FORMULATION OF CONSTRAINED SYSTEM DYNAMICS 
BY ORTHONORMALIZING THE CONFIGURATION SPACE

Wojciech Blajer

Department of Mechanics
Technical University of Radom

An automatic computer code for constructing an orthonormal and differentiable basis of tangent (null) subspace for constrained mechanical systems is proposed. The method uses the Gram-Schmidt vector orthogonalization process, adopted according to the Riemannian space formalism. An interesting and useful peculiarity of the formulation is that the minimal-order (purely kinetic) equations of motion are generated directly in the resolved form (the related mass matrix is the identity matrix). The other problems solved are: generation of a well-posed and sparse supplementary matrix to the constraint matrix, used in the orthonormalization process; diminishing the constraint violation due to numerical errors of integration; estimation of consistent initial values of tangent velocities; and effective determination of constraint reactions.

1. Introduction

A prevalent approach to the dynamic analysis of constrained mechanical systems is to pose initially the problem in terms of a redundant set of coordinates. The related equations of motion are mixed differential-algebraic equations (DAEs). Since the numerical treatment of the DAEs is usually inefficient, several methods have arisen that automatically generate a set of independent variables and reduce the equations of motion to a minimal order for solution. The coordinate partitioning (LU decomposition) method (cf Wehage and Haug, 1982), the zero eigenvalues theorem method (cf Walton and Steeves, 1969; Kamman and Huston, 1984) and closely related singular value decomposition (SVD) method (cf Mani et al., 1985; Singh and Likins, 1985), and methods based on the Householder transformations (cf Kim and Vanderplaeg, 1986; Amirouche et al., 1988) and on the Gram-Schmidt orthogonali-
zation (GSO) process (cf Liang and Lance, 1987: Agrawal and Saigal, 1989) are representative examples of methods of this type. This paper is another contribution to the field.

The idea used herein originates from the works by Liang and Lance (1987), Agrawal and Saigal (1989), where the GSO formulae were used to generate bases of tangent (null) space. As opposed to previous applications, however, the present scheme involves the metric tensor of the system configuration space into the definition of vector length and into the definition of dot product of two vectors. Consequently, a genuine orthonormal basis of tangent space can be built. Projecting the initial dynamic equations into the orthonormal basis of tangent space and expressing them in terms of the corresponding tangent velocities results in a minimal-order set of kinetic equations of motion. An important and useful characteristic of the equations is that the related mass matrix (metric tensor matrix of the orthonormal basis of tangent space) is the unit matrix. In other words, resolved kinetic equations are directly obtained. The other problems solved are: generation of a well-posed and sparse supplementary matrix to the constraint matrix, used in the orthonormalization process; diminishing the constraint violation due to numerical errors of integration; estimation of consistent initial values of tangent velocities; and effective determination of constraint reactions.

2. Preliminary definitions

In order to define properly a vector length and a dot product of two vectors in an n-dimensional configuration space of a system, the critical observation is that the space is a Riemannian space. Accordingly, the concepts and statements of the Riemannian geometry should be used to study dynamics of constrained/multibody systems. Mechanicians make usually little account of these aspects for many problems of mechanics can be described and solved without involving the mathematical tool of Riemannian geometry. Some specialized problems, as the problem undertaken in this paper, will require the use of the tool, however, and ignoring this fact may lead to inconsistencies in mathematical formulation.

The notation used hereafter for vector space and tensor calculus concepts patterns upon Pobedrya (1972) and the constrained system dynamics formulation developed by Blajer (1992a,b); it has also some reference to Papastavridis (1990), Brauchli (1991), Maißer (1991). For details the reader is referred to these works. Here, of special use there will be the distinction between the
contravariant, \( \mathbf{a} = [a_1, \ldots, a_n]^T \), and the covariant, \( \mathbf{a}^* = [a^*_1, \ldots, a^*_n]^T \), respectively, representations of a vector \( \mathbf{a} \) in the \( n \)-dimensional Riemannian space \( \mathcal{R} \), and the interdependence between them is

\[
\mathbf{a}^* = \mathbf{M} \mathbf{a}
\]  

(2.1)

where \( \mathbf{M} \) is the metric tensor (symmetric, positive definite) matrix of the space, referred to the covariant basis in which \( \mathbf{a} \) is defined. Then, a dot product of two vectors \( \mathbf{a} \) and \( \mathbf{b} \) can be written in four possible ways

\[
(\mathbf{a}, \mathbf{b}) = \mathbf{a}^\mathsf{T} \mathbf{M} \mathbf{b} = \mathbf{a}^\mathsf{T} \mathbf{b}^* = \mathbf{a}^* \mathbf{M}^{-1} \mathbf{b}^* = \mathbf{a}^* \mathbf{b}
\]  

(2.2)

and the vector orthogonality condition is \( (\mathbf{a}, \mathbf{b}) = 0 \). A vector length is defined, according to Eq (2.2), as

\[
\|\mathbf{a}\| = \sqrt{(\mathbf{a}, \mathbf{a})}
\]  

(2.3)

Using the above notation, for a system characterized by \( n \) generalized coordinates \( \mathbf{x} = [x_1, \ldots, x_n] \), and subject to \( m \) \((m < n)\) stationary holonomic constraints \( \mathbf{f} = [f_1, \ldots, f_n] \)

\[
\mathbf{f}(\mathbf{x}) = 0
\]  

