J Nonlinear Sci
DOI 10.1007/s00332-008-9030-1

Journal of
Nonlinear
Science

Γ-convergence of Variational Integrators for Constrained Systems

Bernd Schmidt · Sigrid Leyendecker · Michael Ortiz

Received: 23 October 2007 / Accepted: 25 July 2008

Abstract For a physical system described by a motion in an energy landscape under holonomic constraints, we study the Γ-convergence of variational integrators to the corresponding continuum action functional and the convergence properties of solutions of the discrete Euler–Lagrange equations to stationary points of the continuum problem. This extends the results in Müller and Ortiz (J. Nonlinear Sci. 14:279–296, 2004) to constrained systems. The convergence result is illustrated with examples of mass point systems and flexible multibody dynamics.

Keywords Variational integrators · Constrained systems · Gamma-convergence

Mathematics Subject Classification (2000) 37M15 · 37J45 · 49J45 · 65P99 · 70E55 · 70F20

1 Introduction

In this paper, we investigate mechanical systems in an n-dimensional configuration space that can be described by the evolution $t \mapsto u(t) \in \mathbb{R}^n$ with time t subject to

Communicated by Robert V. Kohn.

B. Schmidt (✉)
Zentrum Mathematik, Technische Universität München, Boltzmannstr. 3, 85747 Garching, Germany
e-mail: schmidt@ma.tum.de

S. Leyendecker · M. Ortiz
Division of Engineering and Applied Science, California Institute of Technology, Pasadena, CA 91125, USA

the potential $V : \mathbb{R}^n \to \mathbb{R}$ and subject to holonomic constraints. In Lagrangian mechanics, the physical trajectories of this motion arise as stationary points of the corresponding action functional I which is the kinetic energy minus the potential energy integrated along the trajectory. In the presence of holonomic constraints, modeled by the so-called constraint manifold $M \subset \mathbb{R}^n$, requiring in addition that the trajectories must lie on M, gives rise to the constrained functional I_M.

The theory of discrete variational integrators provides approximations of such systems where now time is viewed as a discrete variable. By discretizing the action functional I, resp., I_M, in time, we arrive at discrete action functionals I^h, resp., I^h_M. (Here, we will concentrate on piecewise linear interpolations of trajectories.) A discrete version of Hamilton's principle of stationary action applied to these action sums leads to discrete Euler–Lagrange equations whose solutions should be approximations to the continuum motion. See Marsden and West (2001), and the references therein, for a general introduction to the theory of variational integrators. Many structure preserving integration schemes for ODEs and their properties are investigated extensively in Hairer et al. (2006). Concerning the symplectic integration of constrained systems, work has been done, e.g., by Jay (1996, 1998), Leimkuhler and Reich (1994), Reich (1996) and for nonholonomic systems, e.g., by McLachlan and Perlmutter (2006).

It has been first noticed by Ortiz and Müller that the theory of Γ-convergence is a convenient tool to investigate the convergence properties of the discrete approximations to the continuum trajectories. (For a general introduction to the theory of Γ-convergence see, e.g., Dal Maso (1993).) In Müller and Ortiz (2004), they show that in fact I^h Γ-converges to I and apply this result to prove that discrete stationary points of I^h converge to continuum stationary points of I. Their work as been extended to more general Lagrangians by Maggi and Morini (2004). Γ-convergence provides a powerful albeit as yet not widely used tool for understanding convergence of dynamical problems. Thus, Γ-convergence establishes convergence of solutions in a global, instead of merely local, sense. In particular, it allows comparing infinite wave trains. This is in contrast to conventional methods of analysis, such as Gronwall's inequality (e.g., Marsden and West 2001) that merely provide exponentially divergent local bounds on discretization errors. The global nature of Γ-convergence is in analogy to the traditional phase-error analysis of time-stepping algorithms for linear systems, which regards convergence in terms of dispersion relations (e.g., Belytschko and Mullen 1976; Belytschko 1981; Hughes 1987). However, Γ-convergence applies to much more general, possibly strongly nonlinear dynamical systems. The main goal of the present paper is to extend the results of Müller and Ortiz (2004) to systems with holonomic constraints, i.e., motions confined to lie on some constraint manifold $M \subset \mathbb{R}^n$. (Since our variational integrators will be derived by piecewise linear interpolation of the continuum trajectories, we will assume a linear ambient configuration space $\mathbb{R}^n$ for M.)

In order to describe such constrained systems, it will often be convenient to work in local coordinates for M—at the expense of a more complicated form of the action functional. The resulting Lagrangians are, in fact, of the form considered in Maggi and Morini (2004). However, note carefully that for the discretized functional I^h_M, the constraint is enforced *only* at the nodal points of the underlying triangulation. As a consequence, our results cannot readily be inferred from the corresponding results in Maggi and Morini (2004).

On the contrary, for constrained systems, we will encounter new phenomena which are related to the fact that discrete trajectories are in general nonunique for given initial positions and velocities. The main novelty introduced here for constrained systems is a selection criterion for the physically relevant solutions that firstly guarantees that discrete trajectories do in fact converge to the continuum motion, and secondly is satisfied in numerical implementations of the scheme. Analytically, the main difficulty to overcome is to obtain improved regularity for the solutions of the discrete Euler–Lagrange equation. Whereas, e.g., an L^∞-bound on positions implies an L^∞-bound on velocities for unconstrained systems, the corresponding result is no longer true for constrained motions without further assumptions.

We complement our analysis by studying two constrained mechanical systems numerically. Holonomic constraints arise naturally in the description of multibody dynamics comprising rigid, elastic, or both types of components. As an easy, yet nontrivial example, we will discuss a double spherical pendulum in some detail. In fact, hard configurational constraints are also obtained for systems in the realm of finite elasticity in the limit of singular geometries, e.g., for plates, beams, etc. We will also give an example showing that our results apply to (finite element approximations of) multibody dynamical systems comprising rigid and elastic components. Note that in both examples, in addition to Theorem 3.4, our numerical computations also provide the rate of convergence of the approximating trajectories.

A more precise account of the results in Müller and Ortiz (2004) and of our set-up is as follows.

Let $X = L^2_{\mathrm{loc}}(\mathbb{R}, \mathbb{R}^n)$, and by $\mathcal{E}$ denote the collection of all open bounded intervals of $\mathbb{R}$. Note that X is a complete metric space when endowed with the distance function inferred from the seminorms $\|u\|_{L^2((-k,k),\mathbb{R}^n)}$, $k \in \mathbb{N}$.

Let $m > 0$ and $V \in C(\mathbb{R}^n)$. The unconstrained action functional $I : X \times \mathcal{E} \to \mathbb{R} \cup \{\infty\}$ is defined by

$$I(u, A) = \begin{cases} \int_A \frac{m}{2} |\dot{u}(t)|^2 - V(u(t))\, \mathrm{d}t, & u \in H^1(A, \mathbb{R}^n), \\ +\infty, & \text{otherwise.} \end{cases}$$

If $V \in C^1$, then the first variation of $I(\cdot, A)$ is given by

$$\begin{aligned} \delta I(u, \varphi, A) &= \frac{\mathrm{d}}{\mathrm{d}r}\bigg|_{r=0} I(u + r\varphi, A) \\ &= \int_A m\dot{u}(t) \cdot \dot{\varphi}(t) - \nabla V\big(u(t)\big) \cdot \varphi(t)\, \mathrm{d}t \end{aligned}$$

for $u \in H^1(A, \mathbb{R}^n)$, $\varphi \in C_c^\infty(A, \mathbb{R}^n)$. (Note that by the Sobolev embedding theorem $u \in H^1(A, \mathbb{R}^n)$ implies that u—modified on a set of measure zero—lies in $C(\overline{A}, \mathbb{R}^n)$.) We call u a stationary point of I if

$$I(u, A) < \infty \quad \text{and} \quad \delta I(u, \varphi, A) = 0$$

for all $A \in \mathcal{E}$ and $\varphi \in C_c^\infty(A, \mathbb{R}^n)$.

Suppose $\mathcal{T}_h$ is a partition of $\mathbb{R}$ of size h, i.e., $\mathcal{T}_h = \{t_i : i \in \mathbb{Z}\}$ for some $\cdots < t_i < t_{i+1} < \cdots$ such that $|t_{i+1} - t_i| \le h$ and $t_i \to \pm\infty$ if $i \to \pm\infty$. Let X_h be the subspace of X consisting of continuous functions such that $u_{|(t_i,t_{i+1})}$ is affine $\forall t_i \in \mathcal{T}_h$. The unconstrained discrete action functionals $I^h : X \times \mathcal{E} \to \mathbb{R}$ are defined to be

$$I^h(u, A) = \begin{cases} I(u, A), & u \in X_h, \\ +\infty, & \text{otherwise}. \end{cases}$$

The stationary points of I^h are elements u_h of X_h such that

$$I(u_h, A) < \infty \quad \text{and} \quad \delta I(u_h, \varphi_h, A) = 0$$

for all $A \in \mathcal{E}$ and $\varphi_h \in X_h$ with $\varphi_h = 0$ on $\mathbb{R} \setminus A$. Note that if $u_h = 0$ on $\mathbb{R} \setminus A$, then setting $u_i := u_h(t_i)$, we can write

$$I^h(u_h, A) = \sum_{i=\mu}^{\nu-1} L_d(u_i, u_{i+1})$$

where $(t_\mu, t_\nu) \subset A$ is the maximal subinterval of A compatible with $\mathcal{T}_h$ and

$$L_d(u_i, u_{i+1}) = \frac{m}{2}\frac{(u_{i+1} - u_i)^2}{t_{i+1} - t_i} - \int_{t_i}^{t_{i+1}} V\left(\frac{t_{i+1} - t}{t_{i+1} - t_i}u_i + \frac{t - t_i}{t_{i+1} - t_i}u_{i+1}\right) \mathrm{d}t. \quad (1)$$

So, stationary points of I^h are solutions of the discrete Euler–Lagrange equations

$$\nabla_2 L_d(u_{i-1}, u_i) + \nabla_1 L_d(u_i, u_{i+1}) = 0, \quad i = \mu + 1, \dots, \nu - 1.$$

The connection between I and its discrete counterpart I^h is studied in detail in Müller and Ortiz (2004), where, in particular, it is shown that if $V \in C(\mathbb{R}^n)$ with $V(s) \le C(1 + |s|^2)$, then:

- For all $A \in \mathcal{E}$, $\Gamma\text{-}\lim_{h\to 0} I^h(\cdot, A) = I(\cdot, A)$ in X. (For the definition of Γ-convergence, see below.)
- If in addition $V \in C^2$ with $|\nabla^2 V| \le C$ for some constant C, then sequences of stationary points of I^h that are uniformly bounded converge—up to subsequences—weakly* in $W^{1,\infty}$ to some $u \in W^{2,\infty}$. Moreover the limiting trajectory u is a stationary point of the continuum action functional I.
- If moreover the Fourier transforms $\hat{u}_h$ of u_h are uniformly bounded Radon measures such that no mass leaks to infinity as $h \to 0$, then $\hat{u}_h \to \hat{u}$ as measures in the flat norm.

