Alma Mater Studiorum · Università di Bologna

SCUOLA DI SCIENZE
Corso di Laurea (Magistrale) in Matematica

Weak approximation of stochastic dynamics on Lie groups

Tesi di Laurea in Matematica

Relatore:
Chiar.mo Prof.
Stefano Pagliarani

Correlatori:
Chiar.mo Prof.
Hans Z. Munthe-Kaas

Chiar.mo Prof.
Adrien Ange André Laurent

Presentata da:
Marco Cacace

Anno Accademico 2022-2023

Abstract

We propose a new methodology for the creation of high weak order stochastic Lie-Runge-Kutta methods for time homogeneous SDE. More precisely, we establish sufficient conditions for a class of second order elliptic operators to generate a Feller semi-group on a general manifold of bounded geometry.

We give an estimate of a Feller semi-group in terms of its infinitesimal generator, and we use that estimate to derive the order conditions for a weak second order Stochastic Lie Runge-Kutta method on a matrix Lie group.

Keywords: **Lie group methods, Magnus expansion, Runge-Kutta methods, matrix-valued SDEs, Feller semi-groups, diffusion processes**

I Introduction

We are interested in the study of numerical methods for on matrix Lie groups. In particular, we are looking for numerical schemes of weak order 2 i.e. if X_t is the solution of an SDE on a Lie group G,X_N is the numerical method, $x \in G$ and $\delta > 0$ we want X_N to stay on G and that for any $h \in (0, \delta)$

$$\mathbb{E}[\phi(X_h) - \phi(X_N)|X_0 = x] = O(h^2)$$

for any $\phi : G \to \mathbb{R}$ belonging to a suitable test functions space.

Sampling SDEs on a manifold has various application in statistics (see e.g. [22]), and molecular dynamics. In particular, sampling from the constrained overdamped Langevin equation

$$dX_t = -\nabla V(X(t))dt + \sigma dW_t \tag{I.1}$$

allows to compute the so-called free energy, which is a key quantity in thermodynamic (see e.g [53],[17],[47]). Moreover SDEs on manifolds appear in finance in the context of interest rate modelling (see [5] and reference therein).

We will consider the extension to SDEs of two classes of methods used for ODE on a Lie group: The Magnus expansion and the Runge-Kutta methods.

The Magnus expansion provides an exponential expansion for the solution of a linear ODE on a Lie group. Given a matrix Lie group G we can express the solution $Y(t)$ of an ODE on G

$$Y'(t) = A(t)Y(t)$$

as $Y(t) = \text{expm}(\Omega(t))$ for some $\Omega(t)$ in the Lie algebra of G. By using the formula for the derivative of the exponential map it is possible to obtain an ODE for $\Omega(t)$ [48]. Solving such ODE by iteration we can obtain a series expansion for $\Omega(t)$ in terms of iterated commutators of $A(t)$. This integrals can be approximated with some quadrature, in particular, a Gauss Legendre quadrature with μ quadrature points gives us convergence up to order 2μ (see [35]).

The Magnus expansion has been generalized for Stratonovich SDE [74],[13] and more recently to Itô SDE [38],[54]. As pointed out in [38] if the system of SDE is autonomous and there is a single (matrix valued) Brownian motion the solution of the Itô SDE can be expressed in terms of Lebesgue integrals only. Moreover, Itô-stochastic Magnus expansion can be used to efficiently solve stochastic partial differential equations (SPDE) with two space variables numerically [39].

Following [38] we will express the Itô-stochastic Magnus expansion in $GL(n, \mathbb{R})$ up to order 2 and express the first momentum of the solution of the ODE. We will verify how, in the case in which the coefficients of the SDE are constant matrices such expansion coincides with the Talay-Tubaro expansion (see [72])

$$\mathbb{E}[\phi(X_h)|X_0 = I_n] = \sum_{i=1}^{N} \frac{h^i}{i!}\mathcal{L}^i + O(h^{N+1}) \tag{I.2}$$

where $\mathcal{L}$ is the infinitesimal generator of the process.

The Runge-Kutta methods are a class of numerical methods used to approximate the solution of non-linear ODE. Given a real ODE

$$y'(t) = f(t, y(t))$$

an implicit Runge-Kutta method has the form

$$y_{N+1} = y_N + h\sum_{i=1}^{s} b_i k_i$$

where, if we call $c_i = \sum_j a_{ij}$

$$k_i = f\left(t_n + c_i h, h\sum_j^s a_{ij} k_j\right)$$

The value of the constants are obtained by confronting the numerical scheme with the Taylor expansion of the exact solution. The value of the coefficients are usually stored in a Butcher tableau

$$\begin{array}{c|cccc} c_1 & a_{11} & a_{12} & \cdots & a_{1s} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_s & a_{s_1} & a_{s2} & \cdots & a_{ss} \\ \hline & b_1 & b_2 & \cdots & b_s \end{array}$$

The calculation of the order condition becomes cumbersome for high order, that's why John Butcher introduced the algebraic tool of the Butcher-series [15]. The Taylor series of the solution of an ODE can be written as

$$y(h) = y_0 + hf + \frac{1}{2}h^2 f^{'}(f) + \frac{h^3}{6} f^{''}(f, f) + \cdots$$

where the derivatives $f^{(k)}(x)$ of the vector field is regarded as a multilinear map $V^k \to V$.The B-series are formal series of the form

$$B(c, f) = c_0 + c_1 hf + c_2 h^2 f'(f) + c_3 h^3 f^{''}(f, f) + \cdots$$

B-series also arise naturally in other fields of mathematics, Brouder [11] pointed out an important link to the work by Connes and Kreimer [18], which was originally written in the context of renormalization.

The formulas for the composition of RK methods can be described in terms of the operation of the Butcher group, the group of the Butcher series [29]. Moreover, the terms $f^{(k)}(f, \cdots, f)$ can be described in terms of rooted trees. This arises an algebraic structure in the Butcher group, indeed the set of linear combination of rooted trees endowed with the operation of grafting is a pre-Lie algebra [51]. The algebra of rooted forests i.e. multisets of rooted trees is called the Connes-Kreimer algebra [18] and has a structure of Hopf algebra (see e.g. [7] or [63] for the definition and basic properties of Hopf algebras).

