



HAL
open science

Step size adjustment and extrapolation for time-stepping schemes in non-smooth dynamics

Christian Studer, Remco Leine, Christoph Glocker

► **To cite this version:**

Christian Studer, Remco Leine, Christoph Glocker. Step size adjustment and extrapolation for time-stepping schemes in non-smooth dynamics. *International Journal for Numerical Methods in Engineering*, 2008, 76 (11), 10.1002/nme.2383 . hal-01352897

HAL Id: hal-01352897

<https://hal.science/hal-01352897>

Submitted on 10 Aug 2016

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.



Distributed under a Creative Commons Attribution 4.0 International License

Step size adjustment and extrapolation for time-stepping schemes in non-smooth dynamics

C. Studer, R. I. Leine and Ch. Glocker

*Institute of Mechanical Systems, Center of Mechanics, Department of Mechanical and Process Engineering,
ETH Zurich, CH-8092 Zurich, Switzerland*

SUMMARY

In this paper we use step size adjustment and extrapolation methods to improve Moreau's time-stepping scheme for the numerical integration of non-smooth mechanical systems, i.e. systems with impact and friction. The scheme yields a system of inclusions, which is transformed into a system of projective equations. These equations are solved iteratively. Switching points are time instants for which the structure of the mechanical system changes, for example, time instants for which a sticking friction element begins to slide. We show how switching points can be localized and how these points can be resolved by choosing a minimal step size. In order to improve the integration of non-smooth systems in the smooth parts, we show how the time-stepping method can be used as a base integration scheme for extrapolation methods, which allow for an increase in the integration order. Switching points are processed by a small time step, while time intervals during which the structure of the system does not change are computed with a larger step size and improved integration order. The overall algorithm, which consists of a time-stepping module, an extrapolation module and a step size adjustment module, is discussed in detail and some examples are given.

KEY WORDS: non-smooth dynamics; contact; time-stepping integration; step size adjustment; extrapolation; higher order

1. INTRODUCTION

The modelling of mechanical systems with unilateral contacts and friction is a challenging task, which is studied in the field of elastic as well as rigid body mechanics. An overwhelming number of methods and approaches exist, of which a complete overview is outside the scope of this paper. For contacts between elastic bodies within a finite element formulation, we refer to the works of Simo and Laursen [1], Laursen [2], Klarbring [3] and Wriggers [4]. Contact between elastic

bodies are not the subject of this paper. However, the methods used in elastic and rigid body contact mechanics are very similar. For contact between rigid bodies, we refer to Brogliato [5], Brogliato *et al.* [6] and Eich-Soellner and Führer [7]. Two main approaches for the modelling of a rigid body system with unilateral contacts and friction exist. One possibility is to regularize the non-smooth force laws that are associated with the unilateral contact and friction, i.e. the contact and the friction forces are defined as a function of the penetration depth or volume and the sliding velocity, respectively, see [8–11]. Such an approach leads to an ordinary differential equation system, which can be simulated straightforwardly using a classical numerical scheme for stiff systems, for example, an implicit Runge–Kutta scheme. The regularization of non-smooth force laws leads to many numerical problems, i.e. stiff differential equations and spurious oscillations due to the stiffness of the contacts. Furthermore, the determination of suitable regularization parameters can pose a problem. Another approach to handle mechanical systems with unilateral contacts and friction is the non-smooth approach, which was originated by Moreau (see [12–14]). Moreau suggested to describe the dynamics using differential measures and to describe the force laws for unilateral frictional contact by set-valued force laws. The dynamics is therefore expressed in Moreau’s approach by measure differential inclusions. The time evolution of the velocities is in this approach not required to be continuous or smooth (but only required to be of locally bounded variation). As a consequence, standard numerical integration techniques cannot be used anymore. Instead, time-stepping schemes are used, which are essentially time discretizations of the measure differential inclusion. Time-stepping schemes, which are used in non-smooth dynamics, are not comparable to standard integration schemes such as Runge–Kutta, Newmark or Galerkin methods. In this paper we restrict ourselves to the non-smooth approach, where we consider only hard contact between rigid bodies. As integrator we use Moreau’s midpoint rule [13], being a time-stepping scheme that has proven to be very robust. Of course, other well working numerical integration schemes for non-smooth systems exist; see [13, 15–26]. The aim of this paper is to improve the accuracy of Moreau’s integration scheme by using step size adjustment and extrapolation [27, 28].

The constitutive behaviour of unilateral frictional contacts is described in the non-smooth approach by set-valued force laws. The impenetrability of the contact and unilaterality of a normal component of the contact force is described by Signorini’s law for normal contact, which is a complementarity inequality relationship between contact distance and normal contact force. Friction is described by Coulomb’s law for dry friction, being a set-valued force law which relates the friction force to the sliding velocity and normal contact force. A frictional unilateral contact is therefore described by *two* set-valued force laws in our approach.

The time evolution of a non-smooth mechanical system can be divided into different ‘smooth parts’, which are time intervals for which the set-valued force laws do not change their mode of operation and the structure of the mechanical system does not change. This implies that closed contacts remain closed and sticking contacts remain sticking during smooth parts of the motion. The so-called switching points separate different smooth parts of the time evolution. In order to administer the operation mode of each set-valued force law, we introduce a state vector $\sigma = \sigma(t)$, which has an entry σ_i for each set-valued law. At switching time instants t_S , the state σ_i of at least one set-valued law changes, i.e. $\sigma^-(t_S) \neq \sigma^+(t_S)$. A straightforward way to simulate non-smooth systems is the event-driven approach [21, 29, 30]. Event-driven methods integrate the differential equation system belonging to a smooth part of the motion with an arbitrary ODE or DAE integrator, e.g. Runge–Kutta, Newmark or Galerkin methods. When encountering a switching time instant t_S , the new state $\sigma^+(t_S)$ is determined in order to set up a new differential equation system with possible

new initial conditions for further integration of the following smooth part. Event-driven methods clearly distinguish between smooth integration and evaluation of impact and contact behaviour at the switching points. The methods have been successfully implemented for systems with few contacts. Event-driven methods are very accurate but bring a lot of administrative effort and are not suited for accumulative switching points, because the methods aim at resolving all switching points. Such an accumulative switching point occurs, for example, in the classical bouncing ball system [5]. After each dissipative impact on the table, the ball will jump less high and the flight time becomes less long. There will be an accumulation of switching points, i.e. infinitely many impacts in a finite time, by the end of the ball's movement. Furthermore, event-driven methods request for each state σ a corresponding differential equation system, which becomes a problem if the number of set-valued laws is large.

In contrast to event-driven approaches, time-stepping integration methods are well suited for systems with many set-valued force laws and accumulative switching points, see [13, 15, 17–26]. The methods determine for each time step a discrete state $\hat{\sigma}$, which applies for the *entire* time step. In each time step the algorithm decides whether the whole time step is computed using the assumption of a closed contact or an open contact and, if the contact is closed, whether it is sticking or slipping. If the step size becomes small, then the evolution of the discrete state $\hat{\sigma}$ and of the state σ becomes similar. Time-stepping methods with a constant step size do not distinguish between the different smooth parts and switching points, and are therefore very robust and simple. The methods are well suited for accumulative switching points, and simulations up to several thousand contacts are possible. As a drawback, time-stepping methods require a very small time step size, and the accuracy is less satisfactory. We will restrict ourselves to the time-stepping method of Moreau, see [13, 19, 26]. To conclude, event-driven methods are very well suited for systems with few contacts and long smooth parts of the motion, while time-stepping schemes have advantages for systems with many contacts and many switching points. Our goal is to improve the time-stepping method such that it enjoys some of the advantages of the event-driven scheme [27, 28]. We choose the time-stepping scheme of Moreau as base scheme, because it can handle systems with many as well as with few contacts, the latter with an accuracy which is worse compared with an event-driven scheme. Similar to the event-driven approach we distinguish between smooth parts and switching points, which results in a step-size-controlled time-stepping algorithm. The step size of time steps in which the system switches must be chosen very small in order to resolve the switching point properly, while the step size of time steps without switching points can be taken much larger. Furthermore, an extrapolation method based on the time-stepping scheme can be used to increase the order of integration in the smooth parts. The resulting method can handle systems with few contacts quite well by refining the step size at switching points and by using extrapolation in the smooth parts of the motion. In addition, systems with many contacts can be handled by the method because the underlying algorithm is still a time-stepping scheme.

2. NON-SMOOTH DYNAMICS

In this section we briefly discuss a possible mathematical formulation of non-smooth mechanical systems. Detailed information is available in [26, 28, 29, 31, 32]. In this paper we use set-valued laws to describe non-smooth interactions between rigid bodies. A unilateral contact is described by a set-valued law of the kind ‘Signorini’s law’, while a friction element is characterized by

a set-valued law of the kind ‘Coulomb’s law for dry friction’. We distinguish between active set-valued force laws, which can be expressed on the velocity level, i.e. set-valued force laws associated with a closed unilateral contact or a friction element, and non-active set-valued force laws, which describe, for example, an open unilateral contact. An active set-valued force law i can be described on the velocity level by an inclusion of the form

$$-\boldsymbol{\gamma}_i \in \mathcal{N}_{\mathcal{C}_i}(\boldsymbol{\lambda}_i) \quad (1)$$

The relative velocities of the i th set-valued law are denoted by $\boldsymbol{\gamma}_i$, and the forces by $\boldsymbol{\lambda}_i$, respectively. The convex set \mathcal{C}_i is the set of all admissible forces $\boldsymbol{\lambda}_i$, whereas \mathcal{C}_i might also be state dependent in case of friction. Note in this context that a unilateral frictional contact is modelled by *two* set-valued force laws, one modelling the unilateral contact and the other characterizing the friction element. Non-active set-valued force laws, i.e. open unilateral contacts, are neglected because it is assumed that their force $\boldsymbol{\lambda}_i$ vanishes.

Example

A unilateral contact has a relative velocity $\boldsymbol{\gamma}_i$ and force $\boldsymbol{\lambda}_i$ pointing in the normal contact-surface direction, and the convex set \mathcal{C}_i is the set of all positive real numbers \mathbb{R}_0^+ . The behaviour for a closed unilateral contact on the velocity level yields

$$-\boldsymbol{\gamma}_i \in \mathcal{N}_{\mathbb{R}_0^+}(\boldsymbol{\lambda}_i) \Leftrightarrow \boldsymbol{\lambda}_i \geq 0, \quad \boldsymbol{\gamma}_i \geq 0, \quad \boldsymbol{\lambda}_i \boldsymbol{\gamma}_i = 0 \quad (2)$$

Either the closed unilateral contact is opening with $\boldsymbol{\lambda}_i = 0$ and $\boldsymbol{\gamma}_i > 0$ or the closed unilateral contact stays closed, i.e. $\boldsymbol{\lambda}_i \geq 0$, $\boldsymbol{\gamma}_i = 0$.

The active set-valued force laws (1) are linked to the equations of motion by Lagrangian multipliers (Lagrange I formulation),

$$\mathbf{M}\dot{\mathbf{u}} - \mathbf{h} - \sum_{i=1}^n \mathbf{W}_i \boldsymbol{\lambda}_i = \mathbf{0} \quad (3)$$

$$-\boldsymbol{\gamma}_i \in \mathcal{N}_{\mathcal{C}_i}(\boldsymbol{\lambda}_i) \quad (4)$$

in which the mass matrix is denoted by $\mathbf{M} = \mathbf{M}(\mathbf{q}, t)$ and the vector of all external and gyroscopic forces is denoted by $\mathbf{h} = \mathbf{h}(\mathbf{q}, \mathbf{u}, t)$. This vector also includes all spring and dashpot forces. The vectors \mathbf{q} and $\mathbf{u} = \dot{\mathbf{q}}$ denote the generalized coordinates and the generalized velocities, respectively. The matrix $\mathbf{W}_i = \mathbf{W}_i(\mathbf{q}, t)$ and the Lagrangian multiplier $\boldsymbol{\lambda}_i$ represent the generalized force direction and the force $\boldsymbol{\lambda}_i$ of the i th active set-valued law. The relative velocity $\boldsymbol{\gamma}_i$ is linked to the generalized velocities \mathbf{u} by $\boldsymbol{\gamma}_i = \mathbf{W}_i^\top \mathbf{u} + \boldsymbol{\zeta}_i$ with $\boldsymbol{\zeta}_i = \boldsymbol{\zeta}_i(\mathbf{q}, t)$. In the case of a unilateral contact with gap function $g_i = g_i(\mathbf{q}, t)$ it holds that $\mathbf{W}_i = \mathbf{w}_i = (\partial g_i / \partial \mathbf{q})^\top$ and $\boldsymbol{\zeta}_i = \boldsymbol{\zeta}_i = \partial g_i / \partial t$.

In the case of an impact, the equations of motion have to be integrated over the instantaneous impact time, which yields the impact equations

$$\mathbf{M}(\mathbf{u}^+ - \mathbf{u}^-) - \sum_{i=1}^n \mathbf{W}_i \boldsymbol{\Lambda}_i = \mathbf{0} \quad (5)$$

Note that the pre- and post-impact velocities are denoted by \mathbf{u}^- and \mathbf{u}^+ , and that the integration of the forces $\boldsymbol{\lambda}_i$ lead to the impulses $\boldsymbol{\Lambda}_i$. In addition to the impact equations, the mathematical description of an impact requires an impact law. We choose Newton’s impact law, which can be expressed as

$$-(\boldsymbol{\gamma}_i^+ + \varepsilon_i \boldsymbol{\gamma}_i^-) \in \mathcal{N}_{\mathcal{D}_i}(\boldsymbol{\Lambda}_i) \quad (6)$$

The scalar ε_i denotes the restitution coefficient, and the convex set \mathcal{D}_i is the set of all admissible impulsive forces Λ_i , whereas \mathcal{D}_i might also be state dependent in the presence of friction. Note that we use the impact law (6) not only for unilateral contacts but also for friction elements, for which the set \mathcal{D}_i then contains all admissible impulsive tangential forces.

Example

As an example we consider Newton's impact law for a closed unilateral contact,

$$-(\gamma_i^+ + \varepsilon_i \gamma_i^-) \in \mathcal{N}_{\mathbb{R}_0^+}(\Lambda_i) \Leftrightarrow \Lambda_i \geq 0, \quad \gamma_i^+ + \varepsilon_i \gamma_i^- \geq 0, \quad \Lambda_i (\gamma_i^+ + \varepsilon_i \gamma_i^-) = 0 \quad (7)$$

The impact law accounts for two phenomena: an inversion of the relative velocity γ_i in the i th unilateral contact due to a positive impulsive force Λ_i acting in the same unilateral contact, and an opening of a closed unilateral contact induced by an impulsive force occurring in another unilateral contact of the mechanical system.

In what follows we discuss an example to introduce differential measures, which will be used to describe a non-smooth system. The example intends to give the reader a short idea about measures. The reader is referred to the literature on distribution and measure theory; see, for example, [33, 34].