(Also compare the results in Maggi and Morini (2004) for more general functionals I.)

Here, we are in particular interested in mechanical systems with holonomic constraints. This can be modeled by requiring that $u \in M$ a.e. for some suitable (k-dimensional) submanifold M of $\mathbb{R}^n$ (the "constraint manifold"), which we will assume to be at least of class C^3. Accordingly, we define the constrained action functional $I_M : X \times \mathcal{E} \to \mathbb{R} \cup \{\infty\}$ by

$$I_M(u, A) = \begin{cases} I(u, A), & u \in M \text{ a.e. on } A, \\ +\infty, & \text{otherwise.} \end{cases}$$

The constrained discrete action functionals $I_M^h : X \times \mathcal{E} \to \mathbb{R}$ are

$$I_M^h(u, A) = \begin{cases} I^h(u, A), & u(t) \in M \ \forall t \in \mathcal{T}_h \cap A, \\ +\infty, & \text{otherwise.} \end{cases}$$

In view of our examples in Sect. 4, let us also mention that our results can be extended in a straightforward way to systems with a general positive definite mass matrix $\boldsymbol{m}$, i.e., whose kinetic energy is given by $\frac{1}{2}\dot{u}^{\mathrm{T}}\boldsymbol{m}\,\dot{u}$ rather than $\frac{m}{2}|\dot{u}|^2$.

The stationary points for constrained systems are most conveniently defined in terms of local coordinates for M. Suppose $A \in \mathcal{E}$ and u is such that $I_M(u, A)$ is finite. (So, in particular, u is continuous and takes values in M on A.) Assume that $u(\overline{A})$ is covered by the domain $U \subset M$ of a single coordinate chart, whose inverse is denoted $\psi : \mathbb{R}^k \supset V \to U$. Define the curve $v : A \to V$ by $u_{|A} = \psi \circ v$. Then u is said to be a stationary point of $I_M(\cdot, A)$ if v is a stationary point of

$$\tilde{J}_\psi(v, A) := I(\psi \circ v, A),$$

which means that

$$\delta \tilde{J}_\psi(v, \varphi, A) = \frac{\mathrm{d}}{\mathrm{d}r}\bigg|_{r=0} \tilde{J}_\psi(v + r\varphi, A) = 0$$

for all $\varphi \in C_c^\infty(A, \mathbb{R}^k)$. (By density, it follows that then, in fact, $\tilde{J}_\psi(v, \varphi, A) = 0$ for all $\varphi \in H_0^1(A, \mathbb{R}^n)$. It is not hard to see that this, in particular, implies that stationary points of $I_M(\cdot, A)$ are well defined.) We say that u is a stationary point for I_M if there exists a covering $\mathbb{R} = \bigcup_{i \in I} A_i$ with $A_i \in \mathcal{E}$ such that $u(\overline{A_i})$ is covered by a single chart and $u_{|A_i}$ is a stationary point of $I(\cdot, A_i)$ for all i. (Using a partition of unity, it is again easy to see that this notion is well defined.)

With the notation introduced above the discrete stationary points of I_M^h are functions in $u_h \in X_h$ such that

$$\nabla_{v_i} \sum_{j=\mu}^{\nu-1} L_d\big(\psi(v_i), \psi(v_{i+1})\big) = 0,$$

where $\psi(v_j) = u_j$ or equivalently,

$$\nabla_2 L_d(u_{i-1}, u_i) + \nabla_1 L_d(u_i, u_{i+1}) \perp T_{u_i} M. \tag{2}$$

2 Γ-convergence

Our first aim is to obtain a Γ-convergence result for constrained systems. Recall that a sequence of functionals $F^h : Y \to [-\infty, \infty]$ on a metric space Y is said to Γ-converge to the functional F if the following two conditions are satisfied.

(i) ("lim inf-inequality") Whenever $y_h \to y$ in Y then

$$\liminf_{h\to 0} F^h(y_h) \geq F(y).$$

(ii) ("recovery sequence") For each $y \in Y$, there exists a sequence $y_h \to y$ such that

$$\lim_{h\to 0} F^h(y_h) = F(y).$$

Proposition 2.1 *Let $V \in C(\mathbb{R}^n)$ with $|V(s)| \leq C(1+|s|^2)$, $A \in \mathcal{E}$. Then $I_M^h(\cdot, A)$ Γ-converges in X and*

$$\Gamma - \lim_{h\to 0} I_M^h(\cdot, A) = I_M(\cdot, A).$$

We will first prove two preparatory results. Assume that $V \in C(\mathbb{R}^n)$ with $|V(s)| \leq C(1+|s|^2)$ throughout this section. Part (i) of the following lemma is contained in Müller and Ortiz (2004).

Lemma 2.2 *Let $A \in \mathcal{E}$.*

(i) *Then $I(\cdot, A)$ is lower semicontinuous in X and continuous in $H^1(A, \mathbb{R}^n)$.*
(ii) *If, for a sequence $h_k \to 0$, $u_{h_k} \in X_{h_k}$ converges to u in X such that $I_M^h(u_{h_k}, A)$ is bounded. Then $u \in M$ a.e. on A.*

Proof (i) Due to our assumptions on V, $u \mapsto \int_A V(u)\,\mathrm{d}t$ is continuous on $L^2(A, \mathbb{R}^n)$, and thus on X and on $H^1(A, \mathbb{R}^n)$. Now, $u \mapsto \int_A \frac{m}{2}\dot{u}^2\,\mathrm{d}t$ is clearly continuous on $H^1(A, \mathbb{R}^n)$, which proves the second claim. But it is also lower semicontinuous on $L^2(A, \mathbb{R}^n)$ since it is lower semicontinuous $H^1(A, \mathbb{R}^n)$ with respect to the weak topology and takes the value ∞ outside $H^1(A, \mathbb{R}^n)$.

(ii) $I_M^h(u_{h_k}, A)$ and $\int_A V(u_{h_k})$ being bounded, in fact, u_{h_k} converges weakly to u in $H^1(A, \mathbb{R}^n)$. So, by the Sobolev embedding theorem, we may assume that $u_{h_k} \to u$ uniformly in $C(\overline{A}, \mathbb{R}^n)$. But again because $I_M^h(u_{h_k}, A)$ is bounded, $u_{h_k}(t) \in M$ for all $t \in \mathcal{T}_h \cap A$, and the claim follows. □

We will also need the following approximation result extending the corresponding assertion in Müller and Ortiz (2004) to our setting of constrained Lagrangians.

Lemma 2.3 *Let $A \in \mathcal{E}$. For every $u \in X$ with $I_M(u, A) < \infty$, there is a sequence $u_h \in X_h$ such that $u_h \to u$ in X and $u_{h|A} \to u_{|A}$ in $H^1(A, \mathbb{R}^n)$ and $u_h(t) \in M$ for all $t \in \mathcal{T}_h \cap A$.*

Proof Let $A = (a, b)$, $\eta \in C_c^\infty(-1, 1)$ be a standard mollifier and define $\eta_h(x) = h^{-1}\eta(x/h)$. Let $N_h w$ denote the nodal interpolation of a function w with respect to the triangulation $\mathcal{T}_h$. By the Sobolev embedding theorem, we may assume that $u \in C([a, b], \mathbb{R}^n)$. As in Müller and Ortiz (2004), we define approximations of u

which are continuous in the slightly larger interval $(a-2h, b+2h)$:

$$v_h(t) = \begin{cases} u(t), & t \le a-2h, \\ u(a), & a-2h < t \le a, \\ u(t), & a < t < b, \\ u(b), & b \le t < b+2h, \\ u(t), & t \ge b+2h. \end{cases}$$

With the help of standard estimates on nodal interpolations and convolutions such as

$$\begin{aligned} &\int_r^s \left| \frac{\mathrm{d}}{\mathrm{d}t}(N_h w - w) \right|^2 \mathrm{d}t \le C \int_{r-h}^{s+h} |\dot{w}|^2 \,\mathrm{d}t, \\ &\int_r^s |\eta_h * w|^2 \,\mathrm{d}t \le C \int_{r-h}^{s+h} |w|^2 \,\mathrm{d}t \end{aligned} \tag{3}$$

(which are included here for later reference), in Müller and Ortiz (2004), it is shown that

$$T_h v_h := N_h(\eta_h * v_h) \to u \quad \text{in } X \quad \text{and in } H^1\big(A, \mathbb{R}^n\big). \tag{4}$$

Let g, to be specified later, be a positive function on $(0,\infty)$ such that $g(h) \to 0$ as $h \to 0$. Since in a suitable neighborhood of M the orthogonal projection P of the ambient space $\mathbb{R}^n$ onto M is well defined, smooth, and globally Lipschitz, we may choose neighborhoods U_h of M and $\alpha : M \to (0,\infty)$ such that

$$\big\{x \in \mathbb{R}^n : \operatorname{dist}(x, M) \le \alpha(Px) g(h)\big\} \subset U_h \subset \big\{x \in \mathbb{R}^n : \operatorname{dist}(x, M) < g(h)\big\},$$

where $\alpha > c(K) > 0$ on compacts $K \subset M$, and functions $p_h \in C^2(\mathbb{R}^n, \mathbb{R}^n)$ that are globally Lipschitz continuous with Lipschitz constant independent of h such that $p_h \equiv P$ on U_h and $p_h \equiv \mathrm{id}$ on $\{x \in \mathbb{R}^n : \operatorname{dist}(x, M) \ge g(h)\}$. Set $S_h w := N_h(p_h \circ (\eta_h * w))$. Then $\|T_h v_h - S_h v_h\|_{L^\infty} \le g(h)$ and thus $S_h v_h \to u$ in X by (4).

