Chapter 7

Theoretical Foundations of the Finite Element Method

Using finite element methods, we need to answer these questions:

- What is the appropriate functional space V for the solution?
- What is the appropriate *weak* or *variational* form of a differential equation?
- What kind of basis functions or finite element spaces should we choose?
- How accurate is the finite element solution?

We briefly address these questions in this Chapter. Recalling that finite element methods are based on *integral forms* and not on the pointwise sense as in finite difference methods. We will generalize the theory corresponding to the pointwise form to deal with integral forms.

7.1 Functional spaces

A *functional Space* is a *set* of functions with operations. For example,

$$C(\Omega) = C^0(\Omega) = \Big\{ \, u(x), \;\; u(x) \text{ is continuous on } \Omega \Big\} \tag{7.1}$$

is a linear space that contains all continuous functions on Ω, the domain where the functions are defined, *i.e.*, $\Omega = [0,\ 1]$. The space is linear because for any real numbers α and β and $u_1 \in C(\Omega)$ and $u_2 \in C(\Omega)$, we have $\alpha u_1 + \beta u_2 \in C(\Omega)$.

The functional space with first order continuous derivatives in 1D is

$$C^1(\Omega) = \Big\{ u(x), \quad u(x),\ u'(x) \text{ are continuous on } \Omega \Big\}\,, \tag{7.2}$$

and similarly

$$C^m(\Omega) = \Big\{ u(x), \quad u(x), u'(x), \cdots, u^{(m)} \text{ are continuous on } \Omega \Big\}\,. \tag{7.3}$$

Obviously,

$$C^0 \supset C^1 \supset \cdots \supset C^m \supset \cdots\,. \tag{7.4}$$

Then as $m \to \infty$, we define

$$C^{\infty}(\Omega) = \{u(x), \quad u(x) \text{ is indefinitely differentiable on } \Omega\} . \tag{7.5}$$

For example, e^x, $\sin x$, $\cos x$, and polynomials, are in $C^{\infty}(-\infty,\infty)$, but some other elementary functions such as $\log x$, $\tan x$, $\cot x$ are not if $x=0$ is in the domain.

7.1.1 Multi-dimensional spaces and multi-index notations

Let us now consider multi-dimensional functions $u(\mathbf{x}) = u(x_1, x_2, \cdots, x_n)$, $\mathbf{x} \in R^n$, and a corresponding multi-index notation that simplifies expressions for partial derivatives. We can write $\alpha = (\alpha_1, \alpha_2, \cdots, \alpha_n)$, $\alpha_i \ge 0$ for an integer vector in R^n, *e.g.*, if $n=5$, then $\alpha = (1,2,0,0,2)$ is one of possible vectors. We can readily represent a partial derivative as

$$D^{\alpha} u(\mathbf{x}) = \frac{\partial^{|\alpha|} u}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2} \cdots \partial x_n^{\alpha_n}} , \quad |\alpha| = \alpha_1 + \alpha_2 + \cdots + \alpha_n , \quad \alpha_i \ge 0 \tag{7.6}$$

which is the so called multi-index notation.

Example 7.1. *For $n=2$ and $u(\mathbf{x}) = u(x_1,x_2)$, all possible $D^{\alpha}u$ when $|\alpha| = 2$ are*

$$\begin{aligned}
\alpha = (2,0), &\quad D^{\alpha} u = \frac{\partial^2 u}{\partial x_1^2} , \\
\alpha = (1,1), &\quad D^{\alpha} u = \frac{\partial^2 u}{\partial x_1 \partial x_2} , \\
\alpha = (0,2), &\quad D^{\alpha} u = \frac{\partial^2 u}{\partial x_2^2} .
\end{aligned}$$

With the multi-index notation, the C^m space in a domain $\Omega \in R^n$ can be defined as

$$C^m(\Omega) = \{u(x_1, x_2, \cdots, x_n), \quad D^{\alpha} u \text{ are continuous on } \Omega, \ |\alpha| \le m\} \tag{7.7}$$

i.e., all possible derivatives up to order m are continuous on Ω.

Example 7.2. *For $n=2$ and $m=3$, we have u, u_x, u_y, u_{xx}, u_{xy}, u_{yy}, u_{xxx}, u_{xxy}, u_{xyy} and u_{yyy} all continuous on Ω if $u \in C^3(\Omega)$, or simply $D^{\alpha}u \in C^3(\Omega)$ for $|\alpha| \le 3$. Note that $C^m(\Omega)$ has infinite dimensions.*

The *distance* in $C^0(\Omega)$ is defined as

$$d(u,v) = \max_{x \in \Omega} |u(x) - v(x)|$$

with the properties (1), $d(u,v) \ge 0$; (2), $d(u,v) = 0$ if and only if $u \equiv v$; and (3), $d(u, v+w) \le d(u,v) + d(u,w)$, the triangle inequality. A linear space with a distance defined is called a *metric space*.

A norm in $C^0(\Omega)$ is a non-negative function of u that satisfies

$$\|u(x)\| = d(u,\theta) = \max_{x\in\Omega} |u(x)|\,, \text{where } \theta \text{ is the zero element,}$$

with the properties (1), $\|u(x)\| \geq 0\,,$ and $\|u(x)\| = 0$ if and only if $u \equiv 0\,;$

(2), $\|\alpha u(x)\| = |\alpha|\|u(x)\|$, where α is a number;

(3), $\|u(x) + v(x)\| \leq \|u(x)\| + \|v(x)\|$, the triangle inequality.

A linear space with a norm defined is called a *normed space*. In $C^m(\Omega)$, the distance and the norm are defined as

$$d(u,v) = \max_{0\leq|\alpha|\leq m}\max_{x\in\Omega} |D^\alpha u(x) - D^\alpha v(x)|\,, \tag{7.8}$$

$$\|u(x)\| = \max_{0\leq|\alpha|\leq m}\max_{x\in\Omega} |D^\alpha u|\,. \tag{7.9}$$

7.2 Spaces for integral forms, $L^2(\Omega)$ and $L^p(\Omega)$

In analogy to pointwise spaces $C^m(\Omega)$, we can define Sobolev spaces $H^m(\Omega)$ in integral forms. The square-integrable space $H^0(\Omega) = L^2(\Omega)$ is defined as

$$L^2(\Omega) = \left\{ u(x), \quad \int_\Omega u^2(x)\,dx < \infty \right\} \tag{7.10}$$

corresponding to the pointwise $C^0(\Omega)$ space. It is easy to see that

$$C(0,1) = C^0(0,1) \subset L^2(0,1)\,.$$

Example 7.3. *It is easy to verify that $y(x) = 1/x^{1/4} \notin C^0(0,1)$, but*

$$\int_0^1 \left(\frac{1}{x^{1/4}}\right)^2 dx = \int_0^1 \frac{1}{\sqrt{x}}dx = 2 < \infty$$

so that $y(x) \in L^2(0,1)$. But it is obvious that $y(x)$ is not in $C(0,1)$ since $y(x)$ blows up as $x \to 0$ from $x > 0$, see Fig.7.1 in which we also show that a piecewise constant function 0 and 1 in $(0,1)$ is in $L^2(0,1)$ but not in $C(0,1)$ since the function is discontinuous (non-removable discontinuity) at $x = 0.5$.

