A direct approach to problems in elasticity

2.1 Introduction

The process of approximating the behaviour of a continuum by 'finite elements' which behave in a manner similar to the real, 'discrete', elements described in the previous chapter can be introduced through the medium of particular physical applications or as a general mathematical concept. We have chosen here to follow the first path, narrowing our view to a set of problems associated with structural mechanics which historically were the first to which the finite element method was applied. In Chapter 3 we shall generalize the concepts and show that the basic ideas are widely applicable.

In many phases of engineering the solution of stress and strain distributions in elastic continua is required. Special cases of such problems may range from two-dimensional plane stress or strain distributions, axisymmetric solids, plate bending, and shells, to fully three-dimensional solids. In all cases the number of interconnections between any 'finite element' isolated by some imaginary boundaries and the neighbouring elements is infinite. It is therefore difficult to see at first glance how such problems may be discretized in the same manner as was described in the preceding chapter for simpler structures. The difficulty can be overcome (and the approximation made) in the following manner:

1. The continuum is separated by imaginary lines or surfaces into a number of 'finite elements'.
2. The elements are assumed to be interconnected at a discrete number of nodal points situated on their boundaries and occasionally in their interior. In Chapter 6 we shall show that this limitation is not necessary. The displacements of these nodal points will be the basic unknown parameters of the problem, just as in simple, discrete, structural analysis.
3. A set of functions is chosen to define uniquely the state of displacement within each 'finite element' and on its boundaries in terms of its nodal displacements.
4. The displacement functions now define uniquely the state of strain within an element in terms of the nodal displacements. These strains, together with any initial strains and the constitutive properties of the material, will define the state of stress throughout the element and, hence, also on its boundaries.
5. A system of ‘forces’ concentrated at the nodes and equilibrating the boundary stresses and any distributed loads is determined, resulting in a stiffness relationship of the form of Eq. (1.3).

Once this stage has been reached the solution procedure can follow the standard discrete system pattern described earlier.

Clearly a series of approximations has been introduced. Firstly, it is not always easy to ensure that the chosen displacement functions will satisfy the requirement of displacement continuity between adjacent elements. Thus, the compatibility condition on such lines may be violated (though within each element it is obviously satisfied due to the uniqueness of displacements implied in their continuous representation). Secondly, by concentrating the equivalent forces at the nodes, equilibrium conditions are satisfied in the overall sense only. Local violation of equilibrium conditions within each element and on its boundaries will usually arise.

The choice of element shape and of the form of the displacement function for specific cases leaves many opportunities for the ingenuity and skill of the engineer to be employed, and obviously the degree of approximation which can be achieved will strongly depend on these factors.

The approach outlined here is known as the displacement formulation.\textsuperscript{1,2}

So far, the process described is justified only intuitively, but what in fact has been suggested is equivalent to the minimization of the total potential energy of the system in terms of a prescribed displacement field. If this displacement field is defined in a suitable way, then convergence to the correct result must occur. The process is then equivalent to the well-known Rayleigh–Ritz procedure. This equivalence will be proved in a later section of this chapter where also a discussion of the necessary convergence criteria will be presented.

The recognition of the equivalence of the finite element method to a minimization process was late.\textsuperscript{2,3} However, Courant in 1943\textsuperscript{4†} and Prager and Synge\textsuperscript{5} in 1947 proposed methods that are in essence identical.

This broader basis of the finite element method allows it to be extended to other continuum problems where a variational formulation is possible. Indeed, general procedures are now available for a finite element discretization of any problem defined by a properly constituted set of differential equations. Such generalizations will be discussed in Chapter 3, and throughout the book application to non-structural problems will be made. It will be found that the processes described in this chapter are essentially an application of trial-function and Galerkin-type approximations to a particular case of solid mechanics.

\section*{2.2 Direct formulation of finite element characteristics}

The ‘prescriptions’ for deriving the characteristics of a ‘finite element’ of a continuum, which were outlined in general terms, will now be presented in more detailed mathematical form.

\footnote{\textsuperscript{†} It appears that Courant had anticipated the essence of the finite element method in general, and of a triangular element in particular, as early as 1923 in a paper entitled ‘On a convergence principle in the calculus of variations.’ König Gesellschaft der Wissenschaften zu Göttingen, Nachrichten, Berlin, 1923. He states: ‘We imagine a mesh of triangles covering the domain \ldots the convergence principles remain valid for each triangular domain.’}
It is desirable to obtain results in a general form applicable to any situation, but to avoid introducing conceptual difficulties the general relations will be illustrated with a very simple example of plane stress analysis of a thin slice. In this a division of the region into triangular-shaped elements is used as shown in Fig. 2.1. Relationships of general validity will be placed in a box. Again, matrix notation will be implied.

