Comput Mech (2011) 47:187–203
DOI 10.1007/s00466-010-0539-4

ORIGINAL PAPER

# Adding kinematic constraints to purely differential dynamics

**Pierangelo Masarati**

Received: 20 July 2009 / Accepted: 25 August 2010 / Published online: 15 September 2010


**Abstract** The dynamics of unconstrained mechanical systems is governed by Ordinary Differential Equations (ODEs). When kinematic constraints need to be accounted for, Differential-Algebraic Equations (DAEs) arise. This work describes the introduction of kinematic constraints, expressed as algebraic relationships between the coordinates of unconstrained mechanical systems, ensuring compliance of the solution with up to the second-order derivative of holonomic constraint equations within the desired accuracy, without altering the ODE structure of the unconstrained problem. This represents a simple, little-intrusive, yet effective means to enforce kinematic constraints into existing formulations and implementations originally intended to address ODE problems, without the complexity of solving DAEs or resorting to implicit numerical integration schemes, and without altering the number and type of equations of the original unconstrained problem. The proposed formulation is compared to known approaches. Numerical applications of increasing complexity illustrate its distinguishing aspects.

**Keywords** Constrained system dynamics · Constraint stabilization · Numerical drift · Multibody dynamics

## 1 Introduction

The dynamics of constrained mechanical systems stems from Newton's intuitions that momentum rate balances applied forces, and reaction forces are needed to grant mutual equilibrium to the parts of a system. Euler extended the idea to rigid body motion, taking rotation and the associated inertia properties into account. Lagrange formalized the equations of motion of constrained systems, exploiting d'Alembert's key intuition that reaction forces related to ideal constraints do not do work for virtual displacements.

A complete review of classic and contemporary methods was recently given in [1,2]. Interested readers should also consult the works of Simeon [3] and Schiehlen [4]. The two main approaches are associated to Lagrange's equations of the first and second kind. Those of the first kind explicitly augment Newton–Euler's equations that describe the dynamics of an unconstrained system of rigid bodies by kinematic constraints $\boldsymbol{\phi} = \mathbf{0}$, and consistently introduce reaction forces as functions of Lagrange's multipliers $\boldsymbol{\lambda}$, resulting in the set of Differential-Algebraic Equations (DAE)

$$\mathbf{M}\dot{\mathbf{x}} = \mathbf{p} \tag{1a}$$

$$\dot{\mathbf{p}} + \boldsymbol{\phi}_{/\mathbf{x}}\boldsymbol{\lambda} = \mathbf{f}^* \tag{1b}$$

$$\boldsymbol{\phi} = \mathbf{0}. \tag{1c}$$

This is the general form of the 'redundant coordinate set' approach in modern multibody dynamics [4].

The equations of the second kind require to express the physical coordinates $\mathbf{x}$, constrained by Eq. (1c), as functions of a smaller set of independent coordinates $\mathbf{q}$, namely $\mathbf{x} = \boldsymbol{\vartheta}(\mathbf{q}, t)$. Lagrange's equations of motion

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{\mathbf{q}}}\right) - \frac{\partial L}{\partial \mathbf{q}} = \frac{\partial \delta W}{\partial \delta \mathbf{q}} \tag{2}$$

depend on the *Lagrangian* function $L = T - U$, the kinetic co-energy $T$ minus the potential energy $U$ related to the conservative forces acting on the system; $\delta W$ is the virtual work of non-conservative forces. The Lagrangian can be easily written in terms of the Cartesian coordinates $\mathbf{x}$ and their derivatives as functions of $\mathbf{q}$.

P. Masarati (✉)
Dipartimento di Ingegneria Aerospaziale, Politecnico di Milano,
via La Masa 34, 20156 Milan, Italy
e-mail: pierangelo.masarati@polimi.it

Equations (2) do not contain reaction forces, as they represent the projection of the dynamics of the system on the manifold of the constraints. The resulting system of Ordinary Differential Equations (ODE) represents the 'minimal coordinate set' approach in modern multibody dynamics [4]. The relationship $\boldsymbol{\vartheta}$ between the Lagrangian and the Cartesian coordinates is closely related to the holonomic constraints $\boldsymbol{\phi}$; in fact, $\boldsymbol{\phi}(\boldsymbol{\vartheta}) \equiv \mathbf{0}$.

Other approaches formulate the same problem from different perspectives: Gauss' principle of least constraint [5], Hertz's principle of least curvature [6], the Gibbs–Appell formula [7,8], and Maggi-Kane equations [9,10], as discussed, for example, in [1,4,11].

The index of the problem in DAE form is higher than 1 [12]; this requires special care in the definition of the initial conditions. The compliance of the numerical solution with the derivatives of the constraints is not guaranteed. The minimal coordinate set approach requires to map the kinematics of the system on the Lagrangian coordinates $\mathbf{q}$. Many authors worked on this topic at the end of the nineteenth century: the approaches of Maggi [9] and Jourdain [13], revitalized by Kane in the second half of the twentieth century [10], propose schemes to reduce the coordinates to an independent subset. Different algorithms have been formulated to compute efficient and robust coordinate reductions, as discussed in [14]. They basically attempt to compute an efficient complement $\mathbf{Q}$ of matrix $\boldsymbol{\phi}_{/\mathbf{x}}$, such that $\delta\mathbf{x}^* = \mathbf{Q}\delta\mathbf{q}$ intrinsically complies with the derivative of the constraint equation, namely $\boldsymbol{\phi}_{/\mathbf{x}}\mathbf{Q}\delta\mathbf{q} = \mathbf{0}\ \forall\delta\mathbf{q}$. Expressing the second time derivative of $\mathbf{x}$ as $\ddot{\mathbf{x}} = \mathbf{Q}\ddot{\mathbf{q}} + (\dot{\mathbf{x}})_{/\mathbf{q}}\dot{\mathbf{q}} + (\dot{\mathbf{x}})_{/t}$, the equations of motion reduce to

$$\mathbf{Q}^T\mathbf{M}\mathbf{Q}\ddot{\mathbf{q}} = \mathbf{Q}^T\mathbf{f} - \mathbf{Q}^T\mathbf{M}((\dot{\mathbf{x}})_{/\mathbf{q}}\dot{\mathbf{q}} + (\dot{\mathbf{x}})_{/t}). \tag{3}$$

The cost and the complexity is related to formulating $\boldsymbol{\vartheta}$ and computing $\mathbf{Q}$. An indubitable advantage is that mechanisms, with few Lagrangian coordinates, reduce to very compact problems.

As an alternative, the problem can be formulated as unconstrained. The motion is projected on the constraint manifold, without explicitly enforcing the constraints. This is the case, for example, of Gauss' principle of least constraint. Many approaches proposed in the literature for index reduction can be cast in this form (see e.g. [15,16]). However, issues arise when the problem is solved numerically. When the time derivatives of the constraint equations are enforced, a systematic, incremental violation of the invariants of the motion, known as 'drift', may occur. Minimal coordinate set approaches may suffer from drift unless special care is used in eliminating constraint violations. This is also true when coordinates are not explicitly reduced, as in the case under discussion.

In Sect. 2 the problem of the dynamics of systems of bodies subjected to kinematic constraints is discussed. A solution is sought that preserves the ODE form of unconstrained systems by projecting the motion on the constraint manifold. Section 3 illustrates a procedure based on a staggered projection on the constraint manifold of the position and velocity resulting from the numerical integration. This procedure allows to simultaneously comply with holonomic constraints and their time derivatives up to second order, and also with non-holonomic constraints and their first-order time derivative. This section also details how the procedure discussed in this work can be exploited to add kinematic constraint enforcement capability to existing dynamic problems formulated as ODEs. In Sect. 4 the procedure is applied to problems of increasing complexity, compared to other approaches available from the literature.

## 2 Constrained system dynamics in ODE form

The dynamics of a generic unconstrained system are described by

$$\mathbf{M}(\mathbf{x})\ddot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \dot{\mathbf{x}}, t). \tag{4}$$

$\mathbf{M} \in \mathbb{R}^{n\times n} > 0$ is the positive-definite mass matrix; $\mathbf{x} \in \mathbb{R}^n$ are the independent coordinates that describe the motion; $\mathbf{f}\colon \mathbb{R}^{2n+1} \mapsto \mathbb{R}^n$ are the forces acting on the system, including internal forces and the contribution $\dot{\mathbf{M}}\dot{\mathbf{x}}$ to inertia forces when the coordinates $\mathbf{x}$ are not inertial. The coordinates $\mathbf{x}$ are either Cartesian coordinates of an ensemble of point masses, or the kinematic parameters that describe the configuration of a set of rigid bodies. The function

$$\boldsymbol{\phi}(\mathbf{x}, t) = \mathbf{0}, \tag{5}$$

with $\boldsymbol{\phi}\colon \mathbb{R}^{n+1} \mapsto \mathbb{R}^c$, describes a set of $c$ ideal bilateral holonomic constraints between the coordinates. Reaction forces $\mathbf{f}_c = \boldsymbol{\phi}_{/\mathbf{x}}^T\boldsymbol{\lambda}$ enforce the constraints, with $\boldsymbol{\lambda} \in \mathbb{R}^c$.

Equation (5) represents an invariant of the motion. It implies that the kinematic variables $\mathbf{x}$ are not independent; in principle, it allows to reduce them to the kinematic coordinates. The time derivatives of $\boldsymbol{\phi}$ are invariants of the motion as well. In exact mathematics, the invariance of the derivatives is implied. However, when the problem is solved numerically, integration errors may result in a violation of the time derivatives of the constraint equation. This is often accepted, in practice, when DAE problems are directly integrated using L-stable integration schemes [12].

The index of the DAE problem can be naively reduced by differentiating the constraint equation. The second derivative of $\boldsymbol{\phi}$,

$$\dot{\boldsymbol{\phi}} = \underbrace{\boldsymbol{\phi}_{/\mathbf{x}}}_{\mathbf{A}}\dot{\mathbf{x}} + \underbrace{\boldsymbol{\phi}_{/t}}_{-\mathbf{b}'} = \mathbf{0} \tag{6a}$$

$$\ddot{\boldsymbol{\phi}} = \underbrace{\boldsymbol{\phi}_{/\mathbf{x}}}_{\mathbf{A}} \ddot{\mathbf{x}} + \underbrace{\left(\dot{\boldsymbol{\phi}}\right)_{/\mathbf{x}} \dot{\mathbf{x}} + \left(\dot{\boldsymbol{\phi}}\right)_{/t}}_{-\mathbf{b}''} = \mathbf{0}, \tag{6b}$$

formulates the holonomic constraint equations at the acceleration level. Non-holonomic constraints can be dealt with in a similar manner; since they are less critical, they will not be considered.

When the constraints are expressed at the acceleration level, the index of the underlying DAE problem of Eqs. (4) and (6b) reduces from 3 to 1. This operation, however, cannot be performed without precautions, since it corresponds to enforcing only the invariance of a derivative of the constraint. The invariance of the first-order derivative of a constraint equation implies the invariance of the constraint, except for a constant violation. The invariance of the second-order derivative no longer implies the invariance of the constraint, not even within exact mathematics. The evidence of this issue is numerical drift of the constraint, which is an erratic growth of the constraint violation.

Instead of exploiting the constraint equation to reduce the problem to the minimal set representation of Eq. (3), another formally pure ODE representation can be obtained by substituting $\ddot{\mathbf{x}}$ from Eq. (4) in Eq. (6b), computing the multipliers from the latter as

$$\boldsymbol{\lambda} = \left(\mathbf{A}\mathbf{M}^{-1}\mathbf{A}^T\right)^+ \left(\mathbf{A}\mathbf{a} - \mathbf{b}''\right), \tag{7}$$

and then eliminating them from Eq. (4),

$$\ddot{\mathbf{x}} = \mathbf{a} - \mathbf{M}^{-1}\mathbf{A}^T \left(\mathbf{A}\mathbf{M}^{-1}\mathbf{A}^T\right)^+ \left(\mathbf{b}'' - \mathbf{A}\mathbf{a}\right), \tag{8}$$

where $\mathbf{a} = \mathbf{M}^{-1}\mathbf{f}$ are the accelerations of the unconstrained system. When $\mathbf{A}$ is not full rank, and thus $\mathbf{A}\mathbf{M}^{-1}\mathbf{A}^T$ is singular, Eq. (7) yields a minimal norm of the multipliers, weighted by the mass of the system, using the Moore-Penrose Generalized Inverse (MPGI). This allows to account for redundant constraints ([17], as reported in [1,18,19]).