Since u is uniformly continuous on $[a, b]$, we can choose $g = g(h)$ to be the modulus of continuity of $u_{|\overline{A}}$ such that if $|t-s| \le h$, then $|u(t) - u(s)| \le g(h)$ and $g(h) \to 0$ as $h \to 0$. But then

$$\operatorname{dist}\big(\eta_h * v_h(t), M\big) \le C g^2(h) \tag{5}$$

on $[a-h, b+h]$. So, by construction of p_h, we obtain that $p_h(\eta_h * v_h) \in M$ on A, and thus indeed $S_h v_h(t) \in M$ for all $t \in \mathcal{T}_h \cap A$.

It remains to estimate

$$\begin{aligned} \int_a^b \left| \frac{\mathrm{d}}{\mathrm{d}t}(S_h v_h - T_h v_h) \right|^2 \mathrm{d}t &\le C \int_{a-h}^{b+h} \left| \frac{\mathrm{d}}{\mathrm{d}t}\big(p_h(\eta_h * v_h) - \eta_h * v_h\big) \right|^2 \mathrm{d}t \\ &= C \int_{a-h}^{b+h} \big| \nabla p_h(\eta_h * v_h) \cdot (\eta_h * \dot{v}_h) - \eta_h * \dot{v}_h \big|^2 \,\mathrm{d}t, \end{aligned} \tag{6}$$

where the inequality followed from (3). Again by (5), $\nabla^2 p_h$ is bounded on $[\eta_h * v_h(t), v_h(t)]$, and thus

$$\big|\nabla p_h\big(\eta_h * v_h(t)\big) - \nabla p_h\big(v_h(t)\big)\big| \le C g(h). \tag{7}$$

For $t \in (a-h, b+h)$, we can decompose $\dot v_h(s)$ into one part $\dot v_h^{\|}(s)$ which lies in $T_{v_h(t)}M$ and an orthogonal part $\dot v_h^{\perp}(s)$. If $|t-s| \le h$, then $|\dot v_h^{\perp}(s)| \le Cg(h)|\dot v_h^{\|}(s)| \le Cg(h)|\dot v_h(s)|$. Noting that $\nabla p_h(v_h(t)) \cdot (\eta_h * \dot v_h^{\|}(t)) = \eta_h * \dot v_h^{\|}(t)$ we can estimate

$$\begin{aligned}
&\big|\nabla p_h\big(v_h(t)\big) \cdot \big(\eta_h * \dot v_h(t)\big) - \eta_h * \dot v_h(t)\big| \\
&\quad = \big|\nabla p_h\big(v_h(t)\big) \cdot \big(\eta_h * \dot v_h^{\perp}(t)\big) - \eta_h * \dot v_h^{\perp}(t)\big| \le Cg(h)\big(\eta_h * |\dot v_h|\big)(t).
\end{aligned}$$

It then follows from (3) that

$$\int_{a-h}^{b+h} \big|\nabla p_h(v_h) \cdot (\eta_h * \dot v_h) - \eta_h * \dot v_h\big|^2 \,\mathrm{d}t \le C g^2(h) \int_{a-2h}^{b+2h} |\dot v_h|^2 \,\mathrm{d}t. \tag{8}$$

Now combine (6), (7), and (8) and apply (3) once more to arrive at

$$\begin{aligned}
&\int_a^b \left|\frac{\mathrm{d}}{\mathrm{d}t}(S_h v_h - T_h v_h)\right|^2 \mathrm{d}t \\
&\le C g^2(h) \int_{a-h}^{b+h} |\eta_h * \dot v_h|^2 \,\mathrm{d}t + C \int_{a-h}^{b+h} \big|\nabla p_h(v_h) \cdot (\eta_h * \dot v_h) - \eta_h * \dot v_h\big|^2 \,\mathrm{d}t \\
&\le C g^2(h) \int_{a-2h}^{b+2h} |\dot v_h|^2 \,\mathrm{d}t \to 0
\end{aligned}$$

as $h \to 0$, so that indeed $S_h v_h \to u$ in $H^1(A, \mathbb{R}^n)$ by (4). □

Remark 2.4 Note that if $\dot u$ is bounded on A, then the above constructed approximations $S_h u_h$ satisfy $\|\frac{\mathrm{d}}{\mathrm{d}t}(S_h u_h)\|_{L^\infty(A,\mathbb{R}^n)} \le C\|\dot u\|_{L^\infty(A,\mathbb{R}^n)}$.

With these preparations, we can now prove Proposition 2.1.

Proof of Proposition 2.1 Let $u_h \in X$ be a sequence converging to $u \in X$. We may assume that $I_M^h(u_h, A)$ is bounded and so $u_h(t) \in M$ for all $t \in \mathcal{T}_h \cap A$ and $u \in M$ on A by Lemma 2.2(ii). Since $I^h(\cdot, A) \ge I(\cdot, A)$, we therefore obtain by Lemma 2.2(i)

$$\begin{aligned}
\liminf_{h\to 0} I_M^h(u_h, A) &= \liminf_{h\to 0} I^h(u_h, A) \ge \liminf_{h\to 0} I(u_h, A) \\
&\ge I(u, A) = I_M(u, A).
\end{aligned}$$

To provide a recovery sequence for $u \in X$ we may w.l.o.g. assume that $I_M(u, A) < \infty$. By Lemma 2.3, there is a sequence $u_h \in X_h$ such that $u_h \to u$ in X, $u_{h|A} \to u_{|A}$ in $H^1(A, \mathbb{R}^n)$ and $u_h(t) \in M$ for all $t \in \mathcal{T}_h \cap A$. So, by Lemma 2.2(i),

$$I_M^h(u_h, A) = I^h(u_h, A) = I(u_h, A) \to I(u, A) = I_M(u, A). \qquad \square$$

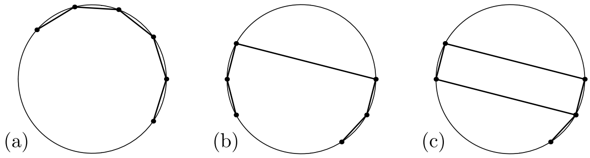


Fig. 1 Motion on the unit circle

3 Stationary Points

In this section, we will investigate the limiting behavior of a sequence u_h of stationary points of I_M^h. For the unconstrained case, it is shown in Müller and Ortiz (2004) (also compare Maggi and Morini 2004) that if u_h is stationary for I^h and $\|u_h\|_{L^\infty}$ is bounded independently of h, then also $\|u\|_{W^{1,\infty}}$ is bounded and in particular—for a subsequence—$u_h \overset{*}{\rightharpoonup} u$ for some u in $W^{1,\infty}$. Furthermore, u is a stationary point of the limiting functional I. We will first see that this does in general not hold for constrained systems. An easy example shows that a sequence of stationary points of I_M^h, although bounded in L^∞, might blow up in $W^{1,\infty}_{\text{loc}}$. However, we will see that this cannot happen if the action $I_M^h(u_h, A)$ over bounded intervals A remains bounded. In this case, our main result will be that in fact—up to subsequences—u_h converges weakly* in $W^{1,\infty}_{\text{loc}}$ to some stationary point u of I_M.

Example 3.1 (Motion on the unit circle) Suppose $M = S^1 \subset \mathbb{R}^2$. Let $\mathcal{T}_h = h\mathbb{Z}$, $m = 1$ and $V \equiv 0$. Let $u_i = u_h(ih)$. Then u_h is stationary if

$$u_{i+1} + u_{i-1} \in \mathbb{R}u_i.$$

The easiest example is given by $u_i = (\cos(i\alpha), \sin(i\alpha))$ where $\alpha = \alpha(h)$ is fixed. However, if $\alpha \gg h$, then $|\dot{u}| = |2\sin(\frac{\alpha}{2})h^{-1}|$ a.e. diverges.

But even if the trajectory behaves nicely initially, $|\dot{u}|$ does not have to remain bounded. Let, e.g., $u_i = (\cos(ih), \sin(ih))$ if $i \le 0$. For positive i, set $u_i = (\cos((i-2)h + \pi), \sin((i-2)h + \pi))$ (see Fig. 1(b)) or, even worse, $u_i = (-\cos(h), \sin(h)), (-1, 0), (\cos(h), -\sin(h)), (1, 0)$ for $i \equiv 1, 2, 3, 4 \bmod 4$, respectively (see Fig. 1(c)). In both cases, $\dot{u}$ blows up in $W^{1,\infty}$ (and in H^1).

Lemma 3.2 *Suppose $u \in X_h$ is a stationary point of I_M^h, $t^* \in \mathbb{R}$, $c > 0$ is a constant and $V \in C^2$. Let $A = (a, b) \in \mathcal{E}$ be a sufficiently small neighborhood of t^* and assume that $\|\dot{u}\|_{L^\infty(A,\mathbb{R}^n)} < c$. Then there exists a constant $C > 0$ such that, for all $u' \in X_h$ with $\|\dot{u}'\|_{L^\infty(A,\mathbb{R}^n)} \le c$, $|u'(a) - u(a)| \le \varepsilon$ and $|u'(b) - u(b)| \le \varepsilon$,*

$$I_M^h\big(u', (a, b)\big) \ge I_M^h\big(u, (a, b)\big) - Ch - C\varepsilon.$$

Proof For $|b-a|$ small enough, the $c|b-a|$-neighborhood of $u(A)$ lies in the domain $U \subset M$ of a single chart $\psi^{-1} : U \to V \subset \mathbb{R}^k$, say. For curves v with values in V, we define $\tilde{J} = \tilde{J}_\psi = I \circ \psi$ as before and J by

$$J(v) := I\big(N_h\psi(v)\big).$$

Note that for $w \in X_h$ with nodal points $w_i \in M$, the linear interpolation v of $\psi^{-1}(w_i)$ satisfies $J(v) = I^h(w)$ and $\|\dot{v}\|_{L^\infty(A,\mathbb{R}^k)} \le C\|\dot{w}\|_{L^\infty(A,\mathbb{R}^n)}$.

Assume first that $a, b \in \mathcal{T}_h$ and that $u'(a) = u(a)$, $u'(b) = u(b)$. We have to show that for all piecewise affine functions v' with $\|\dot{v}'\|_{L^\infty} \le c'$ for some constant c',

$$J\big(v'\big) \ge J(v) - Ch,$$

where $v = N_h \psi^{-1}(u)$.