It is possible to extend the B-series formalism to include also the divergence operation. The resulting set of trees are called aromatic B-series. While B-series are integrators equivariant under all affine maps [50], the aromatic B-series are the maps equivariant under all the invertible invariant map [56], [7](for the definition of equivariant maps see definition4.29).

The Runge-Kutta methods can be generalized to solve ODEs on Lie groups. This class of methods are called Munthe-Kaas (MK) methods, this methods similarly to the Magnus expansion allow to express an approximation of the solution in terms of iterated commutators (see [55],[34] and references therein).

The number of nested commutators needed increase drastically for methods of high order. Using the theory of free Lie algebras and a symmetry argument it is possible to reduce this number (see theorem 2.29 and section 2.8)

As the classical RK methods can be described by linear combinations of rooted trees the RKMK method can be described in terms of linear combinations of ordered trees: the Lie-Butcher series. The space of such series admits the algebraic structure of post-Lie algebra [57].

In the last decades various generalizations of the Runge-Kutta methods for approximating the solutions of SDEs have been presented. In particular we mention Burrage and Burrage [12, 14] and Komori, Mitsui and Sugiura [42] who first introduced stochastic trees and B-series for studying the order conditions of strong convergence of SDEs.

Other contributions were given by Rößler [65, 66, 67, 68] and Debrabant and Kværnø [19, 20, 21] who design high order weak and strong integrators for SDEs on Euclidean spaces and [2] who has extended to SDEs a result from Sanz-Serna and Abia on canonical Runge-Kutta methods, i.e. methods that preserves the symplettic structure of the phase space of Hamiltonian systems [70]. Moreover, [43] applied tree series to a class of stochastic differential algebraic equations (SDAEs) for the computation of strong order conditions.

In [45] a modification of the aromatic trees formalism, called exotic aromatic trees, is introduced for the study of the accuracy of numerical integrators for the invariant measure of a class of ergodic SDEs with additive noise (see also section 4.4). This algebraic formulation is used to simplify the calculations of the order conditions for RK schemes for weak integrators. Algebraic and geometric properties of exotic aromatic trees is actually a subject of research [9, 44, 10].

In [45] the analysis of the order conditions of the weak RK schemes is done by using the Talay-Tubaro expansion (I.2). We will show that, under some condition on the space of the test functions, an abstract version of such expansion holds for any Feller process (see theorem 5.29).

On a compact Riemannian manifold, under some technical conditions diffusion processes are Feller (see e.g. [33]). On a non-compact Riemannian manifold the Feller property of the Brownian motion or, more general diffusion processes is related to some lower bounds for the Ricci tensor (see [62] and references therein). [49] gives a set of sufficient conditions for an elliptic operator to be the infinitesimal generator of a Feller semi-group when the Riemannian manifold is of bounded geometry. These manifolds include all the compact Riemannian manifolds and all the homogeneous spaces equipped with an invariant metric. In particular, any Lie group endowed with a left (or right) invariant metric is of bounded geometry.

Using this result we give a set of sufficient conditions for the Talay-Tubaro expansion to holds for diffusion on general Lie groups. If the generator associated to the diffusion is a uniformly elliptic and C^∞-bounded operator (definition 6.15 and definition 6.20) is the generator of the Feller semi-group. If moreover we impose some conditions on the set of test functions we will recover the expansion of equation (I.2). In particular, if we choose the set of test functions to be the set of smooth, compacty supported functions over the Lie group all the hypothesis of theorem 5.29 are fulfilled.

In chapter 1 we will define Lie groups and Lie algebras, the Lie group exponential map (denoted as exp) and we will give a formula for the differential of the exponential map. We will focus in particular on matrix Lie groups.

In chapter 2 we will describe the deterministic Magnus expansion and the RKMK methods. We will define the Free Lie algebra of a set and the universal enveloping algebra of a Lie algebra. We will also show how to use such algebraic tools to reduce the number of iterated commutators necessary to describe a s-stage q-th order RKMK method.

in chapter 3 we will describe the Levi-Civita connection associated to a Riemannian manifold (M, g) and the corresponding curvature tensor and geodesics. We will define the Riemannian (or geodesic) exponential at the point $p \in M$ (denoted with Exp_p). in section 3.4 we will define the frame bundle of a manifold and the horizontal lift of a vector field. The results of that section will be extended in section 6.3 to give a characterization of an SDE on a manifold in term of its stochastic development. In section 3.5 we will define the manifolds with bounded geometry and show that any Lie group can be equipped with a metric such that the resulting Riemannian manifold will be of bounded geometry (see Example 3.47).

In chapter 4 we will define the Stratonovich and Itô integral of a semi-martingale and describe their properties. We will also define the multiple stochastic integrals and the Itô-taylor expansion of an Itô process. The multiple integral will appear in the Magnus expansion of SDE driven by more than a single Brownian motion (see [74] for the Stratonovich version and [38] for the Itô version). In section 4.4 we will show the exotic aromatic tree formalism defined in [45].

In chapter 5 we will define Markov and Feller processes and their properties. In particular the Kolmogorov backward equation (5.2). We will show that the set of function for which the abstract version of the series (I.2) converges is dense in the space of continuous functions that vanish at infinity, denoted as C_0 and we will give sufficient conditions for the series to be an approximation of order n.

In chapter 6 we will state the main result of [38] and use it to obtain a weak integrator for an SDE with constant coefficients. We will define (semi)martingales and solution of SDE on a manifold by using the Whitney embedding theorem 1.11 and another approach based on the stochastic development of a process. We will define the diffusion processes on a manifold and state the result of [49] for diffusion on manifold with bounded geometry. In section 6.6 we will use the results of the previous chapters to find a second order Lie Runge-Kutta method for diffusion processes over a matrix Lie group.

II Introduzione

Siamo interessati allo studio dei metodi numerici per gruppi di Lie matriciali. In particolare, stiamo cercando schemi numerici di ordine debole 2, ossia se X_t è la soluzione di un'equazione differenziale stocastica su un gruppo di Lie G, X_N è il metodo numerico, $x \in G$ e $\delta > 0$, vogliamo che X_N rimanga su G e che per ogni $h \in (0, \delta)$ valga

$$\mathbb{E}[\phi(X_h) - \phi(X_N)|X_0 = x] = O(h^2)$$

per ogni $\phi : G \to \mathbb{R}$ appartenente a uno spazio di funzioni di test adeguato.

