)^2 dx \\
 &\leq C(1 + \alpha^2 h^2) \sum_{k \in \mathbb{Z}} \int_{I_k^h} g u^2 dx \\
 &= C(1 + \alpha^2 h^2) \int_{\mathbb{R}} g u^2 dx.
 \end{aligned}$$

Thus, combining (5.62) and the above inequality, we have

$$\begin{aligned}
 \|v\|_{1,g^{-1}}^2 &= \int_{\mathbb{R}} g^{-1} v'^2 dx + \bar{\Theta} \int_{\mathbb{R}} g^{-1} v^2 dx \\
 &\leq C(1 + \alpha^2 h^2) \int_{\mathbb{R}} g u'^2 dx + C \alpha^2 \int_{\mathbb{R}} g u^2 dx \\
 &\quad + C(1 + \alpha^2 h^2) \int_{\mathbb{R}} \bar{\Theta} g u^2 dx \\
 &\leq C(1 + \alpha^2 h^2) \int_{\mathbb{R}} g u'^2 dx \\
 &\quad + \left[ \frac{C \alpha^2}{\bar{\Theta}} + C(1 + \alpha^2 h^2) \right] \int_{\mathbb{R}} \bar{\Theta} g u^2 dx \\
 &\leq C \left( 1 + \alpha^2 h^2 + \frac{\alpha^2}{\bar{\Theta}} \right) \|u\|_{1,g}^2 \leq C_2 \|u\|_{1,g}^2. \tag{5.64}
 \end{aligned}$$

Finally, combining (5.58) and (5.64) we get the desired result. □

*Projection with respect to  $B_\Theta(u, v)$*

Suppose  $u \in H_{g,\Theta}^1$  and let  $P_\Theta u$  be the projection of  $u$  onto  $S_h$  defined by

$$B_\Theta(P_\Theta u, v) = B_\Theta(u, v), \quad \text{for all } v \in S_h.$$

The projection  $P_\Theta u$  exists (see Babuška and Aziz (1972)), and it is clear from Lemmas 5.8 and 5.6 that

$$\|P_\Theta u\|_{1,g} \leq C \sup_{v \in S_h} \frac{B_\Theta(u, v)}{\|v\|_{1,g^{-1}}} \leq C \|u\|_{1,g}. \quad (5.65)$$

We first note that, for fixed  $h, \alpha$ , and  $\Theta$ , the polynomials belong to the space  $H_{g,\Theta}^1$ . Moreover, for fixed  $h, \alpha$ , and  $\Theta$ , we can also show, using (5.27) and Remark 37 (page 66), that  $\tilde{\mathcal{I}}_h(x^{k+1}) \in H_{g,\Theta}^1$ , where  $\tilde{\mathcal{I}}_h(x^{k+1})$  is the interpolant of  $x^{k+1}$ , as defined in (4.35).

We now present some simple facts about polynomials and periodic functions.

**Lemma 5.9.** Let the shape functions  $\{\phi_i^h\}_{i \in \mathbb{Z}}$  be reproducing of order  $k$ . Then

$$(a) \quad P_\Theta x^i = x^i, \quad 0 \leq i \leq k, \quad (5.66)$$

$$(b) \quad P_\Theta \tilde{\mathcal{I}}_h(x^{k+1}) = \tilde{\mathcal{I}}_h(x^{k+1}), \quad (5.67)$$

where  $\tilde{\mathcal{I}}_h(x^{k+1})$  is the interpolant of  $x^{k+1}$  as defined in Section 4.

The proofs of these facts are immediate.

**Lemma 5.10.** Suppose  $f \in H_{g,\Theta}^1$  is periodic, i.e.,  $f(x + x_k^h) = f(x)$  for all  $k$ . Then  $P_\Theta f$  is also periodic.

*Proof.* Let  $\tilde{f}(x) = f(x + x_k^h)$ . Then  $[P_\Theta \tilde{f}](x) = [P_\Theta f](x + x_k^h)$ . Now  $f(x) = \tilde{f}(x)$  since  $f$  is periodic, and thus, from the uniqueness of the projection  $P_\Theta$ , we have  $[P_\Theta f](x + x_k^h) = [P_\Theta f](x)$ , i.e.,  $P_\Theta f$  is periodic.  $\square$

**Remark 39.** We note that, if

$$v(x) = \sum_{i \in \mathbb{Z}} c_i^h \phi_i^h(x)$$

is a periodic function, i.e.,  $v(x + x_k^h) = v(x)$  for any  $k$ , then  $v$  is a constant. This could be shown as follows. Since  $v(x + x_k^h) = v(x)$ , we have

$$v(x + x_k^h) = \sum_{i \in \mathbb{Z}} c_i^h \phi_i^h(x + x_k^h) = \sum_{i \in \mathbb{Z}} c_{i+k}^h \phi_i^h(x) = \sum_{i \in \mathbb{Z}} c_i^h \phi_i^h(x) = v(x),$$

which implies that

$$\sum_{i \in \mathbb{Z}} [c_{i+k}^h - c_i^h] \phi_i^h(x) = 0, \quad \text{for all } x \in \mathbb{R}.$$

Using (5.19), we can show that  $\{\phi_i^h\}_{i \in \mathbb{Z}}$  is linearly independent in  $\mathbb{R}$ . Thus we deduce that  $c_{i+k}^h = c_i^h = C$  (constant), for all  $i \in \mathbb{Z}$ . Recalling that  $\{\phi_i^h\}_{i \in \mathbb{Z}}$  forms a partition of unity, we get  $v(x) = C \sum_{i \in \mathbb{Z}} \phi_i^h(x) = C$ .

We now define

$$\xi_{k+1}^\ominus(x) \equiv x^{k+1} - P_\Theta x^{k+1}, \tag{5.68}$$

which, together with the next lemma, will play a central role in the final result of this section.

**Lemma 5.11.** Let  $\xi_{k+1}^\ominus(x)$  be as defined in (5.68) and consider

$$\xi_{k+1}^h(x) = x^{k+1} - \sum_{i \in \mathbb{Z}} (x_i^h)^{k+1} \phi_i^h(x)$$

as defined in (4.36). Then

$$\xi_{k+1}^\ominus '(x) = \xi_{k+1}^h '(x). \tag{5.69}$$

*Proof.* We first note, from the definition of  $\tilde{\mathcal{I}}_h x^{k+1}$ , that  $\xi_{k+1}^h(x) = x^{k+1} - \tilde{\mathcal{I}}_h x^{k+1}$ . Now, using (5.67), we have

$$\begin{aligned} \xi_{k+1}^\ominus &= x^{k+1} - P_\Theta x^{k+1} \\ &= x^{k+1} - \tilde{\mathcal{I}}_h x^{k+1} + \tilde{\mathcal{I}}_h x^{k+1} - P_\Theta x^{k+1} \\ &= \xi_{k+1}^h - P_\Theta [x^{k+1} - \tilde{\mathcal{I}}_h x^{k+1}] \\ &= \xi_{k+1}^h - P_\Theta [\xi_{k+1}^h]. \end{aligned} \tag{5.70}$$

But we know from Lemma 4.1 that  $\xi_{k+1}^h(x)$  is periodic, and therefore from Lemma 5.10 and Remark 39 we infer that  $P_\Theta [\xi_{k+1}^h]$  is a constant. Thus, from (5.70), we get

$$\xi_{k+1}^\ominus '(x) = \xi_{k+1}^h '(x),$$

which is the desired result. □

*Proof of Theorem 5.1.* The proof will be given in several steps.

**1.** Let  $E$  be the extension operator satisfying (3.50). Then, for  $x \in B_H \equiv B_H(0)$ , we have

$$[u_0 - u_h](x) = [u_0 - P_\Theta(Eu_0) - \{Eu_h - P_\Theta(Eu_h)\} + \{P_\Theta(Eu_0) - P_\Theta(Eu_h)\}](x),$$

and therefore

$$(u_0' - u_h')(x) = \{u_0 - P_\Theta(Eu_0)\}'(x) - \delta_h'(x) + \rho_h'(x), \tag{5.71}$$

where

$$\delta_h = Eu_h - P_\Theta(Eu_h), \tag{5.72}$$

$$\rho_h = P_\Theta(Eu_0) - P_\Theta(Eu_h). \tag{5.73}$$

Since  $u_0 = Eu_0$  in  $B_H(0)$ , from Taylor’s theorem we have

$$Eu_0(x) = \sum_{j=0}^k \frac{u_0^{(j)}(0)}{j!} x^j + \frac{u_0^{(k+1)}(0)}{(k+1)!} x^{k+1} + R_{k+1}(Eu_0)(x), \tag{5.74}$$

where  $R_{k+1}(Eu_0)(x)$  is the remainder given by

$$R_{k+1}(Eu_0)(x) = \frac{1}{(k+1)!} \int_0^x (x-t)^{k+1} (Eu_0)^{(k+2)}(t) dt. \tag{5.75}$$

Since  $P_\Theta$  is a linear operator, we have

$$P_\Theta(Eu_0)(x) = \sum_{j=0}^k \frac{u_0^{(j)}(0)}{j!} P_\Theta x^j + \frac{u_0^{(k+1)}(0)}{(k+1)!} P_\Theta x^{k+1} + P_\Theta R_{k+1}(Eu_0)(x). \tag{5.76}$$

We know from (5.66) that  $P_\Theta x^j = x^j, 0 \leq j \leq k$ . Therefore, by first subtracting (5.76) from (5.74), then differentiating the identity, and finally using (5.69), we have

$$\begin{aligned} & \{Eu_0 - P_\Theta(Eu_0)\}'(x) \\ &= \frac{u_0^{(k+1)}(0)}{(k+1)!} \{x^{k+1} - P_\Theta x^{k+1}\}'(x) + [R_{k+1}(Eu_0)]'(x) - [P_\Theta R_{k+1}(Eu_0)]'(x) \\ &= \frac{u_0^{(k+1)}(0)}{(k+1)!} \xi_{k+1}^\Theta'(x) + [R_{k+1}(Eu_0)]'(x) - [P_\Theta R_{k+1}(Eu_0)]'(x) \\ &= \frac{u_0^{(k+1)}(0)}{(k+1)!} \xi_{k+1}^h'(x) + [R_{k+1}(Eu_0)]'(x) - [P_\Theta R_{k+1}(Eu_0)]'(x). \end{aligned} \tag{5.77}$$

Thus from (5.71), (5.77), and using  $e_h(x) \equiv [u_0 - u_h](x)$ , we get for  $x \in B_H$ ,

$$\begin{aligned} e_h'(x) - \frac{u_0^{(k+1)}(0)}{(k+1)!} \xi_{k+1}^\Theta'(x) \\ = [R_{k+1}(Eu_0)]'(x) - [P_\Theta R_{k+1}(Eu_0)]'(x) - \delta'_h + \rho'_h. \end{aligned} \tag{5.78}$$

**2.** From (5.75), we have

$$[R_{k+1}(Eu_0)]'(x) = \frac{1}{k!} \int_0^x (x-t)^k (Eu_0)^{(k+2)}(t) dt,$$

and since  $\|u_0\|_{W_\infty^{k+2}(B_{2H})} \leq C$ , we have, for  $x \in B_{2H}$ ,

$$\begin{aligned} \int_{B_H} |[R_{k+1}(Eu_0)]'|^2 dx &\leq \int_{B_{2H}} g |[R_{k+1}(Eu_0)]'|^2 dx \\ &\leq CH^{2k+2} H \|u_0\|_{W_\infty^{k+2}(B_{2H})}^2. \end{aligned} \tag{5.79}$$

Similarly, again from (5.75), we get

$$\begin{aligned} \int_{B_H} |R_{k+1}(Eu_0)|^2 dx &\leq \int_{B_{2H}} g |R_{k+1}(Eu_0)|^2 dx \\ &\leq CH^{2k+4}H \|u_0\|_{W_\infty^{k+2}(B_{2H})}^2. \end{aligned} \tag{5.80}$$

3. It can be shown from the definition of  $g(x)$  that, for  $0 \leq j \leq k + 1$ ,

$$\int_{2H}^\infty gx^{2j} dx = \int_{2H}^\infty e^{-\alpha(x-H)} x^{2j} dx \leq Ce^{-\alpha H}, \tag{5.81}$$

where  $C$  depends on  $k + 1$ . Now, from (5.65) we get

$$\int_{B_H} |[P_\Theta R_{k+1}(Eu_0)]'|^2 dx \leq \|P_\Theta R_{k+1}(Eu_0)\|_{1,g}^2 \leq C \|R_{k+1}(Eu_0)\|_{1,g}^2. \tag{5.82}$$

We note that, from (5.74), we have

$$[R_{k+1}(Eu_0)]'(x) = (Eu_0)'(x) - \sum_{j=0}^k \frac{u_0^{(j+1)}(0)}{(j+1)!} x^j.$$

Therefore, using (5.81) and the fact that

$$\int_{2H}^\infty g|(Eu_0)'|^2 dx \leq e^{-\alpha H} \int_{2H}^\infty |(Eu_0)'|^2 dx \leq e^{-\alpha H} |Eu_0|_{H^1(\mathbb{R})}^2,$$

we have

$$\begin{aligned} &\int_{2H}^\infty g|[R_{k+1}(Eu_0)]'|^2 dx \\ &\leq C \int_{2H}^\infty g|(Eu_0)'|^2 dx + C \sum_{j=0}^k \left( \frac{u_0^{(j+1)}(0)}{(j+1)!} \right)^2 \int_{2H}^\infty gx^{2j} dx \\ &\leq Ce^{-\alpha H} \{ |Eu_0|_{H^1(\mathbb{R})}^2 + C \|u_0\|_{W_\infty^{k+1}(B_{2H})}^2 \} \\ &\leq Ce^{-\alpha H} \{ \|u_0\|_{H^1(\Omega)}^2 + C \|u_0\|_{W_\infty^{k+2}(B_{2H})}^2 \}, \end{aligned} \tag{5.83}$$

where  $C$  depends on  $k$ . Similarly, we can show that

$$\int_{-\infty}^{-2H} g|[R_{k+1}(Eu_0)]'|^2 dx \leq Ce^{-\alpha H} \{ \|u_0\|_{H^1(\Omega)}^2 + C \|u_0\|_{W_\infty^{k+2}(B_{2H})}^2 \},$$

which together with (5.83) imply that

$$\int_{\mathbb{R}-B_{2H}} g|[R_{k+1}(Eu_0)]'|^2 dx \leq Ce^{-\alpha H} \{ \|u_0\|_{H^1(\Omega)}^2 + C \|u_0\|_{W_\infty^{k+2}(B_{2H})}^2 \}. \tag{5.84}$$

Using similar arguments, we can show that

$$\int_{\mathbb{R}-B_{2H}} g|R_{k+1}(Eu_0)|^2 dx \leq Ce^{-\alpha H} \{ \|u_0\|_{L^2(\Omega)}^2 + C\|u_0\|_{W_\infty^{k+2}(B_{2H})}^2 \}. \quad (5.85)$$

Now, combining (5.79), (5.80), (5.82), (5.84), and (5.85) we get

$$\begin{aligned} & \int_{B_H} |[P_\Theta R_{k+1}(Eu_0)]'|^2 dx \\ & \leq C\|R_{k+1}(Eu_0)\|_{1,g}^2 \\ & = C \int_{\mathbb{R}} g|[R_{k+1}(Eu_0)]'|^2 dx + C\bar{\Theta} \int_{\mathbb{R}} g|R_{k+1}(Eu_0)|^2 dx \\ & \leq C(1 + \bar{\Theta}H^2)H^{2k+2}H\|u_0\|_{W_\infty^{k+2}(B_{2H})}^2 \\ & \quad + C(1 + \bar{\Theta})e^{-\alpha H} \{ \|u_0\|_{H^1(\Omega)}^2 + C\|u_0\|_{W_\infty^{k+2}(B_{2H})}^2 \}. \end{aligned} \quad (5.86)$$

4. We first note from (5.72) that

$$\int_{B_H} \delta'_h{}^2 dx \leq \|\delta_h\|_{1,g}^2 = \|Eu_h - P_\Theta(Eu_h)\|_{1,g}^2. \quad (5.87)$$

Let  $P_\Theta(Eu_h) = \tilde{\mathcal{I}}_h^* Eu_h + \mathcal{E}$ . Then  $\mathcal{E} \in S_h$ . Now from Lemma 5.6 and the definition of  $P_\Theta$ , we have, for all  $v \in S_h$ ,

$$\begin{aligned} B_\Theta(\mathcal{E}, v) &= B_\Theta(P_\Theta(Eu_h) - \tilde{\mathcal{I}}_h^* Eu_h, v) \\ &= B_\Theta(Eu_h - \tilde{\mathcal{I}}_h^* Eu_h, v) \\ &\leq C\|Eu_h - \tilde{\mathcal{I}}_h^* Eu_h\|_{1,g}\|v\|_{1,g^{-1}}, \end{aligned}$$

and hence from Lemma 5.7 we get

$$\|E\|_{1,g} \leq C \sup_{v \in S_h} \frac{B_\Theta(E, v)}{\|v\|_{1,g^{-1}}} \leq C\|Eu_h - \tilde{\mathcal{I}}_h^* Eu_h\|_{1,g}.$$

Thus,

$$\begin{aligned} \|Eu_h - P_\Theta(Eu_h)\|_{1,g} &\leq \|Eu_h - \tilde{\mathcal{I}}_h^* Eu_h\|_{1,g} + \|\mathcal{E}\|_{1,g} \\ &\leq C\|Eu_h - \tilde{\mathcal{I}}_h^* Eu_h\|_{1,g}. \end{aligned} \quad (5.88)$$

We now estimate the right-hand side of the above inequality. We first note that  $Eu_h(x) = u_h(x)$  for  $x \in \Omega$ . Consider  $\underline{\Omega} \subset \Omega$  such that (see Remark 35 on page 61)

$$B_{2H} \subset \underline{\Omega} \quad \text{and} \quad \tilde{\mathcal{I}}_h^* Eu_h|_{\underline{\Omega}} = Eu_h|_{\underline{\Omega}} = u_h|_{\underline{\Omega}}.$$



Therefore, from Lemma 5.3 and using (5.5),

$$\begin{aligned} \int_{\mathbb{R}} g[(Eu_h - \tilde{\mathcal{I}}_h^* Eu_h)']^2 dx &= \int_{\mathbb{R}-\Omega} g[(Eu_h - \tilde{\mathcal{I}}_h^* Eu_h)']^2 dx \\ &\leq e^{-\alpha H} [|Eu_h|_{H^1(\mathbb{R})}^2 + |\tilde{\mathcal{I}}_h^* Eu_h|_{H^1(\mathbb{R})}^2] \\ &\leq Ce^{-\alpha H} [|Eu_h|_{H^1(\mathbb{R})}^2 + \frac{1}{h^2} \|Eu_h\|_{L_2(\mathbb{R})}^2] \\ &\leq \frac{C}{h^2} e^{-\alpha H} \|u_h\|_{H^1(\Omega)}^2 \\ &\leq \frac{C}{h^2} e^{-\alpha H} \|u_0\|_{H^1(\Omega)}^2. \end{aligned} \tag{5.89}$$

Similarly, we can show using Remark 36 (on page 62) that

$$\int_{\mathbb{R}} g[Eu_h - \tilde{\mathcal{I}}_h^* Eu_h]^2 dx \leq Ce^{-\alpha H} \|u_0\|_{L_2(\Omega)}^2,$$

and thus combining this estimate with (5.89) we get

$$\|Eu_h - \tilde{\mathcal{I}}_h^* Eu_h\|_{1,g}^2 \leq \frac{C\bar{\Theta}}{h^2} e^{-\alpha H} \|u_0\|_{H^1(\Omega)}^2.$$

Now, from (5.87), (5.88), and the above estimate, we get

$$\int_{B_H} \delta'_h{}^2 dx \leq \frac{C\bar{\Theta}}{h^2} e^{-\alpha H} \|u_0\|_{H^1(\Omega)}^2. \tag{5.90}$$

5. We first note from (5.73) that

$$\int_{B_H(0)} \rho'_h{}^2 dx \leq \|\rho_h\|_{1,g}^2 = \|P_{\Theta}(Eu_0) - P_{\Theta}(Eu_h)\|_{1,g}^2. \tag{5.91}$$

