

Jim Lambers
ENERGY 281
Spring Quarter 2007-08
Lecture 12 Notes

These notes were originally written by Tara LaForce.

1 General One Dimensional Problems

At this point we will extend the derivation above to include general linear second-order elliptic ODEs of the form

$$a_o(x) \frac{d^2 u(x)}{dx^2} + a_1(x) \frac{du}{dx} + a_2(x) u(x) = f(x). \quad (1)$$

Recall that this equation is elliptic if a_o never changes sign or vanishes, i.e., $|a_o(x)| > \gamma > 0$ for some constant γ . We will also focus on *two-point boundary-value problems*, which are problems where half of the boundary conditions are specified at each endpoint.

1.1 Flow Through Porous Media

One example in which elliptic boundary value problems arise in PE is as one-dimensional flow through porous media. Start by defining the flux σ of the fluid as $\sigma(x) = -k(x) \frac{du}{dx}$. σ can be a general flux function of any type, but for porous media flow, u is hydraulic head or pressure, σ is the flow rate, k is the absolute permeability, $f(x)$ is the fluid source/sink and may represent wells or a boundary condition with flow across it, such as a constant flux boundary with an aquifer. In porous media flow we are also implicitly assuming the additional equation for Darcy's law holds, that is $\sigma = -ku'$. The permeability does not have to be constant in this formulation, but may vary with x .

1.2 One-Dimensional Heat Loss

One example in which elliptic boundary value problems arise in PE is heat loss from a wellbore. For the moment we can consider only 1D heat loss, but radial heat loss can be approximated using FEM as well. In heat loss, u is the temperature, σ is the heat flux given by Fourier's law, $f(x)$ is the heat

source (in this example, the wellbore), and k is the thermal conductivity of the porous medium.

1.3 Boundary Conditions

The most general boundary conditions that we will consider are

$$\begin{aligned}\alpha_0 \frac{du(0)}{dx} + \beta_0 u(0) &= \gamma_0 \\ \alpha_l \frac{du(l)}{dx} + \beta_l u(l) &= \gamma_l\end{aligned}\tag{2}$$

Only the value of u is specified in a *Dirichlet Boundary Condition*. Only the value of $\frac{du(0)}{dx}$ is specified in a *Neumann Boundary Condition*. The flux $-k(0)\frac{du(0)}{dx} = \sigma_0$ or a linear combination of the flux and u are specified in a *natural boundary condition*.

Now we have a general ODE of the form

$$-\frac{d}{dx} \left(k(x) \frac{du(x)}{dx} \right) + c(x) \frac{du}{dx} + b(x) u(x) = f(x)\tag{3}$$

where the second derivative of u has been replaced by the definition of the flux. This ODE doesn't necessarily have a second derivative at every single point in the domain, since $k(x)$ may not be continuous. u must have a second derivative over every smooth sub-domain Ω_i , but there may be discontinuities at the boundary between two subdomains. We didn't worry about the existence of the second derivative in the model problem in the last lecture because the real physical problems we want to solve have the form of (3), with the boundary conditions defined in (2), not the form of (1) from the previous lecture.

As before, the ODE is multiplied by the test function v and integrated to give the variational form of the two-point boundary-value problem on any smooth domain Ω ,

$$-ku'v|_0^l + \int_{\Omega} (ku'v' + cu'v + buv) dx - \int_{\Omega} fv dx = 0\tag{4}$$

If our domain is not smooth, we can solve this problem over a series of subdomains where the ODE is smooth and sum them. There are three types of discontinuities that are possible at the edges of Ω :

- $k(x)$ is discontinuous, $f(x)$ is continuous at x_1 . This gives a jump condition across the boundary of the element, but since k is finite on both sides of the jump the flux is continuous and $[\sigma(x)] = 0$ at the boundary.
- $f(x)$ is discontinuous, $k(x)$ is continuous at x_2 . This gives a jump condition across the boundary of the element, but as long as f is finite on both sides of the jump the flux is continuous and $[\sigma(x)] = 0$ at the boundary.
- $f(x)$ is discontinuous at x_3 and has a concentrated forcing term given by $-\hat{f}\delta(x - x_3)$ which is not finite. This gives a jump condition across the boundary of the element, and the flux is not continuous and $[\sigma(x)] = \hat{f}$ at the boundary.

As a consequence, when we sum over all of the elements, the variational boundary value problem becomes

$$k(0)u'(0)v(0) - k(l)u'(l)v(l) + \int_0^l (ku'v' + cu'v + buv) dx + [\sigma(x_1)]v(x_1) + [\sigma(x_2)]v(x_2) + [\sigma(x_3)]v(x_3) = \int_0^l fv dx. \quad (5)$$

At the points x_1 and x_2 , the jump condition is zero so those terms drop out, but the jump at x_3 is not zero, so we would have to deal with the \hat{f} term in the FEM. For the rest of the derivation we will assume that we have no discontinuities of this type, but it is important to know that even very badly-behaved domains can easily be handled using FEM.

