10

Basic relations for four-node shell elements

In this chapter we describe the basic relations for four-node shell elements related to the FE approximations, numerical integration, and derivation of the tangent matrix and residual vector. The literature on four-node shell elements is vast, see, e.g., [123, 101, 162, 117, 118, 192, 235, 209, 85, 213, 165, 73, 217, 106, 242, 41, 108, 32, 68, 201, 243, 53], and many others.

The finite element method has achieved remarkable sophistication, but also great complexity, see the classical textbooks on FEs, such as [36, 98, 125, 61, 62, 16, 268, 58] and the new ones [160] and [161]. The requirements which new shell elements have to satisfy, are better defined and more demanding than they were some years ago.

10.1 Bilinear isoparametric approximations

Bilinear shape functions. Consider a bilinear function,

$$f(\xi,\eta) \doteq a_0 + a_1\xi + a_2\eta + a_3\xi\eta, \tag{10.1}$$

where $\xi,\eta \in [-1,+1]$ are the *natural coordinates*. Note that the domain is a bi-unit square, spanned by the corner nodes of coordinates $\{\xi_I,\eta_I\} = \{\pm 1,\pm 1\}$, see Fig. 10.1.

The coefficients a_i can be expressed in terms of values of f at corner nodes, using the conditions (i) $f = 1$ at one of the corner nodes, say I, and (ii) $f = 0$ at all other nodes. This yields a set of four equations, from which we can determine a_i for a selected I. The f, with so-determined a_i, is denoted as N_I, and designated as the *shape function*. Repeating this procedure for all corner nodes, we obtain

$$N_1(\xi,\eta) \doteq \frac{1}{4}(1-\xi)(1-\eta), \qquad N_2(\xi,\eta) \doteq \frac{1}{4}(1+\xi)(1-\eta),$$
$$N_3(\xi,\eta) \doteq \frac{1}{4}(1+\xi)(1+\eta), \qquad N_4(\xi,\eta) \doteq \frac{1}{4}(1-\xi)(1+\eta) \quad (10.2)$$

or, in concise form,

$$N_I(\xi,\eta) \doteq \frac{1}{4}(1+\xi_I\,\xi)\,(1+\eta_I\,\eta), \qquad I=1,2,3,4, \quad (10.3)$$

where $(\xi_I,\eta_I) = (\pm 1,\pm 1)$ are coordinates of node I in the bi-unit domain. Note that N_I is a hyperbolic paraboloid (saddle) surface of ξ,η.

Vector of shape functions. Let us define the following vector of shape functions

$$\mathbf{N}(\xi,\eta) \doteq [N_1(\xi,\eta),\, N_2(\xi,\eta),\, N_3(\xi,\eta),\, N_4(\xi,\eta)]\,, \quad (10.4)$$

where N_I are defined in eq. (10.2). This vector can be re-arranged as follows:

$$\mathbf{N}(\xi,\eta) = \frac{1}{4}\,(\mathbf{s} + \boldsymbol{\xi}\,\xi + \boldsymbol{\eta}\,\eta + \mathbf{h}\,\xi\eta)\,, \quad (10.5)$$

where the auxiliary vectors are defined as

$$\mathbf{s} \doteq [1,1,1,1], \qquad \boldsymbol{\xi} \doteq [-1,1,1,-1],$$
$$\boldsymbol{\eta} \doteq [-1,-1,1,1], \qquad \mathbf{h} \doteq [1,-1,1,-1], \quad (10.6)$$

and the subsequent entries correspond to the consecutive nodes. It is easy to check that the vectors $\mathbf{s}$, $\boldsymbol{\xi}$, $\boldsymbol{\eta}$, $\mathbf{h}$ are mutually orthogonal. The vector $\mathbf{s}$ is the translation vector, while the vectors $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$ define ξ and η positions of consecutive nodes. The hourglass vector $\mathbf{h}$ multiplies the bilinear term $\xi\eta$.

Isoparametric approximations for shell elements. In an isoparametric shell FE, the initial position vector $\mathbf{y}_0$, the displacement vector $\mathbf{u}$ and the rotation vector $\boldsymbol{\psi}$, all for the reference surface, are approximated by the same shape functions $N_I(\xi^\alpha)$ as follows:

$$\mathbf{y}_0(\xi^\alpha) = \sum_{I=1}^{nel} N_I(\xi^\alpha)\,\mathbf{y}_{0I}, \qquad \mathbf{u}(\xi^\alpha) = \sum_{I=1}^{nel} N_I(\xi^\alpha)\,\mathbf{u}_I,$$
$$\boldsymbol{\psi}(\xi^\alpha) = \sum_{I=1}^{nel} N_I(\xi^\alpha)\,\boldsymbol{\psi}_I, \qquad \alpha = 1,2, \quad (10.7)$$

where $(\cdot)_I$ denotes a value at node I, and nel is the number of nodes on an element; for four-node elements $nel = 4$. In the sequel, the natural coordinates are designated in two ways,

$$\xi^k \doteq \{\xi^\alpha, \xi^3\} = \{\xi, \eta, \xi^3\}, \qquad \xi^\alpha, \xi^3, \xi, \eta \in [-1, +1], \tag{10.8}$$

where $\alpha = 1, 2$ and $k = 1, 2, 3$.

10.2 Geometry and bases of shell element

The initial geometry of a four-node shell element is defined by (i) positions of four corner nodes and (ii) bilinear shape functions. Using them, we can define the initial position of the reference surface and construct the local vector normal to this surface. The shell as a 3D body is generated by assuming some thickness in this normal direction.

Parametrization of reference surface. The position vector $\mathbf{y}_0$, which defines the reference surface in the initial configuration, see Fig. 10.1, is parameterized in terms of the natural coordinates $\xi, \eta \in [-1, +1]$ in the following way:

$$\mathbf{y}_0(\xi, \eta) = \sum_{I=1}^{4} N_I(\xi, \eta)\ \mathbf{y}_{0I}, \tag{10.9}$$

where $\mathbf{y}_{0I}$ is a position vector of node I, and $N_I(\xi, \eta)$ are shape functions of eq. (10.3). Alternatively, we can write this expression for each component k separately,

$$y_{0k}(\xi, \eta) = \sum_{I=1}^{4} N_I(\xi, \eta)\ y_{0kI}, \qquad k = 1, 2, 3, \tag{10.10}$$

where $y_{0k} \doteq \mathbf{y}_0 \cdot \mathbf{i}_k$ and $y_{0kI} \doteq \mathbf{y}_{0I} \cdot \mathbf{i}_k$. We can use the vector of shape functions of eq. (10.4), to avoid the summation sign,

$$y_{0k}(\xi, \eta) = \mathbf{N}(\xi, \eta)\ \mathbf{y}_{0kI}, \tag{10.11}$$

where $\mathbf{y}_{0kI} \doteq [y_{0k1}, y_{0k2}, y_{0k3}, y_{0k4}]$ is the vector of kth components of the position vectors of all nodes.

The parametrization of the reference surface defined by eq. (10.10) spans either a planar element or a warped element when one of the nodes is shifted out of the plane spanned by the other three nodes. The latter case is illustrated by an example below. More information on warped elements is provided in Sect. 14.

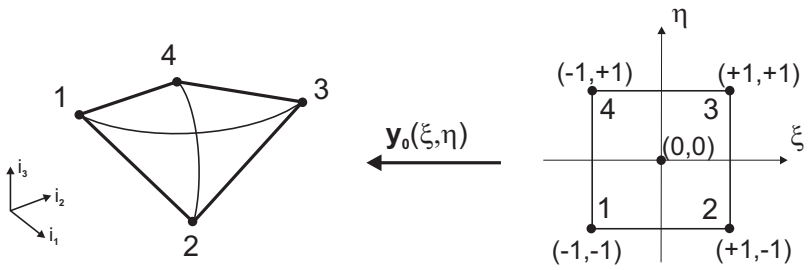

Fig. 10.1 Physical and reference (parent) bi-unit domain of a four-node element.

Example. Consider a square 2×2 element, with nodes 1, 2, 4 located in the X0Y plane, and node 3 elevated in the z-direction by w, see Fig. 10.2a.

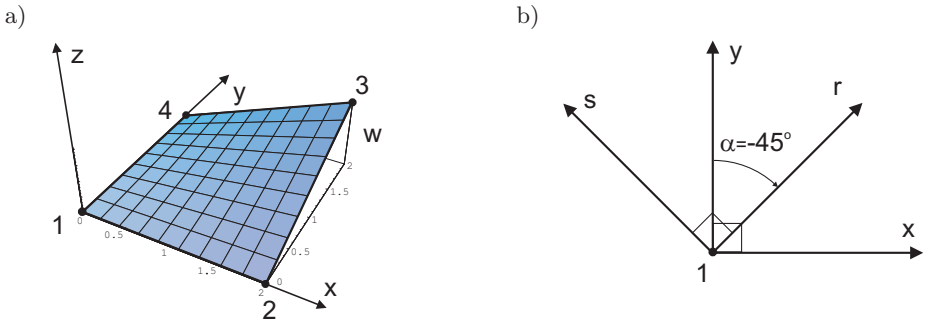

Fig. 10.2 a) The hyperbolic paraboloidal surface spanned by four nodes.
b) Two coordinate systems.

Denote the components of the position vector as follows: $y_{01} \doteq x$, $y_{02} \doteq y$, $y_{03} \doteq z$. The Cartesian coordinates (x, y, z) of corner nodes are assumed to be $p_1 = (0, 0, 0)$, $p_2 = (2, 0, 0)$, $p_3 = (2, 2, w)$, and $p_4 = (0, 2, 0)$. Grouping the x, y, and z components of all nodes as follows::

$$x_I = [0, 2, 2, 0], \qquad y_I = [0, 0, 2, 2], \qquad z_I = [0, 0, w, 0],$$

and using eq. (10.10) for each component, we obtain

$$x = 1 + \xi, \qquad y = 1 + \eta, \qquad z = \frac{w}{4}(1 + \xi)(1 + \eta).$$

From the first two equations, we have $\xi = x - 1$ and $\eta = y - 1$, and the third equation can be expressed solely in Cartesian coordinates,

$$z = \frac{w}{4} xy. \tag{10.12}$$

This is an equation of a hyperbolic paraboloidal (h-p) surface, shown in Fig. 10.2a, having a saddle point at node 1. To obtain this equation in a classical form, we introduce the coordinates r, s, rotated by 45^o w.r.t. $[x, y]$ coordinates, see Fig. 10.2b. Then

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} c & s \\ -s & c \end{bmatrix} \begin{bmatrix} r \\ s \end{bmatrix}, \tag{10.13}$$

where $s \doteq \sin\alpha$, $c \doteq \cos\alpha$ and $\alpha = -45^o$. Using this relation in eq. (10.12), we obtain

$$z = \frac{w}{8}(r^2 - s^2), \tag{10.14}$$

which has the standard form of the h-p surface equation, see [150] p. 545. If we cut this surface using the vertical planes r=const. or s=const., then we obtain parabolas with either minimum or maximum at node 1. If we cut this surface using horizontal planes $z = \text{const.}$, we then obtain hyperbolas, with the asymptotes intersecting at node 1.

Remark 1. For planar (2D) four-node elements, if all angles are smaller than π, then there exists a one-to-one mapping between the element and a bi-unit square spanned on the nodes $\{\xi, \eta\}_I = \{\pm 1, \pm 1\}$, see [98] p. 116. This is also true for planar shell elements, but for warped ones, the question of the inverse mapping becomes complicated. In fact, using the FE method and a numerical integration, we do not need this information, and the Jacobian matrix and its inverse at integration points suffice.

Remark 2. The four-node quadrilateral shell element can also be defined in another way by additionally using the normal vectors at nodes. These vectors must be either computed, e.g. using normals of all elements connected to the node, or be provided as input data, which can be cumbersome. Another possibility is to use a CAD program, in which we can define typical shapes of regular surfaces and directly obtain a normal vector at a selected point.

Natural tangent vectors. The natural vectors tangent to the reference surface are defined as

$$\mathbf{g}_1(\xi, \eta) \doteq \frac{\partial \mathbf{y}_0(\xi, \eta)}{\partial \xi}, \qquad \mathbf{g}_2(\xi, \eta) \doteq \frac{\partial \mathbf{y}_0(\xi, \eta)}{\partial \eta}. \tag{10.15}$$

In general, these vectors are neither unit nor mutually orthogonal. For the bilinear approximation and $\mathbf{y}_0$ in the form of eq. (10.9), we obtain

$$\mathbf{g}_\alpha(\xi,\eta) = \sum_{I=1}^{4} N_{I,\alpha}\,\mathbf{y}_{0I}, \qquad \alpha = 1,2, \tag{10.16}$$

where the derivatives of the shape functions are $N_{I,1} = \frac{1}{4}\xi_I\,(1+\eta_I\,\eta)$ and $N_{I,2} = \frac{1}{4}(1+\xi_I\,\xi)\,\eta_I$. Using the vector of shape functions $\mathbf{N}(\xi,\eta)$ of eq. (10.11), we have

$$(\mathbf{g}_\alpha)_k(\xi,\eta) = \mathbf{N}_{,\alpha}\,\mathbf{y}_{0kI}, \tag{10.17}$$

where $\mathbf{N}_{,1} = \frac{1}{4}\,(\boldsymbol{\xi} + \mathbf{h}\,\eta)$ and $\mathbf{N}_{,2} = \frac{1}{4}\,(\boldsymbol{\eta} + \mathbf{h}\,\xi)$. We see that the tangent vectors vary over the element; $\mathbf{g}_1$ is constant in ξ and linear in η, while $\mathbf{g}_2$ is the opposite way round. At the element's center, $\xi = \eta = 0$, the vectors $\mathbf{g}_\alpha$ are equal to 1/2 of vectors connecting the opposite middle-edge points.

Normal vector. The vector normal to the reference surface is defined as a cross-product of the natural tangent vectors

$$\bar{\mathbf{g}}_3(\xi,\eta) \doteq \mathbf{g}_1(\xi,\eta) \times \mathbf{g}_2(\xi,\eta). \tag{10.18}$$

Note that $\bar{\mathbf{g}}_3$ is perpendicular to the tangent vectors $\mathbf{g}_\alpha$, but is not of unit length. This vector is not associated with the coordinate $\xi^3 \in [-1,+1]$; see eq. (10.30).

Local Cartesian basis. For an irregular geometry of an element, the basis $\{\mathbf{g}_\alpha, \mathbf{g}_3\}$ is normal but skew, which is not convenient, e.g., to define the constitutive relations for non-isotropic materials. Hence, a local Cartesian basis is introduced as described below.

Designate the local Cartesian basis by $\{\mathbf{t}_k\}$ $(k = 1,2,3)$. Define the normal vector as

$$\mathbf{t}_3 \doteq \frac{\bar{\mathbf{g}}_3}{\|\bar{\mathbf{g}}_3\|}, \tag{10.19}$$

where $\bar{\mathbf{g}}_3$ is defined by eq. (10.18). The tangent vectors of the local Cartesian basis can be constructed in several ways; we define them in terms of the auxiliary normalized natural vectors

$$\tilde{\mathbf{g}}_\alpha \doteq \frac{\mathbf{g}_\alpha}{\|\mathbf{g}_\alpha\|}, \tag{10.20}$$

designated by a tilde. Below three types of bases used in shell elements are presented.

Basis 1. *One vector is parallel to the vector of the natural basis*

$$\mathbf{t}_1 \doteq \tilde{\mathbf{g}}_1, \qquad \mathbf{t}_2 \doteq \mathbf{t}_3 \times \mathbf{t}_1, \tag{10.21}$$

where $\mathbf{t}_1$ is identical as $\tilde{\mathbf{g}}_1$, see Fig. 10.3a. This basis was used, e.g., in DYNA3D, see [89], eq. (35).