Now, using (5.4), we have, for all  $v \in S_h$ ,

$$\begin{aligned} B_{\Theta}(\rho_h, v) &= B_{\Theta}(P_{\Theta}Eu_0 - P_{\Theta}Eu_h, v) \\ &= B_{\Theta}(Eu_0 - Eu_h, v) \\ &= B^{\Omega}(Eu_0 - Eu_h, v) + B^{\mathbb{R}-\Omega}(Eu_0 - Eu_h, v) + \Theta D^{\mathbb{R}}(Eu_0 - Eu_h, v) \\ &= B^{\Omega}(u_0 - u_h, v) + B^{\mathbb{R}-\Omega}(Eu_0 - Eu_h, v) + \Theta D^{\mathbb{R}}(Eu_0 - Eu_h, v) \\ &= B^{\mathbb{R}-\Omega}(Eu_0 - Eu_h, v) + \Theta D^{\mathbb{R}-\Omega}(Eu_0 - Eu_h, v) + \Theta D^{\Omega}(Eu_0 - Eu_h, v) \\ &= B_{\Theta}^{\mathbb{R}-\Omega}(Eu_0 - Eu_h, v) + \Theta D^{\Omega}(u_0 - u_h, v). \end{aligned} \tag{5.92}$$

Also, for  $v \in S_h$ ,

$$\begin{aligned}
 & B_{\bar{\Theta}}^{\mathbb{R}-\Omega}(Eu_0 - Eu_h, v) \\
 &= \int_{\mathbb{R}-\Omega} [(Eu_0 - Eu_h)'v' + \bar{\Theta}(Eu_0 - Eu_h)v] dx \\
 &\leq C \|Eu_0 - Eu_h\|_{1,g,\mathbb{R}-\Omega} \|v\|_{1,g^{-1},\mathbb{R}-\Omega} \\
 &\leq C \|Eu_0 - Eu_h\|_{1,g,\mathbb{R}-\Omega} \|v\|_{1,g^{-1}}, \tag{5.93}
 \end{aligned}$$

where

$$\begin{aligned}
 \|v\|_{1,g^{-1},\mathbb{R}-\Omega}^2 &= \int_{\mathbb{R}-\Omega} g^{-1}v'^2 dx + \bar{\Theta} \int_{\mathbb{R}-\Omega} g^{-1}v^2 dx; \\
 \|Eu_0 - Eu_h\|_{1,g,\mathbb{R}-\Omega}^2 &= \int_{\mathbb{R}-\Omega} g(Eu_0 - Eu_h)'^2 dx \\
 &\quad + \bar{\Theta} \int_{\mathbb{R}-\Omega} g(Eu_0 - Eu_h)^2 dx.
 \end{aligned}$$

From the definition of  $g(x)$ , we can show that

$$\begin{aligned}
 \|Eu_0 - Eu_h\|_{1,g,\mathbb{R}-\Omega}^2 &\leq e^{-\alpha H} \bar{\Theta} \|Eu_0 - Eu_h\|_{H^1(\mathbb{R}-\Omega)}^2 \\
 &\leq C e^{-\alpha H} \bar{\Theta} \|u_0 - u_h\|_{H^1(\Omega)}^2 \\
 &\leq C e^{-\alpha H} \bar{\Theta} \|u_0\|_{H^1(\Omega)}^2. \tag{5.94}
 \end{aligned}$$

Now, using the definition of  $g(x)$  and (5.6) with  $R = 2H$ , we get

$$\begin{aligned}
 \int_{\Omega} g(u_0 - u_h)^2 dx &= \int_{B_{2H}} g(u_0 - u_h)^2 dx + \int_{\Omega - B_{2H}} g(u_0 - u_h)^2 dx \\
 &\leq \|u_0 - u_h\|_{L_2(B_{2H})}^2 + e^{-\alpha H} \|u_0 - u_h\|_{L_2(\Omega)}^2 \\
 &\leq Ch^{2k+2} H \|u_0\|_{H^{k+1}(\Omega)}^2 + e^{-\alpha H} \|u_0\|_{H^1(\Omega)}^2,
 \end{aligned}$$

and therefore

$$\begin{aligned}
 & \frac{\Theta}{\|v\|_{1,g^{-1}}} D^{\Omega}(u_0 - u_h, v) \\
 &= \frac{\Theta}{\|v\|_{1,g^{-1}}} \int_{\Omega} (u_0 - u_h)v dx \\
 &\leq \frac{\Theta}{\|v\|_{1,g^{-1}}} \left( \int_{\Omega} g(u_0 - u_h)^2 dx \right)^{1/2} \left( \int_{\Omega} g^{-1}v^2 dx \right)^{1/2} \\
 &\leq \Theta^{\frac{1}{2}} Ch^{k+1} H^{\frac{1}{2}} \|u_0\|_{H^{k+1}(\Omega)} + \Theta^{\frac{1}{2}} e^{-\alpha H/2} \|u_0\|_{H^1(\Omega)}. \tag{5.95}
 \end{aligned}$$

From the inf-sup condition (5.46) and using (5.92) and (5.93), we have

$$\begin{aligned} \|\rho_h\|_{1,g} &\leq C \sup_{v \in S_h} \frac{B_\Theta(\rho_h, v)}{\|v\|_{1,g^{-1}}} \\ &\leq C \|Eu_0 - Eu_h\|_{1,g,\mathbb{R}^{-\Omega}} + \sup_{v \in S_h} \frac{\Theta}{\|v\|_{1,g^{-1}}} D^\Omega(u_0 - u_h, v), \end{aligned}$$

and thus, using (5.91), (5.94) and (5.95), we have

$$\begin{aligned} \int_{B_H} \rho_h'^2 dx &\leq \|\rho_h\|_{1,g}^2 \\ &\leq C e^{-\alpha H} \bar{\Theta} \|u_0\|_{H^1(\Omega)}^2 + C \Theta h^{2k+2} H \|u_0\|_{H^{k+1}(\Omega)}^2. \end{aligned} \tag{5.96}$$

6. We first note from (5.69) that  $\xi_{k+1}^\Theta(x) = \xi_{k+1}^h(x)$ , where  $\xi_{k+1}^h$  is defined in (4.36). Let  $T(u_0) \equiv \frac{u_0^{(k+1)}(0)}{(k+1)!}$ . Then, from (5.78), we have

$$\begin{aligned} e_h'(x) - T(u_0)\xi_{k+1}^h(x) &= [R_{k+1}(Eu_0)]'(x) - [P_\Theta R_{k+1}(Eu_0)]'(x) - \delta_h' + \rho_h', \end{aligned}$$

and therefore, from (5.79), (5.86), (5.90), and (5.96), we have

$$\begin{aligned} &\int_{B_H} (e_h' - T(u_0)\xi_{k+1}^h)' ^2 dx \\ &\leq C \int_{B_H} |[R_{k+1}(Eu_0)]'|^2 dx + C \int_{B_H} |[P_\Theta R_{k+1}(Eu_0)]'|^2 dx \\ &\quad + C \int_{B_H} \delta_h'^2 dx + C \int_{B_H} \rho_h'^2 dx \\ &\leq C H^{2k+2} H \|u_0\|_{W_\infty^{k+2}(B_{2H})}^2 + C(1 + \bar{\Theta} H^2) H^{2k+2} H \|u_0\|_{W_\infty^{k+2}(B_{2H})}^2 \\ &\quad + C(1 + \bar{\Theta}) e^{-\alpha H} \{ \|u_0\|_{H^1(\Omega)}^2 + C \|u_0\|_{W_\infty^{k+2}(B_{2H})}^2 \} + \frac{C \bar{\Theta}}{h^2} e^{-\alpha H} \|u_0\|_{H^1(\Omega)}^2 \\ &\quad + C e^{-\alpha H} \bar{\Theta} \|u_0\|_{H^1(\Omega)}^2 + C \Theta h^{2k+2} H \|u_0\|_{H^{k+1}(\Omega)}^2 \\ &\leq C \left[ H^{2k+2} H + (1 + \bar{\Theta} H^2) H^{2k+2} H + (1 + \bar{\Theta}) e^{-\alpha H} \right. \\ &\quad \left. + \frac{\bar{\Theta}}{h^2} e^{-\alpha H} + \Theta h^{2k+2} H \right] M(u_0), \end{aligned} \tag{5.97}$$

where

$$M(u_0) = \|u_0\|_{H^{k+1}(\Omega)}^2 + \|u_0\|_{W_\infty^{k+2}(B_{2H})}^2.$$

We will now choose  $\alpha$ ,  $\Theta$ , and  $H$ , where  $H = h^\gamma$  and  $\gamma < 1$ . First we choose  $\gamma$  such that

$$H^{k+2} = h^{k+1}, \quad (5.98)$$

which implies that

$$\begin{aligned} h^{\gamma(k+2)} &= h^{k+1}, \quad \text{or} \\ \gamma(k+2) &= k+1, \quad \text{or} \\ \gamma &= \frac{k+1}{k+2} < 1. \end{aligned}$$

Let  $\epsilon > 0$ , which depends on  $\gamma$ , be such that  $\epsilon^* \equiv 1 - \gamma - \epsilon > 0$ . We will now choose  $\alpha$  such that

$$e^{-\alpha H} \leq h^{2k+2} h^{2\gamma+2\epsilon} H = h^{2k+4+3\gamma+2\epsilon}. \quad (5.99)$$

This implies that

$$\alpha \geq C_1 (\ln h^{-1}) h^{-\gamma},$$

where  $C_1 = 2k + 4 + 3\gamma + 2\epsilon$ . Since  $h^{-\epsilon} > \ln h^{-1}$  for small  $h$ , we take

$$\alpha \equiv C_1 h^{-(\gamma+\epsilon)}. \quad (5.100)$$

We now choose

$$\bar{\Theta} \equiv (C_2)^2 h^{-2(\gamma+\epsilon)}, \quad C_2 > C_1. \quad (5.101)$$

We note from (5.100) that  $\alpha h = C_1 h^{1-\gamma-\epsilon} = C_1 h^{\epsilon^*} < 1$  for small  $h$ , and  $\lim_{h \rightarrow 0} \alpha h = 0$ . Thus  $\alpha h$  can be made sufficiently small; this was one of the assumptions in Lemma 5.8. Also, by choosing  $C_2$  sufficiently large in (5.101), we can make  $\frac{\alpha^2}{\bar{\Theta}} = (C_1/C_2)^2 \ll 1$ , *i.e.*, sufficiently small, which was another assumption in Lemma 5.8. Thus the conclusion of Lemma 5.8 is true for the choices of  $\alpha$  and  $\bar{\Theta}$  given in (5.100) and (5.101), respectively.

Now, for these choices of  $\gamma$ ,  $\alpha$ , and  $\bar{\Theta}$ , we have

$$\bar{\Theta} h^{2k+2} = \bar{\Theta} h^{2(\gamma+\epsilon)} h^{2k} h^{2(1-\gamma-\epsilon)} = C_2^2 h^{2k+2\epsilon^*}. \quad (5.102)$$

Using (5.98) and (5.102) we have

$$\bar{\Theta} H^2 H^{2k+2} = \bar{\Theta} H^{2k+4} = \bar{\Theta} h^{2k+2} = C_2^2 h^{2k+2\epsilon^*}. \quad (5.103)$$

Also, from (5.99), we get

$$\bar{\Theta} e^{-\alpha H} \leq h^{2k+4} H \bar{\Theta} h^{2\gamma+2\epsilon} \leq C_2^2 h^{2k+2} H, \quad (5.104)$$

and

$$\frac{\bar{\Theta}}{h^2} e^{-\alpha H} \leq h^{2k+2} H \bar{\Theta} h^{2\gamma+2\epsilon} = C_2^2 h^{2k+2} H. \quad (5.105)$$

Thus, using (5.98) and (5.102)–(5.105) in (5.97), we obtain

$$\|e_h' - T(u_0) \xi_{k+1}^h\|_{L_2(B_H)} \leq C h^{k+\epsilon^*} H^{1/2} M(u_0)^{\frac{1}{2}}, \quad (5.106)$$

and hence, using (5.7) with  $\rho = H$ , we have

$$\frac{\|e_h' - T(u_0)\xi_{k+1}'\|_{L_2(B_H)}}{\|e_h'\|_{L_2(B_H)}} \leq Ch^{\epsilon^*},$$

where  $M(u_0)^{\frac{1}{2}}/\|u_0\|_{H^{k+1}(\Omega)} \leq C$ , which is the desired result. □

**Remark 40.** The balancing of various terms in step 6 of the proof of Theorem 5.1 is similar to the balancing used in the proof of superconvergence of FEM solutions (see Babuška and Strouboulis (2001) and Babuška *et al.* (1996)).

**Remark 41.** Assuming that our superconvergence result is valid in  $L_\infty$ , *i.e.*, assuming that for  $x \in B_H$  there exists  $\epsilon^* > 0$ , such that

$$e_h'(x) = A(u_0)h^k\xi_{k+1}'\left(\frac{x}{h}\right) + O(h^{k+\epsilon^*}),$$

we see that the zeros of  $\xi_{k+1}'(\frac{x}{h})$  are the superconvergence points. In Figure 5.1, we have presented the plot of  $\xi_{k+1}'(y)$  for the RKP shape functions,

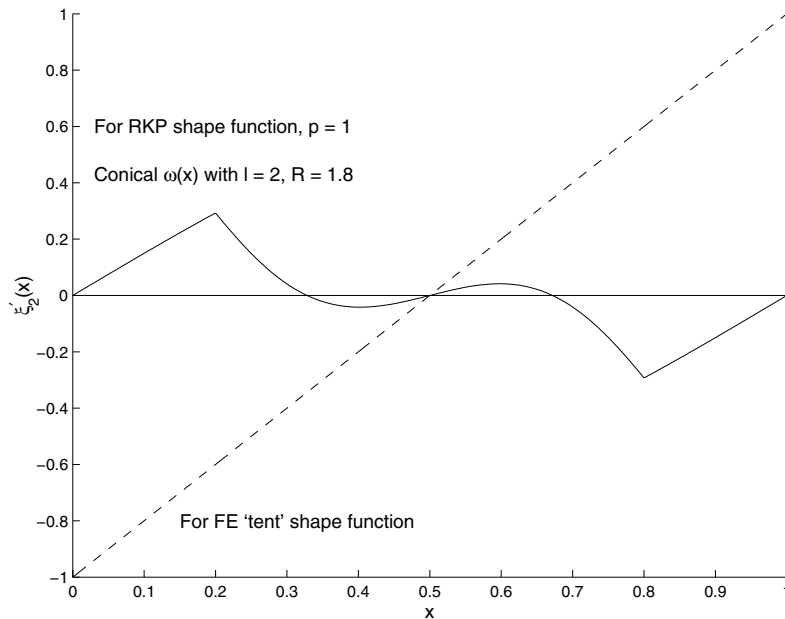


Figure 5.1. The plot of  $\xi_2'(y)$ ,  $0 \leq y \leq 1$  for (a) RKP shape functions, reproducing of order  $k = 1$ , corresponding to the conical weight function with  $l = 2$ ,  $R = 1.8$ , (b) standard ‘tent’ functions used in the FEM.

reproducing of order  $k = 1$ , with respect to the weight function  $w(x)$  given by (4.4) with  $l = 2$  in one dimension. We have also included the plot of  $\xi_{k+1}'(y)$ ,  $k = 1$  (the dashed curve) for the standard tent functions that are used as shape functions in the FEM. We note that  $\xi_2'(y)$  for the tent function has only one zero, whereas  $\xi_2'(y)$  has five zeros. Thus the superconvergence points for the RKP shape function could be distributed quite differently to the corresponding points for standard tent functions in the FEM.

## 6. The generalized finite element method

The idea of the *generalized finite element method* (GFEM) was first introduced in Babuška *et al.* (1994) to address elliptic problems with rough coefficients. This idea was later extended, and called the *partition of unity method* (PUM), in Babuška and Melenk (1997) and Melenk and Babuška (1996). In the current literature, the PUM is referred to as the particle-partition of unity method (Griebel and Schweitzer 2002*a*, 2002*b*, 2002*c*), the method of finite spheres (De and Bathe 2001), the cloud method (Oden, Duarte and Zienkiewicz 1998), the eXtended finite element method (Daux, Moes, Dolbow, Sukumar and Belytschko 2000), and the GFEM (Strouboulis *et al.* 2001*a*, 2001*b*). In this section, we will first describe the GFEM and present the relevant approximation results. We will then discuss the selection of an optimal or near-optimal approximating space, to be used in the GFEM, in certain situations.

### 6.1. Description of the GFEM and related approximation results

In this section we will discuss the GFEM in the context of general particle-shape function systems, which were discussed in Section 3.3. Suppose  $u_0$  is the solution of our model problem (2.1), (2.2) (or (2.3)). We consider a family  $\{\mathcal{M}^\nu\}_{\nu \in N}$  of particle-shape function systems satisfying assumptions A1–A7 with  $k = 0$  and  $\mathcal{A}_x^\nu = I$ ; assumption A5 then reads

$$\sum_{\underline{x} \in X^\nu} \phi_{\underline{x}}^\nu(x) = 1, \quad \text{for all } x \in \mathbb{R}^n. \quad (6.1)$$

The partition of unity (6.1) is the starting point of the GFEM. We will need additional assumptions on  $\{\mathcal{M}^\nu\}_{\nu \in N}$ , namely,

$$\|\phi_{\underline{x}}^\nu\|_{L^\infty(\mathbb{R}^n)} \leq C_1 \quad (6.2)$$

and

$$\|\nabla \phi_{\underline{x}}^\nu\|_{L^\infty(\mathbb{R}^n)} \leq \frac{C_2}{\text{diam}(\eta_x^\nu)}, \quad (6.3)$$

for all  $\underline{x} \in X^\nu$ , and all  $\nu \in N$ . In (6.3), we implicitly assume that  $q > (n/2) + 1$ . We also assume that there is a constant  $C$  such that

$$\text{diam}(\eta_{\underline{x}}^\nu) \leq C, \quad \text{for all } \underline{x} \in X^\nu \text{ and } \nu.$$

For each  $\underline{x} \in X^\nu$ , we assume that we have a finite-dimensional space  $V_{\underline{x}}^\nu$  of functions with good approximation properties. We refer to  $V_{\underline{x}}^\nu$  as *local approximating spaces*. We define a set of particles  $A_\Omega^\nu$ , namely,

$$A_\Omega^\nu = \{\underline{x} \in X^\nu : \eta_{\underline{x}}^\nu \cap \Omega \neq \emptyset\}, \tag{6.4}$$

for each  $\nu \in N$ . From (6.1) we have

$$\sum_{\underline{x} \in A_\Omega^\nu} \phi_{\underline{x}}^\nu(x) = 1, \quad \text{for all } x \in \Omega. \tag{6.5}$$

For an approximating space on  $\Omega$ , we then consider

$$V^\nu = \left\{ v|_\Omega : v = \sum_{\underline{x} \in A_\Omega^\nu} \phi_{\underline{x}}^\nu \psi_{\underline{x}}^\nu, \text{ where } \psi_{\underline{x}}^\nu \in V_{\underline{x}}^\nu \right\}. \tag{6.6}$$

The GFEM is the Galerkin method (2.7) with  $\tilde{B} = B$  and  $S = V^\nu$ , and we will denote the approximate solution  $u_S$ , obtained from the GFEM, by  $u_{\text{GFEM}}$ . When GFEM is used to approximate the solution  $u_0$  of the Neumann problem,  $V_{\underline{x}}^\nu$  can be any finite-dimensional subspace of  $H^1(\eta_{\underline{x}}^\nu)$ . But when the GFEM is used to approximate the solution  $u_0$  of the Dirichlet problem, with the boundary condition (2.3), the functions in  $V_{\underline{x}}^\nu$  are required to satisfy  $v|_{\eta_{\underline{x}}^\nu \cap \partial\Omega} = 0$ , for particles  $\underline{x}$  for which  $|\eta_{\underline{x}}^\nu \cap \Omega| > 0$ . Thus the approximating space  $V^\nu \subset H_0^1(\Omega)$ .

Our next theorem states an approximation result for  $V^\nu$ . We will follow the ideas presented in Babuška *et al.* (2002a, 1994), Babuška and Melenk (1997), Melenk and Babuška (1996), and Strouboulis *et al.* (2001a, 2001b).

**Theorem 6.1.** Suppose  $u \in H^1(\Omega)$  and suppose, for all  $\underline{x} \in A_\Omega^\nu$ , there exists  $\psi_{\underline{x}}^\nu \in V_{\underline{x}}^\nu$  such that

$$\|u - \psi_{\underline{x}}\|_{L_2(\eta_{\underline{x}}^\nu \cap \Omega)} \leq \epsilon_1(\underline{x}), \tag{6.7}$$

$$\|\nabla(u - \psi_{\underline{x}})\|_{L_2(\eta_{\underline{x}}^\nu \cap \Omega)} \leq \epsilon_2(\underline{x}). \tag{6.8}$$

Then the function

$$u_{ap} = \sum_{\underline{x} \in A_\Omega^\nu} \phi_{\underline{x}}^\nu \psi_{\underline{x}}^\nu \in V^\nu \tag{6.9}$$

satisfies

$$\|u - u_{ap}\|_{L_2(\Omega)} \leq \kappa^{1/2} C_1 \left( \sum_{\underline{x} \in A_\Omega^\nu} \epsilon_1^2(\underline{x}) \right)^{1/2} \tag{6.10}$$

and

$$\|\nabla(u - u_{ap})\|_{L_2(\Omega)} \leq (2\kappa)^{1/2} \left( \sum_{\underline{x} \in A_\Omega^\nu} \left( \frac{C_2}{\text{diam}(\eta_{\underline{x}}^\nu)} \right)^2 \epsilon_1^2(\underline{x}) + C_1^2 \epsilon_2^2(\underline{x}) \right)^{1/2}, \tag{6.11}$$

where  $C_1$  and  $C_2$  are given in (6.2) and (6.3).

