BACHELORARBEIT

Structure-Preserving Time Integration of Mechanical Systems on Lie Groups

ausgeführt am

Institut für
Analysis und Scientific Computing
TU Wien

unter der Anleitung von

Prof. Joachim Schöberl

durch

Linus Knoll
Matrikelnummer: 12207062
Donauwörther Straße 27a/2
2380, Perchtoldsdorf

Wien, am 06.07.2025

Eidesstattliche Erklärung

Ich erkläre an Eides statt, dass ich die vorliegende Bachelorarbeit selbstständig und ohne fremde Hilfe verfasst, andere als die angegebenen Quellen und Hilfsmittel nicht benutzt bzw. die wörtlich oder sinngemäß entnommenen Stellen als solche kenntlich gemacht habe. Hilfsmittel der künstlichen Intelligenz wurden nur verwendet, um sprachliche Formulierungen und Zeichensetzung Korrektur zu lesen und gegebenenfalls zu verbessern.

Wien, am 06.07.2025

Linus Knoll

Contents

1 Introduction

In this thesis, we aim to construct and analyze a variational time discretization scheme based on the variational principle of Hamilton–Pontryagin. The principal motivation for this work stems from the need for numerical integrators that preserve the geometric properties of mechanical systems, particularly symplecticity, which is crucial for the long-time behavior of simulations. The Hamilton–Pontryagin principle provides a unifying framework that encapsulates both the Hamiltonian and Lagrangian formulations of classical mechanics, with its strength lying in the natural incorporation of constraints through the use of variational calculus on the Pontryagin bundle.

We first recapitulate the theoretical foundation of Hamilton–Pontryagin mechanics and show its equivalence with the classical Euler–Lagrange formalism. Building upon this, we develop a variational discontinuous Petrov–Galerkin (DPG) method for time discretization. The method is constructed on vector spaces and extended to rigid body dynamics, where the configuration space is a Lie group. By reformulating the Hamilton–Pontryagin principle to allow broken test spaces, we derive a well-posed variational problem with discontinuous test functions, enabling the construction of optimal test spaces and providing rigorous inf-sup stability. Furthermore, we present a second-order integrator arising from a piecewise linear approximation and prove that the resulting discrete flow map is symplectic on vector spaces. The construction is then generalized to constrained mechanical systems such as the motion of a rigid body on the Lie group $\mathbb{R}^3 \times \mathrm{SO}(3)$, where we develop a compatible numerical scheme that respects the manifold structure of the configuration space.

The core contribution of this thesis lies in the formulation and rigorous analysis of a discontinuous Petrov–Galerkin-based symplectic time discretization method for HP mechanics. We derive well-posedness conditions using a Trial-to-Test operator and demonstrate the accuracy and structure-preserving properties of the method. Numerical experiments on the heavy top benchmark problem confirm the theoretical results and show promising performance of the integrator.

2 Hamiltonian-Pontryagin Mechanics

We consider a mechanical system with d degrees of freedom which can be described by generalized coordinates $q = (q_1, q_2, ..., q_d) \in \mathcal{Q}$ with $\mathcal{Q}$ being a real valued vector space and $T\mathcal{Q}$ and $T^*\mathcal{Q}$ its tangent and cotangent bundles.

We define a trajectory of the system over the time $[0, T]$ as a map

$$\mathbf{q}(\cdot) : [0, T] \to \mathcal{Q}. \tag{2.1}$$

The trajectory is indicated by $q_i(t), i = 1, 2, ..., d$, and in the case that $\mathcal{Q}$ is replaced by a manifold, we can define k maps $x_j(q_1, q_2, ..., q_d), j = 1, 2, ..., k$, where x_j gives us the j-th Cartesian coordinate of the system.

Let $L : T\mathcal{Q} \to \mathbb{R}$ denote the systems Lagrangian usually taking the form

$$L(q, \dot{q}) = T(\dot{q}) - V(q), \tag{2.2}$$

where T is the kinetic energy and V the potential energy. In this thesis we only consider nondegenerate Lagrangians of the mentioned form as the focus lies on rigid body motion. However Hamilton-Pontryagin (HP) mechanics can also be applied to degenerate ones.

The motion of the system is determined by the *principle of stationary action*, also known as *Hamilton's principle*:

$$\delta \int_0^T L(q(t), \dot{q}(t))\, dt = 0, \tag{2.3}$$

Given the actual path q we consider variations of the form $q^\varepsilon(t) = q(t) + \varepsilon\delta q(t)$, with $\delta q(0) = \delta q(T) = 0$. The first variation of the action is then given by

$$\begin{aligned}
\frac{d}{d\varepsilon} S[q^\varepsilon]\bigg|_{\varepsilon=0} &= \frac{d}{d\varepsilon} \int_0^T L(q^\varepsilon, \dot{q}^\varepsilon)\, dt \bigg|_{\varepsilon=0} \\
&= \int_0^T \left(\frac{\partial L}{\partial q}(q, \dot{q}) \cdot \delta q + \frac{\partial L}{\partial \dot{q}}(q, \dot{q}) \cdot \delta \dot{q} \right) dt \\
&= \int_0^T \left[\frac{\partial L}{\partial q}(q, \dot{q}) - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}}(q, \dot{q}) \right) \right] \cdot \delta q \, dt \\
&\quad + \frac{\partial L}{\partial \dot{q}}(q, \dot{q}) \cdot \delta q \bigg|_0^T .
\end{aligned}$$

Since $\delta q(0) = \delta q(T) = 0$, the boundary term vanishes. This leads to the Euler–Lagrange equations

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}} \right) - \frac{\partial L}{\partial q} = 0. \tag{2.4}$$

Theorem 2.1. ***(Principle of Hamilton)***
Let L be a smooth Lagrangian defined on $T\mathcal{Q}$ and $q : [0,T] \to \mathcal{Q}$ be a smooth curve satisfying the boundary conditions $q(0) = q_0, q(T) = q_T$ for $q_0, q_T \in \mathcal{Q}$. Then q satisfies the Euler-Lagrange Equations (2.4) if and only if it is a stationary point of the action functional

$$S(q) = \int_0^T L(q(t), \dot{q}(t))\ dt \tag{2.5}$$

By defining the velocity v as an independent variable and demanding the constraint $\dot{q} = v$, we can see that Hamilton's principle is equivalent to extremizing

$$S(q) = \int_0^T L(q(t), v(t))\ dt. \tag{2.6}$$

This can be done by introducing a lagrange multiplier $p \in T^*\mathcal{Q}$, resulting in the **Hamilton-Potryagin Principle**

$$S(q, \dot{q}) = \int_0^T L(q(t), v(t)) + \langle p(t), \dot{q}(t) - v(t) \rangle\ dt \tag{2.7}$$

Which at first seems arbitrary, unifies Hamilton's and Lagrange's viewpoints, accounts for the Legendre Transform and includes the kinematic constraint. To show this, we start with introducing the Hamilton-Pontryagin action integral.

Definition 2.2. ***(Hamilton-Pontryagin action integral)***
The Pontryagin bundle is defined as the Whitney sum $T\mathcal{Q} \oplus T^*\mathcal{Q}$. Given an interval $[0,T]$ and two points q_0, q_T the HP path space is defined as

$$\mathcal{C}(q_0, q_T, [0,T]) = \{(q, v, p) : [0,T] \to T\mathcal{Q} \oplus T^*\mathcal{Q} \mid z = (q, v, p) \in C^2([0,T]), q(0) = q_0, q(T) = q_T\}$$

and the HP action integral $\mathcal{S} : \mathcal{C}(q_0, q_T, [0,T]) \to \mathbb{R}$ by

$$\mathcal{J}(z) = \int_0^T L(q(t), v(t)) + \langle p(t), \dot{q}(t) - v(t) \rangle\ dt$$

The Pontryagin Bundle is a vector bundle over $\mathcal{Q}$ where the fiber in a point $q \in \mathcal{Q}$ is given by the vector space $T_q\mathcal{Q} \times T_q^*\mathcal{Q}$. The HP path space therefore forms a smooth infinite-dimensional manifold.

Theorem 2.3. ***(Variational Principle of Hamilton-Pontryagin)***
Let L be a Lagrangian on $T\mathcal{Q}$ and $q : [0,T] \to \mathcal{Q}$ with continuous partial derivatives of second order with respect to q and v. A curve $c = (q, v, p) \in \mathcal{C}(q_0, q_T, [0,T])$ satisfies the HP equations

$$\dot{q} = v \tag{2.8}$$

$$\dot{p} = \frac{\partial L}{\partial q}(q, v) \tag{2.9}$$

$$p = \frac{\partial L}{\partial v}(q, v) \tag{2.10}$$

if and only if c is a critical point of of the HP action integral $\mathcal{J} : \mathcal{C}(q_0, q_T, [0,T]) \to \mathbb{R}$.

Proof. We start by taking the variation with respect to q

$$\begin{aligned}\frac{d}{d\varepsilon}\mathcal{J}(q^\varepsilon, v, p)\Big|_{\varepsilon=0} &= \frac{d}{d\varepsilon}\int_0^T L(q,v)\cdot\delta q + p\cdot(\dot{q}-v)\,dt\Big|_{\varepsilon=0}\\ &= \int_0^T \frac{\partial L}{\partial q}(q,v)\cdot\delta q + p\cdot\delta\dot{q}\,dt\\ &= \int_0^T \frac{\partial L}{\partial q}(q,v)\cdot\delta q\,dt + p\cdot\delta\dot{q}\Big|_0^T - \int_0^T \dot{p}\cdot\delta q\,dt\\ &= \int_0^T \Big(\frac{\partial L}{\partial q}(q,v) - \dot{p}\Big)\cdot\delta q\,dt\end{aligned}$$

and receive 2.9. Taking the variation with respect to v, and p in a similar way, we receive 2.8, 2.10. [Bou07; YM06a; YM06b; HHIL06]

□

2.1 The HP Equations and the Fiber Derivative

By introducing a kinematic constraint, we were able to formulate an action integral on the configuration space $C(z_1, z_2, [a,b])$, whose extremal satisfies the Hamilton-Pontryagin (HP) equations. By eliminating v using equation 2.10, we derive an initial value problem defined on the cotangent bundle T^*Q. This implies that the extremal can be viewed as an integral curve of a vector field on T^*Q. A useful comparison can be made between this approach of obtaining a vector field on T^*Q and the more conventional method used to derive Hamilton's equations from the second-order Euler-Lagrange equations on the tangent bundle TQ.

In the standard procedure, we starts with a Lagrangian L defined on TQ, and then transitions to T^*Q through the Legendre transformation of L to obtain the Hamiltonian $H : T^*Q \to \mathbb{R}$. The Hamiltonian is given by

$$H(q,p) = \langle p, v(q,p)\rangle - L(q, v(q,p)), \tag{2.11}$$

where $\partial L/\partial v\ (q, v(q,p)) = p$.