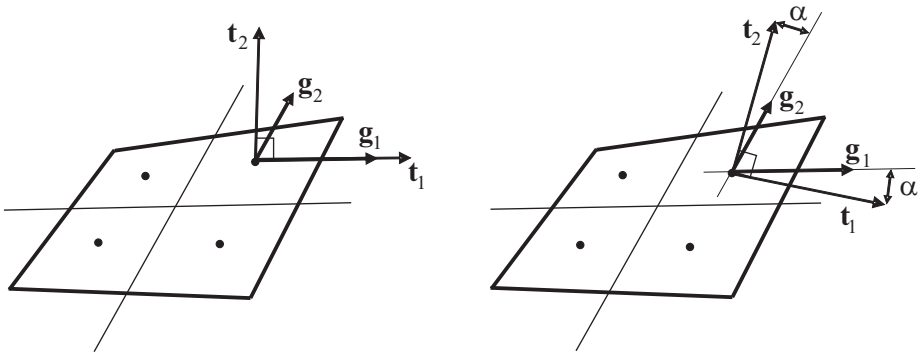

Fig. 10.3 Ortho-normal frames at a Gauss point. a) **Basis 1** and b) **Basis 2**.

Basis 2. *Vectors are equally distant from vectors of the natural basis*

$$\mathbf{t}_1 \doteq \frac{1}{\sqrt{2}}(\tilde{\mathbf{t}}_1 - \tilde{\mathbf{t}}_2), \qquad \mathbf{t}_2 \doteq \frac{1}{\sqrt{2}}(\tilde{\mathbf{t}}_1 + \tilde{\mathbf{t}}_2), \tag{10.22}$$

where the auxiliary vectors are

$$\tilde{\mathbf{t}}_1 \doteq \frac{\tilde{\mathbf{g}}_1 + \tilde{\mathbf{g}}_2}{\|\tilde{\mathbf{g}}_1 + \tilde{\mathbf{g}}_2\|}, \qquad \tilde{\mathbf{t}}_2 \doteq \mathbf{t}_3 \times \tilde{\mathbf{t}}_1. \tag{10.23}$$

This is the most popular basis, see in [98] p. 386 or [50] p. 111 and is shown in Fig. 10.3b.

For this basis, we can show that the average of $\mathbf{t}_1$ and $\mathbf{t}_2$ and the average of $\tilde{\mathbf{g}}_1$ and $\tilde{\mathbf{g}}_2$ are parallel, but have different lengths, i.e.

$$\tfrac{1}{2}(\mathbf{t}_1 + \mathbf{t}_2) = \frac{1}{2\sqrt{2}}\tilde{\mathbf{t}}_1 = a\,\tfrac{1}{2}(\tilde{\mathbf{g}}_1 + \tilde{\mathbf{g}}_2),$$

where $a = 1/(\sqrt{2}\,\|\tilde{\mathbf{g}}_1 + \tilde{\mathbf{g}}_2\|)$.

Besides, the term "equally distant" means that the angle between $\tilde{\mathbf{g}}_1$ and $\mathbf{t}_1$ is equal to the angle between $\tilde{\mathbf{g}}_2$ and $\mathbf{t}_2$. This can be checked in the following way:

$$\begin{aligned}\mathbf{t}_1 \cdot \tilde{\mathbf{g}}_1 &= a\,\{(\tilde{\mathbf{g}}_1 + \tilde{\mathbf{g}}_2) \cdot \tilde{\mathbf{g}}_1 - [\mathbf{t}_3 \times (\tilde{\mathbf{g}}_1 + \tilde{\mathbf{g}}_2)] \cdot \tilde{\mathbf{g}}_1\} \\ &= a\,\{\tilde{\mathbf{g}}_1 \cdot \tilde{\mathbf{g}}_1 - [\mathbf{t}_3 \times \tilde{\mathbf{g}}_2] \cdot \tilde{\mathbf{g}}_1\}\,, \\ \mathbf{t}_2 \cdot \tilde{\mathbf{g}}_2 &= a\,\{(\tilde{\mathbf{g}}_1 + \tilde{\mathbf{g}}_2) \cdot \tilde{\mathbf{g}}_2 + [\mathbf{t}_3 \times (\tilde{\mathbf{g}}_1 + \tilde{\mathbf{g}}_2)] \cdot \tilde{\mathbf{g}}_2\} \\ &= a\,\{\tilde{\mathbf{g}}_2 \cdot \tilde{\mathbf{g}}_2 + [\mathbf{t}_3 \times \tilde{\mathbf{g}}_1] \cdot \tilde{\mathbf{g}}_2\}\,.\end{aligned}$$

These two scalar products are equal, as $\tilde{\mathbf{g}}_1 \cdot \tilde{\mathbf{g}}_1 = \tilde{\mathbf{g}}_2 \cdot \tilde{\mathbf{g}}_2 = 1$ by eq. (10.20).

Finally, we note that for 2D problems formulated in the $\{\mathbf{g}_1, \mathbf{g}_2\}$-plane, we can obtain the components of $\tilde{\mathbf{t}}_2$ as follows::

$$\tilde{\mathbf{t}}_2 = \mathbf{t}_3 \times \tilde{\mathbf{t}}_1 = [0, 0, 1]^T \times [t_1,\, t_2,\, 0]^T = -[t_2,\, t_1,\, 0]^T,$$

where the components of $\tilde{\mathbf{t}}_1 = [t_1, t_2, 0]^T$ are known.

Basis 2. Version 2. *Vectors equally distant from vectors of natural basis*

$$\mathbf{t}_1 \doteq \cos(-\beta/2)\,\tilde{\mathbf{t}}_1 + \sin(-\beta/2)\,\tilde{\mathbf{t}}_2, \quad \mathbf{t}_2 \doteq -\sin(-\beta/2)\,\tilde{\mathbf{t}}_1 + \cos(-\beta/2)\,\tilde{\mathbf{t}}_2, \tag{10.24}$$

where

$$\beta = \arctan \frac{(\mathbf{t}_1 \cdot \mathbf{g}_2)}{(\mathbf{t}_2 \cdot \mathbf{g}_2)}, \qquad \tilde{\mathbf{t}}_1 = \tilde{\mathbf{g}}_1, \qquad \tilde{\mathbf{t}}_2 = \mathbf{t}_3 \times \tilde{\mathbf{t}}_1. \tag{10.25}$$

Derivation. To obtain the same angle between $\mathbf{g}_1$ and $\mathbf{t}_1$ and between $\mathbf{g}_2$ and $\mathbf{t}_2$, we generate an orthonormal basis and then rotate it around the normal vector $\mathbf{t}_3$. We shall use **Basis 1** of eq. (10.21) as the orthonormal basis to start from and denote its vectors as follows:

$$\tilde{\mathbf{t}}_1 = \tilde{\mathbf{g}}_1, \qquad \tilde{\mathbf{t}}_2 = \mathbf{t}_3 \times \tilde{\mathbf{t}}_1.$$

The angle between $\mathbf{t}_2$ and $\mathbf{g}_2$, is denoted by β and we assume that $|\beta| < \pi/2$. To determine β, we can use the following formulas:

$$\sin\beta = \frac{(\tilde{\mathbf{t}}_1 \cdot \mathbf{g}_2)}{\|\mathbf{g}_2\|}, \qquad \cos\beta = \frac{(\tilde{\mathbf{t}}_2 \cdot \mathbf{g}_2)}{\|\mathbf{g}_2\|}, \qquad \tan\beta = \frac{\sin\beta}{\cos\beta} = \frac{(\tilde{\mathbf{t}}_1 \cdot \mathbf{g}_2)}{(\tilde{\mathbf{t}}_2 \cdot \mathbf{g}_2)},$$

from which we obtain

$$\beta = \arctan \frac{(\tilde{\mathbf{t}}_1 \cdot \mathbf{g}_2)}{(\tilde{\mathbf{t}}_2 \cdot \mathbf{g}_2)}.$$

Then we rotate the frame $\{\tilde{\mathbf{t}}_1, \tilde{\mathbf{t}}_2\}$ by the angle $-\beta/2$ around the normal vector $\mathbf{t}_3$. For the rotation tensor defined as

$$\mathbf{R} = \cos(-\beta/2)(\tilde{\mathbf{t}}_1 \otimes \tilde{\mathbf{t}}_1 + \tilde{\mathbf{t}}_2 \otimes \tilde{\mathbf{t}}_2) + \sin(-\beta/2)(\tilde{\mathbf{t}}_2 \otimes \tilde{\mathbf{t}}_1 - \tilde{\mathbf{t}}_1 \otimes \tilde{\mathbf{t}}_2) + \mathbf{t}_3 \otimes \mathbf{t}_3,$$

from $\mathbf{t}_1 = \mathbf{R}\tilde{\mathbf{t}}_1$ and $\mathbf{t}_2 = \mathbf{R}\tilde{\mathbf{t}}_2$, we obtain eq. (10.24).

Note that we can obtain $\sin(-\beta/2)$ and $\cos(-\beta/2)$ in another way, without the use of the arctan function. Then, first we calculate

$$\sin\beta = \tilde{\mathbf{t}}_1 \cdot \tilde{\mathbf{g}}_2, \qquad \cos\beta = \tilde{\mathbf{t}}_2 \cdot \tilde{\mathbf{g}}_2$$

and then, using the half-angle formulas, we obtain

$$\sin(-\beta/2) = s\sqrt{\tfrac{1}{2}(1-\cos\beta)}, \qquad \cos(-\beta/2) = +\sqrt{\tfrac{1}{2}(1+\cos\beta)},$$

where $s = \operatorname{sign}(-\sin\beta)$.

Basis 3. *Vectors related to the reference basis.*

$$\text{If } (\mathbf{t}_3 \cdot \mathbf{i}_1) < (1-\tau) \text{ then } \mathbf{t}_2 = \frac{\mathbf{t}_3 \times \mathbf{i}_1}{\|\mathbf{t}_3 \times \mathbf{i}_1\|} = \frac{1}{\sqrt{t_2^2 + t_3^2}} \begin{bmatrix} 0 \\ t_3 \\ -t_2 \end{bmatrix}, \quad (10.26)$$

$$\text{otherwise } \mathbf{t}_2 = \frac{\mathbf{t}_3 \times \mathbf{i}_2}{\|\mathbf{t}_3 \times \mathbf{i}_2\|} = \frac{1}{\sqrt{t_1^2 + t_2^2}} \begin{bmatrix} t_2 \\ -t_1 \\ 0 \end{bmatrix}, \quad (10.27)$$

and

$$\mathbf{t}_1 = \mathbf{t}_2 \times \mathbf{t}_3, \quad (10.28)$$

where $\mathbf{t}_3 = [t_1, t_2, t_3]^T$ denotes the components of the normal vector in the reference (global) basis $\{\mathbf{i}_k\}$ and τ is a small parameter. The advantage of this basis is that it provides easy identification of directions for complicated curved structures.

The second formula, eq. (10.27), is used when $\mathbf{t}_3$ and $\mathbf{i}_1$ are almost or exactly parallel and the cross-product $\mathbf{t}_3 \times \mathbf{i}_1$ is not well conditioned. Consider the result of the above definitions for two limit cases.

1. If $\mathbf{t}_3 \| \mathbf{i}_1$, then $\mathbf{t}_2 = \mathbf{t}_3 \times \mathbf{i}_2 \| \mathbf{i}_3$ and $\mathbf{t}_1 = \mathbf{t}_2 \times \mathbf{t}_3 \| \mathbf{i}_2$.
2. If $\mathbf{t}_3 \| \mathbf{i}_2$, then $\mathbf{t}_2 = \mathbf{t}_3 \times \mathbf{i}_1 \| -\mathbf{i}_3$ and $\mathbf{t}_1 = \mathbf{t}_2 \times \mathbf{t}_3 \| \mathbf{i}_1$.

Both these cases are shown in Fig. 10.4. Another basis related to the reference basis $\{\mathbf{i}_k\}$ is defined in [71] p. 242.

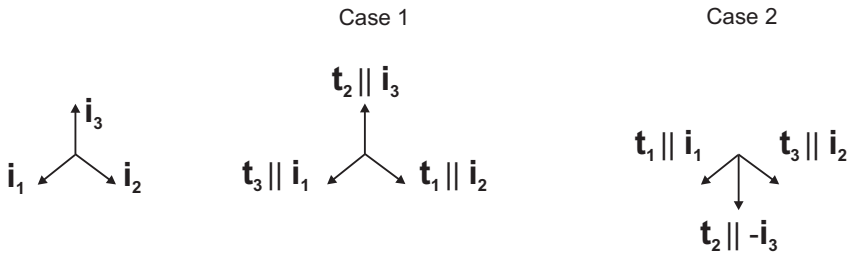

Fig. 10.4 Basis 3. Two limit cases of its position.

Normal vector associated with ξ^3. Note that the vector $\bar{\mathbf{g}}_3$ of eq. (10.18) is not associated with the coordinate $\xi^3 \in [-1,+1]$ and below we derive the proper vector.

The position vector of a shell lamina $\zeta = \text{const.}$ relative to the middle surface is $(\mathbf{y}-\mathbf{y}_0) = \zeta \mathbf{t}_3$, where $\zeta \in [-h/2, +h/2]$. We can parameterize ζ in terms of $\xi^3 \in [-1,+1]$, as $\zeta = (h/2)\,\xi^3$, and define the normal vector as a derivative w.r.t. ξ^3, i.e.

$$\mathbf{g}_3 \doteq \frac{\mathrm{d}(\mathbf{y}-\mathbf{y}_0)}{\mathrm{d}\xi^3} = \frac{h}{2}\mathbf{t}_3. \tag{10.29}$$

This vector stretches from the middle surface to the top surface of a shell, see Fig. 10.5.

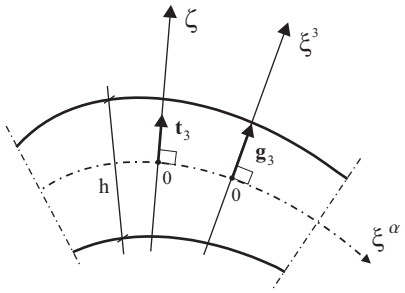

Fig. 10.5 Associated pairs: $(\xi^3, \mathbf{g}_3)$ and $(\zeta, \mathbf{t}_3)$.

Note that the vector $\mathbf{g}_3$ is different from $\bar{\mathbf{g}}_3$ of eq. (10.18) and is associated with ξ^3 because

$$\xi^3 \mathbf{g}_3 = \zeta\, \mathbf{t}_3. \tag{10.30}$$

Hence, we can use either $(\xi^3, \mathbf{g}_3)$ or $(\zeta, \mathbf{t}_3)$, but certainly not $(\xi^3, \bar{\mathbf{g}}_3)$.

Remark. In some works, the normal vector is approximated as

$$\mathbf{t}_3(\xi,\eta) \doteq \sum_{I=1}^{4} N_I(\xi,\eta)\ \mathbf{n}_I, \tag{10.31}$$

where $\mathbf{n}_I$ is the normal unit vector at node I defined as

$$\mathbf{n}_I = \frac{\mathbf{a}_I \times \mathbf{b}_I}{\|\mathbf{a}_I \times \mathbf{b}_I\|}$$

and $\mathbf{a}_I$ and $\mathbf{b}_I$ are the vectors connecting node I with the adjacent corners, see Fig. 10.6. When the element is planar, then this definition is equivalent to eq. (10.19). However, when the element is warped, vectors $\mathbf{n}_I$ are not parallel and $\mathbf{t}_3$ is neither unit nor perpendicular to the local tangent vectors $\mathbf{g}_\alpha$.

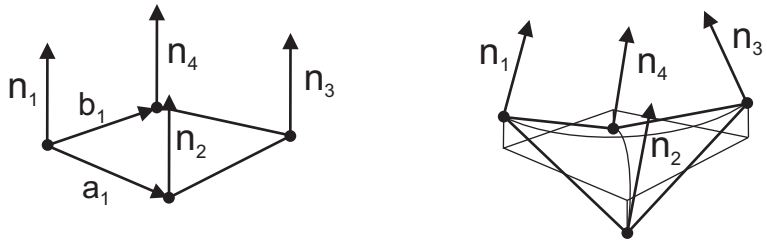

Fig. 10.6 Normal corner vectors for a planar element and a warped element.

