



A graph-theoretic approach to sparse matrix inversion for implicit differential algebraic equations

H. Yoshimura

Department of Applied Mechanics and Aerospace Engineering, Waseda University, Ohkubo, Shinjuku, Tokyo 169-8555, Japan

Correspondence to: H. Yoshimura (yoshimura@waseda.jp)

Received: 15 November 2012 – Revised: 9 March 2013 – Accepted: 5 April 2013 – Published: 6 June 2013

Abstract. In this paper, we propose an efficient numerical scheme to compute sparse matrix inversions for Implicit Differential Algebraic Equations of large-scale nonlinear mechanical systems. We first formulate mechanical systems with constraints by Dirac structures and associated Lagrangian systems. Second, we show how to allocate *input-output relations* to the variables in kinematical and dynamical relations appearing in DAEs by introducing an oriented bipartite graph. Then, we also show that the matrix inversion of Jacobian matrix associated to the kinematical and dynamical relations can be carried out by using the input-output relations and we explain solvability of the sparse Jacobian matrix inversion by using the bipartite graph. Finally, we propose an efficient symbolic generation algorithm to compute the sparse matrix inversion of the Jacobian matrix, and we demonstrate the validity in numerical efficiency by an example of the stanford manipulator.

1 Introduction

Multibody systems such as space structures, manipulators, etc. are known to be represented as implicit mechanical systems with kinematical constraints, holonomic or nonholonomic, which may be eventually expressed by implicit nonlinear Differential-Algebraic Equations (DAEs). In particular, for the numerical integration of such DAEs, we need to employ stiffly stable implicit numerical integrators such as Backward Differentiation Formula (see, Hachtel et al., 1971; Brayton et al., 1972), since the DAEs are to be highly nonlinear and stiff in general. On the other hand, one may face at a serious problem in CPU time for solving the implicit nonlinear algebraic equations, especially, for the case of large-scale systems. Namely, increasing degrees of freedom of the system, it eventually requires much CPU time in computing the matrix inversion of Jacobian matrix of the implicit DAEs in Newton's iteration at each time, since the Jacobian matrix of discretized nonlinear algebraic equations may be random sparse in general.

A major *stumbling block* lies in the fact that the Jacobian matrix has the random sparseness as well as highly nonlinear in terms of generalized coordinates. So far, some nu-

merical technique of sparse matrix inversions for VLSI circuits or networks has been developed by using the block-triangularization of matrices (see, for instance, Orlandea et al., 1977a; Murata et al., 1985), where a structural analysis is effectively made by means of *graph* and *matroid theory* (for instance, refer to Murota, 2000). In these conventional approaches, one may properly find out *pivots* in the Gaussian elimination process at each time step in an ad hoc way (where we note that the choice of pivots is quite relevant with *input-output relations* as will be shown shortly). This eventually requires much CPU time to calculate the inversion of the Jacobian matrix unless utilizing some effective sparse matrix algorithms. Namely, it is almost impossible to figure out a prior *fill-in* and *fill-out* in Gaussian elimination since *topological structure* of such a sparse Jacobian matrix might be so much random and complicated. Thus, we need to develop an efficient numerical algorithm of sparse matrix inversion for solving large-scale implicit DAEs in a systematic way.

In this paper, we develop a graph-theoretic approach to computing sparse matrix inversion for large-scale nonlinear mechanical systems by using Dirac structures and associated implicit Lagrangian systems. The underlying idea is to regard a mechanical system as an interconnected system of

elements, in which systems are comprised of *constitutive relations* of physical elements, *structural relations* among the physical relations, and *causal (input-output) relations* among physical variables. In particular, focusing upon the *input-output relations* associated with all the kinematical and dynamical relations of original mechanical systems, we develop *bipartite graphs* and then we show how the sparse matrix inversion can be made by effectively using the input-output relations. Furthermore, we explain solvability for the sparse Jacobian matrix inversion associated to the DAEs by using the bipartite graph. Finally, we propose an efficient and systematic symbolic generation algorithm to compute the sparse matrix inversion of the Jacobian matrix and we demonstrate its validity in numerical efficiency by an example of the Stanford manipulator.

2 Implicit Lagrangian systems

Let us review *Dirac structures* and associated *implicit Lagrangian systems* by following Yoshimura and Marsden (2006a,b, 2008).

