Chapter 4

VARIATIONAL METHODS

4.1 Introduction to variational principles

Certain classes of partial differential equations possess a variational structure. This means that their solutions u can be interpreted as extremal points over a properly defined function space $\mathcal{U}$, with reference to given functionals $I[u]$. By way of introduction to variational methods, consider a functional $I[u]$ defined as

$$I[u] \;=\; \int_\Omega \big[\frac{k}{2}\left(\frac{\partial u}{\partial x_1}\right)^2 \;+\; \frac{k}{2}\left(\frac{\partial u}{\partial x_2}\right)^2 + fu\big]\, d\Omega \;, \tag{4.1}$$

where $k = k(x_1, x_2) > 0$ and $f = f(x_1, x_2)$ are continuous functions in Ω. In addition, assume that the domain Ω possesses a smooth boundary $\partial\Omega$ with uniquely defined outward unit normal $\mathbf{n}$.

The functional $I[u]$ attains an extremum if, and only if, its first variation vanishes, namely

$$\begin{aligned}\delta I[u] \;&=\; \int_\Omega \big[k\,\frac{\partial u}{\partial x_1}\delta\left(\frac{\partial u}{\partial x_1}\right) \;+\; k\,\frac{\partial u}{\partial x_2}\delta\left(\frac{\partial u}{\partial x_2}\right) \;+\; f\delta u\big]\, d\Omega \\ &=\; \int_\Omega \big[\frac{\partial u}{\partial x_1}k\,\frac{\partial \delta u}{\partial x_1} \;+\; \frac{\partial u}{\partial x_2}k\,\frac{\partial \delta u}{\partial x_2} \;+\; f\delta u\big]\, d\Omega \;=\; 0\;,\end{aligned} \tag{4.2}$$

where u is assumed continuously differentiable. Following the developments of Section 3.2, integration by parts and application of the divergence theorem on (4.2) yields

$$\delta I[u] \;=\; \int_{\partial\Omega} \big[k\left(\frac{\partial u}{\partial x_1}n_1 \;+\; \frac{\partial u}{\partial x_2}n_2\right)\delta u\big]\, d\Gamma - \int_\Omega \big[\frac{\partial}{\partial x_1}(k\frac{\partial u}{\partial x_1}) + \frac{\partial}{\partial x_2}(k\frac{\partial u}{\partial x_2}) - f\big]\delta u\, d\Omega \;=\; 0\,. \tag{4.3}$$

Recalling that $\frac{\partial u}{\partial x_1} n_1 + \frac{\partial u}{\partial x_2} n_2 = \frac{\partial u}{\partial n}$, write

$$\delta I[u] \;=\; \int_{\partial\Omega} k \frac{\partial u}{\partial n} \delta u \, d\Gamma \;-\; \int_{\Omega} \Big[\frac{\partial}{\partial x_1} (k \frac{\partial u}{\partial x_1}) \;+\; \frac{\partial}{\partial x_2} (k \frac{\partial u}{\partial x_2}) \;-\; f \Big] \delta u \, d\Omega \;=\; 0 \; . \tag{4.4}$$

Owing to the arbitrariness of δu, the localization theorem of Section 3.1 implies that

$$\frac{\partial}{\partial x_1} (k \frac{\partial u}{\partial x_1}) \;+\; \frac{\partial}{\partial x_2} (k \frac{\partial u}{\partial x_2}) \;=\; f \quad \text{in } \Omega \tag{4.5}$$

and

$$k \frac{\partial u}{\partial n} \delta u \;=\; 0 \quad \text{on } \partial\Omega \; , \tag{4.6}$$

conditional upon sufficient smoothness of the respective fields. The first of the above two equations is identical to the Laplace-Poisson equation $(3.5)_1$, while the second equation presents three distinct alternatives:

(i) Set

$$\delta u \;=\; 0 \quad \text{on } \partial\Omega \; . \tag{4.7}$$

This condition implies that the dependent variable u is prescribed throughout the boundary $\partial\Omega$. The space of admissible fields u is defined as

$$\mathcal{U} \;=\; \big\{ u \in H^1(\Omega) \quad | \quad u = \bar{u} \text{ on } \partial\Omega \big\} \; , \tag{4.8}$$

where $\bar{u}$ is prescribed independently of the functional $I[u]$, in the sense that the functional contains no information regarding the actual value of u on $\partial\Omega$. Boundary conditions such as $u = \bar{u}$, which appear in the space of admissible fields, are referred to as *essential* (or *geometrical*).

(ii) Set

$$k \frac{\partial u}{\partial n} \;=\; 0 \quad \text{on } \partial\Omega \; . \tag{4.9}$$

In this case, the boundary condition applies on the extremal function u, and is exactly derivable from the functional. Boundary conditions that directly apply to the extremal function (and its derivatives) are referred to as *natural* (or *suppressible*). No boundary restrictions are imposed on $\mathcal{U}$ in the present case.