10.3 Jacobian matrices

Bases for initial configuration. Consider three bases:

1. $\{\mathbf{i}_k\}$ - the reference (global) Cartesian basis, $k = 1, 2, 3$,
2. $\{\mathbf{g}_k\}$ - the local natural basis at the reference surface for the initial configuration with the tangent vectors defined by eq. (10.15) and the normal vector by eq. (10.29).
3. $\{\mathbf{t}_k\}$ - the local Cartesian basis at the reference surface for the initial configuration with the normal vector defined by eq. (10.19) and the tangent vectors generated in one of the three ways described earlier. For simplicity, we denote $S^3 = \zeta$.

The coordinates associated with these bases are designated as y^k, ξ^k, S^k, respectively. Note that the zero of the natural coordinates ξ^α is at the

element's center, while the zero of the Cartesian coordinates S^α is at any considered Gauss point g at which we define the local bases.

The initial position vector $\mathbf{y}$ relative to the position vector of the Gauss point, $\mathbf{y}_g$, can be expressed in the following three ways:

$$\mathbf{y} - \mathbf{y}_g = (y^k - y^k_g)\,\mathbf{i}_k = (\xi^k - \xi^k_g)\,\mathbf{g}_k = S^k\,\mathbf{t}_k. \tag{10.32}$$

This equation links the above-defined bases and coordinates on the tangent plane spanned at a Gauss point.

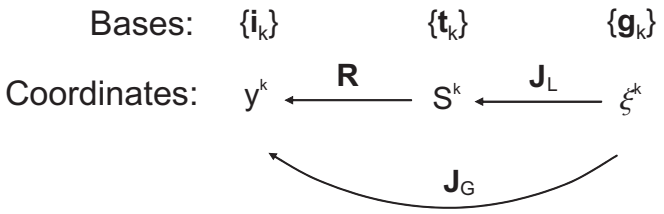

Fig. 10.7 Mappings and Jacobian matrices for the initial configuration.

Jacobian matrices for initial configuration. Let us define the following three types of mappings of coordinates and the Jacobian matrices, see Fig. 10.7:

$$\begin{aligned} \xi^k \mapsto y^l: &\qquad \mathbf{J}_G \doteq \left[\frac{\partial y^l}{\partial \xi^k}\right], \\ S^l \mapsto y^k: &\qquad \mathbf{R} \doteq \left[\frac{\partial y^k}{\partial S^l}\right], \\ \xi^k \mapsto S^l: &\qquad \mathbf{J}_L \doteq \left[\frac{\partial S^l}{\partial \xi^k}\right], \end{aligned} \tag{10.33}$$

where the components of gradients are the matrices arranged as in eq. (2.41).

To obtain the equation linking the above gradients, we use the chain rule of differentiation

$$\frac{\partial y^k}{\partial \xi^l} = \frac{\partial y^k}{\partial S^m}\frac{\partial S^m}{\partial \xi^l}, \quad \text{which yields} \quad \mathbf{J}_G = \mathbf{R}\,\mathbf{J}_L, \tag{10.34}$$

where $k, l, m = 1, 2, 3$.

Rotation matrix $\mathbf{R}$. The angular position of the local basis $\{\mathbf{t}_k\}$ relative to the reference $\{\mathbf{i}_k\}$ is described by the rotation tensor $\mathbf{R} \doteq \mathbf{t}_l \otimes \mathbf{i}_l \in \mathrm{SO}(3)$. This definition implies that $\mathbf{R}\,\mathbf{i}_k = \mathbf{t}_k$, i.e. $\mathbf{t}_k$ is a forward-rotated $\mathbf{i}_k$. The components of $\mathbf{R}$ are

$$R_{jk} = \mathbf{i}_j \cdot (\mathbf{R}\mathbf{i}_k) = \mathbf{i}_j \cdot \mathbf{t}_k \tag{10.35}$$

and, in matrix form,

$$[R_{jk}] = \left[\begin{array}{c|c|c} \mathbf{i}_1 \cdot \mathbf{t}_1 & \mathbf{i}_1 \cdot \mathbf{t}_2 & \mathbf{i}_1 \cdot \mathbf{t}_3 \\ \mathbf{i}_2 \cdot \mathbf{t}_1 & \mathbf{i}_2 \cdot \mathbf{t}_2 & \mathbf{i}_2 \cdot \mathbf{t}_3 \\ \mathbf{i}_3 \cdot \mathbf{t}_1 & \mathbf{i}_3 \cdot \mathbf{t}_2 & \mathbf{i}_3 \cdot \mathbf{t}_3 \end{array}\right] = [\mathbf{t}_1 \,|\, \mathbf{t}_2 \,|\, \mathbf{t}_3]\,, \tag{10.36}$$

where the columns contain components of $\mathbf{t}_k$ in $\{\mathbf{i}_k\}$. The vectors of these components we denote as $\mathbf{t}_k$.

We can show that the Jacobian matrix of the mapping of coordinates $S^l \mapsto y^k$ is equal to the angular position matrix $[R_{jk}]$, i.e.

$$\left[\frac{\partial y^k}{\partial S^j}\right] = [R_{jk}]. \tag{10.37}$$

By eq. (10.32), $(y^j - y_g^j)\mathbf{i}_j = S^k \mathbf{t}_k$, from which we obtain $(y^j - y_g^j) = S^k\,(\mathbf{t}_k \cdot \mathbf{i}_j)$, and the differentiation of both sides w.r.t. S^k yields $\partial y^j / \partial S^k = \mathbf{i}_j \cdot \mathbf{t}_k$, where the r.h.s. is identical as the r.h.s. of eq. (10.35), which ends the proof. □

Global Jacobian matrix $\mathbf{J}_G$. For the mapping of coordinates $\xi^k \mapsto y^i$ of eq. (10.33), the Jacobian matrix is defined as

$$\mathbf{J}_G \doteq \left[\frac{\partial y^i}{\partial \xi^k}\right] = \begin{bmatrix} \frac{\partial y^1}{\partial \xi^1} & \frac{\partial y^1}{\partial \xi^2} & \frac{\partial y^1}{\partial \xi^3} \\ \frac{\partial y^2}{\partial \xi^1} & \frac{\partial y^2}{\partial \xi^2} & \frac{\partial y^2}{\partial \xi^3} \\ \frac{\partial y^3}{\partial \xi^1} & \frac{\partial y^3}{\partial \xi^2} & \frac{\partial y^3}{\partial \xi^3} \end{bmatrix} = [\mathbf{g}_1 \,|\, \mathbf{g}_2 \,|\, \mathbf{g}_3]\,, \tag{10.38}$$

where the columns contain components of $\mathbf{g}_k$ in $\{\mathbf{i}_k\}$. We designate this Jacobian as "global" because the global (reference) Cartesian coordinates y^i are differentiated.

Local Jacobian matrices $\mathbf{J}_L$. For the mapping of coordinates $\xi^k \mapsto S^l$ of eq. (10.33), we define the Jacobian matrix

$$\mathbf{J}_L \doteq \left[\frac{\partial S^i}{\partial \xi^k}\right] = \left[\begin{array}{cc|c} \frac{\partial S^1}{\partial \xi^1} & \frac{\partial S^1}{\partial \xi^2} & 0 \\ \frac{\partial S^2}{\partial \xi^1} & \frac{\partial S^2}{\partial \xi^2} & 0 \\ \hline 0 & 0 & \frac{h}{2} \end{array}\right], \tag{10.39}$$

where the last form is obtained for

$$\frac{\partial S^3}{\partial \xi} = 0, \quad \frac{\partial S^3}{\partial \eta} = 0, \qquad \frac{\partial S^\alpha}{\partial \xi^3} = 0, \quad \frac{\partial S^3}{\partial \xi^3} = \frac{\partial(\xi^3 h/2)}{\partial \xi^3} = \frac{h}{2}. \tag{10.40}$$

We designate this Jacobian as "local", because the local Cartesian coordinates S^i are differentiated.

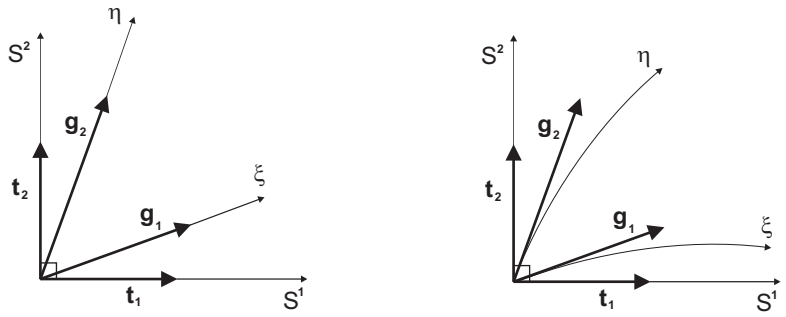

Fig. 10.8 Local bases and coordinates. a) for four-node element, b) for 9-node element with curved boundaries.

Consider only this part of the local mapping $\xi^\beta \mapsto S^\alpha$ $(\alpha, \beta = 1, 2)$ which is related to the tangent plane, see Fig. 10.8. To obtain a Jacobian matrix for this part, we extract the upper 2×2 part of $\mathbf{J}_L$, and denote it as $\mathbf{J}$,

$$\mathbf{J} \doteq \left[\frac{\partial S^\alpha}{\partial \xi^\beta}\right] = \begin{bmatrix} \frac{\partial S^1}{\partial \xi^1} & \frac{\partial S^1}{\partial \xi^2} \\ \frac{\partial S^2}{\partial \xi^1} & \frac{\partial S^2}{\partial \xi^2} \end{bmatrix} = \begin{bmatrix} \mathbf{g}_1 \cdot \mathbf{t}_1 & \mathbf{g}_2 \cdot \mathbf{t}_1 \\ \mathbf{g}_1 \cdot \mathbf{t}_2 & \mathbf{g}_2 \cdot \mathbf{t}_2 \end{bmatrix}, \tag{10.41}$$

where columns contain components of $\mathbf{g}_\beta$ in $\{\mathbf{t}_\alpha\}$.

The last form of $\mathbf{J}$ is obtained as follows. By eq. (10.32), $S^k = (\mathbf{y} - \mathbf{y}_g) \cdot \mathbf{t}_k$, in which $\mathbf{y}$ and $\mathbf{t}_k$ are functions of ξ^α, and the differentiation yields

$$\frac{\partial S^\alpha}{\partial \xi^\beta} = \mathbf{g}_\beta \cdot \mathbf{t}_\alpha + \big(\mathbf{y} - \mathbf{y}_g\big) \cdot \frac{\partial \mathbf{t}_\alpha}{\partial \xi^\beta} \quad \overset{\text{at GP}}{=} \quad \mathbf{g}_\beta \cdot \mathbf{t}_\alpha, \tag{10.42}$$

where $\mathbf{g}_\beta \doteq \partial \mathbf{y}/\partial \xi^\beta$ by eq. (10.15). The last form is valid only for the Gauss point, when $\big(\mathbf{y} - \mathbf{y}_g\big) = \mathbf{0}$, and the second term drops out so we obtain eq. (10.41). The Jacobian at the element center is denoted as $\mathbf{J}_c \doteq \mathbf{J}|_{\xi=\eta=0}$.

Relation between $\mathbf{g}_\alpha$ and $\mathbf{t}_\alpha$. The coordinate gradients imply relations linking the bases with which these coordinates are associated. The natural basis vectors $\mathbf{g}_\alpha$ can be decomposed in the ortho-normal $\{\mathbf{t}_\alpha\}$ as follows:

$$\mathbf{g}_\alpha = (\mathbf{g}_\alpha \cdot \mathbf{t}_1)\, \mathbf{t}_1 + (\mathbf{g}_\alpha \cdot \mathbf{t}_2)\, \mathbf{t}_2, \tag{10.43}$$

in which $(\mathbf{g}_\alpha \cdot \mathbf{t}_\beta)$ are components of $\mathbf{J}$ of eq. (10.41). Hence, we can rewrite

$$\begin{bmatrix} \mathbf{g}_1 \\ \mathbf{g}_2 \end{bmatrix} = \mathbf{J}^T \begin{bmatrix} \mathbf{t}_1 \\ \mathbf{t}_2 \end{bmatrix} \qquad \text{and} \qquad \begin{bmatrix} \mathbf{t}_1 \\ \mathbf{t}_2 \end{bmatrix} = \mathbf{J}^{-T} \begin{bmatrix} \mathbf{g}_1 \\ \mathbf{g}_2 \end{bmatrix}. \tag{10.44}$$

Inverse Jacobian. An inverse of a 2×2 matrix A is given by a simple formula,

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \qquad A^{-1} = \frac{1}{\det A} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}, \tag{10.45}$$

providing $\det A = ad - bc \neq 0$. Applying this formula to the Jacobian matrix $\mathbf{J}$ of eq. (10.41), we obtain

$$\mathbf{J}^{-1} = \frac{1}{\det \mathbf{J}} \begin{bmatrix} \mathbf{g}_2 \cdot \mathbf{t}_2 & -\mathbf{g}_2 \cdot \mathbf{t}_1 \\ -\mathbf{g}_1 \cdot \mathbf{t}_2 & \mathbf{g}_1 \cdot \mathbf{t}_1 \end{bmatrix}, \tag{10.46}$$

where $\det \mathbf{J} \doteq (\mathbf{g}_1 \cdot \mathbf{t}_1)(\mathbf{g}_2 \cdot \mathbf{t}_2) - (\mathbf{g}_1 \cdot \mathbf{t}_2)(\mathbf{g}_1 \cdot \mathbf{t}_1)$.

Another form of the inverse of Jacobian can be obtained with the help of the co-basis $\{\mathbf{g}^\alpha, \mathbf{t}_3\}$. The co-basis vectors $\mathbf{g}^\alpha$ are defined as follows:

$$\mathbf{g}^\alpha : \qquad \mathbf{g}^\alpha \cdot \mathbf{g}_\beta = \delta^\alpha_\beta \qquad \text{and} \qquad \mathbf{g}^\alpha \cdot \mathbf{t}_3 = 0 \tag{10.47}$$

and by analogy with eq. (5.8), they can be calculated as

$$\mathbf{g}^1 = \frac{(\mathbf{g}_2 \times \mathbf{t}_3)}{(\mathbf{g}_2 \times \mathbf{t}_3) \cdot \mathbf{g}_1}, \qquad \mathbf{g}^2 = \frac{(\mathbf{t}_3 \times \mathbf{g}_1)}{(\mathbf{t}_3 \times \mathbf{g}_1) \cdot \mathbf{g}_2}. \tag{10.48}$$

In terms of the co-basis vectors $\mathbf{g}^\alpha$, the inverse of Jacobian is

$$\mathbf{J}^{-1} = \begin{bmatrix} \mathbf{t}_1 \cdot \mathbf{g}^1 & \mathbf{t}_2 \cdot \mathbf{g}^1 \\ \mathbf{t}_1 \cdot \mathbf{g}^2 & \mathbf{t}_2 \cdot \mathbf{g}^2 \end{bmatrix}, \tag{10.49}$$

where columns of $\mathbf{J}^{-1}$ contain components of $\mathbf{t}_\beta$ in $\{\mathbf{g}^\alpha\}$.

