# 7

# DIRECT VARIATIONAL METHODS

## 7.1 INTRODUCTION

In Chapters 5 and 6 we saw how energy principles can be used to obtain governing equations, associated boundary conditions, and, in certain simple cases, solutions for displacements and forces at selective points of a structure. However, the energy methods considered in Chapters 5 and 6 cannot be used, in general, to determine continuous solutions to complex problems.

The present chapter deals with approximate methods that employ the variational statements (i.e., either variational principles or weak formulations) to determine continuous solutions of problems of mechanics. Recall that the energy principles contain, in a single statement, the governing equation(s) and the natural boundary condition(s) of the problem. The energy principles involved setting the first variation of an appropriate functional with respect to the dependent variables to zero. The procedures of the calculus of variations were then used to obtain the governing (Euler–Lagrange) equations of the problem. In contrast, the methods described in this chapter seek a solution in terms of adjustable parameters that are determined by substituting the assumed solution into the functional and finding its extremum or stationary value with respect to the parameters. Such solution methods are called *direct methods*, because the approximate solutions are obtained directly by using the same variational principle that was used to derive the governing equations.

The assumed solutions in the variational methods are in the form of a finite linear combination of *undetermined parameters* with appropriately chosen functions. This amounts to representing a continuous function by a finite linear combination of functions. Since the solution of a continuum problem in general cannot be represented by a finite set of functions, error is introduced into the solution. Therefore, the solution obtained is an *approximation* of the true solution for the equations describing a





physical problem. As the number of linearly independent terms in the assumed solution is increased, the error in the approximation will be reduced, and the assumed solution converges to the desired solution.

The equations governing a physical problem themselves are approximate. The approximations are introduced via several sources, including the geometry, the representation of specified loads and displacements, and the material behavior. In the present study, our primary interest is to determine accurate approximate solutions to appropriate analytical descriptions of physical problems.

The variational methods of approximation described here include the classical methods of Ritz, Galerkin, and Petrov–Galerkin (weighted residuals). Examples of applications of these methods are drawn from the problems of bars, beams, torsion, and membranes [1–23]. Applications of these methods to circular and rectangular plates are considered in Chapter 8. We begin with some mathematical preliminaries.

## 7.2 CONCEPTS FROM FUNCTIONAL ANALYSIS

### 7.2.1 General Introduction

Before we discuss the variational methods of approximation, it is useful to equip ourselves with certain mathematical concepts. These include the vector spaces, norm, inner product, linear independence, orthogonality, and linear and bilinear forms of functions. Since the objective of the present study is to learn about variational methods, we limit our discussion here only to concepts that are pertinent in the context. The reader already familiar with these concepts may browse through the section to gain familiarity with the notation. Others who do not wish to burden themselves with the formalism of functional analysis may skip this section; it would not prevent them from gaining an understanding of the main ideas of variational methods.

A set $X$ is any well-defined collection of things, which are called *members* or *elements* of $X$. In the present study we are concerned with collections of numbers, sequences, functions, and functions of functions. Examples of sets are provided below:

1. The set $\Re$ of all real numbers.
2. The set $C[0, L]$ of all real-valued continuous functions $f(x)$ defined on the closed interval $0 \leq x \leq L$.
3. The collection of all closed intervals, $I_i = [x_i, x_{i+1}]$, on the real line.

The following notation, very standard in mathematics, is adopted here:

$$\begin{aligned} &\subset \text{ means "a subset of"} \\ &\not\subset \text{ means "not a subset of"} \\ &\in \text{ means "an element of"} \\ &\notin \text{ means "not a element of"} \end{aligned} \tag{7.1}$$

$\forall$ means "for all"

$\exists$ means "there exists"

$\ni$ means "such that".

One way of defining a set $S$ is to specify two pieces of information: (1) assume that each element of $S$ is an element of a universal set (i.e., a well-known set), say $X$, and (2) list the properties that elements of the universal set must satisfy in order to be in $S$. For example, let $X$ be the set of all sequences of complex numbers $x = \{x_1, x_2, x_3, \ldots\}$ and $S$ be all elements of $X$ possessing the property

$$\sum_{n=1}^{\infty} |x_n| < \infty.$$

We shall use the following notation

$$S = \left\{ x \in X \colon \sum_{n=1}^{\infty} |x_n| < \infty \right\}, \tag{7.2}$$

which is read "$S$ is the set of all elements of $X$ such that (the colon stands for 'such that') $\sum_{n=1}^{\infty} |x_n| < \infty$."

A set $A \subset \Re$ is said to be *bounded from above* if there exists a real number $\mu$ such that $a \leq \mu$ for all $a \in A$. The real number $\mu$ is said to be an *upper bound* of the set $A$. Similarly, a set $A$ is said to be *bounded from below* if there exists a real number $\gamma$ such that $a \geq \gamma$ for all $a \in A$. The real number $\gamma$ is said to be a *lower bound* of the set $A$. If a set $A$ is bounded from above and from below, we say that $A$ is bounded. An upper (lower) bound $M(m)$ for $A$ is said to be the maximum (minimum) of $A \subset \Re$ if $M \in A$ ($m \in A$). It should be noted that even a bounded set need not have a maximum or a minimum. Every nonempty set of real numbers bounded from above has a "least upper bound," and every nonempty set of real numbers bounded from below has a "greatest lower bound." The least upper bound of a set $A$ is denoted by sup $A$ ("supremum of $A$"), and the greatest lower bound of $A$ is denoted by inf $A$ ("infimum of $A$").

### 7.2.2 Linear Vector Spaces

As we have seen in Chapter 2, the term *vector* is used often to imply a *physical* vector that has "magnitude and direction" and obeys certain rules of vector addition and scalar multiplication. These ideas can be extended to functions, which are also called vectors, provided that the rules of vector addition and scalar multiplication are defined. While the definition of a vector "from a linear vector space" does not require the vector to have a magnitude, in nearly all cases of practical interest the vector is endowed with a magnitude, called the *norm*. In such cases the vector is said to belong to a normed vector space. We begin with a formal definition of an abstract vector space.



A collection of vectors, $u, v, w, \ldots$ is called a *real linear vector space* $V$ over the real number field $\Re$ if the following rules of vector addition and scalar multiplication of a vector are satisfied by the elements of the vector space.

***Vector Addition*** To every pair of vectors $u$ and $v$ there corresponds a unique vector $u + v \in V$, called the *sum* of $u$ and $v$, with the following properties:

(1a) $u + v = v + u$ (commutative);

(1b) $(u + v) + w = u + (v + w)$ (associative);

(1c) there exists a unique vector, $\Theta$, independent of $u$ such that $u + \Theta = u$ for every $u \in V$ (existence of an identity element);

(1d) to every $u$ there exists a unique vector, $-u$ (that depends on $u$), such that $u + (-u) = \Theta$ for every $u \in V$ (existence of the additive inverse element). (7.3)

***Scalar Multiplication*** To every vector $u$ and every real number $\alpha \in \Re$ there corresponds a unique vector $\alpha u \in V$, called the *product* of $u$ and $\alpha$, such that the following properties hold:

(2a) $\alpha(\beta u) = (\alpha\beta)u$ (associative);

(2b) $(\alpha + \beta)u = \alpha u + \beta u$ (distributive w.r.t. the scalar addition);

(2c) $\alpha(u + v) = \alpha u + \alpha v$ (distributive w.r.t. the vector addition);

(2d) $1 \cdot u = u \cdot 1$. (7.4)

Note that in order to prove that a set of vectors qualifies as a vector space, one must define the identity and inverse elements and prove the "closure property" $u + v \in V$ and $\alpha u \in V$ for all $u, v \in V$ and $\alpha \in \Re$.

A subset $S$ of a vector space $V$ is called a *subspace* of $V$, denoted $S \subset V$, if $S$ itself is a vector space with respect to vector addition and scalar multiplication defined over $V$.

**Example 7.1**

1. The set of ordered $n$-tuples $(x_1, x_2, x_3, \ldots, x_n)$ of real numbers $x_1, x_2, \ldots, x_n$ is called the *Cartesian space*, denoted $\Re^n$. A typical element of $\Re^n$ is denoted $\mathbf{x} = (x_1, x_2, x_3, \ldots, x_n)$. The Cartesian space is a linear vector space with respect to the usual rules of addition and scalar multiplication:

   *Vector addition:* $\mathbf{x} + \mathbf{y} = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n) \quad \forall \mathbf{x}, \mathbf{y} \in \Re^n$.

   *Scalar multiplication:* $\alpha\mathbf{x} = (\alpha x_1, \alpha x_2, \ldots, \alpha x_n) \quad \forall \mathbf{x} \in \Re^n$ and $\alpha \in \Re$.

   The identity element is $\mathbf{0} = (0, 0, 0, \ldots)$ ($n$ zeros) and the inverse element is the negative of the vector.

2. Let $\mathcal{P}$ be the set of *all* polynomials in $x$ with real coefficients. A typical element of $\mathcal{P}$ is of the form

$$p(x) = a_0 + a_1 x + a_2 x^2 + \cdots,$$

where $a_0, a_1, \ldots$ are real numbers. Then $\mathcal{P}$ is a linear vector space with respect to the usual rules of addition and scalar multiplication. Also, the set $\mathcal{P}_n$ of polynomials of degree *less than or equal to* degree $n$ is also a linear vector space, as can be verified (by the closure property). Moreover, $\mathcal{P}_n \subset \mathcal{P}$. However, the set of polynomials of degree *equal to* $n$ is not a vector space as the closure property is violated. For example, consider the set of all cubic polynomials. The sum of $p_1(x) = 1 - 2x + 3x^2 + 6x^3$ and $p_2 = -3 + 5x + 2x^2 - 6x^3$ is not a cubic polynomial.

3. Let $C^n[a, b]$, where $n \geq 0$ is an integer, denote the set of all real-valued functions $u(x)$ defined on the interval $a \leq x \leq b$ such that $u$ is continuous, and the derivatives $d^k u/dx^k$ of order $k \leq n$ exist and are continuous on $[a, b]$. It can be shown that $C^n[a, b]$ is a linear vector space with respect to the usual rules of vector addition and scalar multiplication.

4. The set

$$S_0 = \left\{ u : u(x) \in C^2[0, L], -\frac{d}{dx}\left(a(x)\frac{du}{dx}\right) + c(x)u = 0, \qquad 0 < x < L \right\}$$

is a vector space with respect to the usual addition and scalar multiplication. However, the set

$$S = \left\{ u : u(x) \in C^2[0, L], -\frac{d}{dx}\left(a(x)\frac{du}{dx}\right) + c(x)u = f(x), \qquad 0 < x < L \right\}$$

is not a linear vector space (why?).

5. Consider the transverse motion of a cable of length $L$, fixed at its ends (see Fig. 7.1). Let $C[0, L]$ denote the set of all real-valued, continuous functions $u(x, t)$ defined on the closed interval $0 \leq x \leq L$ for any time $t$. The transverse deflection $u(\cdot, t)$ (i.e., configuration) of the cable at any time $t$ can be viewed as an element of $C[0, L]$. However, not every element of $C[0, L]$ is a possible

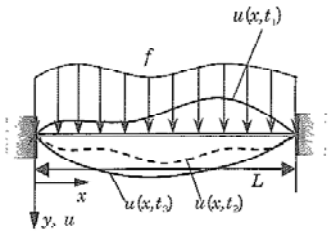


**Figure 7.1** Transverse motion of a cable fixed at both ends.



configuration of the cable because all possible configurations must pass through the points $x = 0$ and $x = L$; i.e., the boundary conditions $u(0, t) = 0$ and $u(L, t) = 0$ must be satisfied. Let $S$ be the subset of $C[0, L]$ made up of all real-valued, continuous functions $u(x, t)$ such that $u(0, t) = 0$ and $u(L, t) = 0$:

$$S = \{u: u(x) \in C[0, L], \qquad u(0, t) = 0, u(L, t) = 0\}.$$

Then $S$ is a subspace of $C[0, L]$, and all possible configurations (i.e., deflections) are contained in this space.

Let $U$ and $V$ be each a linear vector space. An *ordered pair* is a pair of elements $u \in U$ and $v \in V$ where one of the elements is designated as the first member of the pair and the other is designated as the second. We denote ordered pairs by $(u, v)$ with the obvious order. Then $U \times V$ is called a *product space* $W$ with elements $w = (u, v)$, $u \in U$ and $v \in V$:

$$W = \{w: w = (u, v), \ u \in U, \ v \in V\},$$

which is also a linear vector space with respect to the following definitions of vector addition and scalar multiplication of a vector in the product space $U \times V$:

$$(u_1, v_1) + (u_2, v_2) = (u_1 + u_2, v_1 + v_2), \tag{7.5a}$$

$$\alpha(u, v) = (\alpha u, \alpha v), \qquad \alpha \in \Re \tag{7.5b}$$

for $(u_1, v_1), (u_2, v_2) \in W = U \times V$ with $u_1, u_2 \in U$ and $v_1, v_2 \in V$.

Consider an open-bounded domain $\Omega \subset \Re^3$. Note that $\Omega$ is a set of points $\mathbf{x} = (x_1, x_2, x_3)$. A real-valued function $u(\mathbf{x})$ is said to be *square-integrable* in the domain $\Omega$ if the integrals (in the Lebesgue sense)

$$\int_\Omega u(\mathbf{x})\, d\mathbf{x}, \qquad \int_\Omega |u(\mathbf{x})|^2\, d\mathbf{x} \tag{7.6}$$

exist and are finite. The space of square-integrable functions $u$ defined over a domain $\Omega$ is called the $L_2$ space:

$$L_2(\Omega) = \left\{ u(\mathbf{x}): \int_\Omega |u(\mathbf{x})|^2\, d\mathbf{x} \right\}. \tag{7.7}$$

There is a corresponding space $L_\infty(\Omega)$, which consists of all real-valued functions $u(\mathbf{x})$ defined in the domain $\Omega$ such that there exists an $N$ with the property that

$$|u(\mathbf{x})| \leq N.$$

***Linear Independence*** Recall the concepts of coplanar and collinear vectors in Euclidean space from Chapter 2. These concepts can be generalized to function spaces. An expression of the form

$$\alpha_1 u_1 + \alpha_2 u_2 + \cdots = \sum_{i=1}^{n} \alpha_i u_i \tag{7.8}$$

for all functions $u_i(x)$ and scalars $\alpha_i \in \Re$ (real number field) is called a *linear combination* of $u_i$. The equation $\sum_i^n \alpha_i u_i = 0$ is called a *linear relation* among the functions $u_i$. A set of $n$ functions, $u_1, u_2, \ldots, u_n$, is said to be *linearly dependent* if a set of $n$ numbers, $\alpha_1, \alpha_2, \ldots, \alpha_n$, not all of which are zero, can be found such that the following linear relation holds:

$$\sum_{i=1}^{n} \alpha_i u_i = 0. \tag{7.9}$$

If there does not exist at least one nonzero number among $\alpha_i$ such that the above relation is satisfied, the vectors are said to be *linearly independent.*

**Example 7.2**

1. Consider the following set of polynomials, $\{p_i\}$, with

   $$p_1(x) = 1 + x, \qquad p_2(x) = 1 + x^2, \qquad p_3(x) = 1 + x + x^3.$$

   Consider the linear relation

   $$\alpha_1 p_1 + \alpha_2 p_2 + \alpha_3 p_3 = 0$$

   for $\alpha_i \in \Re$. Since the above relation must hold for all $x$, it follows that the coefficients of powers of $x$ must be zero separately. Collecting the coefficients of various powers of $x$ and setting them to zero, we obtain

   $$\alpha_1 + \alpha_2 + \alpha_3 = 0, \qquad \alpha_1 + \alpha_3 = 0, \qquad \alpha_2 = 0, \qquad \alpha_3 = 0.$$

   The solution to these equations is trivial (i.e., all $\alpha_i = 0$); hence, the set $\{p_1, p_2, p_3\}$ is linearly independent.
2. If $p_3$ is replaced by $p_4 = 2 + x + x^2$, we see that the linear relation

   $$\alpha_1 p_1 + \alpha_2 p_2 + \alpha_4 p_4 = 0$$

   requires that

   $$\alpha_1 + \alpha_2 + 2\alpha_4 = 0, \qquad \alpha_1 + \alpha_4 = 0, \qquad \alpha_2 + \alpha_4 = 0.$$

   An infinite number of solutions to the above set of equations exists. For example,

   $$\alpha_4 = 1, \qquad \alpha_1 = \alpha_2 = -1$$



is a solution. Hence, the set $\{p_1, p_2, p_4\}$ is linearly dependent. Indeed, $p_4$ can be expressed as a linear combination of $p_1$ and $p_2$:

$$p_4 = \alpha_1 p_1 + \alpha_2 p_2, \qquad \alpha_1 = \alpha_2 = 1.$$

### 7.2.3 Normed and Inner Product Spaces

***Norm*** The concepts of distance between two points and length of a physical vector can be generalized to abstract vectors, i.e., vectors that are functions. Let $V$ be a linear vector space over the real number field $\Re$. We shall use the notation $\|\cdot\|$ to denote the norm of real-valued functions $u(\mathbf{x})$, $\mathbf{x} \in \Omega \subset \Re^3$. Then, associated with every vector $u \in V$, there exists a real number $\|\cdot\| \in \Re$, called the *norm*, that satisfies certain rules, as discussed below. Thus the norm is the operation, $\|\cdot\| : V \to \Re$.

(1) Nonnegative:

(a) $\|u\| \geq 0$ for all $u$.

(b) $\|u\| = 0$ only if $u = 0$. (7.10)

(2) Homogeneous: $\|\alpha u\| = |\alpha| \, \|u\|$.

(3) Triangle inequality: $\|u + v\| \leq \|u\| + \|v\|$.

If $\|u\|$ satisfies (1a), (2), and (3), it is called a *seminorm*, and is denoted by $|u|$.

A linear vector space endowed with a norm is called a *normed vector space*. A linear subspace $S$ of a normed vector space $V$ is a linear subspace equipped with the norm of $V$.

A norm $\|\cdot\|$ can be used to define a notion of distance between vectors, called *natural metric*:

$$d(u, v) = \|u - v\| \qquad \text{for } u, v \in V. \tag{7.11}$$

Examples of norms will be given shortly.

For $1 \leq p \leq \infty$, we define the *Lebesgue spaces* [see Eq. (7.7)]:

$$L_p(\Omega) = \{u : \|u\|_p < \infty\}, \tag{7.12}$$

where

$$\|u\|_{L_p(\Omega)} \equiv \|u\|_p = \left[\int_\Omega |u(\mathbf{x})|^p \, d\mathbf{x}\right]^{1/p} < \infty. \tag{7.13}$$

For $p = \infty$ we set

$$\|u\|_{L_\infty(\Omega)} = \|u\|_\infty = \sup\{|u(\mathbf{x})| : \mathbf{x} \in \Omega\}. \tag{7.14}$$

This is called the "sup-norm."

Two norms $\|\cdot\|_1$ and $\|\cdot\|_2$ on a normed vector space $V$ are said to be *equivalent* if there exist positive numbers $c_1$ and $c_2$, independent of $u \in V$, such that the following double inequality holds:

$$c_1\|u\|_1 \le \|u\|_2 \le c_2\|u\|_1. \tag{7.15}$$

A normed space $V$ is called *complete* if every Cauchy sequence $\{u_j\}$ of elements of $V$ has a limit $u \in V$. For a normed vector space, a Cauchy sequence is one such that

$$\|u_j - u_k\| \to 0 \qquad \text{as } j, k \to \infty,$$

and completeness means that

$$\|u - u_j\| \to 0 \qquad \text{as } j \to \infty.$$

A normed vector space which is complete in its natural metric is called a *Banach space*. A linear subspace of a Banach space is itself a Banach space if and only if the subspace is complete.

**Example 7.3**

1. The $n$-dimensional Euclidean space $\Re^n$ is a Banach space with respect to the *Euclidean norm*:

$$\|\mathbf{x}\| \equiv \sqrt{\sum_{i=1}^{n} x_i^2}. \tag{7.16}$$

2. The space $C[0, 1]$ of real-valued continuous functions $f(x)$ defined on the closed interval $[0, 1]$ with the sup-norm (7.14) is a Banach space. It is a linear vector space with respect to the vector addition and scalar multiplication defined as

$$(f + g)(x) = f(x) + g(x), \qquad (\alpha f)(x) = \alpha f(x), \qquad \alpha \in \Re.$$

Further, it is complete with respect to the sup-norm in (7.14):

$$\|f\|_\infty \equiv \max |f(x)|.$$

3. *Sobolev space*, $W^{m,p}(\Omega)$. Let $C^m(\Omega)$ denote the set of all real-valued functions with $m$ continuous derivatives defined in $\Omega \in \Re^3$, and let $C^\infty(\Omega)$ denote the set of infinitely differentiable continuous functions. We define on $C^m(\Omega)$ the norm, called the *Sobolev norm*,

$$\|u\|_{m,p} = \left[\int_\Omega \sum_{|\alpha| \le m} |D^\alpha u(\mathbf{x})|^p \, d\mathbf{x}\right]^{1/p}, \tag{7.17}$$



for $1 \le p \le \infty$ and for all $u \in C^m(\Omega)$. In Eq. (7.17), $\alpha$ denotes an $n$-tuple of integers:

$$\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n), \qquad |\alpha| = \sum_i^n \alpha_i, \quad \alpha_i \ge 0,$$
$$D^\alpha = \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2} \cdots \partial x_n^{\alpha_n}}, \tag{7.18}$$

For $m = 1, n = 2$, and $1 \le p < \infty$, we have [$\alpha = (\alpha_1, \alpha_2), \alpha_1, \alpha_2 = 0, 1$], and

$$\|u\|_{1,p} = \left\{\int_{\Omega \subset \Re^2} \left[|u|^p + \left|\frac{\partial u}{\partial x}\right|^p + \left|\frac{\partial u}{\partial y}\right|^p\right] dx dy\right\}^{1/p}. \tag{7.19}$$

The space $C^m(\Omega)$ is not complete with respect to the Sobolev norm $\|\cdot\|_{m,p}$. The completion of $C^m(\Omega)$ with respect to the norm $\|\cdot\|_{m,p}$ is called the *Sobolev space of order* $(m, p)$, denoted by $W^{m,p}(\Omega)$. The completion of $C(\Omega)$ is the $L_2(\Omega)$ space. Hence the Sobolev space is a Banach space. Of course, the Lebesque space $L_p(\Omega)$ is a special case of the Sobolev space $W^{m,p}$ for $m = 0$, and $L_2(\Omega)$ is a special case of $L_p(\Omega)$ for $p = 2$, with the norms defined in (7.13).

If $U$ and $V$ are each normed vector spaces, we can define a norm on the product space $U \times V$ in one of the following ways:

$$\begin{aligned}
&(1)\ \|(u, v)\| = \|u\|_U + \|v\|_V. \\
&(2)\ \|(u, v)\| = \left(\|u\|_U^p + \|v\|_V^p\right)^{1/p}, \qquad p \ge 1. \\
&(3)\ \|(u, v)\| = \max\left(\|u\|_U, \|v\|_V\right).
\end{aligned} \tag{7.20}$$

Then $U \times V$ is a normed vector space with respect to any one of the above norms.

***Inner Product*** Analogous to the scalar product of physical vectors, the *inner product* of a pair of vectors $u$ and $v$ from an abstract vector space $V$ is defined to be a real number, denoted $(u, v)_V$ [i.e., $(\cdot, \cdot)_V : V \times V \to \Re$], which satisfies the following rules for every $u_1, u_2, u, v \in V$ and $\alpha \in \Re$:

$$\begin{aligned}
&(1)\ \text{Symmetry: } (u, v)_V = (v, u)_V. \\
&(2a)\ \text{Homogeneous: } (\alpha u, v)_V = \alpha(u, v)_V. \\
&(2b)\ \text{Additive: } (u_1 + u_2, v)_V = (u_1, v)_V + (u_2, v)_V. \\
&(3)\ \text{Positive-definite: } (u, u)_V > 0 \text{ for all } u \ne 0.
\end{aligned} \tag{7.21}$$

One can define a number of inner products and associated *natural norms* for pairs of functions that, along with their derivatives, are square-integrable. In particular,

the Sobolev space $W^{m,2}(\Omega) \equiv H^m(\Omega)$, which is also known as the *Hilbert space* of order $m$, is endowed with the inner product

$$(u, v)_m = \int_\Omega \sum_{|\alpha| \le m} D^\alpha u(\mathbf{x}) D^\alpha v(\mathbf{x})\, d\mathbf{x}, \tag{7.22}$$

for all $u, v \in H^m(\Omega)$. Note that for $m = 0$, we have $H^0(\Omega) = L_2(\Omega)$. Some special cases of Eq. (7.22) are given by

$$(u, v)_0 = \int_\Omega uv\, dxdy, \qquad \|u\|_0 = \sqrt{(u, u)_0}, \tag{7.23a}$$

$$(u, v)_1 = \int_\Omega \left( uv + \frac{\partial u}{\partial x}\frac{\partial v}{\partial x} + \frac{\partial u}{\partial y}\frac{\partial v}{\partial y} \right) dxdy, \qquad \|u\|_1 = \sqrt{(u, u)_1}, \tag{7.23b}$$

$$(u, v)_2 = \int_\Omega \left( uv + \frac{\partial u}{\partial x}\frac{\partial v}{\partial x} + \frac{\partial u}{\partial y}\frac{\partial v}{\partial y} + \frac{\partial^2 u}{\partial x \partial y}\frac{\partial^2 v}{\partial x \partial y} + \frac{\partial^2 u}{\partial x^2}\frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}\frac{\partial^2 v}{\partial y^2} \right) dxdy, \qquad \|u\|_2 = \sqrt{(u, u)_2}. \tag{7.23c}$$

A linear vector space on which an inner product can be defined is called an *inner product space*. A linear subspace $S$ of an inner product space $V$ is a subspace with the inner product of $V$. Note that the square root of the inner product of a vector with itself satisfies the axioms of a norm. Consequently, one can associate with every inner product in vector space $V$ a norm

$$\|u\|_V = \sqrt{(u, u)_V}. \tag{7.24}$$

The norm thus obtained is called the *norm induced by the inner product*. Since we can associate with each inner product a norm, every inner product space is also a normed vector space. It should be obvious to the reader that the converse does not hold in general.

***Orthogonality*** Two vectors $u, v \in V$ are said to be *orthogonal* if

$$(u, v)_V = 0, \tag{7.25}$$

where $(\cdot, \cdot)_V$ denotes an inner product in $V$. Note that the concept of orthogonality is a generalization of the familiar notion of perpendicularity of one vector to another in Euclidean space. A set of mutually orthogonal vectors is called an orthogonal set.

A sequence of functions $\{\phi_i\}$ in $L_2(\Omega)$ is called *orthonormal* if

$$(\phi_i, \phi_j)_V = \delta_{ij} = \begin{cases} 1, & \text{if } i = j, \\ 0, & \text{if } i \neq j. \end{cases} \tag{7.26}$$

Here $\delta_{ij}$ denotes the Kronecker delta. It can be shown that every orthonormal system is linearly independent.



If two vectors $u$ and $v$ of an inner product space $V$ are orthogonal, then the Pythagorean theorem holds even in function spaces:

$$\|u + v\|_V^2 = (u + v, u + v)_V = (u, u)_V + 2(u, v)_V + (v, v)_V = \|u\|_V^2 + \|v\|_V^2.$$

A set of functions $\{\phi_i\}$ is said to be *complete* in $L_2(\Omega)$ if every piecewise continuous function $f$ can be approximated in $\Omega$ by the sum $\sum_{j=1}^n c_j \phi_j$ in such a way that

$$\mathcal{E}_n \equiv \int_\Omega \left( f - \sum_{j=1}^n c_j \phi_j \right)^2 d\mathbf{x} \tag{7.27}$$

can be made as small as we wish (by increasing $n$). This property of the coordinate functions is the key to the proof of the convergence of the Ritz and Galerkin approximations.

A complete (in its natural metric) inner product space is called a *Hilbert space*. We mention without proof the fact that every inner product space (hence a normed space) has a completion.

The following lemma, referred to as the *fundamental lemma of variational calculus*, plays an important role in variational theory.

**Lemma 7.1** Let $V$ be an inner product space. If $(u, v)_V = 0$ for all $v \in V$, then $u = 0$.

***Proof:*** Since $(u, v)_V = 0$ for all $v$, it must also hold for $v = u$. Then $(u, u)_V = 0$ implies that $u = 0$.

### 7.2.4 Transformations, and Linear and Bilinear Forms

A transformation $T$ from a linear vector space $U$ into another linear vector space $V$ (both vector spaces are defined on the same field of scalars) is a correspondence that assigns to each element $u \in U$ a unique element $v = Tu \in V$. We use the terms *transformation*, *mapping*, and *operator* interchangeably, and the transformation is expressed as $T : U \to V$.

A transformation $T : U \to V$, where $U$ and $V$ vector spaces that have the same scalar field, is said to be *linear* if

1. $T(\alpha u) = \alpha T(u)$, for all $u \in U$, $\alpha \in \Re$ (homogeneous);
2. $T(u_1 + u_2) = T(u_1) + T(u_2)$, for all $u_1, u_2 \in U$, (additive). (7.28)

Otherwise it is said to be a *nonlinear transformation*.

Transformations that map vectors (functions) into real numbers are of special interest in the present study. Such transformations are called *functionals*. A linear transformation $l : V \to \Re$ that maps a linear vector space $V$ into the real number field $\Re$ is called a *linear functional*.

Similarly, a linear transformation that maps pairs of vectors $(u, v) \in V \times V$ into real number field $\Re$, or $B(\cdot, \cdot) : V \times V \to \Re$, is called a *bilinear form*. Examples of linear and bilinear forms are provided by

$$l(u) = \int_a^b fu\, dx, \qquad B(u, v) = \int_a^b \left( \frac{du}{dx}\frac{dv}{dx} + uv \right) dx.$$

A bilinear form is said to be *symmetric* if it is symmetric in its arguments:

$$B(u, v) = B(v, u). \tag{7.29}$$

### 7.2.5 Minimum of a Quadratic Functional

Consider an operator equation of the form

$$Au = f, \tag{7.30}$$

where $A$ is a certain operator (often a differential operator), $A : \mathcal{D}_A \to H$, and $f \in H$ is a given function. Here $\mathcal{D}_A$ denotes a set of elements from a Hilbert space $H$. The denseness of $\mathcal{D}_A$ in $H$ is often assumed, but we will not discuss this topic in the present study.

The differential equations

$$-\frac{d}{dx}\left( EA\frac{du}{dx} \right) = f(x), \qquad EA > 0, \qquad 0 < x < L, \tag{7.31}$$

$$\frac{d^2}{dx^2}\left( EI\frac{d^2u}{dx^2} \right) = f(x), \qquad EI > 0, \qquad 0 < x < L, \tag{7.32}$$

are special cases of the operator equation (7.30) with

$$A = -\frac{d}{dx}\left( EA\frac{d(\cdot)}{dx} \right), \qquad A = \frac{d^2}{dx^2}\left( EI\frac{d^2(\cdot)}{dx^2} \right),$$

respectively. In these cases, $H = L_2(0, L)$. The set $\mathcal{D}_A$ for Eq. (7.31) consists of functions from $C^2(0, L)$ and for (7.32) functions from $C^4(0, L)$.

