Master's thesis

2022

Maja Bjørge Mathiassen

Master's thesis

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

NTNU
Norwegian University of Science and Technology

Maja Bjørge Mathiassen

Coordinate maps for Lie group integrators applied to mechanical systems

December 2022

NTNU
Norwegian University of Science and Technology

NTNU
Norwegian University of
Science and Technology

Coordinate maps for Lie group integrators applied to mechanical systems

Maja Bjørge Mathiassen

Industrial Mathematics
Submission date: December 2022
Supervisor: Brynjulf Owren

Norwegian University of Science and Technology
Department of Mathematical Sciences

Abstract

Lie group integrators are numerical methods used for solving differential equations evolving on a manifold by means of a Lie group action. This thesis introduces the type of Lie group integrators known as Runge-Kutta-Munthe-Kaas methods. It is possible to choose between different coordinate maps and Lie groups. Here we will consider on the three coordinate maps called canonical coordinates of the first and second kind and the Padé(1,1) transform. We shall apply these coordinate maps to the Lie groups $SO(3)$, $Sp(1)$, $SE(3)$, and $\widehat{Sp}(1)$. This thesis provides numerical experiments on Euler's free rigid body and the N-fold three dimensional pendulum.

Sammendrag

Liegruppeintegratorer er numeriske metoder brukt til å løse differensiallikninger som utvikler seg på mangfoldigheter via Liegruppevirkninger. I denne masteroppgaven introduseres en type Liegruppeintegrator som kalles Runge-Kutta-Munthe-Kaas-metoder. Det er mulig å bruke ulike koordinatavbildninger og Liegrupper. Her vil vi vurdere tre forskjellige koordinatavbildninger, som kalles kanoniske koordinater av første og andre slag og Padé(1,1)-transformasjonen. Vi vil bruke disse koordinatavbildningene på Liegruppene $SO(3)$, $Sp(1)$, $SE(3)$ og $\widehat{Sp}(1)$. Denne oppgaven inkluderer numeriske eksperimenter på Eulers frie stivlegeme og på N koplede pendler i tre dimensjoner.

Contents

Figures

Tables

Chapter 1

Introduction

Many physical problems can be described with Lie groups and manifolds. A simple example is Euler's free rigid body that has a configuration space made up of rotation in three dimensions. This is often portrayed using the Lie group $SO(3)$. In this case solution evolves on a sphere, which is a manifold. Many classical methods do not preserve properties of the manifold and the Lie group, like the orthogonality of $SO(3)$.

In the literature there are several different types of methods utilizing manifolds and Lie groups. In 1993, Crouch and Grossman [1] developed a method using flows of vector fields in a Lie algebra. In 1994 Lewis and Simo [2] wrote a paper on integrators on Hamiltonian systems based on Lie groups. This thesis focuses on the methods developed in four papers, two written by Munthe-Kaas in published in 1995 and 1998 [3, 4], one by Munthe-Kaas and Zanna from 1999 [5], and one by Munthe-Kaas in 1999 [6]. There is also Commutator free Lie group integrators [7], and the α method introduced by Hilber, Hughes, and Taylor [8] are generalized for Lie groups in [9–11].

An important component of a Lie group integrator is the coordinate map. The aforementioned papers by Munthe-Kaas all use the exponential map. It is however possible to consider other choices of mappings. In a paper by Müller from 2016 [12], Müller examines two different coordinate mappings, the exponential map and the Cayley transform. Müller describes how these mappings, as well as their differentials, have specific formulations for the Lie groups $SO(3)$ and $SE(3)$. These Lie groups are formed by spacial rotation and rigid body motions respectively, both in three dimensions. Additionally, Müller describes unit quaternions and dual unit quaternions, how they can be used in place of $SO(3)$ and $SE(3)$ respectively, and the specific expressions of the exponential map and the Cayley transform that arise in this setting. This thesis aims to expand upon this paper by also including canonical coordinates of the second kind and the inverses of the differentials, and examining a variation of the classical Cayley transform, which is often referred to as Padé(1,1).

This thesis starts with an introduction to Lie group integrators in chapter 2. Chapter 3 starts with introducing the Lie group $SO(3)$ and finding Lie group specific versions of three coordinate mappings known as the exponential map, canonical coordinates of the second kind, and the Cayley transform, as well as the inverses of their differentials. Then the same is done for unit quaternions. As mentioned, the configuration space of Euler's free rigid body is rotations. In chapter 4, we will see how Runge-Kutta-Munthe-Kaas methods can be applied to this particular problem using $SO(3)$ and unit quaternions, as well as the three coordinate maps. Chapter 5 covers covers the three coordinate mappings for $SE(3)$ and dual unit quaternions, before section 6 shows how this can be applied to the N-fold pendulum.

Chapter 2

Background: Lie group integrators

2.1 Manifolds, tangents, and vector fields

A d-dimensional $\mathcal{M}$ is a topological space covered by a collection of open subsets $U \subset \mathcal{M}$ called coordinate charts and one-to-one and onto maps $\varphi : U \to V \subset \mathbb{R}^d$, where V is an open, connected subset of $\mathbb{R}^d$. Then (U, φ) define coordinates on $\mathcal{M}$. $\mathcal{M}$ is a smooth manifold if the maps $\varphi'' = \varphi' \circ \varphi^{-1}$ are smooth where they are defined, i.e. on $\varphi(U \cap U')$ to $\varphi'(U \cap U')$ [13]. The choice of coordinate map φ plays an important role when preforming computations on a manifold.

An example of a manifold is $\mathbb{R}^d$, which is a d-dimensional manifold covered with a single chart. Another is the unit sphere $S^d = \{x \in \mathbb{R}^{d+1} \mid \|x\|_2 = 1\}$.

A tangent to a manifold can be defined by differentiating a curve. Let $p \in \mathcal{M}$, and let $\rho(t) \in \mathcal{M}$ be a smooth curve where $\rho(0) = p$. Then

$$\boldsymbol{v}_p = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \rho(t).$$

The set of all tangent vectors at p forms the tangent space at p, which is denoted as $T\mathcal{M}|_p$. The union of all tangent spaces at all points on a manifold is called a tangent bundle of $\mathcal{M}$, and is denoted $T\mathcal{M} = \bigcup_{p \in \mathcal{M}} T\mathcal{M}|_p$. The tangent bundle of a d-dimensional manifold is itself a manifold, with dimension $2d$.

A vector field F on $\mathcal{M}$ is a smooth function $F : \mathcal{M} \to T\mathcal{M}$ such that $F(p) \in T\mathcal{M}|_p$ for all $p \in \mathcal{M}$. The collection of all vector fields on $\mathcal{M}$ will be denoted by $\mathfrak{X}(\mathcal{M})$.

2.2 Lie groups, invariant vector fields, and Lie algebras

A Lie group G is simultaneously a group, which has a binary operation and satisfies the axioms of a group, and a manifold. One type of Lie groups are matrix

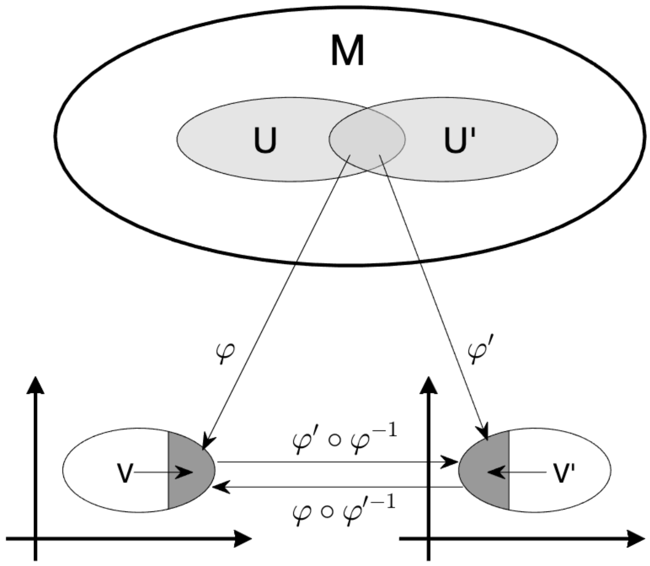

Figure 2.1: Coordinate charts on a manifold

Lie groups, where the elements of the group are matrices. Two of the Lie groups covered in this thesis, $SO(3)$ and $SE(3)$, are matrix Lie groups.

One can define left and right multiplication maps on G. Let $g, h \in G$. Then the left multiplication map $L_g : G \to G$ is given by $L_g(h) = g \cdot h$, or simply $L_g(h) = gh$, where $\cdot$ is the binary operation G is equipped with. The tangent map of the left multiplication are defined as

$$TL_g(v) = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} L_g(\gamma(t)) = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} g\gamma(t),$$

where $\gamma(t) \in G$ is a curve satisfying $\gamma(0) = h$ and $\dot{\gamma}(t) = v \in T_hG$. A vector field v is left invariant if

$$TL_g(v) = v, \quad g \in G.$$

The set of all left invariant vector fields forms a vector space. Similarly, right multiplication $R_g : G \to G$ is $R_g(h) = hg$. From this, the definition of the tangent map of right multiplication and a right invariant vector field follows.

The Lie algebra $\mathfrak{g}$ corresponding to a Lie group G can be defined as the tangent space at the identity e, in which case $\mathfrak{g} = T_eG$. A Lie algebra is equipped with a Lie bracket, which is a bilinear, skew-symmetric mapping $[\cdot,\cdot] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$, that satisfies the Jacobi identity

$$[u,[v,w]] + [w,[u,v]] + [v,[w,u]] = 0 \quad \forall u, v, w \in \mathfrak{g}.$$

An equivalent definition of the Lie algebra is that $\mathfrak{g}$ is the vector space of all left invariant vector fields on G. It is possible to construct a left invariant vector field by using TL_g to translate vectors from one vector space to another. If $X_v|_e = v \in T_e G$, then $X_v|_g = TL_g(v) \in T_g G$. Let v, w be invariant vector fields on G. Then the Lie bracket $[v, w]$ is also an invariant vector field, since

$$TL_g([v,w]) = [TL_g(v), TL_g(w)] = [v,w].$$

This can also be done using right invariant vector fields.

For matrix Lie groups, formula for Lie bracket takes the form

$$[A,B] = AB - BA, \quad A,B \in G.$$

2.3 The adjoint operator

The adjoint operator $\mathrm{Ad}_g : \mathfrak{g} \to \mathfrak{g}$ is defined as

$$\mathrm{Ad}_g(\xi) = TL_g \circ TR_{g^{-1}}(\xi)$$

for $g \in G$ and $\xi \in \mathfrak{g}$. The Lie algebra $\mathfrak{g}$ has a dual space $\mathfrak{g}*$, that refers to the set of linear functions on $\mathfrak{g}$. The dual space is connected to the Lie algebra with a duality pairing $\langle \cdot, \cdot \rangle$. Then the coadjoint operator is then defined as

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

for $\xi \in \mathfrak{g}$ and $\mu \in \mathfrak{g}^*$ [14].

The Lie bracket can be found by differentiating the adjoint operator. Let $g(t)$ be a curve such that $g(0) = e$ and $\dot{g}(0) = X$. Then

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

2.4 Coordinate mappings

In a Lie group G, the neighbourhood around zero in $\mathfrak{g}$ can be used as a coordinate map around $g \in G$. Let $\Phi : \mathfrak{g} \to G$ be a coordinate map where $\Phi(0) = e$. Then $\phi(u)g$ can be used to describe coordinates around g.

A natural choice of coordinate map is the exponential map. For a Lie group, the exponential can be defined as the flow of left invariant vector fields at initial value e. Then, for a curve $a(t) \in G$ and a $v \in \mathfrak{g}$, we have the differential equation

$$\dot{a}(t) = X_v|_{a(t)} = TL_{a(t)}(v), \quad a(0) = e.$$

Then $a(1) = \exp(v)$.

For a mapping Φ, the right trivialized tangent of said mapping is a function $d\Phi_u : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ such that

$$d\Phi_u(v) = \frac{\mathrm{d}}{\mathrm{dt}}\bigg|_{t=0} \Phi(u + tv) \cdot \Phi^{-1}(u) \tag{2.1}$$

for $u, v \in \mathfrak{g}$.

In this thesis we will require $\Phi(0) = e$ and $d\Phi_0^{-1} = e$.

2.5 Lie group actions

A Lie group acts upon a manifold with a group action $\Lambda : G \times \mathcal{M} \to \mathcal{M}$, and satisfies the following properties:

$$\begin{aligned} &\Lambda(e, x) = x \quad \forall x \in \mathcal{M}, \\ &\Lambda(g(\Lambda(h, x)) = \Lambda(g \cdot h, x) \quad \forall x \in \mathcal{M}, \forall g, h \in G. \end{aligned}$$

Common choices includes $\Lambda(g, \xi) = \mathrm{Ad}_g(\xi)$ and $\Lambda(g, \mu) = \mathrm{Ad}^*_{g^{-1}}(\mu)$, both of which will be used in the numerical examples in later chapters.

There is also a Lie algebra action $\lambda : \mathfrak{g} \times \mathcal{M} \to \mathcal{M}$. Suppose these actions are related through a coordinate map $\Phi : \mathfrak{g} \to G$ such that

$$\lambda(y, p) = \Lambda(\Phi(v), p), \quad v \in \mathfrak{g}, p \in \mathcal{M}. \tag{2.2}$$

Using the Lie algebra action, the infinitesimal generator is given by $\lambda_* : \mathfrak{g} \to \mathfrak{X}(\mathcal{M})$ as

$$\lambda_*(v)(p) = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} \lambda(tv, p)$$

for $v \in \mathfrak{g}$ and $p \in \mathcal{M}$. In practise, it is sufficient to note that since $\Phi(0) = e$, one can simply find

$$\lambda_*(v)(p) = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} \Lambda(g(t), p)$$

for a curve $g(t)$ that with $g(0) = e$ and $\dot{g}(0) = v$, thus avoiding having Φ as part of the calculations.

2.6 Lie group methods

A Lie group method aims to solve a differential equation evolving on a manifold $\mathcal{M}$,

$$\dot{y}(t) = F(y(t)), \quad y(t_0) = y_0 \in \mathcal{M}, \tag{2.3}$$

where $F \in \mathfrak{X}(\mathcal{M})$ is a smooth vector field with flow Φ.

In his paper from 1999, Munthe-Kaas [6] assumes that the vector field F is related to a map $f : \mathcal{M} \to \mathfrak{g}$ such that

$$F(y) = \lambda_*(f(y))(y).$$

We can define a vector field $\tilde{f} : \mathfrak{g} \to \mathfrak{g}$ as

$$\tilde{f}(u) = d\Phi_u^{-1}(f(\lambda(u,p)))$$

for a point $p \in \mathcal{M}$. Set $\lambda_p(u) = \lambda(u,p)$. Then $\tilde{f}$ relates to F via

$$\lambda_p' \circ \tilde{f} = F \circ \lambda_p,$$

where the composition applies to the second argument of F. Munthe-Kaas proved this in [6] with $\Phi := \exp$. In [15] there is a proof based on the one by Munthe-Kaas, with a general Φ. This result implies that (2.3) can be replaced with a differential equation $\dot{u} = \tilde{f}(u)$ on $\mathfrak{g}$. The solution of (2.3) is then $y(t) = \lambda_p(u(t))$. Recall definition of λ in (2.2). Then we arrive on the following equation

$$\dot{u} = d\Phi_u^{-1}(f(\Lambda(\Phi(u),p))), \quad u(0) = 0. \tag{2.4}$$

The equation on a Lie algebra can then be solved by a classical integration method. A Runge-Kutta-Munthe-Kaas method, or a RKMK method, solves (2.4) using a Runge-Kutta method. One step of a RKMK method can be summed up as

1. finding an approximation $u_1 \approx u(h)$ of (2.4), using one step of a RK method and $p = y_n$,
2. finding $y_{n+1} = \Phi(\sigma_1)y_n$.

The methods that are used in the numeric sections of this thesis are RKMK methods.

Chapter 3

Rotation in three dimensions

This section builds on section 2 of Müller's paper [12], which focuses on rotation in three dimensions.

3.1 SO(3)

The special orthogonal group in three dimensions is defined as $SO(3) = \{Q \in \mathbb{R}^{3\times3} | Y^T Y = \mathbb{I}_3, \det(Y) = 1\}$, where $\mathbb{I}_3$ is the 3×3 identity matrix. The corresponding Lie algebra is $\mathfrak{so}(3) = \{A \in \mathbb{R}^{3\times3} \mid A + A^T = 0\}$, the group of all skew-symmetric matrices. A matrix $\hat{\boldsymbol{u}} \in \mathfrak{so}(3)$ can be written as

$$\hat{\boldsymbol{u}} = \begin{pmatrix} 0 & -u_3 & u_2 \\ u_3 & 0 & -u_1 \\ -u_2 & u_1 & 0 \end{pmatrix}$$

and can be associated with a vector $\boldsymbol{u} = (u_1, u_2, u_3)$ via the hat map $\boldsymbol{u} \mapsto \hat{\boldsymbol{u}}$.

Let $\hat{\boldsymbol{u}}, \hat{\boldsymbol{v}} \in \mathfrak{so}(3)$. Then

$$TL_{\hat{\boldsymbol{u}}}(\hat{\boldsymbol{v}}) = \boldsymbol{u}\hat{\boldsymbol{v}}, \quad TR_{\hat{\boldsymbol{u}}}(\hat{\boldsymbol{v}}) = \hat{\boldsymbol{v}}\boldsymbol{u}. \tag{3.1}$$

It is also interesting to note that $\boldsymbol{u} \times \boldsymbol{v} = \hat{\boldsymbol{u}}\boldsymbol{v}$. Using (3.1), one can find the adjoint operator

$$\mathrm{Ad}_g(\hat{\boldsymbol{u}}) = TL_g \circ TR_{g^{-1}}(\hat{\boldsymbol{u}}) = g\hat{\boldsymbol{u}}g^{-1} = \widehat{g\boldsymbol{u}}$$

for $g \in G$. Then $\mathrm{Ad}_g(\boldsymbol{u}) = g\boldsymbol{u}$. This also allows us to identify $\mathrm{Ad}^*_{g^{-1}}(\boldsymbol{v}) = g\boldsymbol{v}$. The Lie bracket can then be found by differentiating the adjoint operator, which means that

$$[\boldsymbol{u}, \boldsymbol{v}] = \hat{\boldsymbol{u}}\boldsymbol{v} = \boldsymbol{u} \times \boldsymbol{v}.$$

Since $SO(3)$ is a matrix Lie group, the Lie bracket can also be found by

$$[\hat{\boldsymbol{u}}, \hat{\boldsymbol{v}}] = \hat{\boldsymbol{u}}\hat{\boldsymbol{v}} - \hat{\boldsymbol{v}}\hat{\boldsymbol{u}} = \hat{\boldsymbol{z}}. \tag{3.2}$$

By examining the resulting matrix $\hat{\boldsymbol{z}} \in \mathfrak{so}(3)$, it can be found that $\boldsymbol{z} = \boldsymbol{u} \times \boldsymbol{v}$. Thus $[\boldsymbol{u}, \boldsymbol{v}] = \boldsymbol{u} \times \boldsymbol{v}$, which is the same result as before.

3.1.1 The exponential map

The exponential map, also called canonical coordinates of the first kind, $\exp : \mathfrak{g} \to G$ is a coordinate map that can be defined for any Lie group G [16]. How the exact expression ends up looking might change depending on the Lie group, but it will always be possible to find an expression. This is one of the reasons the exponential map is so commonly used. A downside is that it can be computationally expensive.

