Lecture-1: Weak Derivative and Sobolev Spaces MA 573: Finite Element Methods for PDEs

Dr. Bhupen Deka*

1 Introduction

The mathematical models of science and engineering mainly take the form of differential equations. With the rapid development of high speed computers over the last decades the possibilities of efficiently utilizing these models have dramatically increased. To use mathematical models on computer one needs numerical methods. Only in the very simplest case it is possible to find exact analytical solutions of the equations in the model, and in general one has to rely on numerical techniques finding approximate solutions. One such numerical technique is **Finite Element Method (FEM)** .

In Finite Element Method we seek the approximate solution in finite number of sub-domains instead of solving the problem in whole computational domain. The method was introduced in the late 50's and early 60's by engineers for the numerical solution of differential equations in structural engineering. During mid 60's the mathematical study of finite element method started and then the method was developed by engineers mathematician and numerical analyst into a general method for numerical solution of differential equation with applications in many areas of science and engineering.

In general, we can say that finite element method is a numerical method to solve a boundary value problem which gives the approximate solution at each point of a computational domain. In this method we seek the approximate solution in some sub domain of the computational domain. Such sub-domain are known as elements. The collection of all elements is known as mesh. If the elements are of same size then it is known as uniform mesh otherwise it is called non-uniform mesh.

In this note, we will introduce basic definition and notations which are required to introduce finite element methods for partial differential equations.

2 Weak Derivative

In this section, we shall introduce some basic notations, function spaces and preliminary materials to be used in this thesis. All functions considered here are real valued. For the purpose of introducing notations, we assume Ω to be a convex polygonal domain in $\mathbb{R}^d$ (d-dimensional Euclidean space) and $\partial\Omega$ denote the boundary of Ω. For $x = (x_1, x_2, \ldots, x_d) \in \Omega$, set $dx = dx_1 \ldots dx_d$. Further,

*Department of Mathematics, IIT Guwahati, North Guwahati, India, Guwahati- 781039, India (`bdeka@iitg.ac.in`).

let $\alpha = (\alpha_1, \ldots, \alpha_d)$ be an d-tuple with nonnegative integer component and denote the order of α as $|\alpha| = \alpha_1 + \alpha_2 + \ldots + \alpha_d$. Then, by $D^\alpha \phi$, we shall mean the αth derivative of ϕ defined by

$$D^\alpha \phi = \frac{\partial^{|\alpha|} \phi}{\partial {x_1}^{\alpha_1} \ldots \partial {x_n}^{\alpha_d}}.$$

By support of a function ϕ, denoted by supp(ϕ), we mean the closure of all points x with $\phi(x) \neq 0$, i.e.,

$$\text{supp}(\phi) = \overline{\{x \in \Omega : \phi(x) \neq 0\}}.$$

For any nonnegative integer m, $C^m(\overline{\Omega})$ denotes the space of functions with continuous derivatives upto and including order m in $\overline{\Omega}$. $C_0^m(\Omega)$ is the space of all $C^m(\Omega)$ functions with compact support in Ω and $C_0^\infty(\Omega)$ is the space of all infinitely differentiable functions with compact support in Ω.

Now we introduce the following function spaces which we shall refer frequently. For any domain $\mathcal{M} \subseteq \Omega \subset \mathbb{R}^d,\ d = 1,\ 2,$ with $1 \leq p \leq \infty$, $L^p(\Omega)$ denotes the linear space of equivalence classes of measurable functions ϕ on Ω such that $\|\phi\|_{L^p(\mathcal{M})} < \infty$, where

$$\begin{aligned} \|\phi\|_{L^p(\mathcal{M})} &:= \left(\int_{\mathcal{M}} |\phi(x)|^p dx \right)^{\frac{1}{p}}, \quad 1 \leq p < \infty, \\ \|\phi\|_{L^\infty(\mathcal{M})} &:= \underset{x \in \mathcal{M}}{\text{ess sup}}\ |\phi(x)| < \infty. \end{aligned}$$

When $p = 2$, $L^2(\mathcal{M})$ is a Hilbert space with respect to the inner product

$$(\phi, \psi)_{\mathcal{M}} = \int_{\mathcal{M}} \phi(x)\psi(x)dx.$$

For simplicity of notation, we write the norm $\|.\|_{L^2(\mathcal{M})}$ of $L^2(\mathcal{M})$ by $\|.\|_{\mathcal{M}}$ and remove the subscript $\mathcal{M}$ whenever $\mathcal{M} = \Omega$.

