INTERNATIONAL JOURNAL FOR NUMERICAL METHODS IN ENGINEERING, VOL. 31, 547–572 (1991)

CONTACT-IMPACT BY THE PINBALL ALGORITHM WITH PENALTY AND LAGRANGIAN METHODS

TED BELYTSCHKO AND MARK O. NEAL*

Department of Civil Engineering, The Robert R. McCormick School of Engineering and Applied Science, The Technological Institute, Northwestern University, Evanston, Illinois 60208-3109, U.S.A.

SUMMARY

Contact-impact algorithms, which are sometimes called slideline algorithms, are a computationally time-consuming part of many explicit simulations of non-linear problems because they involve many branches, so they are not amenable to vectorization, which is essential for speed on supercomputers. The pinball algorithm is a simplified slideline algorithm which is readily vectorized. Its major idea is to embed pinballs in surface elements and to enforce the impenetrability condition only to pinballs. It can be implemented in either a Lagrange multiplier or penalty method. It is shown that, in any Lagrange multiplier method, no iterations are needed to define the contact surface. Examples of solutions and running times are given.

1. INTRODUCTION

The interaction of bodies in impact-penetration is treated by special algorithms, often called slideline algorithms, which enforce the constraint that the two bodies cannot occupy the same space at the same time. Lagrange multiplier,[1,2] penalty[2] and projection[3,4] techniques have all been proposed to enforce this constraint. Usually the interpenetration condition is imposed on the piecewise linear or quadratic approximation to the surfaces by the finite element mesh. For problems which include large relative motions between the two bodies and erosion of elements, it becomes difficult and time consuming to keep track of which elements should be involved in the impact calculations. This computational expense is magnified by the fact that these slideline algorithms have many branches, and hence are difficult to vectorize. In dynamic finite element programs with explicit time integration, many of the element and nodal calculations can be vectorized; therefore, if the slideline calculations are not vectorized they can consume a considerable percentage of the total computation time.

In this paper, a new contact-impact procedure called the pinball algorithm is described; a short description was previously given by Belytschko and Neal.[5] The thrust of the pinball algorithm is to allow vectorization of as much of the slideline calculations as possible. This is accomplished by greatly simplifying both the search for the elements involved in the impact and in the enforcement of impenetrability with the use of spheres, or pinballs, embedded in the elements in the slideline calculations. The search then requires only a simple check on the distances between pinballs to determine interpenetration. A similar idea has also been used in the two-dimensional NABOR algorithm,[6] but the NABOR method used an *ad hoc* method based on spheres for the determination of stresses in the continua and did not use a surface normal. In the pinball algorithm the

* Present address: Engineering Mechanics Department, General Motors Research Laboratories, 30500 Mound Road, Warren, MI 48090-9055, U.S.A.

0029–5981/91/030547–26$13.00 *Received 24 May 1990*

element spheres are used only in the contact algorithm, while standard continuum mechanics is used for the continuum elements.

We will begin with a general formulation of the contact-impact problem and its semi-discretization by a weak inequality. The natural discrete form which emanates from this variational inequality is the well known Lagrange multiplier form of the contact problem. This form is used to show some important and useful facts which apply to any Lagrange multiplier problem:

(i) *in an explicit method, the contact surface is defined by the overlap of the contacting bodies when integrated without interaction*; in other words, subject to some restrictions, there is no need for iteration in an explicit method for contact-impact;
(ii) *according to a linearized stability analysis, the stable time step for the central difference method is not decreased by a Lagrange multiplier method.*

A penalty method was used in most of our work with the pinball algorithm. It is shown that the addition of the penalty, in contrast to the Lagrange multiplier method, always decreases the stable time step. In addition, upper bounds on the penalty force which are based on physics and the characteristics of explicit procedures are given; these upper bounds usually decrease the stable time step by less than 25 per cent as compared to the unpenalized problem.

Several numerical examples are given. The first example is one-dimensional; its purpose is to compare projection and various penalty methods and to examine the stability conditions derived here. The remaining examples are quite complex and examine the applicability of the method to problems with erosion and shell buckling(single-surface slideline). Timing studies on these problems show an almost fivefold speed-improvement over the Belytschko–Lin algorithm, which corresponds to approximately a factor of two on total running time.

2. GOVERNING EQUATIONS

We consider the problem of two bodies Ω^A and Ω^B which can impact and contact. The methods to be developed are applicable to any number of bodies but we restrict the formal development to two bodies for simplicity of notation. The spatial (Eulerian) co-ordinates are denoted by x_i, and the velocity v_i is given by

$$v_i = \dot{x}_i \tag{1}$$

where superscript dots denote material time derivatives. The density is denoted by ρ and the body forces by b_i. The stress state is described by the Cauchy (physical) stress σ_{ij} and the velocity strain (rate-of-deformation, stretching tensor) $\dot{\varepsilon}_{ij}$ is used to measure the deformation. The two bodies are governed by the momentum equation, the kinematic relations and the constitutive equation, which are, respectively

$$\sigma_{ij,j} + b_i = \rho \dot{v}_i \qquad \text{in } \Omega^A \cup \Omega^B := \Omega \tag{2}$$

$$\dot{\varepsilon}_{ij} = \tfrac{1}{2}(v_{i,j} + v_{j,i}) := v_{(i,j)} \qquad \text{in } \Omega^A \cup \Omega^B \tag{3}$$

$$\overset{\nabla}{\sigma}_{ij} = S_{ij}(\dot{\varepsilon}_{kl}, \sigma_{kl}, \dots) \qquad \text{in } \Omega^A \cup \Omega^B \tag{4}$$

Equation (3) can be substituted into (4) to provide a set of 5 (9) equations in two (three) dimensions, respectively, in the unknowns v_i and σ_{ij}. The initial conditions on the dependent variables are

$$v_i(0) = v_i^0 \quad \text{in } \Omega^A \cup \Omega^B \tag{5a}$$

$$\sigma_{ij}(0) = \sigma_{ij}^0 \quad \text{in } \Omega^A \cup \Omega^B \tag{5b}$$

Equation (1) can be used to obtain the co-ordinates of all points throughout the evolution of the problem from the velocities.

Impenetrability of the two bodies then requires that

$$\Omega^A \cap \Omega^B = 0 \tag{6}$$

The contact surface at any instant is designated by Γ_c. The boundary conditions are then given by

$$v_i = v_i^* \quad \text{on } \Gamma_{v_i}^A \cup \Gamma_{v_i}^B := \Gamma_{v_i} \quad \text{for } i = 1 \text{ to } 3 \tag{7a}$$

$$\sigma_{ij} n_j = \tau_i^* \quad \text{on } \Gamma_{\tau_i}^A \cup \Gamma_{\tau_i}^B := \Gamma_{\tau_i} \quad \text{for } i = 1 \text{ to } 3 \tag{7b}$$

$$\dot{g} = v_i^A n_i^A + v_i^B n_i^B \leqslant 0 \quad \text{on } \Gamma_c \tag{8}$$

$$\lambda = -\tau_{ni}^A = -\tau_n^B \geqslant 0 \quad \text{on } \Gamma_c \tag{9}$$

where

$$\tau_n^A = \sigma_{ij}^A n_i^A n_j^A \tag{10a}$$

$$\Gamma^A = \Gamma_c \cup \Gamma_{\tau_i}^A \cup \Gamma_{v_i}^A \tag{10b}$$

$$\Gamma_c \cap \Gamma_{\tau_i}^A \cap \Gamma_{v_i}^A = 0 \tag{10c}$$

Similar relations hold for domain B.

3. VARIATIONAL INEQUALITY AND DISCRETE INTERPOLANTS

The weak form of the contact problem is obtained from the principle of virtual work by appending the Lagrange multiplier term $\delta(\lambda \dot{g})$. We consider the trial functions to be kinematically admissible functions, so $v_i \in \mathscr{V}$ and $\lambda \in \Lambda$ where

$$\mathscr{V} = \{v_i : v_i \in C^0(\Omega^A \cup \Omega^B),\ v_i = v_i^* \text{ on } \Gamma_{v_i}\} \tag{11a}$$

$$\Lambda = \{\lambda : \lambda \in C^{-1}(\Gamma_c),\ \lambda \geqslant 0\} \tag{11b}$$

As indicated above, these functions need only be piecewise continuous and satisfy essential boundary conditions. The variations (or test functions) $\delta v_i \in \mathscr{V}_0$, $\delta\lambda \in \Lambda_0$ where

$$\mathscr{V}_0 = \{\delta v_i : \delta v_i \in C^0,\ \delta v_i = 0 \text{ on } \Gamma_{v_i}\} \tag{12a}$$

$$\Lambda_0 = \{\delta\lambda : \delta\lambda \in C^{-1},\ \delta\lambda \leqslant 0 \text{ on } \Gamma_c\} \tag{12b}$$

We define the virtual work by

$$\delta W = \delta W^{\text{int}} + \delta \mathscr{M} - \delta W^{\text{ext}} \tag{13}$$

(see Belytschko[7]) where

$$\delta W^{\text{int}} = \int_\Omega \delta v_{(i,j)} \sigma_{ij} \, d\Omega \tag{14}$$

$$\delta W^{\text{ext}} = \int_\Omega \delta v_i b_i \, d\Omega + \int_{\Gamma_{\tau_j}} \delta v_j \tau_j^* \, d\Gamma \tag{15}$$

$$\delta \mathscr{M} = \int_\Omega \delta v_i \rho \dot{v}_i \, d\Omega \tag{16}$$

The weak form for the contact problem is then given by

If $\mathbf{v} \in \mathcal{V}$, $\lambda \in \Lambda$ and

$$\delta W + \int_{\Gamma_c} \delta(\lambda \dot{g}) \mathrm{d}\Gamma \geqslant 0 \tag{17}$$

for all $\delta\mathbf{v} \in \mathcal{V}_0$, $\delta\lambda \in \Lambda_0$, then the momentum equation (2), traction boundary conditions (7b) and contact surface inequalities (8) are satisfied.

