# IMPLICIT RUNGE–KUTTA METHODS FOR LIPSCHITZ CONTINUOUS ORDINARY DIFFERENTIAL EQUATIONS[*]

XIAOJUN CHEN[†] AND SAYED MAHMOUD[‡]

**Abstract.** Implicit Runge–Kutta (IRK) methods for solving the nonsmooth ordinary differential equation (ODE) involve a system of nonsmooth equations. We show superlinear convergence of the slanting Newton method for solving the system of nonsmooth equations. We prove the slanting differentiability and give a slanting function for the involved function. We develop a new code based on the slanting Newton method and the IRK method for nonsmooth ODEs arising from structural oscillation and pounding. We show that the new code is efficient for solving a nonsmooth ODE model for the collapse of the Tacoma Narrows suspension bridge and simulating 13 different earthquakes.

**Key words.** implicit Runge–Kutta method, slanting Newton method, nonsmooth equations, ordinary differential equation

**AMS subject classifications.** 65H10, 65L05, 65L06

**DOI.** 10.1137/070688699

**1. Introduction.** Let $D \subset R^n$ be a domain, $p : D \to R^n$ be a continuously differentiable function, and $f : D \to R^n$ be a continuous function. The function $f$ is not necessarily differentiable.

We consider the following system of nonsmooth ordinary differential equations of the first order:

$$\text{(1.1)} \qquad \frac{dp(u)}{dt} = f(u), \qquad t \geq 0,$$
$$u(0) = u_0,$$

where $u \in R^n$. Here smoothness refers to continuous differentiability. It is known that if $f$ is Lipschitz continuous, then (1.1) has a unique solution in a certain interval $[0, T]$ [9]. Moreover, it is shown in [2] that $f$ is Lipschitz continuous if and only if $f$ is slantly differentiable. A slanting Newton method can be defined in appropriate function spaces. The concept of slanting differentiability of $f$ means that $f$ has a slanting function $f^o$ such that

$$f(u + \triangle u) - f(u) - f^o(u + \triangle u)\triangle u = o(\|\triangle u\|).$$

This property ensures that the slanting Newton method

$$u^{k+1} = u^k - f^o(u^k)^{-1}f(u^k)$$

has a superlinear convergence rate [2].

Note that, though the function $f$ is nonsmooth, the solution of (1.1) may be smooth (continuously differentiable). See the example of the collapse of the Tacoma

[†]Department of Applied Mathematics, The Hong Kong Polytechnic University, Hung Hom Kowloon, Hong Kong (maxjchen@polyu.edu.hk).

[‡]Department of Mathematical Sciences, Hirosaki University, Hirosaki 036-8561, Japan (elseedy@hotmail.com).

1266

Narrows suspension bridge in section 4. In this paper, we consider numerical solution of the nonsmooth ODEs (1.1).

One step of an $s$-stage implicit Runge–Kutta (IRK) method for solving (1.1) has the following version [8].

**IRK method.** Given a step size $h$, a coefficient matrix $A \in R^{s \times s}$, and a weight vector $b \in R^s$, let $U_0 = u_0$. For $k \geq 0$ do the following.

*Step* 1. Solve the $s \times n$-dimensional system of nonlinear equations

$$(1.2) \qquad H(x) := \begin{pmatrix} p(x_1) - p(U_k) - h \sum_{j=1}^{s} a_{1j} f(x_j) \\ \vdots \\ p(x_s) - p(U_k) - h \sum_{j=1}^{s} a_{sj} f(x_j) \end{pmatrix} = 0$$

to get a solution $x^k = (x_1^k, x_2^k, \ldots, x_s^k)^T \in R^{s \times n}$.

*Step* 2. Solve the $n$-dimensional system of nonlinear equations

$$(1.3) \qquad \tilde{H}(U) := p(U) - p(U_k) - h \sum_{j=1}^{s} b_j f(x_j^k) = 0,$$

and let the solution be $U_{k+1}$.

A practical IRK method can be defined, by choosing appropriate matrix $A$ and vector $b$, such as coefficients of Gauss, Radau IA & IIA, Lobatto IIIA, Burrage, etc. [1, 8]. Recently, much attention has been paid to choosing $A, b$ such that the IRK method has the best properties in some sense of stability [5]. Moreover, Jay [8] showed that for various choices of $A, b$ the use of inexact simplified Newton methods is efficient for solving the system of nonlinear equations (1.2) under the condition that $H$ is continuously differentiable. In this paper, we focus our attention on how to solve the system of nonlinear equations (1.2) efficiently when $f$ is not differentiable. Such nonsmooth problems arise from mathematical models of structural pounding earthquake and structural oscillations [7, 10, 11, 14]. We apply the slanting Newton method [2, 6] to solve (1.2). In section 2, we show that $H$ is slantly differentiable if $f$ is piecewise continuously differentiable. Moreover, we give a simple method to compute a slanting function for $H$. In section 3, we discuss the convergence order of IRK methods for nonsmooth ODEs. In section 4, we illustrate the slanting Newton method by using a simple model for the collapse of the Tacoma Narrows suspension bridge [14]. Furthermore, we develop a code based on the slanting Newton method and the IRK method for nonsmooth ODEs arising from structural oscillation and pounding. A suite of 27 ground motion records from 12 different earthquakes and a record from the Kobe earthquake are used to show that the code is efficient for simulation of a structural pounding earthquake. All data used in the numerical experiments were taken from the PEER Strong Motion Database [15].

**2. Slanting Newton iterations for IRK methods.** Let $X$ and $Y$ be Banach spaces and $D$ be an open domain in $X$. Let $L(X, Y)$ denote the set of all bounded linear operators on $X$ into $Y$.

DEFINITION 2.1 (see [2]). *A function $H : D \subset X \to Y$ is said to be slantly differentiable at $x \in D$ if there exists a mapping $H^o : D \to L(X, Y)$ such that the*

*family $\{H^o(x+\triangle x)\}$ of bounded linear operators is uniformly bounded in the operator norm for h sufficiently small and*

$$\lim_{\triangle x \to 0} \frac{H(x+\triangle x) - H(x) - H^o(x+\triangle x)\triangle x}{\|\triangle x\|} = 0.$$

*The function $H^o$ is called a slanting function for H at x.*

DEFINITION 2.2 (see [2,4]). *A function $H : X \to Y$ is said to be Lipschitz continuous at x if there is a constant K such that for all sufficiently small $\triangle x$*

$$\|H(x+\triangle x) - H(x)\| \le K\|\triangle x\|.$$

LEMMA 2.1 (see [2]). *A function $H : X \to Y$ is slantly differentiable at x if and only if H is Lipschitz continuous at x.*

