Chapter 3

METHODS OF WEIGHTED RESIDUALS

3.1 Introduction

Consider an open and connected set $\Omega \subset \mathbb{R}^n$ with boundary $\partial\Omega$, as in Figure 3.1. A differential operator A involving derivatives up to order p is defined on a function space $\mathcal{U}$, and differential operators B_i, $i = 1, \ldots, k$, involving traces γ_j with $j < p$ are defined on appropriate boundary function spaces. Further, the boundary $\partial\Omega$ is assumed to possess a unique outer unit normal vector $\mathbf{n}$ at every point, and is decomposed (arbitrarily at present) into k parts $\partial\Omega_i$, $i = 1, \ldots, k$, such that

$$\overline{\bigcup_{i=1}^{k} \partial\Omega_i} \;=\; \partial\Omega \;.$$

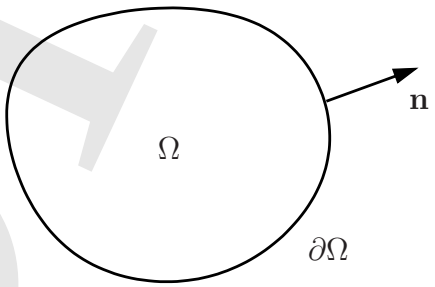


Figure 3.1: *An open and connected domain Ω with smooth boundary written as the union of boundary regions $\partial\Omega_i$*

Given functions f and g_i, $i = 1, \ldots, k$, on Ω and $\partial\Omega_i$, respectively, a mathematical problem associated with a partial differential equation is described by the system

$$\begin{aligned} A[u] &= f \qquad \text{in } \Omega \ , \\ B_i[u] &= g_i \qquad \text{on } \partial\Omega_i \quad , \quad i = 1, \ldots, k \ . \end{aligned} \tag{3.1}$$

With reference to equations (3.1), define functions w_Ω and w_i, $i = 1, \ldots, k$, on Ω and $\partial\Omega_i$, respectively, such that the scalar quantity R, given by

$$R \ = \ \int_\Omega w_\Omega (A[u] \ - \ f) \, d\Omega \ + \ \sum_{i=1}^{k} \int_{\partial\Omega_i} w_i (B_i[u] \ - \ g_i) \, d\Gamma \tag{3.2}$$

be algebraically consistent (i.e., all integrals of the right-hand side have the same units). These functions are called *weighting functions* (or *test functions*).

Equations (3.1) constitute the *strong form* of the differential equation. The scalar equation

$$\int_\Omega w_\Omega (A[u] - f) \, d\Omega \ + \ \sum_{i=1}^{k} \int_{\partial\Omega_i} w_i (B_i[u] - g_i) \, d\Gamma \ = \ 0 \ , \tag{3.3}$$

where functions w_Ω and w_i, $i = 1, \ldots, k$, are arbitrary to within consistency of units and sufficient smoothness for all integrals in (3.3) to exist, is the associated general *weighted-residual form* of the differential equation.

By inspection, the strong form (3.1) implies the general weighted-residual form. The converse is also true, conditional upon sufficient smoothness of the involved fields. The following lemma provides the necessary background for the ensuing proof in the context of $\mathbb{R}^n$.

The localization lemma

Let $f : \ \Omega \mapsto \mathbb{R}$ be a continuous function, where $\Omega \subset \mathbb{R}^n$ is an open set. Then,

$$\int_{\Omega_i} f \, d\Omega \ = \ 0 \ , \tag{3.4}$$

for all open $\Omega_i \subset \Omega$, if, and only if, $f = 0$ everywhere in Ω.

In proving the above lemma, one immediately notes that if $f = 0$, then the integral of f will vanish identically over any Ω_i. To prove the converse, assume by contradiction that there exists a point $\mathbf{x}_0$ in Ω where

$$f(\mathbf{x}_0) \ = \ f_0 \ \neq \ 0 \ , \tag{3.5}$$

and without loss of generality, let $f_0 > 0$. It follows that, since f is continuous, there is an open "sphere" $\mathcal{N} \subset \Omega$ of radius $\delta > 0$ centered at $\mathbf{x}_0$ and defined by

$$\|\mathbf{x} - \mathbf{x}_0\| \ < \ \delta \ , \tag{3.6}$$

such that

$$| \ f(\mathbf{x}) \ - \ f(\mathbf{x}_0) \ | < \ \epsilon \ = \ \frac{f_0}{2} \, , \tag{3.7}$$

for all $\mathbf{x} \in \mathcal{N}$. Thus, it is seen from (3.7) that

$$f(\mathbf{x}) \ > \ \frac{f_0}{2} \tag{3.8}$$

everywhere in $\mathcal{N}$, hence

$$\int_{\mathcal{N}} f \, d\Omega \ > \ \frac{1}{2} \int_{\mathcal{N}} f_0 \, d\Omega \ > \ 0 \ , \tag{3.9}$$

which constitutes a contradiction with the original assumption that the integral of f vanishes identically over all open Ω_i.

Returning to the relation between (3.1) and (3.3), note that since the latter holds for arbitrary choices of w_Ω and w_i, let

$$w_\Omega(\mathbf{x}) \ = \ \begin{cases} 1 & \text{if } \mathbf{x} \in \Omega_i \\ 0 & \text{otherwise} \end{cases} \ , \tag{3.10}$$

for any open $\Omega_i \subset \Omega$, and

$$w_i \ = \ 0 \quad , \quad i = 1, \ldots, k \ . \tag{3.11}$$

Invoking the localization lemma, it is readily concluded that $(3.1)_1$ should hold everywhere in Ω, conditional upon continuity of $A[u]$ and f. Repeating the same process k times (once for each of the boundary conditions) for appropriately defined weighting functions and involving the localization theorem, each one of equations $(3.1)_2$ is recovered on its respective domain.

The equivalence of the strong form and the weighted-residual form plays a fundamental role in the construction of approximate solutions (including finite element solutions) to the underlying problem. Various approximation methods, such as the Galerkin, collocation and least-squares methods, are derived by appropriately restricting the admissible form of the weighting functions and the actual solution.

The above preliminary development applies to linear and non-linear differential operators of any order. A large portion of the forthcoming discussion of weighted-residual methods will involve linear differential equations for which the (linear) operator A contains derivatives of u up to order $p = 2q$, where q is an integer, whereas (linear) operators B_i contain only derivatives of order $0, \ldots, 2q - 1$.

3.2 Galerkin methods

Galerkin methods provide a fairly general framework for the numerical solution of differential equations within the context of the weighted-residual formalization. Here, an introduction to Galerkin methods is attempted by means of their application to the solution of a representative boundary-value problem.

Consider domain $\Omega \subset \mathbb{R}^2$ with smooth boundary $\partial\Omega = \overline{\Gamma_u \cup \Gamma_q}$ and $\Gamma_u \cap \Gamma_q = \emptyset$, as in Figure 3.2. Let the strong form of a boundary-value problem be as follows:

$$\begin{aligned} \frac{\partial}{\partial x_1}(k\,\frac{\partial u}{\partial x_1}) \;+\; \frac{\partial}{\partial x_2}(k\,\frac{\partial u}{\partial x_2}) \;&=\; f \qquad \text{in } \Omega\ , \\ -k\,\frac{\partial u}{\partial n} \;&=\; \bar{q} \qquad \text{on } \Gamma_q\ , \\ u \;&=\; \bar{u} \qquad \text{on } \Gamma_u\ , \end{aligned} \tag{3.5}$$

where $u = u(x_1, x_2)$ is the (yet unknown) solution in Ω. Continuous functions $k = k(x_1, x_2)$

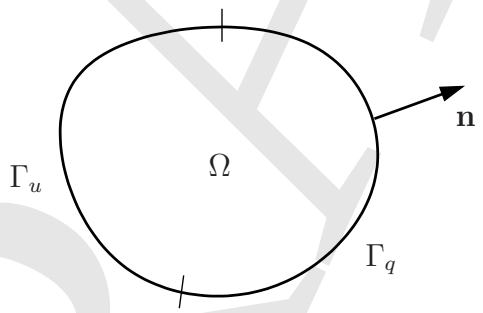


Figure 3.2: *The domain Ω of the Laplace-Poisson equation with Dirichlet boundary Γ_u and Neumann boundary Γ_q*

and $f = f(x_1, x_2)$ defined in Ω, as well as continuous functions $\bar{q} = \bar{q}(x_1, x_2)$ on Γ_q and $\bar{u} = \bar{u}(x_1, x_2)$ on Γ_u are *data* of the problem (i.e., they are known in advance). The boundary conditions $(3.5)_2$ and $(3.5)_3$ are termed *Neumann* and *Dirichlet* conditions, respectively.

It is clear from the statement of the strong form that both the domain and the boundary differential operators are linear in u. This is the *Laplace-Poisson equation*, which has applications in the mathematical modeling of numerous systems in structural mechanics, heat conduction, electrostatics, flow in porous media, etc.

