Second-order sensitivity analysis of multibody systems described by differential/algebraic equations: adjoint variable approach

JIE-YU DING†‡, ZHEN-KUAN PAN‡ and LI-QUN CHEN*§†

†Shanghai Institute of Applied Mathematics and Mechanics, Shanghai University, Shanghai 200072, P.R. China
‡College of Information Engineering, Qingdao University, Qingdao 266071, P.R. China
§Department of Mechanics, Shanghai University, Shanghai 200444, P.R. China

(Received 28 November 2006; revised version received 19 March 2007; accepted 11 June 2007)

A second-order sensitivity analysis has higher accuracy and can improve convergence of optimization design, but its computation is complex. This paper studies the adjoint variable method of second sensitivity analysis for multibody system dynamics described by second-order differential/algebraic equations. Based on general objective function, constraint conditions, initial and end conditions, the adjoint variable equations for first-order sensitivity analysis and design sensitivity formulations are derived first, then second-order sensitivity analysis formulations are set up based on the previous results. For simplification, the second derivative of the objective function is transformed to an initial value problem of an ordinary differential equation with one variable. Finally, the comparison between the results of direct computation, first-order sensitivity and second-order sensitivity validate the method presented in this paper.

Keywords: Multibody system dynamics; Differential/algebraic equations; Optimization design; Sensitivity analysis; Adjoint variable method

AMS Subject Classification: 70E55; 49Q12; 65L80

1. Introduction

With the rapid development of computational multibody dynamics and its applications in optimization design of spacecraft, vehicles, robots, and other mechanisms, sensitivity analysis has received considerable attention in recent years. Sensitivity in optimization design based on multibody analysis includes state sensitivity and design sensitivity, the former is the derivative of state with respect to parameters, the latter is derivative of object function with respect to parameters, which not only can measure the dependence of design on parameters, but also act as a bridge between multibody analysis and optimization.
The divided-difference method, automatic differentiation technique (AD), direct differentiation method, and adjoint variable method have been successfully used for this task. The first method estimates the sensitivity by analysing perturbed design points without explicitly differentiating the equations. Due to the additional set of function evaluations required, this approach may be numerically expensive. The accuracy of this approach is also affected by the magnitude of perturbation as well as the shape of the true design hyperplane at this design point, and is subject to truncation and possible round-off errors. AD [1,2] is based on the implementation of the chain rule of differentiation and has already shown superior performance relative to the divided-difference method in speed and accuracy. Simultaneously, it needs less computer memory and less code, but some research [2] shows that including expert knowledge about the problem structure within the algorithm is extremely hard to achieve with AD tools. Without implementation precautions, a pure syntactical analysis may not be sufficient and could possibly yield wrong results.

The direct differentiation method and adjoint variable method are two major analytical techniques for sensitivity analysis used in multibody optimization design. The direct differentiation method [3] can be easily implemented by straightforward differentiation of system governing equations, performance measures, and/or constraint equations with respect to design variables. The adjoint variable method [3] is used to calculate the design sensitivity after solving a set of adjoint variable equations derived from variations of system equations. For a dynamic system optimization design problem involving $p$ design variables, $n$ generalized coordinates, and $m$ independent algebraic constraint equations, the simultaneous solution of $(n + m)(p + 1)$ and $(n + m + p)$ differential algebraic equations are required for the direct differentiation method and the adjoint variable method, respectively, for first-order sensitivity analysis [4].


Second-order sensitivity can improve convergence efficiency of optimization design and can provide more accurate information. Haug et al. [3] derives a second-order direct differentiation formulation for multibody systems described by first-order ordinary differential equations, but the research based on the second-order differential/algebraic equations [19] needs to be improved. This paper studies the adjoint variable method of second sensitivity analysis for multibody system dynamics described by second-order differential/algebraic equations and sets up second-order sensitivity analysis formulations.
2. Problem formulation

Multibody systems can be described by the following differential/algebraic equations of motion

\[
\begin{cases}
    M(q, b)\ddot{q} + \Phi_q^T \lambda = Q(\dot{q}, q, b, t) \\
    \Phi(q, b, t) = 0
\end{cases}
\]

(1.1) (1.2)

where \( b \in \mathbb{R}^p \) is a vector of design variables, \( q \in \mathbb{R}^n \) is a vector of generalized coordinates. The constraints in algebraic equations (1.2) lead to reaction forces expressed in the differential equations (1.1) by Lagrange multipliers \( \lambda \) and the Jacobian matrix \( \Phi_q \). The holonomic constraint functions \( \Phi \), the generalized mass matrix \( M \), and the applied forces vector \( Q \) depend on the design variables.

It is presumed that the initial and final times \( t_1 \) and \( t_2 \) are determined by

\[
\Omega_i(q^i, b, t_i) = 0, \quad i = 1, 2
\]

(2)

where the superscript 1 denotes the initial time and 2 denotes the final time.

A complete characterization of the motion of the system requires definition of initial conditions on position and velocity as the following form

\[
\begin{align*}
\varphi(q^1, b, t_1) &= 0 \\
\dot{\varphi}(q^1, b, t_1) &= 0
\end{align*}
\]

(3.1) (3.2)

where the matrices \((\Phi_q^{1T}, \varphi_q^{1T})\) and \((\Phi_q^{1T}, \dot{\varphi}_q^{1T})\) must be nonsingular.

A typical objective function of the system for optimization purposes is

\[
\psi(b) = G^1(q^1, q^1, b, t_1) + G^2(q^2, q^2, b, t_2) + \int_{t_1}^{t_2} H(\dot{q}, q, \lambda, b, t) \, dt
\]

(4)

The functions \( G^1, G^2 \) and \( H \) evaluate the performance of the dynamic system at the initial and final time \( t_1, t_2 \) and within an interesting time domain \([t_1, t_2]\).

