Application analysis of time stepping methods for impact problems

Dániel Serfőző and Balázs Pere

Department of Applied Mechanics, Faculty of Information Technology and Electrical Engineering, Széchenyi István University, Győr, Hungary

Received: January 17, 2022 • Revised manuscript received: May 20, 2022 • Accepted: May 23, 2022

ABSTRACT

Dynamic contact and impact problems are widely applicable. An accurate solution method for these kinds of problems can be used in many fields of mechanical engineering (e.g., cutting metalwork, cogwheel drives, etc.). However, the proper handling of the contact is problematic, as there emerges a substantial amount of nonlinearity in the displacement field. Therefore, a spurious high frequency oscillation is present in the solution. These oscillations must be avoided, as divergence can easily occur in the contact algorithm due to them. In order to eliminate this effect, the applied numerical method must be chosen and set properly. In this study, a comprehensive guide is provided for the appropriate selection of the proper numerical method and its parameters.

KEYWORDS

numerical method, contact, impact, oscillation, wave propagation

1. INTRODUCTION

By solving contact problems, one of the main challenges is the proper handling of the emerging nonlinearities. Historically, it was Hertz [1], who firstly dealt with these kinds of problems in 1881. In the field of contact problems, there are only a few simple examples, which can be treated analytically. In most cases, only an approximate solution can be achieved, which is usually provided by the Finite Element Method (FEM) [2, 3]. The Finite Element (FE) analysis of contact problems has a very wide literature [4, 5] in which two main approaches can be distinguished. The first technique is the penalty method [6], which is a simpler, faster, but not so accurate way for solving mechanical contact problems using FEM. The second approach is the Lagrange multiplier method [7] in which the contact pressure emerges as an additional unknown variable. This technique fulfills the contact conditions accurately, but its application is computationally more expensive. A third possibility is the augmented Lagrangian method [8], which is a certain kind of combination of the penalty and the Lagrange multiplier method. In this paper the contact boundary conditions are treated with the Lagrange multiplier method.

In most cases, the solution of a dynamic contact problem requires the time integration of the semi-discretized equation of motion, which is discretized in space, but continuous in time. By an equation of this kind, for the integration in time both forward and backward increment, single- and multi-stage numerical methods [9] can be used. This paper is going to analyze the applicability of four single-stage discrete time integration methods and will demonstrate their specific effects on the solution. The adaptation of the very commonly used Central Difference Method (CDM) [10] to the Lagrange multiplier method was presented by Carpenter et al. [11] in 1991. The main drawback of this method is that it possesses only conditional stability. However, the application of it is computationally cheaper and CDM can be applied effectively when a lumped mass matrix [12] is used instead of a consistent one. Furthermore, the application of the Backward Euler (BE) [13], the NewMark (NM) [14, 15]
and the Hilber-Hughes-Taylor-α (HHT-α) method will be studied, which are widely used backward increment schemes. The main advantage of these methods is that they are unconditionally stable. This is a major aspect; hence these methods are often built in commercial FE software. The Newmark method is the most common implicit method in this kind of software. The formulation of this method contains a parameter through which the numerical damping of the method can be affected directly. The HHT-α method is the improved version of the Newmark method, where the dissipation level can be controlled with two parameters. Nevertheless, the solution produced with a backward increment scheme is more time-consuming to achieve, as a system of equations must be solved in every time step. There are many more time integration schemes, like the Houbolt, the Wilson-θ, the generalized-α, etc. that can be found in [9]. Besides, there exist several multi stage schemes, like Kim’s method and the Kolay-Ricles-α (KR-α) method. In this paper, only the above mentioned four methods are considered, as they proved to be the most useful for contact problems, based on our experiences.

