HAL open science

Higher order event capturing time-stepping schemes for nonsmooth multibody systems with unilateral constraints and impacts

Vincent Acary

To cite this version:

Vincent Acary. Higher order event capturing time-stepping schemes for nonsmooth multibody systems with unilateral constraints and impacts. Applied Numerical Mathematics: an IMACS journal, 2012, Selected Papers from NUMDIFF-12, 62 (10), pp.1259-1275. 10.1016/j.apnum.2012.06.026 . inria-00476398

HAL Id: inria-00476398

https://inria.hal.science/inria-00476398v1

Submitted on 29 Oct 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Higher order event capturing time-stepping schemes for nonsmooth multibody systems with unilateral constraints and impacts

Vincent Acary

INRIA Grenoble Rhône-Alpes, Inovallée de Montbonnot, 655, Avenue de l'Europe. 38334 Saint Ismier Cedex, France

The article is devoted to the study of higher order time integration methods for multibody systems with unilateral constraints. After a brief presentation of the mathematical modeling of nonsmooth multibody systems, several estimates on the local error of consistency of the Moreau time-stepping scheme are given. Based on these estimates, an attempt at an adaptive time-step strategy is presented with academic examples. Finally, higher order event-capturing methods are designed by coupling implicit Runge–Kutta schemes and Moreau's time-stepping scheme.

Keywords:
Multibody systems
Nonsmooth mechanics
Unilateral constraints
Impact
Moreau's sweeping process
Numerical integration scheme
Adaptive time-step strategy
Order of accuracy
Error estimates

0. Notation

The following notation is used throughout the paper. The uniform norm for a function f is denoted by $\|f\|_\infty$ and for a vector $x \in \mathbb{R}^n$ by $\|x\|$. A function f is said to be of class $\mathcal{C}^p$ if it is continuously differentiable up to the order p. Let I denote a real time interval of any sort. The set of functions $f : I \to \mathbb{R}^n$ of bounded variations (BV) is denoted by $BV(I, \mathbb{R}^n)$. For $f \in BV(I, \mathbb{R}^n)$, we denote the right-limit function by $f^+(t) = \lim_{s\to t, s>t} f(s)$, and respectively the left-limit by $f^-(t) = \lim_{s\to t, s<t} f(s)$. We use the following convention introduced in [31]: if I contains its left end, T_l (respectively its right end T_r) we shall agree that $f^-(T_l) = f(T_l)$ (respectively $f^+(T_r) = f(T_r)$). The set of functions $f : I \to \mathbb{R}^n$ of Locally Bounded Variations (LBV) is denoted by $LBV(I, \mathbb{R}^n)$. We denote by $0 = t_0 < t_1 < \cdots < t_k < \cdots < t_N = T$ 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 $h_k = t_{k+1} - t_k$. For simplicity's sake, the schemes are presented in the sequel with a time step shortly denoted by h. The value of a real function $x(t)$ at the time t_k, is approximated by x_k. In the same way, the notation $x_{k+\theta} = (1-\theta)x_k + \theta x_{k+1}$ is used for $\theta \in [0, 1]$. The notation $\mathcal{O}(h)$ is to be understood as $h \to 0$. The notation $\mathrm{d}t$ defines the Lebesgue measure on $\mathbb{R}$.

1. Introduction and motivations

Let us briefly recall what is the context of the modeling and the simulation of multibody systems with unilateral constraints. Let us consider a multibody system described by a generalized coordinates vector $q(t) \in \mathbb{R}^n$ and a generalized velocities vector $v(t) \in \mathbb{R}^n$. In a pure Lagrangian setting, the equations of motion of multibody systems with unilateral constraints may be written as

$$
\begin{cases}
q(t_0) = q_0, \qquad v(t_0) = v_0, & \text{(a)} \\
\dot{q}(t) = v(t), & \text{(b)} \\
M\big(q(t)\big)\dot{v}(t) + F\big(t, q(t), v(t)\big) = G(t,q)\lambda(t), & \text{(c)} \\
g^{\alpha}\big(t, q(t)\big) = 0, \quad \alpha \in \mathcal{E}, & \text{(d)} \\
g^{\alpha}\big(t, q(t)\big) \geqslant 0, \qquad \lambda^{\alpha} \geqslant 0, \qquad \lambda^{\alpha} g^{\alpha}(t,q) = 0, \quad \alpha \in \mathcal{I}, & \text{(e)}
\end{cases}
\tag{1}
$$

where

- the initial conditions are $q_0 \in \mathbb{R}^n$ and $v_0 \in \mathbb{R}^n$,
- the mapping $M : \mathbb{R}^n \to \mathbb{R}^{n\times n}$ is the inertia matrix,
- the mapping $F : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ contains the external forces applied to the system, the internal forces and possibly the gyroscopic forces,
- the mapping $g : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^m$ describes the constraints on the system, the transpose of the Jacobian of the constraints is denoted by the mapping $G(t,q) = \nabla_q g(t,q) : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^{m\times n}$, and
- the sets $\mathcal{E} \subset \mathbb{N}$ and $\mathcal{I} \subset \mathbb{N}$ respectively describe the set of bilateral constraints and unilateral constraints.

The choice of the Lagrangian setting rather than another formulation (e.g. Newton/Euler) is chosen for the sake of simplicity without any loss of generality for the further developments. Similarly, we will consider in this paper that $\mathcal{E} = \emptyset$ and that the constraints are scleronomous constraints, *i.e.* $g(t, q(t)) = g(q(t))$. The methods and results developed in this context extend straightforwardly to the more general case (1). Finally, let us define the following variables relative to the constraints, called *local variables*: the local velocity $U(t)$ and the (local) Lagrange multiplier $\lambda(t)$ which is associated with the generalized reaction forces $r(t)$ such that

$$
U(t) = G^T(q)v(t), \qquad r(t) = G(q)\lambda(t). \tag{2}
$$

By denoting $g(q) = [g_k(q), k \in \mathcal{I}]^T$ and $\lambda = [\lambda_k, k \in \mathcal{I}]^T$, the condition (1)(d) is called the *Signorini condition* or the complementarity condition which will be denoted compactly as

$$
0 \leqslant g(q) \perp \lambda \geqslant 0, \quad \text{or equivalently} \quad -\lambda \in N_{\mathbb{R}^m_+}\big(g(q)\big). \tag{3}
$$

The set $N_K(x)$ stands for the normal cone to a convex set K [28,36] taken at $x \in K$. Using the inclusion in (3), the dynamics can be cast into a Differential Inclusion (DI) form as

$$
\begin{cases}
\dot{q}(t) = v(t), & \text{(a)} \\
M\big(q(t)\big)\dot{v}(t) + F\big(t, q(t), v(t)\big) = r(t) \in -G\big(q(t)\big)N_{\mathbb{R}^m_+}\big(g\big(q(t)\big)\big). & \text{(b)}
\end{cases}
\tag{4}
$$

It is well-known that such dynamics can be nonsmooth and may encounter jumps in the velocities. Therefore, the velocity v usually belongs to $LBV(I, \mathbb{R}^n)$ and the acceleration is a differential measure $\mathrm{d}v$ associated with v. The absolutely continuous generalized coordinates q are integrated from the velocity in a usual way

$$
q(t) = q(t_0) + \int_{t_0}^{t} v(\tau)\,\mathrm{d}\tau, \tag{5}
$$

and the equation of motion (4)(b) is reformulated as a Measure Differential Inclusion (MDI),

$$
M\big(q(t)\big)\,\mathrm{d}v - F\big(t, q(t), v^+(t)\big)\,\mathrm{d}t = -G\big(q(t)\big)\,N_{\mathbb{R}^m_+}\big(g\big(q(t)\big)\big). \tag{6}
$$

The Lagrange multiplier λ is then replaced by a measure $\mathrm{d}I$ such that

$$
-\mathrm{d}I \in N_{\mathbb{R}^m_+}\big(g(q)\big). \tag{7}
$$

Similarly at the generalized forces level, r is replaced by $\mathrm{d}i = G(q)\mathrm{d}I$. To complete the modeling of a multibody system with unilateral constraints, an impact law has to be stated, defining the value of the velocity after a jump. Without entering into deeper details of the impact modeling, the Newton impact law is chosen for its simplicity in order to define the post-impact velocity such that

$$
U^+(t) = -\rho U^-(t), \tag{8}
$$

where ρ is the coefficient of restitution.

For more details on the modeling of multibody systems with unilateral constraints, we refer to [3,35,30] and for the mathematical analysis, we refer to [37,26,39,9].

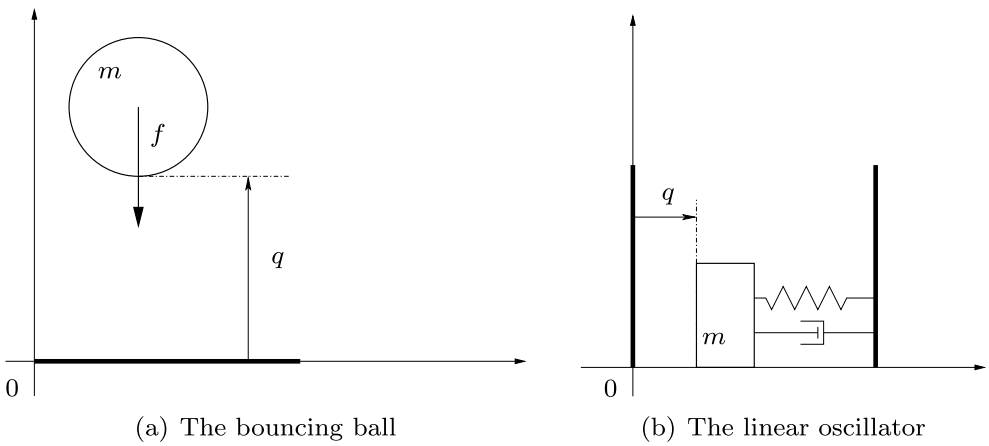


Fig. 1. Simple archetypal test examples.

Remark 1. Note that the formulation (6) encompasses the dynamics of flexible multibody systems which are space-discretized by Finite Element Method (FEM) or any other Galerkin-type method. In the case of standard FEM applications, the generalized coordinates q are the displacements at the nodes of the mesh. In co-rotational approaches [17], the generalized coordinates are a mixture of finite rotations and displacements at the nodes of the mesh in a spatial frame (see [13] for a comprehensive discussion).

Throughout the paper, two test examples will be chosen to highlight the properties of the considered numerical integration scheme.

