ENERGY CONSERVING ALGORITHMS FOR DYNAMIC CONTACT ANALYSIS USING NEWMARK METHODS

H. Zolghadr Jahromi\textsuperscript{1}, B.A. Izzuddin\textsuperscript{2}

Abstract

In this paper, two numerically robust algorithms based on Lagrangian velocity constraint and regularised penalty method are proposed for frictionless dynamic contact analysis utilising the standard Newmark method. It is shown that the Lagrangian velocity constraint method achieves energy conservation and a zero gap constraint approximately, where the errors can be reduced by temporal refinement. In addition to this, a regularised penalty force is devised to ensure exact energy conservation for frictionless contact analysis using the trapezoidal rule. Therefore, unlike the conventional penalty method, the proposed method achieves better accuracy in the impenetrability constraint whilst maintaining the energy stability with no need for extra refinement of the temporal discretisation.

Keywords: Energy conservation; Newmark method; Dynamic contact; Impact; Nonlinear structural analysis; Finite element method

\textsuperscript{1} Research student, Dept. Civil & Env. Eng’g, Imperial College London SW7 2AZ, hamed.jahromi08@imperial.ac.uk.

\textsuperscript{2} Professor of Computational Structural Mechanics, Dept. Civil & Env. Eng’g, Imperial College London, SW7 2AZ, Corresponding author: b.izzuddin@imperial.ac.uk.
1. Introduction

The modelling of highly nonlinear dynamic phenomena in solid and structural mechanics has been of extensive interest over recent years (e.g. [1-3]). In particular, the numerical simulation of dynamic contact problems has been of a great importance due to its frequent application to a wide range of engineering problems. Towards this end, several methods have been developed based on Lagrangian multipliers or penalty functions in an attempt to impose the impenetrability condition of contact analysis.

The kinematic constraints of impenetrability used in finite element analysis procedures work well in conjunction with implicit integration schemes such as the well-known Newmark [4] average acceleration (so called ‘trapezoidal rule’), the mid-point rule (e.g. [1]) and Hilber-Hughes-Taylor methods [5]. While the trapezoidal and mid-point rules are unconditionally stable and energy conserving in linear elastic analysis, much work can be found in the literature regarding the loss of these characteristics for dynamic contact analysis (e.g. [6-7]). It is now well established that, instead of the necessary spectral radius criteria used in linear analysis, a sufficient condition for unconditional stability in nonlinear analysis is realised through the conservation or decay of the total energy within a time-step [8] given as:

\[
U_{n+1} - U_n + V_{n+1} - V_n \leq W_{n+1}^{ext} - W_n^{ext} \quad \text{or} \quad [U]_{n+1}^n + [V]_{n+1}^n \leq [W]_{n+1}^n
\]

where, the variables \( U_i \), \( V_i \) and \( W_i \) are respectively the strain energy, kinetic energy and the work done by the loads at time \( t_i \) for the given temporal discretisation \( \bigcup_{n=0}^{m-1} \{ t_n, t_{n+1} \} \) of the time interval of interest \( T \). And \( m \in \mathbb{N} \) represents the numbers of analysis time-steps. Note that the notation \( [ ]_{n}^{n+1} \) shown in Eq. (1) is defined in this
paper as the incremental difference of the variable within the time-step \([t_n, t_{n+1}]\).

(i.e. \(\chi^{n+1}_n = \chi_{n+1} - \chi_n\)).

Generally, the conservation or decay of the total energy within a time-step can be achieved via three main strategies [8-9], namely a) numerical dissipation, b) enforced energy conservation, c) algorithmic energy control. A brief review on each of these approaches, their relative benefits and the associated shortcomings for application in nonlinear dynamic analysis is provided hereafter.

The first strategy makes use of numerical dissipation to damp out the energy gain under the high frequency response in linear structural analysis. However, while the numerical dissipation of the high frequency range can lead to better accuracy and improved stability in some problems, Erlicher et al. [10] showed that undesirable overshoot and oscillation in the energy can occur in nonlinear analysis due to inaccuracies in the response of the intermediate frequency range. Therefore, numerical dissipation should be considered only as a desired characteristic but not a guarantee for stability in nonlinear dynamic analysis.

The enforced energy conservation was originally developed by Hughes et al. [11] extending the trapezoidal algorithm to achieve energy conservation for nonlinear dynamic analysis. In this approach the energy stability is achieved via a Lagrange multiplier, where the desired energy characteristic is introduced into the equation of total potential energy as a constraint. Hughes et al. [11] originally applied the energy constraint method for dynamic analysis of a hyper-elastic 1D rod problem under free oscillation. It was shown that the modification in the trapezoidal rule results in physically correct energy growth characteristics. However, a study carried out by Kuhl and Ramm [12] illustrated that enforcing energy conservation via a multiplier leads to
final failure in the Newton-Raphson iterations for equilibrium. Furthermore, it was shown that the failure in the numerical procedure occurs at the same stage of the analysis as for the conventional unconstrained algorithm. In this regard, the only benefit of the energy constraint methods can be considered to be firstly the conservation of exact energy using larger time-steps, and secondly the identification of loss of stability through numerical failure in the solution procedure.

Recently, several algorithms based on Lagrangian multipliers [7] and penalty functions [6] were proposed for dynamic contact analysis using the ‘Energy-Momentum Method’ (so called ‘EMM’) introduced by Simo et al. [1]. However, the ‘EMM’ employed and developed previously by researchers (see e.g. [13-15] ) requires solving for a scalar variable either at the integration points over each element at the mid-point of each time-step. Therefore, these methods can become computationally costly and require special care for inclusion in typical nonlinear finite element analysis procedures (see e.g. [16]).

In view of the above, specific consideration is given in this paper to different algorithms enforcing the contact constraints using the Newmark family of methods, which enable energy conservation for frictionless dynamic contact problems. Two algorithms are proposed, employing a Lagrangian velocity constraint and regularised penalty approach, which are shown to ensure unconditional stability and lead to improved accuracy and convergence with temporal and/or spatial refinement.