An operator $A : \mathcal{D}_A \to H$ is called *symmetric* (or *self-adjoint*) if

$$(Au, v)_H = (u, Av)_H \tag{7.33}$$

holds for all $u, v \in \mathcal{D}_A$, where $(\cdot, \cdot)_H$ is the inner product in $H$. An operator $A$ is called *strictly positive* in $\mathcal{D}_A$ if it is symmetric in $\mathcal{D}_A$ and if

$$(Au, u)_H > 0 \text{ holds for all } u \in \mathcal{D}_A \quad \text{and} \quad u \neq 0, \tag{7.34a}$$

$$(Au, u)_H = 0 \text{ if and only if } u \in \mathcal{D}_A \quad \text{and} \quad u = 0. \tag{7.34b}$$



A *quadratic functional* $Q : H \to \Re$ is one that is quadratic in its arguments, $Q(\alpha u) = \alpha^2 Q(u)$ for $\alpha \in \Re$. Every bilinear form $B(\cdot, \cdot)$ can be used to generate a quadratic form $Q$ by setting

$$Q(u) = B(u, u), \qquad u \in H. \tag{7.35}$$

The following results are of fundamental importance for the present study.

**Theorem 7.1** If $A$ is a strictly positive operator in $\mathcal{D}_A$, then

$$Au = f \qquad \text{in } H$$

has at most one solution $u \in \mathcal{D}_A$ in $H$.

***Proof:*** Suppose that there exist two solutions $u_1, u_2 \in \mathcal{D}_A$. Then

$$Au_1 = f \quad \text{and} \quad Au_2 = f \;\to\; A(u_1 - u_2) = 0 \quad \text{in } H$$

and

$$(A(u_1 - u_2), u_1 - u_2)_H = 0 \;\to\; u_1 - u_2 = 0 \quad \text{or} \quad u_1 = u_2 \in \mathcal{D}_A,$$

which was to be proved.

**Theorem 7.2** Let $A$ be a positive operator in $\mathcal{D}_A$, $f \in H$. Let Eq. (7.30) have a solution $u_0 \in \mathcal{D}_A$. Then the quadratic functional

$$I(u) = \tfrac{1}{2}(Au, u)_H - (f, u)_H \tag{7.36}$$

assumes its minimal value in $\mathcal{D}_A$ for the element $u_0$, i.e.,

$$I(u) \geq I(u_0), \qquad \text{and} \qquad I(u) = I(u_0) \qquad \text{only for } u = u_0.$$

Conversely, if $I(u)$ assumes its minimal value, among all $u \in \mathcal{D}_A$, for the element $u_0$, then $u_0$ is the solution of Eq. (7.30) (i.e., $Au_0 = f$).

***Proof:*** First note that $I(u)$ is defined for all $u \in \mathcal{D}_A$. Let $u_0$ be the solution of Eq. (7.30). Then $f = Au_0$. Substituting for $f$ into Eq. (7.36), we obtain for $u \in \mathcal{D}_A$:

$$\begin{aligned} I(u) &= \tfrac{1}{2}(Au, u)_H - (Au_0, u)_H \\ &= \tfrac{1}{2}\left[(Au, u)_H - (Au_0, u)_H - (u, Au_0)_H\right] \\ &= \tfrac{1}{2}\left[(Au, u)_H - (Au_0, u)_H - (Au, u_0)_H\right] \\ &= \tfrac{1}{2}\left[(Au, u)_H - (Au_0, u)_H - (Au, u_0)_H + (Au_0, u_0)_H - (Au_0, u_0)_H\right] \\ &= \tfrac{1}{2}\left[(A(u - u_0), u - u_0)_H - (Au_0, u_0)_H\right], \end{aligned} \tag{7.37}$$

where the linearity and symmetry of $A$, as well as the symmetry of the bilinear form, are used in arriving at the last step. From Eq. (7.37) it follows that

$$I(u_0) = -\tfrac{1}{2}(Au_0, u_0)_H. \tag{7.38}$$

Next, we use the strictly positive property of $A$ to conclude that

$$I(u) \geq I(u_0) \qquad \text{for } u \in \mathcal{D}_A, \qquad \text{and} \qquad I(u) = I(u_0)$$

if and only if $u = u_0$ in $\mathcal{D}_A$. Consequently, if the equation $Au_0 = f$ is satisfied, then the functional $I(u)$ assumes its minimal value in $\mathcal{D}_A$ precisely for the element $u = u_0$.

Now suppose that $I(u)$ assumes its minimal value in $\mathcal{D}_A$ for the element $u_0$. This implies that

$$I(u_0 + \alpha v) \geq I(u_0) \quad \text{for } \alpha \in \mathfrak{R},\ v \in \mathcal{D}_A. \tag{7.39}$$

Using again the symmetry of $A$ and the symmetry of the inner product, one obtains

$$\begin{aligned} I(u_0 + \alpha v) &= \tfrac{1}{2}(A(u_0 + \alpha v), u_0 + \alpha v)_H - (f, u_0 + \alpha v)_H \\ &= \tfrac{1}{2}\left[(Au_0 + \alpha Av, u_0 + \alpha v)_H - 2(f, u_0)_H - 2\alpha(f, v)_H\right] \\ &= \tfrac{1}{2}\big[(Au_0, u_0)_H + \alpha(Av, u_0)_H + \alpha(Au_0, v)_H \\ &\quad + \alpha^2(Av, v)_H - 2(f, u_0)_H - 2\alpha(f, v)_H\big] \\ &\quad - \tfrac{1}{2}\big[(Au_0, u_0)_H + 2\alpha(Au_0, v)_H + \alpha^2(Av, v)_H \\ &\quad - 2(f, u_0)_H - 2\alpha(f, v)_H\big]. \end{aligned} \tag{7.40}$$

Since $u_0 \in \mathcal{D}_A$ and $f \in H$ are fixed elements, it is obvious that for arbitrarily fixed $v \in \mathcal{D}_A$ the function $I(u_0 + \alpha v)$ is a quadratic function in the variable $\alpha$. From Eq. (7.39) it follows that the function has a local minimum at $\alpha = 0$, which implies that its first derivative is equal to zero at $\alpha = 0$ [or, equivalently, the first variation of $I$ is zero; see Eq. (4.89)]:

$$\left[\frac{d}{d\alpha} I(u_0 + \alpha v)\right]_{\alpha=0} = 0,$$

or by Eq. (7.40) that

$$(Au_0, v)_H - (f, v)_H = 0 \quad \text{or} \quad (Au_0 - f, v)_H = 0.$$

Since $v \in H$ is arbitrary, by Lemma 7.1 it follows that $Au_0 - f = 0$ in $H$.

**Example 7.4** Consider the differential equation

$$-\frac{d}{dx}\left(a(x)\frac{du}{dx}\right) = f(x), \qquad a(x) > 0, \qquad 0 < x < L, \tag{7.41}$$



subjected to the boundary conditions

$$u(0) = 0, \qquad u(L) = 0, \tag{7.42}$$

which arise in connection with the transverse deflection of cables. Here $u(x)$ denotes the deflection of a cable of original length $L$, tension $a = a(x)$, and subjected to distributed transverse load $f(x)$ (see Fig. 7.2). The boundary conditions in Eq. (7.42) indicate that the cable is fixed at $x = 0$ and $x = L$.

Let us choose $H = L_2(0, L)$, and define $\mathcal{D}_A$ as the linear set of functions that are continuous with their derivatives up to the second order inclusive in the interval $[0, L]$ and satisfy the end conditions in (7.42). Define the operator $A$ on $\mathcal{D}_A$ by

$$Au = -\frac{d}{dx}\left(a(x)\frac{du}{dx}\right). \tag{7.43}$$

We now set out to prove that $A$ is strictly positive on $\mathcal{D}_A$. First, we note that $A$ is symmetric in $\mathcal{D}_A$: For every $u \in \mathcal{D}_A$ and $v \in \mathcal{D}_A$, we have

$$\begin{aligned} (Au, v)_H &= \int_0^L \left[-\frac{d}{dx}\left(a\frac{du}{dx}\right)\right] v\, dx \\ &= -\left[a\frac{du}{dx}v\right]_0^L + \int_0^L \frac{dv}{dx}\left(a\frac{du}{dx}\right) dx \\ &= \int_0^L a(x)\frac{dv}{dx}\frac{du}{dx}\, dx \qquad (7.44) \\ &\quad - \left[a\frac{dv}{dx}u\right]_0^L + \int_0^L \left[-\frac{d}{dx}\left(a\frac{dv}{dx}\right)\right] u\, dx \\ &= \int_0^L \left[-\frac{d}{dx}\left(a\frac{dv}{dx}\right)\right] u\, dx = (u, Av)_H, \qquad (7.45) \end{aligned}$$

where we have used the fact that $u(0) = u(L) = v(0) = v(L) = 0$. Thus, $A$ is symmetric on $\mathcal{D}_A$. From Eq. (7.44), it follows that

$$(Au, u)_H = \int_0^L a(x)\left(\frac{du}{dx}\right)^2 dx \qquad \text{for all } u \in \mathcal{D}_A. \tag{7.46}$$

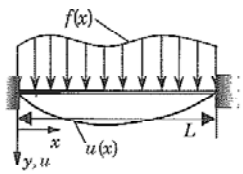


**Figure 7.2** Transverse deflection of a cable fixed at its ends.

Due to the fact that $a(x) > 0$ in $[0, L]$, it follows from (7.46) that

$$(Au, u)_H \geq 0 \qquad \text{for every } u \in \mathcal{D}_A.$$

Moreover, if $(Au, u)_H = 0$, it follows that

$$\frac{du}{dx} = 0 \qquad \text{in } [0, L].$$

This in turn implies that

$$u(x) = c, \text{ constant in } [0, L].$$

Since $u(0) = 0$, it follows that $c = 0$ or $u(x) = 0$ in $[0, L]$. This proves that $A$ is strictly positive in $\mathcal{D}_A$.

The quadratic functional associated with Eqs. (7.41) and (7.42) is given by

$$\Pi(u) = \frac{1}{2}(Au, u)_H - (f, u)_H = \frac{1}{2}\int_0^L a(x)\left(\frac{du}{dx}\right)^2 dx - \int_0^L fu\,dx, \quad (7.47)$$

which represents the total potential energy of the cable. The first term is the elastic strain energy and the second term is the potential energy of the external load $f(x)$.

Let $u_0(x)$ be the solution of Eqs. (7.41) and (7.42). Then $f = Au_0$ and we have

$$\begin{aligned}\Pi(u) &= \frac{1}{2}(Au, u)_H - (Au_0, u)_H \\ &= \frac{1}{2}\int_0^L a(x)\left(\frac{du}{dx}\right)^2 dx - \int_0^L \left[-\frac{d}{dx}\left(a(x)\frac{du_0}{dx}\right)\right] u\,dx \\ &= \frac{1}{2}\int_0^L a(x)\left(\frac{du}{dx}\right)^2 dx - \int_0^L a(x)\frac{du_0}{dx}\frac{du}{dx}\,dx \\ &= \frac{1}{2}\int_0^L a(x)\left(\frac{du}{dx} - \frac{du_0}{dx}\right)^2 dx - \frac{1}{2}\int_0^L a(x)\left(\frac{du_0}{dx}\right)^2 dx. \end{aligned} \quad (7.48)$$

Since $a(x)(u' - u_0')^2 \geq 0$, it is clear from Eq. (7.48) that $\Pi(u)$ is minimal in $\mathcal{D}_A$ if and only if $u' = u_0'$ in $\mathcal{D}_A$. Thus, if $u_0 \in \mathcal{D}_A$ is the solution of Eqs. (7.41) and (7.42), then the functional $\Pi(u)$ assumes its minimum for $u_0 \in \mathcal{D}_A$.

Conversely, let $u_0 \subset \mathcal{D}_A$ be the element minimizing the functional $\Pi(u)$ in (7.47). Let $v \in \mathcal{D}_A$ be an arbitrary element from $\mathcal{D}_A$ and let $\alpha$ be an arbitrary real number. Then the minimum of $\Pi(u)$ implies that $(u = u_0 + \alpha v)$

$$\begin{aligned} 0 &= \frac{d}{d\alpha}\Pi(u_0 + \alpha v)\Big|_{\alpha=0} \\ &= \frac{d}{d\alpha}\left[\frac{1}{2}\int_0^L a(x)\left(\frac{du}{dx}\right)^2 dx - \int_0^L fu\,dx\right]_{\alpha=0} \end{aligned}$$



$$\begin{aligned} &= \int_0^L a(x)\frac{du_0}{dx}\frac{dv}{dx}\,dx - \int_0^L fv\,dx \\ &= \int_0^L \left[\frac{d}{dx}\left(a(x)\frac{du_0}{dx}\right) - f\right] v\,dx. \end{aligned}$$

Since this result must hold for every $v$, it follows that

$$-\frac{d}{dx}\left(a(x)\frac{du_0}{dx}\right) - f = 0 \qquad \text{in } H = L_2(0, L).$$

The above arguments are equivalent to the principle of minimum total potential energy discussed in Chapter 5.

Recall that the operator $A : \mathcal{D}_A \to H$ is symmetric:

$$(Au, v)_H = (u, Av)_H \qquad \text{for } u, v \in \mathcal{D}_A$$

and positive definite on $\mathcal{D}_A$, i.e., there exists a constant $C > 0$ so that

$$(Au, u)_H \geq C\|u\|^2 \qquad \text{holds for every } u \in \mathcal{D}_A. \quad (7.49)$$

Hence, we can define a new inner product $(u, v)_A$ on $\mathcal{D}_A$ as follows:

$$(u, v)_A = (Au, v)_H \qquad \text{for all } u, v \in \mathcal{D}_A. \quad (7.50)$$

It can be easily verified that $(u, v)_A$ satisfies the axioms (1)–(3) in Eq. (7.21) of an inner product. The linear set $\mathcal{D}_A$ with the inner product (7.50) constitutes a linear vector space, called the *energy space*, and denoted by $H_A$. The norm and natural metric follow from the definition in (7.50):

$$\|u\|_A^2 = (u, v)_A, \qquad d(u, v) = \|u - v\|_A. \quad (7.51)$$

The energy space $H_A$ can be shown to be complete with respect to the metric defined in (7.51), and hence it is a Hilbert space. Moreover, it can be shown that the functional $\Pi(u)$ can be extended to all elements of $H_A$, that the functional assumes its minimum at $u_0 \in H_A$, and that the element $u_0$ is uniquely determined by the element $f \in H$. These proofs are beyond the scope of the present study, and interested readers may consult Refs. [4], and [12–15].

## 7.3 THE RITZ METHOD

### 7.3.1 Introduction

As discussed in Chapters 5 and 6, the principles of virtual displacements and forces as applied to continuous systems can be used to determine the governing equations and natural boundary conditions of the problem. The energy methods (e.g., unit-dummy-force and unit-dummy-displacement methods and Castigliano's Theorems I and II)

derived from these principles were used to determine deflections and forces at selected points. Here we consider a powerful method of determining approximate solutions to the governing equations of a problem by directly using the variational statements (i.e., virtual work principles, the principle of total potential energy, or the principle of complementary energy). The method bypasses the derivation of the Euler equations and goes directly from a variational statement of the problem to the solution of the Euler equations. One such direct method was proposed by German engineer W. Ritz (1878–1909).

### 7.3.2 Description of the Method

Consider the linear operator equation

$$Au = f \qquad \text{in } \Omega, \tag{7.52}$$

where $A$ is a strictly positive operator from $H_A$ into $H$, and $f \in H$. The solution $u_0$ of Eq. (7.52) is the element $u_0 \in H_A$ that minimizes the quadratic functional

$$I(u) = \frac{1}{2}(u, u)_A - l(u), \tag{7.53}$$

where $l(\cdot)$ denotes a linear functional. In structural mechanics problems, the functional $I(u)$ represents the total potential energy and $\delta I(u) = 0$ yields Eq. (7.52) as the Euler equation.

We seek an approximation $U_N(x)$ of $u_0(x)$, for a fixed and preselected $N$, in the form

$$u_0(x) \approx U_N(x) = \sum_{i=1}^{N} c_i \phi_i(x) + \phi_0(x), \tag{7.54}$$

where $\phi_i(x)$ are the elements of a base in $H_A$, and $c_i$ are as yet unknown real constants. These constants are determined by the condition that $I(U_N)$ is the minimum. Since $\{\phi_i\}$ is a base in $H_A$, the solution $u_0 \in H_A$ can be approximated to arbitrary accuracy by a suitable linear combination of its elements. Therefore, it can be expected that the approximate solution $U_N$, with constants determined by minimizing the functional $I(U_N)$, will differ sufficiently slightly in $H_A$ from the actual solution $u_0$ if $N$ is selected sufficiently large. This process of determining $U_N$ is known as the Ritz method.

To fully illustrate the basic idea of the Ritz method described above, we consider the axial deformation of a nonuniform bar with an end spring. The governing equation is

$$-\frac{d}{dx}\left(EA(x)\frac{du}{dx}\right) = f(x), \qquad 0 < x < L, \tag{7.55a}$$

subjected to the boundary conditions

$$u(0) = 0, \qquad \left[\left(EA(x)\frac{du}{dx}\right) + ku(x)\right]_{x=L} = P. \tag{7.55b}$$



where $E$ denotes Young's modulus, $A = A(x)$ the area of cross section, $L$ the length, $k$ the spring constant, $f(x)$ the distributed axial load, and $P$ the axial load at $x = L$ (see Fig. 7.3). The space $H_A$ in this case is the completion of the set $\mathcal{D}_A$ of functions that are continuous with their derivatives up to the second order in $[0, L]$.

As discussed earlier, the problem is equivalent to minimizing the total potential energy functional $\Pi$:

$$\Pi(u) = \int_0^L \left[\frac{EA}{2}\left(\frac{du}{dx}\right)^2 - fu\right] dx + \frac{k}{2}[u(L)]^2 - Pu(L). \tag{7.56}$$

The necessary condition for the minimum of $\Pi$ is

$$0 = \delta\Pi = B(\delta u, u) - l(\delta u) \qquad \text{or} \qquad B(\delta u, u) = l(\delta u), \tag{7.57a}$$

where $B(\cdot, \cdot)$ is the bilinear form and $l(\cdot)$ is the linear functional. Equation (7.57a) is known as the *variational problem* associated with Eq. (7.52). The inner product in $H_A$ is defined by

$$(u, v)_A = B(u, v). \tag{7.57b}$$

For the specific case of $\Pi(u)$ in Eq. (7.56), the bilinear and linear forms are

$$B(u, v) = \int_0^L EA\frac{du}{dx}\frac{dv}{dx}dx + ku(L)v(L), \qquad l(u) = \int_0^L fu\,dx + Pu(L). \tag{7.58}$$

The Euler equations and natural boundary conditions associated with the minimization of $\Pi(u)$ in Eq. (7.56) are

$$-\frac{d}{dx}\left(EA\frac{du}{dx}\right) - f = 0 \qquad \text{in } 0 < x < L, \tag{7.59a}$$

$$EA\frac{du}{dx} + ku - P = 0 \qquad \text{at } x = L. \tag{7.59b}$$

The essential boundary condition of the problem is provided by the geometric constraint

$$u(0) = 0. \tag{7.59c}$$

The exact solution $u$ to the problem is one that satisfies Eq. (7.59a) at every $x \in (0, L)$ and the boundary conditions in Eqs. (7.59b,c). Thus, the solution of Eqs. (7.59a–c)

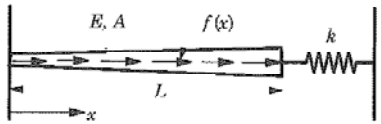


**Figure 7.3** Axial deformation of a nonuniform bar with an end spring.

is equivalent to minimizing $\Pi(u)$ over a set of functions that satisfy the condition in (7.59c).

In the Ritz method, we seek an approximate solution $U_N$ (which may be exact if we choose the right kind of approximate solution) to the problem as a finite linear combination of the form (7.54). The reason for selecting the particular form of the approximate solution will be apparent in the sequel. If we select $\phi_0$ and $\phi_i$ such that $U_N$ satisfies the specified essential boundary condition, $u(0) = 0$, and substitute $U_N$ into the total potential energy functional $\Pi$ in Eq. (7.56), we obtain $\Pi$ as a function of the parameters $c_1, c_2, \ldots, c_N$ (after carrying out the indicated integration with respect to $x$):

$$\Pi = \Pi(c_1, c_2, \ldots, c_N).$$

Then $c_i$ are determined (or adjusted) such that $\delta\Pi = 0$; in other words, we minimize $\Pi$ with respect to $c_i$, $i = 1, 2, \ldots, N$:

$$0 = \delta\Pi = \frac{\partial \Pi}{\partial c_1}\delta c_1 + \frac{\partial \Pi}{\partial c_2}\delta c_2 + \cdots + \frac{\partial \Pi}{\partial c_N}\delta c_N = \sum_{j=1}^{N} \frac{\partial \Pi}{\partial c_j}\delta c_j.$$

Since the set $\{c_i\}$ is linearly independent, it follows that

$$\frac{\partial \Pi}{\partial c_i} = 0 \qquad \text{for } i = 1, 2, \ldots, N \tag{7.60a}$$

or

$$[A]\{c\} = \{b\}. \tag{7.60b}$$

Equations (7.60a,b) represents a set of $N$ linear equations among $c_1, c_2, \ldots, c_N$, whose solution together with Eq. (7.54) yields the approximate solution $U_N(x)$. This completes the description of the Ritz method.

Since the natural boundary conditions of the problem are included in the variational statement, we require the approximate solution $U_N$ to satisfy only the essential boundary conditions. In order for $U_N$ to satisfy the essential boundary conditions *for any* $c_i$, it is convenient to choose the approximation in the form (7.54) and require $\phi_0(x)$ to satisfy the specified essential boundary conditions. For instance, if $u(x)$ is specified to be $\hat{u}$ at $x = 0$, we require $\phi_0(x)$ be such that $\phi_0(0) = \hat{u}$. Then

$$U_N(0) = \sum_{i=1}^{N} c_i\phi_i(0) + \phi_0(0) = \sum_{i=1}^{N} c_i\phi_i(0) + \hat{u}. \tag{7.61a}$$

Since $U_N(0) = \hat{u}$, it follows that

$$\sum_{i-1}^{N} c_i\phi_i(0) = 0 \;\rightarrow\; \phi_i(0) = 0 \qquad \text{for all } i = 1, 2, \ldots, N. \tag{7.61b}$$



Thus, $\phi_i(x)$ must satisfy the *homogeneous form* of specified essential boundary conditions.

Equation (7.54) can be viewed as a representation of $u$ in a component form; the parameters $c_i$ are the *components* (or coordinates) and $\{\phi_i\}$ are the *coordinate functions*. Another interpretation of Eq. (7.54) is provided by the finite Fourier series, in which $c_i$ are known as the *Fourier coefficients*.

### 7.3.3 Properties of Approximation Functions

The *approximation functions* $\phi_0$ and $\phi_i$ should be such that the substitution of Eq. (7.54) into $\delta\Pi$ or its equivalent results in $N$ linearly independent equations for the parameters $c_j$ ($j = 1, 2, \ldots, N$) so that the system has a solution. To ensure that the algebraic equations resulting from the Ritz approximation have a solution, and the approximate solution $U_N(x)$ converges to the true solution $u(x)$ of the problem as the value of $N$ is increased, $\phi_i$ ($i = 1, 2, \ldots, N$) and $\phi_0$ must satisfy certain requirements. Before we list the requirements, it is informative to discuss the concepts of completeness of a set of functions and convergence of a sequence of approximations. To make the ideas presented simple to understand, mathematical rigor is sacrificed.

*Convergence* A sequence $\{U_N\}$ of functions is said to *converge* to $u$ if for each $\epsilon > 0$ there is a number $M > 0$, depending on $\epsilon$, such that

$$\|U_N(\mathbf{x}) - u(\mathbf{x})\| < \epsilon \qquad \text{for all } N > M,$$

where $\|\cdot\|$ denotes a norm of the enclosed quantity and $u$ is called the limit of the sequence. In the above statement, $U_N$ represents the approximate solution and $u$ the true solution. The statement implies that the $N$-parameter solution $U_N$ can be made as close to $u$ as we wish, say within $\epsilon$, by choosing $N$ to be greater than $M = M(\epsilon)$, provided that the approximate solution is convergent. While there is no formula to determine $M$, a series of trials will help determine the value of $N$ for which the approximate solution $U_N$ is within the tolerance.

*Completeness* The concept of convergence of a sequence involves a limit of the sequence. If the limit is not a part of the sequence $\{U_N\}_{N=1}^{\infty}$, then there is no hope of attaining convergence. For example, if the true solution to a certain problem is of the form $u(x) = ax^2 + bx^3 + cx^5$, where $a$, $b$, and $c$ are constants, then the sequence of approximations

$$U_1 = c_1x^3,\ U_2 = c_1x^3 + c_2x^4, \ldots, \qquad U_N = c_1x^3 + c_2x^4 + \cdots + c_Nx^{N+2}$$

*will not* converge to the true solution because the sequence does not contain the $x^2$ term. The sequence is said to be incomplete. As a rule, in selecting an approximate solution one should include all terms up to the highest-order term. If a certain term is not a part of the true solution, like the $x^4$ term, its coefficient will turn out to be zero by the time all terms of the true solution are included in the approximation.

We now list the requirements of a convergent Ritz approximation (7.54):

1. $\phi_0$ must satisfy the *specified* essential boundary conditions. When the specified essential boundary conditions are homogeneous, then $\phi_0(x) = 0$.
2. $\phi_i \in H_A$ $(i = 1, 2, \ldots, N)$ must satisfy the following three conditions:
   (a) be continuous, as required by the variational statement being used;
   (b) satisfy the *homogeneous form* of the specified essential boundary conditions; and
   (c) the set $\{\phi_i\}$ must be linearly independent and complete. (7.62)

### 7.3.4 Ritz Equations for the Parameters

Returning to the problem of determining the approximate solution $U_N(x)$ of the bar problem described by Eqs. (7.59a,b), we substitute Eq. (7.54) into the total potential energy functional $\Pi$:

$$\Pi(c_1, c_2, \ldots, c_N) = \int_0^L \left[ \frac{EA}{2} \left( \sum_{j=1}^N c_j \frac{d\phi_j}{dx} + \frac{d\phi_0}{dx} \right)^2 - f \left( \sum_{j=1}^N c_j \phi_j + \phi_0 \right) \right] dx + \frac{k}{2} \left( \sum_{j=1}^N c_j \phi_j(L) + \phi_0(L) \right)^2 - P \left( \sum_{j=1}^N c_j \phi_j(L) + \phi_0(L) \right).$$

Now differentiating $\Pi$ with respect to $c_i$ $(i = 1, 2, \ldots, N)$, we obtain $N$ linearly independent equations for the unknowns, $c_1, c_2, \ldots, c_N$. We have

$$0 = \frac{\partial \Pi}{\partial c_1} = \int_0^L \left[ EA \frac{d\phi_1}{dx} \left( \sum_{j=1}^N c_j \frac{d\phi_j}{dx} + \frac{d\phi_0}{dx} \right) - f\phi_1 \right] dx + k\phi_1(L) \left( \sum_{j=1}^N c_j \phi_j(L) + \phi_0(L) \right) - P\phi_1(L), \qquad \text{(Eq. 1)}$$

$$0 = \frac{\partial \Pi}{\partial c_2} = \int_0^L \left[ EA \frac{d\phi_2}{dx} \left( \sum_{j=1}^N c_j \frac{d\phi_j}{dx} + \frac{d\phi_0}{dx} \right) - f\phi_2 \right] dx + k\phi_2(L) \left( \sum_{j=1}^N c_j \phi_j(L) + \phi_0(L) \right) - P\phi_2(L), \qquad \text{(Eq. 2)}$$

.....



$$0 = \frac{\partial \Pi}{\partial c_i} = \int_0^L \left[ EA \frac{d\phi_i}{dx} \left( \sum_{j=1}^N c_j \frac{d\phi_j}{dx} + \frac{d\phi_0}{dx} \right) - f\phi_i \right] dx + k\phi_i(L) \left( \sum_{j=1}^N c_j \phi_j(L) + \phi_0(L) \right) - P\phi_i(L), \qquad \text{(Eq. } i\text{)}$$

.....

$$0 = \frac{\partial \Pi}{\partial c_N} = \int_0^L \left[ EA \frac{d\phi_N}{dx} \left( \sum_{j=1}^N c_j \frac{d\phi_j}{dx} + \frac{d\phi_0}{dx} \right) - f\phi_N \right] dx + k\phi_N(L) \left( \sum_{j=1}^N c_j \phi_j(L) + \phi_0(L) \right) - P\phi_N(L). \qquad \text{(Eq. } N\text{)}$$

Note that Eq. $i$ is the $i$th equation of the set of $N$ equations. The $i$th equation can be written in the short form

$$0 = \sum_{j=1}^N a_{ij} c_j - b_i, \qquad (7.63a)$$

where the coefficients $a_{ij}$ and $b_i$ are defined by

$$a_{ij} = \int_0^L EA \frac{d\phi_i}{dx} \frac{d\phi_j}{dx} \, dx + k\phi_i(L)\phi_j(L),$$

$$b_i = -\int_0^L EA \frac{d\phi_0}{dx} \frac{d\phi_i}{dx} \, dx - k\phi_0(L)\phi_i(L) + \int_0^L f\phi_i \, dx + P\phi_i(L). \qquad (7.63b)$$

The $N$ equations can be written in matrix form as

$$[A]\{c\} = \{b\}. \qquad (7.64)$$

Equations (7.64) are called the *equations for the Ritz parameters* $c_i$. Once $c_i$ $(i = 1, 2, \ldots, N)$ are determined from Eq. (7.64), the approximate solution of the problem is given by Eq. (7.54). This displacement can be used to evaluate strains and stresses:

$$\varepsilon_{xx} = \frac{du}{dx} \approx \sum_{i=1}^N c_i \frac{d\phi_i}{dx} + \frac{d\phi_0}{dx}, \qquad \sigma_{xx} = E\varepsilon_{xx} \approx E \left( \sum_{i=1}^N c_i \frac{d\phi_i}{dx} + \frac{d\phi_0}{dx} \right).$$

Equations (7.63a,b) can also be arrived at by substituting the Ritz approximation (7.54) and its variation

$$\delta u \approx \delta U_N = \sum_{i=1}^{N} \delta c_i \phi_i(x) \tag{7.65}$$

in the variational statement $\delta\Pi = 0$, instead of substituting (7.54) in $\Pi$ and then taking the variation with respect to $c_i$. This results in

$$0 = \sum_{i=1}^{N} \delta c_i \left\{ \int_0^L \left[ EA \frac{d\phi_i}{dx} \left( \sum_{j=1}^{N} c_j \frac{d\phi_j}{dx} + \frac{d\phi_0}{dx} \right) - f\phi_i \right] dx \right.$$
$$\left. + k\phi_i(L) \left( \sum_{j=1}^{N} c_j \phi_j(L) + \phi_0(L) \right) - P\phi_i(L) \right\}.$$

Since $\delta c_i$ are arbitrary, we obtain the result in Eq. (7.63a).

Next we discuss the task of selecting the approximation functions $\phi_0$ and $\phi_i$. The properties listed in Eq. (7.62) provide guidelines for selecting the coordinate functions $\phi_0(x)$ and $\phi_i(x)$; they do not, however, give any formulae for generating the functions. Thus, apart from the guidelines, the selection of the coordinate functions is largely arbitrary. As a general rule, the coordinate functions $\phi_i$ should be selected from an admissible set [i.e., those meeting the conditions in Eq. (7.62)], from the lowest order to a desirable order, without missing any intermediate terms (i.e., the completeness property). Also, $\phi_0$ should be any lowest order (including zero) that satisfied the specified essential boundary conditions of the problem; it has no continuity (differentiability) requirement.