Let $\hat{\boldsymbol{u}} \in \mathfrak{so}(3)$. Since this is a square matrix, it is possible to use the matrix exponent

$$\exp(\hat{\boldsymbol{u}}) = \sum_{k=0}^{\infty} \frac{1}{k!}\hat{\boldsymbol{u}}^k.$$

It is unfortunately an infinite sum. It can be approximated by using a finite number of terms, but this method is still inefficient. It would be advantageous to find an alternative that is less expensive to compute.

Note that $\hat{\boldsymbol{u}}^2 = -\alpha^2\mathbb{I}_3$. Combing this with the formula for the matrix exponential, one gets

$$\begin{aligned}\exp(\hat{\boldsymbol{u}}) &= \mathbb{I}_3 + \sum_{n=0}^{\infty} \frac{(-1)^n\alpha^{2n+1}}{(2n+1)!}\hat{\boldsymbol{u}} + \sum_{n=0}^{\infty} \frac{(-1)^n\alpha^{2n+1}}{(2n+1)!}\hat{\boldsymbol{u}}^2 \\ &= \mathbb{I}_3 + \frac{\sin\alpha}{\alpha}\hat{\boldsymbol{u}} + \frac{1-\cos\alpha}{\alpha^2}\hat{\boldsymbol{u}}^2\end{aligned}$$

where $\alpha^2 = \|\boldsymbol{u}\|_2^2$. This formula is often referred to as Rodrigues formula.

There exists a specific formula for the inverse of the differential in this setting. Then $d\exp_{\hat{\boldsymbol{u}}}^{-1}(\hat{\boldsymbol{v}}) = d\exp_{\hat{\boldsymbol{u}}}^{-1}\boldsymbol{v}$, with

$$d\exp_{\hat{\boldsymbol{u}}}^{-1} = \mathbb{I}_3 - \frac{1}{2}\hat{\boldsymbol{u}} + \frac{1-\frac{\alpha}{2}\cot\frac{\alpha}{2}}{\alpha^2}\hat{\boldsymbol{u}}^2 \tag{3.3}$$

for $u, v \in \mathfrak{so}(3)$[16]. A proof of this can be found in [17].

When implementing methods that use $\exp(\hat{\boldsymbol{u}})$ and $d\exp_{\hat{\boldsymbol{u}}}^{-1}$, it is important to keep in mind that both these formulae include division by α. This can cause problems whenever α either is or is very close to zero. A way to deal with this problem is to use Taylor series. For example, the relevant term from (3.3) can be approximated with

$$\frac{1-\frac{\alpha}{2}\cot(\frac{\alpha}{2})}{\alpha^2} = \frac{1}{12} + \frac{\alpha^2}{720} + \frac{\alpha^4}{30240} + O(\alpha^6).$$

Note that $d\exp_{\hat{\boldsymbol{u}}}^{-1}(\hat{\boldsymbol{v}})$ is the vector representation, not the skew symmetric matrix.

Keep in mind that the use of $\hat{\boldsymbol{u}}$ in $\exp(\hat{\boldsymbol{u}})$ and $d\exp_{\hat{\boldsymbol{u}}}^{-1}$ signals that these are the versions specific to $SO(3)$. This will distinguish them from other versions belonging to other Lie groups, which will be covered later on. This holds for the other coordinate maps as well.

3.1.2 Canonical coordinates of the second kind

The exponential map can also be called canonical coordinates of the first kind. Unsurprisingly, canonical coordinated of the second kind is closely related to the exponential map. Let $\boldsymbol{e}_1, ..., \boldsymbol{e}_d$ be a basis for a d-dimensional Lie algebra $\mathfrak{g}$. An element $\boldsymbol{u} \in \mathfrak{g}$ can be written as $\boldsymbol{u} = u_1\boldsymbol{e}_1 + ... + u_d\boldsymbol{e}_d$. Canonical coordinates of the second kind can be defined as

$$\mathrm{ccsk}(\boldsymbol{u}) = \prod_{i=1}^{d} \exp(u_i\boldsymbol{e}_i).$$

As previously mentioned, exponentials can be computationally expensive. However, the elements $u_i\boldsymbol{e}_i$ will consist of mainly zeros and in many cases it is possible to compute $\exp(u_i e_i)$ fairly inexpensively.

A basis for $\mathfrak{so}(3)$, written as vectors, is $\boldsymbol{e}_1 = (1,0,0)$, $\boldsymbol{e}_2 = (0,1,0)$, $\boldsymbol{e}_3 = (0,0,1)$. Then

$$\exp(u_1\hat{\boldsymbol{e}}_1) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(u_1) & -\sin(u_1) \\ 0 & \sin(u_1) & \cos(u_1) \end{pmatrix}$$

which is easily computed. The other exponentials $\exp(u_2\hat{\boldsymbol{e}}_2)$ and $\exp(u_3\hat{\boldsymbol{e}}_3)$ look similar. Let $s_i := \sin(u_i)$ and $c_i := \cos(u_i)$ for $i = 1,2,3$. Then an expression for canonical coordinates of the second kind is

$$\mathrm{ccsk}(\hat{\boldsymbol{u}}) = \begin{pmatrix} c_2c_3 & -c_2s_3 & s_2 \\ c_1s_3 + s_1s_2c_3 & c_1c_3 - s_1s_2s_3 & -s_1c_2 \\ s_1s_3 - c_1s_2c_3 & c_1s_2s_3 + c_1s_1 & c_1c_2 \end{pmatrix}.$$

Recall the definition of the differential of the coordinate map seen in (2.1). Applying this formula leads to

$$d\,\mathrm{ccsk}_{\hat{\boldsymbol{u}}}(\hat{\boldsymbol{v}}) = v_1\,\mathrm{Ad}_{\mathbb{I}_3}(\hat{\boldsymbol{e}}_1) + v_2\,\mathrm{Ad}_{\exp(u_1\hat{\boldsymbol{e}}_1)}(\hat{\boldsymbol{e}}_2) + v_2\,\mathrm{Ad}_{\exp(u_1\hat{\boldsymbol{e}}_1)}\mathrm{Ad}_{\exp(u_2\hat{\boldsymbol{e}}_2)}(\hat{\boldsymbol{e}}_3).$$

The resulting skew symmetric matrix can be reconstructed into the form $A(\boldsymbol{u})\boldsymbol{v}$, where A is a 3×3 matrix only dependent on $\boldsymbol{u}$. Then finding $d\,\mathrm{ccsk}_{\hat{\boldsymbol{u}}}(\hat{\boldsymbol{w}})$ simply means solving the equation $A(\boldsymbol{u})\boldsymbol{v} = \boldsymbol{w}$ with respect to $\boldsymbol{v}$. That means $d\,\mathrm{ccsk}_{\hat{\boldsymbol{u}}}^{-1}(\hat{\boldsymbol{w}}) = A(\boldsymbol{u})^{-1}\boldsymbol{w}$, and we get

$$d\,\mathrm{ccsk}_{\hat{\boldsymbol{u}}}^{-1} = \begin{pmatrix} 1 & \sin u_1 \tan u_2 & -\cos u_1 \tan u_2 \\ 0 & \cos(u_1) & \sin(u_1) \\ 0 & -\sin(u_1)\sec u_2 & \cos(u_1)\sec u_2 \end{pmatrix}.$$

First it is worth noticing that $\mathrm{ccsk}(\hat{\boldsymbol{u}})$, unlike $\exp(\hat{\boldsymbol{u}})$, is well defined for all values of α. Secondly, while $d\,\mathrm{ccsk}_{\hat{\boldsymbol{u}}}^{-1}$ is not defined for $\alpha = (2n+1)\pi/2$, and computations are unreliable in this neighbourhood, this is much less likely to occur than α being in a close neighbourhood of zero. A quick examination of the way an RKMK method is used will reveal that the they always includes a step where $d\Phi_{\hat{0}}^{-1}$, which clearly results in $\alpha = 0$. Thus this coordinate map will lead to much less use of approximations than the exponential requires.

3.1.3 The Cayley map

The Cayley map that will be used here is also known as Padé(1, 1), which is defined as

$$\operatorname{cay}(A) = \left(\mathbb{I}_n - \frac{1}{2}A\right)^{-1}\left(\mathbb{I}_n + \frac{1}{2}A\right)$$

where $A \in \mathbb{R}^{n\times n}$. The more standard version of the Cayley transform is defined as $(\mathbb{I}_n - A)^{-1}(\mathbb{I}_n + A)$. Whenever this version is mentioned it will be called the standard or classic Cayley transform, while Padé(1, 1) will simply be referred to as the Cayley transform.

To justify using this choice of definition, recall that RKMK methods are dependent upon the differential of the coordinate map chosen. For ease of comparison it is desirable to have similar properties for these differentials regardless of the chosen map. In particular, I require $d\Phi_0^{-1} = e$, where 0 and e are, respectively, the zero element and the identity of the chosen Lie algebra. Choosing this version of the Cayley transform ensures that this property is upheld.

The definition of the Cayley transform includes a matrix inverse. These can be computationally expensive, so it is advantageous to seek out an alternative formula. Utilizing that $(\mathbb{I}_3 - \hat{\boldsymbol{u}})^{-1} = \mathbb{I}_3 + \hat{\boldsymbol{u}} + \hat{\boldsymbol{u}}^2 + \hat{\boldsymbol{u}}^3 + \dots$ and $\hat{\boldsymbol{u}}^2 = -\alpha^2\mathbb{I}_3$ it is possible to find

$$\operatorname{cay}(\hat{\boldsymbol{u}}) = \mathbb{I}_3 + \frac{1}{1+\frac{1}{4}\alpha^2}\left(\hat{\boldsymbol{u}} + \frac{1}{2}\hat{\boldsymbol{u}}^2\right).$$

Using (2.1) we can find

$$d\operatorname{cay}_{\hat{\boldsymbol{u}}}^{-1}(\hat{\boldsymbol{v}}) = \left(\mathbb{I}_3 - \frac{\hat{\boldsymbol{u}}}{2}\right)\hat{\boldsymbol{v}}\left(\mathbb{I}_3 + \frac{\hat{\boldsymbol{u}}}{2}\right) = \mathbb{I}_3 - \frac{1}{2}[\hat{\boldsymbol{u}}, \hat{\boldsymbol{v}}] - \frac{1}{4}\hat{\boldsymbol{u}}\hat{\boldsymbol{v}}\hat{\boldsymbol{u}}. \tag{3.4}$$

Note that this point can be reached without making any assumptions about the matrices. By considering that $\hat{\boldsymbol{u}}$ and $\hat{\boldsymbol{v}}$ are skew symmetric matrices, one can find

$$d\operatorname{cay}_{\hat{\boldsymbol{u}}}^{-1} = \left(1 + \frac{1}{4}\alpha^2\right)\mathbb{I}_3 - \frac{1}{2}\hat{\boldsymbol{u}} + \frac{1}{4}\hat{\boldsymbol{u}}^2.$$

One advantage of Cayley over the exponential is that there is no division by α and therefore no need for any approximation using Taylor series.

3.2 Sp(1)

A quaternion $\boldsymbol{Q}$ is commonly represented as $\boldsymbol{Q} = q_0 + q_1\boldsymbol{i} + q_2\boldsymbol{j} + q_3\boldsymbol{k}$, where q_0, q_1, q_2, and q_3 are real numbers and $\boldsymbol{i}, \boldsymbol{j}$ and $\boldsymbol{k}$ are the unit quaternions. The unit quaternions satisfy the property $\boldsymbol{i}^2 = \boldsymbol{j}^2 = \boldsymbol{k}^2 = \boldsymbol{ijk} = -1$. This can also be written as $\boldsymbol{Q} = (q_0, q_1, q_2, q_3)$. The scalar or real part of the quaternion is q_0. The remaining numbers of the quaternion, q_1, q_2, and q_3, form the imaginary part, or the vector part, which is often written as a vector $\boldsymbol{q} = (q1, q2, q3)$. This leads to another way of representing a quaternions, which is $\boldsymbol{Q} = (q_0, \boldsymbol{q})$. These last to versions will be used throughout this thesis. The first version is useful for understanding e.g. the rules for addition and multiplication.

Let $\boldsymbol{Q}$ and $\boldsymbol{P}$ be quaternions. Then $\boldsymbol{Q}+\boldsymbol{P} = (q_0 + p_0, \boldsymbol{q} + \boldsymbol{p})$ and $\boldsymbol{QP} = (q_0 p_0 - \boldsymbol{q} \cdot \boldsymbol{p}, q_0\boldsymbol{p} + p_0\boldsymbol{q} + \boldsymbol{q} \times \boldsymbol{p})$. The inverse of a quaternion is $\boldsymbol{Q}^{-1} = \boldsymbol{Q}^*/||\boldsymbol{Q}||$, where the conjugate is $\boldsymbol{Q}^* = (q_0, -\boldsymbol{q})$ and the norm is $||\boldsymbol{Q}||^2 = q_0^2 + \boldsymbol{q} \cdot \boldsymbol{q}$.

The group of all unit quaternions, $Sp(1) = \{\boldsymbol{Q} \mid ||\boldsymbol{Q}|| = 1\}$, forms a Lie group that has identity element $\boldsymbol{e} = (1, \boldsymbol{0})$. The corresponding Lie algebra is $\mathfrak{sp}(1) = \{\boldsymbol{Q} \mid \boldsymbol{Q} + \boldsymbol{Q}^* = (0, \boldsymbol{0})\}$. These quaternions are called vector quaternions or imaginary quaternions.

Müller provides a simple relation between the Lie groups $\mathfrak{so}(3)$ and $\mathfrak{sp}(1)$ in [12]. Let $\boldsymbol{u} \in \mathbb{R}^3$ be a vector, then $\hat{\boldsymbol{u}} \in \mathfrak{so}(3)$ and $\overline{\boldsymbol{u}} = (0, \frac{1}{2}\boldsymbol{u}) \in \mathfrak{sp}(1)$. Additionally, he includes a way to go from $\boldsymbol{Q} \in Sp(1)$ to $A \in SO(3)$, given by

$$A = (q_0^2 - ||\boldsymbol{q}||_2^2)\mathbb{I}_3 + 2(q_0\hat{\boldsymbol{q}} + \boldsymbol{q}\boldsymbol{q}^T). \tag{3.5}$$

Note that there is no simple way to go from $SO(3)$ to $Sp(1)$.

For this Lie algebra, we have $TL_{\boldsymbol{Q}}(\overline{\boldsymbol{u}}) = \boldsymbol{Q}\overline{\boldsymbol{u}}$ and $TR_{\boldsymbol{Q}}(\overline{\boldsymbol{u}}) = \overline{\boldsymbol{u}}\boldsymbol{Q}$, which results in

$$\mathrm{Ad}_{\boldsymbol{Q}}(\overline{\boldsymbol{u}}) = \boldsymbol{Q}\overline{\boldsymbol{u}}\boldsymbol{Q}^*, \quad Ad^*_{\boldsymbol{Q}^{-1}}(\overline{\boldsymbol{v}}) = \boldsymbol{Q}\overline{\boldsymbol{v}}\boldsymbol{Q}^*.$$

Then the Lie bracket is

$$[\overline{\boldsymbol{u}}, \overline{\boldsymbol{v}}] = \overline{\boldsymbol{u}}\ \overline{\boldsymbol{v}} + \overline{\boldsymbol{v}}\ \overline{\boldsymbol{u}}^* = \left(0, \frac{1}{2}\boldsymbol{u} \times \boldsymbol{v}\right) = \overline{\boldsymbol{u} \times \boldsymbol{v}}. \tag{3.6}$$

Compare (3.6) to the Lie bracket for $\mathfrak{so}(3)$ in (3.2). Both Lie brackets can be written as the Lie algebra representation of the vector $\boldsymbol{u} \times \boldsymbol{v}$. This means that $\mathfrak{so}(3) \cong \mathfrak{sp}(1)$ are isomorphic Lie algebras [12]. It also allows one to argue that both $\mathfrak{so}(3)$ and $\mathfrak{sp}(1)$ are isomorphic to $\mathbb{R}^3$.

3.2.1 The exponential map

Let $\boldsymbol{a}(t) = (a_0(t), a_1(t), a_2(t), a_3(t)) \in Sp(1)$, $\boldsymbol{a}(0) = \boldsymbol{e}$, and $\overline{\boldsymbol{x}} = (0, x_1, x_2, x_3) = (0, \boldsymbol{x}) \in \mathfrak{sp}(1)$. Then $\dot{\boldsymbol{a}}(t) = TL_{\boldsymbol{a}(t)}(\overline{\boldsymbol{x}})$ and $\boldsymbol{a}(1) = \exp(\overline{\boldsymbol{x}})$. First order of business is to identify $TL_{\boldsymbol{a}(t)}(\overline{\boldsymbol{x}})$.

Let $\boldsymbol{b}(t) = (b_0(t), b_1(t), b_2(t), b_3(t))$ be a unit quaternion. We assume that $\boldsymbol{b}(0) = \boldsymbol{e}$, and $\dot{\boldsymbol{b}}(0) = \overline{\boldsymbol{x}}$. Then

$$TL_{\boldsymbol{a}}(\overline{\boldsymbol{x}}) = \frac{\mathrm{d}}{\mathrm{dt}}\bigg|_{t=0} \boldsymbol{a} \cdot \boldsymbol{b}(t) = \boldsymbol{a} \cdot \overline{\boldsymbol{x}} = A_{\overline{\boldsymbol{x}}}\boldsymbol{a}$$

where $A_{\overline{x}}$ is the 4×4 matrix

$$A_{\overline{x}} = \begin{pmatrix} 0 & -x_1 & -x_2 & -x_3 \\ x_1 & 0 & x_3 & -x_2 \\ x_2 & -x_3 & 0 & x_1 \\ x_3 & x_2 & -x_1 & 0 \end{pmatrix}.$$

This results in the differential equation $\dot{\boldsymbol{a}}(t) = A_{\overline{x}}\boldsymbol{a}(t)$, which has solution $\boldsymbol{a}(t) = \exp(tA_{\overline{x}})\boldsymbol{a}(0)$. Since $\boldsymbol{a}(1) = \exp(A_{\overline{x}})\boldsymbol{e}$ and $\exp(A_{\overline{x}})$ is a 4×4 matrix, we can use the matrix exponent to find $\exp(A_{\overline{x}})$.

Let $\overline{\boldsymbol{u}} = (0, \frac{1}{2}\boldsymbol{u}) = \overline{\boldsymbol{x}}$. Then

$$\exp(\overline{\boldsymbol{u}}) = \boldsymbol{a}(1) = \left(\cos(\alpha), \frac{\sin(\alpha)}{\alpha}\boldsymbol{u} \right) \tag{3.7}$$

for $\alpha = \|\frac{1}{2}\boldsymbol{u}\|_2$.

An interesting thing to notice is that using (3.5) to find the $SO(3)$ equivalent of $\exp(\overline{\boldsymbol{u}})$ leads to the same matrix as $\exp(\hat{\boldsymbol{u}})$. This, combined with the relation between $\mathfrak{so}(3)$ and $\mathfrak{sp}(1)$, means that

$$d\exp_{\overline{\boldsymbol{u}}}^{-1}(\overline{\boldsymbol{v}}) = \left(0, \frac{1}{2} d\exp_{\hat{\boldsymbol{u}}}^{-1}(\hat{\boldsymbol{v}})\right)$$

where $d\exp_{\hat{\boldsymbol{u}}}^{-1}(\hat{\boldsymbol{v}})$ is the inverse of the differential of the exponential for $SO(3)$ seen in (3.3).