Example

Let $x(t)$ be the function depicted in Figure 1. The function $x(t)$ has a discontinuity at time t_D , where $t_B < t_D < t_E$. The derivative $\dot{x}(t)$ of $x(t)$ is equal to $a = \text{const}$ almost everywhere. At the discontinuity $t = t_D$, the standard notion of the derivative does not exist. At this time, the function x has a left and a right limit $x(t_D)^+$ and $x(t_D)^-$, and we set $X := x(t_D)^+ - x(t_D)^-$. We aim at determining the function value $x(t_E)$ for given a , X , and $x(t_B)$. We define a measure $dx = \dot{x} dt + (x^+ - x^-) d\eta = a dt + X d\eta$ which has a Lebesgue integrable term $a dt$ and a purely atomic part $X d\eta$. The term $d\eta$ is a Dirac point measure for the discontinuity at time t_D , which means that the integral $\int_{\{t\}} d\eta$ is equal to 1 if $t = t_D$ and 0 otherwise. Calculation of $x(t_E)$ yields

$$\begin{aligned} x(t_E) &= x(t_B) + \int_{[t_B, t_E]} dx = x(t_B) + \int_{t_B}^{t_E} a dt + \int_{[t_B, t_E]} X d\eta \\ &= x(t_B) + a(t_E - t_B) + X \end{aligned} \quad (8)$$

The equations of motion and the impact equation can be combined in an equality of measures

$$\mathbf{M} du - \mathbf{h} dt - \sum_{i=1}^n \mathbf{W}_i d\mathbf{P}_i = 0 \quad (9)$$

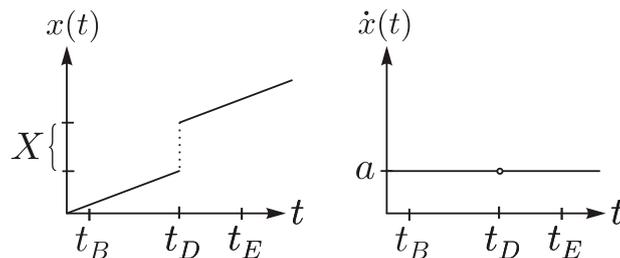


Figure 1. Function $x(t)$ with discontinuity at time t_D . The derivative $\dot{x}(t)$ of $x(t)$ is equal to a almost everywhere, except for a 'single instantaneous infinite peak' at time t_D .

in which $d\mathbf{u} = \dot{\mathbf{u}} dt + (\mathbf{u}^+ - \mathbf{u}^-) d\eta$ and $d\mathbf{P}_i = \dot{\mathbf{P}}_i dt + (\mathbf{P}_i^+ - \mathbf{P}_i^-) d\eta = \boldsymbol{\lambda}_i dt + \boldsymbol{\Lambda}_i d\eta$ are differential measures for velocity \mathbf{u} and percussion \mathbf{P}_i , respectively. The set-valued impact law (6) and the active set-valued force law (1) after integration can be combined in an inclusion

$$-\left((1 + \varepsilon_i) \boldsymbol{\gamma}_{Bi} + \int_{[t_B, t_E]} d\boldsymbol{\gamma}_i \right) \in \mathcal{N}_{\left(\int_{[t_B, t_E]} d\mathcal{A}_i \right)} \left(\int_{[t_B, t_E]} d\mathbf{P}_i \right) \quad (10)$$

in which $d\boldsymbol{\gamma}_i = \dot{\boldsymbol{\gamma}}_i dt + (\boldsymbol{\gamma}_i^+ - \boldsymbol{\gamma}_i^-) d\eta$ is a measure for the relative velocity. The set of differential measures $d\mathcal{A}_i = \mathcal{C}_i dt + \mathcal{D}_i d\eta$ must be understood as

$$d\mathcal{A}_i = \{ da_i \mid da_i = c_i dt + d_i d\eta, c_i \in \mathcal{C}_i, d_i \in \mathcal{D}_i \}, \quad \int d\mathcal{A} := \left\{ \int da \mid a \in \mathcal{A} \right\} \quad (11)$$

This approach must be seen from an engineer's point of view and is only intended to give an idea on how set-valued laws for non-impulsive motion (1) and impulsive events (6) can be combined in a single formulation, which helps to obtain Moreau's time-stepping scheme, see (16). In the case of an impact-free motion, inclusion (10) yields

$$-\left((1 + \varepsilon_i) \boldsymbol{\gamma}_{Bi} + \int_{[t_B, t_E]} \dot{\boldsymbol{\gamma}}_i dt \right) \in \mathcal{N}_{\left(\int_{[t_B, t_E]} \mathcal{C}_i dt \right)} \left(\int_{[t_B, t_E]} \boldsymbol{\lambda}_i dt \right) \quad (12)$$

If $t_E \rightarrow t_B$, then relation (12) reduces to inclusion (1), i.e.

$$-((1 + \varepsilon_i) \boldsymbol{\gamma}_{Bi} + 0) \in \mathcal{N}_{\mathcal{C}_i}(\boldsymbol{\lambda}_i) \Rightarrow -\boldsymbol{\gamma}_{Bi} \in \mathcal{N}_{\mathcal{C}_i}(\boldsymbol{\lambda}_i) \quad (13)$$

In the case of an impact at t_B , the inclusion (10) yields

$$-\left((1 + \varepsilon_i) \boldsymbol{\gamma}_{Bi}^- + \boldsymbol{\gamma}_{Ei}^+ - \boldsymbol{\gamma}_{Bi}^+ \right) = -\left(\boldsymbol{\gamma}_{Ei}^+ + \varepsilon_i \boldsymbol{\gamma}_{Bi}^- \right) \in \mathcal{N}_{\left(\int_{\{t_B\}} d\mathcal{A}_i \right)} \left(\int_{\{t_B\}} d\mathbf{P}_i \right) \equiv \mathcal{N}_{\mathcal{D}_i}(\boldsymbol{\Lambda}_i) \quad (14)$$

which can be simplified to the set-valued impact law (6).

3. TIME STEPPING

Time-stepping methods provide a discrete numerical scheme suitable for the simulation of non-smooth mechanical systems. These methods are widely used, see [13, 15, 17–26], because of their simplicity and their robustness. In contrast to event-driven methods, time-stepping methods enjoy convergence results [12, 35]. Various time-stepping methods exist, and we will restrict ourselves to a brief review of the midpoint method of Moreau, [13, 19, 26, 32]. We consider a time step with step size Δt , of which begin and end points are denoted by the indices B and E , respectively. Moreau's midpoint method calculates a midpoint $\mathbf{q}_M = \mathbf{q}_B + \mathbf{u}_B \Delta t / 2$ and discretizes the measure differential equation (9) as follows:

$$\int_{[t_B, t_E]} \mathbf{M} d\mathbf{u} \rightarrow \mathbf{M}_M (\mathbf{u}_E - \mathbf{u}_B), \quad \mathbf{M}_M = \mathbf{M} \left(\mathbf{q}_M, t_B + \frac{\Delta t}{2} \right)$$

$$\int_{t_B}^{t_E} \mathbf{h} dt \rightarrow \mathbf{h}_M \Delta t, \quad \mathbf{h}_M = \mathbf{h} \left(\mathbf{q}_M, \mathbf{u}_B, t_B + \frac{\Delta t}{2} \right)$$

$$\int_{[t_B, t_E]} \mathbf{W} d\mathbf{P}_i \rightarrow \mathbf{W}_M \hat{\mathbf{P}}_i, \quad \mathbf{W}_M = \mathbf{W} \left(\mathbf{q}_M, t_B + \frac{\Delta t}{2} \right)$$

$$\Rightarrow \mathbf{M}_M (\mathbf{u}_E - \mathbf{u}_B) - \mathbf{h}_M \Delta t - \sum_{i=1}^n \mathbf{W}_{Mi} \hat{\mathbf{P}}_i = 0 \quad (15)$$

The integrals in inclusion (10) can now be evaluated for the small finite time interval $t_E - t_B = \Delta t$,

$$\int_{[t_B, t_E]} d\gamma_i = \gamma_{Ei} - \gamma_{Bi}$$

$$\int_{[t_B, t_E]} d\mathcal{A}_i =: \hat{\mathcal{A}}_i \quad (16)$$

$$\int_{[t_B, t_E]} d\mathbf{P}_i =: \hat{\mathbf{P}}_i$$

which yields a discrete set-valued law

$$-(\gamma_{Ei} + \varepsilon_i \gamma_{Bi}) \in \mathcal{N}_{\hat{\mathcal{A}}_i}(\hat{\mathbf{P}}_i) \quad (17)$$

Note that (17) applies only to active set-valued laws, i.e. set-valued laws that can be described on the velocity level. As friction elements are naturally defined on the velocity level, they are always active and can always be described by (17). Considering unilateral contacts, Moreau's midpoint algorithm calculates the contact distances g_i of all unilateral contacts at the midpoint \mathbf{q}_M in order to evaluate whether these are active ($g_i \leq 0$) or not ($g_i > 0$). Only active unilateral contacts can be modelled by inclusion (17). Unilateral contacts that are non-active, thus open, are disregarded because it is assumed that their contact force contribution is equal to 0.

The discrete scheme (15) and (17) holds for both the smooth motion and the impact case, but only for small time step sizes. Time-stepping schemes are very robust integration schemes and are well suited for mechanical systems with several thousands of contacts. In Figure 2 we exemplarily show thousand disks that are mixed. Between all the disks acts friction.

In what follows, we rewrite methods (15) and (17) in a convenient form and turn the resulting inclusions into projective equations. Various approaches exist to arrive at the projective equations, consider, for example, the natural map of Robinson [36, 37], exact regularization [21] or the saddle-point of the augmented Lagrangian [38, 39]. The resulting projective equations can be solved iteratively, see, for example, [21, 40–44]. The relative velocities γ_i of the i th set-valued law are linked to the generalized velocities \mathbf{u} by $\gamma_i = \mathbf{W}_{Mi}^\top \mathbf{u} + \zeta_{Mi}$ with

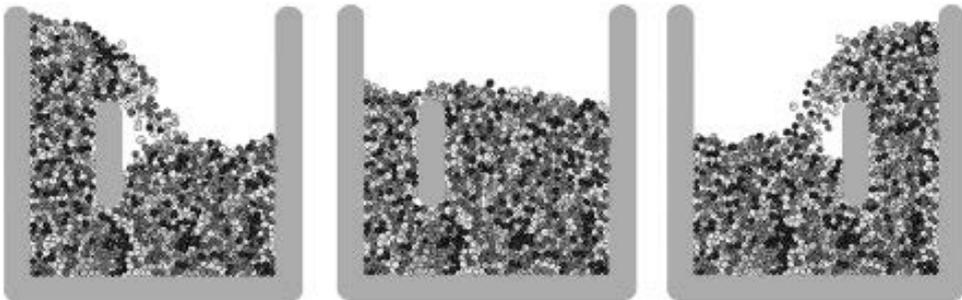


Figure 2. Mixing thousand disks with Moreau's time-stepping scheme.

$\zeta_{Mi} = \zeta_i(\mathbf{q}_M, t_M)$. Using this relation, we can transform the discrete scheme (15) and (17) into relative coordinates γ_i and $\hat{\mathbf{P}}_i$. Elimination of γ_i yields n inclusions describing the individual behaviour of the n set-valued laws

$$-(\gamma_{Ei} + \varepsilon\gamma_{Bi}) = - \left(\sum_{j=1}^n \mathbf{G}_{ij} \hat{\mathbf{P}}_j + \mathbf{c}_i \right) \in \mathcal{N}_{\hat{\mathcal{A}}_i}(\hat{\mathbf{P}}_i), \quad i = 1 \dots n \quad (18)$$

in which $\mathbf{G}_{ij} = \mathbf{W}_{Mi}^\top \mathbf{M}_M^{-1} \mathbf{W}_{Mj}$ denotes the local Delassus matrix and $\mathbf{c}_i = (\mathbf{I} + \varepsilon)\gamma_{Bi} + \mathbf{W}_{Mi}^\top \mathbf{M}_M^{-1} \mathbf{h}_M \Delta t + \zeta_{Mi}$. The matrix ε is a diagonal matrix containing the restitution coefficients ε_i . The inclusions (18) can now be expressed as projective equations [43, 44],

$$\hat{\mathbf{P}}_i = \text{prox}_{\hat{\mathcal{A}}_i} \left(\hat{\mathbf{P}}_i - r_i \left(\sum_{j=1}^n \mathbf{G}_{ij} \hat{\mathbf{P}}_j + \mathbf{c}_i \right) \right), \quad r_i > 0, \quad i = 1 \dots n \quad (19)$$

in which r_i is an arbitrary positive scalar. The $\text{prox}_{\hat{\mathcal{A}}_i}(\mathbf{x})$ function determines the nearest point in the convex set $\hat{\mathcal{A}}_i$ to the function argument, i.e.

$$\text{prox}_{\hat{\mathcal{A}}_i}(\mathbf{x}) = \underset{\hat{\mathbf{x}} \in \hat{\mathcal{A}}_i}{\text{argmin}} \|\mathbf{x} - \hat{\mathbf{x}}\| \quad (20)$$

Solving the n projective equations (19) yields the discrete percussion $\hat{\mathbf{P}}_i$, after which the generalized velocities \mathbf{u}_E and the positions

$$\mathbf{q}_E = \mathbf{q}_B + \frac{\mathbf{u}_B + \mathbf{u}_E}{2} \Delta t \quad (21)$$

can be calculated. It is suggested that the system of projective equations (19) is solved iteratively, for example, by an iterative scheme of the kind

$$\hat{\mathbf{P}}_i^{v+1} = \text{prox}_{\hat{\mathcal{A}}_i} \left(\hat{\mathbf{P}}_i^v - r_i \left(\sum_{j=1}^n \mathbf{G}_{ij} \hat{\mathbf{P}}_j^v + \mathbf{c}_i \right) \right), \quad i = 1 \dots n \quad (22)$$

or

$$\hat{\mathbf{P}}_i^{v+1} = \text{prox}_{\hat{\mathcal{A}}_i} \left(\hat{\mathbf{P}}_i^v - r_i \left(\sum_{j=1}^{i-1} \mathbf{G}_{ij} \hat{\mathbf{P}}_j^{v+1} + \sum_{j=i}^n \mathbf{G}_{ij} \hat{\mathbf{P}}_j^v + \mathbf{c}_i \right) \right), \quad i = 1 \dots n \quad (23)$$

These two schemes (22) and (23) follow straightforwardly from the projective equations (19). However, these two schemes are very closely related to the linear Jacobi and Gauss–Seidel relaxation schemes, which are discussed in the following. We aim at solving the linear system $\mathbf{G}\hat{\mathbf{P}} + \mathbf{c} = \mathbf{0}$ iteratively. Therefore, we split the matrix \mathbf{G} into two matrices \mathbf{B} and \mathbf{C} , i.e. $\mathbf{G} = \mathbf{B} + \mathbf{C}$, which yields the iterative scheme

$$\mathbf{B}\hat{\mathbf{P}}^{v+1} + \mathbf{C}\hat{\mathbf{P}}^v + \mathbf{c} = \mathbf{0} \quad \Rightarrow \quad \hat{\mathbf{P}}^{v+1} = -\mathbf{B}^{-1}(\mathbf{C}\hat{\mathbf{P}}^v + \mathbf{c}) \quad (24)$$

The matrix \mathbf{B} has to be chosen such that it is easily invertible. Two natural choices exist: $\mathbf{B} = \omega^{-1}\mathbf{D}$ and $\mathbf{B} = \mathbf{L} + \omega^{-1}\mathbf{D}$, where ω is a diagonal matrix that contains the relaxation parameters. The

matrices \mathbf{D} and \mathbf{L} contain the diagonal and the strictly lower triangular part of \mathbf{G} , respectively. In the first case we obtain the Jacobi relaxation scheme (JOR)

$$\hat{\mathbf{P}}^{v+1} = \hat{\mathbf{P}}^v - \omega \mathbf{D}^{-1} (\mathbf{G} \hat{\mathbf{P}}^v + \mathbf{c}) \quad (25)$$

and in the second case we obtain the Gauss–Seidel relaxation scheme (SOR), which can be expressed as

$$\hat{\mathbf{P}}^{v+1} = \hat{\mathbf{P}}^v - \omega \mathbf{D}^{-1} (\mathbf{L} \hat{\mathbf{P}}^{v+1} + (\mathbf{G} - \mathbf{L}) \hat{\mathbf{P}}^v + \mathbf{c}) \quad (26)$$

Both schemes can conveniently be transformed into local coordinates $\hat{\mathbf{P}}_i$, i.e.

$$\hat{\mathbf{P}}_i^{v+1} = \hat{\mathbf{P}}_i^v - \omega_{ii} \mathbf{D}_{ii}^{-1} \left(\sum_{j=1}^n \mathbf{G}_{ij} \hat{\mathbf{P}}_j^v + \mathbf{c}_i \right), \quad i = 1 \dots n \quad (27)$$

for the JOR scheme and

$$\hat{\mathbf{P}}_i^{v+1} = \hat{\mathbf{P}}_i^v - \omega_{ii} \mathbf{D}_{ii}^{-1} \left(\sum_{j=1}^{i-1} \mathbf{G}_{ij} \hat{\mathbf{P}}_j^{v+1} + \sum_{j=i}^n \mathbf{G}_{ij} \hat{\mathbf{P}}_j^v + \mathbf{c}_i \right), \quad i = 1 \dots n \quad (28)$$

for the SOR scheme. When comparing the JOR and SOR schemes (27) and (28) with the schemes (22) and (23), it is obvious that the latter two schemes consist of a Jacobi or Gauss–Seidel relaxation step for solving a linear system $\mathbf{G}\hat{\mathbf{P}} + \mathbf{c} = \mathbf{0}$, combined with a projection. We therefore refer to (22) and (23) as projected Jacobi (PJOR) and projected Gauss–Seidel (PSOR) relaxation schemes. The factor r_i is in close relation with the term $\omega_{ii} \mathbf{D}_{ii}^{-1}$, i.e. $r_i \mathbf{I} = \omega_{ii} \mathbf{D}_{ii}^{-1}$, where the entries of ω_{ii} cannot be chosen independently. For a detailed discussion on the iterative solution of projective equations of type (19) we refer to [20, 41–43, 45, 46].

4. ‘CONSTRAINT’ AND ‘IMPRESSED’ MODES

In this section, we investigate the system of projective equations (19) and their iterative solution (22) or (23) more closely. Note that each projective equation has an associated set-valued law that characterizes a unilateral contact, a friction element or an arbitrary other non-smooth element. We will distinguish between set-valued laws in ‘constraint’ and ‘impressed’ modes and relate these two terms to the projection behaviour of the $\text{prox}_{\hat{\mathcal{A}}_i}$ functions. In the end of the section we give a small example.

