10.3.5. *Error estimates for piecewise polynomial interpolation.* In Math 573, we showed that if u_I is the piecewise linear interpolant of u, i.e., the piecewise linear function satisfying $u_h(x_i) = u(x_i)$, $i = 0, 1, \ldots, N$, then if $u \in C^2[a, b]$, then

$$\max_{[a,b]} |u(x) - u_I(x)| \leq \frac{h^2}{8} \max_{[a,b]} |u''(x)|.$$

Hence, we know that this space can approximate smooth functions well, and by taking the mesh spacing sufficiently small, we can approximate to any desired accuracy. To get an estimate for our method, however, we need error estimates of this type in a different norm.

We now use a different approach to get error estimates in L^2 type norms. For this purpose, it will be convenient to define the space

$$H^2[a, b] = \{v : \int_a^b [v^2 + (v')^2 + (v'')^2] \, dx < \infty\},$$

i.e., the space of functions whose derivatives up to order 2 are square-integrable. We then have the following approximation result,

Theorem 32. *If $u \in H^2[a, b]$ and u_I denotes the piecewise linear interpolant of u on a partition of uniform mesh spacing h, then*

$$\|u - u_I\|_{L^2(a,b)} \leq \frac{h^2}{\sqrt{90}} \|u''\|_{L^2(a,b)}, \qquad \|u' - u_I'\|_{L^2(a,b)} \leq \frac{h}{\sqrt{6}} \|u''\|_{L^2(a,b)}.$$

Proof. To get a result of this type, we use the integral form of the remainder for the Taylor series expansion, i.e., if $P_n(x) = \sum_{k=0}^n f^{(k)}(c)(x - c)^k/k!$, then

$$f(x) - P_n(x) = \frac{1}{n!} \int_c^x (x - t)^n f^{(n+1)}(t) \, dt.$$

We first consider a simple interpolation problem on the interval $[0, 1]$. Let $v_I(s)$ denote the linear function that interpolates $v(s)$ at the points $s = 0$ and $s = 1$, i.e., $v_I(s) = (1 - s)v(0) + sv(1)$. Then, letting $n = 1$, $c = s$ and first $x = 0$ and then $x = 1$, we get

$$v(0) = v(s) + v'(s)(0 - s) + \int_s^0 (0 - t)v''(t) \, dt,$$

$$v(1) = v(s) + v'(s)(1 - s) + \int_s^1 (1 - t)v''(t) \, dt.$$

Hence,

$$\begin{aligned} v(s) - v_I(s) &= v(s) - (1 - s)v(0) - sv(1) \\ &= -(1 - s) \int_s^0 (0 - t)v''(t) \, dt - s \int_s^1 (1 - t)v''(t) \, dt \\ &= -\int_0^s (1 - s)tv''(t) \, dt - \int_s^1 s(1 - t)v''(t) \, dt = \int_0^1 g(s, t)v''(t) \, dt, \end{aligned}$$

where

$$g(s, t) = -(1 - s)t, \quad 0 < t < s, \qquad g(s, t) = -s(1 - t), \quad s < t < 1.$$

Applying the Cauchy-Schwarz inequality, we get

$$\int_0^1 [v(s)-v_I(s)]^2\,ds = \int_0^1 \left(\int_0^1 g(s,t)v''(t)\,dt\right)^2 ds$$
$$\leq \int_0^1 \left(\int_0^1 [g(s,t)]^2\,dt \int_0^1 [v''(t)]^2\,dt\right) ds.$$

Now

$$\int_0^1 [g(s,t)]^2\,dt = \int_0^s (1-s)^2t^2dt + \int_s^1 (s)^2(1-t)^2dt = (1-s)^2s^3/3 + (s)^2(1-s)^3/3 = s^2(1-s)^2/3.$$

Combining these results,we get

$$\int_0^1 [v(s)-v_I(s)]^2\,ds \leq \int_0^1 [v''(t)]^2\,dt \int_0^1 [s^2(1-s)^2/3]\,ds \leq \frac{1}{90}\int_0^1 [v''(s)]^2\,ds.$$

The result we want will follow from this basic result by a scaling argument using a change of variable. More specifically, to get an error estimate on the subinterval $[x_{i-1}, x_i]$, we make the change of variable $x = x_{i-1} + hs$. Setting $u(x) = v(s)$, and noting that $ds = h^{-1}dx$ and $(dv^2/ds^2)(s) = h^2(d^2u/dx^2)(x)$, we get

$$\int_{x_{i-1}}^{x_i} [u(x)-u_I(x)]^2\,dx = h\int_0^1 [v(s)-v_I(s)]^2\,ds \leq \frac{h}{90}\int_0^1 [v''(s)]^2\,ds$$
$$= \frac{h}{90}\int_{x_{i-1}}^{x_i} h^4[u''(x)]^2h^{-1}\,dx = \frac{h^4}{90}\int_{x_{i-1}}^{x_i} [u''(x)]^2\,dx.$$

Then

$$\|u-u_I\|_{L^2}^2 = \sum_{i=1}^n \int_{x_{i-1}}^{x_i} [u(x)-u_I(x)]^2\,dx \leq \sum_{i=1}^n \frac{h^4}{90}\int_{x_{i-1}}^{x_i} [u''(x)]^2\,dx = \frac{h^4}{90}\|u''\|_{L^2}^2.$$

The result follows by taking square roots.