Similarly to the exponential for $SO(3)$, the exponential for unit quaternions needs to use Taylor approximations when $\alpha \approx 0$.

In [12], Müller uses another method of finding the exponential for $\mathfrak{sp}(1)$, utilizing unitary matrices. He notes that $\mathfrak{so}(3) \cong \mathfrak{su}(2) \cong \mathfrak{sp}(1)$ and that there exists relation between an element in $SU(2)$ and an element in $Sp(1)$. The resulting unit quaternion is the same as in (3.7).

3.2.2 Canonical coordinates of the second kind

A quaternion $\boldsymbol{Q}$ can be associated with a 4×4 matrix

$$M_{\boldsymbol{Q}} = \begin{pmatrix} q_0 & -\boldsymbol{q}^T \\ \boldsymbol{q} & q_0\mathbb{I}_3 - \hat{\boldsymbol{q}} \end{pmatrix} \tag{3.8}$$

where $\hat{\boldsymbol{q}}$ is found using the hat map [12].

A basis for $\mathfrak{sp}(1)$ is $\overline{\boldsymbol{e}}_1 = (0,1,0,0)$, etc. Using (3.8) they can be written as matrices. This allows the use of the same method for identifying $\text{ccsk}(\overline{\boldsymbol{u}})$ that was used for $\mathfrak{so}(3)$.

Let $\overline{\boldsymbol{u}} = (0, u_1, u_2, u_3) \in \mathfrak{sp}(1)$. Note that $(u_i M_{\overline{e}_i})^2 = -u_i^2\mathbb{I}_4$, which means that $\exp(u_i M_{\overline{e}_i}) = \cos(u_i)\mathbb{I}_4 + \sin(u_i)M_{\overline{e}_i}$, for $i = 1, 2, 3$. This allows us to find $\text{ccsk}(M_{\overline{u}})$.

Define $s_i = \sin(u_i)$ and $c_i = \cos(u_i)$. Then, using (3.8), this can be turned back into the quaternion

$$\text{ccsk}(\overline{\boldsymbol{u}}) = (c_1c_2c_3 - s_1s_2s_3, c_1s_2s_3 + s_1c_2c_3, c_1s_2c_3 - s_1c_2s_3, c_1c_2s_3 + s_1s_2c_3).$$

Similarly to the exponent above, the orthogonal matrix version of $\text{ccsk}(\overline{\boldsymbol{q}})$, found using (3.5), is the same as $\text{ccsk}(\hat{\boldsymbol{q}})$. Then

$$d\,\text{ccsk}_{\overline{\boldsymbol{u}}}^{-1}(\overline{\boldsymbol{v}}) = \left(0, \frac{1}{2} d\,\text{ccsk}_{\hat{\boldsymbol{u}}}^{-1}(\hat{\boldsymbol{v}})\right).$$

3.2.3 Cayley transformation

As previously mentioned, the Cayley transform is defined for matrices. In the case of $SO(3)$, and any matrix Lie group, this is not a problem. Unit quaternions, unfortunately, are not matrices. They do however have a matrix representation.

Let $\overline{\boldsymbol{u}} = (0, \frac{1}{2}\boldsymbol{u}) = (0, \boldsymbol{x}) \in \mathfrak{sp}(1)$. Then $M_{\overline{\boldsymbol{u}}}^2 = -\alpha^2\mathbb{I}_4$ for $\alpha = \|\boldsymbol{x}\|_2$. Using the same method as for skew-symmetric matrices, and then turning the resulting matrix into a quaternion using (3.8), gives us

$$\text{cay}(\overline{\boldsymbol{u}}) = \left(\frac{4-\alpha^2}{4+\alpha^2}, \frac{4}{4+\alpha^2}\boldsymbol{x}\right).$$

Note that, unlike for the exponential and second kind coordinates, this version of Cayley does not correspond to the version for $\mathfrak{so}(3)$.

Recall that (3.4) can be reached without any assumptions about the matrix. By applying (3.4) to the matrix representation of a vector quaternion, and using (3.8) to extract the unit quaternion from the resulting matrix, one can find

$$d\,\text{cay}_{\overline{\boldsymbol{u}}}^{-1}(\overline{\boldsymbol{v}}) = (0, A(\boldsymbol{u})\boldsymbol{v})$$

where

$$A(\boldsymbol{u}) = \left(1 + \frac{\alpha^2}{4}\right)\mathbb{I}_3 - \hat{\boldsymbol{u}} + \frac{\hat{\boldsymbol{u}}^2}{2}.$$

Similarly to the Cayley transform for $\mathfrak{so}(3)$, there is no need for any Taylor approximations.

3.3 Computational cost and computer memory

This section, and a similar section in chapter 5, is influenced by a section in [16] which looks at computational costs of coordinate maps for $\mathfrak{so}(3)$. I have applied a similar system for counting operations, where I count addition and subtraction, multiplication and division, square roots, and trigonometrical functions. The numbers reported in the tables 3.1 and 3.2, and similar tables further on, are my best attempts at reducing costs for these coordinate maps, though there might be ways

	$\Phi(\hat{u})$			$d\Phi_{\hat{u}}^{-1}(\hat{v})$		
	exp	ccsk	cay	exp	ccsk	cay
Add/sub	15	14	15	21	4	15
Mult/div	15	4	14	24	12	18
$\sqrt{\cdot}$	1	–	–	1	–	–
Trig	2	6	–	1	4	–
Total	33	24	29	47	20	33

Table 3.1: Computational cost of the coordinate maps applied to $\mathfrak{so}(3)$.

	$\Phi(\overline{u})$			$d\Phi_{\overline{u}}^{-1}(\overline{v})$		
	exp	ccsk	cay	exp	ccsk	cay
Add/sub	2	4	4	21	4	21
Mult/div	7	12	8	27	15	19
$\sqrt{\cdot}$	1	–	–	1	–	–
Trig	2	6	–	1	4	–
Total	12	22	12	50	23	40

Table 3.2: Computational cost of the coordinate maps applied to $\mathfrak{sp}(1)$.

to bring them down further that I have not considered. For the purpose of comparing the methods I will assume the numbers are fairly accurate. Additionally I assume all four types of operations have the same computational cost, which is not necessarily the case.

For a matrix Lie group, with $n \times n$ matrices as elements, the expected computational cost for both the exponential and Cayley is Cn^3. The constant C tends to be larger for the exponential, where it often is around $20-30$ [15]. While $\mathfrak{so}(3)$ is made up of skew symmetric matrices, attempt at bringing down computational costs by careful examination of various expressions affect the resulting numbers. Nevertheless this tells us that we might expect the exponential to be more computationally expensive than the Cayley transform. Table 3.1 shows that this is in fact the case for $\mathfrak{so}(3)$. The inverse of the differential is also cheaper to compute for the Cayley transform. Canonical coordinates of the second kind is computationally cheaper than both the others.

While we have used matrices when identifying the coordinate maps for $\mathfrak{sp}(1)$, $Sp(1)$ is not a matrix Lie group. From table 3.2 we can see that both the exponential and the Cayley transform are much cheaper to compute for $\mathfrak{sp}(1)$, while the cost of second kind coordinates has changed much less. On the other hand, the inverse of the differential is slightly more expensive for all the coordinate maps. In the case of the exponential and the second kind coordinates this is to be expected, as their differentials are linked to the differentials from $\mathfrak{so}(3)$.

When focusing solely on keeping the computational costs associated with the computations of coordinate maps and their differentials, it seems that $\mathfrak{sp}(1)$ is preferable to $\mathfrak{so}(3)$ for the exponential and the Cayley transform. For canonical coordinates of the second kind there is very little difference.

Another issue to keep in mind is that the choice of Lie group determines how much computer memory is required to record the Lie group and Lie algebra elements. For elements of $\mathfrak{so}(3)$ and $\mathfrak{sp}(1)$ there is not much difference. When keeping in mind that elements of $\mathfrak{sp}(1)$ are vector quaternions, which means the first number in the quaternion always is zero, there are only three numbers that must be included. This is the same number as for $\mathfrak{so}(3)$, when representing a skew symmetric matrix as a vector.

The difference here is for the Lie groups. An element of $SO(3)$ contains nine numbers, while an element of $Sp(1)$ contains only four. While this is a small enough difference to be considered inconsequential, we will see an example in chapter 6 of using multiple Lie group and Lie algebra elements when modelling a physical problem. In that case keeping the memory requirements down becomes more important. When using N elements of a Lie group or Lie algebra that contains c numbers there will be c^N numbers that must to committed to the computer's memory.

Chapter 4

Euler's free rigid body

An example of an equation that can be solved numerically with the application of Lie group methods is the Euler's free rigid body equation (FRB). This equation models a rigid body that spins around a fixed point without any forces acting on it.

The differential equation for FRB is

$$\dot{\boldsymbol{m}} = -\mathcal{I}^{-1}\boldsymbol{m} \times \boldsymbol{m} \quad \boldsymbol{m}(0) = \boldsymbol{m}_0 \tag{4.1}$$

where $\boldsymbol{m}$ is the angular momentum and $\mathcal{I}$ is the inertia tensor $\mathcal{I} = \mathrm{diag}(I_1, I_2, I_3)$. Examination if the differential equation reveals that $\|\boldsymbol{m}\|_2 = (m_1^2 + m_2^2 + m_3^2)^{1/2}$ is conserved. That means that the solution of (4.1) evolves on a sphere S^2 with radius $r = \|\boldsymbol{m}_0\|_2$. In other words, we wish to rotate the vector. Then it is natural to use the Lie group $SO(3)$. Or, as we have previously seen, $Sp(1)$. In any case, the Lie group action is $\Lambda(g, u) = Ad^*_{g^{-1}}(u)$, though how exactly that ends up looking depends on the Lie group choice.

We start by looking at $SO(3)$. Then the Lie group action is $\Lambda(g, \boldsymbol{u}) = g\boldsymbol{u}$, for $g \in SO(3)$ and $\boldsymbol{u} \in \mathbb{R}^3$ as seen in section 3.1. Once the group action is identified, it is possible to find the infinitesimal generator. Let $h(t)$ be a curve such that $h(0) = \mathbb{I}_3$ and $\dot{h}(0) = \hat{\boldsymbol{u}}$. Then

$$\lambda_*(\boldsymbol{u})(\boldsymbol{v}) = \frac{d}{dt}\Lambda(h(t), \boldsymbol{v}) = \hat{\boldsymbol{u}}\boldsymbol{v} = \boldsymbol{u} \times \boldsymbol{v}$$

and (4.1) can be written as

$$\dot{\boldsymbol{m}} = \lambda_*(f(\boldsymbol{m}))(\boldsymbol{m}) = f(\boldsymbol{m}) \times \boldsymbol{m}.$$

By comparing this with (4.1), one obtains

$$f(\boldsymbol{m}) = -\mathcal{I}^{-1}\boldsymbol{m}$$

and we have all necessary components for solving this problem with $SO(3)$.

It is also possible to solve (4.1) using unit quaternions. For $G = Sp(1)$, the group action is $\Lambda(Q,\overline{\boldsymbol{u}}) = Q\overline{\boldsymbol{u}}Q^*$. Then the infinitesimal generator becomes

$$\lambda_*(\overline{\boldsymbol{u}})(\overline{\boldsymbol{v}}) = \left(0, \frac{1}{2}\boldsymbol{u}\times\boldsymbol{v}\right)$$

for $\overline{\boldsymbol{u}} = (0, \frac{1}{2}\boldsymbol{u})$, $\overline{\boldsymbol{v}} = (0, \frac{1}{2}\boldsymbol{v}) \in \mathfrak{sp}(1)$. Recall that $\mathbb{R}^3 \cong \mathfrak{so}(3) \cong \mathfrak{sp}(1)$, as well as the relation between elements of these two Lie algebras. Since $m \in S^2 \subset \mathbb{R}^3$, the angular momentum can also be represented by the vector quaternion $\overline{\boldsymbol{m}} = (0, \frac{1}{2}\boldsymbol{m})$. Since (4.1) evolves on S^2, the equation can also be written as a vector quaternion. Let $f(\overline{\boldsymbol{m}}) = (0, \frac{1}{2}f')$. Then

$$\dot{\overline{\boldsymbol{m}}} = \left(0, -\frac{1}{2}\mathcal{I}^{-1}\boldsymbol{m}\times\boldsymbol{m}\right) = \left(0, \frac{1}{2}f'\times\boldsymbol{m}\right)$$

which results in

$$f(\overline{\boldsymbol{m}}) = \left(0, -\frac{1}{2}\mathcal{I}^{-1}\boldsymbol{m}\right).$$

4.1 Numerical experiments

The Runge Kutta (RK) methods that is used in this thesis are Euler's method, Heun's method, and RK4[1]. The resulting RKMK methods will be referred to as Lie Euler (LE), Lie Heun (LH), and RKMK4, respectively.

All numerical calculations on FRB was done using initial value $\boldsymbol{m}_0 = (3,4,3)$, inertia tensor $\mathcal{I} = \mathrm{diag}(1,2,3)$, and time interval $t_0 = 0$ to $t_f = 5$. Figure 4.1 shows the numerical solution of (4.1) using these values. This figure also includes a sphere with radius $\|\boldsymbol{m}_0\|_2$ to illustrate that the solution does evolve on a sphere[2].

Figure 4.2 show how the global errors for solutions found using different combinations of Lie groups and coordinate mappings develop for different step lengths h for the different solution methods. The step lengths used was determined by setting $k_i \in [k_0, k_f] = [2,3,...,14]$. Then the number of steps are $N_i = 2^{k_i}$ and the step length is $h_i = (t_f - t_0)/N_i$.For each i, the problem was solved using h_i and the solution in the final step, $\boldsymbol{m}_{fi}$, was compared to a reference solution, $\boldsymbol{m}_{ref}$. The reference solution was calculated numerically using $\Phi = \exp$ and $N = 10\cdot 2^{k_f}$. The error $\|\boldsymbol{m}_{fi} - \boldsymbol{m}_{ref}\|_2$ was plotted on a logarithmic scale.

When implemented correctly, a RKMK method inherits the order of the RK method used. Figure 4.2 includes a dotted line $y = x^n$, where n order of the RK method used. If the RKMK method has the correct order, the error should follow this line. Examining figure 4.2 shows that all the methods has the expected order. It is also apparent that solutions found using higher order methods starts converging towards the reference solution for larger step lengths. This is also expected behaviour.

[1]All three Runge Kutta methods are listed in appendix A.

[2]All code used can be found on GitHub at https://github.com/majabm/ThesisCode.git

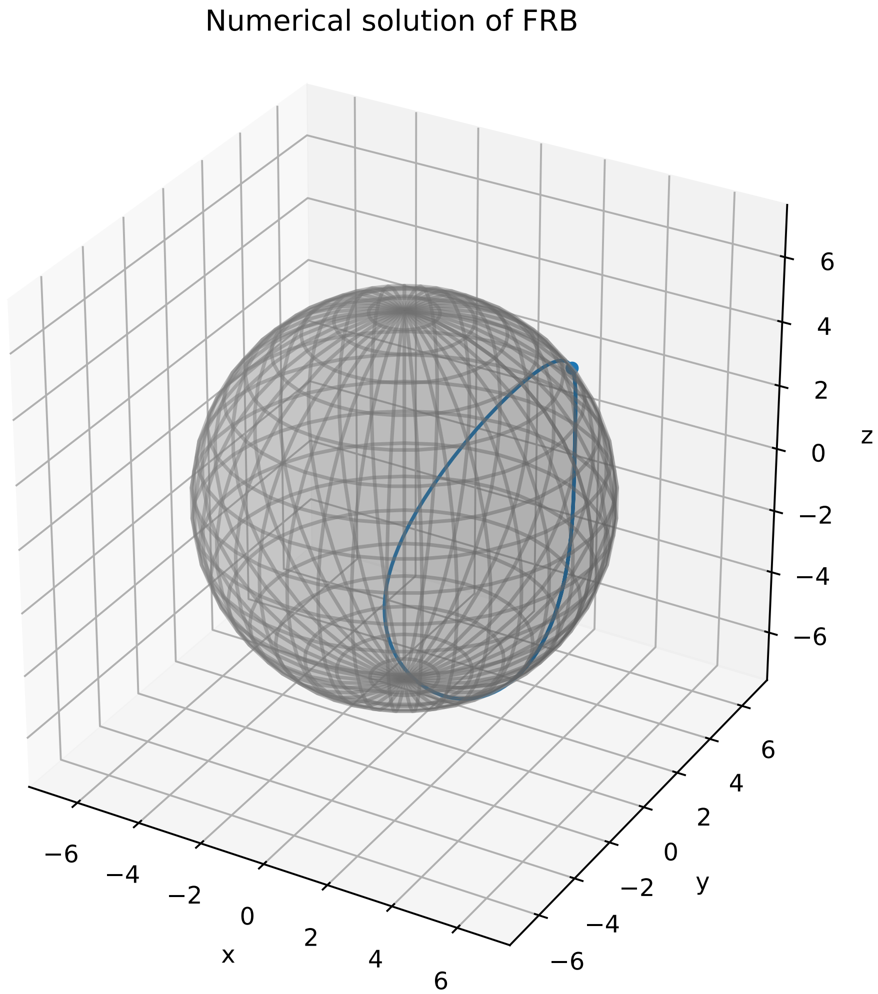

Figure 4.1: Numerical solution of FRB (blue line), solved using RKMK4, with $G = SO(3)$, $\Phi = \exp$, and $N = 1000$. The blue point is the initial solution $\boldsymbol{m}_0$. The grey sphere has radius $\|\boldsymbol{m}_0\|_2$.