Consider the system of projective equations (19) in the solution point $\hat{\mathbf{P}}_j^\infty$ of the iterative scheme, i.e.

$$\hat{\mathbf{P}}_i^\infty = \text{prox}_{\hat{\mathcal{A}}_i} \left(\hat{\mathbf{P}}_i^\infty - r_i \left(\sum_{j=1}^n \mathbf{G}_{ij} \hat{\mathbf{P}}_j^\infty + \mathbf{c}_i \right) \right) \quad (29)$$

This equation can be fulfilled in two ways: either the argument \mathbf{x} of the $\text{prox}_{\hat{\mathcal{A}}_i}$ function lies in the interior of set $\hat{\mathcal{A}}_i$, which yields $\text{prox}_{\hat{\mathcal{A}}_i}(\mathbf{x}) \equiv \mathbf{x}$, or the argument \mathbf{x} of the $\text{prox}_{\hat{\mathcal{A}}_i}$ function does not

lie in the interior of set $\hat{\mathcal{A}}_i$, which yields the $\text{prox}_{\hat{\mathcal{A}}_i}$ function to really project on the boundary of set $\hat{\mathcal{A}}_i$. In the first case we obtain

$$\sum_{j=1}^n \mathbf{G}_{ij} \hat{\mathbf{P}}_j^\infty + \mathbf{c}_i = -(\gamma_{Ei} + \varepsilon \gamma_{Bi}) = \mathbf{0} \quad (30)$$

i.e. $\hat{\mathbf{P}}_i^\infty$ is chosen such that the constraint equation $-(\gamma_{Ei} + \varepsilon \gamma_{Bi}) = \mathbf{0}$ is enforced. Therefore, we denote the associated set-valued law to be in the ‘constraint’ mode. In the second case, the discrete percussion $\hat{\mathbf{P}}_i^\infty$ is impressed due to the projection on the boundary of the set $\hat{\mathcal{A}}_i$, and we denote the associated set-valued law to be in the ‘impressed’ mode. Using the terms ‘constraint’ and ‘impressed’ mode allows for a uniform terminology, which must not distinguish between unilateral contacts, friction elements or other non-smooth elements, for example, a pre-stressed spring. The terminology simply refers to the mathematical structure of the set-valued laws, i.e. whether the laws enforce a constraint equation by a set-valued percussion $\hat{\mathbf{P}}_i$ or whether the laws impress a percussion $\hat{\mathbf{P}}_i$ by an equation of the kind $\hat{\mathbf{P}}_i = \mathbf{f}(\gamma_{Ei}, \gamma_{Bi})$. Note that an integrator only cares about this mathematical structure, i.e. physical criteria such as open or closed unilateral contact are not needed. Using physical criteria such as open or closed unilateral contact is intuitive but offers the following disadvantages: First, one has to guarantee that physical criteria really match the used mathematical structure. This is not as easy as it seems. For example, a sticking friction element will never have a vanishing relative velocity $\gamma_i = 0$ due to numerical reasons. Therefore, tolerances have to be introduced in order to decide whether the friction element sticks or not. In this context it would be, for example, possible that the friction element is detected to be sliding by a physical criterion, but that the mathematical structure remains in the ‘constraint’ mode, which yields a contradiction due to numerical errors. Second, the implementation might not be so easy, as each set-valued law requests its own physical detection procedure. When monitoring the mathematical structure of the set-valued laws, all kinds of set-valued laws follow the same detection procedure, which yields a very short, uniform and extendable implementation.

In practice, we examine the behaviour of the $\text{prox}_{\hat{\mathcal{A}}_i}$ functions at the very end of the iteration in order to determine whether a set-valued law is in the ‘constraint’ mode or the ‘impressed’ mode. We assign to a set-valued law i the state $\hat{\sigma}_i = 1$ if the associated set-valued law is in the ‘impressed’ mode, otherwise the value $\hat{\sigma}_i = 0$.

Example

Consider a point mass that can move on a table in the x -direction. Between the point mass and table acts friction with a maximal friction force $a = \mu mg$. In addition, a force F pulls the point mass in the x -direction. The velocity in the x -direction is denoted by u . The discrete time-stepping scheme for this example reads as

$$m(u_E - u_B) - F\Delta t - \hat{P} = 0, \quad -(u_E + \varepsilon u_B) \in \mathcal{N}_{\hat{\mathcal{A}}}(\hat{P}), \quad \hat{\mathcal{A}} = [-a\Delta t; a\Delta t] \quad (31)$$

Elimination of the velocity u_E yields an inclusion expressed solely in terms of the discrete percussion \hat{P} ,

$$-\left(\frac{F\Delta t}{m} + \frac{\hat{P}}{m} + (1 + \varepsilon)u_B \right) \in \mathcal{N}_{\hat{\mathcal{A}}}(\hat{P}) \quad (32)$$

The projective equation becomes

$$\hat{P} = \text{prox}_{\hat{\mathcal{A}}} \left(\hat{P} - r \left(\frac{F \Delta t}{m} + \frac{\hat{P}}{m} + (1 + \varepsilon) u_B \right) \right) \quad (33)$$

We assume that the velocity u_B at the beginning of the time step is zero and examine the behaviour of the system for different values of the external force F . If F is smaller than the maximal friction force a , which means $|F| < a$, then the point mass will stick on the table and the iterative solution of the projective equation (33) has a fixed point at $\hat{P} = -F \Delta t$,

$$-F \Delta t = \text{prox}_{\hat{\mathcal{A}}}(-F \Delta t) \quad (34)$$

in which the $\text{prox}_{\hat{\mathcal{A}}}$ does not change the value of its argument. The force \hat{P} fulfils the constraint $u = 0$, which can easily be verified by inserting $\hat{P} = -F \Delta t$ in the equations of motion (31). If the external force F is larger than the maximal friction force a , i.e. $|F| > a$, then the point mass will slide on the table and the iterative solution of the projective equation (33) has a fixed point at $\hat{P} = -a \Delta t$,

$$-a \Delta t = \text{prox}_{\hat{\mathcal{A}}} \left(\underbrace{-a \Delta t - r \frac{\Delta t}{m} (F - a)}_{< -a \Delta t} \right) \quad (35)$$

in which the $\text{prox}_{\hat{\mathcal{A}}}$ projects on the set $\hat{\mathcal{A}} = [-a \Delta t, a \Delta t]$ and impresses a discrete percussion $\hat{P} = -a \Delta t$.

5. SWITCHING POINTS

In this section, we discuss how switching points can be detected within a time-stepping scheme. Furthermore, we suggest a step size adjustment in order to resolve these switching points properly.

5.1. Detecting switching points

As already mentioned in Section 1, time-stepping schemes use for each time step m a discrete state $\hat{\mathfrak{f}}^m$, which applies for the *entire* time step. This discrete state can easily be determined by examining the behaviour of the $\text{prox}_{\hat{\mathcal{A}}_i}$ functions, as discussed in Section 4. If two successive time steps m and $m+1$ have different discrete states $\hat{\mathfrak{f}}^m \neq \hat{\mathfrak{f}}^{m+1}$, i.e. if at least one set-valued law changes from ‘impressed’ to ‘constraint’ mode or *vice versa*, then one of the two time steps contains a switching point,

$$\hat{\mathfrak{f}}^m \neq \hat{\mathfrak{f}}^{m+1} \Rightarrow \text{Switching point either in time step } m \text{ or } m+1 \quad (36)$$

In Figure 3 we see, for example, the time evolution of a system with four set-valued laws. We recognize that the second set-valued law switches its discrete state $\hat{\sigma}_2$ from 0 in the second time step to 1 in the third time step. Thus, the corresponding switching point will be located between t_1 and t_3 . Note that the different discrete state $\hat{\mathfrak{f}}$ in the third time step does not necessarily imply that the switching point is located in this time step, as it is also possible that it is located at the end of the second time step.

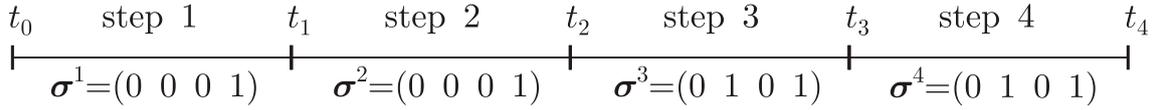


Figure 3. Time evolution of a non-smooth system with four set-valued laws. The behaviour of the set-valued laws is indicated by the values 1 for the ‘impressed’ mode and 0 for the ‘constraint’ mode. The third time step has a discrete state different from its predecessor, and the corresponding switching point will be located between t_1 and t_3 .

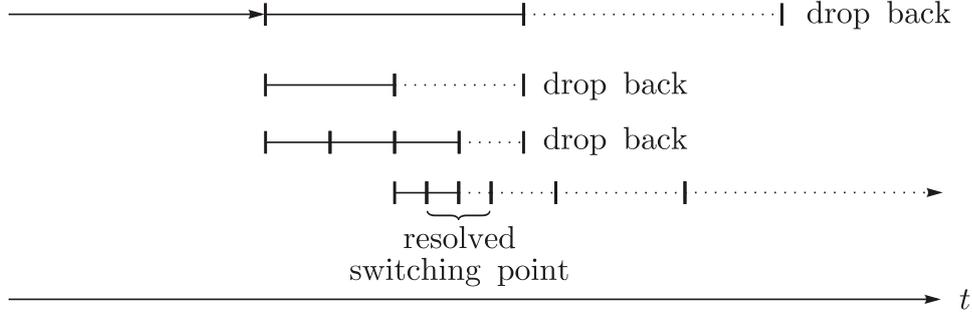


Figure 4. Switching point search by a regula falsi method. A change in the discrete state $\hat{\sigma}$ is indicated by a change in the line style (solid to dotted).

Examples

These two examples aim at emphasizing that $\hat{\sigma}^m \neq \hat{\sigma}^{m+1}$ does not necessarily induce that the switching point is within time step $m + 1$. It can also be located at the end of time step m . Suppose, for example, that a unilateral contact is open at the midpoint \mathbf{q}_M but closed by the end of the time step. Another example is given by the friction element discussed in the example of Section 4. Imagine that the discussed point mass is sliding on the table having a velocity u_B , and imagine that no external force F is present. Owing to the sliding force $-a$ the point mass will stop after a time lapse $\Delta t_1 = u_B m / a$. Examining the corresponding projective equation (35) we recognize that the friction element will only be in the ‘constraint’ mode if the step size is larger than $\Delta t_2 = (1 + \varepsilon) u_B m / a$. For any time step with step size between Δt_1 and Δt_2 , the $\text{prox}_{\hat{\mathcal{A}}}$ function will still project on set $\hat{\mathcal{A}}$ and the friction will be in the ‘impressed’ mode, although the point mass will stick on the table by the end of the time step.

5.2. Resolving switching points with minimal step size

In Section 5.1 we have discussed how time steps that contain switching points can be detected. In what follows we aim at decreasing the step size Δt of such time steps in order to resolve the switching points properly. The main idea is to use a regula falsi approach. In the case of a detected switching point ($\hat{\sigma}^m \neq \hat{\sigma}^{m+1}$), we drop back to the beginning of the predecessor time step m and continue the integration with a smaller step size (for example, $\Delta t \rightarrow \Delta t / 2$) until another switching point is detected. Then we drop back again and repeat this procedure until a minimal step size Δt_{\min} is reached which resolves the switching point properly. After having resolved the switching point, the step size can again be increased until a maximal step size Δt_{\max} is reached, for example, by a factor 2 in each smooth step. This procedure is depicted in Figure 4. Note that, in the case of a detected switching point ($\hat{\sigma}^m \neq \hat{\sigma}^{m+1}$), we drop back to the beginning of the time step m , because the switching point can be located either in the time step m or $m + 1$; see (36). Thus, we

reject both the time steps m and $m + 1$. A more detailed algorithmic description of this procedure is given in Section 9.

6. INTEGRATION ORDER OF MOREAU'S MIDPOINT RULE

In what follows we briefly discuss the integration order of Moreau's time-stepping method. We remember that an integration scheme is called of order p if the absolute value of its local error $e_L(t_{m+1}) = y(t_{m+1}) - \tilde{v}_{m+1}$ is smaller than an arbitrary bounded constant K_L multiplied by Δt^{p+1} . In this case, it holds for the global error $e_G(t_{m+1}) = y(t_{m+1}) - v_{m+1}$ that its absolute value is smaller than an arbitrary constant K_G multiplied by Δt^p . Note that $y(t_{m+1})$ is the exact value at time t_{m+1} , \tilde{v}_{m+1} is the approximation after one integration step started with $y(t_m)$ and v_{m+1} is the approximation gained after several integration steps.

Considering Moreau's midpoint rule, we distinguish between the local integration error e_L of smooth time steps and of switching intervals. For smooth time steps, standard results from numerical integration theory can be used [47–50]. It can be shown that the local integration error in the displacements e_{Lq} and in the velocities e_{Lu} is of order 3 and 2, respectively,

$$|e_{Lq}| \leq K_{Lq} \Delta t^3 \quad (37)$$

$$|e_{Lu}| \leq K_{Lu} \Delta t^2 \quad (38)$$

see [28]. Thus, the global error e_G accumulated in smooth parts of the motion is of order 2 for the displacements and of order 1 for the velocities. The overall global integration error order for the smooth parts is therefore equal to 1. Considering switching intervals, the determination of the integration order is not trivial and not satisfactorily solved up to now. Especially it has to be discussed to which extent the definition of the integration order, which is based on Taylor series expansions, makes sense for the non-smooth case. Also the treatment of accumulative switching points is a challenging problem. In this paper we assume that the local integration order of switching intervals is at least 1, which is a reasonable assumption [28]. Such a local error order would be disastrous considering smooth time steps, because the global error order would become 0 meaning that the global error would not decrease for decreasing step size. In contrast to smooth time steps, the number of switching intervals in a time span does not depend on the chosen step size as soon as a certain refinement has been reached, whereas we exclude the case of accumulative switching points. As a consequence, we assume that the local integration error in the switching interval adds to the global integration error with order 1, but not with order 0 as smooth time steps would do. In the case of dissipative accumulative switching points, the corresponding relative contact velocities are very small just right before the accumulation point. Thus, they should not affect the integration error very much. To conclude, the local errors of all time steps add to the global integration error with at least order 1, and Moreau's midpoint method yields therefore an order 1 integration scheme. Consider in this context also the work of Janin and Lamarque [51], in which the computation of the integration error is given for the case of a finite number of impacts.

7. EXTRAPOLATION METHODS FOR SMOOTH SYSTEMS

In this section we give a brief introduction to the topic of smooth extrapolation methods. Detailed information about these integration schemes can be found in [47, 50, 52, 53]. We will use these

extrapolation methods in Section 8 to improve the integration order of the time-stepping integration during smooth parts of the motion.

Consider an arbitrary base integration scheme of order p . Extrapolation methods allow one to construct an order $(p+k-1)$ approximation using k approximations calculated with the base integration scheme. It can be shown that for smooth functions there exists an asymptotic expansion of the global error $e_{G,\Delta t}(t)$ [47, 50],

$$e_{G,\Delta t}(t) = c_p(t)\Delta t^p + \dots + c_{p+k-2}(t)\Delta t^{p+k-2} + \mathcal{O}(\Delta t^{p+k-1}) \quad (39)$$

Let Δt be the step size of one time step, $\Delta t = t_{m+1} - t_m$. Using the base integration scheme in connection with different step sizes $\Delta t/n_i$, different order p approximations $T_{i,1}$ for $y(t_{m+1})$ can be computed. The index i in $T_{i,j}$ indicates the identifier of the approximation, the index j denotes how many approximations $T_{i,1}$ are used to obtain the approximation $T_{i,j}$. If the base integration scheme is of order 1, then the index j indicates the integration order. The scalar n_i denotes the number of substeps that were taken to obtain the i th first-order approximation $T_{i,1}$. For example, we calculate $T_{1,1}$ by performing one substep Δt , $T_{2,1}$ by performing two substeps $\Delta t/2$ and $T_{3,1}$ by performing four substeps $\Delta t/4$, i.e. $n_3 = 4$, with our base integration scheme. Using the definition of the global error we express

$$\begin{aligned} T_{1,1} &= y - e_{\Delta t} = y - \mathcal{O}(\Delta t^{p+2}) - c_p(\Delta t)^p - c_{p+1}(\Delta t)^{p+1} \\ T_{2,1} &= y - e_{\Delta t/2} = y - \mathcal{O}(\Delta t^{p+2}) - c_p\left(\frac{\Delta t}{2}\right)^p - c_{p+1}\left(\frac{\Delta t}{2}\right)^{p+1} \\ T_{3,1} &= y - e_{\Delta t/4} = y - \mathcal{O}(\Delta t^{p+2}) - c_p\left(\frac{\Delta t}{4}\right)^p - c_{p+1}\left(\frac{\Delta t}{4}\right)^{p+1} \end{aligned} \quad (40)$$

By introducing an order $p+2$ approximation $w = y - \mathcal{O}(\Delta t^{p+2})$ we can solve the above system (40) for the unknowns w , c_p and c_{p+1} . Using the three approximations $T_{1,1}$, $T_{2,1}$ and $T_{3,1}$ of order p we manage to compute an order $p+2$ approximation w of y . To advance the application of the method we interpret the left-hand sides of Equations (40) as polynomials $P(\Delta t)$, $P(\Delta t/2)$ and $P(\Delta t/4)$, where

$$P(\Delta t) = w - c_p\Delta t^p - c_{p+1}\Delta t^{p+1} \quad (41)$$

Solving system (40) is then equal to placing the polynomial $P(\Delta t)$ through the points $(T_{1,1}, \Delta t)$, $(T_{2,1}, \Delta t/2)$ and $(T_{3,1}, \Delta t/4)$. Note that the approximation w is equal to the value of the interpolation polynomial $P(\Delta t)$ at $\Delta t = 0$; hence, we restrict ourselves to calculate only this value. The interpolation and the following extrapolation to the value $w = P(0)$ is done by the Aitken–Neville algorithm [49]. Using this algorithm we can calculate the value $P(0)$ of an interpolating polynomial directly from the node data $(T_{i,1}, \Delta t/n_i)$. Because the algorithm is based on a recursive formula, we can add nodes to the interpolation polynomial and obtain the new value for $P(0)$ without losing the calculation effort already done. This nice property allows us to change easily the order of an extrapolation method. For example, we calculate two-order p approximations $T_{1,1}$ and $T_{2,1}$. Then we compute the value $P(0)$ of the interpolation polynomial defined by $T_{1,1}$ and $T_{2,1}$ with the Aitken–Neville algorithm in order to obtain the order $p+1$ approximation $T_{2,2}$ of y . We compare this new approximation $T_{2,2}$ with the old approximation $T_{1,1}$ to decide whether a further increase in the order is necessary or not. If we decide to increase the order, then we calculate a third-order

p approximation $T_{3,1}$ and obtain the new value $P(0) = T_{3,3}$ of the interpolation polynomial defined by $T_{1,1}$, $T_{2,1}$ and $T_{3,1}$ without losing the calculation effort already done.

