ELSEVIER

Comput. Methods Appl. Mech. Engrg. 171 (1999) 141–171

Computer methods in applied mechanics and engineering

# Geometrically exact 3D beam theory: implementation of a strain-invariant finite element for statics and dynamics

G. Jelenić, M.A. Crisfield*

*Department of Aeronautics, Imperial College of Science, Technology and Medicine, London SW7 2BY, UK*

Received 8 December 1997

Dedicated to the memory of Prof. F.B. Damjanić, who brought the present two authors together

**Abstract**

Geometrically exact 3D beam theory has been used as a basis for development of a variety of finite element formulations. It has recently become apparent that the important requirement of objectivity of adopted strain measures, although provided by the theory itself, does not automatically extend to a finite element formulation. In this paper we present a new finite element formulation of the geometrically exact 3D beam theory, specifically designed to preserve the objectivity of the adopted strain measures. In order to do so the current local rotations are interpolated in a manner similar to that adopted in co-rotational approaches. However, no approximations typical for co-rotational approaches are introduced into the procedure, so in contrast to co-rotational formulations, the present formulation fully preserves the geometric exactness of the theory. A range of numerical examples serves to illustrate the problem and to assess the formulation. 

## 0. Notation

| | |
|---|---|
| $L, x$ | initial length of the beam, arc-length parameter of the beam centroid axis |
| $(x, y, z), (X, Y, Z)$ | coordinates of element and inertial Cartesian frames |
| $\mathcal{R}, \mathcal{N}$ | sets of real numbers and positive integers |
| $SO(3)$ | three-parametric Lie group of 3D rotations |
| $\delta^i_j, \Delta^{ij}_k, e^{ij}_k$ | Kronecker symbol, generalised Kronecker symbol and permutation symbol |
| $\boldsymbol{I}, \boldsymbol{\alpha}$ | $3 \times 3$ unit matrix, any three-dimensional rotational pseudovector |
| $\boldsymbol{E}_i, \boldsymbol{G}_{0,i}, \boldsymbol{G}_{n,i}(x)$ | orthonormal base vectors of inertial coordinate system, element (initial) frame and the frame attached at the cross section at $x$ at time $t_n$ $(i = 1, 2, 3)$ |
| $\boldsymbol{r}, \boldsymbol{\Lambda}$ | position vector and rotation matrix |
| $t, t_0, \Delta t$ | time, initial time, time step length |
| $\alpha, \beta, \gamma$ | Parameter in HHT method and Newmark's integration parameters |
| $\boldsymbol{n}, \boldsymbol{m}$ | vectors of applied distributed forces and torques |
| $\boldsymbol{F}_0, \boldsymbol{T}_0, \boldsymbol{F}_L, \boldsymbol{T}_L$ | applied concentrated forces and torques at ends of the beam |
| $\boldsymbol{N}, \boldsymbol{M}$ | vectors of stress and stress-couple resultants |
| $\boldsymbol{\gamma}, \boldsymbol{\kappa}$ | vectors of translational and rotational strain measures |
| $\boldsymbol{C}_N, \boldsymbol{C}_M, \boldsymbol{c}_N, \boldsymbol{c}_M$ | material and spatial translational and rotational constitutive matrices |
| $\rho, E, G$ (in $\boldsymbol{C}_N$) | density of material, Young's (elastic) modulus and shear modulus |

* Corresponding author. FEA Professor.

0045-7825/99/$ – see front matter 
PII: S0045-7825(98)00249-7

| | |
|---|---|
| $A, A_2, A_3$ | cross-sectional area and shear areas |
| $J_t, I_2, I_3$ | torsional and cross-sectional moments of inertia |
| $\boldsymbol{\pi}, \boldsymbol{J}_\rho$ | specific angular momentum, tensor of mass moments of inertia |
| $G$ (if not in $\boldsymbol{C}_N$) | weak form |
| $\mathscr{F}^h$, $\mathscr{P}^h$ | spaces of approximated translational and rotational test functions |
| $\boldsymbol{f}, \boldsymbol{\varphi}$ | translational and rotational test functions |
| $\mathscr{U}^h$, $\mathscr{S}^h$ | spaces of approximated translational and rotational trial functions |
| $\Delta\boldsymbol{u}, \Delta\boldsymbol{\vartheta}$ | translational and rotational trial functions ($\Delta\boldsymbol{\vartheta}$ are spin variables) |
| $\dot{\boldsymbol{\vartheta}}, \ddot{\boldsymbol{\vartheta}}$ | angular velocity, angular acceleration |
| $\boldsymbol{u}, \dot{\boldsymbol{u}}, \ddot{\boldsymbol{u}}$ | displacement, velocity and acceleration |
| $\boldsymbol{\psi}, \boldsymbol{\theta}, \psi, \theta$ | total and incremental rotational pseudovectors and their norms |
| $\boldsymbol{T}(\boldsymbol{\alpha})$ | matrix relating $\Delta\boldsymbol{\vartheta}$ to $\Delta\boldsymbol{\alpha}$ |
| $\boldsymbol{\Theta}, \boldsymbol{W}, \boldsymbol{A}$ | incremental rotation, angular velocity and angular acceleration with components in the body-attached frame |
| $\tilde{\boldsymbol{q}}, q_0, \boldsymbol{q}$ | quaternion and its scalar and vector parts |
| $N, I^i$ | number of nodes on finite element, polynomial shape functions ($i = 1, \ldots, N$) |
| $\tilde{I}^i$ | generalised shape functions ($i = 1, \ldots, N$) |
| $\boldsymbol{\Lambda}_r, \boldsymbol{\Psi}^I$ | reference triad and local rotation in strain-invariant formulation |
| $\boldsymbol{\phi}_{IJ}$ | relative local rotation between nodes $I$ and $J$ in strain-invariant formulation |
| $I, J$ | fixed nodes in strain-invariant formulation |
| $\boldsymbol{g}^i, \boldsymbol{q}^i_m, \boldsymbol{q}^i_k, \boldsymbol{q}^i_e$ | residual and vectors of inertial, internal and external forces at node $i$ |
| $\boldsymbol{K}^{ij}_k, \boldsymbol{K}^{ij}_m$ | stiffnes matrix and inertia matrix relating the trial functions at node $j$ to the residual at node $i$ |
| $\delta_u, \delta_f$ | displacement and residual norms |
| $\underline{(\bullet)}, (\bullet)_R$ | $(\bullet)$ acted upon by a constant motion, constant quantity $(\bullet)$ |
| $(\bullet)', (\dot{\bullet})$ | derivatives of $(\bullet)$ with respect to arc-length coordinate $x$ and time $t$ |
| $(\hat{\bullet}), (\bullet)^t$ | cross-product operator $(\bullet)\times$, transposed quantity $(\bullet)$ |
| $(\bullet)_0, (\bullet)_n, (\bullet)_{n+1}$ | $(\bullet)$ at times $t_0$, $t_n$ and $t_{n+1}$ |
| $(\bullet)_i, (\bullet)_j$ | $(\bullet)$ at nodes $i$ and $j$ |
| $(\bullet)_I, (\bullet)_J$ | $(\bullet)$ at fixed nodes $I$ and $J$ |
| $(\bullet)^h$, $\Delta(\bullet)$ | approximated quantity $(\bullet)$, variation of $(\bullet)$ |

## 1. Introduction

Geometrically exact beam theory and its finite element implementation have been the subject of extensive research in the recent past. In the present paper, we are concerned with the geometrically exact 3D beam theory of Reissner [1,2] and Simo [3]. This theory has provided the basis of many of the recent finite element formulations for 3D beams [4–9]. In contrast to some earlier formulations, which were based on degenerate continuum concepts [10,11], those based on the geometrically exact theory stem directly from a stress resultant approach. The geometrically exact theory provides the relationships between the configuration and the adopted strain measures which are fully consistent with the virtual work principle and the differential equations of motion, regardless of the magnitude of displacements, rotations and strains involved [1,2].

Geometrically exact beam theory is sometimes referred to as the 'Reissner's beam theory', but strictly speaking the latter is geometrically exact only in 2D [1]. Treatment of rotations in 3D becomes nontrivial primarily because of the nonlinear character of the space of 3D rotations and in [2] Reissner proposed a simplification of the rotation matrix, which enabled the derivation of the required strain–configuration relationship, but unfortunately, also spoiled the geometric exactness of the theory. A different approach was proposed by Simo [3], who utilised a relationship between the variation of the rotation matrix and the infinitesimal rotations to derive the strain-configuration relations without affecting the geometric exactness.

In the implementation proposed by Simo and Vu-Quoc [4], the genuine infinitesimal (iterative) rotational changes were interpolated along the length of the beam. In contrast, Cardona and Géradin [5] regarded the

update of 3D rotations on the basis of the last converged configuration as the most suitable choice and consequently chose to interpolate incremental rotational changes. A third interesting possibility involves updating the 3D rotations on the basis of the initial configuration and leads to the interpolation of the total rotational pseudovectors [5,6]. In all of these cases, the adopted rotational fields were interpolated additively, despite the fact that 3D rotations cannot be additively composed. This problem was analysed in [12,13] and it was shown that the finite element interpolation of all of these rotation variables spoils the objectivity of the adopted strain measures with respect to a rigid rotation. The formulations based on the interpolation of iterative and incremental rotations were additionally proved to be dependent on the history of deformation. Interestingly, the co-rotational technique [14,15] does not necessarily suffer from these drawbacks. This observation allows us to devise a new approach, which combines good characteristics of the geometrically exact and the co-rotational beam theories.

In this paper we elaborate the idea outlined in [12] and derive a new, strain-invariant and path-independent, geometrically exact 3D beam element. The key to the formulation is found in the interpolation of current local rotations in a co-rotational manner, yet without imposing any limitations on their magnitude. The proposed method differs from conventional approaches [4–9] in the interpolation and update, which are now dictated by the objectivity conditions. By introducing the notion of 'generalised shape functions', the proposed method takes a similar form to that of conventional approaches. In particular, this concept enables a highly desirable feature of the kinetic energy being directly related to an inertial frame [8,16], despite the fact that the genuine interpolation is applied to current *local* rotations.

In Section 2, we give a short account of the geometrically exact beam theory and devote Section 3 to the main discussion on the different choices of finite element interpolation for the rotational degrees of freedom. This will be shown to directly influence the objectivity and path-dependence properties of the analysed formulations. The reader may find it useful to complement the derivations in Section 3 with the results of [12], where a more detailed theoretical analysis is given. In Section 4 we show that the concept of generalised shape functions enables a finite element problem definition which takes a very similar form to that of the existing approaches for geometrically exact 3D beams. The numerical assessment of the various formulations is given in Section 5 and the final conclusions are drawn in Section 6.

## 2. Dynamics of geometrically exact 3D beams

This section summarises the application of the geometrically exact 3D beam theory to problems of elastic motion. More detailed accounts can be found in: [3,17] for the kinematics of 3D beams, [5,7,8,18–20] for the numerical time integration, [2–4,12] for the strain measures, [4] for the constitutive equations and [4,7,8] for the weak form of the equations of motion.

### 2.2. Kinematics of 3D beams

The line of centroids of cross sections of the undeformed beam element at time $t_0 = 0$ is, for the sake of simplicity, taken to be a straight line which coincides with the $x$ axis of the element Cartesian frame $(x, y, z)$ with $\boldsymbol{G}_{0,1}, \boldsymbol{G}_{0,2}, \boldsymbol{G}_{0,3}$ as the unit base vectors (Fig. 1). The initial position vector of a material particle $(x, 0, 0)$ on the line of centroids with respect to the inertial frame $(X, Y, Z)$ with $\boldsymbol{E}_1, \boldsymbol{E}_2, \boldsymbol{E}_3$ as the orthonormal base vectors is denoted by $\boldsymbol{r}_0(x)$; $x \in [0, L] \subset \mathscr{R}$, with $L$ being the *initial length of the beam*. Cross sections of the undeformed beam in the coordinate plane $x = \text{const.}$ are assumed to be orthogonal to the line of centroids and their normals coincide with the base vector $\boldsymbol{G}_{0,1}(x) = \boldsymbol{r}_0'(x)$, where a dash $(')$ denotes a derivative with respect to the arc-length parameter $x$. The remaining two base vectors, $\boldsymbol{G}_{0,2}$ and $\boldsymbol{G}_{0,3}$, are directed along the principal axes of inertia of the cross section to make a right-handed orthonormal triad.

The geometric shape of the cross section is assumed to be arbitrary and constant along the axis of the beam. Consequently, an initial position vector $\boldsymbol{R}_0(x, y, z)$ of an arbitrary material particle $(x, y, z)$ would be defined as

$$\boldsymbol{R}_0(x, y, z) = \boldsymbol{r}_0(x) + \boldsymbol{\Lambda}_0(x)\begin{Bmatrix} 0 \\ y \\ z \end{Bmatrix}; \quad x \in [0, L] \subset \mathscr{R}\,,\ (y, z) \in \mathscr{A}\,,$$

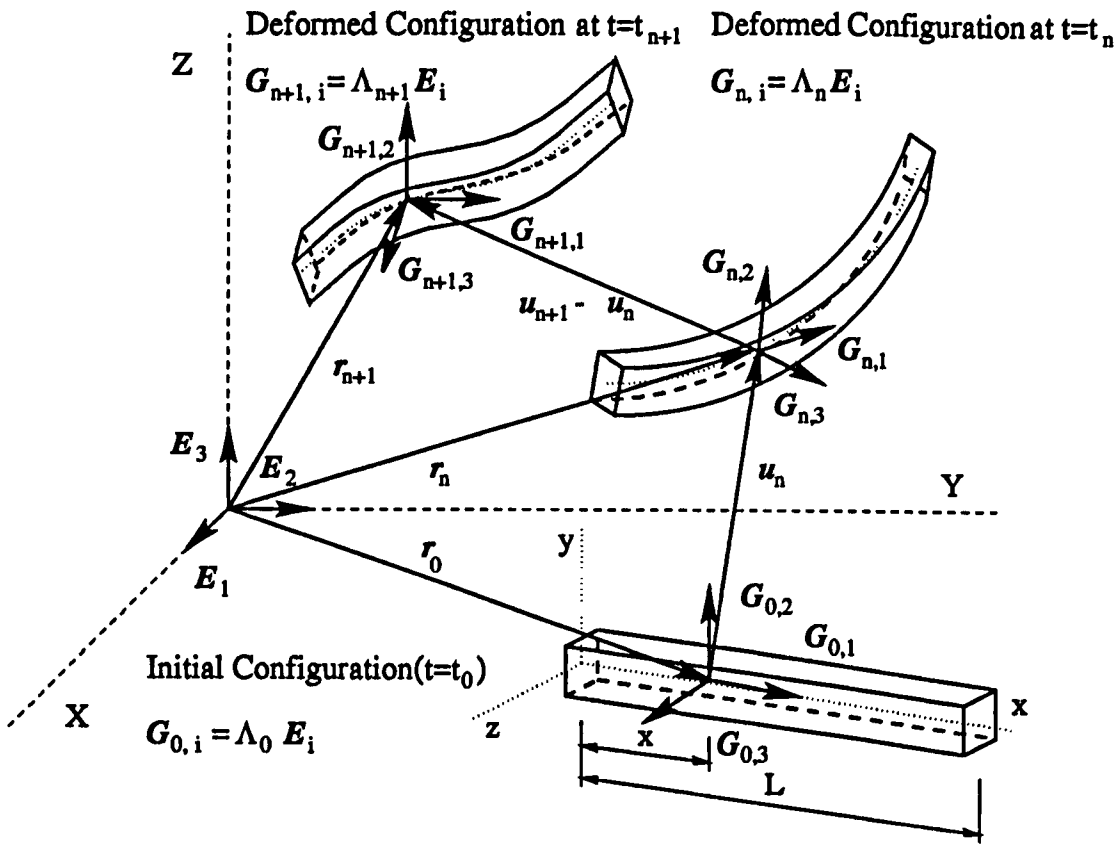



Fig. 1. Initial and deformed configuration of the beam.

where $y$ and $z$ are the coordinates of the material particle within a cross section $\mathscr{A} \subset \mathscr{R}^2$ at $x$ with respect to its centroid and $\Lambda_0$ is an element of the Lie group of proper orthogonal transformations $SO(3)$ with well-known properties $\det \Lambda_0 = 1$; $\Lambda_0^{-1} = \Lambda_0^t$. An orthogonal transformation $\Lambda_0(x)$ defines the relationship between the base vectors of the initial body-attached frame $G_{0,i}$; $i = 1, 2, 3$ and the base vectors of the inertial frame $E_i$; $i = 1, 2, 3$ via $G_{0,i}(x) = \Lambda_0(x)E_i$; $i = 1, 2, 3$. The matrix of components of the initial orthogonal transformation $\Lambda_0$ will be referred to as the *initial rotation matrix*. Obviously, the initial configuration of the beam at $x$ is completely defined by its initial position vector $r_0(x)$ and its initial rotation matrix $\Lambda_0(x)$.

For any time $t_n > 0$ the line of centroids of cross sections in the deformed state is a space curve (Fig. 1) defined by the position vector $r_n(x) = r_0(x) + u_n(x)$. The Bernoulli hypothesis of the plane cross sections remaining planar after deformation and retaining their shape and area is assumed to hold, so the position vector $R_n$ of an arbitrary material particle $(x, y, z)$ of the deformed beam may now be written as

$$R_n(x, y, z) = r(x) + \Lambda_n(x)\begin{Bmatrix} 0 \\ y \\ z \end{Bmatrix},$$

where the *rotation matrix* $\Lambda_n$, in the deformed state at time $t_n$, is also an element of the Lie group of proper orthogonal transformations $SO(3)$, which defines the relationship between the base vectors of the body-attached frame $G_{n,i}(x)$; $i = 1, 2, 3$ in the deformed state at time $t_n$ and the base vectors of the inertial frame $E_i$; $i = 1, 2, 3$ via $G_{n,i}(x) = \Lambda_n(x)E_i$; $i = 1, 2, 3$. Note that, in contrast to the initial configuration, the base vector $G_{n,1}(x)$ in the deformed configuration at time $t_n$ is in general not tangential to the deformed centroid axis, i.e. $G_{n,1}(x) \neq r_n'(x)$. The deformed configuration of the beam at time $t_n$ for any $x \in [0, L]$ is thus completely defined by: (i) the position vector of the deformed line of centroids $r_n(x)$; and (ii) the rotation matrix $\Lambda_n(x)$. When there is no danger of confusion, the time index to a deformed configuration will be dropped and the deformed configuration denoted by $(r, \Lambda)$.