The equivalence of this weak form to the governing partial differential equations is demonstrated in Reference 14.

Semidiscretization and time integration

In the usual manner for finite element discretizations, the velocity field is approximated by using an elementwise separation of variables, so that

$$v_i(\mathbf{x}, t) = \sum_I N_I^e(\mathbf{x})\, v_{iI}^e(t) = \mathbf{N}^e \mathbf{v}^e \tag{18}$$

where $\mathbf{v}^e$ is related to the global matrix of nodal velocities by the Boolean connectivity matrix

$$\mathbf{v}^e = \mathbf{L}^e \mathbf{v} \tag{19}$$

It is also convenient to use the 'assembled' shape functions $\mathbf{N}$ given by

$$\mathbf{N} = \sum_e \mathbf{N}^e \mathbf{L}^e \tag{20}$$

so that we can write

$$v_i(\mathbf{x}, t) = \mathbf{N}(\mathbf{x})\mathbf{v}(t) \tag{21}$$

4. LAGRANGE MULTIPLIER METHOD

Discretization

The Lagrange multiplier field $\lambda(\mathbf{x})$ is approximated by

$$\lambda(\mathbf{x}) = \mathbf{N}^\lambda(\mathbf{x})\boldsymbol{\lambda} \tag{22}$$

Since no derivatives of λ appear in the variational statement, $\mathbf{N}^\lambda$ can be a C^{-1} function. Substituting (21) and (22) into (17) yields

$$\mathbf{f} + \mathbf{M}\dot{\mathbf{v}} + \mathbf{G}^{\mathrm{T}}\boldsymbol{\lambda} = 0 \tag{23}$$

$$\mathbf{G}\mathbf{v} \leqslant 0 \tag{24}$$

where

$$\mathbf{G} = \int_{\Gamma_c} (\mathbf{N}^\lambda)^{\mathrm{T}} \mathbf{n} \cdot \mathbf{N} \mathrm{d}\Gamma \tag{25}$$

In addition, we have the requirement emanating from (9) that

$$\boldsymbol{\lambda} \geqslant \mathbf{0} \tag{26}$$

In writing (26), we assume that the interpolant (22) is of low order (piecewise constant or bilinear), so that (26) is implied by (9). For higher order interpolants, such as bicubics, such a simple correspondence is not available.

Remark 1. In equation (23), $\mathbf{G}^{\mathrm{T}}\boldsymbol{\lambda}$ represents the discrete contact forces, which vanish whenever $\boldsymbol{\lambda} = 0$. Equation (24) is the discrete counterpart of equation (8).

The construction of the Lagrange multiplier (interface traction) interpolants poses an interesting and challenging problem. The main difficulty in the large displacement problem arises from the fact that the nodes of the two bodies are usually not contiguous, as shown in Figure 1. Simo *et al.*[8] have proposed an interesting idea where the interpolant for λ is developed by constructing an intersurface mesh which consists of the projections of the nodes of Ω^A and Ω^B onto Γ_c; see Figure 2. Even in two dimensions, this technique is quite awkward because when nodes become nearly contiguous the sizes of the λ elements become very disparate and it is difficult to determine when a λ element should be annihilated. In three dimensions, such methods would have to be combined with automatic mesh generators such as Dedekind tessellation, for constructing a mesh from the complicated pattern of nodes would be a formidable task without a powerful technique. Thus this process would be useful only in static problems where only order of 10^2 steps are used in a calculation. In dynamic explicit calculations, where the number of steps is generally of order 10^4 to 10^5, simpler techniques are essential.

We have therefore chosen to use a modified master–slave concept to define the λ interpolant. Body A is designated the master surface and on $\Gamma^A \cap \Gamma_c$ we interpolate λ by

$$\lambda = \tilde{\mathbf{N}}\boldsymbol{\lambda} \tag{27}$$

where $\tilde{\mathbf{N}}$ are projections of the element interpolants of body A onto the surface. In the problems solved here, 8-node hexahedra were used, so $\tilde{\mathbf{N}}$ are the two-dimensional bilinear shape functions. The matrix $\mathbf{G}$ is then

$$\mathbf{G} = \int_{\Gamma_c} \tilde{\mathbf{N}}^{\mathrm{T}} \mathbf{n} \cdot \mathbf{N} \mathrm{d}\Gamma = \int_{\Gamma_c} \tilde{\mathbf{N}}^{\mathrm{T}} \tilde{\mathbf{N}} \mathrm{d}\Gamma \tag{28}$$

since $\tilde{\mathbf{N}} = \mathbf{n} \cdot \mathbf{N}$.

The matrix $\mathbf{G}$ can be integrated by Lobatto quadrature; see Hughes.[9] The matrix $\mathbf{G}$ then consists of two submatrices

$$\mathbf{G}_A = \int_{\Gamma_c} (\tilde{\mathbf{N}}^A)^{\mathrm{T}} \tilde{\mathbf{N}}^A \mathrm{d}\Omega \tag{29a}$$

$$\mathbf{G}_{AB} = \int_{\Gamma_c} (\tilde{\mathbf{N}}^A)^{\mathrm{T}} \tilde{\mathbf{N}}^B \mathrm{d}\Omega \tag{29b}$$

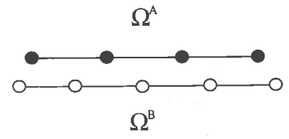


Figure 1. Slideline with non-contiguous nodes

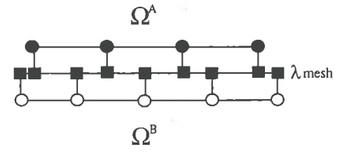


Figure 2. λ mesh

where $\mathbf{G}_A$ is diagonal and positive definite. If the nodes are coincident, $\mathbf{G}_{AB}$ is also diagonal and positive definite.

The central difference, explicit method with variable time steps will be used. The updating formulas for this method are

$$\mathbf{v}^{n+1/2} = \mathbf{v}^{n-1/2} + \Delta t^n \dot{\mathbf{v}}^n \tag{30}$$

$$\mathbf{d}^{n+1} = \mathbf{d}^n + \Delta t^n \mathbf{v}^{n+1/2} \tag{31}$$

$$\Delta t^n = \tfrac{1}{2}(\Delta t^{n-1/2} + \Delta t^{n+1/2}) \tag{32}$$

$$t^n = \sum_n \Delta t^n \tag{33}$$

$$\mathbf{d}^n := \mathbf{d}(t^n) \tag{34a}$$

$$\mathbf{v}^{n+1/2} := \mathbf{v}(t^n + \Delta t^{n+1/2}) \tag{34b}$$

The application of these integration formulas to the semidiscrete Lagrange multiplier form of the contact-impact problem, equations (2), (3) and (26), yields

$$\mathbf{M}(\mathbf{v}^{n+1/2} - \mathbf{v}^{n-1/2}) + \Delta t^n(\mathbf{f}^n + \mathbf{G}^T\boldsymbol{\lambda}^n) = 0 \tag{35a}$$

$$\mathbf{G}\mathbf{v}^{n+1/2} \leqslant 0 \tag{35b}$$

$$\boldsymbol{\lambda}^n \geqslant 0 \tag{35c}$$

In addition, the other dependent variable, the stress field σ_{ij}, must be updated. We will also write this update in central difference form.

$$\sigma^{n+1} = \sigma^n + \Delta t^n \overset{\nabla}{\sigma}{}^{n+1/2}(\mathbf{v}^{n+1/2}, \sigma^n) \tag{36}$$

where the superposed bar indicates these deviations from the central difference formula. In practice the update is not strictly a central difference formula because of techniques such as radial return[9] and due to the asynchronization that comes about since only the stress at t^n is available during the stress update unless an iterative procedure is used.

Non-iterative determination of Lagrange multipliers

In the following, we show that in explicit, central difference integration, the contact surface can be determined by the uncoupled integration of the two bodies without iteration: the surface projected from the volume of interpenetration of the uncoupled update of the two bodies corresponds to the contact surface. Furthermore, the Lagrange multipliers are directly determinable and will have the correct sign. First we show that the Lagrange multipliers are directly determinable on the contact surface. We restrict all matrices to those nodes of Ω^A and Ω^B which lie on Γ_c, since the equations of the remaining nodes are not affected by contact in an explicit algorithm. Furthermore, since friction is not considered, only the velocity components normal to Γ_c appear in the equations. Equation (35b) can be written as

$$\mathbf{G}_A\mathbf{v}_A^{n+1/2} + \mathbf{G}_{AB}\mathbf{v}_B^{n+1/2} \leqslant 0 \tag{37}$$

For the interpenetrated nodes the equality (it will be shown next that the Lagrange multipliers will be of the correct sign) is assumed to hold. Then (37) can be used to express $\mathbf{v}_A^{n+1/2}$ in terms of $\mathbf{v}_B^{n+1/2}$,

$$\mathbf{v}_A^{n+1/2} = -\mathbf{G}_A^{-1}\mathbf{G}_{AB}\mathbf{v}_B^{n+1/2} \tag{38}$$

since $\mathbf{G}_A$ is positive definite; for nodes originally not in contact a constant appears on the RHS but it is ignored here. Equations (35) are now partitioned in the same way as $\mathbf{v}$, so

$$\mathbf{M}_A\Delta\mathbf{v}_A + \Delta t^n(\mathbf{f}_A^n + \mathbf{G}_A\boldsymbol{\lambda}^n) = 0 \tag{39a}$$

$$\mathbf{M}_B\Delta\mathbf{v}_B + \Delta t^n(\mathbf{f}_B^n + \mathbf{G}_{AB}^T\boldsymbol{\lambda}^n) = 0 \tag{39b}$$

$$\Delta\mathbf{v} \equiv \mathbf{v}^{n+1/2} - \mathbf{v}^{n-1/2} \tag{39c}$$

Substituting (38) into (39a) yields

$$\mathbf{M}_A\mathbf{G}_A^{-1}\mathbf{G}_{AB}^T\Delta\mathbf{v}_B - \Delta t^n(\mathbf{f}_A^n + \mathbf{G}_A\boldsymbol{\lambda}^n) = 0 \tag{40a}$$

$$\boldsymbol{\lambda}^n = \mathbf{G}_A^{-1}\left(\frac{1}{\Delta t}\mathbf{M}_A\mathbf{G}_A^{-1}\mathbf{G}_{AB}^T\Delta\mathbf{v}_B - \mathbf{f}_A^n\right) \tag{40b}$$