Remark 2.1 The space $L^2(\mathcal{M})$ is also known as collection of square-integrable functions. For example, a real valued function $f : [0, 1] \to \mathbb{R}$ defined by

$$f(x) = \begin{cases} x, & \text{in } [0, \frac{1}{2}], \\ -x, & \text{in } (\frac{1}{2}, 1], \end{cases}$$

is square-integrable as

$$\|f\|^2_{L^2(0,1)} = \int_0^2 |f|^2 dx = \int_0^1 x^2 dx = \frac{1}{2} < \infty.$$

Therefore, a square-integrable function need not be continuous. On the other hand, a continuous function also need not be square integrable. For example, a real valued function $f : (0, 1) \to \mathbb{R}$ defined by $f(x) = \frac{1}{x},\ x \in (0, 1)$ is continuous, but not square-integrable as

$$\|f\|^2_{L^2(0,1)} = \int_0^1 |f(x)|^2 dx = \int_0^1 x^2 dx = -\Big[\frac{1}{x}\Big]_0^1 \to \infty.$$

Remark 2.2 Throughout this course, we assume that the integration is in the sense of Lebesgue. Note that all Riemann integrable functions in $\Omega \subset \mathbb{R}^d$ are Lebesgue integrable.

Since a square-integrable function need not be continuous, so we can't assume that a square-integrable function has classical derivative. Now, we extend the definition of classical derivative to introduce notion of weak derivative of a function. We begin with following equation

$$\frac{dy}{dx} = g(x) \text{ in } \Omega \subset \mathbb{R}.$$

It is known that if $g \in C(\Omega)$ then $y \in C^1(\Omega)$. The problem arises if $g \in L^2(\Omega)$, then we can not expect the solution in $C^1(\Omega)$. Naturally, we assume that $y \in C(\Omega) \supset C^1(\Omega)$. But functions of $C(\Omega)$ need not be differentiable. Thus, we can not find $\frac{dy}{dx}$ in classical sense. Therefore, we need to attach a meaning to the derivative of y even y is not differentiable or not at all continuous. This is done by introducing weak derivative of such functions.

To motivate towards weak derivative of such function's let us try to define the weak derivative of a **differentiable function** f and then we will generalize the same definition for larger class of functions. Let ${C_0}^\infty(\Omega)$ be the collection of all C^∞ functions define over Ω which vanishes on the boundary $\partial\Omega$ of Ω. Then, for $v \in {C_0}^\infty(\Omega)$, we have

$$\begin{aligned}\int_\Omega \frac{df}{dx} v dx &= fv\Big]_{\partial\Omega} - \int_\Omega f\frac{dv}{dx}dx \\ &= -\int_\Omega f\frac{dv}{dx}dx\end{aligned}$$

Thus, under the sign of integration, derivative of f reduces to the derivative of v. Since, this does not involve the derivative of f, we can define **the weak derivative for those functions f, which not at all differentiable or continuous.** The first order weak derivative of f is given by a function g which satisfies following relation

$$\int_\Omega g v dx = -\int_\Omega f v' dx \ \ \forall v \in {C_0}^\infty(\Omega).$$

Similarly, the second order weak derivative of f is given by a function h which satisfies following relation

$$\int_\Omega h v dx = \int_\Omega f v'' dx \ \ \forall v \in {C_0}^\infty(\Omega).$$

