Chapter 11

The boundary value problem

The boundary value problem in one variable is an ordinary differential equation, for which an initial condition is not enough, instead we need to specify boundary conditions at each end of the interval. Contrary to the initial value problem, the dependent variable does not represent time, but should rather we thought of as a spatial coordinate.

11.1 The boundary value problem

The boundary value problem

We consider the following *boundary value problem*, for which we seek a function $u(x) \in \mathcal{C}^2(0,1)$, such that

$$-u''(x) = f(x), \quad x \in (0,1), \tag{11.1}$$
$$u(0) = u(1) = 0, \tag{11.2}$$

given a source term $f(x)$, and *boundary conditions* at the endpoints of the interval $I = [0,1]$.

We want to find an approximate solution to the boundary value problem in the form of a continuous piecewise polynomial that satisfies the boundary conditions (11.2), that is we seek

$$U \in V_h = \{v \in V_h^{(q)} : v(0) = v(1) = 0\}, \tag{11.3}$$

such that the error $e = u - U$ is small in some suitable norm $\|\cdot\|$.

The *residual* of the equation is defined as

$$R(w) = w'' + f, \tag{11.4}$$

with $R(u) = 0$, for $u = u(x)$ the solution of the boundary value problem. Our strategy is now to find an approximate solution $U \in V_h \subset \mathcal{C}^2(0,1)$ such that $R(U) \approx 0$.

We have two natural methods to find a solution $U \in V_h$ with a minimal residual: (i) the *least squares method*, where we seek the solution with the minimal residual measured in the L_2-norm,

$$\min_{U \in V_h} \|R(U)\|, \tag{11.5}$$

and (ii) *Galerkin's method*, where we seek the solution for which the residual is orthogonal the subspace V_h,

$$(R(U), v) = 0, \quad \forall v \in V_h. \tag{11.6}$$

With an approximation space consisting of piecewise polynomials, we refer to the methods as a *least squares finite element method*, and a *Galerkin finite element method*. With a trigonometric approximation space we refer to Galerkin's method as a *spectral method*.

Galerkin finite element method

The finite element method (FEM) based on (11.6) takes the form: find $U \in V_h$, such that

$$\int_0^1 -U''(x)v(x)\,dx = \int_0^1 f(x)v(x)\,dx, \tag{11.7}$$

for all *test functions* $v \in V_h$. For (11.12) to be well defined, we need to be able to represent the second order derivative U'', which is not obvious for low order polynomials, such as linear polynomials, or piecewise constants.

To reduce this constraint, we can use partial integration to move one derivative from the approximation U to the test function v, so that

$$\int_0^1 -U''(x)v(x)\,dx = \int_0^1 U'(x)v'(x)\,dx - [U'(x)v(x)]_0^1 = \int_0^1 U'(x)v'(x)\,dx,$$

since $v \in V_h$, and thus satisfies the boundary conditions. The finite element method now reads: find $U \in V_h$, such that,

$$\int_0^1 U'(x)v'(x)\,dx = \int_0^1 f(x)v(x)\,dx, \tag{11.8}$$

for all $v \in V_h$.

The discrete problem

We now let V_h be the space of continuous piecewise linear functions, that satisfies the boundary conditions (11.2), that is,

$$U \in V_h = \{v \in V_h^{(1)} : v(0) = v(1) = 0\}, \tag{11.9}$$

so that we can write any function $v \in V_h$ as

$$v(x) = \sum_{i=1}^{n} v_i \phi_i(x), \tag{11.10}$$

over a mesh $\mathcal{T}_h$ with n internal nodes x_i, and $v_i = v(x_i)$ since $\{\phi_i\}_{i=1}^n$ is a nodal basis.

We thus search for an approximate solution

$$U(x) = \sum_{j=1}^{n} U_i \phi_j(x), \tag{11.11}$$

with $U_j = U(x_j)$. If we insert (11.10) and (11.11) into (11.12), we get

$$\sum_{j=1}^{n} U_j \int_0^1 \phi_j'(x)\phi_i'(x)\,dx = \int_0^1 f(x)\phi_i(x)\,dx, \quad i = 1, ..., n, \tag{11.12}$$

which corresponds to the matrix equation

$$Sx = b, \tag{11.13}$$

with $s_{ij} = (\phi_j', \phi_i')$, $x_j = U_j$, and $b_i = (f, \phi_i)$. The matrix S is sparse, since $s_{ij} = 0$ for $|i - j| > 1$, and for large n we need to use an iterative method to solve (11.13).

