

Variational collision integrators in forward dynamics and optimal control

Sigrid Leyendecker*, Carsten Hartmann[⊥], Michael Koch*, Gwen Johnson[#], Michael Ortiz[#]

* Chair of Applied Dynamics, University of Erlangen-Nuremberg, Germany

[⊥] Institut für Mathematik, Freie Universität Berlin, Germany

[#] Computational Solid Mechanics Group, California Institute of Technology

Abstract

The numerical simulation of multibody dynamics often involves constraints of various forms. First of all, we present a structure preserving integrator for mechanical systems with holonomic (bilateral) and unilateral contact constraints, the latter being in the form of a non-penetration condition. The scheme is based on a discrete variant of Hamilton's principle in which both the discrete trajectory and the unknown collision time are varied (cf. [Fete 03]). As a consequence, the collision event enters the discrete equations of motion as an unknown that has to be computed on-the-fly whenever a collision is imminent. The additional bilateral constraints are efficiently dealt with employing a discrete null space reduction (including a projection and a local reparametrisation step) which considerably reduces the number of unknowns and improves the condition number during each time-step as compared to a standard treatment with Lagrange multipliers (cf. [Leye 12]).

In previous works, discrete mechanics and optimal control for constrained systems (DMOCC) has been introduced for the structure preserving simulation of optimal control problems for rigid multibody systems, whereby possible contacts or collisions between the bodies have been disregarded see [Leye 10]. In the formulation presented here, both collision avoidance as well as explicitly planned collisions between non-smooth bodies are included. To this end, a subdifferentiable global contact detection algorithm, the supporting separating hyperplane linear program (SSHLP), based on the signed distance between supporting hyperplanes of two convex sets, is used in the simulation of optimal control problems.

1 Introduction

In the first part of this work (Sections 2 and 3) we present an variational algorithm that handles equality and inequality constraints on the same footing, see [Leye 12]. Our approach closely follows the route described in [Fete 03] where a seamless variational integrator for collision problems is introduced; by 'seamless' we mean that both the integration scheme for the free motion and the contact conditions at the points of collision follow from a single variational principle. Here we extend the seamless variational scheme to the case where, besides the non-penetration condition for a chain of beads, additional holonomic constraints are present during the entire motion. Although our method is an ordinary time-stepping scheme, it shares some of the spirit of event-driven methodology in that it combines an implicit integrator for the free motion of the chain that allows for large time-steps with an exact treatment of the collision events. In particular, the collision times are computed on-the-fly by solving a simple quadratic equation.

For smooth systems, variational integrators with constant time-step preserve various properties of the exact dynamics such as symplecticity or momentum maps (e.g., linear or angular momentum) at the discrete level [Wend 97, Mars 01]. The key idea is simple: rather than discretising the continuous-time equations of motion, variational integrators directly discretise Hamilton's principle by appropriate quadrature rules; the numerical scheme then follows from the now discrete Euler-Lagrange equations. Variational integrators have proven useful, e.g., in understanding the long-term stability of symplectic integrators using backward error analysis [Hair 06]. Moreover, the (discrete) variational principle allows for easy generalisations so as to treat, e.g., infinite-dimensional systems [Brid 06], systems with constraints [Jay 07], or contact problems in continuum mechanics [Cira 05].

The standard molecular dynamics algorithm for systems with holonomic constraints, that is in fact a variational integrator, is the SHAKE/RATTLE algorithm [Ryck 77, Ande 83] which can be considered an augmented version of the Verlet algorithm where the constraints are enforced by suitable Lagrange multipliers.

As an alternative to Lagrange multipliers that may cause stability problems when the integration time-steps are small (e.g., close to a collision point) the discrete null space method eliminates the constraint force and the Lagrange multipliers by projecting the forces in the system onto the space of admissible momenta. It does so by taking advantage of the d'Alembert principle that states that the constraint forces are always acting perpendicular to the constraint manifold and therefore have no component in the tangential direction. The discrete null space method has been introduced in conjunction with an energy-momentum conserving time integration scheme in [Bets 05, Bets 06] and has been transferred to variational time-stepping schemes in [Leye 08]. The new idea in this work is to use the discrete null space methodology also for the treatment of the contact constraints while they are active.

When simulating the optimal control problems for three-dimensional multibody systems, the treatment of contact always imposes a challenge. One simple solution is to formulate a smooth problem where the bodies are allowed to overlap by a certain amount, which is penalised via a penalty potential. Of course, the drawbacks of inadmissible configurations and inexact contact forces that go along with this approach are obvious. However, for the forward dynamics simulation of many applications, the results might be accurate enough, in particular when high penalty parameters in combination with small time steps are tolerable from the computational effort's point of view. If the bodies are not supposed to overlap, the discrete event of contact has to be considered as part of the system's dynamics, thus making the dynamics non-smooth. See e.g. [Cira 05, Gloc 00, Lein 03, Pand 02, Pang 96] and many references on contact formulations in different contexts therein. Non-smooth formulations require the specification of a contact force that (for elastic collisions) reflects the momentum normal to a contact surface at a given time and configuration. In the context of a structure preserving time integration method (being consistent in the evolution of energy and momentum maps) that uses a predefined equidistant time grid, there is no other known procedure except for resolving the collisions in the sense that each contact time (which is likely between the nodes), configuration, and force are exactly computed, see [Fete 03, Leye 12].

The second part of this work (Sections 4 and 5) will show how the described alternatives for the treatment of collisions can be included in the context of optimal control problems. In the smooth formulation with a penalty potential, there arises some difficulty for the optimiser to distinguish between a contact and a control force, since both might point into the same direction. For the non-smooth treatment, the optimal control problem formulation gives more freedom than the forward dynamics problem, since periodic boundary conditions or leaving the exact placement of time nodes (within certain bounds) to the optimiser gives the freedom to assume that contact takes place at a certain time node without loss of generality, i.e. without fixing its physical time. A further challenge is the detection of contact for non-smooth non-convex three-dimensional bodies where a new strategy based on a supporting separating hyperplane linear programming (SSHLP) approach is used, see [John 12]. One major advantage of this strategy is that the subgradient of the SSHLP, supplying the direction of the contact force, can be readily evaluated. Reconfiguration manoeuvres with collision avoidance and with planned collisions are considered as examples.

2 Hamilton's principle for collision problems

Our formulation of the variational collision integrator follows the route taken in [Fete 03] and extends it to the case of a system that, besides a non-penetration condition, is subject to holonomic constraints. For a better understanding of the approach, it is instructive to look at the continuous formulation first. Let $Q \subseteq \mathbb{R}^n$ denote the n -dimensional configuration space of our system where n equals three times the number of particles. We suppose that the system is subject to m holonomic constraints $\mathbf{g}(\mathbf{q}) = \mathbf{0}$ with $\mathbf{g} = (g_1, \dots, g_m)^T$ being the vector of constraints with the requirement that $\mathbf{G} = \nabla \mathbf{g}$ has maximum rank m on the admissible set of configurations, $C = \{\mathbf{q} \in Q : \mathbf{g}(\mathbf{q}) = \mathbf{0}\} \subset Q$, and with $\det \mathbf{G}\mathbf{G}^T \geq a > 0$ being bounded away from zero.¹ The m_c non-penetration conditions for the particles can be expressed in terms of a vector of smooth unilateral constraints $\mathbf{g}_c(\mathbf{q}) \geq \mathbf{0}$ by which the set of admissible configurations turns out to be

$$C_+ = \{\mathbf{q} \in Q : \mathbf{g}(\mathbf{q}) = \mathbf{0}, \mathbf{g}_c(\mathbf{q}) \geq \mathbf{0}\} \subset C.$$

Calling

$$TQ = \{(\mathbf{q}, \mathbf{v}) : \mathbf{q} \in Q, \mathbf{v} \in T_{\mathbf{q}}Q\}$$

¹The latter requires that the function \mathbf{g} has a certain regularity that we can safely assume for most cases of interest.

the state space of the system consisting of the unconstrained positions and velocities, we define the Lagrangian (we use the natural identification of all tangent spaces $T_q Q$ with \mathbb{R}^n)

$$L: TQ \rightarrow \mathbb{R}, \quad L(\mathbf{q}, \mathbf{v}) = \frac{1}{2} \mathbf{v} \cdot \mathbf{M} \mathbf{v} - V(\mathbf{q})$$

with $\mathbf{M} \in \mathbb{R}^{n \times n}$ being the symmetric and positive definite mass matrix and V being a smooth potential energy. Now let $\gamma: [0, T] \rightarrow Q$ be a curve in Q that is everywhere twice continuously differentiable except at an isolated impact point $\mathbf{q}_l = \gamma(t_l)$, $t_l \in (0, T)$ where the curve γ hits the boundary ∂C_+ of the admissible set (i.e., where exactly one of the components of \mathbf{g}_c is zero). The classical action is then of the form

$$S[\gamma, t_l] = \int_0^{t_l} \hat{L}(\gamma(t), \dot{\gamma}(t)) dt + \int_{t_l}^T \hat{L}(\gamma(t), \dot{\gamma}(t)) dt \quad (1)$$

where

$$\hat{L}(\mathbf{q}, \mathbf{v}) = L(\mathbf{q}, \mathbf{v}) - \mathbf{g}(\mathbf{q}) \cdot \boldsymbol{\lambda} \quad (2)$$

is the augmented Lagrangian involving the constraints \mathbf{g} and the Lagrange multiplier $\boldsymbol{\lambda} \in \mathbb{R}^m$ (cf. Lagrange multiplier theorem e.g. in [Zeid 95]). Taking variations with respect to both γ and the unknown collision time t_l , we find

$$\begin{aligned} \delta S[\gamma, t_l] &= \delta \left(\int_0^{t_l} \hat{L}(\gamma(t), \dot{\gamma}(t)) dt + \int_{t_l}^T \hat{L}(\gamma(t), \dot{\gamma}(t)) dt \right) \\ &= \int_0^T \left(\frac{\partial \hat{L}}{\partial \boldsymbol{\gamma}} - \frac{d}{dt} \frac{\partial \hat{L}}{\partial \dot{\boldsymbol{\gamma}}} \right) \cdot \delta \boldsymbol{\gamma} dt - \left[\frac{\partial \hat{L}}{\partial \dot{\boldsymbol{\gamma}}} \cdot \delta \boldsymbol{\gamma} + \hat{L} \delta t_l \right]_{t_l^-}^{t_l^+}. \end{aligned} \quad (3)$$

