Cathrine Thorsen Becker

An introduction to Lie group methods for rigid body systems

Bachelor's project in Mathematical Sciences
Supervisor: Elena Celledoni
May 2021

Bachelor's project

NTNU
Norwegian University of Science and Technology
Faculty of Information Technology and Electrical Engineering
Department of Mathematical Sciences

Kunnskap for en bedre verden

Cathrine Thorsen Becker

An introduction to Lie group methods for rigid body systems

Bachelor's project in Mathematical Sciences
Supervisor: Elena Celledoni
May 2021

Norwegian University of Science and Technology
Faculty of Information Technology and Electrical Engineering
Department of Mathematical Sciences

NTNU
Norwegian University of
Science and Technology

An introduction to Lie group methods for rigid body systems

Cathrine Thorsen Becker

May 2021

Abstract

This thesis is a short introduction to Lie group methods. Some relevant background material of Lie group theory is presented and we give an introduction to Lie-Euler, Runge-Kutta Munthe-Kaas and commutator free methods. Two mechanical systems are studied; the free rigid body and the heavy top. We discuss how the different methods conserve the structures of the manifolds, such as conservation of angular momentum and energy, and the advantages the Lie group integrators have over a classical numerical method.

Contents

1 Introduction

Integration methods on manifolds have for a long time been studied by mathematicians and the study of Lie group methods has over the last 30 years accelerated. They are increasingly being used not only in pure mathematics, but in a wide range of fields such as robotics, computer sciences, mechanical engineering and in particular, physics. Problems in geometric mechanics - for instance those presented in the book by Leok [9] - provide formulations of differential equations where a Lie group integrator is a natural choice. This is because classical numerical methods might not make sense once we are working on vector fields where normal operations like sum and multiplication are not well defined.

This thesis is an elementary introduction to Lie group integrators and is divided into two parts. It is assumed the reader is familiar with some elementary group theory, as well as basic knowledge of numerical mathematics and analysis.

The first part introduces relevant background material to understand the methods and presents examples along the way. This part is mostly theoretical, introducing new definitions and terminology, such as manifolds, vector fields and Lie groups. The examples closely follow the theoretical material and for the most part focuses on the special Euclidean group, $SE(3)$. Additionally, a few Lie group integrators are presented from a theoretical perspective.

The second part focuses on two specific models where the Lie group integrators can be applied; the free rigid body and the heavy top. The Lie group $SE(3)$ is particularly relevant for the heavy top. The results from implementing the integrators for these models are discussed and compared to a classical numerical method, the modified Euler method.

Most of the theoretical material is based on the books by Marsden and Ratiu [3] and Olver [1], and the books by Lee [2, 5] are good supplemental reading for the theoretical part of this thesis. The integration methods and implementations are based on a variety of papers [10, 15, 16], and in particular the work of Munthe-Kaas [12].

2 Manifolds and vector fields

This section is based on the definitions and properties introduced in the first chapter of [1], with some exceptions.

2.1 Manifolds

To understand Lie group integrators we need the notion of manifolds. Formally, we define manifolds in the following way.

Definition 2.1. An m-dimensional manifold $\mathcal{M}$ is a topological space together with a collection of charts (U, φ) such that $U \subset \mathcal{M}$ is an open subset of $\mathcal{M}$ and the map $\varphi : U \to V \subset \mathbb{R}^m$ is continuous, one-to-one and onto. Here, $V \subset \mathbb{R}^m$ is an open, connected subset of $\mathbb{R}^m$. We say $\mathcal{M}$ is smooth if the map $\varphi_{\alpha\beta} = \varphi_\alpha \circ \varphi_\beta^{-1}$ is smooth where it is defined.

Intuitively, we can think about a manifold $\mathcal{M}$ as a space that resembles Euclidean space near every point. It is clear to see that $\mathbb{R}^m$ itself is an m-dimensional manifold, while less trivial is the $m-1$-dimensional manifold of the unit sphere $\mathcal{S}^{m-1} := \left\{\mathbf{x} \in \mathbb{R}^m \mid \sum_{i=1}^m x_i^2 = 1\right\}$. To demonstrate this we look at the *stereographic projection*, which is a map $\sigma : \mathcal{S}^{m-1} \backslash \{N\} \to \mathbb{R}^{m-1}$ where $N = (0, \ldots, 0, 1) \in \mathbb{R}^m$ is the *north pole* in $\mathcal{S}^{m-1}$. This is a widely used example of representing $\mathcal{S}^{m-1}$ as a manifold. It is described in detail in [2] for the sphere, and presented here for the circle, focusing on an intuitive presentation.

Example 2.1. Take a circle $\mathcal{S}$ in $\mathbb{R}^2$ as in figure 1. Now, create a map $\varphi : \mathcal{S} \to \mathbb{R}$ by drawing a line from the point $(0, 1)$ through every point on the circle.

These lines will intersect the line at each point $x \in \mathbb{R}$ and the map φ is onto. However, $(0, 0)$ and $(0, 1)$ will map to the same point on the line, and hence the function is not one-to-one. This is just one way of attempting to create

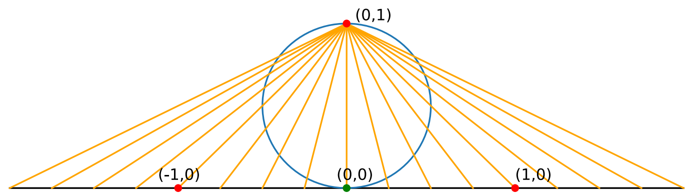


Figure 1: Points on the circle mapped to points on the line.

such a map, but in fact there exists no continuous bijection from $\mathcal{S}$ to $\mathbb{R}$. Because of this, the circle is a manifold needing at least two charts; e.g. one for $(\mathcal{S} \setminus (0,0))$ and one for $(0,0)$. By the same argument, the sphere in any dimension cannot be covered by a single chart.

Informally we define a *submanifold* to be a subset $\mathcal{N} \subset \mathcal{M}$ of a manifold $\mathcal{M}$ which is itself a manifold. More precisely, we state definition 1.12 in [1].

Definition 2.2. A smooth (analytic) n-dimensional *immersed submanifold* of a manifold M is a subset $N \subset M$ parametrized by a smooth (analytic), one-to-one map $F : \widetilde{N} \to N \subset M$, whose domain $\widetilde{N}$, the *parameter space*, is a smooth (analytic) n-dimensional manifold, and such that F is everywhere regular, of maximal rank n.

Additionally, we call a manifold $\mathcal{M}$ *differentiable* if there is a collection of charts such that each point $m \in \mathcal{M}$ is a member of at least one chart, and $\mathcal{M}$ is a union of compatible charts.

With these definitions we can now begin to introduce some key properties, namely the properties of vector fields on manifolds. To obtain these, we need to know about tangent vectors and tangent spaces.

2.2 Tangent spaces and vector fields

Definition 2.3. A tangent vector at a point $x \in \mathcal{M}$ is the tangent to a smooth curve $\phi(t) \in \mathcal{M}$ passing through m. Take $x \in \mathcal{M}$ to be such that $\phi(0) = x$, then

$$\mathbf{v}|_x := \frac{d}{dt}\phi(t)\bigg|_{t=0}$$

is the tangent vector at x in local coordinates.

With this definition we can imagine that given a point $x \in \mathcal{M}$, there exists several curves in $\mathcal{M}$ passing through the point. This results in different tangent vectors, and sets the foundation for tangent spaces.

Definition 2.4. The tangent space to an m-dimensional manifold $\mathcal{M}$ at a point $x \in \mathcal{M}$, denoted $T_x\mathcal{M}$, is the m-dimensional vector space formed by the collection of the tangent vectors at x.

Similarly, we define the *tangent bundle* to be the union of all tangent spaces, i.e. $T\mathcal{M} = \cup_{x\in\mathcal{M}} T_x\mathcal{M}$. Furthermore, we define vector fields as they are defined in [1].

Definition 2.5. A vector field on a manifold $\mathcal{M}$ is a section of the tangent bundle on $\mathcal{M}$, i.e. a smoothly varying assignment of tangent vectors $\mathbf{v} : \mathcal{M} \to T\mathcal{M}$ such that $\mathbf{v}(x) = \mathbf{v}|_x \in T_x\mathcal{M}$. We can express this in local coordinates as

$$\mathbf{v}(x) = \sum_{i=1}^{m} \xi^i(x)\frac{\partial}{\partial x^i}$$

where $\xi^i(x)$ are smooth functions and $\frac{\partial}{\partial x^i}$ denote a basis of the tangent space $T_x\mathcal{M}$.

2.3 Frame vector fields

We define a few new properties as introduced by Crouch and Grossman in [10].