The *distance* in $L^2(\Omega)$ is defined as

$$d(f,g) = \left\{ \int_\Omega |f-g|^2\;dx \right\}^{1/2}, \tag{7.11}$$

which satisfies the three conditions of the distance definition; so $L^2(\Omega)$ is a metric space. We say that two functions f and g are identical ($f \equiv g$) in $L^2(\Omega)$ if $d(f,g) = 0$. For example, the following two functions are identical in $L^2(-2,2)$:

$$f(x) = \begin{cases} 0, & \text{if } -2 \leq x < 0, \\ 1, & \text{if } 0 \leq x \leq 2, \end{cases} \qquad g(x) = \begin{cases} 0, & \text{if } -2 \leq x \leq 0, \\ 1, & \text{if } 0 < x \leq 2. \end{cases}$$

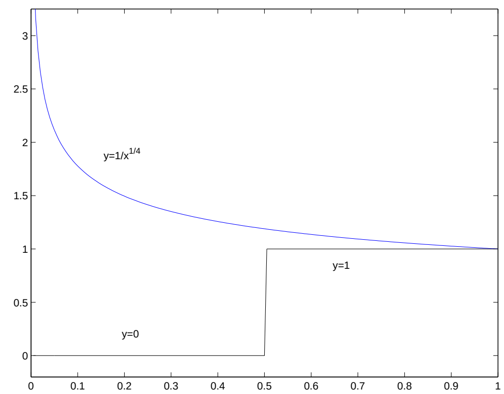

Figure 7.1. *Plot of two functions that are in* $L^2(0,1)$ *but not in* $C[0,1]$*. One function is* $y(x) = \frac{1}{\sqrt[4]{x}}$*. The other one is* $y(x) = 0$ *in* $[0, 1/2)$ *and* $y(x) = 1$ *in* $[1/2, 1]$.

The norm in the $L^2(\Omega)$ space is defined as

$$\|u\|_{L^2} = \|u\|_0 = \left\{ \int_\Omega |u|^2 \, dx \right\}^{1/2} . \tag{7.12}$$

It is straightforward to prove that the usual properties for the distance and the norm hold.

We say that $L^2(\Omega)$ *is a complete space*, meaning that any Cauchy sequence $\{f_n(\Omega)\}$ in L^2 has a limit in $L^2(\Omega)$; *i.e.*, there is a function $f \in L^2(\Omega)$ such that

$$\lim_{n\to\infty} \|f_n - f\|_{L^2} = 0, \quad \text{or} \quad \lim_{n\to\infty} f_n = f.$$

A Cauchy sequence is a sequence that satisfies the property that, for any given positive number ϵ, no matter how small it is, there is an integer N such that

$$\|f_n - f_m\|_{L^2} < \epsilon, \qquad \text{if} \quad m \ge N, \quad n \ge N\,.$$

A complete normed space is called a *Banach* space (a Cauchy sequence converges in terms of the norm), so $L^2(\Omega)$ is a Banach space.

7.2.1 The inner product in L^2

For any two vectors

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \qquad y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

in R^n, we recall that the inner product is

$$(x,y) = x^T y = \sum_{i=1}^{n} x_i y_i = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n \,.$$

Similarly, the inner product in $L^2(\Omega)$ space in R, the real number space, is defined as

$$(f,g) = \int_\Omega f(x) g(x)\, dx \,, \text{ for any } f \text{ and } g \in L^2(\Omega)\,, \tag{7.13}$$

and satisfies the familiar properties

$$\begin{aligned} (f,g) &= (g,f)\,, \\ (\alpha f, g) &= (f, \alpha g) = \alpha (f,g), \quad \forall \alpha \in R, \\ (f, g+w) &= (f,g) + (f,w) \end{aligned}$$

for any f, g and $w \in L^2(\Omega)$. The norm, distance, and inner product in $L^2(\Omega)$ are related as follows:

$$\|u\|_0 = \|u\|_{L^2(\Omega)} = \sqrt{(u,u)} = d(u,\theta) = \left\{ \int_\Omega |u|^2 \, dx \right\}^{1/2} . \tag{7.14}$$

With the $L^2(0,1)$ inner product, for the simple model problem

$$-u'' = f\,, \quad 0 < x < 1\,, \quad u(0) = u(1) = 0,$$

we can rewrite the weak form as

$$(u', v') = (f, v)\,, \quad \forall v \in H_0^1(0,1)\,,$$

and the minimization form as

$$\min_{v \in H_0^1(0,1)} F(v) : \quad F(v) = \frac{1}{2}(v', v') - (f, v)\,.$$

7.2.2 The Cauchy-Schwartz inequality in $L^2(\Omega)$

For a Hilbert space with the norm $\|u\| = \sqrt{(u,u)}$, the Cauchy-Schwartz inequality is

$$|(u,v)| \le \|u\|_0 \|v\|_0 \,. \tag{7.15}$$

Examples of the Cauchy-Schwartz inequality corresponding to inner products in R^n and in L^2 spaces are:

$$\left| \sum_{i=1}^{n} x_i y_i \right| \le \left\{ \sum_{i=1}^{n} x_i^2 \right\}^{1/2} \left\{ \sum_{i=1}^{n} y_i^2 \right\}^{1/2} ,$$

$$\left| \sum_{i=1}^{n} x_i \right| \le \sqrt{n} \left\{ \sum_{i=1}^{n} x_i^2 \right\}^{1/2} ,$$

$$\left| \int_\Omega f g dx \right| \le \left\{ \int_\Omega f^2 \, dx \right\}^{1/2} \left\{ \int_\Omega g^2 \, dx \right\}^{1/2} ,$$

$$\left| \int_\Omega f dx \right| \le \left\{ \int_\Omega f^2 \, dx \right\}^{1/2} \sqrt{V}\,,$$

where V is the volume of Ω.