*Proof.* We will prove only (6.11), since (6.10) can be proved similarly. Since  $\phi_{\underline{x}}^\nu$ , for  $\underline{x} \in A_\Omega^\nu$ , form a partition of unity for  $\Omega$  (see (6.5)), we have

$$\begin{aligned} & \|\nabla(u - u_{ap})\|_{L_2(\Omega)}^2 \\ &= \|\nabla \sum_{\underline{x} \in A_\Omega^\nu} \phi_{\underline{x}}^\nu (u - \psi_{\underline{x}})\|_{L_2(\Omega)}^2 \\ &\leq 2\| \sum_{\underline{x} \in A_\Omega^\nu} (u - \psi_{\underline{x}}) \nabla \phi_{\underline{x}}^\nu \|_{L_2(\Omega)}^2 + 2\| \sum_{\underline{x} \in A_\Omega^\nu} \phi_{\underline{x}}^\nu \nabla (u - \psi_{\underline{x}}) \|_{L_2(\Omega)}^2. \end{aligned} \tag{6.12}$$

For any  $x \in \Omega$ , the sums  $\sum_{\underline{x} \in A_\Omega^\nu} (u - \psi_{\underline{x}}) \nabla \phi_{\underline{x}}^\nu$  and  $\sum_{\underline{x} \in A_\Omega^\nu} \phi_{\underline{x}}^\nu \nabla (u - \psi_{\underline{x}})$  have at most  $\kappa$  nonzero terms (see Remark 16 and (3.60)). Therefore,

$$\left| \sum_{\underline{x} \in A_\Omega^\nu} (u - \psi_{\underline{x}}) \nabla \phi_{\underline{x}}^\nu \right|^2 \leq \kappa \sum_{\underline{x} \in A_\Omega^\nu} |(u - \psi_{\underline{x}}) \nabla \phi_{\underline{x}}^\nu|^2,$$

and

$$\left| \sum_{\underline{x} \in A_\Omega^\nu} \phi_{\underline{x}}^\nu \nabla (u - \psi_{\underline{x}}) \right|^2 \leq \kappa \sum_{\underline{x} \in A_\Omega^\nu} |\phi_{\underline{x}}^\nu \nabla (u - \psi_{\underline{x}})|^2.$$

Hence, from (6.12), (6.7), (6.8), recalling that  $\text{supp}(\phi_{\underline{x}}^\nu) = \eta_{\underline{x}}^\nu$ , we have

$$\begin{aligned} & \|\nabla(u - u_{ap})\|_{L_2(\Omega)}^2 \\ &\leq 2\kappa \sum_{\underline{x} \in A_\Omega^\nu} \|(u - \psi_{\underline{x}}) \nabla \phi_{\underline{x}}^\nu\|_{L_2(\Omega)}^2 + 2\kappa \sum_{\underline{x} \in A_\Omega^\nu} \|\phi_{\underline{x}}^\nu \nabla (u - \psi_{\underline{x}})\|_{L_2(\Omega)}^2 \\ &= 2\kappa \sum_{\underline{x} \in A_\Omega^\nu} \|(u - \psi_{\underline{x}}) \nabla \phi_{\underline{x}}^\nu\|_{L_2(\Omega \cap \eta_{\underline{x}}^\nu)}^2 + 2\kappa \sum_{\underline{x} \in A_\Omega^\nu} \|\phi_{\underline{x}}^\nu \nabla (u - \psi_{\underline{x}})\|_{L_2(\Omega \cap \eta_{\underline{x}}^\nu)}^2 \\ &\leq 2\kappa \sum_{\underline{x} \in A_\Omega^\nu} \left( \left( \frac{C_2}{\text{diam}(\eta_{\underline{x}}^\nu)} \right)^2 \epsilon_1^2(\underline{x}) + C_1^2 \epsilon_2^2(\underline{x}) \right), \end{aligned}$$

which is the desired result. □

**Remark 42.** We note that  $\epsilon_1(\underline{x})$ ,  $\epsilon_2(\underline{x})$  in (6.7), (6.8) depend on the parameter  $\nu$ .



We will now show that both terms of estimate (6.11) are of the same order with additional assumptions on  $V^\nu$ . These additional assumptions depend on the boundary conditions of the approximated function.

**Theorem 6.2.** Suppose  $u_0 \in H^1(\Omega)$  is the solution of the Neumann problem (2.1), (2.2), and suppose there exists  $\psi_{\underline{x}}^\nu \in V_{\underline{x}}^\nu$ ,  $\underline{x} \in A_\Omega^\nu$ , such that (6.7) and (6.8) are satisfied. Moreover, assume that, for  $\underline{x} \in A_\Omega^h$ , the space  $V_{\underline{x}}^\nu$  contains constant functions and that

$$\inf_{\lambda \in \mathbb{R}} \|v - \lambda\|_{L_2(\eta_{\underline{x}}^\nu \cap \Omega)} \leq C (\text{diam}(\eta_{\underline{x}}^\nu)) \|\nabla v\|_{L_2(\eta_{\underline{x}}^\nu \cap \Omega)}, \quad \text{for all } v \in H^1(\eta_{\underline{x}}^\nu \cap \Omega), \tag{6.13}$$

where  $C$  is independent of  $\underline{x} \in X^\nu$  and  $\nu$ . Then there exists  $\tilde{\psi}_{\underline{x}}^\nu \in V_{\underline{x}}^\nu$  so that the corresponding function,

$$\tilde{u}_{ap} = \sum_{\underline{x} \in A_\Omega^\nu} \phi_{\underline{x}}^\nu \tilde{\psi}_{\underline{x}}^\nu \in V^\nu,$$

satisfies

$$\|u_0 - \tilde{u}_{ap}\|_{H^1(\Omega)} \leq C \left( \sum_{\underline{x} \in A_\Omega^h} \epsilon_2^2(\underline{x}) \right)^{1/2}, \tag{6.14}$$

where  $C$  is independent of  $u_0$  and  $\nu$ .

*Proof.* Let  $\psi_{\underline{x}}^\nu \in V_{\underline{x}}^\nu$ ,  $\underline{x} \in A_\Omega^\nu$ , satisfy (6.7) and (6.8). Define  $\tilde{\psi}_{\underline{x}}^\nu = \psi_{\underline{x}}^\nu + r_{\underline{x}}^\nu$ , where  $r_{\underline{x}}^\nu \in \mathbb{R}$  satisfies

$$\|u_0 - \tilde{\psi}_{\underline{x}}^\nu\|_{L_2(\eta_{\underline{x}}^\nu \cap \Omega)} = \inf_{\lambda \in \mathbb{R}} \|u_0 - \psi_{\underline{x}}^\nu - \lambda\|_{L_2(\eta_{\underline{x}}^\nu \cap \Omega)}. \tag{6.15}$$

Since  $V_{\underline{x}}^\nu$  contains constant functions, it is clear that  $\tilde{\psi}_{\underline{x}}^\nu \in V_{\underline{x}}^\nu$ . Also, from (6.15), (6.13) with  $v = u_0 - \psi_{\underline{x}}^\nu$ , and (6.8), we have

$$\begin{aligned} \|u - \tilde{\psi}_{\underline{x}}^\nu\|_{L_2(\eta_{\underline{x}}^\nu \cap \Omega)} &\leq C \text{diam}(\eta_{\underline{x}}^\nu) \|\nabla(u - \psi_{\underline{x}}^\nu)\|_{L_2(\eta_{\underline{x}}^\nu \cap \Omega)} \\ &\leq C \text{diam}(\eta_{\underline{x}}^\nu) \epsilon_2(\underline{x}). \end{aligned} \tag{6.16}$$

Let  $\tilde{u}_{ap} = \sum_{\underline{x} \in A_\Omega^\nu} \phi_{\underline{x}}^\nu \tilde{\psi}_{\underline{x}}^\nu$ . Recall that  $\phi_{\underline{x}}^\nu$ ,  $\underline{x} \in A_\Omega^\nu$ , is a partition of unity for  $\Omega$ . Then, following the arguments in the proof of Theorem 6.1 and using (3.60), (6.2), we can show that

$$\begin{aligned} \|u - \tilde{u}_{ap}\|_{L_2(\Omega)}^2 &= \left\| \sum_{\underline{x} \in A_\Omega^h} \phi_{\underline{x}}^\nu (u - \tilde{\psi}_{\underline{x}}^\nu) \right\|_{L_2(\Omega)}^2 \\ &\leq \kappa \sum_{\underline{x} \in A_\Omega^h} \|\phi_{\underline{x}}^\nu (u - \tilde{\psi}_{\underline{x}}^\nu)\|_{L_2(\Omega)}^2 \end{aligned}$$

$$\begin{aligned}
&= \kappa \sum_{\underline{x} \in A_{\Omega}^h} \|\phi_{\underline{x}}^{\nu}(u - \tilde{\psi}_{\underline{x}})\|_{L_2(\Omega \cap \eta_{\underline{x}}^{\nu})}^2 \\
&\leq C \sum_{\underline{x} \in A_{\Omega}^h} \|(u - \tilde{\psi}_{\underline{x}})\|_{L_2(\Omega \cap \eta_{\underline{x}}^{\nu})}^2, \tag{6.17}
\end{aligned}$$

and using (6.16) in this inequality we get

$$\|u - \tilde{u}_{ap}\|_{L_2(\Omega)}^2 \leq C \sum_{\underline{x} \in A_{\Omega}^h} (\text{diam}(\eta_{\underline{x}}^{\nu}))^2 \epsilon_2^2(\underline{x}). \tag{6.18}$$

Again, following the arguments in the proof of Theorem (6.1), and using (6.2), (6.3), we can show that

$$\begin{aligned}
&\|\nabla(u - \tilde{u}_{ap})\|_{L_2(\Omega)}^2 \\
&\leq 2\kappa \sum_{\underline{x} \in A_{\Omega}^{\nu}} \|(u - \tilde{\psi}_{\underline{x}}^{\nu})\nabla\phi_{\underline{x}}^{\nu}\|_{L_2(\Omega \cap \eta_{\underline{x}}^{\nu})}^2 + 2\kappa \sum_{\underline{x} \in A_{\Omega}^{\nu}} \|\phi_{\underline{x}}^{\nu}\nabla(u - \tilde{\psi}_{\underline{x}}^{\nu})\|_{L_2(\Omega \cap \eta_{\underline{x}}^{\nu})}^2 \\
&\leq C \sum_{\underline{x} \in A_{\Omega}^h} \frac{1}{(\text{diam}(\eta_{\underline{x}}^{\nu}))^2} \|u - \tilde{\psi}_{\underline{x}}^{\nu}\|_{L_2(\Omega \cap \eta_{\underline{x}}^{\nu})}^2 \\
&\quad + C \sum_{\underline{x} \in A_{\Omega}^h} \|\nabla(u - \tilde{\psi}_{\underline{x}}^{\nu})\|_{L_2(\Omega \cap \eta_{\underline{x}}^{\nu})}^2. \tag{6.19}
\end{aligned}$$

By first noting that  $\nabla(u - \tilde{\psi}_{\underline{x}}^{\nu}) = \nabla(u - \psi_{\underline{x}}^{\nu})$ , and then using (6.16) and (6.8) in the above inequality, we get

$$\|\nabla(u - \tilde{u}_{ap})\|_{L_2(\Omega)}^2 \leq C \sum_{\underline{x} \in A_{\Omega}^{\nu}} \epsilon_2^2(\underline{x}). \tag{6.20}$$

Combining this with (6.18) we get (6.14), where we used that  $\text{diam}(\eta_{\underline{x}}^{\nu}) \leq C$  for all  $\underline{x} \in X^{\nu}$  and  $\nu$ .  $\square$

**Theorem 6.3.** Suppose  $u_0 \in H_0^1(\Omega)$  is the solution of the Dirichlet problem (2.1), (2.3), and suppose  $V_{\underline{x}}^{\nu}$ ,  $\underline{x} \in A_{\Omega}^{\nu}$ , satisfy the following assumptions.

- (a) For all  $\underline{x} \in A_{\Omega}^h$  such that  $\eta_{\underline{x}}^{\nu} \cap \partial\Omega = \emptyset$ ,  $V_{\underline{x}}^{\nu}$  contains constant functions, and (6.7), (6.8), and (6.13) hold.
- (b) For all  $\underline{x} \in A_{\Omega}^h$  such that  $|\eta_{\underline{x}}^{\nu} \cap \partial\Omega| > 0$ , functions  $v \in V_{\underline{x}}^{\nu}$  satisfy  $v|_{\eta_{\underline{x}}^{\nu} \cap \partial\Omega} = 0$ , and there is a constant  $C$ , independent of  $\underline{x}$  and  $\nu$ , such that

$$\|v\|_{L_2(\eta_{\underline{x}}^{\nu} \cap \Omega)} \leq C (\text{diam}(\eta_{\underline{x}}^{\nu})) \|\nabla v\|_{L_2(\eta_{\underline{x}}^{\nu} \cap \Omega)}, \tag{6.21}$$

for all  $v \in H^1(\eta_{\underline{x}}^{\nu} \cap \Omega)$  satisfying  $v = 0$  on  $\partial\Omega$ . Moreover (6.7) and (6.8) hold for  $u$  satisfying  $u|_{\eta_{\underline{x}}^{\nu} \cap \partial\Omega} = 0$ .

Then there exists  $\tilde{\psi}_{\underline{x}}^\nu \in V_{\underline{x}}^\nu$  so that the corresponding function,

$$\tilde{u}_{ap} = \sum_{\underline{x} \in A_\Omega^\nu} \phi_{\underline{x}}^\nu \tilde{\psi}_{\underline{x}}^\nu \in V^\nu,$$

satisfies

$$\|u_0 - \tilde{u}_{ap}\|_{H^1(\Omega)} \leq C \left( \sum_{\underline{x} \in A_\Omega^h} \epsilon_2^2(\underline{x}) \right)^{1/2}, \tag{6.22}$$

where  $C$  is independent of  $u_0$  and  $\nu$ .

*Proof.* We first divide the set  $A_\omega^\nu$  into two disjoint sets, namely,

$$\begin{aligned} A_{\Omega,I}^\nu &= \{\underline{x} \in A_\Omega : \eta_{\underline{x}}^\nu \cap \partial\Omega = \emptyset\}, \quad \text{and} \\ A_{\Omega,B}^\nu &= \{\underline{x} \in A_\Omega : \eta_{\underline{x}}^\nu \cap \partial\Omega \neq \emptyset\}. \end{aligned}$$

Let  $\psi_{\underline{x}}^\nu \in V_{\underline{x}}^\nu$ ,  $\underline{x} \in A_\Omega^\nu$ , satisfy (6.7) and (6.8). Define  $\tilde{\psi}_{\underline{x}}^\nu$ , for  $\underline{x} \in A_{\Omega,I}^\nu$ , as in the proof of Theorem 6.2. We know from assumption (a) that, for  $\underline{x} \in A_{\Omega,I}^\nu$ , (6.13) holds and  $V_{\underline{x}}^\nu$  contains constant functions. Therefore, following the argument leading to (6.16) in Theorem 6.2, we get

$$\|u_0 - \tilde{\psi}_{\underline{x}}^\nu\|_{L_2(\eta_{\underline{x}}^\nu \cap \Omega)} \leq C \text{diam}(\eta_{\underline{x}}^\nu) \epsilon_2(\underline{x}), \quad \underline{x} \in A_{\Omega,I}^\nu. \tag{6.23}$$

For  $\underline{x} \in A_{\Omega,B}^\nu$ , we set  $\tilde{\psi}_{\underline{x}}^\nu = \psi_{\underline{x}}^\nu$ . Now,  $u_0|_{\eta_{\underline{x}}^\nu \cap \partial\Omega} = 0$ , and from assumption (b), we know that  $\psi_{\underline{x}}^\nu|_{\eta_{\underline{x}}^\nu \cap \partial\Omega} = 0$  for  $\underline{x} \in A_{\Omega,B}^\nu$ . Thus, using (6.21), with  $v = u_0 - \psi_{\underline{x}}^\nu$ , and (6.8), we have

$$\begin{aligned} \|u_0 - \tilde{\psi}_{\underline{x}}^\nu\|_{L_2(\eta_{\underline{x}}^\nu \cap \Omega)} &= \|u_0 - \psi_{\underline{x}}^\nu\|_{L_2(\eta_{\underline{x}}^\nu \cap \Omega)} \\ &\leq C \text{diam}(\eta_{\underline{x}}^\nu) \epsilon_2(\underline{x}), \quad \underline{x} \in A_{\Omega,B}^\nu. \end{aligned} \tag{6.24}$$

Following the same steps that lead to (6.17) in the proof of Theorem 6.2, and using (6.23) and (6.24), we get

$$\begin{aligned} \|u_0 - u_{ap}\|_{L_2(\Omega)}^2 &\leq C \sum_{\underline{x} \in A_\Omega^\nu} \|u_0 - \tilde{\psi}_{\underline{x}}^\nu\|_{L_2(\eta_{\underline{x}}^\nu \cap \Omega)}^2 \\ &= C \sum_{\underline{x} \in A_{\Omega,I}^\nu} \|u_0 - \tilde{\psi}_{\underline{x}}^\nu\|_{L_2(\eta_{\underline{x}}^\nu \cap \Omega)}^2 + C \sum_{\underline{x} \in A_{\Omega,B}^\nu} \|u_0 - \tilde{\psi}_{\underline{x}}^\nu\|_{L_2(\eta_{\underline{x}}^\nu \cap \Omega)}^2 \\ &\leq C \sum_{\underline{x} \in A_\Omega^\nu} (\text{diam}(\eta_{\underline{x}}^\nu))^2 \epsilon_2^2(\underline{x}). \end{aligned} \tag{6.25}$$

Similarly, following the steps leading to (6.20) in the proof of Theorem 6.2, we get

$$\|\nabla(u - \tilde{u}_{ap})\|_{L_2(\Omega)}^2 \leq C \sum_{\underline{x} \in A_\Omega^\nu} \epsilon_2^2(\underline{x}), \tag{6.26}$$

and combining this with (6.25), we get (6.22), where we used the assumption that  $\text{diam}(\eta_{\underline{x}}^\nu) \leq C$  for all  $\underline{x} \in X^\nu$  and  $\nu$ .  $\square$

**Remark 43.** It is clear from (6.14) and (2.8) that, if  $u_0$  is the solution of (2.1), (2.2), then

$$\|u_0 - u_{\text{GFEM}}\|_{H^1(\Omega)} \leq C \left( \sum_{\underline{x} \in X^\nu} \epsilon_2^2(\underline{x}) \right)^{1/2},$$

provided the local approximation spaces  $V_{\underline{x}}^\nu$  contain constant functions, and (6.13) holds. The above estimate is also true if  $u_0$  is the solution of (2.1), (2.3) provided conditions (a) and (b) of Theorem 6.3 are satisfied. We note that, in the latter case, *i.e.*, when  $u_0$  satisfies the Dirichlet boundary condition,  $u_0|_{\partial\Omega} = 0$ , the space  $V_{\underline{x}}^\nu$ , corresponding to a particle  $\underline{x}$  such that  $\eta_{\underline{x}}^\nu$  intersects  $\partial\Omega$ , does not need to include constant functions, but the functions in  $V_{\underline{x}}^\nu$  have to satisfy the Dirichlet boundary condition on  $\eta_{\underline{x}}^\nu \cap \partial\Omega$ .

**Remark 44.** Conditions (a), (b) in Theorem 6.3, and (6.13) are known as the *uniform Poincaré property*. These conditions put restrictions on the shapes of the  $\{\eta_{\underline{x}}^\nu\}$ . For a detailed discussion of this property, see Babuška and Melenk (1997).

**Remark 45.** The constant  $C_2$  in (6.3) is related to the ratio of the radius of the largest ball contained in  $\eta_{\underline{x}}^\nu$  to the radius of the smallest ball that contains  $\eta_{\underline{x}}^\nu$ . A similar condition is also assumed in the classical FEM. If this ratio is uniformly bounded for all  $\underline{x} \in A_\Omega^\nu$  and  $\nu$ , then (6.13) holds.

