Chapter 3

Weighted residual methods

3.1. Introduction

In this chapter we introduce variational or weighted residual formulation of PDEs. To facilitate the discussion, we limit ourselves to spaces of continuous functions where the notion of pointwise evaluation is well-defined. This can be generalized considerably using Lebesgue integration and Sobolev spaces; however, we defer that development to later chapters.

3.2. Strong formulation

Let $\Omega \subset \mathbb{R}^d$ (open) and consider a PDE of order m in residual form: find $u \in \mathcal{U} \subset \mathcal{C}^m(\Omega)$ such that

$$R[u] = 0, \tag{3.1}$$

where $R : \mathcal{C}^m(\Omega) \to \mathcal{C}^0(\Omega)$ is a differential operator of order m. This is called the *strong formulation* of the PDE because it enforces the governing equations pointwise throughout the domain and has strict regularity requirements on the solution (m continuous derivatives). The PDE is equipped with essential BCs along $\partial\Omega_D \subset \partial\Omega$ and natural BCs along $\partial\Omega_N \subset \partial\Omega$, where $\partial\Omega = \overline{\partial\Omega_D \cup \partial\Omega_N}$; essential and natural boundary conditions are defined in Section 3.5.2. The space $\mathcal{U} \subset \mathcal{C}^m(\Omega)$, called the *solution* or *trial space*, consists of functions in $\mathcal{C}^m(\Omega)$ that satisfy the boundary conditions of the PDE

$$\mathcal{U} := \{u \in \mathcal{C}^m(\Omega) \mid u \text{ satisfies BCs on } \partial\Omega\}. \tag{3.2}$$

Notice that $\mathcal{U}$ is *not*, in general, a linear subspace of $\mathcal{C}^m(\Omega)$ due to the requirement that functions satisfy the BCs, e.g., suppose a boundary condition states $u = g \neq 0$ on $\partial\Omega$, then $u_1, u_2 \in \mathcal{U}$ are such that $u_1(x) = u_2(x) = g(x)$ for $x \in \partial\Omega$, but $(u_1 + u_2)(x) = u_1(x) + u_2(x) = 2g(x) \implies u_1 + u_2 \notin \mathcal{U}$ and therefore $\mathcal{U}$ is not a linear space (not closed under addition). Under certain conditions, e.g., the PDE is linear, the trial space is affine (Section 2.3.8), i.e., $\mathcal{U} = \varphi + \mathcal{U}^0$ where $\varphi \in \mathcal{U}$ is arbitrary and $\mathcal{U}^0$ is a linear space. The linear space $\mathcal{U}^0$ associated with the affine subspace $\mathcal{U}$ is the set of functions that satisfy the *homogeneous form* of the BCs

$$\mathcal{U}^0 := \{u \in \mathcal{C}^m(\Omega) \mid u \text{ satisfies homogeneous BCs on } \partial\Omega\}. \tag{3.3}$$

Example 3.1: Strong formulation of second-order PDE in one dimension

For concreteness consider the following second-order PDE ($m = 2$) in one dimension ($d = 1$) over the unit interval $\Omega := (0, L)$:

$$R[u] := -\frac{d}{dx}\left[a\frac{du}{dx}\right] - f, \tag{3.4}$$

where $a : \Omega \to \mathbb{R}$ and $f : \Omega \to \mathbb{R}$ are known smooth functions, with BCs

$$u(0) = u_0, \qquad \left(a\frac{du}{dx}\right)_{x=L} = Q_L, \tag{3.5}$$

where $u_0, Q_L \in \mathbb{R}$ are known scalars. The first condition is an essential BC and the second is a natural BC, which implies $\partial\Omega_D = \{0\}$ and $\partial\Omega_N = \{1\}$. The affine trial space for this problem is

$$\mathcal{U} := \left\{ u \in \mathcal{C}^2(\Omega) \;\middle|\; u(0) = u_0, \left(a\frac{du}{dx}\right)_{x=L} = Q_L \right\} \tag{3.6}$$

and the corresponding linear space is

$$\mathcal{U}^0 := \left\{ u \in \mathcal{C}^2(\Omega) \;\middle|\; u(0) = 0, \left(a\frac{du}{dx}\right)_{x=L} = 0 \right\}. \tag{3.7}$$

While the strong formulation is easy to understand and usually relates to physical principles (conservation of mass, momentum, energy), it is not always convenient to use as a foundation for numerical methods.

Example 3.2: Strong formulation is not always suitable as foundation of numerical methods

Consider the following problem: find $u \in \mathcal{C}^2(\Omega)$ where $\Omega := (0, \pi/2)$ such that

$$-\frac{d}{dx}\left[e^x \frac{du}{dx}(x)\right] = \sin(x) \tag{3.8}$$

holds for all $x \in \Omega$ and satisfies the boundary conditions

$$u(0) = 1, \quad \left[e^x \frac{du}{dx}(x)\right]_{x=\pi/2} = 0. \tag{3.9}$$

This fits the general form of (3.4) with $a(x) = e^x$, $f(x) = \sin(x)$, $L = \pi/2$, $u_0 = 1$, $Q_L = 0$. The solution of this boundary value problem can be determined by direct integration

$$u(x) = \frac{1}{2}\left[3 + e^{-x}(\sin(x) - \cos(x))\right]. \tag{3.10}$$

However, a numerical method cannot search the infinite-dimensional trial space

$$\mathcal{U} := \left\{ u \in \mathcal{C}^2(\Omega) \;\middle|\; u(0) = 1, \left[e^x \frac{du}{dx}(x)\right]_{x=\pi/2} = 0. \right\} \tag{3.11}$$

for the solution, so we choose to approximate the solution in a finite-dimensional space. We choose an approximation $u_h \in \mathcal{P}^3(\Omega)$ as

$$u(x) \approx u_h(x) := 1 + c_1(x^2 - \pi x) + c_2\left(x^3 - \frac{3\pi^2}{4}x\right), \tag{3.12}$$

where $c_1, c_2 \in \mathbb{R}$ are unknown scalars, to ensure $u_h \in \mathcal{U}$ (satisfies the boundary conditions of (3.8)): $u_h(0) = 1$ and $[e^x u_h'(x)]_{x=\pi/2} = 0$. Since the u_h satisfies the boundary conditions (3.9), if we can determine $c_1, c_2 \in \mathbb{R}$ such that u_h satisfies the PDE in (3.8), u_h will be a solution of the boundary value problem. To determine the unknown scalars, we substitute the expression for u_h into the governing equation to yield the following equation:

$$2\left(\frac{\pi}{2} - 1 - x\right)c_1 - 3\left(x^2 + 2x - \frac{\pi^2}{4}\right)c_2 = e^{-x}\sin(x)$$

Unfortunately these equations are *inconsistent*, i.e., there are no $c_1, c_2 \in \mathbb{R}$ that makes this equation true for all $x \in (0, \pi/2)$. This shows that, by using the strong formulation, we cannot find a solution to (3.8) of the form (3.12). It is not surprising that this approach failed: we are requiring the PDE be satisfied pointwise, but using an approximation that cannot represent its solution. To avoid this issue, we turn to variational formulations of the PDE, which will be the foundation of a number of numerical methods, including the finite element method.