LEMMA 2.2 (see [2]). *Suppose that H is slantly differentiable at a solution $x^*$ of $H(x) = 0$. Let $H^o$ be a slanting function for H at $x^*$ and $\|H^o(x)^{-1}\| \le \Gamma$ in a neighborhood $\mathcal{N}$ of $x^*$, where $\Gamma$ is a positive constant. Then the iterative sequence $\{x^m\}$ generated by the slanting Newton method*

$$(2.1) \qquad\qquad x^{m+1} = x^m - H^o(x^m)^{-1}H(x^m)$$

*superlinearly converges to $x^*$ in a neighborhood of $x^*$.*

DEFINITION 2.3 (see [12]). *We say $f : R^n \to R^n$ is piecewise continuously differentiable if it is continuous and there is a finite collection of continuously differentiable functions $\phi_l : R^n \to R, l \in \mathcal{L} := \{1, 2, \ldots, L\}$, such that*

$$f_i(u) \in \{\phi_l(u), \ l \in \mathcal{L}\}, \ \text{for any } u \in R^n, \quad i \in \{1, 2, \ldots, n\}.$$

*Such a collection is called a representation for f on $R^n$.*

THEOREM 2.1. *The function $H : R^{s \times n} \to R^{s \times n}$ defined in (1.2) is slantly differentiable at any point $x \in R^{s \times n}$ if f is piecewise continuously differentiable.*

*Proof.* We first show that, for a fixed point $\bar{u} \in R^n$ and a fixed index $i \in \{1, 2, \ldots, n\}$, $f_i$ is Lipschitz continuous at $\bar{u} \in R^n$. Let $\mathcal{I}$ and $\mathcal{J}$ be two index sets defined by

$$\mathcal{I} = \{l \mid f_i(\bar{u}) = \phi_l(\bar{u}), l \in \mathcal{L}\}$$

and

$$\mathcal{J} = \{i \mid f_i(\bar{u}) \ne \phi_l(\bar{u}), l \in \mathcal{L}\},$$

respectively. Since $\mathcal{J}$ is a finite set, we have

$$(2.2) \qquad\qquad r = \min_{l \in \mathcal{J}} \|\phi_l(\bar{u}) - f_i(\bar{u})\| > 0.$$

Since $f_i$ is continuous in $R^n$, there are positive constants $\delta_0$ and $\alpha < 1$ such that if $\|u - \bar{u}\| < \delta_0$, then

$$(2.3) \qquad\qquad \|f_i(u) - f_i(\bar{u})\| \le \alpha r.$$

Furthermore, since every $\phi_l, l \in \mathcal{L}$, is continuously differentiable, there are positive constants $K_l$, $\delta_l$ such that if $\|u - \bar{u}\| < \delta_l$, then

$$(2.4) \qquad\qquad \|\phi_l(u) - \phi_l(\bar{u})\| \le K_l\|u - \bar{u}\|, \quad l \in \mathcal{L}.$$

Let

$$K = \max_{l \in \mathcal{L}} K_l \quad \text{and} \quad \delta = \min\left\{\delta_0, \frac{1-\alpha}{K}r, \min_{l \in \mathcal{L}} \delta_l\right\}.$$

For any $u$ satisfying $\|u - \bar{u}\| < \delta$, if $f_i(u) = \phi_l(u), l \in \mathcal{I}$, then we have

$$\|f_i(u) - f_i(\bar{u})\| = \|\phi_l(u) - \phi_l(\bar{u})\| \le K\|u - \bar{u}\|.$$

Otherwise, if $f_i(u) = \phi_j(u), j \in \mathcal{J}$, we apply (2.2)–(2.4) and obtain

$$\begin{aligned}
\alpha r &\ge \|f_i(u) - f_i(\bar{u})\| \\
&= \|\phi_j(u) - f_i(\bar{u})\| \\
&\ge \|\phi_j(\bar{u}) - f_i(\bar{u})\| - \|\phi_j(u) - \phi_j(\bar{u})\| \\
&\ge r - K\|u - \bar{u}\| > \alpha r,
\end{aligned}$$

which implies that there is no $j \in \mathcal{J}$ such that $f_i(u) = \phi_j(u), j \in \mathcal{J}$. Hence

(2.5)                    $$\|f_i(u) - f_i(\bar{u})\| \le K\|u - \bar{u}\|.$$

Now we show that $H$ is Lipschitz continuous at a point $x^* = (x_1^*, \dots, x_s^*) \in R^{s \times n}$ for the norm $\|\cdot\|_\infty$.

Note that $p$ is continuously differentiable in $R^n$, and the point $\bar{u} \in R^n$, the index $i \in \{1, 2, \dots, n\}$, and the norm in (2.2)–(2.5) are arbitrarily chosen. Without loss of generality, we may assume that there are a constant $K$ and a neighborhood $\mathcal{N}^* := \mathcal{N}_1^* \times \mathcal{N}_2^* \dots \mathcal{N}_s^*$ of $x^*$ such that, for any $x \in \mathcal{N}^*$,

$$\|f(x_i) - f(x_i^*)\|_\infty \le K\|x_i - x_i^*\|_\infty$$

and

$$\|p(x_i) - p(x_i^*)\|_\infty \le K\|x_i - x_i^*\|_\infty, \quad i = 1, 2 \dots, s.$$

Therefore, by the definition of $H$, we find

$$\begin{aligned}
\|H(x) - H(x^*)\|_\infty &\le \max_{1 \le i \le s} \|p(x_i) - p(x_i^*)\|_\infty + h\|A\|_\infty s \max_{1 \le i \le s} \|f(x_i) - f(x_i^*)\|_\infty \\
&\le \max_{1 \le i \le s} K\|x_i - x_i^*\|_\infty + hs\|A\|_\infty \max_{1 \le i \le s} K\|x_i - x_i^*\|_\infty \\
&\le (1 + hs\|A\|_\infty)K \max_{1 \le i \le s} \|x_i - x_i^*\|_\infty \\
&\le (1 + hs\|A\|_\infty)K\|x - x^*\|_\infty.
\end{aligned}$$

By Lemma 2.1, $H$ is slantly differentiable at $x^*$.    □

THEOREM 2.2. *Let $f^o \in R^{n \times n}$ be a slanting function for $f$ at $x_i^*$, $i = 1, 2, \dots, s$. Then*

(2.6)      $$H^o(x) = \text{diag}(p'(x_1), \dots, p'(x_s)) - h(A \otimes I_n)\text{diag}(f^o(x_1), \dots, f^o(x_s))$$

*is a slanting function for $H$ at $x^*$, where* diag *denotes block diagonal, $I_n \in R^{n \times n}$ is the identity matrix, and the symbol $\otimes$ denotes the tensor product.*