Let w be a piecewise affine curve with values in V. Then by Taylor expansion

$$\int_a^b \big|N_h\psi(w) - \psi(w)\big|^2 \,\mathrm{d}t \le Ch^2 \left\|\frac{\mathrm{d}}{\mathrm{d}t}\psi(w)\right\|^2_{L^\infty(A,\mathbb{R}^k)} \le Ch^2\|\dot{w}\|^2_{L^\infty(A,\mathbb{R}^k)}$$

and

$$\begin{aligned}\int_a^b \left|\frac{\mathrm{d}}{\mathrm{d}t}\big(N_h\psi(w) - \psi(w)\big)\right|^2 \mathrm{d}t &\le Ch^2 \left\|\frac{\mathrm{d}^2}{\mathrm{d}t^2}\psi(w)\right\|^2_{L^\infty(A,\mathbb{R}^k)} \\ &= Ch^2\big\|\nabla^2\psi(w)(\dot{w},\dot{w})\big\|^2_{L^\infty(A,\mathbb{R}^k)} \le Ch^2\|\dot{w}\|^4_{L^\infty(A,\mathbb{R}^k)},\end{aligned}$$

where $\frac{\mathrm{d}^2}{\mathrm{d}t^2}\psi(w)$ is the absolute continuous part of the second derivative of $\psi \circ w$. Thus, if $\|\dot{w}\|_{L^\infty(A,\mathbb{R}^k)}$ is bounded,

$$\begin{aligned}\big|J(w) - \tilde{J}(w)\big| &= \big|I\big(N_h\psi(w)\big) - I\big(\psi(w)\big)\big| \\ &= \left|\int_a^b \left|\frac{\mathrm{d}}{\mathrm{d}t}N_h\psi(w)\right|^2 - V\big(N_h\psi(w)\big)\,\mathrm{d}t \right. \\ &\qquad \left. - \int_a^b \left|\frac{\mathrm{d}}{\mathrm{d}t}\psi(w)\right|^2 - V\big(\psi(w)\big)\,\mathrm{d}t\right| \\ &\le Ch.\end{aligned}$$

We also need to compare the first variations of J and $\tilde{J}$. Let φ be a piecewise affine curve such that $w(t) + r\varphi(t) \in V$ for all $t \in [a,b]$, $r \in [0,1]$. Then

$$\begin{aligned}&\frac{\mathrm{d}}{\mathrm{d}r}\bigg|_{r=0} \tilde{J}(w + r\varphi) \\ &\quad = \int_a^b m\left(\frac{\mathrm{d}}{\mathrm{d}t}\psi(w)\right) \cdot \left(\frac{\mathrm{d}}{\mathrm{d}t}\big(\nabla\psi(w)\varphi\big)\right) - \nabla V\big(\psi(w)\big)\nabla\psi(w)\varphi\,\mathrm{d}t\end{aligned}$$

and

$$\begin{aligned}&\frac{\mathrm{d}}{\mathrm{d}r}\bigg|_{r=0} J(w + r\varphi) \\ &\quad = \int_a^b m\left(\frac{\mathrm{d}}{\mathrm{d}t}N_h\psi(w)\right) \cdot \left(\frac{\mathrm{d}}{\mathrm{d}t}N_h\big[\nabla\psi(w)\varphi\big]\right) - \nabla V\big(N_h\psi(w)\big)N_h\big[\nabla\psi(w)\varphi\big]\,\mathrm{d}t.\end{aligned}$$

Now similar estimates as above show that

$$\left| \frac{\mathrm{d}}{\mathrm{d}r}\right|_{r=0} \tilde{J}(w+r\varphi) - \frac{\mathrm{d}}{\mathrm{d}r}\Big|_{r=0} J(w+r\varphi) \Big| \le Ch.$$

Now choose φ such that $v+\varphi = v'$. Since v is a stationary point of J, the above estimates show that

$$\begin{aligned} J(v+\varphi) - J(v) &\ge \tilde{J}(v+\varphi) - \tilde{J}(v) - Ch \\ &= \frac{\mathrm{d}}{\mathrm{d}r}\Big|_{r=0} \tilde{J}(v+r\varphi) + \int_0^1 (1-r)\frac{\mathrm{d}^2}{\mathrm{d}r^2}\tilde{J}(v+r\varphi)\,\mathrm{d}r - Ch \\ &\ge \frac{\mathrm{d}}{\mathrm{d}r}\Big|_{r=0} J(v+r\varphi) + \int_0^1 (1-r)\frac{\mathrm{d}^2}{\mathrm{d}r^2}\tilde{J}(v+r\varphi)\,\mathrm{d}r - Ch \\ &= \int_0^1 (1-r)\frac{\mathrm{d}^2}{\mathrm{d}r^2}\tilde{J}(v+r\varphi)\,\mathrm{d}r - Ch. \end{aligned}$$

But

$$\begin{aligned} \frac{\mathrm{d}^2}{\mathrm{d}r^2}\tilde{J}(v+r\varphi) = \int_a^b & m\big|\nabla\psi(v+r\varphi)\cdot\dot{\varphi} + \nabla^2\psi(v+r\varphi)(\dot{v}+r\dot{\varphi},\varphi)\big|^2 \\ & + m\big(\nabla\psi(v+r\varphi)\cdot(\dot{v}+r\dot{\varphi})\big) \\ & \cdot\big(\nabla^3\psi(v+r\varphi)(\dot{v}+r\dot{\varphi},\varphi,\varphi) + 2\nabla^2\psi(v+r\varphi)(\dot{\varphi},\varphi)\big) \\ & - \nabla^2 V\big(\psi(v+r\varphi)\big)\big(\nabla\psi(v+r\varphi)\cdot\varphi, \nabla\psi(v+r\varphi)\cdot\varphi\big) \\ & - \nabla V\big(\psi(v+r\varphi)\big)\nabla^2\psi(v+r\varphi)(\varphi,\varphi)\,\mathrm{d}t \\ \ge \int_a^b & C'|\dot{\varphi}|^2 - C\big(|\varphi|+|\varphi|^2\big)|\dot{\varphi}|^2 - C\big(|\varphi|+|\varphi|^2\big)|\dot{\varphi}| - C|\varphi|^2\,\mathrm{d}t \end{aligned}$$

for some $C' > 0$.

So, if $|b-a|$, and hence $\|\varphi\|_{L^\infty}$ is small enough, we have

$$\begin{aligned} \frac{\mathrm{d}^2}{\mathrm{d}r^2}\tilde{J}(v+r\varphi) &\ge \int_a^b \frac{C'}{2}|\dot{\varphi}|^2 - C|\varphi|\cdot|\dot{\varphi}| - C|\varphi|^2\,\mathrm{d}t \ge \int_a^b \frac{C'}{4}|\dot{\varphi}|^2 - C|\varphi|^2\,\mathrm{d}t \\ &\ge \int_a^b \frac{C'\pi^2}{4|b-a|^2}|\varphi|^2 - C|\varphi|^2\,\mathrm{d}t \ge 0 \end{aligned}$$

by Poincaré's inequality, which concludes the first part of the proof.

Now in the general case, we can choose a maximal interval $(a_h, b_h) \subset (a, b)$ compatible with $\mathcal{T}_h$ and an affine function l_h such that $v'(a_h) + l_h(a_h) = v(a_h)$, $v'(b_h) + l_h(b_h) = v(b_h)$. Then $\|l_h\|_{W^{1,\infty}} \le C(\varepsilon + h)$ and so $|J(v' + l_h, (a_h, b_h)) - J(v', (a_h, b_h))| \le C(\varepsilon + h)$. By our bounds on v' and v, we also have

$$\big|J\big(v', (a_h, b_h)\big) - J\big(v', (a, b)\big)\big| + \big|J\big(v, (a_h, b_h)\big) - J\big(v, (a, b)\big)\big| \le Ch.$$

So, by the first part of the proof, we obtain

$$\begin{aligned} J\big(v', (a,b)\big) &\geq J\big(v', (a_h,b_h)\big) - Ch \geq J\big(v'+l_h, (a_h,b_h)\big) - Ch - C\varepsilon \\ &\geq J\big(v, (a_h,b_h)\big) - Ch - C\varepsilon \geq J\big(v,(a,b)\big) - Ch - C\varepsilon. \end{aligned}$$ □

We now investigate the limiting behavior of a sequence u_h of stationary points of I_M^h. Assuming that (u_h) is locally bounded, we clearly have that for a subsequence $u_h \overset{*}{\rightharpoonup} u$ in L^∞_{loc} for some u. However, as seen in Example 3.1, in general, the limit u does not even have to be continuous or to lie on the constraint manifold M. Note that in all these cases the discrete action functional $I_M^h(u_h, A)$ blows up as $h \to 0$ even over bounded intervals A. To avoid such pathological limiting behavior, we will assume that the action $I_M^h(u_h, A)$ remains bounded for each $A \in \mathcal{E}$. Then we can, in fact, prove a positive result.

Lemma 3.3 *Let $V \in C^1$ such that $|V(s)| \leq C(1+|s|^2)$ and suppose that u_h is a sequence of stationary points for I_M^h such that $|u_h(0)|$ and, for each $A \in \mathcal{E}$, $I_M^h(u_h, A)$ is bounded. Then u_h is bounded in $W^{1,\infty}_{\mathrm{loc}}(\mathbb{R}, \mathbb{R}^n)$.*

Proof Let $a < b$. Using the elementary bounds

$$\|u\|_{L^2((a,b),\mathbb{R}^n)} \leq \sqrt{b-a}\big|u(a)\big| + (b-a)\|\dot{u}\|_{L^2((a,b),\mathbb{R}^n)} \quad \text{and}$$

$$\big|u(t)\big| \leq \sqrt{1+b-a}\sqrt{\big|u(a)\big|^2 + \|\dot{u}\|^2_{L^2((a,b),\mathbb{R}^n)}}, \quad t \in [a,b],$$

we see that

$$\big|u_h(t)\big| \leq 2\sqrt{\big|u_h(a)\big|^2 + \frac{1}{m} I\big(u_h, (a,b)\big) + \frac{C|b-a|}{m}}$$

for $t \in [a,b]$, whenever $b-a$ is so small that $b-a < 1$ and $8C(b-a) \leq (1-b+a)m$. An analogous inequality holds with the roles of a and b interchanged. So, inductively we see that (u_h) is bounded in $L^\infty_{\mathrm{loc}}(\mathbb{R}, \mathbb{R}^n)$, and hence by boundedness of the actions, in $H^1_{\mathrm{loc}}(\mathbb{R}, \mathbb{R}^n)$. In particular, a suitable subsequence converges to some continuous u uniformly on compact subsets of $\mathbb{R}$.