2.1 Dirac structures

Let Q be an n -dimensional configuration manifold, whose kinematical constraints are given by a constraint distribution $\Delta_Q \subset TQ$, given by, at each $q \in Q$,

$$\Delta_Q(q) = \{v \in T_q Q \mid \langle \omega^a(q), v \rangle = 0, a = 1, \dots, m\}, \quad (1)$$

where ω^a are m one-forms on Q . Define the distribution Δ_{T^*Q} on T^*Q by

$$\Delta_{T^*Q} = (T\pi_Q)^{-1}(\Delta_Q) \subset TT^*Q,$$

where $T\pi_Q : TT^*Q \rightarrow TQ$ is the tangent map of the cotangent bundle projection $\pi_Q : T^*Q \rightarrow Q$, while the annihilator of Δ_{T^*Q} can be defined by, for each $z \in T^*_q Q$,

$$\Delta_{T^*Q}^\circ(z) = \{\alpha_z \in T_z^* T^*Q \mid \langle \alpha_z, w_z \rangle = 0 \text{ for all } w_z \in \Delta_{T^*Q}(z)\}.$$

Let Ω be the canonical symplectic structure on T^*Q and $\Omega^b : TT^*Q \rightarrow T^*T^*Q$ be the associated bundle map. Then, a Dirac structure D_{Δ_Q} on T^*Q induced from Δ_Q can be defined by, for each $z \in T^*_q Q$,

$$D_{\Delta_Q}(z) = \{(w_z, \alpha_z) \in T_z T^*Q \times T_z^* T^*Q \mid w_z \in \Delta_{T^*Q}(z) \text{ and } \alpha_z - \Omega^b(z) \cdot w_z \in \Delta_{T^*Q}^\circ(z)\}.$$

2.2 Local representations

Let us choose local coordinates q^i on Q so that Q is locally represented by an open set $W \subset \mathbb{R}^n$. The constraint set Δ_Q defines a subspace of TQ , which we denote by $\Delta(q) \subset \mathbb{R}^n$ at

each point $q \in W$. If the dimension of $\Delta(q)$ is $n - m$, then we can choose a basis $e_{m+1}(q), e_{m+2}(q), \dots, e_n(q)$ of $\Delta(q)$.

Recall that the constraint sets can be also represented by the annihilator of $\Delta(q)$, which is denoted by $\Delta^\circ(q)$ spanned by such one-forms that we write as $\omega^1, \omega^2, \dots, \omega^m$. Using $\pi_Q : T^*Q \rightarrow Q$ locally denoted by $z = (q, p) \mapsto q$ and $T\pi_Q : TT^*Q \rightarrow TQ; (q, p, \dot{q}, \dot{p}) \mapsto (q, \dot{q})$, it follows that

$$\Delta_{T^*Q} \cong \{(q, p, \dot{q}, \dot{p}) \mid q \in U, \dot{q} \in \Delta(q)\}.$$

Let points in T^*T^*Q be locally denoted by (q, p, β, u) , where β is a covector and u is a vector. Then, the annihilator of Δ_{T^*Q} is locally represented as

$$\Delta_{T^*Q}^\circ \cong \{(q, p, \beta, u) \mid q \in U, \beta \in \Delta^\circ(q) \text{ and } u = 0\}.$$

Since we have the local formula $\Omega^b(q, p) \cdot w_{(q,p)} = (q, p, -\dot{p}, \dot{q})$, the condition $\alpha_{(q,p)} - \Omega^b(q, p) \cdot w_{(q,p)} \in \Delta_{T^*Q}^\circ$ reads $\alpha + \dot{p} \in \Delta^\circ(q)$, and $w - \dot{q} = 0$, where $\alpha_{(q,p)} = (q, p, \alpha, w)$ and $w_{(q,p)} = (q, p, \dot{q}, \dot{p})$. Thus, the induced Dirac structure is locally represented by

$$D_{\Delta_Q}(q, p) = \{((\dot{q}, \dot{p}), (\alpha, w)) \mid \dot{q} \in \Delta(q), w = \dot{q}, \alpha + \dot{p} \in \Delta^\circ(q)\}. \quad (2)$$

Representation (I): let us introduce a matrix representation of D_{Δ_Q} given in Eq. (2). First, let $N^T(q)$ be an $n \times m$ matrix whose m -column vectors $\omega^1(q), \dots, \omega^m(q)$ span the basis of $\Delta^\circ(q)$, namely, $N^T(q) = [\omega^1(q), \dots, \omega^m(q)]$ and the distribution $\Delta(q) \subset \mathbb{R}^n \cong T_q Q$ may be represented by

$$\Delta(q) = \{\dot{q} \in \mathbb{R}^n \mid N(q)\dot{q} = 0\}.$$

Using Lagrange multipliers $\lambda = (\lambda_1, \dots, \lambda_m) \in \mathbb{R}^m$, one has

$$\Delta^\circ(q) = \{\beta \in (\mathbb{R}^n)^* \mid \beta = N^T(q)\lambda\}.$$

Thus, the induced Dirac structure can be represented by

$$D_{\Delta_Q}(q, p) = \{((\dot{q}, \dot{p}), (\alpha, w)) \mid N(q)\dot{q} = 0, w = \dot{q}, \alpha + \dot{p} = N^T(q)\lambda\}. \quad (3)$$

Representation (II): as shown in Eq. (3), for Representation (I) for the induced Dirac structure, we utilized the Lagrange multipliers, which represent *constraint forces* in constrained mechanical systems. Here, we develop another representation of D_{Δ_Q} on T^*Q without using the Lagrange multipliers.