(2.4)

the dynamic equations of constrained motion can be represented as follows

\[
\mathbf{M}(\mathbf{x}) \ddot{\mathbf{x}} = \mathbf{h}^*(\mathbf{x}, \dot{\mathbf{x}}, t) + \mathbf{C}^*(\mathbf{x}) \lambda
\]  

(2.5)

where \( \mathbf{M} \) is the \( n \times n \) symmetric positive definite mass matrix (metric tensor matrix of the covariant base vectors spanning the directions of \( \dot{\mathbf{x}} \) components); \( \mathbf{h}^* = [h^*_1, \ldots, h^*_n] \) represents the generalized forces; \( \mathbf{C}^* = (\partial \mathbf{f}/\partial \mathbf{x})^\mathsf{T} \) is the \( n \times m \) constraint Jacobian matrix; \( \lambda = [\lambda_1, \ldots, \lambda_m]^T \) contains the corresponding Lagrange multipliers; and \( t \) represents time. The constraint vectors \( \mathbf{c}_i \) \((i = 1, \ldots, m)\), represented by covariant components as columns of \( \mathbf{C}^* \), span (form a basis of) the \( m \)-dimensional constrained subspace \( \mathcal{C} \) in \( \mathcal{R} \). By assumption, no virtual motion is allowed in the subspace, and the reaction \( \mathbf{r}_i \) of the \( i \)th constraint is spanned along the corresponding constraint vector, \( \mathbf{r}_i^* = \mathbf{c}_i^* \lambda_i \).

Blajer (1992a,b) gave more explanation concerning the above governing equations in the descriptor (DAE) form for a constrained mechanical system. The reasons for using contravariant/covariant vector components are also better detailed there. Moreover, the formulation of those papers is valid for systems subject to holonomic and/or nonholonomic constraints, and for the dynamic analysis carried out in generalized velocities and/or quasi-velocities.
For reasons of simplicity, the formulation of this paper is limited to systems subject to holonomic constraints and to the analysis carried out in generalized velocities, which suffices usually to model and analyse a wide range of problems of multibody dynamics. Nevertheless, the following formulation can easily be modified so as to model and analyse the aforementioned more general cases of mechanics.

The crux of the methods that reduce the DAEs (2.4) and (2.5) to a minimal set of ODEs is constructing an \( n \times k \) full-rank matrix \( \mathbf{D} \) so that the velocity and acceleration equations of constraints (2.4)

\[
\begin{align*}
\mathbf{\ddot{f}} &= \mathbf{C}^* \mathbf{T} \mathbf{\ddot{x}} = \mathbf{0} \\
\mathbf{\dddot{f}} &= \mathbf{C}^* \mathbf{T} \mathbf{\ddot{x}} + \mathbf{\dddot{C}}^* \mathbf{T} \mathbf{\dot{x}} = \mathbf{0}
\end{align*}
\]  

are satisfied when expressed in \( k = n - m \) tangent (independent) velocities \( \mathbf{u} = [u_1, ..., u_k]^T \) defined as

\[
\mathbf{\dot{x}} = \mathbf{D}(\mathbf{x}) \mathbf{u}
\]

In other words, after substituting the relationship (2.8) into Eqs (2.6) and (2.7), the condition of \( \mathbf{\ddot{f}} = \mathbf{0} \) and \( \mathbf{\dddot{f}} = \mathbf{0} \) implies that

\[
\mathbf{C}^* \mathbf{T} \mathbf{D} = \mathbf{0} \quad \text{and} \quad \mathbf{D}^T \mathbf{C}^* = \mathbf{0}
\]

i.e. \( \mathbf{D} \) is an orthogonal complement matrix to \( \mathbf{C}^* \) in \( \mathcal{R} \). Mathematically, Eq (2.6) expresses dot products of the constraint vectors with the velocity vector, \( (c_i, \mathbf{\dot{x}}) = 0 \) \( (i = 1, ..., m) \), and this justifies the covariant representation of column vectors of \( \mathbf{C}^* \). Then, Eq (2.9) expresses \( (c_i, d_j) = 0 \) \( (i = 1, ..., m; \; j = 1, ..., k) \), where the vectors \( d_j \), represented by contravariant components as columns of \( \mathbf{D} \), span (form a basis of) the \( k \)-dimensional tangent (null) space \( \mathcal{T} \) which complements \( \mathcal{C} \) in \( \mathcal{R} \), \( \mathcal{T} \cup \mathcal{C} = \mathcal{R} \) and \( \mathcal{T} \cap \mathcal{C} = \mathbf{0} \).

Projecting the dynamic equations (2.5) into \( \mathcal{T} \) defined by \( d_1, ..., d_k \), and expressing the equations in terms of \( \mathbf{u} \), \( k \) kinetic equations of the constrained motion follow

\[
\mathbf{M}_u(\mathbf{x}) \mathbf{\ddot{u}} = \mathbf{h}_u^*(\mathbf{x}, \mathbf{u}, t)
\]

where

\[
\begin{align*}
\mathbf{M}_u &= \mathbf{D}^T \mathbf{M} \mathbf{D} \\
\mathbf{h}_u^* &= \mathbf{D}^T (\mathbf{h}^* - \mathbf{M} \mathbf{D} \mathbf{u})
\end{align*}
\]

The mass matrix \( \mathbf{M}_u \) of the kinetic equations (2.10) is the metric tensor matrix of \( \mathcal{T} \), referred to the covariant basis formed by \( d_1, ..., d_k \), and Eq (2.11) can
be interpreted as $M(i, j) = (d_i, d_j)$, $i, j = 1, ..., k$. As seen in Eq (2.12), the regular derivation of kinetic equations requires that $\dot{D}$ must be found as well.

