NUMERICAL TIME INTEGRATION OF HIGHER ORDER DYNAMICAL SYSTEMS WITH STATE CONSTRAINTS

Vincent Acary
BIPOP Project
INRIA Rhône-Alpes
France
Vincent.Acary@inrialpes.fr

Bernard.Brogliato
BIPOP Project
INRIA Rhône-Alpes
France
Bernard.Brogliato@inrialpes.fr

Abstract

This work addresses the problem of the numerical time integration of the Moreau's Sweeping Process (MSP) and particularly one of its extensions: The higher order formulation. The goal of this paper is to give some hints and tricks on the ability of the proposed associated time-stepping scheme to deal with non smooth evolutions without explicit event-handling procedures. We present also new algorithms for solving higher order non smooth dynamical systems.

Key words

Moreau's Sweeping process, non smooth dynamics, unilateral constraints, higher order.

1 Introduction

1.1 General position of the problem

In this paper, we are interested in the modeling and the numerical time integration of a dynamical system,

$$\dot{x} = f(x,t), x \in \mathbb{R}^n, t \in [0,T], \tag{1}$$

subjected to a set of constraints on its state[1]:

$$w = h(x) = [h_\alpha(x), \alpha = 1 \dots m]^T \geq 0. \tag{2}$$

The constraints (2) are usually enforced by an external input, let's say, a multiplier $\lambda \in \mathbb{R}^m$, through an input function g such as

$$g : \lambda \in \mathbb{R}^m \mapsto g(\lambda) \in \mathbb{R}^n \tag{3}$$
$$\dot{x} = f(x,t) + g(\lambda) \tag{4}$$

Similarly to the differential index in differential algebraic equations, a fundamental feature of such systems is the relative degree between the output w and the input λ. The relative degree determines the nature of the mathematical solution of the system and the way how to solve them numerically. We will see later that the relative degree plays an important role in the non smoothness of the solution.

Finally, in order to complete the system, additional modeling informations are needed. Particularly, two laws are of utmost importance

a) A generalized equation (Robinson, 1979) between the output w and the multiplier λ, denoted by the following inclusion:

$$0 \in F(w,\lambda) + Q(w,\lambda) \tag{5}$$

where $F : \mathbb{R}^{m\times m} \mapsto \mathbb{R}^{m\times m}$ is assumed to be continuously differentiable and $Q : \mathbb{R}^{m\times m} \rightsquigarrow \mathbb{R}^{m\times m}$ is a multivalued mapping with a closed graph. In the case of relative degree less or equal to one, a complementarity condition is usually introduced for (5):

$$0 \leq w \perp \lambda \geq 0 \Leftrightarrow 0 \in \lambda + \partial\psi_{\mathbb{R}^+}(w) \tag{6}$$

where ψ_K the indicator function of the set K and the symbol ∂ denotes the subdifferential in the sense of the Convex Analysis. For higher order systems or more complicated systems, the simple complementarity is unsuitable, not to say meaningless.

b) A reinitialization mapping or an impact law defining the state of the system after a non smooth event:

$$x(t^+) = \mathcal{F}(x(t^-), t) \tag{7}$$

We will see later that is is also possible to provide us with a compact formulation of the equations (5) and (7) into a single inclusion.

[1] The inequality is to be understood component-wise.

1.2 The case of Lagrangian systems with unilateral constraints

To shed more light on this type of systems, we can consider for instance the well known case of the Lagrangian mechanical systems with unilateral constraints, which model the dynamics of finite dimensional mechanical systems with contact. Let us consider a p-dimensional Lagrangian system in a configuration manifold $\mathcal{M}$, parameterized by a set of p generalized coordinates denoted by $q \in \mathcal{M}$, $M(q)$ the mass matrix and $F(q, \dot{q}, t)$ is the set of forces acting upon the system. Usually, the unilateral constraints are written in the coordinates as:

$$h_\alpha(q(t)) \geq 0, \quad \alpha = 1 \dots \nu, \tag{8}$$

defining an admissible set for the system:

$$\Phi(t) = \{z(t) \in \mathcal{M}, h_\alpha(z(t)) \geq 0, \alpha = 1 \dots \nu\}. \tag{9}$$

With sufficient regularity in time, the Lagrange equations are:

$$M(q)\ddot{q} + F(q, \dot{q}, t) = \textstyle\sum_{\alpha=1}^{\nu} \nabla h_\alpha(q)\lambda_\alpha, \tag{10}$$

where λ_α is the set of the Lagrange multipliers associated with the constraints $h_\alpha(q(t))$ through a complementarity condition:

$$0 \leq h_\alpha(q) \perp \lambda_\alpha \geq 0, \tag{11}$$

In this case, the output function h is defined in terms of the coordinates q and the output function g is related to this function by:

$$g(\lambda) = \begin{bmatrix} 0 \\ \sum_{\alpha=1}^{\nu} \nabla h_\alpha(q)\lambda_\alpha \end{bmatrix} \tag{12}$$

This relation between the input and output function combined with the dynamics is crucial in order to define the mathematical nature of the solutions and to fix the relative degree vector of the Lagrangian mechanical systems to $(2, \dots, 2) \in \mathbb{R}^{1\times\nu}$. The leading Markov parameter of the operator $\lambda \mapsto w$ is precisely equal to the Delassus's matrix, $\nabla^T h(q) M^{-1} \nabla h(q) \in \mathbb{R}^{\nu\times\nu}$, which is full-rank if and only if the constraints are independent.

1.3 The Moreau's sweeping process (MSP)

The MSP is a formulation and a mathematical framework for non linear dynamical systems subjected to unilateral constraints and possibly impacts, initiated and developed by J.J. Moreau. This framework has been applied successfully to several fields in nonlinear mechanics (unilateral contact and friction, plasticity, fluids mechanics, etc ...) and extensively studied from the mathematical point of view (well-posedness, existence and uniqueness of solutions). This formulation is composed of two major parts: a reformulation of the non smooth dynamics in a Measure Differential Inclusion (MDI) framework and if needed the formulation of a consistent reinitialization mapping.