2. The Newmark method

The well-known Newmark method [4] is a single-step implicit time integration scheme commonly used for dynamic analysis of structures. The method can be viewed as a truncated formulation of a Taylor’s series for the displacement and velocity at time
\( t_{n+1} = \Delta t + t_n \), formed about time \( t_n \). The parametric difference equation utilised in Newmark method can be written as follows:

\[
\begin{align*}
\{d_{n+1}\} &= \{d_n\} + \Delta t \{\ddot{d}_n\} + \Delta t^2 \left( \frac{1}{2} - \beta \right) \{\dot{d}_n\} + \beta \{\ddot{d}_{n+1}\} \\
\{\ddot{d}_{n+1}\} &= \{\ddot{d}_n\} + \Delta t \left( (1-\gamma) \{\dot{d}_n\} + (\gamma) \{\ddot{d}_{n+1}\} \right)
\end{align*}
\]

where \( \Delta t = t_{n+1} - t_n \) is the time-step for a given temporal discretisation, \( \{d_n\} \) is the nodal displacement vector evaluated at \( t_n \) (i.e. \( \{d_n\} = \{d(t_n)\} \)), and each superimposed dot represents a time differentiation. \( \beta \) and \( \gamma \) are algorithm parameters, which define the characteristics of the method in terms of accuracy, numerical dissipation and stability.

### 2.1 Incremental energy balance

To investigate the numerical stability of the Newmark method with respect to the dynamic contact inherent nonlinearity, the incremental energy balance is considered for nodal contact of a linear elastic system. For this purpose, assuming a constant symmetric mass matrix during the analysis time interval (\( T \)), the incremental energy balance is written in terms of mean values of the applied force vector and increments of the nodal kinematic vectors, as given by:

\[
[U + V]^{n+1} = (\Delta d)^T \left\{ \frac{f_{n+1}}{2} + \{f_n\} + \left( \frac{1}{2} - \frac{\gamma}{2} \right) \{\Delta f\} \right\} - \left( \beta - \frac{\gamma}{2} \right) \Delta t^2 \left[ \{\ddot{d}\}^T [M] \{\ddot{d}\} \right]^{n+1} \\
- \left( \frac{1}{2} \right) \{\Delta d\}^T [K] \{\Delta d\} + \left( \beta - \frac{\gamma}{2} \right) \Delta t^2 \{\Delta \ddot{d}\}^T [M] \{\Delta \ddot{d}\}
\]

where for all \( i \in \bigcup_{\epsilon=0}^n \{x\} \) the discretised equation of motion is expressed as:

\[
[M] \{\ddot{d}_i\} + [K] \{d_i\} = \{f_i^{\text{ext}}\} + \{f_i^{\text{contact}}\} = \{f_i\}
\]
According to Eqs. (4) and (5), for the case of free vibration (i.e. \( \{ f_n \} = \{ f_{n+1} \} = 0 \)) the incremental energy formulation for \( \gamma = 0.5 \) (i.e. no numerical dissipation), becomes:

\[
[U + V]^{n+1}_n = -\left[ \frac{1}{2} (\beta - 0.25) \Delta t^2 \{ \ddot{d} \}^T M \{ \ddot{d} \} \right]^{n+1}_n
\]  \hspace{1cm} (6)

It is clear from Eq. (6) that for \( 2\beta = \gamma = 0.5 \) (i.e. average acceleration scheme or trapezoidal rule) the total mechanical energy is conserved for a linear elastic system under free vibration. However, the response of this system using Newmark parameters \( \gamma = 0.5 \) and \( \beta \neq \frac{\gamma}{2} \) will suffer from a periodic energy fluctuation. This energy oscillation is caused by the error in estimation of the velocity and displacement which can cause the so called ‘velocity overshoot’ and ‘displacement overshoot’ for cases of \( \beta > \frac{\gamma}{2} \) and \( \beta < \frac{\gamma}{2} \), respectively [17]. Similarly, the Newmark method with \( \gamma \neq 0.5 \) suffers from energy variation, leading to a conditionally stable algorithm for \( \gamma < 0.5 \) and only first order accuracy for \( \gamma > 0.5 \). Therefore, in this study, special consideration is given to the non-dissipative Newmark methods (\( \gamma = 0.5 \)) and particularly the trapezoidal rule to achieve a numerically stable algorithm for frictionless dynamic contact analysis.

3. Normal contact constraints

The simplest type of contact interaction is the case of frictionless contact, where the only non-zero traction component is normal to contact surface of the contacting boundaries. Several methods have been developed over the past decades to introduce kinematic constraints associated with the contact phenomena. Many of these studies attempt to model the dynamic contact by enforcing a constraint on the displacements (so
called ‘gap constraint’) of the nodes of a contacting body (so called ‘master body’) and
the associated nodes/surface of an impacted body (so called ‘slave body’). The
conventional penalty spring and Lagrangian displacement constraint are typical
examples of this.

3.1 Lagrangian displacement constraint

In this method, the so called ‘Kuhn-Tucker’ kinematic conditions are introduced
for frictionless dynamic contact analysis via a multiplier into the equation of total
potential energy [18-19]. These kinematic conditions for normal contact can be written as:

\[
\begin{align*}
(i) \quad & g_N(d^m_i, d^s_i)_\kappa = \bar{g}(d^m_i, d^s_i)_\kappa \bar{\pi}_\kappa \\
(ii) \quad & f^m_{i, \text{contact}} = \tau(d^m_i)_\kappa \bar{\pi}_\kappa \\
(iii) \quad & \left\{ g_N(d^m_i, d^s_i)_\kappa \right\}^T \cdot \left\{ f^m_{i, \text{contact}} \right\} = 0
\end{align*}
\]