Rewriting (5) so that the homogeneous ODE is on the left and the forcing and boundary terms are on the right gives

$$\int_0^l (ku'v' + cu'v + buv) dx = -v(0)k(0)[\gamma_0 - \beta_0u(0)]/\alpha_0 + v(l)k(l)[\gamma_l - \beta_lu(l)]/\alpha_l + \int_0^l fv dx + \hat{f}v(x_3). \quad (6)$$

Note that the boundary conditions have been used to eliminate the values of u' on the boundary.

1.4 Galerkin Approximation

Exactly as in the previous lecture, we are looking for a finite set of basis functions $\{\phi_1, \phi_2, \dots, \phi_M\}$ to approximate $u_M \in H^{(M)}$, for some M . We will assume that the interval $[0, l]$ is divided into N elements; as before, the exact

relation between M and N is determined by the boundary conditions. In this case we are no longer requiring that our basis functions have the property $v(0) = v(l) = 0$, because the solution is not necessarily zero at the endpoints, and we want to have the same test and trial functions. The stiffness matrix for (6) is given by

$$K_{ij} = \int_0^l (k\phi_i'\phi_j' + c\phi_j'\phi_i + b\phi_j\phi_i) dx, \quad i, j = 1, \dots, M. \quad (7)$$

where $j = i$ or $j = i \pm 1$ are the only nonzero entries. The load vector is

$$F_i = \int_0^l f\phi_i(x) dx - \phi_i(0)k(0)\gamma_0/\alpha_0 + \phi_i(l)k(l)\gamma_l/\alpha_l \quad (8)$$

where $\phi_i(0) = 0$ for every element except the first and $\phi_i(l) = 0$ for every element except the last. Once again we will be looking only at linear shape functions, but since we are now using v_N that do not have $v(0) = v(l) = 0$, we are going to re-define the shape functions slightly so that the origin of the shifted coordinate system is at the center of the element and $\xi = -1$ at the left endpoint and $\xi = 1$ at the right endpoint. Now the two functions defined on a generic element Ω_e are given by $\hat{\psi}_1(\xi) = 0.5(1 - \xi)$ and $\hat{\psi}_2(\xi) = 0.5(1 + \xi)$. These are exactly the same functions employed in the previous lecture, except that now on the first and last elements the basis functions are no longer zero. The real purpose of redefining the hat functions in this way is that this formulation allows for the easy definition of higher-order approximations.

Each element makes a contribution to the stiffness matrix of the form

$$k_{ij}^e = \int_{s_1^e}^{s_2^e} (k\psi_i^{e'}(\xi)\psi_j^{e'}(\xi) + c\psi_j^{e'}(\xi)\psi_i^e(\xi) + b\psi_i^e(\xi)\psi_j^e(\xi)) d\xi \quad (9)$$

for $i, j = 1, 2$ and contributes to the load vector

$$f_i^e = \int_{s_1^e}^{s_2^e} (f\psi_i^e(\xi)) d\xi. \quad (10)$$

Typically the integrals are computed numerically. We are ignoring the boundary conditions and any discontinuities in the data at this point.

For linear basis functions each element has two nodes that contribute, so the first element has two equations of the form

$$\begin{aligned} k_{11}^1 u_1 + k_{12}^1 u_2 &= f_1^1 \\ k_{21}^1 u_1 + k_{22}^1 u_2 &= f_2^1 \end{aligned} \quad (11)$$

The i th interior element has two equations of the form

$$\begin{aligned} k_{11}^i u_i + k_{12}^i u_{i+1} &= f_1^i \\ k_{21}^i u_i + k_{22}^i u_{i+1} &= f_2^i \end{aligned} \quad (12)$$

This gives a tridiagonal matrix K and a load vector F such that

$$K = \begin{bmatrix} k_{11}^1 & k_{12}^1 & 0 & 0 \\ k_{21}^1 & k_{22}^1 + k_{11}^2 & k_{12}^2 & 0 \\ 0 & k_{21}^2 & k_{22}^2 + k_{11}^3 & k_{12}^3 \\ 0 & 0 & k_{21}^N & k_{22}^N \end{bmatrix}, \quad F = \begin{bmatrix} f_1^1 \\ f_2^1 + f_1^2 \\ f_2^2 + f_1^3 \\ f_2^N \end{bmatrix}, \quad (13)$$

provided that there are no discontinuities in the initial data (see the reference below for how to handle discontinuities). This formulation isn't complete because it doesn't account for any boundary conditions.

1.5 Natural Boundary Conditions

For general natural boundary conditions of the form

$$\begin{aligned} \alpha_0 \frac{du(0)}{dx} + \beta_0 u(0) &= \gamma_0 \\ \alpha_l \frac{du(l)}{dx} + \beta_l u(l) &= \gamma_l \end{aligned} \quad (14)$$

we have $M = N + 1$, and the systems of equations, including the boundary conditions, is

$$\begin{bmatrix} k_{11}^1 - \frac{k(0)\beta_0}{\alpha_0} & k_{12}^1 & 0 & 0 \\ k_{21}^1 & k_{22}^1 + k_{11}^2 & k_{12}^2 & 0 \\ 0 & k_{21}^2 & k_{22}^2 + k_{11}^3 & k_{12}^3 \\ 0 & 0 & k_{21}^N & k_{22}^N + \frac{k(l)\beta_l}{\alpha_l} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_{N+1} \end{bmatrix} = \begin{bmatrix} f_1^1 - \frac{k(0)\gamma_0}{\alpha_0} \\ f_2^1 + f_1^2 \\ f_2^2 + f_1^3 \\ f_2^N + \frac{k(l)\gamma_l}{\alpha_l} \end{bmatrix} \quad (15)$$

1.6 Neumann Boundary Conditions

For Neumann boundary conditions, the equation is identical to (15) with $\beta_o = \beta_l = 0$.