Let $A = (a,b) \in \mathcal{E}$ and cover a neighborhood of $u([a,b])$ with finitely many domains $U_1, \dots, U_N \subset M$ such that on each U_j

$$|y-x| \leq \big|P_x(y-x)\big| + C'\big|P_x(y-x)\big|^3 \quad \forall x, y \in U_j,$$

where P_x is the projection onto $T_x M$. The discrete Euler–Lagrange equations yield

$$\begin{aligned} P_{u_i}\Bigg[m\bigg(\frac{u_{i+1}-u_i}{t_{i+1}-t_i} - \frac{u_i - u_{i-1}}{t_i - t_{i-1}}\bigg) + \int_{t_i}^{t_{i+1}} \nabla V\big(u_h(t)\big)\frac{t_{i+1}-t}{t_{i+1}-t_i}\,\mathrm{d}t \\ + \int_{t_{i-1}}^{t_i} \nabla V\big(u_h(t)\big)\frac{t-t_{i-1}}{t_i - t_{i-1}}\,\mathrm{d}t \Bigg] = 0 \end{aligned}$$

and so

$$\left|P_{u_i}\left(\frac{u_{i+1}-u_i}{t_{i+1}-t_i}-\frac{u_i-u_{i-1}}{t_i-t_{i-1}}\right)\right| \le C|t_{i+1}-t_{i-1}|.$$

Now, if h is small enough, then for every i, the points u_{i-1}, u_i, and u_{i+1} lie in a single domain U_j, say. But then

$$|u_{i+1}-u_i| \le \left|P_{u_i}(u_{i+1}-u_i)\right| + C'\left|P_{u_i}(u_{i+1}-u_i)\right|^3$$

and thus

$$\begin{aligned}\left|\frac{u_{i+1}-u_i}{t_{i+1}-t_i}\right| \le& \left|\frac{u_i-u_{i-1}}{t_i-t_{i-1}}\right| + C(t_{i+1}-t_{i-1}) \\ &+ C'(t_{i+1}-t_i)^2\left(\left|\frac{u_i-u_{i-1}}{t_i-t_{i-1}}\right| + C(t_{i+1}-t_{i-1})\right)^3. \end{aligned} \tag{9}$$

Since $\|\dot{u}_h\|_{L^2(A,\mathbb{R}^n)}$ is bounded, there is a constant c, independent of h, such that $|\frac{u_{i_0}-u_{i_0-1}}{t_{i_0}-t_{i_0-1}}| \le c$ for some $i_0 = i_0(h)$ with $t_{i_0} \in A$. Set $\gamma_i := |\frac{u_i-u_{i-1}}{t_i-t_{i-1}}|$, $f_i := C(t_i - t_{i-2})$, $g_i := C'(t_i - t_{i-1})^2$. Then (9) reads

$$\gamma_{i+1} \le \gamma_i + f_{i+1} + g_{i+1}(\gamma_i + f_{i+1})^3, \quad \gamma_{i_0} \le c.$$

For $i \ge i_0$ with $t_i \le b + h$, we let

$$F_i = \sum_{j=i_0+1}^{i} f_j \quad \text{and} \quad G_i = \sum_{j=i_0+1}^{i} g_j$$

and claim that for small h,

$$\gamma_i \le c + F_i + G_i(c + F_i + 1)^3. \tag{10}$$

This can be easily seen by induction on i. The case $i = i_0$ is clear. If (10) holds for $i \ge i_0$, then

$$\begin{aligned}\gamma_{i+1} &\le \gamma_i + f_{i+1} + g_{i+1}(\gamma_i + f_{i+1})^3 \\ &\le c + F_{i+1} + G_i(c + F_i + 1)^3 + g_{i+1}\big(c + F_{i+1} + G_i(c + F_i + 1)^3\big)^3 \\ &\overset{!}{\le} c + F_{i+1} + G_{i+1}(c + F_{i+1} + 1)^3,\end{aligned}$$

and the last inequality follows if

$$g_{i+1}\big(c + F_{i+1} + G_i(c + F_i + 1)^3\big)^3 \le g_{i+1}(c + F_{i+1} + 1)^3,$$

i.e., if $G_i(c + F_i + 1)^3 \le 1$. But this is satisfied for small h because

$$F_i = C\sum_{j=i_0+1}^{i}(t_j - t_{j-2}) = C(t_i + t_{i-1} - t_{i_0} - t_{i_0-1}) \le C(b - a + 2h) \tag{11}$$

and

$$G_i = C' \sum_{j=i_0+1}^{i} (t_j - t_{j-1})^2 \le C' \max_{i_0+1 \le j \le i} (t_j - t_{j-1}) \cdot \sum_{j=i_0+1}^{i} (t_j - t_{j-1})$$
$$\le C' h(b - a + h). \tag{12}$$

Now applying (11) and (12) to (10), we see that in fact $|\dot{u}_h| \le C$ on $[t_{i_0-1}, b]$. A similar argument yields that $|\dot{u}_h|$ is bounded on $[a, t_{i_0}]$, too, which concludes the proof. □

We are now in a position to state and prove our main result on the limiting behavior of a sequence of discrete trajectories. Note, in particular, that we also have strong convergence of the velocities.

Theorem 3.4 *Assume that $V \in C^2$ satisfies $|V(s)| \le C(1 + |s|^2)$. Suppose u_h is a sequence of stationary points for I_M^h such that $|u_h(0)|$ and $I_M^h(u_h, A)$ are bounded uniformly in h for all $A \in \mathcal{E}$. Then there exists a subsequence such that $u_h \overset{*}{\rightharpoonup} u$ in $W^{1,\infty}_{\mathrm{loc}}(\mathbb{R}, \mathbb{R}^n)$, $u_h \to u$ in $W^{1,p}_{\mathrm{loc}}(\mathbb{R}, \mathbb{R}^n)$ for all $1 \le p < \infty$ and u is a stationary point of I_M.*

Proof By Lemma 3.3, (u_h) is weak*-precompact in $W^{1,\infty}_{\mathrm{loc}}(\mathbb{R}, \mathbb{R}^n)$, so it has a convergent subsequence (not relabeled) $u_h \overset{*}{\rightharpoonup} u$ (and, in particular, $u_h \to u$ uniformly on compact subsets of $\mathbb{R}$). Let $A \in \mathcal{E}$. By Proposition 2.1

$$\infty > \liminf_{h\to 0} I_M^h(u_h, A) \ge I_M(u, A).$$

In order to prove that u is a stationary point of I_M, we will show that for every $t^* \in \mathbb{R}$, u is, in fact, a $W^{1,\infty}$-local minimizer of $I_M(\cdot, A)$ with respect to its own boundary values on A whenever $A = (a, b)$ is a sufficiently small neighborhood of t^*. By density, this implies that u is a stationary point of $I_M(\cdot, A)$.

Let $\tilde{u}$ be a curve with $\tilde{u}(a) = u(a)$, $\tilde{u}(b) = u(b)$, and $\|\tilde{u} - u\|_{W^{1,\infty}(A)}$ sufficiently small. Choose a recovery sequence $\tilde{u}_h$ for $\tilde{u}$ as in the proof of Proposition 2.1 such that $\tilde{u}_h \to \tilde{u}$ uniformly on $\overline{A}$. Note that by Remark 2.4, we may assume that $\|\dot{\tilde{u}}_h\|_{L^\infty(A,\mathbb{R}^n)}$ is bounded independently of h. Since $u_h \to u$ and $\tilde{u}_h \to \tilde{u}$ uniformly on $\overline{A}$,

$$\big|u_h(a) - \tilde{u}_h(a)\big|, \big|u_h(b) - \tilde{u}_h(b)\big| \le \varepsilon(h)$$

for some $\varepsilon(h) \to 0$ as $h \to 0$. But then, if $|b - a|$ is sufficiently small, we obtain from Lemma 3.2

$$I_M(\tilde{u}, A) = \lim_{h\to 0} I_M^h(\tilde{u}_h, A) \ge \limsup_{h\to 0} \big(I_M^h(u_h, A) - Ch - C\varepsilon(h)\big)$$
$$\ge \liminf_{h\to 0} \big(I_M^h(u_h, A) - Ch - C\varepsilon(h)\big) \ge I_M(u, A),$$

which proves the local minimality of u.

Setting $\tilde{u} = u$, the above argument shows that $\lim_h I_M^h(u_h, A) = I_M(u, A)$, hence

$$\begin{aligned} I_M(u, A) &= \lim_{h\to 0} \int_A \frac{m}{2}\big(|\dot{u}|^2 + 2\dot{u}\cdot(\dot{u}_h - \dot{u}) + |\dot{u}_h - \dot{u}|^2\big) - V(u_h)\,\mathrm{d}t \\ &= I_M(u, A) + 0 + \lim_{h\to 0} \|\dot{u}_h - \dot{u}\|_{L^2}^2. \end{aligned}$$

But then $\dot{u}_h \to \dot{u}$ in $L^2(A, \mathbb{R}^n)$ and so, $\dot{u}_h$ being bounded in $L^\infty(A, \mathbb{R}^n)$, $\dot{u}_h \to \dot{u}$ in $L^p(A, \mathbb{R}^n)$ for all $1 \le p < \infty$. □

4 Numerical Examples

In this section, the statements of Theorem 3.4 will be illustrated by means of numerical examples. The discrete path $\{u_i\}_{i=0}^N \subset X_h$ of a mechanical system is determined as a stationary point of the constrained discrete action functional I_M^h. According to (2), the stationary points are characterized by

$$\nabla_2 L_d(u_{i-1}, u_i) + \nabla_1 L_d(u_i, u_{i+1}) \perp T_{u_i} M. \tag{13}$$

In the sequel, it is assumed that the triangulation $\mathcal{T}_h$ is equispaced, and the constant timestep is denoted by $h = t_{i+1} - t_i$. There are several methods for the implementation of an algorithm that finds discrete paths matching condition (13). Assume that the constraint manifold is specified by the holonomic constraint functions $g(u) = 0 \in \mathbb{R}^m$, $m = n - k$, in particular if zero is a regular value of the constraints, then $= g^{-1}(0) = M$ is a k-dimensional submanifold of $\mathbb{R}^n$. Let $G(u) = \nabla g(u)$ denote the Jacobian of the constraints, then $\mathbb{R}^n = T_{u_i} M \oplus (T_{u_i} M)^\perp = \mathrm{null}(G(u)) \oplus \mathrm{range}(G^\mathrm{T}(u))$. Thus, it is possible to use the Lagrange multiplier method to find appropriate paths. The corresponding discrete Euler–Lagrange equations then read

$$\begin{aligned} &\nabla_2 L_d(u_{i-1}, u_i) + \nabla_1 L_d(u_i, u_{i+1}) + hG^\mathrm{T}(u_i)\cdot\lambda_i = 0, \\ &g(u_{i+1}) = 0. \end{aligned} \tag{14}$$

Since the vector of Lagrange multipliers $\lambda_i \in \mathbb{R}^m$ has to be determined as a variable by the algorithm, the system has to be augmented by the constraint equations $(14)_2$ and is then $n+m$-dimensional. Being a two-step method, the algorithm (14) is not self-starting. From given initial configuration u_0 and initial conjugate momentum p_0, the configuration u_1 (and the Lagrange multiplier λ_0) can be calculated via the constrained discrete Legendre transform

$$\begin{aligned} &p_0 + \nabla_1 L_d(u_0, u_1) + \frac{h}{2} G^\mathrm{T}(u_0)\cdot\lambda_0 = 0, \\ &g(u_1) = 0 \end{aligned} \tag{15}$$

in a way that is consistent with the constrained dynamics.