The right choice of the applied numerical method is crucial, as it has a significant impact on the solution characteristics, the computation time and the achievable accuracy. There is not any universal recipe to reach an ideal solution, many different schemes must be taken into consideration to examine, which method gives the best solution as of the user’s preferences. It is a further problem that due to the spatial discretization a spurious high frequency oscillation occurs in the resulting fields of displacement and contact pressure, which distorts the correct solution. This difficulty can be at least partially eliminated by applying a certain amount of numerical damping by the solution of the equation of motion. Nevertheless, it is also substantial to filter only the disturbance out by leaving the useful component of the signal intact. Thus, the adjustability of the numerical damping is an important aspect to consider, when choosing the right method. The main difficulty is to find an optimal region for the dissipation parameter to achieve a solution with acceptable accuracy, while significantly reducing the spurious oscillations.

In this article, the effectiveness of the aforementioned kinds of methods is going to be analyzed. The accuracy and the practical applicability of the examined methods will be studied from several aspects. Besides, parameter identifications are going to be made in order to determine the right amount of numerical damping for methods with adjustable dissipation level. The main goal of this paper is to provide a comprehensive application analysis about how to choose the right method with the correct parameters for a dynamic contact problem.

2. TIME STEPPING METHODS

Let us consider a structural dynamic problem, where the interacting bodies are linearly elastic. The semi-discretized equation of motion after spatial discretization by finite element method can be written as a system of linear differential equations of second order according to

\[ \mathbf{M}\ddot{u}_t + \mathbf{C}\dot{u}_t + \mathbf{K}u_t = f_t, \]  

where \(\mathbf{M}, \mathbf{C}\) and \(\mathbf{K}\) denote the mass, damping and stiffness matrices, respectively. On the right side \(f_t\) means the applied load vector, while \(u_t, \dot{u}_t, \ddot{u}_t\) are the unknown displacements, velocity and acceleration vectors at time instant \(t\), respectively. Moreover the solution of this differential system of equations requires both initial and boundary conditions from which the former can be written as

\[ u_{t=0} = u_0 \quad \text{and} \quad \dot{u}_{t=0} = \dot{u}_0, \]

where \(u_0\) and \(\dot{u}_0\) denote the initial displacement and velocity vectors, respectively. In the following, four different numerical methods will be introduced for solving numerically the equation of motion (1), hence the deduction will be presented for each method in the next four chapters.

2.1. Central difference method

The CDM is a commonly used forward increment method for various mechanical applications. The main benefits of using this scheme are that it is second order accurate, and computations can be executed very fast due to its explicit formulation. However, the central difference method is only conditionally stable, which reduces the applicable time step region. In this method, the velocity and the acceleration are given using the displacement at the current, previous and next time step.

\[ \dot{u}_t = \frac{u_{t+\Delta t} - u_{t-\Delta t}}{2\Delta t}, \]

\[ \ddot{u}_t = \frac{u_{t+\Delta t} - 2u_t + u_{t-\Delta t}}{\Delta t^2}, \]

\[ (\Delta t^2 C + 2M)u_{t+\Delta t} = (4M - 2\Delta t^2 K)u_t + (\Delta t^2 C - 2M)u_{t-\Delta t} + 2\Delta t^2 f_t. \]

By substituting Eqs (3) and (4) into the Eq. of motion (1) and reordering it, Eq. (5) can be obtained. The main problem with this formula is that it requires the inversion of the term \(\Delta t^2 C + 2M\), which is a computationally expensive procedure,

\[ \ddot{u}_t = \frac{u_t - u_{t-\Delta t}}{\Delta t}. \]

Hence, it is worthwhile to combine Eq. (4) with Eq. (6), as it results a formulation that only requires the inversion of \(M\). As the inversion of a lumped mass matrix is computationally cheap, it is rewarding to diagonalize the mass matrix by the method described in [12]. It is important to note that the formula applied for the \(\dot{u}_t\) velocities originates from the backward Euler method. Thus, the applied explicit technique is a certain kind of combination of the central difference and the backward Euler method,

\[ \mathbf{M}\ddot{u}_{t+\Delta t} = (2M - \Delta t^2 C + \Delta t^2 K)u_t + (\Delta t^2 C - M)u_{t-\Delta t} + \Delta t^2 f_t. \]

Equations (4) and (6) can be substituted into the equation of motion (1). By ordering it the same way, as it was evaluated for Eqs (5) and (7) can be obtained.
2.2. Backward Euler method