Definition 2.6. Let $E = \{\mathcal{E}_1, \dots, \mathcal{E}_d\}$ be a set of vector fields on a manifold $\mathcal{M}$ of dimension $m \leq d$. Then E is a set of frame vector fields if

$$T_x\mathcal{M} = \operatorname{span}\left\{ \mathcal{E}_1|_x, \dots \mathcal{E}_d|_x \right\}, \quad \forall x \in \mathcal{M}.$$

With this definition we can say that any vector field F on $\mathcal{M}$ can be expressed by these frame vector fields as

$$F(y) = \sum_{i=1}^{d} f_i(y)\mathcal{E}_i(y). \tag{1}$$

Furthermore, we introduce frozen vector fields.

Definition 2.7. The vector field F is called frozen at a point p, denoted F_p, if

$$F_p(x) = \sum_{i=1}^{d} f_i(p)\mathcal{E}_i(x).$$

2.4 The exponential and differential map

A curve $\phi(t) : \mathbb{R} \to \mathcal{M}$ when $\phi(t) = x$ is called an *integral curve* of the vector field $\mathbf{v}$ if the tangent to the curve at t coincides with the vector field at x, i.e. $\dot{\phi}(t) = \mathbf{v}(x)$. We can express this in local coordinates by

$$\frac{dx^i}{dt} = \xi^i(x), \quad x^i = \phi_i(t).$$

Additionally, if $\phi(t)$ is a maximal integral curve we write

$$\phi(t) = \exp(t\mathbf{v})x_0, \quad x_0 = \phi(0).$$

Here $\exp(t\mathbf{v})x_0$ is the *flow* generated by $\mathbf{v}$. This exponential mapping has the same properties as the usual exponential function, and we will study this in more detail in section 3.4.

The *differential* or *derivative* map is a map we are well familiar with, and we define it on manifolds.

Definition 2.8. Given a smooth map $F : \mathcal{M} \to \mathcal{N}$ between manifolds $\mathcal{M}$ and $\mathcal{N}$, the differential is given by $dF : T\mathcal{M} \to T\mathcal{N}$ such that

$$dF|_x : T_x\mathcal{M} \to T_{F(x)}\mathcal{N},$$

such that for any curve $\phi(t)$ with $\phi(t)|_{t=0} = x$ and $F(\phi(t))|_{t=0} = F(x)$ we have

$$dF|_x (\mathbf{v}_x) = \mathbf{w}_{F(x)}.$$

Here, $\mathbf{v}_x := \frac{d}{dt}\phi(t)\big|_{t=0}$ and $\mathbf{w}_{F(x)} := \frac{d}{dt}F(\phi(t))\big|_{t=0}$ are tangent vectors.

One can show that the differential is a linear map [1].

3 Lie groups and Lie algebras

All the definitions, properties and examples introduced in this section is based on chapter 9 of Marsden and Ratiu [3], unless otherwise stated. The results of the calculations in example 3.5, 3.6, 3.7 and 3.8 can be verified in chapter 14.7 in [3].

3.1 Lie groups

Now that we have seen the fundamental properties of manifolds and vector fields, we are ready to define a Lie group.

Definition 3.1. A Lie group G is a smooth manifold $\mathcal{M}$ with group structure such that the group multiplication

$$\mu : G \times G \to G, \quad (g, h) \mapsto gh$$

is a C^∞ map.

A Lie group can therefore be said to be a group that is also a differentiable manifold. Some important examples of Lie groups are the special orthogonal group $SO(n)$ and the special Euclidean group $SE(n)$. We will in particular use these two groups for $n = 3$ to implement our Lie group integrators in section 5 and 6, and we will begin to review some properties of $SE(3)$ now.

Example 3.1. The special Euclidean group $SE(3)$ can be defined as an extension of the special orthogonal group $SO(3)$[1]. We describe an element in $SE(3)$ as the pair (g, u), where $g \in SO(3)$ and $u \in \mathbb{R}^3$, and the group operation is given by

$$(g, u) \cdot (h, v) = (gh, gv + u)$$

[1] The Lie group $SO(3)$ and its properties are quite thoroughly described in chapter 9.2 by Marsden and Ratiu, and through examples of the rest of the chapter. We therefore refer the reader to [3] for more information on this.

for $(g, u), (h, v) \in SE(3)$. With this property, one can prove that this set has a group structure, and since it is viewed as a topological space it is a Lie group.

The Lie group $SO(3)$ can be identified as the group of rotations in 3 dimensions, allowing $SE(3)$ to include translations as well. The group $SE(3)$ then corresponds to the set of matrices of the form

$$A = \begin{bmatrix} Q & v \\ 0 & 1 \end{bmatrix}, \quad Q \in SO(3), v \in \mathbb{R}^3.$$

The group operation associated with this set is the usual matrix multiplication. We will use this identification of $SO(3)$ and $SE(3)$ in the following examples and in section 5 and 6.

We will continue to build on this example as we introduce new relevant definitions and results. Additionally, whenever we use manifolds in the following sections, we will assume they are smooth.

3.2 Lie algebras

Given a Lie group, we define Lie algebras.

Definition 3.2. The Lie algebra $\mathfrak{g}$ of a Lie group G is the tangent space at the identity element e, i.e. $\mathfrak{g} := \mathrm{T}_e G$.

Furthermore, we present the algebraic definition of Lie algebras. This definition also introduces the *Lie bracket* which we will use throughout this section, and it is more general than definition 3.2.

Definition 3.3. [1] A Lie algebra $\mathfrak{g}$ is a vector space with a bracket operation $[\cdot, \cdot]$ called the *Lie bracket*, such that

(i) $[\cdot, \cdot] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ is a bilinear map

(ii) It is skew-symmetric, i.e. $[u, v] = -[v, u]$, $\forall u, v \in \mathfrak{g}$

(iii) It satisfies the Jacobi identity, i.e.
$[u,[v,w]] + [w,[u,v]] + [v,[w,u]] = 0, \forall u,v,w \in \mathfrak{g}$

Example 3.2. We will identify the Lie group $SO(3)$ with the matrix Lie group of 3×3 orthogonal matrices with determinant 1. The Lie algebra can be identified with the vector space of 3×3 skew-symmetric matrices with the matrix commutator $[M,N] = MN - NM$, $M,N \in \mathfrak{so}(3)$ as the Lie bracket operation.

We define the Lie algebra of $SE(3)$, denoted $\mathfrak{se}(3)$, to be the pairs $(\hat{\xi},\eta)$, such that $\hat{\xi} \in \mathfrak{so}(3)$ and $\eta \in \mathbb{R}^3$. By identifying $\hat{\xi}$ as a skew-symmetric matrix in $\mathfrak{so}(3)$ we can instead represent it as a vector, and define a vector space isomorphism

$$\xi = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} \mapsto \hat{\xi} = \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix}$$

which is called the *hat map* [3]. With this notation and the properties of skew-symmetric matrices, we can simply write $\hat{\xi}\eta = \xi \times \eta$, using the cross product in $\mathbb{R}^3$. In fact, for $\hat{\xi},\hat{\eta},\hat{\nu} \in \mathfrak{so}(3)$ we have that

$$\begin{aligned}
[\hat{\xi},\hat{\eta}]\nu &= (\hat{\xi}\hat{\eta} - \hat{\eta}\hat{\xi})\nu = \hat{\xi}(\eta\times\nu) - \hat{\eta}(\xi\times\nu) \\
&= \xi\times(\eta\times\nu) - \eta\times(\xi\times\nu) = (\xi\times\eta)\times\nu \\
&= \widehat{\xi\times\eta}\nu.
\end{aligned}$$

Therefore, we can identify $\mathfrak{so}(3)$ with $\mathbb{R}^3$ carrying the cross product as Lie bracket. The hat map becomes a Lie algebra isomorphism [3].

3.3 Left and right invariant vector fields

Definition 3.4. A vector field $\mathbf{v}$ on a Lie group G is called left invariant if

$$dL_g(\mathbf{v}) = \mathbf{v},$$

where L_g : G $\to$ G is the left multiplication of G given by $L_g(h) = gh$ and $dL_g\big|_x : T_xG \to T_{gx}G$ is the derivative mapping.

