Fakultät 4: Mathematik und Naturwissenschaften
Fachgruppe Mathematik und Informatik

BERGISCHE
UNIVERSITÄT
WUPPERTAL

Runge-Kutta–Munthe-Kaas schemes for Stochastic Differential Equations on Manifolds

Dissertation
zur Erlangung des akademischen Grades
Doktor der Naturwissenschaften (Dr. rer. nat.)

vorgelegt von
Michelle Christine Muniz

Prüfungskommission:

Prof. Dr. Michael Günther
Prof. Dr. Matthias Ehrhardt
Dr. Renate Winkler
Prof. Dr. Thomas Kruse
Prof. Dr. Pavel Petrov
Prof. Dr. Bruno Lang

Acknowledgements

Open the door (*idiom*) - to make (something) easier or more likely to happen[1]

How to open a door is a question that accompanied me throughout my years as a PhD student and I am happy to say that I finally know how to answer it now according to the definition above and thanks to the people I am about to mention.

First of all, I would like to thank Prof. Dr. Michael Günther and Prof. Dr. Matthias Ehrhardt for supervising this PhD thesis. I am especially grateful for the time you took to be puzzled about the same problems and trying to solve them with me in what seemed like endless but insightful hours in our seminar room.

Thank you, Michael, for intriguing me early on with the topic of geometric numerical integration and for already supervising my Bachelor's and Master's thesis in this research area applied to Lattice Quantum Chromodynamics. By doing so, I realize now, that you made pursuing a PhD more likely to happen for me. So I thank you once again for opening the door to doctoral studies, where the topics that you taught me have influenced my research interests up to this date and the result can be seen in the thesis presented here.

Thank you, Matthias, for not only opening doors for me but also for (literally) showing me how to open doors for others and for myself. Given a (mathematical) puzzle, that I struggled with to solve alone, you showed me how I can make it easier by sharing and collaborating with other colleagues in our offices, at conferences or within your DAAD projects. With your office being next-door to mine I can confirm the benefits of having short work paths and am grateful for always letting me come by with whatever problem I had.

Next, I thank Dr. Renate Winkler for giving the best lecture in *Einführung in die Numerik gewöhnlicher Differentialgleichungen*. Looking back, this lecture helped me to get my foot in the door of the research I am doing today. Most importantly, I appreciate

[1]Merriam-Webster. (n.d.). Open the door. In Merriam-Webster.com dictionary. Retrieved January 30, 2023, from https://www.merriam-webster.com/dictionary/open%20the%20door

you coming on board when Michael, Matthias and I came knocking on your door with a problem concerning stochastic convergence analysis. Your expertise in this field have made the hours spent in our seminar room working on my problems seem less endless. I am glad to have had you as a companion at the beginning and at the end of my studies.

Having a path beaten to the door of Dr. Kevin Kamm, I would like to express my gratitude for showing me applications of our research in financial mathematics. The joint works with you enriched my journey as a PhD student and have enlightened my further research paths.

Thanks to the *keymaster*, Prof. Dr. Pavel Petrov. Just as a keymaster holds the keys to unlock various doors, you possess the knowledge, expertise and willingness to guide and assist me and others in accessing new knowledge and possibilities that helped me to jump over any obstacle and to break self-constructed loops.

I gratefully acknowledge the support given by the QSL network which made my PhD position possible. More specifically, I extend my gratitude to Chris, Fabia, Kevin, Saskia and Tom for sharing the workload of a doorman's/doorwoman's job among other tasks.

Next, I would like to thank Andreas, Hanna, Long, Michèle and Thomas for always having an open door for discussions. Whenever I was stuck with a problem, you readily took the time to listen and tried to help in all possible ways.

Thanks to my colleagues Anna, Christoph, Daniel, Daniele, Friedemann, Jan, Jens, Kevin, Konrad, Lorenc, Sarah and Tatiana for always keeping the doors open for coffee and puzzle breaks. Despite the shared torments that come with pursuing a PhD we always had fun doing indoor as well as outdoor activities like bowling, hiking, playing Bingo or watching a football match.

I thank my friends Adrian, Chrisi, Deng, Johannes, Niklas, Paula, Sophie, Sranya and Vanessa for helping me forget about the heavy doors at university whenever I needed to and for making me see the perks of living in Wuppertal.

Last but not least, I thank my family, especially my mother Lina and my sister Sarah, for the love and support throughout the years and for not having me think about doors as much as I did while writing these acknowledgements before I left for university.

Contents

Chapter 1

Introduction

We can find stochastic differential equations (SDEs) on manifolds in many problems in different areas ranging from physics to finance. In physics, the examples range from the modelling of rigid bodies such as vehicles and satellites to the modelling of the precessional motion of magnetisation, see [1], [44], [63]. Problems in computer vision, where the motions of objects are estimated from a sequence of projections, are also modelled by stochastic processes on manifolds [68]. Examples in finance include the interest rate models found in [21], [41], [61], [77].

What all of these examples have in common is that they are subjected to some kind of uncertainty and geometric constraints, which are expressed by a stochastic process and a manifold, respectively. We can think of manifolds as curved spaces, e.g. a sphere, such that we know that applying a linear operator on this curved space will give us results that leave the curved space. This phenomenon is called a *drift-off* and an example is visualized in Figure 1.1, where a linear solver was used to approximate a stochastic process on a sphere. The approximations start on the left-hand side of the sphere but dissociate from it

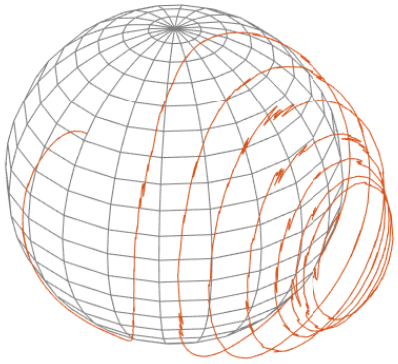

Figure 1.1: Approximation of a stochastic process drifting off a sphere

more and more with every time step taken, which we can see on the right-hand side of the sphere. Since leaving the manifold is equivalent to violating some geometric constraints, our objective is to generate approximations that stay on the manifold.

In this thesis, we extend the linear solvers that are designed for SDEs in linear spaces like stochastic Runge-Kutta methods such that they can be used for SDEs on manifolds by taking up the idea of Munthe-Kaas for solving ordinary differential equations (ODEs) on manifolds. The resulting methods are called *stochastic Runge-Kutta–Munthe-Kaas* (SRKMK) schemes.

1.1 Literature review

SDEs on manifolds have been investigated for more than seven decades now. It started with Itô's paper [31] in 1950 and multiple papers and books followed throughout the years, for example [4], [17], [18], [23], [29], [69] to name a few, which gave us a good foundation on the theory of stochastic analysis on manifolds.

The first essential components for the numerical approximation of the solution of SDEs on manifolds were established around three decades later when stochastic versions of the *Magnus expansion* were analysed, see [3], [12], [37], [79], which were put into practice in [11], [50]. However, it was not until 2008 when SRKMK schemes were first proposed and computed by Malham and Wiese in [44] based on the methods developed by Munthe-Kaas in the course of [56]–[58].

After that, a specific example of SRKMK schemes based on the Euler-Maruyama scheme has been studied in [46], [47], [64], where the method was applied to a special kind of manifolds, namely the matrix Lie groups $\mathrm{SO}(n)$ and $\mathrm{SE}(n)$ for the modelling of rigid body motions. As the considered method was able to preserve the geometry of the Lie groups, the authors called it the *geometric Euler-Maruyama* scheme.

Another particular application of SRKMK schemes can be found in [1], where the authors used Munthe-Kaas methods for a finite-dimensional version of the stochastic Landau–Lifshitz equation.

The results presented in this thesis get in line with the works mentioned above and complete them with an analysis of strong convergence of SRKMK methods. Despite all the applications of SRKMK schemes that can be read about in the cited papers, a proof of strong convergence of Munthe-Kaas methods for SDEs on manifolds was still missing. Researchers interested in convergence proofs will only find the first order convergence in mean-squared error of the geometric Euler-Maruyama scheme applied to linear SDEs

on matrix Lie groups in [46] and the proof of weak convergence of SRKMK schemes on the unit sphere in [1]. For that reason we consider our proof of strong convergence of SRKMK schemes for nonlinear SDEs on manifolds as the main contribution to this stream of research.

Furthermore, we regard our application of SRKMK schemes to SDEs on manifolds considered in financial mathematics as great contributions to geometric modelling in computational finance since we have elaborated novel methodologies to approximate correlation matrices and transition matrices. We present these methodologies at the end of the thesis, where the approximation of correlation matrices is based on the following publications:

- M. M., M. Ehrhardt, M. Günther, *Approximating correlation matrices using stochastic Lie group methods*, Mathematics 9(1) (2021), 94,

- M. M., M. Ehrhardt, M. Günther, *Correlation Matrices driven by Stochastic Isospectral Flows*, in: M. Ehrhardt, M. Günther, *Progress in Industrial Mathematics at ECMI 2021*, The European Consortium for Mathematics in Industry, Springer, 2022, 455–461.

A full-length coverage of our method for the approximation of rating transition matrices can be found in

- K. Kamm and M. M., *A novel approach to rating transition modeling via Machine Learning and SDEs on Lie Groups*, arXiv:2205.15699, May 2022,

- K. Kamm and M. M., *Rating Triggers for Collateral-Inclusive XVA via Machine Learning and SDEs on Lie Groups*, arXiv:2211.00326, November 2022.

This thesis will only cover a small introduction to the approach described in more details in the listed manuscripts.

The proof of strong convergence of SRKMK methods was developed in the course of the three manuscripts

- M. M., M. Ehrhardt, M. Günther, R. Winkler, *Higher Strong Order Methods for linear Itô SDEs on matrix Lie Groups*, BIT Numer. Math. 62 (4) (2022), 1095–1119,

- M. M., M. Ehrhardt, M. Günther, R. Winkler, *Stochastic Runge-Kutta–Munthe-Kaas Methods in the Modelling of Perturbed Rigid Bodies*, Adv. Appl. Math. Mech. 14 (2) (2022), 528–538,

- M. M., M. Ehrhardt, M. Günther, R. Winkler, *Strong stochastic Runge-Kutta-Munthe-Kaas methods for nonlinear Itô SDEs on manifolds*, IMACM-Preprint Wuppertal 22/14, June 2022,

where the first paper only considers linear SDEs on matrix Lie groups, the second deals with a nonlinear SDE on a sphere and the third examines the generalisations to nonlinear SDEs on more abstract manifolds. Here we will present the convergence proof on the basis of these three manuscripts, where the focus is on the latest version of the proof since in hindsight the former two approaches can be considered as special cases of that version.

1.2 Outline

Our goal in this thesis is to present the proof of strong convergence in a comprehensible way for readers that are familiar with the basics of solving ODEs and SDEs in vector spaces. For this purpose, we approach the main result step by step and begin each chapter with a motivational example from the modelling of rigid bodies.

In Chapter 2 we introduce all the vocabulary from differential geometry needed to understand and to approximate the dynamics equation of the rotation of a rigid body, which is an ODE on a sphere. Therefore, we consider manifolds and their tangent spaces to formulate ODEs on manifolds. Then, we regard Lie groups and Lie algebras, which are special manifolds and tangent spaces that allow us to express transport across a manifold. Next, we use these notions of transport across a manifold to set up Runge-Kutta–Munthe-Kaas schemes, which can be used to solve the rigid body equation. Since the formulation of Itô SDEs on manifolds requires a greater concept than tangent spaces and their corresponding tangent bundle, we will also discuss fiber bundles, which will be needed in the third chapter.

The third chapter starts with the set up of a stochastic expansion of the rotation motion of a rigid body, i.e. an SDE on a sphere. In order to fully comprehend this equation we first give an introduction to SDEs in linear spaces and the different definitions of stochastic integrals according to Itô and to Stratonovich. As a next step, we concentrate on how these SDEs in linear spaces can be numerically approximated by stochastic Taylor expansions and stochastic Runge-Kutta methods. Finally, we will be able to connect the notions learned from the second chapter with SDEs to establish the formulation of SDEs on manifolds and their numerical approximation via SRKMK schemes. This is where we will also present our main result of the strong convergence of SRKMK schemes before naming some more methods that could be used to approximate the rigid body problem.

In the fourth chapter we will see the results of applying SRKMK schemes to SDEs considered in rigid body modelling. We visualize the result of our convergence theorem and the preservation of the manifold structure and of other given intrinsic quantities. Moreover, we show how SRKMK methods can be used to model correlations and rating transitions in computational finance.

Lastly, a summary of this work and an outlook to future topics connected to this work will conclude the thesis.

Chapter 2

Numerical Approximation of Ordinary Differential Equations on Manifolds

Motivation: The dynamics of a rigid body

Consider a rigid body such as a satellite rotating around its center of mass. Let $y = (y_1, y_2, y_3)^\top$ be the angular momentum in the body frame and let I_1, I_2, I_3 represent the principal moments of inertia, see [49]. Then, the dynamics of the rotation can be described by the Euler equations, which can be expressed by

$$\begin{pmatrix} \dot{y}_1 \\ \dot{y}_2 \\ \dot{y}_3 \end{pmatrix} = \begin{pmatrix} 0 & y_3/I_3 & -y_2/I_2 \\ -y_3/I_3 & 0 & y_1/I_1 \\ y_2/I_2 & -y_1/I_1 & 0 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}. \tag{2.1}$$

This is an ordinary differential equation (ODE) on the sphere

$$\mathcal{S} = \{(y_1, y_2, y_3) : y_1^2 + y_2^2 + y_3^2 = const.\}. \tag{2.2}$$

Our goal in this chapter is to formulate a method to approximate the solution of the ODE (2.1) such that the structure of the sphere is preserved while at the same time we are introducing vocabulary needed for solving the stochastic counterpart of this problem, which we will consider later on. For this purpose we first introduce some basics of differential geometry in Section 2.1 before establishing the geometry-preserving algorithm by Munthe-Kaas in Section 2.2.

2.1 Manifolds and tangent bundles

Our main scene of action in this thesis are manifolds, which is why we dedicate this entire section to manifolds. Roughly speaking, manifolds are geometrical spaces that locally resemble Euclidean spaces and examples include curves and surfaces like the one given in (2.2). The following more formal definitions of manifolds and attached notions from differential geometry can for example be found in [5], [8], [39], [40], where we used the notations from [24, Chapter 1].

So more formally, we consider a manifold $\mathcal{M}$ to be a topological space such that any point in $\mathcal{M}$ has a neighbourhood homeomorphic to an open ball of the vector space $\mathbb{R}^n$, where $n \in \mathbb{N}$ is the dimension of the manifold. In the following, we will not distinguish between the neighbourhood on $\mathcal{M}$ and the open ball in $\mathbb{R}^n$ and we will both call charts denoted by U_α. Let $m \in \mathcal{M}$ be in the intersection of two charts U_α and U_β, i.e. $m \in U_{\alpha\beta} = U_\alpha \cap U_\beta \neq \emptyset$, then m can be described by the local coordinates in U_α and in U_β. We denote the map that transforms U_α-coordinates to U_β-coordinates by $\varphi_{\beta\alpha}$. A manifold $\mathcal{M}$ is called *smooth* if the changes of coordinates are C^∞-smooth. From here on we consider all manifolds to be smooth and finite dimensional if not stated otherwise.

2.1.1 Tangent spaces and vector fields

Consider a smooth curve $m(t)$ on a manifold $\mathcal{M}$. Visualizing the manifold as a surface, a *tangent vector* can be thought of as a velocity vector to this smooth curve on the surface. So we denote by

$$X_m = \left.\frac{\mathrm{d}m(t)}{\mathrm{d}t}\right|_{t=0}$$

a tangent vector at $m = m(0)$. If we assume that the tangent vector is represented by $X_\alpha = \left.\frac{\mathrm{d}m_\alpha(t)}{\mathrm{d}t}\right|_{t=0}$ in the coordinates of a chart U_α, then we have in another chart U_β the curve $m_\beta(t) = \varphi_{\beta\alpha}(m_\alpha(t))$ and the tangent vector

$$X_\beta = \left.\frac{\mathrm{d}m_\beta(t)}{\mathrm{d}t}\right|_{t=0} = \varphi'_{\beta\alpha} \left.\frac{\mathrm{d}m_\alpha(t)}{\mathrm{d}t}\right|_{t=0} = \varphi'_{\beta\alpha} X_\alpha, \tag{2.3}$$

where $\varphi'_{\beta\alpha}$ is the Jacobi matrix of $\varphi_{\beta\alpha}$. The tangent vectors at a point $m \in \mathcal{M}$ form a vector space $T_m\mathcal{M}$ which is called the *tangent space* at m.

Let $(x^1, \ldots, x^n)$ be coordinates in a chart U_α at m. Then every tangent space $T_m\mathcal{M}$ has a basis denoted by $\frac{\partial}{\partial x^1}, \ldots, \frac{\partial}{\partial x^n}$, where the vector $\frac{\partial}{\partial x^i}$ is the derivative of the i-th coordinate axis passing through m with respect to the corresponding parameter x^i. Hence, every $X_m \in T_m\mathcal{M}$ is represented with coordinates as $X^i \frac{\partial}{\partial x^i}$ using Einstein convention.

The *tangent bundle* $T\mathcal{M}$ in turn is the disjoint union of the tangent spaces $T_m\mathcal{M}$ of $\mathcal{M}$. By assigning a tangent vector to each point $m \in \mathcal{M}$ we get a map, $X\colon \mathcal{M} \to T\mathcal{M}$, which is called a *vector field*. We denote by $\mathfrak{X}(\mathcal{M})$ the set of all smooth vector fields on $\mathcal{M}$. Hence, an *integral curve* $m(t)$ of $X \in \mathfrak{X}(\mathcal{M})$ or rather an ordinary differential equation (ODE) on the manifold $\mathcal{M}$ is described by

$$\dot{m}(t) = X\big(m(t)\big). \tag{2.4}$$

Locally, we can consider the integral curve $m(t)$ as an ODE in a vector space and use classical theorems for showing the existence and uniqueness of a solution $m(t)$ given an initial condition $m(0) = m_0$, see e.g. [24, p. 7] for further details.

Let $f\colon \mathcal{M} \to \mathcal{N}$ be a smooth map between two manifolds $\mathcal{M}$ and $\mathcal{N}$. The *tangent map* (or pushforward) $Tf\colon T\mathcal{M} \to T\mathcal{N}$ sends the tangent vector $X_m = \mathrm{d}m(t)/\mathrm{d}t$, $t = 0$, to a tangent vector in $T\mathcal{N}$ via

$$Tf(X_m) = \left.\frac{\mathrm{d}f(m(t))}{\mathrm{d}t}\right|_{t=0}. \tag{2.5}$$

Setting $\mathcal{N} = \mathbb{R}$ and assuming $m(t)$ to be the integral curve of X with $m(0) = m$, the tangent vector $Tf(X_m)$ is also called the derivative of f in the direction of X at m. The derivative of f along X is obtained by computing $Tf(X_m)$ at all points of $\mathcal{M}$ and denoted by Xf. In local coordinates we have the formula

$$Xf = X^i \frac{\partial f}{\partial x^i}.$$

The *Lie bracket* $[X, Y]$ of two smooth vector fields X and Y is again a vector field on $\mathcal{M}$ such that for any smooth real-valued function f on $\mathcal{M}$ its derivative along $[X, Y]$ is given by $[X, Y]f = X(Yf) - Y(Xf)$, where

$$[X, Y] = X \circ Y - Y \circ X \tag{2.6}$$

is the definition in operator form. Besides skew-symmetry, i.e. $[X, Y] = -[Y, X]$, the Lie bracket also satisfies the so-called *Jacobi identity*:

$$\big[X, [Y, Z]\big] + \big[Y, [Z, X]\big] + \big[Z, [X, Y]\big] = 0 \tag{2.7}$$

for $X, Y, Z \in \mathfrak{X}(\mathcal{M})$. A vector space on which the Lie bracket fulfills the Jacobi identity is called a *Lie algebra*.

2.1.2 Lie groups and Lie algebras

Manifolds with a group structure prove to be an important tool for describing dynamics on manifolds. Therefore, we explore these special manifolds and their tangent spaces based on [27].

Definition 2.1. *A Lie group G is a differentiable manifold equipped with a continuous group product $\cdot\colon G \times G \to G$.*

Hereafter, we will focus on matrix Lie groups, i.e. Lie groups that are also subgroups of the general linear group $\mathrm{GL}(n) = \{A \in \mathbb{R}^{n\times n} : \det(A) \neq 0\}$ such that the group product in the definition above is the matrix multiplication. A typical example for a matrix Lie group that we will encounter frequently throughout this thesis is the group of rotation matrices also called the *special orthogonal group*,

$$\mathrm{SO}(n) = \{Q \in \mathrm{GL}(n) : Q^\top Q = I, \det(Q) = 1\},$$

where I denotes the n-dimensional identity matrix.

Definition 2.2. *A (left) Lie group action is a map $\Lambda\colon G \times \mathcal{M} \to \mathcal{M}$ which satisfies*

1. $\Lambda(I, m) = m$,
2. $\Lambda\big(g_1, \Lambda(g_2, m)\big) = \Lambda(g_1 \cdot g_2, m)$ *for* $g_1, g_2 \in G$.

In the case where $\mathcal{M}$ is the Lie group G itself, a left Lie group action is given by right multiplication $R_m\colon G \to G$, $\Lambda(g, m) = g \cdot m =: R_m(g)$. The tangent map of the right multiplication TR_m can be used to transport a basis in the tangent space at the identity T_IG to a tangent space T_gG at each point $g \in G$. The so-obtained vector field on G is called the *right-invariant vector field*.

The vector space T_IG is called the *Lie algebra of the Lie group* G and is denoted by $\mathfrak{g}$. It is equipped with the Lie bracket (2.6), which in the case of matrix Lie groups and matrix Lie algebras reduces to the *matrix commutator* $[\cdot,\cdot]\colon \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$. The commutator is bilinear, skew-symmetric and satisfies the Jacobi identity (2.7).

An iterated application of the commutator can be expressed by the *adjoint operator*, $\mathrm{ad}_\Omega\colon \mathfrak{g} \to \mathfrak{g}$, $\mathrm{ad}_\Omega(H) = [\Omega, H] = \Omega H - H\Omega$ with

$$\mathrm{ad}_\Omega^0(H) = H, \quad \mathrm{ad}_\Omega^k(H) = \mathrm{ad}_\Omega\left(\mathrm{ad}_\Omega^{k-1}(H)\right) = [\Omega, \mathrm{ad}_\Omega^{k-1}(H)]$$

for $k \geq 1$.

Going back to the matrix Lie group $\mathrm{SO}(n)$ one can easily verify that its corresponding Lie algebra is the space of skew-symmetric matrices,

$$\mathfrak{so}(n) = \{A \in \mathbb{R}^{n\times n} : A + A^\top = \mathbf{0}\},$$

where $\mathbf{0}$ denotes the n-dimensional zero matrix. Note that there is an isomorphism between $\mathbb{R}^3$ and $\mathfrak{so}(3)$, which we call the *hat map*. It is given by $\hat{\cdot}\colon \mathbb{R}^3 \to \mathfrak{so}(3)$,

$$\theta = \begin{pmatrix}\theta_1\\ \theta_2\\ \theta_3\end{pmatrix} \mapsto \hat{\theta} = \begin{pmatrix}0 & -\theta_3 & \theta_2\\ \theta_3 & 0 & -\theta_1\\ -\theta_2 & \theta_1 & 0\end{pmatrix}, \tag{2.8}$$

such that $\hat{\theta}z = \theta \times z$ for $\theta, z \in \mathbb{R}^3$.

The exponential map

The matrix exponential gives us a map from a Lie algebra to its corresponding Lie group, i.e. $\exp\colon \mathfrak{g} \to G$ with

$$\exp(\Omega) = \sum_{k=0}^{\infty} \frac{\Omega^k}{k!}.$$

Moreover, it is a local diffeomorphism in a neighbourhood of $\Omega = \mathbf{0}$, see [25, p. 83ff] for the proof and the formulas regarding the derivative below. The directional derivative of the matrix exponential along an arbitrary matrix H is given by

$$\Big(\frac{\mathrm{d}}{\mathrm{d}\Omega}\exp(\Omega)\Big)H = \frac{\mathrm{d}}{\mathrm{d}t}\exp(\Omega + tH)\Big|_{t=0} = \mathrm{dexp}_\Omega(H)\exp(\Omega) \quad \text{with}$$
$$\mathrm{dexp}_\Omega(H) = \sum_{k=0}^{\infty} \frac{1}{(k+1)!}\,\mathrm{ad}_\Omega^k(H), \tag{2.9}$$

where the series converges for all Ω.

Lemma 2.3 (Baker, 1905)**.** *If the eigenvalues of* ad_Ω *are different from* $2\ell\pi i$ *with* $\ell \in \{\pm 1, \pm 2, \dots\}$*, then* dexp_Ω *is invertible. Let* B_k *denote the Bernoulli numbers defined implicitly by the Taylor series* $\sum_{k=0}^{\infty}(B_k/k!)x^k = x/(e^x - 1)$*, then we have*

$$\mathrm{dexp}_\Omega^{-1}(H) = \sum_{k=0}^{\infty} \frac{B_k}{k!}\,\mathrm{ad}_\Omega^k(H), \tag{2.10}$$

which converges for $\|\Omega\| < \pi$ *in a submultiplicative norm* $\|\cdot\|$*.*

We recall that the first three Bernoulli numbers are $B_0 = 1$, $B_1 = -\frac{1}{2}$, $B_2 = \frac{1}{6}$ and that $B_{2k+1} = 0$ holds for $k \in \mathbb{N}$ such that

$$\operatorname{dexp}_\Omega^{-1}(H) = H - \frac{1}{2}[\Omega, H] + \frac{1}{12}\big[\Omega, [\Omega, H]\big] + \dots.$$

Seeing all these infinite sums that are involved with the matrix exponential one may wonder whether closed-form expressions exist. In the special case, where we are interested in computations with rotations in three dimensions, we can rely on the *Rodrigues formula* (see [49, p. 291]), $\operatorname{expm}\colon \mathfrak{so}(3) \to \mathrm{SO}(3)$,

$$\operatorname{expm}(\Omega) = I + \frac{\sin\theta}{\theta}\Omega + \frac{1}{2}\frac{\sin^2\vartheta}{\vartheta^2}\Omega^2 \tag{2.11}$$

with $\theta = \|\omega\|_2$ and $\vartheta = \theta/2$, where $\hat{\omega} = \Omega \in \mathfrak{so}(3)$. Based on these notations we can also find closed-form expressions for (2.9) and (2.10) for matrices in $\mathfrak{so}(3)$, namely

$$\operatorname{dexp}_\Omega = I + \frac{\sin^2\vartheta}{2\vartheta^2}\Omega + \frac{\theta - \sin\theta}{\theta^3}\Omega^2,$$

where the inverse is given by

$$\operatorname{dexp}_\Omega^{-1} = I - \frac{1}{2}\Omega - \frac{\theta\cot\vartheta - 2}{2\theta^2}\Omega^2, \tag{2.12}$$

see [30, p. 147].

