

A G space theory and a weakened weak (W^2) form for a unified formulation of compatible and incompatible methods: Part I theory

G. R. Liu^{1,2,*},[†]

¹*Centre for Advanced Computations in Engineering Science, Department of Mechanical Engineering, National University of Singapore, 9 Engineering Drive 1, 117576 Singapore, Singapore*

²*Singapore—MIT Alliance (SMA), E4-04-10, 4 Engineering Drive 3, 117576 Singapore, Singapore*

SUMMARY

This paper introduces a G space theory and a weakened weak form (W^2) using the generalized gradient smoothing technique for a unified formulation of a wide class of compatible and incompatible methods. The W^2 formulation works for both finite element method settings and mesh-free settings, and W^2 models can have special properties including softened behavior, upper bounds and ultra accuracy. Part I of this paper focuses on the theory and fundamentals for W^2 formulations. A normed G space is first defined to include both continuous and discontinuous functions allowing the use of much more types of methods/techniques to create shape functions for numerical models. Important properties and a set of useful inequalities for G spaces are then proven in the theory and analyzed in detail. These properties ensure that a numerical method developed based on the W^2 formulation will be spatially stable and convergent to the exact solutions, as long as the physical problem is well posed. The theory is applicable to any problems to which the standard weak formulation is applicable, and can offer numerical solutions with special properties including ‘close-to-exact’ stiffness, upper bounds and ultra accuracy. Copyright © 2009 John Wiley & Sons, Ltd.

Received 17 July 2008; Revised 6 June 2009; Accepted 23 June 2009

KEY WORDS: numerical method; mesh-free method; G space; weakened weak form; point interpolation method; finite element method; solution bound; variational principle; Galerkin weak form; numerical method; compatibility

1. INTRODUCTION

To solve engineering problems, many powerful numerical methods based on weak form formulation have been developed, such as the finite element method (FEM) [1–3] and recently the mesh-free

*Correspondence to: G. R. Liu, Centre for Advanced Computations in Engineering Science, Department of Mechanical Engineering, National University of Singapore, 9 Engineering Drive 1, 117576 Singapore, Singapore.

[†]E-mail: mpeliugr@nus.edu.sg

Contract/grant sponsor: A*STAR, Singapore

Contract/grant sponsor: Open Research Fund Program of the State Key Laboratory of Advanced Technology of Design and Manufacturing for Vehicle Body, Hunan University; contract/grant number: 40915001

methods (see, e.g. [4–6]). The FEM is well developed and is currently the most widely used reliable numerical tool, and many established commercial software packages available. However, there are three major issues related to the FEM. The first issue is the ‘overly stiff’ phenomenon of a fully compatible FEM model of assumed displacement based on the Galerkin weak form, which can have consequences of (1) the so-called ‘locking’ behavior for many problems and (2) inaccuracy in stress solutions. The second issue concerns with the mesh distortion-related problems such as the significant accuracy loss when the element mesh is heavily distorted. The third issue is the mesh generation. We, engineers, often prefer using triangular types of mesh as they can be created much more easily and even automatically for problems of complicated geometry. However, it is well known that the FEM does not like such elements and often give solutions of very poor accuracy. This is the reason for analysts often getting a warning when opting for triangular elements in some commercial software packages.

The overly stiff phenomenon is attributed to nature of the fully compatible displacement approach based on the standard variational principle. Many efforts have been made in resolving this issue, especially in the area of hybrid FEM formulations (see, e.g. [7, 8]). Recently, a smoothed FEM (or SFEM) [9–11] has also been formulated by combining the FEM procedures and the *gradient smoothing operation* known as distributional derivatives in classic sense. The smoothing operation is a very useful numerical tool and has been used in various situations, such as the nonlocal continuum mechanics [12, 13], the smoothed particle hydrodynamics (SPH) [4, 14–16], hybrid FEMs [8], resolving the material instabilities [17] and spatial instability in nodal integrated mesh-free methods [18], and recently obtaining upper bound solution in mesh-free point interpolation methods [19, 20]. The SFEM also uses the smoothing operations based on cells and is found working very effectively for solid mechanics problems and n -sided polygonal elements and very heavily distorted mesh can be used [10]. Detailed theoretical aspects including stability and convergence about SFEM can be found in [11]. The study of SFEM has also clearly shown that the smoothing operation on strains controls the assumed strain field in a proper fashion to ensure the stability and the convergence, and ultimately gives the SFEM some very good features. A more general cell-based smoothed model is the recent cell-based smoothed point interpolation method [21] using general PIM or RPIM shape functions.

In the other front of development related to mesh-free methods, the node-based smoothed point interpolation method (NS-PIM)[‡] has been developed recently [19, 20] using the node-based strain smoothing operations [18] with extensions to discontinuous assumed displacement functions [22]. The NS-PIM is formulated using PIM [5, 23] or RPIM [24] shape functions of Kronecker delta for easy treatment of essential boundary conditions. It was found that NS-PIM or (NS-RPIM [25]) is at least linearly conforming (can always pass the standard patch tests when linear displacements on the boundary are enforced), can produce much better stress solution, much more tolerant to mesh distortion, works very well for triangular cells, and more importantly it provides upper bound solution in energy norm [26]. Following further the idea of NS-PIM and SFEM, a node-based SFEM (or NS-FEM) has also been formulated within the framework of FEM. The NS-FEM can be viewed as a special case of NS-PIM, but based on n -sided polygonal element mesh [27], and has quite similar properties as NS-PIM. It was found that NS-PIM and NS-FEM behave ‘overly soft’ leading to temporal instability when used to solve dynamic problems. To reduce the softening

[‡]The NS-PIM was originally termed as the linearly conforming point interpolation method (LC-PIM), because it is at least linearly conforming. We changed the name because the later formulations of cell-based and edge-based smoothing techniques those are all at least linearly conforming, but distinct in the creation of smoothing domains.

effects, a very effective edge-based smoothed FEM (ES-FEM) for 2D problems [28] and face-based smoothed FEM (FS-FEM) for 3D problems [29] have been recently formulated. The ES-FEM (or FS-FEM) not only produces accurate solution but also is temporally stable and no spurious modes and hence works very well for dynamic problems. The linear ES-FEM using triangular mesh has been found as much as 10 times more accurate in displacement norm than the FEM using the same mesh, and hence known as an ‘ultra-accurate’ model with ‘close-to-exact’ stiffness. A more general edge-based smoothed model is the recent ES-PIM [30] using general PIM or RPIM shape functions.

In examining the above-mentioned works, we find that (1) both compatible and incompatible displacement fields created with FEM and mesh-free settings were used; (2) the traditional weak form is much extended in various manners by changing the bilinear form; and (3) the integration is performed in novel ways far beyond the standard FEM procedures. It is now clearly necessary to establish a new theoretical framework to unify the formulation of all these newly developed element-based or mesh-free methods.

In this work, we attempt to do so by putting together the pieces of recent advances in FEM and mesh-free methods, and establish a new G space theory and a weakened weak form (W^2) formulation as a theoretical framework for all these methods: compatible and incompatible ones. Part I of this paper focuses on the G space theory and fundamentals for W^2 formulation. We first define a G space to include both continuous and discontinuous displacement functions using the generalized gradient smoothing technique [22]. We then prove the important properties of the functions in a G space, including a set of key inequalities that are the foundation for the stability and convergence for a numerical method seeking solution in a G space. These properties ensure that a numerical method developed based on the W^2 formulation will be spatially stable, and convergent to exact solutions, as long as the physical problem is well posed. The theory is applicable to any problems to which the standard weak formulation is applicable, and can offer numerical solutions with special properties including close-to-exact stiffness features, ultra accuracy and upper bounds. The procedure of W^2 formulation and the application of the W^2 formulation to create models for solid mechanics problems will be given in detail in Part II of this paper.

2. FUNCTION APPROXIMATION

We first consider the most basic issue of function approximation in general settings including the FEM settings with a mesh of elements and mesh-free settings. Consider a d -dimensional problem domain of $\Omega \in \mathbb{R}^d$ bounded by Γ that is ‘Lipschitzian’. By default, we speak ‘open’ domain that does not include the boundary of the domain. When we refer to a ‘closed’ domain we will specifically use a box: $\overline{\Omega} = \Omega \cup \Gamma$.

2.1. Mesh: triangulation

In this work, the problem domain is divided into cells (for mesh-free settings) or elements (for FEM settings) of general polygons of multiple sides [11, 27]. However, we prefer triangular types of cells/elements for easy generation for complicated domains using the well-established triangulation algorithms such as the Delaunay algorithm. For 1D problems a cell is defined in \mathbb{R}^1 and is simply a line segment, for 2D problems it is defined in \mathbb{R}^2 and becomes a triangle, and for 3D problems it is defined in \mathbb{R}^3 and is tetrahedron. The domain is divided with N_e cells/elements that are

connected at N_n nodes. The ‘length’ of the cell is denoted *generally* as h and it can be different from cell to cell. For a uniform discretization, h becomes the characteristic dimension of cells. In the case of non-uniform discretization, we let

$$h_{\max} = \max_{i=1, \dots, N_e} (h_i), \quad h_{\min} = \min_{i=1, \dots, N_e} (h_i) \quad (1)$$

and we should assume that the ratio of the smallest and largest cell dimensions is bounded:

$$h_{\max}/h_{\min} = c_{rh} < \infty \quad (2)$$

In such cases, the largest h_{\max} becomes the characteristic dimension of the cells. We require the ratio c_{rh} to be bounded, and also the inner angles θ of the triangles should be strictly larger than zero and less than 180° in theory, and in practice we often require $15 < \theta < 120$. Under such conditions, the largest h_{\max} becomes the characteristic dimension of the cells: controlling h_{\max} puts the entire mesh under control. When we say h approaches zero, the dimensions of all the cells in the entire problem domain approach to zero. Our division is also seamless: $\boxed{\Omega} = \bigcup_{i=1}^{N_e} \boxed{\Omega}_i^e$, where the box stands for closed domains. Such a triangulation is *generally* denoted as T_h : a collection of all Ω_i^e ($i = 1, 2, \dots, N_e$) (exclusive of the boundaries).

2.2. Basis

In any (discrete) numerical method, field functions have to be approximated over the problem domain using a set of nodal values of the functions and the so-called *basis*. Given a linear space S of dimension N_n , a set of N_n members of functions $\phi_n \in S, n = 1, 2, \dots, N_n$ is a basis for S if and only if $\forall w \in S, \exists$ unique $\alpha_n \in \mathbb{R}$ such that

$$w = \sum_{n=1}^{N_n} \alpha_n \phi_n \quad (3)$$

Functions ϕ_n in the *basis* are often given in the form of *nodal* shape functions, and hence the basis is also termed as *nodal basis* in the context of FEM and mesh-free methods. Equation (3) implies that the nodal shape functions must be linearly independent. In the FEM, these linearly independent shape functions are created based on elements using mostly polynomial basis functions, and the linearly independence is ensured by element topology and properly controlled coordinate mapping (see, e.g. [2]). In the mesh-free methods it is based on local nodes using both polynomial and radial basis functions [5], generally no mapping is needed, and the linearly independence is ensured by the use of proper basis functions and/or proper local nodes selection with the help of a background cells [6]. The often used shape functions include: PIM, RPIM, least square (LS), moving least square (MLS), and SPH shape functions [6]. The procedure for the creation of these types of shape functions is rather standard and can be found in great detail in [5, 6], and hence are omitted here. In this paper we assume that there are always a possible ways to obtain a set of independent nodal shape functions for a given set of nodes with a set of elements or background cells in the problem domain Ω .

In addition, for consistence reasons, we require this set of nodal shape functions being *polynomial* linearly complete (or at least when h approaches zero). The linearly completion can be simply achieved by adding in polynomials with constant and linear terms [6], and the resultant models can pass the standard patch tests (see part II). When the pure RBFs are used in constructing the RPIM shape functions, the *polynomial* linearly completion is lost. However, such a linearly completion

can be achieved at the limit of h approaching zero [5]. In this case, the model may not be able to pass the standard patch tests, but will converge when refined [6], if the numerical model is stable.

For the convenience of discussion, we require the shape functions ϕ_n are positions of unity

$$\sum_{n=1}^{N_n} \phi_n(\mathbf{x}) = 1 \quad (4)$$

which ensures correct representation of rigid-body movement, and the Kronecker delta property

$$\phi_i(\mathbf{x}_j) = \begin{cases} 1 & \text{when } i = j \\ 0 & \text{when } i \neq j \end{cases} \quad (5)$$

which allows easy treatments for essential boundary conditions [5, 6]. Once the nodal shape functions are obtained, an assumed (displacement component) function can be given as

$$u^h(\mathbf{x}) = \sum_{n \in S_s} \phi_n(\mathbf{x}) d_n \quad (6)$$

where $\mathbf{x} = \{x_1, x_2\}^T$, S_s is the set of the nodes in the support domain of the cell/element hosting \mathbf{x} , d_n is the nodal value of the (displacement component) function.

2.3. Point interpolation

Given $w \in S$ where S is a linear space, the *interpolant* $\mathcal{I}_h w$ creates a function that lives in a subspace $S_h \subset S$ with N_n dimensions: $\mathcal{I}_h w \in S_h$, where

$$\mathcal{I}_h w(\mathbf{x}) = \sum_{n=1}^{N_n} w(\mathbf{x}_n) \phi_n(\mathbf{x}) \quad (7)$$

With the Delta function property given in Equation (5), we have

$$\mathcal{I}_h w(\mathbf{x}_n) = w(\mathbf{x}_n), \quad n = 1, 2, \dots, N_n \quad (8)$$

All the FEM shape functions and the general PIM and RPIM shape functions can be used for such an interpolation. Note the interpolant generated by the above *point interpolation* does not necessarily live in a desired Hilbert (H^1) space, and it lives in a G^1 space (to be defined later). Therefore, the point interpolation defined here is generally different from those defined in the FEM settings where the interpolation is element based with proper mapping to ensure that the interpolant lives in a desired H^1 space [31]. An interpolant in an H space lives also in the corresponding G space.

