with the form functions

(1.13)
$$N_1(x,y) = \frac{1}{2A}\begin{vmatrix} 1 & 1 & 1 \\ x & x_2 & x_3 \\ y & y_2 & y_3 \end{vmatrix} = \frac{1}{2A}\left(xy_2 - yx_2 - xy_3 + yx_3 + x_2y_3 - x_3y_2\right)$$

(1.14)
$$N_2(x,y) = \frac{1}{2A}\begin{vmatrix} 1 & 1 & 1 \\ x_1 & x & x_3 \\ y_1 & y & y_3 \end{vmatrix} = \frac{1}{2A}\left(yx_1 - xy_1 + xy_3 - yx_3 - x_1y_3 + y_1x_3\right)$$

(1.15)
$$N_3(x,y) = \frac{1}{2A}\begin{vmatrix} 1 & 1 & 1 \\ x_1 & x_2 & x \\ y_1 & y_2 & y \end{vmatrix} = \frac{1}{2A}\left(xy_1 - yx_1 - xy_2 + yx_2 + x_1y_2 - x_2y_1\right)$$

Notice how the properties of simplicia enable the form functions to be easily determined.

1.2.4. Linear along each direction elements on quadrilaterals in 2D. The element E has 4 nodes $\{(x_1,y_1),(x_2,y_2),(x_3,y_3),(x_4,y_4\}$. It is convenient to introduce a local coordinate system (ξ,η) so that the nodes correspond to the local coordinates $(\pm 1, \pm 1)$. The local approximation is then given in the local coordinates by

$$\tilde{q}(\xi,\eta) = \sum_{k=1} N_k(\xi,\eta)Q_k \tag{1.16}$$

with the form functions

$$N_1(\xi,\eta) = \frac{1}{4}(1+\xi)(1+\eta) \tag{1.17}$$

$$N_2(\xi,\eta) = \frac{1}{4}(1-\xi)(1+\eta) \tag{1.18}$$

$$N_3(\xi,\eta) = \frac{1}{4}(1-\xi)(1-\eta) \tag{1.19}$$

$$N_4(\xi,\eta) = \frac{1}{4}(1+\xi)(1-\eta) \tag{1.20}$$

The local transformation of coordinates can also be written in terms of the form functions

$$x(\xi,\eta) = \sum_{k=1} N_k(\xi,\eta)x_k,\ y(\xi,\eta) = \sum_{k=1} N_k(\xi,\eta)y_k \tag{1.21}$$

2. Variational derivation of weighted residual formulations

We now turn to the problem of how to obtain a measure of the error introduced in approximating the exact solution q to the PDE of interest with its piecewise approximation $\tilde{q}$. Some techniques were presented in the general presentation of weighted residual methods carried out in Chapter 2. For a wide class of problems of interest, especially elliptic problems there exist alternative formulations that lead to more efficient numerical algorithms. These are based upon variational and functional analysis and we shall consider the basics of the theory here.

2.1. Variational calculus. Consider the problem of determining the extremum of the integral

$$I(q) = \int_a^b f(x, q, q')dx \tag{2.1}$$

over all functions $q : \mathbb{R} \to \mathbb{R}$ that belong to some class, for example piecewise continuous functions and that satisfy the boundary conditions $q(x = a) = q_a$, $q(x = b) = q_b$. $I(q)$ is called a functional in that it associates a scalar value to each element from a space of functions. We can consider small perturbations of the function q that we denote by δq The perturbations maintain the boundary conditions, i.e.

$$\delta q(x = a) = 0,\ \delta q(x = b) = 0\ . \tag{2.2}$$

The change in I is

$$\delta I = I(q + \delta q) - I(q) = \int_a^b f(x, q + \delta q, q' + \delta q')dx - \int_a^b f(x, q, q')dx\ . \tag{2.3}$$

We shall consider q, q' as independent variables in f and carry out series expansions to obtain

$$\delta I = \int_a^b \left[\left(\frac{\partial f}{\partial q} \right) \delta q + \left(\frac{\partial f}{\partial q'} \right) \delta q' \right] dx\ . \tag{2.4}$$

We can interchange the δ and d/dx operators in the second term and then integrate by parts

$$\int_a^b \left(\frac{\partial f}{\partial q'} \right) \delta q' dx = \int_a^b \left(\frac{\partial f}{\partial q'} \right) \delta \frac{dq}{dx} dx = \int_a^b \left(\frac{\partial f}{\partial q'} \right) \frac{d}{dx} (\delta q)\, dx = \tag{2.5}$$

$$= \left(\frac{\partial f}{\partial q'} \right) (\delta q) \Big|_{x=a}^{x=b} - \int_a^b \frac{d}{dx} \left(\frac{\partial f}{\partial q'} \right) (\delta q)\, dx \tag{2.6}$$