The rotation matrix $\Lambda$ may be parametrised using the *rotational (pseudo)-vector* $\psi = \psi' E_i \Leftrightarrow \hat{\psi} = -\hat{\psi}^{\mathrm{t}} = -e_k^{ij}\psi^k E_i E_j^{\mathrm{t}}$ as [3,17]

$$\Lambda = \exp \hat{\psi}\Lambda_0 \equiv \left(I + \frac{\sin\psi}{\psi}\hat{\psi} + \frac{1-\cos\psi}{\psi^2}\hat{\psi}^2\right)\Lambda_0\,; \qquad \psi = \|\psi\|\,, \tag{2.1}$$

in which $I$ is a $3\times 3$ unit matrix, $\hat{\psi}$ is the skew-symmetric matrix associated with the rotational pseudovector $\psi$, $\psi^i$; $i = 1, 2, 3$ are the components of the rotational pseudovector $\psi$ with respect to the inertial frame, $e_{ij}^k$ is the permutational symbol with all the components being zero apart from $e_{12}^3 = e_{23}^1 = e_{31}^2 = -e_{13}^2 = -e_{32}^1 = -e_{21}^3 = 1$. From here onwards, repeated indices in a product will be regarded as summational over the dimension of a space if they form a subscript-superscript pair. Note that the identity $v \times w = \hat{v}w = -\hat{w}v \quad \forall\, v, w \in \mathcal{R}^3$ can be established, where here, and throughout the paper, the superimposed hat denotes the skew-symmetric matrix of the respective vector quantity. We also note the identity $\Lambda v = \Lambda\hat{v}\Lambda^{\mathrm{t}}$ for any $v \in \mathcal{R}^3$ and any $\Lambda \in SO(3)$.

The rotation matrices at configurations $n$ and $n+1$ (see Fig. 1) are related via the rotational pseudovector $\theta$ which rotates the base vectors $G_{n,i}$ into the base vectors $G_{n+1,i}$ around the axis $\theta/\theta$ for the angle $\theta$ through

$$\Lambda_{n+1} = \exp \hat{\theta}\Lambda_n\,. \tag{2.2}$$

During the linearisation of the nonlinear equations of motion, we will need the variation of a current configuration $(r, \Lambda)$. Variation of the position vector is computed from $r(x) = r_0(x) + u(x)$ as $\Delta r = \Delta u$, while the variation of the rotation matrix is defined through its directional derivative [4] in the direction of a superposed infinitesimal rotation $\Delta\vartheta$ as

$$\Delta\Lambda = \frac{\mathrm{d}}{\mathrm{d}\epsilon}\bigg|_{\epsilon=0} \exp(\epsilon\,\Delta\hat{\vartheta})\Lambda = \Delta\hat{\vartheta}\,\Lambda\,. \tag{2.3}$$

The infinitesimal rotation $\Delta\vartheta$ superposed onto $\Lambda$ is often referred to as the *spin variable*. Unless $\theta \mp 2k\pi$; $k \in \mathcal{N}$, the spin variable $\Delta\vartheta$ at a particular configuration $n+1$ may be related to the variation of the rotational pseudovector between the configurations $n$ and $n+1$, $\Delta\theta$, as

$$\Delta\theta = T(\theta)\,\Delta\vartheta\,, \qquad \Delta\vartheta = T^{-1}(\theta)\,\Delta\theta \tag{2.4}$$

with

$$\begin{aligned} T(\theta) &= \frac{1}{\theta^2}\theta\theta^{\mathrm{t}} + \frac{\frac{\theta}{2}}{\tan\frac{\theta}{2}}\left(I - \frac{1}{\theta^2}\theta\theta^{\mathrm{t}}\right) - \frac{1}{2}\hat{\theta} \\ T^{-1}(\theta) &= \frac{1}{\theta^2}\left(1 - \frac{\sin\theta}{\theta}\right)\theta\theta^{\mathrm{t}} + \frac{\sin\theta}{\theta}I + \frac{1-\cos\theta}{\theta^2}\hat{\theta}\,, \end{aligned} \tag{2.5}$$

where $\theta = \|\theta\|$. See [5,6,8,14] for the derivation of Eqs. (2.4) and (2.5).

### 2.2. Numerical time integration

There are many different schemes for the numerical integration of velocities and accelerations between times $t_n$ and $t_{n+1}$, and here we limit our attention to those from the Newmark's two-parameter family of algorithms, in which the velocities and accelerations at time $t_{n+1}$ are defined using the displacements, velocities and accelerations at time $t_n$, the displacements at time $t_{n+1}$ and the parameters $\beta$ and $\gamma$ as [18]

$$\dot{u}_{n+1} = \frac{\gamma}{\beta\,\Delta t}(u_{n+1} - u_n) + \left(1 - \frac{\gamma}{\beta}\right)\dot{u}_n + \Delta t\left(1 - \frac{\gamma}{2\beta}\right)\ddot{u}_n \tag{2.6}$$

$$\ddot{u}_{n+1} = \frac{1}{\beta\,\Delta t^2}\left[u_{n+1} - u_n - \Delta t\,\dot{u}_n - \Delta t^2\left(\frac{1}{2} - \beta\right)\ddot{u}_n\right], \tag{2.7}$$

where $\dot{u}_n$, $\ddot{u}_n$, $\dot{u}_{n+1}$ and $\ddot{u}_{n+1}$ are the velocities and accelerations at times $t_n$ and $t_{n+1}$, the superimposed dots denote the derivatives with respect to time and $\Delta t = t_{n+1} - t_n$. As pointed out in [20], the application of such an

integration to angular velocities and accelerations only makes sense if performed in a body-attached frame. In this way we obtain

$$W_{n+1} = \frac{\gamma}{\beta\,\Delta t}\,\Theta + \left(1 - \frac{\gamma}{\beta}\right)W_n + \Delta t\left(1 - \frac{\gamma}{2\beta}\right)A_n \tag{2.8}$$

$$A_{n+1} = \frac{1}{\beta\,\Delta t^2}\left[\Theta - \Delta t\,W_n - \Delta t^2\left(\frac{1}{2} - \beta\right)A_n\right], \tag{2.9}$$

where $W_n$, $A_n$, $W_{n+1}$ and $A_{n+1}$ are angular velocities and accelerations at times $t_n$ and $t_{n+1}$, the components of which are given in body-attached frame $G_{n,i}$ (or $G_{n+1,i}$) and $\Theta$ is the rotational pseudovector between configurations at times $t_n$ and $t_{n+1}$ with components in either of the body attached frames. The latter is related to the incremental rotational pseudovector expressed in the inertial frame through $\Theta = \Lambda_n^{\mathrm{t}}\theta = \Lambda_{n+1}^{\mathrm{t}}\theta$.

At any time $t_k$ the angular velocities and accelerations $W_k$ and $A_k$, with the components given in the body-attached frame, are related to the angular velocities and accelerations $\dot{\vartheta}_k$ and $\ddot{\vartheta}_k$ with the components given in the inertial frame through $W_k = \Lambda_k^{\mathrm{t}}\dot{\vartheta}_k$ and $A_k = \Lambda_k^{\mathrm{t}}\ddot{\vartheta}_k$.

*NOTE 1.* It must be noted that neither the infinitesimal rotation $\Delta\vartheta$ superposed onto $\Lambda$ nor the angular velocity $\dot{\vartheta}$ imply the existence of any '$\vartheta$' (see also Reissner [2, p. 735]). The angular velocity $\dot{\vartheta}$ and the first time derivative of the incremental rotation $\dot{\theta}$ are related via $\dot{\vartheta} = T^{-1}(\theta)\dot{\theta}$, with $T^{-1}(\theta)$ as given in Eq. $(2.5)_2$.

The variations of $\ddot{u}_{n+1}$, $W_{n+1}$ and $A_{n+1}$ will be needed later and can be computed from Eqs. (2.7)–(2.9), $(2.4)_1$ and $\Theta = \Lambda^{\mathrm{t}}\theta$ as

$$\Delta\ddot{u}_{n+1} = \frac{1}{\beta\,\Delta t^2}\,\Delta u_{n+1} \tag{2.10}$$

$$\begin{aligned}
\Delta W_{n+1} &= \frac{\gamma}{\beta\,\Delta t}\,\Delta\Theta = \frac{\gamma}{\beta\,\Delta t}\,\Lambda_n^{\mathrm{t}}\,\Delta\theta = \frac{\gamma}{\beta\,\Delta t}\,\Lambda_n^{\mathrm{t}}T(\theta)\,\Delta\vartheta \\
\Delta A_{n+1} &= \frac{1}{\beta\,\Delta t^2}\,\Delta\Theta = \frac{1}{\beta\,\Delta t^2}\,\Lambda_n^{\mathrm{t}}\,\Delta\theta = \frac{1}{\beta\,\Delta t^2}\,\Lambda_n^{\mathrm{t}}T(\theta)\,\Delta\vartheta\,.
\end{aligned} \tag{2.11}$$

*NOTE 2.* In contrast to the incremental displacements, where taking first and second time derivatives results in velocities and accelerations, the first and second time derivatives of the incremental rotation $\Theta$ *are not* angular velocities and accelerations. Cardona and Géradin [5] used Newmark's integration in a different way and substituted the angular velocities and accelerations $W$ and $A$ in Eqs. (2.8) and (2.9) with time derivatives of the incremental rotation, $\dot{\Theta}$ and $\ddot{\Theta}$. In this way the angular accelerations $A$, needed to compute inertial moments, are not immediately supplied by the time integration and they have to be computed from $A_{n+1} = \Lambda_{n+1}^{\mathrm{t}}\ddot{\vartheta}$ after taking the time derivative of $\dot{\vartheta}$ from Note 1. This requires computing the time derivative of $T^{-1}(\theta)$ for the definition of inertial moments and computing the variation of the time derivative of $T^{-1}(\theta)$ for the definition of inertia matrix.

### 2.3. Strain measures

In the Reissner–Simo beam theory, the translational and rotational strain measures $\gamma$ and $\kappa$, defined with respect to the body-attached frame at $x$ are related to the configuration $(r, \Lambda)$ through the equations given in Tables 1 and 2 of [4] as

$$\gamma = \Lambda^{\mathrm{t}}r' - \begin{Bmatrix}1\\0\\0\end{Bmatrix}, \qquad \hat{\kappa} = \Lambda^{\mathrm{t}}\Lambda'\,. \tag{2.12}$$

The variation of the strain–configuration relationships (2.12) is given in Table 3 of [4] and reads

$$\Delta\gamma = \Lambda^{\mathrm{t}}(\Delta u' + r' \times \Delta\vartheta), \qquad \Delta\kappa = \Lambda^{\mathrm{t}}\,\Delta\vartheta'\,. \tag{2.13}$$

A convenient expression for the vector of rotational strains at configuration $k+1$, which is related to the configuration $k$ and the rotational pseudovector $\theta$ between the two configurations, is [5,6,9]

$$\kappa_{k+1} = \kappa_k + \Lambda_k^{\mathrm{t}} T^{-\mathrm{t}}(\theta)\theta' \tag{2.14}$$

where $T^{-1}(\theta)$ is defined by Eq. $(2.5)_2$.

*NOTE 3.* It is a simple exercise [12] to show that the strain measures $\underline{\gamma}$ and $\underline{\kappa}$, related to a configuration $(\underline{r}, \underline{\Lambda})$, obtained by superposing a rigid motion $(r_R, \Lambda_R)$ onto configuration $(r, \Lambda)$ via $\underline{r} = \Lambda_R(r_R + r)$ and $\underline{\Lambda} = \Lambda_R \Lambda$ are equal to the strain measures $\gamma$ and $\kappa$, related to the configuration $(r, \Lambda)$. This guarantees the objectivity of the strain measures (2.12) provided by the geometrically exact 3D beam theory.

### 2.4. Constitutive equations

The constitutive law for linear elastic material is taken to be given by a linear relation between the stress and stress-couple resultants and the adopted strain measures as $N = C_N \gamma$ and $M = C_M \kappa$, where $C_N = \mathrm{diag}[EA \quad GA_2 \quad GA_3]$ and $C_M = \mathrm{diag}[GJ_t \quad EI_2 \quad EI_3]$ are constant constitutive matrices. Here, $E$ and $G$ denote the elastic and shear moduli of the material, $A$ is the cross-sectional area, $A_2$ and $A_3$ are the shear areas in the directions of principal axes of inertia of the cross section, $J_t$ is the torsional inertial moment of the cross section and $I_2$ and $I_3$ are the cross-sectional inertial moments about the principal axes of inertia of the cross section. Since the constitutive matrices $C_N$ and $C_M$ are constant, the variations of the stress and stress-couple resultants are simply

$$\Delta N = C_N \Delta\gamma\,, \qquad \Delta M = C_M \Delta\kappa\,. \tag{2.15}$$

### 2.5. Weak form of the equations of motion

The weak form of the differential equations of motion of a 3D beam is given by Eqs. (2.5) and (2.6) in [8] and for any $t_{n+1}$ reads

$$G_{n+1} \equiv G_{k,n+1} + G_{m,n+1} - G_{e,n+1} = 0\,, \tag{2.16}$$

where

$$G_{k,n+1} = \int_0^L (f'^{\mathrm{t}} \Lambda N - \varphi^{\mathrm{t}} \hat{r}' \Lambda N + \varphi'^{\mathrm{t}} \Lambda M)_{n+1}\, \mathrm{d}x \tag{2.17}$$

is the elastic part of the weak form, while

$$G_{m,n+1} = \int_0^L (f^{\mathrm{t}} A\rho\ddot{u} + \varphi^{\mathrm{t}} \dot{\pi})_{n+1}\, \mathrm{d}x \tag{2.18}$$

is the inertial part of the weak form and

$$G_{e,n+1} = \int_0^L (f^{\mathrm{t}} n + \varphi^{\mathrm{t}} m)_{n+1}\, \mathrm{d}x + (f_0^{\mathrm{t}} F_0 + \varphi_0^{\mathrm{t}} T_0 + f_L^{\mathrm{t}} F_L + \varphi_L^{\mathrm{t}} T_L)_{n+1}\,, \tag{2.19}$$

is the virtual work of the applied loads (note that the boundary term in Eq. (2.19) is not included in [8]). The functions $f \in \mathcal{R}^3$ and $\varphi \in \mathcal{R}^3$ in the weak form are the so-called *test functions*, which are completely arbitrary apart from being continuous and at least once differentiable over the domain of the problem, $[0, L]$, and identically equal to zero at all coordinates $x$ with prescribed kinematic (essential) boundary conditions. In Eq. (2.18) we have denoted the density of the material as $\rho$, while $\dot{\pi} = \Lambda(\hat{W} J_\rho W + J_\rho A)$ is the time derivative of the specific angular momentum $\pi = \Lambda J_\rho W$ with respect to the centroid at $x$ and $J_\rho = \rho\, \mathrm{diag}[I_2 + I_3 \quad I_2 \quad I_3]$ is the tensor of the mass moments of inertia. In Eq. (2.19) we use symbols $n$ and $m$ to denote specific force and moment loads and $F_0, T_0, F_L, T_L$ to denote concentrated forces and torques at $x = 0$ and $x = L$.

The weak form (2.16) leads to the nonlinear equations of motion. They can be solved using an iterative procedure, based on expanding the weak form around the configuration at time $t_{n+1}$ into a Taylor series and neglecting all the second- and higher-order variations, i.e.

$$G_{n+1} + \Delta G_{n+1} = 0\,. \tag{2.20}$$

The variation of the weak form, $\Delta G_{n+1}$, may be expressed as

$$\Delta G_{n+1} = \Delta G_{k,n+1} + \Delta G_{m,n+1} + \Delta G_{e,\Phi,n+1}\,, \tag{2.21}$$

where $\Delta G_{k,n+1}$, $\Delta G_{m,n+1}$ and $\Delta G_{e,\Phi,n+1}$ are the elastic and inertial parts of the variation of the weak form and the part due to the linearisation of the external loads and the adopted test functions. The elastic part of the variation of the weak form at $t_{n+1}$ is obtained by varying $G_{k,n+1}$ in Eq. (2.16) and using the variations of the configuration $\Delta \boldsymbol{r} = \Delta \boldsymbol{u}$ and (2.3), the variations of strain measures (2.13) and the variations of the stress and stress-couple resultants (2.15). In this way and by noting the cross-product operator nature of skew-symmetric matrices (see Section 2.1), we obtain (note that for the sake of clarity the index $n+1$ has been dropped for the rest of this section)

$$\Delta G_k = \int_0^L \begin{Bmatrix} \boldsymbol{f}' \\ \boldsymbol{\varphi} \\ \boldsymbol{\varphi}' \end{Bmatrix}^{\mathrm{t}} \begin{bmatrix} \boldsymbol{c}_N & \boldsymbol{c}_N\hat{\boldsymbol{r}}' - \widehat{\boldsymbol{\Lambda N}} & \boldsymbol{0} \\ -\hat{\boldsymbol{r}}'\boldsymbol{c}_N + \widehat{\boldsymbol{\Lambda N}} & -\hat{\boldsymbol{r}}'\boldsymbol{c}_N\hat{\boldsymbol{r}}' + \hat{\boldsymbol{r}}'\widehat{\boldsymbol{\Lambda N}} & \boldsymbol{0} \\ \boldsymbol{0} & -\boldsymbol{\Lambda M} & \boldsymbol{c}_M \end{bmatrix} \begin{Bmatrix} \Delta \boldsymbol{u}' \\ \Delta \boldsymbol{\vartheta} \\ \Delta \boldsymbol{\vartheta}' \end{Bmatrix} \mathrm{d}x\,, \tag{2.22}$$

with $\boldsymbol{c}_N = \boldsymbol{\Lambda}\boldsymbol{C}_N\boldsymbol{\Lambda}^{\mathrm{t}}$ and $\boldsymbol{c}_M = \boldsymbol{\Lambda}\boldsymbol{C}_M\boldsymbol{\Lambda}^{\mathrm{t}}$. The inertial part of the variation of the weak form at $t_{n+1}$ is obtained by varying $G_{m,n+1}$ in Eq. (2.16) and using the variations of the configuration (2.3) and the variations of velocities and accelerations (2.10) and (2.11) so that

$$\Delta G_m = \int_0^L \begin{Bmatrix} \boldsymbol{f} \\ \boldsymbol{\varphi} \end{Bmatrix}^{\mathrm{t}} \begin{bmatrix} \dfrac{1}{\beta\,\Delta t^2}\boldsymbol{A}\rho\boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{0} & -\hat{\bar{\boldsymbol{\pi}}} + \dfrac{1}{\beta\,\Delta t^2}\boldsymbol{\Lambda}\bar{\boldsymbol{J}}_\rho\boldsymbol{\Lambda}_n^{\mathrm{t}}\boldsymbol{T}(\boldsymbol{\theta}) \end{bmatrix} \begin{Bmatrix} \Delta \boldsymbol{u} \\ \Delta \boldsymbol{\vartheta} \end{Bmatrix} \mathrm{d}x\,, \tag{2.23}$$

with $\bar{\boldsymbol{J}}_\rho = \boldsymbol{J}_\rho + \gamma\,\Delta t(\hat{\boldsymbol{W}}\boldsymbol{J}_\rho - \widehat{\boldsymbol{J}_\rho\boldsymbol{W}})$.