where \( \bar{g}(d^m_i, d^s_i)_\kappa \) represents the closest distance gap vector between a contacting pair
(defined by subscript \( \kappa \)) of the master node \( d^m_i \) and its corresponding contact
node/surface of the slave body \( d^s_i \) or \( \sum N(\zeta, \eta) d^s_i \) evaluated at time \( t_i \) of the
analysis. It should be noted that Eq. (7(i)) will imply \( g_N(d^m_i, d^s_i)_\kappa \) to be positive where
penetration occurs between the contacting boundaries. The function \( \tau(d^m_i)_\kappa \) is the
traction pressure (i.e. positive) applied to the master node in contact at time \( t_i \) and \( \bar{\pi}_\kappa \) is
the unit outward normal to the tangential vector at the current contact slave surface of
the contacting pair considered. Fig. 1 shows a schematic view of the contact between
two bodies with a reference placement \( \Omega_\alpha \ (\alpha = 1, 2) \) undergoing the deformation
\( \phi_\alpha : \Omega_\alpha \times [0, T] \rightarrow \mathbb{R}^{n_{\text{dim}}} \ (n_{\text{dim}} = 1, 2 \ or \ 3) \).
Fig. 1. Schematic view of the contact between two bodies.

Therefore, the total potential energy of the system \( \Pi \), considering the contact constraint can be written as:

\[
\Pi = U + W + \Lambda_d
\]  

where:

\[ U = \text{stored strain energy} = \int_V \left( \int d\varepsilon^T \sigma \right) dV \]  

\[ W = \text{loss of potential energy corresponding to the applied load} = -\int_S (N^T d_i^T p) ds \]  

\[ \Lambda_d = \text{impenetrability constraint} = \sum_{\kappa=1}^q g_N \left( d_i^m, d_i^l \right)_\kappa \lambda_{\kappa} (t_i) = 0 \]  

Note that \( q \) is the total number of master nodes designated as candidates for contact in the search algorithm [20] and the subscript \( \kappa \) denotes the gap and contacting force associated with each contacting pairs. By comparing the Eq. (11) with the normality condition of Eq. (7), it is clear that the multipliers \( \lambda_{\kappa} (t_i) \) represent the normal traction force applied to the corresponding master node at time \( t_i \).

It can be shown that the linearization of the equilibrium between the external forces and internal stresses (using for example the minimum total potential energy principle) leads to a zero diagonal term for each multiplier term [18]. Therefore, for simple contact problems, the use of ‘constraint elimination’ can be more convenient.
3.1.1 **Newmark method**

To investigate the energy stability of the Newmark family of methods using Lagrangian displacement constraints, the energy variation of a non-accelerating particle (i.e. $a_0 = 0$) with initial velocity (i.e. $v_0 \neq 0$) is considered upon its impact with a rigid surface, as depicted in Fig. 2. The variable $\delta \in [0,1]$ shown in Fig. 2 defines the fraction of the analysis time-step in which the actual contact will occur.

![Figure 2. Rigid impact of a non-accelerating particle.](image)

Figure 3 illustrates the energy gain/loss using second order accurate non-dissipative Newmark family method ($\gamma = 0.5$) and considering the unconditional-stability criteria ($\beta \geq \gamma / 2$). It is demonstrated that the trapezoidal rule method ($2\beta = \gamma = 0.5$) combined with Lagrangian displacement constraint can lead to an erroneous energy increase of up to 9 times the initial value prior to contact. Chaudhary and Bathe [21] first suggested that the Newmark method with $\beta = \gamma = 0.5$ (so called ‘Newmark-0.5’) is an effective time integration method for numerical analysis of dynamic contact problems, if the time-step employed is sufficiently small. This is shown in Fig. 3 for rigid impact of non-accelerating impacting point masses (stiffness-free) for Newmark-0.5 method. In addition to this, another non-dissipative Newmark
algorithm with $\beta = 2\gamma = 1$ (so called ‘Newmark-1’) is shown to conserve the total energy for the considered system.

![Diagram](image.png)

**Fig. 3.** Energy gain/loss of particle impact with rigid surface: Newmark ($\gamma = 0.5$).

However, as mentioned previously, the responses obtained using Newmark parameters $\beta \neq \frac{\gamma}{2}$ and $\gamma = 0.5$ will suffer from a periodic energy fluctuation for discrete solids caused by the error in estimation of the velocity or displacement. Therefore, only the trapezoidal rule is considered for further analytical investigation.

### 3.1.2 Analytical investigation of the energy for Newmark method

The analytical study of the total incremental energy for nodal contact of a linear elastic system, as expressed by Eq. (4), shows that the energy gain/loss for the non-dissipative Newmark methods ($\gamma = 0.5$) will occur in two time-steps: 1) initial contact $[t_n, t_{n+1}]$ and 2) rebound $[t_{n+k}, t_{n+k+1}]$, where $k$ is the number of time-steps in persistent contact. From Eq. (4), the sum of incremental energy over the contact duration $[t_n, t_{n+k+1}]$ for one of the contacting bodies can be expressed as:
\[
\sum_{z=0}^{z_{\mathit{zk}}} [U + V]_{n+z} = \frac{1}{2} \left[ \begin{bmatrix} d_{n+z}^{m} \\ d_{n+z}^{s} \end{bmatrix} \right]^T \cdot \left\{ f_{n+1} \right\} + \frac{1}{2} \left[ \begin{bmatrix} d_{n+k}^{m} \\ d_{n+k}^{s} \end{bmatrix} \right]^T \cdot \left\{ f_{n+k} \right\} + \left( \beta - \frac{1}{4} \right) \Delta t^2 \left\{ \ddot{a}_n \right\} M \left\{ \ddot{a}_n \right\} - \left( \beta - \frac{1}{4} \right) \Delta t^2 \left\{ \ddot{a}_{n+k+1} \right\} M \left\{ \ddot{a}_{n+k+1} \right\}
\] (12)