**Remark 46.** In practical computations, one can easily construct particle-shape function systems (with  $k = 0$ ), such that conditions (6.2), (6.3), (6.13), and conditions (a), (b) of Theorem 6.3 are satisfied.

**Remark 47.** We observed that a partition of unity is the starting point for the construction of approximating space for the GFEM. It is important to emphasize that the construction of partition unity for  $k = 0$  is simple: for instance, it could be constructed by Shepard's approach, as discussed in Section 4.

**Remark 48.** We have assumed that our particle-shape function system satisfies A1–A7 with  $k = 0$  (and hence it reproduces polynomials of degree 0), and we have seen that the quality of the approximation in Theorems 6.1–6.3 depends entirely on the approximability properties of the spaces  $V_{\underline{x}}^\nu$ , as quantified by  $\epsilon_1(\underline{x})$  and  $\epsilon_2(\underline{x})$ . If we used a particle-shape function system that reproduced polynomials of degree 1 ( $k = 1$ ), then the space  $V^\nu$  defined in 6.6 would be enlarged, and its approximability would be improved, possibly only marginally, but this improvement would not be directly visible from (6.11) (or (6.14) or (6.18)). Note that Theorems 6.1–6.3 are directed

towards the use of nonpolynomial approximating functions, where the rate of convergence cannot be easily defined.

To clarify this point, suppose that for  $\phi_x^\nu$  we use the usual FE hat functions of degree 1, and  $V_x^\nu$  is the space of constants. Then the GFEM is the classical FEM, with the usual rate of convergence of  $O(h)$ . However, (6.11) (or (6.14) or (6.18)) does not establish this rate. As a second example, let  $V_x^\nu$  be the space of linear polynomials. Then the GFEM is a FE method, but not a usual one. The method has the rate of convergence  $O(h^2)$ , but (6.11) (or (6.14) or (6.18)) only establishes  $O(h)$ .

**Remark 49.** The space  $\mathbb{W}_{\Omega,h}^{k',q}$ , introduced near the end of Section 3.3, is a special case of the space  $V^\nu$ , where we take  $V_x^\nu$  to be  $\mathcal{P}^{k'}(\hat{\eta}_x^h)$ . The proof of Theorem 3.13 is obtained directly from Theorems 6.1–6.3 by considering  $V_x^\nu = \mathcal{P}^{k'}(\hat{\eta}_x^h)$  and applying a standard polynomial approximation result.

The estimates in Theorems 6.2 and 6.3 are quite general, and allow us to employ available information on the approximated function  $u$ . Convergence of the approximation can be obtained by considering  $\nu_i \in N$ ,  $i = 1, 2, \dots$ , such that  $h^{\nu_i} \downarrow 0$ , where  $h^\nu$  is defined in (3.79). This is reminiscent of the  $h$ -version of the FEM. Convergence of the approximation can also be attained by keeping  $\nu$  fixed, and selecting a sequence of spaces  $V_x^{\nu,i}$ ,  $i = 1, 2, \dots$ , so that they are complete in  $H^1(\hat{\eta}_x^\nu)$  or in a space  $\mathcal{W}(\hat{\eta}_x^\nu) \subset H^1(\hat{\eta}_x^\nu)$  that is known to include the approximated function  $u_0$ . This is a generalization of the  $p$ -version of the FEM.

6.2. Selection of  $V_x$  and ‘handbook’ problems

We saw in Section 6.1 that it is important to select spaces  $V_x^\nu$  with good local approximation properties. Principles for selecting shape functions that take advantage of available information on the approximated function were formulated in Babuška, Banerjee and Osborn (2001, 2002b). We will use these ideas to discuss the selection of the space  $V_x^\nu$ . In this section we will suppress  $\nu$  in our notation.

Let  $H_1(\eta_x)$  and  $H_2(\eta_x)$  be two Hilbert spaces, and suppose  $H_2(\eta_x) \subset H_1(\eta_x)$ . Then

$$d_n(H_2, H_1) = \inf_{\substack{S_n \subset H_1 \\ \dim S_n = n}} \sup_{u \in H_2} \inf_{\substack{\chi \in S_n \\ \|u - \chi\|_{H_1}}} \|u - \chi\|_{H_1}$$

is called the  $n$ -width of the  $H_2$ -unit ball in  $H_1$ . Let  $V_x^{(n)}$  be an  $n$ -dimensional subspace of  $H_1$ , and let

$$\Psi(V_x^{(n)}, H_2, H_1) = \sup_{\substack{u \in H_2 \\ \|u\|_{H_2} \leq 1}} \inf_{\chi \in V_x^{(n)}} \|u - \chi\|_{H_1},$$

which is called the *sup-inf*. We will write  $\Psi(V_{\underline{x}}^{(n)})$  for  $\Psi(V_{\underline{x}}^{(n)}, H_2, H_1)$  if the spaces  $H_1, H_2$  are evident from context. It is clear that

$$d_n(H_2, H_1) = \inf_{\substack{V_{\underline{x}}^{(n)} \subset H_1 \\ \dim V_{\underline{x}}^{(n)} = n}} \Psi(V_{\underline{x}}^{(n)}, H_2, H_1).$$

If an  $n$ -dimensional subspace  ${}^0V_{\underline{x}}^{(n)}$  satisfies

$$\Psi({}^0V_{\underline{x}}^{(n)}, H_2, H_1) \leq C d_n(H_2, H_1),$$

where  $C > 1$  is a constant, independent of  $n$ , then we will refer to  ${}^0V_{\underline{x}}^{(n)}$  as a *nearly optimal subspace* relative to  $H_1$  and  $H_2$ . An  $n$ -dimensional subspace  ${}^0\bar{V}_{\underline{x}}^{(n)}$  that satisfies

$$\Psi({}^0\bar{V}_{\underline{x}}^{(n)}, H_2, H_1) = d_n(H_2, H_1),$$

is referred to as an *optimal subspace* relative to  $H_1$  and  $H_2$ . An optimal subspace  ${}^0\bar{V}_{\underline{x}}^{(n)}$  leads to the minimal error that can be achieved with an  $n$ -dimensional space, namely,  $d_n(H_2, H_1)$ ; a nearly optimal subspace leads to essentially the same error,  $d_n(H_2, H_1)$ .

Suppose we are interested in using the GFEM to approximate the solution  $u_0$  of the Dirichlet problem,

$$\begin{cases} \Delta u_0 = 0, & \text{in } \Omega, \\ u_0 = g, & \text{on } \partial\Omega, \end{cases}$$

where  $\Omega$  is a bounded domain in  $\mathbb{R}^2$ . Then, for each  $\underline{x} \in X^\nu$ , we seek a finite-dimensional space  $V_{\underline{x}}$  that contains a good approximation  $\psi_{\underline{x}}$  to  $u_0$  on  $\eta_{\underline{x}}$  (cf. (6.7), (6.8)). This will be done by taking advantage of the available information on  $u_0|_{\eta_{\underline{x}} \cap \Omega}$ , namely that  $u_0|_{\eta_{\underline{x}} \cap \Omega}$  is harmonic. We now illustrate this procedure.

We suppose that  $\eta_{\underline{x}}$  is a disk in  $\mathbb{R}^2$  and, for the sake of simplicity, suppose  $\eta_{\underline{x}}$  is the unit disk. Let  $H_1 = \{u \in H^1(\mathring{\eta}_{\underline{x}}) : u \text{ is harmonic in } \eta_{\underline{x}}\}$ . For the space  $H_2$ , we use  $\mathcal{W}(\mathring{\eta}_{\underline{x}})$ , a (regularity) space known to contain  $u_0$ . More precisely, we suppose  $\mathcal{W}(\mathring{\eta}_{\underline{x}})$  is a linear manifold in  $\{u \in H^1(\mathring{\eta}_{\underline{x}}) : u \text{ is harmonic}\}$  and that  $\|u\|$  is a norm on  $\mathcal{W}(\mathring{\eta}_{\underline{x}})$  that is rotationally invariant and satisfies  $\|u\|_{H^1(\mathring{\eta}_{\underline{x}})} \leq \|u\|$ , for all  $u \in \mathcal{W}(\mathring{\eta}_{\underline{x}})$ . Moreover, we assume  $\mathcal{W}(\mathring{\eta}_{\underline{x}})$  is complete with respect to  $\|\cdot\|$ , i.e.,  $\{\mathcal{W}(\mathring{\eta}_{\underline{x}}), \|\cdot\|\}$  is a Hilbert space. We note that  $\mathcal{W}(\mathring{\eta}_{\underline{x}})$  could be any higher-order (isotropic) Sobolev space.

It is well known that any  $u \in H_1$  is characterized by its trace on the boundary  $I = \partial\eta_{\underline{x}}$ ; these traces will be in

$$S = \{u : \mathbb{R} \rightarrow \mathbb{R} : u \text{ is } 2\pi\text{-periodic, } u \in H^{1/2}(I)\}.$$

Any  $u \in S$  can be expanded in its Fourier series

$$u(\theta) = a_0 + \sum_{k=1}^{\infty} (a_k \cos k\theta + b_k \sin k\theta). \tag{6.27}$$

It is immediate that

$$|u|_{H^{1/2}(I)}^2 = a_0^2 + \sum_{k=1}^{\infty} (a_k^2 + b_k^2)k,$$

where  $|u|_{H^{1/2}(I)}$  is a Sobolev norm of order 1/2 on  $I$ , and the series in (6.27) converges in  $H^{1/2}(I)$ -norm. So we have a one-to-one correspondence between  $u(r, \theta) \in H_1$  ( $(r, \theta)$  are polar coordinates) and  $u(\theta) \in S$ , which we express by writing  $u(r, \theta) \sim u(\theta)$ . We easily find that

$$\|u\|_{H^1(\hat{\eta}_x)}^2 = |u|_{H^{1/2}(I)}^2 = a_0^2 + \sum_{j=1}^{\infty} (a_j^2 + b_j^2)j. \tag{6.28}$$

Thus we identify the space  $H_1$  with  $H^{1/2}(I)$ .

Since  $\|u\|$  is rotationally invariant, the corresponding norm on  $u(\theta)$  will be translation-invariant, and we can thus show that

$$\|u\|^2 = a_0^2 + \sum_{j=1}^{\infty} (a_j^2 + b_j^2)j\beta_j, \tag{6.29}$$

where, since  $\|u\|_{H^1(\hat{\eta}_x)} \leq \|u\|$ , we have  $\beta_j \geq 1$ . If we now define

$$H^\beta(I) = \{u \in S : |u|_\beta < \infty\},$$

where

$$|u|_\beta^2 = a_0^2 + \sum_{k=1}^{\infty} (a_k^2 + b_k^2)k\beta_k, \tag{6.30}$$

then we see that  $u(r, \theta) \in H_2$  if and only if  $u(\theta) \in H^\beta(I)$  and  $\|u\| = |u|_\beta$ . We thus identify the space  $H_2$  with  $H^\beta(I)$ .

We will now find an optimal subspace  ${}^0V_x^{(n)}$  relative to  $H_1$  and  $H_2$ . We will exploit the correspondence  $u(r, \theta) \sim u(\theta)$ , and find  ${}^0V_x^{(n)}$  by first identifying an optimal subspace relative to  $\bar{H}_1 = H^{1/2}(I)$  and  $\bar{H}_2 = H^\beta(I)$ .

Let  $M_n = \{m_1, m_2, \dots, m_n\}$  be a set of  $n$  positive integers, and consider

$$V^{M_n} = \left\{ u \in H^{1/2}(I) : u = a_0 + \sum_{k \in M_n} (a_k \cos k\theta + b_k \sin k\theta) \right\}. \tag{6.31}$$

Clearly,  $V^{M_n}$  is a  $(2n + 1)$ -dimensional space.

**Lemma 6.4.** Let  $\bar{H}_1 = H^{1/2}(I)$ ,  $\bar{H}_2 = H^\beta(I)$ , where  $\beta = (\beta_1, \beta_2, \dots)$ ,  $\beta_k \geq 1$ , and let  $V^{M_n}$  be as defined in (6.31). Then

$$\Psi(V^{M_n}, \bar{H}_2, \bar{H}_1) = (\gamma(V^{M_n}))^{-\frac{1}{2}}, \quad (6.32)$$

where

$$\gamma(V^{M_n}) = \inf_{i \notin M_n} \beta_i.$$

*Proof.* Consider  $u \in \bar{H}_2$  given by

$$u = a_0 + \sum_{k=1}^{\infty} (a_k \cos k\theta + b_k \sin k\theta).$$

Then, from (6.31) we get

$$\inf_{\chi \in V^{M_n}} |u - \chi|_{\bar{H}_1}^2 = \sum_{k \in N - M_n} (a_k^2 + b_k^2)k,$$

where  $N$  is the set of all positive integers. Therefore, from (6.30) and the definition of  $\gamma(V^{M_n})$  we have

$$\begin{aligned} \inf_{\chi \in V^{M_n}} \frac{|u - \chi|_{\bar{H}_1}^2}{|u|_{\bar{H}_2}^2} &= \frac{\sum_{k \in N - M_n} (a_k^2 + b_k^2)k}{a_0^2 + \sum_{k \in N} (a_k^2 + b_k^2)k\beta_k} \\ &\leq \frac{\sum_{k \in N - M_n} (a_k^2 + b_k^2)k}{\sum_{k \in N - M_n} (a_k^2 + b_k^2)k\beta_k} \\ &\leq \frac{1}{\gamma(V^{M_n})}. \end{aligned}$$

Thus,

$$\sup_{u \in \bar{H}_2} \inf_{\chi \in V^{M_n}} \frac{|u - \chi|_{\bar{H}_1}^2}{|u|_{\bar{H}_2}^2} \leq \frac{1}{\gamma(V^{M_n})}. \quad (6.33)$$

Let  $\epsilon > 0$  be arbitrary. Then there is an  $m_0 \notin M_n$ ,  $m_0 \geq 1$ , such that

$$\beta_{m_0} \leq \gamma(V^{M_n}) + \epsilon. \quad (6.34)$$

Consider  $u_{m_0} = \cos m_0\theta$ . Clearly,  $u_{m_0} \notin V^{M_n}$ , and therefore, from (6.27),

$$\inf_{\chi \in V^{M_n}} |u_{m_0} - \chi|_{\bar{H}_1}^2 = |u_{m_0}|_{\bar{H}_1}^2 = m_0.$$

Also, from (6.30), we have  $|u_{m_0}|_{\bar{H}_2}^2 = m_0\beta_{m_0}$ . Therefore, using (6.34), we get

$$\sup_{u \in \bar{H}_2} \inf_{\chi \in V^{M_n}} \frac{|u - \chi|_{\bar{H}_1}^2}{|u|_{\bar{H}_2}^2} \geq \inf_{\chi \in V^{M_n}} \frac{|u_{m_0} - \chi|_{\bar{H}_1}^2}{|u_{m_0}|_{\bar{H}_2}^2} = \frac{1}{\beta_{m_0}} \geq \frac{1}{\gamma(V^{M_n}) + \epsilon}.$$



From this estimate and (6.33), we have

$$\frac{1}{\gamma(V^{M_n}) + \epsilon} \leq \sup_{u \in \bar{H}_2} \inf_{\chi \in V^{M_n}} \frac{|u - \chi|_{\bar{H}_1}^2}{|u|_{\bar{H}_2}^2} \leq \frac{1}{\gamma(V^{M_n})}.$$

Since  $\epsilon$  is arbitrary, we get (6.32). □

**Lemma 6.5.** Let  $\bar{H}_1 = H^{1/2}(I)$  and  $\bar{H}_2 = H^\beta(I)$ , where  $\beta = (\beta_1, \beta_2, \dots)$ ,  $\beta_k \geq 1$ . Then

$$d_{2n}(H_2, H_1) = (\gamma_n^*)^{-\frac{1}{2}},$$

where

$$\gamma_n^* = \sup_{m_1, m_2, \dots, m_n} \inf_{i \notin M_n} \beta_i.$$

The proof of this theorem follows immediately from Lemma 6.4.

**Theorem 6.6.** Suppose  $H_1 = \{u \in H^1(\eta_{\underline{x}}) : u \text{ is harmonic}\}$  and  $H_2 = \mathcal{W}(\eta_{\underline{x}})$  with the norm  $\|u\|_\beta = |u|_\beta$ , given in (6.30), with  $\beta_j \geq 1$ . Suppose in addition that the sequence  $\beta_j$  is non-decreasing. Then the space

$${}^0V_{\underline{x}}^{(2n+1)} = \text{span}\{r^j \cos j\theta, r^j \sin j\theta\}_{j=0}^n,$$

*i.e.*, the span of first  $(2n + 1)$  harmonic polynomials, is optimal relative to  $H_1$  and any  $H_2$  (*i.e.*, any of the spaces  $H_2$  we are considering).

*Proof.* Using the correspondence  $u(r, \theta) \sim u(\theta)$ , we can study the optimality of a finite-dimensional subspace relative to  $H_1$  and  $H_2$ , by studying the optimality of a subspace relative to  $\bar{H}_1$  and  $\bar{H}_2$ . The result follows directly from Lemma 6.5. □

**Remark 50.** Obviously the condition on  $\beta$  in Theorem 6.6 holds for any (isotropic) Sobolev space.

**Remark 51.** Let us return to the solution of the Dirichlet problem mentioned above. Suppose  $\eta_{\underline{x}}$  is far from the boundary of  $\Omega$ . Then, on  $\eta_{\underline{x}}$ , the character of the solution  $u_0$  is approximately the same in any direction. Thus it is appropriate to embed  $u_0$  in a space with a rotationally invariant norm – a usual (isotropic) Sobolev space, for example. Further, we have learned that, on  $\eta_{\underline{x}}$ ,  $u_0$  is well approximated by harmonic polynomials. The situation is, however, somewhat different when  $\eta_{\underline{x}}$  is near the boundary. Then  $u_0$  would be strongly influenced by the boundary values  $g(x)$ . Hence some other shape functions, constructed, for example, by the ‘handbook’ approach (see below), which themselves reflected these boundary values, would be ‘best’.

Thus the optimal shape functions are the solution of the Laplace equation. This approach could also be used in other situations. Vekua (1967) defines and studies analogues of harmonic polynomials for differential equations

with analytic coefficients. Babuška and Melenk (1997) and Laghrouche and Bettess (2000) construct and use special shape functions for the Helmholtz equation,  $-\Delta u - ku = 0$ . Special shape functions for problems in composite materials were used in Strouboulis, Zhang and Babuška (200x).

In this section, we saw an example of choosing an optimal local approximating space  $V_{\underline{x}}$ , which turned out to be the span of first  $(2n + 1)$  harmonic polynomials. In other problems, different local approximating spaces, consisting of optimal or near-optimal approximating functions, are recommended. These optimal or near-optimal approximating functions are solutions of other boundary value problems (posed on  $\eta_{\underline{x}} \cap \Omega$ ). Such locally posed problems are called *handbook problems* and their solutions, which may be available analytically or computed numerically, are called *handbook functions*. This nomenclature is reminiscent of the solved problems and their solutions (via formulae, tables *etc.*) which are used in engineering (Tada, Paris and Irwin 1973). This idea is also used in commercial codes (Szabo, Babuška and Actis 1998).

One of the main advantages of the GFEM is that only simple meshes are used, which need not reflect the boundary, *e.g.*, uniform finite element meshes. Also, in each  $\eta_{\underline{x}}$ , one can use a space  $V_{\underline{x}}$  of arbitrary dimension (depending on  $\underline{x}$ ).  $V_{\underline{x}}$  could be space of polynomials or any other space of functions depending on the local properties of the approximated function.

Choosing  $V_{\underline{x}}$  to be the space of polynomials of low degree  $p$  (and using  $\{\phi_{\underline{x}}\}$  that are reproducing of order  $k$ ), we obtain the  $h$ -version of the FEM. All other classical versions of the FEM – the  $p$  and  $h$ - $p$  versions – are special cases of the GFEM.

The GFEM, with special shape functions, was effectively used to solve differential equations with rough coefficients and, more generally, in problems with micro-structures: see Babuška *et al.* (1994). Babuška and Osborn (2000) showed that, for differential equations with rough coefficients, the classical FEM can converge arbitrarily slowly. With the GFEM, in contrast, with appropriately chosen shape functions, an exponential rate of convergence can be achieved: see Matache, Babuška and Schwab (2000). The GFEM is a very powerful approach for solving problems with micro-structures: see Strouboulis *et al.* (2001b) and Section 9. The GFEM can also be advantageously used in linear and nonlinear crack propagation problems (Moes, Dolbow and Belytschko 1999, Wells and Sluis 2001, Wells, de Borst and Sluis 2002), and problems with boundary layers (Duarte and Babuška 2002).