The backward Euler scheme has an implicit character, so a system of equations (which is often nonlinear) needs to be solved for every time step, resulting high computational costs. This method is unconditionally stable, first order accurate and has a certain amount of numerical damping. The main problem is that the dissipation level is not controllable; hence it often leads to energy dissipation of an undesirable rate,

\[
\begin{align*}
\mathbf{u}_{t+\Delta t} &= \mathbf{u}_t + \Delta t \ddot{\mathbf{u}}_t, \\
\mathbf{u}_{t+\Delta t} &= \mathbf{u}_t + \Delta t \ddot{\mathbf{u}}_t + \frac{\Delta t^2}{2} \left(1 - 2\beta\right) \ddot{\mathbf{u}}_t + \frac{\Delta t^2}{2} \dddot{\mathbf{u}}_t,
\end{align*}
\]

(8) \hspace{1cm} (9)

M\dddot{\mathbf{u}}_{t+\Delta t} + C\dot{\mathbf{u}}_{t+\Delta t} + K\mathbf{u}_{t+\Delta t} = \mathbf{f}_{t+\Delta t}. 
\hspace{1cm} (10)

The general formulation of this method according to [13] can be written as Eqs (8) and (9). Eq. (1) of motion for time step \( t + \Delta t \) leads to formulation (10).

2.3. Newmark method

The Newmark method [14, 15] is a commonly used implicit integration scheme, which is built in many commercial FEM software (Ansys, Abaqus, etc.),

\[
\begin{align*}
\mathbf{u}_{t+\Delta t} &= \mathbf{u}_t + \Delta t \ddot{\mathbf{u}}_t + \frac{\Delta t^2}{2} \left(1 - 2\beta\right) \ddot{\mathbf{u}}_t + \frac{\Delta t^2}{2} \dddot{\mathbf{u}}_t, \\
\dot{\mathbf{u}}_{t+\Delta t} &= \dot{\mathbf{u}}_t + \Delta t \left(1 - \gamma\right) \ddot{\mathbf{u}}_t + \gamma \dddot{\mathbf{u}}_t,
\end{align*}
\]

(11) \hspace{1cm} (12)

\[\beta = \frac{1}{4} (1 + \delta)^2, \text{ and } \gamma = \frac{1}{2} + \delta, \text{ where } \delta \geq 0. \hspace{1cm} (13)\]

Its basic formulation is determined by Eqs (11) and (12). Based on [14], the optimal parameter setting can be given as Eq. (13). By setting the \( \delta \) parameter according to Eq. (13), the integration accuracy and stability can be obtained by only one parameter. By changing this parameter, the amount of numerical damping also alters; hence the problematic oscillations are fully or partially extinguishable. This method is unconditionally stable and first order accurate (except by \( \gamma = 0.5 \) which results second order accuracy). The main disadvantage is the lack of high-frequency damping which limits the reduction of spurious noises.

2.4. HHT-\( \alpha \) method

The HHT-\( \alpha \) method [16] is another scheme, which is widely employed for solving the equation of motion of mechanical systems. This method is the improved version of the Newmark method with controllable numerical dissipation through the \( \alpha \) parameter. The main benefit in relation to the Newmark scheme is that by this method numerical damping can be exerted without degrading the order of accuracy,

\[
\begin{align*}
M\dddot{\mathbf{u}}_{t+\Delta t} + (1 + \alpha)C\dot{\mathbf{u}}_{t+\Delta t} - \alpha C\ddot{\mathbf{u}}_t + (1 + \alpha)K\mathbf{u}_{t+\Delta t} - \alpha K\mathbf{u}_t &= (1 + \alpha)\mathbf{f}_{t+\Delta t} - \alpha \mathbf{f}_t.
\end{align*}
\]

(14)

The basic Equations (11) and (12) of the Newmark method remain valid for this scheme. However, the equation of motion (10) must be modified using the \( \alpha \) parameter according to Eq. (14). In the interval of \( \alpha \in [-0.5, 0] \), the amount of numerical damping can be controlled where the smaller value of \( \alpha \) means the smaller amount of dissipation,

\[
\beta = \frac{1}{4} (1 - \alpha)^2 \text{ and } \gamma = \frac{1}{2} (1 - 2\alpha). 
\hspace{1cm} (15)
\]