The third part of the variation of the weak form in the right-hand side of Eq. (2.21) is zero for configuration-independent test functions $\boldsymbol{f}$ and $\boldsymbol{\varphi}$ and configuration-independent loads. In general, however, it is computed as

$$\Delta G_{e,\Phi} = \int_0^L (D\boldsymbol{\Phi}\cdot\Delta\boldsymbol{U})^{\mathrm{t}} \begin{Bmatrix} \boldsymbol{A}\rho\ddot{\boldsymbol{u}} \\ \boldsymbol{\Lambda N} \\ -\hat{\boldsymbol{r}}'\boldsymbol{\Lambda N} + \dot{\boldsymbol{\pi}} \\ \boldsymbol{\Lambda M} \end{Bmatrix} \mathrm{d}x - \Delta G_e\,, \tag{2.24}$$

where $\boldsymbol{\Phi}$ and $\Delta\boldsymbol{U}$ are shorthand notations for $\langle \boldsymbol{f} \;\; \boldsymbol{f}' \;\; \boldsymbol{\varphi} \;\; \boldsymbol{\varphi}' \rangle^{\mathrm{t}}$ and $\langle \Delta\boldsymbol{u} \;\; \Delta\boldsymbol{u}' \;\; \Delta\boldsymbol{\varphi} \;\; \Delta\boldsymbol{\varphi}' \rangle^{\mathrm{t}}$ and $D\boldsymbol{\Phi}\cdot\Delta\boldsymbol{U}$ is the directional derivative of $\boldsymbol{\Phi}$ along $\Delta\boldsymbol{U}$. Note that $\Delta G_e$ in the right-hand side of Eq. (2.24) includes both the variation of the applied loads at $t_{n+1}$ and the variation of the test functions $\boldsymbol{f}$ and $\boldsymbol{\varphi}$.

## 3. Finite element interpolation

In order to define the weak form of the problem (2.16)–(2.19) with the finite element method, we first need to interpolate the test functions $\boldsymbol{f}(x)$ and $\boldsymbol{\varphi}(x)$. In addition, we must interpolate the deformed geometry, defined by $\boldsymbol{r}(x)$ and $\boldsymbol{\Lambda}(x)$. However, the latter cannot be interpolated in a conventional manner. In order to use the complete Newton–Raphson process based on (2.16)–(2.24) we therefore require the following two steps

(i) Interpolation of the *test functions* $\boldsymbol{f}(x)$ and $\boldsymbol{\varphi}(x)$ and

(ii) Interpolation of the variations of the kinematics, $\Delta\boldsymbol{u}(x)$ and $\Delta\boldsymbol{\vartheta}(x)$

We will consider here step (ii) as leading to *trial functions* (or trial *solutions*) for these variations. Choosing the same interpolation for the test and the trial functions is in many aspects convenient, most notably because it often results in a symmetric tangent stiffness matrix. For that reason, such an interpolation, which leads to the standard Bubnov–Galerkin approach (e.g. [21,22]), is regularly applied at this stage of the analysis [4–5,6,8,9]. In this paper, we depart from such a practice. Our reason for doing so is that, in contrast to the test functions, the interpolation of which need only be governed by the convenience mentioned above, the trial functions must be interpolated carefully in order to maintain the objectivity and the path-independence of the discretised strain

measures. This is so because the variations of the strain measures (2.13) (and hence the strain measures themselves) depend on the trial functions $\Delta\boldsymbol{u}$ and $\Delta\boldsymbol{\vartheta}$.

Following the earlier arguments, we interpolate the test functions in the simplest possible manner (see Section 3.1) and focus on the interpolation of the trial functions, which will be specifically designed to preserve the objectivity and the path-independence of the strain measures. Obviously, the two interpolations may turn out to be different thus leading to a Petrov–Galerkin method [23], in which the tangent stiffness matrix of the problem is genuinely nonsymmetric [7]. Although it is not immediately obvious, this approach may still be more convenient than that of taking the adopted interpolation of the trial functions and applying it to the test functions. This will be explained in more detail later.

### 3.1. Interpolation of test functions

By turning the beam into an $N$-noded finite element we approximate the distribution of test functions $\boldsymbol{f}(x)$ and $\boldsymbol{\varphi}(x)$ along the element using the approximated test functions $\boldsymbol{f}^h(x)$ and $\boldsymbol{\varphi}^h(x)$ defined as

$$\begin{Bmatrix}\boldsymbol{f}(x)\\ \boldsymbol{\varphi}(x)\end{Bmatrix} \doteq \begin{Bmatrix}\boldsymbol{f}^h(x)\\ \boldsymbol{\varphi}^h(x)\end{Bmatrix} = I^i(x)\begin{bmatrix}\boldsymbol{I} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{I}\end{bmatrix}\begin{Bmatrix}\boldsymbol{f}_i\\ \boldsymbol{\varphi}_i\end{Bmatrix}, \tag{3.1}$$

where $\langle \boldsymbol{f}^h(x)\boldsymbol{\varphi}^h(x)\rangle^t \in \mathcal{T}^h = \mathcal{F}^h \times \mathcal{P}^h \subset \mathcal{R}^3 \times \mathcal{R}^3$ and $\mathcal{T}^h$ is a *space of approximated test functions*. As mentioned in Section 2.5, $\langle \boldsymbol{f}(x)\boldsymbol{\varphi}(x)\rangle^t \in \mathcal{R}^3 \times \mathcal{R}^3$. Note that $\mathcal{T}^h$ is an $N$-dimensional vector space with

$$I^i(x)\begin{bmatrix}\boldsymbol{I} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{I}\end{bmatrix};\quad i = 1,\dots,N$$

as its base vectors, while $\mathcal{F}^h$ and $\mathcal{P}^h$ are $N$-dimensional vector spaces with $I^i(x)\boldsymbol{I}$; $i = 1,\dots,N$ as their base vectors. Also note that $\boldsymbol{f}_i = \boldsymbol{f}^h(x_i)$ and $\boldsymbol{\varphi}_i = \boldsymbol{\varphi}^h(x_i)$. The shape functions $I^i(x)$ are polynomials of degree $N-1$, which satisfy the standard conditions $I^i(x_j) = \delta^i_j$, $\Sigma_{i=1}^N I^i(x) = 1 \quad \forall x \in [0, L]$; $i, j = 1,\dots,N$, where $\delta^i_j = 1$ for $i = j$ and $\delta^i_j = 0$ otherwise. Note that the test functions interpolated in this way are *configuration-independent*. Such test functions considerably simplify the definition of the tangent stiffness matrix, because $\Delta G_{e,\Phi}$ in Eq. (2.24) is then dependent only on the applied loads. For most problems the latter are also configuration-independent, so $\Delta G_{e,\Phi}$ vanishes and $\Delta G$ as given by Eq. (2.21) is fully defined by $\Delta G_k$ and $\Delta G_m$ alone.

### 3.2. Interpolation of trial functions

As mentioned earlier, the trial functions should be interpolated carefully in order to preserve the invariance and the path-independence of the adopted strain measures. We understand the interpolated trial functions $\Delta\boldsymbol{u}^h(x)$ and $\Delta\boldsymbol{\vartheta}^h(x)$ as elements of a *space of trial solutions* of the problem $\mathcal{V}^h$ i.e. $\langle \Delta\boldsymbol{u}^h(x)\,\Delta\boldsymbol{\vartheta}^h(x)\rangle^t \in \mathcal{V}^h = \mathcal{U}^h \times \mathcal{S}^h \subset \mathcal{V}$, where $\mathcal{V} = \mathcal{R}^3 \times \mathcal{R}^3$ is the space of the chosen unknowns of the non-discretised problem, $\Delta\boldsymbol{u}(x)$ and $\Delta\boldsymbol{\vartheta}(x)$. The finite element approximations of the strain measures $\boldsymbol{\gamma}(x)$ and $\boldsymbol{\kappa}(x)$ and the configuration $\boldsymbol{r}(x)$, $\boldsymbol{\Lambda}(x)$, based on the finite element approximations of the chosen unknowns, will be denoted as $\boldsymbol{\gamma}^h(x)$, $\boldsymbol{\kappa}^h(x)$, $\boldsymbol{r}^h(x)$ and $\boldsymbol{\Lambda}^h(x)$. We want to know whether or not the objectivity of the strain measures $\boldsymbol{\gamma}$ and $\boldsymbol{\kappa}$, provided by the geometrically exact 3D beam theory (see earlier Note 3 and [12], Proposition 1) is inherited by any finite element discretisation. For that reason we state (see [12] for proof) the following

*PROPOSITION.* Approximated strain measures $\boldsymbol{\gamma}^h$ and $\boldsymbol{\kappa}^h$ associated with an approximated configuration $(\boldsymbol{r}^h, \boldsymbol{\Lambda}^h) \in \mathcal{R}^3 \times SO(3)$ are invariant to a constant (rigid-body) motion $(\boldsymbol{r}_R, \boldsymbol{\Lambda}_R) \in \mathcal{R}^3 \times SO(3)$ in the sense that $\underline{\boldsymbol{\gamma}}^h = \boldsymbol{\gamma}^h$ and $\underline{\boldsymbol{\kappa}}^h = \boldsymbol{\kappa}^h$ with

$$\boldsymbol{\gamma}^h = \boldsymbol{\Lambda}^{h\prime}\boldsymbol{r}^{h\prime} - \begin{Bmatrix}1\\0\\0\end{Bmatrix}, \qquad \underline{\boldsymbol{\gamma}}^h = \underline{\boldsymbol{\Lambda}}^{h\prime}\underline{\boldsymbol{r}}^{h\prime} - \begin{Bmatrix}1\\0\\0\end{Bmatrix}, \qquad \hat{\boldsymbol{\kappa}} = \boldsymbol{\Lambda}^{h\prime}\boldsymbol{\Lambda}^{h\prime} \text{ and } \underline{\hat{\boldsymbol{\kappa}}} = \underline{\boldsymbol{\Lambda}}^{h\prime}\underline{\boldsymbol{\Lambda}}^{h\prime},$$

where $\underline{\boldsymbol{\gamma}}^h$ and $\underline{\boldsymbol{\kappa}}^h$ are approximated strain measures associated with configuration $(\underline{\boldsymbol{r}}^h, \underline{\boldsymbol{\Lambda}}^h) \in \mathcal{R}^3 \times SO(3)$, if and only if this configuration is obtained by superposing the rigid-body motion $(\boldsymbol{r}_R, \boldsymbol{\Lambda}_R)$ onto the configuration $(\boldsymbol{r}^h, \boldsymbol{\Lambda}^h)$ via

$$\underline{\boldsymbol{r}}^h = \boldsymbol{\Lambda}_R(\boldsymbol{r}_R + \boldsymbol{r}^h) \tag{3.2}$$

$$\underline{\boldsymbol{\Lambda}}^h = \boldsymbol{\Lambda}_R \boldsymbol{\Lambda}^h , \tag{3.3}$$

for any constant $(\boldsymbol{r}_R, \boldsymbol{\Lambda}_R) \in \mathcal{R}^3 \times SO(3)$.

The strain-invariance condition (2.2) is satisfied by approximating the displacement trial function $\Delta\boldsymbol{u}(x) \in \mathcal{R}^3$ using $\Delta\boldsymbol{u}^h(x) \in \mathcal{U}^h \subset \mathcal{R}^3$, defined as

$$\Delta\boldsymbol{u}(x) \doteq \Delta\boldsymbol{u}^h(x) = I^i(x)\,\Delta\boldsymbol{u}_i \quad \text{and} \quad \Delta\boldsymbol{u}'(x) \doteq \Delta\boldsymbol{u}^{h'}(x) = I^{i'}(x)\,\Delta\boldsymbol{u}_i , \tag{3.4}$$

with $\Delta\boldsymbol{u}_i = \Delta\boldsymbol{u}^h(x_i)$; $i = 1, \ldots, N$ (see [12] for proof). This interpolation enables a nodal position vector update via

$$\boldsymbol{r}_i = \boldsymbol{r}_{0,i} + \boldsymbol{u}_{i,\text{old}} + \Delta\boldsymbol{u}_i ; \quad i = 1, \ldots, N . \tag{3.5}$$

which in turn enables us to obtain the position vector and its derivative at any coordinate $x \in [0, L]$ as

$$\boldsymbol{r}(x) \doteq \boldsymbol{r}^h(x) = I^i(x)\boldsymbol{r}_i \quad \text{and} \quad \boldsymbol{r}'(x) \doteq \boldsymbol{r}^{h'} = I^{i'}(x)\boldsymbol{r}_i . \tag{3.6}$$

The update of the current velocities and accelerations at the required integration points is performed using Eqs. (2.6) and (2.7).

Before proposing an interpolation for the rotational trial functions $\Delta\boldsymbol{\vartheta}(x)$, which will be invariant in the sense that it will satisfy the strain-invariance condition (3.3) and path-independent in the sense that the approximated strain measures will only be dependent on current approximated configuration, we note that some of the popular and widely used formulations are *path-dependent (relevant only to statics) and/or non-invariant.*

In the total formulation [5,6], the rotation matrix $\boldsymbol{\Lambda}(x)$ is parametrised using the total rotational pseudovector $\boldsymbol{\psi}(x)$ (see Eq. (2.1)). Its finite element approximation $\boldsymbol{\Lambda}^h(x)$ is given as [6]

$$\boldsymbol{\Lambda}(x) \doteq \boldsymbol{\Lambda}^h(x) = \exp \hat{\boldsymbol{\psi}}^h(x)\boldsymbol{\Lambda}_0 = \exp[I^i(x)\hat{\boldsymbol{\psi}}_i]\boldsymbol{\Lambda}_0 , \tag{3.7}$$

where $\boldsymbol{\psi}^h(x)$ is the finite element approximation of the total rotational pseudovector $\boldsymbol{\psi}(x)$ and $\boldsymbol{\psi}_i$; $i = 1, \ldots, N$ are the nodal values of the approximated total rotations. This interpolation fails to satisfy the invariance condition (3.3) and therefore produces a non-invariant finite element formulation [12]. However, the total formulation still provides a path-independent solution [12]. Numerical evidence for both of these claims will be given in Section 5.

In the incremental formulation [5,7], the rotation matrix $\boldsymbol{\Lambda}(x)$ is parametrised using the incremental rotational pseudovector $\boldsymbol{\theta}(x)$ (2.2). Its finite element approximation $\boldsymbol{\Lambda}^h(x)$ is given as [7]

$$\boldsymbol{\Lambda}_{n+1}(x) \doteq \boldsymbol{\Lambda}^h_{n+1}(x) = \exp \hat{\boldsymbol{\theta}}^h(x)\boldsymbol{\Lambda}^h_n(x) = \exp[I^i(x)\hat{\boldsymbol{\theta}}_i]\boldsymbol{\Lambda}^h_n(x) ,$$

where $\boldsymbol{\theta}^h(x)$ is the finite element approximation of the incremental rotational pseudovector $\boldsymbol{\theta}(x)$ and $\boldsymbol{\theta}_i$; $i = 1, \ldots, N$ are the nodal values of the approximated incremental rotations. The incremental formulation is also non-invariant, but in contrast to the total formulation it is also *path-dependent* (note the recursive character of the above equation) [12]. In the total and the incremental formulations, we may either choose the 'spin' variables $\Delta\boldsymbol{\vartheta}_i$ or the 'additive' infinitesimal rotations $\Delta\boldsymbol{\psi}_i$ (or $\Delta\boldsymbol{\theta}_i$) as the nodal unknowns [6,14]. The 'spin' and the 'additive' variables are related via matrix $\boldsymbol{T}(\boldsymbol{\psi}_i)$ (or $\boldsymbol{T}(\boldsymbol{\theta}_i)$) in the way, similar to that indicated in Note 1. For the sake of consistency with other formulations, we will run the numerical examples (see Section 5) using the spin variables, even though when using the total or the incremental formulation on their own it is often more convenient to work with additive infinitesimal rotational variables [5,6]. Various total and incremental formulations [5–7] may also differ to those outlined here in other aspects, but they all share the same definition for the rotational trial functions, which is the real reason for their non-invariance and/or path-dependence.

