# CHAPTER 8 ELEMENT TECHNOLOGY 

by Ted Belytschko<br>Northwestern University<br>@ Copyright 1997

### 8.1 Introduction

Element technology is concerned with obtaining elements with better performance, particularly for large-scale calculations and for incompressible materials. For large-scale calculations, element technology has focused primarily on underintegration to achieve faster elements. For three dimensions, cost reductions on the order of 8 have been achieved through underintegration. However, underintegration requires the stabilization of the element. Although stabilization has not been too popular in the academic literature, it is ubiquitous in large scale calculations in industry. As shown in this chapter, it has a firm theoretical basis and can be combined with multi-field weak forms to obtain elements which are of high accuracy.

The second major thrust of element technology in continuum elements has been to eliminate the difficulties associated with the treatment of incompressible materials. Loworder elements, when applied to incompressible materials, tend to exhibit volumetric locking. In volumetric locking, the displacements are underpredicted by large factors, 5 to 10 is not uncommon for otherwise reasonable meshes. Although incompressible materials are quite rare in linear stress analysis, in the nonlinear regime many materials behave in a nearly incompressible manner. For example, Mises elastic-plastic materials are incompressible in their plastic behavior. Though the elastic behavior may be compressible, the overall behavior is nearly incompressible, and an element that locks volumetrically will not perform well for Mises elastic-plastic materials. Rubbers are also incompressible in large deformations. To be applicable to a large class of nonlinear materials, an element must be able to treat incompressible materials effectively. However, most elements have shortcomings in their performance when applied to incompressible or nearly incompressible materials. An understanding of these shortcomings are crucial in the selection of elements for nonlinear analysis.

To eliminate volumetric locking, two classes of techniques have evolved:

1. multi-field elements in which the pressures or complete stress and strain fields are also considered as dependent variables;
2. reduced integration procedures in which certain terms of the weak form for the internal forces are underintegrated.

Multi-field elements are based on multi-field weak forms or variational principles; these are also known as mixed variational principles. In multi-field elements, additional variables, such as the stresses or strains, are considered as dependent, at least on the element level, and interpolated independently of the displacements. This enables the strain or stress fields to be designed so as to avoid volumetric locking. In many cases, the strain or stress fields are also designed to achieve better accuracy for beam bending problems. These methods cannot improve the performance of an element in general when there are no constraints
such as incompressibility. In fact, for a 4-node quadrilateral, only a 3 parameter family of elements is convergent and the rate of convergence can never exceed that of the 4 -node quadrilateral. Thus the only goals that can be achieved by mixed elements is to avoid locking and to improve behavior in a selected class of problems, such as beam bending.

The unfortunate byproduct of using multi-field variational principles is that in many cases the resulting elements posses instabilities in the additonal fields. Thus most 4-node quadrilaterals based on multi-field weak forms are subject to a pressure instability. This requires another fix, so that the resulting element can be quite complex. The developemnt of truly robust elements is not easy, particularly for low order elements. For this reason, an understanding of element technology is useful to anyone engaged in finite element analysis.

Elements developed by means of underintegration in its various forms are quite similar from a fundamental and practical viewpoint to elements based on multi-fieldvariational principles, and the equivalence was proven by Malkus and Hughes() for certain classes of elements. Therefore, while underintegration is more easily understood than multi-field approaches, the methods suffer from the same shortcomings as multi-field elements: pressure instabilities. Nevertheless, they provide a straightforward way to overcome locking in certain classes of elements.

We will begin the chapter with an overview of element performance in Section 8.1. This Section describes the characteristics of many of the most widely used elements for continuum analysis. The description is limited to elements which are based on polynomials of quadratic order or lower, since elements of higher order are seldom used in nonlinear analysis at this time. This will set the stage for the material that follows. Many readers may want to skip the remainder of the Chapter or only read selected parts based on what they have learned from this Section.

Although the techniques introduced in this Chapter are primarily useful for controlling volumetric locking for incompressible and nearly incompressible materials, they apply more generally to what can collectively be called constrained media problems. Another important class of such problems are structural problems, such as thin-walled shells and beams. The same techniques described in this Chapter will be used in Chapter 9 to develop beam and shell elements.

Section 8.3 describes the patch tests. These are important, useful tests for the performance of an element. Patch tests can be used to examine whether an element is convergent, whether it avoids locking and whether it is stable. Various forms of the patch test are described which are applicable to both static programs and programs with explicit time integration. They test both the underlying soundness of the approximations used in the elements and the correctness of the implementation.

Section 8.4 describes some of the major multi-field weak forms and their application to element development. Although the first major multi-field variational principle to be discovered for elasticity was the Hellinger-Reissner variational principle, it is not considered because it can not be readily used with strain-driven constitutive equations in nonlinear analysis. Therefore, we will confine ourselves to various forms of the HuWashizu principles and some simplifications that are useful in the design of new elements.

We will also describe some limitation principles and stability issues which pertain to mixed elements.

To illustrate the application of element technology, we will focus on the 4-node isoparametric quadrilateral element (QUAD4). This element is convergent for compressible material without any modifications, so none of the techniques described in this Chapter are needed if this element is to be used for compressible materials. On the other hand, for incompressible or nearly incompressible materials, this element locks. We will illustrate two classes of techniques to eliminate volumetric locking: reduced integration and multifield elements. We then show that reduced integration by one-point quadrature is rank deficient, which leads to spurious singular modes. To stabilize these modes, we first consider perturbation hourglass stabilization of Flanagan and Belytschko (1981). We then derive mixed methods for stabilization of Belytschko and Bachrach (1986), and assumed strain stabilization of Belytschko and Bindeman (1991). We show that assumed strain stabilization can be used with multiple-point quadrature to obtain better results when the material response is nonlinear without great increases in cost. The elements of Pian and Sumihara() and Simo and Rifai() are also described and compared. Numerical results are also presented to demonstrate the performance of various implementations of this element. Finally, the extension of these results to the 8-node hexhedron is sketched.

### 8.2. Overview of Element Performance

In this Section, we will provide an overview of characteristics of various widely-used elements with the aim of giving the reader a general idea of how these elements perform, their advantages and their major difficulties. This will provide the reader with an understanding of the consequences of the theoretical results and procedures which are described later in this Chapter. We will concentrate on elements in two dimensions, since the properties of these elements parallel those in three dimensions; the corresponding elements in three dimension will be specified and briefly discussed. The overview is limited to continuum elements; the properties of shell elements are described in Chapter 9.

In choosing elements, the ease of mesh generation for a particular element should be borne in mind. Triangles and tetrahedral elements are very attractive because the most powerful mesh generators today are only applicable to these elements. Mesh generators for quadrilateral elements tend to be less robust and more time consuming. Therefore, triangular and tetrahedral elements are preferable when all other performance characteristics are the same for general purpose analysis.

The most frequently used low-order elements are the three-node triangle and the fournode quadrilateral. The corresponding three dimensional elements are the 4-node tetrahedron and the 8 -node hexahedron. The detailed displacement and strain fields are given later, but as is well-known to anyone familiar with linear finite element theory, the displacement fields of the triangle and tetrahedron are linear and the strains are constant. The displacement fields of the quadrilateral and hexahedron are bilinear and trilinear, respectively. All of these elements can represent a linear displacement field and constant strain field exactly. Consequently they satisfy the standard patch test, which is described in Section 8.3. The satisfaction of the standard patch test insures that the elements converge
in linear analysis, and provide a good guarantee for convergent behavior in nonlinear problems also, although there are no theoretical proofs of this statement.

We will first discuss the simplest elements, the three-node triangle in two dimensions, the four-node tetrahedron in three dimensions. These are also known as simplex elements because a simplex is a set of $n+1$ points in $n$ dimensions. Neither simplex element performs very well for incompressible materials. Constant-strain triangular and tetrahedral elements are characterized by severe volumetric locking in two-dimensional plane strain problems and in three dimensions. They also manifest stiff behavior in many other cases, such as beam bending. For arbitrary arrangements of these elements, volumetric locking is very pronounced for materials such as Mises plasticity. The proviso plane strain is added here because volumetric locking will not occur in plane stress problems, for in plane stress the thickness of the element can change to accommodate incompressible materials. The consequences of volumetric locking are almost a complete lack of convergence. In the presence of volumetric locking, displacements are underpredicted by factors of 5 or more, so the results are completely worthless.

Volumetric locking does not preclude the use of simplex elements for incompressible materials completely, for locking can be avoided by using special arrangements of the elements. For example, the cross-diagonal arrangement of triangles shown in Fig.?? eliminates locking, Naagtegal et al. However, meshing in this arrangement is similar to meshing quadrilaterals, so the benefits arising from triangular and tetrahedral meshing are lost. In addition, this arrangement results in pressure oscillations, such as those described subsequently for quadrilaterals.

When fully integrated, i.e. $2 \times 2$ Gauss quadrature for the quadrilateral, both the 4 -node quadrilateral and the hexahedron lock for incompressible materials. Volumetric locking can be eliminated in these elements by using reduced integration, namely one-point quadrature, or selective-reduced integration, which consists of one-point quadrature on the volumetric terms and $2 \times 2$ quadrature on the deviatoric terms; this is described in detail later. The resulting quadrilateral will then exhibit good convergence properties in the displacements. However, the element still is plagued by one flaw: it exhibits pressure oscillations due to the failure of the quadrilateral with modified quadrature to satisfy the BB-condition, which is described later. As a consequence, the pressure field will often be oscillatory, with a pattern of pressures as shown in fig, ??. This oscillatory pattern in the pressures is often known as checkerboarding. Checkerboarding is sometimes harmless: for example, in materials governed by the Mises law the response is independent of pressure, so pressure oscillations are not very harmful, although they lead to errors in the elastic strains. Checkerboarding can also be eliminated by filtering procedures. Nevertheless it is undesirable, and a user of finite elements should at least be aware of its possibility with these elements. Pressure oscillations also occur for the mixed elements based on multifield variational principles. In fact mixed elements are in many cases identical or very similar in performance to selective reduced integration elements, since theoretically they are in many cases equivalent, Malkus and Hughes(). Some stabilization procedures for BB oscillations have been developed; they are described and discussed in Section ??.

In large scale computations, the fastest form of the quadrilateral and hexahedron is the one-point quadrature element: it is often 3 to 4 times as fast as the selective-reduced quadrature quadrilateral element. In three dimensions, the speedup is of the order of 6 to 8 .

The one-point quadrature element also suffers from pressure oscillations, and in addition possess instabilities in the displacement field. These instabilities are shown in Fig. ??, and have various names: hourglassing, keystoning, kinematic modes, spurious zero energy modes and chickenwiring are some of the appellations for these modes. The control of these modes has been the topic of considerable research, and they can be controlled quite effectively. In fact, the rate of convergence is not decreased by a consistent control of these modes, so for many large scale calculations, one-point quadrature with hourglass control are very effective. Hourglass control is described in Sections ???.

The next highest order elements are the 6 -node triangle and the 8 and 9 node quadrilaterals. The counterparts in three dimensions are the 10 node tetrahedron and the 20 and 27 node quadrilaterals. The 6 -node triangle and 9 -node quadrilateral have a complete quadratic displacement field and complete linear strain field when the edges of the element are straight. Ciarlet and Raviart() in a landmark paper proved that the convergence of these elements is quadratic when the displacement of the midside nodes is small compared to the length of the elements; whether the distortions introduced by a mesh in normal mesh generation are small is often an open question. These elements satisfy the quadratic and linear patch tests when the element sides are straight, but only the linear patch test when the element sides are curved. In other words, these elements cannot reproduce a quadratic displacement field exactly when the sides are not straight. Of course, curved sides are an intrinsic advantage of finite elements, for they enable boundary conditions to be met for higher order elements, but curved sides should only be used for exterior surfaces, since their presence decreases the accuracy of the element. In nonlinear problems with large deformations, the performance of these elements degrades when the midside nodes move substantially; this had already been discussed in the one-dimensional context in Example 2.8.2. Element distortion is a pervasive difficulty in the use of higher order elements for large-deformation analysis: the convergence rate of higher order elements degrades significantly as they are distorted, and in addition solution procedures often fail when distortion becomes excessive.

The 6-node triangle does not lock for incompressible materials, but it fails the BB pressure stability test for incompressible materials. The 9 -node quadrilateral when developed appropriately by a mixed variational principle with a linear pressure field satisfies the pressure stability test and does not lock. It is the only element we have discussed so far which has flawless behavior for incompressible materials.

In summary, element technology deals with two major quirks:

1. volumetric locking, which prevents convergence for incompressible and nearly incompressible materials;
2. pressure oscillations which result from the failure to meet the BB condition.

For low-order elements, the presence of one of these flaws is nearly unavoidable. The quadrilateral with reduced integration and a pressure stabilization or pressure filter appears to be the best of the low-order elements. When speed of computation is a consideration, a stabilized quadrilateral with one-point quadrature appears to be optimal. Only the 9 -node quadrilateral and 27 -node hexahedral are flawless elements for imcompressible materials, and the fact that no flaws have been discovered so far does not preclude that none will ever be discovered. Almost all of these difficulties are driven by incompressibility, and persist
for near-incompressiblity. When the material is compressible, or when considering two dimensional plane stress, standard element procedure can be used.

Error Norms. In order to compare these elements further, it is worthwhile to study a convergence theorem which has been proven for linear problems. Although this theorem has not been proven for the nonlinear regime, it provides insight into element accuracy. For the purpose of studying this convergence theorem, we will first define some norms frequently used in error analysis of finite elements. These will also be used to evaluate some of the element technology developed later in this Chapter.

Errors in finite element analysis are measured by norms. A norm in functional analysis is just a way of measuring the distance between two functions. A norm of the difference between a finite element solution and the exact solution to a problem is a measure of the error in the solution. The most common norms for the evaluating the error in a finite element solution are the $\mathcal{L}_{2}$ norm and the error in energy. The $\mathcal{L}_{2}$ norm of a vector function $f_{i}(\mathbf{x})$ is defined by

$$
\begin{equation*}
\mid f_{i}(\mathbf{x}) \|_{0}=\left(\int_{\Omega} f_{i}(\mathbf{x}) f_{i}(\mathbf{x}) d \Omega\right)^{\frac{1}{2}} \tag{8.2.1}
\end{equation*}
$$

where the subscript nought on the symbol for the norm designates the $\mathcal{L}_{2}$ norm. It can be seen that the $\mathcal{L}_{2}$ norm is always positive, and measures an average or mean value of the function. To use the $\mathcal{L}_{2}$ norm for a measure of error for a finite element solution, we denote the finite element solution for the displacement by $\mathbf{u}^{h}(\mathbf{x})$ and the exact solution by $\mathbf{u}(\mathbf{x})$. The error in the finite element solution at any point can then be expressed by the vector $\mathbf{e}(\mathbf{x})=\mathbf{u}^{h}(\mathbf{x})-\mathbf{u}(\mathbf{x})$. Since we seek a single number for the error, we will use the magnitude of the vector $\mathbf{e}(\mathbf{x})$. which is $\mathbf{e}(\mathbf{x}) \cdot \mathbf{e}(\mathbf{x})$. Thus we can define the error in the displacements by the $\mathcal{L}_{2}$ norm as

$$
\begin{equation*}
\mid \mathbf{e}(\mathbf{x}) \|_{0}=\left(\int_{\Omega} \mathbf{e}(\mathbf{x}) \cdot \mathbf{e}(\mathbf{x}) d \Omega\right)^{\frac{1}{2}} \text { or }\left\|\mathbf{u}-\mathbf{u}^{h}\right\|_{0}=\left(\int_{\Omega}\left(\mathbf{u}-\mathbf{u}^{h}\right) \cdot\left(\mathbf{u}-\mathbf{u}^{h}\right) d \Omega\right)^{\frac{1}{2}} \tag{8.2.2}
\end{equation*}
$$

This error norm measures the average error in the displacements over the domain of the problem. There are many other error norms. We have chosen to use this one because the most powerful and most well known results are expressed in terms of this norm. Furthermore, it gives a measure of error which is useful for engineering purposes.

The second norm we will consider are the norms in Hilbert space. The $\mathcal{H}_{1}$ norm of a vector function is defined by

$$
\begin{equation*}
\mid f_{i}(\mathbf{x}) \|_{1}=\left\{\int_{\Omega} f_{i}(\mathbf{x}) f_{i}(\mathbf{x})+f_{i, j}(\mathbf{x}) f_{i, j}(\mathbf{x}) d \Omega\right)^{\frac{1}{2}} \tag{8.2.3}
\end{equation*}
$$

This norm is a good measure of the error in the derivatives of a function. It includes the terms from the $\mathcal{L}_{2}$ norm given in Eq. (1), but when applied to real approximations, the errors in the derivatives dwarf the errors in the function itself, so they play an insignificant role. If we take the $\mathcal{H}_{1}$ norm of the error in the displacements, i.e. by letting $f_{i}(\mathbf{x})=u_{i}^{h}(\mathbf{x})-u_{i}(\mathbf{x})$ we otain a useful measure of the error in strains. The errors in norm, incidentally, are usually similar to the error in energy, Hughes(,p.273) which is defined by

$$
\begin{equation*}
\|a(\mathbf{u}, \mathbf{u})\|_{1}=\left\{\int_{\Omega} e_{i j} C_{i j k l} e_{k l} d \Omega\right)^{\frac{1}{2}} \quad e_{i j}=\varepsilon_{i j}-\varepsilon_{i j}^{h} \tag{8.2.4}
\end{equation*}
$$

Note the similarity of this form to the strain energy defined in Eq. (). In the above expression, the error in strain replaces the strain in Eq.(), hence the name energy error in strain.

We conclude with a few more facts on norms. The $\mathcal{H}_{\mathrm{r}}$ norm is generated in terms of the $r$ th derivatives of the function. Thus the $\mathcal{H}_{0}$ norm is equivalent to the $\mathcal{L}_{2}$ norm, Eq (1), whereas the $\mathcal{H}_{2}$ norm would involve the squares of the second derivatives. These norms exist, i.e. the integral corresponding to the norms $\mathcal{H}_{\mathrm{r}}$, is integrable, when the function is of continuity $\mathrm{C}^{\mathrm{r}-1}$. This can be seen quite easily for the $\mathcal{H}_{1}$ norm: if the function $f_{i}(\mathbf{x})$ is not $\mathrm{C}^{0}$, i.e. if it is discontinuious, then the derivatives will be Dirac delta functions at the points of the discontinuity. The square of a Dirac delta function cannot be integrated, so the norm can not be evaluated. The kinematic admissibility conditions () are often stated in terms of Hilbert spaces, so in Eq () the requirement could be replaced by . The latter is often found in the literature, but we used the simpler concept since we were not concerned with convergence proofs. For more on norms, seminorms, and other good stuff of this type see Hughes(), Oden and Reddy() or Strang().

Convergence Results for Linear Problems. The fundamental convergence results for linear finite elements is given in the following. If the finite element solution is generated by elements which can reproduce polynomials of order $k$, and if the solution $\mathbf{u}(\mathbf{x})$ is sufficiently smooth for the Hilbert norm $\mathcal{H}_{\mathrm{r}}$ to exist, then

$$
\begin{equation*}
\left|\mathbf{u}-\mathbf{u}^{h}\right|_{m} \leq C h^{\alpha} \mid u \|_{r}, \quad \alpha=\min (k+1-m, r-m) \tag{8.2.5}
\end{equation*}
$$

where $h$ is a measure of element size and C is an arbitrary constant which is independent of $h$ and varies from problem to problem.

We will now examine the implications of this theorem for various elements. The parameter $\alpha$ indicates the rate of convergence of the finite element solution: the greater the value of $\alpha$, the faster the finite element solution converges to the exact solution and therefore the more accurate the element. It is important to note that the rate of convergence is limited by the smoothness of the solution in space. An elastic solution is analytic, i.e. infinitely smooth, if there are no acute corners or cracks, so in that case $r$ tends toward infinity. Therefore, the second term in the definition for $\alpha, r-m$, plays no role for smooth solutions. However, if the solution is not very smooth, i.e. if there are discontinuities in the derivatives or higher order derivatives, then $r$ is finite. For example, if there are discontinuities in the second derivatives, then $r$ is at most 2 , and the second term plays a role.

We first examine what Eq. (5) means for smooth elastic solutions for various elements for the error in displacements. In that case, we consider the $\mathcal{H}_{0}$ norm, which is equivalent to the $L_{2}$ norm, so $m=0$. The 3-node triangle, the 4 -node quadrilateral, the 4 -node tetrahedron and the 8 -node hexahedron all reproduce in linear polynomials exactly; this result is proven for the isoparametric elements in Section 8.?; therefore $k=1$. Therefore, for the elements with linear completeness just listed we obtain that

$$
\alpha=\min (k+1-m, r-m)=\min (1+1-0, \infty-0)=2
$$

This result is illustrated in Fig. (), which shows a log-log plot of the error for the plate with a hole; details of this problem are given in Section ??. In a log-log plot, the graph of error in displacements versus element size according to Eq. (5) is given by a straight line with a slope $\alpha=2$. The solution in this case is said to converge quadratically. The actual numerical results compare with this theoretical result quite closely, although the slope deviates $5 \%$ or so from the theoretical result. Equation 5 is an aymptotic result which should hold only as the element size goes to zero, but remarkably it agrees very well with numerical experiments with realistic meshes.

We next consider the higher order elements, namely the 6 -node triangle, the 9 -node quadrilateral, the tetrahedron and the 27 -node hexahedron with straight edges. In this case $k=2$, and for an elastic solution the remaining constants are unchanged. We find then that $\alpha=3$, so the rate of convergence is cubic in the displacements. This increase of one order in convergence is quite significant, as illustrated in the results shown in Fig. ??. In effect, the choice of a higher order element here buys a tremendous amount of accuracy.

The results for the strains, i.e. the derivatives of the displacement field, are similar. In this case $m=1$ since the error in strains is indicated by the H1-norm. The rates of convergence are then one order lower, $\alpha=1$ for elements with linear completeness, $k=1$, and $\alpha=2$ for elements with quadratic completeness, $k=2$. the results are illustrated for a plate with a hole in fig. ??.

Convergence in Nonlinear Problems. The behavior of elements for nonlinear problems according to this theorem, Eq (), will depend on the smoothness of the constitutive equation. If the constitutive equation is very smooth, such as hyperelastic models for rubber, then the rate of convergence are expected to be the same as for elastic, linear
materials. However, for constitutive equations which are not smooth, such as elasticplastic materials, the second term in the definition of $\alpha$ governs the accuracy. For example in an elastic-plastic material, the relation between stress and strain is C 0 . Therefore the displacements are at most C 1 , and $r=2$. It can now be seen that the rate of convergence of the displacements now is at most of order 2, i.e. $\alpha=2$, regardless of whether the completeness of the element is as indicated by $k$ is linear or quadratic. Thus there appears to be no benefit in going to a higher order element for these materials. Similarly, the rate of convergence in the strains is at most of order $\alpha=1$. Thus, if the constitutive equation is not very smooth, the benefits of a higher order element can be lost.

In summary, for smooth constitutive equations, higher order elements are advantageous because of their higher rate of convergence. If the constitutive equation lacks sufficient smoothness, then there is no advantage in going to higher order elements. These results also are relevant for dynamic problems: when the signals are very smooth, there is some benefit in going to higher order elements, provided that a consistent mass matrix is used. For signals which lack smoothness, there is little advantage to higher order elements.

These statements do not take cognizance of the deterioration of element performance with large deformations. When the deformations are so large that the elements are highly distorted, then the accuracy of the higher order elements also decreases. These provisos pertain to both total and updated Lagrangian meshes, but not to Eulerian meshes. Thus, the amount of element distortion expected should also be considered in the choice of an element for nonlinear analysis.

It could be argued that even elastic problems in practical situations have discontinuities in derivatives due to different materials. However, in linear problems, the element edges are usually aligned with the material interfaces. In that case, the full accuracy of higher order elements can be retained since they can model discontinuities in derivatives effectively along element edges. In elastic-plastic problems, on the other hand, discontinuities float through the model and as the problem evolves, they proliferate. Thus their effects in nonlinear problems are more devastating to accuracy.

It should be stressed that the convergence results (5) has only been proven for linear problems. However, the major impediment to obtaining such convergence results for nonlinear problems is probably the lack of stability of nonlinear solutions. It is likely that the estimates given above, which are based on interpolation error estimates, play a similar role in nonlinear problems. This conjecture appears to be verified by numerical convergence studies which have verified that the estimates () apply in nonlinear problems quite well.

### 8.3. The Patch Tests

The patch tests are an extremely useful for examining the soundness of element formulations, for examining their stability and convergence behavior, and checking the proper implementation of an element in a compute program. The patch test was first conceived by Irons() to examine the soundness of a nonconforming plate element. In this original form, the patch test was primarily a test for polynomial completeness, i.e. the ability to reproduce exactly a polynomial of order k. It has been proposed by Strang() that
the patch test is necessary and sufficient for convergence. Subsequently, the patch test has been generalized and modified so that it can test also for stability in pressures and displacements, simo(), bathe(). Methods for implementing the patch test in explicit programs have also been developed, Belytschko(). Special versions of the patch test to test performance in large displacement analysis can also be constructed. In the following we describe these various forms of the patch test.

Before describing the patch test, it is worthwhile to define a few terms and point out a few overlaps in terms which are at times confusing. In functional analysis, the term completeness refers to the ability of an approximation to approximate a function arbitrarily closely. A sequence of functions $\phi_{I}(\mathbf{x})$ is complete in $\mathcal{H}_{r}$ if for any function $f(\mathbf{x}) \in \mathcal{H}_{r}$,

$$
\begin{equation*}
\mid f(\mathbf{x})-\sum_{I=1}^{n} a_{I} \phi_{I}(\mathbf{x}) \|_{r} \rightarrow 0 \text { as } n \rightarrow \infty \tag{8.2.3}
\end{equation*}
$$

Thus any set of functions is complete if it can approximate any function of a specified continuity arbitrarily closely, when the error is measured by an appropriate norm. The appropriate norm is any norm which exists or a lower order norm.

In the preceding we have referred to polynomial completeness. A better terms which has emerged in wavelet thoery is the reproducing condition. The reproducing condition is defined by the ability of an approximation to reproduce a function exactly. Thus for an interpolant such as a finite element shape function, the reproducing conditions state that if the nodal values of an element are given by $p_{i}\left(\mathbf{x}_{J}\right)$ where $p_{i}(\mathbf{x})$ is an arbitrary function, then

$$
\begin{equation*}
\sum_{J=1}^{m} N_{J}(\mathbf{x}) p_{i}\left(\mathbf{x}_{J}\right)=p_{i}(\mathbf{x}) \tag{8.3.4}
\end{equation*}
$$

This equation is quite subtle and contains more than first meets the eye. It states that when the reproducing condition holds, the shape functions or interpolants are able to exactly reproduce the given function $p_{i}(\mathbf{x})$. For example, if the shape functions are able to reproduce the constant and linear functions, then we have

$$
\begin{equation*}
\sum_{J=1}^{m} N_{J}(\mathbf{x})=1, \quad \sum_{J=1}^{m} N_{J}(\mathbf{x}) x_{j J}=x_{j} \tag{8.3.5}
\end{equation*}
$$

This is called linear completeness by Hughes(), but the term reproducing condition seems more appropriate, since completeness refers to a more general condition described by (4). Therefore, when using completeness in the sense of Hughes we will use the term polynomial completeness.

Any approximation which satisfies the linear reproducing conditions can be shown to be complete. On the other hand, the converse does not hold. Consider for example the Fourier series: they are complete, but they cannot reproduce a linear polynomial.

A third definition pertinent to convergence is the defintion of consistency. Consistency is usually defined in the context of finite difference methods. According to the standard definitions of consiostency, a discretization in space $\mathcal{D}(u)$ is a consistent approximation of a partial differential equation $L(u)=0$ if the error is on the order of the meshsize, i.e. if

$$
\mathcal{L}(u)-\mathcal{D}(u)=o\left(h^{n}\right), \quad n \geq 1
$$

The above states that the truncation error must tend to zero as the nodal spacing, i.e. the element size tends to zero. For time dependent problems, the disretization error will be a function of the time step and the element size h , and the truncation error will depend on both. For a time-independent problem in one dimension

Standard Patch Test. We first describe the standard patch test which checks for polynomial completeness of the displacement field, i.e. the ability of the element to reproduce polynomilas of a specified order. In addition, the test can be used to check the overall implementation of the element in the program; sometimes the shape functions are correct, but the element in a program fails the patch test anyway because of faulty programming.

In the standard patch test, a patch of elements such as shown in Fig.?? is used. The elements should be distorted as shown because the behavior of distorted elements is important and can differ from that of regular elements. No body forces should be applied, and the material properties should be uniform and linear elastic in the patch. The displacements of the nodes on the periphery of the patch are then prescribed according to the order of the patch test. For a linear patch test in two dimensions, the displacement field is given by

$$
\begin{aligned}
& u_{x}=a_{1 x}+a_{2 x} x+a_{3 x} y \\
& u_{y}=a_{1 y}+a_{2 y} x+a_{3 y} y
\end{aligned}
$$

where $a_{I i}$ are constants set by the user; they should all be nonzero to test the reproducing condition completely. This displacement field is used to set the prescribed displacements of the nodes on the periphery of the patch, so the prescribed displacements are

$$
\begin{aligned}
& u_{I x}=a_{1 x}+a_{2 x} x_{I}+a_{3 x} y_{I} \\
& u_{I y}=a_{1 y}+a_{2 y}{ }_{x}+a_{3 y} y_{I}
\end{aligned}
$$

To satisfy the patch test, the finite element solution should be given by () throughout the patch: the nodal displacements at the interior nodes should be given by () and the strains should be constant and given by the application of the strain-displacement equations to the displacement in ():

$$
\begin{aligned}
& \varepsilon_{x}=u_{x, x}=a_{2 x}, \quad \varepsilon_{y}=u_{y, y}=a_{3 y} \\
& 2 \varepsilon_{x y}=u_{x, y}+u_{y, x}=a_{3 x}+a_{2 y}
\end{aligned}
$$

The stresses should similarly be constant and correspond to what would be obtained by multiplying the above strains in the elastic, linear law used in the program. All of these conditions should be met to a high degree of precision, on the order of the precision in the computer used.