2.4. Gradient approximation: 'smoothed' strains

Different from the standard weak formulation, our W^2 formulation also needs to approximate the gradient of functions (strains), in addition to the usual approximation of field (displacement) functions. Smoothing techniques for functions and the gradient of functions are a very useful in numerical methods for function approximation [6]. Such a techniques were widely used in many numerical operations/treatments, such as the nonlocal continuum mechanics [12, 13], the SPH [4, 14, 16], hybrid FEM models [8], stabilizing nodal integrated mesh-free methods [18], and restoring conformability and obtaining upper bound solution in mesh-free point interpolation

methods [19, 26]. Using the smoothing technique and the generalized smoothing technique that works also for discontinuous functions [22], we defined the ‘smoothed’ strain as (Figure 1)

$$\overline{\frac{\partial w_l}{\partial x_i}}(\mathbf{x}) = \begin{cases} \frac{1}{A_{\mathbf{x}}^s} \int_{\Omega_{\mathbf{x}}^s} \frac{\partial w_l}{\partial x_i} d\Omega = \frac{1}{A_{\mathbf{x}}^s} \int_{\Gamma_{\mathbf{x}}^s} w_l(s) n_i ds & \text{when } w_l(\mathbf{x}) \text{ is continuous in } \Omega_{\mathbf{x}}^s \\ \frac{1}{A_{\mathbf{x}}^s} \int_{\Gamma_{\mathbf{x}}^s} w_l(s) n_i ds & \text{when } w_l(\mathbf{x}) \text{ is discontinuous in } \Omega_{\mathbf{x}}^s \end{cases} \quad (9)$$

In the first equation in (9), we used the Green’s Theorem and have both domain integration and boundary line integration for the convenience in later derivation. Since the smoothing domain $\Omega_{\mathbf{x}}^s$ used in the above equations changes (or moves) with \mathbf{x} , it is termed as *moving* smoothing domain. In this work we use *stationary* smoothing domains that are fixed for a point of interest. We do not allow the smoothing domains to overlap: $[\Omega] = \bigcup_{k=1}^{N_s} [\Omega]_k^s$, where Ω_k^s is a smoothing domain bounded by Γ_k^s for point at \mathbf{x}_k , as shown in Figure 2. In this case, Equation (9) becomes:

$$\overline{\frac{\partial w_l}{\partial x_i}}(\mathbf{x}) = \overline{\frac{\partial w_l}{\partial x_i}}(\mathbf{x}_k) = \begin{cases} \frac{1}{A_k^s} \int_{\Omega_k^s} \frac{\partial w_l}{\partial x_i} d\Omega = \underbrace{\frac{1}{A_k^s} \int_{\Gamma_k^s} w_l(s) n_i ds}_{\text{constant in } \Omega_k^s}, & w_l(\mathbf{x}) \text{ is continuous in } \Omega_k^s \\ \underbrace{\frac{1}{A_k^s} \int_{\Gamma_k^s} w_l(s) n_i ds}_{\text{constant in } \Omega_k^s}, & w_l(\mathbf{x}) \text{ is discontinuous in } \Omega_k^s \end{cases} \quad (10)$$

$l = 1, 2, \quad i = 1, 2 \quad \forall \mathbf{x} \in \Omega_k^s$

In carrying out the line integrations on the boundary Γ_k^s in creating a W^2 model, we simply use the standard Gauss integration widely used in FEM [2]. The detailed procedure that leads to Equation (9) and (10) can be found in [22]. Note that when the function is not continuous, the compatible strain does not exist at locations in the problem domain. Therefore, the ‘smoothed’ strain is not exactly the strain obtained by smoothing the compatible strain field for such cases. To be precise, the ‘smoothed’ strain field should be an approximated strain field.

3. G SPACES

We now introduce G spaces of functions of *finite dimensions*. For convenience in discussion, we need to frequently refer to Hilbert or H spaces.

3.1. Brief on H spaces

The H^1 for space used in this work is defined (for 2D cases) as (see, e.g. [32])

$$\mathbb{H}^1(\Omega) = \{v | v \in \mathbb{L}^2(\Omega), \partial v / \partial x_i \in \mathbb{L}^2(\Omega), i = 1, 2\} \quad (11)$$

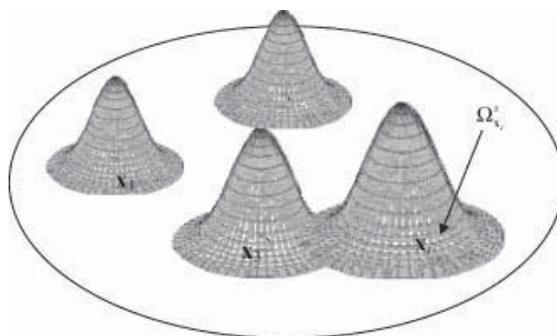


Figure 1. Moving smoothing domains Ω_x^s for the integral representation of a function at \mathbf{x} , over which the smooth function is defined. Note that the smoothing domain can be different for different \mathbf{x} and they can overlap. The smoothing functions can also be different for different \mathbf{x} .

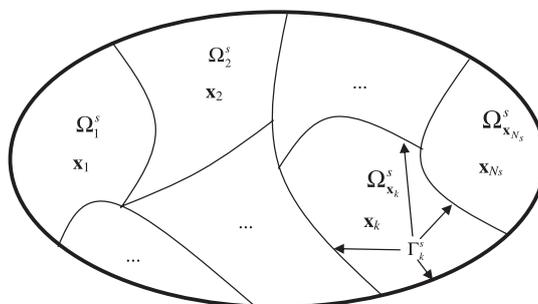


Figure 2. Division of problem domain Ω into non-overlapping stationary smoothing domains Ω_k^s or \mathbf{x}_k bounded by Γ_k^s . The smoothing domain is also used as basis for integration.

and in particular $\mathbb{H}_0^1(\Omega) = \{v \in \mathbb{H}^1(\Omega) | v_i = 0 \text{ on } \Gamma_D\}$. The H^1 full norm is defined as

$$\|v\|_{\mathbb{H}^1(\Omega)}^2 = \underbrace{\int_{\Omega} v^2 d\Omega}_{\|w\|_{L^2(\Omega)}^2} + \underbrace{\int_{\Omega} (\nabla v) \cdot (\nabla v) d\Omega}_{|w|_{\mathbb{H}^1(\Omega)}^2} \tag{12}$$

where

$$\nabla v = \left(\frac{\partial v}{\partial x_1} \quad \frac{\partial v}{\partial x_2} \right) \tag{13}$$

The H^1 semi-norm becomes

$$|v|_{\mathbb{H}^1(\Omega)}^2 = \int_{\Omega} (\nabla v) \cdot (\nabla v) d\Omega \tag{14}$$

In this work, we denote a subspace in H^1 space created using interpolation techniques that ensure compatibility, which can then be defined as

$$\mathbb{H}_h^1(\Omega) = \{v \in \mathbb{H}^1(\Omega) | v(\mathbf{x}) = \boldsymbol{\varphi}^H(\mathbf{x})\mathbf{d}, \mathbf{d} \in \mathbb{R}^{N_n}\} \tag{15}$$

where $\boldsymbol{\phi}^H(\mathbf{x})$ is the matrix of all the (compatible) nodal shape functions that are linearly independent constructed using a proper FEM model, and can be written as

$$\boldsymbol{\phi}^H(\mathbf{x}) = [\phi_1^H(\mathbf{x}) \ \phi_2^H(\mathbf{x}) \ \dots \ \phi_{N_n}^H(\mathbf{x})] \tag{16}$$

Because \mathbb{H}_h^1 is a linear space, each of the nodal shape functions $\phi_i^H(\mathbf{x})$ must also be in \mathbb{H}_h^1 . The linearly independence of shape functions $\phi_i^H(\mathbf{x})$, ($i = 1, 2, \dots, N_n$) is ensured by a standard FEM procedure (element based and proper mapping). In Equation (15) \mathbf{d} is the vector of all the nodal functions values given in the form of

$$\mathbf{d} = \{v_1 \ v_2 \ \dots \ v_{N_n}\}^T \tag{17}$$

Since the values at each node can change independently, we have $\mathbf{d} \in \mathbb{R}^{N_n}$, where \mathbb{R}^{N_n} stands for a real field of N_n dimensions.

Because \mathbb{H}_h^1 is constructed in a discrete form with finite dimensions, it is marked with a subscript ‘ h ’. Functions in \mathbb{H}_h^1 that satisfy the essential (displacement) boundary conditions form a space:

$$\mathbb{H}_{h,0}^1(\Omega) = \{v \in \mathbb{H}_h^1(\Omega) | v = 0 \text{ on } \Gamma_u\} \tag{18}$$

An \mathbb{H}_h^1 space is indeed very exclusive, and the methods that can be used to create functions in an \mathbb{H}_h^1 space are very much limited: the standard FEM technique and the properly performed MLS approximation [6].

3.2. Definitions for G spaces

3.2.1. Smoothing domain creation. Consider a domain Ω discretized by, for example, triangulation with N_e non-overlapping subdomains (called cells in mesh-free context or elements in FEM settings) $\overline{\Omega} = \bigcup_{i=1}^{N_e} \overline{\Omega}_i^e$ with a set of N_n nodes, and N_c^Γ line segments Γ_i^c ($i = 1, \dots, N_c^\Gamma$) which divide the domain Ω into cells/elements. We next divide, in a basically independent way, the domain Ω into N_s non-overlapping subdomains called smoothing domains: $\overline{\Omega} = \bigcup_{k=1}^{N_s} \overline{\Omega}_k^s$ with N_s^Γ line interfaces Γ_i^s ($i = 1, \dots, N_s^\Gamma$) between the smoothing domains. The division of Ω into Ω_k^s is performed in such a way that the interfaces Γ_k^s of Ω_k^s do not share any *finite* portion of the interfaces Γ_i^c on which the function is not square integrable. The interfaces of Ω_k^s can go across Γ_i^c . Only when the function is continuous on Γ_i^c , the sharing of the interfaces of Ω_k^s and Ω_i^e may be permitted. A typical division of domain is given in Figure 3 for 2D domains, where node-based smoothing domains are created on a set of triangular background cells. Figure 4 shows an example for 1D domains.

3.2.2. G^1 space and norms. The \mathbb{G}_h^1 that is relevant to this work can be then defined as follows:

$$\mathbb{G}_h^1(\Omega) = \left\{ \begin{array}{l} v | \mathbf{v}(\mathbf{x}) = \sum_{n=1}^{N_n} \phi_n(\mathbf{x}) d_n = \boldsymbol{\phi}(\mathbf{x}) \mathbf{d}, \ \mathbf{d} \in \mathbb{R}^{N_n} \\ v \in L^2(\Omega), \\ \sum_{k=1}^{N_s} \left(\int_{\Gamma_k^s} v(s) n_i ds \right)^2 > 0 \Leftrightarrow v \neq c \in \mathbb{R}; i = 1, \dots, d \end{array} \right\} \tag{19}$$

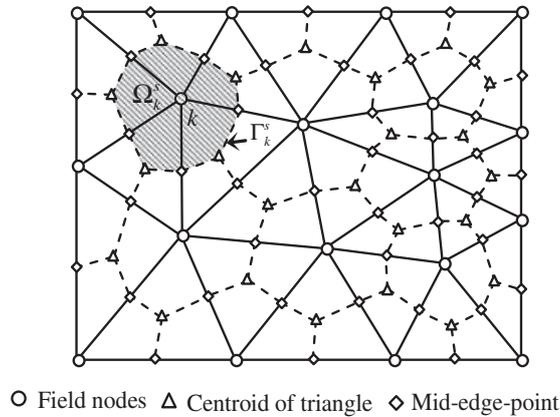


Figure 3. A typical division of domains for G space function. Triangular elements (bounded by solid lines) and the non-overlapping stationary smoothing domain Ω_k^s (bounded by dashed lines Γ_k^s) for node n created by connecting the centroids with the mid-edge-points of the surrounding triangles of a node. Γ_k^s does not share any finite portion of any internal edges of the triangular elements. The smoothing operation is performed over the entire node-based smoothing cell.

where $\phi_n(\mathbf{x})$ is the nodal shape function for node n , and these N_n shape functions form the basis of the \mathbb{G}_h^1 space. It is observed that the \mathbb{G}_h^1 space is of finite dimension (denoted by subscript ‘ h ’) and with a set of functions that are square integrable in Ω formed by point interpolation using a basis. These nodal shape functions are created using nodes selected based on elements/cells using the standard FEM or mesh-free procedures, and hence they are continuous at all these nodes and on these interfaces $\Gamma_k^s (k = 1, \dots, N_s^\Gamma)$. The continuity on all these interfaces Γ_k^s allows a unique evaluation of the generalized smoothed gradient of the functions over the smoothing domains Ω_k^s , so that the variation of the functions over Ω can be captured in a local averaged fashion. We do not restrict on how the basis is created, as long as these nodal shape functions are linearly independent over Ω and hence are capable to form a basis. The norms for \mathbb{G}_h^1 spaces are induced from inner products defined as follows.