In the iterative formulation [4,8,9], the rotation matrix $\boldsymbol{\Lambda}(x)$ is parametrised using the iterative rotation (infinitesimal rotation superposed onto current configuration) i.e. using the rotational trial function $\Delta\boldsymbol{\vartheta}(x)$ itself (2.3), with its finite element approximation $\boldsymbol{\Lambda}^h(x)$ being [4,7–9]

$$\boldsymbol{\Lambda}_{n+1}(x) \doteq \boldsymbol{\Lambda}^h_{n+1}(x) = \exp \Delta\hat{\boldsymbol{\vartheta}}^h(x)\boldsymbol{\Lambda}^h_{n+1,\text{old}}(x) = \exp[I^i(x)\,\Delta\hat{\boldsymbol{\vartheta}}_i]\boldsymbol{\Lambda}^h_{n+1,\text{old}}(x) ,$$

where $\Delta\vartheta^h(x) \in \mathcal{S}^h \subset \mathcal{R}^3$ is the finite element approximation of the rotational trial function $\Delta\vartheta(x) \in \mathcal{R}^3$ and $\Delta\vartheta_i$; $i = 1, \ldots, N$ are the nodal values of iterative rotations. Any formulation based on the interpolation of iterative rotations is also non-invariant and path-dependent [12]. Numerical evidence will be given in Section 5.

In order to devise a strain-invariant and path-independent formulation it is essential to perform the following steps [12]:

(i) the decomposition of the rotation matrix $\Lambda(x)$ using (1) a reference rotation matrix $\Lambda_r$, which is unique for the whole beam and rigidly connected to it, and (2) the rotation matrix defining a local rotation $\Psi^l(x)$ between the reference rotation matrix $\Lambda_r$ and the rotation matrix $\Lambda(x)$, the components of which are given in the coordinate system defined by the reference rotation matrix and

(ii) standard (configuration-independent) interpolation of the local rotations $\Psi^l(x)$.

Following these guidelines, the rotation matrix $\Lambda(x)$ is decomposed and approximated as

$$\Lambda(x) \doteq \Lambda^h(x) = \Lambda_r \exp \hat{\Psi}^{lh}(x) . \tag{3.8}$$

The reference rotation matrix $\Lambda_r$ in Eq. (3.8) must be rigidly attached to a fixed combination of current rotations at two chosen nodes $I$ and $J$ [12]. Here, we limit our attention to a reference matrix which is either rigidly attached to a particular node $I = J$, or else is defined as a 'midway' rotation between the rotations at nodes $I$ and $J$. Both of these possibilities are covered by a definition of the reference matrix via [12]

$$\Lambda_r = \Lambda_I \exp\left(\frac{1}{2}\hat{\phi}_{IJ}\right) \tag{3.9}$$

$$\exp \hat{\phi}_{IJ} = \Lambda_I^t \Lambda_J \to \phi_{IJ} , \tag{3.10}$$

where the extraction of $\phi_{IJ}$ in Eq. (3.10) can be performed using Spurrier's algorithm [24] as given in [4,14] (see also Note 5 at the end of this section). The current local rotations at nodes $I$ and $J$ follow directly from Eqs. (3.8) and (3.9) as $\Psi_I^l = -\frac{1}{2}\phi_{IJ}$ and $\Psi_J^l = \frac{1}{2}\phi_{IJ}$, which implies $\Psi_I^l = -\Psi_J^l$. Note that for $I = J$ the rotation $\phi_{IJ}$ extracted from Eq. (3.10) is $\phi_{IJ} = \mathbf{0}$, which leads to both $\Lambda_r = \Lambda_I \equiv \Lambda_J$ and $\Psi_I^l \equiv \Psi_J^l = \mathbf{0}$.

As mentioned earlier, the standard interpolation has to be applied to the local rotations $\Psi^l(x)$, i.e.

$$\Psi^l(x) \doteq \Psi^{lh}(x) = I^i(x)\Psi_i^l , \tag{3.11}$$

where the nodal local rotations $\Psi_i^l$; $i = 1, \ldots, N$ are extracted from

$$\exp \hat{\Psi}_i^l = \Lambda_r^t \Lambda_i \to \Psi_i^l ; \quad i = 1, \ldots, N . \tag{3.12}$$

It is important to see that *no interpolation* is hidden in the history of the evolution of $\Lambda^h$ as given by Eqs. (3.8)–(3.10), (3.11) and (3.12). The formulation is therefore *path-independent*. A proof that the new formulation is strain-invariant is given in [12]. Numerical evidence will be given in Section 5.

The variation of Eq. $(3.8)_1$ can be computed by using Eq. (2.3) as $\Delta\hat{\vartheta}\,\Lambda = \Delta\hat{\vartheta}_r\,\Lambda_r \exp\hat{\Psi}^l + \Lambda_r\,\Delta(\exp\hat{\Psi}^l)$, which after postmultiplying by $\Lambda^t = (\exp\hat{\Psi}^l)^t\Lambda_r^t$ reduces to $\Delta\hat{\vartheta} = \Delta\hat{\vartheta}_r + \Lambda_r\,\Delta(\exp\hat{\Psi}^l)(\exp\hat{\Psi}^l)^t\Lambda_r^t$. By noting that $T^{-t}(\Psi^l)\,\Delta\Psi^l = \Delta(\exp\hat{\Psi}^l)(\exp\hat{\Psi}^l)^t$ the axial vector associated with the above skew-symmetric matrix is expressed as

$$\Delta\vartheta = \Delta\vartheta_r + \Lambda_r T^{-1}(\Psi^l)\,\Delta\Psi^l ,$$

where the matrix $T^{-1}$ is defined by Eq. $(2.5)_2$.

By inserting Eq. (3.11), we express the approximation of the rotational trial function $\Delta\vartheta(x)$ is

$$\Delta\vartheta(x) \doteq \Delta\vartheta^h(x) = \Delta\vartheta_r + \Lambda_r T^{-1}[\Psi^{lh}(x)]I^i(x)\,\Delta\Psi_i^l . \tag{3.13}$$

In order to express the approximation of the rotational trial function $\Delta\vartheta(x)$ in terms of its nodal values, we first note that $I^i(x_j) = \delta_j^i$; $i, j = 1, \ldots, N$ and substitute into Eq. (3.13) to obtain $\Delta\vartheta_i \equiv \Delta\vartheta^h(x_i) = \Delta\vartheta_r + \Lambda_r T^{-1}(\Psi_i^l)\,\Delta\Psi_i^l$, which can for any $\psi_i^l \neq \mp 2n\pi$; $n \in \mathcal{N}$ be inverted to give

$$\Delta\Psi_i^l = T(\Psi_i^l)\Lambda_r^t(\Delta\vartheta_i - \Delta\vartheta_r) , \tag{3.14}$$

where $T$ is given by Eq. $(2.5)_1$. It should be noted that for the given definition of the reference rotation matrix $\Lambda_r$ with respect to the nodal rotation matrices $\Lambda_I$ and $\Lambda_J$ (Eqs. (3.9) and (3.10)) the variation of the reference rotation matrix can also be expressed in terms of the variation of the rotation matrices at nodes $I$ and $J$. The derivation of $\Delta\vartheta_r$ in terms of $\Delta\vartheta_I$ and $\Delta\vartheta_J$ is fairly straightforward. Eq. (3.14) can be written for nodes $I$ and $J$ and the results related to each other by observing $\Psi_I^I = -\Psi_J^I$ so that $T(\Psi_I^I)\Lambda_r^t(\Delta\vartheta_I - \Delta\vartheta_r) = -T(\Psi_J^I)\Lambda_r^t(\Delta\vartheta_J - \Delta\vartheta_r)$.

By noting that $\Psi_I^I = -\Psi_J^I$ yields $T(\Psi_L^I) = T^t(\Psi_J^I)$, the above equation turns into

$$\Delta\vartheta_r = \Lambda_r[T(\Psi_I^I) + T^t(\Psi_I^I)]^{-1}T(\Psi_I^I)\Lambda_r^t\,\Delta\vartheta_I + \Lambda_r[T(\Psi_J^I) + T^t(\Psi_J^I)]^{-1}T(\Psi_J^I)\Lambda_r^t\,\Delta\vartheta_J\,.$$

This equation can be transformed by noting $(T+T^t)^{-1}T = [T^{-1}(T+T^t)]^{-1} = (I + T^{-1}T^t)^{-1}$ and $T^{-1}(\alpha)T^t(\alpha) = \exp\hat{\alpha}$ for any $\alpha \in \mathcal{R}^3$ [6], and by introducing the notation

$$v_I = (I + \exp\hat{\Psi}_I^I)^{-1} \equiv \left[I + \exp\left(-\frac{1}{2}\hat{\phi}_{IJ}\right)\right]^{-1} = \frac{1}{2}\left(I + \frac{1}{\phi_{IJ}}\tan\frac{\phi_{IJ}}{4}\hat{\phi}_{IJ}\right),$$

$$v_J = (I + \exp\hat{\Psi}_J^I)^{-1} \equiv \left[I + \exp\left(\frac{1}{2}\hat{\phi}_{IJ}\right)\right]^{-1} = \frac{1}{2}\left(I - \frac{1}{\phi_{IJ}}\tan\frac{\phi_{IJ}}{4}\hat{\phi}_{IJ}\right).$$

This process leads to the relationship

$$\Delta\vartheta_r = \Delta_k^{ij}(\delta_I^k + \delta_J^k)\Lambda_r v_j \Lambda_r^t\,\Delta\vartheta_i\,, \tag{3.15}$$

where $\Delta_k^{ij} = \delta_1^i\delta_1^j\delta_k^1 + \cdots + \delta_N^i\delta_N^J\delta_k^N$ can be regarded as a 'generalised' Kronecker symbol with $\Delta_k^{ij} = 1$ for $i = j = k$ and $\Delta_k^{ij} = 0$ otherwise. Finally, by inserting Eq. (3.15) into Eq. (3.14) we obtain

$$\Delta\Psi_i^I = T(\Psi_i^I)\Lambda_r^t[\Delta\vartheta_i - \Delta_k^{mj}(\delta_I^k + \delta_J^k)\Lambda_r v_j \Lambda_r^t\,\Delta\vartheta_m]\,, \tag{3.16}$$

so that the insertion of Eqs. (3.15) and (3.16) into Eq. (3.13) enables us to express $\Delta\vartheta^h(x)$ and $\Delta\vartheta^{h'}(x)$ in terms of $\Delta\vartheta_i$; $i = 1, \ldots, N$ via

$$\begin{aligned}\Delta\vartheta(x) &\doteq \Delta\vartheta^h(x) = \tilde{I}^i(x)\,\Delta\vartheta_i\\ \Delta\vartheta'(x) &\doteq \Delta\vartheta^{h'}(x) = \tilde{I}^{i'}(x)\,\Delta\vartheta_i\,,\end{aligned} \tag{3.17}$$

where the *generalised shape functions* $\tilde{I}^i(x)$; $i = 1, \ldots, N$ and their derivatives are defined as

$$\tilde{I}^i(x) = \Delta_k^{ij}\Lambda_r\{(\delta_I^k + \delta_J^k)[I - T^{-1}(\Psi^{Ih}(x))\tilde{T}]v_j + T^{-1}(\Psi^{Ih}(x))I^k(x)T(\Psi_j^I)\}\Lambda_r^t \tag{3.18}$$

$$\begin{aligned}\tilde{I}^{i'}(x) = \Delta_k^{ij}\Lambda_r\Big\{&-(\delta_I^k + \delta_J^k)\left[\frac{\mathrm{d}}{\mathrm{d}x}T^{-1}(\Psi^{Ih}(x))\tilde{T} + T^{-1}(\Psi^{Ih}(x))\tilde{T}'\right]v_j\\ &+\left[\frac{\mathrm{d}}{\mathrm{d}x}T^{-1}(\Psi^{Ih}(x))I^k(x) + T^{-1}(\Psi^{Ih}(x))I^{k'}(x)\right]T(\Psi_j^I)\Big\}\Lambda_r^t\end{aligned} \tag{3.19}$$

where $\tilde{T}$ and $\tilde{T}'$ are the shorthand notations for $\tilde{T} = I^i(x)T(\Psi_i^I)$ and $\tilde{T}' = I^{i'}(x)T(\Psi_i^I)$. Note that the generalised shape functions (3.18) and their derivatives (3.19) satisfy the conditions $\tilde{I}^i(x_j) = \delta_j^i I$, $\Sigma_{i=1}^N \tilde{I}^i(x) = I$ and $\Sigma_{i=1}^N \tilde{I}^{i'}(x) = 0$; $x \in [0, L]$. By taking the derivative of the matrix $T^{-1}$ (see Eq. $(2.5)_2$) we obtain (see [7] and [14] for the derivation)

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}x}T^{-1}(\Psi^I) = {}&\Psi^{I^t}\Psi^{I'}\frac{\Psi^I\sin\Psi^I - 2(1-\cos\Psi^I)}{\Psi^{I^4}}\hat{\Psi}^I + \frac{1-\cos\Psi^I}{\Psi^{I^2}}\hat{\Psi}^{I'}\\ &+\frac{1}{\Psi^{I^2}}\left(1 - \frac{\sin\Psi^I}{\Psi^I}\right)(\hat{\Psi}^I\hat{\Psi}^{I'} + \hat{\Psi}^{I'}\hat{\Psi}^I) + \Psi^{I^t}\Psi^{I'}\frac{3\sin\Psi^I - \Psi^I(2+\cos\Psi^I)}{\Psi^{I^5}}\hat{\Psi}^{I^2}\,.\end{aligned}$$

For $\Psi^I \to 0$ the above equation reduces to $\mathrm{d}/\mathrm{d}x\,T^{-1}(\Psi^I = 0) = \frac{1}{2}\hat{\Psi}^{I'}$.

*NOTE 4.* The total formulation, whereby $\Lambda(x) \doteq \Lambda(x) = \exp\hat{\psi}^h(x)\Lambda_0$, can be regarded as a special (non-invariant) case of the present formulation. By ignoring the requirement (i) and by defining the reference rotation matrix as $\Lambda_r = \Lambda_0$, it follows that $\Delta\vartheta_r \equiv \Delta\vartheta_0 = 0$, so that Eqs. (3.13) and (3.14) give

$$\Delta\vartheta(x) \doteq \Delta\vartheta^h(x) = \tilde{I}^i(x)\,\Delta\vartheta_i\,; \qquad \tilde{I}^i = \Delta^{ij}_k I^k \Lambda_0 T^{-1}(\Psi^{lh}) T(\Psi^l_j)\Lambda^t_0\,.$$

By comparing Eqs. (3.8) and (3.7) it follows that $\exp\hat{\Psi}^{lh} = \Lambda^t_0 \exp\psi^h \Lambda_0 \equiv \exp\widehat{\Lambda^t_0\psi^h}$ and hence $\Psi^{lh} = A^t_0\psi^h$ and $\tilde{I}^i = \Delta^{ij}_k I^k T^{-1}(\psi^h)T(\psi_j)$.

*NOTE 5.* Spurrier's algorithm [24] as given in [4,14] is used to relate a rotation matrix $\Lambda(\alpha) = \exp\hat{\alpha}$ to a quaternion $\tilde{q} = q_0 + q$, where $q_0$ and $q$ are the scalar and the vector parts of the quaternion, in order to optimise the computational procedure. Here, we also use it to *extract the rotational pseudovector* $\alpha$. It is important to appreciate that while a rotation matrix is associated both with a positive and a negative quaternion (see Eqs (B.4) and (B.5) in [4]), the latter do not necessarily lead to the extraction of the same rotational pseudovector. The rotation matrix $\Lambda(\alpha)$ is an element of the nonlinear differential manifold $SO(3)$, which can be represented by (i.e. is topologically equivalent to) an open sphere $\alpha_1^2 + \alpha_2^2 + \alpha_3^2 < \pi^2 \subset \mathcal{R}^3$ with identified pairs of mutually diametral points on the surface, so that no $\alpha$ with a norm larger than $\pi$ can be *uniquely* extracted from $\Lambda(\alpha)$. A unique extraction of the rotational pseudovector of a magnitude less than $\pi$ is maintained by choosing the quaternion with $q_0 \geqslant 0$ (see Sections 3.5.1 and 3.5.2 in [7]). Round-off errors in the extraction are minimised by applying $\alpha = 2q/\|q\| \arcsin\|q\|$ for $\|q\| < q_0$ and $\alpha = 2q/\|q\| \arccos q_0$ for $\|q\| > q_0$, while $\|q\| = 0 \Rightarrow \alpha = 0$.

The update of nodal rotation matrices is performed using $\Lambda_i = \exp\Delta\hat{\vartheta}_i\,\Lambda_{i,\text{old}}$ at all of the nodes, while the reference rotation matrix $\Lambda_r$ is updated in two stages. First, we compute the relative rotation between the chosen nodes $I$ and $J$ by extracting the rotational pseudovector $\phi_{IJ}$ from Eq. (3.10), after which the reference rotation matrix $\Lambda_r$ is updated from Eq. (3.9). The local nodal rotations are extracted from Eqs. (3.12) and enable the interpolation (3.11), which in turn provides the update of $\Lambda^h(x)$ (3.8) at any $x \in [0, L]$. This enables the update of translational strain measures and angular velocities and accelerations in the usual way. However, inserting Eq. (3.8) into Eq. $(2.12)_2$ shows that the reference (rigid-body) rotation matrix $\Lambda_r$ does not enter the expression for the rotational strain measures, so by bearing in mind links between Eqs. $(2.12)_2$ and (2.14), the update of the rotational strain measures is simply $\kappa(x) = T^{-1}[\Psi^l(x)]\Psi^{l'}(x)$. This means that two configurations, which do not differ in current local rotations but do differ in reference (rigid-body) rotations, have the same rotational strain measures. This is a different way of saying that the formulation is strain-invariant.

*NOTE 6.* The formulation presented in this section can easily be modified into an invariant, but *path-dependent* formulation based on a configuration-independent interpolation of *incremental* local rotations [12]. This formulation should only be applied to genuinely path-dependent problems, such as those in structural dynamics.