That rationale for the standard patch test are the reproducing condition and the fact that () corresponds to an exact solution to the governing equations for linear elasticity. It can easily be seen that () is an exact solution to the elastic problem. Since the strains are constant, and the material properties uniform, the stresses are constant. Therefore, since there are no body forces, the equilibrium equation () is satisfied exactly. Since linear elastic solutions are unique, Eq. () then represents a unique solution to the equations. If the finite element procedure is able to reproduce the linear field, it should be able to replicate this solution exactly because the trial functions include this solution!

When the patch test fails, then the finite element is either not complete, i.e. it can not reproduce the linear field exactly, or there is an error in the program in developing the discrete equations or in the solution of the discrete equations. Whether the reproducing conditions are satisfied can be checked independently by setting the nodal displacements according to () at all nodes and then checking the strains at all quadrature points. This test in fact suffices as a test of the reproducing conditions, and hence of convergence of the element. Going through the solution procedure is primarily a check on the program.

Patch Test in Explicit Programs. The patch test as applied above is not readily applicable to explicit programs because these program do not have a means for solving the linear static equations. However, the patch test can be modified for use in explicit programs as described in elytschko and Chiang(). The basic idea is to prescribe the intial velocities by a linear field identical to Eq.(), so

Here aij are arbitrary constant values, but they should be very small because in most programs the geometric nonlinearities will be triggered otherwise. The program is then used to integrate the equations of motion one time step; no extrenal forces should be applied and a linear, hypoelastic material model such as Eq. () should be used. The rate-ofdeformation or the strains and the accelerations at the end of time step are then checked. The rate-of-deformation should have the correct constant values in all of the elements and the accelerations should vanish at all of the interior nodes. The accelerations should vanish because the stresses should aslo be constant and according to the momentum equation, in the absence of body forces, the accelerations should vanish.

The test should be met to a high degree of precision if the constants aij are small enough. For example, when the constants aij are of order, the accelerations should not be larger than order

Patch Tests for Stability. Simo, Taylor and Z have devised a modified patch test with the aim of checking for stability, primarily in the stability of the displacement field rather than the instability of the pressures. It can also test whether the program treats traction boundary conditions exactly. The main difference from the standard patch test is that the displacements are not prescribed at all nodes. Instead, displacement boundary conditions
are prescribed only for the minimal number of components needed to prevent rigid body rotation. An example of the test is shown in Fig. ??.

This test is not an infallible test for detecting instabilities. Furthermore, it can only detect displacement instabilities, not pressure instabilities. To thoroughly check an element for displacement instabilities, it is also worthwhile to to an eigenvalue analysis on a single free element, i.e. a completely unconstrained element. The number of zero eigenvalues should be equal to the number of rigid body modes. For example, in two dimensional analysis, an element or a patch of elements should posess three zero eigenvalues, which are arise forom two translations and on erotation, whereas in three dimensions, an element should posses six zero eigenvalues, three translational and three rotational rigid body modes. If there are more zero eigenvalues, this indicates an element which may exhibit displacement instabilities; this characteristic is also called rank deficiency of the stiffness matrix, as discussed in Sectiopn 8.?.

### 8.6. Isoparametric Element 4-Node Quadrilateral

In this Section, the isoparametric elements are developed in two dimensions, with an emphasis on the 4 -node quadrilateral. The objective is to present a setting in which we can explain some of the concepts described in the preceding Sections. The displacement field for QUAD4 is given by

$$
\begin{equation*}
\mathrm{u}_{\mathrm{x}}(\xi, \eta)=\sum_{\mathrm{I}=1}^{4} \mathrm{~N}_{1}(\xi, \eta) \mathrm{u}_{\mathrm{xl}} \quad \mathrm{u}_{\mathrm{y}}(\xi, \eta)=\sum_{\mathrm{I}=1}^{4} \mathrm{~N}_{\mathrm{l}}(\xi, \eta) \mathrm{u}_{\mathrm{yl}} \tag{8.2.1}
\end{equation*}
$$

where $\mathrm{N}_{\mathrm{I}}$ is the isoparametric shape function for node I given by

$$
\begin{equation*}
\mathrm{N}(\xi, \eta)=\frac{1}{4}(1+\xi \xi)(1+\eta \eta) \tag{1.2.2}
\end{equation*}
$$

$\mathrm{u}_{\mathrm{xI}}$ and $\mathrm{u}_{\mathrm{yI}}$ are the displacements at node I , and $\mathrm{u}_{\mathrm{x}}(\xi, \eta)$ and $\mathrm{u}_{\mathrm{y}}(\xi, \eta)$ give the displacement field within the element domain. The displacement field is written in terms of a reference coordinate system $(\xi, \eta)$. Within the reference system, the element domain is a bi-unit square as shown in Fig. 8.6.1.


Figure 8.6.1. Element domain in the physical and reference coordinate systems

The transformation (or mapping) between the physical domain and element or parent domain is given by

$$
\begin{align*}
& x(\xi, \eta)=\sum_{\mathrm{I}=1}^{4} \mathrm{~N}_{\mathrm{N}}(\xi, \eta) \mathrm{x}_{\mathrm{l}}  \tag{1.2.3a}\\
& \mathrm{y}(\xi, \eta)=\sum_{\mathrm{I}=1}^{4} \mathrm{~N}_{1}(\xi, \eta) \mathrm{y}_{\mathrm{l}} \tag{1.2.3b}
\end{align*}
$$

where $x_{I}$ and $y_{I}$ are the nodal coordinates. Equations (3a) and (3b) can also be written in the form

$$
\begin{equation*}
\mathrm{x}_{\mathrm{i}}=\sum_{\mathrm{I}=1}^{4} \mathrm{~N}_{\mathrm{I}} \mathrm{X}_{\mathrm{iI}}=\mathbf{N} \mathbf{x}_{\mathrm{i}} \tag{1.2.3c}
\end{equation*}
$$

where $\mathbf{N}$ is a row matrix consisting of the 4 shape functions

$$
\mathbf{N}=\left(\mathbf{N}_{1}, \mathrm{~N}_{2}, \mathrm{~N}_{3}, \mathrm{~N}_{4}\right)
$$

and

$$
\begin{aligned}
& \mathbf{x}_{1}^{\mathrm{t}}=\mathbf{x}^{\mathrm{t}}=\left(\mathrm{x}_{1}, \mathrm{x}_{2}, \mathrm{x}_{3}, \mathrm{x}_{4}\right) \\
& \mathbf{x}_{2}^{\mathrm{t}}=\mathbf{y}^{\mathrm{t}}=\left(\mathrm{y}_{1}, \mathrm{y}_{2}, \mathrm{y}_{3}, \mathrm{y}_{4}\right)
\end{aligned}
$$

Because the same shape functions are used for both the mapping and the displacement interpolation, this element is called an isoparametric element.

The interpolants and mapping, Eq. (2), are bilinear in $(\xi, \eta)$, that is, they contain the following monomials: $(1, \xi, \eta, \xi \eta)$; the last term is called the bilinear term. Thus $u_{x}$ can be written as

$$
\begin{equation*}
\mathrm{u}_{\mathrm{x}}(\xi, \eta)=\alpha_{0 \mathrm{x}}+\alpha_{1 \mathrm{x}} \xi+\alpha_{2 \mathrm{x}} \eta+\alpha_{3 \mathrm{x}} \xi \eta \tag{1.2.4}
\end{equation*}
$$

where $\alpha_{i x}$ are constants. It can easily be verified that the interpolants are linear along each of the edges of the element as follows. Along any of the edges, either $\xi$ or $\eta$ is constant, so the monomial $\xi \eta$ is linear along the edges. Therefore, while the bilinear term is nonlinear within the element, it is linear on the edges. Therefore compatibility, or continuity, of the displacement is assured when elements share two nodes along any edge. QUAD4 can be mixed with linear displacement triangles without any discontinuities.
1.2.1 Strain Field. The strain field is obtained by using Eq. (1). Implicit differentiation is used to evaluate the derivatives because the shape functions are functions of $\xi$ and $\eta$ and the relation (3) can not be inverted explicitly to obtain $\xi$ and $\eta$ in terms of $x$ and $y$. Writing the chain-rule for a shape function in matrix form gives:

$$
\mathbf{J}\left\{\begin{array}{c}
\frac{\partial \mathrm{N}_{\mathrm{I}}}{\partial \mathrm{x}}  \tag{1.2.5a}\\
\frac{\partial \mathrm{~N}_{\mathrm{I}}}{\partial \mathrm{y}}
\end{array}\right\}=\left\{\begin{array}{c}
\frac{\partial \mathrm{N}_{\mathrm{I}}}{\partial \xi} \\
\frac{\partial \mathrm{~N}_{\mathrm{I}}}{\partial \eta}
\end{array}\right\}
$$

where $\mathbf{J}$ is the Jacobian matrix given by

$$
\mathbf{J}=\left[\begin{array}{cc}
\frac{\partial x}{\partial \xi} & \frac{\partial y}{\partial \xi}  \tag{1.2.5b}\\
\frac{\partial x}{\partial \eta} & \frac{\partial y}{\partial \eta}
\end{array}\right]
$$

Its determinant is denoted by $\mathbf{J}$, i.e.

$$
\begin{equation*}
\mathbf{J}=\operatorname{det}(\mathbf{J}) \tag{1.2.6}
\end{equation*}
$$

If we invert (5b), and multiply both sides of (5a) by the inverse, we obtain

$$
\left\{\begin{array}{c}
\frac{\partial \mathrm{N}_{\mathrm{I}}}{\partial \mathrm{x}}  \tag{1.2.7a}\\
\frac{\partial \mathrm{~N}_{\mathrm{I}}}{\partial \mathrm{y}}
\end{array}\right\}=\frac{1}{J}\left[\begin{array}{cc}
\frac{\partial \mathrm{y}}{\partial \eta} & \frac{-\partial \mathrm{y}}{\partial \xi} \\
\frac{-\partial \mathrm{x}}{\partial \eta} & \frac{\partial \mathrm{x}}{\partial \xi}
\end{array}\right]\left\{\begin{array}{c}
\frac{\partial \mathrm{N}_{\mathrm{I}}}{\partial \xi} \\
\frac{\partial \mathrm{~N}_{\mathrm{I}}}{\partial \eta}
\end{array}\right\}
$$

from which we see by the chain rule that

$$
\left[\begin{array}{cc}
\frac{\partial \xi}{\partial x} & \frac{\partial \eta}{\partial x}  \tag{1.2.7b}\\
\frac{\partial \eta}{\partial x} & \frac{\partial \eta}{\partial y}
\end{array}\right]=\frac{1}{J}\left[\begin{array}{cc}
\frac{\partial y}{\partial \eta} & \frac{-\partial y}{\partial \xi} \\
\frac{-\partial x}{\partial \eta} & \frac{\partial x}{\partial \xi}
\end{array}\right]
$$

The derivatives of the spatial coordinates with respect to $\xi$ and $\eta$ can be obtained from (3) and (2). First

$$
\begin{align*}
& \frac{\partial \mathrm{x}}{\partial \xi}=\frac{1}{4} \sum_{\mathrm{I}=1}^{4} \mathrm{x}_{\mathrm{l}} \xi(1+\eta \eta)  \tag{1.2.8a}\\
& \frac{\partial \mathrm{x}}{\partial \eta}=\frac{1}{4} \sum_{\mathrm{I}=1}^{4} \mathrm{x}_{\mathrm{I}} \eta_{1}(1+\xi \xi) \tag{1.2.8b}
\end{align*}
$$

$$
\begin{align*}
& \frac{\partial \mathrm{y}}{\partial \xi}=\frac{1}{4} \sum_{\mathrm{I}=1}^{4} \mathrm{y} \xi\left(\left(1+\eta_{\mathrm{I}} \eta\right)\right.  \tag{1.2.8c}\\
& \frac{\partial \mathrm{y}}{\partial \eta}=\frac{1}{4} \sum_{\mathrm{I}=1}^{4} \mathrm{ym} \eta\left(1+\xi_{\mathrm{I}} \xi\right) \tag{1.2.8d}
\end{align*}
$$

Using the definitions of $\mathbf{J}$ and $\mathbf{J}$ given in Eqs. (5) and (6), respectively, then gives

$$
\begin{align*}
& \mathrm{J}=\frac{1}{8}\left[\mathrm{x}_{24} \mathrm{y}_{31}+\mathrm{x}_{31} \mathrm{y}_{42}+\left(\mathrm{x}_{21} \mathrm{y}_{34}+\mathrm{x}_{34} \mathrm{y}_{12}\right) \xi+\left(\mathrm{x}_{14} \mathrm{y}_{32}+\mathrm{x}_{32} \mathrm{y}_{41}\right) \eta\right]  \tag{1.2.9a}\\
& \mathrm{x}_{\mathrm{IJ}} \equiv \mathrm{x}_{\mathrm{I}}-\mathrm{x}_{\mathrm{J}}  \tag{1.2.9b}\\
& \mathrm{y}_{\mathrm{IJ}} \equiv \mathrm{y}_{\mathrm{I}}-\mathrm{y}_{\mathrm{J}} \tag{1.2.9c}
\end{align*}
$$

Note that the bilinear term is absent in J .
Using the definition of the linear strain gives the following

$$
\left.\begin{array}{l}
\left\{\begin{array}{c}
\varepsilon_{\mathrm{x}} \\
\varepsilon_{\mathrm{y}} \\
2 \varepsilon_{\mathrm{xy}}
\end{array}\right\}=\sum_{\mathrm{I}=1}^{4}\left[\begin{array}{cc}
\frac{\partial \mathrm{N}_{\mathrm{I}}}{\partial \mathrm{x}} & 0 \\
0 & \frac{\partial \mathrm{~N}_{\mathrm{I}}}{\partial \mathrm{y}} \\
\frac{\partial \mathrm{~N}_{\mathrm{I}}}{\partial \mathrm{y}} & \frac{\partial \mathrm{~N}_{\mathrm{I}}}{\partial \mathrm{x}}
\end{array}\right]\left\{\begin{array}{cc}
\mathrm{u}_{\mathrm{xI}} \\
\mathrm{u}_{\mathrm{yI}}
\end{array}\right]=\left[\begin{array}{cc}
\frac{\partial \mathbf{N}}{\partial \mathrm{x}} & \mathbf{0} \\
\mathbf{0} & \frac{\partial \mathbf{N}}{\partial \mathrm{y}} \\
\frac{\partial \mathbf{N}}{\partial \mathrm{y}} & \frac{\partial \mathbf{N}}{\partial \mathrm{x}}
\end{array}\right]\left\{\mathbf{u}_{\mathrm{x}}\right. \\
\mathbf{u}_{\mathrm{y}}
\end{array}\right\} \equiv \mathbf{B d} \text {. }
$$

1.2.2 Linear Reproducing Conditions of Isoparametric Elements. It will now be shown that isoparametric elements of any order reproduce the complete linear velocity (displacement) field. This property is called linear completeness. It guarantees that the element will pass the linear patch test and is essential for the element to be convergent.

A general isoparametric element with $\mathrm{n}_{\mathrm{N}}^{\mathrm{e}}$ nodes is considered because it is easy to demonstrate this property for any isoparametric element. The number of spatial dimensions denoted by $n_{D}^{e}$. The isoparametric transformation is

$$
\begin{equation*}
\mathrm{x}_{\mathrm{i}}=\sum_{\mathrm{I}=1}^{\mathrm{n}_{\mathrm{N}}^{\mathrm{e}}} \mathrm{~N}_{\mathrm{l}}(\mathbf{x}) \mathrm{x}_{\mathrm{il}} \tag{1.2.11}
\end{equation*}
$$

where $\mathrm{i}=1$ to $\mathrm{n}_{\mathrm{D}}^{\mathrm{e}}$. The dependent variable is denoted by u . In the case of two or three dimensional solids, u may refer to any displacement component. For an isoparametric
element, the displacement field is interpolated by the same shape functions used in the mapping (12), so

$$
\begin{equation*}
\mathrm{u}=\sum_{\mathrm{I}=1}^{\mathrm{n}_{\mathrm{N}}^{\mathrm{e}}} \mathrm{~N}_{\mathrm{l}}(\mathbf{x}) \mathrm{u}_{\mathrm{l}} \tag{1.2.12}
\end{equation*}
$$

Consider the situation where the displacement field is linear

$$
\begin{equation*}
\mathrm{u}=\alpha_{\mathrm{o}}+\sum_{\mathrm{i}=1}^{\mathrm{n}_{\mathrm{D}}^{\mathrm{e}}} \alpha_{\mathrm{i}} \mathrm{x}_{\mathrm{i}} \tag{1.2.13}
\end{equation*}
$$

so the nodal displacements are given by

$$
\begin{equation*}
\mathrm{u}_{\mathrm{I}}=\alpha_{\mathrm{o}}+\sum_{\mathrm{i}=1}^{\mathrm{n}_{\mathrm{D}}^{\mathrm{e}}} \alpha_{\mathrm{i}} \mathrm{x}_{\mathrm{il}} \tag{1.2.14}
\end{equation*}
$$

where $\alpha_{o}$ and $\alpha_{i}$ are constants. This can also be written as

$$
\begin{equation*}
\mathrm{u}_{\mathrm{I}}=\alpha_{\mathrm{o}} \mathrm{~s}_{\mathrm{I}}+\sum_{\mathrm{i}=1}^{\mathrm{n}_{\mathrm{D}}^{\mathrm{e}}} \alpha_{\mathrm{i}} \mathrm{x}_{\mathrm{il}} \tag{1.2.15a}
\end{equation*}
$$

or

$$
\begin{equation*}
\mathbf{u}=\alpha_{o} \mathbf{s}+\sum_{\mathrm{i}=1}^{\mathrm{n}_{\mathrm{D}}^{\mathrm{e}}} \alpha_{\mathrm{i}} \mathbf{x}_{\mathrm{i}} \tag{1.2.15b}
\end{equation*}
$$

where $\mathbf{u}$ and $\mathbf{x}_{i}$ are column vectors of the nodal unknowns and coordinates; $\mathbf{s}$ is a column vector of the same dimension consisting of all 1's. Substituting (14) into (12) yields

$$
\begin{equation*}
\mathrm{u}=\sum_{\mathrm{I}=1}^{\mathrm{n}_{\mathrm{N}}^{\mathrm{e}}}\left(\alpha_{\mathrm{o}} \mathrm{~s}_{\mathrm{I}}+\sum_{\mathrm{i}=1}^{\mathrm{n}_{\mathrm{D}}^{\mathrm{e}}} \alpha_{\mathrm{i}} \mathrm{x}_{\mathrm{iI}}\right) \mathrm{N}(\mathbf{x}) \tag{1.2.16}
\end{equation*}
$$

and rearranging the terms

$$
\begin{equation*}
\mathrm{u}=\alpha_{\mathrm{o}} \sum_{\mathrm{I}=1}^{\mathrm{n}_{\mathrm{N}}^{\mathrm{e}}} \mathrm{~s}_{\mathrm{i}} \mathrm{~N}(\mathbf{x})+\sum_{\mathrm{i}=1}^{\mathrm{n}_{\mathrm{D}}^{\mathrm{e}}}\left(\alpha_{\mathrm{i}} \sum_{\mathrm{I}=1}^{\mathrm{n}_{\mathrm{N}}^{\mathrm{e}}} \mathrm{x}_{\mathrm{i}} \mathrm{~N}(\mathbf{x})\right) \tag{1.2.17}
\end{equation*}
$$

It is recognized from (11) that the coefficients of $\alpha_{i}$ on the right hand side of Eq (17) correspond to $\mathrm{x}_{\mathrm{i}}$ so

$$
\begin{equation*}
\mathrm{u}=\alpha_{\mathrm{o}} \sum_{\mathrm{I}=1}^{\mathrm{n}_{\mathrm{N}}^{\mathrm{e}}} \mathrm{~s}_{\mathrm{i}} \mathrm{~N}_{( }(\mathbf{x})+\sum_{\mathrm{i}=1}^{\mathrm{n}_{\mathrm{D}}^{\mathrm{e}}} \alpha_{\mathrm{i}} \mathrm{x}_{\mathrm{i}} \tag{1.2.18}
\end{equation*}
$$

We now make use of the fact that

$$
\begin{equation*}
\sum_{\mathrm{I}=1}^{\mathrm{n}_{\mathrm{N}}^{\mathrm{e}}} \mathrm{~s}_{\mathrm{N}} \mathrm{~N}_{\mathrm{I}}=\sum_{\mathrm{I}=1}^{\mathrm{n}_{\mathrm{N}}^{\mathrm{e}}} \mathrm{~N}(\mathbf{x})=1 \tag{1.2.19}
\end{equation*}
$$

The first equality is obvious since $\mathrm{s}_{\mathrm{I}}=1$. To obtain the second equality consider an element whose nodes are coincident: $\mathrm{x}_{\mathrm{iI}}=\mathrm{x}_{\mathrm{io}}$ for $\mathrm{I}=1$ to $\mathrm{n}_{\mathrm{N}}^{\mathrm{e}}$. The mapping (11) must then yield

$$
\begin{align*}
& \mathrm{x}_{\mathrm{io}}=\sum_{\substack{\mathrm{I}=1 \\
\mathrm{n}_{\mathrm{N}}^{\mathrm{e}}}} \mathrm{~N}(\mathbf{x}) \mathrm{x}_{\mathrm{io}}  \tag{1.2.20a}\\
& =\mathrm{x}_{\mathrm{io}} \sum_{\mathrm{i}=1}^{\mathrm{n}} \mathrm{~N}_{\mathrm{N}} \mathrm{~N}(\mathbf{x}) \tag{1.2.20b}
\end{align*}
$$

Since the above must hold for arbitrary $\mathrm{x}_{\mathrm{i}}$, the second equality in (19) follows.
Making use of (19) then reduces (18) to

$$
\begin{equation*}
\mathrm{u}=\alpha_{\mathrm{o}}+\sum_{\mathrm{i}=1}^{\mathrm{n}_{\mathrm{D}}^{\mathrm{e}}} \alpha_{\mathrm{i}} \mathrm{x}_{\mathrm{i}} \tag{1.2.21}
\end{equation*}
$$

which is precisely the linear field (13) from which the nodal values $u_{I}$ were defined in Eq. (15). Thus any isoparametric element contains the linear field and will exhibit constant strain fields when the nodal displacements emanate from a linear field. As a consequence, it satisfies the linear patch test exactly.

Although this attribute of isoparametric elements appears at first somewhat trivial, its subtlety can be appreciated by noting that the bilinear terms $x y$ will not be represented exactly in a 4-node isoparametric element. Consider for example the case when the nodal displacements are obtained from the bilinear field $u(x, y)=x y$ :

$$
\begin{equation*}
\mathrm{u}(\mathrm{x}, \mathrm{y})=\sum_{\mathrm{I}=1}^{4} \mathrm{u}_{\mathrm{N}} \mathrm{~N}(\xi, \eta)=\sum_{\mathrm{I}=1}^{4} \mathrm{x}_{\mathrm{I}} \mathrm{y} \mathrm{~N}_{\mathrm{N}}(\xi, \eta) \tag{1.2.22}
\end{equation*}
$$

It is impossible to extricate $x y$ from the right hand side of Eq. (22) unless $\mathbf{x}^{t}=\mathrm{a}(-1$, $+1,+1,-1), \mathbf{y}^{\mathrm{t}}=\mathrm{b}(-1,-1,+1,+1)$ where a and b are constants, i.e. when the element is rectangular. Therefore, for an arbitrary quadrilateral, the displacement field is not bilinear when the nodal values are determined from a bilinear field, i.e., when $u_{I}=x_{I} y_{I}$, $u(x, y) \neq x y$.

Similarly, for higher order isoparametrics, such as the 9-node Lagrange element, the distribution within the element is not quadratic when the nodal values of $u$ are obtained from a quadratic field unless the element is rectangular with equispaced nodes. For curved edges, the deviation of the field from quadratic is substantial, and the accuracy diminishes. The convergence proofs of Ciarlet and Raviart (1972) show that the order of convergence for the 9 -node element is better than the 4 -node quadrilateral only when the element midpoint nodes are displaced from the midpoint of the side by a small amount.

The linear completeness of subparametric elements can be shown analogously. In a subparametric element, the mapping is of lower order than the interpolation of the dependent variable. For example, consider the element in Fig. 2 that has a 4-node bilinear mapping with a 9 -node interpolation for $\mathrm{u}(\mathrm{x}, \mathrm{y})$. This is written

$$
\begin{equation*}
\mathrm{u}(\mathrm{x}, \mathrm{y})=\sum_{\mathrm{I}=1}^{9} \mathrm{u}_{\mathrm{N}} \overline{\mathrm{~N}}(\xi, \eta) \tag{1.2.23}
\end{equation*}
$$

$$
\left\{\begin{array}{l}
\mathrm{x}  \tag{1.2.24}\\
\mathrm{y}
\end{array}\right\}=\sum_{\mathrm{I}=1}^{4}\left\{\begin{array}{l}
\mathrm{x}_{\mathrm{I}} \\
\mathrm{y}_{\mathrm{I}}
\end{array}\right\} \mathrm{N}(\xi, \eta)
$$

The 9-node Lagrange interpolant for the dependent variable $u$ is distinguished from the 4node interpolant used for the element mapping by a superposed bar. We now define a set of 9 nodes $\left(\bar{x}_{\mathrm{b}}, \bar{y}_{\mathrm{I}}\right), \quad \mathrm{I}=1$ to 9 , which are obtained by evaluating $(\mathrm{x}, \mathrm{y})$ at the 9 -nodes used for interpolating $\mathrm{u}(\mathrm{x}, \mathrm{y})$ by Eq. (24). Then the mapping can be expressed as

$$
\left\{\begin{array}{l}
\mathrm{x}  \tag{1.2.25}\\
\mathrm{y}
\end{array}\right\}=\sum_{\mathrm{I}=1}^{9}\left\{\begin{array}{l}
\overline{\mathrm{x}}_{\mathrm{I}} \\
\overline{\mathrm{y}}_{\mathrm{I}}
\end{array}\right\} \overline{\mathrm{N}}(\mathrm{\xi}, \eta)
$$

Using (23) and (25), the arguments invoked in going from Eqs. (13) to (21) can be repeated to establish the linear completeness of the subparametric element.

Superparametric elements, in which the mapping is of higher order than the interpolation of the dependent variable, are not complete. This can by shown by considering the element in Fig. 2 with 9-node mapping and 4-node bilinear interpolation for $u(x, y)$. In order to use the previous argument, we would have to use the 4 nodes used for interpolation to do a bilinear mapping, but such a mapping would be unable to reproduce the domain of the element unless it has straight edges with the nodes at the midpoints of the nodes.


Figure 2. Examples of subparametric and superparametric elements
In summary, it has been shown that isoparametric and subparametric elements are linearly complete and consistent in that they represent linear fields exactly. This implies that when the nodal values are prescribed by a linear field, the interpolant is an identical linear field and the derivative of the interpolant has the correct constant value throughout the element. Therefore, for these elements, the correct constant strain state is obtained for a linear displacement field, and the patch test will be satisfied. The element will also represent rigid body translation and rotation exactly. The 4 -node quadrilateral considered here is isoparametric, so it possesses these necessary features. A superparametric element does not have linear completeness, and will therefore fail the patch test.
1.2.3 Element Rank and Rank Deficiency. In order to perform reliably, an element must have the proper rank. If its rank is too small, the global stiffness may be singular or near singular; in the latter case, it will exhibit spurious singular modes. If the rank of an element is too large, it will strain in rigid body motion and either fail to converge or converge very slowly.