Campionare equazioni differenziali stocastiche su una varietà ha varie applicazioni in statistica (vedi ad esempio [22]) e nella dinamica molecolare. In particolare, campionare dalla *overdamped Langevin equation* (I.1) permette di calcolare la cosiddetta energia libera, che è una quantità chiave in termodinamica (vedi ad esempio [53],[17],[47]). Inoltre, le equazioni differenziali stocastiche su varietà compaiono nella finanza nel contesto della modellazione dei tassi di interesse (vedi [5] e le relative referenze)

Esamineremo l'estensione alle SDEs di due classi di metodi utilizzati per le ODE su un gruppo di Lie: l'espansione di Magnus e i metodi di Runge-Kutta.

L'espansione di Magnus fornisce una rappresentazione esponenziale per la soluzione di un'ODE lineare su un gruppo di Lie. Dato un gruppo di Lie matriciale G, possiamo esprimere la soluzione $Y(t)$ di un'ODE su G

$$Y'(t) = A(t)Y(t)$$

come $Y(t) = \text{expm}(\Omega(t))$ per un certo $\Omega(t)$ nell'algebra di Lie di G. Utilizzando la formula per la derivata della mappa esponenziale, è possibile ottenere un'ODE per $\Omega(t)$ [48]. Risolvendo tale ODE per iterazione, possiamo ottenere una espansione in serie per $\Omega(t)$ in termini di commutatori iterati di $A(t)$. Questi integrali possono essere approssimati con una quadratura, in particolare una quadratura di Gauss-Legendre con μ punti di quadratura ci fornisce una convergenza fino all'ordine 2μ (vedi [35]).

L'espansione di Magnus è stata generalizzata per SDE di Stratonovich [74],[13] e più recentemente per SDE di Itô [38],[54]. Come evidenziato in [38], se il sistema di SDE è autonomo e c'è un singolo moto Browniano (a valori matriciali), la soluzione dell'SDE di Itô può essere espressa in termini di integrali di Lebesgue. Inoltre, l'espansione di Magnus stocastica di Itô può essere utilizzata per trovare soluzioni numeriche di equazioni differenziali stocastiche alle derivate parziali parziali (SPDE) con due variabili spaziali [39].

Seguendo [38], esprimeremo l'espansione di Magnus stocastica di Itô in $GL(n, \mathbb{R})$ fino all'ordine 2 e calcoleremo il primo momento della soluzione dell'ODE. Verificheremo come, nel caso in cui i coefficienti dell'SDE siano matrici costanti, tale espansione coincida con l'espansione di Talay-Tubaro (I.2) (vedi [72])

I metodi di Runge-Kutta sono una classe di metodi numerici utilizzati per approssimare la soluzione di ODE non lineari. Data un'ODE reale

$$y'(t) = f(t, y(t))$$

un metodo di Runge-Kutta implicito ha la forma

$$y_{N+1} = y_N + h \sum_{i=1}^{s} b_i k_i$$

dove, se chiamiamo $c_i = \sum_j a_{ij}$

$$k_i = f\left(t_n + c_i h, h \sum_{j}^{s} a_{ij} k_j\right)$$

Il valore delle costanti viene ottenuto confrontando lo schema numerico con l'espansione in serie di Taylor della soluzione esatta. I valori dei coefficienti sono solitamente mostrati in una tabella di Butcher:

$$\begin{array}{c|cccc} c_1 & a_{11} & a_{12} & \cdots & a_{1s} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_s & a_{s1} & a_{s2} & \cdots & a_{ss} \\ \hline & b_1 & b_2 & \cdots & b_s \end{array}$$

Il calcolo delle *order conditions* diventa complicato per ordini elevati, ecco perché John Butcher ha introdotto lo strumento algebrico delle serie di Butcher [15]. La serie di Taylor della soluzione di un'ODE può essere scritta come

$$y(h) = y_0 + hf + \frac{1}{2}h^2 f^{'}(f) + \frac{h^3}{6} f^{''}(f, f) + \cdots$$

dove le derivate $f^{(k)}(x)$ del campo vettoriale sono considerate come mappe multilineari $V^k \to V$.Le B-series sono serie formali della forma

$$B(c, f) = c_0 + c_1 hf + c_2 h^2 f'(f) + c_3 h^3 f^{''}(f, f) + \cdots$$

Le B-series emergono anche in modo naturale in altri campi della matematica, Brouder [11] ha evidenziato un importante collegamento con il lavoro di Connes e Kreimer [18], che è stato originariamente scritto nel contesto della rinormalizzazione.

Le formule per la composizione dei metodi di Runge-Kutta possono essere descritte in termini dell'operazione del gruppo di Butcher, i.e il gruppo delle serie di Butcher [29]. Inoltre, i termini $f^{(k)}(f, \cdots, f)$ possono essere descritti in termini di *rooted trees*. Questo genera una struttura algebrica nel gruppo di Butcher, infatti l'insieme delle combinazioni lineari di *rooted trees* dotato dell'operazione di *grafting* è una pre-Lie algebra [51]. L'algebra dele *rooted forests*, cioè dei multinsiemi di *rooted trees*, è chiamata algebra di Connes-Kreimer [18] ed ha una struttura di algebra di Hopf (vedi ad esempio [7] o [63] per la definizione e le proprietà di base delle algebre di Hopf).

È possibile estendere il formalismo delle B-series per includere l'operazione di divergenza. L'insieme risultante di alberi viene chiamato *aromatic B-series*. Mentre le B-series sono integratori equivarianti rispetto a tutte le mappe affini [50], le *aromatic B-series* sono le mappe equivarianti rispetto a tutte le mappe invertibili invarianti [56], [7](per la definizione di mappa equivariante vedi definizione 4.29).

I metodi di Runge-Kutta possono essere generalizzati per risolvere ODE su gruppi di Lie. Questa classe di metodi è chiamata metodi di Munthe-Kaas (MK), e similmente all'espansione di Magnus, permettono di esprimere un'approssimazione della soluzione in termini di commutatori iterati (vedi [55], [34] e relative referenze).

Il numero di commutatori nidificati necessari aumenta drasticamente per metodi di alto ordine. Utilizzando la teoria delle *Free Lie algebras* e un argomento di simmetria, è possibile ridurre questo numero (vedi teorema 2.29 e sezione 2.8).