In general, m-th order weak derivative of f is denoted by $f^{(m)}$ and defined by following relation

$$\int_\Omega f^{(m)} v dx = (-1)^m \int_\Omega f v^{(m)} dx \ \ \forall v \in {C_0}^\infty(\Omega).$$

Example 2.1 We know that $f(x) = |x|$ is not differentiable in $[0, 1]$ in classical sense. Let us see whether it has a weak derivative or not. If it has a weak derivative f', then by definition it satisfies following relation

$$\int_{-1}^{1} f' v dx = -\int_{-1}^{1} f v' dx \ \ \forall v \in {C_0}^\infty(\Omega). \tag{2.1}$$

For right hand side, use the definition of f and method of parts to have

$$\begin{aligned}
-\int_{-1}^{1} fv'dx &= -\int_{-1}^{0}(-x)v'dx - \int_{0}^{1}(x)v'dx \\
&= \int_{-1}^{0} xv'dx - \int_{0}^{1} xv'dx \\
&= \Big[xv\Big]_{-1}^{0} - \int_{-1}^{0} vdx - \Big[xv\Big]_{0}^{1} + \int_{0}^{1} vdx \\
&= -v(-1) - v(1) + \int_{-1}^{1} gvdx \\
&= \int_{-1}^{1} gvdx \ \ \forall v \in C_0^\infty(-1,1). \qquad (2.2)
\end{aligned}$$

In the last equality, we have used the fact that v vanishes on the boundary of $(-1,1)$. Comparing (2.1) and (2.2), we obtain $f'(x) = g(x) = -1$ in $(-1,0)$ and $f'(x) = g(x) = 1$ in $[0,1)$.

Remark 2.3 In the above example, we can write $f'(x) = g(x) = -1$ in $(-1,0]$ and $f'(x) = g(x) = 1$ in $(0,1)$. Note that weak derivative f' differ at $x = 0$. We assume that both are equal in the sense of almost everywhere.

Now, we proceed to define weak partial derivatives. Before proceeding further, we recall few classical results. Let u, v be scalar functions and $\mathbf{w} = (w_1,\ldots,w_d)$ a vector-valued function of $x \in \mathbb{R}^d$. We define the gradient, the divergence, and the Laplace operator (Laplacian) by

$$\begin{aligned}
\nabla v &= \text{grad} = \Big(\frac{\partial v}{\partial x_1}, \frac{\partial v}{\partial x_2}, \ldots, \frac{\partial v}{\partial x_d}\Big), \\
\nabla \cdot \mathbf{w} &= \text{div } \mathbf{w} = \frac{\partial \mathbf{w_1}}{\partial x_1} + \frac{\partial \mathbf{w_2}}{\partial x_2} + \ldots + \frac{\partial \mathbf{w_d}}{\partial x_d}, \\
\Delta u &= \nabla \cdot \nabla u = \frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2} + \ldots + \frac{\partial^2 u}{\partial x_d^2}.
\end{aligned}$$

We recall the *divergence theorem*

$$\int_\Omega \nabla \cdot \mathbf{w} dx = \int_{\partial\Omega} \mathbf{w}\cdot\mathbf{n} ds, \qquad (2.3)$$

where $\mathbf{n} = (n_1,\ldots,n_d)$ is the outward unit normal to boundary $\partial\Omega$. Applying this to the product $\mathbf{w}v$ we obtain *Greens formula*:

$$\int_\Omega \nabla\cdot\mathbf{w}vdx + \int_\Omega \mathbf{w}\cdot\nabla vdx = \int_{\partial\Omega}\mathbf{w}\cdot\mathbf{n}vds. \qquad (2.4)$$

When applied with $\mathbf{w} = \nabla u$ the formula becomes

$$\int_\Omega \Delta uvdx = -\int_\Omega \nabla u\cdot\nabla udx + \int_{\partial\Omega}\frac{\partial u}{\partial \mathbf{n}}vds, \qquad (2.5)$$

where $\frac{\partial u}{\partial \mathbf{n}} = \nabla u\cdot\mathbf{n}$ is the outward normal derivative of u on $\partial\Omega$.