The proper rank of an element stiffness is given by

$$
\begin{align*}
& \text { proper } \operatorname{rank}\left(\mathbf{K}_{\mathrm{e}}\right)=\operatorname{dim}\left(\mathbf{K}_{\mathrm{e}}\right)-\mathrm{n}_{\mathrm{RB}}  \tag{1.2.26a}\\
& \operatorname{rank} \operatorname{deficiency}\left(\mathbf{K}_{\mathrm{e}}\right)=\operatorname{proper} \operatorname{rank}\left(\mathbf{K}_{\mathrm{e}}\right)-\operatorname{rank}\left(\mathbf{K}_{\mathrm{e}}\right) \tag{1.2.26b}
\end{align*}
$$

where $n_{R B}$ is the number of rigid body modes. Another way of expressing this is that if the element is of proper rank, then

$$
\operatorname{dim}\left(\operatorname{ker}\left(\mathbf{K}_{\mathrm{e}}\right)\right)=\mathrm{n}_{\mathrm{RB}}
$$

where the kernel of $\mathbf{K}_{\mathrm{e}}$ is defined by

$$
\begin{equation*}
\mathbf{x} \in \operatorname{ker}\left(\mathbf{K}_{\mathrm{e}}\right) \text { if } \mathbf{K}_{\mathrm{e}} \mathbf{x}=0 \tag{1.2.27}
\end{equation*}
$$

To determine the rank of an element stiffness which is evaluated by numerical quadrature, consider the quadrature formula

$$
\begin{align*}
\mathbf{K}_{\mathrm{e}} & =\int_{\Omega_{\mathrm{e}}} \mathbf{B}^{\mathrm{t}} \mathbf{C B} \mathrm{~d} \Omega=\int_{-1}^{+1} \int_{-1}^{+1} \mathbf{B}^{\mathrm{t}} \mathbf{C} \mathbf{B} \mathrm{Jd} \xi \mathrm{~d} \eta \\
& =\sum_{\alpha=1}^{\mathrm{n}_{\mathrm{Q}}} \mathrm{w}_{\alpha} \mathrm{J}\left(\xi_{\alpha}\right) \mathbf{B}^{\mathrm{t}}\left(\xi_{\alpha}\right) \mathbf{C}\left(\xi_{\alpha}\right) \mathbf{B}\left(\xi_{\alpha}\right) \tag{1.2.28}
\end{align*}
$$

where $\mathbf{C}$ is a constitutive matrix, $\Omega_{\mathrm{e}}$ is the element domain, $\mathrm{w}_{\alpha}$ are the quadrature weights, and $\xi_{\alpha}$ are the $\mathrm{n}_{\mathrm{Q}}$ quadrature points. In Gauss quadrature, $\mathrm{w}_{\alpha}$ correspond to the products of the one-dimensional weight factors and $\xi_{\alpha}$ are the quadrature points in the reference coordinates. The element domain in (28) and throughout this discussion is assumed to have unit thickness. The above form can be written as

$$
\begin{equation*}
\mathbf{K}_{\mathrm{e}}=\stackrel{\circ}{\mathbf{B}}{ }^{\mathrm{t}} \stackrel{\mathrm{C}}{\mathbf{C}}{ }^{\circ} \tag{1.2.29a}
\end{equation*}
$$

where

$$
\stackrel{\mathrm{o}}{\mathbf{B}}=\left[\begin{array}{c}
\mathbf{B}\left(\mathbf{x}_{1}\right)  \tag{1.2.29b}\\
\mathbf{B}\left(\mathbf{x}_{2}\right) \\
\vdots \\
\mathbf{B}\left(\mathbf{x}_{\mathrm{n}_{\mathrm{Q}}}\right)
\end{array}\right]
$$

$$
\stackrel{\circ}{\mathbf{C}}=\left[\begin{array}{cccc}
\mathrm{w}_{1} \mathrm{~J}\left(\mathbf{x}_{1}\right) \mathbf{C}\left(\mathbf{x}_{1}\right) & \mathbf{0} & \cdots & \mathbf{0}  \tag{1.2.29c}\\
\mathbf{0} & \mathrm{w}_{2} \mathrm{~J}\left(\mathbf{x}_{2}\right) \mathbf{C}\left(\mathbf{x}_{2}\right) & & \mathbf{0} \\
\vdots & & \ddots & \vdots \\
\mathbf{0} & \mathbf{0} & \cdots & \mathrm{w}_{\mathrm{n}_{\mathrm{Q}}} \mathrm{~J}\left(\mathbf{x}_{\mathrm{n}_{\mathrm{Q}}}\right) \mathbf{C}\left(\mathbf{x}_{\mathrm{n}_{\mathrm{Q}}}\right)
\end{array}\right]
$$

A special form of the product-rank theorem is now used, which states that when $\stackrel{\circ}{\mathbf{C}}$ is positive definite

$$
\begin{equation*}
\operatorname{rank} \mathbf{K}_{\mathrm{e}}=\operatorname{rank}\binom{\mathrm{o}}{\mathbf{B}} \tag{1.2.30}
\end{equation*}
$$

Note that $\stackrel{\circ}{\mathbf{C}}$ is positive definite if and only if J and $\mathbf{C}$ are positive definite at all quadrature points. If a material loses ellipticity, as for example in strain softening or non-associative plastic materials, Eq. (30) no longer holds. Similarly, if the element is so distorted that J < 0 , the above may not hold.

Assuming an element domain of unit thickness, the nodal forces are obtained directly from stress field by

$$
\mathbf{f}_{\mathrm{e}}^{\mathrm{int}}=\left\{\begin{array}{l}
\mathbf{f}_{\mathrm{x}}^{\text {int }}  \tag{1.2.31a}\\
\mathbf{f}_{\mathrm{y}}^{\text {int }}
\end{array}\right\}=\int_{\Omega_{\mathrm{e}}} \mathbf{B}^{\mathrm{t}} \mathbf{s} \mathrm{~d} \Omega=\int_{-1}^{+1} \int_{-1}^{+1} \mathbf{B}^{\mathrm{t}} \mathbf{s} \mathrm{~J} \mathrm{~d} \xi \mathrm{~d} \eta
$$

where the stress is written as

$$
\mathbf{s}=\left\{\begin{array}{c}
\sigma_{\mathrm{x}} \\
\sigma_{\mathrm{y}} \\
\sigma_{\mathrm{xy}}
\end{array}\right\}
$$

and

$$
\begin{aligned}
& \left(\mathbf{f}_{\mathrm{x}}^{\mathrm{int}}\right)^{\mathrm{t}}=\left[\mathrm{f}_{\mathrm{x} 1}, \mathrm{f}_{\mathrm{x} 2}, \mathrm{f}_{\mathrm{x} 2}, \mathrm{f}_{\mathrm{x} 4}\right]^{\mathrm{int}} \\
& \left(\mathbf{f}_{\mathrm{y}}^{\mathrm{int}}\right)^{\mathrm{t}}=\left[\mathrm{f}_{\mathrm{y} 1}, \mathrm{f}_{\mathrm{y} 2}, \mathrm{f}_{\mathrm{y} 2}, \mathrm{f}_{\mathrm{y} 4}\right]^{\mathrm{int}}
\end{aligned}
$$

Applying numerical quadrature, this becomes

$$
\begin{align*}
& \mathbf{f}_{\mathrm{e}}^{\mathrm{int}}=\sum_{\alpha=1}^{\mathrm{n}_{\mathrm{Q}}} \mathrm{w}_{\alpha} \mathrm{J}\left(\mathbf{x}_{\alpha}\right) \mathbf{B}^{\mathrm{t}}\left(\mathbf{x}_{\alpha}\right) \mathbf{s}\left(\mathbf{x}_{\alpha}\right)  \tag{1.2.31b}\\
& =\stackrel{\mathbf{B}}{ }_{\mathrm{o}}{ }^{\mathrm{t}} \mathbf{s} \tag{1.2.31c}
\end{align*}
$$

where

$$
\stackrel{\stackrel{o}{\mathbf{s}^{\mathrm{t}}}}{ }=\left[\begin{array}{llll}
\mathrm{w}_{1} \mathrm{~J}\left(\mathbf{x}_{1}\right) \mathbf{s}\left(\mathbf{x}_{1}\right), & \mathrm{w}_{2} \mathrm{~J}\left(\mathbf{x}_{2}\right) \mathbf{s}\left(\mathbf{x}_{2}\right), & \cdots & \mathrm{w}_{\mathrm{n}_{\mathrm{Q}}} \mathrm{~J}\left(\mathbf{x}_{\mathrm{n}_{\mathrm{Q}}}\right) \mathbf{s}\left(\mathbf{x}_{\mathrm{n}_{\mathrm{Q}}}\right) \tag{1.2.31d}
\end{array}\right]
$$

The rank of $\stackrel{\mathbf{B}}{\mathbf{B}}$ can be estimated by the following

$$
\begin{equation*}
\operatorname{rank} \stackrel{\circ}{\mathbf{B}} \leq \min (\text { rows in } \stackrel{\circ}{\mathbf{B}}, \operatorname{dim} \mathcal{D} \mathbf{u}) \tag{1.2.32}
\end{equation*}
$$

where $\mathcal{D}$ is the symmetric gradient operator given by

$$
\mathcal{D}=\left[\begin{array}{cc}
\frac{\partial}{\partial x} & 0 \\
0 & \frac{\partial}{\partial y} \\
\frac{\partial}{\partial y} & \frac{\partial}{\partial x}
\end{array}\right]
$$

and the dimension of $\mathcal{D} \mathbf{u}$ (or $\operatorname{dim} \mathcal{D} \mathbf{u}$ ) is equal to the number of independent functions in $\mathcal{D} \mathbf{u}$. In most cases, the above is an equality, but it is possible, even with regular quadrature schemes and undistorted elements, to lose the equality, i.e. to have linearly dependent rows in the $\mathbf{B}$ matrix.

The rank-sufficiency of QUAD4 will now be examined for various quadrature schemes. The element has 4 nodes with 2 degrees of freedom at each node, so $\operatorname{dim}\left(\mathbf{K}_{\mathrm{e}}\right)=$ 8. The number of rigid body modes is 3: translation in the x and y directions and rotation in the ( $\mathrm{x}, \mathrm{y}$ ) plane. By Eq. (26a), the proper rank of $\mathbf{K}_{\mathrm{e}}=5$.

The most widely used quadrature scheme is $2 \times 2$ Gauss quadrature. The number of quadrature points $\mathrm{n}_{\mathrm{Q}}=4$, the number of rows in each $\mathbf{B}\left(\mathbf{x}_{\alpha}\right)=3$, so the number of rows in $\stackrel{\mathrm{O}}{\mathbf{B}}=12$ This exceeds the proper rank. However, based on the linear completeness of the quadrilateral, it will be shown later that (see Section 1.4)

$$
\begin{align*}
& u_{x}(x, y)=\alpha_{o x}+\alpha_{1 x} x+\alpha_{2 x} y+\alpha_{3 x} h  \tag{1.2.33a}\\
& u_{y}(x, y)=\alpha_{o y}+\alpha_{1 y} x+\alpha_{2 y} y+\alpha_{3 y} h  \tag{1.2.33b}\\
& h \equiv \xi \eta
\end{align*}
$$

Then

$$
\stackrel{o}{\mathbf{e}}=\mathcal{D} \mathbf{u}=\left[\begin{array}{c}
\alpha_{1 x}+\alpha_{3 x} \frac{\partial h}{\partial x}  \tag{1.2.34}\\
\alpha_{2 y}+\alpha_{3 y} \frac{\partial h}{\partial y} \\
\alpha_{2 x}+\alpha_{1 y}+\alpha_{3 x} \frac{\partial h}{\partial y}+\alpha_{3 y} \frac{\partial h}{\partial x}
\end{array}\right]
$$

Examination of the above shows that the strain-rate field contains 5 linearly independent functions: $\alpha_{1 \mathrm{x}}, \alpha_{2 \mathrm{y}}, \alpha_{3 \mathrm{x}} \partial \mathrm{h} / \partial \mathrm{x}, \alpha_{3 \mathrm{y}} \partial \mathrm{h} / \partial \mathrm{y}$, and $\alpha_{2 \mathrm{x}}+\alpha_{1 \mathrm{y}}$. Note that the two constants in the shear strain field must be considered as a single independent field and the function $\alpha_{3 x} \partial h / \partial y+\alpha_{3 y} \partial h / \partial x$ cannot be considered linearly independent because it is a combination of the two functions that have already been included in the list. Thus $\operatorname{dim} \mathcal{D} \mathbf{u}=5$ and since rows in $\stackrel{\mathbf{B}}{\mathbf{B}}=12$, it follows from (32) that

$$
\operatorname{rank}(\stackrel{\mathrm{B}}{\mathbf{B}})=5
$$

It may be concluded that for any quadrature scheme the rank of $\stackrel{\circ}{\mathbf{B}}$ for QUAD4 cannot exceed 5 .

The rank of the element stiffness of QUAD4 for one-point quadrature can be ascertained similarly. In one-point quadrature,,$\stackrel{\circ}{\mathbf{B}}$ consists of $\mathbf{B}$ evaluated at a single point, so its rank is 3 . Therefore rank $\mathbf{K}_{\mathrm{e}}$ is 3 by Eq. (30), and Eq. (26b) indicates that the element has a rank deficiency of 2. This rank deficiency can cause serious difficulties unless it is corrected. Such corrective procedures are described later.
1.2.4. Nodal Forces and B-Matrix for One-Point Quadrature Element. Prior to describing procedures for correcting the rank deficiency of QUAD4 with one point quadrature, it is worthwhile to develop the one-point quadrature formulas in detail. These formulas will then provide the framework for the development of the rank correction procedures, which in QUAD4 are often called hourglass control.

The internal nodal forces are given by (31b). When one-point quadrature is used, the quadrature point is selected to be the origin of the coordinate system in the reference plane: $\xi=\eta=0$. Evaluating the Jacobian at this point yields (see Eq. (9)) gives

$$
\begin{equation*}
J(0)=\frac{1}{8}\left(x_{31} y_{42}+x_{24} y_{31}\right)=\frac{A}{4} \tag{1.2.35}
\end{equation*}
$$

where A is the area of the element. The expression for the internal nodal forces now becomes

$$
\begin{equation*}
\mathbf{f}^{\text {int }}=4 \mathbf{B}^{\mathrm{t}}(\mathbf{0}) \mathbf{s}(\mathbf{0}) \mathrm{J}(\mathbf{0})=\mathrm{A} \mathbf{B}^{\mathrm{t}}(\mathbf{0}) \mathbf{s}(\mathbf{0}) \tag{1.2.36}
\end{equation*}
$$

Evaluating the $\mathbf{B}$ matrix from Eq. (10) at $\xi=\eta=0$ is a simple algebraic process which gives

$$
\mathbf{B}(\mathbf{0})=\left[\begin{array}{cc}
\mathbf{b}_{\mathrm{x}}^{\mathrm{t}} & 0  \tag{1.2.37}\\
0 & \mathbf{b}_{\mathrm{y}}^{\mathrm{t}} \\
\mathbf{b}_{\mathrm{y}}^{\mathrm{t}} & \mathbf{b}_{\mathrm{x}}^{\mathrm{t}}
\end{array}\right]
$$

where

$$
\begin{align*}
\mathbf{b}_{\mathrm{x}}^{\mathrm{t}} & \equiv \mathbf{b}_{1}^{\mathrm{t}} \tag{1.2.38}
\end{align*}=\frac{1}{2 \mathrm{~A}}\left[\mathrm{y}_{24}, \mathrm{y}_{31}, \mathrm{y}_{42}, \mathrm{y}_{13}\right],
$$

Since one-point quadrature is used, the nodal forces are simply the product of the area and the integrand evaluated at $\xi=\eta=0$, which using (37) and (39) gives

$$
\mathbf{f}^{\text {int }}=\mathrm{A}\left[\begin{array}{ccc}
\mathbf{b}_{\mathrm{x}} & \mathbf{0} & \mathbf{b}_{\mathrm{y}}  \tag{1.2.39}\\
\mathbf{0} & \mathbf{b}_{\mathrm{y}} & \mathbf{b}_{\mathrm{x}}
\end{array}\right]\left\{\begin{array}{c}
\sigma_{\mathrm{x}} \\
\sigma_{\mathrm{y}} \\
\sigma_{\mathrm{xy}}
\end{array}\right\}=\mathrm{A} \mathbf{B}^{\mathrm{t}}(\mathbf{0}) \mathbf{s}(\mathbf{0})
$$

The element stiffness matrix for the underintegrated element can be obtained by using the stress-strain law

$$
\begin{equation*}
\mathbf{s}=\mathbf{C e} \tag{1.2.40}
\end{equation*}
$$

in conjunction with Eqs. (37), (39) and $\mathbf{e}=\mathbf{B d}$. This gives

$$
\begin{equation*}
\mathbf{f}^{\mathrm{int}}=\mathbf{K}_{\mathrm{e}} \mathbf{d} \tag{1.2.41}
\end{equation*}
$$

where

$$
\begin{equation*}
\mathbf{K}_{\mathrm{e}}=\mathrm{A} \mathbf{B}^{\mathrm{t}}(\mathbf{0}) \mathbf{C B}(\mathbf{0}) \tag{1.2.42}
\end{equation*}
$$

The element stiffness matrix could also be obtained from (28) by using one-point quadrature and the values of $\mathbf{B}$ and $\mathbf{C}$ at the quadrature point.
1.2.5 Spurious Singular Modes (Hourglass) The presence and shape of the spurious singular modes of the one-point quadrature QUAD4 element will now be demonstrated. Any nodal displacement $\mathbf{d}^{\mathrm{H}}$ that is not a rigid body motion but results in no straining of the element is a spurious singular mode. From (43) it can be seen that such nodal displacements will not generate any nodal forces, i.e. they will not be resisted by the element, since in the absence of strains, the stresses will also vanish, so $\mathbf{f}^{\mathrm{int}}=0$.

Consider the nodal displacements

$$
\begin{align*}
& \mathbf{d}^{\mathrm{Hx}}=\left\{\begin{array}{l}
\mathbf{h} \\
\mathbf{0}
\end{array}\right\} \quad \mathbf{d}^{\mathrm{Hy}}=\left\{\begin{array}{l}
\mathbf{0} \\
\mathbf{h}
\end{array}\right\}  \tag{1.2.43}\\
& \mathbf{h}^{\mathrm{t}}=[+1,-1,+1,-1-1
\end{align*}
$$

It can easily be verified that

$$
\begin{equation*}
\mathbf{b}_{\mathrm{x}}^{\mathrm{t}} \mathbf{h}=0 \quad \mathbf{b}_{\mathrm{y}}^{\mathrm{t}} \mathbf{h}=0 \tag{1.2.44}
\end{equation*}
$$

Therefore, it follows from Eqs. (37), (43) and (44) that

$$
\begin{align*}
& \mathbf{B}(\mathbf{0}) \mathbf{d}^{\mathrm{Hx}}=\mathbf{0}  \tag{1.2.45a}\\
& \mathbf{B}(\mathbf{0}) \mathbf{d}^{\mathrm{Hy}}=\mathbf{0} \tag{1.2.45b}
\end{align*}
$$

Fig. 3 shows rectangular elements with the spurious singular modes of deformation $\mathbf{d}^{\mathrm{Hx}}$ and $\mathbf{d}^{\mathrm{Hy}}$, and also both modes simultaneously. In the rectangle, it can be seen that the hourglass modes are associated with the bilinear term in the displacement field. The deformed configuration of the mesh in the spurious singular modes is shown in Fig. 4. A vertical pair of elements in this x-mode looks like an hourglass, an ancient device for measuring time by the flow of sand from the top element to the bottom. For this reason, this spurious singular mode is often called hourglassing or the hourglass mode. Because each element is in an hourglass mode, the entire mesh can deform as shown without any resisting forces from the element.


Figure 3. Hourglass modes of deformation
The problem of hourglassing first appeared in finite difference hydrodynamics programs in which the derivatives were evaluated by transforming them to contour integrals by means of the divergence theorem; see for example Wilkins and Blum (1975). This procedure tacitly assumed that the derivatives are constant in the domain enclosed by each contour. This assumption is equivalent to the constant strain (and stress) assumption which is associated with one-point quadrature. The equivalence of these contour-integral finite difference methods was demonstrated by Belytschko et al. (1975); also see Belytschko (1983). Many ad hoc procedures for hourglass control were developed by finite difference workers. The procedures focused on controlling the relative rotations of element sides; no consideration was given to maintaining consistency.


Figure 4. Mesh in hourglass mode of deformation

This phenomenon occurs in many other settings, so a variety of names have evolved. For example, they occur frequently in mixed or hybrid elements, where they are called zero-energy modes or spurious zero-energy modes. Hourglass modes are zero-energy modes, since they don't result in any strain at the points in the element which are sampled. Therefore they do no work and $\left(\mathbf{d}^{\mathrm{Hx}}\right)^{\text {trint }}=\left(\mathbf{d}^{\mathrm{Hy}}\right)^{\text {t }} \mathbf{f}^{\text {int }}=0$

In structural analysis, spurious singular modes arise when there is insufficient redundancy, i.e. the number of structural members is insufficient to preclude rigid body motion of part of the structure. Such modes often occur in three dimensional truss structures. In structural analysis, they are called kinematic modes, and because of the close relationship between the structural analysis and finite element communities, this name has also been applied to spurious singular modes. Other names which have been applied to this phenomenon are: keystoning (Key et al. (1978)), chickenwiring, and mesh instability.

For finite element discretizations of partial differential equations, spurious singular modes appears to be the most accurate term for this phenomenon, so we shall use that name. For example, the terms kinematic modes or zero-energy modes are not appropriate for the Laplace equation. In elements where the spurious singular mode has a distinctive appearance, such as the hourglass pattern in QUAD4, we shall also use that name. The condition which leads to spurious singular modes is rank deficiency of the element stiffness matrix.

When rank deficient elements are assembled, the system stiffness will often be singular or nearly singular. Therefore, in matrix methods, the presence of spurious singular modes can be detected by the presence of zero or very small pivots in the total stiffness. If the pivots are zero, the stiffness will be singular and not invertible. If the pivots are very small, the total stiffness is near-singular, and the displacement solutions will be oscillatory in space, i.e., they will exhibit the hourglass mode.

Because a system stiffness matrix is never assembled in explicit methods, the singularity cannot be readily detected. In iterative solvers, the presence of spurious singular modes will often lead to divergence of the solution. With explicit integrators, singular modes are not readily detectable without plots of the deformed configuration. This is also true for matrix dynamic methods, since the mass matrix then renders the system matrix nonsingular even when the stiffness matrix is singular.

The evolution of an hourglass mode in a transient problem is shown in Fig. 5. In this problem, the beam was supported at a single node to facilitate the appearance of the hourglass mode. If all nodes at the left-hand end of the beam were fixed to simulate a clamped support condition, the hourglass mode would not appear. Althoughrank-deficient elements may sometimes appear to work, they should not be used without an appropriate correction.


Figure 5. Four views of a simply supported beam showing the evolution of the hourglass mode (due to symmetry, only half the beam was modeled)

### 1.3 Perturbation Hourglass Stabilization

The simplest way to control spurious singular modes without impairing convergence is to augment the rank of the element stiffness without disturbing the linear completeness (consistency) of the isoparametric element. One approach to this task is to augment the $\mathbf{B}(\mathbf{0})$ matrix of the one-point quadrature element by two rows which are linearly independent of the other three. These additional two rows consist of a $\mathbf{g}$ vector that will be derived subsequently. Adding these two rows corresponds to adding 2 generalized strains. The matrices for the one-point quadrature QUAD4 are then

$$
\begin{align*}
& \widetilde{\mathbf{B}}=\left[\begin{array}{cc}
\mathbf{b}_{\mathrm{x}}^{\mathrm{T}} & 0 \\
0 & \mathbf{b}_{\mathrm{y}}^{\mathrm{T}} \\
\mathbf{b}_{\mathrm{y}}^{\mathrm{T}} & \mathbf{b}_{\mathrm{x}}^{\mathrm{T}} \\
\mathbf{g}^{\mathrm{T}} & 0 \\
0 & \mathbf{g}^{\mathrm{T}}
\end{array}\right] \quad \widetilde{\mathbf{C}}=\left[\begin{array}{ccccc}
\mathrm{C}_{11} & \mathrm{C}_{12} & \mathrm{C}_{13} & 0 & 0 \\
\mathrm{C}_{11} & \mathrm{C}_{22} & \mathrm{C}_{23} & 0 & 0 \\
\mathrm{C}_{13} & \mathrm{C}_{23} & \mathrm{C}_{33} & 0 & 0 \\
0 & 0 & 0 & \mathrm{C}^{\mathrm{Q}} & 0 \\
0 & 0 & 0 & 0 & \mathrm{C}^{\mathrm{Q}}
\end{array}\right]  \tag{1.3.1a}\\
& \widetilde{\mathbf{s}}^{\mathrm{T}}=\left[\sigma_{\mathrm{x}}, \sigma_{\mathrm{y}}, \sigma_{\mathrm{xy}}, \mathrm{Q}_{\mathrm{x}}, \mathrm{Q}_{\mathrm{y}_{-}}\right.  \tag{1.3.1b}\\
& \widetilde{\mathbf{e}}^{\mathrm{T}}=\left[\varepsilon_{\mathrm{x}}, \varepsilon_{\mathrm{y}}, 2 \varepsilon_{\mathrm{xy}}, \mathrm{q}_{\mathrm{x}}, \mathrm{q}_{\mathrm{y}}^{-}\right. \tag{1.3.1c}
\end{align*}
$$

where $\widetilde{\mathbf{C}}$ is the constitutive matrix augmented by two rows and columns in terms of a constant to be determined $\left(C^{\mathrm{Q}}\right) ; \widetilde{\mathbf{s}}$ and $\widetilde{\mathbf{e}}$ are the stress and strain matrices augmented by the generalized stresses and strains ( $\mathrm{Q}_{\mathrm{x}}, \mathrm{Q}_{\mathrm{y}}$ ) and ( $\mathrm{q}_{\mathrm{x}}, \mathrm{q}_{\mathrm{y}}$ ), respectively.

To maintain linear consistency for the element, these additional generalized strains should vanish when the nodal displacements (or velocities) emanate from linear fields. Consistency and stability (rank sufficiency) are essential requirements for a sound numerical method.

The requirement that $q_{x}=q_{y}=0$ for linear fields implies

$$
\begin{equation*}
\mathbf{g}^{\mathrm{T}} \mathbf{u}^{\mathrm{Lin}}=\mathbf{g}^{\mathrm{T}}\left(\alpha_{0} \mathbf{s}+\alpha_{1} \mathbf{x}+\alpha_{2} \mathbf{y}\right)=0 \quad \forall \alpha_{\mathrm{i}} \tag{1.3.2}
\end{equation*}
$$

The above must be satisfied for both $\mathbf{u}_{\mathrm{x}}$ and $\mathbf{u}_{\mathrm{y}}$ so we have not specified the component; $\mathbf{u}^{\text {Lin }}$ is taken from Eq. (1.2.15b). The above can be interpreted as an orthogonality condition: $\mathbf{g}$ must be orthogonal to all linear fields.
1.3.1 The Gamma Vector. We first define a set of four vectors, $\mathbf{b}^{*}$.

$$
\begin{equation*}
\mathbf{b}^{*} \equiv\left(\mathbf{b}_{\mathrm{x}}, \mathbf{b}_{\mathrm{y}}, \mathbf{s}, \mathbf{h}\right) \tag{1.3.3}
\end{equation*}
$$

To obtain $\mathbf{g}$, two properties of $\mathbf{b}^{*}$ are used:

1. the vectors $\mathbf{b}_{i}$ are biorthogonal to $\mathbf{x}_{j}$
2. the vectors $\mathbf{b}_{i}^{*}$ are linearly independent.

The biorthogonality property, given by

$$
\begin{equation*}
\mathbf{b}_{\mathrm{i}}^{\mathrm{t}} \mathbf{x}_{\mathrm{j}}=\delta_{\mathrm{ij}} \quad(\mathrm{i}, \mathrm{j})=1 \text { to } 2 \tag{1.3.4}
\end{equation*}
$$

is an identity which holds for all isoparametric elements:

$$
\begin{equation*}
\frac{\partial \mathbf{N}}{\partial \mathrm{x}_{\mathrm{i}}} \mathbf{x}_{\mathrm{j}}=\delta_{\mathrm{ij}} \tag{1.3.5}
\end{equation*}
$$

The demonstration of this identity is based on the isoparametric mapping, Eq. (1.2.3c), which when combined with (5) gives

$$
\begin{equation*}
\frac{\partial \mathbf{N}}{\partial \mathrm{x}_{\mathrm{i}}} \mathbf{x}_{\mathrm{j}}=\frac{\partial \mathrm{x}_{\mathrm{j}}}{\partial \mathrm{x}_{\mathrm{i}}}=\delta_{\mathrm{ij}} \tag{1.3.6}
\end{equation*}
$$

where the last equality expresses the fact that in two dimensions, for example, $\partial x / \partial x=\partial y / \partial y=1, \partial x / \partial y=\partial y / \partial x=0$. Eq. (5) holds for the derivatives of the shape functions at any point. In particular, it also holds for the point $\xi=\eta=0$ in QUAD4. Additional orthogonality conditions,

$$
\begin{equation*}
\mathbf{b}_{\mathrm{i}}^{\mathrm{t}} \mathbf{s}=0 \quad \mathbf{b}_{\mathrm{i}}^{\mathrm{t}} \mathbf{h}=0 \quad \mathbf{h}^{\mathrm{t}} \mathbf{s}=0 \quad \mathrm{i}=1 \text { to } 2 \tag{1.3.7}
\end{equation*}
$$

can easily be verified by arithmetic using the definitions of these vectors.
The linear independence of the $4 \mathbf{b}_{i}^{*}$ vectors is demonstrated as follows. Assume $\mathbf{b}_{i}^{*}$ are linearly dependent. Then it follows that there exists $\alpha_{i} \neq 0$ such that

$$
\begin{equation*}
\alpha_{1} \mathbf{b}_{x}+\alpha_{2} \mathbf{b}_{y}+\alpha_{3} s+\alpha_{4} \mathbf{h}=0 \tag{1.3.8}
\end{equation*}
$$

Premultiplying the above by $\mathbf{s}^{\mathrm{t}}$, and using (7) yields $\alpha_{3}=0$. Similarly premultiplying by $\mathbf{h}^{\mathrm{t}}$ yields $\alpha_{4}=0$. Then premultiplying $\mathbf{x}^{\mathrm{t}}$ yields

$$
\begin{equation*}
\alpha_{1}+\alpha_{3} \mathbf{x}^{\mathrm{t}} \mathbf{s}+\alpha_{4} \mathbf{x}^{\mathrm{t}} \mathbf{h}=0 \tag{1.3.9}
\end{equation*}
$$

and since it has just been determined that $\alpha_{3}=\alpha_{4}=0$, it follows that $\alpha_{1}=0$. Similarly, premultiplying by $\mathbf{y}^{\mathrm{t}}$ shows that $\alpha_{2}=0$. Thus $\alpha_{i}=0$, for $\mathrm{i}=1$ to 4 , and it follows that the vectors $\mathbf{b}_{i}^{*}$ are linearly independent.