3. First-order design sensitivity analysis

Using the Leibniz rule for the derivative of an integral and the chain rule of differentiation, with integration by parts, the first-order derivative of the objective function \( \psi \) with respect to design variables becomes

\[
\begin{align*}
\psi'(b) &= G_q^1 \dot{q}^1_b + G_q^2 \dot{q}^2_b + (G_q^1 - H_q^1)q^1_b + (G_q^2 + H_q^2)q^2_b + G^1_b + G^2_b \\
&\quad + (\dot{G}^1 - H^1)l^1_b + (\dot{G}^2 + H^2)l^2_b + \int_{t_1}^{t_2} \left[ \left( H_q - \frac{d}{dt} H_q \right) q_b + H_\lambda \lambda_b + H_b \right] \, dt
\end{align*}
\]

(5)

where

\[
\dot{G}^i \triangleq \frac{dG^i}{dt}, \quad i = 1, 2
\]

(6)
Similarly to [3], introducing adjoint variables $\mu, v, \eta^1, \eta^2, \xi^1, \xi^2, \alpha, \beta$, then letting the coefficients of $\ddot{q}_b, \dot{q}_b, \dot{q}_b^T, t^1_b, t^2_b, q_b, \lambda_b$ be zero yields the following adjoint equations

$$M^1 \mu^1 - \Phi_q^T \xi^1 - \Phi_q^T \beta = -G_q^T,$$

(7.1)

$$M^2 \mu^2 + \Phi_q^T \xi^2 = G_q^T,$$

(7.2)

$$M^1 \dot{\mu}^1 + (M^1 + Q_q^T) \mu^1 + \Omega_q^T \xi^1 + \Phi_q^T \eta^1 + \Phi_q^T \xi^1 + \Phi_q^T \alpha = G_q^T - H_q^T,$$

(7.3)

$$M^2 \dot{\mu}^2 + (M^2 + Q_q^T) \mu^2 - \Phi_q^T \eta^2 - \Phi_q^T \xi^2 - \Omega_q^T \xi^2 = -H_q^T - G_q^T,$$

(7.4)

$$\dot{\Phi}^T \xi^1 + \dot{\Phi}^T \xi^1 + \Phi^T \xi^1 + \Phi^T \alpha + \dot{\Phi}^T \beta = \dot{G}_q^T - H_q^T,$$

(7.5)

$$\dot{\Phi}^T \xi^2 + \Phi^T \xi^2 = \dot{G}_q^T + H_q^T,$$

(7.6)

$$M \ddot{\mu} + (2\dot{M} + Q_q^T) \dot{\mu} + (\ddot{M} + \frac{d}{dt}Q_q^T + \Pi_q^T) \mu + \Phi_q^T v = -\frac{d}{dt} H_q^T + H_q^T,$$

(7.7)

$$\Phi_q \mu = H_q^T,$$

(7.8)

with a definition for $\Phi, \dot{\Phi}, \dot{\Omega}, \dot{\varphi}, \dot{\psi}$ similar to equation (6), and

$$\Pi \triangleq M \ddot{q} + \Phi_q^T \lambda - Q = 0$$

(8)

Then equation (5) becomes

$$\psi'(b) = G_b^1 + G_b^2 - \eta^T \Phi_b^1 - \eta^T \Phi_b^1 - \xi^T \Phi_b^1 - \xi^T \Phi_b^2 - \xi^T \Phi_b^2 - \xi^T \Phi_b^2 - \alpha^T \phi_b - \beta^T \phi_b$$

$$+ \int_{t_1}^{t^2} (H_b - \mu^T \Pi_b - v^T \Phi_b) dt$$

$$\triangleq F + \int_{t_1}^{t^2} R dt$$

(9)

Employing the numerical methods of [3], the following sequence of computations yields adjoint variables $\mu, v, \eta^1, \eta^2, \xi^1, \xi^2, \xi^1, \xi^2, \alpha, \beta$.

**Step 1** Equation (7.2) and equation (7.8) of time $t^2$ are

$$
\begin{pmatrix}
M^2 & \Phi_q^T \\
\Phi_q^T & 0
\end{pmatrix}
\begin{pmatrix}
\mu^2 \\
\xi^2
\end{pmatrix}
= 
\begin{pmatrix}
G_q^T \\
H_q^T
\end{pmatrix}
$$

(10)

Since the coefficient matrix is nonsingular, $\mu^2, \xi^2$ are determined.

**Step 2** Equation (7.4) and the first-order derivative of equation (7.8) about time $t$ are

$$
\begin{pmatrix}
M^2 & \Phi_q^T \\
\Phi_q^T & 0
\end{pmatrix}
\begin{pmatrix}
\dot{\mu}^2 \\
\dot{\xi}^2
\end{pmatrix}
= 
\begin{pmatrix}
\Omega_q^T \\
0
\end{pmatrix}
\xi^2 + 
\begin{pmatrix}
-\dot{M}^2 + Q_q^T \\
\frac{d}{dt} H_q^T
\end{pmatrix}
\mu^2 + \begin{pmatrix}
\Phi_q^T \\
\frac{d}{dt} H_q^T
\end{pmatrix}
\xi^2
$$

(11)

Since the coefficient matrix is nonsingular, $\dot{\mu}^2, \dot{\xi}^2$ are uniquely determined as functions of $\xi^2$. With the solution of equation (11), equation (7.6) determines $\xi^2$, hence $\dot{\mu}^2, \dot{\xi}^2$. 

Step 3  Equation (7.7) and equation (7.8) are
\[
\begin{pmatrix}
M & \Phi_q^T \\
\Phi_q & 0
\end{pmatrix}
\begin{pmatrix}
\dot{\mu} \\
\nu
\end{pmatrix}
= \left(-\frac{d}{dt}Q_q^T\right)\dot{\mu}
- \left(\frac{d}{dt}Q_q^T + \Pi_q^T\right)\mu
- \frac{d}{dt}H_q^T + H_q^T \right)
\]  
(12)
Backward integration of equation (12) with \(\mu^2, \dot{\mu}^2\) yields \(\mu, \nu\), hence \(\mu^1, \dot{\mu}^1\).

Step 4  Equation (7.1) is
\[
\begin{pmatrix}
\Phi_q^T \\
\varphi_q^T
\end{pmatrix}
\begin{pmatrix}
\xi^1 \\
\beta
\end{pmatrix}
= M^1 \mu^1 + G_q^T
\]  
(13)
Since the coefficient matrix is nonsingular, \(\xi^1, \beta\) are determined.