**2.2.1 Displacement function**

A typical finite element, $e$, is defined by nodes, $i, j, m$, etc., and straight line boundaries. Let the displacements $u$ at any point within the element be approximated as a column vector, $\hat{u}$:

$$\mathbf{u} \approx \hat{u} = \sum_{k} N_k a^e_k = [N_i, N_j, \ldots] \left\{ \begin{array}{c} a_i \\ a_j \\ \vdots \end{array} \right\}^e = \mathbf{N} \mathbf{a}^e$$  \hspace{1cm} (2.1)

in which the components of $\mathbf{N}$ are prescribed functions of position and $\mathbf{a}^e$ represents a listing of nodal displacements for a particular element.
In the case of plane stress, for instance,
\[ \mathbf{u} = \begin{bmatrix} u(x, y) \\ v(x, y) \end{bmatrix} \]
represents horizontal and vertical movements of a typical point within the element and
\[ \mathbf{a}_i = \begin{bmatrix} u_i \\ v_i \end{bmatrix} \]
the corresponding displacements of a node \( i \).

The functions \( N_i, N_j, N_m \) have to be chosen so as to give appropriate nodal displacements when the coordinates of the corresponding nodes are inserted in Eq. (2.1). Clearly, in general,
\[ N_i(x_i, y_i) = I \] (identity matrix)
while
\[ N_i(x_j, y_j) = N_i(x_m, y_m) = 0, \quad \text{etc.} \]
which is simply satisfied by suitable linear functions of \( x \) and \( y \).

If both the components of displacement are specified in an identical manner then we can write
\[ N_i = N_j I \]
and obtain \( N_i \) from Eq. (2.1) by noting that \( N_i = 1 \) at \( x_i, y_i \) but zero at other vertices.

The most obvious linear function in the case of a triangle will yield the shape of \( N_i \) of the form shown in Fig. 2.2. Detailed expressions for such a linear interpolation are given in Chapter 4, but at this stage can be readily derived by the reader.

The functions \( N \) will be called shape functions and will be seen later to play a paramount role in finite element analysis.

### 2.2.2 Strains

With displacements known at all points within the element the 'strains' at any point can be determined. These will always result in a relationship that can be written in
A direct approach to problems in elasticity

Matrix notation as

\[ \varepsilon \approx \hat{\varepsilon} = S\mathbf{u} \]  

(2.2)

where \( S \) is a suitable linear operator. Using Eq. (2.1), the above equation can be approximated as

\[ \varepsilon \approx \hat{\varepsilon} = B\mathbf{a} \]  

(2.3)

with

\[ B = SN \]  

(2.4)

For the plane stress case the relevant strains of interest are those occurring in the plane and are defined in terms of the displacements by well-known relations which define the operator \( S \):

\[
\begin{align*}
\varepsilon_x &= \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}, \\
\varepsilon_y &= \frac{\partial v}{\partial y} + \frac{\partial u}{\partial x}, \\
\gamma_{xy} &= \frac{\partial u}{\partial y} + \frac{\partial v}{\partial x}.
\end{align*}
\]

With the shape functions \( N_i, N_j, \) and \( N_m \) already determined, the matrix \( B \) can easily be obtained. If the linear form of these functions is adopted then, in fact, the strains will be constant throughout the element.

2.2.3 Stresses

In general, the material within the element boundaries may be subjected to initial strains such as may be due to temperature changes, shrinkage, crystal growth, and so on. If such strains are denoted by \( \varepsilon_0 \) then the stresses will be caused by the difference between the actual and initial strains.

In addition it is convenient to assume that at the outset of the analysis the body is stressed by some known system of initial residual stresses \( \sigma_0 \) which, for instance, could be measured, but the prediction of which is impossible without the full knowledge of the material's history. These stresses can simply be added on to the general definition. Thus, assuming general linear elastic behaviour, the relationship between stresses and strains will be linear and of the form

\[ \sigma = D(\varepsilon - \varepsilon_0) + \sigma_0 \]  

(2.5)

where \( D \) is an elasticity matrix containing the appropriate material properties.

† It is known that strain is a second-rank tensor by its transformation properties; however, in this book we will normally represent quantities using matrix (Voigt) notation. The interested reader is encouraged to consult Appendix B for the relations between tensor forms and matrix quantities.
Again, for the particular case of plane stress three components of stress corresponding to the strains already defined have to be considered. These are, in familiar notation

\[
\sigma = \begin{bmatrix}
\sigma_x \\
\sigma_y \\
\tau_{xy}
\end{bmatrix}
\]

and the D matrix may be simply obtained from the usual isotropic stress–strain relationship

\[
\begin{align*}
\varepsilon_x - (\varepsilon_x)_0 &= \frac{1}{E} \sigma_x - \frac{\nu}{E} \sigma_y \\
\varepsilon_y - (\varepsilon_y)_0 &= -\frac{\nu}{E} \sigma_x + \frac{1}{E} \sigma_y \\
\gamma_{xy} - (\gamma_{xy})_0 &= \frac{2(1+\nu)}{E} \tau_{xy}
\end{align*}
\]

i.e., on solving,

\[
D = \frac{E}{1-\nu^2} \begin{bmatrix}
1 & \nu & 0 \\
\nu & 1 & 0 \\
0 & 0 & (1-\nu)/2
\end{bmatrix}
\]

### 2.2.4 Equivalent nodal forces

Let

\[
q^e = \left\{ q_i^e \right\}
\]

define the nodal forces which are statically equivalent to the boundary stresses and distributed body forces on the element. Each of the forces \( q_i^e \) must contain the same number of components as the corresponding nodal displacement \( a_i \) and be ordered in the appropriate, corresponding directions.

