Computational Mechanics 33 (2004) 174–185 © Springer-Verlag 2003
DOI 10.1007/s00466-003-0516-2

Energy-conserving integration of constrained Hamiltonian systems – a comparison of approaches

S. Leyendecker, P. Betsch, P. Steinmann

Abstract In this paper known results for continuous Hamiltonian systems subject to holonomic constraints are carried over to a special class of discrete systems, namely to discrete Hamiltonian systems in the sense of Gonzalez [6]. In particular the equivalence of the Lagrange Multiplier Method to the Penalty Method (in the limit for increasing penalty parameters) and to the Augmented Lagrange Method (for infinitely many iterations) is shown theoretically. In doing so many features of the different systems, including dimension, condition number, accuracy, etc. are discussed and compared. Two numerical examples are dealt with to illustrate the results.

Keywords Constrained mechanical systems, Lagrange Multiplier Method, Penalty Method, Augmented Lagrange Method, energy-momentum schemes

1 Introduction

We study Hamiltonian systems subject to holonomic constraints in this work. Three well established methods to treat holonomic constraints, the Lagrange Multiplier Method, the Penalty method and the Augmented Lagrange method are presented and the correlations are reviewed. The known equivalences between the solutions of the different continuous systems are carried over to the corresponding discrete systems.

The quality of a numerical integrator can be gauged by its vicinity to physical reality and by its numerical stability. Both are influenced heavily by the inheritance of the conservation properties or first integrals arising from symmetries of the solution of the underlying continuous Hamiltonian system. The use of a discrete derivative defined by Gonzalez [6] leads to an energy-conserving time stepping scheme. Further specification of the discrete derivative by an equivariance property yields a subclass of the energy-conserving schemes, called energy-momentum schemes. Besides the energy, these schemes conserve a momentum map for a mechanical system with symmetry. The correlations between the solutions of the discrete systems using the methods mentioned to enforce the constraints are investigated for the energy-conserving schemes, since they are more general and notationally simpler. Certainly the proofs hold for the special case of energy-momentum schemes.

Motivated by the way Petzold and Loetstedt [11] calculate the order of the condition number of an iteration scheme, we give the condition numbers of the iteration matrices for the three schemes under consideration. It turns out, that for decreasing time steps, the Lagrange Multiplier scheme becomes more and more ill-conditioned, while the condition number of the other two schemes improves.

An outline of the remainder of this paper is as follows: in Sect. 2 the deduction of Hamiltonian systems on symplectic manifolds is shortly outlined and the three methods to treat holonomic constraints and their correlations are reviewed. Section 3 shows the main aspects of the concept of discrete derivatives and sets up the corresponding three discrete Hamiltonian systems. The equivalence of the discrete Penalty system (in the limit for penalty parameters tending to infinity) and of the discrete Augmented Lagrange system (in the limit for infinitely many Augmented Lagrange iterations) to the discrete Lagrange Multiplier system is proved in Sects. 3.1.2 and 3.1.3, respectively. The theoretical results are illustrated numerically with the motion of a double spherical pendulum and an example of rigid body dynamics in Sect. 4. The paper concludes with a short comparison of some aspects of the investigated methods and conclusions.

2 Hamiltonian systems

In the essential part of this work, mechanical systems are considered in a $2n$-dimensional linear phase space $\mathscr{P}$ with the canonical coordinates $\mathbf{z} = (\mathbf{q}, \mathbf{p}) = (q^1, \ldots, q^n, p_1, \ldots, p_n)$, where $\mathbf{q}$ denotes the position, while $\mathbf{p}$ represents the momentum. A *Hamiltonian* is a $\mathscr{C}^1$-function $H : \mathscr{P} \to \mathbb{R}$. Writing DH for the derivative of H, where the components are collected as follows:

$$\left(\frac{\partial H}{\partial q_1}, \ldots, \frac{\partial H}{\partial q_n}, \frac{\partial H}{\partial p_1}, \ldots, \frac{\partial H}{\partial p_n}\right)^{\mathrm{T}} = \left(D_{\mathbf{q}}H(\mathbf{z}), D_{\mathbf{p}}H(\mathbf{z})\right)^{\mathrm{T}} = D^{\mathrm{T}}H(\mathbf{z}) \ ,$$

Hamilton's equations in classical mechanics take the form

$$\begin{aligned} \dot{\mathbf{q}}(t) &= D_{\mathbf{p}}H \\ \dot{\mathbf{p}}(t) &= -D_{\mathbf{q}}H \ . \end{aligned} \tag{1}$$

Received: 1 July 2003 / Accepted: 26 September 2003
Published online: 25 November 2003

S. Leyendecker, P. Betsch, P. Steinmann (✉)
Chair for Applied Mechanics,
Department of Mechanical Engineering,
University of Kaiserlautern, P.O. 3049,
67653 Kaiserslautern, Germany
e-mail: slauer@rhrk.uni-kl.de

It is well known that the Hamiltonian is conserved along the solution $\mathbf{z}(t)$ of (1). Since the Hamiltonian at any point $\mathbf{z} \in \mathscr{P}$ usually consists of the sum of kinetic energy $T : \mathscr{P} \to \mathbb{R}$ and potential energy $U : \mathscr{P} \to \mathbb{R}$, the solution of (1) is called *energy-conserving*.

For our purposes we take the following (usual) assumptions on the energy-functions:

$$T, U \in \mathscr{C}^1(\mathscr{P}, \mathbb{R})$$

$$T(\mathbf{p}) \geq 0 \quad \forall \mathbf{p} \quad \text{and} \quad T(\mathbf{p}) = 0 \iff \mathbf{p} = \mathbf{0}$$

U is bounded from below, i.e.

$$\exists U^* \in \mathbb{R} \text{ with } \inf_{\mathbf{q} \in \mathscr{Q}} U(\mathbf{q}) \geq U^* > -\infty \ . \tag{2}$$

Remark 2.1 In order not to lose the energy-conservation property of the solution of Hamilton's equations when discretizing (1) to a time-stepping scheme, the concept of discrete derivatives by Gonzalez [6] is introduced in Sect. 3. It relies on a more general view of Hamilton's equations. Thus the following description of their embedding into the setting of symplectic manifolds following the presentation of [10] is required.

Let's consider mechanical systems in the n-dimensional configuration manifold $\mathscr{Q}$. In the Hamiltonian view of mechanics, the cotangent bundle $\mathscr{P} = T^*\mathscr{Q}$ with the local coordinates $z = (q^1, \ldots, q^n, p_1, \ldots, p_n)$ and its intrinsic symplectic structure $\omega = dq^i \wedge dp_i$ represents the $2n$-dimensional phase space.

The nondegenerate symplectic structure on $\mathscr{P}$ is used to assign a vector field X_H to the Hamiltonian $H : \mathscr{P} \to \mathbb{R}$ via

$$\mathrm{i}_{X_H}\omega = \mathbf{d}H \ , \tag{3}$$

the *Hamiltonian vector field*. Here $\mathrm{i}_{X_H}\omega$ denotes the interior product (sometimes called contraction) of the vector field X_H (interpreted as a 1-form) and the 2-form ω. $\mathbf{d}H$ denotes the exterior derivative of the Hamiltonian resulting in a 1-form. For $z \in P$, (3) can be written locally as

$$\omega_z^b(X_H(z(t))) = \mathbf{d}H(z) \ . \tag{4}$$

Hamilton's equations are the evolution equations

$$\dot{z}(t) = X_H(z(t)) \ . \tag{5}$$

2.1 Hamiltonian systems subject to holonomic constraints

Let the motion of the mechanical system be restrained by m holonomic constraints on configuration level:

$$\begin{pmatrix} g_1(\mathbf{q}) \\ \vdots \\ g_m(\mathbf{q}) \end{pmatrix} = \mathbf{g}(\mathbf{q}) = \mathbf{0} \quad \text{with } g_i : \mathscr{Q} \to \mathbb{R} \text{ smooth}, \quad i = 1, \ldots, m \ . \tag{6}$$

Remark 2.2 Any constraint on momentum level (especially 'hidden' constraints, obtained by differentiating the configuration-constraints with respect to time) can be dealt with similarly as described in the following, since the variables $\mathbf{q}$ and $\mathbf{p}$ are dealt with on an equal footing. In [2] Betsch and Steinmann investigate the influence of a projection technique to enforce the 'hidden' constraints on the solution of a Hamiltonian system using Lagrange Multipliers. Besides a prevention of oscillatory behaviour in the multipliers and the fulfillment of the 'hidden' constraints themselves, the method did not influence the solution substantially. For these reasons the investigation of constraints on momentum level is not included in this work.

An important assumption for the unique solvability of the constrained problem is that $\mathbf{0}$ is a regular value of the constraints (see, for example, [12, 9]), i.e.

$$\operatorname{rank}(D\mathbf{g}(\mathbf{q})) = m \quad \forall\, \mathbf{q} \in \mathscr{Q} \text{ with } \mathbf{g}(\mathbf{q}) = \mathbf{0} \ . \tag{7}$$

Then $\mathscr{C} = \mathbf{g}^{-1}(\mathbf{0}) \subset \mathscr{P}$ is a submanifold in the phase space, the $(2n - m)$-dimensional constraint manifold.

Remark 2.3 In the following, three alternative methods to treat holonomic constraints, the Lagrange Multiplier Method, the Penalty Method and the Augmented Lagrange Method will be presented and the correlations are reviewed. The possibility to choose generalized coordinates, which reduce the system to the minimal number of equations necessary to describe the dynamics of the system, is skipped here because it has a crucial drawback: a transformation between the canonical coordinates and the generalized coordinates has to be found. This transformation may be very complicated – it might even not exist e.g. in the case of closed loop systems. As a consequence of the introduction of generalized coordinates, the resulting equations of motion are quite involved in general.

The treatment of the constraints shall be represented generally by the scalar valued $\mathscr{C}^1$-function $P : \mathbf{g}(\mathscr{Q}) \to \mathbb{R}$ that is required to be at least linear in $\mathbf{g}$ and of the form

$$P(\mathbf{g}(\mathbf{q})) \geq 0 \ \forall \mathbf{q} \quad \text{and} \quad P(\mathbf{g}(\mathbf{q})) = 0 \iff \mathbf{g}(\mathbf{q}) = \mathbf{0} \ . \tag{8}$$

For constrained Hamiltonian systems P is added to the energies $T + U$, such that the now relevant augmented Hamiltonian $H : \mathscr{P} \to \mathbb{R}$ reads $H(\mathbf{q}, \mathbf{p}) = T(\mathbf{p}) + U(\mathbf{q}) + P(\mathbf{g}(\mathbf{q}))$.

In order to unitise the domains of the functions composing H, we introduce $\tilde{T}, \tilde{U}, \tilde{P} : \mathscr{P} \to \mathbb{R}$ and $\tilde{\mathbf{g}} : \mathscr{P} \to \mathbb{R}^m$ with the properties

$$\tilde{T}(\mathbf{z}) = T(\mathbf{p}), \quad \tilde{U}(\mathbf{z}) = U(\mathbf{q}),$$

$$\tilde{P}(\mathbf{z}) = P(\tilde{\mathbf{g}}(\mathbf{z})) \quad \text{with} \quad \tilde{\mathbf{g}}(\mathbf{z}) = \mathbf{g}(\mathbf{q}) \tag{9}$$

$$\text{and } D_{\mathbf{q}}\tilde{T}(\mathbf{z}) = D_{\mathbf{p}}\tilde{U}(\mathbf{z}) = D_{\mathbf{p}}\tilde{P}(\mathbf{z}) = \mathbf{0}. \tag{10}$$

2.1.1 Lagrange Multiplier Method

If the Lagrange Multiplier Method is used, the additional function P takes the form

$$P_{Lag}(\mathbf{g}(\mathbf{q})) = \langle \boldsymbol{\lambda}, \mathbf{g}(\mathbf{q}) \rangle \ , \tag{11}$$

where $\boldsymbol{\lambda} \in \mathbb{R}^m$ and $\langle ., . \rangle$ denotes the standard inner product in $\mathbb{R}^m$.