According to Gauss' principle [5], Eq. (8) minimizes the *Gaussian*,

$$G = \frac{1}{2}\left(\ddot{\mathbf{x}} - \mathbf{a}\right)^T \mathbf{M}\left(\ddot{\mathbf{x}} - \mathbf{a}\right), \tag{9}$$

subjected to the constraint $\mathbf{A}\ddot{\mathbf{x}} = \mathbf{b}''$. The equivalence of Eqs. (8) and (7) with the corresponding forms initially proposed in [18] by Udwadia and Kalaba is discussed in [20]. The elimination of the multipliers to obtain Eq. (8) may not be trivial when non-ideal constraints need to be considered. However, this is outside the scope of the present work.

Interesting approaches to constraint stabilization have been proposed. Baumgarte, in [21], proposed to enforce a linear combination of the holonomic constraint and its derivatives up to second order. As a consequence, the problem becomes second-order linear differential in the constraint violation, which can be quickly damped by an appropriate choice of the weights. However, the quicker the artificial transient is, the stiffer the problem becomes. Gear et al., in [22], proposed to add the first derivative of holonomic constraints, as a function of generalized velocities, modifying their definition with an additional multiplier. This approach was later extended by Führer and Leimkuhler by also enforcing the second derivative [23]. Bayo et al., in [24], generalized this approach by using weighting matrices. Mattsson and Söderlind proposed an index reduction technique by dummy variable substitution [25]. This technique, based on Pantelides' algorithm [26], is at the core of the Modelica concept [27]. Bauchau proposed a technique bundled in the discretization of the constraint equation and of its derivative, exploiting energy preservation/dissipation properties of the integration algorithm [28]. Borri et al. proposed an Embedded Projection Method (EPM) that structurally reduces arbitrary index DAEs to index 1 form [29]. The resulting problem, consisting in a sequence of ODE and algebraic subproblems, is solved in a staggered manner. Recently, Braun and Goldfarb [30] proposed a technique which essentially consists in formulating the problem in first-order form by adding the definition of the velocities $\dot{\mathbf{q}} = \mathbf{v}$, and in applying an equivalent of Baumgarte's correction.

## 3 Proposed approach

### 3.1 Stabilized constrained dynamics

The constrained dynamics problem in first-order differential form, enforcing both velocity- and acceleration-level constraints, is

$$\dot{\mathbf{x}} + \mathbf{W}^{-1}\mathbf{A}^T\boldsymbol{\mu} = \mathbf{v} \tag{10a}$$

$$\mathbf{M}\dot{\mathbf{v}} + \mathbf{A}^T\boldsymbol{\lambda} = \mathbf{f} \tag{10b}$$

$$\mathbf{A}\dot{\mathbf{x}} = \mathbf{b}' \tag{10c}$$

$$\mathbf{A}\dot{\mathbf{v}} = \mathbf{b}'', \tag{10d}$$

where $\dot{\mathbf{v}}$ is used instead of $\ddot{\mathbf{x}}$ in Eq. (10d). Enforcing the derivatives of the constraint equation to the acceleration level guarantees the reduction to index 1 of the original DAE.

The enforcement of another constraint equation requires to correct the definition of $\mathbf{v}$ with the contribution of another multiplier $\boldsymbol{\mu}$. This removes from $\mathbf{v}$ its projection in a subspace of $\mathbf{x}$ that is skewed with respect to the constraint manifold, yielding Eq. (10a). The arbitrarily chosen weight matrix $\mathbf{W}$ is used to distort the projection. The dimension of $\boldsymbol{\mu}$ is that of a velocity times the dimension of $\mathbf{W}$, [W].

Gear et al. [22] suggest to project the velocity correction in a subspace orthogonal to the constraint manifold, using $\mathbf{W} = \mathbf{I}$. Other authors suggest to use a 'dynamic' projection

with $\mathbf{W} = \mathbf{M}$. Hairer and Wanner modify in this manner the stabilization proposed by Gear et al. "... for the sake of symmetry" [31]. Borri et al. justify it on the basis that in Hamiltonian dynamics the state is composed of the coordinates $\mathbf{x}$ and momenta $\mathbf{p}$, so the correction naturally applies to momenta [29]. Braun and Goldfarb justify the corresponding multiplier $\boldsymbol{\mu}$ in terms of a 'mechanical impulse' [30].

Here a heuristic interpretation in terms of velocity correction that compensates for numerical integration errors is preferred, with $\mathbf{W} = \mathbf{I}$. In fact, in this case Eq. (10a) guarantees that the velocity correction $\mathbf{A}^T\boldsymbol{\mu}$ is orthogonal to any virtual displacement $\delta\mathbf{x}$, which must satisfy $\delta\boldsymbol{\phi} = \mathbf{A}\delta\mathbf{x} = \mathbf{0}$. Similar considerations apply to the reaction forces $\mathbf{f}_c = -\mathbf{A}^T\boldsymbol{\lambda}$, which intrinsically do not do work for a virtual displacement $\delta\mathbf{x}$. However, for generality, a weighting $\mathbf{W}$ is considered throughout the formalization of the approach.

### 3.2 Stabilized constrained dynamics in ODE form

An explicit expression for the derivative of the state can be obtained:

$$\dot{\mathbf{x}} = \mathbf{v} - \mathbf{W}^{-1}\mathbf{A}^T\left(\mathbf{A}\mathbf{W}^{-1}\mathbf{A}^T\right)^+\left(\mathbf{A}\mathbf{v} - \mathbf{b}'\right) \tag{11a}$$

$$\dot{\mathbf{v}} = \mathbf{a} - \mathbf{M}^{-1}\mathbf{A}^T\left(\mathbf{A}\mathbf{M}^{-1}\mathbf{A}^T\right)^+\left(\mathbf{A}\mathbf{a} - \mathbf{b}''\right). \tag{11b}$$

After defining

$$\mathbf{P}_\mathbf{v} = \mathbf{W}^{-1}\mathbf{A}^T\left(\mathbf{A}\mathbf{W}^{-1}\mathbf{A}^T\right)^+\mathbf{A} \tag{12a}$$

$$\mathbf{p}_\mathbf{v} = \mathbf{W}^{-1}\mathbf{A}^T\left(\mathbf{A}\mathbf{W}^{-1}\mathbf{A}^T\right)^+\mathbf{b}' \tag{12b}$$

$$\mathbf{P}_\mathbf{a} = \mathbf{M}^{-1}\mathbf{A}^T\left(\mathbf{A}\mathbf{M}^{-1}\mathbf{A}^T\right)^+\mathbf{A} \tag{12c}$$

$$\mathbf{p}_\mathbf{a} = \mathbf{M}^{-1}\mathbf{A}^T\left(\mathbf{A}\mathbf{M}^{-1}\mathbf{A}^T\right)^+\mathbf{b}'', \tag{12d}$$

Eqs. (11) can be rewritten as

$$\dot{\mathbf{x}} = (\mathbf{I} - \mathbf{P}_\mathbf{v})\,\mathbf{v} + \mathbf{p}_\mathbf{v} \tag{13a}$$

$$\dot{\mathbf{v}} = (\mathbf{I} - \mathbf{P}_\mathbf{a})\,\mathbf{a} + \mathbf{p}_\mathbf{a}. \tag{13b}$$

This form highlights the fact that the velocity- and acceleration-level constraints project the unconstrained acceleration $\mathbf{a}$ and its integral $\mathbf{v}$ on the manifold of the constraints by means of the non-orthogonal projectors $\mathbf{P}_\mathbf{a}$ and $\mathbf{P}_\mathbf{v}$. In case of rheonomic constraints, the unconstrained velocity is also corrected by $\mathbf{p}_\mathbf{v}$. The acceleration is corrected by $\mathbf{p}_\mathbf{a}$ as soon as the constraint at least bends the trajectory of the constrained bodies, or is rheonomic.

One may note that the two projectors are different, and might be tempted to use $\mathbf{P}_\mathbf{a}$ for both projections. This may not be a good choice, since the two projectors act on different types of equations: a kinematic definition, Eq. (10a), involving velocities, and an equilibrium equation, Eq. (10b), involving forces; thus, they may legitimately differ. When $\mathbf{W} = \mathbf{I}$, $\mathbf{P}_\mathbf{v} = \mathbf{A}^+\mathbf{A}$, where the property $\mathbf{A}^T(\mathbf{A}\mathbf{A}^T)^+ = \mathbf{A}^+$ of the MPGI is exploited [32]. This projection is orthogonal; correspondingly, $\mathbf{p}_\mathbf{v} = \mathbf{A}^+\mathbf{b}'$.

To yield Eqs. (11), the multipliers are

$$\boldsymbol{\mu} = \left(\mathbf{A}\mathbf{W}^{-1}\mathbf{A}^T\right)^+\left(\mathbf{A}\mathbf{v} - \mathbf{b}'\right) \tag{14a}$$

$$\boldsymbol{\lambda} = \left(\mathbf{A}\mathbf{M}^{-1}\mathbf{A}^T\right)^+\left(\mathbf{A}\mathbf{a} - \mathbf{b}''\right). \tag{14b}$$

The nature of the multiplier $\boldsymbol{\mu}$ differs from that of $\boldsymbol{\lambda}$ because it represents the velocity correction that is required to bring $\mathbf{v}$ on the constraint manifold. As such, it should be identically zero, were the integration exact.

The proposed approach minimizes the functional

$$G(\dot{\mathbf{v}}, \dot{\mathbf{x}}) = \frac{1}{2}(\dot{\mathbf{v}} - \mathbf{a})^T\mathbf{M}(\dot{\mathbf{v}} - \mathbf{a}) + w\frac{1}{2}(\dot{\mathbf{x}} - \mathbf{v})^T\mathbf{W}(\dot{\mathbf{x}} - \mathbf{v}), \tag{15}$$

where the dimensionality is preserved by the uninfluential scalar coefficient $w$, of units $[\mathrm{m}][\mathrm{t}^{-2}][\mathrm{W}^{-1}]$. Equation (15) expresses a modified first-order form of the Gaussian $G(\ddot{\mathbf{x}})$ of Eq. (9). Its minimization is subjected to the constraints $\mathbf{A}\dot{\mathbf{x}} = \mathbf{b}'$ and $\mathbf{A}\dot{\mathbf{v}} = \mathbf{b}''$. The problem is equivalent to Gauss' principle: when exact mathematics are considered, $\dot{\mathbf{x}} \equiv \mathbf{v}$; the velocity constraint implies the acceleration constraint, and $G(\dot{\mathbf{v}}, \dot{\mathbf{x}}) \equiv G(\ddot{\mathbf{x}})$.

### 3.3 Position drift elimination

In case of holonomic constraints, Eq. (5) is enforced by determining

$$\mathbf{x} = \bar{\mathbf{x}} + \mathbf{W}^{-1}\mathbf{A}^T\boldsymbol{\nu}, \tag{16}$$

where $\bar{\mathbf{x}}$ is obtained integrating $\dot{\mathbf{x}}$ from Eq. (11a), and thus may not comply with the constraint because of numerical drift; the resulting $\mathbf{x}$ is the position that complies with Eq. (5).

This guarantees that the correction $\mathbf{W}^{-1}\mathbf{A}^T\boldsymbol{\nu}$ does not alter the solution in directions that are locally compatible with the constraints. If $\mathbf{W} = \mathbf{I}$, it is a minimal norm correction with respect to the constraint violation. A Newton iteration can be used:

$$\mathbf{x}_0 = \bar{\mathbf{x}} \tag{17a}$$

$$\mathbf{x}_{i+1} = \mathbf{x}_i - \mathbf{W}^{-1}\mathbf{A}^T\left(\mathbf{A}\mathbf{W}^{-1}\mathbf{A}^T\right)^+\boldsymbol{\phi}(\mathbf{x}_i, t). \tag{17b}$$

A solution of Eq. (5) must exist as soon as the constraints are consistent, regardless of being independent or not (constraints are not independent when matrix $\mathbf{A}$ is not full rank; they are inconsistent when they are not independent, and $\boldsymbol{\phi} = \mathbf{0}$ has no solution).

The proposed approach reduces to 1 the index of DAE problems that describe the dynamics of constrained mechanical systems, since it enforces compliance with the constraints

up to the second-order derivative of the coordinates $\mathbf{x}$. In case of holonomic constraints, an embedded compliance iteration allows to substitute the multipliers, yielding a formally pure ODE representation of the constrained dynamics problem.