The distributed body forces \( b \) are defined as those acting on a unit volume of material within the element with directions corresponding to those of the displacements \( u \) at that point.

In the particular case of plane stress the nodal forces are, for instance,

\[
q_i^e = \left\{ \begin{array}{c} U_i \\ V_i \end{array} \right\}^e
\]

with components \( U \) and \( V \) corresponding to the directions of \( u \) and \( v \) displacements, and the distributed body forces are

\[
b = \left\{ \begin{array}{c} b_x \\ b_y \end{array} \right\}
\]

in which \( b_x \) and \( b_y \) are the 'body force' components.
A direct approach to problems in elasticity

To make the nodal forces statically equivalent to the actual boundary stresses and distributed body forces, the simplest procedure is to impose an arbitrary (virtual) nodal displacement and to equate the external and internal work done by the various forces and stresses during that displacement.

Let such a virtual displacement be \( \delta \alpha^e \) at the nodes. This results, by Eqs (2.1) and (2.2), in displacements and strains within the element equal to

\[
\delta u = N \delta \alpha^e \quad \text{and} \quad \delta \varepsilon = B \delta \alpha^e
\]

respectively.

The work done by the nodal forces is equal to the sum of the products of the individual force components and corresponding displacements, i.e., in matrix language

\[
\delta \alpha^T q^e
\]

Similarly, the internal work per unit volume done by the stresses and distributed body forces is

\[
\delta \varepsilon^T \sigma - \delta u^T b
\]
or

\[
\delta \alpha^T (B^T \sigma - N^T b)
\]

Equating the external work with the total internal work obtained by integrating over the volume of the element, \( V^e \), we have

\[
\delta \alpha^T q^e = \delta \alpha^T \left( \int_{V^e} B^T \sigma \, d(vol) - \int_{V^e} N^T b \, d(vol) \right)
\]

As this relation is valid for any value of the virtual displacement, the multipliers must be equal. Thus

\[
q^e = \int_{V^e} B^T \sigma \, d(vol) - \int_{V^e} N^T b \, d(vol)
\]

This statement is valid quite generally for any stress–strain relation. With the linear law of Eq. (2.5) we can write Eq. (2.11) as

\[
q^e = K^e a^e + f^e
\]

where

\[
K^e = \int_{V^e} B^T D B \, d(vol)
\]

and

\[
f^e = -\int_{V^e} N^T b \, d(vol) - \int_{V^e} B^T D \varepsilon_0 \, d(vol) + \int_{V^e} B^T \sigma_0 \, d(vol)
\]

† Note that by the rules of matrix algebra for the transpose of products

\[
(AB)^T = B^T A^T
\]
In the last equation the three terms represent forces due to body forces, initial strain, and initial stress respectively. The relations have the characteristics of the discrete structural elements described in Chapter 1.

If the initial stress system is self-equilibrating, as must be the case with normal residual stresses, then the forces given by the initial stress term of Eq. (2.13b) are identically zero after assembly. Thus frequent evaluation of this force component is omitted. However, if for instance a machine part is manufactured out of a block in which residual stresses are present or if an excavation is made in rock where known tectonic stresses exist a removal of material will cause a force imbalance which results from the above term.

For the particular example of the plane stress triangular element these characteristics will be obtained by appropriate substitution. It has already been noted that the \( \mathbf{B} \) matrix in that example was not dependent on the coordinates; hence the integration will become particularly simple.

The interconnection and solution of the whole assembly of elements follows the simple structural procedures outlined in Chapter 1. In general, external concentrated forces may exist at the nodes and the matrix

\[
\mathbf{r} = \begin{bmatrix}
    r_1 \\
    r_2 \\
    \vdots \\
    r_n
\end{bmatrix}
\]  

(2.14)

will be added to the consideration of equilibrium at the nodes.

A note should be added here concerning elements near the boundary. If, at the boundary, displacements are specified, no special problem arises as these can be satisfied by specifying some of the nodal parameters \( \mathbf{a} \). Consider, however, the boundary as subject to a distributed external loading, say \( \mathbf{f} \) per unit area. A loading term on the nodes of the element which has a boundary face \( A' \) will now have to be added. By the virtual work consideration, this will simply result in

\[
\mathbf{f}' = -\int_{A'} \mathbf{N}^T \mathbf{t} \, d(\text{area})
\]  

(2.15)

with the integration taken over the boundary area of the element. It will be noted that \( \mathbf{f}' \) must have the same number of components as \( \mathbf{u} \) for the above expression to be valid.

Such a boundary element is shown again for the special case of plane stress in Fig. 2.1. An integration of this type is sometimes not carried out explicitly. Often by 'physical intuition' the analyst will consider the boundary loading to be represented simply by concentrated loads acting on the boundary nodes and calculate these by direct static procedures. In the particular case discussed the results will be identical.