3.3. Weighted residual formulation

The *weighted residual* or weighted-integral formulation corresponding to the strong form(ulation) of the PDE over an open domain $\Omega \subset \mathbb{R}^d$ is: find $u \in \mathcal{U} \subset \mathcal{C}^m(\Omega)$ such that

$$\int_\Omega w R[u] \, dV = 0 \tag{3.13}$$

for all $w \in \mathcal{C}_c^\infty(\Omega)$. This weighted residual formulation of the PDE is *equivalent* to the strong formulation, i.e., if u is a solution of the strong form, it is a solution of the weighted residual form and vice versa. This equivalence can be proven by invoking the fundamental lemma of variational calculus (Lemma 3.1) applied to the residual function $R[u]$.

Lemma 3.1 (Fundamental Lemma of Variational Calculus)**.** *Consider an open domain $\Omega \subset \mathbb{R}^d$ and suppose $G \in \mathcal{C}^0(\Omega)$. Then, $G = 0$ on Ω is equivalent to the weighted residual statement:*

$$\int_\Omega G\eta \, dV = 0 \tag{3.14}$$

for all $\eta \in \mathcal{C}_c^\infty(\Omega)$.

Proof. $G = 0$ immediately implies the weighted residual statement. To show the converse is true, suppose the weighted residual statement holds. It can be shown that there exists a sequences of functions $G_n \in \mathcal{C}_c^\infty(\Omega)$ that converges to $G \in L^2(\Omega) \subset \mathcal{C}^0(\Omega)$ (because $\mathcal{C}_c^\infty(\Omega)$ is dense in $L^2(\Omega)$, where $L^2(\Omega)$ is the set of square integrable functions over Ω). From this, we have

$$\lim_{n\to\infty} \int_\Omega G(x) G_n(x) \, dx = \int_\Omega |G(x)|^2 \, dx.$$

From our assumption that the weighted residual statement holds we have

$$\int_\Omega G(x) G_n(x) \, dx = 0,$$

for all $n \in \mathbb{N}$ because $G_n \in \mathcal{C}_c^\infty(\Omega)$. Together these equations imply that $G = 0$. □

The η functions in Lemma 3.1 are called *test functions* in the context of variational formulation of PDEs. It is important to note that the weighted residual formulation is equivalent to enforcing the PDE over its domain Ω, but does not incorporate any of the boundary conditions of the problem; the boundary conditions are enforced *strongly* through the trial space. This implies that any numerical method based on the weighted residual formulation must explicitly enforce all boundary conditions, which can be difficult for complex domains Ω. Also note that this was a choice; a weighted residual statement incorporating the boundary conditions could have been formed by introducing separate test functions over the boundary $\partial\Omega$. In Section 3.5, we introduce the weak formulation of the problem, which relaxes the regularity requirements on the solution u (currently require m times continuously differentiable) and incorporates the Neumann or natural boundary conditions weakly into the integral equation rather than strongly in the trial space.

Example 3.3: Weighted residual formulation of second-order PDE in one dimension

Recall the second-order PDE in one dimension (3.4)-(3.5). The weighted residual formulation is: find $u \in \mathcal{U}$ (3.6) such that

$$\int_0^L w \left(-\frac{d}{dx}\left[a \frac{du}{dx} \right] - f \right) dx = 0 \tag{3.15}$$

for all $w \in \mathcal{C}_c^\infty((0, L))$.

3.4. Method of weighted residuals

With the variational (integral) formulation of partial differential equations introduced in the previous section, we turn to constructing numerical methods based on the weighted residual formulation, called the method of weighted residuals. Recall the weighted residual formulation (3.13) of the general PDE in (3.1): find $u \in \mathcal{U}$ such that

$$\int_\Omega w R[u]\, dV = 0$$

for all $w \in \mathcal{C}_c^\infty(\Omega)$. For simplicity in constructing test function spaces (later), we enforce the weighted residual equation over the larger function space $\mathcal{C}^\infty(\Omega) \supset \mathcal{C}_c^\infty(\Omega)$ (by considering a *larger* space we maintain equivalence to the strong formulation). It is obvious there is no hope of enforcing this condition for all functions in $\mathcal{C}^\infty(\Omega)$ (an infinite-dimensional function space) in a numerical method (intended to be implemented on a computer or computed by hand). Instead, we will settle for enforcing the weighted residual equation on a finite-dimensional subspace $\mathcal{W}_h \subset \mathcal{C}^\infty(\Omega)$. By replacing the infinite-dimensional $\mathcal{C}^\infty(\Omega)$ with the finite-dimensional $\mathcal{W}_h$, the weighted residual formulation is no longer equivalent to the strong formulation, rather it is an approximation.

The subspace $\mathcal{W}_h$ is constructed as the span of a linearly independent set of functions $\{w_1, \dots, w_n\} \subset \mathcal{C}^\infty(\Omega)$, i.e.,

$$\mathcal{W}_h \coloneqq \text{span}\{w_1, \dots, w_n\}.$$

By definition, $\{w_1, \dots, w_n\}$ is a basis of $\mathcal{W}_h$ and $\dim \mathcal{W}_h = n$. By virtue of the weight function appearing linearly in the weighted residual equation, the equation holds for all $w \in \mathcal{W}_h$ if and only if it holds for each w_i, $i = 1, \dots, n$ (Proposition 3.1). Therefore, the finite-dimensional test space approximation of the weighted residual formulation reduces to: find $u \in \mathcal{U}$ such that

$$\int_\Omega w_i R[u]\, dV = 0$$

for $i = 1, \dots, n$.

Proposition 3.1. *Let $\mathcal{Z}$ be any finite-dimensional space ($\dim \mathcal{Z} = n$) of integrable, real-valued functions with domain $\Omega \subset \mathbb{R}^d$. For any function $f : \Omega \to \mathbb{R}$, the following are equivalent*

(i) $\int_\Omega z f\, dV = 0$ *for all* $z \in \mathcal{Z}$

(ii) $\int_\Omega z_i f\, dV = 0$ *for* $i = 1, \dots, n$, *where* $\mathcal{B} = \{z_1, \dots, z_n\}$ *is a basis for* $\mathcal{Z}$.

Proof. Suppose (ii) holds and expand any $z \in \mathcal{Z}$ in the basis $\mathcal{B}$: $z = \alpha_i z_i$. Then we have

$$\int_\Omega z f\, dV = \int_\Omega \alpha_i z_i f\, dV = \alpha_i \int_\Omega z_i f\, dV = 0,$$

which establishes (i). Now suppose (i) holds. Since $\mathcal{B} \subset \mathcal{Z}$, (i) follows trivially by taking $z = z_i \in \mathcal{Z}$ for $i = 1, \dots, n$. □

While this finite-dimensional approximation of the test space has simplified the problem to only needing to test the residual against a finite number of weighting functions, we still must search an infinite-dimensional function space $\mathcal{U}$ (all functions in $\mathcal{C}^m(\Omega)$ that satisfy the BCs) for the solution. To simplify this task to one that can be performed on a computer (or with hand calculations), we restrict the trial space to a finite-dimensional subset $\mathcal{U}_h \subset \mathcal{U}$, which further approximates the original weighted residual formulation. The construction of the finite-dimensional trial space is more delicate than the test space because it is an *affine* space (for linear PDEs) rather than a linear one, i.e., $\mathcal{U} = \varphi + \mathcal{U}^0$ for any $\varphi \in \mathcal{U}$ and $\mathcal{U}^0$ is a linear space (Section 3.2).