Equation (40b) can be used to eliminate $\boldsymbol{\lambda}$ in (29b), yielding

$$(\mathbf{M}_B - \mathbf{G}_{AB}^T\mathbf{G}_A^{-1}\mathbf{M}_A\mathbf{G}_A^{-1}\mathbf{G}_{AB})\Delta\mathbf{v}_B + \Delta t^n(\mathbf{f}_B^n - \mathbf{G}_{AB}\mathbf{G}_A^{-1}\mathbf{f}_A^n) = 0 \tag{41}$$

Equation (41) can be used to obtain $\Delta\mathbf{v}_B$ without knowledge of the Lagrange multipliers.

It is now proven that on the nodes which penetrate according to an uncoupled integration of the two bodies, the Lagrange multipliers will be negative (compression). To demonstrate this, the accelerations of the uncoupled system at which the contact constraint is violated are denoted by $\tilde{\mathbf{a}}_A$ and $\tilde{\mathbf{a}}_B$, so

$$\tilde{\mathbf{a}}_A = -\tilde{\mathbf{M}}_A^{-1}\tilde{\mathbf{f}}_A \tag{42a}$$

$$\tilde{\mathbf{a}}_B = -\tilde{\mathbf{M}}_B^{-1}\tilde{\mathbf{f}}_B \tag{42b}$$

Violation of the constraint (37) implies

$$\tilde{\mathbf{G}}_A\tilde{\mathbf{a}}_A + \tilde{\mathbf{G}}_{AB}\tilde{\mathbf{a}}_B \geqslant 0 \tag{43}$$

where the 'tildes' on the matrices designate those parts of the original matrix at which the constraint is violated. Substituting (42) into (43) gives

$$\tilde{\mathbf{G}}_A\tilde{\mathbf{M}}_A^{-1}\tilde{\mathbf{f}}_A + \tilde{\mathbf{G}}_{AB}\tilde{\mathbf{M}}_B^{-1}\tilde{\mathbf{f}}_B \leqslant 0 \tag{44}$$

Now consider the equations which give the accelerations $\mathbf{a}$ that satisfy the constraints, namely (39):

$$\mathbf{a}_A + \tilde{\mathbf{M}}_A^{-1}(\tilde{\mathbf{f}}_A + \tilde{\mathbf{G}}_A\tilde{\boldsymbol{\lambda}}) = 0 \tag{45a}$$

$$\mathbf{a}_B + \tilde{\mathbf{M}}_B^{-1}(\tilde{\mathbf{f}}_B + \tilde{\mathbf{G}}_{AB}^T\tilde{\boldsymbol{\lambda}}) = 0 \tag{45b}$$

Premultiplying (45a) by $\tilde{\mathbf{G}}_A$ and (45b) by $\tilde{\mathbf{G}}_{AB}$ and summing yields

$$(\tilde{\mathbf{G}}_A\mathbf{a}_A + \tilde{\mathbf{G}}_{AB}\mathbf{a}_B) + (\tilde{\mathbf{G}}_A\tilde{\mathbf{M}}_A^{-1}\tilde{\mathbf{f}}_A + \tilde{\mathbf{G}}_{AB}\tilde{\mathbf{M}}_B^{-1}\tilde{\mathbf{f}}_B) + \mathbf{A}\tilde{\boldsymbol{\lambda}} = 0 \tag{46}$$

where

$$\mathbf{A} = \tilde{\mathbf{G}}_A\tilde{\mathbf{M}}_A^{-1}\tilde{\mathbf{G}}_A + \tilde{\mathbf{G}}_{AB}\tilde{\mathbf{M}}_B^{-1}\tilde{\mathbf{G}}_{AB}^T$$

In the above, the first term vanishes by (43) for nodes which remain in contact, while the second term is negative according to (44), so

$$\mathbf{A}\tilde{\boldsymbol{\lambda}} \geqslant 0 \tag{47}$$

When the meshes in bodies A and B are coincident, $\mathbf{A}$ is diagonal and positive definite. Hence, (47) implies that $\tilde{\lambda}_I \geqslant 0$ for all nodes previously in contact which interpenetrate. Thus if the contact

equality constraint is applied at all nodes where the disjoint analysis of bodies A and B indicates penetration, then the Lagrange multipliers at those nodes will satisfy the contact force inequality (26). Hence, no iterations are needed by explicit procedures under these circumstances. This fact adds to the attractiveness of the ALE and 'adaptive' schemes proposed by Haber[10] and Kulak,[11] respectively, where nodes of the contacting bodies remain coincident.

Remark 2. Note that the stress update is not modified by the contact algorithm, which only changes the velocities at step $n + 1/2$. The stresses used to compute $\mathbf{f}^n$ depend on σ^n, which is a function of $\mathbf{v}^{n-1/2}$ (see (36)).

Remark 3. The lack of symmetry arising from the designation of a master surface can be eliminated by using a two-path algorithm in which A and B are sequentially designated as the masters and the update is obtained by averaging the two updates.

5. STABILITY OF THE LAGRANGE MULTIPLIER METHOD

The critical time step for the Lagrange multiplier method in the case of continuing contact can be obtained by examining the linear homogeneous problem associated with (23) and (24); release does not affect the stiffness of nodes so it need not be considered. In the linear case $\mathbf{f} = \mathbf{Kd}$, so we have the constrained problem

$$\mathbf{M}\dot{\mathbf{v}}^n + \mathbf{K}\mathbf{d}^n + \mathbf{G}\boldsymbol{\lambda}^n = 0 \tag{48a}$$

$$\mathbf{G}^{\mathrm{T}}\dot{\mathbf{v}} = 0 \tag{48b}$$

The central difference method is stable if

$$\Delta t \leqslant \min\left\{\frac{2}{\omega_i}(\sqrt{1+\mu_i^2} - \mu_i)\right\} \tag{49}$$

(see Reference 7) where ω_i and μ_i are the frequencies and fractions of critical damping in the natural modes. The frequencies of the natural modes in the unconstrained problem are determined from the eigenvalue problem

$$\mathbf{K}\mathbf{z} = \omega_i^2 \mathbf{M}\mathbf{z} \tag{50}$$

In the case of two disjoint bodies, Ω^{A} and Ω^{B}, the $\mathbf{K}$ and $\mathbf{M}$ matrices would be the uncoupled stiffness and mass matrices of the two bodies, respectively.

The eigenvalue problem associated with the frequencies of (48) is

$$\begin{bmatrix} (\mathbf{K} - \bar{\omega}_i^2 \mathbf{M}) & \mathbf{G} \\ \mathbf{G}^{\mathrm{T}} & \mathbf{0} \end{bmatrix} \begin{Bmatrix} \mathbf{z} \\ \boldsymbol{\lambda} \end{Bmatrix} = \mathbf{0} \tag{51}$$

It follows immediately from the Rayleigh nesting theorem that the frequencies of (51), $\bar{\omega}_i$, are nested by the frequencies ω_i, so

$$\bar{\omega}_{\max} \leqslant \omega_{\max} \tag{52}$$

Hence, on the basis of a linear analysis, the Lagrange multiplier method should not decrease the stable time step and the stable time step for the mesh can be obtained by element-based procedures (see Reference 7). However, numerical experience suggests that the non-linearities of contact-impact do reduce the time step slightly.

6. PENALTY METHOD

We first consider a generalized variational form that encompasses both the Lagrange multiplier form and a penalty. It is obtained by adding penalties to (17):

$$\delta W + \int_{\Gamma_c} \left[p_2 \delta \dot{g} g + \delta\left(p_0 \lambda \dot{g} + \frac{p_1}{2} \dot{g}^2 \right) \right] \mathrm{d}\Gamma \geqslant 0 \tag{53}$$

where p_i, $i = 0$ to 2, are parameters to be selected. Rather than employing slack variables, we tacitly append Heaviside step functions to g, i.e.

$$g \leftarrow gH(g) \tag{54}$$

The terms in the above contact integral can be identified as (i) p_0, the coefficient of the Lagrange multiplier term; (ii) p_1, the penalty on the rate of the interpenetration; (iii) p_2, the penalty on the interpenetration.

We first consider the case when $p_2 = 0$. The equations are then

$$\mathbf{f} + \mathbf{M}\dot{\mathbf{v}} + p_0 \mathbf{G}^{\mathrm{T}}\boldsymbol{\lambda} + p_1 \mathbf{G}\mathbf{v} = 0 \tag{55a}$$

$$\mathbf{G}\mathbf{v} \leqslant 0 \tag{55b}$$

The contact force is now given by

$$\mathbf{f}_c = p_0 \mathbf{G}^{\mathrm{T}}\boldsymbol{\lambda} + p_1 \mathbf{G}\mathbf{v} \tag{56}$$

In the Lagrange multiplier method, the contact force is given by (56) with $p_0 = 1$, $p_1 = 0$. The above is an augmented Lagrangian of the type extensively described by Bertsekas.[12] In these methods, the constraint (55b) can be satisfied by an iterative procedure on λ, where

$$\lambda^{\upsilon+1} = \lambda^{\upsilon} + \frac{p_1^{\upsilon}}{p_0} \mathbb{G}\mathbb{v}^{\upsilon} \tag{57}$$

In contrast to Bertsekas,[12] we do not introduce slack variables for the inequality constraints, because, as described previously, when explicit time integration is used, the nodes at which the inequality holds automatically separate. Thus the constrained problem needs to be dealt with only at the nodes which are *not* separating, i.e. where $\dot{g} \leqslant 0$.