If \( \alpha \) is treated independently, then Eq. (13) parameters from the Newmark method remain unchanged, so the result of the numerical computation can be influenced by two separate parameters, \( \delta \) and \( \alpha \). However, if \( \beta \) and \( \gamma \) are chosen according to Eq. (15), the scheme becomes second order accurate, although this does not cause the improvement of the numerical results necessarily. In this case the practical range \( \alpha \in [-1/3, 0] \). In practice, it is usually worthwhile to try both Eq. (15) and the independent choice of \( \delta \) and \( \alpha \).

3. MODELING OF THE IMPACT PROBLEM

3.1. FEM model

The methods that have been described above should be tested using an example by which it can be judged, how well they can handle the sudden changes in characteristic indicators like the velocity or the contact pressure. For this purpose, a 1D impact problem is an ideal choice as it is a widely used test example in the literature of contact problems [11, 17, 18]. The main benefits of this example are that its solution can be made easily and fast and it has an exact analytical solution, which provides a good reference to assess the results. Besides, it challenges the achievable accuracy of the numerical method at a sufficient level. Thus, if an accurate solution can be produced for this test example, then the applied method can presumably treat more complex geometries too. Therefore, a linearly elastic rod is considered which is moving towards a rigid wall (Fig. 1).

In order to solve this test example, a FEM model must be constructed based on the mechanical model shown in Fig. 1. This FEM model is built using 1D truss elements, which possess 2 nodes with 1 Degree of Freedom (DoF) per node, which is the longitudinal displacement. The FEM model is shown in Fig. 2, where \( L = 20 \) [mm]; \( h = 0.1 \) [mm]; \( n = 200 \) [-]; and \( \rho_0 = 1000 \) [mm/s]. The model contains uniform elements with \( E = 90 \) [MPa] and \( \rho = 7.85 \times 10^{-9} \) [kg/mm^3] material parameters. Based on the performed calculations, above the applied element number the results do not improve substantially.

![Fig. 1. The mechanical model of the examined 1D impact problem](image-url)
3.2. Treatment of the contact

The proper handling of the contact is an important aspect by the modeling.

As it has been mentioned earlier, by the Lagrange multiplier method, the penetration between the contacted bodies is completely prohibited. Because of this condition, certain displacement constraints are prescribed the following way:

$$G(u + X) = 0. \quad (16)$$

In the equation above, $X$ means the initial configuration (vector of length $n + 1$ filled with the initial coordinates of nodes), $u$ is the displacement vector, and $G$ is the contact constraint matrix. This equation is a certain kind of manifestation of the Hertz-Signorini-Moreau (HSM) conditions [4] in contact mechanics. The detailed description of here applied formulation can be found in [11]. Thus, Eq. (1) must be modified with regard to the contact force acting between the elastic rod and the rigid wall

$$\ddot{u} + C\dot{u} + Ku + G^T \lambda = f. \quad (17)$$

So, the equation of motion can be written in the form of Eq. (17) where $\lambda$ denotes the Lagrange multiplier at time instant $t$.

In this case, that quantity is equivalent with the surface contact pressure. Thus, the Lagrange multiplier method proceeds at the here presented example by treating $\lambda$ as unknown and solving Eqs (16) and (17) simultaneously. The numerical methods examined in Section 2 can easily be deducted for the contacting case by replacing the equation of motion (1) with (17).

4. PARAMETER IDENTIFICATION

In the case of the Newmark and the HHT-\(\alpha\) method, the quality of the numerical solution highly depends on the proper choice of the alterable parameter(s). Hence, a parameter identification is needed in order to test the applicability for a contact problem, while $p_i^{\text{FEM}}$ means the variable calculated from the FEM model at the $i$-th time step. On the other hand, the amount of the spurious oscillations is reviewed by the number of wave peaks in the time evolution of the considered variable denoted as $\text{NoP}_p$ (Number of Peaks).