This additive decomposition of the trial space suggests a convenient means to construct general elements of $\mathcal{U}_h$: define a function $\varphi \in \mathcal{U}$, i.e., $\varphi \in \mathcal{C}^m(\Omega)$ that satisfies all BCs (called a *particular solution*), and

introduce a finite-dimensional subspace of the linear space $\mathcal{U}^0$ ($\mathcal{C}^m(\Omega)$ functions satisfying the homogeneous BCs), denoted $\mathcal{U}_h^0$ ($\dim \mathcal{U}_h^0 = n$), and define the approximation space

$$\mathcal{U}_h := \varphi + \mathcal{U}_h^0. \tag{3.16}$$

We define $\mathcal{U}_h^0$ as the span of a linearly independent set of functions $\{\phi_1, \dots, \phi_n\}$

$$\mathcal{U}_h^0 := \text{span}\{\phi_1, \dots, \phi_n\}. \tag{3.17}$$

Then elements $u_h \in \mathcal{U}_h$ take the form

$$u_h = \varphi + \alpha_i \phi_i, \tag{3.18}$$

where $\alpha \in \mathbb{R}^n$. Finally, to define a meaningful trial space of a PDE of order m, the basis functions $\{\phi_1, \dots, \phi_n\}$ should possess m non-zero derivative functions, which we justify at the end of this section.

Example 3.4: Construction of finite-dimensional trial space

Let us formally construct the finite-dimensional trial space used in Example 3.2. The infinite-dimensional trial space is defined in (3.11). We choose the particular solution to be $\varphi : \mathbb{R} \to \mathbb{R}$ to be

$$\varphi := 1, \tag{3.19}$$

which is easy to verify satisfies the BCs: $\varphi(0) = 1$, $[e^x \varphi'(x)]_{x=\pi/2} = 0$. We define the homogeneous portion of the trial space to be the two-dimensional linear space: $\mathcal{U}_h^0 := \text{span}\{\phi_1, \phi_2\}$ where

$$\phi_1(x) := x^2 - \pi x, \quad \phi_2(x) := x^3 - \frac{3\pi^2}{4} x. \tag{3.20}$$

Because $\phi_1(0) = \phi_2(0) = 0$, they satisfy the homogeneous BCs at $x = 0$. Furthermore, we have $[e^x \phi_1'(x)]_{x=\pi/2} = [e^x(2x - \pi)]_{x=\pi/2} = 0$ and $[e^x \phi_2'(x)]_{x=\pi/2} = \left[e^x(3x^2 - 3\pi^2/4)\right]_{x=\pi/2} = 0$, which confirms they satisfy the homogeneous BCs at $x = \pi/2$. Furthermore the functions that comprise the trial space $\mathcal{U}_h = \varphi + \mathcal{U}_h^0$ (φ and $\mathcal{U}_h^0$ defined above) are smooth and have at least $m = 2$ non-zero derivative. Therefore $\mathcal{U}_h$ is a valid, two-dimensional affine trial space.

Finally, we substitute the affine trial space approximation (3.18) into the weighted residual formulation to obtain its finite-dimensional approximation: find $\alpha \in \mathbb{R}^n$, where $\alpha = (\alpha_1, \dots, \alpha_n)$ such that

$$\int_\Omega w_i R[\varphi + \alpha_j \phi_j] \, dV = 0 \tag{3.21}$$

for $i = 1, \dots, n$. In the special case where R is a linear operator, the above equation reduces to

$$\alpha_j \int_\Omega w_i R[\phi_j] \, dV = -\int_\Omega w_i R[\varphi], \tag{3.22}$$

which can be written as the linear system of equations $K\alpha = F$, where $K \in M_{n,n}(\mathbb{R})$ and $F \in \mathbb{R}^n$ are defined as

$$K_{ij} = \int_\Omega w_i R[\phi_j] \, dV, \qquad F_i = -\int_\Omega w_i R[\varphi] \, dV. \tag{3.23}$$

This linear system makes the requirement that ϕ_i for $i = 1, \dots, n$ possess m non-zero derivatives. Otherwise ϕ_j could be in the null space of R, i.e., $R[\phi_j] = 0$, and the matrix K would have a row of all zeros (singular). Intuitively this means the basis vector ϕ_j would not be contributing to the approximation so any value of the corresponding coefficient α_j would yield the same approximation. If the trial space basis functions possess m non-zero derivatives, the matrix K is nonsingular (invertible) due to the linear independence of the basis vectors $\{w_1, \dots, w_n\}$ and $\{\phi_1, \dots, \phi_n\}$.

Example 3.5: Petrov-Galerkin weighted residual method

To close this section we return to Example 3.2 and apply the weighted residual method to approximate (3.8) where directly using the strong formulation failed. We use the trial space constructed in Example 3.4 (validity in terms of BC enforcement, smoothness, and non-zero derivative functions established). For simplicity we choose the finite-dimensional test space to be $\mathcal{W}_h = \{w_1, w_2\} \subset \mathcal{C}^\infty((0,1))$, where

$$w_1(x) := 1, \qquad w_2(x) := x. \tag{3.24}$$

Since the PDE is linear, we form the matrix in (3.23)

$$K_{ij} = \int_0^{\pi/2} w_i R[\phi_j]\,dx \implies \begin{cases} K_{11} = \displaystyle\int_0^{\pi/2} w_1 R[\phi_1]\,dx = \int_0^{\pi/2} \left[-2xe^x - \sin(x)\right]\,dx = -8.49 \\ K_{12} = \displaystyle\int_0^{\pi/2} w_1 R[\phi_2]\,dx = \int_0^{\pi/2} \left[3(-x^2-2x+1)e^x - \sin(x)\right]\,dx = -25.18 \\ K_{21} = \displaystyle\int_0^{\pi/2} w_2 R[\phi_1]\,dx = \int_0^{\pi/2} \left[-2x^2e^x - x\sin(x)\right]\,dx = -9.76 \\ K_{22} = \displaystyle\int_0^{\pi/2} w_2 R[\phi_2]\,dx = \int_0^{\pi/2} \left[-3x(x^2+2x-1)e^x - x\sin(x)\right]\,dx = -32.56 \end{cases}$$

and the right-hand side vector

$$F_i = -\int_0^{\pi/2} w_i R[\varphi]\,dx \implies \begin{cases} F_1 = -\displaystyle\int_0^{\pi/2} w_1 R[\varphi]\,dx = \int_0^L \sin(x)\,dx = 1 \\ F_2 = -\displaystyle\int_0^{\pi/2} w_2 R[\varphi]\,dx = \int_0^L x\sin(x)\,dx = 1. \end{cases}$$

We solve the resulting linear system of equations to find $\alpha_1 = -0.24$ and $\alpha_2 = 0.041$, which results in the following approximation to the solution of the PDE (Figure 3.1):

$$u_h(x) = 1 - 0.24(x^2 - \pi x) + 0.041(x^3 - 3\pi^2 x/4). \tag{3.25}$$

Unlike the strong form, the weighted residual form resulted in a solvable linear system of equations and a valid approximation of the PDE; however, the solution provides a poor approximation to the exact solution (Figure 3.1). In Example 3.6 we will use the Ritz method to obtain an accurate approximation using a trial space of the same dimension (2).