Come i metodi classici di Runge-Kutta possono essere descritti da combinazioni lineari di *rooted trees*, il metodo RKMK può essere descritto in termini di combinazioni lineari di *ordered trees*: le Lie-Butcher series. Lo spazio di tali serie ammette una struttura algebrica di post-Lie algebra [57].

Negli ultimi decenni sono state presentate varie generalizzazioni dei metodi di Runge-Kutta per approssimare le soluzioni delle SDE. In particolare, menzioniamo Burrage and Burrage [12, 14] e Komori, Mitsui and Sugiura [42] che per primi hanno introdotto gli *stochastic trees* e le *B-series* per lo studio delle *order conditions* della convergenza forte delle SDE. Altri contributi sono stati dati da Rößler [65, 66, 67, 68] e Debrabant and Kværnø[19, 20, 21] che hanno progettato integratori deboli e forti di alto ordine per SDEs e [2] che ha esteso alle SDEs un risultato di Sanz-Serna e Abia sui *canonical Runge-Kutta methods*, ovvero metodi che preservano la struttura simplattica dello spazio delle fasi dei sistemi Hamiltoniani [70]. Inoltre, [43] ha applicato tree series a una classe di stochastic differential algebraic equations (SDAEs) per il calcolo delle *strong order conditions*.

In [45] viene introdotta una modifica del formalismo degli *aromatic trees*, chiamata *exotic aromatic trees*, per lo studio dell'accuratezza degli integratori numerici per la misura invariante di una classe di ergodic SDEs con rumore additivo (vedi sezione 4.4). Questa formulazione algebrica viene utilizzata per semplificare i calcoli delle *order conditions* di integratori deboli Runge-Kutta. Le proprietà algebriche e geometriche degli exotic aromatic trees sono attualmente oggetto di ricerca [9, 44, 10].

In [45] l'analisi delle *order conditions* degli schemi RK deboli viene effettuata utilizzando l'espansione di Talay-Tubaro (I.2). Dimostreremo che, sotto alcune condizioni sullo spazio delle funzioni di test, una versione astratta di tale espansione vale per qualsiasi processo Feller (vedi teorema 5.29).

In una varietà riemanniana compatta, sotto alcune condizioni tecniche, i processi di diffusione sono Feller (vedi ad esempio [33]). In una varietà riemanniana non compatta, la proprietà di Feller del moto Browniano o, più in generale, dei processi di diffusione è correlata a determinati limiti inferiori per il tensore di Ricci (vedi [62] e riferimenti al suo interno). [49] fornisce un insieme di condizioni sufficienti affinché un operatore ellittico sia il generatore infinitesimale di un semigruppo di Feller quando la varietà riemanniana ha geometria limitata. Queste varietà includono tutte le varietà riemanniane compatte e tutti gli spazi omogenei dotati di una metrica invariante. In particolare, ogni gruppo di Lie dotato di una metrica invariante a sinistra (o a destra) ha geometria limitata.

Utilizzando questo risultato, forniamo un insieme di condizioni sufficienti affinché l'espansione di Talay-Tubaro valga per la diffusione su gruppi di Lie generali. Se il generatore associato alla diffusione è un operatore uniformemente ellittico e C^∞-bounded (definizione 6.15 e definizione 6.20) è il generatore del semigruppo di Feller. Se inoltre imponiamo alcune condizioni sull'insieme delle funzioni di test, otterremo l'espansione dell'equazione (I.2). In particolare, se scegliamo l'insieme delle funzioni di test come l'insieme delle funzioni lisce a supporto compatto sul gruppo di Lie, tutte le ipotesi del teorema 5.29 sono soddisfatte.

Nel capitolo 1 definiremo i gruppi di Lie e le algebre di Lie, la mappa esponenziale del gruppo di Lie (denotata come exp) e forniremo una formula per la differenziale della mappa esponenziale. Ci concentreremo in particolare sui gruppi di Lie matriciali.

Nel capitolo 2 descriveremo l'espansione di Magnus deterministica e i metodi RKMK. Definiremo la *Free Lie algebra* di un insieme e la *universal enveloping algebra* di una algebra di Lie. Mostreremo anche come utilizzare tali strumenti algebrici per ridurre il numero di commutatori iterati necessari per descrivere un metodo RKMK di ordine q con s stadi.

Nel capitolo 3 descriveremo la connessione di Levi-Civita associata a una varietà riemanniana (M, g) e il tensore di curvatura corrispondente e le geodetiche. Definiremo l'esponenziale riemanniano nel punto $p \in M$ (indicato con Exp_p). Nella sezione 3.4 definiremo il *frame bundle* di una varietà e la *horizontal development* di un campo vettoriale. I risultati di questa sezione verranno estesi nella sezione 6.3 per fornire una caratterizzazione di una SDE su una varietà in termini del suo *stochastic development*. Nella sezione 3.5 definiremo le varietà con geometria limitata e mostreremo che ogni gruppo di Lie può essere dotato di una metrica tale che la varietà riemanniana risultante sarà di geometria limitata (vedi Esempio 3.47).

Nel capitolo 4 definiremo l'integrale di Stratonovich e Itô di una semimartingala e ne descriveremo le proprietà. Definiremo anche gli integrali stocastici multipli e l'espansione di Itô-Taylor di un processo Itô. L'integrale multiplo comparirà nell'espansione di Magnus delle SDE guidate da più di un solo moto Browniano (vedi [74] per la versione di Stratonovich e [38] per la versione di Itô). Nella sezione 4.4 mostreremo il formalismo degli *exotic aromatic trees* definito in [45].

Nel capitolo 5 definiremo i processi di Markov e Feller e le loro proprietà. In particolare, l'equazione di Kolmogorov backward (5.2). Mostreremo che l'insieme delle funzioni per le quali la versione astratta della serie (I.2) converge è denso nello spazio delle funzioni continue che si annullano all'infinito, indicato come C_0, e daremo condizioni sufficienti affinché la serie sia un'approssimazione di ordine n.