(iii) Admit a decomposition of boundary $\partial\Omega$ into parts Γ_u and Γ_q, such that $\partial\Omega = \overline{\Gamma_u \cup \Gamma_q}$. Subsequently, set

$$\begin{aligned} \delta u \; &= 0 \quad \text{on } \Gamma_u \; , \\ k \frac{\partial u}{\partial n} \; &= 0 \quad \text{on } \Gamma_q \; . \end{aligned} \tag{4.10}$$

Here, essential and natural boundary conditions are enforced on mutually disjoint portions of the boundary. In this case, the problem is said to involve *mixed* boundary conditions, and the space of admissible fields is defined as

$$\mathcal{U} \;=\; \left\{ u \in H^1(\Omega) \quad | \quad u = \bar{u} \text{ on } \Gamma_u \right\} \;. \tag{4.11}$$

It can be concluded from the above, with reference to (4.4) that essential boundary conditions appear on variations of u and, possibly, its derivatives (and therefore place restrictions on the space of admissible fields), while natural boundary conditions appear directly on derivatives of the extremal function u. Equation (4.4) reveals that extremization of the functional in (4.1) yields a function u which satisfies the differential equation $(3.5)_1$ and boundary conditions selected in conjunction to the space of admissible fields $\mathcal{U}$.

Following case (iii), note that the space of admissible variations $\mathcal{U}_0$ is defined as

$$\mathcal{U}_0 \;=\; \left\{ u \in H^1(\Omega) \quad | \quad u = 0 \text{ on } \Gamma_u \right\} \;=\; H^1_0(\Omega) \;. \tag{4.12}$$

It can be easily seen that the option of non-homogeneous natural boundary conditions of the form

$$-k\frac{\partial u}{\partial n} \;=\; \bar{q} \quad \text{on } \Gamma_q \tag{4.13}$$

can be accommodated, if the original functional is amended so that it reads

$$\bar{I}[u] \;=\; \int_\Omega \Big[\frac{k}{2}\left(\frac{\partial u}{\partial x_1}\right)^2 \;+\; \frac{k}{2}\left(\frac{\partial u}{\partial x_2}\right)^2 + fu\Big]\, d\Omega \;+\; \int_{\Gamma_q} \bar{q} u \, d\Gamma \;, \tag{4.14}$$

where $\bar{q} = \bar{q}(x_1, x_2)$ is a continuous function on Γ_q.

Vanishing of the first variation of $\bar{I}[u]$ implies that

$$\int_\Omega \Big[\frac{\partial u}{\partial x_1} k \frac{\partial \delta u}{\partial x_1} \;+\; \frac{\partial u}{\partial x_2} k \frac{\partial \delta u}{\partial x_2} \;+\; f\delta u\Big]\, d\Omega \;+\; \int_{\Gamma_q} \bar{q}\delta u \, d\Gamma \;=\; 0 \;. \tag{4.15}$$

Equation (4.15) is termed the weak (variational) form of boundary-value problem (3.5). Comparing the above equation to (3.14), it is obvious that they are identical provided that the space of admissible field $\mathcal{W}$ for w in (3.14) is identical to that of δu in (4.15).

The nature of the extremum point u (i.e., whether it renders $I[u]$ minimum, maximum or merely stationary) can be determined by means of the second variation of $I[u]$. Specifically, write

$$\delta^2 \bar{I}[u] \;=\; \delta\left(\delta \bar{I}[u]\right) \;=\; \int_\Omega \Big[k\left(\frac{\partial \delta u}{\partial x_1}\right)^2 \;+\; k\left(\frac{\partial \delta u}{\partial x_2}\right)^2\Big]\, d\Omega \;, \tag{4.16}$$

and note that $\delta^2 \bar{I}[u] > 0$, for all $\delta u \neq 0$, provided $\Gamma_u \neq \emptyset$. This is true because, if δu is assumed to be constant throughout the domain, it has to vanish everywhere, by definition of $\mathcal{U}_0$. It turns out that the conditions $\delta \bar{I}[u] = 0$ and $\delta^2 \bar{I}[u] > 0$ are sufficient for any $I[u]$ to attain a local minimum at u, provided that $\delta^2 \bar{I}[u]$ is also bounded from below at u, namely that

$$\delta^2 \bar{I}[u] \ \geq c \|\delta u\|^2 \ , \tag{4.17}$$

where c is a positive constant.

The weak (variational) form of problem (3.5) can be stated operationally as follows: find $u \in \mathcal{U}$, such that for all $\delta u \in \mathcal{U}_0$

$$B(\delta u, u) \ + \ (\delta u, f) \ + \ (\delta u, \bar{q})_{\Gamma_q} \ = \ 0 \ , \tag{4.18}$$

where the bilinear form $B(\delta u, u)$ is defined as

$$B(\delta u, u) \ = \ \int_\Omega \left(\frac{\partial \delta u}{\partial x_1} k \frac{\partial u}{\partial x_1} \ + \ \frac{\partial \delta u}{\partial x_2} k \frac{\partial u}{\partial x_2} \right) d\Omega \ , \tag{4.19}$$

and the linear forms $(\delta u, f)$ and $(\delta u, \bar{q})_{\Gamma_q}$ are defined, respectively, as

$$(\delta u, f) \ = \ \int_\Omega \delta u f \, d\Omega \tag{4.20}$$

and

$$(\delta u, \bar{q})_{\Gamma_q} \ = \ \int_{\Gamma_q} \delta u \bar{q} \, d\Gamma \ . \tag{4.21}$$

The correspondence of the above operational form with that of Section 3.2 is noted for the purpose of the forthcoming comparison between the Galerkin method and the variational method, when applied to problem (3.5).

In addition to the above weak (variational) form, there exists a variational *principle* associated with the solution u of problem (3.5). This can be stated as follows: find $u \in \mathcal{U}$, such that

$$\bar{I}[u] \ \leq \ \bar{I}[v] \ , \tag{4.22}$$

for all $v \in \mathcal{U}$, where $\bar{I}[u]$ is defined in (4.14).

