

Bipenalty method for time domain computational dynamics

BY HARM ASKES^{1,*}, MIGUEL CARAMÉS-SADDLER¹
AND ANTONIO RODRÍGUEZ-FERRAN²

¹*University of Sheffield, Department of Civil and Structural Engineering,
Mappin Street, Sheffield S1 3JD, UK*

²*Universitat Politècnica de Catalunya, Laboratori de Càlcul Numèric,
Jordi Girona 1-3, 08034 Barcelona, Spain*

Penalty functions can be used to add constraints to systems of equations. In computational mechanics, stiffness-type penalties are the common choice. However, in dynamic applications stiffness penalties have the disadvantage that they tend to *decrease* the critical time step in conditionally stable time integration schemes, leading to increased CPU times for simulations. In contrast, inertia penalties *increase* the critical time step. In this paper, we suggest the simultaneous use of stiffness penalties and inertia penalties, which is denoted as the *bipenalty method*. We demonstrate that the accuracy of the bipenalty method is at least as good as (and usually better than) using either stiffness penalties or inertia penalties. Furthermore, for a number of finite elements (bar elements, beam elements and square plane stress/plane strain elements) we have derived ratios of the two penalty parameters such that their combined effect on the critical time step is neutral. The bipenalty method is thus superior to using stiffness penalties, because the decrease in critical time step can be avoided. The bipenalty method is also superior to using inertia penalties, since the constraints are realized with higher accuracy.

Keywords: penalty method; constraints; critical time step; explicit time integration; bipenalty method

1. Introduction

The penalty method is a popular technique to impose constraints, such as Dirichlet conditions, tyings or contact conditions, to a system of equations. In its common format, a penalty function is an addition to the potential energy of the system that consists of the constraint that is to be enforced together with a user-defined number known as the penalty parameter. The interpretation of such penalty parameters is that of (very) stiff springs, and we will refer to these penalties as *stiffness penalties*. If the penalty parameter is infinitely large, the constraint is satisfied exactly; however, in practical applications finite magnitudes of the penalty parameter must be used, implying that the constraint is only enforced approximatively. Thus, there is an inherent dilemma in using penalty functions: for *modelling accuracy* the penalty parameter must be taken as large

as possible, but for *numerical accuracy* (related to the machine precision of computers) the penalty parameter must not be too large, see for instance the discussion in Bathe (1996).

There is another drawback of using stiffness penalties, which manifests itself in time domain dynamic applications. As is well known, the speed of sound of a material is proportional to the square root of stiffness divided by mass. The use of stiffness penalties thus leads to an increase of the speed of sound. When a time integration algorithm is used that is only conditionally stable (such as explicit time integration), a critical time step exists such that numerical instabilities may occur if the time step for numerical time integration is chosen larger than this critical time step. The critical time step is inversely proportional to the speed of sound, thus there is an important implication of using stiffness penalties: stiffness penalties lead to a decrease of the critical time step, which implies that numerical simulations require longer CPU times (Belytschko & Neal 1991; Carpenter *et al.* 1991).

In recent years research has been carried out aimed at using lower values of the penalty parameters, while maintaining accuracy of constraint imposition, or aimed at penalty formulations that have a less adverse effect on the critical time step. Ilanko and coworkers (Ilanko & Dickinson 1999; Ilanko 2002*a,b*, 2003, 2005*a*; Askes & Ilanko 2006; Askes *et al.* 2008) proposed the use of positive as well as negative values for the penalty parameters and they showed that the exact solution is bracketed by the results obtained with either positive or negative penalties. This result can be used to interpolate between results obtained with moderately large (but not very large) positive and negative penalty parameters. More recently, it was suggested to use inertia penalties instead of stiffness penalties (Ilanko 2005*b*; Williams & Ilanko 2005; Hetherington & Askes 2009). Whereas stiffness penalties can be interpreted as stiff springs, inertia penalties can be regarded as heavy masses; both types of penalties limit the motion of the relevant degree of freedom (d.f.). There is an additional advantage of using inertia penalties instead of stiffness penalties: since the speed of sound is lowered by increasing the mass, the critical time step of a finite element penalized with inertia is *increased* (Hetherington & Askes 2009).

The opposite effects of stiffness penalties and inertia penalties on the critical time step naturally leads to the idea that they can be combined: the motion of a d.f. that is to be constrained can be penalized *simultaneously* with stiffness penalties and with inertia penalties. This approach will be denoted here as the *bipenalty method* and its development is the main aim of this paper. We will focus on two major aspects of the proposed bipenalty method:

- *Effects on accuracy.* Engineering judgement suggests that the simultaneous use of stiffness penalties and inertia penalties provides similar or better accuracy of constraint imposition than using either of the two penalties with equal magnitude on its own. We will formulate a mathematical proof that underpins this notion.
- *Effects on stability.* Since stiffness penalties and inertia penalties have opposite effects on the critical time step, there are certain combinations of stiffness penalties and inertia penalties for which there is a *zero* net effect on the critical time step. For a number of finite element types it is possible to derive closed-form expressions for the ratio of stiffness penalty

parameter over inertia penalty parameter for which the net effect on the critical time step is neutral—such a ratio will be denoted as the ‘critical penalty ratio’.

After reviewing penalty functions based on stiffness and inertia in §2, accuracy and stability of the bipenalty method will be treated in §§3 and 4, respectively. Numerical examples are presented in §5 to illustrate and corroborate the analytical results of the earlier sections. Throughout the paper, the emphasis is on *time domain* analysis; possible extension of the findings to *frequency domain* analysis are beyond the scope of this study.

2. Stiffness penalties and inertia penalties

The potential energy \mathcal{U} of a structure is usually written as

$$\mathcal{U} = \frac{1}{2} \mathbf{u}^T \mathbf{K} \mathbf{u} - \mathbf{u}^T \mathbf{f}, \quad (2.1)$$

where \mathbf{K} the stiffness matrix, \mathbf{u} is a vector of displacements and \mathbf{f} the vector of external forces. A constraint $u_n - \bar{u} = 0$ is applied to the n th d.f. and \bar{u} is its prescribed value. The penalized potential energy is accordingly written as

$$\mathcal{U} = \frac{1}{2} \mathbf{u}^T \mathbf{K} \mathbf{u} - \mathbf{u}^T \mathbf{f} + \frac{1}{2} \alpha_s (u_n - \bar{u})^2, \quad (2.2)$$

where α_s is a penalty parameter of the stiffness type.

The usual expression for the kinetic energy \mathcal{T} reads

$$\mathcal{T} = \frac{1}{2} \dot{\mathbf{u}}^T \mathbf{M} \dot{\mathbf{u}}, \quad (2.3)$$

where \mathbf{M} is the mass matrix. The constraint introduced above is written in rate form and added to the kinetic energy as

$$\mathcal{T} = \frac{1}{2} \dot{\mathbf{u}}^T \mathbf{M} \dot{\mathbf{u}} + \frac{1}{2} \alpha_m (\dot{u}_n - \dot{\bar{u}})^2, \quad (2.4)$$

where α_m is an inertial penalty parameter that penalizes the difference in actual velocity \dot{u}_n and prescribed velocity $\dot{\bar{u}}$. The equations of motion follow from

$$\frac{d}{dt} \frac{\partial \mathcal{T}}{\partial \dot{\mathbf{u}}^T} + \frac{\partial \mathcal{U}}{\partial \mathbf{u}^T} = [\mathbf{M} + \mathbf{M}^P] \ddot{\mathbf{u}} + [\mathbf{K} + \mathbf{K}^P] \mathbf{u} - \mathbf{f} - \mathbf{f}^P = \mathbf{0}. \quad (2.5)$$

The components of the mass penalty matrix \mathbf{M}^P , stiffness penalty matrix \mathbf{K}^P and penalty force vector \mathbf{f}^P are zero except for the entries related to the n th d.f.:

$$M_{nn}^P = \alpha_m, \quad (2.6)$$

$$K_{nn}^P = \alpha_s \quad (2.7)$$

and

$$f_n^P = \alpha_m \ddot{\bar{u}} + \alpha_s \bar{u}. \quad (2.8)$$

To keep the subsequent derivations simple, in the remainder of the paper we will assume that the prescribed displacement \bar{u} is constant in time, by which $\dot{\bar{u}} = \ddot{\bar{u}} = 0$. However, varying values of \bar{u} would only change the penalty force vector, which turns out to be of minor importance in our derivations. It is also

useful to introduce dimensionless penalty factors p_m and p_s that relate the penalty parameters α_m and α_s to the corresponding diagonal entries of \mathbf{M} and \mathbf{K} via $\alpha_m = p_m M_{nn}$ and $\alpha_s = p_s K_{nn}$, respectively.