The non-degeneracy of L together with the implicit function theorem ensure that the velocity v can be expressed as a function of the generalized coordinates q and momenta p. This relationship is known as the fiber derivative of L. Hamilton's equations are then derived from the phase space principle, which yields the following system

$$\dot{q} = \frac{\partial H}{\partial p}(q,p) \tag{2.12}$$

$$\dot{p} = -\frac{\partial H}{\partial q}(q,p). \tag{2.13}$$

However, these equations are not yet in the form of the HP equations. To put 2.12 into the correct form, we perform another Legendre transform on the Hamiltonian H with respect to p. Similarly, in order to put equation 2.13 in the correct form, we differentiate the Legendre transform of L with respect to the velocity v obtaining

$$L(q, v(q,p)) = \langle p, v(q,p)\rangle - H(q,p), \quad \frac{\partial H}{\partial p}(q,p) = v(q,p).$$

The kinematic constraint then emerges naturally from the fiber derivative of the Hamiltonian. In summary, to derive the HP equations (which describe a vector field on T^*Q) from a Lagrangian L on TQ, one must perform two successive Legendre transforms, or equivalently, compute two fiber derivatives. Contrary, when deriving the HP equations from the HP principle itself, the Legendre transformation arises naturally from the principle, and no prior transformation is necessary. [Bou07; LR04; KBS23b; KBS23a]

2.2 Symplecticness of Flow

Symplecticness is defined by so-called structure matrices in $\mathbb{R}^{2d\times 2d}$, which can be (and in our case, will be) skew-symmetric, bilinear forms. Such a form can be described by a matrix J and gives the following definition.

A linear map $L : \mathbb{R}^{2d} \to \mathbb{R}^{2d}$ is called symplectic with regard to the structure matrix J if

$$L^T J L = J.$$

A differential map $\varphi : \mathbb{R}^{2d} \to \mathbb{R}^{2d}$ is called symplectic with regard to the structure matrix J if the Jacobian matrix $\varphi'(p,q)$ is symplectic

$$\varphi'(p,q)^T J \varphi'(p,q) = J.$$

On $\mathbb{R}^{2d}$, a skew-symmetric bilinear form Ω is induced by the matrix

$$J = \begin{pmatrix} 0 & I_d \\ -I_d & 0 \end{pmatrix} \in \mathbb{R}^{2d\times 2d}$$

For $d = 1$, we can consider $\Omega(\xi,\eta) = \xi^p\eta^q - \xi^q\eta^p$ as the oriented area of the parallelogram spanned by $\xi = \begin{pmatrix}\xi^q\\ \xi^p\end{pmatrix}$ and $\eta = \begin{pmatrix}\eta^q\\ \eta^p\end{pmatrix}$. In higher dimensions this leads to the sum of the oriented areas of the projections onto the coordinate planes (q_i, p_i) by

$$\Omega(\xi,\eta) = \sum_{i=1}^{d} \xi_i^p\eta_i^q - \xi_i^q\eta_i^p$$

If we now take a 2-dimensional sub-manifold M of $\mathbb{R}^{2d}$, given by a parameterization as $M = \psi(K), K \subset \mathbb{R}^2$, with $\psi(s,t)$ being a continuously differentiable function, we can consider M as the limit of a union of small parallelograms spanned by

$$\frac{\partial\psi}{\partial s}(s,t)ds \text{ and } \frac{\partial\psi}{\partial t}(s,t)dt$$

For one such parallelogram, taking the sum over the oriented areas of the projections as given above yields

$$\int_M \Omega = \iint_K \Omega(\frac{\partial \psi}{\partial s}(s,t), \frac{\partial \psi}{\partial t}(s,t))\, ds\, dt$$

The definition of a symplectic mapping $\varphi : \mathbb{R}^{2d} \to \mathbb{R}^{2d}$ then becomes

$$\varphi^*\Omega(\xi,\eta) := \Omega(\varphi'(p,q)\xi, \varphi'(p,g)\eta) = \Omega(\xi,\eta)$$

according to the definition before. Hence there holds

$$\int_M \Omega = \iint_K \Omega(\varphi'(p,q)\frac{\partial \psi}{\partial s}(s,t), \varphi'(p,q)\frac{\partial \psi}{\partial t}(s,t))\, ds\, dt.$$

This implies for $d = 1$ the conservation of area under a symplectic map and for $d > 1$ the conservation of the sum of oriented areas of projections of M on to the (q_i, p_i)-coordinate planes.

The HP equations give us a system of differential-algebraic equations from which v can be eliminated by a function of q and p. The resulting system can then be viewed as an initial value problem. Given values $(q(0), p(0))$ we can define a flow map $\varphi_{HP_t} : T^*\mathcal{Q} \to T^*\mathcal{Q}, \quad (q(0), p(0)) \mapsto (q(t), p(t))$ for a fixed time step, simply by integrating the equations.

As symplecticity is an important property of a mechanical systems we want to show that the HP flow preserves it. For that it suffices to show that there exists a twice continuously differentiable function $H(q,p)$ fulfilling 2.12 and 2.13

Theorem 2.4. *(Poincaré, 1899)*
Let $H(p,q)$ be a twice continuously differentiable function for which 2.12 and 2.13 holds on an open set $U \subset \mathbb{R}^{2d}$. Then, for each fixed t, the Hamiltonian flow φ_t is a symplectic transformation wherever it is defined.

We can now define the smooth vector field $F_{HP} : T^*\mathcal{Q} \to T(T^*\mathcal{Q})$ as

$$F_{HP}(q,p) = \frac{d}{dt}(p,-q) = (\frac{\partial L}{\partial q}(q, v(q,p)), -v(q,p)) \tag{2.14}$$

Computing the Jacobian Matrix of F_{HP} using 2.10 and the implicit function theorem, we get

$$\nabla F_{HP}(q,p) = \begin{pmatrix} L_{qq} + L_{qv}v_q & L_{qv}v_p \\ -v_q & -v_p \end{pmatrix} = \begin{pmatrix} L_{qq} - L_{qv}L_{vv}^{-1}L_{qv} & -L_{qv}L_{vv}^{-1} \\ -L_{vv}^{-1}L_{qv} & -L_{vv}^{-1} \end{pmatrix} \tag{2.15}$$

which is symmetric. From the Integrability Lemma follows, that for every (q_0, p_0) there exists a neighborhood and a function $H(q,p)$ such that $F_{HP} = \nabla H$.

Theorem 2.5. *(Integrability Lemma)*
Let $D \subset \mathbb{R}^n$ be open and $f : D \to \mathbb{R}^n$ be continuously differentiable. Assume that the Jacobian matrix $\nabla f(y)$ is symmetric for all $y \in D$. Then, for every $y_0 \in D$, there exists a neighborhood of y_0 and a function $H(y)$ such that

$$f(y) = \nabla H(y)$$

on this neighborhood. In other words, the differential form

$$f_1(y)\, dy_1 + \cdots + f_n(y)\, dy_n = dH$$

is a total differential. [Bou07; LR04; KBS23b; KBS23a; HHIL06]

3 A Variational Discontinuous Petrov-Galerkin Time Discretization on Vector Spaces

In this section, we propose a variational time discretization of the Hamilton-Pontryagin (HP) functional 2.7, based on a discontinuous Petrov–Galerkin (DPG) approach. Our objective is to construct a numerically stable and structure-preserving method for the time integration of mechanical systems. To this end, and to allow for a analytical treatment, we restrict our attention to linear rigid body systems, in which of the system is given by a quadratic form. The variational formulation corresponding to the HP principle yields the system

$$\begin{cases} (q, v, p) \in \mathcal{C}(q_0, q_1, [0, T]) \\ \int_0^T (\dot{q} - v) \cdot \delta p \, dt = 0, \\ \int_0^T \left(p - \frac{\partial L}{\partial v}(q, v)\right) \cdot \delta v \, dt = 0, \quad \forall (\delta q, \delta v, \delta p) \in \mathcal{C}(0, 0, [0, T]), \\ \int_0^T \left(\frac{\partial L}{\partial q}(q, v) - \dot{p}\right) \cdot \delta q \, dt = 0. \end{cases} \tag{3.1}$$

To enable a practical numerical treatment, we discretize the time interval $[0, T]$ by introducing a uniform partition

$$0 = t_0 < t_1 < \cdots < t_N = T,$$

with subintervals $I_n = (t_n, t_{n+1})$ and fixed time step size $h = t_{n+1} - t_n$.

3.1 Breaking the Spaces

In the Petrov-Galerkin framework, we introduce a larger, broken trial space. We begin by relaxing the boundary conditions for δq, thereby allowing discontinuities at the interval junctions. These broken test functions introduce an additional variable, as integrating the third equation by parts yields

$$\sum_{i=0}^{N-1} \int_{t_i}^{t_{i+1}} p \cdot \dot{\delta q} + \frac{\partial L}{\partial q}(q, v) \cdot \delta q \, dt - p \cdot \delta q \Big|_{t_i}^{t_{i+1}} = \sum_{i=0}^{N-1} \int_{t_i}^{t_{i+1}} p \cdot \dot{\delta q} + \frac{\partial L}{\partial q}(q, v) \cdot \delta q \, dt - \hat{p} \cdot \delta q \Big|_{t_i}^{t_{i+1}},$$

where $\hat{p} = p|_{\{t_0, t_1, \ldots, t_N\}}$ is defined only at the interval boundaries.

Our updated formulation is

$$\begin{cases} (q,v,p) \in \mathcal{C}(q_0, q_N, [0,T]), \quad \hat{p} \in \bigtimes_{i=0}^{N} T^*_{q(t_i)}\mathcal{Q}, \\ \sum_{i=0}^{N-1} \int_{t_i}^{t_{i+1}} (\dot{q} - v) \cdot \delta p \, dt = 0, \\ \sum_{i=0}^{N-1} \int_{t_i}^{t_{i+1}} \left(p - \frac{\partial L}{\partial v}(q,v)\right) \cdot \delta v \, dt = 0, \quad \forall (\delta q, \delta v, \delta p) \in \mathcal{C}(\cdot,\cdot,[0,T]), \\ \sum_{i=0}^{N-1} \int_{t_i}^{t_{i+1}} \frac{\partial L}{\partial q}(q,v) \cdot \delta q + p \cdot \dot{\delta q} \, dt - \hat{p} \cdot \delta q \Big|_{t_i}^{t_{i+1}} = 0. \end{cases} \tag{3.2}$$

To simplify the formulation in the following chapter, we focus only on a single interval I_0. Assuming that $\mathcal{Q}$ is a d-dimensional vector space over $\mathbb{R}$ (i.e., $\mathcal{Q} \cong \mathbb{R}^d$), we can extend our test and trial spaces to $H^1([0,t_1],\mathbb{R}^d)$. Together with the assumption that our potential is a quadratic form we get

$$\begin{cases} q \in H^1([0,t_1],\mathbb{R}^d), q(0) = q_0 \\ \hat{p}_1 \in \mathbb{R}^d, \\ \dot{q} = v \\ p = Mv \\ \int_0^{t_1} -Kq \cdot \delta q + p \cdot \dot{\delta q} \; dt - \hat{p}_1 \cdot \delta q(t_1) = -\hat{p}_0 \cdot \delta q(0), \quad \forall \delta q \in H^1([0,t_1],\mathbb{R}^d). \end{cases} \tag{3.3}$$

where K is the semi positive definite matrix defining $V(q) = 1/2 q^T K q$ and M is the positive definite mass matrix defining $T(v) = 1/2 v^T M v$. Using the first two equations to substitute p by $M\dot{q}$ in the third equation yields

$$\begin{cases} q \in H^1([0,t_1],\mathbb{R}^d), q(0) = q_0 \\ \hat{p}_1 \in \mathbb{R}^d, \\ \int_0^{t_1} -Kq \cdot \delta q + p \cdot \dot{\delta q} \; dt - \hat{p}_1 \cdot \delta q(t_1) = -\hat{p}_0 \cdot \delta q(0), \quad \forall \delta q \in H^1([0,t_1],\mathbb{R}^d). \end{cases} \tag{3.4}$$

In the subsequent chapters we will proof that 3.4 is uniquely solvable.