Example 1 *(The bouncing ball).* This is the standard bouncing ball under gravity depicted in Fig. 1(a). The dynamics is constant with a forcing term equal to f together with a unilateral contact on the ground,

$$\begin{cases} \dot{v}(t) = f(t) + \lambda(t), & \dot{q}(t) = v(t), \\ 0 \leqslant q(t) \perp \lambda(t) \geqslant 0, & v^+(t) = -\rho v^-(t), \quad \text{if } q(t) = 0. \end{cases} \tag{9}$$

The interesting feature of the bouncing ball example is the presence of a finite accumulation of impact when $0 < e < 1$ and $f < 0$. The analytical solution of this example can be found in [12]. A more pleasant analytical solution due to Ballard [10] is provided in the sequel. It will be used as a benchmark in the further sections. The parameters are chosen as $f = -2$, $\rho = 1/2$ and the initial data as $t_0 = 0$, $q_0 = 1$ and $v_0 = 0$. The analytical solution reads

$$\begin{cases} q(t) = -t^2 + 1, \\ v(t) = -2t, \end{cases} \qquad t \in [0, 1),$$

$$\begin{cases} q(t) = -(t-3)^2 - \dfrac{3}{2^n}(t-1) + \dfrac{1}{2^{n-1}}\left(3 - \dfrac{1}{2^n}\right), \\ v(t) = -2(t-3) - \dfrac{3}{2^n}, \end{cases} \qquad t \in \left[3 - \frac{1}{2^{n-1}}, 3 - \frac{1}{2^n}\right),$$

$$\begin{cases} q(t) = 0, \\ v(t) = 0, \end{cases} \qquad t \in [3, +\infty). \tag{10}$$

Example 2 *(The linear oscillator example).* The dynamics of this one-degree-of-freedom system depicted in Fig. 1(b) example is similar to the dynamics (9) but with a linear spring-damper internal force, that is

$$m\dot{v}(t) + cv(t) + kq(t) = \lambda(t). \tag{11}$$

The explicit analytical solution can be found in [22]. The previous trivial free dynamics (9) with a null or a constant forcing term are exactly integrated with any first order scheme. With the linear, but nontrivial, dynamical term in (11), the order of accuracy of higher order schemes can be exhibited.

Two types of methods are currently available to numerically integrate nonsmooth multibody systems. We briefly summarize their properties in the following paragraphs.

Nonsmooth event tracking method. These methods are also called *event-driven* methods, where the time of discontinuities in the velocity or in its derivatives, also called a *nonsmooth event* or shortly an event, is detected and accurately located.

Between two events, the system is integrated with any standard Differential Algebraic Equations (DAE) solver with a suitable order according to the regularity of the system. Such a method, detailed in [35,1,3] can be very efficient for simple problems but suffers from several drawbacks. If the number of events is large or worse infinite, in a finite time interval, the time integration cannot efficiently advance in time. This is particularly true when a finite accumulation is encountered. Secondly, event-driven methods are in practice very sensitive to the numerical tolerances used in the detection procedure of events. Thirdly, these methods require a reformulation of the constraints at higher kinematic levels (velocity, acceleration, ...). Due to the intrinsic unilateral character of the constraints, the derivatives of the constraints involve some additional conditional statements. On the numerical point of view, this index-like reduction implies the introduction of new numerical tolerances to trigger the conditional statements and their associated difficulties. To summarize, event-driven approaches are well-suited when nonsmooth events are rare and well-separated in time and in small-scale nonsmooth multibody systems.

Nonsmooth event capturing method. These methods are also called briefly *time-stepping methods*. The two main nonsmooth event capturing methods are due to Moreau [30] and to Paoli and Schatzman [33,34]. In such methods, the time-integration is performed with a time step, whose length does not depend on the exact location of nonsmooth events. The advantages of this class of methods are: a) their convergence proof, b) their efficiency even in the case of finite accumulation of impacts and c) their ability to work without any accurate event detection. Finally, another advantage of this method is that it does not require higher derivation of the unilateral constraints (velocity level for the Moreau scheme and direct coordinate level for the Schatzman–Paoli scheme). However, the major drawback is their intrinsic low order of accuracy. When events are encountered, the local error of consistency is at best $\mathcal{O}(h)$. Over smooth periods, the order $\mathcal{O}(h^2)$ is expected to be as for the numerical integration of index-2 DAEs with the backward Euler method.

The objective of this paper is to propose several alternative nonsmooth event capturing methods, *i.e.* time-stepping methods with higher-order accuracy results and better efficiency. The efficiency is measured by the ratio of the global error and the CPU effort. The targeted applications are mechanical systems with a small number of bodies, for which unilateral contact and free motion play an important role for the global behavior of the system. The rigid multibody systems in circuit breakers [1], robotic and control applications [5,21], transport applications [16] are the favorite ones. In such applications, the accurate treatment of a large number of events and finite accumulation due to the proper dynamics or clearances in joints is crucial. The quest for high accuracy in large-scale systems such as granular materials or large scale finite element applications is very expensive and useless most of the time. Nevertheless, the methods developed in this paper should apply to this type of systems but with an extreme CPU effort.

Finally, the work in [41,40] has to be cited as the first attempt to increase the efficiency of Moreau's scheme by an extrapolation method. However in the latter, no proof of order of accuracy can be found and the authors made the assumption that Moreau's scheme has always a local truncation error equal to $\mathcal{O}(h)$. We will see that the assumption fails to be satisfied for the velocities in most cases.

2. Moreau's sweeping process and time-stepping scheme

Moreau's scheme [29,30,32] for scleronomous holonomic perfect unilateral constraints is based on a formulation of unilateral constraints in terms of local velocities together with the Newton impact law (see [26,9,38] for details).

Moreau [30] proposed a compact formulation of the impact law as an MDI,

$$-\mathrm{d}I \in N_{T_{\mathbb{R}^m_+}(g(q(t)))}\big(U^+(t) + \rho U^-(t)\big) \tag{12}$$

where $T_{\mathbb{R}^m_+}(y)$ stands for the tangent cone to $\mathbb{R}^m_+$ at y [28,36]. Finally, we obtain an MDI, the so-called *Moreau sweeping process*,

$$M\big(q(t)\big)\,\mathrm{d}v - F\big(t, q(t), v^+(t)\big)\,\mathrm{d}t \in -G\big(q(t)\big)N_{T_{\mathbb{R}^m_+}(y(t))}\big(U^+(t) + \rho U^-(t)\big). \tag{13}$$

Remark 2. This formulation of the unilateral constraints together with Newton's impact law can be interpreted as an index reduction technique in DAE theory. If the constraints on the generalized coordinates are satisfied for the initial conditions, they are also satisfied at any time.

Well-posedness assumptions. The following assumptions are made to ensure the well-posedness of the problem.

Assumption 1 *(Existence and uniqueness).* A unique global solution over $[0, T]$ for Moreau's sweeping process is assumed such that $q(\cdot)$ is absolutely continuous and admits a right velocity $v^+(\cdot)$ at every instant t of $[0, T]$ and such that the function $v^+ \in LBV([0, T], \mathbb{R}^n)$.

Assumption 1 is ensured in the framework introduced by Ballard [9] who proves the existence and uniqueness of a solution in a general framework mainly based on the analyticity of data. The following assumption is standard for the applicability of time-stepping schemes of order $p \geqslant 1$.

Assumption 2 *(Smoothness of data).* The following smoothness on the data will be assumed: a) the inertia operator $M(q)$ is assumed to be of class $\mathcal{C}^p$ and definite positive, b) the force mapping $F(t,q,v)$ is assumed to be of class $\mathcal{C}^p$, c) the constraint functions $g(q)$ are assumed to be of class $\mathcal{C}^{p+1}$ and d) the Jacobian matrix $G(q) = \nabla_q^T g(q)$ is assumed to have full-row rank.

Throughout the paper, the definition of a smooth period is as follows

Definition 1 *(Smooth period of evolution).* Let us assume that the data satisfies Assumption 2. The system undergoes a smooth evolution over the so-called *smooth period* denoted by $S \subset I \subset \mathbb{R}$ if the local velocity $U^-(t) = U^+(t) \geqslant 0$ for all $t \in S$.

Over a smooth period, solving (1) amounts to solving an index-3 DAE. Under Assumption 2, a unique maximal solution of class $\mathcal{C}^{p+1}$ in the smooth period is ensured [9]. Under Assumption 2, the problem (13) can be stated in terms of the local variables U and $\mathrm{d}I$ such that

$$\begin{cases} \mathrm{d}U = W\big(q(t)\big)\,\mathrm{d}I + G^T\big(q(t)\big)M^{-1}\big(q(t)\big)F\big(t, q(t), v^+(t)\big)\,\mathrm{d}t, \\ -\mathrm{d}I \in N_{T_{\mathbb{R}^m_+}(y(t))}\big(U^+(t) + \rho U^-(t)\big), \end{cases} \tag{14}$$

where $W(q) = G^T(q)M^{-1}(q)G(q)$ is called the *Delassus operator* which is also of class $\mathcal{C}^p$ and invertible. Finally, we introduce the following notation

$$H(q) = M^{-1}(q)G(q)W^{-1}(q), \tag{15}$$

which is also assumed to be of class $\mathcal{C}^p$.

Moreau's time-stepping scheme extended with a θ-method. The numerical time integration of the MDI (13) is performed on an interval $(t_k, t_{k+1}]$ of length h as follows ($\theta \in [0,1]$):

$$\begin{cases} M(q_{k+\theta})(v_{k+1} - v_k) - hF(t_{k+\theta}, q_{k+\theta}, v_{k+\theta}) = p_{k+1} = G(q_{k+\theta})P_{k+1}, & \text{(a)} \\ q_{k+1} = q_k + hv_{k+\theta}, & \text{(b)} \\ U_{k+1} = G^T(q_{k+\theta})v_{k+1}, & \text{(c)} \\ -P_{k+1} \in N_{T_{\mathbb{R}^m_+}(\bar{g}_{k+\gamma})}(U_{k+1} + \rho U_k), & \text{(d)} \\ \bar{g}_{k+\gamma} = g(q_k) + h\gamma U_k, \quad \gamma \in [0,1], & \text{(e)} \end{cases} \tag{16}$$

where the following approximations are considered

$$v_{k+1} \approx v^+(t_{k+1}), \qquad U_{k+1} \approx U^+(t_{k+1}), \qquad p_{k+1} \approx \mathrm{d}i\big(]t_k, t_{k+1}]\big), \qquad P_{k+1} \approx \mathrm{d}I\big(]t_k, t_{k+1}]\big). \tag{17}$$