Remark:

☛ Directional derivatives can be used in deriving the weak (variational) equation (4.15) from functional $\bar{I}[u]$. Indeed, write

$$D_v \bar{I}[u] \;=\; 0 \;\Rightarrow \Big[\frac{d}{d\omega}\bar{I}[u+\omega v]\Big]_{\omega=0} \;=\; 0 \;\Rightarrow \\ \int_\Omega \big[\frac{\partial u}{\partial x_1}k\,\frac{\partial v}{\partial x_1} \;+\; \frac{\partial u}{\partial x_2}k\,\frac{\partial v}{\partial x_2} \;+\; fv\big]\,d\Omega \;+\; \int_{\Gamma_q}\bar{q}v\,d\Gamma \;=\; 0\;, \quad (4.23)$$

where $v \in \mathcal{U}_0$.

4.2 Weak (variational) forms and variational principles

The analysis in Section 4.1, as applied to the model problem (3.5), provides an attractive perspective to the solution of certain partial differential equations: the solution is identified with a "point", which minimizes an appropriately constructed functional over an admissible function space. Weak (variational) forms can be made fully equivalent to respective strong forms, as evidenced in the discussion of the weighted residual methods, under certain smoothness assumptions. However, the equivalence between weak (variational) forms and variational principles is not guaranteed: indeed, there exists no general method of constructing functionals $I[u]$, whose extremization recovers a desired weak (variational) form. In this sense, only certain partial differential equations are amenable to analysis and solution by variational methods.

Vainberg's theorem provides the necessary and sufficient condition for the equivalence of a weak (variational) form to a functional extremization problem. If such an equivalence holds, the functional is referred to as a *potential*.

Theorem (Vainberg)
Consider a weak (variational) form

$$G(u,\delta u) \;=\; B(u,\delta u) \;+\; (f,\delta u) \;+\; (\bar{q},\delta u)_{\Gamma_q} \;=\; 0\;, \quad (4.24)$$

where $u \in \mathcal{U}$, $\delta u \in \mathcal{U}_0$, and f and $\bar{q}$ are independent of u. Assume that G possesses a Gâteaux derivative in a neighborhood $\mathcal{N}$ of u, and the Gâteaux differential $D_{\delta u_1}B(u,\delta u_2)$ is continuous in u at every point of $\mathcal{N}$. Then, the necessary and sufficient condition for the above weak form to be derivable from a potential in $\mathcal{N}$ is that

$$D_{\delta u_1} G(u, \delta u_2) \;=\; D_{\delta u_2} G(u, \delta u_1) \;, \tag{4.25}$$

namely that $D_{\delta u_1} G(u, \delta u_2)$ be symmetric for all δu_1, $\delta u_2 \in \mathcal{U}_0$ and all $u \in \mathcal{N}$.

Preliminary to proving the above theorem, introduce the following two lemmas:

Lemma 1

Show that

$$D_v I[u] \;=\; \lim_{\Delta\omega \to 0} \frac{I[u \;+\; \Delta\omega\, v] \;-\; I[u]}{\Delta\omega} \;. \tag{4.26}$$

To prove that Lemma 1 holds, use the definition of the directional derivative of I in the direction v, so that

$$\begin{aligned} D_v I[u] \;&=\; \Big\{ \frac{d}{d\omega} I[u \;+\; \omega v] \Big\}_{\omega=0} \\ &=\; \Big\{ \lim_{\Delta\omega \to 0} \frac{I[u \;+\; \omega v \;+\; \Delta\omega\, v] \;-\; I[u \;+\; \omega v]}{\Delta\omega} \Big\}_{\omega=0} \\ &=\; \lim_{\Delta\omega \to 0} \Big\{ \frac{I[u \;+\; \omega v \;+\; \Delta\omega\, v] \;-\; I[u \;+\; \omega v]}{\Delta\omega} \Big\}_{\omega=0} \\ &=\; \lim_{\Delta\omega \to 0} \frac{I[u \;+\; \Delta\omega\, v] \;-\; I[u]}{\Delta\omega} \;. \end{aligned} \tag{4.27}$$

In the above derivation, note that operations $\frac{d}{d\omega}$ and $|_{\omega=0}$ are not interchangeable (as they both refer to the same variable ω), while $\lim_{\Delta\omega \to 0}$ and $|_{\omega=0}$ <u>are</u> interchangeable, conditional upon sufficient smoothness of $I[u]$.

Lemma 2 (Lagrange's formula)

Let $I[u]$ be a functional with Gâteaux derivatives everywhere, and u, $u + \delta u$ be any points of $\mathcal{U}$. Then,

$$I[u + \delta u] \;-\; I[u] = D_{\delta u} I[u \;+\; \epsilon\, \delta u] \qquad , \qquad 0 < \epsilon < 1 \;. \tag{4.28}$$

To prove Lemma 2, fix u and $u + \delta u$ in $\mathcal{U}$, and define function f on $\mathbb{R}$ as

$$f(\omega) \;=\; I[u \;+\; \omega\, \delta u] \;. \tag{4.29}$$

It follows that

$$\begin{aligned} f' \;&=\; \frac{df}{d\omega} \;=\; \lim_{\Delta\omega \to 0} \frac{f(\omega \;+\; \Delta\omega) \;-\; f(\omega)}{\Delta\omega} \\ &=\; \lim_{\Delta\omega \to 0} \frac{I[u \;+\; \omega\, \delta u \;+\; \Delta\omega\, \delta u] \;-\; I[u \;+\; \omega\, \delta u]}{\Delta\omega} \;=\; D_{\delta u} I[u \;+\; \omega\, \delta u] \;, \end{aligned} \tag{4.30}$$

where Lemma 1 was invoked. Then, using the standard mean-value theorem of calculus, write

$$\begin{aligned} I[u + \delta u] - I[u] &= f(1) - f(0) \\ &= \frac{f(1) - f(0)}{1 - 0} = f'(\epsilon) = D_{\delta u} I[u + \epsilon\,\delta u] , \end{aligned} \tag{4.31}$$

where $0 < \epsilon < 1$.