*Proof.* Since $p$ is continuously differentiable and $f^o$ is a slanting function for $f$ at $x_i^*$, $i = 1, 2, \dots, s$, we have that, if $\|x - x^*\|$ is sufficiently small,

$$p(x_i) - p(x_i^*) = p'(x_i)(x_i - x_i^*) + o(\|x_i - x_i^*\|), \quad i = 1, 2, \dots s,$$

and

$$f(x_i) - f(x_i^*) = f^o(x_i)(x_i - x_i^*) + o(\|x_i - x_i^*\|), \quad i = 1, 2, \ldots, s.$$

Therefore, we get

$$H(x) - H(x^*)$$

$$= \begin{pmatrix} p(x_1) - p(x_1^*) \\ \vdots \\ p(x_s) - p(x_s^*) \end{pmatrix} - h \begin{pmatrix} \sum_{j=1}^{s} a_{1j}(f(x_j) - f(x_j^*)) \\ \vdots \\ \sum_{j=1}^{s} a_{sj}(f(x_j) - f(x_j^*)) \end{pmatrix}$$

$$= \begin{pmatrix} p'(x_1)(x_1 - x_1^*) + o(\|x_1 - x_1^*\|) \\ \vdots \\ p'(x_s)(x_s - x_s^*) + o(\|x_s - x_s^*\|) \end{pmatrix} - h \begin{pmatrix} \sum_{j=1}^{s} a_{1j}(f^o(x_j)(x_j - x_j^*) + o(\|x_j - x_j^*\|)) \\ \vdots \\ \sum_{j=1}^{s} a_{sj}(f^o(x_j)(x_j - x_j^*) + o(\|x_j - x_j^*\|)) \end{pmatrix}$$

$$= \begin{pmatrix} p'(x_1) & & \\ & \ddots & \\ & & p'(x_s) \end{pmatrix} \begin{pmatrix} x_1 - x_1^* \\ \vdots \\ x_s - x_s^* \end{pmatrix} - h \begin{pmatrix} a_{11}f^o(x_1) & \ldots & a_{1s}f^o(x_s) \\ \vdots & & \vdots \\ a_{s1}f^o(x_1) & \ldots & a_{ss}f^o(x_s) \end{pmatrix} \begin{pmatrix} x_1 - x_1^* \\ \vdots \\ x_s - x_s^* \end{pmatrix}$$
$$+ o(\|x - x^*\|)$$

$$= (\mathrm{diag}(p'(x_1), \ldots, p'(x_s)) - h(A \otimes I_n)\mathrm{diag}(f^o(x_1), \ldots, f^o(x_s)))(x - x^*) + o(\|x - x^*\|).$$

Hence $H^o$ defined in (2.6) is a slanting function for $H$ at $x^*$. $\quad\square$

Now we consider how to compute a slanting function $f^o$ for a piecewise continuously differentiable function $f$. From Theorem 2.1, $f$ is a locally Lipschitzian function. By the Rademacher theorem, $f$ is differentiable almost everywhere. Hence, we can define the Clarke generalized Jacobian [3]

$$\partial f(y) = \mathrm{co}\{ \lim_{\substack{y^k \to y \\ y^k \in D_f}} f'(y^k) \},$$

where the symbol co denotes the convex hull and $D_f$ is the set of points where $f$ is differentiable. In [2], it is shown that any single valued selection of $\partial f(y)$ is a slanting function of $f$ at $y$. Furthermore, from Lemma 2 in [12], if $f_i$ is differentiable at $y$, then there exists $\phi_l$ in the representation for $f$ such that

$$f_i(y) = \phi_l(y) \quad \text{and} \quad f_i'(y) = \phi_l'(y).$$

Therefore, we can define a slanting function of $f$ as

$$(f^o(y))_i = \phi_l'(y), \qquad \text{where } f_i(y) = \phi_l(y).$$

For such a slanting function, by Theorems 2.1 and 2.2 and Lemma 2.1, we can easily obtain the following corollary.

COROLLARY 2.1. *Let $x^*$ be a solution of $H(x) = 0$. If $f$ is piecewise continuously differentiable, and $p'(x_i^*)$, $i = 1, 2, \ldots, s$, are nonsingular, then for small $h$ the slanting Newton method (2.1) is well defined and superlinearly converges to $x^*$.*

*Proof.* By the continuity of $p'$ and the nonsingularity of $p'(x_1^*)$, there are a positive constant $\gamma_1$ and a neighborhood $\mathcal{N}_1$ of $x^*$ such that, for any $x \in \mathcal{N}_1$, $p'(x_i)$ are all nonsingular and $\|p'(x_i)^{-1}\|_\infty \leq \gamma_1$, $i = 1, 2, \ldots, s$. Moreover, by the locally Lipschitzian continuity of $f$, there are a positive constant $\gamma_2$ and a neighborhood $\mathcal{N}_2$ of $x^*$ such that, for any $x \in \mathcal{N}_2$, $\|f^o(x_i)\|_\infty \leq \gamma_2$, $i = 1, 2, \ldots, s$. Let

$$B = h \mathrm{diag}(p'(x_1)^{-1}, \ldots, p'(x_s)^{-1})(A \otimes I)\mathrm{diag}(f^o(x_1), \ldots, f^o(x_s)),$$

$$\gamma = s\gamma_1\gamma_2\|A\|_\infty,$$

and $\mathcal{N} = \mathcal{N}_1 \cap \mathcal{N}_2$. Then, for $h \leq \lambda/\gamma$ ($\lambda < 1$) and $x \in \mathcal{N}$, we have

$$\|B\|_\infty \leq \lambda < 1.$$

Hence, for $h \leq \lambda/\gamma$ ($\lambda < 1$) and $x \in \mathcal{N}$, $H^o(x)$ is nonsingular and

$$
\begin{aligned}
\|H^o(x)^{-1}\| &= \|\mathrm{diag}(p'(x_i)^{-1})(I - B)\|_\infty \\
&\leq \|\mathrm{diag}(p'(x_i)^{-1})\|_\infty \|I + B + B^2 + \cdots\|_\infty \\
&\leq \frac{\gamma}{1 - \lambda} \\
&=: \Gamma.
\end{aligned}
$$

By Lemma 2.2, the slanting Newton method (2.1) is well defined and superlinearly converges to $x^*$.  □

