NTNU
Norwegian University of
Science and Technology

Lie Group Integrators for Cotangent Bundles of Lie Groups and Their Application to Systems of Dipolar Soft Spheres

Vanje Rebni Kjer

Master of Science in Physics and Mathematics
Submission date: June 2017
Supervisor: Brynjulf Owren, IMF

Norwegian University of Science and Technology
Department of Mathematical Sciences

Preface

This thesis completes my studies in the Master's degree programme in Applied Physics and Mathematics, with specialisation in Industrial Mathematics, at the Norwegian University of Science and Technology (NTNU).

I would like to thank my supervisor, Professor Brynjulf Owren at the Department of Mathematical Sciences at NTNU, for all his guidance, insightful discussions and helpful explanations and comments. Also, special thanks to my partner Torbjørn for his support and willingness to proofread my thesis, and for his love and encouragement.

Vanje Rebni Kjer,
Trondheim, June 2017

Abstract

The objective of this thesis is to study numerical integrators and their application to solving ordinary differential equations arising from mechanical systems. Many mechanical problems are naturally formulated on Lie groups or on the groups' tangent or cotangent bundle, especially those where the equations of motion are derived from variational principles. The bundles inherit the structure of the original Lie group and can be considered Lie groups themselves. Often these systems are highly complicated and one will need to use numerical methods for solving them. One class of numerical methods which is particularly suitable for such problems is Lie group integrators. We shall apply the Runge-Kutta-Munthe-Kaas (RKMK) methods in such a setting.

We will give an introduction to Lie group theory, which we utilize to identify a Lie group structure on the cotangent bundles of Lie groups. Next, we provide a short overview of Lagrangian and Hamiltonian mechanics. Thereafter, we will conduct numerical experiments on two Hamiltonian systems using the RKMK methods. The first system concerns the motion of a free rigid body, and the second system the motion of a system of identical dipolar soft spheres in a molecular dynamics setting.

Samandrag

Formålet med denne oppgåva er å studere numeriske integratorar og deira evne til å løyse ordinære differensiallikningar frå mekaniske system. Mange mekaniske problem kan naturleg uttrykkjast på Lie-grupper eller på gruppa sin tangent- eller kotangentbunt, spesielt dersom rørslelikningane er utleidd frå variasjonsprinsipp. Buntane arvar strukturen av den opphavelege Lie-gruppa og kan sjølve bli sett på som Lie-grupper. Ofte er desse systema svært kompliserte og ein vil måtte ta i bruk numeriske metodar for å løyse dei. Ein klasse av numeriske metodar spesielt eigna for slike problem er Lie-gruppe-integratorar. Vi skal ta i bruk Runge-Kutta-Munthe-Kaas-metodane (RKMK) i ein slik situasjon.

Vi vil gi ein introduksjon til Lie-gruppeteori, som vi nyttar til å identifisere Lie-gruppestrukturen på kotangentbuntar til Lie-grupper. Deretter vil vi gi ei kort oversikt av Lagrange- og Hamilton-mekanikk. Etter det vil vi utføre numeriske eksperiment på to Hamiltonske system ved å bruke RKMK-metodar. Det første systemet handlar om rørsla til ein fri stiv lekam, og det andre om rørsla til eit system av identiske dipolare mjuke sfærer i ein molekylærdynamisk situasjon.

Contents

Chapter 1

Introduction

The idea of a manifold was developed over years by several mathematicians. Among these was Carl Friedrich Gauss, who in 1827 published his outstanding work where he used local coordinates on a surface and considered a surface as an abstract space without being embedded in a Euclidean space. Later, in 1854, Bernhard Riemann used the word manifold to describe objects of higher dimensional differential geometry. Still, the modern definition of a manifold did not appear before in 1931 [28].

Manifolds are analogues of smooth curves and surfaces to higher dimensions. More precisely, one can say that manifolds are spaces that are locally Euclidean in a small neighbourhood around every point. If a manifold has enough structure, the basic concepts of calculus, such as differentiation and integration, can be defined. An example of a manifold is the sphere which is a three dimensional object, but looks locally like $\mathbb{R}^2$. Other examples are Euclidean spaces, smooth plane curves such as circles and parabolas and smooth surfaces such as paraboloids and ellipsoids [1, 15, 17, 30].

A Lie group is a manifold with a group structure in which the group multiplication and inversion are smooth maps. Giving the tangent space at the identity a Lie bracket makes it a Lie algebra, which contains important information about the group. Some examples of Lie groups are the general and special linear group, the orthogonal group, the unitary group, and the symplectic group [15, 17, 28, 30].

A lot of our current understanding of Lie groups is thanks to the Norwegian mathematician Sophus Lie, who in the late nineteenth century began his work on Lie groups and Lie algebras. He became a professor in Leipzig in 1886, and he published later the three volume treatise "Theorie der Transformationsgruppe" which he and his assistant Friedrich Engel wrote together. The theory of Lie groups and Lie algebras is an intersection of group theory, topology and linear algebra, and is hence a rich and vibrant branch of mathematics [17, 28].

In many cases from mechanics, the phase space is the tangent or cotangent bundle of a Lie group [12]. This is typical when the equations of motions are derived from variational principles. The bundles are also manifolds, and inherit a Lie group structure from their configuration Lie group. In such situations we can apply Lie group integrators, which guarantee that both the numerical and analytical solutions evolve on the same manifold [5]. In order to apply Lie group integrators one must tailor them to the Lie group at hand. How to do this will be explained in this thesis.

There are several Lie group integrators available, but we will focus on Runge-Kutta-Munthe-Kaas methods, introduced by Munthe-Kaas in [20]. The idea behind the RKMK methods is to solve ODEs in the Lie algebra of the Lie group instead of on the manifold. Then, exploiting the vector space structure of the Lie algebra, any classical Runge-Kutta method can be used to determine the solution [5, 19, 20]. For the interested reader, we refer to the literature for more information about other Lie group methods; see e.g. [7, 16, 18] and references therein.

We consider Hamiltonian dynamics on cotangent bundles of Lie groups and implement Lie group integrators to numerically find the solution. A well-known problem is the free rigid body rotating around its center of mass, a problem we will consider and solve using the RKMK methods.

Another particularly interesting application is that of the motion of interacting dipolar soft spheres, for instance water molecules. We will set up a formulation of this problem using multiple copies of $\mathbb{R}^3 \times SO(3)$ that describe translations and rotations of dipoles. We demonstrate that the Lie group integrators can be applied with excellent results through numerical experiments.

The thesis is organised as follows. In Chapter 2, we present theory on manifolds and Lie groups as background material for the numerical methods discussed later. Section 2.4 introduces the Runge-Kutta-Munthe-Kaas methods. Sections 2.5 - 2.8 discuss mechanical systems and the application of RKMK methods to such systems. Chapter 3 presents numerical experiments and their results. A conclusion and topics of future works are presented in Chapter 4.

Chapter 2

Theory

2.1 Manifolds

An n-dimensional manifold is a topological space $\mathcal{M}$ that is locally Euclidean of dimension n in a neighbourhood around every point $m \in \mathcal{M}$. By denoting the neighbourhood $U \subset \mathcal{M}$, we may define a chart as the pair (U, ϕ) such that

$$\phi : U \mapsto \phi(U) \subset \mathbb{R}^n,$$

meaning that ϕ is homeomorphic to an open subset of $\mathbb{R}^n$. The subset U is the domain of ϕ, and ϕ is a bijective map. The local coordinates of m are denoted $\phi(m) = (x_1, \ldots, x_n)$, and if $\phi(m) = 0$, we say that the chart (ϕ, U) is centered at $m \in \mathcal{M}$ [1, 20, 22, 28].

Two charts, $(U_i, \phi_i), (U_j, \phi_j) : \mathbb{R}^n \mapsto \mathbb{R}^n$, $i, j \in \mathbb{N}$, are C^∞ compatible if the functions

$$\phi_i \circ \phi_j^{-1} : \phi_j(U_i \cap U_j) \subset \mathbb{R}^n \mapsto \phi_i(U_i \cap U_j) \subset \mathbb{R}^n,$$
$$\phi_j \circ \phi_i^{-1} : \phi_i(U_i \cap U_j) \subset \mathbb{R}^n \mapsto \phi_j(U_i \cap U_j) \subset \mathbb{R}^n,$$

are smooth. Equivalently, (U_i, ϕ_i) and (U_j, ϕ_j) are compatible if they are overlapping and are diffeomorphic, that is, bijective C^∞ maps whose inverses ϕ_i^{-1} and ϕ_j^{-1} are also C^∞ [1, 4, 20, 28].

A collection of compatible charts $A = \{(U_i, \phi_i)\}$ is called an atlas on $\mathcal{M}$

if

$$\mathcal{M} = \bigcup_i \{U_i\}. \tag{2.1}$$

A manifold $\mathcal{M}$ is smooth if it is a manifold with a maximal atlas, that is, the atlas is not contained in any other atlas [1, 28].

The manifolds we will work with in this thesis are n-dimensional and differentiable, meaning that a manifold $\mathcal{M}$ is a union of compatible charts (2.1), where every $m \in \mathcal{M}$ is a member of at least one chart [22, 30].

2.1.1 Tangent spaces

Let $\mathcal{M}$ be a manifold, and assume we have a curve $\gamma(t) \in \mathcal{M}$, where

$$\gamma(0) = m \in \mathcal{M},\ \dot{\gamma}(0) = v_m.$$

Then, v_m is the tangent vector at m. If two different curves go through m, having the same "direction", they give the same tangent vector in m, such that for $\zeta(t) \in \mathcal{M}$

$$\zeta(0) = \gamma(0),\ \dot{\zeta}(0) = \dot{\gamma}(0).$$

Thus, a tangent vector may be defined as an equivalence class of curves [22].

The tangent space at m is a vector space denoted $T_m\mathcal{M}$, and contains every tangent vector from differentiating all the differentiable curves that go through m [22, 28]. Thus, for all t, the tangent space at m can be defined as

$$T_m\mathcal{M} = \{v_m = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} \gamma(t)\ :\ \gamma(t)\ \in \mathcal{M},\ \gamma(0) = m\}.$$

Every point on the manifold has its own tangent space, and the collection of all the tangent spaces of $\mathcal{M}$ is called the tangent bundle, denoted

$$T\mathcal{M} = \bigcup_{m \in \mathcal{M}} T_m\mathcal{M}, \tag{2.2}$$

meaning that the tangent bundle is a collection of tangent vectors [28].

The tangent bundle is in general not a linear space; in fact, it is a manifold that has twice the dimension of $\mathcal{M}$, where the local coordinates on TM are induced by coordinates on $\mathcal{M}$. We explain this as following: Assume the coordinates of $m \in \mathcal{M}$ are $(m_1, \ldots, m_n)$. By differentiating curves $\gamma(t)$ that go through m, we get the tangent vectors $\dot{\gamma}(0) = v_m = (v_{m,1}, \ldots, v_{m,n})$. The local coordinates for $T\mathcal{M}$ are set as $(m_1, \ldots, m_n, v_{m,1}, \ldots, v_{m,n})$, which are of size $2n$ [22].

2.1.2 Differentials

Let $\mathcal{M}$ and $\mathcal{N}$ be two manifolds, where $\phi : \mathcal{M} \mapsto \mathcal{N}$ is a smooth map between them. At every point $m \in \mathcal{M}$, ϕ induces a linear tangent map between the tangent spaces of $\mathcal{M}$ and $\mathcal{N}$, called the differential of ϕ at m [20, 28].

Assume $\gamma(t)$ is a smooth curve in $\mathcal{M}$, where

$$\gamma(0) = m \in \mathcal{M},\ \dot{\gamma}(0) = v_m \in T_m\mathcal{M}.$$

Then, the differential of ϕ at m is $\phi_* : T_m\mathcal{M} \mapsto T_{\phi(m)}\mathcal{N}$, defined as

$$\phi_*(v_m) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \phi(\gamma(t)) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} (\phi \circ \gamma)(t),$$

where $\phi_*(v_m)$ is the velocity vector of the image curve $(\phi \circ \gamma)$ at $\phi(m)$, that is, the tangent vector in $T_{\phi(m)}\mathcal{N}$.

2.1.3 Vector fields and flows

Let $\mathcal{M}$ be a manifold and $T\mathcal{M}$ its tangent bundle. Assume we have a curve $\gamma(t) \in \mathcal{M}$, such that

$$\gamma(0) = m \in \mathcal{M}, \quad \dot{\gamma}(0) = v_m.$$

A map $\pi : T\mathcal{M} \mapsto M$ where $\pi(m, v_m) = m$ is called the natural projection. We define a section of the tangent bundle to be a map $X : \mathcal{M} \mapsto T\mathcal{M}$ such that $\pi \circ X = \mathrm{Id}_{\mathcal{M}}$ [22]. Let X be a vector field on a differentiable manifold $\mathcal{M}$, then X, is a section of the tangent bundle and a function that assigns to every point $m \in \mathcal{M}$ a tangent vector v_m. If X is a C^∞ map, the vector field is smooth [20, 28].

If the curve $\gamma : \mathbb{R} \mapsto \mathcal{M}$ is smooth and satisfies

$$\dot{\gamma}(t) = X(\gamma(t)),\ t \in (a, b),$$

then it is an integral curve of X. In addition, if X is smooth, there exists, for each $m \in \mathcal{M}$, an interval $(a(m), b(m)) \subset \mathbb{R}$ and a smooth curve, $\gamma_m : (a(m), b(m)) \mapsto \mathcal{M}$ with $\gamma(0) = m$ such that $\gamma_m(t)$ is an integral curve of X.

By holding t fixed, we can obtain a diffeomorphism of $\mathcal{M}$ by setting

$$\begin{aligned} \phi_{X,t} : \mathcal{D}_t &\mapsto \mathcal{M}, \text{where } \mathcal{D}_t = \{m \in \mathcal{M} : t \in (a(m), b(m))\} \\ m &\mapsto \gamma_m(t). \end{aligned}$$

This map is called the flow of the vector field X [25].

Let $\mathfrak{X}(\mathcal{M})$ and $\mathfrak{X}(\mathcal{N})$ denote the sets of all vector fields on $\mathcal{M}$ and on $\mathcal{N}$, another differentiable manifold. If $\phi : \mathcal{M} \mapsto \mathcal{N}$ is a smooth map, and $X \in \mathfrak{X}(\mathcal{M})$ and $Y \in \mathfrak{X}(\mathcal{N})$ satisfy

$$\phi_* \circ X = Y \circ \phi, \tag{2.3}$$

they are called ϕ-related vector fields, denoted $X \sim_\phi Y$. This means that ϕ maps the flow of X to the flow of Y [1, 23, 30, 28].

2.1.4 Dual spaces

Let V be a (linear) vector space of dimension n, having a basis $e_1, \ldots, e_n$ such that any element $x \in V$ can be uniquely expressed as

$$x = x_1 e_1 + \cdots + x_n e_n, \tag{2.4}$$

for all $x_i \in \mathbb{R}$. The dual space of V, denoted V^*, has the same dimension as V and is also a vector space. It consists of the set of linear functionals on V, and we can express a basis on V^* by means of linear functions $\mu_1, \ldots, \mu_n$, that satisfy

$$\mu_i(x) = x_i,$$

for x in (2.4), or equivalently

$$\mu_i(e_j) = \delta_{ij},\ i, j = 1, \ldots, n,$$

where δ is the Kronecker delta [28].

We call $x \in V$ a primal vector and $\mu \in V^*$ a dual vector or cotangent vector. Usually, we say that a dual vector acts on a primal vector, and a common notation is

$$\langle \mu, x \rangle = \mu(x) \in \mathbb{R}, \tag{2.5}$$

where $\langle \cdot, \cdot \rangle$ denotes the duality pairing [22].

Moreover, since the dual space V^* is a vector space, it will also have a dual space $(V^*)^*$. However, $(V^*)^* \cong V$ so if $y \in (V^*)^*$ it is identified with $x \in V$ whenever $\langle y, \mu \rangle = \langle \mu, x \rangle$. Thus, under the identification just seen, the ordering of the arguments in the duality pairing is irrelevant, hence a primal vector may act on a dual vector. Then the duality pairing becomes a symmetric, bilinear form on $V \times V^*$, namely $\langle \mu, x \rangle = \langle x, \mu \rangle$ [22].

As earlier mentioned, tangent spaces are vector spaces, so if m is a point on the manifold $\mathcal{M}$, its tangent space, $T_m\mathcal{M}$, has a dual space, denoted $T_m^*\mathcal{M}$. The family of all the cotangent spaces of $\mathcal{M}$ is called the cotangent bundle, and can be defined in a similar way as the tangent bundle in (2.2), namely

$$T\mathcal{M}^* = \bigcup_{m \in \mathcal{M}} T_m^*\mathcal{M}.$$

2.2 Lie groups and Lie algebras

Definition 2.1. A group is a pair $(G, *)$ where G is a set and $*$ is a binary operation $* : G \times G \mapsto G$, where $*$ and G must satisfy certain properties. The binary operation $*$ must be associative, that is, for all $a, b, c \in G$:

$$a * (b * c) = (a * b) * c. \tag{2.6}$$

There must also exist an identity element e, and for all $a \in G$ there must exist an inverse $a^{-1} \in G$, such that:

$$\begin{aligned} a * e = e * a = a, \\ a * a^{-1} = a^{-1} * a = e. \end{aligned} \tag{2.7}$$

If $*$ is commutative, i.e. $a * b = b * a$ for all $a, b \in G$, then $(G, *)$ is called an Abelian group.

If G is a differentiable manifold and the maps $*$ and $a \mapsto a^{-1}$ are smooth, we call $(G, *)$ a Lie group. Often a Lie group is denoted only as G, a notation we from now on will adopt [4, 14, 15].

2.2.1 Multiplication on Lie groups

Let G be a Lie group and for some fixed $g \in G$, denote

$$\begin{aligned} L_g &: G \mapsto G, \\ R_g &: G \mapsto G, \end{aligned}$$

as the left and right multiplication (or synonymously left and right translations) on G, respectively [28].