Residual functions R_Ω, R_q and R_u are defined according to

$$\begin{aligned} R_\Omega(x_1,x_2) &= \frac{\partial}{\partial x_1}(k\,\frac{\partial u}{\partial x_1}) + \frac{\partial}{\partial x_2}(k\,\frac{\partial u}{\partial x_2}) - f \qquad \text{in } \Omega\ , \\ R_q(x_1,x_2) &= -k\,\frac{\partial u}{\partial n} - \bar{q} \qquad \text{on } \Gamma_q\ , \\ R_u(x_1,x_2) &= u - \bar{u} \qquad \text{on } \Gamma_u\ . \end{aligned} \tag{3.6}$$

Introducing arbitrary functions $w_\Omega = w_\Omega(x_1,x_2)$ in Ω, $w_q = w_q(x_1,x_2)$ on Γ_q and $w_u = w_u(x_1,x_2)$ on Γ_u, the weighted-residual form (3.3) reads

$$\int_\Omega w_\Omega R_\Omega\,d\Omega + \int_{\Gamma_q} w_q R_q\,d\Gamma + \int_{\Gamma_u} w_u R_u\,d\Gamma = 0\ , \tag{3.7}$$

and, as argued earlier, is equivalent to the strong form of the boundary-value problem, provided that the weighting functions are arbitrary to within unit consistency and proper definition of the integrals in (3.7).

A series of assumptions are introduced in deriving the Galerkin method. First, assume that boundary condition $(3.5)_3$ is satisfied at the outset, namely that the solution u is sought over a set of candidate functions that already satisfy $(3.5)_3$. Hence, the third integral of the left-hand side of (3.7) vanishes and the choice of function w_u becomes irrelevant.

Observing that the two remaining integral terms in (3.7) are consistent unit-wise, provided that w_Ω and w_q have the same units, introduce the second assumption leading to a so-called *Galerkin formulation*: this is a particular choice of functions w_Ω and w_q according to which

$$\begin{aligned} w_\Omega &= w \qquad \text{in } \Omega\ , \\ w_q &= w \qquad \text{on } \Gamma_q\ . \end{aligned} \tag{3.8}$$

Substitution of the above expressions for the weighting functions into the reduced form of (3.7) yields

$$\int_\Omega w\left[\frac{\partial}{\partial x_1}\left(k\,\frac{\partial u}{\partial x_1}\right) + \frac{\partial}{\partial x_2}\left(k\,\frac{\partial u}{\partial x_2}\right) - f\right]\,d\Omega - \int_{\Gamma_q} w\left[k\,\frac{\partial u}{\partial n} + \bar{q}\right]\,d\Gamma = 0\ , \tag{3.9}$$

which, after integration by parts and use of the divergence theorem[1] , is rewritten as

$$-\int_\Omega \left[\frac{\partial w}{\partial x_1}\, k\, \frac{\partial u}{\partial x_1} \;+\; \frac{\partial w}{\partial x_2}\, k\, \frac{\partial u}{\partial x_2} \;+\; wf\right] d\Omega \;+\; \int_{\partial\Omega} wk \left[\frac{\partial u}{\partial x_1} n_1 + \frac{\partial u}{\partial x_2} n_2\right] d\Gamma \\ -\int_{\Gamma_q} w\left[k\,\frac{\partial u}{\partial n} + \bar{q}\right] d\Gamma \;=\; 0\,. \quad (3.10)$$

Recall that the projection of the gradient of u in the direction of the outward unit normal $\mathbf{n}$ is given by

$$\frac{\partial u}{\partial n} \;=\; \frac{du}{d\mathbf{x}}\cdot\mathbf{n} \;=\; \frac{\partial u}{\partial x_1}\, n_1 + \frac{\partial u}{\partial x_2}\, n_2\;, \quad (3.11)$$

and, thus, the above weighted-residual equation is also written as

$$-\int_\Omega \left[\frac{\partial w}{\partial x_1}\, k\, \frac{\partial u}{\partial x_1} \;+\; \frac{\partial w}{\partial x_2}\, k\, \frac{\partial u}{\partial x_2} \;+\; wf\right] d\Omega \;+\; \int_{\Gamma_u} wk\,\frac{\partial u}{\partial n}\, d\Gamma \;-\; \int_{\Gamma_q} w\bar{q}\, d\Gamma \;=\; 0\,. \quad (3.12)$$

Here, an additional assumption is introduced, namely

$$w \;=\; 0 \qquad \text{on } \Gamma_u\,. \quad (3.13)$$

This last assumption leads to the weighted residual equation

$$\int_\Omega \left[\frac{\partial w}{\partial x_1}\, k\, \frac{\partial u}{\partial x_1} \;+\; \frac{\partial w}{\partial x_2}\, k\, \frac{\partial u}{\partial x_2} \;+\; wf\right] d\Omega \;+\; \int_{\Gamma_q} w\bar{q}\, d\Gamma \;=\; 0\;, \quad (3.14)$$

which is identified with the Galerkin formulation of the original problem.

Alternatively, it is possible to assume that <u>both</u> $(3.5)_{2,3}$ are satisfied at the outset and write the weighted residual statement for $w_\Omega = w$ as

$$\int_\Omega w\left[\frac{\partial}{\partial x_1}\left(k\,\frac{\partial u}{\partial x_1}\right) + \frac{\partial}{\partial x_2}\left(k\frac{\partial u}{\partial x_2}\right) \;- f\right] d\Omega \;=\; 0\,. \quad (3.15)$$

Again, integration by parts and use of the divergence theorem transform the above equation into

$$-\int_\Omega \left[\frac{\partial w}{\partial x_1}\, k\, \frac{\partial u}{\partial x_1} \;+\; \frac{\partial w}{\partial x_2}\, k\, \frac{\partial u}{\partial x_2} \;+\; wf\right] d\Omega \;+\; \int_{\partial\Omega} wk\,\frac{\partial u}{\partial n}\, d\Gamma \;=\; 0\;, \quad (3.16)$$

which, in turn, becomes identical to (3.14) by imposing restriction (3.13) and making explicit use of the assumed condition $(3.5)_2$.

[1] This theorem states that given a closed smooth surface $\partial\Omega$ with interior Ω and a C^1 function $f\;:\;\Omega\to\mathbb{R}$, then

$$\int_\Omega f_{,i}\; d\Omega \;=\; \int_{\partial\Omega} f n_i\, d\Gamma\;,$$

where n_i denotes the i-th component of the outer unit normal to $\partial\Omega$.

The weighted residual problem associated with equation (3.14) can be expressed operationally as follows: find $u \in \mathcal{U}$, such that, for all $w \in \mathcal{W}$,

$$B(w,u) \;+\; (w,f) \;+\; (w,\bar{q})_{\Gamma_q} \;=\; 0 \;, \tag{3.17}$$

where

$$\mathcal{U} \;=\; \left\{ u \in H^1(\Omega) \quad | \quad u = \bar{u} \text{ on } \Gamma_u \right\} \;, \tag{3.18}$$

$$\mathcal{W} \;=\; \left\{ w \in H^1(\Omega) \quad | \quad w = 0 \text{ on } \Gamma_u \right\} \;. \tag{3.19}$$

In the above, $B(w,u)$ is a (symmetric) bi-linear form defined as

$$B(w,u) \;=\; \int_\Omega \left(\frac{\partial w}{\partial x_1}\, k\, \frac{\partial u}{\partial x_1} \;+\; \frac{\partial w}{\partial x_2}\, k\, \frac{\partial u}{\partial x_2} \right) d\Omega \;, \tag{3.20}$$

whereas (w,f) and $(w,\bar{q})_{\Gamma_q}$ are linear forms defined respectively as

$$(w,f) \;=\; \int_\Omega w f \, d\Omega \tag{3.21}$$

and

$$(w,\bar{q})_{\Gamma_q} \;=\; \int_{\Gamma_q} w \bar{q} \, d\Gamma \;. \tag{3.22}$$

The identification of admissible solution fields $\mathcal{U}$ and weighting function fields $\mathcal{W}$ is dictated by restrictions placed during the derivation of (3.14) and by the requirement that the bi-linear form $B(w,u)$ be computable (i.e., that the integral be well-defined). Clearly, alternative definitions of $\mathcal{U}$ and $\mathcal{W}$ (with regards to smoothness) can also be acceptable.

A finite-dimensional *Galerkin approximation* of (3.14) is obtained by restating the weighted-residual problem as follows: find $u_h \in \mathcal{U}_h$, such that, for all $w_h \in \mathcal{W}_h$,

$$B(w_h,u_h) + (w_h,f) + (w_h,\bar{q})_{\Gamma_q} = 0 \;, \tag{3.23}$$

where $\mathcal{U}_h$ and $\mathcal{W}_h$ are subspaces of $\mathcal{U}$ and $\mathcal{W}$, respectively, so that

$$\begin{aligned} u \doteq u_h \;&=\; \sum_{I=1}^{N} \alpha_I \, \varphi_I(x_1,x_2) \;+\; \varphi_0(x_1,x_2) \;, \\ w \doteq w_h \;&=\; \sum_{I=1}^{N} \beta_I \, \psi_I(x_1,x_2) \;. \end{aligned} \tag{3.24}$$

In the above, $\varphi_I(x_1,x_2)$ and $\psi_I(x_1,x_2)$, $I = 1,2,\ldots,N$, are given functions (called *interpolation* or *basis functions*), which vanish on Γ_u, and $\varphi_0(x_1,x_2)$ is chosen so that u_h satisfy

boundary condition $(3.5)_3$. Parameters $\alpha_I \in \mathbb{R}$, $I = 1, 2, \ldots, N$, are to be determined by invoking (3.14), while parameters $\beta_I \in \mathbb{R}$, $I = 1, 2, \ldots, N$, are arbitrary.

A *Bubnov-Galerkin* approximation is obtained from (3.24) by setting $\psi_I = \varphi_I$ for all $I = 1, 2, \ldots, N$. This is the most popular version of the Galerkin method. Use of functions $\psi_I \neq \varphi_I$ yields a so-called *Petrov-Galerkin* approximation.

