
Multibody System Dynamics **7:** 265–284, 2002.

Elimination of Constraint Violation and Accuracy Aspects in Numerical Simulation of Multibody Systems

WOJCIECH BLAJER
Institute of Applied Mechanics, Technical University of Radom, ul. Krasickiego 54, 26-600 Radom, Poland; E-mail: wblajer@kiux.man.radom.pl

(Received: 22 November 1999; accepted in revised form: 26 October 2001)

Abstract. Multibody systems are often modeled as constrained systems, and the constraint equations are involved in the dynamics formulations. To make the arising governing equations more tractable, the constraint equations are differentiated with respect to time, and this results in unstable numerical solutions which may violate the lower-order constraint equations. In this paper we develop a methodology for numerically exact elimination of the constraint violations, based on appropriate corrections of the state variables (after each integration step) without any modification in the motion equations. While the elimination of violation of position constraints may require few iterations, the violation of velocity constraints is removed in one step. The total energy of the system is sometimes treated as another measure of the integration process inaccuracy. An improved scheme for one-step elimination of the energy constraint violation is proposed as well. The conclusion of this paper is, however, that the energy conservation is of minor importance as concerns the improvement of accuracy of numerical simulations. Some test calculations are reported.

Key words: multibody dynamics, constraint violation, numerical accuracy.

1. Introduction

Multibody systems, both holonomic and nonholonomic, are often modeled as constrained systems. Some (or all) kinematic joints are first cut off in order to obtain a more tractable system whose equations of motion can be derived by using automated procedures, and then the closing constraint conditions are imposed on the system. The arising governing equations are composed of the constraint reaction-induced equations of motion and the constraint equations. In simulations, the latter are commonly applied in the time-differentiated forms, and this results in unstable numerical solutions, i.e. the original constraint equations are violated by the solutions burdened with the integration truncation errors. Special procedures must be followed to avoid/minimize the phenomenon. A popular technique is Baumgarte's constraint violation stabilization method [3], where the second-order differential constraint equations are used in a stabilized form. However, the method does not provide full constraint satisfaction, and has some ambiguity in determining optimal

feedback gains. The involvement of time integral of constraint violations improves the constraint satisfaction [16]. A variety of suggestions have then been proposed to eliminate the feedback gain ambiguity and extend Baumgarte's idea, see, e.g., [1, 2, 9, 12, 14, 21]. The other developments are based on a penalty approach [17] and a mass-orthogonal projection [4], resulting in remarkable constraint satisfaction and accuracy in long simulations. A different solution was proposed in [5, 20] that uses a geometric elimination of constraint violations. Improved schemes of the latter type are developed and tested in this contribution.

Without any modification in the motion equations, the state variables are corrected geometrically to eliminate the constraint violation with a numerical accuracy after each step of integration or a sequence of steps. The corrections are performed in the directions orthogonal to the constraint manifold [5, 6], and as such do not affect the system motion on the manifold. While the exact elimination of violation of position constraints may require few iterations, the violation of velocity constraints is removed in one step. Then, if the total energy of the system can be computed from the initial energy plus the energy input rate due to the external or dissipative forces, it can be regarded as another measure of accuracy of numerical simulation and treated as an artificial constraint on the system [20]. Assumed the violation of kinematic constraints is removed, an improved scheme is developed that assures exact elimination of the energy constraint by appropriate correction of the system velocity in one step. The conclusion of this paper is, however, that the energy conservation may be of minor importance as concerns the improvement of accuracy of numerical simulations. Some test calculations are reported to illustrate the effectiveness of the constraint violation elimination schemes and to conclude on accuracy of the numerical simulations.

2. The Governing Equations

The mathematical models for the dynamic analysis of rigid multibody systems fall into two main categories as concerns the used governing equations. The first group formulations are characterized by dependent state variables whose number exceeds the system state number, i.e. the system is treated as a constrained system. The arising equations of motion form a mixed set of differential-algebraic equations (DAEs), which results in a computationally inefficient algorithms burdened furthermore with the constraint violation problem. The other approach is to use a minimal number of (independent) state variables for a unique representation of motion by means of pure ordinary differential equations (ODEs). The numerical integration of these equations is often more efficient compared to integration of equations expressed in terms of dependent variables, and the numerical solution is released from the problem of constraint violation. However, the minimal-form formulations of multibody dynamics may be quite complex, i.e. may require many symbolic manipulations. Let us shortly review some basic formulations of the two types.

2.1. A CONSTRAINED SYSTEM

Consider an n-degrees-of-freedom autonomous system characterized by n generalized coordinates $\mathbf{p} = [p_1 \dots p_n]^T$ and n velocities (either generalized velocities, quasi-velocities or both) $\mathbf{v} = [v_1 \dots v_n]^T$. The system equations of motion can be written in the following matrix form [5, 6]:

$$\dot{\mathbf{p}} = \mathbf{A}(\mathbf{p})\mathbf{v}, \tag{1}$$

$$\mathbf{M}(\mathbf{p})\dot{\mathbf{v}} + \mathbf{d}(\mathbf{p}, \mathbf{v}) = \mathbf{f}(\mathbf{p}, \mathbf{v}, t), \tag{2}$$

where, in the kinematic differential equation (1), $\mathbf{A}$ is an $n \times n$ transformation matrix (frequently $\dot{\mathbf{p}} = \mathbf{v}$, i.e. $\mathbf{A}$ is an identity matrix), and, in the dynamic equation (2), $\mathbf{M}$ is the $n \times n$ symmetric positive definite generalized mass matrix, $\mathbf{d}$ represents the centrifugal, Coriolis and gyroscopic dynamic terms, $\mathbf{f}$ represents the applied forces, and t is the time. We refer to the system as an *unconstrained system*, which can be a collection of unconstrained particles/bodies, an open-loop (tree structure) system, or a combination of such subsystems.

Let the above system be subject to m_{H} holonomic (H) and m_{NH} nonholonomic (NH) constraints, $m = m_{\mathrm{H}} + m_{\mathrm{NH}}$, both assumed scleronomic for simplicity:

$$\mathbf{\Phi}_{\mathrm{H}}(\mathbf{p}) = \mathbf{0}, \tag{3}$$

$$\mathbf{\Psi}_{\mathrm{NH}}(\mathbf{p}, \mathbf{v}) \equiv \mathbf{C}_{\mathrm{NH}}(\mathbf{p})\mathbf{v} = \mathbf{0}, \tag{4}$$

where $\mathbf{C}_{\mathrm{NH}}$ is the $m_{\mathrm{NH}} \times n$ NH constraint matrix. By differentiating with respect to time the H constraints (3), m unified first-order differential constraint equations, $\mathbf{\Psi} = [\dot{\mathbf{\Phi}}_{\mathrm{H}}^T \;\; \mathbf{\Psi}_{\mathrm{NH}}^T]^T$, are obtained

$$\mathbf{\Psi}(\mathbf{p}, \mathbf{v}) \equiv \mathbf{C}(\mathbf{p})\mathbf{v} = \mathbf{0}, \tag{5}$$

where $\mathbf{C} = [\mathbf{C}_{\mathrm{H}}^T \;\; \mathbf{C}_{\mathrm{NH}}^T]^T$ and $\mathbf{C}_{\mathrm{H}} = (\partial \mathbf{\Phi}_{\mathrm{H}}/\partial \mathbf{p})\mathbf{A}$. The second-order differential constraint equations are then obtained as

$$\dot{\mathbf{\Psi}}(\mathbf{p}, \mathbf{v}) \equiv \mathbf{C}(\mathbf{p})\dot{\mathbf{v}} - \boldsymbol{\xi}(\mathbf{p}, \mathbf{v}) = \mathbf{0}, \tag{6}$$

where $\boldsymbol{\xi} = -\dot{\mathbf{C}}\mathbf{v}$.

According to the Lagrange multiplier method, the dynamic equations (2) are now

$$\mathbf{M}(\mathbf{p})\dot{\mathbf{v}} + \mathbf{d}(\mathbf{p}, \mathbf{v}) = \mathbf{f}(\mathbf{p}, \mathbf{v}, t) - \mathbf{C}^T(\mathbf{p})\boldsymbol{\lambda}, \tag{7}$$

where $\mathbf{C}^T\boldsymbol{\lambda}$ are the generalized forces due to the constraint reactions $\boldsymbol{\lambda} = [\lambda_1 \dots \lambda_n]^T$ (Lagrange multipliers) related to the H and NH constraints, respectively, $\boldsymbol{\lambda} = [\boldsymbol{\lambda}_{\mathrm{H}}^T \;\; \boldsymbol{\lambda}_{\mathrm{NH}}^T]^T$. By assuming that $\mathbf{\Phi}_{\mathrm{H}}$ expresses prohibited translations and rotations, and $\mathbf{\Psi}_{\mathrm{NH}}$ denote vanishing translational and rotational velocities in the joints, $\boldsymbol{\lambda}$ contains respective physical forces and moments.