The values $T_{i,j}$ are calculated according to Aitken–Neville using the formula

$$T_{i,j+1} = T_{i,j} + \frac{T_{i,j} - T_{i-1,j}}{n_i/n_{i-j} - 1} \quad (42)$$

and are stored in an extrapolation tableau of the form

$$\begin{array}{cccc} T_{1,1} & & & \\ T_{2,1} & T_{2,2} & & \\ T_{3,1} & T_{3,2} & T_{3,3} & \\ \vdots & \vdots & \vdots & \ddots \end{array} \quad (43)$$

To build up the extrapolation tableau (43), we first calculate the approximations $T_{1,1}$ and $T_{2,1}$ in order to get $T_{2,2}$. The approximations $T_{2,1}$ and $T_{3,1}$ give $T_{3,2}$, which results together with $T_{2,2}$ in the value $T_{3,3}$, and so on. The extrapolation tableau is successively built up until the required accuracy is reached.

Example

This small example illustrates the use of the extrapolation method. We aim at the numerical solution of the differential equation $y'(x) = 3x^2$ with $y(0) = 0$, the exact solution of which is $y(x) = x^3$. As a base scheme we take the explicit Euler rule, $y_{i+1} = y_i + y'(x_i)\Delta x$, which becomes in our case $y_{i+1} = y_i + 3x_i^2\Delta x$. In order to obtain $y(1)$, we perform one time step with $\Delta x = 1$, two time steps with $\Delta x = \frac{1}{2}$ and four time steps with $\Delta x = \frac{1}{4}$. The extrapolation tableau yields

$$\begin{aligned} T_{1,1} &= 0 \\ T_{2,1} &= \frac{3}{8}, \quad T_{2,2} = \frac{3}{8} + \frac{\frac{3}{8} - 0}{\frac{2}{1} - 1} = \frac{3}{4} \\ T_{3,1} &= \frac{21}{32}, \quad T_{3,2} = \frac{21}{32} + \frac{\frac{21}{32} - \frac{3}{8}}{\frac{4}{2} - 1} = \frac{15}{16}, \quad T_{3,3} = \frac{15}{16} + \frac{\frac{15}{16} - \frac{3}{4}}{\frac{4}{1} - 1} = 1 \end{aligned} \quad (44)$$

We recognize that the third-order approximation $T_{3,3} = 1$ gives the exact result $y(1) = 1$.

8. EXTRAPOLATION METHODS FOR NON-SMOOTH SYSTEMS

As mentioned in Section 1, the time evolution of non-smooth systems can be divided into smooth parts and switching points, or, in terms of the time-stepping algorithm, in smooth time steps and switching intervals. In this section, we show how extrapolation methods can be applied to smooth time steps. First, we remark that it does not make sense to apply extrapolation or any other higher-order integration schemes to a switching interval, because its non-smoothness contradicts the asymptotic expansion of the integration error. Switching intervals have to be run with a minimal step size Δt_{\min} , see Section 5, while smooth time steps can be treated with an increased step size and higher-order integration schemes, which is the aim of this section. Extrapolation methods

allow either a variable order or a variable step size. In our case we use the variable step size to resolve switching points and the variable order to obtain a good accuracy in the smooth parts of the motion.

We use the time-stepping method of Moreau as the base integration scheme, which is assumed to be of order 1, see also Section 6. In smooth parts, different approximations $T_{i,1}$ for the main time step m are gained by dividing this main step in different substeps, which are processed by the time-stepping method. Based on these approximations $T_{i,1}$, higher-order approximations can be computed using extrapolation. All integration steps within a main time step m must have the discrete state $\hat{\mathbf{c}}^{m-1}$ of the predecessor time interval $m-1$. If the state $\hat{\mathbf{c}}$ changes, then a switching point is present, and extrapolation cannot be applied. The step size of such switching intervals is refined according to Section 5.

When we started to use extrapolation in combination with Moreau's time-stepping scheme, we encountered serious oscillation problems that caused extrapolation to fail. We observed short spontaneous openings of unilateral contacts, which have no physical reasons. The reason for this numerical problem is discussed in Section 8.1. In Section 8.2 we discuss modifications to circumvent the problem.

8.1. Failure of extrapolation

Problems with extrapolation can occur for the case of enduring unilateral contact, for which the associated set-valued law is in the 'constraint' mode, i.e.

$$\gamma_{Ei} = -\varepsilon_i \gamma_{Bi} \quad (45)$$

Note that this equation must hold for both impact and enduring contacts. If we have a switching interval m with the 'constraint' mode, then the time step is looked upon as non-smooth and processed with a minimal step size Δt_{\min} . This corresponds to our idea of an impact. If we have a smooth time step with the 'constraint' mode, i.e. the time step's predecessor and successor time step are also in the 'constraint' mode, then we would speak of enduring contact. Consider, for example, a ball falling on a table with restitution coefficient $0 < \varepsilon < 1$. When the ball is not in contact with the table, then the contact force is zero and we assume the 'impressed' mode. If the ball touches the table, then the contact changes to the 'constraint' mode and the velocity of the ball is inverted according to Equation (45). The contact will open again, the contact force will be zero and we will have 'impressed' mode until the ball touches the table again with a smaller velocity. After some elapsed time, the contact will remain in the 'constraint' mode at a certain step. The relative contact velocity at the end of this time step will be so small that the next time step will also be in the 'constraint mode'. The system does not change anymore, we speak of enduring contact and we might increase the step size.

It is very important to realize that Equation (45) is independent of the step size Δt . Let t_E and t_B be two points in time with $t_E - t_B = \Delta t$. Applying Equation (45) on this time step yields $\gamma_{Ei} = -\varepsilon \gamma_{Bi}$. However, if we decide to split this time step into two successive time steps with step size $\Delta t/2$, we obtain $\gamma_{Ei} = (-\varepsilon)^2 \gamma_{Bi}$, which is a completely different result with even a *different* sign! Strictly speaking, Equation (45) is a mapping, but not a valid consistent integration scheme for enduring contact. The reason why the time-stepping scheme also works for enduring contact is that the relative contact velocities in the case of enduring contact are very small, and therefore the error does not affect the integration very much. The mapping is also a contraction due to $\varepsilon \leq 1$, and the very small relative contact velocity will therefore remain very small. It is assumed that

$\gamma_{Ei} \approx (-\varepsilon)^2 \gamma_{Bi} \approx -\varepsilon \gamma_{Bi} \approx 0$. When using the time-stepping scheme as the base integrator for an extrapolation method, this argumentation does not hold anymore in general. One has to keep in mind that the values $T_{i,1}$ define the nodes for the interpolating polynomial $P(\Delta t)$ and also that very small values might affect the extrapolation to $P(0)$ dramatically. Let Equation (45) be our base integration scheme for an enduring contact. We use i substeps $\Delta t/i$ to obtain $T_{i,1}$, i.e. one substep to obtain $T_{1,1}$, two substeps $\Delta t/2$ to obtain $T_{2,1}$ and so on. We choose, for example, the restitution coefficient $\varepsilon=0.7$. Calculating different first-order approximations for $T_{i,1}$ yields

$$\begin{aligned}\frac{T_{1,1}}{\gamma_{Bi}} &= (-0.7)^1 = -0.7 \\ \frac{T_{2,1}}{\gamma_{Bi}} &= (-0.7)^2 = 0.49 \\ \frac{T_{3,1}}{\gamma_{Bi}} &= (-0.7)^3 = -0.343 \\ \frac{T_{4,1}}{\gamma_{Bi}} &= (-0.7)^4 = 0.2401\end{aligned}\tag{46}$$

We recognize that all the approximations are still contractive, that is, a very small relative contact velocity $\gamma_B \approx 0$ will remain small. Making an extrapolation with the base approximations (46) yields

$$\begin{aligned}\frac{T_{1,1}}{\gamma_{Bi}} &= -0.7 \\ \frac{T_{2,1}}{\gamma_{Bi}} &= 0.49, \quad \frac{T_{2,2}}{\gamma_{Bi}} = 1.68 \\ \frac{T_{3,1}}{\gamma_{Bi}} &= -0.343, \quad \frac{T_{3,2}}{\gamma_{Bi}} = -2.009, \quad \frac{T_{3,3}}{\gamma_{Bi}} = -3.8535 \\ \frac{T_{4,1}}{\gamma_{Bi}} &= 0.2401, \quad \frac{T_{4,2}}{\gamma_{Bi}} = 1.9894, \quad \frac{T_{4,3}}{\gamma_{Bi}} = 5.9878, \quad \frac{T_{4,4}}{\gamma_{Bi}} = 9.2682\end{aligned}\tag{47}$$

The contractivity is not fulfilled, and the relative contact velocity and thus the energy grows dramatically. In Figure 5(a) the values $T_{i,i}/\gamma_{Bi}$ are depicted as a function of the restitution coefficient ε . Note that if the absolute value of $T_{i,i}/\gamma_{Bi}$ is larger than 1, the velocity and thus the energy increases. The second-order approximation $T_{2,2}$ increases the energy for restitution coefficient larger than one half, $\varepsilon > 0.5$. The third- and fourth-order approximations show an even worse behaviour. Note that the Aiken–Neville instruction (42) is of the form

$$T_{i,j+1} = T_{i,j} + \frac{1}{\alpha}(T_{i,j} - T_{i-1,j}), \quad \alpha > 0\tag{48}$$

If the signs of $T_{i,j}$ and $T_{i-1,j}$ are unequal, then the absolute value $|T_{i,j+1}|$ will always be larger than the absolute value of $|T_{i,j}|$.

8.2. Modification to prevent failure of extrapolation

In Section 8.1 we encountered the problem that extrapolation increases the energy dramatically in the case of enduring contact. This increase is due to the changing signs of the approximations $T_{i,1}$,

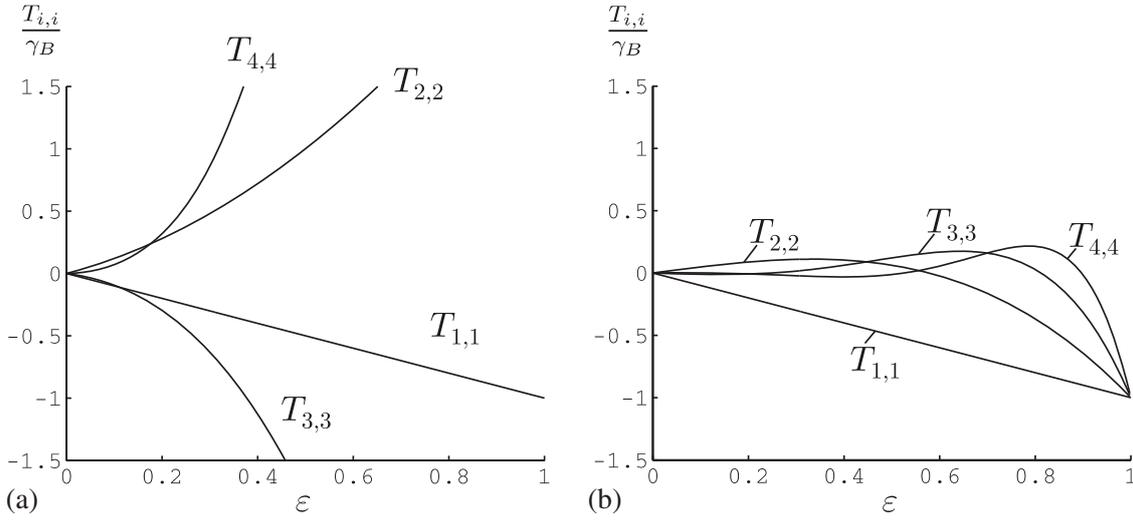


Figure 5. (a) The values $T_{i,i}$ that were obtained with $n_1=1$, $n_2=2$, $n_3=3$ and $n_4=4$. (b) The number of substeps were chosen as $n_1=1$, $n_2=3$, $n_3=5$ and $n_4=7$. The resulting approximations $T_{i,i}/\gamma_B$ are shown as a function of the restitution coefficient ϵ .

which are used for the extrapolation process. If we only use an uneven number $n_i = (i-1) \cdot 2 + 1$ of substeps to obtain $T_{i,1}$, then the sign of all approximations $T_{i,1}$ will be the same. Increasing the order will decrease the velocity and the energy. This can be verified by investigating the Aiken–Neville instruction in the general form (48). If the signs of $T_{i,j}$ and $T_{i-1,j}$ are equal, then the absolute value $|T_{i,j+1}|$ will always be smaller than the absolute value of $|T_{i,j}|$. In Figure 5(b) the values $T_{i,i}/\gamma_B$ are depicted as a function of the restitution coefficient ϵ , where $n_1=1$, $n_2=3$, $n_3=5$ and $n_4=7$ substeps were used. We recognize that the absolute value of $T_{i,i}/\gamma_B$ is always smaller than 1.

Another approach to solve the problem is to use a restitution coefficient $\epsilon=0$ for an enduring unilateral contact. In the case of partially elastic impacts, the restitution coefficient ϵ then has to be set back to its original value. The approach therefore requires a distinction between enduring unilateral contact and impact, which is not as trivial as it seems. Note that the step size adjustment uses $\Delta t \gg \Delta t_{\min}$ for smooth time steps (enduring contact) and $\Delta t = \Delta t_{\min}$ for switching intervals (impacts). Using a modified restitution coefficient $\epsilon_{\text{mod}} = \epsilon^{\Delta t / \Delta t_{\min}}$ would guarantee a small or even vanishing ϵ_{mod} in the smooth case (enduring contact) and $\epsilon_{\text{mod}} = \epsilon$ in the switching intervals (impacts). Changing the restitution coefficient during a simulation might be hazardous. Therefore, the authors encourage to use a restriction to an uneven number of substeps, as discussed in the beginning of this section.

The extrapolation for non-smooth systems described above can now be applied to the time-stepping algorithm. We take as the base integration scheme Moreau’s midpoint rule and apply the extrapolation on both the velocities \mathbf{u} and the displacements \mathbf{q} .

9. OVERALL ALGORITHM

In this section we present an overall algorithm that incorporates step size adjustment and extrapolation. The algorithm consists of three modules: the step size adjustment module, the extrapolation module and the time-stepping module. The step size adjustment module controls the simulation and strings together the different time steps. It calls the extrapolation module, which performs one

Table I. Overview of the time-stepping module.

	Performs one time step Δt according to Moreau
Pre	$\mathbf{q}_B, \mathbf{u}_B, t_B, \Delta t, \text{guess } \hat{\mathbf{P}}$
Post	$\mathbf{q}_E, \mathbf{u}_E, t_E, \hat{\mathbf{P}}, \hat{\boldsymbol{\sigma}}$

time step. For smooth time steps extrapolation is used to obtain a higher integration order. In order to obtain the approximations $T_{i,1}$, the extrapolation module calls the time-stepping module, which performs one time step according to Moreau. We will discuss the three modules in the following, starting with the basic time-stepping module.

9.1. Time-stepping module

This module is a one-step time-stepping solver, which takes the generalized coordinates \mathbf{q}_B , the generalized velocities \mathbf{u}_B , the time t_B and the step size Δt as input arguments. An initial guess for the discrete percussions $\hat{\mathbf{P}}$ should also be provided. The time-stepping module performs one time step with step size Δt . The behaviour of the prox _{\mathcal{J}} functions at the very end of the iteration is observed and returned in the vector $\hat{\boldsymbol{\sigma}}$, which describes the discrete state of the different set-valued laws. The value 1 is assigned to a set-valued law, which is in the ‘impressed’ mode (projection), and the value 0 is for a set-valued law, which is in the ‘constraint’ mode (no projection). Furthermore, the generalized coordinates \mathbf{q}_E , the generalized velocities \mathbf{u}_E and the time t_E at the end of the time step are returned. In addition, the discrete percussion $\hat{\mathbf{P}}$ is provided. A short overview of the module is given in Table I.