In (Moreau, 1971; Moreau, 1972; Moreau, 1977), the problem of the existence and the uniqueness of the following first order sweeping process is addressed:

$$-du \in \partial\psi_{K(t)}(u(t)) \tag{13}$$

where $K(t)$ is a moving convex set. The measure du is a differential measure (Stieltjes) associated with the state u considered as a right continuous function of local bounded variations (RCLBV).

In the case of the second order dynamics, for instance, for Lagrangian mechanical systems, Schatzman (Schatzman, 1978) and Moreau (Moreau, 1983) have reformulated the equation of motion (10) in terms of measure differential equations:

$$M(q)dv + F(q, v, t)dt = \sum_{\alpha=1}^{\nu} \nabla h_\alpha(q)\lambda_\alpha, \tag{14}$$

where dt is the Lebesgue measure, dv is a differential measure associated with the velocity $v = \dot{q}^+(t)$ considered as a RCLBV function and finally, λ is henceforth a measure. To complete this measure differential equation, Schatzman introduced a purely elastic impact law and Moreau proposed a compact formulation of an inelastic impact law as a measure inclusion:

$$-H^T(q(t))\lambda \in \partial\psi_{V(q(t))}(v(t^+) + ev(t^-)) \tag{15}$$

where $V(q)$ is the tangent cone to $\Phi(t)$ at q, e is the coefficient of restitution and $H^T(q)\lambda = \sum_{\alpha=1}^{\nu} \nabla h_\alpha(q)\lambda_\alpha$.

The definition of the inclusion of a measure into a cone may be found in (Monteiro Marques, 1993, p 76). Roughly speaking, this inclusion ensures that if the evolution is continuous, the absolutely continuous part of the measure which can be identified to a function, belongs to the cone. At an instant of jumps, the amplitude of the jump must also belong to the cone. Finally, we obtain a MDI, the so-called Sweeping process:

$$M(q)dv + F(q, v, t)dt \in -\partial\psi_{V(q)}(v(t^+) + ev(t^-)) \tag{16}$$

In Section 2, we give some hints about the time integration of the MSP.

1.4 Recent works

Recently, several extensions have been proposed, particularly in order to deal with higher order systems, and more precisely, to deal with systems with arbitrary relative degree.

In (Heemels *et al.*, 2000), the authors study the so-called Linear Complementarity Problem (LCS) defined by:

$$\begin{cases} \dot{x} = Ax + B\lambda \\ w = Cx + D\lambda \\ 0 \leq w \perp \lambda \geq 0 \end{cases} \tag{17}$$

and try to give a meaning to the solution of such systems adding a huge amount of assumptions. It is clear that such type of systems cannot provide us with a consistent modeling without a set of strong conditions and may be meaningless in a general case. For instance, solutions of such systems may be distributions of higher degree. Giving conditions on the sign of a distribution is meaningless. The application of standard numerical techniques on this ill-posed system leads to a great number of inconsistenies (Camlibel *et al.*, 2002). In (Heemels *et al.*, 2000), a pseudo-algorithm for constructing a solution is given and is based on a fully hybrid approach in which the LCS is considered a multimodal linear system.

In (Acary *et al.*, 2005; Acary and Brogliato, 2003), a completely different approach is chosen, which belongs to the theory of non smooth analysis. Starting from a linear dynamical system and a linear output subjected to unilateral constraints:

$$\begin{cases} \dot{x} = Ax + B\lambda \\ w = Cx + D\lambda \geq 0 \end{cases} \quad , \tag{18}$$

we give a meaning to the solution for all $t \geq 0$by some existence and uniqueness proofs; we reformulate the system in a convenient way in order to obtain a self-contained formulation and finally we propose a consistent numerical scheme. This formulation will be briefly sum up in the Section 3.

1.5 Outline of the paper.

The paper is organised as follows. In the Section 2, we recall briefly the basics on the MSP and its time integration. This seminal work concerns the case of a dynamical system submitted to state constraints with a relative degree less or equal than 2. In the section 3, we present an extension of the sweeping process for dynamical systems of relative degree greater than 2. Particularly, the method of numerical time-integration is detailed in the Sections 3.1 and 3.2. A comparison is made in the Section 3.4 with an backward Euler scheme used in (Camlibel *et al.*, 2002) is made in order to avoid confusions. In the Section 3.5, we give a flavour of the order of the scheme on a particular example.

2 The Numerical time integration of the MSP

In this section, we give some details about the seminal work of Moreau on the numerical time integration of the sweeping process. These details provide tools which lead us to the design of the extended algorithm presented in Section 3.

Let us consider the following notations. We denote by $0 = t_0 < t_1 < \ldots < t_k < t_N = T$ a finite partition (or a subdivision) of the time interval $[0, T], T > 0$. The integer N stands for the number of time intervals in the subdivision. The length of a time step is denoted by $h = t_{k+1} - t_k$. The approximation of $f(t_k)$, the value of a real function f at the time t_k, is denoted by f_k .

2.1 First order sweeping process

Under suitable hypothesis on the multivalued function $t \mapsto K(t)$, numerous convergence and consistency results (Monteiro Marques, 1993; Kunze and Monteiro Marques, 2000) were given for this algorithm together with well-posedness results for the sweeping process (13). Using the so-called "Catching-up algorithm" defined as (Moreau, 1977):

$$-(u(t^+_{k+1}) - u(t^+_k)) \in \partial\psi_{K(t^+_{k+1})}(u(t^+_{k+1})) \tag{19}$$

By elementary convex analysis, and using the convention that $u_{k+1} = u(t^+_{k+1})$ this is equivalent to:

$$u_{k+1} = prox(K_{k+1}, u_k) \tag{20}$$

Contrary to the standard backward Euler scheme with which it might be confused, the catching-up algorithm is based on the evaluation of the measure du on the interval $(t_k, t_{k+1}]$, i.e. $du((t_k, t_{k+1}]) = u(t^+_{k+1}) - u(t^+_k)$. The Euler scheme is based on the approximation of $\dot{u}(t)$ which is not defined in a classical sense everywhere for our case. When the time step vanishes, the approximation of the measure du tends to a finite value corresponding to the jump of u. This remark is crucial for the consistency of the scheme. Particularly, this fact ensures that we handle only finite values. Furthermore using higher order numerical schemes is at best useless, more often it is dangerous. Basically, a general way to obtain a finite difference-type scheme of order n is to write a Taylor expansion of order n or higher. Such a scheme is meant to approximate the n-th derivative of the discretized function. If the solution we are dealing with is obviously not differentiable, what is the meaning of using a scheme with order $n \geq 2$ which tries to approximate derivatives which do not exist ? In summary, standard higher-order numerical schemes are inadequate for time-stepping discretization of dynamical systems subjected to unilateral constraints.