3. Accuracy of the bipenalty method

For the time integration of the equations of motion (2.5) we will use the Newmark method. Denoting the time-discretized counterparts of the displacements \mathbf{u} , velocities $\dot{\mathbf{u}}$ and accelerations $\ddot{\mathbf{u}}$ with \mathbf{d} , \mathbf{v} and \mathbf{a} , respectively, the Newmark equations can be written as

$$\mathbf{v}^{t+\Delta t} = \mathbf{v}^t + (1 - \gamma)\Delta t \mathbf{a}^t + \gamma \Delta t \mathbf{a}^{t+\Delta t} \quad (3.1)$$

and

$$\mathbf{d}^{t+\Delta t} = \mathbf{d}^t + \Delta t \mathbf{v}^t + \left(\frac{1}{2} - \beta\right) \Delta t^2 \mathbf{a}^t + \beta \Delta t^2 \mathbf{a}^{t+\Delta t}, \quad (3.2)$$

where γ and β are user-defined parameters. Expressed in terms of the time-discretized accelerations \mathbf{a} at time $t + \Delta t$, the equations of motion read

$$\begin{aligned} [\mathbf{M} + \mathbf{M}^P + \beta \Delta t^2 (\mathbf{K} + \mathbf{K}^P)] \mathbf{a}^{t+\Delta t} &= \mathbf{f} + \mathbf{f}^P - [\mathbf{K} + \mathbf{K}^P] (\mathbf{d}^t + \Delta t \mathbf{v}^t \\ &+ \left(\frac{1}{2} - \beta\right) \Delta t^2 \mathbf{a}^t). \end{aligned} \quad (3.3)$$

Furthermore, we will assume initial values $\mathbf{a}^0 = \mathbf{v}^0 = \mathbf{0}$ as well as $\mathbf{d}^0 = \mathbf{0}$ except $d_n^0 = \bar{u}$.

(a) Convergence of the bipenalty method

Adopting the initialization indicated above, the penalized (n th) equation at time $t = \Delta t$ can be written as

$$\sum_i (M_{ni} + \beta \Delta t^2 K_{ni}) a_i^{\Delta t} + (\alpha_m + \alpha_s \beta \Delta t^2) a_n^{\Delta t} = f_n - K_{nn} \bar{u}. \quad (3.4)$$

We introduce $\alpha_m = \eta_m p$ and $\alpha_s = \eta_s p$, where η_m and η_s are user-defined numbers to set the magnitude of inertia penalties relative to the magnitude of stiffness penalties, while p is a large number. Thus, convergence is to be verified for the case $p \rightarrow \infty$. Dividing equation (3.4) by p and rearranging yields

$$(\eta_m + \eta_s \beta \Delta t^2) a_n^{\Delta t} = \frac{1}{p} \left(f_n - K_{nn} \bar{u} - \sum_i (M_{ni} + \beta \Delta t^2 K_{ni}) a_i^{\Delta t} \right) \quad (3.5)$$

by which

$$\lim_{p \rightarrow \infty} a_n^{\Delta t} = \lim_{p \rightarrow \infty} \frac{1}{(\eta_m + \eta_s \beta \Delta t^2) p} \left(f_n - K_{nn} \bar{u} - \sum_i (M_{ni} + \beta \Delta t^2 K_{ni}) a_i^{\Delta t} \right) = 0, \quad (3.6)$$

where we have used that all $a_i^{\Delta t}$ are finite, since they are obtained from equation (3.3) that contains a positive definite system matrix. Thus, convergence to the same (exact) solution is observed irrespective of whether one uses stiffness penalties, inertia penalties or both.

In order to assess the influence of the error in the penalized d.f. on the other d.f., we compare the penalized system of equations given in equation (2.5) with the exact solution $\tilde{\mathbf{u}}$ obtained via Lagrange multipliers. The latter is obtained from

$$\begin{bmatrix} \mathbf{M} & \mathbf{0} \\ \mathbf{0} & 0 \end{bmatrix} \begin{bmatrix} \ddot{\tilde{\mathbf{u}}} \\ \ddot{\tilde{\lambda}} \end{bmatrix} + \begin{bmatrix} \mathbf{K} & \mathbf{A}^T \\ \mathbf{A} & 0 \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{u}} \\ \tilde{\lambda} \end{bmatrix} = \begin{bmatrix} \mathbf{f} \\ g \end{bmatrix}, \quad (3.7)$$

where λ is the Lagrange multiplier of the constrained n th d.f. (that is, λ is the reaction force conjugated to u_n). Furthermore, \mathbf{A} and g are the usual Lagrange multiplier contributions to left-hand-side and right-hand-side: all $A_i = 0$ except $A_n = 1$, and $g = \bar{u}$. Subtracting the first set of equations of expression (3.7) from equation (2.5) yields

$$\mathbf{M}\ddot{\mathbf{e}} + \mathbf{M}^P\ddot{\mathbf{u}} + \mathbf{K}\mathbf{e} + \mathbf{K}^P\mathbf{u} - \mathbf{A}^T\lambda = \mathbf{f}^P, \quad (3.8)$$

where the error \mathbf{e} is defined as $\mathbf{e} = \mathbf{u} - \tilde{\mathbf{u}}$. For the case of a single constraint on the n th d.f., the n th equation of expression (3.8) can be simplified as

$$\sum_i (M_{ni}\ddot{e}_i + K_{ni}e_i) = \lambda - (\eta_s(u_n - \bar{u}) + \eta_m\ddot{u}_n)p. \quad (3.9)$$

Applying the Newmark time integration to the error \mathbf{e} , equation (3.9) can be written for the first time evaluation $t = \Delta t$ as

$$\sum_i \left(\frac{1}{\beta\Delta t^2} M_{ni} + K_{ni} \right) e_i = \lambda - (\eta_s(u_n - \bar{u}) + \eta_m\ddot{u}_n)p \equiv r_n. \quad (3.10)$$

This last expression shows that the error is obtained from

$$\mathbf{e} = \left[\frac{1}{\beta\Delta t^2} \mathbf{M} + \mathbf{K} \right]^{-1} \mathbf{r}, \quad (3.11)$$

where the vector of residuals \mathbf{r} is defined as $r_i = 0$ for all i except $i = n$, for which r_n is defined in equation (3.10). Therefore, the errors of all d.f. e_i are governed by the same scalar, namely the residual r_n of the constrained d.f. We can now make the following observations:

- In equation (3.6) it was shown that $a_n^{\Delta t} = 0$ for $p \rightarrow \infty$, therefore $e_n \rightarrow 0$ for $p \rightarrow \infty$. In the context of equation (3.10), this can only hold if $r_n \rightarrow 0$ for $p \rightarrow \infty$.
- Since $r_n \rightarrow 0$ for $p \rightarrow \infty$, it follows from equation (3.10) that $e_i \rightarrow 0$ for $p \rightarrow \infty$, which is valid for all i .

Thus, we have shown that the bipenalty method converges for all d.f. Results for subsequent time instants can be obtained with induction.

(b) Rate of convergence of the bipenalty method

To compare the accuracy of the bipenalty method with that of using either stiffness penalties or inertia penalties, we investigate the rate of convergence of the penalized d.f., that is $da_n^{\Delta t}/dp$, for the case $p \rightarrow \infty$. Since it was established in equation (3.6) that all penalty methods converge to the same solution, it can be said that a smaller value of $da_n^{\Delta t}/dp$ corresponds to a more accurate method.

From equation (3.4) it follows that

$$\sum_i (M_{ni} + \beta \Delta t^2 K_{ni}) \frac{da_i^{\Delta t}}{dp} + (\eta_m + \eta_s \beta \Delta t^2) a_n^{\Delta t} + p(\eta_m + \eta_s \beta \Delta t^2) \frac{da_n^{\Delta t}}{dp} = 0 \quad (3.12)$$

by which

$$\frac{da_n^{\Delta t}}{dp} = - \frac{\sum_i (M_{ni} + \beta \Delta t^2 K_{ni}) da_i^{\Delta t}/dp + (\eta_m + \eta_s \beta \Delta t^2) a_n^{\Delta t}}{p(\eta_m + \eta_s \beta \Delta t^2)}. \quad (3.13)$$

The magnitude of $da_n^{\Delta t}/dp$ is governed by the denominator in equation (3.13). Taking the same magnitude for each penalty parameter, it can be seen that the bipenalty method is at least as accurate as (or, in case $\beta > 0$, higher than) using either stiffness penalties or inertia penalties. Again, induction may be used to extend this result to subsequent time steps.

4. Stability of the bipenalty method

The critical time step Δt_{crit} that is relevant for conditionally stable time integrators is inversely proportional to the so-called ‘maximum sampling frequency’ ω_{max} that can be monitored by the finite element discretization (Hughes 2000), namely

$$\Delta t_{\text{crit}} = \frac{\Omega_{\text{crit}}}{\omega_{\text{max}}} \quad \text{with} \quad \Omega_{\text{crit}} = \frac{1}{\sqrt{\frac{1}{2}\gamma - \beta}}, \quad (4.1)$$

where the critical sampling ratio Ω_{crit} is given for the Newmark time integrator and for a system without physical damping. This maximum sampling frequency is found by solving a dynamic eigenvalue problem for one finite element written as

$$\det(-\omega^2[\mathbf{M} + \mathbf{M}^P] + [\mathbf{K} + \mathbf{K}^P]) = 0 \quad (4.2)$$

and selecting the largest root of ω . We aim to establish a relationship between the two penalties α_m and α_s such that the maximum sampling frequency is not affected. A penalty ratio r is defined in terms of the dimensionless penalty factors as $r = p_s/p_m$. For large values of r , the stiffness penalty dominates which could lead to an increase in ω_{max} and, thus, an undesired decrease in the critical time step. Conversely, for small values of r the inertia penalty dominates, which lowers ω_{max} and increases the critical time step. We will seek the upper limit of r such that the maximum sampling frequency of a bipenalized element is identical to the maximum sampling frequency of an unpenalized element—this upper limit of r is called the ‘critical penalty ratio’ and denoted as r_{crit} .