*Remark* 2.1. When $H$ is continuously differentiable, Jay [8] suggested to use the simplified Newton method

$$(2.7) \qquad x^{m+1} = x^m - H'(x^0)^{-1}H(x^m)$$

for solving the system of nonlinear equations in IRK methods. If $x^0$ is sufficiently close to a solution $x^*$ of $H(x) = 0$,

$$\Phi(x) := x - H'(x^0)^{-1}H(x)$$

is locally contractive in a neighborhood of $x^*$ containing $x^0$. By using the contractive property, Jay proved the convergence of an inexact simplified Newton method. Note that, when $H$ is continuously differentiable, $H'(x)$ is a slanting function $H^o(x)$ for $H$. However, if $H$ is not differentiable at $x^*$,

$$\Psi(x) := x - H^o(x^0)^{-1}H(x)$$

is not locally contractive in any neighborhood of $x^*$ containing $x^0$. For example,

$$H(x) = \begin{cases} x^2 - 1, & x \geq 1, \\ 1 - x, & x < 1. \end{cases}$$

The function $H$ is not differentiable at the solution $x^* = 1$ of $H(x) = 0$. It is easy to get a slanting function for $H$:

$$H^o(x) = \begin{cases} 2x, & x \geq 1, \\ -1, & x < 1. \end{cases}$$

For $x^0 < 1$ and $x \geq 1$,

$$|\Psi(x) - \Psi(x^*)| = |x - H^o(x^0)^{-1}H(x) - x^*| = |x + (x^2 - 1) - 1| = |x + 2||x - x^*| > |x - x^*|.$$

For $x^0 \geq 1$ and $0 < x < 1$,

$$|\Psi(x) - \Psi(x^*)| = |x - H^o(x^0)^{-1}H(x) - x^*|$$

$$= \left| x - \frac{1-x}{2x^0} - 1 \right| = \left| 1 + \frac{1}{2x^0} \right| |x - x^*| > |x - x^*|.$$

Therefore, the simplified Newton method cannot be applied to solve (1.2) when $H$ is not differentiable.

*Remark* 2.2. Analysis in this section can be easily generalized to the system of nonautonomous ordinary differential equations

$$(2.8) \qquad \frac{dp(t, u)}{dt} = f(t, u), \qquad t \geq 0,$$

$$u(0) = u_0.$$

**3. Convergence order of IRK methods.** We have tested various IRK methods with the slanting Newton method and the bisection method on numerous problems in structural oscillation and pounding. From our numerical experiments, we observe that the order of the IRK methods for Lipschitz continuous ODEs can be preserved if there are finite discontinuous times, and we can find these discontinuous times sufficiently accurately. However, in many cases the order of convergence may drop to one. In theory, we can show that the order of convergence is at least one for the Lipschitz continuous ODEs under mild conditions. This is done for the solution $u(t)$ to (1.1) in a fixed interval $[0, T]$ with the number $n$ of steps chosen such that $t_n = nh = T$. Let

$$e_k(h) = U_k - u(t_k), \qquad k = 0, 1, \ldots, n,$$

and

$$E(h) = \max_{k=0,\ldots,n} \|e_k(h)\|.$$

We assume that the IRK method is well defined; that is, for any $U_k$ there are $x^k$, $U_{k+1}$ such that $H(x^k) = 0$ and $\tilde{H}(U_{k+1}) = 0$. For simplicity, we consider $p(u) = u$. Let

$$G(x, U) := \begin{pmatrix} x_1 - U - h \sum_{j=1}^{s} a_{1j} f(x_j) \\ \vdots \\ x_s - U - h \sum_{j=1}^{s} a_{sj} f(x_j) \end{pmatrix}.$$

Since $f$ is Lipschitz continuous, the function $G$ is Lipschitz continuous. By the Rademacher theorem, $G$ is differentiable almost everywhere. Hence we can define the Clarke generalized Jacobian [3]

$$\partial G(x, U) = [\pi_x \partial G(x, U), \pi_U \partial G(x, U)],$$

where $\pi_x \partial G(x, U)$ signifies the set of all $(s \times n) \times (s \times n)$ matrices $M$ such that, for some $(s \times n) \times n$ matrix $N$, the $(s \times n) \times (s \times n + n)$ matrix $[M, N]$ belongs to $\partial G(x, U)$.

For small $h$, $\pi_x \partial G(x^k, U_k)$ is of maximal rank; i.e., every matrix $M$ in $\pi_x \partial G(x^k, U_k)$ is nonsingular. By the implicit function theorem for Lipschitz continuous function in [3], there exist a neighborhood $\mathcal{U}_k$ of $U_k$ and a Lispchitz function $\psi_k(\cdot; h) : \mathcal{U}_k \to R^{s \times n}$ such that $x^k = \psi_k(U_k; h)$ and, for every $U \in \mathcal{U}_k$, $G(\psi_k(U; h), U) = 0$. Therefore, we have

$$(3.1) \qquad U_{k+1} = U_k + h \sum_{j=1}^{s} b_j f((\psi_k(U_k; h))_j) =: U_k + h \phi_k(U_k; h).$$

Obviously $\phi_k(\cdot; h) : R^n \to R^n$ is a locally Lipschitz continuous function.

THEOREM 3.1. *Suppose that there are positive constants $K_1$ and $K_2$ such that*

$$(3.2) \qquad \|\phi_k(u(t_k); h) - \phi_k(U_k; h)\| \le K_1 \|u(t_k) - U_k\|$$

*and*

$$(3.3) \qquad \|\phi_k(u(t_k); h) - u'(t_k)\| \le K_2 h.$$

*Then there is a constant $\alpha > 0$ such that*

$$E(h) \le \alpha h.$$

*Proof.* By the Lipschitz continuity of $f$, $u'$ is Lipschitz continuous. This implies that there is a constant $K_3 > 0$ such that

$$(3.4) \qquad \left\| \frac{u(t_k + h) - u(t_k)}{h} - u'(t_k) \right\| \le K_3 h.$$

By (3.1), we have

$$e_{k+1} - e_k = U_{k+1} - U_k - (u(t_{k+1}) - u(t_k)) = h \left( \phi_k(U_k; h) - \frac{u(t_k + h) - u(t_k)}{h} \right).$$

Hence from (3.2)–(3.4), we obtain

$$\|e_{k+1}\| \le (1 + hK_1)\|e_k\| + (K_2 + K_3)h^2$$

$$= (1 + hK_1)^{k+1}\|e_0\| + \frac{K_2 + K_3}{K_1} \left( K_1 h \sum_{j=0}^{k} (1 + K_1 h)^j \right) h,$$

which, together with $\|e_0\| = 0$, implies that

$$E(h) \le \frac{K_2 + K_3}{K_1}(e^{K_1 T} - 1)h. \qquad \square$$

If $s = 1$, $a_{11} = 0$, and $b_1 = 1$, then $\psi_k(U_k; h) = U_k$, $\phi_k(U_k; h) = f(U_k)$, and (3.1) reduces to the Euler method which has convergence order one. Obviously, (3.2) and (3.3) hold for $\phi_k(U_k; h) = f(U_k)$ and $\phi_k(u(t_k); h) = f(u(t_k))$. Theorem 3.1 shows that the IRK methods for Lispchitz continuous ODEs have convergence order one as the Euler method if $u'(t)$ is not differentiable at some points in the interval $[t_k, t_{k+1}]$ for some $k$. It is worth noting that the implicit RK methods are numerically stable but require more computational time to solve a system of nonsmooth equations at each step, while explicit methods are faster but may cause a numerical stability problem for stiff ODEs. It will be interesting to study a hybrid method that takes advantage of both implicit and explicit RK methods for solving Lipschitz continuous ODEs.