3.2.3. G^1 norms for 1D scalar fields. The associated inner product is given by:

$$(w, v)_{\mathbb{G}^1(\Omega)} = \int_{\Omega} wv \, d\Omega + \sum_{k=1}^{N_s} \Omega_k^s \overline{w'} \cdot \overline{w'} = \underbrace{\int_{\Omega} wv \, d\Omega}_{(w,v)} + \underbrace{\sum_{k=1}^{N_s} A_k^s \bar{g}(w) \bar{g}(v)}_{(\overline{w'}, \overline{w'})} \tag{20}$$

Note the summation is possible because the division of Ω into Ω_k^s is performed in such a way that the interfaces Γ_k^s of Ω_k^s do not share any finite portion of the interfaces Γ_i^c on which the function is not square integrable: no energy loss in the interface of the smoothing domains. In Equation (20) the (approximated) smoothed gradient is denoted as

$$\overline{w'} = \frac{\partial \overline{w}}{\partial x} = \frac{1}{A_k^s} \int_{\Gamma_k^s} w(s) n_x \, ds = \underbrace{\frac{1}{A_k^s} (w_{k+1/2} - w_{k-1/2})}_{=\bar{g}, \text{ constant in } \Omega_k^s} = \bar{g}(w) \tag{21}$$

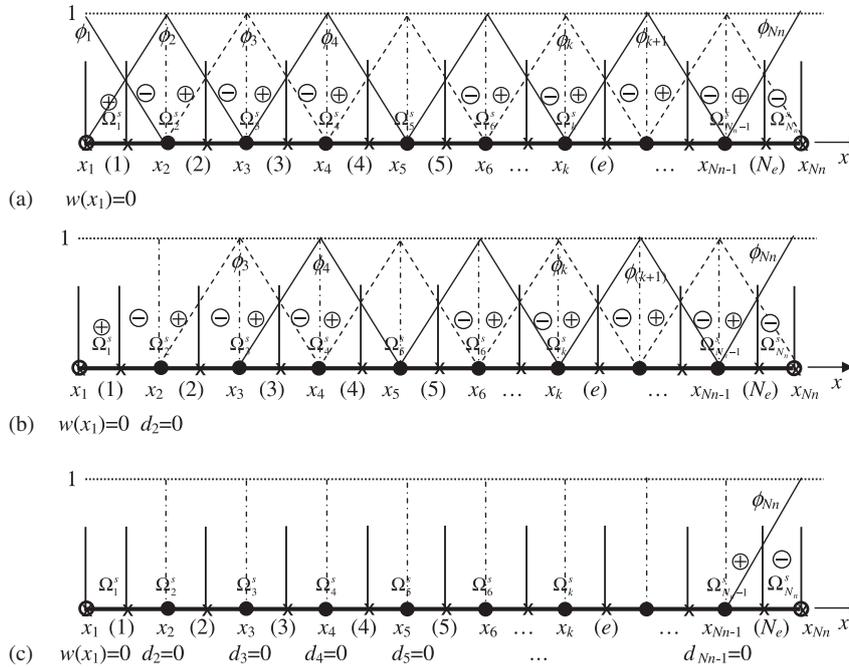


Figure 4. Division of a 1D problem domain Ω into N_e elements with N_e nodes at \mathbf{x}_k , and the linear finite element shape functions are used for creating function w . The domain is then divided into $N_s = N_n$ non-overlapping node-based smoothing domains Ω_k^s for \mathbf{x}_k bounded by $x_{k-1/2}$ and $x_{k+1/2}$. The smoothing domain is used as a base for integration. The interpolation of function w and integration using node-based smoothing domains create a positivity relay for the semi-norm of w starting from the Dirichlet boundary, resulting in the second inequality. (a) Function w is fixed at the left end by the given Dirichlet boundary condition. The semi-norm contributed from Ω_1^s gives surely a positive value, as long as $d_2 \neq 0$; (b) when $d_2 = 0$, Ω_2^s now gives surely a positive value, as long as $d_3 \neq 0$; and (c) when $d_2 = d_3 = \dots = d_{N_n-1} = 0$, both $\Omega_{N_n-1}^s$ and $\Omega_{N_n}^s$ give positive values, as long as $d_{N_n} \neq 0$. If $d_{N_n} = 0$, it becomes a case of fixed boundary and Ω_{N_n} will still be positive. The ‘positivity relay’ completes.

where $\bar{g}(w)$ denotes the smoothed derivatives of w with respect to x , and the smoothing domains Ω_k^s ‘centered’ at $x_k (=x_n$, in this 1D case) is bounded by $x_{k-1/2}$ and $x_{k+1/2}$, as shown in Figure 4. The $G^1(\Omega)$ semi-norm is next defined as

$$|w|_{G^1(\Omega)}^2 = \sum_{k=1}^{N_s} A_k^s |\bar{w}'|^2 = \underbrace{\sum_{k=1}^{N_s} A_k^s \bar{g}^2(w)}_{(\bar{w}', \bar{w}')} \tag{22}$$

and the G^1 full norm becomes

$$\|w\|_{G^1(\Omega)}^2 = \underbrace{\int_{\Omega} w^2 d\Omega}_{(w,w)=\|w\|_{L^2}^2} + \underbrace{|w|_{G^1(\Omega)}^2}_{(\bar{w}', \bar{w}')} = \|w\|_{L^2}^2 + |w|_{G^1(\Omega)}^2 \tag{23}$$

which is induced from the inner product equation (20).

3.2.4. *G¹ norms for 2D scalar fields.* The associated inner product is given by

$$(w, v)_{\mathbb{G}^1(\Omega)} = \int_{\Omega} wv \, d\Omega + \underbrace{\sum_{k=1}^{N_s} A_k^s \overline{\nabla w} \cdot \overline{\nabla v}}_{(w,v)} = \underbrace{\int_{\Omega} wv \, d\Omega}_{(w,v)} + \underbrace{\sum_{k=1}^{N_s} A_k^s (\bar{g}_1(w)\bar{g}_1(v) + \bar{g}_2(w)\bar{g}_2(v))}_{(\overline{\nabla w}, \overline{\nabla v})} \quad (24)$$

where the (approximated) smoothed gradient is denoted as

$$\overline{\nabla w} = \begin{pmatrix} \overline{\frac{\partial w}{\partial x_1}} & \overline{\frac{\partial w}{\partial x_2}} \end{pmatrix} = \begin{pmatrix} \underbrace{\frac{1}{A_k^s} \int_{\Gamma_k^s} w(s)n_1 \, ds}_{=\bar{g}_{1,\text{constant in } \Omega_k^s}} & \underbrace{\frac{1}{A_k^s} \int_{\Gamma_k^s} w(s)n_2 \, ds}_{=\bar{g}_{2,\text{constant in } \Omega_k^s}} \end{pmatrix} = (\bar{g}_1(w) \ \bar{g}_2(w)) \quad (25)$$

where $\bar{g}_i(w)$ denotes the smoothed derivatives of w with respect to x_i .

The $G^1(\Omega)$ semi-norm is next defined as

$$|w|_{\mathbb{G}^1(\Omega)}^2 = \sum_{k=1}^{N_s} A_k^s |\overline{\nabla w}|^2 = \underbrace{\sum_{k=1}^{N_s} A_k^s (\bar{g}_1^2(w) + \bar{g}_2^2(w))}_{(\overline{\nabla w}, \overline{\nabla w})} \quad (26)$$

and the G^1 full norm becomes

$$\|w\|_{\mathbb{G}^1(\Omega)}^2 = \underbrace{\int_{\Omega} w^2 \, d\Omega}_{(w,w)=\|w\|_{L^2}^2} + \underbrace{|w|_{\mathbb{G}^1(\Omega)}^2}_{(\overline{\nabla w}, \overline{\nabla w})} = \|w\|_{L^2}^2 + |w|_{\mathbb{G}^1(\Omega)}^2 \quad (27)$$

which is induced from the inner product equation (24). The definitions for 3D scalar fields are natural extension and hence are omitted here.

3.2.5. *G¹ norms for 2D vector fields.* For vector fields, we need to use vectors of functions. For example, when the function has two components, we should have $w = (w_1 \ w_2)$, where $w_1, w_2 \in \mathbb{G}_h^1$ are the two component functions. In this case, we have the smoothed gradient for the k th smoothing domain in the following form:

$$\overline{\nabla w} = \begin{pmatrix} \overline{\frac{\partial w_1}{\partial x_1}} & \overline{\frac{\partial w_1}{\partial x_2}} \\ \overline{\frac{\partial w_2}{\partial x_1}} & \overline{\frac{\partial w_2}{\partial x_2}} \end{pmatrix} = \begin{pmatrix} \underbrace{\frac{1}{A_k^s} \int_{\Gamma_k^s} w_1(s)n_1 \, ds}_{=\bar{g}_{11,\text{constant in } \Omega_k^s}} & \underbrace{\frac{1}{A_k^s} \int_{\Gamma_k^s} w_1(s)n_2 \, ds}_{=\bar{g}_{12,\text{constant in } \Omega_k^s}} \\ \underbrace{\frac{1}{A_k^s} \int_{\Gamma_k^s} w_2(s)n_1 \, ds}_{=\bar{g}_{21,\text{constant in } \Omega_k^s}} & \underbrace{\frac{1}{A_k^s} \int_{\Gamma_k^s} w_2(s)n_2 \, ds}_{=\bar{g}_{22,\text{constant in } \Omega_k^s}} \end{pmatrix} = \begin{pmatrix} \bar{g}_{11}(w_1) & \bar{g}_{12}(w_1) \\ \bar{g}_{21}(w_2) & \bar{g}_{22}(w_2) \end{pmatrix} \quad (28)$$

where $\bar{g}_{ij}(w)$ denotes the smoothed derivatives of w_i with respect to x_j . We notice here that the (smoothed) gradient is now a matrix, and hence there can be many *equivalent* ways to define the associated inner product. In this work, we decide to have the definition associated with the type of physical problems to be studied for convenience of proving necessary theories for that type of the problems. Considering 2D solid mechanics problems, we define the associated inner product in the form of

$$(w, v)_{\mathbb{G}^1(\Omega)} = \underbrace{\int_{\Omega} (w_1 v_1 + w_2 v_2) \, d\Omega}_{(w, v)} + \underbrace{\sum_{k=1}^{N_s} A_k^s [\bar{g}_{11}(w_1)\bar{g}_{11}(v_1) + \bar{g}_{22}(w_2)\bar{g}_{22}(v_2) + (\bar{g}_{12}(w_1) + \bar{g}_{21}(w_2))(\bar{g}_{12}(v_1) + \bar{g}_{21}(v_2))]}_{(\nabla w, \nabla v)} \tag{29}$$

The induced $G^1(\Omega)$ semi-norm is first defined as

$$|w|_{\mathbb{G}^1(\Omega)}^2 = \underbrace{\sum_{k=1}^{N_s} A_k^s (\bar{g}_{11}^2(w_1) + \bar{g}_{22}^2(w_2) + (\bar{g}_{12}(w_1) + \bar{g}_{21}(w_2))^2)}_{(\nabla w, \nabla w)} \tag{30}$$

It is clear that in our definition of the inner product and hence the induced the semi-norm we have intentionally related to the strain components, and hence the L^2 norm of the vector of strains.

The associated G^1 full norm can now be defined as

$$\|w\|_{\mathbb{G}^1(\Omega)}^2 = \underbrace{\int_{\Omega} (w_1^2 + w_2^2) \, d\Omega}_{(w, w) = \|w\|_{L^2}^2} + \underbrace{|w|_{\mathbb{G}^1(\Omega)}^2}_{(\nabla w, \nabla w)} = \|w\|_{L^2}^2 + |w|_{\mathbb{G}^1(\Omega)}^2 \tag{31}$$

3.2.6. G^1 norms for 3D vector fields. For vector fields with three-component functions in three-dimensions (3D), such as the 3D solid mechanics problems, we shall have $w = (w_1 \ w_2 \ w_3)$, where $w \in (\mathbb{G}_h^1)^3$. In this case we define, naturally, the associated inner product as

$$(w, v)_{\mathbb{G}^1(\Omega)} = \int_{\Omega} (w_1 v_1 + w_2 v_2 + w_3 v_3) \, d\Omega + \sum_{k=1}^{N_s} A_k^s \begin{bmatrix} \bar{g}_{11}(w_1)\bar{g}_{11}(v_1) + \bar{g}_{22}(w_2)\bar{g}_{22}(v_2) + \bar{g}_{33}(w_3)\bar{g}_{33}(v_3) \\ + (\bar{g}_{12}(w_1) + \bar{g}_{21}(w_2))(\bar{g}_{12}(v_1) + \bar{g}_{21}(v_2)) \\ + (\bar{g}_{13}(w_1) + \bar{g}_{31}(w_3))(\bar{g}_{13}(v_1) + \bar{g}_{31}(v_3)) \\ + (\bar{g}_{23}(w_2) + \bar{g}_{32}(w_3))(\bar{g}_{23}(v_2) + \bar{g}_{32}(v_3)) \end{bmatrix} \tag{32}$$

The associated $G^1(\Omega)$ semi-norm is defined as

$$|w|_{G^1(\Omega)}^2 = \sum_{k=1}^{N_s} A_k^s \begin{pmatrix} \bar{g}_{11}^2(w_1) + \bar{g}_{22}^2(w_2) + \bar{g}_{33}^2(w_3) \\ + (\bar{g}_{12}(w_1) + \bar{g}_{21}(w_2))^2 \\ + (\bar{g}_{13}(w_1) + \bar{g}_{31}(w_3))^2 \\ + (\bar{g}_{23}(w_2) + \bar{g}_{32}(w_3))^2 \end{pmatrix} \tag{33}$$

and the G^1 full norm becomes

$$\|w\|_{G^1(\Omega)}^2 = \int_{\Omega} (w_1^2 + w_2^2 + w_3^2) \, d\Omega + |w|_{G^1(\Omega)}^2 \tag{34}$$

We finally define a space for functions that are fixed on the Dirichlet boundaries and hence the functions cannot ‘float’

$$G_{h,0}^1 = \{v \in G_h^1(\Omega) \mid v_i = 0 \text{ on } \Gamma_u\} \tag{35}$$

The G spaces defined in this section are ‘unusual’ in two ways: first, we do not use derivatives of functions because we want to accommodate discontinuous functions that can be generated much easily in both the mesh-free or finite element settings and second, the Frobenius or trace norms are not used as induced matrix norms, and we intentionally define the inner product in such a way that the inner product induced norms are related to the L^2 norms of the vector of strains, which facilitate a smoother process in the later part of the derivation of some key inequalities. Let us now examine the G_h^1 space in comparison with the H_h^1 space.

3.3. The difference between G_h^1 and H_h^1 spaces

The major differences between a G_h^1 space and the corresponding Hilbert space or H_h^1 space are as follows:

- (1) The H_h^1 space requires the function and the first derivatives of the function all square integrable, but in the G_h^1 space we require only the function itself square integrable.
- (2) The requirement on function is now further weakened upon the already weakened requirement for functions in an H_h^1 space, and hence G_h^1 spaces can be viewed as spaces of functions with weakened weak (W²) continuity.
- (3) The first derivatives of functions in an H_h^1 space need to be bounded from above: $\int_{\Omega} (\partial v / \partial x_i)^2 \, d\Omega < \infty$. This is because the energy in the weak formulation needs to be bounded. We do not worry about the possibility of $\int_{\Omega} (\partial v / \partial x_i)^2 \, d\Omega = 0$, because it will never happen as long as the function is not zero everywhere (the well-known Poincare–Friedrichs inequality). On the other hand, for functions in a G_h^1 space, however, we need the ‘positivity’ to bound the functions from below: $\sum_{k=1}^{N_s} \left(\int_{\Gamma_k^s} v(s) n_i \, ds \right)^2 > 0$. This is because we need to ensure the stability of our W² formulations. The energy in the W² formulation is automatically bounded from above, because the function itself is square integrable, and we use only the function values to evaluate the energy.
- (4) Because a member in an H_h^1 space is a member in the corresponding G_h^1 space (using the same mesh with proper smoothing domains), we shall have $H_h^1 \subset G_h^1$. In addition, because

a function in a \mathbb{G}_h^1 space is also a member of the L^2 space; therefore, a \mathbb{G}_h^1 space is a subspace of L^2 space. We shall then have

$$\mathbb{H}_h^1(\Omega) \subset \mathbb{G}_h^1(\Omega) \subset \mathbb{L}^2(\Omega) \quad (36)$$

3.4. Basic properties

Because the \mathbb{G}^1 spaces are defined in the above-mentioned ‘unusual’ manner, we have to show that they possess all the necessary basic properties.