To get the estimate for the derivative, we differentiate the equation $v(s) - v_I(s) = \int_0^1 g(s,t)v''(t)\,dt$ to obtain

$$v'(s) - v_I'(s) = \int_0^1 \frac{\partial g}{\partial s}(s,t)v''(t)\,dt.$$

Then, proceeding as before, we get

$$\int_0^1 [v'(s)-v_I'(s)]^2\,ds = \int_0^1 \left(\int_0^1 \frac{\partial g}{\partial s}(s,t)v''(t)\,dt\right)^2 ds$$
$$\leq \int_0^1 \left(\int_0^1 [\frac{\partial g}{\partial s}(s,t)]^2\,dt \int_0^1 [v''(t)]^2\,dt\right) ds.$$

Now

$$\int_0^1 [\frac{\partial g}{\partial s}(s,t)]^2\,dt = \int_0^s t^2\,dt + \int_s^1 (1-t)^2\,dt = s^3/3 + (1-s)^3/3.$$

Combining these results, we get

$$\int_0^1 [v'(s) - v_I'(s)]^2\,ds \le \int_0^1 [v''(t)]^2\,dt \int_0^1 [s^3 + (1-s)^3]/3\,ds \le \frac{1}{6}\int_0^1 [v''(s)]^2\,ds.$$

Changing variables as before, we obtain

$$\begin{aligned}\int_{x_{i-1}}^{x_i} [u'(x) - u_I'(x)]^2\,dx = h^{-2}h\int_0^1 [v'(s) - v_I'(s)]^2\,ds &\le \frac{1}{6h}\int_0^1 [v''(s)]^2\,ds \\ &= \frac{1}{6h}\int_{x_{i-1}}^{x_i} h^4[u''(x)]^2 h^{-1}\,dx = \frac{h^2}{6}\int_{x_{i-1}}^{x_i} [u''(x)]^2\,dx.\end{aligned}$$

Then

$$\|u' - u_I'\|_{L^2}^2 = \sum_{i=1}^n \int_{x_{i-1}}^{x_i} [u'(x) - u_I'(x)]^2\,dx \le \sum_{i=1}^n \frac{h^2}{6}\int_{x_{i-1}}^{x_i} [u''(x)]^2\,dx = \frac{h^2}{6}\|u''\|_{L^2}^2.$$

The result follows by taking square roots. □

10.3.6. *Error estimates for the finite element method.* We can now easily obtain error estimates for the finite element solution by combining this approximation result with Lemma 6.

Theorem 33. *If $u \in H^2[a,b]$ is the exact solution of the boundary value problem (Problem P) and u_h is the solution of the approximate problem P_h with V_h chosen to be the space of continuous piecewise linear functions, then*

$$\|u - u_h\|_E \le Ch\|u''\|_{L^2(a,b)}.$$

Proof. Choosing $w_h = u_I$ in Lemma 6, we get $\|u - u_h\|_E \le \|u - u_I\|_E$. Let $p^* = \max_{[a,b]} |p(x)|$ and $q^* = \max_{[a,b]} |q(x)|$. Then

$$\|v\|_E^2 = \int_a^b [p(v')^2 + qv^2]\,dx \le p^*\|v'\|_{L^2(a,b)}^2 + q^*\|v\|_{L^2(a,b)}^2.$$

Hence, using Theorem 32, we have for all $0 < h \le h_0$

$$\|u - u_I\|_E^2 \le \left[p^*\frac{h^2}{6} + q^*\frac{h^4}{90}\right]\|u''\|_{L^2(a,b)}^2 \le C^2h^2\|u''\|_{L^2(a,b)}^2,$$

where C depends on p^*, q^*, and h_0, but is independent of u and h. □

Since $\|u - u_I\|_{L^2} \le Ch^2\|u''\|_{L^2}$, one could ask whether the finite element solution also satisfies an estimate of this type.

Lemma 7. $\|u - u_h\|_{L^2} \le Ch\|u - u_h\|_E$.

Proof. We define $w \in V^0$ as the solution of

$$a(w,z) = (u - u_h, z), \quad \text{for all } z \in V^0.$$

One can show that w satisfies $\|w''\|_{L^2} \le C\|u-u_h\|_{L^2}$. Then, since $a(u-u_h, v_h) = 0$ for all $v_h \in V_h^0$, we choose $z = u - u_h$ and $v_h = w_I$ to obtain

$$\begin{aligned}\|u-u_h\|_{L^2}^2 = a(w, u-u_h) = a(w-w_I, u-u_h) &\le \|w-w_I\|_E\|u-u_h\|_E \\ &\le Ch\|w''\|_{L^2}\|u-u_h\|_E \le Ch\|u-u_h\|_{L^2}\|u-u_h\|_E.\end{aligned}$$

Dividing by $\|u-u_h\|_{L^2}$ gives the result. □

Combining results, we get

Theorem 34. $\|u-u_h\|_{L^2} \le Ch^2\|u''\|_{L^2}$.