Check. We can check that the inverse matrices of eqs. (10.46) and (10.49) are identical. Let us transform the 12-component of the matrix of eq. (10.49) as follows:

$$\mathbf{t}_2 \cdot \mathbf{g}^1 = \frac{\mathbf{t}_2 \cdot (\mathbf{g}_2 \times \mathbf{t}_3)}{(\mathbf{g}_2 \times \mathbf{t}_3) \cdot \mathbf{g}_1} = \frac{\mathbf{g}_2 \cdot (\mathbf{t}_3 \times \mathbf{t}_2)}{(\mathbf{g}_1 \times \mathbf{g}_2) \cdot \mathbf{t}_3} = \frac{-\mathbf{g}_2 \cdot \mathbf{t}_1}{\det \mathbf{J}}, \tag{10.50}$$

where we used $\mathbf{g}_1 \times \mathbf{g}_2 = \mathbf{t}_3 \det \mathbf{J}$, see eq. (10.106) for details. Hence, the obtained expression is identical to the 12-component of eq. (10.46). For the other components of $\mathbf{J}^{-1}$, we can proceed similarly.

Relation between $\mathbf{g}^\alpha$ and $\mathbf{t}_\alpha$. As previously in the derivation of eq. (10.44), we can use the fact that the coordinate gradients imply relations linking the bases with which these coordinates are associated. The co-basis vectors $\mathbf{g}^\alpha$ can be decomposed in the ortho-normal $\{\mathbf{t}_\alpha\}$ as follows:

$$\mathbf{g}^\alpha = (\mathbf{g}^\alpha \cdot \mathbf{t}_1)\, \mathbf{t}_1 + (\mathbf{g}^\alpha \cdot \mathbf{t}_2)\, \mathbf{t}_2, \tag{10.51}$$

in which $(\mathbf{g}^\alpha \cdot \mathbf{t}_\beta)$ are components of the inverse Jacobian matrix $\mathbf{J}^{-1}$ of eq. (10.49). Hence, we can rewrite

$$\begin{bmatrix} \mathbf{g}^1 \\ \mathbf{g}^2 \end{bmatrix} = \mathbf{J}^{-1} \begin{bmatrix} \mathbf{t}_1 \\ \mathbf{t}_2 \end{bmatrix} \qquad \text{and} \qquad \begin{bmatrix} \mathbf{t}_1 \\ \mathbf{t}_2 \end{bmatrix} = \mathbf{J} \begin{bmatrix} \mathbf{g}^1 \\ \mathbf{g}^2 \end{bmatrix}. \tag{10.52}$$

Co-basis definition expressed by Jacobian matrices. The condition defining the co-basis $\mathbf{g}^\alpha \cdot \mathbf{g}_\beta = \delta^\alpha_\beta$ can be rewritten as

$$(\mathbf{t}_i \cdot \mathbf{g}^\alpha)(\mathbf{g}_\beta \cdot \mathbf{t}_i) = \delta^\alpha_\beta, \qquad i = 1, 2, \tag{10.53}$$

where $\mathbf{g}^\alpha = (\mathbf{g}^\alpha \cdot \mathbf{t}_i)\, \mathbf{t}_i$ and $\mathbf{g}_\beta = (\mathbf{g}_\beta \cdot \mathbf{t}_k)\, \mathbf{t}_k$ $(i, k = 1, 2)$, and we can calculate

$$\mathbf{g}^\alpha \cdot \mathbf{g}_\beta = (\mathbf{g}^\alpha \cdot \mathbf{t}_i)(\mathbf{g}_\beta \cdot \mathbf{t}_k)\, \mathbf{t}_i \cdot \mathbf{t}_k = (\mathbf{t}_i \cdot \mathbf{g}^\alpha)(\mathbf{g}_\beta \cdot \mathbf{t}_i). \tag{10.54}$$

We note that $\mathbf{g}_\beta \cdot \mathbf{t}_i = \partial S^i / \partial \xi^\beta$ by eq. (10.42) and, hence, on the basis of eq. (10.53), we can define the gradient

$$\frac{\partial \xi^\alpha}{\partial S^i} \doteq \mathbf{t}_i \cdot \mathbf{g}^\alpha. \tag{10.55}$$

On the other hand, by eq. (10.49), $\mathbf{g}^\alpha \cdot \mathbf{t}_i = \left[\mathbf{J}^{-1}\right]_{\alpha i}$ and, hence,

$$\mathbf{J}^{-1} = \begin{bmatrix} \mathbf{t}_1 \cdot \mathbf{g}^1 & \mathbf{t}_2 \cdot \mathbf{g}^1 \\ \mathbf{t}_1 \cdot \mathbf{g}^2 & \mathbf{t}_2 \cdot \mathbf{g}^2 \end{bmatrix} = \begin{bmatrix} \frac{\partial \xi^1}{\partial S^1} & \frac{\partial \xi^1}{\partial S^2} \\ \frac{\partial \xi^2}{\partial S^1} & \frac{\partial \xi^2}{\partial S^2} \end{bmatrix} \doteq \left[\frac{\partial \xi^\alpha}{\partial S^i}\right]. \tag{10.56}$$

Therefore, eq. (10.53) can be rewritten simply as $\mathbf{J}^{-1}\mathbf{J} = \mathbf{I}$, where $\mathbf{I}$ is the identity matrix.

Example. Note that the procedure of calculation of $\mathbf{J}^{-1}$ allows us to avoid expressing explicitly the natural coordinates ξ^α in terms of the ortho-normal coordinates S^α. This is an advantage because such relations can be quite complicated. For instance, for **Basis 1** attached at the element center, these relations are as follows:

$$\mathbf{g}_1\, \xi^1 = \mathbf{t}_1\, (S^1 - S^2 \tan\beta), \qquad \mathbf{g}_2\, \xi^2 = \mathbf{t}_1 S^2 \tan\beta + \mathbf{t}_2 S^2, \tag{10.57}$$

where $\beta < \pi/2$ is the angle between $\mathbf{t}_2$ and $\mathbf{g}_2$ and $\tan\beta = (\mathbf{t}_1 \cdot \mathbf{g}_2)/(\mathbf{t}_2 \cdot \mathbf{g}_2)$. Then

$$\mathbf{J}^{-1} = \begin{bmatrix} \frac{\partial \xi^1}{\partial S^1} & \frac{\partial \xi^2}{\partial S^1} \\ \frac{\partial \xi^1}{\partial S^2} & \frac{\partial \xi^2}{\partial S^2} \end{bmatrix} = \begin{bmatrix} 1/\sqrt{g_{11}} & 0 \\ -\tan\beta/\sqrt{g_{11}} & 1/(\cos\beta\sqrt{g_{22}}) \end{bmatrix}, \tag{10.58}$$

where $g_{\alpha\alpha} \doteq \mathbf{g}_\alpha \cdot \mathbf{g}_\alpha$ and $\cos\beta = (\mathbf{t}_2 \cdot \mathbf{g}_2)/\sqrt{g_{22}}$.

Local Jacobian and its inverse for the vector of shape functions. We can approximate the relative vector $(\mathbf{y} - \mathbf{y}_g)$, see eq. (10.32), using a vector of shape functions $\mathbf{N}(\xi^1, \xi^2)$ of eq. (10.5), and write

$$S^\alpha = \frac{1}{4}\left[(\mathbf{s}\mathbf{S}^\alpha) + (\boldsymbol{\xi}\mathbf{S}^\alpha)\,\xi + (\boldsymbol{\eta}\mathbf{S}^\alpha)\,\eta + (\mathbf{h}\mathbf{S}^\alpha)\,\xi\eta\right], \tag{10.59}$$

where $\mathbf{S}^\alpha \doteq (\mathbf{y}_I - \mathbf{y}_g)\cdot\mathbf{t}_\alpha = [S_1^\alpha,\, S_2^\alpha,\, S_3^\alpha,\, S_4^\alpha]^T$ is the vector of projections of nodal relative position vectors on $\mathbf{t}_\alpha$. Differentiating eq. (10.59) w.r.t. the natural coordinates, we obtain

$$\frac{\partial S^\alpha}{\partial \xi} = \frac{1}{4}\left[\,(\boldsymbol{\xi}\mathbf{S}^\alpha) + (\mathbf{h}\mathbf{S}^\alpha)\,\eta\,\right], \qquad \frac{\partial S^\alpha}{\partial \eta} = \frac{1}{4}\left[\,(\boldsymbol{\eta}\mathbf{S}^\alpha) + (\mathbf{h}\mathbf{S}^\alpha)\,\xi\,\right] \tag{10.60}$$

and, hence, the Jacobian matrix of eq. (10.41) is

$$\mathbf{J} = \begin{bmatrix} \frac{1}{4}\left[\,(\boldsymbol{\xi}\mathbf{S}^1) + (\mathbf{h}\mathbf{S}^1)\,\eta\,\right] & \frac{1}{4}\left[\,(\boldsymbol{\eta}\mathbf{S}^1) + (\mathbf{h}\mathbf{S}^1)\,\xi\,\right] \\ \frac{1}{4}\left[\,(\boldsymbol{\xi}\mathbf{S}^2) + (\mathbf{h}\mathbf{S}^2)\,\eta\,\right] & \frac{1}{4}\left[\,(\boldsymbol{\eta}\mathbf{S}^2) + (\mathbf{h}\mathbf{S}^2)\,\xi\,\right] \end{bmatrix}. \tag{10.61}$$

Note that column 1 varies linearly with η, while column 2 varies linearly with ξ.

An inverse of the Jacobian can be obtained by eq. (10.45) and it is as follows:

$$\mathbf{J}^{-1} = \frac{1}{\det \mathbf{J}} \begin{bmatrix} \frac{1}{4}\left[\,(\boldsymbol{\eta}\mathbf{S}^2) + (\mathbf{h}\mathbf{S}^2)\,\xi\,\right] & -\frac{1}{4}\left[\,(\boldsymbol{\eta}\mathbf{S}^1) + (\mathbf{h}\mathbf{S}^1)\,\xi\,\right] \\ -\frac{1}{4}\left[\,(\boldsymbol{\xi}\mathbf{S}^2) + (\mathbf{h}\mathbf{S}^2)\,\eta\,\right] & \frac{1}{4}\left[\,(\boldsymbol{\xi}\mathbf{S}^1) + (\mathbf{h}\mathbf{S}^1)\,\eta\,\right] \end{bmatrix}, \tag{10.62}$$

where

$$\det \mathbf{J} = J(\xi, \eta) = J_0 + J_1\xi + J_2\eta, \tag{10.63}$$

and its components are

$$J_0 \doteq \frac{1}{16}\left[(\boldsymbol{\xi}\mathbf{S}^1)\,(\boldsymbol{\eta}\mathbf{S}^2) - (\boldsymbol{\eta}\mathbf{S}^1)\,(\boldsymbol{\xi}\mathbf{S}^2)\right],$$

$$J_1 \doteq \frac{1}{16}\left[(\boldsymbol{\xi}\mathbf{S}^1)\,(\mathbf{h}\mathbf{S}^2) - (\mathbf{h}\mathbf{S}^1)\,(\boldsymbol{\xi}\mathbf{S}^2)\right],$$

$$J_2 \doteq \frac{1}{16}\left[(\mathbf{h}\mathbf{S}^1)\,(\boldsymbol{\eta}\mathbf{S}^2) - (\boldsymbol{\eta}\mathbf{S}^1)\,(\mathbf{h}\mathbf{S}^2)\right].$$

The bilinear term $J_{12}\,\xi\eta$ is not present in the expansion eq. (10.63) because

$$J_{12} \doteq \frac{1}{16}\left[(\mathbf{h}\mathbf{S}^1)\,(\mathbf{h}\mathbf{S}^2) - (\mathbf{h}\mathbf{S}^1)\,(\mathbf{h}\mathbf{S}^2)\right] = 0.$$

It can be checked for parallelograms that only $J_0 \neq 0$, while $J_1 = J_2 = 0$.

Example. Jacobian matrices for basic shapes of element. The Jacobian matrix contains information about the initial shape of the element. In Fig. 10.9, we show several basic shapes of a planar four-node element. The Jacobian matrix and its determinant for these shapes are as follows:

a) square 2×2: $\mathbf{J} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $\det \mathbf{J} = 1$,

b) rectangle: $\mathbf{J} = \begin{bmatrix} 5/2 & 0 \\ 0 & 1 \end{bmatrix}$, $\det \mathbf{J} = 5/2$,

c) parallelogram: $\mathbf{J} = \begin{bmatrix} 5/2 & 0 \\ 1 & 1 \end{bmatrix}$, $\det \mathbf{J} = 5/2$,

d) trapezoid: $\mathbf{J} = \begin{bmatrix} 5/2+\eta & 0 \\ \xi & 1 \end{bmatrix}$, $\det \mathbf{J} = 5/2+\eta$,

e) trapezoid: $\mathbf{J} = \begin{bmatrix} (5+\eta)/2 & 0 \\ (1+\xi)/2 & 1 \end{bmatrix}$, $\det \mathbf{J} = (5+\eta)/2$,

f) irregular: $\mathbf{J} = \begin{bmatrix} (5+\eta)/2 & (1+\eta)/4 \\ (1+\xi)/2 & (5+\xi)/4 \end{bmatrix}$, $\det \mathbf{J} = (6+\xi+\eta)/2$.

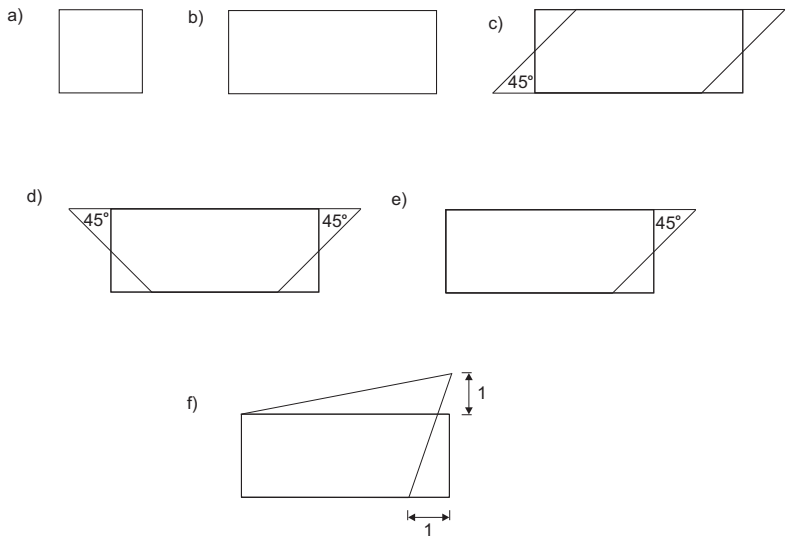

Fig. 10.9 Typical shapes of four-node elements.
Elements c), d), e), f) are obtained from the rectangle of size 5×2.

We see that $\mathbf{J}$ is a diagonal matrix only for the square and the rectangle. For the square, the rectangle, and the parallelogram, $\mathbf{J}$ is constant over the element.

For trapezoids and irregular elements, $\mathbf{J}$ is non-diagonal and non-constant and linearly depends on ξ, η. Note that for non-rectangular shapes, especially when the element aspect ratio is far from 1, the accuracy of four-node elements substantially decreases, see tests of Sects. 15.2.4 and 15.2.6.

10.4 Deformation gradient, $\mathbf{F}^T\mathbf{F}$ and $\mathbf{Q}^T\mathbf{F}$ products

Using various bases in FE computations. To enable linking of finite elements of various spatial orientation, the displacement vector $\mathbf{u}$ and the rotation vector $\boldsymbol{\psi}$ are represented in the reference Cartesian basis $\{\mathbf{i}_k\}$. However, in computations on the level of an element, we have a choice and one of the following two bases can be used:

1. the reference Cartesian basis $\{\mathbf{i}_k\}$. In order to perform the local operations, we have to transform strain components and components of the $\mathbf{Q}^T\mathbf{F}$ product to the local basis $\{\mathbf{t}_k\}$ at a Gauss point.
2. the elemental Cartesian basis $\{\mathbf{t}_k^c\}$ at the element center. Then, first, the displacement and rotation components must be transformed from the reference basis $\{\mathbf{i}_k\}$ to this basis. In order to perform the local operations, we transform the strain components and components of the $\mathbf{Q}^T\mathbf{F}$ product to the local basis $\{\mathbf{t}_k\}$ at a Gauss point. Afterwards, the tangent stiffness matrix at the residual vector must be transformed to the reference basis $\{\mathbf{i}_k\}$.
   The use of this basis requires additional transformations but, contrary to expectations, can significantly improve the efficiency of a four-node element if zero values are accounted for in the implementation.
   Note that the use of the elemental basis $\{\mathbf{t}_k^c\}$ is indispensable in the case of warped four-node elements if the substitute flat element and the warpage correction are used, see Sect. 14.3 for details.