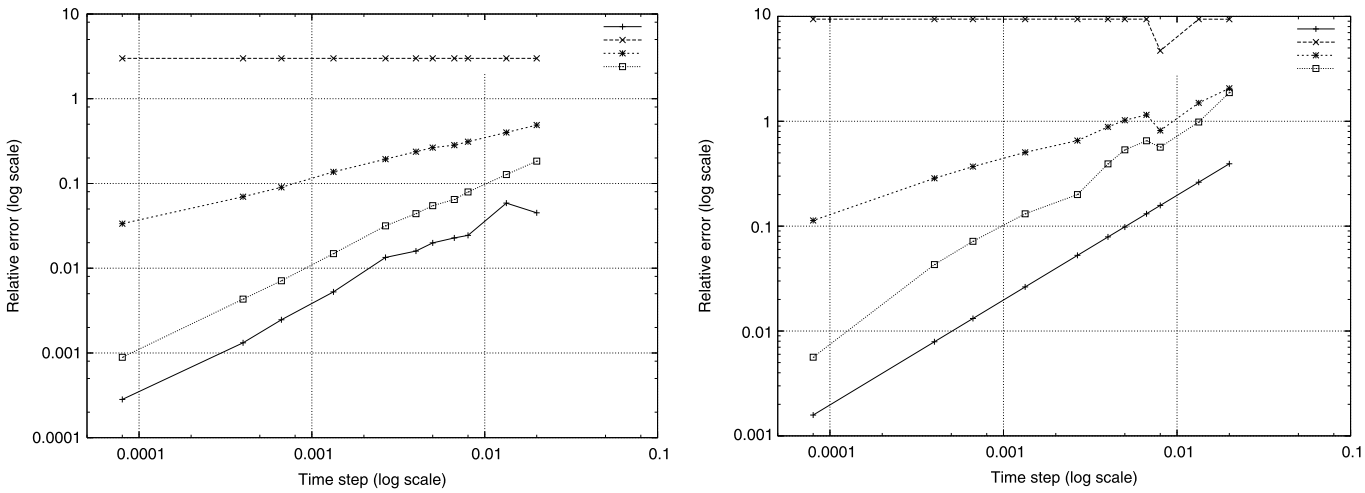
The value $\bar{g}_{k+\gamma}$ is a prediction of the constraint which allows the computation of the tangent cone $T_{\mathbb{R}^m_+}(\bar{g}_{k+1})$. The inclusion can be stated equivalently as a *conditional complementarity condition* for all $\alpha \in [1 \dots m] \in \mathcal{I}$,

$$\text{if } \bar{g}^\alpha_{k+\gamma} \leqslant 0 \quad \text{then } 0 \leqslant U^\alpha_{k+1} + \rho U^\alpha_k \perp P^\alpha_{k+1} \geqslant 0, \quad \text{otherwise } P^\alpha_{k+1} = 0. \tag{18}$$

Remark 3. Remark 2 can be completed by the following comment. As we said earlier, the inclusion (12) appears as an index reduction in DAE theory. Its time discretized counterpart (16)(d) implies some drift in the constraints on q. Nevertheless, it ensures the satisfaction of Newton's impact law at each time-step and possesses very good stability properties so that the chattering free stabilization on the constraints is achieved [4].

The convergence of Moreau's time stepping scheme has been shown in [26,24,38,14] under various assumptions mainly with $\theta \in \{0,1\}$. For the sake of readability, we will introduce the following short notation in the sequel $M_{k+\theta} = M(q_{k+\theta})$, $W_{k+\theta} = W(q_{k+\theta})$, $H_{k+\theta} = H(q_{k+\theta})$ and $F_{k+\theta} = F(t_{k+\theta}, q_{k+\theta}, v_{k+\theta})$.

Remark 4. The projection of the velocity onto the tangent cone of $\mathbb{R}^m_+$ yields a slight violation of the constraints in generalized coordinates which occurs at the impact. The violation of the discrete coordinate can be corrected by adding another multiplier on the position level leading to the following scheme:



(a) The bouncing ball (Example 1) (b) The linear oscillator (Example 2)

Fig. 2. Empirical order of convergence of Moreau's time-stepping scheme. × uniform norm, + Hausdorff norm, ✳ $\|\cdot\|_2$ norm, □ $\|\cdot\|_1$ norm.

$$\begin{cases} M_{k+\theta}(v_{k+1}-v_k) - hF_{k+\theta} = p_{k+1} = G_{k+\theta}P_{k+1}, & \text{(a)} \\ q_{k+1} = q_k + hv_{k+\theta} + G(q_{k+\theta})\tau_{k+1}, & \text{(b)} \\ U_{k+1} = G_{k+\theta}^T v_{k+1}, & \text{(c)} \\ -P_{k+1} \in N_{T_{\mathbb{R}^m_+}(\bar{g}_{k+1})}(U_{k+1} + \rho U_k), & \text{(d)} \\ -\tau_{k+1} \in N_{\mathbb{R}^m_+}(g_{k+1}), & \text{(e)} \\ \bar{g}_{k+1} = g(q_k) + h\gamma U_k, \quad \gamma \in [0,1]. & \text{(f)} \end{cases} \tag{19}$$

The multiplier τ_{k+1} has no physical meaning. It is an artificial projection embedded into Moreau's time-stepping scheme to satisfy the constraints at the discrete times. It is very similar to the GGL algorithm proposed by [15] in solving the drift problem in DAE theory after an index reduction procedure. It is noteworthy that this scheme has no rigorous convergence proof and can exhibit some unstable behavior on large-scale examples with quite large time-steps.

2.1. Empirical order of convergence

Although the Moreau scheme enjoys some convergence results, no general result has been proved concerning its local order of consistency and global order of convergence. We conclude this section with an empirical measure of the error of convergence on Examples 1 and 2.

Measuring errors. In order to evaluate the order of accuracy on simple examples, we need to use a norm which is consistent with the BV functions and then to introduce a notion of convergence with provides us with a reasonable substitute to the uniform convergence for the uniformly continuous functions: the convergence in the sense of filled-in graph introduced by Moreau [27]. Shortly, for an LBV function $f : [0,T] \mapsto \mathbb{R}^n$, we define the filled-in graph, $gr^\star(f)$ by adding line segments to the graph of f in such a way that all the gaps are filled:

$$\mathrm{gr}^\star(f) = \left\{(t,x) \in [0,T] \times \mathbb{R}^n,\ 0 \leqslant t \leqslant T \text{ and } x \in \left[f\left(t^-\right), f\left(t^+\right)\right]\right\}. \tag{20}$$

Such graphs are closed bounded subsets of $[0,T] \times \mathbb{R}^n$, hence, we can use the Hausdorff distance $h^\star(\mathrm{gr}^\star(f), \mathrm{gr}^\star(g))$ between two such sets with a suitable metric $d((t,x),(s,y)) = \max\{|t-s|, \|x-y\|\}$. Thanks to this Hausdorff distance we are able to measure the error with respect to a reference solution given on our examples by an analytical result. When an analytical solution is known, an equivalent grid-function norm to the function norm in $\mathcal{L}_p(I, \mathbb{R}^n)$ is defined and denoted by

$$\|e\|_p = \left(h \sum_{i=0}^{N} |e_i|^p\right)^{1/p}, \quad 1 \leqslant p < +\infty, \text{ with } e = \left[e_i = f_i - f(t_i),\ i = 0 \ldots N\right]^T. \tag{21}$$

The computational effort for $\|e\|_p$ is smaller than the Hausdorff distance for piecewise continuous analytical functions. Although one of the examples has an accumulation of impacts, it is possible to check on Fig. 2 that the empirical order of Moreau's scheme is near to 1 with the Hausdorff norm $h^\star(\cdot)$ and the $\|\cdot\|_1$ norm. With the uniform norm, convergence cannot be observed.

3. Local order estimates for Moreau's time-stepping scheme

The goal of this section is to evaluate the local truncation error for Moreau's scheme as

$$e = \begin{bmatrix} e_v \\ e_q \end{bmatrix} = \begin{bmatrix} v^+(t_k+h) - v_{k+1} \\ q(t_k+h) - q_{k+1} \end{bmatrix}, \tag{22}$$

starting with the exact solution as initial data, *i.e.*, $q_k = q(t_k)$ and $v_k = v^+(t_k)$. Under Assumption 1, the error can be estimated by

$$e_q = \int_{t_k}^{t_{k+1}} v(s)\,\mathrm{d}s - h\big(\theta v_{k+1} + (1-\theta)v_k\big), \tag{23}$$

and

$$e_v = \int_{(t_k,t_{k+1}]} \mathrm{d}v - hM_{k+\theta}^{-1}F_{k+\theta} - M_{k+\theta}^{-1}p_{k+1}. \tag{24}$$

Example 3 *(Bouncing ball continued).* Let us consider the bouncing ball example (Example 1) and a time interval such that the impacting time t_* belongs to $(t_k, t_{k+1}]$. The error is

$$\begin{cases} e_v = -(1+\rho)[v_k + hf\sigma], \\ e_q = -q_k - h\big(\rho(1-\sigma+1)\big)v_k - fh^2\left[\rho(1-\sigma)\sigma + \frac{1}{2}(1-\sigma)^2 + \theta\right] \end{cases} \quad \text{if } p_{k+1} = 0,$$

$$\begin{cases} e_v = -hf[1-\sigma-\rho\sigma], \\ e_q = -q_k - h\big((1+\rho)(1-\theta) - e\sigma\big)v_k - fh^2\left(\rho(1-\sigma)\sigma + \frac{1}{2}(1-\sigma)^2\right) \end{cases} \quad \text{if } p_{k+1} > 0, \tag{25}$$

where $\sigma = (t_k - t_*)/h \in (0,1]$. The approximate solution of the Moreau scheme depends on the forecast of the active constraints, *i.e.* $\bar{g}_{k+1} = q_k + \gamma h v_k$. Using the fact that $q(t_*) = q_k + v_k\sigma h + \frac{1}{2}(\sigma h)^2 = 0$, we obtain that $q_k = -\sigma v_k h - \frac{1}{2}f(\sigma h)^2$ and

$$\begin{cases} e_v = -(1+\rho)[v_k + hf\sigma], \\ e_q = -h\big(\rho(1-\sigma+1) - \sigma\big)v_k - fh^2\left[\rho(1-\sigma)\sigma + \frac{1}{2}(1-\sigma)^2 - \frac{1}{2}(\sigma)^2 + \theta\right] \end{cases} \quad \text{if } p_{k+1} = 0,$$

$$\begin{cases} e_v = -hf[1-\sigma-\rho\sigma], \\ e_q = -h\big((1+\rho)(1-\theta-\sigma)\big)v_k - fh^2\left(\rho(1-\sigma)\sigma + \frac{1}{2}(1-\sigma)^2 - \frac{1}{2}(\sigma)^2\right) \end{cases} \quad \text{if } p_{k+1} > 0. \tag{26}$$

Near the finite accumulation of impact at time $t = 3$, we can also try to evaluate the error. Let us consider a time step such that $[t_k, t_{k+1}] = [3-h, 3+h]$ and n_0 such that $h \in [1/2^{n_0}, 1/2^{n_0-1}]$. The local error in velocity is given if the impact is detected $p_{k+1} > 0$ by

$$e_v = v(3+h) - v_{k+1} = -2h - \frac{3}{2^{n_0}}. \tag{27}$$

As $h \to 0$, we have $n_0 \to \infty$, and $\frac{1}{2^{n_0}} = \mathcal{O}(h)$ and then $e_v = \mathcal{O}(h)$.