Equation (12) shows analytically the extent and source of energy instability upon initial contact and release for a linear elastic system. For \( \beta = 0.25 \) and considering both contacting bodies, where the contact forces on the two bodies are equal and opposite (i.e. \( f_{n+1}^s = -f_{n+1}^m \)), the following incremental energy is obtained:

\[
\sum_{z=0}^{z_{\mathit{zk}}} [U + V]_{n+z} = -\frac{1}{2} \left\{ g_N \left( d_{n}^{m}, d_{n}^{s} \right) \right\}^T \cdot \left\{ \lambda \left( t_{n+1} \right) \right\} + \frac{1}{2} \left\{ g_N \left( d_{n+k}^{m}, d_{n+k}^{s} \right) \right\}^T \cdot \left\{ \lambda \left( t_{n+k} \right) \right\}
\] (13)

From Eq. (13), it can be concluded that the source of energy loss is associated with the initial contact whereas the energy gain occurs at the rebound. Furthermore, the energy variation cannot be reduced by temporal refinement, since the traction force and normal gap at the initial and rebound time-steps have inverse and direct relationships, respectively, with the time-step. Bearing this in mind, for persistent contact analysis of a discretised multi-degree of freedom system, the enforcement of the gap constraint via a Lagrange multiplier leads to severe velocity and displacement oscillation of the contacting nodes which becomes more severe as the spatial mesh is refined (e.g. [6]).

While spatial refinement reduces the value of the contacting masses and hence the magnitude of the contact forces, the effect of discretisation is that persistent contact transforms into multiple contacts over the contact duration. Accordingly, for the trapezoidal rule, the cumulative energy gain/loss upon each contact will eventually cause numerical instability of the response which cannot be resolved by refinement of spatial and/or temporal discretisation, as illustrated in a later example.
3.2 Penalty method

The penalty method is an alternative approach for enforcing contact constraints which is computationally straightforward and is commonly used in finite element analysis programs. In this method, the normal traction force emerges as a result of a high stiffness spring located at the point of contact to prevent significant penetration. To introduce the penalty constraint into the equation of the total potential energy, the Lagrangian multiplier function in Eq. (8) (i.e. $\Lambda_d$) is replaced by a penalty spring energy function (i.e. $\Lambda_p$):

$$\Pi = U + W + \Lambda_p$$

where:

$$\Lambda_p = \text{penalty gap constraint} = \frac{1}{2} \sum_{\kappa=1}^{q} K_{\kappa}^p g_{\kappa} \left( d_{m}^{i}, d_{s}^{i} \right)^2,$$

$K_{\kappa}^p$ is the stiffness of the penalty spring, the subscript kappa ($\kappa$) denotes the gap and penalty stiffness associated with each pair of the contacting nodes/segments, and $q$ is the total number of master nodes designated as candidates for contact in the search algorithm.

The penalty method also suffers from a non-physical energy gain/loss upon contact, though, as discussed later in detail, better accuracy in energy conservation may be achieved through temporal refinement. In terms of computational solution procedure, the penalty method has the advantage of avoiding the difficulties in dealing with zero diagonal terms in the system of equations compared to the Lagrangian multiplier approach. However, the main drawback of this method is that small penetrations between the impacting bodies are inevitable. Therefore, the reliability and accuracy of the results depends significantly on the assumption made for the contact spring stiffness.
and the analysis time-step size. There is also the potential for ill-conditioning with the use of excessively large penalty stiffness, leading to sensitivity of the solution to round-off errors.

3.2.1 Analytical investigation of energy conservation with Newmark trapezoidal rule

Similar to the Lagrangian displacement constraint, undesirable energy alteration is observed for contact analysis with the use of trapezoidal rule and the penalty spring. The incremental energy variation for a linear elastic system over \( k \) steps of persistent contact (i.e. \([t_n, t_{n+k+1}]\)) is expressed below:

\[
\sum_{z=0}^{z=k} [U + V]_{n+z} = \frac{1}{2} \left[ g_N \left( \mathbf{d}^m_{n}, \mathbf{d}^r_{n} \right) \right]_{n+z}^T \{ K^p g_N \left( \mathbf{d}^m_{n+1}, \mathbf{d}^r_{n+1} \right) \} + \frac{1}{2} \left[ g_N \left( \mathbf{d}^m_{n+k}, \mathbf{d}^r_{n+k} \right) \right]_{n+k}^T \{ K^p g_N \left( \mathbf{d}^m_{n+k+1}, \mathbf{d}^r_{n+k+1} \right) \} \tag{16}
\]

According to Eq. (16), the energy balance upon impact is related to the variation of the gap function for the initial contact \( g_N \left( \mathbf{d}^m_{n}, \mathbf{d}^r_{n} \right) \) and rebound \( g_N \left( \mathbf{d}^m_{n+k}, \mathbf{d}^r_{n+k} \right) \) as well as the stiffness of the penalty spring. Therefore, with the use of suitable penalty stiffness, better accuracy in energy conservation and contact response can be achieved through refinement of analysis time-steps. However, defining appropriate penalty stiffness remains the main drawback of this method.

4. Proposed methods

Encouraged by some previous studies [6-7, 22-23], Laursen and Chawla [7] first used the Energy-Momentum Method [1] combined with a Lagrangian velocity constraint (zero gap-rate constraint). It was shown that the method achieves conservation of both energy and momentum for frictionless dynamic contact problems.
Nevertheless, the method still suffers from the requisite condition for evaluating the internal forces at the integration points or over each element in an average form, which makes this method difficult to apply in typical finite element analysis programs and renders it computationally costly.

The focus of the present work is to use the Lagrangian velocity constraint as well as a newly developed regularised penalty formulation for suitable and effective time integration schemes such as Newmark method, with the aim of achieving a stable and accurate response for the analysis of dynamic contact problems.