9.2. Extrapolation module

This module consists of an extrapolation routine with some additional features. It performs one main time step and uses, if possible, extrapolation to increase the integration order. The module takes the generalized coordinates \mathbf{q}_B , the generalized velocities \mathbf{u}_B , the time t_B , the step size Δt and the minimal step size Δt_{\min} as input arguments. An initial guess for the discrete percussions $\hat{\mathbf{P}}$ should also be provided. In addition, an accepted discrete state $\hat{\boldsymbol{\sigma}}_{\text{acc}}$ should be passed as an input argument. This vector $\hat{\boldsymbol{\sigma}}_{\text{acc}}$ defines the discrete state on which the extrapolation should be based. The extrapolation module splits the main time step into an uneven number of substeps $\Delta t / (2 \cdot (i-1) + 1)$, $i = 1, 2, \dots$. These substeps are passed to the time-stepping module to obtain the different order 1 approximations $T_{i,1}$. Based on the order 1 approximations $T_{i,1}$, the extrapolation module calculates higher-order approximations $T_{i,i}$ by building up an extrapolation tableau as described in Section 7. The extrapolation is stopped when the error criterion $\|T_{i,i} - T_{i-1,i-1}\| \leq \text{atol} + \text{rtol} \|T_{i,i}\|$ is satisfied, when a maximal user-defined integration order n_{\max} is reached or when a further order increase requests smaller time steps than Δt_{\min} . When calling the time-stepping module for computing the approximations $T_{i,1}$, the extrapolation module has always to check whether the returned discrete state $\hat{\boldsymbol{\sigma}}$ is identical to the accepted discrete state $\hat{\boldsymbol{\sigma}}_{\text{acc}}$. If this is not the case, then a switching point is detected and the extrapolation aborts. The extrapolation module exits with one of the following flags.

Successful smooth extrapolation. All substeps were performed on the accepted discrete state $\hat{\boldsymbol{\sigma}}_{\text{acc}}$, and the required integration accuracy was achieved without using substeps with step size

smaller than Δt_{\min} and without exceeding the maximal user-defined order n_{\max} . In this case the extrapolation was successful and the generalized coordinates \mathbf{q}_E , the generalized velocities \mathbf{u}_E and the time t_E at the end of the time step are returned. Note that the discrete percussion $\hat{\mathbf{P}}$ of the time step is not subjected to extrapolation and only an approximate value based on a first-order approximation is returned. In addition, the used discrete state is returned, which is identical to the accepted discrete state $\hat{\boldsymbol{\sigma}}_{\text{acc}}$.

Smooth extrapolation, required accuracy not achieved. All substeps were performed on the accepted discrete state $\hat{\boldsymbol{\sigma}}_{\text{acc}}$, but the required integration accuracy was not achieved because a substep with step size smaller than Δt_{\min} would have been necessary to increase the order, or because the maximal user-defined order n_{\max} was reached. The same data as in the case of ‘successful smooth extrapolation’ are returned together with the flag that the required accuracy was not achieved.

Detected switching point. A substep with a discrete state $\hat{\boldsymbol{\sigma}}$ different from the accepted discrete state $\hat{\boldsymbol{\sigma}}_{\text{acc}}$ is identified. The step size Δt of the main time step is larger than Δt_{\min} . A switching point is assumed and the extrapolation is aborted.

Resolved switching point. A substep with a discrete state $\hat{\boldsymbol{\sigma}}$ different from the accepted discrete state $\hat{\boldsymbol{\sigma}}_{\text{acc}}$ is identified. The step size Δt of the main time step is equal to Δt_{\min} , which means that there was only one substep with step size Δt_{\min} . The switching point is resolved. The generalized coordinates \mathbf{q}_E , the generalized velocities \mathbf{u}_E and the time t_E at the end of the time step are returned together with the discrete percussion $\hat{\mathbf{P}}$. The used discrete state is returned as a new accepted discrete state $\hat{\boldsymbol{\sigma}}_{\text{acc}}$.

A short overview of the module is given in Table II, and a flow chart is depicted in Figure 6.

9.3. Step size adjustment module

The third module is responsible for the step size adjustment. It interprets the results of the extrapolation module in order to steer the step size of the main time steps. The module strings all the main time steps together and is responsible for the output of the simulation data. The basic task of the module is to decrease the step size if a switching point is detected and to increase the step size in smooth parts of the time evolution of the system. The minimal and maximal step sizes are Δt_{\min} and Δt_{\max} , respectively. Switching points are located as described in Section 5. A short overview of the adjustment module is given in Table III.

Table II. Overview of the extrapolation module.

	Performs one time step using an extrapolation method. Calls the time-stepping module.	
Pre	$\mathbf{q}_B, \mathbf{u}_B, t_B, \Delta t, \Delta t_{\min}, \text{guess } \hat{\mathbf{P}}, \hat{\boldsymbol{\sigma}}_{\text{acc}}, n_{\max}$	
Post	Successful smooth extrapolation	$\mathbf{q}_E, \mathbf{u}_E, t_E, \hat{\mathbf{P}}, \text{unchanged } \hat{\boldsymbol{\sigma}}_{\text{acc}}$
	Smooth extrapolation, required accuracy not achieved	$\mathbf{q}_E, \mathbf{u}_E, t_E, \hat{\mathbf{P}}, \text{unchanged } \hat{\boldsymbol{\sigma}}_{\text{acc}}$
	Detected switching point	Abort
	Resolved switching point	$\mathbf{q}_E, \mathbf{u}_E, t_E, \hat{\mathbf{P}}, \text{new } \hat{\boldsymbol{\sigma}}_{\text{acc}}$

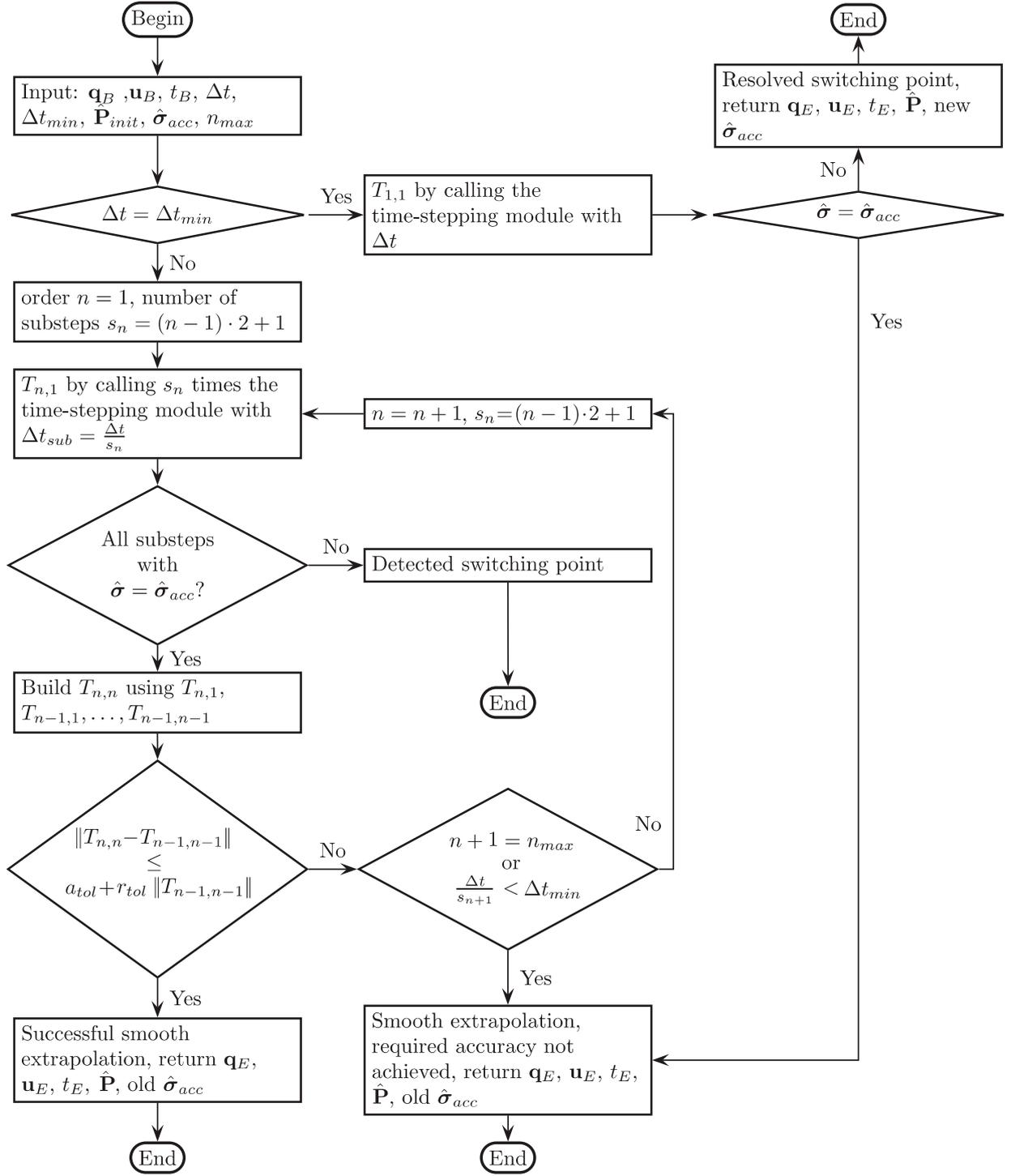


Figure 6. Flow chart of the extrapolation algorithm.

Table III. Overview of the time step adjustment module.

	Performs the simulation, controls the step size length, calls the extrapolation module.
Pre	Initial conditions $\mathbf{q}(t_0)$, $\mathbf{u}(t_0)$, t_0 , time step sizes Δt_{\min} , Δt_{\max}
Post	$\mathbf{q}(t)$, $\mathbf{u}(t)$, t , $\hat{\mathbf{P}}(t)$, switching points

When decreasing and increasing the step size, one has to keep in mind that the extrapolation procedure needs at least a step size $\Delta t = 3 \cdot \Delta t_{\min}$. Otherwise, the step size of the three substeps needed to obtain $T_{2,1}$ would be smaller than Δt_{\min} . A step size Δt with $\Delta t_{\min} < \Delta t < 3 \cdot \Delta t_{\min}$ is set equal to Δt_{\min} if the step size is subjected to diminishment. Otherwise the step size is set equal to $3 \cdot \Delta t_{\min}$.

Special attention has to be paid to ensure that step sizes are not increased while searching for a switching point. It has also to be pointed out that a change in the discrete state $\hat{\mathbf{c}}$ may not necessarily lead to a detection of a switching point, because time steps with smaller step size might lead to another solution without a switching point. These problems are solved by introducing a time t_{noInc} , which corresponds to the end time of the time step that changed its contact configuration. Because the switching point is assumed to be before this time instant, the step size should not be increased until t_{noInc} is passed.

The module stores the data of three time steps, the so-called accepted time step, the previous time step and the actual time step. The accepted time step is a time step that cannot be rejected anymore. A time step becomes accepted if its successor time step has the same discrete state $\hat{\mathbf{c}}$, or if the time step has a minimal step size Δt_{\min} . The previous time step is a time step that has been successfully processed by the extrapolation module, but which cannot yet be accepted because the step size Δt is larger than Δt_{\min} and because its successor time step has not yet been calculated. Note that the previous time step still might be rejected due to an undetected switching point at the end of the step, as already discussed in Section 5. The actual time step is a time step that is subjected to calculation.

A flow chart of the step size adjustment algorithm is depicted in Figure 7. A calculation step i takes an accepted, a previous and an actual time step i as the input, whereas the actual time step is subjected to calculation. The actual time step is processed by the extrapolation module and, according to the flags of the extrapolation module, a new accepted and a new previous time step $i + 1$ is defined. Furthermore, the initial data for a new actual time step $i + 1$ are provided. These time steps serve as the input argument for the next calculation step $i + 1$. In order to start the integration, the first actual time step of the simulation is run with the minimal step size and therefore becomes an accepted time step.

As mentioned before, the new accepted and previous time step $i + 1$ as well as the initial data for the actual time step $i + 1$ are chosen according to the flags of the extrapolation of the actual time step i . In what follows we will discuss the different possibilities.

Successful smooth extrapolation. Neither a switching point nor a critical smooth time step was detected, and the integration is continued. The actual time step i becomes the previous time step $i + 1$, and the previous time step i becomes the accepted time step $i + 1$. The initial data for the actual time step $i + 1$ follow from the end data of the actual time step i . The accepted time step i is sent to the output. The step size is increased if $t_{B,i+1}^{\text{act}} < t_{\text{noInc}}$. See Table IV for more detailed information.

Smooth extrapolation, required accuracy not reached. In this case one can either accept the results and proceed like in the case of ‘successful smooth integration’, or one can reject the actual time step i and recalculate it with a diminished step size. The latter is done in the examples presented in Section 10. An actual time step with step size Δt_{\min} cannot be rejected. The accepted and previous time steps remain unaffected.

Detected switching point. A switching point has been detected by the extrapolation method. The time t_{noInc} is set equal to the end time $t_{E,i}^{\text{act}}$ of the actual time step i . The integration drops back to the

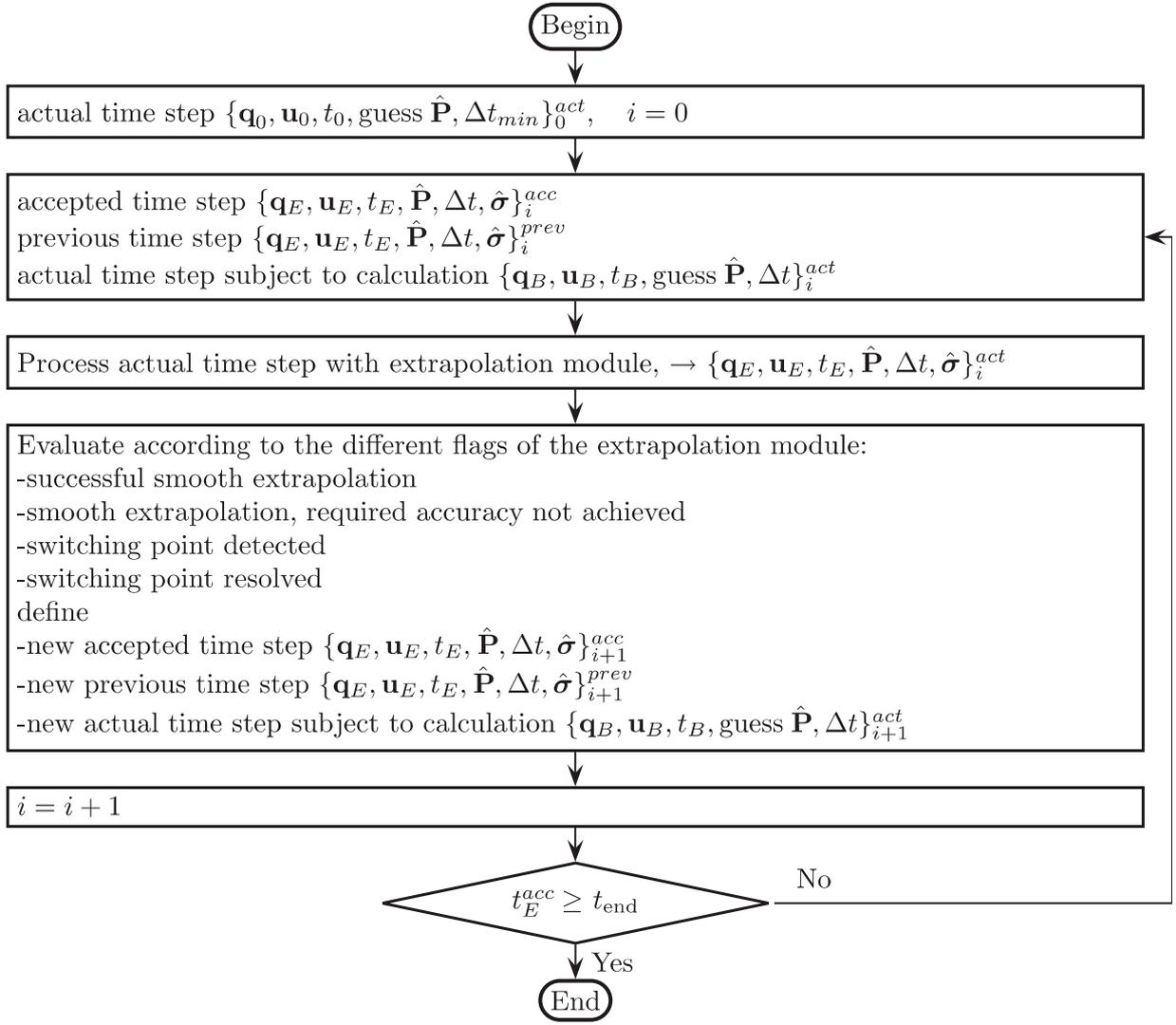


Figure 7. Flow chart of the step size adjustment algorithm.

Table IV. Successful smooth extrapolation.