2.2 Second order MSP: Overview of the Non Smooth Contact Dynamics (NSCD) method

The NSCD method is the numerical discretization of the second order MSP introduced by Moreau in

(Moreau, 1988). This numerical scheme is devoted to the numerical integration of Lagrangian mechanical systems (16). In this section we briefly present the major features of this numerical method. For details, see (Moreau, 1999; Moreau, 2003; Jean, 1999).

The NSCD method performs the numerical time integration of the MDI (16) on an interval $(t_k, t_{k+1}]$. Using the notation

$$v_{k+1} \approx v(t^+_{k+1}); \quad \mu_{k+1} \approx \lambda((t_k, t_{k+1}]) \tag{21}$$

it may be written down as follows:

$$\begin{cases} -M(v_{k+1} - v_k) - h((1-\theta)F(v_{k+1}, q_{k+1}, t_{k+1}) \\ \qquad -\theta F(v_k, q_k, t_k)) = R_{k+1} \\ \\ R_{k+1} = H^T(\tilde{q}_{k+1})\mu_{\alpha\, k+1} \in \partial\psi_{T_\Phi(\tilde{q}_{k+1})}(v_{k+1}) \\ \\ q_{k+1} = q_k + h((1-\theta)v_{k+1} + \theta v_k), \quad \theta \in [\frac{1}{2}, 1] \\ \\ \tilde{q}_{k+1} = q_k + h v_k \end{cases} \tag{22}$$

The inclusion in (22) can be stated equivalently as a complementarity problem:

$$\text{if } h(\tilde{q}_{k+1}) \leq 0 \text{ then } 0 \leq H(\tilde{q}_{k+1})v_{k+1} \perp \mu_{k+1} \geq 0 \tag{23}$$

The value $\tilde{q}_{k+1}$ is a prediction of the position which allows the computation of the tangent cone T_Φ. A $\theta-$method is used for the integration of the position assuming that q is absolutely continuous. The same approximation is made with the term $F(v(t^+), q(t), t)$.

2.3 Comments

From a numerical point of view, two major lessons can be learned form this work:

1. First, the various terms manipulated by the numerical algorithm are of finite value. The use of differential measures of the time interval $(t_k, t_{k+1}]$,. i.e., $dv((t_k, t_{k+1}]) = v(t^+_{k+1}) - v(t^+_k)$ and $= \mu_{k+1} = d\lambda((t_k, t_{k+1}])$ is fundamental and allows a rigorous treatment of the non smooth evolutions. When the time-step h vanishes, it allows one to deal with finite jumps. When the evolution is smooth, the scheme is equivalent to a backward Euler scheme. We can remark that nowhere an approximation of the acceleration $\ddot{q}$ is used.
2. Secondly, the inclusion in terms of velocity allows to treat the displacement as a secondary variable. A viability lemma ensures that the constraints on q will be respected at convergence. We will see further that this formulation gives more stability to the scheme.

These remarks on the contact dynamics method might be viewed only as some numerical tricks. In fact, the mathematical study of the second order MDI by Moreau provides a sound mathematical ground to this numerical scheme. It is noteworthy that convergence results have been proved for such time-stepping schemes (Monteiro Marques, 1993; Stewart, 1998).

2.4 Example on the bouncing ball

The algorithm (22) provides a numerical scheme with very nice properties. The reader may convince his/herself of this by studying the simple bouncing ball on a rigid plane subject to gravity and with elastic restitution. The proposed time-discretization of the motion of this ball is

$$\begin{aligned} -m(v(t_{k+1}) + v(t_k)) - hmg \\ \in \partial\psi_{V(\tilde{q}_{k+1})}(v(t_{k+1}) + ev(t_k)) \end{aligned} \tag{24}$$

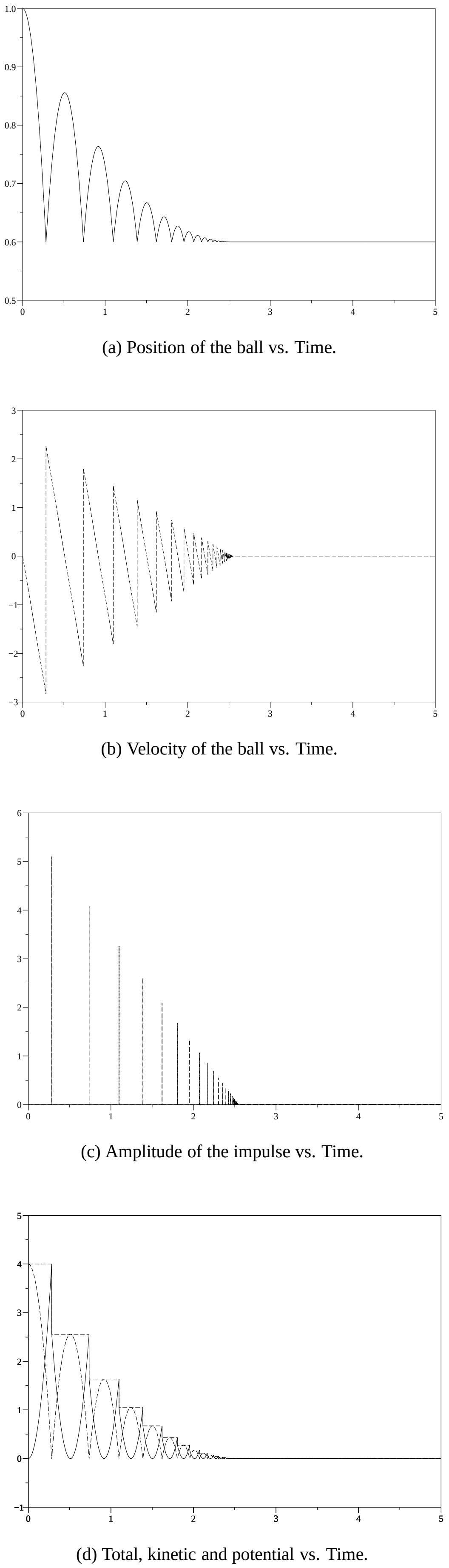
where m is the mass of the ball, e the coefficient of restitution and g the gravity. One notes that the dissipativity property shows through the power of h in the term hg which has the dimension of an impulse (there is no h pre-multiplying the right-hand-side since this is a cone).