## 7. Solutions of elliptic boundary value problems

In this section we will discuss the approximate solution of the model problem (2.1)–(2.2) (or (2.3)), introduced in Section 2, by a meshless method. We will address the Neumann boundary condition (2.2) and the Dirichlet

boundary condition (2.3) separately. These problems have the variational formulation (2.4).

For  $0 < h \leq 1$ , we consider a family of particle-shape function systems

$$\{\mathcal{M}^h\}_{0 < h \leq 1} = \{X^h, \{h_{\underline{x}}^h, \omega_{\underline{x}}^h, \phi_{\underline{x}}^h\}_{\underline{x} \in X^h}\}_{0 < h \leq 1},$$

satisfying assumptions A1–A7 in Section 3.3 and (3.63). Recall that (3.63) is trivially satisfied if the shape functions are reproducing of order  $k$ . The family  $\{\mathcal{M}^h\}_{0 < h \leq 1}$  was introduced in Section 3.3; recall that

$$\sup_{\underline{x} \in X^h} h_{\underline{x}}^h \leq h.$$

We will be interested in assessing the approximation error as  $h \downarrow 0$ .

Let  $u_0$  be the solution of (2.4), where  $\Omega$  is a bounded domain with Lipschitz-continuous boundary. In this section, we will sometimes assume that the boundary of  $\Omega$  is smooth. We will use the space  $\mathbb{V}_{\Omega,h}^{k,q}$ , defined in (3.85), to approximate  $u_0$ . It was shown in Theorem 3.11 that  $\mathbb{V}_{\Omega,h}^{k,q}$  is  $(k + 1, q)$ -regular. Moreover,  $\mathbb{V}_{\Omega,h}^{k,q}$  satisfies the local assumption LA. Recall that  $k$  is the order of the quasi-reproducing shape functions considered in  $\{\mathcal{M}^h\}$  and  $q$  is the smoothness index of these shape functions. The parameters  $k$  and  $q$  are in assumptions A1–A7 and we assume that  $q \leq k + 1$ . We also recall that  $\mathbb{V}_{\Omega,h}^{k,q}$  does not involve all the particles in  $X^h$ ; it only involves particles in the set

$$A_{\Omega}^h = \{\underline{x} \in X^h : \hat{\eta}_{\underline{x}}^h \cap \Omega \neq \emptyset\}. \tag{7.1}$$

Various classes of shape functions can be used for  $\phi_{\underline{x}}^h$  in the system  $\{\mathcal{M}^h\}$ . In Section 4, one such class of shape functions, namely RKP shape functions, were discussed, and references related to other classes of shape functions used in practice were provided.

We note that it is possible to construct particle-shape function systems  $\mathcal{M}^h$ , satisfying A1–A7, such that the set of particles  $X^h \subset \Omega$  and the corresponding  $\mathbb{V}_{\Omega,h}^{k,q}$  have the desired approximation properties. We do not consider such  $\mathbb{V}_{\Omega,h}^{k,q}$  in this section, and we will further remark on this issue in the next subsection.

Let  $u_S = u_h \in \mathbb{V}_{\Omega,h}^{k,q}$  be the approximate solution defined by (2.7) with  $S = \mathbb{V}_{\Omega,h}^{k,q}$ . Since  $\mathbb{V}_{\Omega,h}^{k,q}$  is  $(k + 1, q)$ -regular, we note that  $\mathbb{V}_{\Omega,h}^{k,q} \subset H = H^1(\Omega)$  provided  $q \geq 1$ . Thus  $u_h$  is the solution of

$$\begin{cases} u_h \in \mathbb{V}_{\Omega,h}^{k,q} \\ \tilde{B}(u_0, v) = \int_{\Omega} f v \, dx, \quad \text{for all } v \in \mathbb{V}_{\Omega,h}^{k,q}, \end{cases} \tag{7.2}$$

where the bilinear form  $\tilde{B}$  is either  $B$ , given in (2.5), or a perturbation of  $B$ . Clearly,  $u_h$  is the solution of a Galerkin method. This Galerkin method is a *meshless method* since the construction of the test and the trial space, *i.e.*,  $\mathbb{V}_{\Omega,h}^{k,q}$ , does not require a mesh. As we remarked in Section 1, avoiding mesh generation is one of the main features and advantages of meshless methods.

In this section, we will consider  $u_h$  as an approximation of  $u_0$  and primarily study the error  $u_0 - u_h$ . We set some notation that will be used in this study in the following sections. We define

$$E_{\underline{x}}^h \equiv \overset{\circ}{\eta}_{\underline{x}}^h \cap \partial\Omega, \quad \underline{x} \in A_{\Omega}^h, \quad (7.3)$$

and

$$A_{\partial\Omega}^h = \{\underline{x} \in A_{\Omega}^h : E_{\underline{x}}^h \neq \emptyset\}. \quad (7.4)$$

Thus  $A_{\partial\Omega}^h$  is the set of particles  $\{\underline{x}\}$  such that  $\overset{\circ}{\eta}_{\underline{x}}^h$  has non-empty intersection with  $\partial\Omega$ .

### 7.1. A meshless method for Neumann boundary conditions

In this section we will address the approximation of solution  $u_0$  of (2.1) and (2.2) by the meshless method. The analysis presented here is based on the ideas and results in Babuška (1971) and Babuška and Aziz (1972). See also the references listed in these articles.

The solution  $u_0$  of (2.1), (2.2) can be variationally characterized by (2.4), which is

$$\begin{cases} u_0 \in H^1(\Omega) \\ B(u_0, v) = \int_{\Omega} f v \, dx, \quad \text{for all } v \in H^1(\Omega). \end{cases} \quad (7.5)$$

We wish to approximate  $u_0$  by  $u_h$ , the solution of (7.2) with  $\tilde{B} = B$ . For an error estimate, from (2.8) we have

$$\|u_0 - u_h\|_{H^1(\Omega)} \leq \inf_{\chi \in \mathbb{V}_{\Omega,h}^{k,q}} \|u_0 - \chi\|_{H^1(\Omega)}.$$

Suppose  $u_0 \in H^l(\Omega)$ . Then, since  $\mathbb{V}_{\Omega,h}^{k,q}$  is  $(k+1, q)$ -regular and  $q \geq 1$ , we have

$$\|u_0 - u_h\|_{H^1(\Omega)} \leq Ch^{\mu} \|u_0\|_{H^l(\Omega)},$$

where  $\mu = \min(k, l-1)$ . We summarize this in the following theorem.

**Theorem 7.1.** Suppose  $u_0 \in H^l(\Omega)$ , with  $l \geq 1$ , is the solution of (7.5), where  $\partial\Omega$  is Lipschitz-continuous. Let  $u_h \in \mathbb{V}_{\Omega,h}^{k,q}$ , with  $q \geq 1$ , be the

approximate solution given by (7.2) with  $\tilde{B} = B$ . Then

$$\|u_0 - u_h\|_{H^1(\Omega)} \leq h^\mu \|u_0\|_{H^l(\Omega)}, \tag{7.6}$$

where

$$\mu = \min(k, l - 1). \tag{7.7}$$

We note that the computation of  $u_h$ , in Theorem 7.1, depends on the definition of  $\mathbb{V}_{\Omega,h}^{k,q}$  and involves particles that are also outside  $\Omega$ . In the literature, especially in the engineering literature,  $(t, k^*)$ -regular particle spaces are constructed using particles inside  $\Omega$ , but the support of some of the corresponding particle shape functions could be partly outside  $\Omega$ . The apparent reason for such construction is that the approximate solution is viewed as an interpolant with respect to data inside  $\Omega$ , and hence only the particles that are inside  $\Omega$  are considered. This is certainly not necessary.

The construction of the approximation space  $S$  (in (2.7)) using particles only inside  $\Omega$  may sometimes lead to better conditioning of the underlying linear system. On the other hand, such construction is more expensive and the approximations could show boundary layer behaviour (Babuška *et al.* 200x).

### 7.2. Meshless methods for Dirichlet boundary conditions

In this section we consider the approximation of the solution  $u_0$  of the Dirichlet boundary value problem (2.1) and (2.3) by meshless methods. The variational characterization of  $u_0$  is given by

$$\begin{cases} u_0 \in H_0^1(\Omega) \\ B(u_0, v) = \int_{\Omega} f v \, dx, \quad \text{for all } v \in H_0^1(\Omega). \end{cases} \tag{7.8}$$

The Galerkin method (2.7) to approximate  $u_0$  would require that the approximating space  $S$  be a subspace of  $H = H_0^1(\Omega)$  and thus that the approximating functions satisfy the essential homogeneous Dirichlet boundary condition. Unlike shape functions used in the FEM, the particle shape functions  $\phi_{\underline{x}}^h$  (we consider  $h$  as the parameter), considered in Section 3.3, do not in general satisfy the so-called ‘Kronecker delta’ property, *i.e.*,  $\phi_{\underline{x}}^h(\underline{y}) \neq \delta_{\underline{x}, \underline{y}}$ ,  $\underline{x}, \underline{y} \in X^h$ . This is also true for translation-invariant particle shape functions discussed in Section 3.2 (see Section 4.2). Thus it is difficult to construct a subspace  $S \subset \mathbb{V}_{\Omega,h}^{k,q}$  such that  $S$  could be used in (2.7) as the approximation space and the functions in  $S$  satisfy the Dirichlet boundary condition.

In the literature, several meshless methods have been proposed to approximate the solutions of Dirichlet boundary value problems. They are meshless

methods in the sense that they use  $\mathbb{V}_{\Omega,h}^{k,q}$  as the approximating space. These methods are:

- (1) the penalty method,
- (2) the Lagrange multiplier method,
- (3) the Nitsche and related methods,
- (4) the collocation method,
- (5) combination of meshless and finite element methods,
- (6) the characteristic function method.

In this section we will describe these methods. We note that the GFEM, discussed in Section 6, uses an approximating space different from  $\mathbb{V}_{\Omega,h}^{k,q}$ , and can also be used to approximate the solution of a Dirichlet boundary value problem.

We will assume that the boundary  $\partial\Omega$  of  $\Omega$  is sufficiently smooth. The smoothness assumption on the boundary simplifies the arguments presented here, but various results could be obtained when the boundary is not smooth.

#### *The penalty method*

The main idea of the penalty method is to use a perturbed variational principle. For  $\sigma > 0$ , we consider the bilinear form

$$\tilde{B}(u, v) \equiv B_\sigma(u, v) \equiv B(u, v) + h^{-\sigma} D(u, v), \quad (7.9)$$

where

$$B(u, v) = \int_{\Omega} (\nabla u \cdot \nabla v + uv) \, dx, \quad (7.10)$$

$$D(u, v) = \int_{\partial\Omega} uv \, dx. \quad (7.11)$$

We note that (7.10) is the bilinear form given in (2.5). We consider the solution  $u_h = u_{\sigma,h} \in \mathbb{V}_{\Omega,h}^{k,q}$  of (7.2), namely

$$B_\sigma(u_{\sigma,h}, v) = \int_{\Omega} f v \, dx, \quad \text{for all } v \in \mathbb{V}_{\Omega,h}^{k,q}. \quad (7.12)$$

We note that  $u_{\sigma,h}$  is  $u_S$ , where  $u_S$  is defined in (2.7). For  $v \in H^1(\Omega)$ , let

$$Q_\sigma(v) = B(v, v) + h^{-\sigma} D(v, v) - 2 \int_{\Omega} f v \, dx. \quad (7.13)$$

It is well known that

$$Q_\sigma(u_{\sigma,h}) = \min_{v \in \mathbb{V}_{\Omega,h}^{k,q}} Q_\sigma(v). \quad (7.14)$$

We now present a convergence result for the penalty method.

**Theorem 7.2.** Suppose  $u_0 \in H^l(\Omega) \cap H_0^1(\Omega)$ ,  $l > 3/2$ , is the solution of (7.8). Let  $u_{\sigma,h} \in \mathbb{V}_{\Omega,h}^{k,q}$  be the solution of (7.12). Then, for any  $0 < \epsilon < \min(l - 3/2, 1/2)$ , we have

$$\|u_0 - u_{\sigma,h}\|_{H^1(\Omega)} \leq C(\epsilon)h^\mu \|u_0\|_{H^l(\Omega)}, \tag{7.15}$$

where

$$\mu = \min\left(k, l - 1, \frac{\sigma}{2}, k + \frac{1}{2} - \frac{\sigma}{2} - \epsilon, l - \frac{1}{2} - \frac{\sigma}{2} - \epsilon\right), \tag{7.16}$$

and  $C(\epsilon)$  depends on  $\epsilon$ , but is independent of  $h$  and  $u_0$ .

*Proof.* For any  $v \in H^1(\Omega)$ , we define

$$R_\sigma(v) = B(u_0 - v, u_0 - v) + h^{-\sigma} D\left(\frac{\partial u_0}{\partial n} h^\sigma + v, \frac{\partial u_0}{\partial n} h^\sigma + v\right). \tag{7.17}$$

Then, from Green’s theorem,

$$\begin{aligned} R_\sigma(v) &= B(u_0, u_0) + B(v, v) - 2B(u_0, v) \\ &\quad + h^\sigma D\left(\frac{\partial u_0}{\partial n}, \frac{\partial u_0}{\partial n}\right) + h^{-\sigma} D(v, v) + 2D\left(\frac{\partial u_0}{\partial n}, v\right) \\ &= B(u_0, u_0) + h^\sigma D\left(\frac{\partial u_0}{\partial n}, \frac{\partial u_0}{\partial n}\right) \\ &\quad + B(v, v) + h^{-\sigma} D(v, v) - 2 \int_\Omega f v \, dx \\ &= B(u_0, u_0) + h^\sigma D\left(\frac{\partial u_0}{\partial n}, \frac{\partial u_0}{\partial n}\right) + Q_\sigma(v), \quad \text{for all } v \in H^1(\Omega), \end{aligned}$$

where  $Q_\sigma(v)$  is given by (7.13). Therefore,

$$\min_{v \in \mathbb{V}_{\Omega,h}^{k,q}} R_\sigma(v) = B(u_0, u_0) + h^\sigma D\left(\frac{\partial u_0}{\partial n}, \frac{\partial u_0}{\partial n}\right) + \min_{v \in \mathbb{V}_{\Omega,h}^{k,q}} Q_\sigma(v),$$

and thus from (7.14) we get

$$R_\sigma(u_{\sigma,h}) = \min_{v \in \mathbb{V}_{\Omega,h}^{k,q}} R_\sigma(v).$$

Hence, from (7.17) and the above relation, we have

$$\begin{aligned} \|u_0 - u_{\sigma,h}\|_{H^1(\Omega)}^2 &= B(u_0 - u_{\sigma,h}, u_0 - u_{\sigma,h}) \\ &\leq R_\sigma(u_{\sigma,h}) \\ &\leq R_\sigma(v), \quad \text{for all } v \in \mathbb{V}_{\Omega,h}^{k,q}. \end{aligned} \tag{7.18}$$

Since  $\mathbb{V}_{\Omega,h}^{k,q}$  is  $(k + 1, q)$ -regular with  $q \geq 1$ , there is a  $g_h \in \mathbb{V}_{\Omega,h}^{k,q}$  such that

$$\|u_0 - g_h\|_{H^s(\Omega)} \leq Ch^\mu \|u_0\|_{H^l(\Omega)}, \tag{7.19}$$

where  $\mu = \min(k + 1 - s, l - s)$  and  $0 \leq s \leq 1$ . Now, from (7.17) with  $v = g_h$  and using the Schwartz inequality, we have

$$R_\sigma(g_h) \leq C \left( \|u_0 - g_h\|_{H^1(\Omega)}^2 + h^\sigma \int_{\partial\Omega} \left( \frac{\partial u_0}{\partial n} \right)^2 ds + h^{-\sigma} \int_{\partial\Omega} g_h^2 ds \right). \tag{7.20}$$

We will estimate the right-hand side of the above inequality. We first note that  $u_0 = 0$  on  $\partial\Omega$ . Let  $0 < \epsilon < \min(l - \frac{3}{2}, \frac{1}{2})$ . Then, using a trace inequality and (7.19) with  $s = (1/2) + \epsilon$ , we get

$$\begin{aligned} \|g_h\|_{L_2(\partial\Omega)}^2 &= \|u_0 - g_h\|_{L_2(\partial\Omega)}^2 \leq C(\epsilon) \|u_0 - g_h\|_{H^{\frac{1}{2}+\epsilon}(\Omega)}^2 \\ &\leq C(\epsilon) h^{2\mu_1} \|u_0\|_{H^l(\Omega)}^2, \end{aligned} \tag{7.21}$$

where  $\mu_1 = \min(k + \frac{1}{2} - \epsilon, l - \frac{1}{2} - \epsilon)$ . Also from a trace inequality, we have

$$\left\| \frac{\partial u_0}{\partial n} \right\|_{L_2(\partial\Omega)}^2 \leq C(\epsilon) \|u_0\|_{H^{\frac{3}{2}+\epsilon}(\Omega)}^2. \tag{7.22}$$

Now using (7.21), (7.22), and (7.19), with  $s = 1$ , in (7.20), we get

$$\begin{aligned} R_\sigma(g_h) &\leq C(\epsilon) (h^{2\min(k,l-1)} + h^\sigma + h^{2\mu_1-\sigma}) \|u_0\|_{H^l(\Omega)}^2 \\ &\leq C(\epsilon) h^{2\mu} \|u_0\|_{H^l(\Omega)}^2, \end{aligned} \tag{7.23}$$

where  $\mu = \min(k, l - 1, \frac{\sigma}{2}, k + \frac{1}{2} - \frac{\sigma}{2} - \epsilon, l - \frac{1}{2} - \frac{\sigma}{2} - \epsilon)$ . Finally, combining (7.18) and (7.23) we get the desired result.  $\square$

**Remark 52.** If we consider  $\mathbb{V}_{\Omega,h}^{k,q}$  in Theorem 7.2 such that  $k + 1 \geq l > 3/2$ , then with  $\sigma = l - \frac{1}{2} - \epsilon$  it is easy to see that (7.15) holds with

$$\mu = \frac{1}{2} \left( l - \frac{1}{2} - \epsilon \right).$$

Estimate (7.15) can be improved. We present the following result, based on the analysis in Babuška (1970, 1971, 1972, 1973b), without proof.

**Theorem 7.3.** Suppose  $u_0 \in H^l(\Omega) \cap H_0^1(\Omega)$  is the solution of (7.8). Let  $u_{\alpha,h} \in \mathbb{V}_{\Omega,h}^{k,q}$  be the solution of (7.12). If  $k + 1 \geq l \geq 2$ , then, for any  $\epsilon > 0$ , we have

$$\|u_0 - u_{\alpha,h}\|_{H^1(\Omega)} \leq C(\epsilon) h^{\mu-\epsilon} \|u_0\|_{H^l(\Omega)}, \tag{7.24}$$

where  $C(\epsilon)$  is independent of  $u_0$  and  $h$  but depends on  $\epsilon$ , and  $\mu$  is given by

$$\mu = \min \left( \sigma, l + \sigma - 2, l + \frac{\sigma}{2} - \frac{3}{2}, \frac{k + 1 - \kappa}{k} (l - 1) \right), \tag{7.25}$$

where

$$\kappa = \max \left( 1, \frac{1 + \sigma}{2} \right). \tag{7.26}$$



**Remark 53.** If  $\sigma \geq 1$  in Theorem 7.3, then  $l + \sigma - 2 \geq l + \frac{\sigma}{2} - \frac{3}{2}$ , and therefore  $\mu$  in (7.24) is given by

$$\mu = \min\left(\sigma, l + \frac{\sigma}{2} - \frac{3}{2}, \frac{k + 1 - \kappa}{k}(l - 1)\right), \tag{7.27}$$

where  $\kappa$  is as in (7.26).

**Example.** Consider  $l = 2$  and  $\sigma = 1$  in Theorem 7.3. Then, from (7.26), we have  $\kappa = 1$ , and (7.27) yields  $\mu = \min(1, 1, 1) = 1$ . This is the optimal rate of convergence. For higher values of  $l$  and  $k + 1 \geq l$ , there is a loss in the rate of convergence and we get a sub-optimal rate of convergence, *i.e.*,  $\mu < \min(l - 1, k)$ .

Thus, from Theorem 7.3 we conclude the following.