Successful smooth extrapolation of actual time step $i \rightarrow \{\mathbf{q}_E, \mathbf{u}_E, t_E, \hat{\mathbf{P}}, \Delta t, \hat{\boldsymbol{\sigma}}\}_i^{\text{act}}$	
Previous $i + 1 = \text{actual } i$	$\{\mathbf{q}_E, \mathbf{u}_E, t_E, \hat{\mathbf{P}}, \Delta t, \hat{\boldsymbol{\sigma}}\}_{i+1}^{\text{pre}} = \{\mathbf{q}_E, \mathbf{u}_E, t_E, \hat{\mathbf{P}}, \Delta t, \hat{\boldsymbol{\sigma}}\}_i^{\text{act}}$
Accepted $i + 1 = \text{previous } i$	$\{\mathbf{q}_E, \mathbf{u}_E, t_E, \hat{\mathbf{P}}, \Delta t, \hat{\boldsymbol{\sigma}}\}_{i+1}^{\text{acc}} = \{\mathbf{q}_E, \mathbf{u}_E, t_E, \hat{\mathbf{P}}, \Delta t, \hat{\boldsymbol{\sigma}}\}_i^{\text{pre}}$
Initial data for actual $i + 1$	$\{\mathbf{q}_B, \mathbf{u}_B, t_B\}_{i+1}^{\text{act}} = \{\mathbf{q}_E, \mathbf{u}_E, t_E\}_i^{\text{act}}$
Step size increase	$t_{B,i+1}^{\text{acc}} > t_{\text{noInc}} \rightarrow \Delta t_{i+1}^{\text{act}} = \max(2 \cdot \Delta t_i^{\text{act}}, 3 \cdot \Delta t_{\text{min}})$ $t_{B,i+1}^{\text{acc}} \leq t_{\text{noInc}} \rightarrow \Delta t_{i+1}^{\text{act}} = \Delta t_i^{\text{act}}$

accepted time step i , i.e. the actual and the previous time steps i are rejected. The previous time step $i + 1$ is set equal to the accepted time step i . The initial data for the actual time step $i + 1$ follow from the end data of the accepted time step i . The step size is diminished. See Table V for more detailed information.

Resolved switching point. A switching point has been resolved by the extrapolation method. The actual time step i becomes both the accepted and previous time steps $i + 1$, and provides the new

Table V. Switching point detected.

Switching point detected, reject actual and previous time steps i	
Previous $i + 1 =$ accepted i	$\{\mathbf{q}_E, \mathbf{u}_E, t_E, \hat{\mathbf{P}}, \Delta t, \hat{\boldsymbol{\sigma}}\}_{i+1}^{\text{pre}} = \{\mathbf{q}_E, \mathbf{u}_E, t_E, \hat{\mathbf{P}}, \Delta t, \hat{\boldsymbol{\sigma}}\}_i^{\text{acc}}$
Accepted $i + 1 =$ accepted i	$\{\mathbf{q}_E, \mathbf{u}_E, t_E, \hat{\mathbf{P}}, \Delta t, \hat{\boldsymbol{\sigma}}\}_{i+1}^{\text{acc}} = \{\mathbf{q}_E, \mathbf{u}_E, t_E, \hat{\mathbf{P}}, \Delta t, \hat{\boldsymbol{\sigma}}\}_i^{\text{acc}}$
Initial data for actual $i + 1$	$\{\mathbf{q}_B, \mathbf{u}_B, t_B\}_{i+1}^{\text{act}} = \{\mathbf{q}_E, \mathbf{u}_E, t_E\}_i^{\text{acc}}$
Diminish step size	$\Delta t_{i+1}^{\text{act}} = \frac{1}{2} \Delta t_i^{\text{act}}$, if $\Delta t_{i+1}^{\text{act}} < 3 \cdot \Delta t_{\min} \Rightarrow \Delta t_{i+1}^{\text{act}} = \Delta t_{\min}$

Table VI. Resolved switching point.

Switching point resolved $\rightarrow \{\mathbf{q}_E, \mathbf{u}_E, t_E, \hat{\mathbf{P}}, \Delta t, \hat{\boldsymbol{\sigma}}\}_i^{\text{act}}$	
Previous $i + 1 =$ actual i	$\{\mathbf{q}_E, \mathbf{u}_E, t_E, \hat{\mathbf{P}}, \Delta t, \hat{\boldsymbol{\sigma}}\}_{i+1}^{\text{pre}} = \{\mathbf{q}_E, \mathbf{u}_E, t_E, \hat{\mathbf{P}}, \Delta t, \hat{\boldsymbol{\sigma}}\}_i^{\text{act}}$
Accepted $i + 1 =$ actual i	$\{\mathbf{q}_E, \mathbf{u}_E, t_E, \hat{\mathbf{P}}, \Delta t, \hat{\boldsymbol{\sigma}}\}_{i+1}^{\text{acc}} = \{\mathbf{q}_E, \mathbf{u}_E, t_E, \hat{\mathbf{P}}, \Delta t, \hat{\boldsymbol{\sigma}}\}_i^{\text{act}}$
Initial data for actual $i + 1$	$\{\mathbf{q}_B, \mathbf{u}_B, t_B\}_{i+1}^{\text{act}} = \{\mathbf{q}_E, \mathbf{u}_E, t_E\}_i^{\text{act}}$
Step size increase	$\Delta t_{i+1}^{\text{act}} = \max(2 \cdot \Delta t_i^{\text{act}}, 3 \cdot \Delta t_{\min})$

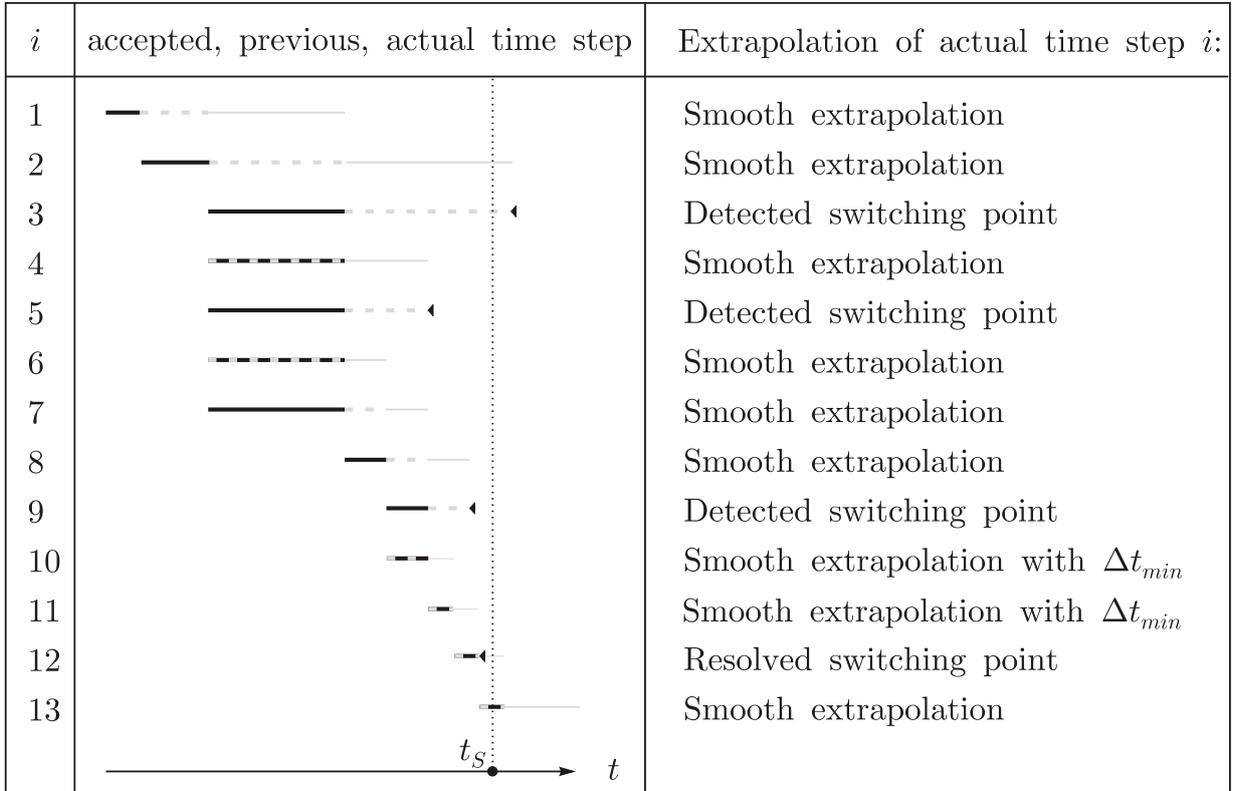


Figure 8. Switching point search by a regula falsi method. The accepted, previous and actual time steps are plotted as a black bar, a dashed grey bar and as a thin grey bar, respectively. The exact position of the switching point is denoted by t_S . Black triangles indicate that the actual time step i was not successful, because its discrete state $\hat{\boldsymbol{\sigma}}$ differs from the accepted time step i . The right column lists the extrapolation results of the actual time step i . Note that the previous and accepted time steps may coincide in some situations (black/grey dotted bars).

accepted discrete state σ_{acc} as well as the initial data for the actual time step $i + 1$. The time t_{noInc} is set equal to the end time $t_{E,i}^{\text{act}}$ of the actual time step i to enable an increase in the step size. The data of the accepted time step i are sent to the output. See Table VI for more detailed information.

Example

Figure 8 shows a typical switching point search procedure.

10. EXAMPLES

In this section we present some examples of non-smooth systems whose time evolution has been obtained by using the algorithm described above. We first show three simple examples, i.e. a point mass falling on a table, a point mass sliding on a table and a single DOF impact oscillator. Furthermore, we will discuss the woodpecker toy, which is a well-known benchmark problem in non-smooth mechanics.

10.1. Point mass falling on a table

In this first example we examine a point mass falling on a table. When the point mass hits the table, an impact occurs and the point mass is reflected. We choose the restitution coefficient $\varepsilon = 0.7$, which causes the point mass to jump less high after each impact. The time lapse between the impacts becomes smaller and smaller and ends up in an accumulation point, i.e. infinitely many impacts in a finite time. After the accumulation point, the point mass will remain on the table. The point mass is dropped from an initial height $z_0 = 0.07$ m. The mass is chosen as $m = 1$ kg. The minimal and maximal step sizes are $\Delta t_{\text{min}} = 10^{-5}$ s and $\Delta t_{\text{max}} = 0.05$ s, respectively. To underline the significance of an uneven number of extrapolation substeps, we first run the simulation with $\{1, 2, 3, 4, \dots\}$ substeps. In Figure 9(a) we see the time evolution of the step size Δt . We recognize that the step size decreases to the minimal step size in the case of impacts. After having passed the accumulation point, the step size does not increase anymore due to small velocity oscillations induced by the extrapolation; see Figure 9(b). When using only an uneven number of extrapolation substeps, this problem does not occur anymore. In Figure 10 the time evolution of the step size Δt and of the integration order as well as the time evolution of the height z and of its derivative \dot{z} are shown. We recognize that the step size increases after the simulation has passed the accumulation point. In the smooth part of the motion, the integration order is 2, which yields the exact result for the parabolic path of $z(t)$.

10.2. Point mass sliding on a table

We consider a point mass that slides on a table; see Figure 11(a). The mass and the maximal friction force are $m = 1$ kg and $a = \mu mg = 2$ N, respectively. The tangential restitution coefficient ε is set to zero. The point mass has an initial velocity $\mathbf{u}_0 = (2.0, -0.2)^\top$ m/s, and an external force $F = (-2.5, -0.25)^\top$ N acts on the point mass during the first 3 s. A plot of the path in the xy -plane is depicted in Figure 11(b). Furthermore, the time evolutions of the integration order, the step size Δt , the friction force λ and the absolute velocity of the point mass are depicted in Figures 11(c)–(f). Owing to the external force F the point mass will change its sliding direction almost to the opposite direction. The absolute velocity decreases, and a small step size

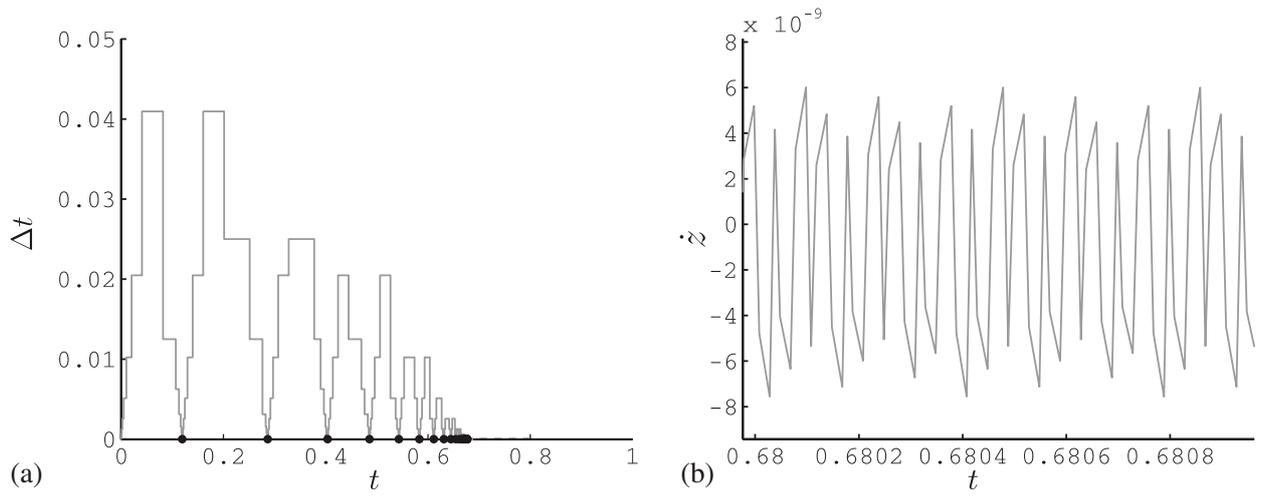


Figure 9. Simulation results of a point mass falling on a table. Computations have been done *without* the restriction of using only an uneven number of extrapolation substeps: (a) the time evolution of the step size Δt and (b) high-velocity oscillations that prevent the step size from being increased.

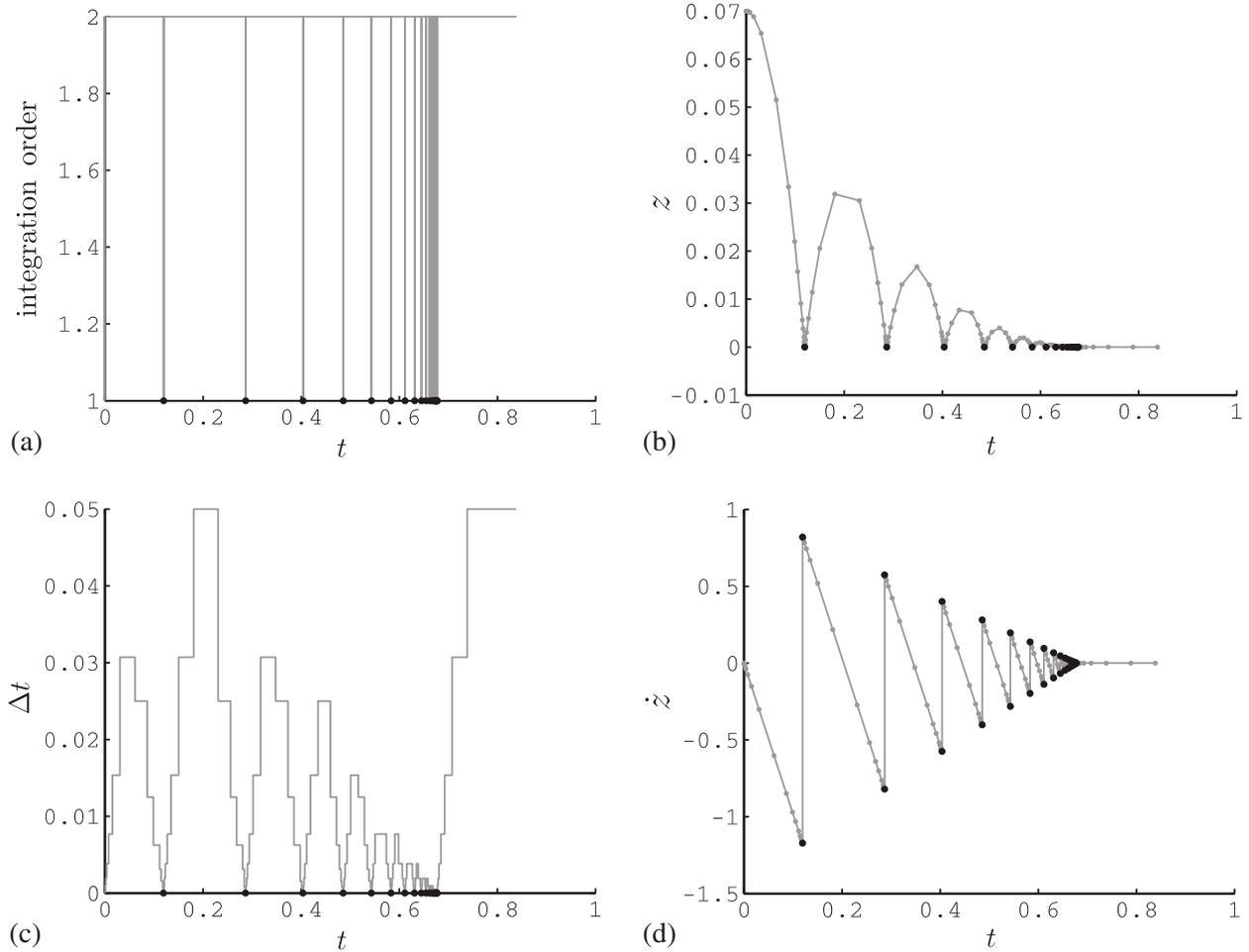


Figure 10. Simulation results of a point mass falling on a table. (a)–(d) The time evolution of the step size Δt , the integration order, the height z and its derivative \dot{z} .

together with a high integration order is used in this part of the simulation. We recognize that the sliding force in the x -direction changes its direction, and that the sliding force in the y -direction has a peak at the turning point. The absolute value of the sliding force remains 2 N. After the point mass has turned its direction, it is accelerated in the direction of the external force. The