Step 5  Equation (7.3) is
\[
\begin{pmatrix}
\Phi_q^T \\
\varphi_q^T
\end{pmatrix}
\begin{pmatrix}
\eta^1 \\
\alpha
\end{pmatrix}
= -\Omega_q^T \xi^1 - M^1 \dot{\mu}^1 - (M^1 + Q_q^T)\mu^1 - \Phi_q^T \xi^1 + G_q^T - H_q^T
\]  
(14)
Since the coefficient matrix is nonsingular, \(\eta^1, \alpha\) are uniquely determined as functions of \(\xi^1\). With the solution of equation (14), equation (7.5) determines \(\xi^1\), hence \(\eta^1, \alpha\).

Step 6  Using results of Steps 1–5, the first derivative of \(\psi\) with respect to \(b\) can be evaluated by equation (14) without computing \(\dot{q}_b, \ddot{q}_b, q_b^1, t_b^1, t_b^2, \dot{q}_b, \lambda_b\).

4. Second-order design sensitivity analysis

Using the Leibniz rule, chain rule and integration by parts, the derivative of \(\psi'(b)\) with respect to design variables becomes
\[
\psi''(b) = \left(\dot{F}_q^1 + \dot{R}_q^1\right)q_b^1 + \left(\ddot{F}_q^2 - \ddot{R}_q^2\right)q_b^2 + \left(\dot{R}_q^1 - \ddot{R}_q^1 + \ddot{R}_q^2\right)q_b^1
+ \left(\dddot{F}_q^2 + \dddot{R}_q^2\right)q_b^2
+ \sum_{i=1}^{2} (\Phi_i^T \eta_i^1 + \Phi_i^T \xi_i^1) - \eta_i^1 \alpha_i - \Phi_i^T \beta_i + \ddot{b}
+ \int_{t_1}^{t_2} [(\dot{R}_q - \dddot{R}_q)q_b + \dddot{R}_q \lambda_b - \Pi_b^T \mu_b - \Phi_b^T \nu_b + \dddot{R}_b] dt
\]  
(15)
where
\[
\dot{F} \triangleq G_b^T + G_b^2 - \Phi_b^T \eta^1 \Phi_b^T - \Phi_b^T \xi^1 \Phi_b^T - \Phi_b^T \xi^2 - \Phi_b^T \xi^1 - \Phi_b^T \xi^2 - \Omega_b^T \xi^1 - \Omega_b^T \xi^2 - \Phi_b^T \alpha - \Phi_b^T \beta
\]  
(16)
The symbol \(\sim\) above a variable denotes the variable that is to be held fixed for the partial differentiation indicated.

In order to evaluate the second derivatives of equation (15), terms involving \(q_b^1, \dot{q}_b, q_b^2, t_b^1, t_b^2, \eta_b, \xi_b, \mu_b, \mu_b, \alpha_b, \beta_b, q_b, \lambda_b, \mu_b, \nu_b\) must be rewritten explicitly in terms of computable quantities.
The derivative of equation (1.1) with respect to design variables is

\[ M\ddot{q}_b - Q_i\dot{q}_b + \Pi q_i + \Phi^T_q \lambda_b + \Pi_b = 0 \]  

(17)

Both sides of equation (17) are multiplied by the transpose of an introduced adjoint variable \( \kappa \) and integrated from \( t^1 \) to \( t^2 \) to yield

\[ \int_{t^1}^{t^2} \kappa^T (M\ddot{q}_b - Q_i\dot{q}_b + \Pi q_i + \Phi^T_q \lambda_b + \Pi_b) dt = 0 \]  

(18)

Integrating the first and second term of equation (18) by parts gives

\[ \int_{t^1}^{t^2} \left\{ \left[ \kappa^T M + \kappa^T (2\dot{M} + Q_i) \right] + \kappa^T \left( \dot{M} + \frac{d}{dt}Q_i + \Pi q_i \right) \right\} q_i + \kappa^T \Phi^T_q \lambda_b + \kappa^T \Pi_b \right\} dt \]

\[ - \kappa^T M^1 \dot{q}_b^1 + \kappa^T M^2 \dot{q}_b^2 + [\kappa^T M^1 + \kappa^T (\dot{M}^1 + Q_i^1)]q_b^1 \]

\[ - [\kappa^T M^2 + \kappa^T (\dot{M}^2 + Q_i^2)]q_b^2 = 0 \]  

(19)

The derivative of equation (1.2) with respect to design variables is

\[ \Phi_q q_i + \Phi_b = 0 \]  

(20)

Both sides of equation (20) are multiplied by the transpose of an introduced adjoint variable \( \chi \) and integrated from \( t^1 \) to \( t^2 \) to yield

\[ \int_{t^1}^{t^2} (\chi^T \Phi_q q_i + \chi^T \Phi_b) dt = 0 \]  

(21)

Equation (1.2) and its velocity equations of time \( t^1 \) and \( t^2 \) are

\[ \Phi_i (q_i, b, t^i) = 0 \quad i = 1, 2 \]  

(22.1)

\[ \Phi_i \dot{q}_i + \Phi_{i'} = 0 \]  

(22.2)

Derivatives of equation (22.1) and equation (22.2) with respect to design variables with the introduced adjoint variables \( \tau^i \), \( \tau^i \) \( (i = 1, 2) \) are

\[ t^i T \Phi_i q_i^i + t^i T \Phi_i t_i^i + t^i T \Phi_b^i = 0, \quad i = 1, 2 \]  

(23.1)

\[ t^i T \Phi_i q_i^i + t^i T \Phi_i t_i^i + t^i T \Phi_b^i = 0 \]  

(23.2)

Derivatives of equation (2) with respect to design variables with the introduced adjoint variables \( \varepsilon^i \) \( (i = 1, 2) \) are

\[ \varepsilon^i \Omega_i q_i^i + \varepsilon^i \dot{\lambda}_i t_i^i + \varepsilon^i \Omega_b^i = 0, \quad i = 1, 2 \]  

(24)