Substitution of u_h and w_h, as defined in (3.24), into the weak form (3.14) results in

$$\sum_{I=1}^{N} \beta_I \int_\Omega \left[\psi_{I,1}\ \psi_{I,2}\right] k \Big(\sum_{J=1}^{N} \begin{bmatrix} \varphi_{J,1} \\ \varphi_{J,2} \end{bmatrix} \alpha_J \ + \ \begin{bmatrix} \varphi_{0,1} \\ \varphi_{0,2} \end{bmatrix} \Big) \, d\Omega + \sum_{I=1}^{N} \beta_I \int_\Omega \psi_I f \, d\Omega + \sum_{I=1}^{N} \beta_I \int_{\Gamma_q} \psi_I \, \bar{q} \, d\Gamma \ = \ 0 \ , \quad (3.25)$$

or, alternatively,

$$\sum_{I=1}^{N} \beta_I \Big(\sum_{J=1}^{N} K_{IJ} \, \alpha_J \ - \ F_I \Big) \ = \ 0 \ , \quad (3.26)$$

where

$$K_{IJ} \ = \ \int_\Omega \left[\psi_{I,1}\ \psi_{I,2}\right] k \begin{bmatrix} \varphi_{J,1} \\ \varphi_{J,2} \end{bmatrix} d\Omega \ , \quad (3.27)$$

and

$$F_I \ = \ - \int_\Omega \psi_I f \, d\Omega \ - \ \int_\Omega \left[\psi_{I,1}\ \psi_{I,2}\right] k \begin{bmatrix} \varphi_{0,1} \\ \varphi_{0,2} \end{bmatrix} d\Omega \ - \ \int_{\Gamma_q} \psi_I \, \bar{q} \, d\Gamma \ . \quad (3.28)$$

Since the parameters β_I are arbitrary, it follows that

$$\sum_{J=1}^{N} K_{IJ} \alpha_J \ = \ F_I \quad , \quad I = 1, 2, \ldots, N \ , \quad (3.29)$$

or, in matrix form,

$$\mathbf{K}\boldsymbol{\alpha} \ = \ \mathbf{F} \ , \quad (3.30)$$

where $\mathbf{K}$ is the $N \times N$ *stiffness matrix* with components given by (3.27), $\mathbf{F}$ is the $N \times 1$ *forcing vector* with components as in (3.28), and $\boldsymbol{\alpha}$ is the $N \times 1$ vector of parameters α_I introduced in $(3.24)_1$.

It is important to note that the Galerkin approximation (3.24) transforms the integro-differential equation (3.14) into a system of linear algebraic equations to be solved for $\boldsymbol{\alpha}$.

Remarks:

- It should be noted that, strictly speaking, $\mathcal{U}$ is not a linear space, since it violates the closure property (see Section 2.1). However, it is easy to reformulate equations (3.5) so that they only involve homogeneous Dirichlet boundary conditions, in which case $\mathcal{U}$ is formally a linear space and $\mathcal{U}_h$ a linear subspace of it. Indeed, any linear partial differential equation of the form

$$A[u] \;=\; f$$

with non-homogeneous boundary conditions

$$u \;=\; \bar{u}$$

on a part of its boundary Γ_u, can be rewritten without loss of generality as

$$A[v] \;=\; f \;-\; A[u_0]$$

with homogeneous boundary conditions on Γ_u, where u_0 is any given function in the domain of u, such that $u_0 = \bar{u}$ on Γ_u.

- It can be easily seen from (3.27) that the stiffness matrix $\mathbf{K}$ is symmetric for a Bubnov-Galerkin approximation. For the same type of approximation, it can be shown that, under mild assumptions, $\mathbf{K}$ is also positive-definite (therefore non-singular), so that the system (3.30) possesses a unique solution.

- Generally, there exists no precisely defined set of assumptions that guarantee the non-singularity of the stiffness matrix $\mathbf{K}$ emanating from a Petrov-Galerkin approximation.

- The terminology "stiffness" matrix and "forcing" vector originates in structural engineering and is associated with the physical interpretation of these quantities in the context of linear elasticity.

Example:
Consider a one-dimensional counterpart of the Laplace-Poisson equation in the form

$$\begin{aligned} \frac{d^2u}{dx^2} \;&=\; 1 \qquad \text{in } \Omega = (0,1)\ , \\ -\frac{du}{dx} \;&=\; 2 \qquad \text{on } \Gamma_q = \{1\}\ , \\ u \;&=\; 0 \qquad \text{on } \Gamma_u = \{0\}\ . \end{aligned}$$

Hence, equation (3.14) takes the form

$$\int_0^1 \left(\frac{dw}{dx}\frac{du}{dx} \;+\; w\right) dx \;+\; 2w\Big|_{x=1} \;=\; 0\;. \qquad (\dagger)$$

A one-parameter Bubnov-Galerkin approximation can be obtained by setting $N = 1$ in equations (3.24) and choosing

$$\varphi_0(x) \;=\; 0$$

and

$$\varphi_1(x) \;=\; x\;.$$

Substituting u_h and w_h into (†) gives

$$\int_0^1 (\beta_1\alpha_1 \;+\; \beta_1 x)\, dx \;+\; 2\beta_1 \;=\; 0\;,$$

and, since β_1 is an arbitrary parameter, it follows that

$$\alpha_1 \;=\; -\frac{5}{2}\;.$$

Thus, the one-parameter Bubnov-Galerkin approximation of the solution to the above differential equation is

$$u_h(x) \;=\; -\frac{5}{2}\,x\;.$$

Similarly, a two-parameter Bubnov-Galerkin approximation is obtained by choosing

$$\varphi_0(x) \;=\; 0$$

and

$$\varphi_1(x) \;=\; x \qquad , \qquad \varphi_2(x) \;=\; x^2\;.$$

Again, (†) implies that

$$\int_0^1 \left[(\beta_1 \;+\; 2\beta_2 x)(\alpha_1 \;+\; 2\alpha_2 x) \;+\; (\beta_1 x \;+\; \beta_2 x^2)\right] dx \;+\; 2(\beta_1 \;+\; \beta_2) \;=\; 0\;,$$

and due to the arbitrariness of β_1 and β_2, one may write

$$\begin{aligned}
\int_0^1 \beta_1(\alpha_1 \;+\; 2\alpha_2 x)\, dx \;&=\; -2\beta_1 \;-\; \int_0^1 \beta_1 x\, dx\;,\\
\int_0^1 \beta_2\, 2x(\alpha_1 \;+\; 2\alpha_2 x)\, dx \;&=\; -2\beta_2 \;-\; \int_0^1 \beta_2 x^2\, dx\;,
\end{aligned}$$

from where it follows that

$$\begin{aligned}
\alpha_1 \;+\; \alpha_2 \;&=\; -\frac{5}{2}\;,\\
\alpha_1 \;+\; \frac{4}{3}\alpha_2 \;&=\; -\frac{7}{3}\;.
\end{aligned}$$

Solving the above linear system yields $\alpha_1 = -3$ and $\alpha_2 = \frac{1}{2}$, so that

$$u_h(x) \;=\; -3x \;+\; \frac{1}{2}\,x^2 \;.$$

It can be easily confirmed by direct integration that the exact solution of the differential equation is identical to the one obtained by the above two-parameter Bubnov-Galerkin approximation. It can be concluded that in this particular problem, the two-dimensional subspace $\mathcal{U}_h$ of all admissible functions $\mathcal{U}$ contains the exact solution, and, also, that the Bubnov-Galerkin method is capable of recovering it.

In the remainder of this section, the Galerkin method is summarized in the context of the model problem

$$\begin{aligned} A[u] \;&=\; f \qquad \text{in } \Omega \;, \\ B[u] \;&=\; g \qquad \text{on } \Gamma_q \;, \\ u \;&=\; \bar{u} \qquad \text{on } \Gamma_u \;, \end{aligned} \tag{3.31}$$

where A is a linear second-order differential operator on a space of admissible domain functions u, and B is a linear first-order differential operator on the space of the traces of u. In addition, it is assumed that $\partial\Omega = \overline{\Gamma_u \cup \Gamma_q}$ and $\Gamma_u \cap \Gamma_q = \emptyset$. The method is based on the construction of a weighted integral form written as

$$\int_\Omega w_\Omega (A[u] \;-\; f)\, d\Omega \;+\; \int_{\Gamma_q} w_q (B[u] \;-\; g)\, d\Gamma \;=\; 0 \;, \tag{3.32}$$

where the space of admissible solutions u satisfies $(3.31)_3$ at the outset. In addition, w_q is chosen to vanish identically on Γ_u and, depending on unit consistency and the particular form of $(3.31)_2$, is chosen to be equal to w (or $-w$) on Γ_q.

3.3 Collocation methods

Collocation methods are based on the idea that an approximate solution to a boundary- or initial-value problem can be obtained by enforcing the underlying equations at suitably chosen points in the domain of analysis. Starting from the general weighted-residual form given in (3.3), assume, as in the Galerkin method, that boundary condition $(3.31)_3$ will be explicitly satisfied by the admissible functions u_h, and obtain the reduced form

$$\int_\Omega w_\Omega (A[u] \;-\; f)\, d\Omega \;+\; \int_{\Gamma_q} w_q (B[u] \;-\; g)\, d\Gamma \;=\; 0 \;, \tag{3.33}$$

for arbitrary functions w_Ω on Ω and w_q on Γ_q. A finite-dimensional admissible field for u_h can be constructed according to

$$u(\mathbf{x}) \doteq u_h(\mathbf{x}) \;=\; \sum_{I=1}^{N} \alpha_I \, \varphi_I(\mathbf{x}) \;+\; \varphi_0(\mathbf{x}) \;, \tag{3.34}$$

with conditions on Γ_u set to $\varphi_I(\mathbf{x}) = 0$, $I = 1, \ldots, N$, and $\varphi_0(\mathbf{x}) = \bar{u}$.