First, a \mathbb{G}_h^1 space is a *normed linear* space because it is clear that $\forall w_1, w_2 \in \mathbb{G}_h^1$, we have $(w_1 + w_2) \in \mathbb{G}_h^1$; and for $\forall w \in \mathbb{G}_h^1$ and $\forall \alpha \in \mathbb{R}$, we also have $\alpha w \in \mathbb{G}_h^1$.

Second, from the definition, we observe the *positivity*

$$\|w\|_{\mathbb{G}^1} > 0 \quad \forall w \in \mathbb{G}_h^1, w \neq 0 \quad (37)$$

and the *scalability*

$$\|\alpha w\|_{\mathbb{G}^1} = |\alpha| \|w\|_{\mathbb{G}^1} \quad \forall \alpha \in \mathbb{R} \quad \forall w \in \mathbb{G}_h^1 \quad (38)$$

Third, we have the *triangular inequality*

$$\|w + v\|_{\mathbb{G}^1} \leq \|w\|_{\mathbb{G}^1} + \|v\|_{\mathbb{G}^1} \quad \forall w \in \mathbb{G}_h^1 \quad \forall v \in \mathbb{G}_h^1 \quad (39)$$

which can be proven as follows. We first proof this for 2D scalar functions:

$$\begin{aligned} \|w + v\|_{\mathbb{G}^1} &= \left[\int_{\Omega} (w + v)^2 d\Omega + \sum_{k=1}^{N_s} A_k^s |\overline{\nabla}(w + v)|^2 \right]^{1/2} \\ &= \left[\int_{\Omega} w^2 d\Omega + \int_{\Omega} v^2 d\Omega + 2 \int_{\Omega} wv d\Omega \right. \\ &\quad \left. + \sum_{k=1}^{N_s} A_k^s ((\bar{g}_1(w) + \bar{g}_1(v))^2 + (\bar{g}_2(w) + \bar{g}_2(v))^2) \right]^{1/2} \\ &= \left[\int_{\Omega} w^2 d\Omega + \int_{\Omega} v^2 d\Omega + 2 \int_{\Omega} wv d\Omega \right. \\ &\quad \left. + \sum_{k=1}^{N_s} A_k^s (\bar{g}_1^2(w) + 2\bar{g}_1(w)\bar{g}_1(v) + \bar{g}_1^2(v) + \bar{g}_2^2(w) + 2\bar{g}_2(w)\bar{g}_2(v) + \bar{g}_2^2(v)) \right]^{1/2} \end{aligned} \quad (40)$$

$$\begin{aligned}
 & \left[\underbrace{\int_{\Omega} w^2 \, d\Omega + \sum_{k=1}^{N_s} A_k^s (\bar{g}_1^2(w) + \bar{g}_2^2(w))}_{\|w\|_{G^1}^2} + \underbrace{\int_{\Omega} v^2 \, d\Omega + \sum_{k=1}^{N_s} A_k^s (\bar{g}_1^2(v) + \bar{g}_2^2(v))}_{\|v\|_{G^1}^2} \right]^{1/2} \\
 = & \left[\underbrace{2 \left(\int_{\Omega} wv \, d\Omega + \sum_{k=1}^{N_s} A_k^s (\bar{g}_1(w)\bar{g}_1(v) + \bar{g}_2(w)\bar{g}_2(v)) \right)}_{(w,v)_{G^1} \leq \|w\|_{G^1} \|v\|_{G^1}} \right] \\
 \leq & [\|w\|_{G^1}^2 + 2\|w\|_{G^1} \|v\|_{G^1} + \|v\|_{G^1}^2]^{1/2} = \|w\|_{G^1} + \|v\|_{G^1} \quad \forall w \in \mathbb{G}_h^1 \quad \forall v \in \mathbb{G}_h^1 \quad (41)
 \end{aligned}$$

In the above proof process, we used the well-known Cauchy–Schwarz inequality for our inner product induced norms.

The exact same procedure can be applied to prove the triangular inequality for vector functions, but it will be a little lengthy. We prove it here for the 2D case, by examining first the semi-norm of the sum of two functions $w, v \in \mathbb{G}_h^1$ based on the definition equation (30):

$$\begin{aligned}
 |w+v|_{\mathbb{G}^1(\Omega)}^2 &= \sum_{k=1}^{N_s} A_k^s (\bar{g}_{11}^2(w_1+v_1) + \bar{g}_{22}^2(w_2+v_2) + (\bar{g}_{12}(w_1+v_1) + \bar{g}_{21}(w_2+v_2))^2) \\
 &= \sum_{k=1}^{N_s} A_k^s \left(\bar{g}_{11}^2(w_1) + \bar{g}_{11}^2(v_1) + 2\bar{g}_{11}(w_1)\bar{g}_{11}(v_1) + \bar{g}_{22}^2(w_2) + \bar{g}_{22}^2(v_2) \right. \\
 &\quad \left. + 2\bar{g}_{22}(w_2)\bar{g}_{22}(v_2) + ((\bar{g}_{12}(w_1) + \bar{g}_{21}(w_2)) + (\bar{g}_{12}(v_1) + \bar{g}_{21}(v_2)))^2 \right) \\
 &= \left(\underbrace{\sum_{k=1}^{N_s} A_k^s (\bar{g}_{11}^2(w_1) + \bar{g}_{22}^2(w_2) + (\bar{g}_{12}(w_1) + \bar{g}_{21}(w_2))^2)}_{|w|_{\mathbb{G}^1(\Omega)}^2} \right. \\
 &\quad \left. + \underbrace{\sum_{k=1}^{N_s} A_k^s (\bar{g}_{11}^2(v_1) + \bar{g}_{22}^2(v_2) + (\bar{g}_{12}(v_1) + \bar{g}_{21}(v_2))^2)}_{|v|_{\mathbb{G}^1(\Omega)}^2} \right. \\
 &\quad \left. + 2 \sum_{k=1}^{N_s} A_k^s \left(\bar{g}_{11}(w_1)\bar{g}_{11}(v_1) + \bar{g}_{22}(w_2)\bar{g}_{22}(v_2) \right. \right. \\
 &\quad \left. \left. + (\bar{g}_{12}(w_1) + \bar{g}_{21}(w_2))(\bar{g}_{12}(v_1) + \bar{g}_{21}(v_2)) \right) \right) \quad (42)
 \end{aligned}$$

We then examining the full norm of the sum of two functions $w, v \in \mathbb{G}_h^1$ based on the definition equation (31):

$$\begin{aligned}
 \|w+v\|_{\mathbb{G}^1(\Omega)}^2 &= \int_{\Omega} ((w_1^+ v_1)^2 + (w_2 + v_2)^2) \, d\Omega + |w+v|_{\mathbb{G}^1(\Omega)}^2 \\
 &= \int_{\Omega} (w_1^2 + v_1^2 + w_2^2 + v_2^2 + 2w_1 v_1 + 2w_2 v_2) \, d\Omega + |w+v|_{\mathbb{G}^1(\Omega)}^2 \\
 &= \int_{\Omega} (w_1^2 + v_1^2) \, d\Omega + \int_{\Omega} (w_2^2 + v_2^2) \, d\Omega + 2 \int_{\Omega} (w_1 v_1 + w_2 v_2) \, d\Omega + |w+v|_{\mathbb{G}^1(\Omega)}^2 \quad (43)
 \end{aligned}$$

Substituting Equation (42) into (43), gives

$$\begin{aligned} \|w+v\|_{\mathbb{G}^1}^2 &= \int_{\Omega} (w_1^2+v_1^2) \, d\Omega + \int_{\Omega} (w_2^2+v_2^2) \, d\Omega + 2 \int_{\Omega} (w_1 v_1 + w_2 v_2) \, d\Omega \\ &\quad + |w|_{\mathbb{G}^1}^2 + |v|_{\mathbb{G}^1}^2 + 2 \sum_{k=1}^{N_s} A_k^s \left(\begin{aligned} &+\bar{g}_{11}(w_1)\bar{g}_{11}(v_1) + \bar{g}_{22}(w_2)\bar{g}_{22}(v_2) \\ &+(\bar{g}_{12}(w_1) + \bar{g}_{21}(w_2))(\bar{g}_{12}(v_1) + \bar{g}_{21}(v_2)) \end{aligned} \right) \\ &= \|w\|_{\mathbb{G}^1}^2 + \|v\|_{\mathbb{G}^1}^2 + 2 \underbrace{\left(\int_{\Omega} (w_1 v_1 + w_2 v_2) \, d\Omega \right. \\ &\quad \left. + \sum_{k=1}^{N_s} A_k^s \left(\begin{aligned} &\bar{g}_{11}(w_1)\bar{g}_{11}(v_1) + \bar{g}_{22}(w_2)\bar{g}_{22}(v_2) \\ &+(\bar{g}_{12}(w_1) + \bar{g}_{21}(w_2))(\bar{g}_{12}(v_1) + \bar{g}_{21}(v_2)) \end{aligned} \right) \right)}_{(w,v) \leq \|w\|_{\mathbb{G}^1(\Omega)} \|v\|_{\mathbb{G}^1(\Omega)}} \\ &\leq \|w\|_{\mathbb{G}^1}^2 + \|v\|_{\mathbb{G}^1}^2 + 2\|w\|_{\mathbb{G}^1} \|v\|_{\mathbb{G}^1} = (\|w\|_{\mathbb{G}^1} + \|v\|_{\mathbb{G}^1})^2 \end{aligned} \tag{44}$$

which is Equation (39). Note here that we used again the Cauchy–Schwarz inequality.

Finally, comparing Equations (23) with (22), we obtain

$$|w|_{\mathbb{G}^1(\Omega)} \leq \|w\|_{\mathbb{G}^1(\Omega)} \quad \forall w \in \mathbb{G}_h^1 \tag{45}$$

meaning that the \mathbb{G}^1 full norm is always larger than the \mathbb{G}^1 semi-norm.

3.5. Convergence property for functions in H^1 space

Remark 3.1

Convergence property: For $w, v \in \mathbb{H}^1$, when $N_s \rightarrow \infty$ and all $\Omega_k^s \rightarrow 0$, \bar{W} becomes Delta functions and the integral representation becomes exact. At such a limit, we have $\overline{\nabla w} \rightarrow \nabla w$, $(w, v)_{\mathbb{G}^1(\Omega)} \rightarrow (w, v)_{\mathbb{H}^1(\Omega)}$, $\|w\|_{\mathbb{G}^1(\Omega)} \rightarrow \|w\|_{\mathbb{H}^1(\Omega)}$, $|w|_{\mathbb{G}^1(\Omega)} \rightarrow |w|_{\mathbb{H}^1(\Omega)}$, $\|w\|_{\mathbb{G}^2(\Omega)} \rightarrow \|w\|_{\mathbb{H}^2(\Omega)}$, and $|w|_{\mathbb{G}^2(\Omega)} \rightarrow |w|_{\mathbb{H}^2(\Omega)}$.

Remark 3.1 ensures that all the bound properties for \mathbb{G}_h^1 norms convergence to the corresponding \mathbb{H}_h^1 norms defined in the same manner at the limit of $N_s \rightarrow \infty$ and all $\Omega_k^s \rightarrow 0$ for all functions in the H space. In this work, however, we need the inequalities for finite smoothing domains and for all functions in \mathbb{G}_h^1 spaces, which are termed as G inequalities to be derived in the next sections.

3.6. First inequality for functions in H^1 space

The first inequality relates the (full) \mathbb{G}^1 norm of a function to L^2 norm of the nodal values of the function when the function is approximated based on an approximation method using local nodes scattered in the problem domain. For easy analysis and comprehension, we first consider a one-dimensional (1D) problem defined in Ω discretized using the linear finite elements, as shown in Figure 4. We will prove that a \mathbb{G}^1 space can be created using node-based smoothing domains. Over the 1D domain, we use N_e elements with $N_n (= N_e + 1)$ nodes. The standard linear finite element procedure is employed to formulate a set of linear shape (basis) functions that are then used to create a set of functions in an \mathbb{H}^1 space. In this case, the set of shape functions will be

clearly linearly independent, form a basis, and a function created using this set of nodal shape functions and the nodal values of the function are in an H¹ space, and can be expressed as

$$w(x) = \sum_{n \in S_e} \phi_n d_n \tag{46}$$

where d_n is the nodal function value at node n and ϕ_n is the linear shape function for node n , shown as straight lines in Figure 4.

Using Equation (23), the full G¹ norm of a function for 1D problems is given by

$$\|w\|_{G^1(\Omega)} = \left[\underbrace{\int_{\Omega} w^2 d\Omega}_{>0, \forall w \neq 0} + \underbrace{\sum_{k=1}^{N_s} A_k^s(\bar{g}^2(w))}_{\geq 0, \text{ for any } w \text{ and } \Omega_k^s} \right]^{1/2} \tag{47}$$

where the second term in the right-hand side (RHS) of the above equation will always be no-less than zero for any w regardless how the smoothing domains are created. The first term is always larger than zero, as long as w is not zero.

Since the set of FEM shape functions is linearly independent and hence can form the basis for the G¹ space. By the definition of Equation (3), a $w(x)$ created using Equation (46) (hence $\|w\|_{G^1(\Omega)}$) will never be zero everywhere in Ω unless all the nodal values d_n in Ω (hence $\|d\|_{L^2(\Omega)}$) are zero. This means that any nonzero $\|d\|_{L^2(\Omega)}$ will surely produce a nonzero positive $\|w\|_{G^1(\Omega)}$.

On the other hand, if $\|w\|_{G^1(\Omega)}$ is zero for a $w \in \tilde{S} \subset H^1$, from the norm definition equation (47), we immediately see that w must be zero everywhere in Ω , otherwise a nonzero positive value will be generated by the first term on LHS of Equation (47). Hence all the nodal displacements d_n in Ω (hence $\|d\|_{L^2(\Omega)}$) must be zero, if $\|w\|_{G^1(\Omega)}$ is zero. If $\|w\|_{G^1(\Omega)}$ is nonzero for a $w \in \tilde{S}$, w must be nonzero *somewhere* in Ω , and there will be at least a nonzero d_n (hence nonzero $\|d\|_{L^2(\Omega)}$).