If $q_0 > 0$ then the ball falls down until penetration is detected at step k^* (i.e. $q_{k^*-1} > 0$ while $q_{k^*} < 0$). Then the velocity is reversed, i.e. $v_{k^*+1} = -ev_{k^*}$ while $q_{k^*+1} = q_{k^*}$. Thus the system is re-initialized at each impact, with the same velocity and at the same position. There are no errors introduced by the numerical scheme and one can simulate several billions of such cycles without energy gain nor losses. Clearly this is not possible with an event-driven scheme, even if a very accurate detection procedure is used. The unavoidable penetration is not a major issue, since anyway the discretized system cannot be exactly at $q = 0$. What is crucial is that the penetration goes to zero when $h \to 0$. In the case $e \in (0, 1)$, an infinity of rebounds in finite time occur in the continuous time model. This Zeno behaviour is correctly integrated as depicted on the Figure 1.

3 The Higher Order MSP (HOMSP)

Starting from (18) and denoting the relative degree between w and y by r, we perform the following state-space transformation $z = Wx$, W square full-rank of order n, and:$z^T = (w, \dot{w}, \ddot{w}, ..., w^{(r-1)}, \xi^T) = (\bar{z}^T, \xi^T)$, $\xi \in \mathrm{I\!R}^{n-r}$. The so-called zero-dynamical (ZD) canonical form is obtained:

$$\begin{cases} \dot{z}_i(t) = z_{i+1}(t) \ \ (t \geq 0), i = 1 \ldots r-1 \\ \dot{z}_r(t) = CA^rW^{-1}z(t) + CA^{r-1}B\lambda(t) \ \ (t \geq 0) \\ \dot{\xi}(t) = A_\xi \xi(t) + B_\xi z_1(t) \ \ (t \geq 0) \\ w(t) = z_1(t) \geq 0 \ \ (t \geq 0) \end{cases} \tag{25}$$



(a) Position of the ball vs. Time.

(b) Velocity of the ball vs. Time.

(c) Amplitude of the impulse vs. Time.

(d) Total, kinetic and potential vs. Time.

Figure 1. Bouncing Ball on a rigid plane. $e = 0.8, g = 10m.s^{-2}, m = 1kg, h = 5.10^{-3}s$

Here $A_\xi \in \mathbb{R}^{n-r\times n-r}$ and $B_\xi \in \mathbb{R}^{n-r\times 1}$. The transition matrix of the ZD form is

$$WAW^{-1} = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 & 0_{n-r} \\ 0 & 0 & 1 & \dots & 0 & 0_{n-r} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \dots & 1 & 0_{n-r} \\ d_1 & d_2 & d_3 & \dots & d_r & d_\xi^T \\ B_\xi & 0^{n-r} & 0^{n-r} & \dots & 0^{n-r} & A_\xi \end{pmatrix} \tag{26}$$

where $(0^{n-r})^T = 0_{n-r} = (0, \dots, 0) \in \mathbb{R}^{1\times(n-r)}$.

The ZD form (25) outlines the role of the relative degree. For instance, if the variable z_1 jumps, the $r-1$-order derivative z_r is a Dirac distribution of degree r. If the variable $z_1(t) = 0$, and its derivatives $z_2(t), z_3(t)$ are negative, we need to apply a jump to this derivatives to respect the constraint. This action is performed by a set of input multipliers. The evolution may be clearly non smooth. Therefore, following the work of Moreau, we propose in (Acary *et al.*, 2005) to state a higher order measure differential equation as:

$$\begin{cases} dz_i - z_{i+1}\, dt = d\nu_i, i = 1 \dots r-1 \\ dz_r - CA^r W^{-1} z\, dt = CA^{r-1} B d\nu_r \\ \dot{\xi} - A_\xi \xi - B_\xi z_1 = 0 \end{cases} \tag{27}$$

where dz_i are the Stieltjes measure associated with the RCLBV function $z_i(t)$ and the measure $d\nu_i$ is the multiplier which enforces the unilateral constraints. Finally, we need to add a reinitialization mapping and extending the MSP we write:

$$d\nu_i \in -\partial\psi_{T^{i-1}_{\mathbb{R}^+}(Z_{i-1})}(z_i(t^+)), i = 1 \dots r \tag{28}$$

where $Z_{i-1} = [z_1(t^-), \dots z_{i-1}(t^-)]$. The cone $T^{i-1}_{\mathbb{R}^+}(Z_{i-1})$ is defined by induction such that:

$$\begin{aligned} &T^0\ (z_1) = \mathbb{R}^+ \\ &T^1\ _{\mathbb{R}^+}(Z_1) = T_{\mathbb{R}^+}(z_1) \\ &T^2\ _{\mathbb{R}^+}(Z_2) = T_{T_{\mathbb{R}^+}(z_1)}(z_2) \\ &\quad \vdots \\ &T^i\ _{\mathbb{R}^+}(Z_i) = T_{T^{i-1}_{\mathbb{R}^+}(Z_{i-1})}(z_i) \end{aligned} \tag{29}$$

where $T_C(z)$ denotes the tangent cone to the convex set C taken at z. For more details on the properties of this formulation, we refer to (Acary *et al.*, 2005).