In Example 3, the consistency error in generalized coordinates e_q is always in $\mathcal{O}(h)$ and it is difficult to obtain a better approximation except for very particular choices of e and σ which cannot be chosen a priori by the user. The consistency error in velocity e_v is $\mathcal{O}(1)$ if the impact is not correctly forecast. In this case, there is no chance to reduce the amplitude of the consistency error with the time-step. The situation may happen if the activation of the constraint based on the prediction $\bar{g}_{k+1}$ is not correct, which systematically occurs in Example 3 with $\gamma = 0$.

3.1. General estimates for the local error

Lemma 1. *Let $I = [t_k, t_{k+1}]$. Let us assume that the function $f \in BV(I, \mathbb{R}^n)$. Then we have the following inequality for the θ-method, $\theta \in [0,1]$,*

$$\left\| \int_{t_k}^{t_{k+1}} f(s)\,\mathrm{d}s - h\big(\theta f(t_{k+1}) + (1-\theta)f(t_k)\big) \right\| \leqslant C(\theta)(t_{k+1} - t_k)\,\mathrm{var}(f, I), \tag{28}$$

where $\mathrm{var}(f, I) \in \mathbb{R}$ *is the variation of* f *on* I *and* $C(\theta) = \theta$ *if* $\theta \geqslant 1/2$ *and* $C(\theta) = 1 - \theta$ *otherwise. Furthermore, the value of* $C(\theta)$ *yields a sharp bound in* (28).

Proof. Let us consider the function $\varphi_\theta : [t_k, t_{k+1}] \to \mathbb{R}$ such that

$$\varphi_\theta(t) = (t - t_{k+1}) + h\theta \quad \text{for } \theta \in [0, 1] \text{ with } h = t_{k+1} - t_k. \tag{29}$$

Using the integration by parts for the differential measure $\mathrm{d}f$ associated with f (see [31]), we get

$$\int_{[t_k, t_{k+1}]} \varphi_\theta(t)\,\mathrm{d}f = h\big(\theta f(t_{k+1}) + (1-\theta) f(t_k)\big) - \int_{t_k}^{t_{k+1}} f(s)\,\mathrm{d}s. \tag{30}$$

The integral term involving the differential measure $\mathrm{d}f$ can be bounded as follows (see [31])

$$\left\| \int_{[t_k, t_{k+1}]} \varphi_\theta(t)\,\mathrm{d}f \right\| \leqslant \max_{t \in [t_k, t_{k+1}]} \big|\varphi_\theta(t)\big| \,\mathrm{var}(f, I) \tag{31}$$

since $\varphi_\theta(\cdot)$ is uniformly continuous on $[t_k, t_{k+1}]$ and $f \in BV(I, \mathbb{R}^n)$. Computing the value of the maximum value $C(\theta)$ gives,

$$\max_{t \in [t_k, t_{k+1}]} \big|\varphi_\theta(t)\big| = \max_{t \in [t_k, t_{k+1}]} \big|(t - t_{k+1}) + h\theta\big| = \begin{cases} h\theta, & \text{if } \theta \geqslant \frac{1}{2} \\ h(1-\theta), & \text{if } \theta \leqslant \frac{1}{2} \end{cases} \triangleq C(\theta)h. \tag{32}$$

The result (28) is obtained from (30) and (31). In order to prove that the bound is sharp, let us consider for instance that f is given by $f(t) = \alpha$ for $t \in (t_k, t_{k+1})$, $f(t_k) = 1$, and $f(t_{k+1}) = \beta$. Choosing $\alpha = 1$, $0 \leqslant \beta < 1$ for $1/2 \leqslant \theta \leqslant 1$ and $\beta = \alpha$, $0 \leqslant \beta < 1$ for $0 \leqslant \theta < 1/2$, a straightforward calculus shows that the bound is attained. □

Proposition 1. *Under Assumptions* 1 *and* 2*, the local order of consistency of the Moreau time-stepping scheme for the generalized coordinates is* $e_q = \mathcal{O}(h)$ *and at least for the velocities* $e_v = \mathcal{O}(1)$.

Proof. The estimate e_v on the velocity is trivial if we recall that $M_{k+\theta}^{-1}(F_{k+\theta} + p_{k+1})$ is bounded on $[t_k, t_{k+1}]$. The BV function $v(\cdot)$ is also bounded on $[t_k, t_{k+1}]$ then we have that e_v is also bounded. Therefore, we obtain $e_v = \mathcal{O}(1)$. Using Lemma 1 for $v^+ \in BV(I, \mathbb{R}^n)$, we get

$$\left\| \int_{t_k}^{t_{k+1}} v(s)\,\mathrm{d}s - h\big(\theta v^+(t_{k+1}) + (1-\theta) v^+(t_k)\big) \right\| \leqslant C(\theta) h\,\mathrm{var}\big(v^+, I\big). \tag{33}$$

Since $v_k = v^+(t_k)$ and $v_{k+1} = v(t_{k+1}) + \mathcal{O}(1)$, we obtain for (33)

$$\left\| \int_{t_k}^{t_{k+1}} v(s)\,\mathrm{d}s - h\big(\theta v_{k+1} + (1-\theta) v_k\big) - \theta\mathcal{O}(h) \right\| = \big\| e_q - \theta\mathcal{O}(h) \big\| \leqslant C(\theta) h\,\mathrm{var}\big(v^+, I\big), \tag{34}$$

which completes the proof. □

3.2. Definition of index sets

In order to improve the estimate in Proposition 1, we need to introduce several index sets of constraints given by the following definition.

Definition 2 *(Index sets* $\mathcal{I}_*$*,* $\mathcal{I}(t)$ *and* $\mathcal{I}_k$*).* Three specific index sets of constraints are defined:

1. the index set $\mathcal{I}_*$ such that the exact solution has an impact at t_*
$$\mathcal{I}_* = \big\{\alpha \in \mathcal{I} \;\big|\; g^\alpha(q) = 0,\ P^\alpha \geqslant 0,\ U^{\alpha,+}(t_*) - U^{\alpha,-}(t_*) = -(1+\rho) U^{\alpha,-}(t_*) > 0\big\} \subset \mathcal{I}, \tag{35}$$
2. the index set $\mathcal{I}(t)$ such that the exact solution has a persistent contact
$$\mathcal{I}(t) = \big\{\alpha \in \mathcal{I} \;\big|\; g^\alpha(q) = 0,\ \lambda^\alpha(t) \geqslant 0,\ U^{\alpha,+}(t) = U^{\alpha,-}(t) = 0\big\} \subset \mathcal{I}, \tag{36}$$
3. the index set $\mathcal{I}_k$ such that the approximate solution with the exact initial data $q(t_k)$, $v(t_k)$ encounters an impact in the time-step $(t_k, t_{k+1}]$
$$\mathcal{I}_k = \big\{\alpha \in \mathcal{I} \;\big|\; \bar{g}_{k+1}^\alpha \leqslant 0,\ P_{k+1}^\alpha \geqslant 0,\ U_{k+1}^\alpha - U_k^\alpha = -(1+\rho) U_k^\alpha\big\} \subset \mathcal{I}. \tag{37}$$

3.3. Smooth motion with persistent contacts

Assumption 3. Let us assume that the system evolves in a smooth period with only persistent contacts in $(t_k, t_{k+1}]$. In particular, we assume a constant index set $\mathcal{I}(t)$ for all $t \in (t_k, t_{k+1}]$ and

$$\mathrm{d}I^\alpha = \lambda^\alpha(t)\,\mathrm{d}t, \quad \alpha \in \mathcal{I}(t) \quad \text{or equivalently} \quad \mathrm{d}i = r(t)\mathrm{d}t. \tag{38}$$

Proposition 2. *Let us assume that Assumptions 2 and 3 hold. If $\mathcal{I}(t) = \mathcal{I}_k$ for all $t \in (t_k, t_{k+1}]$ and $U_k^\alpha = 0$ for all $\alpha \in \mathcal{I}_k$, then the local order of consistency of the scheme is $e_v = \mathcal{O}(h^2)$ and $e_q = \mathcal{O}(h^2)$.*

Proof. In the case that $\mathcal{I}(t) = \mathcal{I}_k$ for all $t \in (t_k, t_{k+1}]$, the Moreau time-stepping amounts to solving

$$\begin{cases} M_{k+\theta}(v_{k+1} - v_k) - hF_{k+\theta} = G_{k+\theta}P_{k+1}, \\ q_{k+1} = q_k + hv_{k+\theta}, \\ U_{k+1} = G_{k+\theta}^T v_{k+1} = 0. \end{cases} \tag{39}$$

Thanks to Assumption 2, a unique solution of class $\mathcal{C}^{p+1}$ is expected with a multiplier of class $\mathcal{C}^p$. The time-stepping scheme can therefore be studied as an application of the backward Euler scheme or a θ-method for an index-2 DAE. The results on the local order of convergence obtained for implicit Runge–Kutta method [19,18] or for backward differentiation formulas [23,11] can be straightforwardly extended to complete the proof. □

3.4. Continuous Lagrange multiplier with a single impact in the time-step

In this section, the following assumptions are stated

Assumption 4. Let us assume that only a single impact at $t_* \in (t_k, t_{k+1}]$ occurs in the time interval. Neglecting the singular continuous part of the decomposition of $\mathrm{d}i$ and $\mathrm{d}v$, the following decomposition is assumed

$$\mathrm{d}i = p\delta_{t_*} + r(t)\,\mathrm{d}t, \quad \text{and} \quad \mathrm{d}v = \big(v^+(t_*) - v^-(t_*)\big)\delta_{t_*} + u'(t)\,\mathrm{d}t, \tag{40}$$

where the multiplier $r(\cdot)$ and the velocity $u(\cdot)$ are assumed to be absolutely continuous on $[t_k, t_{k+1}]$.