2.2. DEPENDENT VARIABLE FORMULATIONS

The state variables $\mathbf{p}$ and $\mathbf{v}$ of the above defined system are dependent. The direct formulation of the governing equations in terms of these variables follows after combining the kinematic differential equations (1), the constraint reaction-induced dynamic equations (7), and the constraint equations (6), which results in $2n + m$ DAEs in $\mathbf{p}$, $\mathbf{v}$, and $\boldsymbol{\lambda}$,

$$\begin{aligned} \dot{\mathbf{p}} &= \mathbf{A}\mathbf{v} \\ \begin{bmatrix} \mathbf{M} & \mathbf{C}^T \\ \mathbf{C} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \dot{\mathbf{v}} \\ \boldsymbol{\lambda} \end{bmatrix} &= \begin{bmatrix} \mathbf{f} - \mathbf{d} \\ \boldsymbol{\xi} \end{bmatrix} \end{aligned} \quad \Longleftrightarrow \quad \begin{aligned} \dot{\mathbf{p}} &= \mathbf{A}(\mathbf{p})\mathbf{v} \\ \mathbf{G}(\mathbf{p}) \begin{bmatrix} \dot{\mathbf{v}} \\ \boldsymbol{\lambda} \end{bmatrix} &= \mathbf{g}(\mathbf{p}, \mathbf{v}, t) \end{aligned} \tag{8}$$

and the initial values of the state variables must satisfy the lower-order constraint conditions (3) and (5), $\boldsymbol{\Phi}_{\mathrm{H}}(\mathbf{p}_0) = \mathbf{0}$ and $\boldsymbol{\Psi}(\mathbf{p}_0, \mathbf{v}_0) \equiv \mathbf{C}(\mathbf{p}_0)\mathbf{v}_0 = \mathbf{0}$. Since matrix $\mathbf{G}$ is invertible if only the row-rank of $\mathbf{C}$ is maximal, $\mathrm{rank}(\mathbf{C}) = \max$ (constraints (5) are independent), DAEs (8) can be solved for $\mathbf{p}$ and $\mathbf{v}$ using standard ODE methods.

The other popular formulation introduces an orthogonal complement matrix $\mathbf{D}(\mathbf{p})$ (an $n \times r$ full column-rank matrix, where $r = n - m$ is the number of degrees of freedom of the system) to the $m \times n$ constraint matrix $\mathbf{C}$, such that $\mathbf{D}^T\mathbf{C}^T = \mathbf{0}$. As shown, e.g., in [1, 5, 6, 12], the premultiplication of Equations (7) by $\mathbf{D}^T$ means the projection of the dynamic equations into the tangent directions with respect to constraints (3) and (4). The projected dynamic equations, released from the constraint reactions due to $\mathbf{D}^T\mathbf{C}^T\boldsymbol{\lambda} = \mathbf{0}$, combined with the constraint equations (6), lead to the following $2n$ ODEs in $\mathbf{p}$ and $\mathbf{v}$ [5, 6]

$$\begin{aligned} \dot{\mathbf{p}} &= \mathbf{A}\mathbf{v} \\ \begin{bmatrix} \mathbf{D}^T\mathbf{M} \\ \mathbf{C} \end{bmatrix} \dot{\mathbf{v}} &= \begin{bmatrix} \mathbf{D}^T(\mathbf{f} - \mathbf{d}) \\ \boldsymbol{\xi} \end{bmatrix} \end{aligned} \quad \Longleftrightarrow \quad \begin{aligned} \dot{\mathbf{p}} &= \mathbf{A}(\mathbf{p})\mathbf{v} \\ \mathbf{H}(\mathbf{p})\dot{\mathbf{v}} &= \mathbf{h}(\mathbf{p}, \mathbf{v}, t) \end{aligned} \tag{9}$$

An explicit (by guess/inspection) choice of $\mathbf{D}$ for a given $\mathbf{C}$, which is not unique in general, can be done only for simple systems. For more complex systems, $\mathbf{D}$ is usually determined numerically by using various techniques [1, 12, 13]. A popular one is based on variable partitioning [18], which will be reported in Section 2.3.

Using the block-matrix inversion scheme [8]

$$\begin{bmatrix} \mathbf{P} & \mathbf{Q} \\ \mathbf{R} & \mathbf{S} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{P}^{-1} + \mathbf{E}\boldsymbol{\Delta}^{-1}\mathbf{F} & -\mathbf{E}\boldsymbol{\Delta}^{-1} \\ -\boldsymbol{\Delta}^{-1}\mathbf{F} & \boldsymbol{\Delta}^{-1} \end{bmatrix}, \tag{10}$$

where $\boldsymbol{\Delta} = \mathbf{S} - \mathbf{R}\mathbf{P}^{-1}\mathbf{Q}$, $\mathbf{E} = \mathbf{P}^{-1}\mathbf{Q}$, and $\mathbf{F} = \mathbf{R}\mathbf{P}^{-1}$, the inversion of the leading matrix $\mathbf{G}$ in Equation (8) can be represented as

$$\begin{aligned} \mathbf{G}^{-1} &\equiv \begin{bmatrix} \mathbf{M} & \mathbf{C}^T \\ \mathbf{C} & \mathbf{0} \end{bmatrix}^{-1} \\ &= \begin{bmatrix} \mathbf{M}^{-1} - \mathbf{M}^{-1}\mathbf{C}^T(\mathbf{C}\mathbf{M}^{-1}\mathbf{C}^T)^{-1}\mathbf{C}\mathbf{M}^{-1} & \mathbf{M}^{-1}\mathbf{C}^T(\mathbf{C}\mathbf{M}^{-1}\mathbf{C}^T)^{-1} \\ (\mathbf{C}\mathbf{M}^{-1}\mathbf{C}^T)^{-1}\mathbf{C}\mathbf{M}^{-1} & -(\mathbf{C}\mathbf{M}^{-1}\mathbf{C}^T)^{-1} \end{bmatrix}, \end{aligned} \tag{11}$$

On the other hand, in [7] it is shown that

$$\mathbf{H}^{-1} = \begin{bmatrix} \mathbf{D}^T\mathbf{M} \\ \mathbf{C} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{D}(\mathbf{D}^T\mathbf{M}\mathbf{D})^{-1} & \mathbf{M}^{-1}\mathbf{C}^T(\mathbf{C}\mathbf{M}^{-1}\mathbf{C}^T)^{-1} \end{bmatrix} \tag{12}$$

and then $\mathbf{D}(\mathbf{D}^T\mathbf{M}\mathbf{D})^{-1}\mathbf{D}^T = \mathbf{M}^{-1} - \mathbf{M}^{-1}\mathbf{C}^T(\mathbf{C}\mathbf{M}^{-1}\mathbf{C}^T)^{-1}\mathbf{C}\mathbf{M}^{-1}$.

Applying the schemes (11) and (12) to Equations (8) and (9), respectively, one obtains:

$$\begin{aligned} \dot{\mathbf{p}} &= \mathbf{A}\mathbf{v}, \\ \dot{\mathbf{v}} &= \mathbf{M}^{-1}(\mathbf{f} - \mathbf{d}) + \mathbf{M}^{-1}\mathbf{C}^T(\mathbf{C}\mathbf{M}^{-1}\mathbf{C}^T)^{-1}[\boldsymbol{\xi} - \mathbf{C}\mathbf{M}^{-1}(\mathbf{f} - \mathbf{d})]. \end{aligned} \tag{13}$$

Incidentally, the above dependent variable formulation can be obtained directly [3] by substituting $\dot{\mathbf{v}}$ from the dynamic equation (7) into the constraint equation (6), resolving the followed relation for $\boldsymbol{\lambda}(\mathbf{p}, \mathbf{v}, t) = (\mathbf{C}\mathbf{M}^{-1}\mathbf{C}^T)^{-1}[\mathbf{C}\mathbf{M}^{-1}(\mathbf{f} - \mathbf{d}) - \boldsymbol{\xi}]$, and then using the result back in the dynamic equation.

All the three dependent variable formulations, (8), (9) and (13), involve the second-order differential constraint equations (6). Thus, in simulation, the exact realization of only these constraint equations is assured by assumption. The lower-order constraint equations (3) and (5) may be violated by the numerical solutions $\tilde{\mathbf{p}}(t)$ and $\tilde{\mathbf{v}}(t)$ burdened with the numerical error of integration, $\tilde{\boldsymbol{\Phi}}_{\mathrm{H}} = \boldsymbol{\Phi}_{\mathrm{H}}(\tilde{\mathbf{p}}) \neq \mathbf{0}$ and $\tilde{\boldsymbol{\Psi}} = \boldsymbol{\Psi}(\tilde{\mathbf{p}}, \tilde{\mathbf{v}}) \neq \mathbf{0}$. The situation differs when independent state variables are used.