Requiring that the integral vanishes, yields Newton's law

$$\mathbf{M} \ddot{\mathbf{q}} = -\nabla V(\mathbf{q}) - \mathbf{G}(\mathbf{q})^T \boldsymbol{\lambda}, \quad \mathbf{g}(\mathbf{q}) = \mathbf{0} \quad (4)$$

for the motion away from the collision. The remaining boundary terms provide the contact conditions at the point of impact. To make this precise we call g_c the scalar component of \mathbf{g}_c that is zero at the moment of collision (cf. 3.2 for the treatment of multiple collisions at a time) and note that $g_c(\gamma(t_l)) = 0$ entails $\delta g_c(\gamma(t_l)) = 0$, i.e.,

$$G_c(\gamma(t_l)) \cdot (\delta \mathbf{q}_l + \dot{\boldsymbol{\gamma}}(t_l) \delta t_l) = 0$$

where $\mathbf{q}_l = \gamma(t_l)$ denotes the impact point. The last equation is satisfied if either

$$\delta \mathbf{q}_l = -\dot{\boldsymbol{\gamma}}(t_l) \delta t_l$$

or

$$\delta \mathbf{q}_l \perp G_c(\mathbf{q}_l) \quad \text{for} \quad \delta t_l = 0$$

which determines the admissible variations of the curve at the collision point. Taking joint variations of \mathbf{q}_l and t_l under the constraint $\delta \mathbf{q}_l = -\dot{\boldsymbol{\gamma}}(t_l) \delta t_l$ implies

$$\left[\hat{L} - \frac{\partial \hat{L}}{\partial \dot{\boldsymbol{\gamma}}} \cdot \dot{\boldsymbol{\gamma}} \right]_{t_l^-}^{t_l^+} = 0 \quad (5)$$

which, using that

$$E = \frac{\partial \hat{L}}{\partial \dot{\boldsymbol{\gamma}}} \cdot \dot{\boldsymbol{\gamma}} - \hat{L},$$

is equivalent to conservation of energy E . Conversely, if we vary \mathbf{q}_l orthogonal to the contact surface while keeping t_l fixed (i.e., $\delta t_l = 0$), then

$$\left[\frac{\partial \hat{L}}{\partial \dot{\boldsymbol{\gamma}}} \cdot \delta \boldsymbol{\gamma} \right]_{t_l^-}^{t_l^+} = 0 \quad \text{for} \quad \delta \boldsymbol{\gamma}(t_l) \perp G_c(\mathbf{q}_l)$$

implies

$$\left[\frac{\partial \hat{L}}{\partial \dot{\boldsymbol{\gamma}}} \right]_{t_{\iota}^-}^{t_{\iota}^+} = G_c^T \lambda_c. \quad (6)$$

Here $\lambda_c \in \mathbb{R}$ is an unknown Lagrange multiplier that must be determined by solving (5). Defining the conjugate momentum $\boldsymbol{p} = \partial L / \partial \dot{\boldsymbol{\gamma}}$ in the usual way, equations (5) and (6) can be recast as

$$\boldsymbol{p}(t_{\iota}^+) = \boldsymbol{p}(t_{\iota}^-) + G_c^T(\boldsymbol{q}_{\iota}) \lambda_c, \quad \boldsymbol{p}(t_{\iota}^+) \cdot \boldsymbol{M}^{-1} \boldsymbol{p}(t_{\iota}^+) = \boldsymbol{p}(t_{\iota}^-) \cdot \boldsymbol{M}^{-1} \boldsymbol{p}(t_{\iota}^-) \quad (7)$$

But the last equation is simply the momentum reflection law for an elastic collision. That is, the change in momentum occurs normal to the contact surface where the amount by which the momentum changes is determined from the conservation of energy during the collision (note that the potential V is not needed to determine the unknown multiplier λ_c). As a consequence of (7), both total linear and angular momentum and the total energy are conserved during the collision.

We stress that the contact condition (7) follows seamlessly from the boundary terms arising in (3). In the next section we will show, again following the approach described in [Fete 03], how the discrete variant of (1) naturally gives rise to a fully variational collision integrator.

3 Variational collision integrator for constrained problems

Our derivation below follows closely the route taken in [Fete 03], but extends it to the case of systems subject to holonomic constraints. The constraints are treated fully variationally using the discrete null space method that has been introduced in [Leye 08]. The new idea here is to use the discrete null space methodology also for the treatment of the contact forces while they are active.

3.1 Discrete variational principle and discrete null space reduction

Assume that the time nodes $t_0, t_1, \dots, t_{\iota-1}, t_{\iota+1}, \dots, t_N$ with a constant basic time-step $h = t_{n+1} - t_n$ are given, however the collision time t_{ι} with $t_{\iota-1} \leq t_{\iota} \leq t_{\iota+1}$ is unknown. Let the discrete trajectory be denoted by $\boldsymbol{q}_d = \{\boldsymbol{q}_n\}_{n=0}^N$ with $\boldsymbol{q}_n \approx \boldsymbol{q}(t_n)$, and let $\boldsymbol{\lambda}_d = \{\boldsymbol{\lambda}_n\}_{n=0}^N$ with $\boldsymbol{\lambda}_n \approx \boldsymbol{\lambda}(t_n)$ approximate the Lagrange multipliers. As usual in the context of discrete variational principles (e.g., see [Mars 01]), the discrete Lagrangian is an approximation to the action integral of the continuous Lagrangian over one time interval. In accordance with (2), this yields away from a collision

$$\hat{L}_d(\boldsymbol{q}_n, \boldsymbol{q}_{n+1}, t_n, t_{n+1}) \approx \int_{t_n}^{t_{n+1}} L(\boldsymbol{q}, \boldsymbol{v}) - \boldsymbol{g}(\boldsymbol{q}) \cdot \boldsymbol{\lambda} dt$$

In this work, a midpoint discrete Lagrangian

$$L_d(\boldsymbol{q}_n, \boldsymbol{q}_{n+1}, t_n, t_{n+1}) = (t_{n+1} - t_n) L \left(\frac{\boldsymbol{q}_n + \boldsymbol{q}_{n+1}}{2}, \frac{\boldsymbol{q}_{n+1} - \boldsymbol{q}_n}{t_{n+1} - t_n} \right) \quad (8)$$

is used in the following discrete augmented Lagrangian

$$\hat{L}_d(\boldsymbol{q}_n, \boldsymbol{q}_{n+1}, t_n, t_{n+1}) = L_d(\boldsymbol{q}_n, \boldsymbol{q}_{n+1}, t_n, t_{n+1}) - \frac{t_{n+1} - t_n}{2} (\boldsymbol{g}(\boldsymbol{q}_n) \cdot \boldsymbol{\lambda}_n + \boldsymbol{g}(\boldsymbol{q}_{n+1}) \cdot \boldsymbol{\lambda}_{n+1}).$$

For the clarity of exposition, the dependence of the discrete Lagrangian on given time nodes is not stated explicitly in the sequel and it is assumed that only one collision $g_c(\boldsymbol{q}_{\iota}) = 0$ occurs at t_{ι} during the time interval $[t_{\iota-1}, t_{\iota+1}]$; see Figure 1 for an illustration. The extension to multiple collisions is formally straightforward, however more involved from the implementation point of view (see 3.2). An approximation to (1) is given

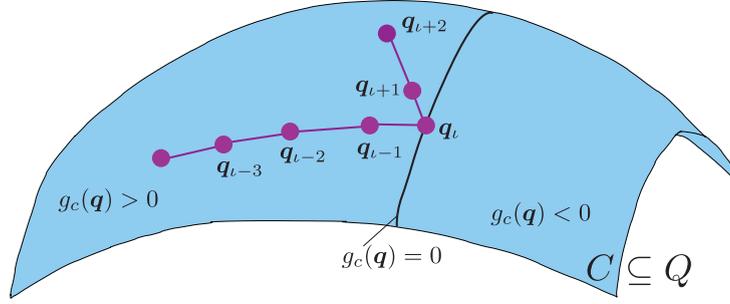


Figure 1. Discrete trajectory in constraint manifold.

by the discrete action sum

$$\begin{aligned}
S_d = & \sum_{n=0}^{\iota-2} L_d(\mathbf{q}_n, \mathbf{q}_{n+1}) - \frac{t_{n+1} - t_n}{2} (\mathbf{g}(\mathbf{q}_n) \cdot \boldsymbol{\lambda}_n + \mathbf{g}(\mathbf{q}_{n+1}) \cdot \boldsymbol{\lambda}_{n+1}) \\
& + L_d(\mathbf{q}_{l-1}, \mathbf{q}_l, t_{l-1}, t_l) - \frac{t_l - t_{l-1}}{2} (\mathbf{g}(\mathbf{q}_{l-1}) \cdot \boldsymbol{\lambda}_{l-1} + \mathbf{g}(\mathbf{q}_l) \cdot \boldsymbol{\lambda}_l + g_c(\mathbf{q}_l) \cdot \lambda_c) \\
& + L_d(\mathbf{q}_l, \mathbf{q}_{l+1}, t_l, t_{l+1}) - \frac{t_{l+1} - t_l}{2} (\mathbf{g}(\mathbf{q}_l) \cdot \boldsymbol{\lambda}_l + \mathbf{g}(\mathbf{q}_{l+1}) \cdot \boldsymbol{\lambda}_{l+1} + g_c(\mathbf{q}_l) \cdot \lambda_c) \\
& + \sum_{n=\iota+1}^{N-1} L_d(\mathbf{q}_n, \mathbf{q}_{n+1}) - \frac{t_{n+1} - t_n}{2} (\mathbf{g}(\mathbf{q}_n) \cdot \boldsymbol{\lambda}_n + \mathbf{g}(\mathbf{q}_{n+1}) \cdot \boldsymbol{\lambda}_{n+1}).
\end{aligned} \tag{9}$$