Otherwise, if $\|w\|_{G^1(\Omega)}$ is nonzero for a $w \in \tilde{S}$ that is zero *everywhere* in Ω , $\sum_{k=1}^{N_s} A_k^s(\bar{w}_1^2)$ in Equation (47) must be nonzero. This can only happen when \bar{w}_1 (see Equation (21)) is nonzero for *some* smoothing domains, this is however contradicting to the fact that w is zero *everywhere* in Ω . Therefore, any nonzero $\|w\|_{G^1(\Omega)}$ will surely require a nonzero $\|d\|_{L^2(\Omega)}$. In summary, there must exist a nonzero positive constant c_{dw}^f , such that

$$\|d\|_{L^2(\Omega)} \geq c_{dw}^f \|w\|_{G^1(\Omega)} \quad \forall w \in \tilde{S} \tag{48}$$

or equivalently there must exist a nonzero positive constant c_{wd}^f , such that

$$\|w\|_{G^1(\Omega)} \geq c_{wd}^f \|d\|_{L^2(\Omega)} \quad \forall w \in \tilde{S} \tag{49}$$

Equations (48) or (49) is called, in this work, the first inequality for functions in a Hilbert space, which states that these norms $\|w\|_{G^1(\Omega)}$ and $\|d\|_{L^2(\Omega)}$ are equivalent, as long as w is created using linearly independent FEM shape functions and nodal values d_n in the form of Equation (46).

3.7. First inequality for functions in G¹ space

In the above analysis, it is clear that we always have the inequality in Equation (48), as long as these nodal shape functions are linearly independent, because this is only the condition that we

used in the above derivation. Therefore, mesh-free shape functions such as the ones created using the point interpolation procedure (PIM and RPIM) [5, 6] can also be used, and a function can be created in the form of Equation (6). In such a situation, the nodes used in the shape function creation are not confined within the cell/element, and is in fact usually beyond the cell/element. Therefore, the w generated is generally not continuous and hence not belongs to an \mathbb{H}_h^1 space, but a \mathbb{G}_h^1 space. In Reference [6], typical examples of discontinuous functions generated using the PIM and RPIM shape functions are plotted and analyzed in detail in Figure 5.47 for 1D cases, and Figures 8.1 and 8.2 for 2D cases. In our analysis here, however, the discontinuity is irrelevant, because for the inequality in Equations (48) or (49) to hold, all we need is the linearly independence of these shape functions. We thus immediately have

$$\|d\|_{L^2(\Omega)} \geq c_{dw}^f \|w\|_{\mathbb{G}^1(\Omega)} \quad \forall w \in \mathbb{G}_h^1 \tag{50}$$

or equivalently

$$\|w\|_{\mathbb{G}^1(\Omega)} \geq c_{wd}^f \|d\|_{L^2(\Omega)} \quad \forall w \in \mathbb{G}_h^1 \tag{51}$$

which is called the first inequality for functions in a G space. We now record the following remark.

Remark 3.2

Functions in a \mathbb{G}_h^1 space satisfy the first inequality equations (50) or (51): the full G norm of a function in a \mathbb{G}_h^1 space is equivalent to the L^2 norm of the nodal values of the function.

3.8. Extension to higher dimensions and vector functions

It is clear that all these arguments used in the above discussion also applicable to scalar fields of any higher dimensions, because one node still carries one nodal unknown, and the field function is created using these nodal values and shape functions. Therefore, we conclude that Remark 3.2 is valid for 2D and 3D cases. For vector functions, the same arguments also hold, but we need linearly independent shape functions, respectively, for each of the component functions w_1, w_2 and maybe w_3 .

3.9. Second inequality

The second inequality relates the \mathbb{G}^1 semi-norm of a function to the L^2 norm of the nodal values of the function in a \mathbb{G}^1 with a set of smoothing domains created for evaluating the \mathbb{G}^1 semi-norm. Again, for easy analysis and comprehension, we first consider a 1D problem defined in Ω discretized using the finite elements. We now further construct a set of node-based smoothing domains: one smoothing domain for each of these nodes is created with two boundary points located at the centers of the two neighboring elements, as shown in Figure 4.

3.9.1. Positivity relay. Using Equation (22), the \mathbb{G}^1 semi-norm for 1D problems can be written as

$$|w|_{\mathbb{G}^1(\Omega)} = \left[\sum_{k=1}^{N_s} A_k^s \bar{g}^2(w) \right]^{1/2} = \left[\sum_{k=1}^{N_s} \frac{1}{A_k^s} (w(x_{k+1/2})n(x_{k+1/2}) + w(x_{k-1/2})n(x_{k-1/2}))^2 \right]^{1/2} \tag{52}$$

where $w \in \tilde{\mathbf{S}}_0 \subset \mathbb{H}_0^1$, $x_{k+1/2}$ and $x_{k-1/2}$ denote the x coordinates at, respectively, the right and left boundary points of the k th smoothing domain (at the middles of the two neighboring elements). For any nonzero function $w \in \tilde{\mathbf{S}}$ with at least one $d_n \neq 0$, the semi-norm is evaluated as

$$|w|_{\mathbb{G}^1(\Omega)} = \left[\begin{aligned} & \frac{1}{A_1^s} (w(x_{1+1/2}) \times 1 + w(x_1) \times (-1))^2 \\ & + \frac{1}{A_2^s} (w(x_{2+1/2}) \times 1 + w(x_{1+1/2}) \times (-1))^2 + \dots \\ & + \frac{1}{A_k^s} (w(x_{k+1/2}) \times 1 + w(x_{k-1/2}) \times (-1))^2 + \dots \\ & + \frac{1}{A_{N_n}^s} (w(x_{N_n}) \times 1 + w(x_{N_n-1/2}) \times (-1))^2 \end{aligned} \right]^{1/2} \tag{53}$$

Using Equation (46), we obtain

$$|w|_{\mathbb{G}^1(\Omega)} = \left[\begin{aligned} & \frac{1}{A_1^s} \left(\underbrace{\phi_1^{1+1/2}}_{\neq 0} \underbrace{d_1}_{=0, BC} + \underbrace{\phi_2^{2-1/2}}_{\neq 0} d_2 - \underbrace{\phi_1^1}_{=1} \underbrace{d_1}_{=0, BC} \right)^2 \\ & + \frac{1}{A_2^s} \left(\phi_2^{2+1/2} d_2 + \phi_3^{3-1/2} d_3 - \phi_1^{1+1/2} \underbrace{d_1}_{=0, BC} - \phi_2^{2-1/2} d_2 \right)^2 + \dots \\ & + \frac{1}{A_k^s} (\phi_k^{k+1/2} d_s + \phi_{k+1}^{k+1/2} d_{s+1} - \phi_{k-1}^{k-1/2} d_{k-1} - \phi_k^{k-1/2} d_k)^2 + \dots \\ & + \frac{1}{A_{N_n}^s} (\phi_{N_n}^{N_n} d_{N_n} - \phi_{N_n-1}^{N_n-1/2} d_{N_n-1} - \phi_{N_n}^{N_n-1/2} d_{N_n})^2 \end{aligned} \right]^{1/2} \tag{54}$$

where $\phi_i^j = \phi_i(x_j)$. Invoking the Dirichlet boundary condition at x_1 , if $d_2 \neq 0$, we have

$$|w|_{\mathbb{G}^1(\Omega)} \geq c_{wd2} d_2^2 \quad \forall w \in \tilde{\mathbf{S}} \quad \text{and} \quad d_2 \neq 0 \tag{55}$$

where $c_{wd2} = 1/(2A_1) > 0$. If $d_2 = 0$ but $d_3 \neq 0$, we then have

$$|w|_{\mathbb{G}^1(\Omega)} \geq \underbrace{c_{wd3}}_{=1/(2A_2^s) > 0} d_3^2 \quad \forall w \in \tilde{\mathbf{S}} \quad \text{with} \quad d_2 = 0 \quad \text{and} \quad d_3 \neq 0 \tag{56}$$

We see clearly a ‘positivity relay’ initiated by the Dirichlet boundary condition at the left most point of the 1D domain, as shown in Figure 4. When this positivity relay continues, and we shall have at the end

$$|w|_{\mathbb{G}^1(\Omega)} \geq \underbrace{c_{wdN_n}}_{=1/(2A_{N_n-1}^s) > 0} d_{N_n}^2 \quad \forall w \in \tilde{\mathbf{S}} \quad \text{with} \quad d_2 = d_3 = \dots = d_{N_n-1} = 0 \quad \text{and} \quad d_{N_n} \neq 0 \tag{57}$$

Let

$$c_{wd}^{\min} = \min(c_{wd2}, c_{wd3}, \dots, c_{wdN_n}) \tag{58}$$

We then obtain

$$|w|_{\mathbb{G}^1(\Omega)} \geq \underbrace{\frac{c_{wd}^{\min}}{N_n - 1}}_{c_{wd}^s} (\underbrace{d_1^2}_{=0} + d_2^2 + d_3^2 + \dots + d_{N_n}^2) \geq c_{wd} \|d\|_{\mathbb{L}^2(\Omega)} \quad \forall w \in \tilde{\mathbf{S}} \tag{59}$$

Which means that we can always find a nonzero positive c_{wd}^s , such that

$$|w|_{\mathbb{G}^1(\Omega)} \geq c_{wd}^s \|d\|_{\mathbb{L}^2(\Omega)} \quad \forall w \in \tilde{\mathbf{S}} \tag{60}$$

or equivalently, there always exist a nonzero positive c_{dw}^s , such that

$$|w|_{\mathbb{G}^1(\Omega)} \geq c_{dw}^s \|d\|_{\mathbb{L}^2(\Omega)} \quad \forall w \in \tilde{\mathbf{S}} \tag{61}$$

This equivalence is clearly a result from the above process of positivity relay. This equivalence of the semi-norm $|w|_{\mathbb{G}^1(\Omega)}$ and the norm $\|d\|_{\mathbb{L}^2(\Omega)}$ is clearly a result from the above process of positivity relay. Such a positivity relay always occurs because the particular way of partitioning the problem domain into smoothing domains and the way that the semi-norm is defined: it sums up the squared differences of the value of w at two boundary points of the smoothing domain that contains a node. When the value of w at any one of the boundary points (it does not have to be on the left most boundary of the problem domain) vanishes, it will trigger a positivity relay leading to the second inequality in Equations (60) or (61).

3.9.2. *Independence of rows in the smoothed gradient matrix.* The equivalence of the semi-norm $|w|_{\mathbb{G}^1(\Omega)}$ and the L^2 norm $\|d\|_{\mathbb{L}^2(\Omega)}$ can also be derived from another point of view. From Equation (54), we see that for the \mathbb{G}^1 semi-norm of w to vanish, all the following equations have to be satisfied for the smoothing domains:

$$\begin{aligned} \text{For } \Omega_1^s: & \frac{1}{A_1^s} \phi_2^{1+1/2} d_2 = 0 \\ \text{For } \Omega_2^s: & \frac{1}{A_2^s} (\phi_2^{2+1/2} d_2 + \phi_3^{3-1/2} d_3) = 0 \\ & \vdots \\ \text{For } \Omega_k^s: & \frac{1}{A_k^s} (\phi_k^{k+1/2} d_k + \phi_{k+1}^{(k+1)-1/2} d_{k+1}) = 0 \\ & \vdots \\ \text{For } \Omega_{N_n}^s: & \frac{1}{A_{N_n}^s} \phi_{N_n}^{N_n} d_{N_n} = 0 \end{aligned} \tag{62}$$

or in matrix form

$$\begin{array}{l}
 \Omega_1^s \rightarrow \\
 \Omega_2^s \rightarrow \\
 \vdots \\
 \Omega_k^s \rightarrow \\
 \vdots \\
 \Omega_{N_n}^s \rightarrow
 \end{array}
 \left[\begin{array}{cccccc}
 \frac{\phi_2^{1+1/2}}{A_2^s} & 0 & 0 & 0 & 0 & 0 \\
 \frac{\phi_2^{2+1/2}}{A_2^s} & \frac{\phi_3^{2+1/2}}{A_2^s} & 0 & \dots & \dots & 0 \\
 \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
 \vdots & \vdots & \frac{\phi_k^{k+1/2}}{A_k^s} & \frac{\phi_{k+1}^{k+1/2}}{A_k^s} & \vdots & \vdots \\
 \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
 0 & 0 & \dots & \dots & 0 & \frac{\phi_{N_n}^{N_n}}{A_{N_n}^s}
 \end{array} \right]
 \underbrace{\left[\begin{array}{c} d_2 \\ d_3 \\ \vdots \\ d_k \\ \vdots \\ d_{N_n} \end{array} \right]}_{\mathbf{d}} = \mathbf{B}_{N_n \times (N_n-1)} \mathbf{d} = \mathbf{0} \quad (63)$$

B

where **B** is the smoothed gradient matrix for node-based smoothing domains. It is a tall–slim matrix for this set of node-based smoothing domains, and the equation system given Equation (63) is an ‘over-determined’ system. Simple observation (perform simple Gauss elimination) reveals that the first $(N_n - 1)$ rows in **B** are clearly linearly independent. In fact, because each row in **B** represents an interpolation at a point in a particular element using the shape functions that are linearly independent, there must be $(N_n - 1)$ rows in **B** being linearly independent, and hence any choice of $(N_n - 1)$ rows from **B** will be linearly independent. Therefore, the only possibility for Equation (63) to be satisfied is $\mathbf{d} = \mathbf{0}$. In other words, any nonzero \mathbf{d} will surely produce a nonzero positive $|w|_{G^1(\Omega)}$. On the other hand, if $|w|_{G^1(\Omega)}$ is nonzero, there must be at least one equation in Equation (62) is nonzero, hence there is at least one nonzero entry in the RHS of Equation (63). Choosing $(N_n - 1)$ rows from Equation (63) with at least one row corresponding to the nonzero entry in the RHS of Equation (63), we can then form the following equation:

$$\mathbf{B}_r \mathbf{d} = \underbrace{\mathbf{b}}_{\neq 0} \quad (64)$$

Because the rows in the reduced matrix \mathbf{B}_r are clearly linearly independent, thus \mathbf{B}_r is invertible, and Equation (64) can be solved for a nonzero \mathbf{d} resulting in a nonzero $\|d\|_{L^2(\Omega)}$. Therefore, the semi-norms $|w|_{G^1(\Omega)}$ and norm $\|d\|_{L^2(\Omega)}$ are equivalent, and hence Equations (60) and (61) hold.