When $p_2 = 0$, the contact problem is governed by the *relative* velocities between the impacting bodies. The objective is to bring the relative velocities $\dot{g}$ to zero, for it is this condition which is most important to determining the stress waves which emanate from impact; any interpenetration is secondary since wave generation does not depend on interpenetration. In fact, it is undesirable to induce oscillations in the velocity on the contact interface in the process of exactly satisfying the interpenetration condition.

7. PINBALL ALGORITHM

The main idea of the pinball algorithm is to enforce the impenetrability condition and define the interpenetration g via a set of spheres, or pinballs which are embedded in the finite elements, as shown in Figure 3. By enforcing the contact constraint on the spheres rather than the elements themselves, the time required by the contact algorithm can be greatly reduced because: (i) the determination of whether interpenetration has occurred becomes a simple check of the distance between two pinballs, (ii) when combined with a penalty method, it involves almost no iterative calculations or conditional statements, so it is much more amenable to vectorization.

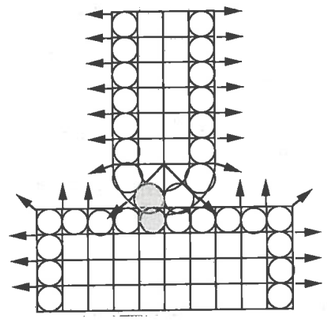


Figure 3. The pinball concept shown in two dimensions

The pinball algorithm is used in conjunction with the assembled surface normal algorithm of Belytschko and Lin,[4] which assembles an approximate normal to outside surfaces. The normal is non-zero only on nodes on outside surfaces, and pinballs are placed only in elements with at least one node with a non-zero normal.

The hexahedral elements used in this formulation are described in detail by Flanagan and Belytschko.[13] A sphere, or pinball, is embedded in each of these hexahedral elements of the mesh. These pinballs are then be used to determine which elements are involved in the contact. The centre and radii of the sphere are given in element e by

$$C_i = \frac{1}{8}\sum_{I=1}^{8} x_{Ii}^e \tag{58a}$$

$$R = \sqrt[3]{\frac{3V^e}{4\pi}} \tag{58b}$$

respectively, where C_i are the co-ordinates of the centre of the sphere, x_{Ii}^e are the co-ordinates of node I of element e, R is the radius of the pinball and V^e is the volume of element e.

The centre of each sphere is simply the average of its nodal co-ordinates while the radius is determined by setting the volume of the resulting sphere equal to the volume of the element itself. For elastic–plastic problems most of the element deformation can be considered nearly incompressible; therefore, the element volume, and also the radii of the pinballs, will change little over the course of the simulation. For this reason, we calculate the radii only once and consider them to be constant thereafter. The centres of the pinballs, however, are calculated every time step. For materials with substantial compressibility this assumption of nearly constant volume would be incorrect and the radius for each element would have to be recalculated every few steps.

The detection of the impacting pairs is, computationally, a very simple procedure. The distance between the centres of each slave pinball and each master pinball is calculated and then compared with the sum of the radii of the two elements. Interpenetration has occurred when

$$d < R_1 + R_2 \tag{59}$$

where d is the distance between the centres of elements 1 and 2 and R_1, R_2 are the radii of elements 1 and 2. Note that in this procedure the masters and the slaves may be penetrated by more than one element during a time step.

The penetration depth of the two elements is easily calculated. Consider two interpenetrating pinballs, 1 and 2, in Figure 4, with the velocities $\mathbf{v}_1$ and $\mathbf{v}_2$; the normals of the associated surfaces are $\mathbf{n}_1$ and $\mathbf{n}_2$. The position vectors of the centres of the two pinballs are given by $\mathbf{C}_1$ and $\mathbf{C}_2$. The penetration is given by g and is defined as the relative displacement of the centres of the pinballs in the average normal direction needed to eliminate interpenetration, so that the following equation determines g:

$$\mathbf{d}^{\mathrm{T}}\mathbf{d} = (R_1 + R_2)^2 \tag{60}$$

where

$$\mathbf{d} = \mathbf{C}_1 - \mathbf{C}_2 + g\mathbf{n} \tag{61}$$

where

$$\mathbf{n} = (\mathbf{n}_2 - \mathbf{n}_1)/\|\mathbf{n}_2 - \mathbf{n}_1\| \tag{62}$$

where $\|\cdot\|$ designates the length of a vector. The penetration depth g can be determined by

$$g = -b + \sqrt{b^2 - c} \tag{63}$$

where

$$b = \mathbf{n}^{\mathrm{T}}(\mathbf{C}_1 - \mathbf{C}_2) \tag{64}$$

$$c = \|\mathbf{C}_1 - \mathbf{C}_2\|^2 - (R_1 + R_2)^2 \tag{65}$$

Note that only the positive sign on the radical in (63) need be considered; the negative root corresponds to a negative value of g which is irrelevant.

The rate of penetration is computed by

$$\dot{g} = \Delta g/\Delta t \tag{66}$$

where $\mathbf{v}_I^{(j)}$, $I = 1$ to 8, are the nodal velocities of element j.

The quantity g can also be considered to be given by the time integral of $\dot{g}$:

$$g = \int_{t_1}^{t} \dot{g}\,\mathrm{d}t \tag{67}$$

where t_1 is the time when penetration begins. In a surface-based slideline algorithm, $\dot{g}$ is not path independent. To see this, consider the path shown in Figure 5. If node A traverses the path (1 to 3)

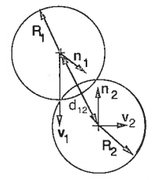


Figure 4. Interpenetration of two pinballs

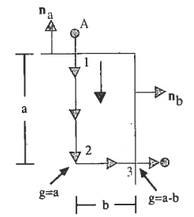


Figure 5. Penetrating node with path dependent penetration depth

and $\mathbf{n}_a$ is chosen as the normal along (1 to 2) and $\mathbf{n}_b$ as the normal along (2 to 3), then when node A exits the penetrated body $g = a - b$. In a more complicated path, the value of g at exit would depend on the point at which the normal used in computing g switches from $\mathbf{n}_a$ to $\mathbf{n}_b$, but in general it would not vanish at exit. The pinball algorithm provides a natural and unique way of computing g: it is simply and uniquely computed by (61).

8. PENALTY PARAMETERS AND STABILITY

One of the major disadvantages of penalty methods is that few guidelines are available to provide a good penalty parameter. In this section a procedure for choosing the appropriate penalty parameter is presented based on limiting the impact to plastic impact (zero coefficient of restitution on the pinball level) and minimizing the decrease in the stable time step.

To obtain estimates on the stable time step, the element eigenvalue inequality,[13] which bounds the maximum eigenvalue of the system, will be used. This eigenvalue inequality theorem is not limited to element level submatrices. The submatrices may be smaller assemblages of elements such as groups of two elements. For any subsystem or element the eigenproblem is

$$\mathbf{K}^s\mathbf{x}^s = (\omega^s)^2\mathbf{M}^s\mathbf{x}^s \tag{68}$$

where ω^s is the subsystem frequency, $\mathbf{x}^s$ is the subsystem eigenvector and $\mathbf{K}^s$, $\mathbf{M}^s$ are the subsystem stiffness and mass matrices, respectively. The system eigenvalues are then bounded by the subsystem eigenvalues.

$$\omega_{\max} \leqslant \omega^s_{\max} \tag{69}$$

where $\omega^s_{\max}$ is the maximum of all the eigenvalues of all the subsystems. Furthermore, if these subsystem matrices are assemblages of element matrices then the element eigenvalue inequality gives

$$\omega^s_{\max} \leqslant \omega^e_{\max} \tag{70}$$

To obtain a sharp bound on the stable time step, two-element assemblages rather than single elements are used.

$$[\mathbf{K}^s_{\mathrm{D}} + \mathbf{K}^s_{\mathrm{P}}]\mathbf{x}^s = (\omega^s)^2\mathbf{M}^s\mathbf{x}^s \tag{71}$$

where $\mathbf{K}^s_{\mathrm{D}}$ is the stiffness of the assemblage (or subsystem) of contacting elements and $\mathbf{K}^s_{\mathrm{P}}$ is the penalty stiffness; we have ignored damping.