Given Lemma 2, one may proceed directly to the proof of Vainberg's theorem. First, prove the necessity of (4.25), namely that if an appropriate $I[u]$ exists, then (4.25) holds. To this end, fix $u \in \mathcal{N}$ and define the scalar quantity Δ as

$$\Delta = I[u + a\,\delta u_1 + b\,\delta u_2] - I[u + a\,\delta u_1] - I[u + b\,\delta u_2] + I[u] , \tag{4.32}$$

where a, b are non-zero scalars, such that $u + \eta\,\delta u_1 + \theta\,\delta u_2 \in \mathcal{N}$, for all $0 \leq \eta \leq a$ and $0 \leq \theta \leq b$. Also, define functional $J[u]$ as

$$J[u] = I[u + a\,\delta u_1] - I[u] , \tag{4.33}$$

so that

$$\Delta = J[u + b\,\delta u_2] - J[u] . \tag{4.34}$$

Using Lemma 2 and the above definitions of $J[u]$ and Δ, write

$$\begin{aligned} \Delta = J[u + b\,\delta u_2] - J[u] &= D_{b\delta u_2} J[u + \epsilon_1 b\,\delta u_2] \\ &= D_{b\delta u_2} I[u + \epsilon_1 b\,\delta u_2 + a\,\delta u_1] - D_{b\delta u_2} I[u + \epsilon_1 b\,\delta u_2] \\ &= B(u + \epsilon_1 b\,\delta u_2 + a\,\delta u_1, b\delta u_2) + (f, b\,\delta u_2) + (\bar{q}, b\,\delta u_2)_{\Gamma_q} \\ &\quad - B(u + \epsilon_1 b\,\delta u_2, b\,\delta u_2) - (f, b\,\delta u_2) - (\bar{q}, b\,\delta u_2)_{\Gamma_q} \\ &= b\,B(a\delta u_1, \delta u_2) = ab\,B(\delta u_1, \delta u_2) . \end{aligned} \tag{4.35}$$

Alternatively, let functional $K[u]$ be defined as

$$K[u] = I[u + b\,\delta u_2] - I[u] , \tag{4.36}$$

so that Δ is also written as

$$\Delta = K[u + a\,\delta u_1] - K[u] . \tag{4.37}$$

Using the steps followed in the derivation of (4.35), it can be readily concluded that

$$\Delta = ab\,B(\delta u_2, \delta u_1) , \tag{4.38}$$

so that (4.35) and (4.38) lead to

$$B(\delta u_1, \delta u_2) \;=\; B(\delta u_2, \delta u_1)\ . \tag{4.39}$$

Noting that, due to the linearity of B,

$$\begin{aligned} D_{\delta u_1} B(u, \delta u_2) \;&=\; \Big\{ \frac{d}{d\omega} B(u \;+\; \omega\, \delta u_1, \delta u_2) \Big\}_{\omega=0} \\ &=\; \Big\{ \frac{d}{d\omega} \big[B(u, \delta u_2) \;+\; \omega B(\delta u_1, \delta u_2) \big] \Big\}_{\omega=0} =\; B(\delta u_1, \delta u_2)\ , \end{aligned} \tag{4.40}$$

and, similarly,

$$D_{\delta u_2} B(u, \delta u_1) \;=\; B(\delta u_2, \delta u_1)\ , \tag{4.41}$$

it follows that condition (4.25) holds.

In order to show the sufficiency of (4.25), namely prove that (4.25) implies the existence of an appropriate functional $I[u]$, define

$$I[u] \;=\; \int_0^1 B(tu, u)\, dt \;+\; (f, u) \;+\; (\bar{q}, u)_{\Gamma_q}\ . \tag{4.42}$$

Since

$$\frac{d}{d\omega} B(tu, u \;+\; \omega\, \delta u) \;=\; B(tu, \delta u)\ , \tag{4.43}$$

and

$$\frac{d}{d\omega} B(\omega\, \delta u, u \;+\; \omega\, \delta u) \;=\; B(\delta u, u) \;+\; 2\omega B(\delta u, \delta u)\ , \tag{4.44}$$

the directional derivative of $I[u]$ in the direction δu is written as

$$D_{\delta u} I[u] \;=\; D_{\delta u} \Big(\int_0^1 B(tu, u)\, dt \Big) \;+\; (f, \delta u) \;+\; (\bar{q}, \delta u)_{\Gamma_q}\ . \tag{4.45}$$