3.3.1 Point-collocation method

First, identify n interior points in Ω with coordinates $\mathbf{x}_i$, $i = 1, \ldots, n$, and $N - n$ boundary points on Γ_q with coordinates $\mathbf{x}_i$, $i = n+1, \ldots, N$. These are referred to as *domain* and *boundary collocation points*, respectively, and are shown schematically in Figure 3.3.

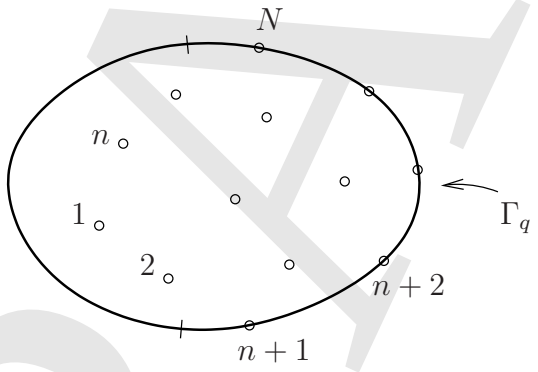


Figure 3.3: *The point-collocation method*

The interior and boundary weighting functions are respectively defined according to

$$w_{\Omega h}(\mathbf{x}) \;=\; \sum_{i=1}^{n} \beta_i \, \delta(\mathbf{x} - \mathbf{x}_i) \tag{3.35}$$

and

$$w_{qh}(\mathbf{x}) \;=\; \rho^2 \sum_{i=n+1}^{N} \beta_i \, \delta(\mathbf{x} - \mathbf{x}_i) \;, \tag{3.36}$$

where the scalar parameter ρ^2 is introduced in w_{qh} for unit consistency. Substitution of

(3.34-3.36) into the weak form (3.33) yields

$$\sum_{i=1}^{n} \beta_i \Big(A[\sum_{I=1}^{N} \alpha_I \, \varphi_I(\mathbf{x}_i) \; + \; \varphi_0(\mathbf{x}_i)] \; - \; f(\mathbf{x}_i) \Big) \\ + \; \rho^2 \sum_{i=n+1}^{N} \beta_i \Big(B[\sum_{I=1}^{N} \alpha_I \, \varphi_I(\mathbf{x}_i) \; + \; \varphi_0(\mathbf{x}_i)] \; - \; g(\mathbf{x}_i) \Big) \; = \; 0 \; . \quad (3.37)$$

Recalling that A and B are linear in u, it follows that

$$\sum_{i=1}^{n} \beta_i \Big(\sum_{I=1}^{N} \alpha_I A[\varphi_I(\mathbf{x}_i)] \; + \; A[\varphi_0(\mathbf{x}_i)] \; - \; f(\mathbf{x}_i) \Big) \\ + \; \rho^2 \sum_{i=n+1}^{N} \beta_i \Big(\sum_{I=1}^{N} \alpha_I B[\varphi_I(\mathbf{x}_i)] \; + \; B[\varphi_0(\mathbf{x}_i)] \; - \; g(\mathbf{x}_i) \Big) \; = \; 0 \; . \quad (3.38)$$

Since the parameters β_i are arbitrary, the above scalar equation results in a system of N linear algebraic equations of the form

$$\sum_{I=1}^{N} K_{iI} \, \alpha_I \; = \; F_i \quad , \quad i = 1, \ldots, N \; , \quad (3.39)$$

where

$$K_{iI} \; = \; \begin{cases} A[\varphi_I(\mathbf{x}_i)] \quad , & 1 \le i \le n \; , \\ \rho^2 B[\varphi_I(\mathbf{x}_i)] \quad , & n+1 \le i \le N \end{cases} \quad , \quad I = 1, \ldots, N \; , \quad (3.40)$$

and

$$F_i \; = \; \begin{cases} -A[\varphi_0(\mathbf{x}_i)] + f(\mathbf{x}_i) \quad , & 1 \le i \le n \; , \\ -\rho^2 \big(B[\varphi_0(\mathbf{x}_i)] - g(\mathbf{x}_i) \big) \quad , & n+1 \le i \le N \end{cases} \; . \quad (3.41)$$

These equations are solved for the parameters α_I, so that the approximate solution u_h is obtained from (3.34).

The particular choice of admissible fields renders the integrals in (3.33) well-defined, since products of Dirac-delta functions (from w_h) and smooth functions (from u_h) are always properly integrable.

Example:
Consider the partial differential equation

$$\frac{\partial^2 u}{\partial x_1^2} \; + \; \frac{\partial^2 u}{\partial x_2^2} \; = \; -1 \qquad \text{in } \Omega = \{(x_1, x_2) \quad | \quad | \, x_1 \, | \le 1 \; , \; | \, x_2 \, | \le 1\} \; , \\ \frac{\partial u}{\partial n} \; = \; 0 \qquad \text{on } \partial\Omega \; .$$

The domain of the problem is sketched in Figure 3.4. It is immediately concluded that the boundary $\partial\Omega$ does not possess a unique outward unit normal at points $(\pm 1, \pm 1)$. It can be shown, however, that this difficulty can be surmounted by a limiting process, thus rendering the present method of analysis valid.

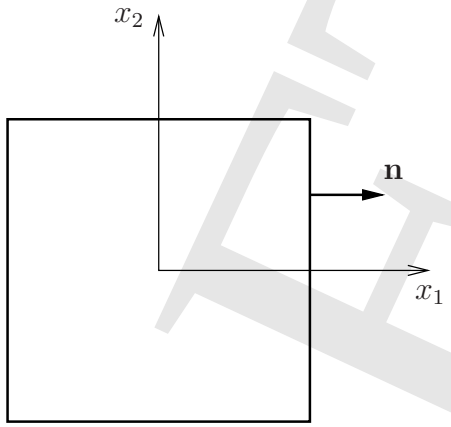


Figure 3.4: *The point collocation method in a square domain*

The above boundary-value problem is referred to as a *Neumann problem*. It is easily concluded that the solution of this above problem is defined only to within an arbitrary constant, i.e., if $u(x_1, x_2)$ is a solution, then so is $u(x_1, x_2) + c$, where c is any constant.

In order to simplify the analysis, use is made of a one-parameter space of admissible solutions which satisfies all boundary conditions. To this end, write u_h as

$$u_h(x_1, x_2) \;=\; \alpha_1(1 \;-\; x_1^2)^2(1 \;-\; x_2^2)^2 \;.$$

It is easy to show that

$$\frac{\partial^2 u_h}{\partial x_1^2} \;+\; \frac{\partial^2 u_h}{\partial x_2^2} \;=\; -4\alpha_1\big[(1-3x_1^2)(1-x_2^2)^2 \;+\; (1-x_1^2)^2(1-3x_2^2)\big] \;. \qquad (\dagger)$$

Noting that the solution should be symmetric with respect to axes $x_1 = 0$ and $x_2 = 0$, pick the single interior collocation point to be located at the intersection of these axes, namely at $(0, 0)$. It follows from (†) that

$$K_{11}\,\alpha_1 \;=\; F_1 \;,$$

where $K_{11} = -8$ and $F = -1$, so that $\alpha_1 = \frac{1}{8}$ and the approximate solution is

$$u_h \;=\; \frac{1}{8}(1 \;-\; x_1^2)^2(1 \;-\; x_2^2)^2 \;.$$

Alternatively, one may choose to start with a one-parameter space of admissible solutions which satisfies the domain equation everywhere, and enforce the boundary conditions at a single point on the boundary. For example, let

$$u_h(x_1, x_2) \;=\; -\frac{1}{4}(x_1^2 \;+\; x_2^2) \;+\; \alpha_1(x_1^4 \;+\; x_2^4 - 6x_1^2x_2^2) \;,$$

and choose to satisfy the boundary condition at point $(1,0)$ (thus, due to symmetry, also at point $(-1,0)$). It follows that

$$\frac{\partial u_h}{\partial n}(1,0) \;=\; \frac{\partial u_h}{\partial x_1}(1,0) \;=\; -\frac{1}{2} \;+\; 4\alpha_1 \;=\; 0 \;,$$

hence $\alpha_1 = \frac{1}{8}$, and

$$u_h(x_1,x_2) \;=\; -\frac{1}{4}(x_1^2 \;+\; x_2^2) \;+\; \frac{1}{8}(x_1^4 \;+\; x_2^4 - 6x_1^2x_2^2) \;.$$

A combined domain and boundary point collocation solution can be obtained by starting with a two-parameter approximation function

$$u_h(x_1,x_2) \;=\; \alpha_1(x_1^2+x_2^2) + \alpha_2(1 \;-\; x_1^2)(1 \;-\; x_2^2)$$

and selecting one interior and one boundary collocation point. In particular, taking $(0,0)$ to be the interior collocation point leads to the algebraic equation

$$\alpha_1 - \alpha_2 \;=\; -1/4 \;.$$

Subsequently, choosing $(1,0)$ as the boundary collocation point yields

$$\alpha_1 - \alpha_2 \;=\; 0 \;.$$

Clearly the system of the preceding two equations is singular, which means here that the two collocations points, in effect, generate conflicting restrictions for the two-parameter approximation function. In such a case, one may choose an alternative boundary collocation point, e.g., $(1,1/\sqrt{2})$, which results in the equation

$$2\alpha_1 - \alpha_2 \;=\; 0 \;,$$

which, when solved simultaneously with the equation obtained from interior collocation, leads to

$$\alpha_1 \;=\; 1/4 \qquad , \qquad \alpha_2 \;=\; 1/2 \;,$$

hence,

$$u_h(x_1,x_2) \;=\; \frac{1}{4}(x_1^2+x_2^2) + \frac{1}{2}(1 \;-\; x_1^2)(1 \;-\; x_2^2) \;.$$