## 4. Finite element problem definition

### *4.1. Dynamic residual and equivalent stiffness matrix for a finite element*

By inserting the approximations for the test functions (3.1) into the weak form (2.16), we obtain the approximated weak form $G^h_{n+1}$ as

$$G^h_{n+1} = \langle f^t_i \varphi^t_i \rangle g^i_{n+1}\,, \tag{4.1}$$

where $g^i_{n+1}$ is the algorithmic contribution of all the forces acting on the beam at node $i$ of the beam finite element, and may be expressed as

$$g^i_{n+1} = q^i_{k,n+1} + q^i_{m,n+1} - q^i_{e,n+1}\,, \tag{4.2}$$

where the vectors of inertial, internal and external forces, $q^i_{k,n+1}$, $q^i_{m,n+1}$ and $q^i_{e,n+1}$, follow from Eqs. (2.16)–(2.19), (3.1), (4.1) and (4.2) as (the time index has been dropped)

$$q_k^i = \int_0^L \begin{bmatrix} I^{i'}I & 0 \\ -I^i\hat{r}' & I^{i'}I \end{bmatrix} \begin{Bmatrix} \Lambda N \\ \Lambda M \end{Bmatrix} \mathrm{d}x\,, \tag{4.3}$$

$$q_m^i = \int_0^L \begin{Bmatrix} I^i A\rho\ddot{u} \\ I^i\Lambda(\hat{W}J_\rho W + J_\rho A) \end{Bmatrix} \mathrm{d}x \tag{4.4}$$

$$q_e^i = \int_0^L \begin{Bmatrix} I^i n \\ I^i m \end{Bmatrix} \mathrm{d}x + \delta_1^i \begin{Bmatrix} F_0 \\ T_0 \end{Bmatrix} + \delta_N^i \begin{Bmatrix} F_L \\ T_L \end{Bmatrix}. \tag{4.5}$$

By inserting the approximations for the test functions (3.1) and the approximations for the trial functions (3.4) and (3.17) into the variation of the weak form (2.21) and assuming configuration-independent loads we obtain the variation of the approximated weak form $\Delta G_{n+1}^h$ as

$$\Delta G_{n+1}^h = \langle f_i^t \varphi_i^t \rangle (K_{k,n+1}^{ij} + K_{m,n+1}^{ij}) \begin{Bmatrix} \Delta u_j \\ \Delta\vartheta_j \end{Bmatrix}, \tag{4.6}$$

where $K_{k,n+1}^{ij}$ the stiffness matrix (obtained by the linearisation of internal forces) and $K_{m,n+1}^{ij}$ is the inertia matrix (due to the linearisation of inertial forces) at time $t_{n+1}$ relating the nodal trial solutions at node $j$ to the nodal forces at node $i$. These matrices are fully defined by Eqs. (2.21), (2.22), (2.23), (3.1), (3.4) and (3.17) as

$$K_k^{ij} = \int_0^L \begin{bmatrix} I^{i'}I^{j'}c_N & I^{i'}(c_N\hat{r}' - \widehat{\Lambda N})\tilde{I}^j \\ I^jI^{j'}(\widehat{\Lambda N} - \hat{r}'c_N) & I^{i'}c_M\tilde{I}^{j'} - I^{i'}\widehat{\Lambda M}\tilde{I}^j + I^i\hat{r}'(\Lambda N - c_N\hat{r}')\tilde{I}^j \end{bmatrix} \mathrm{d}x \tag{4.7}$$

$$K_m^{ij} = \int_0^L \begin{bmatrix} \dfrac{1}{\beta\,\Delta t^2} I^iI^jA\rho I & 0 \\ 0 & I^i\left(-\hat{\pi} + \dfrac{1}{\beta\,\Delta t^2}\Lambda\bar{J}_\rho\Lambda_n^tT(\theta)\right)\tilde{I}^j \end{bmatrix} \mathrm{d}x\,. \tag{4.8}$$

In order to alleviate locking problems, the vector of internal forces (4.3) and the tangent stiffness matrix (4.7) are numerically integrated using a reduced Gaussian quadrature rule with the number of integration points equal to the degree of polynomials used for the test functions. In contrast, full quadrature (with the number of integration points being equal to the number of nodes) is used for the integration of the vectors of inertial and external forces (4.4) and (4.5) and for the numerical integration of the inertia matrix (4.8).

The solution procedure follows by forcing the approximated weak form (4.1) rather than the non-discretised weak form (2.16) to vanish ($G_{n+1}^h = 0$) and applying the Taylor expansion (2.20) to approximated weak form (4.1) and its variation (4.6). Since the nodal values $f_i$; $i = 1,\ldots,N$ and $\varphi_i$; $i = 1,\ldots,N$ of the interpolated test functions in Eqs. (4.1) and (4.6) are arbitrary, Eqs. (4.1), (4.6) and $G_{n+1}^h = 0$ yield the Newton–Raphson equation

$$K^{ij}\,\Delta p_j = -g_{n+1}^i\,; \quad i = 1,\ldots,N\,, \tag{4.9}$$

where $\Delta p_j = \langle \Delta u_j \quad \Delta\vartheta_j \rangle^t$ and $K^{ij} = K_k^{ij} + K_m^{ij}$ is the $ij$-submatrix of an *equivalent tangent stiffness matrix.*

*NOTE* 7. The present strain-invariant formulation may be embedded within a Bubnov–Galerkin approach by providing the configuration-dependent interpolation of approximated rotational test functions $\varphi^h(x)$ via $\varphi^h(x) = \tilde{I}^i(x)\varphi_i$ and $\varphi^{h'}(x) = \tilde{I}^{i'}(x)\varphi_i$ with $\tilde{I}^i(x)$ and $\tilde{I}^{i'}(x)$ being defined by Eqs. (3.18) and (3.19). The main advantage of a Bubnov–Galerkin approach is that, if consistently linearised, it leads to a more symmetric (although in systems with 3D rotations if spin variables are used not a fully symmetric [4,5,14,15]) stiffness matrix. An artificial symmetrisation of the stiffness matrix in a Bubnov–Galerkin approach often does not impair the quadratic rate of convergence and therefore enables the use of a symmetric system solver [25,26]. However, the inertia matrix is strongly non-symmetric due to the gyroscopic effect so this advantage is limited to static analyses. Furthermore, a tremendous amount of additional work is required in order to linearise the configuration-dependent test functions (2.24). A non-consistent Bubnov–Galerkin approach, in which the variation of configuration-dependent test functions would be neglected, would be simply obtained by replacing $I^i$ and $I^{i'}$ in the second row of Eqs. (4.3)–(4.5), (4.7) and (4.8) by $\tilde{I}^i$ and $\tilde{I}^{i'}$.

### 4.2. Update procedure

The update of the kinematics, using the iterative translations and rotations obtained by solving Eq. (4.9), is performed using the procedure described in Section 3.2. Note that the strain measures are updated at the integration points used for the computation of the internal force vector (4.3) and the tangent stiffness matrix (4.7), while the velocities and accelerations are updated at different integration points, which are used for the computation of the inertial and external force vectors (4.4) and (4.5) and the tangent inertia matrix (4.8).

## 5. Examples

The described algorithm, along with the conventional total, incremental and iterative algorithms, have been coded and mounted in the research version of the finite element system LUSAS [27]. In this section we use this code to run representative numerical simulations and to further assess our theoretical findings. A nonsymmetric frontal solver will be utilised. The reference rotation matrix $\boldsymbol{\Lambda}_r$ in the new formulation will be defined depending on the chosen interpolation. For elements with a linear interpolation, it will be associated with the 'midway' rotation between the nodes ($I = 1, J = 2$), for the elements with quadratic interpolation, it will be identified with the mid-node rotation matrix ($I = J = 2$) and for the elements with cubic interpolation, it will be associated with the midway rotation between the two internal nodes ($I = 2, J = 3$) (see Section 3.2 for details). Where appropriate, the examples will also be run using an alternative iterative formulation [9] (for which a full numerical integration is used unless otherwise stated and there is no interpolation for displacements) and the co-rotational formulation [14].

In the numerical examples we use either or both of the following two convergence criteria:

(i) the displacement norm, defined as the square root of the sum of the squares of the nodal iterative displacements $\Delta \boldsymbol{u}_i$ over all the nodes in the structure $M$ as percentage of the square root of the sum of the squares of the total nodal displacements $\boldsymbol{u}_i$ over all the nodes in the structure, must be less than a prescribed tolerance $\delta_u$, i.e.

$$100\left(\frac{\sum_{i=1}^{M} \Delta \boldsymbol{u}_i^{\mathrm{t}} \Delta \boldsymbol{u}_i}{\sum_{i=1}^{M} \boldsymbol{u}_i^{\mathrm{t}} \boldsymbol{u}_i}\right)^{1/2} < \delta_u , \tag{5.1}$$

(ii) the residual norm, defined as the square root of the sum of the squares of the nodal residual forces $\boldsymbol{g}_f^i$ (defined as the force part of the nodal residual vector (4.2), i.e. $\boldsymbol{g}^i = \langle \boldsymbol{g}_f^i \boldsymbol{g}_t^i \rangle^{\mathrm{t}}$) over all the nodes in the structure as percentage of the square root of the sum of the squares of the nodal forces and reactions $\boldsymbol{q}_{e,f}^i$ (defined as the force part of the nodal vector of applied loads (4.5), i.e. $\boldsymbol{q}_e^i = \langle \boldsymbol{q}_{e,f}^i \boldsymbol{q}_{e,t}^i \rangle^{\mathrm{t}}$) over all the nodes in the structure, must be less than a prescribed tolerance $\delta_f$, i.e.

$$100\left(\frac{\sum_{i=1}^{M} \boldsymbol{g}_f^{i^{\mathrm{t}}} \boldsymbol{g}_f^i}{\sum_{i=1}^{M} \boldsymbol{q}_{e,f}^{i^{\mathrm{t}}} \boldsymbol{q}_{e,f}^i}\right)^{1/2} < \delta_f . \tag{5.2}$$

Note that rotational degrees of freedom do not feature in the computation of the above norms.

### 5.1. Example 1: Single-element path-dependence test

In this simple example, a beam of length $L = 1$ with Young's modulus $E = 1.2 \times 10^8$ and Poisson's ratio $\nu = 0.3$ and the geometric properties of cross sections $A = A_2 = A_3 = 0.1$, $I_2 = I_3 = 8.3 \times 10^{-5}$ and $J_t = 1.6 \times 10^{-4}$ is subject to prescribed end-point rotations

$$\psi_1 = \begin{Bmatrix} 1.00 \\ -0.50 \\ 0.25 \end{Bmatrix} \quad \text{and} \quad \psi_2 = \begin{Bmatrix} -0.40 \\ 0.70 \\ 0.10 \end{Bmatrix}. \tag{5.3}$$

The displacements at the first node are fixed to zero and the beam is modelled using a single element with linear interpolation for displacements and formulation-dependent rotations. The equilibrium is defined as the state in which relation (5.1) is satisfied for $\delta_u = 10^{-5}$.

The example is run by first (i) applying the prescribed end-point rotations in a single increment. Due to the imposed boundary constraints, the problem is solved in only two iterations regardless of the formulation used (note that there are no force loads and force reactions in this example and that in the geometrically exact isoparametric formulation the rotational strains depend only on rotations and their derivatives). For the same reason, there is no difference in the results obtained using either the total, incremental or iterative formulations. The values for the rotational strains at the integration point, $\kappa^h(L/2)$ and the associated displacements of the second node, $u_2$ are given in Table 1 under the incrementation sequence (i).

In order to check the path-dependence properties of the analysed formulations, we next (ii) apply the prescribed end-point rotations in two increments. In order to simulate practical applications, where nodal rotations are independent degrees of freedom, we scale the prescribed end-point rotations in the two increments by different factors. Here, we apply $0.775\psi_1$ and $0.4\psi_2$ in the first increment and $0.225\psi_1$ and $0.6\psi_2$ in the second increment. The values for the rotational strains at the integration point, $\kappa^h(L/2)$, and the associated displacements of the second node are given in Table 1 under the incrementation sequence (ii).

Clearly, the results in the two runs are different for the incremental and the iterative formulations, i.e. these two formulations are *path-dependent*. The total and the new, invariant, formulations are *path-independent*, and so is the co-rotational formulation, which provided a key idea behind the new formulation.

### 5.2. Example 2: 3D bend as a practical path-dependence test

The example chosen here to investigate the implications of path-dependence on practical applications is the well-known 3D cantilever bend, which has been extensively analysed and documented by a number of researchers [4–6,10,11,15]. The bend is modelled using eight finite elements with a linear interpolation of the displacements and the formulation-dependent rotations. One-point integration is used for all the formulations apart from the locking-free iterative formulation [9], where two-point integration is employed.

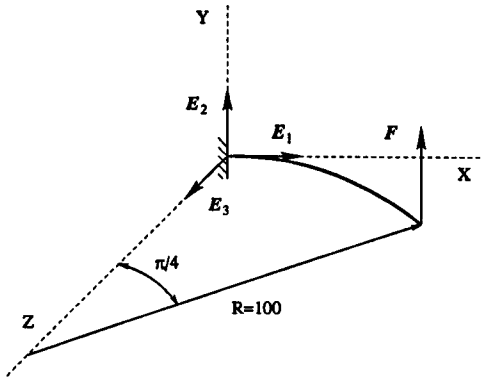
The bend is clamped at one end and is subject to a vertical force $\boldsymbol{F} = F\boldsymbol{E}_2$ of constant magnitude $F = 600$ at the other end. The geometry and the material and geometric properties of the structure are shown in Fig. 2. Equilibrium is taken to be established when the convergence criteria (5.1) and (5.2) are satisfied for $\delta_u = \delta_f = 10^{-7}$.

This problem is solved for the tip displacements $U$, $V$ and $W$ along the coordinate axes $X$, $Y$, $Z$ and is analysed for three different incrementation sequences:

Table 1
Components of the rotational strains $\kappa^h(L/2)$ and the displacements at second node $u_2$ as obtained by using different incrementation sequences: (i) single increment, (ii) two increments

| Formulation | $\kappa_1^h(L/2)$ | $\kappa_2^h(L/2)$ | $\kappa_3^h(L/2)$ | $u_{2,1}$ | $u_{2,2}$ | $u_{2,3}$ |
|---|---|---|---|---|---|---|
| Total (i) | −1.27464 | 1.26756 | −0.40350 | −0.02009 | 0.18605 | −0.07187 |
| Total (ii) | −1.27464 | 1.26756 | −0.40350 | −0.02009 | 0.18605 | −0.07187 |
| Incremental (i) | −1.27464 | 1.26756 | −0.40350 | −0.02009 | 0.18605 | −0.07187 |
| Incremental (ii) | −1.28872 | 1.25182 | −0.41280 | −0.01701 | 0.16401 | −0.08261 |
| Iterative (i) | −1.27464 | 1.26756 | −0.40350 | −0.02009 | 0.18605 | −0.07187 |
| Iterative (ii) | −1.28872 | 1.25182 | −0.41280 | −0.01701 | 0.16401 | −0.08261 |
| Invariant (i) | −1.26383 | 1.27102 | −0.42294 | −0.02408 | 0.20094 | −0.08490 |
| Invariant (ii) | −1.26383 | 1.27102 | −0.42294 | −0.02408 | 0.20094 | −0.08490 |
| Co-rotational [14] (i) | −1.09293 | 1.09861 | −0.36557 | −0.02398 | 0.20046 | −0.08490 |
| Co-rotational [14] (ii) | −1.09293 | 1.09861 | −0.36557 | −0.02398 | 0.20046 | −0.08490 |
| Iterative [9] (i)[a] | −1.27464 | 1.26756 | −0.40350 | | No converged solution | |
| Iterative [9] (ii) | | | No converged solution | | | |

[a] The given rotational strains are as obtained in the first iteration.




Stiffness properties: $EA = 10^7$

$GA_2 = GA_3 = 5 \cdot 10^6$

$GI_t = EI_2 = EI_3 = 10^7/12$

Finite element mesh: eight elements with linear interpolation

Concentrated load: $F = 600$

Fig. 2. 3D cantilever bend subject to a fixed vertical force at the free end.

(i) three equal increment sizes ($F/3 + F/3 + F/3$),
(ii) three different increment sizes ($F/2 + F/4 + F/4$) and
(iii) ten equal increment sizes ($10 \times F/10$).

The results for the components of the tip displacement are given in Table 2. We note that all of the results are very similar and also very close to the results given in [10,11].

We again observe that the incremental and iterative formulations are dependent on the history of incrementation, while the total, new and co-rotational formulations are not. However, the effect of path-dependence is, in this example, far smaller than in the previous one because of the finer finite-element mesh.