With the aid of (4.43) and (4.44), the first term on the right-hand side of the above equation may be rewritten as

$$\begin{aligned} D_{\delta u} \Big(\int_0^1 B(tu, u)\, dt \Big) \;&=\; \int_0^1 D_{\delta u} B(tu, u)\, dt \\ &=\; \int_0^1 \Big\{ \frac{d}{d\omega} B(tu \;+\; t\omega\, \delta u, u \;+\; \omega\, \delta u) \Big\}_{\omega=0} dt \\ &=\; \int_0^1 \Big\{ \frac{d}{d\omega} B(tu, u \;+\; \omega\, \delta u) \;+\; \frac{d}{d\omega} B(t\omega\, \delta u, u \;+\; \omega\, \delta u) \Big\}_{\omega=0} dt \\ &=\; \int_0^1 \big[B(tu, \delta u) \;+\; B(t\, \delta u, u) \big]\, dt\ . \end{aligned} \tag{4.46}$$

Exploiting the assumed symmetry of B, it follows from the above that

$$\begin{aligned} D_{\delta u} \int_0^1 B(tu, u)\, dt \;&=\; 2 \int_0^1 tB(u, \delta u)\, dt \\ &=\; 2B(u, \delta u) \big[\frac{t^2}{2} \big]_0^1 \;=\; B(u, \delta u) \;, \end{aligned} \tag{4.47}$$

which proves that $I[u]$, as defined in (4.42), is indeed an appropriate functional.

<u>Remarks</u>:

- ☞ Apart from some technicalities, Vainberg's theorem can be proved following the above general procedure, even when B is non-linear in u.
- ☞ Checking condition (4.25) is typically an easy task.
- ☞ Vainberg's theorem not only establishes a condition for potentiality of a weak from, but also provides a direct definition of the potential in the form of (4.42).

Example:
Recall the weak (variational) form of Section 3.2, which is associated with boundary-value problem (3.5). In this case,

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

$$(f, \delta u) \;=\; \int_\Omega f\, \delta u\, d\Omega \;,$$

and

$$(\bar{q}, \delta u)_{\Gamma_q} \;=\; \int_{\Gamma_q} \bar{q}\, \delta u\, d\Gamma \;.$$

Using Vainberg's theorem, it can be immediately concluded that, since B is symmetric, there exists a potential $I[u]$, which, according to (4.42), is given by

$$\begin{aligned} I[u] \;&=\; \frac{1}{2} B(u, u) \;+\; (f, u) \;+\; (\bar{q}, u)_{\Gamma_q} \\ &=\; \int_\Omega \big[\frac{k}{2} \left(\frac{\partial u}{\partial x_1} \right)^2 \;+\; \frac{k}{2} \left(\frac{\partial u}{\partial x_2} \right)^2 + fu \big]\, d\Omega \;+\; \int_{\Gamma_q} u \bar{q}\, d\Gamma \;, \end{aligned}$$

whose extremization yield the above weak (variational) form.

4.3 Rayleigh-Ritz method

The Rayleigh-Ritz method provides approximate solutions to partial differential equations, whose weak (variational) form is derivable from a functional $I[u]$. The central idea of the Rayleigh-Ritz method is to extremize $I[u]$ over a properly constructed subspace $\mathcal{U}_h$ of the space of admissible fields $\mathcal{U}$. To this end, write

$$u \doteq u_h \ = \ \sum_{I=1}^{N} \alpha_I \, \varphi_I \ + \ \varphi_0 \ , \tag{4.48}$$

where φ_I, $I = 1, \ldots, N$, is a specified family of interpolation functions that vanish where essential boundary conditions are enforced. In addition, function φ_0 is defined so that u_h satisfy identically the essential boundary conditions. Consequently, a proper N-dimensional subspace $\mathcal{U}_h$ is completely defined by (4.48). Extremization of $I[u]$ over $\mathcal{U}_h$ yields

$$\delta I[u_h] \ = \ \delta I[\sum_{I=1}^{N} \alpha_I \, \varphi_I \ + \ \varphi_0] \ = \ 0 \ . \tag{4.49}$$

Instead of directly obtaining the weak (variational) form of the problem by determining the explicit form of $\delta I[u_h]$ as a function of u_h, one may rewrite the extremization statement as a function of parameters α_I, $I = 1, \ldots, N$, namely

$$\delta I(\alpha_1, \ldots, \alpha_N) \ = \ 0 \ . \tag{4.50}$$

It follows from (4.50) that $I(\alpha_1, \ldots, \alpha_N)$ is an integral scalar function which attains an extremum over $\mathcal{U}_h$ if, and only if,

$$\frac{\partial I}{\partial \alpha_1} \, \delta\alpha_1 \ + \ \frac{\partial I}{\partial \alpha_2} \, \delta\alpha_2 \ + \ \ldots \ + \ \frac{\partial I}{\partial \alpha_N} \, \delta\alpha_N \ = \ 0 \ . \tag{4.51}$$

Since the variations $\delta\alpha_I$, $I = 1, \ldots, N$, are arbitrary, it may be immediately concluded that

$$\frac{\partial I}{\partial \alpha_I} \ = \ 0 \qquad , \qquad I = 1, \ldots, N \ . \tag{4.52}$$

Equations (4.52) may be solved for parameters α_I, so that an approximate solution to the variational problem is expressed by means of (4.48).