3.2 Abstract Variational Formulation

From now on we interpret problem 3.4 in an abstract variational framework. Let $b : X \times Y \to \mathbb{R}$ be a bilinear form and $l \in Y^*$ a linear functional. We seek $x \in X$ such that

$$b(x,v) = l(v), \quad \forall v \in Y, \tag{3.5}$$

where X and Y are Hilbert spaces, and Y^* denotes the space of continuous linear functionals on Y. A standard approach would be to use the lemma of Lax Milgram to check the solvability. However this theorem is limited to coercive bilinear forms with $X = Y$. But fortunately this is not the only way of proving stable solvability. One alternative way arises from the well-known theory of mixed systems

Theorem 3.1. *In the above setting of Hilbert space X and Y, 3.5 holds if and only if*

$$\inf_{x \in X\setminus\{0\}} \sup_{y \in Y\setminus\{0\}} \frac{|b(x,y)|}{\|x\|_X \|y\|_Y} \geq \gamma, \quad \textit{and} \tag{3.6}$$

$$\{y \in Y \ : \ b(x,y) = 0 \textit{ for all } x \in X\} = \{0\}, \tag{3.7}$$

or equivalently

$$\inf_{y \in Y\setminus\{0\}} \sup_{x \in X\setminus\{0\}} \frac{|b(x,y)|}{\|x\|_X \|y\|_Y} \geq \gamma, \quad \textit{and} \tag{3.8}$$

$$\{x \in X \ : \ b(x,y) = 0 \textit{ for all } y \in Y\} = \{0\}. \tag{3.9}$$

To realize this computationally, we use finite-dimensional subspaces $X_h \subseteq X$ and $Y_h \subseteq Y$ with $dim(X_h) = dim(Y_h)$ and pose the discrete problem as follows

Find $x_h \in X_h$ such that

$$b(x_h, v_h) = l(v_h), \quad \forall v_h \in Y_h. \tag{3.10}$$

However, the solvability of the finite dimensional equation does not follow from the well-posedness of the original problem. We still have to show that the discrete problem fulfills the inf-sup condition as well.

Theorem 3.2. *In the above setting of Hilbert space X, Y and finite dimensional subspaces $X_h \subset X$ and $Y_h \subset Y$ with $dim(X_h) = dim(Y_h)$. If 3.5 holds and provided there exists a $\gamma_h > 0$ such that*

$$\inf_{x_h \in X_h\setminus\{0\}} \sup_{y_h \in Y_h\setminus\{0\}} \frac{|b(x_h,y_h)|}{\|x_h\|_X \|y_h\|_Y} \geq \gamma_h, \tag{3.11}$$

then there is a unique $x_h \in X_h$ satisfying

$$b(x_h, v_h) = l(v_h), \quad \forall v_h \in Y_h. \tag{3.12}$$

and

$$\|x - x_h\|_X \leq \frac{\|b\|}{\gamma_h} \inf_{z_h \in X_h} \|x - z_h\|_X \tag{3.13}$$

This theorem addresses a fundamental difficulty in establishing discrete stability solely based on the well-posedness of the continuous variational problem. One of the central challenges in analyzing Petrov–Galerkin discretizations of the form 3.10 is that the continuous inf-sup condition 3.6 does not, in general, imply the validity of the discrete inf-sup condition 3.11, in contrast to coercive bilinear forms $a(\cdot,\cdot)$, for which discrete stability is inherited from the continuous setting. Consequently, it is possible for Petrov-Galerkin methods to be unstable, even when the variational formulation 3.5 is well-posed.

A further notable observation drawn from theorem 3.2 is that the test space Y_h does not appear in the error estimate 3.13. Its influence is restricted to the discrete inf-sup condition, which governs the stability of the method. The approximation properties reflected in 3.13 depend exclusively on the choice of the trial space X_h. This decoupling permits the design of X_h to prioritize approximation quality, while the construction of Y_h can be directed solely towards fulfilling the stability criterion. As a consequence, methods for constructing discrete test spaces Y_h that satisfy the discrete inf-sup condition independently of the dimension or structure of X_h are of particular relevance. [Bab71; Bre74; EG+21; DG25; Bab72; Neč62]

3.3 Optimal Test Spaces

The problem of constructing such test spaces is exactly which has given rise to the ideal Petrov–Galerkin Method. For a given continuous linear form $b(\cdot,\cdot) : X \times Y \to \mathbb{R}$ we define a Trial-to-test Operator $T : X \to Y$ as

$$\langle Tx, y\rangle_Y = b(x,y) \quad \text{for all } y \in Y \tag{3.14}$$

This operator is well defined due to the Riesz Representation Theorem and Tz is called an optimal test function because of the following characteristic.

Proposition 3.3. ***(Optimizer)*** *For any $z \in X$ the maximum of*

$$f_z(y) = \frac{|b(z,y)|}{\|y\|_Y} \tag{3.15}$$

over all nonzero $y \in Y$ is attained at $y = Tz$.

Proof. Using 3.14 to rewrite f_z together with the Cauchy-Schwarz Inequality we get

$$\sup_{y \in Y\setminus\{0\}} f_z(y) = \sup_{y \in Y\setminus\{0\}} \frac{|\langle T(u,p), y\rangle_Y|}{\|y\|_Y} = \|Tz\|_Y$$

with $f_z(Tz) = \|Tz\|_Y$. □

As can be seen easily, now we can define our optimal test space Y_h^{opt} as $Y_h^{opt} = T(X_h)$ because then the discrete inf-sup condition 3.11 is implied by the continuous case 3.6.

Proposition 3.4. *If the inf-sup condition 3.6 holds with some $\gamma > 0$, then the discrete inf-sup condition 3.11 holds with some $\gamma_h \geq \gamma > 0$ when we set $Y_h \supseteq Y$.*

Proof. For any $x_h \in X_H$ with

$$s_1 = \sup_{y\in Y\setminus\{0\}} \frac{|b(z_h,y)|}{\|y_h\|_Y}, \quad s_2 = \sup_{y_h\in Y_h\setminus\{0\}} \frac{|b(z_h,y_h)|}{\|y_h\|_Y}$$

By the definition of the supremum it holds that $s_1 \geq s_2$. With Proposition 3.3 we can reformulate s_1 as

$$s_1 = \|Tz_h\|_Y = \frac{|\langle Tz_h, Tz_h\rangle_Y|}{\|Tz_h\|_Y} \leq \sup_{y_h\in Y_h^{opt}\setminus\{0\}} \frac{|\langle Tz_h, y_h\rangle_Y|}{\|y_h\|_Y} \leq \sup_{y_h\in Y_h\setminus\{0\}} \frac{|\langle Tz_h, y_h\rangle_Y|}{\|y_h\|_Y} = s_2$$

Hence $s_1 = s_2$ and the discrete inf-sup condition 3.11 holds with $\gamma_h \geq \gamma > 0$.

□

3.4 Well-posedness

We have now developed the theory to proof that formulation 3.4 is well-posed. We start by reducing 3.4 to a problem with zero boundary value. If the homogeneous problem is well posed then we continue by posing the problem as finding $u \in H^1([0,t_1];\mathbb{R}^d)$ with $u(0) = u_0 \neq 0$ such that

$$b(u,v) = l(v), \quad \forall v \in Y.$$

This is uniquely solved by $u = u_0 + \tilde{u}$ with $\tilde{u} \in \{v \in H^1([0,t_1];\mathbb{R}^d) \mid v(0) = 0\}$ and $\tilde{u}$ solving

$$b(\tilde{u},v) = l(v) - b(u_0,v), \quad \forall v \in Y.$$

Therefore it suffices to show, that for $X = \{v \in H^1([0,t_1];\mathbb{R}^d) \mid v(0) = 0\} \times \mathbb{R}^d$, $Y = H^1([0,t_1];\mathbb{R}^d)$ the bilinear form $b : X \times Y \to \mathbb{R}$ defined as

$$b((u,p),v) = \int_0^{t_1} -Ku \cdot v + M\dot{u} \cdot \dot{v} \; dt - p \cdot v(t_1) \tag{3.16}$$

fulfills the conditions 3.6 and 3.7. On X we define the norm $\|(u,p)\|_X^2 = \|u\|_{H^1}^2 + |p|^2$ and on Y as $\|v\|_Y^2 = \|\dot{v}\|_{L^2}^2 + |v(t_1)|^2$. Following the Riesz Representation Theorem, for each $(u,p) \in X$ there exists a $y_{(u,p)} \in Y$ such that $\forall\ v \in Y : b((u,p),v) = \langle y_{(u,p)}, v\rangle_Y$. We can now solve for the Trial-to-Test operator analytically

Lemma 3.5. *A Trial-to-Test operator $T : X \to Y$ for the bilinear form 3.16 is given by*

$$T(u,p)(t) = M(u(t) - u(t_1)) - p + K\int_0^t u(\tau)(t-\tau)\; d\tau - K\int_0^{t_1} u(\tau)(1+t_1-\tau)\; d\tau \tag{3.17}$$

Proof. On Y we have defined a scalar product as $\langle v,u\rangle_Y = \langle \dot{v},\dot{v}\rangle_{L^2}^2 + v(t_1)\cdot v(t_1)$. The first time derivative and boundary value for $T(u,p)$ are given by

$$T(\dot{u},p)(t) = M\dot{u}(t) + K\int_0^t u(\tau)\; d\tau$$

$$T(u,p)(t_1) = -p - K\int_0^{t_1} u(\tau)\; d\tau$$

From this follows that

$$\begin{aligned}\langle T(u,p),v\rangle_Y &= \int_0^{t_1} M\dot{u}(t)\cdot\dot{v}(t) + (K\int_0^t u(\tau)\; d\tau)\cdot\dot{v}(t)\; dt - p\cdot v(t_1) - (K\int_0^{t_1} u(\tau)d\tau)\cdot v(t_1)\\ &= \int_0^{t_1} M\dot{u}(t)\cdot\dot{v}(t) - Ku(t)\cdot v(t)\; dt - p\cdot v(t_1)\\ &= b((u,p),v).\end{aligned}$$

□

With proposition 3.3 we can now rewrite 3.6

$$\inf_{(u,p)\,\in\,X\setminus\{0\}}\ \sup_{v\,\in\,Y\setminus\{0\}}\frac{|b((u,p),v)|}{\|(u,p)\|_X\|v\|_Y}=\inf_{(u,p)\,\in\,X\setminus\{0\}}\frac{\|Tz\|_Y}{\|(u,p)\|_X}\geq\gamma$$

and see that it suffices to show that

$$\|T(u,p)\|_Y\geq\gamma\|(u,p)\|_X \tag{3.18}$$

holds. To prove this, we will first show that it holds for the dense subspace $X_0=\{u\in C^\infty([0,t_1];\mathbb{R}^d)\}\times\mathbb{R}^d$ of X and then use a limiting argument. But before this, we have to prove a few important properties of operator 3.17.

Lemma 3.6. *The operator $T:X\to Y$ defined by 3.17 is linear and continuous. Furthermore $T|_{X_0}$ is injective.*

Proof. (i) As T only consists of linear terms, it is linear itself.