Once the nodal displacements have been determined by solution of the overall 'structural' type equations, the stresses at any point of the element can be found from the relations in Eqs (2.3) and (2.5), giving

\[
\mathbf{\sigma} = \mathbf{D} \mathbf{a} + \mathbf{D} \mathbf{e}_0 + \mathbf{c}_0
\]  

(2.16)
A direct approach to problems in elasticity

in which the typical terms of the relationship of Eq. (1.4) will be immediately recognized, the element stress matrix being

$$Q' = DB$$  \hspace{1cm} (2.17)

To this the stresses

$$\sigma_e^0 = -D\varepsilon_0 \quad \text{and} \quad \sigma_0$$  \hspace{1cm} (2.18)

have to be added.

2.2.5 Generalized nature of displacements, strains, and stresses

The meaning of displacements, strains, and stresses in the illustrative case of plane stress was obvious. In many other applications, shown later in this book, this terminology may be applied to other, less obvious, quantities. For example, in considering plate elements the 'displacement' may be characterized by the lateral deflection and the slopes of the plate at a particular point. The 'strains' will then be defined as the curvatures of the middle surface and the 'stresses' as the corresponding internal bending moments (see Volume 2).

All the expressions derived here are generally valid provided the sum product of displacement and corresponding load components truly represents the external work done, while that of the 'strain' and corresponding 'stress' components results in the total internal work.

2.3 Generalization to the whole region — internal nodal force concept abandoned

In the preceding section the virtual work principle was applied to a single element and the concept of equivalent nodal force was retained. The assembly principle thus followed the conventional, direct equilibrium, approach.

The idea of nodal forces contributed by elements replacing the continuous interaction of stresses between elements presents a conceptual difficulty. However, it has a considerable appeal to 'practical' engineers and does at times allow an interpretation which otherwise would not be obvious to the more rigorous mathematician. There is, however, no need to consider each element individually and the reasoning of the previous section may be applied directly to the whole continuum.

Equation (2.1) can be interpreted as applying to the whole structure, that is,

$$\mathbf{u} = \tilde{N}\mathbf{a}$$  \hspace{1cm} (2.19)

in which \(\mathbf{a}\) lists all the nodal points and

$$\tilde{N}_i = N_i^e$$  \hspace{1cm} (2.20)

when the point concerned is within a particular element \(e\) and \(i\) is a point associated with that element. If point \(i\) does not occur within the element (see Fig. 2.3)

$$\tilde{N}_i = 0$$  \hspace{1cm} (2.21)
Matrix $\bar{B}$ can be similarly defined and we shall drop the bar superscript, considering simply that the shape functions, etc., are always defined over the whole region $V$.

For any virtual displacement $\delta a$ we can now write the sum of internal and external work for the whole region as

$$-\delta a^T r = \int_V \delta u^T b \, dV + \int_A \delta u^T \dot{t} \, dA - \int_V \delta \epsilon^T \sigma \, dV$$

(2.22)

In the above equation $\delta a$, $\delta u$, and $\delta \epsilon$ can be completely arbitrary, providing they stem from a continuous displacement assumption. If for convenience we assume they are simply variations linked by the relations (2.19) and (2.3) we obtain, on substitution of the constitutive relation (2.5), a system of algebraic equations

$$K a + f = r$$

(2.23)

where

$$K = \int_V B^T D B \, dV$$

(2.24a)

and

$$f = - \int_V N^T b \, dV - \int_A N^T \dot{t} \, dA - \int_V B^T D \epsilon_0 \, dV + \int_V B^T \sigma_0 \, dV$$

(2.24b)

The integrals are taken over the whole volume $V$ and over the whole surface area $A$ on which the tractions are given.

It is immediately obvious from the above that

$$K_{ij} = \sum \kappa_{ij} \quad f_i = \sum f'_i$$

(2.25)

by virtue of the property of definite integrals requiring that the total be the sum of the parts:

$$\int_V (\ ) \, dV = \sum \int_{V'} (\ ) \, dV$$

(2.26)
A direct approach to problems in elasticity

The same is obviously true for the surface integrals in Eq. (2.25). We thus see that the 'secret' of the approximation possessing the required behaviour of a 'standard discrete system of Chapter 1' lies simply in the requirement of writing the relationships in integral form.

The assembly rule as well as the whole derivation has been achieved without involving the concept of 'interelement forces' (i.e., $q^i$). In the remainder of this book the element superscript will be dropped unless specifically needed. Also no differentiation between element and system shape functions will be made.

However, an important point arises immediately. In considering the virtual work for the whole system [Eq. (2.22)] and equating this to the sum of the element contributions it is implicitly assumed that no discontinuity in displacement between adjacent elements develops. If such a discontinuity developed, a contribution equal to the work done by the stresses in the separations would have to be added.

Fig. 2.4 Differentiation of a function with slope discontinuity ($C_0$ continuous).
Put in other words, we require that the terms integrated in Eq. (2.26) be finite. These terms arise from the shape functions $N_i$ used in defining the displacement $u$ [by Eq. (2.19)] and its derivatives associated with the definition of strain [viz. Eq. (2.3)]. If, for instance, the ‘strains’ are defined by first derivatives of the functions $N$, the displacements must be continuous. In Fig. 2.4 we see how first derivatives of continuous functions may involve a ‘jump’ but are still finite, while second derivatives may become infinite. Such functions we call $C_0$ continuous.