- It is advantageous to use  $\mathbb{V}_{\Omega,h}^{k,q}$  with higher values of  $k$ , since it leads to higher accuracy. For example, if  $l = 4$  and  $\sigma = 3$ , then  $\kappa = 2$  and (7.27) yields  $\mu = 3(\frac{k-1}{k})$ . Thus higher values of  $k$  will increase accuracy. But higher values of  $k$  reduce the sparsity of the resulting linear system.
- Too small or too large a value of  $\sigma$  may decrease the accuracy significantly. For example, if  $\sigma = 2k + 1$  then  $\kappa = k + 1$ , and (7.27) yields  $\mu = 0$ .

The use of penalty methods was recently suggested in the literature, *e.g.*, Atluri and Shen (2002), Liu (2002) and Zhu and Atluri (1998), without any theoretical analysis. An empirical penalty value of  $\sigma$ , unrelated to  $l$  or  $k$ , was suggested in Zhu and Atluri (1998).

*The Lagrange multiplier method*

The theory of Lagrange multiplier methods, in the context of finite element methods, was developed in Babuška (1973a) (see also Babuška and Aziz (1972)). This theory can also be extended to meshless methods.

It is known (*cf.* Babuška and Aziz (1972)) that the effectiveness of this method depends on a delicate relationship between the approximating space  $S_h^{t,k^*}(\Omega)$  and the space of Lagrange multipliers  $S_{h_1}^{t,k^*}(\partial\Omega)$ , where both  $S_h^{t,k^*}(\Omega)$  and  $S_{h_1}^{t,k^*}(\partial\Omega)$  satisfy an inverse assumption. In the context of meshless methods, we let the approximating space to be the particle space  $\mathbb{V}_{\Omega,h}^{k,q}$ . We know that  $\mathbb{V}_{\Omega,h}^{k,q}$  is  $(k + 1, q)$ -regular, and satisfies the inverse assumption, IA, under additional hypotheses given at the end of Section 3.3. The space of Lagrange multipliers  $S_{h_1}^{k+1,q}(\partial\Omega)$  has the same  $(t, k^*)$ -regularity as the approximating space, and the functions in  $S_{h_1}^{k+1,q}(\partial\Omega)$  are defined only on  $\partial\Omega$  with respect to particles on  $\partial\Omega$ . Thus,  $\partial\Omega$  must contain enough particles. We note that the functions in  $S_{h_1}^{k+1,q}(\partial\Omega)$  are not restrictions of functions in

$\mathbb{V}_{\Omega,h}^{k,q}$  on  $\partial\Omega$ . Then, following the analysis in Babuška (1973a) and Babuška and Aziz (1972), one can show that, if the size of the supports of the basis elements of  $S_{h_1}^{k+1,q}(\partial\Omega)$  is of the same order as  $|E_{\underline{x}}^h|$ ,  $\underline{x} \in A_{\partial\Omega}^h$  ( $E_{\underline{x}}^h$  and  $A_{\partial\Omega}^h$  defined in (7.3) and (7.4), respectively), then the approximate solution obtained from the Lagrange multiplier method converges. If the size of the supports of the basis elements of  $S_{h_1}^{k+1,q}(\partial\Omega)$  is smaller than  $\eta_{\underline{x}}^h$ ,  $\underline{x} \in A_{\partial\Omega}^h$ , then the method is unstable. This relationship was further analysed in Pitkäranta (1979, 1980).

The Lagrange multiplier technique leads to the optimal rate of convergence in comparison to the penalty method, where the optimal rate of convergence is usually not attained. But, as we mentioned before, the sufficient conditions for convergence are quite delicate.

Recently, the Lagrange multiplier technique was applied in the context of meshless methods without any theoretical analysis, in Belytschko, Lu and Gu (1994), Liu (2002), Lu, Belytschko and Gu (1994) and Mukherjee and Mukherjee (1997).

#### *The Nitsche and related methods*

Because of the delicate nature of Lagrange multiplier methods, there has been some interest in looking for other methods to deal with the issue of the imposition of Dirichlet boundary conditions, and to avoid complications that are present in Lagrange multiplier methods. To that end, certain methods were proposed in Barbara and Hughes (1991) and Stenberg (1995). But a similar method was proposed much earlier by Nitsche (1970–1971). We will discuss the Nitsche method, following the presentation in Stenberg (1995). We will still assume that  $\partial\Omega$  is smooth.

To approximate the solution  $u_0$  of (7.8) by the Nitsche method, we consider the particle space  $\mathbb{V}_{\Omega,h}^{k,q}$  with  $q \geq 2$ . We also assume that:

- $\text{card}(A_{\partial\Omega}^h) \leq \kappa$  and

$$C_1 h \leq h_{E_{\underline{x}}^h} \leq C_2 h, \quad \underline{x} \in A_{\partial\Omega}^h, \quad (7.28)$$

where  $h_{E_{\underline{x}}^h} \equiv |E_{\underline{x}}^h|$ , for  $\underline{x} \in A_{\partial\Omega}^h$  ( $E_{\underline{x}}^h$  and  $A_{\partial\Omega}^h$  are defined in (7.3) and (7.4), respectively);

- there exists  $0 < \mathcal{K} < \infty$ ,  $\mathcal{K} = \mathcal{K}(X^h)$ , such that

$$\left\| \frac{\partial v}{\partial n} \right\|_{-\frac{1}{2},h} \leq \mathcal{K} [B(v,v)]^{1/2}, \quad \text{for all } v \in \mathbb{V}_{\Omega,h}^{k,q}, \quad (7.29)$$

where  $B(u,v)$  was defined in (7.10) and

$$\left\| \frac{\partial v}{\partial n} \right\|_{-\frac{1}{2},h}^2 = \sum_{\underline{x} \in A_{\partial\Omega}^h} h_{E_{\underline{x}}^h} \left\| \frac{\partial v}{\partial n} \right\|_{H^0(E_{\underline{x}}^h)}^2. \quad (7.30)$$

We define the bilinear form  $\tilde{B}(u, v) = B_\gamma(u, v)$ , where

$$B_\gamma(u, v) = B(u, v) - D\left(\frac{\partial u}{\partial n}, v\right) - D\left(\frac{\partial v}{\partial n}, u\right) + \gamma \sum_{\underline{x} \in A_{\partial\Omega}^h} h_{E_{\underline{x}}^h}^{-1} \int_{E_{\underline{x}}^h} uv \, ds,$$

with  $\gamma > 0$ ;  $B(u, v)$ ,  $D(u, v)$  are as defined in (7.10), (7.11), respectively. The approximate solution  $u_{h,\gamma} \in \mathbb{V}_{\Omega,h}^{k,q}$ , obtained from the Nitsche method, is given by

$$B_\gamma(u_{h,\gamma}, v) = \int_{\Omega} f v \, dx, \quad \text{for all } v \in \mathbb{V}_{\Omega,h}^{k,q}. \tag{7.31}$$

For  $u \in H^2(\Omega)$ , we define the norm

$$\|u\|^2 = B(u, u) + \left\| \frac{\partial u}{\partial n} \right\|_{-\frac{1}{2},h}^2 + \|u\|_{\frac{1}{2},h}^2,$$

where

$$\|u\|_{\frac{1}{2},h}^2 = \sum_{\underline{x} \in A_{\partial\Omega}^h} h_{E_{\underline{x}}^h}^{-1} \|u\|_{H^0(E_{\underline{x}}^h)}^2,$$

and  $\left\| \frac{\partial u}{\partial n} \right\|_{-\frac{1}{2},h}^2$  is given by (7.30) with  $v$  replaced by  $u$ . We first note that, from the Schwartz inequality, we have

$$\begin{aligned} \sum_{\underline{x} \in A_{\partial\Omega}^h} h_{E_{\underline{x}}^h}^{-1} \int_{E_{\underline{x}}^h} uv \, ds &\leq \sum_{\underline{x} \in A_{\partial\Omega}^h} \left( \int_{E_{\underline{x}}^h} h_{E_{\underline{x}}^h}^{-1} u^2 \, ds \right)^{1/2} \left( \int_{E_{\underline{x}}^h} h_{E_{\underline{x}}^h}^{-1} v^2 \, ds \right)^{1/2} \\ &\leq \left[ \sum_{\underline{x} \in A_{\partial\Omega}^h} h_{E_{\underline{x}}^h}^{-1} \|u\|_{H^0(E_{\underline{x}}^h)}^2 \right]^{1/2} \left[ \sum_{\underline{x} \in A_{\partial\Omega}^h} h_{E_{\underline{x}}^h}^{-1} \|v\|_{H^0(E_{\underline{x}}^h)}^2 \right]^{1/2} \\ &= \|u\|_{\frac{1}{2},h} \|v\|_{\frac{1}{2},h}. \end{aligned} \tag{7.32}$$

Also,

$$\begin{aligned} D\left(u, \frac{\partial v}{\partial n}\right) &\leq \sum_{\underline{x} \in A_{\partial\Omega}^h} \int_{E_{\underline{x}}^h} \left| h_{E_{\underline{x}}^h}^{-1/2} u h_{E_{\underline{x}}^h}^{1/2} \frac{\partial v}{\partial n} \right| \, ds \\ &\leq \sum_{\underline{x} \in A_{\partial\Omega}^h} h_{E_{\underline{x}}^h}^{-1/2} \|u\|_{H^0(E_{\underline{x}}^h)} h_{E_{\underline{x}}^h}^{1/2} \left\| \frac{\partial v}{\partial n} \right\|_{H^0(E_{\underline{x}}^h)} \\ &\leq \|u\|_{\frac{1}{2},h} \left\| \frac{\partial v}{\partial n} \right\|_{-\frac{1}{2},h}. \end{aligned} \tag{7.33}$$

Using the same arguments used to obtain (7.33), we get

$$D\left(\frac{\partial u}{\partial n}, v\right) \leq \left\| \frac{\partial u}{\partial n} \right\|_{-\frac{1}{2}, h} \|v\|_{\frac{1}{2}, h}. \quad (7.34)$$

Now, using (7.32)–(7.34), it can easily be shown that

$$B_\gamma(u, v) \leq (1 + \gamma) \|u\| \|v\|. \quad (7.35)$$

We now show that, for a proper value of  $\gamma$ , the bilinear form  $B_\gamma(u, v)$  is coercive.

**Lemma 7.4.** Suppose  $\mathcal{K}^2 < \gamma$ , where  $\mathcal{K}$  satisfies (7.29). Then,

$$B_\gamma(v, v) \geq C_1^* \|v\|^2, \quad \text{for all } v \in \mathbb{V}_{\Omega, h}^{k, q}, \quad (7.36)$$

where  $C^* = C^*(X^h) > 0$ .

*Proof.* Let  $v \in \mathbb{V}_{\Omega, h}^{k, q}$ , and choose any  $\epsilon > 0$ . From the definition of  $B_\gamma(u, v)$  and (7.33) with  $u = v$ , we have

$$\begin{aligned} B_\gamma(v, v) &= B(v, v) - 2D\left(v, \frac{\partial v}{\partial n}\right) + \gamma \|v\|_{\frac{1}{2}, h}^2 \\ &\geq B(v, v) - 2\|v\|_{\frac{1}{2}, h} \left\| \frac{\partial v}{\partial n} \right\|_{-\frac{1}{2}, h} + \gamma \|v\|_{\frac{1}{2}, h}^2 \\ &\geq B(v, v) - \epsilon \|v\|_{\frac{1}{2}, h}^2 - \frac{1}{\epsilon} \left\| \frac{\partial v}{\partial n} \right\|_{-\frac{1}{2}, h}^2 + \gamma \|v\|_{\frac{1}{2}, h}^2 \\ &= B(v, v) - \frac{1}{\epsilon} \left\| \frac{\partial v}{\partial n} \right\|_{-\frac{1}{2}, h}^2 + (\gamma - \epsilon) \|v\|_{\frac{1}{2}, h}^2 \\ &\geq \left(1 - \frac{\mathcal{K}^2}{\epsilon}\right) B(v, v) + (\gamma - \epsilon) \|v\|_{\frac{1}{2}, h}^2. \end{aligned}$$

Therefore, considering  $\epsilon = \frac{1}{2}(\mathcal{K}^2 + \gamma)$  in the above inequality, we get

$$B_\gamma(v, v) \geq C_1 [B(v, v) + \|v\|_{\frac{1}{2}, h}^2], \quad (7.37)$$

where  $C_1 = \min\left(\frac{\gamma - \mathcal{K}^2}{\gamma + \mathcal{K}^2}, \frac{\gamma - \mathcal{K}^2}{2}\right)$ . Now, from the definition of  $\|\cdot\|$  and using (7.29), we get

$$\begin{aligned} \|v\|^2 &\leq B(v, v) + \mathcal{K}^2 B(v, v) + \|v\|_{\frac{1}{2}, h}^2 \\ &\leq (1 + \mathcal{K}^2) [B(v, v) + \|v\|_{\frac{1}{2}, h}^2]. \end{aligned}$$

Thus, combining the above inequality with (7.37), we get

$$B_\gamma(v, v) \geq C^* \|v\|^2,$$

where  $C^* = \frac{1}{1+\mathcal{K}^2} \min(\frac{\gamma-\mathcal{K}^2}{\gamma+\mathcal{K}^2}, \frac{\gamma-\mathcal{K}^2}{2})$ , which is (7.36). □

We now present the following result.

**Theorem 7.5.** Suppose  $u_0 \in H^l(\Omega)$ , for  $l \geq 2$ , is the solution of (7.8). Let  $u_{h,\gamma} \in \mathbb{V}_{\Omega,h}^{k,q}$ , with  $q \geq 2$ , be the solution of (7.31), where  $\mathbb{V}_{\Omega,h}^{k,q}$  satisfies (7.28) and (7.29). Then

$$\|u_0 - u_{h,\gamma}\|_{H^1(\Omega)} \leq \frac{C(1 + \gamma)}{C^*(X^\nu)} h^\mu \|u_0\|_{H^l(\Omega)}, \quad \mu = \min(k, l - 1), \quad (7.38)$$

where  $C^*(X^\nu)$  is as in (7.36).

*Proof.* It is easy to see that

$$B_\gamma(u_0, v) = \int_\Omega f v \, dx, \quad \text{for all } v \in H^1(\Omega),$$

and therefore,

$$B_\gamma(u_0 - u_{h,\gamma}, v) = 0, \quad \text{for all } v \in \mathbb{V}_{\Omega,h}^{k,q}. \quad (7.39)$$

Now, for any  $g_h \in \mathbb{V}_{\Omega,h}^{k,q}$ , using (7.36), (7.39) and (7.35), we have

$$\begin{aligned} \|g_h - u_{h,\gamma}\|^2 &\leq \frac{1}{C^*} \hat{B}_\gamma(g_h - \hat{u}_{h,\gamma}, g_h - u_{h,\gamma}) \\ &\leq \frac{1}{C^*} \hat{B}_\gamma(g_h - u_0, g_h - u_{h,\gamma}) \\ &\leq \frac{(1 + \gamma)}{C^*} \|u_0 - g_h\| \|g_h - u_{h,\gamma}\|, \end{aligned}$$

and hence

$$\|g_h - u_{h,\gamma}\| \leq \frac{(1 + \gamma)}{C^*} \|u_0 - g_h\|.$$

Therefore,

$$\begin{aligned} \|u_0 - u_{h,\gamma}\| &\leq \|u_0 - g_h\| + \|g_h - u_{h,\gamma}\| \\ &\leq C \|u_0 - g_h\|, \quad \text{for all } g \in \mathbb{V}_{\Omega,h}^{k,q}. \end{aligned} \quad (7.40)$$

Now, using (7.28) and a trace inequality, we have

$$\begin{aligned} \|u_0 - g_h\|_{\frac{1}{2},h}^2 &= \sum_{\underline{x} \in A_{\partial\Omega}^h} h_{E_{\underline{x}}^h}^{-1} \|u_0 - g_h\|_{H^0(E_{\underline{x}}^h)}^2 \\ &\leq Ch^{-1} \|u_0 - g_h\|_{H^0(\partial\Omega)}^2 \end{aligned}$$

$$\begin{aligned} &\leq Ch^{-1} \left( \frac{1}{h} \|u_0 - g_h\|_{H^0(\Omega)}^2 + h \|u_0 - g_h\|_{H^1(\Omega)}^2 \right) \\ &= C(h^{-2} \|u_0 - g_h\|_{H^0(\Omega)}^2 + \|u_0 - g_h\|_{H^1(\Omega)}^2), \end{aligned} \quad (7.41)$$

where  $C$  depends on  $\kappa$ . Also, using a similar argument, we have

$$\left\| \frac{\partial(u_0 - g_h)}{\partial n} \right\|_{-\frac{1}{2},h}^2 \leq C(h^2 \|u_0 - g_h\|_{H^2(\Omega)}^2 + \|u_0 - g_h\|_{H^1(\Omega)}^2). \quad (7.42)$$

where  $C$  depends on  $\kappa$ . Thus from (7.41) and (7.42) we get

$$\begin{aligned} \|u_0 - g_h\|^2 &= \|u_0 - g_h\|_{H^1(\Omega)}^2 + \left\| \frac{\partial(u_0 - g_h)}{\partial n} \right\|_{-\frac{1}{2},h}^2 + \|u_0 - g_h\|_{\frac{1}{2},h}^2 \\ &\leq Ch^{-2} \sum_{j=0}^2 h^{2j} \|u_0 - g_h\|_{H^j(\Omega)}^2, \end{aligned}$$

and hence, from the definition of  $\|\cdot\|$  and (7.40), we have

$$\begin{aligned} \|u_0 - \hat{u}_{h,\gamma}\|_{H^1(\Omega)} &\leq \|u_0 - \hat{u}_{h,\gamma}\| \\ &\leq Ch^{-1} \sum_{j=0}^2 h^j \|u_0 - g_h\|_{H^j(\Omega)}, \quad \text{for all } g_h \in \mathbb{V}_{\Omega,h}^{k,q}. \end{aligned} \quad (7.43)$$

Finally, we choose  $g \in \mathbb{V}_{\Omega,h}^{k,q}$  such that

$$\|u_0 - g\|_{H^s(\Omega)} \leq Ch^{\mu_1 - s} \|u_0\|_{H^l(\Omega)}, \quad 0 \leq s \leq 2,$$

where  $\mu_1 = \min(k+1, l)$ . Using this in (7.43) we get the desired result.  $\square$

We now discuss situations where the assumption required to prove Theorem 7.5 is valid. The major problem is to estimate  $\mathcal{K}(X^h)$  given in (7.29). We would like to have  $\mathcal{K}(X^h) \leq C$ , uniformly for all  $0 < h \leq 1$ . If the supports  $\eta_{\underline{x}}^h$  of the particle function  $\phi_{\underline{x}}^h$  are ‘reasonable’, *e.g.*, in  $\mathbb{R}^2$  or in  $\mathbb{R}^3$ , then it is easy to see that the necessary condition for  $\mathcal{K}(X^h) \leq C$  is that  $h_{E_{\underline{x}}^h} \geq \alpha |\eta_{\underline{x}}^h|$  for  $\underline{x} \in A_{\partial\Omega}^h$ . This can be enforced by properly selecting the set of particles  $X^h$ . This aspect can also affect the design of adaptive meshless (Nitsche) methods. Since the estimates of these constants are difficult to estimate accurately, we may select larger values of  $\gamma$  in (7.31) so that Theorem 7.5 is valid.

The Nitsche method presented here is superior to both the penalty method and Lagrange multiplier methods. The Nitsche method, in the framework of meshless methods, was addressed in Babuška *et al.* (2002*a*) and implemented in Schweitzer (200*x*) and Griebel and Schweitzer (2002*e*).

*The collocation method*

The collocation method, in the framework of meshless methods, was recently proposed in Atluri and Shen (2002), Zhu and Atluri (1998) and Hagen (1996). The method consists of adding constraint equation, at certain points of the boundary  $\partial\Omega$ , to the stiffness matrix. No analysis was presented to address the convergence of the approximate solution obtained from this method. Collocation using radial basis functions was analysed in Franke and Schaback (1998).

*Combination of meshless methods and the finite element method*

This method was proposed, *e.g.*, in Krongauz and Belytschko (1996). The main idea in this approach is to use classical finite elements (which could also be interpreted as particle functions) in a neighbourhood of the boundary  $\partial\Omega$ , and to select other particle functions such that their supports do not intersect  $\partial\Omega$ .

*The characteristic function method*

The method was proposed in connection to the Ritz method when the approximating functions were global polynomials (see Mikhlin (1971) and Kantorovich and Krylov (1958)). If a domain  $\Omega$  has a smooth boundary  $\partial\Omega$ , there exists a smooth function  $\Phi$  such that

$$\begin{aligned} \Phi(x) &> 0, & x \in \Omega, \\ \Phi(x) &= 0, & x \in \partial\Omega, \\ \text{and } |\nabla\Phi(x)| &\geq \alpha > 0, & x \in \partial\Omega. \end{aligned}$$

Let  $S_h^\Phi = \{u : u = \Phi v, v \in \mathbb{V}_{\Omega,h}^{k,q}\}$ . Then it is obvious that  $S_h^\Phi \subset H_0^1(\Omega)$ . We approximate the solution  $u_0$  of (2.1) and (2.3) by  $u_h \in S_h^\Phi$ , where  $u_h$  is the solution  $u_S$  of (2.7) with  $S = S_h^\Phi$ .