3.3.2 Subdomain-collocation method

A generalization of the point-collocation method is obtained as follows: let Ω_i, $i = 1,\ldots,n$, and $\Gamma_{q,i}$, $i = n+1,\ldots,N$, be mutually disjoint connected subsets of the domain Ω and the boundary Γ_q, respectively, as in Figure 3.5. It follows that

$$\bigcup_{i=1}^{n} \Omega_i \;\subset\; \Omega \tag{3.42}$$

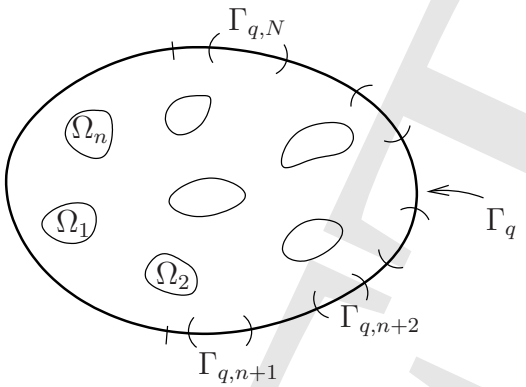


Figure 3.5: *The subdomain-collocation method*

and

$$\bigcup_{i=n+1}^{N} \Gamma_{q,i} \;\subset\; \Gamma_q \;. \tag{3.43}$$

Recall the weighted residual form (3.33) and define the weighting function on Ω as

$$w_{\Omega h}(\mathbf{x}) \;=\; \sum_{i=1}^{n} \beta_i w_{\Omega,i}(\mathbf{x}) \;, \tag{3.44}$$

with

$$w_{\Omega,i}(\mathbf{x}) \;=\; \begin{cases} 1 & \text{if } \mathbf{x} \in \Omega_i \;, \\ 0 & \text{otherwise} \end{cases} \;. \tag{3.45}$$

Similarly, write on Γ_q

$$w_{qh}(\mathbf{x}) \;=\; \sum_{i=n+1}^{N} \beta_i w_{q,i}(\mathbf{x}) \;, \tag{3.46}$$

with

$$w_{q,i}(\mathbf{x}) = \begin{cases} \rho^2 & \text{if } \mathbf{x} \in \Gamma_{q,i} \;, \\ 0 & \text{otherwise} \end{cases} \;. \tag{3.47}$$

Given the above weighting functions, the weighted-residual form (3.33) is rewritten as

$$\sum_{i=1}^{n} \int_{\Omega_i} \beta_i (A[u] \;-\; f)\, d\Omega \;+\; \sum_{i=n+1}^{N} \rho^2 \int_{\Gamma_{q,i}} \beta_i (B[u] \;-\; g)\, d\Gamma \;=\; 0 \;. \tag{3.48}$$

Substitution of u_h from (3.34) into the above weak form yields

$$\sum_{i=1}^{n} \beta_i \int_{\Omega_i} \big(A[\sum_{I=1}^{N} \alpha_I \, \varphi_I(\mathbf{x}) \;+\; \varphi_0(\mathbf{x})] \;-\; f \big) \, d\Omega \\ +\; \rho^2 \sum_{i=n+1}^{N} \beta_i \int_{\Gamma_{q,i}} \big(B[\sum_{I=1}^{N} \alpha_I \varphi_I(\mathbf{x}) \;+\; \varphi_0(\mathbf{x})] \;-\; g \big) \, d\Gamma \;=\; 0 \; . \quad (3.49)$$

Invoking the linearity of A and B in u, the above equation can be also written as

$$\sum_{i=1}^{n} \beta_i \int_{\Omega_i} \big(\sum_{I=1}^{N} \alpha_I A[\varphi_I(\mathbf{x})] \;+\; A[\varphi_0(\mathbf{x})] \;-\; f \big) \, d\Omega \\ +\; \rho^2 \sum_{i=n+1}^{N} \beta_i \int_{\Gamma_{q,i}} \big(\sum_{I=1}^{N} \alpha_I B[\varphi_I(\mathbf{x})] \;+\; B[\varphi_0(\mathbf{x})] \;-\; g \big) \, d\Gamma \;=\; 0 \; , \quad (3.50)$$

from where it can be concluded that, since β_i are arbitrary,

$$\sum_{I=1}^{N} K_{iI} \, \alpha_I \;=\; F_i \quad , \quad i = 1, \ldots, N \; , \quad (3.51)$$

where

$$K_{iI} \;=\; \begin{cases} \displaystyle\int_{\Omega_i} A[\varphi_I(\mathbf{x}_i)] \, d\Omega \quad , & 1 \le i \le n \; , \\ \rho^2 \displaystyle\int_{\Gamma_{q,i}} B[\varphi_I(\mathbf{x}_i)] \, d\Gamma \quad , & n+1 \le i \le N \end{cases} \quad , \quad I = 1, \ldots, N \; , \quad (3.52)$$

and

$$F_i \;=\; \begin{cases} \displaystyle\int_{\Omega_i} \big(-A[\varphi_0(\mathbf{x}_i)] + f(\mathbf{x}_i) \big) \, d\Omega \quad , & 1 \le i \le n \; , \\ \rho^2 \displaystyle\int_{\Gamma_{q,i}} \big(-B[\varphi_0(\mathbf{x}_i)] + g(\mathbf{x}_i) \big) \, d\Gamma \quad , & n+1 \le i \le N \end{cases} \; . \quad (3.53)$$

Remarks:

- The point-collocation method requires very small computational effort to form the stiffness matrix and the forcing vector.
- Collocation methods generally lead to unsymmetric stiffness matrices.
- Choice of collocation points is not arbitrary – for certain classes of differential equations, one may identify collocation points that yield optimal accuracy of the approximate solution.

3.4 Least-squares methods

Consider the model problem (3.31) of the previous section, and assuming that $(3.31)_3$ is satisfied by the space of admissible solutions u, form the "least-squares" functional $I[u]$, defined as

$$I[u] = \int_\Omega (A[u] \ - \ f)^2 \, d\Omega \ + \ \rho^2 \int_{\Gamma_q} (B[u] - g)^2 \, d\Gamma \ , \tag{3.54}$$

where ρ is a consistency parameter. Clearly, the functional attains an absolute minimum at the solution of (3.31). In order to find the extrema of the functional defined in (3.54), determine its first variation as

$$\delta I[u] \ = \ 2 \int_\Omega (A[u] \ - \ f) \, \delta(A[u] \ - \ f) \, d\Omega \ + \ 2\rho^2 \int_{\Gamma_q} (B[u] \ - \ g) \, \delta(B[u] \ - \ g) \, d\Gamma \ . \tag{3.55}$$

Since A and B are linear in u, it is easily seen that

$$\begin{aligned} \delta A[u] \ &= \ \lim_{\omega \to 0} \frac{A[u \ + \ \omega \, \delta u] \ - \ A[u]}{\omega} \\ &= \ \lim_{\omega \to 0} \frac{A[u] \ + \ \omega A[\delta u] \ - \ A[u]}{\omega} \ = \ A[\delta u] \ . \end{aligned} \tag{3.56}$$

Consequently, extremization of $I[u]$ requires that

$$\int_\Omega A[\delta u](A[u] \ - \ f) \, d\Omega \ + \ \rho^2 \int_{\Gamma_q} B[\delta u](B[u] \ - \ g) \, d\Gamma \ = \ 0 \ . \tag{3.57}$$

At this stage, introduce the finite-dimensional approximation for u_h as in (3.34), and, in addition, write

$$\delta u(\mathbf{x}) \doteq \delta u_h(\mathbf{x}) \ = \ \sum_{I=1}^{N} \delta\alpha_I \, \varphi_I(\mathbf{x}) \ , \tag{3.58}$$

with $\delta\alpha_I$, $I = 1, \ldots, N$, being arbitrary scalar parameters. Use of u_h and δu_h in (3.54) results in

$$\begin{aligned} &\int_\Omega A[\sum_{I=1}^{N} \delta\alpha_I \, \varphi_I(\mathbf{x})] \big(A[\sum_{J=1}^{N} \alpha_J \, \varphi_J(\mathbf{x}) \ + \ \varphi_0(\mathbf{x})] \ - \ f \big) \, d\Omega \\ &\qquad + \ \rho^2 \int_{\Gamma_q} B[\sum_{I=1}^{N} \delta\alpha_I \, \varphi_I(\mathbf{x})] \big(B[\sum_{J=1}^{N} \alpha_J \, \varphi_J(\mathbf{x}) \ + \ \varphi_0(\mathbf{x})] \ - \ g \big) \, d\Gamma \ = \ 0 \ . \end{aligned} \tag{3.59}$$