Let us choose an $n \times (n - m)$ matrix $B(q) = [e_{m+1}(q), \dots, e_n(q)]$, whose column vectors span the basis of $\Delta(q)$. Then, it follows that the distribution $\Delta(q) \subset \mathbb{R}^n \cong T_q Q$ can be also represented by

$$\Delta(q) = \{\dot{q} \in \mathbb{R}^n \mid \dot{q} = B(q)u\},$$

where $u = (u^{m+1}, \dots, u^n) \in \mathbb{R}^{n-m}$. Note that the orthogonality condition between $N(q)$ and $B(q)$ holds:

$$B^T(q)N^T(q) = 0.$$

The above condition naturally comes from the fact that Δ° is the annihilator of the distribution Δ ; namely, in other words, the basis $e_{m+1}(q), \dots, e_n(q)$ is orthogonal to the dual basis $\omega^1(q), \dots, \omega^m(q)$ at each $q \in Q$. Therefore, one can read that

$$\Delta^\circ(q) = \{ \beta \in (\mathbb{R}^n)^* \mid B^T(q)\beta = 0 \}.$$

Thus, the induced Dirac structure $D_{\Delta_Q} \subset TT^*Q \oplus T^*T^*Q$ can be represented without using the Lagrange multipliers as

$$D_{\Delta_Q}(q, p) = \{ ((\dot{q}, \dot{p}), (\alpha, w)) \mid N(q)\dot{q} = 0, w = \dot{q}, B^T(q)(\alpha + \dot{p}) = 0 \}.$$

2.3 Ehresmann connection and structural relations

We briefly review an Ehresmann connection associated with nonholonomic mechanical systems; as to the details, for example, refer to Yoshimura and Marsden (2006b).

Assume that there is a bundle structure with a projection $\pi : Q \rightarrow \mathcal{R}$ for Q ; that is, there exists another manifold \mathcal{R} called the base. We call the kernel of $T_q\pi$ at any point $q \in Q$ the *vertical space* denoted by \mathcal{V}_q . An *Ehresmann connection* A is a vertical vector-valued one-form on Q , which satisfies

1. $A_q : T_qQ \rightarrow \mathcal{V}_q$ is a linear map at each point $p \in Q$,
2. A is a projection : $A(v_q) = v_q$, for all $v_q \in \mathcal{V}_q$.

Thus, we can split the tangent space at q such that $T_qQ = \mathcal{H}_q \oplus \mathcal{V}_q$, where $\mathcal{H}_q = \text{Ker } A_q$ is the horizontal space at q .

Suppose there exist nonholonomic constraints $\Delta_Q \subset TQ$, which are given by $m (< n)$ algebraic equations for n generalized velocity vector $v = \dot{q} = (\dot{q}^1, \dots, \dot{q}^n) \in \Delta(q) \subset T_qQ$ as in Eq. (1). Let us choose an Ehresmann connection A such that $\mathcal{H}_q = \Delta_Q(q)$ or we assume that the connection is chosen such that the constraints are written as $A \cdot v_q = 0$, where the constraint distribution Δ_Q is spanned by a set of m independent one-forms, which is given, in local coordinates $q^i = (r^\alpha, s^a)$ for Q , by

$$\omega^a = ds^a - J_\alpha^a(r, s)dr^\alpha.$$

In a matrix representation,

$$N(q)v = \left[\begin{array}{c|c} I_m & -\mathcal{J}(q) \end{array} \right] \begin{bmatrix} v^\circ \\ v^* \end{bmatrix} = o_m, \tag{4}$$

where v is locally split into *dependent* velocity $v^\circ = \dot{q}^\circ = (\dot{q}^1, \dots, \dot{q}^m)$ and *independent* velocity $v^* = \dot{q}^* = (\dot{q}^{m+1}, \dots, \dot{q}^n)$ and \mathcal{J} is a submatrix associated with the constraints. Geometrically speaking, this splitting corresponds to a choice of *Ehresmann connections* for the given constraints (see Yoshimura and Marsden, 2006b).

Corresponding to the annihilator, one has the dynamical relations associated to the generalized force vector $Q = (Q_1, \dots, Q_n) \in \Delta^\circ(q) \subset T^*Q$ dual to v as

$$B^T(q)Q = \left[\begin{array}{c|c} \mathcal{J}(q)^T & I_{n-m} \end{array} \right] \begin{bmatrix} Q^\circ \\ Q^* \end{bmatrix} = o_{n-m}, \tag{5}$$

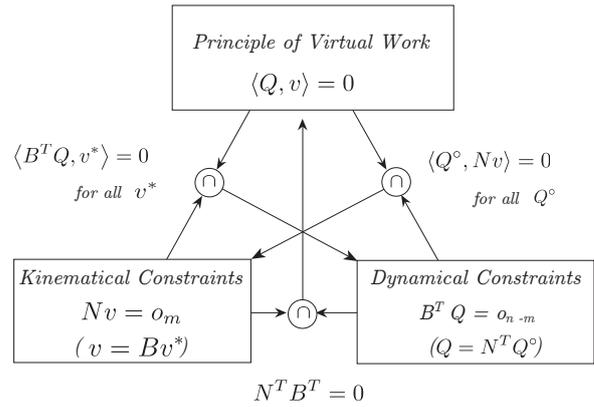


Figure 1. Duality principle.