Example 2.2. For matrix Lie groups the binary operation, $*$, is just standard matrix multiplication such that for $p \in G$

$$L_g p = gp, \quad R_g p = pg.$$

■

Since the maps $*$ and $a \mapsto a^{-1}$ are smooth, both L_g and R_g map a neighbourhood of the identity to a neighbourhood of g and are bijective C^∞ maps whose inverses $(L_g)^{-1} = L_{g^{-1}}$ and $R_g)^{-1} = R_{g^{-1}}$ are also smooth. So, when studying the structure of a Lie group, we only need to examine the neighbourhood of the identity, since all the local information about the group is concentrated there.

This also makes the tangent space at the identity important, because when taking the identity element e to g, L_g and R_g induce an isomorphism of tangent spaces, denoted $L_{g*} : T_e G \mapsto T_g G$ and $R_{g*} : T_e G \mapsto T_g G$, defined as

$$\begin{aligned} L_{g*}\xi &= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} L_g h(t) = L_g \dot{h}(0), \\ R_{g*}\xi &= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} R_g h(t) = R_g \dot{h}(0), \end{aligned}$$

where $h(t)$ is a curve in G, with $h(0) = e \in G$ and $\dot{h}(0) = \xi \in T_e G$ [28]. Another name for L_{g*} and R_{g*} are the differentials of L_g and R_g, respectively.

Example 2.3. For matrix Lie groups both elements in G and in TG are matrices such that we may multiply them together using matrix multiplication, that is

$$L_{g*}\xi = g\xi, \quad R_{g*}\xi = \xi g.$$

■

By choosing any vector ξ in the tangent space at the identity, the induced left and right tangent maps describe the tangent space T_gG at a point $g \in G$. We call this "trivialisation". Hence, every tangent vector $v_g \in T_gG$ can be left or right trivialised, that is, written on the form $v_g = L_{g*}\xi$ or $v_g = R_{g*}\xi$, respectively [18, 19, 29, 30].

As earlier explained, tangent spaces are vector spaces, meaning that they have dual spaces. This leads to the additional and important property of L_g and R_g, namely; they also induce the dual maps $L_g^* : T_g^*G \mapsto T_e^*G$ and $R_g^* : T_g^*G \mapsto T_e^*G$. By using the duality pairing (2.5), we can define the dual maps, from here on called cotangent maps, as following. Assume that the dual vector of v_g is $p_g \in T_g^*G$, then

$$\langle p_g, v_g \rangle = \langle p_g, L_{g*}\xi \rangle := \left\langle L_g^* p_g, \xi \right\rangle,$$

and

$$\langle p_g, v_g \rangle = \langle p_g, R_{g*}\xi \rangle := \left\langle R_g^* p_g, \xi \right\rangle,$$

where L_{g*} and R_{g*} are isomorphisms between T_g^*G and T_e^*G [7].

Then, for $\mu \in T_e^*G$ the dual vector p_g can be represented as

$$p_g = L_{g^{-1}}^* \mu,$$

or

$$p_g = R_{g^{-1}}^* \mu,$$

that is, in a left or right trivialised way, respectively.

2.2.2 Lie group actions

A left Lie group action of a Lie group G on a manifold $\mathcal{M}$ is a smooth map

$$\Lambda^L : G \times \mathcal{M} \mapsto \mathcal{M} \tag{2.8}$$

which satisfies the properties

$$\Lambda^L(e, m) = m,$$
$$\Lambda^L(g_1 g_2, m) = \Lambda^L(g_1, \Lambda^L(g_2, m)),$$

for all $m \in \mathcal{M}$ and $g_1, g_2 \in G$, where e denotes the identity element in G [18, 20].

Similarly, a right Lie group action is a smooth map

$$\Lambda^R : \mathcal{M} \times G \mapsto \mathcal{M} \tag{2.9}$$

which satisfies the properties

$$\Lambda^R(m, e) = m,$$
$$\Lambda^R(m, g_1 g_2) = (\Lambda^R(m, g_1), g_2).$$

The group action is called locally transitive if for all $m \in \mathcal{M}$, there exists an open neighbourhood $U \subset \mathcal{M}$ such that $m \in U$ and

$$\Lambda_m := \Lambda(\cdot, m) : G \mapsto U$$

maps G onto U [19, 23]. Note that a Lie group action is coordinate independent which makes it suitable for discussing numerical methods on Lie groups in general terms.

Example 2.4. Consider the case when $\mathcal{M} = G$, the Lie group itself. Then, the left group action Λ^L is equal to right multiplication on G, and the right group action Λ^R is equal to left multiplication on G, that is

$$\Lambda^L(g, m) = R_m(g),$$
$$\Lambda^R(m, g) = L_m(g).$$

■

2.2.3 Lie algebras

A vector space $\mathfrak{g}$ over $\mathbb{R}$ together with the bracket

$$[\cdot, \cdot] : \mathfrak{g} \times \mathfrak{g} \mapsto \mathfrak{g}, \tag{2.10}$$

forms a Lie algebra $\mathfrak{g}$ if the bracket satisfies the following properties: For all $U, V, W \in \mathfrak{g}$ and $a, b \in \mathbb{R}$ the bracket must be [1, 30]

1. Bilinear:
   $[aU + bV, W] = a[U, W] + b[V, W]$, $[W, aU + bV] = a[W, U] + b[W, V]$.
2. Skew-symmetric: $[V, U] = -[U, V]$.
3. Satisfying the Jacobi identity: $[U, [V, W]] + [V, [W, U]] + [W, [U, V]] = 0$.

Every Lie group has a Lie algebra associated to it. The underlying vector space is the tangent space to G at e, denoted $\mathfrak{g}$ such that $\mathfrak{g} = T_e G$.
Let u and v be elements in $\mathfrak{g}$ and consider the curves $g(t)$ and $h(t)$ in G, where

$$g(0) = h(0) = e, \quad \dot{h}(0) = v, \quad \dot{g}(0) = u.$$

The bracket (2.10) of a Lie group is defined as [20, 28]

$$[u, v] = \frac{\partial^2}{\partial t \partial s}\bigg|_{t=s=0} g(t) \cdot h(s) \cdot g(t)^{-1}.$$

2.2.4 Lie algebra actions

Assume G is a Lie group acting on a manifold $\mathcal{M}$, where

$$\phi : \mathfrak{g} \mapsto G \tag{2.11}$$

is a smooth map such that $\phi(0) = e$, the identity element in G. Together with the left or right Lie group action (2.8) or (2.9), the map ϕ induces left and right Lie algebra actions, $\lambda^L : \mathfrak{g} \times \mathcal{M} \mapsto \mathcal{M}$ or $\lambda^R : \mathcal{M} \times \mathfrak{g} \mapsto \mathcal{M}$ defined as:

$$\begin{aligned} \lambda^L(v, m) &= \Lambda^L(\phi(v), m), \\ \lambda^R(v, m) &= \Lambda^R(\phi(v), m), \end{aligned} \tag{2.12}$$

where $m \in \mathcal{M}$ and $v \in \mathfrak{g}$. These maps are non-unique since changing ϕ also changes λ^L and λ^R [20].

The differentials of λ^L and λ^R are called the infinitesimal generators of the Lie algebra actions. They define directions tangent to the manifold, and are Lie algebra homomorphisms denoted $\lambda^L_* : \mathfrak{g} \mapsto \mathfrak{X}(\mathcal{M})$ and $\lambda^R_* : \mathfrak{g} \mapsto \mathfrak{X}(\mathcal{M})$, defined

$$\begin{aligned} (\lambda^L_* v)(m) &= \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} \lambda^L(tv, m), \\ (\lambda^R_* v)(m) &= \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} \lambda^R(tv, m), \end{aligned} \tag{2.13}$$

respectively. Note that these homomorphisms are uniquely defined in that they do not depend on ϕ [20].

2.2.5 Left-invariant vector fields on a Lie group

If $X \in \mathfrak{X}(G)$ is a vector field on a Lie group G, then X is left-invariant if

$$L_{g*}X = X,$$

for all $g \in G$. That is, X is left-invariant if and only if it is L_g-related (2.3) to itself, meaning that for all $h \in G$,

$$L_{g*}(X_h) = X_{gh}.$$

Figure 2.1 shows the commutative diagram of a left-invariant vector field.

Example 2.5. A tangent vector $\xi \in \mathfrak{g}$ can be used to generate a left-invariant vector field X^L_ξ on G by letting $X^L_\xi(g) = L_{g*}\xi$, since

$$L_{g*}(X^L_\xi(h)) = L_{g*}(L_{h*}\xi) = (L_g \circ L_h)_*\xi = X^L_\xi(gh),$$

for all $h \in G$. Analogously, ξ can generate a right-invariant vector field by letting $X^R_\xi(g) = R_{g*}\xi$ [18, 28]. ■

$$\begin{tikzcd} TG \arrow[r, "L_{g*}"] & TG \\ G \arrow[u, "X^L"] \arrow[r, "L_g"'] & G \arrow[u, "X^L"'] \end{tikzcd}$$

Figure 2.1: Commutative diagram of a left-invariant vector field.

2.2.6 Adjoint representations

Assume we have a Lie group G, and its corresponding Lie algebra $\mathfrak{g}$. The adjoint representation $\mathrm{Ad}_g : \mathfrak{g} \mapsto \mathfrak{g}$ is a linear and well-defined operator frequently used in the setting of Lie group theory. It is defined as

$$\mathrm{Ad}_g\xi = \mathrm{L}_{g*}\ \mathrm{R}_{g^{-1}*}\ \xi, \tag{2.14}$$

where $g \in G$ and $\xi \in \mathfrak{g}$. It has a dual operator called the coadjoint representation, that is, the map $\mathrm{Ad}_g^* : \mathfrak{g}^* \mapsto \mathfrak{g}^*$, defined as

$$\mathrm{Ad}_g^*\mu = (\mathrm{R}_{g^{-1}*}\,\mathrm{L}_{g*})^*\ \mu,$$

where $\mathfrak{g}^*$ denotes the dual space of $\mathfrak{g}$, and $\mu \in \mathfrak{g}^*$ [3, 16].

Example 2.6. If G is a matrix Lie group, and $h(t)$ is a curve in G, where $h(0) = e \in G$ and $\dot{h}(0) = \xi \in \mathfrak{g}$, we may give an explicit formula for the adjoint representation, namely

$$\mathrm{Ad}_g\xi = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} L_g R_{g^{-1}} h(t) = g\dot{h}(t)g^{-1}\big|_{t=0} = g\xi g^{-1}.$$

Similarly, for $\mu \in \mathfrak{g}^*$,

$$\mathrm{Ad}_g^*\mu = g^{-1}\mu g,$$

is an explicit expression for the coadjoint representation [11]. ∎

In the rest of the section we assume G is a matrix Lie group. By differentiating $\mathrm{Ad}_{g(t)}$ and $\mathrm{Ad}_{g(t)}^*$ we can determine their tangent maps, called the adjoint and coadjoint operators, respectively. Let $g(t) = \exp(tv)$ be a curve in G, where $g(0) = e \in G$ and $\dot{g}(0) = v \in \mathfrak{g}$. Then the ad operators are

$$\begin{aligned} \mathrm{ad}_v\xi &= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \mathrm{Ad}_{g(t)}\xi = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} g(t)\xi g(t)^{-1} = v\xi - \xi v, \\ \mathrm{ad}_v^*\mu &= -\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \mathrm{Ad}_{g(t)}^*\mu = -\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} g(t)^{-1}\mu g(t) = v\mu - \mu v. \end{aligned} \tag{2.15}$$

In fact, $\mathrm{ad}_v\xi = [v, \xi]$, which for matrix Lie groups is their matrix commutator [21]. The following lemma states some properties of the adjoint and coadjoint representations that we will use in the following sections.

Lemma 2.7.

a) $\mathrm{Ad}_{gh}\xi = \mathrm{Ad}_g\mathrm{Ad}_h\xi$.

b) $(\mathrm{Ad}_g)^{-1} = \mathrm{Ad}_{g^{-1}}$.

c) $(\mathrm{Ad}_g^*)^{-1} = \mathrm{Ad}_{g^{-1}}^*$.

d) $\mathrm{Ad}_{gh}^*\mu = \mathrm{Ad}_h^*\mathrm{Ad}_g^*\mu$.

Proof.

a) $\mathrm{Ad}_{gh}\xi = gh\xi(gh)^{-1} = gh\xi h^{-1}g^{-1} = g\mathrm{Ad}_h\xi g^{-1} = \mathrm{Ad}_g\mathrm{Ad}_h\xi.$

b) $\mathrm{Ad}_{g^{-1}}\mathrm{Ad}_g\xi = g^{-1}g\xi g^{-1}g = \xi \Rightarrow (\mathrm{Ad}_g)^{-1} = \mathrm{Ad}_{g^{-1}}.$

c) $\left\langle (\mathrm{Ad}_g^*)^{-1}\mu, v\right\rangle = \left\langle \mu, (\mathrm{Ad}_g)^{-1}v\right\rangle = \langle \mu, \mathrm{Ad}_{g^{-1}}v\rangle = \left\langle \mathrm{Ad}_{g^{-1}}^*\mu, v\right\rangle.$

d) $\left\langle \mathrm{Ad}_{gh}^*\mu, v\right\rangle = \langle \mu, \mathrm{Ad}_{gh}v\rangle = \langle \mu, \mathrm{Ad}_g\mathrm{Ad}_h v\rangle = \left\langle \mathrm{Ad}_g^*\mu, \mathrm{Ad}_h v\right\rangle = \left\langle \mathrm{Ad}_h^*\mathrm{Ad}_g^*\mu, v\right\rangle.$

□

2.2.7 Identifying $\mathfrak{so}(3)$ with $\mathbb{R}^3$

Assume G is the Lie group $SO(3)$, the special orthogonal group of dimension 3, consisting of 3×3 orthogonal matrices with unit determinant. The corresponding Lie algebra is denoted $\mathfrak{so}(3)$ and consists of 3-dimensional skew-symmetric matrices. We can identify $\mathfrak{so}(3)$ with $\mathbb{R}^3$ via the "hat map"

$$v = [v_1, v_2, v_3]^T \mapsto \hat{v} = \begin{bmatrix} 0 & -v_3 & v_2 \\ v_3 & 0 & -v_1 \\ -v_2 & v_1 & 0 \end{bmatrix},$$

where $\hat{v} \in \mathfrak{so}(3)$ and $v \in \mathbb{R}^3$ [6, 9, 18].

Lemma 2.8. Let $g \in SO(3)$ and $\hat{v} \in \mathfrak{so}(3)$, then using the hat map the adjoint representation can be expressed on the form

$$\mathrm{Ad}_g v = gv, \quad v \in \mathbb{R}^3.$$

Proof. Let $w \in \mathbb{R}^3$. Then

$$(\mathrm{Ad}_g\hat{v})(w) = g\hat{v}(g^{-1}w) = g(v \times g^{-1}w) = gv \times w,$$

such that

$$(\mathrm{Ad}_g\hat{v}) = (gv)^\wedge,$$

where $\wedge$ means applying the hat map [18]. □

Note that for $v, w \in \mathbb{R}^3$

$$\mathrm{ad}_v(w) = \hat{v}w = v \times w,$$

where $\times$ is the familiar cross product between vectors in $\mathbb{R}^3$.

2.2.8 The exponential map

The exponential map $\exp : \mathfrak{g} \mapsto G$ is a local, smooth map defined through the flow of left or right invariant vector fields on G. If G is a matrix Lie group, the exponential mapping is defined by the well-known power series

$$\exp(v) = \sum_{k=0}^{\infty} \frac{v^k}{k!}, \tag{2.16}$$

where $v \in \mathfrak{g}$ [15, 17, 20, 28].

In (2.11) we considered a coordinate map ϕ which, together with the Lie group action Λ, induced the Lie algebra action λ. The exponential map is an example of such a map, and it is the coordinate map we will consider in this thesis. For other examples of coordinate maps, see e.g. [8] and references therein.

The exponential map defines a connection

$$\lambda^L(v, m) = \Lambda^L(\exp(v), m),$$
$$\lambda^R(v, m) = \Lambda^R(\exp(v), m),$$

where $m \in \mathcal{M}$ and $v \in \mathfrak{g}$. It is valid only locally, near the points $m \in \mathcal{M}$ and $0 \in \mathfrak{g}$.

In addition, a relation between the infinitesimal generators of the Lie algebra actions (2.13) and the group actions (2.8) and (2.9) can be expressed as

$$(\lambda_*^L v)(m) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \Lambda^L(\exp(tv), m),$$
$$(\lambda_*^R v)(m) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \Lambda^R(\exp(tv), m).$$

The derivative of the exponential map is an important part of the RKMK methods, which we will see later. In the following section we therefore determine the derivative of the exponential map, but to do this, we need the following definition and lemma.

Definition 2.9.

$$\exp(\mathrm{ad}_u)(v) = \sum_{k=0}^{\infty} \frac{1}{k!} \mathrm{ad}_u^k(v).$$

Lemma 2.10.

$$L_{*\exp(u)}R_{*\exp(-u)}v = \mathrm{Ad}_{\exp(u)}v = \exp(u)v\exp(-u) = \exp(\mathrm{ad}_u)(v).$$

Proof. Let

$$y(t) = \exp(tu)v\exp(-tu),$$

and

$$z(t) = \exp(t\mathrm{ad}_u)(v).$$

Differentiating $y(t)$ gives

$$\begin{aligned}\dot{y}(t) &= u\exp(tu)v\exp(-tu) + \exp(tu)v(-u)\exp(-tu)\\ &= u\exp(tu)v\exp(-tu) - \exp(tu)v\exp(-tu)u = uy - yu = [u, y],\\ y(0) &= \exp(0)v\exp(0) = v.\end{aligned}$$

Next, we differentiate $z(t) = \sum_{k=0}^{\infty} \frac{t^k}{k!}\mathrm{ad}_u^k v$. Then,

$$\begin{aligned}\dot{z}(t) &= \frac{\mathrm{d}}{\mathrm{d}t}\sum_{k=0}^{\infty}\frac{t^k}{k!}\mathrm{ad}_u^k v = \sum_{k=1}^{\infty}\frac{kt^{k-1}}{(k-1)!}\mathrm{ad}_u^k v = \sum_{k=0}^{\infty}\frac{kt^k}{k!}\mathrm{ad}_u(\mathrm{ad}_u^k v)\\ &= \mathrm{ad}_u(\sum_{k=0}^{\infty}\frac{t^k}{k!}\mathrm{ad}_u^k v) = \mathrm{ad}_u z = [u, z],\\ z(0) &= \mathrm{ad}_u^0(v) = v.\end{aligned}$$

We observe that both $y(t)$ and $z(t)$ satisfy the same differential equation with the same initial value, hence $y(t) = z(t)$, where especially $y(1) = z(1)$, such that by definition 2.9

$$\exp(u)v\exp(-u) = \exp(\mathrm{ad}_u)(v).$$

□

2.2.9 The derivative of the exponential map

This section is based on [23], and we intend to find the derivative of the exponential map (2.16).