3.1 Principle of the Numerical Time integration

Let us start with the generic equation for $i = 1 \dots r-1$ of the measure differential formalism (27). The evaluation of this MDI on the time interval $(t_k, t_{k+1}]$ yields:

$$dz_i((t_k, t_{k+1}]) - \int_{(t_k,t_{k+1}]} z_{i+1}(\tau)d\tau = d\nu_i((t_k, t_{k+1}]). \tag{30}$$

The values of the measures $dz_i((t_k, t_{k+1}])$ and $\mu_{i,k+1} \triangleq d\nu_i((t_k, t_{k+1}])$ are kept as primary variables and this fact is crucial for the consistency of the method while a non-smooth evolution. The integral term is approximated thanks to:

$$\int_{(t_k,t_{k+1}]} z_{i+1}(\tau)d\tau \approx hz_{i+1}(t^+_{k+1}) = hz_{i+1,k+1}, \tag{31}$$

and then we obtain:

$$z_{i,k+1} - z_{i,k} - hz_{i+1,k+1} = \mu_{i,k+1}. \tag{32}$$

The approximation of the inclusion (28) is performed in the following way:

$$\mu_{i,k+1} \in -\partial\psi_{T^{i-1}_\Phi(Z_{i-1,k})}(z_{i,k+1}). \tag{33}$$

Finally, the time integration of a generic equation of the MDI for $i = 1 \ldots r-1$ in (27) is given by:

$$\begin{cases} z_{i,k+1} - z_{i,k} - hz_{i+1,k+1} = \mu_{i,k+1} \\ \mu_{i,k+1} \in -\partial\psi_{T^{i-1}_\Phi(Z_{i-1,k})}(z_{i,k+1}) \end{cases}. \tag{34}$$

The second equation in (27) is discretized as:

$$z_{r,k+1} - z_{r,k} - hCA^rW^{-1}z_{k+1} = CA^{r-1}B\ \mu_{r,k+1}$$

$$\mu_{r,k+1} \in -\partial\psi_{T^{r-1}_\Phi(z_{1,k},\ldots,z_{r-1,k})}(z_{r,k+1}). \tag{35}$$

For the zero dynamics defined in the last equation of (27), we use for the sake of simplicity[2] an Euler Backward scheme:

$$\xi_{k+1} - \xi_k - hA_\xi\xi_{k+1} - hB_\xi z_{1,k+1} = 0. \tag{36}$$

The numerical time integration of the HOMSP is defined as the inclusion in (34), (35) and (36).

3.2 General algorithm

In this section, we provide a short and possibly pedagogical overview of the implementation of the numerical algorithm of the HOMSP. The state vector is denoted by

$$z_{k+1} = [z_{1,k+1}, \ldots, z_{r,k+1}, \xi^T_{k+1}]^T = [\bar{z}^T_{k+1}, \xi^T_{k+1}]^T$$

To be more explicit in the computation of the state vector, we introduce the matrix $P \in \mathbb{R}^{r\times n}$ such that

$$\bar{z}_{k+1} = Pz_{k+1}. \tag{37}$$

Matrix formulation of the ZD form in view of numerical integration. We perform the numerical integration trough the ZD canonical form. Given $W \in \mathbb{R}^{n\times n}$, the linear transformation of the state space, we introduce the following matrix notations:

$$[I - h\bar{A}]z_{k+1} = z_k + \bar{B}\mu_{k+1} \tag{38}$$

where the matrices $\bar{A} = WAW^{-1} \in \mathbb{R}^{n\times n}$ is defined as (26) and $\bar{B} \in \mathbb{R}^{n\times r}$ is defined as follows:

$$\bar{B} = \left[\begin{array}{ccccc} 1 & 0 & \ldots & \ldots & 0 \\ 0 & 1 & \ddots & (0) & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & 1 & 0 \\ 0 & \ldots & \ldots & 0 & M \\ \hline & & & & \\ & & 0 & & \\ & & & & \end{array}\right] \text{ with } M = CA^{r-1}B \tag{39}$$

Expression of the inclusions (33) **in terms of nested complementarity problems.** Let us consider the following inclusion:

$$\mu_{1,k+1} \in -\partial\psi_\Phi(z_{1,k+1}).$$

Let Φ be $\mathbb{R}^+$ in the sequel. This inclusion may be stated equivalently as a complementarity problem[3]:

$$0 \le z_{1,k+1} \perp \mu_{1,k+1} \ge 0$$

If $r > 1$ then we must handle the second inclusion:

$$\mu_{2,k+1} \in -\partial\psi_{T^1_\Phi(z_{1,k})}(z_{2,k+1})$$

which can be reformulated in terms of a complementarity problem:

$$\text{If } z_{1,k} \le 0, \text{ then } 0 \le z_{2,k+1} \perp \mu_{2,k+1} \ge 0$$

In this way, for $r > 2$, we get the following complementarity problem:

$$\text{If } z_{1,k} \le 0 \text{ and } z_{2,k} \le 0, \text{ then } 0 \le z_{3,k+1} \perp \mu_{3,k+1} \ge 0$$

[2] Depending on the regularity of z_1, a higher order scheme might be used for the time-integration of the zero dynamics.

[3] In a more general setting, a cone complementarity problem has to be written $K \ni z_{1,k+1} \perp -\mu_{1,k+1} \in K^\star$, with $K^\star$ the dual cone of K.

In the general case, we search the integer $1 \leq r^\star \leq r$ satisfying the following condition:

$$r^\star = \begin{cases} 1, \text{ if } z_{1,k} > 0 \\ 1 + \max\{j \leq r-1 : \quad z_{i,k} \leq 0, \forall i \leq j\} \end{cases}$$

Then we obtain the following set of nested complementarity problems:

$$0 \leq z_{i,k+1} \perp \mu_{i,k+1} \geq 0, \quad 1 \leq i \leq r^\star \tag{40}$$

We define the vectors collecting the state and the multiplier for the "active" constraints by:

$$\begin{aligned} z^\star_{k+1} &= [z_{1,k+1}, \ldots, z_{r^\star,k+1}]^T \\ \mu^\star_{k+1} &= [\mu_{1,k+1}, \ldots, \mu_{r^\star,k+1}]^T. \end{aligned} \tag{41}$$

and we introduce the matrix $R \in \mathbb{R}^{r^\star \times r}$ describing the relation between $z^\star_{k+1}$ and z_{k+1}:

$$z^\star_{k+1} = R\bar{z}_{k+1}. \tag{42}$$

Assuming that $\mu_{i,k+1} = 0, i > r^\star$, we get the relation between $\mu_{i,k+1}$ and $\mu^\star_{k+1}$:

$$\mu_{k+1} = R^T \mu^\star_{k+1}. \tag{43}$$

Formulation of the one-step LCP problem Assuming that $r^\star$ is computed at each step and that z_k is known, the following set of discretized equations has to be solved to advance from step k to step $k+1$:

$$\begin{cases} [I - h\bar{A}]z_{k+1} = z_k + \bar{B}\mu_{k+1} \\ \bar{z}_{k+1} = Pz_{k+1} \\ z^\star_{k+1} = R\bar{z}_{k+1} \\ \mu_{k+1} = R^T\mu^\star_{k+1} \\ 0 \leq \mu^\star_{k+1} \perp z^\star_{k+1} \geq 0. \end{cases} \tag{44}$$