The Hamiltonian system (1) is augmented by the constraint equations and reads

$$\begin{aligned}\dot{\mathbf{q}} &= \frac{\partial T}{\partial \mathbf{p}} \\ \dot{\mathbf{p}} &= -\frac{\partial U}{\partial \mathbf{q}} - D^{\mathrm{T}}\mathbf{g}(\mathbf{q})\boldsymbol{\lambda} \\ \mathbf{0} &= \mathbf{g}(\mathbf{q}) \ .\end{aligned} \tag{12}$$

The Lagrange Multiplier Theorem guarantees the equivalence of the system (12) to (1) subject to $\mathbf{g}(\mathbf{q}) = \mathbf{0}$. Obviously, for a solution of (12), the constraints are fulfilled exactly. The system consists of $2n + m$ equations and is to be solved for the $2n + m$ variables $(\mathbf{q}, \mathbf{p}, \boldsymbol{\lambda})$. For consistent initial data (12) has a unique solution, see [12, 8].

2.1.2 Penalty Method

The Penalty Method approximates the constrained problem by an unconstrained one. A potential is introduced (in terms of the extra function P), which grows large when the system deviates from the constraint manifold. The penalty parameter determines the severity of the violation of the constraints.

Beyond (8), other conditions on P to be a penalty function are:

$$\begin{aligned}&P \text{ must be convex} \ , \\ &P \text{ must be at least quadratic in } \mathbf{g} \ .\end{aligned} \tag{13}$$

The penalty function we will work with is of the form $P_{Pen}(\mathbf{g}(\mathbf{q})) = \mu R(\mathbf{g}(\mathbf{q}))$ with the penalty parameter $\mu \in \mathbb{R}^+$ and the function R such that (8) and (13) are fulfilled. According to (1) the Hamiltonian system reads:

$$\begin{aligned}\dot{\mathbf{q}} &= \frac{\partial T}{\partial \mathbf{p}} \\ \dot{\mathbf{p}} &= -\frac{\partial U}{\partial \mathbf{q}} - \mu D^{\mathrm{T}}\mathbf{g}(\mathbf{q}) DR(\mathbf{g}(\mathbf{q})) \ .\end{aligned} \tag{14}$$

A widely-used example of a penalty function is $P_{Pen}(\mathbf{g}(\mathbf{q})) = \mu \|\mathbf{g}(\mathbf{q})\|^2, \mu \in \mathbb{R}^+$.

Under the assumptions taken, we know from standard ODE-theory that the $2n$-dimensional system (14) is uniquely solvable for given initial data.

In [13] Rubin and Ungar prove an important result for constrained motion described by Euler–Lagrange equations. They show that for a sequence of penalty parameters $(\mu_s)_{s\in\mathbb{N}}$ with $\lim_{s\to\infty} \mu_s = \infty$, the limit point $(\bar{\mathbf{q}}(t), \dot{\bar{\mathbf{q}}}(t)) = \lim_{s\to\infty}(\mathbf{q}^s(t), \dot{\mathbf{q}}^s(t))$ of the sequence of solutions – where $(\mathbf{q}^s(t), \dot{\mathbf{q}}^s(t))$ solve the penalty system corresponding to μ_s – fulfills the constraints. Furthermore they show that there exists a multiplier $\boldsymbol{\lambda}$, such that $(\bar{\mathbf{q}}(t), \dot{\bar{\mathbf{q}}}(t), \boldsymbol{\lambda})$ solve the corresponding Lagrange Multiplier system.

In [5] Bornemann and Schütte carry on with that issue and offer an abstract approach relying on the weak convergence in the sense of distributions. They even give explicitly the sequence converging (weakly in the sense of distributions) to the correct Lagrange Multiplier.

In Sect. 3.1.2, we show a corresponding result for a certain class of discrete Hamiltonian systems.

In practice, the Penalty Method entails a number of drawbacks: (i) Moderate penalty parameters do usually allow unacceptable constraint violation, (ii) for fixed time step size, increasing penalty parameters result in increasingly ill-conditioned systems and (iii) large penalty parameters lead to stiff systems which may cause unstable numerical solutions.

2.1.3 Augmented Lagrange Method

The third approach investigated herein is the Augmented Lagrange Method. In this connection we refer to [4] for a physically insightful interpretation in the context of dynamical systems and [1, 9] for a general introduction. The Augmented Lagrange Method can be regarded as a combination of a Penalty Method and a dual method. Dual methods are based on the idea that it is the Lagrange Multipliers which are the fundamental unknowns in a constrained problem. They have meaningful interpretations such as the costs to keep the system on the constraint manifold.

The extra function to treat the constraints by the Augmented Lagrange Method is a combination of those used in the previous methods:

$$P_{Aug}(\mathbf{g}(\mathbf{q})) = \langle \boldsymbol{\lambda}_k, \mathbf{g}(\mathbf{q}) \rangle + \mu R(\mathbf{g}(\mathbf{q})) \ , \tag{15}$$

with the difference that $\mu \in \mathbb{R}^+$ needs not tend to infinity to fulfill the constraints. Instead of that it may remain of relatively moderate value and the improvement in the constraints is achieved by passing through an extra loop. The multipliers are not determined as an unknown variable, but during an iteration process. Starting with $\boldsymbol{\lambda}_0 = \mathbf{0}$, in each iteration the $2n$-dimensional Hamiltonian system $(16)_{1,2}$ is solved for $(\mathbf{q}_k, \mathbf{p}_k)$ with a fixed value of $\boldsymbol{\lambda}_k$. Then the multiplier is updated according to the Uzawa-like rule $(16)_3$.

$$\begin{aligned}\dot{\mathbf{q}}_k &= \frac{\partial T}{\partial \mathbf{p}_k} \\ \dot{\mathbf{p}}_k &= -\frac{\partial U}{\partial \mathbf{q}_k} - D^{\mathrm{T}}\mathbf{g}(\mathbf{q}_k)\boldsymbol{\lambda}_k - \mu D^{\mathrm{T}}\mathbf{g}(\mathbf{q}_k) DR(\mathbf{g}(\mathbf{q}_k)) \\ \boldsymbol{\lambda}_{k+1} &= \boldsymbol{\lambda}_k + \mu \mathbf{g}(\mathbf{q}_k) \ .\end{aligned} \tag{16}$$

In each iteration the constraints are less violated and the iterations stop as soon as the constraints are fulfilled satisfactorily.

It is known (see e.g. [1]) that in nonlinear constrained optimization problems, a sequence $(x_k)_{k\in\mathbb{N}}$ of minimizing solutions, obtained by Augmented Lagrange iterations, converges to a value $\bar{x}$ which, together with the limit $\bar{\lambda}$ of the corresponding sequence of multipliers $(\lambda_k)_{k\in\mathbb{N}}$, solves the optimization problem, where the constraints have been enforced by Lagrange Multipliers.

In Sect. 3.1.3, we show a corresponding result for a certain class of discrete Hamiltonian systems.

3 Discrete Hamiltonian systems

In order not to lose the conservation properties under discretization, Gonzalez [6] introduced the concept of a discrete derivative leading to implicit second-order conserving time-stepping schemes for general Hamiltonian systems. It is intimately correlated to the more general view of Hamilton's equations described in Remark 2.2. The main ideas of this concept shall be outlined here before it is applied to constrained Hamiltonian systems.

Let $(\mathscr{P}, \omega)$ be a symplectic space (where $\mathscr{P}$ is open in an $2n$-dimensional Euclidian space), serving as phase space for the Hamiltonian $H : \mathscr{P} \to \mathbb{R}$. A discrete approximation of Hamilton's differential equations (5) is given by

$$\mathbf{z}_{n+1} - \mathbf{z}_n = h \mathsf{X}_{\mathscr{H}}(\mathbf{z}_n, \mathbf{z}_{n+1}) \tag{17}$$

where $h > 0$ is the time step and $\mathsf{X}_{\mathscr{H}} : \mathscr{P} \to T\mathscr{P}$ is the discrete Hamiltonian vector field, defined via the discrete analogue of (4)

$$\omega^b_{\mathbf{z}_{n+\frac{1}{2}}}(\mathsf{X}_{\mathscr{H}}(\mathbf{z}_n, \mathbf{z}_{n+1})) = \mathsf{D}H(\mathbf{z}_n, \mathbf{z}_{n+1}) \ , \tag{18}$$

where the discrete derivative (defined below) of the Hamiltonian is used on the right hand side.

Definition 3.1 The discrete derivative $\mathsf{D}f : \mathscr{P} \times \mathscr{P} \to T^*\mathscr{P}$ for a smooth function $f : \mathscr{P} \to \mathbb{R}$ is defined by the following properties:

(i) Directionality: $\mathsf{D}f(\mathbf{x}, \mathbf{y}) \cdot (\mathbf{y} - \mathbf{x}) = f(\mathbf{y}) - f(\mathbf{x})$ for all $\mathbf{x}, \mathbf{y} \in \mathscr{P}$.
(ii) Consistency: $\mathsf{D}f(\mathbf{x}, \mathbf{y}) = Df(\mathbf{w}) + O(\|\mathbf{y} - \mathbf{x}\|)$ for all $\mathbf{x}, \mathbf{y} \in \mathscr{P}$ with $\|\mathbf{y} - \mathbf{x}\|$ sufficiently small.

Here $\mathbf{w} = \frac{1}{2}(\mathbf{x} + \mathbf{y})$ and $\|\cdot\|$ denotes the standard Euclidian norm in $\mathbb{R}^{2n}$.

To interpret (i) correctly, note that for all $\mathbf{x} \in \mathscr{P}$ the tangent space $T_{\mathbf{x}}\mathscr{P}$ as well as the cotangent space $T^*_{\mathbf{x}}\mathscr{P}$ at $\mathbf{x}$ to the linear space $\mathscr{P}$ can be identified with $\mathscr{P}$ itself, as it is usual for linear spaces.

A general example of a discrete derivative is given by

$$\mathsf{D}f(\mathbf{x}, \mathbf{y}) = Df(\mathbf{w}) + \frac{f(\mathbf{y}) - f(\mathbf{x}) - Df(\mathbf{w}) \cdot (\mathbf{y} - \mathbf{x})}{\|\mathbf{y} - \mathbf{x}\|^2} (\mathbf{y} - \mathbf{x}) \ . \tag{19}$$

It is a second-order approximation to the exact derivative at the midpoint.

With this construction, the Hamiltonian H is conserved along a solution sequence $(\mathbf{z}_n)_{n \in \mathbb{N}}$ of (17) in the sense that $H(\mathbf{z}_{n+1}) - H(\mathbf{z}_n) = 0$ for all $n \in \mathbb{N}$.