Assume we have a Lie group G with a corresponding Lie algebra $\mathfrak{g}$. Suppose there is a curve $\sigma(t) = \sigma_0 + tv \in \mathfrak{g}$, where

$$\begin{aligned} \sigma(0) &= \sigma_0 \in \mathfrak{g}, \\ \dot{\sigma}(0) &= v \in T_{\sigma_0}\mathfrak{g} \cong \mathfrak{g}. \end{aligned}$$

A segment of σ gives rise to a curve g in G via the exponential map, such that

$$\begin{aligned} g(t) &= \exp(\sigma(t)) = \exp(\sigma_0 + tv) \in G, \\ g(0) &= \exp(\sigma(0)) = \exp(\sigma_0) \in G, \\ \dot{g}(0) &= T_{\sigma_0}\exp(v) = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} \exp(\sigma_0 + tv) \in T_{\exp(\sigma_0)}G. \end{aligned}$$

Here, $T_{\sigma_0}\exp$ denotes the differential of exp at σ_0. We know that by using left or right trivialisation, every tangent space of a Lie group can be identified with the Lie algebra. So, for $\xi_L, \xi_R \in \mathfrak{g}$

$$T_{\sigma_0}\exp(v) = L_{\exp(\sigma_0)*}\xi_L = R_{\exp(\sigma_0)*}\xi_R, \tag{2.17}$$

where $\xi_R = \mathrm{Ad}_{\exp(\sigma_0)}\xi_R$. The subscripts L and R on ξ_L, ξ_R denote that we have used left and right multiplication, respectively.

Let $y_t(s) = \exp(s(\sigma_0 + tv))$, such that

$$T_{\sigma_0}\exp(v) = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} y_t(1).$$

Differentiating with respect to s gives

$$\dot{y}_t(s) = \frac{\mathrm{d}}{\mathrm{d}s} y_t(s) = (\sigma_0 + tv)y_t(s).$$

Noticing that $y_t(s) = \exp(s\sigma_0) + \mathcal{O}(t)$ as $t \to 0$, we get [7]

$$\dot{y}_t(s) - \sigma_0 y_t(s) = tv\exp(s\sigma_0) + \mathcal{O}(t^2).$$

Introducing the integrating factor $\exp(-t\sigma_0)$ gives

$$\frac{\mathrm{d}}{\mathrm{d}s}(\exp(-s\sigma_0)y_t(s)) = t\exp(-s\sigma_0)v\exp(s\sigma_0) + \mathcal{O}(t^2).$$

We integrate and use that $y_t(0) = \mathrm{Id}$, hence

$$\exp(-s\sigma_0)y_t(s) - \mathrm{Id} = t\int_0^s \exp(-r\sigma_0)v\exp(r\sigma_0)\mathrm{d}r + \mathcal{O}(t^2),$$

such that

$$
\begin{aligned}
y_t(s) &= \exp(s\sigma_0) + t\int_0^s \exp((s-r)\sigma_0)v\exp(r\sigma_0)\mathrm{d}r + \mathcal{O}(t^2) \\
&= \exp(s\sigma_0) + t\int_0^s \exp(k\sigma_0)v\exp(-k\sigma_0)\exp(s\sigma_0)\mathrm{d}k + \mathcal{O}(t^2),
\end{aligned}
$$

where we have used the substitution $k = s - r$.

Differentiating $y_t(1)$ and using definition (2.9) gives

$$
\begin{aligned}
\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} y_t(1) &= \int_0^1 \exp(k\sigma_0)v\exp(-k\sigma_0)\mathrm{d}k\ \exp(\sigma_0) \\
&= \int_0^1 \exp(k\mathrm{ad}_{\sigma_0})(v)\mathrm{d}k\ \exp(\sigma_0) \\
&= \int_0^1 \exp(kz)|_{z=\mathrm{ad}_{\sigma_0}}(v)\mathrm{d}k\ \exp(\sigma_0) \\
&= \left(\left.\frac{\exp(z)-1}{z}\right|_{z=\mathrm{ad}_{\sigma_0}}(v)\right)\exp(\sigma_0).
\end{aligned}
$$

We define $\mathrm{dexp}^R_{\sigma_0} : \mathfrak{g} \mapsto \mathfrak{g}$ by

$$
\mathrm{dexp}^R_{\sigma_0} v := \int_0^1 \exp(r\mathrm{ad}_{\sigma_0})(v)\mathrm{d}r = \left.\frac{\exp(z)-1}{z}\right|_{z=\mathrm{ad}_{\sigma_0}}(v).
$$

We let ξ_L and ξ_R in (2.17) be $\mathrm{dexp}^L_{\sigma_0}(v)$ and $\mathrm{dexp}^R_{\sigma_0}(v)$, respectively. Then,

$$
T_{\sigma_0}\exp(v) = L_{\exp(\sigma_0)*}\mathrm{dexp}^L_{\sigma_0}(v) = R_{\exp(\sigma_0)*}\mathrm{dexp}^R_{\sigma_0}(v),
$$

where [16]

$$
\begin{aligned}
\mathrm{dexp}^L_{\sigma_0} &= \mathrm{Ad}_{\exp(-\sigma_0)}\mathrm{dexp}^R_{\sigma_0} \\
&= \exp(\mathrm{ad}_{-\sigma_0})\left(\frac{\exp(\mathrm{ad}_{\sigma_0}) - I}{\mathrm{ad}_{\sigma_0}}\right) \\
&= \frac{I - \exp(\mathrm{ad}_{-\sigma_0})}{\mathrm{ad}_{\sigma_0}} \\
&= \mathrm{dexp}^R_{-\sigma_0}.
\end{aligned}
$$

The superscripts L and R on $\mathrm{dexp}^L_{\sigma_0}$ and $\mathrm{dexp}^R_{\sigma_0}$ denote that we have used left and right multiplication, respectively.

Example 2.11. For matrix Lie groups,

$$T_{\sigma_0}\exp(v) = \mathrm{dexp}^R_{\sigma_0}(v)\exp(\sigma_0) = \exp(\sigma_0)\mathrm{dexp}^L_{\sigma_0}(v).$$

■

The dexp map can also be considered as an infinite series of nested commutators [23]

$$\mathrm{dexp}^R_{\sigma_0}(v) = (I + \frac{1}{2!}\mathrm{ad}_{\sigma_0} + \frac{1}{3!}\mathrm{ad}^2_{\sigma_0} + \cdots)(v) = v + \frac{1}{2!}[\sigma_0, v] + \frac{1}{3!}[\sigma_0, [\sigma_0, v]] + \cdots$$
$$\mathrm{dexp}^L_{\sigma_0}(v) = (I - \frac{1}{2!}\mathrm{ad}_{\sigma_0} + \frac{1}{3!}\mathrm{ad}^2_{\sigma_0} + \cdots)(v) = v - \frac{1}{2!}[\sigma_0, v] + \frac{1}{3!}[\sigma_0, [\sigma_0, v]] + \cdots ,$$

which is often quite useful.

However, it is the inverse of the dexp map which is most important for Lie group integrators. Observe that the function

$$\phi_1(z) = \frac{\exp(z) - 1}{z}$$

is analytic in all of $\mathbb{C}$, meaning that its inverse

$$\frac{z}{\exp(z) - 1}$$

is analytic when $\phi_1(z) \neq 0$. This has a converging Taylor series expansion

$$\frac{z}{\exp(z) - 1} = 1 - \frac{z}{2} + \sum_{k=1}^{\infty} \frac{B_{2k}}{(2k)!} z^{2k},$$

around $z = 0$ in the open disk $|z| < 2\pi$, where B_{2k} are the Bernoulli numbers. Thus, the map is given by

$$\begin{aligned} v = \mathrm{dexp}^{-1}_{\sigma_0}(w) &= \left.\frac{z}{\exp(z) - 1}\right|_{z=\mathrm{ad}_{\sigma_0}} (w) \\ &= w - \frac{1}{2}[\sigma_0, w] + \frac{B_2}{2!}[\sigma_0, [\sigma_0, w]] + \cdots \\ &= \sum_{n=0}^{\infty} \frac{B_n}{n!} \mathrm{ad}^n_{\sigma_0}, \end{aligned} \tag{2.18}$$

where $v = \frac{\mathrm{d}}{\mathrm{d}t}\big|_{t=0}\sigma(t) = \dot{\sigma}\ \big|_{t=0} \in \mathfrak{g}$, such that by solving (2.18) we actually solve an ODE in the Lie algebra. In section 2.4 we will see that this is a fundamental step in the RKMK methods. Figure 2.2 shows an overview of the mappings explained above.

$$
\begin{tikzcd}
\mathfrak{g} \arrow[r, "R_{\exp(\sigma_0)*}"] & T_{\exp(\sigma_0)}G \\
\mathfrak{g} \arrow[u, "\mathrm{dexp}^R_{\sigma_0}"] \arrow[ur, "T_{\exp(\sigma_0)}"] \arrow[r, "\mathrm{dexp}^L_{\sigma_0}"'] & \mathfrak{g} \arrow[u, "L_{\exp(\sigma_0)*}"']
\end{tikzcd}
$$

Figure 2.2: Overview of the left and right dexp map.

2.3 Runge-Kutta methods

Runge-Kutta methods are well-studied and popular methods for numerical integration of ODEs. A general s-stage Runge-Kutta method is on the form:

$$
\begin{aligned}
y_{n+1} &= y_n + h\sum_{i=1}^{s} b_i k_i, \\
k_i &= f\left(t_n + c_i h,\ y_n + h\sum_{j=1}^{s} a_{ij}k_j\right), \quad i = 1, 2 \ldots, s.
\end{aligned}
$$

The coefficients, $b_i,\ a_{ij}$ and $c_i = \sum_{j=1}^{s} a_{ij},\ i, j = 1, 2, \ldots, s$ are real numbers, and can for convenience be arranged in a Butcher tableau, see table 2.1. The matrix $A = (a_{ij}) \in \mathbb{R}^{s\times s}$, and the two vectors $b = [b_1, b_2, \ldots, b_s]^T \in \mathbb{R}^s$

$$
\begin{array}{c|ccc}
c_1 & a_{11} & \cdots & a_{1s} \\
\vdots & \vdots & & \vdots \\
c_s & a_{s1} & \cdots & a_{ss} \\
\hline
 & b_1 & \cdots & b_s
\end{array}
$$

Table 2.1: A general Butcher tableau.

and $c = [c_1, c_2, \ldots, c_s]^T \in \mathbb{R}^s$ define the method. For more information on Runge-Kutta methods, see the literature, for instance [14, 26].

2.4 Runge-Kutta-Munthe-Kaas methods

Let $\mathcal{M}$ be a differentiable manifold, where a given ODE is evolving. Assume that the ODE is written on the form

$$\dot{y} = F(t, y), \quad y(0) = p \in \mathcal{M}, \tag{2.19}$$

where $F : \mathbb{R} \times \mathcal{M} \mapsto T\mathcal{M}$. The function $F(t, \cdot) \in \mathfrak{X}(\mathcal{M})$ is a vector field, $y(t) \in \mathcal{M}$ is the flow, and $\dot{y} \in T_y\mathcal{M}$ is the tangent vector at y [20]. In general the manifold is nonlinear, hence, solving the ODE and determining the numerical flow is not necessarily trivial. However, suppose we have a Lie group G acting on $\mathcal{M}$, with $\mathfrak{g}$ its corresponding Lie algebra. As earlier mentioned, the Lie algebra is a vector space, which makes it highly suitable for solving ODEs numerically. In fact, solving the ODE in the Lie algebra of the Lie group instead of on the manifold is the main idea behind the Runge-Kutta-Munthe-Kaas methods [8, 12, 19, 23].

We will now give a derivation of the RKMK methods. Suppose we have an arbitrary curve $\gamma(t) \in \mathcal{M}$, where

$$\gamma(0) = m \in \mathcal{M}, \quad \dot{\gamma}(0) = v_m \in T_m\mathcal{M}.$$

By applying the Lie group action Λ (2.8), we can represent some segment of γ by a curve $g(t)$ on G, where

$$g(0) = e \in G, \quad \dot{g}(0) = \xi \in \mathfrak{g}.$$

Assuming Λ is locally transitive we get

$$\gamma(t) = \Lambda_m(g(t)), \quad \gamma(0) = \Lambda_m(g(0)).$$

We want to represent $v_m = \dot{\gamma}(0)$ by means of elements in the Lie algebra of the Lie group, thus exploiting the linear structure of the vector space. From section 2.1.1 we know that for any element $\xi \in \mathfrak{g}$ there exists a vector field on $\mathcal{M}$, namely [7, 22]

$$X_\xi(m) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \gamma(t) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \Lambda(g(t), m) = (\lambda_* \xi)(m).$$

Thus, we have a mapping from $\mathfrak{g}$ to the vector field on $\mathcal{M}$, namely the infinitesimal generators of the Lie algebra actions

$$\lambda_* : \mathfrak{g} \mapsto \mathfrak{X}(\mathcal{M}).$$

The space of all vector fields X_ξ of the above form is finite dimensional and linear, while $\mathfrak{X}(\mathcal{M})$ is infinite dimensional. Since $F \in \mathfrak{X}(\mathcal{M})$, we do not expect F to be on the same form as X_ξ. We may fix this by exchanging $\xi \in \mathfrak{g}$ with some smooth function $f : \mathbb{R} \times \mathcal{M} \mapsto \mathfrak{g}$. Then, any element in $\mathfrak{X}(\mathcal{M})$ can be written on the form

$$X_{f(t,m)}(m) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \gamma(t) = (\lambda_* f(t,m))(m).$$

If the ODE (2.19) has the form $F(t,m) = (\lambda_* f(t,m))(m)$, we say that it is written in the generic format [23].

If $\tilde{f}(t,\sigma) \in \mathfrak{X}(\mathfrak{g})$ is the vector field

$$\tilde{f}(t,\sigma) = \mathrm{dexp}_\sigma^{-1}(f(t, \lambda(\sigma, m))), \tag{2.20}$$

then F is λ_m-related (2.3) to $\tilde{f}$. For a proof we refer to the literature, e.g. [20].

Recall section 2.2.9, where we showed that segment of a curve $\sigma(t) \in \mathfrak{g}$ could be represented via the exponential map on G, such that $g(t) = \exp(\sigma(t))$. Using (2.20) we may, for t small enough, locally represent the solution $\gamma(t)$ close to $\gamma(0) = m$ as

$$\gamma(t) = \Lambda(g(t), m) = \Lambda(\exp(\sigma(t)), m) = \lambda(\sigma(t), m), \tag{2.21}$$

where $\sigma(t)$ satisfies the ODE

$$\begin{aligned} \dot{\sigma}(t) &= \tilde{f}(t,\sigma) = \mathrm{dexp}_\sigma^{-1}(f(t,\lambda(\sigma,m))) = \sum_{k=0}^{\infty} \frac{B_k}{k!} \mathrm{ad}_\sigma^k f(t, \lambda(\sigma, m)), \\ \sigma(0) &= 0 \in \mathfrak{g}. \end{aligned} \tag{2.22}$$

We can approximate $\mathrm{dexp}_\sigma^{-1}$ by a truncated sum, which we then use to numerically solve the differential equation (2.22) using any classical Runge–Kutta (RK) methods (implicit or explicit). The truncated sum has the expression

$$\mathrm{dexp}_\sigma^{-1}(f(t,\lambda(\sigma,m))) \approx \sum_{k=0}^{N-1} \frac{B_k}{k!} \mathrm{ad}_\sigma^k f(t, \lambda(\sigma, m)),$$

where the order of the Runge-Kutta method is equal to N.

From the previous discussion we can now formulate a numerical method for ODEs on manifolds. That is, we construct a vector field $\dot{\sigma}(t) = \tilde{f}(t,\sigma)$

on $\mathfrak{g}$ the Lie algebra of the acting group G through the map λ_m. Then, we take one step with a classical Runge-Kutta method to the ODE on $\mathfrak{g}$ and determine the numerical flow on the Lie algebra. Next, we map the flow on $\mathfrak{g}$ back to $\mathcal{M}$ via λ_m, and determine the numerical flow on $\mathcal{M}$. We do this repeatedly until the desired number of steps has been taken. This is called the Runge-Kutta-Munthe-Kaas methods [14, 16].

Recall that the generic format for ODEs to be solved using RKMK methods is [20]

$$\dot{\gamma}(t) = (\lambda_* f(t, y))(\gamma), \quad \gamma(0) = m, \tag{2.23}$$

which also can be seen from the diagram in figure 2.3 [19].

$$\begin{tikzcd}
T\mathfrak{g} \arrow[r, "\lambda_{m*}"] & T\mathcal{M} \\
\mathfrak{g} \arrow[u, "\tilde{f}"] \arrow[r, "\lambda_m"'] & \mathcal{M} \arrow[u, "F"']
\end{tikzcd}$$

Figure 2.3: Commutative diagram.

Summing up:

To solve an ODE on a Lie group, $\dot{y} = F(t, y)$, $y(0) = m$, using an RKMK method:

1. Determine the manifold $\mathcal{M}$, the Lie group G and the Lie algebra $\mathfrak{g}$.
2. Using the connection between the Lie group action Λ and the Lie algebra action λ, write the ODE on the form $F(y) = (\lambda_* f(y))(y)$, $f : \mathbb{R} \times \mathcal{M} \mapsto \mathcal{M}$.
3. Determine the function f from the given ODE.
4. Map the ODE from the vector field on $\mathcal{M}$ to the vector field on the Lie algebra $\mathfrak{g}$.
5. Determine the flow on $\mathfrak{g}$ using a classical RK-method.