A proof of the Cauchy-Schwartz inequality

Noting that $(u,u) = \|u\|^2$, we construct a quadratic function of α given u and v:

$$f(\alpha) = (u+\alpha v, u+\alpha v) = (u,u) + 2\alpha(u,v) + \alpha^2(v,v) \geq 0\,.$$

The quadratic function is non-negative; hence the discriminant of the quadratic form satisfies

$$\Delta = b^2 - 4ac \leq 0, \quad i.e., \quad 4(u,v)^2 - 4(u,u)(v,v) \leq 0 \;\text{ or }\; (u,v)^2 \leq (u,u)(v,v)\,,$$

yielding the Cauchy-Schwartz inequality $|(u,v)| \leq \|u\|\|v\|$, on taking the square root of both sides.

A complete Banach space with an inner product defined is called a *Hilbert space*. Hence $L^2(\Omega)$ is a Hilbert space (linear space, inner product, complete).

Relationships between the spaces

The relationships (and relevant additional properties) in the hierarchy of defined spaces may be summarized diagrammatically:

Metric Space (distance) $\Longrightarrow$ Normed Space (norm) $\Longrightarrow$ Banach space (complete) $\Longrightarrow$ Hilbert space (inner product).

7.2.3 $L^p(\Omega)$ spaces

An $L^p(\Omega)$ space is defined as

$$L^p(\Omega) = \left\{ u(x), \int_\Omega |u(x)|^p \, dx < \infty \right\}, \tag{7.16}$$

and the distance in $L^p(\Omega)$ is defined as

$$d(f,g) = \left\{ \int_\Omega |f-g|^p \, dx \right\}^{1/p}. \tag{7.17}$$

An $L^p(\Omega)$ space has a distance and is complete, so it is a Banach space. However, it is not a Hilbert space, because no corresponding inner product is defined unless $p = 2$.

7.3 Sobolev spaces and weak derivatives

Similar to $C^m(\Omega)$ spaces, we use Sobolev spaces $H^m(\Omega)$ to define function spaces with derivatives involving integral forms. If there is no derivative, then the relevant Sobolev space is

$$H^0(\Omega) = L^2(\Omega) = \left\{ v(x), \; \int_\Omega |v|^2 dx < \infty \right\}. \tag{7.18}$$

7.3.1 Definition of weak derivatives

If $u(x) \in C^1[0,1]$, then for any function $\phi \in C^1(0,1)$ such that $\phi(0) = \phi(1) = 0$ we recall

$$\int_0^1 u'(x)\phi(x) = u\phi \Big|_0^1 - \int_0^1 u(x)\phi'(x)\,dx = -\int_0^1 u(x)\phi'(x)\,dx\,, \tag{7.19}$$

where $\phi(x)$ is a test function in $C_0^1(0,1)$. The first order weak derivative of $u(x) \in L^2(\Omega) = H^0(\Omega)$ is defined to be a function $v(x)$ satisfying

$$\int_\Omega v(x)\phi(x)dx = -\int_\Omega u(x)\phi'(x)dx \tag{7.20}$$

for all $\phi(x) \in C_0^1(\Omega)$ with $\phi(0) = \phi(1) = 0$. If such a function exists, then we write $v(x) = u'(x)$.

Example 7.4.

Consider the following function $u(x)$

$$u(x) = \begin{cases} \dfrac{x}{2}, & \text{if } 0 \le x < \dfrac{1}{2}, \\[2ex] \dfrac{1-x}{2}, & \text{if } \dfrac{1}{2} \le x \le 1, \end{cases}$$

It is obvious that $u(x) \in C[0,1])$ but $u'(x) \notin C(0,1)$ since the classic derivative does not exist at $x = \frac{1}{2}$. Let $\phi(x) \in C^1(0,1)$ be any function that vanishes at two ends, *i.e.*, $\phi(0) = \phi(1) = 0$, and has first order continuous derivtive on (0, 1). We carry out the following integration by parts.

$$\begin{aligned}
\int_0^1 \phi' u dx &= \int_0^{\frac{1}{2}} \phi' u dx + \int_{\frac{1}{2}}^1 \phi' u dx \\
&= \phi(x)\, u(x)\Big|_0^{\frac{1}{2}} + \phi(x)\, u(x)\Big|_{\frac{1}{2}}^1 - \int_0^{\frac{1}{2}} \phi u' dx + \int_{\frac{1}{2}}^1 \phi u' dx \\
&= \phi(1/2)\Big(u(1/2-) - u(1/2+)\Big) - \int_0^{\frac{1}{2}} \frac{\phi(x)}{2} dx - \int_{\frac{1}{2}}^1 \frac{-\phi(x)}{2} dx \\
&= -\int_0^1 \psi(x)\phi(x)dx,
\end{aligned}$$

where we have used the property that $\phi(0) = \phi(1) = 0$ and $\psi(x)$ is defined as

$$\psi(x) = \begin{cases} \dfrac{1}{2}, & \text{if } 0 < x < \dfrac{1}{2}, \\[2ex] -\dfrac{1}{2}, & \text{if } \frac{1}{2} < x < 1, \end{cases}$$

which is what we would expect. In other words, we define the weak derivative of $u(x)$ as $u'(x) = \psi(x)$ which is an $H^1(0,1)$ but not a $C(0,1)$ function.

Similarly, the m-th order weak derivative of $u(x) \in H^0(\Omega)$ is defined as a function $v(x)$ satisfying

$$\int_\Omega v(x)\phi(x)dx = (-1)^m \int_\Omega u(x)\phi^{(m)}(x)dx \tag{7.21}$$

for all $\phi(x) \in C_0^m(\Omega)$ with $\phi(x) = \phi'(x) = \cdots = \phi^{(m-1)}(x) = 0$ for all $x \in \partial\Omega$. If such a function exists, then we write $v(x) = u^{(m)}(x)$.

7.3.2 Definition of Sobolev spaces $H^m(\Omega)$

The Sobolev space $H^1(\Omega)$ defined as

$$H^1(\Omega) = \left\{ v(x)\,,\ D^\alpha v \in L^2(\Omega)\,,\ |\alpha| \le 1 \right\} \tag{7.22}$$

involves first order derivatives, *e.g.*,

$$H^1(a,b) = \left\{ v(x)\,,\ a < x < b\,,\ \int_a^b v^2 dx < \infty\,,\ \int_a^b (v')^2 dx < \infty \right\},$$

and in two space dimensions, $H^1(\Omega)$ is defined as

$$H^1(\Omega) = \left\{ v(x,y)\,,\ v \in L^2(\Omega)\,,\ \frac{\partial v}{\partial x} \in L^2(\Omega)\,,\ \frac{\partial v}{\partial y} \in L^2(\Omega) \right\}.$$

The extension is immediate to the Sobolev space of general dimension

$$H^m(\Omega) = \left\{ v(\mathbf{x})\,,\ D^\alpha v \in L^2(\Omega)\,,\ |\alpha| \le m \right\}. \tag{7.23}$$