Nel capitolo 6 enunceremo il risultato principale di [38] e lo utilizzeremo per ottenere un integratore debole per una SDE con coefficienti costanti. Definiremo (semi)martingale e soluzioni di SDEs su una varietà utilizzando il teorema di immersione di Whitney 1.11 e un altro approccio basato sullo *stochastic development* di un processo. Definiremo i processi di diffusione su una varietà e enunceremo il risultato di [49] per i processi di diffusione su varietà con geometria limitata. Nella sezione 6.6 utilizzeremo i risultati dei capitoli precedenti per trovare un metodo di Runge-Kutta di Lie di secondo ordine per i processi di diffusione su un gruppo di Lie matriciale.

Contents

1 Theory of Lie groups and Lie algebras

1.1 Differentiable manifolds

We start with a rapid survey on the basic results on differentiable manifolds.

Definition 1.1. *A topological space M is said a n-manifold if it is Haussdorff, with second countable components and locally euclidean, i.e. given any point $p \in M$ $\exists (U, \phi)$ where U is a neighborhood of p and $\phi : U \to \mathbb{R}^n$ is an homeomorphism onto its image.*

*(U,ϕ) is called a **chart** $\phi(p) := (x^1(p), ..., x^n(p))$ is called **system of coordinates** at U. If the system of coordinates has n components the manifold is said to have dimension n*

A topological space X is said paracompact if any open cover has a locally finite refinement i.e. every point of X has a neighbourhood which intersect only finite many sets of the cover. We have the following theorem.

Proposition 1.2 ([46] Theorem 1.15)**.** *Any manifold is paracompact*

We can endow a topological manifold with a differential (or analytical) structure.

Definition 1.3. *Given a topological manifold M an assignment $\mathcal{D} : U \to \mathcal{D}(U)$, where U is an open set in the topology of M is called a differential (analytic) structure if:*

1. *$\forall$ $U \subseteq M$ open set $\mathcal{D}(U)$ is an algebra of real (complex) function containing the constant function*
2. *$\forall$ $U, V : V \subseteq U$ if $f \in \mathcal{D}(U)$, $f|_V \in \mathcal{D}(V)$*

   if $\{V_i\}_{i \in I}$ are open sets and f is a real (complex) function defined in $\bigcup_{i \in I} V_i$ such that $f|_{V_i} \in \mathcal{D}(V_i)$ so $f \in \mathcal{D}(\bigcup_{i \in I} V_i)$
3. *given any $x \in M$ given any point $x \in M$ there is a chart $(U, \phi = \{x^i\})$ such that $x \in U$ and the components of its system of coordinates belongs to $\mathcal{D}(U)$.*

   Moreover given any open set V contained in U and any function f defined in V $f \in \mathcal{D}(V) \Leftrightarrow f \circ \phi^{-1} \in C^\infty(\phi(V))$ ($f \in \mathcal{D}(V) \Leftrightarrow f \circ \phi^{-1}$ is analytic)

$(M, \mathcal{D})$ is called a smooth (analytic) manifold.
*The collection of charts of 3 is called **smooth atlas***

Definition 1.4. *a map $f : M \to N$ between manifolds is said to be smooth at p if given a chart (U, ϕ) at p and a chart (V, ψ) at $f(p)$ such that $f(U) \subseteq V$ and $\psi \circ f \circ \phi^{-1}$ is a smooth map in the sense of the usual calculus.*

By using the paracompactness condition it is possible to prove the existence of a useful family of functions on the manifold

Lemma 1.5 ([46] theorem 2.23)**.** *Given a smooth manifold M and an open cover $(U_\alpha)_{\alpha \in I}$ there exists a family of smooth function $(\psi_\alpha : M \to \mathbb{R})_{\alpha \in I}$ such that*

1. *$0 \leq \psi_\alpha(p) \leq 1$ for any $\alpha \in I$, for any $p \in M$*
2. *$\operatorname{supp}(\psi_\alpha) \subseteq U_\alpha$ for any $\alpha \in I$*
3. *$(\operatorname{supp}(\psi_\alpha))_{\alpha \in I}$ is locally finite*
4. *$\sum_{\alpha \in I} \psi_\alpha = 1$*

*such family is called **smooth partition of unity** subordinate to the open cover.*

Any open subset of a manifold is itself a manifold

Example 1.6 (**Open submanifolds**)**.** *Let M be a manifold $\{(U_a, \phi_a)_{a \in I}\}$ an atlas for M and U an open set in the topology of M. So U is a manifold and $\{(U_a \cap U, \phi_a|_U)_{a \in I}\}$ is an atlas for U*

An important feature of smooth manifold is the existence of a tangent space at any point. This can be viewed as a generalization of the concept of tangent plane of a parametric surface of $\mathbb{R}^3$ and gives us a relation between derivations and "vectors".

Definition 1.7. *Let p be a point of a manifold M, $f : M \to \mathbb{R}$ a smooth map. The **germ** of f at p is defined as $f_p = \{g \in C^\infty(M) : g(x) = f(x)$ for any x in a neighbourhood of p $\}$. The space of germs at p forms an algebra D_p.*

A tangent vector at p is an element $v \in (D_p)^*$ (the algebraic dual of D_p) which is real and satisfies the Leibnitz rule:

$$v(f_p g_p) = f(p)v(g_p) + v(f_p)g(p)$$

The space of tangent vectors at p is denoted as T_pM

We have an useful characterization of the elements of the tangent space in term of smooth curves.

Theorem 1.8 ([46] chapter 3)**.** *If M is an n-dimensional manifold and $(U, \phi = (x^i)_{i=1}^n)$ a chart on p.*

The space T_pM is an n-dimensional vector space with bases

$$\{\frac{\partial}{\partial x^i}|_p : f \in D_p \to \frac{\partial f}{\partial t^i}|_{t^i = x^i(p)} \in \mathbb{R}\}_{i=1..n}$$

Suppose $\rho : [-\epsilon, \epsilon] \to M$ is a smooth curve such that $\rho(0) = p$ so $\rho'(0) := \frac{d\rho^i}{dt}|_{t=0}\frac{\partial}{\partial x^i}$ is a tangent vector at p.

Conversely any $v \in T_pM$ can be written as the derivative of some curve $\rho : \rho(0) = p$

We end this section with a characterization of the manifolds given by the Whitney embedding theorem. This will allow us to consider any manifold as an embedded subspace of $\mathbb{R}^N$ for N big enough.