To get better approximations, we use higher order piecewise polynomials. One can show that if u_I is a continuous piecewise polynomial of degree $\le k$ interpolating u at the mesh points $x_0, \cdots, x_n$ and at $k-1$ distinct points in the interior of each subinterval (x_{i-1}, x_i), $i = 1, \ldots, n$, e.g.,

$$u_I(x_i) = u(x_i), \quad i = 0, 1, \ldots, n,$$
$$u_I(x_{i-1+j/k}) = u(x_{i-1+j/k}), \quad j = 1, 2, \ldots, k-1, \quad i = 1, \ldots, n,$$

then for some constant C independent of h and u,

$$\|u-u_I\|_{L^2(a,b)} \le Ch^{k+1}\|u^{(k+1)}\|_{L^2(a,b)}, \qquad \|u'-u_I'\|_{L^2(a,b)} \le Ch^k\|u^{(k+1)}\|_{L^2(a,b)}.$$

This result easily leads to the error estimates

$$\|u-u_h\|_E \le Ch^k\|u^{(k+1)}\|_{L^2(a,b)}, \qquad \|u-u_h\|_{L^2} \le Ch^{k+1}\|u^{(k+1)}\|_{L^2(a,b)}.$$

10.3.7. *Other boundary conditions.* We now consider how to solve the same equation, but with a different type of boundary condition, i.e., we replace the boundary condition $u(b) = g_b$ by the boundary condition $p(b)u'(b) = g_b$. We could also do a similar replacement at the left end point, but it is more instructive to keep the boundary condition $u(a) = g_a$ to illustrate the difference in the way different types of boundary conditions are handled by the finite element method.

We begin from the integration by parts formula:

$$(Lu, v) = \int_a^b \left[-\frac{d}{dx}\left(p(x)\frac{du}{dx}\right) + q(x)u\right] v\,dx = -pu'v|_a^b + \int_a^b (p(x)u'v' + quv]\,dx.$$

Then,

$$a(u,v) \equiv \int_a^b (p(x)u'v' + quv]\,dx = (Lu, v) + pu'v|_a^b.$$

Now if u is a solution of our new boundary value problem, $Lu = f$ and $p(b)u'(b) = g_b$. Hence, for all sufficiently smooth v satisfying $v(a) = 0$,

$$a(u,v) = (f,v) + g_b v(b). \tag{10.2}$$

Note we no longer take $v(b) = 0$ since if u is the solution of the boundary value problem, we can evaluate the term $p(b)u'(b)v(b)$ for any v, namely it is equal to $g_b v(b)$.

We now reverse this process. If u satisfies (10.2) for all (sufficiently smooth) functions satisfying $v(a) = 0$, then integrating by parts, we see that

$$(Lu, v) + p(b)u'(b)v(b) = (f, v) + g_b v(b).$$

By first considering choices of v for which $v(b) = 0$, we can conclude that $Lu = f$, $a < x < b$. Then by choosing $v(b) = 1$, we conclude that $p(b)u'(b) = g_b$. So, in this case, we recover both the differential equation and one of the boundary conditions. To get u to satisfy the boundary condition $u(a) = g_a$, we must impose it directly.

So we see that if the boundary condition for u in the integration by parts formula is known, (i.e., is one of the boundary conditions in the boundary value problem), then we do not have to impose any condition on the function v and do not have to build that boundary condition into the space in which we look for the solution u, while if the boundary condition for u in the integration by parts formula is not one of the boundary conditions in the boundary value problem, we take $v = 0$ at that point and build that boundary condition into the space in which we look for the solution u.

For example, to define a variational problem corresponding to the boundary value problem we are now considering, we define the spaces

$$W^0 = \{v \in V : v(a) = 0\}, \qquad W^g = \{v \in V : v(a) = g_a\}.$$

Then our new variational formulation is:

Problem Q: Find $u \in W^g$ such that $a(u, v) = (f, v) + g_b v(b)$ for all $v \in W^0$.

The approximate problem is then defined in a similar way. We first define

$$W_h^0 = \{v \in V_h : v(a) = 0\}, \qquad W_h^g = \{v \in V_h : v(a) = g_a\}.$$

Then the approximate problem is:

Problem Q_h: Find $u_h \in W_h^g$ such that $a(u_h, v) = (f, v) + g_b v(b)$ for all $v \in W_h^0$.

The analysis is the same as for the previous problem. The linear system we have to solve will involve one additional unknown, $u_h(b)$.