Thus far, we have introduced and defined the test basis $\{w_1, \ldots, w_n\}$ and trial basis $\{\phi_1, \ldots, \phi_n\}$ independently, which is commonly referred to as a *Petrov-Galerkin* method. In the remainder of this section, we introduce three common choices to define the test function basis in terms of the trial basis functions.

3.4.1 Bubnov-Galerkin method

The first approach, called the Bubnov-Galerkin or Galerkin method, takes the test space $\mathcal{W}_h$ to be the same as the homogeneous trial space $\mathcal{U}_h^0$, which is usually accomplished by using the same basis, i.e., $w_i = \phi_i$ for $i = 1, \ldots, n$. The general form of the finite-dimensional weighted residual equation (3.21) reduces to

$$\int_\Omega \phi_i R[\varphi + \alpha_j \phi_j]\,dV = 0. \tag{3.26}$$

In the linear case, the system matrix and right-hand side (3.23) reduce to

$$K_{ij} = \int_\Omega \phi_i R[\phi_j]\,dV, \qquad F_i = -\int_\Omega \phi_i R[\varphi]\,dV. \tag{3.27}$$

Example 3.6: Galerkin weighted residual method

We return to Example 3.2 and apply the weighted residual method to approximate (3.8) using a Galerkin choice for test space, using the trial space constructed in Example 3.4. Since the PDE is linear, we form the matrix in (3.27)

$$K_{ij} = \int_0^{\pi/2} \phi_i R[\phi_j]\, dx \implies \begin{cases} K_{11} = \int_0^{\pi/2} \phi_1 R[\phi_1]\, dx \approx 7.35 \\ K_{12} = \int_0^{\pi/2} \phi_1 R[\phi_2]\, dx \approx 22.32 \\ K_{21} = \int_0^{\pi/2} \phi_2 R[\phi_1]\, dx \approx 13.38 \\ K_{22} = \int_0^{\pi/2} \phi_2 R[\phi_2]\, dx \approx 39.99 \end{cases}$$

and the right-hand side vector

$$F_i = -\int_0^{\pi/2} \phi_i R[\varphi]\, dx \implies \begin{cases} F_1 = -\int_0^{\pi/2} \phi_1 R[\varphi]\, dx \approx -0.86 \\ F_2 = -\int_0^{\pi/2} \phi_2 R[\varphi]\, dx \approx -1.60. \end{cases}$$

We solve the resulting linear system of equations to find $\alpha_1 = -0.28$ and $\alpha_2 = 0.052$, which results in the following approximation to the solution of the PDE (Figure 3.1):

$$u_h(x) = 1 - 0.28(x^2 - \pi x) + 0.052(x^3 - 3\pi^2 x/4). \tag{3.28}$$

The solution is similar to the Petrov-Galerkin approximation in that it provides a poor approximation to the PDE solution; however, unlike directly using the strong formulation, we obtain a consistent approximation. In Example 3.6 we will use the Ritz method to obtain an accurate approximation using a trial space of the same dimension (2).

3.4.2 Collocation method

Another common approach, called the collocation method, takes the test basis functions

$$w_i(x) = \delta(x - x_i), \tag{3.29}$$

for $i = 1, \ldots, m$, where $x_i \in \Omega$ are selected collocation points throughout the domain and δ is the Dirac delta function. The Dirac delta function $\delta : \Omega \to \mathbb{R}$ is defined such that for any function $f \in \mathcal{C}^1(\Omega)$ and point $\xi \in \Omega$:

$$\int_\Omega f(x)\delta(x - \xi)\, dV = f(\xi). \tag{3.30}$$

Substituting the test functions (3.29) into the finite-dimensional weight-integral formulation (3.21), we see that the collocation method is equivalent to requiring the residual function be zero at the collocation nodes (instead of in a weighted integral sense)

$$R[\varphi + \phi_j c_j](x_i) = 0. \tag{3.31}$$

In the linear case, the system matrix and right-hand side (3.23) reduce to

$$K_{ij} = R[\phi_j](x_i), \qquad F_i = -R[\varphi](x_i). \tag{3.32}$$

Example 3.7: Collocation method

Let us revisit Example 3.2 in the context of the collocation method using the same trial space constructed in Example 3.4. The collocation method simply enforces the residual at specified points throughout the domain. For simplicity we choose these points to be equally spaced away from the boundary: $x_1 = \pi/6$ and $x_2 = \pi/3$. This lead to the linear system (3.32) with matrix

$$K_{ij} = R[\phi_j](x_i) \implies \begin{cases} K_{11} = R[\phi_1](x_1) \approx -2.27 \\ K_{12} = R[\phi_2](x_1) \approx -2.13 \\ K_{21} = R[\phi_1](x_2) \approx -6.83 \\ K_{22} = R[\phi_2](x_2) \approx -19.60 \end{cases} \tag{3.33}$$

and right-hand side vector

$$F_i = -R[\varphi](x_i) \implies \begin{cases} F_1 = -R[\varphi](x_1) \approx 0.50 \\ F_2 = -R[\varphi](x_2) \approx 0.87. \end{cases} \tag{3.34}$$

We solve the linear system to find $\alpha_1 = -0.27$ and $\alpha_2 = 0.049$, which lead to the following approximation to the solution of the PDE

$$u_h(x) = 1 - 0.27(x^2 - \pi x) + 0.049(x^3 - 3\pi^2 x/4). \tag{3.35}$$

The solution is similar to the Petrov-Galerkin and Galerkin approximations in that it provides a poor approximation to the PDE solution; however, unlike directly using the strong formulation, we obtain a consistent approximation. In Example 3.6 we will use the Ritz method to obtain an accurate approximation using a trial space of the same dimension (2).

3.4.3 Least-squares method

Finally, the least-squares method defines the solution coefficients $\alpha \in \mathbb{R}^m$ to be the solution of the minimization problem

$$\underset{\alpha \in \mathbb{R}^m}{\text{minimize}} \quad f(\alpha) := \int_\Omega R[\varphi + \alpha_j \phi_j]^2 \, dV. \tag{3.36}$$

The first-order optimality condition states that f is stationary with respect to α_i, i.e., $\dfrac{\partial f}{\partial \alpha_i} = 0$, which leads to

$$\int_\Omega \phi_i R'[\varphi + \alpha_j \phi_j] R[\varphi + \alpha_j \phi_j] \, dV = 0 \tag{3.37}$$

for $i = 1, \dots, m$. This is fits the form of a weighted residual method with $w_i = \phi_i R'[\varphi + \alpha_j \phi_j]$ for $i = 1, \dots, n$.

3.5. Weak formulation

The construction of the weak formulation of a partial differential equation begins with the weighted residual formulation of the PDE (3.13) and, assuming the PDE is of order $m = 2r$ for $r \in \mathbb{N}$ (even), moves r derivatives from the PDE solution variable u onto the test function w using integration-by-parts. The final step in the derivation of the weak form is to incorporate the natural boundary conditions from the problem statement into boundary terms that arise. This approaches *improves upon* the weighted residual formulation in two keys ways. First, it *weakens* the regularity requirements on the trial space since solutions only need to be continuously differentiable r times to define the weak formulation. Furthermore, the trial space is no longer constrained to solutions that satisfy natural BCs because they are incorporated into the integral equation and imposed weakly.