6. Map the flow on $\mathfrak{g}$ back to the vector field on $\mathcal{M}$.

Algorithm 1: The Runge-Kutta-Munthe-Kaas methods

Initialize y_0

for $n = 1, 2, \ldots$
  for $i = 1, 2, \ldots, s$
    $u_i = h\sum_{j=1}^{i-1} a_{i,j}k_j$
    $v_i = f(\lambda(u_i, y_n))$
    $k_i = \text{dexpinv}(u_i,\ v_i,\ N)$
  end
  $\sigma = h\sum_{i=1}^{s} b_i k_i$
  $y_{n+1} = \lambda(\sigma, y_n)$.
end

2.5 Lagrangian mechanics

In Lagrangian mechanics one considers a configuration space which has the structure of a differentiable manifold, in addition to having a Lagrangian function evolving on its tangent bundle. We consider cases when the configuration space is a Lie group, and when the equations of motion are derived from variational principles. Then, the phase space is typically the tangent or cotangent bundle of the Lie group, and Lie group methods can be used to determine the numerical solutions [7].

By introducing a change of variables known as the Legendre transform a Lagrangian system may be referred to as Hamiltonian. Then, the phase space is the cotangent bundle of the configuration space, which is an even-dimensional manifold with a symplectic structure. In addition, there exists a Hamiltonan function which is the Legendre transform of the Lagrangian function. In a dynamical system, the configuration space usually represents the positions and the phase space represents all possible values of position and momentum variables [3].

The following theory of Lagrangian mechanics is based on [24].

In a mechanical system, assume G is the configuration manifold of the generalized position coordinates. In particular, suppose G is a Lie group and $\mathfrak{g}$ its corresponding Lie algebra. Assume we have a Lagrangian function defined on the tangent bundle TG of G, that is, $L : TG \mapsto \mathbb{R}$. We define the Lagrangian action functional on curves $g(t) \in G$ by

$$S[g] = \int_{t_0}^{t_1} L(t, g(t), \dot{g}(t))\ \mathrm{d}t, \quad t_0 \leq t \leq t_1. \tag{2.24}$$

We introduce a perturbation curve $\delta g(t) \in T_{g(t)}G$, and take the variation (where by abuse of notation we write $g + \epsilon\delta g$ for a perturbed version of g)

$$\delta S[g] = \left.\frac{\mathrm{d}}{\mathrm{d}\epsilon}\right|_{\epsilon=0} S[g + \epsilon\delta g] \tag{2.25}$$

for all perturbations $\delta g(t)$.

Then, by combining (2.24) and (2.25) we get

$$\begin{aligned} \delta S[g] &= \left.\frac{\mathrm{d}}{\mathrm{d}\epsilon}\right|_{\epsilon=0} \int_{t_0}^{t_1} L(t,\ g(t) + \epsilon\delta g(t),\ \dot{g}(t) + \epsilon\delta\dot{g}(t))\ \mathrm{d}t \\ &= \int_{t_0}^{t_1} \left\langle \frac{\partial L}{\partial g}(t, g(t), \dot{g}(t)), \delta g(t) \right\rangle + \left\langle \frac{\partial L}{\partial \dot{g}}(t, g(t), \dot{g}(t)), \delta\dot{g}(t) \right\rangle\ \mathrm{d}t, \end{aligned} \tag{2.26}$$

where both $\frac{\partial L}{\partial g}(t, g(t), \dot{g}(t))$ and $\frac{\partial L}{\partial \dot{g}}(t, g(t), \dot{g}(t)) \in T_g^*G$. Using integration by parts in the last term gives

$$\begin{aligned} \left\langle \frac{\partial L}{\partial \dot{g}}(t, g(t), \dot{g}(t)), \delta\dot{g}(t) \right\rangle\ \mathrm{d}t = &\left.\left\langle \frac{\partial L}{\partial \dot{g}}(t, g(t), \dot{g}(t)), \delta g(t) \right\rangle\right|_{t=t_0}^{t=t_1} \\ &- \int_{t_0}^{t_1} \left\langle \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{g}}(t, g(t), \dot{g}(t)), \delta g(t) \right\rangle\ \mathrm{d}t, \end{aligned}$$

which we insert into (2.26).

Hamilton's principle states that the motion of the system follows the curve g extremizing S, that is, $\delta S[g] = 0$. Assuming this holds for all $\delta g(t)$, a lemma due to Lagrange [27] implies that the variation $\delta g(t)$ vanishes at the end points. Hence, we obtain the Euler-Lagrange equations

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{g}} = \frac{\partial L}{\partial g},$$

also called the Lagrangian equations of motion. We have omitted the dependent variables for aesthetic reasons.

We define the conjugate momentum variable $p \in T_g^*G$, via a change of variables known as the Legendre transform

$$p = \frac{\partial L}{\partial \dot{g}}(t, g(t), \dot{g}(t)). \tag{2.27}$$

To make the equations more readable, we will omit the time dependency from now on.

The Legendre transform is well-defined if it is invertible with respect to $\dot{g}$. So, assuming invertibility we can write $\dot{g}$ as a function of g and p, thus obtaining the ODE system

$$\dot{g} = \phi(g, p), \quad \dot{p} = \frac{\partial L}{\partial g}(g, \phi(g, p)),$$

for some function ϕ.

The Hamiltonian function is defined as

$$H(g, p) = \langle p, \dot{g} \rangle - L(g, \dot{g}).$$

We insert the new variables (2.27) into the Hamiltonian and get

$$H(g, p) = \langle p, \phi(g, p) \rangle - L(g, \phi(g, p)). \tag{2.28}$$

By differentiating (2.28) with respect to g and p we observe that

$$
\begin{aligned}
\frac{\partial H}{\partial p}(g,p) &= \phi(g,p) + \frac{\partial \phi}{\partial p}(g,p)p - \frac{\partial \phi}{\partial p}(g,p)\frac{\partial L}{\partial \dot{g}}(g,\phi(g,p)) \\
&= \phi(g,p) + \frac{\partial \phi}{\partial p}(g,p)p - \frac{\partial \phi}{\partial p}(g,p)p \\
&= \phi(g,p) \\
&= \dot{g}, \\
\\
\frac{\partial H}{\partial g}(g,p) &= \frac{\partial \phi}{\partial g}(g,p)p - \frac{\partial L}{\partial g}(g,p) - \frac{\partial \phi}{\partial g}(g,p)\frac{\partial L}{\partial \dot{g}}(g,\phi(g,p)) \\
&= \frac{\partial \phi}{\partial g}(g,p)p - \frac{\partial L}{\partial g}(g,p) - \frac{\partial \phi}{\partial g}(g,p)p \\
&= -\frac{\partial L}{\partial g}(g,p) \\
&= -\dot{p},
\end{aligned}
\tag{2.29}
$$

where we take the point of view that $\frac{\partial \phi}{\partial p} : T_g^*G \mapsto T_gG$ and $\frac{\partial \phi}{\partial g} : T_g^*G \mapsto T_g^*G$.

For simplicity, we rewrite (2.29) without the dependent variables, thus, the Hamiltonian equations of motion are

$$
\dot{g} = \frac{\partial H}{\partial p}, \quad \dot{p} = -\frac{\partial H}{\partial g}.
$$

As explained earlier, a tangent vector $v_g \in T_gG$ of any element $g \in G$ can be represented by either left or right trivialisation. Hence, by abuse of notation

$$
v_g = L_{g*}\xi = g\xi, \text{ or } v_g = R_{g*}\xi = \xi g,
$$

where $\xi \in \mathfrak{g}$.

Consider the left-trivialised case. Then,

$$
L(g,\dot{g}) = L(g,g\xi) := \ell(g,\xi), \quad \xi = g^{-1}\dot{g}.
$$

We introduce a perturbation curve $\delta g(t) \in T_{g(t)}G$ as above which we also trivialise, such that $\delta g(t) = g\eta$. Then, we take the variation

$$\begin{aligned}
\delta S[g] &= \int_{t_0}^{t_1} \left.\frac{\mathrm{d}}{\mathrm{d}\epsilon}\right|_{\epsilon=0} \ell(g+\epsilon\delta g,\ (g+\epsilon\delta g)^{-1})(\dot g + \epsilon\delta\dot g)\ \mathrm{d}t \\
&= \int_{t_0}^{t_1} \left\langle \frac{\partial \ell}{\partial g}(g,\xi), \delta g \right\rangle + \left\langle \frac{\partial \ell}{\partial \xi}(g,\xi), g^{-1}\ \delta\dot g - g^{-1}\ \delta g\ g^{-1}\dot g \right\rangle\ \mathrm{d}t \\
&= \int_{t_0}^{t_1} \left\langle \frac{\partial \ell}{\partial g}(g,\xi), \delta g \right\rangle + \left\langle \frac{\partial \ell}{\partial \xi}(g,\xi), g^{-1}\ \delta\dot g - \eta\xi \right\rangle\ \mathrm{d}t.
\end{aligned}$$

Next, we calculate

$$g^{-1}\ \delta\dot g = g^{-1}\frac{\mathrm{d}}{\mathrm{d}t}(g\eta) = g^{-1}\ \dot g\eta + \dot\eta = \dot\eta + \xi\eta,$$

such that, using integration by parts, we get

$$\begin{aligned}
\delta S[g] &= \int_{t_0}^{t_1} \left\langle L_g^*\frac{\partial \ell}{\partial g}(g,\xi), \eta \right\rangle + \left\langle \frac{\partial \ell}{\partial \xi}(g,\xi), \dot\eta + \mathrm{ad}_\xi \eta \right\rangle\ \mathrm{d}t \\
&= \left.\left\langle \frac{\partial \ell}{\partial \xi}(g,\xi), \eta \right\rangle\right|_{t=t_0}^{t=t_1} + \int_{t_0}^{t_1} \left\langle L_g^*\frac{\partial \ell}{\partial g}(g,\xi) - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \ell}{\partial \xi}(g,\xi) + \mathrm{ad}_\xi^*\frac{\partial \ell}{\partial \xi}(g,\xi), \eta \right\rangle\ \mathrm{d}t.
\end{aligned}$$

Assuming this holds for all $\delta g(t)$ as above, we once again apply the Lagrange lemma [27], which implies that the variation $\delta g(t)$ vanishes at the end points, and thus, so will η. Then we get the left-trivialised Euler-Lagrange equations

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \ell}{\partial \xi}(g,\xi) = L_g^*\frac{\partial \ell}{\partial g}(g,\xi) + \mathrm{ad}_\xi^*\frac{\partial \ell}{\partial \xi}(g,\xi).$$

Applying the Legendre transform once more, we get

$$\mu = \frac{\partial \ell}{\partial \xi}(g,\xi) \in \mathfrak{g}^*,$$

such that, assuming invertibility,

$$\xi = \phi(g,\mu),\ \ \phi : G\times\mathfrak{g}^* \mapsto \mathfrak{g},$$

for some function ϕ. Hence, we obtain the ODE system

$$\dot g = g\xi = L_{g*}\phi(g,\mu) = g\phi(g,\mu), \quad \dot\mu = L_g^*\frac{\partial \ell}{\partial g}(g,\phi(g,\mu)) + \mathrm{ad}_{\phi(g,\mu)}^*\mu. \quad (2.30)$$

The right-trivialised case follows in similar steps, resulting in the right-trivialised Euler-Lagrange equations

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \ell}{\partial \xi}(g,\xi) = R_g^*\frac{\partial \ell}{\partial g}(g,\xi) - \mathrm{ad}_\xi^*\frac{\partial \ell}{\partial \xi}(g,\xi).$$

Using the Legendre transform, we get

$$\mu = \frac{\partial \ell}{\partial \xi}(g,\xi) \in \mathfrak{g}^*, \text{ such that } \xi = \phi(g,\mu), \phi : G \times \mathfrak{g}^* \mapsto \mathfrak{g},$$

for some function ϕ, assuming the map is invertible. Hence, we obtain the ODE system

$$\dot{g} = \xi g = R_{g*}\phi(g,\mu) = \phi(g,\mu)g, \quad \dot{\mu} = R_g^*\frac{\partial \ell}{\partial g}(g,\phi(g,\mu)) - \mathrm{ad}^*_{\phi(g,\mu)}\mu.$$

The ODE systems derived from the Euler-Lagrange equations may be interpreted as left and right trivialised Hamiltonian systems. We find that by choosing the functional derivatives of the Hamiltonian as

$$\begin{aligned}\frac{\delta H}{\delta g} &= -L_g^*\frac{\partial \ell}{\partial g}(g,\phi(g,\mu)),\\ \frac{\delta H}{\delta \mu} &= \phi(g,\mu),\end{aligned}$$

the Hamiltonian vector field in the left trivialised case is [12]

$$X_H^L(g,\mu) = \left(L_{g*}\frac{\delta H}{\delta \mu}, -\frac{\delta H}{\delta g} + \mathrm{ad}^*_{\frac{\delta H}{\delta \mu}}\mu\right), \tag{2.31}$$

and similarly, by choosing

$$\begin{aligned}\frac{\delta H}{\delta g} &= -R_g^*\frac{\partial \ell}{\partial g}(g,\phi(g,\mu)),\\ \frac{\delta H}{\delta \mu} &= \phi(g,\mu),\end{aligned}$$

the Hamiltonian vector field in the right trivialised case is [12]

$$X_H^R(g,\mu) = \left(R_{g*}\frac{\delta H}{\delta \mu}, -\frac{\delta H}{\delta g} - \mathrm{ad}^*_{\frac{\delta H}{\delta \mu}}\mu\right). \tag{2.32}$$

In section 2.7 we will show how to connect this to the RKMK methods.

2.6 The Lie group structure of cotangent bundles of Lie groups

In the following sections we give a detailed description of how to solve mechanical systems using the RKMK methods when the configuration space is a Lie group and the phase space is its cotangent bundle. We start by showing how to consider the cotangent bundle as a Lie group in its own right by first determining its binary operation.

Assume we have a Lie group G with its corresponding Lie algebra $\mathfrak{g}$. Let $(g, p) \in T^*G$ such that $g \in G$ and $p \in T_g^*G$. We can associate p with $\mu \in \mathfrak{g}^*$, by using left or right trivialisation, that is, by using the induced cotangent maps we can write, $\mu = L_g^* p$ or $\mu = R_g^* p$, respectively. Hence, we have a mapping

$$(g, p) \mapsto (g, \mu),$$

which is in fact an isomorphism of tangent spaces [7].

Engø explains in [12] that group multiplication on the tangent or cotangent bundle of a Lie group is induced by the group multiplication in G, and therefore can be interpreted as a lift of the binary operation in G up to TG or T^*G, respectively. In addition, he states that even though the lifted and the trivial group structures are fundamentally different, the lifted multiplication inherit the original geometric structure of the Lie group G. Engø also gives an approach on how to derive an induced group product on the tangent bundle. For a wider understanding, we will repeat some of this method.

Assume the elements (g, v) and (h, u) belong to $G \times \mathfrak{g}$. Let $g(t)$ and $h(t)$ be curves in G, where $g(0) = g$ and $h(0) = h$. Then, we may associate $g(t)$ to (g, v) and $h(t)$ to (h, u). By using either left or right trivialisation we can give a representation of the curves. We only do this in the left trivialised case, then

$$\begin{aligned} g(t) &= g + tgv + \mathcal{O}(t^2), \\ h(t) &= h + thu + \mathcal{O}(t^2), \end{aligned}$$

where gv and hu are the tangent vectors in T_gG and T_hG, respectively. Multiplying the curves together using the group multiplication in G gives

$$\begin{aligned} g(t)h(t) &= (g + tgv + \mathcal{O}(t^2))(h + thu + \mathcal{O}(t^2)) \\ &= gh + tgh(u + \mathrm{Ad}_{h^{-1}}v) + \mathcal{O}(t^2), \end{aligned}$$

where $\mathrm{Ad}_{h^{-1}}$ is the adjoint representation defined in (2.14). Thus, in the left trivialised case, the group multiplication on G induces the following group multiplication on T^*G, namely

$$(g, v) \cdot (h, u) = (gh, u + \mathrm{Ad}_{h^{-1}} v).$$

Also, in a similar way the group multiplication on T^*G in the left and right trivialised cases can be determined. However, since we are dealing with cotangent vectors instead of tangent vectors, we must do some adjustments to the product. Let μ and η be cotangent vectors in g^*, such that (g, μ) and (h, η) are elements in T^*G. Then, in the left trivialised case

$$(g, \mu) \cdot (h, \eta) = (gh, \eta + \mathrm{Ad}^*_h \mu), \tag{2.33}$$

where $\cdot$ denotes multiplication between elements in $G \times \mathfrak{g}^*$.

Similarly, in the right trivialised case,

$$(g, \mu) \cdot (h, \eta) = (gh, \mu + \mathrm{Ad}^*_{h^{-1}} \eta). \tag{2.34}$$

We call (2.33) and (2.34) the left and right semidirect product on T^*G, respectively. Hence, we can associate T^*G with $G \ltimes \mathfrak{g}^*$ and write $\mathcal{G} = G \ltimes \mathfrak{g}^* \cong T^*G$, where $\ltimes$ denotes the semidirect group structure. For more information see the literature [3, 18].