(ii) We want to show that $\|T(u,p)\|_Y^2=\|T(\dot{u},p)\|_{L^2}^2+|T(u,p)(t_1)|^2\leq C(\|u\|_{H^1}^2+|p|^2)$. The derivative and boundary values of $T(u,p)$ are

$$T(\dot{u},p)(t)=M\dot{u}(t)+K\int_0^t u(\tau)\,d\tau$$

$$T(u,p)(t_1)=-p-K\int_0^{t_1}u(\tau)\,d\tau$$

and therefore, using the Cauchy-Schwarz inequality together with $(a+b)^2\leq 2a^2+2b^2$ yields

$$\begin{aligned}\|T(\dot{u},p)\|_{L^2}^2&\leq\int_0^{t_1}(\|M\|_{max}|\dot{u}(t)|+\sqrt{t_1}\|K\|_{max}\|u\|_{L^2})^2\,dt\\&\leq 2\|M\|_{max}^2\|\dot{u}\|_{L^2}^2+2{t_1}^2\|K\|_{max}^2\|u\|_{L^2}^2\end{aligned}$$

$$\begin{aligned}|T(u,p)(t_1)|^2&\leq 2|p|^2+2|K\int_0^{t_1}u(\tau)\,d\tau|^2\\&\leq 2|p|^2+2t_1\|K\|_{max}^2\|u\|_{L^2}^2.\end{aligned}$$

Combining both parts we get

$$\|T(u,p)\|_Y^2\leq 2\max\{1,t_1\|K\|_{max}^2,\|M\|_{max}^2,{t_1}^2\|K\|_{max}^2\}(\|u\|_{H^1}^2+|p|^2)$$

(iii) To prove that $T|_{X_0}$ is injective, we want show that $\ker T|_{X_0}=\{0\}$ as T is linear. Assume $T(u,p)=0$, then

$$T(\dot{u},p)(0)=Mu(0)=0\implies u(0)=0$$

$$T(\ddot{u},p)(t)=M\ddot{u}(t)+Ku(t)$$

This is a system of second order differential equations with initial values $u(0)=\dot{u}(0)=0$. We can transform this system into a first order system as M is positive-definite and therefore know that the only solution is $u(t)=0$. □

Now that we have established some properties of T, we can go on with proving 3.18. Consider the unit sphere $S = \{(u,p) \in X_0 \mid \|(u,p)\|_X = 1\}$ and the functional

$$\Phi : X_0 \to \mathbb{R},\ (u,p) \mapsto \|T(u,p)\|_Y$$

which is continuous as $\|T(u,p)\|_Y \leq C\|(u,p)\|_X$ .Furthermore, because $T|_{X_0}$ is injective it holds that

$$\Phi(u,p) > 0 \iff (u,p) \neq 0$$

Assume now, that

$$\inf_{(u,p)\in S} \Phi(u,p) = 0.$$

Then there exists a sequence $(u_k, p_k) \in S^{\mathbb{N}}$ such that $\lim_{k\to\infty} \Phi(u_k, p_k) = 0$. We can define the energy functionals

$$E_k(t) = \frac{1}{2}\dot{u}_k(t)^T M \dot{u}_k(t) + \frac{1}{2} u_k(t)^T K u_k(t)$$

with the according time derivative

$$\dot{E}_k(t) = \dot{u}_k(t)^T (M\ddot{u}_k(t) + K u_k(t))$$

As $\|T(u_k, p_k)\|_Y \to 0$ it follows that

$$T(\dot{u_k}, p_k)(t) = M\dot{u}_k(t) + K\int_0^t u_k(\tau)\ d\tau \to 0 \quad \text{in } L^2$$

$$T(\dot{u_k}, p_k)(t_1) = M\dot{u}_k(t_1) + K\int_0^{t_1} u_k(\tau)\ d\tau \to 0 \quad \text{in } \mathbb{R}^d$$

and therefore

$$T(\ddot{u_k}, p_k) = M\ddot{u}_k + K u_k \to 0 \quad \text{in } L^2$$

holds for the second derivative as $\|T(\dot{u_k}, p_k)\|_Y \to 0$. Using the Trace Operator, we see that the boundary value

$$T(\dot{u_k}, p_k)(0) = M\dot{u}_k(0)$$

converges to zero as well. Putting all the pieces together gives us

$$E_k(0) = \frac{1}{2}\dot{u}_k(0)^T M \dot{u}_k(0) + \frac{1}{2} u_k(0)^T K u_k(0) \to 0$$

and furthermore, as $\|\dot{u}\|_{L^2} \leq 1$, it follows that $\dot{E}_k(t) \to 0$ in L^2.
If we now reformulate $E_k(t)$ as the Integral of $\dot{E}$ we get that

$$\begin{aligned} |E_k(t)| &= |E_k(0) + \int_0^t \dot{E}(\tau)\ d\tau| \\ &\leq |E_k(0)| + |\int_0^t \dot{E}(\tau)\ d\tau| \\ &\leq |E_k(0)| + \|\dot{E}\|_{L^2} \to 0 \quad \text{uniformly on } [0, t_1]. \end{aligned}$$

As $E_k(t)$ converges uniformly and

$$|E_k(t)| \geq 1/2|\dot{u}(t)^T M \dot{u}(t)^T| \geq 1/2\lambda_{min}(M)|\dot{u}(t)|^2$$

$\dot{u}$ we also get uniform convergence for $\dot{u}$ and therefore with the following theorem (H.W. Engel S.232 Satz 9.10) $\|\dot{u}\|_{L^2} \to 0$ and $\|u\|_{L^2} \to 0$.

Theorem 3.7. *Let for all $n \in \mathbb{N}$: $f_n : [a,b] \to \mathbb{R}^d$ be differentiable, (f_n') converges uniformly to $g : [a,b] \to \mathbb{R}^d$ and there exist $x_0 \in [a,b]$ such that $f_n(x_0) \to y_0$. Then there exists $f : [a,b] \to \mathbb{R}^d$, such that $f = lim_{n\to\infty} f_n$ uniformly and is differentiable on (a,b) with*

$$f'(x) = \lim_{n\to\infty} f_n'(x) = g(x) \quad \textit{for all } x \in (a,b).$$

As $p_k = -T(u_k, p_k)(t_1) - K\int_0^{t_1} u_k(\tau)d\tau$, it also converges to zero resulting in

$$\|(u_k, p_k)\|_X^2 = \|u_k\|_{H^1}^2 + |p_k|^2 \to 0,$$

which is a contradiction to $\|(u_k, p_k)\|_X^2 = 1,\ \forall\ k \in \mathbb{N}$.

It therefore holds that

$$\|T(u,p)\|_Y \geq \gamma\|(u,p)\|_X, \forall\ (u,p) \in X_0$$

As X_0 is dense in X, for each $(u,p) \in X$ there exists a sequence $((u_k, p_k))_{k\in\mathbb{N}} \in X_0^{\mathbb{N}}$ such that $(u_k, p_k) \to (u,p)$ in X. As $\|T(u_k,p_k)\|_Y \geq \gamma\|(u_k,p_k)\|_X$ holds for every $k \in \mathbb{N}$ it also holds for (u,p).

We have now shown that our bilinear form b fulfills the "inf-sup" condition 3.6 but we are still left with showing that $N = \{v \in Y\ |b((u,p),v) = 0,\ \forall\ (u,p) \in X\} = \{0\}$.
Let $v \in N$, then taking elements of form $(0,p) \in X$ yields

$$b((0,p),v) = -p \cdot v(t_1) = 0,\ \forall\ p \in \mathbb{R}^d$$

and therefore, $v(t_1) = 0$. For every $\phi \in C_b^\infty([0,t_1];\mathbb{R})$ we can find exactly one $(u,p) \in X$ solving $M\ddot{u}(t) + Ku(t) = -\phi(t), u(0) = 0, \dot{u}(0) = 0$. Using integration by parts we get that

$$\begin{aligned} b((u,p),v) &= \int_0^{t_1} -(Ku(t)) \cdot v(t) + (M\dot{u}(t)) \cdot \dot{v}(t)\ dt \\ &= \int_0^{t_1} -(Ku(t)) \cdot v(t) - (M\ddot{u}(t)) \cdot v(t)\ dt\ + \dot{u}(t_1) \cdot v(t_1) - \dot{u}(0) \cdot v(0) \\ &= \int_0^{t_1} \phi(t) \cdot v(t)\ dt = 0 \end{aligned}$$

By the fundamental lemma of the calculus of variations it must hold that $v = 0$. Therefore we have shown that our problem 3.4 is well-posed.

3.5 The Discrete System

For our discrete problem, we consider the spaces

$$X_n = \{u \in P_n([0,t_1];\mathbb{R}^d) \mid u(0) = 0\} \times \mathbb{R}^d, \quad Y_n = P_n([0,t_1];\mathbb{R}^d),$$

and again only focus on the homogeneous boundary value problem.

As stated in theorem 3.2, we only need to prove that the discrete inf-sup condition (3.11) holds. Unfortunately, the straightforward approach of showing $Y_n \supseteq Y_n^{\text{opt}} = T(X_n)$ does not work. Analyzing the optimal test space $T(X_n)$, we find that for $u_n \in X_n$,

$$T(u_n,p)(t) = M(u_n(t) - u_n(t_1)) - p + K\int_0^t u_n(\tau)(t-\tau)\,d\tau - K\int_0^{t_1} u_n(\tau)(1+t_1-\tau)\,d\tau$$

is a polynomial of degree at most $n+2$, due to the convolution term $\int_0^t u_n(\tau)(t-\tau)\,d\tau$ and therefore have to take another approach.

Let B_n denote the operator defined by

$$B_n : X_n \to Y_n^*, \quad B_n(u_n,p) = b((u_n,p),\cdot).$$

We start by proving some important properties of B_n.

Proposition 3.8. *For $h < C_{K,M}$, where $C_{K,M}$ depends on K and M, the operator B_n defined by 3.5 is a bijection. Therefore, its inverse $B_n^{-1} : Y_n^* \to X_n$ is continuous.*

Proof. To show that B_n is a bijection, it suffices to show that it is injective, since B_n is linear and $\dim(X_n) = \dim(Y_n) = \dim(Y_n^*)$.

Assume $B_n(u,p)(v) = b((u,p),v) = 0$ for all $v \in Y_n$. We want to show that $(u,p) = (0,0)$.