The Cayley map

Besides the exponential map there are more local diffeomorphisms from a Lie algebra $\mathfrak{g}$ to a corresponding Lie group G near $\Omega = \mathbf{0}$. Here, we will only have a closer look on the Cayley transform,

$$\operatorname{cay}\colon \mathfrak{g} \to G, \quad \operatorname{cay}(\Omega) = (I - \Omega)^{-1}(I + \Omega), \tag{2.13}$$

which can be used in case of *quadratic Lie groups*. A quadratic Lie group is of the form

$$G = \{Q : Q^\top P Q = P\}$$

for a given constant matrix P. Note that the Lie group $\mathrm{SO}(n)$ is quadratic by setting $P = I$.

For the derivative of $\operatorname{cay}(\Omega)$ we have

$$\Big(\frac{\mathrm{d}}{\mathrm{d}\Omega}\operatorname{cay}(\Omega)\Big)H = \operatorname{dcay}_\Omega(H)\operatorname{cay}(\Omega)$$

with

$$\operatorname{dcay}_\Omega(H) = 2(I-\Omega)^{-1}H(I+\Omega)^{-1}$$

and its inverse given by

$$\operatorname{dcay}_\Omega^{-1}(H) = \frac{1}{2}(I-\Omega)H(I+\Omega), \tag{2.14}$$

see [25, p. 128].

Examining the Cayley map we notice that compared to the matrix exponential, one is not required to evaluate any infinite series.

2.1.3 Fiber bundles

In the following, we would like to piece together the objects evolving around manifolds that we have learned so far in order to understand a greater concept, namely *fiber bundles*. Fiber bundles are spaces that can locally be identified by a Cartesian product of topological spaces. We use the notations from [24, Section 1.3] to give a more formal definition.

Definition 2.4 (Fiber bundle)**.** *A fiber bundle is determined if the following five objects are given*

- *a manifold $\mathcal{M}$, the base,*
- *a manifold E, the total space,*
- *a manifold F, the standard fiber,*
- *a Lie group G, the structure group,*
- *a smooth projection $\pi_{\mathcal{M}}\colon E \to \mathcal{M}$*

and if the following interrelations hold

- *a certain left action of G on F is given,*
- *for any $m \in \mathcal{M}$ the set $E_m = \pi_{\mathcal{M}}^{-1}m$ (the fiber at m) is homeomorphic to F,*

- *for any chart U_α in $\mathcal{M}$ the set $\pi_{\mathcal{M}}^{-1}U_\alpha$ is homeomorphic to $U_\alpha \times F$,*
- *for $U_{\alpha\beta} = U_\alpha \cap U_\beta \neq \emptyset$ the 'change of coordinates' from $U_\alpha \times F$ to $U_\beta \times F$ is given by the pair $(\varphi_{\beta\alpha}, g_{\beta\alpha}(m))$ where $\varphi_{\beta\alpha}\colon U_{\alpha\beta} \to U_{\alpha\beta}$ and $g_{\beta\alpha}(m)\colon F \to F$ is a diffeomorphism of F according to the left action of G.*

A fiber bundle is often denoted by its total space E over its base manifold $\mathcal{M}$. The reader might recognize that we have already encountered a special kind of fiber bundle, namely the tangent bundle $T\mathcal{M}$.

The tangent bundle $T\mathcal{M}$ of a manifold $\mathcal{M}$ is a fiber bundle over $\mathcal{M}$, where fibers at $m \in \mathcal{M}$ are the tangent spaces $T_m\mathcal{M}$ which are homeomorphic to the standard fiber $F = \mathbb{R}^n$. An element of the tangent bundle $T\mathcal{M}$ is a tangent vector which we may denote by

- (m, X) if we want to highlight that we are considering a point in $T\mathcal{M}$ (locally homeomorphic to $U_\alpha \times \mathbb{R}^n$) or
- X_m if we want to highlight that we are considering it as a tangent vector to $\mathcal{M}$ at m (as done above).

Moreover, we see that for $E = T\mathcal{M}$ we have the structure group $G = \mathrm{GL}(n)$ and $g_{\beta\alpha} = \varphi'_{\beta\alpha}$ according to (2.3).

Definition 2.5 (Section)**.** *A section X of the fiber bundle E is a map $X\colon \mathcal{M} \to E$ such that $\pi_{\mathcal{M}} \circ X = \mathrm{id}\colon \mathcal{M} \to \mathcal{M}$, i.e. $X(m) \in E_m$ for any $m \in \mathcal{M}$.*

Given this definition we can now specify a vector field $X\colon \mathcal{M} \to T\mathcal{M}$ as a section of the tangent bundle $T\mathcal{M}$.

More examples of fiber bundles include

- the *cotangent bundle* $T^*\mathcal{M}$, which is the set of all cotangent spaces $T_m^*\mathcal{M}$ and by definition the dual space of the tangent bundle $T\mathcal{M}$,
- the *second order tangent bundle* $\tau\mathcal{M}$, which consists of second order tangent vectors, i.e. second order differential operators on $\mathcal{M}$ with zero constant term and a symmetric matrix of coefficients at second order derivatives in local coordinates, and
- the *second tangent bundle* of $\mathcal{M}$ denoted by $TT\mathcal{M}$, which is the tangent bundle of the manifold $T\mathcal{M}$.

For more information on these fiber bundles we refer to [24].

2.2 Runge-Kutta–Munthe-Kaas schemes

For the demonstration of a numerical approximation scheme for ODEs on manifolds we first take a closer look on manifolds that are *homogeneous spaces* in Section 2.2.1, where the main references are [58], [59]. A homogeneous manifold is a manifold $\mathcal{M}$ with a *transitive* Lie group action $\Lambda\colon G \times \mathcal{M} \to \mathcal{M}$ (see Definition 2.2), which means that there is an element g in the Lie group G such that starting from one point on the manifold we can reach any other point on the manifold,

$$\Lambda(g, m_1) = m_2$$

for all $m_1, m_2 \in \mathcal{M}$. Examples for such kind of manifolds will be given below in Section 2.2.2.

2.2.1 ODEs on homogeneous manifolds

We already know that ODEs on manifolds $\mathcal{M}$ can be described by (2.4) and that the right hand side of this ODE is a right-invariant vector field in the case where $\mathcal{M}$ is a Lie group G. Considering homogeneous manifolds $\mathcal{M}$ this representation of an ODE can again be specified by using *Lie algebra actions* on $\mathcal{M}$.

The Lie algebra action

Given a (left) Lie group action $\Lambda\colon G\times\mathcal{M} \to \mathcal{M}$ a (left) Lie algebra action $\lambda\colon \mathfrak{g}\times\mathcal{M} \to \mathcal{M}$ can be defined by

$$\lambda(\Omega, y) = \Lambda(\exp(\Omega), y),$$

see [58]. We will call Lie algebra actions, that originate from transitive Lie group actions, transitive as well.

Note that a Lie algebra action λ is not uniquely defined by a Lie group action Λ. Other local diffeomorphisms of the Lie group can also be used instead of the matrix exponential, e.g. the Cayley map in the case of quadratic Lie groups.

Now we can use Lie algebra actions $\lambda\colon \mathfrak{g} \times \mathcal{M} \to \mathcal{M}$ to map the Lie algebra $\mathfrak{g}$ onto tangent spaces of $\mathcal{M}$ or rather to use elements of the Lie algebra to generate vector fields on the homogeneous manifold $\mathcal{M}$. We define $\lambda_*\colon \mathfrak{g} \to \mathfrak{X}(\mathcal{M})$ by

$$(\lambda_* v)(y) = \frac{\mathrm{d}}{\mathrm{d}t}\lambda(tv, y)\bigg|_{t=0} \tag{2.15}$$

where $v \in \mathfrak{g}$ and $y \in \mathcal{M}$. We remark that at a fixed point $y_0 \in \mathcal{M}$ this locally corresponds to the definition of the pushforward (2.5) of $\lambda_{y_0} \colon \mathfrak{g} \to \mathcal{M}$ with $\lambda_{y_0}(v) = \lambda(v, y_0)$ since $T\mathfrak{g} \cong \mathfrak{g}$ and $(\lambda_{y_0})_* v \in T_{y_0}\mathcal{M}$.

The generic representation of ODEs on manifolds

Then, assuming there is a smooth function $f \colon \mathcal{M} \to \mathfrak{g}$, an ODE on a homogeneous manifold can be represented by

$$\dot{y} = (\lambda_* f(y))\,(y). \tag{2.16}$$

Such a representation can always be found on homogeneous manifolds and at least locally on other manifolds as the linear span of the vector fields $\frac{\partial}{\partial x^1}, \dots, \frac{\partial}{\partial x^n}$ creates a transitive algebra action on $\mathcal{M}$, see [58].

Theorem 2.6. *On a given manifold $\mathcal{M}$ assume that there is a (left) Lie algebra action $\lambda \colon \mathfrak{g} \times \mathcal{M} \to \mathcal{M}$ and a function $f \colon \mathcal{M} \to \mathfrak{g}$ such that an ODE for $y(t) \in \mathcal{M}$ is represented by (2.16) with the initial value $y(0) = y_0$. For sufficiently small t, the solution is $y(t) = \lambda\left(\Omega(t), y_0\right)$, where $\Omega(t) \in \mathfrak{g}$ satisfies the ODE*

$$\dot{\Omega} = \operatorname{dexp}_\Omega^{-1}\left(f\big(\lambda(\Omega, y_0)\big)\right), \quad \Omega(0) = \mathbf{0}. \tag{2.17}$$

For a fixed $y_0 \in \mathcal{M}$, let $\lambda_{y_0}(\Omega) = \lambda(\Omega, y_0)$ and similarly $\Lambda_{y_0}\big(\exp(\Omega)\big) = \Lambda\big(\exp(\Omega), y_0\big)$. This theorem can be proven by showing that the vector fields in (2.16) and (2.17) are λ_{y_0}-*related* (see [58, Lemma 8]).

We would like to give an intuition on how a proof would look like based on the notions discussed in the previous sections. The solution curve $y(t) = \lambda_{y_0}\big(\Omega(t)\big) = \Lambda_{y_0}\big(\exp(\Omega(t))\big)$ is a smooth curve on $\mathcal{M}$. On the one hand, we are locally mapping the Lie algebra $\mathfrak{g}$ onto tangent spaces of $\mathcal{M}$ via the Lie algebra action λ_{y_0}. Therefore, we can use the formula (2.3) to transform a tangent vector described in $\mathfrak{g}$ to a tangent vector described in $T_{y_0}\mathcal{M}$. Let $v^\Omega \coloneqq \operatorname{dexp}_\Omega^{-1}\left(f\big(\lambda(\Omega, y_0)\big)\right)$ denote the right hand side of (2.17), then we compute the directional derivative of λ_{y_0} along v^Ω to change the coordinates and get

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\Lambda_{y_0}\big(\exp(\Omega + tv^\Omega)\big)\bigg|_{t=0} &= \Lambda'_{y_0}\big(\exp(\Omega)\big)\,\frac{\mathrm{d}}{\mathrm{d}t}\exp(\Omega + tv^\Omega)\bigg|_{t=0} \\
&= \Lambda'_{y_0}\big(\exp(\Omega)\big)\operatorname{dexp}_\Omega(v^\Omega)\exp(\Omega) \\
&= \Lambda'_{y_0}\big(\exp(\Omega)\big) f\big(\lambda(\Omega, y_0)\big)\exp(\Omega).
\end{aligned} \tag{2.18}$$

On the other hand, we can compute the right hand side of (2.16) using the properties of

the Lie group action to obtain

$$\begin{aligned}\Big(\lambda_* f\big(\lambda_{y_0}(\Omega)\big)\Big)\big(\lambda_{y_0}(\Omega)\big) &= \frac{\mathrm{d}}{\mathrm{d}t}\Lambda\left(\exp\left(tf\big(\lambda_{y_0}(\Omega)\big)\right), \Lambda\big(\exp(\Omega), y_0\big)\right)\Big|_{t=0} \\ &= \frac{\mathrm{d}}{\mathrm{d}t}\Lambda\left(\exp\left(tf\big(\lambda_{y_0}(\Omega)\big)\right)\exp(\Omega), y_0\right)\Big|_{t=0} \\ &= \Lambda'_{y_0}\big(\exp(\Omega)\big) f\big(\lambda_{y_0}(\Omega)\big)\exp(\Omega).\end{aligned}$$

We conclude that both approaches give the same result.

To sum up, we can transform a vector field on the manifold $\mathcal{M}$ into a vector field in the Lie algebra $\mathfrak{g}$ and solve the ODE (2.16) on $\mathcal{M}$ by locally solving the ODE (2.17) in $\mathfrak{g}$. The ODE (2.17) might look more complicated but has the benefit that it is an ODE in a vector space such that well-known ODE solvers like *Runge-Kutta methods* can be used.

Runge-Kutta schemes

We remind the reader that for the numerical approximation of the initial value problem

$$\dot{y} = F(t, y), \quad y(t_0) = y_0,$$

with $y\colon \mathbb{R} \to \mathbb{R}^n$ and $F\colon \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ an s-stage Runge-Kutta method is given by

$$\begin{aligned} k_i &= F\big(t_0 + c_i h, y_0 + h\sum_{j=1}^{s} a_{ij} k_j\big), \quad i = 1, \dots, s, \\ y_1 &= y_0 + h\sum_{i=1}^{s} b_i k_i, \end{aligned}$$

where h is an uniform step size between two successive discrete time points t_ℓ and $t_{\ell+1}$, $\ell = 0, 1, \dots$. The coefficients b_i, a_{ij} are real numbers with $c_i = \sum_{j=1}^{s} a_{ij}$, which can be arranged in a Butcher tableau:

c_1	a_{11}	$\dots$	a_{1s}
$\vdots$	$\vdots$	$\ddots$	$\vdots$
c_s	a_{s1}	$\dots$	a_{ss}
	b_1	$\dots$	b_s

Table 2.1: Butcher Tableau

A Runge-Kutta method is said to have *convergence order* p if

$$y_1 - y(t_0 + h) = \mathcal{O}(h^{p+1}) \quad \text{for } h \to 0,$$

i.e. if the Taylor series for y_1 and for the exact solution $y(t_0 + h)$ coincide up to the term h^p, see e.g. [26, p. 134]. The most simple Runge-Kutta method with $s = 1$ is the *explicit Euler scheme* of convergence order $p = 1$ with coefficients:

0	0
	1

Table 2.2: Butcher Tableau of the explicit Euler scheme

We notice that we introduced Runge-Kutta methods for a nonautonomous initial value problem although we have regarded autonomous ODEs on manifolds for simplicity before. However, it is straightforward to extend the results for ODEs on manifolds to the nonautonomous case by considering $f\colon \mathbb{R} \times \mathcal{M} \to \mathfrak{g}$, $(t, y) \mapsto f(t, y)$ in (2.16) as done in [58].

Runge-Kutta–Munthe-Kaas schemes

Now, we apply a Runge-Kutta method to (2.17) and denote a truncated approximation of (2.9) by

$$\operatorname{dexpinv}(\Omega, H, q) = \sum_{k=0}^{q} \frac{B_k}{k!} \operatorname{ad}_\Omega^k(H) \tag{2.19}$$

to get the following *Runge-Kutta–Munthe-Kaas* (RKMK) scheme for solving the ODE (2.16) on a manifold $\mathcal{M}$.

Algorithm 1 RKMK

for $\ell = 0, 1, \ldots, L-1$ **do**
  for $i = 1, 2, \ldots, s$ **do**
    $\bar{\Omega}_i = h \sum_{j=1}^{s} a_{ij} k_j$
    $k_i = \operatorname{dexpinv}\big(\bar{\Omega}_i, f\big(\lambda(\bar{\Omega}_i, y_\ell)\big), q\big)$
  end for
  $\Omega_1 = h \sum_{j=1}^{s} b_j k_j$
  $y_{\ell+1} = \lambda(\Omega_1, y_\ell)$
end for

A statement on the convergence of this algorithm is given in the following theorem from [25, p. 126].

Theorem 2.7. *If the Runge-Kutta method is of order p and if the truncation index in* (2.19) *satisfies $q \geq p-2$, then the method of Algorithm 1 is of order p.*

In other words, RKMK schemes inherit the convergence order p of the underlying Runge-Kutta method, where we do not have to consider the infinite sum in (2.10). The number of summands in (2.19) that have to be evaluated can be chosen correspondingly to the desired convergence order p.

2.2.2 Examples

Now, let us consider some exemplary manifolds with corresponding Lie algebra actions and finally solve our motivating example of the dynamics of a rigid body problem. The following and more examples can be found in [30], [58], [59].

Example 1

Setting $\mathcal{M} = \mathbb{R}^n$ and $\lambda(v, y) = v + y$ we see that Algorithm 1 reduces to the classical Runge-Kutta method since $(\lambda_* v)(y) = v$.

Example 2

Assume that there is a set of vector fields $E_1, \dots, E_n \in \mathfrak{X}(\mathcal{M})$, which span the tangent space $T_m\mathcal{M}$ at each point $m \in \mathcal{M}$. Such vector fields are called *frames* on $\mathcal{M}$ and are assumed to be 'easily' integrated such that an ODE on $\mathcal{M}$ can be written as

$$\dot{y} = \sum_{i=1}^{n} f_i(y) E_i,$$

where $f_i \colon \mathcal{M} \to \mathbb{R}$ are smooth. We set $\mathfrak{g}$ as the Lie subalgebra of $\mathfrak{X}(\mathcal{M})$ spanned by E_i, $i = 1, \dots, n$ and set the Lie algebra action $\lambda \colon \mathfrak{g} \times \mathcal{M} \to \mathcal{M}$ as the flow operator such that $y(t) = \lambda(tF, y_0)$ is the solution of

$$\dot{y} = F\big(y(t)\big), \quad y(0) = y_0,$$

for $F \in \mathfrak{g}$. Since the pushforward of the Lie algebra action is

$$(\lambda_* F)(y_0) = \frac{\mathrm{d}}{\mathrm{d}t} \lambda(tF, y_0) \bigg|_{t=0} = F\big(y(t)\big)\big|_{t=0} = F(y_0)$$

we see that the above ODE written in terms of rigid frames is of the form (2.16) by setting $f\colon \mathcal{M} \to \mathfrak{g}$ as $f(y) = \sum_{i=1}^{n} f_i(y)E_i$. For solving this ODE with the RKMK algorithm one has to follow a fixed flow in the Lie subalgebra generated by $\{E_i\}$, $i = 1, \dots, n$. This is the same setting as considered for the *Crouch-Grossman* schemes [16].

Example 3

Let $\mathcal{M}$ be the $(n-1)$-dimensional sphere,

$$S^{n-1} = \{y \in \mathbb{R}^n : y^\top y = 1\}.$$

Then we can use rotation matrices $R \in \mathrm{SO}(n)$ to move around the sphere, such that we have the transitive Lie group action $\Lambda\colon \mathrm{SO}(n) \times S^{n-1} \to S^{n-1}$ with $\Lambda(R, y) = Ry$. For the corresponding Lie algebra action $\lambda\colon \mathfrak{so}(n) \times S^{n-1} \to S^{n-1}$, $\lambda(v, y) = \exp(v)y$ we have $(\lambda_* v)(y) = vy$ such that an ODE on S^{n-1} can be described by

$$\dot{y} = f(y)y, \quad y(0) = y_0,$$

where $f\colon S^{n-1} \to \mathfrak{so}(n)$. Note that in this case the Lie algebra action could also be defined via the Cayley map (2.13).

We see that the rigid body equation (2.1) is of this form with $n = 3$, where $f\colon S^2 \to \mathfrak{so}(3)$ reads

$$y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \mapsto \begin{pmatrix} 0 & y_3/I_3 & -y_2/I_2 \\ -y_3/I_3 & 0 & y_1/I_1 \\ y_2/I_2 & -y_1/I_1 & 0 \end{pmatrix}$$

for given constants I_1, I_2 and I_3. The dynamics evolve on $\mathcal{M} = \mathcal{S} = S^2$ if the initial value is normalized such that $\|y_0\| = 1$. Applying the explicit Euler scheme we can truncate the series (2.19) at $q = 0$ (according to Theorem 2.7) and get the following RKMK scheme,

$$\begin{aligned} \Omega_1 &= hf(y_\ell), \\ y_{\ell+1} &= \exp(\Omega_1)y_\ell, \end{aligned}$$

for $\ell = 0, 1, \dots$, which is also called the *Lie-Euler scheme*. Higher order RKMK schemes can be constructed in the same way and are given explicitly for example in [30, Appendix A].

A more thorough investigation of Lie group methods applied to rigid body dynamics can be found in [14].

Chapter 3

Numerical Approximation of Stochastic Differential Equations on Manifolds

Motivation: The dynamics of a perturbed rigid body

Consider again a rotating rigid body such as a satellite with dynamics as given in (2.1). Now, assume that there is some perturbance caused by e.g. measurement uncertainties. We suppose that the perturbation is of Gaussian nature and can be added to the right hand side of (2.1) such that the dynamics of the rigid body is described by

$$d\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 0 & y_3/I_3 & -y_2/I_2 \\ -y_3/I_3 & 0 & y_1/I_1 \\ y_2/I_2 & -y_1/I_1 & 0 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} dt + \begin{pmatrix} 0 & y_3/J_3 & -y_2/J_2 \\ -y_3/J_3 & 0 & y_1/J_1 \\ y_2/J_2 & -y_1/J_1 & 0 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \circ dW_t \quad (3.1)$$

with some more moments of inertia $J_1,\ J_2,\ J_3$ and a scalar Wiener process W_t with $dW_t \sim \mathcal{N}(0, dt)$. This is a (Stratonovich) stochastic differential equation (SDE) on the sphere $\mathcal{S}$ defined in (2.2).

In this chapter we aim at understanding this SDE on $\mathcal{S}$ by building upon the theoretical results of the previous chapter. We expand the theory of the deterministic case to set up an algorithm for solving SDEs on manifolds. Therefore, we first deal with important

properties of Itô and Stratonovich SDEs in vector spaces in Section 3.1 and their strong approximation in Section 3.2. After that we focus on the formulation of SDEs on manifolds in Section 3.3 and analyse the stochastic Runge-Kutta–Munthe-Kaas (SRKMK) scheme in Section 3.4.

3.1 Stochastic Differential Equations in vector spaces

In this section we intend to introduce SDEs as ODEs subjected to random perturbation as done in the motivational example above but in a more formal and general way. For the depiction of SDEs we focus on the Itô calculus in Section 3.1.1 but also give a brief insight on the formulation according to Stratonovich in Section 3.1.2.

As we assume the reader to be familiar with stochastic processes and the theory of measure and integration we only summarize theoretical results gathered from [60] in order to introduce properties of SDEs which we will need later on for the formulation of SDEs on manifolds and for the proof of strong convergence of SRKMK schemes.

3.1.1 Itô Stochastic Differential Equations

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space. In equation (3.1) we introduced a random perturbation to the ODE (2.1) by using a *Wiener process* W_t, which we would like to characterize now. A Wiener process (or *Brownian motion*) W_t is a stochastic process with the following properties:

1. $W_0 = 0$ $\mathbb{P}$-almost surely (a.s.).
2. W_t has Gaussian increments: $W_t - W_s$ is normally distributed with mean 0 and variance $t - s$, i.e. $W_t - W_s \sim \mathcal{N}(0, t-s)$ for all $0 \leq s \leq t$.
3. W_t has independent increments: $W_t - W_s$ and $W_u - W_r$ are independent for all $0 \leq r < u < s < t$.

Now, let us define a stochastic integral corresponding to the Wiener process for functions $b(t,\omega)\colon [0,\infty) \times \Omega \to \mathbb{R}$ with the following properties:

(i) $(t,\omega) \mapsto b(t,\omega)$ is $\mathcal{B} \times \mathcal{F}$-measurable, where $\mathcal{B}$ denotes the σ-algebra on $[0,\infty)$,

(ii) $b(t,\omega)$ is $\mathcal{F}_t$-adapted,

(iii) $\mathbb{E}\big[\int_0^t b(s,\omega)^2 ds\big] < \infty$.

For the sake of a simple notation we drop the dependence on ω and keep the dependence on the time t as an index whenever it is more convenient or drop it as well when the dependence on time is clear.

Definition 3.1 (Itô integral)**.** *The Itô integral of a stochastic process b_t (that fulfills the conditions (i)–(iii) above) with respect to the Brownian motion W_t is given by*

$$\int_0^t b_s \, dW_s = \lim_{L\to\infty} \sum_{\ell=0}^{L-1} b_{t_\ell}(W_{t_{\ell+1}} - W_{t_\ell})$$

for partitions $\{0 = t_0 < t_1 < \cdots < t_L = t\}$ of the time interval $[0, t]$ with $\max_{0\le\ell\le L-1}\{|t_{\ell+1} - t_\ell|\} \to 0$ for $L \to \infty$.

An important feature of this stochastic integral is that it satisfies the *Itô isometry*,

$$\mathbb{E}\left[\left(\int_0^t b_s \, dW_s\right)^2\right] = \mathbb{E}\left[\int_0^t b_s^2 \, ds\right]. \tag{3.2}$$

see [60, p. 29]. To motivate the formulation of an SDE we consider it as an expansion of an ODE. Let us look at the ODE $\dot{x}(t) = a\big(t, x(t)\big)$ with the initial value $x(t_0) = x_0$ in integral form

$$x(t) = x_0 + \int_{t_0}^t a\big(s, x(s)\big) \, ds,$$

where we assume $a(\cdot,\cdot)\colon [0,\infty) \times \mathbb{R} \to \mathbb{R}$ to be a continuous function fulfilling a local Lipschitz condition. Adding a random perturbation to the right hand side in form of an Itô integral gives us an interpretation of the *Itô process*.

Definition 3.2 (Itô process)**.** *An Itô process is a stochastic process X_t of the form*

$$X_t = X_{t_0} + \int_{t_0}^t a(s, X_s) \, ds + \int_{t_0}^t b(s, X_s) \, dW_s,$$

where we assume the stochastic process $b(\cdot,\cdot)$ to have the same properties as above and the function $a(\cdot,\cdot)$ to be $\mathcal{F}_t$-adapted with $\int_{t_0}^t |a(s,x)| \, ds < \infty$ for all $t \ge 0$ $\mathbb{P}$-almost surely. A shorter notation is given in differential form as

$$dX_t = a(t, X_t) \, dt + b(t, X_t) \, dW_t, \tag{3.3}$$

which is called an Itô stochastic differential equation with drift $a(t, X_t) \, dt$ and diffusion part $b(t, X_t) \, dW_t$.