Similarly, one can define right invariant vector fields. An important observation about left (and right) invariant vector fields, is that by knowing the vector field at the identity element e, we can reproduce it everywhere else. In particular,

$$\mathbf{v}(e) = A \Longrightarrow dL_g\big|_e (\mathbf{v}(e)) = \mathbf{v}(g). \tag{2}$$

Example 3.3. The identity element of the Lie group $SE(3)$ is the pair $(I, 0)$, where I is the 3×3 identity matrix and 0 is the zero vector in $\mathbb{R}^3$. We denote the left invariant vector field, $X_{(\hat{\xi},\eta)}$, and from equation (2) we know that the vector field at the identity should be $(\hat{\xi}, \eta)$, i.e.

$$X_{(\hat{\xi},\eta)}\Big|_{(I,0)} = (\hat{\xi}, \eta).$$

We can calculate $X_{(\hat{\xi},\eta)}$ everywhere by using the definition of left invariance,

$$X_{(\hat{\xi},\eta)} = dL_{(\hat{\xi},\eta)} X_{(\hat{\xi},\eta)},$$

and at a given point $(g, u) \in SE(3)$, we calculate

$$\begin{aligned} X_{(\hat{\xi},\eta)}\Big|_{(g,u)} &= TL_{(g,u)}(\hat{\xi}, \eta) = \frac{d}{dt}\Big|_{t=0} L_{(g,u)}(h(t), v(t)) \\ &= \frac{d}{dt}\Big|_{t=0} (g, u) \cdot (h(t), v(t)) = \frac{d}{dt}\Big|_{t=0} (gh(t), gv(t) + u) \\ &= (g\dot{h}(t), g\dot{v}(t))\Big|_{t=0} = (g\hat{\xi}, g\eta) \end{aligned}$$

where $(h(t), v(t))$ is a curve in $SE(3)$ such that $(\dot{h}(0), \dot{v}(0)) = (\hat{\xi}, \eta)$ and $(h(0), v(0)) = (I, 0)$. With an analogous calculation we can obtain the right invariant vector field, given by

$$TR_{(g,u)}(\hat{\xi}, \eta) = (\hat{\xi}g, \hat{\xi}u + \eta).$$

As expected, one can easily confirm that the vector field at the identity is $(\hat{\xi}, \eta)$, both for the left and right invariant case.

3.4 The exponential map

As introduced in section 2.4, the exponential map related to integral curves and the flow of a vector field is important to the study of Lie groups.

Definition 3.5. Given a Lie group G and its corresponding Lie algebra $\mathfrak{g}$, the exponential map $\exp : \mathfrak{g} \to G$ is given by

$$\exp(\xi) = \gamma(1)$$

where $\xi \in \mathfrak{g}$ and $\gamma : \mathbb{R} \to G$ is the integral curve of the left or right invariant vector field associated with ξ. Furthermore, we can also write $\exp(t\xi) = \gamma(t)$.

In particular, for all matrix Lie groups the exponential map coincides with the matrix exponential [4]. Additionally, one can prove that $\gamma(0) = \exp 0 = e$ and $\dot{\gamma}(0) = \xi$ where $e \in G$ is the identity element [3][2].

Example 3.4. We can find the exponential map of $SE(3)$ by calculating the flow of the left invariant vector field we have previously obtained. We recall the definition of an integral curve in our case; $(g(t), u(t))$ is the integral curve of the vector field $X_{(\hat{\xi},\eta)}$ if the tangent to the curve at t coincides with the vector field at $(\hat{\xi}, \eta)$, i.e.

$$(\dot{g}(t), \dot{u}(t)) = X_{(\hat{\xi},\eta)}\Big|_{(g(t),u(t))}.$$

Since we have already calculated the left invariant vector field in example 3.3, we get the differential equation

$$(\dot{g}(t), \dot{u}(t)) = (g(t)\hat{\xi}, g(t)\eta).$$

Solving for $g(t)$ and $u(t)$ is straightforward using the exponential and basic knowledge of differential equations, and we obtain

$$g(t) = \exp(\hat{\xi}t), \quad u(t) = \frac{\exp(\hat{\xi}t) - I}{\hat{\xi}}\eta$$

[2]This result uses *one-parameter subgroups* which are not defined in this text. They are, however, well described in the books by Marsden and Ratiu [3] and Lee [5], and we therefore refer the reader to these for more information.

using the initial condition $(g(0), u(0)) = (I, 0)$. Now, the final expression for the exponential is given at $t = 1$, i.e.

$$\exp(\hat{\xi}, \eta) = (g(1), u(1)) = \left(\exp \hat{\xi}, \frac{\exp \hat{\xi} - I}{\hat{\xi}} \eta\right)$$

[11]. One can verify that the exact same exponential will be obtained by using the right invariant vector field instead.

3.5 Adjoint and coadjoint representations

Definition 3.6. For an element $g \in G$ in a Lie group, we define a map

$$A_g : G \to G, \quad A_g(h) = ghg^{-1} = L_g \circ R_{g^{-1}}(h)$$

where L_g and R_g is the left and right multiplication of G, respectively, and $h \in G$ is any element in G. The adjoint representation of G on $\mathfrak{g}$, $Ad_g : \mathfrak{g} \to \mathfrak{g}$, is then the tangent map $TA_g\big|_e$ at identity, i.e. for an element $\xi \in \mathfrak{g}$ in the Lie algebra,

$$\mathrm{Ad}_g(\xi) = TA_g\big|_e(\xi) = TL_g \circ TR_{g^{-1}}(\xi).$$

Assuming we are working with a matrix Lie group that it is a subgroup of $GL(n)$[3], one can show that the adjoint representation is given by $\mathrm{Ad}_g(\xi) = g\xi g^{-1}$ [4].

Definition 3.7. Given a curve $g(t) \in G$ such that $g(0) = e$ and $\dot{g}(0) = \eta$, the Lie bracket $[\cdot, \cdot]$ can be defined as

$$[\eta, \xi] = \left.\frac{d}{dt}\right|_{t=0} \mathrm{Ad}_{g(t)}(\xi)$$

where $\mathrm{Ad}_{g(t)}(\xi)$ is the adjoint representation of G.

[3]Chapter 9 in [3] provides a detailed presentation of this Lie group, called the *General Linear Group*.

Definition 3.8. The adjoint representation $\mathrm{ad}_\xi : \mathfrak{g} \to \mathfrak{g}$ of the Lie algebra $\mathfrak{g}$ of a Lie group G, is given by

$$\mathrm{ad}_\xi(\eta) = [\xi, \eta]$$

where $\xi \in \mathfrak{g}$ is any element of $\mathfrak{g}$. This operator is also called the *ad operator*.

Proposition 3.1. [12] Powers of the adjoint representation ad_ξ are iterated commutators, i.e.

$$\begin{aligned}\mathrm{ad}_\xi^2(\eta) &= \mathrm{ad}_\xi([\xi, \eta]) = [\xi, [\xi, \eta]],\\ \mathrm{ad}_\xi^3(\eta) &= \mathrm{ad}_\xi([\xi, [\xi, \eta]]) = [\xi, [\xi, [\xi, \eta]]],\end{aligned}$$

and so on.

Example 3.5. By using definition 3.6 we can calculate the adjoint representation of $SE(3)$. Take $(g, u) \in SE(3)$ and $(\hat{\xi}, \eta) \in \mathfrak{se}(3)$, then

$$\begin{aligned}\mathrm{Ad}_{(g,u)}(\hat{\xi}, \eta) &= TL_{(g,u)} \circ TR_{(g,u)^{-1}}(\hat{\xi}, \eta) = L_{(g,u)} \circ TR_{(g^T, -g^T u)}(\hat{\xi}, \eta)\\ &= TL_{(g,u)}(\hat{\xi} g^T, \hat{\xi}(-g^T u) + \eta) = (g\hat{\xi} g^T, -g\hat{\xi} g^T u + g\eta)\\ &= (\widehat{g\xi}, -\widehat{g\xi} u + g\eta).\end{aligned}$$

Here we have used that $(g, u)^{-1} = (g^{-1}, -g^{-1}u) = (g^T, -g^T u)$ (since $g^{-1} = g^T$ for $g \in SE(3)$) and the property that $g\hat{\xi} g^{-1} = \widehat{g\xi}$.

Definition 3.9. The coadjoint representation $\mathrm{Ad}_\mathrm{g}^* : \mathfrak{g}^* \to \mathfrak{g}^*$ for $g \in G$ is a map from the dual of the Lie algebra $\mathfrak{g}$ to itself, such that

$$\left\langle \mathrm{Ad}_g^*(\mu), \xi \right\rangle = \langle \mu, \mathrm{Ad}_g(\xi) \rangle$$

for all $\mu \in \mathfrak{g}^*, \xi \in \mathfrak{g}$.

Recall that the dual space V^* of a linear space V consists of all linear functions on V, and for an operator $A : V \to V$ we can define the adjoint $A^* : V^* \to V^*$ with

$$\langle A^* \mu, v \rangle = \langle \mu, Av \rangle$$

for all $\mu \in V^*$ and $v \in V$, and $\langle \cdot, \cdot \rangle$ is the duality pairing [6]. The previous definition of the coadjoint representation is analogous, and we note that for real matrix groups $A^* = A^T$.