3.1 Constrained discrete Hamiltonian systems

As descibed at the beginning of Sect. 2.1, the treatment of m holonomic constraints $\mathbf{g} : \mathscr{Q} \to \mathbb{R}^m$ for a Hamiltonian system can be covered by the Hamiltonian $H(\mathbf{q}, \mathbf{p}) = T(\mathbf{p}) + U(\mathbf{q}) + P(\mathbf{g}(\mathbf{q})) = \tilde{T}(\mathbf{z}) + \tilde{U}(\mathbf{z}) + \tilde{P}(\mathbf{z})$, assuming that the conditions (2), (7–10) hold. The function P will be defined according to the method used to treat the constraints.

Denoting the midpoint by $\mathbf{z}_{n+\frac{1}{2}} = \frac{1}{2}(\mathbf{z}_{n+1} + \mathbf{z}_n)$ and using the example (19) a discrete derivative for H is

$$\begin{aligned}
\mathsf{D}H(\mathbf{z}_n, \mathbf{z}_{n+1}) &= D\tilde{T}(\mathbf{z}_{n+\frac{1}{2}}) + D\tilde{U}(\mathbf{z}_{n+\frac{1}{2}}) + D\tilde{P}(\mathbf{z}_{n+\frac{1}{2}}) \\
&+ \left(\frac{\tilde{T}(\mathbf{z}_{n+1}) - \tilde{T}(\mathbf{z}_n) - \langle D\tilde{T}(\mathbf{z}_{n+\frac{1}{2}}), \mathbf{z}_{n+1} - \mathbf{z}_n \rangle}{\|\mathbf{z}_{n+1} - \mathbf{z}_n\|^2} \right. \\
&+ \frac{\tilde{U}(\mathbf{z}_{n+1}) - \tilde{U}(\mathbf{z}_n) - \langle D\tilde{U}(\mathbf{z}_{n+\frac{1}{2}}), \mathbf{z}_{n+1} - \mathbf{z}_n \rangle}{\|\mathbf{z}_{n+1} - \mathbf{z}_n\|^2} \\
&+ \left. \frac{\tilde{P}(\mathbf{z}_{n+1}) - \tilde{P}(\mathbf{z}_n) - \langle D\tilde{P}(\mathbf{z}_{n+\frac{1}{2}}), \mathbf{z}_{n+1} - \mathbf{z}_n \rangle}{\|\mathbf{z}_{n+1} - \mathbf{z}_n\|^2} \right) (\mathbf{z}_{n+1} - \mathbf{z}_n) \\
&= D\tilde{T}(\mathbf{z}_{n+\frac{1}{2}}) + D\tilde{U}(\mathbf{z}_{n+\frac{1}{2}}) + D\tilde{P}(\mathbf{z}_{n+\frac{1}{2}}) \\
&+ \left(S_{\tilde{T}}(\mathbf{z}_n, \mathbf{z}_{n+1}) + S_{\tilde{U}}(\mathbf{z}_n, \mathbf{z}_{n+1}) + S_{\tilde{P}}(\mathbf{z}_n, \mathbf{z}_{n+1}) \right) \\
&(\mathbf{z}_{n+1} - \mathbf{z}_n)
\end{aligned} \tag{20}$$

where $\langle ., . \rangle$ denotes the standard inner product in $\mathbb{R}^{2n}$ and the scalars $S_{\tilde{T}}(\mathbf{z}_n, \mathbf{z}_{n+1}), S_{\tilde{U}}(\mathbf{z}_n, \mathbf{z}_{n+1}), S_{\tilde{P}}(\mathbf{z}_n, \mathbf{z}_{n+1})$ belonging to the scalar valued terms in the functions $\tilde{T}, \tilde{U}, \tilde{P}$ respectively have been introduced in an obvious way.

Using the canonical symplectic structure for the definition of the Hamiltonian vector field (18) and inserting this in the discrete equations of motion (17) we arrive at the discrete constrained Hamiltonian system

$$\begin{aligned}
\frac{\mathbf{q}_{n+1} - \mathbf{q}_n}{h} &= D_{\mathbf{p}} \tilde{T}(\mathbf{z}_{n+\frac{1}{2}}) \\
&+ \left(S_{\tilde{T}}(\mathbf{z}_n, \mathbf{z}_{n+1}) + S_{\tilde{U}}(\mathbf{z}_n, \mathbf{z}_{n+1}) \right. \\
&\left. + S_{\tilde{P}}(\mathbf{z}_n, \mathbf{z}_{n+1}) \right) (\mathbf{p}_{n+1} - \mathbf{p}_n) \\
\frac{\mathbf{p}_{n+1} - \mathbf{p}_n}{h} &= -D_{\mathbf{q}} \tilde{U}(\mathbf{z}_{n+\frac{1}{2}}) - D_{\mathbf{q}} \tilde{P}(\mathbf{z}_{n+\frac{1}{2}}) \\
&- \left(S_{\tilde{T}}(\mathbf{z}_n, \mathbf{z}_{n+1}) + S_{\tilde{U}}(\mathbf{z}_n, \mathbf{z}_{n+1}) \right. \\
&\left. + S_{\tilde{P}}(\mathbf{z}_n, \mathbf{z}_{n+1}) \right) (\mathbf{q}_{n+1} - \mathbf{q}_n) \ .
\end{aligned} \tag{21}$$

3.1.1 Lagrange Multiplier Method

As described in Sect. 2.1.1, the use of Lagrange Multipliers to enforce m constraints enlarges the number of equations as well as the number of unknowns by m. The function P_{Lag} is the inner product of $\boldsymbol{\lambda} \in \mathbb{R}^m$ and $\mathbf{g}(\mathbf{q})$:

$$\tilde{P}_{\text{Lag}}(\mathbf{z}) = P_{Lag}(\tilde{\mathbf{g}}(\mathbf{z})) = \langle \tilde{\mathbf{g}}(\mathbf{z}), \boldsymbol{\lambda} \rangle \ . \tag{22}$$

Consequently, the discrete Hamiltonian system takes the form

$$\begin{aligned}
\frac{\mathbf{q}_{n+1} - \mathbf{q}_n}{h} &= D_{\mathbf{p}} \tilde{T}(\mathbf{z}_{n+\frac{1}{2}}) + \Bigg(S_{\tilde{T}}(\mathbf{z}_n, \mathbf{z}_{n+1}) \\
&+ S_{\tilde{U}}(\mathbf{z}_n, \mathbf{z}_{n+1}) \\
&- \frac{\langle D^T \tilde{\mathbf{g}}(\mathbf{z}_{n+\frac{1}{2}}) \boldsymbol{\lambda}_{n+1}, \mathbf{z}_{n+1} - \mathbf{z}_n \rangle}{\|\mathbf{z}_{n+1} - \mathbf{z}_n\|^2} \Bigg) (\mathbf{p}_{n+1} - \mathbf{p}_n) \\
\frac{\mathbf{p}_{n+1} - \mathbf{p}_n}{h} &= -D_{\mathbf{q}} \tilde{U}(\mathbf{z}_{n+\frac{1}{2}}) - D^T \tilde{\mathbf{g}}(\mathbf{z}_{n+\frac{1}{2}}) \lambda_{n+1} \\
&- \Bigg(S_{\tilde{T}}(\mathbf{z}_n, \mathbf{z}_{n+1}) + S_{\tilde{U}}(\mathbf{z}_n, \mathbf{z}_{n+1}) \\
&- \frac{\langle D^T \tilde{\mathbf{g}}(\mathbf{z}_{n+\frac{1}{2}}) \boldsymbol{\lambda}_{n+1}, \mathbf{z}_{n+1} - \mathbf{z}_n \rangle}{\|\mathbf{z}_{n+1} - \mathbf{z}_n\|^2} \Bigg) (\mathbf{q}_{n+1} - \mathbf{q}_n) \\
\tilde{\mathbf{g}}(\mathbf{z}_{n+1}) &= \mathbf{0} \ .
\end{aligned} \tag{23}$$

Analogous to the continuous case, under the assumptions taken, for given $\mathbf{z}_n$ with $\tilde{\mathbf{g}}(\mathbf{z}_n) = \mathbf{0}$, (23) has a unique solution $(\mathbf{z}_{n+1}, \boldsymbol{\lambda}_{n+1})$. Obviously, $\mathbf{z}_{n+1}$ satisfies the position constraints such that $\mathbf{z}_{n+1} \in \mathscr{C}$.

Solving the nonlinear system of equations (23) by Newton–Raphson iteration involves a tangent matrix of rather poor structure (neither symmetric nor banded). Furthermore for decreasing time steps h, the system becomes more and more ill-conditioned. Following the ideas in [11], it can be easily calculated that the condition number of the Jacobian of the residual (23) is of the order $O(\frac{1}{h^2})$. Obviously, (23) yields a solution that fulfills the constraints exactly (up to a numerical tolerance) and provides information about the constraint forces in the mechanical system.

3.1.2 Penalty Method

Setting up (21) for the Hamiltonian $H(\mathbf{q},\mathbf{p}) = \tilde{T}(\mathbf{z}) + \tilde{U}(\mathbf{z}) + \mu\tilde{R}(\mathbf{z})$ with $\mu \in \mathbb{R}^+$ and $\tilde{R}(\mathbf{z}) = R(\tilde{\mathbf{g}}(\mathbf{z}))$, leads to the discrete Hamiltonian system

$$\begin{aligned}
\frac{\mathbf{q}_{n+1}-\mathbf{q}_n}{h} &= D_{\mathbf{p}}\tilde{T}(\mathbf{z}_{n+\frac{1}{2}}) + (\mathbf{p}_{n+1}-\mathbf{p}_n)\Big(S_{\tilde{T}}(\mathbf{z}_n,\mathbf{z}_{n+1}) + S_{\tilde{U}}(\mathbf{z}_n,\mathbf{z}_{n+1}) \\
&\quad + \frac{\mu R(\tilde{\mathbf{g}}(\mathbf{z}_{n+1})) - \mu R(\tilde{\mathbf{g}}(\mathbf{z}_n)) - \langle \mu D^T\tilde{\mathbf{g}}(\mathbf{z}_{n+\frac{1}{2}}) D_{\tilde{\mathbf{g}}} R(\tilde{\mathbf{g}}(\mathbf{z}_{n+\frac{1}{2}})), \mathbf{z}_{n+1}-\mathbf{z}_n\rangle}{\|\mathbf{z}_{n+1}-\mathbf{z}_n\|^2}\Big) \\
\frac{\mathbf{p}_{n+1}-\mathbf{p}_n}{h} &= -D_{\mathbf{q}}\tilde{U}(\mathbf{z}_{n+\frac{1}{2}}) - \mu D^T\tilde{\mathbf{g}}(\mathbf{z}_{n+\frac{1}{2}}) D_{\tilde{\mathbf{g}}} R(\tilde{\mathbf{g}}(\mathbf{z}_{n+\frac{1}{2}})) - (\mathbf{q}_{n+1}-\mathbf{q}_n)\Big(S_{\tilde{T}}(\mathbf{z}_n,\mathbf{z}_{n+1}) + S_{\tilde{U}}(\mathbf{z}_n,\mathbf{z}_{n+1}) \\
&\quad + \frac{\mu R(\tilde{\mathbf{g}}(\mathbf{z}_{n+1})) - \mu R(\tilde{\mathbf{g}}(\mathbf{z}_n)) - \langle \mu D^T\tilde{\mathbf{g}}(\mathbf{z}_{n+\frac{1}{2}}) D_{\tilde{\mathbf{g}}} R(\tilde{\mathbf{g}}(\mathbf{z}_{n+\frac{1}{2}})), \mathbf{z}_{n+1}-\mathbf{z}_n\rangle}{\|\mathbf{z}_{n+1}-\mathbf{z}_n\|^2}\Big)
\end{aligned} \tag{24}$$

For all $\mu \in \mathbb{R}^+$ (24) has a unique solution $\mathbf{z}_{n+1}$ for given $\mathbf{z}_n$ with $\tilde{\mathbf{g}}(\mathbf{z}_n) = \mathbf{0}$.