As before, the discrete variational principle for the constrained motion requires that $\delta S_d = 0$ for all admissible variations $\delta \mathbf{q}_1, \dots, \delta \mathbf{q}_l, \dots, \delta \mathbf{q}_{N-1}, \delta \boldsymbol{\lambda}_1, \dots, \delta \boldsymbol{\lambda}_N, \delta \lambda_c, \delta t_l$. This then yields discrete equations of motion for the dynamics off the contact surface with additional boundary conditions at the collision points. The equations for the collision-free motion are stated first:

Pre- and post-collision As long as no collision takes place, i.e., for $n = 1, \dots, \iota-2$ the discrete variational principle yields the following system

$$D_2 L_d(\mathbf{q}_{n-1}, \mathbf{q}_n) + D_1 L_d(\mathbf{q}_n, \mathbf{q}_{n+1}) - \frac{t_{n+1} - t_{n-1}}{2} \mathbf{G}^T(\mathbf{q}_n) \boldsymbol{\lambda}_n = \mathbf{0} \tag{10a}$$

$$\mathbf{g}(\mathbf{q}_{n+1}) = \mathbf{0} \tag{10b}$$

which is solved for $\mathbf{q}_2, \dots, \mathbf{q}_{l-1}, \boldsymbol{\lambda}_1, \dots, \boldsymbol{\lambda}_{l-2}$. Here, $D_i L_d$ denotes the derivative of the discrete Lagrangian with respect to the i -th argument. Note that just as in the continuous case described in Section 2, there exists a Lagrange multiplier theorem relating stationary points of the discrete action (9) to the finite dimensional system (10), see [Mars 01]. Once the collision configuration \mathbf{q}_l and time t_l and the first post-collision configuration \mathbf{q}_{l+1} have been determined (as will be described below), normal time-stepping continues for $n = \iota + 1, \dots, N - 1$, i.e., (10) is solved for $\mathbf{q}_{l+2}, \dots, \mathbf{q}_N, \boldsymbol{\lambda}_{l+1}, \dots, \boldsymbol{\lambda}_{N-1}$. Obviously, (10) is a two-step method, thus special care must be given to the initialisation of the simulation. Assuming that no collision takes place during the first time-step, the following equations determine \mathbf{q}_1 and $\boldsymbol{\lambda}_0$ from given initial data $(\mathbf{q}(0), \dot{\mathbf{q}}(0)) \in TC$. First, one sets $\mathbf{q}_0 = \mathbf{q}(0)$ and $\mathbf{p}_0^- = (\partial L / \partial \dot{\mathbf{q}})(\mathbf{q}(0), \dot{\mathbf{q}}(0))$ and then solves

$$\begin{aligned}
\mathbf{p}_0^- + D_1 L_d(\mathbf{q}_0, \mathbf{q}_1) - \frac{t_1 - t_0}{2} \mathbf{G}^T(\mathbf{q}_0) \boldsymbol{\lambda}_0 = \mathbf{0} \\
\mathbf{g}(\mathbf{q}_1) = \mathbf{0}
\end{aligned}$$

Collisions in the first time-step can be handled analogously to later collisions described below.

Discrete null space reduction Details on the reduction of the discrete variational equations of motion with constraints via the discrete null space method with local reparametrisation can be found in [Leye 08].

The main idea is the elimination of the constraint forces from the discrete system via the premultiplication with an appropriate null space matrix $\mathbf{P}(\cdot) : \mathbb{R}^{n-m} \rightarrow TC$, i.e., the null space matrix fulfils $\text{range } \mathbf{P}(\mathbf{q}_n) = \ker \mathbf{G}(\mathbf{q}_n)$. Secondly, a local reparametrisation of the constraint manifold $\mathbf{q}_{n+1} = \mathbf{F}_d(\mathbf{u}_{n+1}, \mathbf{q}_n) \in C$ in terms of the discrete generalised coordinates $\mathbf{u}_{n+1} \in \mathbb{R}^{n-m}$ representing the system's change during one time-step ensures that the constraints are fulfilled and (10b) becomes superfluous. The reduced equations read

$$\mathbf{P}^T(\mathbf{q}_n) [D_2 L_d(\mathbf{q}_{n-1}, \mathbf{q}_n) + D_1 L_d(\mathbf{q}_n, \mathbf{F}_d(\mathbf{u}_{n+1}, \mathbf{q}_n))] = \mathbf{0} \quad (11)$$

Away from the collision, they are solved for $\mathbf{u}_2, \dots, \mathbf{u}_{\ell-1}$ and $\mathbf{u}_{\ell+2}, \dots, \mathbf{u}_N$ while in the very first step,

$$\mathbf{P}^T(\mathbf{q}_0) [\mathbf{p}_0^- + D_1 L_d(\mathbf{q}_0, \mathbf{F}_d(\mathbf{u}_1, \mathbf{q}_0))] = \mathbf{0}$$

is solved for \mathbf{u}_1 . In contrast to an absolute parametrisation in generalised coordinates with respect to the initial configuration \mathbf{q}_0 reading $\mathbf{q}_{n+1} = \mathbf{F}_d(\mathbf{u}_{n+1}, \mathbf{q}_0)$, locality of the discrete parametrisation avoids singularities present, e.g., when dealing with large rotations. The described procedure reduces the $(d+m)$ -dimensional system (10) to the $(d-m)$ -dimensional system (11). Depending on the particular problem under consideration, this can reduce the computational costs substantially. Due to the elimination of the Lagrange multipliers from the set of unknowns, the well known condition problem associated with discretisations of index 3 DAEs is removed. While the condition number of the Jacobian matrix in the linearisation of (10) is of the order h^{-3} , the corresponding condition number in (11) is independent of the time-step (see [Leye 08]). Note that after solving (11), the Lagrange multipliers can always be determined as a post-processing step if one is interested in the constraint forces. This is particularly important when the contact forces themselves are eliminated using the discrete null space reduction (cf. equation (15) below).

Collision Before integrating forward in a new time-step, the contact inequality condition is checked. If it is violated, i.e., if $g_c(\mathbf{q}_\ell) < 0$, then \mathbf{q}_ℓ is discarded and the collision configuration \mathbf{q}_ℓ and time t_ℓ as well as $\lambda_{\ell-1}$ are determined by solving

$$D_2 L_d(\mathbf{q}_{\ell-2}, \mathbf{q}_{\ell-1}) + D_1 L_d(\mathbf{q}_{\ell-1}, \mathbf{q}_\ell, t_{\ell-1}, t_\ell) - \frac{t_\ell - t_{\ell-2}}{2} \mathbf{G}^T(\mathbf{q}_{\ell-1}) \lambda_{\ell-1} = \mathbf{0} \quad (12a)$$

$$\mathbf{g}(\mathbf{q}_\ell) = \mathbf{0} \quad (12b)$$

$$g_c(\mathbf{q}_\ell) = 0 \quad (12c)$$

which have been obtained from $\delta S_d = 0$ in (9) for variations $\delta \mathbf{q}_{\ell-1}, \delta \lambda_\ell, \delta \lambda_c$. After that, taking admissible variations $\delta \mathbf{q}_\ell, \delta \lambda_{\ell+1}, \delta t_\ell$ yields

$$D_2 L_d(\mathbf{q}_{\ell-1}, \mathbf{q}_\ell, t_{\ell-1}, t_\ell) + D_1 L_d(\mathbf{q}_\ell, \mathbf{q}_{\ell+1}, t_\ell, t_{\ell+1}) - \frac{t_{\ell+1} - t_{\ell-1}}{2} (\mathbf{G}^T(\mathbf{q}_\ell) \lambda_\ell + \mathbf{G}_c^T(\mathbf{q}_\ell) \lambda_c) = \mathbf{0} \quad (13a)$$

$$\mathbf{g}(\mathbf{q}_{\ell+1}) = \mathbf{0} \quad (13b)$$

$$D_4 L_d(\mathbf{q}_{\ell-1}, \mathbf{q}_\ell, t_{\ell-1}, t_\ell) + D_3 L_d(\mathbf{q}_\ell, \mathbf{q}_{\ell+1}, t_\ell, t_{\ell+1}) = 0 \quad (13c)$$

from which $\mathbf{q}_{\ell+1}, \lambda_\ell, \lambda_c$ follow. Note that (13c) is resulting from the variation with respect to the time node t_ℓ , thus it is a conservation condition for the discrete energy.

Discrete null space reduction Equivalent to (12), the reduced system yielding \mathbf{u}_ℓ, t_ℓ reads

$$\mathbf{P}^T(\mathbf{q}_{\ell-1}) [D_2 L_d(\mathbf{q}_{\ell-2}, \mathbf{q}_{\ell-1}) + D_1 L_d(\mathbf{q}_{\ell-1}, \mathbf{q}_\ell, t_{\ell-1}, t_\ell)] = \mathbf{0} \quad (14a)$$

$$g_c(\mathbf{q}_\ell) = 0 \quad (14b)$$

and instead of solving (13), the unknown $\mathbf{u}_{\ell+1}$ can be obtained from

$$\mathbf{P}_c^T(\mathbf{q}_\ell) \mathbf{P}^T(\mathbf{q}_\ell) [D_2 L_d(\mathbf{q}_{\ell-1}, \mathbf{q}_\ell, t_{\ell-1}, t_\ell) + D_1 L_d(\mathbf{q}_\ell, \mathbf{q}_{\ell+1}, t_\ell, t_{\ell+1})] = \mathbf{0} \quad (15a)$$

$$D_4 L_d(\mathbf{q}_{\ell-1}, \mathbf{q}_\ell, t_{\ell-1}, t_\ell) + D_3 L_d(\mathbf{q}_\ell, \mathbf{q}_{\ell+1}, t_\ell, t_{\ell+1}) = 0 \quad (15b)$$

As mentioned before, the Lagrange multipliers and in particular the contact forces $\mathbf{G}_c^T(\mathbf{q}_\ell) \lambda_c$ can be computed as a post-processing step.