Since the relationship (2.8) is, in general, nonintegrable (the integral of $u$ may have no physical meaning), the minimal-order governing equations of the constrained motion are then composed of Eqs (2.8) and (2.10).

3. Revised orthonormalization process

Application of the GSO process to generation of a tangent subspace basis requires that the constraint matrix $C^*$ must be appended to an $n \times k$ matrix $E^*$ so that the produced $n \times n$ matrix $P^*$

$$P^* = [C^* E^*] \quad (3.1)$$

is of maximal rank. Formulation of $E^*$ is not unique, and the task can be completed by any method. The problem of effective and well-conditioned formulation of $E^*$ is discussed in Section 5.

The condition $\text{rank}(P^*) = n$ assures that the vectors $p_i$ ($i = 1, ..., n$), represented by covariant components as columns of $P^*$, are independent in $R$. They span thus the whole $R$ and form a basis of the space. Consequently, the vectors $e_j$ ($j = 1, ..., k$), represented as columns of $E^*$, cover $T$ but do not, in general, span the subspace. The vectors may give projections into the constrained subspace as well, $E^{*T}M^{-1}C^* \neq 0$. The physical interpretation of $E^*$ is that it stands for the Jacobian matrix of supplementary constraints that freeze the system, i.e. $P^*T\dot{x} = 0$.

Fig. 1. Illustration of the orthonormalization process

Applying a properly modified GSO process to the base vectors $p_i$ ($i = 1, ..., n$), an orthonormal basis of $R$ can be produced. The recursion
formulae of the revised orthonormalization process are

\[ w_i = M^{-1}w_i^* \]

\[ w_i^* = v_i^* \|v_i\| \]

\[ v_1^* = p_1^* \]

\[ v_{i+1}^* = p_{i+1}^* - \sum_{j=1}^{i} (w_j, p_{i+1}) w_j^* \]  \hspace{1cm} (3.2)

\[ \|v_1\|^2 = \|p_1\|^2 \]

\[ \|v_{i+1}\|^2 = \|p_{i+1}\|^2 - \sum_{j=1}^{i} (w_j, p_{i+1})^2 \]

where (it is convenient to use)

\[ \|p_{i+1}\|^2 = (p_{i+1}, p_{i+1}) = p_{i+1}^* M^{-1} p_{i+1}^* \]

\[ (w_j, p_{i+1}) = w_j^* p_{i+1}^* \]

As a result, the matrix \( P^* \) transforms to a matrix \( W \) which, according to Eq (3.1), can be split into

\[ W = [W_c \ W_d] \]  \hspace{1cm} (3.3)

An attribute of the \( n \times m \) matrix \( W_c \) is that its columns \( w_i \) \((i = 1, ..., m)\) are linear combinations of \( c_i^* \) \((i = 1, ..., m)\). The vectors \( w_1, ..., w_m \) span thus \( C \), and form an orthonormal basis of the subspace. Since, by assumption, all the orthonormal vectors \( w_i \) \((i = 1, ..., n)\) are orthogonal to each other, it follows then

\[ W_d^T M W_c = 0 \quad \text{and} \quad W_d^T C^* = 0 \]  \hspace{1cm} (3.4)

\( W_d \) is thus an orthogonal complement matrix to \( C^* \).

Agrawal and Saigal (1989), Liang and Lance (1987) also used the GSO process to produce orthogonal complement matrices. However, since the metric of \( R \) had not been involved in those schemes (the schemes had a form given Eq (3.2) after substituting for \( M \) the identity matrix \( I \), the tangent subspace base vectors obtained were neither orthonormal nor orthogonal to each other in the meaning used in this paper. They would be orthogonal only if the mass matrix \( M \) of the initial dynamic equations (2.5) was \( M = \alpha I \), \( \alpha \) being a real constant, and they would be orthonormal if \( M = I \) (the contravariant and covariant representations of vectors are identical in the latter case). As will be seen in Section 4, the genuine orthonormality of vectors \( w_i \) \((i = 1, ..., n)\) produced by the scheme of Eq (3.2) yields useful peculiarities of the kinetic equations of motion that follow.
In order to obtain the time derivative of $\mathbf{W}$, one can use the following recursion formulae

\[
\dot{\mathbf{w}}_i = (\mathbf{M}^{-1})^i \mathbf{w}_i^* + \mathbf{M}^{-1} \left( \ddot{\mathbf{v}}_i^* \| \mathbf{v}_i^* \|^{-1} + \mathbf{v}_i^*(\| \mathbf{v}_i^* \|^{-1})' \right)
\]

\[
\ddot{\mathbf{v}}_1^* = \dot{\mathbf{p}}_1^*
\]

\[
(\| \mathbf{v}_1^* \|^{-1})' = \frac{1}{2} \| \mathbf{v}_1^* \|^{-\frac{3}{2}} \langle \mathbf{p}_1, \mathbf{p}_1 \rangle
\]