The following proposition may be viewed as the discrete counterpart of the result by Rubin and Ungar [13] associated with the energy-conserving scheme under investigation.

Proposition 3.2 Let $\mathbf{z}_n \in \mathscr{C}$ be consistent coordinates at time t_n, $n \in \mathbb{N}$ arbitrary. Let $(\mu_s)_{s\in\mathbb{N}} \subset \mathbb{R}^+$ be an arbitrary sequence with $\mu_s \to \infty$ as $s \to \infty$ and denote the solution of the system (24) corresponding to μ_s by $\mathbf{z}^s_{n+1}$. Then the sequence of solutions $(\mathbf{z}^s_{n+1})_{s\in\mathbb{N}}$ of (24) converges to the solution $\mathbf{z}_{n+1}$ of the Lagrange Multiplier Method (23) as $s \to \infty$.

Proof:
Let $n \in \mathbb{N}$ and $s \in \mathbb{N}$ be arbitrary. Since we integrate the system of equations by an energy-conserving method and we start with a consistent value, i.e. $\tilde{\mathbf{g}}(\mathbf{z}_n) = \mathbf{0}$, we know that the correct energy of the mechanical system $\tilde{T}(\mathbf{z}_n) + \tilde{U}(\mathbf{z}_n) = H(\mathbf{z}_n) = H_0$ is conserved along the solution of (24):

$$\tilde{T}(\mathbf{z}^s_{n+1}) + \tilde{U}(\mathbf{z}^s_{n+1}) + \mu_s\tilde{R}(\mathbf{z}^s_{n+1}) = H_0 \ . \tag{25}$$

Since $\tilde{T}$ and $\tilde{R}$ are non-negative and $\tilde{U}$ is bounded from below, it follows that there exists $J^s_{n+1} \in \mathbb{R}^+$ with

$$\begin{aligned}
&\mu_s\tilde{R}(\mathbf{z}^s_{n+1}) \le J^s_{n+1} \quad s \in \mathbb{N} \text{ arbitrary} \\
&\quad \Rightarrow \exists J_{n+1} \in \mathbb{R}^+ \quad \text{with} \ \lim_{s\to\infty} \mu_s\tilde{R}(\mathbf{z}^s_{n+1}) \le J_{n+1} \\
&\quad \Rightarrow_{\mu_s\to\infty} \lim_{s\to\infty} \tilde{R}(\mathbf{z}^s_{n+1}) = 0 \\
&\quad \Rightarrow_{(8)} \lim_{s\to\infty} \tilde{\mathbf{g}}(\mathbf{z}^s_{n+1}) = \mathbf{0}
\end{aligned}$$

Let $\bar{\mathbf{z}}_{n+1}$ be the limit point of $(\mathbf{z}^s_{n+1})_{s\in\mathbb{N}}$, then it follows by the continuity of the constraint functions that $\tilde{\mathbf{g}}(\bar{\mathbf{z}}_{n+1}) = \mathbf{g}(\bar{\mathbf{q}}_{n+1}) = \mathbf{0}$, which means that the constraints are fulfilled by $\bar{\mathbf{z}}_{n+1} = \lim_{s\to\infty} \mathbf{z}^s_{n+1}$.

Together with (8) this implies $\tilde{P}_{Pen}(\bar{\mathbf{z}}_{n+1}) = \lim_{s\to\infty} \mu_s R(\tilde{\mathbf{g}}(\mathbf{z}^s_{n+1})) = 0$. In particular the growth of $\tilde{P}_{Pen}$ is bounded, i.e. there is a constant $K_{n+1} \in \mathbb{R}^+$ with $\|D\tilde{P}_{Pen}(\bar{\mathbf{z}}_{n+1})\| \le K_{n+1}$.

By the boundedness of H (that implies the boundedness of $\tilde{T}$, $\tilde{U}$ and $\tilde{P}$ via (25)), we can conclude that for all $n \in \mathbb{N}$ the solutions $(\mathbf{z}^s_{n+1})_{s\in\mathbb{N}}$ as well as the limit point $\bar{\mathbf{z}}_{n+1}$ lie in a bounded region in $\mathscr{P}$. Consequently, $D\tilde{P}_{Pen}(\bar{\mathbf{z}}_{n+\frac{1}{2}})$ is bounded for $\bar{\mathbf{z}}_{n+\frac{1}{2}} = \frac{1}{2}(\bar{\mathbf{z}}_{n+1} + \mathbf{z}_n)$.

$$\begin{aligned}
D\tilde{P}_{Pen}(\bar{\mathbf{z}}_{n+\frac{1}{2}}) &= \lim_{s\to\infty} \mu_s D^T\tilde{\mathbf{g}}(\mathbf{z}^s_{n+\frac{1}{2}}) D_{\tilde{\mathbf{g}}} R(\tilde{\mathbf{g}}(\mathbf{z}^s_{n+\frac{1}{2}})) \\
&= D^T\tilde{\mathbf{g}}(\bar{\mathbf{z}}_{n+\frac{1}{2}}) \lim_{s\to\infty} \mu_s D_{\tilde{\mathbf{g}}} R(\tilde{\mathbf{g}}(\mathbf{z}^s_{n+\frac{1}{2}})) \ .
\end{aligned} \tag{26}$$

Since $\mathbf{0}$ is a regular value of the constraints, $D^T\tilde{\mathbf{g}}$ has full rank in $\mathbf{z}_n$ and $\bar{\mathbf{z}}_{n+1}$. So, if the time step is small enough, $D^T\tilde{\mathbf{g}}(\bar{\mathbf{z}}_{n+\frac{1}{2}}) : T_{\tilde{\mathbf{g}}(\bar{\mathbf{z}}_{n+\frac{1}{2}})}\mathbb{R}^m \to T_{\bar{\mathbf{z}}_{n+\frac{1}{2}}}\mathscr{P}$ has also full rank, particularly it is injective. Hence there exist $(\boldsymbol{\theta}^s_{n+1})_{s\in\mathbb{N}} \in \mathbb{R}^m$ and $\bar{\boldsymbol{\theta}}_{n+1} \in \mathbb{R}^m$ with

$$\begin{aligned}
&\bar{\boldsymbol{\theta}}_{n+1} = \lim_{s\to\infty} \boldsymbol{\theta}^s_{n+1} = \lim_{s\to\infty} \mu_s D_{\tilde{\mathbf{g}}} R(\tilde{\mathbf{g}}(\mathbf{z}^s_{n+\frac{1}{2}})) \\
&\quad \text{and} \ \|\bar{\boldsymbol{\theta}}_{n+1}\| < \infty \ .
\end{aligned} \tag{27}$$

With this preliminaries and abbreviations, for $\mathbf{z}_n \in \mathscr{P}$ with $\tilde{\mathbf{g}}(\mathbf{z}_n) = \mathbf{0}$ given, $\bar{\mathbf{z}}_{n+1}$ fulfills the equations

$$\frac{\bar{\mathbf{q}}_{n+1}-\mathbf{q}_n}{h} = D_{\mathbf{p}}\tilde{T}(\bar{\mathbf{z}}_{n+\frac{1}{2}}) + \Big(S_{\tilde{T}}(\mathbf{z}_n,\bar{\mathbf{z}}_{n+1}) + S_{\tilde{U}}(\mathbf{z}_n,\bar{\mathbf{z}}_{n+1}) - \frac{\langle D^T\tilde{\mathbf{g}}(\bar{\mathbf{z}}_{n+\frac{1}{2}})\bar{\boldsymbol{\theta}}_{n+1}, \mathbf{z}_{n+1}-\mathbf{z}_n\rangle}{\|\bar{\mathbf{z}}_{n+1}-\mathbf{z}_n\|^2}\Big)(\bar{\mathbf{p}}_{n+1}-\mathbf{p}_n)$$

$$\frac{\bar{\mathbf{p}}_{n+1}-\mathbf{p}_n}{h} = -D_{\mathbf{q}}\tilde{U}(\bar{\mathbf{z}}_{n+\frac{1}{2}}) - D^T\tilde{\mathbf{g}}(\bar{\mathbf{z}}_{n+\frac{1}{2}})\bar{\boldsymbol{\theta}}_{n+1} - \Big(S_{\tilde{T}}(\mathbf{z}_n,\bar{\mathbf{z}}_{n+1}) + S_{\tilde{U}}(\mathbf{z}_n,\bar{\mathbf{z}}_{n+1}) - \frac{\langle D^T\tilde{\mathbf{g}}(\bar{\mathbf{z}}_{n+\frac{1}{2}})\bar{\boldsymbol{\theta}}_{n+1}, \bar{\mathbf{z}}_{n+1}-\mathbf{z}_n\rangle}{\|\bar{\mathbf{z}}_{n+1}-\mathbf{z}_n\|^2}\Big)(\bar{\mathbf{q}}_{n+1}-\mathbf{q}_n)$$

$$\tilde{\mathbf{g}}(\bar{\mathbf{z}}_{n+1}) = \mathbf{0} \ . \qquad (28)$$

This system equals exactly (23). Because of the uniqueness of the solution $(\mathbf{z}_{n+1},\boldsymbol{\lambda}_{n+1})$ of (23) we have $(\bar{\mathbf{z}}_{n+1},\bar{\boldsymbol{\theta}}_{n+1}) = (\mathbf{z}_{n+1},\boldsymbol{\lambda}_{n+1})$, or loosely speaking, at each time step the solution sequence of the Penalty Method $(\mathbf{z}^s_{n+1})_{s\in\mathbb{N}}$ converges to the solution $\mathbf{z}_{n+1}$ of the Lagrange Multiplier Method. Furthermore, the product of the penalty parameter and the derivative of R at the point $\mathbf{z}^s_{n+\frac{1}{2}}$ converges to the corresponding Lagrange Multiplier $\boldsymbol{\lambda}_{n+1}$.

Remark 3.3 This proposition holds for the class of energy-conserving integrators using the discrete derivative (3.1). The main argument in the proof is the conservation of energy along the solution of the discrete system. The statement could have been proved directly for the subclass of energy-momentum schemes, but here the more general (and notationally simpler) case has been preferred.

Remark 3.4 In [13] and [5], the special penalty function $P(\mathbf{g}(\mathbf{q}(t))) = \mu\|\mathbf{g}(\mathbf{q}(t))\|^2$ is considered in the continuous (∞-dimensional) case and only the weak convergence of $\mu_s\mathbf{g}(\mathbf{q}^s(t))$ to the correct Lagrange Multiplier is discussed. Since the discrete system is finite-dimensional, weak and strong convergence are equivalent for
$\bar{\boldsymbol{\theta}}_{n+1} = \lim_{s\to\infty}\mu_s D_{\tilde{\mathbf{g}}}R(\tilde{\mathbf{g}}(\mathbf{z}^s_{n+\frac{1}{2}}))$.