3.2 Multiple collisions

If multiple collisions occur at the same time t_l , i.e. $\mathbf{g}_c(\mathbf{q}) \geq \mathbf{0} \in \mathbb{R}^{m_c}$ is really vector valued with $m_c \in \mathbb{R}$, then $(\sum_{k=1}^{m_c} g_{c_k}(\mathbf{q}_l)) \lambda_c$ is used in the discrete action (9). Note that corresponding to the single unknown collision time t_l , the Lagrange multiplier λ_c is scalar. Variation with respect to λ_c yields $(\sum_{k=1}^{m_c} g_{c_k}(\mathbf{q}_l)) = 0$ in (12c) which, together with the positiveness of all components, is equivalent to $g_{c_k}(\mathbf{q}_l) = 0$ for $k = 1, \dots, m_c$. Accordingly, $(\sum_{k=1}^{m_c} \mathbf{G}_{c_k}^T(\mathbf{q}_l)) \lambda_c$ appears in (13a) and then, as for single collisions, the discrete energy conservation condition for the complete system (13c) determines λ_c . The discrete null space reduction works analogously to the single collision case.

If collisions follow each other in quick succession, the algorithm does not return to the regular grid immediately but computes another collision time node within a basic time-step h .

3.3 Example: chain of four beads in a box

As an example, we consider a linear chain of four beads (of the same mass, length and radius) in a three-dimensional box of size $8 \times 10 \times 14$ (relative to the bead size with radius $r = 0.1$). Depending on the initial conditions, planar or real three-dimensional, periodic or non-periodic motion takes place where multiple collisions (between multiple beads or between multiple beads and multiple walls) occur. We consider an example of non-integrable three-dimensional motion. The initial conditions

$$\mathbf{q}_0 = \left(-\frac{3}{2\sqrt{2}}, -\frac{1}{2\sqrt{2}}, 0, -\frac{1}{2\sqrt{2}}, \frac{1}{2\sqrt{2}}, 0, \frac{1}{2\sqrt{2}}, -\frac{1}{2\sqrt{2}}, 0, \frac{3}{2\sqrt{2}}, \frac{1}{2\sqrt{2}}, 0\right)$$

$$\dot{\mathbf{q}}_0 = \left(-\frac{1}{10}, \frac{1}{10}, -\frac{1}{10}, 0, 0, 0, 0, 0, 0, 0, 0, 0\right)$$

lead to chaotic motion of the chain. Figure 3 illustrates different configurations, where beads in contact with other beads are coloured in blue while the red beads are in contact with a wall of the box. Note that the third row depicts a bead-bead collision happening a very short time before a bead-wall collision and the fourth row shows a bead-wall collision followed immediately by a bead-bead collision. In particular, the time between the described collisions is shorter than the basic time-step $h = 0.1$. Obviously, the algorithm is capable to handle multiple collisions (at one time node) and collisions following each other rapidly in a stable way exhibiting good energy behaviour (fluctuations are of the order of magnitude 10^{-4} , they do not show any drift, see Figure 2).

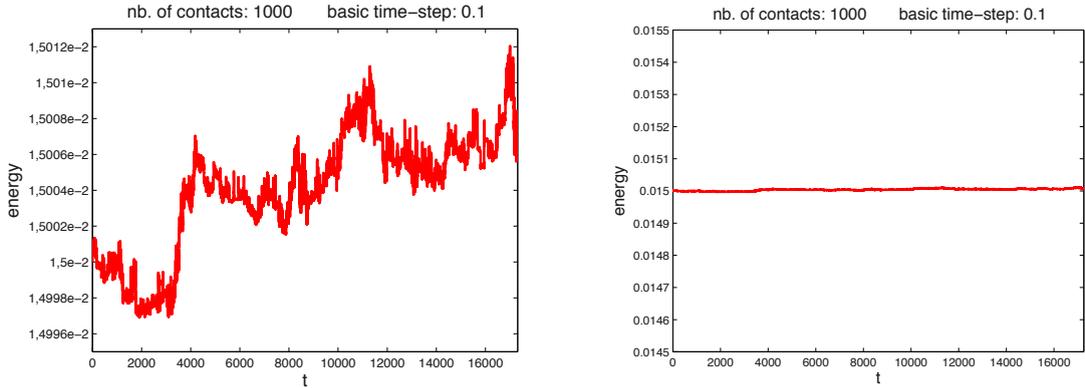


Figure 2. Chain of four beads, chaotic motion: evolution of energy during a simulation of 1000 collisions, including various collisions following each other in quick succession.

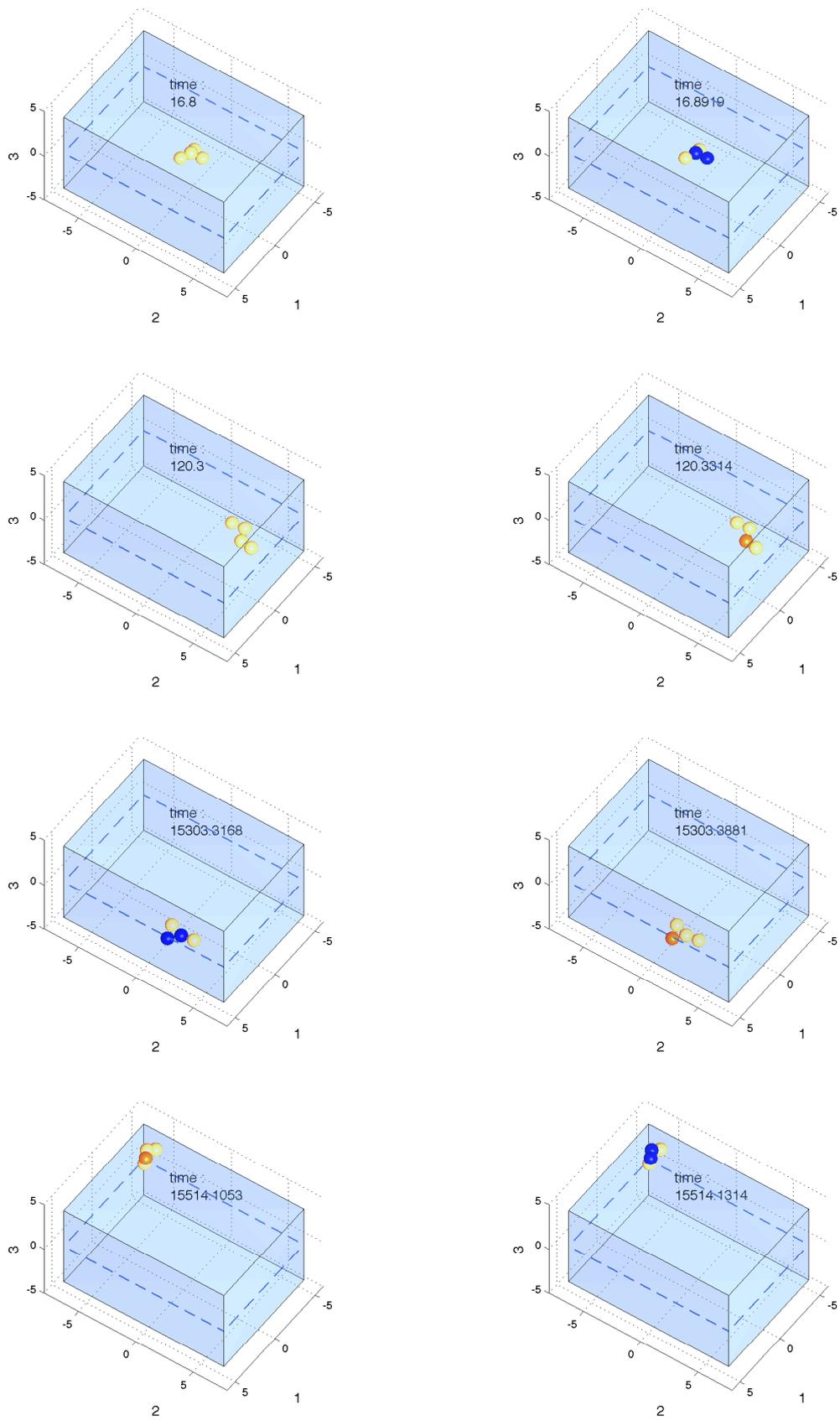


Figure 3. Chain of four beads, chaotic motion: different configurations including bead-bead and bead-wall collisions following each other rapidly.

4 Optimal control with collision avoidance

In order to avoid collisions or even any contact between bodies (including grazing collisions, i.e. the mere touching contact without a velocity component in the direction normal to the contact surface), the first step is to detect contact or overlapping. This can be quite a challenge for three-dimensional non-smooth and maybe non-convex bodies. Then, the second step is to enforce the avoidance of collisions in the simulation of optimal control problems.

4.1 Contact detection for non-smooth bodies via SSHLP

To detect contact between convex bodies, a subdifferentiable global contact detection algorithm, the so called supporting separating hyperplane (SSH) algorithm, developed in [John 12] is used. Based on theorems from convex and affine geometry, this algorithm determines the signed distance between supporting hyperplanes of two convex sets. In its first formulation, the algorithm yields a quadratically constrained linear program. However, the problems's structure is such that an equivalent linear program (LP) (with explicit expressions of the subderivative supplying the contact force) can always be formulated.