Since the determination of the eigenvalue of such a subsystem would be quite difficult, it is desirable to find a form of $\mathbf{K}^s_{\mathrm{P}}$ that gives an upper bound on the frequency, ω^s, and which can be obtained by assembling element level penalty stiffnesses. Such a matrix is given by the diagonal absolute row sum matrix

$$\hat{\mathbf{K}}^s_{ij} = \delta_{ij}\sum_k |(\mathbf{K}^s_{\mathrm{P}})_{ik}| \quad \text{(no sum on } i\text{)} \tag{72}$$

Since $\hat{\mathbf{K}}^s$ is a diagonal matrix we can find element penalty stiffnesses for each of the two elements that, when assembled, will form $\hat{\mathbf{K}}^s$, so the stability problem reduces to finding the maximum frequency of the element matrices.

$$[\mathbf{K}^e_{\mathrm{D}} + \hat{\mathbf{K}}^e]\mathbf{x}^e = (\omega^e)^2\mathbf{M}^e\mathbf{x}^e \tag{73}$$

A proof that the maximum frequency of (73) will bound the maximum frequency is given by Neal.[14]

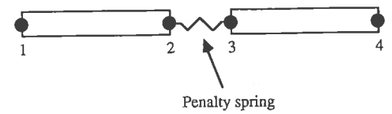


Figure 6. One-dimensional impact problem of two elements

To illustrate the method, consider two one-dimensional 2-node bar elements. The two elements and a penalty spring are shown in Figure 6. The mass and stiffness matrices of the subsystem are given by

$$\mathbf{M}^s = \frac{\rho AL}{2}\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{74}$$

$$\mathbf{K}^s_{\mathrm{D}} = \frac{EA}{L}\begin{bmatrix} 1 & -1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & -1 & 1 \end{bmatrix} \tag{75}$$

where ρ, A, L and E are the density, cross-sectional area, length and modulus of the elements, respectively.

For the particular case of the two bar elements with the penalty between nodes 2 and 3, the penalty stiffness for the standard penalty method is given by

$$\mathbf{K}^s_{\mathrm{P}} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & p_2 & -p_2 & 0 \\ 0 & -p_2 & p_2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \tag{76}$$

where p_2 is the penalty parameter as defined in the previous section. The eigenproblem of equation (71) results in a cubic equation for the eigenvalues. The diagonal row sum form of the penalty stiffness matrix is given by

$$\hat{\mathbf{K}}^s = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 2p_2 & 0 & 0 \\ 0 & 0 & 2p_2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \tag{77}$$

It is apparent that this matrix can be assembled from element level penalty stiffnesses given by

$$\hat{\mathbf{K}}^1 = \begin{bmatrix} 0 & 0 \\ 0 & 2p_2 \end{bmatrix} \quad \hat{\mathbf{K}}^2 = \begin{bmatrix} 2p_2 & 0 \\ 0 & 0 \end{bmatrix} \tag{78}$$

The element eigenvalue problem is

$$\left(\frac{AE}{L}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} + \begin{bmatrix} 2p_2 & 0 \\ 0 & 0 \end{bmatrix}\right)\mathbf{x} = (\omega^e)^2 \frac{\rho AL}{2}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\mathbf{x} \tag{79}$$

so

$$\omega^e = 2\frac{c}{L}\sqrt{\frac{1+\beta+\sqrt{1+\beta^2}}{2}} \tag{80}$$

where

$$\beta = \frac{p_2 L}{EA} \tag{81}$$

From equation (80) we can see that if β is zero then the maximum frequency reduces to $2c/L$, which is the maximum frequency for the unconstrained problem. Substituting equation (80) into equation (49) with $\mu = 0$ gives the stable time step with the penalty

$$\Delta t = \frac{L}{c}\sqrt{\frac{2}{1+\beta+\sqrt{1+\beta^2}}} \tag{82}$$

As can be seen, the introduction of the penalty always decreases the stable time step: if $\beta > 0$, then $\Delta t < L/c$.

The stability condition provides one guideline as to the magnitude of the penalty parameter. By solving equation (82) for the penalty parameter β we get

$$\beta \leqslant \frac{2f(1-f)}{(1-2f)} \tag{83}$$

where $f = 1/f_c^2$ and f_c is the Courant number.

To test the stability criteria derived above, the one-dimensional example described in Section 10 was tested with the penalty formulation. The minimum penalty parameter, β, that results in an instability as predicted by equation (83) is compared to numerical detection of instability in Figure 7. As can be seen from Figure 7, the stability analysis given above will not guarantee stability.

The second condition on the penalty force emanates from energy conservation. In order to preserve stability the penalty force cannot increase the energy. This corresponds to requiring that the coefficient of restitution does not exceed 1·0, and in fact it will be noted that for a semidiscrete

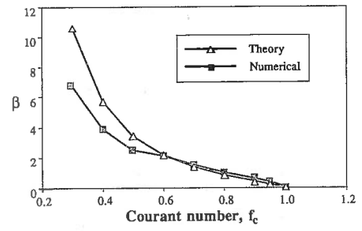


Figure 7. Linearized stability compared to numerical stability

system plastic impact is appropriate. To ensure energy conservation, note that

$$T^{\mathrm{P}} = e^2 T \tag{84}$$

where T is the kinetic energy before impact, T^{P} is the kinetic energy after impact (relative to the centre of mass), and e is the coefficient of restitution. If we consider impact as shown in Figure 6, equation (84) is equivalent to

$$e(v_2 - v_3) = (v_3^{\mathrm{P}} - v_2^{\mathrm{P}}) \tag{85}$$

(see Greenwood[15]), where v_2 and v_3 are the velocities before impact and v_2^{P} and v_3^{P} are the velocities after impact. Since $e \leqslant 1$,

$$(v_2 - v_3) \geqslant (v_3^{\mathrm{P}} - v_2^{\mathrm{P}}) \tag{86}$$

The post impact velocities for the penalty method are given by

$$v_2^{\mathrm{P}} = v_2 - \Delta t \frac{F^{\mathrm{P}}}{m_2} \tag{87a}$$

$$v_3^{\mathrm{P}} = v_3 + \Delta t \frac{F^{\mathrm{P}}}{m_3} \tag{87b}$$

where F^{P} is the penalty force given by

$$F^{\mathrm{P}} = p\,(x_2 - x_3) = \frac{\beta EA}{L}(x_2 - x_3) \tag{88}$$

where x_2 and x_3 are the positions of nodes 2 and 3. Substituting (87) into (86) and solving for the penalty force yields the condition that

$$F^{\mathrm{P}} \leqslant \frac{2m_2 m_3 (v_2 - v_3)}{\Delta t(m_2 + m_3)} := 2F^{\mathrm{P}}_{\mathrm{crit}} \tag{89}$$

Furthermore, if we restrict the coefficient of restitution to $e = 0$ (plastic impact), we obtain the condition

$$F^{\mathrm{P}} \leqslant F^{\mathrm{P}}_{\mathrm{crit}} \tag{90}$$

Note that this condition cannot be established at the beginning of a problem since the constraint depends on the nodal velocities at the time of impact. This condition must be satisfied each time two nodes initially impact. The stable p_2 is calculated first by equation (83). Then the penalty force is calculated by equation (88). The maximum allowable penalty force which does not increase kinetic energy is determined by equation (89) and the minimum of these two forces is then used.

This energy bound on the penalty force was tested numerically for the example described in Section 10; the results are shown in Figure 8. It can be seen from this figure that the predicted stability limits match the numerical results very closely if the coefficient of restitution is limited to values between 0 and 1.

The maximum frequency of the hexahedral element is given by Flanagan and Belytschko[16] in terms of a reduced eigenvalue problem.

$$\omega_{\max} = \frac{8}{\rho V} k_{\max} \tag{91}$$

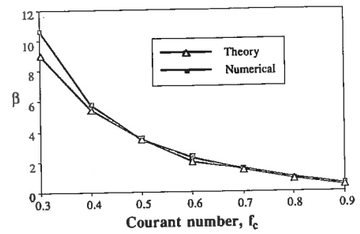


Figure 8. Linearized stability compared to numerical stability with zero coefficient of restitution

where

$$k_{\max} \leqslant \frac{\lambda + 2\mu}{V} a_{ii} \tag{92}$$

$$a_{ij} = B_{iI} B_{jI} \tag{93}$$

and where $\omega_{\max}$ is the maximum frequency, ρ and V are the element density and volume, λ and μ are the Lamé constants, respectively, and B_{iI} is the gradient operator where the small indices ranges over the spatial co-ordinates and the upper case indices range over the nodes of the element. The original eigenvalue problem in terms of the stiffness matrix is written as

$$K_{iIjJ} u_{jJ} = k u_{iI} \tag{94}$$

In Neal[14] we derive a diagonal form of the element penalty stiffness given by

$$\hat{K}_{iIjJ} = \tfrac{1}{2} p_2 \delta_{IJ} \delta_{ij} \tag{95}$$

To find the eigenvalues of the element with penalty force effects, the following eigenvalue problem is written:

$$(K_{iIjJ} + \hat{K}_{iIjJ}) u_{jJ} = \hat{k} u_{iI} \tag{96}$$

where $\hat{k}$ are the eigenvalues of this new problem. Since this penalty stiffness is diagonal the addition to the usual stiffness will result in a shift of the usual eigenvalues k.

$$\hat{k}_{\max} = k_{\max} + \tfrac{1}{2} p_2 \tag{97}$$

The frequency of the element with penalty stiffness effects is then given by (91) in terms of $\hat{k}_{\max}$.

Remark 4. The inability of the linearized stability analysis to predict the correct critical time step for a penalty method, as in the Lagrange multiplier method, stems from the fact that the impact process is inherently *non-linear*. If the penalty force exceeds $F^{\mathrm{p}}_{\mathrm{crit}}$, the impacting nodes will impact and release within the same time step with an apparent coefficient of restitution which exceeds 1. This violates the physics of impact and leads to violation of conservation of energy. Therefore, a linearized stability analysis is adequate only when the nodes remain in contact.

Remark 5. Although imposing (89) satisfies conservation of energy, the penalty force should be limited by (90) because the coefficient of restitution for impact of explicitly integrated semi-discretized systems should be plastic with $e = 0$. This results from the physics of wave propagation and the properties of explicit integration. Two surfaces which impact cannot release until the

rarefaction waves reach the impacting surface. Since the rarefaction waves are generated at free surfaces by the reflection of the compressive waves generated by impact, and the free surfaces are at least one node away from the area of impact, in the discrete model the Courant stability condition implies that rarefaction waves cannot reach the impact area in fewer than two time steps from the time of impact. Hence, in an explicitly integrated discrete mesh, impact is plastic with a vanishing coefficient of restitution.