For a twice continuously differentiable function f, i.e. $f(t,x) \in C^2([0,\infty)\times\mathbb{R},\mathbb{R})$ and an Itô process given by (3.3) it holds that $Y_t = f(t,X_t)$ is again an Itô process with

$$dY_t = \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial x}dX_t + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(dX_t)^2,$$

where the coefficients are evaluated at (t,X_t) and $(dX_t)^2 = (dX_t)\cdot(dX_t)$ is computed according to the rules

$$dt\cdot dt = 0, \quad dt\cdot dW_t = dW_t\cdot dt = 0 \quad \text{and} \quad dW_t\cdot dW_t = dt,$$

such that we have

$$dY_t = \left(\frac{\partial f}{\partial t} + \frac{\partial f}{\partial x}a + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}b^2\right)dt + \frac{\partial f}{\partial x}b\,dW_t, \tag{3.4}$$

which is called the (one-dimensional) *Itô formula* [60, p. 44].

Next, we want to extend the results regarding the Itô process and the Itô formula to the multidimensional case. Assume that $W_t = (W_t^1,\dots,W_t^d)$ is a d-dimensional standard Brownian motion with respect to a filtration $(\mathcal{F}_t)_{t\geq 0}$ satisfying the usual conditions. We set the time interval $I = [t_0,T]$ for some $0 \leq t_0 < T < \infty$. Consider the n-dimensional Itô process

$$X_t = X_{t_0} + \int_{t_0}^t a(s,X_s)\,ds + \sum_{j=1}^d \int_{t_0}^t b^j(s,X_s)\,dW_s^j \tag{3.5}$$

for $t\in I$, where we now assume the components a^i and $b^{i,j}$ of the drift $a\colon I\times\mathbb{R}^n\to\mathbb{R}^n$ and the diffusion coefficient $b\colon I\times\mathbb{R}^n\to\mathbb{R}^{n\times d}$, respectively, to fulfill the conditions from Definition 3.2 correspondingly. By b^j we denote the j-th column of the $n\times d$-matrix function $b = b^{i,j}$ for $j = 1,\dots,d$.

Let $g(t,x) = \big(g^1(t,x),\dots,g^p(t,x)\big)$ be a C^2 map from $[0,\infty)\times\mathbb{R}^n$ into $\mathbb{R}^p$. Then the process $Y_t = g(t,X_t)$ with X_t given as in (3.5) is again an Itô process, whose k-th component reads

$$dY_t^k = \frac{\partial g^k}{\partial t}dt + \sum_{i=1}^n \frac{\partial g^k}{\partial x^i}dX_t^i + \frac{1}{2}\sum_{i,j=1}^n \frac{\partial^2 g^k}{\partial x^i\partial x^j}dX_t^i dX_t^j, \tag{3.6}$$

where the coefficients are again evaluated at (t,X_t), see [60, p. 49]. One can use similar Itô rules as above, i.e. $dW_t^i dW_t^j = \delta_{ij}dt$ and $dW_t^i dt = dt dW_t^i = 0$, where δ_{ij} denotes the Kronecker delta, to simplify the formula.

The following theorem from [60, p. 66] tells us under which conditions there exists an unique solution for (3.5).

Theorem 3.3. *Let $T > 0$ and $a(\cdot,\cdot)\colon [0,T] \times \mathbb{R}^n \to \mathbb{R}^n$, $b(\cdot,\cdot)\colon [0,T] \times \mathbb{R}^n \to \mathbb{R}^{n\times d}$ be measurable functions satisfying*

$$\|a(t,x) - a(t,y)\| + \|b(t,x) - b(t,y)\| \leq D\|x-y\|, \quad x,y \in \mathbb{R}^n, t \in [0,T] \tag{3.7}$$

for some constant D, and such that

$$\|a(t,x)\| + \|b(t,x)\| \leq C(1+\|x\|), \quad x \in \mathbb{R}^n, t \in [0,T] \tag{3.8}$$

for some constant C. Let Z be a random variable with $\mathbb{E}\left[\|Z\|^2\right] < \infty$. Then the SDE (3.5) with $t_0 = 0$ and $X_0 = Z$ has an unique solution X_t with $\mathbb{E}\Big[\int_0^T \|X_t\|^2 dt\Big] < \infty$.

The norm considered above is the Euclidean norm for processes in $\mathbb{R}^n$ and the Frobenius norm for processes in $\mathbb{R}^{n\times d}$, i.e. $\|A\|_F = \sqrt{\sum_{i=1}^n \sum_{j=1}^d |a_{ij}|^2}$ for an arbitrary matrix $A \in \mathbb{R}^{n\times d}$. We note that additional to a global Lipschitz condition (3.7), a linear growth bound (3.8), which in this context is also called *Itô condition* [24, p. 131], is a sufficient condition to show the existence and uniqueness of the solution.

3.1.2 Stratonovich Stochastic Differential Equations

The Itô stochastic integral in Definition 3.1 is constructed by evaluating the function b at the left-end point of the time interval $[t_\ell, t_{\ell+1}]$. Evaluating the time interval at the midpoint will give us the *Stratonovich integral* of b,

$$\int_0^t b_s \circ dW_s = \lim_{L\to\infty} \sum_{\ell=0}^{L-1} b_{\frac{1}{2}(t_\ell + t_{\ell+1})}(W_{t_{\ell+1}} - W_{t_\ell})$$

for partitions of the time interval $[0,t]$.

Similar to Itô processes we can define a stochastic process as an expansion of an ODE by assuming an additive perturbation in terms of a Stratonovich integral [69]. Let the functions $\underline{a}$ and b fulfill some regularity assumptions as above. A corresponding stochastic differential equation is given by

$$X_t = X_{t_0} + \int_{t_0}^t \underline{a}(s, X_s)\, ds + \int_{t_0}^t b(s, X_s) \circ dW_s$$

with the differential notation as

$$dX_t = \underline{a}(t, X_t)\, dt + b(t, X_t) \circ dW_t. \tag{3.9}$$

An advantage of SDEs based on the Stratonovich integral is that they transform using the change of variables as in the deterministic case. There is no Itô formula with second order derivatives needed as we can use the ordinary chain rule for the stochastic process $Y_t = f(t, X_t)$, i.e. $dY_t = \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial x}dX_t$, with dX_t given by (3.9) to obtain

$$dY_t = \left(\frac{\partial f}{\partial t} + \frac{\partial f}{\partial x}\underline{a}\right) dt + \frac{\partial f}{\partial x} b \circ dW_t, \tag{3.10}$$

where the coefficients are again evaluated at (t, X_t).

This makes Stratonovich SDEs the more natural choice to extend ODEs on manifolds such that they account for random perturbation since constructions based on the change of variables formula (2.3) can be applied in a straightforward manner. For this reason, we set up the SDE on the sphere $\mathcal{S}$ in our motivational example (3.1) as a Stratonovich SDE.

Nevertheless, Stratonovich integrals have the disadvantage that they are not *martingales* as Itô integrals are, see [60, p. 37] for more details. As a consequence, it is more difficult to use the Stratonovich integral for proofs, for example for the proof of Theorem 3.3. Additionally, the evaluation of the time interval at the midpoint gives the Stratonovich integral the character of "looking into the future" [60, p. 24]. This is an undesirable feature for the application in financial mathematics since stock prices for example are only known at the beginning of a given time interval. Therefore, literature concerning finance use solely the Itô integral.

However, it is always possible to switch from one notation to the other one to take advantage of the corresponding stochastic integral. In order to transform the Stratonovich SDE (3.9) to the Itô notation (3.3) one can modify the drift as follows

$$a(t, x) = \underline{a}(t, x) + \frac{1}{2}\frac{\partial b}{\partial x}(t, x) b(t, x), \tag{3.11}$$

see [36, p. 157], where we call the second summand the *Itô correction term*. For the transformation of Itô to Stratonovich SDEs the equation can be rearranged correspondingly.

Below, we will focus on the Itô notation since it appears to be a better fit for our purposes but we will also give corresponding remarks regarding the Stratonovich notation.

3.2 Strong Approximation of SDEs

As we have seen in the previous chapter numerical approximation schemes for ODEs in vector spaces can be used to solve ODEs on manifolds by applying them in a corresponding Lie algebra. For this reason, we consider the strong approximation of SDEs in vector spaces by means of Itô-Taylor expansions in Section 3.2.1 and Stochastic Runge-Kutta (SRK) methods in Section 3.2.2 to use them to approximate SDEs on manifolds later on.

But first let us concretise the definition of convergence in a strong sense according to [36], [67]. For the numerical approximation we discretize the time interval $I = [t_0, T]$ by $I_h = \{t_0, t_1, \dots, t_L\}$ with $t_0 < t_1 < \cdots < t_L = T$ with step sizes $h_\ell = t_{\ell+1} - t_\ell$ for $\ell = 0, 1, \dots, L-1$. Further, let $h = \max_{0 \le \ell < L} h_\ell$ denote the maximum step size.

An approximating process X^h is said to *converge in a strong sense with order* $\gamma > 0$ to the Itô process X_t if there exists a finite constant C and a $\delta_0 > 0$ such that

$$\mathbb{E}\big[\|X_T - X^h(T)\|\big] \le Ch^\gamma$$

for any time discretization with maximum step size $h \in (0, \delta_0)$. For further investigations we will use the mean-square convergence, which implies strong convergence since the absolute error can be estimated by the root mean-square error via the Lyapunov inequality:

$$\mathbb{E}\left[\|X_T - X^h(T)\|\right] \le \Big(\mathbb{E}\left[\|X_T - X^h(T)\|^2\right]\Big)^{1/2}.$$

Definition 3.4 (Mean-square Convergence)**.** *A sequence of approximation processes* $X^h = \big(X(t)\big)_{t \in I_h}$ *converges in the mean-square sense with order* γ *to the solution* X *of SDE* (3.5) *at time* T *if there exists a constant* $C > 0$ *and some* $\delta_0 > 0$ *such that for each* $h \in (0, \delta_0)$

$$\Big(\mathbb{E}\left[\|X_T - X^h(T)\|^2\right]\Big)^{1/2} \le Ch^\gamma. \tag{3.12}$$

To give a complete overview of notions regarding the convergence, we mention that there also exists the concept of *weak convergence*. An approximating process X^h *converges in the weak sense with order* $\beta > 0$ if there exists a finite constant C and a $\delta_0 > 0$ such that

$$\big|\mathbb{E}\big(g(X_T)\big) - \mathbb{E}\big(g(X^h(T))\big)\big| \le Ch^\beta \tag{3.13}$$

for any polynomial g and time discretization with maximum step size $h \in (0, \delta_0)$.

Below, we will limit our analysis of approximating processes to their convergence in a strong sense as it implies weak convergence. Nevertheless, an analysis of weak convergence is still meaningful since higher orders can usually be achieved, to name only one benefit.

3.2.1 Itô-Taylor expansions

Stochastic Taylor expansions can be viewed as an extension of Taylor expansions in the deterministic case. There are stochastic Taylor expansions according to the Itô and to the Stratonovich integral. For our purpose we limit the presentation here on Itô-Taylor expansions, where we use the notations from [36, Chapter 5]. The reader can learn about Stratonovich-Taylor expansions in [36, Section 5.6].

Let the row vector $\alpha = (j_1, j_2, \ldots, j_l)$ with $j_i \in \{0, 1, \ldots, d\}$ for $i = 1, 2, \ldots, l$ be a multi-index of length $l := l(\alpha)$ and let M be the set of all multi-indices α including a multi-index v of length zero, i.e. $l(v) = 0$. We denote by $-\alpha$ and $\alpha-$ the multi-index obtained by deleting the first and last component of α, respectively.

Multiple Itô integrals

Let ρ and τ be two stopping times with $t_0 \leq \rho(\omega) \leq \tau(\omega) \leq T$ almost surely. Then, a *multiple Itô integral* is defined recursively by

$$I_\alpha[f(\cdot)]_{\rho,\tau} := \begin{cases} f(\tau), & l = 0, \\ \int_\rho^\tau I_{\alpha-}[f(\cdot)]_{\rho,s}\, ds, & l \geq 1 \text{ and } j_l = 0, \\ \int_\rho^\tau I_{\alpha-}[f(\cdot)]_{\rho,s}\, dW_s^{j_l}, & l \geq 1 \text{ and } j_l \geq 1, \end{cases}$$

for adapted right continuous with left limits (càdlàg) stochastic processes $f = \{f(t), t \geq 0\}$ in $\mathcal{H}_\alpha$, which is defined as follows. The set $\mathcal{H}_v$ contains all the càdlàg processes that satisfy $|f(t,\omega)| < \infty$ almost surely. The next set $\mathcal{H}_{(0)}$ is the totality of all càdlàg processes with

$$\int_0^t |f(s,\omega)|\, ds < \infty$$

almost surely and $\mathcal{H}_{(1)}$ is defined as the set of all such processes with

$$\int_0^t |f(s,\omega)|^2\, ds < \infty$$

almost surely. For further multi-indices α with $l(\alpha) = 1$ we write $\mathcal{H}_{(j)} = \mathcal{H}_{(1)}$ for each $j \in \{2, \ldots, d\}$ if $d \geq 2$. If $l(\alpha) > 1$ then $\mathcal{H}_\alpha$ is defined as the set of all càdlàg processes that satisfy

$$I_{\alpha-}[f(\cdot)]_{\rho,\cdot} \in \mathcal{H}_{(j_l)}.$$

Some examples for multiple Itô integrals are the following:

$$\begin{aligned} I_{(0)}[f(\cdot)]_{0,t} &= \int_0^t f(s)\,ds, \\ I_{(1)}[f(\cdot)]_{0,t} &= \int_0^t f(s)\,dW_s^1, \\ I_{(1,1)}[f(\cdot)]_{0,t} &= \int_0^t \int_0^{s_2} f(s_1)\,dW_{s_1}^1 dW_{s_2}^1, \\ I_{(0,1)}[f(\cdot)]_{0,t} &= \int_0^t \int_0^{s_2} f(s_1)\,ds_1 dW_{s_2}^1. \end{aligned}$$

For simpler notation we write $I_\alpha = I_\alpha[1]_{0,t}$ and $W_t^0 = t$ for $\alpha \in M$ and $t \geq 0$.

Itô coefficient functions

Consider the differential operators

$$\begin{aligned} L^0 &= \frac{\partial}{\partial t} + \sum_{k=1}^n a^k \frac{\partial}{\partial x^k} + \frac{1}{2} \sum_{k,l=1}^n \sum_{j=1}^d b^{k,j} b^{l,j} \frac{\partial^2}{\partial x^k \partial x^l}, \\ L^j &= \sum_{k=1}^n b^{k,j} \frac{\partial}{\partial x^k} \quad \text{for } j \in \{1, \ldots, d\}. \end{aligned}$$

The *Itô coefficient function* is defined recursively by

$$f_\alpha = \begin{cases} f, & l = 0, \\ L^{j_1} f_{-\alpha}, & l \geq 1, \end{cases}$$

for $\alpha = (j_1, \ldots, j_l)$ and $f \in C^p(\mathbb{R}_+ \times \mathbb{R}^n, \mathbb{R})$ where $p = l(\alpha) + n(\alpha)$ with $n(\alpha)$ being the number of components of α which are equal to zero.

Setting $f(t, x) \equiv x$ in the one-dimensional case $n = d = 1$ we get for example

$$f_{(0)} = a, \quad f_{(1)} = b, \quad f_{(1,1)} = bb', \quad f_{(1,0)} = ba', \quad f_{(0,1)} = ab' + \frac{1}{2} b^2 b''.$$

Itô-Taylor expansions

A subset $\mathcal{A} \subset M$ is called a *hierarchical set* if $\mathcal{A}$ is nonempty, the multi-indices in $\mathcal{A}$ are uniformly bounded in length and if

$$-\alpha \in \mathcal{A} \quad \text{for each } \alpha \in \mathcal{A} \setminus \{v\}.$$

The *remainder set* $\mathcal{B}(\mathcal{A})$ of $\mathcal{A}$ is defined by

$$\mathcal{B}(\mathcal{A}) = \{\alpha \in M \setminus \mathcal{A} : -\alpha \in \mathcal{A}\}.$$

Theorem 3.5. *Let $\mathcal{A} \subset M$ be a hierarchical set, $f\colon \mathbb{R}_+ \times \mathbb{R}^n \to \mathbb{R}$ and let ρ and τ be two stopping times with $t_0 \leq \rho(\omega) \leq \tau(\omega) \leq T$ almost surely. Then the Itô-Taylor expansion*

$$f(\tau, X_\tau) = \sum_{\alpha \in \mathcal{A}} I_\alpha \left[f_\alpha(\rho, X_\rho)\right]_{\rho,\tau} + \sum_{\alpha \in \mathcal{B}(\mathcal{A})} I_\alpha \left[f_\alpha(\cdot, X_\cdot)\right]_{\rho,\tau}$$

holds, provided all of the derivatives of f, a and b and all of the multiple Itô integrals exist.

Strong Convergence of truncated Itô-Taylor expansions

We consider the *truncated Itô-Taylor expansion*

$$X_k(t) = \sum_{\alpha \in \Lambda_k} I_\alpha \left[f_\alpha(t_0, X_{t_0})\right]_{t_0,t} \tag{3.14}$$

for $t \in [t_0, T]$, $k = 0, 1, \ldots$ and $f(t, x) \equiv x$ and assume that the necessary derivatives and multiple integrals exist for all $\alpha \in \Lambda_k \cup \mathcal{B}(\Lambda_k)$ with $\Lambda_k = \{\alpha \in M : l(\alpha) + n(\alpha) \leq k\}$ and $\mathcal{B}(\Lambda_k) = \{\alpha \in M \setminus \Lambda_k : -\alpha \in \Lambda_k\}$.

The simplest Itô-Taylor approximation for (3.5) is the *Euler-Maruyama scheme* with strong order $\gamma = 0.5$, which reads

$$X_{\ell+1} = X_\ell + a\, h + \sum_{j=1}^{d} b^j \Delta W^j \tag{3.15}$$

with $\Delta W^j = W^j_{t_{\ell+1}} - W^j_{t_\ell} = I_{(j)}$. Note that contrary to above the index indicates the time step.

An approximation scheme converging with strong order $\gamma = 1.0$ is the *Milstein scheme*,

$$X_{\ell+1} = X_\ell + a\, h + \sum_{j=1}^{d} b^j \Delta W^j + \sum_{j_1, j_2 = 1}^{d} L^{j_1} b^{j_2} I_{(j_1, j_2)}. \tag{3.16}$$

For $n = d = 1$ an *Itô-Taylor scheme with* $\gamma = 1.5$ is given by

$$\begin{aligned} X_{\ell+1} = X_\ell + a\,h + b\,\Delta W + \frac{1}{2}bb'\left((\Delta W)^2 - h\right) \\ + a'b\,\Delta Z + \frac{1}{2}\left(aa' + \frac{1}{2}b^2a''\right)h^2 \\ + \left(ab' + \frac{1}{2}b^2b''\right)(h\,\Delta W - \Delta Z) \\ + \frac{1}{2}b\left(bb'' + (b')^2\right)\left(\frac{1}{3}(\Delta W)^2 - h\right)\Delta W \end{aligned} \tag{3.17}$$

where $\Delta Z = I_{(1,0)}$. The coefficients a and b and their derivatives in the three methods above are all evaluated at $t = t_\ell$.

The strong convergence order of other truncated Itô-Taylor expansions can be determined via the following theorem from [36, p. 206].

Theorem 3.6. *Suppose that* $f_\alpha(t_0, X_{t_0}) \in \mathcal{H}_\alpha$ *for all* $\alpha \in \Lambda_k$ *and that* $f_\alpha(\cdot, X_\cdot) \in \mathcal{H}_\alpha$ *with*

$$\sup_{t_0 \le t \le T} \mathbb{E}\left[|f_\alpha(t, X_t)|^2\right] \le C_1 C_2^{l(\alpha)+n(\alpha)} \left\lfloor \frac{1}{2}(l(\alpha) + n(\alpha)) \right\rfloor !$$

for all $\alpha \in \mathcal{B}(\Lambda_k)$. *Then*

$$\mathbb{E}\left[\|X_t - X_k(t)\|^2\right] \le C_3 \frac{(C_4(t - t_0))^{k+1}}{\left\lfloor \frac{1}{2}(k+1) \right\rfloor !}$$

for all $t \in [t_0, T]$, *so the truncated Itô-Taylor expansion* (3.14) *converges to the Itô process* X_t *in the mean-square sense.*

3.2.2 Stochastic Runge-Kutta methods

Numerical schemes for SDEs that are derived from Itô-Taylor expansions but do not require the computations of derivatives can be constructed by expanding deterministic Runge-Kutta methods to the stochastic case. However, coming up with a general formula for the stochastic counterpart of Runge-Kutta methods is challenging since different strategies were developed to obtain order conditions. While Burrage & Burrage generalized B-series to the stochastic case [9], Rößler used a coloured rooted tree analysis [66] to come up with order conditions for stochastic Runge-Kutta (SRK) methods. We present the explicit SRK schemes created by Rößler in [65], [67] but we would also like to point out that one could use different SRK methods for example the one developed in [10], [73],

[74].

For the strong numerical approximation of SDEs it is necessary to simulate multiple stochastic integrals. The Itô stochastic integrals

$$I_{(j),\ell} = \int_{t_\ell}^{t_{\ell+1}} dW_s^j, \quad I_{(j_1,j_2),\ell} = \int_{t_\ell}^{t_{\ell+1}} \int_{t_\ell}^{s_2} dW_{s_1}^{j_1} dW_{s_2}^{j_2}$$

for $t_\ell, t_{\ell+1} \in I_h$ and $1 \leq j_1, j_2 \leq d$ are simulated as follows:

- $I_{(j)} = I_{(j),\ell} \sim \mathcal{N}(0, h_\ell)$,
- $I_{(j,j)} = \frac{1}{2}(I_{(j)}^2 - h_\ell)$,
- $I_{(j,0)} = \frac{1}{2} h_\ell (I_{(j)} + \frac{1}{\sqrt{3}} \zeta_j)$ with $\zeta_j \sim \mathcal{N}(0, h_\ell)$ independent from $I_{(j)}$ for all $1 \leq j \leq d$,
- $I_{(0,j)} = h_\ell I_{(j)} - I_{(j,0)}$.

Further, let $I_{(1,1,1)} = \frac{1}{6}(I_{(1)}^3 - 3 I_{(0)} I_{(1)})$ be the approximation of

$$I_{(1,1,1),\ell} = \int_{t_\ell}^{t_{\ell+1}} \int_{t_\ell}^{s_3} \int_{t_\ell}^{s_2} dW_{s_1}^1 dW_{s_2}^1 dW_{s_3}^1.$$

We are interested in methods with higher strong order than the Euler-Maruyama scheme (3.15), which can be considered as a SRK method of strong order $\gamma = 0.5$. A strong order $\gamma = 1$ SRK method for the numerical approximation of (3.5) is given by

$$\begin{aligned} X_{\ell+1} = X_\ell + \sum_{i=1}^{s} \alpha_i a\big(t_\ell + c_i^{(0)} h_\ell, H_i^{(0)}\big) h_\ell \\ + \sum_{k=1}^{d} \sum_{i=1}^{s} \Big(\beta_i^{(1)} I_{(k)} + \beta_i^{(2)} \sqrt{h_\ell}\Big) b^k\big(t_\ell + c_i^{(1)} h_\ell, H_i^{(k)}\big) \end{aligned} \tag{3.18}$$

for $\ell = 0, 1, \ldots, L-1$ with stages

$$\begin{aligned} H_i^{(0)} &= X_\ell + \sum_{j=1}^{s} A_{ij}^{(0)} a\big(t_\ell + c_j^{(0)} h_\ell, H_j^{(0)}\big) h_\ell + \sum_{l=1}^{d} \sum_{j=1}^{s} B_{ij}^{(0)} b^l\big(t_\ell + c_j^{(1)} h_\ell, H_j^{(l)}\big) I_{(l)}, \\ H_i^{(k)} &= X_\ell + \sum_{j=1}^{s} A_{ij}^{(1)} a\big(t_\ell + c_j^{(0)} h_\ell, H_j^{(0)}\big) h_\ell + \sum_{l=1}^{d} \sum_{j=1}^{s} B_{ij}^{(1)} b^l\big(t_\ell + c_j^{(1)} h_\ell, H_j^{(l)}\big) \frac{I_{(l,k)}}{\sqrt{h_\ell}} \end{aligned}$$

for $i = 1, \ldots, s$ and $k = 1, \ldots, d$.

For SDEs with scalar noise, i.e. (3.5) with $d = 1$, an efficient SRK with strong order $\gamma = 1.5$ can be formulated as

$$X_{\ell+1} = X_\ell + \sum_{i=1}^{s} \alpha_i a\big(t_\ell + c_i^{(0)} h_\ell, H_i^{(0)}\big) h_\ell + \sum_{i=1}^{s} \left(\beta_i^{(1)} I_{(1)} + \beta_i^{(2)} \frac{I_{(1,1)}}{\sqrt{h_\ell}} + \beta_i^{(3)} \frac{I_{(1,0)}}{h_\ell} + \beta_i^{(4)} \frac{I_{(1,1,1)}}{h_\ell} \right) b\big(t_\ell + c_i^{(1)} h_\ell, H_i^{(1)}\big) \quad (3.19)$$

for $\ell = 0, 1, \ldots, L-1$ with stages

$$H_i^{(0)} = X_\ell + \sum_{j=1}^{s} A_{ij}^{(0)} a\big(t_\ell + c_j^{(0)} h_\ell, H_j^{(0)}\big) h_\ell + \sum_{j=1}^{s} B_{ij}^{(0)} b\big(t_\ell + c_j^{(1)} h_\ell, H_j^{(1)}\big) \frac{I_{(1,0)}}{h_\ell},$$
$$H_i^{(1)} = X_\ell + \sum_{j=1}^{s} A_{ij}^{(1)} a\big(t_\ell + c_j^{(0)} h_\ell, H_j^{(0)}\big) h_\ell + \sum_{j=1}^{s} B_{ij}^{(1)} b\big(t_\ell + c_j^{(1)} h_\ell, H_j^{(1)}\big) \sqrt{h_\ell}$$

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

The coefficients of the SRK methods can be arranged in an extended version of the Butcher tableau from Table 2.1:

$c^{(0)}$	$A^{(0)}$	$B^{(0)}$	
$c^{(1)}$	$A^{(1)}$	$B^{(1)}$	
	α	$\beta^{(1)}$	$\beta^{(2)}$
		$\beta^{(3)}$	$\beta^{(4)}$

Table 3.1: Extended Butcher tableau

Coefficients of SRK methods with strong order $\gamma = 1.0$ and order $\gamma = 1.5$ can be found in the Butcher tableaus presented in Table 3.2 and Table 3.3, respectively. The construction of higher strong order SRK schemes is computationally complex and are known to be inefficient since the evaluation of multiple stochastic integrals is costly.

As already mentioned a way-out may be offered by weak SRK methods if one is interested in the approximation of distributional characteristics of the solution of (3.5). Weak approximations do not require information on the driving Wiener process, such that random variables with distributions that are easy to simulate can be used. We refer to [66], [74] for the analysis of weak SRK schemes.