Derivatives of equation (3.1) and equation (3.2) with respect to design variables with the introduced adjoint variables \( \pi \), \( \rho \) are

\[ \pi^T \Phi_i q_i^i + \pi^T \Phi t_i^i + \pi^T \Phi_b = 0 \]  

(25.1)

\[ \rho^T \Phi_i q_i^i + \rho^T \Phi t_i^i + \rho^T \Phi_b = 0 \]  

(25.2)
Derivatives of equations (7.1)–(7.8) with respect to design variables with the introduced adjoint variables \( \delta^1, \delta^2, \gamma^1, \gamma^2, \theta^1, \theta^2, \omega, \nu \) are

\[
\begin{align*}
\delta^T \tilde{E}_{1,q} q_b^1 + \delta^T \tilde{E}_{1,q} q_b^1 + \delta^T \tilde{E}_{1,b} + \delta^T M \mu_b^1 & = \delta^T \Phi_q^T \xi_b^1 \\
- \delta^T \phi^T \beta_b + \delta^T \tilde{E}_{1,b} & = 0 \\
\delta^T \tilde{E}_{2,q} q_b^2 + \delta^T \tilde{E}_{2,q} q_b^2 + \delta^T \tilde{E}_{2,b} & = 0 \\
\gamma^T \tilde{E}_{3,q} q_b^3 + \gamma^T \tilde{E}_{3,q} q_b^3 + \gamma^T \tilde{E}_{3,b} + \gamma^T \Phi_q^T \eta_b^1 + \gamma^T \Phi_q^T \xi_b^1 + \gamma^T \phi^T \alpha_b & = 0 \\
\gamma^T \tilde{E}_{4,q} q_b^4 + \gamma^T \tilde{E}_{4,q} q_b^4 + \gamma^T \tilde{E}_{4,b} & = 0 \\
\theta^1 \tilde{E}_{5,q} q_b^5 + \theta^1 \tilde{E}_{5,q} q_b^5 + \theta^1 \tilde{E}_{5,b} + \theta^1 H^1 \lambda_b^1 + \theta^1 \Phi^T \eta_b^2 & = 0 \\
\theta^1 \phi^T \beta_b + \theta^1 \tilde{E}_{5,b} & = 0 \\
\theta^2 \tilde{E}_{6,q} q_b^6 + \theta^2 \tilde{E}_{6,q} q_b^6 + \theta^1 \tilde{E}_{6,b} & = 0 \\
\int_{t_1}^{t_2} \left[ \omega^T (\tilde{E}_{7,q} - \tilde{E}_{7,q} + \dot{\tilde{E}}_{7,q}) - \omega^T (\tilde{E}_{7,q} - 2 \dot{\tilde{E}}_{7,q}) + \omega^T \tilde{E}_{7,q} \right] q_b^1 + (\omega^T M - \omega^T \Omega_q^T \nu_b^1 \\
+ \omega^T \Phi_q^T v_b^1 + \omega^T \tilde{E}_{7,\lambda} \lambda_b^2 + \omega^T (\tilde{E}_M \mu_b^1 + \omega^T \tilde{E}_{7,b}) dt & - \omega^T \tilde{E}_{7,q} q_b^1 \\
+ \omega^T \tilde{E}_{7,\lambda} \lambda_b^2 + \omega^T (\tilde{E}_M \mu_b^1 + \omega^T \tilde{E}_{7,b}) dt & - \omega^T \tilde{E}_{7,q} q_b^1 \\
- \omega^2 T \tilde{E}_{7,\lambda} \lambda_b^2 + \omega^2 T \tilde{E}_{7,b} & = 0 \\
\int_{t_1}^{t_2} \left[ \nu^T \Phi_q \mu_b^1 + \nu^T \tilde{E}_{8,q} + \nu^T (\tilde{E}_{8,q} - \dot{\tilde{E}}_{8,q}) \right] q_b^1 + \nu^T H_{\lambda,\lambda} \lambda_b^2 + \nu^T \tilde{E}_{8,b} dt & = 0
\end{align*}
\]

where

\[
\tilde{E}_1 \triangleq \tilde{M}^1 \tilde{\mu}_1 - \Phi_q^T \tilde{\xi}_1 - \phi_q^T \tilde{\beta} + \Omega_q^T \nu_b^1 \\
\tilde{E}_2 \triangleq \tilde{M}^2 \tilde{\mu}_2 + \Phi_q^T \tilde{\xi}_2 - \Omega_q^T \nu_b^1 \\
\tilde{E}_3 \triangleq \tilde{M}^1 \tilde{\mu}_1 + (\tilde{M}^1 + \Omega_q^T \nu_b^1) \tilde{\mu}_1 + \Omega_q^T \nu_b^1 + \Phi_q^T \tilde{\xi}_1 + \phi_q^T \tilde{\alpha} - G_q^1 + H_q^1
\]
\[ \mathbf{E}_4 \triangleq M^2 \ddot{\mu}^2 + (M^2 + Q_{q}^T) \ddot{\mu}^2 - \Phi_{q}^T \gamma^2 - \Phi_{q}^T \xi^2 - \Omega_{q}^T \ddot{\gamma}^2 + G_{q}^T + H_{q}^T \]
\[ \mathbf{E}_5 \triangleq \Phi_{T}^T \ddot{\eta}^1 + \Phi_{T}^T \ddot{\xi}^1 + \Omega_{T}^T \ddot{\gamma}^1 + \Phi_{q}^T \ddot{\alpha} + \Phi_{q}^T \ddot{\beta} - \dot{G}_{T}^1 + H_{T}^T \]
\[ \mathbf{E}_6 \triangleq \Phi_{T}^2 \ddot{\eta}^2 + \Phi_{T}^2 \ddot{\xi}^2 + \Omega_{T}^2 \ddot{\gamma}^2 - \dot{G}_{T}^2 - H_{T}^2 \]
\[ \mathbf{E}_7 \triangleq M \ddot{\mu} + (2M + Q_{q}^T) \ddot{\mu} + \left( \ddot{M} + \frac{d}{dt} Q_{q}^T + \Pi_{q}^T \right) \ddot{\mu} + \Phi_{q}^T \ddot{\nu} + \frac{d}{dt} H_{q}^T - H_{q}^T \]
\[ \mathbf{E}_8 \triangleq \Phi_{q} \ddot{\mu} - H_{\lambda}^T \]

and \( \overline{E}_{i,x} \triangleq \delta \overline{E}_i / \partial x \) \( (i = 1, \ldots, 8) \).