2.3. INDEPENDENT VARIABLE FORMULATIONS

Let us first recall a general methodology [5, 6, 12] of converting the dynamic equations of motion to a minimal set in independent *virtual speeds* $\mathbf{u} = [u_1 \dots u_r]^T$, $r = n - m$, called also *independent kinematic parameters* in Maggi's and Gibbs–Appell methods [15], and *generalized speeds* in Kane's method [13]. Using $\mathbf{u}$, the *implicit* first-order and second-order differential constraint equations (5) and (6) can be replaced by their *explicit* forms

$$\mathbf{v} = \mathbf{D}(\mathbf{p})\mathbf{u}, \tag{14}$$

$$\dot{\mathbf{v}} = \mathbf{D}(\mathbf{p})\dot{\mathbf{u}} + \boldsymbol{\gamma}(\mathbf{p}, \mathbf{u}), \tag{15}$$

where $\boldsymbol{\gamma} = \dot{\mathbf{D}}\mathbf{u}$, and the implicit constraint equations (5) and (6) are satisfied identically when their explicit forms (14) and (15) are substituted [5, 6]. This means also that $\mathbf{D}$ is an orthogonal complement matrix to $\mathbf{C}$, $\mathbf{C}\mathbf{D} = \mathbf{0} \Leftrightarrow \mathbf{D}^T\mathbf{C}^T = \mathbf{0}$, and $\mathbf{C}\boldsymbol{\gamma} = \boldsymbol{\xi}$.

The components of $\mathbf{u}$ may be a subset of $\mathbf{v}$, a set of new velocities, and/or a set of kinematic parameters which may have no physical meaning [5, 6], and the choice of $\mathbf{u}$ is closely related to the determination of $\mathbf{D}$. While an analytical, by guess or inspection, formulation of Equations (14) and (15) is usually feasible

only for simple systems, an automatic computer-oriented code for obtaining the relations is provided by the *variable partitioning method* [6, 7, 19]. Using $\mathbf{v} = [\mathbf{u}^T \ \ \mathbf{w}^T]^T$, where r independent $\mathbf{u}$ and m dependent $\mathbf{w}$ velocities are chosen so that the respective constraint matrix factorization $\mathbf{C} = [\mathbf{U} \ \vdots \ \mathbf{W}]$ is characterized by $\det(\mathbf{W}) \neq 0$, from the factorized constraint equations (5) and (6), $\boldsymbol{\Psi} \equiv \mathbf{U}\mathbf{u}+\mathbf{W}\mathbf{w} = \mathbf{0}$ and $\dot{\boldsymbol{\Psi}} \equiv \mathbf{U}\dot{\mathbf{u}} + \mathbf{W}\dot{\mathbf{w}} - \boldsymbol{\xi} = \mathbf{0}$, one receives after simple manipulations:

$$\begin{aligned} \mathbf{v} &= \begin{bmatrix} \mathbf{I} \\ -\mathbf{W}^{-1}\mathbf{U} \end{bmatrix} \mathbf{u} \equiv \mathbf{D}\mathbf{u}, \\ \dot{\mathbf{v}} &= \begin{bmatrix} \mathbf{I} \\ -\mathbf{W}^{-1}\mathbf{U} \end{bmatrix} \dot{\mathbf{u}} + \begin{bmatrix} \mathbf{0} \\ \mathbf{W}^{-1}\boldsymbol{\xi} \end{bmatrix} \equiv \mathbf{D}\dot{\mathbf{u}} + \boldsymbol{\gamma}. \end{aligned} \tag{16}$$

Applying the explicit constraint equations (14) and (15), the governing equations (9) transform to $n + r$ ODEs in $\mathbf{p}$ and $\mathbf{u}$ (see [5, 6] for more details),

$$\begin{aligned} \dot{\mathbf{p}} &= \mathbf{A}\mathbf{D}\mathbf{u} \\ \mathbf{D}^T\mathbf{M}\mathbf{D}\dot{\mathbf{u}} + \mathbf{D}^T(\mathbf{M}\boldsymbol{\gamma} + \mathbf{d}) &= \mathbf{D}^T\mathbf{f} \end{aligned} \iff \begin{aligned} \dot{\mathbf{p}} &= \mathbf{A}_u(\mathbf{p})\mathbf{u} \\ \mathbf{M}_u(\mathbf{p})\dot{\mathbf{u}} + \mathbf{d}_u(\mathbf{p}, \mathbf{u}) &= \mathbf{f}_u(\mathbf{p}, \mathbf{v}, t) \end{aligned} \tag{17}$$

The above procedure is valid for both H and NH systems. By virtue of the formulation, the unified first-order differential constraint equations (5) are satisfied. The original H constraints (3) may however still be violated by the numerical solution $\tilde{\mathbf{p}}(t)$, $\tilde{\boldsymbol{\Phi}}_{\mathrm{H}} = \boldsymbol{\Phi}_{\mathrm{H}}(\tilde{\mathbf{p}}) \neq \mathbf{0}$. This can be removed by expressing the motion equations in terms k independent generalized coordinates $\mathbf{q} = [q_1 \dots q_k]^T$, $k = n - m_{\mathrm{H}}$. The H constraints (3) and their differentiated forms can then be given explicitly by

$$\mathbf{p} = \mathbf{g}(\mathbf{q}) \ \Rightarrow \ \mathbf{v} = \mathbf{D}'_{\mathrm{H}}(\mathbf{q})\dot{\mathbf{q}} \ \Rightarrow \ \dot{\mathbf{v}} = \mathbf{D}'_{\mathrm{H}}(\mathbf{q})\ddot{\mathbf{q}} + \boldsymbol{\gamma}'_{\mathrm{H}}(\mathbf{q}, \dot{\mathbf{q}}), \tag{18}$$

where $\mathbf{D}'_{\mathrm{H}} = \mathbf{A}^{-1}(\partial\mathbf{g}/\partial\mathbf{q})$ is now of dimension $n \times k$, and $\gamma'_{\mathrm{H}} = \mathbf{D}'_{\mathrm{H}}\dot{\mathbf{q}}$. Using these relations and the above described procedure, the dynamic equations (2) can be transformed to

$$\mathbf{M}_q(\mathbf{q})\ddot{\mathbf{q}} + \mathbf{d}_q(\mathbf{q}, \dot{\mathbf{q}}) = \mathbf{f}_q(\mathbf{q}, \dot{\mathbf{q}}, t), \tag{19}$$

where $\mathbf{M}_q = \mathbf{D}'^T_{\mathrm{H}}\mathbf{M}\mathbf{D}'_{\mathrm{H}}$, $\mathbf{d}_q = \mathbf{D}'^T_{\mathrm{H}}(\mathbf{M}\boldsymbol{\gamma}'_{\mathrm{H}} + \mathbf{d})$, and $\mathbf{f}_q = \mathbf{D}'_{\mathrm{H}}\mathbf{f}$. The number of these equations is equal to $k = n - m_{\mathrm{H}}$, and the system described in this way is subject to only m_{NH} NH constraints, which, given *implicitly* and *explicitly*, are now:

$$\begin{aligned} \boldsymbol{\Psi}'_{\mathrm{NH}} &\equiv \mathbf{C}'_{\mathrm{NH}}(\mathbf{q})\dot{\mathbf{q}} = \mathbf{0}, \\ \dot{\boldsymbol{\Psi}}'_{\mathrm{NH}} &\equiv \mathbf{C}'_{\mathrm{NH}} \equiv \mathbf{C}'_{\mathrm{NH}}(\mathbf{q})\ddot{\mathbf{q}} - \boldsymbol{\xi}'_{\mathrm{NH}}(\mathbf{q}, \dot{\mathbf{q}}) = \mathbf{0}, \end{aligned} \tag{20}$$

$$\begin{aligned} \dot{\mathbf{q}} &= \mathbf{D}'_{\mathrm{NH}}(\mathbf{q})\mathbf{u}', \\ \ddot{\mathbf{q}} &= \mathbf{D}'_{\mathrm{NH}}(\mathbf{q})\dot{\mathbf{u}}' + \boldsymbol{\gamma}'_{\mathrm{NH}}(\mathbf{q}, \mathbf{u}'). \end{aligned} \tag{21}$$

The minimal-form governing equations are then the following $k+r$ ODEs in $\mathbf{q}$ and $\mathbf{u}'$

$$\begin{array}{c} \dot{\mathbf{q}} = \mathbf{D}'_{\mathrm{NH}}\mathbf{u}' \\ \mathbf{D}'^{T}_{\mathrm{NH}}\mathbf{M}_q\mathbf{D}'_{\mathrm{NH}}\dot{\mathbf{u}}' + \mathbf{D}'^{T}_{\mathrm{NH}}(\mathbf{M}_q\boldsymbol{\gamma}'_{\mathrm{NH}} + \mathbf{d}_q) = \mathbf{D}'^{T}_{\mathrm{NH}}\mathbf{f}_q \end{array} \Longleftrightarrow \begin{array}{c} \dot{\mathbf{q}} = \mathbf{D}'_{\mathrm{NH}}(\mathbf{q})\mathbf{u}' \\ \mathbf{M}'_u(\mathbf{q})\dot{\mathbf{u}}' + \mathbf{d}'_u(\mathbf{q}, \mathbf{u}') = \mathbf{f}'_u(\mathbf{p}, \mathbf{v}, t) \end{array} \tag{22}$$

Summarizing, the different governing equations introduced above can be grouped as follows:

(a) The dependent variable formulations (8), (9) and (13) in which both the position **p** and velocity **v** variables are dependent. The formulations are valid for systems with H and/or NH constraints, and the numerical solution of the governing equations tends to violate the original H constraint equations (3) as well as the unified first-order differential constraint equations (5).
(b) The mixed dependent-independent formulation (17) in which the dependent position **p** and independent velocity **u** variables are used. The formulation is valid for systems with H and/or NH constraints, and the numerical solution of the governing equations tends to violate only the original H constraint equations (3).
(c) The independent variable formulations (19) (for H systems) and (22) (for NH systems) in which the independent state variables: **q** and $\dot{\mathbf{q}}$, and **q** and $\mathbf{u}'$, respectively, are used. The numerical solutions of these equations are released from the kinematic constraint violation problem.