**4. Numerical experiment.** Jay [8] developed a code based on 3-stage IRK methods for the numerical solution of (1.1) where $f$ is continuously differentiable. To solve (1.1) where $f$ is not differentiable, we developed a new code based on the slanting Newton method and Jay's code. In our code, the function $f$ can be defined by a finite collection of continuously differentiable functions $\{\phi_l, l \in \mathcal{L}\}$. See Definition 2.3 and Theorem 2.1. The set of points where $f$ is not differentiable is

$$\mathcal{S}_f = \{u \in D \mid \phi_i(u) - \phi_j(u) = 0, \phi'_i(u) - \phi'_j(u) \neq 0, i, j \in \mathcal{L}\}.$$

If a point in $\mathcal{S}_f$ is contained within an integration interval, then the order of the IRK method may drop to one. If we can compute exactly the discontinuity times, then the order of the IRK method can be preserved. See the example in subsection 4.1. In our code, a root-finding process is used to find the discontinuities in time. We tested the new code by using many nondifferentiable ODEs. Numerical results show that the new code is efficient. In this section, we report numerical results of IRK methods with the 2-stage Burrage coefficient [1]

$$A = \begin{pmatrix} 1/4 & 0 \\ 1/2 & 1/4 \end{pmatrix} \quad \text{and} \quad b = \begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix},$$

the 2-stage Radau IA coefficient [13]

$$A = \begin{pmatrix} 1/4 & -1/4 \\ 1/4 & 5/12 \end{pmatrix} \quad \text{and} \quad b = \begin{pmatrix} 1/4 \\ 3/4 \end{pmatrix},$$

the 2-stage Radau IIA coefficient [13]

$$A = \begin{pmatrix} 5/12 & -1/12 \\ 3/4 & 1/4 \end{pmatrix} \quad \text{and} \quad b = \begin{pmatrix} 3/4 \\ 1/4 \end{pmatrix},$$

and the 3-stage Lobatto IIIA coefficient [8]

$$A = \begin{pmatrix} 0 & 0 & 0 \\ 5/24 & 1/3 & -1/24 \\ 1/6 & 2/3 & 1/6 \end{pmatrix} \quad \text{and} \quad b = \begin{pmatrix} 1/6 \\ 2/3 \\ 1/6 \end{pmatrix}$$

for solving a nondifferentiable ODE model for the collapse of the Tacoma Narrows suspension bridge and simulating 13 different earthquakes with a suite of 28 ground motion records.

**4.1. Bridge collapse.** The collapse of the Tacoma Narrows suspension bridge in 1940 left many open questions about the collapse. Lazer and McKenna [10] contend that the nonlinear effects were the main factors leading to the large oscillations of the bridge. The following ODE is a simple version of their model [14]:

(4.1)                $m\ddot{u} + q(u) = g(t), \qquad u(0) = 0, \quad \dot{u}(0) = \gamma > 0,$

where

$$q(u) = \begin{cases} \alpha u, & u \geq 0, \\ \beta u, & u < 0. \end{cases}$$

Here $m$ is the mass of the section of the roadway, $g$ is the applied force, $q$ is an upward restoring force when $u \geq 0$ and a downward restoring force when $u < 0$, and $\alpha$ and

$\beta$ are Hooke's constants for the tension and compression, respectively. Note that the function $q(u)$ can be written as

$$q(u) = \alpha \max(0, u) - \beta \max(0, -u).$$

It is shown in [6] that the mapping $\max(0, \cdot) : L^q(\Omega) \to L^p(\Omega)$, with $1 \le p < q \le \infty$, is slantly differentiable on $L^q(\Omega)$ and

$$G(u)(t) = \begin{cases} 1, & u(t) > 0, \\ 0, & u(t) < 0, \\ \delta, & u(t) = 0, \end{cases}$$

is a slanting function for the mapping $\max(0, \cdot)$, where $\delta$ is a fixed arbitrary number. By the linear combination [2], the function $q : L^q(\Omega) \to L^p(\Omega)$, with $1 \le p < q \le \infty$, is slantly differentiable on $L^q(\Omega)$ and

$$q^o(u)(t) = \begin{cases} \alpha, & u(t) > 0, \\ \beta, & u(t) < 0, \\ \delta, & u(t) = 0, \end{cases}$$

is a slanting function for $q$. Furthermore, this problem can be equivalently written as (2.8), by resetting

$$u = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}, \qquad p(u) = \begin{pmatrix} u_1 \\ mu_2 \end{pmatrix}, \quad \text{and} \quad f(t, u) = \begin{pmatrix} u_2 + \gamma \\ g(t) - q(u_1) \end{pmatrix}.$$

Hence, we can apply the IRK method with the slanting Newton method to obtain numerical solution $U^k, k \ge 0$, of this problem. Moreover, if $q(U^k)q(U^{k-1}) < 0$, then there may be a $\tilde{t} \in [t_{k-1}, t_k]$ such that $q(u(\tilde{t})) = 0$. We used the bisection method to find the discontinuity time $\tilde{t}$ and recomputed $U^k$ by the IRK method with $\tilde{t}$ and $\tilde{h} = \tilde{t} - t_{k-1}$.

For $\alpha = 4, \beta = 1, \gamma = 1$, and $g(t) = \sin 4t$ (Example 1, pp. 264 in [14]), the exact solution in the interval $[0, 3\pi]$ is

$$u^*(t) = \begin{cases} (\frac{2}{3} - \frac{1}{6}\cos 2t)\sin 2t, & 0 \le t \le \frac{\pi}{2}, \\ (\frac{7}{5} - \frac{4}{15}\sin t \cos 2t)\cos t, & \frac{\pi}{2} \le t \le \frac{3\pi}{2}, \\ (-\frac{11}{15} - \frac{1}{6}\cos 2t)\sin 2t, & \frac{3\pi}{2} \le t \le 2\pi, \\ (-\frac{23}{15} - \frac{4}{15}\cos t \cos 2t)\sin t, & 2\pi \le t \le 3\pi, \end{cases}$$

which is twice continuously differentiable in $[0, 3\pi]$.