Remark 2.4 Setting $\mathbf{w} = (u,0,\ldots,0)$ in (2.4), we obtain

$$\int_\Omega \frac{\partial u}{\partial x_1}vdx = -\int_\Omega u\frac{\partial v}{\partial x_1}dx + \int_{\partial\Omega} un_1vds.$$

In general,

$$\int_\Omega \frac{\partial u}{\partial x_i} v dx = -\int_\Omega u \frac{\partial v}{\partial x_i} dx + \int_{\partial\Omega} u n_i v ds. \tag{2.6}$$

Relation (2.6) motivates to define first order weak partial derivative of u with respect to variable x_i, which is denoted by $\frac{\partial u}{\partial x_i}$ and defined by

$$\int_\Omega \frac{\partial u}{\partial x_i} v dx = -\int_\Omega u \frac{\partial v}{\partial x_i} dx \ \ \forall v \in C_0^\infty(\Omega). \tag{2.7}$$

For a given multi-index $\alpha = (\alpha_1, \ldots, \alpha_d)$ with nonnegative integer, the α-th order weak derivative of a function u is denoted by $D^\alpha u$ and defined by

$$\int_\Omega D^\alpha u v dx = (-1)^{|\alpha|} \int_\Omega u D^\alpha v dx \ \ \forall v \in C_0^\infty(\Omega), \tag{2.8}$$

where $|\alpha| = \alpha_1 + \alpha_2 + \ldots + \alpha_d$ and recall that $D^\alpha v$ is given by

$$D^\alpha v = \frac{\partial^{|\alpha|} v}{\partial {x_1}^{\alpha_1} \ldots \partial {x_n}^{\alpha_d}}.$$

Remark 2.5 Suppose $\Omega \subset \mathbb{R}^2$. Then first order partial derives of v are given by multi-index $\alpha = (1, 0)$ and $\alpha = (0, 1)$ as follows

$$D^\alpha v = \frac{\partial^{|\alpha|} v}{\partial {x_1}^{\alpha_1} \ldots \partial {x_n}^{\alpha_d}} = \frac{\partial v}{\partial x_1} \ \ \& \ \ D^\alpha v = \frac{\partial^{|\alpha|} v}{\partial {x_1}^{\alpha_1} \ldots \partial {x_n}^{\alpha_d}} = \frac{\partial v}{\partial x_2}.$$

Therefore $\alpha = (1, 0)$ and $\alpha = (0, 1)$ in (2.8) gives weak partial derivatives $\frac{\partial u}{\partial x_1}$ and $\frac{\partial u}{\partial x_2}$, respectively. Similarly, second order partial derivatives of v are given by multi-index $\alpha = (2, 0),\ \alpha = (0, 2)$ and $\alpha = (1, 1)$ as follows

$$\begin{aligned} D^\alpha v &= \frac{\partial^{|\alpha|} v}{\partial {x_1}^{\alpha_1} \ldots \partial {x_n}^{\alpha_d}} = \frac{\partial^2 v}{\partial x_1^2}, \\ D^\alpha v &= \frac{\partial^{|\alpha|} v}{\partial {x_1}^{\alpha_1} \ldots \partial {x_n}^{\alpha_d}} = \frac{\partial^2 v}{\partial x_2^2}, \\ D^\alpha v &= \frac{\partial^{|\alpha|} v}{\partial {x_1}^{\alpha_1} \ldots \partial {x_n}^{\alpha_d}} = \frac{\partial^2 v}{\partial x_1 \partial x_2}. \end{aligned}$$

Therefore $\alpha = (2, 0), \alpha = (0, 2)$ and $\alpha = (1, 1)$ in (2.8) gives weak partial derivatives $\frac{\partial^2 u}{\partial x_1^2}$, $\frac{\partial^2 u}{\partial x_2^2}$ and $\frac{\partial^2 u}{\partial x_1 \partial x_2}$, respectively.