Since A and B are linear in u, it follows that the above equation can be also written as

$$\int_\Omega \sum_{I=1}^{N} \delta\alpha_I A[\varphi_I(\mathbf{x})]\big(\sum_{J=1}^{N} \alpha_J A[\varphi_J(\mathbf{x})] \;+\; A[\varphi_0(\mathbf{x})] \;-\; f\big)\, d\Omega \\ +\; \rho^2 \int_{\Gamma_q} \sum_{I=1}^{N} \delta\alpha_I B[\varphi_I(\mathbf{x})]\big(\sum_{J=1}^{N} \alpha_J B[\varphi_J(\mathbf{x})] \;+ B[\varphi_0(\mathbf{x})] \;-\; g\big)\, d\Gamma \;=\; 0\;, \quad (3.60)$$

which, in turn, invoking the arbitrariness of $\delta\alpha_I$, gives rise to a system of linear algebraic equations of the form

$$\sum_{J=1}^{N} K_{IJ}\,\alpha_J \;=\; F_I \quad , \quad I = 1, \ldots, N\;, \quad (3.61)$$

where

$$K_{IJ} \;=\; \int_\Omega A[\varphi_I] A[\varphi_J]\, d\Omega \;+\; \rho^2 \int_{\Gamma_q} B[\varphi_I] B[\varphi_J]\, d\Gamma \quad , \quad I, J = 1, \ldots, N \quad (3.62)$$

and

$$F_I \;=\; \int_\Omega A[\varphi_I]\big(f \;-\; A[\varphi_0]\big)\, d\Omega \;+\; \rho^2 \int_{\Gamma_q} B[\varphi_I]\big(g \;-\; B[\varphi_0]\big)\, d\Gamma \quad , \quad I = 1, \ldots, N\;. \quad (3.63)$$

It is important to note that the smoothness requirements for the admissible functions u are governed by the integrals that appear in (3.54). It can be easily deduced that if A is a differential operator of second order (i.e., maps functions u to partial derivatives of second order), then for (3.54) to be well-defined, it is necessary that $u \in H^2(\Omega)$. This requirement can be contrasted to the one obtained in the Galerkin approximation of (3.5), where it was concluded that both u and w need only belong to $H^1(\Omega)$.

Remarks:

- The stiffness matrix that emanates from the least-squares functional is symmetric by construction and, may be positive-definite, conditional upon the particular form of the boundary conditions.

- A slightly more general weighted-residual formulation of the least-squares problem based directly on (3.33) is recovered as follows: choose $w_\Omega = A[w]$, $w_q = B[w]$ and set $w = 0$ on Γ_u. Then, the weak form in (3.57) is reproduced, where w appears in place of δu.

3.5 Composite methods

The Galerkin, collocation and least-squares methods can be appropriately combined to yield composite weighted residual methods. The choice of admissible weighting functions defines the degree and form of blending between the above methods. Without attempting to provide an exhaustive presentation, note that a simple Galerkin/collocation method can be obtained for the model problem (3.31), with associated weighted-residual form (3.32), by defining the admissible solutions as in (3.34) and setting

$$w_\Omega(\mathbf{x}) \doteq w_{\Omega h}(\mathbf{x}) = \sum_{I=1}^{m} \beta_I \psi_I(\mathbf{x}) + \sum_{I=m+1}^{n} \beta_I \rho_1^2 \delta(\mathbf{x} - \mathbf{x}_I) , \tag{3.64}$$

where ψ_I, $I = 1, \ldots, N$ vanish on Γ_u. In addition, on Γ_q,

$$w_q(\mathbf{x}) \doteq w_{qh}(\mathbf{x}) = \sum_{I=1}^{m} \beta_I \psi_I(\mathbf{x}) + \sum_{I=n+1}^{N} \beta_I \rho_2^2 \delta(\mathbf{x} - \mathbf{x}_I) , \tag{3.65}$$

where ρ_1 and ρ_2 are scaling factors.

A simple collocation/least-squares method can be similarly obtained by defining the domain and boundary weighting functions according to

$$w_\Omega(\mathbf{x}) \doteq w_{\Omega h}(\mathbf{x}) = \sum_{I=1}^{m} \beta_I A[\psi_I(\mathbf{x})] + \sum_{I=m+1}^{n} \beta_I \rho_1^2 \delta(\mathbf{x} - \mathbf{x}_I) \tag{3.66}$$

and

$$w_q(\mathbf{x}) \doteq w_{qh}(\mathbf{x}) = \sum_{I=1}^{m} \beta_I B[\psi_I(\mathbf{x})] + \sum_{I=n+1}^{N} \beta_I \rho_2^2 \delta(\mathbf{x} - \mathbf{x}_I) , \tag{3.67}$$

respectively, where, again, ψ_I, $I = 1, \ldots, N$, vanish on Γ_u.

3.6 An interpretation of finite difference methods

Finite-difference methods can be interpreted as weighted residuals methods. In particular, the difference operators can be viewed as differential operators over appropriately chosen polynomial spaces. As a demonstration of this interpretation, consider the boundary-value problem (1.1), and let grid points x_i, $i = 1, \ldots, N$, be chosen in the interior of the domain

$(0, L)$, as in Section 1.2.2. The system of equations

$$\begin{aligned} u_2 - 2u_1 &= \frac{f_1 \Delta x^2}{k} - u_0 \ , \\ u_{l+1} - 2u_l + u_{l-1} &= \frac{f_l \Delta x^2}{k} \quad , \quad l = 2, \ldots, N-1 \ , \\ -2u_N + u_{N-1} &= \frac{f_N \Delta x^2}{k} - u_L \end{aligned} \tag{3.68}$$

is obtained by applying the centered-difference operator

$$\frac{d^2 u}{dx^2} \doteq \frac{u_{l+1} - 2u_l + u_{l-1}}{\Delta x^2} \tag{3.69}$$

at all interior points. Note that in the above equations $(\cdot)_l = (\cdot)(x_l)$.

In order to analyze the above finite-difference approximation, consider the domain-based weighted-residual form

$$\int_0^L w \left(k \frac{d^2 u}{dx^2} - f \right) dx = 0 \ , \tag{3.70}$$

where boundary conditions $(1.1)_2$ and $(1.1)_3$ are assumed to hold at the outset. Subsequently, define the approximate solution u_h *within each sub-domain* $(x_l - \frac{\Delta x}{2}, x_l + \frac{\Delta x}{2}]$, $l = 2, \ldots, N-1$, as

$$u_h(x) = \sum_{i=l-1}^{l+1} N_i(x) \alpha_i \ , \tag{3.71}$$

where N_i are polynomials of degree 2 defined as

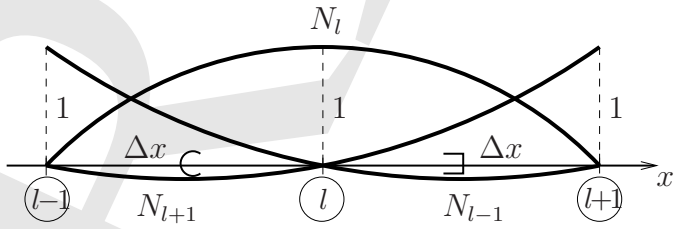


Figure 3.6: *Polynomial interpolation functions used in the weighted-residual interpretation of the finite difference method*

$$\begin{aligned} N_{l-1}(x) &= \frac{(x - x_l)(x - x_{l+1})}{2\Delta x^2} \ , \\ N_l(x) &= -\frac{(x - x_{l-1})(x - x_{l+1})}{\Delta x^2} \ , \\ N_{l+1}(x) &= \frac{(x - x_{l-1})(x - x_l)}{2\Delta x^2} \ . \end{aligned} \tag{3.72}$$

Figure 3.6 illustrates the three interpolation functions in the representative sub-domain. It can be readily seen from the definition of N_{l-1}, N_l and N_{l+1} that $\alpha_l = u_l$, which implies that the parameters α_l can be interpreted as the values of the dependent variable at $x = x_l$, therefore

$$u_h(x) \;=\; \sum_{i=l-1}^{l+1} N_i(x) u_i \;. \tag{3.73}$$

Likewise, the function u_h is given in the sub-domains $[0, x_1 + \frac{\Delta x}{2}]$ and $(x_N - \frac{\Delta x}{2}, L]$ by

$$\begin{aligned} u_h(x) \;&=\; N_0(x) u_0 \;+\; \sum_{i=1}^{2} N_i(x) u_i \;, \\ u_h(x) \;&=\; \sum_{i=N-1}^{N} N_i(x) u_i \;+\; N_{N+1}(x) u_L \;, \end{aligned} \tag{3.74}$$

respectively, so that the boundary conditions are satisfied at both end-points.

The approximation for w_h over the domain is taken in the form

$$w_h \;=\; \sum_{l=1}^{N} \beta_l \, \delta(x \;-\; x_l) \;, \tag{3.75}$$

so that equation (3.70) is written as

$$\sum_{l=1}^{N} \beta_l \left(k \frac{d^2 u_h}{dx^2}(x_l) \;-\; f(x_l) \right) \;=\; 0 \;. \tag{3.76}$$

It follows from (3.73) and (3.72) that in the representative sub-domain

$$\frac{d^2 u_h}{dx^2} \;=\; \frac{u_{l-1}}{\Delta x^2} \;-\; 2\frac{u_l}{\Delta x^2} \;+\; \frac{u_{l+1}}{\Delta x^2} \;. \tag{3.77}$$

This, in turn, implies that at $x = x_l$

$$\frac{k}{\Delta x^2}(u_{l-1} \;-\; 2u_l \;+\; u_{l+1}) \;-\; f_l \;=\; 0 \;, \tag{3.78}$$

owing to the arbitrariness of parameters β_l, $l = 1, \ldots, N$, in (3.76). Similarly, in sub-domain $[0, x_1 + \frac{\Delta x}{2}]$,

$$\frac{k}{\Delta x^2}(u_0 \;-\; 2u_1 \;+\; u_2) \;-\; f_1 \;=\; 0 \;, \tag{3.79}$$

and, in sub-domain $(x_N - \frac{\Delta x}{2}, L]$,

$$\frac{k}{\Delta x^2}(u_{N-1} \;-\; 2u_N \;+\; u_L) \;-\; f_N \;=\; 0 \;. \tag{3.80}$$

Thus, the finite difference equations are recovered exactly at all interior grid points.