We chose step sizes

$$h_l = \frac{3\pi}{N_l}, \quad N_l = 100 \times 2^l, \quad l = 0, \dots, 6,$$

so that

$$\log h_{l-1} - \log h_l = \log 2, \quad l = 1, \dots, 6.$$

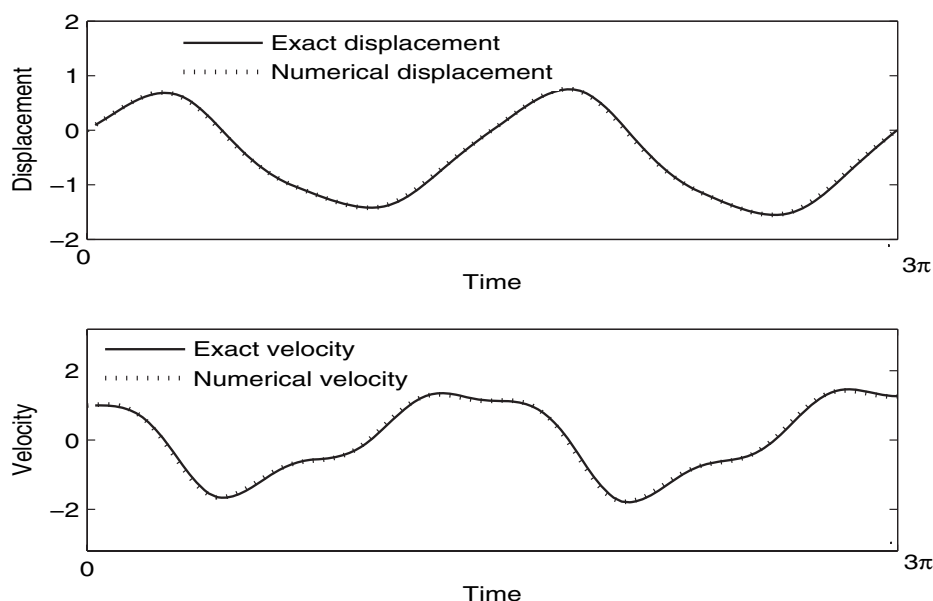We obtain approximate solutions $\{U^k\}_{k=0}^{N_l}$, $l = 0, \dots, 6$.

FIG. 4.1. *Exact solution and numerical solution for the collapse of the Tacoma Narrows suspension bridge.*

Figure 4.1 shows what the exact solution and the velocity look like and the numerical solution obtained by the IRK method with the Radau IA and $h_l = \frac{3\pi}{100}$. Figure 4.2 shows the error of the approximate solution

$$\|U_{N_l} - u^*(t_{N_l})\|_2, \quad l = 0, \dots, 6,$$

at the end point $t_{N_l} = 3\pi$ to compare the 2-stage Burrage, Radau IA, and Radau IIA coefficients. The logarithmic scale is used for the step sizes and errors so that the order of convergence can be seen clearly, when $\|U_{N_l} - u^*(t_{N_l})\|_2 = O(h^l)$. Figure 4.2 shows that the IRK method has order two of convergence for this nonsmooth ODE.

**4.2. Seismic pounding.** We consider the following nonlinear model of seismic pounding between two adjacent structures called $B1$ and $B2$.

For $i = 1, 2$, let $m_i$ be the masses, $r_i$ be the viscous damping coefficients and $k_i$ be the initial stiffness for $B1$ and $B2$, respectively. The coupling equation of motion for two adjacent buildings subjected to horizontal ground motion $\ddot{u}_g$ has the following version [7, 11]:

$$(4.2) \qquad \begin{array}{rcl} m_1\ddot{u}_1 + r_1\dot{u}_1 + k_1 u_1 + q(u_1, u_2, \dot{u}_1, \dot{u}_2) &=& -m_1\ddot{u}_g, \\ m_2\ddot{u}_2 + r_2\dot{u}_2 + k_2 u_2 - q(u_1, u_2, \dot{u}_1, \dot{u}_2) &=& -m_2\ddot{u}_g, \end{array}$$

where $q$ is the pounding force

$$q(u_1, u_2, \dot{u}_1, \dot{u}_2) = \begin{cases} \alpha(u_1 - u_2 - d)^\gamma + \beta(\dot{u}_1 - \dot{u}_2) & \text{if } u_1 - u_2 > d, \ \dot{u}_1 - \dot{u}_2 > 0, \\ \alpha(u_1 - u_2 - d)^\gamma & \text{if } u_1 - u_2 > d, \ \dot{u}_1 - \dot{u}_2 \le 0, \\ 0 & \text{if } u_1 - u_2 \le d. \end{cases}$$

Here $d$ is the initial separation distance between $B1$ and $B2$, and $\alpha > 0$, $\beta > 0$, and $\gamma > 1$ are fixed parameters.
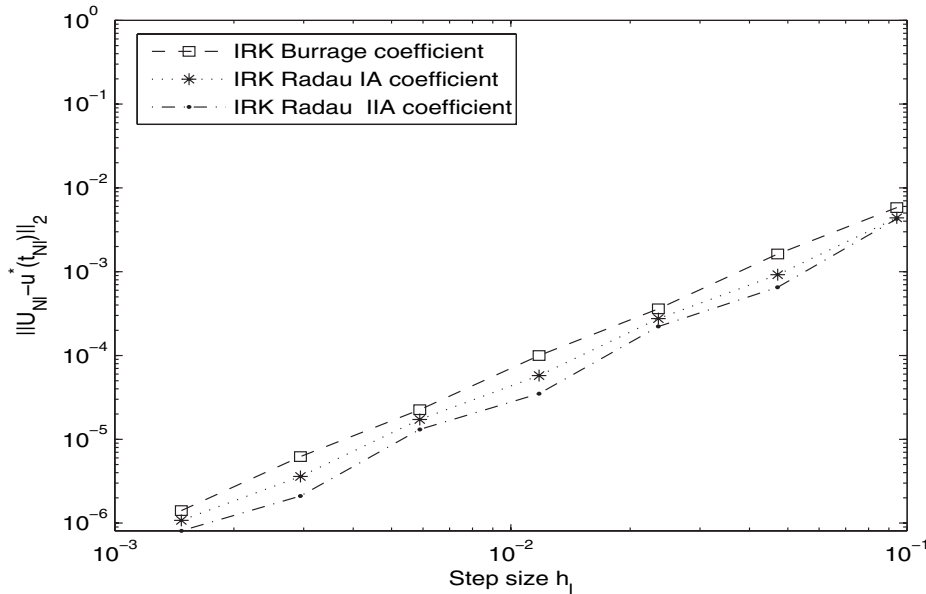
FIG. 4.2. *Error at $t_{N_l} = 3\pi$ of the IRK method with Burrage and Radau coefficients and various step sizes $h_l$ for the collapse of the Tacoma Narrows suspension bridge.*