We will now check if $\mathcal{G}$ is a Lie group. First we examine whether the semidirect products satisfy the properties of a group product, that is, it is associative (2.6) if there exists an identity element and if every element in $\mathcal{G}$ has an inverse element (2.7). We start by testing the right semidirect product (2.34). Let $g_1, g_2, g_3 \in G$ and $\mu_1, \mu_2, \mu_3 \in \mathfrak{g}^*$, then

- Associativity:

$$\begin{aligned}
((g_1, \mu_1) \cdot (g_2, \mu_2)) \cdot (g_3, \mu_3) &= (g_1 g_2, \mu_1 + \mathrm{Ad}^*_{g_1^{-1}} \mu_2) \cdot (g_3, \mu_3) \\
&= (g_1 g_2 g_3, \mu_1 + \mathrm{Ad}^*_{g_1^{-1}} \mu_2 + \mathrm{Ad}^*_{g_1^{-1}} \mathrm{Ad}^*_{g_2^{-1}} \mu_3) \\
&= (g_1, \mu_1) \cdot ((g_2, \mu_2) \cdot (g_3, \mu_3)).
\end{aligned}$$

- Existence of an identity element $(e, 0) \in \mathcal{G}$:

$$\begin{aligned}
(g, \mu) \cdot (e, 0) &= (ge, \mu + \mathrm{Ad}^*_{g^{-1}} 0) = (g, \mu), \\
(e, 0) \cdot (g, \mu) &= (eg, 0 + \mathrm{Ad}^*_{e^{-1}} \mu) = (g, \mu).
\end{aligned}$$

- Existence of an inverse element of $(g, \mu) \in \mathcal{G}$:

$$
\begin{aligned}
(g, \mu) \cdot (h, \eta) &= (gh, \mu + \mathrm{Ad}^*_{g^{-1}}\eta) = (e, 0), \\
&\Rightarrow h = g^{-1},\ \eta = -\mathrm{Ad}^*_g \mu, \\
(h, \eta) \cdot (g, \mu) &= (hg, \eta + \mathrm{Ad}^*_{h^{-1}}\mu) = (e, 0), \\
&\Rightarrow g = h^{-1},\ \mu = -\mathrm{Ad}^*_h \eta,
\end{aligned}
$$

where we have used lemma 2.7.

Hence, (2.34) satisfies all the properties of a group product and we call it a right semidirect product on $\mathcal{G}$. Also, since G is a Lie group and Ad^*_g is a linear operator, both the semidirect product and the inversion map are smooth, such that $\mathcal{G}$ is indeed a Lie group. The same applies for the left trivialised semidirect product (2.33).

In section 2.2.1, which discussed multiplication on Lie groups, we learned about the importance of the identity element e, and that the left and right group multiplication induced tangent maps between the Lie algebra and the other spaces in the tangent bundle. Since we have just seen that $\mathcal{G} = G \ltimes \mathfrak{g}^*$ is a Lie group with a group product, we will now determine its tangent map.

A way of determining the tangent map is to choose an arbitrary curve $(h(t), \eta(t))$ on $\mathcal{G}$, where

$$
\begin{aligned}
h(0) &= h,\ \dot{h}(0) = v, \\
\eta(0) &= \eta,\ \dot{\eta}(0) = \zeta.
\end{aligned}
$$

Observing that $(v, \zeta) \in T_{(h,\eta)}\mathcal{G}$ is the tangent vector of $(h(0), \eta(0))$, we can search for a tangent vector $\xi \in \mathfrak{g}$ such that $v = R_{h*}\xi$. Then, we can represent any vector in $T\mathcal{G}$ by using vectors from the Lie algebra. By letting $(h(t), \eta(t)) = (e(t)h, \eta(t))$, where $e(0) = e, \dot{e}(0) = \xi$, the right tangent mapping on $T\mathcal{G}$ becomes

$$
\begin{aligned}
R_{(g,\mu)*}(v,\zeta) &= \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} R_{(g,\mu)}(e(t)h,\eta(t)) \\
&= \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} (e(t)h,\eta(t))\cdot(g,\mu) \\
&= \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} (e(t)hg,\eta(t)+\mathrm{Ad}^*_{(e(t)h)^{-1}}\mu) \\
&= (R_{hg*}\xi,\zeta+\frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0}\mathrm{Ad}^*_{h^{-1}e(t)^{-1}}\mu) \\
&= (R_{hg*}\xi,\zeta-\mathrm{ad}^*_\xi\mathrm{Ad}^*_{h^{-1}}\mu),
\end{aligned}
$$

where we have used lemma 2.7 and the coadjoint map.

It is the tangent map at the identity $(e,0)\in\mathcal{G}$ that is most applicable, which we find by setting $(h,\eta)=(e,0)$ in the above equation. That is, the right tangent map in the identity is

$$R_{(g,\mu)*}(\xi,\zeta) = (R_{g*}\xi,\zeta-\mathrm{ad}^*_\xi\mu). \tag{2.35}$$

Similarly, we can determine the left tangent map in the identity derived from the left trivialised semidirect product (2.33), that is

$$L_{(g,\mu)*}(\xi,\zeta) = (L_{g*}\xi,\zeta+\mathrm{ad}^*_\xi\mu). \tag{2.36}$$

2.7 RKMK methods on the cotangent bundle

The following text is based on the theory of Runge-Kutta-Munthe-Kaas methods given in section 2.4, and our intention is to tailor the RKMK methods to problems where the phase space is T^*G. We will use the approach given under the head line “Summing up” in section 2.4.

Assume we have a Lie group $\mathcal{G}=T^*G\cong G\ltimes\mathfrak{g}^*$ that acts on itself, where its corresponding Lie algebra is $\tilde{\mathfrak{g}}=\mathfrak{g}\ltimes\mathfrak{g}^*$. If an ODE

$$(\dot{g},\dot{\mu}) = F(g,\mu), \tag{2.37}$$

where $F:\mathcal{G}\mapsto\tilde{\mathfrak{g}}$, is evolving on $\mathcal{G}$, we first need to rewrite the ODE on the generic form (2.23)

$$(\dot{g},\dot{\mu}) = (\lambda_* f(g,\mu))(g,\mu),$$

where $(g, \mu) \in \mathcal{G}$, $(\dot{g}, \dot{\mu}) \in \tilde{\mathfrak{g}}$ and $f : \mathcal{G} \mapsto \tilde{\mathfrak{g}}$ is a smooth function.

We do this by using the definition of λ_* and its connection to the group multiplication, that is

$$(\lambda_* f(g,\mu))(g,\mu) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \lambda(tf(g,\mu),(g,\mu)) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \Lambda(\exp(tf(g,\mu)),(g,\mu)).$$

The derivative of the group multiplication in the latter term is equal to either of the tangent maps (2.35) or (2.36) depending on whether right or left trivialisation are applied, respectively.

We can split the function f into two components

$$f_g : \mathcal{G} \mapsto \mathfrak{g},$$
$$f_\mu : \mathcal{G} \mapsto \mathfrak{g}^*,$$

and then, using the right tangent map, the generic form of (2.37) becomes

$$(\dot{g}, \dot{\mu}) = (R_{g*} f_g, f_\mu - \mathrm{ad}^*_{(f_g)} \mu). \tag{2.38}$$

Similarly, using the left tangent map, the generic form of (2.37) becomes

$$(\dot{g}, \dot{\mu}) = (L_{g*} f_g, f_\mu + \mathrm{ad}^*_{f_g} \mu), \tag{2.39}$$

where we omit the dependent variables to make the equations more readable.

Example 2.12. For matrix Lie groups, the generic format of the ODE (2.37) becomes either

$$(\dot{g}, \dot{\mu}) = (f_g g, f_\mu - \mathrm{ad}^*_{f_g} \mu),$$

or

$$(\dot{g}, \dot{\mu}) = (g f_g, f_\mu + \mathrm{ad}^*_{f_g} \mu),$$

depending on whether we use right and left trivialisation, respectively. ■

Note that we also have a connection between the RKMK methods and the trivialised Euler-Lagrange equations including the Hamiltonian vector fields described in the end of section 2.5. That is, by comparing (2.39) with (2.30) and (2.38) with (2.32), we can express the Euler-Lagrange equations on $\mathfrak{g} \ltimes \mathfrak{g}^*$ by setting

$$
\begin{aligned}
f_g(g,\mu) &= \phi(g,\mu) = \frac{\delta H}{\delta \mu},\\
f_\mu(g,\mu) &= L_g^* \frac{\partial \ell}{\partial g}(g,\phi(g,\mu)) = -\frac{\delta H}{\delta g},
\end{aligned}
$$

using right trivialisation, and similarly

$$
\begin{aligned}
f_g(g,\mu) &= \phi(g,\mu) = \frac{\delta H}{\delta \mu},\\
f_\mu(g,\mu) &= R_g^* \frac{\partial \ell}{\partial g}(g,\phi(g,\mu)) = -\frac{\delta H}{\delta g}
\end{aligned}
$$

using left trivialisation [12, 24].

The coordinate map ϕ in (2.11), later taken to be the exponential function, also undergoes a substantial change in the current setting. The map takes elements in the Lie algebra to elements in the Lie group, that is

$$\exp : \mathfrak{g} \ltimes \mathfrak{g}^* \mapsto G \ltimes \mathfrak{g}^*.$$

Let $(\xi,\alpha) \in \mathfrak{g} \ltimes \mathfrak{g}^*$. To find the flow of (2.38) or (2.39), setting $f_g = \xi$ and $f_\mu = \alpha$, e.g. solve

$$\dot{g} = R_{g*}\xi,\ \dot{\mu} = \alpha - \mathrm{ad}_\xi^*\mu,\ g(0) = e, \mu(0) = 0$$

to obtain curves $g(t), \mu(t)$. Then

$$\exp(\xi,\alpha) = (g(1),\mu(1)) = (\exp_G, \mathrm{dexp}_\xi^*\alpha),$$

where dexp_ξ^* is the dual of dexp_ξ [12].

The ad operator defined in (2.15) is needed in the RKMK methods and we will now derive it for the Lie algebra $\mathfrak{g} \ltimes \mathfrak{g}^*$. Let $(g(t),\mu(t))$ and $(h(s),\eta(s))$ be curves in $\mathcal{G}$ where

$$
\begin{aligned}
(g(0),\mu(0)) &= (h(0),\eta(0)) = (e,0) \in \mathcal{G},\\
(\dot{g}(0),\dot{\mu}(0)) &= (\xi,\zeta) \in \tilde{\mathfrak{g}},\\
(\dot{h}(0),\dot{\eta}(0)) &= (v,\beta) \in \tilde{\mathfrak{g}}.
\end{aligned}
$$

The ad operator is the same independent of whether we use left or right multiplication, so we only compute it using left multiplication. First, observe that

$$
\begin{aligned}
(g(t),\mu(t))&\cdot(h(s),\eta(s))\cdot(g(t),\mu(t))^{-1}\\
&= (g(t)(h(s),\mathrm{Ad}^*_{h(s)}\mu(t)+\eta(s))\cdot(g(t)^{-1},-\mathrm{Ad}^*_{g(t)^{-1}}\mu(t))\\
&= (g(t)h(s)g(t)^{-1},\mathrm{Ad}^*_{g(t)^{-1}}(\mathrm{Ad}^*_{h(s)}\mu(t)+\eta(s))-\mathrm{Ad}^*_{g(t)^{-1}}\mu(t))
\end{aligned}
$$

Then, by differentiating, the ad operator becomes

$$
\begin{aligned}
\mathrm{ad}_{(\xi,\zeta)}(v,\beta) &= \frac{\partial^2}{\partial t\partial s}\bigg|_{t=s=0}(g(t)h(s)g(t)^{-1},\mathrm{Ad}^*_{g(t)^{-1}}(\mathrm{Ad}^*_{h(s)}\mu(t)+\eta(s))-\mathrm{Ad}^*_{g(t)^{-1}}\mu(t))\\
&= (\mathrm{ad}_\xi v,-\mathrm{ad}^*_\xi\beta+\mathrm{ad}^*_v\zeta).
\end{aligned}
$$

The next step is stating the flow of (2.37) on $\mathcal{G}$, similarly to what was done in section 2.4 with $(\sigma_g,\sigma_\mu)\in\mathfrak{g}\ltimes\mathfrak{g}^*$ taking the place of σ. That is, for small values of t

$$
\begin{aligned}
(g(t),\mu(t)) &= \lambda^L_{(g,\mu)}(\sigma_g,\sigma_\mu)\\
&= \Lambda^L_{(g,\mu)}\exp(\sigma_g,\sigma_\mu)\\
&= R_{(g,\mu)}(\exp(\sigma_g),\mathrm{dexp}^*_{\sigma_g}(\sigma_\mu))\\
&= (\exp(\sigma_g)g,\mathrm{Ad}^*_{\exp(-\sigma_g)}\mu+\mathrm{dexp}^*_{\sigma_u}(\sigma_\mu)),
\end{aligned}
\tag{2.40}
$$

when using right trivialisation, and

$$
\begin{aligned}
(g(t),\mu(t)) &= \lambda^R_{(g,\mu)}(\sigma_g,\sigma_\mu)\\
&= \Lambda^R_{(g,\mu)}\exp(\sigma_g,\sigma_\mu)\\
&= L_{(g,\mu)}(\exp(\sigma_g),\mathrm{dexp}^*_{\sigma_g}(\sigma_\mu))\\
&= (g\ \exp(\sigma_g),\mathrm{Ad}^*_{\exp(\sigma_g)}\mu+\mathrm{dexp}^*_{\sigma_g}\sigma_\mu),
\end{aligned}
\tag{2.41}
$$

when using left trivialisation [20].

From the discussion in section 2.4, we know that $(\sigma_g(t),\sigma_\mu(t))$ satisfies the ODE

$$
\begin{aligned}
(\dot\sigma_g(t),\dot\sigma_\mu(t)) &= \mathrm{dexp}^{-1}_{(\sigma_g,\sigma_\mu)}(f(t,\lambda((\sigma_g,\sigma_\mu),(g_0,\mu_0))))\\
&= \sum_{k=0}^{\infty}\frac{B_k}{k!}\mathrm{ad}^k_{(\sigma_g,\sigma_\mu)}f(t,(g_0,\mu_0)),\\
(\sigma_g(0),\sigma_\mu(0) &= (0,0)\in\tilde{\mathfrak{g}},
\end{aligned}
$$

where $\mathrm{dexp}^{-1}_{(\sigma_g,\sigma_\mu)}$ can be approximated by the truncated sum

$$\sum_{k=0}^{N-1} \frac{B_k}{k!} \mathrm{ad}^k_{(\sigma_g,\sigma_\mu)}(f_g, f_\mu).$$

Algorithm 2:
An explicit RKMK algorithm on $\mathcal{G} = G \ltimes \mathfrak{g}^*$

Initialize (g_0, μ_0).

for $n = 1, 2, \ldots$
  for $i = 1, 2, \ldots, s$

$$(u_{g,i}, u_{\mu,i}) = h \sum_{j=1}^{i-1} a_{i,j}(k_{g,j}, k_{\mu,j})$$

$$(v_{g,i}, v_{\mu,i}) = f(\lambda((u_{g,i}, u_{\mu,i}), (g_n, \mu_n)))$$

$$(k_{g,i}, k_{\mu,i}) = \mathrm{dexpinv}((u_{g,i}, u_{\mu,i}), (v_{g,i}, v_{\mu,i}),\ N)$$

  end

$$(\sigma_g, \sigma_\mu) = h \sum_{i=1}^{s} b_i(k_{g,i}, k_{\mu,i})$$

$$(g_{n+1}, \mu_{n+1}) = \lambda((\sigma_g, \sigma_\mu), (g_n, \mu_n)).$$

end

2.8 A specific Lie group example

In a numerical example later we will consider the motion of soft dipolar spheres whose phase space is the cotangent bundle described in this section. First, recall from the previous section the Lie group $\mathcal{G} = G \ltimes \mathfrak{g}^*$ and its Lie algebra $\tilde{\mathfrak{g}} = \mathfrak{g} \ltimes \mathfrak{g}^*$. We will now specify the Lie group $\mathcal{G}$. Let $G = \mathbb{R}^3 \times SO(3)$ and $\mathfrak{g}^* = \mathbb{R}^3 \times \mathfrak{so}(3)^*$. Then

$$\mathcal{G} = (\mathbb{R}^3 \times SO(3)) \ltimes (\mathbb{R}^3 \times \mathfrak{so}(3)^*),$$
$$\tilde{\mathfrak{g}} = (\mathbb{R}^3 \times \mathfrak{so}(3)) \ltimes (\mathbb{R}^3 \times \mathfrak{so}(3)^*).$$

We assume that an ODE is evolving on $\mathcal{G}$, and that $\mathcal{G}$ acts on itself. The setting is the same as what was just discussed in section 2.7, except that we now have a specific Lie group.

Suppose the ODE is on the form

$$(\dot{q}, \dot{Q}, \dot{p}, \dot{m}) = F(q, Q, p, m),$$

where $F : \mathcal{G} \mapsto \tilde{\mathfrak{g}}$, and

$$\begin{aligned}((q, Q), (p, m)) &\in \mathcal{G},\\ ((\dot{q}, \dot{Q}), (\dot{p}, \dot{m})) &\in \tilde{\mathfrak{g}}.\end{aligned}$$

Taking the same approach as before, we rewrite the ODE on the generic form (2.23)

$$(\dot{q}, \dot{Q}, \dot{p}, \dot{m}) = (\lambda_* f(q, Q, p, m))(q, Q, p, m), \tag{2.42}$$

where $f : \mathcal{G} \mapsto \tilde{\mathfrak{g}}$ is a smooth function. Assume f has four components, such that $f = (f_q, f_Q, f_p, f_m)$, where

$$\begin{aligned}f_q &: \mathcal{G} \mapsto \mathbb{R}^3,\\ f_Q &: \mathcal{G} \mapsto \mathfrak{so}(3),\\ f_p &: \mathcal{G} \mapsto \mathbb{R}^3,\\ f_m &: \mathcal{G} \mapsto \mathfrak{so}(3)^*.\end{aligned}$$

Then, the explicit representation of the ODE on the generic form becomes

$$\begin{aligned}(\dot{q}, \dot{Q}, \dot{p}, \dot{m}) &= ((f_q, f_Q)(q, Q), (f_p, f_m) - \mathrm{ad}^*_{(f_q, f_Q)}(p, m))\\ &= (f_q, f_Q Q, (f_p, f_m) - (0, \mathrm{ad}^*_{f_Q} m))\\ &= (f_q, f_Q Q, f_p, f_m - \mathrm{ad}^*_{f_Q} m),\end{aligned} \tag{2.43}$$

using (2.38), that is, the right tangent map. Similarly using (2.39), that is, the left tangent map we get

$$\begin{aligned}(\dot{q}, \dot{Q}, \dot{p}, \dot{m}) &= ((q, Q)(f_q, f_Q), (f_p, f_m) + \mathrm{ad}^*_{(f_q, f_Q)}(p, m))\\ &= (f_q, Q f_Q, (f_p, f_m) + (0, \mathrm{ad}^*_{f_Q} m))\\ &= (f_q, Q f_Q, f_p, f_m + \mathrm{ad}^*_{f_Q} m),\end{aligned} \tag{2.44}$$

where as before, we have omitted the dependent variables.