7.3.3 Inner products in $H^m(\Omega)$ spaces

The inner product in $H^0(\Omega)$ is the same as that in $L^2(\Omega)$, *i.e.*,

$$(u,v)_{H^0(\Omega)} = (u,v)_0 = \int_\Omega uv\,dx\,.$$

The inner product in $H^1(a,b)$ is defined as

$$(u,v)_{H^1(a,b)} = (u,v)_1 = \int_a^b \left(uv + u'v' \right) dx\,;$$

the inner product in $H^1(\Omega)$ of two variables is defined as

$$(u,v)_{H^1(\Omega)} = (u,v)_1 = \iint_\Omega \left(uv + \frac{\partial u}{\partial x}\frac{\partial v}{\partial x} + \frac{\partial u}{\partial y}\frac{\partial v}{\partial y} \right) dxdy\,; \quad \text{and}$$

the inner product in $H^m(\Omega)$ (of general dimension) is

$$(u,v)_{H^m(\Omega)} = (u,v)_m = \iint_\Omega \sum_{|\alpha| \le m} \left(D^\alpha u(\mathbf{x}) \right) \left(D^\alpha v(\mathbf{x}) \right) d\mathbf{x}\,. \tag{7.24}$$

The norm in $H^m(\Omega)$ (of general dimension) is

$$\|u\|_{H^m(\Omega)} = \|u\|_m = \left\{ \int_\Omega \sum_{|\alpha| \le m} |D^\alpha u(\mathbf{x})|^2 \, d\mathbf{x} \right\}^{1/2}, \tag{7.25}$$

therefore $H^m(\Omega)$ is a Hilbert space. A norm can be defined from the inner product, *e.g.*, in $H^1(a,b)$, the norm is

$$\|u\|_1 = \left\{ \int_a^b \left(u^2 + {u'}^2 \right) dx \right\}^{1/2}.$$

The distance in $H^m(\Omega)$ (of general dimension) is defined as

$$d(u,v)_m = \|u - v\|_m \,. \tag{7.26}$$

7.3.4 Relations between $C^m(\Omega)$ and $H^m(\Omega)$ — the Sobolev embedding theorem

In 1D spaces, we have

$$H^1(\Omega) \subset C^0(\Omega), \quad H^2(\Omega) \subset C^1(\Omega), \quad \cdots, \; H^{1+j}(\Omega) \subset C^j(\Omega)\,.$$

Theorem 7.1. *The Sobolev embedding theorem: If $2m > n$, then*

$$H^{m+j} \subset C^j, \quad j = 0, 1, \cdots, \tag{7.27}$$

where n is the dimension of the independent variables of the elements in the Sobolev space.

Example 7.5. *In 2D spaces, we have $n = 2$. The condition $2m > n$ means that $m > 1$. From the embedding theorem, we have*

$$H^{2+j} \subset C^j, \quad j = 0 \Longrightarrow H^2 \subset C^0, \quad j = 1 \Longrightarrow H^3 \subset C^1, \cdots. \tag{7.28}$$

If $u(x,y) \in H^2$, which means that u, u_x, u_y, u_{xx}, u_{xy} and u_{yy} all belong to L^2, then $u(x,y)$ is continuous, but u_x and u_y may not be continuous!

Example 7.6. *In 3D spaces, we have $n = 3$ and the condition $2m > n$ means that $m > 3/2$ whose closest integer is number two, leading to the same result as in 2D:*

$$H^{2+j} \subset C^j, \quad j = 0 \Longrightarrow H^2 \subset C^0\,, \quad j = 1 \;\Longrightarrow H^3 \subset C^1\,, \; \cdots. \tag{7.29}$$

We regard the *regularity* of a solution as the degree of smoothness for a class of problems measured in C^m or H^m space. Thus for $u(x) \in H^m$ or $u(x) \in C^m$, the larger the m the smoother the function.

7.4 FE analysis for 1D boundary value problems

For the simple 1D model problem

$$-u'' = f\,, \quad 0 < x < 1\,, \quad u(0) = u(1) = 0\,,$$

we know that the weak form is

$$\int_0^1 u'v'\,dx = \int_0^1 fv\,dx \quad \text{or} \quad (u',v') = (f,v)\,.$$

Intuitively, because v is arbitrary we can take $v = f$ or $v = u$ to get

$$\int_0^1 u'v'\,dx = \int_0^1 {u'}^2\,dx\,, \qquad \int_0^1 fv\,dx = \int_0^1 f^2\,dx\,,$$

so u, u', f, v and v' should belong to $L^2(0,1)$; *i.e.*, we have $u \in H_0^1(0,1)$ and $v \in H_0^1(0,1)$; so the solution is in the Sobolev space $H_0^1(0,1)$. We should also take v in the same space for a conforming finite element method. From the Sobolev embedding theorem, we also know that $H^1 \subset C^0$, so the solution is continuous.

7.4.1 Conforming FE methods

Definition 7.2. *If the finite element space is a subspace of the solution space, then the finite element space is called a* conforming finite element *space, and the finite element method is called a* conforming FE method*.*

For example, the piecewise linear function over a given a mesh is a conforming finite element space for the model problem. We mainly discuss conforming finite element methods in this book. On including the boundary condition, we define the solution space as

$$H_0^1(0,1) = \Big\{ v(x)\,, \quad v(0) = v(1) = 0\,,\ v \in H^1(0,1) \Big\}\,. \tag{7.30}$$

When we look for a finite element solution in a finite dimensional space V_h, it should be a subspace of $H_0^1(0,1)$ for a conforming finite element method. For example, given a mesh for the 1D model, we can define a finite dimensional space using piecewise continuous linear functions over the mesh:

$$V_h = \{ v_h\,,\ v_h(0) = v_h(1) = 0\,,\ v_h \text{ is continuous piecewise linear} \}\,.$$

The finite element solution would be chosen from the finite dimensional space V_h, a subspace of $H_0^1(0,1)$. If the solution of the weak form is in $H_0^1(0,1)$ but not in the V_h space; then an error is introduced on replacing the solution space with the finite dimensional space. Nevertheless, the finite element solution is the best approximation in V_h in some norm, as discussed later.