3.9.3. *Types of smoothing domains.* Next, we examine other types of smoothing domains. Because the first inequality holds for any types of smoothing domain, we only need to examine the second inequality for different types of smoothing domains.

For 1D problems with linear FEM shape functions, the simplest alternative to the already discussed node-based smoothing domains is the cell-based smoothing domains: we simply use an

element as a smoothing domain. In this case, Equation (63) becomes

$$\underbrace{\begin{bmatrix} \frac{\phi_2^2}{A_1} & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{\phi_3^3}{A_2} & 0 & \dots & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \frac{\phi_{k+1}^{k+1}}{A_k^s} & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \dots & 0 & \frac{\phi_{N_n}^{N_n}}{A_{N_n}^s} \end{bmatrix}}_{\mathbf{B}} \underbrace{\begin{Bmatrix} d_2 \\ d_3 \\ \vdots \\ d_{N_n} \end{Bmatrix}}_{\mathbf{d}} = \mathbf{B}_{(N_n-1) \times (N_n-1)} \mathbf{d} = \mathbf{0} \tag{65}$$

where the smoothed gradient matrix \mathbf{B} that is a diagonal matrix, obviously invertible, and hence the semi-norms $|w|_{\mathbb{G}^1(\Omega)}$ and norm $\|d\|_{\mathbb{L}^2(\Omega)}$ must be equivalent. Therefore, Equations (60) and (61) hold. This is no surprise, because when we use cell (element)-based smoothing domain and linear shape functions, the model becomes exactly the FEM model [19, 26]. Essentially what is happening here is that the linearly independent set of shape functions provides a basis for creating a set of $(N_n - 1)$ independent function values, and the element-based smoothing domains ‘isolate’ and ‘lock-in’ these values, and the semi-norm squares these values individually ensuring the positivity in the semi-norm.

When higher order of elements (hence higher-order shape functions) is used, the use of one whole element as one smoothing domain will not work. Because in such a case, the smoothed gradient matrix \mathbf{B} becomes a short-fat matrix (not enough ‘isolations’), and the equations system given in Equation (63) becomes ‘under determined’, although these shape functions are still linearly independent. Hence, there will be a plenty of nonzero \mathbf{d} that satisfies Equation (63) to produce zero semi-norm, and hence we cannot proof the equivalency of the semi-norms $|w|_{\mathbb{G}^1(\Omega)}$ and norm $\|d\|_{\mathbb{L}^2(\Omega)}$. Therefore, when higher order of elements is used, we have to use more smoothing domains than the number of the elements. The use of node-based smoothing domains will surely work (with enough ‘isolations’) as discussed earlier. We can also use two or more smoothing domains in each element (to create more ‘isolations’), as in the SFEM [9–11]. The question is now how should we form the smoothing domains and how many smoothing domains we have to use.

3.9.4. Linearly independent smoothing domains. Linearly independent smoothing domain is defined as a smoothing domain whose row in Equation (63) is linearly independent of the rest of the rows in the smoothed gradient matrix \mathbf{B} for other smoothing domains. For 1D problems with linear FEM shape functions, both the node-based and cell-based smoothing domains are linearly independent.

The use of more smoothing domains will leads to a tall–slim \mathbf{B} matrix as seen in Equation (63), but it does not necessarily guarantee a sufficient number of independent rows in \mathbf{B} . This is because

the division of the smoothing domains will affect the independency of the rows in the \mathbf{B} matrix. For example, in the node-based smoothing domain case, one may further divide some of the node-based smoothing domains into more and smaller smoothing domains. This will produce more equations in Equation (63), but some of them will be linearly dependent. Therefore, the division of smoothing domains should be performed in relation to the node distribution or the background cells. For example, we should not have any smoothing domain that includes two nodes or more. Typical proven divisions that give sufficient number of independent smoothing domains are cell/element based as in the smoothed FEM or SFEM [7, 10] and CS-PIM [21], node based as in the NS-PIM [19, 26] and NS-FEM [27], as well as the edge/face based as in the ES-FEM [28, 29].

3.9.5. A discussion on soft and stiff modes: upper and lower bounds. When the number of linearly independent smoothing domains satisfies the minimum number given in [22], the use of more (hence finer division) smoothing domains will not change the fact of *positivity*, but will affect the *degree* of the positivity, and it generally leads to a ‘stiffer’ model. When the number of smoothing domains becomes infinite and the size of *all* the smoothing domains becomes infinitely small, we arrived at an FEM model (see Remark 3.1 and [10, 11]). Because a fully compatible FEM model is known overly stiff, and the solution will be lower bound to the exact solution in energy norm (see,), and hence we do not have any reason to make an even stiffer model.

We do, however, have many good reasons to make softer models, and one of which reasons is to obtain an upper bound solution. To produce a softer model, Liu discovered that we can simply create a model that uses a smaller number (but larger than the minimum number) of smoothing domains, as it is done in the NS-PIM [19, 26] and in NS-RPIM [33]. Theoretically, one can make a model as softer as desired and can even make a model that is singular! This implies that it is always possible to make a model soft enough to produce upper bound solution known as the existence theory. However, when the model gets too soft, there can be nonzero energy modes that can lead to temporal instability and the model cannot be used for dynamics problems, because a certain amount of kinetic energy can excite these modes and produce erroneous solutions. When we want to remove spurious modes and to have temporal stability, we should use more smoothing domains, as we do in the ES-FEM [28] and SFEM [10, 11]. The ES-FEM model was found to possess very ‘close-to-exact’ stiffness [28].

3.9.6. A discussion on hourglass modes. It is well known that there can be spatially instable modes called hourglass modes when we use reduced integration in quadrilateral elements in FEM. The hourglass modes are zero-energy nodes and hence are spatially instable (any amount of inputs can create instability in the numerical solution), and hence is different from the nonzero energy spurious modes (mentioned in Section 3.9.5) that are temporally instable (you need a certain amount of kinetic energy to excite these modes). There have been many discussions on the hourglass modes in the open literature. It is known that an hourglass model appears when one Gauss sampling point per quadrilateral element is used, which is quite similar (in term of stability behavior) to our case of using cell-based smoothing domain with one smoothing domain for each of the elements. When one smoothing domain per element is used, the equation system can be singular, because the minimum integration points of $2n_t/3$ may not be satisfied, and hence the positivity of the semi-norm can be lost [22]. For example, when one quadrilateral element is used for the entire problem domain with two nodes fixed, we shall have $8 - 4 = 4$ total unknowns and need a minimum of $\frac{4}{3}$ smoothing domains. If only one smoothing domain for this element is used, the equation system will be singular. To prevent this kind of situations from happening, we need to use more

smoothing domains in at least some of the elements in a model. A detailed discussion on this scenario can be found in [10] where cell-based smoothing models are established and examined in the FEM framework. When the node-based smoothing domains are used, however, such hourglass modes will not appear, because the number of smoothing domains becomes now n_t that is always larger the minimum integration points of $2n_t/3$. A detailed discussion on this analysis can be found in [27] where node-based smoothing models are established and examined in the FEM framework.

From the above analysis, we conclude that for any finite model as long as a minimum number of linearly independent smoothing domains are used, functions created using any set of linearly independent nodal shape functions will satisfy the inequalities equations (60) and (61). Therefore, mesh-free shape functions such as the ones created using the point interpolation procedure (PIM and RPIM) [5] can also be used, because the discontinuity is irrelevant. We now present the following remark.

Remark 3.3

If at least a minimum number of linearly independent smoothing domains are used, we should have the second inequality

$$|w|_{\mathbb{G}^1(\Omega)} \geq c_{wd}^s \|d\|_{L^2(\Omega)} \quad \forall w \in \mathbb{G}_{h,0}^1(\Omega) \quad (66)$$

or equivalently

$$\|d\|_{L^2(\Omega)} \geq c_{dw}^s |w|_{\mathbb{G}^1(\Omega)} \quad \forall w \in \mathbb{G}_{h,0}^1(\Omega) \quad (67)$$

meaning that the full \mathbb{G}^1 semi-norm of a function in a \mathbb{G} space is equivalent to the L^2 norm of the nodal values of the function.

Based on the same argument presented in Section 3.8, we conclude that the second inequality also holds for 2D and 3D cases. All we need is to use the smoothing domains with minimum number of linearly independent smoothing domains and a set of shape functions that are linearly independent. In other words, a set of linearly independent nodal shape functions (created using FEM and/or mesh-free settings) can form functions in a normed \mathbb{G}^1 space, as long as a set of minimum number of linearly independent smoothing domains are used to evaluate the semi-norms. In actual practice, we have successfully used node-based, edge-based, and cell-based smoothing domains, in both FEM and mesh-free settings, as discussed in [22].

3.10. Third inequality

We now ready to present the third inequality stated in the following theorem.

Theorem 3.1

Equivalence of G norms: When atleast a minimum number of independent node-based smoothing domains are used to evaluate the \mathbb{G}^1 norms, there exists a positive nonzero constant c_G such that

$$c_G \|w\|_{\mathbb{G}^1(\Omega)} \leq |w|_{\mathbb{G}^1(\Omega)} \quad \forall w \in \mathbb{G}_{h,0}^1 \quad (68)$$

meaning that the \mathbb{G}^1 full norm and the \mathbb{G}^1 semi-norm of any function in a $\mathbb{G}_{h,0}^1$ space are equivalent.

Proof

The combination of the inequality equation (50), and the inequality equation (66) gives

$$|w|_{\mathbb{G}^1(\Omega)} \geq c_{wd}^s \|d\|_{\mathbb{L}^2(\Omega)} \geq \underbrace{c_{wd}^s c_{dw}^f}_{c_G} \|w\|_{\mathbb{G}^1(\Omega)} \geq c_G \|w\|_{\mathbb{G}^1(\Omega)} \quad \forall w \in \mathbb{G}_{h,0}^1(\Omega) \tag{69}$$

which is the third inequality. □

Combining Equations (45) and (69), we arrived at the following chain inequalities:

$$c_G \|w\|_{\mathbb{G}^1(\Omega)} \leq |w|_{\mathbb{G}^1(\Omega)} \leq \|w\|_{\mathbb{G}^1(\Omega)} \quad \forall w \in \mathbb{G}_{h,0}^1 \tag{70}$$

The third inequality equation (68) is a generalized version of the well-known Poincare–Friedrichs inequality. It is the foundation of the W² formulation, ensuring the stability of the solution. Equation (70) is essential to ensure both the uniqueness and convergence of a W² formulation of a *physically* stable problem. For solid mechanic problems, for example, we need the material being stable (see, part II for definition).

3.11. Softening effects

We further examine some of the important properties of functions in G spaces.

Remark 3.4

For a function in an H¹ space, the G¹ semi-norm of the function is no larger than the H¹ semi-norm (of same type) of the function

$$|w|_{\mathbb{G}^1(\Omega)} \leq |w|_{\mathbb{H}^1(\Omega)} \quad \forall w \in \mathbb{H}^1 \tag{71}$$

meaning that the smoothing operation results in a smaller semi-norm measure. This is the fundamental inequality of the so-called softening effects [26].

Proof

From Equation (14), the H¹ semi-norm is defined by

$$|w|_{\mathbb{H}^1(\Omega)}^2 = \int_{\Omega} (\nabla w) \cdot (\nabla w) \, d\Omega = \sum_{k=1}^{N_s} \int_{\Omega_k^s} \left(\frac{\partial w}{\partial x_1} \frac{\partial w}{\partial x_1} + \frac{\partial w}{\partial x_2} \frac{\partial w}{\partial x_2} \right) \, d\Omega \tag{72}$$

The summation is possible because *w* is in an H space. For such a *w*, the G¹ semi-norm can be written as follows using Equations (22) and (21):

$$|w|_{\mathbb{G}^1(\Omega)}^2 = \sum_{k=1}^{N_s} A_k^s |\overline{\nabla w}|^2 = \sum_{k=1}^{N_s} A_k^s \left(\overline{\frac{\partial w}{\partial x_1} \frac{\partial w}{\partial x_1}} + \overline{\frac{\partial w}{\partial x_2} \frac{\partial w}{\partial x_2}} \right) \tag{73}$$

To prove Equation (71), we need only to prove

$$A_k^s \overline{\frac{\partial w}{\partial x_i} \frac{\partial w}{\partial x_i}} \leq \int_{\Omega_k^s} \frac{\partial w}{\partial x_i} \frac{\partial w}{\partial x_i} \, d\Omega \quad \forall w \in \mathbb{H}^1 \tag{74}$$

where $i = 1$ or 2 . Because w is in an H space, the first equation in Equation (10) is applicable, and hence Equation (74) becomes

$$\left(\int_{\Omega_k^s} \frac{1}{A_k^s} \frac{\partial w}{\partial x_i} \right) d\Omega \left(\int_{\Omega_k^s} \frac{1}{A_k^s} \frac{\partial w}{\partial x_i} d\Omega \right) \leq \frac{1}{A_k^s} \int_{\Omega_k^s} \frac{\partial w}{\partial x_i} \frac{\partial w}{\partial x_i} d\Omega \quad \forall w \in \mathbb{H}^1 \tag{75}$$

or

$$\left(\int_{\Omega_k^s} \frac{1}{A_k^s} \frac{\partial w}{\partial x_i} d\Omega \right)^2 \leq \frac{1}{A_k^s} \int_{\Omega_k^s} \left(\frac{\partial w}{\partial x_i} \right)^2 d\Omega \quad \forall w \in \mathbb{H}^1 \tag{76}$$

Dividing Ω_k^s into n_q subdomains of equal areas $\Omega_k^s = \cup_{q=1}^{n_q} \Omega_{k,q}^s$, we then have

$$\frac{1}{A_k^s} \int_{\Omega_k^s} \left(\frac{\partial w}{\partial x_i} \right)^2 d\Omega = \lim_{\substack{n_q \rightarrow \infty \\ \Omega_{k,q}^s \rightarrow 0}} \left(\frac{1}{A_k^s} \sum_{q=1}^{n_q} A_{k,q}^s \left(\frac{\partial w}{\partial x_i} \Big|_q \right)^2 \right) \tag{77}$$

Using the inequality of arithmetic and geometric means (AM-GM inequality), which states that the AM of a list of non-negative real numbers is greater than or equal to the GM of the same list, we arrive at

$$\begin{aligned} \frac{1}{A_k^s} \int_{\Omega_k^s} \left(\frac{\partial w}{\partial x_i} \right)^2 d\Omega &= \lim_{\substack{n_q \rightarrow \infty \\ \Omega_{k,q}^s \rightarrow 0}} \left(\frac{1}{A_k^s} \sum_{q=1}^{n_q} \underbrace{\frac{A_{k,q}^s}{A_k^s}}_{\frac{1}{n_q}} \underbrace{\left(\frac{\partial w}{\partial x_i} \Big|_q \right)^2}_{\geq 0} \right) \geq \lim_{\substack{n_q \rightarrow \infty \\ \Omega_{k,q}^s \rightarrow 0}} \left(\frac{1}{A_k^s} \sum_{q=1}^{n_q} \frac{A_{k,q}^s}{A_k^s} \frac{\partial w}{\partial x_i} \Big|_q \right)^2 \\ &= \lim_{\substack{n_q \rightarrow \infty \\ \Omega_{k,q}^s \rightarrow 0}} \left(\frac{1}{(A_k^s)^2} \left(\sum_{q=1}^{n_q} A_{k,q}^s \frac{\partial w}{\partial x_i} \Big|_q \right)^2 \right) = \left(\frac{1}{A_k^s} \int_{\Omega_k^s} \frac{\partial w}{\partial x_i} d\Omega \right)^2 \end{aligned} \tag{78}$$

which is the inequality in Equation (76) leading to Equation (74) and then Equation (71). □

Because of Equation (71), we immediately have

Remark 3.5

For a function in an H^1 space, the G^1 full norm of the function is no larger than the H^1 full norm of the function

$$\|w\|_{G^1(\Omega)} \leq \|w\|_{H^1(\Omega)} \quad \forall w \in \mathbb{H}^1 \tag{79}$$

meaning that the smoothing operation results in a smaller (full) norm measure. This is because the first term in the RHS of Equation (12) and that of Equation (23) are exactly the same.