For the problem at hand, $\phi_0 = 0$ since the specified essential boundary condition is homogeneous. Next, we find $\phi_1(x)$ such that $\phi_1(0) = 0$ and differentiable at least once with respect to $x$ because $\Pi$ involves the first derivatives of $u \approx U_N$. If an algebraic polynomial is to be selected, the lowest-order polynomial that has a nonzero first derivative is

$$\phi_1(x) = a + bx,$$

where $a$ and $b$ are constants to be determined. The condition $\phi_1(0) = 0$ gives $a = 0$. Since $b$ is arbitrary, we take it to be unity (any nonzero constant will be absorbed into $c_1$). When $N > 1$, property 2(c) in Eq. (7.62) requires that $\phi_i$, $i > 1$, should be selected such that the set $\{\phi_i\}_{i=1}^{N}$ is linearly independent and makes the set complete. In the present case, this is done by choosing $\phi_2$ to be $x^2$. Clearly, $\phi_2(x) = x^2$ meets the conditions $\phi_2(0) = 0$, linearly independent of $\phi_1(x) = x$ (i.e., $\phi_2$ is not a constant multiple of $\phi_1$), and the set $\{x, x^2\}$ is complete (i.e., no other admissible term up to quadratic is omitted). In other words, in selecting coordinate functions of a given degree, one should not omit any lower-order terms that are admissible. Otherwise the approximate solution will never converge to the exact solution, no matter how many



terms are used in the Ritz approximation, as the exact solution may have those lower-order terms that were omitted in the approximate solution. Note that $\hat{\phi}_2(x) = x + x^2$ is also an admissible function that meets all requirements. Then

$$U_2(x) = c_1\phi_1 + c_2\hat{\phi}_2 = \hat{c}_1 x + \hat{c}_2 x^2, \qquad \text{with } \hat{c}_1 = c_1 + c_2,\ \hat{c}_2 = c_2,$$

which is equivalent to

$$U_2(x) = c_1\phi_1 + c_2\phi_2.$$

Thus, one may select $\phi_i = x^i$, $i = 1, 2, \ldots, N$.

If trigonometric functions are to be selected, one may be tempted to select $\phi_1 = \sin(\pi x/L)$, which satisfies the condition $\phi_1(0) = 0$. However, this choice also gives $U_1(L) = 0$ since $\phi_1(L) = 0$. A better choice would be to select $\phi_1(x) = \sin(\pi x/2L)$, or for $N > 1$, select $\phi_i = \sin[(2i-1)\pi x/2L]$; this choice will yield a good solution.

For the choice of algebraic polynomials, the $N$-parameter Ritz approximation for the bar problem is

$$U_N(x) = \sum_{i=1}^{N} c_i \phi_i(x), \qquad \phi_i(x) = x^i, \tag{7.66}$$

and the coefficients $a_{ij}$ and $b_i$ for

$$EA = a_0\left(2 - \frac{x}{L}\right), \qquad f = f_0 \text{ (a constant)}, \qquad P = P_0, \tag{7.67}$$

are given by

$$a_{ij} = a_0 \int_0^L \left(1 - \frac{x}{L}\right) \frac{d\phi_i}{dx}\frac{d\phi_j}{dx}\, dx + k\phi_i(L)\phi_j(L)$$
$$= a_0 ij \int_0^L \left(1 - \frac{x}{L}\right) x^{i+j-2}\, dx + k(L)^{i+j}$$
$$= a_0 \frac{ij(1+i+j)}{(i+j-1)(i+j)}(L)^{i+j-1} + k(L)^{i+j}, \tag{7.68a}$$

$$b_i = \int_0^L f\phi_i\, dx + P\phi_i(L) = \frac{f_0}{i+1}(L)^{i+1} + P_0(L)^i. \tag{7.68b}$$

For one-term approximation ($N = 1$ and $k = 0$), we have

$$a_{11} = \frac{3}{2}a_0 L, \qquad b_1 = \frac{1}{2}f_0 L^2 + P_0 L,$$

$$c_1 = \frac{b_1}{a_{11}} = \frac{6}{9a_0 L}\left(\frac{3}{6}f_0 L^2 + PL\right) = \frac{f_0 L + 2P_0}{3a_0},$$

and

$$U_1(x) = \frac{f_0 L + 2P_0}{3a_0} x.$$

For $N = 2$ and $k = 0$, we have

$$a_{11} = \frac{3}{2} a_0 L, \qquad a_{12} = a_{21} = \frac{4}{3} a_0 L^2, \qquad a_{22} = \frac{5}{3} a_0 L^3,$$
$$b_1 = \frac{1}{2} f_0 L^2 + P_0 L, \qquad b_2 - \frac{1}{3} f_0 L^3 + P_0 L^2.$$

The Ritz equations can be written in matrix form as

$$\frac{a_0 L}{6} \begin{bmatrix} 9 & 8L \\ 8L & 10L^2 \end{bmatrix} \begin{Bmatrix} c_1 \\ c_2 \end{Bmatrix} = \frac{f_0 L^2}{6} \begin{Bmatrix} 3 \\ 2L \end{Bmatrix} + P_0 L \begin{Bmatrix} 1 \\ L \end{Bmatrix},$$

whose solution by Cramer's rule is

$$c_1 = \frac{1}{a_0} \left( \frac{7}{13} f_0 + \frac{6}{13} P_0 \right), \qquad c_2 = \frac{3}{13 a_0 L} \left( -f_0 L + P_0 \right).$$

Hence the two-parameter Ritz solution is

$$U_2(x) = \frac{7 f_0 L + 6 P_0}{13 a_0} x + \frac{3(P_0 - f_0 L)}{13 a_0 L} x^2.$$

The exact solution of Eqs. (7.55a,b) with $u(0) = 0$, $k = 0$, $EA = a_0[2 - (x/L)]$, and $f = f_0$ is

$$u(x) = \frac{f_0 L}{a_0} x + \frac{(f_0 L - P_0) L}{a_0} \log\left(1 - \frac{x}{2L}\right) \tag{7.69a}$$
$$\approx \frac{f_0 L + P_0}{2 a_0} x + \frac{P_0 - f_0 L}{8 a_0 L} x^2 + \frac{P_0 - f_0 L}{24 a_0 L^2} x^3 + \cdots \tag{7.69b}$$

Table 7.1 contains a comparison of the Ritz coefficients $c_i$ for $N = 1, 2, \ldots, 8$ with the exact coefficients in Eq. (7.69b) for $L = 10$ ft, $a_0 = 180 \times 10^6$ lb, $f_0 = 0$, and $P_0 = 10$ kip. Clearly the Ritz coefficients $c_i$ converge to the exact ones as $N$ goes from 1 to 8.

### 7.3.5 General Features of the Method

Some general features of the Ritz method are listed below:

1. If the approximate functions $\phi_i(x)$ are selected to satisfy Eq. (7.62), the assumed approximation $U_N(x)$ normally converges to the actual solution $u(x)$ with an increase in the number of parameters (i.e., $N \to \infty$). A mathematical proof of such an assertion is not given here, but interested readers may consult the references at the end of the chapter.



**Table 7.1 The Ritz coefficients[a] for the axial deformation of an isotropic elastic bar subjected to axial force**

| $n$ | $\bar{c}_1$ | $\bar{c}_2$ | $\bar{c}_3$ | $\bar{c}_4$ | $\bar{c}_5$ | $\bar{c}_6$ | $\bar{c}_7$ | $\bar{c}_8$ |
|---|---|---|---|---|---|---|---|---|
| 1 | 37.037 | | | | | | | |
| 2 | 25.641 | 12.821 | | | | | | |
| 3 | 28.219 | 4.409 | 4.879 | | | | | |
| 4 | 27.691 | 7.788 | 0.000 | 3.029 | | | | |
| 5 | 27.794 | 6.701 | 3.389 | −1.040 | 1.664 | | | |
| 6 | 27.775 | 7.009 | 1.904 | 2.012 | −1.142 | 0.952 | | |
| 7 | 27.778 | 6.929 | 2.453 | 0.320 | 1.447 | −0.980 | 0.560 | |
| 8 | 27.778 | 6.948 | 2.272 | 1.094 | −0.287 | 1.136 | −0.769 | 0.336 |
| Exact | 27.778 | 6.944 | 2.315 | 0.868 | 0.347 | 0.145 | 0.062 | 0.027 |

[a] $\bar{c}_i = c_i \times 10^{5+i}$.

2. For increasing values of $N$, the previously computed coefficients of the algebraic equations (7.60b) remain unchanged (provided the previously selected coordinate functions are not changed), and one must add newly computed coefficients to the system of equations.
3. If the set of approximation functions $\{\phi_i\}$ chosen is an orthogonal set in the sense $B(\phi_i, \phi_j) = a_{(ij)}\delta_{ij}$ (no sum on $i$ and $j$), then one need not invert the system of equations, and the solution is obtained as $c_i = b_i / a_{ii}$.
4. The Ritz method applies to all problems, linear or nonlinear, as long as the variational problem

$$B(\delta u, u) = l(\delta u) \tag{7.70}$$

is equivalent to the governing equation and natural boundary conditions. In general, $B(\delta u, u)$ is linear in $\delta u$ but may be nonlinear in $u$ (and thus $B(\cdot, \cdot)$ may not be symmetric), and $l(\delta u)$ is a linear functional.
5. If the variational problem used in the Ritz approximation is such that its bilinear form is symmetric (in $u$ and $\delta u$), the resulting algebraic equations are also symmetric and, therefore, only elements above or below the main diagonal of the coefficient matrix need to be computed.
6. If $\delta\Pi$ (or $B(\delta u, u)$) is nonlinear in $u$, the resulting algebraic equations $[A(\{c\})]\{c\} = \{b\}$ will also be nonlinear in the parameters $c_i$. To solve such nonlinear equations, a variety of numerical methods are available (e.g., the Newton–Raphson method). Generally, there is more than one solution to the set of nonlinear equations.
7. Since the strains are computed from the approximate displacements, the strains and stresses are generally less accurate than the displacements.
8. The governing equation and natural boundary conditions of the problem are satisfied only in the variational sense, and not in the differential equation sense.

Therefore, the displacements obtained from the Ritz approximation generally do not satisfy the equations of equilibrium pointwise.

9. Since a continuous system is approximated by a finite number of coordinates (or degrees of freedom), the approximate system is less flexible than the actual system (recall that $\Pi(u) < \Pi(U)$ for any $U$ that is not exact). Consequently, the displacements obtained by the Ritz method using the principle of minimum total potential energy converge to the exact displacement from below:

$$U_1 < U_2 < \cdots < U_n < U_m \cdots < u, \qquad \text{for } m > n, \tag{7.71}$$

where $U_N$ denotes the $N$-parameter Ritz approximation of $u$ obtained from the principle of minimum total potential energy. The displacements obtained from the Ritz approximations based on the total complementary energy principle provide an upper bound for the exact solution.

10. Although the discussion of the Ritz method in this section thus far is confined to a linear solid mechanics problem, the method can be employed for any equation that admits a variational formulation (in the sense discussed in Comment 4), as will be illustrated through several examples shortly. However, the bounds mentioned above do not hold unless the variational problem is based on a minimum variational principle.

### 7.3.6 Examples

In this section we illustrate the application of the Ritz method to a variety of problems. These include static, eigenvalue, and transient problems. As will be shown in the sequel, the Ritz method can also be applied to problems that either do not admit a quadratic functional or where one knows only the governing equations of the problem. In the latter case, a way to develop the so-called weak form is discussed in Section 7.4.

**Example 7.5** Consider a uniform cross-section bar of length $L$, with the left end fixed and the right end connected to a rigid support via a linear elastic spring (with spring constant $k$), as shown in Fig. 7.4. We wish to determine the first two natural axial frequencies of the bar using the Ritz method.

The kinetic energy $K$ and the strain energy $U$ associated with the axial motion of the member are given by

$$K = \int_0^L \frac{\rho A}{2}\left(\frac{\partial u}{\partial t}\right)^2 dx, \qquad U = \int_0^L \frac{EA}{2}\left(\frac{\partial u}{\partial x}\right)^2 dx + \frac{k}{2}[u(L,t)]^2. \tag{7.72}$$

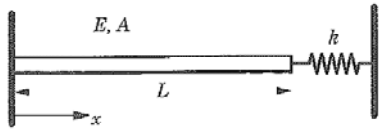


**Figure 7.4** Natural vibrations of a bar with an end spring.



Substituting for $K$ and $U$ from Eq. (7.72), and $V = 0$ in Hamilton's principle, we obtain [$\delta u(x, t_1) = \delta u(x, t_2) = 0$ and $\delta u(0, t) = 0$]:

$$\begin{aligned} 0 &= \int_{t_1}^{t_2} \delta(K - U)\, dt \\ &= \frac{1}{2}\int_{t_1}^{t_2} \delta\left\{\int_0^L \left[\rho A\left(\frac{\partial u}{\partial t}\right)^2 - EA\left(\frac{\partial u}{\partial x}\right)^2\right] dx - k[u(L,t)]^2\right\} dt \qquad (7.73) \\ &= \int_{t_1}^{t_2}\left[\int_0^L \left(\rho A \frac{\partial u}{\partial t}\frac{\partial \delta u}{\partial t} - EA\frac{\partial u}{\partial x}\frac{\partial \delta u}{\partial x}\right) dx - ku(L,t)\delta u(L,t)\right] dt \\ &= \int_{t_1}^{t_2}\left[\int_0^L \left(-\rho A \frac{\partial^2 u}{\partial t^2}\delta u - EA\frac{\partial u}{\partial x}\frac{\partial \delta u}{\partial x}\right) dx - ku(L,t)\delta u(L,t)\right] dt. \qquad (7.74) \end{aligned}$$

We seek the periodic motion of the form

$$u(x,t) = u_0(x)e^{i\omega t}, \qquad i = \sqrt{-1}, \tag{7.75}$$

where $\omega$ is the frequency of natural vibration, and $u_0(x)$ is the amplitude. Substituting Eq. (7.75) into Eq. (7.74), we obtain

$$0 = \int_0^L \left(\rho A\omega^2 u_0 \delta u_0 - EA\frac{du_0}{dx}\frac{d\delta u_0}{dx}\right) dx - ku_0(L)\delta u_0(L), \tag{7.76}$$

where $(i\omega)^2 = -\omega^2$, and $\int_{t_1}^{t_2} e^{2i\omega t} dt$, being nonzero, is factored out. We use Eq. (7.76) to determine the values of $\omega$. Note that the Rayleigh quotient for the problem at hand is given by

$$\omega^2 = \frac{\int_0^L EA(du_0/dx)(d\delta u_0/dx)\, dx + ku_0(L)\delta u_0(L)}{\int_0^L \rho A u_0 \delta u_0\, dx}.$$

The Euler equation and natural boundary condition associated with Eq. (7.76) are

$$-\frac{d}{dx}\left(EA\frac{du_0}{dx}\right) - \rho A\omega^2 u_0 = 0, \qquad 0 < x < L, \tag{7.77a}$$

$$EA\frac{du_0}{dx} + ku_0 = 0 \qquad \text{at } x = L. \tag{7.77b}$$

The essential boundary condition is $u_0(0) = 0$.

A nondimensionalization of the variables is used for simplicity:

$$\bar{x} = \frac{x}{L}, \qquad \bar{u} = \frac{u_0}{L}, \qquad \alpha = \frac{kL}{EA}, \qquad \lambda = \frac{\omega^2 \rho L^2}{E}. \tag{7.78}$$

Then Eq. (7.76) becomes

$$0 = \int_0^1 \left( \lambda \bar{u} \delta u - \frac{d\bar{u}}{dx} \frac{d\delta \bar{u}}{dx} \right) dx - \alpha \bar{u}(1) \delta u(1). \tag{7.79}$$

The bar over the nondimensional variables will be omitted in the interest of brevity. Further, in the following discussions, we shall assume that $\alpha = 1$.

Substituting an $N$-parameter Ritz approximation (obviously, we have $\phi_0 = 0$):

$$\bar{u}(x) \approx U_N(x) = \sum_{i=1}^{N} c_i \phi_i(x)$$

into Eq. (7.79), we obtain

$$0 = \sum_{i=1}^{N} \left\{ \sum_{j=1}^{N} \left[ \lambda \int_0^1 \phi_i \phi_j \, dx - \left( \int_0^1 \frac{d\phi_i}{dx} \frac{d\phi_j}{dx} \, dx + \alpha \phi_i(1) \phi_j(1) \right) \right] c_j \right\} \delta c_i.$$

Because of the indcpendent nature of $\delta c_i$, we obtain

$$0 = \sum_{j=1}^{N} \left[ \lambda \int_0^1 \phi_i \phi_j \, dx - \left( \int_0^1 \frac{d\phi_i}{dx} \frac{d\phi_j}{dx} \, dx + \alpha \phi_i(1) \phi_j(1) \right) \right] c_j, \tag{7.80a}$$

and in matrix form

$$([A] - \lambda [M]) \{c\} = \{0\}, \tag{7.80b}$$

where

$$a_{ij} = \int_0^1 \frac{d\phi_i}{dx} \frac{d\phi_j}{dx} \, dx + \alpha \phi_i(1) \phi_j(1), \qquad m_{ij} = \int_0^1 \phi_i \phi_j \, dx. \tag{7.80c}$$

Equation (7.80b) represents a matrix eigenvalue problem, and we obtain $N$ eigenvalues, $\lambda_i$, $i = 1, 2, \ldots, N$. An analytical method for finding eigenvalues and eigenvectors was discussed in Chapter 2 (see Section 2.3.4).

For the problem at hand, the approximation functions can be taken as

$$\phi_i(x) = x^i. \tag{7.81a}$$

Substituting $\phi_i = x^i$ into Eq. (7.80c), we obtain

$$m_{ij} = \int_0^1 \phi_i \phi_j \, dx = \frac{1}{i + j + 1},$$
$$a_{ij} = \int_0^1 \frac{d\phi_i}{dx} \frac{d\phi_j}{dx} \, dx + \phi_i(1) \phi_j(1) = \frac{ij}{i + j - 1} + 1. \tag{7.81b}$$



Since we wish to determine two eigenvalues, we take $N = 2$ and obtain

$$m_{11} = \tfrac{1}{3}, \qquad m_{12} = \tfrac{1}{4}, \qquad m_{22} = \tfrac{1}{5}, \qquad a_{11} = 2, \qquad a_{12} = 2, \qquad a_{22} = \tfrac{7}{3},$$

and the matrix eigenvalue problem (7.80b) becomes

$$\left( \begin{bmatrix} 2 & 2 \\ 2 & \frac{7}{3} \end{bmatrix} - \lambda \begin{bmatrix} \frac{1}{3} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{5} \end{bmatrix} \right) \begin{Bmatrix} c_1 \\ c_2 \end{Bmatrix} = \begin{Bmatrix} 0 \\ 0 \end{Bmatrix}. \tag{7.82}$$

For a nontrivial solution (i.e., $c_1 \neq 0, c_2 \neq 0$), the determinant of the coefficient matrix in Eq. (7.82) is set to zero:

$$\begin{vmatrix} 2 - \frac{\lambda}{3} & 2 - \frac{\lambda}{4} \\ 2 - \frac{\lambda}{4} & \frac{7}{3} - \frac{\lambda}{5} \end{vmatrix} = 0$$

or

$$15\lambda^2 - 640\lambda + 2400 = 0.$$

The quadratic equation has two roots:

$$\lambda_1 = 4.1545, \qquad \lambda_2 = 38.512 \rightarrow \omega_1 = \frac{2.038}{L} \sqrt{\frac{E}{\rho}}, \qquad \omega_2 = \frac{6.206}{L} \sqrt{\frac{E}{\rho}}. \tag{7.83}$$

The eigenvectors are given by

$$U_2^{(i)} = c_1^{(i)} x + c_2^{(i)} x^2,$$

where $c_1^{(i)}$ and $c_2^{(i)}$ are calculated from Eq. (7.82) for $\lambda = \lambda_i$, $i = 1, 2$ (see Example 2.8).

The exact values of $\lambda$ are the roots of the transcendental equation

$$\lambda + \tan \lambda = 0, \tag{7.84}$$

whose first two roots are ($\omega^2 = \lambda$):

$$\omega_{01} = \frac{2.02875}{L} \sqrt{\frac{E}{\rho}}, \qquad \omega_{02} = \frac{4.91318}{L} \sqrt{\frac{E}{\rho}}. \tag{7.85}$$

Note that the first approximate frequency is closer to the exact than the second.

If one selects $\phi_0$ and $\phi_i$ to satisfy the natural boundary condition also, the degree of polynomials will inevitably go up. For example, the lowest-order function that satisfies the homogeneous form (we still have $\phi_0 = 0$) of the natural boundary

condition $u'(1) + u(1) = 0$ is

$$\hat{\phi}_1 = 3x - 2x^2. \tag{7.86}$$

The one-parameter solution with the choice of $\hat{\phi}_1$ in Eq. (7.84) gives $\lambda_1 = 50/12 = 4.1667$, which is no better than the two-parameter solution computed using $\phi_1 = x$ and $\phi_2 = x^2$. Of course, solution $c_1\hat{\phi}_1$ would yield a more accurate value for $\lambda_1$ than the solution $c_1\phi_1$. Although $c_1\hat{\phi}_1$ and $c_1\phi_1 + c_2\phi_2$ are of the same degree (polynomials), the latter gives better accuracy for $\lambda_1$ because the number of parameters is greater, which provides greater freedom to adjust the parameters.

**Example 7.6** Consider a uniform cross-section bar of length $L$ with the left end fixed and the right end connected to a rigid support via a linear elastic spring with spring constant $k$. Suppose that the bar is subjected to a body force $f(x, t)$ (see Fig. 7.4). We wish to determine the transient response of the bar under the assumption that the motion starts from rest, i.e., the initial conditions of the problem are

$$u(x, 0) = 0, \qquad \dot{u}(x, 0) - 0. \tag{7.87}$$

The kinetic and strain energies associated with the axial motion of the bar are given in Eq. (7.70). The potential energy due to $f(x, t)$ is

$$V = -\int_{t_1}^{t_2}\int_0^L fu\,dx.$$

Then Eq. (7.73) becomes

$$0 = \int_{t_1}^{t_2}\left[\int_0^L \left(-\rho A\frac{\partial^2 u}{\partial t^2}\delta u - EA\frac{\partial u}{\partial x}\frac{\partial \delta u}{\partial x} + f\delta u\right)dx - ku(L, t)\delta u(L, t)\right]dt. \tag{7.88}$$

The Euler–Lagrange equations associated with Eq. (7.88) are

$$\frac{\partial}{\partial x}\left(EA\frac{\partial u}{\partial x}\right) - \frac{\partial}{\partial t}\left(\rho A\frac{\partial u}{\partial t}\right) + f = 0, \qquad 0 < x < L; \quad t > 0, \tag{7.89a}$$

$$EA\frac{\partial u}{\partial x} + ku = 0, \qquad \text{at } x = L; \quad t \geq 0. \tag{7.89b}$$

When one is interested in determining the time-dependent solution $u(x, t)$ under applied load $f(x, t)$, the Ritz solution is sought in the form

$$u(x, t) \approx \sum_{j=1}^{n} c_j(t)\phi_j(x), \qquad \phi_i = x^i, \tag{7.90}$$



where $c_j$ are now time-dependent parameters to be determined for all times $t > 0$. Substituting Eq. (7.90) into (7.88), we obtain

$$\begin{aligned} 0 &= -\sum_{j=1}^{n}\left[\left(\int_0^1 \phi_i\phi_j\,dx\right)\frac{d^2c_j}{dt^2} + \left(\int_0^1 \frac{d\phi_i}{dx}\frac{d\phi_j}{dx}\,dx + \phi_i(1)\phi_j(1)\right)c_j\right] \\ &\quad + \int_0^1 \phi_i\bar{f}\,dx \\ &= -\sum_{j=1}^{n}\left(m_{ij}\frac{d^2c_j}{dt^2} + a_{ij}c_j\right) + b_i, \end{aligned} \tag{7.91a}$$

where $x$, $u$, $f$, and $t$ are nondimensionalized as

$$\bar{x} = \frac{x}{L}, \qquad \bar{u} = \frac{u}{(f_0L^2/EA)}, \qquad \bar{f} = \frac{f}{f_0}, \qquad \bar{t} = \frac{t}{L\sqrt{\rho/E}}, \tag{7.91b}$$

$f_0$ being a constant, and

$$\begin{aligned} m_{ij} &= \int_0^1 \phi_i\phi_j\,dx, \\ a_{ij} &= \int_0^1 \frac{d\phi_i}{dx}\frac{d\phi_j}{dx}\,dx + \phi_i(1)\phi_j(1), \\ b_i &= \int_0^1 \phi_i\bar{f}(x, t)\,dx. \end{aligned} \tag{7.91c}$$

For $N = 1$ and $f = f_0$ (or $\bar{f} = 1$), we have

$$m_{11}\frac{d^2c_1}{dt^2} + a_{11}c_1 = b_1 \qquad \text{or} \qquad \frac{1}{3}\frac{d^2c_1}{dt^2} + 2c_1 = \frac{1}{2}.$$

The solution to the second-order differential equation is

$$c_1(t) = A\sin\sqrt{6}t + B\cos\sqrt{6}t + \tfrac{1}{4}.$$

Hence the one-parameter Ritz solution is given by

$$U_1(x, t) = \left(A\sin\sqrt{6}t + B\cos\sqrt{6}t + \tfrac{1}{4}\right)x,$$

where $A$ and $B$ are constants to be determined using the initial conditions. For zero initial conditions

$$u(x, 0) = 0, \qquad \dot{u}(x, 0) = 0, \tag{7.92}$$

we can determine the constants as $B = -1/4$ and $A = 0$. The solution becomes

$$U_1(x, t) = \tfrac{1}{4}\left(1 - \cos\sqrt{6}t\right)x.$$

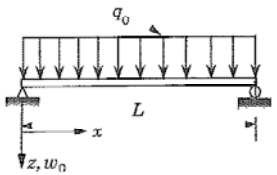


**Figure 7.5** A simply supported beam under uniform load.

For $N \geq 2$, the resulting system of differential equations in time, Eq. (7.91a), may be solved for $c_i(t)$ using the Laplace transform method or a numerical method, the latter being more practical. In general, the initial conditions (7.92) cannot be satisfied exactly, requiring approximation. Further, the numerical solution may be obtained only for discrete values of time. For example, at time $t_s = s\Delta t$ (i.e., the total time interval is divided into a finite number of time steps of size $\Delta t$), we would have

$$u_N(x, t_s) = \sum_{j=1}^{N} c_j(t_s)\phi_j(x). \tag{7.93}$$

For more details of the ideas discussed here, see Example 7.17.

**Example 7.7** Consider a simply supported beam of length $L$. We wish to find the transverse deflection of the beam under uniformly distributed transverse load $q_0$ (see Fig. 7.5) using the Euler–Bernoulli beam theory. The principle of virtual displacements for the problem becomes

$$0 = \int_0^L \left( EI \frac{d^2\delta w_0}{dx^2} \frac{d^2 w_0}{dx^2} - \delta w_0 q_0 \right) dx. \tag{7.94a}$$

The essential boundary conditions are

$$w_0(0) = w_0(L) = 0. \tag{7.94b}$$

We choose a two-parameter approximation of the form

$$w_0 \approx W_2(x) = c_1\phi_1 + c_2\phi_2 + \phi_0, \qquad \delta W_2 = \delta c_1\phi_1 + \delta c_2\phi_2, \tag{7.95a}$$

where

$$\phi_0 = 0, \qquad \phi_1 = x(L - x), \qquad \phi_2 = x^2(L - x). \tag{7.95b}$$

Substituting Eq. (7.95a) into Eq. (7.94a), we obtain

$$0 = \int_0^L \left[ EI\left(\delta c_1\phi_1'' + \delta c_2\phi_2''\right)\left(c_1\phi_1'' + c_2\phi_2''\right) - \left(\delta c_1\phi_1 + \delta c_2\phi_2\right) \right] q_0\, dx$$



$$= \sum_{i=1}^{2} \left( \sum_{j=1}^{2} a_{ij}c_j - b_i \right) \delta c_i \qquad \text{or} \qquad 0 = \sum_{j=1}^{2} a_{ij}c_j - b_i,$$

where

$$a_{ij} = \int_0^L EI \frac{d^2\phi_i}{dx^2} \frac{d^2\phi_j}{dx^2}\, dx, \qquad b_i = \int_0^L q_0\phi_i\, dx.$$

For the particular choice of $\phi_i$ in Eq. (7.95b), we have

$$EIL \begin{bmatrix} 4 & 2L \\ 2L & 4L^2 \end{bmatrix} \begin{Bmatrix} c_1 \\ c_2 \end{Bmatrix} = \frac{q_0 L^3}{12} \begin{Bmatrix} 2 \\ L \end{Bmatrix},$$

and the solution of these equations yields the result

$$c_1 = \frac{q_0 L^2}{24EI}, \qquad c_2 = 0,$$

so that the two-parameter Ritz solution becomes

$$W_2(x) = \frac{q_0 L^4}{24EI} \left( \frac{x}{L} - \frac{x^2}{L^2} \right). \tag{7.96}$$

The exact solution of the problem is given by

$$w_0(x) = \frac{q_0 L^4}{24EI} \left( \frac{x}{L} - 2\frac{x^3}{L^3} + \frac{x^4}{L^4} \right). \tag{7.97}$$

The maximum deflections according to the exact and Ritz solutions are

$$w_0\left(\frac{L}{2}\right) = \frac{5}{384}\frac{q_0 L^4}{EI} = 0.01302\frac{q_0 L^4}{EI}, \quad W_2\left(\frac{L}{2}\right) = \frac{37}{2352}\frac{q_0 L^4}{EI} = 0.01042\frac{q_0 L^4}{EI}.$$

Thus the two-parameter Ritz approximation is about 20% in error.

The three-parameter Ritz approximation with $\phi_3 = x^3(L - x)$ yields

$$EIL \begin{bmatrix} 4 & 2L & 2L^2 \\ 2L & 4L^2 & 4L^3 \\ 2L^2 & 4L^3 & 4.8L^4 \end{bmatrix} \begin{Bmatrix} c_1 \\ c_2 \\ c_3 \end{Bmatrix} = \frac{q_0 L^3}{12} \begin{Bmatrix} 2 \\ L \\ 0.6L^2 \end{Bmatrix}, \tag{7.98}$$

and we obtain

$$c_1 = c_2 L = -c_3 L^2 = \frac{q_0 L^2}{24EI}.$$

The three-parameter Ritz solution coincides with the exact solution in Eq. (7.97).

*Remark 1* If the beam is subjected to a point load $F_0$ at $x = L/2$, instead of a uniform load throughout the span of the beam (see Fig. 7.6), the exact solution will be in two parts:

$$w_0(x) = \begin{cases} \dfrac{F_0 L^3}{48EI}\left(3\dfrac{x}{L} - 4\dfrac{x^3}{L^3}\right), & 0 \le x \le \dfrac{L}{2}, \\ \dfrac{F_0 L^3}{48EI}\left[3\dfrac{x}{L} - 4\dfrac{x^3}{L^3} + \left(2\dfrac{x}{L} - 1\right)^3\right], & \dfrac{L}{2} \le x \le L. \end{cases} \tag{7.99}$$

Then it is clear that we must seek the Ritz solution also in two parts.