In 1993, Eich [15] proposed a projection formally identical to that of Eq. (12a), with $\mathbf{W} = \mathbf{I}$, to project the integrated coordinates $\bar{\mathbf{x}}$ onto the manifold of a linear constraint $\boldsymbol{\phi}(\mathbf{x}, t) = \mathbf{A}\mathbf{x} - \mathbf{b}(t) = \mathbf{0}$.

In 2000, Yu and Chen [33] proposed an interesting algorithm, that consists in integrating Eqs. (4) and (6b) directly, applying some correction to the resulting $\dot{\mathbf{x}}$ and $\mathbf{x}$ only when the constraint violation exceeds a given tolerance. The correction to $\mathbf{x}$ is analogous to that of Eq. (17), with $\mathbf{W} = \mathbf{I}$. The proposed approach differs from that of Yu and Chen by embedding the correction to $\dot{\mathbf{x}}$ in the problem, and by always applying the correction of Eq. (17).

### 3.4 Application to ODE problems

Consider a typical problem in fully implicit ODE form

$$\mathbf{g}\,(\mathbf{y}, \dot{\mathbf{y}}, t) = \mathbf{0}, \tag{18}$$

where $\mathbf{y}$ are the state variables, not necessarily related to kinematics or kinetics. The portion of $\mathbf{y}$ not related to mechanics is called $\mathbf{w}$ in the following.

The need to apply kinematic constraints in form of algebraic equations would change the nature of the problem from ODE to DAE, possibly requiring to entirely reformulate the problem. On the contrary, the proposed approach would allow to preserve the purely differential nature of the problem, moving the part related to kinematic constraints to an external loop that:

- corrects the force and velocity deficit required to constrain the motion of the system on the constraint manifold, Eqs. (11);
- optionally corrects the position resulting from the integration of Eq. (11a) according to Eq. (16), if holonomic constraints are considered.

From the point of view of developing and implementing a formulation for time-accurate integration of constrained dynamics problem, Eqs. (11) can be rewritten in the equivalent form

$$\dot{\mathbf{x}} = \mathbf{v} - \mathbf{W}^{-1}\mathbf{A}^T \left(\mathbf{A}\mathbf{W}^{-1}\mathbf{A}^T\right)^+ \left(\mathbf{A}\mathbf{v} - \mathbf{b}'\right) \tag{19a}$$

$$\mathbf{M}\dot{\mathbf{v}} = \mathbf{f} - \mathbf{A}^T \left(\mathbf{A}\mathbf{M}^{-1}\mathbf{A}^T\right)^+ \left(\mathbf{A}\mathbf{M}^{-1}\mathbf{f} - \mathbf{b}''\right), \tag{19b}$$

typical of implicit formulations of mechanical problems. This form clearly shows how the constraints contribute to correcting the velocities in Eq. (19a), and significantly the forces in Eq. (19b).

The dynamics of a mechanical system are usually expressed in the form of Eq. (19b), unless the mass matrix $\mathbf{M}$ is trivially diagonal. The inverse of the mass matrix (actually, its factorization) is still required to compute the reduced mass matrix

$$\tilde{\mathbf{M}} = \left(\mathbf{A}\mathbf{M}^{-1}\mathbf{A}^T\right)^+, \tag{20}$$

and the reduced unconstrained acceleration

$$\tilde{\mathbf{a}} = \mathbf{A}\mathbf{M}^{-1}\mathbf{f}, \tag{21}$$

but in many practical cases either only a fraction of it needs to be computed and factorized, or its computation and factorization are not excessively demanding.

#### *3.4.1 Rigid bodies*

When $b$ independent rigid bodies are considered, the mass matrix $\mathbf{M}$ is block-diagonal instead of diagonal, i.e. $\mathbf{M} = \mathrm{diag}(\mathbf{M}_1, \mathbf{M}_2, \ldots, \mathbf{M}_b)$, with

$$\mathbf{M}_i = \begin{bmatrix} m_i\mathbf{I} & \mathbf{s}_i \times^T \\ \mathbf{s}_i \times & \mathbf{J}_i \end{bmatrix}, \tag{22}$$

where $m_i$, $\mathbf{s}_i$ and $\mathbf{J}_i$ are the 0th, 1st, and 2nd order inertia moments of the $i$th body. The computation of the reduced mass matrix of Eq. (20) only requires the inverse of the blocks related to the constrained nodes. For example, if the holonomic constraint connects only nodes 1 and 2, and thus $\mathbf{A} = [\mathbf{A}_1, \mathbf{A}_2, \mathbf{0}, \ldots, \mathbf{0}]$, then $\tilde{\mathbf{M}} = (\mathbf{A}_1\mathbf{M}_1^{-1}\mathbf{A}_1^T + \mathbf{A}_2\mathbf{M}_2^{-1}\mathbf{A}_2^T)^+$. Also, Eq. (21) reduces to $\tilde{\mathbf{a}} = \mathbf{A}_1\mathbf{M}_1^{-1}\mathbf{f}_1 + \mathbf{A}_2\mathbf{M}_2^{-1}\mathbf{f}_2$, if $\mathbf{f}_1$ and $\mathbf{f}_2$ are the portions of $\mathbf{f} = \{\mathbf{f}_1; \mathbf{f}_2; \ldots; \mathbf{f}_b\}$ that act on bodies 1 and 2. The state vector of Eq. (18) is $\mathbf{y} = \{\mathbf{x}; \mathbf{v}; \mathbf{w}\}$.

#### *3.4.2 Ritz forms*

When the kinematics is modeled using $r$ Ritz forms (i.e. independent, compliant with boundary conditions) $\mathbf{x}_i \in \mathbb{R}^n$, with $i = 1, r$ and $r \leq n$, namely $\mathbf{x} = \sum_{i=1,r} \mathbf{x}_i q_i = \mathbf{X}\mathbf{q}$, with $\mathbf{X} \in \mathbb{R}^{n\times r}$ full column rank, and $\mathbf{q} \in \mathbb{R}^r$ the corresponding set of reduced coordinates, the generalized mass matrix is $\hat{\mathbf{M}} = \mathbf{X}^T\mathbf{M}\mathbf{X}$. By analogy, a generalized weight matrix $\hat{\mathbf{W}} = \mathbf{X}^T\mathbf{W}\mathbf{X}$ is introduced. When the kinematic constraints $\boldsymbol{\phi}$ are written as functions of the physical kinematics expressed by $\mathbf{x}$, Eqs. (19) become

$$\dot{\mathbf{q}} = \mathbf{r} - \hat{\mathbf{W}}^{-1}\mathbf{X}^T\mathbf{A}^T \left(\mathbf{A}\mathbf{X}\hat{\mathbf{W}}^{-1}\mathbf{X}^T\mathbf{A}^T\right)^+ \left(\mathbf{A}\mathbf{X}\mathbf{r} - \mathbf{b}'\right) \tag{23a}$$

$$\hat{\mathbf{M}}\dot{\mathbf{r}} = \mathbf{X}^T\mathbf{f} - \mathbf{X}^T\mathbf{A}^T\left(\mathbf{A}\mathbf{X}\hat{\mathbf{M}}^{-1}\mathbf{X}^T\mathbf{A}^T\right)^+ \left(\mathbf{A}\mathbf{X}\hat{\mathbf{M}}^{-1}\mathbf{X}^T\mathbf{f} - \mathbf{b}''\right), \tag{23b}$$

which corresponds to projecting the constraint matrix $\mathbf{A}$ and the forces $\mathbf{f}$ onto the space of the Ritz forms. So, after defining $\hat{\mathbf{A}} = \mathbf{A}\mathbf{X}$ and $\hat{\mathbf{f}} = \mathbf{X}^T\mathbf{f}$, Eqs. (23) become

$$\dot{\mathbf{q}} = \mathbf{r} - \hat{\mathbf{W}}^{-1}\hat{\mathbf{A}}^T\left(\hat{\mathbf{A}}\hat{\mathbf{W}}^{-1}\hat{\mathbf{A}}^T\right)^+\left(\hat{\mathbf{A}}\mathbf{r} - \mathbf{b}'\right) \tag{24a}$$

$$\hat{\mathbf{M}}\dot{\mathbf{r}} = \hat{\mathbf{f}} - \hat{\mathbf{A}}^T\left(\hat{\mathbf{A}}\hat{\mathbf{M}}^{-1}\hat{\mathbf{A}}^T\right)^+\left(\hat{\mathbf{A}}\hat{\mathbf{M}}^{-1}\hat{\mathbf{f}} - \mathbf{b}''\right). \tag{24b}$$

In some cases, the Ritz forms might not allow to describe the kinematics of constrained points with enough detail. In this case matrix $\hat{\mathbf{A}}$ might not be full row-rank, although matrix $\mathbf{X}$ is full column-rank, even in case matrix $\mathbf{A}$ is full row-rank as well. This is yet another reason to require the use of the MPGI for matrices $\hat{\mathbf{A}}\hat{\mathbf{W}}^{-1}\hat{\mathbf{A}}^T$ and $\hat{\mathbf{A}}\hat{\mathbf{M}}^{-1}\hat{\mathbf{A}}^T$ in Eqs. (24).

The correction of Eq. (16) becomes

$$\mathbf{q} = \overline{\mathbf{q}} + \hat{\mathbf{W}}^{-1}\hat{\mathbf{A}}^T\boldsymbol{\nu}, \tag{25}$$

where $\overline{\mathbf{q}}$ are the (possibly non-compliant) coordinates resulting from the integration of $\dot{\mathbf{q}}$, and the correction iteration becomes

$$\mathbf{q}_0 = \overline{\mathbf{q}} \tag{26a}$$

$$\mathbf{q}_{i+1} = \mathbf{q}_i - \hat{\mathbf{W}}^{-1}\hat{\mathbf{A}}^T\left(\hat{\mathbf{A}}\hat{\mathbf{W}}^{-1}\hat{\mathbf{A}}^T\right)^+\boldsymbol{\phi}\left(\mathbf{X}\mathbf{q}_i, t\right). \tag{26b}$$

If matrix $\hat{\mathbf{A}}$ is not full row-rank, the constraint equation might not be satisfied within the desired tolerance; thus constraints become mildly inconsistent.

To summarize, the use of Ritz forms does not alter the formulation based on Cartesian coordinates, provided the mass matrix, the constraints matrix, and the loads vector are reduced using the forms $\mathbf{X}$. The state vector of Eq. (18) is $\mathbf{y} = \{\mathbf{q}; \mathbf{r}; \mathbf{w}\}$.

### 3.5 Kinematic constraint addition procedure

The addition of a kinematic constraint to the generic problem $\mathbf{g}$ of Eq. (18), while preserving its fully implicit ODE form, according to the proposed procedure requires to:

(a) if the constraint is holonomic, project the value of $\mathbf{x}$ onto the invariant manifold by means of the iteration of Eq. (16);
(b) compute the Jacobian matrix of the constraint, $\mathbf{A}$, using the corrected value of $\mathbf{x}$, if the constraint is holonomic;
(c) compute the right-hand side of the velocity-level constraint, $\mathbf{b}'$, using the corrected value of $\mathbf{x}$, if the constraint is holonomic;
(d) project the value of $\mathbf{v}$ tangent to the invariant manifold by means of Eq. (11a);
(e) compute the right-hand side of the acceleration-level constraint, $\mathbf{b}''$, using the corrected value of $\dot{\mathbf{x}}$ and $\mathbf{x}$, as needed;
(f) compute the reduced mass matrix of Eq. (20) and the unconstrained acceleration vector of Eq. (21), exploiting the simplifications allowed by structure and sparsity of $\mathbf{A}$ and $\mathbf{M}$;
(g) compute the force correction $\mathbf{A}^T\tilde{\mathbf{M}}\left(\tilde{\mathbf{a}} - \mathbf{b}''\right)$; or
(h) the acceleration correction $\mathbf{M}^{-1}\mathbf{A}^T\tilde{\mathbf{M}}\left(\tilde{\mathbf{a}} - \mathbf{b}''\right)$, exploiting the simplifications allowed by structure and sparsity of $\mathbf{A}$ and $\mathbf{M}$;
(i) integrate the original system of ODE with corrected velocity and acceleration or force.

Matrices $\mathbf{M}$ and $\mathbf{A}$, and vector $\mathbf{f}$ need to be replaced by their projections $\hat{\mathbf{M}}$, $\hat{\mathbf{A}}$, and $\hat{\mathbf{f}}$ if the coordinates are not Cartesian.