Local operations for shell element. The shell assumptions and several techniques related to the formulation of a finite element require a local definition of directions, e.g.

- the Reissner hypothesis,
- imposition of the zero normal stress (ZNS) condition,
- modifications of transverse shear strains (using the ANS method),
- formulation of the drilling RC equation,
- integration of the strain energy, when it is separated into the integration in the normal (fiber) direction and the integration in the tangent (lamina) direction.

Hence, no matter whether the basis $\{\mathbf{i}_k\}$ or $\{\mathbf{t}_k^c\}$ is used on the element's level, we have to transform the strain components and components of the $\mathbf{Q}^T\mathbf{F}$ product to the local basis $\{\mathbf{t}_k\}$ at a Gauss point. Note that instead of transforming components from one basis to another, we can operate on the backward-rotated objects, as described in Sect. 2.

Remarks on use of a skew basis at element's center. The skew basis at the element center, $\{\mathbf{g}^c_\alpha, \mathbf{t}^c_3\}$, is used in mixed elements based on the Hellinger–Reissner functional and the Hu–Washizu functional, see Sect. 11.5. Representations of stress are assumed in this basis, while representations of strain for the Hu–Washizu functional are assumed either in this basis or in its co-basis. These representations are next transformed to the local orthonormal basis at the element center $\{\mathbf{t}^c_k\}$.

The formulas for a transformation between a non-orthogonal basis and a Cartesian basis are derived in Sect. 2. In the case of the in-plane $(\alpha\beta)$ components, for non-symmetric tensors we use eqs. (2.21) and (2.25), while for symmetric tensors we use eq. (2.29) with $\mathbf{T}^*$ replaced by $\mathbf{T}^{**}$ of eq. (2.30) or eq. (2.31). For the transverse $(\alpha 3)$ components, we use eqs. (2.26) and (2.27).

Deformation gradient, $\mathbf{F}^T\mathbf{F}$ and $\mathbf{Q}^T\mathbf{F}$ products. Below, we derive matrices of components for the deformation gradient $\mathbf{F}$, the Cauchy–Green deformation tensor $\mathbf{C} \doteq \mathbf{F}^T\mathbf{F}$, and the $\mathbf{Q}^T\mathbf{F}$ product. They are derived for the reference basis $\{\mathbf{i}_k\}$ and, subsequently, the latter two matrices are transformed to the local bases $\{\mathbf{t}_k\}$. As mentioned earlier, instead of the reference basis $\{\mathbf{i}_k\}$, the elemental basis $\{\mathbf{t}^c_k\}$ can be used as well. Two ways of derivation are presented below in which the position vectors are treated as functions of either (A) the natural coordinates, or (B) the local Cartesian coordinates.

(A) Natural coordinates. For the coordinates $\{\xi^\alpha, \zeta\}$, we take $\zeta = \frac{h}{2}\xi^3$, express $\zeta \in [-h/2, +h/2]$ in terms of $\xi^3 \in [-1, +1]$, and use the natural coordinates $\{\xi^k\}$ $(k = 1, 2, 3)$.

Then the position vector in the initial configuration of eq. (5.1) is as follows:

$$\mathbf{y}(\xi^k) = \mathbf{y}_0(\xi^\alpha) + \frac{h}{2}\xi^3\ \mathbf{t}_3(\xi^\alpha), \qquad \alpha = 1, 2, \tag{10.64}$$

and the current position vector is $\mathbf{x} = \mathbf{x}(\xi^k(\mathbf{y}))$. The deformation gradient of eq. (5.15) can be written simply as

$$\mathbf{F} \doteq \frac{\partial \mathbf{x}}{\partial \mathbf{y}} = \frac{\partial \mathbf{x}}{\partial \xi^k} \otimes \frac{\partial \xi^k}{\partial \mathbf{y}}, \tag{10.65}$$

with ξ^k serving as intermediate variables. Let us use the components of $\mathbf{y}$ and $\mathbf{x}$ in the reference basis $\{\mathbf{i}_k\}$. Then $\mathbf{y} = y^m\,\mathbf{i}_m$ and $\mathbf{x} = x^l\,\mathbf{i}_l$, $(m, l = 1, 2, 3)$ and we differentiate only the components,

$$\frac{\partial \mathbf{x}}{\partial \xi^k} = \frac{\partial x^l}{\partial \xi^k}\,\mathbf{i}_l, \qquad \frac{\partial \mathbf{y}}{\partial \xi^k} = \frac{\partial y^m}{\partial \xi^k}\,\mathbf{i}_m. \tag{10.66}$$

Hence, the inverse derivative is

$$\frac{\partial \xi^k}{\partial \mathbf{y}} = \frac{\partial \xi^k}{\partial y^m}\,\mathbf{i}_m \tag{10.67}$$

and the deformation gradient becomes

$$\mathbf{F} = \frac{\partial \mathbf{x}}{\partial \xi^k} \otimes \frac{\partial \xi^k}{\partial \mathbf{y}} = \frac{\partial x^l}{\partial \xi^k}\frac{\partial \xi^k}{\partial y^m}\,\mathbf{i}_l \otimes \mathbf{i}_m = F_{lm}\,\mathbf{i}_l \otimes \mathbf{i}_m, \tag{10.68}$$

where

$$F_{lm} \doteq \frac{\partial x^l}{\partial \xi^k}\frac{\partial \xi^k}{\partial y^m}. \tag{10.69}$$

Let us introduce the matrices of components

$$\mathbf{F} \doteq [F_{lm}], \qquad \mathbf{J}_G^{\text{curr}} \doteq \left[\frac{\partial x^l}{\partial \xi^k}\right], \qquad \mathbf{J}_G \doteq \left[\frac{\partial y^m}{\partial \xi^k}\right], \tag{10.70}$$

where $\mathbf{J}_G$ is the global Jacobian of eq. (10.33). Then the deformation gradient matrix can be computed as

$$\mathbf{F} = \mathbf{J}_G^{\text{curr}}\,\mathbf{J}_G^{-1}. \tag{10.71}$$

Now we can compute the components of the Cauchy–Green tensor $\mathbf{C} \doteq \mathbf{F}^T\mathbf{F}$ and of the $\mathbf{Q}_0^T\mathbf{F}$ product, and transform them to the local orthonormal basis $\{\mathbf{t}_k\}$. This can be done, as derived in Sect. 2, by the transformation of components $(\,\cdot\,)_{\text{local}} = \mathbf{R}^T\,(\,\cdot\,)_{\text{global}}\,\mathbf{R}$, see eq. (2.13). Note that $\mathbf{R}$ is the rotation matrix of eq. (10.33).

1. Components of the Cauchy–Green deformation tensor $\mathbf{C} \doteq \mathbf{F}^T\,\mathbf{F}$,

$$\mathbf{C}_* = \mathbf{R}^T\,\mathbf{C}\,\mathbf{R} = \bar{\mathbf{F}}^T\bar{\mathbf{F}}, \tag{10.72}$$

2. Components of the $\mathbf{Q}^T\mathbf{F}$ product, i.e. $\mathbf{Q}^T\mathbf{F}$,

$$(\mathbf{Q}^T\mathbf{F})_* = \mathbf{R}^T\,(\mathbf{Q}^T\mathbf{F})\,\mathbf{R} = (\mathbf{Q}\,\mathbf{R})^T\,\bar{\mathbf{F}}, \tag{10.73}$$

where

$$\bar{\mathbf{F}} \doteq \mathbf{F}\mathbf{R} = \mathbf{J}_G^{\text{curr}}\mathbf{J}_G^{-1}\,\mathbf{R} = \mathbf{J}_G^{\text{curr}}\mathbf{J}_L^{-1}, \tag{10.74}$$

with the last form obtained on use of eq. (10.34).

(B) Orthonormal local coordinates. We can express $\zeta \in [-h/2, +h/2]$ in terms of $\xi^3 \in [-1, +1]$ as $\zeta = (h/2)\,\xi^3$, and define $S^3 \doteq \xi^3$. Then we can use the Cartesian coordinates $\{S^k\}$ $(k = 1, 2, 3)$ instead of $\{S^\alpha, \zeta\}$.

The position vector in the initial configuration of eq. (5.1) is now as follows:

$$\mathbf{y}(S^k) = \mathbf{y}_0(S^\alpha) + \frac{h}{2} S^3\ \mathbf{t}_3(S^\alpha), \qquad \alpha = 1, 2, \tag{10.75}$$

and the current position vector is $\mathbf{x} = \mathbf{x}(S^k(\mathbf{y}))$. The deformation gradient of eq. (5.15) can be written simply as

$$\mathbf{F} \doteq \frac{\partial \mathbf{x}}{\partial \mathbf{y}} = \frac{\partial \mathbf{x}}{\partial S^k} \otimes \frac{\partial S^k}{\partial \mathbf{y}}, \tag{10.76}$$

with S^k being intermediate variables. Let us use the components of $\mathbf{y}$ and $\mathbf{x}$ in the reference basis $\{\mathbf{i}_k\}$. Then $\mathbf{y} = y^m\,\mathbf{i}_m$ and $\mathbf{x} = x^l\,\mathbf{i}_l$ and we can differentiate the components

$$\frac{\partial \mathbf{x}}{\partial S^k} = \frac{\partial x^l}{\partial S^k}\,\mathbf{i}_l, \qquad \frac{\partial \mathbf{y}}{\partial S^k} = \frac{\partial y^m}{\partial S^k}\,\mathbf{i}_m. \tag{10.77}$$

Hence, the inverse derivative is

$$\frac{\partial S^k}{\partial \mathbf{y}} = \frac{\partial S^k}{\partial y^m}\,\mathbf{i}_m \tag{10.78}$$

and the deformation gradient becomes

$$\mathbf{F} = \frac{\partial \mathbf{x}}{\partial S^k} \otimes \frac{\partial S^k}{\partial \mathbf{y}} = \frac{\partial x^l}{\partial S^k}\frac{\partial S^k}{\partial y^m}\,\mathbf{i}_l \otimes \mathbf{i}_m = F_{lm}\,\mathbf{i}_l \otimes \mathbf{i}_m, \tag{10.79}$$

where

$$F_{lm} \doteq \frac{\partial x^l}{\partial S^k}\frac{\partial S^k}{\partial y^m}. \tag{10.80}$$

Let us define the following matrices of components

$$\mathbf{F} \doteq [F_{lm}], \qquad \nabla\mathbf{x} \doteq \left[\frac{\partial x^l}{\partial S^k}\right], \qquad \mathbf{R} \doteq \left[\frac{\partial y^m}{\partial S^k}\right], \tag{10.81}$$

where $\mathbf{R}$ is the rotation matrix of eq. (10.33). Then the deformation gradient matrix can be computed as

$$\mathbf{F} = \nabla\mathbf{x}\ \mathbf{R}^T. \tag{10.82}$$

Now we can compute the components of the Cauchy–Green tensor $\mathbf{C} \doteq \mathbf{F}^T\mathbf{F}$ and of the $\mathbf{Q}_0^T\mathbf{F}$ product and transform them to the local orthonormal basis $\{\mathbf{t}_k\}$. This can be done, as derived in Sect. 2, by the transformation of components $(\,\cdot\,)_{\mathrm{L}} = \mathbf{R}^T\,(\,\cdot\,)_{\mathrm{G}}\,\mathbf{R}$, see eq. (2.13).

1. Components of the Cauchy–Green deformation tensor $\mathbf{C} \doteq \mathbf{F}^T\,\mathbf{F}$,

$$\mathbf{C}_* = \mathbf{R}^T\,\mathbf{C}\,\mathbf{R} = (\nabla\mathbf{x})^T\nabla\mathbf{x}, \tag{10.83}$$

2. Components of the $\mathbf{Q}^T\mathbf{F}$ product, i.e. $\mathbf{Q}^T\mathbf{F}$,

$$(\mathbf{Q}^T\mathbf{F})_* = \mathbf{R}^T\,(\mathbf{Q}^T\mathbf{F})\,\mathbf{R} = (\mathbf{Q}\,\mathbf{R})^T\,\nabla\mathbf{x}. \tag{10.84}$$

Note that, formally, $\nabla\mathbf{x}$ plays the same role as $\bar{\mathbf{F}}$ in case (A).

The formulation based on the coordinates S^k makes sense when the derivatives of shape functions are expressed in terms of ortho-normal S^α, as, e.g., for the one-integration point element. Besides, it is an analogue of the formulation used in the analytical studies in Sect. 6.

Increment of Green strain. Version 1. For $\mathbf{x} = \mathbf{x}_n + \Delta\mathbf{u}$, where n refers to the last known configuration and Δ to the increment from the last known configuration to the current one, the deformation gradient can be multiplicatively decomposed as follows:

$$\mathbf{F} \doteq \frac{\partial\mathbf{x}}{\partial\mathbf{y}} = \frac{\partial\mathbf{x}}{\partial\mathbf{x}_n}\frac{\partial\mathbf{x}_n}{\partial\mathbf{y}} = \Delta\mathbf{F}\,\mathbf{F}_n. \tag{10.85}$$

Then, the Green strain can be rewritten as

$$\mathbf{E} \doteq \tfrac{1}{2}(\mathbf{F}^T\mathbf{F} - \mathbf{I}) = \tfrac{1}{2}\left[\mathbf{F}_n^T\,(\Delta\mathbf{F}^T\Delta\mathbf{F})\,\mathbf{F}_n - \mathbf{I}\right], \tag{10.86}$$

where

$$\Delta\mathbf{F} \doteq \frac{\partial\mathbf{x}}{\partial\mathbf{x}_n} = \nabla_n\mathbf{x} = \mathbf{I} + \nabla_n(\Delta\mathbf{u}) \tag{10.87}$$

and $\nabla_n(\cdot) \doteq \partial(\cdot)/\partial\mathbf{x}_n$ denotes the gradient w.r.t. the known position vector. We can linearize the $\Delta\mathbf{F}^T\Delta\mathbf{F}$ product w.r.t. $\Delta\mathbf{u}$, which yields

$$\begin{aligned}(\Delta\mathbf{F}^T\Delta\mathbf{F}) &= \mathbf{I} + \nabla_n(\Delta\mathbf{u}) + \nabla_n^T(\Delta\mathbf{u}) + \underbrace{\nabla_n^T(\Delta\mathbf{u})\nabla_n(\Delta\mathbf{u})}_{\text{neglected}}\\ &\approx \mathbf{I} + 2\Delta\boldsymbol{\varepsilon},\end{aligned} \tag{10.88}$$

where $\Delta\boldsymbol{\varepsilon} \doteq \mathrm{sym}\nabla_n(\Delta\mathbf{u})$ is the infinitesimal strain increment. Hence, the Green strain can be expressed as

$$\mathbf{E} = \tfrac{1}{2}\left[\mathbf{F}_n^T\,(\mathbf{I} + 2\Delta\boldsymbol{\varepsilon})\,\mathbf{F}_n - \mathbf{I}\right] = \mathbf{E}_n + \mathbf{F}_n^T\,\Delta\boldsymbol{\varepsilon}\,\mathbf{F}_n, \tag{10.89}$$

where $\mathbf{E}_n \doteq \frac{1}{2}(\mathbf{F}_n^T\mathbf{F}_n - \mathbf{I})$, and the increment of the Green strain can be obtained as the pull-back of the infinitesimal strain increment, $\Delta\mathbf{E} \doteq \mathbf{E} - \mathbf{E}_n = \mathbf{F}_n^T\,\Delta\boldsymbol{\varepsilon}\,\mathbf{F}_n$.