Suppose that we use the virtual work statement (or $\delta\Pi = 0$)

$$0 = \int_0^L EI \frac{d^2\delta w_0}{dx^2}\frac{d^2 w_0}{dx^2}\,dx - F_0\delta w_0\left(\frac{L}{2}\right) \tag{7.100}$$

with the three-parameter approximation

$$W_3(x) = c_1 x(L-x) + c_2 x^2(L-x) + c_3 x^3(L-x), \tag{7.101}$$

we obtain the same coefficient matrix as in Eq. (7.94), because the bilinear form did not change but the linear form changed, and the right-hand side is given by

$$\frac{F_0 L^2}{4}\begin{Bmatrix} 1 \\ \dfrac{L}{2} \\ \dfrac{L^2}{4} \end{Bmatrix}.$$

The solution for the Ritz coefficients gives

$$c_1 = \frac{F_0 L}{12EI}, \qquad c_2 = -\frac{F_0}{12EI}, \qquad c_3 = -\frac{5F_0}{64EIL}.$$

The Ritz solution *does not* coincide with the exact solution. In particular, the maximum deflection predicted by the three-parameter Ritz approximation (7.101) is

$$W_3\left(\frac{L}{2}\right) = \frac{7}{384}\frac{F_0 L^3}{EI} = \frac{F_0 L^3}{54.86EI},$$

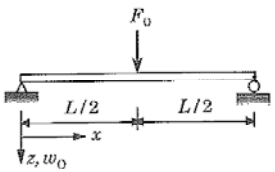


**Figure 7.6** A simply supported beam under center point load.



whereas the exact value from Eq. (7.99) is

$$w_0\left(\frac{L}{2}\right) = \frac{F_0 L^3}{48EI}. \tag{7.102}$$

The reason for the Ritz solution based on the variational problem (7.100) not being exact even for a point load is that (7.100) does not account for the discontinuity in the load. Note that the exact shear force, $Q = EI(d^3 w_0/dx^3)$, is discontinuous at $x = L/2$, but the approximate one is continuous. Thus, we must modify the variational problem, as discussed in Example 7.8.

*Remark 2* One can also use trigonometric polynomials in place of algebraic polynomials for the approximation functions $\phi_i$. For instance, the deflection of a simply supported beam subjected to continuous distributed load $q(x)$ can be represented by

$$w_0 \approx c_1 \sin\frac{\pi x}{L} + c_2 \sin\frac{3\pi x}{L} + \cdots + c_N \sin(2N-1)\frac{\pi x}{L}. \tag{7.103}$$

The functions $\phi_i = \sin(2i-1)(\pi x/L)$ are linearly independent, and are complete if all lower functions up to $\sin(2N-1)(\pi x/L)$ are included. When the load is sinusoidal,

$$q(x) = q_0 \sin\frac{m\pi x}{L} \quad (\text{for fixed } m), \tag{7.104}$$

we obtain the exact solution ($c_1 = c_2 = \cdots = c_{m-1} = c_{m+1} = \cdots = c_N = 0$):

$$w_0(x) = c_m \sin\frac{m\pi x}{L}, \qquad c_m = \frac{q_0 L^4}{EIm^4\pi^4}. \tag{7.105}$$

This solution cannot be represented using a finite set of algebraic functions $\{\phi_i\}$, although the Ritz solution with a finite number of terms may be very close to the exact when evaluated at a point.

On the other hand, if the load is representable by an algebraic polynomial (e.g., $q(x)$ is a constant, linear, or higher-order function of $x$), then the Ritz solution (7.103) will not coincide with the exact solution (7.97) for any finite value of $N$, because the sine-series representation of such a load is an infinite series. For example, when $q = q_0$, a constant, then

$$q(x) = q_0 - \sum_{i=1,3,\ldots}^{\infty} \frac{16q_0}{\pi i}\sin\frac{i\pi x}{L}. \tag{7.106}$$

However, the Ritz solution (7.103) converges rapidly, giving an accurate solution, especially away from the ends, for a finite value of $N$.

Thus, in general, a judicious choice of approximation functions $\phi_i$ based on the source term $q(x)$ will not only make the computational effort minimal, but also gives an accurate solution.

**Example 7.8** Here we consider a beam with discontinuous loading. As an example, consider the beam shown in Fig. 7.6. We divide the beam into as many parts as there are regions with continuous loading. Consider the $i$th part, located between $x = x_i$ and $x = x_{i+1}$. Within each part of the beam, the differential equation for $w_0(x)$ is

$$\frac{d^2}{dx^2}\left(EI\frac{d^2w_0}{dx^2}\right) = q(x), \qquad x_i < x < x_{i+1}, \tag{7.107}$$

where $q(x)$ is any distributed load in the part. We isolate the $i$th part and set up a free-body diagram depicting the internal forces $(Q_1^{(i)}, Q_2^{(i)})$ and moments $(M_1^{(i)}, M_2^{(i)})$, as shown in Fig. 7.7. The total potential energy for the $i$th part is

$$\Pi_i(w_0) = \int_{x_i}^{x_{i+1}} \left[\frac{EI}{2}\left(\frac{d^2w_0}{dx^2}\right)^2 - qw_0\right] dx - M_1^{(i)}\left(-\frac{dw_0}{dx}\right)_{x_i} - M_2^{(i)}\left(-\frac{dw_0}{dx}\right)_{x_{i+1}} - Q_1^{(i)}w_0(x_i) - Q_2^{(i)}w_0(x_{i+1}). \tag{7.108}$$

where the left end is labeled as 1 and the right end as 2.

To use the Ritz method, we must select approximation functions based on specified boundary conditions. If we assume for the moment that all of the boundary conditions are of natural type, then the Ritz approximation over the $i$th part may be assumed in the form (see Example 5.6):

$$w_0^i(x) = c_1 + c_2x + c_3x^2 + c_4x^3 = \varphi_1^{(i)}(x)w_1^{(i)} + \varphi_2^{(i)}(x)\theta_1^{(i)} + \varphi_3^{(i)}(x)w_2^{(i)} + \varphi_4^{(i)}(x)\theta_4^{(i)}, \tag{7.109}$$

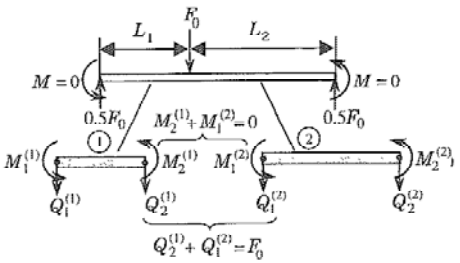


**Figure 7.7** Boundary, continuity, and equilibrium conditions for the beam of Example 7.8.



where ($\bar{x} = x - x_i$ and $L_i = x_{i+1} - x_i$)

$$\varphi_1^{(i)} = 1 - 3\left(\frac{\bar{x}}{L_i}\right)^2 + 2\left(\frac{\bar{x}}{L_i}\right)^3, \qquad \varphi_2^{(i)} = -\bar{x}\left(1 - \frac{\bar{x}}{L_i}\right)^2,$$
$$\varphi_3^{(i)} = 3\left(\frac{\bar{x}}{L_i}\right)^2 - 2\left(\frac{\bar{x}}{L_i}\right)^3, \qquad \varphi_4^{(i)} = -\bar{x}\left[\left(\frac{\bar{x}}{L_i}\right)^2 - \frac{\bar{x}}{L_i}\right]. \tag{7.110}$$

Substitution of Eq. (7.109) into $\delta\Pi_i = 0$ and carrying out the integration yields the Ritz equations for the $i$th part:

$$\frac{2E_iI_i}{L_i^3}\begin{bmatrix} 6 & -3L_i & -6 & -3L_i \\ -3L_i & 2L_i^2 & 3L_i & L_i^2 \\ -6 & 3L_i & 6 & 3L_i \\ -3L_i & L_i^2 & 3L_i & 2L_i^2 \end{bmatrix}\begin{Bmatrix} w_1^{(i)} \\ \theta_1^{(i)} \\ w_2^{(i)} \\ \theta_2^{(i)} \end{Bmatrix} = \begin{Bmatrix} q_1^{(i)} \\ q_2^{(i)} \\ q_3^{(i)} \\ q_4^{(i)} \end{Bmatrix} + \begin{Bmatrix} Q_1^{(i)} \\ M_1^{(i)} \\ Q_2^{(i)} \\ M_2^{(i)} \end{Bmatrix}, \tag{7.111a}$$

where

$$q_j^{(i)} = \int_{x_i}^{x_{i+1}} q^{(i)}(x)\varphi_j^{(i)}(x)\,dx. \tag{7.111b}$$

Here the superscript or subscript $i$ on variables indicates that they belong to the $i$th part. Note that the modulus of elasticity and moment of inertia are assumed to be constant within each part, while the load is arbitrary but continuous in each part. For uniformly distributed load $q^{(i)}(x) = q_0^{(i)}$ on part $i$, we have

$$\begin{Bmatrix} q_1^{(i)} \\ q_2^{(i)} \\ q_3^{(i)} \\ q_4^{(i)} \end{Bmatrix} = \frac{q_0^{(i)}L_i}{12}\begin{Bmatrix} 6 \\ L_i \\ 6 \\ L_i \end{Bmatrix}. \tag{7.111c}$$

Equations (7.111a) relate the four generalized displacements $(w_1^{(i)}, \theta_1^{(i)}, w_2^{(i)}, \theta_2^{(i)})$ to the four generalized forces $(Q_1^{(i)}, M_1^{(i)}, Q_2^{(i)}, M_2^{(i)})$. Obviously, four of the eight variables should be known in order to solve the four equations. If there are $n$ parts, there will be a total of $4n$ equations in $8n$ variables. Hence the remaining $4n$ variables should be eliminated through known conditions (e.g., boundary conditions, continuity conditions, and equilibrium conditions).

To illustrate the ideas, consider a simply supported beam with a point load $F_0$ at the center. Obviously, the beam needs to be divided into two parts, $0 \le x \le L/2$ and $L/2 \le x \le L$. The Ritz equations for the two parts are given below ($E_1 = E_2 = E$, $I_1 = I_2 = I$, $L_1 = L_2 = L/2$, and $q(x) = 0$).

Part 1:

$$\frac{16EI}{L^3}\begin{bmatrix} 6 & -1.5L & -6 & -1.5L \\ -1.5L & 0.5L^2 & 1.5L & 0.25L \\ -6 & 1.5L & 6 & 1.5L \\ -1.5L & 0.25L & 1.5L & 0.5L^2 \end{bmatrix}\begin{bmatrix} w_1^{(1)} \\ \theta_1^{(1)} \\ w_2^{(1)} \\ \theta_2^{(1)} \end{bmatrix} = \begin{bmatrix} Q_1^{(1)} \\ M_1^{(1)} \\ Q_2^{(1)} \\ M_2^{(1)} \end{bmatrix}. \tag{7.112a}$$

Part 2:

$$\frac{16EI}{L^3}\begin{bmatrix} 6 & 1.5L & 6 & -1.5L \\ -1.5L & 0.5L^2 & 1.5L & 0.25L \\ -6 & 1.5L & 6 & 1.5L \\ -1.5L & 0.25L & 1.5L & 0.5L^2 \end{bmatrix}\begin{Bmatrix} w_1^{(2)} \\ \theta_1^{(2)} \\ w_2^{(2)} \\ \theta_2^{(2)} \end{Bmatrix} = \begin{Bmatrix} Q_1^{(2)} \\ M_1^{(2)} \\ Q_2^{(2)} \\ M_2^{(2)} \end{Bmatrix}. \tag{7.112b}$$

There are a total of eight equations in 16 variables. However, some of the variables are duplicative and others are related. In particular, we have the following eight conditions (see Fig. 7.7):

*Boundary Conditions*

$$w_1^{(1)} = 0, \qquad w_2^{(2)} = 0, \qquad M_1^{(1)} = 0, \qquad M_2^{(2)} = 0. \tag{7.113a}$$

*Continuity Conditions*

$$w_2^{(1)} = w_1^{(2)}, \qquad \theta_2^{(1)} = \theta_1^{(2)}. \tag{7.113b}$$

*Equilibrium Conditions*

$$Q_2^{(1)} + Q_1^{(2)} = F_0, \qquad M_2^{(1)} + M_1^{(2)} = 0. \tag{7.113c}$$

Imposition of conditions in Eqs. (7.113a,b) is straightforward. However, to impose the equilibrium conditions (7.113c), we must add the third equation in (7.112a) to the first equation in (7.112b) and then, using the first condition of Eq. (7.113c) [conditions in Eqs. (7.113a,b) are also used], we obtain

$$\frac{16EI}{L^3}(1.5L\Theta_1 + 12W_2 - 1.5L\Theta_3) = F_0, \tag{7.114a}$$

where $w_2^{(1)} = W_2$, $\theta_1^{(1)} = \Theta_1$, $\theta_2^{(1)} = \Theta_2$, and $\theta_2^{(2)} = \Theta_3$. Similarly, adding the fourth equation in (7.112a) to the second equation in (7.112b) and then using the second condition in (7.113c), we obtain

$$\frac{16EI}{L^3}(0.25L\Theta_1 + L^2W_2 + 1.5L\Theta_3) = 0. \tag{7.114b}$$



Thus, we have a total of six equations: first two equations of Eq. (7.112a), the last two equations of Eq. (7.112b), and two equations in Eqs. (7.114a,b) in six unknowns $Q_1^{(1)}$, $\Theta_1$, $W_2$, $\Theta_2$, $\Theta_3$, and $Q_2^{(2)}$. Solving the middle four equations for ($\Theta_1$, $W_2$, $\Theta_2$, $\Theta_3$), we obtain

$$\Theta_1 = -\frac{F_0L^2}{16EI}, \qquad W_2 = \frac{F_0L^3}{48EI}, \qquad \Theta_2 = 0.0, \qquad \Theta_3 = \frac{F_0L^2}{16EI}, \tag{7.115}$$

and $Q_1^{(1)}$ and $Q_2^{(2)}$ can be computed to be $-0.5F_0$ from the first equation of (7.112a) and the last equation of (7.112b). The rotation $\Theta_2$ at the center of the beam is correctly predicted to be zero (why?). The four-parameter Ritz solution [but with the same polynomial degree as in the two-parameter solution (7.95a,b)] becomes

$$W_4(x) = \begin{cases} \dfrac{F_0L^3}{16EI}\left(\dfrac{x}{L} - 4\dfrac{x^2}{L^2} + 4\dfrac{x^3}{L^3}\right) + \dfrac{F_0L^3}{48EI}\left(12\dfrac{x^2}{L^2} - 16\dfrac{x^3}{L^3}\right), & 0 \le x \le \dfrac{L}{2}, \\ \dfrac{F_0L^3}{48EI}\left[1 - 12\left(\dfrac{\bar{x}}{L}\right)^2 + 16\left(\dfrac{\bar{x}}{L}\right)^3\right] + \dfrac{F_0L^3}{16EI}\left[2\left(\dfrac{\bar{x}}{L}\right)^2 - 4\left(\dfrac{\bar{x}}{L}\right)^3\right], & \dfrac{L}{2} \le x \le L, \end{cases} \tag{7.116}$$

where $\bar{x} = x - L/2$. Simplification of Eq. (7.116) gives the expression in Eq. (7.99); thus, the Ritz solution matches with the exact solution.

The procedure described in this example closely resembles that of the finite element method, which will be discussed in detail in Chapter 9. The procedure is valid for *all* beam problems irrespective of the nature of the loading and boundary conditions. In addition, the procedure gives exact values of the deflection $w_0(x)$ and rotation $-(dw_0/dx)$ at the end points of each part for all loads and boundary conditions, provided that the flexural rigidity $EI$ is constant within each part.

## 7.4 GENERAL BOUNDARY-VALUE PROBLEMS

All the examples presented in Section 7.3.3 utilized the minimum total potential energy principle or Hamilton's principle. For problems outside the field of solid and structural mechanics, the construction of an analog of the minimum total potential energy principle is needed to derive the Ritz equations. Here the procedure for constructing *weak forms* from differential equations and use of these statements in the Ritz approximation are studied.

### 7.4.1 Variational Formulations

The steps involved in the weak formulation of differential equations are described with the aid of three model equations: (1) a second-order equation in one dimension, (2) a fourth-order equation in one dimension, and (3) a second-order equation in two dimensions. These equations are quite general and they arise in a number of fields in engineering and applied sciences. Most of the ideas in developing a weak form are

presented in connection with Model Equation 1. Model Equation 2 is used to illustrate the weak form development for a higher-order equation, and Model Equation 3 is used to extend the ideas to two dimensions.

***Model Equation 1*** Consider the problem of finding the function $u(x)$ that satisfies the differential equation

$$-\frac{d}{dx}\left(a(x)\frac{du}{dx}\right)+c(x)u-f=0 \qquad \text{for } 0<x<L, \tag{7.117a}$$

and the boundary conditions

$$u(0)=u_0, \qquad \left(a\frac{du}{dx}\right)\Big|_{x=L}=P, \tag{7.117b}$$

where $a=a(x)$, $c=c(x)$, $f=f(x)$, $u_0$, and $P$ are the data (i.e., known quantities) of the problem. Equation (7.117a) arises in connection with the analytical description of many physical processes. For example, conduction and convection heat transfer in a plane wall or fin (1D heat transfer), flow through channels and pipes, transverse deflection of cables, axial deformation of bars, and many others. Table 7.2 contains a list of several field problems described by (7.117a) when $c(x)=0$. The mathematical structure common to different fields is brought out in this table. Thus, if we can develop a numerical procedure by which Eq. (7.117a) can be solved for all physically possible boundary conditions, the procedure can be used to solve all field problems listed in Table 7.2, as well as many others. This fact provides us with the motivation to use Eq. (7.2) as a model second-order equation in one dimension. The motivation for the development of the weak form and a step-by-step procedure for the weak form development are discussed next.

In the Ritz method, we seek an approximate solution to Eq. (7.117a) in the form

$$U_N(x)=\sum_{j=1}^{N}c_j\phi_j(x)+\phi_0(x), \tag{7.118}$$

and determine the unknown parameters $c_j$ such that Eqs. (7.117a) and (7.117b) are satisfied by the $N$-parameter approximate solution $U_N$. For example, suppose that $L=1$, $a=x$, $c=1$, $u_0=1$, $f=0$, and $P=0$; then we could take $N=2$ and write the approximate solution of (7.117a) in the form ($\phi_1=x^2-2x$, $\phi_2=x^3-3x$, $\phi_0=1$):

$$u\approx U_N=c_1(x^2-2x)+c_2(x^3-3x)+1,$$

which satisfies the boundary conditions (7.117b) of the problem for any values of $c_1$ and $c_2$. Then the constants $c_1$ and $c_2$ must be determined such that the approximate solution $U_N$ satisfies Eq. (7.117a):

$$-\frac{d}{dx}\left(x\frac{du}{dx}\right)+u=0 \tag{a}$$



**Table 7.2 Some examples of second-order equation (7.117a) in one dimension:**

$$-\frac{d}{dx}\left(a\frac{du}{dx}\right)=f \qquad \text{for } 0<x<L$$

$$\text{EBC: } u(0)=u_0; \qquad \text{NBC: } \left(a\frac{du}{dx}\right)_{x=L}=P$$

| Field | Primary Variable $u$ | Coefficient[a] $a$ | Source Term $f$ | Secondary Variable $P$ |
|---|---|---|---|---|
| 1. Cables | Transverse deflection | $T$ | Distributed vertical force | Axial load |
| 2. Bars | Longitudinal displacement | $EA$ | Distributed axial force | Axial load |
| 3. Heat transfer | Temperature | $k$ | Internal heat generation | Heat flux Heat |
| 4. Pipe flows | Hydrostatic pressure | $\frac{\pi D^4}{128\mu}$ | Flow source | Flow rate |
| 5. Viscous flows | Velocity | $\mu$ | Pressure gradient | Stress |
| 6. Seepage | Fluid head | $\varepsilon$ | Fluid flux | Flow |
| 7. Electrostatics | Electrical potential | $\epsilon$ | Charge density | Electric flux |

[a] $E$ = Young's modulus; $A$ = area of cross section; $D$ = diameter of the pipe; $k$ = thermal conductivity; $\mu$ = viscosity; $T$ = tension; $\varepsilon$ = permeability; $\epsilon$ = dielectric constant.

in some sense. If we require $U_N$ to satisfy the above equation in the exact sense, we obtain

$$-\frac{dU_N}{dx}-x\frac{d^2U_N}{dx^2}+U_N=-2c_1(x-1)-3C_2(x^2-1)-2c_1x-6c_2x^2$$
$$+c_1(x^2-2x)+c_2(x^3-3x)+1=0.$$

Since this expression must be zero at all $x$, the coefficients of the various powers of $x$ must be zero:

$$\begin{aligned}1+2c_1+3c_2&=0,\\ -(6c_1+3c_2)&=0,\\ c_1-9c_2&=0,\\ c_2&=0.\end{aligned}$$

The above relations are inconsistent; hence there is *no solution* to the equations. If we were able to find a unique solution to these equations, then $U_N=c_1(x^2-2x)+c_2(x^3-3x)+1$ is the exact solution of the problem. It is not always possible for

arbitrary data of the problem to find the exact solution. An alternative is that we may require the approximate solution $U_N$ to satisfy the differential equation (a) in a weighted-integral sense,

$$\int_0^1 w(x)R\,dx = 0, \tag{b}$$

where $R$ is the *residual* (i.e., error) in the differential equation,

$$R \equiv -\frac{dU_N}{dx} - x\frac{d^2U_N}{dx^2} + U_N \neq 0,$$

and $w(x)$ is a function, called the *weight function*, which is introduced to provide as many independent relations among $c_j$ as there are unknown parameters $c_j$ ($j = 1, 2, \ldots, N$). This is accomplished by selecting $N$ independent functions for $w(x)$. For example, if we take the two choices $w(x) = 1$ and $w(x) = x$, which are linearly independent, we obtain

$$0 = \int_0^1 1 \cdot R\,dx = (1 + 2c_1 + 3c_2) + \tfrac{1}{2}(-6c_1 - 3c_2) + \tfrac{1}{3}(c_1 - 9c_2) + \tfrac{1}{4}c_2,$$

$$0 = \int_0^1 x \cdot R\,dx = \tfrac{1}{2}(1 + 2c_1 + 3c_2) + \tfrac{1}{3}(-6c_1 - 3c_2) + \tfrac{1}{4}(c_1 - 9c_2) + \tfrac{1}{5}c_2,$$

or

$$\tfrac{2}{3}c_1 + \tfrac{5}{4}c_2 = 1, \qquad \tfrac{3}{4}c_1 + \tfrac{31}{20}c_2 = \tfrac{1}{2}. \tag{c}$$

These equations provide two linearly independent relations for $c_1$ and $c_2$ that can be solved to obtain unique solution, $c_1 = 222/23$ and $c_2 = -100/23$, and the approximate solution $U_2(x)$ becomes

$$U_2(x) = c_1(x^2 - 2x) + c_2(x^3 - 3x) + 1 = \tfrac{222}{23}(x^2 - 2x) - \tfrac{100}{23}(x^3 - 3x) + 1. \tag{d}$$

Thus, integral statements of the type in (b) provide means for obtaining as many algebraic equations as there are unknown coefficients in the approximation. There are several variational methods, in addition to the Ritz method, in which approximate solutions of the type $u \approx \sum c_j\phi_j + \phi_0$ are sought, and the coefficients $c_j$ are determined, as shown above, using an integral statement. These methods differ from each other in the choice of the weight function $w(x)$ and the integral statement used, which in turn dictates the choice of the approximation functions $\phi_j$ and $\phi_0$. For the moment we deal with the Ritz method of approximation.

As discussed above, the necessary and sufficient number of algebraic relations among the $c_j$'s can be obtained by recasting the differential equation (7.117) in a weighted-integral form:

$$0 = \int_0^L w(x)\left[-\frac{d}{dx}\left(a\frac{du}{dx}\right) + cu - f\right]dx, \tag{7.119}$$



where $w(x)$ denotes the weight function, which for the moment is arbitrary. For $u(x) \approx U_N(x)$ and each independent choice of $w(x)$, we obtain an independent algebraic equation relating all $c_j$. A total of $N$ independent equations are required to solve for the $N$ parameters $c_j$. When a weighted-integral statement like Eq. (7.119) is used to obtain the $N$ equations among $c_j$, the method is known as the *weighted-residual method*, and it will be discussed in Section 7.5. Note that the use of Eq. (7.119) precludes that $U_N(x)$ satisfies *all* specified boundary conditions and is differentiable as many times as required in the differential equation.

To weaken the continuity (i.e., differentiability) required of $U_N(x)$ and therefore of $\phi_j(x)$, we trade the differentiation in (7.119) from $u$ to $w$ such that both $u$ and $w$ are differentiated equally—once each in the present case. The resulting integral form is termed the *weak form* of (7.117). This form is not only equivalent to (7.117) but it also contains the natural boundary conditions of the problem, and therefore $U_N(x)$ need not satisfy the natural boundary conditions. The three-step procedure of constructing the weak form of (7.117) is discussed next.

The first step is to multiply the governing differential equation with a weight function $w$ and integrate over domain $(0, L)$, giving Eq. (7.119). The second step is to trade differentiation from $u$ to $w$, using integration by parts. This is achieved as follows. Consider the identity

$$-w\left[\frac{d}{dx}\left(a\frac{du}{dx}\right)\right] = -\frac{d}{dx}\left(wa\frac{du}{dx}\right) + a\frac{dw}{dx}\frac{du}{dx}, \tag{7.120a}$$

which is simply the product rule of differentiation applied to the product of two functions, $a(du/dx)$ and $w$. Integrating this identity over the domain, we obtain

$$-\int_0^L w\left[\frac{d}{dx}\left(a\frac{du}{dx}\right)\right]dx = -\int_0^L \frac{d}{dx}\left(wa\frac{du}{dx}\right)dx + \int_0^L a\frac{dw}{dx}\frac{du}{dx}\,dx$$

$$= -\left[wa\frac{du}{dx}\right]_0^L + \int_0^L a\frac{dw}{dx}\frac{du}{dx}\,dx. \tag{7.120b}$$

Substituting (7.120b) into (7.119), we arrive at the result

$$0 = \int_0^L \left(a\frac{dw}{dx}\frac{du}{dx} + cwu - wf\right)dx - \left[w \cdot a\frac{du}{dx}\right]_0^L. \tag{7.121}$$

The third and last step is to identify the primary and secondary variables of the variational (or weak) form. This requires us to classify the boundary conditions of each differential equation into *essential* (or geometric) and *natural* (or force) boundary conditions. The classification is made uniquely by examining the boundary term appearing in the second step of the weak form development, namely, Eq. (7.121):

$$\left[w \cdot a\frac{du}{dx}\right]_0^L.$$

As a rule, the coefficient of the weight function in the boundary expression is called the *secondary variable*, and its specification constitutes the *natural* or *Neumann* boundary condition. The dependent unknown *in the same form as the weight function* in the boundary expression is termed the *primary variable*, and its specification constitutes the *essential* or *Dirichlet* boundary condition. For the model equation at hand, the primary and secondary variables are

$$u \qquad \text{and} \qquad a\frac{du}{dx} \equiv Q.$$

Next, we denote the secondary variables at the end points by some symbols:

$$-Q_0 = \left(a\frac{du}{dx}\right)\Big|_0, \qquad Q_L = \left(a\frac{du}{dx}\right)\Big|_L = P. \tag{7.122a}$$

With the notation in (7.122a), the variational form becomes

$$0 = \int_0^L \left(a\frac{dw}{dx}\frac{du}{dx} + cwu - wf\right)dx - w(0)Q_0 - w(L)Q_L. \tag{7.122b}$$

Now is the time to discuss the conditions on the weight function $w(x)$. Clearly, it should be differentiable at least once (like $u(x)$ is). The Ritz method uses the weak form, with $w = \phi_i$ to obtain the $i$th equation of the set of $N$ relations among $c_j$'s. Thus, we require $w(x)$ to satisfy the homogeneous form of specified essential boundary conditions (i.e., to belong to the set of admissible variations). In the present case, $w(x)$ should be once differentiable and vanish at $x = 0$. Hence, the final weak form is

$$0 = \int_0^L \left(a\frac{dw}{dx}\frac{du}{dx} + cwu - wf\right)dx - w(L)P. \tag{7.123}$$

This completes the three-step procedure of constructing the weak form. The weak form in (7.123) contains two types of expressions: those containing both $w$ and $u$; and those containing only $w$. We group the former type into a single expression, called the *bilinear form*:

$$B(w,u) \equiv \int_0^L \left(a\frac{dw}{dx}\frac{du}{dx} + cwu\right)dx. \tag{7.124a}$$

We denote all terms containing only $w$ (but not $u$) by $l(w)$, called the *linear form*:

$$l(w) = \int_0^L wf\,dx + w(L)P. \tag{7.124b}$$

The statement (7.123) can now be expressed as one of finding $u$ from the set of admissible functions (i.e., differentiable at least once and satisfying the essential boundary conditions) such that the variational problem

$$B(w,u) = l(w) \tag{7.125}$$



is satisfied for all $w$ from the set of admissible variations. As seen before, the bilinear form results directly in the coefficient matrix, while the linear form gives rise to the right-hand-side column vector of the Ritz equations. Since $B(\cdot,\cdot)$ is symmetric, the functional associated with the variational problem (7.125) is given by Eq. (7.53) (which represents the total potential energy in the case of bars or cables):

$$\begin{aligned} I(u) &= \frac{1}{2}B(u,u) - l(u) \\ &= \int_0^L \left[\frac{a}{2}\left(\frac{du}{dx}\right)^2 + \frac{c}{2}u^2\right]dx - \int_0^L wf\,dx - w(L)P. \end{aligned} \tag{7.126}$$

The weak form development up to Eq. (7.125) is valid even for the case in which the coefficients $a$ and $c$ are functions of the dependent variable $u$, making the problem nonlinear. In that case, $B(w,u)$ will be no longer a bilinear form and the functional $I(u)$ may not exist. However, to use the Ritz method one needs only the variational statement (7.125), which always exists.

***Model Equation 2*** Here we consider a fourth-order differential equation of the form

$$\frac{d^2}{dx^2}\left(b(x)\frac{d^2u}{dx^2}\right) + c(x)u = f, \qquad 0 < x < L, \tag{7.127}$$

We will not be concerned with any specific boundary conditions, as the weak form development naturally leads to the classification of the variables into primary and secondary type, and their specification constitutes, respectively, the essential and the natural type. This equation arises, for example, in connection with the bending of straight beams, where $u(x)$ denotes the transverse deflection, $b(x) = E(x)I(x)$ the bending rigidity, $c(x) = k$ the foundation modulus (if any), and $f(x)$ the distributed transverse load.

The first two steps in the development of the weak form of the equation are summarized below. Note that in this case we must transfer two derivatives to the weight function so that both $w$ and $u$ are required to have the same order of continuity (or differentiability).