(3.5)

\[
\ddot{\mathbf{v}}_{i+1}^* = \dot{\mathbf{p}}_{i+1}^* - \sum_{j=1}^{i} \left( \langle \mathbf{w}_j, \mathbf{p}_{i+1} \rangle \mathbf{w}_j^* + \langle \mathbf{w}_j, \mathbf{p}_i \rangle \dot{\mathbf{w}}_j^* \right)
\]

\[
(\| \mathbf{v}_{i+1}^* \|^{-1})' = \frac{1}{2} \| \mathbf{v}_{i+1}^* \|^{-\frac{3}{2}} \left( \langle \mathbf{p}_{i+1}, \mathbf{p}_{i+1} \rangle' - 2 \sum_{j=1}^{i} \langle \mathbf{w}_j, \mathbf{p}_{i+1} \rangle \langle \mathbf{w}_j, \mathbf{p}_i \rangle' \right)
\]

where

\[
(\mathbf{M}^{-1})' = -\mathbf{M}^{-1} \dot{\mathbf{M}} \mathbf{M}^{-1}
\]

\[
\langle \mathbf{p}_{i+1}, \mathbf{p}_{i+1} \rangle' = 2 \mathbf{p}_{i+1}^{\mathbf{p}_{i+1}^*} \mathbf{M}^{-1} \mathbf{p}_{i+1}^* + \mathbf{p}_{i+1}^{\mathbf{p}_{i+1}^*} (\mathbf{M}^{-1})' \mathbf{p}_{i+1}^*
\]

\[
(\langle \mathbf{w}_j, \mathbf{p}_{i+1} \rangle)' = \mathbf{w}_j^T \dot{\mathbf{p}}_{i+1}^* + \mathbf{w}_j^T \mathbf{p}_{i+1}^*
\]

Using Eqs (3.2) and (3.5), $\mathbf{W}_d$ and $\dot{\mathbf{W}}_d$, corresponding to $\mathbf{D}$ and $\dot{\mathbf{D}}$ in Eqs (2.10) $\div$ (2.12), can be constructed. For better understanding of the mechanism of the orthonormalization process, one can refer to simple illustration in the two-dimensional Cartesian space ($\mathbf{M} = \mathbf{I}$), shown in Fig.1. The two nonorthogonal vectors $\mathbf{p}_1$ and $\mathbf{p}_2$ are orthonormalized to $\mathbf{w}_1$ and $\mathbf{w}_2$. The direction of $\mathbf{p}_1$ can also represent the constrained subspace (vector $\mathbf{c}$), and the direction of $\mathbf{p}_2$ – the subspace defined by $\mathbf{E}$ (vector $\mathbf{e}$). The vector $\mathbf{w}_2$ is thus orthogonal to the constrained direction, and represents an orthonormal basis of the tangent subspace. The orthonormalization process of Eq (3.12) is a generalization of this simple case to the $n$-dimensional Riemannian space.

4. Peculiarities of the orthonormal basis formulation

As the vectors defined by $\mathbf{W}$ form an orthonormal basis of $\mathcal{R}$, it follows that

\[
\mathbf{W}^T \mathbf{M} \mathbf{W} = \mathbf{W}^T \mathbf{W}^* = \mathbf{W}^* \mathbf{M}^{-1} \mathbf{W}^* = \mathbf{W}^* \mathbf{W} = \mathbf{I}^{(n)}
\]

(4.1)
where $I^{(n)}$ is the $n \times n$ identity matrix (the metric tensor matrix of the orthonormal basis). In other words, for $i, j = 1, \ldots, n$

\[
(w_i, w_j) = \begin{cases} 
1 & \text{for } i = j \\
0 & \text{for } i \neq j 
\end{cases} \tag{4.1a}
\]

Making use of the division of $W$ as in Eq (3.3), or formulating Eq (4.1a) for $i, j = m+1, \ldots, n$, it is evident that

\[
W_d^T M W_d = I^{(k)} \tag{4.2}
\]

and the $k \times k$ identity matrix $I^{(k)}$ is the metric tensor matrix of the orthonormal basis of $T$, defined by $w_{m+1}, \ldots, w_n$. This observation is of paramount importance for the formulation of kinetic equations. Projecting the initial dynamic equations (2.5) into the orthonormal basis of $T$ and expressing them in terms of the corresponding tangent velocities $u$, it follows that $M_u = W_d^T M W_d = I^{(k)}$, and the resolved form of governing equations of the constrained motion is obtained as follows

\[
\dot{x} = W_d(x) u \tag{4.3}
\]

\[
\dot{u} = W_d^T (h^* - M \dot{W}_d u) = h_u^*(x, u, t) \tag{4.4}
\]

Note that the above result cannot be achieved by using the formulations given by Agrawal and Saigal (1989), Liang and Lance (1987). The tangent subspace bases produced there satisfy the condition $W_d^T C^* = 0$, but $M_u(x) = W_d^T M W_d$ are general symmetric positive-definite matrices.

Three more aspects may be of importance when using the approach proposed above. The first is that the formulation of Eqs (4.3) and (4.4) assures, by assumption, maintenance of the higher-order constraint Eqs (2.6) and (2.7); which does not protect the analysis from violation of the constraint Eqs (2.4) due to the numerical errors of integration of Eqs (4.3) and (4.4), however. Even if the initial position of the system satisfies the constraint equations, during the simulation the integrated position $x^i$ may violate the constraints, $f(x^i) \neq 0$, and the violation gradient accelerates usually in the course of time. The same concerns the integration process of Eqs (2.8) and (2.10).