Applying the boundary conditions and then replacing the above result in (2.4) leads to

$$\delta I = \int_a^b \left[\left(\frac{\partial f}{\partial q} \right) - \frac{d}{dx} \left(\frac{\partial f}{\partial q'} \right) \right] \delta q dx\ . \tag{2.7}$$

For I to be at an extremum δI must maintain the same sign under any perturbation of the extremum. This is only possible if the factor multiplying δq in the above integral is zero everywhere. If it were not then δq_1 would give some value δI_1 and $-\delta q_1$ would lead to the opposite value $-\delta I_1$ and I would not be at an extremum. We therefore have

$$\left(\frac{\partial f}{\partial q} \right) - \frac{d}{dx} \left(\frac{\partial f}{\partial q'} \right) = 0 \tag{2.8}$$

as the condition for I to be at an extremum. This is known as the Euler variational principle. At the extremum we obviously have $\delta I = 0$.

The importance of the Euler variational principle for numerical solution of PDE's rests upon the link it furnishes between an integral formulation $I(q)$ and a differential equation (2.8). We can write down specific forms of f that lead to PDE's of great practical interest. For example replacing

$$f(x, q, q') = \frac{1}{2} \left(\frac{dq}{dx} \right)^2 - gq \tag{2.9}$$

in (2.8) leads to the differential equation

$$q'' = g \tag{2.10}$$

with the boundary conditions $q(x=a) = q_a$, $q(x=b) = q_b$. This is the standard 2 point boundary problem for a second order ODE. Recall that this can be solved by either direct discretization leading to the linear system of equations

$$Q_{j-1} - 2Q_j + Q_{j+1} = h^2 g_j, \; j = 1, \ldots, N-1 \tag{2.11}$$

or by using a shooting method combined with an initial value solve in which we seek $z = q'(x=a)$ that leads to $q(x=b; z) = q_b$. The variational formulation above suggests a third approach. Instead of directly solving the ODE we can seek q that minimizes $I(q)$ with f given by (2.9). This is extremely useful in constructing finite element approximations as we will see below.

Other important expressions of the Euler variational principle can be derived for various situations. Let us consider the ones most often encountered.

(1) Functional of two functions in 1D. The functional is

$$I(p,q) = \int_a^b f(x,p,p',q,q')dx \tag{2.12}$$

and the Euler variational principle leads to

$$\left(\frac{\partial f}{\partial p}\right) + \left(\frac{\partial f}{\partial q}\right) - \frac{d}{dx}\left(\frac{\partial f}{\partial p'}\right) - \frac{d}{dx}\left(\frac{\partial f}{\partial q'}\right) = 0 \tag{2.13}$$

(2) Functional of a 2D function.

$$I(q) = \int_c^d \int_a^b f(x,y,q,q_x,q_y)\,dxdy \tag{2.14}$$

$$\left(\frac{\partial f}{\partial q}\right) - \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial q_x}\right) - \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial q_y}\right) = 0 \tag{2.15}$$

(3) Functional involving second order derivatives in 1D.

$$I(p,q) = \int_a^b f(x,p,p',p'')dx \tag{2.16}$$

$$\left(\frac{\partial f}{\partial q}\right) - \frac{d}{dx}\left(\frac{\partial f}{\partial q'}\right) + \frac{d^2}{dx^2}\left(\frac{\partial f}{\partial q''}\right) = 0 \tag{2.17}$$

2.2. Ritz methods. In the Ritz formulation of the finite element method we seek a piecewise approximation that minimizes the functional associated with the PDE of interest. The piecewise local approximation can be expressed as

$$\tilde{q}(x) = \sum_e \sum_k Q_k^e N_k^e(x) \tag{2.18}$$

where the e sum is over all elements and the k sum is over all nodes within an element. The unknowns of the problem are the nodal values Q_k^e. The form functions $N_k^e(x)$ correspond to some chosen approximation scheme. Let f be associated with the PDE we are interested in solving. The problem reduces to finding $\{Q_k^e\}$ that minimizes

$$I(\tilde{q}) = \int_a^b f(x,\tilde{q},\tilde{q}_x)dx\ . \tag{2.19}$$

This can be solved by finding the solution to the system of equations

$$\frac{\partial}{\partial Q_k^e} I(\bar{q}) = 0 \tag{2.20}$$

with e going over all elements and k over all element nodes.