Given a function between manifold we define its **differential** at p as

$$df_p : T_pM \to T_{f(p)}N \quad : df_p(v)(g) = v(g \circ f) \tag{1.1}$$

Definition 1.9. *Define the total differential of the map f is defines as $df : TM \to TN$ such that $df((p, X)) = df_p(X_p)$*

Definition 1.10. *Let $f : M \to N$ a smooth map between manifold.*

- *It is called an **immersion** if its differential is everywhere injective.*
- *It is called an **submersion** if its differential is everywhere surjective.*
- *It is called an **embedding** if it is an immersion and an homeomorphism onto its image.*

  in this case the image $f(M)$ is called an embedded submanifold.

Theorem 1.11 (**Whitney embedding theorem**, ([46] theorem 6.15))**.** *Let M a manifold of dimension n so there is an embedding between M and $\mathbb{R}^{2n}$*

1.2 Vector fields and flows

The disjoint union of all the tangent spaces $TM = \bigsqcup_{p \in M} T_pM$ is called the tangent bundle

Definition 1.12. *A function $X : M \to TM$ such that $X(p) \in T_pM$ $\forall p \in M$ is called vector field. The space of all the vector fields over M is denoted as $\mathcal{H}(M)$ If there exist smooth vector fields $\{X_1, \cdots X_n\}$ such that, for any $p \in M$ $X_1(p), \cdots X_n(p)$ span T_pM the manifold is called parallelilzable.*

We give some notable examples of vector fields:

Example 1.13. *If $(U, (x^i))$ is a chart of M we have that $\mathcal{H}(U)$ is spanned by $\{\frac{\partial}{\partial x^i} \in TM\}$, where $\frac{\partial}{\partial x^i}(p) = \frac{\partial}{\partial x^i}|_p$. This set of vector fields is called the coordinate vector fields for X in U*

Example 1.14. *Given $X, Y \in \mathcal{H}(M)$ the lie bracket of the two vector fields is defined as $[X, Y] = XY - YX$. Let $p \in M$. $(U, (x^i))$ a chart containing p. If $X_p = X^i \frac{\partial}{\partial x^i}|_p$, $Y_p = Y^j \frac{\partial}{\partial x^j}|_p$ the coordinate expression for the Lie bracket is:*

$$[X, Y] = (X^i Y^i - Y^i X^i)\frac{\partial}{\partial x^i} \tag{1.2}$$

So the Lie bracket of two vector fields is a vector field. This implies that $(\mathcal{H}(M), [., .])$ is an algebra.

Here we have usedthe Einstein summation convention i.e. if $\alpha = (i_1, \cdots, i_N)$ is a multi-index

$$A_\alpha B^\alpha := \sum_{i_1} \cdots \sum_{i_N} A_{(i_1,\cdots,i_N)} B^{(i_1,\cdots i_N)}$$

unless specified we will always use this convection. We can now define what it means for an ODE to evolve on the manifold

Definition 1.15. *Let M a manifold and $X = X \in \mathcal{H}(M)$. A differential equation evolving on the manifold is an equation of the form:*

$$y' = X(t, y) \tag{1.3}$$

with $y(0) = y_0 \in M$

$y(t)$ is called an **integral curve** of X passing through y_0. By expressing the vector field in coordinates in a neighbourhood of y_0 we obtain an ODE in $\mathbb{R}^n$. If the vector field is smooth in such neighbourhood we have, for the standard results of existence and uniqueness of Odes:

Theorem 1.16. *Given a smooth vector field for each point $y_0 \in M$ exists $\epsilon > 0$ and a curve $y : (-\epsilon, \epsilon) \to M$ that is solution of (1.3). Any other solution of such equation agreed in their common domain.*

If all the integral curves of a vector field can be extended over all $\mathbb{R}$ it is called a **complete vector field**.

Theorem 1.17 ([46] theorem 9.16)**.** *Every compactly supported smooth vector field is complete. In particular every smooth vector field on a compact manifold is complete.*

Let return to equation (1.3). there exists an operator $\Psi_{t,X} : M \to M$ such that

$$y(t) = \Psi_{t,X}(y_0)$$

This operator is called the **flow** of the vector field X.

Proposition 1.18. *The flow of a vector field satisfies the following properties:*

$$\Psi_{0,X}(y_0) = y_0$$

$$\Psi_{t,X} \circ \Psi_{s,X} = \Psi_{t+s,X}$$

$$\Psi_{t,X} = \Psi_{1,tX}$$

In particular properties the first and the second properties imply that if X is complete for any fixed y_0, $\Psi_{t,X}$ is a group in the time parameter.

Proof. The first property follows by the definition of flow, while the second follows because $\Psi_{t,X} \circ \Psi_{s,X}(y_0)$ and $\Psi_{t+s,X}(y_0)$ are both the integral curve passing through y_0 at the time $t + s$.

The third property follows by rescaling the vector field. □

Moreover given any operator $\theta : \mathbb{R} \times M \to M$ which satisfies the first and the second properties for any $p \in M$ we have that such operator is the flow of a vector field, called the **infinitesimal generator** of the flow. This vector field can be obtained by differentiation

$$X(y_0) = \frac{d}{dt}\theta(t, y_0)|_{t=0}$$

We conclude this section with a theorem on the commutator of flows of vector fields.

Theorem 1.19 ([34] equation 2.4)**.** *Let X, Y two complete vector fields,$\Psi_{s,X}, \Psi_{t,Y}$ their flows.*

Define $\Phi_{s,t}$ as

$$\Phi_{s,t} = \Psi_{s,X} \circ \Psi_{t,Y} \circ \Psi_{-s,X} \circ \Psi_{-t,Y}$$

For small s, t we have:

$$\Phi_{s,t}(y_0) = y_0 + st[X, Y] + O(s^2 t) + O(st^2)$$

Corollary 1.20. *The flows associated to two complete vector fields X and Y commute if and only if $[X, Y] = 0$*

1.3 Lie groups and Lie algebras

We start with some basic definition and theorems on Lie groups and Lie algebras.

Definition 1.21. *A Lie group is a group $(G, +, i)$ endowed with a differential structure that is compatible with the operation of the group, i.e. the addition and the inverse map are smooth maps in the differentiable structure of G*

Definition 1.22. *A lie algebra $(\mathfrak{g}, [.,.])$ over the field K is an algebra whose operation satisfies the following conditions:*
for each x,y,$z \in \mathfrak{g}$

- ***skew-symmetry**: $[x, y] = -[y, x]$*
- ***Jacobi identity** $[x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0$*

The Lie algebra is said finite dimensional if it is finite dimensional as a vector space. Unless specified in this paper any Lie algebra is supposed to be finite dimensional.