The preceding developments are now used as tools for the construction of $\mathbf{g}$, via the consistency requirement (2). Since the vectors $\mathbf{b}_{i}^{*}$ are linearly independent, they span $R^{4}$. It follows that any vector in $R^{4}$, including $\mathbf{g}$ can be expressed as a linear combination of $\mathbf{b}_{\mathrm{i}}^{*}$ :

$$
\begin{equation*}
\mathbf{g}=\beta_{1} \mathbf{b}_{x}+\beta_{2} \mathbf{b}_{y}+\beta_{3} \mathbf{h}+\beta_{4} \mathbf{s} \tag{1.3.10}
\end{equation*}
$$

where $\beta_{i}$ are constants to be determined by the linear consistency requirement (2). Substituting (10) into (2) and collecting the coefficients of $\alpha_{i}$ yields

$$
\begin{align*}
& \alpha_{0}\left(\beta_{1} \mathbf{b}_{\mathrm{x}}^{\mathrm{t}} \mathbf{s}+\beta_{2} \mathbf{b}_{\mathrm{y}}^{\mathrm{t}} \mathbf{s}+\beta_{3} \mathbf{h}^{\mathrm{t}} \mathbf{s}+\beta_{4} \mathbf{s}^{\mathbf{t}} \mathbf{s}\right) \\
& +\alpha_{1}\left(\beta_{1} \mathbf{b}_{\mathrm{x}}^{\mathrm{t}} \mathbf{x}+\beta_{2} \mathbf{b}_{\mathrm{y}}^{\mathrm{t}} \mathbf{x}+\beta_{3} \mathbf{h}^{\mathrm{t}} \mathbf{x}+\beta_{4} \mathbf{s}^{\mathrm{t}} \mathbf{x}\right)  \tag{1.3.11}\\
& +\alpha_{2}\left(\beta_{1} \mathbf{b}_{\mathrm{x}}^{\mathrm{t}} \mathbf{y}+\beta_{2} \mathbf{b}_{\mathrm{y}}^{\mathrm{t}} \mathbf{y}+\beta_{3} \mathbf{h}^{\mathrm{t}} \mathbf{y}+\beta_{4} \mathbf{s}^{\mathrm{t}} \mathbf{y}\right)=0
\end{align*}
$$

Since the above must vanish for all $\alpha_{i}$, each coefficient of $\alpha_{i}$ must vanish. Taking the coefficient of $\alpha_{0}$ and simplifying by means of Eqs. (4) and (7) gives

$$
\begin{equation*}
\beta_{4} \mathbf{s}^{\mathbf{t}} \mathbf{s}=4 \beta_{4}=0 \tag{1.3.12}
\end{equation*}
$$

Using (12) and (6) to simplify the coefficient of $\alpha_{1}$ in (11) gives

$$
\begin{equation*}
\beta_{1}+\beta_{3} \mathbf{h}^{t} \mathbf{x}=0 \tag{1.3.13}
\end{equation*}
$$

and a similar procedure for the coefficient of $\alpha_{2}$ gives

$$
\begin{equation*}
\beta_{2}+\beta_{3} \mathbf{h}^{\mathrm{t}} \mathbf{y}=0 \tag{1.3.14}
\end{equation*}
$$

Using Eqs. (13) and (14) in (10) to express $\beta_{1}$ and $\beta_{2}$ in terms of $\beta_{3}$ and using (12) yields

$$
\begin{equation*}
\mathbf{g}=\beta_{3}\left[\mathbf{h}-\left(\mathbf{h}^{t} \mathbf{x}\right) \mathbf{b}_{\mathbf{x}}-\left(\mathbf{h}^{\mathbf{t}} \mathbf{y}\right) \mathbf{b}_{\mathbf{y}}\right] \tag{1.3.15}
\end{equation*}
$$

The constant $\beta_{3}$ remains undetermined, since for any value of $\beta_{3}$ the vector $\mathbf{g}$ is orthogonal to all linear fields. It will be convenient later to have $\mathbf{g}^{t} \mathbf{h}=1$, so we set $\beta_{3}=1 / 4$. The value of $\beta_{3}=1$ was used in Flanagan and Belytschko (1981) because the reference element was a unit square; this changes some of the subsequent constants but not the substance of the development. In this description we choose $\beta_{3}=1 / 4$, which gives

$$
\begin{equation*}
\mathbf{g}=\frac{1}{4}\left[\mathbf{h}-\left(\mathbf{h}^{\mathrm{t}} \mathbf{x}\right) \mathbf{b}_{\mathrm{x}}-\left(\mathbf{h}^{\mathrm{t}} \mathbf{y}\right) \mathbf{b}_{\mathrm{y}}\right] \tag{1.3.16a}
\end{equation*}
$$

The above expression can be written in indicial notation as

$$
\begin{equation*}
\mathbf{g}=\frac{1}{4}\left[\mathbf{h}-\left(\mathbf{h}^{\mathrm{t}} \mathbf{x}_{\mathrm{i}}\right) \mathbf{b}_{\mathrm{i}}\right] \tag{1.3.16b}
\end{equation*}
$$

or

$$
\begin{equation*}
\gamma_{\mathrm{I}}=\frac{1}{4}\left[\mathrm{~h}_{\mathrm{I}}-\left(\mathrm{h}_{\mathrm{J}} \mathrm{x}_{\mathrm{iJ}}\right) \mathrm{b}_{\mathrm{iI}}\right] \tag{1.3.16c}
\end{equation*}
$$

Using (4), (7), and (16a) the following are easily verified by:

$$
\begin{equation*}
\mathbf{g}^{\mathbf{t}} \mathbf{x}=\mathbf{g}^{\mathbf{t}} \mathbf{y}=\mathbf{g}^{\mathbf{t}} \mathbf{s}=0 \quad \mathbf{g}^{\mathbf{t}} \mathbf{h}=1 \tag{1.3.17}
\end{equation*}
$$

1.3.2 Stabilization Forces and Stiffness Matrix. Replacing $\mathbf{B}(\mathbf{0})$ and $\mathbf{s}(\mathbf{0})$ in Eq. (1.2.39) by the augmented matrices $\widetilde{\mathbf{B}}$ and $\widetilde{\mathbf{s}}$ gives the nodal forces

$$
\begin{align*}
& \mathbf{f}^{\mathrm{int}}=\mathrm{A}\left[\begin{array}{ccc}
\mathbf{b}_{\mathrm{x}} & \mathbf{0} & \mathbf{b}_{\mathrm{y}} \\
\mathbf{0} & \mathbf{b}_{\mathrm{y}} & \mathbf{b}_{\mathrm{x}}
\end{array}\right]\left\{\begin{array}{c}
\sigma_{\mathrm{x}} \\
\sigma_{\mathrm{y}} \\
\sigma_{\mathrm{xy}}
\end{array}\right\}+\mathrm{A}\left\{\begin{array}{l}
\mathrm{Q}_{\mathrm{x}} \mathbf{g} \\
\mathrm{Q}_{\mathrm{y}} \mathbf{g}
\end{array}\right\}  \tag{1.3.18}\\
& =\mathrm{AB}(\mathbf{0}) \underline{\mathbf{s}(\mathbf{0})+\mathbf{f}^{\text {stab }}} \tag{1.3.19}
\end{align*}
$$

The first term in the internal force is obtained by one-point quadrature. The generalized stresses and are strains are obtained by the stress-strain law, $\widetilde{\mathbf{s}}=\widetilde{\mathbf{C}} \widetilde{\mathbf{e}}$, and the straindisplacement equation $\widetilde{\mathbf{e}}=\widetilde{\mathbf{B}} \mathbf{d}$ :

$$
\begin{align*}
& \mathrm{Q}_{\mathrm{x}}=\mathrm{C}^{\mathrm{Q}} \mathrm{q}_{\mathrm{x}} \quad \mathrm{Q}_{\mathrm{y}}=\mathrm{C}^{\mathrm{Q}_{\mathrm{y}}}  \tag{1.3.20}\\
& \mathrm{q}_{\mathrm{x}}=\mathbf{g}^{\mathrm{t}} \mathbf{u}_{\mathrm{x}} \quad \mathrm{q}_{\mathrm{y}}=\mathbf{g}^{\mathrm{t}} \mathbf{u}_{\mathrm{y}}
\end{align*}
$$

The stiffness matrix is obtained by substituting replacing $\mathbf{B}(\mathbf{0})$ and $\mathbf{C}$ in (1.2.42) by the augmented matrices $\widetilde{\mathbf{B}}$ and $\widetilde{\mathbf{C}}$ which gives

$$
\begin{equation*}
\mathbf{K}_{\mathrm{e}}=\mathbf{K}_{\mathrm{e}}^{1 \mathrm{pt}}+\mathbf{K}_{\mathrm{e}}^{\mathrm{stab}} \tag{1.3.21a}
\end{equation*}
$$

where

$$
\begin{align*}
& \mathbf{K}_{\mathrm{e}}^{1 \mathrm{pt}}=\mathrm{AB}  \tag{1.3.21b}\\
& \mathbf{K}_{\mathrm{e}}^{\mathrm{t}}(\mathbf{0}) \mathbf{C B}(\mathbf{0})  \tag{1.3.21c}\\
& \mathrm{stab}^{\mathrm{stab}}\left[\begin{array}{cc}
\mathbf{g g} & \mathbf{0} \\
\mathbf{0} & \mathbf{g} \mathbf{g}^{\mathrm{t}}
\end{array}\right]
\end{align*}
$$

$\mathbf{K}_{\mathrm{e}}^{1 \mathrm{pt}}$ is the stiffness matrix obtained from one-point quadrature. $\mathbf{K}_{\mathrm{e}}^{\text {stab }}$ is obtained from the additional generalized strains which were introduced to stabilize the element and is often called a stabilization matrix. The stabilization matrix is of rank 2 . Combined with the onepoint quadrature stiffness, it yields a matrix of rank 5, which is the correct rank for the QUAD4.

This form of the linearly consistent generalized strains occurs in many stabilization procedures for underintegrated elements, and will be designated as the $\mathbf{g}$ vector. Note that the vector is not completely determined by linear consistency: an unspecified constant $\beta_{3}$ remains. This vector is orthogonal to the nodal displacements which emanate from a linear
field for arbitrarily shaped quadrilaterals. When the element is rectangular, $\mathbf{h}^{\mathrm{t}} \mathbf{x}=\mathbf{h}^{\mathrm{t}} \mathbf{y}=0$ and $\mathbf{g}=\beta_{3} \mathbf{h}$.
1.3.3 Scaling the Stabilization Forces. Since the constants $C^{Q}$ in Eq. (1) are not true material constants, it is important to provide formulas for these constants which provide approximately the same degree of stabilization regardless of the geometry and material properties of the element. The basic objective is to obtain a scaling which perturbs the element sufficiently to insure the correct rank but not to overwhelm the one-point quadrature stiffness.

One procedure for selecting CQ is to scale the maximum eigenvalue of the stabilization stiffness to the maximum eigenvalue of the underintegrated stiffness. In fact, it would be desirable to shift eigenvalues associated with hourglass modes out of the spectrum that is of interest in the response. The hourglass modes in a fully integrated element are smaller than the 1-point quadrature element eigenvalues. To avoid locking, the stabilization should be a small fraction of the one point quadrature eigenvalue.

According to Flanagan and Belytschko (1981), the maximum eigenvalue for the 1point quadrature version of QUAD4 for an isotropic material is bounded by

$$
\begin{align*}
& \frac{1}{2} \mathrm{Ac}^{2} \mathrm{~b} \leq \lambda_{\max } \leq \mathrm{Ac}^{2} \mathrm{~b}  \tag{1.3.22a}\\
& \mathrm{~b}=\sum_{\mathrm{i}=1}^{2} \mathbf{b}_{\mathrm{i}}^{\mathrm{T}} \mathbf{b}_{\mathrm{i}} \quad \mathrm{c}^{2}=\lambda+2 \mu \tag{1.3.22b}
\end{align*}
$$

The eigenvalues of $\mathbf{K}_{\mathrm{e}}$ are given by the eigenvalue problem

$$
\begin{equation*}
\mathbf{K x}=\lambda \mathbf{x} \tag{1.3.23}
\end{equation*}
$$

The eigenvalue associated with the stabilization can be estimated by letting $\mathbf{x}=\mathbf{d}^{\mathrm{Hx}}$ in the Rayleigh quotient, which with (20) and the orthogonality properties (4) and (7) gives

$$
\begin{equation*}
\lambda=\frac{\mathbf{x}^{\mathrm{t}} \mathbf{K} \mathbf{x}}{\mathbf{x}^{\mathrm{t}} \mathbf{x}}=\frac{\mathrm{AC}^{\mathrm{Q}^{\mathrm{t}} \mathbf{h}^{\mathrm{g}}} \mathbf{g}^{\mathrm{t}} \mathbf{h}}{\mathbf{h}^{\mathrm{t}} \mathbf{h}} \tag{1.3.24}
\end{equation*}
$$

where the second equality follows because $\mathbf{K}_{\mathrm{e}}^{1 \mathrm{tt}} \mathbf{h}=\mathbf{0}$. Using Eq. (17), it can be seen that

$$
\begin{equation*}
\lambda=\mathrm{AC}^{\mathrm{Q} / 4} \tag{1.3.25}
\end{equation*}
$$

Using Eqs. (22) and (25) it follows that the eigenvalue associated with the stabilization is scaled to the lower bound on the maximum eigenvalue of the element if

$$
\begin{equation*}
C^{\mathrm{Q}}=2 \alpha_{s} c^{2} \mathrm{~b}=2 \alpha_{\mathrm{s}}(\lambda+2 \mu) \sum_{i=1}^{2} \mathbf{b}_{i}^{t} \mathbf{b}_{\mathrm{i}} \tag{1.3.26}
\end{equation*}
$$

where $\alpha_{\mathrm{s}}$ is a scaling parameter.
In Flanagan and Belytschko (1981), the hourglass control parameter was scaled by the dynamic eigenvalue, i.e., the frequency, of the element. However, since the hourglass control is strictly an element-stiffness related stress, this seems inconsistent and a pure stiffness scaling is more appropriate.

### 1.4 Mixed Method Hourglass Stabilization

The mixed variational principles are another vehicle for developing four-node quadrilaterals which do not lock. Furthermore, they can be used to develop rankcompensation procedures for underintegrated elements (stabilization matrices) which do not involve any arbitrary perturbation parameters and are based on the material properties and geometry of the element.
1.4.1 Displacement Field of QUAD4 in Terms of Biorthogonal Basis. Before developing elements based on a mixed method, the displacement field in the element is expressed in a form in which the parts which cause locking can easily be identified and corrected. Furthermore, when this expression for the displacement field is used, the stiffness matrix can be obtained in closed form without any numerical integration. This is useful for understanding its properties and for implementation in vector method programs.

The procedure described here is based on Belytschko and Bachrach (1986). As a preliminary to developing this expression (which will be called the Belytschko-Bachrach form), the basis vectors $\mathbf{b}^{* *}$ and $\mathbf{x}^{*}$ are defined so they are biorthogonal in $\mathrm{R}^{4}$ :

$$
\begin{align*}
& \left(\mathbf{x}_{\beta}^{*}\right)^{t} \mathbf{b}_{\alpha}^{* *}=\delta_{\alpha \beta} \quad(\alpha, \beta)=1 \text { to } 4  \tag{1.4.1}\\
& \mathbf{b}^{* *}=\left(\mathbf{b}_{x}, \mathbf{b}_{\mathbf{y}}, \mathbf{g}, \mathbf{s}^{*}\right)  \tag{1.4.2a}\\
& \mathbf{x}^{*}=(\mathbf{x}, \mathbf{y}, \mathbf{h}, \mathbf{s}) \tag{1.4.2b}
\end{align*}
$$

where

$$
\begin{equation*}
\mathbf{s}^{*}=\frac{1}{4}\left[\mathbf{s}-\left(\mathbf{s}^{\mathrm{t}} \mathbf{x}\right) \mathbf{b}_{\mathrm{x}}-\left(\mathbf{s}^{\mathrm{t}} \mathbf{y}\right) \mathbf{b}_{\mathrm{y}}\right] \tag{1.4.3}
\end{equation*}
$$

The vector $\mathbf{s}^{*}$ is obtained by orthogonalizing $\mathbf{s}$ to $\mathbf{x}$ and $\mathbf{y}$. The arbitrary constant which emerges is chosen to be $1 / 4$ so that $\mathbf{s}^{\boldsymbol{t}} \mathbf{s}^{*}=1$. Most of the identities involved in (1) have already been proven; see Eqs. (1.3.4-7); the remaining ones are easily verified using (3) and (1.3.7).

To develop the Belytschko-Bachrach form, we start by expressing the displacement field as

$$
\begin{equation*}
\mathrm{u}(\mathrm{x}, \mathrm{y})=\alpha_{0}+\alpha_{1} \mathrm{x}+\alpha_{2} \mathrm{y}+\alpha_{3} \xi \eta \tag{1.4.4}
\end{equation*}
$$

Only a single component is considered since the procedure for both components is identical. Evaluating the above at the 4 nodes gives

$$
\begin{equation*}
\mathrm{u}_{\mathrm{I}}=\mathrm{u}\left(\mathrm{x}_{\mathrm{I}}, \mathrm{y}_{\mathrm{I}}\right)=\alpha_{\mathrm{o}}+\alpha_{1} \mathrm{x}_{\mathrm{I}}+\alpha_{2} \mathrm{y}_{\mathrm{I}}+\alpha_{3} \xi_{\mathrm{I}} \eta_{\mathrm{I}} \tag{1.4.5}
\end{equation*}
$$

It is easily shown that $\xi_{\mathrm{I}} \eta_{\mathrm{I}}=\mathrm{h}_{\mathrm{I}}$, so writing the above in matrix form gives

$$
\begin{equation*}
\mathbf{u}=\alpha_{o} \mathbf{s}+\alpha_{1} \mathbf{x}+\alpha_{2} \mathbf{y}+\alpha_{3} \mathbf{h} \tag{1.4.6}
\end{equation*}
$$

which is a linear combination of the linearly independent $\mathbf{x}^{*}$ vectors. Linear independence of $\mathbf{x}^{*}$ follows from the biorthogonality of $\mathbf{x}^{*}$ and $\mathbf{b}^{* *}$.

We now exploit this biorthogonality to evaluate $\alpha_{i}$. Premultiplying (6) by $s^{* t}$ and invoking the orthogonality conditions, $\mathbf{s}^{* t} \mathbf{x}=\mathbf{s}^{* t} \mathbf{y}=\mathbf{s}^{* t} \mathbf{h}=0$ yields

$$
\begin{equation*}
\alpha_{o}=\mathbf{s}^{* t} \mathbf{u} \tag{1.4.7}
\end{equation*}
$$

Similarly premultiplying respectively by $\mathbf{b}_{\mathrm{x}}^{\mathrm{t}}, \mathbf{b}_{\mathrm{y}}^{\mathrm{t}}$ and $\mathbf{g}^{\mathrm{t}}$ and using the biorthogonality (4) yields

$$
\begin{align*}
& \alpha_{1}=\mathbf{b}_{\mathrm{x}}^{\mathrm{t}} \mathbf{u}  \tag{1.4.8a}\\
& \alpha_{2}=\mathbf{b}_{\mathrm{y}}^{\mathrm{t}} \mathbf{u}  \tag{1.4.8b}\\
& \alpha_{3}=\mathbf{g}^{\mathrm{t}} \mathbf{u} \tag{1.4.8c}
\end{align*}
$$

Substituting the above into (4) yields

$$
\begin{equation*}
u(x, y)=\left(\mathbf{s}^{* t}+x \mathbf{b}_{x}^{\mathrm{t}}+\mathrm{y} \mathbf{b}_{\mathrm{y}}^{\mathrm{t}}+\xi \eta \mathbf{g}^{\mathrm{t}}\right) \mathbf{u} \tag{1.4.9a}
\end{equation*}
$$

The two components of the displacement field can be expressed in the same form

$$
\begin{align*}
& \mathrm{u}_{\mathrm{x}}(\mathrm{x}, \mathrm{y})=\left(\mathbf{s}^{\mathrm{t}}+\mathrm{x} \mathbf{b}_{\mathrm{x}}^{\mathrm{t}}+\mathrm{y} \mathbf{b}_{\mathrm{y}}^{\mathrm{t}}+\mathrm{h} \mathbf{g}^{\mathrm{t}}\right) \mathbf{u}_{\mathrm{x}}  \tag{1.4.9b}\\
& \mathrm{u}_{\mathrm{y}}(\mathrm{x}, \mathrm{y})=\left(\mathbf{s}^{\mathrm{t}}+\mathrm{x} \mathbf{b}_{\mathrm{x}}^{\mathrm{t}}+\mathrm{y} \mathbf{b}_{\mathrm{y}}^{\mathrm{t}}+\mathrm{h} \mathbf{g}^{\mathrm{t}}\right) \mathbf{u}_{\mathbf{y}}  \tag{1.4.9c}\\
& \mathrm{h}=\xi \eta \tag{1.4.9d}
\end{align*}
$$

This is the same interpolation as the standard isoparametric form (1.2.1), however, this expression will more clearly reveal what causes locking and how to eliminate it.
1.4.2 Volumetric Locking. The four node quadrilateral locks in plane strain for incompressible materials when it is fully integrated. The cause of volumetric locking can be explained by considering a mesh of elements in plane strain with fixed boundaries on two sides as shown in Fig. 6. Consider the element in the lower left-hand corner, element 1. The nodal displacements of the element for an incompressible material must preserve the total volume of the element (or to be specific, the area in plane strain, since constant volume implies that the area be constant). If we consider small displacements, the only displacements of node 3 which maintain constant area are

$$
\begin{align*}
& u_{\mathrm{x} 3}=-\alpha \mathrm{a}  \tag{1.4.10}\\
& \mathrm{u}_{\mathrm{y} 3}=+\alpha \mathrm{b}
\end{align*}
$$

where $\alpha$ is an arbitrary parameter; all other nodal displacements of element 1 are zero due to the boundary conditions. Differentiating (9b) and (9c), we obtain the dilatation throughout the element.

$$
\begin{equation*}
\Delta=u_{x, x}+u_{y, y}=\mathbf{b}_{\mathrm{x}}^{\mathrm{t}} \mathbf{u}_{\mathrm{x}}+\mathbf{b}_{\mathrm{y}}^{\mathrm{t}} \mathbf{u}_{\mathrm{y}}+\frac{\partial \mathrm{h}}{\partial \mathrm{x}} \mathbf{g}^{\mathrm{t}} \mathbf{u}_{\mathrm{x}}+\frac{\partial \mathrm{h}}{\partial \mathrm{y}} \mathbf{g}^{\mathrm{t}} \mathbf{u}_{\mathrm{y}} \tag{1.4.11a}
\end{equation*}
$$

Substituting Eq. (10) into (11a), the constant part of the dilatation drops out leaving

$$
\begin{equation*}
\Delta=\frac{1}{4} \alpha\left(\frac{\partial h}{\partial y}-\frac{\partial h}{\partial x}\right) \tag{1.4.11b}
\end{equation*}
$$

which only vanishes everywhere except along a line that passes through the origin.
For rectangular elements as in Fig. 7, Eq. (11b) simplifies to

$$
\begin{equation*}
\Delta=\frac{\alpha}{\mathrm{ab}}(\mathrm{~b} \overline{\mathrm{x}}-\mathrm{ay}) \tag{1.4.12}
\end{equation*}
$$

where ( $\overline{\mathrm{x}}, \overline{\mathrm{y}}$ ) is a local coordinate system. The volumetric strain is non zero everywhere except along the line $\bar{y}=(b / a) \bar{x}$; therefore, for an incompressible material the volumetric energy will be infinite if the strain energy is evaluated exactly as is the case in a fully integrated element. Thus node 3 will not be able to move; nodes 2 and 3 then provide a rigid boundary for the left hand side of element 2 , and it can similarly be shown that by using these arguments for element 2 that node 6 cannot move. This argument can then be repeated for all nodes of the mesh to show that deformation of the mesh is impossible. This argument also applies to meshes of skewed elements as in Fig. 6.


Figure 6. Mesh of quadrilateral elements with fixed boundaries on two sides


Figure 7. Partial mesh of rectangular elements fixed on two sides
Another way to examine this behavior is consider an arbitrary element deformation as expressed by Eq. (4).

$$
u_{x}=\alpha_{0 x}+\alpha_{1 x} x+\alpha_{2 x} y+\alpha_{3 x} h
$$

$$
\begin{equation*}
u_{y}=\alpha_{0 y}+\alpha_{1 y} x+\alpha_{2 y} y+\alpha_{3 y} h \tag{1.4.13}
\end{equation*}
$$

We can evaluate the change in area of the element by integrating the dilatation over the element domain.

$$
\begin{equation*}
\int_{\Omega \mathrm{e}} \Delta \mathrm{dA}=\int_{\Omega \mathrm{e}}\left(\alpha_{1 \mathrm{x}}+\alpha_{2 \mathrm{y}}+\alpha_{3 \mathrm{x}} \frac{\partial \mathrm{~h}}{\partial \mathrm{x}}+\alpha_{3 \mathrm{y}} \frac{\partial \mathrm{~h}}{\partial \mathrm{y}}\right) \mathrm{dA} \tag{1.4.15}
\end{equation*}
$$

We can show algebraically using (1.2.7b) and (1.2.8a-d) an important property of $\mathrm{h}(\xi, \eta)$ :

$$
\begin{equation*}
\int_{\Omega_{e}} \frac{\partial \mathrm{~h}}{\partial \mathrm{x}} \mathrm{~d} \Omega=\int_{\Omega_{\mathrm{e}}} \frac{\partial \mathrm{~h}}{\partial \mathrm{y}} \mathrm{~d} \Omega=0 \tag{1.4.16}
\end{equation*}
$$

Therefore (15) is trivial to integrate and the change in element area is

$$
\begin{equation*}
\int_{\Omega \mathrm{e}} \Delta \mathrm{dA}=\left(\alpha_{1 \mathrm{x}}+\alpha_{2 \mathrm{y}}\right) \mathrm{A} \tag{1.4.17}
\end{equation*}
$$

which is zero only for $\alpha_{2 y}=-\alpha_{1 x}$. If we now consider volume preserving element deformation, i.e. $\alpha_{2 y}=-\alpha_{1 x}$, the dilatation is

$$
\begin{equation*}
\Delta=\alpha_{3 x} \frac{\partial \mathrm{~h}}{\partial \mathrm{x}}+\alpha_{3 \mathrm{y}} \frac{\partial \mathrm{~h}}{\partial \mathrm{y}} \tag{1.4.18}
\end{equation*}
$$

This dilatation will be non zero everywhere within the element except along the curve $\alpha_{3 x} \partial h / \partial x=-\alpha_{3 y} \partial h / \partial y$, even though the overall element deformation is volume preserving. Thus it becomes apparent that locking arises from the inability of the element to represent the isochoric field associated with the hourglass mode, as reflected by $\alpha_{3 x}$ and $\alpha_{3 y}$ in the above equations. To eliminate locking, the strain field must be designed, or projected, so that in the hourglass mode the dilatation in the projected strain field vanishes throughout the element. In more general terms this may be stated as follows: to avoid locking, the strain field must be isochoric throughout the element for any displacement field which preserves the total volume of the element. In particular, in the quadrilateral, the dilatation must vanish in the entire element for the hourglass mode, because this displacement mode is equivoluminal.
1.4.3 Variational principle The weak form corresponding to the Hu-Washizu variational principle is given for a single element domain by

$$
0=\delta \pi(\mathbf{u}, \mathbf{e}, \mathbf{s})=\int_{\Omega_{\mathrm{e}}} \delta \mathbf{e}^{\mathrm{t}} \mathbf{C e d} \Omega+\delta \int_{\Omega_{\mathrm{e}}} \mathbf{s}^{\mathrm{t}}(\mathbf{e}-\mathcal{D} \mathbf{u}) \mathrm{d} \Omega-\delta \mathrm{W}^{\mathrm{ext}}
$$

$$
\begin{equation*}
=\int_{\Omega_{\mathrm{e}}}\left[\delta \mathbf{e}^{\mathrm{t}}(\mathbf{C e}-\mathbf{s})-\delta \mathbf{s}^{\mathrm{t}}(\mathbf{e}-\mathcal{D} \mathbf{u})+\delta(\mathcal{D} \mathbf{u})^{\mathrm{t}} \mathbf{s}\right] \mathrm{d} \Omega-\delta \mathbf{d}^{\mathrm{t}} \mathrm{e}^{\mathrm{ext}} \tag{1.4.19}
\end{equation*}
$$

where $\delta$ denotes a variation, $\mathbf{e}$ is the interpolated strain, and $\mathbf{s}$ the interpolated stress. $\mathcal{D u}$ is the symmetric displacement gradient which would be equivalent to the strain in a displacement method. In mixed elements that are derived from the Hu-Washizu variational principle, the displacement gradient is projected on a smaller space to avoid locking. The term $\delta \mathrm{W}^{\text {ext }}$ designates the external work, $\mathbf{f}^{\text {ext }}$ the external nodal forces. The domain chosen for (19) is a single element, but an arbitrary domain can also be assumed if connectivity is introduced into the subsequent development.