Alternatively, the discrete null space method can be used. Besides leading to a k-dimensional system of equations, it also bears the advantage of removing the conditioning problems associated with the use of the Lagrange multiplier method. The discrete null space method relies on a transformation of (14) in two steps. First of all, $(14)_1$ is premultiplied with the transpose of the null space matrix $P(u):\mathbb{R}^k \to T_u M$, having the property $\text{range}(P(u)) = \text{null}(G(u))$. Secondly, $(14)_2$ is redundantized introducing the nodal reparametrization $u_{i+1} = F_d(w_{i+1}, u_i) \in M$. The resulting time-stepping scheme

$$P^{\mathrm{T}}(u_i)\cdot\big[\nabla_2 L_d(u_{i-1}, u_i) + \nabla_1 L_d\big(u_i, F_d(w_{i+1}, u_i)\big)\big] = 0 \tag{16}$$

has to be solved for w_{i+1} whereupon the redundant configuration variable can be updated using the nodal reparametrization F_d. A consistent configuration u_1 can be found from the equation

$$P^{\mathrm{T}}(u_0)\cdot\big[p_0 + \nabla_1 L_d\big(u_0, F_d(w_1, u_0)\big)\big] = 0. \tag{17}$$

It is important to note that (16) and (14) are equivalent and both schemes are variational. See Leyendecker et al. (2008b) for a detailed investigation of the discrete null space method for the variational integration of constrained dynamical systems. Previous works on the discrete null space method in conjunction with an energy-momentum conserving scheme are Betsch (2005), Betsch and Leyendecker (2006), Leyendecker et al. (2008a).

In the following examples, the midpoint rule has been used to approximate the integral of the potential energy over one timestep in the discrete Lagrangian given in (1), thus the discrete Lagrangian reads

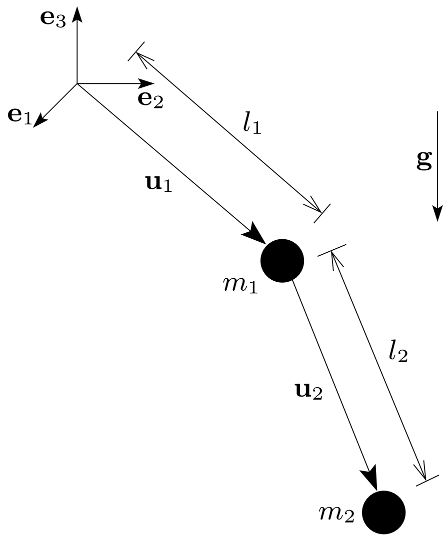
$$L_d(u_i, u_{i+1}) = \frac{m}{2}\frac{(u_{i+1}-u_i)^2}{h} - hV\left(\frac{u_{i+1}+u_i}{2}\right). \tag{18}$$

Using the variational integrator with Lagrange multipliers for constraint enforcement (14), the resulting implicit scheme is similar to the SHAKE algorithm. However, SHAKE is based on a trapezoidal rule (instead of the midpoint rule) yielding the evaluation of the potential gradient at one given configuration only, which can have unfavorable consequences on the robustness of simulations, in particular in the context of stiff nonlinear elasticity problems as beam dynamics. In contrast to the so far mentioned algorithms that enforce configuration constraints only, the velocity Verlet integrator RATTLE enforces the temporally differentiated form of the constraints on velocity level as well; see Andersen (1983).

4.1 Double Spherical Pendulum

The motion of a double spherical pendulum in three dimensional space has been simulated using the discrete null space method, i.e., (16) is solved to determine the stationary points of the constrained discrete action functional I_M^h. The double spherical pendulum in Fig. 2 is suspended at the origin of the inertial frame $\{e_I\}$. Massless rigid rods of lengths $l_1, l_2 \in \mathbb{R}$ connect the masses $m_1, m_2 \in \mathbb{R}$ to each other and

Fig. 2 Double spherical pendulum



to the origin, respectively. The gravitational acceleration with value g points in the negative e_3-direction. The configuration variable $u \in \mathbb{R}^6$ is composed of the placement in space $u^1 \in \mathbb{R}^3$ of the first mass and the placement $u^2 \in \mathbb{R}^3$ of the second mass $m^2 \in \mathbb{R}$ relative to the first one. The constraints represent the constancy of the lengths of the rigid rods. They restrict possible configurations to the constraint manifold $M = S^2_{l_1} \times S^2_{l_2}$ consisting of two spheres: one about the origin with radius l_1 and one about the first mass with radius l_2.

In the simulation of the double spherical pendulum's motion, the following parameters have been used. The masses are $m_1 = 10$ and $m_2 = 5$ and the rigid rods have the lengths $l_1 = 1$ and $l_2 = 1.5$. The gravitational acceleration is given by $g = 9.81$. The initial positions of the point masses are $u^1(0) = l_1 e_1$ and $u^2(0) = l_2 e_1$ and initial velocities are given by $\dot{u}^1(0) = e_2$ and $\dot{u}^2(0) = e_3$. These initial velocities are consistent with the constraints, i.e., they lie in the tangent space $T_{u_0}M$. The considered motion takes place in the time interval $\bar{A} = [0, 10]$.

Snapshots of the motion of the double spherical pendulum are shown in Fig. 3 on the left. The diagram on the right confirms the algorithmic conservation of the component L_3 of the angular momentum corresponding to the gravitational direction. Furthermore, the good energy behavior of the variational integrator is revealed by the evolution of the total energy which appears to be conserved. In fact, the total energy oscillates around the correct value with amplitudes in the range of 10^{-3}.

Table 1 shows that for the considered example, velocities are indeed bounded independently of h.

A reference solution $u_{h_{\text{ref}}}$ has been calculated using the time step $h_{\text{ref}} = 10^{-5}$. The convergence statement of Theorem 3.4 is illustrated by Fig. 4. It reveals the well-known fact that variational integrators based on a discrete Lagrangian given in (18) are second order convergent also holds for constrained problems.

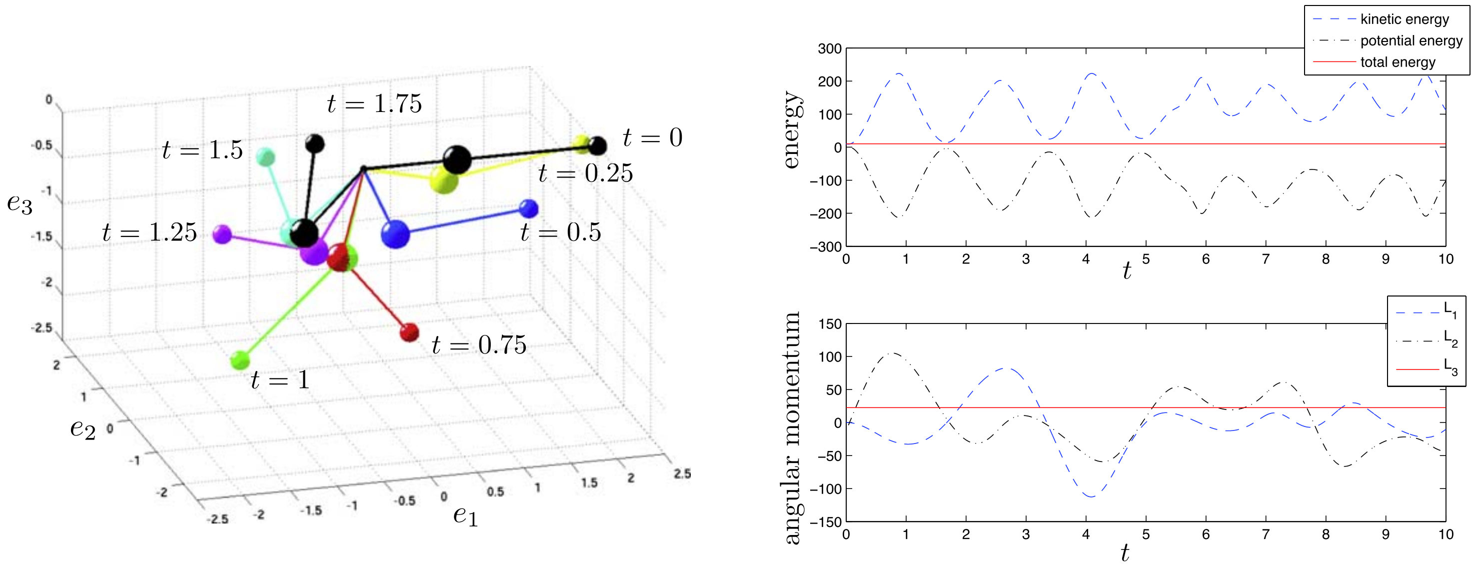


Fig. 3 Double spherical pendulum: snapshots of the motion and energy and components of angular momentum $L = L_i e_i$ $(h = 10^{-5})$