For  $u_0 \in H^l(\Omega) \cap H_0^1(\Omega)$ ,  $l \geq 2$ , we define  $w_0 = \frac{u_0}{\Phi}$ . Then, using Hardy's inequality (Theorem 329 of Hardy, Littlewood and Polya (1952)), one can show that  $w_0 \in H^{l-1}(\Omega)$ . Using this result, we obtain the following theorem.

**Theorem 7.6.** Suppose  $u_0 \in H^l(\Omega) \cap H_0^1(\Omega)$ , and suppose  $l \geq 2$ . Then there exists  $w_h \in \mathbb{V}_{\Omega,h}^{k,q}$  such that  $g_h = \Phi w_h$  satisfies

$$\|u_0 - g_h\|_{H^1(\Omega)} \leq Ch^\mu \|u_0\|_{H^l(\Omega)}, \quad \mu = \min(k, l - 2). \tag{7.44}$$

*Proof.* Recall that  $\mathbb{V}_{\Omega,h}^{k,q}$  is  $(k + 1, q)$ -regular with  $q \geq 1$ . Then there exists  $w_h \in \mathbb{V}_{\Omega,h}^{k,q}$  such that

$$\|w_0 - w_h\|_{H^1(\Omega)} \leq Ch^\mu \|w_0\|_{H^{l-1}(\Omega)} \leq Ch^\mu \|u_0\|_{H^l(\Omega)}, \tag{7.45}$$

where  $\mu = \min(k, l - 2)$ . Now, from the definition of  $w_0$ , we have

$$u_0 - \Phi w_h = u_0 - \Phi w_0 + \Phi(w_0 - w_h) = \Phi(w_0 - w_h),$$

and hence, using (7.45), we have

$$\|u_0 - \Phi w_h\|_{H^1(\Omega)} \leq C \|w_0 - w_h\|_{H^1(\Omega)} \leq Ch^\mu \|u_0\|_{H^l(\Omega)}, \quad \mu = \min(k, l - 2),$$

which is the desired result.  $\square$

**Remark 54.** It is clear from (2.8) and (7.44) that, for  $l \geq 2$ ,

$$\|u_0 - u_h\|_{H^1(\Omega)} \leq Ch^\mu \|u_0\|_{H^l(\Omega)}, \quad \mu = \min(k, l - 2).$$

We further note that this order of convergence cannot, in general, be improved.

### *The generalized finite element method*

We note that all the methods described so far primarily use  $\mathbb{V}_{\Omega,h}^{k,q}$  as the approximating space. The GFEM, on the other hand, uses different approximating spaces, as we have seen in Section 6. The use of these approximating spaces makes the GFEM extremely flexible.

We recall that in the GFEM we start with a partition of unity with respect to a simple mesh that need not conform to the boundary of the domain. This partition unity could be the particle shape functions defined in Section 3. Then ‘handbook’ functions are used as local approximating spaces. The Dirichlet boundary condition can be implemented by choosing the local approximation space  $V_{\underline{x}}$ , for  $\underline{x} \in A_{\partial\Omega}^h$ , such that the functions in  $V_{\underline{x}}$  satisfy the Dirichlet boundary conditions.

We have presented a few approaches on how to use meshless approximation to approximate solutions of PDEs. To impose Dirichlet boundary conditions on meshless approximation is a challenge, and we looked into some methods that can overcome this difficulty. While discussing these methods, we assumed that the boundary of the domain is smooth, for simplicity. But the results presented here can be generalized to cover nonsmooth boundaries, especially piecewise smooth boundaries.

Some methods were implemented and reported in the literature, but lacked rigorous theoretical analysis. All the methods reported here have certain advantages as well as disadvantages. If the particle space  $\mathbb{V}_{\Omega,h}^{k,q}$  is used as an approximating space, then in our opinion the Nitsche method is very promising, because it is relatively robust and easy to implement. But we note that  $\mathbb{V}_{\Omega,h}^{k,q}$  is difficult to construct for higher values of  $k$ , and the use of  $\mathbb{V}_{\Omega,h}^{k,q}$  with lower values of  $k$  reduces the accuracy of the method. On the other hand, the GFEM uses a partition of unity (the basis functions of  $\mathbb{V}_{\Omega,h}^{k,q}$  with  $k = 0$ ), which is easy to construct, and higher accuracy can be attained by using suitable local approximation spaces.



## 8. Implementational aspects of meshless methods

In this section, we will briefly discuss the implementational aspects of meshless methods and the GFEM. As for the finite element method, the implementation of meshless methods and the GFEM has four major parts:

- (1) construction of particle shape functions,
- (2) construction of the stiffness matrix,
- (3) solution of the linear system of equations,
- (4) *a posteriori* error estimation, adaptivity, and computation of data of interest.

We now discuss these items in turn.

### *Construction of particle shape functions*

In the classical finite element method, we start with a mesh that is related to the domain, and then shape functions are defined with respect to the chosen mesh. In a meshless method, we start with particles  $\{\underline{x}\}$ . Corresponding to each particle  $\underline{x}$ , a particle shape function with compact support  $\eta_{\underline{x}}$  is constructed, such that the  $\{\eta_{\underline{x}}\}$  form an open cover of the domain  $\Omega$ . The construction of shape functions that are reproducing of order  $k = 0$  or  $1$  is not difficult. For  $k = 0$ , one may use Shepard's approach (Shepard 1968) as described in Section 4.1. For  $k = 1$  and for an appropriate particle distribution, one may first construct a mesh using tetrahedra such that the particles are the nodes of the mesh (*i.e.*, the vertices of the tetrahedra). This procedure is not difficult, as there are efficient codes available for constructing such a mesh. The shape function corresponding to the particle  $\underline{x}$  can be taken to be the standard hat functions, whose support is the union of all the simplices with  $\underline{x}$  as one of its vertices. We note, however, that, for  $k = 1$ , smoother shape functions can also be constructed (see Han and Meng (2001)). For  $k = 0, 1$ , we have to check that  $\text{card}(S_{\underline{x}}) \leq \kappa$ ,  $\kappa$  is independent of  $\underline{x}$ , where  $S_{\underline{x}} = \{\underline{y} : \eta_{\underline{y}} \cap \eta_{\underline{x}} \neq \emptyset\}$ . For the Nitsche method, described in Section 7.2, we also have to check that  $\mathcal{K}$ , defined in (7.29), is bounded. The construction of particle shape functions for  $k \geq 2$  is more expensive than for  $k = 0, 1$ , and it may be more difficult to check assumptions A1–A7 and (3.63), which ensure convergence.

In contrast, the GFEM uses only a partition of unity, and thus particle shape functions with  $k = 0, 1$ , described in the last paragraph, can be used for this purpose. Also, a simple regular distribution of particles could be used to construct the partition of unity. The space of local shape functions,  $V_{\underline{x}}$ , could be created analytically or through 'handbook' solutions. Dirichlet boundary conditions are also treated by appropriate selection of  $V_{\underline{x}}$ , and hence we do not have to use special methods, *e.g.*, the penalty method, the Nitsche method, *etc.*, which simplifies the implementation.

### *Construction of the stiffness matrix*

The construction of the stiffness matrix for a meshless method is laborious and delicate. In fact, this is where we pay the price for avoiding mesh generation. The elements of the resulting stiffness matrix are integrals, which have to be numerically evaluated over various regions. These regions are not simple tetrahedra as in the finite element method, where they naturally come from a mesh. These regions, for a meshless method, are of the form  $\eta_{\underline{x}} \cap \eta_{\underline{y}} \cap \Omega$ ,  $\underline{x}, \underline{y} \in X^\nu$ , and can be extremely complicated. Also, the integrals have to be evaluated accurately, as it is known that inaccurate numerical integration leads to very poor results (see, for example, Chen, Wu, Yoon and You (2001)). A special numerical integration scheme is given in De and Bathe (2001), where the  $\{\eta_{\underline{x}}\}$  are balls and the region of integration is the intersection of two balls. The problem of effective integration has also been addressed in Dolbow and Belytschko (1999), Griebel and Schweitzer (2002b), Schweitzer (200x) and Strouboulis *et al.* (2001a, 2001b). The numerical integration poses additional problems in the GFEM when singular functions are included in the local approximating space  $V_{\underline{x}}$ . Standard integration schemes in this situation yield poor accuracy. This problem in the GFEM was handled in Strouboulis *et al.* (2001a) by using adaptive numerical integration. Because of this sensitivity to numerical integration, the use of adaptive integration is preferred in GFEMs.

Thus we see that an accurate and effective numerical integration scheme to approximate the elements of the stiffness matrix is essential for the success of meshless methods. We will remark further on this issue in the next subsection. We note that numerical integration and construction of stiffness matrices in these methods are parallelizable.

### *Solution of the linear system*

The exact stiffness matrix (without numerical integration) obtained from a meshless method could be positive definite with a large condition number. This is caused by using shape functions with large overlap between their supports, which makes the shape functions ‘almost’ linearly dependent. Moreover, the exact stiffness matrix obtained from the GFEM could be positive semi-definite, as shown in Strouboulis *et al.* (2001a). But the underlying linear system obtained from the GFEM is always consistent, *i.e.*, the linear system has non-unique solutions. The lack of unique solvability of the linear system does not imply that the GFEM produces non-unique solutions. In fact, if the vector  $\{c_{\underline{x},j}\}_{1 \leq j \leq n_{\underline{x}}}$ , with  $\dim V_{\underline{x}} = n_{\underline{x}}$ , is a solution of this linear system, then the solution

$$u_h = \sum_{\underline{x}} \sum_{j=1}^{n_{\underline{x}}} \phi_{\underline{x}} c_{\underline{x},j} \psi_{\underline{x}}^j,$$

obtained from the GFEM, where  $\psi_{\underline{x}}^j$  is a basis of  $V_{\underline{x}}$ , is unique.

We have already mentioned the importance of numerical integration in evaluating the elements of the stiffness matrix obtained from a meshless method. We further note that the elements of the load vector is also evaluated by numerical integration. To obtain a consistent linear system (after the use of numerical integration), the numerical integration scheme applied to compute an element of the load vector should be same as the scheme used to compute the corresponding row of the stiffness matrix.

To find the solution of the linear system obtained from a meshless method (or from the GFEM), one can use a special direct solver based on elimination or an iterative solver. Strouboulis *et al.* (2001a) used direct solvers, *e.g.*, subroutines MA27 and MA47 of the Harwell Subroutine Library, to solve the sparse positive semi-definite linear system obtained from the GFEM. The use of these solvers was successful even when the nullity of the stiffness matrix was large. It was also shown in Strouboulis *et al.* (2001a) that round-off errors did not play a significant role in solving the linear system, *i.e.*, the round-off error was almost the same as when the standard finite element linear system is solved by the elimination method.

An iterative algorithm for solving such linear systems was given in Strouboulis *et al.* (2001a). The idea of this algorithm, which has been used in many situations, is to perturb the stiffness matrix by adding a small multiple of the identity matrix. The perturbed matrix, say  $P$ , is positive definite and any solver could be used to solve  $Px = b$ . Using this fact and a few iterations of a simple iterative technique, a solution of the original linear system could be obtained. We refer to Strouboulis *et al.* (2001a) for a complete description of the effectiveness of this iterative algorithm.

We have noted before that the linear system obtained from the meshless method is consistent even if the stiffness matrix is positive semi-definite. In this situation, a solver based on conjugate gradient methods can also be used. The convergence in this situation is similar to the convergence of conjugate gradient methods when applied to solve the linear system obtained from the standard finite element method. The multigrid method is not directly applicable to the linear system when the stiffness matrix is positive semi-definite, since the eigenfunction corresponding to the zero eigenvalue of the stiffness matrix is global and oscillatory. The same is also true when the stiffness matrix is ‘almost’ singular. However, a special version of the multigrid method was proposed in Schweitzer (200x) and Griebel and Schweitzer (2002c), when the underlying partition of unity is reproducing of order  $k = 0$ . For another multigrid method, see Xu and Zikatanov (2002).

#### *A posteriori error estimation, adaptivity and programming*

The rigorous theory of *a posteriori* error estimation originated in Babuška and Melenk (1997) and other estimates, based on various averaging, were also used. These estimators can be used as error indicators for adaptivity

purposes. For adaptive approaches in meshless methods, we refer readers to Schweitzer (200x) and Belytschko, Liu and Singer (1998b).

We finally mention that programming the meshless method is an important issue and it requires specific concepts. For this aspect of the meshless method, we refer to Dolbow and Belytschko (1998), Griebel and Schweitzer (2002d), and Strouboulis *et al.* (2001a, 2001b).

## 9. Examples

Meshless methods have been applied to linear and nonlinear elliptic problems, as well as to problems related to other differential equations: we refer to Li and Liu (2002). However, it is essential to characterize the types of problems where this method is, or could be, superior to standard methods (Belytschko, Gerlach, Krongauz, Krysl and Dolbow 1998a).

In this paper, we address only the application of the meshless method on a class of linear, elliptic problems. As stated in the Introduction, one of the main advantages of meshless methods is that it avoids mesh generation. This is essential when the domain is complex. Another advantage of this method is that it allows the use of various ‘special’ local shape functions to improve the accuracy.

The generalized finite element method (GFEM) was discussed in detail by Strouboulis *et al.* (2001b), and it was shown that the method is effective. Three types of meshes with successive refinements were used in that paper, and we present one of these meshes in Figure 9.1. This is a simple finite element mesh and it *does not* reflect the geometry of the underlying domain. Then, using the linear finite elements as a partition of unity, and special functions for local approximation, an improvement in the rate of convergence was achieved. Detailed numerical data, with comments on various aspects of the method, for instance numerical integration, *etc.*, were presented in

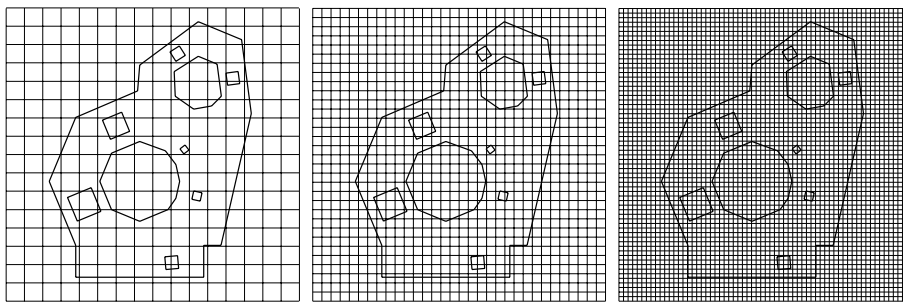


Figure 9.1. A mesh used in Strouboulis *et al.* (2001b) for the construction of a partition of unity in the context of the GFEM to solve a problem posed on the domain.

Strouboulis *et al.* (2001*b*). We note, however, that although the domain considered in this example (*i.e.*, the domain in Figure 9.1) was simple, and classical finite element methods with mesh refinement or an adaptive procedure could have been used, the analysis and data presented in Strouboulis *et al.* (2001*b*) clearly show the scope and potential of the GFEM.

As mentioned above, the power of the GFEM lies in handling problems where the underlying domain has complex geometry. Three types of complex domain, shown in Figure 9.2, were analysed by the GFEM in Strouboulis *et al.* (2001*b*). Another complex domain with fibres, analysed in the same paper, is shown in Figure 9.3, where the fibre distribution was taken from Babuška, Andersson, Smith and Levin (1999). To construct finite element meshes for these domains is very complex and nearly impossible. ‘Handbook’ problems that characterize the local behaviour of the approximated solution (*e.g.*, in the neighbourhood of a crack, fibres, *etc.*) were used to construct special shape functions for these problems.

The GFEM has an advantage in dealing with problems with singularities (in the neighbourhood of geometric edges) in three dimensions. When the basic finite element tetrahedral mesh is used in such problems, it is well known that classical edge refinement is cumbersome. This problem was handled using the GFEM in Duarte, Babuška and Oden (2000), where a refinement by special functions, at positions indicated by an error indicator, was performed. The GFEM was also used to handle difficulties stemming from orthotropic problems in Duarte and Babuška (2002).

There are other types of problem where the GFEM is quite effective. They include multisite problems, where many crack configurations are present, and crack propagation problems, where the geometry of the domain changes. The GFEM could be used in such problems by considering local approximating spaces consisting of functions that are discontinuous over the cracks,

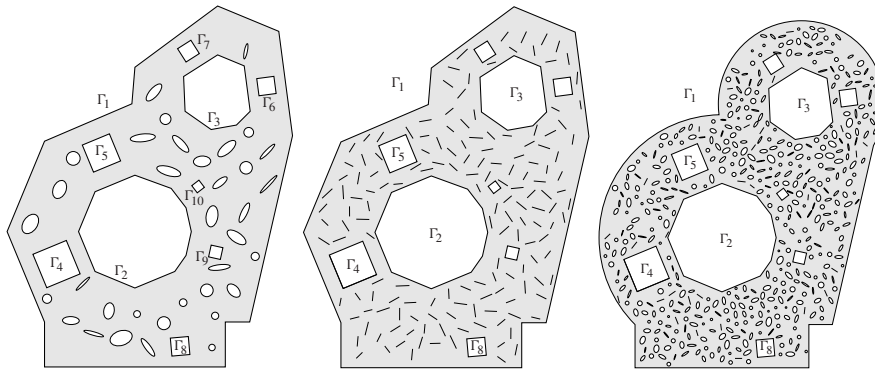


Figure 9.2. The three types of domain analysed by the GFEM in Strouboulis *et al.* (2001*b*).

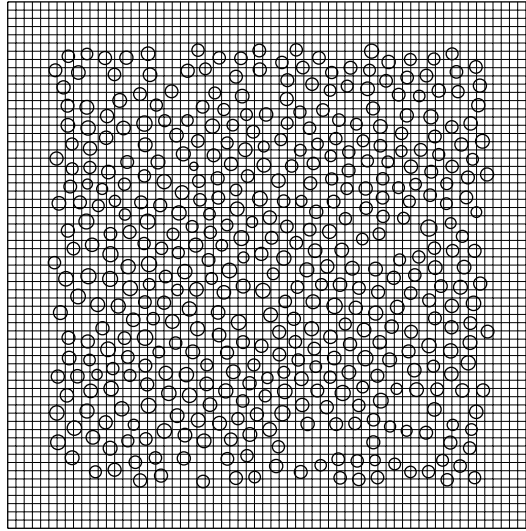


Figure 9.3. The problem of fibre composite type analysed in Strouboulis *et al.* (2001*b*).

and including a singular function with respect to the tip of the crack into the local approximating space. Then the propagation of the crack is computed (using stress intensity factors); the old singular function in the approximating space is replaced by a new singular function, new discontinuous functions are added in the same space, and the process of computing the propagation of crack and changing the local approximation spaces is repeated. In this process, the matrix of the underlying linear system at a particular step can be obtained by augmenting the matrix corresponding to the previous step with new rows and columns, and it is possible to solve the new linear system using the Schur complement, which uses the previously computed data. This general idea was used in Li and Liu (2002) and Moes, Gravouil and Belytschko (2002) without using the previously computed data.

The GFEM is an important tool in approximating solutions of elliptic problems with rough coefficients as well as homogenization problems. We mention that the usual finite element method may give extremely poor results when applied to elliptic problems with rough coefficients, as shown in Babuška and Osborn (2000). It was shown in Babuška *et al.* (1994), using a detailed analysis, that the GFEM leads to the same rate of convergence for problems with rough coefficients as when the coefficients are smooth. The GFEM is also related to upscaling (Arbogast 2000) and stabilization (Hughes 1995).

We emphasize that in this paper we have considered only a small (but important) family of problems. We have shown that the use of meshless

methods, particularly the use of the GFEM, on these problems is advantageous in comparison to the standard finite element method. Of course, there are other types of problems, especially nonlinear problems, which we have not addressed in this paper.

## 10. Future challenges

We have addressed some mathematical problems concerning meshless methods. This method is, in fact, a family of approaches that sometimes differ in implementational details, and are referred to by a variety of names. Particular forms of meshless methods are used for solving important engineering problems in solid and fluid mechanics – linear and nonlinear, stationary and time-dependent problems, with fixed and changing boundary conditions, for differential and integral equations of various types. There is a rapidly growing literature on the subject, which is mainly directed towards engineering and focuses on particular problems, whose mathematical and theoretical aspects are not completely understood. For a description of the state of the art we refer to Atluri and Shen (2002), Babuška *et al.* (2002*a*), Griebel and Schweitzer, eds (2002*a*), Li and Liu (2002), G. R. Liu (2002) and Zhang, Liu, Li, Qian and Hao (2002).