4. AN ERROR ESTIMATION

4.1. Error in G¹ norm measure

In this section we only examine the error of point interpolation used to approximate a ‘target’ function, which is related to the errors in a W² numerical model. We need to find out the *h*-dependence of the error in both G_{*h*}¹ and L² norms in relation to the ‘strong’ norm: the infinity norm of the target function to be approximated by interpolation. For simplicity, we only consider linear interpolation and the target function should be smoother, and hence the error norms will be bounded by the infinity norm of the second derivative of the target function. To establish such bound relations, we need to know exactly the relationship between the element/cell mesh and smoothing domains. Figure 4 shows clearly such a relationship for our 1D problem. We first define the interpolation error for the given target function *w* as

$$e^I(x) = w(x) - \mathcal{I}_h w(x) \quad \forall x \in T_q \tag{80}$$

where *T_q* is defined in Section 2.1. We then state

Theorem 4.1

If the target function $w \in \mathbb{G}_{h,0}^1$ and $w|_{T_q} \in \mathbb{C}^2(T_q)$, $q = 1, \dots, N_e$ where $w|_{T_q}$ means *w* restricted within *T_q*, the linear interpolation error in G¹ norm satisfies, in general:

$$|e^I|_{\mathbb{G}^1(\Omega)} = |w - \mathcal{I}_h w|_{\mathbb{G}^1(\Omega)} \leq h_{\max} \frac{3c_{rh}}{4} \left(\max_{q=1, \dots, N_e} \max_{\xi \in T_q} |w'(\xi)| \right) \tag{81}$$

where $c_{rh} = h_{\max}/h_{\min}$, and the error in L² norm satisfies

$$|e^I|_{\mathbb{L}^2(\Omega)} = |w - \mathcal{I}_h w|_{\mathbb{L}^2(\Omega)} \leq h_{\max}^2 \left(\max_{q=1, \dots, N_e} \max_{\xi \in T_q} |w''(\xi)| \right) \tag{82}$$

In particular, when uniform mesh is used, and when the second derivative of *w* is constant in the cells sharing the smoothing domains we further have

$$|e^I|_{\mathbb{G}^1(\Omega)} = |w - \mathcal{I}_h w|_{\mathbb{G}^1(\Omega)} = h^{1.5} \frac{1}{4} |w''_{\max}| \tag{83}$$

and the error in L² norm satisfies

$$|e^I|_{\mathbb{L}^2(\Omega)} = |w - \mathcal{I}_h w|_{\mathbb{L}^2(\Omega)} = \frac{h^2}{\sqrt{120}} |w''_{\max}|^2 \tag{84}$$

Proof

Consider a target function $w \in \mathbb{G}_{h,0}^1$ and $w|_{T_q} \in \mathbb{C}^2(T_q)$, $q = 1, \dots, N_e$. Using Taylor’s expansion with respect to *x_k* (see Figure 4), there are exist a $\xi \in T_q$ such that

$$w(x) = w(x_k) + w'(x_k)(x - x_k) + \frac{1}{2} w''(\xi(x))(x - x_k)^2 \quad \forall x \in T_q \tag{85}$$

In Equation (85), we note the fact that ξ is in fact dependent on *x*. We also note that we do not require the existence of *w''* on the boundary of *T_q*: meaning that the first derivative of *w* can

‘jump’ there. Using Equation (8), the linear interpolant has to pass through the two nodes at x_k and x_{k+1} , and hence should be given as

$$\mathcal{I}_h w(x) = w(x_k) + [w'(x_k) + \frac{1}{2}w''(\xi(x_{k+1}))h_k](x - x_k) \quad \forall x \in T_q \tag{86}$$

Based on the definition equation (21), for our 1D problem and the k th smoothing cell, we have

$$\bar{g}(w_k) = \frac{\overline{\partial w}}{\partial x} = \frac{1}{A_k^s} \int_{\Gamma_k^s} w(s)n_1 ds = \frac{1}{A_k^s} (w_{k+1/2} - w_{k-1/2}) = \frac{2}{h_k + h_{k-1}} (w_{k+1/2} - w_{k-1/2}) \tag{87}$$

where $w_k = w(x_k)$, and $h_0 = h_{N_n} = 0$. Substituting Equations (85) and (86) into (87), we then have for the k th smoothing cell:

$$\begin{aligned} \bar{g}(\underbrace{w_k - \mathcal{I}_h w_k}_{e^I(x_k)}) &= \frac{2}{h_k + h_{k-1}} \left\{ \begin{aligned} &w''(\xi(x_{k+1/2}))(x_{k+1/2} - x_k)^2 - \frac{1}{2}w''(\xi(x_{k+1}))h_k(x_{k+1/2} - x_k) \\ &-\frac{1}{2}w''(\xi(x_{k-1/2}))(x_{k-1/2} - x_{k-1})^2 \\ &+\frac{1}{2}w''(\xi(x_k))h_{k-1}(x_{k-1/2} - x_{k-1}) \end{aligned} \right\} \\ &= \frac{2}{h_k + h_{k-1}} \left\{ \begin{aligned} &\frac{h_k^2}{4} \left[\frac{1}{2}w''(\xi(x_{k+1/2})) - w''(\xi(x_{k+1})) \right] \\ &-\frac{h_{k-1}^2}{4} \left[\frac{1}{2}w''(\xi(x_{k-1/2})) - w''(\xi(x_k)) \right] \end{aligned} \right\} \tag{88} \end{aligned}$$

Let

$$w''_{\max} = \max_{i=1, \dots, N_e} \max_{\xi \in T_q} |w''(\xi)| \tag{89}$$

and hence we shall have $w''_{\max} \neq 0$, otherwise the second derivative of w will be zero everywhere in the problem domain, and the interpolation will be exact: no need for error estimation. Using Equations (1) and (2), Equation (88) becomes

$$\begin{aligned} \left| \bar{g}(\underbrace{w_k - \mathcal{I}_h w_k}_{e^I(x_k)}) \right| &= \frac{h_{\max}^2}{4} \underbrace{\frac{2}{h_k + h_{k-1}}}_{\leq \frac{1}{h_{\min}}} |w''_{\max}| \underbrace{\left[\begin{aligned} &\frac{1}{2} \frac{h_k^2}{h_{\max}^2} \frac{w''(\xi(x_{k+1/2}))}{|w''_{\max}|} - \frac{h_k^2}{h_{\max}^2} \frac{w''(\xi(x_{k+1}))}{|w''_{\max}|} \\ &-\frac{1}{2} \frac{h_{k-1}^2}{h_{\max}^2} \frac{w''(\xi(x_{k-1/2}))}{|w''_{\max}|} + \frac{h_{k-1}^2}{h_{\max}^2} \frac{w''(\xi(x_k))}{|w''_{\max}|} \end{aligned} \right]}_{c_{hw} \leq 3} \\ &\leq \frac{3h_{\max}^2}{4} \underbrace{\frac{h_{\max}}{h_{\min}}}_{c_{rh} < \infty} |w''_{\max}| = h_{\max} \frac{3c_{rh}}{4} |w''_{\max}| \tag{90} \end{aligned}$$

The error defined in \mathbb{G}^1 norm in the global problem domain becomes

$$\begin{aligned}
 |w - \mathcal{I}_h w|_{\mathbb{G}^1}^2 &= \sum_{k=1}^{N_s} A_k^s |\bar{g}(w_s - \mathcal{I}_h w_s)|^2 \leq \sum_{k=1}^{N_s} h_{\max} \left(h_{\max} \frac{3c_{rh}}{4} |w''_{\max}| \right)^2 \\
 &= \underbrace{N_s}_{\leq \frac{1}{h_{\max}}} h_{\max} \left(h_{\max} \frac{3c_{rh}}{4} |w''_{\max}| \right)^2 \leq \left(h_{\max} \frac{3c_{rh}}{4} |w''_{\max}| \right)^2
 \end{aligned} \tag{91}$$

Therefore, we have Equation (81).

We now examine the error defined in \mathbb{L}^2 norm in the global problem domain. Using Equations (85) and (86), we have

$$\begin{aligned}
 |w - \mathcal{I}_h w|_{T_q} &= \frac{1}{2} |w''(\xi(x))(x - x_k)^2 - w''(\xi(x_{k+1}))h_k(x - x_k)| \\
 &= \frac{h_{\max}^2}{2} |w''_{\max}| \underbrace{\left| \frac{w''(\xi(x))}{|w''_{\max}|} \frac{(x - x_k)^2}{h_{\max}^2} - \frac{w''(\xi(x_{k+1}))}{|w''_{\max}|} \frac{h_k(x - x_k)}{h_{\max}^2} \right|}_{\leq 2} \\
 &\leq h_{\max}^2 |w''_{\max}|
 \end{aligned} \tag{92}$$

The error defined in \mathbb{L}^2 norm becomes:

$$\begin{aligned}
 |w - \mathcal{I}_h w|_{\mathbb{L}^2}^2 &= \sum_{q=1}^{N_e} \int_{\Omega_q^e} (w - \mathcal{I}_h w)^2 dx \leq \sum_{q=1}^{N_e} h_q (h_{\max}^2 |w''_{\max}|)^2 \leq \sum_{q=1}^{N_e} h_{\max} (h_{\max}^2 |w''_{\max}|)^2 \\
 &= \underbrace{N_e}_{\leq \frac{1}{h_{\max}}} h_{\max} (h_{\max}^2 |w''_{\max}|)^2 \leq (h_{\max}^2 |w''_{\max}|)^2
 \end{aligned} \tag{93}$$

which is Equation (82).

In Equation (90), we need to estimate c_{hw} and gave a very sloppy bound of 3. This is in fact a very big overestimate in a usual situation. If a uniform mesh is used, and the target function has a constant second derivative in the cells sharing a smoothing domain, c_{hw} should be zero for all the inner smoothing domains. In such a situation, Equation (90) becomes

$$|\bar{g}(w_k - \mathcal{I}_h w_k)| = \begin{cases} \frac{h}{4} |w''_{\max}|, & k = 1, N_n \\ 0, & k = 2, 3, \dots, N_n - 1 \end{cases} \tag{94}$$

and Equation (91) becomes

$$\begin{aligned}
 |w - \mathcal{I}_h w|_{\mathbb{G}^1}^2 &= \sum_{k=1}^{N_s} A_k^s |\bar{g}(w_k - \mathcal{I}_h w_k)|^2 = A_1^s |\bar{g}(w_1 - \mathcal{I}_h w_1)|^2 + A_{N_n}^s |\bar{g}(w_{N_n} - \mathcal{I}_h w_{N_n})|^2 \\
 &= \frac{h}{2} \left(\frac{h}{4} |w''_{\max}| \right)^2 + \frac{h}{2} \left(\frac{h}{4} |w''_{\max}| \right)^2 = h \left(\frac{h}{4} |w''_{\max}| \right)^2 = \frac{h^3}{4^2} |w''_{\max}|^2
 \end{aligned} \tag{95}$$

which is Equation (83). Under the same condition, Equation (92) becomes

$$(w - \mathcal{I}_h w)|_{T_q} = \frac{1}{2} w'' ((x - x_k)^2 - h_k(x - x_k)) \quad (96)$$

and

$$\begin{aligned} |w - \mathcal{I}_h w|_{\mathbb{L}^2}^2 &= \sum_{q=1}^{N_e} \int_{T_q} (w - \mathcal{I}_h w)^2 dx = \frac{1}{4} \sum_{q=1}^{N_e} \int_{T_q} (w'')^2 ((x - x_k)^2 - h(x - x_k))^2 dx \\ &= \frac{1}{4} (w'')^2 N_e \int_0^h (x^2 - hx)^2 dx = \frac{1}{4} (w'')^2 N_e \int_0^h (x^4 - 2hx^3 + h^2x^2) dx \\ &= \frac{1}{4} (w'')^2 N_e \left(\frac{1}{5} x^5 - \frac{1}{2} hx^4 + \frac{1}{3} h^2x^3 \right)_0^h = \frac{h^5}{4} (w'')^2 N_e \left(\frac{1}{5} - \frac{1}{2} + \frac{1}{3} \right) \\ &= \frac{h^5}{4} \frac{1}{30} (w'')^2 \underbrace{N_e}_{=\frac{1}{h}} = \frac{h^4}{120} (w'')^2 \end{aligned} \quad (97)$$

which is Equation (84). This completes the proof. \square

4.2. Comparison with H^1 norm measures

In the standard weak form formulation such as FEM, we should have the following bounds. For $w \in \mathbb{H}_0^1$ and uniform mesh [21, 34]

$$|w - \mathcal{I}_h w|_{\mathbb{H}^1} \leq h \left(\max_{q=1, \dots, N_e} \max_{\xi \in T_q} |w''(\xi)| \right) \quad (98)$$

and

$$|w - \mathcal{I}_h w|_{\mathbb{L}^2} \leq h^2 \left(\max_{q=1, \dots, N_e} \max_{\xi \in T_q} |w''(\xi)| \right) \quad (99)$$

The proof of these bounds in FEM was based on Rolle's Theorem. Although an exact and physically meaningful comparison is difficult, because the space difference between \mathbb{H}_0^1 and $\mathbb{G}_{h,0}^1$, and the errors in H^1 and G^1 norms carry different physical meanings. But an 'indicative' comparison can be useful in some ways. For errors in semi-norm measure, Equation (98) is quite close to Equation (81) for uniform mesh ($c_{rh} = 1.0$): they all give a convergence rate of 1.0. Equation (83) shows, however, the W^2 formulation can provide a convergence rate in G^1 semi-norm measure of 1.5 at least for cases of even division of node-based smoothing domains. Compared with Equation (98), the convergence rate is 50% higher.