Note that the entire procedure rests upon the ability to determine a function f that corresponds to a PDE of practical interest. In many situations we have physical guidance that such a variational principle formulation exists. The basic underpinning is furnished by analytical mechanics and the physical principle of least action which finds various expressions in different disciplines. The principle of least action asserts that of all the generalized trajectories $(p, q) = \{q_k(t), p_k(t) \mid k = 1, \ldots, 3N\}$ of a system of N particles, the one actually followed minizes the action S

$$S = \int_{t_0}^{t_1} \mathrm{L}(t, p, q)\, dt \tag{2.21}$$

with L being the Lagrangean of the system. Here q_k denote generalized coordinates and p_k generalized momenta. Though not always immediately apparent this leads to other expressions typically called minimum energy functionals. These can be written for systems with no dissipative effects. Here are some examples of functions f linked to important PDE's:

(1) Poisson equation in 2D

$$f = \frac{1}{2}\left(q_x^2 + q_y^2\right) - gq \tag{2.22}$$

for which (2.15) gives

$$q_{xx} + q_{yy} = g \tag{2.23}$$

(2) Poisson equation in 3D

$$f = \frac{1}{2}\left(q_x^2 + q_y^2 + q_z^2\right) - gq \tag{2.24}$$

for which the Euler variational principle

$$\left(\frac{\partial f}{\partial q}\right) - \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial q_x}\right) - \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial q_y}\right) - \frac{\partial}{\partial z}\left(\frac{\partial f}{\partial q_z}\right) = 0 \tag{2.25}$$

gives

$$q_{xx} + q_{yy} + q_{zz} = g \tag{2.26}$$

2.3. **Galerkin methods.** The Ritz formulation typically leads to a system of equations which has nice numerical properties. However there are many systems for which a variational formulation is not possible typically because the system has dissipative behavior. In such situations we can again use an integral reformulation of the PDE of interest based upon the concept of a weak solution already introduced in the study of hyperbolic problems. Suppose we're looking for a solution to the problem

$$\mathrm{A}q = g \tag{2.27}$$

with A some differential operator. A function q that directly satisfies (2.27) is called a classical solution. Consider now some space of test functions v and a scalar product defined for the functions q and v. From (2.27) we can derive

$$(\mathsf{A}q, v) = (g, v) \tag{2.28}$$

where $(\cdot,\cdot)$ denotes the scalar product, e.g.

$$(u, v) = \int_a^b u(x)v(x)\,dx \;. \tag{2.29}$$

In (2.28) we can apply integration by parts to obtain

$$(q, \mathsf{A}^* v) = (g, v) \tag{2.30}$$

where A^* is the adjoint operator of A. This typically enables us to avoid differentiating functions q that might be discontinuous. We can now use (2.30) to determine the unknown coefficients of a finite element approximation

$$\tilde{q}(x) = \sum_e \sum_k Q_k^e N_k^e(x) \tag{2.31}$$

by requiring

$$\sum_e \sum_k Q_k^e \left(N_k^e(x), \mathsf{A}^* v\right) = (g, v) \;. \tag{2.32}$$

The only piece missing is how we choose the test functions v. In a Galerkin formulation these are chosen to be the form functions themselves leading to

$$\sum_e \sum_k Q_k^e \left(N_k^e(x), \mathsf{A}^* N_j^e(x)\right) = (g, N_j^e(x)) \;, \tag{2.33}$$

thus defining a linear system

$$\mathsf{A}Q = b \tag{2.34}$$

$$A_{jk} = \left(N_k^e(x), \mathsf{A}^* N_j^e(x)\right) \;. \tag{2.35}$$

2.4. A detailed example. Let us now carry out the steps involved in solving a Poisson equation in 2D using a Ritz formulation and quadrilateral elements. The mathematical statement of the problem is

$$\left\{ \begin{array}{ll} q_{xx} + q_{yy} = g & (x, y) \in \Omega \\ q = b & (x, y) \in \partial\Omega \end{array} \right. \tag{2.36}$$

with the domain $\Omega = [a, b] \times [c, d]$ and $\partial\Omega$ denoting the boundary of Ω on which Dirichlet conditions are given. The element form functions are given by (1.17)-(1.20) and the function f is given by (2.22). The function $I(\tilde{q})$ is

$$I(\tilde{q}) = \int_c^d \int_a^b f(x, y, \tilde{q}, \tilde{q}_x, \tilde{q}_y)\,dxdy = \int_c^d \int_a^b \left[\frac{1}{2}\left(\tilde{q}_x^2 + \tilde{q}_y^2\right) - g\tilde{q} \right] dxdy \;. \tag{2.37}$$

The finite element approximation is determined by the chosen form functions and the nodal values $\{Q_k^e\}$. The extremum of $I(\tilde{q})$ is attained when

$$\frac{\partial}{\partial Q_k^e} I(\tilde{q}) = 0 \tag{2.38}$$

which leads to

$$\int_c^d \int_a^b \left[\left(\tilde{q}_x \frac{\partial \tilde{q}_x}{\partial Q_k^e} + \tilde{q}_y \frac{\partial \tilde{q}_y}{\partial Q_k^e} \right) - g \frac{\partial \tilde{q}}{\partial Q_k^e} \right] dxdy = 0 \tag{2.39}$$