A map $f : \mathfrak{g} \to \mathfrak{h}$ between two Lie algerbas is said a Lie algebra homomorphism if it preserves the algebra operation i.e.

$$[f(X), f(Y)] = f([X, Y])$$

Suppose $\{e_i\}_{i=1}^n$ are a bases for a Lie algebra (thought as a vector space). It is possible to express the element $[e_i, e_j]$ as a linear combination of the elements of the bases. Combining the definition of a bases of a vector space and definition 1.22 we obtain the following:

Proposition 1.23. *Given a lie algebra $(\mathfrak{g}, [.,.])$ over a field K, given a bases $\{e_i\}_{i=1}^n$ for $\mathfrak{g}$, $\exists!\{c_{ij}^k\}$ such that*

$$[e_i, e_j] = c_{ij}^k e_k$$

Moreover such constants satisfies

- $c_{ij}^p + c_{ji}^p = 0$
- $c_{mr}^p c_{kh}^r + c_{kr}^p c_{hm}^r + c_{hr}^p c_{mk}^r = 0$

Example 1.24. *The algebra $(\mathcal{H}(M), [.,.])$ of example 1.14 is a Lie algebra*

Example 1.25. *Let V a finite dimensional vector space over K. The space $\mathfrak{gl}(V)$ of the endomorphisms of V endowed with the Lie brackets is a Lie algebra over K. In particular if $V = M(n, K)$, the space of $n \times n$ matrices in K The space is called $\mathfrak{gl}(n, K)$*

It is possible to associate to any Lie group G a lie algebra g. There are two equivalent characterizations of that: as the algebra of left invariant vector fields over G or as the tangent space at the identity T_eG.

Definition 1.26. *Let G a Lie group and call $L_x : G \to G \ : \ L_x(y) = xy$. A vector field X is left invariant if $dL_p(X) = X$ for any $p \in G$, where dL_p is the differential of definition 1.9 ($(dL_p)|_h : T_hG \to T_{ph}G$).*

By using the coordinate expression of the Lie bracket we obtain that

$$dL_p([X, Y]) = [dL_p(X), dL_p(Y)] = [X, Y]$$

So the set of all the left invariant vector fields forms a Lie subalgebra of $(\mathcal{H}(G), [.,.]) =$ that we call the Lie algebra associated to the Lie group.

We have the following theorem

Theorem 1.27 ([73] theorem 2.3.1)**.** *Let G be a Lie group, $e \in G$ the identity and $\mathfrak{g}$ its Lie algebra. The map $X \to X_e$ from $\mathfrak{g}$ to T_eG is an isomorphism. In particular, $\dim(G) = \dim(\mathfrak{g})$.*

using theorem 1.27 it is possible to define a global section of the tangent bundle of the Lie group, indeed if we consider a basis $\{S_i\}$ for the Lie algebra the vectors

$$(X_i)_p = dL_p(S_i)$$

forms a bases for T_pG for any $p \in G$. So we have proven that

Proposition 1.28. *Any Lie group is parallelizable*

by using the results of section 1.2 it is possible to express the commutator of a Lie algebra in terms of derivative of integral curves of the vector fields.

Proposition 1.29. *Let $u, v \in \mathfrak{g}$ and X, Y the associated left invariant vector fields. If $f(t)$, $h(s)$ the integral curves of that vector fields passing through the identity*

$$[u, v] = [X, Y] = \frac{\partial}{\partial t \partial s}(f(t)h(s)f(t)^{-1})_{t=s=0}$$

(this can be proven by using the Leibniz rule and $\frac{\partial}{\partial t}(f(t)^{-1})_{t=0} = -X$)

We end this section with some basic results on the homomorphisms of Lie groups and Lie algebras

Proposition 1.30 ([69] proposition 2.9)**.** *Let G_1,G_2,G_3 Lie groups and $\mathfrak{g}_1$,$\mathfrak{g}_2$,$\mathfrak{g}_3$ the corresponding Lie algebras. Call e_i the identity in the Lie group G_i.*

Let $\phi : G_1 \to G_2$, $\psi : G_2 \to G_3$ Lie group homomorphisms.

1. *$d\phi_{e_1} : \mathfrak{g}_1 \to \mathfrak{g}_2$ is a Lie algebra homomorphism.*
2. *$d(\psi \circ \phi)_{e_2} = d\psi_{e_2} \circ d\phi_{e_1}$*
3. *$\mathrm{id} : G_i \to G_i$ is the identity implies that $d(\mathrm{id})_{e_i}$ is the identity of $\mathfrak{g}_i$*
4. *If ϕ is an isomorphism of Lie groups, then $d\phi$ is an isomorphism of Lie algebras*
5. *$If d\phi$ is isomorphism of Lie algebras, then $\ker(\phi)$ is a discrete abelian subgroup of G*

Corollary 1.31. *The Lie algebra of a subgroup H of G is a subalgebra of the Lie algebra of G*

Proof. The inclusion $i : H \to G$ is a Lie group homomorphism and restricted to its image is the identity. For proposition 1.30 its derivative is a Lie algebra homomorphism. Restricted to its image is the identity so $\mathfrak{h} \subseteq \mathfrak{g}$. □

1.4 Matrix Lie groups

We now characterize an important class of Lie groups and correspondent lie algebras, called the matrix lie groups or the linear Lie groups We will use these 2 technical lemmas:

Lemma 1.32 ([46] chapter 3)**.** *Let U an open set of an m dimensional manifold M. For each $p \in M$ we have:*

$$T_pU \cong T_pM \cong \mathbb{R}^m$$

Lemma 1.33. *Let M an n-dimensional manifold, N an k-dimensional manifold with $n > k$ Let $f : M \to N$ a submersion and let $c \in N$ a regular value i.e. a value such that the differential of x is not zero for any value of $f^{-1}(c)$. So $f^{-1}(c)$ is a manifold*

Proof. Let $p \in f^{-1}(c)$

By the submersion theorem ([46] theorem 4.12) there are U neighbourhood of p in M, V neighbourhood of c such that $f(U) \subseteq V$, $f(x^1, ...x^n) = (x^1, ..., x^k)$ and c correspond to 0 in this coordinate system.