The conceptually simplest method for diminishing the constraint violation is, after each step of integration or a sequence of steps, to treat $x^i$ as a trial root of $f(x) = 0$, and then to solve iteratively for a numerically exact root. As $\text{dim}(f) < \text{dim}(x)$, the following modified Newton-Raphson formula is proposed

\[
\begin{bmatrix}
C^T(x^i) \\
W_d^T(x^i M^T(x^i))
\end{bmatrix}
\Delta x = T^T(x^i) M(x^i) \Delta x = -
\begin{bmatrix}
f(x^i) \\
0
\end{bmatrix} \tag{4.5}
\]
Fig. 2. Reduction of constraint violation

where \( \Delta x = x - x^i \), and \( T = [M^{-1}C^* \ W_d]^\top \) is the transformation matrix between the covariant representations of unit vectors \( k_i \) \((i = 1, ..., n)\) spanning the directions of \( x \) components (the covariant components of vectors are contained in the \( n \times n \) identity matrix, \( K^* = I^{(n)} \), so that \( (x, k_i) = x_i \) and the covariant components of base vectors \( c_1, ..., c_m, w_{m+1}, ..., w_n \); refer to Blajer (1992a,b) for details. The formula (4.5) translates the \( m \)-vector \( f(x^i) \) into the \( n \)-vector \( \Delta x \) required for diminishing the violation, and assures that the system position is corrected in the constrained directions only (the position in \( T \) is not changed). Since \( M_t = T^\top MT = \text{diag}(C^{*\top}M^{-1}C^*, I^{(k)}) \) is the matrix of the metric tensor of basis \( c_1, ..., c_m, w_{m+1}, ..., w_n \), it is easy to show that \( (T^\top M)^{-1} = TM_t^{-1} \), which applied to Eq (4.5) yields

\[
\Delta x = -M^{-1}C^*(C^{*\top}M^{-1}C^*)^{-1}f
\]

Note that \( M_c^{-1}f \) are the contravariant components of the required corrections expressed in the basis of \( C \) defined by \( c_1, ..., c_m \), and \( M_c = C^{*\top}M^{-1}C^* \) is the metric tensor matrix of the basis. The constraint violation measure \( f(x^i) \) corresponds thus to the forces which are orthogonal to the constraint manifold and assure that the system moves on the manifold (cf Arnold, 1978).

The correction of constraint violation can also be included directly into the integration process of Eqs (4.3) and (4.4) by adopting the Baumgarte method (cf Baumgarte, 1972; Östermeyer, 1990). Using a PI-controller scheme, one can write

\[
M_t \begin{bmatrix} 0 \\ u \end{bmatrix} = T^\top M \dot{x} + \begin{bmatrix} G_1 f + G_0 \int f \, dt \\ 0 \end{bmatrix}
\]
where $M_t$ and $T$ have been defined above, and $G_1$ and $G_0$ are diagonal matrices of feedback gains. Then after premultiplying Eq (4.5) by $(T^T M)^{-1}$, one obtains

$$\dot{x} = W_d u - M^{-1} C^*(C^{*T} M^{-1} C^*)^{-1} \left( G_1 f + G_0 \int f dt \right)$$  \hspace{1cm} (4.6a)

The stabilized governing equations are composed then of Eqs (4.6a) and (4.4). It may be worth noting that the stabilizing terms appear in the kinematic equations, whereas they are commonly placed in the dynamic equations (cf Baumgarte, 1972). In the present formulation, however, as the corrections are represented in $C$ only, they will not affect the kinetic Eqs (4.4) defined in $T$. Since involvement of the stabilizing terms into the integration process may be computationally expensive, analysts will probably prefer to check acceptability of the current constraint violation and correct occasionally the system position according to the scheme of Eq (4.5a).

A closely related problem is determination of the initial values of $u$ for given initial values $x_0$ and $\dot{x}_0$. As $f(x_0) = 0$ and $\dot{f}(x_0, \dot{x}_0) = 0$, the initial values $u_0$ can be found from Eq (4.6) as

$$u_0 = W_d^T (x_0) M(x_0) \dot{x}_0$$  \hspace{1cm} (4.7)

For Eqs (2.8) and (2.10) the corresponding formula would read

$$u_0 = M_u^{-1} (x_0) D^T(x_0) M(x_0) \dot{x}_0$$  \hspace{1cm} (4.7a)

where $M_u$ is defined in Eq (2.11). The problem of consistent initial values of tangent velocities is seldom undertaken in the literature.

The third aspect that may also be of importance is an effective determination of constraint reactions. The classical scheme (Wittenburg, 1977)

$$\lambda(x, u, t) = - (C^{*T} M^{-1} C^*)^{-1} \left( C^{*T} W_d u + C^{*T} M^{-1} h^* \right)$$  \hspace{1cm} (4.8)

may be computationally expensive. Moreover, $\lambda = [\lambda_1, \ldots, \lambda_m]^T$ should not, in general, be identified with the values of physical forces and moments of constraint reactions. In fact, a particular constraint of a system can be formulated in different ways, provided that the system's motion is prohibited in the same constrained subspace. We can write thus

$$C^{*T} \lambda = B^* \lambda_f = W_c^* \lambda_w$$  \hspace{1cm} (4.9)

where $B^*(x)$ is the $n \times m$ full-rank matrix of distribution of the physical constraint reaction values $\lambda_f = [\lambda_{f1}, \ldots, \lambda_{fm}]^T$; $W_c^*$ contains in columns
$w_i^* (i = 1, \ldots, m)$ produced by the orthonormalization process of Eq (3.2); and $\lambda_w$ conserves the corresponding Lagrange multipliers. By substituting $W_c^*$ for $C^*$ into Eq (4.8), and noting that $W_c^{\top} M^{-1} W_c^* = I^m$ and $\dot{W}_d^* W_d = -W_c^{\top} \dot{W}_d$, one obtains

$$\lambda_w(x, u, t) = \left( W_c^{\top} \dot{W}_d u - W_c^{\top} h^* \right)$$  \hspace{1cm} (4.10)