3. State and Energy Stabilization

The numerical solution of equations of motion is always burdened with the integration truncation errors. For a constrained system modeled in dependent state variables, when the governing equations involve the second-order differential constraint equations, one consequence of the numerical inaccuracy is violation of the lower-order constraint equations, measured in the directions orthogonal to the respective constraints [5, 6]. The independent variable formulations assure the constraint consistency by assumption. The other consequence of truncation errors, relevant to both dependent and independent variable formulations, is possible difference from the 'exact' solution on the constraint manifold. For a system with H constraints these concepts are illustrated in Figure 1. The interpretation for a NH constraint case is more intuitive. The state inaccuracy tends to increase with simulation time and, not controlled, may make the numerical analysis unreliable or even worthless.

3.1. ELIMINATION OF KINEMATIC CONSTRAINT VIOLATIONS

A legitimate technique for stabilizing the kinematic constraint violations is Baumgarte's method [3], in which the second-order differential constraint equations (6) are replaced by their stabilized form (see also [1, 2, 9, 12, 14, 16, 20, 21])

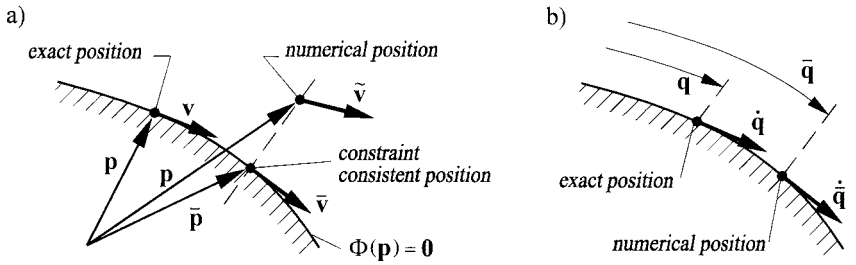

Figure 1. Illustration of state inaccuracy for a system with holonomic constraints; (a) dependent variable formulation, (b) independent coordinate formulation.

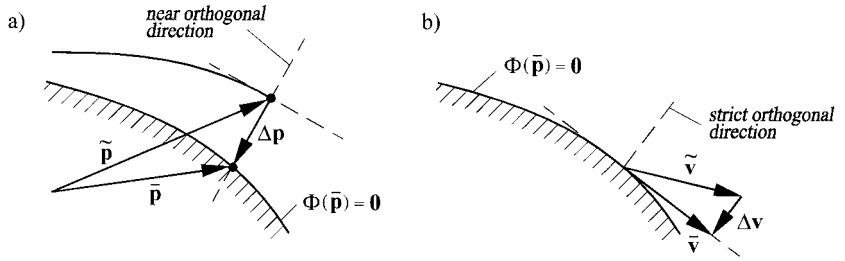

Figure 2. Two-step elimination of kinematic constraint violations.

$$\dot{\boldsymbol{\Psi}} + \boldsymbol{\alpha}\boldsymbol{\Psi} + \boldsymbol{\beta}\begin{bmatrix} \boldsymbol{\Phi}_{\mathrm{H}} \\ \mathbf{0} \end{bmatrix} = \mathbf{0} \quad \Leftrightarrow \quad \boldsymbol{\xi}_{\mathrm{stab}} = \boldsymbol{\xi} - \boldsymbol{\alpha}\boldsymbol{\Psi} - \boldsymbol{\beta}\begin{bmatrix} \boldsymbol{\Phi}_{\mathrm{H}} \\ \mathbf{0} \end{bmatrix}, \tag{23}$$

where $\boldsymbol{\Psi} = [\dot{\boldsymbol{\Phi}}_{\mathrm{H}}^{T} \ \boldsymbol{\Psi}_{\mathrm{NH}}^{T}]^{T}$, and $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are diagonal matrices of feedback gains. The governing equations (8), (9) and (13) are then modified by applying $\boldsymbol{\xi}_{\mathrm{stab}}$ instead of only $\boldsymbol{\xi}$. The method does not work well for relatively complicated systems, however. It does not provide full constraint satisfaction either, and has some ambiguity in determining the optimal feedback gains [2, 9, 14, 16, 21].

The other approach [5] is to correct directly the state variables so that to eliminate the constraint violations after each integration step or a sequence of steps, without any modification in the motion equations. The schemes proposed there use an advanced geometrical formalism, and an important feature of the state corrections is that they are performed in the orthogonal-to-constraint directions, and as such do not affect the system kinetic motion (see Figure 2). Improved schemes of this type are developed and tested in this contribution.

For a current numerical position $\tilde{\mathbf{p}}(t)$, $\tilde{\boldsymbol{\Phi}}_{\mathrm{H}} = \boldsymbol{\Phi}_{\mathrm{H}}(\tilde{\mathbf{p}}) \neq \mathbf{0}$ is a measure of deflection of $\tilde{\mathbf{p}}(t)$ from the virtual H constraint manifold, and the components of $\tilde{\boldsymbol{\Phi}}(t)$ denote distances from $\tilde{\mathbf{p}}(t)$ to the manifold measured along the constraint gradients. Then, $\boldsymbol{\Psi} = \boldsymbol{\Psi}(\tilde{\mathbf{p}}, \tilde{\mathbf{v}}) \neq \mathbf{0}$ indicates that the numerical velocity $\tilde{\mathbf{v}}(t)$ is not pointed in a constraint allowed direction (for H constraints, is not tangent to the constraint manifold). The components of $\tilde{\boldsymbol{\Psi}}$ are measures of projections of $\tilde{\mathbf{v}}(t)$ onto the constrained directions; see [5, 6] for more details. The problem at hand is

then to find formulae for 'translating' the violations into appropriate position and velocity corrections,

$$\Delta\mathbf{p} = \overline{\mathbf{p}}(t) - \tilde{\mathbf{p}}(t) \quad \text{and} \quad \Delta\mathbf{v} = \overline{\mathbf{v}}(t) - \tilde{\mathbf{v}}(t), \tag{24}$$

where $\overline{\mathbf{q}}(t)$ and $\overline{\mathbf{v}}(t)$ are the corrected, constraint consistent position and velocity of the system, $\mathbf{\Phi}_{\mathrm{H}}(\overline{\mathbf{p}}) = \mathbf{0}$ and $\mathbf{\Psi}(\overline{\mathbf{p}}, \overline{\mathbf{v}}) = \mathbf{0}$.

According to the geometrical interpretation [5, 6], $\boldsymbol{\xi}$ are constraint induced accelerations of the system, and are pointed in the constrained directions (in the basis of constrained subspace defined by constraint vectors contained in $\mathbf{C}$ as rows). The governing equations (8), (9) and (13) involve then variant formulae for transformation of $\boldsymbol{\xi}$ from the constrained directions to the directions related to $\mathbf{v}$. As previously said, the constraint violations $\tilde{\mathbf{\Phi}}_{\mathrm{H}}$ and $\tilde{\mathbf{\Psi}}$ are also represented in the constrained directions. This yields the following correction schemes:

$$\begin{bmatrix} \mathbf{M} & \mathbf{C}^T \\ \mathbf{C} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{A}^{-1}\Delta\mathbf{p} \\ \mathbf{0} \end{bmatrix} = - \begin{bmatrix} \mathbf{0} \\ \tilde{\mathbf{\Phi}}_{\mathrm{H}} \\ \mathbf{0} \end{bmatrix}, \tag{25a}$$

$$\begin{bmatrix} \mathbf{M} & \mathbf{C}^T \\ \mathbf{C} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \Delta\mathbf{v} \\ \mathbf{0} \end{bmatrix} = - \begin{bmatrix} \mathbf{0} \\ \tilde{\mathbf{\Psi}} \end{bmatrix}, \tag{25b}$$

$$\begin{bmatrix} \mathbf{D}^T\mathbf{M} \\ \mathbf{C} \end{bmatrix} \mathbf{A}^{-1}\Delta\mathbf{p} = - \begin{bmatrix} \mathbf{0} \\ \tilde{\mathbf{\Phi}}_{\mathrm{H}} \\ \mathbf{0} \end{bmatrix}, \tag{26a}$$

$$\begin{bmatrix} \mathbf{D}^T\mathbf{M} \\ \mathbf{C} \end{bmatrix} \Delta\mathbf{v} = - \begin{bmatrix} \mathbf{0} \\ \tilde{\mathbf{\Psi}} \end{bmatrix}, \tag{26b}$$