The procedure we will follow is summarized as follows:

- Select an element type and specify the mass matrix, which is taken here as either the lumped mass matrix or the consistent mass matrix (these are related to discrete masses and distributed masses, respectively). In this paper, we will present results for a number of finite elements with increasing complexity, namely two-noded bar elements, two-noded beam elements and four-noded square plane stress/plane strain elements.

Table 1. Summary of obtained critical penalty ratios—all values of r_{crit} have been derived analytically except those indicated with an asterisk, which were postulated and verified numerically.

element type	mass matrix	penalized d.f.	Poisson's ratio	critical penalty ratio
bar	lumped	displacement	n.a.	$r_{\text{crit}} = 2$
	consistent	displacement	n.a.	$r_{\text{crit}} = 4$
beam	lumped	displacement	n.a.	$r_{\text{crit}} = 8$
		rotation	n.a.	$r_{\text{crit}} = 2$
		displacement and rotation	n.a.	$r_{\text{crit}} \downarrow 2$
	consistent	displacement	n.a.	$r_{\text{crit}} = 260$
		rotation	n.a.	$r_{\text{crit}} = 20$
		displacement and rotation	n.a.	$r_{\text{crit}} \downarrow 20$
square/plane stress	lumped	displacement	all values	$r_{\text{crit}} = \frac{6(1+\nu)}{3-\nu}$
				$r_{\text{crit}} \geq \frac{8(1+\nu)}{3-\nu}$ (*)
	consistent	displacement	$\nu < \frac{1}{3}$	$r_{\text{crit}} \geq \frac{8(1+\nu)}{3-\nu}$ (*)
			$\nu \geq \frac{1}{3}$	$r_{\text{crit}} = \frac{8(1+\nu)}{3-\nu}$
square/plane strain	lumped	displacement	all values	$r_{\text{crit}} = \frac{6}{3-4\nu}$
				$r_{\text{crit}} \geq \frac{8}{3-4\nu}$ (*)
	consistent	displacement	$\nu < \frac{1}{4}$	$r_{\text{crit}} \geq \frac{8}{3-4\nu}$ (*)
			$\nu \geq \frac{1}{4}$	$r_{\text{crit}} = \frac{8}{3-4\nu}$

- Solve equation (4.2) with $p_m = p_s = 0$. This gives the maximum sampling frequency of an unpenalized element.
- Substitute the unpenalized value of ω_{max} into equation (4.2) and solve again, but now with $p_m \neq 0$ and $p_s \neq 0$. This gives a relation between p_m and p_s , from which r_{crit} can be determined.

Readers who wish to skip the technical details may want to refer directly to table 1, where we have summarized all obtained values for the critical penalty ratio.

(a) Bar elements

The lumped mass matrix, consistent mass matrix and stiffness matrix of a two-noded bar element read

$$\mathbf{M}_{\text{lum}} = \frac{\rho Ah}{2} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \mathbf{M}_{\text{con}} = \frac{\rho Ah}{6} \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \quad \text{and} \quad \mathbf{K} = \frac{EA}{h} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}, \quad (4.3)$$

where ρ , E , A and h are the mass density, Young's modulus, cross-sectional area and element size, respectively. Penalties are imposed by multiplying one of the diagonal terms with $1 + p_m$ (for the mass matrices) or with $1 + p_s$ (for the stiffness matrix).

When the lumped mass matrix is taken, the penalized eigenvalue problem reads

$$\omega^2 \left(\frac{1}{4} \omega^2 - \frac{c_e^2}{h^2} \right) + \left(\frac{1}{2} \omega^2 p_m - \frac{c_e^2}{h^2} p_s \right) \left(\frac{1}{2} \omega^2 - \frac{c_e^2}{h^2} \right) = 0, \quad (4.4)$$

where $c_e = \sqrt{E/\rho}$ is the elastic bar velocity (i.e. the speed of sound). The unpenalized eigenvalue problem is retrieved by taking $p_m = p_s = 0$, which gives $\omega_{\max} = 2c_e/h$. This value for ω_{\max} is next substituted into the penalized eigenvalue problem. This renders $2p_m - p_s = 0$ and consequently $r_{\text{crit}} = 2$.

With a consistent mass matrix, the penalized eigenvalue problem can be written as

$$\omega^2 \left(\frac{1}{4} \omega^2 - 3 \frac{c_e^2}{h^2} \right) + \left(\omega^2 p_m - 3 \frac{c_e^2}{h^2} p_s \right) \left(\frac{1}{3} \omega^2 - \frac{c_e^2}{h^2} \right) = 0. \quad (4.5)$$

The unpenalized maximum sampling frequency equals $\omega_{\max} = 2\sqrt{3} c_e/h$. When this value is used for ω in equation (4.5) with $p_m \neq 0$ and $p_s \neq 0$, we obtain $4p_m - p_s = 0$ so that $r_{\text{crit}} = 4$.

(b) Euler–Bernoulli beam elements

The consistent mass matrix and stiffness matrix of a two-noded Euler–Bernoulli beam element read

$$\mathbf{M}_{\text{con}} = \frac{\rho Ah}{420} \begin{bmatrix} 156 & 22h & 54 & -13h \\ 22h & 4h^2 & 13h & -3h^2 \\ 54 & 13h & 156 & -22h \\ -13h & -3h^2 & -22h & 4h^2 \end{bmatrix} \quad (4.6)$$

and

$$\mathbf{K} = \frac{EI}{h^3} \begin{bmatrix} 12 & 6h & -12 & 6h \\ 6h & 4h^2 & -6h & 2h^2 \\ -12 & -6h & 12 & -6h \\ 6h & 2h^2 & -6h & 4h^2 \end{bmatrix},$$

where I is the second moment of area of the cross section. Although there is no unique definition of the lumped mass matrix, most authors (Ishihara 1981; Kim 1993; Bathe 1996) use the following expression:

$$\mathbf{M}_{\text{lum}} = \frac{\rho Ah}{24} \begin{bmatrix} 12 & 0 & 0 & 0 \\ 0 & h^2 & 0 & 0 \\ 0 & 0 & 12 & 0 \\ 0 & 0 & 0 & h^2 \end{bmatrix}. \quad (4.7)$$

In deriving the penalized eigenvalue problems, we will distinguish between three cases, namely (i) penalized deflection, (ii) penalized rotation, and (iii) penalized deflection as well as penalized rotation. In terms of modifying the element matrices, this amounts to multiplying the first diagonal component, the second diagonal component, or both, with factors $1 + p_m$ and $1 + p_s$. However, derivations will only be presented for the (most general) third case.

Using a lumped mass matrix and with penalties applied on deflection as well as rotation, the polynomial equation of the eigenvalue problem contains not only linear terms in p_m and p_s but also quadratic terms in p_m and p_s :

$$\begin{aligned} & \omega^4 \left(\omega^2 - 192c_e^2 \frac{I}{Ah^4} \right) \left(\omega^2 - 48c_e^2 \frac{I}{Ah^4} \right) + 2\omega^4 \left(\omega^4 - 180\omega^2 c_e^2 \frac{I}{Ah^4} + 4608c_e^4 \frac{I^2}{A^2h^8} \right) p_m \\ & - 24\omega^2 c_e^2 \frac{I}{Ah^4} \left(5\omega^4 - 792\omega^2 c_e^2 \frac{I}{Ah^4} + 12672c_e^4 \frac{I^2}{A^2h^8} \right) p_s \\ & + \left(\omega^4 p_m^2 - \frac{120\omega^2 c_e^2 I}{Ah^4} p_m p_s + \frac{2304c_e^4 I^2}{A^2h^8} p_s^2 \right) \left(\omega^4 - 120\omega^2 c_e^2 \frac{I}{Ah^4} + 576 \frac{I^2}{A^2h^8} \right) = 0 \end{aligned} \quad (4.8)$$

For an unpenalized element the maximum sampling frequency $\omega_{\max} = c_e \sqrt{192I/h^2} \sqrt{A}$. With this value substituted back into equation (4.8) we obtain the following equation for the relation between p_m and p_s :

$$25p_s^2 + 384p_m - 156p_s + 400p_m^2 - 250p_m p_s = 0. \quad (4.9)$$

When p_s is resolved from this last expression, the critical penalty ratio is found as a function of the inertia penalty factor p_m as

$$r_{\text{crit}} = 5 + \frac{78}{25p_m} \pm 3 \sqrt{\frac{676}{625p_m^2} + \frac{1100}{625p_m} + 1}. \quad (4.10)$$

For $p_m \rightarrow \infty$ we obtain $r_{\text{crit}} \uparrow 8$ or $r_{\text{crit}} \downarrow 2$. These two asymptotic values coincide with the critical penalty ratios found for penalized deflection only and penalized rotation only, respectively (the derivations of which are not shown here). In case both deflection and rotation are penalized, the lower value for r_{crit} as obtained from equation (4.10) must be taken; since the relevant convergence is from above it is safe to use $r = 2$.