It is worth noticing that all the corrections to $\mathbf{x}$ and $\mathbf{v}$ can be performed prior to computing the actual value of $\mathbf{f}$. As a consequence, this will be performed based on the corrected value of the state, as its value is required to compute the unconstrained accelerations of the constrained parts of the system for either of Eqs. (11b) or (19b).

Notice also that the computations required for the projection of the mechanical states that participate in the algebraic constraints affects only that subset of the state related to the mechanical portion of the problem, and does not affect any non-mechanical part $\mathbf{w}$ of the state.

These considerations suggest that the intrusiveness of the proposed approach is minimal, and its impact on the modification of existing implementations of purely differential problems can be limited as well.

## 4 Applications

### 4.1 Simple pendulum

Consider a simple rigid point mass pendulum, of mass $m = 1$ kg and length $L = 1$ m, whose dynamics are described by

$$m\ddot{x} + \frac{x}{\sqrt{x^2+z^2}}\lambda = 0 \tag{27a}$$

$$m\ddot{z} + \frac{z}{\sqrt{x^2+z^2}}\lambda = -mg \tag{27b}$$

$$\sqrt{x^2+z^2} = L, \tag{27c}$$

with $g = 9.81$ m/s$^2$.

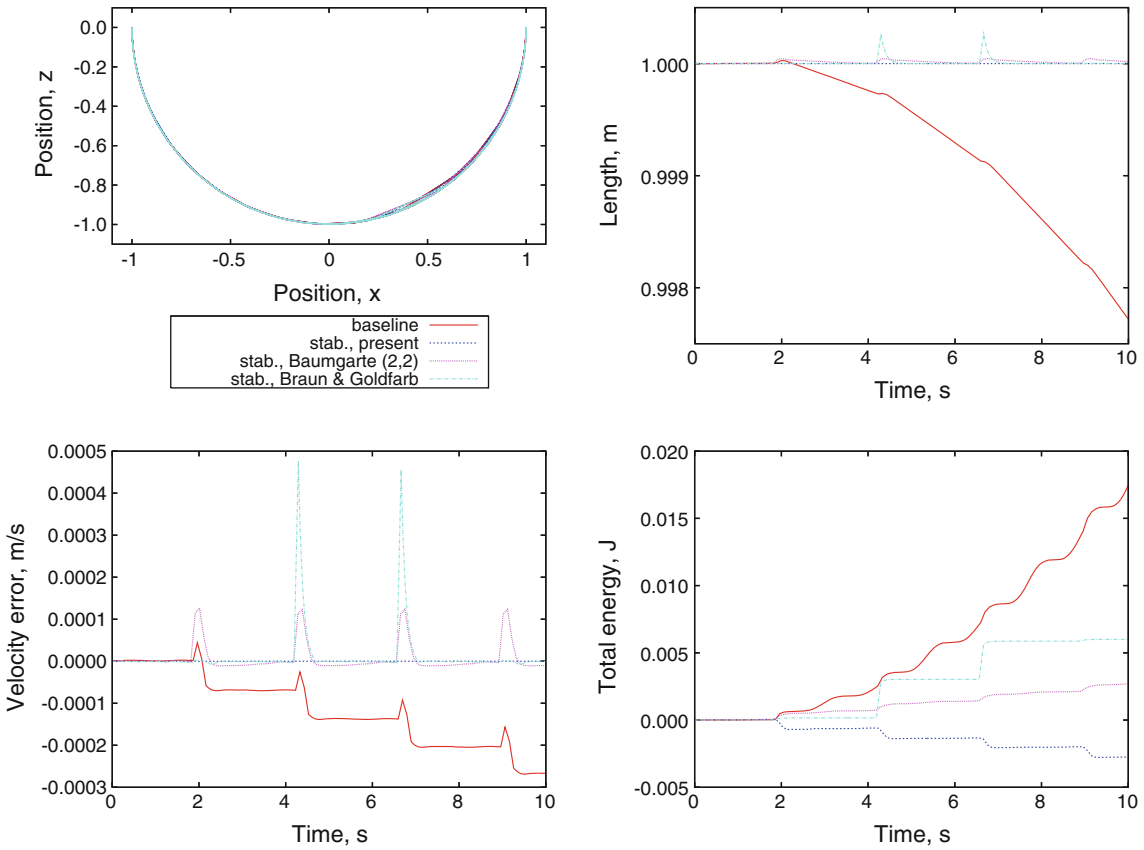


**Fig. 1** Pendulum: trajectory (*top left*), length (*top right*), velocity error (*bottom left*) and total energy (*bottom right*); relative tolerance $10^{-3}$, absolute tolerance $10^{-6}$

Figures 1 and 2 show the results obtained by numerical integration with different tolerances when the pendulum starts in horizontal position with no initial velocity. The solution indicated as 'baseline' is obtained by integrating the problem in the form of Eq. (8), rewritten in first-order form, using an explicit Runge–Kutta integration scheme with 5th order truncation error [34], implemented in the `ode45` function provided by Octave [35]. This solution is of little practical use; it is presented to highlight how constraint stabilization techniques based on penalty methods make the problem comparatively stiffer. The solution indicated as 'present' is obtained using the outlined procedure, integrated using the above mentioned integration algorithm. The solution indicated as 'Baumgarte (2,2)' in the figures uses Baumgarte's correction [21], with coefficients $\alpha_b = 2$ and $\beta_b = 2$, using the same integration algorithm. The solution indicated as 'Braun and Goldfarb' uses the correction proposed in [30].

Since this work is not dedicated to the discussion of how to tune the parameters of the Baumgarte algorithm, results obtained with different combinations of $\alpha_b$, $\beta_b$ are not presented. However, it is worth noticing that larger values of $\beta_b$, with $\alpha_b = \beta_b$, yield a faster transient, and thus a smaller error on the position constraint, although with a locally larger error in the velocity constraint. In those cases, a much shorter time step is required, as the problem becomes stiffer.

The problem is integrated over 10 s, with initial conditions $\{x, z\} = \{1, 0\}$. The results of Fig. 1 have been integrated with a relative tolerance of $10^{-3}$ and an absolute tolerance of $10^{-6}$ (the default), while those of Fig. 2 used a relative and an absolute tolerance of $10^{-9}$. A maximum time step of 0.1 s and an initial time step of 0.01 s were used; this required about 200 and 850 time steps respectively with the looser and the tighter tolerance. The only notable exception is that of Braun and Goldfarb's method. It behaves much like Baumgarte's, but the choice of $\alpha_b = \beta_b = 1/\mathrm{d}t$ results in a much stiffer problem than in the case of Baumgarte's with $\alpha_b = \beta_b = 2$. As a consequence, as soon as the tolerance is tightened, the solution is much closer to that of the proposed method, but at the cost of nearly twice the time steps. A similar result would be obtained using Baumgarte's method with larger gains.

The results clearly show that the proposed approach complies with the constraint with the accuracy of the inner nonlinear algebraic loop required to enforce the position constraint ($10^{-12}$ in this case). The total energy is not preserved (nor any provisions to enforce this property are put in place). However, its change is limited (as a reference, the potential

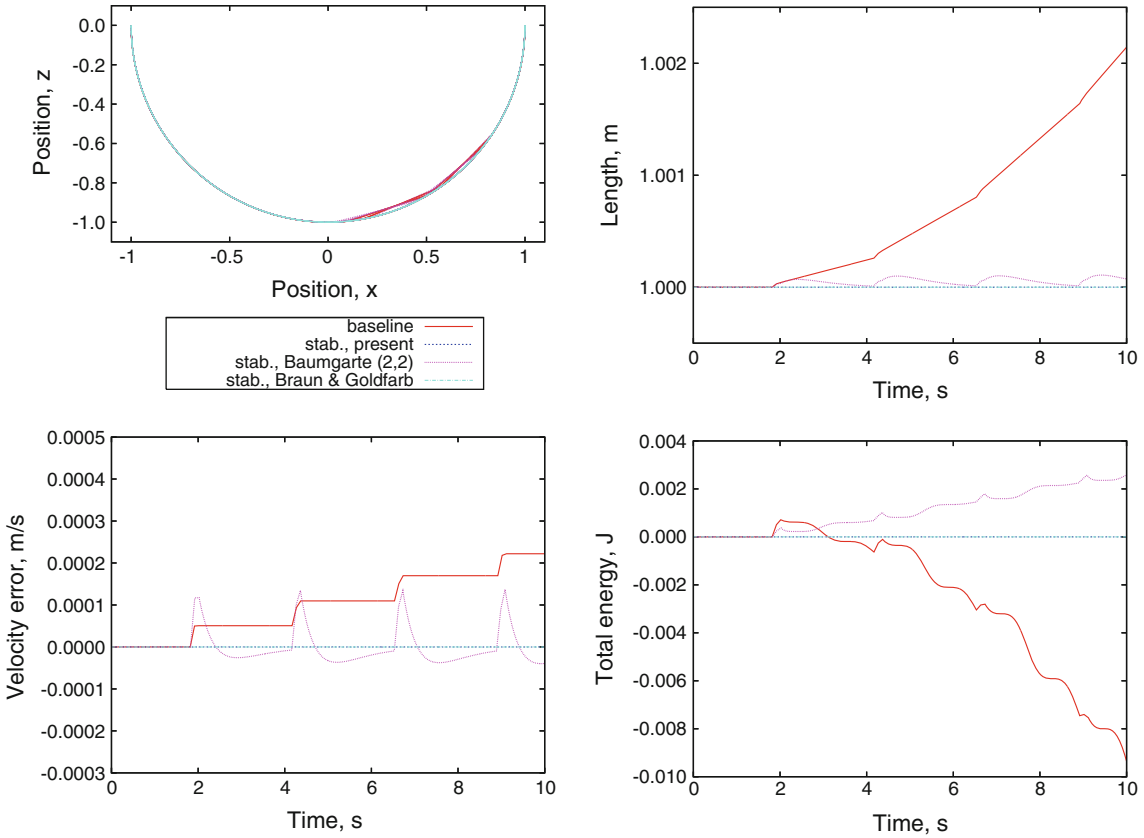


**Fig. 2** Pendulum: trajectory (*top left*), length (*top right*), velocity error (*bottom left*) and total energy (*bottom right*); relative tolerance $10^{-9}$, absolute tolerance $10^{-9}$

and the kinetic energy change by 9.81 J during a period, so the reduction is less than 0.01% per period), and happens to consist in a reduction, as opposed to the other methods. Note that only the proposed stabilization results in essentially preserving the total energy, provided the tolerance is strict enough.

Of course, no general rule should be inferred from this specific case, as nothing in the method enforces neither the preservation nor the dissipation of the total energy. If the tolerance on the error is tightened, the error on the energy is not significantly affected when the drift is either ignored (baseline) or compensated using Baumgarte's method, while it essentially vanishes using the proposed approach.

### 4.2 The swing problem

This problem, proposed in [36], consists in a mix of rigid and deformable elements, the latter modeled as beams, with significant numerical stiffness related to both the presence of kinematic constraints and a variety of dynamics ranging from the very low bending modes to the extremely high axial modes of the beams.

The model is sketched in Fig. 3. The beam is made of aluminum alloy ($E = 73$ MPa, $\nu = 0.3$, $\rho = 2700$ kg/m$^3$). The section is $1 \times 5$ mm$^2$. The length is 0.72 m. It carries a 0.5 kg mass lumped at mid span, in point C. It is pinned to two rigid links with negligible mass, one pinned to the ground on the vertical of point A, and the other one pinned to the ground along a line crossing point E and rotated 45° to the right, respectively 0.36 and $0.36\sqrt{2}$ m long.

The system is excited by a triangular force that starts at time $t = 0$ s, grows to 2 N at time $t = 0.128$ s, and returns to 0 at $t = 0.256$ s. At some point, between 0.6 and 0.7 s, the first link and the beam are nearly aligned. As such, the beam is subjected to an abrupt excitation, since the leftmost end of the beam receives a nearly impulsive excitation that triggers its very high frequency axial modes. What follows is close to chaos, as refining the mesh or reducing the time step would respectively add higher frequency dynamics or integrate it more accurately.