Remark 6. The restriction (90), i.e. that impact is plastic, forces the discrete impact process to be non-conservative, i.e. to dissipate energy. Thus, even elastic impact becomes a dissipative process in the semidiscrete model. This anomaly is a consequence of the fact that the impact condition $\dot{g} = 0$ applies only to a set of measure zero (the surface of the impacting bodies) in the PDE's, but applies to a finite volume in the semidiscrete system.

Remark 7. Equation (89) provides a natural way for providing a penalty force which is a function of $\dot{g}$. Using $F^{\mathrm{p}}_{\mathrm{crit}}$ as defined by (89) is excessive when the meshes are not coincident, but applying a fraction of this value in combination with a penalty based on g is advantageous. In fact, the relative magnitudes can be adjusted to critically damp the impact process. However, a rate-based penalty does retard the release process.

9. PENALTY IMPLEMENTATION OF PINBALL METHOD

The implementation of the penalty method in the pinball method is described here. The penalty force on any pinball is applied to all nodes of each element. The force is proportional to the penetration depth and is given by

$$\mathbf{F}^{\mathrm{p}} = (p_1 \dot{g} + p_2 g)\mathbf{n} \tag{98}$$

where $\mathbf{n}$ is given in equation (62) and

$$p_2 = \frac{\beta B A^2}{V} \tag{99}$$

and B, A and V are the bulk modulus, area of the impacted surface and volume of the element, $\mathbf{F}^{\mathrm{p}}$ is the penalty force on the pinball. In the present context, the properties of two pinballs must be considered, so the penalty parameter will be given by

$$p_2 = \tfrac{1}{2}\beta(B_1 R_1 + B_2 R_2) \tag{100}$$

where B_1, B_2 are the bulk moduli of the impacting pinballs, and R_1, R_2 are the radii of the impacting pinballs. Equation (98) gives the contact force that will be applied in opposite directions to each of the two impacting pinballs. This force is then divided among the eight nodes of each element.

$$\mathbf{F}^{e\mathrm{p}}_n = \tfrac{1}{8}\mathbf{F}^{\mathrm{p}} \qquad n = 1, 8 \tag{101}$$

where $\mathbf{F}^{e\mathrm{p}}_n$ are the element level penalty forces on local node n of the element. These forces are then assembled to the global force vector as usual. A flowchart of the impact algorithm is given in Table I.

The penalty force is divided among the eight nodes of the hexahedron to preserve the symmetry of the underlying linearized system. Since the position of the pinball depends on the eight nodes of the hexahedron, the linearized equations would not be symmetric if the force were subdivided only among the surface nodes; an alternative algorithm where $\mathbf{C}$ depends only on the surface node

Table I. Slideline algorithm

1. If this is the first step, use the element volume to calculate a radius for all elements on the slideline.
2. Calculate element normals for all elements. Elements with zero normals are eliminated from consideration in the contact search.
3. Calculate the centre of all elements with non-zero normals.
4. Put elements into appropriate cells.
5. Loop through elements of each cell to determine the penetrating pairs of elements.
6. Calculate the contact forces to be applied to the nodes of impacting element pairs.
7. Return to main driving routine.

Table II. Explicit algorithm with slideline

1. Initialization.
2. Calculate the external nodal forces.
3. Compute the internal nodal force array.
   a. Calculate the element stresses.
   b. Compute the element nodal forces arising from the element stresses.
   c. Assemble the element nodal forces to the internal nodal force array.
4. Call the slideline algorithm to calculate the contact forces and add them to the external force array.
5. Compute the nodal accelerations.
6. Integrate the accelerations to obtain the nodal velocities and displacements.
7. Go to 2.

velocities and hence the penalty forces are distributed only to the surface nodes is now under investigation, Belytschko and Bindeman.[17]

The penalty forces, along with the forces arising from element stresses and externally applied loads, are used in the calculation of the nodal accelerations. Therefore the contact routine appears in the algorithm immediately before the nodal accelerations are calculated. The flowchart of the complete explicit algorithm with the contact algorithm is given in Table II.

10. NUMERICAL EXAMPLES

In order to test the accuracy and efficiency of the proposed contact procedure several problems were examined. The first problem considered was the impact of two one-dimensional bars consisting of ten elements each. This problem was considered in order to compare two different methods of enforcing the impenetrability constraint: the penalty method and the projection method. This contact constraint is the only non-linearity that appears in the problem. As can be seen in Figure 9, one of the bars is given an initial velocity so that it impacts with the second bar. The material properties are such that the wave speed in the two bars is 10·0 m/sec.

Figures 10 to 12 give the velocity time histories for the nodes at the midpoint of the first rod ($x = 5{\cdot}0$), at the interface on the first rod ($x = 10{\cdot}0$) and at the midpoint of the second rod ($x = 15{\cdot}5$). As can be seen from Figure 11, the penalty method gives a rather noisy solution at the contact interface yet this does not appear to have much effect away from the contact zone (see Figures 10 and 12). The results for the Lagrange multiplier method were nearly identical to the projection method and therefore were not included in the results.

The second problem considered was of a copper rod impacting a steel plate at high velocity. The rod projectile consisted of 414 elements while the plate or target was modelled by 1428 elements. The geometry and the material properties for each of the objects are given in Table III.

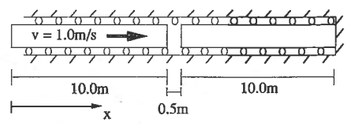


Figure 9. One-dimensional impact problem

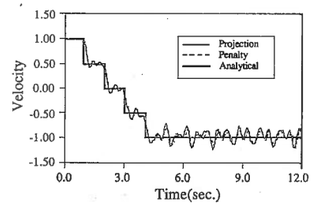


Figure 10. Projection and penalty methods at midpoint of first rod

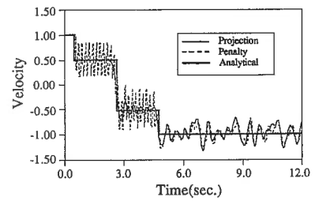


Figure 11. Projection and penalty methods at interface on first rod

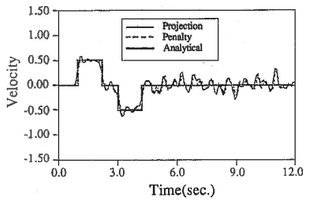


Figure 12. Projection and penalty methods at midpoint of second rod

Table III. Geometry and material properties of penetration problem 1

	Projectile (rod with a round nose)	Target (plate)
Length	= 4·900 in	3·950 in
Width	= —	7·900 in (half plate is modelled)
Thickness	= —	0·375 in
Radius	= 0·500 in	—
Density	= 8·31e − 3 lb-sec^2/in^4	7·34e − 3 lb-sec^2/in^4
Bulk modulus	= 2·0739e + 7 psi	2·4200e + 7 psi
Shear modulus	= 6·3800e + 6 psi	9·3000e + 6 psi
Plastic modulus	= —	1·5000e + 5 psi
1·4300e + 5 psi		
Yield stress	= 2·0300e + 4 psi	1·6000e + 5 psi
Ultimate stress	=	6·5300e + 4 psi
2·0300e + 5 psi		
Initial velocity	= 5·5566e + 4 in/sec (x-component) − 5·5566e + 4 in/sec (z-component)	0·0

The evolution of the problem is shown in Figure 13. In this example problem and the one that follows a von Mises yield criterion is used with isotropic, piecewise linear hardening. The yield stress and plastic modulus for each material are given in Table III. The material hardens until the effective stress is equal to the ultimate stress, at which point the material is considered perfectly plastic. When the effective strain of an element reaches the maximum allowable effective strain, the element is eroded, that is the stress in the element is considered zero from that time on. The maximum allowable effective strain used for steel is 1·0 while that used for copper is 2·0.

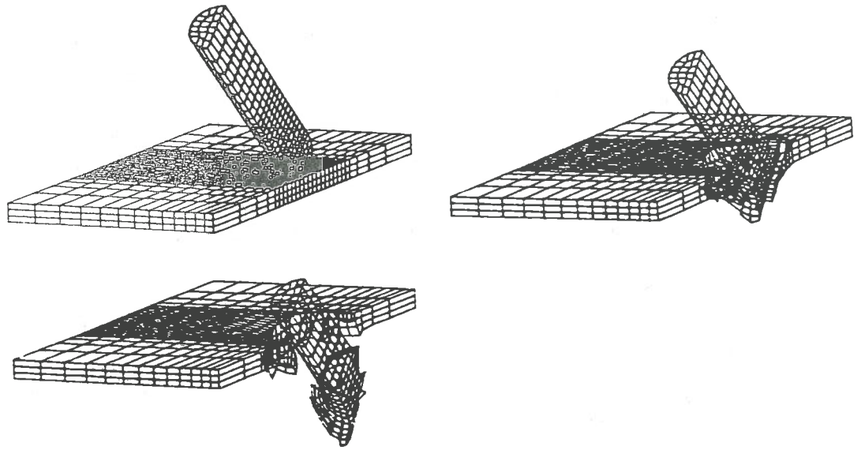


Figure 13. Example problem 1 at times 0, 25 and 54 μsec

Table IV. Timing studies for penetration problems

Algorithm	Example 1	Example 2, mesh 1	Example 2, mesh 2
Previous method	34·7 sec	94·4 sec	302·2 sec
Pinball algorithm	22·0 sec	40·4 sec	143·0 sec

This problem was also examined by Belytschko and Lin[4] with their projection method and a comparison of running times for both the methods is given in the first column of Table IV. As can be seen from this table, the new algorithm is substantially more efficient than the previous one on a vectorized machine. For this comparison, both algorithms were implemented into the three-dimensional finite element code WHAM3D and run on a Cray X-MP/14 with the CFT77 compiler. The differences in running times are due only to the different slideline algorithms. For unvectorized runs, the new algorithm is only marginally more efficient than the previous method. When the vectorized compiler is used, however, the old version of the slideline algorithm consumes nearly 50 per cent of the total CPU; for the new procedure this value has been reduced to only 15 per cent in the vectorized run (see Figures 14 and 15).