For the positive integer n , let $\alpha \neq \mathbf{0} \in \mathbb{R}^n$ and $a \in \mathbb{R}$, then the affine and convex set

$$H_{\alpha,a} = \{\mathbf{x} \in \mathbb{R}^n \mid \langle \alpha, \mathbf{x} \rangle - a = 0\}$$

is called a hyperplane in \mathbb{R}^n . Here, α is the normal vector to the plane while $a = \langle \alpha, \mathbf{a} \rangle$ for some point $\mathbf{a} \in \mathbb{R}^n$ belonging to the plane. The signed distance between a point $\mathbf{y} \in \mathbb{R}^n$ and a hyperplane can be computed as $\langle \alpha, \mathbf{y} \rangle - a$. Clearly, a hyperplane has two distinct sides. For $\mathbf{b} \in \mathbb{R}^n$ and $b = \langle \alpha, \mathbf{b} \rangle$, the signed distance between parallel hyperplanes with normal α reads $d(H_{\alpha,a}, H_{\alpha,b}) = b - a$.

Let $K_1, K_2 \subset \mathbb{R}^n$ be compact convex sets and let $\text{ext}K_1, \text{ext}K_2$ denote the sets of their extreme points (corners for polyhedral sets) and $\text{bd}K_1, \text{bd}K_2$ their boundaries, respectively. A hyperplane $H_{\alpha,a}$ is said to support the set K_1 at $\mathbf{x} \in \text{bd}K_1$, when $\mathbf{x} \in H_{\alpha,a}$ and K_1 is entirely contained in one of the half-spaces associated with distinct sides of $H_{\alpha,a}$, i.e. for all $\mathbf{y} \in K_1$ either $\langle \alpha, \mathbf{y} \rangle - a \leq 0$ or $\langle \alpha, \mathbf{y} \rangle - a \geq 0$ holds.

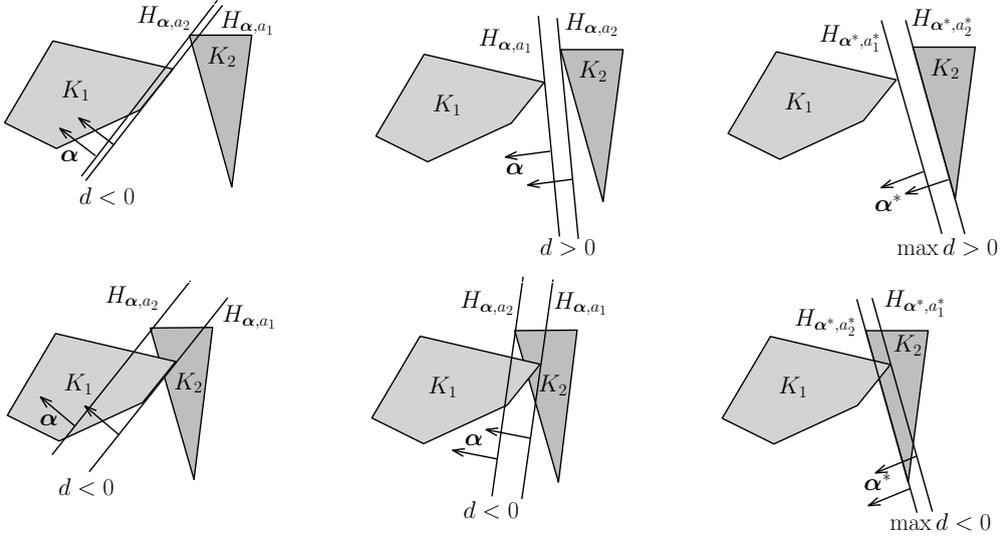


Figure 4. Strictly separable (first row) and non-separable (second row) polyhedral sets K_1, K_2 with signed distance between different supporting hyperplanes and maximum signed distance (Figure from [John 12]).

The maximum signed distance between parallel supporting hyperplanes of each set can be computed either as a solution $h(K_1, K_2)$ of the following quadratically constrained linear program (QCLP), or as the solution $g(K_1, K_2)$ of the following linearly constrained linear program (LCLP) for some unit vector $\beta \in S^{n-1}$.

QCLP	LCLP
$h(K_1, K_2) = \max_{\substack{\alpha \in \mathbb{R}^n \\ a_1, a_2 \in \mathbb{R}}} a_1 - a_2$	$g(K_1, K_2) = \max_{\substack{\alpha \in \mathbb{R}^n \\ a_1, a_2 \in \mathbb{R}}} a_1 - a_2$
subject to $\langle \alpha, x \rangle - a_1 \geq 0 \quad \forall x \in \text{ext}K_1$ $\langle \alpha, y \rangle - a_2 \leq 0 \quad \forall y \in \text{ext}K_2$ $\langle \alpha, \alpha \rangle = 1$	subject to $\langle \alpha, x \rangle - a_1 \geq 0 \quad \forall x \in \text{ext}K_1$ $\langle \alpha, y \rangle - a_2 \leq 0 \quad \forall y \in \text{ext}K_2$ $\langle \beta, \alpha \rangle = 1$

Using of the following three results shown in [John 12]

- i) **Theorem 1.** *Two compact convex sets K_1 and K_2 are strictly separable, i.e. $K_1 \cap K_2 = \emptyset$ (properly separable, i.e. $K_1 \cap K_2 = \emptyset$ or $K_1 \cap K_2 \subset H_{\alpha, a_1} \equiv H_{\alpha, a_2}$) if and only if $h > 0$ (if and only if $h \geq 0$).*
- ii) **Theorem 2,** giving conditions under which an optimal solution of the QCLP is equivalent to an optimal solution of the LCLP
- iii) **Remark 3** and remarks following the proof of Theorem 2 on the choice of β in praxis

we solve the LCLP to detect contact or overlapping between non-smooth convex bodies. If non-convex bodies are present, they are subdivided into convex bodies for which then contact or overlapping is tested.

4.2 Discrete mechanics and optimal control for constrained systems with collision avoidance

The goal of an optimal control problem is to determine an optimal state trajectory with the corresponding optimal control trajectory actuating the dynamical system such that an optimality criterion is reached. Thus, an objective functional is minimised with respect to the state and control trajectory while the equations of motion, initial and final conditions as well as path constraints have to be fulfilled. To simulate an optimal control problem numerically via a direct method, the problem is transformed into a finite dimensional constrained optimisation problem which can be solved e.g. by a standard SQP or interior point method. Thereby, the particular form of the discrete equations of motion (which serve as constraints for the optimisation) plays a crucial role for the resulting approximate solution. As described in the previous sections, in this discrete mechanics approach, they are derived via a discrete variational principle, leading to structure preservation (symplecticity, consistency in momentum maps in the presence of symmetry and good energy behaviour) along the discrete trajectories, see [Mars 01, Leye 08, Leye 10].

Let the degrees of freedom are actuated by the discrete control sequence $\tau_d = \{\tau_n\}_{n=0}^{N-1}$, approximating the control trajectory $\tau(t) : [t_0, t_N] \rightarrow \mathbb{R}^{n-m}$ by a constant $\tau_n \in \mathbb{R}^{n-m}$ in each time interval $[t_n, t_{n+1}]$. Using $\tau_{n-1}^+ = \frac{h}{2}\tau_{n-1}$ and $\tau_n^- = \frac{h}{2}\tau_n$, redundant n -dimensional control forces $B^T(q_n) \cdot (\tau_{n-1}^+ + \tau_n^-)$ can be computed using the configuration dependent input transformation matrix $B^T(q_n) \in \mathbb{R}^{n \times (n-m)}$. In the presence of non conservative actuation forces and constraints, a discrete constrained Lagrange-d'Alembert principle yields the variational integrator – in this case the constrained forced discrete Euler-Lagrange equations. Altogether, discrete mechanics and optimal control for constrained systems (DMOCC) yields the following finite dimensional constrained optimisation problem for the simulation of the optimal control problem with boundary values $(q(t_0), \dot{q}(t_0)) = (q^0, \dot{q}^0)$ and $(q(t_N), \dot{q}(t_N)) = (q^N, \dot{q}^N)$. Note that J_d, C_d are discrete approximations to the objective and cost functionals, respectively, while s_d, h_d, r_d are discrete versions of the initial, final and path constraints.

Let the contact constraint function be denoted by $g_c : Q \rightarrow \mathbb{R}$ and assume that it is computed as the solution of the LCLP. Then collision avoidance requires to augment DMOCC by the inequality constraint $g_c(q_n) > 0$ for $n = 1, \dots, N-1$.

DMOCC with collision avoidance	
minimise discrete objective function	
	$\min_{\mathbf{u}_d, \boldsymbol{\tau}_d} J_d(\mathbf{u}_d, \boldsymbol{\tau}_d) = \min_{\mathbf{u}_d, \boldsymbol{\tau}_d} \sum_{n=0}^{N-1} C_d(\mathbf{u}_n, \mathbf{u}_{n+1}, \boldsymbol{\tau}_n)$
subject to the constraints for $n = 1, \dots, N - 1$	
discrete equations of motion	$\mathbf{P}^T(\mathbf{q}_n) \cdot [D_1 L_d(\mathbf{q}_n, \mathbf{F}(\mathbf{u}_{n+1}, \mathbf{q}_n)) + D_2 L_d(\mathbf{q}_{n-1}, \mathbf{q}_n) + \mathbf{B}^T(\mathbf{q}_n) \cdot (\boldsymbol{\tau}_{n-1}^+ + \boldsymbol{\tau}_n^-)] = \mathbf{0}$
collision avoidance	$g_c(\mathbf{q}_n) > 0$
initial value constraints	$\mathbf{s}_d(\mathbf{u}_0, \mathbf{u}_1, \boldsymbol{\tau}_0, \mathbf{q}^0, \dot{\mathbf{q}}^0) = \mathbf{0}$
path constraints	$\mathbf{h}_d(\mathbf{u}_n, \mathbf{u}_{n+1}, \boldsymbol{\tau}_n) \leq \mathbf{0}$
final point constraints	$\mathbf{r}_d(\mathbf{u}_{N-1}, \mathbf{u}_N, \boldsymbol{\tau}_{N-1}, \mathbf{q}^N, \dot{\mathbf{q}}^N) = \mathbf{0}$

The constrained optimisation problem is solved in Matlab using the `fmincon` function (SQP algorithm with active set strategy or interior point method).