The eigenvalue problem with a consistent mass matrix and penalties on deflection as well as rotation of the first node leads to an equation for ω as

$$\begin{aligned} & \omega^4 \left(\omega^2 - 8400c_e^2 \frac{I}{Ah^4} \right) \left(\omega^2 - 720c_e^2 \frac{I}{Ah^4} \right) + \frac{32}{35} \omega^4 \left(19\omega^4 - 58380\omega^2 c_e^2 \frac{I}{Ah^4} \right. \\ & \left. + 10584000c_e^4 \frac{I^2}{A^2h^8} \right) p_m - 384\omega^2 c_e^2 \frac{I}{Ah^4} \left(13\omega^4 - 27060\omega^2 c_e^2 \frac{I}{Ah^4} \right. \end{aligned}$$

$$\begin{aligned}
& +1512000c_e^4 \frac{I^2}{A^2 h^8} \Big) p_s + \left(\frac{832}{49} \omega^4 p_m^2 - 7680 \omega^2 c_e^2 \frac{I}{Ah^4} p_m p_s + 230400 c_e^4 \frac{I^2}{A^2 h^8} p_s^2 \right) \\
& \times \left(\omega^4 - 1224 \omega^2 c_e^2 \frac{I}{Ah^4} + 15120 c_e^4 \frac{I^2}{A^2 h^8} \right) = 0
\end{aligned} \tag{4.11}$$

With $p_m = p_s = 0$ the maximum sampling frequency of an unpenalized element is found as $\omega_{\max} = c_e \sqrt{8400I/h^2} \sqrt{A}$. Substituting this value into equation (4.11) leads to

$$1709p_s^2 + 6832000p_m - 274400p_s + 8886800p_m^2 - 478520p_m p_s = 0. \tag{4.12}$$

Resolving p_s and dividing the result by p_m yields

$$r_{\text{crit}} = 140 + \frac{137200}{1709p_m} \pm 120 \sqrt{\frac{11764900}{p_m^2} + \frac{6580}{5127p_m} + 1}. \tag{4.13}$$

For $p_m \rightarrow \infty$ it can be seen that $r_{\text{crit}} \uparrow 260$ or $r_{\text{crit}} \downarrow 20$. Similar to the beam element with lumped mass, these two asymptotic values coincide with the critical penalty ratios for penalized deflection only and penalized rotation only, respectively (again, these derivations are not presented here for the sake of brevity). For the case where both deflection and rotation are penalized, the value $r = 20$ can be used safely.

(c) Square plane stress/plane strain elements

Finally, we will derive critical penalty ratios for one particular continuum element, namely a four-noded quadrilateral element, with one penalized d.f. If this element is taken to be isoparametric with arbitrary shape, the Jacobian will appear in the derivations. To avoid this, we will assume that the element is square, which is in any case the recommended shape of the element for optimal behaviour. Our results for the critical time steps of this element are in accordance with those presented earlier in Ling & Cherukuri (2002).

Assuming unit thickness in the out-of-plane direction, the lumped mass, consistent mass and stiffness matrices for a four-node square plane stress element are given by

$$\mathbf{M}_{\text{lum}} = \frac{\rho h^2}{4} \mathbf{I}_{8 \times 8}, \tag{4.14}$$

$$\mathbf{M}_{\text{con}} = \frac{\rho h^2}{36} \begin{bmatrix} 4 & 0 & 2 & 0 & 1 & 0 & 2 & 0 \\ 0 & 4 & 0 & 2 & 0 & 1 & 0 & 2 \\ 2 & 0 & 4 & 0 & 2 & 0 & 1 & 0 \\ 0 & 2 & 0 & 4 & 0 & 2 & 0 & 1 \\ 1 & 0 & 2 & 0 & 4 & 0 & 2 & 0 \\ 0 & 1 & 0 & 2 & 0 & 4 & 0 & 2 \\ 2 & 0 & 1 & 0 & 2 & 0 & 4 & 0 \\ 0 & 2 & 0 & 1 & 0 & 2 & 0 & 4 \end{bmatrix} \tag{4.15}$$

and

$$\mathbf{K} = \begin{bmatrix} \frac{E(3-\nu)}{6(1-\nu^2)} & \frac{E}{8(1-\nu)} & \frac{-E(3+\nu)}{12(1-\nu^2)} & \frac{-E(1-3\nu)}{8(1-\nu^2)} & \frac{-E(3-\nu)}{12(1-\nu^2)} & \frac{-E}{8(1-\nu)} & \frac{E\nu}{6(1-\nu^2)} & \frac{E(1-3\nu)}{8(1-\nu^2)} \\ \frac{E}{8(1-\nu)} & \frac{E(3-\nu)}{6(1-\nu^2)} & \frac{E(1-3\nu)}{8(1-\nu^2)} & \frac{E\nu}{6(1-\nu^2)} & \frac{-E}{8(1-\nu)} & \frac{-E(3-\nu)}{12(1-\nu^2)} & \frac{-E(1-3\nu)}{8(1-\nu^2)} & \frac{-E(3+\nu)}{12(1-\nu^2)} \\ \frac{-E(3+\nu)}{12(1-\nu^2)} & \frac{E(1-3\nu)}{8(1-\nu^2)} & \frac{E(3-\nu)}{6(1-\nu^2)} & \frac{-E}{8(1-\nu)} & \frac{E\nu}{6(1-\nu^2)} & \frac{-E(1-3\nu)}{8(1-\nu^2)} & \frac{-E(3-\nu)}{12(1-\nu^2)} & \frac{E}{8(1-\nu)} \\ \frac{-E(1-3\nu)}{8(1-\nu^2)} & \frac{E\nu}{6(1-\nu^2)} & \frac{-E}{8(1-\nu)} & \frac{E(3-\nu)}{6(1-\nu^2)} & \frac{E(1-3\nu)}{8(1-\nu^2)} & \frac{-E(3+\nu)}{12(1-\nu^2)} & \frac{E}{8(1-\nu)} & \frac{-E(3-\nu)}{12(1-\nu^2)} \\ \frac{-E(3-\nu)}{12(1-\nu^2)} & \frac{-E}{8(1-\nu)} & \frac{E\nu}{6(1-\nu^2)} & \frac{E(1-3\nu)}{8(1-\nu^2)} & \frac{E(3-\nu)}{6(1-\nu^2)} & \frac{E}{8(1-\nu)} & \frac{-E(3+\nu)}{12(1-\nu^2)} & \frac{-E(1-3\nu)}{8(1-\nu^2)} \\ \frac{-E}{8(1-\nu)} & \frac{-E(3-\nu)}{12(1-\nu^2)} & \frac{-E(1-3\nu)}{8(1-\nu^2)} & \frac{-E(3+\nu)}{12(1-\nu^2)} & \frac{E}{8(1-\nu)} & \frac{E(3-\nu)}{6(1-\nu^2)} & \frac{E(1-3\nu)}{8(1-\nu^2)} & \frac{E\nu}{6(1-\nu^2)} \\ \frac{E\nu}{6(1-\nu^2)} & \frac{-E(1-3\nu)}{8(1-\nu^2)} & \frac{-E(3-\nu)}{12(1-\nu^2)} & \frac{E}{8(1-\nu)} & \frac{-E(3+\nu)}{12(1-\nu^2)} & \frac{E(1-3\nu)}{8(1-\nu^2)} & \frac{-E(3-\nu)}{6(1-\nu^2)} & \frac{-E}{8(1-\nu)} \\ \frac{E(1-3\nu)}{8(1-\nu^2)} & \frac{-E(3+\nu)}{12(1-\nu^2)} & \frac{E}{8(1-\nu)} & \frac{-E(3-\nu)}{12(1-\nu^2)} & \frac{-E(1-3\nu)}{8(1-\nu^2)} & \frac{E\nu}{6(1-\nu^2)} & \frac{-E}{8(1-\nu)} & \frac{E(3-\nu)}{6(1-\nu^2)} \end{bmatrix}, \quad (4.16)$$

where $\mathbf{I}_{8 \times 8}$ is an 8×8 identity matrix. The plane strain case can be retrieved from the plane stress case by using a fictitious Poisson ratio ν' instead of ν and a fictitious Young's modulus E' instead of E , which are given by

$$\nu' = \frac{\nu}{1-\nu} \quad \text{and} \quad E' = \frac{E}{1-\nu^2}. \quad (4.17)$$

In order to obtain more compact expressions, we define a dimensionless frequency ζ as $\zeta = \omega h / c_e$.