This problem can be very critical for numerical integration schemes. The proposed approach is perhaps one of the least suitable, like any other based on the explicit integration of ODEs, when a finite element model of the beam is considered, because the high frequency axial modes will dictate the time step that needs to be used by explicit integration schemes, without appreciably contributing to the dynamics of the problem. However, it is presented to highlight how the

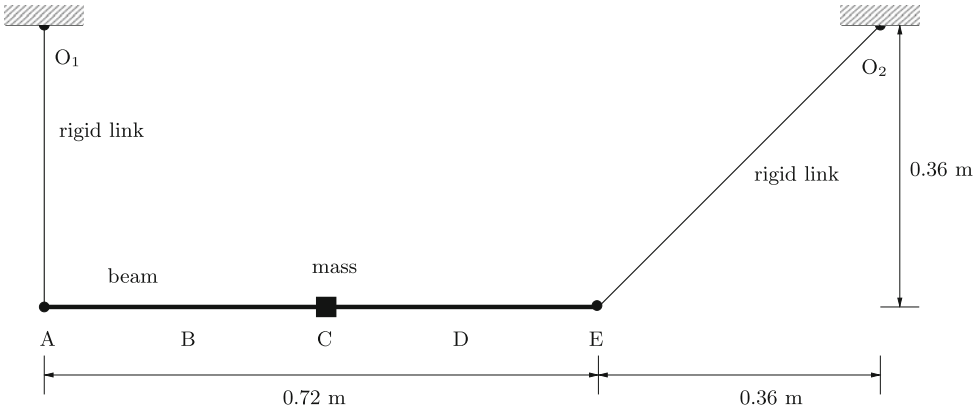


**Fig. 3** Sketch of the swing model (from [36])

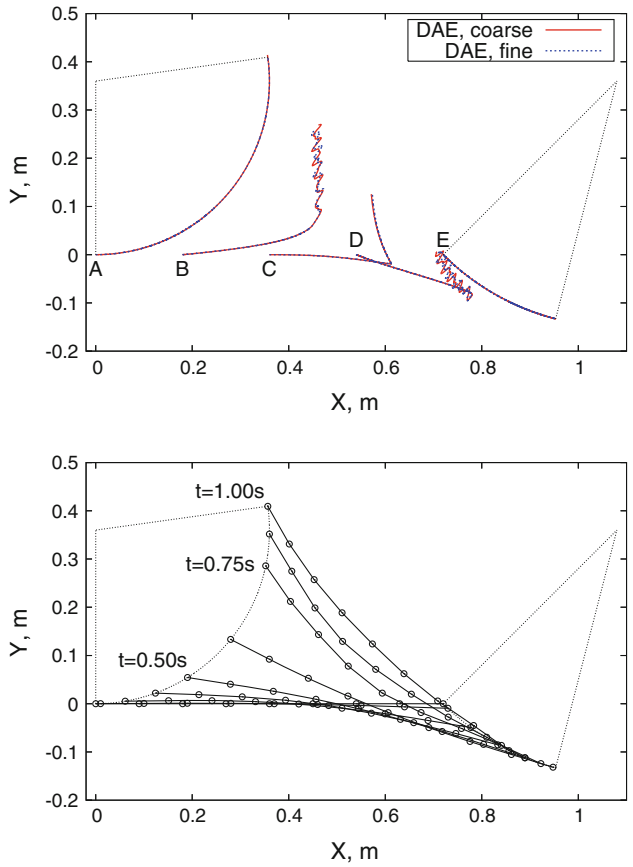


**Fig. 4** Swing: pictorial views of the trajectory of marker points A, B, C, D, and E

proposed approach can be at least marginally beneficial also in cases where explicit time integration is not appropriate.

The reference analysis has been performed using a model based on implicit integration of the problem in differential algebraic form, using an original finite volume implementation of geometrically exact beam elements [37]. Figure 4 illustrates the trajectory of points A–E integrated with the free general-purpose multibody solver MBDyn (http://www.mbdyn.org/), which solves the DAE problem using an implicit integration scheme. The figure shows how

**Table 1** Swing: eigenanalysis

| Mode | Frequency (Hz) |
|---|---|
| First symmetric bending | 0.443 |
| First antisymmetric bending | 18.17 |
| Second symmetric bending | 28.46 |
| … | |
| First antisymmetric axial | 319.5 |

a mesh refinement, from 2 to 4 three-node beam elements, results in a motion that is split into two parts: the regular one, prior to the nearly-impulsive loading, that is insensitive to the discretization, and the chaotic part, where mesh refinement leads to different behavior of marker points B and D. Marker points A and E are minimally affected, as they are constrained to move along a circular path. The motion of C is "regularized" by the lumped mass.

A previously unpublished two-node variant of the formulation proposed in [37] has been implemented to solve the problem using explicit integration with the proposed procedure. It is detailed in Appendix.

In order to gain further insight into the dynamic properties of the system, and in view of its analysis using a Ritz model, the eigenanalysis of a refined FE model of the simply supported beam, modeled with 40 two-node beam elements, is reported in Table 1, while Fig. 5 presents the corresponding mode shapes (the transverse displacement of bending modes, and the axial displacement of the axial mode). All displacement components of the two ends have been disallowed, so that mode shapes do not affect the motion of the constraints with the rigid links.

The shape of the first antisymmetric axial mode is essentially a triangle, because the lumped mass dominates over the distributed inertia of the beam, so the behavior of the beam is essentially static. Similarly, the first symmetric bending mode is basically the cubic that describes the static response to a lumped transverse load at mid-span. The axial mode is expected to dominate the transient when the mass abruptly reverts its motion, while the bending modes are expected to

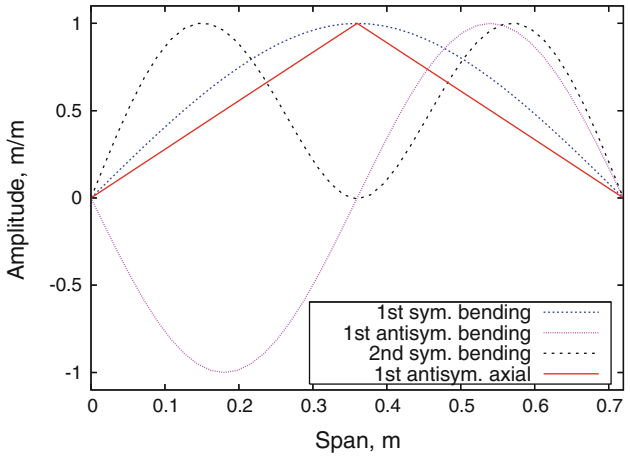


**Fig. 5** Swing: shapes of relevant modes

dominate the free vibrations of the beam, superimposed to its rigid motion and essentially unaffected by the forces exerted by the constraints.

The following analyses have been performed:

(0) a reference analysis of a FE model, using MBDyn; this model consists of 2 three-node beam elements; a 4 three-node model has been considered as well in Fig. 4;
(1a) a baseline analysis of a FE model, using Eq. (8) rewritten in first-order form; this model consists of 4 two-node beam elements;
(1b) an analysis using the same FE model, with the proposed corrections;
(2a) a baseline analysis of a modal model, consisting of the 4 modes described in Table 1, using Eq. (8) rewritten in first-order form;
(2b) an analysis of the same modal model, with the proposed corrections;
(3a) a simplified analysis using the same model but omitting the high-frequency axial mode;
(3b) an analysis of the same simplified modal model, with the proposed corrections.

All models have been integrated using the same explicit Runge–Kutta algorithm mentioned earlier, with more relaxed tolerances.

Table 2 shows how the modal models performed much better than the FE ones, since they selectively allowed to include in the analysis only the desired dynamics.

It is worth stressing that the FE models could not converge after the nearly-impulsive event at about 0.65 s without excessively reducing the time step. As a consequence, the analysis had to be stopped. However, those numbers provide some important insight into the behavior of the proposed correction.

**Table 2** Swing: computation related information

| Case | Last time (s) | Steps |
|---|---|---|
| (1a) Beam FE, baseline | 0.69 | 69,033 |
| (1b) Beam FE, present | 0.69 | 69,033 |
| (2a) Modal, baseline | 1.00 | 23,075 |
| (2b) Modal, present | 1.00 | 28,103 |
| (3a) Modal, baseline (no axial mode) | 1.00 | 1,103 |
| (3b) Modal, present (no axial mode) | 1.00 | 1,103 |

When considering the FE analysis, the proposed algorithm requires exactly the same number of time steps of the baseline case. Figure 6 shows that the error between the baseline and the present analysis is negligible. This occurs because the time step required to fall within the region of stability of the integration scheme applied to this very stiff problem yields enough accuracy to limit the drift until stiffness actually comes into play, at $t \cong 0.67$ s. The order of magnitude of the average time step drops to $10^{-5}$, and a much shorter one would probably be required to complete the simulation. Note that the largest natural frequency of the free-free model of the beam used in the coarse FE model analysis is of the order of $10^4$ Hz.

On the contrary, when considering the modal model, the proposed algorithm requires only little more time steps than the baseline one. This is due to the fact that enforcing the algebraic constraints with reduced problem stiffness improves the solution process, or at least does not disturb it. The order of magnitude of the average time step is close to $10^{-4}$. Note that the violation of the constraint and of its derivative resulting from the integration of the modal model appear to be much larger than the corresponding ones obtained by integrating the beam FE model, as one can infer by comparing Figs. 6 and 7.

Finally, note that removing the only relatively stiff mode, the axial one, results in basically hitting the requested upper time step limit of $10^{-3}$ most of the times, since the problem is no longer stiff. The differences in the motion when the axial mode is omitted are negligible, as illustrated in Fig. 8.

Figure 8 allows to capture how the proposed approach performs. The direct beam FE model fails to converge as soon as the stiffness of the problem dominates the response, regardless of constraint stabilization. The modal model, with the proposed correction, does not suffer from the severe drift that characterizes the baseline integration. Even when the axial mode is removed, the proposed stabilization performs relatively well.

### 4.3 Helicopter blade pitch control

This application shows the opportunity of using a kinematic constraint to replace a mechanical component whose high