Table 1 Double spherical pendulum: boundedness of velocity

h	$\max_{t_j \in \{t_i\}_{i=0}^N} \lVert \dot{u}_h(t_j) \rVert$
10^{-1}	14.437579674951671
10^{-2}	11.992045771547241
10^{-3}	11.979460050283279
10^{-4}	11.979355188823591
10^{-5}	11.979353929835233

Fig. 4 Double spherical pendulum: second order convergence to reference solution for $h \in [10^{-1}, \ldots, 10^{-4}]$

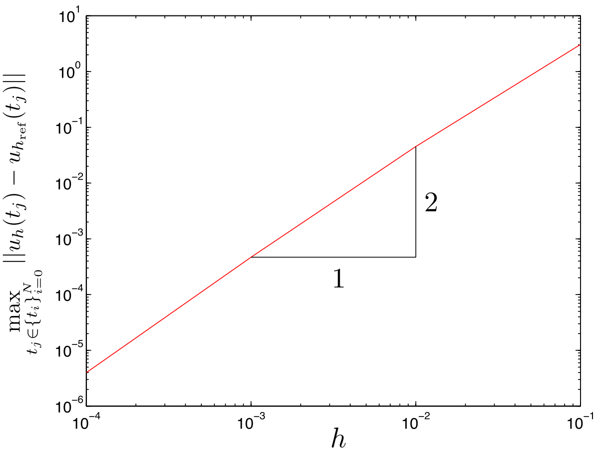
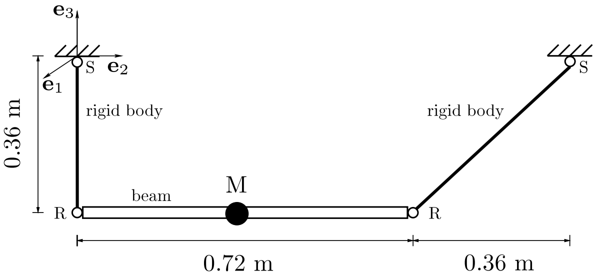


Fig. 5 Three-bar swing comprising a flexible beam with midspan mass hinged (by revolute joints R) to rigid bodies fixed in space (by spherical joints S)



4.2 Three-Bar Swing

The second example deals with the swing shown in Fig. 5. It consists of an elastic beam hinged at its ends to rigid bodies by revolute joints. The rigid bodies are fixed in space by spherical joints. An additional point mass is concentrated at the beam's midpoint. This example has been investigated previously in Bauchau et al. (1995) using an energy-conserving scheme and the generalized-α method. In Ibrahimbegović and Mamouri (2002) results from an energy-conserving and an energy-decaying scheme are presented. The purpose of the presentation of this example here is twofold. First of all, to the author's knowledge, the (orthonormality constrained) director based formulation of geometrically exact beams introduced in Betsch and Steinmann (2002)

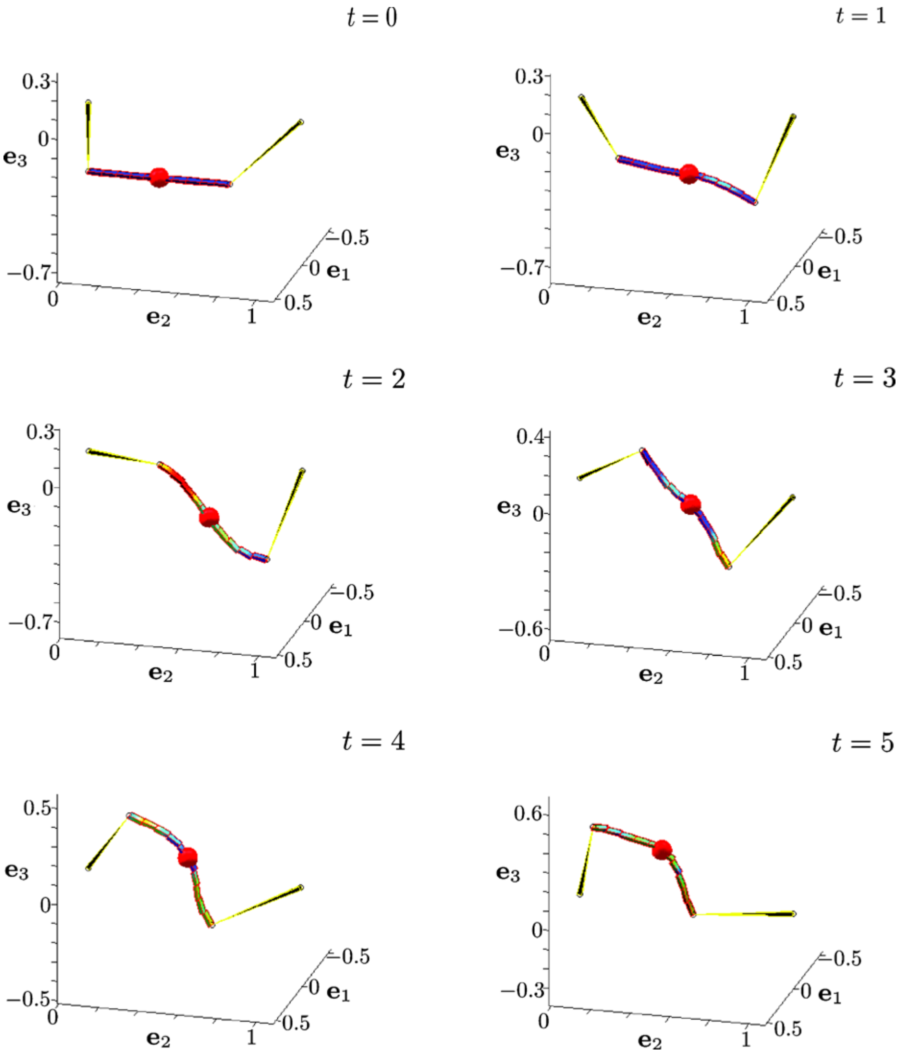


Fig. 6 Three-bar swing: snapshots of the motion and deformation ($h = 5 \cdot 10^{-4}$)

and Romero and Armero (2002) is used here in the framework of a variational time-stepping scheme for the first time. This constrained formulation is beneficial to the description of multibody dynamics consisting of rigid and elastic components, see Betsch and Leyendecker (2006), Leyendecker et al. (2008a). Even though an energy-momentum conserving time integration scheme is used there, the expressions for the mass matrix corresponding to the spatial discretization of the beam by finite elements, constraints, constraint Jacobian, and especially for the null space matrix can be transformed according to the requirements of the variational schemes (14) and (16) in a straightforward way. Different methods for the enforcement of the orthonormality constraints for the director triads have been compared in Leyendecker

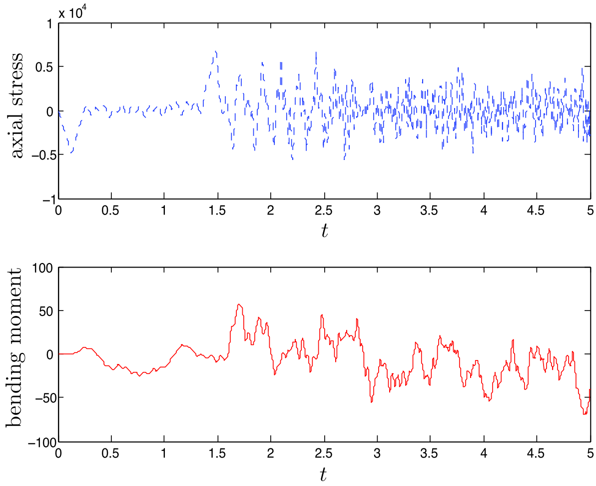


Fig. 7 Three-bar swing: axial force and bending moment with respect to an axis parallel to e_1 in the element to the right of the concentrated mass ($h = 5 \cdot 10^{-4}$)

Table 2 Three-bar swing: boundedness of velocity

h	$\max_{t_j \in \{t_i\}_{i=0}^{N}} \|\dot{u}_h(t_j)\|$
10^{-2}	21.264358970618083
$5 \cdot 10^{-3}$	21.502895414634740
10^{-3}	29.211974341839312
$5 \cdot 10^{-4}$	29.414409469970764
10^{-4}	28.986258982641335
$5 \cdot 10^{-5}$	28.970130936433204
10^{-5}	28.964966706946093

et al. (2004, 2006) for rigid bodies and elastic beams, respectively. Secondly, the example illustrates that Theorem 3.4 is applicable in the presence of highly nonlinear elastic behavior where, despite the fact that the elastic energy is given by a fourth order term in the configuration variables, V satisfies the quadratic energy bound $|V(s)| \leq C(1 + |s|^2)$ in a neighborhood of the constraint manifold M.

Both rigid bodies' mass is 0.01 kg and they have the shape of pyramids with a square ground face of edge length 0.02 m and the height of 0.36 m and $0.36\sqrt{2}$ m, respectively. A concentrated mass of $M = 5$ kg is rigidly connected at the midspan node of the beam, which is discretized by 10 linear finite beam elements. The semi-discrete beam's response to loading is based on hyperelastic material behavior with stiffness parameters GA $= 175\,480.7692$ N, EA $= 547\,500$ N, $\mathrm{EI}_1 = 114.0625\ \mathrm{N\,m}^2$, $\mathrm{EI}_2 = 10.2656\ \mathrm{N\,m}^2$, and GJ $= 13.7401\ \mathrm{N\,m}^2$. The sectional mass properties are $A_\rho = 7500\ \mathrm{kg\,m}^{-1}$, $M_\rho^1 = 1.5625$ kg m and $M_\rho^2 = 0.1406$ kg m. The cross section is oriented such that the smaller of the two bending stiffnesses is with respect to the axis parallel to e_1. (Note that the numbering of the bending stiffnesses corresponds to the numbering of the nodal director triads which differ from the inertial frame.) The loading is a triangular pulse in e_2-direction which is applied at the midspan mass.