7.4.2 FE analysis for 1D Sturm-Liouville problems

A 1D Sturm-Liouville problem on (x_l, x_r) with a Dirichlet boundary condition at two ends is

$$\begin{aligned} -(p(x)u'(x))' + q(x)u(x) &= f(x)\,, \quad x_l < x < x_r\,, \\ u(x_l) = 0\,, \qquad & u(x_r) = 0\,, \\ p(x) \ge p_{min} > 0\,, \qquad & q(x) \ge q_{min} \ge 0\,. \end{aligned} \tag{7.31}$$

The conditions on $p(x)$ and $q(x)$ guarantee the problem is well-posed, such that the weak form has a unique solution. It is convenient to assume $p(x) \in C(x_l, x_r)$ and $q(x) \in C(x_l, x_r)$. Later we will see that these conditions together with $f(x) \in L^2(x_l, x_r)$, guarantee the unique solution to the weak form of the problem. To derive the weak form, we multiply both sides of the equation by a test function $v(x)$, $v(x_l) = v(x_r) = 0$ and integrate from x_l to x_r to get

$$\begin{aligned} \int_{x_l}^{x_r} \Big(-(p(x)u')' + qu\Big)\, v\, dx &= -pu'v\Big|_{x_l}^{x_r} + \int_{x_l}^{x_r} \Big(pu'v' + quv\Big)\, dx \\ &= \int_{x_l}^{x_r} fv\, dx \\ \Longrightarrow \int_{x_l}^{x_r} \Big(pu'v' + quv\Big)\, dx &= \int_{x_l}^{x_r} fv\, dx\ , \ \forall v \in H_0^1(x_l, x_r) \quad \text{or} \quad a(u,v) = L(v). \end{aligned}$$

7.4.3 The bilinear form

The integral

$$a(u,v) = \int_{x_l}^{x_r} \Big(pu'v' + quv\Big)\, dx \tag{7.32}$$

is a bilinear form, because it is linear for both u and v from the following

$$\begin{aligned} a(\alpha u + \beta w, v) &= \int_{x_l}^{x_r} \Big(p(\alpha u' + \beta w')v' + q(\alpha u + \beta w)v\Big)\, dx \\ &= \alpha \int_{x_l}^{x_r} \Big(pu'v' + quv\Big)\, dx + \beta \int_{x_l}^{x_r} \Big(w'v' + qwv\Big)\, dx \\ &= \alpha a(u,v) + \beta a(w,v)\,, \end{aligned}$$

where α and β are scalars; and similarly,

$$a(u, \alpha v + \beta w) = \alpha a(u,v) + \beta a(u,w)\,.$$

It is noted that this bilinear form is an inner product, usually different from the L^2 and H^1 inner products, but if $p \equiv 1$ and $q \equiv 1$ then

$$a(u,v) = (u,v).$$

Since $a(u,v)$ is an inner product, under the conditions: $p(x) \ge p_{min} > 0$, $q(x) \ge 0$, we can define the *energy norm* as

$$\|u\|_a = \sqrt{a(u,u)} = \left\{ \int_{x_l}^{x_r} \Big(p(u')^2 + qu^2\Big)\, dx \right\}^{\frac{1}{2}}, \tag{7.33}$$

where the first term may be interpreted as the *kinetic energy* and the second term as the *potential energy.* The Cauchy-Schwartz inequality implies $|a(u,v)| \le \|u\|_a \|v\|_a$.

The bilinear form combined with linear form often simplifies the notation for the weak and minimization forms, *e.g.*, for the above Sturm-Liouville problem the weak form becomes

$$a(u,v) = L(v)\ ,\ \forall v \in H_0^1(x_l, x_r)\,, \tag{7.34}$$

and the minimization form is

$$\min_{v \in H_0^1(x_l,x_r)} F(v) = \min_{v \in H_0^1(x_l,x_r)} \left\{ \frac{1}{2}\, a(v,v) - L(v) \right\}. \tag{7.35}$$

Later we will see that all self-adjoint differential equations have both weak and minimization forms, and that the finite element method using the Ritz form is the same as with the Galerkin form.

7.4.4 The FE method for 1D Sturm-Liouville problems using hat basis functions

Consider *any* finite dimensional space $V_h \subset H_0^1(x_l, x_r)$ with the basis

$$\phi_1(x) \in H_0^1(x_l, x_r)\,,\ \phi_2(x) \in H_0^1(x_l, x_r)\,,\ \cdots\,,\ \phi_M(x) \in H_0^1(x_l, x_r),$$

that is,

$$\begin{aligned} V_h &= span\,\{\phi_1, \phi_2, \cdots, \phi_M\} \\ &= \left\{ v_s, \quad v_s = \sum_{i=1}^{M} \alpha_i \phi_i \right\} \subset H_0^1(x_l, x_r)\,. \end{aligned}$$

The Galerkin finite method assumes the approximate solution to be

$$u_s(x) = \sum_{j=1}^{M} \alpha_j \phi_j(x)\,, \tag{7.36}$$

and the coefficients $\{\alpha_j\}$ are chosen such that the weak form

$$a(u_s, v_s) = (f, v_s), \quad \forall\ v_s \in V_h$$

is satisfied. Thus we enforce the weak form in the finite dimensional space V_h instead of the solution space $H_0^1(x_l, x_r)$, which introduces some error.

Since *any* element in the space is a linear combination of the basis functions, we have

$$a\,(u_s, \phi_i) = (f, \phi_i)\,, \quad i = 1, 2, \cdots, M\,,$$

or

$$a\left(\sum_{j=1}^{M} \alpha_j\, \phi_j, \phi_i \right) = (f, \phi_i)\,, \quad i = 1, 2, \cdots, M\,,$$

or

$$\sum_{j=1}^{M} a(\phi_j, \phi_i)\, \alpha_j = (f, \phi_i)\,, \quad i = 1, 2, \cdots, M\,.$$

In the matrix-vector form $AX = F$, this system of algebraic equations for the coefficients is

$$\begin{bmatrix} a(\phi_1,\phi_1) & a(\phi_1,\phi_2) & \cdots & a(\phi_1,\phi_M) \\ a(\phi_2,\phi_1) & a(\phi_2,\phi_2) & \cdots & a(\phi_2,\phi_M) \\ \vdots & \vdots & \vdots & \vdots \\ a(\phi_M,\phi_1) & a(\phi_M,\phi_2) & \cdots & a(\phi_M,\phi_M) \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_M \end{bmatrix} = \begin{bmatrix} (f,\phi_1) \\ (f,\phi_2) \\ \vdots \\ (f,\phi_M) \end{bmatrix},$$

and the system has some attractive properties.

- The coefficient matrix A is symmetric, *i.e.*, $\{a_{ij}\} = \{a_{ji}\}$ or $A = A^T$, since $a(\phi_i, \phi_j) = a(\phi_j, \phi_i)$. Note that this is only true for a self-adjoint problem such as the above, with the second order ODE

$$-(pu')' + qu = f\,.$$

  For example, the similar problem involving the ODE

$$-u'' + u' = f$$

  is not self-adjoint; and the Galerkin finite element method using the corresponding weak form

$$(u', v') + (u', v) = (f, v) \text{ or } (u', v') + (u, v') = (f, v)$$

  produces terms such as (ϕ_i', ϕ_j) that differ from (ϕ_j', ϕ_i), so that the coefficient matrix A is not symmetric.