In the third example problem, the impact of a steel rod into a thick steel plate was considered. The geometry and material properties are given in Table V. Two meshes were considered for this problem; the two meshes are described by Table VI. The evolution of the problem is shown in Figure 16 and the second two columns of Table IV give the comparisons of CPU requirements with the Belytschko–Lin method. The difference in the improvement compared to the previous example is probably due to the fact that a higher percentage of elements are in the target.

The dynamic Hertz problem, which consists of the impact of an elastic sphere with a rigid wall, was examined to test the accuracy of the new algorithm. Ten degrees of the sphere are modelled with 499 elements as shown in Figure 17. All of the nodes in this model are constrained in the

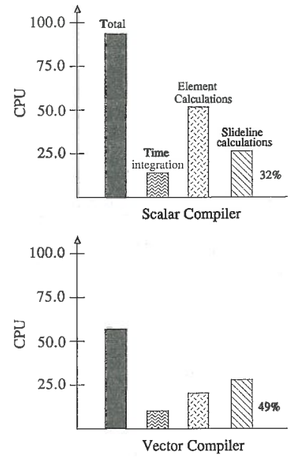


Figure 14. CPU requirements of the Belytschko–Lin algorithm

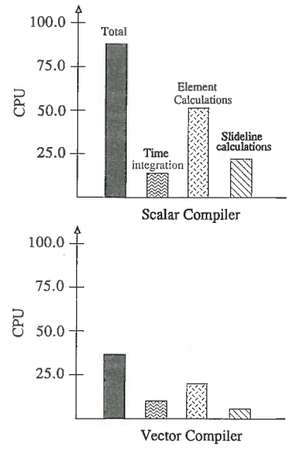


Figure 15. CPU requirements of the pinball algorithm

Table V. Geometry and material properties of thick target penetration problem

	Projectile (rod with a round nose)	Target (plate)
Length	= 4·040 in	0·600 in
Width	= —	5·600 in (half plate is modelled)
Thickness	= —	1·000 in
Radius	= 0·201 in	—
Density	= 7·30e − 4 lb-sec^2/in^4	7·30e − 4 lb-sec^2/in^4
Bulk modulus	= 2·3810e + 7 psi	2·778e + 7 psi
Shear modulus	= 1·1630e + 7 psi	1·136e + 7 psi
Plastic modulus	= —	1·5000e + 5 psi
1·4300e + 5 psi		
Yield stress	= 2·5000e + 5 psi	1·6000e + 5 psi
Ultimate stress	=	3·1000e + 5 psi
1·8500e + 5 psi		
Initial velocity	= 5·5566e + 4 in/sec (x-component) − 5·5566e + 4 in/sec (z-component)	0·0

Table VI. Meshes used for thick penetration problem

Mesh number	Projectile elements	Target elements
1	128	2688
2	918	4320

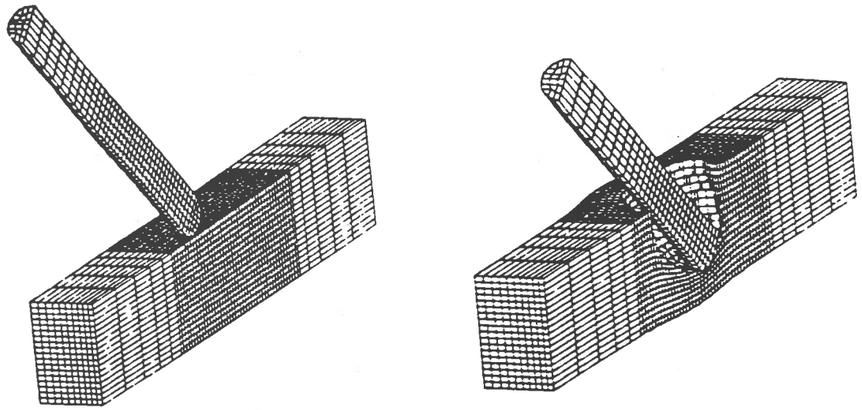


Figure 16. Evolution of example problem 2 at times 0·0 and 1·5e – 4 sec

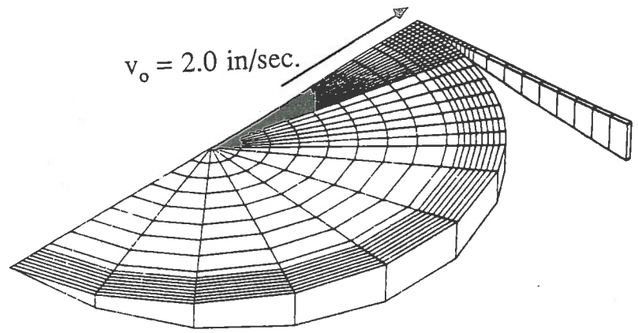


Figure 17. Elastic impact of sphere (dynamic Hertz problem) with rigid wall

circumferential direction; the nodes along the diameter are also constrained in the radial direction. Figure 17 also gives the material properties and dimensions of the sphere. The contact radius as a function of time is compared for the numerical simulation and the analytical result[18] in Figure 18.

The final example problem is that of a box beam impacting a rigid wall as shown in Figure 19.[19,20] The mesh consists of 756 shell elements described in Reference 21 and due to

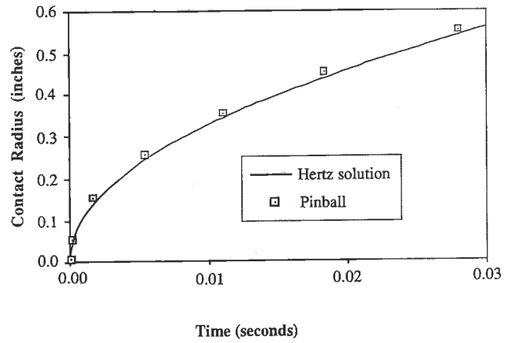
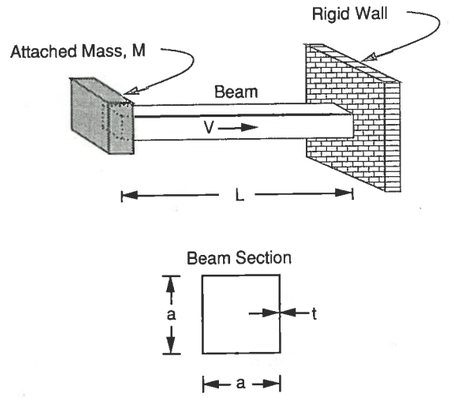


Figure 18. Radius of contact as function of time for dynamic Hertz problem



INITIAL CONDITIONS	MATERIAL	GEOMETRY
V = 15.64 m/s	E = 2.06 x 10^{11} N/m²	L = 0.15 m
	n = 0.3	a = 0.03 m
	r_o = 7840 kg/m³	t = 0.0015 m
	s_y = 2.0 x 10^8 N/m²	
	E_p = 6.3 x 10^8 N/m²	
	M = 1400 kg	

Figure 19. Box beam problem

symmetry only one quarter of the beam is modelled. This example was performed to demonstrate the capability of the new algorithm to simulate the contact of shell elements. The modifications of the pinball algorithm which are necessary for this application to shell elements are described by Sarwas.[22] Figure 20 gives the acceleration of the rigid mass as a function of time. Figure 21 shows

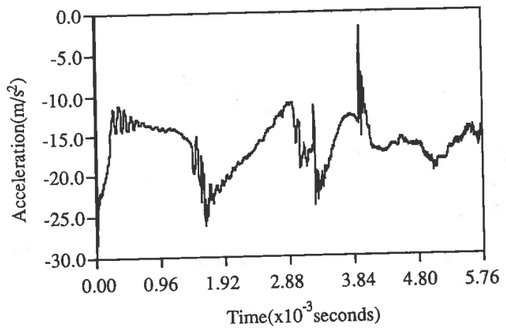


Figure 20. Acceleration of the attached mass

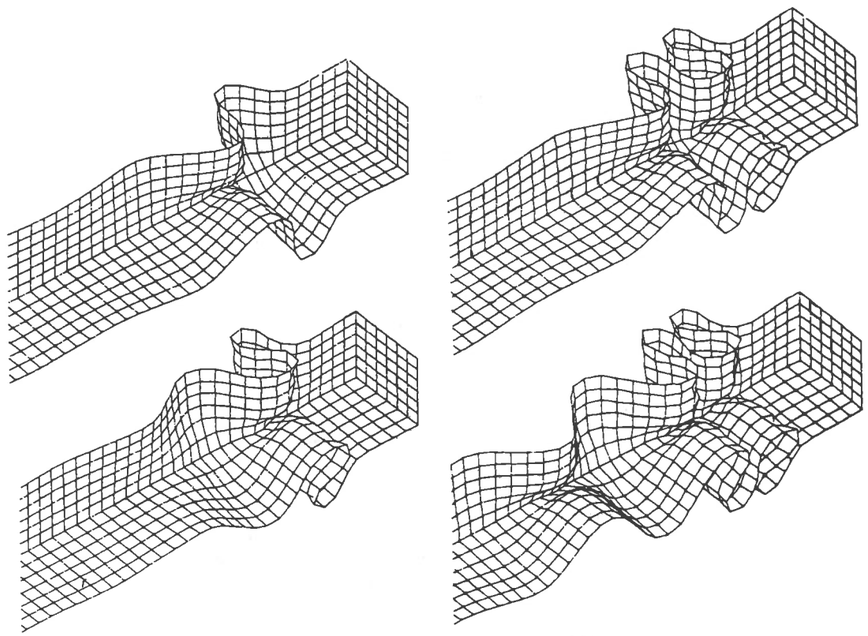


Figure 21(a). Evolution of box beam to 3·84 msec

that as the beam buckles the shell elements come into contact with one another. The times of these plots correspond to the times given on the x-axis of Figure 20. For this problem the entire mesh is considered as a single surface slideline, so contact between any two elements in the mesh is possible.