Table 2
Components of the tip displacement of the 3D cantilever bend (Fig. 2) obtained by using different incrementation sequences: (i) three equal increments, (ii) three different increments, (iii) ten equal increments

| Formulation | $U$ | $V$ | $W$ |
|---|---|---|---|
| Total (i) | −23.47905 | 53.37118 | −13.48336 |
| Total (ii) | −23.47905 | 53.37118 | −13.48336 |
| Total (iii) | −23.47905 | 53.37118 | −13.48336 |
| Incremental (i) | −23.47881 | 53.36998 | −13.48730 |
| Incremental (ii) | −23.47886 | 53.37010 | −13.48687 |
| Incremental (iii) | −23.47877 | 53.36984 | −13.48773 |
| Iterative (i) | −23.47801 | 53.36609 | −13.49232 |
| Iterative (ii) | −23.47872 | 53.36956 | −13.49115 |
| Iterative (iii) | −23.47876 | 53.36980 | −13.48785 |
| Invariant (i) | −23.47948 | 53.37149 | −13.48282 |
| Invariant (ii) | −23.47948 | 53.37149 | −13.48282 |
| Invariant (iii) | −23.47948 | 53.37149 | −13.48282 |
| Co-rotational [14] (i) | −23.73911 | 53.73249 | −13.67271 |
| Co-rotational [14] (ii) | −23.73911 | 53.73249 | −13.67271 |
| Co-rotational [14] (iii) | −23.73911 | 53.73249 | −13.67271 |
| Iterative [9] (i) | −23.49673 | 53.28004 | −13.48511 |
| Iterative [9] (ii) | −23.49675 | 53.27987 | −13.48533 |
| Iterative [9] (iii) | −23.49670 | 53.28026 | −13.48469 |

Indeed, as argued in [12], the path-dependence and non-invariance vanish if the number of nodes in the mesh tends to infinity. It should also be noted that, in contrast to the simple Example 5.1, the manifestation of the path dependence is different in the incremental and iterative formulations. While the difference between the incrementation sequences (i) and (ii) is in the incremental formulation entirely due to the different increment sizes, the (more pronounced) difference between the two sequences in the iterative formulation is additionally due to a different iterative path that leads to the solution. In particular, 13, 9 and 7 iterations are needed for the incrementation sequence (i) and 14, 17 and 9 iterations are needed for the incrementation sequence (ii). In the alternative iterative formulation [9], the results for the two incrementation sequences differ to a far lesser extent. This is very probably due to the more robust structure of the formulation, which provides a monotonic iteration towards the solution (four iterations are needed in each increment for both incrementation sequences; even if the load is applied in a single increment the problem is solved in four iterations).

### 5.3. Example 3: Single-element objectivity test

The objectivity of the different formulations is tested by superposing a constant rotation $\boldsymbol{\psi}_R^{\mathrm{t}} = \langle 0.2 \quad 1.2 \quad -0.5 \rangle^{\mathrm{t}}$ onto the prescribed rotations $\boldsymbol{\psi}_1$ and $\boldsymbol{\psi}_2$ given by Eq. (5.3) for the beam analysed in a single increment in Example 1. The components of the rotational strain $\boldsymbol{\kappa}^h(L/2)$ at the integration point for that problem are given in Table 1 under (i). For convenience, they are also reproduced in the left-hand part of Table 3. The new prescribed rotations at the beam ends are extracted from $\exp \hat{\underline{\boldsymbol{\psi}}}_1 = \exp \hat{\boldsymbol{\psi}}_R \exp \hat{\boldsymbol{\psi}}_1$ and $\exp \hat{\underline{\boldsymbol{\psi}}}_2 = \exp \hat{\boldsymbol{\psi}}_R \exp \hat{\boldsymbol{\psi}}_2$ by using, e.g. Spurrier's algorithm (see [14,24] and Note 5) as

$$\underline{\boldsymbol{\psi}}_1 = \begin{Bmatrix} 1.00145\,66233\,24399 \\ 0.34679\,74254\,22351 \\ -0.83717\,18210\,05534 \end{Bmatrix} \quad \text{and} \quad \underline{\boldsymbol{\psi}}_2 = \begin{Bmatrix} 0.08849\,14860\,02004 \\ 1.93320\,47713\,48018 \\ -0.08186\,60178\,89401 \end{Bmatrix}. \tag{5.4}$$

The deformed state of the beam associated with prescribed rotations (5.4) is characterised by the rotational strain $\underline{\boldsymbol{\kappa}}^h(L/2)$ given in the right-hand part of Table 3. The results in Table 3 show that the rotational strains associated with the two configurations, which only differ by a rigid-body rotation, are different for the total, the incremental and the iterative formulations. This illustrates the *non-invariance* of these formulations. It can be seen from Table 3 that the new formulation is indeed invariant and so is the co-rotational formulation.

The results of Examples 1 and 3, here obtained by running a finite element code, correspond with the hand-calculated results given in [12].

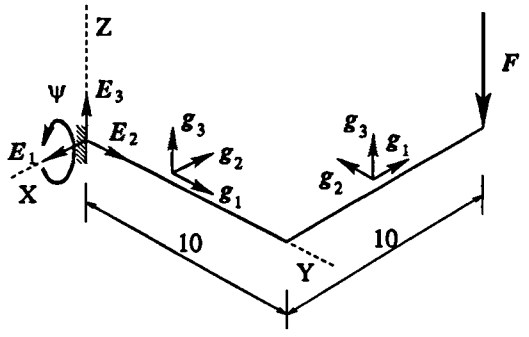
### 4.4. Example 4: Elbow cantilever subject to prescribed rotation and point load

In this example, we analyse a rectangular elbow cantilever subject to a concentrated vertical force at its free end and a prescribed rotation at its fixed end (Fig. 3). The frame is composed of two equally long mutually perpendicular legs, the orientation of which is defined by the orthonormal element frames $\boldsymbol{g} = [\boldsymbol{g}_1 \quad \boldsymbol{g}_2 \quad \boldsymbol{g}_3]$, where the unit base vector $\boldsymbol{g}_1$ is always directed along the centroidal axis and the base vectors $\boldsymbol{g}_2$ and $\boldsymbol{g}_3$ are directed along the principal axes of inertia. The stiffness properties of the legs and the initial orientation of each element frame with respect to the orthonormal global frame $\boldsymbol{E} = [\boldsymbol{E}_1 \quad \boldsymbol{E}_2 \quad \boldsymbol{E}_3]$ are given in Fig. 3.

Each leg is modelled by a single isoparametric element, with linear, quadratic or cubic interpolation for the displacements and formulation-dependent rotations. The applied force $\boldsymbol{F} = -F\boldsymbol{E}_3$ is of a fixed magnitude $F = 5$

Table 3
Components of the rotational strains $\boldsymbol{\kappa}_2^h(L/2)$ (Example 5.1 (i)) and $\underline{\boldsymbol{\kappa}}^h(L/2)$ (Example 5.3)

| Formulation | Example 5.1 (i) | | | Example 5.3 | | |
|---|---|---|---|---|---|---|
| | $\boldsymbol{\kappa}_1^h(L/2)$ | $\boldsymbol{\kappa}_2^h(L/2)$ | $\boldsymbol{\kappa}_3^h(L/2)$ | $\underline{\boldsymbol{\kappa}}_1^h(L/2)$ | $\underline{\boldsymbol{\kappa}}_2^h(L/2)$ | $\underline{\boldsymbol{\kappa}}_3^h(L/2)$ |
| Total | −1.27464 | 1.26756 | −0.40350 | −1.26399 | 1.31371 | −0.33751 |
| Incremental | −1.27464 | 1.26756 | −0.40350 | −1.26399 | 1.31371 | −0.33751 |
| Iterative | −1.27464 | 1.26756 | −0.40350 | −1.26399 | 1.31371 | −0.33751 |
| Invariant | −1.26383 | 1.27102 | −0.42294 | −1.26383 | 1.27102 | −0.42294 |
| Co-rotational [14] | −1.09293 | 1.09861 | −0.36557 | −1.09293 | 1.09861 | −0.36557 |
| Iterative [9] | −1.27464 | 1.26756 | −0.40350 | −1.26399 | 1.31371 | −0.33751 |




Stiffness properties: $EA = GA_2 = GA_3 = 10^6$

$GI_t = EI_2 = EI_3 = 10^3$

Finite element mesh: one isoparametric element per leg

Concentrated load: $F = 5$

Fig. 3. 3D cantilever frame subject to a fixed vertical force at the free end and a prescribed rotation around base vector $E_1$ at the fixed end.

and constant vertical direction. An applied prescribed rotation $\psi = \psi E_1$ rotates the whole structure around the unit vector $E_1$ of the global frame. This rotation does not change the direction of the applied load $F$.

Equilibrium is taken to be established when the convergence criteria (5.1) and (5.2) are satisfied for $\delta_u = \delta_f = 10^{-6}$. If convergence has not been achieved within 30 iterations, the load step is halved. This problem is solved for the displacements $U$, $V$ and $W$ along the coordinate axes $X$, $Y$, $Z$ of the global frame $E$ for different amounts of the prescribed rotation, $\psi$.

The problems of non-objectivity become obvious both by applying: (i) the force $F = 5$ in the first increment and a prescribed rotation $\psi = \pi/2$ in the second increment; and (ii) the force $F = 5$ and the prescribed rotation $\psi = \pi/2$ in a single increment. In the deformed configuration the displacement component $V$ of the tip is expected to be exactly $-10$. However, as shown in Table 4, it is only so if the problem is solved using the new formulation. The total, incremental and iterative formulations all produce wrong results with different amounts of error. The co-rotational formulation [14] cannot solve this problem and the code to run the alternative

Table 4
Tip displacement $V$ after applying (i) vertical tip force $F = 5$ in the first increment and prescribed rotation $\psi = \pi/2$ in the second increment and (ii) vertical tip force $F = 5$ and prescribed rotation $\psi = \pi/2$ in a single increment by using different element types

| Formulation | Interpolation | | |
|---|---|---|---|
| | Linear | Quadratic | Cubic |
| Total (i) | −9.89719 | −9.99977 | −9.99995 |
| Total (ii) | −9.89719 | −9.99977 | −9.99995 |
| Incremental (i) | −9.46626 | −9.98876 | −9.99974 |
| Incremental (ii) | −9.89719 | −9.99977 | −9.99995 |
| Iterative (i) | −9.44092 | −9.98218 | −9.99961 |
| Iterative (ii) | −9.60206 | −9.98500 | −9.99966 |
| Invariant (i) | −10.00000 | −10.00000 | −10.00000 |
| Invariant (ii) | −10.00000 | −10.00000 | −10.00000 |
| Co-rotational [14] (i) | No convergence | N/A | N/A |
| Co-rotational [14] (ii) | No convergence | N/A | N/A |
| Iterative [9] (i) | | No facilities present | |
| Iterative [9] (ii)* | −9.60691 | −9.54986 | −9.89848 |

* In order to achieve the results accurate to five decimal points it was necessary to apply 4-point, 8-point and 11-point quadrature rule for the respective interpolation.

Table 5
Tip displacement $V$ after applying vertical tip force $F = 5$ in the first increment and prescribed rotation $\psi = \pi/2$ in the second increment by using different numbers of elements with linear interpolation

| Formulation | Number of elements per leg | | | |
|---|---|---|---|---|
| | One | Two | Four | Eight |
| Total | −9.89719 | −9.97179 | −9.99276 | −9.99818 |
| Incremental | −9.46626 | −9.85135 | −9.96174 | −9.99037 |
| Iterative | −9.44092 | −9.84199 | −9.95804 | −9.98967 |
| Invariant | −10.00000 | −10.00000 | −10.00000 | −10.00000 |
| Co-rotational [14] | No convergence | | | |

iterative formulation [9] is not endowed with the facilities to apply prescribed rotations in subsequent increments. When the problem is run for incrementation sequence (ii), the incremental and iterative formulations produce results which are different from those obtained for the incrementation sequence (i) because of the path-dependence of these formulations (see Examples 5.1 and 5.2). Also, since the incrementation sequence (ii) is a single-increment one, the results of the total and the incremental formulations coincide.

The error in the non-invariant formulations reduces with $p$-refinement, i.e. by increasing the order of interpolation within elements. As shown in Table 5, the error also reduces with $h$-refinement, i.e. by increasing the number of elements in the mesh.

The errors due to the non-objectivity are prone to accumulation. If the example is run for: (i) $\psi = 0$; and (ii) $\psi = 2\pi$ with $\psi$ being applied in four equal increments and $F$ being kept constant throughout the incrementation, the results in the two runs should be the same, because these runs only differ in the rotation of one revolution. The results are collected in Table 6. As pointed out in [6] the total formulation is unable to handle total rotations larger than $2\pi$. The co-rotational formulation cannot cope with such incremental rotations and there are no facilities for running the alternative iterative formulation, so for these three formulations, only the results for the first run are available. For the other three formulations, a comparison between the results for the two runs in

Table 6
Tip displacement $W$ for (i) $\psi = 0$ and (ii) $\psi = 2\pi$ by using different element types

| Formulation | Interpolation | | |
|---|---|---|---|
| | Linear | Quadratic | Cubic |
| Total (i) | −6.18621 | −6.76756 | −6.76841 |
| Total (ii) | N/A | N/A | N/A |
| Incremental (i) | −6.18621 | −6.76756 | −6.76841 |
| Incremental (ii) | −7.15263 | −6.77303 | −6.76844 |
| Iterative (i) | −6.18296 | −6.76793 | −6.76840 |
| Iterative (ii) | −7.35661 | −6.78935 | −6.76860 |
| Invariant (i) | −6.18601 | −6.76754 | −6.76841 |
| Invariant (ii) | −6.18601 | −6.76754 | −6.76841 |
| Co-rotational [14] (i) | −7.04422 | N/A | N/A |
| Co-rotational [14] (ii) | No convergence | N/A | N/A |
| Iterative [9] (i) | −6.11815 | −6.76800 | −6.76847 |
| Iterative [9] (ii) | No facilities present | | |

Table 7
Tip displacement $W$ for $\psi = n\pi/2$. Results for the elements with quadratic interpolation

| $n$ | Incremental | Iterative | Invariant |
|---|---|---|---|
| 0 | −6.76756 | −6.76793 | −6.76754 |
| 4 | −6.77303 | −6.78935 | −6.76754 |
| 8 | −6.78386 | −6.81773 | −6.76754 |
| 12 | −6.79468 | −6.84694 | −6.76754 |
| 16 | −6.80549 | −6.87699 | −6.76754 |
| 20 | −6.81631 | −6.90787 | −6.76754 |

Table 6 clearly shows that the incremental and the iterative formulations introduce errors and that these errors are more severe for lower-order elements.

If the frame is left spinning as described above and the displacement components at the free end are monitored at the end of each revolution, it can be observed that the errors produced by the incremental and the iterative formulations are growing. In contrast, the new formulation, which has been designed to guarantee the objectivity of strain measures, preserves the displacement components at the end of each revolution. The results for the displacement component $W$ for the first five revolutions, for elements with quadratic interpolation, are given in Table 7.

The graphical comparison of the results for the incremental and the invariant formulations for the first 200 revolutions are presented in Fig. 4, while the graphical comparison of the results for the iterative and the invariant formulations for the first 200 revolutions are presented in Fig. 5. Clearly, a coarse mesh of lower-order elements is exceedingly prone to error accumulation due to the non-objectivity of the strain measures (see Figs. 4(a) and 5(a)). As the mesh is $p$-refined, this error reduces so that for the elements with cubic interpolation it is barely discernible (see Figs. 4(c) and 5(c)).

This phenomenon can also be observed by monitoring the strain energy at the end of each revolution. These values are expected to be constant, but the incremental and iterative formulations are unable to provide such results. In contrast, the new formulation retains the value of the strain energy regardless of the number of revolutions applied (see Fig. 6). Again, the error in the two non-invariant formulations is reduced by applying a higher-order interpolation.

It is interesting to note that the incremental and iterative formulations not only fail to provide objective strain measures, but are also less robust in coping with large increments, once the error accumulates more considerably. On several occasions, the solution procedure cannot achieve equilibrium for a given prescribed incremental rotation and the problem has to be overcome by halving the size of the incremental rotation. This happens in all the runs, using the incremental and the iterative formulations, and for all element types. Fig. 7 shows the evolution of the applied incremental rotation against the total applied prescribed rotation for the incremental formulation for different element types. Corresponding results for the iterative formulation are given in Fig. 8. In contrast, no reduction of the applied incremental rotation is ever needed for the new formulation. The total number of increments needed to solve the problem, using the different formulations and element types, is given in Table 8.

It is worth commenting on the strange kink between the applied prescribed rotations $\psi = 596.9026$ and $\psi = 603.18579$ (revolutions 95 and 96) as shown in Figs. 5(a) and 6(a) for the iterative formulation using elements with a linear interpolation. As mentioned earlier, this formulation is non-invariant and path-dependent with respect to the *iterative path* toward a converged state. This means that if the iterative path leans towards a divergence, but convergence is eventually achieved, this will have an effect on the solution. Indeed, by analysing the evolution of the number of iterations needed to achieve equilibrium against the increment number, plotted in Fig. 9, it can be observed that the convergence at increment 381 (note that increments 380 and 384 correspond to completed revolutions 95 and 96) is achieved after as many as 28 iterations, which is roughly five times the average number of iterations per increment. Fig. 8(a) shows that very soon after that stage, the increment size has to be halved, as the incremental rotations $\pi/2$ become increasingly difficult to handle (note that the procedure is set up to cut the increment size in half if the convergence has not been achieved within 30 iterations). The iterative procedure at increment 381 is obviously struggling to find the bowl of quadratic convergence and, in so doing, the large iterative rotations are encountered, which strongly influences the path-dependent solution.