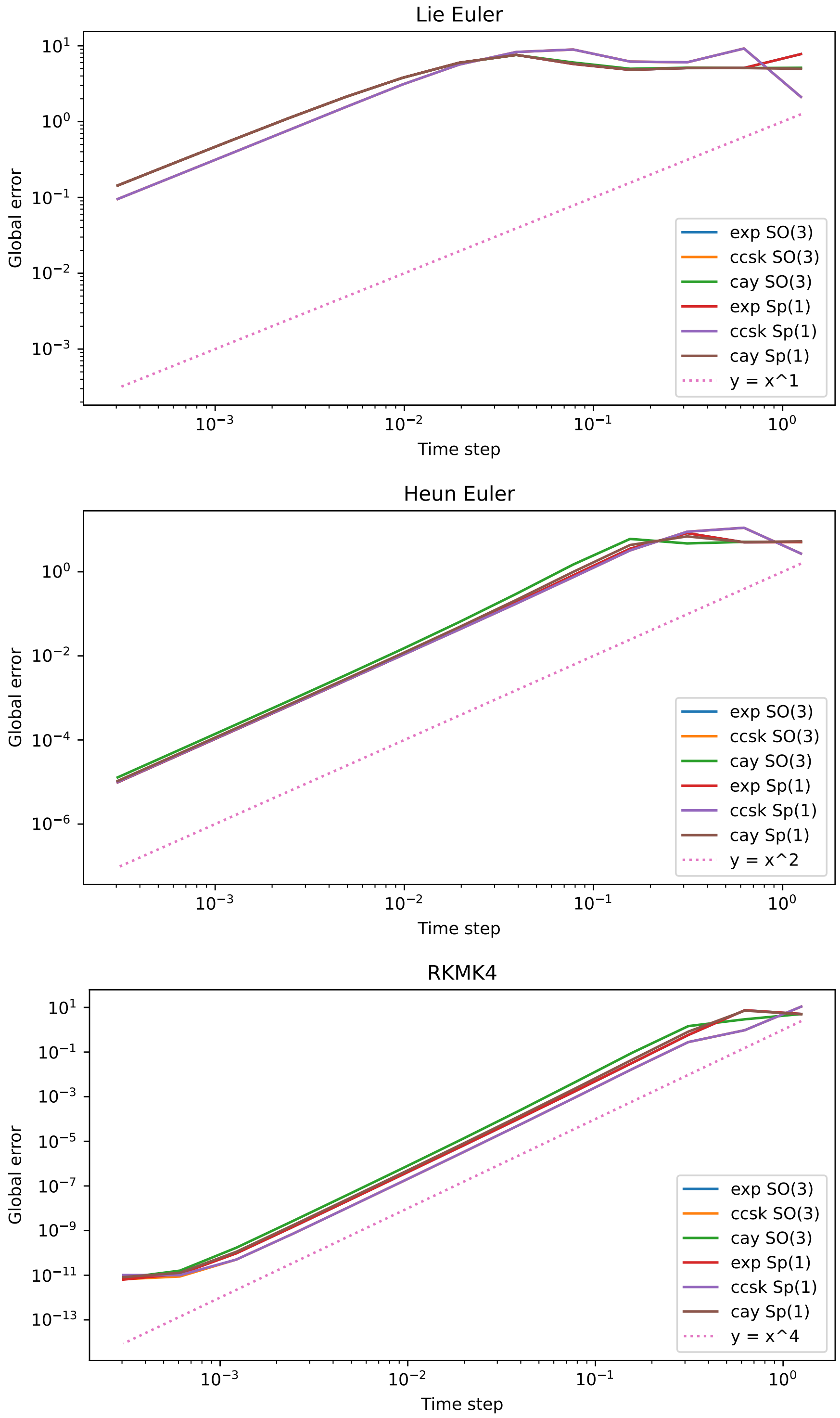

Figure 4.2: Convergence rate for the implemented methods, based on global error. The reference solution used RKMK4, $SO(3)$, and the exponential map.

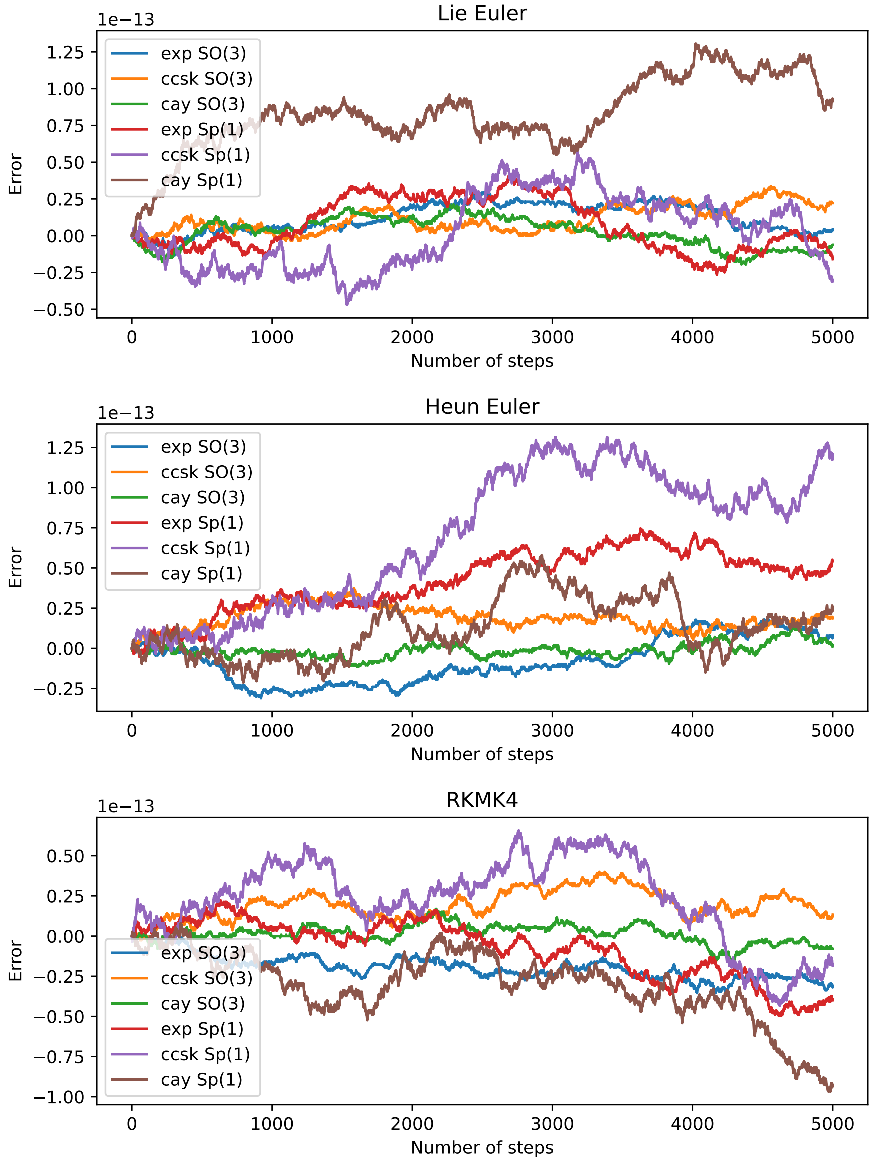

Figure 4.3: Changes in angular momentum for each step when $h = 1 \cdot 10^{-3}$.

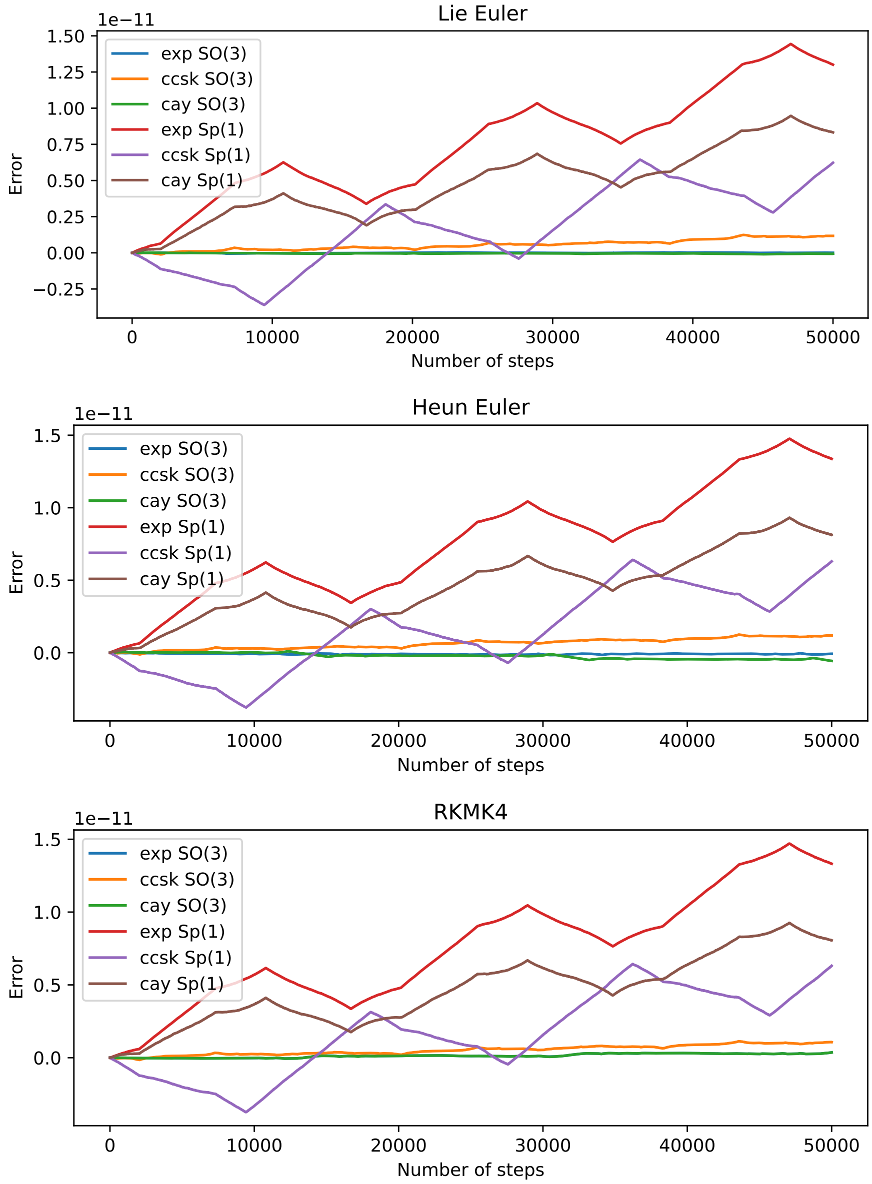

Figure 4.4: Changes in angular momentum for each step when $h = 1 \cdot 10^{-4}$.

Another area of interest is how well the numerical solution stays on the manifold. Figure 4.3 shown how far from $\|\boldsymbol{m}_0\|_2$ the Euclidean norm of the numerical solution differs for each step when the step length is $h = 1 \cdot 10^{-3}$. The error is of order 10^{-13}, which means all the methods preserve the angular momentum.

Figure 4.4 shows the same as figure 4.3, but with $h = 1 \cdot 10^{-4}$. Here the error is of order 10^{-11}, which is slightly higher than when using $h = 1 \cdot 10^{-3}$, but still small enough to claim the angular momentum is preserved. What is worth noting is the difference between the methods that use $SO(3)$ and the ones using $Sp(1)$. The error found using $Sp(1)$ tends to be larger and varies more than the error found using $SO(3)$.

Chapter 5

Rigid body motions

This section, similarly to section 3, builds on [12].

5.1 SE(3)

The special euclidean group, $SE(3)$, is often used to model rigid body motions. One way of defining this Lie group is as a semi-direct product of $SO(3)$ and $\mathbb{R}^3$, $SE(3) = SO(3) \rtimes \mathbb{R}^3$. Let $(U, \boldsymbol{u}) \in SE(3)$, where $U \in SO(3)$ and $\boldsymbol{u} \in \mathbb{R}^3$. The product of two elements is then defined

$$(U, \boldsymbol{u}) \cdot (V, \boldsymbol{v}) = (UV, U\boldsymbol{v} + \boldsymbol{u}), \quad (U, \boldsymbol{u}), (V, \boldsymbol{v}) \in SE(3).$$

The identity element of $SE(3)$ is $e = (\mathbb{I}_3, \boldsymbol{0}_3)$, where $\boldsymbol{0}_3$ is the zero vector in $\mathbb{R}^3$. The inverse of $(U, \boldsymbol{u})$ is

$$(U, \boldsymbol{u})^{-1} = (U^{-1}, -U^{-1}\boldsymbol{u}) = (U^T, -U^T\boldsymbol{u})$$

since $U^{-1} = U^T$ is a property of orthogonal matrices. The Lie algebra of $SE(3)$ is $\mathfrak{se}(3)$, which has elements $(\hat{\xi}, \boldsymbol{\eta})$. Here $\hat{\xi} \in \mathfrak{so}(3)$ and $\boldsymbol{\eta} \in \mathbb{R}^3$.

An alternative way of representing $SE(3)$ and $\mathfrak{se}(3)$ is via 4×4 matrices

$$M_{(\hat{\xi}, \eta)} = \begin{pmatrix} \hat{\xi} & \boldsymbol{\eta} \\ \boldsymbol{0}_3 & 0 \end{pmatrix} \in \mathfrak{se}(3), \qquad M_{(U,u)} \begin{pmatrix} U & \boldsymbol{u} \\ \boldsymbol{0}_3 & 1 \end{pmatrix} \in SE(3).$$

Both these versions have their uses. The matrix representation allows the use of matrix specific formulae. The other version, especially when $\hat{\xi}$ is represented as a vector, does not contain any numbers known to be either one or zero. This is why the version of $SE(3)$ that is first presented is the preferred version in this thesis.

Here $TL_{(U,\boldsymbol{u})}(\xi, \boldsymbol{\eta}) = (U\hat{\xi}, U\boldsymbol{\eta}) = (U \times \xi, U\boldsymbol{\eta})$ and $TR_{(U,\boldsymbol{u})}(\xi, \boldsymbol{\eta}) = (\xi \times U, \xi \times \boldsymbol{u} + \boldsymbol{\eta})$, which means that

$$\mathrm{Ad}_{(U,\boldsymbol{u})}(\xi, \boldsymbol{\eta}) = (U\xi, -U\xi \times \boldsymbol{u} + U\boldsymbol{\eta}) \tag{5.1}$$

and

$$[(\xi, \boldsymbol{\eta}), (\boldsymbol{\mu}, \boldsymbol{\nu})] = (\xi \times \boldsymbol{\mu}, -\boldsymbol{\mu} \times \boldsymbol{\eta} + \xi \times \boldsymbol{\nu}). \tag{5.2}$$

5.1.1 The exponential map

The preferred way of looking at $SE(3)$ means elements that are not matrices, which means using the method described for $Sp(1)$ to determine an expression for the exponential. Define a curve $(U(t), \boldsymbol{u}(t))$ where $(U(0), \boldsymbol{u}(0)) = (\mathbb{I}_3, \mathbf{0})$. Then, for $(\hat{\xi}, \boldsymbol{\eta}) \in \mathfrak{se}(3)$,

$$(\dot{U}(t), \dot{\boldsymbol{u}}(t) = TL_{(U(t),\boldsymbol{u}(t))}(\xi, \boldsymbol{\eta}) = (U(t)\hat{\xi}, U(t)\eta)$$

and $\exp(\hat{\xi}, \boldsymbol{\eta}) = (U(1), \boldsymbol{u}(1))$. The exponential map becomes

$$\exp(\hat{\xi}, \boldsymbol{\eta}) = (\exp(\hat{\xi}), \phi(\hat{\xi})\boldsymbol{\eta})$$

with $\phi(z) = (e^z - 1)/z$. We cannot use $\phi(\hat{\xi})$ directly, since $\hat{\xi}$ not invertible. Going via an interpretation using $z \in \mathbb{C}$, and by noting that $\hat{\xi}^3 = -\alpha^2\hat{\xi}$ with $\alpha = \|\xi\|_2$, we get

$$\phi(\hat{\xi}) = \left.\frac{e^z - 1}{z}\right|_{z=\hat{\xi}} = \sum_{k=0}^{\infty} \frac{\hat{\xi}^k}{(k+1)!} = \mathbb{I}_3 + \frac{1-\cos(\alpha)}{\alpha^2}\hat{\xi} + \frac{\alpha - \sin(\alpha)}{\alpha^3}\hat{\xi}^2.$$

In [14], there is an exact expression for the inverse of the differential for $\mathfrak{se}(3)$. Let $(\hat{\xi}, \boldsymbol{\eta})$ and $(\hat{\boldsymbol{\mu}}, \boldsymbol{\nu})$ be elements in $\mathfrak{se}(3)$. Then

$$d\exp^{-1}_{(\hat{\xi},\eta)}(\hat{\boldsymbol{\mu}}, \boldsymbol{\nu}) = (\boldsymbol{\zeta}, \boldsymbol{\theta}),$$

where

$$\begin{aligned}
\boldsymbol{\zeta} &= \boldsymbol{\mu} - \frac{1}{2}\xi \times (\xi \times \boldsymbol{\mu}), \\
\boldsymbol{\theta} &= \boldsymbol{\nu} - \frac{1}{2}(\boldsymbol{\eta} \times \boldsymbol{\mu} + \xi \times \boldsymbol{\nu}) + \rho g_2(\alpha)\xi \times (\xi \times \boldsymbol{\mu}) \\
&\quad + g_1(\alpha)(\boldsymbol{\eta} \times (\xi \times \boldsymbol{\mu}) + \xi \times (\boldsymbol{\eta} \times \boldsymbol{\mu}) + \xi \times (\xi \times \boldsymbol{\nu})),
\end{aligned}$$

and

$$g_1(z) = \frac{1 - \frac{z}{2}\cot\frac{z}{2}}{z^2}, \quad g_2(z) = \frac{1}{z}\frac{d}{dz}g_1(z),$$

$\rho = \boldsymbol{\xi}^T\boldsymbol{\eta}$, and $\alpha = \|\xi\|_2$.

As before, be aware that these expressions for $\exp(\hat{\xi}, \boldsymbol{\eta})$ and $d\exp^{-1}_{(\hat{\xi},\eta)}(\hat{\boldsymbol{\mu}}, \boldsymbol{\nu})$ includes division by α and that some terms might need a Taylor approximation.

5.1.2 Canonical coordinates of the second kind

Let $\boldsymbol{x} = (x_1, ..., x_6) = (\xi, \boldsymbol{\eta})$, and $\tilde{\boldsymbol{x}} = (\hat{\xi}, \boldsymbol{\eta})$. A basis for $\mathfrak{se}(3)$ is $\tilde{\boldsymbol{e}}_1, ..., \tilde{\boldsymbol{e}}_6$, where $\boldsymbol{e}_1 = (1,0,0,0,0,0)$ etc. Let $E_i = \exp(x_i M_{\tilde{e}_i})$.

It would be natural to define $\mathrm{ccsk}(\tilde{\boldsymbol{x}}) = E_1 E_2 \cdots E_6$, following the pattern from section 3. Calculating $d\,\mathrm{ccsk}^{-1}_{\tilde{x}}$ using this ordering is perfectly possible, but will unfortunately result in a fairly complex expression. However, it is not necessary to use this ordering. There are other choices that results in better versions of $d\,\mathrm{ccsk}_{\tilde{x}}^{-1}$, while not affecting complexity of the expression for the second kind coordinates in any meaningful way. I have therefore chosen to define

$$\mathrm{ccsk}(\tilde{\boldsymbol{x}}) = E_4 E_5 E_6 E_1 E_2 E_3.$$

Note that $E_i = \exp(x_i \tilde{\boldsymbol{e}}_i) = (\exp(x_i M_{\hat{e}_i}), \mathbf{0}_3)$ for $i = 1, 2, 3$, and that $E_4 E_5 E_6 = (\mathbb{I}_3, (x_4, x_5, x_6))$. Then

$$\mathrm{ccsk}(\hat{\xi}, \boldsymbol{\eta}) = (\mathrm{ccsk}(\hat{\xi}), \boldsymbol{\eta})$$

Similarly to the result for $\mathfrak{so}(3)$, the differential takes the form

$$d\,\mathrm{ccsk}_{\tilde{u}}(\tilde{\boldsymbol{v}}) = \sum_{i=1}^{6} v_i B_i(\tilde{\boldsymbol{e}}_i)$$

with $B_1 = \mathbb{I}_4$ and $B_i = \mathrm{Ad}_{\exp(u_1 \tilde{e}_1)} \dots \mathrm{Ad}_{\exp(u_{i-1} \tilde{e}_{i-1})}$, $i = 2, ..., 6$. Then

$$d\,\mathrm{ccsk}_{(\hat{\xi},\eta)}^{-1} = \begin{pmatrix} d\,\mathrm{ccsk}_{\hat{\xi}}^{-1} & 0_{3\times 3} \\ -\hat{\boldsymbol{\eta}} & \mathbb{I}_3 \end{pmatrix}$$

where $0_{3\times 3}$ is a 3×3 zero matrix.