We presented basic approximation results in  $L_2$ -based spaces, and discussed their use in the approximate solution of linear elliptic boundary value problems. The approximation theory we developed is applicable to virtually all variationally formulated problems, provided that the stability of the variational method is guaranteed (the inf-sup condition, sometimes called the BB condition, is satisfied). In contrast to coercive problems, proving stability for non-coercive problems is a delicate issue.

Meshless methods, in particular the GFEM, permit the use of non-polynomial and non-smooth shape functions with a certain special character. This feature was successfully utilized in solving partial differential equations with rough coefficients and, more generally, in problems with microstructures. The GFEM can be directly related to upscaling and to stabilization. Special shape functions related to the oscillatory behaviour of the solutions, as in the case of the Helmholtz equation, have been successfully used. The GFEM was used in problems in which the discontinuities propagate, as in linear and nonlinear crack problems. Meshless methods and GFEM can be used for solving higher-order differential equations because it is possible to construct shape functions with high regularity. There are, of course, many remaining issues to be addressed: we mention, in particular, the challenging problems of adaptive selection of shape functions and of proving *a posteriori* error estimates.

Meshless methods and the GFEM are generalizations of the classical FEM, in its  $h$ ,  $p$  and  $hp$  versions. Hence all the theoretical and implementational

problems for the  $h$ ,  $p$  and  $hp$  versions of the FEM have analogues for meshless methods; and there are additional problems, having special features, with meshless methods.

Today, many versions of meshless methods, based on various discretization principles, for example, variational principles, collocation, *etc.*, are used in an *ad hoc* way, with emphasis on constructive results. Thus it is important to create an effective framework in which to study these methods, and to assess their effectiveness.

### Acknowledgements

The authors would like to thank Elsevier Science for allowing the use of Figures 9.1, 9.2 and 9.3 in this paper. They were published in the journal *Computer Methods in Applied Mechanics and Engineering*. The authors would also like to thank Professor Weimin Han for reading and commenting on a preliminary version of this paper.

### REFERENCES

- T. Arbogast (2000), Numerical subgrid upscaling of two-phase flow in porous media, in *Numerical Treatment of Multiphase Flows in Porous Media* (Z. Chen, ed.), Vol. 552 of *Lecture Notes in Physics*, Springer, pp. 35–49.
- M. G. Armentano (2002), ‘Error estimates in Sobolev spaces for moving least square approximation’, *SIAM J. Numer. Anal.* **39**, 38–51.
- M. G. Armentano and R. G. Duran (2001), ‘Error estimates for moving least square approximation’, *Appl. Numer. Math.* **37**, 397–416.
- S. N. Atluri and S. Shen (2002), *The Meshless Local Petrov–Galerkin Method*, Tech Science Press.
- I. Babuška (1970), ‘Approximation by hill functions’, *Comment Math. Univ. Carolinae* **11**, 787–811.
- I. Babuška (1971), The finite element method for elliptic differential equations, in *Numerical Solution of Partial Differential Equations II, SYNSPADE 1970* (B. Hubbard, ed.), Academic Press, London, pp. 69–106.
- I. Babuška (1972), ‘Approximation by hill functions II’, *Comment Math. Univ. Carolinae* **13**, 1–22.
- I. Babuška (1973a), ‘The finite element method with Lagrange multiplier’, *Numer. Math.* **20**, 179–192.
- I. Babuška (1973b), ‘The finite element method with penalty’, *Math. Comp.* **27**, 221–228.
- I. Babuška (1994), Courant element: Before and after, in *Finite Element Methods: Fifty Years of the Courant Element*, Vol. 164 of *Lecture Notes in Pure and Applied Mathematics*, Marcel Dekker, pp. 37–51.
- I. Babuška and A. K. Aziz (1972), Survey lectures on the mathematical foundations of the finite element method, in *Mathematical Foundations of the Finite Element Method with Applications to Partial Differential Equations* (A. K. Aziz, ed.), Academic Press, pp. 3–345.



- I. Babuška and J. M. Melenk (1997), ‘The partition of unity finite element method’, *Internat. J. Numer. Meth. Engr.* **40**, 727–758.
- I. Babuška and J. Osborn (2000), ‘Can a finite element method perform arbitrarily badly?’, *Math. Comp.* **69**, 443–462.
- I. Babuška and T. Strouboulis (2001), *The Finite Element Method and its Reliability*, Clarendon Press, Oxford.
- I. Babuška, B. Andersson, P. J. Smith and K. Levin (1999), ‘Damage analysis of fiber composites, Part 1: Statistical analysis on fiber scale’, *Comput. Meth. Appl. Mech. Engr.* **172**, 27–77.
- I. Babuška, U. Banerjee and J. Osborn (2001), On principles for the selection of shape functions for the generalized finite element method, Technical Report 01-16, TICAM, University of Texas at Austin.
- I. Babuška, U. Banerjee and J. Osborn (2002a), Meshless and generalized finite element methods: A survey of major results, in *Meshfree Methods for Partial Differential Equations* (M. Griebel and M. A. Schweitzer, eds), Springer.
- I. Babuška, U. Banerjee and J. Osborn (2002b), ‘On principles for the selection of shape functions for the generalized finite element method’, *Comput. Methods Appl. Mech. Engr.* **191**, 5595–5629.
- I. Babuška, U. Banerjee and J. Osborn (200x), ‘On the approximability and the selection of particle shape functions’. Submitted to *Numer. Math.*
- I. Babuška, G. Caloz and J. Osborn (1994), ‘Special finite element methods for a class of second order elliptic problems with rough coefficients’, *SIAM J. Numer. Anal.* **31**, 945–981.
- I. Babuška, T. Strouboulis, C. S. Upadhyay and S. K. Gangaraj (1996), ‘Computer-based proof of existence of superconvergence points in the finite element method: Superconvergence of the derivatives in the finite element solution of Laplace’s, Poisson’s, and elasticity equations’, *Numerical Methods for PDEs* **12**, 347–392.
- J. C. Barbara and T. J. R. Hughes (1991), ‘The finite element method with Lagrange multipliers on the boundary circumventing the Babuška–Brezzi condition’, *Comput. Methods Appl. Mech. Engr.* **85**, 109–128.
- T. Belytschko, Y. Krongauz, D. Organ, M. Fleming and P. Krysl (1996), ‘Meshless methods: An overview and recent developments’, *Comput. Methods Appl. Mech. Engr.* **139**, 3–47.
- T. Belytschko, C. Gerlach, Y. Krongauz, P. Krysl and J. Dolbow (1998a), Why meshfree methods, in *Proc. IV World Congress on Computational Mechanics, Buenos Aires, Argentina*.
- T. Belytschko, W. K. Liu and M. Singer (1998b), On adaptivity and error criteria for meshfree methods, in *Advances in Adaptive Computational Methods in Mechanics* (P. Ladeveze and J. T. Oden, eds), pp. 217–228.
- T. Belytschko, Y. Y. Lu and L. Gu (1994), ‘Element-free Galerkin methods’, *Internat. J. Numer. Meth. Engr.* **37**, 229–256.
- J. H. Bramble and S. R. Hilbert (1970), ‘Estimation of linear functionals on Sobolev spaces with application to Fourier transforms and spline interpolation’, *SIAM J. Numer. Anal.* **7**, 113–124.
- J. H. Bramble and S. R. Hilbert (1971), ‘Bounds on a class of linear functionals with applications to Hermite interpolation’, *Numer. Math.* **16**, 362–369.

- M. D. Buhmann (2000), Radial basis functions, in *Acta Numerica*, Vol. 9, Cambridge University Press, pp. 1–38.
- J. S. Chen, C. T. Wu, S. Yoon and Y. You (2001), ‘A stabilized conforming nodal integration for Galerkin meshfree methods’, *Internat. J. Numer. Meth. Engr.* **50**, 435–466.
- P. G. Ciarlet (1980), *Finite Element Methods for Elliptic Problems*, North-Holland.
- C. Daux, N. Moes, J. Dolbow, N. Sukumar and T. Belytschko (2000), ‘Arbitrary branched and intersecting cracks with eXtended finite element method’, *Internat. J. Numer. Meth. Engr.* **48**, 1741–1760.
- S. De and K. J. Bathe (2001), ‘The method of finite spheres with improved numerical integration’, *Computers and Structures* **79**, 2183–2196.
- P. Dierckx (1995), *Curve and Surface Fitting with Splines*, Clarendon Press, Oxford.
- J. Dolbow and T. Belytschko (1998), ‘An introduction to programming the meshless element free Galerkin method’, *Arch. Comput. Methods Engr.* **5**, 207–241.
- J. Dolbow and T. Belytschko (1999), ‘Numerical integration of the Galerkin weak form in meshfree methods’, *Comput. Mech.* **23**, 219–230.
- C. A. M. Duarte (1995), A review of some meshless methods to solve partial differential equations, Technical Report 95-06, TICAM, University of Texas at Austin.
- C. A. M. Duarte and I. Babuška (2002), ‘Mesh independent  $p$ -orthotropic enrichment using generalized finite element method’, *Internat. J. Numer. Meth. Engr.* **55**, 1477–1492.
- C. A. M. Duarte, I. Babuška and J. T. Oden (2000), ‘Generalized finite element method for three dimensional structural problems’, *Computers and Structures* **77**, 219–232.
- C. Franke and R. Schaback (1998), ‘Convergence order estimates of meshless collocation methods using radial basis functions’, *Adv. Comp. Math.* **8**, 381–399.
- R. Franke (1978), Smooth surface approximation by a local method of interpolation at scattered points, Technical Report NPS-53-78-002, Naval Postgraduate School.
- R. Franke (1979), A critical comparison of some methods for interpolation of scattered data, Technical Report NPS-53-79-003, Naval Postgraduate School. Available from NTIS AD-A081 688/4.
- R. A. Gingold and J. J. Monaghan (1977), ‘Smoothed particle hydrodynamics: Theory and application to non-spherical stars’, *Mon. Not. R. Astr. Soc.* **181**, 375–389.
- R. A. Gingold and J. J. Monaghan (1982), ‘Kernel estimates as a basis for general particle methods in hydrodynamics’, *J. Comp. Phys.* **46**, 429–453.
- W. J. Gordon and K. Wixon (1978), ‘Shepard’s method of metric interpolation to bivariate and multivariate data’, *Math. Comp.* **32**, 253–264.
- M. Griebel and M. A. Schweitzer, eds (2002a), *Meshfree Methods for Partial Differential Equations*, Vol. 26 of *Lecture Notes in Computational Science and Engineering*, Springer.
- M. Griebel and M. A. Schweitzer (2002b), ‘A particle-partition of unity method, Part II: Efficient cover construction and reliable integration’, *SIAM J. Sci. Comp.* **23**, 1655–1682.

- M. Griebel and M. A. Schweitzer (2002*c*), 'A particle-partition of unity method, Part III: A multilevel solver', *SIAM J. Sci. Comp.* **24**, 377–409.
- M. Griebel and M. A. Schweitzer (2002*d*), A particle-partition of unity method, Part IV: Parallelization, in *Meshfree Methods for Partial Differential Equations* (M. Griebel and M. A. Schweitzer, eds), Springer.
- M. Griebel and M. A. Schweitzer (2002*e*), A particle-partition of unity method, Part V: Boundary conditions, in *Geometric Analysis and Nonlinear Partial Differential Equations* (S. Hildebrand and H. Karcher, eds), Springer.
- D. Hagen (1996), 'Element-free Galerkin methods in combination with finite element approaches', *Comput. Methods Appl. Mech. Engr.* **139**, 237–262.
- W. Han and X. Meng (2001), 'Error analysis of the reproducing kernel particle method', *Comput. Methods Appl. Mech. Engr.* **190**, 6157–6181.
- G. H. Hardy, J. E. Littlewood and G. Polya (1952), *Inequalities*, Cambridge University Press.
- T. J. R. Hughes (1995), 'Multiscale phenomena, Green's functions, the Dirichlet to Neumann formulation, subgrid scale models, bubbles and the origins of stabilized methods', *Comput. Methods Appl. Mech. Engr.* **127**, 387–401.
- I. V. Kantorovich and V. I. Krylov (1958), *Approximate Methods of Higher Analysis*, Interscience Publishers.
- Y. Krongauz and T. Belytschko (1996), 'Enforcement of essential boundary conditions in meshless approximation using finite elements', *Comput. Methods Appl. Mech. Engr.* **131**, 133–145.
- O. Laghrouche and P. Bettess (2000), Solving short wave problems using special finite elements: Towards an adaptive approach, in *Mathematics of Finite Elements and Applications X* (J. Whiteman, ed.), Elsevier, pp. 181–195.
- P. Lancaster and K. Salkauskas (1981), 'Surfaces generated by moving least squares method', *Math. Comp.* **37**, 141–158.
- P. Lancaster and K. Salkauskas (1986), *Curve and Surface Fitting: An Introduction*, Academic Press, London.
- S. Li and W. K. Liu (1996), 'Moving least squares reproducing kernel particle method, Part II: Fourier analysis', *Comput. Methods Appl. Mech. Engr.* **139**, 159–194.
- S. Li and W. K. Liu (2002), 'Meshfree and particle methods and their application', *Applied Mechanics Review* **55**, 1–34.
- G. R. Liu (2002), *Meshless Methods*, CRC Press, Boca Raton.
- W. K. Liu, S. Jun and Y. F. Zhang (1995), 'Reproducing kernel particle methods', *Int. J. Numer. Meth. Fluids* **20**, 1081–1106.
- W. K. Liu, S. Li and T. Belytschko (1997), 'Moving least square reproducing kernel particle method: Methodology and convergence', *Comput. Methods Appl. Mech. Engr.* **143**, 422–453.
- Y. Y. Lu, T. Belytschko and L. Gu (1994), 'A new implementation of the element free Galerkin method', *Comput. Methods Appl. Mech. Engr.* **113**, 397–414.
- A. M. Matache, I. Babuška and C. Schwab (2000), 'Generalized  $p$ -FEM in homogenization', *Numer. Math.* **86**, 319–375.
- D. H. McLain (1974), 'Drawing contours from arbitrary data points', *Comp. J.* **17**, 89–97.

- J. M. Melenk and I. Babuška (1996), ‘The partition of unity finite element method: Theory and application’, *Comput. Methods Appl. Mech. Engr.* **139**, 289–314.
- S. G. Mikhlin (1971), *Numerical Performance of Variational Methods*, Nordhoff.
- A. Moes, J. Dolbow and T. Belytschko (1999), ‘A finite element method for crack growth without remeshing’, *Internat. J. Numer. Engr.* **46**, 131–150.
- N. Moes, A. Gravouil and T. Belytschko (2002), ‘Non-planar 3D crack growth by extended finite element and level sets, Part 1: Mechanical model’, *Internat. J. Numer. Meth. Engr.* **53**, 2549–2568.
- J. J. Monaghan (1982), ‘Why particle methods work’, *SIAM J. Sci. Stat. Comp.* **3**, 422–433.
- J. J. Monaghan (1988), ‘An introduction to SPH’, *Comp. Phys. Comm.* **48**, 89–96.
- Y. K. Mukherjee and S. Mukherjee (1997), ‘On boundary conditions in the element-free Galerkin method’, *Computational Mechanics* **19**, 264–270.
- K. Nanbu (1980), ‘Direct simulation scheme derived from the Boltzmann equation’, *J. Phys. Soc. Japan* **49**, 20–40.
- H. Neunzert and J. Struckmeier (1995), Particle methods for the Boltzmann equation, in *Acta Numerica*, Vol. 4, Cambridge University Press, pp. 417–457.
- J. Nitsche (1970–1971), ‘Über ein Variationsprinzip zur Lösung von Dirichlet-Problemen bei Verwendung von Teilräumen, die keinen Randbedingungen unterworfen sind’, *Abh. Math. Univ. Hamburg* **36**, 9–15.
- J. Nitsche and A. H. Schatz (1974), ‘Interior estimates for Ritz–Galerkin methods’, *Math. Comp.* **28**, 937–958.
- J. T. Oden, C. A. Duarte and O. C. Zienkiewicz (1998), ‘A new cloud based hp finite element method’, *Comput. Methods Appl. Mech. Engr.* **153**, 117–126.
- J. Pitkäranta (1979), ‘Boundary subspaces for finite element method with Lagrange multipliers’, *Numer. Math.* **33**, 273–289.
- J. Pitkäranta (1980), ‘Local stability conditions for the Babuska method of Lagrange multipliers’, *Math. Comp.* **35**, 1113–1129.
- M. J. D. Powell (1992), The theory of radial basis function approximation in 1990, in *Advances in Numerical Analysis, II* (W. Light, ed.), Clarendon Press, pp. 105–210.
- W. Rudin (1991), *Functional Analysis*, McGraw-Hill.
- A. H. Schatz and L. B. Wahlbin (1995), ‘Maximum norm estimates for finite element methods, Part II’, *Math. Comp.* **64**, 907–928.
- M. A. Schweitzer (200x), *A Parallel Multilevel Partition of Unity Method*. Book manuscript.
- D. Shepard (1968), A two-dimensional interpolation function for irregularly spaced points, in *Proc. ACM Nat. Conf.*, pp. 517–524.
- E. M. Stein (1970), *Singular Integrals and Differentiability Properties of Functions*, Princeton University Press.
- R. Stenberg (1995), ‘On some techniques for approximating boundary conditions in the finite element method’, *J. Comp. Appl. Math.* **63**, 139–148.
- G. Strang (1971), The finite element method and approximation theory, in *Numerical Solution of Partial Differential Equations II, SYNPADE 1970* (B. Hubbard, ed.), Academic Press, London, pp. 547–584.

- G. Strang and G. Fix (1973), A Fourier analysis of finite element variational method, in *Constructive Aspects of Functional Analysis*, Edizioni Cremonese, pp. 795–840.
- T. Strouboulis, I. Babuška and K. Copps (2001a), ‘The design and analysis of the generalized finite element method’, *Comput. Methods Appl. Mech. Engr.* **181**, 43–69.
- T. Strouboulis, K. Copps and I. Babuška (2001b), ‘The generalized finite element method’, *Comput. Methods Appl. Mech. Engr.* **190**, 4081–4193.
- T. Strouboulis, L. Zhang and I. Babuška (200x), Generalized finite element method using mesh-based handbooks: Applications in domains with many voids. Manuscript.
- N. Sukumar, N. Moes, B. Moran and T. Belytschko (2000), ‘Extended finite element method for three dimensional crack modelling’, *Internat. J. Numer. Meth. Engr.* **48**, 1549–1570.
- B. Szabo, I. Babuška and R. Actis (1998), ‘A simpler way to handle complex analysis’, *Machine Design*.
- H. Tada, P. C. Paris and G. R. Irwin (1973), *The Stress Analysis of Cracks Handbook*, Del Research Corporation, Hellertown, PA.
- R. L. Taylor, O. C. Zienkiewicz and E. Onate (1998), ‘A hierarchical finite element method based on partition of unity’, *Comput. Methods Appl. Mech. Engr.* **152**, 73–84.
- R. S. Varga (1971), *Functional Analysis and Approximation Theory in Numerical Analysis*, Regional Conference Series in Applied Mathematics, SIAM, Philadelphia, PA.
- I. N. Vekua (1967), *New Methods for Solving Elliptic Equations*, North-Holland.
- L. B. Wahlbin (1995), *Superconvergence in Galerkin Finite Element Methods*, Vol. 1605 of *Lecture Notes in Mathematics*, Springer.
- G. N. Wells and L. J. Sluis (2001), ‘A new method for modelling cohesive finite elements’, *Internat. J. Numer. Meth. Engr.* **50**, 2667–2682.
- G. N. Wells, R. de Borst and L. J. Sluis (2002), ‘A consistent geometrical nonlinear approach for delamination’, *Internat. J. Numer. Meth. Engr.* **54**, 1333–1355.
- J. Xu and L. T. Zikatanov (2002), On multigrid methods for generalized finite element methods, in *Meshfree Methods for Partial Differential Equations* (M. Griebel and M. A. Schweitzer, eds), Springer, pp. 401–418.
- L. T. Zhang, W. K. Liu, S. F. Li, D. Qian and S. Hao (2002), Survey of multi-scale meshfree particle methods, in *Meshfree Methods for Partial Differential Equations* (M. Griebel and M. A. Schweitzer, eds), Springer, pp. 441–458.
- T. Zhu and S. N. Atluri (1998), ‘A modified collocation method and a penalty formulation for enforcing essential boundary conditions’, *Computational Mechanics* **21**, 165–178.