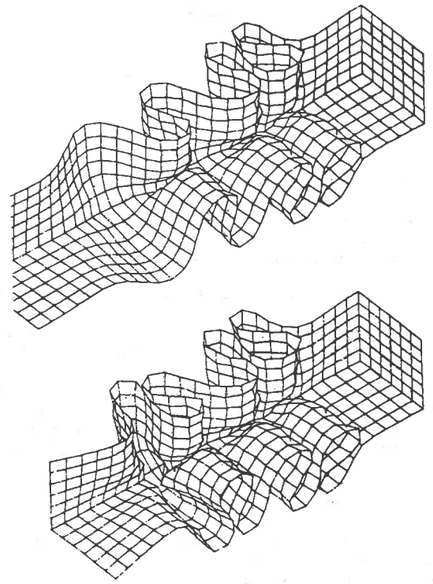


Figure 21(b). Box beam at 4·8 msec and 5·76 msec

11. CONCLUSIONS

The problem of contact-impact has been investigated, starting with the weak inequality which leads to the Lagrange multiplier form. The following are the major findings.

1. For contact of coincident meshes, no iteration is needed in an explicit integration procedure to determine the contact surface.
2. A useful upper bound can be set on the penalty by the constraint that the impact of nodes is plastic.
3. Very good bounds for the stable time step can be obtained by diagonalizing the penalty stiffnesses and adding them to the element.

The major breakthrough of this paper is the demonstration that a contact-impact algorithm can be simplified dramatically by interpreting the interpenetration g between the bodies as the interpenetration of spheres embedded in the elements. This simplifies the contact-impact algorithm and facilitates vectorization. Computer times for large three-dimensional problems show a fivefold speedup in the slideline algorithm and as much as a factor of two in total running time. The method is primarily intended for problems where sliding and friction are not crucial, such as penetration and crashworthiness. The effect of oscillations normal to the interface which will undoubtedly accompany sliding of two plane bodies has not been studied.

ACKNOWLEDGEMENT

The support of the Army Research Office under grant no. DAAL03–87–K–0035 to Northwestern University is gratefully acknowledged.

REFERENCES

1. T. J. R. Hughes, R. L. Taylor, J. L. Sackman, A. Curnier and W. Kanoknukulchai, 'A finite element method for a class of contact-impact problems', *Comp. Methods Appl. Mech. Eng.*, **8**, 249–276 (1976).
2. J. O. Hallquist, G. L. Goudreau and D. J. Benson, 'Sliding interfaces with contact-impact in large-scale Lagrangian computations', *Comp. Methods Appl. Mech. Eng.*, **51**, 107–137 (1985).
3. R. L. Taylor, N. J. Carpenter and M. G. Katona, 'Analysis of nonlinear transient problems with constraints', *Int. j. numer. methods eng.*, to appear.
4. T. Belytschko and J. I. Lin, 'A three-dimensional impact-penetration algorithm with erosion', *Comp. Struct.*, **25**, 95–104 (1987).
5. T. Belytschko and M. O. Neal, 'The vectorized pinball contact impact routine', in A. H. Hadjian (ed.), *Trans. 10th Int. Conf. on Structural Mechanics in Reactor Technology, Vol. B*, Anaheim, CA. AASMiRT, Los Angeles, CA, 1989, pp. 161–166.
6. G. R. Johnson, R. A. Stryk and J. G. Dodd, 'Dynamic Lagrangian computations for solids, with variable nodal connectivity for severe distortions', *Int. j. numer. methods eng.*, **23**, 509–522 (1986).
7. T. Belytschko, 'An overview of semidiscretization and time integration procedures', in T. Belytschko and T. J. R. Hughes (eds.), *Computational Methods for Transient Analysis*, North-Holland, Amsterdam, 1983, pp. 1–65.
8. J. C. Simo, P. Wriggers and R. L. Taylor, 'A perturbed Lagrangian formulation for the finite element solution of contact problems', *Comp. Methods Appl. Mech. Eng.*, **50**, 163–180 (1985).
9. T. J. R. Hughes, 'Analysis of transient algorithms with particular reference to stability behavior', in T. Belytschko and T. J. R. Hughes (eds.), *Computational Methods for Transient Analysis*, North-Holland, Amsterdam, 1983, pp. 67–155.
10. R. B. Haber, 'A mixed Eulerian–Lagrangian displacement model for large deformation analysis in solid mechanics', *Comp. Methods Appl. Mech. Eng.*, **43**, 277–292 (1984).
11. R. F. Kulak, 'Adaptive contact elements for three-dimensional explicit transient analysis', *Comp. Methods Appl. Mech. Eng.*, **72**, 125–151 (1989).
12. D. P. Bertsekas, *Constrained Optimization and Lagrange Multiplier Methods*, Academic Press, New York, 1982.
13. D. P. Flanagan and T. Belytschko, 'A uniform strain hexahedron and quadrilateral with orthogonal hourglass control', *Int. j. numer. methods eng.*, **17**, 679–706 (1981).
14. M. O. Neal, 'Contact-impact by the pinball algorithm with penalty, projection, and augmented Lagrangian methods', *Ph.D. Dissertation*, Northwestern University, Evanston, IL, 1989.
15. D. T. Greenwood, *Principles of Dynamics*, Prentice-Hall, Englewood Cliffs, N.J., 1965, pp. 153–155.
16. D. P. Flanagan and T. Belytschko, 'Eigenvalues and stable time steps for the uniform strain hexahedron and quadrilateral', *J. Appl. Mech. ASME*, **51**, 35–40 (1984).
17. T. Belytschko and L. P. Bindeman, in preparation.
18. S. P. Timoshenko and J. N. Goodier, *Theory of Elasticity*, McGraw-Hill, New York, 1934, pp. 420–422.
19. D. J. Benson and J. O. Hallquist, 'A single surface contact algorithm for the postbuckling analysis of shell structures', *Report* to the University of California at San Diego, CA, 1987.
20. M. Oldenburg, 'Finite element analysis of thin walled structures subjected to impact loading', *Doctoral Thesis*, Lulea University of Technology, 1988.
21. T. Belytschko, J. I. Lin and C.-S. Tsay, 'Explicit algorithms for the nonlinear dynamics of shells', *Comp. Methods Appl. Mech. Eng.*, **42**, 225–251 (1984).
22. R. E. Sarwas, 'Hidden line elimination and a modified pinball algorithm for finite element contact-impact problems with shell elements', *M.S. Thesis*, Northwestern University, Evanston, IL, 1990.

INTERNATIONAL JOURNAL FOR NUMERICAL METHODS IN ENGINEERING, VOL. 31, 573–592 (1991)

SOLID ELEMENTS WITH ROTATIONAL DEGREES OF FREEDOM: PART 1—HEXAHEDRON ELEMENTS

SHAH M. YUNUS AND TIMOTHY P. PAWLAK

Swanson Analysis Systems, Inc., P.O. Box 65, Johnson Road, Houston, Pennsylvania, U.S.A.

ROBERT D. COOK

Department of Engineering Mechanics, University of Wisconsin-Madison, Madison, Wisconsin 53706, U.S.A.

SUMMARY

This is the first of a two part paper on three-dimensional finite elements with rotational degrees of freedom (DOF). Part I introduces an 8-node solid hexahedron element having three translational and three rotational DOF per node. The corner rotations are introduced by transformation of the midside translational DOF of a 20-node hexahedron element. The new element produces a much smaller effective band width of the global system equations than does the 20-node hexahedron element having midside nodes.

A small penalty stiffness is introduced to augment the usual element stiffness so that no spurious zero energy modes are present. The new element passes the patch test and demonstrates greatly improved performance over elements of identical shape but having only translational DOF at the corner nodes.

INTRODUCTION

The successful use of corner rotation to improve the performance of two-dimensional (2-D) membrane elements has been presented in References 1–5. To further enhance these 2-D elements a hybrid formulation has also been studied in References 4 and 6–8. In Reference 8 a three-dimensional (3-D) hybrid element with rotational DOF is developed.

The present paper develops an 8-node displacement-based hexahedron including translations and rotations as nodal DOF for 3-D elasticity problems. The intent is to introduce an 8-noded element that is computationally attractive when compared with the 20- or 27-node hexahedron elements with translations only as DOF.

Although the added rotational DOF decrease the computational efficiency as compared with the conventional 8-node hexahedron, improved accuracy outweighs this shortcoming. Other attempts to improve the performance of the 8-node hexahedron by use of incompatible modes[9,10] or selective choice of stress or strain fields[11] have been reasonably successful. However, such elements[9–11] are found to be highly sensitive to the element configuration, producing the best results when in right hexahedron form. The relative computational cost of various element types for large 3-D meshes is approximately proportional to the cubic power of the DOF added by each element. In large meshes each additional element is introduced by adding only one corner node, and a relative estimate of CPU times is given in Table I.

The proposed hexahedron element is developed by transforming the midside nodal DOF of a 20-noded isoparametric element in terms of corner nodal translations and rotations. The inclusion of corner rotational DOF allows the normal motion along element edges to be quadratic while the tangential displacements remain only linear. Such discrepancy in tangential

0029–5981/91/030573–20$10.00

Received 25 October 1989
Revised 24 April 1990