The condition of the Jacobian of (24) is of the order $O(\mu^2h^2)$. In contrast to (23), here the condition improves when the time step size decreases, provided that μ is fixed.

3.1.3 Augmented Lagrange Method

As explained in Sect. 2.1.3, the function P, that covers the enforcement of the constraints in the Augmented Lagrange Method is a combination of those used in the two previous sections.

$$P_{Aug}(\mathbf{g}(\mathbf{q})) = P_{Lag}(\mathbf{g}(\mathbf{q})) + P_{Pen}(\mathbf{g}(\mathbf{q})) = \langle\boldsymbol{\lambda}_k,\mathbf{g}(\mathbf{q})\rangle + \mu R(\mathbf{g}(\mathbf{q}))$$

where $\mu\in\mathbb{R}^+$ is fixed and $\boldsymbol{\lambda}_k$ is updated during an extra iteration according to the formula

$$\boldsymbol{\lambda}^0 = \mathbf{0}$$
$$\boldsymbol{\lambda}^{k+1} = \boldsymbol{\lambda}^k + \mu D_{\mathbf{g}}R(\mathbf{g}(\mathbf{q}^k)) \ .$$

For $\mathbf{z}_n$ with $\tilde{\mathbf{g}}(\mathbf{z}_n) = \mathbf{0}$ and $\boldsymbol{\lambda}^k_{n+1}$ given, we consider the discrete Hamiltonian system

$$\frac{\mathbf{q}^k_{n+1}-\mathbf{q}_n}{h} = D_{\mathbf{p}}\tilde{T}(\mathbf{z}^k_{n+\frac{1}{2}}) + \Big(S_{\tilde{T}}(\mathbf{z}_n,\mathbf{z}^k_{n+1}) + S_{\tilde{U}}(\mathbf{z}_n,\mathbf{z}^k_{n+1}) + S_{\tilde{P}_{Aug}}(\mathbf{z}_n,\mathbf{z}^k_{n+1})\Big)(\mathbf{p}^k_{n+1}-\mathbf{p}_n)$$

$$\frac{\mathbf{p}^k_{n+1}-\mathbf{p}_n}{h} = -D_{\mathbf{q}}\tilde{U}(\mathbf{z}^k_{n+\frac{1}{2}}) - D^T\tilde{\mathbf{g}}(\mathbf{z}^k_{n+\frac{1}{2}})\big(\boldsymbol{\lambda}^k_{n+1} + \mu D_{\tilde{\mathbf{g}}}R(\tilde{\mathbf{g}}(\mathbf{z}^k_{n+1}))\big) - \Big(S_{\tilde{T}}(\mathbf{z}_n,\mathbf{z}^k_{n+1}) + S_{\tilde{U}}(\mathbf{z}_n,\mathbf{z}^k_{n+1}) + S_{\tilde{P}_{Aug}}(\mathbf{z}_n,\mathbf{z}^k_{n+1})\Big)(\mathbf{q}^k_{n+1}-\mathbf{q}_n) \qquad (29)$$

with

$$S_{\tilde{P}_{Aug}}(\mathbf{z}_n,\mathbf{z}^k_{n+1}) = \|\mathbf{z}^k_{n+1}-\mathbf{z}_n\|^{-2}\Big(\langle\boldsymbol{\lambda}^k_{n+1},\tilde{\mathbf{g}}(\mathbf{z}^k_{n+1})\rangle + \mu\tilde{R}(\mathbf{z}^k_{n+1}) - \langle D^T\tilde{\mathbf{g}}(\mathbf{z}^k_{n+\frac{1}{2}})\big(\boldsymbol{\lambda}^k_{n+1} + \mu D_{\tilde{\mathbf{g}}}R(\tilde{\mathbf{g}}(\mathbf{z}^k_{n+1}))\big), \mathbf{z}^k_{n+1}-\mathbf{z}_n\rangle\Big) \qquad (30)$$

where the example (19) of a discrete derivative has been used for $\tilde{T},\tilde{U},\tilde{P}_{Lag}$, while it has been slightly modified for $\tilde{P}_{Pen}$:

$$\hat{\mathrm{D}}\tilde{P}_{Pen}(\mathbf{z}_n,\mathbf{z}^k_{n+1}) = D^T\tilde{\mathbf{g}}(\mathbf{z}^k_{n+\frac{1}{2}})D_{\tilde{\mathbf{g}}}P_{Pen}(\tilde{\mathbf{g}}(\mathbf{z}^k_{n+1})) + \frac{\tilde{P}_{Pen}(\mathbf{z}^k_{n+1}) - \tilde{P}_{Pen}(\mathbf{z}_n) - \langle D^T\tilde{\mathbf{g}}(\mathbf{z}^k_{n+\frac{1}{2}})D_{\tilde{\mathbf{g}}}P_{Pen}(\tilde{\mathbf{g}}(\mathbf{z}^k_{n+1})), \mathbf{z}^k_{n+1}-\mathbf{z}_n\rangle}{\|\mathbf{z}^k_{n+1}-\mathbf{z}_n\|^2}(\mathbf{z}^k_{n+1}-\mathbf{z}_n) \qquad (31)$$

This is done for technical reasons. Using $\hat{\mathrm{D}}$ allows to show the fulfillment of the constraints at $\bar{\mathbf{z}}_{n+1}$, whereas using D would lead to the fulfillment of the constraints at the midpoint $\bar{\mathbf{z}}_{n+\frac{1}{2}}$. $\hat{\mathrm{D}}$ is a discrete derivative as well, so (29) is a consistent, energy-conserving system.

Proposition 3.5 Let $\mathbf{z}_n\in\mathscr{C}$ be consistent coordinates at time t_n, $n\in\mathbb{N}$ arbitrary. Let $\mu\in\mathbb{R}^+$ be arbitrary and denote the solution of the system (29) corresponding to $\boldsymbol{\lambda}^k_{n+1}$ by $\mathbf{z}^k_{n+1}$. Then the sequence of solutions $(\mathbf{z}^k_{n+1})_{k\in\mathbb{N}}$ of (29) converges to the solution $\mathbf{z}_{n+1}$ of the Lagrange Multiplier Method (23) as $k\to\infty$. Furthermore the sequence of Lagrange Multipliers $(\boldsymbol{\lambda}^k_{n+1})_{k\in\mathbb{N}}$ converges to the correct Lagrange Multiplier $\boldsymbol{\lambda}_{n+1}$ that solves (23) together with $\mathbf{z}_{n+1}$.

Proof:
Let $n \in \mathbb{N}$ and $k \in \mathbb{N}$ be arbitrary. Since the energy is conserved along the solutions $\mathbf{z}_{n+1}^k$ of (29) and $\tilde{\mathbf{g}}(\mathbf{z}_n) = \mathbf{0}$, we have

$$H(\mathbf{z}_{n+1}^k) = \tilde{T}(\mathbf{z}_{n+1}^k) + \tilde{U}(\mathbf{z}_{n+1}^k) + \langle \boldsymbol{\lambda}_{n+1}^k, \tilde{\mathbf{g}}(\mathbf{z}_{n+1}^k)\rangle + \mu \tilde{R}(\mathbf{z}_{n+1}^k) = H_0 \ . \qquad (32)$$

With the assumptions (2), (8), (13) it follows that there is $K_{n+1}^k \in \mathbb{R}^+$ such that $\langle \boldsymbol{\lambda}_{n+1}^k, \tilde{\mathbf{g}}(\mathbf{z}_{n+1}^k)\rangle \leq K_{n+1}^k$. Let $\bar{\mathbf{z}}_{n+1} = \lim_{k\to\infty} \mathbf{z}_{n+1}^k$ be the limit point of the solutions of (29) with the corresponding $\boldsymbol{\lambda}_{n+1}^k$ and let

$$\bar{\boldsymbol{\lambda}}_{n+1} = \lim_{k\to\infty} \boldsymbol{\lambda}_{n+1}^k = \boldsymbol{\lambda}_{n+1}^0 + \sum_{k=1}^{\infty} \mu D_{\tilde{\mathbf{g}}} R(\tilde{\mathbf{g}}(\mathbf{z}_{n+1}^k)) \ . \qquad (33)$$

Since k is arbitrary, there is $K_{n+1} \in \mathbb{R}^+$ such that $\langle \bar{\boldsymbol{\lambda}}_{n+1}, \tilde{\mathbf{g}}(\bar{\mathbf{z}}_{n+1})\rangle \leq K_{n+1}$, which comprehends two cases:

- $\bar{\boldsymbol{\lambda}}_{n+1} = \infty \Rightarrow \tilde{\mathbf{g}}(\bar{\mathbf{z}}_{n+1}) = \lim_{k\to\infty} \tilde{\mathbf{g}}(\bar{\mathbf{z}}_{n+1}^k) = \mathbf{0}$
- $\bar{\boldsymbol{\lambda}}_{n+1} < \infty \Rightarrow \lim_{k\to\infty} D_{\tilde{\mathbf{g}}} R(\tilde{\mathbf{g}}(\mathbf{z}_{n+1}^k)) = \mathbf{0}$
  $\underset{R\text{ convex}}{\Rightarrow} \tilde{\mathbf{g}}(\bar{\mathbf{z}}_{n+1}) = \mathbf{0}$

In any case this states that the constraints are fulfilled in the limit point $\bar{\mathbf{z}}_{n+1}$. The same arguments as in the proof of Proposition 3.2 imply (i) that the solutions $\left(\mathbf{z}_{n+1}^k\right)_{k\in\mathbb{N}}$ as well as the limit point $\bar{\mathbf{z}}_{n+1}$ lie in a bounded region in $\mathscr{P}$ and (ii) the injectivity of $D^T\tilde{\mathbf{g}}(\bar{\mathbf{z}}_{n+\frac{1}{2}})$. A look at the system (29) then excludes the first case.

Summing up we find that $(\bar{\mathbf{z}}_{n+1}, \bar{\boldsymbol{\lambda}}_{n+1})$ fulfill the equations

$$\frac{\bar{\mathbf{q}}_{n+1} - \mathbf{q}_n}{h} = D_{\mathbf{p}}\tilde{T}(\bar{\mathbf{z}}_{n+\frac{1}{2}}) + \left(S_{\tilde{T}}(\mathbf{z}_n, \bar{\mathbf{z}}_{n+1}) + S_{\tilde{U}}(\mathbf{z}_n, \bar{\mathbf{z}}_{n+1}) - \frac{\langle D^T\tilde{\mathbf{g}}(\bar{\mathbf{z}}_{n+\frac{1}{2}})\bar{\boldsymbol{\lambda}}_{n+1}, \mathbf{z}_{n+1} - \mathbf{z}_n\rangle}{\left\|\bar{\mathbf{z}}_{n+1} - \mathbf{z}_n\right\|^2} \right)(\bar{\mathbf{p}}_{n+1} - \mathbf{p}_n)$$

$$\frac{\bar{\mathbf{p}}_{n+1} - \mathbf{p}_n}{h} = -D_{\mathbf{q}}\tilde{U}(\bar{\mathbf{z}}_{n+\frac{1}{2}}) - D^T\tilde{\mathbf{g}}(\bar{\mathbf{z}}_{n+\frac{1}{2}})\bar{\boldsymbol{\lambda}}_{n+1} - \left(S_{\tilde{T}}(\mathbf{z}_n, \bar{\mathbf{z}}_{n+1}) + S_{\tilde{U}}(\mathbf{z}_n, \bar{\mathbf{z}}_{n+1}) - \frac{\langle D^T\tilde{\mathbf{g}}(\bar{\mathbf{z}}_{n+\frac{1}{2}})\bar{\boldsymbol{\lambda}}_{n+1}, \bar{\mathbf{z}}_{n+1} - \mathbf{z}_n\rangle}{\left\|\bar{\mathbf{z}}_{n+1} - \mathbf{z}_n\right\|^2} \right)(\bar{\mathbf{q}}_{n+1} - \mathbf{q}_n)$$

$$\tilde{\mathbf{g}}(\bar{\mathbf{z}}_{n+1}) = \mathbf{0} \ . \qquad (34)$$

This system equals exactly (23). Because of the uniqueness of the solution $(\mathbf{z}_{n+1}, \boldsymbol{\lambda}_{n+1})$ of (23) we have $(\bar{\mathbf{z}}_{n+1}, \bar{\boldsymbol{\lambda}}_{n+1}) = (\mathbf{z}_{n+1}, \boldsymbol{\lambda}_{n+1})$, i.e. the sequence $\left(\mathbf{z}_{n+1}^k, \boldsymbol{\lambda}_{n+1}^k\right)_{k\in\mathbb{N}}$ of solutions of the Augmented Lagrange Method together with the corresponding multipliers converges to the solution of the Lagrange Multiplier Method at each time step.