Example 3.6. We now derive the left coadjoint representation of $SE(3)$, $\mathrm{Ad}^*_{(g,u)^{-1}}$. From the previous example we can write

$$\begin{aligned}\mathrm{Ad}_{(g,u)^{-1}}(\hat{\xi},\eta) &= \mathrm{Ad}_{(g^T,-g^Tu)}(\hat{\xi},\eta) = (\widehat{g^T\xi}, -\widehat{g^T\xi}(-g^Tu) + g^T\eta) \\ &= (\widehat{g^T\xi}, g^T\xi \times g^Tu + g^T\eta) = (\widehat{g^T\xi}, g^T(\xi \times u) + g^T\eta).\end{aligned}$$

The duality pairing of $(\hat{\xi},\eta)$ with $(\hat{\mu},\nu) \in \mathfrak{se}(3)$ is

$$\langle(\hat{\mu},\nu),(\hat{\xi},\eta)\rangle = \mu^T\xi + \nu^T\eta$$

and so we use the definition above and obtain

$$\begin{aligned}\left\langle(\hat{\mu},\nu), \mathrm{Ad}_{(g,u)^{-1}}(\hat{\xi},\eta)\right\rangle &= \left\langle(\hat{\mu},\nu),(\widehat{g^T\xi}, g^T(\xi\times u) + g^T\eta)\right\rangle \\ &= \mu^Tg^T\xi + \nu^Tg^T(\xi \times u) + \nu^Tg^T\eta \\ &= \mu^Tg^T\xi - \nu^Tg^T(u \times \xi) + \nu^Tg^T\eta \\ &= \mu^Tg^T\xi - \nu^Tg^T\hat{u}\xi + \nu^Tg^T\eta \\ &= \left\langle((\mu^Tg^T \widehat{- \nu^T} g^T\hat{u})^T, (u^Tg^T)^T), (\hat{\xi},\eta)\right\rangle \\ &= \left\langle(\widehat{g\mu} + \widehat{u \times g\nu}, g\nu), (\hat{\xi},\eta)\right\rangle\end{aligned}$$

$$\Longrightarrow \mathrm{Ad}^*_{(g,u)^{-1}}(\hat{\mu},\nu) = (\widehat{g\mu} + \widehat{u \times g\nu}, g\nu).$$

Here we have used the fundamental property of skew-symmetric matrices, that $\hat{u}^T = -\hat{u}$ for $\hat{u} \in \mathfrak{so}(3)$.

The derivation we have done of the coadjoint of $SE(3)$ is important for working on the heavy top equations, since it has $\mathfrak{se}(3)^*$ as solution space [13]. We will see the applications of this in section 6.

Example 3.7. We now calculate the Lie bracket from definition 3.7. Given a curve $(g(t),u(t)) \in SE(3)$ and $(\hat{\xi},\eta),(\hat{\mu},\nu) \in \mathfrak{se}(3)$ such that $(g(0),u(0)) =$

$(I, 0)$ and $(\dot{g}(0), \dot{u}(0)) = (\hat{\xi}, \eta)$ we find

$$\begin{aligned}
\left[(\hat{\xi}, \eta), (\hat{\mu}, \nu)\right] &= \left.\frac{d}{dt}\right|_{t=0} \mathrm{Ad}_{(g(t), u(t))}\,(\hat{\mu}, \nu) \\
&= \left.\frac{d}{dt}\right|_{t=0} (\widehat{g(t)\mu}, -\widehat{g(t)\mu} u(t) + g(t)\nu) \\
&= \left.\frac{d}{dt}\right|_{t=0} (\widehat{\dot{g}(t)\mu}, -\widehat{\dot{g}(t)\mu} u(t) - \widehat{g(t)\mu}\dot{u}(t) + \dot{g}(t)\nu) \\
&= (\widehat{\hat{\xi}\mu}, -\hat{\mu}\eta + \hat{\xi}\nu) = (\widehat{\xi \times \mu}, -\mu \times \eta + \xi \times \nu).
\end{aligned}$$

3.6 The derivative of the exponential map

As we have previously defined the derivative and exponential in section 2.4, we want to add these properties together and define the derivative of the exponential map.

Definition 3.10. Let $u \in \mathfrak{g}$. Then the derivative of the exponential map $\mathrm{dexp}_u : \mathfrak{g} \to \mathfrak{g}$ is given by

$$\mathrm{dexp}_u(v) = TR_{\exp(-u)} \circ T\exp_u(v).$$

Proposition 3.2. [12] The derivative of the exponential map can be written as

$$\mathrm{dexp}_u(v) = \sum_{k=0}^{\infty} \frac{1}{(k+1)!}\,\mathrm{ad}_u^k(v) = \left.\frac{\exp(z) - 1}{z}\right|_{z=\mathrm{ad}_u} (v)$$

where $\mathrm{ad}_u : \mathfrak{g} \to \mathfrak{g}$ is the adjoint representation.

The proof of this uses the fact that powers of adjoint representation are iterated commutators as presented in proposition 3.1.

Furthermore, proposition 3.2 allows us to obtain the inverse of the derivative of the exponential, $\mathrm{dexp}_u^{-1}(v)$, by inverting $\mathrm{dexp}_u(v)$. We obtain

$$\mathrm{dexp}_u^{-1}(v) = \left.\frac{z}{\exp(z) - 1}\right|_{z=\mathrm{ad}_u} (v) = \sum_{k=0}^{\infty} \frac{B_k}{k!}\,\mathrm{ad}_u^k(v)$$

[12] where the coefficients B_k are the Bernoulli numbers.

3.7 Lie group actions

The adjoint and coadjoint representation are two examples of Lie group actions [3], and these are formally defined in the following definition.

Definition 3.11. Let $\mathcal{M}$ be a manifold and G a Lie group. Then a left Lie group action $\Lambda : G \times \mathcal{M} \to \mathcal{M}$ of G on $\mathcal{M}$ is a smooth map with the properties that

(i) $\Lambda(e, m) = m, \forall m \in \mathcal{M}$

(ii) $\Lambda(gh, m) = \Lambda(g, \Lambda(h, m)), \forall g, h \in G, m \in \mathcal{M}$

Similarly, we define a *right* Lie group action by changing the second property to $\Lambda(gh, m) = \Lambda(h, \Lambda(g, m))$. We can also denote the Lie group action as $\Lambda_g : \mathcal{M} \to \mathcal{M}$ for $g \in G$. Furthermore, we define a few important properties of Lie group actions.

Definition 3.12. A Lie group action $\Lambda : G \times \mathcal{M} \to \mathcal{M}$ is

(i) *Transitive* if for every $x, y \in \mathcal{M}$ there exists a $g \in G$ such that $y = \Lambda(g, x)$.

(ii) *Free* if for any $m \in \mathcal{M}$, $\Lambda(g, m) = m \Rightarrow g = e$.

3.8 The infinitesimal generator

Definition 3.13. Let $\Lambda : G \times \mathcal{M} \to \mathcal{M}$ be a Lie group action of G on $\mathcal{M}$, and for $\xi \in \mathfrak{g}$ and $m \in \mathcal{M}$, define an $\mathbb{R}$-action $\Lambda^\xi : \mathbb{R} \times \mathcal{M} \to \mathcal{M}$ such that $\Lambda^\xi(t, m) = \Lambda(\exp t\xi, m)$. Then the infinitesimal generator of the action Λ is given by

$$\lambda_*(\xi)|_m = \left.\frac{d}{dt}\right|_{t=0} \Lambda(\exp t\xi, m).$$

$\lambda_*(\xi)$ is a vector field on $\mathcal{M}$ for every $\xi \in \mathfrak{g}$.

As we have previously remarked, the adjoint and coadjoint representations are Lie group actions, and so we can find their infinitesimal generators.

Proposition 3.3. The infinitesimal generator of the adjoint action of G on $\mathfrak{g}$, $Ad_g : \mathfrak{g} \to \mathfrak{g}$ for $g \in G$, is given by

$$\lambda_*(\xi)\big|_\eta = [\xi, \eta] = \mathrm{ad}_\xi(\eta).$$

Here, $\mathrm{ad}_\xi : \mathfrak{g} \to \mathfrak{g}$ for $\xi \in \mathfrak{g}$ is the adjoint representation of $\mathfrak{g}$. Furthermore, the infinitesimal generator of the coadjoint action of G on $\mathfrak{g}^*$, $Ad^*_{g^{-1}} : \mathfrak{g}^* \to \mathfrak{g}^*$ for $g \in G$, is given by

$$\gamma_*(\xi)\big|_\eta = -\,\mathrm{ad}^*_\xi(\eta).$$

Proof. [3] For the first statement we simply have

$$\lambda_*(\xi)\big|_\eta = \mathrm{Ad}_{\exp t\xi}(\eta) = [\xi, \eta] = \mathrm{ad}_\xi(\eta),$$

following directly from definition 3.7, 3.8 and 3.13. For the second statement we take $\mu \in \mathfrak{g}$ and obtain

$$\begin{aligned}
\left\langle \gamma_*(\xi)\big|_\eta, \mu \right\rangle &= \left\langle \frac{d}{dt}\bigg|_{t=0} \mathrm{Ad}^*_{\exp(-t\xi)}(\eta), \mu \right\rangle \\
&= \frac{d}{dt}\bigg|_{t=0} \left\langle \mathrm{Ad}^*_{\exp(-t\xi)}(\eta), \mu \right\rangle \\
&= \frac{d}{dt}\bigg|_{t=0} \left\langle \eta, \mathrm{Ad}_{\exp(-t\xi)}(\mu) \right\rangle \\
&= \left\langle \eta, \frac{d}{dt}\bigg|_{t=0} \mathrm{Ad}_{\exp(-t\xi)}(\mu) \right\rangle \\
&= \langle \eta, -\,\mathrm{ad}_\xi(\mu)\rangle = \left\langle -\,\mathrm{ad}^*_\xi(\eta), \mu \right\rangle \\
\implies \gamma_*(\xi)\big|_\eta &= -\,\mathrm{ad}^*_\xi(\eta).
\end{aligned}$$