Proposition 3. *Let us assume that Assumptions 1, 2 and 4 hold.*

1. *If $\mathcal{I}_* = \mathcal{I}_k$, then the local order of consistency of the Moreau time stepping scheme for the velocity is $e_v = \mathcal{O}(h)$.*
2. *If $\mathcal{I}_* \neq \mathcal{I}_k$, then we retrieve the rough estimate of Proposition 1, that is $e_v = \mathcal{O}(1)$.*

Proof. Under Assumption 4, the MDI (14) can be decomposed to obtain

$$M\big(q(t)\big)u'(t) = F\big(t, q(t), v^+(t)\big) + r(t), \quad \mathrm{d}t\text{-almost everywhere} \tag{41}$$

and

$$M\big(q(t_*)\big)\big(v^+(t_*) - v^-(t_*)\big) = p, \quad \text{for } t = t_*, \tag{42}$$

where the differential measure $\mathrm{d}v$ has been decomposed as

$$\mathrm{d}v = u'(t)\,\mathrm{d}t + \big(v^+(t_*) - v^-(t_*)\big)\delta_{t_*} = M\big(q(t)\big)^{-1}\big[F\big(t, q(t), v^+(t)\big) + r(t)\big]\,\mathrm{d}t + M\big(q(t_*)\big)^{-1}p\delta_{t_*}. \tag{43}$$

The function $u'(\cdot)$ is equal to $v'(\cdot)$ $\mathrm{d}t$-almost everywhere. In this case, the error e_v is

$$e_v = \int_{t_k}^{t_{k+1}} M^{-1}\big(q(t)\big)\big[F\big(t, q(t), v^+(t)\big) + r(t)\big]\,\mathrm{d}t - hM_{k+\theta}^{-1}F_{k+\theta} + M\big(q(t_*)\big)^{-1}p - M_{k+\theta}^{-1}p_{k+1}. \tag{44}$$

Since the function $t \mapsto M(q(t))^{-1}r(t)$ is assumed to be continuous, it is bounded on $[t_k, t_{k+1}]$. The mappings F and M are assumed to be of class C^p, $p \geqslant 1$ and $v^+ \in BV([t_k, t_{k+1}], \mathbb{R}^n)$, therefore the function $t \mapsto M^{-1}(q(t))F(t, q(t), v^+(t))$ is also bounded on $[t_k, t_{k+1}]$. The following estimate for the integral term of (44) is then deduced

$$\int_{t_k}^{t_{k+1}} M^{-1}\big(q(t)\big)\big[F\big(t, q, v^+\big) + r(t)\big]\,\mathrm{d}t - hM_{k+\theta}^{-1}F_{k+\theta} = \mathcal{O}(h). \tag{45}$$

The remaining term in (44) can be computed as

$$M\big(q(t_*)\big)^{-1}p - M_{k+\theta}^{-1}p_{k+1} = M\big(q(t_*)\big)^{-1}G\big(q(t_*)\big)W^{-1}\big(q(t_*)\big)\big[U^+(t_*) - U^-(t_*)\big] \\ - M_{k+\theta}^{-1}G_{k+\theta}W_{k+\theta}^{-1}\big[U_{k+1} - U_k - hG_{k+\theta}^T M_{k+\theta}^{-1}F_{k+\theta}\big], \tag{46}$$

and using (15), we have

$$M\big(q(t_*)\big)^{-1}p - M_{k+\theta}^{-1}p_{k+1} = H\big(q(t_*)\big)\big[U^+(t_*) - U^-(t_*)\big] - H_{k+\theta}\big[U_{k+1} - U_k - hG_{k+\theta}^T M_{k+\theta}^{-1}F_{k+\theta}\big]. \tag{47}$$

The mapping $H(q)$ is assumed to be of class $\mathcal{C}^p$, $p \geqslant 1$ therefore

$$H(q_{k+\theta}) - H(q) = \nabla_q^T H(q)(q_{k+\theta} - q) + \mathcal{O}\big(\|q_{k+\theta} - q\|^2\big), \tag{48}$$

where the gradient of the matrix $H(q)$, which is a 3-order tensor, is denoted by $\nabla_q^T H(q)$. Let us now evaluate $q_{k+\theta} - q(t_*) = (1-\theta)q(t_k) + \theta q_{k+1} - q(t_*)$. By Proposition 1, we have $q_{k+1} = q(t_{k+1}) + \mathcal{O}(h)$. Since v is bounded, we have also $q(t_{k+1}) = q(t_k) + \mathcal{O}(h)$ and then $q_{k+1} = q(t_k) + \mathcal{O}(h)$. We can write $q_{k+\theta} - q(t_*) = q(t_k) - q(t_*) + \mathcal{O}(h)$. The boundedness of v gives

$$q_{k+\theta} - q(t*) = \mathcal{O}(h). \tag{49}$$

Using (48) and (49) and the regularity assumption on H, we obtain

$$H(q_{k+\theta}) - H\big(q(t_*)\big) = \mathcal{O}(h). \tag{50}$$

For the expression (47), we get

$$M\big(q(t_*)\big)^{-1}p - M_{k+\theta}^{-1}p_{k+1} = H_{k+\theta}\big[U^+(t_*) - U^-(t_*) - \big[U_{k+1} - U_k - hG_{k+\theta}^T M_{k+\theta}^{-1}F_{k+\theta}\big]\big] + \mathcal{O}(h). \tag{51}$$

Let us define in (51) the error e_U in terms of local velocity at contact as,

$$e_U = \big[U^+(t_*) - U^-(t_*) - \big[U_{k+1} - U_k - hG_{k+\theta}^T M_{k+\theta}^{-1}F_{k+\theta}\big]\big]. \tag{52}$$

Depending on the constraint $\alpha \in \mathcal{I}$ belongs or not to the index sets $\mathcal{I}_*$ and $\mathcal{I}_k$, the error e_U can be estimated as follows.

Case 1. Let us assume that $\mathcal{I}_* = \mathcal{I}_k$.

For all $\alpha \in \mathcal{I}_* \cap \mathcal{I}_k$, we have $U^{+,\alpha}(t_*) = -\rho^\alpha U^{-,\alpha}(t_*)$ and $U_{k+1}^\alpha = -\rho^\alpha U_k^\alpha$. The error term e_U^α is written

$$\begin{aligned} e_U^\alpha &= -\big(1+\rho^\alpha\big)\big(U^{-,\alpha}(t_*) - U_k^\alpha\big) + hG_{k+\theta}^{T,\alpha}M_{k+\theta}^{-1}F_{k+\theta} \\ &= -\big(1+\rho^\alpha\big)\int_{t_k}^{t_*} U'^\alpha(t)\,\mathrm{d}t + hG_{k+\theta}^{T,\alpha}M_{k+\theta}^{-1}F_{k+\theta}. \end{aligned} \tag{53}$$

Since $U'^\alpha(t) = G^T(q(t))M^{-1}(q(t))[F(t,q(t),v^+(t)) + r(t)]$ $\mathrm{d}t$-almost everywhere, we have

$$\int_{t_k}^{t_*} U'^\alpha(t)\,\mathrm{d}t = \int_{t_k}^{t_*} G^T\big(q(t)\big)M^{-1}\big(q(t)\big)\big[F\big(t,q(t),v^+(t)\big) + r(t)\big]\,\mathrm{d}t. \tag{54}$$

With the same argument of boundedness of $G^T(q(t))M^{-1}(q(t))[F(t,q(t),v^+(t)) + r(t)]$, we get

$$\int_{t_k}^{t_*} U'^\alpha(t)\,\mathrm{d}t = \mathcal{O}(h), \tag{55}$$

and therefore $e_U^\alpha = \mathcal{O}(h)$. By Definition 2, we have $P^\beta = P_{k+1}^\beta = 0$ for all $\beta \notin \mathcal{I}_* \cap \mathcal{I}_k$. We conclude that $e_U^\beta = 0$. From (45), (51) and (55), the consistency error is $e_v = \mathcal{O}(h)$.

Case 2. Let us assume that $\mathcal{I}_* \neq \mathcal{I}_k$. The error term e_U^α is

$$\begin{aligned} e_U^\alpha &= -\big(1+\rho^\alpha\big)\big(U_k^\alpha\big) + hG_{k+\theta}^{T,\alpha}\Phi_{v,k} = \mathcal{O}(1) \quad \text{for all } \alpha \notin \mathcal{I}_* \text{ and } \alpha \in \mathcal{I}_k, \\ e_U^\alpha &= -\big(1+\rho^\alpha\big)\big(U^-(t_*)\big) = \mathcal{O}(1) \quad \text{for all } \alpha \in \mathcal{I}_* \text{ and } \alpha \notin \mathcal{I}_k. \end{aligned} \tag{56}$$

From (45), (47), and (56), the consistency error for velocities is $e_v = \mathcal{O}(1)$. □

3.5. Comments on the local error estimates of Propositions 1 and 3

The local error estimates obtained in this section are quite rough. In summary, the following points may be outlined:

- The local error in coordinates is $e_q = \mathcal{O}(h)$ and it cannot be improved as the bouncing ball example shows. Note that even though the velocity is exactly integrated in time, we cannot expect a better order for the numerical integration of q as Lemma 1 shows.
- The local error in velocity is at least $e_v = \mathcal{O}(1)$ if the impact is not well-forecast. In practice, this situation is usual. For instance, if $\gamma = 0$ is chosen in (16)(e), the impact is not forecast in the correct time-interval for the academic examples presented in this paper. It illustrates the possible convergence problem that we can have in uniform norm as we mentioned in Section 2.1.
- The local error in velocity is shown to be at least $e_v = \mathcal{O}(h)$ with only one impact in the time-interval and a well-forecast impact. If there is no right accumulation of impacts at any points, this proof is sufficient in theory. Indeed, we can always find a sufficiently small time-step such that there is only one impact inside $(t_k, t_{k+1}]$. However, in numerical practice, this result is not satisfactory because such a time step will vanish at the accumulation of impacts. An open issue is to prove that we have $e_v = \mathcal{O}(h)$ in a time-step which contains a well-forecast finite accumulation as we made in Example 3.

4. An attempt at adjusting the time-step size for Moreau's scheme

4.1. Practical local error estimates in standard smooth case and automatic step-size control

In order to control the time step, a practical estimation of the consistency error is needed. Let us consider the standard case of ODE given by $\dot{x} = f(x,t)$. A standard error estimation for a scheme with a local order in $\mathcal{O}(h^{p+1})$ may be based on an extrapolation with halved time-steps (see [20, pages 164–172]) and yields a practical error estimate such that:

$$e_2 = x(t_0 + h) - x_2 = \frac{x_{1/2} - x_1}{2^p - 1} + \mathcal{O}\big(h^{p+2}\big), \tag{57}$$

where $x_1, x_{1/2}$ and x_2 are the values obtained by the numerical time-integration with halved time-steps.