Example 2.2 The function u defined on $\Omega = (-1, 1) \times (-1, 1)$ by

$$u = \begin{cases} x, & x > 0 \\ 0, & x \leq 0 \end{cases}$$

has first order weak partial derivatives. Determine them.

Solution. Suppose $\frac{\partial u}{\partial y}$ is the first order weak partial derivative of u with respect to variable y. Then, for all $v \in C_0^\infty(\Omega)$, we obtain

$$\begin{aligned} \int_\Omega \frac{\partial u}{\partial y} v dx dy &= -\int_\Omega u \frac{\partial v}{\partial y} dx dy \\ &= -\int_{-1}^1 \int_{-1}^1 u(x, y) \frac{\partial v}{\partial y} dx dy \\ &= -\int_0^1 x \Big[\int_{-1}^1 \frac{\partial v}{\partial y} dy \Big] dx = -\int_0^1 x \Big[v(x, 1) - v(x, -1) \Big] dx = 0. \end{aligned}$$

Here, we have used the fact that v vanishes along $x = 1,\ x = -1,\ y = 1$ and $y = -1$. Therefore,

$$\int_\Omega \frac{\partial u}{\partial y} v dxdy = \int_\Omega 0 \times v dxdy \ \ \forall v \in C_0^\infty(\Omega),$$

so that weak partial derivative $\frac{\partial u}{\partial y} = 0$ in Ω. Similarly, for all $v \in C_0^\infty(\Omega)$, we obtain

$$\begin{aligned}
\int_\Omega \frac{\partial u}{\partial x} v dxdy &= -\int_\Omega u \frac{\partial v}{\partial x} dxdy \\
&= -\int_{-1}^{1}\int_{-1}^{1} u(x,y)\frac{\partial v}{\partial x} dxdy \\
&= -\int_{-1}^{1}\Big(\int_0^1 x\frac{\partial v}{\partial x} dx\Big) dx = -\int_{-1}^{1}\Big([xv]_0^1 - \int_0^1 v dx\Big) dy \\
&= \int_{-1}^{1}\int_0^1 v dxdy = \int_{-1}^{1}\int_{-1}^{1} H(x,y) v dxdy,
\end{aligned}$$

where H is the weak first order partial derivative $\frac{\partial u}{\partial x}$ and defined by

$$H(x,y) = \begin{cases} 1, & x > 0, \\ 0, & x \leq 0. \end{cases}$$

3 Sobolev Spaces

In this section, we will introduce Hilbertian Sobolev spaces.

Definition 3.1 *Let $m > 0$ be an integer and let $1 \leq p < \infty$. Then the Sobolev space $W^{m,p}(\Omega)$ is defined by*

$$W^{m,p}(\Omega) = \{u \in L^p(\Omega) : D^\alpha \in L^p(\Omega)\ \forall\ |\alpha| \leq m\},$$

where $D^\alpha u$ is the α th order weak derivative of u. In other words, $W^{m,p}(\Omega)$ is the collection of all functions in $L^p(\Omega)$ such that that all its distributional derivatives upto order m are also in $L^p(\Omega)$.

For $p = 2$, $W^{m,p}(\Omega)$ is a Hilbert space and it is denoted by $H^m(\Omega)$. More precisely, in $\Omega \subset \mathbb{R}$, we have

$$H^m(\Omega) = \Big\{u \in L^2(\Omega) :\ u,\ u',\ \ldots, u^{(m)} \in L^2(\Omega)\Big\}$$

with the inner product

$$\langle u, v\rangle_{H^m(\Omega)} = \int_\Omega \Big\{uv + u'v' + \ldots + u^{(m)}v^{(m)}\Big\} dx$$

and induced norm

$$\|u\|_{H^m(\Omega)} = \Big[\|u\|^2_{L^2(\Omega)} + \|u'\|^2_{L^2(\Omega)} + \ldots + \|u^{(m)}\|^2_{L^2(\Omega)}\Big]^{\frac{1}{2}}.$$

Here, $u^{(m)}$ denotes the m-th order weak derivative of u.