4.1 Lagrangian velocity constraint

As discussed before, the numerical instability observed with the Lagrangian displacement multiplier is caused by the cumulative energy gain during multiple contacts of a discretised system. For the persistent contact case, this energy instability can be linked to the lack of kinematic constraints which cause severe velocity and displacement oscillation of the contacting nodes. Accordingly, by differentiation of the normality ‘Kuhn-Tucker’ condition with respect to time and holding the equality for a zero gap constraint, a condition for zero velocity-gap (i.e. gap-rate) is achieved for persistent dynamic contact analysis. The zero gap-rate constraint, referred to in this paper as velocity constraint, is generally not captured with time-step algorithms using the Lagrangian displacement constraint or the penalty springs. In this study, a velocity constraint multiplier is suggested with the trapezoidal rule to achieve a robust algorithmic solution procedure for frictionless dynamic contact analysis, which is expressed as:

\[ r_x \left( \ddot{d}_i^n, \dot{d}_i^r, d_i^n, d_i^r \right)_x \dot{\lambda}_x(t_i) = 0 \]  

(17)
where \( r_N \left( \mathbf{d}^m_i, \mathbf{d}^n_i, \mathbf{d}^m_i, \mathbf{d}^n_i \right) \) is the normal gap-rate function for the pair of contacting segments \( \kappa \) evaluated at time \( t_i \), and \( \lambda_\kappa (t_i) \) is the corresponding Lagrangian multiplier.

The Lagrangian velocity constraint is thus introduced into the equation of total potential energy as follows:

\[
\Pi = U + W + \Lambda_v
\]

where:

\[
\Lambda_v = \text{persistency constraint} = \sum_{\kappa=1}^d r_N \left( \mathbf{d}^m_i, \mathbf{d}^n_i, \mathbf{d}^m_i, \mathbf{d}^n_i \right)_\kappa \lambda_\kappa (t_i) = 0,
\]

\[
r_N \left( \mathbf{d}^m_i, \mathbf{d}^n_i, \mathbf{d}^m_i, \mathbf{d}^n_i \right)_\kappa = \text{normal gap-rate} = \overline{\mathbf{r}} \left( \mathbf{d}^m_i, \mathbf{d}^n_i, \mathbf{d}^m_i, \mathbf{d}^n_i \right)_\kappa \overline{\mathbf{r}}
\]

and the condition for contact is identified based on both displacement and gap-rate penetration between the boundaries. In contrast to the Lagrangian gap-constraint approach, where the multiplier \( \lambda_\kappa (t_i) \) represents the traction force for the corresponding segments in contact, the traction force vector in the velocity constraint method can be obtained by differentiating the persistency constraint (i.e. \( \Lambda_v \)) with respect to the master node’s displacement vector. This is shown in equation below with further simplification using the chain rule:

\[
\{ \mathbf{Z}^m \} = \left\{ \frac{\partial \Lambda_v}{\partial \mathbf{d}^m_i} \right\} \otimes \left\{ \frac{\partial \mathbf{d}^m_i}{\partial \mathbf{d}^m_i} \right\} + \left\{ \frac{\partial \Lambda_v}{\partial \mathbf{d}^m_i} \right\}
\]

In the above, \( \{ \mathbf{Z}^m \} \) represents the applied traction forces to the master nodes normal to the current slave surface, and operator \( \otimes \) represents the entrywise vector product.
4.1.1 Analytical investigation of energy conservation with Newmark trapezoidal rule

The analytical energy balance for nodal contact of a linear elastic system between
the time interval \([t_n, t_{n+k+1}]\) is expressed below for the trapezoidal rule with the
persistency constraint:

\[
\sum_{z=0}^{z=k} [U + V]_{n+z} = -\frac{1}{2} \left\{ r_N \left( \dot{d}_n^m, \dot{d}_n^r, \dot{d}_n^m, \dot{d}_n^r \right) \right\}^T \left\{ \lambda \left( t_{n+1} \right) \right\} \\
- \frac{1}{2} \left\{ r_N \left( \dot{d}_{n+k+1}^m, \dot{d}_{n+k+1}^r, \dot{d}_{n+k+1}^m, \dot{d}_{n+k+1}^r \right) \right\}^T \left\{ \lambda \left( t_{n+k} \right) \right\}
\]

Based on Eq. (22), it is observed that the energy variation is related to the
multipliers and the gap-rate at time-steps \(t_n\) and \(t_{n+k+1}\). Noting that the widely used
Newmark method inherits the conservation of linear momentum characteristics, the
following can be written:

\[
\int_{t_n}^{t_{n+k}} \mathcal{F} \cdot d\mathbf{t} = ct. \quad (23)
\]

Therefore, based on Eqs. (21) and (23), refinement in the temporal discretisation
results in a reduction of the multiplier and hence better energy conservation.
Additionally, it can be shown that, with the zero gap-rate constraint, temporal
refinement leads to better accuracy in the impenetrability constraint. These desirable
characteristics for the Lagrangian velocity constraint approach are highlighted in several
numerical examples along with the advantages and disadvantages of the previously
discussed methods based on displacement constraints.

4.2 Regularised penalty method

Based on the energy balance formulation shown in Eq. (16), and motivated by the
regularised penalty force developed for the Energy-Momentum Method [6], a new
energy conserving algorithm is proposed here for the trapezoidal rule using the regularised penalty force applied to the master node. This is expressed as:

\[
\begin{align*}
\text{For } H\left(g_{N,n+1}\right)_\kappa &\neq H\left(g_{N,n}\right)_\kappa \\
 f_{m,\text{contact}}^{n+1}_\kappa &= \frac{K^p\left(g_{N,n+1}\right)_\kappa . H\left(g_{N,n+1}\right)_\kappa - K^p\left(g_{N,n}\right)_\kappa . H\left(g_{N,n}\right)_\kappa}{\left(g_{N,n+1}\right)_\kappa - \left(g_{N,n}\right)_\kappa} - f_{m,\text{contact}}^{n}_\kappa \\
\text{For } H\left(g_{N,n+1}\right)_\kappa &= H\left(g_{N,n}\right)_\kappa \\
 f_{m,\text{contact}}^{n+1}_\kappa &= K^p\left(g_{N,n+1}\right)_\kappa . H\left(g_{N,n+1}\right)_\kappa + \left(g_{N,n}\right)_\kappa . H\left(g_{N,n}\right)_\kappa - f_{m,\text{contact}}^{n}_\kappa
\end{align*}
\]

