10.3.2. *Discretized equations.* Before considering finite element spaces using higher order piecewise polynomials, let us look at how one uses piecewise linear functions to get an approximation to the solution to the boundary value problem. Returning to Problem P_h, we need a general expression for the approximate solution $u_h \in V_h^g$. But such a function can be written in the form

$$u_h(x) = \sum_{j=0}^{n} \alpha_j \phi_j(x) = \sum_{i=j}^{n-1} \alpha_j \phi_j(x) + \alpha_0 \phi_0(x) + \alpha_n \phi_n(x) = \sum_{j=1}^{n-1} \alpha_j \phi_j(x) + g_a \phi_0(x) + g_b \phi_n(x),$$

where we have used the facts that $\alpha_j = u_h(x_j)$ and $u_h(a) = g_a$, $u_h(b) = g_b$. Thus, to determine u_h, we need to determine the constants $\alpha_1, \ldots, \alpha_{n-1}$, which are the values of the approximate solution at the interior mesh points. We also need a general expression for a function $v \in V_h^0$. By the same reasoning, this can be written as

$$v(x) = \sum_{i=1}^{n-1} \beta_i \phi_i(x).$$

Next we observe that the variational equation in Problem P_h being true for all $v \in V_h^0$ is equivalent to it being true for the choices $v = \phi_i$, $i = 1, \ldots, n-1$. Since each of these functions $\in V_h^0$, if the equation hold for all $v \in V_h^0$, it certainly holds for these particular choices. On the other hand, if

$$a(u_h, \phi_i) = (f, \phi_i), \quad i = 1, \ldots, n-1,$$

then since any $v \in V_h^0$ can be written as $\sum_{i=1}^{n-1} \beta_i \phi_i(x)$, for some constants β_i, we have

$$a(u_h, v) = \sum_{i=1}^{n-1} \beta_i a(u_h, \phi_i) = \sum_{i=1}^{n-1} \beta_i (f, \phi_i) = (f, v).$$

In the above, we have used the fact that for all $u, v, w \in V$ and all constants α and β,

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

Thus, we may recast our problem in the form: Find $\alpha_1, \cdots, \alpha_{n-1}$ such that

$$\sum_{j=1}^{n-1} a(\phi_j, \phi_i) \alpha_j = (f, \phi_i) - g_a a(\phi_0, \phi_i) - g_b a(\phi_n, \phi_i), \qquad i = 1, \ldots, n-1.$$

If we define a vector $\alpha = (\alpha_1, \ldots, \alpha_{n-1})^T$, an $n-1 \times n-1$ matrix A with entries $A_{ij} = a(\phi_j, \phi_i)$ and a vector F with entries $F_i = (f, \phi_i) - g_a a(\phi_0, \phi_i) - g_b a(\phi_n, \phi_i)$, then Problem P_h is equivalent to solving the linear system $A\alpha = F$.

To see how this system compares to the type of linear system produced by the finite difference method, we compute the entries of the matrix A. The key observation in this computation is that $\phi_i(x) \neq 0$ only on the subintervals $[x_{i-1}, x_i]$ and $[x_i, x_{i+1}]$. Hence $a(\phi_j, \phi_i) = 0$, unless $j = i-1$, $j = i$, or $j = i+1$, since otherwise either ϕ_i or ϕ_j will be zero at each x.

Thus, the matrix A will be tridiagonal, just as it is in the case of finite differences. Now

$$\begin{aligned} a(\phi_{i-1},\phi_i) &= \int_{x_{i-1}}^{x_i} [p\phi'_{i-1}\phi'_i + q\phi_{i-1}\phi_i]\,dx = \int_{x_{i-1}}^{x_i} [p(-1/h^2) + q\phi_{i-1}\phi_i]\,dx, \\ a(\phi_i,\phi_i) &= \int_{x_{i-1}}^{x_{i+1}} [p\phi'_i\phi'_i + q\phi_i\phi_i]\,dx = \int_{x_{i-1}}^{x_{i+1}} [p(1/h^2) + q\phi_i\phi_i]\,dx, \\ a(\phi_{i+1},\phi_i) &= \int_{x_i}^{x_{i+1}} [p\phi'_{i+1}\phi'_i + q\phi_{i+1}\phi_i]\,dx = \int_{x_i}^{x_{i+1}} [p(-1/h^2) + q\phi_{i+1}\phi_i]\,dx. \end{aligned}$$

Note that the limits of integration have been simplified by using the facts that $\phi_{i-1} = 0$ for $x \ge x_i$, $\phi_{i+1} = 0$ for $x \le x_i$, and $\phi_i = 0$ for $x \le x_{i-1}$ and $x \ge x_{i+1}$. When p and q are constants, we find that

$$a(\phi_{i-1},\phi_i) = -p/h + qh/6, \qquad a(\phi_i,\phi_i) = 2p/h + 4qh/6, \qquad a(\phi_{i+1},\phi_i) = -p/h + qh/6.$$

To evaluate these integrals for general p and q, we need to use numerical integration formulas. Two formulas that are exact for linear polynomials are:

$$\begin{aligned} \int_a^b f(x)\,dx &= (b-a)f([a+b]/2) + \frac{(b-a)^3}{24} f''(\eta), \qquad \text{midpoint rule}, \\ \int_a^b f(x)\,dx &= (b-a)[f(a)+f(b)]/2 - \frac{(b-a)^3}{12} f''(\eta), \qquad \text{trapezoidal rule}. \end{aligned}$$

If we use the midpoint rule to evaluate integrals involving p and the trapezoidal rule for the q and f integrals, we get that

$$a(\phi_{i-1},\phi_i) = -p(x_i - h/2)/h, \qquad a(\phi_i,\phi_i) = [p(x_i - h/2) + p(x_i + h/2)]/h + q(x_i)h,$$
$$a(\phi_{i+1},\phi_i) = -p(x_i + h/2)/h, \qquad (f,\phi_i) = hf(x_i).$$

Hence, we get exactly the same equations as in the finite difference scheme.