In $\Omega \subset \mathbb{R}^2$, we have

$$H^1(\Omega) = \Big\{ u \in L^2(\Omega) : \ u, \ \frac{\partial u}{\partial x} \ \& \ \frac{\partial u}{\partial y} \in L^2(\Omega) \Big\}$$

$$H^2(\Omega) = \Big\{ u \in L^2(\Omega) : \ u, \ \frac{\partial u}{\partial x}, \ \frac{\partial u}{\partial y}, \ \frac{\partial^2 u}{\partial x^2}, \frac{\partial^2 u}{\partial x \partial y} \ \& \ \frac{\partial^2 u}{\partial y^2} \in L^2(\Omega) \Big\}$$

and so on. The inner product in $H^1(\Omega)$ is given by

$$\langle u, v \rangle_{H^1(\Omega)} = \int_\Omega \Big\{ uv + \frac{\partial u}{\partial x}\frac{\partial v}{\partial x} + \frac{\partial u}{\partial y}\frac{\partial v}{\partial y} \Big\} dxdy = \int_\Omega \Big\{ uv + \nabla u \cdot \nabla v \Big\} dxdy$$

and the induced norm

$$\|u\|_{H^1(\Omega)} = \left[\int_\Omega \Big\{ u^2 + \Big(\frac{\partial u}{\partial x}\Big)^2 + \Big(\frac{\partial u}{\partial y}\Big)^2 \Big\} dxdy \right]^{\frac{1}{2}} = \left[\int_\Omega \Big\{ u^2 + |\nabla u|^2 \Big\} dxdy \right]^{\frac{1}{2}}.$$

Similarly, inner product in $H^2(\Omega)$ is given by

$$\langle u, v \rangle_{H^2(\Omega)} = \int_\Omega \Big\{ uv + \nabla u \cdot \nabla v + \Delta u \Delta u \Big\} dxdy$$

and the induced norm

$$\|u\|_{H^2(\Omega)} = \left[\int_\Omega \Big\{ u^2 + |\nabla u|^2 + |\Delta u|^2 \Big\} dxdy \right]^{\frac{1}{2}}.$$

Definition 3.2 The collection of all H^1 functions vanishing on the boundary is a closed subspace of H^1 and it is denoted by $H_0^1(\Omega) = \{v \in H^1(\Omega) : v|_{\partial\Omega} = 0\}$.

Remark 3.1 Consider following boundary value problem

$$y'' + y = f \quad \text{in } \Omega = (a, b) \quad \& \quad y(a) = 0 = y(b). \tag{3.1}$$

Multiply the above equation by $v \in H_0^1(\Omega)$ and then integrate by parts to have

$$\int_a^b \{y'v' + yv\} dx = \int_a^b fv dx \ \text{ since } v(a) = 0 = v(b).$$

Therefore, solution y satisfies following equation

$$A(y, v) = (f, v) \ \ \forall v \in H_0^1(\Omega), \tag{3.2}$$

where

$$A(y, v) = \int_a^b \{y'v' + yv\} dx \ \ \& \ \ (f, v) = \int_a^b fv dx \ \ \forall y, \ v \in H_0^1(\Omega).$$

Equation (3.2) is known as variational formulation (problem)/integral formulation/weak formulation of the original problem (3.1).

Now, we give a formal definition for the variational problem.

Definition 3.3 Suppose H is a given Hilbert space. Let $A : H \times H \to \mathbf{R}$ and $L : H \to \mathbf{R}$ is a functional. Then a variational problem is defined as: Find $u \in H$ such that

$$A(u, v) = L(v) \ \ \forall \ v \in H. \tag{3.3}$$

In finite element algorithm, we approximate u in a finite dimensional space $\tilde{H}$. If $\tilde{H} \subset H$, then we call conforming finite element method. Throughout this course, we will discuss **conforming FEM**.