□

Example 3.8. Take $(\hat{\xi}, \eta), (\hat{\mu}, \nu) \in \mathfrak{se}(3)$, then the infinitesimal generator induced by the left coadjoint action will map $(\hat{\xi}, \eta)$ to a vector field on $\mathfrak{se}(3)^*$. We write

$$\lambda_*(\hat{\xi}, \eta)\Big|_{(\hat{\mu}, \nu)} = \frac{d}{dt}\bigg|_{t=0} \Lambda(\exp(t(\hat{\xi}, \eta)), (\hat{\mu}, \nu)),$$

where $\exp(t(\hat{\xi}, \eta)) = (g(t), u(t))$ is a curve in $SE(3)$. We simplify the expression by using that $(g(t), u(t))$ passes through the identity at $t = 0$ with tangent $(\hat{\xi}, \eta)$, i.e. $(\dot{g}(0), \dot{u}(0)) = (\hat{\xi}, \eta)$. This gives us

$$\begin{aligned}
\lambda_*(\hat{\xi}, \eta)\Big|_{(\hat{\mu}, \nu)} &= \frac{d}{dt}\bigg|_{t=0} \mathrm{Ad}^*_{(g(t), u(t))^{-1}}(\hat{\mu}, \nu) \\
&= \frac{d}{dt}\bigg|_{t=0} (\widehat{g(t)\mu} + \widehat{u(t) \times g(t)}\nu, g(t)\nu) \\
&= (\widehat{\dot{g}(t)\mu} + \widehat{\dot{u}(t) \times g(t)}\nu + \widehat{u(t) \times \dot{g}(t)}\nu, \dot{g}(t)\nu)\Big|_{t=0} \\
&= (\widehat{\hat{\xi}\mu} + \widehat{\eta \times \nu}, \hat{\xi}\nu) = (\widehat{\xi \times \mu} + \widehat{\eta \times \nu}, \xi \times \nu).
\end{aligned}$$

4 Lie group integrators

4.1 Introduction

We now have all the information necessary to begin derive our Lie group integrators. The basic idea is to have a smooth manifold $\mathcal{M}$, a vector field F on $\mathcal{M}$ (the tangent space $F(y(t)) \in T_{y(t)}\mathcal{M} \quad \forall t$), a transitive Lie group action on $\mathcal{M}$ and its infinitesimal generator. Our Lie group is denoted G with the Lie algebra $\mathfrak{g}$. Then for $m \in \mathcal{M}$, we can construct a map $f : \mathcal{M} \to \mathfrak{g}$ such that

$$F|_m = \lambda_*(f(m))|_m$$

[12]. Using this idea, we want to solve the initial value problem

$$\begin{aligned} \dot{y} &= F(y) \\ y\left(t_0\right) &= y_0 \end{aligned} \tag{3}$$

for $y_0 \in \mathcal{M}$. The reason for this approach to such a problem is that classical integration methods are not going to make sense once we are working on vector fields where normal operations like sum and multiplication might not be well defined.

4.2 Lie-Euler method

The explicit Euler method, given by the equation

$$y_{n+1} = y_n + hf\left(y_n\right),$$

is a well known first-order method to solve initial value ODE's [7]. We want to convert this method to a method on manifolds, usually called the Lie-Euler method.

The idea is to look at the problem locally. This means we define $\dot{z} = f\left(y_n\right)$, $z(0) = y_n$ and set $y_{n+1} = z(h)$, $t_{n+1} = t_n + h$.

We look at the case for matrix Lie groups. Consider the Lie group equation $\dot{y} = A(y)y$ [14] where A is a matrix. Locally this becomes $\dot{z} = A(y_n)\,z$, $z(0) = y_n$. The solution is then given by $z(t) = \mathrm{e}^{tA(y_n)}y_n$ meaning that we obtain the Lie-Euler method [15]

$$y_{n+1} = \mathrm{e}^{hA(y_n)}y_n, \tag{4}$$

where e is the matrix exponential.

4.3 Runge-Kutta Munthe-Kaas methods

Runge-Kutta methods are well known classical integration methods [8], and we can create a similar, yet relevant method for working on manifolds. As mentioned in the introduction of this section, we need a transitive Lie group action on $\mathcal{M}$. We denote it $\Lambda : G \times \mathcal{M} \to \mathcal{M}$. For a neighborhood of $y_0 \in \mathcal{M}$ we can write $y(t)$ as

$$y(t) = \Lambda\left(\exp(\phi(t)), y_0\right)$$

where $\phi(t) \in \mathfrak{g}$ is a curve in $\mathfrak{g}$. We are essentially converting the problem from $\mathcal{M}$ to $\mathfrak{g}$. One can show that the differential equation for ϕ in $\mathfrak{g}$ becomes

$$\dot{\phi} = \mathrm{dexp}_{\phi}^{-1}\left(f\left(\Lambda_{y_0}(\exp(\phi))\right)\right) \tag{5}$$

where $\Lambda_{y_0}(g) = \Lambda\left(g, y_0\right)$ for $g \in G$ and $f(x) := \sum_{i=1}^{d} f_i(x)E_i \in \mathfrak{g}$ such that $E_1, \ldots, E_d$ is a basis of the Lie algebra $\mathfrak{g}$. See [12] for details on this derivation.

We can now write one step of the Runge-Kutta Munthe-Kaas (RKMK) method, given y_n, step size h and the function $f : \mathcal{M} \to \mathfrak{g}$.

$$\begin{aligned}
&\text{for i = 1 to s:}\\
&\qquad \phi_i = h\sum_{j=1}^{i-1} a_{i,j}\,\hat{\mathrm{dexp}}_{\phi_j}^{-1}\left(f\left(\Lambda_{y_n}\left(\exp\left(\phi_j\right)\right)\right)\right)\\
&\tilde{\phi} = h\sum_{i=1}^{s} b_i\,\mathrm{dexp}_{\phi_i}^{-1}\left(f\left(\Lambda_{y_n}\left(\exp\left(\phi_i\right)\right)\right)\right)\\
&Y_1 = \Lambda_{y_n}(\exp(\tilde{\phi}))
\end{aligned} \tag{6}$$

[12]. Here $a_{i,j}$ and b_i for $i = 1, \ldots, s$, $j = 1, \ldots, s$ are the parameters of a classical Runge-Kutta method, and $\hat{\mathrm{dexp}}_\phi^{-1}$ is the truncation (to the correct consistency order) of the expansion

$$\mathrm{dexp}_{\phi(t)}^{-1}(u) = \left.\frac{z}{e^z - 1}\right|_{z=\mathrm{ad}_\phi}(u) = \sum_{k=0}^{\infty} \frac{B_k}{k!} \mathrm{ad}_\phi^k(u). \tag{7}$$

We are simply using a classical Runge-Kutta method on equation (5).

When working on the free rigid body and heavy top models, we will in particular use the 4th order Runge-Kutta Munthe-Kaas method. Relevant is then the Butcher tableau of the 4th order Runge-Kutta method given by table 1, which gives us the coefficients $a_{i,j}$ and b_i.

$$\begin{array}{c|cccc} 0 & & & & \\ \frac{1}{2} & \frac{1}{2} & & & \\ \frac{1}{2} & 0 & \frac{1}{2} & & \\ 1 & 0 & 0 & 1 & \\ \hline & \frac{1}{6} & \frac{1}{3} & \frac{1}{3} & \frac{1}{6} \end{array}$$

Table 1: Butcher tableau of 4th order Runge-Kutta method [8].

4.4 Commutator free methods

We now present commutator free methods. We refer the reader to [16] for a more detailed presentation of this.

These methods do not involve commutators and use as few exponentials as possible. The idea is to use frames and frozen vector fields as introduced in section 2.3, and the fact that a smooth vector field F on $\mathcal{M}$ can be expressed in terms of the frames (equation (1)).

We present one step of the method as in [16].

$$
\begin{aligned}
&\text{for i = 1 to r:}\\
&\qquad Y_r = \exp\left(\sum_k \alpha_{rJ}^k F_k\right) \cdots \exp\left(\sum_k \alpha_{r1}^k F_k\right) p\\
&\qquad F_r = hF_{Y_r} = h\sum_i f_i\left(Y_r\right) E_i\\
&y_1 = \exp\left(\sum_k \beta_J^k F_k\right) \cdots \exp\left(\sum_k \beta_1^k F_k\right) p
\end{aligned}
\tag{8}
$$

For a commutator free method of order 4, a generalized version of the classical Runge-Kutta method of order 4 is given in the Butcher tableau in table 2. Observe that this method uses five exponentials per step.

$$
\begin{array}{c|cccc}
0 & & & & \\
\frac{1}{2} & \frac{1}{2} & & & \\
\frac{1}{2} & 0 & \frac{1}{2} & & \\
1 & \frac{1}{2} & 0 & 0 & \\
 & -\frac{1}{2} & 0 & 1 & \\
\hline
 & \frac{1}{4} & \frac{1}{6} & \frac{1}{6} & -\frac{1}{12} \\
 & -\frac{1}{12} & \frac{1}{6} & \frac{1}{6} & \frac{1}{4}
\end{array}
$$

Table 2: Butcher tableau of 4th order commutator free method [16].