The penalized eigenvalue problem for a square plane stress element with lumped mass is found to be

$$\begin{aligned} & \zeta^4 \{ \zeta^2(\nu+1) - 4 \} \{ 3\zeta^2(\nu^2-1) - 2\nu + 6 \} \{ -9(1-\nu^2)^3(1+p_m)\zeta^8 \\ & + 6(1-\nu^2)^2(12p_m + (3-\nu)p_s + 15-\nu)\zeta^6 + 3(1-\nu^2)(p_m(31\nu^2+6\nu-57) \\ & + 16p_s(\nu-3) + 48\nu^2 + 16\nu - 96)\zeta^4 + 2(\nu-3)(18p_m(\nu^2-1) + p_s(31\nu^2 \\ & + 6\nu-57) + 48(\nu^2-1))\zeta^2 - 24(\nu-3)^2p_s \} = 0. \end{aligned} \quad (4.18)$$

Taking $p_m = p_s = 0$ yields $\zeta_{\max} = 2/\sqrt{1-\nu}$, irrespective of the magnitude of ν . We substitute this value for ζ into equation (4.18) to obtain $6(\nu+1)p_m - (3-\nu)p_s = 0$, hence $r_{\text{crit}} = 6(1+\nu)/(3-\nu)$ for plane stress. The equivalent plane strain case is found by using the expression for ν' as given in equation (4.17). As a result, $r_{\text{crit}} = 6/(3-4\nu)$ for plane strain.

Using the consistent mass matrix, the eigenvalue problem for a square plane stress element leads to a polynomial equation in ζ as

$$\begin{aligned} & \zeta^4 \{ \zeta^2(\nu+1) - 12 \} \{ \zeta^2(\nu^2-1) - 6\nu + 18 \} \{ -(1-\nu^2)^3(9 + 16p_m)\zeta^8 - 6(1-\nu^2)^2 \\ & \times (4p_m(\nu-19) + 4p_s(\nu-3) + 9\nu - 63)\zeta^6 - 36(1-\nu^2)(p_m(-49\nu^2 - 10\nu + 103) \\ & + p_s(\nu^2 - 22\nu + 57) - 36\nu^2 - 36\nu + 144)\zeta^4 + 54(\nu-3)(40p_m(\nu^2-1) \\ & + p_s(49\nu^2 + 10\nu - 103) + 144\nu^2 - 144)\zeta^2 - 3240p_s(\nu-3)^2 \} = 0. \end{aligned} \quad (4.19)$$

The maximum dimensionless sampling frequency depends on the value of the Poisson ratio. For $\nu \geq \frac{1}{3}$ we obtain $\zeta_{\max} = \sqrt{12/(1-\nu)}$, which implies for the penalized system that $8(1+\nu)p_m - (3-\nu)p_s = 0$, therefore $r_{\text{crit}} = 8(1+\nu)/(3-\nu)$ for plane stress. The counterpart transition value for ν in plane strain is $\nu = \frac{1}{4}$. For $\nu \geq \frac{1}{4}$ the maximum dimensionless sampling frequency of the unpenalized system is $\zeta_{\max} = \sqrt{12/(1+\nu)(1-2\nu)}$. For the penalized eigenvalue problem this value leads to $8p_m - (3-4\nu)p_s = 0$ and therefore $r_{\text{crit}} = 8/(3-4\nu)$ for plane strain. No solutions for r_{crit} could be found for the case $\nu < \frac{1}{3}$ in plane stress and $\nu < \frac{1}{4}$ in plane strain. In the numerical examples we will use the expressions for r_{crit} given above irrespective of ν , and verify these are safe for $\nu < \frac{1}{3}$ and $\nu < \frac{1}{4}$, respectively.

5. Results

In this section we will present a selection of results on stability and accuracy of the bipenalty method. To report on stability and accuracy for all possible combinations of element types, mass matrices, time integration schemes and penalty ratios would be excessive, therefore the discussion is limited to the following issues:

- *For bar elements.* Illustrate the destabilizing effects of choosing the penalty ratio r larger than its critical value. This is done for two combinations of mass matrix and time integration scheme.
- *For bar elements.* Illustrate how the penalty ratio r affects the accuracy of constraint imposition. This is again done for two combinations of mass matrix and time integration scheme.
- *For beam elements.* Verify the effects on stability of penalizing deflection, rotation or both. This is done for explicit central difference time integration with a lumped mass matrix only.
- *For square two-dimensional elements.* Demonstrate that the analytically obtained expressions for the critical penalty ratio r_{crit} are valid for values of the Poisson ratio $\nu \in [0, \frac{1}{2})$. This is done for the central differences time integration scheme combined with a consistent mass matrix.

(a) Bar problem—stability analysis

The first problem is a one-dimensional bar of length $L = 100$ m subjected to a tensile traction $\sigma = 1 \text{ N m}^{-2}$ on the left end. The traction is applied from time $t = 0$ s onwards and kept constant for the duration of the simulation up to $t = 150$ s. The right end of the bar is fixed, and this support condition is modelled with penalty functions. The material parameters are taken as $E = 1 \text{ N m}^{-2}$ and $\rho = 1 \text{ kg m}^{-3}$. A finite element mesh consisting of 100 two-noded bar elements is used. We will show results for two combinations of Newmark parameters and mass matrices, namely (i) the explicit central differences scheme obtained via $\beta = 0$ and $\gamma = \frac{1}{2}$ with a lumped mass matrix, so that the critical time step $\Delta t_{\text{crit}} = 1$ s, and (ii) the Fox–Goodwin scheme found with $\beta = \frac{1}{12}$ and $\gamma = \frac{1}{2}$ in combination with a consistent mass matrix, so that $\Delta t_{\text{crit}} = 1/\sqrt{2}$ s. In either case, we apply a time step

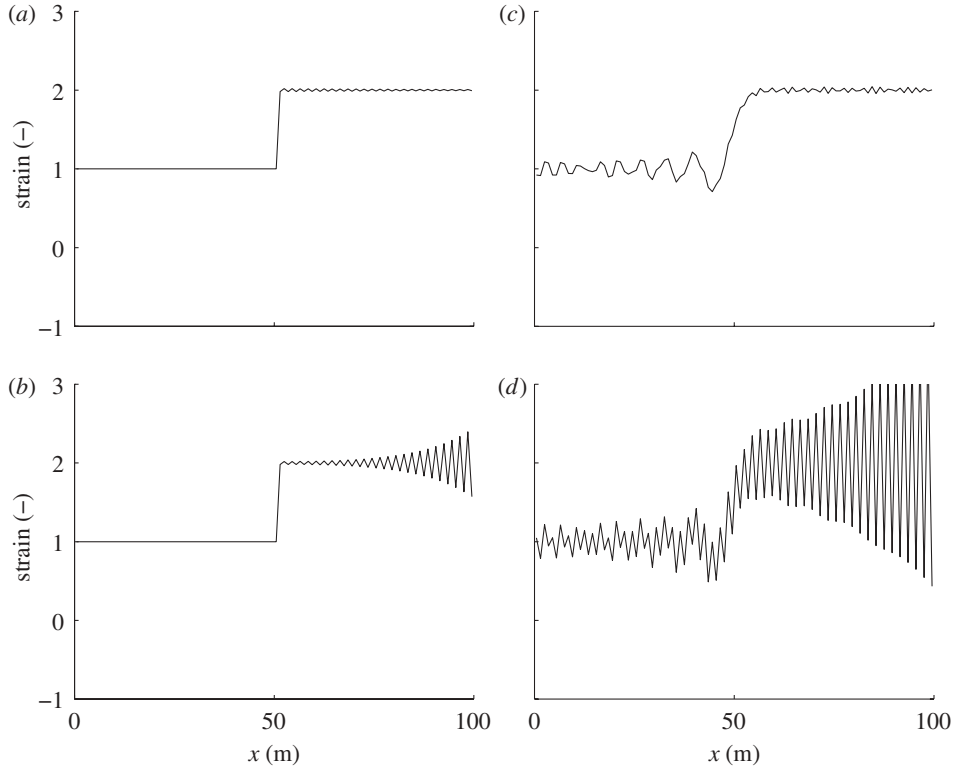


Figure 1. Stability analysis of the bar problem with inertia penalty $p_m = 10^2$ —strain profiles across the bar at time $t = 150$ s: (a) lumped mass with $r = 2 = r_{\text{crit}}$, (b) lumped mass with $r = 2.004 > r_{\text{crit}}$, (c) consistent mass with $r = 4 = r_{\text{crit}}$ and (d) consistent mass with $r = 4.004 > r_{\text{crit}}$.

$\Delta t = \Delta t_{\text{crit}}$ so that no instabilities are expected while the wave propagates through the bar—any instabilities that arise are thus entirely due to the penalization of the right end of the bar. Our purpose is not to cross compare these two sets of results, but instead to show that our findings on the critical penalty ratio hold for each combination of mass matrix and time integration scheme.