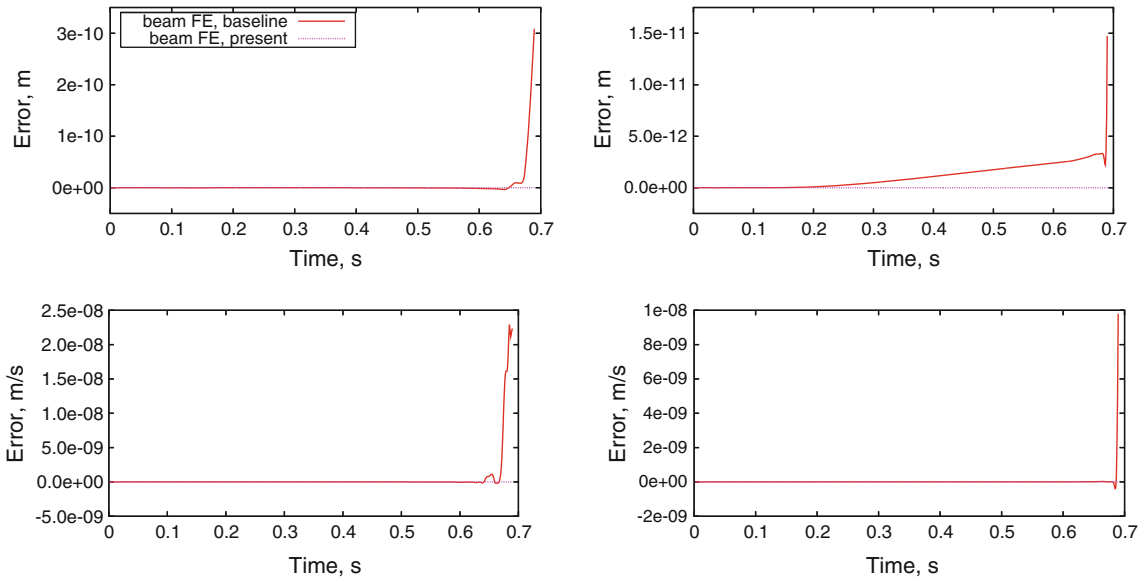


**Fig. 6** Swing: beam FE model; constraint (*top*), constraint derivative (*bottom*) error

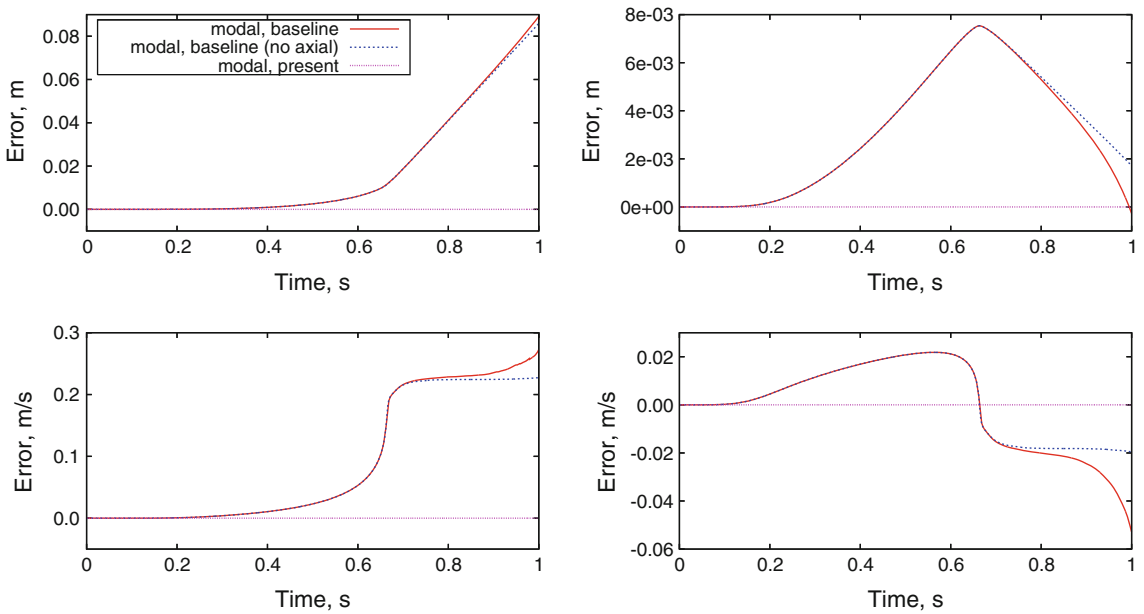


**Fig. 7** Swing: modal model; constraint (*top*), constraint derivative (*bottom*) error

stiffness would otherwise drive the stability of the analysis without significantly contributing to its accuracy, if not deteriorating it.

Figure 9 shows a sketch of the problem, which is a simplification of the blade root articulation and pitch control of an articulated helicopter rotor.

The location of the upper end of the pitch link, point A, depends on the flap angle, $\beta$, and on the pitch angle, $\theta$; according to the figure,

$$\mathbf{x}_{\mathrm{A}} = \begin{Bmatrix} x_\beta \\ 0 \\ 0 \end{Bmatrix} + \begin{bmatrix} \cos\beta & 0 & -\sin\beta \\ 0 & 1 & 0 \\ \sin\beta & 0 & \cos\beta \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix} \times \begin{Bmatrix} x_\theta \\ y_\theta \\ 0 \end{Bmatrix} = \mathbf{x}_\beta + \mathbf{R}_\beta \mathbf{R}_\theta \mathbf{x}_\theta . \tag{28}$$

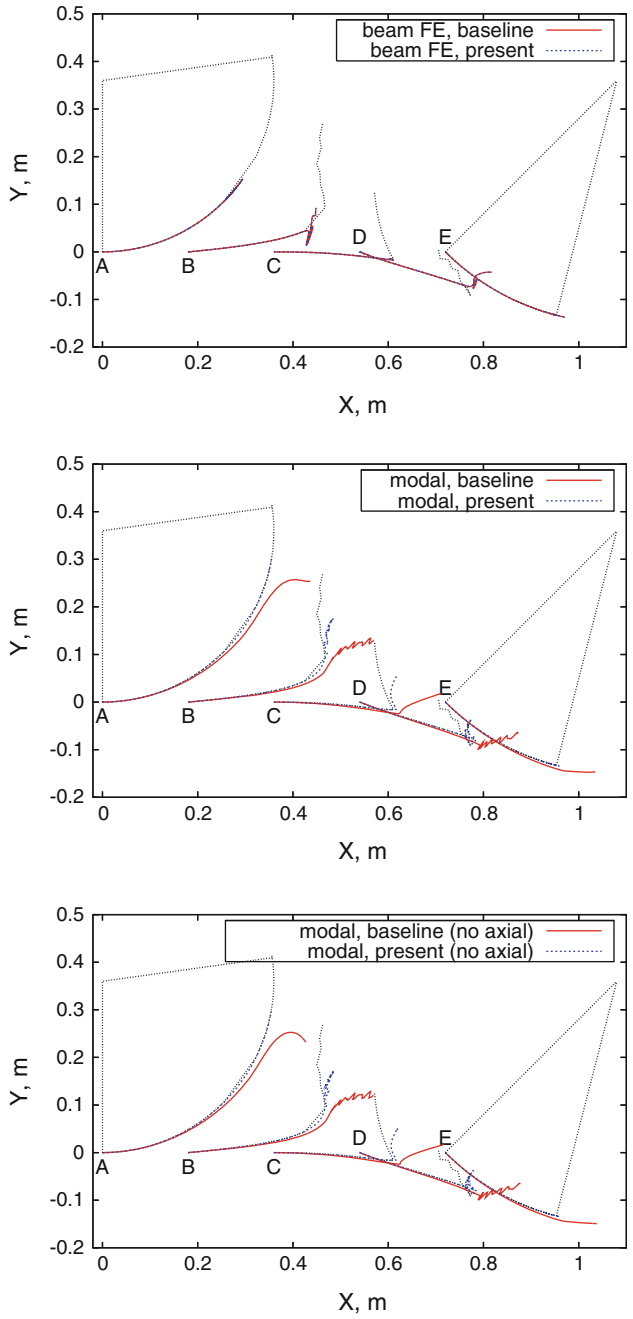


**Fig. 8** Swing: motion of markers A, B, C, D, and E; beam FE (*top*) and modal (*mid, bottom*) models, the last without the axial mode

The location of the lower end, point B, is fixed on the swashplate.

The pitch link is usually quite stiff. Some compliance may be present because of the ball joints that connect it to the blade in A and to the swashplate in B. As a consequence, modeling it as a deformable body may result in very stiff equations.

In usual helicopter rotors, the natural flapping frequency of the blade is close to the rotating speed $\Omega$. The natural pitch frequency of an unconstrained blade would be close to $\Omega$ as well. However, when constrained by the pitch link, the blade pitch motion is no longer free, but rather kinematically determined.

The first torsional frequency of a blade connected to the control system is essentially dominated by the torsion of the blade. By design, it is usually higher than $b\Omega$, where $b$ is the number of blades. When the deformability of the pitch link is considered, this frequency reduces a little bit, but in many cases the deformability of the pitch link has minimal impact on the low-frequency dynamics of the blade. As a consequence, a model designed to address very low-frequency dynamics would be negatively impacted by the need to model a very stiff pitch link.

To model a rigid pitch link, one needs to enforce the constraint

$$\sqrt{(\mathbf{x}_\mathrm{A}-\mathbf{x}_\mathrm{B})^T(\mathbf{x}_\mathrm{A}-\mathbf{x}_\mathrm{B})}-L=0; \tag{29}$$

$L$ is the length of the pitch link; $\mathbf{x}_\mathrm{B}=\{x_\mathrm{B};y_\mathrm{B};z_\mathrm{B}\}$ is the location of point B in a frame rotating with the hub. Enforcing Eq. (29) may not be trivial, depending on the relationship between the location of points A and B and the coordinates $\beta$ and $\theta$ of the blade.

The effect of applying a pitch control to the helicopter can be approximated as a constant plus a 1/rev periodic change of $z_\mathrm{B}$ (actually, by tilting the swashplate, all the components of $\mathbf{x}_\mathrm{B}$ change). At the same time, the blade flaps and thus participates in the kinematic relationship of Eq. (29).

To illustrate the effectiveness and the minimal intrusiveness of the proposed approach, the independent flap and pitch dynamics of a rigid rotor blade, governed by two mildly coupled second-order ODEs, is constrained using Eq. (29).

According to Johnson [38], the rigid flap and twist dynamics of a rotating blade, linearized about a steady trim condition, are described by

$$I_b\left(\ddot{\beta}+\Omega^2\nu_\beta^2\beta\right)-I_x\left(\ddot{\theta}+\Omega^2\theta\right)=M_\beta \tag{30a}$$

$$I_f\left(\ddot{\theta}+\Omega^2\theta\right)-I_x\left(\ddot{\beta}+\Omega^2\beta\right)=M_\theta +K_c\left(\theta_c-\theta-K_p\beta\right), \tag{30b}$$

where meaning and value of the symbols are illustrated in Table 3, and:

- equations are dimensional, as opposed to [38];
- the aerodynamic moments about flap, $M_\beta$, and pitch, $M_\theta$, are not made explicit but rather replaced by generic periodic forcing terms,

$$M_\beta=-\frac{\gamma}{16}\Omega I_b\dot{\beta}+M_{\beta_0}+M_{\beta_1}\cos\left(\Omega t+\varphi_{\beta_1}\right)$$

$$M_\theta=M_{\theta_0}+M_{\theta_1}\cos\left(\Omega t+\varphi_{\theta_1}\right).$$

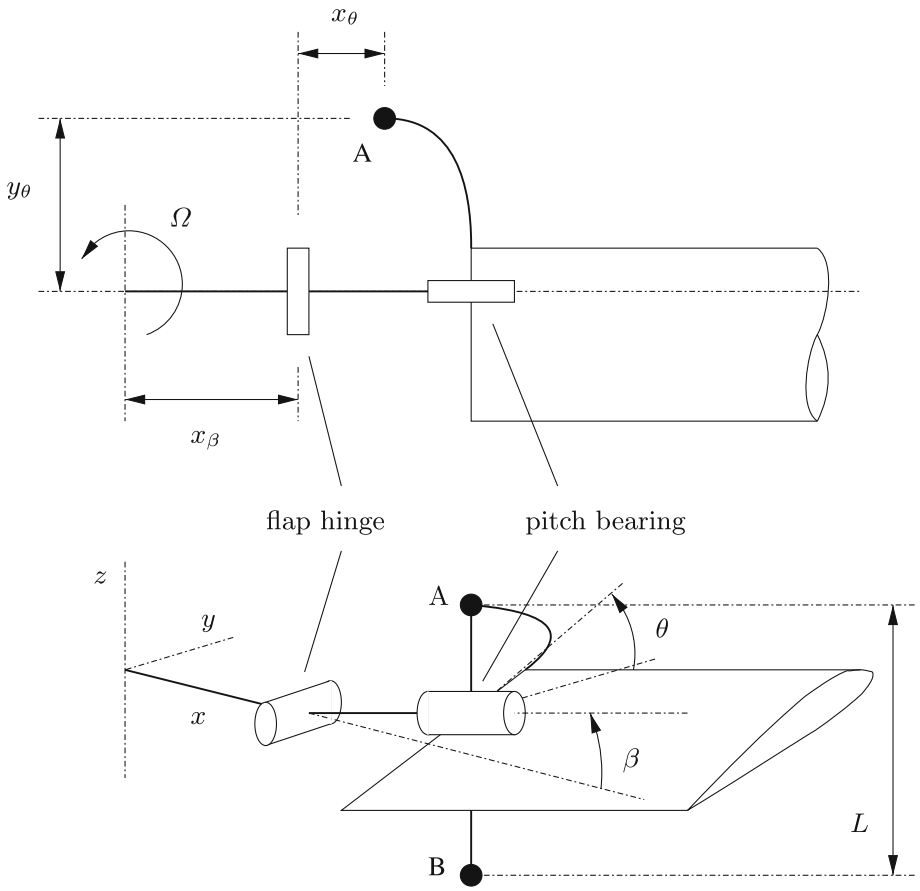


**Fig. 9** Helicopter blade pitch control system sketch

The flap moment includes some significant damping of the order of $(-\gamma \Omega I_b/16)$, where $\gamma$ is the Lock number, a ratio between aerodynamic and inertial forces involved in blade flapping. Coupling between flap and pitch through the aerodynamics is neglected;

– the control stiffness $K_c$ is omitted when the control system is modeled by the algebraic constraint.

The system is integrated for 2 s, subjected to a periodic flapping moment $M_{\beta_1} = 10$ kN m, $\varphi_{\beta_1} = \pi/2$, considering:

1. *baseline*: a model based on Eqs. (8), with $K_c = 0$;
2. *present*: a model based on the proposed approach, with $K_c = 0$;
3. *penalty*: a model with $K_c$ set to the value of Table 3, corresponding to a blade pitch frequency of 10.5 Ω, and without any kinematic constraints projection.