0									
0	0			0					
0	0	0		0	0				
0									
0	0			1					
0	0	0		−1	0				
	1	0	0	1	0	0	0	$\frac{1}{2}$	$-\frac{1}{2}$

Table 3.2: Coefficients of SRK method with strong order $\gamma = 1.0$ [67]

0											
$\frac{3}{4}$	$\frac{3}{4}$			$\frac{3}{2}$							
0	0	0		0	0						
0	0	0	0	0	0	0					
0											
$\frac{1}{9}$	$\frac{1}{9}$			$\frac{1}{3}$							
$-\frac{2}{9}$	$-\frac{5}{9}$	$\frac{1}{3}$		$-\frac{1}{3}$	1						
$\frac{1}{3}$	−1	$\frac{1}{3}$	1	1	−1	1					
	$\frac{1}{3}$	$\frac{2}{3}$	0	0	$\frac{13}{4}$	$-\frac{9}{4}$	$-\frac{9}{4}$	$\frac{9}{4}$	$-\frac{15}{4}$	$\frac{15}{4}$	$\frac{3}{4}$	$-\frac{3}{4}$
					$-\frac{9}{4}$	$\frac{9}{4}$	$\frac{9}{4}$	$-\frac{9}{4}$	6	−9	0	3

Table 3.3: Coefficients of SRK method with strong order $\gamma = 1.5$ [65]

3.3 Stochastic Differential Equations on manifolds

Having acquired the basics of SDEs in linear spaces we now bring this knowledge together with the notions of fiber bundles. We execute this by combining the contents reviewed in the second chapter about ODEs on manifolds with the contents of Section 3.1 in order to formulate SDEs on manifolds. Classical references on this topic include [17], [18], [29], which we identify as references for readers with a strong background in stochastic analysis. We choose to introduce SDEs on manifolds based on [22], [24] since these references build upon basics in differential geometry and add theoretical results of stochastic processes based on this background. This approach reflects our level of knowledge and the one we assume the reader to have. Therefore, we first formulate Itô SDEs on manifolds in terms of sections of the Itô bundle in Section 3.3.1, which is followed by a representation of SDEs on homogeneous manifolds in Section 3.3.2 based on the depiction in [44].

3.3.1 The Itô bundle

Let $\mathcal{M}$ be a smooth manifold of dimension n. The benefit of considering SDEs on manifolds based on the definition of Stratonovich is that they can be described as sections of the tangent bundle $T\mathcal{M}$ like ODEs on manifolds, see (2.4). Under the change of the coordinates $\varphi_{\beta\alpha}$ a Stratonovich process (3.9) on $\mathcal{M}$ transforms into

$$d\varphi_{\beta\alpha}(X_t) = \varphi'_{\beta\alpha}\big(\underline{a}(t, X_t)\, dt + b(t, X_t) \circ dW_t\big)$$

according to the transformation rule (3.10). This formula coincides with the transformation of a tangent vector (2.3), which makes working with Stratonovich SDEs more similar to the deterministic case.

For the definition of Itô SDEs on manifolds some more consideration has to be taken into account in order to present the Itô formula (3.4) correctly. An Itô SDE on $\mathcal{M}$ is a section of a different fiber bundle. We introduce this fiber bundle according to the Definition 2.4 by first presenting its structure group, where we use the notations from [24, Chapter 7].

Let $L(\mathbb{R}^d, \mathbb{R}^n)$ be the space of linear operators from $\mathbb{R}^d$ to $\mathbb{R}^n$ and $L^2(\mathbb{R}^n)$ be the set of bilinear mappings $\alpha\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$.

Definition 3.7. *The Itô group G_I is the set of pairs (B, β) where $B \in \mathrm{GL}(n)$ and $\beta \in L^2(\mathbb{R}^n)$ with the operation defined by*

$$(B, \beta) \cdot (C, \gamma) = \big(B \circ C, B \circ \gamma(\cdot, \cdot) + \beta\big(C(\cdot), C(\cdot)\big)\big).$$

Note that G_I is indeed a group with the unit element $(I, 0)$, where I is the unit operator and 0 the zero bilinear mapping, and inverse $(B, \beta)^{-1} = \big(B^{-1}, -B^{-1} \circ \beta\big(B^{-1}(\cdot), B^{-1}(\cdot)\big)\big)$, which can be verified by direct calculations.

We recall that the trace operator "tr" of a bilinear mapping $\Psi(\cdot, \cdot)$ is defined by

$$\operatorname{tr} \Psi = \sum_{i=1}^{n} \Psi(e_i, e_i),$$

where $e_1, \dots, e_n$ denotes an arbitrary orthonormal frame.

Definition 3.8. *The Itô bundle $I(\mathcal{M})$ over a manifold $\mathcal{M}$ is a fiber bundle with the standard fiber $F = \mathbb{R}^n \times L(\mathbb{R}^d, \mathbb{R}^n)$ and structure group G_I that acts on F from the left by*

$$(B, \beta) \cdot (X, A) = \left(BX + \frac{1}{2} \operatorname{tr} \beta\big(A(\cdot), A(\cdot)\big), B \circ A\right). \tag{3.20}$$

It follows that an element $(m_\alpha, (a_\alpha, b_\alpha))$ of the Itô bundle in the chart U_α is transformed to another chart U_β by

$$(m_\beta, (a_\beta, b_\beta)) = \left(\varphi_{\beta\alpha} m_\alpha, \left(\varphi'_{\beta\alpha} a_\alpha + \frac{1}{2} \operatorname{tr} \varphi''_{\beta\alpha}(b_\alpha, b_\alpha), \varphi'_{\beta\alpha} b_\alpha\right)\right). \tag{3.21}$$

Definition 3.9. *An Itô equation is a section of the Itô bundle $I(\mathcal{M})$.*

So assuming an Itô process (3.5) on the manifold $\mathcal{M}$ as a section of $I(\mathcal{M})$ it transforms under a coordinate change $\varphi_{\beta\alpha}$ into

$$d\varphi_{\beta\alpha}(X_t) = \varphi'_{\beta\alpha}\big(a(t, X_t)\, dt + b(t, X_t)\, dW_t\big) + \frac{1}{2} \operatorname{tr} \varphi''_{\beta\alpha}\big(b(t, X_t), b(t, X_t)\big)\, dt,$$

which obeys the Itô formula (3.6). We refer to [24, Theorem 7.20] for the existence of an unique solution of an Itô equation on $\mathcal{M}$.

There is a correspondence of the Itô equations defined as sections of $I(\mathcal{M})$ and Itô equations derived from the second order tangent bundle $\tau\mathcal{M}$ via a *local connector* $\Gamma_m(\cdot, \cdot)$ in the second tangent bundle $TT\mathcal{M}$ of $\mathcal{M}$, where $\tau\mathcal{M}$ and $TT\mathcal{M}$ are fiber bundles as mentioned in Section 2.1.3. These constructions seem to be beneficial when dealing with *Riemannian* manifolds. More details can be found in [24, Section 7.3].

Although some terminologies differ from that used here, the idea of defining Itô SDEs on manifolds as sections of a new fiber bundle can be dated back to Itô himself [31]. Based on this idea the Itô bundle was first constructed by Belopolskaya and Dalecky in [4].

Another definition of Itô SDEs on manifolds as 2-*jets* of smooth functions was developed by Armstrong and Brigo in [2], where an informal description of this approach is given by writing down a system of difference equations using a coordinate-free notation.

3.3.2 Stochastic Differential Equations on homogeneous manifolds

Let us now specify the description of an SDE on a homogeneous manifold. We recall that a homogeneous space is a manifold $\mathcal{M}$ with a transitive Lie group action $\Lambda\colon G \times \mathcal{M} \to \mathcal{M}$. Let $\lambda\colon \mathfrak{g} \times \mathcal{M} \to \mathcal{M}$ be the corresponding Lie algebra action and $f_j\colon \mathcal{M} \to \mathfrak{g}$ for $j = 0, 1, \ldots, d$. Based on the representation of an ODE on a homogeneous manifold (2.16) we assume a Stratonovich SDE on a homogeneous manifold $\mathcal{M}$ to be given by

$$dy = V_0(y)\, dt + \sum_{j=1}^{d} V_j(y) \circ dW_t^j, \quad y(0) = y_0,$$

where we assume $V_j \colon \mathcal{M} \to T\mathcal{M}$ to be vector fields that can be described by the transitive Lie algebra action,

$$V_j(y) = \big(\lambda_* f_j(y)\big)(y) = \frac{\mathrm{d}}{\mathrm{d}t}\lambda\big(tf_j(y), y\big)\Big|_{t=0} = \frac{\mathrm{d}}{\mathrm{d}t}\Lambda\Big(\exp\big(tf_j(y)\big), y\Big)\Big|_{t=0}$$

for $j = 0, 1, \ldots, d$, see the definition in (2.15). This description of an SDE on $\mathcal{M}$ corresponds to the one given in [44]. Note that we return to the autonomous notation of differential equations for simplicity. The extension to the nonautonomous case can be constructed straightforwardly by considering $f_j \colon \mathbb{R}_+ \times \mathcal{M} \to \mathfrak{g}$, $(t, y) \mapsto f_j(t, y)$ for $j = 0, 1, \ldots, d$ above.

The corresponding Itô notation of this SDE looks as follows,

$$dy = \left(V_0(y) + \frac{1}{2}\sum_{j=1}^{d} \nabla_{V_j} V_j(y)\right) dt + \sum_{j=1}^{d} V_j(y)\, dW_t^j, \quad y(0) = y_0, \tag{3.22}$$

which is evident by applying the Itô/Stratonovich conversion formula (3.11). By $\nabla_Y X$ for two vector fields $X, Y \in \mathfrak{X}(\mathcal{M})$ we denote the derivative of X along the vector field Y to express the Itô correction term. In some cases it might be possible to formulate the Itô version (3.22) without the Itô correction term, but this will lead to some *unnatural* geometric restrictions on the drift coefficient of the SDE. An example for this will be given below.

As done in the previous chapter we formulate a differential equation in the Lie algebra $\mathfrak{g}$ that corresponds to the representation (3.22) on the manifold $\mathcal{M}$.

Theorem 3.10. *Let $\lambda \colon \mathfrak{g} \times \mathcal{M} \to \mathcal{M}$ be a Lie algebra action and $f_j \colon \mathcal{M} \to \mathfrak{g}$ for $j = 0, 1, \ldots, d$. Assume that an Itô SDE for $y(t) \in \mathcal{M}$ is given by (3.22). For t small enough and up to a stopping time T_* the solution of this SDE is given by $y(t) = \lambda\big(\Omega(t), y_0\big)$ where $\Omega(t) \in \mathfrak{g}$ satisfies*

$$d\Omega = \operatorname{dexp}_\Omega^{-1}\Big(f_0\big(\lambda(\Omega, y_0)\big)\Big)\, dt + \sum_{j=1}^{d} \operatorname{dexp}_\Omega^{-1}\Big(f_j\big(\lambda(\Omega, y_0)\big)\Big)\, dW_t^j,\ \Omega(0) = \mathbf{0}, \tag{3.23}$$

on the time interval $[0, T_*)$.

The theorem can be proved by applying Itô's formula to $y(t) = \lambda_{y_0}\big(\Omega(t)\big)$, where $\lambda_{y_0}\big(\Omega(t)\big) = \lambda(\Omega(t), y_0)$ for $y_0 \in \mathcal{M}$ fixed, and identifying the result with (3.22). Here we will only report the first and second order directional derivative of the Lie algebra action, which are needed for the transformation according to (3.21), and carry out the

computations for a specific manifold $\mathcal{M} = \mathrm{Sym}(n)$ below. For a more general proof we refer to [44].

Let v_j^Ω denote the coefficients of (3.23), i.e. $v_j^\Omega = \mathrm{dexp}_\Omega^{-1}\Big(f_j\big(\lambda_{y_0}(\Omega)\big)\Big)$ for $j = 0, 1, \ldots, d$. For the directional derivatives we have

$$\begin{aligned}
\lambda'_{y_0}(\Omega)v_j^\Omega &= \left(\frac{\mathrm{d}}{\mathrm{d}\Omega}\lambda_{y_0}(\Omega)\right) v_j^\Omega = \Lambda'_{y_0}\big(\exp(\Omega)\big) f_j\big(\lambda_{y_0}(\Omega)\big)\exp(\Omega),\\
\lambda''_{y_0}(\Omega)(v_j^\Omega, v_j^\Omega) &= \Lambda''_{y_0}\big(\exp(\Omega)\big)\Big(f_j\big(\lambda_{y_0}(\Omega)\big)\exp(\Omega), f_j\big(\lambda_{y_0}(\Omega)\big)\exp(\Omega)\Big)\\
&\quad + \Lambda'_{y_0}\big(\exp(\Omega)\big) f_j'\big(\lambda_{y_0}(\Omega)\big)\Lambda'_{y_0}\big(\exp(\Omega)\big) f_j\big(\lambda_{y_0}(\Omega)\big)\exp(\Omega)\exp(\Omega)\\
&\quad + \Lambda'_{y_0}\big(\exp(\Omega)\big) f_j\big(\lambda_{y_0}(\Omega)\big) f_j\big(\lambda_{y_0}(\Omega)\big)\exp(\Omega),
\end{aligned} \tag{3.24}$$

where the second directional derivative only needs to be considered for $j = 1, \ldots, d$. See (2.18) for the computation of the first directional derivative and the Appendix for the derivation of the second.

Examples

Before advancing to the numerical approximation we provide examples of the representation of the SDE (3.22) for some specific manifolds. One can also reconsider the exemplary manifolds discussed before in Section 2.2.2 and extend the given ODEs accordingly to the stochastic case as done in our motivational example for the rigid body problem. Here we examine different manifolds which we will also reencounter in the next chapter. The first two examples are extensions of examples found in [58], [59] and the third example can also be found in [44].

We remind the reader that Itô SDEs on manifolds that are not globally homogeneous spaces can always at least locally be expressed by (3.22) since a transitive Lie algebra action can be constructed by means of the vector fields $\frac{\partial}{\partial x^1}, \ldots, \frac{\partial}{\partial x^n}$ corresponding to some local coordinates $(x_1, \ldots, x_n)$ and using the transformation (3.21) from one chart to another.

Example 1 If the considered manifold $\mathcal{M}$ is a matrix Lie group G we can choose a right Lie group action $\Lambda\colon \mathcal{M} \times G \to \mathcal{M}$, which is defined analogously to Definition 2.2, and set it as $\Lambda(m, g) = mg$. A corresponding right Lie algebra action $\lambda\colon \mathcal{M} \times \mathfrak{g} \to \mathfrak{g}$ may be given by $\lambda(y_0, \Omega) = y_0 \exp(\Omega)$ such that we obtain the left-invariant vector field $\big(\lambda_* f_j(y)\big)(y) = y f_j(y)$ for $j = 0, 1, \ldots, d$. An SDE with a solution evolving on $\mathcal{M} = G$

can then be formulated as

$$dy = \left(yf_0(y) + \frac{1}{2} \sum_{j=1}^{d} \left(yf_j^2(y)y + yf'(y)yf(y) \right) \right) dt + \sum_{j=1}^{d} yf_j(y)\, dW_t^j. \tag{3.25}$$

We will consider this representation for an Itô SDE on the Lie group of stochastic matrices in Section 4.2.2, where we simulate rating transition matrices. Left-invariant vector fields are chosen in order to preserve the *Chapman-Kolmogorov equation* for Markov chains.

Assuming that Itô SDEs on manifolds $\mathcal{M}$ are instead formulated without the Itô correction term, which in the case of (3.25) is given by

$$dy = yf_0(y)\, dt + \sum_{j=1}^{d} yf_j(y)\, dW_t^j,$$

leads to restrictions on the choice of $f_0 \colon \mathcal{M} \to \mathfrak{g}$, $y \mapsto f_0(y)$. Let us visualize this by setting $\mathcal{M} = G = \mathrm{SO}(n)$ which means that the solution should satisfy $y^\top y = I$ and $d(y^\top y) = dI = \mathbf{0}$. Applying Itô's formula we get

$$d(y^\top y) = \left(f_0(y)^\top + f_0(y) + \sum_{j=1}^{d} f_j(y)^\top f_j(y) \right) dt + \sum_{j=1}^{d} \left(f_j(y)^\top + f_j(y) \right) dW_t^j.$$

The identification of this result with $\mathbf{0}$ gives us the conditions that $f_j(y)$ must be skew-symmetric, $f_j(y)^\top + f_j(y) = \mathbf{0}$, for $j = 1, \ldots, d$ and that

$$f_0(y)^\top + f_0(y) = \sum_{j=1}^{d} f_j^2(y), \tag{3.26}$$

whereas considering the SDE representation (3.25) on $\mathrm{SO}(n)$ leads to the assumption that $f_j(y) \in \mathfrak{so}(n)$ for all indices $j = 0, 1, \ldots, d$. The condition (3.26) can be used to one's advantage as done e.g. in [51] for the approximation of correlation matrices. Nevertheless, we consider this small example as a motivation to formulate Itô SDEs on manifolds including the Itô correction term from here on.

Example 2 On the manifold of symmetric matrices $\mathcal{M} = \mathrm{Sym}(n)$ we can choose a Lie group action based on $G = \mathrm{SO}(n)$ with the corresponding Lie algebra $\mathfrak{g} = \mathfrak{so}(n)$. The corresponding Lie algebra action can be set as $\lambda_{y_0}(\Omega) = \exp(\Omega) y_0 \exp(-\Omega)$. Computing

the directional derivatives, we get

$$\begin{aligned}\lambda'_{y_0}(\Omega)v_j^\Omega &= [f_j(y), y]\,,\\ \lambda''_{y_0}(\Omega)(v_j^\Omega, v_j^\Omega) &= \big[f_j'(y)\,[f_j(y), y]\,, y\big] + \big[f_j(y), [f_j(y), y]\,\big],\end{aligned}$$

which coincides with the result obtained by directly using (3.24) and the representation in (3.22) since we have the vector field

$$\begin{aligned}V_j(y) = \big(\lambda_* f_j(y)\big)(y) &= \frac{\mathrm{d}}{\mathrm{d}t}\lambda\big(tf_j(y), y\big)\Big|_{t=0}\\ &= \frac{\mathrm{d}}{\mathrm{d}t}\exp\big(tf_j(y)\big)y\exp\big(-tf_j(y)\big)\Big|_{t=0}\\ &= f_j(y)y - yf_j(y) = [f_j(y), y]\end{aligned}$$

and for the differentiation of the vector field along itself

$$\begin{aligned}\nabla_{V_j}V_j(y) &= \left(\frac{\mathrm{d}}{\mathrm{d}y}\,[f_j(y), y]\right)[f_j(y), y]\\ &= \left[\left(\frac{\mathrm{d}}{\mathrm{d}y}f_j(y)\right)[f_j(y), y]\,, y\right] + \left[f_j(y), \left(\frac{\mathrm{d}}{\mathrm{d}y}y\right)[f_j(y), y]\right]\\ &= \big[f_j'(y)\,[f_j(y), y]\,, y\big] + \big[f_j(y), [f_j(y), y]\,\big]\end{aligned}$$

for $j = 0, 1\ldots, d$. The SDE (3.22) deployed with these coefficients preserves the eigenvalues of y_0, i.e. we have an *isospectral flow*. More specifically, if the eigenvalues of y_0 are nonnegative, then the solution of this SDE produces symmetric and positive semi-definite matrices. This is a property that comes in handy if one is interested in modelling covariance and correlation matrices, see Section 4.2.1.

Example 3 Consider the Lie group of rigid body motions, the special Euclidean group $\mathrm{SE}(3) \cong \mathrm{SO}(3) \times \mathbb{R}^3$, that consists of the rotation motions $R \in \mathrm{SO}(3)$ and the translations $r \in \mathbb{R}^3$ of a rigid body stacked into a 4×4-matrix. An exemplary representative of this Lie group is given by

$$g = \begin{pmatrix} R & r\\ O^\top & 1\end{pmatrix}$$

with $O = (0, 0, 0)^\top$. The corresponding Lie algebra $\mathfrak{se}(3) \cong \mathfrak{so}(3) \times \mathbb{R}^3$ consists of matrices that look like

$$v = \begin{pmatrix} \hat{w} & u\\ O^\top & 0\end{pmatrix}$$

where we use the hat map (2.8). Due to this isomorphism between $\mathfrak{so}(3)$ and $\mathbb{R}^3$ we can use the shorthand notation $v = (w, u)$ with $w, u \in \mathbb{R}^3$ to represent an element of $\mathfrak{se}(3)$. Now, we want to formulate an SDE on the manifold $\mathcal{M} = \mathfrak{se}(3)^*$, the dual space of the Lie algebra $\mathfrak{se}(3)$. A Lie group action with $G = \mathrm{SE}(3)$ is specified by

$$\Lambda\colon \mathrm{SE}(3) \times \mathfrak{se}(3)^* \to \mathfrak{se}(3)^*, \quad \Lambda(g, y) = \big(R\pi + r \times (R\rho), R\rho\big)$$

for $y = (\pi, \rho) \in \mathfrak{se}(3)^*$ and $g = (R, r) \in \mathrm{SE}(3)$ using a similar shorthand notation as above. To compute the vector fields in (3.22) with $f_j\colon \mathfrak{se}(3)^* \to \mathfrak{se}(3)$, $f_j(y) = \big(w_j(y), u_j(y)\big)$, we have

$$V_j(y) = \big(\lambda_* f_j(y)\big)(y) = \big(\pi \times w_j(y) + \rho \times u_j(y), \rho \times w_j(y)\big)$$

for $j = 0, 1, \ldots, d$. This will be used to compute the dynamics of an autonomous underwater vehicle in Section 4.1.2.

3.4 Stochastic Runge-Kutta–Munthe-Kaas schemes

Being familiar with different versions of the SDE (3.22) on a homogeneous manifold $\mathcal{M}$, let us devise a procedure for solving it numerically such that the approximations stay on $\mathcal{M}$. Such kind of procedure, where a numerical approximation scheme is applied to (3.23), was first formulated in [44] based on the Algorithm 1 by Munthe-Kaas.

Algorithm 3.11. *Let $\lambda\colon \mathfrak{g} \times \mathcal{M} \to \mathcal{M}$ be a transitive Lie algebra action on $\mathcal{M}$. Divide the time interval $[0, T]$ uniformly into L subintervals $[t_\ell, t_{\ell+1}]$, $\ell = 0, 1, \ldots, L-1$ and define an uniform step size $h = t_{\ell+1} - t_\ell$. Starting from $y_0 \in \mathcal{M}$ at $t_0 = 0$ repeat the following steps until $t_{\ell+1} = T$.*

1. ***Initialization step:*** *Let y_ℓ be the approximation of y_t at time $t = t_\ell$.*
2. ***Numerical method step:*** *Compute an approximation of the solution of (3.23) after one time step, $\Omega_{t_1} = \Omega_{t_1 - t_0} = \Omega_h$, by applying an Itô-Taylor scheme or a stochastic Runge-Kutta method to the SDE in the Lie algebra $\mathfrak{g}$ and denote the obtained approximation of Ω_h by Ω_1.*
3. ***Action step:*** *Advance on the manifold $\mathcal{M}$ using the Lie algebra action as $y_{\ell+1} = \lambda(\Omega_1, y_\ell)$ to get a numerical solution of (3.22).*

Denote by $v_j\colon \mathfrak{g} \to \mathfrak{g}$ the coefficients of (3.23), $v_j(\Omega) = \exp_\Omega^{-1}\Big(f_j\big(\lambda_{y_0}(\Omega)\big)\Big)$ for $j = 0, 1, \ldots, d$. The first numerical method that comes into one's mind is probably the

Euler-Maruyama scheme (3.15) for the second step of this algorithm such that

$$\Omega_1 = \Omega_0 + v_0(\Omega_0)\,h + \sum_{j=1}^{d} v_j(\Omega_0)\,\Delta W^j = f_0\big(\lambda(\mathbf{0}, y_\ell)\big)\,h + \sum_{j=1}^{d} f_j\big(\lambda(\mathbf{0}, y_\ell)\big)\,\Delta W^j,$$

where the evaluation at $\Omega_0 = \mathbf{0}$ simplifies the computations immensely. The application of this method to manifolds $\mathcal{M}$ that are matrix Lie groups with $y_{\ell+1} = \lambda(\Omega_1, y_\ell) = y_\ell \exp(\Omega_1)$ was intensively analysed in [46], [47], [64], where it was called the *geometric Euler-Maruyama* scheme.

Applying the Milstein scheme (3.16) with strong order $\gamma = 1$ or the Itô-Taylor scheme (3.17) with strong order $\gamma = 1.5$ as numerical methods in the second step of this algorithm requires the computation of the directional derivatives of v_j, $j = 0, 1, \ldots, d$. See the Appendix for the derivations of these derivatives.

This motivates the implementation of stochastic Runge-Kutta (SRK) methods in the second step of the algorithm since the computation of derivatives is avoided.

We adapt the notations of Rößler's explicit s-stage SRK scheme (3.19) with coefficients given in Table 3.1 such that we get a representation as for the RKMK schemes in Algorithm 1 for ODEs on manifolds. A stochastic RKMK (SRKMK) method for $d = 1$ as a specification of the procedure above is given in Algorithm 2.

Algorithm 2 SRKMK

for $\ell = 0, 1, \ldots, L-1$ **do**
  for $i = 1, 2, \ldots, s$ **do**
    $\bar{\Omega}_i = \sum_{j=1}^{i-1} A_{ij}^{(0)} v_0(\bar{\Omega}_j)h + \sum_{j=1}^{i-1} B_{ij}^{(0)} v_1(\tilde{\Omega}_j)\frac{I_{(1,0)}}{h}$
    $\tilde{\Omega}_i = \sum_{j=1}^{i-1} A_{ij}^{(1)} v_0(\bar{\Omega}_j)h + \sum_{j=1}^{i-1} B_{ij}^{(1)} v_1(\tilde{\Omega}_j)\sqrt{h}$
    $v_0(\bar{\Omega}_i) = \mathrm{dexpinv}\,\big(\bar{\Omega}_i, f_0\big(\lambda(\bar{\Omega}_i, y_\ell)\big), q\big)$
    $v_1(\tilde{\Omega}_i) = \mathrm{dexpinv}\,\big(\tilde{\Omega}_i, f_1\big(\lambda(\tilde{\Omega}_i, y_\ell)\big), q\big)$
  end for
  $\Omega_1 = \sum_{i=1}^{s} \alpha_i v_0(\bar{\Omega}_i)h + \sum_{i=1}^{s} \beta_i^{(1)} v_1(\tilde{\Omega}_i) I_{(1)} + \sum_{i=1}^{s} \beta_i^{(2)} v_1(\tilde{\Omega}_i)\frac{I_{(1,1)}}{h} + \sum_{i=1}^{s} \beta_i^{(3)} v_1(\tilde{\Omega}_i)\frac{I_{(1,0)}}{h} + \sum_{i=1}^{s} \beta_i^{(4)} v_1(\tilde{\Omega}_i)\frac{I_{(1,1,1)}}{h}$
  $y_{\ell+1} = \lambda(\Omega_1, y_\ell)$
end for

The linear operator $\mathrm{dexpinv}\colon \mathfrak{g} \to \mathfrak{g}$ gives us truncated versions of the coefficients in (3.23) and is defined as in (2.19).