Example: Consider the functional $I[u]$ defined in the domain $(0, 1)$ as

$$I[u] \ = \ \int_0^1 [\frac{1}{2} \left(\frac{du}{dx}\right)^2 \ + \ u] \, dx \ + \ 2u|_{x=1} \ ,$$

and the associated essential boundary condition $u(0) = 0$. The above functional is associated with the one-dimensional version of the Laplace-Poisson equation discussed in Section 3.2. In particular, it can be readily established that extremization of $I[u]$ recovers the solution to a boundary-value problem of the form

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

In order to obtain a Rayleigh-Ritz approximation to the solution of the preceding boundary-value problem, write u_h as

$$u_h(x) \ = \ u_N(x) \ = \ \sum_{I=1}^{N} \alpha_I \varphi_I(x) \ + \ \varphi_0(x) \ ,$$

and set, for simplicity, $\varphi_0 = 0$, so that the homogeneous essential boundary condition at $x = 0$ be satisfied. A one-parameter Rayleigh-Ritz approximation can be determined by choosing $\varphi_1(x) = x$. Then,

$$\begin{aligned} I[u_1] &= \int_0^1 \big[\frac{1}{2}\alpha_1^2 \ + \ \alpha_1 x\big]\, dx \ + \ 2\alpha_1 \\ &= \frac{1}{2}\alpha_1^2 \ + \ \frac{5}{2}\alpha_1 \ . \end{aligned}$$

Setting the first variation of $I[u_1]$ to zero, it follows that

$$\alpha_1 \ + \ \frac{5}{2} \ = \ 0 \ ,$$

from where it is concluded that $\alpha_1 = -\frac{5}{2}$, and

$$u_1(x) \ = \ -\frac{5}{2}x \ .$$

Similarly, one may consider a two-parameter polynomial Rayleigh-Ritz approximation by choosing $\varphi_1(x) = x$ and $\varphi_2(x) = x^2$. In this case, $I[u]$ takes the form

$$\begin{aligned} I[u_2] &= \int_0^1 \big[\frac{1}{2}\left(\alpha_1 \ + \ 2\alpha_2 x\right)^2 \ + \ \left(\alpha_1 x \ + \ \alpha_2 x^2\right)\big]\, dx \ + \ 2(\alpha_1 \ + \ \alpha_2) \\ &= \frac{1}{2}\alpha_1^2 \ + \ \alpha_1\alpha_2 \ + \ \frac{2}{3}\alpha_2^2 \ + \ \frac{5}{2}\alpha_1 \ + \ \frac{7}{3}\alpha_2 \ . \end{aligned}$$

Setting the first variation of $I[u_2]$ to zero, results in the system of equations

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

whose solution gives $\alpha_1 = -3$ and $\alpha_2 = \frac{1}{2}$, hence

$$u_2(x) \;=\; -3x \;+\; \frac{1}{2}x^2 \;.$$

The approximate solution $u_2(x)$ coincides with the exact solution of the boundary-value problem. Furthermore, $u_1(x)$ and $u_2(x)$ coincide with the respective solutions obtained in Section 3.2 using the Bubnov-Galerkin method with the same interpolation functions.

A different approximate solution $\tilde{u}_2$ can be obtained using the Rayleigh-Ritz method in connection with *piece-wise* linear polynomial interpolation functions of the form

$$\varphi_1(x) \;=\; \begin{cases} 2x & \text{if } 0 \leq x \leq 0.5 \\ 2(1 \;-\; x) & \text{if } 0.5 < x \leq 1 \end{cases}$$

and

$$\varphi_2(x) \;=\; \begin{cases} 0 & \text{if } 0 \leq x \leq 0.5 \\ 2(x \;-\; \frac{1}{2}) & \text{if } 0.5 < x \leq 1 \end{cases} \;,$$

where functions φ_1 and φ_2 are depicted in Figure 4.1. Then, $I[\tilde{u}_2]$ is written as

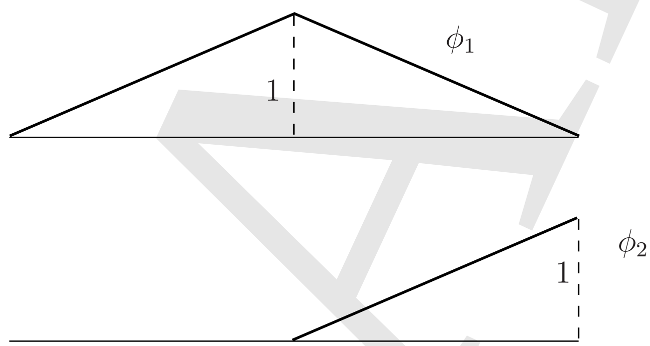


Figure 4.1: *Piecewise linear interpolations functions in one dimension*

$$\begin{aligned} I[\tilde{u}_2] \;&=\; \int_0^{0.5} \big[\frac{1}{2}\,(2\alpha_1)^2 \;+\; 2\alpha_1 x\big]\,dx \\ &\quad+\; \int_{0.5}^{1} \big[\frac{1}{2}\,(-2\alpha_1 \;+\; 2\alpha_2)^2 \;+\; 2\alpha_1(1 \;-\; x) \;+\; 2\alpha_2(x \;-\; \frac{1}{2})\big]\,dx \;+\; 2\alpha_2 \\ &=\; 2\alpha_1^2 \;-\; 2\alpha_1\alpha_2 \;+\; \alpha_2^2 \;+\; \frac{1}{2}\alpha_1 \;+\; \frac{9}{4}\alpha_2 \;. \end{aligned}$$