where Δ_Q° denotes the annihilator of Δ_Q , and $Q^\circ = (Q_1, \dots, Q_m)$ and $Q^* = (Q_{m+1}, \dots, Q_n)$ are the generalized force vectors dual to v° and v^* respectively. On the other hand, the input-output relation between Q° and Q^* is reverse to v° and v^* ; namely, Q° is the input and Q^* the output. In the above, the *orthogonality condition* holds:

$$B^T(q)N^T(q) = 0.$$

The matrices $N(q)$ and $B(q)$ are called *connection matrices* (see Yoshimura, 1995). This orthogonality condition denotes *principle of virtual work*, which is given by

$$\langle Q, v \rangle = 0, \text{ for all } v \in \Delta(q) \text{ and } Q \in \Delta^\circ(q),$$

where \langle, \rangle denotes a duality pairing.

The dual set of constraints given by Eqs. (4) and (5) indicates *structural relations*, namely, it represents how physical elements are interconnected. Thus, we sometimes call the structural relation an **interconnection** among the physical elements. In circuit theory, it is known that Eqs. (4) and (5) correspond to KCL and KVL and also that the virtual work principle is known as *Tellegen's theorem*. Furthermore, there exists a relation called **duality principle** as in Fig. 1, which is known as *Planck-Okada-Arsove principle* in circuit theory (see Yoshimura, 1995).

2.4 Implicit DAEs for Lagrangian systems

Here, we show how the notion of Dirac structures can be fit into the formulation of implicit Lagrangian systems (see Yoshimura and Marsden, 2006a,b, 2008). Let L be a Lagrangian on TQ , which is given by

$$L(q, \dot{q}) = \frac{1}{2} \langle \dot{q}^\circ, M \dot{q}^\circ \rangle - U(q^\circ),$$

where we assume that L is only associated to \dot{q}° , M is a mass matrix whose components are functions of q° , and U denotes a potential energy function of q° . This implies that L is possibly *degenerate*.

The **Lagrange-d'Alembert principle** is given by

$$\int_a^b \left\langle \frac{d}{dt} \frac{\partial L}{\partial \dot{q}} - \frac{\partial L}{\partial q}, \delta q \right\rangle dt + \int_a^b \langle F, \delta q \rangle dt = 0,$$

where δq satisfies the constraint

$$N(q)\delta q = o_m.$$

So, one can obtain the dual dynamical relation

$$B^T(q)Q = o_{n-m},$$

where

$$Q = \frac{d}{dt} \frac{\partial L}{\partial \dot{q}} - \frac{\partial L}{\partial q} + F$$

and it directly induces

$$Q = \begin{bmatrix} Q^\circ \\ Q^* \end{bmatrix} = \begin{bmatrix} M\dot{v}^\circ + f(q^\circ, v^\circ) - \frac{\partial U}{\partial q^\circ} \\ \tau \end{bmatrix}, \quad (6)$$

where

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}} - \frac{\partial L}{\partial q} = \begin{bmatrix} M\dot{v}^\circ + f(q^\circ, v^\circ) - \frac{\partial U}{\partial q^\circ} \\ 0 \end{bmatrix}$$

and

$$F = \begin{bmatrix} 0 \\ \tau \end{bmatrix}.$$

Notice that τ indicates the external forces. Of course, equations of motion can be written as

$$B^T(q) \left(\frac{d}{dt} \frac{\partial L}{\partial \dot{q}} - \frac{\partial L}{\partial q} + F \right) = o_{n-m}.$$

Furthermore, one has the *second-order condition* (see Marsden and Ratiu, 1999):

$$\dot{q}^\circ - v^\circ = o_m, \quad (7)$$

$$\dot{q}^* - v^* = o_{n-m}. \quad (8)$$

From Eqs. (4), (5), (6), (7) and (8), we can obtain the following *local differential-algebraic equations*

$$G(x(t), \dot{x}(t); u(t)) = 0, \quad (9)$$

where $x = (q, v, Q) = (q^\circ, q^*, v^\circ, v^*, Q^\circ, Q^*) \in W \times W \times W^*$ denotes the state variables and $u = \tau \in (W^{n-m})^* = (\mathbb{R}^{n-m})^*$ the input variables. In the above, we locally set $TQ \cong W \times W$ and $T^*Q \cong W \times W^*$, and hence $TQ \oplus T^*Q \cong W \times W \times W^*$, where $Q \cong W = W^m \times W^{n-m} = \mathbb{R}^m \times \mathbb{R}^{n-m}$ is an n -dimensional vector space which is a model space for Q . Thus, the mathematical model of the Lagrangian system is given by the implicit DAEs:

$$G = \begin{bmatrix} G_1 \\ G_2 \\ G_3 \\ G_4 \\ G_5 \\ G_6 \end{bmatrix} = \begin{bmatrix} \dot{q}^\circ - v^\circ \\ \dot{q}^* - v^* \\ v^\circ - \mathcal{J}v^* \\ \mathcal{J}^T Q^\circ + Q^* \\ Q^\circ - M\dot{v}^\circ - f(q^\circ, v^\circ) - \partial U / \partial q^\circ \\ Q^* - \tau \end{bmatrix}. \quad (10)$$

3 Sparse tableau approach

From the viewpoint of numerical analysis for mechanical systems, there exist two kinds of dynamical problems; namely, the *forward dynamics* and *inverse dynamics*.