Equation (83) was obtained under the conditions of (1) uniform division of elements/cells and (2) constant second derivative of w in the cells sharing a smoothing domain. The first condition is essentially the same 'symmetrical condition' of smoothing domains for the integral representation to produce the first gradient of a function exactly [5]. When this condition is satisfied, the entire interior smoothing domains become symmetrical, and the smoothing operation will reproduce the first derivative exactly. The only error will be on the boundary where the symmetry condition

cannot be satisfied. The second condition of constant second derivative of the target function seems to be very strong, but it can be rather easy to be quite closely satisfied, because all we need is the second derivative of the target function being constant in the cells sharing a smoothing domain. When the mesh is refined, we can often expect the second derivative being approximately constant locally. Therefore, the rate given in Equation (83) can be expected when the mesh is reasonably fine. In practical applications, on the other hand, the first condition is rather very difficult to meet, simply because it is rare to have uniform division of element/cells for practical problems of complicated geometry. However, we can in fact to expect the smoothing domains to be approximately ‘symmetric’ locally ($h_k \approx h_{k-1}$ for node k for our 1D problems). In such cases, we can still expect Equation (83) holds approximately and hence a convergence rate of 1.5 in G^1 semi-norm. This has been confirmed in many numerical examples presented in [22, 26, 33], where numerical rates of about 1.4 were often found. We were excited about the higher convergence rate, but could not give a good explanation then. We now have at least one possible explanation: the use of (symmetric) smoothing domain reduces significantly the error measured in G^1 semi-norm. More in depth analysis on the errors is needed.

Let us now further examine for general problem with reasonably smoothing second derivative. The coefficient c_{hw} in Equation (90) becomes

$$\begin{aligned}
 c_{hw} &= \left[\frac{1}{2} \frac{h_k^2}{h_{\max}^2} \underbrace{\frac{w''(\xi(x_{k+1/2}))}{|w''_{\max}|}}_{-1 \leq \cdot \leq 1} - \frac{h_k^2}{h_{\max}^2} \underbrace{\frac{w''(\xi(x_{k+1}))}{|w''_{\max}|}}_{-1 \leq \cdot \leq 1} - \frac{1}{2} \frac{h_{k-1}^2}{h_{\max}^2} \underbrace{\frac{w''(\xi(x_{k-1/2}))}{|w''_{\max}|}}_{-1 \leq \cdot \leq 1} + \frac{h_{k-1}^2}{h_{\max}^2} \underbrace{\frac{w''(\xi(x_k))}{|w''_{\max}|}}_{-1 \leq \cdot \leq 1} \right] \\
 &\approx \frac{1}{2} \frac{1}{h_{\max}^2} \underbrace{\frac{|w''(\xi(x_k))|}{|w''_{\max}|}}_{\leq 1} |h_k^2 - h_{k-1}^2| = \frac{1}{2} \frac{h_k(h_k + h_{k-1})}{h_{\max}^2} \frac{|w''(\xi(x_k))|}{|w''_{\max}|} \left| 1 - \underbrace{\frac{h_{k-1}}{h_k}}_{\approx 1} \right| \tag{100}
 \end{aligned}$$

If the lengths of the two neighboring elements of the interior nodes are different, the contribution of these nodes to the error norm can be in control (when mesh is refined), and there the convergence rate will be about 1.0 and not 1.5. However, c_{hw} in Equation (90) can be a very small number. If, for example, there is a 20% length difference in the two neighboring elements of all the interior nodes ($h_{k-1}/h_k = 0.8$), we shall have $c_{hw} \approx 0.2$. This means that even the rate of convergence cannot be improved, the results will still be about 5 times more accurate. This was also observed in the ES-FEM where the edge-based smoothing domains are not quite symmetric, but were often found about 2–10 times more accurate in G^1 norm compared with that of FEM measured in H^1 using the same mesh [28]. In extreme cases, where all the smoothing domain are not symmetric at all, the results of the W^2 formulation will still be better than the standard weak formulation, as shown in Equations (98) and (81).

For errors in L^2 norms, the convergence rates given by Equations (82) and (84) are the same, and it is also the same as in the standard weak form formulation such as FEM given in Equation (99) where Rolle’s theorem is used the in proof process. The bound given by Equation (84) is, however, about 10 times tighter than Equation (82). Note again that such comparisons are made based on different error norms, and hence not exactly comparable. In addition, the interpolation error and

the solution error are different, and hence whether the solution error of a model can be bounded by the interpolation errors need further investigations. More in depth error analyses are needed.

5. CONCLUSION

In this paper we have established for the first time G^1 spaces with a set of important inequalities, which leads to the following major findings:

1. If a function is constructed using a set of linearly independent nodal shape functions for all the nodes in a discrete model, and at least a minimum number of linearly independent smoothing domains are used to evaluate the semi norm, the function is in a \mathbb{G}_h^1 space.
2. The full G^1 norm of a function in a G^1 space is equivalent to the L^2 norm of the nodal values of the function (First inequality).
3. The semi G^1 norm of a function in a \mathbb{G}_h^1 space is equivalent to the L^2 norm of the nodal values of the function (Second inequality).
4. The full norm of a function in a \mathbb{G}_h^1 space is equivalent to the semi-norm (Third inequality).
5. The smoothing operation results in a smaller G^1 semi-norm measure compared with the H^1 semi-norm measure. This is the fundamental inequality for the so-called softening effects.
6. Based on the G^1 space theory, a W^2 formulation of any physically stable problem will be stable and converge. It is applicable to any problems to which the standard weak formulation is applicable.

Some additional analyses on G space theory can be found in [35]. The author is aware of that we have in fact created much more questions than the answers and conclusions, such as what would be the generalization of G^1 to G^m (where m is non-negative integer), what would be the dual space G^{-1} , what kind of linear functionals can we allow, etc. At this stage, we can expect that the dual space G^{-1} should be ‘smaller’ than the H^{-1} , and G^{-1} is properly only a little ‘larger’ than G^1 (in contrast we know that H^{-1} is much ‘larger’ than H^1). In addition, h -dependence of the solution error as well as the regularity issues are not yet clear. We know now the G space offers a theoretical foundation for W^2 formulations that work well (stable and convergent), but how well is yet to be answered. To provide answers to all these and many other questions one may have, much more works and efforts are required. The authors hope this paper can initiate a study in the G spaces and W^2 formulation related areas and hopeful many can join us to accomplish a class of more effective computational methods. Helps from the mathematical community is greatly appreciated on these theoretical issue. Part II of this paper will present W^2 formulation for solid mechanics problems with 2D and 3D examples.

ACKNOWLEDGEMENTS

The author would like to thank Dr Zhang GY for his help in running of these example problems for this paper. The authors also thank Dr Xu X, Mr T. Nguyen-Thoi, Dr H. Nguyen-Xuan, Mr Cheng Lei and many other members at ACES for their comments and suggestions to this work. This work is partially supported by A*STAR, Singapore. It is also partially supported by the Open Research Fund Program of the State Key Laboratory of Advanced Technology of Design and Manufacturing for Vehicle Body, Hunan University, People’s Republic of China under the grant number 40915001.

REFERENCES

1. Zienkiewicz OC, Taylor RL. *The Finite Element Method* (5th edn). Butterworth, Heinemann: Oxford, 2000.
2. Liu GR, Quek SS. *Finite Element Method: A Practical Course*. Butterworth, Heinemann: Burlington, MA, 2003.
3. Oliveira Eduardo R De Arantes E. Theoretical foundations of the finite element method. *International Journal of Solids and Structures* 1968; **4**:929–952.
4. Liu GR, Liu MB. *Smoothed Particle Hydrodynamics—A Mesh Free Particle Method*. World Scientific: Singapore, 2003.
5. Liu GR, Gu YT. *An Introduction to MFree Methods and their Programming*. Springer: Berlin, 2005.
6. Liu GR. *Meshfree Methods: Moving Beyond the Finite Element Method* (2nd edn). CRC Press: Boca Raton, U.S.A., 2009.
7. Pian THH, Wu CC. *Hybrid and Incompatible Finite Element Methods*. CRC Press: Boca Raton, 2006.
8. Quarteroni A, Valli A. *Numerical Approximation of Partial Differential Equations*, Chapter 7. Springer: Berlin, 1994.
9. Liu GR, Dai KY, Nguyen TT. A smoothed finite element method for mechanics problems. *Computational Mechanics* 2007; **39**:859–877.
10. Liu GR, Nguyen TT, Dai KY, Lam KY. Theoretical aspects of the smoothed finite element method (SFEM). *International Journal for Numerical Methods in Engineering* 2007; **71**:902–930.
11. Dai KY, Liu GR, Nguyen TT. An n -sided polygonal smoothed finite element method (nSFEM) for solid mechanics. *Finite Elements in Analysis and Design* 2007; **43**:847–860.
12. Eringen AC, Edelen DGB. On nonlocal elasticity. *International Journal of Engineering Sciences* 1972; **10**: 233–248.
13. Zhang YQ, Liu, GR, Han X. Effect of small length scale on elastic buckling of multi-walled carbon nanotubes under radial pressure. *Physics Letters A* 2006; **349**:370–376.
14. Lucy LB. Numerical approach to testing the fission hypothesis. *Astronomical Journal* 1977; **82**:1013–1024.
15. Liu MB, Liu GR, Zong Z. An overview on smoothed particle hydrodynamics. *International Journal of Computational Methods* 2008; **5**:135–188.
16. Monaghan JJ. Why particle methods work (hydrodynamics). *SIAM Journal on Scientific and Statistical Computing* 1982; **3**:422–433.
17. Chen JS, Wu CT, Belytschko T. Regularization of material instabilities by mesh free approximations with intrinsic length scales. *International Journal for Numerical Methods in Engineering* 2000; **47**:1303–1322.
18. Chen JS, Wu CT, Yoon S, You Y. A stabilized conforming nodal integration for Galerkin mesh-free methods. *International Journal for Numerical Methods in Engineering* 2001; **50**:435–466.
19. Liu FR, Zhang GY, Dai KY, Wang YY, Zhong ZH, Li GY, Han X. A linearly conforming point interpolation method (LC-PIM) for 2D mechanics problems. *International Journal for Computational Methods* 2005; **2**:645–665.
20. Zhang GY, Liu GR, Wang YY, Huang HT, Zhong ZH, Li GY, Han X. A linearly conforming point interpolation method (LC-PIM) for three-dimensional elasticity problems. *International Journal for Numerical Methods in Engineering* 2007; **72**:1524–1543.
21. Liu GR, Zhang GY. A normed G space and weakened weak (W_2) formulation of a cell-based smoothed point interpolation method. *International Journal of Computational Methods* 2009; **6**(1):147–179.
22. Liu GR. A generalized gradient smoothing technique and the smoothed bilinear form for Galerkin formulation of a wide class of computational methods. *International Journal of Computational Methods* 2008; **5**(2):199–236.
23. Liu GR, Gu YT. A point interpolation method for two-dimensional solids. *International Journal for Numerical Methods in Engineering* 2001; **50**:937–951.
24. Wang JG, Liu GR. A point interpolation meshless method based on the radial basis functions. *International Journal for Numerical Methods in Engineering* 2002; **54**:1623–1648.
25. Liu GR, Li Y, Dai KY, Luan MT, Xue W. A linearly conforming radial point interpolation method for solid mechanics problems. *International Journal of Computational Methods* 2006; **3**:401–428.
26. Liu GR, Zhang GY. Upper bound solution to elasticity problems: a unique property of the linearly conforming point interpolation method (LC-PIM). *International Journal for Numerical Methods in Engineering* 2008; **74**:1128–1161.
27. Liu GR, Nguyen TT, Nguyen XH, Lam KY. A node-based smoothed finite element method (N-SFEM) for upper bound solutions to solid mechanics problems. *Computers and Structures* 2009; **87**:14–26.
28. Liu GR, Nguyen TT, Lam KY. An edge-based smoothed finite element method (E-SFEM) for static and dynamic problems of solid mechanics. *Journal of Sound and Vibration* 2009; **320**:1100–1130.

29. Nguyen TT, Liu GR, Lam KY, Zhang GY. A face-based smoothed finite element method (F-SFEM) for 3D linear and nonlinear solid mechanics problems using 4-node tetrahedral elements. *International Journal for Numerical Methods in Engineering* 2009; **78**:324–353. DOI: 10.1002/nme.2491.
30. Liu GR, Zhang GY. Edge-based smoothed point interpolation method (ES-PIM). *International Journal of Computational Methods* 2008; **5**(4):621–646.
31. Peraire J. *Lecture Notes on Finite Element Methods for Elliptic Problems*. MIT: Cambridge, MA, 1999.
32. Hughes TJR. *The Finite Element Method*. Prentice-Hall: London, 1987.
33. Zhang GY, Liu GR, Nguyen TT, Song CX, Han X, Zhong ZH, Li GY. The upper bound property for solid mechanics of the linearly conforming radial point interpolation method (LC-RPIM). *International Journal of Computational Methods* 2007; **4**(3):521–541.
34. Strang G, Fix GJ. *An Analysis of the Finite Element Method*. Prentice-Hall: Englewood Cliffs, 1973.
35. Liu GR. On a G space theory. *International Journal of Computational Methods* 2009; **6**(2):257–289.