Table 8
Total number of increments needed to solve the problem with $F = 5$ and $\psi = 400\pi$ (initially applied in multiples of $\pi/2$)

| Formulation | Interpolation | | |
|---|---|---|---|
| | Linear | Quadratic | Cubic |
| Incremental | 2448 | 14080 | 1440 |
| Iterative | 1280 | 25856 | 944 |
| Invariant | 800 | 800 | 800 |

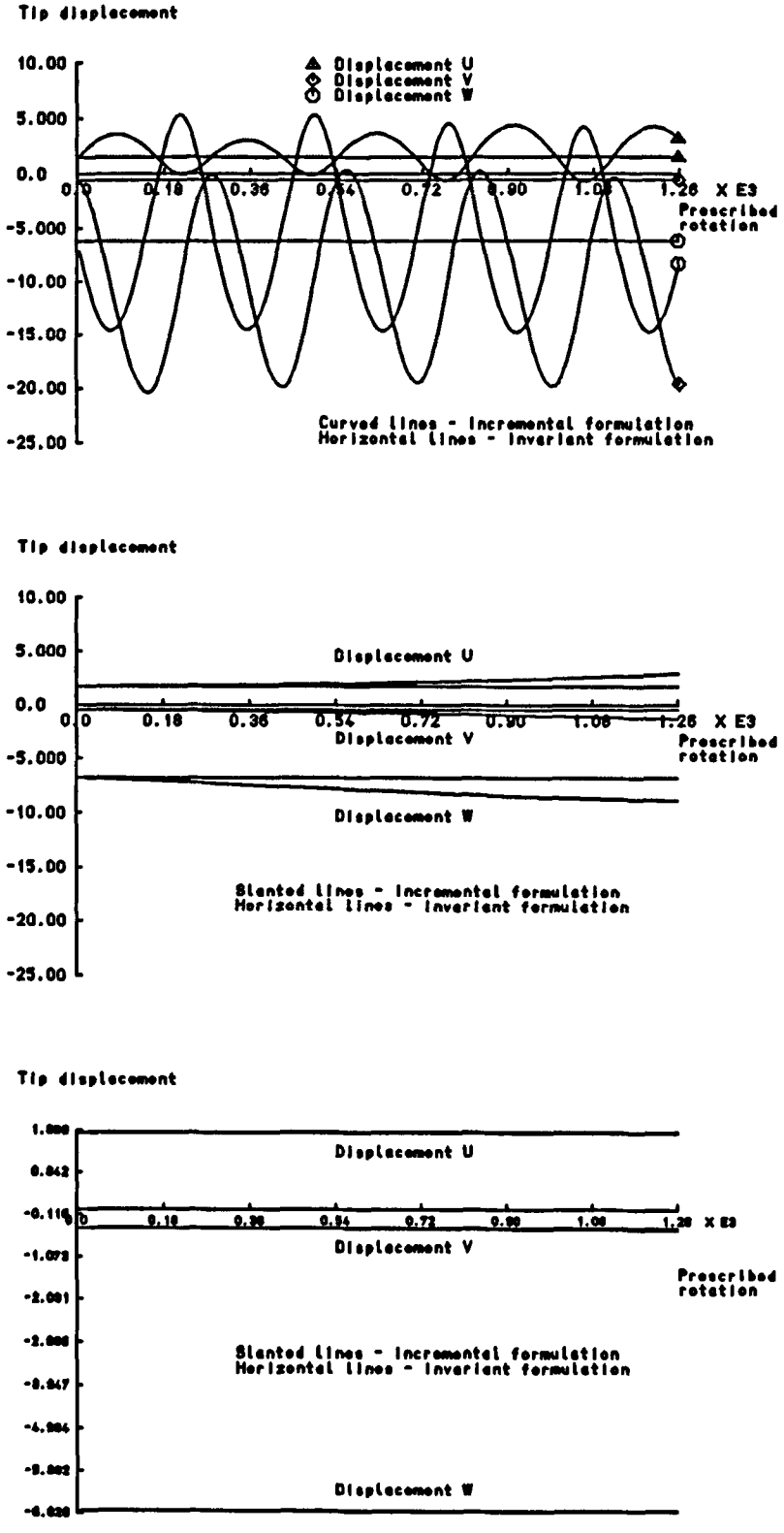



Fig. 4. Displacement components at the free end of the cantilever for first 200 revolutions for the incremental and invariant formulations using the elements with: (a) linear interpolation; (b) quadratic interpolation; (c) cubic interpolation.

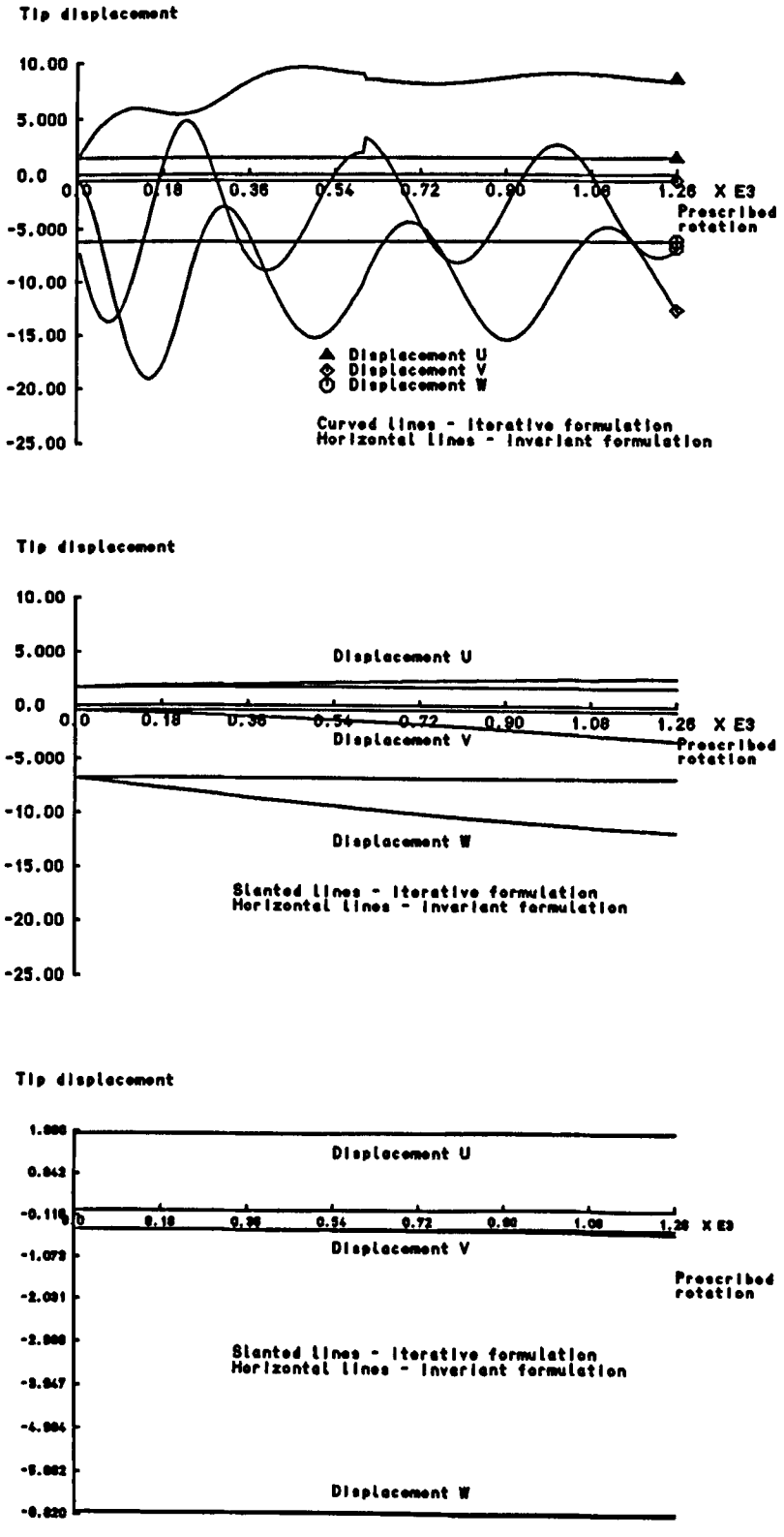
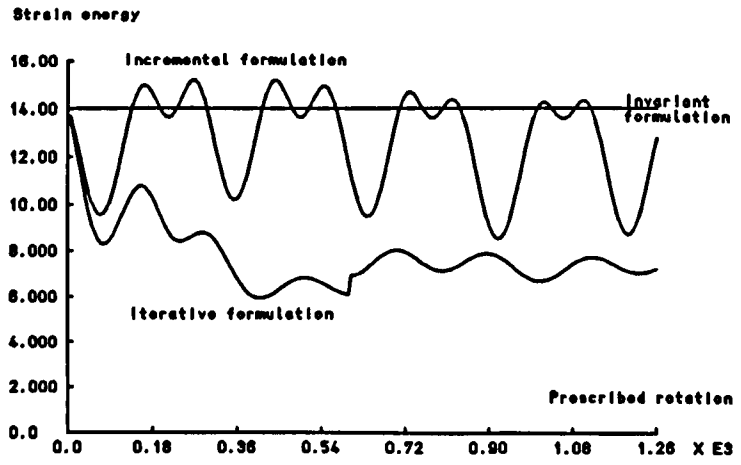
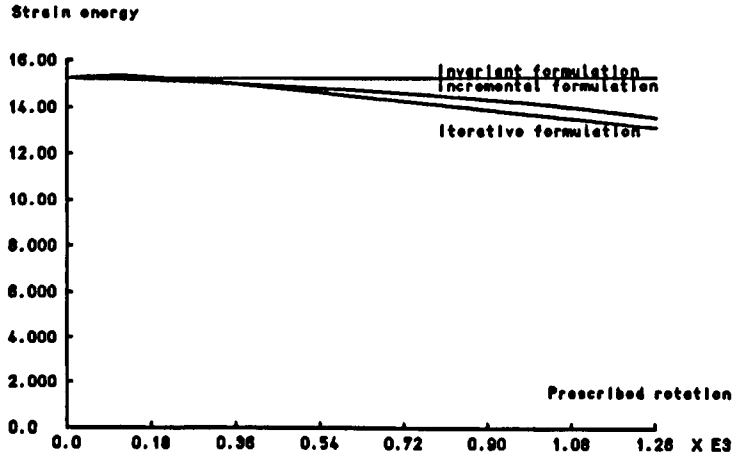



Fig. 5. Displacement components at the free end of the cantilever for first 200 revolutions for the iterative and invariant formulations using the elements with (a) linear interpolation; (b) quadratic interpolation; (c) cubic interpolation.



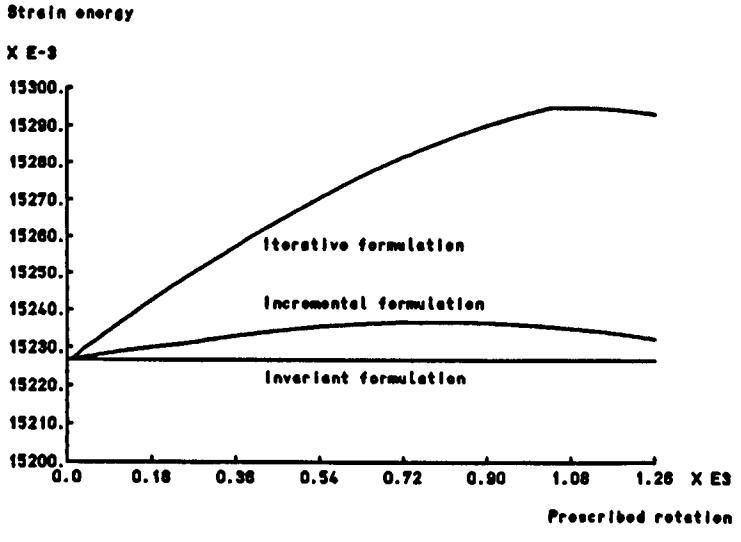




Fig. 6. Strain energy for first 200 revolutions for the incremental, iterative and invariant formulations using the elements with: (a) linear interpolation; (b) quadratic interpolation; (c) cubic interpolation.

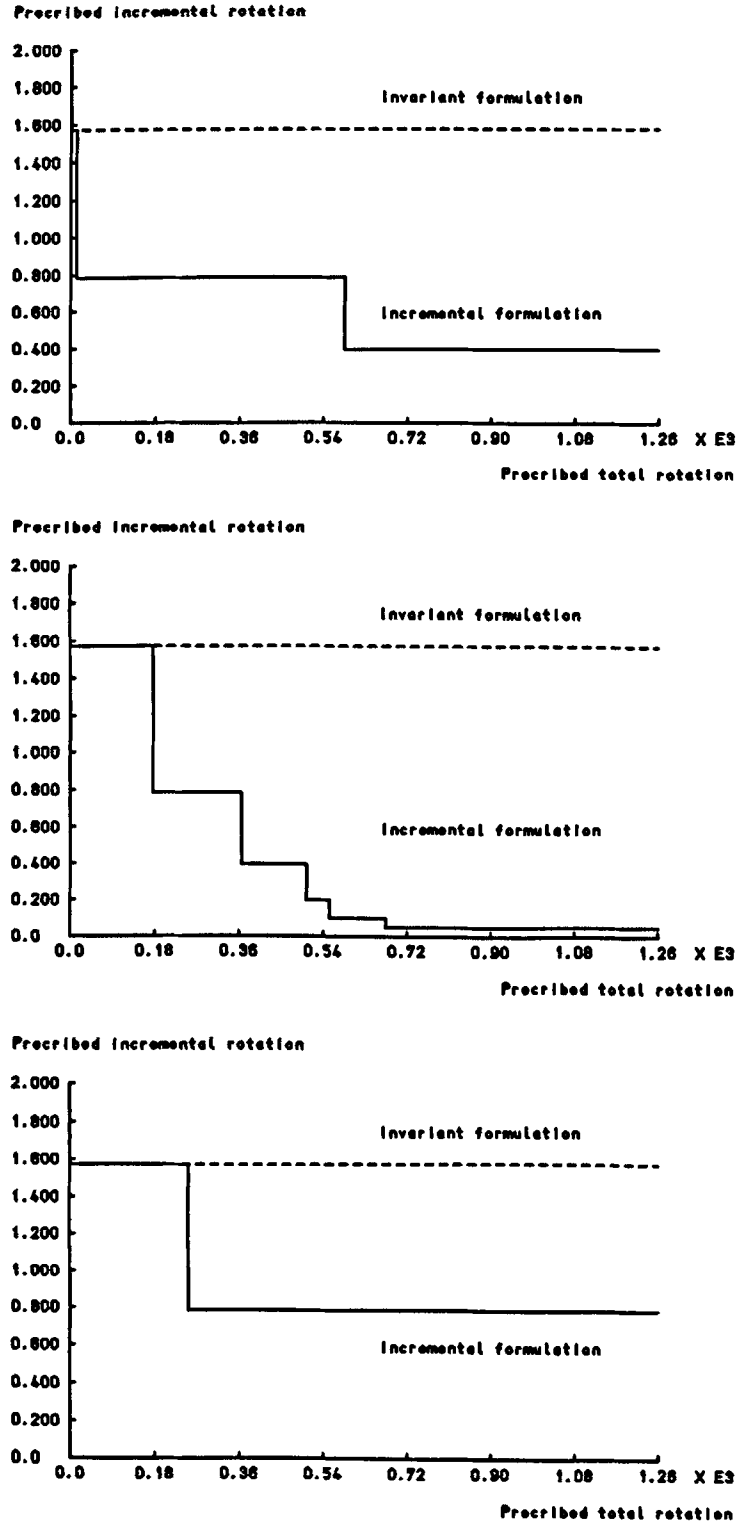



Fig. 7. Incremental rotation against the total prescribed rotation for first 200 revolutions for the incremental and invariant formulations using the elements with: (a) linear interpolation; (b) quadratic interpolation; (c) cubic interpolation.

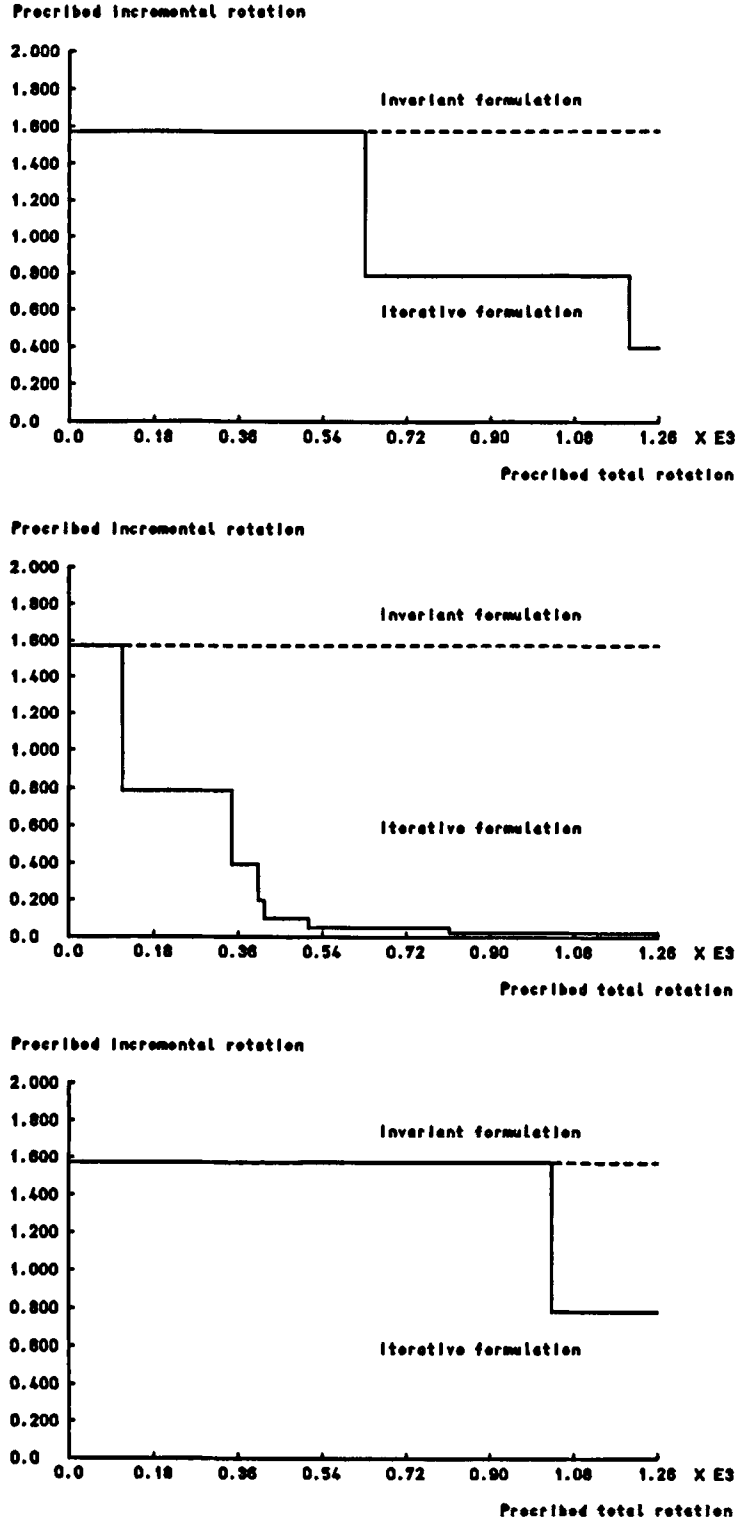



Fig. 8. Intremental rotation against the total prescribed rotation for first 200 revolutions for the iterative and invariant formulations using the elements with: (a) linear interpolation; (b) quadratic interpolation; (c) cubic interpolation.