Based on a local error estimate, the time-step selection method can be designed following the standard procedure set out in [20]. For a given absolute tolerance vector $atol \in \mathbb{R}^n$ and a relative tolerance vector $rtol \in \mathbb{R}^n$, the optimal step size is chosen as

$$h_{\text{opt}} = h\left(\frac{1}{\text{error}}\right)^{1/(p+1)} \tag{58}$$

with

$$\text{error} = \left\| \left[\frac{e_k}{atol_k + rtol_k \max(|x_{0,k}|, |x_k|)}, \; k \in 1 \ldots n \right] \right\|. \tag{59}$$

Usually, the step size is not allowed to decrease or increase too fast by means of the following heuristic rule

$$h_{\text{new}} = h \min\left(\alpha_{\max}, \max\left(\alpha_{\min}, \alpha\left(\frac{1}{\text{error}}\right)^{1/(p+1)}\right)\right) \tag{60}$$

where $\alpha, \alpha_{\min}$ and $\alpha_{\max}$ are some user parameters.

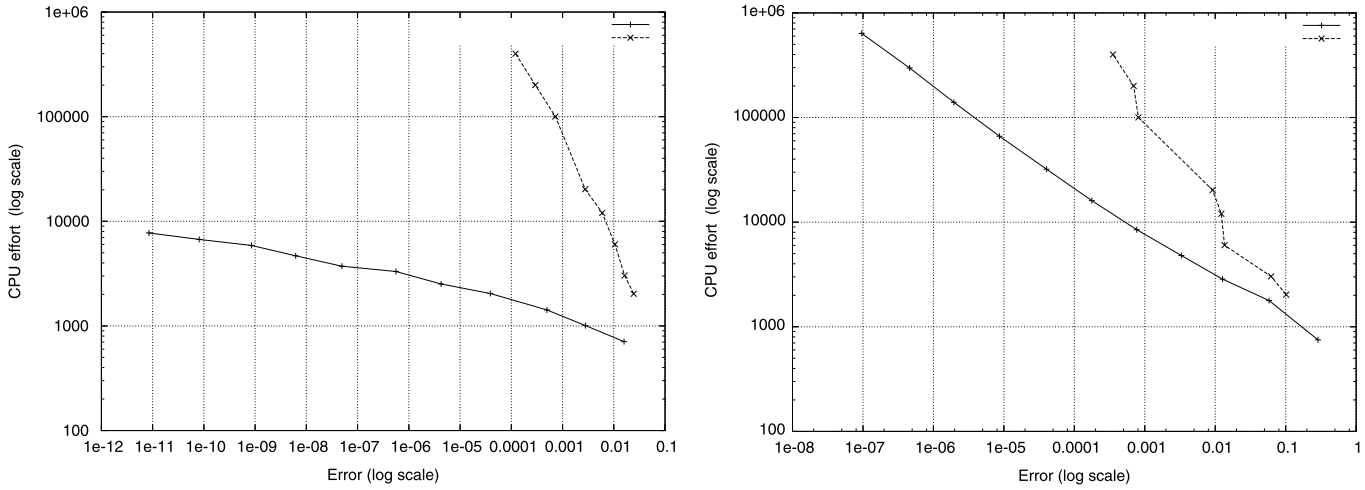
4.2. What can be done in nonsmooth situation?

It seems to be obvious that an implementation which relies on the local error in velocity e_v is not a good idea since $e_v = \mathcal{O}(1)$ in practical situations. Indeed, there is no chance that the error in $v(\cdot)$ will reduce with the time-step h.

Let us consider the estimate on the coordinates $e_q = \mathcal{O}(h)$. In this case, the error e_q will be proportional to h for sufficiently small h. Unfortunately, for such an estimate we cannot apply standard error based on an extrapolation with halved time-steps mainly because $p = 0$ which reflects that fact that the error is not smoothly transported. The proposed approach is heuristic and is only justified in several examples presented in this paper. Clearly, a more thorough study is needed to confirm the good behavior of the approach on more intricate systems. We propose to evaluate the error by

$$e_{q,1/2} \approx (q_{1/2} - q_1). \tag{61}$$

This practical estimation of the error recalls the oldest device of Runge cited by [20, page 164]. For Examples 1 (the bouncing ball) and 2 (the linear oscillator), the results are depicted on Fig. 3 where the CPU effort is plotted with respect to the global error $\|e\|_1$. The CPU effort is compared to the error with respect to the analytical solution. The CPU effort is counted as the



(a) The bouncing ball example (b) The linear oscillator example

Fig. 3. Precision work diagram for Moreau's time-stepping scheme with the heuristic (61). + adaptive time-step strategy, × constant time-steps.

total number of numerical evaluations of the right-hand side of (16). The parameters of the automatic step size control are $\alpha_{\min} = 0.5$, $\alpha_{\max} = 5$ and $\alpha = 1.0$ and the simulation is performed for $T = 5$.

The preliminary conclusion is that we are able to reach reasonable accuracy without a huge CPU effort. However, this preliminary result has to be confirmed by a more thorough study on more complex examples.

Another attempt has been made, based on the inspection of the possible events in the time-interval. If an event is present, the local order of accuracy of the scheme is given by Proposition 1. If there is no event in the time-interval, the local order of accuracy of the scheme is chosen as the same as the backward Euler scheme for index-2 DAE. This approach which is detailed in [2] improves a little bit the standard adaptive time-step strategy.

5. Higher order event-capturing time-stepping scheme

In this section, we propose a method to design a higher order time-stepping scheme which extends an original idea of Mannshardt [25] for ordinary differential equations with discontinuities. The key idea of such schemes is based on a rough localization of the impact or events such as activations or deactivations of constraints into the so-called *critical time-step*. Choosing a method of order p with a time step h for integrating the smooth dynamics, the integration over the critical time-step is performed with a method of order q. The length of the critical time-step denoted by $\bar{h}$ is chosen such that $\bar{h}^{q+1} = \mathcal{O}(h^{p+1})$.

5.1. Integration of the smooth dynamics

Mainly for the sake of simplicity, the numerical integration over a smooth period is made with a Runge–Kutta (RK) method on the following index-1 DAE,

$$\begin{cases} M\big(q(t)\big)\dot{v}(t) = F\big(t, q(t), v(t)\big) + G(q)\lambda(t), \\ \dot{q}(t) = v(t), \\ \gamma(t) = G\big(q(t)\big)\dot{v}(t) = 0. \end{cases} \tag{62}$$

In practice, the time-integration is performed for the following system

$$\begin{cases} M\big(q(t)\big)\dot{v}(t) = F\big(t, q(t), v(t)\big) + G(q)\lambda(t), \\ \dot{q}(t) = v(t), \\ 0 \leqslant \gamma(t) = G\big(q(t)\big)\dot{v}(t) \perp \lambda(t) \geqslant 0 \end{cases} \tag{63}$$

on the time-interval I where the index set $\mathcal{I}(t)$ of active constraints is assumed to be constant on I and $\lambda(t) > 0$ for all $t \in I$. Using (63) rather than (62) allows one to detect events if the acceleration γ becomes nonnegative and/or if the multiplier λ

vanishes. Using the standard notation for the RK methods (see [20] for details), the complementarity problem that we have to solve at each time-step reads

$$\begin{cases} t_{ki} = t_k + c_i h, \\ v_{k+1} = v_k + h \sum_{i=1}^{s} b_i V'_{ki}, \\ q_{k+1} = q_k + h \sum_{i=1}^{s} b_i V_{ki}, \\ V'_{ki} = M^{-1}(Q_{ki})\big[F(t_{ki}, Q_{ki}, V_{ki}) + G(Q_{ki})\lambda_{ki}\big], \\ V_{ki} = v_k + h \sum_{j=1}^{s} a_{ij} V'_{nj}, \\ Q_{ki} = q_k + h \sum_{j=1}^{s} a_{ij} V_{nj}, \\ 0 \leqslant \gamma_{ki} = G(Q_{ki}) V'_{ki} \perp \lambda_{ki} \geqslant 0. \end{cases} \tag{64}$$

Assumption 5. Let I be a smooth period time-interval (see Definition 1). We assume that

1. the local order of the RK method (64) is p that is

$$e_q = e_v = \mathcal{O}\big(h^{p+1}\big), \tag{65}$$

2. starting from inconsistent initial value $\tilde{q}_k$ such that $\tilde{q}_k - q_k = \mathcal{O}(h^{p+1})$, the error made by the RK method (64) is

$$\tilde{q}_{k+1} - q_{k+1} = \mathcal{O}\big(h^{p+1}\big). \tag{66}$$

Assumption 5.1 is ensured by the result in [18, Theorem 1.1, Sec. IV.1]. Assumption 5.2 should be obtained by extending results of [19] on the convergence of RK methods and the influence of perturbations. Rather than giving a long textual explanation of an implementation of such a scheme, we propose to outline the main features of the method in Algorithm 1.

Algorithm 1 Higher order event-capturing time-stepping scheme for the step k.

Require: DAE solver of order p for the smooth index-1 dynamics with bilateral constraints (62)
Require: h time-step, $I = [t_k, t_{k+1}]$
Require: q_k, v_k initial conditions of the step
Require: $C \in \mathbb{R}_+$, user defined positive constant
  // Initialization
  $i \leftarrow 0$
  $I_{k,i} \leftarrow [t_k, t_{k+1}], t_{k,i} \leftarrow t_k, t_{k,N} \leftarrow t_{k+1}$
  $q_{k,i} \leftarrow q_k, v_{k,i} \leftarrow v_k$
  Compute $\mathcal{I}_{k,i} = \{\alpha \in 1 \dots m \mid g^\alpha(q_{k,i}) \leqslant 0\}$ the index set of active constraints at time $t_{k,i}$
  // Integration
  while $t_{k,i} < t_{k+1}$ **do**
    Compute q_{k+1}, v_{k+1} by the numerical integration over I_i of (62) with a DAE solver of order p and a constant index set $\mathcal{I}_i$
    Compute $\mathcal{I}_{k+1} = \{\alpha \in 1 \dots m \mid g^\alpha(q_{k+1}) \leqslant 0\}$
    if $\mathcal{I}_i \neq \mathcal{I}_{k+1}$ **then**
      Locate roughly the first event time $t_{i,*}$ such that

$$t_{i,*} \in [t_{k,i+1}, t_{k,i+2}], \quad |t_{k,i+2} - t_{k,i+1}| \leqslant Ch^{p+1}, \ t_{k,i+2} \leqslant t_{k+1}$$