1.7 Dirichlet Boundary Conditions

For Dirichlet boundary conditions, the matrix problem reduces to a smaller problem. Since $u(0)$ and $u(l)$ are both specified, we don't have to solve for them. Therefore, $M = N - 1$, but for consistency with the discussion of the other boundary conditions, we number the basis functions ϕ_2, \dots, ϕ_N , so that the same basis functions are always associated with the same elements. The first and last rows do not need to be included in the system of equations, but the rows of the load vector corresponding to the second and N th basis functions must be adjusted, so the new system is

$$\begin{bmatrix} k_{22}^1 + k_{11}^2 & k_{12}^2 & 0 & 0 \\ k_{21}^2 & k_{22}^2 + k_{11}^3 & k_{12}^3 & 0 \\ 0 & k_{21}^3 & k_{22}^3 + k_{11}^4 & k_{12}^{N-1} \\ 0 & 0 & k_{21}^{N-1} & k_{22}^{N-1} + k_{11}^N \end{bmatrix} \begin{bmatrix} u_2 \\ u_3 \\ u_4 \\ u_N \end{bmatrix} = \begin{bmatrix} f_2^1 + f_1^2 - \frac{k_{21}^1 \gamma_0}{\beta_0} \\ f_2^2 + f_1^3 \\ f_2^3 + f_1^4 \\ f_2^{N-1} + f_1^N - \frac{k_{12}^N \gamma_l}{\beta_l} \end{bmatrix} \quad (16)$$

Any combination of boundary conditions is also possible, and each boundary can be set up as described above independently of the other.

2 Higher-Order Approximations

In general it is possible to use any polynomial function to approximate the function on each element. In practice it is rarely desirable to use much higher than quadratic basis functions because higher-order functions have too much oscillation. In order to define basis functions of order n each element must have $n + 1$ nodes. The i th shape functions for an n th order approximation for the basis functions is

$$\hat{\psi}_i(\xi) = \frac{(\xi - \xi_1)(\xi - \xi_2) \dots (\xi - \xi_{i-1})(\xi - \xi_{i+1}) \dots (\xi - \xi_{n+1})}{(\xi_i - \xi_1)(\xi_i - \xi_2) \dots (\xi_i - \xi_{i-1})(\xi_i - \xi_{i+1}) \dots (\xi_i - \xi_{n+1})} \quad (17)$$

Each of these functions is one at the node ξ_i and zero at ξ_j for $i \neq j$, which implies that they are all linearly independent on the element. There are $n + 1$ linearly independent shape functions on each element. Using this definition with $n = 1$ gives the linear basis functions discussed in section 1.4. For $n = 2$ we have to define three nodes per element, two are at the ends of

the element and one is in the center. Now the three shape functions are $\hat{\psi}_1(\xi) = \frac{1}{2}\xi(\xi - 1)$, $\hat{\psi}_2(\xi) = 1 - \xi^2$, and $\hat{\psi}_3(\xi) = \frac{1}{2}\xi(\xi + 1)$.

For higher-order approximations the matrix K is defined in the same way as for linear elements except that now we have

$$u_M^e(x) = \sum_{j=1}^{N_e} u_j^e \psi_j^e(x) \quad (18)$$

on each element, where N_e is the number of nodes per element. This means that now instead of having to find u_M at two points on every element we need to find u_M at $n + 1$ points, which makes our matrix-vector problem considerably larger. Frequently higher order methods are worth the cost because a higher-order approximation has error of order h^{n+1} , and is still computationally cheaper than dividing the grid into enough elements to get the same error from a linear approximation.

Consider, for example, if the interval $[0, 1]$ is divided up into four elements. A linear approximation would contain 8 basis functions and have an error on the order of $h^2 = 0.25^2 = 0.0625$. A quadratic approximation would contain 12 basis functions and have an error on the order of $h^3 = 0.25^3 = 0.015625$. In order to get an error this small using linear approximations we would have to divide the domain into eight elements so that $h^2 = 0.125^2 = 0.015625$, and we would have to compute 16 basis functions. Though the matrix-vector problem is the same size for linear elements with $h = 0.125$ and quadratic elements with $h = 0.25$, the quadratic formulation requires only 3/4 as many computations in order to assemble the matrix.

References

- [1] Becker, E. B., G. F. Carey, and J. T. Oden, *Finite Elements: an Introduction*, Texas Institute for Computational Mechanics, UT Austin, 1981.