4.1. Newmark method

Based on these two indicators the practical range of $\delta$ parameter for a contact problem may be ascertained. After determining them for the contact pressure for multiple values of $\delta$ inside the interval $\delta \in [0, 1]$, these indicators can be plotted in the function of $\delta$ according to Fig. 3a and b, respectively. As it can be seen in these figures, $\text{NoP}_p$ is decreasing strictly monotonic, while $\epsilon_p$ is starting to grow after a certain value of $\delta$ is stepped over. Hence, an ideal parameter value cannot be ascertained, by which both of the applied indicators are reaching their minimum, so an acceptable compromise must be made. Based on Fig. 3, a so called “practical stage” may be assigned, inside, which both measures remain at sufficiently low values. In accordance with the numerical tests using the FEM model, this practical stage can be determined as $\alpha \in [0.3, 1]$.

4.2. HHT-\(\alpha\) method

The parameter identification for this method is more complicated, as there are two independent parameters affecting the numerical solution. The applied indicators are the same as by the Newmark method, but by reason of the one more ruling parameter, a surface plot must be applied (Fig. 4a and b). In order to establish a beneficial correlation between $\alpha$ and $\delta$, the Period Error (PE) [19] function must be analyzed,

$$\delta_{\text{minimal PE}} = -1.5\alpha. \quad (19)$$

Based on this examination, Eq. (19) may be concluded to reach a parameter setting that results minimum for the period error. Considering Fig. 4, this connection proves to be a practical compromise between the accuracy and the elimination of the spurious oscillations. Hence, it is beneficial to choose the parameter values according to Eq. (19), especially as it makes the parameter setting much easier. Based on Fig. 4 and in accordance with the concrete numerical tests with the FEM model, the “practical stage” can be ascertained as $\alpha \in [-0.5, -0.3]$. 

![Fig. 2. FEM model](image)

![Fig. 3. a) Relative error of the pressure for varying values of $\delta$, b) number of wave peaks in the time for varying values of $\delta$](image)
5. APPLICATION IN A 1D EXAMPLE

In order to compare the applied methods and check the results of the parameter identification, a concrete function must be plotted that characterizes the 1D impact. As the spurious oscillations are mainly problematic in the contact pressure and the velocity functions, the methods shall be compared using the time evolution of these quantities (Figs 5–6). The magnified area on the left side shows the oscillations emerging directly after the moment of impact. The enlarged region on the right side of Fig. 5 indicates how steep the pressure decreases to zero which illustrates the accuracy of the solutions. The prominence in the zoomed plot on the right side of Fig. 6 is caused by the wave propagation in the rod and must be as narrow as possible.

Besides, the indicators that have been used in section 4 are illustrated by identical dissipation for each method in Figs 7 and 8 for the pressure and the velocity, respectively. By studying them, it can be assessed that there cannot be found an ideal method which to use is beneficial in all aspects.

The central difference method offers a fast and accurate result, but the large amount of oscillations makes it problematic to apply. In contrast, the backward Euler method offers the best performance in the elimination of oscillations, but this feature is coupled with the highest relative error. Based on this study, the Newmark and the HHT-α prove to have the highest potential at this kind of application. Both of them are controllable with one or more parameters, by which an acceptable compromise may be made. Even from these two methods, the HHT-α seems to possess more beneficial damping characteristics, as by identical dissipation it results a better solution. As the Newmark method is built in the ANSYS mechanical APDL software, the performed calculations were verified by that comparison. Applying the same FEM model with identical parameters in ANSYS,
the results show a very precise match, proving that the
calculations made in Julia language are accurate.

6. CONCLUSION

In this study a thorough analysis was made about the appli-
cation of numerical methods in contact problems. Four
different, widely used numerical schemes were deducted and
applied for a 1D example. In the parameter identification,
different aspects were presented, through which a correct
parameter setting may be determined for numerical methods
with one or more control parameters. The minimization of
the period error proved to be a benefit of the
HHT-α method regarding the examined measures. The results
of the application in a concrete 1D example also confirm
that this approach is correct. Nevertheless, a numerical
method that is ideal to solve a contact problem in all aspects
cannot be ascertained, as the right choice highly depends on
the user’s actual priorities. The main function of this paper is
to provide a comprehensive guide for the correct selection
of the numerical method for a concrete contact problem.