Increment of Green strain. Version 2. The increment of the Green strain can be defined as $\Delta\mathbf{E} \doteq \mathbf{E}_{n+1} - \mathbf{E}_n$ and expressed as

$$\Delta\mathbf{E} = \tfrac{1}{2}(\mathbf{F}_{n+1}^T\mathbf{F}_{n+1} - \mathbf{F}_n^T\mathbf{F}_n). \tag{10.90}$$

The deformation gradients at t^n and t^{n+1} can be expressed by the mid-point deformation gradient $\mathbf{F}_{n+1/2} \doteq \partial\mathbf{x}_{n+1/2}/\partial\mathbf{y}$ as follows:

$$\mathbf{F}_{n+1} \doteq \frac{\partial\mathbf{x}_{n+1}}{\partial\mathbf{X}} = \frac{\partial\mathbf{x}_{n+1}}{\partial\mathbf{x}_{n+1/2}}\frac{\partial\mathbf{x}_{n+1/2}}{\partial\mathbf{y}} = \left[\frac{\partial\mathbf{x}_{n+1}}{\partial\mathbf{x}_{n+1/2}}\right]\mathbf{F}_{n+1/2}, \tag{10.91}$$

$$\mathbf{F}_n \doteq \frac{\partial\mathbf{x}_n}{\partial\mathbf{X}} = \frac{\partial\mathbf{x}_n}{\partial\mathbf{x}_{n+1/2}}\frac{\partial\mathbf{x}_{n+1/2}}{\partial\mathbf{y}} = \left[\frac{\partial\mathbf{x}_{n+1/2}}{\partial\mathbf{x}_n}\right]^{-1}\mathbf{F}_{n+1/2}. \tag{10.92}$$

Then, the increment of the Green strain can be written as

$$\Delta\mathbf{E} = \mathbf{F}_{n+1/2}^T\,\Delta\boldsymbol{\varepsilon}\;\mathbf{F}_{n+1/2}, \tag{10.93}$$

where the part which is pushed-forward to the mid-point position is

$$2\Delta\boldsymbol{\varepsilon} = \left[\frac{\partial\mathbf{x}_{n+1}}{\partial\mathbf{x}_{n+1/2}}\right]^T\left[\frac{\partial\mathbf{x}_{n+1}}{\partial\mathbf{x}_{n+1/2}}\right] - \left[\frac{\partial\mathbf{x}_n}{\partial\mathbf{x}_{n+1/2}}\right]^T\left[\frac{\partial\mathbf{x}_n}{\partial\mathbf{x}_{n+1/2}}\right]. \tag{10.94}$$

We note that

$$\frac{\partial\mathbf{x}_{n+1}}{\partial\mathbf{x}_{n+1/2}} = \frac{\partial(\mathbf{x}_{n+1/2} + \frac{1}{2}\Delta\mathbf{u})}{\partial\mathbf{x}_{n+1/2}} = \mathbf{I} + \tfrac{1}{2}\frac{\partial\Delta\mathbf{u}}{\partial\mathbf{x}_{n+1/2}} \tag{10.95}$$

and

$$\frac{\partial\mathbf{x}_n}{\partial\mathbf{x}_{n+1/2}} = \frac{\partial(\mathbf{x}_{n+1/2} - \frac{1}{2}\Delta\mathbf{u})}{\partial\mathbf{x}_{n+1/2}} = \mathbf{I} - \tfrac{1}{2}\frac{\partial\Delta\mathbf{u}}{\partial\mathbf{x}_{n+1/2}}. \tag{10.96}$$

Using the above relations in eq. (10.94), we obtain

$$\Delta\boldsymbol{\varepsilon} = \tfrac{1}{2}\left[\frac{\partial\Delta\mathbf{u}}{\partial\mathbf{x}_{n+1/2}} + \left(\frac{\partial\Delta\mathbf{u}}{\partial\mathbf{x}_{n+1/2}}\right)^T\right]. \tag{10.97}$$

Formula (10.93) is used in finite strain plasticity, e.g., in [97, 250].

Rate of Green strain. Differentiation of the Green strain $\mathbf{E} \doteq \frac{1}{2}(\mathbf{F}^T\mathbf{F} - \mathbf{I})$ w.r.t. time t, yields

$$\begin{aligned} 2\dot{\mathbf{E}} &= \dot{\mathbf{F}}^T\mathbf{F} + \mathbf{F}^T\dot{\mathbf{F}} = \mathbf{F}^T(\mathbf{F}^{-T}\dot{\mathbf{F}}^T + \dot{\mathbf{F}}\mathbf{F}^{-1})\mathbf{F} \\ &= \mathbf{F}^T(\nabla\mathbf{v}^T + \nabla\mathbf{v})\mathbf{F} = 2\mathbf{F}^T\mathbf{d}\,\mathbf{F}, \end{aligned} \tag{10.98}$$

where the spatial velocity gradient $\nabla\mathbf{v} \doteq \dot{\mathbf{F}}\mathbf{F}^{-1}$, and the rate of deformation $\mathbf{d} \doteq \frac{1}{2}(\nabla\mathbf{v}^T + \nabla\mathbf{v})$. Using this formula, we can obtain an interpretation of the above two forms of increment of the Green strain.

Writing eq. (10.98) at time instant t^n and multiplying by Δt, we obtain

$$\Delta\mathbf{E} = \mathbf{F}_n^T\,\Delta\boldsymbol{\varepsilon}\,\mathbf{F}_n, \tag{10.99}$$

where $\Delta\mathbf{E} \doteq \dot{\mathbf{E}}_n\Delta t$ and $\Delta\boldsymbol{\varepsilon} \doteq \mathbf{d}_n\Delta t$ by the forward Euler finite-difference scheme, which is first-order accurate. This equation corresponds to eq. (10.89).

Writing eq. (10.98) at time instant $t^{n+1/2}$ and multiplying by Δt, we obtain

$$\Delta\mathbf{E} = \mathbf{F}_{n+1/2}^T\,\Delta\boldsymbol{\varepsilon}\,\mathbf{F}_{n+1/2}, \tag{10.100}$$

where $\Delta\mathbf{E} \doteq \dot{\mathbf{E}}_{n+1/2}\Delta t$ and $\Delta\boldsymbol{\varepsilon} \doteq \mathbf{d}_{n+1/2}\Delta t$ by the central finite-difference scheme, which is second-order accurate. This equation corresponds to eq. (10.93).

10.5 Numerical integration of shell elements

Infinitesimal volume and area of shell element. Below we consider the formulas suitable for (i) the ortho-normal coordinates S^k and (ii) the natural coordinates ξ^k. The latter are actually used in our computations.

(i) Ortho-normal coordinates S^k. The differential of the position vector $\mathbf{y}$ can be written as

$$\mathrm{d}\mathbf{y} \doteq \mathbf{t}_1\,\mathrm{d}S^1 + \mathbf{t}_2\,\mathrm{d}S^2 + \mathbf{t}_3\,\mathrm{d}S^3, \tag{10.101}$$

where $\mathrm{d}S^3 \doteq \zeta \in [-h/2, +h/2]$. An infinitesimal volume of the rectangular parallelepiped spanned by vectors $(\mathrm{d}\mathbf{y})_i = \mathbf{t}_i\,\mathrm{d}S^i$ (no summation) is as follows:

$$\mathrm{d}V \doteq (\mathbf{t}_1\,\mathrm{d}S^1 \times \mathbf{t}_2\,\mathrm{d}S^2)\cdot(\mathbf{t}_3\,\mathrm{d}S^3) = (\mathbf{t}_1 \times \mathbf{t}_2)\cdot\mathbf{t}_3\,\mathrm{d}S^1\mathrm{d}S^2\mathrm{d}S^3 = \mathrm{d}S^1\mathrm{d}S^2\mathrm{d}S^3, \tag{10.102}$$

as $\mathbf{t}_1 \times \mathbf{t}_2 = \mathbf{t}_3$. The infinitesimal area of a rectangle spanned by the tangent vectors $(\mathrm{d}\mathbf{y})_\alpha$ $(\alpha = 1, 2)$ is

$$\mathrm{d}A \doteq (\mathbf{t}_1\,\mathrm{d}S^1 \times \mathbf{t}_2\,\mathrm{d}S^2) \cdot \mathbf{t}_3 = \mathrm{d}S^1\mathrm{d}S^2. \tag{10.103}$$

(ii) Natural coordinates ξ^k. The differential of the position vector $\mathbf{y}$ can be written as

$$\mathrm{d}\mathbf{y} \doteq \frac{\partial \mathbf{y}_0}{\partial \xi^1}\,\mathrm{d}\xi^1 + \frac{\partial \mathbf{y}_0}{\partial \xi^2}\,\mathrm{d}\xi^2 + \mathbf{t}_3\frac{h}{2}\,\mathrm{d}\xi^3 = \mathbf{g}_1\,\mathrm{d}\xi^1 + \mathbf{g}_2\,\mathrm{d}\xi^2 + \mathbf{t}_3\frac{h}{2}\,\mathrm{d}\xi^3, \tag{10.104}$$

where $\mathbf{g}_\alpha \doteq \partial \mathbf{y}_0/\partial \xi^\alpha$, $\xi^3 \doteq 2\zeta/h$, and $\xi^k \in [-1, +1]$. An infinitesimal volume of the parallelepiped spanned by the component vectors is

$$\mathrm{d}V \doteq \left(\mathbf{g}_1\,\mathrm{d}\xi^1 \times \mathbf{g}_2\,\mathrm{d}\xi^2\right) \cdot \left(\mathbf{t}_3\frac{h}{2}\,\mathrm{d}\xi^3\right) = \frac{h}{2}(\mathbf{g}_1 \times \mathbf{g}_2) \cdot \mathbf{t}_3\ \mathrm{d}\xi^1\mathrm{d}\xi^2\mathrm{d}\xi^3. \tag{10.105}$$

Note that

$$\mathbf{g}_1 \times \mathbf{g}_2 = J\,\mathbf{t}_3, \qquad J = (\mathbf{g}_1 \cdot \mathbf{t}_1)(\mathbf{g}_2 \cdot \mathbf{t}_2) - (\mathbf{g}_2 \cdot \mathbf{t}_1)(\mathbf{g}_1 \cdot \mathbf{t}_2), \tag{10.106}$$

for $\mathbf{g}_\alpha$ decomposed in $\{\mathbf{t}_i\}$ as $\mathbf{g}_\alpha = (\mathbf{g}_\alpha \cdot \mathbf{t}_1)\,\mathbf{t}_1 + (\mathbf{g}_\alpha \cdot \mathbf{t}_2)\,\mathbf{t}_2$ and $\mathbf{t}_1 \times \mathbf{t}_2 = \mathbf{t}_3$. Besides, $J \doteq \det \mathbf{J}$, where $\mathbf{J}$ is the Jacobian matrix of eq. (10.41). Using the above relation $(\mathbf{g}_1 \times \mathbf{g}_2) \cdot \mathbf{t}_3 = J\,\mathbf{t}_3 \cdot \mathbf{t}_3 = J$ and

$$\mathrm{d}V = \frac{h}{2}\,J\ \mathrm{d}\xi^1\mathrm{d}\xi^2\mathrm{d}\xi^3. \tag{10.107}$$

We note that $(h/2)\,J = (h/2)\,\det \mathbf{J} = \det \mathbf{J}_L$, for $\mathbf{J}_L$ of eq. (10.39). Besides, $\mathbf{J}_L = \mathbf{R}^T\mathbf{J}_G$, by eq. (10.34), and hence $\det \mathbf{J}_L = \det \mathbf{J}_G$, as $\det \mathbf{R} = 1$. The infinitesimal area of the parallelogram spanned by vectors $(\mathrm{d}\mathbf{y})_\alpha$ $(\alpha = 1, 2)$ is

$$\mathrm{d}A \doteq (\mathbf{g}_1\,\mathrm{d}\xi^1 \times \mathbf{g}_2\,\mathrm{d}\xi^2) \cdot \mathbf{t}_3 = (\mathbf{g}_1 \times \mathbf{g}_2) \cdot \mathbf{t}_3\,\mathrm{d}\xi^1\mathrm{d}\xi^2 = J\ \mathrm{d}\xi^1\mathrm{d}\xi^2. \tag{10.108}$$

Remark. In the above derivations, we assumed that the element's geometry is approximately flat, see the restriction of eq. (5.19). To account for curvature, the infinitesimal parallelepiped should be spanned by the vectors $\hat{\mathbf{g}}_\alpha(\zeta)$ of eq. (5.5) and integrated over the thickness.

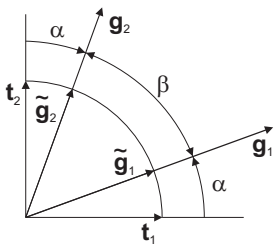

Fig. 10.10 Sign of the Jacobian determinant depends on α or β.

Example. We show that the sign of the Jacobian determinant J is a function of the angle between $\mathbf{t}_1$ and $\mathbf{g}_1$, denoted as α in Fig. 10.10. We assume that the basis $\{\mathbf{t}_i\}$ is constructed as **Basis 2**, see eq. (10.22).

The natural basis vectors $\mathbf{g}_\alpha$ can be expressed as $\mathbf{g}_\alpha = \tilde{\mathbf{g}}_\alpha \|\mathbf{g}_\alpha\|$, where $\tilde{\mathbf{g}}_\alpha$ are unit vectors given in $\{\mathbf{t}_\alpha\}$ as follows:

$$\tilde{\mathbf{g}}_1 = \cos\alpha\ \mathbf{t}_1 - \sin\alpha\ \mathbf{t}_2, \qquad \tilde{\mathbf{g}}_2 = -\sin\alpha\ \mathbf{t}_1 + \cos\alpha\ \mathbf{t}_2.$$

Then the Jacobian determinant becomes

$$\begin{aligned} J &= (\mathbf{g}_1 \cdot \mathbf{t}_1)(\mathbf{g}_2 \cdot \mathbf{t}_2) - (\mathbf{g}_2 \cdot \mathbf{t}_1)(\mathbf{g}_1 \cdot \mathbf{t}_2) \\ &= \|\mathbf{g}_1\|\|\mathbf{g}_2\|(\cos^2\alpha - \sin^2\alpha) = \|\mathbf{g}_1\|\|\mathbf{g}_2\|(1 - 2\sin^2\alpha). \end{aligned} \tag{10.109}$$

Because $\|\mathbf{g}_1\|\|\mathbf{g}_2\| > 0$, the sign of J depends on the angle α, i.e.

$$\begin{aligned} &J > 0, \ \text{for} \ \ |\sin\alpha| < 1/\sqrt{2}, \quad \text{or} \ \ |\alpha| < 45^\text{o}, \\ &J = 0, \ \text{for} \ \ |\sin\alpha| = \pm 1/\sqrt{2}, \ \text{or} \ \ |\alpha| = 45^\text{o}, \\ &J < 0, \ \text{for} \ \ |\sin\alpha| > 1/\sqrt{2}, \quad \text{or} \ \ |\alpha| > 45^\text{o}. \end{aligned} \tag{10.110}$$

We can rewrite these conditions in terms of the angle between $\mathbf{g}_1$ and $\mathbf{g}_2$, defined as $\beta \doteq 90^\text{o} - 2\alpha$, see Fig. 10.10, as follows:

$$\begin{aligned} &J > 0, \ \text{for} \ \ 0^\text{o} < \beta < 180^\text{o}, \\ &J = 0, \ \text{for} \ \ \beta = 0^\text{o} \ \text{or} \ \ \beta = 180^\text{o}, \\ &J < 0, \ \text{for} \ \ \beta > 180^\text{o}\ . \end{aligned} \tag{10.111}$$

We see that J is singular when $\mathbf{g}_1$ and $\mathbf{g}_2$ are co-linear and is negative if they are inclined at the angle greater than 180^o.