3.4.1 Strong Convergence of SRKMK schemes

Since the SRKMK scheme in Algorithm 2 uses the map dexpinv: $\mathfrak{g} \to \mathfrak{g}$, this raises again the question (as in the deterministic case, see Theorem 2.7) of how to choose the truncation index q such that the SRKMK scheme inherits the strong convergence order γ of the underlying SRK method. The following theorem is based on results from [53]–[55] and gives us an answer to this question.

Theorem 3.12 (Strong convergence of SRKMK schemes)**.** *Let q denote the truncation index in* (2.19)*, and let the stochastic Runge-Kutta scheme applied to the SDE* (3.23) *in the Lie algebra $\mathfrak{g}$ be of strong order γ. Furthermore, assume that $(f_j \circ \lambda_{y_0})\colon \mathfrak{g} \to \mathfrak{g}$ fulfills a linear growth condition, i.e.*

$$\|(f_j \circ \lambda_{y_0})(\Omega)\|_F \leq a_j + b_j \|\Omega\|_F \quad \text{for } a_j, b_j < \infty, \tag{3.27}$$

where we use the notation $\lambda_{y_0}\colon \mathfrak{g} \to \mathcal{M}$, $\lambda_{y_0}(\Omega) = \lambda(\Omega, y_0) = \Lambda(\exp(\Omega), y_0)$ for the Lie algebra action, $\|\cdot\|_F$ denotes the Frobenius norm and $j = 0, 1, \ldots, d$. If the truncation index q satisfies $q \geq 2\gamma - 2$, then the SRKMK scheme for solving the SDE (3.22) *on the manifold $\mathcal{M}$ is also of strong order γ.*

Proof. Since we assume the Lie algebra action $\lambda_{y_0}\colon \mathfrak{g} \to \mathcal{M}$, $\lambda_{y_0}(\Omega) = \Lambda(\exp(\Omega), y_0)$ to be a smooth mapping, the order of convergence on $\mathcal{M}$ is as high as the order of the scheme used for the corresponding equation in $\mathfrak{g}$. Therefore, we analyse the strong convergence order of a SRK scheme applied to (3.23). Let us denote by Ω_h the exact solution of (3.23) at time $t = t_1 = h$, and by Ω_h^q the exact solution of the truncated version of (3.23) after one time step, namely

$$d\Omega = \sum_{k=0}^{q} \frac{B_k}{k!} \operatorname{ad}_\Omega^k \left((f_0 \circ \lambda_{y_0})(\Omega)\right) dt + \sum_{j=1}^{d} \sum_{k=0}^{q} \frac{B_k}{k!} \operatorname{ad}_\Omega^k \left((f_j \circ \lambda_{y_0})(\Omega)\right) dW_t^j, \quad \Omega(0) = \mathbf{0}.$$

We prove the statement of the theorem in the following seven steps.

Step 1: Numerical error

We inspect the mean-squared error as given in the Definition 3.4 and apply the Minkowski inequality:

$$\left(\mathbb{E}\left[\|\Omega_h - \Omega_1\|_F^2\right]\right)^{1/2} \leq \left(\mathbb{E}\left[\|\Omega_h - \Omega_h^q\|_F^2\right]\right)^{1/2} + \left(\mathbb{E}\left[\|\Omega_h^q - \Omega_1\|_F^2\right]\right)^{1/2}.$$

The result can be interpreted as a split of the mean-squared error into a modelling error (the first summand) and a numerical error (the second summand). Since we are assuming that we are using a SRK scheme of strong order γ, the numerical error satisfies

$$\left(\mathbb{E}\left[\|\Omega_h^q - \Omega_1\|_F^2\right]\right)^{1/2} \leq C_1 h^\gamma$$

for some $C_1 < \infty$ by construction and it remains to prove that the modelling error satisfies

$$\left(\mathbb{E}\left[\|\Omega_h - \Omega_h^q\|_F^2\right]\right)^{1/2} \leq C_2 h^{(q+2)/2}$$

for some $C_2 < \infty$.

Step 2: Itô isometry

Let us analyse the modelling more thoroughly by again using the Minkowski inequality, the Itô isometry (3.2) and properties of the matrix norm, such that

$$\begin{aligned}
&\left(\mathbb{E}\left[\|\Omega_h - \Omega_h^q\|_F^2\right]\right)^{1/2} \\
&\leq \left(\mathbb{E}\Big[\Big\| \int_0^h \sum_{k=q+1}^{\infty} \frac{B_k}{k!} \operatorname{ad}_{\Omega_s}^k \left((f_0 \circ \lambda_{y_0})(\Omega_s)\right) ds + \right. \\
&\qquad \left. \sum_{j=1}^{d} \int_0^h \sum_{k=q+1}^{\infty} \frac{B_k}{k!} \operatorname{ad}_{\Omega_s}^k \left((f_j \circ \lambda_{y_0})(\Omega_s)\right) dW_s^j \Big\|_F^2 \Big]\right)^{1/2} \\
&\leq \left(\mathbb{E}\left[\left\| \int_0^h \sum_{k=q+1}^{\infty} \frac{B_k}{k!} \operatorname{ad}_{\Omega_s}^k \left((f_0 \circ \lambda_{y_0})(\Omega_s)\right) ds \right\|_F^2\right]\right)^{1/2} + \\
&\qquad \sum_{j=1}^{d} \left(\mathbb{E}\left[\left\| \int_0^h \sum_{k=q+1}^{\infty} \frac{B_k}{k!} \operatorname{ad}_{\Omega_s}^k \left((f_j \circ \lambda_{y_0})(\Omega_s)\right) dW_s^j \right\|_F^2\right]\right)^{1/2} \\
&\leq \left(\int_0^h \mathbb{E}\left[\left\| \sum_{k=q+1}^{\infty} \frac{B_k}{k!} \operatorname{ad}_{\Omega_s}^k \left((f_0 \circ \lambda_{y_0})(\Omega_s)\right) \right\|_F^2\right] ds\right)^{1/2} + \\
&\qquad \sum_{j=1}^{d} \left(\int_0^h \mathbb{E}\left[\left\| \sum_{k=q+1}^{\infty} \frac{B_k}{k!} \operatorname{ad}_{\Omega_s}^k \left((f_j \circ \lambda_{y_0})(\Omega_s)\right) \right\|_F^2\right] ds\right)^{1/2} \\
&\leq \sum_{j=0}^{d} \left(\int_0^h \mathbb{E}\left[\left(\sum_{k=q+1}^{\infty} \frac{|B_k|}{k!} \left\|\operatorname{ad}_{\Omega_s}^k \left((f_j \circ \lambda_{y_0})(\Omega_s)\right)\right\|_F\right)^2\right] ds\right)^{1/2}. \qquad (3.28)
\end{aligned}$$

Step 3: Adjoint operator

The Frobenius norm is submultiplicative such that we get the following estimate for the adjoint operator of an arbitrary matrix H,

$$\|\mathrm{ad}_\Omega(H)\|_F = \|[\Omega, H]\|_F = \|\Omega H - H\Omega\|_F \leq 2\|\Omega\|_F \|H\|_F .$$

By induction it follows for the k-th iterate of the adjoint operator that

$$\left\|\mathrm{ad}_\Omega^k(H)\right\|_F = \left\|\left[\Omega, \mathrm{ad}_\Omega^{k-1}(H)\right]\right\|_F \leq 2^k \|\Omega\|_F^k \|H\|_F$$

such that the input of the expected value in (3.28) can be estimated by

$$\left(\|(f_j \circ \lambda_{y_0})(\Omega_s)\|_F \sum_{k=q+1}^{\infty} \frac{|B_k|}{k!} 2^k \|\Omega_s\|_F^k\right)^2 \tag{3.29}$$

for $j = 0, 1, \ldots, d$, which is evident by setting $H = (f_j \circ \lambda_{y_0})(\Omega_s)$ above.

Step 4: Estimate for the remainder

We know from Lemma 2.3 that the Bernoulli numbers are implicitly defined by

$$\sum_{k=0}^{\infty} \frac{B_k}{k!} x^k = \frac{x}{e^x - 1}.$$

Taking the absolute values of the Bernoulli numbers on the left-hand side leads to

$$\sum_{k=0}^{\infty} \frac{|B_k|}{k!} x^k = \frac{x}{2}\left(1 + \cot\left(\frac{x}{2}\right)\right) + 2, \tag{3.30}$$

which is an identity that has also been used for example in [30, p. 48]. The right-hand side of this identity is defined on $I = \{x \in \mathbb{R} : \frac{x}{2\pi} \notin \mathbb{Z}\}$ such that we can set

$$g\colon I \to \mathbb{R}, \quad x \mapsto \frac{x}{2}\left(1 + \cot\left(\frac{x}{2}\right)\right) + 2.$$

Let us apply Taylor's theorem to the function g at the point 0 to obtain

$$g(x) = \sum_{k=0}^{q} \frac{g^{(k)}(0)}{k!} x^k + R_q(x), \quad R_q(x) = \frac{g^{(q+1)}(\xi)}{(q+1)!} x^{q+1},$$

where we consider the Lagrange form of the remainder $R_q(x)$ for some real number ξ between 0 and x.

Comparing this representation of the function g to (3.30) and to (3.29), we see that the remainder $R_q(x)$ corresponds to the summation in (3.29) if we set $x = 2\|\Omega_s\|_F$. Let us recall that the expression (2.10) only converges for $\|\Omega\| < \pi$.

Therefore, we limit the domain of the function g to

$$\tilde{I} = \{x \in \mathbb{R} : |x| < 2\pi\}.$$

The restriction of g to $\tilde{I}$ gives us the benefit that we can find upper bounds for g and its derivatives. In particular, there exists an upper bound dependent on the truncation index, $M_q > 0$, such that

$$\left| g|_{\tilde{I}}^{(q+1)}(\xi) \right| \leq M_q$$

for all ξ between 0 and x. Hence, we can find the following estimate for the remainder,

$$|R_q(x)| = \left| \frac{g|_{\tilde{I}}^{(q+1)}(\xi)}{(q+1)!} x^{q+1} \right| \leq \frac{M_q}{(q+1)!} |x|^{q+1},$$

and the following estimate for (3.29) by identifying x with $2\|\Omega_s\|_F$,

$$\left(\frac{2^{q+1} M_q}{(q+1)!} \left\| (f_j \circ \lambda_{y_0})(\Omega_s) \right\|_F \|\Omega_s\|_F^{q+1} \right)^2.$$

Our next task is to find an estimate for the expected value of this expression.

Step 5: Linear growth

Assuming that the linear growth condition (3.27) holds and applying a simple Binomial formula, we have

$$\begin{aligned}
&\mathbb{E}\Big[\big\| f_j\big(\lambda_{y_0}(\Omega_s)\big) \big\|_F^2 \big\|\Omega_s\big\|_F^{2(q+1)} \Big] \\
&\leq \mathbb{E}\Big[\big(a_j + b_j \|\Omega_s\|_F\big)^2 \big\|\Omega_s\big\|_F^{2(q+1)} \Big] \\
&\leq a_j^2 \mathbb{E}\Big[\big\|\Omega_s\big\|_F^{2(q+1)} \Big] + 2a_j b_j \mathbb{E}\Big[\big\|\Omega_s\big\|_F^{2(q+3/2)} \Big] + b_j^2 \mathbb{E}\Big[\big\|\Omega_s\big\|_F^{2(q+2)} \Big],
\end{aligned} \tag{3.31}$$

where $a_j, b_j < \infty$ for $j = 0, 1, \ldots, d$. We recognize that we are left with three similar terms, where the first summand is the one of lowest order.

Step 6: Itô-Taylor expansion

The only thing left to do is inserting an estimate for the Itô-Taylor expansion according to Theorem 3.6 (see also Proposition 5.9.1 in [36]) to get

$$\Omega_s = \Omega_0 + R_s = R_s, \quad \mathbb{E}\left[\|R_s\|_F^2\right] \leq C_1 s$$

for some $C_1 < \infty$ such that

$$\mathbb{E}\left[\|\Omega_s\|_F^{2(q+1)}\right] = \mathbb{E}\left[\|R_s\|_F^{2(q+1)}\right] \leq C_1 s^{q+1}.$$

Similar estimates but of higher order hold for the other summands in (3.31).

Step 7: Overall estimate

This last step serves as a summary of the previous steps:

$$\begin{aligned}
&\left(\mathbb{E}\left[\|\Omega_h - \Omega_h^q\|_F^2\right]\right)^{1/2} \\
&\leq \sum_{j=0}^{d} \left(\int_0^h \mathbb{E}\left[\Big(\sum_{k=q+1}^{\infty} \frac{|B_k|}{k!} \big\| \operatorname{ad}_{\Omega_s}^k \left(f_j(\lambda_{y_0}(\Omega_s)) \right) \big\|_F \Big)^2 \right] ds \right)^{1/2} \\
&\leq \sum_{j=0}^{d} \left(\int_0^h \mathbb{E}\left[\left\| f_j(\lambda_{y_0}(\Omega_s)) \right\|_F^2 \Big(\sum_{k=q+1}^{\infty} \frac{|B_k|}{k!} 2^k \|\Omega_s\|_F^k \Big)^2 \right] ds \right)^{1/2} \\
&\leq \frac{2^{q+1} M_q}{(q+1)!} \sum_{j=0}^{d} \left(\int_0^h \mathbb{E}\left[\left\| f_j(\lambda_{y_0}(\Omega_s)) \right\|_F^2 \|\Omega_s\|_F^{2(q+1)} \right] ds \right)^{1/2} \\
&\leq \frac{2^{q+1} M_q}{(q+1)!} \sum_{j=0}^{d} \left(\int_0^h \left(a_i^2 \mathbb{E}\left[\|\Omega_s\|_F^{2(q+1)} \right] + 2 a_i b_i \mathbb{E}\left[\|\Omega_s\|_F^{2(q+3/2)} \right] + b_i^2 \mathbb{E}\left[\|\Omega_s\|_F^{2(q+2)} \right] \right) ds \right)^{1/2} \\
&\leq \frac{2^{q+1} M_q}{(q+1)!} \sum_{j=0}^{d} \left(\int_0^h \mathcal{O}(s^{q+1}) ds \right)^{1/2} = \mathcal{O}\left(h^{(q+2)/2} \right).
\end{aligned}$$

We achieved our objective from the first step, which concludes our proof.

□

After proving the main result of this thesis we would like to make the following remarks:

1. The linear growth condition (3.27) is a sufficient condition for the existence and uniqueness of the solution of (3.23), which can be confirmed by the comparison with (3.8) in Theorem 3.3. In the case where the considered manifold $\mathcal{M}$ is the unit

sphere S^2 and the corresponding Lie algebra is $\mathfrak{g} = \mathfrak{so}(3)$, which is for example the case when rotation motions of rigid bodies are regarded, the linear growth condition can be omitted since properties of skew-symmetric matrices and estimates on the unit sphere can be used instead. We refer to [54] for more details.

2. As mentioned in Section 3.2.2 the construction of SRK schemes of higher strong order than $\gamma = 1.5$ is computationally costly. However, assuming that there will be efficient higher order SRK methods in the future, Theorem 3.12 gives an instruction on how to construct higher order SRKMK methods. Given a strong convergence order γ one has to evaluate the coefficients of the SDE in the Lie algebra $\mathfrak{g}$ only up to the index $2\gamma - 2$. We refer the reader to [45] and the references therein for more details on the approximation of iterated stochastic integrals.

3. If the considered Lie algebra action $\lambda\colon \mathfrak{g}\times\mathcal{M} \to \mathcal{M}$ is based on a quadratic Lie group G, then there is no formulation of a restrictive summation needed since one can apply the Cayley map instead of the matrix exponential which does not introduce any modelling error.

4. In the first step of the proof we have split the mean-squared error into a numerical and a modelling error, where we assumed the numerical error to have the convergence order of the applied SRK method. We have shown that the modelling error is of the same order if the truncation index q is chosen to fulfill $q \geq 2\gamma - 2$. A violation of this inequality will thus lead to an order reduction, e.g. if the truncation index is set as $q = 0$ when a SRK method of order $\gamma = 1.5$ is applied. A visualization of this order reduction will be provided in Chapter 4.

3.4.2 Comparison to other numerical schemes

SRKMK methods are certainly not the only possibility to approximate SDEs on manifolds. Therefore, we present some other ideas.

Stochastic Runge-Kutta schemes

We already know that applying a SRK method directly to (3.22) will result in a drift-off, which means that the numerical approximations do not stay on the manifold $\mathcal{M}$ since the linear operations of SRK schemes do not comply with curved spaces. As a redemption of the violation of the manifold structure one can think of a projection step in the following

way

$$\|y_1 - \tilde{y}_1\| \to \min$$

such that coming from the approximation y_1 obtained by the SRK method one computes the closest approximation $\tilde{y}_1 \in \mathcal{M}$. One can set up $\tilde{y}_1$ using a Lie algebra action $\lambda_{y_0} \colon \mathfrak{g} \to \mathcal{M}$ for a fixed $y_0 \in \mathcal{M}$, i.e. $\tilde{y}_1 = \lambda_{y_0}(\tilde{\Omega})$, and define $\tilde{\Omega}$ as a linear combination of the basis matrices $\mathcal{E}_i$ of the Lie algebra $\mathfrak{g}$, i.e. $\tilde{\Omega} = \sum_{i=1}^{n} \alpha_i \mathcal{E}_i$ assuming n to be the dimension of the Lie algebra $\mathfrak{g}$. Then the coefficients $\alpha_i \in \mathbb{R}$ are the parameters to be optimized in the minimization problem defined above. However, we can already presume this method to be much more computationally intense than SRKMK schemes since an optimization problem has to be solved in every discretization step for every path.

Stochastic Magnus expansion

The original idea of Magnus [43] was to solve the linear matrix ODE

$$\dot{y} = f(t)y(t), \quad y(0) = y_0,$$

by assuming that the solution can be written as $y(t) = \exp\big(\Omega(t)\big)y_0$ and applying the Picard iteration to the logarithm, i.e. to (2.17) such that we have

$$\Omega(t) = \int_0^t f(t_1)\, dt_1 - \frac{1}{2}\int_0^t \Big[\int_0^{t_1} f(t_2)\, dt_2, f(t_1)\Big]\, dt_1 + \dots.$$

This approach has been expanded to nonlinear ODEs (see [6] and the references therein) and to Stratonovich SDEs on manifolds (see [11], [77]). A derivation of the stochastic Magnus expansion for linear Itô SDEs can be found in [34]. An extension of this approach to nonlinear Itô SDEs is not straightforward. Since the linear Itô SDE in [34] is set up without the Itô correction term, one has to deal with a corresponding correction in (3.23):

$$v_0(\Omega) = \operatorname{dexp}_{\Omega}^{-1}\left(f_0(t) - \frac{1}{2}\sum_{j=1}^{d}\sum_{k=0}^{\infty}\sum_{l=0}^{\infty}\left(\frac{\operatorname{ad}_{\Omega}^{k}(v_j)}{(k+1)!}\frac{\operatorname{ad}_{\Omega}^{l}(v_j)}{(l+1)!} + \frac{\big[\operatorname{ad}_{\Omega}^{k}(v_j), \operatorname{ad}_{\Omega}^{l}(v_j)\big]}{(k+l+2)(k+1)!\,l!}\right)\right).$$

As the main idea of the Magnus expansion is to solve (3.23) via a Picard iteration, we expect a high computational effort for the evaluation of the coefficients with this drift in a nonlinear case. Moreover, we have seen that this approach will lead to unnatural geometric restrictions if applied to manifolds that are not linear spaces. Nevertheless, a comparison of SRKMK schemes and stochastic Magnus integrators for nonlinear Stratonovich SDEs is still outstanding.

Geometric Castell-Gaines methods

Let $\hat{\Omega}_t$ denote a truncated Magnus expansion. Then an approximation of the solution of (3.22) is given by $y_t = \exp(\hat{\Omega}_t)y_0$. On a time interval $[t_\ell, t_{\ell+1}]$ we get the Magnus integrator

$$y_{\ell+1} = \exp(\hat{\Omega}_{\ell,\ell+1})y_\ell,$$

where $\hat{\Omega}_{\ell,\ell+1}$ is the truncated Magnus series across the interval $[t_\ell, t_{\ell+1}]$. Thus, the solution of the SDE (3.22) can be approximated by a solution of the ODE

$$\dot{u}(\tau) = \hat{\Omega}_{\ell,\ell+1}u(\tau)$$

for $\tau \in [0, 1]$, see the approach of Castell & Gaines in [13] and [44] for this geometric expansion of that approach. Computing a solution of this ODE will give us $u(1) \approx y_{\ell+1}$ if $u(0) = y_\ell$.

Numerical experiments where classical Runge-Kutta methods were applied to solve the ODE and therefore approximate the solution of (3.22) can be found in [44]. The results have shown that these schemes perform better than the direct application of stochastic Taylor schemes but worse than SRKMK methods in terms of preserving the manifold structure.

Applying a RKMK method (see Algorithm 1) to the ODE would ensure that $y_{\ell+1} \in \mathcal{M}$. However, a practical procedure for determining $f\colon \mathcal{M} \to \mathfrak{g}$ such that $\hat{\Omega}_{\ell,\ell+1} = \lambda_* f(y)$ is still missing.

Chapter 4

Applications

In this chapter we finally solve the SDE on the sphere $\mathcal{S}$ that describes the dynamics of a randomly perturbed rigid body, e.g. the dynamics of a satellite taking measurement uncertainties into account. We consider (3.1) again or rather its counterpart in Itô form, namely

$$dy = \left(f_0(y)y + \frac{1}{2} \left(\frac{\mathrm{d}}{\mathrm{d}y} f_1(y)y \right) f_1(y)y \right) dt + f_1(y)y \, dW_t, \quad y(0) = y_0, \tag{4.1}$$

with $f_0, f_1 \colon \mathcal{S} \to \mathfrak{so}(3)$ given as

$$f_0(y) = \begin{pmatrix} 0 & y_3/I_3 & -y_2/I_2 \\ -y_3/I_3 & 0 & y_1/I_1 \\ y_2/I_2 & -y_1/I_1 & 0 \end{pmatrix}, \quad f_1(y) = \begin{pmatrix} 0 & y_3/J_3 & -y_2/J_2 \\ -y_3/J_3 & 0 & y_1/J_1 \\ y_2/J_2 & -y_1/J_1 & 0 \end{pmatrix}$$

for $y = (y_1, y_2, y_3)^\top \in \mathcal{S} = S^2$ and constants I_1, I_2, I_3 and J_1, J_2, J_3 that represent the principal moments of inertia. We approximate the solution by applying SRKMK methods which we have derived and analysed in the previous chapters. The results of the simulation will be presented first in the forthcoming Section 4.1, where we will also have a look on the dynamics of an autonomous underwater vehicle subjected to random perturbations.

Moreover, we will present applications of SRKMK schemes to SDEs on manifolds found in financial mathematics in Section 4.2. In these examples we will only look at linear Itô SDEs on manifolds to make the calibration of the models to historical data as easy as possible.

All examples shown below have been implemented in the software package MATLAB.

4.1 Examples in rigid body modelling

The dynamics of a rigid body are usually described by an ODE on a manifold that describes rotations and translations of the rigid body, see e.g. [14] and the references therein. We consider these dynamics and assume them to be affected by random perturbations such that we are dealing with SDEs on the corresponding manifolds as done for example in [44], [77].

As these equations for physical interrelations were established as extensions of ODEs on manifolds, one will find them usually denoted in Stratonovich form. However, we choose to continue using the Itô representation since we conducted the analysis in the previous chapter in Itô notation but we keep in mind that we are also solving the Stratonovich counterpart of the SDEs like the one given in (3.1).

4.1.1 A perturbed rigid body

We approximate the solution of (4.1) by using Algorithm 3.11, where we have set $y_0 = (\cos(0.9), 0, \sin(0.9))^\top$ as the initial value in $\mathcal{M} = \mathcal{S} = S^2$ and $(I_1, I_2, I_3) = (3, 1, 2)$ and $(J_1, J_2, J_3) = (1, 0.5, 1.5)$ as the moments of inertia. Note that by this choice of constants we are dealing with noncommutative vector fields in (4.1). We remind the reader that we set $G = \mathrm{SO}(3)$ for the Lie group action $\Lambda\colon \mathrm{SO}(3) \times S^2 \to S^2$, $\Lambda(R, y) = Ry$ for an arbitrary rotation matrix R, and $\mathfrak{g} = \mathfrak{so}(3)$ for the Lie algebra action $\lambda\colon \mathfrak{so}(3) \times S^2 \to S^2$, $\lambda(\Omega, y) = \exp(\Omega)y$ for a skew-symmetric matrix Ω (as described in Example 3 in Section 2.2.2).

For the Numerical method step of Algorithm 3.11 we used the Euler-Maruyama scheme (3.15), the SRK method (3.18) of strong order $\gamma = 1$ with the coefficients from Table 3.2 and the SRK method (3.19) of strong order $\gamma = 1.5$ with the coefficients from Table 3.3, where we chose the truncation index q in (2.19) according to Theorem 3.12, i.e. $q = 0$ for the schemes of strong order $\gamma = 0.5$ and $\gamma = 1$ and we chose $q = 1$ for Rößler's scheme of order $\gamma = 1.5$. Appending one step with the corresponding Lie algebra action $\lambda(\Omega_1, y_\ell) = \exp(\Omega_1)y_\ell$ to the one step taken with the SRK method applied to (3.23) in the Lie algebra $\mathfrak{g} = \mathfrak{so}(3)$ to obtain Ω_1 extends the SRK to a SRKMK method and preserves the geometry of the manifold $\mathcal{M} = S^2$. For this reason the name *geometric Euler-Maruyama* (gEM) scheme was established in [47] for the simplest SRKMK method. We take up this way of naming the geometry-preserving numerical schemes and call the schemes elaborated above *geometric SRK* methods and use the abbreviations *gSRK1* and *gSRK1.5* correspondingly to indicate the higher order geometric schemes.

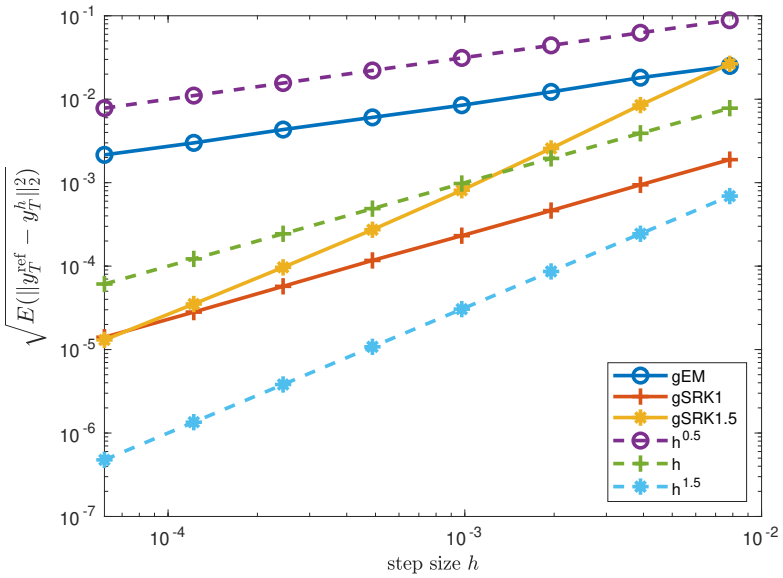

Figure 4.1: Simulation of the mean-square convergence order for $M = 1000$ paths in the rigid body problem.