5 Free rigid body

5.1 The free rigid body equations

The first application of the theoretical material we have introduced will be the free rigid body model. A free rigid body is a rigid body rotating freely about its center of mass with no external forces involved, and the distance between two points of the body remains unchanged during the motion. We assume a smooth motion, implying that we can describe the rotation as a function $f(x,t) = A(t)x$. Here, $A(t)$ is a proper rotation, meaning $A(t) \in SO(3)$ for time t and $x \in \mathbb{R}^3$ is a point in space. See chapter 9 and 15 in the book by Marsden and Ratiu [3] for more information on this system.

We have presented examples of $SE(3)$ in section 3, and we will use it more in section 6, but the subgroup $SO(3)$ is more appropriate for the free rigid body considered in this section.

We begin by presenting the Euler free rigid body equations

$$\dot{m} = -\mathbb{I}^{-1}m \times m \tag{9}$$

[3]. $m \in \mathbb{R}^3$ is the angular momentum in body coordinates and $\mathbb{I} \in \mathbb{R}^{3\times 3}$ is the inertia tensor

$$\mathbb{I} = \begin{bmatrix} I_1 & 0 & 0 \\ 0 & I_2 & 0 \\ 0 & 0 & I_3. \end{bmatrix}$$

As usual, $\times$ is the standard cross product in $\mathbb{R}^3$. We can simplify equation (9) into a more useful matrix equation

$$\dot{m} = \begin{bmatrix} \dot{m}_1 \\ \dot{m}_2 \\ \dot{m}_3 \end{bmatrix} = \begin{bmatrix} 0 & \frac{m_3}{I_3} & -\frac{m_2}{I_2} \\ -\frac{m_3}{I_3} & 0 & \frac{m_1}{I_1} \\ \frac{m_2}{I_2} & -\frac{m_1}{I_1} & 0 \end{bmatrix} \begin{bmatrix} m_1 \\ m_2 \\ m_3 \end{bmatrix} = A(m) \cdot m. \tag{10}$$

We see that this equation is a Lie group equation as introduced in section 4.2. We will implement the Lie Euler method, RKMK and commutator free method for this model.

The implementations will use arbitrary values for the inertia vector $I = [I_1, I_2, I_3]$ (the diagonal of the inertia tensor $\mathbb{I}$) and the initial condition m_0. We choose $I = [1, 5, 6]$ and $m_0 = [0.7, 0.3, 0.5]$ and integrate over the time interval $[0, 1]$ with 100 steps, implying step size $h = 0.01$.

5.2 Lie Euler method

We begin with a method of order 1, the Lie Euler method as described in section 4.2. This is perhaps one of the simplest Lie group methods, and hence it is a great place to start. The single step of this method is given by equation (4). Here h is the step size and e is the exponential map given by the matrix exponential since $SO(3)$ is a matrix Lie group. Since $A(m_n)$ is a skew-symmetric matrix we use the Rodrigues' formula [3]

$$\mathrm{e}^{A} = I + \frac{\sin\alpha}{\alpha}A + \frac{1-\cos\alpha}{\alpha^2}A^2 \tag{11}$$

where $\alpha = \|a\| = \sqrt{a_1^2 + a_2^2 + a_3^2}$ and a is the vector mapping to A through the hat map as defined in section 3.2.

5.3 Runge-Kutta Munthe-Kaas

The next Lie group method we want to implement is the RKMK method of order 4. This is a better method than Lie Euler. One can show that it is defined in the following way.

$$\begin{aligned}
A_1 &= hA\left(m_n\right)\\
A_2 &= hA\left(\exp\left(\frac{1}{2}A_1\right)\cdot m_n\right)\\
A_3 &= hA\left(\exp\left(\frac{1}{2}A_2 - \frac{1}{8}\left[A_1, A_2\right]\right)\cdot m_n\right)\\
A_4 &= hA\left(\exp\left(A_3\right)\cdot m_n\right)\\
m_{n+1} &= \exp\left(\frac{1}{6}\left(A_1 + 2A_2 + 2A_3 + A_4 - \frac{1}{2}\left[A_1, A_4\right]\right)\right)\cdot m_n
\end{aligned}$$

[12]. Here $[\cdot,\cdot]$ is the Lie bracket of $SO(3)$ or the matrix commutator as for every matrix Lie group, h is the step size and exp is the matrix exponential.

5.4 Commutator free method

To avoid large matrix commutator calculations we implement a third method, a commutator free method of order 4 as introduced in section 4.4. One step of the method is given by

$$\begin{aligned}
M_1 &= m_n \\
M_2 &= \exp\left(\frac{1}{2}hA_1\right)\cdot m_n \\
M_3 &= \exp\left(\frac{1}{2}hA_2\right)\cdot m_n \\
M_4 &= \exp\left(hA_3 - \frac{1}{2}hA_1\right)\cdot M_2 \\
m_{n+\frac{1}{2}} &= \exp\left(\frac{1}{12}h\left(3A_1 + 2A_2 + 2A_3 - A_4\right)\right)\cdot m_n \\
m_{n+1} &= \exp\left(\frac{1}{12}h\left(-A_1 + 2A_2 + 2A_3 + 3A_4\right)\right)\cdot m_{n+\frac{1}{2}}
\end{aligned}$$

using the Butcher tableau in table 2. Here $A_i = A\left(M_i\right)$ for every i.

5.5 Convergence orders

As we have implemented the three different methods, we want to show that the convergence orders are as stated; 1 for Lie Euler and 4 for both RKMK and commutator free method.

To check the orders we define an error function. This will compare the angular momentum m of the Lie group method with a reference solution at time T. It will apply the Lie group method with different step sizes, find the 2-norm between the end values and return a list of these errors. We use `scipy.integrate.odeint` in Python to obtain a reference solution to the problem. To obtain good results we need a low tolerance and we use a relative and absolute tolerance of 10^{-13} and 10^{-14}, respectively.

We produce a loglog-plot of the convergence orders next to $\mathcal{O}(h^i)$ for $i = 1, 2, 3, 4$ to compare, and obtain the plot in figure 2.

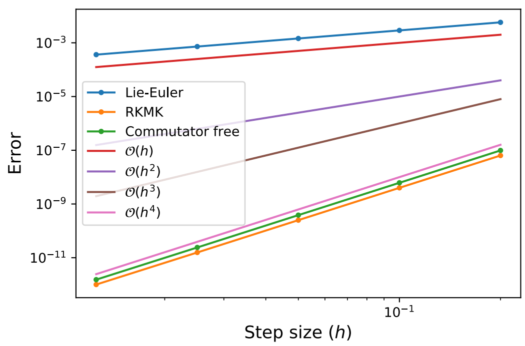


Figure 2: Convergence orders of Lie Euler method, RKMK and commutator free method alongside $\mathcal{O}(h^i)$ for $i = 1, 2, 3, 4$. The step sizes are chosen to be $h = \frac{1}{5}, \frac{1}{10}, \frac{1}{20}, \frac{1}{40}, \frac{1}{80}$.

Clearly the convergence orders are as expected, seeing that the loglog-plot of the Lie Euler method is parallel to $\mathcal{O}(h)$ and the loglog-plots of RKMK and commutator free method are parrallel to $\mathcal{O}(h^4)$.

The importance of showing the convergence orders comes from the fact that a higher convergence order means that the error approaches zero faster as the step size decreases. It is therefore clear that the methods of order 4 will be more useful than the Lie Euler method; they will require fewer steps and hence converge to the solution quicker. We cannot say more about the advantages and disadvantages of using RKMK over commutator free method from the order of convergence, however the commutator free method is - as the name implies - commutator free. This means that the method does not take use of the matrix commutator, i.e. two matrix multiplications. It does not require much extra time to compute these when we work with small matrices like those in $SO(3)$, but with problems on much larger matrices it might make a difference.