Equations (19), (21), (23.1), (23.2), (24), (25.1), (25.2), (26.1)–(26.8) can be subtracted from equation (15) without changing the value of \( \psi''(b) \) for any choice of adjoint variables. For the coefficients of \( \phi_1^1, \phi_1^2, q_1^2, q_b^1, t_1^1, t_1^2, \eta_1^2, \eta_1^2, \xi_1^1, \xi_1^2, \lambda_1^1, \lambda_2^1, \mu_1^1, \mu_2^1, \alpha_1, \beta_1, q_2^1, \lambda_b, \mu_b, \nu_b \) to be zeros, it is required that the adjoint variables satisfy the following equations.

\[ \Phi_{q}^{T,1} \tau^1 + \phi_{q}^{T,1} \rho - M^1 \kappa^1 + E_{1,1}^T \gamma^1 + E_{3,1}^T \gamma^1 + E_{5,1}^T \gamma^1 - \overline{E}_{7,1} \gamma^1 = \overline{F}_{q}^T + \overline{R}_{q}^T \]  
\[ M^2 \kappa^2 + \Phi_{q}^{T,2} \tau^2 + \overline{E}_{2,2} \gamma^2 + \overline{E}_{4,2} \gamma^2 + \overline{E}_{6,2} \gamma^2 + \overline{E}_{7,2} \gamma^2 = \overline{F}_{q}^T - \overline{R}_{q}^T \]  
\[ \Phi_{q}^{T,1} \kappa^1 + \phi_{q}^{T,1} \kappa^1 + M^1 \kappa^1 + (M^1 + Q_{q}^T) \kappa^1 + \Phi_{q}^{T,1} \kappa^1 + \overline{E}_{1,1}^T \kappa^1 + E_{3,1}^T \kappa^1 + E_{5,1}^T \kappa^1 + (E_{7,1}^T - \overline{E}_{7,1}^T) \kappa^1 + \overline{E}_{7,1}^T \kappa^1 + \overline{E}_{7,1}^T \kappa^1 - H_{\lambda}^T \kappa^1 \kappa^1 + \overline{F}_{q}^T + \overline{R}_{q}^T + \overline{R}_{q}^T \]  
\[ M^2 \kappa^2 - \Phi_{q}^{T,2} \kappa^2 - \Omega_{q}^{T,2} \kappa^2 + (M^2 + Q_{q}^T) \kappa^2 + \Phi_{q}^{T,2} \kappa^2 - \overline{E}_{2,2} \kappa^2 - \overline{E}_{4,2} \kappa^2 + \overline{E}_{6,2} \kappa^2 = \overline{F}_{q}^T - \overline{R}_{q}^T - \overline{R}_{q}^T \]  
\[ \Phi_{q}^{T,1} \kappa^1 + \phi_{q}^{T,1} \kappa^1 + \overline{E}_{1,1}^T \kappa^1 + \overline{E}_{3,1}^T \kappa^1 + \overline{E}_{5,1}^T \kappa^1 - \overline{E}_{7,1} \omega^1 - \overline{E}_{8,1} \nu^1 = \overline{F}_{q}^T - \overline{R}_{q}^T \]  
\[ \Phi_{q}^{T,2} \kappa^2 + \phi_{q}^{T,2} \kappa^2 + \overline{E}_{2,2} \kappa^2 + \overline{E}_{4,2} \kappa^2 + \overline{E}_{6,2} \kappa^2 + \overline{E}_{7,2} \kappa^2 = \overline{F}_{q}^T - \overline{R}_{q}^T - \overline{R}_{q}^T \]  

\[ H_{q}^{T,1} \gamma^1 + H_{\lambda}^{T,1} = 0 \]  
\[ H_{q}^{T,2} \gamma^2 + H_{\lambda}^{T,2} = 0 \]  
\[ \Phi_{q}^{T,1} \phi^1 \phi^1 + \phi_{q}^{T,1} \phi^1 \phi^1 = \Phi_{q}^1, \quad i = 1, 2 \]  
\[ \Phi_{q}^{T,1} \phi^1 \phi^1 - \phi_{q}^{T,1} \phi^1 \phi^1 = \Phi_{q}^1, \quad i = 1, 2 \]  
\[ M^1 \phi^1 = 0, \quad i = 1, 2 \]  
\[ M^1 \phi^1 = (M^1 + Q_{q}^T) \phi^1 + M^1 \omega^1 - (M^1 + Q_{q}^T) \omega^1 = 0, \quad i = 1, 2 \]  
\[ \phi_{q}^{T,1} \phi^1 + \phi_{q}^{T,1} \phi^1 = \phi_{b} \]  
\[ \phi_{q}^{T,1} \phi^1 + \phi_{q}^{T,1} \phi^1 = \phi_{b} \]
Then the second-order derivative of objective function with respect to design variables becomes:

$$
\psi''(b) = \bar{F}_b - \sum_{i=1}^{2} \left[ \iota^T \Phi^i_b + \tau^i \Phi^i_b + \varepsilon^i \Omega^i_b \right] - \pi^T \varphi_b - \rho^T \varphi_b - \delta^T \bar{E}_{1,b} \nonumber
$$

$$- \delta^2 \bar{E}_{2,b} - \gamma^1 \bar{E}_{3,b} - \gamma^2 \bar{E}_{4,b} - \theta^1 \bar{E}_{5,b} - \theta^2 \bar{E}_{6,b} \nonumber
$$

$$+ \int_{t_1}^{t_2} \left[ \bar{R}_b - \kappa^T \Pi_b - \chi^T \Phi_b - \omega^T \bar{E}_{7,b} - \upsilon^T \bar{E}_{8,b} \right] dt \nonumber
$$

$$\Delta = S + \int_{t_1}^{t_2} T dt \nonumber
$$

Equations (28.1)–(28.18) determine adjoint variables $\kappa, \chi, \iota^1, \iota^2, \tau^1, \tau^2, \epsilon^1, \epsilon^2, \pi, \rho, \delta^1, \delta^2, \gamma^1, \gamma^2, \theta^1, \theta^2, \omega, \upsilon$. Thus, the second derivative of $\psi$ with respect to $b$ can be evaluated.