In some problems the ‘strain’ in a generalized sense may be defined by second derivatives. In such cases we shall obviously require that both the function $N$ and its slope (first derivative) be continuous. Such functions are more difficult to derive but we shall make use of them in plate and shell problems (see Volume 2). The continuity involved now is called $C_1$ continuity.

### 2.4 Displacement approach as a minimization of total potential energy

The principle of virtual displacements used in the previous sections ensured satisfaction of equilibrium conditions within the limits prescribed by the assumed displacement pattern. Only if the virtual work equality for all, arbitrary, variations of displacement was ensured would the equilibrium be complete.

As the number of parameters of $a$ which prescribes the displacement increases without limit then ever closer approximation of all equilibrium conditions can be ensured.

The virtual work principle as written in Eq. (2.22) can be restated in a different form if the virtual quantities $\delta a$, $\delta u$, and $\delta e$ are considered as variations of the real quantities.

Thus, for instance, we can write

$$\delta \left( a^T r + \int_V u^T b \, dV + \int_A u^T t \, dA \right) = -\delta W$$

(2.27)

for the first three terms of Eq. (2.22), where $W$ is the potential energy of the external loads. The above is certainly true if $r$, $b$, and $t$ are conservative (or independent of displacement).

The last term of Eq. (2.22) can, for elastic materials, be written as

$$\delta U = \int_V \delta \varepsilon^T \sigma \, dV$$

(2.28)

where $U$ is the ‘strain energy’ of the system. For the elastic, linear material described by Eq. (2.5) the reader can verify that

$$U = \frac{1}{2} \int_V \varepsilon^T D \varepsilon \, dV - \int_V \varepsilon^T D \varepsilon_0 \, dV + \int_V \varepsilon^T \sigma_0 \, dV$$

(2.29)

will, after differentiation, yield the correct expression providing $D$ is a symmetric matrix. (This is indeed a necessary condition for a single-valued $U$ to exist.)

Thus instead of Eq. (2.22) we can write simply

$$\delta (U + W) = \delta (\Pi) = 0$$

(2.30)

in which the quantity $\Pi$ is called the total potential energy.
A direct approach to problems in elasticity

The above statement means that for equilibrium to be ensured the total potential energy must be stationary for variations of admissible displacements. The finite element equations derived in the previous section [Eqs (2.23)–(2.25)] are simply the statements of this variation with respect to displacements constrained to a finite number of parameters \( a \) and could be written as

\[
\frac{\partial \Pi}{\partial a} = \begin{bmatrix}
\frac{\partial \Pi}{\partial a_1} \\
\frac{\partial \Pi}{\partial a_2} \\
\vdots
\end{bmatrix} = 0
\]  

(2.31)

It can be shown that in stable elastic situations the total potential energy is not only stationary but is a minimum. Thus the finite element process seeks such a minimum within the constraint of an assumed displacement pattern.

The greater the degrees of freedom, the more closely will the solution approximate the true one, ensuring complete equilibrium, providing the true displacement can, in the limit, be represented. The necessary convergence conditions for the finite element process could thus be derived. Discussion of these will, however, be deferred to subsequent sections.

It is of interest to note that if true equilibrium requires an absolute minimum of the total potential energy, \( \Pi \), a finite element solution by the displacement approach will always provide an approximate \( \Pi \) greater than the correct one. Thus a bound on the value of the total potential energy is always achieved.

If the functional \( \Pi \) could be specified, a priori, then the finite element equations could be derived directly by the differentiation specified by Eq. (2.31).

The well-known Rayleigh–Ritz\(^8–9\) process of approximation frequently used in elastic analysis uses precisely this approach. The total potential energy expression is formulated and the displacement pattern is assumed to vary with a finite set of undetermined parameters. A set of simultaneous equations minimizing the total potential energy with respect to these parameters is set up. Thus the finite element process as described so far can be considered to be the Rayleigh–Ritz procedure. The difference is only in the manner in which the assumed displacements are prescribed. In the Ritz process traditionally used these are usually given by expressions valid throughout the whole region, thus leading to simultaneous equations in which no banding occurs and the coefficient matrix is full. In the finite element process this specification is usually piecewise, each nodal parameter influencing only adjacent elements, and thus a sparse and usually banded matrix of coefficients is found.

By its nature the conventional Ritz process is limited to relatively simple geometrical shapes of the total region while this limitation only occurs in finite element analysis in the element itself. Thus complex, realistic, configurations can be assembled from relatively simple element shapes.

A further difference in kind is in the usual association of the undetermined parameter with a particular nodal displacement. This allows a simple physical interpretation invaluable to an engineer. Doubtless much of the popularity of the finite element process is due to this fact.
2.5 Convergence criteria

The assumed shape functions limit the infinite degrees of freedom of the system, and the true minimum of the energy may never be reached, irrespective of the fineness of subdivision. To ensure convergence to the correct result certain simple requirements must be satisfied. Obviously, for instance, the displacement function should be able to represent the true displacement distribution as closely as desired. It will be found that this is not so if the chosen functions are such that straining is possible when the element is subjected to rigid body displacements. Thus, the first criterion that the displacement function must obey is as follows:

Criterion 1. The displacement function chosen should be such that it does not permit straining of an element to occur when the nodal displacements are caused by a rigid body motion.

