Acta Numerica (2003), pp. 1–125
DOI: 10.1017/S0962492902000090

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

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

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

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

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

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,

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

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 (200x).

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

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

and

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

or

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

where $f \in L_2(\Omega)$ is given. We will assume that Ω 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

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

where

$$B(u, v) \equiv \int_\Omega (\nabla u \cdot \nabla v + uv)\, \mathrm{d}x \text{ and } \mathcal{F}(v) \equiv \int_\Omega fv\, \mathrm{d}x. \tag{2.5}$$

The energy norm of u_0 is defined by

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

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

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

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 Ω 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 $j = (j_1, \ldots, j_n) \in \mathbb{Z}^n$ and $0 < h \leq 1$, let

$$x_j^h = (j_1 h, \ldots, j_n h) = hj.$$

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 $\eta \equiv$ supp ϕ, 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 η). Then define

$$\phi_j^h(x) = \phi_j^h(x_1, \ldots, x_n) = \phi\left(\frac{x - jh}{h}\right) = \phi\left(\frac{x_1 - j_1 h}{h}, \ldots, \frac{x_n - j_n h}{h}\right), \tag{3.1}$$

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 ϕ is called the *basic shape function*. In this section we will be interested in the approximation properties of

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

which is the linear span of the associated shape functions, as $h \to 0$. 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

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

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

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

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 ϕ).

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.