Remark 3.6 (Equal to Remark 3.3). This proposition holds for the class of energy-conserving integrators using the discrete derivative (3.1). The main argument in the proof is the conservation of energy along the solution of the discrete system. The statement could have been proved directly for the subclass of energy-momentum schemes, but here the more general (and notationally simpler) case has been preferred.

The condition of the iteration matrix in the Newton–Raphson iteration for (29) is also of the order $O(\mu^2h^2)$, but the numerical examples show, that for the same penalty parameter and time step it is slightly worse than in the Penalty Method, since the influence of the penalty parameter is a little stronger. The advantage of the Augmented Lagrange Method is that the fulfillment of the constraints is improved during the extra iteration while μ is kept constant, so if the time step is small enough to balance the penalty parameter, the Newton-Raphson iteration runs with a very good-conditioned tangent matrix.

4 Numerical examples

Although the use of a special equivariant discrete derivative (see [6]) complicates the setup of the general discrete constrained Hamiltonian system (21), it leads to less complicated systems for the relatively simple examples treated in the following. Besides that the solution is more realistic due to the additional conservation of momentum maps, energy-momentum schemes exhibit improved numerical behaviour. All the numerical results presented here are obtained by employing energy-momentum schemes. For the double spherical pendulum, the relevant equations are given explicitely.

4.1 Double spherical pendulum

In the first numerical example we consider the motion of the double spherical pendulum in Fig. 1. It is suspended at the origin of the 3-dimensional cartesian coordinate system. Massless rigid rods of lengths l_1 and l_2 connect the masses m_1 and m_2 to each other and to the origin, respectively. The gravitational acceleration with absolute value g points in the negative $\mathbf{e}_2$-direction. The kinetic energy T and the potential energy U are given by the following expressions:

$$T(\mathbf{p}) = \frac{1}{2}\mathbf{p}^T\mathbf{M}^{-1}\mathbf{p}, \qquad U(\mathbf{q}) = g\begin{pmatrix}\mathbf{e}_2\\ \mathbf{e}_2\end{pmatrix}^{\mathrm{T}}\mathbf{Mq} \ , \qquad (35)$$

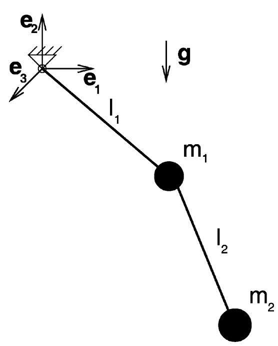

Fig. 1. Double spherical pendulum

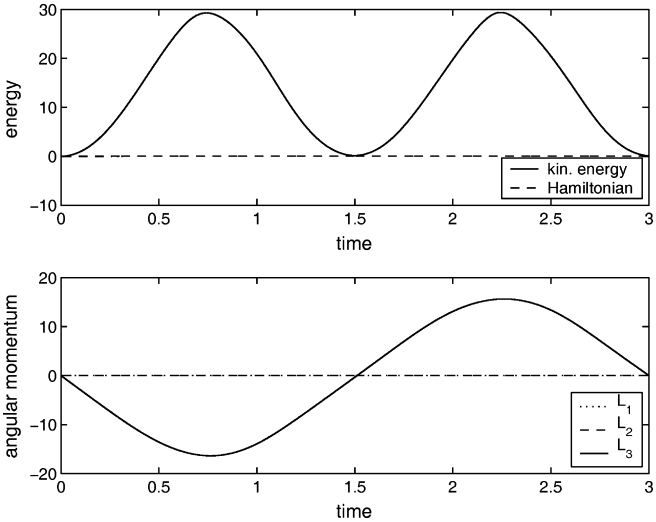

Fig. 2. Algorithmic evolution of kinetic energy, Hamiltonian and angular momentum, $h = 10^{-3}$

with $\mathbf{q} = \begin{pmatrix} \mathbf{q}^1 \\ \mathbf{q}^2 \end{pmatrix}, \mathbf{p} = \begin{pmatrix} \mathbf{p}^1 \\ \mathbf{p}^2 \end{pmatrix} \in \mathbb{R}^6$ and the 6×6 diagonal mass matrix

$$\mathbf{M} = \begin{pmatrix} m_1 \mathbf{I}_3 & \mathbf{0} \\ \mathbf{0} & m_2 \mathbf{I}_3 \end{pmatrix} \quad . \tag{36}$$

The constraints are related to the constancy of the lengths of the rigid rods:

$$\begin{aligned} g_1(\mathbf{q}) &= \frac{1}{2}\left(\mathbf{q}^1 \cdot \mathbf{q}^1 - l_1^2\right) = 0 \\ g_2(\mathbf{q}) &= \frac{1}{2}\left((\mathbf{q}^2 - \mathbf{q}^1) \cdot (\mathbf{q}^2 - \mathbf{q}^1) - l_2^2\right) = 0 \end{aligned} \tag{37}$$

The constraints restrict possible configurations to the manifold $S^2_{l_1} \times S^2_{l_2}$ consisting of two spheres, one about the origin with radius l_1 and one about the first mass with radius l_2.

In the following the motion of the pendulum with unit masses and rods of unit length is calculated. The pendulum starts at a horizontal initial position with zero initial velocity. The absolute value of the gravitational acceleration is $g = 9.81$.

Remark 4.1 All schemes investigated conserve the total energy and the angular momentum with respect to the gravitational axis up to numerical errors. This is shown exemplarily in Fig. 2.

4.1.1 Lagrange Multiplier Method

For the Lagrange Multiplier Method, the discrete energy-momentum system (being a special case of (23)) for the motion of the double spherical pendulum takes the form

$$\begin{aligned} \frac{\mathbf{q}_{n+1} - \mathbf{q}_n}{h} &= \mathbf{M}^{-1} \mathbf{p}_{n+\frac{1}{2}} \\ \frac{\mathbf{p}_{n+1} - \mathbf{p}_n}{h} &= \mathbf{M} g \begin{pmatrix} \mathbf{e}_2 \\ \mathbf{e}_2 \end{pmatrix} - \begin{pmatrix} \mathbf{q}^1_{n+\frac{1}{2}} & -\mathbf{q}^2_{n+\frac{1}{2}} + \mathbf{q}^1_{n+\frac{1}{2}} \\ \mathbf{0} & \mathbf{q}^2_{n+\frac{1}{2}} - \mathbf{q}^1_{n+\frac{1}{2}} \end{pmatrix} \boldsymbol{\lambda}_{n+1} \\ \mathbf{g}(\mathbf{q}_{n+1}) &= \mathbf{0} \quad . \end{aligned} \tag{38}$$

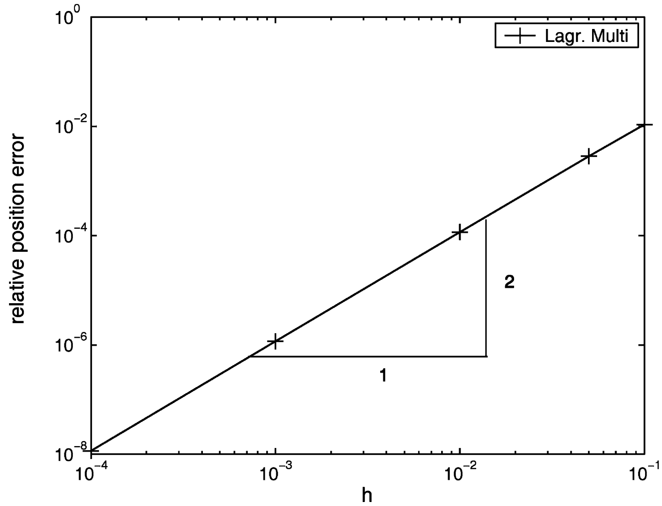

Fig. 3. Relative position error for the Lagrange Multiplier Method, reference solution calculated with $h = 10^{-5}$

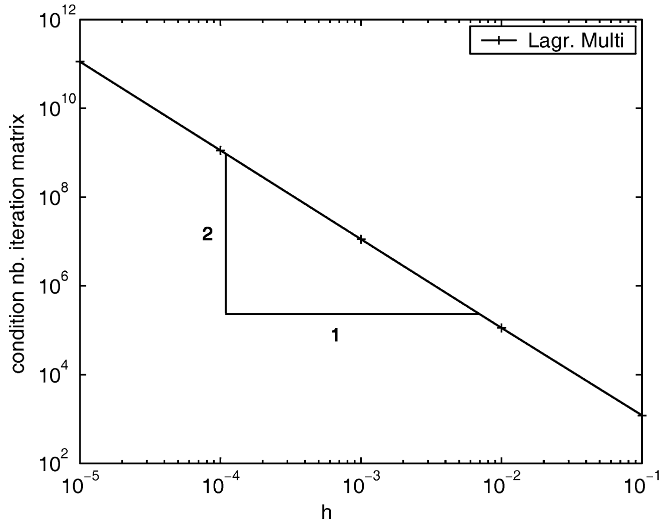

Fig. 4. Proportionality of the condition number of the iteration matrix to $1/h^2$

This system is in accordance with the example in [7] and with the mG(1) method in [3]. According to $(38)_3$, for the Lagrange Multiplier Method the constraints are fulfilled numerically exact (up to the order $\mathcal{O}(10^{-16})$). This scheme is second order accurate, one can see in Fig. 3 that the calculated solutions converge to a reference solution, which has been calculated with a time step $h = 10^{-5}$, quadratically as the time step decreases.

The dependence of the condition number of the iteration matrix on the time step is depicted in Fig. 4.