This self-evident condition can be violated easily if certain types of function are used; care must therefore be taken in the choice of displacement functions.

A second criterion stems from similar requirements. Clearly, as elements get smaller nearly constant strain conditions will prevail in them. If, in fact, constant strain conditions exist, it is most desirable for good accuracy that a finite size element is able to reproduce these exactly. It is possible to formulate functions that satisfy the first criterion but at the same time require a strain variation throughout the element when the nodal displacements are compatible with a constant strain solution. Such functions will, in general, not show good convergence to an accurate solution and cannot, even in the limit, represent the true strain distribution. The second criterion can therefore be formulated as follows:

Criterion 2. The displacement function has to be of such a form that if nodal displacements are compatible with a constant strain condition such constant strain will in fact be obtained. (In this context again a generalized 'strain' definition is implied.)

It will be observed that Criterion 2 in fact incorporates the requirement of Criterion 1, as rigid body displacements are a particular case of constant strain – with a value of zero. This criterion was first stated by Bazeley et al. in 1965. Strictly, both criteria need only be satisfied in the limit as the size of the element tends to zero. However, the imposition of these criteria on elements of finite size leads to improved accuracy, although in certain situations (such as illustrated by the axisymmetric analysis of Chapter 5) the imposition of the second one is not possible or essential.

Lastly, as already mentioned in Sec. 2.3, it is implicitly assumed in this derivation that no contribution to the virtual work arises at element interfaces. It therefore appears necessary that the following criterion be included:

Criterion 3. The displacement functions should be chosen such that the strains at the interface between elements are finite (even though they may be discontinuous).

This criterion implies a certain continuity of displacements between elements. In the case of strains being defined by first derivatives, as in the plane stress example quoted here, the displacements only have to be continuous. If, however, as in the
plate and shell problems, the 'strains' are defined by second derivatives of deflections, first derivatives of these have also to be continuous.2

The above criteria are included mathematically in a statement of 'functional completeness' and the reader is referred elsewhere for full mathematical discussion.11–16 The 'heuristic' proof of the convergence requirements given here is sufficient for practical purposes in all but the most pathological cases and we shall generalize all of the above criteria in Section 3.6 and more fully in Chapter 10, where we shall present a universal test which justifies convergence even if some of the above criteria are violated.

2.6 Discretization error and convergence rate

In the foregoing sections we have assumed that the approximation to the displacement as represented by Eq. (2.1) will yield the exact solution in the limit as the size $h$ of elements decreases. The arguments for this are simple: if the expansion is capable, in the limit, of exactly reproducing any displacement form conceivable in the continuum, then as the solution of each approximation is unique it must approach, in the limit of $h \to 0$, the unique exact solution. In some cases the exact solution is indeed obtained with a finite number of subdivisions (or even with one element only) if the polynomial expansion is used in that element and if this can fit exactly the correct solution. Thus, for instance, if the exact solution is of the form of a quadratic polynomial and the shape functions include all the polynomials of that order, the approximation will yield the exact answer.

The last argument helps in determining the order of convergence of the finite element procedure as the exact solution can always be expanded in the vicinity of any point (or node) $i$ as a polynomial:

$$ u = u_i + \left( \frac{\partial u}{\partial x} \right) (x - x_i) + \left( \frac{\partial u}{\partial y} \right) (y - y_i) + \cdots $$ (2.32)

If within an element of 'size' $h$ a polynomial expansion of degree $p$ is employed, this can fit locally the Taylor expansion up to that degree and, as $x - x_i$ and $y - y_i$ are of the order of magnitude $h$, the error in $u$ will be of the order $O(h^p+1)$. Thus, for instance, in the case of the plane elasticity problem discussed, we used a linear expansion and $p = 1$. We should therefore expect a convergence rate of order $O(h^2)$, i.e., the error in displacement being reduced to $\frac{1}{4}$ for a halving of the mesh spacing.

By a similar argument the strains (or stresses) which are given by the $m$th derivatives of displacement should converge with an error of $O(h^{p+1-m})$, i.e., as $O(h)$ in the example quoted, where $m = 1$. The strain energy, being given by the square of the stresses, will show an error of $O(h^{2(p+1-m)})$ or $O(h^2)$ in the plane stress example.

The arguments given here are perhaps a trifle 'heuristic' from a mathematical viewpoint – they are, however, true and correctly give the orders of convergence, which can be expected to be achieved asymptotically as the element size tends to zero and if the exact solution does not contain singularities. Such singularities may result in infinite values of the coefficients in terms omitted in the Taylor expansion of Eq. (2.32) and invalidate the arguments. However, in many well-behaved problems the mere determination of the order of convergence often suffices to extrapolate the
solution to the correct result. Thus, for instance, if the displacement converges at $O(h^2)$ and we have two approximate solutions $u^1$ and $u^2$ obtained with meshes of size $h$ and $h/2$, we can write, with $u$ being the exact solution,

\[
\frac{u^1 - u}{u^2 - u} = \frac{O(h^2)}{O(h/2)^2} = 4
\]

(2.33)

From the above an (almost) exact solution $u$ can be predicted. This type of extrapolation was first introduced by Richardson and is of use if convergence is monotonic and nearly asymptotic.