This yields the following closed form for the one-step LCP problem:

$$\begin{cases} z^\star_{k+1} = RP[I - h\bar{A}]^{-1}z_k \\ \qquad\qquad +RP[I - h\bar{A}]^{-1}\bar{B}R^T\mu^\star_{k+1} \\ 0 \leq \mu^\star_{k+1} \perp z^\star_{k+1} \geq 0. \end{cases} \tag{45}$$

3.3 Properties of the scheme

Measures of time intervals as primary variables. As we have seen earlier, the measures of the time interval $(t_k, t_{k+1}]$, i.e. $dz((t_k, t_{k+1}])$ and $\mu_{i,k+1} \triangleq d\nu_i((t_k, t_{k+1}])$ are kept as primary variables. This fact ensures that the various terms manipulated by the numerical algorithm have finite values. The use of differential measures of the time interval $(t_k, t_{k+1}]$ allows a rigorous treatment of the nonsmooth evolutions. When the time-step h vanishes, it allows to deal with finite jumps. When the evolution is smooth, the scheme is equivalent to a backward Euler scheme. We can remark that nowhere a direct approximation of the density z'_t with respect to the Lebesgue measure is made. The use of a first order algorithm is not chosen as usual through the approximation of the integral term (31) but required by the evaluation of the differential measure.

Boundedness, local bounded variation and convergence In (Acary *et al.*, 2005), some results have been proved under a condition of dissipativity that pave the way to the convergence of the proposed time-stepping scheme. Particularly, the boundedness of the approximate solution has been proved. Starting from the approximation $z_{i,k}$ of the RCLBV function $z_i(t)$, we construct a family of step functions $z^N_i(t)$ such that:

$$z^N_i(t) = z_{i,k}, \forall t \in [t_k, t_{k+1}), i = 1 \ldots r \tag{46}$$

The local variation of these set of RCLBV functions, $z^N_i(t)$ is proved and thanks to the Helly's theorem, we are able to prove that there exists a subsequence that converges to some limit RCLBV function.

3.4 Comparison with a backward Euler scheme

For the first and the second order sweeping process, the time integration method is often confused with a standard backward Euler scheme. To highlight the difference with the numerical time integration of the HOMSP, we consider several examples of inconsistencies introduced in (Camlibel *et al.*, 2002). In that paper, the authors consider the most naive way of integrating a LCS (17) by applying directly a backward Euler scheme:

$$\begin{cases} \dfrac{x_{k+1} - x_k}{h} = Ax_{k+1} + B\lambda_{k+1} \\ w_{k+1} = Cx_{k+1} + D\lambda_{k+1} \\ 0 \leq \lambda_{k+1} \perp w_{k+1} \geq 0 \end{cases} \tag{47}$$

which can be reduced to the LCP $(y_{k+1}, \lambda_{k+1}) = LCP(M, b_{k+1})$ where

$$M = hC(I - hA)^{-1}B \tag{48}$$

$$b_{k+1} = C(I - hA)^{-1}x_k \tag{49}$$

They claim some consistency and convergence results. Shortly, under the assumption that D is nonnegative definite or that the triplet (A,B,C) is a minimal representation and (A,B,C,D) is passive, they exhibit that a subsequence of y_k, λ_k, x_k converges weakly to a solution of the LCS. Such assumptions imply that the relative degree r is less or equal to 1. In the case of the relative degree 0, the LCS is a equivalent to a standard system of ordinary equation with a Lipschitz-continuous vector field (Goeleven and Brogliato, 2004). The result of convergence is then the standard result of convergence for the Euler backward scheme. In the case of a relative degree equal to 1, these results corroborate the previous results of (Brezis, 1973).

As we said earlier, they also exhibit several examples for which the backward Euler scheme does not work at all. We will consider below two similar examples and comment the difference between the Backward Euler scheme and our approach.

Example 1 Let us consider a LCS (17) with the following matrix definition:

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}; B = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}; C = \begin{pmatrix} 1 & 0 & 0 \end{pmatrix}; D = 0 \quad (50)$$

The relative degree r of this LCS is equal to 2 ($D = 0, CB = 0, CAB \neq 0$). If we apply the time discretization given by (47), we can remark that:

$$\lim_{h \longrightarrow 0} \frac{1}{h} M = \lim_{h \longrightarrow 0} C(I - hA)^{-1} B = 0 \quad (51)$$

It is clear that if h is taken very small, which may be needed in many practical cases or for the convergence analysis of the scheme, then the LCP matrix for (17) has little chance to be well conditioned due to the fact that $CB = 0$.

If we consider the initial data $x_0 = (0, -1, 0)^T$, we obtain by a straightforward application of the scheme the following solution:

$$x_k = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}; \forall k \geq 1 \quad (52)$$

$$\lambda_1 = \frac{1}{h}; \quad \lambda_k = 0, \forall k \geq 2 \quad (53)$$

We can remark that the multiplier λ_1 which is the solution of the LCP at the first step, tends towards $+\infty$ when h vanishes. In this example, the state x seems to be well approximated but both the LCP matrix and the multiplier tends to inconsistent value when h vanishes. This inconsistency is just the result of an attempt to approximate the point value of a distribution which is non sense.

If we consider now the initial data $x_0 = (-1, -1, 0)^T$, we obtain the following numerical solution:

$$x_k = \begin{pmatrix} k \\ \frac{1}{h} \\ 0 \end{pmatrix}; \forall k \geq 1 \quad (54)$$

$$\lambda_1 = \frac{1}{h^2}; \quad \lambda_k = 0, \forall k \geq 2 \quad (55)$$

With a such initial data, the exact solution should be $x_k = 0, \forall k \geq 1$. We can see that there is an inconsistency in the result because the first component of the approximate state does not depend on the time-step. We can not expect that this approximation converges to any solution.