In the computation of (2.43) and (2.44) we have used that

$$\begin{aligned}
\mathrm{ad}_{(f_q,f_Q)}(q,Q) &= ((f_q,f_Q)(q,Q)-(q,Q)(f_q,f_Q)) \\
&= (f_q+q, f_Q Q) - (q+f_q, Q f_Q) \\
&= (f_q+q-q-f_q, f_Q Q - Q f_Q) \\
&= (0,[f_Q,Q]) \\
&= (0,\mathrm{ad}_{f_Q} Q),
\end{aligned}$$

and, thus

$$\begin{aligned}
\left\langle \mathrm{ad}^*_{(f_q,f_Q)}(p,m),(q,Q)\right\rangle &= \left\langle (p,m),\mathrm{ad}_{(f_q,f_Q)}(q,Q)\right\rangle \\
&= \left\langle (p,m),(0,\mathrm{ad}_{f_Q} Q)\right\rangle \\
&= \left\langle (0,\mathrm{ad}^*_{f_Q} m),(q,Q)\right\rangle .
\end{aligned}$$

The flow of (2.42) on $\mathcal{G}$ can now be stated, that is, for small values of t using (2.40) and right multiplication the solution becomes

$$\begin{aligned}
(q(t),Q(t),p(t),m(t)) &= \lambda((\sigma_q,\sigma_Q,\sigma_p,\sigma_m),(q,Q,p,m)) \\
&= (\exp(\sigma_q,\sigma_Q)(q,Q), \mathrm{dexp}^*_{(\sigma_q,\sigma_Q)}(\sigma_p,\sigma_m) + \mathrm{Ad}^*_{\exp(-\sigma_q,-\sigma_Q)}(p,m)) \\
&= (\sigma_q+q, \exp(\sigma_Q)Q, \sigma_p+p, \mathrm{dexp}^*_{\sigma_Q}(\sigma_m)+\mathrm{Ad}^*_{\exp(-\sigma_Q)}m).
\end{aligned}$$

Similarly, using (2.41) and left multiplication, the flow on $\mathcal{G}$ becomes

$$(q(t),Q(t),p(t),m(t)) = (q+\sigma_q, Q\ \exp(\sigma_Q), p+\sigma_p, \mathrm{Ad}^*_{\exp(\sigma_Q)}m + \mathrm{dexp}^*_{\sigma_Q}\sigma_m).$$

In the previous computations, we have used that

$$\begin{aligned}
\mathrm{Ad}_{(\sigma_q,\exp(\sigma_Q))}(q,Q)) &= (\sigma_q,\exp(\sigma_Q))\cdot(q,Q)\cdot(-\sigma_q,\exp(-\sigma_Q)) \\
&= (\sigma_q+q-\sigma_q, \exp(\sigma_Q)Q\exp(-\sigma_Q)) \\
&= (q,\mathrm{Ad}_{\exp(\sigma_Q)}(Q)),
\end{aligned}$$

hence

$$\begin{aligned}
\left\langle \mathrm{Ad}^*_{(\sigma_q,\exp(\sigma_Q))})(p,m),(q,Q)\right\rangle &= \left\langle (p,m),\mathrm{Ad}_{(\sigma_q,\exp(\sigma_Q))})(q,Q)\right\rangle \\
&= \left\langle (p,m),(q,\mathrm{Ad}_{\exp(\sigma_Q)}(Q))\right\rangle \\
&= \left\langle (p,\mathrm{Ad}_{\exp(\sigma_Q)}(m)),(q,Q)\right\rangle .
\end{aligned}$$

The ad operator becomes

$$\mathrm{ad}_{(\sigma_q,\sigma_Q,\sigma_p,\sigma_m)}(f_q,f_Q,f_p,f_m) = (0,\mathrm{ad}_{\sigma_Q} f_Q, 0, \mathrm{ad}^*_{f_Q}\sigma_m - \mathrm{ad}^*_{\sigma_Q}\sigma_m).$$

Using this, we can determine $(\sigma_q(t), \sigma_Q(t), \sigma_p(t), \sigma_m(t))$ by solving

$$\begin{aligned}(\dot\sigma_q, \dot\sigma_Q, \dot\sigma_p, \dot\sigma_m) &= \mathrm{dexp}^{-1}_{(\sigma_q,\sigma_Q,\sigma_p,\sigma_m)}(f_q, f_Q, f_p, f_m) \\ &\approx \sum_{k=0}^{N-1} \frac{B_k}{k!} \mathrm{ad}^k_{(\sigma_q,\sigma_Q,\sigma_p,\sigma_m)}(f_q, f_Q, f_p, f_m), \\ (\sigma_q(0), \sigma_Q(0), \sigma_p(0), \sigma_m(0)) &= (0,0,0,0) \in \tilde{\mathfrak{g}},\end{aligned}$$

using a classical Runge-Kutta method of order N.

Algorithm 3:

The RKMK algorithm on $\mathcal{G} = SO(3) \ltimes \mathfrak{so}(3)^*$.

Initialize (q_0, Q_0, p_0, m_0).

for $n = 1, 2, \ldots$

  for $i = 1, 2, \ldots, s$

$$(u_{q,i}, u_{Q,i}, u_{p,i}, u_{m,i}) = h \sum_{j=1}^{i-1} a_{i,j}(k_{q,j}, k_{Q,j}, k_{q,j}, k_{Q,j})$$

$$\begin{aligned}&(v_{q,i}, v_{Q,i}, v_{p,i}, v_{m,i}) \\ &= f(q_n + u_{q,i}, Q\mathrm{exp}(u_{Q,i}), p_n + u_{p,i}, \mathrm{Ad}^*_{\exp(u_{Q,i})} m_n + \mathrm{dexp}^*_{u_{Q,i}}(u_{m,i}))\end{aligned}$$

$$\begin{aligned}&(k_{q,j}, k_{Q,j}, k_{p,j}, k_{m,j}) \\ &= \left(v_{q,i}, \textstyle\sum_{j=0}^{N-1} \mathrm{ad}^j_{u_{Q,i}} v_{Q,i}, v_{p,i}, \textstyle\sum_{j=0}^{N-1} \left(\mathrm{ad}^*_{v_{Q,i}} u_{m,i} - \mathrm{ad}^*_{u_{Q,i}} v_{m,i} \right)^j \right)\end{aligned}$$

  end

$$(\sigma_q, \sigma_Q, \sigma_p, \sigma_m) = h \sum_{i=1}^{s} b_i(k_{q,j}, k_{Q,j}, k_{p,j}, k_{m,j})$$

$$\begin{aligned}&(q_{n+1}, Q_{n+1}, p_{n+1}, m_{n+1}) \\ &= (q_n + \sigma_q, Q\mathrm{exp}(\sigma_Q), p_n + \sigma_p, \mathrm{Ad}^*_{\exp(\sigma_Q)} m_n + \mathrm{dexp}^*_{\sigma_Q} \sigma_m).\end{aligned}$$

end

Chapter 3

Numerical experiments

3.1 Rigid bodies

Arnold defined in [2] that a rigid body is a system of point masses where the distance between any two points of the body always is the same. The body's configuration space is $\mathbb{R}^3 \times SO(3)$, a six-dimensional Lie-group consisting of both translations and rotations, and the phase space is the cotangent bundle $\mathbb{R}^3 \times \mathfrak{so}(3)^*$.

Consider a rigid body rotating around its center of mass. Its movement will be uniform and linear, rotating like its center of mass were fixed at a stationary point. Assume that the rigid body is "free", meaning no external forces are acting on it. In this case, the configuration space reduces to the three-dimensional Lie group $SO(3)$, and the phase space is the cotangent bundle $\mathfrak{so}(3)^*$ [8, 18].

Points on a rigid body can be given in coordinates relative to the body or relative to an axis system fixed in space, where the different coordinates come from left and right trivialisation of the cotangent bundle [8, 12].

Let $Q \in SO(3)$ be the configuration of the body, and denote by $m_b \in \mathfrak{so}(3)^*$ and $m_r \in \mathfrak{so}(3)^*$ the body's angular momentum relative to the body and to the reference frame, respectively. Note that m_b is obtained using either left trivialisation or m_r using right trivialisaton [8, 12].

Euler published in 1765 the equations of motion of a rigid body, where the two important theorems considering the rigid body's motion were stated [3]. The first theorem states that the angular momentum relative to the reference frame is preserved under motion, that is

$$\dot{m}_r = 0.$$

The second theorem states that the angular momentum relative to the body satisfies

$$\dot{m}_b = \mathrm{ad}^*_{\omega_b} m_b,$$

where $\omega_b = I^{-1} m_b \in \mathfrak{so}(3)$ is the angular velocity relative to the body, and $I = [I_1, I_2, I_3]^T$ are the principal moments of inertia [18]. The latter theorem is often refered to as the Euler equations [2, 3, 8]. The left and right trivialised equations of motions may be derived using the trivialised Hamiltonian ODEs (2.31) and (2.32). We will only focus on the left trivialised case, but the right trivialised case follows similar steps [8, 12].

To ease the notation, we from now on omit the subscript b in m_b. We also write the angular momentum in the body frame on component form. Then the Euler equations can be expressed as

$$\begin{aligned} \dot{m}_1 &= \left(\frac{1}{I_3} - \frac{1}{I_2}\right) m_2 m_3, \\ \dot{m}_2 &= \left(\frac{1}{I_1} - \frac{1}{I_3}\right) m_1 m_3, \\ \dot{m}_3 &= \left(\frac{1}{I_2} - \frac{1}{I_1}\right) m_1 m_2. \end{aligned} \tag{3.1}$$

The physical motion of (3.1) is equal to rotating the coordinate axis in $\mathbb{R}^3$. Then, both the angular momentum and the total energy will be preserved [2, 18].

The angular momentum and total energy functions

$$C(m) = \frac{1}{2}\left(m_1^2 + m_2^2 + m_3^2\right), \tag{3.2}$$

and

$$H(m) = \frac{1}{2}\left(\frac{m_1^2}{I_1} + \frac{m_2^2}{I_2} + \frac{m_3^2}{I_3}\right), \tag{3.3}$$

are called the Casimir and the Hamiltonian functions, respectively. Differentiating (3.2) and (3.3), we get

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}C(m(t)) &= \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\left(m_1(t)^2 + m_2(t)^2 + m_3(t)^2\right) \\ &= m_1(t)\dot{m}_1(t) + m_2(t)\dot{m}_2(t) + m_3(t)\dot{m}_3(t) \\ &= 0,\end{aligned}$$

and

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}H(m(t)) &= \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{m_1(t)^2}{I_1} + \frac{m_2(t)^2}{I_2} + \frac{m_3(t)^2}{I_3}\right) \\ &= \frac{m_1(t)\dot{m}_1(t)}{I_1} + \frac{m_2(t)\dot{m}_2(t)}{I_2} + \frac{m_3(t)\dot{m}_3(t)}{I_3} \\ &= 0.\end{aligned}$$

Hence, the Casimir and the Hamiltonian functions are conserved along the solutions. This leads to solutions that are closed curves lying on the intersection of a sphere with an ellipsoid [13].

Observe that the functional derivatives of the Hamiltonian are

$$\begin{aligned}\frac{\delta H}{\delta Q} &= 0, \\ \frac{\delta H}{\delta m} &= I^{-1}m,\end{aligned}$$

which we insert into (2.31). That is, we get the ODE system

$$X_H^L(Q,m) = \left(L_{Q*}I^{-1}m, 0 + \mathrm{ad}^*_{I^{-1}m}m\right).$$

We get the same ODE by inserting $f_Q = I^{-1}m = \omega$ and $f_m = 0$ in the left tangent map (on the generic form) (2.39), that is

$$(\dot{Q}, \dot{m}) = (L_{Q*}f_Q, f_m + \mathrm{ad}^*_{f_Q}m).$$

Then, for $G = SO(3)$ the equations of motion are [12, 24]

$$\begin{aligned}(\dot{Q}, \dot{m}) &= (QI^{-1}m, \mathrm{ad}^*_{I^{-1}m}m) \\ &= (QI^{-1}m, -\widehat{I^{-1}m}m) \\ &= (Q\omega, \widehat{m}\omega).\end{aligned}$$

We see that the equations are decoupled, hence we may solve the them separately of each others. Also, we see that the ODE system is on the generic

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

Table 3.1: The Butcher tableau for RK4.

form (2.23), thus, we can solve it by applying the RKMK methods.

We use a Runge-Kutta method of order 4, with coefficients given in the Butcher tableau 3.1, that is Applying normalized Gaussian initial values for the angular momentum, we integrate the system from $t = 0$ to $t = 100$, using the step size $h = 0.01$. Figure 3.1 shows ten solution curves with different initial conditions on a sphere. Figures 3.3 and 3.4 show the error of the Hamiltonian and Casimir functions, respectively, when integrating the system from $t = 0$ to $t = 1000$, using the step size $h = 10^{-4}$. The results look promising, in figure 3.1 we see that the solution curves look closed, and in figures 3.3 and 3.4 we see that the errors are almost on machine precision level. Figure 3.2 shows the order of convergence for a fourth order Runge-Kutta method. We see that the convergence order is as expected, which suggests that the implementation is correct.

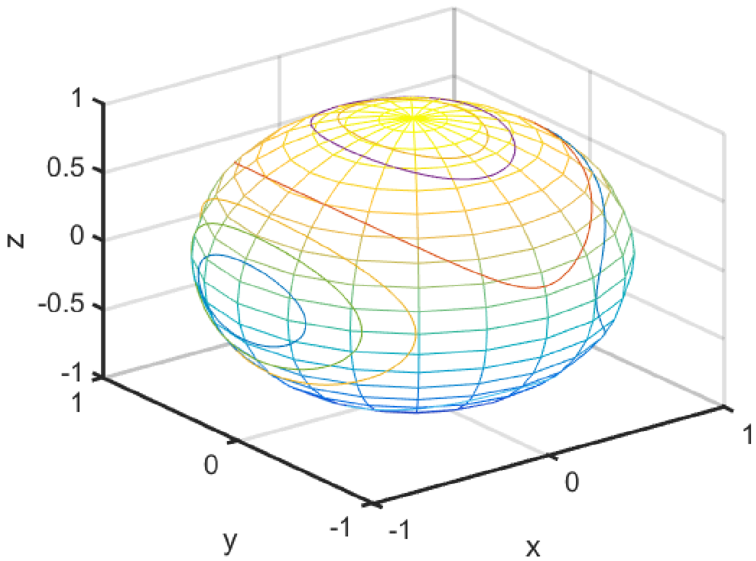

Figure 3.1: Ten solution curves on the sphere, using normalized random Gaussian initial values, integrated from $t = 0$ to $t = 100$, taking 10000 steps, step length $h = 0.01$.

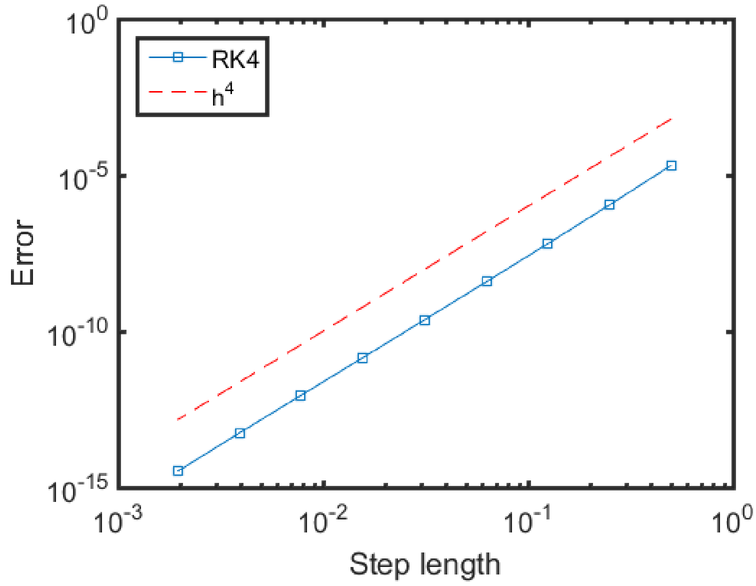

Figure 3.2: Order plot using a fourth order Runge-Kutta method. Step length $h = 2^i, i = 1, \ldots, 10$, $t = 0$ to $t = 1$. The dashed lines are h^4.

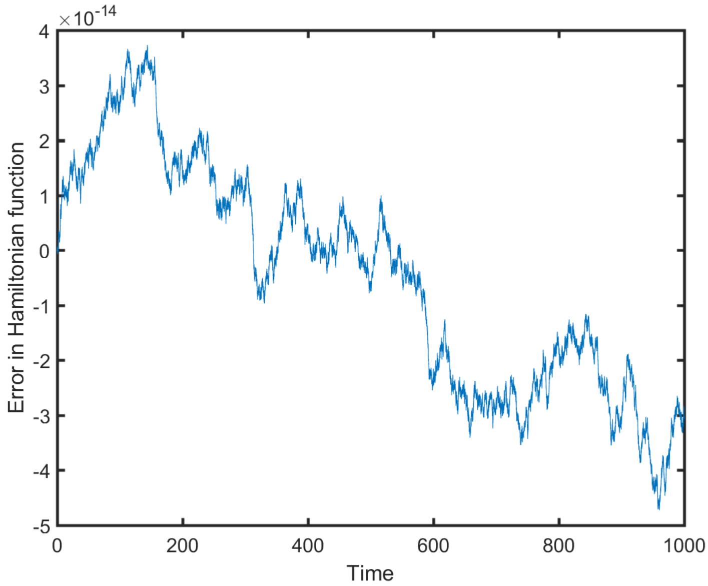

Figure 3.3: Error in the Hamiltonian function, $H(m_k) - H(m_1)$, from $t = 0$ to $t = 1000$, taking 1000000 steps, i.e. step length $h = 0.001$.