We shall return to the important question of estimating errors due to the discretization process in Chapter 14 and will show that much more precise methods than those arising from convergence rate considerations are possible today. Indeed automatic mesh refinement processes are being introduced so that the specified accuracy can be achieved (viz. Chapter 15).

Discretization error is not the only error possible in a finite element computation. In addition to obvious mistakes which can occur when using computers, errors due to round-off are always possible. With the computer operating on numbers rounded off to a finite number of digits, a reduction of accuracy occurs every time differences between 'like' numbers are being formed. In the process of equation solving many subtractions are necessary and accuracy decreases. Problems of matrix conditioning, etc., enter here and the user of the finite element method must at all times be aware of accuracy limitations which simply do not allow the exact solution ever to be obtained. Fortunately in many computations, by using modern machines which carry a large number of significant digits, these errors are often small.

The question of errors arising from the algebraic processes will be stressed in Chapter 20 dealing with computation procedures.

### 2.7 Displacement functions with discontinuity between elements – non-conforming elements and the patch test

In some cases considerable difficulty is experienced in finding displacement functions for an element which will automatically be continuous along the whole interface between adjacent elements.

As already pointed out, the discontinuity of displacement will cause infinite strains at the interfaces, a factor ignored in this formulation because the energy contribution is limited to the elements themselves.

However, if, in the limit, as the size of the subdivision decreases continuity is restored, then the formulation already obtained will still tend to the correct answer. This condition is always reached if

(a) a constant strain condition automatically ensures displacement continuity and

(b) the constant strain criterion of the previous section is satisfied.

To test that such continuity is achieved for any mesh configuration when using such non-conforming elements it is necessary to impose, on an arbitrary patch of elements, nodal displacements corresponding to any state of constant strain. If
A direct approach to problems in elasticity

Nodal equilibrium is simultaneously achieved without the imposition of external, nodal, forces and if a state of constant stress is obtained, then clearly no external work has been lost through interelement discontinuity.

Elements which pass such a patch test will converge, and indeed at times non-conforming elements will show a superior performance to conforming elements.

The patch test was first introduced by Irons and has since been demonstrated to give a sufficient condition for convergence. The concept of the patch test can be generalized to give information on the rate of convergence which can be expected from a given element.

We shall return to this problem in detail in Chapter 10 where the test will be fully discussed.

2.8 Bound on strain energy in a displacement formulation

While the approximation obtained by the finite element displacement approach always overestimates the true value of $U$, the total potential energy (the absolute minimum corresponding to the exact solution), this is not directly useful in practice. It is, however, possible to obtain a more useful limit in special cases.

Consider in particular the problem in which no 'initial' strains or initial stresses exist. Now by the principle of energy conservation the strain energy will be equal to the work done by the external loads which increase uniformly from zero. This work done is equal to $-\frac{1}{2} W$ where $W$ is the potential energy of the loads.

Thus

$$U + \frac{1}{2} W = 0 \quad (2.34)$$

or

$$\Pi = U + W = -U \quad (2.35)$$

whether an exact or approximate displacement field is assumed.

Thus in the above case the approximate solution always underestimates the value of $U$ and a displacement solution is frequently referred to as the lower bound solution.

If only one external concentrated load $R$ is present the strain energy bound immediately informs us that the deflection under this load has been underestimated (as $U = -\frac{1}{2} W = \frac{1}{2} r^T a$). In more complex loading cases the usefulness of this bound is limited as neither local deflections nor stresses, i.e., the quantities of real engineering interest, can be bounded.

It is important to remember that this bound on strain energy is only valid in the absence of any initial stresses or strains.

The expression for $U$ in this case can be obtained from Eq. (2.29) as

$$U = \frac{1}{2} \int_V \epsilon^T D \epsilon \, d(vol) \quad (2.36)$$

which becomes by Eq. (2.2) simply

$$U = \frac{1}{2} a^T \left[ \int_V B^T D B \, d(vol) \right] a = \frac{1}{2} a^T K a \quad (2.37)$$

a 'quadratic' matrix form in which $K$ is the 'stiffness' matrix previously discussed.
The above energy expression is always positive from physical considerations. It follows therefore that the matrix $K$ occurring in all the finite element assemblies is not only symmetric but is 'positive definite' (a property defined in fact by the requirements that the quadratic form should always be greater than or equal to zero).

This feature is of importance when the numerical solution of the simultaneous equations involved is considered, as simplifications arise in the case of 'symmetric positive definite' equations.