The isoparametric shape functions are used to interpolate the displacement field, which when integrated, gives the symmetric displacement gradient as

$$
\begin{equation*}
\mathcal{D} \mathbf{u}=\mathbf{B} \mathbf{d} \tag{1.4.20}
\end{equation*}
$$

We introduce additional interpolants for the strains and stresses.

$$
\begin{align*}
& \mathbf{e}=\mathbf{E} \mathbf{e}  \tag{1.4.21}\\
& \mathbf{s}=\mathbf{S} \mathbf{s} \tag{1.4.22}
\end{align*}
$$

where the interpolation matrices, $\mathbf{E}$ and $\mathbf{S}$, and the augmented strains and stresses, $\mathbf{e}$ and $\mathbf{s}$ will be defined subsequently. Substituting (20), (21), and (22) into (19), we obtain

$$
\begin{equation*}
0=\int_{\Omega_{\mathrm{e}}}\left[\delta \mathbf{e}^{\mathrm{t}} \mathbf{E}^{\mathrm{t}}(\mathbf{C E e - S} \mathbf{s})-\delta \mathbf{s}^{\mathrm{t}} \mathbf{S}^{\mathrm{t}}(\mathbf{E e}-\mathbf{B d})+\delta \mathbf{d}^{\mathrm{t}} \mathbf{B}^{\mathrm{t}} \mathbf{S} \mathbf{s}\right] \mathrm{d} \Omega-\delta \mathbf{d}^{\mathrm{t}} \mathbf{f}_{\mathrm{ext}} \tag{1.4.23}
\end{equation*}
$$

By invoking the stationary condition on (19), we obtain

$$
\begin{align*}
& \widetilde{\mathbf{C}}=\widetilde{\mathbf{E}}^{\mathrm{t}} \mathbf{s}  \tag{1.4.24a}\\
& \widetilde{\mathbf{E}} \mathbf{e}=\widetilde{\mathbf{B}} \mathbf{d}  \tag{1.4.24b}\\
& \widetilde{\mathbf{B}}^{\mathrm{t}} \mathbf{s}=\mathbf{f}^{\mathrm{ext}} \tag{1.4.24c}
\end{align*}
$$

where

$$
\begin{align*}
& \widetilde{\mathbf{C}} \equiv \int_{\Omega_{\mathrm{e}}} \mathbf{E}^{\mathrm{t}} \mathbf{C E} \mathrm{~d} \Omega  \tag{1.4.25a}\\
& \widetilde{\mathbf{E}} \equiv \int_{\Omega_{\mathrm{e}}} \mathbf{S}^{\mathrm{t} \mathbf{E} \mathrm{~d} \Omega}  \tag{1.4.25b}\\
& \widetilde{\mathbf{B}} \equiv \int_{\Omega_{\mathrm{e}}} \mathbf{S}^{\mathrm{t} \mathbf{B} \mathrm{~d} \Omega} \tag{1.4.25c}
\end{align*}
$$

We obtain expressions for $\mathbf{e}, \mathbf{s}$, and the stiffness matrix from (24a-c).

$$
\begin{align*}
& \mathbf{e}=\widetilde{\mathbf{E}}^{-1} \widetilde{\mathbf{B}} \mathbf{d}  \tag{1.4.26a}\\
& \mathbf{s}=\widetilde{\mathbf{E}}^{-1} \widetilde{\mathbf{C}} \mathbf{e} \tag{1.4.26b}
\end{align*}
$$

$$
\begin{equation*}
\mathbf{f}^{\mathrm{ext}}=\widetilde{\mathbf{B}}^{\mathrm{t}} \mathbf{s}=\widetilde{\mathbf{B}}^{\mathrm{t}} \widetilde{\mathbf{E}}^{-1} \widetilde{\mathbf{C}}^{-1} \widetilde{\mathbf{B}} \mathbf{d} \equiv \mathbf{K}_{\mathrm{e}} \mathbf{d} \tag{1.4.26c}
\end{equation*}
$$

1.4.4 Strain Interpolations to Avoid Locking. The strain-field associated with the displacement field (9) can be obtained by straightforward differentiation which gives:

$$
\begin{align*}
\mathcal{D} \mathbf{u} & =\left\{\begin{array}{c}
\frac{\partial u_{x}}{\partial \mathrm{x}} \\
\frac{\partial u_{y}}{\partial \mathrm{y}} \\
\frac{\partial u_{x}}{\partial y}+\frac{\partial u_{y}}{\partial \mathrm{x}}
\end{array}\right\}=\left[\begin{array}{cc}
\mathbf{b}_{\mathrm{x}}^{\mathrm{t}}+\frac{\partial \mathrm{h}}{\partial \mathrm{x}} \mathbf{g}^{\mathrm{t}} & 0 \\
0 & \mathbf{b}_{\mathrm{y}}^{\mathrm{t}}+\frac{\partial \mathrm{h}}{\partial \mathrm{y}} \mathbf{g}^{\mathrm{t}} \\
\mathbf{b}_{\mathrm{y}}^{\mathrm{t}}+\frac{\partial \mathrm{h}}{\partial \mathrm{y}} \mathbf{g}^{\mathrm{t}} & \mathbf{b}_{\mathrm{x}}^{\mathrm{t}}+\frac{\partial \mathrm{h}}{\partial \mathrm{x}} \mathbf{g}^{\mathrm{t}}
\end{array}\right]\left(\left\{\begin{array}{l}
\mathbf{u}_{\mathrm{x}} \\
\mathbf{u}_{\mathrm{y}}
\end{array}\right\}=\mathbf{B d}\right.  \tag{1.4.27a}\\
& =\left\{\begin{array}{c}
\varepsilon_{\mathrm{x}}^{\mathrm{o}}+\mathrm{q}_{\mathrm{x}} \frac{\partial \mathrm{~h}}{\partial \mathrm{x}} \\
\varepsilon_{\mathrm{y}}^{\mathrm{o}}+\mathrm{q}_{\mathrm{y}} \frac{\partial \mathrm{~h}}{\partial \mathrm{y}} \\
2 \varepsilon_{\mathrm{xy}}^{\mathrm{o}}+\mathrm{q}_{\mathrm{x}} \frac{\partial \mathrm{~h}}{\partial \mathrm{y}}+\mathrm{q}_{\mathrm{y}} \frac{\partial \mathrm{~h}}{\partial \mathrm{x}}
\end{array}\right\} \tag{1.4.27b}
\end{align*}
$$

where the naughts indicate the constant part of the strain field.
In Section 1.4.2, it was demonstrated that QUAD4 elements of incompressible material can lock when the dilatational energy at any point other than the origin is included in the stiffness matrix. It was also shown that this is caused by the dilatational field associated with the hourglass modes, $\mathbf{d}^{\mathrm{Hx}}$ and $\mathbf{d}^{\mathrm{Hy}}$, which in a fully integrated element always leads to non-vanishing dilatation. From Eq. (27b), it can be seen that the hourglass modes generate the nonconstant part of the volumetric field.

In constructing a strain interpolant which will not lock volumetrically, we then have two alternatives:

1. the nonconstant terms of the first two rows of Eq. (27b) can be dropped
2. the first two rows can be modified so that no volumetric strains occur in the hourglass modes.

The first alternative leads to the strain field

$$
\mathbf{e}=\left\{\begin{array}{c}
\varepsilon_{x}^{o}  \tag{1.4.28a}\\
\varepsilon_{y}^{o} \\
2 \varepsilon_{x y}^{o}+q_{x} \partial h / \partial y+q_{y} \partial h / \partial x
\end{array}\right\}
$$

This can be written in the form of Eq. (17) by letting

$$
\begin{align*}
& \mathbf{E}=\left[\begin{array}{ccccc}
1 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 \\
0 & 0 & 1 & \partial \mathrm{~h} / \partial \mathrm{y} & \partial \mathrm{~h} / \partial \mathrm{x}
\end{array}\right]  \tag{1.4.28b}\\
& \mathbf{e}^{\mathrm{t}}=\left[\begin{array}{lllll}
\varepsilon_{\mathrm{x}}^{\mathrm{o}}, & \varepsilon_{\mathrm{y}}^{0}, & 2 \varepsilon_{\mathrm{xy}}^{0}, & \mathrm{q}_{\mathrm{x}}, & \mathrm{q}_{\mathrm{y}}
\end{array}\right. \tag{1.4.28c}
\end{align*}
$$

The second alternative is to define the strain field by

$$
\mathbf{e}=\left\{\begin{array}{c}
\varepsilon_{x}^{o}+q_{x} \partial h / \partial x-q_{y} \partial h / \partial y  \tag{1.4.29}\\
\varepsilon_{y}^{o}+q_{y} \partial h / \partial y-q_{x} \partial h / \partial x \\
2 \varepsilon_{x y}^{o}+q_{x} \partial h / \partial y+q_{y} \partial h / \partial x
\end{array}\right\}
$$

In Eq. (29), the dilatation $\varepsilon_{x}+\varepsilon_{\mathrm{y}}$ still vanishes in the hourglass mode, since regardless of the value of $q_{x}$ and $q_{y}$, Eq. (29) yields

$$
\begin{equation*}
\varepsilon_{\mathrm{x}}+\varepsilon_{\mathrm{y}}=\varepsilon_{\mathrm{x}}^{\mathrm{o}}+\varepsilon_{\mathrm{y}}^{0} \tag{1.4.30}
\end{equation*}
$$

The question then arises as to which of the two alternatives, (28) or (29), is preferable. The field in (29) is frame invariant whereas (28) is not. However, the computations associated with (28) are simpler. However, neither of these are particularly attractive for most problems from the viewpoints of accuracy and efficiency.

For elements which involve beam bending, the performance of the element can be improved strikingly by omitting the nonconstant part of the shear field. This shear strain field cannot be combined with the extensional strains in (28) because the strain field would then only contain three independent functions, and the element would be rank deficient. Therefore this shear strain field is combined with the extensional strains in (29), which gives

$$
\mathbf{e}=\left\{\begin{array}{c}
\varepsilon_{x}^{o}+q_{x} \partial h / \partial x-q_{y} \partial h / \partial y  \tag{1.4.31a}\\
\varepsilon_{y}^{o}+q_{y} \partial h / \partial y-q_{x} \partial h / \partial x \\
\varepsilon_{x y}^{o}
\end{array}\right\}
$$

For this element

$$
\mathbf{E}=\left[\begin{array}{ccccc}
1 & 0 & 0 & \partial h / \partial x & -\partial h / \partial y  \tag{1.4.31b}\\
0 & 1 & 0 & -\partial h / \partial x & \partial h / \partial y \\
0 & 0 & 1 & 0 & 0
\end{array}\right]
$$

with e given in (28c). The strain field (31) leads to the "Optimal Incompressible" or OI element in Belytschko and Bachrach. This element performs well in beam bending problems when one set of element sides are parallel to the axis of the beam and the elements are not too distorted.

The performance of QUAD4 in bending can be enhanced even further for isotropic, elastic problems by using a strain field which depends on Poisson's ratio as follows:

$$
\mathbf{e}=\left\{\begin{array}{c}
\varepsilon_{x}^{o}+q_{x} \frac{\partial h}{\partial x}-\bar{v} q_{y} \frac{\partial h}{\partial y}  \tag{1.4.32}\\
\varepsilon_{y}^{o}+q_{y} \frac{\partial h}{\partial y}-\bar{v} q_{x} \frac{\partial h}{\partial x} \\
2 \varepsilon_{x y}^{o}
\end{array}\right\}
$$

$$
\text { where } \bar{v} \equiv\left\{\begin{array}{c}
v \text { for plane stress } \\
v /(1-v) \text { for plane strain }
\end{array}\right.
$$

This is the field called "Quintessential Bending and Incompressible" or QBI in Belytschko and Bachrach (1986), which has great accuracy in bending with linear elastic material. For a rectangle, as shown by Fröier et al. (1974), this element corresponds to the incompatible quadrilateral of Wilson et al. (1973).
1.4.5 Stiffness Matrix for OI Element. In order to gain more insight into these mixed elements and to see how they are used to construct stabilization (rank-compensating) matrices which do not involve arbitrary parameters, the stiffness matrix for the OI element, which is based on the strain field (31b) will be developed.

The stress field is chosen to be

$$
\begin{align*}
& \mathbf{s}=\left[\begin{array}{ccccc}
1 & 0 & 0 & \partial \mathrm{~h} / \partial \mathrm{x} & 0 \\
0 & 1 & 0 & 0 & \partial \mathrm{~h} / \partial \mathrm{y} \\
0 & 0 & 1 & 0 & 0
\end{array}\right] \mathbf{s} \equiv \mathbf{S} \mathbf{~ s}  \tag{1.4.33.a}\\
& \mathbf{s}^{\mathrm{t}}=\left[\begin{array}{lllll}
\sigma_{\mathrm{x}}^{\mathrm{o}}, & \sigma_{y}^{\mathrm{o}}, & \sigma_{\mathrm{xy}}^{\mathrm{o}}, & \mathrm{Q}_{\mathrm{x}} & \mathrm{Q}_{\mathrm{y}}
\end{array}\right. \tag{1.4.33b}
\end{align*}
$$

Using (16), and integrating (25b) and (25c), we obtain $\widetilde{\mathbf{E}}$ and $\widetilde{\mathbf{B}}$ as follows:

$$
\widetilde{\mathbf{E}}=\left[\begin{array}{cc}
A \mathbf{I}_{3 \times 3} & \mathbf{0}_{3 \times 2}  \tag{1.4.34}\\
\mathbf{0}_{2 \mathrm{x} 3} & {\left[\begin{array}{cc}
\mathrm{H}_{\mathrm{xx}} & -\mathrm{H}_{\mathrm{xy}} \\
-\mathrm{H}_{\mathrm{xy}} & \mathrm{H}_{\mathrm{yy}}
\end{array}\right]}
\end{array}\right]
$$

where A is the area of the element, and

$$
\begin{equation*}
\mathrm{H}_{\mathrm{ij}}=\int_{\Omega_{\mathrm{e}}}\left(\partial \mathrm{~h} / \partial \mathrm{x}_{\mathrm{i}}\right)\left(\partial \mathrm{h} / \partial \mathrm{x}_{\mathrm{j}}\right) \mathrm{d} \Omega \tag{1.4.35}
\end{equation*}
$$

The $\widetilde{\mathbf{B}}$ matrix is given by

$$
\widetilde{\mathbf{B}}=\left[\begin{array}{cc}
\mathrm{A} \mathbf{b}_{\mathrm{x}}^{\mathrm{t}} & \mathbf{0}  \tag{1.4.36}\\
\mathbf{0} & \mathrm{~A} \mathbf{b}_{\mathrm{y}}^{\mathrm{t}} \\
\mathrm{~A} \mathbf{b}_{\mathrm{y}}^{\mathrm{t}} & \mathrm{~A} \mathbf{b}_{\mathrm{x}}^{\mathrm{t}} \\
\mathrm{H}_{\mathrm{xx}} \mathbf{g}^{\mathrm{t}} & \mathbf{0} \\
\mathbf{0} & \mathrm{H}_{\mathrm{yy}} \mathbf{g}^{\mathrm{t}}
\end{array}\right]
$$

The inverse of $\widetilde{\mathbf{E}}$ is given by

$$
\widetilde{\mathbf{E}}^{-1}=\left[\begin{array}{cc}
\frac{1}{\mathrm{~A}} \mathbf{I}_{3 \times 3} & \mathbf{0}_{3 \times 2}  \tag{1.4.37}\\
& \frac{1}{\mathrm{H}}\left[\begin{array}{cc}
\mathrm{H}_{\mathrm{yy}} & \mathrm{H}_{\mathrm{xy}} \\
\mathrm{H}_{\mathrm{xy}} & \mathrm{H}_{\mathrm{xx}}
\end{array}\right]
\end{array}\right]
$$

where

$$
\mathrm{H}=\mathrm{H}_{\mathrm{xx}} \mathrm{H}_{\mathrm{yy}}-\mathrm{H}_{\mathrm{xy}}^{2}
$$

Using (26a), (36) and (37), we can evaluate $\mathbf{e}$ to be

$$
\begin{align*}
& \varepsilon_{\mathrm{x}}^{\mathrm{o}}=\mathbf{b}_{\mathrm{x}}^{\mathrm{t}} \mathbf{u}_{\mathrm{x}} \\
& \varepsilon_{\mathrm{y}}^{\mathrm{o}}=\mathbf{b}_{\mathrm{y}}^{\mathrm{t}} \mathbf{u}_{\mathrm{y}} \\
& 2 \varepsilon_{\mathrm{y}}^{\mathrm{o}}=\mathbf{b}_{\mathrm{x}}^{\mathrm{t}} \mathbf{u}_{\mathrm{y}}+\mathbf{b}_{\mathrm{y}}^{\mathrm{t}} \mathbf{u}_{\mathrm{x}}  \tag{1.4.39}\\
& \mathrm{q}_{\mathrm{x}}=\frac{1}{\mathrm{H}}\left(\mathrm{H}_{\mathrm{xx}} \mathrm{H}_{\mathrm{yy}} \mathbf{g}^{\mathrm{t}} \mathbf{u}_{\mathrm{x}}+\mathrm{H}_{\mathrm{xy}} \mathrm{H}_{\mathrm{yy}} \mathbf{g}^{\mathrm{t}} \mathbf{u}_{\mathrm{y}}\right) \\
& \mathrm{q}_{\mathrm{y}}=\frac{1}{\mathrm{H}}\left(\mathrm{H}_{\mathrm{xy}} \mathrm{H}_{\mathrm{xx}} \mathbf{g}^{\mathrm{t}} \mathbf{u}_{\mathrm{x}}+\mathrm{H}_{\mathrm{xx}} \mathrm{H}_{\mathrm{yy}} \mathbf{g}^{\mathrm{t}} \mathbf{u}_{\mathrm{y}}\right)
\end{align*}
$$

From (39) we can see that in the mixed element, the constant parts of the strain field corresponds exactly to the constant part which emanates from the displacement field $\mathcal{D} \mathbf{u}$ given in (27). The nonconstant part depends strictly on the hourglass mode (any component of $\mathbf{u}_{x}$ or $\mathbf{u}_{y}$ which is not orthogonal to $\mathbf{g}$ ). The effect of the projection is to modify this part of the strain field so that the volumetric strains vanish.

To complete the evaluation of the element stiffness, we obtain $\widetilde{\mathbf{C}}$ by integrating (25a).

$$
\widetilde{\mathbf{C}}=\left[\begin{array}{cc}
\mathrm{AC}_{3 \times 3} & \mathbf{0}_{3 \mathrm{x} 2}  \tag{1.4.40a}\\
\mathbf{0}_{2 \mathrm{x} 3} & {\left[\begin{array}{cc}
4 \mu \mathrm{H}_{\mathrm{xx}} & -4 \mu \mathrm{H}_{\mathrm{xy}} \\
-4 \mu \mathrm{H}_{\mathrm{xy}} & 4 \mu \mathrm{H}_{\mathrm{yy}}
\end{array}\right]}
\end{array}\right]
$$

For a linear isotropic material, $\mathbf{C}$ is given by

$$
\mathbf{C}=\left[\begin{array}{ccc}
\bar{\lambda}+2 \mu & \bar{\lambda} & 0  \tag{1.4.40b}\\
\bar{\lambda} & \bar{\lambda}+2 \mu & 0 \\
0 & 0 & \mu
\end{array}\right]
$$

For plane strain, $\lambda=\lambda=2 \mu v /(1-2 v)$; for plane stress $\lambda=2 \mu \nu /(1-v)$. Substituting (28c), (37), and (40a) into (26b), we obtain

$$
\left\{\begin{array}{l}
\mathrm{Q}_{\mathrm{x}}  \tag{1.4.41}\\
\mathrm{Q}_{\mathrm{y}}
\end{array}\right\}=4 \mu\left\{\begin{array}{l}
\mathrm{q}_{\mathrm{x}} \\
\mathrm{q}_{\mathrm{y}}
\end{array}\right\}
$$

It can be seen already that because of the way the assumed strain field was designed, the nonconstant part depends only on the shear modulus $\mu$ and is independent of the bulk modulus.

Evaluating the stiffness by (26c),

$$
\begin{align*}
& \mathbf{K}_{\mathrm{e}}=\mathbf{B}^{\mathrm{t}}(\mathbf{0}) \mathbf{C B}(\mathbf{0})+\mathbf{K}_{\mathrm{e}}^{\text {stab }}  \tag{1.4.42a}\\
& \mathbf{K}_{\mathrm{e}}^{\text {stab }}=\left[\begin{array}{cc}
\mathrm{c}_{1} \mathbf{g} \boldsymbol{g} & \mathrm{c}_{2} \mathbf{g} \boldsymbol{g} \\
\mathrm{c}_{2} \mathbf{g} \boldsymbol{g} & \mathrm{c}_{3} \mathbf{g} \boldsymbol{g}
\end{array}\right. \tag{1.4.42b}
\end{align*}
$$

$\mathbf{B}(\mathbf{0})$ is given by (1.2.39). Constants for OI stabilization as well as those for QBI are given in Table 1. QBI stabilization is derived in the same way as OI with $\mathbf{E}$ given by (32).

Table 1. Constants for the mixed method stabilization matrix

| Stabilization | $\mathrm{c}_{1}$ | $\mathrm{c}_{2}$ | $\mathrm{c}_{3}$ |
| :--- | :--- | :--- | :--- |
| OI | $4 \mu \mathrm{H}_{\mathrm{xx}} \mathrm{H}^{*}$ | $4 \mu \mathrm{H}_{\mathrm{xy}} \mathrm{H}^{*}$ | $4 \mu \mathrm{H}_{\mathrm{yy}} \mathrm{H}^{*}$ |
| QBI | $2 \mu(1+\bar{v}) \mathrm{H}_{\mathrm{xx}} \mathrm{H}^{* *}$ | $2 \mu \bar{v}(1+\bar{v}) \mathrm{H}_{\mathrm{xy}} \mathrm{H}^{* *}$ | $2 \mu(1+\bar{v}) \mathrm{H}_{\mathrm{yy}} \mathrm{H}^{* *}$ |

Note. $\bar{v} \equiv v$ for plane stress and $\bar{v} \equiv v /(1-v)$ for plane strain;

$$
\mathrm{H}^{*} \equiv \frac{\mathrm{H}_{\mathrm{xx}} \mathrm{H}_{\mathrm{yy}}}{\mathrm{H}_{\mathrm{xx}} \mathrm{H}_{\mathrm{yy}}-\mathrm{H}_{\mathrm{xy}}^{2}} \quad \mathrm{H}^{* *} \equiv \frac{\mathrm{H}_{\mathrm{xx}} \mathrm{H}_{\mathrm{yy}}}{\mathrm{H}_{\mathrm{xx}} \mathrm{H}_{\mathrm{yy}}-\overline{\mathrm{v}} \mathrm{H}_{\mathrm{xy}}^{2}}
$$

Explicit approximate expressions for $\mathrm{H}_{\mathrm{ij}}$ can be obtained by integrating (28) in closed form assuming the Jacobian to be constant within the element domain.

$$
\begin{align*}
& \mathrm{H}_{\mathrm{xx}}=\frac{\left(\mathbf{L}_{1}^{\mathrm{t}} \mathbf{y}\right)^{2}+\left(\mathbf{L}_{2}^{\mathrm{t}} \mathbf{y}\right)^{2}}{3 \mathrm{~A}} \\
& \mathrm{H}_{\mathrm{yy}}=\frac{\left(\mathbf{L}_{1}^{\mathrm{t}} \mathbf{x}\right)^{2}+\left(\mathbf{L}_{2}^{\mathrm{t}} \mathbf{x}\right)^{2}}{3 \mathrm{~A}}  \tag{1.4.43}\\
& \mathrm{H}_{\mathrm{xy}}=\frac{-\left(\mathbf{L}_{1}^{\mathrm{t}} \mathbf{x}\right)\left(\mathbf{L}_{1}^{\mathrm{t}} \mathbf{y}\right)-\left(\mathbf{L}_{2}^{\mathrm{t}} \mathbf{x}\right)\left(\mathbf{L}_{2}^{\mathrm{t}} \mathbf{y}\right)}{3 \mathrm{~A}}
\end{align*}
$$

where

$$
\begin{align*}
& \mathbf{L}_{1}^{\mathrm{t}}=(-1,+1,+1,-1)  \tag{1.4.44}\\
& \mathbf{L}_{2}^{\mathrm{t}}=(-1,-1,+1,+1) .
\end{align*}
$$

The quantities used to evaluate $\mathrm{H}_{\mathrm{ij}}$ can also be used to evaluate the element area by

$$
\begin{equation*}
\mathrm{A}=\frac{1}{4}\left[\left(\mathbf{L}_{1}^{\mathrm{t}} \mathbf{x}\right)\left(\mathbf{L}_{2}^{\mathrm{t}} \mathbf{y}\right)-\left(\mathbf{L}_{1}^{\mathrm{t}} \mathbf{y}\right)\left(\mathbf{L}_{2}^{\mathrm{t}} \mathbf{x}\right)\right] \tag{1.4.45}
\end{equation*}
$$

The first part of the stiffness corresponds to the one-point quadrature stiffness. The second part is the stabilization or rank compensating stiffness, which is of rank 2 and thus increases the rank of the total stiffness from 3 (rank of one-point quadrature stiffness) to 5 . This is the correct rank of QUAD4 according to Eq. (1.2.26a).

The form given in (35) can be considered a canonical form for the stiffness matrix of QUAD4 if the constants $c_{i}$ are arbitrary. For any $c_{i}$, this element stiffness will satisfy the patch test. The constants $c_{i}$ can be varied to improve the performance of the element for specific problem classes, but as shown by numerical studies in Belytschko and Bachrach (1986), the rate of convergence will always be the same, provided the element does not lock. When the stabilization matrix is independent of the bulk modulus $\lambda+\frac{2}{3} \mu$, the element will not lock volumetrically.

This development also provides guidance about the design of stabilization procedures in nonlinear problems which are based on material properties. If the current linearized value of $\mu$ can be estimated, then (34) provides a stress-strain relation between $\dot{Q}_{i}$ and $\dot{\mathrm{q}}_{\mathrm{i}}$.
1.4.6 Frame Invariance. The elimination of the nonconstant part of the shear strain as is done with the OI and QBI elements improves their performance in bending problems. The cantilever test problems of the next section demonstrates the excellent coarse mesh bending accuracy of QBI; however, elimination of the nonconstant shear strain also causes the element to lose frame invariance. For most problems, the effect is negligible, but for coarse mesh bending, the effect can be significant.

Elements based on the OI or QBI assumed strain field can be made frame invariant by evaluating the stabilization matrix using an orthogonal, local coordinate system that is aligned with the element. The local coordinate system, called the ( $(\hat{x}, \widehat{y})$ system, is related to the global coordinate system by a rotation matrix $\mathbf{R}$.

The nodal coordinates, $\mathbf{x}$ and $\mathbf{y}$, evaluated in the local ( $(\hat{x}, \hat{y})$ coordinate system, are renamed $\widehat{\mathbf{x}}$ and $\widehat{\mathbf{y}}$. They are obtained by

$$
\left\{\begin{array}{l}
\widehat{\mathbf{x}}  \tag{1.4.46}\\
\widehat{\mathbf{y}}
\end{array}\right\}=\mathbf{R}\left\{\begin{array}{l}
\mathbf{x} \\
\mathbf{y}
\end{array}\right\}
$$

where $\mathbf{R}$ consists of the standard two dimensional rotation matrix arranged in an $8 \times 8$ matrix to transform each pair of nodal coordinates, $\mathrm{x}_{\mathrm{I}}$ and $\mathrm{y}_{\mathrm{I}}$.

$$
\mathbf{R}=\left[\begin{array}{cc}
\cos \theta \mathbf{I}_{4 \times 4} & \sin \theta \mathbf{I}_{4 \times 4}  \tag{1.4.47}\\
-\sin \theta \mathbf{I}_{4 \times 4} & \cos \theta \mathbf{I}_{4 \times 4}
\end{array}\right]
$$

where $\mathbf{I}_{4 \times 4}$ is a rank 4 identity matrix.
The angle between the global and local coordinate system can be defined by

$$
\begin{equation*}
\tan \theta=\mathbf{L}_{1}^{\mathrm{t}} \mathbf{y} / \mathbf{L}_{1}^{\mathrm{t}} \mathbf{x} \tag{1.4.48}
\end{equation*}
$$

This definition aligns the $\hat{x}$ axis with the referential $\xi$ axis of the element as shown in Fig. 8. This definition may not be appropriate for anisotropic material. This point is discussed further in the explicit formulation that follows.