5.1.3 The Cayley map

For $z \in \mathbb{C}$, $\mathrm{cay}(z)$ has Taylor series

$$\frac{1 + z/2}{1 - z/2} = 1 + z + \frac{1}{2} z^2 + \cdots = \sum_{k=0}^{\infty} f_k z^k$$

where f_k are the coefficients of this Taylor series. We assume a similar expression of this holds true when z is a square matrix. Recall that an element of $\mathfrak{se}(3)$ can be written as a 4×4 matrix, and let $Z = M_{(\hat{\xi},\eta)}$. Then

$$\mathrm{cay}(Z) = \frac{\mathbb{I}_4 + Z/2}{\mathbb{I}_4 - Z/2} = \begin{pmatrix} \sum_{k=0}^{\infty} f_k \hat{\xi}^k & \sum_{k=1}^{\infty} f_k \hat{\xi}^k \boldsymbol{\eta} \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} \mathrm{cay}(\hat{\xi}) & \sum_{k=1}^{\infty} f_k \hat{\xi}^k \boldsymbol{\eta} \\ 0 & 1 \end{pmatrix}.$$

Let $S(\hat{\xi}) := \sum_{k=1}^{\infty} f_k \hat{\xi}^k$. Then

$$S(\hat{\xi}) = \frac{\mathrm{cay}(\hat{\xi}) - 1}{\hat{\xi}}.$$

Note that everything until this point can be done whether or not you use the Padé(1,1) version of the Cayley transform or not, as this choice only affects which values the Taylor coefficients f_k has.

In the case of the Padé(1,1) version of the Cayley transform, it is also possible to find

$$S(\hat{\xi}) = \left(\mathbb{I}_3 - \frac{1}{2}\hat{\xi}\right)^{-1}.$$

This results in

$$\text{cay}(\hat{\xi}, \boldsymbol{\eta}) = \left(\text{cay}(\hat{\xi}), \left(\mathbb{I}_3 - \frac{1}{2}\hat{\xi}\right)^{-1} \boldsymbol{\eta}\right). \tag{5.3}$$

To avoid inverting a matrix, it is also possible to write $S(\hat{\xi})$ as

$$S(\hat{\xi}) = \mathbb{I}_3 + \frac{1}{4+\alpha^2}(\hat{\xi}^2 + A(\xi))$$

where

$$A(\xi) = \begin{pmatrix} 0 & -2\xi_3 & 2\xi_2 \\ 2\xi_3 & 0 & 2\xi_1 - 2\xi_1\xi_3 \\ -2\xi_2 & 2\xi_1 & 0 \end{pmatrix}.$$

Using (5.3), the inverse of the differential can be found the following way. Define $\phi(\hat{\xi}, \boldsymbol{\eta}) = \text{cay}(\hat{\xi}, \boldsymbol{\eta})$, $A := \mathbb{I}_3 - \frac{1}{2}\hat{\xi}$, and $B := \mathbb{I}_3 - \frac{1}{2}\hat{\xi}$. Then

$$\begin{aligned} \phi_{(\hat{\xi},\boldsymbol{\eta})}(\boldsymbol{\zeta}, \boldsymbol{\theta}) &= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \phi(\hat{\xi} + t\hat{\boldsymbol{\zeta}}, \boldsymbol{\eta} + t\boldsymbol{\theta}) \\ &= \left(A^{-1}\hat{\boldsymbol{\zeta}}A^{-1}, A^{-1}\boldsymbol{\theta} + \frac{1}{2}A^{-1}\hat{\boldsymbol{\zeta}}A^{-1}\boldsymbol{\eta}\right) \end{aligned}$$

and

$$\phi^{-1}(\hat{\xi}, \boldsymbol{\eta}) = (B^{-1}A, -B^{-1}\boldsymbol{\eta}).$$

Note that $B - 2\mathbb{I}_3 = -A$ and $A^{-1}B^{-1} = B^{-1}A^{-1}$. The differential of second kind coordinates becomes

$$\begin{aligned} d\,\text{cay}_{(\hat{\xi},\boldsymbol{\eta})}(\hat{\boldsymbol{\zeta}}, \boldsymbol{\theta}) &= \phi_{(\hat{\xi},\boldsymbol{\eta})}(\hat{\boldsymbol{\zeta}}, \boldsymbol{\theta})\phi^{-1}(\hat{\xi}, \boldsymbol{\eta}) \\ &= \left(A^{-1}\hat{\boldsymbol{\zeta}}B^{-1}, A^{-1}\boldsymbol{\theta} - \frac{1}{2}A^{-1}\hat{\boldsymbol{\zeta}}B^{-1}\boldsymbol{\eta}\right) =: (\hat{\boldsymbol{\mu}}, \boldsymbol{\nu}). \end{aligned}$$

Solving this for $(\boldsymbol{\zeta}, \boldsymbol{\theta})$ in terms of $(\hat{\boldsymbol{\mu}}, \boldsymbol{\nu})$, and substituting A and B results in

$$d\,\text{cay}^{-1}_{(\hat{\xi},\boldsymbol{\eta})}(\hat{\boldsymbol{\mu}}, \boldsymbol{\nu}) = \left(\left(\mathbb{I}_3 - \frac{1}{2}\hat{\xi}\right)\hat{\boldsymbol{\mu}}\left(\mathbb{I}_3 + \frac{1}{2}\hat{\xi}\right), \left(\mathbb{I}_3 + \frac{1}{2}\hat{\xi}\right)\left(\boldsymbol{\nu} + \frac{1}{2}\hat{\boldsymbol{\mu}}\boldsymbol{\eta}\right)\right).$$

Note that in this particular result includes the skew symmetric matrix, not the vector version of the matrix.

Details about the standard version of the Cayley transform can be found in [18].

5.2 Unit dual quaternions

Dual quaternions build on quaternions using dual numbers. A dual quaternion can be written as $\check{\boldsymbol{Q}} = \boldsymbol{Q} + \varepsilon \boldsymbol{Q}_\varepsilon = (\boldsymbol{Q}, \boldsymbol{Q}_\varepsilon)$, where $\boldsymbol{Q}$ and $\boldsymbol{Q}_\varepsilon$ are quaternions and $\varepsilon^2 = 0$. We call ε the dual unit. A dual quaternion can also be written as $\check{\boldsymbol{Q}} = (q_0 + \varepsilon q_{\varepsilon 0}, \boldsymbol{q} + \varepsilon \boldsymbol{q}_\varepsilon)$ or $\check{\boldsymbol{Q}} = (q_0, q_1, q_2, q_3, q_{\varepsilon 0}, q_{\varepsilon 1}, q_{\varepsilon 2}, q_{\varepsilon 3})$. It is more common to denote a dual quaternion as $\hat{\boldsymbol{Q}}$, but doing so here, risks confusing dual quaternions with skew symmetric matrices.

Let $\check{\boldsymbol{Q}}, \check{\boldsymbol{P}}$ be dual quaternions. Then $\check{\boldsymbol{Q}} + \check{\boldsymbol{P}} = (\boldsymbol{Q} + \boldsymbol{P}, \boldsymbol{Q}_\varepsilon + \boldsymbol{P}_\varepsilon)$ and $\check{\boldsymbol{Q}}\check{\boldsymbol{P}} = (\boldsymbol{Q}\boldsymbol{P}, \boldsymbol{Q}\boldsymbol{P}_\varepsilon + \boldsymbol{Q}_\varepsilon \boldsymbol{P})$. A dual quaternion has conjugate $\check{\boldsymbol{Q}}^* = (\boldsymbol{Q}^*, \boldsymbol{Q}_\varepsilon^*)$, norm $\|\check{\boldsymbol{Q}}\| = \check{\boldsymbol{Q}}\check{\boldsymbol{Q}}^* = (\boldsymbol{Q}\boldsymbol{Q}^*, \boldsymbol{Q}\boldsymbol{Q}_\varepsilon^* + \boldsymbol{Q}_\varepsilon \boldsymbol{Q}^*)$, and inverse: $\check{\boldsymbol{Q}}^{-1} = \check{\boldsymbol{Q}}^* / \|\check{\boldsymbol{Q}}\|$.

Dual quaternions, i.e. dual quaternions with norm equal to one, form a Lie group $\widehat{Sp}(1)$ with the binary operation being dual quaternion multiplication [12]. This group has identity element $\check{\boldsymbol{e}} = (\boldsymbol{e}, \boldsymbol{0})$, where $\boldsymbol{e}$ is the identity element from $Sp(1)$ and $\boldsymbol{0} = (0, \boldsymbol{0}_3)$ is the zero quaternion.

The Lie algebra is $\widehat{\mathfrak{sp}}(1) = \{\check{\boldsymbol{Q}} \mid \check{\boldsymbol{Q}} + \check{\boldsymbol{Q}}^* = \check{\boldsymbol{0}}\}$, which is all vector dual quaternions. They take the form $\check{\bar{\boldsymbol{Q}}} = (0, \boldsymbol{q}) + \varepsilon(0, \boldsymbol{q}_\varepsilon)$. For elements in the Lie algebra, I will also use the notation $\check{\bar{\boldsymbol{Q}}} = (0, \boldsymbol{q} + \varepsilon \boldsymbol{q}_\varepsilon)$.

Let $(\boldsymbol{\xi}, \boldsymbol{\eta}) \in \mathbb{R}^6$, then $(\hat{\boldsymbol{\xi}}, \boldsymbol{\eta}) \in SE(3)$ and $\frac{1}{2}(0, \boldsymbol{\xi} + \varepsilon \boldsymbol{\eta}) \in \widehat{\mathfrak{sp}}(1)$ are equivalent. In this setting $TL_{\check{\boldsymbol{Q}}}(\check{\bar{\boldsymbol{P}}}) = \check{\boldsymbol{Q}}\check{\bar{\boldsymbol{P}}}$ and $TR_{\check{\boldsymbol{Q}}}(\check{\bar{\boldsymbol{P}}}) = \check{\bar{\boldsymbol{P}}}\check{\boldsymbol{Q}}$ which results in

$$\mathrm{Ad}_{\check{\boldsymbol{Q}}}(\check{\bar{\boldsymbol{P}}}) = \check{\boldsymbol{Q}}\check{\bar{\boldsymbol{P}}}\check{\boldsymbol{Q}}^*$$

and

$$\begin{aligned} [\check{\bar{\boldsymbol{Q}}}, \check{\bar{\boldsymbol{P}}}] &= \check{\bar{\boldsymbol{Q}}}\,\check{\bar{\boldsymbol{P}}} + \check{\bar{\boldsymbol{P}}}\,\check{\bar{\boldsymbol{Q}}}^* \\ &= \frac{1}{2}(0, \boldsymbol{q} \times \boldsymbol{q} + \varepsilon(\boldsymbol{q} \times \boldsymbol{q}_\varepsilon - \boldsymbol{q} \times \boldsymbol{q}_\varepsilon)) =: \frac{1}{2}(0, \boldsymbol{r} + \varepsilon \boldsymbol{r}_\varepsilon). \end{aligned}$$

Comparing this to the Lie bracket for $\mathfrak{se}(3)$ in (5.2), and letting Q, P be the corresponding $\mathfrak{se}(3)$ elements of $\check{\bar{\boldsymbol{Q}}}, \check{\bar{\boldsymbol{P}}}$, $[Q, P] = (\hat{\boldsymbol{r}}, r_\varepsilon)$, while $[\check{\bar{\boldsymbol{Q}}}, \check{\bar{\boldsymbol{P}}}] = \check{\overline{(r, r_\varepsilon)}}$, in which case we can conclude that $\mathfrak{se}(3) \cong \widehat{\mathfrak{sp}}(1)$. It is also possible to argue that $\mathbb{R}^6 \cong \mathfrak{se}(3)$.

5.2.1 The exponential map

In [12], the exponential map for $\mathfrak{se}(3)$ is given as

$$\begin{aligned} \exp(0, \xi + \varepsilon \boldsymbol{\eta}) = &\left(\cos(\alpha), \frac{\sin(\alpha)}{\alpha} \xi \right) \\ &+ \varepsilon \left(-\xi \cdot \boldsymbol{\eta} \frac{\sin(\alpha)}{\alpha}, \frac{\xi \cdot \boldsymbol{\eta}}{\alpha^2} \cos(\alpha) \xi - \frac{1}{\alpha^2} \frac{\sin(\alpha)}{\alpha} \hat{\xi}^2 \boldsymbol{\eta} \right) \end{aligned}$$

with $\alpha^2 = \|\xi\|$. Define $\boldsymbol{n} = \xi/\alpha$. Then

$$\begin{aligned}\exp(0, \boldsymbol{n} + \varepsilon\boldsymbol{\eta}) =& (\cos(\alpha), \sin(\alpha)\boldsymbol{n}) \\ &+ \varepsilon\Big(-\boldsymbol{n}\cdot\boldsymbol{\eta}\sin(\alpha), \boldsymbol{n}\cdot\boldsymbol{\eta}\cos(\alpha)\boldsymbol{n} - \frac{\sin(\alpha)}{\alpha}\hat{\boldsymbol{n}}^2\boldsymbol{\eta}\Big).\end{aligned}$$

When necessary, the term $\sin(\alpha)/\alpha$ can be approximated by a Taylor series. For sufficiently small α, $\boldsymbol{n}$ also needs to be approximated. A simple way of doing this is to note that $\boldsymbol{n} \approx \boldsymbol{\xi}$ for small α.

Let $d\exp^{-1}_{(\hat{\xi},\eta)}(\hat{\mu}, \nu) = (\zeta, \boldsymbol{\theta})$ for $\mathfrak{se}(3)$, and let $\check{\overline{\boldsymbol{Q}}} = \frac{1}{2}(0, \xi + \varepsilon\boldsymbol{\eta})$ and $\check{\overline{\boldsymbol{P}}} = \frac{1}{2}(0, \boldsymbol{\mu} + \varepsilon\boldsymbol{\nu})$. Then

$$d\exp^{-1}_{\check{\overline{\boldsymbol{Q}}}}(\check{\overline{\boldsymbol{P}}}) = \frac{1}{2}(0, \zeta + \varepsilon\boldsymbol{\theta}).$$

5.2.2 Canonical coordinates of the second kind

Similarly to unit quaternions, dual unit quaternions can also be associated with a matrix. In this case it becomes an 8×8 matrix. Let $\check{\boldsymbol{Q}} = (\boldsymbol{Q}, \boldsymbol{Q}_\varepsilon)$ be a dual quaternion. Recall the matrix associated with a regular quaternion from equation (3.8). Then

$$M_{\check{Q}} = \begin{pmatrix} M_{\overline{Q}} & 0_{4\times 4} \\ M_{\overline{Q}_\varepsilon} & M_{\overline{Q}} \end{pmatrix}$$

where $0_{4\times4}$ is the 4×4 zero matrix.

A basis for $\widehat{\mathfrak{sp}}(1)$ is $\boldsymbol{e}_1 = (1, 0, 0, 0, 0, 0, 0, 0)$, etc. Let $\check{\overline{\boldsymbol{u}}} = (0, u_2, u_3, u_4, 0, u_6, u_7, u_8) \in \widehat{\mathfrak{sp}}(1)$ and $E_i = \exp(u_i M_{e_i})$ for $i = 1, ..., 8$. Similarly as in the section for second kind coordinates for $\mathfrak{se}(3)$ I have chosen to reorder the exponents, and will use

$$\mathrm{ccsk}(\check{\overline{\boldsymbol{u}}}) = E_5E_6E_7E_8E_1E_2E_3E_4.$$

Let $s_i = \sin(u_i)$ and $c_i = \cos(u_i)$ for $i = 1, ..., 8$. Then

$$\begin{aligned}\mathrm{ccsk}(\check{\overline{\boldsymbol{u}}}) =& (c_2c_3c_4 - s_2s_3s_4, c_2s_3s_4 + s_2c_3c_4, c_2s_3c_4 - s_2c_3s_4, c_2c_3s_4 + s_2s_3c_4, \\ & (-s_2c_3a_6 - (c_2a_7 + s_2a_8)s_3)c_4 + (-c_2s_3a_6 + (s_2a_7 - c_2a_8)c_3)s_4, \\ & (c_2c_3a_6 + (s_2a_7 - c_2a_8)s_3)c_4 + (-s_2s_3a_6 + (s_2a_8 + c_2a_7)c_3)s_4, \\ & (-s_2s_3a_6 + (s_2a_8 + c_2a_7)c_3)c_4 + (-c_2c_3a_6 + (-s_2a_7 + c_2a_8)s_3)s_4, \\ & (c_2s_3a_6 + (-s_2a_7 + c_2a_8)c_3)c_4 + (-s_2c_3a_6 - (c_2a_7 + s_2a_8)s_3)s_4).\end{aligned}$$

Let $d\,\mathrm{ccsk}^{-1}_{(\hat{\xi},\eta)}(\hat{\mu}, \nu) = (\zeta, \boldsymbol{\theta})$ for $\mathfrak{se}(3)$. Then

$$d\,\mathrm{ccsk}^{-1}_{\check{\overline{u}}}(\check{\overline{\boldsymbol{v}}}) = \frac{1}{2}(0, \zeta + \varepsilon\boldsymbol{\theta})$$

for $\check{\overline{\boldsymbol{u}}} = \frac{1}{2}(0, \xi + \varepsilon\boldsymbol{\eta})$ and $\check{\overline{\boldsymbol{v}}} = \frac{1}{2}(0, \boldsymbol{\mu} + \varepsilon\boldsymbol{\nu})$.

5.2.3 The Cayley map

Just as a quaternion is not a matrix, neither is a dual quaternion. Unlike for quaternions, however, attempting to find an expression for the Cayley transform using the matrix representation does not appear to work in this case. The resulting matrix $M_{\check{\boldsymbol{Q}}}$ has no relation on the form $(M_{\check{\boldsymbol{Q}}})^2 = c\mathbb{I}_8$ for some constant c. Nor have I been successful in identifying a similar relation or pattern for powers of $M_{\check{\boldsymbol{Q}}}$. Thus the method for finding the Cayley transform used previously is rendered unusable. This might imply that Cayley is not not defined for $\widehat{\mathfrak{sp}}(1)$, or that a different approach is necessary. I did not pursue the matter any further.

5.3 Computational cost and computer memory

Table 5.1 shows computational costs for the coordinate maps on $\mathfrak{se}(3)$. While we do not represent elements as 4×4 matrices, $SE(3)$ can still be considered a matrix Lie group. In that case we expect the exponential to be more computationally expensive than the Cayley transform, which does appear to be the case. Similarly to coordinate maps on $\mathfrak{so}(3)$, canonical coordinates of the second kind preforms the better than the two other options.

For $\widehat{\mathfrak{sp}}(1)$, the first thing to keep in mind is that we have no expression for Cayley in this case. From table 5.2, the two remaining maps both preform best in one case each. For $\Phi(\check{\boldsymbol{u}})$, the exponential map is less expensive to compute, while it is the most expensive choice for $d\Phi^{-1}_{\check{\boldsymbol{u}}}(\check{\boldsymbol{v}})$.

In this particular case, when using the exponential map, $\widehat{\mathfrak{sp}}(1)$ seems like the most appropriate choice, while second kind coordinates is computationally less expensive for $\mathfrak{se}(3)$.

For elements of $\mathfrak{se}(3)$ and $\widehat{\mathfrak{sp}}(1)$ it is necessary to keep track of six numbers. Again the difference is in the Lie groups, where $SE(3)$ requires twelve numbers and $\widehat{Sp}(1)$ requires eight.