Recall that $W = W^m \times W^{n-m} = \mathbb{R}^m \times \mathbb{R}^{n-m}$ is the model space for Q . The forward dynamics analysis is the case in which given a smooth input vector

$$u(t) := \tau(t) \in (W^{n-m})^* = (\mathbb{R}^{n-m})^*$$

as a *vector function* of time t , numerically integrate Eq. (9) in terms of t to obtain

$$x(t) = (q^\circ(t), q^*(t), v^\circ(t), v^*(t), Q^\circ(t), Q^*(t))$$

as the *output*, where $x \in W \times W \times W^*$. On the other hand, the inverse dynamics analysis is the case in which given a smooth input vector

$$u(t) := q^*(t) \in W^{n-m} = \mathbb{R}^{n-m}$$

as a *vector function* of time t , then compute

$$x(t) = (q^\circ(t), v^\circ(t), v^*(t), Q^\circ(t), Q^*(t), \tau(t))$$

as the output, in which case $x \in W^m \times W \times W^* \times (W^{n-m})^*$.

In this paper, we explore the case of the inverse dynamics analysis by the sparse tableau approach, where the state vector is given by $x = (q^\circ, v^\circ, v^*, Q^\circ, Q^*, \tau)$. To do this, let us first discretize Eq. (9) at time $t = t_n$. By using the *Backward Differentiation Formula (BDF)* developed by Gear (see, for instance, Brennen et al., 1995), the time-derivative term $\dot{x}_n = \dot{x}(t_n)$ associated to the state vector $x_n = x(t_n)$ may be approximately discretized by the backwards $x_{n-i} = x(t_{n-i})$ as

$$\dot{x}_n = - \sum_{i=0}^k \frac{1}{h} \alpha_i x_{n-i} \quad (1 \leq k \leq 6), \quad (11)$$

where $h = t_n - t_{n-1}$ denotes a time step, k is a backward order and α_i indicates the coefficient associated to the i -th backward order. Substituting Eq. (11) into Eq. (9), we obtain the discretized nonlinear algebraic equations as follows:

$$G(x_n, \dot{x}(x_n); u(t_n)) = 0. \quad (12)$$

Recall the algorithm of the *Sparse Tableau Approach* is given in Fig. 2 (see Hachtel et al., 1971; Brayton et al., 1972), where we linearize Eq. (12) at each time step $t = t_n$ as

$$J(x_n^{(r)}) \Delta x_n^{(r)} = -G(x_n^{(r)}), \quad (13)$$

and where $J = [\partial G_i / \partial x_j]_{t=t_n}$ is the *Jacobian matrix* and $\Delta x_n^{(r)}$ denotes the r -th iterated *corrector vector* at t_n . Then, it follows

$$\begin{aligned} x_n^{(r+1)} &= x_n^{(r)} + \Delta x_n^{(r)} \\ &= x_n^{(r)} - J^{-1}(x_n^{(r)})G(x_n^{(r)}). \end{aligned} \quad (14)$$

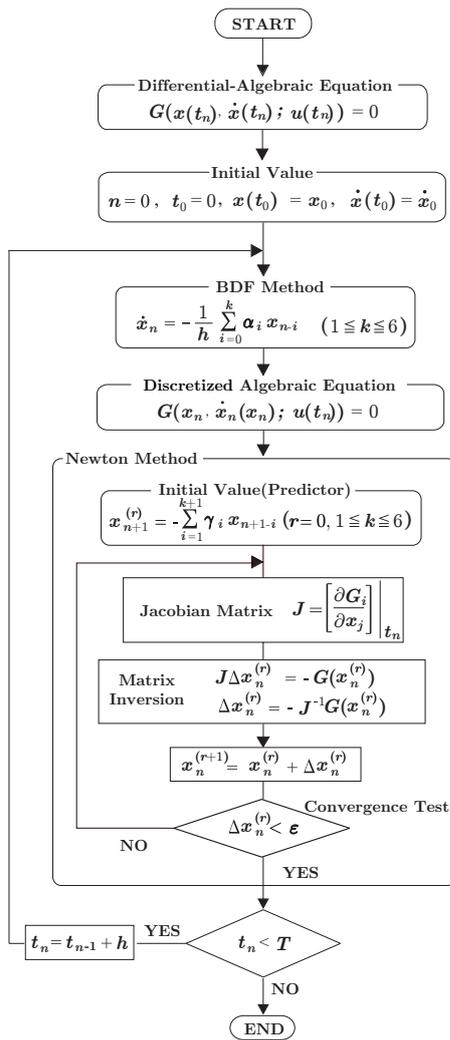


Figure 2. Sparse tableau method.