Figures 1 (using an inertia penalty factor $p_m = 10^2$) and 2 (using an inertia penalty factor $p_m = 10^4$) show the strain profiles across the bar at time $t = 150$ s for the various combinations of Newmark parameters and mass matrices, while also stable and unstable values for the ratio $r = p_s/p_m$ were taken. At time $t = 150$ s the wave front has propagated through the entire bar, been reflected on the penalized right end, and returned halfway along the bar. It can be verified that values of $r = r_{\text{crit}}$ result in stable wave reflection on the penalized boundary, whereas values of $r > r_{\text{crit}}$ lead to destabilizing wave reflection. The particular unstable values of r in figures 1 and 2 have been selected such that the instabilities are visible at time $t = 150$ s but not grown out of proportion as yet. Nevertheless, it is emphasized that as long as $\Delta t = \Delta t_{\text{crit}}$ and $r = p_s/p_m > r_{\text{crit}}$, all instabilities are initiated when the wave front reaches the penalized boundary. This observation holds irrespective of the magnitude of the penalty factors p_m and p_s .

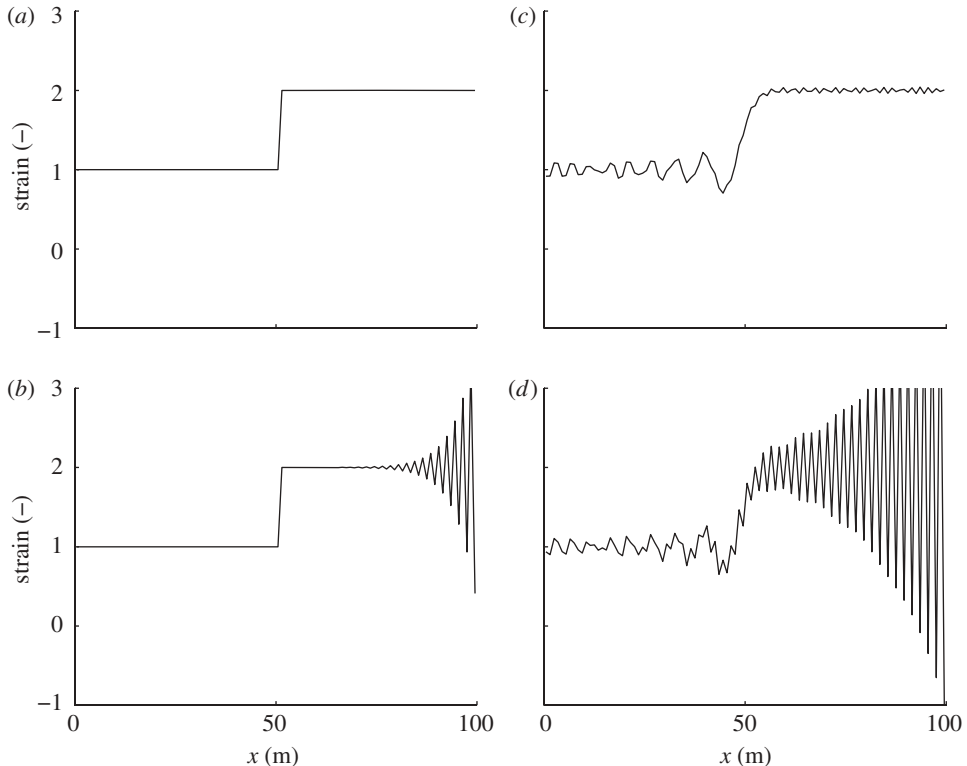


Figure 2. Stability analysis of the bar problem with inertia penalty $p_m = 10^4$ —strain profiles across the bar at time $t = 150$ s: (a) lumped mass with $r = 2 = r_{\text{crit}}$, (b) lumped mass with $r = 2.02 > r_{\text{crit}}$, (c) consistent mass with $r = 4 = r_{\text{crit}}$, and (d) consistent mass with $r = 4.01 > r_{\text{crit}}$.

(b) Bar problem—accuracy analysis

Next, the accuracy aspects of the bipenalty method are studied for the same bar problem as in §5a. The inertia penalty factor $p_m = 10^4$ throughout and different stiffness penalty factors p_s are taken. According to equation (3.13), this may affect the accuracy with which the constraint is modelled. The same combinations of mass matrix and time integration scheme are used, and $\Delta t = \Delta t_{\text{crit}}$ in all cases.

Figure 3 shows the time history plots of displacement (figure 3a–c) and acceleration (figure 3d–f) of the constrained d.f. for the case where lumped mass and central difference time integration ($\beta = 0$, $\gamma = \frac{1}{2}$) is used. Three values for the penalty ratio r are taken, namely $r = 0$ (figure 3a,d), $r = 1 = \frac{1}{2}r_{\text{crit}}$ (figure 3b,e) and $r = 2 = r_{\text{crit}}$ (figure 3c,f). For this case, the Newmark parameter $\beta = 0$ and equation (3.13) predicts that all three values of r should lead to the same convergence rates for the constrained acceleration. This seems to be verified in figure 3, in which the maximum error in the acceleration is about $-4 \times 10^{-4} \text{ m s}^{-2}$, irrespective of the value for r . However, the error in the acceleration is monotonic in case $r = 0$, which leads to strongly accumulating error in the displacement, see figure 3a. On the other hand, taking $r > 0$ leads to oscillating errors in the acceleration, the result of which is that the displacement error does not

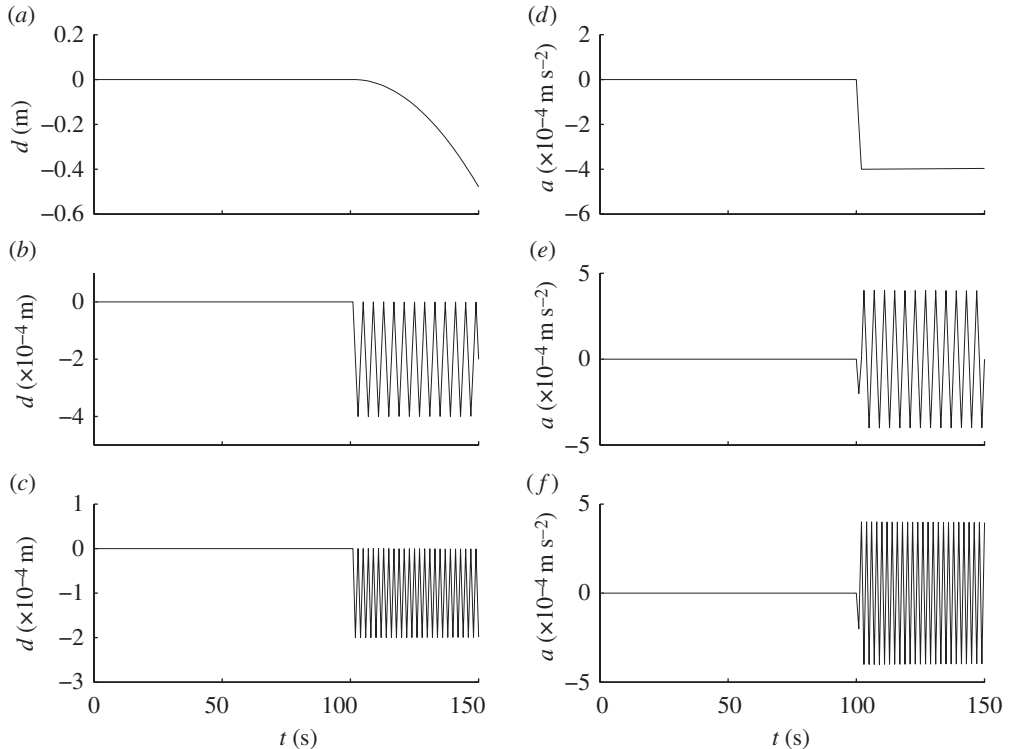


Figure 3. Accuracy analysis of the bar problem with lumped mass and central difference scheme—penalized displacement (a–c) and penalized acceleration (d–f), (a,d) $r=0$, (b,e) $r=\frac{1}{2}r_{\text{crit}}$ and (c,f) $r=r_{\text{crit}}$.

accumulate but instead remains bounded. These oscillations in the acceleration error are of higher frequency for larger r (i.e. for larger values of the stiffness penalty), which also reduces the magnitude of the displacement error.

Figure 4 shows similar results for the combination of consistent mass with Fox–Goodwin time integration ($\beta = \frac{1}{12}$, $\gamma = \frac{1}{2}$). The three values of the penalty ratio are here taken as, from top to bottom, $r=0$, $r=2 = \frac{1}{2}r_{\text{crit}}$ and $r=4 = r_{\text{crit}}$. Since now $\beta = \frac{1}{12}$, there is a slight effect of r on the acceleration error, cf. equation (3.13): larger values for r imply larger values for the stiffness penalty and lower error in the acceleration. The relation between r and the displacement error is similar to that in figure 3. Zero stiffness penalty ($r=0$) leads to significant accumulation of the displacement error, whereas finite stiffness penalty ($r>0$) gives bounded displacement error. Moreover, the larger the stiffness penalty, the smaller the maximum error in the displacement.