- A is positive definite, *i.e.*,

$$x^T A\, x > 0 \text{ if } x \neq 0\,, \text{ and all eigenvalues of } A \text{ are positive.}$$

Proof: For any $\eta \neq 0$, we show that $\eta^T A\,\eta > 0$ as follows

$$
\begin{aligned}
\eta^T A\,\eta = \eta^T (A\eta) &= \sum_{i=1}^{M} \eta_i \sum_{j=1}^{M} a_{ij}\eta_j \\
&= \sum_{i=1}^{M} \eta_i \sum_{j=1}^{M} a(\phi_i, \phi_j)\eta_j \\
&= \sum_{i=1}^{M} \eta_i \sum_{j=1}^{M} a(\phi_i, \eta_j\phi_j) \\
&= \sum_{i=1}^{M} \eta_i\, a\left(\phi_i, \sum_{j=1}^{M} \eta_j\phi_j\right) \\
&= a\left(\sum_{i=1}^{M} \eta_i\phi_i, \sum_{j=1}^{M} \eta_j\phi_j\right) \\
&= a\,(v_s, v_s) = \|v_s\|_a^2 > 0\,,
\end{aligned}
$$

since $v_s = \displaystyle\sum_{i=1}^{M} \eta_i\phi_i \neq 0$ because η is a nonzero vector and the $\{\phi_i\}$'s are linear independent.

7.4.5 Local stiffness matrix and load vector using the hat basis functions

The local stiffness matrix using the hat basis functions is a 2×2 matrix of the following,

$$
K_i^e = \begin{bmatrix} \int_{x_i}^{x_{i+1}} p(x)\,(\phi_i')^2 dx & \int_{x_i}^{x_{i+1}} p(x)\,\phi_i'\phi_{i+1}' dx \\ \int_{x_i}^{x_{i+1}} p(x)\,\phi_{i+1}'\phi_i' dx & \int_{x_i}^{x_{i+1}} p(x)\,(\phi_{i+1}')^2 dx \end{bmatrix} + \begin{bmatrix} \int_{x_i}^{x_{i+1}} q(x)\,\phi_i^2 dx & \int_{x_i}^{x_{i+1}} q(x)\,\phi_i\phi_{i+1} dx \\ \int_{x_i}^{x_{i+1}} q(x)\,\phi_{i+1}\phi_i dx & \int_{x_i}^{x_{i+1}} q(x)\,\phi_{i+1}^2 dx \end{bmatrix},
$$

and the local load vector is

$$
F_i^e = \begin{bmatrix} \int_{x_i}^{x_{i+1}} f\phi_i dx \\ \int_{x_i}^{x_{i+1}} f\phi_{i+1} dx \end{bmatrix}.
$$

The global stiffness matrix and load vector can be assembled element by element.

7.5 Error analysis of the FE method

Error analysis for finite element methods usually includes two parts:

1. error estimates for an intermediate function in V_h, often the interpolation function; and
2. convergence analysis, a limiting process that shows the finite element solution converges to the true solution of the weak form in some norm, as the mesh size h approaches zero.

We first recall some notations and setting up:

1. Given a weak form $a(u, v) = L(v)$ and a space V, which usually has infinite dimension, the problem is to find a $u \in V$ such that the weak form is satisfied for any $v \in V$. Then u is called the solution of the weak form.

2. A *finite dimensional* subspace of V denoted by V_h (i.e. $V_h \subset V$) is adopted for a conforming finite element method and it does not have to depend on h, however.

3. The solution of the weak form in the subspace V_h is denoted by u_h, *i.e.*, we require $a(u_h, v_h) = L(v_h)$ for any $v_h \in V_h$.

4. The global error is defined by $e_h = u(x) - u_h(x)$, and we seek a sharp upper bound for $\|e_h\|$ using certain norm.

It was noted that error is introduced when the finite dimensional space replaces the solution space, as the weak form is usually only satisfied in the sub-space V_h and not in the solution space V. However, we can prove that the solution satisfying the weak form in the sub-space V_h is the best approximation to the exact solution u in the finite dimensional space in the energy norm.

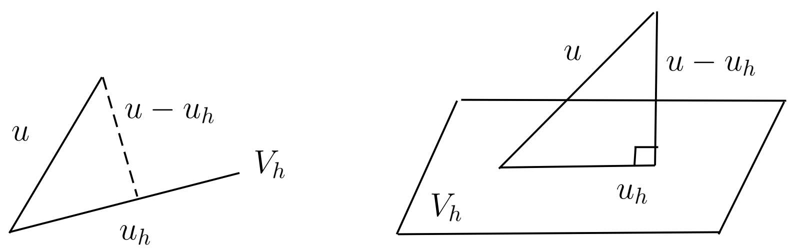

Figure 7.2. *A diagram of FE approximation properties. The finite element solution is the best approximation to the solution u in the finite dimensional space V_h in the energy norm; and the error $u - u_h$ is perpendicular to the finite dimensional space V_h in the inner product of $a(u, v)$.*

Theorem 7.3.

1. *u_h is the* projection *of u onto V_h through the inner product $a(u, v)$, i.e.,*

$$u - u_h \perp V_h \quad or \quad u - u_h \perp \phi_i\,, \; i = 1, 2, \cdots, M\,, \tag{7.37}$$

$$a(u - u_h, v_h) = 0 \;\; \forall\, v_h \in V_h \quad or \quad a(u - u_h, \phi_i) = 0\,, \; i = 1, 2, \cdots, M\,, \tag{7.38}$$

where $\{\phi_i\}$'s are the basis functions.