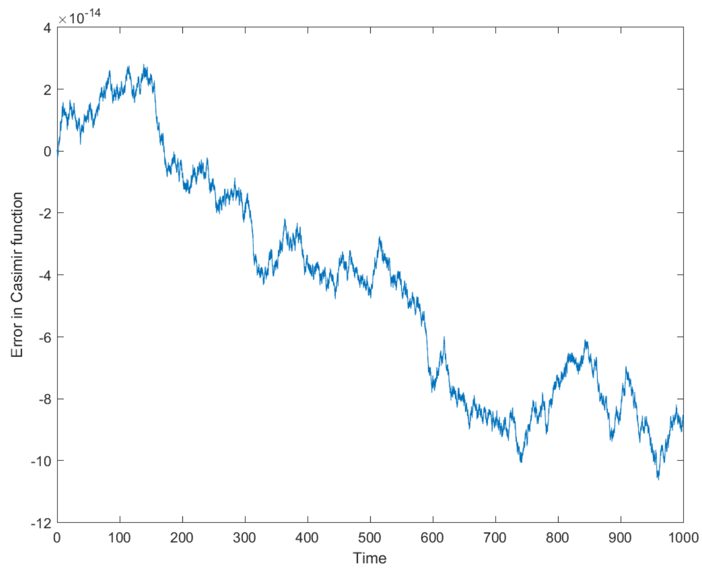

Figure 3.4: Error in the Casimir function, $C(m_k) - C(m_1)$, from $t = 0$ to $t = 1000$, taking 1000000 steps, i.e. step length $h = 0.001$.

3.2 Molecular dynamics - dipolar soft spheres

Consider a large system of identical coupled dipolar soft spheres, where every sphere is a rigid body with configuration space $G = \mathbb{R}^3 \times SO(3)$ and phase space T^*G. We trivialise the cotangent bundle such that $T^*G \approx G \ltimes \mathfrak{g}^*$, leading to the total phase space $\mathcal{G} = (G \ltimes \mathfrak{g}^*)^N$, where N denotes the number of particles. That is, the Lie group under consideration is

$$\mathcal{G} = ((\mathbb{R}^3 \times SO(3)) \ltimes ((\mathbb{R}^3 \times \mathfrak{so}(3)^*))^N$$

with the corresponding Lie algebra

$$\tilde{\mathfrak{g}} = ((\mathbb{R}^3 \times \mathfrak{so}(3)) \ltimes (\mathbb{R}^3 \times \mathfrak{so}(3)^*))^N.$$

Assume the dipolar soft spheres are water molecules, and let M_i denote the total mass of the ith particle, $\mathbf{q}_i$ the position of its center of mass, $\mathbf{p}_i$ its linear momentum, $\mathbf{Q}_i$ its orientation, and $\mathbf{m}_i$ its angular momentum in the body frame [9, 10].

The Hamiltonian function for the system is the total energy, and is given by

$$H(\mathbf{q}, \mathbf{Q}, \mathbf{p}, \mathbf{m}) = T(\mathbf{p}, \mathbf{m}) + V(\mathbf{q}, \mathbf{Q}),$$

where T and V are the total kinetic and potential energies, respectively.

The kinetic energy is the sum of $T_i^{trans}(\mathbf{m}_i)$ and $T_i^{rot}(\mathbf{p}_i)$, that is, the translational and the rotational kinetic energy of each body, respectively. These are given as

$$\begin{aligned} T_i^{trans}(\mathbf{p}_i) &= \frac{\|\mathbf{p}_i\|^2}{2M_i}, \\ T_i^{rot}(\mathbf{m}_i) &= \frac{1}{2}\mathbf{m}_i \cdot (I^{-1}\mathbf{m}_i), \end{aligned}$$

where $I = \mathrm{diag}(I_i)$, $i = 1, 2, 3$, the inertia matrix.

The potential energy is the sum of short range and the dipole interactions, $V_{i,j}^{short}$ and $V_{i,j}^{dip}$, respectively, where $V_{i,j}$ denotes the interaction between molecule i and j. We have

$$\begin{aligned} V(\mathbf{q}, \mathbf{Q}) &= \sum_{j>i} V_{i,j}(\mathbf{q}_i, \mathbf{Q}_i, \mathbf{q}_j, \mathbf{Q}_j) \\ &= \sum_{j>i} (V_{i,j}^{short} + V_{i,j}^{dip}), \end{aligned}$$

where

$$
\begin{aligned}
V_{i,j}^{short} &= 4\epsilon\left(\frac{\sigma}{r_{i,j}}\right)^{12}, \\
V_{i,j}^{dip} &= \frac{1}{r_{i,j}^3}\boldsymbol{\mu}_i\cdot\boldsymbol{\mu}_j - \frac{3}{r_{i,j}^5}(\boldsymbol{\mu}_i\cdot\mathbf{r}_{i,j})(\boldsymbol{\mu}_j\cdot\mathbf{r}_{i,j}),
\end{aligned}
$$

for $\mathbf{r}_{i,j} = \mathbf{q}_i - \mathbf{q}_j$, $r_{i,j} = \|\mathbf{r}_{i,j}\|$ and $\boldsymbol{\mu}_i$ the orientation vector for the ith molecule. Let $\bar{\boldsymbol{\mu}}$ be an initial fixed reference orientation for the molecule, such that $\boldsymbol{\mu}_i = \mathbf{Q}_i\bar{\boldsymbol{\mu}}$.

The Hamiltonian function can now be written as

$$
\begin{aligned}
H(\mathbf{q},\mathbf{p},\mathbf{m},\mathbf{Q}) = &\sum_i\left(\frac{\|\mathbf{p}_i\|^2}{2M_i} + \frac{1}{2}\mathbf{m}_i\cdot(I_i^{-1}\mathbf{m}_i)\right) \\
&+\sum_{j>i}\left(4\epsilon\left(\frac{\sigma}{r_{i,j}}\right)^{12} + \frac{1}{r_{i,j}^3}\boldsymbol{\mu}_i\cdot\boldsymbol{\mu}_j - \frac{3}{r_{i,j}^5}(\boldsymbol{\mu}_i\cdot\mathbf{r}_{i,j})(\boldsymbol{\mu}_j\cdot\mathbf{r}_{i,j})\right).
\end{aligned}
$$

This leads the the Hamiltonian ODE system

$$
\begin{aligned}
\dot{\mathbf{q}}_i &= \frac{\mathbf{p}_i}{M_i}, \\
\dot{\mathbf{Q}}_i &= \mathbf{Q}_i\widehat{(I_i^{-1}\mathbf{m}_i)} \\
\dot{\mathbf{p}}_i &= -\frac{\partial V}{\partial \mathbf{q}_i}, \\
\dot{\mathbf{m}}_i &= \mathbf{m}_i\times(I_j^{-1}\mathbf{m}_i) - \mathrm{rot}\left(\mathbf{Q}_i^T\frac{\partial V}{\partial \mathbf{Q}_i}\right),
\end{aligned}
\tag{3.4}
$$

which constitutes the equations of motion that we intend to solve. The position, linear and angular momenta, $\mathbf{q}_i$, $\mathbf{p}_i$, and $\mathbf{m}_i$, are all represented by vectors in $\mathbb{R}^3$, while the orientation $\mathbf{Q}_i$ is a (3×3) orthogonal matrix. The function rot maps a (3×3) matrix A to its skew-symmetric part $(A - A^T)$ before associating the result to a vector in $\mathbb{R}^3$. That is, if $A = \{a_{i,j}\}$, then [6, 9, 10]

$$
\mathrm{rot}(A) = \begin{bmatrix} a_{3,2}-a_{2,3} \\ a_{1,3}-a_{3,1} \\ a_{2,1}-a_{1,2} \end{bmatrix}.
$$

We see that (3.4) fits on the left trivialised generic format (2.44) if we define for any particle i

$$
\begin{aligned}
f_{q,i} &= \frac{\mathbf{p}_i}{M_i}, \\
f_{Q,i} &= I^{-1}\mathbf{m}_i, \\
f_{p,i} &= -\frac{\partial V}{\partial \mathbf{q}_i}(\mathbf{q}_i, \mathbf{Q}_i), \\
f_{m,i} &= -\mathrm{rot}(\mathbf{Q}_i^T \frac{\partial V}{\partial \mathbf{Q}_i}(\mathbf{q}_i, \mathbf{Q}_i)), \\
\mathrm{ad}^*_{f_{Q,i}}\mathbf{m}_i &= \mathbf{m}_i \times I^{-1}\mathbf{m}_i.
\end{aligned}
$$

Hence, we can apply the RKMK methods to this particular system.

3.2.1 Implementation

We follow the steps in the RKMK algorithm presented in section 2.8, and use

$$
\begin{aligned}
\frac{\partial V}{\partial \mathbf{q}_i} &= \sum_{j \neq i} \left[\left(-\frac{48}{r_{i,j}^{14}} - \frac{3\mu_i^T\mu_i}{r_{i,j}^5} + \frac{15\mu_i^T\mathbf{r}_{ij}\mu_j^T\mathbf{r}_{ij}}{r_{i,j}^7} \right)\mathbf{r}_{ij} - \left(\frac{3\mu_j^T\mathbf{r}_{ij}}{r_{ij}^5} \right)\mu_i - \left(\frac{3\mu_i^T\mathbf{r}_{ij}}{r_{ij}^5} \right)\mu_j \right] \\
\frac{\partial V}{\partial \mathbf{Q}_i} &= \sum_{j \neq i} \left[\frac{1}{r_{i,j}^3}\mu_j\bar{\mu}_i^T - \left(\frac{3\mu_j^T\mathbf{r}_{ij}}{r_{i,j}^5} \right)\mathbf{r}_{ij}\bar{\mu}_i^T \right].
\end{aligned}
$$

For the computation of the exponential of real skew-symmetric matrices, we apply Rodrigues' formula which is an inexpensive and explicit formula [7, 9]:

$$
\exp(u) = I + \frac{\sin(\alpha)}{\alpha}u + \frac{1-\cos(\alpha)}{\alpha^2}u^2, \ \ \alpha^2 = u^T u, u \in \mathbb{R}^3.
$$

For more information, see the proof in [18]. A similar expression exists for evaluating dexp_u [9, 16]

$$
\mathrm{dexp}_u = I + \frac{\sin(\alpha/2)}{(\alpha^2/2)}\hat{u} + \frac{\alpha - \sin(\alpha)}{\alpha^3}\hat{u}^2, \ \ \alpha^2 = u^T u, u \in \mathbb{R}^3.
$$

3.2.2 Results

First, we simulate a system of $N = 100$ particles, using the same initial values as in [6]. That is, we choose $M_i = 1$, $\sigma = \epsilon = 1$ and the moments of inertia

$(I_1, I_2, I_3) = (1, 1.88, 2.88)$, which corresponds to those of water molecules. The positions $\mathbf{q}_i(0)$ are chosen as vectors with Gaussian distributed random components between $-N$ and N. For the linear and angular momenta we set $\mathbf{p}_i(0) = \mathbf{m}_i(0) = 0$. Finally, the orientation matrices $\mathbf{Q}_i(0)$ are chosen as random orthogonal matrices and the fixed reference orientation for the dipoles is $\bar{\mu}_i = [0, 1, 1]^T$.

Figure 3.5 shows the order of convergence for a third and a fourth order Runge-Kutta method. We see that the convergence order is as expected, which suggests that the implementation is correct.

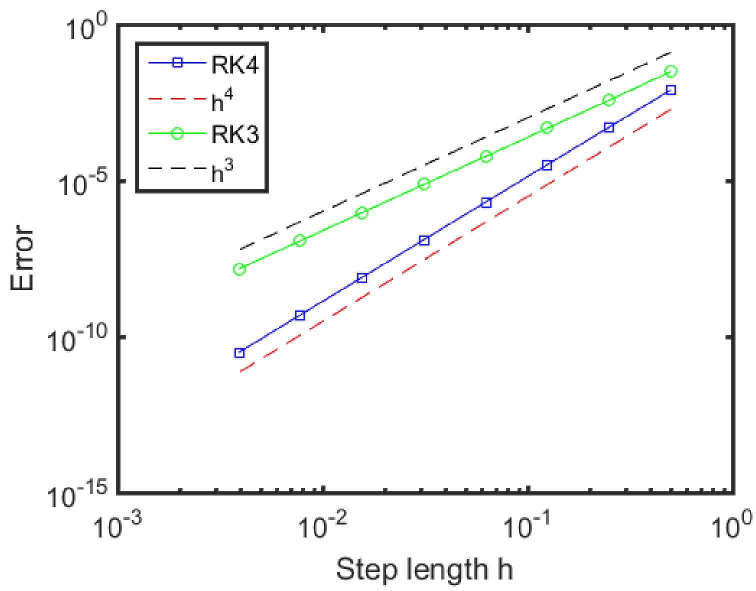

Figure 3.5: Order of convergence for a third and a fourth order Runge-Kutta method. Dashed lines show h^3 and h^4. Integrated from $t = 0$ to $t = 1$ with varying step size h.

Figure 3.6 shows the orthogonality error $\|Q_i^T Q_i - I\|$ for each of the 100 particles in every time step, integrated from $t = 0$ to $t = 100$, using a step size $h = 0.1$. We observe that the orthogonality errors have minor fluctuations, but still are on the order 10^{-14}. We consider this a good indication of preservation of orthogonality, which is to be expected from Lie group methods on $SO(3)$.

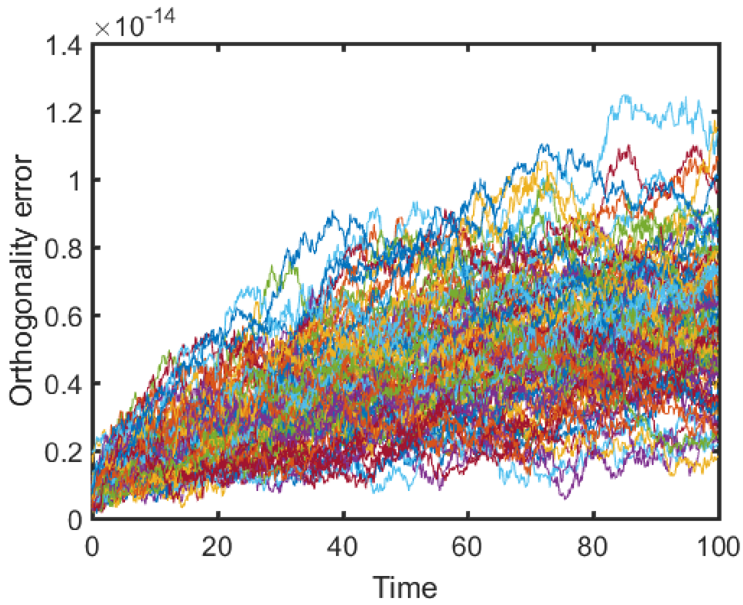

Figure 3.6: Orthogonality for each of the 100 particles, using 1000 steps between $t = 0$ and $t = 100$.

Figures 3.7 – 3.10 show the the relative errors in the Hamiltonian function after the nth time step, that is

$$\frac{H(\mathbf{q}^n, \mathbf{Q}^n, \mathbf{p}^n, \mathbf{m}^n) - H(\mathbf{q}^0, \mathbf{Q}^0, \mathbf{p}^0, \mathbf{m}^0)}{H(\mathbf{q}^0, \mathbf{Q}^0, \mathbf{p}^0, \mathbf{m}^0)}.$$

The final time T=1 in both figures 3.7 and 3.8, but the step length is different. The difference in shape of the errors may suggest that the second error is on machine precision level.

Figures 3.9 and 3.10 have the same step length but different final times $T = 100$ and $T = 1000$, respectively. Due to the size of the systems, the number of particles is chosen differently in these two simulations; using 100 particles with a final time $t = 1000$ requires a simulation time of more than a week. Figure 3.10 shows the relative Hamiltonian error using $N = 100$ particles, and figure 3.9 using $N = 10$ particles, respectively.

We observe that there are several error spikes, and after closer inspection of the simulations, it seems as though these spikes occur when two particles come close enough to each other. We suspect that this causes the short range interaction term to become very large, resulting in a floating point error which causes the spikes. If these suspicions are correct, increasing the floating point precision to e.g. 128 bits may fix these spikes. In addition, since the simulation in figure 3.9 is a system ten times as large as in the simulation in figure 3.10, it makes sense that the spikes will occur more often.

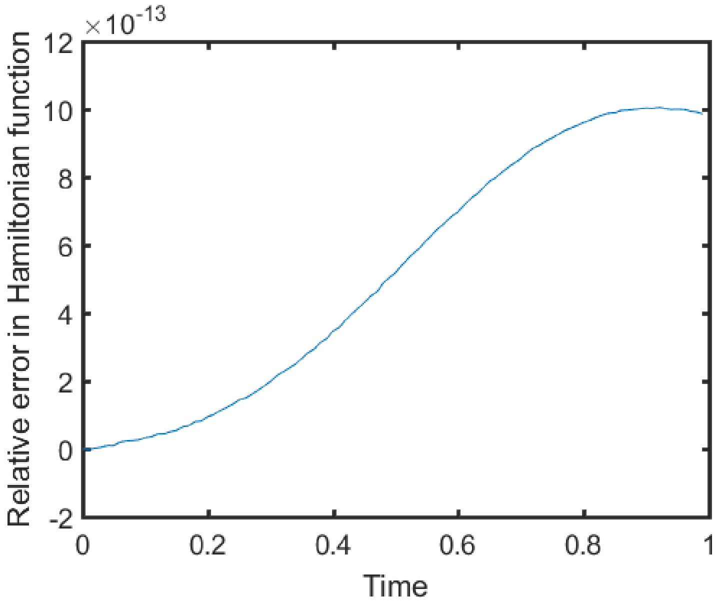

Figure 3.7: Relative error in the Hamiltonian function of a system of 100 particles, using 100 steps between $t = 0$ and $t = 1$, i.e. step length $h = 0.01$.