Convergence

The result of our simulation of the mean-square error,

$$\left(\mathbb{E}\left(\|y_T - y_T^h\|_2^2\right)\right)^{1/2} \approx \left(\frac{1}{M}\sum_{j=1}^{M}\left\|y_{T,j}^{\mathrm{ref}} - y_{T,j}^h\right\|_2^2\right)^{1/2},$$

for $M = 1000$ paths is depicted in a log-log-plot in Figure 4.1. The approximations y_T^h were computed by following the steps of Algorithm 3.11 with the step sizes $h = 2^{-i}$ for $i = 7, 8, \ldots, 14$.

For the reference solution y_T^{ref} we applied the SRKMK method gSRK1.5 using a smaller step size, $h = 2^{-16}$, and the Lie algebra action $\lambda(\Omega, y) = \mathrm{cay}(\Omega)y$ to get a more accurate solution since applying (2.14) does not introduce any modelling error. We remark that we could also have used the closed-form expression (2.12) for the matrix exponential based on the Rodrigues formula (2.11). Figure 4.1 shows that the chosen truncation indices are sufficient for the SRKMK schemes to inherit the convergence order γ of the SRK method chosen in the second step of Algorithm 3.11.

We can assume that the weak convergence order of the respective schemes is at least as high as the strong convergence order as this is usually the case for numerical schemes for SDEs in vector spaces. However, checking this assumption numerically includes finding

a suitable class of functions g in the definition of weak convergence (3.13) that maps elements from the manifold $\mathcal{M}$ to $\mathbb{R}$. Setting g simply as the identity function would lead to the problem of finding a suitable approximation of the expected value of elements on a manifold since the arithmetic mean would give results drifting off the manifold. We view the analysis and simulation of the weak convergence of SRKMK methods as future work and give an idea of how this investigation can be approached in the Conclusion.

Geometric properties

The geometry-preserving property of SRKMK schemes can be viewed in Figure 4.2. We computed $L = 450$ steps with a step size of $h = 0.1$ applying the SRK method (3.18) of order $\gamma = 1$, which we abbreviate by SRK1, directly to (4.1) and its extension to a SRKMK method gSRK1. The sample path of SRK1 clearly drifts off the unit sphere S^2 and would give us incorrect dynamics of the rigid body since keeping

$$C(y) = y_1^2 + y_2^2 + y_3^2$$

constant represents the preservation of the angular momentum. In contrast to SRK1, the sample path of gSRK1 remains on S^2 and preserves the manifold structure. Another visualization of the drift-off of the SRK scheme opposed to the SRKMK scheme is given in Figure 4.3, which shows that gSRK1 preserves the manifold within machine precision.

Sample paths for the case where the two vector fields in (4.1) commute, i.e. where $(J_1, J_2, J_3) = 2\dot{(}I_1, I_2, I_3)$, are shown in Figure 4.4. We point out that the red sample path in Figure 4.4 corresponds to the sample path from our motivating example in Figure 1.1 in the Introduction of this thesis and that we achieved our goal from this example with the simulation of the blue sample path (gSRK1 in Figure 4.4).

4.1.2 Autonomous underwater vehicle

Let us consider an ellipsoidal rigid body immersed in an ideal fluid, i.e. in an infinitely large volume of incompressible, irrotational and inviscid fluid that is at rest at infinity, and assume that its center of gravity and its center of buoyancy coincide. One is interested in such kind of rigid bodies when control laws for unmanned underwater vehicles that must manage their own motions are examined.

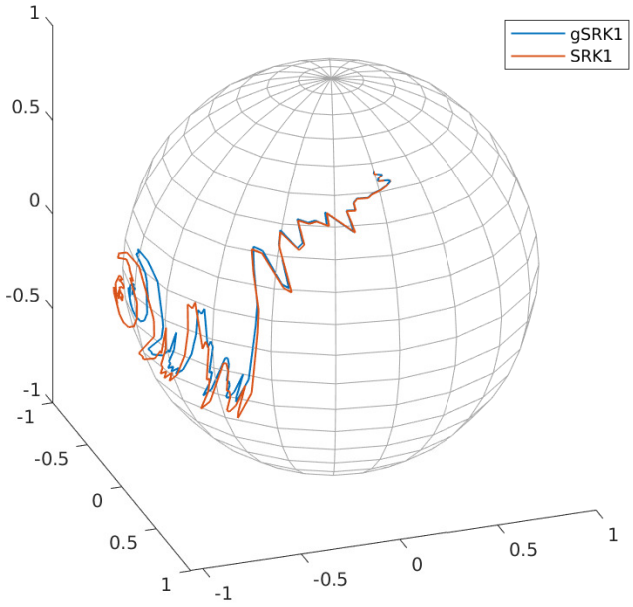

Figure 4.2: Sample path of the SRK method (3.18) of strong order $\gamma = 1$ and its geometry-preserving counterpart gSRK1 on S^2.

ODE on $\mathfrak{se}(3)^*$

The dynamics of such autonomous underwater vehicles (AUV) can be derived using Kirchhoff's equations. We denote by w and u the angular and linear velocity vectors, respectively, and by π and ρ the angular and linear momentum vectors, respectively. Then the equations of motion can be described by

$$\dot{\pi} = \pi \times w + \rho \times u, \quad \dot{\rho} = \rho \times w, \tag{4.2}$$

which can be viewed as Lie-Poisson (non-canonical Hamiltonian) dynamics on $\mathfrak{se}(3)^*$, the dual space of the Lie algebra $\mathfrak{se}(3)$, resulting from a reduction of the *full* dynamics on the phase space $T^*\mathrm{SE}(3)$, the cotangent bundle of the special Euclidean group $\mathrm{SE}(3)$. Conserved quantities along these equations of motion, that reflect the preservation of the angular and linear momentum, are given by the Casimir functions $C_i \colon \mathfrak{se}(3)^* \to \mathbb{R}$,

$$C_1 = \pi^\top \rho, \quad C_2 = |\rho|^2.$$

We refer to [28] for the derivation and more details.

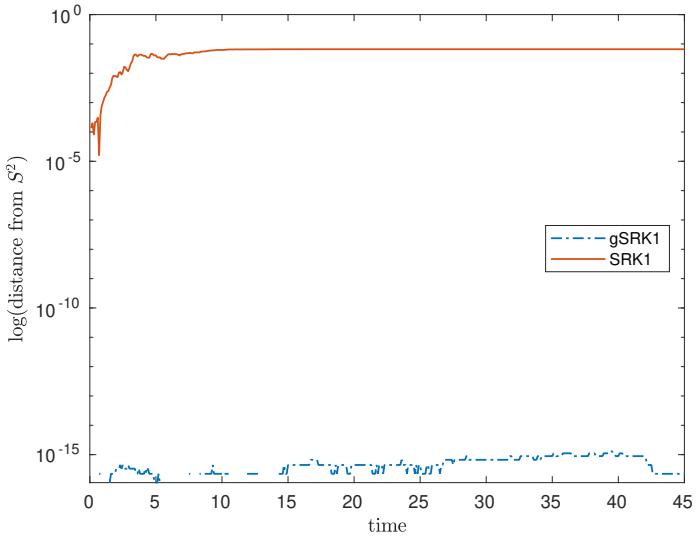

Figure 4.3: Logarithmic distance of the numerical solutions from the unit sphere S^2.

Elements of the Lie group of rigid body motions, $(R, r) \in \mathrm{SE}(3)$ with $R \in \mathrm{SO}(3)$ and $r \in \mathbb{R}^3$ (see Example 3 in Section 3.3.2 for this notation), represent the orientation and position of the AUV in the inertial frame and can be reconstructed by solving

$$\dot{R} = R\hat{w}, \quad \dot{r} = Ru,$$

given the angular and linear velocity, w and u, and using the hat map $\hat{\cdot}\colon \mathbb{R}^3 \to \mathfrak{so}(3)$ defined by (2.8).

SDE on $\mathfrak{se}(3)^*$

Now, let us assume that the dynamics (4.2) are perturbed by a Wiener process such that a model of the AUV for $y = (\pi, \rho) \in \mathfrak{se}(3)^*$ is given by (3.22) with $d = 1$ and the coefficients given by

$$V_j(y) = \big(\lambda_* f_j(y)\big)(y) = \big(\pi \times w_j(y) + \rho \times u_j(y), \rho \times w_j(y)\big)$$

(as stated in Example 3 in Section 3.3.2) for $j = 0, 1$. The Lie group and Lie algebra acting on $\mathcal{M} = \mathfrak{se}(3)^*$ are set as $G = \mathrm{SE}(3)$ and $\mathfrak{g} = \mathfrak{se}(3)$, respectively, and the functions $f_j\colon \mathcal{M} \to \mathfrak{g}$, $f_j(y) = (w_j(y), u_j(y))$ are defined by the angular velocity $w_j(y) = I_j^{-1}\pi$ and the linear velocity $u_j(y) = M_j^{-1}\rho$, where I_j and M_j are diagonal matrices for $j = 0, 1$ representing the moments of inertia and mass matrices, respectively. An expansion of the dynamics of the AUV according to Kirchhoff's equations to the stochastic case has been first considered in [44].

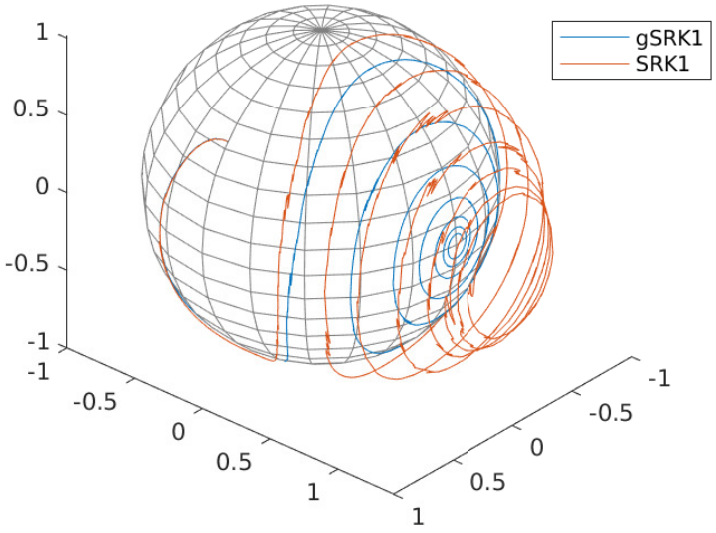

Figure 4.4: Sample paths of SRK1 and gSRK1 on S^2 for commuting vector fields.

The approximation of the solution of this SDE on the manifold $\mathcal{M} = \mathfrak{se}(3)^*$ via SRKMK methods requires the computation of the exponential map, $\exp_{\mathfrak{se}(3)} \colon \mathfrak{se}(3) \to \mathrm{SE}(3)$. A closed-form expression can be found in [62] and reads as

$$\exp_{\mathfrak{se}(3)}(\Omega) = \begin{pmatrix} \Theta & \frac{1}{\|\theta\|_2^2}\big((I-\Theta)(\theta\times\zeta) + \theta\theta^\top\zeta\big) \\ O^\top & 1 \end{pmatrix}$$

with $O = (0,0,0)^\top$ for an arbitrary $\Omega = (\theta, \zeta) \in \mathfrak{se}(3)$ using a similar shorthand notation as described earlier. The submatrix $\Theta = \exp_{\mathfrak{so}(3)}(\hat{\theta}) \in \mathrm{SO}(3)$ can be computed via the Rodrigues formula (2.11).

Simulation

We applied Algorithm 3.11 to solve (3.22) with the initial value $y_0 = (\pi_0, \rho_0)$ with the initial angular momentum $\pi_0 = (\sqrt{2}, \sqrt{2}, 0)^\top$ and linear momentum $\rho_0 = (0, \sqrt{2}, \sqrt{2})^\top$, the moments of inertia $I_0 = \mathrm{diag}(3,1,2)$ and $I_1 = \mathrm{diag}(1, 0.5, 1.5)$ and the mass matrices $M_0 = \mathrm{diag}(20, 55, 101)$ and $M_1 = \mathrm{diag}(55, 78, 120)$.

In the Numerical method step of the Algorithm 3.11 we used the Euler-Maruyama scheme (3.15) of strong order $\gamma = 0.5$, the SRK scheme (3.18) of strong order $\gamma = 1$ and

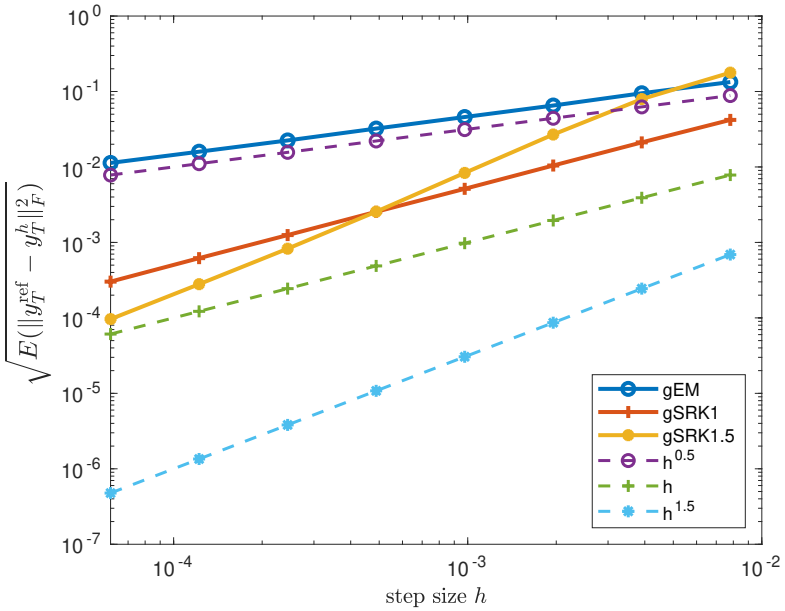

Figure 4.5: Simulation of the mean-square convergence for $M = 1000$ paths in the AUV problem.

the SRK scheme (3.19) of strong order $\gamma = 1.5$ with truncation indices chosen according to Theorem 3.12 to obtain approximations y_T^h for the step sizes $h = 2^{-i}$ for $i = 7, 8, \ldots, 14$.

Similar to the rigid body problem in the previous example we simulated the mean-square convergence again using the highest-order scheme as a reference solution with the step size $h = 2^{-16}$. The result can be viewed in the log-log-plot in Figure 4.5. It shows a similar outcome as obtained previously, which verifies our Theorem 3.12 and leads us to the conclusion that the convergence behaviour of SRKMK methods does not depend on the underlying problem they are applied to.

The SRKMK methods preserve the Casimir functions $C_1 = \pi^\top \rho$ and $C_2 = |\rho|^2$ whereas the underlying SRK methods applied directly to (3.22) fail at this task, which has already been shown in [44]. An example for this can be found in Figure 4.6, where we have plotted the logarithmic distance of sample paths using the SRK scheme (3.18) of strong order $\gamma = 1$ and its geometric extension denoted by gSRK1.

As the second Casimir function $C_2 = |\rho|^2$ can be depicted as a sphere, we visualized the preservation or rather the violation of the manifold by the sample paths in Figure 4.7.

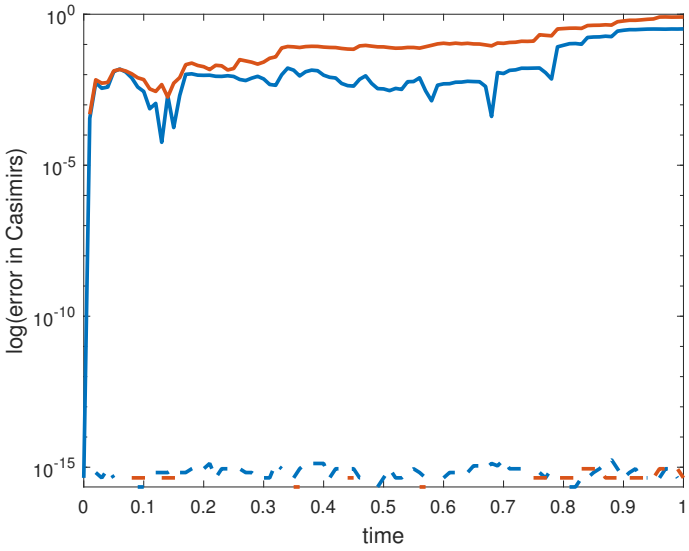

Figure 4.6: Distance of a sample path of Rößler's order 1 scheme (SRK1, solid lines) applied directly to (3.22) and its geometry-preserving counterpart (gSRK1, dashed lines). Blue lines correspond to the error in C_1 while red lines indicate the error in C_2.

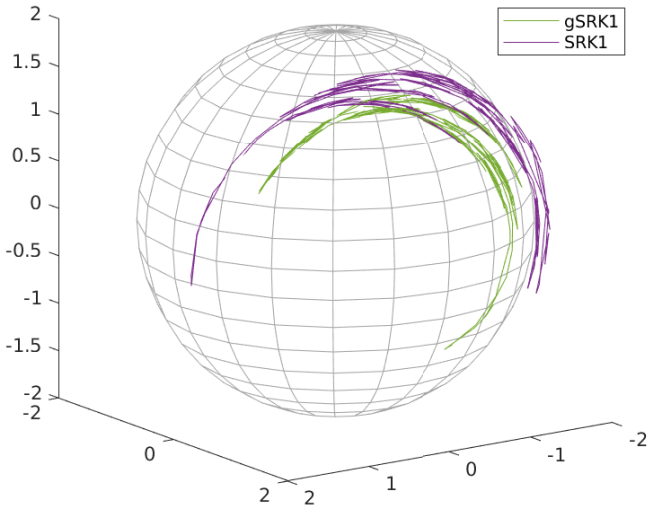

Figure 4.7: Sample paths of approximations obtained by using gSRK1 and SRK1 on the sphere of the second Casimir function.

4.2 Examples in computational finance

Practitioners of financial mathematics usually do not work on manifolds that are not also linear spaces. At least there is not much literature documenting this. The few examples that we could find include [21], [41], [61], which consider interest rate models on manifolds.

It is a pity that there are only this little applications of manifolds in a financial mathematical context so far since some considered quantities in finance can be analysed more efficiently by acknowledging their properties on a curved space.

Here we will only examine stochastic processes that evolve on the manifold of symmetric matrices in the first example and on the manifold of transition matrices in the second example. We hope that this will motivate the development of more models in computational finance taking advantage of the manifold structure in the future.

Modelling in financial mathematics often involves a calibration to some historical data, e.g. past stock prices, if we assume that future stock prices depend on past stock prices. In order to make the calibration to the historical data as easy as possible we suppose that the functions $f_j \colon \mathcal{M} \to \mathfrak{g}$, $j = 1, \dots, d$, are independent of the solution of the SDE on the manifold (3.22). These functions are needed for the definition of the vector fields on the right-hand side of (3.22).

This assumption allows us to construct the elements in the Lie algebra $\mathfrak{g}$ as constant or time-dependent linear combinations of the basis matrices of the vector space $\mathfrak{g}$. The coefficients of this linear combination will then give us the degrees of freedom needed for the calibration to the given data.

We will see in our first example that solving the resulting SDE (3.22) with the geometric Euler-Maruyama scheme is equivalent to approximating the following SDE in the Lie algebra,

$$d\Omega_t = A_t \, dt + \sum_{j=1}^{d} B_{j,t} \, dW_t^j, \quad \Omega_0 = \mathbf{0}, \tag{4.3}$$

where the coefficients A_t and $B_{j,t}$, $j = 1, \dots, d$, are the constructed linear combinations of the basis matrices of the Lie algebra $\mathfrak{g}$.

Taking advantage of this simplified version of the SDE (3.23) we reverse the construction of the SDE on the manifold $\mathcal{M}$ in our second example such that it is set up based on (4.3).

4.2.1 Covariance matrices as an isospectral flow

We assume a scenario in which a risk manager is given a density function of historical correlations between two or more entities for a specific time range and the initial correlation matrix of these entities at the beginning of the considered time interval retrieved from the middle office's reporting system. Based on these information the risk manager is given the task to generate valid time-dependent correlation matrices that take the stochastic nature of correlations into account while trying to match the density function of the historical data.

Stochastic correlation models for these kind of scenarios can for example be found in [71] and [19] and a comparison of these and more stochastic correlation models is given in [48]. Using the techniques acquired in the previous chapters we would advise the risk manager to model the underlying covariance matrices as a stochastic process on the manifold of symmetric matrices and to solve the resulting SDE with a SRKMK scheme. This methodology was developed in the course of [51] and [52] and is an extension of the deterministic model considered in [72]. We substantiate this approach below and show its effectiveness with the aid of a simple low-dimensional example.

SDE on the manifold of covariance matrices

Correlation matrices R_t, $t \geq 0$, are symmetric and positive semi-definite matrices with diagonal elements equal to one and absolute values of non-diagonal elements less than or equal to one. They can be recovered from their underlying covariance matrices P_t via the relation

$$R_t = \Sigma_t^{-1} P_t \Sigma_t^{-1}, \quad \Sigma_t = \big(\operatorname{diag}(P_t)\big)^{1/2},$$

where $\operatorname{diag}(P_t)$ are diagonal matrices with entries from the diagonal of P_t. The geometric property of symmetry allows us to consider covariance matrices as elements of the manifold

$$\operatorname{Sym}(n) = \{P \in \operatorname{GL}(n) : P = P^\top\}.$$

Note that the product of two symmetric matrices is not necessarily symmetric, which is why this manifold is not a Lie group. However, we can use the Lie group of rotation matrices $\operatorname{SO}(n)$ to define transport across the manifold $\mathcal{M} = \operatorname{Sym}(n)$ via the Lie group action

$$\Lambda \colon \operatorname{SO}(n) \times \operatorname{Sym}(n) \to \operatorname{Sym}(n), \quad \Lambda(Q, P) = QPQ^\top,$$

where Q is an arbitrary orthogonal matrix. Let P_0 be a fixed initial covariance matrix, then $\Lambda_{P_0}(Q) = \Lambda(Q, P_0)$ is similar to P_0, in particular $\Lambda_{P_0}(Q)$ and P_0 have the same eigenvalues. It follows that if P_0 is positive semi-definite, then $\Lambda_{P_0}(Q)$ is positive semi-definite as well. The corresponding Lie algebra action reads as

$$\lambda\colon \mathfrak{so}(n) \times \mathrm{Sym}(n) \to \mathrm{Sym}(n), \quad \lambda(\Omega, P) = \exp(\Omega) P \exp(-\Omega)$$

for a skew-symmetric Ω.

An SDE on $\mathcal{M} = \mathrm{Sym}(n)$ for $P_t = \lambda_{P_0}(\Omega_t) = \exp(\Omega_t) P_0 \exp(-\Omega_t)$ can be formulated as

$$dP_t = \left([Y_{0,t}, P_t] + \frac{1}{2} \sum_{j=1}^{d} \big[Y_{j,t}, [Y_{j,t}, P_t] \big] \right) dt + \sum_{j=1}^{d} [Y_{j,t}, P_t]\, dW_{j,t}. \tag{4.4}$$

This is an isospectral flow in $\mathrm{Sym}(n)$, where $Y_{j,t} \in \mathfrak{so}(n)$ for $j = 0, 1, \ldots, d$ (see Example 2 in Section 3.3.2).

Creating a correlation flow

Based on this SDE for covariance matrices and assuming that an initial correlation matrix R_0^{hist} and that a density function f^{hist} of the historical data are given we propose the risk manager to follow the subsequent steps in order to create valid, time-dependent correlation matrices that reflect the stochastic nature of correlations:

1. Compute from R_0^{hist} an initial covariance matrix P_0 and set the skew-symmetric matrices $Y_{0,t}, \ldots, Y_{d,t}$ in the *covariance flow* (4.4) such that parameters as degrees of freedom can be incorporated.

2. Solve the covariance flow (4.4) with a SRKMK method and transform the obtained covariance matrices to correlation matrices.

3. Estimate the density function from the so-obtained *correlation flow* and calibrate the involved parameters such that the density function of the correlation flow f^{flow} matches the density function of the historical correlation f^{hist}.

Let us concretise these steps for $n = 2$ and $d = 2$. Assume that the risk manager retrieved from the middle office's reporting system the initial correlation matrix

$$R_0^{\mathrm{hist}} = \begin{pmatrix} 1 & -0.0159 \\ -0.0159 & 1 \end{pmatrix} \tag{4.5}$$

of the moving correlations between the S&P 500 index and the Euro/US-Dollar exchange rate on a daily basis computed with a window size of 30 days from January 3, 2005 to January 6, 2006 seen in Figure 4.8. Since the correlation shown in this Figure is only

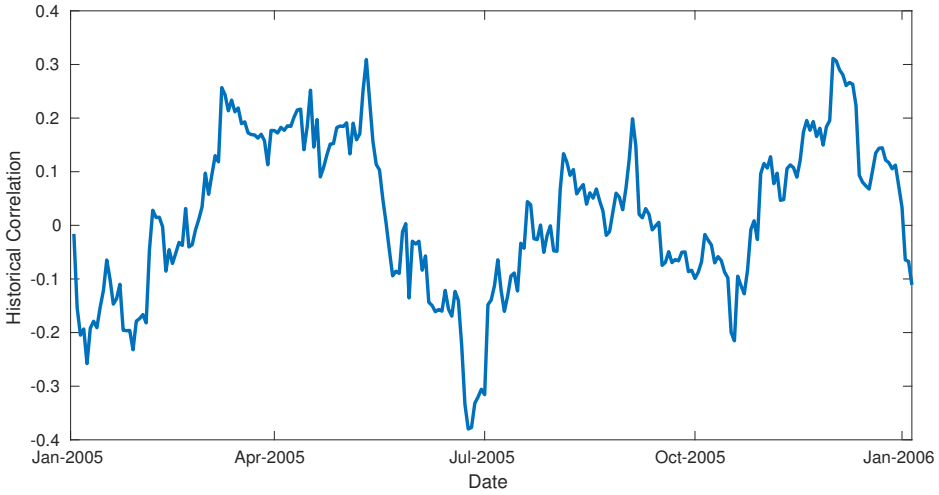

Figure 4.8: The 30-day historical correlations between S&P 500 and Euro/US-Dollar exchange rate, source of data: www.yahoo.com.

one of many possible realizations we assume the risk manager to be aware of the density function f^{hist} of the considered correlation shown in Figure 4.9, which was computed using kernel smoothing functions [7].

