Appendix B
Galerkin Method with Flux Boundary Conditions

Larry C. Young, tildentechnologies.com

Consider a problem approximated by the following trial functions:

$$y(x) = \sum_{i=0}^{n+1} \ell_i(x)\, y_i \tag{B.1}$$

where $\ell_i(x)$ are the Lagrange interpolating polynomials. These functions are then substituted into the differential equation to form the residual as in Eq. (1.3). Assume the boundary conditions are:

$$y = 0 \quad at\ x = 0$$
$$\frac{dy}{dx} + \alpha y = 0 \quad at\ x = 1$$

The usual procedure in orthogonal collocation or pseudospectral methods is to satisfy both of these conditions exactly, i.e. use boundary collocation [Finlayson (1972), p. 101; Villadsen and Michelsen (1978), p. 137; Trefethen (2000), p. 137; Boyd (2000), p. 111, and numerous others].

$$y_0 = 0$$
$$\sum_{i=0}^{n+1} A_{n+1,i} y_i + \alpha y_{n+1} = 0$$

where $\boldsymbol{A}$ is the first derivative matrix. If the boundary values are eliminated, Eq. (B.1) becomes:

$$y(x) = \sum_{i=1}^{n} [\ell_i(x) + b_i \ell_{n+1}(x)]\, y_i \tag{B.2}$$

where $b_i = -A_{n+1,i}/(A_{n+1,n+1} + \alpha)$. The trial functions are now the terms in the brackets of Eq. (B.2) and the Galerkin method weights the residual, R, by these trial functions. If the integrals are approximated by quadrature using all $n + 2$ collocation points, the result is:

$$\sum_{k=0}^{n+1} W_k R(x_k, \boldsymbol{y})[\ell_i(x_k) + b_i \ell_{n+1}(x_k)] = W_i R(x_i, \boldsymbol{y}) + W_{n+1} R(1, \boldsymbol{y}) b_i = 0 \tag{B.3}$$

Eq. (B.3) is not equivalent to a collocation method because of the second term involving the residual at the boundary. Collocation sets only the first interior residual term to zero, so the nonzero second term seriously violates the Galerkin method. The boundary weight, W_{n+1}, is O(n^2), so this inconsistency creates an error over and above that caused by the approximate quadrature. Numerical experiments suggest the error introduced is significant. Since the residual at the boundary decreases exponentially with n, the method with boundary collocation converges exponentially, but often at a much slower rate, especially for fluxes. Boundary collocation also causes errors in the overall mass or energy balance (see Section 3.1.4).

Boundary collocation works well with Gauss points or Radau points with $W_{n+1} = 0$. However, Lobatto points usually produce greater accuracy if boundary conditions are properly treated. Chebyshev points use Clenshaw-Curtis quadrature, which also has a nonzero boundary weight. Although it is a less accurate quadrature, boundary collocation should not be used with Chebyshev points either. W_{n+1} for Clenshaw-Curtis

quadrature is roughly half that for Lobatto quadrature, so the inconsistency is smaller, but still significant.

A far better solution is to treat the boundary condition as a natural condition, so that it becomes part of the approximation. This approach is standard for most Galerkin applications and is described more fully in sections 3.1.3 and 3.1.4. All methods are conservative with a natural treatment. The examples demonstrate that it works very well. Others have proposed similar alternatives to boundary collocation [Canuto, *et al*. (1988), Funaro (1992), Shen and Tang (2006)]; however, these methods appear not to have gained much traction. Mathematicians often pay little attention to the accuracy of fluxes, whereas it is one of the most important results for engineers, since it determines the exchange between the objects being modeled.