In the Newton method, it is necessary to take initial values near from the solution x_n and the k -th prediction formula x_n^{Pr} is given by

$$x_n^{Pr} = x_n^{(0)} = - \sum_{i=1}^k \gamma_i x_{n-i},$$

where γ_i is the i -th coefficient.

In the inverse dynamics analysis, the state vector is given by $x = (q^\circ, v^\circ, v^*, Q^\circ, Q^*, \tau)$ and it follows from Eq. (10) that the Jacobian matrix is given as in Fig. 3, where I_n stands for the n -th degree unit matrix.

4 Bipartite graphs

4.1 Input-output relations

The Jacobian matrix $J(x_n^{(r)})$ obtained in Eq. (13) apparently has the characteristic of random sparseness. So, we develop

$$J = \begin{bmatrix} q^\circ & v^\circ & v^* & Q^\circ & Q^* & \tau \\ \left(-\frac{\alpha_0}{h}\right) I_m & -I_m & & & & \\ & & -I_{n-m} & & & \\ -\frac{\partial(\mathcal{J}v^*)}{\partial q^\circ} & I_m & -\mathcal{J} & & & \\ \frac{\partial(\mathcal{J}^T Q^\circ)}{\partial q^\circ} & & & \mathcal{J}^T I_{n-m} & & \\ -\frac{\partial f}{\partial q^\circ} & \left(\frac{\alpha_0}{h} M - \frac{\partial f}{\partial v^\circ}\right) & & I_m & & \\ & & & & I_{n-m} & -I_{n-m} \end{bmatrix}$$

Figure 3. Jacobian matrix.

an efficient symbolic generation for computations of the sparse Jacobian matrix inversion for Newton’s iterations. To do this, let us consider an **input-output relation** among state variables for every relation in (10). Now, we can uniquely allocate the input-output relation to the kinematical and dynamical relations in Eqs. (4) and (5) as follows:

$$v^\circ - \mathcal{J}v^* = o_m \quad (\text{output : } v^\circ, \text{ input : } v^*),$$

$$\mathcal{J}^T Q^\circ + Q^* = o_{n-m} \quad (\text{input : } Q^\circ, \text{ output : } Q^*).$$

Similarly, for the second-order conditions between $v = (v^\circ, v^*)$ and $\dot{q} = (\dot{q}^\circ, \dot{q}^*)$, one has

$$\dot{q}^\circ - v^\circ = o_m \quad (\text{output : } \dot{q}^\circ, \text{ input : } v^\circ),$$

$$\dot{q}^* - v^* = o_{n-m} \quad (\text{input : } \dot{q}^*, \text{ output : } v^*).$$

Furthermore, as to the equations of motion, it follows

$$Q^\circ - M \dot{v}^\circ - f(q^\circ, v^\circ)$$

$$- \partial_{q^\circ} U = o_m \quad (\text{output : } Q^\circ, \text{ input : } q^\circ, v^\circ),$$

$$Q^* - \tau = o_{n-m} \quad (\text{input : } Q^*, \text{ output : } \tau).$$

The input-output relations in the mentioned above can be determined by physical **causality**. Needless to say that the time derivative terms $\dot{x}(t_n)$ are expressed in terms of the backwards $x_{n-i}(t_n)$, $i = 0, \dots, 6$ by using the BDF as in equation (11). Corresponding to the Jacobian matrix J in (13), define the **causal Jacobian matrix** \hat{J} by assigning -1 to the input and $+1$ to the output as to the j -th variable in the i -th relation associated to J_{ij} as

$$\hat{J}_{ij} = \begin{cases} -1 & : \text{ the } j\text{-th variable is input,} \\ +1 & : \text{ the } j\text{-th variable is output,} \\ 0 & : \text{ otherwise.} \end{cases}$$

Thus, the causal Jacobian matrix \hat{J} is given in Fig. 4.

In the above, \widehat{I}_m indicates the m -th unit matrix. Note that there exists an element with $+1$ in each row, which plays a role of the pivot in the Gaussian elimination. Further, $\widehat{I}_{m,n-m}$ is the $m \times (n-m)$ matrix, in which $+1$ are allocated to non-zero components of the $m \times (n-m)$ matrix \mathcal{J} .