(24)

where, for mathematical simplification, \( g_{N,i} \) stands for \( g_N\left(d_i^n, d_i^r\right)_\kappa \) and \( H \) represents the common form of unit-step Heaviside function:

\[
H\left(x\right) = \begin{cases} 
0, & x \leq 0 \\
1, & x > 0 \\
\forall x \in \mathbb{R}
\end{cases}
\]

(25)

The algorithm deals with the energy gain and loss at initial contact and rebound by regularising the contact force. According to Eq. (24), the proposed regularised penalty method with the well-known trapezoidal rule is devised such that it conserves precisely the total energy of a linear elastic system upon impact regardless of the analysis time-step size. This can be easily verified by inserting the contact force obtained in Eq. 24, at both initial and rebound stage, into the equation of total potential energy given by Eq. (13). The precise algorithmic energy conservation with the proposed regularised algorithm ensures numerical stability, and therefore achieves better accuracy in the impenetrability constraint with a larger time-step size compared to the conventional penalty method. These desirable characteristics are demonstrated in the next section with some illustrative numerical examples.
5. Numerical examples

Two examples are considered here to demonstrate the numerical robustness, applicability and computational superiority of the proposed algorithms compared with conventional algorithms.

Regarding computational efficiency, it is noted that the computational cost of the contact procedures per time step is relatively small compared to that of the finite element computations, regardless of the contact algorithm. Accordingly, the relative overall efficiency of the dynamic contact simulations can be measured in terms of the time-step required for accuracy and stability. Since the CPU time is inversely proportional to the time-step size, particularly when the step is sufficiently small to require one iteration for nonlinear analysis, the comparison of computational time is based hereafter on direct comparison of the time-step sizes.

5.1 Rod impact

A simple but insightful impact problem of two identical elastic rods, as depicted in Fig. 4, is considered to illustrate the relative advantages of the Lagrangian velocity method and regularised penalty method. No initial applied loading is considered, and the rods are uniformly discretised along their length \( L = 1.0m \) by 2-node 1D linear elastic elements. The rods have a unit cross section \( A = 1.0m^2 \) as well as unit Young’s modulus \( E = 1.0N/m^2 \) and material density \( \rho = 1.0kg/m^3 \).
A program is developed in MATLAB for the analysis of 1D impacting rod problems allowing for spatial and temporal discretisation. The program is written for the general time-integration schemes using a Newton-Raphson solution procedure, where focus is given here to the Newmark family of methods for temporal discretisation. Consistent mass matrices are considered here, though it was observed that analogous responses are obtained with consistent and lumped mass matrices for a sufficiently fine discretisation [24]. Convergence criteria are based on the condition \( \left\{ G^{(j+1)} \right\}^T \left\{ G^{(j+1)} \right\} < \varepsilon \), where \( \left\{ G^{(j+1)} \right\} \) is the vector of out of balance in the current iteration \((j+1)\) and the tolerance parameter \( \varepsilon \) is chosen as \(10^{-7}\).

5.1.1 Conventional methods

Some of the various integration schemes and contact algorithms previously described are considered for this numerical study to support the analytical studies on the energy instability of conventional methods and to highlight the advantages of the proposed methods. The analysis time-steps for the numerical simulations presented for the rod example are considered to be less than 12% of the minimum period obtained from eigenvalue analysis of individual elements \(T_{\text{min,el.}}\) for all numerical simulations.
[25]. For spatial discretisation, linear elastic elements with $\frac{L}{n_{el.}}$ element length are considered, where $n_{el.}$ represents the number of elements used for modelling each rod.

5.1.1.1 Lagrangian displacement constraint

To elaborate on the numerical instability associated with the use of the trapezoidal rule and a Lagrangian gap constraint, the two rods are discretised with a fine mesh ($n_{el.} = 100$) and a time-step size of $\Delta t = 2 \times 10^{-3} sec$ through the analysis time interval of interest $T = 5 sec$. Figure 5 depicts the displacement, velocity and multiplier history obtained at the contacting ends of the rods as well as the alteration in the kinetic and potential energy of the system during contact. As mentioned before, severe velocity oscillations between the contacting nodes (Fig. 5(b)) leads to cumulative energy gain, standing in this case at almost 400% (i.e. 0.5 vs ~2.5 N.m), and numerical instability (Fig. 5(c)) of the response.
To avoid the high-frequency oscillations between the contacting nodes, the numerically dissipative Newmark algorithm with $\gamma > 0.5$ and $\beta = \frac{(\gamma + 0.5)}{2}$ is considered. It is shown in Fig. 6 that the response obtained with this method achieves energy decay (i.e. numerically stable) because of its high numerical dissipation. However, the non-physical energy decay cannot be achieved for problems involving persistent contacts with a low frequency range oscillations for the contacting segments as well as problems with multiple contacts (e.g. surface to surface contact).

Furthermore, the particular dissipative scheme considered is only first-order accurate in predicting the acceleration term which is undesirable for long duration dynamic analysis.
Fig. 6. Impact response using Newmark method ($\beta = 0.55, \ \gamma = 0.6$) with Lagrangian displacement constraint ($n_{el} = 100, \ \Delta t = 2 \times 10^{-3}$).

5.1.1.2 Conventional penalty method

For the conventional penalty method, non-physical energy variation is also observed using the Newmark algorithm with $2\beta = \gamma = 0.5$ (i.e. trapezoidal rule). However, based on Eq. (16), better accuracy in energy conservation and the dynamic response can be achieved through temporal refinement. This comparison is carried out for the current example, and the responses are shown in Figs. 7 and 8. Note that the energy variation upon contact is considerably reduced by refining the analysis time-step from $\Delta t = 2 \times 10^{-3} \ sec$ (Fig. 7) to $\Delta t = 3 \times 10^{-5} \ sec$ (Fig. 8). However, satisfying both the acceptable penetration criterion ($\max\left(g_N \left(d_i^m, d_i^r\right)\right) \leq \frac{L}{50 \cdot n_{el}}$) and energy stability leads...
to an iterative process with an excessively small analysis time-step size 

\( \Delta t = 3.0 \times 10^{-5} \ll T_{\text{min,el.}} / 20 \).