All models have been integrated using the explicit Runge–Kutta scheme mentioned earlier, with relative tolerance set to $10^{-2}$, absolute tolerance set to $10^{-4}$, and a maximum allowed time step of 0.01 s. The resulting motion is shown in Fig. 10.

It is not surprising that the first two cases, that enforce a kinematic constraint, ran the whole analysis at the maximum allowed time step, thus requiring 200 time steps. However, it is interesting to notice that the last case required nearly 700 steps to get to a solution that is qualitatively comparable to that of the first two cases. If the stiffness of the control system is increased further, the analysis does not converge.

The 'penalty' case does not drift, while the 'baseline' one does, as shown in Fig. 11 on the right. However, the 'penalty' case oscillates significantly about the expected solution, as shown in the same figure on the left.

Figure 12 shows the axial force in the pitch link. The 'penalty' case is dominated by the high-frequency torsional mode, and thus is not representative of the actual behavior of the system. The cases with kinematic constraint enforcement, on the contrary, clearly show how the response of the system at a frequency close to resonance is basically counteracted by the damping in the blade flapping. As a consequence, the pitch link is essentially unloaded, except for a small second harmonic. This is entirely related to the nonlinear kinematics

**Table 3** Helicopter blade and pitch control properties [38]

| Property | Value |
|---|---|
| Rotor angular velocity, $\Omega$ | 40.0 (radian/s) |
| Rotor radius, $R$ | 5.0 (m) |
| Blade flap inertia, $I_b$ | 416.67, (kg m$^2$) |
| Blade twist inertia, $I_f$ | $9.375 \times 10^{-2}$ (kg m$^2$) |
| Blade centrifugal inertia, $I_x$ | $1.875 \times 10^{-2}$ (kg m$^2$) |
| Lock number, $\gamma$ | 8.0 (adim.) |
| Flap hinge radial offset, $x_\beta$ | 0.2 (m) |
| Point A radial offset, $x_\theta$ | 0.05 (m) |
| Point A chordwise offset, $y_\theta$ | 0.2 (m) |
| Point B radial offset, $x_\mathrm{B}$ | 0.25 (m) |
| Point B chordwise offset, $y_\mathrm{B}$ | 0.2 (m) |
| Point B out-of-plane off., $z_\mathrm{B}$ | $-0.15$ (m) |
| Pitch link length, $L$ | 0.15 (m) |
| Blade flap freq. ratio[a], $\nu_\beta$ | 1.0625 (adim.) |
| Pitch-flap coupling[b,c], $K_p$ | 0.25 (adim.) |
| Control system stiff.[c], $K_c$ | $1.6538 \times 10^4$ (N m) |

[a] Computed: $\nu_\beta = \sqrt{1 + \frac{3}{2}\frac{x_\beta}{R - x_\beta}}$

[b] Computed: $K_p = \frac{x_\theta}{y_\theta}$

[c] Only used for analysis without kinematic constraints

of the constraint, since the uncoupled flap and twist equations are linear. Note that the multiplier $\mu$ is essentially zero; nonetheless, its presence is instrumental in eliminating drift from the motion and the reaction force.

If the stiffness of the control system is reduced (e.g. to 5.5 $\Omega$, a reasonable value for the first torsional frequency of a helicopter blade), the angular velocity oscillations (Fig. 10, right) can be much less pronounced, but at the cost of a much larger violation of the constraint, although oscillating at nearly half the frequency.

## 5 Conclusions

This work discussed how kinematic constraints can be proficiently enforced in problems described by ODEs, without altering the nature or the number of equations and coordinates of the problem. Their introduction requires minimal implementation overhead: only the possibility to manipulate the (tentative) solution and the right-hand side before the highest order derivatives are computed is required.

The resulting problem is no longer explicit when holonomic constraints are considered. However, to comply with the constraints, the projection of the positions on the invariant manifold can be performed separately by computing the

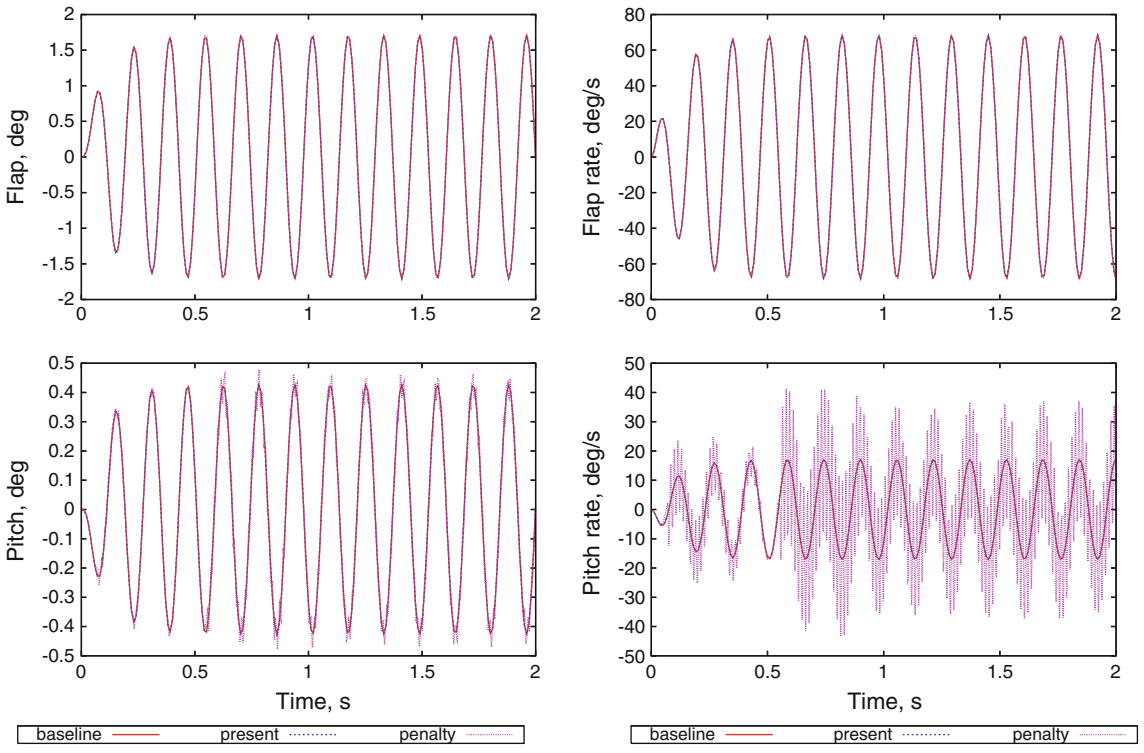


**Fig. 10** Helicopter blade motion: flap (*top*) and pitch (*bottom*); angles (*left*) and angular velocities (*right*)

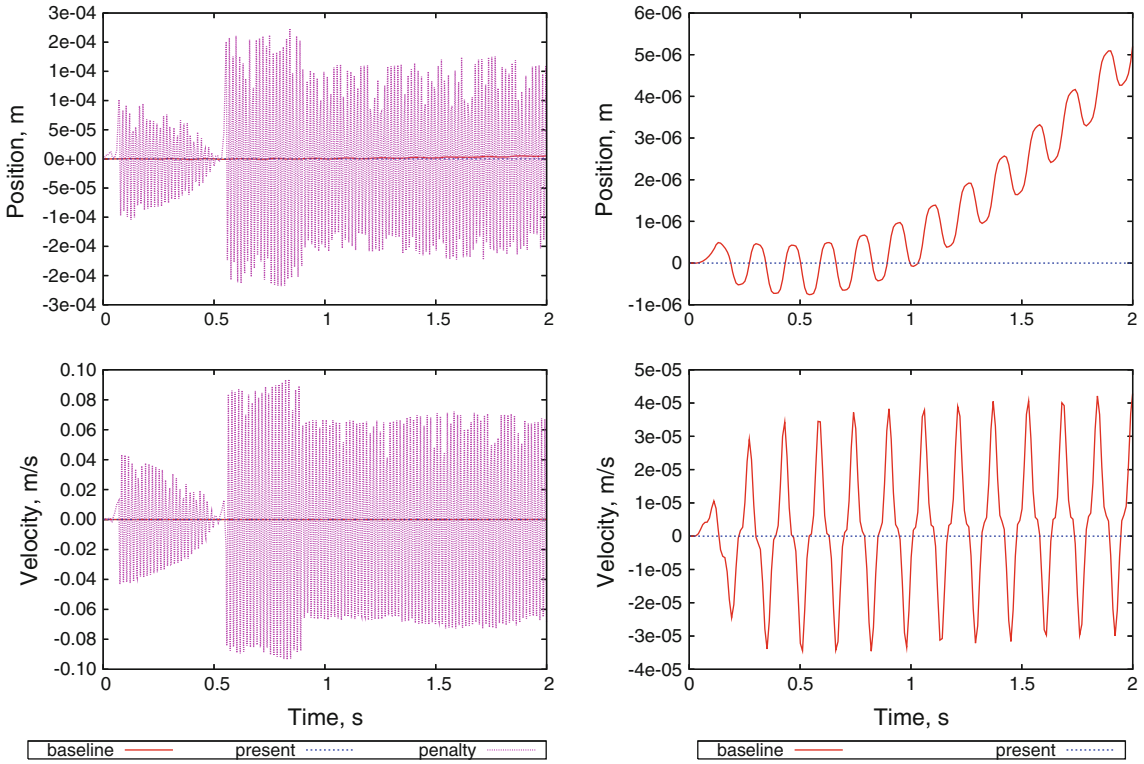


**Fig. 11** Helicopter blade error: position (*top*) and velocity (*bottom*); 'penalty' omitted on the *right*

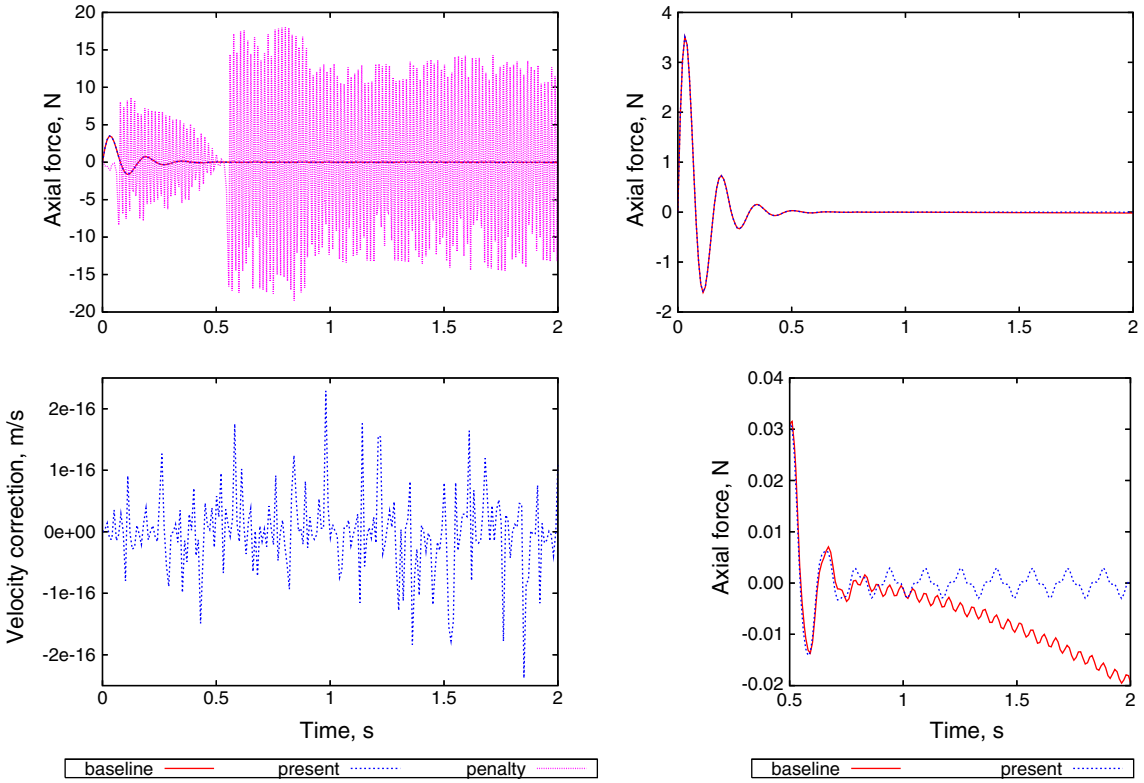


**Fig. 12** Helicopter blade multipliers: pitch link axial force (*top left, top and bottom right*), velocity correction (*bottom left*)

solution of a nonlinear algebraic problem. The original problem remains explicit.