4.3 Example: puzzle assembly with collision avoidance

The considered puzzle consists of three non-smooth non-convex rigid bodies that are initially at rest in fully specified initial configurations. All bodies are fully actuated and supposed to reconfigure such that at the final time $t_N = 0.29$ (where $h = 0.01$ and $N = 29$), they are in a prescribed relative placement and orientation to each other. In other words, the puzzle manoeuvre ends at rest in a fully assembled configuration whose absolute placement and orientation in space is left free. During the manoeuvre, contact is detected using the SSHLP described in Section 4.1 and collisions are avoided. The initial guess for the optimisation is determined via an inverse dynamics problem, i.e. collision free reconfiguration trajectories have been guessed for the three bodies and the corresponding actuation has been determined solving the discrete equations of motion (see DMOCC with collision avoidance in Section 4.2) for the discrete control sequence $\boldsymbol{\tau}_d$. The goal of the optimisation is to minimise the control effort, i.e. $C_d = h t \boldsymbol{\tau}_n^T \cdot \boldsymbol{\tau}_n$ while bounds on the optimisation variables $\mathbf{u}_d, \boldsymbol{\tau}_d$ ensure that a local minimum with relatively small displacements is found. The control effort is reduced from a value of $J_d = 89$ at the initial guess to $J_d = 0.0630$ for the optimised solution. See Figure 5 for snapshots of different configurations during the puzzle's assembly.

5 Optimal control problems with contact

In many optimal control tasks like e.g. docking manoeuvres or walking and jumping motion, contact can not be avoided or is even necessary. However, in most cases, it is not known in advance when or where contact takes place. In particular, one can optimise the contact time and configuration with respect to specified goals.

5.1 Smooth penalty formulation

A penalty potential is an easy way to treat contact without losing the generality of the problem formulation while retaining its smoothness, i.e. in DMOCC with collision avoidance described in Section 4.2, the collision avoidance inequality condition is left away. In case that the bodies overlap, i.e. $g_c(\mathbf{q}_n) < 0$, an n -dimensional penalty force $\mathbf{f}_n^c = \nabla(\frac{\mu}{2} g_c^2(\mathbf{q}_n))$ with the positive penalty parameter $\mu \in \mathbb{R}$ is included in the discrete equations of motion in DMOCC, which then read

$$\mathbf{P}^T(\mathbf{q}_n) \cdot [D_1 L_d(\mathbf{q}_n, \mathbf{F}(\mathbf{u}_{n+1}, \mathbf{q}_n)) + D_2 L_d(\mathbf{q}_{n-1}, \mathbf{q}_n) + \mathbf{B}^T(\mathbf{q}_n) \cdot (\boldsymbol{\tau}_{n-1}^+ + \boldsymbol{\tau}_n^-) + \mathbf{f}_n^c] = \mathbf{0}$$

Note that an explicit expression for the subderivative of the SSHLP can be readily evaluated such that ∇g_c can always be computed.

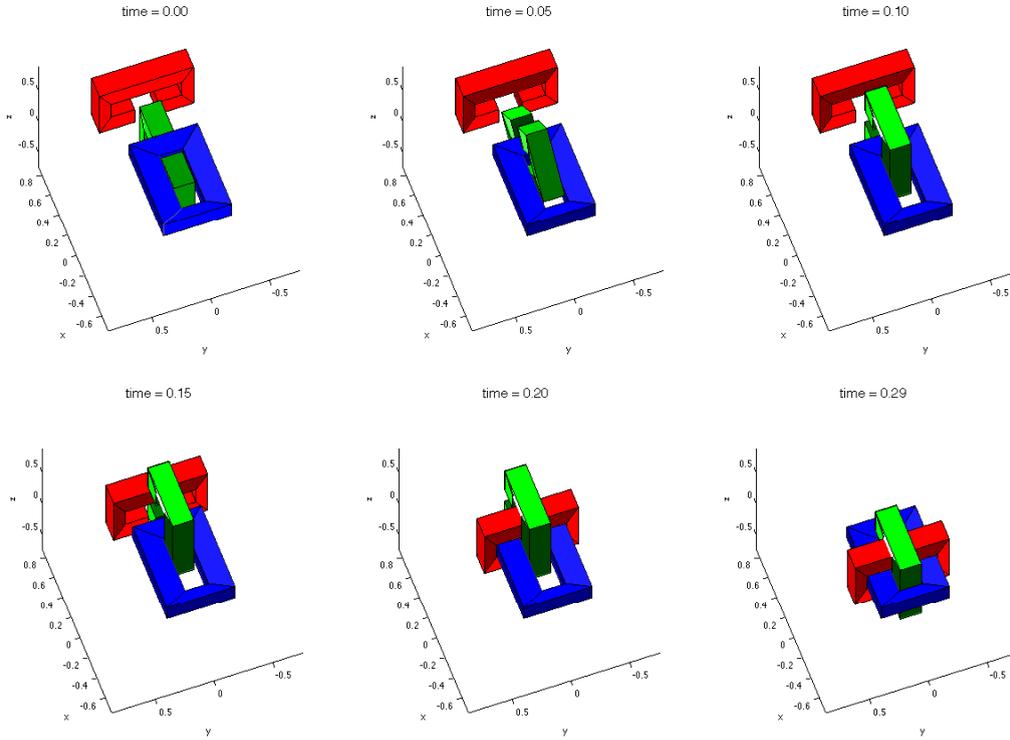


Figure 5. Puzzle assembly with collision avoidance: snapshots of different configurations.

5.2 Non-smooth contact formulation

Suppose that a mechanical system is to be optimally controlled from an initial to a final state in such a way that precisely one collision takes place during the manoeuvre at a time node, say at $t_\iota \in [t_0, t_N]$ with $g_c(\mathbf{q}_\iota) = 0$. If the physical contact time was prescribed, the problem formulation would lose its generality. To retain generality, we let the contact time be determined as an optimisation variable. One way to realise this is to introduce positive scaling factors $\sigma_1, \sigma_2 \in \mathbb{R}$ for the time step before and after the contact time, see Figure 6. To ensure that time steps do not degenerate, the scaling factors can be bounded in the path constraints. Furthermore, the total manoeuvre time $t_N = (\iota\sigma_1 + (N - \iota)\sigma_2)ht$ can either be left free, or bounded or fixed by a scalar valued function $w(\sigma_1, \sigma_2, t_N) \leq 0$ as desired.

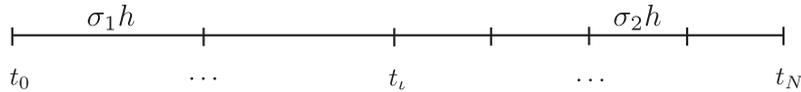


Figure 6. Time grid with differently scaled time step $\sigma_1 ht$ and $\sigma_2 ht$ before and after the contact time node t_ι .

The most crucial point in treating the contact is to determine the correct change in momentum due to the collision. Lets first assume that the collision is perfectly elastic and frictionless, i.e. the kinetic energy (and therewith the total energy) does not change when computed immediately before and after the impact. When a point mass hits a surface, its linear momentum component in the direction of the surface normal at the contact point is reversed while the tangential component remains unchanged. However, the present situation is more complex since we deal with systems of rigid bodies having translational and rotational degrees of freedom on the one hand, and on the other hand, their kinematics is described using a constrained formulation reflecting their rigidity and possible interconnections by joints. These configuration constraints

give rise to hidden constraints on velocity and on momentum level. Using the following discrete reduced Legendre transforms

$$\begin{aligned} {}^P \mathbf{p}_n^- &= \mathbf{P}^T(\mathbf{q}_n) \cdot [-D_1 L_d(\mathbf{q}_n, \mathbf{F}(\mathbf{u}_{n+1}, \mathbf{q}_n)) - \mathbf{B}^T(\mathbf{q}_n) \cdot \boldsymbol{\tau}_n^-] \\ {}^P \mathbf{p}_n^+ &= \mathbf{P}^T(\mathbf{q}_n) \cdot [D_2 L_d(\mathbf{q}_{n-1}, \mathbf{q}_n) + \mathbf{B}^T(\mathbf{q}_n) \cdot \boldsymbol{\tau}_{n-1}^+] \end{aligned}$$

the discrete equations of motion in DMOCC can be interpreted as ${}^P \mathbf{p}_n^+ - {}^P \mathbf{p}_n^- = \mathbf{0}$ which is a balance of discrete momenta at t_n , where the discrete momenta have been computed from the past (${}^P \mathbf{p}_n^+$) and the following time interval (${}^P \mathbf{p}_n^-$), respectively, and both are mapped by the null space matrix into the appropriate reduced cotangent space (being consistent with the constraints on momentum level). When a contact configuration has been reached, i.e. $g_c(\mathbf{q}_t) = 0$, the component of ${}^P \mathbf{p}_t^+$ in the direction of the surface normal has to be reflected. Note that in the reduced formulation, this surface normal is given by $({}^P \nabla g_c(\mathbf{q}_t))^T = \mathbf{P}^T(\mathbf{q}_t) \cdot (\nabla g_c(\mathbf{q}_t))^T$. Thus, using the reduced mass matrix ${}^P \mathbf{M}(\mathbf{q}_t) = \mathbf{P}^T(\mathbf{q}_t) \cdot \mathbf{M} \cdot \mathbf{P}(\mathbf{q}_t)$ the component of ${}^P \mathbf{p}_t^+$ in the direction of the surface tangent, denoted by $({}^P \mathbf{p}_t^+)_{\parallel}$, fulfills