5. Second-order design sensitivity analysis algorithm

Using the values of $\ddot{q}, \dot{q}, q, \lambda$ and adjoint variables $\mu, \upsilon, \eta^1, \eta^2, \zeta^1, \zeta^2, \xi^1, \xi^2, \alpha, \beta$ have been obtained from the first-order design sensitivity analysis, the following sequence of computations yields all adjoint variables that are required for evaluating the sensitivity vector in equation (29).

**Step 1** Equation (28.9) and equation (28.13) yields

$$
\begin{pmatrix}
\Phi^i \\
\varphi^i 
\end{pmatrix}
\gamma^1 = - \begin{pmatrix}
\hat{\Phi} \\
\hat{\varphi} 
\end{pmatrix} \theta^1 + \begin{pmatrix}
\Phi^i_b \\
\varphi^i_b 
\end{pmatrix} \nonumber
$$

where $\gamma^1$ is determined as a function of $\theta^1$. With the solution of equation (30), equation (28.7) determines $\theta^1$, hence $\gamma^1$. With the value of $\gamma^1$, equation (28.11) determines $\omega^1$.

**Step 2** Equation (28.10) and equation (28.14) yields

$$
\begin{pmatrix}
\Phi^i \\
\varphi^i 
\end{pmatrix}
\delta^1 = \begin{pmatrix}
\hat{\Phi} \\
\hat{\varphi} 
\end{pmatrix} \theta^1 + \begin{pmatrix}
-\Phi^i_\gamma \gamma^1 \Phi^i_b \\
\varphi^i_b 
\end{pmatrix} \nonumber
$$

Since the coefficient matrix is nonsingular, $\delta^1$ is determined. With the values of $\delta^1, \gamma^1$ and $\omega^1$, equation (28.12) determines $\dot{\omega}^1$. 


Step 3 The first- and second-order derivatives of equation (28.18) about time $t$ are

$$\Phi_q \dot{\omega} + \left( \frac{d}{dt} \Phi_q \right) \omega = -\frac{d}{dt} \Phi_b^T$$

(32.1)

$$\Phi_q \ddot{\omega} + 2 \left( \frac{d}{dt} \Phi_q \right) \dot{\omega} + \left( \frac{d^2}{dt^2} \Phi_q \right) \omega = -\frac{d^2}{dt^2} \Phi_b^T$$

(32.2)

Equation (28.17) and equation (32.2) are

$$\left( \begin{array}{cc} M & \Phi_q^T \\ \Phi_q & 0 \end{array} \right) \left( \begin{array}{c} \dot{\omega} \\ \dot{v} \end{array} \right) = \left( \begin{array}{c} Q_q \ddot{\omega} - \Pi_q \omega - \Pi_b \\ -2 \left( \frac{d}{dt} \Phi_q \right) \dot{\omega} - \left( \frac{d^2}{dt^2} \Phi_q \right) \omega + \frac{d^2}{dt^2} \Phi_b^T \end{array} \right)$$

(33)

With the initial values $\omega^1$, $\dot{\omega}^1$, equation (33) determines $\omega$, $v$, hence $\omega^2$, $\dot{\omega}^2$ and $\gamma^2$, $\delta^2$. With the value of $\gamma^2$, equation (28.8) determines $\theta^2$.

Step 4 Equation (28.2) and equation (28.16) of time $t^2$ are

$$\left( \begin{array}{cc} M^2 & \Phi_{q2}^2 \\ \Phi_{q2}^2 & 0 \end{array} \right) \left( \begin{array}{c} \kappa^2 \\ \tau^2 \end{array} \right) = \left( \begin{array}{c} \tilde{E}_{2,q^2} \delta^2 - \tilde{E}_{4,q^2} T^2 - \tilde{E}_{6,q^2} T^2 - \tilde{E}_{7,q^2} T^2 + \tilde{F}_{q^2} T - \tilde{R}_{q^2} \\ -\tilde{E}_{7,\lambda^2} T^2 + H_{\lambda^2}^2 T + \bar{R}_{\lambda^2} \end{array} \right)$$

(34)

Since the coefficient matrix is nonsingular, $\kappa^2$, $\tau^2$ are determined.

Step 5 The first- and second-order derivatives of equation (28.16) about time $t$ are

$$\Phi_q \ddot{\kappa} + \left( \frac{d}{dt} \Phi_q \right) \dot{\kappa} = \frac{d}{dt} \left[ \tilde{E}_{7,\lambda} T^2 + H_{\lambda} T^2 + \bar{R}_{\lambda} \right]$$

(35.1)

$$\Phi_q \dddot{\kappa} + 2 \left( \frac{d}{dt} \Phi_q \right) \ddot{\kappa} + \left( \frac{d^2}{dt^2} \Phi_q \right) \dot{\kappa} = \frac{d^2}{dt^2} \left[ \tilde{E}_{7,\lambda} T^2 + H_{\lambda} T^2 + \bar{R}_{\lambda} \right]$$

(35.2)

Equation (28.4) and equation (35.1) of time $t^2$ are

$$\left( \begin{array}{cc} M^2 & \Phi_{q2}^2 \\ \Phi_{q2}^2 & 0 \end{array} \right) \left( \begin{array}{c} \kappa^2 \\ -\tau^2 \end{array} \right) = \left( \Omega_{q^2}^T \right) \varepsilon^2$$

where

$$\begin{align*}
\{\left( M^2 + Q_{q2}^T \right) \kappa^2 + \Phi_{q2}^T T^2 + \tilde{E}_{2,q^2} T^2 + \tilde{E}_{4,q^2} T^2 + \tilde{E}_{6,q^2} T^2 + \tilde{E}_{7,q^2} T^2 + \tilde{F}_{q^2} T + \tilde{R}_{q^2} \\
+ \tilde{E}_{7,\lambda^2} T^2 + H_{\lambda^2}^2 T + \bar{R}_{\lambda^2} \}\left( \frac{d}{dt} \Phi_q \right) \dot{\kappa} - \frac{d}{dt^2} \left[ \tilde{E}_{7,\lambda} T^2 + H_{\lambda}^2 T^2 + \bar{R}_{\lambda^2} \right]
\end{align*}$$