10.3.3. *Existence and uniqueness of the finite element solution.* As in the case of finite differences, discretizing by finite elements leads to a square linear system of equations. Thus, to show existence, we need to show there can be at most one solution. However, if there are two solutions u_h^1 and u_h^2, then $u_h^1 - u_h^2 \in V_h^0$, and subtracting equations, we find that

$$a(u_h^1 - u_h^2, v) = 0, \qquad \text{for all } v \in V_h^0.$$

Choosing $v = u_h^1 - u_h^2$, and using the facts that $p(x) \ge p_* > 0$ and $q(x) \ge 0$, we have that

$$0 = a(v,v) = \int_a^b [p(v')^2 + qv^2]\,dx \ge p_* \int_a^b (v')^2\,dx.$$

Hence $\int_a^b (v')^2\,dx = 0$ and so $v' = 0$. Hence, v is a constant. But $v \in V_h^0$, so $v(a) = 0$. Hence, $v \equiv 0$. We note that this result will be true for any choice of finite element spaces, since we did not use anything that was particular to piecewise linear functions.

10.3.4. *Error estimates.* For the finite difference method, we derived an estimate for the maximum error at the mesh points. For finite elements, it is more natural to derive an error estimate in the "energy" norm defined by

$$\|v\|_E = [a(v,v)]^{1/2}.$$

However, one can also obtain estimates for $\|u - u_h\|_{L^2(a,b)}$ and for $\max_{[a,b]} |u(x) - u_h(x)|$ and the rate of convergence of the method will depend on the norm in which the error is measured. For example, it is not hard to check that if $f(x) = (1/\sqrt{m})\sin(mx)$, then

$$\|f\|_{L^2(0,\pi)} = \sqrt{\pi/(2m)}, \qquad \|f'\|_{L^2(0,\pi)} = \sqrt{m\pi/2}.$$

Hence, as m increases, the L^2 norm of f decreases, while the L^2 norm of f' (and hence $\|f\|_E$) increases.

It is useful to note that the bilinear form $a(u,v)$ is an inner product on the space V, i.e., it satisfies the following properties for all $u, v, w \in V$ and all scalars c: (i) $a(u,v) = a(v,u)$, (ii) $a(u, v+w) = a(u,v) + a(u,w)$, (iii) $a(cu, v) = ca(u,v)$, (iv) $a(u,u) \geq 0$, and $a(u,u) = 0$ if and only if $u = 0$. Thus, we are led to the energy norm on V in a natural way. Whenever we have an inner product, we have the Cauchy-Schwarz inequality relating the inner product to the norm, i.e.,

Lemma 5.

$$|a(u,v)| \leq \|u\|_E \|v\|_E, \qquad \textit{for all } v \in V.$$

Proof. If u or v is zero, then the lemma is obvious. Assume both are not zero. Since

$$a(\|v\|_E u \pm \|u\|_E v, \|v\|_E u \pm \|u\|_E v) \geq 0,$$

we have by (ii) and (iii) that

$$\|v\|_E^2 a(u,u) \pm \|u\|_E \|v\|_E a(u,v) \pm \|u\|_E \|v\|_E a(v,u) + \|u\|_E^2 a(v,v) \geq 0.$$

Now using (i),

$$2\|v\|_E^2 \|u\|_E^2 \geq \pm 2\|u\|_E \|v\|_E a(u,v)$$

The result follows by dividing by $2\|u\|_E \|v\|_E$. □

Example:

If we apply the Cauchy-Schwarz inequality to the L^2 inner product, $(u,v) = \int_a^b u(x)v(x)\,dx$, we get

$$\left| \int_a^b u(x)v(x)\,dx \right| \leq \left(\int_a^b [u(x)]^2\,dx \right)^{1/2} \left(\int_a^b [v(x)]^2\,dx \right)^{1/2}.$$

We shall use this result frequently in our analysis.

To obtain an error estimate, the first step is to relate the error between u and u_h to the error in the best approximation of u by functions in V_h^g.

Lemma 6.

$$\|u - u_h\|_E \leq \|u - w_h\|_E, \qquad \textit{for all } w_h \in V_h^g.$$

The lemma says that u_h, the approximation produced by the finite element method, is the best approximation to u among elements of V_h^g, if we measure the error in the energy norm.

Proof. The true and approximate solutions u and u_h satisfy the respective equations

$$a(u, v) = (f, v), \qquad \text{for all } v \in V^0,$$
$$a(u_h, v_h) = (f, v_h), \qquad \text{for all } v_h \in V_h^0.$$

Since $V_h^0 \subseteq V^0$, we choose $v = v_h$ and subtract equations to get

$$a(u - u_h, v_h) = 0, \qquad \text{for all } v_h \in V_h^0.$$

Then

$$a(u - u_h, u - u_h) = a(u - u_h, u - w_h) + a(u - u_h, w_h - u_h) = a(u - u_h, u - w_h),$$

since $w_h - u_h \in V_h^0$ (i.e., both functions are in V_h^g and hence their difference is in V_h^0). Hence,

$$\|u - u_h\|_E^2 = a(u - u_h, u - u_h) = a(u - u_h, u - w_h) \le \|u - u_h\|_E \|u - w_h\|_E.$$

The lemma follows by dividing by $\|u - u_h\|_E$. □

This lemma reduces the question of error estimates for the finite element method to a question in approximation theory, i.e., how well can the exact solution u be approximated by functions in V_h^g. Clearly, unless the space is chosen so that it has good approximation properties, the finite element method will not be able to produce a good approximation.