REFERENCES

[1] H. Hertz, “On the contact of elastic solids” (in German), J. Reine
Angew. Math., vol. 92, pp. 156–171, 1881.
[2] K. J. Bathe, Finite Element Procedures for Solids and Structures.
MIT Open Courseware, MIT, 1982–1986.
[3] M. Fahad, R. Nagy, and P. Fuleki, “Creep model to determine rut
development by autonomous truck axles on pavement,” Pollack
Period., vol. 17, no. 1, pp. 66–77, 2022.
[4] P. Wriggers, Computational Contact Mechanics. John Wiley &
Sons Ltd, 2002.
[5] T. A Laursen, Computational Contact and Impact Mechanics,
Fundamentals of Modeling Interfacial Phenomena in Nonlinear
Finite Element Analysis. Springer Science & Business Media, 2013.
[6] K Willner and L Gaul, “A penalty approach for contact descrip-
tion by fem based on interface physics,” Trans. Eng. Sci., vol. 7,
pp. 257–264, 1970.
[7] P. Papadopoulos and J. M. Solberg, “A Lagrange multiplier
method for the finite element solution of frictionless contact
problems,” Math. Comput. Model., vol. 28, nos 4–8, pp. 373–384,
1998.
[8] J. C. Simo and T. A. Laursen, “An augmented Lagrangian treat-
ment of contact problems involving friction,” Comput. Struct., vol.
42, no. 1, pp. 97–116, 1992.
[9] T. Belytschko and T. J. R. Hughes, Eds, Computational Methods
for Transient Analysis. North-Holland, 1983.
[10] B. Wu, H. Bao, J. Ou, and S. Tian, “Stability and accuracy analysis
of the central difference method for real-time substructure
testing,” Earthquake Eng. Struct. Dyn., vol. 34, no. 7, pp. 705–718,
2005.
[11] N. J. Carpenter, R. L. Taylor, and M. G. Katona, “Lagrange con-
straints for transient finite element surface contact,” Int. J. Numer.
Methods Eng., vol. 32, no. 1, pp. 103–128, 1991.
[12] S. R. Wu, “Lumped mass matrix in explicit finite element method
for transient dynamics of elasticity,” Comput. Methods Appl.
Mech. Eng., vol. 195, nos 44–47, pp. 5983–5994, 2006.
[13] B. Sumithra and A. Tamilselvan, “Numerical solution of stiff
system by backward Euler method,” Appl. Math. Sci., vol. 9, no.
67, pp. 3303–3311, 2015.
[14] N. M. Newmark, “A method of computation for structural
dynamics,” J. Eng. Mech. Division, vol. 85, no. 3, pp. 67–94,
1959.
[15] H. Maourane and T. Szabó, “Linear and nonlinear dynamical
analysis of a Crane model,” Pollack Period., vol. 15, no. 2, pp.
82–93, 2020.
[16] H. M. Hilber, T. J. R. Hughes, and R. L. Taylor, “Improved nu-
merical dissipation for time integration algorithms in structural
dynamics,” Earthquake Eng. Struct. Dyn., vol. 5, no. 3, pp. 283–
292, 1977.
[17] W. Kim, “A new family of two-stage explicit time integration
methods with dissipation control capability for structural dy-
namics,” Eng. Struct., vol. 195, pp. 358–372, 2019.
[18] C. Kolay and J. M. Ricles, “Development of a family of uncondi-
tionally stable explicit direct integration algorithms with controll-
able numerical energy dissipation,” Earthquake Eng. Struct. Dyn.,
vol. 43, no. 9, pp. 1361–1380, 2014.
[19] G. Noh, S. Ham, and K. J. Bathe, “Performance of an implicit time
integration scheme in the analysis of wave propagations,” Comput.
Struct., vol. 123, pp. 93–105, 2013.