2. *u_h is the best approximation in the energy norm, i.e.,*

$$\|u - u_h\|_a \le \|u - v_h\|_a\,, \;\forall\, v_h \in V_h\,.$$

Proof:

$$\begin{aligned}
a(u,v) &= (f,v)\,,\ \forall\, v \in V,\\
\to a(u,v_h) &= (f,v_h)\,,\ \forall\, v_h \in V_h \ \text{ since } V_h \subset V\,,\\
a(u_h,v_h) &= (f,v_h)\,,\ \forall\, v_h \in V_h \ \text{ since } u_h \text{ is the solution in } V_h\,,\\
\text{subtract} \to a(u-u_h,v_h) &= 0 \ \text{ or } a(e_h,v_h) = 0\,,\ \forall\, v_h \in V_h\,.
\end{aligned}$$

Now we prove that u_h is the best approximation in V_h.

$$\begin{aligned}
\|u-v_h\|_a^2 &= a(u-v_h,u-v_h)\\
&= a(u-u_h+u_h-v_h,u-u_h+u_h-v_h)\\
&= a(u-u_h+w_h,u-u_h+w_h)\,,\ \text{on letting} \quad w_h = u_h - v_h \in V_h\,,\\
&= a(u-u_h,u-u_h+w_h) + a(w_h,u-u_h+w_h)\\
&= a(u-u_h,u-u_h) + a(u-u_h,w_h) + a(w_h,u-u_h) + a(w_h,w_h)\\
&= \|u-u_h\|_a^2 + 0 + 0 + \|w_h\|_a^2, \quad \text{since,} \quad a(e_h,u_h) = 0\\
&\ge \|u-u_h\|_a^2
\end{aligned}$$

i.e., $\|u-u_h\|_a \le \|u-v_h\|_a$. Fig. 7.2 is a diagram to illustrate the theorem.

Example: For the Sturm-Liouville problem,

$$\begin{aligned}
\|u-u_h\|_a^2 &= \int_a^b \Big(p(x)\,(u'-u_h')^2 + q(x)\,(u-u_h)^2 \Big)\,dx\\
&\le p_{max}\int_a^b (u'-u_h')^2 dx + q_{max}\int_a^b (u-u_h)^2\,dx\\
&\le \max\{p_{max},q_{max}\}\int_a^b \Big((u'-u_h')^2 + (u-u_h)^2 \Big)\,dx\\
&= C\|u-u_h\|_1^2,
\end{aligned}$$

where $C = \max\{p_{max},q_{max}\}$. Thus we obtain

$$\begin{aligned}
&\|u-u_h\|_a \le \hat{C}\|u-u_h\|_1\,,\\
&\|u-u_h\|_a \le \|u-v_h\|_a \le \bar{C}\|u-v_h\|_1\,.
\end{aligned}$$

7.5.1 Interpolation functions and error estimates

Usually the solution is unknown; so in order to get the error estimate we choose a special $v_h^* \in V_h$, for which we can get a good error estimate. We may then use the error estimate $\|u-u_h\|_a \le \|u-v_h^*\|_a$ to get an error estimate for the finite element solution (maybe over-estimated). Usually we can choose a *piecewise interpolation function* for this purpose. That is another reason that we choose piecewise linear, quadratic or cubic functions over the given mesh in finite element methods.

Linear 1D piecewise interpolation function

Given a mesh $x_0, x_1, x_2, \cdots, x_M$, the linear 1D piecewise interpolation function is defined as

$$u_I(x) = \frac{x - x_i}{x_{i-1} - x_i} u(x_{i-1}) + \frac{x - x_{i-1}}{x_i - x_{i-1}} u(x_i)\,, \quad x_{i-1} \le x \le x_i\,.$$

It is obvious that $u_I(x) \in V_h$, where $V_h \subset H^1$ is the set of continuous piecewise linear functions that have the first order weak derivative, so

$$\|u - u_h\|_a \le \|u - u_I\|_a\,.$$

Since $u(x)$ is unknown, then so is $u_I(x)$. Nevertheless, we know the *upper error bound* of the interpolation functions.

Theorem 7.4. *Given a function $u(x) \in C^2[a, b]$ and a mesh $x_0, x_1, x_2, \cdots, x_M$, the continuous piecewise linear function u_I has the error estimates*

$$\|u - u_I\|_\infty = \max_{x\in[a,b]} |u(x) - u_I(x)| \le \frac{h^2}{8}\|u''\|_\infty\,, \tag{7.39}$$

$$\|u'(x) - u_I'(x)\|_{L^2(a,b)} \le h\sqrt{b-a}\|u''\|_\infty\,. \tag{7.40}$$

Proof: If $\tilde{e_h} = u(x) - u_I(x)$, then $\tilde{e_h}(x_{i-1}) = \tilde{e_h}(x_i) = 0$. From Rolle's theorem, there must be at least one point z_i between x_{i-1} and x_i such that $\tilde{e_h}'(z_i) = 0$, hence

$$\begin{aligned} \tilde{e_h}'(x) &= \int_{z_i}^{x} \tilde{e_h}''(t)\,dt \\ &= \int_{z_i}^{x} \big(u''(t) - u_I''(t)\big)\,dt \\ &= \int_{z_i}^{x} u''(t)\,dt\,. \end{aligned}$$

Therefore, we obtain the error estimates below

$$\begin{aligned} |\tilde{e_h}'(x)| &\le \int_{z_i}^{x} |u''(t)|dt \le \|u''\|_\infty \int_{z_i}^{x} dt \le \|u''\|_\infty h\,, \quad \text{and} \\ \|\tilde{e_h}'\|_{L^2(a,b)} &= \|\tilde{e_h}'\|_0 \le \left\{ \|u''\|_\infty^2 \int_a^b h^2 dt \right\}^{\frac{1}{2}} \le \sqrt{b-a}\|u''\|_\infty h\,; \end{aligned}$$

so we have proved the second inequality. To prove the first, assume that $x_{i-1} + h/2 \le z_i \le x_i$, otherwise we can use the other half interval. From the Taylor expansion

$$\begin{aligned} \tilde{e_h}(x) &= \tilde{e_h}(z_i + x - z_i)\,, \text{ assuming } x_{i-1} \le x \le x_i\,, \\ &= \tilde{e_h}(z_i) + \tilde{e_h}'(z_i)(x - z_i) + \frac{1}{2}\tilde{e_h}''(\xi)(x - z_i)^2, \;\; x_{i-1} \le \xi \le x_i\,, \\ &= \tilde{e_h}(z_i) + \frac{1}{2}\tilde{e_h}''(\xi)(x - z_i)^2\,, \end{aligned}$$

so at $x = x_i$ we have

$$
\begin{aligned}
0 = \tilde{e_h}(x_i) &= \tilde{e_h}(z_i) + \frac{1}{2}\tilde{e_h}''(\xi)(x_i - z_i)^2 \,, \\
\tilde{e_h}(z_i) &= -\frac{1}{2}\tilde{e_h}''(\xi)(x_i - z_i)^2 \,, \\
|\tilde{e_h}(z_i)| &\le \frac{1}{2}\|u''\|_\infty (x_i - z_i)^2 \le \frac{h^2}{8}\|u''\|_\infty \,.
\end{aligned}
$$

Note that the largest value of $\tilde{e_h}(x)$ has to be the z_i where the derivative is zero.