Figure 8. Local ( $\widehat{\mathrm{x}}, \widehat{\mathrm{y}}$ ) coordinate system aligned with $\xi$ axis of an element
We evaluate $\widehat{\mathbf{b}}_{\mathrm{x}}, \widehat{\mathbf{b}}_{\mathrm{y}}$, and $\hat{\mathbf{g}}$ by substituting $\widehat{\mathbf{x}}$ and $\widehat{\mathbf{y}}$ for $\mathbf{x}$ and $\mathbf{y}$ in (1.2.40) and (1.3.15).

$$
\begin{align*}
& \hat{\mathbf{b}}_{\mathrm{x}}^{\mathrm{t}}=\frac{1}{2 \mathrm{~A}}\left[\widehat{\mathrm{y}}_{24}, \widehat{\mathrm{y}}_{31}, \widehat{\mathrm{y}}_{42}, \widehat{\mathrm{y}}_{13}\right]  \tag{1.4.49a}\\
& \widehat{\mathbf{b}}_{\mathrm{y}}^{\mathrm{t}}=\frac{1}{2 \mathrm{~A}}\left[\widehat{\mathrm{x}}_{42}, \widehat{\mathrm{x}}_{13}, \widehat{\mathrm{x}}_{24}, \widehat{\mathrm{x}}_{31}\right]  \tag{1.4.49b}\\
& \widehat{\mathrm{y}}_{\mathrm{IJ}}=\widehat{\mathrm{y}}_{\mathrm{F}} \widehat{\mathrm{y}}_{\mathrm{J}} \widehat{\mathrm{x}}_{\mathrm{IJ}}=\widehat{\mathrm{x}}_{\mathrm{r}}-\widehat{\mathrm{x}}_{\mathrm{J}}  \tag{1.4.49c}\\
& \widehat{\mathbf{g}}=\frac{1}{4}\left[\mathbf{h}+\left(\mathbf{h}^{t} \mathbf{x}\right) \widehat{\mathbf{b}}_{\mathrm{x}}+\left(\mathbf{h}^{\mathrm{t}} \widehat{\mathbf{y}}^{\mathbf{y}} \widehat{\mathbf{b}}_{\mathrm{y}}\right]\right. \tag{1.4.50}
\end{align*}
$$

Hats are added to terms to indicate that local coordinates are used in their evaluation. $\widehat{\mathrm{H}}_{\mathrm{xx}}$, $\widehat{\mathrm{H}}_{\mathrm{yy}}$, and $\widehat{\mathrm{H}}_{\mathrm{xy}}$, are evaluated by Eq. (44) with $\widehat{\mathbf{x}}$ and $\widehat{\mathbf{y}}$ substituted for $\mathbf{x}$ and $\mathbf{y}$. Likewise, the constants in Table 1 are evaluated in terms of $\widehat{\mathrm{H}}_{\mathrm{xx}}, \widehat{\mathrm{H}}_{\mathrm{yy}}$, and $\widehat{\mathrm{H}}_{\mathrm{xy}}$ and are renamed $\hat{\mathrm{c}}_{1}$, $\hat{\mathbf{c}}_{2}$, and $\hat{\mathbf{c}}_{3}$. The stabilization matrix is analogous to (42b) and is given by

$$
\widehat{\mathbf{K}}_{\mathrm{e}}^{\text {stab }}=\left[\begin{array}{ll}
\hat{\mathbf{c}}_{1} \widetilde{\mathbf{g}}^{\mathrm{t}} & \hat{\mathbf{c}}_{2} \widetilde{\mathbf{g}}^{\mathrm{t}}  \tag{1.4.51}\\
\hat{\mathrm{c}}_{2} \widehat{\mathbf{g}}^{\mathrm{g}} & \hat{\mathbf{c}}_{3} \widehat{\mathbf{g}}^{\mathrm{g}}
\end{array}\right]
$$

In order to add the element stabilization matrix to the global stiffness matrix, it must be transformed back to the global coordinate system by

$$
\begin{equation*}
\mathbf{K}_{\mathrm{e}}^{\mathrm{stab}}=\mathbf{R}^{\mathrm{t}} \widehat{\mathbf{K}}_{\mathrm{e}}^{\mathrm{stab}} \mathbf{R} \tag{1.4.52}
\end{equation*}
$$

It is simple enough to evaluate $\mathbf{K}_{\mathrm{e}}^{\text {stab }}$ in closed form as

$$
\mathbf{K}_{\mathrm{e}}^{\text {stab }}=\left[\begin{array}{ll}
\hat{\mathbf{c}}_{1}^{*} \widehat{\mathbf{g g}}^{\mathrm{t}} & \hat{\mathbf{c}}_{2}^{*} \widehat{\mathbf{g g}}^{\mathrm{t}}  \tag{1.4.53}\\
\hat{\mathbf{c}}_{2}^{*} \widehat{\mathbf{g g}}^{\mathrm{t}} & \hat{\mathbf{c}}_{3}^{*} \widehat{\mathbf{g g}}^{\mathrm{t}}
\end{array}\right]
$$

where

$$
\begin{align*}
& \hat{\mathbf{c}}_{1}^{*}=\hat{\mathbf{c}}_{1} \cos ^{2} \theta+\hat{\mathbf{c}}_{3} \sin ^{2} \theta-2 \hat{\mathbf{c}}_{2} \sin \theta \cos \theta \\
& \hat{\mathbf{c}}_{2}^{*}=\hat{\mathbf{c}}_{2}\left(\cos ^{2} \theta-\sin ^{2} \theta\right)+\left(\hat{\mathrm{c}}_{1}-\hat{\mathbf{c}}_{3}\right) \sin \theta \cos \theta  \tag{1.4.54}\\
& \hat{\mathbf{c}}_{3}^{*}=\hat{\mathbf{c}}_{3} \cos ^{2} \theta+\hat{\mathbf{c}}_{1} \sin ^{2} \theta+2 \hat{\mathbf{c}}_{2} \sin \theta \cos \theta
\end{align*}
$$

1.4.7 Hourglass Control Procedure. We seek to evaluate internal forces directly by the first equality in Eq. (26c) which in corotational coordinates is

$$
\begin{equation*}
\hat{\mathbf{f}}^{\text {int }}=\hat{\widetilde{B}}^{\mathrm{t}} \widehat{\mathbf{s}} \tag{1.4.55}
\end{equation*}
$$

For the OI and QBI strain and stress fields, Eq. (55) can be shown to take the form of a one-point element plus stabilization forces:

$$
\begin{equation*}
\hat{\mathbf{f}}^{\text {int }}=\mathrm{A} \widehat{\mathbf{B}}^{\mathrm{t}}(\mathbf{0}) \widehat{\mathbf{s}}(\mathbf{0})+\hat{\mathbf{f}}^{\text {stab }} \tag{1.4.56}
\end{equation*}
$$

A procedure for large deformation, nonlinear problems based on the previous analysis of the mixed element is described. The mixed field OI given in Section 1.4.5 will be used for this purpose. Implementations based on other assumed strain fields can be developed similarly.

The development hinges on the fact that the linear theory developed in Section 1.4.5 is identical to the nonlinear theory if all variables are interpreted as rates. Thus components of the generalized hourglass strain rates can be obtained by Eq. (26a) written in the form

$$
\begin{align*}
& \dot{\hat{\mathrm{q}}}_{\mathrm{x}}=\frac{1}{\widehat{\mathrm{H}}}\left(\widehat{\mathrm{H}}_{\mathrm{xx}} \widehat{\mathrm{H}}_{\mathrm{yy}} \hat{\mathbf{g}}^{\mathrm{t}} \widehat{\mathbf{v}}_{\mathrm{x}}+\widehat{\mathrm{H}}_{\mathrm{xy}} \widehat{\mathrm{H}}_{\mathrm{yy}} \hat{\mathbf{g}}^{\mathrm{t}} \hat{\mathbf{v}}_{\mathrm{y}}\right)  \tag{1.4.57}\\
& \dot{\mathrm{q}}_{\mathrm{y}}=\frac{1}{\widehat{\mathrm{H}}}\left(\widehat{\mathrm{H}}_{\mathrm{xy}} \widehat{\mathrm{H}}_{\mathrm{xx}} \hat{\mathbf{g}}^{t} \mathbf{v}_{\mathrm{x}}+\widehat{\mathrm{H}}_{\mathrm{xx}} \widehat{\mathrm{H}}_{\mathrm{yy}} \hat{\mathbf{g}}^{\mathrm{t}} \widehat{\mathbf{v}}_{\mathrm{y}}\right)
\end{align*}
$$

where the superposed carets indicate that the quantities are evaluated using a corotational coordinate system. This formulation results in a frame invariant element. The corotational coordinate system is equivalent to the local coordinate system presented in the previous section, however it is embedded in the element and rotates with the element as the element deforms. The corotational coordinate system can be embedded by various techniques, and for anisotropic materials, it is advantageous to embed the coordinate system so that it coincides with axes of orthotropy or other directional features of the material.

The corotational coordinate system can also be used to evaluate the rate-ofdeformation and update the stress at the quadrature point. An advantage of the corotational system is that a frame invariant stress rate is not needed for large deformation problems.

The stress-strain law for the generalized hourglass strain rates and stress rates is given by

$$
\begin{equation*}
\dot{\hat{\mathrm{Q}}}_{\mathrm{i}}=4 \mu \dot{\hat{\mathrm{q}}}_{\mathrm{i}} \tag{1.4.58}
\end{equation*}
$$

This relation involves the shear modulus $\mu$ and assumes an isotropic material response. The shear modulus, $\mu$ is obtained by taking the ratio of the effective deviatoric stress rates and strain rates.

$$
\begin{align*}
& 2 \mu=\left(\frac{\dot{\hat{\mathrm{s}}}_{\mathrm{ij}} \dot{\hat{\mathrm{~s}}}_{\mathrm{ij}}}{\frac{\hat{\mathrm{e}}_{\mathrm{ij}} \dot{\mathrm{e}}_{\mathrm{ij}}}{i}}\right)^{\frac{1}{2}}  \tag{1.4.59a}\\
& \dot{\hat{\mathrm{~s}}}_{\mathrm{ij}}=\dot{\hat{\sigma}}_{\mathrm{ij}}-\frac{1}{3} \dot{\mathrm{p}} \delta_{\mathrm{ij}}  \tag{1.4.59b}\\
& \dot{\hat{\mathrm{e}}}_{\mathrm{ij}}=\dot{\hat{\varepsilon}}_{\mathrm{ij}}-\frac{1}{3} \dot{\hat{\varepsilon}}_{\mathrm{kk}} \delta_{\mathrm{ij}} \tag{1.4.59c}
\end{align*}
$$

In two-dimensional plane stress problems

$$
\begin{align*}
& \dot{\hat{\mathrm{s}}}_{\mathrm{ij}} \dot{\hat{\mathrm{~s}}}_{\mathrm{ij}}=\dot{\dot{\sigma}}_{\mathrm{x}}^{2}-\dot{\hat{\sigma}}_{\mathrm{x}} \dot{\hat{\sigma}}_{\mathrm{y}}+\dot{\hat{\sigma}}_{\mathrm{x}}^{2}+3 \dot{\hat{\sigma}}_{\mathrm{xy}}^{2}  \tag{1.4.60a}\\
& \dot{\hat{\mathrm{e}}}_{\mathrm{ij}} \dot{\hat{e}}_{\mathrm{ij}}=\dot{\hat{\varepsilon}}_{\mathrm{x}}^{2}-\dot{\hat{\varepsilon}}_{\mathrm{x}} \dot{\hat{\varepsilon}}_{\mathrm{y}}+\dot{\hat{\varepsilon}}_{\mathrm{x}}^{2}+3 \dot{\hat{\varepsilon}}_{\mathrm{xy}}^{2} \tag{1.4.60b}
\end{align*}
$$

The stabilization stresses are then updated by

$$
\begin{equation*}
\widehat{\mathrm{Q}}_{\mathrm{i}}^{\mathrm{n}+1}=\hat{\mathrm{Q}}_{\mathrm{i}}^{\mathrm{n}}+\int_{\mathrm{t}^{\mathrm{n}}}^{\mathrm{t}^{\mathrm{n}+1}} \dot{\hat{\mathrm{Q}}}_{\mathrm{i}} \mathrm{dt} \tag{1.4.61a}
\end{equation*}
$$

which for the central difference method gives

$$
\begin{equation*}
\widehat{\mathrm{Q}}_{\mathrm{i}}^{\mathrm{n}+1}=\widehat{\mathrm{Q}}_{\mathrm{i}}^{\mathrm{n}}+\Delta \mathrm{t} \dot{\mathrm{Q}}_{\mathrm{i}}^{\mathrm{n}+1 / 2} \tag{1.4.61b}
\end{equation*}
$$

The stabilization internal forces, evaluated by (26b), are then given by

$$
\begin{align*}
& \hat{\mathbf{f}}_{\mathrm{x}}^{\mathrm{stab}}=\left(\widehat{\mathrm{H}}_{\mathrm{xx}} \widehat{\mathrm{Q}}_{\mathrm{x}}-\widehat{\mathrm{H}}_{\mathrm{xy}} \widehat{\mathrm{Q}}_{\mathrm{y}}\right) \hat{\mathbf{g}}  \tag{1.4.62a}\\
& \hat{\mathbf{f}}_{\mathrm{y}}^{\text {stab }}=\left(\widehat{\mathrm{H}}_{\mathrm{yy}} \widehat{\mathrm{Q}}_{\mathrm{y}}-\widehat{\mathrm{H}}_{\mathrm{xy}} \widehat{\mathrm{Q}}_{\mathrm{x}}\right) \hat{\mathbf{g}} \tag{1.4.62b}
\end{align*}
$$

Not that the stress and strain which is used to evaluate the shear modulus is marked with hats to indicate that these are corotational quantities. This is not necessary since the shear modulus is an invariant quantity for isotropic material.

The assumptions made in this development is that the material response is uniform over the element and the deviatoric response is isotropic. The second assumption can be avoided by using a $\widetilde{\mathbf{C}}$ matrix based on a fully anisotropic $\mathbf{C}$. However, this entails availability of $\mathbf{C}$ in the computational process, and in procedures such as radial return for elastoplasticity, $\mathbf{C}$ is not available. The first assumption is more troublesome; as the elastic-plastic front passes across an element, one-point quadrature is not as effective in resolving the behavior along the boundary. This effect has been noted in Liu et al. (1988). Usually, however, the substantially reduced cost of one-point quadrature elements allows more elements to compensate for this effect. Adaptive schemes with automatic mesh refinement in zones of rapidly varying material behavior are also effective. To avoid these difficulties, assumed strain methods with 2 or more quadrature points as described in Section 1.5.6 can be used.

### 1.5 Assumed Strain Hourglass Stabilization

In this section, the stabilization procedure for the quadrilateral will be developed by means of the assumed strain methodology. The arguments used in constructing the assumed strain field for this procedure are identical to those used with the Hu-Washizu principle. However, the implementation is much simpler because many of the intermediate matrices which are required in the Hu-Washizu approach can be bypassed. Nevertheless, the results obtained by the assumed strain procedure differ very little from the results obtained by the corresponding Hu-Washizu elements.

The assumed strain approach can also be used in conjunction with quadrature schemes which use more than one point. This avoids the use of stabilization schemes, but does require substantially more effort if the constitutive equations are complex.

In addition to describing the assumed strain method, the notion of projection of strains is examined further in this section. It is shown that the assumed strain fields which eliminate volumetric locking and excessive stiffness in bending problems correspond to projections of the higher order terms in the strain field.
1.5.1 Variational principle Assumed strain elements herein are based on a simplified form of the Hu-Washizu variational principle as described by Simo and Hughes (1986) in which the interpolated stress is assumed to be orthogonal to the difference between the symmetric part of the velocity gradient and the interpolated rate-of-deformation. Therefore, the second term of (1.4.19) drops out leaving

$$
\begin{equation*}
0=\delta \pi(\mathbf{e})=\int_{\Omega_{\mathrm{e}}} \delta \mathbf{e}^{\mathrm{t}} \operatorname{Ced} \Omega-\delta \mathbf{d}^{\mathrm{t}} \mathbf{e}^{\mathrm{ext}} \tag{1.5.1}
\end{equation*}
$$

In this form, the interpolated stress does not need to be defined since it no longer appears in the variational principle.

The discrete equations then require only the interpolation of the strain, which we relate to the nodal displacements by $\overline{\mathbf{B}}$ which will be defined later.

$$
\begin{equation*}
\mathbf{e}(\mathbf{x}) \equiv \overline{\mathbf{B}}(\mathbf{x}) \mathbf{d} \tag{1.5.2}
\end{equation*}
$$

Substituting (2) into (1) gives

$$
\begin{equation*}
0=\delta \mathbf{d}^{\mathrm{t}} \int_{\Omega_{\mathrm{e}}} \overline{\mathbf{B}}^{\mathrm{t}} \mathbf{C} \overline{\mathbf{B}} \mathrm{~d} \Omega \mathbf{d}-\delta \mathbf{d}^{\mathrm{t}} \mathbf{f}^{\mathrm{ext}} \tag{1.5.3}
\end{equation*}
$$

so the arbitrariness of $\delta \mathbf{d}$ leads to

$$
\begin{equation*}
\mathbf{f}^{\mathrm{int}}=\mathbf{f}^{\mathrm{ext}} \tag{1.5.4}
\end{equation*}
$$

where

$$
\begin{equation*}
\mathbf{f}^{\mathrm{int}}=\mathbf{K}_{\mathrm{e}} \mathbf{d} \tag{1.5.5}
\end{equation*}
$$

and

$$
\begin{equation*}
\mathbf{K}_{\mathrm{e}}=\int_{\Omega_{\mathrm{e}}} \overline{\mathbf{B}}^{\mathrm{t}} \mathbf{C} \cdot \overline{\mathbf{B}} \mathrm{~d} \Omega \tag{1.5.6}
\end{equation*}
$$

The stiffness matrix of the fully integrated isoparametric element is found by (1.2.28). The application of the assumed strain method to the development of a stabilization procedure for an underintegrated element then involves the construction of an appropriate form for the $\overline{\mathbf{B}}$ matrix which avoids locking.
1.5.2 Elimination of Volumetric Locking. To eliminate volumetric locking, the strain field must be projected so that the volumetric strain energy always vanishes in the hourglass mode. For this purpose, we consider a general form of the assumed strain

$$
\begin{align*}
& \mathbf{e}=\left\{\begin{array}{c}
\varepsilon_{x}^{o}+q_{x} e_{1} h,{ }_{x}+q_{y} e_{2} h, y \\
\varepsilon_{y}^{o}+q_{x} e_{2} h,{ }_{x}+q_{y} e_{1} h, y \\
2 \varepsilon_{x y}^{o}+q_{x} e_{3} h, y+q_{y} e_{3} h, x
\end{array}\right\} \equiv\left\{\begin{array}{c}
\varepsilon_{x}^{o}+\widetilde{\varepsilon}_{x} \\
\varepsilon_{y}^{o}+\widetilde{\varepsilon}_{y} \\
2 \varepsilon_{x y}^{o}+2 \widetilde{\varepsilon}_{x y}
\end{array}\right\}  \tag{1.5.7a}\\
& \mathrm{q}_{\mathrm{x}}=\mathbf{g}^{\mathrm{t} \mathbf{u}_{x} \quad q_{y}=\mathbf{g}^{\mathrm{t}} \mathbf{u}_{y}} \tag{1.5.7b}
\end{align*}
$$

where $e_{1}, e_{2}$, and $e_{3}$ are arbitrary constants, and $q_{x}$ and $q_{y}$ are the magnitudes of the hourglass modes, which vanish except when the element is in the hourglass mode. In (7a) and subsequent equations, commas denote derivatives with respect to the variables that follow. Substituting (1.4.39) and (7b) into (7a), the assumed strain field is put into $\overline{\mathbf{B}}$ form as follows:

$$
\begin{align*}
\mathbf{e} & =\overline{\mathbf{B}} \mathbf{d} \\
\overline{\mathbf{B}} & =\left[\begin{array}{cc}
\mathbf{b}_{\mathrm{x}}^{\mathrm{t}}+\mathrm{e}_{1} \mathrm{~h},{ }_{\mathrm{x}} \mathbf{g}^{\mathrm{t}} & \mathrm{e}_{2} \mathrm{~h},{ }_{y} \mathbf{g}^{\mathrm{t}} \\
\mathrm{e}_{2} \mathrm{~h},{ }_{\mathrm{x}} \mathbf{g}^{\mathrm{t}} & \mathbf{b}_{\mathrm{y}}^{\mathrm{T}}+\mathrm{e}_{1} \mathrm{~h}, \mathbf{g}^{\mathrm{t}} \\
\mathbf{b}_{\mathrm{y}}^{\mathrm{T}}+\mathrm{e}_{3} \mathrm{~h},{ }_{\mathrm{y}} \mathbf{g}^{\mathrm{t}} & \mathbf{b}_{\mathrm{x}}^{\mathrm{T}}+\mathrm{e}_{3} \mathrm{~h},{ }_{\mathrm{x}} \mathbf{g}^{\mathrm{t}}
\end{array}\right] \tag{1.5.8}
\end{align*}
$$

For the purpose of illustrating the projections, the symmetric displacement gradient (1.4.27b) is written as

$$
\mathcal{D} \mathbf{u}=\left\{\begin{array}{c}
u_{x, x}^{o}+\widetilde{u}_{x, x}  \tag{1.5.9}\\
u_{y, y}^{o}+\widetilde{u}_{y, y} \\
u_{x, y}^{o}+\widetilde{u}_{x, y}+u_{y, x}^{o}+\widetilde{u}_{y, x}
\end{array}\right\}=\left\{\begin{array}{c}
\varepsilon_{x}^{o}+q_{x} h, x \\
\varepsilon_{y}^{o}+q_{y} h, y \\
2 \varepsilon_{x y}^{o}+q_{x} h, y+q_{y} h, x
\end{array}\right\}
$$

The dilatation of the assumed strain field given by Eq. (2a), which is denoted by $\bar{\Delta}$, vanishes in the hourglass mode if $\mathrm{e}_{1}=-\mathrm{e}_{2}$. This is shown as follows. Consider the nodal displacements that correspond to the hourglass mode of deformation.

$$
\begin{equation*}
\mathbf{u}_{\mathrm{x}}=\alpha_{3 x} \mathbf{h} \quad \mathbf{u}_{\mathrm{y}}=\alpha_{3 y} \mathbf{h} \tag{1.5.10}
\end{equation*}
$$

Evaluating the strain by (2), we obtain the dilatation as

$$
\begin{equation*}
\bar{\Delta}=\varepsilon_{\mathrm{x}}+\varepsilon_{\mathrm{y}}=\left(\mathrm{e}_{1}+\mathrm{e}_{2}\right)\left(\alpha_{3 \mathrm{x}} \mathrm{~h}, \mathrm{x}+\alpha_{3 \mathrm{y}} \mathrm{~h}, \mathrm{y}\right) \tag{1.5.11}
\end{equation*}
$$

So for $\mathrm{e}_{1}=-\mathrm{e}_{2}, \bar{\Delta}=0$. Thus, with this projected strain, the dilatation vanishes throughout the element in the hourglass mode. Furthermore, it can easily be shown that for the meshes in Figs. 6 and 7 with the nodal displacements given by (1.4.10), the dilatation $\bar{\Delta}$ vanishes throughout the element.

For linear elastic material with a constitutive matrix given by (1.4.40b), and the nodal displacements given in Eq. (10), the strain energy of the assumed strain element with $e_{2}=-e_{1}$ is

$$
\begin{align*}
\mathrm{U} & =\frac{1}{2} \int_{\Omega_{e}} \mathrm{e}^{\mathrm{t}} \mathbf{C e d} \Omega  \tag{1.5.12}\\
& =\mu \mathrm{e}_{1}^{2}\left(\alpha_{3 \mathrm{x}}^{2} \mathrm{H}_{\mathrm{xx}}+\alpha_{3 \mathrm{x}} \alpha_{3 y} H_{x y}+\alpha_{3 y}^{2} H_{y y}\right)+\frac{1}{2} \mu \mathrm{e}_{3}^{2}\left(\alpha_{3 y}^{2} \mathrm{H}_{\mathrm{xx}}+\alpha_{3 \mathrm{x}} \alpha_{3 y} H_{x y}+\alpha_{3 x}^{2} \mathrm{H}_{y y}\right)
\end{align*}
$$

which is independent of the bulk modulus. Thus the volumetric energy in this element is always finite and the element will not be subject to volumetric locking.

The portion of the volumetric strain which has been eliminated by this projection is often called "spurious" or "parasitic" volumetric strain. Whatever the name, it is certainly undesirable for the treatment of incompressible materials. Since in the nonlinear range, many materials are incompressible, its elimination from the element is crucial.

The character of this projection for various values of $e_{1}$ (when $e_{1}=-e_{2}$ ) is shown in Fig. 9. The two axes represent the nonconstant terms in $u_{x, x}$ and $u_{y, y}$, which are denoted by $\widetilde{\mathrm{u}}_{\mathrm{x}, \mathrm{x}}$ and $\widetilde{\mathrm{u}}_{\mathrm{y}, \mathrm{y}}$, and the corresponding terms in the assumed strain (compare Eqs. (8) and (9)) $\widetilde{\varepsilon}_{\mathrm{x}}$ and $\widetilde{\varepsilon}_{\mathrm{y}}$ respectively. The square represents an example of a point in ( $\widetilde{\mathrm{u}}_{\mathrm{x}, \mathrm{x}}, \widetilde{\mathrm{u}}_{\mathrm{y}, \mathrm{y}}$ )
space, while the circles represent corresponding points in $\left(\widetilde{\varepsilon}_{\mathrm{x}}, \widetilde{\varepsilon}_{\mathrm{y}}\right)$ space. From the formula relating these quantities, namely

$$
\begin{align*}
& \widetilde{\varepsilon}_{\mathrm{x}}=\mathrm{e}_{1}\left(\widetilde{\mathrm{u}}_{\mathrm{x}, \mathrm{x}}-\widetilde{\mathrm{u}}_{\mathrm{y}, \mathrm{y}}\right)  \tag{1.5.13a}\\
& \widetilde{\varepsilon}_{\mathrm{y}}=\mathrm{e}_{1}\left(\mathrm{u}_{\mathrm{u}, \mathrm{y}}-\widetilde{\mathrm{u}}_{\mathrm{x}, \mathrm{x}}\right) \tag{1.5.13b}
\end{align*}
$$

it can be seen the $\mathrm{e}_{1}=\frac{1}{2}$ corresponds to a normal projection of the functions $\widetilde{\mathrm{u}}_{\mathrm{x}, \mathrm{x}}, \widetilde{\mathrm{u}}_{\mathrm{y}, \mathrm{y}}$ onto the line $\widetilde{\varepsilon}_{\mathrm{x}}+\widetilde{\varepsilon}_{\mathrm{y}}=0$, which is the line on which the higher order terms in the assumed strain field posses no dilatation. Other values of $e_{1}$ shift the higher order terms of the assumed strain along the same line.

is a point representing $\widetilde{\mathrm{u}}_{\mathrm{x}, \mathrm{x}}, \widetilde{\mathrm{u}}_{\mathrm{y}, \mathrm{y}}$
$\bigcirc$ are points representing $\widetilde{\varepsilon}_{\mathrm{x}}, \widetilde{\varepsilon}_{\mathrm{y}}$
Figure 9. Depiction of projection of nonconstant part of displacement gradient $\widetilde{\mathrm{u}}_{\mathrm{x}, \mathrm{x}}, \widetilde{\mathrm{u}}_{\mathrm{y}, \mathrm{y}}$ onto isochoric assumed strain fields
1.5.3 Shear Locking and its Elimination. Shear locking in the four-node quadrilateral may be explained and eliminated by projection in a similar manner. It should be mentioned, and this will become clear from the results, that the effect of "spurious" shear is somewhat different than that of "spurious" strains in volumetric locking. In volumetric locking, the results completely fail to converge; with spurious shear, the solutions converge but rather slowly. Thus the term "excessive shear stiffness" is probably more precise, but the term shear locking is also a useful description.

To understand shear locking and its elimination, consider a beam represented by a single row of elements which is in pure bending as shown in Fig. 10. In pure bending, the moment field is constant and as is well known to structural engineers, the shear must vanish, since the shear is the derivative of the moment with respect to $x: s=m, x$.


Figure 10. A beam in pure bending showing that the deformation is primarily into the hourglass mode

To eliminate shear locking, the portion of the shear field which is triggered by any nodal displacements which are not orthogonal to $\mathbf{g}$ must be eliminated. Since only $\mathbf{h}$ is not orthogonal to $\mathbf{g}$, this is another way of saying that the shear associated with the hourglass mode must be eliminated. This can be accomplished by letting $e_{3}=0$ in Eq. (8). In pure bending, the nodal displacements in the local coordinate system of the element defined as shown in Fig. 10 are given by

$$
\begin{equation*}
\widehat{\mathbf{u}}_{\mathrm{x}}=\mathrm{ch} \quad \widehat{\mathbf{u}}_{\mathrm{y}}=\mathbf{0} \tag{1.5.14}
\end{equation*}
$$

where c is an arbitrary constant. If the strain energy is computed using Eq. (8) for arbitrary $e_{3}$, we find that the shear strain energy

$$
\begin{equation*}
U_{\text {shear }}=\frac{1}{2} \mu \mathrm{e}_{3}^{2} \mathrm{c}^{2} \mathrm{H}_{\mathrm{yy}}=0 \tag{1.5.15}
\end{equation*}
$$

so it vanishes as expected when $\mathrm{e}_{3}=0$; parasitic shear in bending is thus eliminated. This corresponds to the projection illustrated in Fig. 11. The shear field emanating from the displacement field can be written in terms of the 3 parameters $\varepsilon_{\mathrm{xy}}^{0}, q_{\mathrm{x}}$, and $\mathrm{q}_{\mathrm{y}}$. The second and third parameters are associated with the parts of the shear which are triggered only by the hourglass mode of deformation. The assumed shear strain field, $\varepsilon_{\mathrm{xy}}$, is the projection of the strain field emanating from the displacement field onto the line of constant shear strain fields, as shown in Fig. 11.


Figure 11. Projection of higher order shear terms in assumed strain elements
Table 2 lists the arbitrary constants for Eq. (8) for the assumed strain elements considered in this paper. Note that the fully integrated QUAD4 element can be obtained by stabilization with one point quadrature for linear materials. It can be shown that ASMD stabilization is identical to the mean dilatation approach of Nagtegaal et al. (1974) for linear materials. ASQBI and ASOI are identical to the mixed method QBI and OI stabilization of Belytschko and Bachrach (1986) for rectangular elements.

Table 2. Constants to define the assumed strain field

| Element | $\mathrm{e}_{1}$ | $\mathrm{e}_{2}$ | $=:$ |
| :--- | :--- | :--- | :--- |
| QUAD | 1 | 0 | 1 |
| ASMD | $\frac{1}{2}$ | $-\frac{1}{2}$ | 1 |
| ASQBI | 1 | $-\bar{v}$ | 0 |
| ASOI | 1 | -1 | 0 |
| ADS | $\frac{1}{2}$ | $-\frac{1}{2}$ | 0 |

1.5.4 Stiffness Matrices for Assumed Strain Elements. The stiffness matrix for all of the assumed strain elements can be obtained by (16). If we take advantage of (1.4.16), then

$$
\begin{equation*}
\mathbf{K}_{\mathrm{e}}=\mathbf{K}_{\mathrm{e}}^{1 \mathrm{pt}}+\mathbf{K}_{\mathrm{e}}^{\mathrm{stab}} \tag{1.5.16}
\end{equation*}
$$

where $\mathbf{K}_{\mathrm{e}}^{1}{ }^{\mathrm{pt}}$ is the stiffness obtained by one-point quadrature with the quadrature point $\xi=\eta=0$, and $\mathbf{K}_{\mathrm{e}}^{\text {stab }}$ is the rank 2 stabilization stiffness, which is given by

$$
\mathbf{K}_{\mathrm{e}}^{\text {stab }}=2 \mu\left[\begin{array}{cc}
\left(\mathrm{c}_{1} \mathrm{H}_{\mathrm{xx}}+\mathrm{c}_{2} \mathrm{H}_{\mathrm{yy}}\right) \gamma \gamma^{\mathrm{t}} & \mathrm{c}_{3} \mathrm{H}_{\mathrm{xy}} \gamma \gamma^{\mathrm{t}}  \tag{1.5.17}\\
\mathrm{c}_{3} \mathrm{H}_{\mathrm{xy}} \gamma \gamma^{t} & \left(\mathrm{c}_{1} \mathrm{H}_{\mathrm{yy}}+\mathrm{c}_{2} \mathrm{H}_{\mathrm{xx}}\right) \gamma \gamma^{\mathrm{t}}
\end{array}\right]
$$

where the constants $c_{1}, c_{2}$, and $c_{3}$ are given by Table 3 . Constants are given not only for the elements listed in Table 2, but also for the ASSRI stabilization which behaves like the

SRI element of Hughes (1987) with elastic material. SRI stabilization cannot be derived by the assumed strain approach. It is obvious that the plane strain QUAD4 element will lock for nearly incompressible materials since $c_{1}$ and $c_{3}$ get very large. The projection to eliminate excessive shear stiffness corresponds to $c_{2}=0$. When both projections are made, then $\mathrm{c}_{1}=-\mathrm{c}_{3}$.