## Appendix

### Finite-volume beam model

The Finite-Volume beam model proposed in [37] is here summarized and specialized to the case of a two-node beam element in order to make this work as self-contained as possible. This formulation has never been published before, although being available since 2001 in the free multibody software MBDyn.

As illustrated in [37], the differential equilibrium equations of a beam can be integrated using piecewise constant weight functions. This corresponds to cutting the beam in finite portions, centered at the nodes, and writing the equilibrium of each chunk about the node taking into account the internal force and moment at the interfaces between chunks. Considering the problem element-wise, namely the portion of beam between two nodes, each element is divided in two half-chunks, each contributing to the equilibrium of its respective node. The two half-chunks exchange the same internal force and moment, which are applied with opposite sign to each node, and evaluated at a point halfway between the two nodes.

The forces and moments at nodes $n = 1, 2$, located in $\mathbf{x}_n$, as functions of the internal force $\mathbf{f}_I$ and moment $\mathbf{m}_I$ at mid-point $\mathbf{x}_I$, are

$$\mathbf{f}_n = (-1)^n \mathbf{f}_I \tag{31a}$$

$$\mathbf{m}_n = (-1)^n \left( (\mathbf{x}_I - \mathbf{x}_n) \times \mathbf{f}_I + \mathbf{m}_I \right). \tag{31b}$$

The internal force and moment are formulated in the reference frame of the beam section, $\mathbf{R}_I$, as functions of the linear and angular strains, $\tilde{\boldsymbol{\varepsilon}}_I$ and $\tilde{\boldsymbol{\kappa}}_I$, as

$$\mathbf{f}_I = \mathbf{R}_I \tilde{\mathbf{f}} \left( \tilde{\boldsymbol{\varepsilon}}, \tilde{\boldsymbol{\kappa}} \right) \tag{32a}$$

$$\mathbf{m}_I = \mathbf{R}_I \tilde{\mathbf{m}} \left( \tilde{\boldsymbol{\varepsilon}}, \tilde{\boldsymbol{\kappa}} \right), \tag{32b}$$

based on the desired constitutive properties.

The linear and angular strains represent a measure of the straining of the beam when its configuration changes from a reference (subscript '0') to an arbitrary one (no subscript),

$$\tilde{\boldsymbol{\varepsilon}} = \mathbf{R}^T \mathbf{x}_{/\eta} - \mathbf{R}_0^T \mathbf{x}_{0/\eta} \tag{33a}$$

$$\tilde{\boldsymbol{\kappa}} = \mathbf{R}^T \boldsymbol{\rho}_\eta - \mathbf{R}_0^T \boldsymbol{\rho}_{0\eta}, \tag{33b}$$

where $\eta$ is a curvilinear abscissa along the beam, and $\boldsymbol{\rho}_\eta = \mathrm{ax}(\mathbf{R}_{/\eta}\mathbf{R}^T)$ is the beam curvature.[1]

They can be expressed as functions of the kinematics of the nodes by separately interpolating the positions and the orientations with linear shape functions $N_n(\xi) = (1+(-1)^n\xi)/2$, $\xi \in [-1, 1]$, namely

$$\mathbf{x}(\xi) = \sum_{n=1,2} N_n \mathbf{x}_n. \tag{34}$$

Differentiation with respect to $\eta$ yields

$$(\cdot)_{/\eta} = (\cdot)_{/\xi} \frac{\mathrm{d}\xi}{\mathrm{d}\eta}, \tag{35}$$

with $\mathrm{d}\eta/\mathrm{d}\xi = (\mathbf{x}_{/\xi} \cdot \mathbf{x}_{/\xi})^{1/2}$, and $\mathbf{x}_{/\xi} = \sum_{n=1,2} N_{n/\xi} \mathbf{x}_n$.

The interpolation of the orientation deserves special care. In fact, the orientation matrix cannot be directly interpolated. However, one can exploit the helicoidal strain assumption [39], mutuated from the screw theory used to describe rigid body dynamics, and interpolate the orientation as

$$\begin{aligned} \mathbf{R}(\xi) &= \mathbf{R}_1 \exp\left( (N_2(\xi)\boldsymbol{\theta}) \times \right) \\ &= \mathbf{R}_2 \exp\left( (N_1(\xi)\boldsymbol{\theta}) \times \right), \end{aligned} \tag{36}$$

with

$$\boldsymbol{\theta} = \mathrm{ax}\left( \exp^{-1}\left( \mathbf{R}_1^T \mathbf{R}_2 \right) \right). \tag{37}$$

The rationale is that any orientation intermediate between $\mathbf{R}_1$ and $\mathbf{R}_2$ is expressed as $\mathbf{R}_1$ itself, multiplied by the corresponding fraction $N_2(\xi)$ of the relative orientation between the nodes, expressed by the Euler vector $\boldsymbol{\theta}$, thus yielding a helix with pitch $\|\mathbf{x}_2 - \mathbf{x}_1\| / \|\boldsymbol{\theta}\|$.

The optimal location of $\mathbf{x}_I$ that minimizes shear locking (an underestimation of the compliance to transverse loads related to overintegration of the shear strain) is $\xi = 0$, corresponding to mid-span. Slender beams exhibit some residual shear locking which can be eliminated, when considering a linear elastic constitutive law, by correcting the transverse shear stiffness to account for the slenderness of the beam element.

[1] The ax $(\cdot)$ operator is the inverse of the $(\cdot)\times$ operator: ax $(\mathbf{z}\times) = \mathbf{z}$.

## References


1. Laulusa A, Bauchau OA (2008) Review of classical approaches for constraint enforcement in multibody systems. J Comput Nonlinear Dyn 3(1). doi:10.1115/1.2803257
2. Bauchau OA, Laulusa A (2008) Review of contemporary approaches for constraint enforcement in multibody systems. J Comput Nonlinear Dyn 3(1). doi:10.1115/1.2803258
3. Simeon B (1995) MBSPACK—numerical integration software for constrained mechanical motion. Surv Math Ind 5:169–202
4. Schiehlen W (1997) Multibody system dynamics: roots and perspectives. Multibody Syst Dyn 1(2):149–188 (1997). doi:10.1023/A:1009745432698
5. Gauss CF (1829) Ueber ein neues algemeines Grundgesetz der Mechanik. J fuer die reine und angewandte Mathematik 4: 232–235 (in German)
6. Hertz H (1894) Die Prinzipien der Mechanik. Gesammelte Werke 3 (in German)
7. Gibbs JW (1879) On the fundamental formulas of dynamics. Am J Math 2:49–64

8. Appell P (1899) Sur les mouvements de roulment; equations du mouvement analougues a celles de Lagrange. Comptes Rendus 129:317–320 (in French)
9. Maggi GA (1903) Principii di stereodinamica: Corso sulla formazione, l'interpretazione e l'integrazione delle equazioni del movimento dei solidi. Hoepli, Milano (in Italian)
10. Kane TR, Wang CF (1965) On the derivation of equations of motion. J Soc Ind Appl Math 13(2):487–492
11. Eberhard P, Schiehlen W (2006) Computational dynamics of multibody systems: History, formalisms, and applications. J Comput Nonlinear Dyn 1:3–12. doi:10.1115/1.1961875
12. Brenan KE, La Vern Campbell S, Petzold LR (1989) Numerical solution of initial-value problems in differential-algebraic equations. North-Holland, New York
13. Jourdain PEB (1909) Note on an analogue of Gauss principle of least constraint. Q J Pure Appl Math 40:153–197
14. Pennestrì E, Vita L (2004) Strategies for the numerical integration of DAE systems in multibody dynamics. Comput Appl Eng Educ 12(2):106–116. doi:10.1002/cae.20005
15. Eich E (1993) Convergence results for a coordinate projection method applied to mechanical systems with algebraic constraints. SIAM J Numer Anal 30(5):1467–1482. doi:10.1137/0730076
16. Lubich Ch, Engstler Ch, Nowak U, Pöhle U (1995) Numerical integration of constrained mechanical systems using MEXX. Mech Based Des Struct Mach 23(4):473–495. doi:10.1080/08905459508905248
17. Lilov L, Lorer M (1982) Dynamic analysis of multirigid-body systems based on Gauss principle. Z Angew Math Mech 62:539–545. doi:10.1002/zamm.19820621006
18. Udwadia FE, Kalaba RE (1996) Analytical dynamics. Cambridge University Press, New York
19. Neto MA, Ambrósio J (2003) Stabilization methods for the integration of dae in the presence of redundant constraints. Multibody Syst Dyn 10(1):81–105. doi:10.1023/A:1024567523268
20. Fumagalli A, Masarati P (2009) Efficient application of Gauss' principle to generic mechanical systems. Proc IMechE Part K: J Multi-body Dyn 223:121–131. doi:10.1243/14644193JMBD182
21. Baumgarte J (1972) Stabilization of constraints and integrals of motion in dynamical systems. Comput Methods Appl Mech 1:1–36. doi:10.1016/0045-7825(72)90018-7
22. Gear CW, Leimkuhler B, Gupta GK (1985) Automatic integration of Euler–Lagrange equations with constraints. J Comput Appl Math 12: 77–90. doi:10.1016/0377-0427(85)90008-1
23. Führer C, Leimkuhler BJ (1991) Numerical solution of differential-algebraic equations for constrained mechanical motion. Numer Math 59(1):55–69. doi:10.1007/BF01385770
24. Bayo E, García de Jalón J, Serna MA (1988) A modified Lagrangian formulation for the dynamic analysis of constrained mechanical systems. Comput Methods Applied Mechanics and Engineering, 71(2):183–195. doi:10.1016/0045-7825(88)90085-0
25. Mattsson SE, Söderlind G (1993) Index reduction in differential-algebraic equations using dummy derivatives. SIAM J Scientific Comput 14(3):677–692. doi:10.1137/0914043
26. Pantelides CC (1988) The consistent initialization of differential-algebraic systems. SIAM J Scientific Comput 9(2):213–231. doi:10.1137/0909014
27. Tiller M (2001) Introduction to physical modeling with modelica. Kluwer Academic Publishers, Dordrecht
28. Bauchau OA (2003) A self-stabilized algorithm for enforcing constraints in multibody systems. Int J Solids Struct 40:3253–3271. doi:10.1016/S0020-7683(03)00159-8
29. Borri M, Trainelli L, Croce A (2006) The embedded projection method: a general index reduction procedure for constrained system dynamics. Comput Methods Appl Mech Eng 195:6974–6992. doi:10.1016/j.cma.2005.03.010
30. Braun DJ, Goldfarb M (2009) Eliminating constraint drift in the numerical simulation of constrained dynamical systems. Comput Methods Appl Mech Engrg doi:10.1016/j.cma.2009.05.013
31. Hairer E, Wanner G (1996) Solving ordinary differential equations, vol II. 2nd rev. edn. Springer, Berlin
32. Ben-Israel A, Greville TNE (1974) Generalized inverses: theory and applications. Wiley, New York
33. Yu Q, Chen I-M (2000) A direct violation correction method in numerical simulation of constrained multibody systems. Comput Mech 26:52–57. doi:10.1007/s004660000149
34. Dormand JR, Prince PJ (1980) A family of embedded Runge–Kutta formulae. J Comput Appl Math 6(1):19–26. doi:10.1016/0771-050X(80)90013-3
35. Hindmarsh AC (1983) ODEPACK, a systematized collection of ODE solvers. In: Stepleman RS et al (eds) Scientific computing. North-Holland, Amsterdam, pp 55–64
36. Bauchau OA, Theron NJ (1996) Energy decaying scheme for nonlinear elastic multi-body systems. Comput Struct 59(2):317–331. doi:10.1016/0045-7949(95)00250-2
37. Ghiringhelli GL, Masarati P, Mantegazza P (2000) A multi-body implementation of finite volume beams. AIAA J 38(1):131–138. doi:10.2514/2.933
38. Johnson W (1980) Helicopter theory. Princeton University Press, Princeton
39. Borri M, Bottasso CL (1994) An intrinsic beam model based on a helicoidal approximation. Part I. Formulation. Int J Numer Methods Eng 37:2267–2289. doi:10.1002/nme.1620371308