If we apply the proposed numerical scheme in (34)–(35)–(36), we obtain the following solution:

$$x_k = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}; \forall k \geq 1 \quad (56)$$

$$\mu_{1,1} = 1; \quad \mu_{2,1} = 1, \quad (57)$$

$$\mu_{1,k} = 0, \mu_{2,k} = 0 \forall k \geq 2 \quad (58)$$

which converges to the the time-continuous solution of the (28).

Example 2 Let us consider this second very simple example:

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}; B = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}; C = \begin{pmatrix} 1 & 0 & 0 \end{pmatrix}; D = 0 \quad (59)$$

In this case, the relative degree, r is equal to 3. The direct discretization of the system leads to the same problem as in the previous example even in the case where the initial data satisfies the constraints. Let us consider the case of $x_0 = (0, -1, 0)^T$, we obtain the following numerical solution:

$$x_k = \begin{pmatrix} \frac{k(k+1)}{2h} \\ k \\ \frac{1}{h} \end{pmatrix}; \forall k \geq 1 \quad (60)$$

$$\lambda_1 = \frac{1}{h^2}; \quad \lambda_k = 0, \forall k \geq 2 \quad (61)$$

Even in the case of some satisfying initial data with respect to constraints, we can not expect that this solution converges to the exact solution. The solution given by the proposed numerical (34)–(35)–(36) is

$$\begin{aligned} & x_k = \begin{pmatrix}0\\0\\0\end{pmatrix} ;\forall k \geq 1 \\ & \mu_{1,1} = 1; \quad \mu_{2,1} = 1, \quad \mu_{3,1} = 0 \\ & \mu_{i,k} = 0, \ ,\forall k \geq 2, i = 1 \dots 3 \end{aligned} \tag{62}$$

Interest of the ZD canonical form from the numerical point of view. Let us consider now the ZD form with $r = n$ for simplicity sake. A direct discretization with the Euler backward scheme leads to:

$$\begin{cases} \dfrac{z_{k+1} - z_k}{h} = \bar{A} z_{k+1} + \hat{B}\lambda_{k+1} \\ 0 \leq z_{1,k+1} \perp \lambda_{k+1} \geq 0 \end{cases} \tag{63}$$

with $\hat{B} = (0, \dots, 0, CA^{r-1}B)^T$. Using (63), we get the following complementarity problem $(z_{1,k+1}, \Lambda_{k+1}) = LCP(\frac{1}{h^r}M, b_{k+1})$ with the LCP matrix given by

$$M = CA^{r-1}[hW(I - h\bar{A})^{-1}\bar{B} + B]. \tag{64}$$

It is clear from (64) that if h vanishes, then the LCP matrix is close to the matrix $CA^{r-1}B$ which is the LCP matrix of the time-continuous ZD form,. So if $CA^{r-1}B$ assures that the LCP(λ) has a unique solution, this should be the case for the discretized LCP as well, for small enough step h.

The interest of working with the ZD dynamics lies in the fact that this allows one to keep the properties of the LCP from the continuous time t formulation to the discretized formulation. This seems to be some kind of minimal requirement for the discrete algorithm, since in any case λ_{k+1} has to be calculated[4].

3.5 Empirical order of the scheme

In order to conclude, we give in this paragraph an estimation on the order of the scheme by studying empirically on a fourth order example.

In this example, we illustrate the role of the zero dynamics on the behaviour of the following system

$$\begin{cases} \dot{z}_1(t) = z_2(t) \\ \dot{z}_2(t) = z_3(t) \\ \dot{z}_3(t) = -z_1(t) - z_2(t) - z_3(t) - d_\xi^T \xi(t) + \lambda(t) \\ \dot{\xi}_1(t) = \xi_2(t) \\ \dot{\xi}_2(t) = -\xi_1(t) + z_1(t) \\ w(t) = z_1(t) \geq 0 \end{cases} \tag{65}$$

with the initial condition $z(0) = (1, 0, 0, 0, 0)^T$ and $d_\xi = (0, -1)$. All the simulations are performed with Scilab©. In this case, we have a sequence of non trivial intervals where the constraints remain active, see Figure 2. The time interval is $[0, 10]$ and the time step is equal to $h = 10^{-1}$.

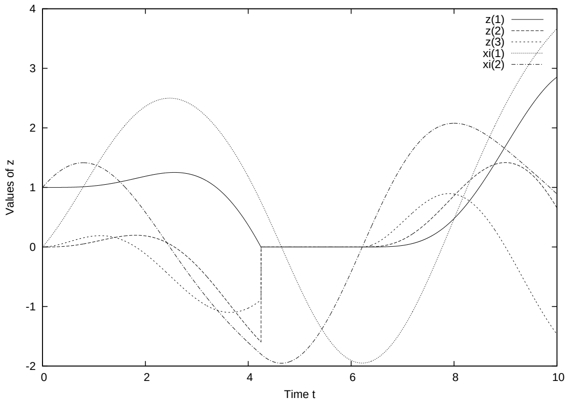


Figure 2. Trajectories of z and ξ given by the proposed numerical scheme

If we want to evaluate the order of accuracy of the scheme on this simple example, we need to use a norm consistent with the set of RCLBV functions and to introduce a notion of convergence providing a reasonable substitute to the uniform convergence of the continuous function. To overcome this difficulty, the convergence in the sense of filled-in graph has been introduced by Moreau (Moreau, 1978). Shortly, for a RCLBV function $f : [0, T] \mapsto \mathbb{R}^n$, we define the filled-in graph, $gr^\star f$ by adding some line segments to the graph of f in such a way that all the gap are filled:

$$\begin{aligned} gr^\star f = \{ & (t, x) \in [0, T] \times \mathbb{R}^n, \\ & 0 \leq t \leq T \text{ and } x \in [f(t^-), f(t^+)]) \} \end{aligned} \tag{66}$$

Such graphs are closed bounded subset of $[0, T] \times \mathbb{R}^n$, hence, we can use the Hausdorff distance between two such sets with a suitable metric:

$$d((t, x), (s, y)) = \max(|t - s|, \|x - y\|_{\mathbb{R}^n}) \tag{67}$$

Defining the excess of separation between two graphs by

$$e(gr^\star f, gr^\star g) = \sup_{(t,x) \in gr^\star f} \ \inf_{(s,y) \in gr^\star g} d((t, x), (s, y)), \tag{68}$$

the Hausdorff distance between two filled-in graphs $h^\star$ is defined by

$$h^\star(gr^\star f, gr^\star g) = \max(e(gr^\star f, gr^\star g), e(gr^\star g, gr^\star f)) \tag{69}$$

[4] Some schemes and some dynamical formulation, do not use an explicit calculation of the multiplier. But they necessarily use underlying arguments equivalent to having a well-posed LCP.