4.2 Oriented bipartite graphs

Let us illustrate the input-output relations as to the Jacobian matrix $J = [J_{ij}(x)] = [\partial G_i / \partial x_j]_{t=t_n}$ by introducing

$$\hat{J} = \begin{bmatrix} q^\circ & v^\circ & v^* & Q^\circ & Q^* & \tau \\ +I_m & -I_m & & & & \\ & & +I_{n-m} & & & \\ -\hat{I}_m & +I_m & -\hat{I}_{m,n-m} & & & \\ -\hat{I}_{n-m,m} & & & -\hat{I}_{n-m,m} + I_{n-m} & & \\ -\hat{I}_m & -\hat{I}_m & & +I_m & & \\ & & & & -I_{n-m} + I_{n-m} & \end{bmatrix}$$

Figure 4. Causal Jacobian matrix.

bipartite graphs. Recall that the i -th row of the Jacobian matrix $J = [J_{ij}(x)] = [\partial G_i / \partial x_j] \Big|_{t=t_n}$ corresponds to the i -th vector function $G_i(x, \dot{x}; u)$ and the j -th column corresponds to j -th vector x_j respectively, where $i, j = 1, \dots, N (= 3n)$. Here, $G_i(x, \dot{x}; u) = 0$ are the N -th order DAEs and $x = (x_1, \dots, x_6) = (q^\circ, v^\circ, v^*, Q^\circ, Q^*, \tau)$ is the N -vector.

Given the causal Jacobian matrix \hat{J} associated to J , let $V_r = \{G_1, G_2, \dots, G_N\}$ be the **row-set** of \hat{J} and let $V_c = \{\Delta x_1, \dots, \Delta x_6\}$ be the **column-set** of \hat{J} . Here, we assign V_r and V_c to **vertex set** $V := V_r \times V_c$.

Now, there are nonzero elements (± 1) in the i -th row of the causal Jacobian matrix and suppose that they are in the k -th and l -th columns. For instance, as to the first equation $G_1 = 0$, one has the kinematical relation between $x_1 = q^\circ$ and $x_2 = v^\circ$ as $\dot{x}_1 - x_2 = 0$. By the BDF discretization, the corresponding linearized equation is given as

$$-\frac{1}{h}(\alpha_0 \Delta x_1) - \Delta x_2 = 0.$$

In order to illustrate input-output relations among the state variables by *graphs*, let us define **arc-set** by

$$A = \{(j, i) \mid \hat{J}_{ij} \neq 0, i \in V_c, j \in V_r\},$$

where the arcs represent some **relations** between vertices in V_c (state variables) as to every vertex (equation) in V_r . Furthermore, the direction of an arc indicates *causality* or an input-output relation among state variables associated to the column vertices. Let $a \in A$ be an arc. Let us introduce a map $s : A \rightarrow V$, which is given by, for an arc $a \in A$, $s(a)$ denotes the *initial vertex* of a . Similarly, define a map $t : A \rightarrow V$, which is given by, for an arc $a \in A$, $t(a)$ indicates the *terminal vertex* of a . So, we can define the set of *vertices incident* to a by $\{s(a), t(a)\}$. Sometimes, $s(a)$ is called the **source** of a and $t(a)$ the **target** of a . Hence, we can define a **directed bipartite graph** by

$$B = (V_r, V_c, A, s, t),$$

by which we can illustrate the *input-output relations associated to the Jacobian matrix* as shown in Fig. 5.

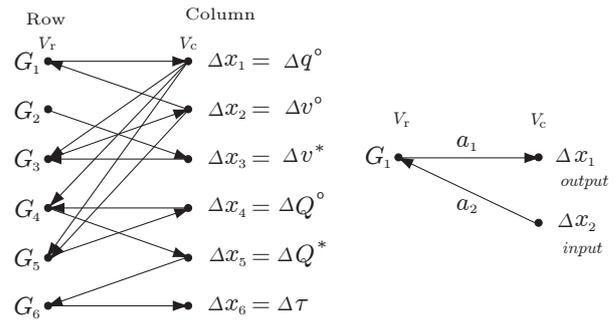


Figure 5. Directed bipartite graph.

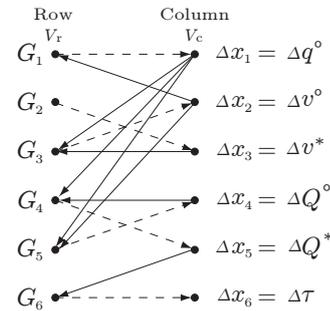


Figure 6. Perfect matching.

4.3 Perfect matching

A **matching** $M \subset A$ of a bipartite graph $B = (V_r, V_c, A, s, t)$ is defined by a *set of arcs without common vertices*. Especially, M is called a **perfect matching** if it is a matching which matches all vertices of the graph, namely, every vertex of the graph is incident to exactly one edge of the matching: $s(M) = V_r$ and $t(M) = V_c$.

In Fig. 6, the arcs that are drawn by the broken lines consist of the perfect matching. If a perfect matching exists in B , the corresponding Jacobian matrix J is a *square and nonsingular* matrix since $|V_r| = |V_c|$. Here, $|V|$ means the size of V , namely, the number of elements in the set of vertices V .

Let us see that $|V^+| = |V^-|$ is equivalent with the fact that the Jacobian matrix is *diagonalizable*. First, let us show the following relation:

Perfect matchings exist in $B \Rightarrow J$ is diagonalizable.

This is clear because a perfect matching can be detected by the *elementary row and column operations* in matrix; namely, (1) interchanging two rows or columns; (2) adding a multiple of one row or column to another; (3) multiplying any row or column by a nonzero element, although there might be several perfect matchings for a given bipartite graph. Next, let us show the following relation:

J is diagonalizable \Rightarrow Perfect matchings exist in B .