Fig. 7. Impact response using trapezoidal rule with conventional penalty spring 

\( \left( K^p = 10^6 \text{ N} / m, \ n_{el.} = 100, \ \Delta t = 2 \times 10^{-3} \right) \).
Fig. 8. Impact response using trapezoidal rule with conventional penalty spring

\[
K_p = 10^6 \text{ N/m, } n_{el.} = 100, \Delta t = 3 \times 10^{-5}
\]

Application of the numerically dissipative algorithm with the penalty method leads to a similar response to that previously shown in Fig. 6 and is hence omitted herein.

5.1.2 Proposed methods

The robustness of the proposed Lagrangian velocity constraint and regularised penalty methods as well as their effectiveness in modelling frictionless dynamic contact are illustrated hereafter considering the impacting rods problem.
5.1.2.1 *Lagrangian velocity constraint*

The performance of the proposed Lagrangian velocity constraint method utilising the Newmark average acceleration scheme is considered herein for the frictionless impacting rods problem. Figure 9 depicts the dynamic response obtained using this algorithm, where each rod is modelled with half of the number of elements used in previous analysis (i.e. \( n_{el} = 50 \)).

![Displacement history](image1.png)

![Velocity history](image2.png)

![Energy variation](image3.png)

**Fig. 9.** Impact response using trapezoidal rule with zero gap-rate constraint

\[
\left( n_{el} = 50, \ \Delta t = 2 \times 10^{-3} \right).
\]

Figure 9 clearly shows the robustness of the suggested algorithmic method and its energy conservation characteristic. Furthermore, it is shown in Figs. 10 and 11 that, the
method achieves even better accuracy in terms of energy conservation and dynamic response through refining the mesh and/or reducing the time-step size, respectively.

![Fig. 10. Effect of spatial refinement on energy (\(n_{el.} = 100, \Delta t = 2 \times 10^{-3}\)).](image1)

![Fig. 11. Effect of temporal refinement on energy (\(n_{el.} = 100, \Delta t = 1 \times 10^{-3}\)).](image2)
5.1.2.2  *Regularised penalty method*

The same analysis carried out before for the impacting rods using the conventional penalty method, as shown in Fig. 7, is repeated here with the proposed regularised penalty method so as to highlight its benefits. The accuracy of predicting the energy of the system as well as that of displacement and velocity predictions are illustrated in Fig. 12 using an identical time-step size to the analysis carried out with the conventional penalty method (Fig. 7). Evidently, the regularised penalty method is shown to be superior, achieving energy conservation with a larger time-step than the conventional penalty method.

---

**Fig. 12.** Impact response using trapezoidal rule with regularised penalty method

\[
\left(n_{el} = 100, \Delta t = 2 \times 10^{-3}\right).
\]
5.2 Sphere impact

To further illustrate the advantages of the proposed methods for a realistic impact problem, the impact of a hollow steel sphere with a rigid plane surface is considered, as shown in Fig. 13(a). This problem involves multiple contact points between the boundaries of the sphere and the planar surface as the sphere undergoes significant deformations.

The hollow sphere has an outer diameter of 2.05\(m\) and a uniform thickness \(t = 0.05m\). Typical elastic material properties for steel are considered with Young’s modulus \(E = 210\, GPa\), Poisson’s ratio \(\nu = 0.3\) and material density of \(\rho = 8 \times 10^3\, kg/m^3\). Note that the proposed contact algorithms can be used with finite elements accounting for both geometric and material nonlinearity, though in this example a linear material response is considered solely for the purpose of examining energy conservation in the system. No initial applied loading is considered, and the initial velocity of the thin spherical shell is considered to be equal to 75\(m/sec\).

Fig. 13. Geometric configuration and discretisation of impacting elastic sphere.
To analyse complex structures under impact loading, the conventional and proposed contact algorithms are implemented into the nonlinear FE software ADAPTIC [26]. For the hollow sphere problem considered here, 9-node co-rotational quadrilateral shell elements [27] are used for spatial discretisation.

The sphere is discretised using incremental spherical coordinates $\Delta \theta$ and $\Delta \varphi$ using the 9-noded quadrilateral elements, as shown in Fig. 13(b). With $\Delta \theta = \pi / 22$ and $\Delta \varphi = 2\pi / 20$, two small circular holes closed by a rigid circular plates are considered at the contacting and opposite ends, thus implying a total of 400 elements (Fig. 14). Regarding mass modelling, the row sum procedure is used for lumping the exact consistent mass matrix for each individual curved element based on the equations shown in Fig. 13(b).

Temporal discretisation is again considered with the trapezoidal rule method using the different algorithms for contact analysis. Note that the proposed algorithms possess no numerical dissipation during contact; therefore, numerical stability with the proposed algorithms is not affected by the time-step size. To model impact, contact elements are provided between the sphere nodes on the rings parallel to the x-y plane and the planar surface (Fig. 14). It is shown later that due to the relatively high initial velocity, the sphere experiences high deformation, and several contact points are detected along the different rings. However, due to rotational rigidity imposed by the plates over the circumference of the bottom circular hole, nodes on the second ring are not subject to contact with the planar surface, hence their results are not reported hereafter.
5.2.1 Conventional methods

Although implicit time integration scheme is used throughout, the analysis time-steps for the numerical simulations of the sphere impact problems are considered to be less than 0.5% of the period for radial vibration of the hollow sphere, which is obtained with eigenvalue analysis (Fig. 15) as $T_{\text{radial mode}} = 7.65 \times 10^{-3} \text{ sec}$. This is close to the theoretical value of $T_{\text{radial mode}} = \pi \sqrt{\frac{2 \rho(1-v)}{E}} = 7.25 \times 10^{-3} \text{ sec}$ derived by Love [28] assuming a very thin spherical shell of radius $\bar{r}$.
Fig. 15. Initial configuration (dashed line) and mode shape (solid line) for sphere.