(c) Beam problem—stability analysis

Next, a horizontal Euler–Bernoulli beam of length $L=50$ m is considered. The cross-sectional area $A=1$ m², the second moment of area $I=\frac{1}{12}$ m⁴, the Young’s modulus $E=1$ N m⁻² and the mass density $\rho=1$ kg m⁻³. An upward force

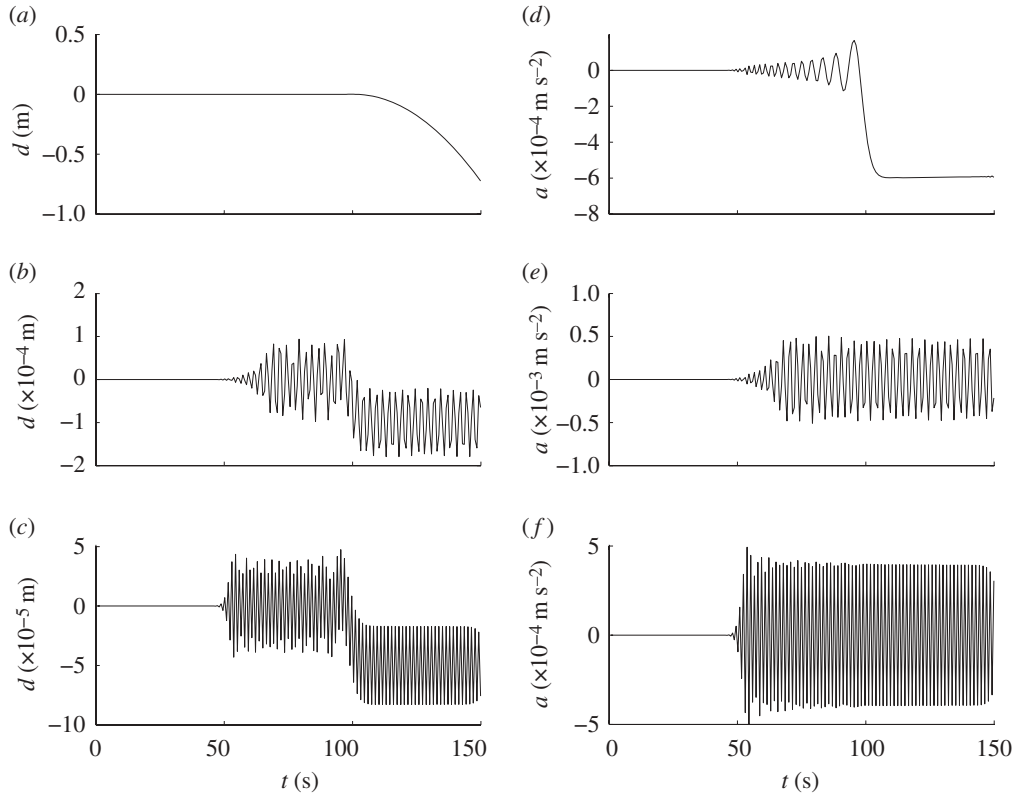


Figure 4. Accuracy analysis of the bar problem with consistent mass and Fox–Goodwin scheme—penalized displacement (a–c) and penalized acceleration (d–f), (a,d) $r = 0$, (b,e) $r = \frac{1}{2}r_{\text{crit}}$ and (c,f) $r = r_{\text{crit}}$.

$F = 0.01/\Delta t N$ is applied at the left end of the beam during the first time step, and support conditions are applied via penalty functions at the right end of the beam—see below for specific details on the support conditions. The beam is subdivided into 50 equal-sized elements. The lumped mass matrix is used. Time integration is carried out with the central differences scheme and $\Delta t = \Delta t_{\text{crit}}$. The inertia penalty factor $p_m = 10^4$.

In §4b critical penalty ratios were derived for different support conditions, namely $r_{\text{crit}} = 8$ for a penalized deflection, $r_{\text{crit}} = 2$ for a penalized rotation and $\lim_{p_m \rightarrow \infty} r_{\text{crit}} = 2$ for the combination of penalized deflection with penalized rotation. However, we will simply assume $r_{\text{crit}} = 2$ for the latter case, since in practice $p_m \gg 1$. In figure 5 displacement profiles across the beam are plotted at time $t = 500$ s. The effectiveness of using the bipenalty method on imposing the constraints can be verified qualitatively in figure 5a–c. In figure 5d–f the occurrence of instabilities can be verified in case the penalty ratio is chosen just 0.1 per cent larger than its critical value—this indicates that the derived values for the critical penalty ratio r_{crit} are crisp, even the approximative value of r_{crit} for the combination penalized deflection/penalized rotation.

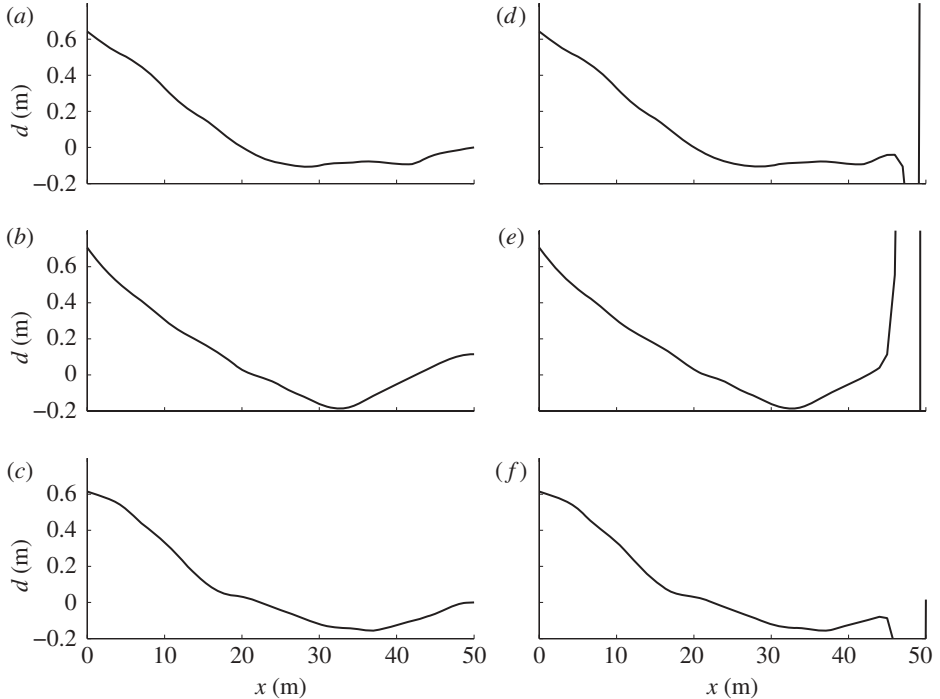


Figure 5. Stability analysis of the beam problem—critical penalty ratio $r = r_{\text{crit}}$ (*a-c*) and supercritical penalty ratio $r = 1.001r_{\text{crit}}$ (*d-f*); (*a, d*) penalized deflection, (*b, e*) penalized rotation and (*c, f*) penalized deflection together with penalized rotation.

(*d*) Plane strain problem—stability analysis

The final example uses square plane strain elements. A bar-type structure of height $H = 1$ m and length $L = 10$ m is modelled with 10 square four-noded finite elements. The Young’s modulus $E = 1 \text{ N m}^{-2}$, the mass density $\rho = 1 \text{ kg m}^{-3}$ and a range of Poisson ratios were used (see below). Two leftward forces of 0.01 N each are applied on the two leftmost nodes from time $t = 0$ onwards. The horizontal displacement of the top-right node is constrained using the bipenalty method. The consistent mass matrix is used with central difference time integration and $\Delta t = \Delta t_{\text{crit}}$.

In §4*c* an expression for the critical penalty factor r_{crit} was derived that is valid for a certain range of values for the Poisson ratio, namely $\frac{1}{4} \leq \nu < \frac{1}{2}$ in plane strain. As mentioned in §4*c*, an expression for r_{crit} could not be found for $\nu < \frac{1}{4}$. It is therefore of interest to verify whether the expression for $\frac{1}{4} \leq \nu < \frac{1}{2}$ could be used as well for $\nu < \frac{1}{4}$. Figure 6 shows the deformed mesh at time $t = 50$ s, whereby the displacements have not been magnified. Various values were used for the Poisson ratio, namely $\nu \in \{0.1, 0.2, 0.3, 0.4\}$. Figure 6*a-d* shows results obtained with using a critical penalty ratio, whereas the results in figure 6*e-h* were found with a larger value for the penalty ratio, that is $r = 1.004r_{\text{crit}}$. Taking $r > r_{\text{crit}}$ leads to instabilities in case $\nu = 0.3$ and, more spectacularly, $\nu = 0.4$, which are within the range of the derived expression for r_{crit} as given in §4*c*.