Since we are working in finite-dimensional spaces, we can define equivalent norms on X_n and Y_n as

$$\|(u,p)\|_{X_n} = \|\dot{u}\|_{L^2} + |p|, \quad \|v\|_{Y_n} = \|\dot{v}\|_{L^2} + |v(t_1)|.$$

We consider two cases:

(i) $|p| \geq \|\dot{u}\|_{L^2}$**:**

Let $v(t) \equiv v_c := -\operatorname{sign}(p)\frac{p}{|p|} \in Y_n$. Then

$$|u(t)| = \left|\int_0^t \dot{u}(\tau)\,d\tau\right| \leq \sqrt{t_1}\|\dot{u}\|_{L^2},$$

$$\|u\|_{L^2}^2 \leq \int_0^{t_1} |u(t)|^2\,dt \leq {t_1}^2\|\dot{u}\|_{L^2}^2,$$

$$\left|\int_0^{t_1} Ku\cdot v\,dt\right| \leq \|K\|_{\max}\|u\|_{L^2}\|v\|_{L^2} \leq \|K\|_{\max}\|\dot{u}\|_{L^2}{t_1}^{3/2}.$$

Hence, we find a lower bound for $|b((u,p),v)|$, by

$$\begin{aligned}|b((u,p),v)| &= \left|\int_0^{t_1} Ku\cdot v\,dt + v\cdot p\right| \\ &= \left|\int_0^{t_1} Ku\cdot v\,dt + |p|\right| \\ &\geq |p| - \|K\|_{\max}\|\dot{u}\|_{L^2}{t_1}^{3/2} \\ &\geq |p|(1-\|K\|_{\max}{t_1}^{3/2}).\end{aligned}$$

For t_1 small enough such that ${t_1}^{3/2} < \frac{1}{\|K\|_{\max}} := \beta_1$, and using $|p| \geq \|\dot{u}\|_{L^2}$, we conclude with

$$|b((u,p),v)| \geq \frac{c}{2}(|p| + \|\dot{u}\|_{L^2}). \tag{3.19}$$

(ii) $|p| < \|\dot{u}\|_{L^2}$**:**
Let $v = u/\|\dot{u}\|_{L^2}$. Then

$$v(0) = 0, \quad \|\dot{v}\|_{L^2} = 1, \quad |v(t_1)| \leq \sqrt{t_1}, \quad \|v\|_{Y_n} = 1 + \sqrt{t_1},$$

$$\left|\int_0^{t_1} M\dot{u}\cdot\dot{v}\,dt\right| = \frac{1}{\|\dot{u}\|_{L^2}}\int_0^{t_1} M\dot{u}\cdot\dot{u}\,dt \geq \lambda_{\min}(M)\|\dot{u}\|_{L^2},$$

$$\left|\int_0^{t_1} Ku\cdot v\,dt\right| \leq \|K\|_{\max}\|\dot{u}\|_{L^2}{t_1}^{3/2}, \quad |p\cdot v(t_1)| \leq \|\dot{u}\|_{L^2}\sqrt{t_1}.$$

Thus, the lower bound is

$$\begin{aligned}|b((u,p),v)| &\geq \lambda_{\min}(M)\|\dot{u}\|_{L^2} - \|K\|_{\max}\|\dot{u}\|_{L^2}{t_1}^{3/2} - \|\dot{u}\|_{L^2}\sqrt{t_1} \\ &= \|\dot{u}\|_{L^2}\left(\lambda_{\min} - \|K\|_{\max}{t_1}^{3/2} - \sqrt{t_1}\right).\end{aligned}$$

If $\lambda_{\min} - \|K\|_{\max}h^{3/2} - \sqrt{t_1} > 0$, i.e., for $t_1 < \beta_2$, then

$$|b((u,p),v)| \geq \frac{c}{2}(\|\dot{u}\|_{L^2} + |p|). \tag{3.20}$$

Since in both cases we obtain a lower bound for $|b((u,p),v)|$ for sufficiently small h, we conclude that B_n is injective, and therefore bijective. □

We have now developed a variational framework for solving rigid body problems in time, based on the Variational Principle of Hamilton-Pontryagin. As shown in chapter one 2 the flow preserves the symplectic two form and it is desirable that our discrete flow of $(q,\hat{p})$ preserves the symplectic two form as well.[HHIL06; Bab72; Bet16; DG25]

3.6 Symplecticity of Discrete Flow

We consider the solution $(q, \hat{p}_{n+1})$ to our discrete problem for given initial conditions $(q_n, \hat{p}_n)$. We start by showing that for any two initial condition pairs $(q^1_{n+1}, \hat{p}^1_{n+1})$, $(q^2_n, \hat{p}^2_n)$, the mapping $\Phi_h(q_n, \hat{p}_n) = (q_{n+1}, \hat{p}_{n+1})$, where q_{n+1} is just $q(t_{n+1})$, fulfills

$$(q^1_n, \hat{p}^1_n)^T \; J \; (q^2_n, \hat{p}^2_n) = \Phi_h(q^1_n, \hat{p}^1_n)^T \; J \; \Phi_h(q^2_n, \hat{p}^2_n). \tag{3.21}$$

For our initial conditions $(q^1_{n+1}, \hat{p}^1_{n+1})$ it holds that there exists p^1 and a $\hat{p}^1_{n+1}$ such that

$$\int_{t_n}^{t_{n+1}} -Kq^1 \cdot \delta q + M\dot{q}^1 \cdot \dot{\delta q} \; dt - \hat{p}^1_{n+1} \cdot \delta q(t_1) = -\hat{p}^1_n \cdot \delta q(0), \quad \forall \delta q \in H^1([t_n, t_{n+1}], \mathbb{R}^d)$$

Therefore we can use the solution of our second initial value problem q^2 as a test function for the variational problem of q^1 and get

$$\int_{t_n}^{t_{n+1}} -Kq^1 \cdot q^2 + M\dot{q}^1 \cdot \dot{q}^2 \; dt - \hat{p}^1_{n+1} \cdot q^2_{n+1} = -\hat{p}^1_n \cdot q^2_n$$

Using the same trick for the variational problem of q^2 yields

$$\int_{t_n}^{t_{n+1}} -Kq^1 \cdot q^2 + M\dot{q}^1 \cdot \dot{q}^2 \; dt - \hat{p}^2_{n+1} \cdot q^1_{n+1} = -\hat{p}^2_n \cdot q^1_n$$

By subtracting expression one from expression two we achieve

$$\hat{p}^2_{n+1} \cdot q^1_{n+1} - \hat{p}^1_{n+1} \cdot q^2_{n+1} = \hat{p}^2_n \cdot q^1_n - \hat{p}^1_n \cdot q^2_n,$$

and together with the following Lemma we see that Φ_h is symplectic.

Lemma 3.9. *Consider a map $\varphi : \mathbb{R}^{2d} \to \mathbb{R}^{2d}, (q_n, p_n) \mapsto (q_{n+1}, p_{n+1})$. If for any $(q, p), (\tilde{q}, \tilde{p}) \in \mathbb{R}^{2d}$ it holds that*

$$(q, p)^T \; J \; (\tilde{q}, \tilde{p}) = \varphi(q, p)^T \; J \; \varphi(\tilde{q}, \tilde{p}).$$

Then φ is a symplectic map.

Proof. We want to verify the general definition of a symplectic map

$$\varphi'(p, q)^T J \varphi'(p, q) = J$$

Let $(q, p), (\tilde{q}, \tilde{p}) \in \mathbb{R}^{2d}$ be arbitrary but fixed and let

$$(q_\varepsilon, p_\varepsilon) = (q, p) + \varepsilon(\delta q, \delta p)$$

$$(\tilde{q}_\mu, \tilde{p}_\mu) = (\tilde{q}, \tilde{p}) + \mu(\delta\tilde{q}, \delta\tilde{p})$$

where $(\delta q, \delta p), \; (\delta\tilde{q}, \delta\tilde{p}) \in \mathbb{R}^{2d}$ arbitrary. Then we know that

$$(q_\varepsilon, p_\varepsilon)^T \; J \; (\tilde{q}_\mu, \tilde{p}_\mu) = \varphi(q_\varepsilon, p_\varepsilon)^T \; J \; \varphi(\tilde{q}_\mu, \tilde{p}_\mu)$$

Differentiating both sides first with respect to ε at $\varepsilon = 0$ and then with respect to μ at $\mu = 0$ gives us

$$(q, p)^T \; J \; (\tilde{q}, \tilde{p}) = [(\delta q, \delta p)^T \; \varphi(q, p)^T] \; J \; [\varphi'(\tilde{q}, \tilde{p}) \; (\delta\tilde{q}, \delta\tilde{p})]$$

As $(\delta q, \delta p), \; (\delta\tilde{q}, \delta\tilde{p}) \in \mathbb{R}^{2d}$ are arbitrary, by choosing $(q, p) = (\tilde{q}, \tilde{p})$ we get that φ is a symplectic map. □

3.7 Order of accuracy

In this section, we want to prove that a numerical scheme which solves our discrete problem 3.5 with polynomials of degree k is of order $2k$. Let u_k be the solution of the discrete problem. We are interested in the error of our approximated solution at the end of our time step. For given initial conditions, we already know by theorem 3.2 that the error estimate

$$\|u - u_k\|_X \leq \frac{\|b\|}{\gamma_k} \inf_{z_k \in X_k} \|u - z_k\|_X = \frac{\|b\|}{\gamma_k} \|\mathcal{I}_k u - u_k\|_X$$

holds, where $\mathcal{I}_k$ denotes the polynomial interpolation operator of order k. It is well known that the polynomial interpolation error is of order k.

Furthermore, as the continuous problem is well-posed we get by theorem 3.1 that the dual problem to 3.5 is well-posed. Therefore, we can find a $w \in Y$ such that

$$b(x, w) = f(x) \quad \forall x \in X,$$

with $f(x) = x(t_1)$ being the point evaluation functional at the end point t_1. Together with the fact that

$$b(u - u_k, v_k) = b(u, v_k) - b(u_k, v_k) = l(v_k) - l(v_k) = 0$$

for all $v_k \in Y_k$ we get

$$|u(t_1) - u_k(t_1)| = |b(u - u_k, w)| = |b(u - u_k, w - \mathcal{I}_k w)| \leq \|b\| \|u - u_k\|_X \|w - \mathcal{I}_k w\|_Y = \mathcal{O}(t_1^{2k})$$

and therefore, the numerical scheme is of order $2k$. [Bab72; Bet16; DG25]

3.8 Formulation of a Second-Order Integrator

In this section, we want to show how to construct such a time stepping scheme for the case where we consider polynomials of degree at most one.

For the construction, we will come back to the larger formulation

$$\begin{cases} q \in \mathcal{P}^1([0, t_1], \mathbb{R}^d), q(0) = q_0 \\ \hat{p}_1 \in \mathbb{R}^d, \\ \int_0^{t_1} (\dot{q} - v) \cdot \delta p \; dt, \quad \forall \delta p \in P^0([0, t_1], \mathbb{R}^d) \\ \int_0^{t_1} (p - Mv) \cdot \delta v, \quad \forall \delta v \in P^0([0, t_1], \mathbb{R}^d) \\ \int_0^{t_1} -Kq \cdot \delta q + p \cdot \dot{\delta q} \; dt - \hat{p}_1 \cdot \delta q(t_1) = -\hat{p}_0 \cdot \delta q(0), \quad \forall \delta q \in \mathcal{P}^1([0, t_1], \mathbb{R}^d). \end{cases} \tag{3.22}$$

The first integral expression can be integrated directly to

$$\int_0^{t_1} (\dot{q} - v) \cdot \delta p \; dt = t_1 (\frac{q(t_1) - q_0}{t_1} - v) \cdot \delta p = 0 \iff q(t_1) - q_0 - v = 0. \tag{3.23}$$

The second integral expression can be integrated directly as well and yields

$$\int_0^{t_1} (p - Mv) \cdot \delta v \; dt = t_1 (p - Mv) \cdot \delta v = 0 \iff p - Mv = 0. \tag{3.24}$$

For the third equation, we have the quadratic Term $Kq \cdot \delta q$ for which we will use a trapezoidal rule to get

$$\int_0^{t_1} -Kq \cdot \delta q + p \cdot \dot{\delta q} \; dt - \hat{p}_1 \cdot \delta q(t_1) + \hat{p}_0 \cdot \delta q(0) =$$
$$\int_0^{t_1} -Kq \cdot \delta q \; dt + p \cdot (\delta q(t_1) - \delta q(0)) - \hat{p}_1 \cdot \delta q(t_1) + \hat{p}_0 \cdot \delta q(0) \approx$$
$$-\frac{t_1}{2}(Kq(t_1) \cdot \delta q(t_1) + Kq(0) \cdot \delta q(0)) + p \cdot (\delta q(t_1) - \delta q(0)) - \hat{p}_1 \cdot \delta q(t_1) + \hat{p}_0 \cdot \delta q(0)$$

As $\delta q \in \mathcal{P}^1([0, t_1], \mathbb{R}^d)$ is arbitrary it follows that $\delta q(0),\ \delta q(t_1) \in \mathbb{R}^d$ are arbitrary and we can split our equation into two parts, one for $\delta q(0)$ and one for $\delta q(t_1)$. We can therefore solve the two equations separately

$$p - \hat{p}_0 - \frac{t_1}{2} Kq(0) = 0 \tag{3.25}$$

$$\hat{p}_1 - p - \frac{t_1}{2} Kq(t_1) = 0 \tag{3.26}$$

Putting all equations together results in a system of linear equations, where we have $4d$ unknowns coming from $q(t_1), \hat{p}_1, p$, and v, as well as $4d$ equations

$$q(t_1) - q_0 - v = 0 \tag{3.27}$$
$$p - Mv = 0 \tag{3.28}$$
$$p - \hat{p}_0 - \frac{t_1}{2} Kq(0) = 0 \tag{3.29}$$
$$\hat{p}_1 - p - \frac{t_1}{2} Kq(t_1) = 0. \tag{3.30}$$

It is important to note that although we have used the trapezoidal rule, which is not exact for $Kq \cdot \delta q$ our time stepping scheme is still symplectic and has an accuracy of order two, as the trapezoidal rule itself is of order three. To see that symplecticity is not impaired, one can follow again exactly the steps in section 3.6.