5.2.1.1 Lagrangian displacement constraint

To highlight once more the numerical instability occurring during impact analysis using the trapezoidal rule and a Lagrangian gap constraint, the sphere impact problem is analysed here for the time interval of $T = 0.007\, sec$ using a time-step of $\Delta t = 1 \times 10^{-3}\, sec$. Figure 16 depicts the gap distance, gap-rate and multiplier history obtained at the contacting rings on the sphere parallel to the global x-y plane, as well as the evolution of the energy in the system during multiple contacts. Figure 16(b) shows the velocity shoot-up and severe velocity oscillations between the contacting nodes of the first ring and the surface. This results in cumulative energy gain of around 340%, as shown in Fig. 16(c), which can lead to numerical instability with successive contact.
5.2.1.2 Conventional penalty method

The impact of the sphere is considered here with the conventional penalty contact algorithm, where the penalty spring stiffness for the contact elements attached to the ring nodes is taken as $K^p = 10^{11}\, kN/m$. The analysis is carried out for two time-step sizes, $\Delta t = 1 \times 10^{-5}\, sec$ and $\Delta t = 1 \times 10^{-6}\, sec$, and the results obtained are shown in Figs.
17 and 18, respectively. As discussed in detail before, it is shown that the energy variation can be controlled by refining the time-step; however, satisfying both the acceptable penetration criterion and energy stability requires an iterative process and leads to an excessively small analysis time-step size.

![Graphs showing impact response](image)

**Fig. 17.** Impact response using trapezoidal rule with conventional penalty spring ($K_r = 10^{14} N/m, \Delta t = 1 \times 10^{-5}$).
Fig. 18. Impact response using trapezoidal rule with conventional penalty spring
\[ (K^p = 10^{14} \text{ N/m}, \Delta t = 1 \times 10^{-6}). \]

5.2.2 Proposed methods

In this section, the Lagrangian velocity constraint and regularised penalty methods are used to simulate the sphere impact problem so as to demonstrate their effectiveness in modelling frictionless dynamic contact for realistic problems involving multiple contact points.
5.2.2.1 **Lagrangian velocity constraint**

The performance of the proposed Lagrangian velocity constraint method utilising the Newmark average acceleration scheme is considered herein for the sphere impact problem using $\Delta t = 1 \times 10^{-5}\text{sec}$. It is shown that the sphere remains in contact for a duration of around $0.006\text{sec}$ (Fig. 19), and over this duration the total energy in the system is conserved as should be for normal elastic impact problem. Importantly, this is achieved with a time-step which is more than 10 times that required with the penalty method.
5.2.2.2 Regularised penalty method

The proposed regularised penalty method is used here to model the sphere impact, again using $\Delta t = 1 \times 10^{-5}$ sec. It is observed from the results in Fig. 20 that this method is also numerically robust and achieves energy conservation at a time-step which is more than 10 times that required by the conventional penalty method.
Comparison between the responses of the two energy conserving solutions obtained with regularised penalty method and Lagrangian velocity constraint shows good agreements in terms of the displacement field. Moreover, the two methods compare very well throughout the analysis considering the vertical displacement at the upper end of the sphere, as shown in Fig. 21. The hollow sphere becomes fully compressed at $t = 2.8\,\text{sec}$, which corresponds to a maximum vertical deflection $\Delta$. 

$$ (K_r = 10^{14} \, N/m, \Delta t = 1 \times 10^{-3}) $$
$\Delta_{\text{max}}=0.148m$, The deformed shapes for the impacting sphere, for the three stages of initial contact, fully compressed and rebound, are shown in Fig. 22 with the distribution of the inner longitudinal strain $\varepsilon_{\theta \theta}$.

![Vertical displacement history at the top of the sphere.](image)

**Fig. 21.** Vertical displacement history at the top of the sphere.

![Longitudinal strain distribution and deformed shape of impacting sphere.](image)

**Fig. 22.** Longitudinal strain distribution and deformed shape of impacting sphere.

### 6. Conclusion

It is widely established that the conservation or decay in the total energy of a structural system can ensure the desirable unconditional stability in nonlinear dynamic analysis. In this paper, a thorough energy assessment of the Newmark family of methods is undertaken for dynamic contact of linear elastic structures. It is shown that
with a Lagrangian displacement constraint, the well-known trapezoidal rule can lead to erroneous energy estimation for the case of rigid impact of a non-accelerating particle with a rigid surface, which cannot be redressed by using temporal refinement.

To further investigate energy stability, the incremental energy formulation of the trapezoidal rule method is considered for multi-degree of freedom systems using both the Lagrangian displacement constraint and penalty methods. It is shown analytically for the Lagrangian displacement constraint approach that the cumulative energy error cannot be reduced by further refinement in the time-step or mesh, due to high frequency oscillation in the displacements of the contacting nodes. In contrast, the penalty method can achieve energy stability by temporal refinement, though the main drawback of this method is that small penetrations between the impacting bodies are inevitable. Furthermore, the reliability and accuracy of the results depends significantly on the assumptions made for the contact spring stiffness and the time-step size, while the use of an excessively large penalty stiffness can potentially lead to an ill-conditioned system.

In this study, algorithmic energy conserving methods based on a Lagrangian velocity constraint and regularised penalty stiffness combined with the well-known trapezoidal rule are proposed for robust frictionless dynamic contact analysis. It is shown that the zero gap-rate constraint efficiently achieves an approximate displacement constraint, the magnitude of which can be reduced by temporal refinement. Both methods are also shown to fulfil the conservation law of energy whilst maintaining applicability and efficiency for conventional FE solution procedures. Furthermore, unlike the Lagrangian displacement constraint approach, the methods
achieve better accuracy in the response upon refining the temporal and/or spatial
discretisation.
References