$$\Delta\mathbf{p} = -\mathbf{A}\mathbf{M}^{-1}\mathbf{C}^T(\mathbf{C}\mathbf{M}^{-1}\mathbf{C}^T)^{-1} \begin{bmatrix} \tilde{\mathbf{\Phi}}_{\mathrm{H}} \\ \mathbf{0} \end{bmatrix}, \tag{27a}$$

$$\Delta\mathbf{v} = -\mathbf{M}^{-1}\mathbf{C}^T(\mathbf{C}\mathbf{M}^{-1}\mathbf{C}^T)^{-1}\tilde{\mathbf{\Psi}}, \tag{27b}$$

which correspond to the governing equations (8), (9) and (13), respectively. The equivalence of the schemes can be shown by using Equations (11) and (12). As shown in [5], the position correction in Equations (27) can also be stated as $\Delta\mathbf{p} = -\mathbf{A}\mathbf{M}^{-1}\mathbf{C}_{\mathrm{H}}^T(\mathbf{C}_{\mathrm{H}}\mathbf{M}^{-1}\mathbf{C}_{\mathrm{H}}^T)^{-1}\tilde{\mathbf{\Phi}}_{\mathrm{H}}$, and similar modifications are possible for schemes (25) and (26) as well. The present schemes are more convenient, however, since they use the same matrices as in the respective governing equations. Moreover, as compared to [5], the new schemes (25) and (26) seem to be of special importance when the formulations (8) or (9) are used.

In computations, the position correction $\overline{\mathbf{p}} = \tilde{\mathbf{p}} + \Delta\mathbf{p}$ (Figure 2a) should be done first. Since the coefficient matrices in schemes (25), (26) and (27) are estimated

using the deflected position $\tilde{\mathbf{p}}$, the correction precision may be diminished. The exact correction may thus be an iterative process. Nevertheless, usually two or at most three iterations are needed to achieve $\mathbf{\Phi}_{\mathrm{H}}(\overline{\mathbf{p}}) = \mathbf{0}$ with a numerical accuracy. Then, having the system position revised, the velocity correction $\overline{\mathbf{v}} = \tilde{\mathbf{v}} + \Delta\mathbf{v}$ (Figure 2b) can be done in one step, where now $\tilde{\mathbf{\Psi}} = \mathbf{\Psi}(\overline{\mathbf{p}}, \tilde{\mathbf{v}}) \equiv \mathbf{C}(\overline{\mathbf{p}})\tilde{\mathbf{v}}$.

In the case of the governing equations (17) expressed in dependent coordinates $\mathbf{p}$ and independent velocities $\mathbf{u}$, the violations of only the original H constraints (3) may occur, which may be eliminated according to the position corrections of the above schemes.

It might be worth noting that resembling schemes for elimination of kinematic constraint violations were developed in [20]. For a system modeled in generalized coordinates $\mathbf{p}$ ($\dot{\mathbf{p}} = \mathbf{v}$) and subject to H constraints, the correction schemes were $\Delta\mathbf{p} = -\mathbf{C}^T(\mathbf{C}\mathbf{C}^T)^{-1}\tilde{\mathbf{\Phi}}$ and $\Delta\mathbf{v} = -\mathbf{C}^T(\mathbf{C}\mathbf{C}^T)^{-1}\tilde{\mathbf{\Psi}}$. Such schemes are somewhat inconsistent, however. As $\mathbf{p}$ may in general be translational and/or rotational coordinates, the entries of constraint gradients (rows of $\mathbf{C}$) may have different dimensions. Consequently, the entries of the $m \times m$ matrix $\mathbf{C}\mathbf{C}^T$ may be calculated by summing up addends of different dimensions. By using $\mathbf{C}\mathbf{M}^{-1}\mathbf{C}^T$, the correct dot products of constraint gradients [5, 6], the inconsistency is removed. Moreover, only the present schemes (25–27) assure that the corrections are performed in the virtually orthogonal directions. The schemes proposed in [20] do not satisfy this condition in a general case.

3.2. ASPECTS OF ACCURACY OF CONSTRAINT-CONSISTENT SOLUTIONS

The numerical integration errors continuously disturb also the system kinetic motion, i.e. the motion consistent with the constraint conditions (on the constraint manifold), and the problem relates both the dependent and independent variable formulations. By inaccuracy of numerical integration in this sense we mean the difference between the 'exact' and the numerical (constraint-consistent) solution to the equations of motion. Since an analytical (exact by assumption) solution to the equations is usually unattainable, one must resort to their numerical solutions. The reference solution can then be obtained using a very small step size, and a difference between the solutions obtained for moderate step sizes and the reference solution are always observed irrespective of the method chosen to integrate the equations of motion.

A popular measure of the integration process inaccuracy is also the total energy of the system, computed from the initial energy plus the energy input rate due to external and dissipative forces [10, 21] (reflected also in the developments of energy preserving integrators, see, e.g., [11, 12, 18]). The total energy can then be treated as an artificial constraint on the system

$$\Psi_E = E - E_0 - \int_{t_0}^{t} \dot{E}\,\mathrm{d}t = 0, \tag{28}$$

where $E = T + V$ is the sum of the kinetic T and potential V energies, and $\dot{E}$ denotes the energy input rate to the system (for conservative systems $\dot{E} = 0$). For the dependent variable formulations (8), (9) and (13), and the mixed dependent-independent formulation (17), the kinetic energy functions are, respectively, $T(\mathbf{p}, \mathbf{v}) = \mathbf{v}^T\mathbf{M}(\mathbf{p})\mathbf{v}/2$ and $T(\mathbf{p}, \mathbf{u}) = \mathbf{u}^T\mathbf{M}_u(\mathbf{p})\mathbf{u}/2$, while $V(\mathbf{p})$. For the independent variable formulations (19) (for H systems) and (22) (for NH systems) we have, respectively, $T(\mathbf{q}, \dot{\mathbf{q}}) = \dot{\mathbf{q}}^T\mathbf{M}_q(\mathbf{q})\dot{\mathbf{q}}/2$ and $T(\mathbf{q}, \mathbf{u}') = \mathbf{u}'^T\mathbf{M}'_u(\mathbf{q})\mathbf{u}'/2$, and in both cases $V(\mathbf{q})$. At a given instant of time we can then write respectively: $\Psi_E(\mathbf{p}, \mathbf{v})$, $\Psi_E(\mathbf{p}, \mathbf{u})$, $\Psi_E(\mathbf{q}, \dot{\mathbf{q}})$ and $\Psi_E(\mathbf{q}, \mathbf{u}')$, and treat condition (28) as an additional nonlinear nonholonomic constraint on the system.

Let us focus for a while on the dependent variable formulations (in terms of $\mathbf{p}$ and $\mathbf{v}$), and assume that the system's degree of freedom is equal to one, $r = n - m = 1$. Having the kinematic constraint violations removed, $\boldsymbol{\Phi}(\overline{\mathbf{p}}) = \mathbf{0}$ and $\boldsymbol{\Psi}(\overline{\mathbf{p}}, \overline{\mathbf{v}}) = \mathbf{0}$, the energy constraint violation $\overline{\Psi}_E = \Psi_E(\overline{\mathbf{p}}, \overline{\mathbf{v}}) \neq 0$ means that, compared to the 'exact' solution, either the calculated position of the system on the constraint manifold is somewhat advanced/backward, the system velocity is a little increased/decreased, or both. The simplest way of elimination of the energy constraint violation is to correct the system velocity $\overline{\overline{\mathbf{v}}} = \overline{\mathbf{v}} + \Delta\overline{\mathbf{v}}$ so that $\Psi_E(\overline{\mathbf{p}}, \overline{\overline{\mathbf{v}}}) = 0$. The correction suggested in [20] is $\Delta\overline{\mathbf{v}} = -\mathbf{C}_E^T(\mathbf{C}_E\mathbf{C}_E^T)^{-1}\overline{\Psi}_E$, where $\mathbf{C}_E = \partial\Psi_E/\partial\mathbf{v}$ can be interpreted as an energy constraint vector. As motivated before, and by analogy to Equation (27), the scheme should rather be $\Delta\overline{\mathbf{v}} = -\mathbf{M}^{-1}\mathbf{C}_E^T(\mathbf{C}_E\mathbf{M}^{-1}\mathbf{C}_E^T)^{-1}\overline{\Psi}_E$. Then, since $\mathbf{C}_E = \mathbf{v}^T\mathbf{M}$, the correction simplifies to $\Delta\overline{\mathbf{v}} = -(\overline{\Psi}_E/2\overline{T})\overline{\mathbf{v}}$, where $2\overline{T} = \overline{\mathbf{v}}^T\mathbf{M}(\overline{\mathbf{p}})\overline{\mathbf{v}}$. However, the consequent corrected velocity $\overline{\overline{\mathbf{v}}} = (1 - \overline{\Psi}_E/2\overline{T})\overline{v}$ does not assure exact elimination of the energy constraint violation. Namely, using $\overline{\overline{\mathbf{v}}}$, after some manipulations, one arrives at $\Psi_E(\overline{\mathbf{p}}, \overline{\overline{\mathbf{v}}}) = \overline{\Psi}_E^2/4\overline{T} \neq 0$. The exact satisfaction of the energy constraint can then be achieved recursively [20].