4 Formulation of Rigid Body Motion

In this chapter we want to develop a numerical scheme for solving the motion of a motion of a rigid body. As will be shown in this chapter, the configuration space of a mechanical system must not be a linear vector space, but instead be a Lie Group. The theory of HP mechanics in 2 can be generalized, simply by replacing the configuration vector space $\mathcal{Q}$ by a Lie Group. However, for the extensive proof we are referencing on [Bou07]. The only important theorem we will us is the following

Theorem 4.1. *Let $G = \{x \in \mathbb{R}^d \mid \ \varphi(x) = 0\}$ with $\varphi : \mathbb{R}^d \to \mathbb{R}$ and rank$(\nabla\varphi(x)) = d$ for all $x \in G$. Let L be a Lagrangian on TG with continuous partial derivatives of second order with respect to q and v, then the following are equivalent*

The HP principle,

- *using generalized coordinates*

$$\delta \int_0^T L(g, v) + \langle p, \dot{g} - v\rangle \ dt = 0 \tag{4.1}$$

  holds where $(g(t), v(t), p(t)) \in TG \oplus T^*G$ can be varied arbitrarily and independently.

- *using constrained coordinates*

$$\delta \int_a^b L(g, v) + \langle p, \dot{g} - v\rangle + \lambda\varphi(g) \ dt = 0$$

  holds where $(g(t), v(t), p(t)) \in T\mathbb{R}^{12} \oplus T^*\mathbb{R}^{12}$ can be varied arbitrarily and independently.

This illustrates that the HP principle in terms of generalized coordinates is equivalent to the HP principle using constrained coordinates and the Lagrange multiplier method. Using constrained ccordinates the HP equations change to

$$\dot{q} = v \tag{4.2}$$

$$\dot{p} = \frac{\partial L}{\partial q}(q, v) + \lambda\nabla\varphi(q) \tag{4.3}$$

$$p = \frac{\partial L}{\partial v}(q, v) \tag{4.4}$$

Working with constrained coordinates we can define a variational formulation similar to 3.4 as

$$\begin{cases} q \in H^1([t_n, t_n+h], \mathbb{R}^{12}), q(0) = q_0 \in G \\ \hat{p}_1 \in T_{q_0}G, \\ \int_{t_n}^{t_n+h} -Kq \cdot \delta q + M\dot{q} \cdot \dot{\delta q} + \lambda \nabla\varphi(q)\delta q \; dt - \hat{p}_1 \cdot \delta q(t_n+h) = -\hat{p}_0 \cdot \delta q(0), \\ \forall \delta q \in H^1([t_n, t_n+h], \mathbb{R}^d). \\ \int_{t_n}^{t_n+h} \delta\lambda\varphi(q) \; dt, \quad \forall \delta\lambda \in H^1([t_n, t_n+h], \mathbb{R}^d). \end{cases} \tag{4.5}$$

The goal of the subsequent chapter is to define a configuration space for mechanical systems and generalize the numerical method we have developed on vector spaces to Lie Groups. [Bou07; HHIL06]

4.1 The Configuration Space of a Rigid Body in $\mathbb{R}^3$

The position of a rigid body in an inertial reference frame is represented by a vector $x \in \mathbb{R}^3$, meaning it belongs to a linear space. In addition to the position, there are three more degrees of freedom that describe the orientation of the rigid body. However, these orientation variables cannot, in general, be globally represented as elements of a three-dimensional linear space.

In engineering, small deviations from a nominal configuration are often described using three rotation angles such as Euler angles or Bryant angles. These representations, however, are prone to singularities when dealing with large rotations.

To avoid these singularities, alternative representations are used—such as Euler parameters, or the rotation matrix

$$R \in \mathrm{SO}(3) := \{R \in \mathbb{R}^{3\times 3} : RR^\top = I_3, \; \det(R) = +1\}.$$

The set SO(3) forms a three-dimensional differentiable manifold embedded in $\mathbb{R}^{3\times 3}$.

To describe the configuration of a rigid body, this rotation matrix can be combined with the position vector $x \in \mathbb{R}^3$, resulting in a six-dimensional quantity $q := (x, R)$. This can be done using the direct product group $G = \mathbb{R}^3 \times \mathrm{SO}(3)$, where the group operation $\circ$ is defined as

$$(x, R) \circ (\tilde{x}, \tilde{R}) = (x + \tilde{x}, R\tilde{R}).$$

Consider an element $R \in SO(3)$. The variations of RR^T and R^TR satisfy the following relations:

$$\delta(R^TR) = \delta R^T R + R^T \delta R = W + W^T = 0,$$

and

$$\delta(RR^T) = \delta R R^T + R\delta R^T = W + W^T = 0,$$

from which it follows that W and W^T are skew-symmetric matrices. Moreover, we have that

$$\delta R = WR = RW, \tag{4.6}$$

where clearly δR belongs to the tangent space to $SO(3)$ at R, denoted by $T_R SO(3)$.

The tangent space at the identity element I forms the Lie algebra of $SO(3)$, which is denoted by

$$\mathfrak{so}(3) = T_I SO(3).$$

From equation 4.6, it follows that $\mathfrak{so}(3)$ corresponds to the linear space of skew-symmetric tensors of the form

$$\tilde{\omega} = \begin{bmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{bmatrix} = \text{skew}[\omega]. \tag{4.7}$$

Since the Lie algebra $\mathfrak{so}(3)$ is isomorphic to $\mathbb{R}^3$, every element $W \in \mathfrak{so}(3)$ can be represented by a vector

$$\omega = (\omega_1, \omega_2, \omega_3) \in \mathbb{R}^3.$$

Consider now a rigid body B and attach to it a coordinate frame that moves with the body. Assuming the origin of this body-fixed frame remains stationary, Euler's theorem tells us that every instantaneous motion of the body is a rotation about some axis. Let ω be a vector indicating the direction of this rotation axis, where the magnitude $\|\omega\|$ represents the angular velocity.

For a material point x within the body, the velocity induced by this rotation is given by the cross product

$$v = \omega \times x = \begin{pmatrix} \omega_2 x_3 - \omega_3 x_2 \\ \omega_3 x_1 - \omega_1 x_3 \\ \omega_1 x_2 - \omega_2 x_1 \end{pmatrix} = \begin{pmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \tag{5.1}$$

This velocity vector is orthogonal to both ω and x, and its magnitude is given by $\|\omega\| \cdot \|x\| \cdot \sin\gamma$, where γ is the angle between ω and x. To compute the kinetic energy, we integrate the energy density over the entire body

$$T = \frac{1}{2} \int_B \|\omega \times x\|^2 \, \mathrm{d}m \tag{5.2}$$

Expanding the squared norm of the cross product yields

$$T = \frac{1}{2} \int_B \left[(\omega_2 x_3 - \omega_3 x_2)^2 + (\omega_3 x_1 - \omega_1 x_3)^2 + (\omega_1 x_2 - \omega_2 x_1)^2 \right] \, \mathrm{d}m$$

After distributing the terms, the kinetic energy takes the matrix form

$$T = \frac{1}{2} \omega^\top \boldsymbol{\Theta} \omega, \quad \text{where} \quad \Theta_{ii} = \int_B (x_k^2 + x_l^2) \, \mathrm{d}m, \quad \Theta_{ik} = -\int_B x_i x_k \, \mathrm{d}m \quad (i \neq k) \tag{5.3}$$

The indices k and l in the expression for Θ_{ii} denote the two indices different from i.

Euler, through elaborate trigonometric manipulations, demonstrated the existence of special directions—so-called principal axes—in which the kinetic energy simplifies to a diagonal expression

$$T = \frac{1}{2} I_1 \omega_1^2 + \frac{1}{2} I_2 \omega_2^2 + \frac{1}{2} I_3 \omega_3^2 \tag{5.4}$$

This result was the earliest known diagonalization of a 3×3 quadratic form. Furthermore for a diagonal matrix $D = \text{diag}(d_1, d_2, d_3)$ and any skew symmetric matrix W it holds that $\text{tr}(WDW^T) = (d_2 + d_3)w_1^2 + (d_3 + d_1)w_2^2 + (d_2 + d_1)w_3^2$. If we now choose $d_k = \int_B x_k^2 dm$, together with the identity that $\dot{R} = RW$ we can reformulate the kinetic energy of the body as

$$T = \frac{1}{2}\text{tr}(WDW^T) = \frac{1}{2}\text{tr}(\dot{R}D\dot{R}^T).$$

With this expression of the kinetic energy we can go on and show that for the conjugate momenta then holds

$$P = \frac{\partial T}{\partial \dot{R}} = \dot{R}D = RWD.$$

This means that the conjugate momenta as well as the velocity of a rigid body can be written as the product of an orthogonal and skew symmetric matrix.[Bet16; HHIL06; Hol11; HAG24]

4.2 A Second-Order Lie Group Time Stepping Scheme

We now want to show that the scheme developed 3.8 can be extended to the Lie Group $\mathbb{R}^3 \times \text{SO}(3)$. If we use the same approach for discretizing 4.5 as in 3.8 we get the following equations

$$\begin{aligned} hv_n &= q_{n+1} - q_n \\ p_n &= Mv_n \\ p_n &= \hat{p}_n + \frac{h}{2}Kq_n - \lambda_n \frac{h}{2}\nabla\varphi(q_n) \\ 0 &= g(q_{n+1}) \\ \hat{p}_{n+1} &= p_n + \frac{h}{2}Kq_{n+1} - \lambda_{n+1}\frac{h}{2}\nabla\varphi(q_{n+1}) \end{aligned} \tag{4.8}$$

As we can see the key difference to the vector space scheme is that we are enforcing $q(t_1 + h)$ to be an element of the Lie Group G. However we are interested in propagating $(q_0, \hat{p}_0) \in T^*G$ to $(q(h), \hat{p}_1) \in T^*G$ with

$$T^*G = \{(q,p)| \;\; q \in G, p \in T_q^*G\} = \{(q,p)| \;\; \varphi(q) = 0, \; G(q)M^{-1}p = 0\}.$$

Therefore we have to replace the last line with a projection step

$$\begin{aligned} hv_n &= q_{n+1} - q_n \\ p_n &= Mv_n \\ p_n &= \hat{p}_n + \frac{h}{2}Kq_n - \lambda_n \frac{h}{2}G(q_n) \\ 0 &= g(q_{n+1}) \\ \hat{p}_{n+1} &= p_n + \frac{h}{2}Kq_{n+1} - \lambda_{n+1}\frac{h}{2}G(q_{n+1}) \\ 0 &= G(q_{n+1})M^{-1}\hat{p}_{n+1} \end{aligned} \tag{4.9}$$

to enforce $(q(t_1 + h), \hat{p}_1) \in T^*G$. Now let us discuss the properties of this method.