*Step 1:*

$$0 = \int_0^L w(x)\left[\frac{d^2}{dx^2}\left(b\frac{d^2u}{dx^2}\right) + cu - f\right]dx. \tag{7.128}$$

*Step 2:*

$$0 = \int_0^L \left[-\frac{dw}{dx}\frac{d}{dx}\left(b\frac{d^2u}{dx^2}\right) + cwu - wf\right]dx + \left[w\cdot\frac{d}{dx}\left(b\frac{d^2u}{dx^2}\right)\right]_0^L$$

$$= \int_0^L \left( b\frac{d^2w}{dx^2}\frac{d^2u}{dx^2} + cwu - wf \right) dx + \left[ w \cdot \frac{d}{dx}\left( b\frac{d^2u}{dx^2} \right) \right]_0^L - \left[ \frac{dw}{dx} \cdot b\frac{d^2u}{dx^2} \right]_0^L. \tag{7.129}$$

It is clear from the boundary expressions that the secondary variables are

$$\frac{d}{dx}\left( b\frac{d^2u}{dx^2} \right), \qquad b\frac{d^2u}{dx^2}. \tag{7.130a}$$

In the case of beams they represent the shear force and bending moment in the beam. The primary variables are [$w \to u$ and $(dw/dx) \to (du/dx)$]

$$u, \qquad \frac{du}{dx}. \tag{7.130b}$$

*Step 3:* Denoting the shear forces and bending moments at the two ends of the beam as (proper signs are inserted to make all of the $Q$'s and $M$'s have the negative sign in the weak form; this also happens to be the correct definition of the bending moments and shear forces on the left and right ends of the beam):

$$\begin{aligned} \left[ \frac{d}{dx}\left( b\frac{d^2u}{dx^2} \right) \right]_{x=0} &= Q_0, \qquad \left[ b\frac{d^2u}{dx^2} \right]_{x=0} = M_0, \\ \left[ -\frac{d}{dx}\left( b\frac{d^2u}{dx^2} \right) \right]_{x=L} &= Q_L, \qquad \left[ -b\frac{d^2u}{dx^2} \right]_{x=0} = M_L. \end{aligned} \tag{7.131}$$

Then Eq. (7.129) takes the final form

$$\begin{aligned} 0 = \int_0^L \left( b\frac{d^2w}{dx^2}\frac{d^2u}{dx^2} + cwu - wf \right) dx - w(0)Q_0 - w(L)Q_L \\ - \left( -\frac{dw}{dx} \right)\Big|_0 M_0 - \left( -\frac{dw}{dx} \right)\Big|_L M_L. \end{aligned} \tag{7.132}$$

The bilinear form, linear form, and functional for the problem are

$$\begin{aligned} B(w,u) &= \int_0^L \left( b\frac{d^2w}{dx^2}\frac{d^2u}{dx^2} + cwu \right) dx, \\ l(w) &= \int_0^L wf\,dx + w(0)Q_0 + w(L)Q_L + \left( -\frac{dw}{dx} \right)\Big|_0 M_0 + \left( -\frac{dw}{dx} \right)\Big|_L M_L, \\ I(u) &= \int_0^L \left[ \frac{b}{2}\left( \frac{d^2u}{dx^2} \right)^2 + \frac{c}{2}u^2 - fu \right] dx \\ &\quad - u(0)Q_0 - u(L)Q_L - \left( -\frac{du}{dx} \right)\Big|_0 M_0 - \left( -\frac{du}{dx} \right)\Big|_L M_L. \end{aligned} \tag{7.133}$$

In the case of beam bending, $I(u)$ is nothing but the total potential energy $\Pi(u)$.



The weight function $w(x)$ is required to be twice differentiable and to vanish at the points where $u$ and to $du/dx$ are specified. Equation (7.132) can be specialized to any beam with specific boundary conditions and loading.

***Model Equation 3*** Lastly, we consider the problem of determining the solution $u(x, y)$ to the partial differential equation

$$-\frac{\partial}{\partial x}\left( a_1\frac{\partial u}{\partial x} \right) - \frac{\partial}{\partial y}\left( a_2\frac{\partial u}{\partial y} \right) + a_0u = f \qquad \text{in } R \tag{7.134}$$

in a two-dimensional domain $R$. Here $a_0$, $a_1$, $a_2$, and $f$ are known functions of position $(x, y)$ in $R$. The function $u$ is required to satisfy, in addition to the differential equation (7.134), certain boundary conditions on the boundary $S$ of $R$. The variational formulation to be presented tells us the precise form of the essential and natural boundary conditions of the equation. Equation (7.134) arises in many fields of engineering, including in 2D heat transfer, stream function or velocity potential formulation of inviscid flows, transverse deflections of a membrane, and torsion of a cylindrical member.

The three-step procedure applied to Eq. (7.134) results in the following equations:

*Step 1:*

$$0 = \int_R w\left[ -\frac{\partial}{\partial x}\left( a_1\frac{\partial u}{\partial x} \right) - \frac{\partial}{\partial y}\left( a_2\frac{\partial u}{\partial y} \right) + a_0u - f \right] dxdy. \tag{7.135}$$

*Step 2:*

$$\begin{aligned} 0 = \int_R \left( a_1\frac{\partial w}{\partial x}\frac{\partial u}{\partial x} + a_2\frac{\partial w}{\partial y}\frac{\partial u}{\partial y} + a_0wu - wf \right) dxdy \\ - \oint_S w\left( a_1\frac{\partial u}{\partial x}n_x + a_2\frac{\partial u}{\partial y}n_y \right) ds, \end{aligned} \tag{7.136}$$

where we used integration by parts [or the Green–Gauss theorem, Eq. (2.89)] to transfer differentiation to $w$ so that both $u$ and $w$ have the same order derivatives. The boundary term shows that $u$ is the primary variable while

$$a_1\frac{\partial u}{\partial x}n_x + a_2\frac{\partial u}{\partial y}n_y$$

is the secondary variable.

*Step 3:* The last step in the procedure is to impose the specified boundary conditions. Suppose that $u$ is specified on portion $S_1$ and the natural boundary condition is

specified on the remaining portion $S_2$ of the boundary:

$$u = \hat{u} \quad \text{on } S_1, \qquad a_1 \frac{\partial u}{\partial x} n_x + a_2 \frac{\partial u}{\partial y} n_y = \hat{g} \quad \text{on } S_2. \tag{7.137}$$

Then $w$ is arbitrary on $S_2$ and equal to zero on $S_1$. Consequently, Eq. (7.136) simplifies to

$$0 = \int_R \left( a_1 \frac{\partial w}{\partial x}\frac{\partial u}{\partial x} + a_2 \frac{\partial w}{\partial y}\frac{\partial u}{\partial y} + a_0 wu - wf \right) dxdy - \int_{S_2} w\hat{g}\, ds. \tag{7.138}$$

The bilinear form, linear form, and functionals are

$$B(w,u) = \int_R \left( a_1 \frac{\partial w}{\partial x}\frac{\partial u}{\partial x} + a_2 \frac{\partial w}{\partial y}\frac{\partial u}{\partial y} + a_0 wu \right) dxdy, \tag{7.139a}$$

$$l(w) = \int_R wf\, dxdy + \int_{S_2} w\hat{g}\, ds, \tag{7.139b}$$

$$I(u) = \frac{1}{2}\int_R \left[ a_1 \left(\frac{\partial u}{\partial x}\right)^2 + a_2 \left(\frac{\partial u}{\partial y}\right)^2 + a_0 u^2 \right] dxdy - \int_R uf\, dxdy - \int_{S_2} \hat{g}u\, ds. \tag{7.139c}$$

Once again, $I(u)$ represents the total potential energy in the case of a membrane problem.

### 7.4.2 Ritz Approximations

The Ritz method can be applied directly to the weak forms (7.124), (7.132), and (7.138)—or into the variational problem

$$B(w,u) = l(w). \tag{7.140}$$

Here we consider the case in which $B(\cdot,\cdot)$ is a bilinear form. Substituting

$$u \approx U_N = \sum_{j=1}^{N} c_j \phi_j + \phi_0, \qquad w = \phi_i, \tag{7.141}$$

into Eq. (7.140) to obtain the $i$th equation

$$\sum_{j=1}^{N} a_{ij} c_j = b_i, \qquad j = 1, 2, \ldots, N, \tag{7.142a}$$

where

$$a_{ij} = B(\phi_i, \phi_j), \qquad b_i = l(\phi_i) - B(\phi_0, \phi_i). \tag{7.142b}$$



The specific expressions of $A_{ij}$ and $b_i$ of each model equation can be written using the respective bilinear and linear forms. For example, for the third model equation, we have

$$a_{ij} = \int_R \left( a_1 \frac{\partial \phi_i}{\partial x}\frac{\partial \phi_j}{\partial x} + a_2 \frac{\partial \phi_i}{\partial y}\frac{\partial \phi_j}{\partial y} + a_0 \phi_i \phi_j \right) dxdy,$$

$$b_i = \int_R \left( \phi_i f - a_1 \frac{\partial \phi_i}{\partial x}\frac{\partial \phi_0}{\partial x} - a_2 \frac{\partial \phi_i}{\partial y}\frac{\partial \phi_0}{\partial y} - a_0 \phi_i \phi_0 \right) dxdy + \int_{S_2} \phi_i \hat{g}\, dS. \tag{7.143}$$

Once $\phi_i$ and $\phi_0$ are selected, subject to the conditions stated in Section 7.2, the coefficients of matrix $[A]$ and column vector $\{b\}$ can be computed, and the linear algebraic equations in Eq. (7.142a) can be solved for the Ritz coefficients. We consider specific examples next.

**Example 7.9** We wish to solve the partial differential equation (Laplace's equation)

$$-\left( \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \right) = 0 \qquad \text{in } 0 < (x, y) < 1, \tag{7.144a}$$

subject to the boundary conditions (see Fig. 7.8):

$$\begin{aligned} u &= 0 \qquad \text{on } x = 0, 1 \text{ and } y = 0, \\ u &= \sin \pi x \qquad \text{on } y = 1. \end{aligned} \tag{7.144b}$$

The weak form of the equation can be obtained as a special case from Eq. (7.138) by setting $f = 0$, $a_1 = a_2 = 1$, $a_0 = 0$, and $S_2 = 0$ ($S_1 = S$):

$$0 = \int_0^1 \int_0^1 \left( \frac{\partial \delta u}{\partial x}\frac{\partial u}{\partial x} + \frac{\partial \delta u}{\partial y}\frac{\partial u}{\partial y} \right) dxdy. \tag{7.145}$$

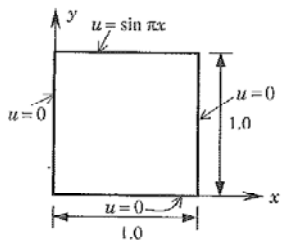


**Figure 7.8** Domain and boundary conditions of the problem in Example 7.9.

The Ritz equations are given by Eq. (7.142a) with $a_{ij}$ and $b_i$ given by

$$a_{ij} = \int_0^1 \int_0^1 \left( \frac{\partial \phi_i}{\partial x}\frac{\partial \phi_j}{\partial x} + \frac{\partial \phi_i}{\partial y}\frac{\partial \phi_j}{\partial y} \right) dx dy,$$
$$b_i = -\int_0^1 \int_0^1 \left( \frac{\partial \phi_i}{\partial x}\frac{\partial \phi_0}{\partial x} + \frac{\partial \phi_i}{\partial y}\frac{\partial \phi_0}{\partial y} \right) dx dy. \tag{7.146}$$

Next we select $\phi_0$ and $\phi_i$ for the problem. The function $\phi_0$ is required to satisfy the boundary conditions in Eq. (7.144b) because all of them are of the essential type. The following choice for $\phi_0$ meets the conditions:

$$\phi_0 = y \sin \pi x. \tag{7.147a}$$

The functions $\phi_i (i = 1, 2, \ldots, n)$ are required to satisfy the homogenous form of Eq. (7.144b) and to be linearly independent. We choose

$$\phi_1 = \sin \pi x \sin \pi y, \qquad \phi_2 = \sin \pi x \sin 2\pi y, \qquad \phi_3 = \sin 2\pi x \sin \pi y, \text{ etc.} \tag{7.147b}$$

From Eq. (7.146) we obtain

$$a_{ij} = \begin{cases} \frac{\pi^2}{2}, & \text{if } i = j, \\ 0, & \text{if } i \neq j \end{cases}, \quad \text{(i.e., [A] is diagonal)},$$
$$b_i = \begin{cases} -\frac{\pi}{2}, & \text{if } i = 1, \\ 0, & \text{if } i \neq 1. \end{cases} \tag{7.147c}$$

Hence, the Ritz solution becomes ($c_1 = -1/\pi$ and all other $c_i = 0$):

$$U_1(x, y) = c_1 \phi_1(x, y) + \phi_0(x, y)$$
$$= y \sin \pi x - \frac{1}{\pi} \sin \pi x \sin \pi y = \sin \pi x \left( y - \frac{1}{\pi} \sin \pi y \right). \tag{7.148a}$$

The exact solution of Eq. (7.144a) is given by

$$u(x, y) = \frac{\sin \pi x \sinh \pi y}{\sinh \pi}. \tag{7.148b}$$

**Example 7.10** Consider the following pair of coupled differential equations, which arise in connection with the Timoshenko beam theory (see Exercise 6.15):

$$\frac{\partial}{\partial x}\left[ S\left( \frac{\partial w_0}{\partial x} + \phi_x \right) \right] + q = I_0 \frac{\partial^2 w_0}{\partial t^2}, \tag{7.149}$$

$$\frac{\partial}{\partial x}\left( D \frac{\partial \phi_x}{\partial x} \right) - S\left( \frac{\partial w_0}{\partial x} + \phi_x \right) = I_2 \frac{\partial^2 \phi_x}{\partial t^2}, \tag{7.150}$$



where $S$ is the shear stiffness ($S = K_s GA$; $K_s$ is the shear correction coefficient, $G$ the shear modulus, and $A$ the area of cross section), $D$ the bending stiffness, $w_0$ the transverse deflection, $\phi_x$ the rotation, $q$ the distributed transverse load, and $I_0$ and $I_2$ are mass inertias. Assume that $D$, $S$, $I_0$, and $I_2$ are constants. The "specified" boundary conditions are of the form (as will be clear from the third step of the weak form)

$$-\left( D \frac{\partial \phi_x}{\partial x} \right)_{x=0} = M_1, \qquad \left( D \frac{\partial \phi_x}{\partial x} \right)_{x=L} = M_2, \tag{7.151a}$$

$$-S\left( \frac{\partial w_0}{\partial x} + \phi_x \right)_{x=0} = Q_1, \qquad S\left( \frac{\partial w_0}{\partial x} + \phi_x \right)_{x=L} = Q_2. \tag{7.151b}$$

We wish to derive the Ritz equations of the problem.

First we develop the weak form of the equations using the three-step procedure. Multiply the first equation with weight function $v_1$ and the second one with weight function $v_2$ and integrate over the length of the beam to obtain

$$0 = \int_0^L v_1 \left\{ \frac{\partial}{\partial x}\left[ S\left( \frac{\partial w_0}{\partial x} + \phi_x \right) \right] + q - I_0 \frac{\partial^2 w_0}{\partial t^2} \right\} dx$$
$$- \int_0^L \left\{ -S \frac{\partial v_1}{\partial x}\left( \frac{\partial w_0}{\partial x} + \phi_x \right) + v_1 q - I_0 v_1 \frac{\partial^2 w_0}{\partial t^2} \right\} dx$$
$$+ \left[ S\left( \frac{\partial w_0}{\partial x} + \phi_x \right) v_1 \right]_0^L$$
$$= \int_0^L \left\{ -S \frac{\partial v_1}{\partial x}\left( \frac{\partial w_0}{\partial x} + \phi_x \right) + v_1 q - I_0 v_1 \frac{\partial^2 w_0}{\partial t^2} \right\} dx$$
$$+ Q_1 v_1(0) + Q_2 v_1(L), \tag{7.152a}$$

$$0 = \int_0^L v_2 \left\{ \frac{\partial}{\partial x}\left( D \frac{\partial \phi_x}{\partial x} \right) - S\left( \frac{\partial w_0}{\partial x} + \phi_x \right) - I_2 \frac{\partial^2 \phi_x}{\partial t^2} \right\} dx$$
$$= \int_0^L \left\{ -D \frac{\partial v_2}{\partial x}\frac{\partial \phi_x}{\partial x} - S v_2 \left( \frac{\partial w_0}{\partial x} + \phi_x \right) - I_2 v_2 \frac{\partial^2 \phi_x}{\partial t^2} \right\} dx$$
$$+ \left[ D \frac{\partial \phi_x}{\partial x} v_2 \right]_0^L$$
$$= \int_0^L \left\{ -D \frac{\partial v_2}{\partial x}\frac{\partial \phi_x}{\partial x} - S v_2 \left( \frac{\partial w_0}{\partial x} + \phi_x \right) - I_2 v_2 \frac{\partial^2 \phi_x}{\partial t^2} \right\} dx$$
$$+ M_1 v_2(0) + M_2 v_2(L). \tag{7.152b}$$

Note that integration by parts was used such that the expression $w_{0,x} + \phi_x$ is preserved, as it enters the boundary term representing the shear force. Such considerations can

only be used by knowing the mechanics of the problem at hand. Also, note that the pair of weight functions $(v_1, v_2)$ (from a product space of admissible variations) satisfy the homogeneous form of specified essential boundary conditions on the pair $(w_0, \phi_x)$ (with the correspondence $v_1 \sim w_0$ and $v_2 \sim \phi_x$). Writing the bilinear form for the problem is a little involved; one may treat $u = (w_0, \phi_x)$ and $v = (v_1, v_2)$ as vectors (from a vector space) and then write the bilinear form $B(v, u)$.

We assume Ritz approximations of the form

$$w_0(x,t) \approx \sum_{j=1}^{M} d_j(t)\psi_j(x) + \psi_0(x), \qquad \phi_x(x,t) \approx \sum_{j=1}^{N} c_j(t)\theta_j(x) + \theta_0(x), \tag{7.153}$$

and *derive* the ordinary differential equations involving the time derivatives of $d_i$ and $c_i$. From weak forms (7.152a,b) it is clear that all of the specified boundary conditions are of the natural type. Hence $\psi_0(x) = 0$ and $\theta_0(x) = 0$.

Next we substitute the approximations into the weak forms for $w_0(x)$ and $\phi_x(x)$, and $v_1(x) = \psi_i(x)$ and $v_2(x) = \theta_i(x)$, and obtain

$$\begin{aligned} 0 &= \int_0^L \left[ -S\frac{d\psi_i}{dx}\left(\sum_{j=1}^{m}\frac{d\psi_j}{dx}d_j + \sum_{j=1}^{n} c_j\theta_j\right) + \psi_i q - I_0\psi_i\left(\sum_{j=1}^{m}\frac{d^2d_j}{dt^2}\psi_j\right)\right] dx \\ &\quad + Q_1\psi_i(0) + Q_2\psi_i(L) \\ &= -\sum_{j=1}^{m} A_{ij}d_j - \sum_{j=1}^{n} B_{ij}c_j - \sum_{j=1}^{m} M_{ij}^1\frac{d^2d_j}{dt^2} + F_i^1, \end{aligned} \tag{7.154}$$

$$\begin{aligned} 0 &= \int_0^L \left[ -D\frac{d\theta_i}{dx}\left(\sum_{j=1}^{n}\frac{d\theta_j}{dx}c_j\right) - S\theta_i\left(\sum_{j=1}^{m}\frac{d\psi_j}{dx}d_j + \sum_{j=1}^{n}\theta_j c_j\right)\right. \\ &\qquad \left. -I_2\theta_i\left(\sum_{j=1}^{n}\frac{d^2c_j}{dt^2}\theta_j\right)\right] dx + M_1\theta_i(0) + M_2\theta_i(L) \\ &= -\sum_{j=1}^{m} C_{ij}d_j - \sum_{j=1}^{n} D_{ij}c_j - \sum_{j=1}^{n} M_{ij}^2\frac{d^2c_j}{dt^2} + F_i^2, \end{aligned} \tag{7.155}$$

where

$$A_{ij} = \int_0^L S\frac{d\psi_i}{dx}\frac{d\psi_j}{dx}\,dx, \qquad B_{ij} = \int_0^L S\frac{d\psi_i}{dx}\theta_j\,dx,$$

$$M_{ij}^1 = \int_0^L I_0\psi_i\psi_j\,dx, \qquad F_i^1 = \int_0^L \psi_i q\,dx + Q_1\psi_i(0) + Q_2\psi_i(L),$$



$$C_{ij} = \int_0^L S\theta_i\frac{d\psi_j}{dx}\,dx, \qquad D_{ij} = \int_0^L \left(D\frac{d\theta_i}{dx}\frac{d\theta_j}{dx} + S\theta_i\theta_j\right) dx, \tag{7.156}$$

$$M_{ij}^2 = \int_0^L I_2\theta_i\theta_j\,dx, \qquad F_i^2 = M_1\theta_i(0) + M_2\theta_i(L).$$

In matrix form, we can write this as

$$\begin{bmatrix} [A] & [B] \\ [C] & [D] \end{bmatrix}\begin{Bmatrix} \{d\} \\ \{c\} \end{Bmatrix} + \begin{bmatrix} [M^1] & [0] \\ [0] & [M^2] \end{bmatrix}\begin{Bmatrix} \{\ddot{d}\} \\ \{\ddot{c}\} \end{Bmatrix} = \begin{Bmatrix} \{F^1\} \\ \{F^2\} \end{Bmatrix}, \tag{7.157a}$$

or

$$[K]\{\Delta\} + [M]\{\ddot{\Delta}\} = \{F\}. \tag{7.157b}$$

In closing this section, we make a couple of additional comments on the Ritz method. In developing the weak form, one should bear in mind that the boundary terms obtained from the integration by parts should be physically meaningful. The variational form used for the Ritz method does not have to be a quadratic functional, but it should be a form that includes the natural boundary conditions of the problem. Therefore, the Ritz method can be applied even to nonlinear problems. Of course, the resulting simultaneous algebraic equations are nonlinear and there can be more than one solution to the equations (see Example 7.20). The selection of approximation functions becomes increasingly difficult with the dimension and shape of the domain.

## 7.5 WEIGHTED-RESIDUAL METHODS

### 7.5.1 Introduction

As discussed earlier, weighted-residual methods are those in which we seek approximate solutions using a weighted-integral statement of the equation(s). To fix the ideas, consider a boundary-value problem described by the operator equation

$$A(u) = f \quad \text{in } \Omega, \tag{7.158a}$$

subjected to boundary conditions

$$B_1(u) = \hat{u} \quad \text{on } \Gamma_1, \qquad B_2(u) = \hat{g} \quad \text{on } \Gamma_2, \tag{7.158b}$$

where $A$ is a linear or nonlinear differential operator, $u$ is the dependent variable, $f$ is a given force term in the domain $\Omega$, $B_1$ and $B_2$ are boundary operators associated with essential and natural boundary conditions of the operator $A$, and $\hat{u}$ and $\hat{g}$ are specified values on the portions $\Gamma_1$ and $\Gamma_2$ of the boundary $\Gamma$ of the domain. An example of Eqs. (7.159a) and (7.159b) is given by

$$A(u) = -\frac{d}{dx}\left(a\frac{du}{dx}\right), \qquad B_1(u) = u, \qquad B_2(u) = a\frac{du}{dx},$$

$\Gamma_1$ is the point $x = 0$, and $\Gamma_2$ is the point $x = L$. The weighted-integral statement of Eq. (7.159a) is

$$0 = \int_\Omega w(x)\,[A(u) - f]\,dx. \tag{7.159}$$

We seek a solution $U_N$ that satisfies the specified boundary conditions (7.158b) such that the above equation is also satisfied. A complete discussion of the procedure will be given shortly.

The weak form of the operator equation (7.158a) can be constructed whenever the operator $A$ permits the use of integration by parts [or gradient/divergence theorems in Eqs. (2.89a,b)] to transfer part of the differentiation from the dependent variable $u$ to the weight function $w$ and incorporate the natural boundary conditions of the problem. In general, the weak form can be constructed if the operator $A$ is expressible as a product of two operators:

$$A = T^*(aT) \tag{7.160a}$$

where operator $T^*$ is called the *adjoint* of $T$ and is related to $T$ by

$$\int_\Omega T(u)v\,d\mathbf{x} = \int_\Omega uT^*(v)\,d\mathbf{x} + \oint_\Gamma C_1(u)C_2(v)\,dS \tag{7.160b}$$

for all $u$ and $v$. Here $\Omega$ is domain with boundary $\Gamma$. The boundary operators $C_1$ and $C_2$ depend on the operator $A$. For example, when $A = -(d/dx)[a(du/dx)]$, operators $T$, $T^*$, $C_1$, and $C_2$ are given by

$$T = \frac{d}{dx}, \qquad T^* = -\frac{d}{dx}, \qquad C_1 = 1, \qquad C_2 = 1.$$

If $A = -\nabla^2 = -\nabla \cdot \nabla$, then we have [$u$ must be a scalar and $\mathbf{v}$ a vector function of position in Eq. (7.160b)]:

$$T = \nabla\ (\text{grad}), \qquad T^* = -\nabla \cdot (\text{div}), \qquad C_1 = 1, \qquad C_2 = \hat{\mathbf{n}}\cdot\,.$$

The weak form of Eq. (7.158a,b), with $A$ given by Eq. (7.160a), can be derived as follows:

$$\begin{aligned} 0 &= \int_\Omega w\,[A(u) - f]\,d\mathbf{x} \\ &= \int_\Omega w\,[T^*(aT(u)) - f]\,d\mathbf{x} \\ &= \int_\Omega [T(w)(aT(u)) - wf]\,d\mathbf{x} - \oint_\Gamma B_1(w)B_2(u)\,dS, \end{aligned}$$



where $C_1 = B_1$ and $B_2 = C_2(aT(u))$. Since $B_1(w) = 0$ on portion $\Gamma_1$ (where $B_1(u)$ is specified), and $B_2(u) = \hat{g}$ on $\Gamma_2$, with $\Gamma_1 + \Gamma_2 = \Gamma$, we have

$$0 = \int_\Omega [aT(w)T(u) - wf]\,d\mathbf{x} - \int_{\Gamma_2} B_1(w)\hat{g}\,dS,$$

which is the weak form we set to derive.

Returning to the weighted-residual methods, we seek an approximate solution of $u$ in the form (as in the Ritz method)

$$U_N(x) = \sum_{j=1}^{N} c_j\phi_j(x) + \phi_0(x), \tag{7.161}$$

where the parameters $c_j$ are determined by requiring that the residual in the governing equation due to the approximation

$$\mathcal{R}_N = A\left(\sum_{j=1}^{N} c_j\phi_j + \phi_0\right) - f \neq 0 \tag{7.162}$$

be orthogonal to a set of $N$ linearly independent *weight functions* $\psi_i(x)$, which in general are different from the approximation functions $\phi_i$:

$$\int_\Omega \psi_i\mathcal{R}_N(\mathbf{x}, \{c\}, \{\phi\}, f)\,d\mathbf{x} = 0, \qquad (i = 1, 2, \ldots, N). \tag{7.163}$$

Equation (7.163) is the same as that obtained by substituting approximation (7.161) into the weighted-residual statement (7.159). It provides $N$ linearly independent equations for the determination of the parameters $c_i$. If $A$ is a nonlinear operator, the resulting algebraic equations will be nonlinear.

The approximation functions ($\phi_0$, $\phi_i$) and weight functions $\psi_i$ in a weighted-residual method must satisfy the following conditions:

1. $\phi_j (j = 1, 2, \ldots, N)$ should satisfy three conditions:
   (a) Each $\phi_j$ is *continuous* as required in the weighted-residual statement; i.e., $\phi_j$ should be such that $U_N$ yields a nonzero value of $A(U_N)$.
   (b) Each $\phi_j$ satisfies the *homogeneous form* of *all* specified (i.e., essential as well as natural) boundary conditions.
   (c) The set $\{\phi_j\}$ is *linearly independent* and *complete*.
2. $\phi_0$ has the main purpose of satisfying *all* specified boundary conditions associated with the equation. It is necessarily zero when the specified boundary conditions are homogeneous.
3. The set $\{\psi_i\}$ should be linearly independent. (7.164)

There are two main differences between the approximation functions used in the Ritz method and those used in weighted-residual methods:

1. *Continuity.* The approximation functions used in the weighted-residual methods are required to have the same differentiability as in the differential equation, whereas those used in the Ritz method must be differentiable as required by the weak form.
2. *Boundary conditions.* The approximation functions used in the weighted-residual method must satisfy the homogeneous form of both geometric and force boundary conditions, whereas those used in the Ritz method must satisfy the homogeneous form of only the essential boundary conditions, since the natural boundary conditions are already included in the weak form.

Both of these differences require $\phi_i$ to be of a higher order than those used in the Ritz method.

Various special cases of the weighted-residual method differ from each other due to the choice of the weight function, $\psi_i$. The most commonly used weight functions are:

$$\begin{aligned} &\textit{The Petrov–Galerkin method:} && \psi_i \neq \phi_i. \\ &\textit{Galerkin's method:} && \psi_i = \phi_i. \\ &\textit{Least-squares method:} && \psi_i = A(\phi_i). \\ &\textit{Collocation method:} && \psi_i = \delta(\mathbf{x} - \mathbf{x}_i). \end{aligned} \tag{7.165}$$

Here $\delta(\cdot)$ denotes the Dirac delta function. Although the least-squares method is listed as a special case of the weighted-residual method here, it is based on the concept of minimizing an integral statement. In general, the least-squares method is *not* a special case of the weighted-residual method. These remarks will be discussed in more detail below. In addition to the methods listed above, there are other variational methods (methods in which the unknown parameters $c_i$ are adjusted such that the governing equations are satisfied in a certain sense). These include the subdomain method and Trefftz method. These methods will also be discussed briefly in this chapter.

### 7.5.2 Galerkin's Method

The Galerkin method is a special case of the Petrov–Galerkin method in which the approximation functions and the weighted functions are the same ($\phi_i = \psi_i$). Hence, the Galerkin integral is given by

$$\int_\Omega \phi_i \mathcal{R}_N(\mathbf{x}, \{c\}, \{\phi\}, f)\, d\mathbf{x} = 0, \qquad (i = 1, 2, \ldots, N). \tag{7.166}$$

If the Galerkin method is used for second-order or higher-order equations, it would involve the use of higher-order coordinate functions and the solution of nonsymmetric equations.



The Ritz and Galerkin methods yield the same set of algebraic equations for the following two cases:

1. The specified boundary conditions of the problem are all of the essential type, and therefore the requirements on $\phi_i$ and $\phi_0$ in both methods are the same.
2. The problem has both essential and natural boundary conditions, but the coordinate functions used in the Galerkin method are also used in the Ritz method.

### 7.5.3 Least-Squares Method

The least-squares method is based on the idea of minimizing the integral of the square of the residual:

$$\text{minimize } I(c_1, c_2, \ldots, c_N) = \int_\Omega \mathcal{R}_N^2(\mathbf{x}, \{c\}, \{\phi\}, f)\, d\mathbf{x}. \tag{7.167a}$$

We obtain (from $\delta I = 0$)

$$0 = \int_\Omega \frac{\partial \mathcal{R}_N}{\partial c_i} \mathcal{R}_N(\mathbf{x}, \{c\}, \{\phi\}, f)\, d\mathbf{x}, \tag{7.167b}$$

which, when $A$ is a linear operator, becomes

$$0 = \int_\Omega A(\phi_i) \mathcal{R}_N(\mathbf{x}, \{c\}, \{\phi\}, f)\, d\mathbf{x}. \tag{7.167c}$$

Clearly, Eq. (7.167c) is a special case of Eq. (7.163) with $\psi_i = A(\phi_i)$. The least-squares method is more suitable for first-order equations. For eigenvalue problems and time-dependent problems it is not suitable, as shown below. On the other hand, the least-squares method is the only other method, in addition to the Ritz method, that is based on the minimization of a functional. The least-squares method also results in a positive-definite coefficient matrix.