In order to develop the above scheme to one-step elimination scheme of the energy constraint violation, let us introduce a modified corrected velocity $\overline{\overline{\mathbf{v}}} = (1 - \overline{\Psi}_E/2\overline{T} - \varepsilon)\overline{\mathbf{v}}$. Substituting $\overline{\overline{\mathbf{v}}}$ into $\Psi_E(\overline{\mathbf{p}}, \overline{\overline{\mathbf{v}}}) = 0$, one obtains $\overline{T}\varepsilon^2 + (\overline{\Psi} - 2\overline{T})\varepsilon + (\overline{\Psi}^2/4\overline{T}) = 0$, and then $\varepsilon = 1 - \overline{\Psi}/2\overline{T} - (1 - \overline{\Psi}_E/\overline{T})^{1/2}$. The improved velocity correction scheme is finally

$$\overline{\overline{\mathbf{v}}} = \sqrt{1 - \overline{\Psi}_E/\overline{T}}\,\overline{\mathbf{v}}. \tag{29a}$$

Using formula (29a), one can easily check that $\Psi_E(\overline{\mathbf{p}}, \overline{\overline{\mathbf{v}}}) = 0$. By analogy, for the mixed dependent-independent formulation (17), and the independent formulations (19) and (22), the correction schemes are respectively:

$$\overline{\overline{\mathbf{u}}} = \sqrt{1 - \overline{\Psi}_E/\overline{T}}\,\overline{\mathbf{u}}; \tag{29b}$$

$$\dot{\overline{\overline{\mathbf{q}}}} = \sqrt{1 - \overline{\Psi}_E/\overline{T}}\,\dot{\overline{\mathbf{q}}}; \tag{29c}$$

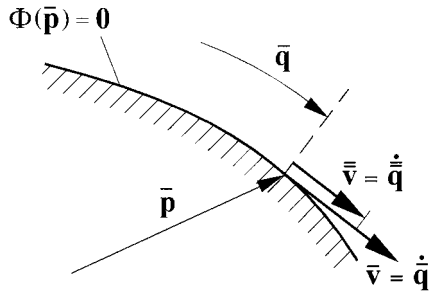

Figure 3. Illustration of velocity correction for elimination of energy constraint violation (a holonomic system case).

$$\overline{\overline{\mathbf{u}}}' = \sqrt{1 - \overline{\Psi}_E / \overline{T}}\,\overline{\mathbf{u}}', \tag{29d}$$

where $\overline{\Psi}_E$ and $\overline{T}$ are respective functions as described in Equation (28).

The geometrical interpretation of the energy constraint violation elimination, for a system with H constraints, is illustrated in Figure 3. Note that the correction affects only the velocity of the system, the position on the constraint manifold is not directly corrected. An evident drawback of the energy constraint control schemes is the precision of evaluation of $\overline{\Psi}_E$ for nonconservative systems ($\dot{E} \neq 0$). The accurate determination of the current energy input due to the external forces, possibly velocity, position and time dependent, may be a cumbersome task [11]. There are, however, two more essential objections raised against the energy correction as a means to improve accuracy of the numerical solution:

1. For many-degrees-of-freedom systems we deal with one condition (28) on r independent velocities, and the 'proportional' correction (29)/(30) of the velocities may not secure the 'exact' motion satisfaction.
2. At a given instant of time, the current total energy E is the sum of the kinetic T and potential V energies, the former depending on both the current system velocity and position, the latter on the position only. It may then happen that inaccuracies in T and V for the current numerical state of motion, compared to their exact values, will be of opposite signs and very similar in values. The resulting energy constraint violation may thus be very small compared to the inaccuracies in T and V. The velocity correction required to eliminate the energy constraint violation may therefore be insignificant for improving the accuracy of the numerical solution in the abovementioned sense.

In sum, the violation of energy constraint (28), which is caused by the inaccuracy of numerical integration process, may be a deficient measure of inaccuracy of the numerical solution. Consequently, the elimination of the violation may not improve the accuracy considerably. In the author's opinion, the conclusion can also be extended to the energy preserving integrators of equations of motion, including the

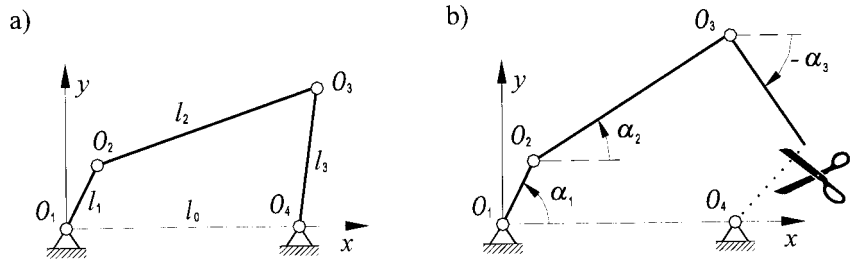

Figure 4. The four-bar mechanism (a) and its open-loop representation (b).

popular non-dissipative implicit trapezoidal rule of Newmark method [11, 12, 18]. We will come back to this problem and illustrate it with test experiments in the following.

4. Test Simulations

Results of numerical simulation of two simple systems are examined. The first system is a constrained system to illustrate the effectiveness of the proposed schemes for elimination of geometric and energy constraints. The other example is to illustrate a possible deficiency of energy constraint violation as a measure of inaccuracy of the numerical solution.

4.1. EXAMPLE 1

A four-bar mechanism shown in Figure 4a is governed by gravity force, i.e. it is a conservative system. The open-loop mechanism representation (Figure 4b) is characterized by three joint coordinates $\mathbf{p} = [\alpha_1 \ \alpha_2 \ \alpha_3]^T$, and the kinematic and dynamic equations of the interim system, (1) and (2), are $\dot{\mathbf{p}} = \mathbf{v}$ and $\mathbf{M}(\mathbf{p})\dot{\mathbf{v}} + \mathbf{d}(\mathbf{p}, \mathbf{v}) = \mathbf{f}(\mathbf{p})$, where

$$
\begin{aligned}
M_{1,1} &= J_1 + (0.25m_1 + m_2 + m_3)l_1^2, \quad M_{2,2} = J_2 + (0.25m_2 + m_3)l_2^2,\\
M_{1,2} &= (0.5m_2 + m_3)l_1 l_2 \cos(\alpha_2 - \alpha_1), \quad M_{2,3} = 0.5m_3 l_2 l_3 \cos(\alpha_3 - \alpha_2),\\
M_{1,3} &= 0.5m_3 l_1 l_3 \cos(\alpha_3 - \alpha_1), \quad M_{3,3} = J_3 + 0.25m_3 l_3^2,
\end{aligned}
$$

$$
\begin{aligned}
d_1 &= -(0.5m_2 + m_3)l_1 l_2 \dot{\alpha}_2^2 \sin(\alpha_2 - \alpha_1) - 0.5m_3 l_1 l_3 \dot{\alpha}_3^2 \sin(\alpha_3 - \alpha_1),\\
d_2 &= (0.5m_2 + m_3)l_1 l_2 \dot{\alpha}_1^2 \sin(\alpha_2 - \alpha_1) - 0.5m_3 l_2 l_3 \dot{\alpha}_3^2 \sin(\alpha_3 - \alpha_2),\\
d_3 &= 0.5m_3 l_1 l_3 \dot{\alpha}_1^2 \sin(\alpha_3 - \alpha_1) + 0.5m_3 l_2 l_3 \dot{\alpha}_2^2 \sin(\alpha_3 - \alpha_2),
\end{aligned}
$$

$$
\begin{aligned}
f_1 &= -(0.5m_1 + m_2 + m_3)g l_1 \cos\alpha_1,\\
f_1 &= -(0.5m_2 + m_3)g l_2 \cos\alpha_2,\\
f_1 &= -0.5m_3 g l_3 \cos\alpha_3.
\end{aligned}
$$

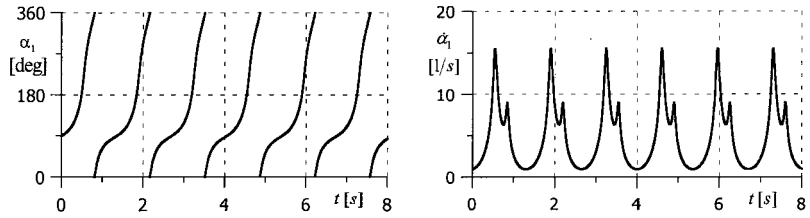

Figure 5. Evolution of the position and velocity of link 1.