Theorem 4.2. *For $(q_n, \hat{p}_n) \in T^*G$ there exists a locally unique solution of the numerical method 4.9.*

Proof. At first we will show that the system defined by the first 4 equations has a locally unique solution $(q_{n+1}, v_n, p_n, \lambda_n)$ and then go on to the last two equations to show that then they also have a locally unique solution $(\hat{p}_{n+1}, \lambda_{n+1})$.

We can rewrite the fourth equation to

$$0 = g(q_{n+1}) = g(q_n) + \int_0^1 \nabla\varphi(q_n + \tau(q_{n+1} - q_n))(q_{n+1} - q_n)\, d\tau.$$

As our initial value q_n lies in our Lie-Group $g(q_n)$ evaluates to zero. Inserting the definition of q_{n+1} by the first equation of 4.9 and dividing by h together with the third equation yields the system $F(p_{n+1}, h\lambda_n, h) = 0$ with

$$F(p, \nu, h) = \begin{pmatrix} p - \hat{p}_n - \frac{h}{2} K q_n + \nu \frac{1}{2} \nabla\varphi(q_n) \\ \int_0^1 \nabla\varphi(q_n + \tau h M^{-1} p) M^{-1} p \, d\tau \end{pmatrix}$$

Since $(q_n, \hat{p}_n) \in T^*G$ we have $F(\hat{p}_n, 0, 0) = 0$ and

$$\frac{\partial F}{\partial(p, \nu)}(\hat{p}_n, 0, 0) = \begin{pmatrix} I & \frac{1}{2}\nabla\varphi(q_n)^T \\ \nabla\varphi(q_n) M^{-1} & 0 \end{pmatrix},$$

which is invertible because $\nabla\varphi(q_n)$ and M^{-1} are regular. Therefore, an application of the implicit function theorem proves, that a locally unique solution $(p_n, v_n, h\lambda_n)$ exists. The projection step now poses a nonlinear system for $\hat{p}_{n+1}$ and $h\lambda_{n+1}$, to which the implicit function theorem can be applied as well. □

Theorem 4.3. *The numerical method 4.9 is symmetric, symplectic, and convergent of order two*

Proof. The symmetry of the method follows directly if we add the consistency terms $\varphi(q_n) = 0$ and $\nabla\varphi(q_n) M^{-1} \hat{p}_n = 0$.

The implicit function theorem used in 4.2 yields the estimates

$$p_n = \hat{p}_n + \mathcal{O}(h), \quad v_n = M^{-1}\hat{p}_n + \mathcal{O}(h) \quad h\lambda_n = \mathcal{O}(h), \quad h\lambda_{n+1} = \mathcal{O}(h)$$

and therefore, together with the equations 4.9, give

$$q_{n+1} = q(t_{n+1}) + \mathcal{O}(h^2), \quad \hat{p}_{n+1} = p(t_{n+1}) - \nabla\varphi(q(t_{n+1}))\nu + \mathcal{O}(h^2)$$

with $(q(t), p(t))$ being the solution of 4.2-4.4 passing through $(q_n, \hat{p}_n)$ at t_n. If we now use the last equation of 4.9 we see that

$$0 = \nabla\varphi(q(t_{n+1})) M^{-1} p(t_{n+1}) + \nabla\varphi(q(t_{n+1})) M^{-1} \nabla\varphi(q(t_{n+1}))^T \nu + \mathcal{O}(h^2).$$

We know that $\nabla\varphi(q(t_{n+1})) M^{-1} p(t_{n+1}) = 0$ and $\nabla\varphi(q(t_{n+1})) M^{-1} \nabla\varphi(q(t_{n+1}))^T$ is invertible, meaning $\nu = \mathcal{O}$. Therefore the local error is of size $\mathcal{O}(h^2)$. To establish convergence, we

follow a classical argument. The numerical scheme defines a mapping $\Phi_h : T^*G \to T^*G$. Consider again the exact solution $(q(t), p(t))$ of system 4.2-4.4, which passes through the numerical values $(q_n, \hat{p}_n) \in T^*G$ at time t_n. We estimate the local error introduced at each time step and analyze how these errors propagate over time. By summing the accumulated errors across the steps, we obtain a bound on the global error. Specifically, we find that $\hat{p}_n - p(t_n) = \mathcal{O}(h)$ and $q_n - q(t_n) = \mathcal{O}(h)$, provided that the final time $t_n = nh$ remains bounded. This approach follows standard convergence analysis. Since symmetric methods are always of even order, convergence of order two follows.

To show that 4.9 is symplectic we start again by showing that the mapping $(q_n, \hat{p}_n) \mapsto (p_n, q_{n+1})$ is symplectic. We therefore consider λ as a function of $(q_n, \hat{p}_n)$ to calculate the derivative

$$\begin{pmatrix} I & 0 \\ -\frac{h}{2}M^{-1} & I \end{pmatrix} \cdot \frac{\partial(p_n, q_{n+1})}{\partial(\hat{p}_n, q_n)} = \begin{pmatrix} I - \frac{h}{2}\lambda_p \nabla\varphi & \frac{h}{2}K - \frac{h}{2}\lambda g_{qq} - \frac{h}{2}\lambda_q \nabla\varphi \\ 0 & I \end{pmatrix} \tag{4.10}$$

and therefore

$$\left(\frac{\partial(p_n, q_{n+1})}{\partial(\hat{p}_n, q_n)}\right)^T J \left(\frac{\partial(p_n, q_{n+1})}{\partial(\hat{p}_n, q_n)}\right) - J = \begin{pmatrix} 0 & I - \frac{h}{2}\lambda_p \nabla\varphi \\ -I + \frac{h}{2}\lambda_p \nabla\varphi & \frac{h}{2}(\lambda_q \nabla\varphi - \lambda_q \nabla\varphi) \end{pmatrix}. \tag{4.11}$$

Now test the relation with two tangent vectors $\xi_1 \in T_{(\hat{p}_n, q_n)}T^*G$ and $\xi_2 \in T_{(\hat{p}_n, q_n)}T^*G$. Decomposing $\xi = (\xi_{\hat{p}_n}, \xi_{q_n})$, and using the fact that $\nabla\varphi(q_n)\xi_{q_n,j} = 0$ for $j = 1, 2$, the resulting expression simplifies to

$$\xi_1^{\mathrm{T}} J \xi_2.$$

This confirms that the transformation $(q_n, \hat{p}_n) \mapsto (q_{n+1}, p_n)$ is symplectic. The proof for the projection step $(q_{n+1}, p_n) \mapsto (q_{n+1}, \hat{p}_{n+1})$ works analogous to the one before. □

For our configuration space $G = \mathrm{SO}(3) \times \mathbb{R}^3$ we can further simplify 4.9. We start by introducing $G = \{(x_Q, Q \in \mathbb{R}^{12} \mid \varphi(Q) = Q^T Q - I = 0\}$, where the constraint $\varphi(Q) = 0$ calls for 6 Lagrange multipliers, since $Q^T Q$ is symmetric. Furthermore, for $(x_Q, Q) \in G$ we can express an any element $(x_V, V) \in T_{(x_Q,Q)}G$ as (x_V, QW) where W is a skew symmetric matrix. Of course the same holds for any element in $T^*_{(x_Q,Q)}G$. Making use of that we can rewrite the equations of 4.9 for the rotational part as

$$\begin{aligned} hV_n &= Q_{n+1} - Q_n \\ P_n &= MV_n \\ P_n &= \hat{P}_n + \frac{h}{2}KQ_n - \Lambda_n \frac{h}{2}Q_n = Q_n \hat{W}_n + \frac{h}{2}KQ_n - \Lambda_n \frac{h}{2}Q_n \\ 0 &= g(Q_{n+1}) \\ Q_{n+1}\hat{W}_{n+1} &= \hat{P}_{n+1} = P_n + \frac{h}{2}Kq_{n+1} - \Lambda_{n+1}\frac{h}{2}Q_{n+1} \\ 0 &= G(Q_{n+1})M^{-1}Q_{n+1}\hat{W}_{n+1} \end{aligned} \tag{4.12}$$

We multiply the third equation with Q_n^T, the fifth equation with Q_{n+1}^T and take only the skew symmetric part of the equations to eliminate the symmetric lagrange multiplier matrices Λ_n

and Λ_{n+1}. Furthermore the last equation equates to zero as $Q_{n+1}\hat{W}_{n+1} \in T_{Q^*_{n+1}}G$ yielding

$$\begin{aligned} hV_n &= Q_{n+1} - Q_n \\ P_n &= MV_n \\ skew(\hat{W}_n) &= skew(Q_n^T P_n - \frac{h}{2} Q_n^T K Q_n) \\ 0 &= g(Q_{n+1}) \\ skew(\hat{W}_{n+1}) &= skew(Q_{n+1}^T P_n + \frac{h}{2} Q_{n+1}^T K Q_{n+1}) \end{aligned} \tag{4.13}$$

The equations for the translational part stay the same as in the vector space. The full times stepping scheme for our rigid body problem therefore states

$$\begin{aligned} hx_{V_n} &= x_{Q_{n+1}} - x_{Q_n} \\ hV_n &= Q_{n+1} - Q_n \\ \\ x_{P_n} &= mx_{V_n} \\ P_n &= MV_n \\ \\ x_{P_n} &= x_{\hat{P}_n} + \frac{h}{2} K x_{Q_n} \\ skew(\hat{W}_n) &= skew(Q_n^T P_n - \frac{h}{2} Q_n^T K Q_n) \\ \\ x_{\hat{P}_{n+1}} &= x_{P_n} + \frac{h}{2} K x_{Q_{n+1}} \\ skew(\hat{W}_{n+1}) &= skew(Q_{n+1}^T P_n + \frac{h}{2} Q_{n+1}^T K Q_{n+1}) \\ \\ 0 &= g(Q_{n+1}) \end{aligned} \tag{4.14}$$

In the course of this work, the time stepping scheme 4.13 was implemented in C++. To support general multibody systems composed of interconnected rigid bodies, the implementation allows for arbitrary system topologies in which components are connected via beams and springs. Furthermore, a custom automatic differentiation framework was developed specifically for this purpose, enabling the consistent and accurate computation of derivatives required by the Newton method to solve the nonlinear systems arising at each time step. The implementation is based on a self-developed basic linear algebra library that uses high-performance computing concepts and expression templates for efficient execution. Thanks to its modular design, the implementation supports flexible definitions of system components and constraint types, providing a robust and extensible platform for simulating complex multibody dynamics governed by nonlinear interactions and geometric constraints. The full implementation can be found on github. [Bet16; HHIL06; Hol11; HAG24; KBS23b]