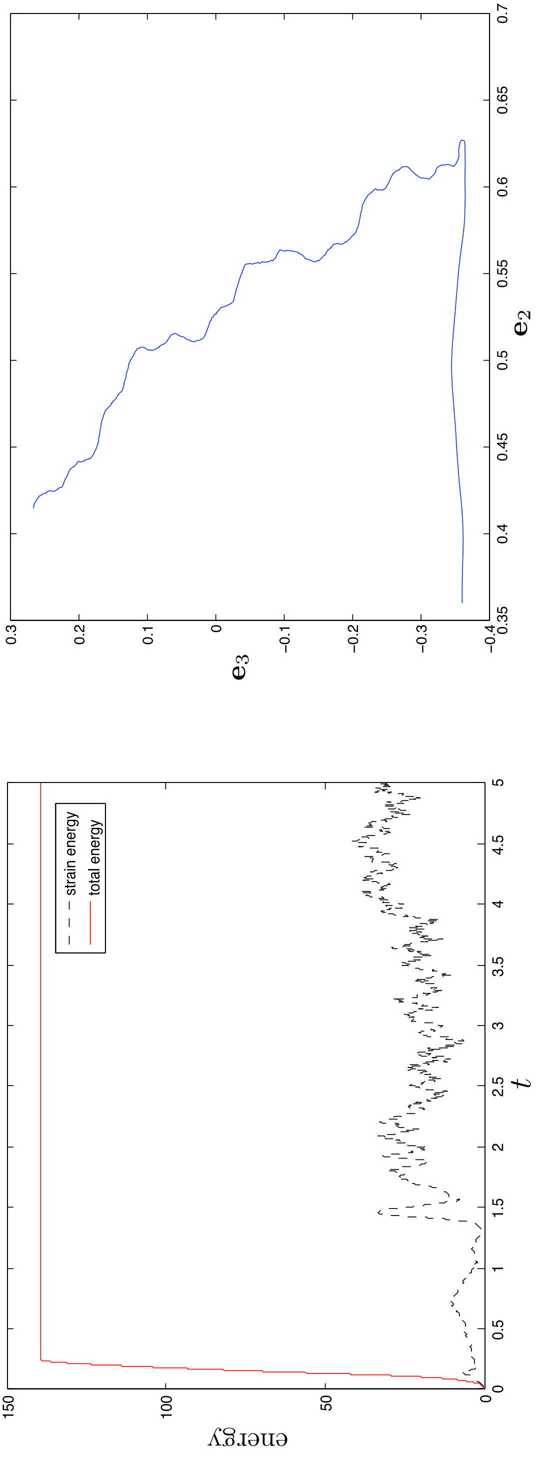


Fig. 8 Three-bar swing: energy and orbit of the concentrated mass in the (e_2, e_3)-plane ($h = 5 \cdot 10^{-4}$)

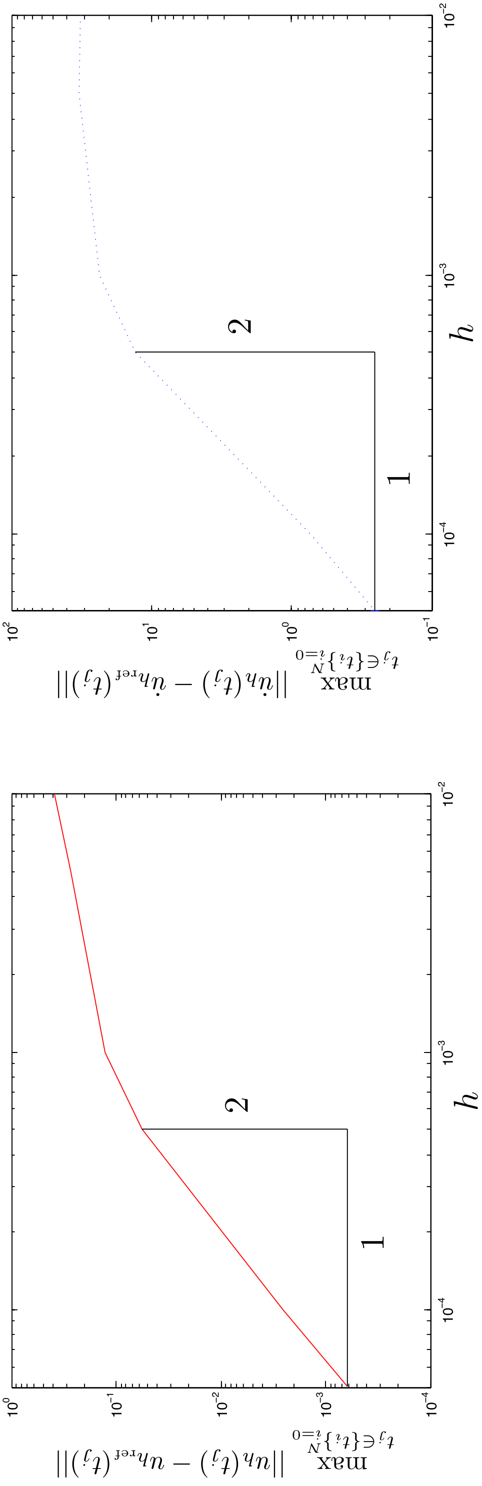


Fig. 9 Three-bar swing: convergence of configuration and velocity to reference solution for $h \in [10^{-2}, \ldots, 5 \cdot 10^{-5}]$

It starts with 0 N at $t = 0$ s, peaks with 10 000 N at $t = 0.125$ s, and ends with 0 N at $t = 0.25$ s.

Snapshots of the motion and deformation are depicted in Fig. 6. The elements' colors represent a linear interpolation of the sum of the resulting axial and shear forces norm and the resulting bending and torsional moments norm. Thereby blue (dark grey) represents zero and red (brighter grey) represents 4000. According to the loading in axial direction of the beam, the axial forces dominate the stress resultants thus the colors are representing the axial force distribution qualitatively. The evolution of the axial force and bending moments with respect to the axis e_1 in the element to the right of the concentrated mass can be observed from Fig. 7. The left-hand diagram in Fig. 8 reveals again the good energy behavior of the variational scheme. After the vanishing of the external load, energy fluctuations are of the order of 10^{-4} for the time step $h = 5 \cdot 10^{-4}$. The diagram on the right-hand side illustrates the orbit of the concentrated mass in the (e_2, e_3)-plane. One can see clearly how the beam's deformation superposes the overall rigid motion of the multibody system. Boundedness of the velocities can be observed from Table 2.

References

Andersen, H.C.: RATTLE: A velocity version of the SHAKE algorithm for molecular dynamics calculations. J. Comput. Phys. **52**(1), 24–34 (1983)

Bauchau, O.A., Damilano, G., Theron, N.J.: Numerical integration of non-linear elastic multi-body systems. Int. J. Numer. Methods Eng. **38**, 2727–2751 (1995)

Belytschko, T., Mullen, R.: Mesh partitions of explicit–implicit time integrators. In: Bathe, K.-J., Oden, J.T., Wunderlich, W. (eds.) Formulations and Computational Algorithms in Finite Element Analysis, pp. 673–690. MIT Press, Cambridge (1976)

Belytschko, T.: Partitioned and adaptive algorithms for explicit time integration. In: Wunderlich, W. (ed.) Nonlinear Finite Element Analysis in Structural Mechanics, pp. 572–584. Springer, New York (1981)

Betsch, P.: The discrete null space method for the energy consistent integration of constrained mechanical systems. Part I: Holonomic constraints. Comput. Methods Appl. Mech. Eng. **194**(50–52), 5159–5190 (2005)

Betsch, P., Leyendecker, S.: The discrete null space method for the energy consistent integration of constrained mechanical systems. Part II: Multibody dynamics. Int. J. Numer. Methods Eng. **67**(4), 499–552 (2006)

Betsch, P., Steinmann, P.: Frame-indifferent beam finite elements based upon the geometrically exact beam theory. Int. J. Numer. Methods Eng. **54**, 1775–1788 (2002)

Dal Maso, G.: An Introduction to Γ-Convergence. Birkhäuser, Boston (1993)

Hairer, E., Lubich, C., Wanner, G.: Geometric Numerical Integration: Structure-Preserving Algorithms for Ordinary Differential Equations, 2nd edn. Springer, Berlin (2006)

Hughes, T.J.R.: The Finite Element Method: Linear Static and Dynamic Finite Element Analysis. Prentice-Hall, Englewood Cliffs (1987)

Ibrahimbegović, A., Mamouri, S.: Energy conserving/decaying implicit time-stepping scheme for nonlinear dynamics of three-dimensional beams undergoing finite rotations. Comput. Methods Appl. Mech. Eng. 4241–4258 (2002)

Jay, L.O.: Symplectic partitioned Runge–Kutta methods for constrained Hamiltonian systems. SIAM J. Numer. Anal. **33**(1), 368–387 (1996)

Jay, L.O.: Structure preservation for constrained dynamics with super partitioned additive Runge–Kutta methods. SIAM J. Sci. Comput. **20**(2), 416–446 (1998)

Leimkuhler, B., Reich, S.: Symplectic integration of constrained Hamiltonian systems. Math. Comput. **63**, 589–605 (1994)

Leyendecker, S., Betsch, P., Steinmann, P.: Energy-conserving integration of constrained Hamiltonian systems—a comparison of approaches. Comput. Mech. **33**, 174–185 (2004)

Leyendecker, S., Betsch, P., Steinmann, P.: Objective energy–momentum conserving integration for the constrained dynamics of geometrically exact beams. Comput. Methods Appl. Mech. Eng. **195**, 2313–2333 (2006)

Leyendecker, S., Betsch, P., Steinmann, P.: The discrete null space method for the energy consistent integration of constrained mechanical systems. Part III: Flexible multibody dynamics. Multibody Syst. Dyn. **19**, 45–72 (2008a)

Leyendecker, S., Marsden, J.E., Ortiz, M.: Variational integrators for constrained dynamical systems. ZAMM **88**, 677–708 (2008b)

Maggi, F., Morini, M.: A Γ-convergence result for variational integrators of Lagrangians with quadratic growth. ESAIM Control Optim. Calc. Var. **10**, 656–665 (2004)

Marsden, J.E., West, M.: Discrete mechanics and variational integrators. Acta Numer. **10**, 357–514 (2001)

McLachlan, R.I., Perlmutter, M.: Integrators for nonholonomic mechanical systems. J. Nonlinear Sci. **16**, 283–328 (2006)

Müller, S., Ortiz, M.: On the gamma-convergence of discrete dynamics and variational integrators. J. Nonlinear Sci. **14**(4), 153–212 (2004)

Reich, S.: Symplectic integrators for systems of rigid bodies. Fields Inst. Commun. **10**, 181–191 (1996)

Romero, I., Armero, F.: An objective finite element approximation of the kinematics of geometrically exact rods and its use in the formulation of an energy–momentum scheme in dynamics. Int. J. Numer. Methods Eng. **54**, 1683–1716 (2002)