4.1.2 Penalty Method

For the double spherical pendulum the Penalty Method can be interpreted as a replacement of the rods by springs of stiffness μ. The discrete energy-momentum-conserving equations of motion with $P_{Pen}(\mathbf{g}(\mathbf{q})) = \mu \|\mathbf{g}(\mathbf{q})\|^2$ are given by

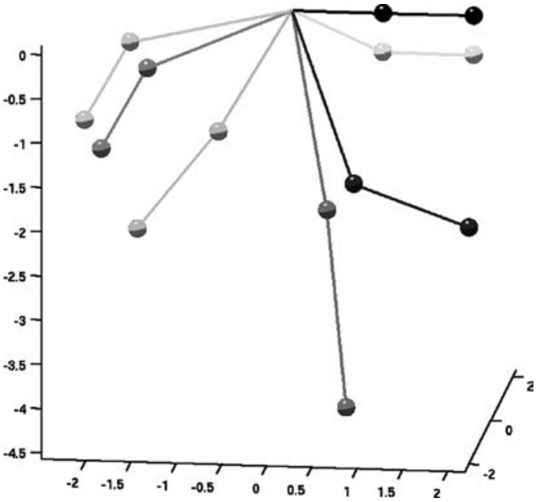

Fig. 5. Motion of the double spherical pendulum, $\mu = 5$

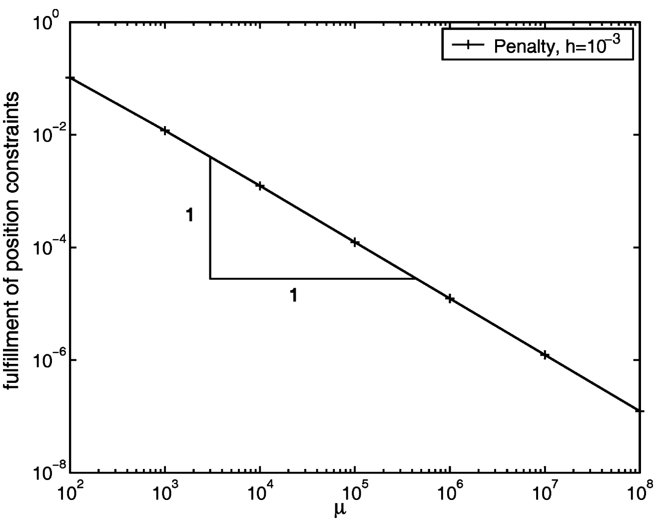

Fig. 6. Influence of μ on constraint fulfillment, $h = 10^{-3}$

$$
\begin{aligned}
\frac{\mathbf{q}_{n+1}-\mathbf{q}_n}{h} &= \mathbf{M}^{-1}\mathbf{p}_{n+\frac{1}{2}} \\
\frac{\mathbf{p}_{n+1}-\mathbf{p}_n}{h} &= \mathbf{M}g\begin{pmatrix}\mathbf{e}_2\\ \mathbf{e}_2\end{pmatrix} \\
&\quad - 2\mu\frac{(g_1(\mathbf{q}_{n+1}))^2-(g_1(\mathbf{q}_n))^2}{\left\|\mathbf{q}_{n+1}^1\right\|^2-\left\|\mathbf{q}_n^1\right\|^2}\begin{pmatrix}\mathbf{q}_{n+\frac{1}{2}}^1\\ \mathbf{0}\end{pmatrix} \\
&\quad - 2\mu\frac{(g_2(\mathbf{q}_{n+1}))^2-(g_2(\mathbf{q}_n))^2}{\left\|\mathbf{q}_{n+1}^2-\mathbf{q}_{n+1}^1\right\|^2-\left\|\mathbf{q}_n^2-\mathbf{q}_n^1\right\|^2}\begin{pmatrix}-\mathbf{q}_{n+\frac{1}{2}}^2+\mathbf{q}_{n+\frac{1}{2}}^1\\ \mathbf{q}_{n+\frac{1}{2}}^2-\mathbf{q}_{n+\frac{1}{2}}^1\end{pmatrix}
\end{aligned}
\tag{39}
$$

The extension of the springs with stiffness $\mu = 5$ can be seen clearly in Fig. 5, Figs. 6 and 7 show the statements of Proposition 3.2. The fulfillment of the constraints improves and the solution of the penalty system for the double spherical pendulum (39) converges to that of the corresponding Lagrange Multiplier system (38) as the penalty parameter increases.

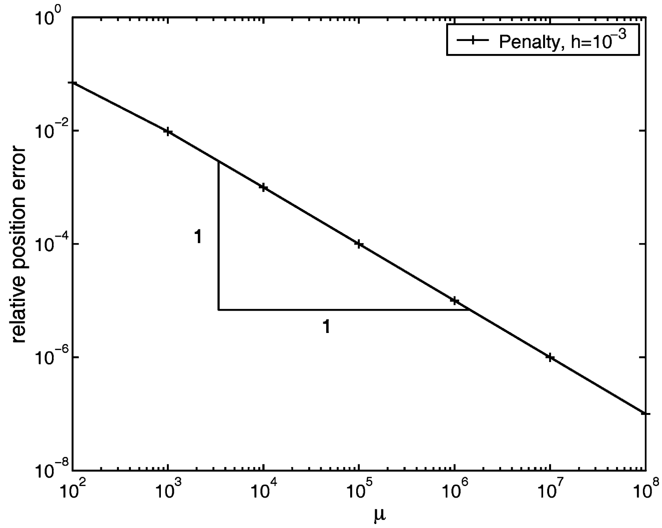

Fig. 7. Relative error of the position calculated by the Penalty Method with increasing μ, $h = 10^{-3}$. Reference solution calculated by the Lagrange Multiplier Method, $h = 10^{-3}$

4.1.3 Augmented Lagrange Method

Calculating the motion with the Augmented Lagrange Method means to solve the system

$$
\begin{aligned}
\frac{\mathbf{q}_{n+1}^k-\mathbf{q}_n}{h} &= \mathbf{M}^{-1}\mathbf{p}_{n+\frac{1}{2}}^k \\
\frac{\mathbf{p}_{n+1}^k-\mathbf{p}_n}{h} &= \mathbf{M}g\begin{pmatrix}\mathbf{e}_2\\ \mathbf{e}_2\end{pmatrix} - \Bigg[\lambda_{n+1}^{1,k} \\
&\qquad + 2\mu\frac{(g_1(\mathbf{q}_{n+1}^k))^2-(g_1(\mathbf{q}_n))^2}{\left\|\mathbf{q}_{n+1}^{k,1}\right\|^2-\left\|\mathbf{q}_n^1\right\|^2}\Bigg]\begin{pmatrix}\mathbf{q}_{n+\frac{1}{2}}^{1,k}\\ \mathbf{0}\end{pmatrix} \\
&\qquad - \left[\lambda_{n+1}^{2,k} + 2\mu\frac{(g_2(\mathbf{q}_{n+1}^k))^2-(g_2(\mathbf{q}_n))^2}{\left\|\mathbf{q}_{n+1}^{2,k}-\mathbf{q}_{n+1}^{1,k}\right\|^2-\left\|\mathbf{q}_n^2-\mathbf{q}_n^1\right\|^2}\right] \\
&\qquad \times \begin{pmatrix}-\mathbf{q}_{n+\frac{1}{2}}^{2,k}+\mathbf{q}_{n+\frac{1}{2}}^{1,k}\\ \mathbf{q}_{n+\frac{1}{2}}^{2,k}-\mathbf{q}_{n+\frac{1}{2}}^{1,k}\end{pmatrix} \\
\boldsymbol{\lambda}_{n+1}^{k+1} &= \boldsymbol{\lambda}_{n+1}^k + \mu 2\mathbf{g}(\mathbf{q}_{n+1}^k)
\end{aligned}
\tag{40}
$$

iteratively until the desired accuracy has been reached for the constraint fulfillment. The greater the penalty parameter is, the fewer iterations are required to reach this accuracy, since for high penalty parameters, the constraints are already fulfilled to some degree in the first iteration. We depict the results for the penalty parameter $\mu = 10^7$ and time step $h = 10^{-3}$ to corroborate the statements of Proposition 3.5.

Within two AL-iterations the error in the fulfillment of the constraints drops under the required tolerance of 10^{-10}, see Fig. 8. Furthermore, the table shows the convergence of the multipliers to the exact Lagrangian Multiplier and the convergence of the calculated position to that of the corresponding Lagrange Multiplier system.

AL-iteration	error of constraints	multipliers	position
1	$2.6064 \cdot 10^{-10}$	$3.625 \cdot 10^{-3}$	$4.161 \cdot 10^{-10}$
2	$0.8824 \cdot 10^{-10}$	$2.7431 \cdot 10^{-3}$	$1.4269 \cdot 10^{-10}$

Fig. 8. Convergence of the solution of the Augmented Lagrange Method, $\mu = 10^7, h = 10^{-3}$, to the solution of the Lagrange Multiplier Method, $h = 10^{-3}$

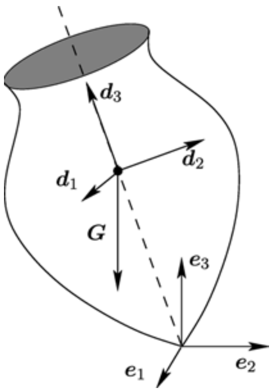

Fig. 9. Heavy symmetrical top with director triad in the center of gravity

4.2
Rigid body motion

This example deals with a heavy symmetrical top in Fig. 9. One point on the symmetry axis is fixed at the origin and the gravitational force acts in the direction of $-\mathbf{e}_3$. The equations of motion for the rigid body are formulated as a Hamiltonian system subject to holonomic constraints. The configuration variables are the position of the center of mass $\boldsymbol{\varphi}(t) \in \mathbb{R}^3$ and a triad $\{\mathbf{d}_1, \mathbf{d}_2, \mathbf{d}_3\}$ of directors $\mathbf{d}_i(t) \in \mathbb{R}^3$. See [2] for a detailed exposition of this formulation, the specific parameters and the initial conditions for this example. For simplicity we assume that the directions of the principal axes of inertia of the top coincide with those of the director triad. The assumption of rigidity gives rise to $m = 6$ orthonormality constraints:

$$\mathbf{0} = \mathbf{g}(\mathbf{d}_1, \mathbf{d}_2, \mathbf{d}_3) = \begin{pmatrix} \frac{1}{2}(\mathbf{d}_1 \cdot \mathbf{d}_1 - 1) \\ \frac{1}{2}(\mathbf{d}_2 \cdot \mathbf{d}_2 - 1) \\ \frac{1}{2}(\mathbf{d}_3 \cdot \mathbf{d}_3 - 1) \\ \mathbf{d}_1 \cdot \mathbf{d}_2 \\ \mathbf{d}_1 \cdot \mathbf{d}_3 \\ \mathbf{d}_2 \cdot \mathbf{d}_3 \end{pmatrix} \tag{41}$$

4.2.1
Penalty Method

The motion of the center of gravity, calculated with $\mu = 10^7, h = 10^{-3}$ with the Penalty Method is depicted in Fig. 10. It coincides with the results documented in [2], which have been calculated using the Lagrange Multiplier Method.