The closing constraint equations (1) are

$$\mathbf{\Phi}_{\mathrm{H}} = \begin{bmatrix} l_1 \cos\alpha_1 + l_2 \cos\alpha_2 + l_3 \cos\alpha_3 - l_0 \\ l_1 \sin\alpha_1 + l_2 \sin\alpha_2 + l_3 \sin\alpha \end{bmatrix} = \mathbf{0}$$

and the differential constraint equations (5) and (6) are defined by

$$\mathbf{C} = \begin{bmatrix} -l_1 \sin\alpha_1 & -l_2 \sin\alpha_2 & -l_3 \sin\alpha_3 \\ l_1 \cos\alpha_1 & l_2 \cos\alpha_2 & l_3 \cos\alpha_3 \end{bmatrix},$$

$$\boldsymbol{\xi} = \begin{bmatrix} l_1\dot{\alpha}_1^2 \cos\alpha_1 + l_2\dot{\alpha}_2^2 \cos\alpha_2 + l_3\dot{\alpha}_3^2 \cos\alpha_3 \\ l_1\dot{\alpha}_1^2 \sin\alpha_1 + l_2\dot{\alpha}_2^2 \sin\alpha_2 + l_3\dot{\alpha}_3 62 \sin\alpha_3 \end{bmatrix}.$$

In this way, the governing equations (8) are completed.

The mechanism parameters used in calculations were:

$$l_0 = 1\ \mathrm{m} \quad l_1 = 0.3\ \mathrm{m}, \quad l_2 = 1\ \mathrm{m}, \quad l_3 = 0.6\ \mathrm{m},$$

$$m_1 = 1.5\ \mathrm{kg}, \quad m_2 = 5\ \mathrm{kg}, \quad m_3 = 3\ \mathrm{kg},$$

while the moments of inertia were $J_{C_i} = m_i l_1^2/12$, with C_i at the link midpoints ($i = 1, 2, 3$).

With the initial conditions defined by $\alpha_{10} = \pi/2$ and $\dot{\alpha}_{10} = 1\ \mathrm{sec}^{-1}$, using Runge–Kutta fourth-order algorithm with a very small step size ($\Delta t_r = 0.0005$ sec), first we obtained a reference ('numerically exact') solution shown in Figure 5, with negligible geometric and energy constraint violation for the simulation time period from 0 to 8 sec. Then applying the time step $\Delta t = 0.05$ sec, test simulations have been carried out for:

(a) no constraint control,
(b) Baumgarte's stabilization (23) with $\alpha = 10$ and $\beta = 25$ ($\beta = \alpha^2/4$),
(c) the projective elimination (25)/(27) of geometric constraints violation,
(d) the projective elimination (25)/(27) of geometric constraints violation as above plus the energy constraint elimination (28).

The time-variations of kinematic constraint violations are shown in Figure 6. As seen, the simulation of motion with no constraint control, case (a), leads to abrupt

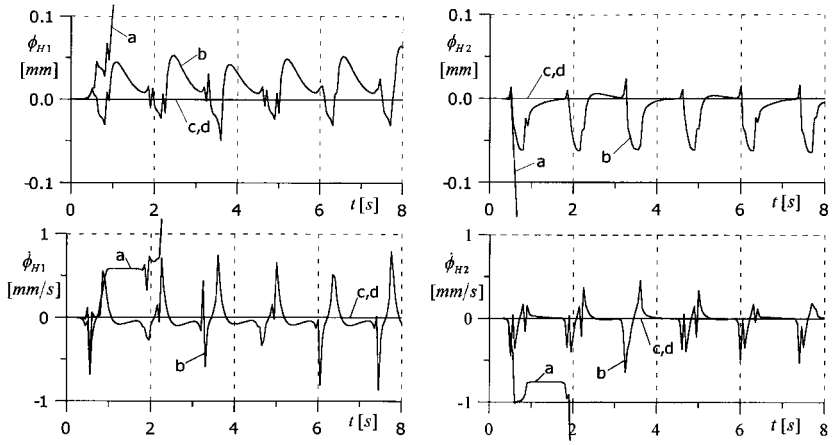

Figure 6. Violations of kinematic constraints.

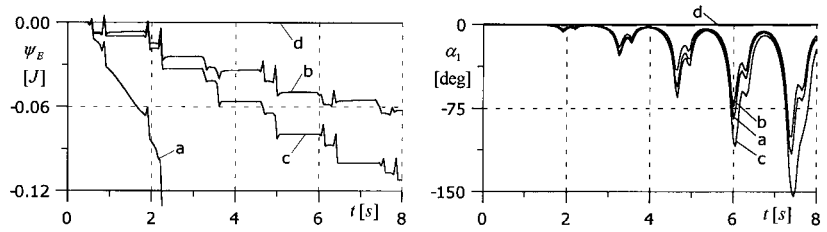

Figure 7. Violation of energy constraint and inaccuracy in position of link 1.

constraint violation. Baumgarte's stabilization, case (b), confines the violations to some limits. The projective method, cases (c) and (d), assures no constraint violation (the violations are less than 10^{-15}). During the whole simulation period the number of evaluation for the position constraint satisfaction was at most two, and the elimination of velocity constraint violations were then achieved in one evaluation.

The violation of energy constraint and the difference between the calculated and the reference solutions for link 1 position are shown in Figure 7. As seen, only the combination of both kinematic and energy constraint elimination, case (d), assures an accurate numerical integration. The author wants to bring out into strong relief, however, that one should not be very enthusiastic about the result. The case studied above is very specific, let us say: favorable. Namely, according to the author's experience, the numerical solution of equations of motion (for both position and velocity) is 'delayed' as compared to the exact solution, which yields that the current numerical values of the kinetic and potential energies differ from their exact values. In the case at hand, in that parts of motion when the mechanism raises up (when the link 1 moves from approximately $-\pi/2$ to $\pi/2$) the inaccuracies in T and V are both negative. The elimination (29) of energy constraint violation 'accelerate' then the numerical solution, and the accuracy of the corrected numerical solution

can be achieved. The situation is different when the mechanism goes down. A delayed numerical solution causes a negative inaccuracy in T while the inaccuracy in V is positive. In these conditions the velocity correction (29) eliminates only the difference between the inaccuracies in T and V, and the solution accuracy (the difference between the exact and the numerical solutions) may not be improved. Thus, the simulated motion of the mechanism is a succession of periods in which the solution accuracy recurrently improves (owing to the velocity correction (29) which works) and spoils due to the numerical truncation errors (when the velocity correction (29) is ineffective). Since the latter process in disturbing the exact solution is rather slow (see Figure 7), and the correction (29) is very effective, the global accuracy of numerical solution is maintained.

Above we described a privileged situation in which the use of energy constraint violation as a means to improve accuracy of numerical integration if effective. This does not hold in a general case, however. In the following we report a very simple numerical experiment to show that the energy constraint violation criterion may be defective.

4.2. EXAMPLE 2

Consider the equation of motion that governs the linear free vibration of a single degree of freedom system

$$m\ddot{x} + kx = 0 \;\Rightarrow\; \ddot{x} + \omega^2 x = 0,$$

where $\omega^2 = k/m$. The analytical closed form solution of the equation is

$$x(t) = A_1 \cos \omega t + A_2 \sin \omega t$$

with $A_1 = x_0$ and $A_2 = \dot{x}_0/\omega$. Having the exact reference solution we can estimate accuracy of numerical integration. The state form of the equation of motion used in numerical simulations were the following two first-order ODEs in x and v:

$$\dot{x} = v; \quad \dot{v} = -\omega^2 x.$$

The system data and initial state of motion were:

$$m = 1 \text{ kg}, \quad k = 10 \text{ N/M}, \quad x_0 = 0.1 \text{ m}, \quad v_0 = 0 \text{ m/s}.$$

Using Runge–Kutta fourth-order algorithm with the time step $\Delta t = 0.1$ sec, two test simulation have been carried out:

(a) without energy constraint violation control,
(b) with the energy constraint violation elimination (29).

Then, using the time step $\Delta t = 0.025$ sec, we applied

(c) the trapezoidal rule of Newmark method

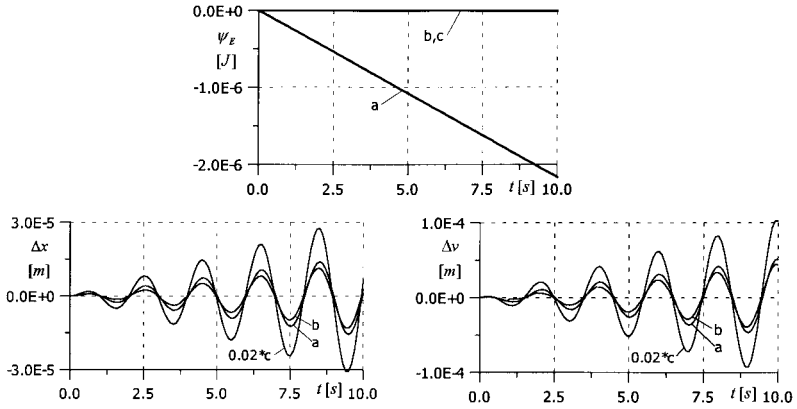

Figure 8. The energy constraint violation and inaccuracy in numerical solution.

to integrate the equation of motion. The trapezoidal rule was chosen for it is energy preserving for linear unconstrained problems [10–12, 18]. The classical scheme of this type, for the case at hand, yields the following system of linear equations in x_{n+1} and v_{n+1}:

$$\begin{aligned} x_{n+1} - x_n &= \Delta t(v_{n+1} + v_n)/2, \\ v_{n+1} - v_n &= -\Delta t\omega^2(x_{n+1} + x_n)/2. \end{aligned}$$

By solving the equations, the solution can be advanced from time t_n to $t_{n+1} = t_n + \Delta t$.