### 7.5.4 Collocation Method

In the collocation method we require the residual to vanish at a selected number of points $\mathbf{x}^i$ in the domain:

$$\mathcal{R}_N(\mathbf{x}^i, \{c\}, \{\phi\}, f) = 0, \qquad (i = 1, 2, \ldots, N), \tag{7.168a}$$

which can be written, with the help of the Dirac delta function, as

$$\int_\Omega \delta(\mathbf{x} - \mathbf{x}^i) \mathcal{R}_N(\mathbf{x}, \{c\}, \{\phi\}, f)\, d\mathbf{x} = 0, \qquad (i = 1, 2, \ldots, N). \tag{7.168b}$$

Thus, the collocation method is a special case of the weighted-residual method (7.163) with $\psi_i(\mathbf{x}) = \delta(\mathbf{x} - \mathbf{x}^i)$. In the collocation method, one must choose as many collocation points as there are undetermined parameters. In general, these points should be distributed uniformly in the domain. Otherwise, ill-conditioned equations among $c_j$ may result.

### 7.5.5 Eigenvalue and Time-Dependent Problems

It should be noted that if the problem at hand is an eigenvalue problem or a time-dependent problem, the operator equation (7.158a) takes the following alternative forms:

*Eigenvalue Problem*

$$A(u) - \lambda C(u) = 0. \tag{7.169}$$

*Time-Dependent Problem*

$$A_t(u) + A(u) = f(x, t). \tag{7.170}$$

In Eq. (7.169), $\lambda$ is the eigenvalue, which is to be determined along with the eigenvector $u(x)$, and $A$ and $C$ are spatial differential operators. An example of the equation is provided by the buckling of a beam column

$$\frac{d^2}{dx^2}\left(EI\frac{d^2u}{dx^2}\right) + P\frac{d^2u}{dx^2} = 0,$$

where $u$ denotes the lateral deflection and $P$ is the axial compressive load. The problem involves determining the value of $P$ and mode shape $u(x)$ such that the governing equation and certain end conditions of the beam are satisfied. The minimum value of $P$ is called the *critical buckling load*. Comparing the above equation with Eq. (7.169), we see that

$$A(u) = \frac{d^2}{dx^2}\left(EI\frac{d^2u}{dx^2}\right), \qquad \lambda = P, \qquad C(u) = -\frac{d^2u}{dx^2}.$$

In Eq. (7.170), $A$ is a spatial differential operator and $A_t$ is a temporal differential operator. Examples of Eq. (7.170) are provided by the equations of heat transfer in a plane wall and axial motion of a bar, respectively:

$$\rho c_v \frac{\partial u}{\partial t} - \frac{\partial}{\partial x}\left(k\frac{\partial u}{\partial x}\right) = f(x, t),$$

$$-\rho\frac{\partial^2 u}{\partial t^2} - \frac{\partial}{\partial x}\left(EA_0\frac{\partial u}{\partial x}\right) = f(x, t).$$



In the first equation, $u$ denotes temperature, $\rho$ the density, $c_v$ specific heat at constant volume, $k$ the conductivity, and $f$ is internal heat generation. Clearly, we have

$$A_t(u) = \rho c_v \frac{\partial u}{\partial t}, \qquad A(u) = -\frac{\partial}{\partial x}\left(k\frac{\partial u}{\partial x}\right).$$

In the second equation, $u$ denotes the axial displacement, $\rho$ the density, $E$ Young's modulus, $A_0$ area of cross section, and $f$ body force per unit length. In this case, we have

$$A_t(u) = -\rho\frac{\partial^2 u}{\partial t^2}, \qquad A(u) = -\frac{\partial}{\partial x}\left(EA_0\frac{\partial u}{\partial x}\right).$$

Application of the weighted-residual method to Eqs. (7.169) and (7.170) follows the same idea, i.e., Eq. (7.163) holds. However, two comments are in order.

1. In the case of time-dependent problems, the integral in Eq. (7.163) is still over the spatial domain and the weight function $\psi$ is a function of the spatial coordinate only. Thus, Eq. (7.163) leads to a set of ordinary differential equations in time among $c_j(t)$. These need to be further approximated using time-approximation schemes.
2. In the case of the least-squares method, the question arises as to what should be the weight function $\psi_i$. Let us examine the least-squares method first for the linear eigenvalue problem (i.e., $A$ and $C$ are linear operators). We have

$$\begin{aligned} 0 &= \delta\int_\Omega \mathcal{R}_N^2\, d\mathbf{x}, \qquad \mathcal{R}_N = A(U_N) - \lambda C(U_N) \\ &= 2\int_\Omega [A(\phi_i) - \lambda C(\phi_i)]\,\mathcal{R}_N\, d\mathbf{x}, \qquad (i = 1, 2, \ldots, N) \\ &= 2\sum_{j=1}^{N}\left\{\int_\Omega [A(\phi_i) - \lambda C(\phi_i)]\,[A(\phi_j) - \lambda C(\phi_j)]\, d\mathbf{x}\right\} c_j. \end{aligned}$$

Clearly, the eigenvalue problem becomes quadratic in $\lambda$, which is not desirable from a computational viewpoint.

In the case of time-dependent problems, we have

$$\begin{aligned} 0 &= \delta\int_\Omega \mathcal{R}_N^2\, d\mathbf{x}, \qquad \mathcal{R}_N = A(U_N) + A_t(U_N) \\ &= \int_\Omega \left[A(\phi_i) + \sum_{k=1}^{N}\phi_k \frac{\partial}{\partial c_i} A_t(c_k)\right]\mathcal{R}_N\, d\mathbf{x}, \qquad (i = 1, 2, \ldots, N). \end{aligned}$$

which is also complicated. Thus, the least-squares method leads to complicated systems of equations for eigenvalue or time-dependent problems. An alternative is to use $\psi_i = A(\phi_i)$ in all cases. This avoids the problems seen above.

It is possible to develop the so-called space-time approximations, i.e., to use the variational methods, treating time as an additional coordinate. It is found that such approaches are complicated, and they are not used in practice. Therefore, we will not consider the space-time approximations in the present study.

### 7.5.6 Equations for Undetermined Parameters

Here we develop the discrete equations for the equilibrium, eigenvalue, and time-dependent problems under the assumption that $A$ is a linear operator, as is the case with most problems considered in this study. The equations are valid for all special cases of the weighted-residual method except for the least-squares method. They are also valid for the least-squares method if one accepts the use of $A(\phi_i)$ for $\psi_i$, which is certainly true for equilibrium problems. For eigenvalue problems we assume that all boundary conditions are homogeneous and therefore $\phi_0 = 0$.

***Equilibrium Problems*** We have

$$A\left(\sum_{j=1}^{N} c_j\phi_j + \phi_0\right) = \sum_{j=1}^{N} c_j A(\phi_j) + A(\phi_0), \tag{7.171}$$

and Eq. (7.163) for the equilibrium equation (7.158a) becomes

$$\sum_{j=1}^{N}\left[\int_\Omega \psi_i A(\phi_j)\,d\mathbf{x}\right] c_j - \int_\Omega \psi_i[f - A(\phi_0)]\,d\mathbf{x} = 0$$

or

$$\sum_{j=1}^{N} a_{ij}c_j - b_i = 0 \;\; (i = 1, 2, \ldots, N), \qquad [A]\{c\} = \{b\}, \tag{7.172a}$$

where

$$a_{ij} = \int_\Omega \psi_i A(\phi_j)\,d\mathbf{x}, \qquad b_i = \int_\Omega \psi_i\,[f - A(\phi_0)]\,d\mathbf{x}. \tag{7.172b}$$

Note that $a_{ij}$ is *not symmetric* in general, even when $\psi_i = \phi_i$ (Galerkin's method). It is symmetric in the least-squares method because $\psi_i = A(\phi_i)$. Also, if (a) $A$ is an operator of the form in Eq. (7.160a), (b) $\psi_i = \phi_i$, and (c) $\phi_i$ satisfy the homogeneous form of specified essential and natural boundary conditions, it can be shown that

$$a_{ij} = \int_\Omega T(\phi_i)aT(\phi_j)\,d\mathbf{x}, \qquad b_i = \int_\Omega \phi_i f\,d\mathbf{x} + \int_{\Gamma_2} \phi_i \hat{g}\,dS. \tag{7.172c}$$



***Eigenvalue Problems*** For linear eigenvalue problem of the form in Eq. (7.169), we have

$$\sum_{j=1}^{N}\left\{\int_\Omega \psi_i[A(\phi_j) - \lambda C(\phi_j)]d\mathbf{x}\right\} c_j = 0$$

or

$$\sum_{j=1}^{N}(a_{ij} - \lambda g_{ij})\,c_j = 0 \;\; (i = 1, 2, \ldots, N), \qquad ([A] - \lambda[G])\{c\} = \{0\}, \tag{7.173a}$$

where $a_{ij}$ are defined by Eq. (7.172b) and

$$g_{ij} = \int_\Omega \psi_i C(\phi_j)\,d\mathbf{x}. \tag{7.173b}$$

Note that using Eqs. (7.173a,b) for the least-squares method amounts to using $\psi_i = A(\phi_i)$; $a_{ij}$ is symmetric but $g_{ij}$ is unsymmetric.

***Time-Dependent Problems*** For linear time-dependent problem of the form in Eq. (7.170), we have

$$\sum_{j=1}^{N}\left\{\int_\Omega \psi_i[A(\phi_j)c_j + \phi_j A_t(c_j)]d\mathbf{x}\right\} - \int_\Omega \psi_i[f - A(\phi_0)]\,d\mathbf{x} = 0$$

or

$$\sum_{j=1}^{N}(a_{ij}c_j + m_{ij}A_t(c_j)) = 0 \;\; (i = 1, 2, \ldots, N), \qquad [A]\{c\} + \lambda[M]\{A_t(c)\} = \{0\}, \tag{7.174a}$$

where $a_{ij}$ are defined by Eq. (7.172b) and

$$m_{ij}A_t(c_j) = \int_\Omega \psi_i\phi_j A_t(c_j)\,d\mathbf{x}. \tag{7.174b}$$

Again recall that using Eqs. (7.174a,b) for the least-squares method amounts to using $\psi_i = A(\phi_i)$.

### 7.5.7 Examples

A number of examples are considered here to illustrate the use of various weighted residual methods. Equilibrium, eigenvalue, and time-dependent problems are considered.

**Example 7.11** Consider the eigenvalue problem of Example 7.5 in a nondimensional form [see Eqs. (7.77a,b)]:

$$-\frac{d^2u}{dx^2} - \lambda u = 0, \qquad u(0) = 0, \qquad \frac{du}{dx} + u = 0 \text{ at } x = 1.$$

In the weighted-residual method, $\phi_i$ must satisfy not only the condition $\phi_i(0) = 0$ but also the condition $\phi_i'(1) + \phi_i(1) = 0$. The lowest-order function that satisfies the two conditions is

$$\phi_1(x) = 3x - 2x^2. \tag{7.175}$$

The one-parameter Galerkin's solution for the natural frequency (see Example 7.5) can be computed using

$$0 = c_1 \int_0^1 \phi_1 \left( \frac{d^2\phi_1}{dx^2} + \lambda\phi_1 \right) dx \qquad \text{or} \qquad \left( -\frac{10}{3} + \frac{4}{5}\lambda \right) c_1 = 0, \tag{a}$$

which gives (for nonzero $c_1$) $\lambda = 50/12 = 4.167$. If the same function is used for $\phi_1$ in the one-parameter Ritz solution, we obtain, as discussed in Example 7.5, the same result as in the one-parameter Galerkin solution.

If we use the one-parameter collocation method with the collocation point at $x = 0.5$, we obtain [$\phi_1(0.5) = 1.0$ and $(d^2\phi_1/dx^2) = -4.0$]:

$$0 - c_1\phi_1(0.5) \left[ \left( \frac{d^2\phi_1}{dx^2} \right) \Big|_{x=0.5} + \lambda\phi_1(0.5) \right] \qquad \text{or} \qquad (-4 + \lambda)\, c_1 = 0,$$

which gives $\lambda = 4$.

The one-parameter least-squares approximation with $\psi_1 = A(\phi_1)$ gives

$$0 = c_1 \int_0^1 \frac{d^2\phi_1}{dx^2} \left( \frac{d^2\phi_1}{dx^2} + \lambda\phi_1 \right) dx \qquad \text{or} \qquad \left( -4 + \tfrac{5}{6}\lambda \right) c_1 = 0$$

which gives $\lambda = 4.8$. If we use $\psi_1 = A(\phi_1) - \lambda\phi_1$, we obtain

$$\begin{aligned} 0 &= c_1 \int_0^1 \left( \frac{d^2\phi_1}{dx^2} + \lambda\phi_1 \right) \left( \frac{d^2\phi_1}{dx^2} + \lambda\phi_1 \right) dx \\ &= \left( \tfrac{4}{5}\lambda^2 - \tfrac{20}{3}\lambda + 16 \right) c_1, \end{aligned} \tag{b}$$

whose roots are

$$\lambda_{1,2} = \tfrac{25}{6} \pm \tfrac{1}{6}\sqrt{445} \;\rightarrow\; \lambda_1 = 7.6825, \qquad \lambda_2 = 0.6508. \tag{c}$$

Neither root is closer to the exact value of 4.116. This indicates that the least-squares method with $\psi_i = A(\phi_i)$ is perhaps more suitable than $\psi_i = A(\phi_i) - \lambda C(\phi_i)$.



Let us consider a two-parameter weighted-residual solution to the problem

$$U_2(x) = c_1\phi_1(x) + c_2\phi_2(x), \tag{d}$$

where $\phi_1(x)$ is given by Eq. (7.175). To determine $\phi_2(x)$, we begin with a polynomial that is one degree higher than that used for $\phi_1$:

$$\phi_2(x) = a + bx + cx^2 + dx^3,$$

and obtain

$$\phi_2(0) = 0 \rightarrow a = 0; \qquad \phi_2'(1) + \phi_2(1) = 0 \rightarrow 2b + 3c + 4d = 0 \text{ or } d = -\tfrac{2}{4}b - \tfrac{3}{4}c.$$

We can arbitrarily pick the values of $b$ and $c$, except that they are not both equal to zero (for obvious reasons). Thus we have an infinite number of possibilities. If we pick $b = 0$ and $c = 4$, we have $d = -3$, and $\phi_2$ becomes

$$\phi_2(x) = a + bx + cx^2 + dx^3 = 4x^2 - 3x^3. \tag{7.176a}$$

On the other hand, if we choose $b = 1$ and $c = 2$, we have $d = -2$, and $\phi_2$ becomes

$$\phi_2(x) = a + bx + cx^2 + dx^3 = x + 2x^2 - 2x^3 = \varphi_2(x). \tag{7.176b}$$

The set $\{\phi_1, \phi_2\}$ is equivalent to the set $\{\phi_1, \varphi_2\}$. Note that

$$\begin{aligned} U_2(x) &= c_1\phi_1(x) + c_2\phi_2(x) \\ &= c_1(3x - 2x^2) + c_2(4x^2 - 3x^3) \\ &= 3c_1x + (-2c_1 + 4c_2)x^2 - 3c_2x^3, \\ U_2(x) &\approx c_1\phi_1(x) + c_2\varphi_2(x) \\ &= \bar{c}_1(3x - 2x^2) + \bar{c}_2(x + 2x^2 - 2x^3) \\ &= (3\bar{c}_1 + \bar{c}_2)\, x + (-2\bar{c}_1 + 2\bar{c}_2)\, x^2 - 2\bar{c}_2x^3. \end{aligned}$$

Comparing the two relations, we can show that

$$\bar{c}_1 = c_1 - 0.5c_2, \qquad \bar{c}_2 = 1.5c_2.$$

Hence, either set will yield the same final solution for $U_2(x)$ or $\lambda$.

Using $\phi_1$ and $\phi_2$ [from Eq. (7.176a)], we compute the residual of the approximation as

$$\begin{aligned} R = -\frac{d^2U_2}{dx^2} - \lambda U_2 &= -c_1\frac{d^2\phi_1}{dx^2} - c_2\frac{d^2\phi_2}{dx^2} - \lambda\,(c_1\phi_1 + c_2\phi_2) \\ &= c_1\left( -\frac{d^2\phi_1}{dx^2} - \lambda\phi_1 \right) + c_2\left( -\frac{d^2\phi_2}{dx^2} - \lambda\phi_2 \right). \end{aligned} \tag{e}$$

For the Galerkin method, we set the integral of the weighted residual to zero and obtain

$$0 = \int_0^1 \phi_1(x) R\,dx = \int_0^1 \phi_1(x)\left[-c_1\frac{d^2\phi_1}{dx^2} - c_2\frac{d^2\phi_2}{dx^2} - \lambda(c_1\phi_1 + c_2\phi_2)\right]dx$$
$$= K_{11}c_1 + K_{12}c_2 - \lambda(M_{11}c_1 + M_{12}c_2),$$
$$0 = \int_0^1 \phi_2(x) R\,dx = \int_0^1 \phi_2(x)\left[-c_1\frac{d^2\phi_1}{dx^2} - c_2\frac{d^2\phi_2}{dx^2} - \lambda(c_1\phi_1 + c_2\phi_2)\right]dx$$
$$= K_{21}c_1 + K_{22}c_2 - \lambda(M_{21}c_1 + M_{22}c_2).$$

In matrix form, we have

$$[K]\{c\} - \lambda[M]\{c\} = \{0\},$$

where

$$K_{ij} = -\int_0^1 \phi_i\frac{d^2\phi_j}{dx^2}\,dx, \qquad M_{ij} = \int_0^1 \phi_i\phi_j\,dx.$$

First, for the choice of functions in Eqs. (7.176a) and (7.176b), we have

$$\frac{d^2\phi_1}{dx^2} = -4, \qquad \frac{d^2\phi_2}{dx^2} = 8 - 18x.$$

Evaluating the integrals, we obtain

$$K_{11} = -\int_0^1 \phi_1\frac{d^2\phi_1}{dx^2}\,dx = \int_0^1 (3x - 2x^2)(4)\,dx = \frac{10}{3},$$
$$K_{12} = -\int_0^1 \phi_1\frac{d^2\phi_2}{dx^2}\,dx = \int_0^1 (3x - 2x^2)(-8 + 18x)\,dx = \frac{7}{3},$$
$$K_{21} = -\int_0^1 \phi_2\frac{d^2\phi_1}{dx^2}\,dx = \int_0^1 (4x^2 - 3x^3)(4)\,dx = \frac{7}{3},$$
$$K_{22} = -\int_0^1 \phi_2\frac{d^2\phi_2}{dx^2}\,dx = \int_0^1 (4x^2 - 3x^3)(-8 + 18x)\,dx = \frac{38}{15},$$
$$M_{11} = \int_0^1 \phi_1\phi_1\,dx = \int_0^1 (3x - 2x^2)(3x - 2x^2)\,dx = \frac{4}{5},$$
$$M_{12} = \int_0^1 \phi_1\phi_2\,dx = \int_0^1 (3x - 2x^2)(4x^2 - 3x^3)\,dx = \frac{3}{5} = M_{21},$$
$$M_{22} = \int_0^1 \phi_2\phi_2\,dx = \int_0^1 (4x^2 - 3x^3)(4x^2 - 3x^3)\,dx = \frac{17}{35},$$



and

$$\left(\frac{1}{15}\begin{bmatrix}50 & 35\\ 35 & 38\end{bmatrix} - \frac{\lambda}{35}\begin{bmatrix}28 & 21\\ 21 & 17\end{bmatrix}\right)\begin{Bmatrix}c_1\\ c_2\end{Bmatrix} = \begin{Bmatrix}0\\ 0\end{Bmatrix}.$$

For nontrivial solution, $c_1 \neq 0$ and $c_2 \neq 0$, we set the determinant of the coefficient matrix to zero to obtain the characteristic polynomial

$$675 - \frac{1332}{7}\lambda + \frac{315}{49}\lambda^2 = 0 \qquad \text{or} \qquad 525 - 148\lambda + 5\lambda^2 = 0, \tag{f}$$

which gives

$$\lambda_1 = 4.121, \qquad \lambda_2 = 25.479. \tag{g}$$

Clearly, the value of $\lambda_1$ has improved over that computed using the one-parameter approximation. The exact value of the second eigenvalue is 24.139.

If we were to use the collocation method, we may select $x = 1/3$ and $x = 2/3$ as the collocation points, among other choices. We leave this as an exercise to the reader.

**Example 7.12** Next we consider the transient response of the problem discussed in Example 7.6. The governing equations are

$$-\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial t^2} = 1, \qquad u(0,t) = 0, \qquad \frac{\partial u}{\partial x} + u = 0 \ \text{ at } x = L \text{ for all } t > 0,$$

with zero initial conditions. We use the one-parameter approximation $u(x,t) \approx c_1(t)\phi_1(x)$ with $\phi_1(x)$ defined in Eq. (7.175).

For the Galerkin method, we obtain

$$0 = \int_0^1 \phi_1\left(-c_1\frac{d^2\phi_1}{dx^2} + \frac{d^2c_1}{dt^2}\phi_1 - 1\right)dx \qquad \text{or} \qquad \frac{4}{5}\frac{d^2c_1}{dt^2} + \frac{10}{3}c_1 = \frac{5}{6}, \tag{a}$$

whose (exact) solution is ($\sqrt{50/12} \approx 2.0412$):

$$c_1(t) = A\sin 2.04t + B\cos 2.04t + \tfrac{1}{4}.$$

For zero initial conditions, $u(x,0) = \dot{u}(x,0) = 0$ [or $c_1(0) = 0$ and $\dot{c}_1(0) = 0$], the total solution becomes ($A = 0$ and $B = -1/4$):

$$u_1(x,t) = \tfrac{1}{4}(1 - \cos 2.04t)(3x - 2x^2). \tag{b}$$

For the one-parameter collocation method with the collocation point at $x = 0.5$, we obtain

$$0 = \phi_1(0.5)\left[-c_1\left(\frac{d^2\phi_1}{dx^2}\right)\Big|_{0.5} + \frac{d^2c_1}{dt^2}\phi_1(0.5) - 1\right] \qquad \text{or} \qquad \frac{d^2c_1}{dt^2} + 4c_1 = 1, \tag{c}$$

so that

$$c_1(t) = A \sin 2t + B \cos 2t + \tfrac{1}{4},$$

and the one-parameter collocation solution becomes

$$u_1(x, t) = \tfrac{1}{4}(1 - \cos 2t)\left(3x - 2x^2\right). \tag{d}$$

**Example 7.13** Consider the simply supported beam problem of Example 7.7. Since all specified boundary conditions are homogeneous, again we have $\phi_0 = 0$. In the Galerkin method, $\phi_i$ must satisfy the homogeneous form of all specified boundary conditions ($w_0 = M_{xx} = 0$ at $x = 0, L$):

$$\phi_i = 0, \qquad \frac{d^2\phi_i}{dx^2} = 0. \tag{a}$$

For the choice of algebraic polynomials, we assume a five-parameter polynomial because there are four conditions in Eq. (a):

$$\phi_1(x) = a + bx + cx^2 + dx^3 + ex^4.$$

Using the boundary conditions, we find that

$$a = c = 0, \qquad bL + dL^3 + eL^4 = 0, \qquad 6dL + 12eL^2 = 0.$$

Thus we have $b = eL^3$ and $d = -2eL$. The function $\phi_1$ is given by (taking $eL^4 = 1$)

$$\phi_1 = \frac{x}{L}\left(1 - 2\frac{x^2}{L^2} + \frac{x^3}{L^3}\right). \tag{b}$$

Substituting the one-parameter Galerkin approximation $W_1 = c_1\phi_1$ into the residual,

$$\mathcal{R} = EI\frac{d^4W_1}{dx^4} - q_0 = \frac{24EI}{L^4}c_1 - q_0. \tag{c}$$

Since the residual is already a constant (which implies that the solution is exact), there is no need to integrate the weighted residual over the domain. By setting $\mathcal{R}$ to zero we obtain $c_1 = q_0L^4/(24EI)$, and the solution becomes

$$W_1(x) = \frac{q_0L^4}{24EI}\left[\left(\frac{x}{L}\right) - 2\left(\frac{x}{L}\right)^3 + \left(\frac{x}{L}\right)^4\right], \tag{d}$$

which coincides with the exact solution. Note that the solution obtained is independent of the particular weighted residual method. It can be shown that a one-parameter Ritz solution with $\phi_1$ given by Eq. (b) also yields the same exact solution (d).

Note that the solution to this problem is symmetric about $x = L/2$. Hence, we can use a half-beam model to solve the problem. In using a half beam, we must address the boundary conditions at $x = L/2$. The boundary conditions at this point



are that the slope is zero, $(dw_0/dx) = 0$, and shear force is zero. Hence, for this case, the approximation functions $\phi_i$ in the weighted-residual method must satisfy the conditions

$$\text{at } x = 0\text{: } \phi_i = 0, \quad \frac{d^2\phi_i}{dx^2} = 0, \qquad \text{and at } x = \frac{L}{2}\text{: } \frac{d\phi_i}{dx} = 0, \quad \frac{d^3\phi_i}{dx^3} = 0.$$

Obviously, the function $\phi_1(x)$ of Eq. (b) satisfies the above conditions. Another choice of $\phi_i$ is provided by

$$\phi_i(x) = \sin\frac{(2i-1)\pi x}{L}.$$

**Example 7.14** Another weighted-residual method that has not been introduced formally in this study is the *subdomain method*. In the subdomain method, we divide the domain of the problem into as many subdomains as there are undetermined parameters, $c_i$, and then on each subdomain we require the integral of the residual $\mathcal{R}_N$ to be zero:

$$\int_{\Omega_i} \mathcal{R}_N(\mathbf{x}, \{c\}, \{\phi\}, f)\, d\mathbf{x} = 0, \qquad (i = 1, 2, \ldots, N), \tag{7.177a}$$

where $\Omega_i$ is the $i$th subdomain. The method can be viewed as a piecewise application of the weighted residual method with $\psi_i = 1$:

$$\psi_i = \begin{cases} 1, & \text{if } \mathbf{x} \in \Omega_i, \\ 0, & \text{otherwise.} \end{cases} \tag{7.177b}$$

Obviously, in this method, negative errors can cancel positive errors to give zero net error (unless $\mathcal{R}_N$ is a positive function in $\Omega_i$), although the sum of the absolute values of the errors may be very large. The finite volume method that is popular in fluid dynamics is based on ideas similar to the subdomain method.

As an example, we consider the beam problem of Example 7.13 (see Fig. 7.5). Using the half-beam model with the following two-parameter approximation:

$$W_2(x) = c_1 \sin\frac{\pi}{L} + c_2 \sin\frac{3\pi x}{L}, \tag{a}$$

we determine approximate solutions using both the collocation method and the subdomain method.

*Collocation Method* Using collocation points at $x = L/4$ and $x = L/2$, we obtain

$$EI\left[c_1\left(\frac{\pi}{L}\right)^4 \sin\frac{\pi}{4} + c_2\left(\frac{3\pi}{L}\right)^4 \sin\frac{3\pi}{4}\right] - q_0 = 0,$$

$$EI\left[c_1\left(\frac{\pi}{L}\right)^4 \sin\frac{\pi}{2} + c_2\left(\frac{3\pi}{L}\right)^4 \sin\frac{3\pi}{2}\right] - q_0 = 0,$$

which yield

$$EI\left(\frac{\pi}{L}\right)^4\begin{bmatrix}1/\sqrt{2} & 81/\sqrt{2}\\ 1 & -81\end{bmatrix}\begin{Bmatrix}c_1\\ c_2\end{Bmatrix}=q_0\begin{Bmatrix}1\\ 1\end{Bmatrix}.$$

or

$$c_1=\frac{\sqrt{2}+1}{2\pi^4}\frac{q_0L^4}{EI},\qquad c_2=\frac{\sqrt{2}-1}{162\pi^4}\frac{q_0L^4}{EI}$$

The two-parameter collocation solution becomes

$$W_2(x)=\frac{q_0L^4}{162EI\pi^4}\left(195.55\sin\frac{\pi x}{L}+0.414\sin\frac{3\pi x}{L}\right). \tag{b}$$

The maximum deflection, $W_2(L/2)=1.205(q_0L^4/EI\pi^4)=(q_0L^4/80.87EI)$, is 5% in error compared to the exact value $(q_0L^4/76.8EI)$.

*Subdomain Method* We consider two subdomains, $\Omega_1=(0, L/4)$ and $\Omega_2=(L/4, L/2)$, in the first half of the beam. For the same choice of coordinate functions as in the collocation method, we obtain

$$\int_0^{L/4}\left[c_1EI\left(\frac{\pi^4}{L}\right)^4\sin\frac{\pi x}{L}+c_2EI\left(\frac{3\pi}{L}\right)^4\sin\frac{3\pi x}{L}-q_0\right]dx=0,$$

$$\int_{L/4}^{L/2}\left[c_1EI\left(\frac{\pi^4}{L}\right)^4\sin\frac{\pi x}{L}+c_2EI\left(\frac{3\pi}{L}\right)^4\sin\frac{3\pi x}{L}-q_0\right]dx=0,$$

or

$$EI\left(\frac{\pi}{L}\right)^4\frac{L}{\pi}\left(1-\frac{1}{\sqrt{2}}\right)c_1+EI\left(\frac{3\pi}{L}\right)^4\frac{L}{3\pi}\left(1+\frac{1}{\sqrt{2}}\right)c_2=\frac{q_0L}{4},$$

$$EI\left(\frac{\pi}{L}\right)^4\frac{L}{\pi}\frac{1}{\sqrt{2}}c_1+EI\left(\frac{3\pi}{L}\right)^4\frac{L}{3\pi}\left(-\frac{1}{\sqrt{2}}\right)c_2=\frac{q_0L}{4}.$$

The solution of these equations is

$$c_1=\frac{\sqrt{2}+1}{4\sqrt{2}\pi^3}\frac{q_0L^4}{EI},\qquad c_2=\frac{\sqrt{2}-1}{108\sqrt{2}\pi^3}\frac{q_0L^4}{EI},$$

and the solution $W_2$ becomes

$$W_2(x)=\frac{q_0L^4}{108\sqrt{2}EI\pi^3}\left(65.184\sin\frac{\pi x}{L}+0.414\sin\frac{3\pi x}{L}\right). \tag{c}$$

The center deflection obtained in the subdomain method, $W_2(L/2)=(q_0L^4/73.12EI)$, is $-5\%$ in error compared to the exact value.



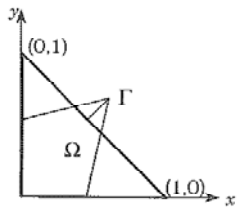


**Figure 7.9** A triangular membrane.

**Example 7.15** Consider the equation

$$A(u)=\nabla^2u+\lambda u=0\quad\text{in }\Omega,\qquad u=0\quad\text{on }\Gamma, \tag{a}$$

where $\Omega$ is the triangular domain shown in Fig. 7.9 and $\Gamma$ is its boundary. Equation (a) describes a nondimensional form of the equation governing the natural vibration of a triangular membrane of side $a$, mass density $\rho$, and tension $T$ ($\lambda=\rho a^2\omega^2/T$, $\omega$ the frequency of vibration). We wish to determine the fundamental frequency (i.e., determine $\lambda$) of vibration by using a one-parameter Galerkin approximation of the problem.