The above formula requires no matrix inversion and uses $W_c^*$, $W_c$, and $\dot{W}_d$ produced in the orthonormalization process and used in the formulation of kinetic equations (4.3) and (4.4). Having $\lambda_w$ determined, $\lambda$ (or $\lambda_f$) can be calculated from Eq (4.9). As this is an over-determined system of $n$ linear equations in $m$ unknowns $\lambda$ (or $\lambda_f$), only $m$ equations should be chosen, provided that the corresponding $m \times m$ submatrix matrix of $C^*$ (or $B^*$) is invertible. A rational choice of the submatrix can be made by using the projective criterion described in the next section.

5. Formulation of the supplementary matrix $E^*$

As said, the matrix $E^*$ defined in Eq (3.1) can be determined by any method, provided that $\text{rank}([C^* E^*]) = n$. For a given matrix $C^*$, the formulation of $E^*$ is not unique and may vary as $C^*$ depends on $x$. The robustness of the orthonormalization process may require that the projections of vectors $e_j (j = 1, \ldots, k)$ into $C$ are relatively small. Referring to the illustration shown in Fig.1, even if the directions of $e$ and $c$ are very close to each other, the two vectors still span over (form a basis of) the two-dimensional space. In such a case, however, the orthonormalization process may be ill-conditioned with its accuracy diminished. The condition $\text{det}(P^*) \neq 0$ says nothing about how the problem is conditioned, and the value of $\text{det}(P^*)$ is not a criterion of the conditioning either.

For small systems the supplementary matrix $E^*$ can sometimes be guessed or stated by intuition. In general, and especially for large systems, a formal procedure is needed. It is suggested by Liang and Lance (1987) to use SVD or LU factorization of $C^*$ to obtain $E^*$. At the point of application, a matrix $E = D$, see Eqs (2.8) $\div$ (12) defining $T$ would just be obtained, and the contravariant representation of the vectors $e_j (j = 1, \ldots, k)$ plays no role only if the metric of $R$ is disregarded. Since $C^*$ varies during the simulation process, the procedure should be occasionally repeated in order to avoid the
singularity of \( P^* \). By contrast, Agrawal and Saigal (1989) included the formulation of \( E^* \), directly in their orthogonalization process. After having the constraint vectors orthonormalized (in the meaning of their definition), they choose consecutively \( e^*_j \) \((j = 1, ..., k)\) and update the orthonormalized subspace. A chosen vector \( e^*_j \) is a unit vector along a particular \( \hat{x}_i \) \((i = 1, ..., n)\) direction, and a criterion of the \( i \)-direction choice acceptability is that \( e^*_j \) does not project totally into the so far orthonormalized subspace of \( \mathcal{R} \).

The approach proposed in this paper is conceptually close to that used by Agrawal and Saigal (1989); \( E^* \) is built at one stage, however, before the orthonormalization process is initialized. The present idea is to choose those column vectors \( k^*_i \) \((i = 1, ..., n)\) of \( K^* \)

\[
K^* = \begin{bmatrix}
1 & 0 & \cdots & 0 \\
0 & 1 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & 1
\end{bmatrix}
\]

(5.1)

for \( e^*_j \) \((j = 1, ..., k)\), which give the relatively smallest projections into \( C \) (or biggest projections into \( T \)). A particular \( k^*_i \) is a covariant representation of a unit vector \( k_i \) spanned along the \( i \)-th generalized velocity direction, \( (\hat{x}, k_i) = \hat{x}^T k^*_i = \hat{x}_i \). The squared length of \( k_i \) is

\[
\|k_i\|^2 = k^*_i M^{-1} k^*_i = M_{ii}^{-1}
\]

(5.2)

where \( M_{ii}^{-1} \) is the \((i, i)\)th element of \( M^{-1} \). Note that the vectors \( k_i \) \((i = 1, ..., n)\) defined by the unit matrix \( K^* \) are, in general, neither orthonormal nor orthogonal in \( \mathcal{R} \) as \( (k_i, k_j) = k^*_i M^{-1} k^*_j = M_{ij}^{-1} \). The projections of \( k_i \) into \( C \) and \( T \) can be stated as

\[
k^*_i^{(c)} = W^T c \cdot k^*_i = (W^T c)_i
\]

(5.3)

\[
k^*_i^{(d)} = W^T d \cdot k^*_i = (W^T d)_i
\]

where \((\cdot)_i\) denotes the \( i \)-th column of the corresponding matrix. The squared lengths of the projections are

\[
\|k^*_i^{(c)}\|^2 = (W^T c W^T c)_i
\]