$$\left\langle ({}^P \nabla g_c(\mathbf{q}_t))^T, ({}^P \mathbf{p}_t^+)_{\parallel} \right\rangle_{({}^P \mathbf{M}(\mathbf{q}_t))^{-1}} = 0 \quad (16)$$

where $\langle \mathbf{a}, \mathbf{b} \rangle_{\mathbf{A}} = \mathbf{a}^T \cdot \mathbf{A} \cdot \mathbf{b}$ is a norm for the vectors \mathbf{a}, \mathbf{b} and the square matrix \mathbf{A} of appropriate dimension. Inserting the momentum decomposition into the tangential and the normal direction ${}^P \mathbf{p}_t^+ = ({}^P \mathbf{p}_t^+)_{\parallel} + ({}^P \mathbf{p}_t^+)_{\perp}$ into Equation (16), the normal momentum component can be found explicitly via

$$({}^P \mathbf{p}_t^+)_{\perp} = \frac{\left\langle ({}^P \nabla g_c(\mathbf{q}_t))^T, {}^P \mathbf{p}_t^+ \right\rangle_{({}^P \mathbf{M}(\mathbf{q}_t))^{-1}}}{\left\langle ({}^P \nabla g_c(\mathbf{q}_t))^T, ({}^P \nabla g_c(\mathbf{q}_t))^T \right\rangle_{({}^P \mathbf{M}(\mathbf{q}_t))^{-1}}} ({}^P \nabla g_c(\mathbf{q}_t))^T$$

Finally, for frictionless collisions with a coefficient of restitution $e \in [0, 1]$, where $e = 1$ represents a perfectly elastic and $e = 0$ a perfectly plastic collision, the post collision momentum can be computed according to the following momentum reflection in normal direction at the contact surface

$${}^P \mathbf{p}_{t,post}^+ = {}^P \mathbf{p}_{t,pre}^+ - (1 + e) ({}^P \mathbf{p}_{t,pre}^+)_{\perp}$$

With these preliminaries, the DMOCC with contact problem can be formulated as follows.

DMOCC with contact	
minimise discrete objective function	
$\min_{\mathbf{u}_d, \boldsymbol{\tau}_d, \sigma_1, \sigma_2}$	$J_d(\mathbf{u}_d, \boldsymbol{\tau}_d, \sigma_1, \sigma_2) = \min_{\mathbf{u}_d, \boldsymbol{\tau}_d} \sum_{n=0}^{N-1} C_d(\mathbf{u}_n, \mathbf{u}_{n+1}, \boldsymbol{\tau}_n, \sigma_1, \sigma_2)$
subject to constraints for $n = 1, \dots, N - 1$	
discrete equations of motion	${}^P \mathbf{p}_n^+ - {}^P \mathbf{p}_n^- = \mathbf{0}$
contact	$g_c(\mathbf{q}_t) = 0$
momentum reflection	${}^P \mathbf{p}_{t,pre}^+ - (1 + e) ({}^P \mathbf{p}_{t,pre}^+)_{\perp} = {}^P \mathbf{p}_{t,post}^+$
initial value constraints	$\mathbf{s}_d(\mathbf{u}_0, \mathbf{u}_1, \boldsymbol{\tau}_0, \dot{\mathbf{q}}^0) = \mathbf{0}$
path constraints	$\mathbf{h}_d(\mathbf{u}_n, \mathbf{u}_{n+1}, \boldsymbol{\tau}_n, \sigma_1, \sigma_2) \leq \mathbf{0}$
final point constraints	$\mathbf{r}_d(\mathbf{u}_{N-1}, \mathbf{u}_N, \boldsymbol{\tau}_{N-1}, \dot{\mathbf{q}}^N) = \mathbf{0}$
total time condition	$w(\sigma_1, \sigma_2, t_N) \leq 0$

In particular, collision avoidance can be required at all other time nodes in the path constraints vector. If multiple, say $N_c \in \mathbb{N}$, collisions are planned at the time node numbers t_1, \dots, t_{N_c} , then multiple contact conditions $g_c(\mathbf{q}_{t_1}) = 0, \dots, g_c(\mathbf{q}_{t_{N_c}}) = 0$ with the corresponding momentum reflection conditions are constraining the optimisation in DMOCC with contact. Furthermore, $N_c + 1$ scaling factors are present in the total time condition $w(\sigma_1, \dots, \sigma_{N_c+1}, t_N) \leq 0$.

ι	4	9	14
σ_1	2.1795	0.9653	0.6186
σ_2	0.8185	1.0149	1.3337
t_ι	0.2180	0.2172	0.2165
J_d	$1.0244 \cdot 10^{-6}$	$1.0194 \cdot 10^{-6}$	$1.0005 \cdot 10^{-6}$

Table 1. Cube hitting a wall: values of scaling factors, contact time and control effort for different contact node numbers ι .

5.3 Example: cube hitting the wall

In the first example, a cube starts in a prescribed initial state with a translational velocity towards a close by wall. In the prescribed final state, the final velocity is reversed and bounds on the actuation are set such that the final state can not be reached without a collision of the cube with the wall. Here, $N = 30$ time nodes and a time step of $ht = 0.025$ are used and the control effort is minimised. The final time is prescribed as $t_N = 0.75$. This example serves as numerical evidence that the resulting optimal state and control trajectories and the resulting optimal contact time are independent of the chosen contact node number. Table 1 and Figure 7 shows that for $\iota \in \{4, 9, 14\}$ qualitatively the same contact time, control effort and evolution of actuating force and torque are obtained (of course the scaling factors yielding this contact time are different).

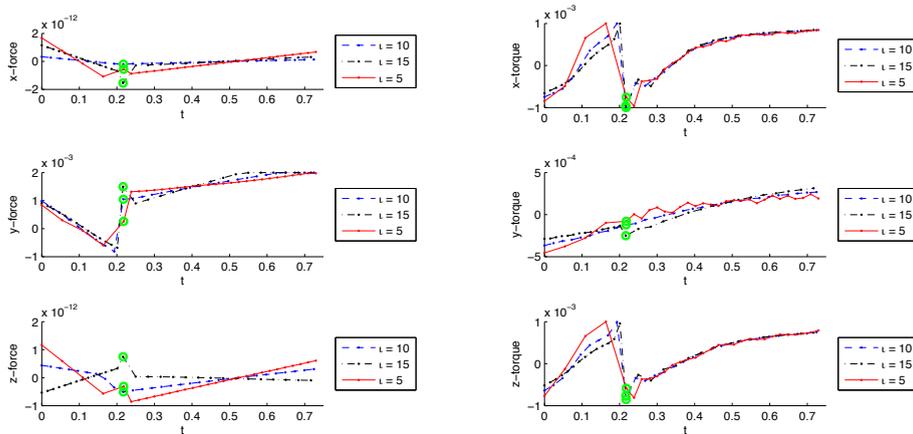


Figure 7. Cube hitting a wall: evolution of actuating force and torque for different contact node numbers ι . The contact time is marked in green.

5.4 Example: two cubes with two planned collisions

The following problem considers two cubes with fully specified initial configurations. The green cube is initially at rest. It is not actuated, thus its final state is not prescribed. The blue cube has an initial velocity and is required to be steered into a prescribed final rest position that overlaps with the green cube's initial configuration. Thus, the blue cube must push the green cube out of the way, wherefore two collisions are planned at the time node numbers $\iota_1 = 20, \iota_2 = 39$ and collisions are avoided at all other time nodes. Altogether, $N = 50$ time nodes and $ht = 0.01$ are used, while the final time is prescribed as $t_N = 0.5$, thus the total time condition $w(\sigma_1, \sigma_2, \sigma_3, t_N) = (\sigma_1 \iota_1 + \sigma_2(\iota_2 - \iota_1) + \sigma_3(N - \iota_2))ht - t_N = 0$ is used in DMOCC with contact described in Section 5.2. As a result of the control problem minimising the control effort to $J_d = 2.8573 \cdot 10^{-5}$, collisions happen at $t_{\iota_1} = 0.1975, t_{\iota_2} = 0.3900$, thus $\sigma_1 = 0.98725, \sigma_2 = 1.0132, \sigma_3 = 1.0004$. The initial guess for the optimisation is obtained via a forward dynamics simulation (treating contact according to the decomposition contact response formulation given in [Cira 05], which has been extended for the case of oriented bodies and constrained dynamics) yielding also a guess for possible

contact times. The left hand side plot in Figure 8 shows the evolution of the actuating force and torque for the blue cube. The corresponding net torque evolution as well as the evolution of angular momentum for the complete system are shown in the right hand side plot, where the lowest plot shows the consistency of the angular momentum evolution, i.e. the change in angular momentum exactly represents the actuation (up to the numerical tolerance) giving numerical evidence of the structure preservation properties of the discrete equations of motion. Snapshots of different configurations are depicted in Figure 9.