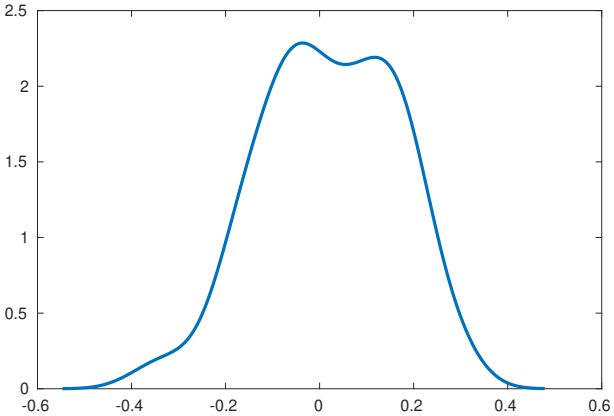

Figure 4.9: Empirical density function of the historical correlation between S&P 500 and Euro/US-Dollar exchange rate, computed with the MATLAB function `ksdensity`.

1. Setting P_0 and $Y_{0,t}, Y_{1,t}, Y_{2,t}$: For the derivation of an initial covariance matrix from the given correlation matrix R_0^{hist} we set $P_0 = HDH^\top$, where D is a diagonal matrix, which entries are the eigenvalues of the estimated covariance matrix of the whole historical data, and H is an orthogonal matrix such that $\|R_0 - R_0^{\text{hist}}\|_F$ is minimized, where $R_0 = \Sigma_0^{-1} P_0 \Sigma_0^{-1}$ with $\Sigma_0 = \big(\text{diag}(P_0)\big)^{1/2}$ (see [51], [72]). This procedure gives us the initial covariance matrix

$$P_0 = \begin{pmatrix} 0.0233 & -0.0005 \\ -0.0005 & 0.0427 \end{pmatrix}. \tag{4.6}$$

For the construction of time-dependent, skew-symmetric matrices $Y_{j,t}$ we multiply time-dependent functions $g_j(t)$, $j = 0, 1, 2$, with the basis matrix of $\mathfrak{so}(2)$, namely

$$C = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

By experimenting with different functions in $Y_{j,t} = g_j(t)C$ we found that

$$g_0(t) = x_1 t \sin(x_2 t), \quad g_1(t) = x_3 + x_4 t, \quad g_2(t) = x_5 + x_6 t, \tag{4.7}$$

worked best with respect to the regarded historical data, where the parameters $x_1, \ldots, x_6 \in \mathbb{R}$ can be associated with the degrees of freedom of this stochastic correlation model. Different approaches for the construction of the matrices P_0 and $Y_{j,t}$ are of course possible and can be adapted according to the given data.

2. SRKMK scheme: Given this initial covariance matrix and the coefficient matrices we can now solve (4.4) by the following variation of Algorithm 3.11.

Algorithm 4.1. *Divide the time interval $[0, T]$ uniformly into L subintervals $[t_\ell, t_{\ell+1}]$, $\ell = 0, 1, \ldots, L-1$ and define $h = t_{\ell+1} - t_\ell$ as the uniform step size. Starting with $t_0 = 0$ and $\Omega_0 = \mathbf{0}$ the following steps are repeated until $t_{\ell+1} = T$.*

(i) *Let P_ℓ be the approximation of P_t at time $t = t_\ell$.*

(ii) *Compute Ω_1 by applying a SRK or an Itô-Taylor scheme to the SDE (4.3).*

(iii) *Define a numerical solution of (4.4) as $P_{\ell+1} = \exp(\Omega_1) P_\ell \exp(-\Omega_1)$.*

We can append the computation of the correlation flow as an additional step:

(iv) Set $R_{\ell+1} = \Sigma_{\ell+1}^{-1} P_{\ell+1} \Sigma_{\ell+1}^{-1}$ with $\Sigma_{\ell+1} = \big(\text{diag}(P_{\ell+1})\big)^{1/2}$.

3. Calibration and Results: From the resulting correlation flow we estimate the density function f^{flow} again by using the MATLAB function `ksdensity`. Then we can calibrate the parameters $x_1, \dots, x_6$ in (4.7) such that the mean-squared error of the density functions is minimized, i.e.

$$\frac{1}{N}\sum_{i=1}^{N}\left(f^{\text{hist}}(z_i) - f^{\text{flow}}(z_i)\right)^2 \to \min,$$

where $f^{\text{hist}}(z)$ and $f^{\text{flow}}(z)$ are evaluated at $N = 100$ equally spaced points.

Following these three steps of creating a correlation flow, where we used the Euler-Maruyama scheme (3.15) in the second step of Algorithm 4.1, we found a mean-squared error of $9.57 \cdot 10^{-4}$ for the parameters

$$(x_1, x_2, x_3, x_4, x_5, x_6) = (6.22, -5.22, 9.88, -5.19, -0.62, -16.63).$$

Figure 4.10 shows that the so-found density function f^{flow} approximates the historical data quite well.

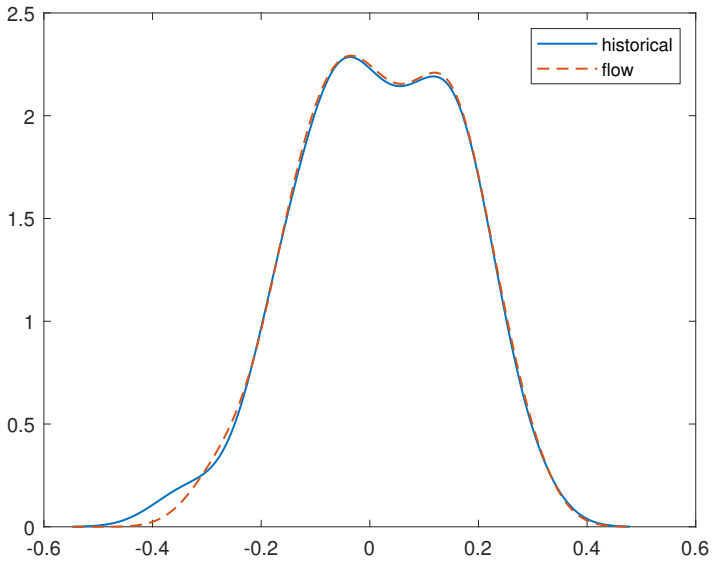

Figure 4.10: Empirical density function of the historical correlation and the correlation flow between S&P 500 and Euro/US-Dollar exchange rate, computed with the MATLAB function `ksdensity`.

Convergence

Applying the Euler-Maruyama scheme in Algorithm 4.1 (ii) gives us a higher strong convergence order than in the previous examples in rigid body modelling due to the diffusion coefficients depending only on time in (4.3). According to Theorem 3.12 we choose the truncation index $q = 0$ in (2.19) which leads to

$$B_{j,t} \approx \operatorname{dexpinv}(\Omega, Y_{j,t}, 0) = \sum_{k=0}^{0} \frac{B_k}{k!} \operatorname{ad}_{\Omega}^{k}(Y_{j,t}) = Y_{j,t}$$

for $j = 1, 2$. In this case we are dealing with *additive* noise since the coefficients are independent of Ω and $Y_{j,t}$ are constructed by multiplying a time-dependent function with the generator of the Lie algebra $\mathfrak{g} = \mathfrak{so}(2)$. This leads to a strong convergence order of $\gamma = 1$ for the geometric Euler-Maruyama scheme (see Figure 4.11).

Moreover, a strong order of $\gamma = 1$ is also obtained if the condition on the truncation index $q \geq 2\gamma - 2$ from Theorem 3.12 is violated in such a way that the truncation index is chosen as $q = 0$ for a SRK method that is supposed to have the strong order $\gamma = 1.5$. This order reduction is also depicted in Figure 4.11, where we applied Rößler's scheme (3.19) with $q = 0$ instead of $q = 1$.

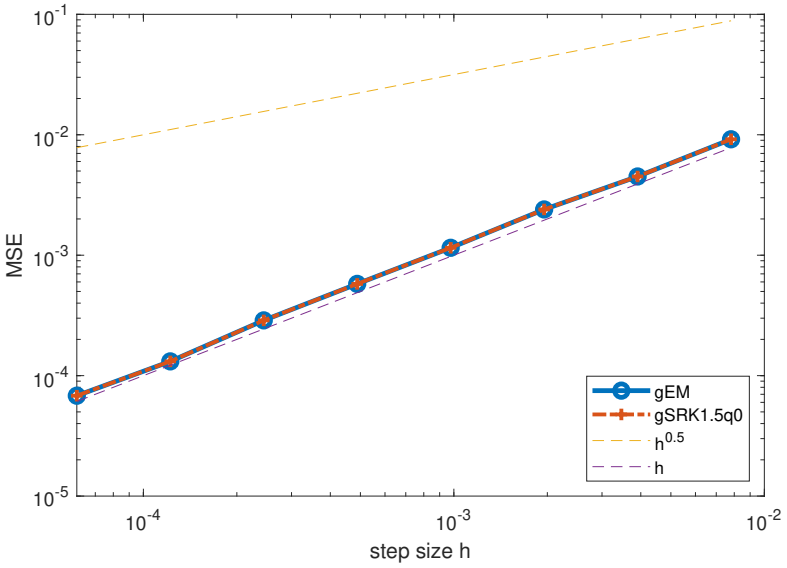

Figure 4.11: Simulation of the mean-square convergence of schemes with $\gamma = 1$ for $M = 1000$ paths.

The fact that setting $q = 1$ in the same scheme leads to the correct convergence order of $\gamma = 1.5$ can be viewed in Figure 4.12. This Figure also shows the strong convergence

order of the SRKMK method that uses the truncated Itô-Taylor scheme (3.17) as the numerical integrator for (4.3) in the second step of Algorithm 4.1, where the needed derivatives for the truncated Itô-Taylor expansion are derived in the Appendix.

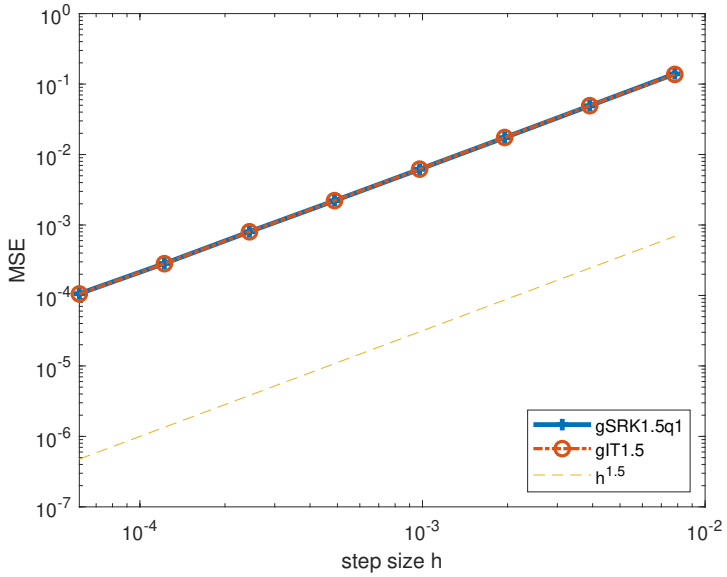

Figure 4.12: Simulation of the mean-square convergence of schemes with $\gamma = 1.5$ for $M = 1000$ paths.

Our reference solution in all plots was obtained by using the SRKMK method of strong order $\gamma = 1.5$ with $q = 1$ and the Lie algebra action $\lambda(\Omega, P) = \operatorname{cay}(\Omega) P \operatorname{cay}(-\Omega)$ for $h = 2^{-16}$.

4.2.2 Rating transition modelling

A rating is an indicator of creditworthiness and agencies like S&P, Moody's and Fitch are required to publish the history of rating changes of some entities of different sectors. After applying the *Aalen-Johansen estimator* (see [38]) or a stochastic reconstruction for the *cohort method* (see [33]) to the historical rating data, one obtains a valid rating transition matrix.

The publication of rating changes happens on a discrete time scale and can be downloaded from the websites of these agencies. However, for many applications, e.g. credit and debit valuation adjustments (CVA and DVA, respectively), it is important to know rating changes on a continuous time scale. For this reason, we model rating transition matrices with a stochastic process to take the uncertainty of the ratings into account and calibrate our model to the historical data published by rating agencies.

Let us focus on the data set from S&P for the corporate sector and group the different ratings to **A**, **B**, **C** and **D**, where **A** is the best and **D** the worst rating. An example of an one year rating transition matrix is given in Table 4.1. It shows for example that

From \ To	**A**	**B**	**C**	**D**
A	0.9395	0.0566	0.0037	2.7804e-04
B	0.0092	0.9680	0.0211	0.0017
C	6.2064e-04	0.0440	0.8154	0.1400
D	0	0	0	1

Table 4.1: Example of an one year rating transition matrix.

the probability of transitioning from rating **A** today to rating **B** in one year is 5.66%. We assume that once an entity has defaulted and has the worst rating **D** that it cannot recover from it and stays in that rating.

Most importantly, we see that the rows of these rating transition matrices sum up to one. Hence, we are dealing with stochastic matrices which form the Lie group $G = \{R \in \mathrm{GL}(n) : R\mathbf{1} = \mathbf{1}\}$ with the corresponding Lie algebra $\mathfrak{g} = \{L \in \mathbb{R}^{n\times n} : L\mathbf{1} = \mathbf{0}\}$, where $\mathbf{1} = (1, \dots, 1)^\top$ and $\mathbf{0} = (0, \dots, 0)^\top$ are vectors in $\mathbb{R}^n$. We refer to [15] for the proof of these geometric properties of transition matrices.

SDE on the Lie group of stochastic matrices

Since we are only interested in stochastic matrices that resemble the transition matrix in Table 4.1, i.e. elements of

$$G_{\geq 0} \coloneqq \{R \in G : R_{ij} \in [0, 1], R_{nj} = (0, \dots, 0, 1) \text{ for } i, j = 1, \dots, n\},$$

we look at the following subset of the Lie algebra,

$$\mathfrak{g}_{\geq 0} \coloneqq \{L \in \mathfrak{g} : L_{ii} \leq 0, L_{ij} \geq 0, i \neq j, L_{nj} = \mathbf{0}^\top\},$$

as applying the matrix exponential to matrices in this subset will give us matrices in $G_{\geq 0}$, see [70, p. 86]. We notice that for a scale of $n = 4$ different ratings, an arbitrary matrix

$L \in \mathfrak{g}_{\geq 0}$ can be represented as a linear combination of $(n-1)^2 = 9$ matrices, namely

$$\mathcal{E}_1 = \begin{pmatrix} -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \mathcal{E}_2 = \begin{pmatrix} -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \mathcal{E}_3 = \begin{pmatrix} -1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

$$\mathcal{E}_4 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \mathcal{E}_5 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \mathcal{E}_6 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

$$\mathcal{E}_7 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \mathcal{E}_8 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \mathcal{E}_9 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

where the coefficients of the linear combination must be nonnegative. Thus, a stochastic process described by (4.3) with A_t and B_t set as linear combinations of $\mathcal{E}_1, \ldots, \mathcal{E}_9$ must have positive and pathwise-increasing coefficients for A_t and B_t to evolve in $\mathfrak{g}_{\geq 0}$. These properties can be fulfilled for example by using jump processes with positive jumps only or processes with stochastic coefficients. We choose the latter and set

$$\begin{aligned} dL_t^{(k)} &= \left| Y_t^{(k)} \right|^{a_k} dt, \\ dY_t^{(k)} &= b_k dt + \sigma_k dW_t, \quad Y_0^{(k)} = 0, \end{aligned} \tag{4.8}$$

such that $L_t = \sum_{k=1}^{(n-1)^2} L_t^{(k)} \mathcal{E}_k$ are continuous stochastic processes in $\mathfrak{g}_{\geq 0}$, where we assume a constant drift b_k, power a_k and volatility σ_k to be positive parameters. Note that if one is interested in the calibration in the risk-neutral measure instead of the historical measure one can apply *Girsanov's theorem* to (4.8) in a well-known linear setting rather than having to apply it in the Lie group of stochastic matrices. More information on the change of measure in this setting of rating matrices can be found in [33].

Now, applying Itô's formula to $\lambda_{R_0} \colon \mathfrak{g}_{\geq 0} \to G_{\geq 0}$, $\lambda_{R_0}(L_t) = R_0 \exp(L_t) =: R_t$ for a fix $R_0 \in G_{\geq 0}$ gives us a stochastic process evolving on the Lie group of stochastic matrices, i.e. valid rating transition matrices R_t. We use the most simple SRKMK scheme, the geometric Euler-Maruyama, to solve the resulting SDE and denote the approximations by $R_{t_{k+1}}^{\mathrm{gEM}} = R_{t_k} \exp(L_{t_{k+1}})$. Note that our approximations fulfill the *Chapman-Kolmogorov* equation by construction.

Calibration and Results

Due to the scarcity of rating data we let a Deep Neural Network, called TimeGAN [78], learn rating distributions from historical data from 2011 until the end of 2019 and denote the output by R_t^{GAN}. For the calibration of the parameters in (4.8) we match the first four moments of R_t^{GAN} and R_t^{gEM}. For more details on the method, calibration and an application to pricing CVA and DVA dependent on rating triggers we refer to [32], [33].

Here, we only take a brief look at some results generated by R_t^{gEM}. Figure 4.13 shows some trajectories of each entry of the calibrated R_t^{gEM} in the course of one year, where the first graph in the first row corresponds to the change of rating from **A** to **A**, the second graph corresponds to the change from **A** to **B** and so on. We omitted the last row since it is always a unit row vector. The cloud of grey lines represents 1000 trajectories with

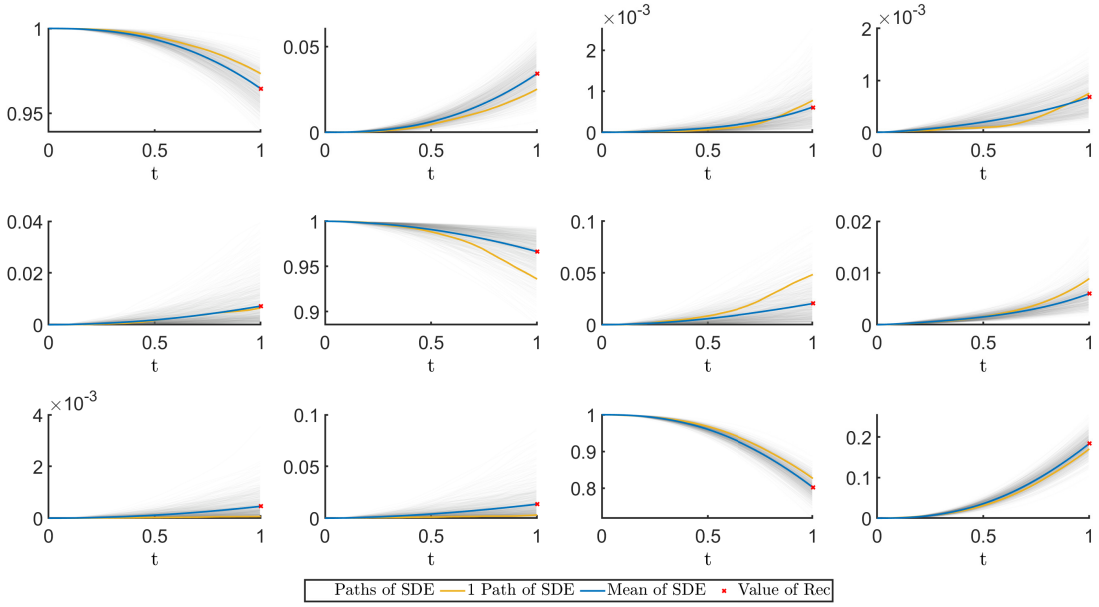

Figure 4.13: Trajectories of calibrated R_t

an exemplary trajectory given as a yellow line. Furthermore, we see a good fit of the mean trajectory (in blue) to the values of the rating transition matrix obtained by the stochastic reconstruction of the cohort method at $t = 1$ year (red dots).

Let us list some reasonable properties of rating transitions with a time span up to one year:

1. It is more likely to stay in the initial rating than changing to another one.

2. It is more likely to be downgraded than upgraded.

3. Lower rated entities are more likely to default.

4. The rating spreads more over time.

We claim that our generated rating transition matrices R_t^{gEM} fulfill these properties perfectly since we were able to make the following observations on the basis of Figure 4.13 and 4.14:

1. The rating transition matrices are strongly diagonally dominant:

$$[R_t(\omega)]_{ii} \geq \sum_{j \neq i} [R_t(\omega)]_{ij}.$$

2. The sum of the upper triangular matrix is bigger than the sum of the lower triangular matrix:

$$\sum_{i<j} [R_t(\omega)]_{ij} \geq \sum_{i>j} [R_t(\omega)]_{ij}.$$

3. The last column is increasing from the best starting rating to the lowest:

$$[R_t(\omega)]_{1n} \leq [R_t(\omega)]_{2n} \leq \cdots \leq [R_t(\omega)]_{nn}.$$

4. The diagonal elements are decreasing over time:

$$[R_s(\omega)]_{ii} \geq [R_t(\omega)]_{ii}$$

for all $s < t$.

The preference of downgrading compared to upgrading can especially be confirmed by viewing Figure 4.14. It shows the histograms of transition probabilities after one year in the calibrated R_t^{gEM} model with a beta distribution fitted to the histograms indicated by the dashed blue line. As expected we see skewed rating distributions with one tail being fatter than the other one.

Overall, we are satisfied with the outcome of our model for rating transitions. Of course one can think of different properties that rating transition matrices should fulfill. However, our model can still be modified and expanded in order to fulfill more properties, e.g. by using time-dependent parameters or adding more diffusion parts in (4.8).

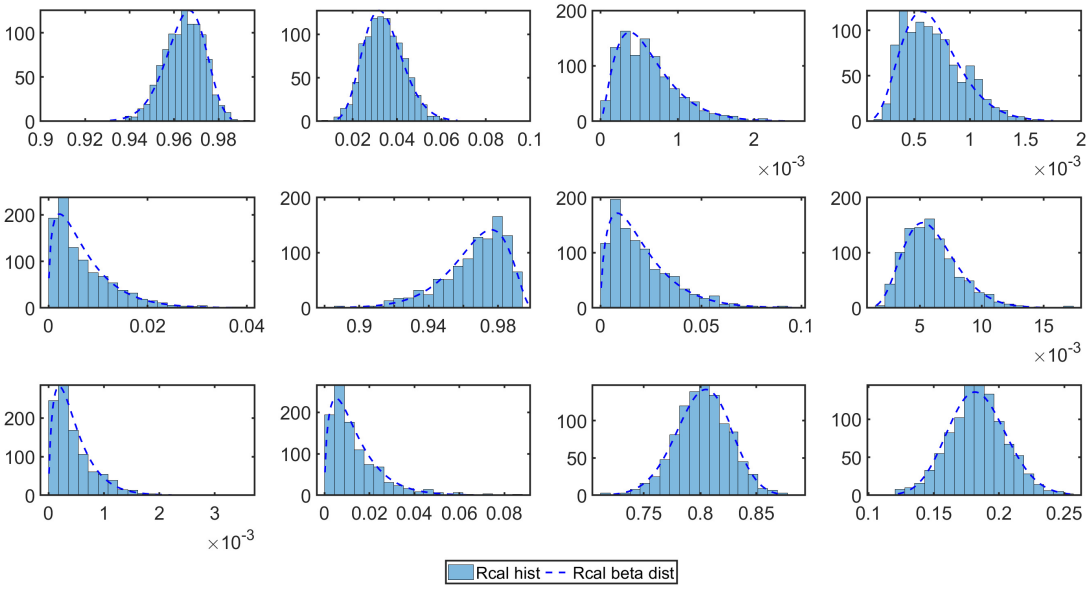

Figure 4.14: Histograms of ratings transition probabilities at 12 months.

Chapter 5

Conclusion

In this thesis we gave an introduction to ODEs on manifolds and how they can be solved numerically via Munthe-Kaas schemes in order to ease the set up of their stochastic counterpart, namely SDEs on manifolds and their approximation by a stochastic expansion of Munthe-Kaas schemes. We illustrated our approach by connecting theoretic results to the application to rigid body motions in each chapter.

Our main result is the analysis of strong convergence of these stochastic Runge-Kutta–Munthe-Kaas (SRKMK) schemes, which we conducted in a seven-steps proof. By analysing an SDE in a corresponding Lie algebra we were able to set up a condition of how many summands of the drift and diffusion part have to be evaluated to obtain a desired strong convergence order.

To demonstrate these theoretic results we applied the stochastic RKMK schemes to SDEs on four exemplary manifolds, namely a sphere, the dual space of the Lie algebra of rigid body motions, the space of symmetric and positive definite matrices and the matrix Lie group of transitions, where the former two manifolds were considered in the modelling of rigid bodies and the latter two in a context of financial mathematics.

In the examples of rigid body modelling we extended the ODEs on manifolds, that are usually considered for the dynamics, such that they account for uncertainties (e.g. in measurements). We visualized that the application of SRKMK schemes compared to conventional SRK schemes has the great benefit of preserving the structure of the manifold. Additional to the manifold structure we have shown that SRKMK schemes can conserve other intrinsic quantities of the underlying system like the Casimir functions whereas the corresponding SRK methods fail at this task.

Regarding the applications in computational finance our contribution is focused on the formulation of stochastic processes such that they intrinsically preserve desired geometric

attributes, for example the symmetry and positive definiteness of correlation matrices or the unit row sums of rating transition matrices. The SDEs on these described manifolds could then be solved by the SRKMK schemes which preserve these geometric properties by construction.

In these examples we have shown the importance of SRKMK schemes and their superiority to the underlying SRK method when it comes to the preservation of geometric characteristics. In a next step one could conduct a comparison to different numerical approximation schemes that are known to preserve geometric properties as well like the stochastic Magnus expansion [34], [77].

Moreover, one could analyse the SRKMK schemes with regard to the preservation of other geometric characteristics like symplecticity or time-reversibility, where we have a stochastic expansion of the work done in [75] based on approaches similar to [42], [76], [80] in mind.

So far, we have only examined manifolds with a matrix Lie group action. We see the expansion to manifolds with Lie group actions that are not necessarily based on subgroups of $\mathrm{GL}(n)$ as future work. This would generalize the application of Munthe-Kaas schemes to a broader class of manifolds, like the one considered in the second example of Section 2.2.2. In this example one could also think of a stochastic expansion of Crouch-Grossman schemes [16].

Another aspect that needs more investigation in the future is the stability and weak convergence of SRKMK schemes. For the analysis of weak convergence one could start from the setting considered in [1], where a proof of weak convergence of Munthe-Kaas methods for the stochastic Landau-Lifshitz equation on a sphere can be found. Based on this proof a generalization to other manifolds can be established.

Thinking of further applications of SRKMK methods in computational finance it might be worth it to look deeper into backward stochastic differential equations (BSDEs) on manifolds since the authors of [35] express a necessity of preserving the positivity of the approximated solution of some pricing problems in finance. As a starting point for constructing SRKMK methods for BSDEs on manifolds one may consider SRKMK methods for BSDEs on Lie groups, which are analysed in [20].