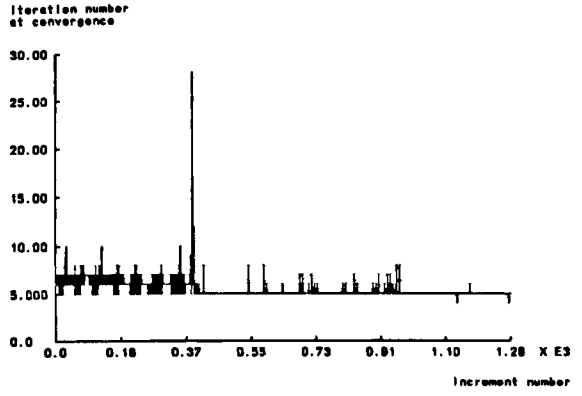



Fig. 9. Number of iterations needed to achieve converged solution against the increment number for the iterative formulation using the elements with linear interpolation.

### 5.5. Example 5: Large displacement vibration of the elbow cantilever

All of the formulations discussed in this paper can be used to run dynamic applications. However, all of these formulations suffer from numerical instabilities typical for the Newmark family of algorithms when applied to nonlinear problems [28–30]. In our subsequent work we aim to address these problems in detail and propose a suitable remedy. However, the present invariant dynamic formulation is an important first step and hence we now detail the results for a dynamic example.

The iterative solution procedure is set up in such a way that if the equilibrium for a particular time step has not been obtained within a prescribed number of iterations the time-step length is halved and the iterative procedure for that time step is repeated. Such a reduced time step will then be gradually increased by the factor of 1.2 until the original time-step length is restored. Dynamic equilibrium at any time step is taken to be satisfied when the convergence criteria (5.1) and (5.2) are satisfied for $\delta_u = \delta_f = 10^{-6}$.

We examine the elbow cantilever in Fig. 3, but now we subject it to an out-of-plane force load $\boldsymbol{F} = F\boldsymbol{E}_3$ applied at its elbow and pointing upwards in contrast to the downward force at the tip, shown in Fig. 3. We do not apply any prescribed rotation at the fixed end. The force is linearly increased from $F = 0$ at $t = 0$ to $F = 50$ at $t = 1$ and then linearly decreased to $F = 0$ at $t = 2$. From then onwards the frame is only subject to large-displacement vibration. In order to emphasise the influence of the rotational degrees of freedom, we artificially exaggerate the mass moments of inertia of the cross sections by choosing $\rho A = 1$ and $\rho I_2 = \rho I_3 = 10$. This example was originally solved by Simo and Vu-Quoc [8], who adopted a time-step $\Delta t = 0.25$ and analysed the motion during the first 30 units of time using the trapezoidal rule ($\beta = 0.25$ and $\gamma = 0.5$) for two different meshes.

We reproduce the results of Simo and Vu-Quoc [8] for the coarse mesh (composed of one quadratic element per leg) by utilising the current iterative formulation. The results obtained by running this example using the incremental, the total and the new, invariant, formulation are very similar. In particular, Fig. 10(a) shows that the results for the vertical displacements of the elbow and the tip for the first 50 units of time graphically coincide for the iterative and the incremental formulation. The fact that the results for the total and the invariant formulation also almost coincide (see Fig. 10(b)) should be qualified by the observation that, for this example, the rigid-body rotations are small. We recall that it is the handling of the rigid-body rotations that basically distinguishes the two formulations.

It can be seen from Fig. 11 that the total energy in all the formulations begins to oscillate towards the end of the analysed period of time, which is an indication of the numerical instability. The influence of the higher modes of vibration, which are generally responsible for this instability, can be reduced by applying some numerical damping. While this often prevents an onset of numerical instability it also reduces the total energy in the system and is not particularly useful in long term dynamic analyses.

Fig. 12(a) shows that if the current example is run with the maximum recommended numerical damping

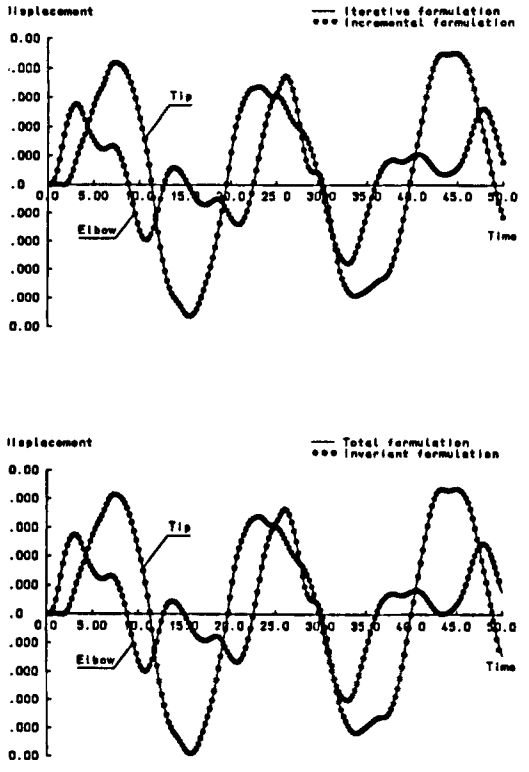



Fig. 10. Vertical displacements of the elbow and the tip of the right-angle elbow cantilever for the first 50 units of time: (a) iterative and incremental formulations; (b) total and invariant formulations.

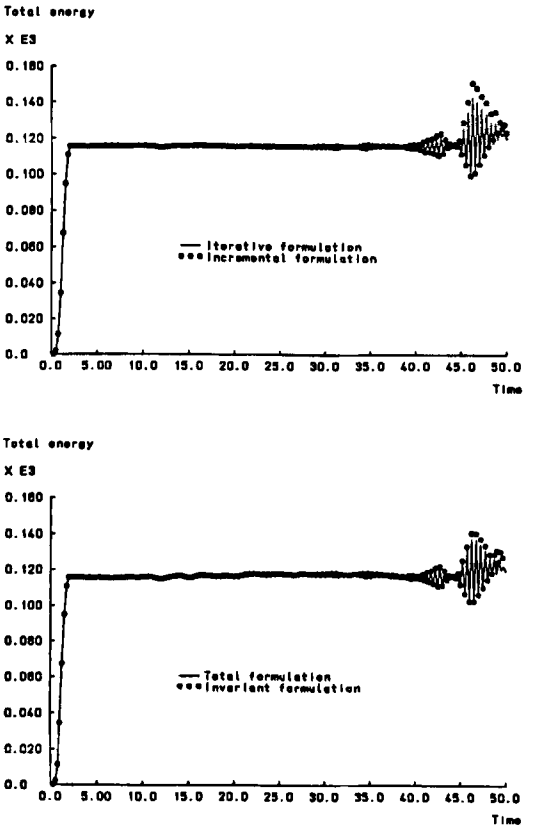



Fig. 11. Total energy of the right-angle elbow cantilever for the first 50 units of time: (a) iterative and incremental formulations; (b) total and invariant formulations.

($\alpha = -0.033$, $\beta = 0.442225$ and $\gamma = 0.83$) for 14 400 units of time (57 600 time steps), the final total energy in all the formulations drops well below 10% of its actual value (see [7] for the application of the $\alpha$-method to the problems of dynamics of geometrically exact 3D beams). The invariant and, to a lesser extent, the iterative formulation appear to preserve more of the energy, but they also experience some undesirable energy growth. However, the latter should not be attributed to particular formulations but is rather a deficiency of the $\alpha$-method (which is originally developed for linear problems) when applied to nonlinear problems. Indeed, if the same example is run by using the coarsest possible mesh, in which each leg is modelled by a single element with linear interpolation, all the formulations experience energy growth and are thus susceptible to eventual energy blow-up. The results are given in Fig. 12(b).

The energy growth for the invariant formulation in Fig. 12(b) should not be attributed to the new concepts, but rather to the interpolation of total quantities. Indeed, if the example is run by using the incremental invariant formulation, briefly mentioned in Note 6, the energy graphs for both meshes (Fig. 12(a) and 12(b)) graphically coincide with the energy graphs of the standard incremental formulation (marked as '3'). This is again so because of the lack of particularly large rigid-body rotations in this example. It is worth noting that this incremental invariant formulation fits more naturally with dynamic procedures featuring numerical time integration in an incremental sense. The path-dependence of this formulation, as indeed a path-dependence of any other dynamic formulation, does not have the importance it has for static analyses, because dynamic time-stepping procedures are fundamentally path-dependent.

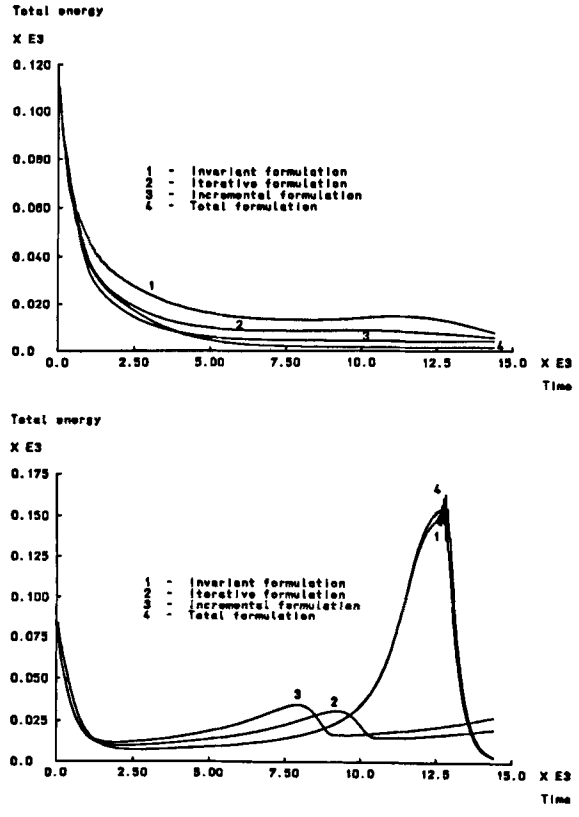



Fig. 12. Total energy of the right-angle elbow cantilever for the first 14 400 units of time: (a) one quadratic element per leg; (b) one linear element per leg.

## 6. Conclusions

In the presence of large 3D rotations the concept of objectivity of strain measures does not extend naturally from the theory to the finite element formulation. Consequently, the finite element implementation of a strain-invariant theory has to be specifically designed to inherit this important property. A new geometrically exact strain-invariant 3D beam finite element formulation is proposed in the paper. This formulation has been used together with some popular, but non-invariant and sometimes even path-dependent formulations, to run a series of illustrative numerical examples. In general, the following conclusions can be drawn:

(1) Formulations based on the standard (configuration-independent) interpolation of total rotations are non-invariant.
(2) Formulations based on the standard interpolation of incremental rotations are non-invariant and dependent on the history of incrementation.
(3) Formulations based on the standard interpolation of iterative rotations are non-invariant and dependent on the iterative path toward a converged solution.

The non-invariance and path-dependence in these formulations decrease with both $p$-refinement and $h$-refinement and in practical applications cannot always be easily spotted. While the resulting errors may sometimes be small, it is important to bear these deficiencies in mind. The formulation proposed to alleviate these problems is based on a standard (configuration-independent) interpolation of current *local* rotations with respect to a reference triad, which is constant for an element. While this principle resembles the one adopted in co-rotational techniques, it must be noted that with respect to co-rotational techniques, the proposed formulation is advantageous in the following features:

(4) The geometric exactness of the theory is preserved in the sense that the current local rotations can take any values between $-\pi$ and $\pi$.
(5) The introduction of the concept of configuration-dependent generalised shape functions to interpolate the trial solutions allows a definition of kinetic energy directly with respect to the inertial frame, thus avoiding non-desirable couplings due to centrifugal and Coriolis effects.

A fully consistent, quadratically convergent, approach is provided by applying a Petrov–Galerkin methodology, in which the test functions are interpolated using standard interpolation polynomials. In this approach, the consistent linearisation is straightforward and results in a non-symmetric equivalent tangent stiffness matrix. Work is in progress on the introduction of the current strain-invariant concepts into an energy and momentum conserving formulation for dynamics of 3D beams.

## Acknowledgement

This work has been financially supported by the Engineering and Physical Sciences Research Council of Great Britain under the contract AEST 2504.

## References

[1] E. Reissner, On one-dimensional finite-strain beam theory: The plane problem, J. Appl. Math. Phys. (ZAMP) 23 (1972) 795–804.
[2] E. Reissner, On finite deformations of space-curved beams, J. Appl. Math. Phys. (ZAMP) 32 (1981) 734–744.
[3] J.C. Simo, A finite strain beam formulation. The three-dimensional dynamic problem. Part I, Comput. Methods Appl. Mech. Engrg. 49 (1985) 55–70.
[4] J.C. Simo and L. Vu-Quoc, A three-dimensional finite-strain rod model. Part II: Computational Aspects, Comput. Methods Appl. Mech. Engrg. 58 (1986) 79–116.
[5] A. Cardona and M. Géradin, A beam finite element non-linear theory with finite rotations, Int. J. Numer. Methods Engrg. 26 (1988) 2403–2438.
[6] A. Ibrahimbegović, F. Frey and I. Kožar, Computational aspects of vector-like parametrization of three-dimensional finite rotations, Int. J. Numer. Methods Engrg. 38 (1995) 3653–3673.
[7] G. Jelenić and M.A. Crisfield, Interpolation of rotational variables in nonlinear dynamics of 3D beams, Int. J. Numer. Methods Engrg. 43 (1998) 1193–1222.
[8] J.C. Simo and L. Vu-Quoc, On the dynamics in space of rods undergoing large motions—a geometrically exact approach, Comput. Methods Appl. Mech. Engrg. 66 (1988) 125–161.
[9] G. Jelenić and M. Saje, A kinematically exact space finite strain beam model—finite element formulation by generalized virtual work principle, Comput. Methods Appl. Mech. Engrg. 120 (1995) 131–161.
[10] K.-J. Bathe and S. Bolourchi, Large displacement analysis of three-dimensional beam structures, Int. J. Numer. Methods Engrg. 14 (1979) 961–986.
[11] E.N. Dvorkin, E. Oñate and J. Oliver, On a non-linear formulation for curved Timoshenko beam elements considering large displacement/rotation increments, Int. J. Numer. Methods Engrg. 26 (1988) 1597–1613.
[12] M.A. Crisfield and G. Jelenić, Objectivity of strain measures in geometrically exact 3D beam theory and its finite element implementation, Proc. Roy. Soc. Lond. A (1998) to be published.
[13] M.A. Crisfield, G. Jelenić and U. Galvanetto, Finite elements with nonlinear statics and dynamics, Proc. IUTAM/IACM Symposium on Discretization Methods in Structural Mechanics, Vienna, Austria, 2–6 June 1997.
[14] M.A. Crisfield, Non-linear Finite Element Analysis of Solids and Structures, Vol. 2, Advanced Topics (John Wiley & Sons, Chichester–New York–Weinheim–Brisbane–Singapore–Toronto, 1997).
[15] M.A. Crisfield, A consistent co-rotational formulation for non-linear, three-dimensional, beam-elements, Comput. Methods Appl. Mech. Engrg. 81 (1990) 131–150.
[16] J.C. Simo and L. Vu-Quoc, On the dynamics of flexible beams under large overall motions—The plane case: Part I. J. Appl. Mech. 53 (1986) 849–854.
[17] J. Argyris, An excursion into large rotations, Comput. Methods Appl. Mech. Engrg. 32 (1982) 85–155.
[18] N.M. Newmark, A method of computation for structural dynamics, J. Engrg. Mech. Div., ASCE, EM3 85 (1959) 67–94.
[19] H.M. Hilber, T.J.R. Hughes and R.L. Taylor, Improved numerical dissipation for time integration algorithms in structural dynamics, Earth. Engrg. Struct. Dyn. 5 (1977) 283–292.
[20] J.C. Simo and K.K. Wong, Unconditionally stable algorithms for rigid body dynamics that exactly preserve energy and momentum, Int. J. Numer. Methods Engrg. 31 (1991) 19–52.
[21] B.A. Finlayson, The Method of Weighted Residuals and Variational Principles (Academic Press, New York and London, 1972).
[22] T.J.R. Hughes, The Finite Element Method (Prentice-Hall, Inc., Englewood Cliffs, NJ, 1987).

[23] C. Johnson, Numerical Solution of Partial Differential Equations by the Finite Element Method (Cambridge University Press, Cambridge, 1995).
[24] R.A. Spurrier, Comments on 'Singularity-free extraction of a quartonian from a discrete cosine matrix', J. Spacecraft 15 (1978) 255.
[25] N. Buechter and E. Ramm, Shell theory versus degeneration—a comparison in large rotation finite-element analysis, Int. J. Numer. Methods Engrg. 34 (1992) 39–59.
[26] J.C. Simo, The (symmetric) Hessian for geometrically nonlinear models in solid mechanics: Intrinsic definition and geometric interpretation, Comput. Methods Appl. Mech. Engrg. 96 (1992) 189–200.
[27] LUSAS User Manual, FEA Ltd., Kingston upon Thames, 1993.
[28] J.C. Simo and N. Tarnow, The discrete energy-momentum method. Conserving algorithms for nonlinear elastodynamics, J. Appl. Math. Phys. (ZAMP) 43 (1992) 757–792.
[29] M.A. Crisfield and J. Shi, A co-rotational element/time-integration strategy for non-linear dynamics, Int. J. Numer. Methods Engrg. 37 (1994) 1897–1913.
[30] J.C. Simo, N. Tarnow and M. Doblare, Nonlinear dynamics of three-dimensional rods: exact energy and momentum conserving algorithms, Int. J. Numer. Methods Engrg. 38 (1995) 1431–1473.