The traditional distinction between the finite difference and the finite element method is summarized by noting that finite differences approximate differential operators by (algebraic) difference operators which apply on admissible fields $\mathcal{U}$, whereas finite elements use the exact differential operators which apply only on subspaces of these admissible fields. The weighted-residual framework allows for a unified interpretation of both methods.

Remark:

- ☛ The choice of admissible fields $\mathcal{U}_h$ and $\mathcal{W}_h$ is legitimate, since the integral on the left-hand side of (3.70) is always well-defined.

It is instructive at this point to review a finite element solution of the same problem (1.1), which can be deduced using a Bubnov-Galerkin formulation. To this end, assume that the interpolation functions in (x_{l-1}, x_{l+1}) are piecewise-linear polynomials, namely

$$\begin{aligned}
\varphi_{l-1}(x) &= \begin{cases} -\dfrac{x}{\Delta x} & x < 0 \\ 0 & x \geq 0 \end{cases} , \\
\varphi_{l}(x) &= \begin{cases} 1 + \dfrac{x}{\Delta x} & x < 0 \\ 1 - \dfrac{x}{\Delta x} & x \geq 0 \end{cases} \quad , \quad , \\
\varphi_{l+1}(x) &= \begin{cases} 0 & x < 0 \\ \dfrac{x}{\Delta x} & x \geq 0 \end{cases} ,
\end{aligned} \tag{3.81}$$

see Figure 3.7. Note that here the coordinate system x is centered at point l, without any loss of generality. Then, letting

$$u_h = \sum_{i=l-1}^{l+1} u_i \varphi_i(x) \quad , \quad w_h = \sum_{i=l-1}^{l+1} w_i \varphi_i(x) \tag{3.82}$$

in (x_{l-1}, x_{l+1}), it follows that

$$\frac{du_h}{dx} = \begin{cases} \dfrac{u_l - u_{l-1}}{\Delta x} & x < 0 \\ \dfrac{u_{l+1} - u_l}{\Delta x} & x \geq 0 \end{cases} \tag{3.83}$$

and, likewise,

$$\frac{dw_h}{dx} = \begin{cases} \dfrac{w_l - w_{l-1}}{\Delta x} & x < 0 \\ \dfrac{w_{l+1} - w_l}{\Delta x} & x \geq 0 \end{cases} . \tag{3.84}$$

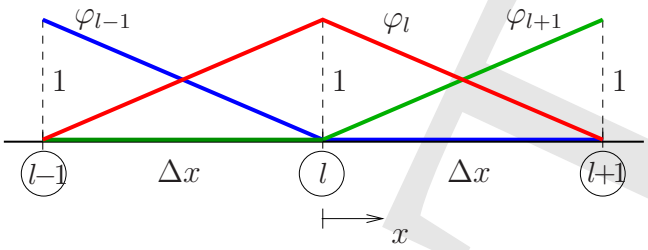


Figure 3.7: *Interpolation functions for a finite element approximation of a one-dimensional two-cell domain*

Neglecting any boundary conditions at x_{l-1} and x_{l+1}, one may write the weak form as

$$\int_{-\Delta x}^{\Delta x} (\frac{dw_h}{dx} k \frac{du_h}{dx} + w_h f)\, dx \;=\; 0 \;, \tag{3.85}$$

which, upon substituting (3.83) and (3.84) in (3.85), becomes

$$\begin{aligned} &\int_{-\Delta x}^{0} \left[\frac{(w_l - w_{l-1})k(u_l - u_{l-1})}{\Delta x^2} + \left\{ -\frac{x}{\Delta x} w_{l-1} + \left(1 + \frac{x}{\Delta x}\right) w_l \right\} f \right] dx \\ &\quad + \int_{0}^{\Delta x} \left[\frac{(w_{l+1} - w_l)k(u_{l+1} - u_l)}{\Delta x^2} + \left\{ \left((1 - \frac{x}{\Delta x}\right) w_l + \frac{x}{\Delta x} w_{l+1} \right\} f \right] dx \;=\; 0 \;. \end{aligned} \tag{3.86}$$

Setting $w_{l-1} = w_{l+1} = 0$, it follows that

$$\begin{aligned} w_l \int_{-\Delta x}^{0} \left[\frac{k}{\Delta x^2}(u_l - u_{l-1}) + (1 + \frac{x}{\Delta x}) f \right] dx \\ + w_l \int_{0}^{\Delta x} \left[-\frac{k}{\Delta x^2}(u_{l+1} - u_l) + (1 - \frac{x}{\Delta x}) f \right] dx \;=\; 0 \;. \end{aligned} \tag{3.87}$$

Next, define the equivalent force f_l as

$$\Delta x f_l \;=\; \int_{-\Delta x}^{0} (1 + \frac{x}{\Delta x}) f \, dx + \int_{0}^{\Delta x} (1 - \frac{x}{\Delta x}) f \, dx \;. \tag{3.88}$$

Subsequently, recalling that w_l is arbitrary and integrating (3.87) leads to (3.78). This means that, under the definition in (3.88), the finite element solution of (1.1) with piecewise linear polynomial interpolations coincides with the finite difference solution that uses the classical difference operator (1.2). The latter can be equivalently thought of as a weighted residual method with piecewise quadratic approximations for the dependent variable u and delta function approximations for the corresponding weighting functions. This observation is consistent with the findings in Sections 1.2.2 and 1.2.3.

3.7 Exercises

Problem 1

Consider the boundary-value problem

$$\begin{aligned} \frac{d^2u}{dx^2} + u + x &= 0 \quad \text{in } \Omega = (0,1) , \\ u &= 1 \quad \text{on } \Gamma_u = \{0\} , \\ \frac{du}{dx} &= 0 \quad \text{on } \Gamma_q = \{1\} . \end{aligned}$$

Assume a general three-parameter polynomial approximation to the exact solution, in the form

$$u_h(x) = \alpha_0 + \alpha_1 x + \alpha_2 x^2 . \qquad (\dagger)$$

(a) Place a restriction on parameters α_i by enforcing only the Dirichlet boundary condition, and obtain a *Bubnov-Galerkin* approximation of the solution.

(b) Starting from the general quadratic form of u_h in (†), place a restriction on parameters α_i as in part (a), and determine a *Petrov-Galerkin* approximation of the solution assuming

$$w_h(x) = \beta_1\psi_1(x) + \beta_2\psi_2(x) ,$$

where functions $\psi_1(x)$ and $\psi_2(x)$ are defined as

$$\psi_1(x) = \begin{cases} 0 & \text{for } 0 \le x \le \frac{1}{2} \\ 1 & \text{for } \frac{1}{2} < x \le 1 \end{cases} ,$$

$$\psi_2(x) = \begin{cases} x & \text{for } 0 \le x \le \frac{1}{2} \\ 0 & \text{for } \frac{1}{2} < x \le 1 \end{cases} .$$

Clearly justify the admissibility of w_h for the proposed approximation.

(c) Starting again from the general quadratic form of u_h in (†), enforce all boundary conditions on u_h and uniquely determine a one-parameter *point-collocation* approximation.

Problem 2

The stiffness matrix $\boldsymbol{K} = [K_{IJ}]$ emanating from a Bubnov-Galerkin approximation of the boundary-value problem

$$\begin{aligned} \frac{\partial}{\partial x_1}\Big(k\frac{\partial u}{\partial x_1}\Big) + \frac{\partial}{\partial x_2}\Big(k\frac{\partial u}{\partial x_2}\Big) &= f \quad \text{in } \Omega \subset \mathbb{R}^2 , \\ u &= \bar{u} \quad \text{on } \Gamma_u , \\ -k\frac{\partial u}{\partial n} &= \bar{q} \quad \text{on } \Gamma_q \end{aligned}$$

has components K_{IJ} given by

$$K_{IJ} = \int_\Omega \left\{ \begin{array}{cc} \varphi_{I,1} & \varphi_{I,2} \end{array} \right\} k \left\{ \begin{array}{c} \varphi_{J,1} \\ \varphi_{J,2} \end{array} \right\} d\Omega ,$$

where the approximation for u is of the general form

$$u(x_1, x_2) \doteq u_h(x_1, x_2) = \sum_{I=1}^{N} \alpha_I \varphi_I(x_1, x_2) + \varphi_0(x_1, x_2) ,$$

and the functions φ_I, $I = 1, \ldots, N$, are assumed linearly independent. Show that $\boldsymbol{K}$ is positive-definite in $\mathbb{R}^N$, provided $k > 0$ and $\Gamma_u \neq \emptyset$.