### 2.9 Direct minimization

The fact that the finite element approximation reduces to the problem of minimizing the total potential energy $\Pi$ defined in terms of a finite number of nodal parameters led us to the formulation of the simultaneous set of equations given symbolically by Eq. (2.31). This is the most usual and convenient approach, especially in linear solutions, but other search procedures, now well developed in the field of optimization, could be used to estimate the lowest value of $\Pi$. In this text we shall continue with the simultaneous equation process but the interested reader could well bear the alternative possibilities in mind.24,25

### 2.10 An example

The concepts discussed and the general formulation cited are a little abstract and readers may at this stage seek to test their grasp of the nature of the approximations derived. While detailed computations of a two-dimensional element system are performed using the computer, we can perform a simple hand calculation on a one-dimensional finite element of a beam. Indeed, this example will allow us to introduce the concept of generalized stresses and strains in a simple manner.

Consider the beam shown in Fig. 2.5. The generalized 'strain' here is the curvature. Thus we have

$$\varepsilon \equiv \kappa = -\frac{d^2 w}{dx^2}$$

where $w$ is the deflection, which is the basic unknown. The generalized stress (in the absence of shear deformation) will be the bending moment $M$, which is related to the 'strain' as

$$\sigma \equiv M = -EI \frac{d^2 w}{dx^2}$$

Thus immediately we have, using the general notation of previous sections,

$$D \equiv EI$$

If the displacement $w$ is discretized we can write

$$w \equiv Na$$

for the whole system or, for an individual element, $ij$. 
In this example the strains are expressed as the second derivatives of displacement and it is necessary to ensure that both \( w \) and its slope

\[
w_x \equiv \frac{dw}{dx} = \theta
\]

be continuous between elements. This is easily accomplished if the nodal parameters are taken as the values of \( w \) and the slope, \( \theta \). Thus,

\[
a_i = \begin{bmatrix} w_i \\ \theta_i \end{bmatrix}
\]

The shape functions will now be derived. If we accept that in an element two nodes (i.e., four variables) define the deflected shape we can assume this to be given by a cubic

\[
w = \alpha_1 + \alpha_2 s + \alpha_3 s^2 + \alpha_4 s^3 \quad \text{where} \quad s = \frac{x}{L}.
\]

This will define the shape functions corresponding to \( w_i \) and \( \theta_i \) by taking for each a cubic giving unity for the appropriate points \( (x = 0, L \) or \( s = 0,1) \) and zero for other quantities, as shown in Fig. 2.5.

The expressions for the shape functions can be written for the element shown as

\[
N_i = [1 - 3s^2 + 2s^3, L(s - 2s^2 + s^3)]
\]

\[
N_j = [3s^2 - 2s^3, L(-s^2 + s^3)]
\]

Immediately we can write

\[
B_i = -\frac{d^2 N_i}{dx^2} = \frac{1}{L^2} [6 - 12s, L(4 - 6s)]
\]

\[
B_j = -\frac{d^2 N_j}{dx^2} = \frac{1}{L^2} [-6 + 12s, L(2 - 6s)]
\]
and the stiffness matrices for the element can be written as

\[
K'_{ij} = \int_0^L B_i^T E I B_j \, dx = \frac{EI}{L^3} \begin{bmatrix}
12 & 6L & -12 & 6L \\
6L & 4L^2 & -6L & 2L^2 \\
-12 & -6L & 12 & -6L \\
6L & 2L^2 & -6L & 4L^2
\end{bmatrix}
\]

We shall leave the detailed calculation of this and the 'forces' corresponding to a uniformly distributed load \( p \) (assumed constant on \( ij \) and zero elsewhere) to the reader. It will be observed that the final assembled equations for a node \( i \) are of the form linking three nodal displacements \( i,j,k \). Explicitly these equations are for elements of equal length \( L \):

\[
EI \begin{bmatrix}
-12/L^3 & -6/L^2 \\
6/L^2 & 2/L
\end{bmatrix} \begin{bmatrix}
w_k \\
\theta_k
\end{bmatrix} + EI \begin{bmatrix}
24/L^3 & 0 \\
0 & 8/L
\end{bmatrix} \begin{bmatrix}
w_i \\
\theta_i
\end{bmatrix} + \begin{bmatrix}
pL/2 \\
-pL^2/12
\end{bmatrix} = 0
\]

It is of interest to compare these with the exact form represented by the so-called 'slope-deflection' equations which can be found in standard texts on structural analysis.

Here it will be found that the finite element approximation has achieved the exact solution at nodes for a uniform load. We show in Chapter 3 and in Appendix H reasons for this unexpected result.

### 2.11 Concluding remarks

The 'displacement' approach to the analysis of elastic solids is still undoubtedly the most popular and easily understood procedure. In many of the following chapters we shall use the general formulae developed here in the context of linear elastic analysis (Chapters 4, 5, and 6). These are also applicable in the context of non-linear analysis, the main variants being the definitions of the stresses, generalized strains, and other associated quantities. It is thus convenient to summarize the essential formulae, and this is done in Appendix C.

In Chapter 3 we shall show that the procedures developed here are but a particular case of finite element discretization applied to the governing equilibrium equations written in terms of displacements.\(^{26}\) Clearly, alternative starting points are possible. Some of these will be mentioned in Chapters 11 and 12.

### References