We compute the entries of the matrix S, referred to as a *stiffness matrix*, from the definition of the basis functions (10.31), starting with the diagonal entries,

$$\begin{aligned} s_{ii} &= (\phi_i', \phi_i') = \int_0^1 (\phi_i')^2(x)\,dx = \int_{x_{i-1}}^{x_i} (\lambda_{i,1}')^2(x)\,dx + \int_{x_i}^{x_{i+1}} (\lambda_{i+1,0}')^2(x)\,dx \\ &= \int_{x_{i-1}}^{x_i} \left(\frac{1}{h_i}\right)^2 dx + \int_{x_i}^{x_{i+1}} \left(\frac{1}{h_{i+1}}\right)^2 dx = \frac{1}{h_i} + \frac{1}{h_{i+1}}, \end{aligned}$$

and similarly we compute the off-diagonal entries,

$$s_{ii+1} = (\phi_i', \phi_{i+1}') = \int_0^1 \phi_i'(x)\phi_{i+1}'(x)\,dx = \int_{x_i}^{x_{i+1}} \frac{-1}{h_{i+1}} \frac{1}{h_{i+1}}\,dx = -\frac{1}{h_{i+1}},$$

and

$$s_{ii-1} = (\phi_i', \phi_{i-1}') = \int_0^1 \phi_i'(x)\phi_{i-1}'(x)\,dx = ... = -\frac{1}{h_i}.$$

The variational problem

Galerkin's method is based on the *variational formulation*, or *weak form*, of the boundary value problem, where we search for solution in a vector space V, for which the variational form is well defined: find $u \in V$, such that

$$\int_0^1 u'(x)v'(x)\,dx = \int_0^1 f(x)v(x)\,dx, \tag{11.14}$$

for all $v \in V$.

To construct an appropriate vector space V for (11.1) to be well defined, we need to extend L^2 spaces to include also derivatives, which we refer to as *Sobolev spaces*. We introduce the vector space $H^1(0,1)$, defined by,

$$H^1(0,1) = \{v \in L^2(0,1) : v' \in L^2(0,1)\}, \tag{11.15}$$

and the vector space that also satisfies the boundary conditions (11.2),

$$H_0^1(0,1) = \{v \in H^1(0,1) : v(0) = v(1) = 0\}. \tag{11.16}$$

The variational form (11.1) is now well defined for $V = H_0^1(0,1)$, since

$$\int_0^1 u'(x)v'(x)\,dx \leq \|u'\|\|v'\| < \infty \tag{11.17}$$

by Cauchy-Schwarz inequality, and

$$\int_0^1 f(x)v(x)\,dx \leq \|f\|\|v\| < \infty, \tag{11.18}$$

for $f \in L^2(0,1)$.

Optimality of Galerkin's method

Galerkin's method (11.12) corresponds to searching for an approximate solution in a finite dimensional subspace $V_h \subset V$, for which (11.1) is satisfied for all test functions $v \in V_h$.

The Galerkin solution U is the best possible approximation in V_h, in the sense that,

$$\|u - U\|_E \leq \|u - v\|_E, \quad \forall v \in V_h, \tag{11.19}$$

with the *energy norm* defined by

$$\|w\|_E = \left(\int_0^1 |w'(x)|^2\,dx\right)^{1/2}. \tag{11.20}$$

Thus $U \in V_h$ represents a projection of $u \in V$ onto V_h, with respect to the inner product defined on V,

$$(v, w)_E = \int_0^1 v'(x)w'(x)\, dx, \tag{11.21}$$

with $\|w\|_E^2 = (w, w)_E$. The *Galerkin orthogonality*,

$$(u - U, v)_E = 0, \quad \forall v \in V_h, \tag{11.22}$$

expresses the optimality of the approximation U, as

$$\|u - U\|_E \leq \|u - v\|_E, \quad \forall v \in V_h, \tag{11.23}$$

which follows by

$$\begin{aligned} \|u - U\|_E^2 &= (u - U, u - u_h)_E = (u - U, u - v)_E + (u - U, v - u_h)_E \\ &= (u - U, u - v)_E \leq \|u - U\|_E \|u - v\|_E, \end{aligned}$$

for any $v \in V_h$.