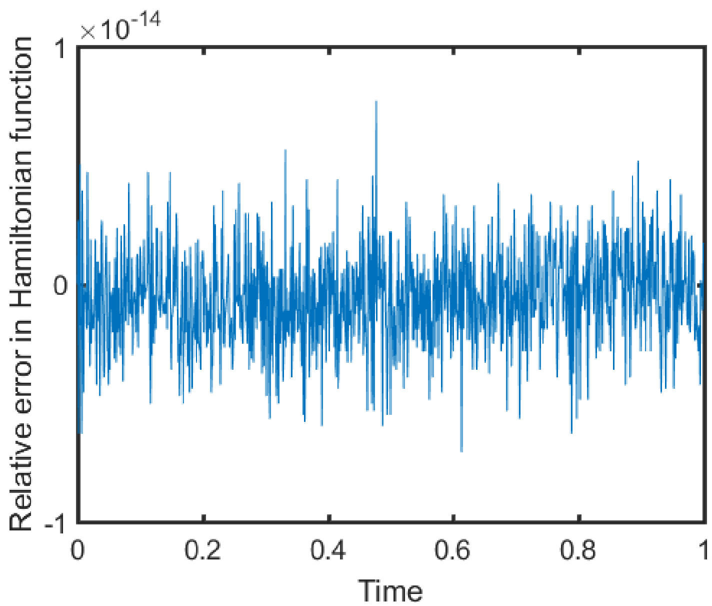

Figure 3.8: Relative error in the Hamiltonian function of a system of 100 particles, using 1000 steps between $t = 0$ and $t = 1$, i.e. step length $h = 0.001$.

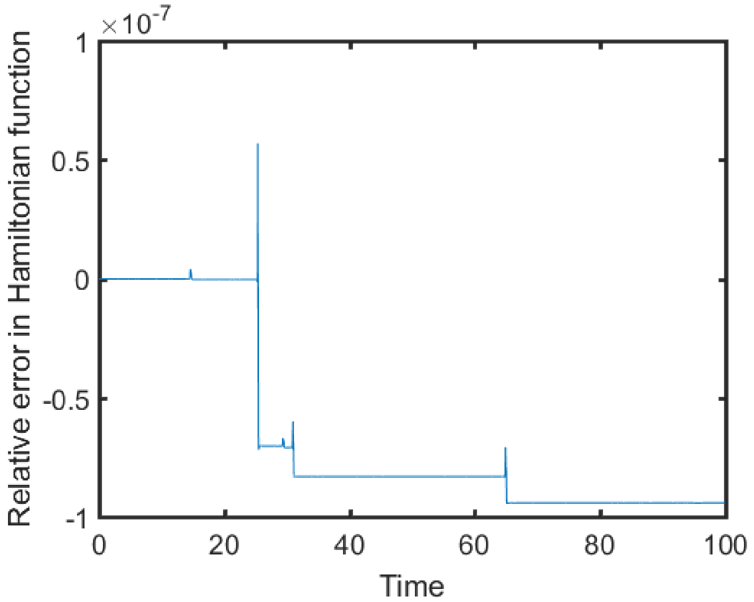

Figure 3.9: Relative error in the Hamiltonian function of a system of 100 particles, using 10000 steps between $t = 0$ and $t = 100$, i.e. step length $h = 0.01$.

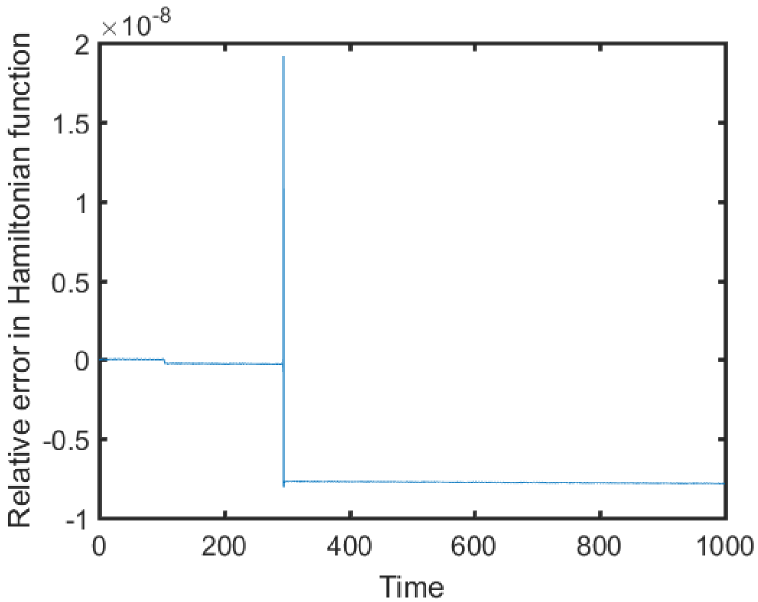

Figure 3.10: Relative error in the Hamiltonian function of a system of 10 particles, using 100000 steps between $t = 0$ and $t = 1000$, i.e. step length $h = 0.01$.

It is hard to visualise the solutions in $\mathbb{R}^3$ without animations, so instead we construct a 2D example, the results of which are shown in figures 3.11 – 3.14. This requires changes in the initialising process, that is, we set every 3rd component in the position vector of each particle to zero. The orientation vector $\bar{\mu}$ is set to $[1, 1, 0]^T$. The orientation matrices $\mathbf{Q}_i$ are as before random orthogonal matrices, but restricted to rotations of the first and second components.

Figure 3.11 shows the solution curves of two particles interacting between $t = 0$ and $t = 100$, using 1000 time steps. The circles mark every 100 time steps, starting at $t = 0$ where the black dots marking the starting positions. We observe that the particles seem to attract each other. However, after approximately 450 time steps, the distance between the particles does not change and stays approximately the same for the rest of the time. This is reminiscent of dipole bonds. Figure 3.12 shows the orientation of the two particles. As long as the particles' orientation differ by more than 90 degrees, they repel each other, otherwise they attract each other.

Figures 3.13 and 3.14 show the motion and orientation of ten particles in 2D between $t = 0$ and $t = 50$. We can see that some particles stay close to each other and form bond-like structures during the entire simulation, while others are pushed out of the system. This situation may not be physically correct since it is reasonable to assume that there would be some constraints on the positions in a real life situation. Nevertheless, the plots do look credible from a physical standpoint.

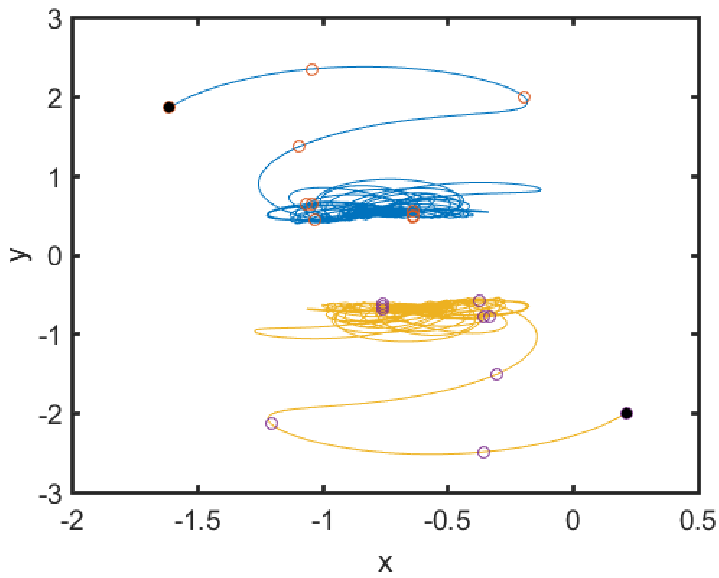

Figure 3.11: Two particles in 2D and their positions when interacting, using 1000 steps between $t = 0$ and $t = 100$.

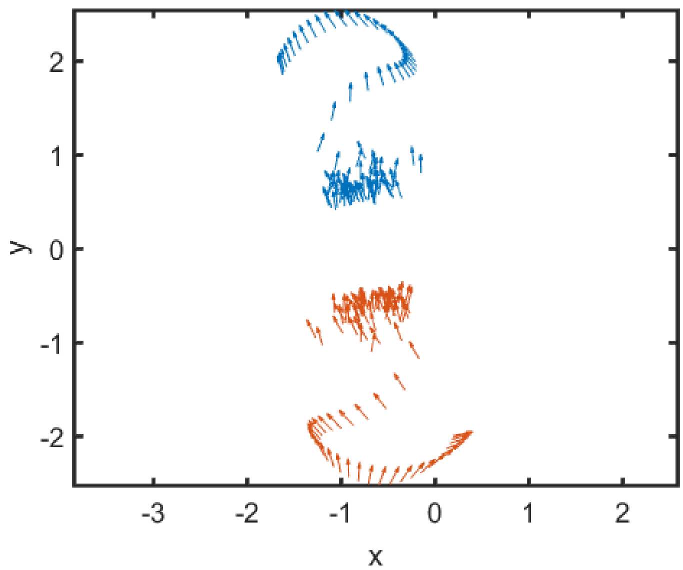

Figure 3.12: Two particles in 2D and their directions when interacting, using 1000 steps between $t = 0$ and $t = 100$.

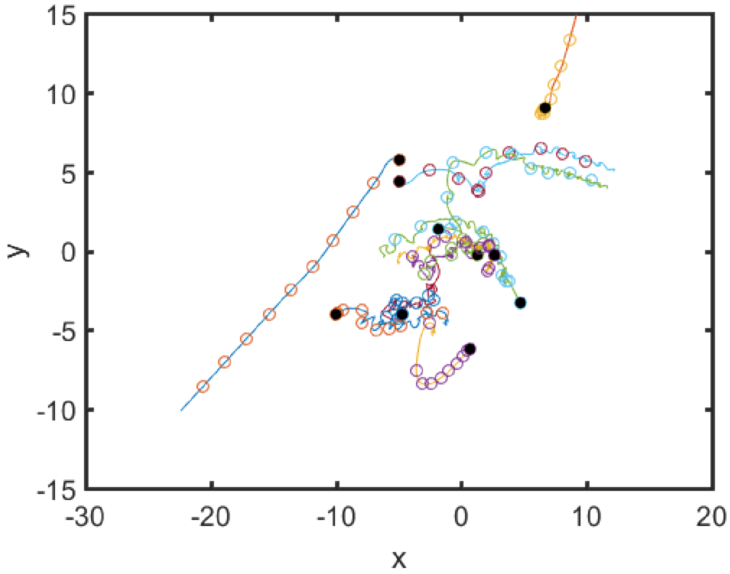

Figure 3.13: Ten particles in 2D and their positions when interacting, using 1000 steps between $t = 0$ and $t = 50$.

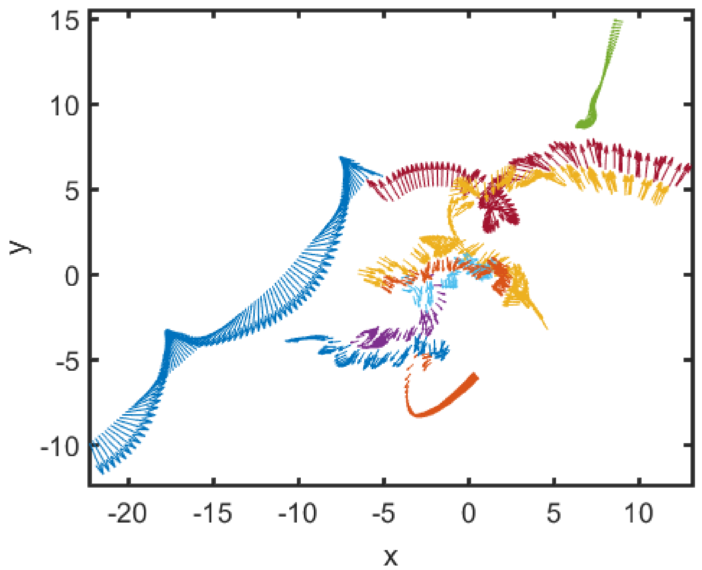

Figure 3.14: Ten particles in 2D and their directions when interacting, using 1000 steps between $t = 0$ and $t = 50$.

Chapter 4

Conclusion and future work

We have given a theoretical background on manifolds and Lie groups and used this to present the Runge-Kutta-Munthe-Kaas methods. The Euler-Lagrange equations of motion have been derived from variational principles, and we have seen that such cases typically lead to the phase space being either the tangent or the cotangent bundle of a Lie group [12]. We have applied RKMK methods to solve the equation of motion in their Hamiltonian formulation, for a free rigid body and a large system of coupled dipolar spheres. We obtained excellent results in the numerical experiments.

Future work:

The particles in the water molecule experiment are in an unbounded domain. It could be an interesting test for the methods to see what would happen if they were in a bounded domain. Would the state of the system go towards a (dynamic or static) equilibrium?

Another interesting direction is to consider model reduction; even in a 100 particle system the calculations take a long time and so it might be of interest to use the techniques from [9] to reduce the size of the system and thus, also computational costs. This must however be done in such a way that it does not interfere with the nice properties of Lie group methods such as e.g. orthogonality preservation.

It could also be interesting to implement other Lie group integrators, and perhaps also try other coordinate maps than the exponential map. Then by comparing numerical results we could determine the integrators that best suit the problem. In addition, the flows of the ODEs we have considered are symplectic, thus, preserving this structure by applying a symplectic integrator might give an even better long time behaviour [5].

Bibliography

[1] R. Abraham, J. E. Marsden, and T. Ratiu. *Manifolds, tensor analysis, and applications*, volume 75 of *Applied Mathematical Sciences*. Springer-Verlag, New York, second edition, 1988.

[2] V. I. Arnold. *Mathematical methods of classical mechanics*, volume 60 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1989. Translated from Russian by K. Vogtmann and A. Weinstein.

[3] V. I. Arnold and B. A. Khesin. *Topological methods in hydrodynamics*, volume 125 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 1998.

[4] E. D. Bloch. *Proofs and fundamentals*. Undergraduate Texts in Mathematics. Springer, New York, 2011. A first course in abstract mathematics, Second edition.

[5] G. Bogfjellmo and H. Marthinsen. High-order symplectic partitioned lie group methods. *Foundations of computational mathematics*, 16(2):493–530, 2016.

[6] E. Celledoni, F. Fassò, N. Säfström, and A. Zanna. The exact computation of the free rigid body motion and its use in splitting methods. *SIAM J. Sci. Comput.*, 30(4):2084–2112, 2008.

[7] E. Celledoni, H. Marthinsen, and B. Owren. An introduction to Lie group integrators - basics, new developments and applications. *J. Comput. Phys.*, 257(part B):1040–1061, 2014.

[8] E. Celledoni and B. Owren. Lie group methods for rigid body dynamics and time integration on manifolds. *Computer methods in applied mechanics and engeneering*, 192(3-4):421–438, 2003.

[9] E. Celledoni, B. Owren, and H. Parks. Structure preserving model reduction on Lie groups. Unpublished article, 2015.

[10] A. Dullweber, B. Leimkuhler, and R. McLachlan. Symplectic splitting methods for rigid body molecular dynamics. *Journal of chemical physics*, 107(15):5840–5851, 1997.

[11] K. Engø. On the construction of geometric integrators in the RKMK class. *BIT*, 40(1):41–61, 2000.

[12] K. Engø. Partitioned Runge-Kutta methods in Lie-group setting. *BIT*, 43(1):21–39, 2003.

[13] E. Hairer. Solving differential equations on manifolds. *Lecture Notes, Universite de Geneve, Geneva, Switzerland*, 2011.

[14] E. Hairer, C. Lubich, and G. Wanner. *Geometric numerical integration*, volume 31 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, second edition, 2006.

[15] B. Hall. *Lie groups, Lie algebras, and representations*, volume 222 of *Graduate Texts in Mathematics*. Springer, Cham, second edition, 2015.

[16] A. Iserles, H. Munthe-Kaas, S. P. Nørsett, and A. Zanna. Lie-group methods. In *Acta numerica, 2000*, volume 9 of *Acta Numer.*, pages 215–365. Cambridge Univ. Press, Cambridge, 2000.

[17] J. M. Lee. *Introduction to smooth manifolds*, volume 218 of *Graduate Texts in Mathematics*. Springer, New York, second edition, 2013.

[18] J. E. Marsden and T. Ratiu. *Introduction to mechanics and symmetry: a basic exposition of classical mechanical systems*, volume 17. Springer Science & Business Media, 2013.

[19] H. Munthe-Kaas. Runge-Kutta methods on Lie groups. *BIT*, 38(1):92–111, 1998.

[20] H. Munthe-Kaas. High order Runge-Kutta methods on manifolds. In *Proceedings of the NSF/CBMS Regional Conference on Numerical Analysis of Hamiltonian Differential Equations (Golden, CO, 1997)*, volume 29, pages 115–127, 1999.

[21] P. J. Olver. *Applications of Lie groups to differential equations*, volume 107. Springer Science & Business Media, 2000.

[22] B. Owren. A crash course on manifolds. 2015. Lecture notes.

[23] B. Owren. The foundations of Runge-Kutta-Munthe-Kaas methods. 2015. Lecture notes.

[24] B. Owren. Notes for variational Lie group integrators. 2016. Lecture notes.

[25] B. Owren and A. Marthinsen. Runge-Kutta methods adapted to manifolds and based on rigid frames. *BIT*, 39(1):116–142, 1999.

[26] E. Süli and D. F. Mayers. *An introduction to numerical analysis*. Cambridge University Press, Cambridge, 2003.

[27] J. L. Troutman. *Variational calculus and optimal control*. Undergraduate Texts in Mathematics. Springer-Verlag, New York, second edition, 1996.

[28] L. W. Tu. *An introduction to manifolds*. Universitext. Springer, New York, second edition, 2011.

[29] V. S. Varadarajan. *Lie groups, Lie algebras, and their representations*, volume 102 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1984. Reprint of the 1974 edition.

[30] F. W. Warner. *Foundations of differentiable manifolds and Lie groups*, volume 94 of *Graduate Texts in Mathematics*. Springer-Verlag, New York-Berlin, 1983. Corrected reprint of the 1971 edition.