Again, setting the variation of $I[\tilde{u}_2]$ to zero yields

$$\begin{aligned} 4\alpha_1 \;-\; 2\alpha_2 \;&=\; -\frac{1}{2} \;, \\ -2\alpha_1 \;+\; 2\alpha_2 \;&=\; -\frac{9}{4} \;, \end{aligned}$$

so that $\alpha_1 = -\frac{11}{8}$ and $\alpha_2 = -\frac{5}{2}$, and

$$\tilde{u}_2(x) \;=\; \begin{cases} -\frac{11}{4}x & \text{if } 0 \leq x \leq 0.5 \\ -\frac{1}{4}(1 \;+\; 9x) & \text{if } 0.5 < x \leq 1 \end{cases} \;.$$

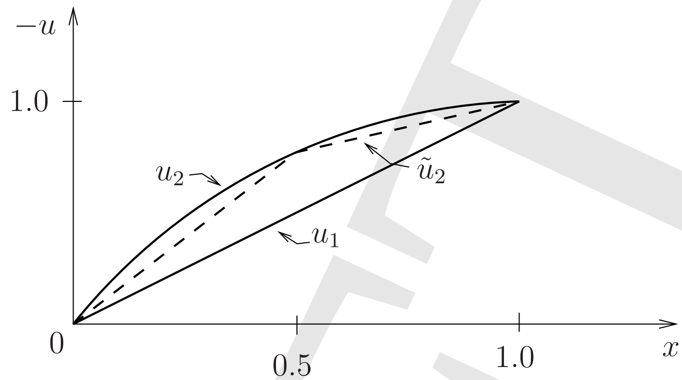


Figure 4.2: *Comparison of exact and approximate solutions*

Solutions u_1, u_2 and $\tilde{u}_2$ are plotted in Figure 4.2

The Rayleigh-Ritz method is related to the Bubnov-Galerkin method, in the sense that, whenever the former is applicable, it yields identical approximate solutions with the latter, when using the same interpolation functions. However, it should be understood that, even in these cases, the methods are fundamentally different in that the former is a variational method, whereas the latter is not.

4.4 Exercises

Problem 1
Consider the fourth-order ordinary differential equation

$$\frac{d^4u}{dx^4} \;=\; f \quad \text{in} \;\; \Omega = (0,1) \; ,$$

where f is a function of x. No boundary conditions are prescribed at this stage on $\partial\Omega = \big\{\{0\},\{1\}\big\}$.

(a) Multiply the differential equation by a function v and subsequently integrate over the domain Ω (note that other than the standard integrability requirement, no restrictions are placed on v, since no boundary conditions have been specified).

(b) Perform two successive integrations by parts on the above integral to obtain

$$\int_\Omega \Big(\frac{d^4u}{dx^4} \;-\; f\Big) v \, dx \;=\; D_v \Big\{ \int_\Omega \Big(\frac{1}{2} \Big(\frac{d^2u}{dx^2}\Big)^2 \;-\; fu \Big) \, dx \Big\} \;+$$
$$\frac{d^3u}{dx^3}(1)\, v(1) \;-\; \frac{d^3u}{dx^3}(0)\, v(0) \;-\; \frac{d^2u}{dx^2}(1)\, \frac{dv}{dx}(1) \;+\; \frac{d^2u}{dx^2}(0)\, \frac{dv}{dx}(0) \, .$$

(c) Conclude from part (b) that stationarity of the functional $I[u]$, defined as

$$I[u] \;=\; \int_\Omega \Big(\frac{1}{2} (\frac{d^2u}{dx^2})^2 \;-\; fu \Big) \, dx \;,$$

implies that the given differential equation is satisfied in Ω and, moreover,

$$\begin{aligned} \frac{d^3u}{dx^3}(1)\, v(1) &= 0 \;, \\ \frac{d^3u}{dx^3}(0)\, v(0) &= 0 \;, \\ \frac{d^2u}{dx^2}(1)\, \frac{dv}{dx}(1) &= 0 \;, \\ \frac{d^2u}{dx^2}(0)\, \frac{dv}{dx}(0) &= 0 \;. \end{aligned}$$

(d) Identify all possible essential and natural boundary conditions on $\partial\Omega$. Note that essential boundary conditions appear on the variations v (and therefore restrict the admissible fields), while natural boundary conditions appear directly on derivatives of the extremal function u.

(e) Consider an expanded functional $I_1[u]$ given by

$$I_1[u] \;=\; \int_\Omega \Big(\frac{1}{2} (\frac{d^2u}{dx^2})^2 \;-\; fu \Big) \, dx \;+\; q_1(0) u(0) \;+\; q_1(1) \frac{du}{dx}(1) \;,$$

where q_1 is defined on $\partial\Omega$, and derive the boundary equations associated with stationarity of $I_1[u]$. Again, identify all possible essential and natural boundary conditions on $\partial\Omega$. Can the functional be further amended so that it read

$$I_2[u] \;=\; \int_\Omega \Big(\frac{1}{2} (\frac{d^2u}{dx^2})^2 \;-\; fu \Big) \, dx \;+\; q_1(0) u(0) \;+\; q_1(1) \frac{du}{dx}(1) \;+\; q_2(1) \frac{d^2u}{dx^2}(1) \;,$$

where q_2 is defined at $x = 1$? Clearly explain your answer.

<u>Problem 2</u>

Consider the initial-value problem

$$\begin{aligned} \frac{\partial}{\partial x} (k \frac{\partial u}{\partial x}) \;-\; f &= l \, \frac{\partial u}{\partial t} \quad \text{in} \;\; \Omega \times (0, T) \;, \\ u &= \bar{u}(x, t) \quad \text{on} \;\; \partial\Omega \times (0, T) \;, \\ u(x, 0) &= u_0(x) \quad \text{in} \;\; \Omega \;, \end{aligned}$$

for the determination of $u = u(x, t)$, where $\Omega \subset \mathbb{R}$, and k, l and f are given non-vanishing functions of x. Starting from the above strong form, obtain the weak (variational) form of the problem and use Vainberg's theorem to show that there exists no variational theorem associated with the weak form.