(5.4)

\[
\|k^*_i^{(d)}\|^2 = (W^T d W^T d)_i
\]

where \((\cdot)_i\) denotes the \((i, i)\)th entry of the corresponding matrix. The above formulae are relatively simple due to: \( W^T_c M W_c = I^{(m)} \) and
\[ W_d^T M W_d = I^{(k)} \]. The relative volume of the projections can be measured by the generalized cosine and sine of the angle between \( k_i \) and \( k_i^{(c)} \)

\[
\cos \alpha_i = \frac{\| k_i^{(c)} \|}{\| k_i \|} \quad \sin \alpha_i = \frac{\| k_i^{(d)} \|}{\| k_i \|}
\] (5.5)

and it is easy to show that \( \cos^2 \alpha_i + \sin^2 \alpha_i = 1 \). If \( \cos \alpha_i = 1 \) (\( \sin \alpha_i = 0 \)), \( k_i \) is totally sunk in \( C \) (give no projection into \( T \)), and the corresponding \( k_i^* \) cannot be used in \( E^* \). The condition \( \cos \alpha_i < 1 \) (\( \sin \alpha_i > 0 \)) indicates that \( k_i \) is represented in \( T \), and there are at least \( k \) such vectors. Since all the vectors \( k_i \) (\( i = 1, ..., n \)) are linearly independent, any matrix \( E^* \) composed of \( k_j^* \) (\( j = 1, ..., k \)) such that \( \cos \alpha_j < 1 \) (\( \sin \alpha_j > 0 \)) is acceptable. An optimal choice of \( E^* \) is to gather those columns \( k_i^* \) (\( i = 1, ..., n \)) whose corresponding \( \cos \alpha_i \) (\( \sin \alpha_i \)) have the smallest (biggest) values, i.e. which give the relatively smallest (biggest) projections into \( C \) (\( T \)).

At a particular instant of simulation, the above projective criterion for choosing \( E^* \) uses the matrix \( W \) generated previously by the orthonormalization process. The criterion cannot thus be applied to initialize the process. At the initial point, \( \| k_i^{(c)} \| \) used in Eq (5.5) should thus be redefined according to \( \| k_i^{(c)} \|^2 = (C^* T M^{-1})^T M c^{-1} (C^* T M^{-1})_i \), where \( (C^* T M^{-1})_i \) is the \( i \)th column of \( C^* T M^{-1} \), and \( M_c = C^* T M^{-1} C^* \) is the metric tensor matrix of \( C \) defined by \( c_1, ..., c_m \). The other possibility is to apply the orthonormalization process only to the constraint vectors in order to generate \( W_c \). In fact, the latter approach could be applied at any instant of simulation, i.e. the orthonormalization process could be stopped after the constrained subspace is orthonormalized and, using Eqs (5.4), the matrix \( E^* \) could be chosen to finish the process. For \( k < m \), however, the sine criterion as in Eqs (5.4) may be computationally cheaper. Note also that the criterion can be used only occasionally in order to check the acceptability or redefine the current choice of \( E^* \).

If the orthonormalization process has been once initialized, at a particular instant of simulation one can take the current \( W_d^* \) for \( E^* \), and use it in the next step (or a sequence of steps) of integration of Eqs (4.3) and (4.4). This assures the best-conditioned matrix \( P^* \) to be obtained. However, \( W_d^* \) is a general matrix, whereas \( E^* \) defined as a set of columns of \( K^* \) is a sparse matrix having only one nonzero (unity) entry in each column. The latter feature benefits simplifications in the orthonormalization process for \( i = m + 1, ..., n \), and for large systems with few constraints this may be of importance. Note also that \( \dot{e}_j^* = 0 \), which yields further simplifications in the scheme (3.5).

The projective criterion defined in equations (5.2) \( \div (5.5) \) can also be applied for indicating the best conditioned \( m \) equations from \( n \) Eqs (4.9) to
determine $\lambda$ (or $\lambda_f$) as a function of $\lambda_w$. In this case, however, the chosen equations should correspond to those $m$ vectors from $k_i$ ($i = 1, ..., n$) which give the relatively biggest (smallest) projections into $C$ (T), i.e. whose $\cos \alpha_i$ ($\sin \alpha_i$) have the biggest (smallest) values.

6. Simple illustration

![Diagram of a moving physical pendulum]

Fig. 3. A moving physical pendulum

Consider a moving physical pendulum as shown in Fig.3. Setting $\mathbf{x} = [x_1, x_2, x_3]^T$ as indicated, the mass matrix of the system (the metric tensor of the related basis of the three-dimensional configuration space $\mathcal{R}$) is

$$
\mathbf{M} = \text{diag}(m, m, J)
$$

(6.1)

where $m$ and $J$ are the mass and the moment of inertia of the pendulum. The constraint equation is

$$
\mathbf{f} = x_2 - \rho \sin x_3 = 0
$$

(6.2)

The matrix $\mathbf{P}^*$ defined in Eq (3.1) can be constructed as

$$
\mathbf{P}^* = \begin{bmatrix}
0 & 1 & 0 \\
1 & 0 & 0 \\
-\rho \cos x_3 & 0 & 1
\end{bmatrix}
$$

(6.3)
where the first column is the constraint vector. Application of the orthonormalization scheme of Eq (3.2) and (3.5) yields

\[
\mathbf{W}_d^T = \begin{bmatrix}
\mu_1^{-1} & 0 & 0 \\
0 & \mu_2^{-1} \rho \cos x_3 & \mu_2^{-1} \\
0 & -\mu_2^{-3} J \rho \dot{x}_3 \sin x_3 & \mu_2^{-3} \rho \dot{x}_3 \sin x_3 \cos x_3
\end{bmatrix}
\]  

(6.4)

\[
\dot{\mathbf{W}}_d^T = \begin{bmatrix}
0 & 0 & 0 \\
0 & -\mu_2^{-3} J \rho \dot{x}_3 \sin x_3 & \mu_2^{-3} \rho \dot{x}_3 \sin x_3 \cos x_3
\end{bmatrix}
\]  