3.5.1 Model problem

For concreteness, consider the canonical second-order PDE over the unit interval $\Omega := (0, L) \subset \mathbb{R}$: find $u \in \mathcal{U} \subset \mathcal{C}^2(\Omega)$ such that

$$\begin{aligned} -\frac{d}{dx}\left(a\frac{du}{dx}\right) + cu &= f \qquad \text{in } \Omega \\ u(0) &= u_0 \\ \left(a\frac{du}{dx}\right)_{x=L} &= Q_L, \end{aligned} \tag{3.38}$$

where $a, c, f \in \mathcal{F}_{\Omega\to\mathbb{R}}$ are given functions of sufficient smoothness, $u_0, Q_L \in \mathbb{R}$ are given constants, and the trial space is given in (3.6). The weighted residual form of the PDE reads: find $u \in \mathcal{U}$ such that

$$\int_0^L w\left[-\frac{d}{dx}\left(a\frac{du}{dx}\right) + cu - f\right] dx, \tag{3.39}$$

for all $w \in \mathcal{C}_c^\infty(\Omega)$. Apply integration-by-parts (2.47) to move one derivative from u to w

$$\int_0^L \left[\frac{dw}{dx}a\frac{du}{dx} + w(cu - f)\right] dx - \left[wa\frac{du}{dx}\right]_0^L = 0. \tag{3.40}$$

This is the *weak formulation* of (3.38) *without boundary conditions*. To incorporate natural boundary conditions into the weak formulation, we need a concrete definition of essential and natural boundary conditions, which stems from the definition of primary and secondary variables.

3.5.2 Essential and natural boundary conditions

Essential, or Dirichlet, boundary conditions are conditions on *primary* variables along boundaries, while *natural*, or Neumann, boundary conditions are conditions on *secondary* variables along boundaries. Primary and secondary variables can be identified from the weak formulation without boundary conditions, e.g., (3.40). Secondary variables multiply the test functions (or their derivatives) in the boundary terms, whereas primary variables are identified by replacing the test function (w in this case) with the PDE solution variable (u in this case). A number of examples of primary vs. secondary variables and essential vs. natural boundary conditions are provided in the next section.

From these definitions, it is clear that PDEs of order $2r$ will have r primary and secondary variables since there will be r boundary terms resulting from r applications of integration-by-parts to move derivatives from the solution variable to the test variable. There will also be a secondary variable *corresponding* to each primary variable. The secondary variable corresponding to a primary variable can be identified by replacing the primary variable with its corresponding test function in the boundary terms of the weak form and identifying the variable that multiplies it. With these definitions, it follows that u is the only primary variable for the PDE in (3.38) and $a\frac{du}{dx}$ is the corresponding secondary variable. This implies the boundary conditions in (3.38) are classified as

$$u(0) = u_0 \quad \text{(essential BC)}, \qquad \left(a\frac{du}{dx}\right)_{x=L} = Q_L \quad \text{(natural BC)}.$$

3.5.3 Complete weak formulation

To complete the weak formulation we incorporate the natural BCs into (3.40) by substituting them into the boundary term of the weighted residual formulation (integrated-by-parts) to yield

$$\int_0^L \left[\frac{dw}{dx}a\frac{du}{dx} + w\left(cu - f\right)\right] dx + \left[wa\frac{du}{dx}\right]_{x=0} - w(L)Q_L = 0, \tag{3.41}$$

which must hold for all $\mathcal{C}_c^\infty(\Omega)$. However, since functions in $\mathcal{C}_c^\infty(\Omega)$ vanish on $\partial\Omega$, both boundary terms are zero, which eliminates our ability to enforce the natural BCs. For this reason, we *extend* our space of test

functions by only requiring the test functions vanish on $\partial\Omega_D$, i.e., the portion of the boundary with essential BCs, to yield the test space $\mathcal{W} \coloneqq \{w \in \mathcal{C}^\infty(\Omega) \mid w(0) = 0\}$. It can be verified that $\mathcal{W}$ is a linear space for any PDE. Notice that this does not destroy equivalence to the strong formulation because we are still enforcing the integral equation for all $w \in \mathcal{C}_c^\infty(\Omega)$; we have just added functions to *also* enforce the natural BCs. The complete weak formulation is: find $u \in \tilde{\mathcal{V}} \subset \mathcal{C}^2(\Omega)$ such that

$$\int_0^L \left[\frac{dw}{dx} a \frac{du}{dx} + w(cu - f) \right] dx - w(L) Q_L = 0 \tag{3.42}$$

for all $w \in \mathcal{W}$. Since the integral equation incorporates the natural BCs, the trial space does not have to be restricted to functions that satisfy the natural BCs (only essential BCs). Therefore the trial space for the weak form $\tilde{\mathcal{V}} \coloneqq \{u \in \mathcal{C}^2(\Omega) \mid u(0) = u_0\}$ extends the trial space from the strong form (3.6).

We have assumed sufficient regularity of the solution and test function to *pose* the strong formulation of the PDE and *derive* the weak formulation. Under these regularity conditions, the strong, weighted residual, and weak formulations are equivalent. However, we can *weaken* the regularity requirements by directly considering the weak form of a PDE since lower-order derivatives of the PDE variable appear in the weak formulation than in the strong and weighted residual formulations. In the context of our model problem, this means we can search for solutions in $\mathcal{C}^1(\Omega)$ instead of requiring $\mathcal{C}^2(\Omega)$: find $u \in \mathcal{V}$ such that

$$\int_0^L \left[\frac{dw}{dx} a \frac{du}{dx} + w(cu - f) \right] dx - w(L) Q_L = 0 \tag{3.43}$$

for all $w \in \mathcal{W}$, where the trial space is $\mathcal{V} \coloneqq \{u \in \mathcal{C}^1(\Omega) \mid u(0) = u_0\}$. This is a more general form of (3.8) that allows for solutions with weaker regularity. The term *weak* comes from the weakening of the regularity requirements of the solution (trial) space. Under these weaker regularity requirements, the strong and weak form are not equivalent (since the strong form may not even be defined if, e.g., $u \notin \mathcal{C}^2(\Omega)$). We will extend these concepts in later chapters by introducing a notion of regularity based on *integrability* rather than smoothness (differentiability). While the weakened regularity requirements may seem pedantic, it plays a significant role in the construction of finite element spaces.