The Galerkin method is based on the weighted-integral statement

$$0=\int_\Omega c_1\left(\frac{\partial^2\phi_1}{\partial x^2}+\frac{\partial^2\phi_1}{\partial y^2}+\lambda\phi_1\right)\phi_1\,dxdy. \tag{b}$$

The function $\phi_1(x, y)$ must vanish on the boundary $\Gamma$. Thus, we have

$$\phi_1(x,0)=0,\qquad \phi_1(0,y)=0,\qquad \phi_1(x,y)=0\text{ on the line }x+y-1=0. \tag{c}$$

Hence the choice for $\phi_1(x, y)$ is

$$\phi_1(x,y)=(x-0)(y-0)(x+y-1),\qquad \frac{\partial^2\phi_1}{\partial x^2}+\frac{\partial^2\phi_1}{\partial y^2}=2(x+y). \tag{d}$$

Hence we have

$$0=c_1\int_0^1\int_0^{1-y}[2(x+y)+\lambda xy(x+y-1)]xy(x+y-1)\,dxdy$$

or

$$\lambda=-\frac{\int_0^1\int_0^{1-y}2(x+y)xy(x+y-1)\,dxdy}{\int_0^1\int_0^{1-y}x^2y^2(x+y-1)^2\,dxdy}=56. \tag{e}$$

**Example 7.16** Consider the differential equation (Poisson's equation)

$$A(u) \equiv -\left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}\right) = f_0 \qquad \text{in} \quad -a \le x,\ y \le a$$
$$u = 0 \qquad \text{on the boundary} \tag{7.178}$$

in a square region. The origin of the coordinate system is taken at the center of the domain, as shown in Fig. 7.10. The equation arises, among others, in connection with (1) the transverse deflection of a membrane fixed on all sides and subjected to uniform pressure $f_0$, (2) the study of the torsion of a square cross-section prismatic bar with $f_0 = 2G\theta$, where $G$ denotes the shear modulus and $\theta$ is the angle of twist per unit length, and (3) conduction heat transfer in a square region with internal heat generation of $f_0$ unit area. The function $u$ denotes the deflection $u$ in the case of a membrane, the Prandtl stress function $\Phi$ in the case of the torsion problem, and the temperature $T$ in the case of conduction heat transfer.

The quadratic functional associated with Eq. (7.178) is given by

$$I(u) = \int_{-a}^{a}\int_{-a}^{a} \frac{1}{2}\left[\left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2 - 2f_0 u\right] dxdy, \tag{7.179}$$

which represents the total potential energy functional for the membrane problem and the total complementary energy functional for the torsion problem. For the heat conduction problem, $I(u)$ does not have a physical meaning. The functional $I(u)$ can be derived by the weak form procedure outlined in Section 7.4.1. The boundary condition in Eq. (7.178) is of the essential type. The natural boundary condition for the equation involves the specification of $\partial u/\partial n$, which is not the specified boundary condition of the problem, unless a quadrant is used.

For the Galerkin method, we assume ($\phi_0 = 0$)

$$U_N(x, y) = \sum_{j=1}^{N} c_j \phi_j(x, y) \tag{7.180}$$

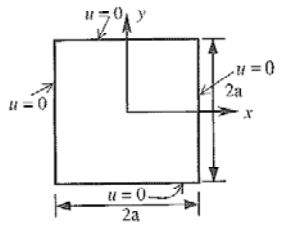


**Figure 7.10** The square domain of Example 7.16.



where

$$\phi_1 = (a^2 - x^2)(a^2 - y^2), \qquad \phi_2 = (x^2 + y^2)\phi_1, \ldots. \tag{7.181}$$

Clearly, $\phi_i$ are twice differentiable with respect to $x$ and $y$ and satisfy the boundary conditions. From Eq. (7.172b), we have

$$a_{ij} = \int_{-a}^{a}\int_{-a}^{a} \phi_i A(\phi_j)\, dxdy, \qquad b_i = \int_{-a}^{a}\int_{-a}^{a} \phi_i f_0\, dxdy. \tag{7.182}$$

For $N = 1$, we have

$$\frac{256}{45} c_1 = \frac{16}{9}\frac{f_0}{a^2}$$

and the one-parameter solution is given by

$$U_1(x, y) - \frac{5f_0 a^2}{16}\left(1 - \frac{x^2}{a^2}\right)\left(1 - \frac{y^2}{a^2}\right). \tag{7.183a}$$

For the two-parameter approximation, we have

$$a^8 \begin{bmatrix} \frac{256}{45} & \frac{1024}{525}a^2 \\ \frac{1024}{525}a^2 & \frac{11264}{4725}a^4 \end{bmatrix} \begin{Bmatrix} c_1 \\ c_2 \end{Bmatrix} = f_0 a^6 \begin{Bmatrix} \frac{16}{9} \\ \frac{32}{45}a^2 \end{Bmatrix}, \tag{a}$$

whose solution yields

$$c_1 = \frac{1295 f_0}{4432a^2}, \qquad c_2 = \frac{525 f_0}{8864a^4}. \tag{b}$$

The two-parameter Galerkin solution of Eq. (7.178) is given by

$$U_2(x, y) = \frac{f_0 a^2}{8864}[2590 + 525(\bar{x}^2 + \bar{y}^2)](1 - \bar{x}^2)(1 - \bar{y}^2), \tag{7.183b}$$

where $\bar{x} = x/a$ and $\bar{y} = y/a$.

For the Ritz method, the same approximation functions as used in the Galerkin method must be used (why?), and we obtain

$$a_{ij} = \int_{-a}^{a}\int_{-a}^{a}\left(\frac{\partial \phi_i}{\partial x}\frac{\partial \phi_j}{\partial x} + \frac{\partial \phi_i}{\partial y}\frac{\partial \phi_j}{\partial y}\right) dxdy$$
$$b_i = \int_{-a}^{a}\int_{-a}^{a} f_0 \phi_i\, dxdy. \tag{7.184}$$

Calculations will show that $a_{ij}$ and $b_i$ are the same as those in Eq. (a). Therefore, the Ritz and Galerkin solutions coincide.

The exact solution to Eq. (7.178) can be obtained using the separation of variables method, and it is given by

$$u(x, y) = \frac{16 f_0 a^2}{\pi^3} \sum_{n=1,3,5\ldots}^{\infty} \frac{1}{n^3}(-1)^{(n-1)/2}\left[1 - \frac{\cosh(n\pi y/2a)}{\cosh(n\pi/2)}\right]\cos\frac{n\pi x}{2a}. \tag{7.185}$$

The exact solution at the center of the region is

$$u_0(0, 0) = 0.2942 f_0 a^2,$$

whereas the two-parameter Galerkin/Ritz solution is $0.2922 f_0 a^2$, which is only 0.68% in error.

Examples 7.17 and 7.18 presented below illustrate the use of variational methods for the solution of problems in two dimensions and time-dependent problems. The approximation is selected such that the parameters $c_i$ are functions of one coordinate (say, time) and $\phi_i$ are functions of the remaining (say, spatial) coordinates. The procedure used in Examples 7.17 and 7.18 to obtain ordinary differential equations from partial differential equation(s) is termed the *semidiscretization method* or *Kantorovich method*. The Lévy method of solution to be discussed in Chapter 8 is also similar to these methods, and all of them are based on the separation of variables concept.

**Example 7.17** Consider the following nondimensionalized partial differential equation (such as the one arising in transient heat transfer):

$$\begin{aligned} &\frac{\partial u}{\partial t} - \frac{\partial^2 u}{\partial x^2} = 0, && 0 < x < 2, \\ &u(0, t) = u(2, t) = 0 && \text{for } t > 0, \\ &u(x, 0) = 1.0 && \text{for } 0 < x < 2. \end{aligned}$$

Owing to the symmetry about $x = 1$, we solve the following equivalent problem:

$$\frac{\partial u}{\partial t} - \frac{\partial^2 u}{\partial x^2} = 0, \qquad 0 < x < 1, \tag{7.186a}$$

$$u(0, t) = 0, \qquad \frac{\partial u}{\partial x}(1, t) = 0 \qquad \text{for } t > 0, \tag{7.186b}$$

$$u(x, 0) = 1.0 \qquad \text{for } 0 < x < 1. \tag{7.187}$$

Let us consider the following form of a one-parameter Galerkin approximation:

$$U_1(x, t) = c_1(t)\phi_1(x) = c_1(t)(2x - x^2). \tag{7.188}$$



The function $\phi_1(x)$, which is a function of $x$ only, satisfies the boundary conditions in Eq. (7.186b). We should determine $c_1(t)$ such that the initial condition (7.187) is satisfied. This requires that

$$c_1(0) = 1. \tag{7.189}$$

Substituting Eq. (7.188) into Eq. (7.186a) and setting up the Galerkin integral, we obtain

$$\begin{aligned} 0 &= \int_0^1 \left(\frac{\partial u_1}{\partial t} - \frac{\partial^2 u_1}{\partial x^2}\right)\phi_1\, dx \\ &= \int_0^1 \left[\frac{dc_1}{dt}(2x - x^2) - c_1(-2)\right](2x - x^2)\, dx \\ &= \frac{8}{15}\frac{dc_1}{dt} + \frac{4}{3}c_1. \end{aligned} \tag{7.190}$$

The solution of Eq. (7.190) is given by

$$c_1(t) = Ae^{-(5/2)t}, \qquad c_1(0) = 1 \;\rightarrow\; A = 1,$$

and the one-parameter Galerkin solution becomes

$$U_1(x, t) = e^{-2.5t}(2x - x^2). \tag{7.191}$$

The exact solution of Eqs. (7.186)–(7.187) is given by

$$u(x, t) = 2\sum_{n=0}^{\infty} \frac{e^{-\lambda_n^2 t}\sin\lambda_n x}{\lambda_n}, \qquad \lambda_n = \frac{(2n+1)\pi}{2}. \tag{7.192}$$

For a two-parameter Ritz approximation, we seek the solution in the form

$$u(x, t) \approx U_2(x, t) = c_1(t)\phi_1(x) + c_2(t)\phi_2(x), \tag{7.193}$$

where $\phi_i$ satisfy only the essential boundary condition, $\phi_i(0) = 0$. Thus the functions

$$\phi_1(x) = x, \qquad \phi_2(x) = x^2 \tag{7.194}$$

are admissible.

We use the weak form (for the Ritz approximation)

$$0 = \int_0^1 \phi_i\left(\frac{\partial u}{\partial t} - \frac{\partial^2 u}{\partial x^2}\right)dx = \int_0^1 \left(\phi_i\frac{\partial u}{\partial t} + \frac{\partial \phi_i}{\partial x}\frac{\partial u}{\partial x}\right)dx. \tag{7.195}$$

Substituting Eqs. (7.193) and (7.194) into Eq. (7.195) and carrying out the integration, we obtain

$$\frac{1}{3}\frac{dc_1}{dt} + \frac{1}{4}\frac{dc_2}{dt} + c_1 + c_2 = 0, \tag{7.196a}$$

$$\frac{1}{4}\frac{dc_1}{dt} + \frac{1}{5}\frac{dc_2}{dt} + c_1 + \frac{4}{3}c_2 = 0, \tag{7.196b}$$

which must be solved subject to the initial condition in Eq. (7.187). It is clear that there is no way in which the initial condition will be satisfied exactly by the selected approximation (7.193). Therefore, we satisfy the initial condition in the Galerkin-integral sense:

$$\int_0^1 [u(x,0) - 1]\phi_i \, dx = 0, \qquad i = 1, 2, \tag{7.197a}$$

which gives

$$\tfrac{1}{3}c_1(0) + \tfrac{1}{4}c_2(0) = \tfrac{1}{2}, \qquad \tfrac{1}{4}c_1(0) + \tfrac{1}{5}c_2(0) = \tfrac{1}{3}. \tag{7.197b}$$

Now we use the Laplace transform method to solve Eqs. (7.196a,b) and (7.197b). Let $\bar{c}_i(s)$ denote the Laplace transform of $c_i(t)$:

$$\bar{c}_i = L[c_i(t)] \equiv \int_{-\infty}^{\infty} c_i(t) e^{-st} \, dt. \tag{7.198}$$

The Laplace transform of the derivative of $c_i(t)$ is given by

$$L\left[\frac{dc_i}{dt}\right] = s\bar{c}_i(s) - c_i(0). \tag{7.199}$$

Using Eq. (7.199), Eqs. (7.196a,b) can be transformed to

$$\left(\tfrac{1}{3}s + 1\right)\bar{c}_1 + \left(\tfrac{1}{4}s + 1\right)\bar{c}_2 = \tfrac{1}{3}c_1(0) + \tfrac{1}{4}c_2(0), \tag{7.200a}$$

$$\left(\tfrac{1}{4}s + 1\right)\bar{c}_1 + \left(\tfrac{1}{5}s + \tfrac{4}{3}\right)\bar{c}_2 = \tfrac{1}{4}c_1(0) + \tfrac{1}{5}c_2(0). \tag{7.200b}$$

The right-hand side can be substituted from Eq. (7.197b), and the resulting equations can be solved for $\bar{c}_1(s)$ and $\bar{c}_2(s)$:

$$\bar{c}_1(s) = \frac{\frac{1}{2}\left(\frac{1}{5}s + \frac{4}{3}\right) - \frac{1}{3}\left(\frac{1}{4}s + 1\right)}{\left(\frac{1}{5}s + \frac{4}{3}\right)\left(\frac{1}{3}s + 1\right) - \left(\frac{1}{4}s + 1\right)^2} = \frac{12s + 240}{3s^2 + 104s + 240}, \tag{7.201a}$$

$$\bar{c}_2(s) = \frac{\frac{1}{3}\left(\frac{1}{3}s + 1\right) - \frac{1}{2}\left(\frac{1}{4}s + 1\right)}{\left(\frac{1}{5}s + \frac{4}{3}\right)\left(\frac{1}{3}s + 1\right) - \left(\frac{1}{4}s + 1\right)^2} = -\frac{10s + 120}{3s^2 + 104s + 240}. \tag{7.201b}$$

The inverse transform of Eqs. (7.201a,b) can be obtained by using the identity

$$L^{-1}\left[\frac{s + a_1}{(s + a)(s + b)}\right] = \frac{a_1 - a}{b - a}e^{-at} + \frac{a_1 - b}{a - b}e^{-bt}. \tag{7.202}$$



Hence we have

$$c_1(t) = 1.6408e^{-32.1807t} + 2.3592e^{-2.486t}, \tag{7.203a}$$

$$c_2(t) = -\left(2.265e^{-32.1807t} + 1.068e^{-2.486t}\right). \tag{7.203b}$$

Equations (7.193) and (7.203a,b) together give the two-parameter Ritz solution.

The two variational solutions in Eqs. (7.191) and (7.193) are compared, for various values of $(x, t)$, with the series solution (7.192) in Table 7.3. The two-parameter Ritz

**Table 7.3 Comparison of the variational solutions with the series solution of a parabolic equation with initial condition, $u(x,0) = 1$ [see Eqs. (7.186)–(7.187)]**

| | | Series Solution | Variational Solution | |
|---|---|---|---|---|
| $t$ | $x$ | Eq. (7.192) | Eq. (7.191) | Eq. (7.193) |
| 0.05 | 0.2 | 0.4727 | 0.3177 | 0.4265 |
| | 0.4 | 0.7938 | 0.5648 | 0.7413 |
| | 0.6 | 0.9418 | 0.7413 | 0.9443 |
| | 0.8 | 0.9880 | 0.8472 | 1.0357 |
| | 1.0 | 0.9965 | 0.8825 | 1.0154 |
| 0.25 | 0.2 | 0.2135 | 0.1927 | 0.2306 |
| | 0.4 | 0.4052 | 0.3426 | 0.4152 |
| | 0.6 | 0.5560 | 0.4496 | 0.5538 |
| | 0.8 | 0.6520 | 0.5139 | 0.6466 |
| | 1.0 | 0.6848 | 0.5353 | 0.6934 |
| 0.50 | 0.2 | 0.1145 | 0.1031 | 0.1238 |
| | 0.4 | 0.2177 | 0.1834 | 0.2230 |
| | 0.6 | 0.2996 | 0.2407 | 0.2975 |
| | 0.8 | 0.3522 | 0.2751 | 0.3473 |
| | 1.0 | 0.3703 | 0.2865 | 0.3725 |
| 0.75 | 0.2 | 0.0617 | 0.0552 | 0.0665 |
| | 0.4 | 0.1174 | 0.0981 | 0.1198 |
| | 0.6 | 0.1616 | 0.1288 | 0.1600 |
| | 0.8 | 0.1900 | 0.1472 | 0.1866 |
| | 1.0 | 0.1997 | 0.1534 | 0.2001 |
| 1.0 | 0.2 | 0.0333 | 0.0296 | 0.0357 |
| | 0.4 | 0.0633 | 0.0525 | 0.0643 |
| | 0.6 | 0.0872 | 0.0690 | 0.0858 |
| | 0.8 | 0.1025 | 0.0788 | 0.1002 |
| | 1.0 | 0.1077 | 0.0821 | 0.1075 |
| 1.50 | 0.2 | 0.0097 | 0.0085 | 0.0103 |
| | 0.4 | 0.0184 | 0.0151 | 0.0186 |
| | 0.6 | 0.0254 | 0.0198 | 0.0248 |
| | 0.8 | 0.0298 | 0.0226 | 0.0289 |
| | 1.0 | 0.0313 | 0.0235 | 0.0310 |

solution is more accurate than the one-parameter Galerkin solution and agrees more closely with the series solution for large values of time.

**Example 7.18** Consider the membrane/torsion problem considered in Example 7.16. Here we seek a one-parameter approximation of the form

$$U_1(x, y) = c_1(x)\phi_1(y) = c_1(x)(y^2 - a^2). \tag{7.204a}$$

Since $u = 0$ on $x = \pm a$ and on $y = \pm a$, it follows that we must determine $c_1(x)$ such that it satisfies the conditions

$$c_1(-a) = c_1(a) = 0. \tag{7.204b}$$

Substituting Eq. (7.204a) into the Galerkin integral ($u \approx U_1$),

$$0 = \int_{-a}^{a}\int_{-a}^{a}\left[-\left(\frac{\partial^2 U_1}{\partial x^2} + \frac{\partial^2 U_1}{\partial y^2}\right) - f_0\right]\phi_1\, dxdy, \tag{7.205a}$$

we obtain

$$\begin{aligned} 0 &= \int_{-a}^{a}\int_{-a}^{a}\left(-\frac{d^2c_1}{dx^2}\phi_1 - c_1\frac{d^2\phi_1}{dy^2} - f_0\right)\phi_1\, dxdy \\ &= -\int_{-a}^{a}\left\{\int_{-a}^{a}\left[(y^2 - a^2)\frac{d^2c_1}{dx^2} + 2c_1 + f_0\right](y^2 - a^2)\, dy\right\} dx. \end{aligned} \tag{7.205b}$$

Performing the integration with respect to $y$ and dividing throughout by the coefficient of $d^2c_1/dx^2$, we obtain

$$0 = \int_{-a}^{a}\left(\frac{d^2c_1}{dx^2} - \frac{5}{2a^2}c_1 - \frac{5f_0}{4a^2}\right)dx. \tag{7.206a}$$

An examination of the integrand in (7.206a) shows that $c_1(x)$ is *not* a periodic function. Hence, the integral vanishes only if the integrand is identically zero:

$$\frac{d^2c_1}{dx^2} - \frac{5}{2a^2}c_1 - \frac{5f_0}{4a^2} = 0. \tag{7.206b}$$



This completes the application of the Galerkin method. We can solve the ordinary differential equation (7.206) either exactly or by an approximate method, such as the Ritz or Galerkin method. We consider the exact solution of Eq. (7.206) subjected to the boundary conditions in Eq. (7.204b).

The exact solution of Eq. (7.206) is given by

$$c_1(x) = A\cosh kx + B\sinh kx - \frac{f_0}{2}, \qquad k = \sqrt{\frac{5}{2a^2}}. \tag{7.207}$$

Using the conditions in Eq. (7.204b), the constants of integration, $A$ and $B$, are evaluated to be

$$A = \frac{f_0}{2\cosh ka}, \qquad B = 0.$$

The solution in Eq. (7.204a) becomes

$$U_1(x, y) = \frac{f_0a^2}{2}\left(1 - \frac{y^2}{a^2}\right)\left[1 - \frac{\cosh kx}{\cosh ka}\right]. \tag{7.208}$$

A two-parameter approximation of the form

$$U_2(x, y) = (y^2 - a^2)[c_1(x) + c_2(x)y^2] \tag{7.209}$$

gives the differential equations

$$\begin{aligned} &\frac{8}{15}a^2\frac{d^2c_1}{dx^2} + \frac{8}{105}a^4\frac{d^2c_2}{dx^2} - \left(\frac{4}{3}c_1 + \frac{4}{15}a^2c_2\right) = \frac{2}{3}f_0 \\ &\frac{8}{105}a^4\frac{d^2c_1}{dx^2} + \frac{8}{315}a^6\frac{d^2c_2}{dx^2} - \left(\frac{4a^2}{15}c_1 + \frac{44}{105}a^4c_2\right) = \frac{2}{15}f_0a^2, \end{aligned} \tag{7.210}$$

with the boundary conditions

$$c_1(-a) = c_1(a) = c_2(-a) = c_2(a) = 0. \tag{7.211}$$

The simultaneous differential equations in Eq. (7.210) can be solved as follows: let $D = d/dx$, $D^2 = d^2/dx^2$, and so on. Then we can write Eq. (7.210) in the operator form

$$\begin{bmatrix} \frac{8a^2}{15}D^2 - \frac{4}{3} & \frac{8a^4}{105}D^2 - \frac{4a^2}{15} \\ \frac{8a^4}{105}D^2 - \frac{4a^2}{15} & \frac{8a^6}{315}D^2 - \frac{44a^4}{105} \end{bmatrix}\begin{Bmatrix} c_1 \\ c_2 \end{Bmatrix} = \frac{2}{15}f_0\begin{Bmatrix} 5 \\ a^2 \end{Bmatrix}. \tag{7.212}$$

Using Cramer's rule, but keeping in mind that $D$'s operate on the quantities in front of them, we obtain

$$L(c_1) = \begin{vmatrix} \frac{2}{3}f_0 & \frac{8a^4}{105}D^2 - \frac{4a^2}{15} \\ \frac{2}{15}f_0a^2 & \frac{8a^6}{315}D^2 - \frac{44a^4}{105} \end{vmatrix}$$

$$= \left(\frac{8}{315}a^6D^2 - \frac{44}{105}a^4\right)\frac{2}{3}f_0 \quad \left(\frac{8}{105}a^4D^2 - \frac{4a^2}{15}\right)\frac{2}{15}f_0a^2,$$

$$= -\frac{384}{1575}f_0a^4, \tag{7.213a}$$

$$L(c_2) = \begin{vmatrix} \frac{8a^2}{15}D^2 - \frac{4}{3} & \frac{2}{3}f_0 \\ \frac{8}{105}a^4D^2 - \frac{4a^2}{15} & \frac{2}{15}f_0a^2 \end{vmatrix} = 0, \tag{7.213b}$$

where $L(\cdot)$ is the determinant of the operator matrix in Eq. (7.212),

$$L = \frac{256a^8}{33075}\left(D^4 - \frac{28}{a^2}D^2 + \frac{63}{a^4}\right). \tag{7.214}$$

The general solutions of Eqs. (7.213a,b) are

$$\begin{aligned} c_1(x) &= -\frac{f_0}{2} + A_1\cosh k_1x + A_2\sinh k_1x + A_3\cosh k_2x + A_4\sinh k_2x, \\ c_2(x) &= B_1\cosh k_1x + B_2\sinh k_1x + B_3\cosh k_2x + B_4\sinh k_2x, \end{aligned} \tag{7.215}$$

where $A_i$ and $B_i (i = 1, 2, 3, 4)$ are constants to be determined, and $k_1^2$ and $k_2^2$ are the roots of the quadratic equation

$$\begin{aligned} k^4 - \frac{28}{a^2}k^2 + \frac{63}{a^4} &= 0, \\ k_1^2 = 14 - \sqrt{133}, \qquad k_2^2 &= 14 + \sqrt{133}. \end{aligned} \tag{7.216}$$

Substituting Eq. (7.215) into the first equation in (7.213a), we obtain the relationships

$$\begin{aligned} \left(\frac{8a^2}{15}k_1^2 - \frac{4}{3}\right)A_i &= \left(\frac{4a^2}{15} - \frac{8}{105}a^4k_1^2\right)B_i, \qquad i = 1, 2, \\ \left(\frac{8a^2}{15}k_2^2 - \frac{4}{3}\right)A_i &= \left(\frac{4a^2}{15} - \frac{8}{105}a^4k_2^2\right)B_i, \qquad i = 3, 4. \end{aligned} \tag{7.217}$$

Use of the boundary conditions in Eq. (7.211) gives $A_2 = A_4 = B_2 = B_4 = 0$, and

$$\begin{aligned} A_1\cosh k_1a + A_3\cosh k_2a &= 0.5f_0, \\ B_1\cosh k_1a + B_3\cosh k_2a &= 0, \end{aligned} \tag{7.218}$$

which can be solved along with Eq. (7.217) for $A_1$, $B_1$, $A_3$, and $B_3$, and we have

$$\begin{aligned} c_1(x) &= 0.5f_0 + 0.516f_0\frac{\cosh k_1x}{\cosh k_1a} - 0.0156f_0\frac{\cosh k_2x}{\cosh k_2a}, \\ c_2(x) &= -0.1138f_0\frac{\cosh k_1x}{\cosh k_1a} + 0.1138f_0\frac{\cosh k_2x}{\cosh k_2a}. \end{aligned} \tag{7.219}$$



**Table 7.4 Comparison of the solution $\bar{u}(x, 0)$ of Eq. (7.178) obtained by the Galerkin/Ritz and semidiscretization methods with the series solution[a]**

| | | Ritz–Galerkin Solution[b] Eqs. (7.183a,b) | | Semidiscretization Eq. (7.209) | |
|---|---|---|---|---|---|
| $x/a$ $(y = 0)$ | Series Solution Eq. (7.185) | One Parameter | Two Parameter | One Parameter | Two Parameter |
| 0.0 | 0.29445 | 0.31250 | 0.29219 | 0.30261 | 0.29473 |
| 0.1 | 0.29194 | 0.30937 | 0.28986 | 0.30014 | 0.29222 |
| 0.2 | 0.28437 | 0.30000 | 0.28278 | 0.29266 | 0.28462 |
| 0.3 | 0.27158 | 0.28437 | 0.27075 | 0.27999 | 0.27177 |
| 0.4 | 0.25328 | 0.26250 | 0.25340 | 0.26180 | 0.25339 |
| 0.5 | 0.22909 | 0.23437 | 0.23025 | 0.23765 | 0.22909 |
| 0.6 | 0.19854 | 0.20000 | 0.20065 | 0.20693 | 0.19839 |
| 0.7 | 0.16104 | 0.15937 | 0.16382 | 0.16886 | 0.16072 |
| 0.8 | 0.11591 | 0.11250 | 0.11884 | 0.12250 | 0.11546 |
| 0.9 | 0.06239 | 0.05937 | 0.06463 | 0.06667 | 0.06203 |
| 1.0 | −0.00038 | 0.00000 | 0.00000 | −0.00447 | 0.00000 |
| | | *Shear stress, $\bar{\sigma}_{yz}(a, 0)$* | | | |
| 1.0 | 0.67500 | 0.62500 | 0.70284 | 0.72636 | 0.66416 |

[a] $\bar{u} = u/q_0a^2$; $\bar{\sigma}_{yz} = \sigma_{yz}/f_0a$.
[b] The Ritz and Galerkin solutions are the same for the same coordinate functions.

Equations (7.219) and (7.208a) together define the two-parameter solution of the torsion problem.

A comparison of the solutions obtained using the Galerkin method and the semidiscretization method (with one and two parameters) with the series solution of the torsion problem is presented in Table 7.4. The table also includes the shearing stress, $\sigma_{yz} = -(\partial u/\partial x)$, at $x = a$ and $y = 0$. It is clear from the results that the two-parameter solutions are more accurate than the one-parameter solutions.

**Example 7.19** *(The Trefftz method)* In all the variational methods discussed in this chapter, the coordinate functions were selected such that they satisfied the boundary conditions of the problem, and the unknown parameters were determined using a variational procedure, such as the minimization of a quadratic functional or setting the weighted residual to zero. An alternative approach that is complementary to the above approach is to select the approximation functions to satisfy the governing differential equation and determine the unknown parameters such that the boundary conditions are satisfied in a variational/integral sense. This method is known as the Trefftz method, and its application is illustrated via the torsion problem of Example 7.16.

The torsion of cylindrical members can be formulated alternatively in terms of the conjugate function $\Psi$, which is related to the Prandtl stress function $\Phi$:

$$\Phi = \Psi - \frac{f_0}{4}(x^2 + y^2) \qquad \text{in } \Omega, \tag{7.220}$$

where $f_0 = 2G\theta$, as in Example 7.16. Consequently, $\Psi$ is governed by the Laplace equation

$$\nabla^2\Psi = 0 \qquad \text{in } \Omega = \{(x, y): -a < x, y < a\}, \tag{7.221}$$

subjected to the boundary condition

$$\Psi = \frac{f_0}{4}(x^2 + y^2) \qquad \text{on } \Gamma. \tag{7.222}$$

Note that a nonhomogeneous equation (i.e., Poisson's equation) with homogeneous boundary conditions is transformed to a homogeneous equation (i.e., Laplace's equation) with nonhomogeneous boundary conditions. For the square domain of Example 7.16, we wish to determine an approximate solution of Eqs. (7.221) and (7.222) using the Trefftz method.

We select an $N$-parameter approximation of the form

$$\Psi \approx \Psi_N = \sum_{j=1}^{N} c_j\phi_j(x, y), \tag{7.223}$$

where $\phi_j$ are selected such that $\Psi_N$ satisfies the governing equation $\nabla^2\Psi_N = 0$ (i.e., $\phi_j$ should be harmonic functions). The parameters $c_j$ are determined by the requirement that the boundary condition (7.222) is satisfied in an integral sense:

$$\begin{aligned} 0 &= \oint_\Gamma \left[\Psi_N - \frac{f_0}{4}(x^2 + y^2)\right]\frac{\partial\phi_i}{\partial n}\,ds \\ &= \oint_\Gamma \left[\sum_{j=1}^{N} c_j\phi_j - \frac{f_0}{4}(x^2 + y^2)\right]\frac{\partial\phi_i}{\partial n}\,ds \\ &\equiv \sum_{j=1}^{N} t_{ij}c_j - b_i, \end{aligned} \tag{7.224}$$

where

$$t_{ij} = \oint_\Gamma \frac{\partial\phi_i}{\partial n}\phi_j\,ds, \qquad b_i = \oint_\Gamma \frac{f_0}{4}(x^2 + y^2)\frac{\partial\phi_i}{\partial n}ds. \tag{7.225}$$

As a specific example, we choose a one-parameter approximation with

$$\phi_1 = x^4 - 6x^2y^2 + y^4, \tag{7.226}$$



which satisfies the equation $\nabla^2\phi_1 = 0$. We have

$$\begin{aligned} t_{11} &= 4\int_{-a}^{a}(4a^3 - 12ax^2)(a^4 - 6a^2x^2 + x^4)\,dx = \frac{1536}{35}a^8, \\ b_1 &= -\int_{-a}^{a} f_0(-12x^2a + 4a^3)(x^2 + a^2)\,dx = -\frac{64}{30}a^6 f_0, \end{aligned} \tag{7.227}$$

and the solution is given by

$$\Psi_1 = -\frac{7f_0}{144a^2}(x^4 - 6x^2y^2 + y^4). \tag{7.228}$$

The Prandtl stress function $\Phi$ becomes

$$\Phi = \Psi - \frac{f_0}{4}(x^2 + y^2) = -\frac{7f}{144a^2}(x^4 - 6x^2y^2 + y^4) - \frac{f}{4}(x^2 + y^2). \tag{7.229}$$

The maximum shear stress $\sigma_{yz}$ is

$$\sigma_{yz}(a, 0) = -\left(\frac{\partial\Phi_1}{\partial x}\right)_{(a,0)} = \frac{5}{6}f_0a = 0.833\,f_0a.$$

The maximum shear stress is about 23.4% greater than the exact solution.