To compute a reference solution, the number of time-steps is chosen as $N = 10^6$, i.e., for a time step $h = 10^{-5}$. The results of the distance in the sense of filled-in graph is displayed in log scale on the Figure 3. On this example, the order of accuracy of the time-stepping scheme (34)–(35)–(36) is close to 1 as expected.

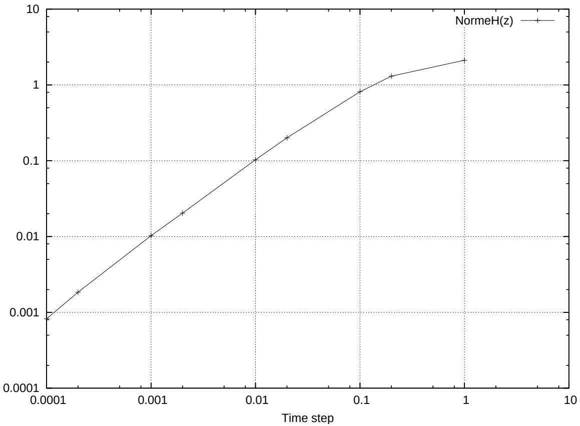


Figure 3. Empirical order of the scheme

4 Conclusion

In this paper, several basic facts have been presented about the problem of the numerical time integration of the Moreau's Sweeping Process (MSP). We have paid closed attentions to one its generalizations: the Higher Order Moreau's Sweeping Process (HOMSP)in proposing a formulation and an associated time-stepping scheme to deal with non smooth evolutions without explicit event-handling procedures.

References

Acary, V. and B. Brogliato (2003). Higher order moreau's sweeping process. In: *Non smooth Mechanics and Analysis: theoretical and numerical advances, Colloquium in the honor of the 80th Birthday of J.J. Moreau (2003)* (P. Alart, O. Maisonneuve and R.T. Rockafellar, Eds.). Kluwer.

Acary, V., B. Brogliato and D. Goeleven (2005). Higher order moreau's sweeping process : Mathematical formulation and numerical simulation. *In revision for Mathematical Programming A.* Draft version in INRIA Research report 5236, 2004, `www.inria.fr/rrrt/rr-5236.html`.

Brezis, H. (1973). *Opérateurs maximaux monotones et semi-groupe de contraction dans les espaces de Hilbert*. North Holland. Amsterdam.

Camlibel, K., W.P.M.H. Heemels and J.M. Schumacher (2002). Consistency of a time-stepping method for a class of piecewise-linear networks.. *IEEE Trans. Circuits and systems I* **49**, 349–357.

Goeleven, D. and B. Brogliato (2004). Stability and instability matrices for linear evolution variational inequalities. *IEEE Transactions on Automatic Control* **49**(4), 521–534.

Heemels, W.P.M.H., J.M. Schumacher and S. Weiland (2000). Linear complementarity problems. *S.I.A.M. Journal of applied mathematics* **60**(4), 1234–1269.

Jean, M. (1999). The non smooth contact dynamics method. *Computer Methods in Applied Mechanics and Engineering* **177**, 235–257. Special issue on computational modeling of contact and friction, J.A.C. Martins and A. Klarbring, editors.

Kunze, M. and M.D.P. Monteiro Marques (2000). An introduction to moreau's sweeping process. In: *Impact in Mechanical systems: Analysis and Modelling* (B. Brogliato, Ed.). Vol. 551 of *Lecture Notes in Physics*. pp. 1–60. Springer.

Monteiro Marques, M. D. P. (1993). *Differential Inclusions in NonSmooth Mechanical Problems : Sh ocks and Dry Friction*. Birkhauser, Verlag.

Moreau, J.J. (1971). Rafle par un convexe variable (premire partie), exposé n 15. *Séminaire d'analyse convexe*.

Moreau, J.J. (1972). Rafle par un convexe variable (deuxime partie) exposé n3. *Séminaire d'analyse convexe*.

Moreau, J.J. (1977). Evolution problem associated with a moving convex set in a Hilbert space. *Journal of Differential Equations* **26**, 347–374.

Moreau, J.J. (1978). Approximation en graphe d'une évolution discontinue. *RAIRO Analyse numérique/ Numerical Analysis* **12**, 75–84.

Moreau, J.J. (1983). Liaisons unilatérales sans frottement et chocs inélastiques. *Comptes Rendus de l'Acadmie des Sciences* **296 serie II**, 1473–1476.

Moreau, J.J. (1988). Unilateral contact and dry friction in finite freedom dynamics. In: *Nonsmooth mechanics and applications* (J.J. Moreau and P.D. Panagiotopoulos, Eds.). pp. 1–82. Number 302 In: *CISM, Courses and Lectures*. Springer Verlag.

Moreau, J.J. (1999). Numerical aspects of the sweeping process. *Computer Methods in Applied Mechanics and Engineering* **177**, 329–349. Special issue on computational modeling of contact and friction, J.A.C. Martins and A. Klarbring, editors.

Moreau, J.J. (2003). An introduction to unilateral dynamics. In: *Novel Approaches in Civil Engineering* (M. Frémond and F. Maceri, Eds.). Springer Verlag.

Robinson, S.M. (1979). Generalized equations and their solutions. I. Basic theory. *Mathematical programming study* **10**, 128–141.

Schatzman, M. (1978). A class of nonlinear differential equations of second order in time.. *Non linear Analysis* **2**(3), 355–373.

Stewart, D. (1998). Convergence of a time-stepping scheme for rigid-body dynamics and resolution of Painlevé's problem. *Archives for Rational Mechanics and Analysis* **145**, 215–260.