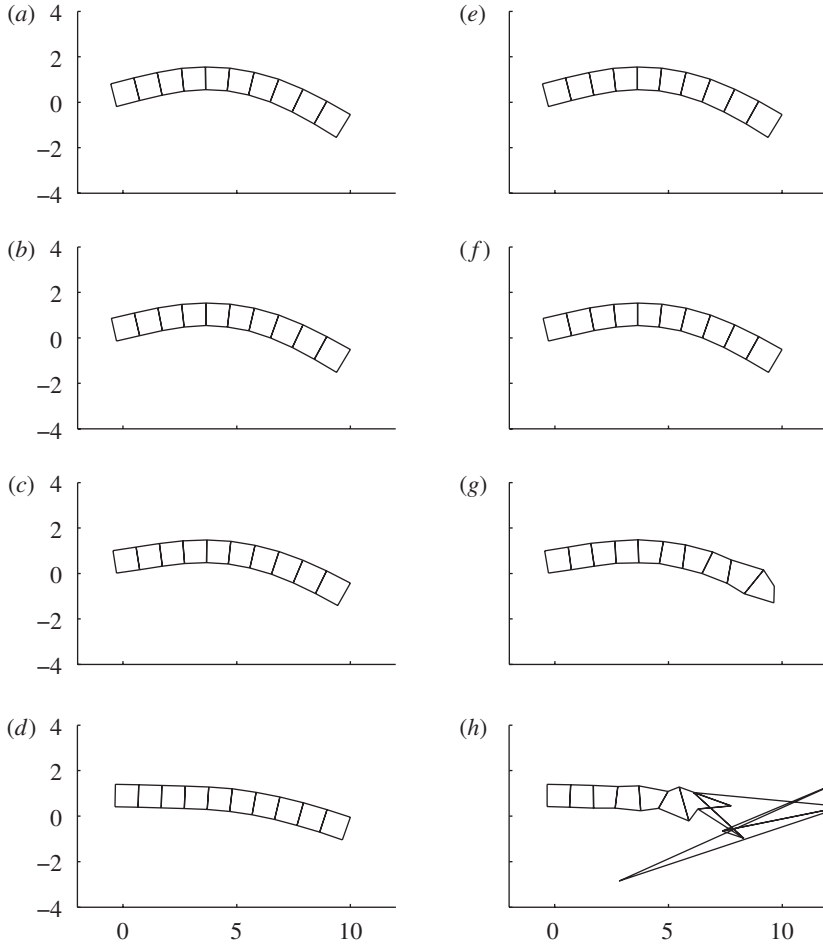


Figure 6. Stability analysis of the two-dimensional plane strain problem—critical penalty ratio $r = r_{\text{crit}}$ (a–d) and supercritical penalty ratio $r = 1.004 \cdot r_{\text{crit}}$ (e–h); Poisson’s ratio: (a,e) $\nu = 0.1$, (b,f) $\nu = 0.2$, (c,g) $\nu = 0.3$ and (d,h) $\nu = 0.4$.

Taking $r > r_{\text{crit}}$ does not seem to lead to instabilities for values of Poisson’s ratio $\nu < \frac{1}{4}$, which is an indication that the suggested expression for r_{crit} is safe for $\nu < \frac{1}{4}$.

6. Discussion and concluding remarks

In this paper we have introduced the bipenalty method, which consists of the simultaneous use of stiffness-type penalties and inertia-type penalties in the enforcement of constraints. The bipenalty method has significant advantages over the use of either stiffness penalties or inertia penalties. Compared with stiffness penalties, the bipenalty method can be tuned such that the critical time step of conditionally stable time integration schemes is not affected—this can lead

to significant savings in computing time. Compared with inertia penalties, the bipenalty method is more accurate when penalty parameters of equal magnitude are taken.

The performance of the bipenalty method is assessed in terms of the ratio of stiffness penalty over inertia penalty, which is termed the ‘penalty ratio’. For reasons of numerical stability in conditionally stable time integration schemes, the penalty ratio is bounded from above. Closed form expressions for the so-called ‘critical penalty ratios’ have been derived for a range of element types, namely bar elements, beam elements and square plane stress/plane strain elements, including a distinction between lumped mass matrices and consistent mass matrices. These values are summarized in table 1.

For bar elements and beam elements, all the derived expressions for the critical penalty ratio have been confirmed by numerical experimentation (although not all individual results were presented). We have found that results are always stable if the penalty ratio is smaller than the critical penalty ratio, and results destabilize if the penalty ratio is larger than the critical penalty ratio. For the square plane stress/plane strain elements, two open issues must be mentioned:

- When using a consistent mass matrix, expressions for the critical penalty ratio could not be found for all values of the Poisson ratio. However, numerical experimentation shows that this does not seem to be a problem; the expressions for the critical penalty ratio are safe for the whole range of tested Poisson’s ratios.
- The theoretical derivations and the numerical simulations were carried out with a single penalized d.f. However, in practice it is common that a single finite element has multiple constrained d.f. Preliminary derivations have not led to meaningful expressions for the critical penalty ratio, and further research is required to resolve this.

To summarize, the bipenalty method is superior to stiffness penalties (for reasons of stability; moreover, since time step restrictions can be avoided in the bipenalty method CPU times tend to be smaller than with stiffness penalties) and superior to inertia penalties (for reasons of accuracy). Although the term critical penalty ratio refers to the *stability* aspects of the bipenalty method, it was also found that the most *accurate* results are obtained if the penalty ratio is as close as possible to the critical penalty ratio—thus, the critical penalty ratio for stability is also the optimum penalty ratio for accuracy.

References

- Askes, H. & Ilanko, S. 2006 The use of negative penalty functions in linear systems of equations. *Proc R. Soc. A* **462**, 2965–2975. (doi:10.1098/rspa.2006.1716)
- Askes, H., Piercy, S. & Ilanko, S. 2008 Tying in linear systems of equations modelled with positive and negative penalty functions. *Commun. Numer. Meth. Eng.* **24**, 1163–1169.
- Bathe, K.-J. 1996 *Finite element procedures*. Englewood Cliffs, NJ: Prentice-Hall.
- Belytschko, T. & Neal, M. 1991 Contact-impact by the pinball algorithm with penalty and Lagrangian methods. *Int. J. Numer. Meth. Eng.* **31**, 547–572. (doi:10.1002/nme.1620310309)
- Carpenter, N., Taylor, R. & Katona, M. 1991 Lagrange constraints for transient finite element surface contact. *Int. J. Numer. Meth. Eng.* **32**, 103–128. (doi:10.1002/nme.1620320107)

- Hetherington, J. & Askes, H. 2009 Penalty methods for time domain computational dynamics based on positive and negative inertia. *Comput. Struct.* **87**, 1474–1482.
- Hughes, T. 2000 *The finite element method: linear static and dynamic finite element analysis*. New York, NY: Dover.
- Ilanko, S. 2002a Existence of natural frequencies of systems with imaginary restraints and their convergence in asymptotic modelling. *J. Sound Vib.* **255**, 883–898. (doi:10.1006/jsvi.2001.4191)
- Ilanko, S. 2002b The use of negative penalty functions in constrained variational problems. *Commun. Numer. Meth. Eng.* **18**, 659–668. (doi:10.1002/cnm.525)
- Ilanko, S. 2003 The use of asymptotic modelling in vibration and stability analysis of structures. *J. Sound Vib.* **263**, 1047–1054. (doi:10.1016/S0022-460X(03)00273-6)
- Ilanko, S. 2005a Asymptotic modelling theorems for the static analysis of linear elastic structures. *Proc. R. Soc. A* **461**, 3525–3542. (doi:10.1098/rspa.2005.1519)
- Ilanko, S. 2005b Introducing the use of positive and negative inertial functions in asymptotic modelling. *Proc. R. Soc. A* **461**, 2545–2562. (doi:10.1098/rspa.2005.1472)
- Ilanko, S. & Dickinson, S. 1999 Asymptotic modelling of rigid boundaries and connections in the Rayleigh–Ritz method. *J. Sound Vib.* **219**, 370–378. (doi:10.1006/jsvi.1998.1847)
- Ishihara, K. 1981 A finite element lumped mass scheme for solving eigenvalue problems of circular arches. *Numer. Math.* **36**, 267–290. (doi:10.1007/BF01396655)
- Kim, K.-O. 1993 A review of mass matrices for eigenproblems. *Comput. Struct.* **46**, 1041–1048. (doi:10.1016/0045-7949(93)90090-Z)
- Ling, X. & Cherukuri, H. 2002 Stability analysis of an explicit finite element scheme for plane wave motions in elastic solids. *Comput. Mech.* **29**, 430–440. (doi:10.1007/s00466-002-0353-8)
- Williams, F. & Ilanko, S. 2005 The use of reciprocals of positive and negative inertial functions in asymptotic modelling. *Proc. R. Soc. A* **462**, 1909–1915. (doi:10.1098/rspa.2005.1643)