That the Jacobian matrix J is diagonalizable means that one can properly select a pivot in each row or column from

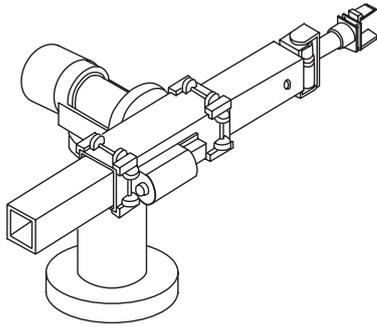


Figure 9. Stanford manipulator.

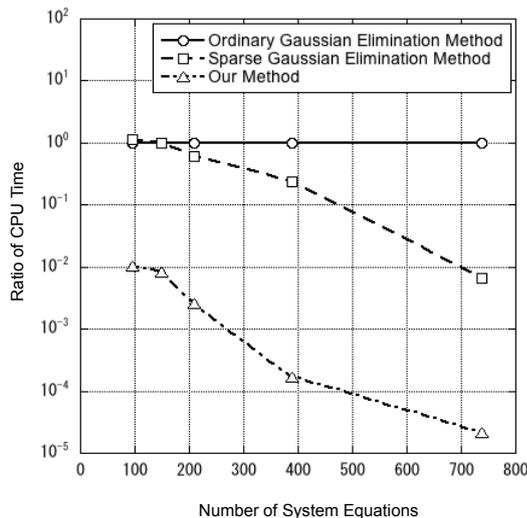


Figure 10. Comparison of CPU time.

has a great advantage in the CPU time efficiency in comparison with other approaches.

6 Conclusions

We have shown symbolic generation of sparse matrix inversion for large-scale mechanical systems. We have set up Implicit DAE models in the context of Lagrangian systems and we have shown the input-output relations as to the Jacobian matrix associated with linearized equations. Then, we have shown the solvability of the linearized equations by using the bipartite graph. Furthermore, we have proposed symbolic generation of the random sparse matrix inversion for the Jacobian matrix. Finally, we have demonstrated the validity of our approach by numerical analysis with an example of the Stanford manipulator comparing with the inner-product sparse matrix Gaussian elimination algorithm as well as the standard Gaussian elimination algorithm.

Acknowledgements. This research is partially supported by JSPS Grant-in-Aid 23560269, JST-CREST, Waseda University Grant for SR 2012A-602 and IRSES project Geomech-246981.

Edited by: A. Müller

Reviewed by: two anonymous referees

References

- Brayton, R. K., Gustavson, F. G., and Hachtel, G. D.: A new efficient algorithm for solving differential-algebraic systems using implicit backward differentiation formulas, *Proc. IEEE*, 60, 98–108, 1972.
- Brennen, K. E., Campbell, S. L., and Petzold, L. R.: *Numerical Solution of Initial-Value Problems in Differential-Algebraic Equations*, Philadelphia, SIAM, 1995.
- Hachtel, G. D., Brayton, R. K., and Gustavson, F. G.: The sparse tableau approach to network analysis and design, *IEEE Trans. Circuit Theory*, CT-18, 101–113, 1971.
- Marsden, J. E. and Ratiu, T. S.: *Introduction to Mechanics and Symmetry*, volume 17 of *Texts in Applied Mathematics*, Springer-Verlag, 2nd Edn., 1999.
- Murata, T., Oguni, T., and Karaki, Y.: *Supercomputer-Application to Scientific Computing*, Maruzen, 1985 (in Japanese).
- Murota, K.: *Matrices and Matroids for Systems Analysis*, Springer-Verlag, Berlin, 2000.
- Orlandea, N., Chace, M. A., and Calahan, D. A.: A sparsity-oriented approach to the dynamic analysis and design of mechanical systems – Part 1, Part 2, *Trans. ASME, J. Eng. Ind.*, 99, 773–779, 780–784, 1977.
- Yoshimura, H.: *Dynamics of Flexible Multibody Systems*, Ph.D. thesis, Waseda University, Tokyo, Japan, 1995.
- Yoshimura, H. and Imai, N.: A sparse tableau approach to dynamics of large-scale nonlinear mechanical systems, *IEICE Nonlinear Problems*, 103, 39–43, 2004.
- Yoshimura, H. and Marsden, J. E.: Dirac structures in Lagrangian mechanics Part I: Implicit Lagrangian systems, *J. Geom. Phys.*, 57, 133–156, 2006a.
- Yoshimura, H. and Marsden, J. E.: Dirac structures in Lagrangian mechanics Part II: Variational structures, *J. Geom. Phys.*, 57, 209–250, 2006b.
- Yoshimura, H. and Marsden, J. E.: Representations of Dirac Structures and Implicit Port-controlled Lagrangian Systems, *Proc. International Symposium on Mathematical Theory of Networks and Systems 2008*, Blacksburg, Virginia, 1–12, 2008.