The Trefftz method has limited use, because it can only be utilized for Dirichlet-type boundary-value problems (i.e., boundary-value problems in which the function is specified on the boundary). Furthermore, it is not easy to find approximation functions that satisfy the governing differential equations. Some of the hybrid finite element models are based on ideas similar to the Trefftz method.

**Example 7.20** Consider the nonlinear differential equation

$$-\frac{d}{dx}\left(u\frac{du}{dx}\right) + 1 = 0, \qquad 0 < x < 1, \tag{7.230a}$$

subject to the boundary conditions

$$u(1) = \sqrt{2}, \qquad \left(\frac{du}{dx}\right)\Big|_{x=0} = 0. \tag{7.230b}$$

We wish to determine a one-parameter Ritz solution to the problem.

The weak form of Eq. (7.230a) is given by

$$0 = \int_0^1 \left[u\frac{dw}{dx}\frac{du}{dx} + w \cdot 1\right]dx.$$

The boundary term is zero because $w(1) = 0$ and $(du/dx)(0) = 0$.

Let

$$u \approx U_1 = c_1\phi_1 + \phi_0, \qquad \text{with } \phi_0 = \sqrt{2}, \qquad \phi_1 = 1 - x.$$

Substituting into the weak form, we obtain

$$0 = \int_0^1 \left\{ \left[ c_1(1-x) + \sqrt{2} \right] (-1)(-c_1) + (1-x) \right\} dx$$
$$= c_1 \left[ c_1 \cdot \tfrac{1}{2} + \sqrt{2} \right] + \tfrac{1}{2}$$

or

$$c_1^2 + 2\sqrt{2}c_1 + 1 = 0, \qquad (c_1)_{1,2} = -\sqrt{2} \pm \sqrt{2-1} = -\sqrt{2} \pm 1.$$

Thus, there are two approximate solutions to the nonlinear problem. We must choose one value of $c_i$ using some meaningful criterion. We shall take the value of $c_1$ that yields the smallest integral of the residual in the differential equation:

$$\int_0^1 \left[ -\frac{d}{dx}\left( u \frac{du}{dx} \right) + 1 \right] dx = \left( c_1^2 + 1 \right).$$

Clearly, the smaller (in absolute value) root gives the smaller value of the residual. Hence, we choose $(c_1)_1 = -\sqrt{2} + 1$. The solution becomes

$$U_1 = \left(1 - \sqrt{2}\right)(1-x) + \sqrt{2} = 1 + \left(\sqrt{2} - 1\right)x.$$

The exact solution is $u(x) = \sqrt{1+x^2}$. At $x = 0$ the approximate solution matches with the exact.

## 7.6 SUMMARY

In this chapter, the Ritz and weighted-residual (e.g., Galerkin, least-squares, and collocation) methods were presented and their application to simple problems was illustrated. The Ritz method makes use of the weak form provided by the principle of virtual displacements, the principles of minimum total potential energy, or the one developed from the governing equations of the problem as discussed in Sections 7.4.1 and 7.5.1. The key feature of the weak form is that it includes the governing equation(s) as well as the natural boundary condition(s) of the problem. Hence, the Ritz approximation is not required to satisfy the natural boundary conditions. All weighted-residual methods, except the least-squares method, are based on an weighted-integral statement of the governing equation(s), whereas the least-squares method is based on the minimization of the square of the governing equation(s). Thus, the integral statements used in all weighted-residual methods do not include any boundary conditions as a part of the statements. Hence, the approximations chosen for the weighted-residual methods are required to satisfy *all* (both natural and essential) specified boundary conditions of the problem. Consequently, the approximation functions are of higher order than those used in the Ritz method. The Ritz as well as the weighted-residual methods may be used for linear and nonlinear differential equations. However, in the case of the least-squares method, there are some limitations, as discussed in Section 7.5.5. Although the least-squares method can be interpreted as a special case of the weighted-residual methods for linear static problems, it is based on an integral statement whose Euler equations, if derived, are *not* the same as the governing equations. Thus, the least-squares method is quite different from the other weighted-residual methods. Overall, the Ritz method is the most efficient method, especially for solid and structural mechanics problems.



The single most difficult step in using the variational methods presented in this chapter is the selection of the approximation functions. The requirements (7.62) and (7.164) on the approximation functions merely provide the guidelines for their selection. The selection of the approximation functions becomes even more difficult for problems with irregular domains (i.e., noncircular and nonrectangular) or discontinuous data (loading as well as geometry). Further, the generation of coefficient matrices for the resulting algebraic equations cannot be automated for a *class* of problems that differ from each other only in the geometry of the domain, boundary conditions, or loading. These limitations of the classical variational methods can be overcome by representing a given domain as a collection of geometrically simple subdomains for which we can systematically generate the approximation functions (see Example 7.8). One such technique, namely, the finite element method, is discussed in Chapter 9. The finite element method is based on ideas similar to the classical variational methods, especially in developing the system of algebraic equations for the unknown coefficients, but the method views a given domain as a collection of conveniently chosen subdomains that allow a systematic generation of the approximation functions.

## EXERCISES

**7.1** Let $\mathcal{P}$ be the vector space of all polynomials with real coefficients. Determine which of the following subsets of $\mathcal{P}$ are subspaces:

(a) $S_1 = \{p(x)\colon \quad p(1) = 0\}$.
(b) $S_2 = \{p(x)\colon \quad \text{degree of } p(x) = 3\}$.
(c) $S_3 = \{p(x)\colon \quad \text{degree of } p(x) \leq 3\}$.
(d) $S_4 = \{p(x)\colon \quad \text{constant term is zero}\}$.

**7.2** Determine which of the following sets of vectors in $\Re^3$ are linearly independent over the real number field $\Re$:

(a) $\{(-1, 1, 0), (-1, 1, 1), (-2, -1, 1), (1, 1, 1)\}$.
(b) $\{(1, 0, 0), (1, 1, 0), (1, 1, 1)\}$.
(c) $\{(1, 1, 1), (1, 2, 3), (2, -1, 1)\}$.

**7.3** Determine if the following sets of functions are linearly independent.

(a) $\{\sin(n\pi x/L)\}_{n=1}^{3}$, $0 \leq x \leq L$.
(b) $\{x^n(1-x)\}_{n=0}^{3}$, $0 \leq x \leq 1$.
(c) $\{1 + x + x^2, 1 + 2x + 3x^2, 2 - 3x + x^3\}$, $0 \leq x \leq 1$.

**7.4** Compute the $L_2$ norm and the sup-norm of the following functions in the interval indicated:

(a) $u(x) = \sin \pi x - x$, on $0 \le x \le 1$.

(b) $u(x) = x^{1/3}$, on $0 \le x \le 1$.

(c) $u(x) = \cos \pi x + 2x - 1$, on $0 \le x \le 1$.

(d) $u(x) = \sqrt{1 + x^2}$, on $0 \le x \le 1$.

(e) $u(x) = \begin{cases} 100 \sin 100\pi x, & 0 \le x \le 0.01 \\ 0, & 0.01 \le x \le 1. \end{cases}$

**7.5** Prove the following relations in a real inner product space:

(a) Parallelogram law: $\|u + v\|^2 + \|u - v\|^2 = 2(\|u\|^2 + \|v\|^2)$.

(b) $(u, v) = \frac{1}{4}[\|u + v\|^2 - \|u - v\|^2]$.

(c) $|\|u\| - \|v\|| \le \|u - v\|$.

**7.6** Let $\mathbf{u} = \mathbf{u}(x)$ and $\mathbf{v} = \mathbf{v}(x)$ be two vector functions of $x$. Which one of the following products qualifies as an inner product?

(a) $(\mathbf{u}, \mathbf{v}) \equiv \int_0^L \mathbf{u} \cdot \mathbf{v}\, dx$.

(b) $(\mathbf{u}, \mathbf{v}) \equiv \int_0^L \mathbf{u}(L - x) \cdot \mathbf{v}(x)\, dx$.

**7.7** Compute the inner product of the following pairs of functions on the interval indicated. Use the $L_2$-inner product and the $H^1$-inner product.

(a) $u = x - x^2$, $v = \sin \pi x$, on $0 \le x \le 1$.

(b) $u = (1 + x)$, $v = 3x^2 - 1$, on $-1 \le x \le 1$.

(c) $u = \sin \pi x$, $v = \cos \pi x$, on $0 \le x \le 1$.

(d) $u = \sin \pi x$, $v = a + bx + cx^2$, on $0 \le x \le 1$.

(e) $u = \sin \pi x \sin \pi y$, $v = (1 - x^2 - y^2)$, on $0 \le x, y \le 1$.

(f) $u = (x^2 - a^2)(y^2 - b^2)$, $0 \le x \le a$ and $0 \le y \le b$.

**7.8** Find the distance between the following pair of functions using the inner product in Eq. (7.23a): $u = x^3 - 3x + 2$ and $v = (x - 1)^2$, $0 \le x \le 1$.

**7.9** Check whether the following pair of functions are orthogonal in the $L_2(0, 1)$-space: $u(x) = 2 + 3x^2 - x$ and $v(x) = \frac{1}{3} + 3x - 5x^2$.

**7.10** Check whether the pair of functions in Exercise 7.9 are orthogonal in the $H^1(0, 1)$-space.

**7.11** Determine the constants $a$ and $b$ such that $w(x) = a + bx + 3x^2$ is orthogonal in $L_2(0, 1)$ to both $u(x)$ and $v(x)$ of Exercise 7.9.

**7.12** Determine $C$ if the indicated pairs of vectors are orthogonal in $\mathfrak{R}^4$:

(a) $(1, C - 1, 2, 1 + C)$, $(2C, 4, C, 1)$.

(b) $(2, 1 - C, 4, 3C)$, $(C, 1 + C, 3C, -7)$.



**7.13** Determine a vector in the space $\mathcal{P}$ of polynomials of degree 2 such that the vector is orthogonal to the polynomials $p_1 = 1 + x - 2x^2$ and $p_2 = -2 + 4x + x^2$ in $L_2(1,0)$.

**7.14** Which of the following operators qualify as linear operators?

(a) $A(u) = -\dfrac{d}{dx}\left(a\dfrac{du}{dx}\right)$.

(b) $T(u) = \nabla^2 u + 1$.

(c) $I(u) = \int_0^a K(x)\dfrac{du}{dx}\, dx - u(0)$.

**7.15** If $T_1$ and $T_2$ are linear operators defined on $\mathfrak{R}^3$ by $T_1(\mathbf{x}) = (0, x_1 - x_3, x_2 + x_3)$, $T_2(\mathbf{x}) = (x_2 - x_3, 2x_1 - x_2, x_1)$, determine:

(a) $T_1 + T_2$.

(b) $T_1 T_2$.

(c) $T_2 T_1$.

**7.16** Find the matrix representing the linear transformation from $\mathfrak{R}^4$ into $\mathfrak{R}^3$ with respect to the standard bases of $\mathfrak{R}^4$ and $\mathfrak{R}^3$:

$$T(x_1, x_2, x_3, x_4) = (x_1 + 2x_2, 2x_2 + x_3, x_3 + x_1).$$

**7.17** Identify the linear and bilinear functionals:

(a) $I(u, v) = \int_\Omega \nabla u \cdot \nabla v\, dx dy$.

(b) $I(u, v) = \displaystyle\int_0^a \left(b\frac{d^2u}{dx^2}\frac{d^2v}{dx^2} + fv\right) dx$.

**7.18** Determine which of the following operators represent bilinear forms:

(a) $B$: $\mathfrak{R}^2 \times \mathfrak{R}^2 \to \mathfrak{R}$, $B(\mathbf{x}, \mathbf{y}) = (x_1 + y_1)^2 - (x_1 - y_1)^2$.

(b) $B$: $\mathfrak{R}^2 \times \mathfrak{R}^2 \to \mathfrak{R}$, $B(\mathbf{x}, \mathbf{y}) = x_1 y_2 - x_2 y_1$.

**7.19** Consider the functional

$$B(u, v) = \tfrac{1}{4}(\|u + v\|^2 - \|u - v\|^2)$$

on an inner product space. Show that:

(a) $B(\cdot, \cdot)$ is bilinear in $u$ and $v$.

(b) For fixed $u$, $f_u(v) = B(u, v)$ is a (bounded) linear functional.

(c) $B(u, v)$ satisfies the axioms of an inner product.

**7.20** If $I(u)$ is a quadratic functional, show that (a) its first variation is a bilinear functional of $u$ and $\delta u$, and (b) its first variation is a linear functional of $\delta u$.

Give admissible approximation functions, either algebraic or trigonometric, for a two-parameter Ritz approximation of problems in Exercises 7.21–7.27. Assume

that the total potential energy principle or weak form is used to construct the Ritz approximation.

**7.21** A cable suspended between points $A : (0, 0)$ and $B : (L, h)$ and subjected to uniformly distributed transverse load of intensity $f_0$.

**7.22** A cantilever beam subjected to uniformly distributed load of intensity $q_0$.

**7.23** The symmetric half of the simply supported beam problem considered in Example 7.7 ($0 \le x \le L/2$).

**7.24** A beam clamped at the left end and simply supported at the right end, and subjected to point load $F_0$ at $x = L/2$.

**7.25** A simply supported beam with a spring support at $x = L/2$, and subjected to uniformly distributed load of intensity $q_0$.

**7.26** A square elastic membrane fixed on all its sides and subjected to a uniformly distributed load of intensity $f_0$.

**7.27** A quadrant model (because of the biaxial symmetry) of the membrane problem of Exercise 7.26.

**7.28–7.34** Find the two-parameter Ritz approximation using algebraic polynomials of the problems in Exercises 7.21–7.27 and compare the solutions with the exact solutions when possible.

**7.35** Find the first two natural frequencies of a cantilever beam. Take $EI = (a + bx)^{-1}$ where $a$ and $b$ are constants.

**7.36** Find a two-parameter Ritz approximation of the transverse deflection of a simply supported beam on an elastic foundation that is subjected to uniformly distributed load. Use (a) algebraic polynomials and (b) trigonometric polynomials.

**7.37** Derive the matrix equations corresponding to the $N$-parameter Ritz approximation

$$W_N = c_1 x^2 + c_2 x^3 + \cdots + c_N x^{N+1}$$

of a cantilever beam with a uniformly distributed load, $q_0$. Compute $a_{ij}$ and $b_j$ in explicit form in terms of $i$, $j$, $L$, $EI$, and $q_0$.

**7.38** Use a two-parameter Ritz approximation with trigonometric functions to determine the critical buckling load $P$ of a simply supported beam.

**7.39** Consider the buckling of a uniform beam according to the Timoshenko beam theory. The total potential energy functional for the problem can be written as

$$\Pi(w_0, \phi_x) = \frac{1}{2}\int_0^L \left[ D\left(\frac{d\phi_x}{dx}\right)^2 + S\left(\frac{dw_0}{dx} + \phi_x\right)^2 - N\left(\frac{dw_0}{dx}\right)^2 \right] dx,$$

where $w_0(x)$ is the transverse deflection, $\phi_x$ is the rotation, $D$ is the flexural stiffness, $S$ is the shear stiffness, and $N$ is the axial compressive load. We wish to determine the critical buckling load $N_{cr}$ of a simply supported beam using



the Ritz method. Assume a one-parameter approximation of $w_0$ and $\phi_x$ and determine the critical buckling load.

**7.40** Determine the $N$-parameter Ritz solution for the transient response of a simply supported beam under step loading $q(x, t) = q_0 H(t - t_0)$, where $H(t)$ denotes the Heaviside step function. Use trigonometric functions for $\phi_i(x)$.

**7.41** Show that the two-parameter Ritz solution for the transient response of the bar considered in Example 7.6 yields the equations

$$\begin{bmatrix} \frac{1}{3} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{5} \end{bmatrix} \begin{Bmatrix} \ddot{c}_1 \\ \ddot{c}_2 \end{Bmatrix} + \begin{bmatrix} 2 & 2 \\ 2 & \frac{7}{3} \end{bmatrix} \begin{Bmatrix} c_1 \\ c_2 \end{Bmatrix} = \begin{Bmatrix} \frac{1}{2} \\ \frac{1}{3} \end{Bmatrix}. \tag{a}$$

Use the Laplace transform method to determine the solution of these equations.

**7.42** Derive the weak forms of the following nonlinear equations:

$$-\frac{dN_{xx}}{dx} = f(x), \qquad 0 \le x \le L, \tag{a}$$

$$-\frac{d}{dx}\left(\frac{dw_0}{dx} N_{xx}\right) - \frac{d^2 M_{xx}}{dx^2} = q(x), \qquad 0 \le x \le L, \tag{b}$$

where

$$N_{xx} = EA\left[\frac{du_0}{dx} + \frac{1}{2}\left(\frac{dw_0}{dx}\right)^2\right], \qquad M_{xx} = -EI\frac{d^2 w_0}{dx^2}.$$

**7.43** Consider a uniform beam fixed at one end and supported by an elastic spring (spring constant $k$) in the vertical direction. Assume that the beam is loaded by a uniformly distributed load $q_0$. Determine the one-parameter Ritz solution using algebraic functions.

**7.44** Consider the problem of finding the fundamental frequency of a circular membrane of radius $a$, fixed at its edge. The governing equation for axisymmetric vibration is

$$-\frac{1}{r}\frac{d}{dr}\left(r\frac{du}{dr}\right) - \lambda u = 0, \qquad 0 < r < a,$$

where $\lambda$ is the frequency parameter and $u$ is the deflection of the membrane. (a) Construct the weak form, (b) use one-parameter Ritz approximation to determine $\lambda$, and (c) use two-parameter Ritz approximation to determine $\lambda$. Select trigonometric functions for $\phi_i(r)$.

**7.45** Consider the problem of finding the solution of the equation

$$\frac{d^2 u}{dx^2} + u + x = 0, \qquad 0 < x < 1; \qquad u(0) = u(1) = 0.$$

(a) Develop the weak form, (b) assume $N$-parameter Ritz approximation of the form

$$U_N(x) = x(1-x)\left(c_1 + c_2 x + \cdots + c_N x^{N-1}\right)$$

and obtain the Ritz equations for the unknown coefficients, and (c) determine the two-parameter solution and compare it with the exact solution

$$u(x) = \frac{\sin x}{\sin 1} - x.$$

**7.46** Consider the Bessel equation

$$x^2\frac{d^2u}{dx^2} + x\frac{du}{dx} + (x^2-1)u = 0, \qquad 1 < x < 2$$

subject to the boundary conditions $u(1) = 1$ and $u(2) = 2$. Assume $u = w + x$ and reduce the equation to

$$-\frac{d}{dx}\left(x\frac{dw}{dx}\right) - \frac{x^2-1}{x}w = x^2, \qquad 1 < x < 2,$$

subject to the boundary conditions $w(1) = w(2) = 0$. Determine a one-parameter Ritz approximation of the problem and compare it with the analytical solution

$$u(x) = 3.6072J_1(x) + 0.75195Y_1(x),$$

where $J_1$ and $Y_1$ are Bessel functions of the first and second kind, respectively.

**7.47** Derive the weak form and obtain a one-parameter Ritz solution of the problem

$$-\left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}\right) = 1 \text{ in a unit square},$$
$$u(1, y) = u(x, 1) = 0,$$
$$\frac{\partial u}{\partial x}(0, y) = \frac{\partial u}{\partial y}(x, 0) = 0.$$

The origin of the coordinate system is taken at the lower left corner of the unit square.

**7.48** Determine the two-parameter Petrov–Galerkin solution (with $\psi_1 = 1$ and $\psi_2 = x$) of the following differential equation and compare, if possible, with the exact solution:

$$-\frac{d^2u}{dx^2} = \frac{1}{1+x}, \qquad 0 < x < 1; \qquad u(0) = u'(1) = 0.$$



**7.49** Determine the two-parameter Petrov–Galerkin solution (with $\psi_1 = 1$ and $\psi_2 = x$) of the following differential equation:

$$-\frac{d}{dx}\left[(1+x^2)\frac{du}{dx}\right] + u = \sin \pi x + 3x - 1, \quad 0 < x < 1; \quad u(0) = u'(1) = 0.$$

**7.50** Determine the one-parameter Galerkin solution of the equation

$$\frac{d^2}{dx^2}\left[\left(2 + \frac{x}{L}\right)\frac{d^2w_0}{dx^2}\right] + kw_0 = q_0\frac{x}{L}$$

that governs a cantilever beam on an elastic foundation and subjected to linearly varying load (from zero at the free end to $q_0$ at the fixed end). Take $k = L = 1$ and $q_0 = 3$, and use algebraic polynomials.

**7.51** Determine a two-parameter Petrov–Galerkin solution (with $\psi_1 = 1$ and $\psi_2 = xy$) of the problem

$$\begin{aligned} -\nabla^2 u &= 0 && \text{in } 0 < (x, y) < 1, \\ u &= \sin \pi x && \text{on } y = 0, \\ u &= 0 && \text{on all other sides.} \end{aligned}$$

**7.52** Solve the nonlinear differential equation

$$-\frac{d}{dx}\left(u\frac{du}{dx}\right) + 1 = 0, \qquad 0 < x < 1,$$
$$u(1) = \sqrt{2}, \qquad u'(0) = 0.$$

Use a one-parameter Petrov–Galerkin approximation with (a) $\psi_1 = 1$ and (b) $\psi_1 = x$. Compare the results with the exact solution, $u(x) = \sqrt{1+x^2}$. See Example 7.20.

**7.53** Find the first two eigenvalues associated with the differential equation

$$-\frac{d^2u}{dx^2} = \lambda u, \qquad 0 < x < 1,$$
$$u(0) = 0, \qquad u(1) + u'(1) = 0.$$

Use the least-squares method. Use the operator definition $A = -(d^2/dx^2)$ to avoid increasing the degree of the characteristic polynomial for $\lambda$.

**7.54** Solve the Poisson equation

$$\begin{aligned} -\nabla^2 u &= f_0 && \text{in a unit square}, \\ u &= 0 && \text{on the boundary}. \end{aligned}$$

using the following $N$-parameter Galerkin approximation:

$$U_N = \sum_{i,j=1}^{N} c_{ij} \sin i\pi x \sin j\pi y.$$

**7.55** Solve the nonlinear equation in Exercise 7.52 by the Galerkin method.

**7.56** Find a two-parameter Galerkin solution of a clamped (at both ends) beam under uniformly distributed load.

**7.57** Solve the equation in Exercise 7.49 using the least-squares method.

**7.58** Solve the problem of Exercise 7.52 using the least-squares method.

**7.59** Solve the equation in Exercise 7.53 using the least-squares method.

**7.60** Solve the equation in Exercise 7.54 using the least-squares method.

**7.61** Consider a cantilever beam of variable flexural rigidity, $EI = a_0[2-(x/L)^2]$, and carrying a distributed load, $q = q_0[1 - (x/L)]$. Find a three-parameter solution using the collocation method.

**7.62** Repeat Exercise 7.52 using the collocation method.

**7.63** Solve the differential equation in Exercise 7.53 by the collocation method.

**7.64** Solve the problem in Exercise 7.54 using the one-point collocation method.

**7.65** Consider the Laplace equation

$$-\nabla^2 u = 0, \qquad 0 < x < 1, \qquad 0 < y < \infty,$$
$$u(0, y) = u(1, y) = 0 \quad \text{for } y > 0,$$
$$u(x, 0) = x(1 - x), \qquad u(x, \infty) = 0, \qquad 0 \le x \le 1.$$

Assuming an approximation of the form

$$u(x, y) = c_1(y)x(1 - x),$$

find the differential equation for $c_1(y)$ and solve it exactly.

**7.66** Use the semidiscretization method to find a two-parameter approximation of the form

$$u(x, y) = c_1(y) \cos \frac{\pi x}{2a} + c_2(y) \cos \frac{3\pi x}{2a},$$

to determine an approximate solution of the torsion problem in Example 7.16.

**7.67** Use the semidiscretization method to find a two-parameter approximation of the form

$$w(x, t) = c_1(t)(1 - \cos 2\pi x),$$



to determine an approximate solution of the equation

$$\frac{\partial^4 w}{\partial x^4} + \frac{\partial^2 w}{\partial t^2} = 0, \qquad 0 < x < 1, \qquad t > 0,$$
$$w = \frac{\partial w}{\partial x} = 0 \qquad \text{at } x = 0, 1 \text{ and } t > 0,$$
$$w = \sin \pi x - \pi x(1 - x), \qquad \frac{\partial w}{\partial t} = 0 \text{ at } t = 0, \quad 0 < x < 1.$$

Use the Galerkin method to satisfy the initial conditions.

**7.68** Consider the problem of finding the eigenvalues associated with the equation

$$\nabla^2 u + \lambda u = 0 \quad \text{in } \Omega; \qquad u = 0 \quad \text{on } \Gamma,$$

where $\Omega$ is a rectangle, $-a < x < a$ and $-b < y < b$, and $\Gamma$ is its boundary (see Example 7.18). Assuming approximation of the form

$$U_1 = c_1(x)\phi_1(y) = c_1(x)(y^2 - b^2),$$

determine the first eigenvalue $\lambda_1$.

**7.69** Repeat Exercise 7.68 with a two-parameter approximation of the form

$$U_2 = c_1(x)\phi_1(y) + c_2(x)\phi_2(y) = [c_1(x) + c_2(x)y^2](y^2 - b^2).$$

**7.70** Solve the problem in Example 7.19 using a two-parameter approximation of the form (which is more complete than the one-parameter approximation used in Example 7.19, although it has no effect on the derivative of the solution)

$$\Psi_2 = c_1 + c_2(x^4 - 6x^2y^2 + y^4).$$

## REFERENCES


1. Oden, J. T., and Reddy, J. N., *Variational Methods in Theoretical Mechanics*, 2nd edition, Springer-Verlag, Berlin (1982).
2. Oden, J. T., and Ripperger, E. A., *Mechanics of Elastic Structures*, 2nd edition, Hemisphere, New York (1981).
3. Reddy, J. N., *Energy and Variational Methods in Applied Mechanics*, John Wiley, New York (1984).
4. Reddy, J. N., *Applied Functional Analysis and Variational Methods in Engineering*, McGraw-Hill, New York, 1986; reprinted by Krieger, Melbourne, FL (1992).
5. Reddy, J. N., *An Introduction to the Finite Element Method*, 2nd edition, McGraw-Hill, New York (1993).
6. Reddy, J. N., *Theory and Analysis of Elastic Plates*, Taylor & Francis, Philadelphia (1999).
7. Reddy, J. N., and Rasmussen, M. L., *Advanced Engineering Analysis*, John Wiley, New York (1982); reprinted by Krieger, Melbourne, FL (1991).

8. Washizu, K., *Variational Methods in Elasticity and Plasticity*, 3rd edition, Pergamon Press, New York (1982).
9. Timoshenko, S. P., and Woinowsky-Krieger, S., *Theory of Plates and Shells*, McGraw-Hill, Singapore (1970).
10. Lanczos, C., *The Variational Principles of Mechanics*, University of Toronto Press, Toronto (1964).
11. Langhaar, H. L., *Energy Methods in Applied Mechanics*, John Wiley, New York (1962).
12. Mikhlin, S. G., *Variational Methods in Mathematical Physics*, translated from the 1957 Russian edition by T. Boddington, Macmillan, New York (1964).
13. Mikhlin, S. G., *The Problem of the Minimum of a Quadratic Functional*, translated from the 1952 Russian edition by A. Feinstein, Holden-Day, San Francisco (1965).
14. Mikhlin, S. G., *An Advanced Course of Mathematical Physics*, Elsevier, New York (1970).
15. Rektorys, K., *Variational Methods in Mathematics, Science and Engineering*, 2nd edition, Reidel, Boston (1980).
16. Shaw, F. S., *Virtual Displacements and Analysis of Structures*, Prentice-Hall, Englewood Cliffs, NJ (1972).
17. Dym, C. L., and Shames, I. H., *Solid Mechanics: A Variational Approach*, McGraw-Hill, New York (1973).
18. Charlton, T. M., *Energy Principles in Theory of Structures*, Oxford University Press, Oxford (1973).
19. Richards, T. H., *Energy Methods in Stress Analysis*, Ellis Horwood (distributed by John Wiley & Sons), Chichester, UK (1977).
20. Davies, G. A. O., *Virtual Work in Structural Analysis*, John Wiley, Chichester, UK (1982).
21. McGuire, W., Gallagher, R. H., and Ziemian, R. D., *Matrix Structural Analysis*, 2nd edition, John Wiley, New York (2000).
22. Kantorovich, L. V., "Some remarks on Ritz's method" (in Russian), *Tr. Vyssh. Voen. Morsk. Inzh. Stroit Uchil.*, No. 3 (1941).
23. Kantorovitch, L. V., and Krylov, V. I., *Approximate Methods of Higher Analysis*, translated by C. D. Benster, Noordhoff, Groningen (1958).
24. Galerkin, B. G., "Series-solutions of some cases of equilibrium of elastic beams and plates" (in Russian), Vest. Inzh. Teck., **1**, 897–903 (1915).
25. Galerkin, B. G., "Berechung der frei gelagerten elliptschen Platte auf Biegung," *Z. Agnew. Math. Mech.*, **3**(2), 113–117 (1923).
26. Rayleigh, Lord (John William Strutt), "Some general theorems relating to vibrations," *Proc. Lond. Math. Soc.*, **4**, 357–368 (1873).
27. Rayleigh, Lord (John William Strutt), *Theory of Sound*, 1st edition (1877); 2nd revised edition, Dover, New York (1945).
28. Ritz, W., "Über eine neue Methode zur Lösung gewisser Variationsprobleme der mathematischen Physik," *Journal für reine und angewandte Mathematik*, **135**, 1–61 (1908).
29. Trefftz, E., "Zur Theorie der Stabilität des elastischen Gleichgewichts," *Z. Angew. Math. Mech.*, **13**, 160–165 (1923).
30. Trefftz, E., "Ein Gegenstück zum Ritzschen Verfahren," *Proc. 2nd Int. Congr. Appl. Mech.*, E. Meissner (ed.), Zürich, pp. 131–137 (1926).

# 8

# THEORY AND ANALYSIS OF PLATES

## 8.1 INTRODUCTION

### 8.1.1 General Comments

In this chapter we study bending, buckling, and natural vibration of plate structures. A *plate* is a flat structural element with planform dimensions much larger than its thickness and is subjected to loads that cause bending deformation in addition to stretching (see Fig. 8.1). Street manhole covers, table tops, side panels and roofs of buildings and transportation vehicles, glass window panels, turbine disks, bulkheads, and tank bottoms provide familiar examples of plate structures. A *shell* is initially a curved structural element with thickness much smaller than the other dimensions. Like a plate, a shell is subjected to loads that cause stretching and bending deformations. Examples of shell structures are provided by pressure vessels, pipes, curved panels of a variety of structures including automobiles and aerospace vehicles, tires, and roof domes and sheds.

In most cases, the thickness of plate and shell structures is about one-tenth or less of the smallest in-plane (i.e., within the plane) dimension. When the thickness is one-twentieth of an in-plane dimension or less, they are termed *thin*; otherwise they are said to be *thick*. Because of the smallness of the thickness dimension, it is often not necessary to model plate and shell structures using 3D elasticity equations. Simple 2D theories that are based on some kinematic assumptions can be developed to study deformation, stresses, natural frequencies, and global buckling loads of plate and shell structures. In the present study, the governing equations of the 2D theories are derived using the principle of virtual displacements for circular plates in cylindrical coordinates and for noncircular plates in rectangular Cartesian coordinates. Bending, buckling, and vibration solutions are obtained for particular cases of