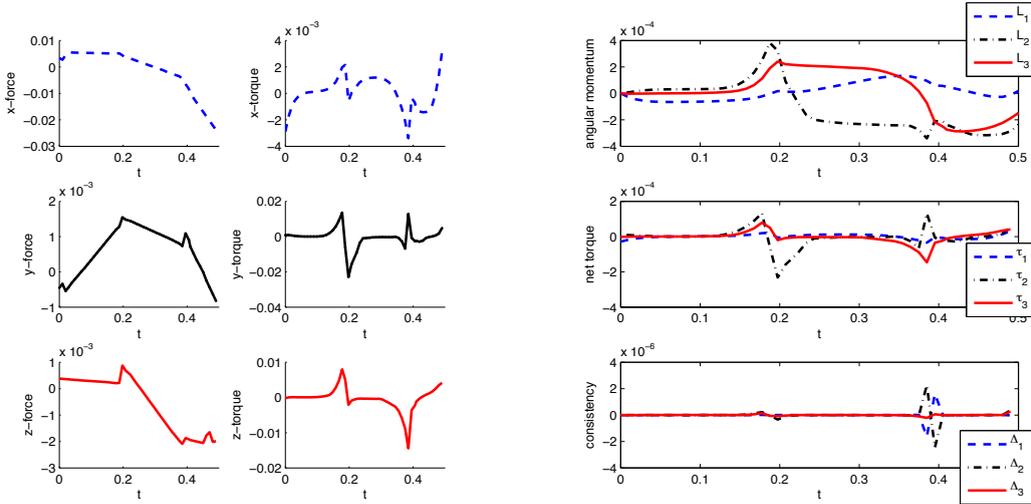


Figure 8. Two cubes with two planned collisions: evolution of actuating force and torque on blue cube (left) and evolution of angular momentum, net torque and consistency of angular momentum for the complete system (right).

6 Conclusion

Starting from a discrete version of Hamilton’s principle, we have derived a structure preserving integrator for mechanical systems including holonomic (bilateral) constraints as well as unilateral contact (inequality) constraints. The basic idea of the discrete variational principle for systems involving collisions is to treat both the discrete-time trajectory and unknown collision time as unknowns; the resulting discrete Euler-Lagrange equations plus boundary conditions then lead to a time-stepping algorithm that includes the contact force and the constraint forces corresponding to the holonomic constraints. As we have shown all reaction forces can be eliminated from the discrete equations of motion using a discrete null space reduction (involving a projection and a local reparametrisation step) which considerably reduces the number of unknowns as compared to the standard treatment with Lagrange multipliers. As a consequence, the condition number during iteration stays low. This fact, together with the possibility to determine the contact time by solving an algebraic equation rather than searching it via bisection strategies, leads to much lower computational costs while, at the same time, it increases the accuracy to which the contact time is determined.

With the supporting separating hyperplane linear programming (SSHLP) approach, a very efficient method for detecting contact between non-smooth bodies has been described briefly. The resulting signed distance between the bodies is first of all used as an inequality constraint in optimal control simulations with collision avoidance. Secondly, optimal control manoeuvres with planned contacts and collisions are considered. Here it is an advantage that the subgradient of the SSHLP, supplying the direction of the contact force, can be readily evaluated. For the simulation of contact manoeuvres, the time grid is variable, such that the contact time can be determined as part of the optimisation while the contact node number can be fixed, what facilitates the implementation substantially. The presented formulation and results constitute first steps of ongoing work and many theoretical aspects and more complex examples will be considered in the future. Furthermore, a comparison of the SSHLP’s efficiency with standard contact detection methods is required.

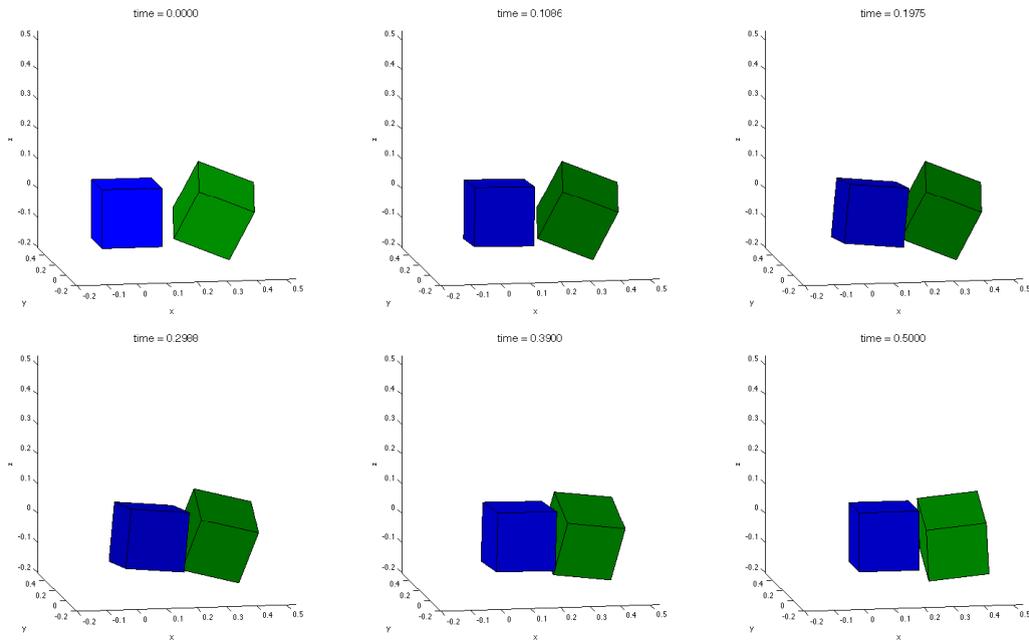


Figure 9. Two cubes with two planned collisions: snapshots of different configurations.

REFERENCES

- [Ande 83] H. Andersen. “RATTLE: A velocity version of the SHAKE algorithm for molecular dynamics calculations”. *J. Comput. Physics*, Vol. 52, pp. 24–34, 1983.
- [Bets 05] P. Betsch. “The discrete null space method for the energy consistent integration of constrained mechanical systems. Part I: Holonomic constraints”. *Comput. Methods Appl. Mech. Engrg.*, Vol. 194, No. 50-52, pp. 5159–5190, 2005.
- [Bets 06] P. Betsch and S. Leyendecker. “The discrete null space method for the energy consistent integration of constrained mechanical systems. Part II: Multibody dynamics”. *Int. J. Numer. Meth. Engrg.*, Vol. 67, No. 4, pp. 499–552, 2006.
- [Brid 06] T. Bridges and S. Reich. “Numerical methods for Hamiltonian PDEs”. *J. Phys. A: Math. Gen.*, Vol. 39, pp. 5287–5320, 2006.
- [Cira 05] F. Cirak and M. West. “Decomposition contact response (DCR) for explicit finite element dynamics”. *Int. J. Numer. Meth. Engrg.*, Vol. 64, pp. 1078–1110, 2005.
- [Fete 03] R. Fetecau, J. Marsden, M. Ortiz, and M. West. “Nonsmooth Lagrangian Mechanics and Variational Collision Integrators”. *Siam J. applied dynamical systems*, Vol. 2, No. 3, pp. 381–416, 2003.
- [Gloc 00] C. Glocker. “Scalar force potentials in rigid multibody systems”. In: F. Pfeiffer and C. Glocker, Eds., *Multibody Dynamics with Unilateral Contacts*, Vol. 421 of *CISM Courses and Lectures*, pp. 69–146, Springer, 2000.
- [Hair 06] E. Hairer, G. Wanner, and C. Lubich. *Geometric Numerical Integration: Structure-Preserving Algorithms for Ordinary Differential Equations*. Springer, 2006.
- [Jay 07] L. Jay. “Beyond conventional Runge-Kutta methods in numerical integration of ODEs and DAEs by use of structures and local models”. *J. Comput. Appl. Math.*, Vol. 204, pp. 56–76, 2007.

- [John 12] G. Johnson, M. Ortiz, and S. Leyendecker. “A linear programming-based algorithm for the signed separation of (non-smooth) convex bodies”. *Int. J. Numer. Meth. Engng.*, 2012. accepted for publication.
- [Lein 03] R. Leine, D. van Campen, and C. Glocker. “Nonlinear dynamics and modeling of various wooden toys with impact and friction”. *Journal of vibration and control*, Vol. 9, No. 1-2, pp. 25–78, 2003.
- [Leye 08] S. Leyendecker, J. Marsden, and M. Ortiz. “Variational integrators for constrained dynamical systems”. *ZAMM*, Vol. 88, No. 9, pp. 677–708, 2008.
- [Leye 10] S. Leyendecker, S. Ober-Blöbaum, J. Marsden, and M. Ortiz. “Discrete mechanics and optimal control for constrained systems”. *Optimal Control Applications & Methods*, Vol. 31, No. 6, pp. 505–528, 2010. DOI: 10.1002/oca.912.
- [Leye 12] S. Leyendecker, C. Hartmann, and M. Koch. “Variational collision integrator for polymer chains”. *Journal of Computational Physics*, Vol. 231, 2012.
- [Mars 01] J. Marsden and M. West. “Discrete mechanics and variational integrators”. *Acta Numerica*, Vol. 10, pp. 357–514, 2001.
- [Pand 02] A. Pandolfi, C. Kane, J. Marsden, and M. Ortiz. “Time-discretized variational formulation of nonsmooth frictional contact”. *Int. J. Numer. Meth. Engng.*, Vol. 53, pp. 1801–1829, 2002.
- [Pang 96] J. Pang and J. Trinkle. “Complementarity formulations and existence of solutions of dynamic multi-rigid-body contact problems with Coulomb friction”. *Math. Program.*, Vol. 73, pp. 199–226, 1996.
- [Ryck 77] J. Ryckaert, G. Ciccotti, and H. Berendsen. “Numerical integration of the Cartesian equations of motion of a system with constraints: Molecular dynamics of n-alkanes”. *J. Comput. Phys.*, Vol. 23, pp. 327–341, 1977.
- [Wend 97] J. Wendlandt and J. Marsden. “Mechanical Integrators Derived from a Discrete Variational Principle”. *Physica D*, Vol. 106, pp. 223–246, 1997.
- [Zeid 95] E. Zeidler. *Applied Functional Analysis: Main Principles and Their Applications (Applied Mathematical Sciences, Vol. 109)*. Springer, Berlin, 1995.