Problem 3
Consider the boundary-value problem

$$\begin{aligned} -\frac{d}{dx}\big((1+x)\frac{\partial u}{\partial x}\big) &= 0 \quad \text{in} \quad (0,1) \;, \\ u(0) &= 0 \;, \\ u(1) &= 1 \;. \end{aligned}$$

Construct the weak (variational) form of this problem and use Vainberg's theorem to ascertain that there exists a variational principle associated with the problem. Also, construct the relevant functional $I[u]$ and specify the space over which it attains a minimum. Obtain a Rayleigh-Ritz approximation to the solution of the above problem, assuming a two-parameter polynomial approximation. Submit a plot of the exact and the approximate solution.

Problem 4
Consider the boundary-value problem

$$\begin{aligned} \left(\frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2}\right) &= f \qquad \text{in } \Omega = \{(x_1, x_2) \;|\; 0 < x_1, x_2 < 1\} \;, \\ \frac{\partial u}{\partial n} &= \bar{0} \qquad \text{on } \Gamma_q = \{(x_1, x_2) \;|\; x_1 = 0 \text{ or } x_2 = 0\} \;, \\ u &= 0 \qquad \text{on } \Gamma_u = \{(x_1, x_2) \;|\; x_1 = 1 \text{ or } x_2 = 1\} \;, \end{aligned}$$

where $f = f(x_1, x_2)$. In particular, minimize an appropriate functional $I[u]$ over a properly defined functional space $\mathcal{U}$ (treat boundary conditions on $x_1 = 0$ and $x_2 = 0$ as natural) using a one-parameter polynomial approximation

$$u_1(x_1, x_2) \;=\; \alpha_1 \varphi_1 \;,$$

where $\varphi_1(x_1, x_2) = (1-x_1)(1-x_2)$. Also, suggest a two-parameter polynomial approximation

$$u_2(x, y) \;=\; \alpha_1 \varphi_1 \;+\; \alpha_2 \varphi_2$$

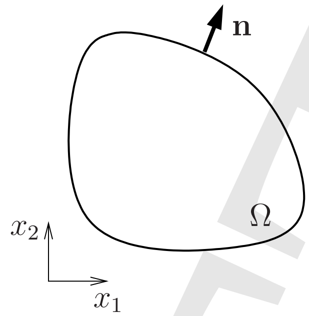
by specifying φ_2 to be the next to φ_1 in the hierarchy of admissible polynomials in $\mathcal{U}$. In this part, do not solve for the coefficients α_1 and α_2.

Problem 5
Consider the homogeneous boundary-value problem

$$\frac{\partial^4 u}{\partial x_1^4} \;+\; 2\frac{\partial^4 u}{\partial x_1^2 \partial x_2^2} \;+\; \frac{\partial^4 u}{\partial x_2^4} \;=\; f \quad \text{in} \quad \Omega \subset \mathbb{R}^2 \;,$$

$$u \;=\; 0 \;\;, \;\; \frac{\partial u}{\partial n} \;=\; 0 \quad \text{on} \quad \partial\Omega \;,$$

where $f = f(x_1, x_2)$.



Show that the solution u of the above problem extremizes the functional $I[u]$ defined as

$$I[u] \;=\; \int_\Omega \Big[\frac{1}{2}\Big\{\big(\frac{\partial^2 u}{\partial x_1^2}\big)^2 \;+\; 2\,\frac{\partial^2 u}{\partial x_1^2}\frac{\partial^2 u}{\partial x_2^2} \;+\; \big(\frac{\partial^2 u}{\partial x_2^2}\big)^2\Big\} \;-\; fu\Big]\, d\Omega\ ,$$

over the set of all admissible functions $\mathcal{U}$. Verify that the same conclusion can be reached for the functional $I_1[u]$ defined as

$$I_1[u] \;=\; \int_\Omega \Big[\frac{1}{2}\Big\{\big(\frac{\partial^2 u}{\partial x_1^2}\big)^2 \;+\; 2\,\big(\frac{\partial^2 u}{\partial x_1 \partial x_2}\big)^2 \;+\; \big(\frac{\partial^2 u}{\partial x_2^2}\big)^2\Big\} \;-\; fu\Big]\, d\Omega\ .$$

4.5 Suggestions for further reading

Sections 4.1

[1] K. Washizu. *Variational Methods in Elasticity & Plasticity.* Pergamon Press, Oxford, 1982. [*This is a classic book on variational methods with emphasis on structural and solid mechanics*].

[2] H. Sagan. *Introduction to the Calculus of Variations.* Dover, New York, 1992. [*This book contains a complete discussion of the theory of first and second variation*].

Section 4.2

[1] M.M. Vainberg. *Variational Methods for the Study of Nonlinear Operators.* Holden-Day, San Francisco, 1964. [*This book contains many important mathematical results, including a non-linear version of Vainberg's theorem*].

Section 4.3

[1] B.A. Finlayson and L.E. Scriven. The method of weighted residuals – a review. *Appl. Mech. Rev.*, 19:735–748, 1966. [*This review article contains on page 741 an exceptionally clear discussion of the relationship between Rayleigh-Ritz and Galerkin methods*].