Table 3. Constants for assumed strain stabilization

| Element | $\mathrm{c}_{1}$ | $\mathrm{c}_{2}$ | $\mathrm{c}_{3}$ |
| :---: | :---: | :---: | :---: |
| QUAD4 (plane strain) | 1-v | 1/2 | 1 |
|  | 1-2v |  | 2(1-2v) |
| QUAD4 (plane stress) | 1 | 1/2 | $1+v$ |
|  | 1-v |  | 2(1-v) |
| ASSRI | 1 | 1/2 | 1/2 |
| ASMD | 1/2 | 1/2 | 0 |
| ASQBI | $1+\bar{v}$ | 0 | $-\bar{v}(1+\bar{v})$ |
| ASOI | 2 | 0 | -2 |
| ADS | 1/2 | 0 | -1/2 |

1.5.5 Nonlinear Hourglass Control. The nonlinear counterpart of the Simo-Hughes (1986) principle has been given by Fish and Belytschko (1988) as the following weak form:

$$
\begin{equation*}
0=\delta \Pi=\int_{\Omega_{\mathrm{e}}} \delta \dot{e}^{\mathrm{t}} \mathbf{s}(\stackrel{\circ}{\mathbf{e}}, \mathbf{s}, \ldots) \mathrm{d} \Omega+\delta \int_{\Omega_{\mathrm{e}}} \mathbf{t}^{\mathrm{t}}(\mathcal{D} \mathbf{v}-\stackrel{\circ}{\mathbf{e}}) \mathrm{d} \Omega-\delta \mathbf{v}^{\mathrm{t}} \mathbf{f}^{\mathrm{ext}} \tag{1.5.18}
\end{equation*}
$$

where $\boldsymbol{e}$ is the interpolated velocity strain (rate-of-deformation), $\mathbf{s}$ the Cauchy stress which is computed from the velocity strain and other state variables by the constitutive equation, $\mathbf{t}$ the interpolated Cauchy stress, and $\mathcal{D} \mathbf{v}$ is the symmetric part of the velocity gradient; the latter would be equivalent to the rate-of-deformation in a standard displacement method, but in mixed methods, the velocity gradient is projected on a smaller space to avoid locking. Note that $\mathbf{s}$ was the symbol for the interpolated stress in Section 1.4, but has a new meaning here. The superposed circle on the symbol for the rate of deformation, $\stackrel{\circ}{\mathbf{e}}$ does not indicate a time derivative. The velocity and strain-rate (rate-of-deformation) are interpolated by

$$
\begin{align*}
& \mathbf{v}=\sum_{\mathrm{I}=1}^{\mathrm{n}_{\mathrm{N}}^{\mathrm{e}}} \mathbf{N}_{\mathrm{I}}(\xi, \eta) \mathbf{v}_{\mathrm{I}} \equiv \mathbf{N} \mathbf{v}  \tag{1.5.19}\\
& \stackrel{\circ}{\mathbf{e}}=\sum_{\mathrm{I}=1}^{\mathrm{n}_{\mathrm{N}}^{\mathrm{e}}} \overline{\mathbf{B}}_{\mathrm{I}}(\xi, \eta) \mathbf{v}_{\mathrm{I}} \equiv \overline{\mathbf{B}} \mathbf{v} \tag{1.5.20}
\end{align*}
$$

where $\mathrm{n}_{\mathrm{N}}^{\mathrm{e}}$ is the number of nodes per element. In addition, we define the standard $\mathbf{B}$ matrix by

$$
\begin{equation*}
\mathcal{D} \mathbf{v}=\sum_{\mathrm{I}=1}^{\mathrm{n}_{\mathrm{N}}^{\mathrm{e}}} \mathcal{D}\left(\mathbf{N}_{\mathrm{I}} \mathbf{v}_{\mathrm{I}}\right) \equiv \mathbf{B} \mathbf{v} \tag{1.5.21}
\end{equation*}
$$

Substituting Eqs. (20) into (18) and using the orthogonality condition for $\mathbf{t}$ as before gives

$$
\begin{equation*}
0=\delta \Pi=\delta \mathbf{v}^{\mathrm{T}} \int_{\Omega} \overline{\mathbf{B}}^{\mathrm{T}} \mathbf{s} \mathrm{~d} \Omega-\delta \mathbf{v}^{\mathrm{T}} \mathbf{f}^{\mathrm{ext}} \tag{1.5.22}
\end{equation*}
$$

Exploiting the arbitrariness of $\delta \mathbf{v}$ we obtain the discrete equilibrium equations

$$
\begin{align*}
& \mathbf{f}^{\text {int }}-\mathbf{f}^{\mathrm{ext}}=0  \tag{1.5.23}\\
& \mathbf{f}^{\mathrm{int}}=\int_{\Omega} \overline{\mathbf{B}}^{\mathrm{t}} \mathbf{s d} \Omega \tag{1.5.24}
\end{align*}
$$

where the stress is given by some nonlinear constitutive equation

$$
\begin{equation*}
\mathbf{s}=\mathbf{S}(\dot{\mathbf{e}}, \mathbf{s}, \ldots)=\mathbf{S}(\overline{\mathbf{B}} \mathbf{v}, \mathbf{s}, \ldots) \tag{1.5.25}
\end{equation*}
$$

The above formulation is applicable to problems with both material and geometric nonlinearities. In applying the assumed strain stabilization procedure, it is convenient to use a corotational formulation as discussed in Section 1.4.7, where the Cauchy stresses and velocity strains are expressed in terms of a coordinate system ( $\hat{x}, \hat{y}$ ) which rotates with the element. As with mixed method stabilization of Section 1.4, a corotational coordinate system also assure that the element is frame invariant.

The internal forces in a corotational formulation are given by

$$
\begin{equation*}
\hat{\mathbf{f}}^{\mathrm{int}}=\int_{\Omega} \hat{\overline{\mathbf{B}}}^{\mathrm{t}} \hat{\mathbf{s}} \mathrm{~d} \Omega \tag{1.5.26}
\end{equation*}
$$

where the superposed tildes indicate quantities expressed in terms of the corotational coordinates. The counterpart of (20) is

$$
\begin{equation*}
\hat{\hat{\mathbf{e}}}=\hat{\overline{\mathbf{B}}} \widehat{\mathbf{v}} \tag{1.5.27}
\end{equation*}
$$

and the rate form of the constitutive equation can be written

$$
\begin{equation*}
\dot{\hat{\mathbf{s}}}=\widehat{\mathbf{C}} \hat{\hat{\mathbf{e}}} \tag{1.5.28}
\end{equation*}
$$

where $\widehat{\mathbf{C}}$ is a matrix which depends on the stress and other state variables; for an incrementally isotropic hypoelastic material, $\widehat{\mathbf{C}}$ is given by (1.4.40b).

The above form of a stress-strain law is objective (frame-invariant). The spin is then given by

$$
\begin{equation*}
\omega=\frac{1}{2}\left(\frac{\partial \widehat{\mathrm{v}}_{\mathrm{y}}}{\partial \widehat{\mathrm{x}}}-\frac{\partial \widehat{\mathrm{v}}_{\mathrm{x}}}{\partial \widehat{\mathrm{y}}}\right) \tag{1.5.29}
\end{equation*}
$$

In developing the hourglass resistance based on physical parameters, two assumptions must be made:

1. the spin is constant within the element
2. the material response tensor $\widehat{\mathbf{C}}$ is constant within the element.

The velocity for the 4 node quadrilateral is given by a form identical to (1.4.9b)

$$
\begin{equation*}
v_{i}=\left(\mathbf{s}^{\mathrm{t}}+\mathrm{x} \hat{\mathbf{b}}_{\mathrm{x}}^{\mathrm{t}}+\mathrm{y} \hat{\mathbf{b}}_{\mathrm{y}}^{\mathrm{t}}+\mathrm{h} \hat{\mathbf{g}}^{\mathrm{t}}\right) \widehat{\mathbf{v}}_{\mathrm{i}} \tag{1.5.30}
\end{equation*}
$$

The spin (29) is then given by

$$
\begin{align*}
& \left.\omega=\frac{1}{2}\left(\hat{\mathbf{b}}_{\mathrm{x}}^{\mathrm{t}} \widehat{\mathbf{v}}_{\mathrm{y}}-\hat{\mathbf{b}}_{\mathrm{y}}^{\mathrm{t}} \widehat{\mathbf{v}}_{\mathrm{x}}+\dot{\hat{\mathrm{q}}} \mathrm{y}^{\hat{\mathrm{h}}_{\mathrm{x}}}-\dot{\hat{\mathrm{q}}_{\mathrm{x}}} \hat{\mathrm{~h}}_{\mathrm{y}}\right)\right)  \tag{1.5.31}\\
& \dot{\hat{\mathrm{q}}}_{\mathrm{x}}=\hat{\mathbf{g}}^{\mathrm{t}} \widehat{\mathrm{v}}_{\mathrm{x}} \quad \dot{\hat{\mathrm{q}}}_{\mathrm{y}}=\hat{\mathbf{g}}^{\mathrm{t}} \hat{\mathbf{v}}_{\mathrm{y}} \tag{1.5.32}
\end{align*}
$$

Because of the orthogonality property (1.4.16), the average spin is given by

$$
\begin{equation*}
\omega^{\mathrm{o}}=\frac{1}{\mathrm{~A}} \int_{\Omega_{\mathrm{e}}} \omega \mathrm{~d} \Omega=\frac{1}{2}\left(\hat{\mathbf{b}}_{\mathrm{x}}^{\mathrm{t}} \widehat{\mathbf{v}}_{\mathrm{y}}-\widehat{\mathbf{b}}_{\mathrm{y}}^{\mathrm{t}} \widehat{\mathbf{v}}_{\mathrm{x}}\right) \tag{1.5.33}
\end{equation*}
$$

This corresponds to the spin at the center of the element. It can be seen from (31) that the stronger the hourglass mode, the more assumption 1 is violated.

To illustrate the remainder of the development, the special case, $e_{2}=-e_{1}, e_{3}=0$ is considered. The corotational components of the velocity strain are then given by the counterpart of (8).

For an anisotropic material, the stress rate is then given by

It can be seen that the corotational stress rate always has the same distribution within the element, so the stress also has the same form; at any point in time, $\hat{\mathbf{s}}^{0}$ is the constant part of the element stress field evaluated at the quadrature point, and $\hat{\mathbf{s}}^{1}$ is the nonconstant part.

Taking advantage of this form of the stress field, and inserting (34) and (35) into (26), and taking advantage of the orthogonality properties of $\widehat{\mathrm{h}}_{\mathrm{x}}$ and $\widehat{\mathrm{h}}_{\mathrm{y}}$ (1.4.16) and the fact that $\mathbf{C}$ is constant in the element, gives

$$
\begin{equation*}
\hat{\mathbf{f}}^{\mathrm{int}}=\mathrm{A} \widehat{\mathbf{B}}^{\mathrm{o}} \hat{\mathbf{s}}^{\mathrm{o}}+\hat{\mathbf{f}}^{\mathrm{stab}} \tag{1.5.36}
\end{equation*}
$$

where $\hat{\mathbf{f}}^{\text {stab }}$ are the hourglass (stabilization) nodal forces, which are given by

$$
\hat{\mathbf{f}}^{\text {stab }}=\left\{\begin{array}{l}
\hat{\mathrm{Q}}_{\mathrm{x}} \hat{\mathbf{g}}  \tag{1.5.37}\\
\hat{\mathrm{Q}}_{\mathrm{y}} \hat{\mathbf{g}}
\end{array}\right\}
$$

where

$$
\left\{\begin{array}{l}
\dot{\hat{Q}}_{x}  \tag{1.5.38}\\
\dot{\hat{Q}}_{y}
\end{array}\right\}=\mathrm{e}_{1}^{2}\left(\widehat{\mathrm{C}}_{11}-\widehat{\mathrm{C}}_{12}-\widehat{\mathrm{C}}_{21}+\widehat{\mathrm{C}}_{22}\right)\left\{\begin{array}{l}
\widehat{\mathrm{H}}_{\mathrm{xx}} \dot{\hat{\mathrm{q}}}_{\mathrm{x}}-\widehat{\mathrm{H}}_{\mathrm{xy}} \dot{\hat{\mathrm{q}}}_{\mathrm{y}} \\
\widehat{\mathrm{H}}_{\mathrm{yy}} \dot{\hat{\mathrm{q}}}_{\mathrm{y}}-\widehat{\mathrm{H}}_{\mathrm{xy}} \dot{\hat{\mathrm{q}}}_{\mathrm{x}}
\end{array}\right\}
$$

and $\widehat{\mathbf{B}}^{\text {o }}$ is the constant part of $\widehat{\mathbf{B}}$ which is given by

$$
\begin{align*}
& \widehat{\mathbf{B}}^{\mathrm{o}}=\left[\begin{array}{cc}
\hat{\mathbf{b}}_{\mathrm{x}}^{\mathrm{t}} & \mathbf{0} \\
\mathbf{0} & \widehat{\mathbf{b}}_{\mathrm{y}}^{\mathrm{t}} \\
\hat{\mathbf{b}}_{\mathrm{y}}^{\mathrm{t}} & \widehat{\mathbf{b}}_{\mathrm{x}}^{\mathrm{t}}
\end{array}\right]  \tag{1.5.39a}\\
& \widehat{\mathbf{b}}_{\mathrm{x}}^{\mathrm{t}}=\frac{1}{2 \mathrm{~A}}\left[\hat{\mathrm{y}}_{24}, \hat{\mathrm{y}}_{31}, \hat{\mathrm{y}}_{42}, \hat{\mathrm{y}}_{13}\right] \quad \hat{\mathbf{b}}_{\mathrm{y}}^{\mathrm{t}}=\frac{1}{2 \mathrm{~A}}\left[\hat{\mathrm{x}}_{42}, \hat{\mathrm{x}}_{13}, \hat{\mathrm{x}}_{24}, \hat{\mathrm{x}}_{31}\right] \tag{1.5.39b}
\end{align*}
$$

The nodal force vector is arranged by components:

$$
\begin{equation*}
\mathbf{f}^{\mathrm{t}}=\left[\mathrm{f}_{\mathrm{x}}^{\mathrm{t}}, \mathbf{f}_{\mathrm{y}}^{\mathrm{t}}\right] \quad \mathbf{f}_{\mathrm{x}}^{\mathrm{t}}=\left[\mathrm{f}_{\mathrm{x} 1}, \mathrm{f}_{\mathrm{x} 2}, \mathrm{f}_{\mathrm{x} 3}, \mathrm{f}_{\mathrm{x} 4}\right] \quad \mathbf{f}_{\mathrm{y}}^{\mathrm{t}}=\left[\mathrm{f}_{\mathrm{y} 1}, \mathrm{f}_{\mathrm{y} 2}, \mathrm{f}_{\mathrm{y} 3}, \mathrm{f}_{\mathrm{y} 4}\right] \tag{1.5.40}
\end{equation*}
$$

For an isotropic material, $\dot{\hat{\mathrm{Q}}}_{\mathrm{x}}$ and $\dot{\hat{\mathrm{Q}}}_{\mathrm{y}}$ can be written in terms of the constants given in Table 3.

$$
\left\{\begin{array}{l}
\dot{\hat{Q}}_{\mathrm{x}}  \tag{1.5.41}\\
\dot{\hat{\mathrm{Q}}}_{\mathrm{y}}
\end{array}\right\}=2 \mu\left\{\begin{array}{l}
\left(\mathrm{c}_{1} \widehat{\mathrm{H}}_{\mathrm{xx}}+\mathrm{c}_{2} \widehat{\mathrm{H}}_{\mathrm{yy}}\right) \dot{\hat{\mathrm{q}}}_{\mathrm{x}}+\mathrm{c}_{3} \widehat{\mathrm{H}}_{\mathrm{xy}} \dot{\hat{\mathrm{q}}}_{\mathrm{y}} \\
\left(\mathrm{c}_{1} \widehat{\mathrm{H}}_{\mathrm{yy}}+\mathrm{c}_{2} \widehat{\mathrm{H}}_{\mathrm{xx}}\right) \dot{\mathrm{q}}_{\mathrm{y}}+\mathrm{c}_{3} \widehat{\mathrm{H}}_{\mathrm{xy}} \dot{\mathrm{q}}_{\mathrm{x}}
\end{array}\right\}
$$

As with mixed method stabilization, the shear modulus in a nonlinear isotropic process is given by (1.4.9a).

Table 4 is a flowchart outlining the procedure to evaluate nodal forces in an explicit program with time step $\Delta \mathrm{t}$. Implementation in a static program simply requires the replacement of the products of rates and $\Delta \mathrm{t}$ by an increment in the corresponding integral; for example, $\Delta \hat{\mathbf{s}}$ replaces $\dot{\hat{\mathbf{s}}} \Delta \mathrm{t}$.

Table 4. Element nodal force calculation

1. update corotational coordinate system
2. transform nodal velocities $\mathbf{v}$ and coordinates $\mathbf{x}$ to corotational coordinate system
3. compute strain-rate at quadrature the point by $\dot{\hat{\mathbf{e}}}=\widehat{\mathbf{B}}^{\mathrm{o}} \hat{\mathbf{v}}$ (Eq. (39) gives $\widehat{\mathbf{B}}^{\mathrm{o}}$ )
4. compute stress-rate by constitutive law and update stress (note: $\dot{\mathbf{s}}=\Delta \hat{\mathbf{s}} / \Delta \mathrm{t}$ )
5. compute generalized hourglass strain rates by Eq. (32)
6. compute the generalized hourglass stresses rates by (38) and update the generalized hourglass stresses
7. compute $\hat{\mathbf{f}}^{\text {int }}$ by Eqs. (36) and (37)
8. transform $\hat{\mathbf{f}}^{\text {int }}$ to global system and assemble

Remark 3.1 The stress rate in (36) corresponds to the Green-Naghdi rate if the corotational coordinate system is rotated by $\omega \Delta \mathrm{t}$ in each time step.
Remark 3.2 If the Jaumann rate is used in conjunction with a fixed coordinate system, the stress field loses the form of (35) and other approximations are needed.
Remark 3.3 Because of the assumption of a constant spin and material response in the element, deviations from this assumption are directly proportional to the strength of the hourglass modes (see for example (31-33)); thus in h-adaptive methods, it is advantageous to refine by fission those elements which exhibit substantial hourglass energy, as advocated in Belytschko, Wong, and Plaskacz (1989).
Remark 3.4 If a Jaumann rate is used in a fixed coordinate system, the stress field does not maintain the distribution (35) This is one reason that the corotational form is preferred.
1.5.6 Assumed Strain with Multiple Integration Points. In the development above, stabilization forces are obtained for a reduced one-point integration element. One-point integration was chosen because it is usually advantageous to keep the number of stress evaluations to a minimum; however, there is a correlation between the number of integration points needed in a mesh and the nonlinearity of the stress field. An example of this is the dynamic cantilever beam of Section 1.6.3. For elastic material, a very accurate solution can be obtained with only one element through the depth of the beam, because the axial stress varies linearly through the depth. For elastic-plastic material, many elements are need through the depth to obtain a reasonably accurate solution, because the axial stress varies nonlinearly through the depth. The number of integration points can be increased by refining the mesh, or by increasing the number of integration points in each element. The latter method has the advantage of being able to increase the number of quadrature points without reducing the stable time step of an explicit method.

The assumed strain fields developed above can be used with any number of integration points without encountering locking since the strain fields have zero dilatational strain throughout the element domain for incompressible material. The element force vector for multi-point integration using an assumed strain field is analogous to (1.2.31b) and is given by

$$
\begin{equation*}
\hat{\mathbf{f}}_{\mathrm{e}}^{\mathrm{int}}=\sum_{\alpha=1}^{\mathrm{n}_{\mathrm{Q}}} \mathrm{w}_{\alpha} \mathrm{J}\left(\mathbf{x}_{\alpha}\right) \hat{\overline{\mathbf{B}}}^{\mathrm{t}}\left(\mathbf{x}_{\alpha}\right) \hat{\mathbf{s}}\left(\mathbf{x}_{\alpha}\right) \tag{1.5.42}
\end{equation*}
$$

where $\hat{\overline{\mathbf{B}}}\left(\mathbf{x}_{\alpha}\right)$ and $\hat{\mathbf{s}}\left(\mathbf{x}_{\alpha}\right)$ are the corotational counterparts of (8) and (25) evaluated at a quadrature point, $\mathbf{x}_{\alpha}$. Stabilization forces, may or may not be necessary with (42) depending on the location of the integration points.

The $\mathbf{g}$ terms in (8) assure rank sufficiency, as long as $\widehat{\mathrm{h}}, \mathrm{x}$ and $\widehat{\mathrm{h}}, \mathrm{y}$ are not too small. If we consider the rectangular element in Fig. 12 with a corotational coordinate system, the referential axes are parallel to the corotational axes, so

$$
\begin{array}{lll}
\hat{\xi}_{, x}=\frac{1}{a} & \hat{\eta}_{, y}=\frac{1}{b} \quad \hat{\eta}_{, x}=\hat{\xi}_{, y}=0  \tag{1.5.43a}\\
\hat{\mathrm{~h}}_{\mathrm{x}}=\frac{1}{\mathrm{a}} \eta & \hat{\mathrm{~h}}_{\mathrm{y}}=\frac{1}{\mathrm{~b}} \xi & (1.5 .43 \mathrm{~b})
\end{array}
$$

From (43b), it is apparent that $\widehat{h}_{\mathrm{x}_{\mathrm{x}}}=0$ along the $\eta$ axis and $\widehat{\mathrm{h}}, \mathrm{y}=0$ along the $\xi$ axis. Therefore if the integration points are all located on one of the referential axes, stabilization forces will be needed in either the $\widehat{x}$ or $\hat{y}$ directions to maintain rank sufficiency.


Figure 12. A rectangular element in the corotational coordinate system
Full $2 \times 2$ integration using Eq. (42) is rank sufficient, but nearly the same results are obtained with two integration points using a modified form of Eq. (42) given by

$$
\begin{equation*}
\hat{\mathbf{f}}_{\mathrm{e}}^{\mathrm{int}}=2 \mathrm{~J}(\mathbf{0}) \sum_{\alpha=1}^{2} \hat{\overline{\mathbf{B}}}^{\mathrm{t}}\left(\mathbf{x}_{\alpha}\right) \hat{\mathbf{s}}\left(\mathbf{x}_{\alpha}\right) \tag{1.5.44}
\end{equation*}
$$

In Eq. (44), $\mathbf{J}(\mathbf{0})$ is the Jacobian evaluated at the origin of the referential coordinate system, and the two integration points are either $\mathbf{x}_{1}=(-1 / \sqrt{3},-1 / \sqrt{3}), \mathbf{x}_{2}=(+1 / \sqrt{3}$, $+1 / \sqrt{3})$, or else $\mathbf{x}_{1}=(-1 / \sqrt{3},+1 / \sqrt{3}), \mathbf{x}_{2}=(+1 / \sqrt{3},-1 / \sqrt{3})$. The choice of the pair of integration points makes little difference in the solution. This 2-point integration scheme is similar to the IPS2 element reported in Liu et al. (1988). The formulation here differs by using an assumed strain field is used to improve accuracy. Using, the QBI strain field, a flexural-superconvergent 2-point element is obtained.

In Section 1.6.3, we observe that the ASQBI element with 1-point integration provides an accurate coarse mesh solution with elastic material; however, with elasticplastic material, the coarse mesh solution is poor. We can therefore attribute the error in the elastic-plastic solution to an insufficient number of integration points. This large error is not surprising if we consider the nature of the solution. The plastic deformation of a beam in bending initiates at the top and bottom surfaces of a beam where the axial stress is greatest. With 1-point integration, the only stress evaluation is at the center of the element, so while the stress state at the integration point remains within the yield surface, the stress state may be outside the yield surface at other points in the element domain. For coarse mesh bending, the error is large, resulting in too little plastic deformation.

The 2-point integration scheme of Eq. (44), and $2 \times 2$ integration by Eq. (42) improve on 1-point integration by placing integration points nearer the edge of the element. In Section 1.6.3, the effect of multiple stress evaluations is demonstrated by the solution of an elastic-plastic cantilever beam. Results for the 2 and 4-point integration schemes are given in Tables 11a through 11d. Both use the QBI strain field, so the 2 point scheme is called $\operatorname{ASQBI}(2 \mathrm{pt})$, and $2 \times 2$ integration is called $\operatorname{ASQBI}(2 \times 2)$. Both of these elements have flexural-superconvergence as does the 1-point element with ASQBI stabilization, so the elastic part of the solution is solved very accurately. Therefore, the difference in the solutions of these three elements with elastic-plastic material can be attributed to the effect of multiple stress evaluations on the nonlinear part of the solution.

### 1.6 Numerical Results

The numerical examples reported here include linear and nonlinear problems. The linear problems were studied to examine the convergence rate of various forms of these and competing elements. Table 5 gives a complete listing of the names associated with the elements tested in this section. All use 1-point integration in the nonlinear problems except for QUAD4, ASQBI (2pt), ASQBI(2x2).

Table 5. Names and descriptions of elements tested in this section

| Name | Section | Description |
| :--- | :---: | :--- |
| QUAD4 | 1.2 | Standard isoparametric element with full 2x2 integration. |
| FB (0.1) | 1.3 | Perturbation hourglass stabilization with the hourglass control <br> factor, $\alpha_{s}=0.1$. (Flanagan and Belytschko (1981)) |
| FB (0.3) | 1.3 | Same as FB (0.1), except with $\alpha_{s}=0.3$ |
| OI | 1.4 | Mixed method Optimal Incompressible stabilization (Belytschko and <br> Bachrach (1986)). |
| QBI | 1.4 | Mixed method Quintessential Bending and Incompressible <br> stabilization (Belytschko and Bachrach). |
| ASOI | 1.5 | Assumed strain stabilization using the OI strain field |
| ASQBI | 1.5 | Assumed strain stabilization using the QBI strain filed |
| ADS | 1.5 | Assumed deviatoric strain stabilization |
| ASMD | 1.5 | Assumed strain stabilization using the strain field associated with the <br> mean dilatation element (Nagtegaal et al.(1974)). |
| ASSRI | 1.5 | Assumed strain stabilization using the strain field associated with <br> selective reduced integration (Hughes(1980)) |
| ASQBI(2 pt) | 1.5 .6 | The QBI strain field is used with two stress evaluations per element |
| ASQBI(2x2) | 1.5 .6 | The QBI strain filed is used with four stress evaluations per element |
| Pian-Sumihara | The Pian-Sumihara (1984) hybrid element (the formulation does not <br> appear in this paper) |  |

1.6.1 Static Beam. A linear, elastic cantilever with a load at its end is shown in Fig. 13. M and P at the left end of the cantilever are reactions at the support.


Figure 13. Static cantilever beam
This problem is identical to that used by Belytschko and Bachrach (1986). The analytical solution from Timoshenko and Goodier (1970) is

$$
\begin{align*}
& u_{x}(x, y)=\frac{-P y}{6 \overline{\mathrm{E}} I}\left[(6 L-3 x) x+(2+\bar{v})\left(y^{2}-\frac{1}{4} D^{2}\right)\right]  \tag{1.5.44a}\\
& u_{y}(x, y)=\frac{P}{6 \bar{E} I}\left[3 \bar{v} y^{2}(L-x)+\frac{1}{4}(4+5 \bar{v}) D^{2} x+(3 L-x) x^{2}\right] \tag{1.5.44b}
\end{align*}
$$

where $\quad \mathrm{I}=\frac{1}{12} \mathrm{D}^{3}$

$$
\begin{align*}
& \overline{\mathrm{E}}=\left\{\begin{array}{cc}
E & \text { for plane stress } \\
E /\left(1-v^{2}\right) & \text { for plane strain }
\end{array}\right.  \tag{1.5.45a}\\
& \bar{v}=\left\{\begin{array}{cl}
v & \text { for plane stress } \\
v /(1-v) & \text { for plane strain }
\end{array}\right. \tag{1.5.45b}
\end{align*}
$$

The displacements at the support end, $\mathrm{x}=0,-\frac{1}{2} \mathrm{D} \leq \mathrm{y} \leq \frac{1}{2} \mathrm{D}$ are nonzero except at the top, bottom, and midline (as shown in Fig. 14). Reaction forces are applied at the support based on the stresses corresponding to $(1.5 .46)$ at $x=0$, which are

$$
\begin{align*}
& \sigma_{x}=-\frac{P y}{I}(L-x)  \tag{1.5.46a}\\
& \sigma_{y}=0 .  \tag{1.5.46b}\\
& \tau_{x y}=\frac{P}{2 I}\left(\frac{1}{4} D^{2}-y^{2}\right) \tag{1.5.46c}
\end{align*}
$$

The distribution of applied load to the nodes at $\mathrm{x}=\mathrm{L}$ is also obtained from the closed-form stress fields. The coarsest mesh used is shown in Fig. 14. This problem is symmetric, so only half the cantilever is modeled.