The ordinary differential equation (4.2) is equivalent to the following system:

$$(4.3) \quad \begin{aligned} \dot{u}_1 &= u_3, \\ \dot{u}_2 &= u_4, \\ m_1 \dot{u}_3 &= -r_1 u_3 - k_1 u_1 - q(u_1, u_2, u_3, u_4) - m_1 \ddot{u}_g, \\ m_2 \dot{u}_4 &= -r_2 u_4 - k_2 u_2 + q(u_1, u_2, u_3, u_4) - m_2 \ddot{u}_g. \end{aligned}$$

Let us denote

$$u = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{pmatrix}, \quad p(u) = \begin{pmatrix} u_1 \\ u_2 \\ m_1 u_3 \\ m_2 u_4 \end{pmatrix}, \quad \text{and} \quad f(t, u) = \begin{pmatrix} u_3 \\ u_4 \\ -r_1 u_3 - k_1 u_1 - q(u) - m_1 \ddot{u}_g \\ -r_2 u_4 - k_2 u_2 + q(u) - m_2 \ddot{u}_g \end{pmatrix}.$$

Together with the initial condition $u(0) = 0$, we obtain a system of nonsmooth ordinary differentiable equations (2.8).

We choose

$$m_1 = m_2 = 7.8, \ r_1 = 16.34, \ r_2 = 8.17, \ k_1 = 3421.5, \ k_2 = 855.4,$$
$$\alpha = 25000, \ \gamma = 3/2, \ d = 0.5, \ T = 10.$$

Table 4.1 reports numerical results of the use of a suite of 27 ground motion records with different peak ground accelerations (PGA) from 12 different earthquakes and difference parameters $\beta$. The values

$$e(h_1) = ||U^{h_1}(T) - U^{h_4}(T)||_2 \quad \text{and} \quad e(h_3) = ||U^{h_3}(T) - U^{h_4}(T)||_2$$

show the difference of numerical solutions at the end point, for approximate solutions $U^{h_l}$, with different step sizes

$$h_1 = 2 \times 10^{-3}, \quad h_3 = 2 \times 10^{-4}, \quad h_4 = 10^{-4}.$$

TABLE 4.1
*Numerical results for a suite of 27 ground motion records.*

| | Earthquake | PGA | $\beta$ | Burrage | | Lobatto IIIA | |
|---|---|---|---|---|---|---|---|
| | | | | $e(h_1)$ | $e(h_3)$ | $e(h_1)$ | $e(h_3)$ |
| I | Northridge, 1994 | 0.11 | 113.77 | 0.0430 | 0.0010 | 0.0676 | 0.0010 |
| | Imperial valley, 1979 | 0.10 | 118.90 | 0.0183 | 0.0015 | 0.0564 | 0.0004 |
| | San Fernando, 1971 | 0.09 | 104.88 | 0.0286 | 0.0032 | 0.1902 | 0.0017 |
| | Loma Prieta, 1989 | 0.19 | 119.17 | 0.0380 | 0.0030 | 0.2140 | 0.0012 |
| | Morgan Hill, 1984 | 0.19 | 103.95 | 0.0337 | 0.0027 | 0.1407 | 0.0004 |
| | N.Palm Springs, 1986 | 0.21 | 126.02 | 0.0276 | 0.0030 | 0.1184 | 0.0003 |
| | Whittier Narrows, 1987 | 0.30 | 143.88 | 0.0104 | 0.0017 | 0.0229 | 0.0005 |
| | Landers, 1992 | 0.28 | 141.05 | 0.0210 | 0.0008 | 0.1645 | 0.0061 |
| | Morgan Hill, 1984 | 0.29 | 138.12 | 0.0341 | 0.0034 | 0.1951 | 0.0022 |
| II | Loma Prieta, 1989 | 0.37 | 121.05 | 0.0870 | 0.0077 | 0.1180 | 0.0074 |
| | Northridge, 1994 | 0.42 | 130.49 | 0.1980 | 0.0066 | 0.0871 | 0.0088 |
| | Cape Mendocino, 1992 | 0.39 | 128.33 | 0.0183 | 0.0032 | 0.3519 | 0.0032 |
| | Northridge, 1994 | 0.51 | 161.44 | 0.1408 | 0.0128 | 0.3350 | 0.0311 |
| | Loma Prieta, 1989 | 0.48 | 150.10 | 0.0485 | 0.0053 | 0.5086 | 0.0049 |
| | Northridge, 1994 | 0.48 | 163.14 | 0.1893 | 0.0111 | 0.4833 | 0.0104 |
| | Loma Prieta, 1989 | 0.51 | 146.00 | 0.0883 | 0.0036 | 0.5095 | 0.0027 |
| | N.Palm Springs, 1986 | 0.59 | 146.00 | 0.0392 | 0.0028 | 0.1813 | 0.0005 |
| | Cape Mendocino, 1992 | 0.59 | 166.23 | 0.0683 | 0.0022 | 0.4249 | 0.0013 |
| III | Loma Prieta, 1989 | 0.61 | 153.37 | 0.3305 | 0.0194 | 0.3815 | 0.0106 |
| | Coalinga, 1983 | 0.60 | 145.43 | 0.0570 | 0.0070 | 0.3204 | 0.0084 |
| | Northridge, 1994 | 0.59 | 160.51 | 0.3654 | 0.0380 | 0.7291 | 0.0387 |
| | Cape Mendocino, 1999 | 0.66 | 157.76 | 0.0726 | 0.0048 | 0.8559 | 0.0059 |
| | Duzce, 1983 | 0.82 | 149.00 | 0.0525 | 0.0026 | 0.1905 | 0.0032 |
| | Coalinga, 1983 | 0.84 | 154.89 | 0.0127 | 0.0016 | 0.1759 | 0.0007 |
| | Northridge, 1994 | 0.84 | 140.00 | 0.0021 | 0.0002 | 0.3798 | 0.0005 |
| | Superstition Hills, 1987 | 0.89 | 141.05 | 0.0941 | 0.0038 | 0.3936 | 0.0084 |
| | Cape Mendocino, 1992 | 1.04 | 142.30 | 0.0469 | 0.0033 | 0.0721 | 0.0023 |
| CPU time (sec) for the 27 records | | | | 514 | 33532 | 573 | 34017 |
| I: Low intensity | | II: Moderate intensity | | | | III: High intensity | |

Figure 4.3 reports numerical results of the 3-stage Lobatto IIIA coefficient with parameter $\beta = 112$ in a simulation of the 1995 Great Hanshin earthquake, commonly referred to as the Kobe earthquake. The exact solution of this problem is unknown. In the first subplot, we show the difference of simulation solutions at end point $T = 20$:

$$\|U^{h_l}(T) - U^{h_5}(T)\|_2, \quad l = 0, \ldots, 4,$$

where

$$h_l = \frac{20}{N_l}, \quad N_l = 1000 \times 2^l, \quad l = 0, \ldots, 5.$$

We use a logarithmic scale for the step sizes and the difference. The 3-stage Lobatto IIIA method seems to maintain its fourth order of convergence for this problem. The remainder of Figure 4.3 uses step size $h = 0.002$, which shows the displacement $u_1$, $u_2$ and the force $q$ for the first 20 seconds of the earthquake. Moreover, we enlarge the plot around the maximum force in the first 10 seconds and the following 10 seconds to show what the pounding force looks like.