7.5.2 Error estimates of the finite element methods using the interpolation function

Theorem 7.5. *For the 1D Sturm-Liouville problem, the following error estimates hold,*

$$
\begin{aligned}
\|u - u_h\|_a &\le Ch\|u''\|_\infty, \\
\|u - u_h\|_1 &\le \hat{C}h\|u''\|_\infty \,,
\end{aligned}
$$

where C and $\hat{C}$ are two constants.

Proof:

$$
\begin{aligned}
\|u - u_h\|_a^2 &\le \|u - u_I\|_a^2 \\
&\le \int_a^b \Big(p(x)\,(u' - u_I')^2 + q(x)\,(u - u_I)^2 \Big)\, dx \\
&\le \max\{p_{max}, q_{max}\} \int_a^b \Big((u' - u_I')^2 + (u - u_I)^2 \Big)\, dx \\
&\le \max\{p_{max}, q_{max}\}\, \|u''\|_\infty^2 \int_a^b \big(h^2 + h^4/64\big)\, dx \\
&\le Ch^2\|u''\|_\infty \,.
\end{aligned}
$$

The second inequality is obtained because $\| \ \|_a$ and $\| \ \|_1$ are equivalent, so

$$
c\|v\|_a \le \|v\|_1 \le C\|v\|_a \,, \quad \hat{c}\|v\|_1 \le \|v\|_a \le \hat{C}\|v\|_1 \,.
$$

7.5.3 Error estimate in the pointwise norm

We can easily prove the following error estimate.

Theorem 7.6. *For the 1D Sturm-Liouville problem,*

$$
\|u - u_h\|_\infty \le Ch\|u''\|_\infty \,, \tag{7.41}
$$

where C is a constant. The estimate is not sharp (or optimal); or simply it is over-estimated.

We note that u_h' is discontinuous at nodal points, and the infinity norm $\|u' - u_h'\|_\infty$ can only be defined for continuous functions.

Proof:

$$
\begin{aligned}
e_h(x) = u(x) - u_h(x) &= \int_a^x e_h'(t)dt \\
|e_h(x)| &\le \int_a^b |e_h'(t)|dt \\
&\le \left\{ \int_a^b |e_h'|^2 dt \right\}^{1/2} \left\{ \int_a^b 1\, dt \right\}^{1/2} \\
&\le \sqrt{b-a} \left\{ \int_a^b \frac{p}{p_{min}} |e_h'|^2 dt \right\}^{1/2} \\
&\le \sqrt{\frac{b-a}{p_{min}}}\, \|e_h\|_a \\
&\le \sqrt{\frac{b-a}{p_{min}}}\, \|\tilde{e_h}\|_a \\
&\le Ch\, \|u''\|_\infty .
\end{aligned}
$$

Remark 7.1. *Actually, we can prove a better inequality*

$$\|u - u_h\|_\infty \le Ch^2 \|u''\|_\infty ,$$

so the finite element method is second order accurate. This is an optimal (sharp) error estimate.

7.6 Exercises

1. Assuming the number of variables $n = 3$, describe the Sobolev space $H^3(\Omega)$ (*i.e.*, for $m = 3$) in terms of $L^2(\Omega)$, retaining all the terms but not using the multi-index notation. Then using the multi-index notation when applicable, represent the inner product, the norm, the Schwartz inequality, the distance, and the Sobolev embedding theorem in this space.

2. Consider the function $v(x) = |x|^\alpha$ on $\Omega = (-1, 1)$ with $\alpha \in \mathcal{R}$. For what values of α is $v \in H^0(\Omega)$? (Consider both positive and negative α.) For what values is $v \in H^1(\Omega)$? in $H^m(\Omega)$? For what values of α is $v \in C^m(\Omega)$?

   Hint: Make use of the following

   $$|x|^\alpha = \begin{cases} x^\alpha & \text{if } x \ge 0 \\ (-x)^\alpha & \text{if } x < 0\,, \end{cases}$$

   and when k is a non-negative integer note that

   $$|x|^\alpha = \begin{cases} x^{2k} & \text{if } \alpha = 2k \\ 1 & \text{if } \alpha = 0\,; \end{cases}$$

also

$$\lim_{x\to 0} |x|^\alpha = \begin{cases} 0 & \text{if } \alpha > 0 \\ 1 & \text{if } \alpha = 0 \\ \infty & \text{if } \alpha < 0 \end{cases} \quad \text{and} \quad \int_{-1}^{1} |x|^\alpha dx = \begin{cases} \dfrac{2}{\alpha+1} & \text{if } \alpha > -1 \\ \infty & \text{if } \alpha \le -1 \,. \end{cases}$$

3. Are each of the following statements true or false? Justify your answers.

   (a) If $u \in H^2(0,1)$, then u' and u'' are both continuous functions.

   (b) If $u(x,y) \in H^2(\Omega)$, then $u(x,y)$ may not have continuous partial derivatives $\partial u/\partial x$ and $\partial u/\partial y$.

   Does $u(x,y)$ have first and second order *weak* derivatives?
   Is $u(x,y)$ continuous in Ω?

4. Consider the Sturm-Liouville problem

$$-\big(p(x)u(x)'\big)' + q(x)u(x) = f(x)\,, \quad 0 < x < \pi\,,$$

$$\alpha u(0) + \beta u'(0) = \gamma\,, \quad u'(\pi) = u_b\,,$$

$$\text{where} \qquad 0 < p_{min} \le p(x) \le p_{max} < \infty\,,$$

$$0 \le q_{min} \le q(x) \le q_{max} < q_\infty\,.$$

   (a) Derive the weak form for the problem. Define a bilinear form $a(u,v)$ and a linear form $L(v)$ to simplify the weak form. What is the energy norm?

   (b) What kind of restrictions should we have for α, β, and γ in order that the weak form has a solution?

   (c) Determine the space where the solution resides under the weak form.

   (d) If we look for a finite element solution in a finite dimensional space V_h using a conforming finite element method, should V_h be a subspace of C^0, or C^1, or C^2?

   (e) Given a mesh $x_0 = 0 < x_1 < x_2 \cdots < x_{M-1} < x_M = \pi$, if the finite dimensional space is generated by the hat functions, what kind of structure do the local and global stiffness matrix and the load vector have? Is the resulting linear system of equations formed by the global stiffness matrix and the load vector symmetric, positive definite and banded?

5. Consider the two-point BVP

$$-u''(x) = f(x), \quad a < x < b\,, \qquad u(a) = u(b) = 0\,.$$

   Let $u_h(x)$ be the finite element solution using the piecewise linear space (in $H_0^1(a,b)$) spanned by a mesh $\{x_i\}$. Show that

$$\|u - u_h\|_\infty \le Ch^2,$$

   where C is a constant.
   Hint: First show $\|u_h - u_I\|_a = 0$, where $u_I(x)$ is the interpolation function in V_h.