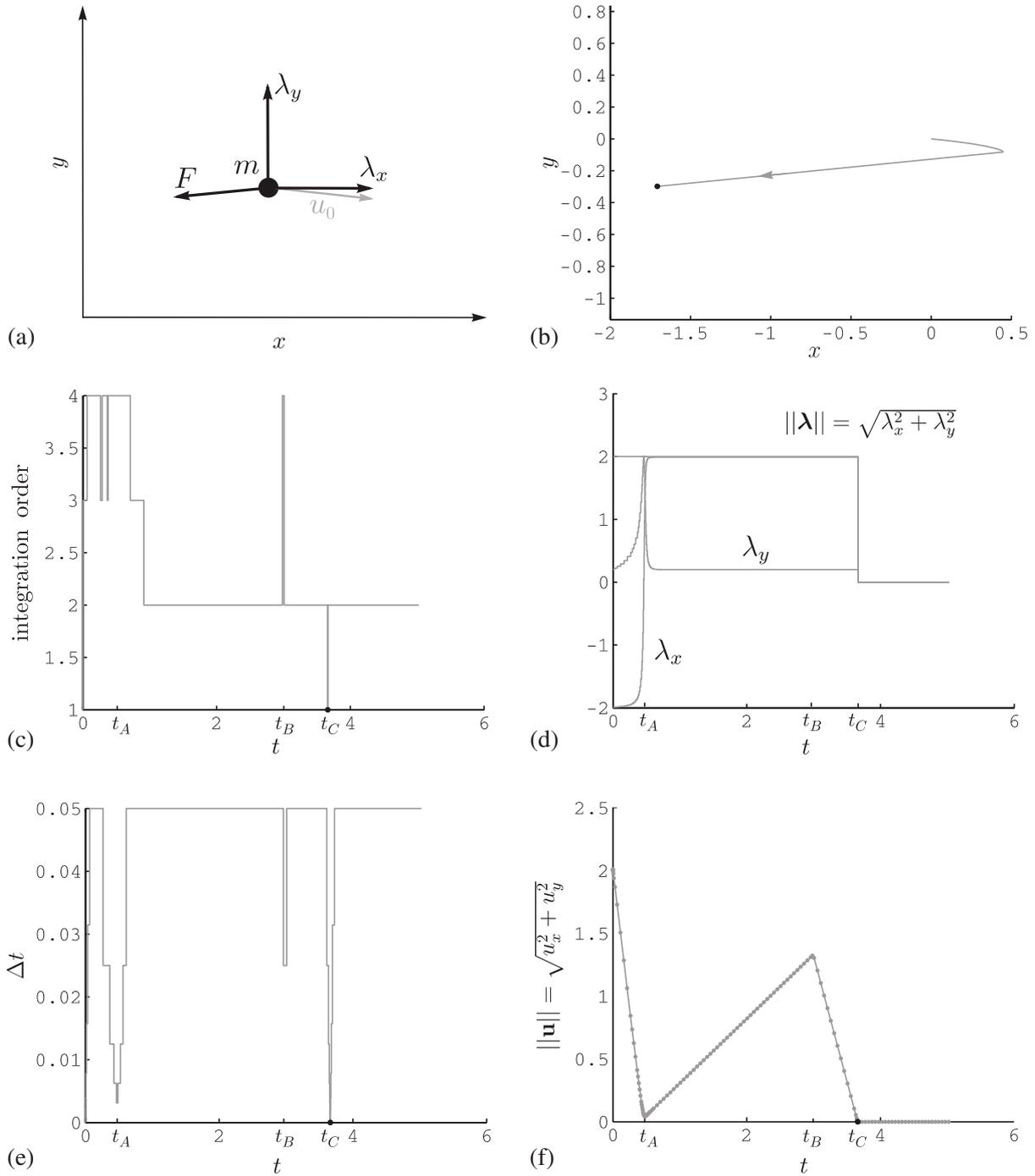


Figure 11. Point mass sliding on a table. (b) The path in the xy -plane. (c)–(f) The time evolutions of the integration order, the step size Δt , the friction force λ and the absolute velocity of the point mass. At time instants t_A , t_B and t_C the point mass turns around, the force is switched off and a slip-stick transition occurs, respectively.

displacements are described by a parabolic polynomial; thus, an integration order of 2 is sufficient. After 3 s the external force vanishes, and the friction force decelerates the point mass until it sticks.

10.3. Single DOF impact oscillator

We discuss a single DOF impact oscillator, which consists of a mass, a spring and a unilateral contact that disturbs the oscillation of the mass. The mass is $m=0.1$ kg, the spring stiffness

is $c = 20 \text{ N/m}$ and the unilateral contact's restitution coefficient is $\varepsilon = 0.6$. The position of the point mass is addressed by the coordinate x , and the velocity by its derivative \dot{x} . The spring is unstressed for $x = -0.15 \text{ m}$. The initial conditions are $x = -0.5 \text{ m}$ and $\dot{x} = 0.2 \text{ m/s}$. Figure 12(b) depicts the time evolution of the contact impulse Λ . Figures 12(c) and (d) show the position x and the velocity \dot{x} of the mass.

In contrast to the other examples, the time evolution of this example has been obtained by using a time step adjustment module, which uses a fixed integration order in the smooth time intervals, i.e. the integration order is *not* determined by an error estimation $T_{n,n} - T_{n-1,n-1}$ but is given by the user. We aim at investigating the relationship between the global error e_G , the integration order p and the maximal steps size Δt_{\max} , see Figure 12(e). Note that if the global error e_G is smaller than $K_G \Delta t_{\max}^p$, then a double-logarithmic plot of the global error e_G versus the maximal step size Δt_{\max} should be more or less a straight line with slope $s = p$. The global error e_G is the sum of the local errors e_{L1} of the smooth time steps and the local errors e_{L2} of the switching intervals. Let the propagation of the local integration error during the simulation be characterized by a scalar P , which depends on the total integration time. The smooth time steps are processed with a maximal step size Δt_{\max} and an integration order p , thus the local error e_{L1} is smaller than $K_{L1} \Delta t_{\max}^{p+1}$. The switching intervals are processed with a minimal step size Δt_{\min} and a local integration error of order 1, thus the local error e_{L2} is smaller than $K_{L2} \Delta t_{\min}$. Consider a time-stepping integration with n smooth time steps and m switching intervals. The total time span of all smooth time steps is denoted by $\Delta t_1 = n \Delta t_{\max}$, the total time span of all switching intervals by $\Delta t_2 = m \Delta t_{\min}$ and the total time span of the whole integration by $\Delta t_3 = \Delta t_1 + \Delta t_2$. Note that the m switching intervals resolve the m switching incidents and that therefore the number of switching intervals does not depend on the minimal step size Δt_{\min} . On the other hand, the number of smooth time steps increases if the maximal step size Δt_{\max} is reduced. The global integration error e_G can be estimated as

$$\begin{aligned} e_G &= P \left(n e_{L1} + m e_{L2} \right) \leq P \left(\frac{\Delta t_3 - m \Delta t_{\min}}{\Delta t_{\max}} K_{L1} \Delta t_{\max}^{p+1} + m K_{L2} \Delta t_{\min} \right) \\ &= P \left((\Delta t_3 - m \Delta t_{\min}) K_{L1} \Delta t_{\max}^p + m K_{L2} \Delta t_{\min} \right) \end{aligned} \quad (49)$$

Choosing Δt_{\min} to be equal to Δt_{\max}^p and assuming that $\Delta t_3 - m \Delta t_{\min} \approx \Delta t_3$ yield the estimation

$$e_G \leq K_G \Delta t_{\max}^p \quad (50)$$

which shows that the global error is of order p with respect to Δt_{\max} . The suggested choice of $\Delta t_{\min} = \Delta t_{\max}^p$ is problematic in view of rounding errors. If the minimal step size Δt_{\min} is too small, i.e. around the working precision of the processor, rounding errors destroy the integration order of the method. On the other hand, a maximal step size Δt_{\max} , which is too large, may cause problems in the detection of closed unilateral contacts.

In the authors' opinion it is questionable to speak of an overall integration order p in the classical sense, because we still use integration steps with a local integration order of 1 for switching points. Only the piecewise smooth parts of the motion are approximated by an order p scheme. However, the minimal step size, which resolves switching points, can be chosen small enough to ensure that the local integration error of these steps does not destroy the accuracy gained in the smooth steps. If we use the maximal step size Δt_{\max} to estimate the local integration errors of switching intervals, then we obtain a pseudo-higher-order estimation with respect to Δt_{\max} . To conclude, we achieve an overall integration order of p with respect to Δt_{\max} by restricting the local error of switching

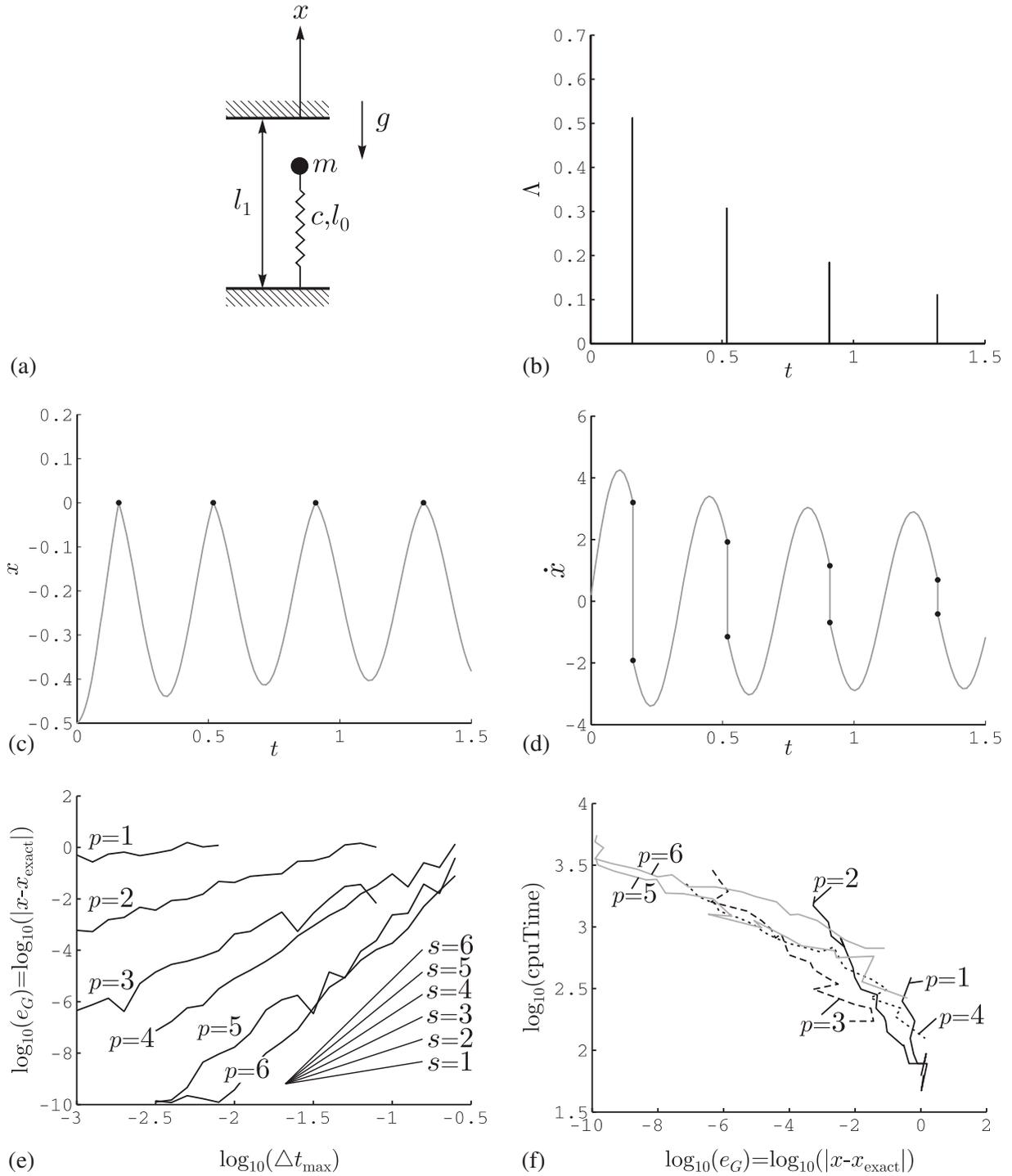


Figure 12. Single DOF impact oscillator. (b) The contact impulses Λ . (c) and (d) The position x and the velocity \dot{x} of the mass. (e) The global error $e_G = |x - x_{\text{exact}}|$ is depicted as a function of the maximal step size Δt_{max} for different calculations with integration order p . The results are approximately straight lines with slope $s = p$. (f) A double-logarithmic plot of the global error *versus* the cpu time, which is in our case linearly dependent on the number of performed successful and unsuccessful PJOR/PSOR iteration steps. When using a high integration order, the error can significantly be reduced with only a small additional effort. On the other hand, low-order integration is fast but does not allow for a significant reduction in the integration error.

intervals by a minimal step size $\Delta t_{\text{min}} = \Delta t_{\text{max}}^p$ and by restricting the local error of smooth time steps by a higher integration order p in the classical sense.

Figure 12(f) shows a double-logarithmic plot of the global error *versus* the cpu time, which in our case linearly depends on the number of performed successful and unsuccessful PJOR/PSOR

iteration steps. When using a high integration order, the error can significantly be reduced with only a small additional effort. On the other hand, low-order integration is fast but does not allow for a significant reduction in the integration error.

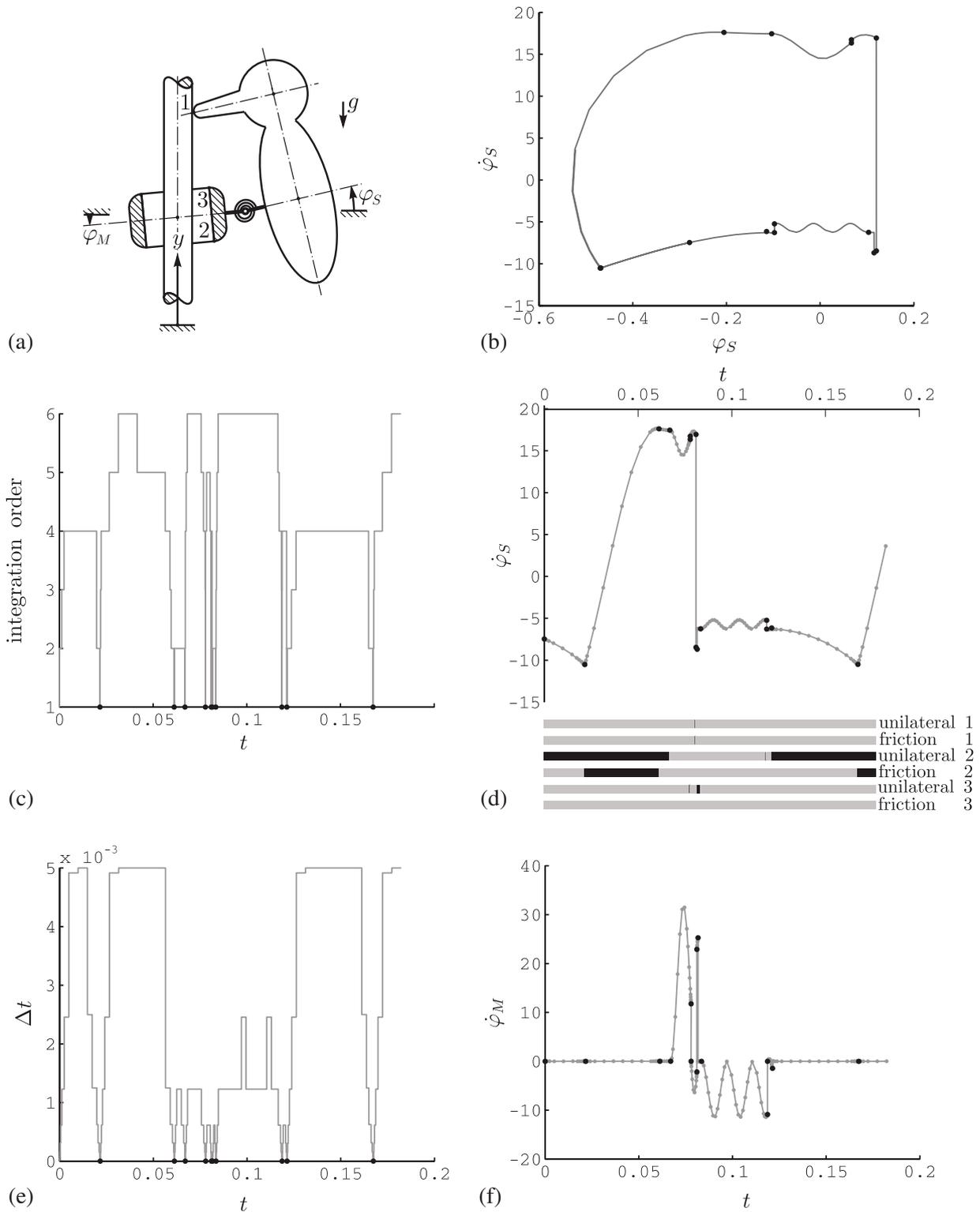


Figure 13. Woodpecker toy together with the results of the simulation. The time evolutions of $\dot{\varphi}_M$ and $\dot{\varphi}_S$ are complemented by a diagram, which shows the discrete state $\hat{\sigma}$ of the six different set-valued laws. Black parts correspond to the ‘constraint’ mode; grey parts correspond to the ‘impressed’ mode.