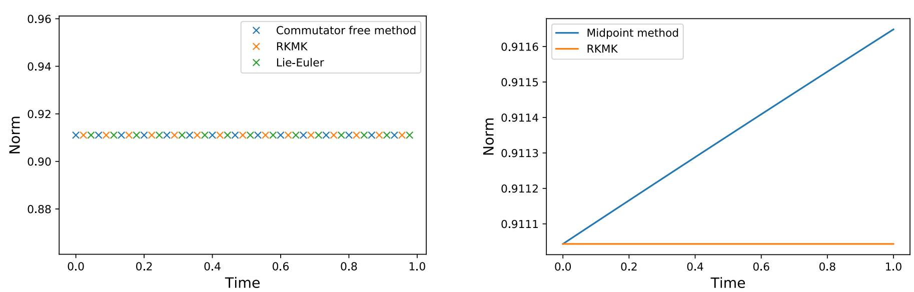
5.6 Comparison with classical integration methods

An important note on the free rigid body system is that the angular momentum is conserved [3], i.e. the norm of the vector m is conserved over time. This is valuable from a physical perspective and we will see that in this case the Lie group integrators have an advantage over classical integration methods.

We take the modified Euler method as an example [7]. This is an explicit numerical method and one step is given by

$$y_{n+1} = y_n + hf\left(t_n + \frac{h}{2}, y_n + \frac{h}{2}f\left(t_n, y_n\right)\right).$$

We can can produce two plots; one to compare the norms of the Lie group integrators and one to compare with the modified Euler method. Figure 3a demonstrates that the norm indeed is conserved for the Lie group integrators



(a) (b)

Figure 3: Norm of the angular momentum vector m for the free rigid body model. (a) Shows shows the norms using Lie group integrators, while (b) compares the Lie group method (RKMK) against a classical numerical method (the modified Euler method).

we have implemented. Furthermore, we plot the angular momentum norms with RKMK against the modified Euler method in figure 3b, obtaining a clear indication that the modified Euler method does not conserve angular momentum.

6 The heavy top

6.1 The heavy top equations

The heavy top is a rigid body moving with a fixed point under the influence of gravity [3]. This means that we can describe the motion by the special Euclidean group $SE(3)$. We also mentioned in section 3 that the solution space is $\mathfrak{se}(3)^*$, so the work we have done in the examples will be relevant to solve these equations and apply Lie group integrators.

From Marsden and Ratiu [3] we formulate the heavy top equations. In our own notation, they are given by

$$\begin{aligned} \dot{\mu} &= \mu \times \mathbb{I}^{-1}\mu + \nu \times c\chi \\ \dot{\nu} &= \nu \times \mathbb{I}^{-1}\mu \end{aligned} \tag{12}$$

where $\mathbb{I}$ is the inertia tensor as a 3×3 diagonal matrix with the inertia elements I_1, I_2 and I_3 on the diagonal. Furthermore, c is mass times the gravitational acceleration g times the line segment of length l connecting the fixed point to the center of mass of the body. ν is a vector pointing in the direction from the fixed point of the heavy top to the center of mass (in body coordinates) and χ is the constant unit vector along the line segment of length l. The solution consists of the pair $(\mu(t), \nu(t))$ where $\mu(t) \in \mathfrak{so}(3)^*$ is the angular momentum and $\nu \in \mathbb{R}^3$ is a vector.

We want to use these equations to find a function $f : \mathfrak{se}(3)^* \to \mathfrak{se}(3)$ introduced in equation (5) in section 4.3. This can be done by using the infinitesimal generator, λ_* and calculating $(\dot{\mu}, \dot{\nu}) = \lambda_*(f(\mu, \nu))|_{(\mu,\nu)}$. We get the following calculation.[4]

$$\begin{aligned} (\dot{\mu}, \dot{\nu}) = \lambda_*(f(\mu, \nu))|_{(\mu,\nu)} &= \lambda_*(f_1(\mu, \nu), f_2(\mu, \nu))|_{(\mu,\nu)} \\ &= (f_1(\mu, \nu) \times \mu + f_2(\mu, \nu) \times \nu, f_1(\mu, \nu) \times \nu) \end{aligned}$$

[4]We remove the notation of hats since it is apparent which elements are vectors and which are matrices.

We substitute $\dot{\mu}$ and $\dot{\nu}$ with the equations in (12), and obtain

$$\begin{aligned} f_1(\mu, \nu) &= -\mathbb{I}^{-1}\mu \\ f_2(\mu, \nu) &= -c\chi. \end{aligned} \tag{13}$$

The inertia vector $I = [I_1, I_2, I_3]$, the vector χ and the initial value y_0 are chosen to be $I = [1, 5, 6]$, $\chi = [0, 0, 1]$ and $y_0 = [\sin 1.1, 0, \cos 1.1, 1, 0.2, 3]$ in our implementation. Here, we represent a value $y = (\mu, \nu) \in \mathfrak{se}(3)^*$ as a 6-dimensional vector such that the first and the last three elements are the elements of the vector μ and ν, respectively. For simplicity we set $c = 1$ and integrate over the time interval $[0, 1]$ with 100 steps.

6.2 Runge-Kutta Munthe-Kaas

We begin by implementing the RKMK method of order 4 for the heavy top model. By section 4.3 we need to implement functions for dexp_ϕ^{-1}, f, Λ_{y_n} and exp, where $\phi(t)$ is a curve in $\mathfrak{se}(3)$. f is the function in equation (13), dexp_ϕ^{-1} is given by equation (7) and Λ_{y_n} is a transitive Lie group action, chosen to be the left coadjoint action of $SE(3)$. Then we can implement the method given in equation (6), and one step of the algorithm is simplified to

$$\begin{aligned} K_1 &= F(0, y_n) \\ K_2 &= F\left(\frac{1}{2}hK_1, y_n\right) \\ K_3 &= F\left(\frac{1}{2}hK_2, y_n\right) \\ K_4 &= F(hK_3, y_n) \\ y_{n+1} &= \Lambda_{y_n}\left(\exp\left(\frac{1}{6}hK_1 + \frac{1}{3}hK_2 + \frac{1}{3}hK_3 + \frac{1}{6}hK_4\right)\right) \end{aligned}$$

where $F(\phi, y) := \dot{\phi}$ is the vector field defined in equation (5). There is a similar formula to the Rodrigues' formula for calculating the exponential of elements in $\mathfrak{se}(3)$ [17]. Given $\mathbf{v} = (\hat{\xi}, \eta) \in \mathfrak{se}(3)$,

$$\exp(\mathbf{v}) = (\exp(\hat{\xi}), \mathbf{V}\eta)$$

where

$$\mathbf{V} = I + \frac{1-\cos\alpha}{\alpha^2}\hat{\xi} + \frac{\alpha-\sin\alpha}{\alpha^3}\hat{\xi}^2$$

such that $\alpha = \|\xi\|$. $\exp(\hat{\xi})$ is the exponential map for elements in $\mathfrak{so}(3)$ defined in equation (11) of the previous section.

6.3 Commutator free method

As for the 4th order RKMK method for the heavy top, we use the vector field $F(\phi, y) := \dot{\phi}$ and the transitive Lie group action Λ_{y_n} to simplify the commutator free method in equation (8) to the following algorithm.

$$\begin{aligned}
K_1 &= y_n \\
A_1 &= F(0, K_1) \\
K_2 &= \Lambda_{y_n}\left(\exp\left(\frac{1}{2}hA_1\right)\right) \\
A_2 &= F(0, K_2) \\
K_3 &= \Lambda_{y_n}\left(\exp\left(\frac{1}{2}hA_2\right)\right) \\
A_3 &= F(0, K_3) \\
K_4 &= \Lambda_{K_2}\left(\exp\left(hA_3 - \frac{1}{2}hA_1\right)\right) \\
A_4 &= F(0, K_4) \\
y_{n+\frac{1}{2}} &= \Lambda_{y_n}\left(\exp\left(\frac{1}{12}h\left(3A_1 + 2A_2 + 2A_3 - A_4\right)\right)\right) \\
y_{n+1} &= \Lambda_{y_{n+\frac{1}{2}}}\left(\exp\left(\frac{1}{12}h\left(-A_1 + 2A_2 + 2A_3 + 3A_4\right)\right)\right)
\end{aligned}$$

6.4 Convergence orders

Similar to the free rigid body model, we want to study the convergence orders and confirm what we have previously stated. We use `scipy.integrate.odeint` in Python to obtain a reference solution to the problem, with a relative and absolute tolerance of 10^{-13} and 10^{-14}, respectively. By the same method as in section 5.5, we define an error function and obtain a loglog-plot of the error in figure 4.

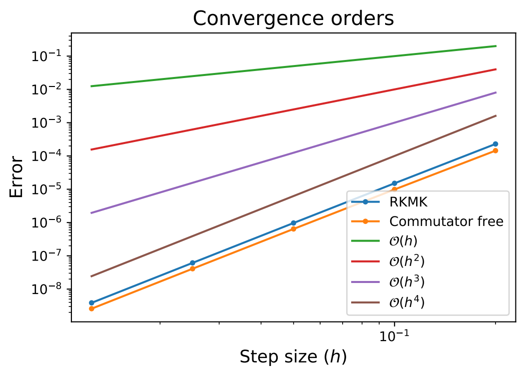


Figure 4: Convergence order of RKMK and commutator free method alongside $\mathcal{O}(h^i)$ for $i = 1, 2, 3, 4$. The step sizes are $h = \frac{1}{5}, \frac{1}{10}, \frac{1}{20}, \frac{1}{40}, \frac{1}{80}$.