Finally, we note that the Jacobian determinant J is computed at the Gauss points and it should be positive to avoid negative volumes which are non-physical.

Remark. Note that the condition requiring that the internal angles between adjacent edges of a four-node element be within the range $[45^{\circ}, 135^{\circ}]$ is motivated by accuracy concerns, as it is more restrictive than necessary to avoid $J < 0$. For instance, if node 3 is placed on the line linking node 2 and 4, then the angle at node 3 is 180°, far beyond the above range. Still, $J > 0$ everywhere except at node 3, where $J = 0$.

Volume and area of shell element. Below we consider the formulas for integration suitable for the natural coordinates ξ^k.

The volume of a shell element is defined as an integral,

$$V \doteq \int_V \mathrm{d}V = \int_{-1}^{+1}\int_{-1}^{+1}\int_{-1}^{+1} \frac{h}{2}\, J \; \mathrm{d}\xi^1 \mathrm{d}\xi^2 \mathrm{d}\xi^3, \tag{10.112}$$

where we used $\mathrm{d}V$ of eq. (10.107). If the thickness h is constant in the element, then we obtain $V = hA$, where the shell element area is defined as

$$A \doteq \int_A \mathrm{d}A = \int_{-1}^{+1}\int_{-1}^{+1} J \; \mathrm{d}\xi^1 \mathrm{d}\xi^2, \tag{10.113}$$

where $\mathrm{d}A$ is defined in eq. (10.108). For a four-node element and the bilinear approximation, $\det \mathbf{J} = J_0 + J_1\,\xi^1 + J_2\,\xi^2$, see eq. (10.63), and the area $A = 4J_0$.

Equation (10.112) is general, can be applied to elements of arbitrary shape, also to the warped ones, while a simpler expression can be found for a flat four-node element. We can divide the four-node elements into two triangles, e.g., by the diagonal 1-3, and calculate its area as follows:

$$A \doteq \tfrac{1}{2}\,(\mathbf{y}_{32} \times \mathbf{y}_{12} - \mathbf{y}_{34} \times \mathbf{y}_{14}) \cdot \mathbf{t}_3, \tag{10.114}$$

where $\mathbf{t}_3$ is a unit normal vector and $\mathbf{y}_{KL} \doteq \mathbf{y}_K - \mathbf{y}_L$, i.e. the vector connecting nodes K and L $(K, L \in \{1, 2, 3, 4\})$. Then $\mathbf{y}_{KL} = S^{\alpha}_{KL}\,\mathbf{t}_{\alpha}$, where $S^{\alpha}_{KL} \doteq S^{\alpha}_K - S^{\alpha}_L$ and $S^{\alpha}_I \doteq (\mathbf{y}_I - \mathbf{y}_c) \cdot \mathbf{t}_{\alpha}$, where $\mathbf{y}_c$ is the position of the element center. Finally, the area can be expressed as

$$A = \tfrac{1}{2}\left(S^1_{31}\, S^2_{42} - S^2_{31}\, S^1_{24}\right) \tag{10.115}$$

or

$$A = \frac{1}{4}\left[(\boldsymbol{\xi}\mathbf{S}^1)(\boldsymbol{\eta}\mathbf{S}^2) - (\boldsymbol{\xi}\mathbf{S}^2)(\boldsymbol{\eta}\mathbf{S}^1)\right]. \tag{10.116}$$

To prove the correctness of the last form, we have to perform multiplications and introduce the differences of coordinates S^{α}_{KL}, which yields eq. (10.115). Comparing eq. (10.116) with the definition of J_0 of eq. (10.63), we obtain $A = 4J_0$.

Integration over the element volume. The volume of the shell element is mapped onto a unit cube and a numerical integration is performed as follows:

$$\int_V F \, \mathrm{d}V = \int_{-1}^{+1}\int_{-1}^{+1}\int_{-1}^{+1} \bar{F}(\xi^k) \, \frac{h}{2} \, J \, \mathrm{d}\xi^1 \mathrm{d}\xi^2 \mathrm{d}\xi^3 = \sum_{n=1}^{N_{IP}} w_n \, \bar{F}(\xi_n^k), \tag{10.117}$$

where $\mathrm{d}V$ is defined in eq. (10.107), $\bar{F}(\xi^k) \doteq \frac{1}{2} f(\xi^k) \, h(\xi^\alpha) \, J(\xi^\alpha)$, and w_n denotes the weighing factor for the integration point n.

We can separately specify the integration rule for the reference lamina (l) and for the fiber (f), see Fig. 5.2, as follows:

$$\int_V F \, \mathrm{d}V = \sum_{l=1}^{N_{IP}^l} w_l \sum_{f=1}^{N_{IP}^f} w_f \, \bar{F}(\xi_l^\alpha, \xi_f^3). \tag{10.118}$$

The order in which the above summations are performed can have a significant effect on the speed of the FE code and the effect can be contrary to our expectations; this issue is discussed in [89]. Note that, generally, it is better to write two integration loops instead of one, as usually it makes a difference to the compiler's optimizer.

Integration of strain energy over thickness. For the Reissner kinematics, the deformation gradient $\mathbf{F}$ is a linear function of the normal coordinate $\zeta \in [-h/2, +h/2]$. Hence, the strain $\mathbf{E}$ and the strain energy density function $\mathcal{W}$ are polynomials of ζ,

$$\mathbf{E}(\zeta) = \sum_{n=0}^{N} \frac{1}{n!} \mathbf{E}^{(n)} \, \zeta^n, \qquad \mathcal{W}(\zeta) = \sum_{n=0}^{N} \mathcal{W}_n \, \zeta^n, \tag{10.119}$$

where $\mathbf{E}^{(n)}$ denotes the nth derivative w.r.t ζ at the middle surface, $\zeta = 0$. The shell-type strain energy is defined as $\mathcal{W}_{\mathrm{sh}} \doteq \int_{-h/2}^{+h/2} \mathcal{W}(\zeta) \, \mathrm{d}\zeta$, and involves the integration through-the-thickness.

In Table 10.1 are (1) the analytically integrated $\mathcal{W}_{\mathrm{sh}}$, and (2) the minimum number of integration points to obtain exact $\mathcal{W}_{\mathrm{sh}}$ for two types of numerical quadratures, Gauss and Newton–Cotes (NC). Various forms of strain are assumed. Note that $\mathcal{W}_2^*$ depends on all derivatives of the strain, while $\mathcal{W}_2$ only on $\mathbf{E}^{(1)}$.

Table 10.1 Integration of SVK strain energy over ζ for various forms of strain. MNIP=minimum no. of integration points

Form of strain $\mathbf{E}(\zeta)$	Coefficients of strain energy $\mathcal{W}(\zeta)$	analytically integr. $\mathcal{W}_{\text{sh}}$	MNIP Gauss	MNIP NC
$\mathbf{E}^{(0)}$	$\mathcal{W}_0(\mathbf{E}^{(0)})$	$h\mathcal{W}_0$	1	1
$\mathbf{E}^{(0)}+\zeta\mathbf{E}^{(1)}$	$\mathcal{W}_0(\mathbf{E}^{(0)}), \mathcal{W}_1(\mathbf{E}^{(0)},\mathbf{E}^{(1)}),$ $\mathcal{W}_2(\mathbf{E}^{(1)})$	$h\mathcal{W}_0+\frac{h^3}{12}\mathcal{W}_2$	2	3
$\mathbf{E}^{(0)}+\zeta\mathbf{E}^{(1)}+\frac{\zeta^2}{2}\mathbf{E}^{(2)}$	$\mathcal{W}_0(\mathbf{E}^{(0)})$, $\mathcal{W}_1(\mathbf{E}^{(0)},\mathbf{E}^{(1)}),$ $\mathcal{W}_2^*(\mathbf{E}^{(0)},\mathbf{E}^{(1)},\mathbf{E}^{(2)}),$ $\mathcal{W}_3(\mathbf{E}^{(1)},\mathbf{E}^{(2)})$, $\mathcal{W}_4(\mathbf{E}^{(2)})$	$h\mathcal{W}_0+\frac{h^3}{12}\mathcal{W}_2^*+\frac{h^5}{80}\mathcal{W}_4$	3	5

Numerical integration of four-node shell elements over thickness. In the four-node shell element, we use a 2×2 Gauss rule for integration over the reference lamina. (The analytical integration over the lamina is used, e.g., in the so-called one-integration point element.) The integration over the fiber is performed either analytically or one of the following 1D integration rules is used:

1. the 2-point Gauss rule. The locations of sampling points for the interval $\xi\in[-1,+1]$ and weighing factors for the 2-point Gauss rule are given in Table 10.2.

Table 10.2 2-point Gauss integration rule.

m	1	2
ξ_m	$-1/\sqrt{3}$	$1/\sqrt{3}$
w_m	1	1

2. the 5-point Simpson rule. In the Simpson method, the interval is divided into an even number of intervals and within each pair of intervals the function is approximated by a parabola. The method is exact for polynomials of degree at most 3. The locations of sampling points for the interval $\xi\in[-1,+1]$ and weighing factors for the 5-point Simpson rule are given in Table 10.3. In the context of shells, the Simpson rule has the advantage that the sampling points are also located at the ends of the interval, $\xi_m=\pm 1$, i.e. at the most external laminas.

Table 10.3 5-point Simpson integration rule.

m	1	2	3	4	5
ξ_m	-1	-1/2	0	1/2	1
w_m	1/6	4/6	3/6	4/6	1/6

10.6 Newton method and tangent operator

Newton method. Consider the potential energy functional defined as

$$F(\mathbf{z}) \doteq \int_V \mathcal{W}(\mathbf{z})\,\mathrm{d}V - \ F_{\mathrm{ext}}, \tag{10.120}$$

where $\mathcal{W}(\mathbf{z})$ is the strain energy expressed by $\mathbf{z} \doteq \{\mathbf{u}, \boldsymbol{\psi}\}$ and F_{ext} is a functional of external forces. The below-described procedure is analogous for other governing functionals.

We write the stationarity condition of F as

$$\delta F \doteq DF(\bar{\mathbf{z}}) \cdot \delta\mathbf{z} = \left.\frac{\mathrm{d}}{\mathrm{d}t} F(\bar{\mathbf{z}} + t\,\delta\mathbf{z})\right|_{t=0} = 0, \tag{10.121}$$

where $DF(\bar{\mathbf{z}}) \cdot \delta\mathbf{z}$ is the directional derivative of F at $\bar{\mathbf{z}}$ in the direction $\delta\mathbf{z}$, t is a scalar parameter, and $\bar{\mathbf{z}}$ denotes the known (last computed) solution. This is the virtual work (VW) equation.

We designate $G \doteq \delta F$ and rewrite the VW equation as $G(\mathbf{z}) = 0$. It can be linearized and solved iteratively, e.g., using the Newton scheme defined as follows:

$$\left.\begin{aligned} DG(\bar{\mathbf{z}}) \cdot \Delta\mathbf{z} &= -G(\bar{\mathbf{z}}) \\ \mathbf{z} &= \bar{\mathbf{z}} + \Delta\mathbf{z} \end{aligned}\right\}, \tag{10.122}$$

where

$$DG(\bar{\mathbf{z}}) \cdot \Delta\mathbf{z} \doteq \left.\frac{\mathrm{d}}{\mathrm{d}t} G(\bar{\mathbf{z}} + t\,\Delta\mathbf{z})\right|_{t=0} \tag{10.123}$$

is the directional derivative of G at $\bar{\mathbf{z}}$ in the direction $\Delta\mathbf{z}$. This derivative provides the tangent operator $\mathbf{K}$.

Tangent operator for the linear material. Consider the strain energy $\mathcal{W}(\mathbf{z}) \doteq \frac{1}{2}\boldsymbol{\varepsilon} \cdot (\mathbb{C}\,\boldsymbol{\varepsilon})$, where $\mathbb{C} \doteq \partial\boldsymbol{\sigma}/\partial\boldsymbol{\varepsilon}$ is the tangent constitutive matrix. For simplicity, we omit the integral and F_{ext} in eq. (10.120).

Then $F(\mathbf{z}) \doteq \mathcal{W}(\mathbf{z})$ and its variation is $G \doteq \delta F = \delta\mathcal{W} = \boldsymbol{\sigma} \cdot \delta\boldsymbol{\varepsilon}$, where the constitutive equation $\boldsymbol{\sigma} \doteq \partial\mathcal{W}/\partial\boldsymbol{\varepsilon} = \mathbb{C}\,\boldsymbol{\varepsilon}$. The directional derivative is calculated as

$$DG(\bar{\mathbf{z}}) \cdot \Delta\mathbf{z} = \frac{\partial G}{\partial \mathbf{z}} \Delta\mathbf{z} = \Delta\boldsymbol{\sigma} \cdot \delta\boldsymbol{\varepsilon} + \boldsymbol{\sigma} \cdot \Delta(\delta\boldsymbol{\varepsilon}). \tag{10.124}$$

The first component of eq. (10.124) can be transformed using the constitutive relation in the incremental form $\Delta\boldsymbol{\sigma} = \mathbb{C}\,\Delta\boldsymbol{\varepsilon}$, where $\mathbf{B} \doteq \partial\boldsymbol{\varepsilon}/\partial\mathbf{z}$ is the kinematical strain-displacement matrix, to obtain

$$\Delta\boldsymbol{\sigma} \cdot \delta\boldsymbol{\varepsilon} = (\mathbb{C}\,\Delta\boldsymbol{\varepsilon}) \cdot \delta\boldsymbol{\varepsilon} = (\mathbb{C}\,\mathbf{B}\,\Delta\mathbf{z}) \cdot (\mathbf{B}\,\delta\mathbf{z}) = \delta\mathbf{z} \cdot (\mathbf{B}^T \mathbb{C}\,\mathbf{B}\,\Delta\mathbf{z}). \tag{10.125}$$

The second component of eq. (10.124) can be rewritten as

$$\boldsymbol{\sigma} \cdot \Delta(\delta\boldsymbol{\varepsilon}) = \Delta(\boldsymbol{\sigma}^* \cdot \delta\boldsymbol{\varepsilon}) = \Delta\delta(\boldsymbol{\sigma}^* \cdot \boldsymbol{\varepsilon}), \tag{10.126}$$

provided that, in differentiation, $\boldsymbol{\sigma}$ is treated as independent of $\mathbf{z}$, which is indicated by the asterisk, i.e. $\boldsymbol{\sigma}^*$. Note that the scalar $(\boldsymbol{\sigma}^* \cdot \boldsymbol{\varepsilon})$ is differentiated in the last form of this equation. Then the second component becomes

$$\Delta\delta(\boldsymbol{\sigma}^* \cdot \boldsymbol{\varepsilon}) = \delta\mathbf{z} \cdot \left[\frac{\partial^2(\boldsymbol{\varepsilon} \cdot \boldsymbol{\sigma}^*)}{\partial \mathbf{z}^2} \Delta\mathbf{z} \right]. \tag{10.127}$$

Finally,

$$DG(\bar{\mathbf{z}}) \cdot \Delta\mathbf{z} = \delta\mathbf{z} \cdot (\,\mathbf{K}\,\Delta\mathbf{z}\,), \qquad \mathbf{K} \doteq \mathbf{B}^T\mathbf{C}\mathbf{B} + \frac{\partial^2(\boldsymbol{\varepsilon} \cdot \boldsymbol{\sigma}^*)}{\partial \mathbf{z}^2}, \tag{10.128}$$

where $\mathbf{K}$ is the tangent stiffness operator.

Computation of tangent matrix. The major part of computation of the tangent matrix is the computation of derivatives and for non-linear functions these derivatives are more complicated than the function itself. There are three ways to compute the tangent matrix:

1. Analytic, i.e. by hand or using one of the symbolic manipulators such as *Mathematica*, *Maple*, and others, which can be used for manipulating equations and obtaining expressions for partial derivatives. This way is exact but laborious.
2. Numerical, i.e. by finite difference (FD) approximations and either two-sided or one-sided differences can be used. This yields an inefficient code, which can be also inaccurate.
3. Automatic (or algorithmic) differentiation of the computer program. The automatic differentiation (AD) programs can deal with constructs such as branches and loops and derivatives are correct up to the machine precision. The AD has strong theoretical foundations and is a mature computational technology, which can be used with confidence, see, e.g., [185, 82, 84, 83].