      // Smooth time integration
      Compute $q_{k,i+1}, v_{k,i+1}$ by the numerical integration over $[t_{k,i}, t_{k,i+1}]$ of (62) with a DAE solver of order p and a constant $\mathcal{I}_i$
      // Nonsmooth time integration
      Compute $q_{k,i+2}, v_{k,i+2}$ by the numerical integration over $[t_{k,i+1}, t_{k,i+2}]$ of (13) with Moreau's time stepping scheme
      $i \leftarrow i + 2$
      Update $\mathcal{I}_{k,i} = \{\alpha \in 1 \dots m \mid g^\alpha(q_{k,i}) \leqslant 0\}$
      $I_i \leftarrow [t_{k,i}, t_{k+1}]$
    end if
  end while

The local error on the whole time-step is given by the following theorem

Theorem 1. *Let us assume that Assumptions 1, 2 and 5 hold. The local error of consistency of the scheme described by Algorithm 1 is of order p in the generalized coordinates that is*

$$e_q = \mathcal{O}\big(h^{p+1}\big). \tag{67}$$

Proof. If there is no event in the interval $I_k = [t_k, t_{k+1}]$, the result is trivial. Otherwise, the proof is given by induction on the finite sequence of the time intervals $[t_{k,i}, t_{k,i+1}]$ generated by Algorithm 1. Let us assume that the error by the scheme up to the time $t_{k,i}$ is

$$e_{k,i} = q_{k,i} - q(t_{k,i}) = \mathcal{O}\big(h^{p+1}\big). \tag{68}$$

We denote by $\tilde{q}_{k,i+1}$ and $\tilde{v}_{k,i+1}$ the values obtained by the time integration method on $[t_{k,i}, t_{k,i+1}]$ starting from the exact values $q(t_{k,i})$ and $v(t_{k,i})$. The time-integration on the time-interval $[t_{k,i}, t_{k,i+1}]$ is performed by one of the following schemes:

1. *Moreau's time stepping scheme.* We have in this case that

$$e_{k,i+1} = q_{k,i+1} - q(t_{k,i+1}) = q_{k,i+1} - \tilde{q}_{k,i+1} + \tilde{q}_{k,i+1} - q(t_{k,i+1}). \tag{69}$$

According to Proposition 1, we get $q(t_{k,i+1}) - \tilde{q}_{k,i+1} = \mathcal{O}(\bar{h})$. For the remaining part of the error, we can write

$$\begin{aligned} e_{k,i+1} &= q_{k,i} + \bar{h}\big(\theta v_{k,i+1} + (1-\theta) v_{k,i}\big) - \big(q(t_{k,i}) + \bar{h}\big(\theta \tilde{v}_{k,i+1} + (1-\theta)\tilde{v}_{k,i}\big)\big) + \mathcal{O}(\bar{h}), \\ &= e_{k,i} + \bar{h}\big(\theta(v_{k,i+1} - \tilde{v}_{k,i+1}) + (1-\theta)(v_{k,i} - \tilde{v}_{k,i})\big) + \mathcal{O}(\bar{h}), \\ &= \mathcal{O}\big(h^{p+1}\big) + \mathcal{O}(\bar{h}) = \mathcal{O}\big(h^{p+1}\big). \end{aligned} \tag{70}$$

2. *Index-1 DAE solver* (64). We have in this case that

$$e_{k,i+1} = q_{k,i+1} - q(t_{k,i+1}) = q_{k,i+1} - \tilde{q}_{k,i+1} + \tilde{q}_{k,i+1} - q(t_{k,i+1}). \tag{71}$$

Using Assumption 5, we obtain $e_{k,i+1} = \mathcal{O}(h^{p+1})$.

By induction on the finite number of time-intervals inside $[t_k, t_{k+1}]$ the result is obtained. □

In practice, a lower bound for $|t_{k,i+2} - t_{k,i+1}|$ is added and chosen with respect to the machine accuracy.

Remark 5. There are certain similarities between the method in [41,40] and the proposed approach. Higher order accuracy is sought by roughly locating the events inside a critical time-step which is also integrated by Moreau's scheme. Nevertheless, important differences can be pointed out. Firstly, there is no proof of the order of accuracy of the Moreau scheme in [41,40]. Consequently, the length of the critical time-step is set *a priori* by the user. In this way, the expected higher order accuracy on the global error is not ensured. Secondly, extrapolation methods are used over smooth phases. They are based on Moreau's scheme with an impact rule given by (8). This leads to several difficulties which require the modification of standard extrapolation methods, as well as the value of the coefficient of restitution, in order to retrieve consistency. In our approach, standard IRK codes for index-1 DAE can be used. Standard BDF methods might be also considered even though frequent restarts after an event would penalize the efficiency.

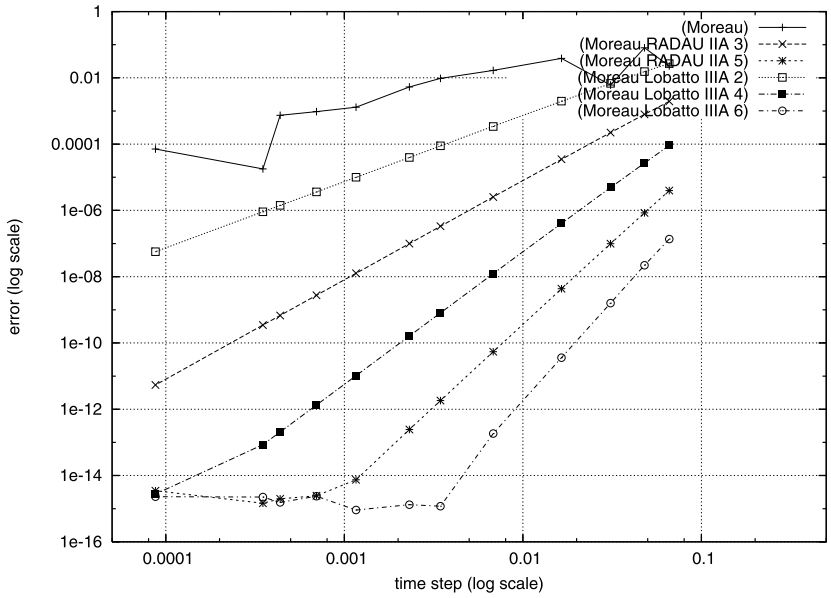
5.2. Numerical applications

Theorem 1 is illustrated on the benchmark Examples 1 and 2 in Fig. 4 for standard implicit RK methods. The implicit RK methods are the well-known RADAU IIA methods of order 3 and 5 and Lobatto IIIA methods of order 2, 4 and 6. Details on these methods can be found in [20,18]. Similar results for half-explicit RK methods can be found in Fig. 5.

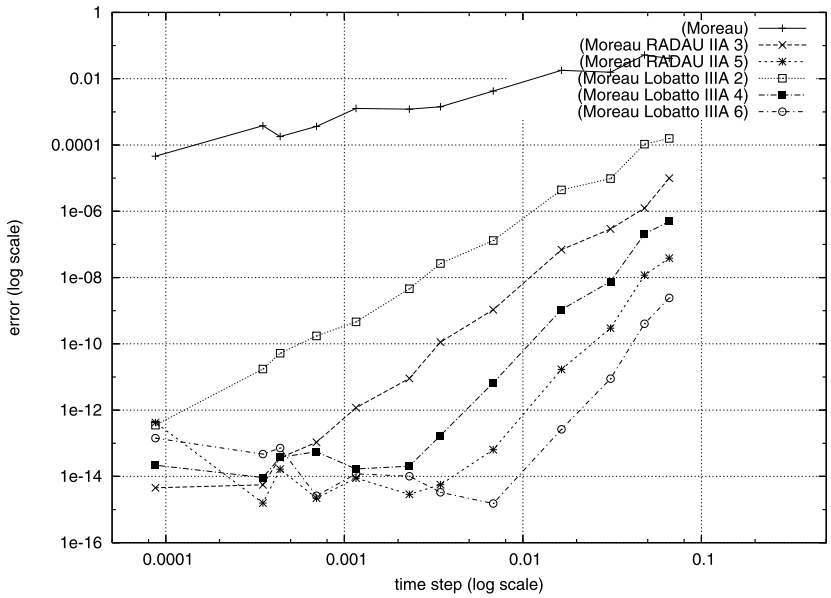
Fig. 4 presents the global error in generalized coordinates with respect to the time-step. The conclusions are quite encouraging since we retrieve the global order of the associated RK methods even in presence of finite accumulation of impacts.

6. Conclusions and perspectives

In this paper, several approaches have been proposed to improve the resolution (ratio computational cost/error) and the order of accuracy of the Moreau time-stepping scheme. As far as we know, the results on the order of consistency are original. Unfortunately, the estimates on the accuracy of the method, which are very low and attained on very simple examples, do not allow the use of sophisticated variable time-step strategies. This is one of the main motivations to design



(a) The linear oscillator example



(b) The bouncing ball example

Fig. 4. Order of accuracy. Implicit Runge–Kutta methods coupled with Moreau's time-stepping scheme.

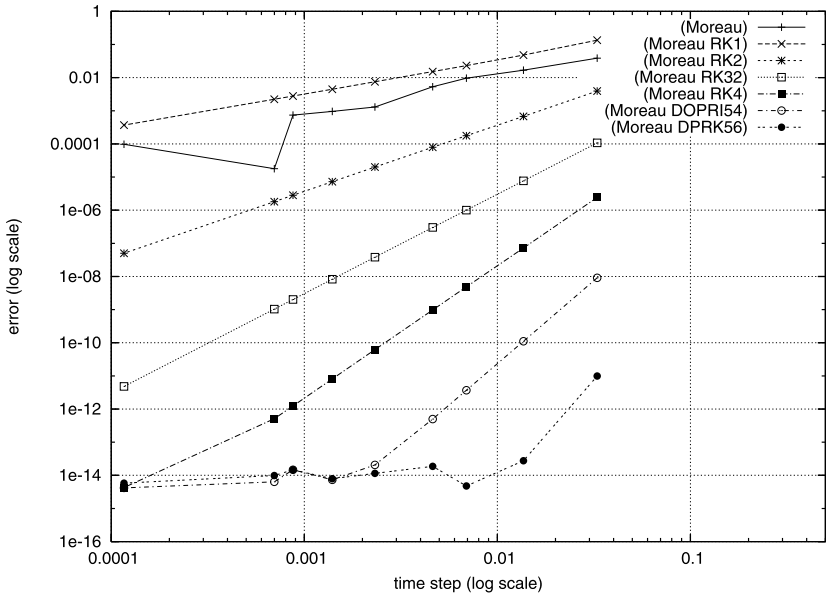
higher order event-capturing schemes. A first attempt to build such a scheme is proposed in this paper by coupling standard Runge–Kutta scheme for DAE with Moreau's scheme. This new scheme behaves well in simple academic examples and has to be tested on intricate nonlinear multibody systems.