All of the ground motion records used in the numerical experiments were taken from the PEER Strong Motion Database [15]. The numerical experiments were performed by using MATLAB 7.0 on a Dell PC with 2 MB memory and 800 MHz.

**5. Conclusion.** Systems of slantly differentiable ODEs arise from many applications in earthquake engineering and structural dynamics. We proposed IRK methods
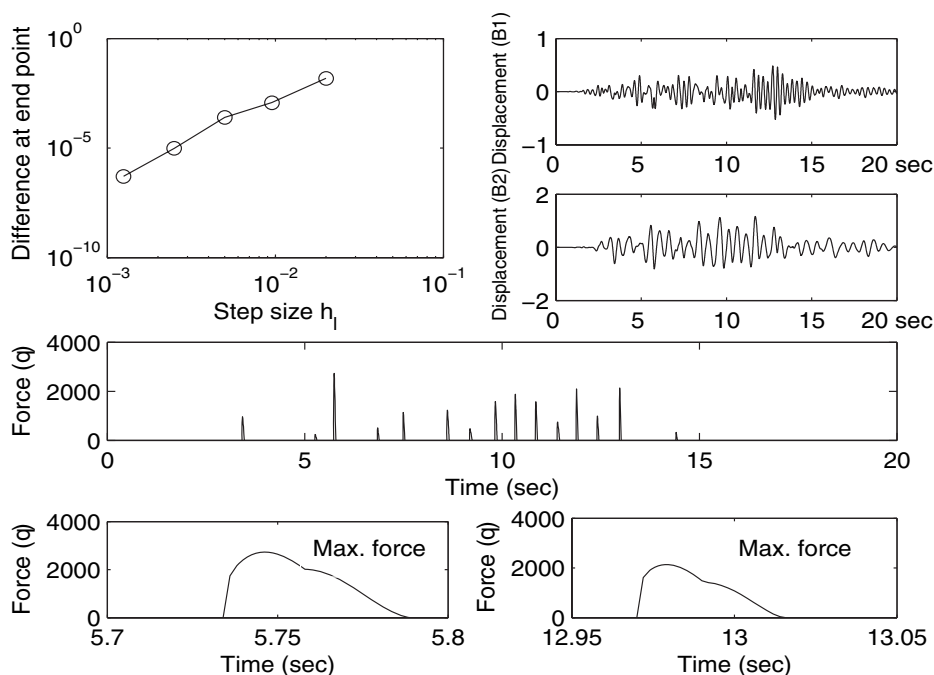
Fig. 4.3. *Difference of simulation solutions. Displacement $u_1$, $u_2$, force $q$, and maximum force during the first* 10 *seconds and the following* 10 *seconds.*

with the slanting Newton method to solve the system of nonsmooth ODEs. We studied the slanting differentiability of the nonsmooth functions involved in the ODEs and proved the superlinear convergence of the slanting Newton method for solving the system of nonsmooth equations in the IRK method. Moreover, we developed a code for nonsmooth ODEs based on Jay's code [8] and the analysis in section 2. Numerical experiments show that the code is efficient. In section 4, we reported some numerical results for the collapse of the Tacoma Narrows suspension bridge and simulating 13 different earthquakes.

**Acknowledgments.** The authors are grateful to Professor Robert Russell and two anonymous referees for their many helpful comments.

## REFERENCES

[1] K. BURRAGE, *Efficiently implementable algebraically stable Runge–Kutta methods*, SIAM J. Numer. Anal., 19 (1982), pp. 245–258.
[2] X. CHEN, Z. NASHED, AND L. QI, *Smoothing methods and semismooth methods for nondifferentiable operator equations*, SIAM J. Numer. Anal., 38 (2000), pp. 1200–1216.
[3] F.H. CLARKE, *Optimization and Nonsmooth Analysis*, Wiley, New York, 1983.
[4] J.E. DENNIS AND R.B. SCHNABEL, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, SIAM, Philadelphia, 1996.
[5] L. FERRACINA AND M.N. SPLIKER, *An extension and analysis of the Shu-Osher representation of Runge-Kutta methods*, Math. Comp., 74 (2004), pp. 201–219.
[6] M. HINTERMÜLLER, K. ITO, AND K. KUNISCH, *The primal-dual active set strategy as a semismooth Newton method*, SIAM J. Optim., 13 (2003), pp. 865–888.
[7] R. JANKOWSKI, *Non-linear viscoelastic modelling of earthquake-induced structural pounding*, Earthquake Engng. Struct. Dyn., 34 (2005), pp. 595–611.

[8]  L.O. Jay, *Inexact simplified Newton iterations for implicit Runge-Kutta methods*, SIAM J.
     Numer. Anal., 38 (2000), pp. 1369–1388.
[9]  R. Kress, *Numerical Analysis*, Springer-Verlag, New York, 1998.
[10] A.C. Lazer and P.J. McKenna, *Large-amplitude periodic oscillations in suspension bridges:
     Some new connections with nonlinear analysis*, SIAM Rev., 32 (1990), pp. 537–578.
[11] S. Muthukumar and R. DesRoches, *A Hertz contact model with non-linear damping for
     pounding simulation*, Earthquake Engng. Struct. Dyn., 35 (2006), pp. 811–828.
[12] R.T. Rockafellar, *A property of piecewise smooth functions*, Comput. Optim. Appl., 25
     (2003), pp. 247–250.
[13] B. Tasic, *Numerical Methods for Solving ODE Flow*, Eindhoven University Press, The Nether-
     lands, 2004.
[14] D.G. Zill, *Differential Equations with Modeling Applications*, 7th ed., Brooks/Cole Thompson
     Learning, Pacific Grove, CA, 2001.
[15] http://peer.berkeley.edu/smcat/.