4.3 Numerical Tests for the Heavy Top Benchmark Problem

In this section, we numerically investigate the convergence behavior of our Lie group time-stepping scheme. The method is applied to the equations of motion of the heavy

top benchmark problem formulated on the configuration spaces $G = \mathrm{SO}(3) \times \mathbb{R}^3$. In the benchmark problem, the top rotates about a fixed point as shown in Fig. 1. Therefore,

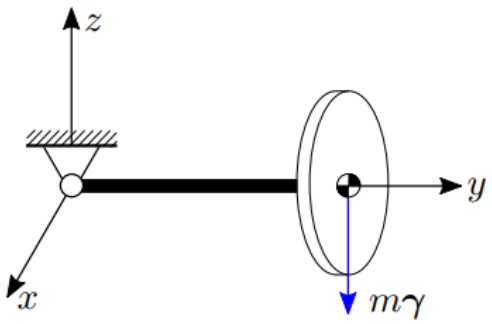

Figure 1: Benchmark problem Heavy top, see e.g. [BC10]

the configuration variables (x, R) are subject to holonomic constraints $x = RX$. We consider an initial configuration defined by $R(0) = I_3$ with an angular velocity $\boldsymbol{\omega}(0) = (0,\ 150,\ -4.61538)^\top$. All other Model parameters and initial conditions are summarized in Table 1. In the numerical experiments, the integrator is initialized with the starting values

$$q_0 := q(t_0), \quad \hat{p}_0 = Mv_0 = Mv(t_0)$$

and we are comparing solutions computed for step sizes

$$h = 1.25 \times 10^{-4},\ 2.5 \times 10^{-4},\ 5.0 \times 10^{-4},\ \ldots,\ 4.0 \times 10^{-3}$$

against a reference solution computed with the very fine step size $h = 2.5 \times 10^{-5}$, shown below in Fig 2. We are interested in the asymptotic behavior of the global errors in q_n, v_n, and λ_n as the time step size h goes to zero. The quantities shown in Fig. 3 are the maximum of the relative errors $\max_n \|e_n^{(\cdot)}\|/\|(\cdot)_n\|$ over the time interval $[t_0, t_{\text{end}}] = [0, 1]$ as well as the absolute error $\max_n \|e_n^{(\cdot)}\|$.

On a double-logarithmic scale, the error plots for both q_n and v_n appear as straight lines with slope 2, indicating second-order convergence with respect to the variables q and v. [Bet16; HAG24]

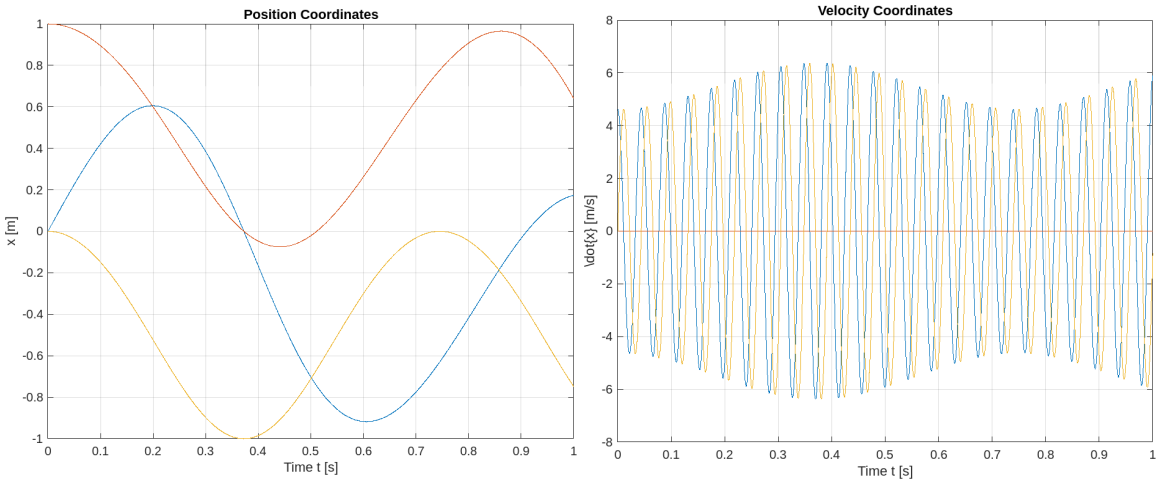

Figure 2: Heavy top benchmark: Reference solution

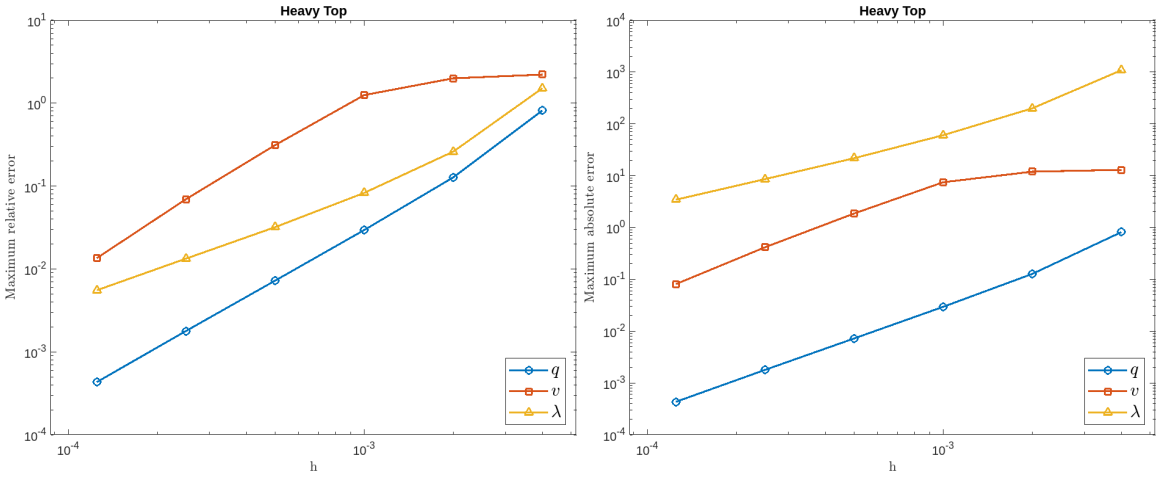

Figure 3: Global error of integrator versus h for $t \in [0, 1]$. Left plot relative error, right plot absolute error

Parameter	Value	Notes
m [kg]	15	mass of the heavy top
J_{xx} [kg·m^2]	0.234375	moment of inertia w.r.t. COM for x axes
J_{yy} [kg·m^2]	0.468750	moment of inertia w.r.t. COM for y axes
J_{zz} [kg·m^2]	0.234375	moment of inertia w.r.t. COM for z axes
$\mathbf{x}$ [m]	$[0.0,\ 1.0,\ 0.0]$	position vector of the COM w.r.t the fixing point represented in the body-attached frame
$\boldsymbol{\gamma}$ [m/s^2]	$[0.0,\ 0.0,\ -9.81]$	vector of gravity acceleration
$\boldsymbol{\omega}_0$ [rad/s]	$[0.0,\ 150.0,\ -4.61538]$	initial angular velocity
$\boldsymbol{\psi}_0$ [rad]	$[0.0,\ 0.0,\ 0.0]$	initial orientation (rotation vector)

Table 1: Model parameters and initial conditions of the Heavy Top, see [HAG24]

5 Outlook

In this thesis we have established a mathematically rigorous and computational approach to time discretization for mechanical systems governed by the Hamilton–Pontryagin principle. By employing a variational discontinuous Petrov–Galerkin formulation, we developed a discretization scheme that preserves the geometric structure of the underlying continuous system on vector spaces.

A central outcome of this thesis is the development of a general framework for constructing symplectic time stepping schemes of order $2k$ on vector spaces. While the framework is fully established in the setting of vector spaces, its extension to configuration spaces being Lie groups remains an open problem. In particular, although we have shown that the second-order time stepping scheme preserves symplecticity in the Lie group setting, a general proof for higher-order schemes is still lacking. Thus, the adaptation of the full $2k$-order construction to Lie groups constitutes an important direction for future research.

Nevertheless, the successful derivation of a second-order symplectic integrator on Lie groups demonstrates the potential applicability of the method beyond vector spaces and provides a promising first step towards a unified theory of geometric time integration on manifolds.

Bibliography

[Bab71] I. Babuška. Error-bounds for finite element method. *Numerische Mathematik*, 16(4):322–333, 1971.

[Bab72] I. Babuska. Survey lectures on the mathematical foundations of the finite element method. *The Mathematical Foundations of the Finite Element Method with Applicaions to Partial Differential Equations*:3–359, 1972.

[BC10] O. Brüls and A. Cardona. On the use of Lie group time integrators in multibody dynamics. *Journal of Computational and Nonlinear Dynamics*, 5(3):1–13, 2010.

[Bet16] P. Betsch. *Structure-preserving integrators in nonlinear structural dynamics and flexible multibody dynamics*. Springer, 2016.

[Bou07] N. M. Bou-Rabee. *Hamilton-Pontryagin integrators on Lie groups*. PhD thesis, California Institute of Technology, 2007.

[Bre74] F. Brezzi. On the existence, uniqueness and approximation of saddle-point problems arising from Lagrangian multipliers. *Publications des séminaires de mathématiques et informatique de Rennes*, (S4):1–26, 1974.

[DG25] L. Demkowicz and J. Gopalakrishnan. The discontinuous Petrov–Galerkin method. *Acta Numerica*, 34:293–384, 2025.

[EG+21] A. Ern, J.-L. Guermond, et al. *Finite elements II*. Springer, 2021.

[HAG24] S. Holzinger, M. Arnold, and J. Gerstmayr. Evaluation and implementation of Lie group integration methods for rigid multibody systems. *Multibody System Dynamics*:1–34, 2024.

[HHIL06] E. Hairer, M. Hochbruck, A. Iserles, and C. Lubich. Geometric numerical integration. *Oberwolfach Reports*, 3(1):805–882, 2006.

[Hol11] D. D. Holm. *Geometric mechanics-Part II: Rotating, translating and rolling*. World Scientific, 2011.

[KBS23a] P. L. Kinon, P. Betsch, and S. Schneider. Structure-preserving integrators based on a new variational principle for constrained mechanical systems. *Nonlinear Dynamics*, 111(15):14231–14261, 2023.

[KBS23b] P. L. Kinon, P. Betsch, and S. Schneider. The GGL variational principle for constrained mechanical systems. *Multibody System Dynamics*, 57(3):211–236, 2023.

[LR04] B. Leimkuhler and S. Reich. *Simulating hamiltonian dynamics*, number 14. Cambridge university press, 2004.

[Neč62] J. Nečas. Sur une méthode pour résoudre les équations aux dérivées partielles du type elliptique, voisine de la variationnelle. *Annali della Scuola Normale Superiore di Pisa-Scienze Fisiche e Matematiche*, 16(4):305–326, 1962.

[YM06a] H. Yoshimura and J. E. Marsden. Dirac structures in Lagrangian mechanics part I: implicit Lagrangian systems. *Journal of Geometry and Physics*, 57(1):133–156, 2006.

[YM06b] H. Yoshimura and J. E. Marsden. Dirac structures in Lagrangian mechanics Part II: Variational structures. *Journal of Geometry and Physics*, 57(1):209–250, 2006.