(36)

Since the coefficient matrix is nonsingular, $\kappa^2$, $\tau^2$ are uniquely determined as functions of $\varepsilon^2$. With the solution of equation (36), equation (28.6) determines $\varepsilon^2$, hence $\kappa^2$, $\tau^2$. 
Step 6  Equation (28.15) and equation (35.2) are
\[
\begin{pmatrix} M & \Phi_q^T \\ \Phi_q & 0 \end{pmatrix} \begin{pmatrix} \dot{\kappa} \\ \chi \end{pmatrix} = \begin{pmatrix} -(2\dot{M} + Q_q^T)\dot{k} - (\ddot{M} + \frac{d}{dt}Q_q^T + \Pi_q^T)\kappa - \bar{E}_{7,q}\dot{\omega} \\ (\bar{E}_{7,q} + 2\bar{E}_{7,q}^T)\dot{\omega} - (\bar{E}_{7,q}^T - \bar{E}_{7,q} + \bar{E}_{7,q}^T)\omega \\ -\bar{E}_{8,q}\dot{\nu} - (\bar{E}_{8,q} + \dot{E}_{8,q})\nu + \bar{R}_q - \bar{R}_q + \bar{R}_q \\ -2 (\frac{d}{dt}\Phi_q) \dot{k} - (\frac{d^2}{dt^2}\Phi_q) \kappa + \frac{d^2}{dt^2} \left[ -\bar{E}_{7,\lambda}\omega + H_{\lambda\lambda}^T\nu + \bar{R}_q \right] \end{pmatrix}
\]  
(37)

Backward integration of equation (37) with \( \kappa^2, \dot{k} \) yields \( \kappa, \chi \), hence \( \kappa^1, \dot{k}^1 \).

Step 7  Equation (28.1) is
\[
\begin{pmatrix} \Phi_{q^1}^T & \Phi_{q^2}^T \end{pmatrix} \begin{pmatrix} \tau^1 \\ \rho \end{pmatrix} = M^1 \kappa^1 - \bar{E}_{1,q}^T \delta^1 - \bar{E}_{3,q}^T \gamma^1 - \bar{E}_{5,q}^T \theta^1 + \bar{E}_{7,q}^T \omega^1 + \bar{F}_{q} + \bar{R}_q^T
\]  
(38)

Since the coefficient matrix is nonsingular, \( \tau^1, \rho \) are determined.

Step 8  Equation (28.3) is
\[
\begin{pmatrix} \Phi_{q^1}^T & \varphi_{q^1}^T \end{pmatrix} \begin{pmatrix} t^1 \\ \pi \end{pmatrix} = -\Omega_q^1 \dot{\epsilon}^1 - M^1 \dot{\kappa}^1 - (\dot{M} + Q_q^T)\kappa^1 - \Phi_q^T \tau^1 - \bar{E}_{1,q}^T \delta^1
\]  
\[
- \bar{E}_{3,q}^T \gamma^1 - \bar{E}_{5,q}^T \theta^1 + (\bar{E}_{7,q}^T + \bar{E}_{7,q}^T)\omega^1 - \bar{E}_{7,q}^T \dot{\omega}^1
\]  
\[
- H_{\lambda\lambda}^T \nu^1 + \bar{F}_{q} + \bar{R}_q^T + \bar{R}_q^T
\]  
(39)

Since the coefficient matrix is nonsingular, \( t^1, \pi \) are uniquely determined as functions of \( \dot{\epsilon}^1 \). With the solution of equation (39), equation (28.5) determines \( \dot{\epsilon}^1 \), hence \( t^1, \pi \).

Step 9  Using results of Steps 1–8, the second-order design sensitivity vector \( \psi''(b) \) can be evaluated by equation (29).

6. Computation of \( \psi''(b) \)

Computation of the second-order derivative of the objective function with respect to design variables can be translated into an initial problem of a typical differential equation.

Function \( F(t) \) is defined to be
\[
F(t) = S + \int_{t_1}^{t} T dt, \quad t \in [t^1, t^2]
\]  
(40)

\( F(t) \) is a function about time \( t \) from \( t^1 \) to \( t^2 \), its initial value about \( t^1 \) is \( S \) and its final value about \( t^2 \) is the second-order derivative of \( \psi \) with respect to \( b \), and its derivative about time \( t \) is \( T \). Then a typical differential equation with initial value is found to be
\[
F(t^1) = S, \quad \frac{dF}{dt} = T
\]  
(41)

Equation (2) and equation (3.1) determine the initial time \( t^1 \), and the solution of equation (41) yields \( F(t^2) \), that is \( \psi''(b) \).
7. Reliability of design sensitivity analysis

A small perturbation $\delta b$ is given to the design variation $b$, and then the new design variation $b^*$ is

$$b^* = b + \delta b$$

(42)

Using Taylor’s formula, second-order approximations of the objective function $\psi$ may be calculated, i.e.,

$$\psi(b^*) \approx \psi(b) + \psi'(b)\delta b + \frac{1}{2!}\delta b^T\psi''(b)\delta b$$

(43)

The actual change in the value of the constraint function $\Delta \psi$ and the design sensitivity prediction of the change in functional value $\delta \psi$ are given by

$$\Delta \psi \triangleq \psi(b^*) - \psi(b) \approx \psi'(b)\delta b + \frac{1}{2!}\delta b^T\psi''(b)\delta b \triangleq \delta \psi$$

(44)

If the value of $\Delta \psi$ is approximately equal to the value of $\delta \psi$, the design sensitivity analysis should be regarded as reliable.

8. Numerical example

Figure 1 shows a slider-crank mechanism in its initial position. The loading in this case is only the weight of the members. The state variables and the design variables are presumed to be $q = [x_1, y_1, \theta_1, x_2, y_2, \theta_2, x_3]^T$ and $b = [l_1, l_2, m_1, m_2, m_3]^T$.