For the sake of generality in the remainder of this document, we use *functionals* (mapping from functions to scalars) to represent the weak formulation of a general PDE (3.1) of order $m = 2r$, $r \in \mathbb{N}$ over a domain (open) $\Omega \subset \mathbb{R}^d$ with essential BCs prescribed on $\partial\Omega_D$ and natural BCs on $\partial\Omega_N$ such that $\partial\Omega = \overline{\partial\Omega_D \cup \partial\Omega_N}$: find $u \in \mathcal{V} \subset \mathcal{C}^r(\Omega)$ such that

$$B(w, u) = \ell(w) \tag{3.44}$$

for all $w \in \mathcal{W}$, where $B : \mathcal{W} \times \mathcal{V} \to \mathbb{R}$ and $\ell : \mathcal{W} \to \mathbb{R}$ are functionals defining the (weak) integral equations and the trial and test spaces are

$$\begin{aligned} \mathcal{V} &\coloneqq \{u \in \mathcal{C}^r(\Omega) \mid u \text{ satisfies essential BCs on } \partial\Omega_D\} \\ \mathcal{W} &\coloneqq \{w \in \mathcal{C}^\infty(\Omega) \mid w(x) = 0 \text{ for } x \in \partial\Omega_D\}. \end{aligned} \tag{3.45}$$

Since $\mathcal{C}^\infty(\Omega)$ is a linear space and $\mathcal{W}$ is closed under addition and scalar multiplication, $\mathcal{W}$ is a linear space. On the other hand, $\mathcal{V}$ is an *affine* space, i.e., $\mathcal{V} = \varphi + \mathcal{V}^0$ where $\varphi \in \mathcal{V}$ is arbitrary and $\mathcal{V}^0$ is a linear space. Unlike the weighted residual formulation, the trial space of the weak form $\mathcal{V}$ is an affine space for any PDE due to the definition of an essential BC (the trial space of the weighted residual method is only an affine space for special PDEs, e.g., linear PDEs). The linear space $\mathcal{V}^0$ is the collection $\mathcal{C}^r(\Omega)$ functions that satisfy the homogeneous essential boundary conditions

$$\mathcal{V}^0 = \{u \in \mathcal{C}^r(\Omega) \mid u \text{ satisfies homogeneous essential BCs on } \partial\Omega_D\}. \tag{3.46}$$

Equation (3.44) is called a *bilinear form* for linear PDEs. For our model problem ($r = 1$), the functionals in (3.44) are

$$B(w, u) \coloneqq \int_0^L \left[\frac{dw}{dx} a \frac{du}{dx} + wcu \right] dx, \qquad \ell(w) \coloneqq \int_0^L wf \, dx + w(L) Q_L. \tag{3.47}$$

We close this section with the derivation of the weak formulation for some more complicated PDEs. For each problem we will apply the three central steps to derive the weak formulation:

1) derive weighted residual formulation from strong formulation,

2) integrate-by-parts (r times for PDE of order $2r$) to move r derivatives onto the test function, and

3) simplify boundary terms by enforcing natural BCs and that all test functions vanish at essential BCs.

Example 3.8: Fourth-order PDE in one dimension

Let $\Omega := (0, L) \subset \mathbb{R}$ and consider the fourth-order partial differential equation: find $w \in \mathcal{C}^4(\Omega)$ such that

$$\frac{d^4 w}{dx^4} = 0 \tag{3.48}$$

with boundary conditions $w(0) = w_0$, $w(L) = w_L$, $\frac{dw}{dx}(0) = w_0'$, and $\frac{d^2w}{dx^2}(L) = w_L''$. This boundary value problem describes the deflection of a (Euler-Bernoulli) beam subject to prescribed deflection at the left (w_0) and right end (w_L), prescribed rotation at the left end (w_0'), and applied moment at the right end (w_L'').

To construct the weak formulation, multiply the governing equations by a test function $v \in \mathcal{C}^2(\Omega)$, integrate over the domain, and apply integration-by-parts twice to move two derivatives onto the test function

$$0 = \int_0^L v \frac{d^4 w}{dx^4}\, dx = \int_0^L -\frac{dv}{dx}\frac{d^3 w}{dx^3}\, dx + \left[v \frac{d^3 w}{dx^3}\right]_0^L = \int_0^L \frac{d^2 v}{dx^2}\frac{d^2 w}{dx^2}\, dx + \left[v \frac{d^3 w}{dx^3}\right]_0^L - \left[\frac{dv}{dx}\frac{d^2 w}{dx^2}\right]_0^L.$$

From examining the boundary terms, we identify the primary variables as w and $\frac{dw}{dx}$ (identify test function v in boundary terms and replace with PDE function w) and the corresponding secondary variables as $\frac{d^3w}{dx^3}$ and $\frac{d^2w}{dx^2}$ (terms multiplying the test functions in the boundary terms), respectively. Since both primary variables are specified at $x = 0$, we set the corresponding test functions to zero $v(0) = \frac{dv}{dx}(0) = 0$. In addition, the primary variable w is specified as $x = L$ so we set the corresponding test function to zero $v(L) = 0$. Using these choices for the test function and incorporating the natural boundary condition $\frac{d^2w}{dx^2}(L) = w_L''$ into the boundary terms above, we arrive at the weak form

$$\int_0^L \frac{d^2 v}{dx^2}\frac{d^2 w}{dx^2}\, dx - \frac{dv}{dx}(L) w_L'' = 0,$$

which can be formulated for solutions in $\mathcal{C}^2(\Omega)$.

Example 3.9: Timoshenko beam

Let $\Omega := (0, L) \subset \mathbb{R}$ and consider the system of second-order partial differential equations that govern the deflection of a beam using Timoshenko theory: find $w \in \mathcal{C}^2(\Omega)$ and $\phi_x \in \mathcal{C}^2(\Omega)$ such that

$$\begin{aligned} -\frac{d}{dx}\left[S\left(\frac{dw}{dx} + \phi_x\right)\right] + c_f w &= q \\ -\frac{d}{dx}\left(D\frac{d\phi_x}{dx}\right) + S\left(\frac{dw}{dx} + \phi_x\right) &= 0 \end{aligned} \tag{3.49}$$

holds in Ω with boundary conditions $w(0) = \phi_x(0) = 0$, $\left[S\left(\frac{dw}{dx} + \phi_x\right)\right]_{x=L} = F_L$, and $\left[D\frac{d\phi_x}{dx}\right]_{x=L} = M_0$. The remaining terms are known (sufficiently smooth) functions $S, D, c_f, q \in \mathcal{F}_{\Omega\to\mathbb{R}}$ and scalars $M_0, F_L \in \mathbb{R}$.

Since the governing equation is a *system* of PDEs, we introduce a test function for each equation: $v_1 \in \mathcal{C}^1(\Omega)$ for the first PDE (for w) and $v_2 \in \mathcal{C}^1(\Omega)$ for the second PDE (for ϕ_x). To construct the weak formulation, we begin by constructing the weighted residual formulation: multiply each PDE by its own test function, integrate each over the domain, and add them:

$$\int_0^L \left(v_1\left\{-\frac{d}{dx}\left[S\left(\frac{dw}{dx} + \phi_x\right)\right] + c_f w - q\right\} + v_2\left\{-\frac{d}{dx}\left(D\frac{d\phi_x}{dx}\right) + S\left(\frac{dw}{dx} + \phi_x\right)\right\}\right) dx = 0.$$

Apply integration-by-parts to move a derivative from w onto v_1 and from ϕ_x onto v_2

$$\int_0^L \left\{ \frac{dv_1}{dx} S\left(\frac{dw}{dx} + \phi_x\right) + \frac{dv_2}{dx} D \frac{d\phi_x}{dx} + v_1(c_f w - q) + v_2 S\left(\frac{dw}{dx} + \phi_x\right) \right\} dx \\ - \left[v_1 S\left(\frac{dw}{dx} + \phi_x\right) \right]_0^L - \left[v_2 D \frac{d\phi_x}{dx} \right]_0^L = 0.$$

From this form, we can identify the primary variables as w (replace v_1 with w in the boundary terms to identify) and ϕ_x (replace v_2 with ϕ_x in the boundary terms to identify) and the corresponding secondary variables as $S\left(\frac{dw}{dx} + \phi_x\right)$ and $D\frac{d\phi_x}{dx}$, respectively. Since both primary variables are specified at $x = 0$, we take $v_1(0) = v_2(0) = 0$ and substitute the natural boundary conditions at $x = L$ to arrive at the final version of the weak formulation

$$\int_0^L \left\{ \frac{dv_1}{dx} S\left(\frac{dw}{dx} + \phi_x\right) + \frac{dv_2}{dx} D \frac{d\phi_x}{dx} + v_1(c_f w - q) + v_2 S\left(\frac{dw}{dx} + \phi_x\right) \right\} dx - v_1(L) F_L - v_2(L) M_L = 0,$$

which can be formulated for solutions in $\mathcal{C}^1(\Omega)$.