As seen from the plots in Figure 8, both the energy constraint violation elimination scheme (29) (case b) and the trapezoidal rule (case c) assure the energy conservation. The exact numerical solution is not assured, however. The application of the correction scheme (29) improves the accuracy inconsiderably, while the inaccuracy of numerical integration with the use of Newmark scheme is 100 times bigger as compared to the solution with no energy control obtained with the Runge–Kutta method.

The simultaneous energy preservation and solution inaccuracy can be explained following the previous comments. As said, the numerical solution, both in position and velocity, is 'delayed' compared to the exact solution. This yields differences in current numerical values of T and V as compared to the exact values, ΔT and ΔV. The signs of ΔT and ΔV are opposite while their absolute values are close to each other. As a consequence, the energy constraint violation, $\Delta E = \Delta T + \Delta V$ may be relatively small compared to the absolute values of ΔT and ΔV. The situation, for the numerical integration with no constraint control, is illustrated in the first plot in Figure 9. The correction scheme (29) results then in vanishing ΔE, while ΔT and ΔV remain almost unchanged (the second plot in Figure 9). A similar situation is observed for the Newmark integration scheme, and the ΔT and ΔV are much bigger in values (the third plot in Figure 9).

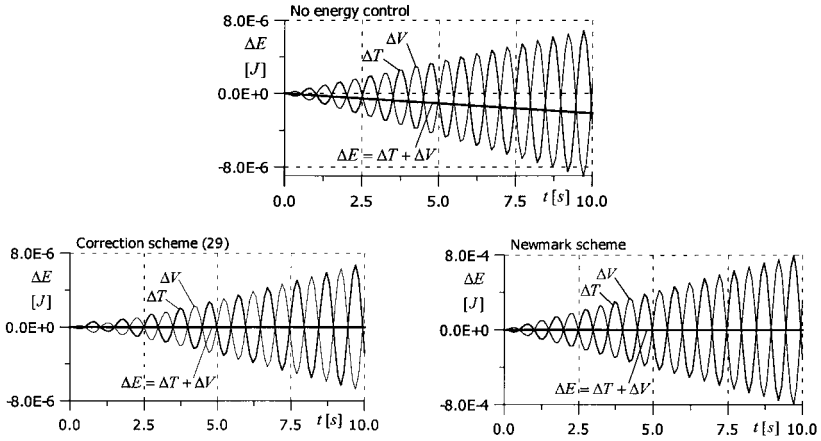

Figure 9. The kinetic and potential energies violations.

Only these simple numerical experiments show that the total energy constraint (28) may be an unreliable criterion for the inaccuracy measure of numerical integration of equations of motion. Consequently, the accuracy of numerical simulation may not improve considerably by elimination of the energy constraint violation according to scheme (29) (neither by application of an energy preserving integration algorithm). The uncertainty of the energy criterion for many degree of freedom systems is a separate problem, and an accurate determination of the energy input rate (for nonconservative systems) may be difficult in practical applications.

5. Concluding Remarks

The simplest way to improve accuracy of numerical simulation of multibody systems is to diminish the integration time step. However, using a very small step size may make the analysis numerically expensive. This may exclude the real-time simulations, either, and may still be insecure for very long simulation times. Applying larger integration steps improves the numerical integration efficiency, but aggravates at the same time the numerical inaccuracy.

In the case of constrained systems an important measure of numerical integration inaccuracy is the kinematic constraint violation. Leaving the constraint violation uncontrolled may impair reliability of the numerical simulation. In this paper three virtually equivalent schemes (25), (26) and (27), associated with different formulations of constrained system dynamics, have been developed for the numerically exact elimination of the constraint violations. The elimination is achieved by appropriate correction of state of the system, and the procedures assure that the corrections are performed in the directions orthogonal to the constraint manifolds. In this way, the system effective (constraint consistent) motion is not disturbed by the corrections.

The other consequence of numerical truncation errors is inaccuracy in the constraint consistent solution to equations of motion, the problem related to both constrained and unconstrained system dynamics formulations. A seemingly reasonable measure of the inaccuracy is the energy constraint violation (28). In this paper we develop the projective scheme proposed in [20] for elimination of the constraint violation. The proposed procedures (29) assure that the violation can be removed in one step by simple velocity correction. On the other hand, we argue the energy constraint violation to be a possibly defective measure of the numerical solution inaccuracy. Consequently, the removal of energy constraint violation may not remarkably improve the solution accuracy. The assertion is justified first by theoretical considerations and then numerical experiments. It is also demonstrated that the application of the energy preserving integration Newmark algorithm may not improve the accuracy of numerical simulation, either.

References

1. Amirouche, F.M.L., *Computational Methods for Multibody Dynamics*, Prentice-Hall, Englewood Cliffs, NJ, 1992.
2. Bae, D.-S. and Yang, S.-M., 'A stabilization method for kinematic and kinetic constraint equations', in *Real-Time Integration Methods for Mechanical System Simulation*, E.J. Haug and R.C. Deyo (eds), NATO ASI Series, Vol. F69, Springer-Verlag, Berlin, 1990, 209–232.
3. Baumgarte, J., 'Stabilization of constraints and integrals of motion in dynamical systems', *Computer Methods in Applied Mechanics and Engineering* **1**, 1972, 1–16.
4. Bayo, E. and Ledesma, R., 'Augmented Lagrangian and mass-orthogonal projection methods for constrained multibody dynamics', *Nonlinear Dynamics* **9**, 1996, 113–130.
5. Blajer, W., 'A geometric unification of constrained system dynamics', *Multibody System Dynamics* **1**, 1997, 3–21.
6. Blajer, W., 'A geometrical interpretation and uniform matrix formulation of multibody system dynamics', *Zeitschrift für angewandte Mathematik und Mechanik* **81**, 2001, 247–259.
7. Blajer, W., Schiehlen, W. and Schirm, W., 'A projective criterion to the coordinate partitioning method for multibody dynamics', *Archive of Applied Mechanics* **64**, 1994, 215–222.
8. Campbell, S.L. and Meyer, C.D. Jr., *Generalized Inverses of Linear Transformations*, Pitman, London, 1979.
9. Chang, C.O. and Nikravesh, P.E., 'An adaptive constraint violation stabilization method for dynamic analysis of mechanical systems', *Journal of Mechanisms, Transmissions, and Automation in Design* **107**, 1985, 488–492.
10. Chen, S., Handen, J.M. and Tortorelli, D.A., 'Unconditionally energy stable implicit time integration: Application to multibody system analysis and design', *International Journal for Numerical Methods in Engineering* **48**, 2000, 791–822.
11. Chung, J. and Hulbert, G.M., 'A time integration algorithm for structural dynamics with improved numerical dissipation: The generalized-α method', *Journal of Applied Mechanics* **60**, 1993, 371–375.
12. García de Jalón, J. and Bayo, E., *Kinematic and Dynamic Simulation of Multibody Systems*, Springer-Verlag, New York, 1994.
13. Kane, T.R. and Levinson, D.A., *Dynamics: Theory and Applications*, McGraw-Hill, New York, 1986.
14. Lin, S.-T. and Hong, M.-C., 'Stabilization method for numerical integration of multibody mechanical systems', *Journal of Mechanical Design* **120**, 1998, 565–572.

15. Neimark, J.I. and Fufaev, N.A., *Dynamics of Nonholonomic Systems*, Moscow University Publishers, Moscow, 1967 [in Russian].
16. Ostermeyer, G.-P., 'On Baumgarte stabilization for differential algebraic equations', in *Real-Time Integration Methods for Mechanical System Simulation*, E.J. Haug and R.C. Deyo (eds), NATO ASI Series, Vol. F69, Springer-Verlag, Berlin, 1990, 193–207.
17. Park, K.C. and Chiou J.C., 'Stabilization of computational procedures for constrained dynamical systems', *Journal of Guidance, Control, and Dynamics* **11**, 1988, 365–370.
18. Simo, J.C. and Wong, K.K., 'Unconditionally stable algorithms for rigid body dynamics that exactly preserve energy and momentum', *International Journal for Numerical Methods in Engineering* **31**, 1991, 19–52.
19. Wehage, R.A. and Haug, E.J., 'Generalized coordinate partitioning for dimension reduction in analysis of constrained dynamic systems', *Journal of Mechanical Design* **104**, 1982, 247–255.
20. Yoon, S., Howe, R.M. and Greenwood, D.T., 'Geometric elimination of constraint violations in numerical simulation of Lagrangian equations', *Journal of Mechanical Design* **116**, 1994, 1058–1064.
21. Yoon, S., Howe, R.M. and Greenwood, D.T., 'Stability and accuracy analysis of Baumgarte's constraint violation stabilization method', *Journal of Mechanical Design* **117**, 1995, 446–453.