(6.5)

where \( \mu_1 = \sqrt{m} \), and \( \mu_2 = \sqrt{J + \rho \dot{x}_3 \sin x_3} \). Using Eqs (6.1) and (6.4) it is easy to show that \( \mathbf{W}_d^T \mathbf{M} \mathbf{W}_d = \mathbf{I}^{(2)} \), and the final governing equations are

\[
\dot{x}_1 = \mu_1^{-1} u_1
\]

\[
\dot{x}_2 = \mu_2^{-1} \rho u_2 \cos x_3
\]

\[
\dot{x} = \mu_2^{-1} u_2
\]

\[
\dot{u}_1 = \mu_1^{-1} h_1^*
\]

\[
\dot{u}_2 = \mu_2^{-1} \left( \rho h_2^* \cos x_3 + h_3^* \right)
\]

where \( h^* = [h_1^*, h_2^*, h_3^*]^T \) are the applied forces and moments on the pendulum. For given \( \mathbf{x}_0 \) and \( \dot{\mathbf{x}}_0 \), the initial values of \( \mathbf{u} \) are

\[
u_{10} = \mu_1 \dot{x}_{10}
\]

\[
u_{20} = \mu_1 \mu_2 \dot{x}_{20} + \mu_2 \dot{J} \dot{x}_{30}
\]

(6.7)

7. Conclusions

The main achievements of this paper are as follows:

- The vector orthonormalization process has been redefined according to the Riemannian space formulation.

- By supplementing the constraint matrix and employing the modified orthonormalization formulae, an orthonormal and differentiable basis of the tangent subspace can be constructed.
• A projective criterion for choosing the supplementary matrix is proposed.

• Resolved kinetic equations defined in the orthonormal basis of tangent subspace and expressed in terms of the corresponding tangent velocities, are obtained.

• A formulation is proposed for diminishing the constraint violation due to numerical errors of integration of the minimal-order equations of motion.

• A formula for exact determination of consistent initial values of the tangent speeds is provided.

• The problem of effective determination of constraint reactions is discussed.

As compared with the other methods (mentioned in Section 1) for constructing the tangent subspace bases, the present method seems to be computationally efficient. The obtained orthonormal basis of \( T \) is differentiable and benefits by the resolved kinetic Eqs (4.4). This feature of the formulation may be especially useful in applications. The only inconvenience is the necessity for determining \( \mathbf{M}^{-1} \), used frequently in the formulation process, and \( (\mathbf{M}^{-1})^\cdot = -\mathbf{M}^{-1}\mathbf{M}\mathbf{M}^{-1} \) applied to Eq (3.5). In the case of absolute coordinate formulation of multibody dynamics, \( \mathbf{M} \) is a constant diagonal matrix; determination of \( \mathbf{M}^{-1} \) is then a trivial task and \( \dot{\mathbf{M}} = 0 \). In a general case, however, \( \dot{\mathbf{M}} \) should be obtained symbolically at the stage of modelling of the problem, and \( \mathbf{M}^{-1} \) computed during the simulation process.

A final remark is that the role of Riemannian space formalism for the dynamic analysis of constrained/multibody systems is commonly underestimated. The formalism is indeed a powerful mathematical tool of the analysis that clarifies and automates many mathematical transformations. What is of greater importance, the formalism sets the analysis in physical order. As an illustration of the latter statement let us consider again the constraint vector \( \mathbf{c}^* = [0 \ 1 \ -\rho \cos x_3]^T \) introduced in Section 6. Following the common definition, the squared length of the vector is \( \|\mathbf{c}\|^2 = 1 + \rho^2 \cos^2 x_3 \), and the summands are of different dimensions. The Riemannian geometry adjusts the dimensions, \( \|\mathbf{c}\|^2 = \frac{1}{m} + \frac{1}{J} \rho^2 \cos^2 x_3 \). Evidently, in computations we play with numbers, and the dimensions are not seen. Consequently, many problems may be solved when no attention is payed to these aspects. In particular, all the methods mentioned in Section 1 produce the orthogonal complement matrices without taking the metric of the configuration space into consideration. Involving the space metric would probably benefit better conditioned formulations of the methods, however.
Acknowledgment

The research leading to this paper was supported in part by the State Committee for Scientific Research, grant No. 3 0955 91 01.

References


Formulowanie dynamiki układów nieswobodnych poprzez ortonormalizację przestrzeni konfiguracji

**Streszczenie**

W pracy zaproponowano algorytm automatycznego formulowania ortonormalnej i różniczkowej bazy podprzestrzeni stycznej dla nieswobodnych układów mechanicznych. Metoda wykorzystuje reguły Grama Schmidt ortogonalizacji wektorów, adoptując je zgodnie z formalizmem przestrzeni Riemanna. Interesującą i użyteczną własnością proponowanego sformulowania jest to, że kanoniczne równania ruchu układu (liczba ich jest równa ilości stopni swobody) generowane są bezpośrednio w rozwikłanej formie (macierz ma tych równań jest macierzą jednostkową). Innymi poruszonymi zagadnieniami są: generowanie dobrze uwarunkowanej i rzadko upakowanej macierzy uzupełniającej macierz więzów, wykorzystywanej następnie w procesie ortogonalizacji; korygowanie naruszenia więzów w wyniku kumulowania się niedokładności całkowania numerycznego równań ruchu; wyznaczanie zgodnych z warunkami więzów wartości początkowych prędkości stycznych; oraz efektywne wyznaczanie reakcji więzów.

*Manuscript received May 5, 1993; accepted for print June 23, 1993*