Figure 14. Coarse mesh of rectangular elements
All meshes use elements with an aspect ratio of 2 . Only the top half of the cantilever is modeled since the problem is antisymmetric. The following isotropic elastic materials were used:

1. Plane stress, $v=0.25$
2. Plane strain, $v=0.4999$

The displacement and energy error norms are plotted in Figs. 15 and 16. for $v=0.25$, the rate of convergence of the displacement error norm is around 1.8 for all of the elements except for QBI, ASQBI and Pian-Sumihara which converge at a rate of 2. All have a rate of convergence of the energy error norm of 1 . For $v=0.4999$, the rate of convergence of the displacement error norms is around 1.7 to 1.8 and the and rate of convergence of the energy error norms is 1.0 for all elements except QUAD4 which locks as expected and exhibits very slow convergence. For incompressible material, QBI and ASQBI are almost identical to OI and ASOI, whereas ASMD has less absolute accuracy. For rectangular elements and any linear material, OI and ASOI are identical. Likewise, the Pian and Sumihara (1984) element is identical to QBI and ASQBI.



| $\longrightarrow-$ | $\longrightarrow-$ | $\longrightarrow$ | $\longrightarrow-$ | $\longrightarrow$ |
| :---: | :---: | :---: | :---: | :---: |
| QUAD4 | ASMD | QBI, ASQBI, <br> and <br> Pian-Sumihara | ASOI <br> and OI | ADS |

Figure 15. Convergence of displacement and energy error norms; $v=0.25$, plane stress



| $\longrightarrow-$ | $\longrightarrow-$ | $\longrightarrow$ | $\longrightarrow-$ | $\longrightarrow-$ |
| :---: | :---: | :---: | :---: | :---: |
| QUAD4 | ASMD | QBI, ASQBI, <br> and <br> Pian-Sumihara | ASOI <br> and OI | ADS |

Figure 16. Convergence of displacement and energy error norms; $v=0.4999$, plane strain
To assess the coarse mesh accuracy of the elements, the normalized end displacements (point A if Fig. 14) for the 1 x 4 element mesh are shown in Table 6. A coarse $1 \times 4$ element mesh of skewed elements was also run and the normalized end displacements (point A in Fig. 17) are shown in Table 7. Pian-Sumihara is slightly better than ASQBI for the skewed elements, but the difference is minor.


Figure 17. Skewed coarse mesh with $\theta=9.462^{\circ}$

Table 6. $\mathrm{d}_{\mathrm{yFEM}} / \mathrm{d}_{\mathrm{yAnalytical}}$ at point A of mesh in Fig. 14 (rectangular elements)

| Material | QUAD4 | ASMD | QBI, ASQBI, and Pian-Sumihara | $\begin{aligned} & \overline{\mathrm{OI}} \\ & \mathrm{ASOI} \end{aligned}$ | and ADS |
| :---: | :---: | :---: | :---: | :---: | :---: |
| 1 | 0.708 | 0.797 | 0.986 | 0.862 | 1.155 |
| 2 | 0.061 | 0.935 | 0.982 | 0.982 | 1.205 |

Table 7. $d_{y F E M} / d_{y A n a l y t i c a l}$ at point A of mesh in Fig. 17 (skewed elements)

| Material | QUAD4 | ASMD | ASQBI | Pian- <br> Sumihara | ASOI | ADS |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 1 | 0.689 | 0.776 | 0.948 | 0.955 | 0.834 | 1.112 |
| 2 | 0.061 | 0.915 | 0.957 | 0.960 | 0.957 | 1.170 |

1.6.2 Circular Hole in Plate. This problem was considered to evaluate the performance of these elements in a different setting. A plate with a hole, solved by R. C. J. Howland (1930) is shown in Fig. 18. The solution is in the form of an infinite series and gives the stress field around the circular hole in the center of an axially loaded plane stress plate of finite width and of infinite length. The series converges only within a circular region around the hole. The diameter of this circular area is equal to the plate width. The displacement field is not given so convergence of the displacement norm could not be checked.

The shaded area indicates
the region of convergence


Figure 18. Plate of finite width with a circular hole
For the finite element meshes, the plate length was taken to be twice the plate width. The nodes at which the load is applied are outside the region in which the analytical solution converges, so the analytical solution could not be used to determine the load distribution on the end of the plate. The nodal forces were therefore calculated by assuming the analytical stress field at infinity, which is uniaxial. The error due to the finite length was checked by running meshes with lengths of 2 and 5 times the plate width. The difference between these solutions was found to be negligible. Four different meshes were used which are summarized in Table 8. Fig. 19 shows the dimensions and boundary conditions of the finite element model, and Fig. 20 shows the discretization for mesh 3 with 320 elements. The problem is symmetric, so only one fourth of the plate was modeled.

Table 8. Meshes used for Howland plate with hole problem

| Mesh number | Number of elements |  |
| :--- | :--- | :--- |
|  | Total in mesh | In portion of mesh used to <br> calculate the energy norm |
| 1 | 20 | 12 |
| 2 | 80 | 48 |
| 3 | 320 | 192 |
| 4 | 1280 | 768 |



Figure 19. Finite element model of plate with a circular hole


Figure 20. Mesh 3 discretization
The circular hole is approximated by elements with straight edges, so the hole is actually a polygon. As the number of elements is increased, the shape and area of the hole changes slightly.

Because the analytical solution only converges in a region around the hole, a subset of the total number of elements in the mesh was used to calculate the energy norm. This area, shaded in Fig. 19, was held constant as the mesh was refined, except for the change in the area of the hole.

Table 9 shows the calculated stress concentration factor at point A on Fig. 19 normalized by the analytical solution. At point $\mathrm{A}, \sigma_{\mathrm{x}}=3.0361$ according to the analytical solution. The stress concentration factor depends on both the constant and non-constant part of the stress field. None of the elements can represent exactly the nonlinear stress field in the area near the hole; however, some are better than others. The ASQBI element was shown earlier to represent the pure bending mode of deformation better than the ASOI elements. This ability seems to help also in the calculation of the stress concentrationfactor at point A. For the ASMD and ADS elements $\left(e_{1}=1 / 2\right)$, the non-constant part of the strain is only half the magnitude that of the ASOI element $\left(e_{1}=1\right)$, so the stress concentration factor is lower.

Table 9. $\sigma_{\mathrm{xFEM}} / \sigma_{\mathrm{xAnalytical}}$ at point A in Fig. 19

| Mesh | QUAD4 | ASMD | ASQBI | Pian-S | ASOI | $\overline{\text { ADS }}$ |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 1 | 0.888 | 0.721 | 0.885 | 0.778 | 0.772 | 0.733 |
| 2 | 0.973 | 0.838 | 0.961 | 0.914 | 0.874 | 0.831 |
| 3 | 0.994 | 0.900 | 0.988 | 0.971 | 0.926 | 0.902 |
| 4 | 1.000 | 0.946 | 0.997 | 0.993 | 0.963 | 0.947 |

Table 10 shows the normalized x-component of stress at the center of the element that is nearest to the point of maximum stress (point A on Fig. 19). This value is independent of the nonconstant part of the stress field, so there is much less variation between the elements. The coordinates of the element center change as the mesh is refined, so the analytical stress used to normalize the solutions is included in Table 10.

Table 10. $\sigma_{\text {xFEM }} / \sigma_{\text {xAnalytical }}$ at the center of the element nearest point A in Fig. 19

| Mesh | Analytical stress | QUAD4 | ASMD | ASQBI | Pian-S | ASOI | ADS |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 1 | 1.671 | 1.000 | 1.031 | 1.009 | 1.056 | . 982 | 1.040 |
| 2 | 2.089 | 1.010 | 1.029 | 1.013 | 1.038 | . 995 | 1.031 |
| 3 | 2.462 | 1.005 | 1.015 | 1.006 | 1.016 | . 997 | 1.012 |
| 4 | 2.717 | 1.002 | 1.007 | 1.003 | 1.007 | . 999 | 1.008 |

Fig. 21 shows the convergence of the error in the energy norm. All elements were found to have convergence rates ranging from 0.92 to 0.98 . Theoretically, the convergence rate of the energy norm should go to 1 as the element size $\mathrm{H} \rightarrow 0$. Note that the differences in the errors for the various elements are much smaller than in the beam problem. This is expected, since the nonconstant mode of deformation in this problem is much less significant than it is in bending.


| $\square$ | QUAD4 |
| :--- | :--- |
| $\square$ | ASMD |
| $\square$ | ASQBI |
| $\square$ | ASOI |
| $\square$ | ADS |
| $\square-$ | Pian-Sumihara |

Figure 21. Convergence of the error in the energy norm
1.6.3 DynamicCantilever The rate form of stabilization was implemented in the two dimensional version of WHAMS (Belytschko and Mullen (1978)). An end loaded cantilever was modeled with both elastic and elastic-plastic materials as shown in Fig. 22. A similar problem is reported in Liu et al. (1988). Two plane-strain isotropic materials were used with $v=0.25, \mathrm{E}=1 \times 10^{4}$, and the material density, $\rho=1$.
(1) elastic
(2) elastic-plastic with 1 plastic segment $\left(\sigma_{y}=300 ; \mathrm{E}_{\mathrm{t}}=0.01 \mathrm{E}\right)$
where $\sigma_{\mathrm{y}}$ is the yield stress, $\mathrm{E}_{\mathrm{t}}$ is the plastic hardening modulus; a Mises yield surface and isotropic hardening were used.


Figure 22. Dynamic cantilever beam
Ten meshes were considered. Six of them are composed of rectangular elements, while the other four are skewed. A coarse mesh called the $1 \times 6$ mesh has one element through the beam depth and 6 along the length. The aspect ratio of these elements is nearly 1. Meshes of $2 \times 12,4 \times 24$, and $8 x 48$ elements are generated from the $1 \times 6$ mesh by
subsequent divisions of each element into 4 smaller elements. Two meshes of elongated elements, $2 \times 6$ (E) and $4 \times 12$ (E) were made of elements with aspect ratios of slightly more than 2. Finally four meshes are made up of skewed elements. Two of them, 2x12(S) and $4 \times 24(S)$, are formed by skewing $2 \times 12$ and $4 \times 24$; the other two, $2 \times 6(E S)$ and $4 \times 12(E S)$, are formed by skewing $2 \times 6(E)$ and $4 x 12(E)$. Figures ( $23 a-g$ ) show 7 of the meshes.


Figure 23a. 1x6 mesh

Figure 23b. 4x24 mesh


Figure 23c. $4 \times 12$ (E) mesh


Figure 23d. 2x12(S) mesh


Figure 23e. $4 \times 24(\mathrm{~S})$ mesh


Figure 23f. 2 x 6 (ES) mesh


Figure 23g. 4x12(ES) mesh
The problem involves very large displacement (of order one third the length of the beam). No analytical solutions is available, so the results are not normalized; however, a more refined meshes of $32 \times 192$ elements were run using a 1 -point element with ADS stabilization in an attempt to find a converged solution. The end displacements at point A in Fig. 23(a) are listed in Tables 11a through 11d. Fig. 24 is a typical deformed mesh which shows the large strain and rotation that occurs. Figs. (25a-e) are time plots of the ycomponent of the displacement at the end of the cantilever. The first three demonstrate the convergence of the elastic-plastic solution with mesh refinement for ASQBI and ADS stabilization, and for the $\operatorname{ASQBI}(2 \mathrm{pt})$ element. These plots also include the elastic solution and the $32 \times 192$ element elastic-plastic solution using ADS stabilization for comparison. The last two time plots each show a solution of a single mesh by ADS and ASQBI stabilization, and the ASQBI (2pt) and ASQBI (2x2) elements. These plots also include the elastic and $32 \times 192$ element solution for comparison.

Table 12 lists the percentage of the strain energy that is associated with the hourglass mode of deformation at the time of maximum end displacement for some of the runs with elastic-plastic material. As expected, nearly all the strain energy is in the hourglass mode for the coarse ( $1 \times 6$ ) mesh. As the mesh is refined, the percentage of strain energy in the hourglass mode decreases rapidly, so the importance of accurately calculating the hourglass strains also decreases.


Figure 24. Deformed $4 \times 24$ mesh showing maximum end displacement (elastic-plastic material)

With all of the elements, the onset of plastic deformation is significantly retarded when the mesh is too coarse. This is most evident in the QBI elements which are flexuralsuperconvergent for elastic material. The ADS or FB (0.1) elastic solutions are too flexible, which tends to mask the error caused by too few integration points. The only sure way to reduce the error in solutions that involve elastic-plastic bending is to increase the number of integration points. This can be accomplished by mesh refinement or by using multiple integration points in each element, as with the 2 point and $2 \times 2$ integration. If the mesh is refined, not only are the number of integration points increased, but the amount of strain energy that is in the hourglass mode of deformation decreases (Table 12), so the accuracy of the coarse mesh solution becomes less relevant. When multiple integration points are used, the energy in the nonconstant modes of deformation remains significant, so an accurate strain field such as ASQBI is more important.

With two and four stress evaluations per element respectively, $\operatorname{ASQBI}(2 \mathrm{pt})$ and $\operatorname{ASQBI}(2 \mathrm{x} 2)$ give similar results to ADS stabilization when the mesh is refined to $8 \times 48$ elements. These elements are also have flexural-superconvergence with elastic material. The improvement over a 1-point element with ASQBI stabilization is similar to the improvement obtained by one level of mesh refinement, and it is significantly less computationally expensive. Each level of mesh refinement slows the run by a factor of 8, while additional integration points slow it by less than 2 for ASQBI ( 2 pt ) and 4 for ASQBI ( $2 \times 2$ ). For this problem with a fairly simple constitutive relationship, the additional c.p.u time needed for an a second stress evaluation is largely offset by the elimination of the need for stabilization, so $\operatorname{ASQBI}(2 \mathrm{pt})$ solutions are less than $10 \%$ slower than the stabilized 1point element.


Figure 25a. End displacement of elastic-plastic cantilever; ASQBI stabilization


Figure 25b. End displacement of elastic-plastic cantilever; ASQBI (2 pt) element


Figure 25c. End displacement for elastic-plastic cantilever; ADS stabilization


Figure 25d. End displacement for $4 \times 12$ (E) mesh (elastic-plastic)


Time
Figure 25e. End displacement for $4 \times 24$ element mesh (elastic-plastic)

Table 11a. Maximum end displacement of elastic cantilever

| Element | $\overline{1 \times 6}$ | $: \overline{2 \times 12}$ | $\overline{4 \times 24}$ | $: \overline{8 \times 48}$ | $\overline{2 \times 6(\mathrm{E})}$ |
| :--- | :--- | :--- | :--- | :--- | :--- |
| QUAD4 (2x2) | 4.69 | 6.14 | 6.68 | 6.84 | 4.92 |
| FB (0.1) | 15.9 | 8.12 | 7.17 | 6.97 | 7.22 |
| FB (0.3) | 7.68 | 7.04 | 6.93 | 6.91 | 5.35 |
| OI | 4.78 | 6.17 | 6.70 | 6.85 | 6.11 |
| ASOI | 4.78 | 6.17 | 6.70 | 6.85 | 6.11 |
| QBI | 6.89 | 6.86 | 6.88 | 6.90 | 6.66 |
| ASQBI | 16.89 | 6.86 | 6.88 | 6.90 | 6.79 |
| ASQBI (2x2) | 16.89 | 6.86 | 6.88 | 6.90 | 6.79 |
| ASQBI (2 pt) | 16.89 | 6.85 | 6.88 | 6.90 | 6.86 |
| ADS | 14.2 | 7.95 | 7.13 | 6.96 | 7.87 |
| ASMD | 8.49 | 7.20 | 6.97 | 6.92 | 5.86 |
| ASSRI | 6.05 | 6.63 | 6.82 | 6.88 | 5.23 |

Table 11b. Maximum end displacement of elastic cantilever for the meshes of skewed elements; solutions are normalized by the solutions from Table 11a for the corresponding meshes of rectangular elements

| Element | $\overline{2 \times 12(S)}$ | $=\overline{4 \times 24(\mathrm{~S})}$ | $\overline{2 \times 6(\mathrm{ES})}$ | $\overline{4 \times 12(\mathrm{ES})}$ |
| :--- | :--- | :--- | :--- | :--- |
| QUAD4 (2X2) | 0.99 | 0.99 | 0.97 | 0.98 |
| FB (0.1) | 1.01 | 1.00 | 0.99 | 0.99 |
| FB (0.3) | 0.99 | 1.00 | 0.99 | 0.99 |
| OI | 0.99 | 0.99 | 0.97 | 0.98 |
| ASOI | 1.00 | 1.00 | 0.98 | 0.99 |
| QBI | 0.99 | 0.99 | 0.97 | 0.98 |
| ASQBI | 0.99 | 0.99 | 0.97 | 0.98 |
| ASQBI (2x2) | 0.99 | 0.99 | 0.97 | 0.98 |
| ASQBI (2 pt) | 0.99 | 0.99 | 0.96 | 0.98 |
| ADS | 1.00 | 1.00 | 0.99 | 0.99 |
| ASMD | 1.00 | 1.00 | 0.99 | 0.99 |
| ASSRI | 0.99 | 1.00 | 0.99 | 0.99 |

Table 11c. Maximum end displacement and residual displacement (in parentheses) of elastic-plastic cantilever; a solution by ADS stabilization with a $32 \times 192$ element mesh gives a maximum displacement of 8.17, and a residual displacement of 5.24.

| Element | 1x6 | 2x12 | $4 \times 24$ | $8 \times 48$ | 2x6(E) | 4x12(E) |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| QUAD4 (2x2) | 4.69 | 6.30 | 7.31 | 7.85 | 4.94 | 6.61 |
|  | (0.11) | (1.79) | (3.69) | (4.65) | (0.78) | (2.76) |
|  | 15.9 | 8.39 | 8.18 | 8.14 | 7.22 | 7.67 |
| FB (0.1) | (0.00) | (3.40) | (4.88) | (5.04) | (1.05) | (3.82) |
|  | 7.68 | 7.05 | 7.59 | 7.92 | 5.35 | 6.69 |
| FB (0.3) | (0.12) | (1.15) | (3.74) | (4.67) | (0.13) | (2.41) |
|  | 4.78 | 6.17 | 7.17 | 7.76 | 6.11 | 7.00 |
| OI | (0.05) | (0.20) | (3.13) | (4.41) | (0.16) | (2.63) |
|  | 4.78 | 6.17 | 7.17 | 7.76 | 6.11 | 7.00 |
| ASOI | (0.05) | (0.20) | (3.16) | (4.40) | (0.16) | (2.63) |
|  | 6.89 | 6.86 | 7.53 | 7.90 | 6.79 | 7.34 |
| QBI | (0.11) | (0.89) | (3.69) | (4.64) | (0.34) | (3.16) |
|  | 6.89 | 6.86 | 7.54 | 7.90 | 6.79 | 7.34 |
| ASQBI | (0.11) | (0.87) | (3.72) | (4.64) | (0.34) | (3.16) |
|  | 6.98 | 7.52 | 7.86 | 8.05 | 7.27 | 7.68 |
| ASQBI (2x2) | (1.79) | (3.62) | (4.53) | (4.99) | (3.10) | (4.17) |
|  | 7.00 | 7.53 | 7.87 | 8.06 | 7.28 | 7.69 |
| ASQBI (2 pt) | (1.75) | (3.54) | (4.57) | (5.01) | (3.14) | (4.21) |
|  | 14.2 | 8.15 | 8.12 | 8.12 | 7.94 | 7.94 |
| ADS | (0.00) | (3.03) | (4.77) | (5.01) | (1.89) | (4.19) |
|  | 8.49 | 7.21 | 7.73 | 7.97 | 5.59 | 6.83 |
| ASMD | (0.13) | (1.38) | (4.05) | (4.77) | (0.14) | (2.58) |
|  | 6.05 | $6.63$ | $7.42$ | $7.86$ | 5.23 | $6.60$ |
| ASSRI | (0.09) | (0.60) | (3.54) | (4.57) | (0.12) | (2.21) |

Table 11d. Maximum end displacement and residual end displacement (in parentheses) of elastic-plastic cantilever for the meshes of skewed elements; solutions are normalized by the solutions from Table 11c for the corresponding meshes of rectangular elements

| Element | 2x12(S) | 4x24(S) | 2x6(ES) | 4x12(ES) |
| :---: | :---: | :---: | :---: | :---: |
| QUAD4 (2x2) | 1.08 | 0.98 | 0.98 | 0.98 |
|  | (0.62) | (0.96) | (1.21) | (1.02) |
|  | 1.04 | 0.99 | 1.02 | 0.99 |
| FB (0.1) | (1.18) | (0.99) | (1.78) | (1.05) |
|  | 1.00 | 0.99 | 0.99 | 0.99 |
| FB (0.3) | (1.23) | (0.99) | (2.28) | (1.04) |
|  | 0.99 | 0.98 | 0.97 | 0.98 |
| OI | (2.40) | (0.98) | (3.61) | (0.97) |
|  | 1.00 | 0.99 | 0.98 | 0.98 |
| ASOI | (2.45) | (0.98) | (3.66) | (0.96) |
|  | 0.99 | 0.98 | 0.98 | 0.98 |
| QBI | (1.21) | (0.99) | (3.07) | (0.97) |
|  | 0.99 | 0.98 | 0.98 | 0.98 |
| ASQBI | (1.28) | (0.97) | (3.12) | (0.98) |
|  | 0.98 | 0.98 | 0.98 | 0.98 |
| ASQBI (2x2) | (0.96) | (0.97) | (1.03) | (0.98) |
|  | 0.98 | 0.98 | 0.96 | 0.98 |
| ASQBI ( 2 pt ) | (1.03) | (0.97) | (0.96) | (0.97) |
|  | 1.03 | 0.99 | 1.03 | 0.99 |
| ADS | (1.17) | (0.99) | (1.48) | (1.02) |
|  | 1.00 | 0.98 | 0.99 | 0.99 |
| ASMD | (1.30) | (0.98) | (3.15) | (1.04) |
|  | 0.99 | 0.98 | 0.99 | 0.99 |
| ASSRI | (1.62) | (0.97) | (1.53) | (1.05) |

Table 12. Hourglass energy in the mesh when the end displacements maximum (normalized by total strain energy)

| Mesh | $\overline{\text { FB (0.1) }}$ | $\overline{\text { ASOI }}$ | $\overline{\text { ASQBI }}$ | $\overline{\text { ADS }}$ |
| :--- | :--- | :--- | :--- | :--- |
| $1 \times 6$ | 0.982 | 0.975 | 0.981 | 0.988 |
| 2x12 | 0.108 | 0.327 | 0.247 | 0.124 |
| $4 \times 24$ | 0.033 | 0.110 | 0.079 | 0.207 |
| $8 \times 48$ | 0.011 | 0.035 | 0.026 | $\underline{0.036}$ |

REMARK 6.1 The QUAD4 element performs no better that the stabilized one-point elements
REMARK 6.2 The value of $\alpha_{s}$ has a significant effect on the solution of bending problems using perturbation stabilization (FB) when the mesh is coarse
REMARK 6.3 Those elements that do not project out the nonconstant part of the strain field, (QUAD4, ASMD, and ASSRI) stiffen significantly more than the others when the elements are elongated as with $2 \times 6(\mathrm{E})$ and $4 \times 12(\mathrm{E})$ solutions. Perturbation stabilization $(\mathrm{FB})$ is also sensitive since it is not responsive to the element aspect ratio.

REMARK 6.4 Skewing the elements seems to have little effect on any of the elements. This may be a little deceptive since this is a large deformation problem. The elements of all the meshes skew noticeably when deformed (Fig. 24) so the initially skewed meshes only introduce additional skewing. The elastic-plastic $2 \times 6$ (ES) results are of dubious significance, since the elastic-plastic $2 \times 6$ (E) solutions are quite inaccurate.
REMARK 6.5 Another set of runs was made using an elastic-plastic material with a larger plastic modulus $\left(\mathrm{E}_{\mathrm{t}}=0.1 \mathrm{E}\right)$. The results were similar to those for $\left(\mathrm{E}_{\mathrm{t}}=0.01 \mathrm{E}\right)$ and are not shown.
1.6.4 Cylindrical Stress Wave. A two dimensional domain with a circular hole at its center was modeled with 4876 quadrilateral elements as shown in Figs 26 and 27. A compressive load with the time history shown in Fig. 28 was applied to the hole and the dynamic evolution was obtained until $t=0.09$. The domain is large enough to prevent the wave from reflecting from the outer boundary. Elastic and elastic plastic materials were used.

To provide an estimate of the error in the 2 D results, solutions were obtained for the same domain and load history using 3600 axisymmetric, 1D elements. The radial strain $\varepsilon_{\mathrm{rr}}$ for the elastic and elastic-plastic solutions at $t=0.09$ is shown in Fig. 29. The normalized $\mathrm{L}_{2}$ norms of the error in displacements at time $\mathrm{t}=0.09$ along the radial lines at $\theta=0$ and $\theta=\pi / 4$ are given in Tables 13a and 13b. All of the elements have the same magnitude of error.


Elastic material:
Young's modulus, $\mathrm{E}=1 \times 10^{6}$
Density, $\rho=1.0$

Elastic-plastic material:
Young's modulus, $\mathrm{E}=1 \times 10^{6}$
Density, $\rho=1.0$
Yield stress, $\sigma_{y}=1 \times 10^{4}$
Plastic modulus, $\mathrm{E}_{\mathrm{t}}=\mathrm{E} 16$

Figure 26. 4 node quad. mesh dimensions

Figure 27. Discretization of infinite domain with a hole

Pressure load at $r=10$ in. (ksi)


Time
Figure 28. Load history


Figure 29. Radial strain at $\mathrm{t}=0.09$

Table 13a. Normalized $L_{2}$ norms of error in displacements for material 1 (elastic)

| $\theta$ | QUAD4 | $\overline{\mathrm{FB}}$ (0.1) | ASMD | ASQBI | ASOI | ADS |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| $0^{\circ}$ | . 014 | . 014 | . 014 | . 014 | . 013 | . 014 |
| $45^{\circ}$ | . 022 | . 022 | . 019 | . 019 | . 012 | . 021 |

Table 13b. Normalized $L_{2}$ norms of error in displacements for material 2 (elastic-plastic)

| $\theta$ | QUAD4 | FB (0.1) | ASMD | ASQBI | ASOI | ADS |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| $0^{\circ}$ | . 0063 | . 0063 | . 0061 | . 0061 | . 0061 | . 0063 |
| $45^{\circ}$ | . 0069 | . 0069 | . 0086 | . 0088 | . 0073 | . 0088 |

1.6.5 Static Cantilever. The solutions to the test problems of Sections 1.6.1 and 1.6.2 were obtained using a local coordinate formulation of the stabilization matrix: Likewise, the solutions to the test problems of Sections 1.6.3 and 1.6.4 were obtained using a corotational coordinate formulation. The need for these local and corotationalformulations to obtain a frame invariant element is discussed in Section 1.4.6. The following solutions to a static cantilever demonstrate this need.

A cantilever with a shear load at its end was solved by two versions of the linear static finite element code using QBI stabilization. One version had a local coordinate formulation, and the other did not. These are called the "local" and "global" formulations respectively. A total of seven solutions were obtained with three meshes as shown in Figs. (30a-c). Each was solved with the longitudinal axis of the undeformed beam aligned with the global x axis, and also with the beam initially rotated before applying the load.


Figure 30a. 1x6 element mesh


Figure 30b. 1x3 element mesh


Figure 30c. $4 \times 12$ element mesh
Table 14 lists the end displacement in the direction of the load for the seven solutions normalized by the solutions of the unrotated meshes. Therefore, these numbers do not demonstrate absolute accuracy, but the variation in the element stiffness that occurs with rigid body rotation. The results show that the global formulation is sensitive to rigid body rotation when the elements are elongated and the mesh is coarse. When the aspect ratio 1 , both formulations are frame invariant. Also, when the mesh is refined, the lack of frame invariance is less noticeable. The local formulation is always frame invariant.

Table 14. End displacements in the direction of the applied load normalized by the $0^{\circ}$ solution

| Mesh | Initial <br> rotation <br> (degrees) | Global | Local |
| :--- | :--- | :--- | :--- |
|  | 0 | 1.00 | 1.00 |
| $1 \times 6$ | 45 | 1.00 | 1.00 |
|  | 0 | 1.00 | 1.00 |
| $1 \times 3$ | 22.5 | 0.71 | 1.00 |
|  | 45 | 0.49 | 1.00 |
|  | 0 | 1.00 | 1.00 |
| $4 \times 12$ | 45 | 0.94 | 1.00 |

### 1.7 Discussion and Conclusions

The bilinear quadrilateral element is a good choice for solving two dimensional continuum problems with explicit methods, because the mass matrix can be lumped with little loss of accuracy. There are two major benefits to 1-point integration with the quadrilateral. The first is the elimination of volumetric locking which plagues the fully integrated element. The second is a reduction in the computational effort for such elements. A drawback of 1-point integration is that spurious modes will occur if they are not stabilized. We have examined some ways of stabilizing the spurious modes in this chapter.

With all the methods considered, the stabilization forces are proportional to a $\mathbf{g}$ vector which is orthogonal to the constant strain modes of deformation, so the stabilization forces do not contribute to the constant strain field. Therefore, all have a quadratic rate of convergence in the displacement error norm. The major difference between the methods is in the way the evaluation of the magnitude of the stabilization forces.

Flanagan and Belytschko (1981) were motivated by the desire to keep the stabilization forces small so they would not interfere with the solution or cause locking. This stabilization has the drawback of requiring a user specified parameter. A bending dominated solution can depend significantly on the value of the parameter which is undesirable.

Using mixed methods, Belytschko and Bachrach (1986) chose strain and stress fields that more closely resemble the strength of materials solution of elastic deformation. Thus, they were able to use stabilization to improve to the accuracy of bending solutions. They obtain very accurate bending solutions with very few elements with elastic material. Mixed method stabilization is dependent only on material properties and element geometry; no user specified parameter is needed.

The Simo-Hughes form of the assumed strain method has also been used to develop stabilization. The assumed strain fields are motivated in the same way as the mixed method elements, and the resulting stabilization is nearly the same. As with mixed method stabilization, no user specified parameter is needed. The most noticeable difference between assumed strain and mixed-method stabilization is in the derivation. Assumed strain stabilization is much simpler. As we will see in Chapter 2, a major benefit of this simplification is the ability to derive stabilization for the three dimensional 8 node hexahedral element.

The relative performance of these elements is problem dependent; thus QBI and ASQBI are very accurate for elastic bending, but they do not perform as well for elastic-
plastic problems. Although it is not so accurate for elastic bending, ADS may be a good choice since it is very simple to implement and does not require knowledge of the material's Poisson's ratio. It's performance should exceed that of the other 1-point elements for elastic-plastic solutions. If the Poisson's ratio of the material is known, the ASQBI strain field with 2-point integration will provide both accurate elastic bending and reasonable elastic-plastic performance at a slightly higher cost.