	$\Phi(\hat{\xi}, \eta)$			$d\Phi^{-1}_{(\hat{\xi},\eta)}(\hat{\mu}, \nu)$		
	exp	ccsk	cay	exp	ccsk	cay
Add/sub	35	4	27	47	10	27
Mult/div	35	14	35	69	18	45
$\sqrt{\cdot}$	1	–	–	1	–	–
Trig	2	6	–	3	4	–
Total	73	24	62	120	32	72

Table 5.1: Computational cost of the coordinate maps applied to $\mathfrak{se}(3)$.

	$\Phi(\check{\bar{u}})$			$d\Phi^{-1}_{\check{\bar{u}}}(\check{\bar{v}})$		
	exp	ccsk	cay	exp	ccsk	cay
Add/sub	11	22	–	35	10	–
Mult/div	26	44	–	41	24	–
$\sqrt{\cdot}$	1	–	–	1	–	–
Trig	2	6	–	2	4	–
Total	40	72	–	79	38	–

Table 5.2: Computational cost of the coordinate maps applied to $\widehat{\mathfrak{sp}}(1)$.

Chapter 6

The N-fold 3D pendulum

The N-fold pendulum is a system of N connected three dimensional pendulums. This version assumes ideal spherical joints between the pendulums, and that the pendulums cannot interact with one another. The masses of the pendulums are $m_1, ..., m_N$ and the lengths are $L_1, ..., L_N$. For more details on the system of equations, see [19] for the modelling and [14] where the problem was re-framed into a Lie group setting. This section was adapted from the latter.

Pendulum i in the chain has position $\boldsymbol{q}_i$, which will lie on a sphere centered around the pendulum above, or, in the case $i = 1$, a fixed suspension point. Then $\dot{\boldsymbol{q}}_i \in T_{\boldsymbol{q}_i}S^2 = \{\boldsymbol{v} \in \mathbb{R}^3 \mid \boldsymbol{v}^T\boldsymbol{q}_i = 0\}$. The angular velocity $\boldsymbol{\omega}_i$ is given by $\dot{\boldsymbol{q}}_i = \boldsymbol{\omega}_i \times \boldsymbol{q}_i$, and since we assume that $\boldsymbol{\omega}_i^T\boldsymbol{q}_i = 0$, we get $\boldsymbol{\omega}_i \in T_{\boldsymbol{q}_i}S^2$.

In the case $N = 1$, we have a single pendulum with position $\boldsymbol{q}$ and angular velocity $\boldsymbol{\omega}$. When looking at a single pendulum, choosing $G = SO(3)$ is a natural choice. However, we want to extend this to a general $N > 1$. In that case $G = SE(3)$ is more appropriate, since all but the first pendulum in the chain have a non stationary suspension point.

The group action is defined from the adjoint operator $Ad_h : \mathfrak{se}(3) \to \mathfrak{se}(3)$ for $h \in SE(3)$, seen in (5.1). Since $\mathfrak{se}(3) \simeq \mathbb{R}^6$, we can also define the group action on $\mathbb{R}^6$. Additionally, since the group action then maps points of

$$TS^2_{|\boldsymbol{q}|} = \{(\tilde{\boldsymbol{q}}, \tilde{\boldsymbol{\omega}}) \in \mathbb{R}^3 \times \mathbb{R}^3 \mid \tilde{\boldsymbol{\omega}}^T\tilde{\boldsymbol{q}} = 0, |\tilde{\boldsymbol{q}}| = |\boldsymbol{q}|\} \subset \mathbb{R}^6$$

into points of $TS^2_{|\boldsymbol{q}|}$ for all points in $\mathbb{R}^6$, the group action can be defines as an action on $TS^2 = TS^2_{|\boldsymbol{q}|=1}$. Then $\Lambda : SE(3) \times TS^2 \to TS^2$ is

$$\Lambda((R, \boldsymbol{r}), (\boldsymbol{u}, \boldsymbol{v})) = Ad_{(R,\boldsymbol{r})}(\boldsymbol{u}, \boldsymbol{v}) = (R\boldsymbol{u}, R\boldsymbol{v} + \hat{\boldsymbol{r}}R\boldsymbol{u})$$

and the infinitesimal generator is

$$\lambda_*(\boldsymbol{u}, \boldsymbol{v})(\boldsymbol{q}, \boldsymbol{\omega}) = (\boldsymbol{u} \times \boldsymbol{q}, \boldsymbol{u} \times \boldsymbol{\omega} + \boldsymbol{v} \times \boldsymbol{q}).$$

This can be extended to a general $N > 1$. In this case the Lie group is $G = (SE(3))^N$. This group has binary operation $\cdot$ given by

$$A \cdot B := ((A_1, \boldsymbol{a}_1)(B_1, \boldsymbol{b}_1), ..., (A_N, \boldsymbol{a}_N)(B_N, \boldsymbol{b}_N))$$

where $A = (A_1, \boldsymbol{a}_1, .., A_N, \boldsymbol{a}_n)$, $B = (B_1, \boldsymbol{b}_1, ..., B_N, \boldsymbol{b}_N) \in (SE(3))^N$ and $(A_i, \boldsymbol{a}_i)(B_i, \boldsymbol{b}_i)$ is the product on $SE(3)$. In this case the group action $\Lambda : (SE(3))^N \times (TS^2)^N \to (TS^2)^N$ is

$$\begin{aligned}&\Lambda((A_1, \boldsymbol{a}_1, ..., A_N, \boldsymbol{a}_n)(\boldsymbol{q}_1, \boldsymbol{\omega}_1, ..., \boldsymbol{q}_N, \boldsymbol{\omega}_N))\\&\quad = (A_1\boldsymbol{q}_1, A_1\boldsymbol{\omega}_1 + \hat{\boldsymbol{a}}_1 A_1 \boldsymbol{q}_1, ..., A_N\boldsymbol{q}_N, A_N\boldsymbol{\omega}_N + \hat{\boldsymbol{a}}_N A_N \boldsymbol{q}_N)\end{aligned}$$

and the infinitesimal generator is

$$\lambda_*(\xi)(\boldsymbol{m}) = (\boldsymbol{u}_1 \times \boldsymbol{q}_1, \boldsymbol{u}_1 \times \boldsymbol{\omega}_1 + \boldsymbol{v}_1 \times \boldsymbol{q}_1, ..., \boldsymbol{u}_N \times \boldsymbol{q}_N, \boldsymbol{u}_N \times \boldsymbol{\omega}_N + \boldsymbol{v}_N \times \boldsymbol{q}_N)$$

for $\boldsymbol{\xi} = (\boldsymbol{u}_1, \boldsymbol{v}_1, ..., \boldsymbol{u}_N, \boldsymbol{v}_N) \in \mathfrak{se}(3)^N$ and $\boldsymbol{m} = (\boldsymbol{q}_1, \boldsymbol{\omega}_1, ..., \boldsymbol{q}_N, \boldsymbol{\omega}_N) \in (TS^2)^N$.

All the pendulums can be described with $(\boldsymbol{q}, \boldsymbol{\omega}) = (\boldsymbol{q}_1, \boldsymbol{\omega}_1, \boldsymbol{q}_2, \boldsymbol{\omega}_2, ..., \boldsymbol{q}_N, \boldsymbol{\omega}_N) \in (TS^2)^N$. In the case of a general N, the system of equations is

$$\begin{aligned}\dot{\boldsymbol{q}}_i &= \boldsymbol{\omega} \times \boldsymbol{q}_i, \quad i = 1, ..., N,\\ \dot{\boldsymbol{\omega}} &= A_{\boldsymbol{q}}^{-1}([\boldsymbol{g}_1, ..., \boldsymbol{g}_N]^T)\\ &= \begin{pmatrix} \boldsymbol{h}_1(\boldsymbol{q}, \boldsymbol{\omega}) \\ \dots \\ \boldsymbol{h}_N(\boldsymbol{q}, \boldsymbol{\omega}) \end{pmatrix} = \begin{pmatrix} \boldsymbol{a}_1(\boldsymbol{q}, \boldsymbol{\omega}) \times \boldsymbol{q}_1 \\ \dots \\ \boldsymbol{a}_N(\boldsymbol{q}, \boldsymbol{\omega}) \times \boldsymbol{q}_N \end{pmatrix}, \quad i = 1, ..., N.\end{aligned} \tag{6.1}$$

Here, $A_{\boldsymbol{q}} : T_{\boldsymbol{q}_1}S^2 \times \cdot T_{\boldsymbol{q}_N}S^2 \to T_{\boldsymbol{q}_1}S^2 \times \cdot T_{\boldsymbol{q}_N}S^2$ is a linear map $A_{\boldsymbol{q}}(\boldsymbol{\omega}) := R(\boldsymbol{q})\boldsymbol{\omega}$, where $R(\boldsymbol{q}) \in R^{3N\times 3N}$ is a symmetric block matrix

$$\begin{aligned}R(\boldsymbol{q})_{ii} &= \left(\sum_{j=i}^{N} m_j\right) L_i^2 \mathbb{I}_3 \in R^{3\times 3},\\ R(\boldsymbol{q})_{ij} &= \left(\sum_{k=j}^{N} m_k\right) L_i L_j \hat{\boldsymbol{q}}_i^T \hat{\boldsymbol{q}}_j = R(\boldsymbol{q})_{ji}^T \in R^{3\times 3}, \quad i < j.\end{aligned}$$

The elements $\boldsymbol{g}_i$ are given by

$$\boldsymbol{g}_i = \boldsymbol{g}_i(\boldsymbol{q}, \boldsymbol{\omega}) = \sum_{\substack{j=1\\ j\neq i}}^{N} M(\boldsymbol{q})_{ij} |\boldsymbol{\omega}_j|^2 \hat{\boldsymbol{g}}_i \boldsymbol{q}_j - \left(\sum_{j=i}^{N} m_j\right) g L_i \hat{\boldsymbol{q}}_i \boldsymbol{e}_3, \quad i = 1, ..., N$$

and $\boldsymbol{a}_1, ..., \boldsymbol{a}_N : (TS^2)^N \to \mathbb{R}^3$, are N functions that can be set to be

$$\boldsymbol{a}_i := \boldsymbol{q}_i \times \boldsymbol{h}_i(\boldsymbol{q}, \boldsymbol{\omega}).$$

Then

$$\lambda_*(f(\boldsymbol{q}, \boldsymbol{\omega}))(\boldsymbol{q}, \boldsymbol{\omega}) = F(\boldsymbol{q}, \boldsymbol{\omega})$$

for $F \in \mathfrak{X}((TS^2)^N)$. This allows us to identify the map $f : (TS^2)^N \to (\mathfrak{se}(3))^N$ as

$$f(\boldsymbol{q}, \boldsymbol{\omega}) = \begin{pmatrix} \boldsymbol{\omega}_1 \\ \boldsymbol{q}_1 \times h_1 \\ \cdots \\ \cdots \\ \boldsymbol{\omega}_N \\ \boldsymbol{q}_N \times h_N \end{pmatrix} \in \mathfrak{se}(3)^N.$$

This system of equations can also be solved using $\widehat{Sp}(1)$, by utilizing that $\mathbb{R}^6 \cong \mathfrak{se}(3) \cong \widehat{\mathfrak{sp}}(1)$. Let $\boldsymbol{\xi} = (\boldsymbol{u}_1, \boldsymbol{v}_1, ..., \boldsymbol{u}_N, \boldsymbol{v}_N) \in \mathfrak{se}(3)^N$ and $\boldsymbol{m} = (\boldsymbol{q}_1, \boldsymbol{\omega}_1, ..., \boldsymbol{q}_N, \boldsymbol{\omega}_N) \in (TS^2)^N$. Then $\check{\bar{\boldsymbol{\xi}}} = (A_1, ..., A_N)$ and $\check{\bar{\boldsymbol{m}}} = (B_1, ..., B_N)$ contains the dual vector quaternions $A_i = \frac{1}{2}(0, \boldsymbol{u}_i + \varepsilon \boldsymbol{v}_i)$ and $B_i = \frac{1}{2}(0, \boldsymbol{q}_i + \varepsilon \boldsymbol{\omega}_i)$ for $i = 1, ..., N$. The infinitesimal generator is

$$\lambda_*(\check{\bar{\boldsymbol{\xi}}})(\check{\bar{\boldsymbol{m}}}) = (C_1, C_2, ..., C_N)$$

where $C_i = \frac{1}{2}(0, \boldsymbol{u}_i \times \boldsymbol{q}_i + \varepsilon(\boldsymbol{u}_i \times \boldsymbol{\omega}_i + \boldsymbol{v}_i \times \boldsymbol{q}_i))$.

Let $\check{\bar{\boldsymbol{m}}}_i = \frac{1}{2}(0, \boldsymbol{q}_i + \varepsilon \boldsymbol{\omega}_i)$. The system of equations can be rewritten as

$$\dot{\check{\bar{\boldsymbol{m}}}}_i = \frac{1}{2}(0, \dot{\boldsymbol{q}}_i + \varepsilon \dot{\boldsymbol{\omega}}_i), \quad i = 1, ..., N$$

where $\dot{\boldsymbol{q}}_i$ and $\dot{\boldsymbol{\omega}}_i$ are defined in (6.1). This leads to the map

$$f(\check{\bar{\boldsymbol{q}}}, \check{\bar{\boldsymbol{\omega}}}) = (f_1, f_2, ..., f_3) \in (\widehat{\mathfrak{sp}}(1))^N,$$

where

$$f_i = \frac{1}{2}(0, \boldsymbol{\omega}_i + \varepsilon(\boldsymbol{q}_i \times \boldsymbol{h}_i)), \quad i = 1, ..., N.$$

6.1 Numerical experiments on the 2-fold pendulum

All numerical computations in this section is done using $t_0 = 0$, $t_f = 5$, $L_i = 1$, $m_i = 1$, and $g = 10$.

In addition to Taylor expansions some of the methods use, the exponential of the $\widehat{Sp}(1)$ needs an approximation for the vector $\boldsymbol{n}$. This is done using $TOL = 1e-9$, and then $\boldsymbol{n} = \boldsymbol{\xi}$ when $\alpha < TOL$.

The numerical solution of the 2-fold pendulum, found using the exponential on $\mathfrak{se}(3)$, RKMK4, and $N = 1000$, is shown in figure 6.1. Here, the initial solution is $\boldsymbol{q}_i(0) = (1/\sqrt{2}, 0, 1/\sqrt{2})$ and $\boldsymbol{\omega}_i(0) = (0, 1, 0)$ for $i = 1, 2$.

From figure 6.2, we can see that all three methods have the correct convergence rate. Similarly to figure 4.2 for FRB, the reference solution here uses the exponential map and $SE(3)$. In this case there does not appear to be much of a

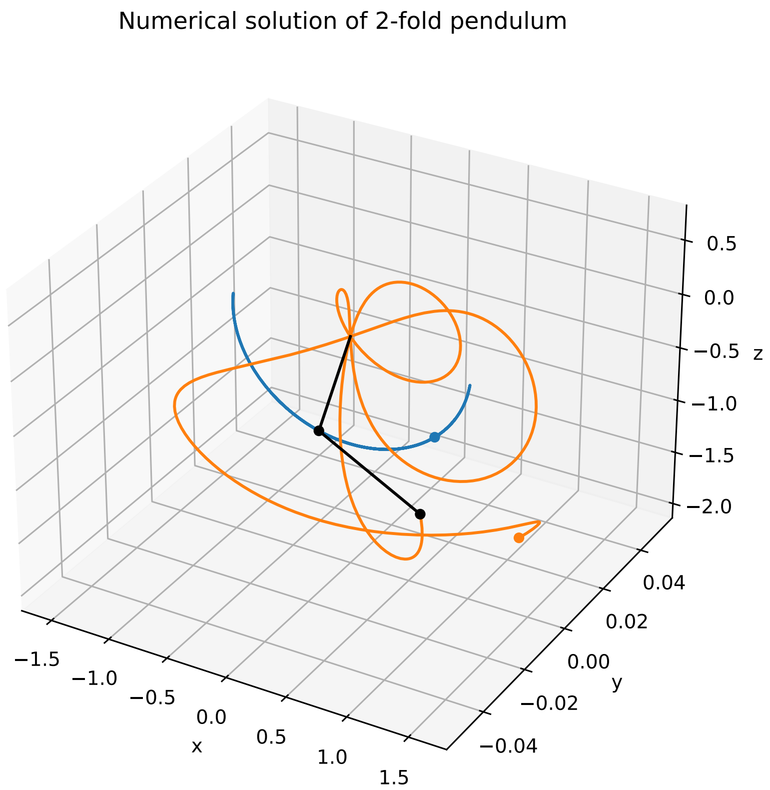

Figure 6.1: Numerical solution of the 2-fold pendulum, with $SE(3)$, RKMK4, the exponential map, and $N = 1000$. The blue and orange lines represents the paths of the pendulums, and the blue and orange points are the initial positions. The black line and points are of the pendulums at $t_f = 5$. The initial condition is $\boldsymbol{q}_i(0) = (1/\sqrt{2}, 0, 1/\sqrt{2})$, $\boldsymbol{\omega}_i(0) = (0, 1, 0)$, $i = 1, 2$.

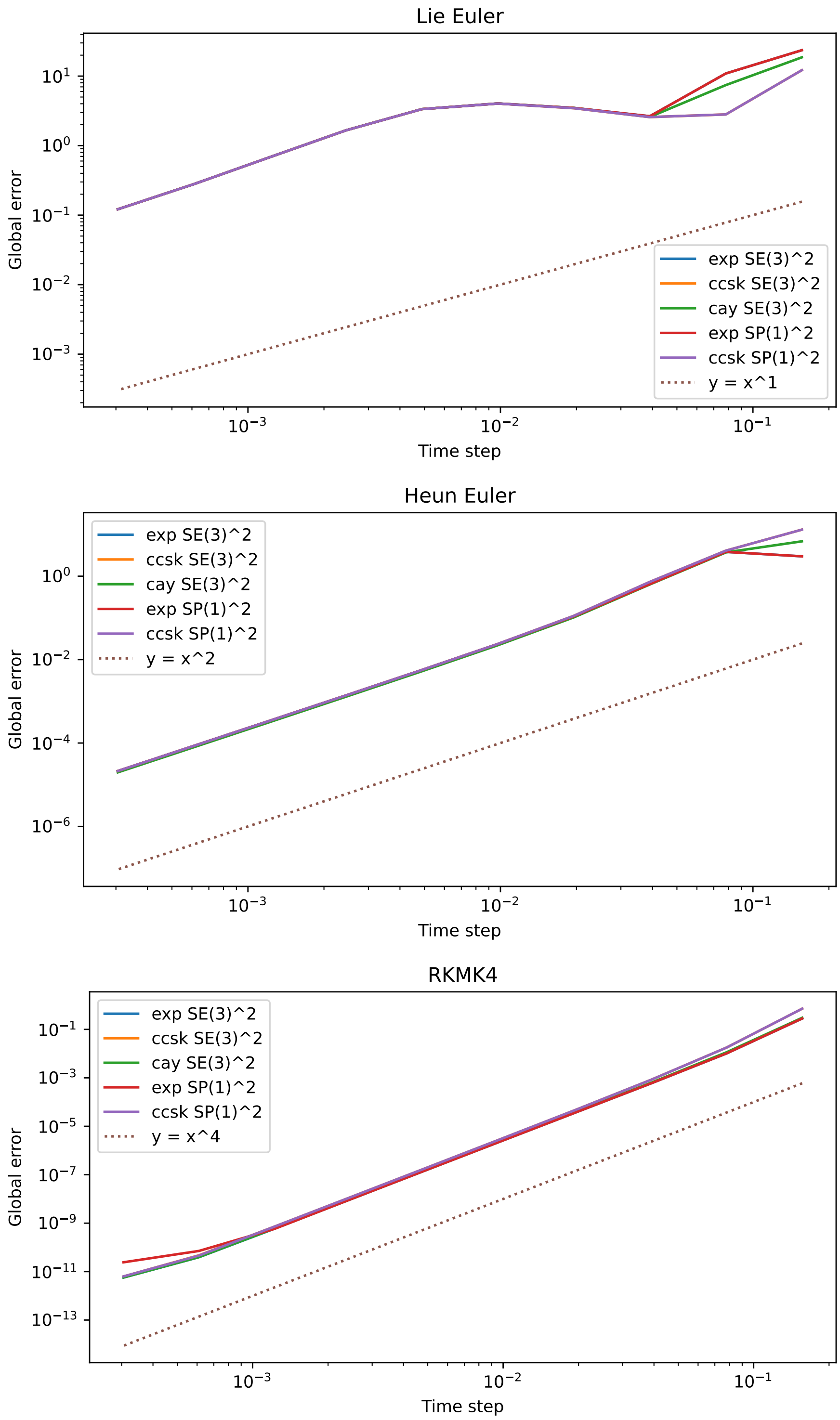

Figure 6.2: Convergence rate for the implemented methods, based on global error. The reference solution used RKMK4, $SE(3)$, and the exponential map. The initial condition is $\boldsymbol{q}_i(0) = (1/\sqrt{2}, 0, 1/\sqrt{2})$, $\boldsymbol{\omega}_i(0) = (0, 1, 0)$, $i = 1, 2$.

difference between the different choices, especially for small h. Note that in all figures, the Lie group $\widehat{Sp}(1)$ is written as $SP(1)$.