Example 3.10: Poisson equation in d dimensions

Let $\Omega \subset \mathbb{R}^d$ (open) and consider the Poisson equation: find $u \in \mathcal{C}^2(\Omega)$ such that

$$\begin{aligned} -\Delta u &= 0 && \text{in } \Omega \\ u &= g && \text{on } \partial\Omega_D \\ \nabla u \cdot n &= h && \text{on } \partial\Omega_N, \end{aligned} \tag{3.50}$$

where the boundary of the domain $\partial\Omega$ is partitioned into $\partial\Omega_D$ (essential/Dirichlet condition applied) and $\partial\Omega_N$ (natural/Neumann condition applied), i.e., $\partial\Omega = \overline{\partial\Omega_1 \cup \partial\Omega_2}$. For convenience, we convert this equation to indicial notation: $-u_{,ii} = 0$ in Ω, $u = g$ on $\partial\Omega_D$, $u_{,i} n_i = h$ on $\partial\Omega_N$.

To derive the weak formulation, we follow the standard procedure and setup the weighted residual equation by multiplying by a test function $w \in \mathcal{C}^1(\Omega)$ and integrating over the domain

$$\int_\Omega w(-u_{,ii})\, dV = 0.$$

Applying integration-by-parts (use the identity $(wu_{,i})_{,i} = w_{,i}u_{,i} + wu_{,ii}$ and apply the divergence theorem; see (2.48)) yields

$$\int_\Omega w(-u_{,ii})\, dV = \int_\Omega \left(w_{,i}u_{,i} - (wu_{,i})_{,i}\right) dV = \int_\Omega w_{,i}u_{,i}\, dV - \int_{\partial\Omega} wu_{,i}n_i\, dS = 0.$$

By examining the boundary terms, we see that u is the primary variable (from replacing the test function with u in the boundary term) and $u_{,i}n_i$ is the secondary variable (multiplies the test function in the boundary term). Next, we choose $w(x) = 0$ for $x \in \partial\Omega_D$ because the primary variable is specified on $\partial\Omega_D$. This causes the integral over the entire boundary to become an integral over only $\partial\Omega_N$ because of the additive property of integration

$$\int_{\partial\Omega} wu_{,i}n_i\, dS = \int_{\partial\Omega_D} wu_{,i}n_i\, dS + \int_{\partial\Omega_N} wu_{,i}n_i\, dS = \int_{\partial\Omega_N} wu_{,i}n_i\, dS,$$

where the last equality used $w = 0$ on $\partial\Omega_D$. Finally, we substitute the natural boundary condition $u_{,i}n_i = h$ on $\partial\Omega_N$ into the weak form to yield

$$\int_\Omega w_{,i}u_{,i}\, dV - \int_{\partial\Omega_N} w\, h\, dS = 0,$$

which can be formulated for solutions in $\mathcal{C}^1(\Omega)$.

3.6. Ritz method

While the method of weighted residuals did not suffer from the same drawbacks as methods based on the strong formulation, they have their own disadvantages. In particular, the approximation functions used must have $2r$ non-zero derivatives for PDEs of order $2r$, which eliminates a number of useful and efficient families of approximations. In addition, the weighted residual form does not incorporate any of the boundary conditions so the solution basis must account for them. To avoid these issues, we construct a numerical method based on the weak formulation of the PDE, which only requires the solution basis have m non-zero derivatives (because half of the derivatives of the PDE were moved onto the test functions) and only need to satisfy the essential boundary conditions (the natural boundary conditions are embedded in the weak form).

Recall the weak form of a general PDE of order $m = 2r$ (3.44) with test and trial space defined in (3.45). Following our derivation of the method of weighted residuals, we approximate the infinite-dimensional test and trial spaces using finite-dimensional spaces. Construction of a finite-dimensional trial space closely follows the corresponding procedure for the method of weighted residuals. First, since the trial space for the weak formulation is an affine space, we write it as

$$\mathcal{V} = \varphi + \mathcal{V}^0, \tag{3.51}$$

where $\varphi \in \mathcal{V}$ (particular solution) and $\mathcal{V}^0$ is a linear space of $\mathcal{C}^r(\Omega)$ functions that satisfy the homogeneous essential BCs. Then we define the finite-dimensional approximation to $\mathcal{V}$ as

$$\mathcal{V}_h \coloneqq \varphi + \mathcal{V}_h^0, \tag{3.52}$$

where $\mathcal{V}_h^0 \subset \mathcal{V}_h$ and $\dim \mathcal{V}_h^0 = n$. We define $\mathcal{V}_h^0$ as the span of a linearly independent set of functions $\{\phi_1, \dots, \phi_n\}$. Then elements $u_h \in \mathcal{V}_h$ take the form

$$u_h = \varphi + \alpha_i \phi_i, \tag{3.53}$$

where $\alpha \in \mathbb{R}^n$. Similar to the weighted residual method, the basis functions of the trial space should posses r non-zero derivatives for a PDE of order $m = 2r$ (half as many as required by the weighted residual method) to ensure the resulting system has a unique solution. This plays a significant role in the construction of finite element spaces and enables the use of piecewise linear basis functions, by far the most widely used finite element solution space, for second-order PDEs.

To define the finite-dimensional test space, the Ritz method employs a Galerkin approximation, i.e., $\mathcal{W}_h \coloneqq \mathcal{V}_h^0$. With these approximations, the finite-dimensional (Ritz) approximation of the weak form (3.44) is: find $u_h \in \mathcal{V}_h$ such that

$$B(w_h, u_h) = \ell(w_h) \tag{3.54}$$

for all $w_h \in \mathcal{V}_h^0$. From Proposition 3.1, enforcing (3.54) for all $w_h \in \mathcal{V}_h^0$ is equivalent to enforcing it for all vectors in a basis (since $\mathcal{V}_h^0$ finite-dimensional), which reduces the Ritz formulation to: find $\alpha \in \mathbb{R}^n$ such that

$$B(\phi_i, \varphi + \alpha_j \phi_j) = \ell(\phi_i) \tag{3.55}$$

for $i = 1, \dots, n$. In the special case where B is bilinear and ℓ is linear, (3.55) becomes

$$\alpha_j B(\phi_i, \phi_j) = \ell(\phi_i) - B(\phi_i, \varphi), \tag{3.56}$$

which can be written as the linear system of equations $K\alpha = F$, where $K \in M_{n,n}(\mathbb{R})$ and $F \in \mathbb{R}^n$ are defined as

$$K_{ij} = B(\phi_i, \phi_j), \qquad F_i = \ell(\phi_i) - B(\phi_i, \varphi). \tag{3.57}$$