The perspectives for this work are to improve the theoretical framework for the proposed scheme. For instance, Assumption 5 should be improved or proved by standard arguments. The use of index-2 DAE form in smooth periods should also be considered together with the possible study of order reduction due to propagation of error in the multiplier [6,7].

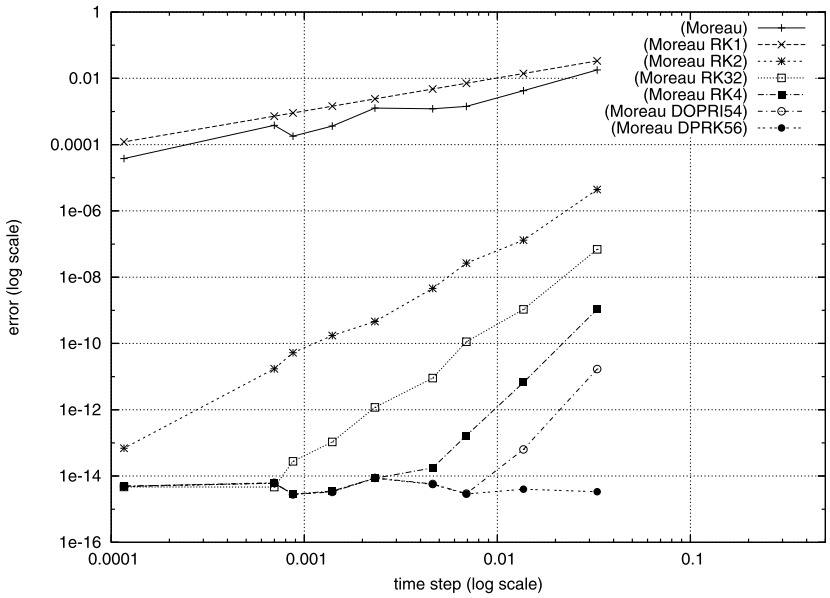
The case of Coulomb's friction is originally included in the seminal work of Moreau [30] and convergence was proven by [26,38] under certain assumptions. Formally, Coulomb's friction may also be treated with the proposed scheme without any major technical difficulties. Finally, the main open issue for higher order time-stepping schemes is to show whether the global error can be extrapolated so that we can efficiently use standard adaptive time-step strategies.

Acknowledgements

This work has been supported by the French National Research Agency (ANR) trough COSINUS program (project SALADYN ANR-08-COSI-014) http://saladyn.inria.gforge.fr/. The author is indebted to the reviewers for their remarks and comments which help a lot to improve the contents.



(a) The linear oscillator example



(b) The Bouncing Ball example

Fig. 5. Order of accuracy. Half-explicit Runge–Kutta methods coupled with Moreau's time-stepping scheme.

References

[1] M. Abadie, Dynamic simulation of rigid bodies: Modelling of frictional contact, in: B. Brogliato (Ed.), Impacts in Mechanical Systems: Analysis and Modelling, in: Lecture Notes in Phys., vol. 551, Springer, 2000, pp. 61–144.

[2] V. Acary, Toward higher order event-capturing schemes and adaptive time-step strategies for nonsmooth multibody systems, Technical report, INRIA, 2009.

[3] V. Acary, B. Brogliato, Numerical Methods for Nonsmooth Dynamical Systems: Applications in Mechanics and Electronics, Lect. Notes Appl. Comput. Mech., vol. 35, Springer Verlag, 2008.

[4] V. Acary, B. Brogliato, Implicit Euler numerical scheme and chattering-free implementation of sliding mode systems, Systems Control Lett. 59 (5) (May 2010) 284–293, http://dx.doi.org/10.1016/j.sysconle.2010.03.002.

[5] V. Acary, C.I. Morarescu, F. Pérignon, B. Brogliato, Numerical simulation of nonsmooth systems and switching control with the siconos/control toolbox, in: 6th Euromech Nonlinear Dynamics Conference, ENOC 2008, St. Petersburg, 29 June 2008.

[6] M. Arnold, A perturbation analysis for the dynamical simulation of mechanical multibody systems, Appl. Numer. Math. 18 (1995) 37–56.

[7] M. Arnold, Numerical methods for simulation in applied dynamics, in: Arnold, and Schiehlen [8], pp. 191–246.

[8] M. Arnold, W. Schiehlen (Eds.), Simulation Techniques in Applied Dynamics, CISM Courses and Lectures, vol. 507, Springer, 2008.

[9] P. Ballard, The dynamics of discrete mechanical systems with perfect unilateral constraints, Arch. Ration. Mech. Anal. 154 (2000) 199–274.

[10] P. Ballard, Bounded variations and measure theory on the line, Lectures Notes of the Second Summer on Nonsmooth Dynamics Held in Autrans, France, 2003.

[11] K.E. Brenan, S. Campbell, L.R. Petzold, Numerical Solution of Initial-Value Problems in Differential-Algebraic Equations, North-Holland, 1989.

[12] B. Brogliato, Nonsmooth Mechanics: Models, Dynamics and Control, 2nd edition, Springer Verlag, London, 1999.

[13] O. Brüls, A. Cardona, M. Géradin, Modeling, simulation and control of flexible multibody systems, in: Arnold, and Schiehlen [8], pp. 21–74.
[14] R. Dzonou, M.D.P. Monteiro Marques, A sweeping process approach to inelastic contact problems with general inertia operators, Eur. J. Mech. A Solids 26 (3) (2007) 474–490.
[15] C.W. Gear, B. Leimkuhler, G.K. Gupta, Automatic integration of Euler–Lagrange equations with constraints, J. Comput. Appl. Math. 12–13 (1985) 77–90.
[16] Ch. Glocker, E. Cataldi-Spinola, R.I. Leine, Curve squealing of trains: Measurement, modelling and simulation, J. Sound Vibration 324 (1–2) (2009) 365–386.
[17] M. Géradin, A. Cardona, Flexible Multibody Dynamics: A finite Element Approach, J. Wiley & Sons, New York, 2001, 340 p.
[18] E. Hairer, G. Wanner, Solving Ordinary Differential Equations II. Stiff and Differential-Algebraic Problems, Springer, 1996.
[19] E. Hairer, Ch. Lubich, M. Roche, The Numerical Solution of Differential-Algebraic Systems by Runge–Kutta Methods, Springer Verlag, 1987.
[20] E. Hairer, S.P. Norsett, G. Wanner, Solving Ordinary Differential Equations I. Nonstiff Problems, Springer, 1993.
[21] Andrei Herdt, Holger Diedam, Pierre-Brice Wieber, Dimitar Dimitrov, Katja Mombaur, Moritz Diehl, Online walking motion generation with automatic foot step placement, Advanced Robotics 24 (5–6) (2010) 719–737, Brill Academic Publishers, Utrecht, http://hal.inria.fr/inria-00391408/en/.
[22] O. Janin, C.H. Lamarque, Comparison of several numerical methods for mechanical systems with impacts, Internat. J. Numer. Methods Engrg. 51 (9) (2001) 1101–1132.
[23] P. Lötstedt, L.R. Petzold, Numerical solution of nonlinear algebraic equations with algebraic constraints: I – convergence results for backward differentiation formulas, Math. Comp. 46 (174) (1986) 491–516.
[24] M. Mabrouk, A unified variational for the dynamics of perfect unilateral constraints, Eur. J. Mech. A Solids 17 (1998) 819–842.
[25] R. Mannshardt, One-step methods of any order for ordinary differential equations with discontinuous right-hand sides, Numer. Math. 31 (1978) 131–152.
[26] M.D.P. Monteiro Marques, Differential Inclusions in Nonsmooth Mechanical Problems. Shocks and Dry Friction, Progr. Nonlinear Differential Equations Appl., vol. 9, Birkhäuser, Basel, 1993.
[27] J.J. Moreau, Approximation en graphe d'une évolution discontinue, RAIRO, Anal. Numér. 12 (1978) 75–84.
[28] J.J. Moreau, Fonctionnelles Convexes. Séminaire sur les équations aux dérivées partielles, subventionné par le CNRS, Collège de France, Paris, 1967.
[29] J.J. Moreau, Liaisons unilatérales sans frottement et chocs inélastiques, C. R. Acad. Sci. Ser. II 296 (1983) 1473–1476.
[30] J.J. Moreau, Unilateral contact and dry friction in finite freedom dynamics, in: J.J. Moreau, P.D. Panagiotopoulos (Eds.), Nonsmooth Mechanics and Applications, in: CISM Courses and Lectures, vol. 302, Springer Verlag, Wien, New York, 1988, pp. 1–82.
[31] J.J. Moreau, Bounded variation in time, in: J.J. Moreau, P.D. Panagiotopoulos, G. Strang (Eds.), Topics in Nonsmooth Mechanics, Birkhäuser, Basel, 1988, pp. 1–74.
[32] J.J. Moreau, Numerical aspects of the sweeping process, in: J.A.C. Martins, A. Klarbring (Eds.), Special Issue on Computational Modeling of Contact and Friction, Comput. Methods Appl. Mech. Engrg. 177 (1999) 329–349.
[33] L. Paoli, M. Schatzman, A numerical scheme for impact problems I: The one-dimensional case, SIAM J. Numer. Anal. 40 (2) (2002) 702–733.
[34] L. Paoli, M. Schatzman, A numerical scheme for impact problems II: The multi-dimensional case, SIAM J. Numer. Anal. 40 (2) (2002) 734–768.
[35] F. Pfeiffer, C. Glocker, Multibody Dynamics with Unilateral Contacts. Non-linear Dynamics, J. Wiley & Sons, 1996.
[36] R.T. Rockafellar, Convex Analysis, Princeton University Press, 1970.
[37] M. Schatzman, A class of nonlinear differential equations of second order in time, Nonlinear Anal. 2 (3) (1978) 355–373.
[38] D. Stewart, Convergence of a time-stepping scheme for rigid-body dynamics and resolution of Painlevé's problem, Arch. Ration. Mech. Anal. 145 (1998) 215–260.
[39] D. Stewart, Rigid body dynamics with friction and impact, SIAM Rev. 42 (1) (2000) 3–39.
[40] C. Studer, Numerics of Unilateral Contacts and Friction. Modeling and Numerical Time Integration in Non-Smooth Dynamics, Lect. Notes Appl. Comput. Mech., vol. 47, Springer Verlag, 2009.
[41] C. Studer, R.I. Leine, Ch. Glocker, Step size adjustment and extrapolation for time stepping schemes in non-smooth dynamics, Internat. J. Numer. Methods Engrg. 76 (11) (2008) 1747–1781.