We conclude that despite all these open questions and the related open problems SRKMK methods perform well applied to SDEs on manifolds and are worth to be investigated more in the future.

References

[1] M. Ableidinger and E. Buckwar, "Weak stochastic Runge–Kutta Munthe-Kaas methods for finite spin ensembles," *Applied Numerical Mathematics*, vol. 118, pp. 50–63, 2017.

[2] J. Armstrong and D. Brigo, "Intrinsic stochastic differential equations as jets," *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 474, no. 2210, p. 20 170 559, 2018.

[3] G. B. Arous, "Flots et séries de Taylor stochastiques," *Probability Theory and Related Fields*, vol. 81, pp. 29–77, 1989.

[4] Y. I. Belopolskaya and Y. L. Dalecky, *Stochastic Equations and Differential Geometry*. Dordrecht: Springer Netherlands, 1990, pp. 1–34.

[5] M. Berger and B. Gostiaux, *Differential Geometry: Manifolds, Curves, and Surfaces*. Springer Science & Business Media, 2012, vol. 115.

[6] S. Blanes, F. Casas, J. Oteo, and J. Ros, "The Magnus expansion and some of its applications," *Physics Reports*, vol. 470, no. 5, pp. 151–238, 2009.

[7] A. Bowman and A. Azzalini, *Applied Smoothing Techniques for Data Analysis: The Kernel Approach with S-Plus Illustrations* (Oxford Statistical Science Series). Oxford University Press, 1997.

[8] W. L. Burke, *Applied differential geometry*. Cambridge University Press, 1985.

[9] K. Burrage and P. M. Burrage, "Order conditions of stochastic Runge-Kutta methods by B-series," *SIAM Journal on Numerical Analysis*, vol. 38, no. 5, pp. 1626–1646, 2000.

[10] K. Burrage and P. M. Burrage, "High strong order explicit Runge-Kutta methods for stochastic ordinary differential equations," *Applied Numerical Mathematics*, vol. 22, no. 1-3, pp. 81–101, 1996.

[11] K. Burrage and P. M. Burrage, "High strong order methods for non-commutative stochastic ordinary differential equation systems and the Magnus formula," *Physica D: Nonlinear Phenomena*, vol. 133, no. 1-4, pp. 34–48, 1999.

[12] F. Castell, "Asymptotic expansion of stochastic flows," *Probability theory and related fields*, vol. 96, pp. 225–239, 1993.

[13] F. Castell and J. Gaines, "An efficient approximation method for stochastic differential equations by means of the exponential Lie series," *Mathematics and computers in simulation*, vol. 38, no. 1-3, pp. 13–19, 1995.

[14] E. Celledoni and B. Owren, "Lie group methods for rigid body dynamics and time integration on manifolds," *Computer Methods in Applied Mechanics and Engineering*, vol. 192, no. 3, pp. 421–438, 2003.

[15] C. Coletti, R. Carneiro, and S. Yepes, "Some geometric properties of stochastic matrices," in *Proceeding Series of the Brazilian Society of Computational and Applied Mathematics, 2020*.

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

[17] K. D. Elworthy, *Stochastic Differential Equations on Manifolds* (London Mathematical Society Lecture Note Series). Cambridge University Press, 1982.

[18] M. Emery, *Stochastic Calculus on Manifolds*. Springer Berlin, Heidelberg, 1989.

[19] C. van Emmerich, "Modelling correlation as a stochastic process," *preprint*, vol. 6, no. 03, 2006.

[20] A. Estrade and M. Pontier, "Backward stochastic differential equations in a lie group," *Séminaire de Probabilités XXXV*, pp. 241–259, 2001.

[21] D. Filipović and J. Teichmann, "On the geometry of the term structure of interest rates," *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, vol. 460, no. 2041, pp. 129–167, 2004.

[22] Y. E. Gliklikh, "Stochastic differential equations on manifolds," in *Ordinary and Stochastic Differential Geometry as a Tool for Mathematical Physics*. Dordrecht: Springer Netherlands, 1996, pp. 75–98.

[23] Y. E. Gliklikh, *Global Analysis in Mathematical Physics: Geometric and Stochastic Methods*. New York, NY: Springer New York, 1997.

[24] Y. E. Gliklikh, *Global and stochastic analysis with applications to mathematical physics*. Springer, 2011.

[25] E. Hairer, C. Lubich, and G. Wanner, *Geometric numerical integration* (Springer Series in Computational Mathematics), Second. Springer-Verlag, Berlin, 2006, vol. 31, Structure-preserving algorithms for ordinary differential equations.

[26] E. Hairer, G. Wanner, and S. P. Nørsett, "Runge-Kutta and extrapolation methods," in *Solving Ordinary Differential Equations I: Nonstiff Problems*. Springer Berlin Heidelberg, 1993, pp. 129–353.

[27] B. C. Hall, *Lie Groups, Lie Algebras, and Representations: An Elementary Introduction* (Graduate Texts in Mathematics). New York: Springer, 2010, vol. 222.

[28] P. Holmes, J. Jenkins, and N. E. Leonard, "Dynamics of the Kirchhoff equations I: Coincident centers of gravity and buoyancy," *Physica D: Nonlinear Phenomena*, vol. 118, no. 3, pp. 311–342, 1998.

[29] E. P. Hsu, *Stochastic analysis on manifolds*. American Mathematical Soc., 2002.

[30] A. Iserles, H. Z. Munthe-Kaas, S. P. Nørsett, and A. Zanna, "Lie group methods," *Acta Numerica*, vol. 9, pp. 215–365, 2005.

[31] K. Itô, "Stochastic differential equations in a differentiable manifold," *Nagoya Mathematical Journal*, vol. 1, pp. 35–47, 1950.

[32] K. Kamm and M. Muniz, *A novel approach to rating transition modelling via machine learning and SDEs on Lie groups*, 2022. [Online]. Available: `https://arxiv.org/abs/2205.15699`.

[33] K. Kamm and M. Muniz, *Rating triggers for collateral-inclusive XVA via machine learning and SDEs on Lie groups*, 2022. [Online]. Available: `https://arxiv.org/abs/2211.00326`.

[34] K. Kamm, S. Pagliarani, and A. Pascucci, "On the stochastic Magnus expansion and its application to SPDEs," *Journal of Scientific Computing*, vol. 89, 2020.

[35] L. Kapllani and L. Teng, *Deep learning algorithms for solving high dimensional nonlinear backward stochastic differential equations*, 2022. arXiv: `2010.01319 [math.NA]`.

[36] P. Kloeden and E. Platen, *Numerical Solution of Stochastic Differential Equations* (Stochastic Modelling and Applied Probability). Springer Berlin Heidelberg, 1992.

[37] H. Kunita, "On the representation of solutions of stochastic differential equations," in *Séminaire de Probabilités XIV 1978/79*, Springer, 1980, pp. 282–304.

[38] D. Lando and T. M. Skødeberg, "Analyzing rating transitions and rating drift with continuous observations," *Journal of Banking & Finance*, vol. 26, no. 2, pp. 423–444, 2002.

[39] S. Lang, *Fundamentals of differential geometry*. Springer Science & Business Media, 2012, vol. 191.

[40] J. Lee, *Manifolds and Differential Geometry* (Graduate studies in mathematics). American Mathematical Society, 2009.

[41] N. Lim and N. Privault, "Analytic bond pricing for short rate dynamics evolving on matrix Lie groups," *Quantitative Finance*, vol. 16, no. 1, pp. 119–129, 2016.

[42] Q. Ma, D. Ding, and X. Ding, "Symplectic conditions and stochastic generating functions of stochastic Runge–Kutta methods for stochastic Hamiltonian systems with multiplicative noise," *Applied Mathematics and Computation*, vol. 219, no. 2, pp. 635–643, 2012.

[43] W. Magnus, "On the exponential solution of differential equations for a linear operator," *Communications on Pure and Applied Mathematics*, vol. 7, pp. 649–673, 1954.

[44] S. J. Malham and A. Wiese, "Stochastic Lie group integrators," *SIAM Journal on Scientific Computing*, vol. 30, no. 2, pp. 597–617, 2008.

[45] S. J. Malham and A. Wiese, "Efficient almost-exact Lévy area sampling," *Statistics & Probability Letters*, vol. 88, pp. 50–55, 2014.

[46] G. Marjanovic, M. J. Piggott, and V. Solo, "A simple approach to numerical methods for stochastic differential equations in Lie groups," in *2015 IEEE 54th Conference on Decision and Control (CDC)*, IEEE, 2015, pp. 7143–7150.

[47] G. Marjanovic, M. J. Piggott, and V. Solo, "Numerical methods for stochastic differential equations in the Stiefel manifold made simple," in *2016 IEEE 55th Conference on Decision and Control (CDC)*, IEEE, 2016, pp. 2853–2860.

[48] L. Márkus and A. Kumar, "Comparison of stochastic correlation models," *Journal of Mathematical Sciences*, vol. 237, pp. 810–818, 2019.

[49] J. E. Marsden and T. S. Ratiu, *Introduction to Mechanics and Symmetry: A Basic Exposition of Classical Mechanical Systems*, Second. Springer Publishing Company, Incorporated, 1999.

[50] T. Misawa, "A Lie algebraic approach to numerical integration of stochastic differential equations," *SIAM Journal on Scientific Computing*, vol. 23, no. 3, pp. 866–890, 2001.

[51] M. Muniz, M. Ehrhardt, and M. Günther, "Approximating correlation matrices using stochastic Lie group methods," *Mathematics*, vol. 9, no. 1, 2021.

[52] M. Muniz, M. Ehrhardt, and M. Günther, "Correlation matrices driven by stochastic isospectral flows," in *Progress in Industrial Mathematics at ECMI 2021*, M. Ehrhardt and M. Günther, Eds., Springer International Publishing, 2022, pp. 455–461.

[53] M. Muniz, M. Ehrhardt, M. Günther, and R. Winkler, "Higher strong order methods for linear Itô SDEs on matrix Lie groups," *BIT Numerical Mathematics*, vol. 62, no. 4, pp. 1095–1119, Jan. 2022.

[54] M. Muniz, M. Ehrhardt, M. Günther, and R. Winkler, "Stochastic Runge-Kutta–Munthe-Kaas methods in the modelling of perturbed rigid bodies," *Advances in Applied Mathematics and Mechanics*, vol. 14, no. 2, pp. 528–538, 2022.

[55] M. Muniz, M. Ehrhardt, M. Günther, and R. Winkler, "Strong stochastic Runge-Kutta-Munthe-Kaas methods for nonlinear Itô SDEs on manifolds," *IMACM-Preprint Wuppertal 22/14*, 2022.

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

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

[58] H. Munthe-Kaas, "High order Runge-Kutta methods on manifolds," *Applied Numerical Mathematics*, vol. 29, no. 1, pp. 115–127, 1999.

[59] H. Munthe-Kaas and A. Zanna, "Numerical integration of differential equations on homogeneous manifolds," in *Foundations of Computational Mathematics*, F. Cucker and M. Shub, Eds., Springer Berlin Heidelberg, 1997, pp. 305–315.

[60] B. Oksendal, *Stochastic Differential Equations (5th Ed.): An Introduction with Applications*. Berlin, Heidelberg: Springer-Verlag, 2000.

[61] F. Park, C. Chun, C. Han, and N. Webber, "Interest rate models on Lie groups," *Quantitative Finance*, vol. 11, no. 4, pp. 559–572, 2011.

[62] J. Park and W.-K. Chung, "Geometric integration on Euclidean group with application to articulated multibody systems," *IEEE Transactions on Robotics*, vol. 21, no. 5, pp. 850–863, 2005.

[63] W. Park, Y. Liu, Y. Zhou, M. Moses, and G. S. Chirikjian, "Kinematic state estimation and motion planning for stochastic nonholonomic systems using the exponential map," *Robotica*, vol. 26, no. 4, pp. 419–434, 2008.

[64] M. J. Piggott and V. Solo, "Geometric Euler-Maruyama schemes for stochastic differential equations in SO(n) and SE(n)," *SIAM Journal on Numerical Analysis*, vol. 54, no. 4, pp. 2490–2516, 2016.

[65] A. Rößler, "Explicit order 1.5 schemes for the strong approximation of Itô stochastic differential equations," *Proceedings in Applied Mathematics and Mechanics*, vol. 5, no. 1, pp. 817–818, 2005.

[66] A. Rößler, "Rooted tree analysis for order conditions of stochastic Runge-Kutta methods for the weak approximation of stochastic differential equations," *Stochastic analysis and applications*, vol. 24, no. 1, pp. 97–134, 2006.

[67] A. Rößler, "Strong and weak approximation methods for stochastic differential equations—some recent developments," in *Recent Developments in Applied Probability and Statistics: Dedicated to the Memory of Jürgen Lehn*, L. Devroye, B. Karasözen, M. Kohler, and R. Korn, Eds. Heidelberg: Physica-Verlag HD, 2010, pp. 127–153.

[68] S. Soatto, R. Frezza, and P. Perona, "Motion estimation via dynamic vision," *IEEE Transactions on Automatic Control*, vol. 41, no. 3, pp. 393–413, 1996.

[69] R. Stratonovich, "A new representation for stochastic integrals and equations," *SIAM Journal on Control*, vol. 4, no. 2, pp. 362–371, 1966.

[70] D. Stroock, *An Introduction to Markov Processes* (Graduate Texts in Mathematics). Springer Berlin Heidelberg, 2006.

[71] L. Teng, M. Ehrhardt, and M. Günther, "Modelling stochastic correlation," *Journal of Mathematics in Industry*, vol. 6, pp. 1–18, 2016.

[72] L. Teng, X. Wu, M. Günther, and M. Ehrhardt, "A new methodology to create valid time-dependent correlation matrices via isospectral flows," *ESAIM: Mathematical Modelling and Numerical Analysis (ESAIM: M2AN)*, vol. 54, no. 2, pp. 361–371, 2020.

[73] A. Tocino and R. Ardanuy, "Runge-Kutta methods for numerical solution of stochastic differential equations," *Journal of Computational and Applied Mathematics*, vol. 138, no. 2, pp. 219–241, 2002.

[74] A. Tocino and J. Vigo-Aguiar, "Weak second order conditions for stochastic Runge-Kutta methods," *SIAM Journal on Scientific Computing*, vol. 24, no. 2, pp. 507–523, 2002.

[75] M. Wandelt, M. Günther, and M. Muniz, "Geometric integration on Lie groups using the Cayley transform with focus on Lattice QCD," *Journal of Computational and Applied Mathematics*, vol. 387, p. 112 495, 2021.

[76] P. Wang, J. Hong, and D. Xu, "Construction of symplectic Runge-Kutta methods for stochastic Hamiltonian systems," *Communications in Computational Physics*, vol. 21, no. 1, pp. 237–270, 2017.

[77] Z. Wang, Q. Ma, Z. Yao, and X. Ding, "The Magnus Expansion for Stochastic Differential Equations," *Journal of NonLinear Science*, vol. 30, no. 1, pp. 419–447, Sep. 2019.

[78] J. Yoon, D. Jarrett, and M. van der Schaar, "Time-series generative adversarial networks," in *Advances in Neural Information Processing Systems*, H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett, Eds., vol. 32, Curran Associates, Inc., 2019.

[79] Y. Yuiti, "Stochastic differential equations and nilpotent Lie algebras," *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, vol. 47, no. 2, pp. 213–229, 1979.

[80] W. Zhou, J. Zhang, J. Hong, and S. Song, "Stochastic symplectic Runge–Kutta methods for the strong approximation of Hamiltonian systems with additive noise," *Journal of Computational and Applied Mathematics*, vol. 325, pp. 134–148, 2017.

Appendix A

Derivatives

Let y_0 be fixed in $\mathcal{M}$. Then using the chain rule and the derivative of the matrix exponential $\exp\colon \mathfrak{g} \to G$ found in Section 2.1.2 the first derivative of the Lie algebra action $\lambda_{y_0}\colon \mathfrak{g} \to \mathcal{M}$, $\lambda_{y_0}(\Omega) = \Lambda_{y_0}\big(\exp(\Omega)\big)$ in the direction of an arbitrary matrix $H \in \mathfrak{g}$ is

$$\begin{aligned}\left(\frac{\mathrm{d}}{\mathrm{d}\Omega}\lambda_{y_0}(\Omega)\right) H &= \Lambda'_{y_0}\big(\exp(\Omega)\big)\left(\frac{\mathrm{d}}{\mathrm{d}\Omega}\exp(\Omega)\right) H \\ &= \Lambda'_{y_0}\big(\exp(\Omega)\big)\,\mathrm{dexp}_\Omega(H)\exp(\Omega).\end{aligned} \tag{A.1}$$

For the second derivative we have

$$\begin{aligned}\left(\frac{\mathrm{d}^2}{\mathrm{d}\Omega^2}\lambda_{y_0}(\Omega)\right)(H,\tilde{H}) &= \left(\frac{\mathrm{d}}{\mathrm{d}\Omega}\Lambda'_{y_0}\big(\exp(\Omega)\big)\,\mathrm{dexp}_\Omega(H)\exp(\Omega)\right)\tilde{H} \\ &= \left(\frac{\mathrm{d}}{\mathrm{d}\Omega}\Lambda'_{y_0}\big(\exp(\Omega)\big)\right)\tilde{H}\,\mathrm{dexp}_\Omega(H)\exp(\Omega) \\ &\quad + \Lambda'_{y_0}\big(\exp(\Omega)\big)\left(\frac{\mathrm{d}}{\mathrm{d}\Omega}\mathrm{dexp}_\Omega(H)\right)\tilde{H}\exp(\Omega) \\ &\quad + \Lambda'_{y_0}\big(\exp(\Omega)\big)\,\mathrm{dexp}_\Omega(H)\left(\frac{\mathrm{d}}{\mathrm{d}\Omega}\exp(\Omega)\right)\tilde{H} \\ &= \Lambda''_{y_0}\big(\exp(\Omega)\big)\,\mathrm{dexp}_\Omega(\tilde{H})\exp(\Omega)\,\mathrm{dexp}_\Omega(H)\exp(\Omega) \\ &\quad + \Lambda'_{y_0}\big(\exp(\Omega)\big)\left(\frac{\mathrm{d}}{\mathrm{d}\Omega}\mathrm{dexp}_\Omega(H)\right)\tilde{H}\exp(\Omega) \\ &\quad + \Lambda'_{y_0}\big(\exp(\Omega)\big)\,\mathrm{dexp}_\Omega(H)\,\mathrm{dexp}_\Omega(\tilde{H})\exp(\Omega).\end{aligned} \tag{A.2}$$

Evaluating the derivatives (A.1) and (A.2) in the direction of $v_1^\Omega = \mathrm{dexp}_\Omega^{-1}\left(f_1\big(\lambda_{y_0}(\Omega)\big)\right)$,

i.e. inserting $H = \tilde{H} = v_1^\Omega$, we get

$$\begin{aligned}
\left(\frac{\mathrm{d}}{\mathrm{d}\Omega}\lambda_{y_0}(\Omega)\right) v_1^\Omega &= \Lambda'_{y_0}\big(\exp(\Omega)\big) f_1\big(\lambda_{y_0}(\Omega)\big)\exp(\Omega),\\
\left(\frac{\mathrm{d}^2}{\mathrm{d}\Omega^2}\lambda_{y_0}(\Omega)\right)(v_1^\Omega, v_1^\Omega) &= \Lambda''_{y_0}\big(\exp(\Omega)\big) f_1\big(\lambda_{y_0}(\Omega)\big)\exp(\Omega) f_1\big(\lambda_{y_0}(\Omega)\big)\exp(\Omega)\\
&\quad + \Lambda'_{y_0}\big(\exp(\Omega)\big) f_1'\big(\lambda_{y_0}(\Omega)\big)\Lambda'_{y_0}\big(\exp(\Omega)\big) f_1\big(\lambda_{y_0}(\Omega)\big)\exp(\Omega)\exp(\Omega)\\
&\quad + \Lambda'_{y_0}\big(\exp(\Omega)\big) f_1\big(\lambda_{y_0}(\Omega)\big) f_1\big(\lambda_{y_0}(\Omega)\big)\exp(\Omega).
\end{aligned}$$

As mentioned in Algorithm 3.11 we can also use an Itô-Taylor scheme to approximate the SDE (3.23) in the Lie algebra $\mathfrak{g}$. We provide the derivation of the derivatives needed in the Itô-Taylor scheme of strong order $\gamma = 1.5$ in (3.17). Therefore, we compute the derivatives for the case where the functions $f_j\colon \mathcal{M} \to \mathfrak{g}$, $j = 0, 1, \ldots, d$ are independent of the solution of (3.22). In this case we need the derivatives of the drift and diffusion coefficient in (3.23) in the direction of an arbitrary matrix $H \in \mathfrak{g}$. Let V be an arbitrary matrix in the Lie algebra $\mathfrak{g}$ as well, then we obtain for a manifold $\mathcal{M}$ with a quadratic Lie group acting on it the following first and second derivatives of the Cayley map:

$$\begin{aligned}
\left(\frac{\mathrm{d}}{\mathrm{d}\Omega}\operatorname{dcay}_\Omega^{-1}(V)\right) H &= \left(\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}\Omega}(V + V\Omega - \Omega V - \Omega V\Omega)\right) H\\
&= \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\big(V + V(\Omega + tH) - (\Omega + tH)V\\
&\qquad -(\Omega + tH)V(\Omega + tH)\big)\big|_{t=0}\\
&= \frac{1}{2}(VH - HV - \Omega VH - HV\Omega),\\
\left(\frac{\mathrm{d}^2}{\mathrm{d}\Omega^2}\operatorname{dcay}_\Omega^{-1}(V)\right)(H, H) &= -HVH.
\end{aligned}$$

If we evaluate the first derivative in the direction of $H = \operatorname{dcay}_\Omega^{-1}(V)$, we obtain

$$\left(\frac{\mathrm{d}}{\mathrm{d}\Omega}\operatorname{dcay}_\Omega^{-1}(V)\right)\operatorname{dcay}_\Omega^{-1}(V) = \frac{1}{2}(-V\Omega V - V\Omega V\Omega + \Omega V\Omega V + \Omega V\Omega V\Omega).$$

For non-quadratic Lie groups, we require the derivatives of the drift and diffusion coefficient of (3.23) according to the matrix exponential. Thus, we compute the derivative

of the first five summands of (2.10), namely

$$\begin{aligned}\sum_{k=0}^{4}\frac{B_k}{k!}\operatorname{ad}_\Omega^k(H) &= H-\frac{1}{2}[\Omega,H]+\frac{1}{12}\big[\Omega,[\Omega,H]\big]-\frac{1}{720}\Big[\Omega,\Big[\Omega\big[\Omega,[\Omega,H]\big]\Big]\Big]\\ &= H-\frac{1}{2}(\Omega H-H\Omega)+\frac{1}{12}(\Omega^2H+H\Omega^2-2\Omega H\Omega)\\ &\quad-\frac{1}{720}(\Omega^4H-4\Omega^3H\Omega+6\Omega^2H\Omega^2-4\Omega H\Omega^3+H\Omega^4)\end{aligned}$$

and we will see later why this is enough.

Computing the first derivative we get

$$\begin{aligned}\left(\frac{\mathrm{d}}{\mathrm{d}\Omega}\sum_{k=0}^{4}\frac{B_k}{k!}\operatorname{ad}_\Omega^k(H)\right)\tilde{H} = &-\frac{1}{2}(\tilde{H}H-H\tilde{H})\\ &+\frac{1}{12}(\Omega\tilde{H}H+\tilde{H}\Omega H+H\Omega\tilde{H}+H\tilde{H}\Omega-2\tilde{H}H\Omega-2\Omega H\tilde{H})\\ &-\frac{1}{720}\big(\tilde{H}\Omega^3H+\Omega\tilde{H}\Omega^2H+\Omega^2\tilde{H}\Omega H+\Omega^3\tilde{H}H\\ &-4(\Omega\tilde{H}\Omega H\Omega+\tilde{H}\Omega^2H\Omega+\Omega^2\tilde{H}H\Omega+\Omega^3\tilde{H}H)\\ &+6(\Omega^2H\Omega\tilde{H}+\Omega^2H\tilde{H}\Omega+\Omega\tilde{H}H\Omega^2+\tilde{H}\Omega H\Omega^2)\\ &-4(\Omega H\Omega\tilde{H}\Omega+\Omega H\tilde{H}\Omega^2+\Omega H\Omega^2\tilde{H}+\tilde{H}H\Omega^3)\\ &+H\tilde{H}\Omega^3+H\Omega\tilde{H}\Omega^2+H\Omega^2\tilde{H}\Omega+H\Omega^3\tilde{H}\big).\end{aligned}$$

Whereas the second directional derivative is given by

$$\begin{aligned}&\left(\frac{\mathrm{d}^2}{\mathrm{d}\Omega^2}\sum_{k=0}^{4}\frac{B_k}{k!}\operatorname{ad}_\Omega^k(H)\right)\tilde{H}^2\\ &\quad=\frac{1}{6}(\tilde{H}^2H+H\tilde{H}^2-2\tilde{H}H\tilde{H})\\ &\qquad-\frac{1}{360}\big((\Omega\tilde{H}\Omega\tilde{H}H+\Omega\tilde{H}^2\Omega H+\tilde{H}^2\Omega^2H+\Omega^2\tilde{H}^2H+\tilde{H}\Omega\tilde{H}\Omega H+\tilde{H}\Omega^2\tilde{H}H)\\ &\qquad-4(\Omega\tilde{H}\Omega H\tilde{H}+\Omega\tilde{H}^2H\Omega+\tilde{H}^2\Omega H\Omega+\tilde{H}\Omega^2H\tilde{H}+\tilde{H}\Omega H\Omega+\Omega^2\tilde{H}H\tilde{H})\\ &\qquad+6(\Omega^2H\tilde{H}^2+\Omega\tilde{H}H\Omega\tilde{H}+\tilde{H}\Omega H\Omega\tilde{H}+\Omega\tilde{H}H\tilde{H}\Omega+\tilde{H}\Omega H\tilde{H}\Omega+\tilde{H}^2H\Omega^2)\\ &\qquad-4(\Omega H\Omega\tilde{H}^2+\Omega H\tilde{H}^2\Omega+\tilde{H}H\Omega\tilde{H}\Omega+\Omega H\tilde{H}\Omega\tilde{H}+\tilde{H}H\tilde{H}\Omega^2+\tilde{H}H\Omega^2\tilde{H})\\ &\qquad+(H\tilde{H}\Omega\tilde{H}\Omega+H\tilde{H}^2\Omega^2+H\tilde{H}\Omega^2\tilde{H}+H\Omega\tilde{H}\Omega\tilde{H}+H\Omega\tilde{H}^2\Omega+H\Omega^2\tilde{H}^2)\big).\end{aligned}$$

The evaluation of the derivatives at $\Omega_0=\mathbf{0}$ causes many summands to become zero, which makes computing higher summands ($k>4$) unnecessary.