10.4. The woodpecker toy

The woodpecker toy consists of a pole, a sleeve with a hole that is slightly larger than the diameter of the pole, a spring and the woodpecker body. In operation, the woodpecker moves down the pole performing some kind of pitching motion, which is controlled by the sleeve [19]. The system is modelled by six set-valued laws, which describe the unilateral contact and friction element between beak and pole, the unilateral contact and friction element between the lower end of the sleeve and the pole and the unilateral contact and friction element between the upper end of the sleeve and the pole. The mechanical model as well as the initial conditions for the simulation can be found in [19]. The minimal and the maximal step sizes are chosen as $\Delta t_{\min} = 10^{-7}$ s and $\Delta t_{\max} = 5 \times 10^{-3}$ s, respectively. The maximal integration order is 6. In Figure 13(a) we see the model of the woodpecker [29]. In Figures 13(b)–(f), the phase diagram φ_S versus $\dot{\varphi}_S$ is provided together with the time evolution of $\dot{\varphi}_M$, $\dot{\varphi}_S$, Δt and of the used integration order. Switching points are depicted as small black dots. Note that the integration order and the step size decrease to 1 and Δt_{\min} , respectively, if a switching point is present. Smooth parts of the motion are run with a larger step size and a higher integration order. Around the time instant $t = 0.1$ s high oscillations are present and the step size does not reach Δt_{\max} because otherwise the integration order would be larger than the maximal user-defined order. The time evolutions of $\dot{\varphi}_M$ and $\dot{\varphi}_S$ are complemented by a diagram, which shows the discrete state $\hat{\mathbf{b}}$ of the six different set-valued laws. Black parts correspond to the ‘constraint’ mode, grey parts correspond to the ‘impressed’ mode. The different switching incidents are shown in Table VII. In analogy to the example of the single DOF impact oscillator in Section 10.3, the time evolution of the woodpecker has also been computed with a step size adjustment module, which uses a fixed integration order. In contrast to the single DOF impact oscillator, the same minimal step size $\Delta t_{\min} = 10^{-10}$ s has been chosen for all computations, as otherwise the maximal step size Δt_{\max} would have become too large. The different approximations have been compared with a reference solution that has been computed using an absolute error tolerance of $\text{atol}_\mu = 10^{-10}$ m/s. The minimal and maximal step sizes of the reference solution are $\Delta t_{\min} = 10^{-10}$ s and $\Delta t_{\max} = 9 \times 10^{-4}$ s, respectively. Figure 14 shows the double-logarithmic plot of the maximal time step Δt_{\max} versus the global integration error e_G of the coordinate φ_S . We recognize that the double-logarithmic plot consists of more or less straight lines with slope

Table VII. The different switching incidents of the woodpecker toy.

Contact		From	To
Friction contact	2	Slip	Stick
Friction contact	2	Stick	Slip
Unilateral contact	2	Closed	Open
Unilateral contact	3	Open	Closed
Unilateral contact	3	Closed	Open
Unilateral contact	1	Open	Closed
Unilateral contact	1	Closed	Open
Unilateral contact	3	Open	Closed
Unilateral contact	3	Closed	Open
Unilateral contact	2	Open	Closed
Unilateral contact	2	Closed	Open
Unilateral contact	2	Open	Closed

The corresponding switching points are shown in Figure 13.

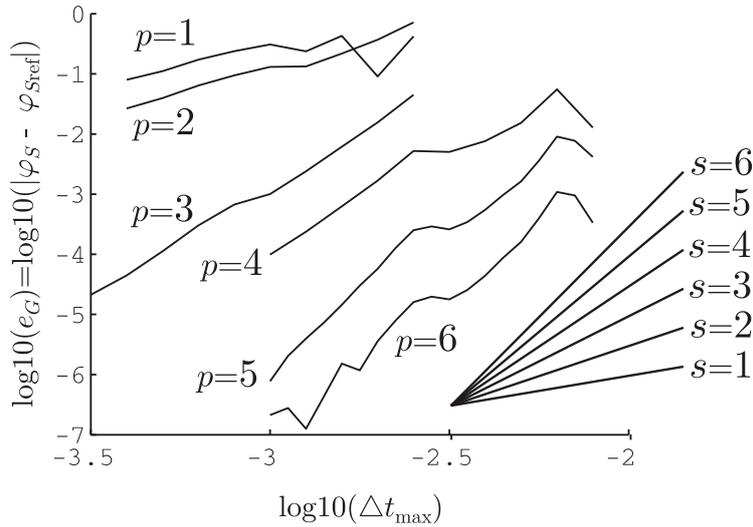


Figure 14. Double-logarithmic plot of the maximal step size Δt_{\max} versus the global integration error e_G of the coordinate φ_s .

$s = p$ even if the minimal step size is not equal to $\Delta t_{\min} = \Delta t_{\max}^p$. A possible explanation might be that the few local errors of the switching intervals can be neglected compared with the accumulation of the many local errors during the smooth part of the motion. In addition, the approximations are compared with a reference solution, which uses the same minimal step size $\Delta t_{\min} = 10^{-10}$ s.

11. CONCLUSIONS

In this paper we discussed how the concepts of step size adjustment and extrapolation can be used to improve the time-stepping method of Moreau. The presented algorithm provides switching point search and higher integration order in the smooth parts of the motion, but because the base scheme is still the time-stepping scheme, it is also suited for systems with more than a few contacts. An overall integration order of p with respect to the maximal step size Δt_{\max} can be gained if the minimal step size Δt_{\min} is chosen so small that the corresponding order 1 steps do not destroy the accuracy gained by the higher-order integration of the smooth time steps. There are two main ideas behind the presented algorithm. The first idea is to use extrapolation based on a time-stepping scheme. This allows us to use the same base integrator for all time steps. Of course, it would also be possible to implement an arbitrary higher-order integration scheme for the smooth parts of the motion, for example, a Runge–Kutta scheme, but in this case we still would have to apply the time-stepping scheme to manage time steps in which switching points occur. We have chosen extrapolation because at every time step we have as base a time-stepping scheme, i.e. if extrapolation or switching point detection would fail for some reasons, then we would still have a stable and robust integrator that processes the time step. The switching point detection can be combined straightforwardly with the extrapolation process, i.e. when doing the extrapolation substeps, one can monitor whether the mathematical structure on which extrapolation is based changes or not. When using the time-stepping scheme as the base scheme for the extrapolation, only an uneven number of extrapolation substeps has to be applied, as otherwise energy would increase in the case of enduring contact. The second idea is to observe the projection behaviour of the projective equations, which yields a very simple procedure for detecting switching points. Of

course, there exist other ways to detect switching points. For example, one could check whether the normal contact force of a unilateral contact becomes zero, or if the friction force belonging to a frictional contact becomes equal to the maximal friction force. There are two drawbacks when doing this. Firstly, each kind of set-valued law would require an own detection procedure, which would make the implementation cumbersome. Secondly, due to numerical inaccuracies, it would be, for example, difficult to tell whether a unilateral contact force is zero or not. Observing the projection behaviour of the projective equations provides a binary decision tool, which is independent of the kind of set-valued law, elegant in implementation and very robust in addition. Note that the integrator cares only about the changes in the mathematical structure. From our point of view, it is more natural to detect directly these structural changes than to monitor some related physical criteria. The presented algorithm is well suited for non-smooth systems with a small number of set-valued laws. By choosing a very small minimal step size, the algorithm tends toward the accuracy of an event-driven algorithm. Nevertheless, systems with many set-valued laws can also be treated by the same algorithm, because the base integration scheme is still a time-stepping method.

REFERENCES

1. Simo JC, Laursen TA. An augmented Lagrangian treatment of contact problems involving friction. *Computers and Structures* 1992; **42**(1):97–116.
2. Laursen TA. *Computational Contact and Impact Mechanics*. Springer: New York, 2002.
3. Klarbring A. A mathematical-programming approach to 3-dimensional contact problems with friction. *Computer Methods in Applied Mechanics and Engineering* 1986; **58**(2):175–200.
4. Wriggers P. *Computational Contact Mechanics* (2nd edn). Springer: Berlin, 2006.
5. Brogliato B. *Nonsmooth Impact Mechanics*. Lecture Notes in Control and Information Sciences. Springer: Berlin, Heidelberg, New York, 1996.
6. Brogliato B, ten Dam AA, Paoli L, Génot F, Abadie M. Numerical simulation of finite dimensional multibody nonsmooth mechanical systems. *Applied Mechanics Reviews* (ASME) 2002; **55**(2):107–150.
7. Eich-Soellner E, Führer C. *Numerical Methods in Multibody Dynamics*. B.G. Teubner: Stuttgart, 1998.
8. Eberhard P. *Kontaktuntersuchungen durch hybride Mehrkörpersystem/Finite Elemente Simulationen*. Shaker Verlag: Aachen, 2000.
9. Izadbakhsh A, McPhee J, Birkett S. Dynamic modeling and experimental testing of a piano action mechanism with a flexible hammer shank. *Proceedings of the IDETC/CIE-2007*, Las Vegas, U.S.A., 2007.
10. Muth B, Eberhard P. Investigation of large systems consisting of many spatial polyhedral bodies. *Proceedings of the ENOC Conference*, Eindhoven, The Netherlands, 2005.
11. Stronge WJ, James R, Ravani B. Oblique impact with friction and tangential compliance. *Philosophical Transactions of the Royal Society of London, Series A* 2001; **359**(1789):2447–2465.
12. Monteiro Marques MDP. *Differential Inclusions in Nonsmooth Mechanical Problems. Shocks and Dry Friction*. Progress in Nonlinear Differential Equations and their Applications, vol. 9. Birkhauser: Basel, 1993.
13. Moreau JJ. *Unilateral Contact and Dry Friction in Finite Freedom Dynamics*. Non-smooth Mechanics and Applications CISM Courses and Lectures, vol. 302. Springer: Wien, 1988.
14. Stewart D. Reformulations of measure differential inclusions and their closed graph property. *Journal of Differential Equations* 2001; **175**(1):108–129.
15. Anitescu M, Potra FA, Stewart DE. Time-stepping for three-dimensional rigid body dynamics. *Computer Methods in Applied Mechanics and Engineering* 1999; **177**(3–4):183–197.
16. Ebrahimi S, Eberhard P. Contact of planar deformable bodies using a linear complementarity formulation. *Proceedings in Applied Mathematics and Mechanics* (PAMM) 2005; **5**(1):197–198.
17. Förg M, Pfeiffer F, Ulbrich H. Simulation of unilateral constrained systems with many bodies. *Multibody System Dynamics* 2005; **14**(137–154):137–154.
18. Funk K. *Simulation eindimensionaler Kontinua mit Unstetigkeiten*. VDI-Fortschrittberichte Mechanik/Bruchmechanik, vol. 18/294. VDI-Verlag: Düsseldorf, 2004.

19. Glocker Ch, Studer C. Formulation and preparation for numerical evaluation of linear complementarity systems in dynamics. *Multibody System Dynamics* 2005; **13**(4):447–463.
20. Jean M. The non-smooth contact dynamics method. *Computer Methods in Applied Mechanics and Engineering* 1999; **177**(3–4):235–257.
21. Leine RI, Nijmeijer H. *Dynamics and Bifurcations of Non-smooth Mechanical Systems*. Lecture Notes in Applied and Computational Mechanics, vol. 18. Springer: Berlin, 2004.
22. Paoli L, Schatzman M. A numerical scheme for impact problems I: the one-dimensional case. *SIAM Journal on Numerical Analysis* 2002; **40**(2):702–733.
23. Potra FA, Anitescu M, Gavrea B, Trinkle J. A linearly implicit trapezoidal method for integrating stiff multibody dynamics with contacts, joints and friction. *International Journal for Numerical Methods in Engineering* 2006; **66**(7):1079–1124.
24. Stewart DE, Trinkle JC. An implicit time-stepping scheme for rigid body dynamics with inelastic collisions and Coulomb friction. *International Journal for Numerical Methods in Engineering* 1996; **39**(15):2673–2691.
25. Stieglmeier A. *Zur numerischen Berechnung strukturvarianter Mehrkörpersysteme*. VDI-Fortschrittberichte Mechanik/Bruchmechanik, vol. 18/271. VDI-Verlag: Düsseldorf, 2001.
26. Studer C, Glocker Ch. Simulation of non-smooth mechanical systems with many unilateral constraints. *Proceedings of the ENOC Conference*, Eindhoven, The Netherlands, 2005.
27. Studer C, Glocker Ch. Augmented time stepping by step-size adjustment and extrapolation. *Proceedings of the IDETC/CIE-2007*, Las Vegas, U.S.A., 2007.
28. Studer C. Augmented time-stepping integration of non-smooth dynamical systems. *Ph.D. Thesis*, ETH Zurich, Zurich, 2008.
29. Glocker Ch. *Dynamik von Starrkörpersystemen mit Reibung und Stößen*. VDI-Fortschrittberichte Mechanik/Bruchmechanik, vol. 18/182. VDI-Verlag: Düsseldorf, 1995.
30. Pfeiffer F, Glocker Ch. *Multibody Dynamics with Unilateral Contacts*. Wiley: New York, 1996.
31. Glocker Ch. *Set-valued Force Laws—Dynamics of Non-smooth Systems*. Lecture Notes in Applied Mechanics, vol. 1. Springer: Berlin, Heidelberg, 2001.
32. Glocker Ch. *Simulation von harten Kontakten mit Reibung: Eine iterative Projektionsmethode*. VDI-Berichte No. 1968: Schwingungen in Antrieben 2006 Tagung Fulda, vol. 1968. VDI-Verlag: Düsseldorf, 2006.
33. Elstrodt J. *Maß- und Integrationstheorie*. Springer: Berlin, Heidelberg, New York, 1996.
34. Moreau JJ. Bounded variation in time. *Topics in Nonsmooth Mechanics*. Birkhäuser Verlag: Basel, 1988; 1–74.
35. Dzonou R. A numerical scheme for finite dimensional frictionless dynamics with inelastic contact and a general inertia operator. *Proceedings of the ENOC-2005*, Eindhoven, 2005.
36. Facchinei F, Pang JS. *Finite-dimensional Variational Inequalities and Complementarity Problems*. Springer Series in Operations Research, vols I & II. Springer: New York, 2003.
37. Robinson SM. Generalized equations and their solution. I. Basic theory. *Mathematical Programming Study* 1979; **10**:128–141.
38. Bertsekas DP. *Constrained Optimization and Lagrange Multiplier Methods*. Academic Press: New York, 1982.
39. Rockafellar RT. Lagrange multipliers and variational inequalities. In *Variational Inequalities and Complementarity Problems*, Lions G, Cottle X (eds). Wiley: New York, 1979.
40. Alart P, Curnier A. A mixed formulation for frictional contact problems prone to Newton like solution methods. *Computer Methods in Applied Mechanics and Engineering* 1991; **92**(3):353–375.
41. Förg M, Geier T, Neumann L, Ulbrich H. r-Factor strategies for the augmented Lagrangian approach in multi-body contact mechanics. *Proceedings of the III European Conference on Computational Mechanics*, Lisbon, Portugal, 2006.
42. Jourdan F, Alart P, Jean M. A Gauss–Seidel like algorithm to solve frictional contact problems. *Computer Methods in Applied Mechanics and Engineering* 1998; **155**(1):31–47.
43. Studer C, Glocker Ch. Representation of normal cone inclusion problems in dynamics via nonlinear equations. *Archive of Applied Mechanics* 2006; **76**(5–6):327–348.
44. Studer C, Glocker Ch. Solving normal cone inclusion problems in contact mechanics by iterative methods. *Journal of System Design and Dynamics* 2007; **1**(3):458–467.
45. De Saxcé G, Feng ZQ. New inequality and functional for contact with friction: the implicit standard material approach. *Mechanics of Structures and Machines* 1991; **19**:301–325.
46. Murty KG. *Linear Complementarity, Linear and Nonlinear Programming*. Heldermann: Berlin, 1988.
47. Hairer E, Norsett SP, Wanner G. *Solving Ordinary Differential Equations I*. Springer: Berlin, Heidelberg, New York, 2000.

48. Hairer E, Norsett SP, Wanner G. *Solving Ordinary Differential Equations II*. Springer: Berlin, Heidelberg, New York, 2000.
49. Schwarz HR. *Numerische Mathematik*. B.G. Teubner: Stuttgart, 1997.
50. Strehmel K, Weiner R. *Numerik gewöhnlicher Differentialgleichungen*. B.G. Teubner: Stuttgart, 1995.
51. Janin O, Lamarque CH. Comparison of several numerical methods for mechanical systems with impacts. *International Journal for Numerical Methods in Engineering* 2001; **51**(9):1101–1132.
52. Deuffhard P, Hairer E, Zugck J. One-step and extrapolation methods for differential–algebraic systems. *Numerische Mathematik* 1987; **51**(5):501–516.
53. Lubich C. Linearly implicit extrapolation methods for differential–algebraic-systems. *Numerische Mathematik* 1989; **55**(2):197–211.