The Penalty Method results show qualitatively the same behaviour as for the double spherical pendulum

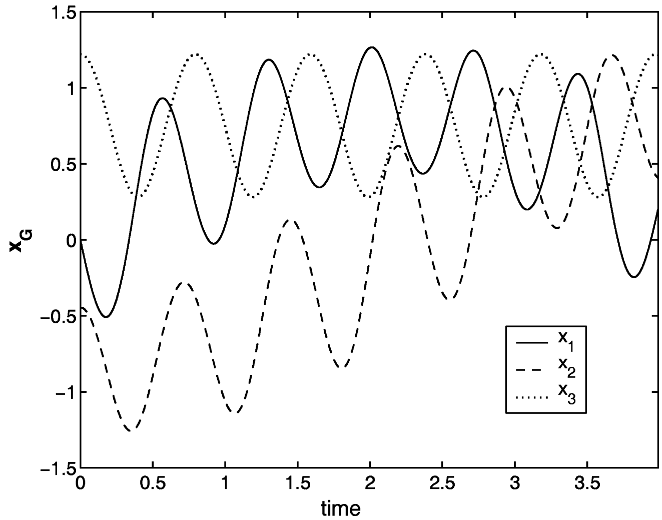

Fig. 10. Motion of the center of mass, $\mu = 10^7, h = 10^{-3}$

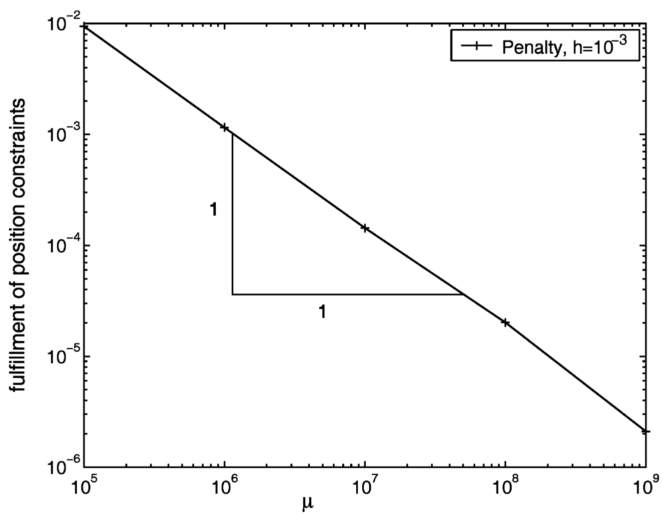

Fig. 11. Influence of μ on constraint fulfillment, $h = 10^{-3}$

(particularly they affirm the statements of Proposition 3.2), but they are of minor accuracy. Comparing Fig. 11 to Fig. 6 and Fig. 12 to Fig. 7 it turns out, that the fulfillment of the rigid body constraints as well as the relative position error are about three orders of magnitude worse than for the double spherical pendulum. This is one drawback of the Penalty Method, the more involved the problems get, the greater are the penalty parameters required to enforce the constraints.

4.2.2
Augmented Lagrange Method

Also the results of the Augmented Lagrange Method for the motion of the heavy symmetrical top corroborate the statements of Proposition 3.5. For comparison to the double spherical pendulum, we depict the results for $\mu = 10^7$ and $h = 10^{-3}$. For a wide range of time steps, the algorithm for the more complicated problem requires six Augmented Lagrange iterations to reach the accuracy of 10^{-10} for the fulfillment of the position constraints, see

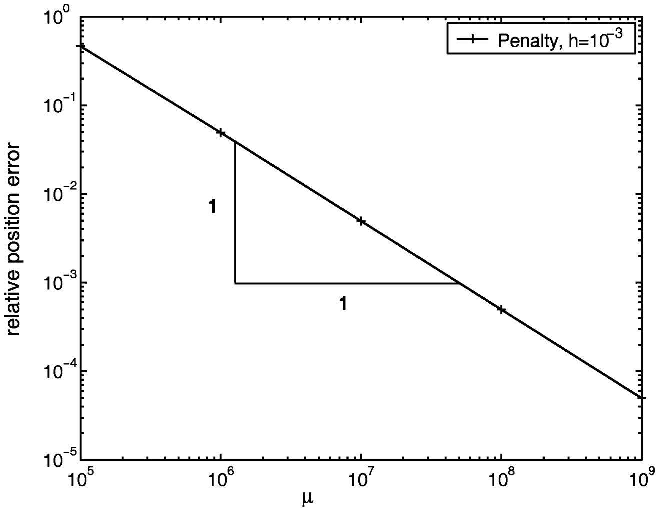

Fig. 12. Relative error of the position calculated by the Penalty Method. Reference solution calculated by the Lagrange Multiplier Method, $h = 10^{-3}$

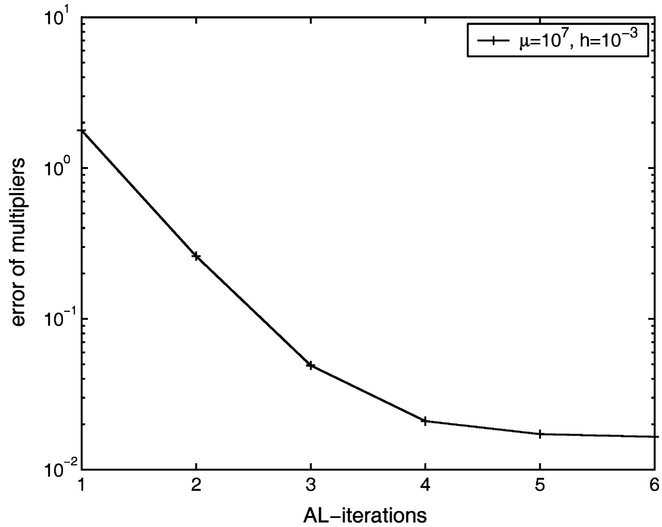

Fig. 14. Error of the multipliers calculated with the Augmented Lagrange Method, with respect to the exact Lagrangian Multiplier, $h = 10^{-3}$

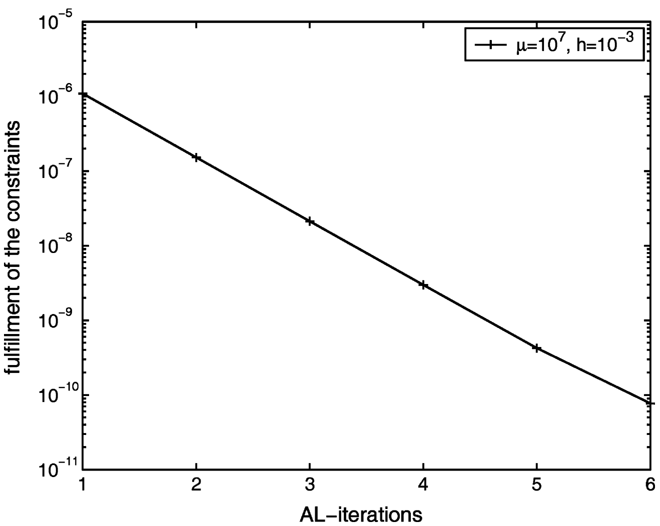

Fig. 13. Improvement of the fulfillment of the constraints during AL-iterations

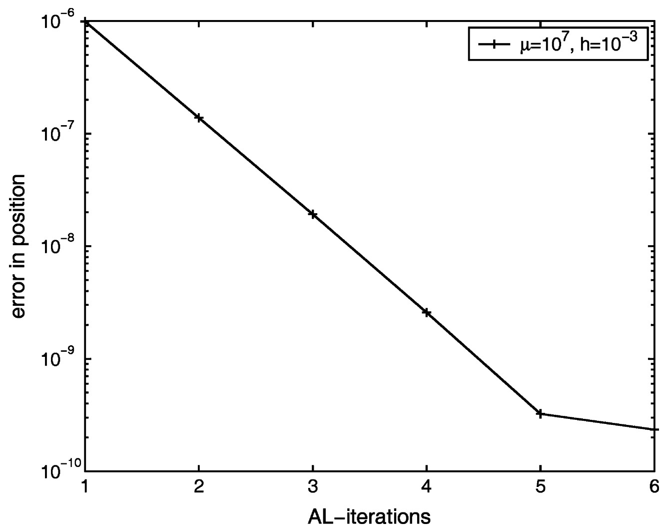

Fig. 15. Error of the position calculated with the Augmented Lagrange Method, with respect to the position calculated by the Lagrangian Multiplier Method, $h = 10^{-3}$

Fig. 13. The convergence of the multipliers to the exact Lagrangian Multiplier and the convergence of the position to that of the corresponding Lagrange Multiplier system can be seen in Fig. 14 and Fig. 15, respectively.

	unknowns	constraints	condition
Lagrange Multipliers	$2n + m$	num. zero	$\mathcal{O}\left(\frac{1}{h^2}\right)$
Penalty	$2n$	not known	$\mathcal{O}(\mu^2 h^2)$
Augmented Lagrange	$2n$	num. tol	$\mathcal{O}(\mu^2 h^2)$

Fig. 16. Theoretical aspects of the methods

5 Comparison and conclusions

Figure 16 shows some aspects for comparison of the three methods. The number of unknowns using the Lagrange Multiplier Method is generally higher than for the other two methods. For a physical system in n-dimensional configuration space subject to m holonomic constraints, the Lagrange Multiplier systems consists of $2n + m$ equations whereas the dimension of the Penalty system and the Augmented Lagrange system is $2n$.

The advantage of the Lagrange Multiplier Method is certainly that the constraints are fulfilled numerical exactly. For the Penalty Method, accuracy can hardly be predicted, it is strongly problem dependent as could be seen in the examples. For the Augmented Lagrange Method, any reasonable accuracy can be reached but this must be paid with a certain number of extra iterations, which is again dependent on the complexity of the problem under consideration.

In the limit for decreasing time steps, the condition number of the iteration matrix of the Lagrange Multiplier system is certainly worse than that of the other two systems, but since the choice of a penalty parameter and a time step for a concrete problem depends strongly on ones intention and on the problem itself, the condition numbers can hardly be compared theoretically.

Concerning the numerical costs our tests indicated that the Penalty Method can compete with the Lagrange Multiplier Method whereas the Augmented Lagrange Method, in most cases required considerably more computational time.

References

1. **Bertsekas DP** (1995) Nonlinear Programming. Athena Scientific, MAT Bert.
2. **Betsch P, Steinmann P** (2001) Constrained integration of rigid body dynamics. Comput. Meth. Appl. Mech. Eng. 191: 467–488
3. **Betsch P, Steinmann P** (2002) Conserving properties of a time FE method – Part III: Mechanical systems with holonomic constraints. Int. J. Numer. Meth. Eng. 53: 2271–2304
4. **Blajer W** (2002) Augmented Lagrangian formulation: geometrical interpretation and application to systems with singularities and redundancy. Multibody System Dynamics 8: 141–159
5. **Bornemann FA, Schütte C** (1996) Homogenization of Hamiltonian systems with a strong constraining potential. Physica D 102: 57–77
6. **Gonzalez O** (1996) Time integration and discrete Hamiltonian systems. J. Nonlinear Science 6: 449–467
7. **Gonzalez O** (1999) Mechanical systems subject to holonomic constraints: differential-algebraic formulations and conservative integration. Physica D 132: 165–174
8. **Knoerr AP** (1998) A dynamical proof of the method of Lagrange. Soc. Industrial Appl. Math. 40(4): 941–944
9. **Luenberger DG** (1984) Linear and nonlinear programming. Addison-Wesley
10. **Marsden JE, Ratiu TS** (1994) Introduction to mechanics and symmetry. A basic exposition of classical mechaincal systems. Texts in Applied Mathematics 17. Springer-Verlag
11. **Petzold L, Loetstedt P** (1986) Numerical solution of nonlinear differential equations with algebraic constraints II: practical implications. Soc. Industrial Appl. Math. 7(3): 720–733
12. **Rheinboldt WC** (1984) Differential-algebraic systems as differential equations on manifolds. Math. Computation 43(168): 473–482
13. **Rubin H, Ungar P** (1957) Motion under a strong constraining force. Commun. Pure Appl. Math. 10(1): 65–87