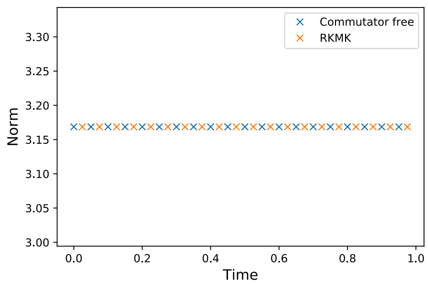
We observe that both the convergence order of RKMK and commutator free method is 4. However, it is also worth noting that the two methods use a different number of exponentials. RKMK uses 4, while commutator free method uses 5.

6.5 Comparison with classical integration methods

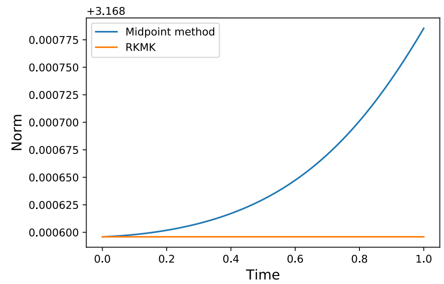
We choose to work with the classical modified Euler method to compare with RKMK and commutator free method.

In the case of the heavy top equations, the angular momentum μ will not be conserved [3]. However, the Euclidean norm of the vector ν is constant, since we have that

$$\begin{aligned}\frac{d}{dt}\|\nu\|^2 &= \frac{d}{dt}\nu^T\nu = 2\nu^T\dot{\nu} = 2\nu^T(\nu \times \mathbb{I}^{-1}\mu) \\ &= 2(\nu \times \nu)^T\mathbb{I}^{-1}\mu = 0.\end{aligned}$$



(a)



(b)

Figure 5: Norm of the vector ν for the heavy top model. (a) Shows shows the norms using the Lie group integrators RKMK and commutator free method, while (b) compares the Lie group method RKMK against a classical numerical method (the modified Euler method).

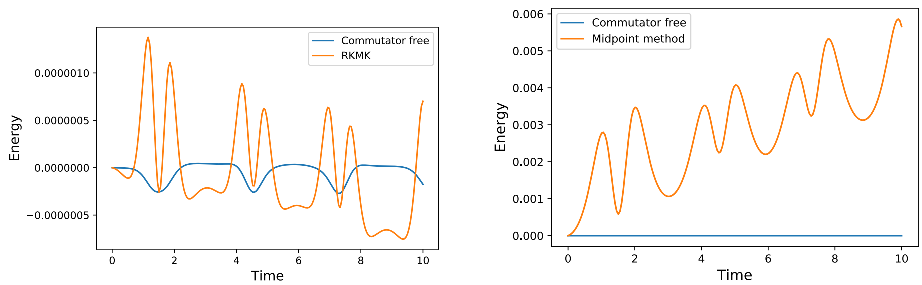
We can verify that this is true for our implementation, and figure 5 shows the results. The figure also shows that the modified Euler method does not conserve the vector ν, meaning this method is clearly less suitable for the problem.

Furthermore, we study the conservation of energy of the system. The Hamiltonian of the heavy top equations is presented in [3] and is given by

$$H(\mu, \nu) = \frac{1}{2} \left\langle \mu, \mathbb{I}^{-1} \mu \right\rangle + \langle \nu, c\mathcal{X} \rangle ,$$

where $\langle \cdot, \cdot \rangle$ is the standard inner product on $\mathbb{R}^3$. Since the Hamiltonian energy of our system should be conserved, we plot the energy error over time by comparing the energy at time t with the energy at time $t = 0$. We use the Lie group integrators RKMK and commutator free method, and compare against each other and the classical modified Euler method. Observe the resulting plots in figure 6. We have increased the time interval to $[0, 10]$ with 200 steps such that the step size is $h = 0.05$.

The Hamiltonian energy is not conserved for any of the chosen methods, however from figure 6 it is clear that the energy error is a significantly smaller



(a) **(b)**

Figure 6: Hamiltonian energy error of the heavy top model using Lie group integrators (RKMK and commutator free method) and the modified Euler method. (a) shows the two Lie group integrators against each other, while (b) shows the commutator free method against the modified Euler method.

for the Lie group integrators compared to the modified Euler method. Additionally, the figure suggests that the commutator free method is better than RKMK, because of a smaller deviation in energy.

We can argue once again that from a physical perspective, the chosen Lie group integrators are more reliable than the classical integration method. This is clear since neither the vector ν nor the Hamiltonian energy is conserved by using the modified Euler method, and the energy error is unquestionably larger than that of the Lie group integrators.

7 Conclusion

In this thesis we have studied some elementary material of Lie group methods and seen some applications by simulating two mechanical systems. The theoretical material has provided a basis for understanding the Lie group integrators. We have in particular focused on studying 4th order Runge-Kutta Munthe-Kaas and commutator free methods, and we have briefly introduced the Lie-Euler method. However, there are several methods, e.g. Magnus methods [14] and RKMK and commutator free methods of orders different from 4, that we have not treated in this thesis. We refer the reader to [11, 14, 16] for further reading on these topics.

We have seen that the implementation of the two models, free rigid body and heavy top, have favourable results by using Lie group integrators over classical numerical methods. This is mainly because the Lie group methods have an unquestionable smaller error in conserving the structures of the manifolds, such as angular momentum and energy. Both the angular momentum for the free rigid body and the norm of the vector ν for the heavy top are only conserved using Lie group methods as opposed to the modified Euler method. Furthermore, the Hamiltonian energy of the heavy top system is not completely conserved for the Lie group methods, however, compared to the modified Euler method, the energy error is undoubtedly smaller.

The models we have presented are only two examples of mechanical systems of which we can apply Lie group integrators. In addition to $SO(3)$ and $SE(3)$, the book by by Marsden and Ratiu [3] introduces several other Lie groups that occur in mechanics, and provide short introductions to their applications. Additionally, several different rigid and multi-body systems are presented in the book by Leok [9] (in particular in chapter 9 and 10), and we refer the reader to this for further reading on formulations of more complex systems.

References

[1] P. Olver. *Equivalence Invariants and Symmetries*, Cambridge University Press, Cambridge, 1995.

[2] J. M. Lee. *Introduction to Topological Manifolds*, Graduate Texts in Mathematics, Springer-Verlag, 2000.

[3] J. E. Marsden and T. S. Ratiu. *Introduction to Mechanics and Symmetry*, Texts in Applied Mathematics vol. 17, Springer-Verlag, 1999.

[4] B. C. Hall. *Lie Groups, Lie Algebras, and Representations*, Graduate Texts in Mathematics, Springer-Verlag, 2015.

[5] J. M. Lee. *Introduction to Smooth Manifolds*, Graduate Texts in Mathematics, Springer-Verlag, 2012.

[6] S. Axler. *Linear Algebra Done Right*, Undergraduate Texts in Mathematics, Springer-Verlag, 2015.

[7] E. Süli and D. Mayers. *An Introduction to Numerical Analysis*, Cambridge University Press, 2003.

[8] E. Hairer, G. Wanner and S. P. Nørsett. *Solving Ordinary Differential Equations I*, Nonstiff Problems, Springer-Verlag, 1993.

[9] T. Lee, M. Leok and N. H. McClamroch. *Global Formulations of Lagrangian and Hamiltonian Dynamics on Manifolds*, Interaction of Mechanics and Mathematics, Springer-Verlag, 2018.

[10] P. E. Crouch and R. Grossman. *Numerical Integration of Ordinary Differential Equations on Manifolds*, J. Nonlinear Sci., **3** (1993), 1-33.

[11] C. Curry and B. Owren. *Variable step size commutator free Lie group integrators*, Numerical Algorithms, **82** (2019), 1359-1376.

[12] H. Munthe-Kaas. *High order Runge-Kutta methods on manifolds*, Appl. Num. Math., **29** (1999), 115-127.

[13] E. Celledoni, E. Çokaj, A. Leone, D. Murari and B. Owren. *Lie Group integrators for mechanical systems*, arXiv:2102.12778 [math.NA], February 2021.

[14] A. Iserles, H. Munthe-Kaas, S. P. Nørsett and A. Zanna. *Lie-group methods*, Cambridge University Press, 1999.

[15] E. Celledoni, H. Marthinsen and B. Owren. *An introduction to Lie group integrators – basics, new developments and applications*, Journal of Computational Physics **257** (2014), Part B: 1040-1061.

[16] E. Celledoni, A. Marthinsen and B. Owren. *Commutator-free Lie group methods*, Future Generation Computer Systems, **19** (2003), 341–352.

[17] C. Altafini. *The de Casteljau algorithm on SE(3)*, Nonlinear control in the year 2000, Lecture Notes in Control and Information Sciences, **258** (2001), 23-34, Springer-Verlag.

NTNU

Kunnskap for en bedre verden