Once this system (3.55) (or (3.57) in the linear case) has been solved for the coefficients α, they are substituted back into (3.53) to obtain our approximation of the PDE

$$u \approx u_h = \varphi + \alpha_j \phi_j. \tag{3.58}$$

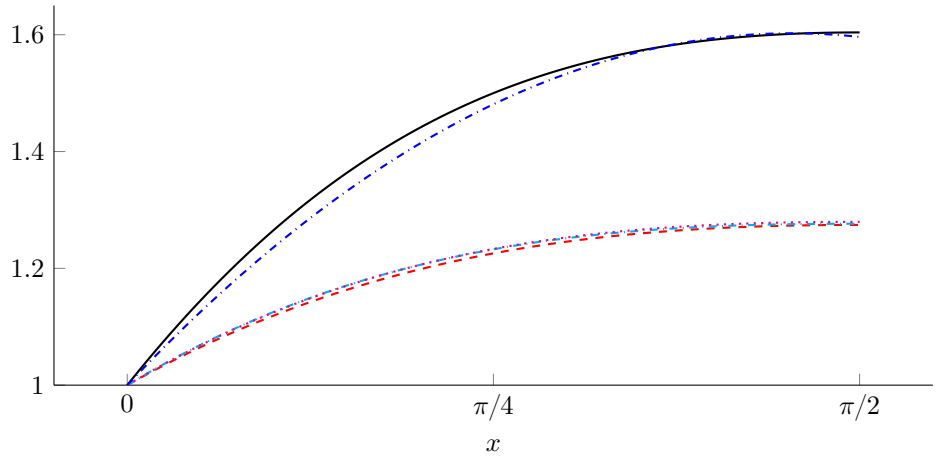

Figure 3.1: The solution to (3.8) (——), its approximation using the Ritz method (-·-·-) in Example 3.6 and the method of weighted residuals: Petrov-Galerkin in Example 3.4 (- - -), Bubnov-Galerkin in Example 3.4.1 (-··-··-), and collocation in Example 3.4.2 (·····).

Example 3.11: Ritz method

To conclude this section, we return to Example 3.2 and apply the Ritz method to approximate the solution of (3.8). The weak formulation of (3.8) is: find $u \in \mathcal{V}$ such that:

$$\int_0^{\pi/2} \left[\frac{dw}{dx} a \frac{du}{dx} - wf \right] dx = 0 \tag{3.59}$$

for all $w \in \mathcal{W}$, where the trial and test space are

$$\mathcal{V} := \left\{ u \in \mathcal{C}^1((0,\pi/2)) \mid u(0) = 1 \right\}, \qquad \mathcal{W} := \{ w \in \mathcal{C}^\infty((0,\pi/2)) \mid w(0) = 0 \} \tag{3.60}$$

and $a(x) := e^x$ and $f(x) := \sin x$. The weak formulation can be writtin in bilinear form (3.44) as

$$B(w,u) = \int_0^{\pi/2} \frac{dw}{dx} a \frac{du}{dx}\, dx, \qquad \ell(w) = \int_0^{\pi/2} wf\, dx \tag{3.61}$$

Given the poor approximation of the trial space in Examples 3.4,3.4.2, we use the flexibility afforded by the weak formulation to construct an improved one

$$\mathcal{V}_h := \varphi + \text{span}\{\phi_1, \dots, \phi_N\}, \tag{3.62}$$

where $\varphi := 1$ and $\phi_k = x^k$ for $k = 1, \dots, N$. This implies the Ritz linear system (3.57) is ($N = 2$)

$$K_{ij} = B(\phi_i, \phi_j) = \int_0^{\pi/2} ijx^{i+j-2} e^x\, dx \implies \begin{cases} K_{11} = e^{-\pi/2} - 1 \approx 3.81 \\ K_{12} = K_{21} = 2 + (\pi - 2)e^{\pi/2} \approx 7.49 \\ K_{22} = -8 + (\pi^2 - 4\pi + 8)e^{\pi/2} \approx 17.51 \end{cases} \tag{3.63}$$

$$F_i = \ell(\phi_i) - B(\phi_i, \varphi) = \int_0^{\pi/2} x^i \sin(x)\, dx \implies \begin{cases} F_1 = 1 \\ F_2 = \pi - 2 \approx 1.14, \end{cases}$$

which can be solved to obtain the coefficients $\alpha_1 = 0.845$ and $\alpha_2 = -0.296$. This leads to the following approximation to the solution of the PDE

$$u_h(x) = 1 + 0.845x - 0.296x^2. \tag{3.64}$$

From Figure 3.1 we see the Ritz solution is a far better approximation to the true solution (3.10) of the PDE in (3.8), which largely comes from the flexibility afforded by the weakened conditions on the trial space. To improve the approximation we simply include more terms in the polynomial expansion, i.e., larger N.

As we will see in the remainder of the course, the finite element method is a Ritz method with a particularly convenient choice/construction of the trial space.

3.7. Summary

This chapter introduced variational formulations (weighted residual and weak) of the partial differential equations and numerical methods based on them (weighted residual and Ritz):

1) The strong or differential formulation of a partial differential equation is not always amenable to approximation in a finite-dimensional trial space.

2) An equivalent formulation recasts the PDE as a weighted integral statement that must hold for arbitrary weighting functions.

3) Integrating the weighted residual statement by parts yields another variational formulation of the PDE, known as the weak formulation. The term *weak* comes from the weaker regularity requirements on the trial space (only need to be r times differentiable for a PDE of order $2r$, whereas weighted residual form requires $2r$ times differentiable).

4) The weak formulation defines a systematic procedure to identify primary vs. secondary variables of a boundary value problem, which leads to a formal definition of essential vs. natural boundary conditions.

5) Since the weighted residual formulation does not incorporate the boundary conditions of the boundary value problem, the trial space must only contain solutions that satisfy all essential and natural boundary conditions, which can be a difficult task.

6) The weak formulation incorporates natural boundary conditions so the trial space is only required to contain solutions that satisfy the essential boundary conditions.

7) The three step procedure for deriving the weak formulation of a PDE from its strong formulation is:

   1) derive weighted residual formulation from strong formulation,

   2) integrate-by-parts (r times for PDE of order $2r$) to move r derivatives onto the test function, and

   3) simplify boundary terms by enforcing natural BCs and that all test functions vanish at essential BCs.

8) The method of weighted residuals is a numerical method for approximating boundary value problems based on their weighted residual formulation. It approximates the infinite-dimensional test and trial spaces with finite-dimensional subspaces (affine space for trial space). Common approaches to choose the test space include:

   - Petrov-Galerkin: independent test/trial spaces
   - (Bubnov-)Galerkin: test space taken to be homogeneous part of trial space
   - Collocation: enforce PDE at selected points throughout the domain
   - Least-squares: solution defined such that L^2 norm of the residual function is minimized

9) The Ritz method is a Galerkin method based on the weak formulation where the finite-dimensional trial space contains functions satisfying the essential boundary conditions and the test space is taken as the homogeneous part of the trial space (which is a linear space).