Then the generalized mass matrix is

$$M = \text{diag}[m_1, m_1, J_1, m_2, m_2, J_2, m_3], \quad J_1 = \frac{1}{12}m_1l_1^2, \quad J_2 = \frac{1}{12}m_2l_2^2$$

the generalized force matrix is

$$Q = [0, -m_1g, 0, 0, -m_2g, 0, 0]^T$$

and the constraint equations of the system are

$$\Phi = [x_1 - \frac{l_1}{2}\cos\theta_1, \; y_1 - \frac{l_1}{2}\sin\theta_1, \; x_2 - l_1\cos\theta_1 - \frac{l_2}{2}\cos\theta_2, \; y_2 - l_1\sin\theta_1 - \frac{l_2}{2}\sin\theta_2, \; x_3 - l_1\cos\theta_1 - l_2\cos\theta_2, \; l_1\sin\theta_1 + l_2\sin\theta_2]^T = 0$$

Figure 1. A slider-crank mechanism. $l_1$ and $l_2$ are the lengths of the crank and the pole, $m_1$, $m_2$ and $m_3$ are the masses of the members respectively, $x_i, y_i (i = 1, 2, 3)$ are coordinates of the centre of the members, and $\theta_i (i = 1, 2)$ are the attitude angles of the crank and the pole.
The input data are given, in kg-m units, as follows

\[ m_1 = m_2 = 1 \text{ kg}, \quad m_3 = 2 \text{ kg}, \quad l_1 = 1 \text{ m}, \quad l_2 = \sqrt{3} \text{ m}, \quad \theta_1 = 30^\circ \]

and the function selected for analysis is

\[ \psi = \int_0^1 (x_3 - x_3^1)^2 \, dt \]

Using the adjoint variable method, the objective function and the first- and second-order derivative of it with respect to design variables are found to be

\[ \psi = 0.2306, \quad \psi'(\mathbf{b}) = [0.3069 \quad -0.2797 \quad -0.0640 \quad -0.0643 \quad 0.0129] \]

\[ \psi''(\mathbf{b}) = \begin{bmatrix}
-1.7663 & -3.3445 & 1.8253 & 1.8593 & -0.2492 \\
-2.5468 & -3.7673 & 2.4402 & 2.6126 & -0.1994 \\
0.0957 & -0.2716 & 0.0163 & -0.0180 & -0.0790 \\
0.0532 & -0.1555 & 0.0094 & -0.0094 & -0.0439 \\
-0.0751 & 0.2141 & -0.0130 & 0.0140 & 0.0622
\end{bmatrix} \]

Let \( \delta b_i = 0.001, (i = 1, \ldots, 5) \). Perturbation and reanalysis yield

\[ \psi = 0.2305, \quad \psi'(\mathbf{b}) = [0.3052 \quad -0.2789 \quad -0.0640 \quad -0.0642 \quad 0.0129] \]

\[ \psi''(\mathbf{b}) = \begin{bmatrix}
-1.7556 & -3.3455 & 1.8210 & 1.8541 & -0.2496 \\
-2.5323 & -3.7654 & 2.4336 & 2.6050 & -0.1991 \\
0.0951 & -0.2712 & 0.0163 & -0.0179 & -0.0789 \\
0.0528 & -0.1552 & 0.0093 & -0.0094 & -0.0438 \\
-0.0746 & 0.2138 & -0.0129 & 0.0140 & 0.0622
\end{bmatrix} \]

The predicted cost function variation is

\[ \Delta \psi = -1.0660 \times 10^{-4} \]

The comparable result is

\[ \delta \psi = -8.9843 \times 10^{-5} \]

which shows good agreement, so the results are reliable.

Let \( \mathbf{b}^0 = [1, \sqrt{3}, 1, 1, 2]^T \), and \( \mathbf{b}^j = \mathbf{b}^0 + i \cdot \delta \mathbf{b} \), with \( \delta b_j = 0.001, (j = 1, \ldots, 5) \). Using Taylor’s formula, first- and second-order approximations \( \psi(\mathbf{b}^j) \) are found to be

\[ \psi(\mathbf{b}^j) \approx \psi(\mathbf{b}^0) + \psi'(\mathbf{b}^0)(\mathbf{b}^j - \mathbf{b}^0) \]

\[ \psi(\mathbf{b}^j) \approx \psi(\mathbf{b}^0) + \psi'(\mathbf{b}^0)(\mathbf{b}^j - \mathbf{b}^0) + \frac{1}{2!}(\mathbf{b}^j - \mathbf{b}^0)^T \psi''(\mathbf{b}^0)(\mathbf{b}^j - \mathbf{b}^0) \]

Approximations are compared with the results that were calculated directly in figure 2. The real line in the figure denotes the results that were calculated directly through \( \psi \), the dotted line denotes the first-order approximation of \( \psi \), and the dash-dot line denotes the second-order approximation of \( \psi \). The labeled dots denote the results about each design variable \( \mathbf{b}^j (i = 0, \ldots, 10) \). It shows that the second-order approximation is much more accurate than the first-order approximation.
9. Conclusions

The adjoint variable method has a higher efficiency for sensitivity analysis of multibody system dynamics with a large number of parameters. For a system involving \( p \) design variables, \( n \) generalized coordinates, and \( m \) independent algebraic constraint equations, the solution of \( 2(n + m) \) differential algebraic equations are required for first-order sensitivity analysis and \( 4(n + m) \) differential algebraic equations are required for second-order sensitivity analysis with the method presented in this paper. While for the direct differentiation method, \( (n + m)(p + 1) \) differential algebraic equations must be solved for first-order sensitivity analysis [13] and about \( (n + m)(p^2 + p + 1) \) differential algebraic equations need to be solved for second-order sensitivity analysis [3]. The conclusion is relatively clear. For large numbers of design variables, the adjoint variable method is superior.

During the implementation, we simply use the same fixed step method to synchronize the forward and backward process of dynamic analysis and adjoint variable equations. Application of adaptive step methods with interpolation for synchronization will improve the implementation greatly of such an adjoint variable method.

References