We want these methods to preserve geometries of S^2 and TS^2. In other words we require the methods to preserve the properties

- $\boldsymbol{q}_i(t)^T\boldsymbol{q}_i(t) = 1, i = 1, 2,$
- $\boldsymbol{q}_i(t)^T\boldsymbol{\omega}_i(t) = 0, i = 1, 2.$

We start with the property $\boldsymbol{q}_1^T\boldsymbol{q}_1 = 1$. Figure 6.3 and figure 6.4 show how $1 - \boldsymbol{q}_1^T\boldsymbol{q}_1$ evolves for $N = 5000$ and $N = 50000$. Here, we see the same trend as we saw for FRB. For the smaller choice of N, there is little difference between methods, and the error is of order 10^{-14}, while for the larger N the error is of order 10^{-12}, and there is a bit more difference between the methods. Since the error is very small, these methods seems to preserve the property $\boldsymbol{q}_i^T\boldsymbol{q}_i = 1$, $i = 1, 2$, i.e. the solutions stay on S^2.

Figure 6.5 and figure 6.6 show the evolution of $\boldsymbol{q}_1^T\boldsymbol{\omega}_1$. Here the error is also of order 10^{-14} for $N = 5000$ and 10^{-12} for $N = 50000$. Thus $\boldsymbol{\omega}_1 \in T_{\boldsymbol{q}_1}S^2$.

Figures showing that the properties $\boldsymbol{q}_2^T\boldsymbol{q}_2 = 1$ and $\boldsymbol{q}_2^T\boldsymbol{\omega}_2 = 0$ also are preserved, are in appendix B. They show the same overarching trends, with error between 10^{-12} and 10^{-14}.

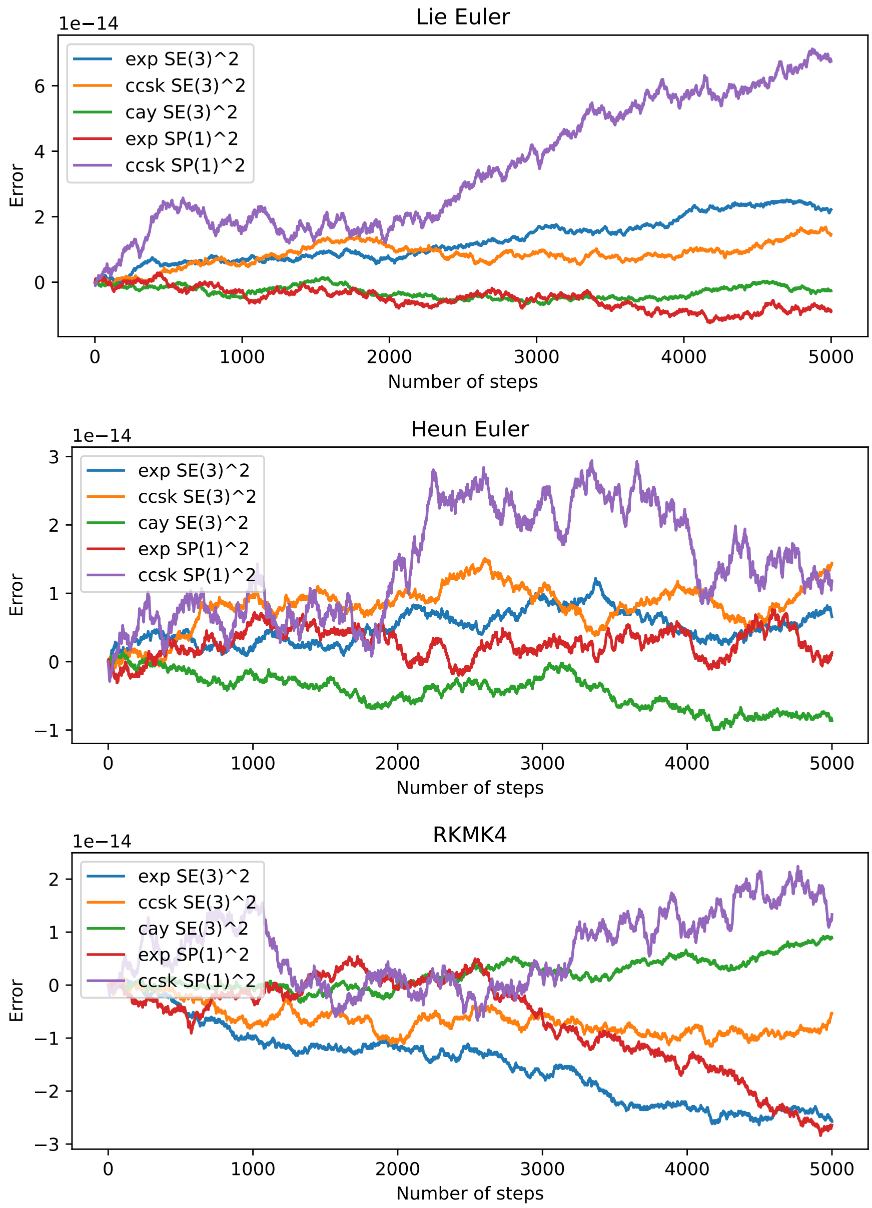

Figure 6.3: Changes in $1 - \boldsymbol{q}_1^T \boldsymbol{q}_1$ for each step when $h = 1 \cdot 10^{-3}$.The initial condition is $\boldsymbol{q}_i(0) = (0, 1, 0)$, $\boldsymbol{\omega}_i(0) = (1, 0, 1)$, $i = 1, 2$.

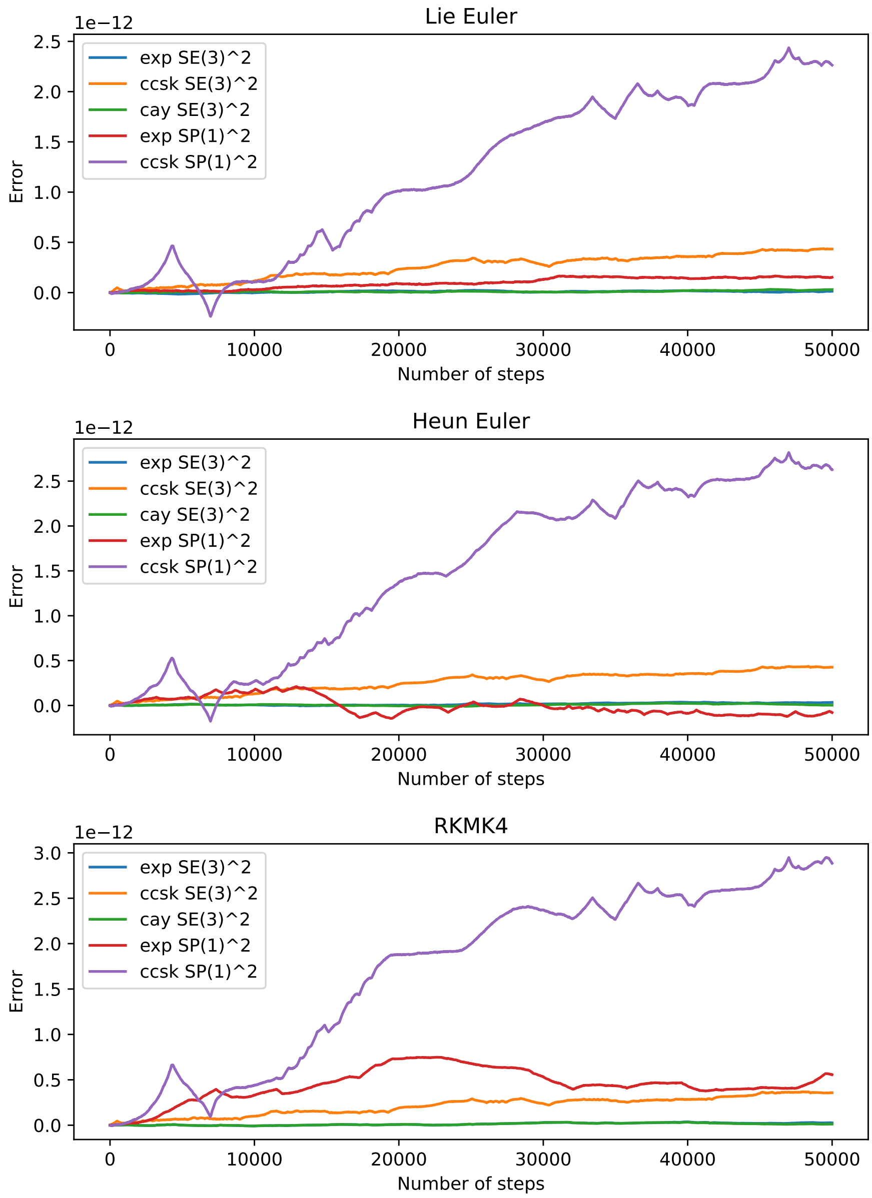

Figure 6.4: Changes in $1 - \boldsymbol{q}_1^T \boldsymbol{q}_1$ for each step when $h = 1 \cdot 10^{-4}$.The initial condition is $\boldsymbol{q}_i(0) = (0, 1, 0)$, $\boldsymbol{\omega}_i(0) = (1, 0, 1)$, $i = 1, 2$.

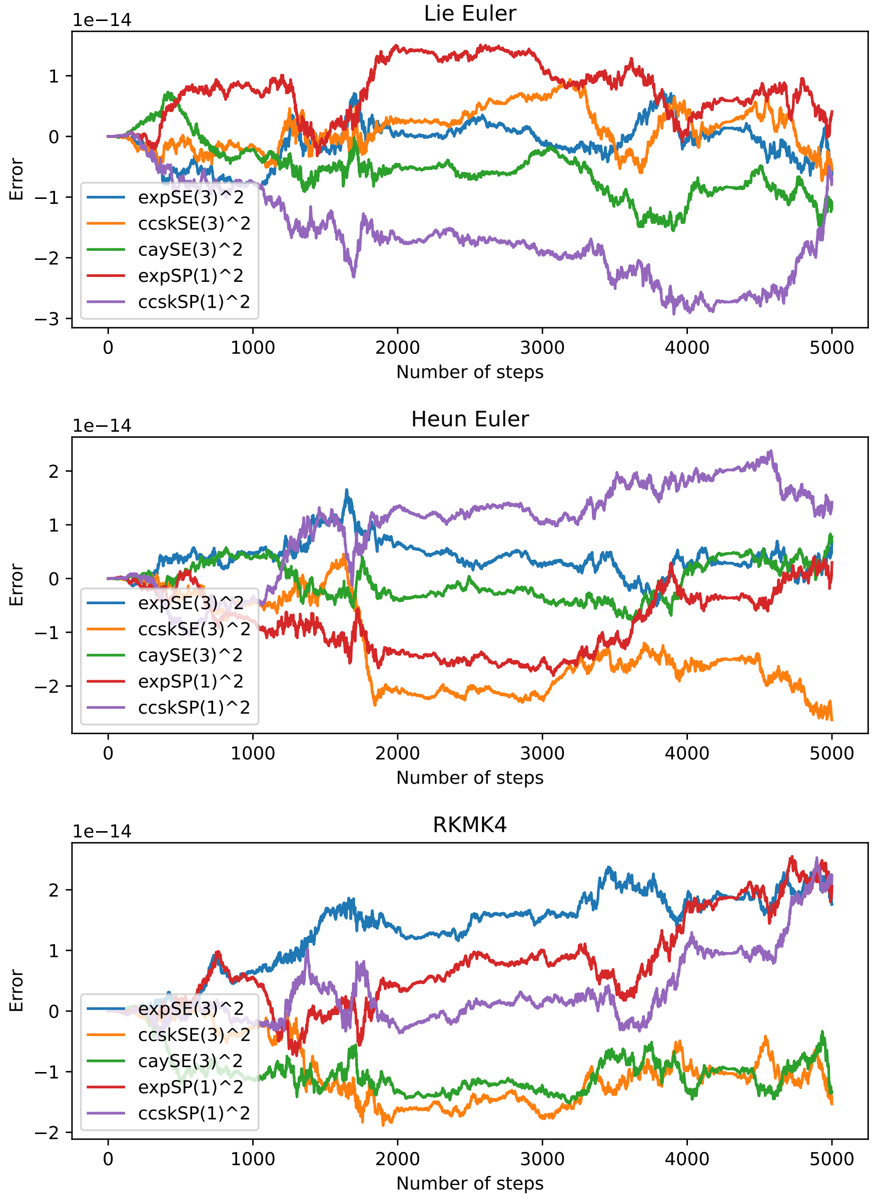

Figure 6.5: Changes in $\boldsymbol{q}_1^T \boldsymbol{\omega}_1$ for each step when $h = 1 \cdot 10^{-3}$

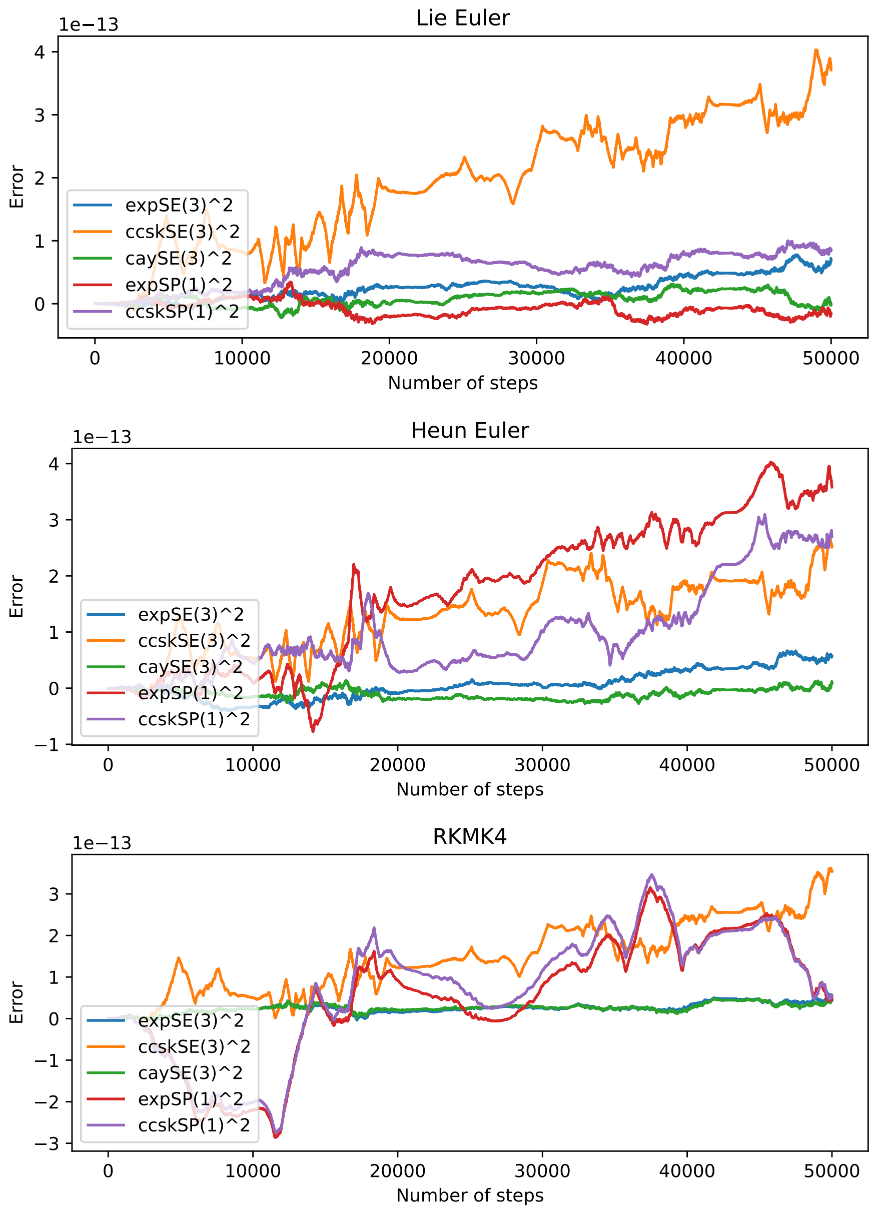

Figure 6.6: Changes in $q_1^T \omega_1$ for each step when $h = 1 \cdot 10^{-4}$

Chapter 7

Conclusion

In this thesis, I have considered how different Lie groups and coordinate mappings can be utilized in Lie group integrators. The Lie groups $SO(3)$, $Sp(1)$, $SE(3)$, and $\widehat{Sp}(1)$ can be used for rotation and rigid body motions in three dimensions. In this thesis we have seen that the exponential, canonical coordinates of the second kind, and the variant of the Cayley transform known as Padé$(1,1)$ can be applied for these Lie groups, with the possible exception of Cayley for $\widehat{Sp}(1)$.

Numerical experiments on Euler's free rigid body and the N-fold three dimensional pendulum show that the RKMK methods with relevant combinations of Lie groups and coordinate mappings preserve the geometries properties of the problems.

There has also been a brief discussion on the computations costs of the various coordinate maps, and a mention of the fact that the choice of Lie group will influence the amount of computer memory that is required. The Lie groups based on quaternions, $Sp(1)$ and $\widehat{Sp}(1)$, require less memory than $SO(3)$ and $SE(3)$. In most cases the exponential is the most computationally expensive option while canonical coordinates of the second kind is notably less expensive. However, examining only computational costs and memory requirements connected to the choice of coordinate map and Lie group is not enough to declare that some combination is always better than the others. Other elements of the Lie group method will also affect the overall performance of the numerical solver. An example is the Lie group action, which depends on the Lie group used and the mechanical system the Lie group method is applied to.