Problem 3

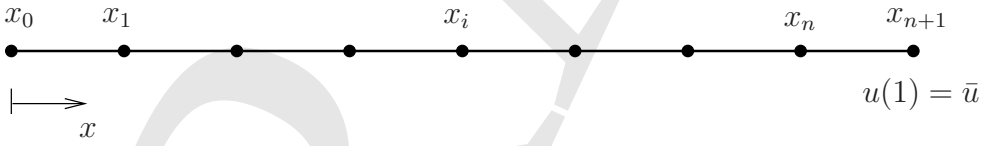
Consider the weak form of the one-dimensional Laplace equation

$$\int_\Omega k w_{,x} u_{,x} \, d\Omega + \int_\Omega w f \, d\Omega + w(0)\bar{q} = 0 ,$$

where $\Omega = (0, 1)$ and $k > 0$ is a constant. Assume that the admissible fields $\mathcal{U}$ and $\mathcal{W}$ for u and w, respectively, allow the first derivative of u and w to exhibit finite jumps at points $x_i \in \Omega$, $i = 1, 2, \ldots, n$, and show that

$$\sum_{i=0}^{n} \int_{x_i}^{x_{i+1}} w(k\, u_{,xx} - f) \, d\Omega + w(0)[k u_{,x}(0) - \bar{q}] + \sum_{i=1}^{n} w(x_i) k [u_{,x}(x_i^+) - u_{,x}(x_i^-)] = 0 .$$

From the above equation derive the strong form of the problem. What can you conclude regarding smoothness of the exact solution across the x_i's?



Problem 4

Consider the boundary-value problem

$$\frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2} = -1 \quad \text{in} \quad \Omega = (-1, 1) \times (-1, 1) ,$$
$$u + \frac{\partial u}{\partial n} = 0 \quad \text{on} \quad \partial\Omega ,$$

with reference to a fixed Cartesian coordinate system $x_1 - x_2$. The boundary condition on $\partial\Omega$ is referred to as a *Robin* (or *third type*) boundary condition.

(a) Conclude that the (unknown) exact solution is symmetric with respect to lines $x_1 = 0$, $x_2 = 0$, $x_1 - x_2 = 0$ and $x_1 + x_2 = 0$.

(b) Start from the general two-dimensional polynomial field which is complete up to degree 6 and show that using only the aforementioned symmetries of the exact solution, the polynomial approximation is reduced to

$$\begin{aligned} u_h \;&=\; \alpha_0 \;+\; \alpha_1\,(x_1^2 \;+\; x_2^2) \;+\; \alpha_2\,x_1^2x_2^2 \;+\; \alpha_3\,(x_1^4 \;+\; x_2^4) \\ &\quad +\; \alpha_4\,(x_1^4x_2^2 \;+\; x_1^2x_2^4) \;+\; \alpha_5\,(x_1^6 \;+\; x_2^6)\;, \end{aligned}$$

where $\alpha_i,\ i = 0, 1, ..., 5$, are arbitrary constants. Subsequently, apply the boundary conditions on u_h and, thus, place restrictions on parameters α_i.

(c) Find two non-trivial approximate solutions to the above problem by means of a one-parameter and a two-parameter interior collocation method using appropriate approximation functions from the family of functions obtained in part (b). Pick collocation points judiciously for both approximations.

Problem 5

Consider the initial-value problem

$$\begin{aligned} \frac{du}{dt} \;+\; u \;&=\; t \quad \text{in } \Omega = (0,1)\;, \\ u(0) \;&=\; 1\;, \end{aligned}$$

and assume a general three-parameter polynomial approximation u_h written as

$$u_h(t) \;=\; \alpha_0 \;+\; \alpha_1 t \;+\; \alpha_2 t^2\;,$$

where α_i, $i = 0, 1, 2$, are scalar parameters to be determined.

(a) Place a restriction on u_h by directly enforcing the initial condition, and, subsequently, obtain an approximate solution to the problem using the point-collocation method. Select the collocation points judiciously.

(b) Place the same restriction on u_h as in part (a), and obtain an approximate solution to the problem using a Bubnov-Galerkin method.

Problem 6

Consider the non-linear second-order ordinary differential equation of the form

$$A[u] \;=\; f \quad \text{in } \Omega = (0,1)\;,$$

where

$$A[u] \;=\; -2u\,\frac{d^2u}{dx^2} \;+\; \Big(\frac{du}{dx}\Big)^2$$

and

$$f \;=\; 4\;,$$

with boundary conditions $u(0) = 1$ and $u(1) = 0$. Find each of the approximate polynomial solutions u_h, as instructed in the problem statement, and compute the *residual* error norm $\mathcal{E}$ defined as

$$\mathcal{E}(u_h) \;=\; \|A[u_h] \;-\; f\|_{L_2}\;.$$

Compare the approximate solutions by means of $\mathcal{E}$. Comment on the results of your analysis.

Problem 7

Consider the partial differential equation

$$\frac{\partial^2 u}{\partial x^2} \; - \; \frac{\partial u}{\partial t} \; = \; 0 \quad \text{for } (x,t) \in (0,1) \times (0,T) \; , \tag{‡}$$

subject to boundary conditions

$$u(0,t) \; = \; 0 \quad \text{for } t \in (0,T) \; , \tag{††}$$

$$\frac{\partial u}{\partial x}(1,t) \; = \; 0 \quad \text{for } t \in (0,T) \; , \tag{‡‡}$$

and initial condition

$$u(x,0) \; = \; 1 \quad \text{for } x \in (0,1) \; , \tag{†‡}$$

where $T > 0$. Let a family of approximations $u_h(x,t)$ be written as

$$u_h(x,t) \; = \; \{\alpha_0 \; + \; \alpha_1 x \; + \; \alpha_2 x^2\} \, \theta(t) \; , \tag{‡†}$$

where $\theta(t)$ is a (yet unknown) function of time, and α_0, α_1 and α_2 are scalar parameters to be determined.

(a) Obtain a reduced form of $u_h(x,t)$ by enforcing boundary conditions (††) and (‡‡).

(b) Determine an initial condition $\theta(0)$ for function $\theta(t)$ by forcing the reduced form of $u_h(x,t)$ obtained in part (a) to satisfy equation (†‡) in the sense of the least-squares method.

(c) Arrive at a first-order ordinary differential equation for function $\theta(t)$ by applying to differential equation (‡) a Petrov-Galerkin method in the spatial domain. Use the approximation function $u_h(x,t)$ of part (a) and a weighting function $w_h(x,t)$ given by

$$w_h(x,t) \; = \; x \; .$$

Find a closed-form expression for $\theta(t)$ by solving the differential equation analytically and using the initial condition obtained in part (b).

The solution procedure outlined above, where a partial differential equation is reduced into an ordinary differential equation by an approximation of the form (‡†), is referred to as the *Kantorovich method*.

Problem 8

Consider the differential equation

$$\frac{\partial^2 u}{\partial x_1^2} \; + \; 2 \, \frac{\partial^2 u}{\partial x_2^2} \; = \; -2$$

in the square domain $\Omega = \{(x_1, x_2) \quad | \quad \mid x_1 \mid < 1 \; , \; \mid x_2 \mid < 1\}$, with homogeneous Dirichlet boundary condition $u = 0$ everywhere on $\partial\Omega$.

(a) Reformulate the above boundary-value problem in terms of a new dependent variable v defined as

$$v \;=\; u \;+\; \frac{1}{2}x_1^2 \;+\; \frac{1}{4}x_2^2 \;.$$

Verify that the resulting partial differential equation in v is homogeneous, while the boundary condition is non-homogeneous and expressed as

$$v \;=\; \bar{v} \;=\; \frac{1}{2}x_1^2 \;+\; \frac{1}{4}x_2^2 \quad \text{on } \partial\Omega \;.$$

(b) Introduce a one-parameter family of approximate solutions v_h of the form

$$v_h \;=\; \alpha\,\varphi \;,$$

where

$$\varphi \;=\; x_1^2 \;-\; \frac{1}{2}x_2^2 \;,$$

and show that v_h satisfies the homogeneous partial differential equation obtained in part (a). Subsequently, determine the scalar parameter α using a weighted-residual method on $\partial\Omega$ by requiring that

$$\int_{\partial\Omega} w_u\,(v_h \;-\; \bar{v})\,d\Gamma \;=\; 0 \;,$$

where $w_u = \beta$, and β is an arbitrary scalar.

3.8 Suggestions for further reading

Sections 3.1-3.5

[1] B.A. Finlayson and L.E. Scriven. The method of weighted residuals – a review. *Appl. Mech. Rev.*, 19:735–748, 1966. [*This is an excellent review of weighted residual methods, including a discussion of their relation to variational methods*].

[2] G.F. Carey and J.T. Oden. *Finite Elements: a Second Course*, volume II. Prentice-Hall, Englewood Cliffs, 1983. [*This volume discusses the Galerkin method in Chapter 1 and the other weighted residual methods in Chapter 4*].

[3] O.D. Kellogg. *Foundations of Potential Theory*. Dover, New York, 1953. [*Chapter IV of this book contains an excellent discussion of the divergence theorem for domains with boundaries that possess corners*].

Section 3.4

[1] P.P. Lynn and S.K. Arya. Use of the least-squares criterion in the finite element formulation. *Int. J. Num. Meth. Engr.*, 6:75–88, 1973. [*This article uses the least-squares method for the solution of the two-dimensional Laplace-Poisson equation, in conjunction with the finite element method for the construction of the admissible fields*].

Section 3.6

[1] O.C. Zienkiewicz and K. Morgan. *Finite Elements and Approximation.* Wiley, New York, 1983. [*The relation between finite element and finite difference methods is addressed in Section 3.10*].

[2] K.W. Morton. Finite Difference and Finite Element Methods. *Comp. Phys. Comm.*, 12:99-108, (1976). [*This article presents a comparison between finite difference and finite element methods from a finite difference viewpoint*].