Below, we discuss these three ways of generating the tangent matrix.

1. Stiffness tangent matrix derived analytically. Again, consider the VW equation in the simplified form, i.e. $G(\mathbf{z}) \doteq \delta\mathcal{W} \doteq \boldsymbol{\sigma} \cdot \delta\boldsymbol{\varepsilon}$, where $\boldsymbol{\sigma} \doteq \partial\mathcal{W}/\partial\boldsymbol{\varepsilon}$. We can split

$$\delta\boldsymbol{\varepsilon} = [\delta_1\boldsymbol{\varepsilon}, ..., \delta_N\boldsymbol{\varepsilon}], \qquad \delta_i\boldsymbol{\varepsilon} = \frac{\partial\boldsymbol{\varepsilon}}{\partial z_i}\,\delta q_i, \qquad i = 1, ..., N, \tag{10.129}$$

where $z_i \in \mathbf{z}_I$ denotes a nodal variable of a discrete FE model and, for simplicity, we take, only one of its components, i.e. $G_i(\mathbf{z}) \doteq \boldsymbol{\sigma} \cdot \delta_i\,\boldsymbol{\varepsilon}$. We shall calculate a derivative of this component w.r.t. one component, $z_j \in \mathbf{z}_I$,

$$\frac{\partial G_i(\mathbf{z}_I)}{\partial z_j} = \frac{\partial}{\partial z_j}\,(\boldsymbol{\sigma} \cdot \delta_i\boldsymbol{\varepsilon}) = \boldsymbol{\sigma}_{,j} \cdot (\mathbf{B}_i\,\delta z_i) + \boldsymbol{\sigma} \cdot (\mathbf{B}_{i,j}\,\delta z_i), \tag{10.130}$$

where

$$\boldsymbol{\sigma}_{,j} = \frac{\partial\boldsymbol{\sigma}}{\partial\boldsymbol{\varepsilon}}\frac{\partial\boldsymbol{\varepsilon}}{\partial z_j} = \mathbb{C}\,\mathbf{B}_j, \quad \mathbf{B}_i = \frac{\partial\boldsymbol{\varepsilon}}{\partial z_i}, \quad \mathbf{B}_{i,j} = \frac{\partial\mathbf{B}_i}{\partial z_j} = \frac{\partial^2\boldsymbol{\varepsilon}}{\partial z_i\,\partial z_j}, \tag{10.131}$$

and $\mathbb{C} \doteq \partial\boldsymbol{\sigma}/\partial\boldsymbol{\varepsilon}$ is the constitutive tangent operator. Note that

1. If components of $\boldsymbol{\sigma}$ and $\boldsymbol{\varepsilon}$ are written as matrices, then $\mathbf{B}_i$ and $\mathbf{B}_{i,j}$ are also matrices but $\mathbb{C}$ must be a 4D matrix. Using the identity $\mathbf{T}_1 \cdot \mathbf{T}_2 = \mathrm{tr}(\mathbf{T}_2^T\mathbf{T}_1)$, we can rewrite eq. (10.130) as

   $$\frac{\partial G_i(\mathbf{z}_I)}{\partial z_j} = \mathrm{tr}\,(\mathbf{B}_i^T\,\mathbb{C}\,\mathbf{B}_j + \mathbf{B}_{i,j}^T\,\boldsymbol{\sigma})\;\delta z_i. \tag{10.132}$$

   This form is not used in computations because of the inconvenient form of $\mathbb{C}$.
2. If components of $\boldsymbol{\sigma}$ and $\boldsymbol{\varepsilon}$ are written as vectors, then $\mathbf{B}_i$ and $\mathbf{B}_{i,j}$ are also vectors, while $\mathbb{C}$ can be written as a 2D matrix; this is the so-called Voigt's notation. Then we can rewrite eq. (10.130) as follows:

   $$\frac{\partial G_i(\mathbf{z}_I)}{\partial z_j} = (\mathbf{B}_i\,\delta z_i) \cdot (\mathbb{C}\,\mathbf{B}_j) + (\mathbf{B}_{i,j}\,\delta z_i) \cdot \boldsymbol{\sigma}. \tag{10.133}$$

   For all components taken into account, we have $i, j = 1, ..., N$ and the vectors $\mathbf{B}_i$ can be arranged as a matrix $\mathbf{B} = [\mathbf{B}_1, ..., \mathbf{B}_N]$. Then

   $$\frac{\partial G(\mathbf{z}_I)}{\partial\mathbf{z}} \doteq \left[\frac{\partial G_i(\mathbf{z})}{\partial z_j}\right]_{i,j=1,...,N} = \delta\mathbf{B}^T\mathbb{C}\,\mathbf{B} \;+\; \left(\frac{\mathrm{d}\delta\mathbf{B}}{\mathrm{d}\mathbf{z}}\right)^T\boldsymbol{\sigma}, \tag{10.134}$$

where $\delta\mathbf{B} = [\mathbf{B}_1\,\delta z_1, ..., \mathbf{B}_N\,\delta z_N]$ and

$$\left(\frac{\mathrm{d}\delta\mathbf{B}}{\mathrm{d}\mathbf{z}}\right)^T = \begin{bmatrix} \mathbf{B}_{1,1}\,\delta z_1 & ... & \mathbf{B}_{1,N}\,\delta z_1 \\ ... & ... & ... \\ \mathbf{B}_{N,1}\,\delta z_N & ... & \mathbf{B}_{N,N}\,\delta z_N \end{bmatrix}. \tag{10.135}$$

Equation (10.134) is used to generate the tangent stiffness matrix which is defined as follows:

$$\mathbf{K} \doteq \int_A \left[\mathbf{B}^T\mathbb{C}\,\mathbf{B} \;+\; \left(\frac{\partial\mathbf{B}}{\partial\mathbf{z}}\right)^T\boldsymbol{\sigma}\right]\,\mathrm{d}V, \tag{10.136}$$

where the displacement matrix and the initial stress matrix are defined as

$$\mathbf{K}_0 + \mathbf{K}_L \doteq \int_A \mathbf{B}^T\mathbb{C}\mathbf{B}\,\mathrm{d}V, \qquad \mathbf{K}_\sigma \doteq \int_A \left(\frac{\partial\mathbf{B}}{\partial\mathbf{z}}\right)^T\boldsymbol{\sigma}\,\mathrm{d}V. \tag{10.137}$$

Here, $\mathbf{K}_0$ is the infinitesimal (linear) stiffness matrix, while $\mathbf{K}_u$ and $\mathbf{K}_\sigma$ are the parts which appear for non-linear strains and/or non-linear constitutive relations, see the classical textbooks on FEs.

2. Stiffness tangent matrix derived by finite difference method. The stiffness matrix can be approximated by a secant operator obtained using the Finite Difference (FD) method. This is a very inefficient method which can be accelerated by deriving $\mathbf{B}$ analytically and using it to compute the initial stress matrix $\mathbf{K}_\sigma$ by the FD method. This is obtained as follows:

1. First, the analytical formula for $\mathbf{B}$ is derived and compared with the FD approximation

$$\mathbf{B}_i^{\mathrm{FD}} = \frac{\boldsymbol{\varepsilon}(z_i+\tau) - \boldsymbol{\varepsilon}(z_i+\tau)}{2\tau}, \qquad z_i \in \mathbf{z}_I, \tag{10.138}$$

where $\tau = 10^{-8}$ for double precision and $\mathbf{B}_i^{\mathrm{FD}}$ is the ith column of $\mathbf{B}^{\mathrm{FD}}$. This verification should be done for $z_i \neq 0$.
2. The so-verified analytical $\mathbf{B}$ is used to compute the initial stress matrix $\mathbf{K}_\sigma$ by the FD method,

$$(\mathbf{K}_\sigma)_i^{\mathrm{FD}} = \frac{\mathbf{B}^T(z_i+\tau)\,\boldsymbol{\sigma} - \mathbf{B}^T(z_i-\tau)\,\boldsymbol{\sigma}}{2\tau}, \tag{10.139}$$

where $(\mathbf{K}_\sigma)_i^{\mathrm{FD}}$ is the ith column of $(\mathbf{K}_\sigma)^{\mathrm{FD}}$. If components of $\boldsymbol{\sigma}$ are written as a vector, then $\mathbf{B}$ is a matrix and, in the nominator, we obtain a difference of two vectors.

Remark. Note that the use of the analytically derived $\mathbf{B}$ is very important for efficiency. Otherwise we have to compute the second derivative of strain because

$$\mathbf{K}_\sigma = \left(\frac{\mathrm{d}\mathbf{B}}{\mathrm{d}\mathbf{z}}\right)^T \boldsymbol{\sigma} = \left(\frac{\mathrm{d}^2\boldsymbol{\varepsilon}}{\mathrm{d}\mathbf{z}^2}\right)^T \boldsymbol{\sigma}. \tag{10.140}$$

Computation of the second derivative is time-consuming for multidimensional $\mathbf{z}$ because, first, a second-order hyper-surface must be spanned. For instance, for each strain component ε_{kl}, we have to span

$$\varepsilon_{kl} = a_0 + \sum_{i=1}^{N} a_i\, z_i + \sum_{i=1}^{N-1} \sum_{j=1+i}^{N} a_{ij}\, z_i\, z_j + \sum_{i=1}^{N} a_{ii}\, z_i^2, \tag{10.141}$$

where the base functions are polynomials of up to the second order. The number of coefficients $a_0,\, a_i, a_{ij}, a_{ii}$ which have to be calculated is

$$p = \underbrace{1}_{\text{constant}} + \underbrace{N}_{\text{linear}} + \underbrace{(N^2 - N)/2}_{\text{mixed quadratic}} + \underbrace{N}_{\text{pure quadratic}} = (N^2 + 3N + 2)/2, \tag{10.142}$$

and for $N = 8$ (four-node element $\times$ 2 dofs/node), we obtain $p = 45$. The semi-analytical method is more efficient but even this method can be used only when the efficiency of the FE is not important.

Finally, we note that the formulas given above may also be used to verify the correctness of $\mathbf{B}$ and $\mathbf{K}_\sigma$ derived either analytically or by an automatic differentiation program.

3. Stiffness tangent matrix by automatic differentiation. Even for a single shell element, we have to use many independent variables, e.g. in case of shell elements with six dofs/node, we use 24 variables in the four-node element and 54 in the 9-node element. This means that the functional F must be differentiated w.r.t. this number of variables, which produces thousands of formulas.

In consequence, the process of derivation of this matrix and coding is time-consuming and error prone. Controlling and modifications of a code become difficult because of its size. For this reason, the programs in which we can write operations in a very compact way and perform automatically differentiation are very useful.

In the automatic differentiation (AD) programs, there are several options which can be used to calculate the residual vector and the stiffness tangent matrix. The simplest possibility is

1. Apply the FE approximations and integrate the functional F, which yields an algebraic function $\hat{F}(\mathbf{z}_I)$ of the set of nodal variables $\mathbf{z}_I$,

$$F \quad \xrightarrow{\text{FE}} \quad \hat{F}(\mathbf{z}_I). \tag{10.143}$$

2. The residual vector and the stiffness tangent matrix are calculated as derivatives of $\hat{F}(\mathbf{z}_I)$ w.r.t. nodal variables

$$\mathbf{r} \doteq \frac{\mathrm{d}\hat{F}(\mathbf{z}_I)}{\mathrm{d}\mathbf{z}_I}, \qquad \mathbf{K} \doteq \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}\mathbf{z}_I}. \tag{10.144}$$

   These operations of differentiation can be coded in a few lines, and the form of results depends on the features of a particular AD program.

We develop finite elements using two programs: *FEAP* [183] and *AceGen* [181] and they are combined in the following way. In *AceGen*, we derive the algebraic function $\hat{F}(\mathbf{z}_I)$ and we code the automatic differentiation operations of eq. (10.144), to obtain the tangent matrix and the residual for an element. The resulting subroutine is in Fortran and is included into *FEAP* to build an executable program which is the subject of tests.

- *AceGen* is a fully reliable system enabling automatic derivation of formulae for numerical procedures developed by J. Korelc[1]. It is written as an add-on package for *Mathematica* and uses the symbolic language of *Mathematica*. The approach implemented in *AceGen* combines several techniques such as: (1) symbolic and algebraic capabilities of *Mathematica*, (2) automatic differentiation (forward and backward mode), (3) automatic code generation, (4) simultaneous optimization of expressions, and many other techniques. For details, see [130, 131] and the manuals.
- *FEAP* is a research finite element environment developed by R.L. Taylor[2], and its source is distributed by the University of California at Berkeley. *FEAP* has an open architecture which allows us to connect user subroutines through a pre-defined interface, see [268]. This program is used in many universities as an excellent environment for developing new finite elements.

[1] Prof. Jože Korelc, University of Ljubljana, Ravnikova 4, SI-1000, Ljubljana, Slovenia. E-mail: AceProducts@fgg.uni-lj.si (http://www.fgg.uni-lj.si/Symech/).

[2] Prof. Robert L. Taylor, Department of Civil Engineering, University of California at Berkeley, Berkeley, CA 94720. E-mail: rlt@ce.berkeley.edu (http://www.ce.berkeley.edu/~rlt/feap/).

Example: forward and backward automatic differentiation. Automatic differentiation can be performed in two ways, designated as *forward* and *backward*. The basic idea is well illustrated by the simple example used, e.g., in the manual of *AceGen* which shows two ways of differentiation of a composite function, $f_3(z_i, f_1(z_i), f_2(z_i, f_1(z_i)))$, depending on the variables z_i, $i = 1, .., n$.

The *forward* differentiation of the function f_3 yields the following formulas:

$$\begin{aligned}
&v_1 = f_1(z_i), && \frac{\partial v_1}{\partial z_i} = \frac{\partial f_1}{\partial z_i}, && i = 1, ..., n,\\
&v_2 = f_2(z_i, v_1), && \frac{\partial v_2}{\partial z_i} = \frac{\partial f_2}{\partial z_i} + \frac{\partial f_2}{\partial v_1}\frac{\partial v_1}{\partial z_i}, && i = 1, ..., n,\\
&v_3 = f_3(z_i, v_1, v_2), && \frac{\partial v_3}{\partial z_i} = \frac{\partial f_3}{\partial z_i} + \frac{\partial f_3}{\partial v_1}\frac{\partial v_1}{\partial z_i} + \frac{\partial f_3}{\partial v_2}\frac{\partial v_2}{\partial z_i}, && i = 1, ..., n,
\end{aligned} \tag{10.145}$$

where v_1, v_2, v_3 are the intermediate variables generated during differentiation.

The *backward* differentiation looks like

$$\begin{aligned}
&v_3 = f_3(z_i, v_1, v_2), && \overline{v_3} = \frac{\partial v_3}{\partial v_3} = 1,\\
&v_2 = f_2(z_i, v_1), && \overline{v_2} = \frac{\partial v_3}{\partial v_2} = \frac{\partial f_3}{\partial v_2}\overline{v_3},\\
&v_1 = f_1(z_i), && \overline{v_1} = \frac{\partial v_3}{\partial v_1} = \frac{\partial f_3}{\partial v_1}\overline{v_3} + \frac{\partial f_2}{\partial v_1}\overline{v_2},\\
&z_i && \frac{\partial v_3}{\partial z_i} = \frac{\partial f_3}{\partial z_i}\overline{v_3} + \frac{\partial f_2}{\partial z_i}\overline{v_2} + \frac{\partial f_1}{\partial z_i}\overline{v_1}, \quad i = 1, ..., n.
\end{aligned} \tag{10.146}$$

The *backward* differentiation yields more effective algorithms for large n, although is more time consuming.