Future work on this subject could be to determine whether or not the Cayley transform exists for $\widehat{Sp}(1)$. There is also the interesting possibility that a beam can be modelled with a generalization of the N-fold pendulum.

Bibliography

[1] P. E. Crouch and R. L. Grossman, 'Numerical integration of ordinary differential equations on manifolds,' *Journal of Nonlinear Science*, vol. 3, pp. 1–33, 1993.

[2] D. Lewis and J. C. Simo, 'Conserving algorithms for the dynamics of Hamiltonian systems on Lie groups,' *Journal of Nonlinear Science*, vol. 4, pp. 253–299, 1994.

[3] H. Munthe-Kaas, 'Lie-Butcher theory for Runge-Kutta methods,' *BIT*, vol. 35, pp. 572–587, 1995.

[4] H. Z. Munthe-Kaas, 'Runge-Kutta methods on Lie groups,' *BIT Numerical Mathematics*, vol. 38, pp. 92–111, 1998.

[5] H. Munthe-Kaas and B. Owren, 'Computations in a free Lie algebra,' *Philosophical Transactions of the Royal Society of London Series A*, vol. 357, p. 957, 1999.

[6] H. Munthe-Kaas, 'High order Runge-Kutta methods on manifolds,' *Applied Numerical Mathematics*, vol. 29, pp. 115–127, 1999.

[7] E. Celledoni, A. Marthinsen and B. Owren, 'Commutator-free Lie group methods,' *Future Generation Computer Systems*, vol. 19, pp. 341–352, 2003, Special Issue on Geometric Numerical Algorithms.

[8] H. M. Hilber, T. J. R. Hughes and R. L. Taylor, 'Improved numerical dissipation for time integration algorithms in structural dynamics,' *Earthquake Engineering & Structural Dynamics*, vol. 5, pp. 283–292, 1977.

[9] O. Bruls and A. Cardona, 'On the use of Lie group time integrators in multibody dynamics,' *Journal of Computational and Nonlinear Dynamics*, vol. 5, 2010.

[10] M. Arnold and O. Brüls, 'Convergence of the generalized-α scheme for constrained mechanical systems,' *Multibody System Dynamics*, vol. 18, pp. 185–202, Aug. 2007.

[11] M. Arnold, O. Brüls and A. Cardona, 'Error analysis of generalized-α Lie group time integration methods for constrained mechanical systems,' *Numerische Mathematik*, vol. 129, pp. 149–179, 2015.

[12] A. Müller, 'Coordinate mappings for rigid body motions,' *Journal of Computational and Nonlinear Dynamics*, vol. 12, 2016.

[13] P. J. Olver, *Equivalence, Invariants, and Symmetry*. Cambridge University Press, 1995.

[14] E. Celledoni, E. Çokaj, A. Leone, D. Murari and B. Owren, 'Lie group integrators for mechanical systems,' *International Journal of Computer Mathematics*, pp. 1–31, 2021.

[15] B. Owren and A. Marthinsen, 'Integration methods based on canonical coordinates of the second kind,' *Numerische Mathematik*, vol. 87, pp. 763–790, 2001.

[16] E. Celledoni and B. Owren, 'Lie group methods for rigid body dynamics and time integration on manifolds,' *Computer Methods in Applied Mechanics and Engineering*, vol. 192, pp. 421–438, 2003.

[17] A. Iserles, 'Solving linear ordinary differential equations by exponentials of iterated commutators,' *Numerische Mathematik*, vol. 45, 1984.

[18] A. Müller, 'Review of the exponential and Cayley map on se(3) as relevant for Lie group integration of the generalized Poisson equation and flexible multibody systems,' *Proceedings of The Royal Society A*, vol. 477, 2021.

[19] T. Lee, M. Leok and N. Mcclamroch, *Global Formulations of Lagrangian and Hamiltonian Dynamics on Manifolds: A Geometric Approach to Modeling and Analysis* (Interaction of Mechanics and Mathematics). Springer, 2018.

Appendix A

Runge-Kutta methods

For a differential equation

$$\dot{y}(t) = f(t, y(t))$$

the three Runge-Kutta methods used in this paper can be written as follows. Euler's method (first order method):

$$y_{n+1} = y_n + hf(t_n, y_n),$$

Heun's method (second order method):

$$\begin{aligned}
\tilde{y}_{n+1} &= y_n + hf(t_n, y_n),\\
y_{n+1} &= y_n + \frac{h}{2}\big[f(t_n, y_n) + f(t_{n+1}, \tilde{y}_{n+1})\big],
\end{aligned}$$

RK4 (fourth order method):

$$\begin{aligned}
k_1 &= hf(t_n, y_n),\\
k_2 &= hf(t_n + \frac{h}{2}, y_n + \frac{1}{2}k_1),\\
k_3 &= hf(t_n + \frac{h}{2}, y_n + \frac{1}{2}k_2),\\
k_4 &= hf(t_{n+1}, y_n + k_3),\\
y_{n+1} &= y_n + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4).
\end{aligned}$$

Appendix B

Additional figures for section 6.1

The figures B.1, B.2, B.3, and B.4 show how $1-\boldsymbol{q}_2^T\boldsymbol{q}_2$ and $\boldsymbol{q}_2^T\boldsymbol{\omega}_2$ develop for the 2-fold pendulum.

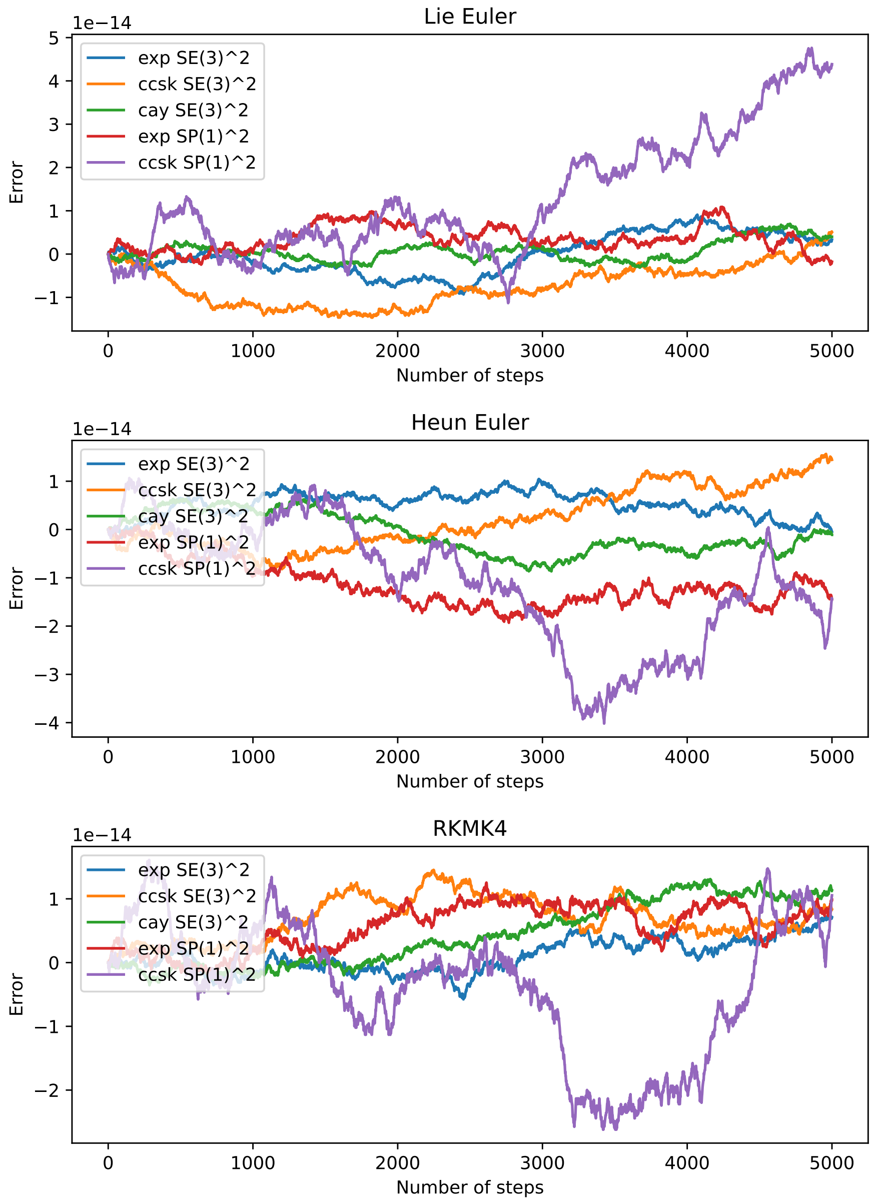

Figure B.1: Changes in $1-\boldsymbol{q}_2^T\boldsymbol{q}_2$ for each step when $h = 1 \cdot 10^{-3}$

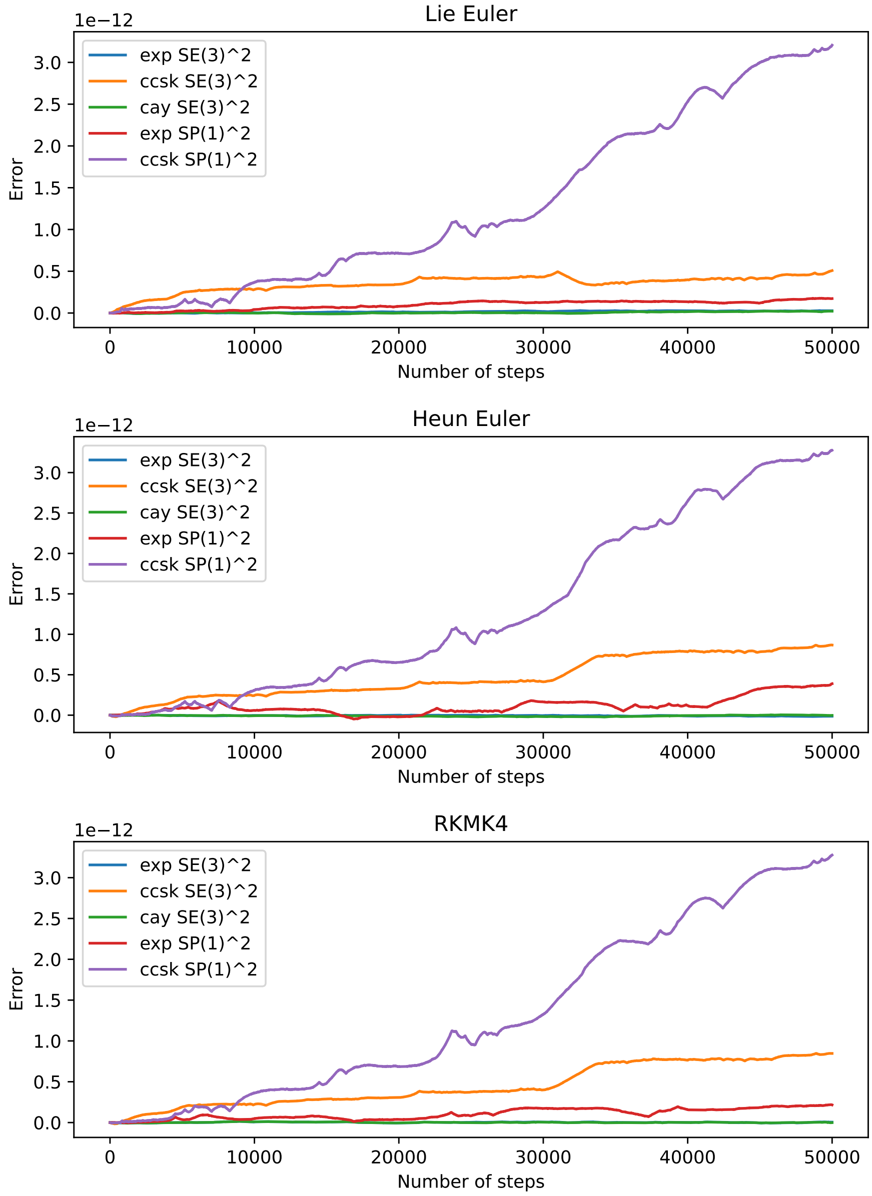

Figure B.2: Changes in $1 - \boldsymbol{q}_2^T \boldsymbol{q}_2$ for each step when $h = 1 \cdot 10^{-4}$

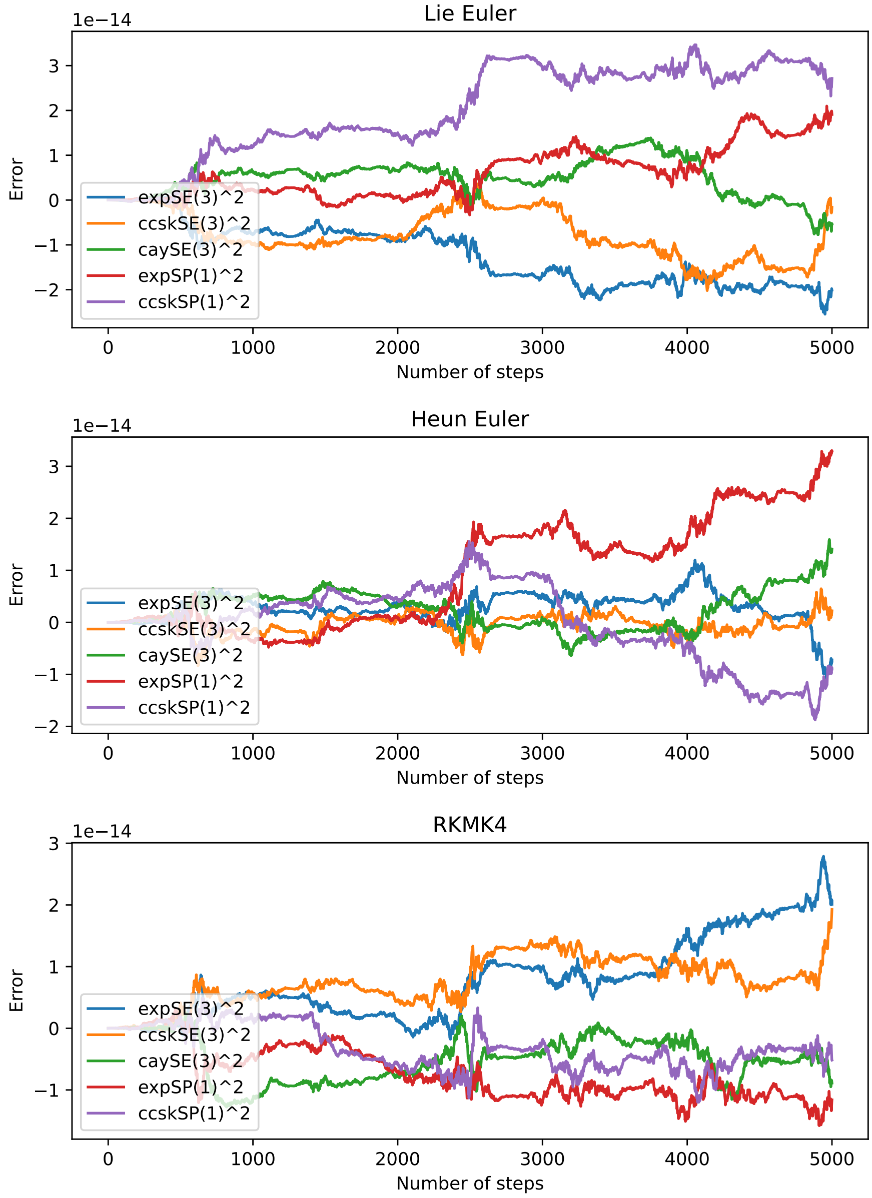

Figure B.3: Changes in $\boldsymbol{q}_2^T \boldsymbol{\omega}_2$ for each step when $h = 1 \cdot 10^{-3}$

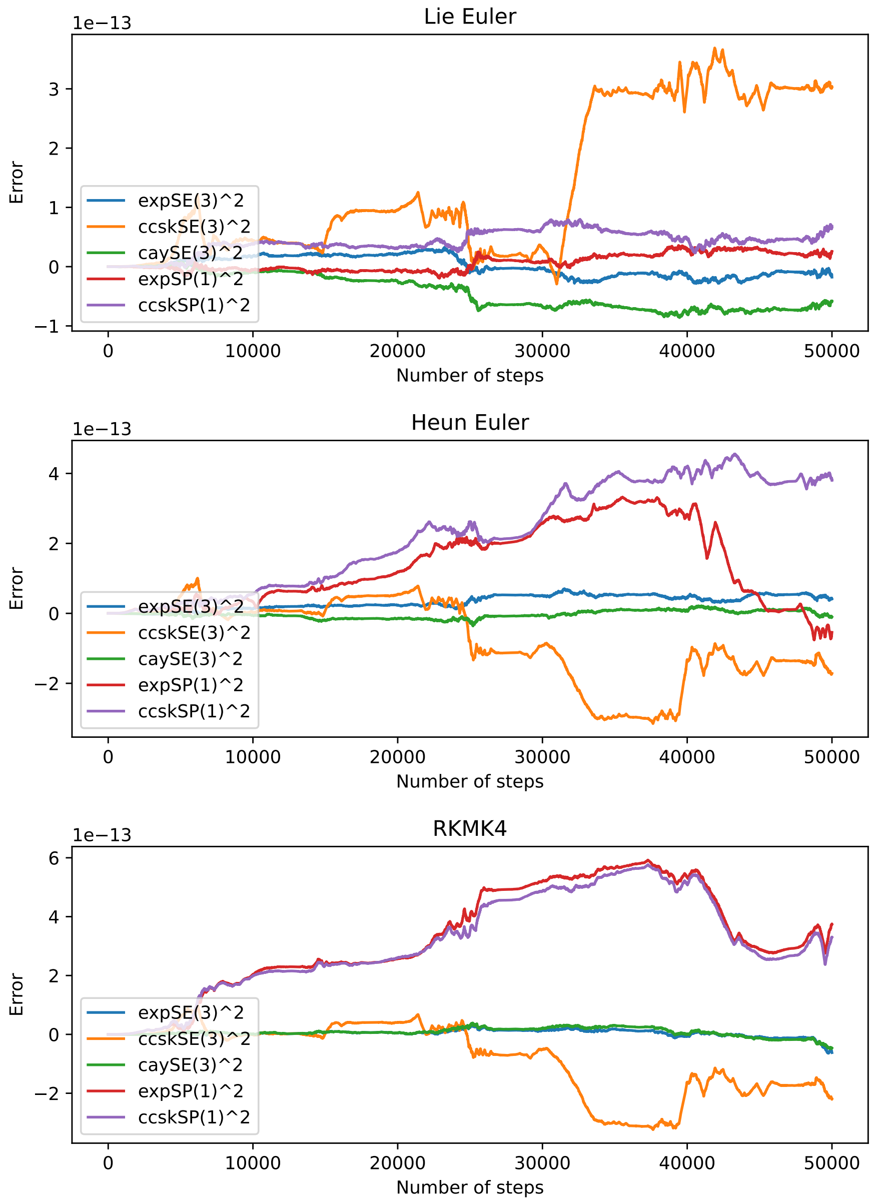

Figure B.4: Changes in $\boldsymbol{q}_2^T \boldsymbol{\omega}_2$ for each step when $h = 1 \cdot 10^{-4}$