In this coordinate system the first k coordinates of $f^{-1}(c) \cap U$ are 0. So If we call $\phi = (x^{k+1}, ..x^n)$ We obtain a diffeomorphism between $f^{-1}(c) \cap U$ and an euclidean space. □

Example 1.34. ***The general linear group*** $GL(n, \mathbb{R})$

$$GL(n, \mathbb{R}) = \{A \in M_n(\mathbb{R}) : \det(A) \neq 0\}$$

Consider the space of the $n \times n$ matrices $M_n(\mathbb{R})$. There is an isomorphism between it and $\mathbb{R}^{n^2}$ (this two spaces can be identified by stringing all the coefficents of the matrices on a single row).

So $M_n(\mathbb{R})$ is a n^2 dimensional manifold. $GL(n, \mathbb{R})$ is an open subset of that so it is a n^2 dimensional manifold as well. As it is a group w.r.to matrix multiplication it is a Lie group.

By lemma 1.32 we have that its Lie algebra $T_1(GL(n, \mathbb{R})$ is isomorphic (as a vector space) to $M_n(\mathbb{R})$ (here identified as $\mathbb{R}^{n^2}$). To find the multiplication consider the characterization of the lie algebras of theorem 1.27. If $A = (a_i^j) \in M_n(\mathbb{R})$ we have that the left invariant vector field *corresponding to A can be written as $A_p = dL_p A_{I_n}$ for each $p \in GL(n,\mathbb{R})$. Writing them in coordinates we obtain*

$$A_p = x^{ij} a_j^k \frac{\partial}{\partial x^{ik}}$$

So the the Lie bracket induced by $\mathcal{H}(M)$ (equation (1.2)) is

$$[A,B]_p = x^{ij}(a_j^r b_r^k - b_j^r a_r^k)\frac{\partial}{\partial x^{ik}} \tag{1.4}$$

So the multiplication of $M_n(\mathbb{R})$ is the commutator of the two matrices i.e. we have $T_{I_n}(GL(n,\mathbb{R})) \cong \mathfrak{gl}(n,\mathbb{R})$ as a lie algebra.

Example 1.35. ***The special linear group*** $SL(n,\mathbb{R})$

$$SL(n) = \{A \in GL(n,\mathbb{R}) : \det(A) = 1\}$$

Consider the function $f : M_n(\mathbb{R}) \to \mathbb{R} : f(A) = \det(A) - 1$. It is a smooth function because $\det$ is a polynomial function in the coordinates and $f^{-1}(0) = SL(n)$. To show that 0 is a regular value we want to calculate $d(\det)_A(X)$, $A \in SL(n,\mathbb{R})$. If I_n is the identity matrix we have

$$\det(A - tX) = \det(A^{-1})t^n \det(t^{-1}I_n - A^{-1}X) = t^n P_{A^{-1}X}(t^{-1}) \tag{1.5}$$

Where $P_{A^{-1}X}(t)$ is the characteristic polynomial of $A^{-1}X$.

Because $df_A(X) = \frac{\partial}{\partial t}|_{t=0} \det(A - tX) = tr(A^{-1}X)$ we have that 0 is a regular value for f (take $X = A$).

So by lemma 1.33 $SL(n,\mathbb{R})$ is a smooth manifold.

Equation (1.5) also give us that $d(\det)_{I_n}(X) = tr(X)$. So the associated lie algebra is

$$\mathfrak{sl}(n,\mathbb{R}) = T_{I_n}(SL(n,\mathbb{R})) = \{A \in GL(n,\mathbb{R}) : tr(A) = 0\}$$

Example 1.36. ***The orthogonal group*** $O(n,\mathbb{R})$

$$O(n) = \{A \in GL(n,\mathbb{R}) : AA^T = I_n\}$$

The space of symmetrical matrices $S_n(\mathbb{R})$ is a $\frac{n(n+1}{2}$-dimensional vector space so we can define of it a differential structure as done for $M_n(\mathbb{R})$.

Consider the function $f : M_n(\mathbb{R}) \to S_n(\mathbb{R}) : f(A) = AA^T$. This function is smooth and $f^{-1}(I_n) = O(n)$. Let $B \in O(n)$. By taking a curve $\gamma : (-1,1) \to O(n)$: $\gamma(0) = B$ we obtain.

$$df_B(X) = XB^T + BX^T \tag{1.6}$$

So $df_B \neq 0$ and $O(n)$ is a manifold.

Applying equation 1.6 we obtain that the Lie algebra is the set of skew symmetric matrices.

$$T_{I_n}(O(n)) := \mathfrak{so}(n) = \{A \in Gl(n) : A + A^T = 0\}$$

Example 1.37. ***The special orthogonal group*** $SO(n)$

$$SO(n) = SL(n) \cap O(n)$$

By the continuity of the determinant function and the fact that $\det(A) = \pm 1$ for any matrix in $O(n)$ we obtain that $SO(n)$ is one of the two connected components of $O(n)$. So it is a manifold.

The Lie algebra is the same of $O(n)$.

We have the following topological result on the Lie groups $O(n)$ and $S = (n)$.

Proposition 1.38 ([31] lemma 2.1.4)**.** *The group $O(n)$ and the group $SO(n)$ are compact*

We end the section with the characterization of the derivative of the left(right) action of a matrix Lie group

Lemma 1.39. *Let G a matrix Lie group, $p, q \in G$. $L_p(q) = pq$, $R_p(q) = qp$. For any $A \in M_n(\mathbb{R})$*

$$d(L_p)_e(A) = pA$$

$$d(R_p)_e(A) = Ap$$

Proof.

$$d(L_p)_e(A) = \lim_{t\to 0}\frac{1}{t}(L_p(e+tv) - L_p(e)) = pv$$

The proof for the right action can be done in similar way ◻

1.5 The exponential map

Lemma 1.40 ([69] chapter 3 proposition 2.11)**.** *Let G a Lie group with Lie algebra $\mathfrak{g}$ for any $v \in \mathfrak{g}$ $\exists!\ \phi : \mathbb{R} \to G$ Lie group homomorphism such that $d(\phi)_e(\frac{d}{dt}) = v$*

A first consequence of this lemma is that any left invariant vector field is complete.

Proposition 1.41 ([69] 3 chapter corollary 2.12)**.** *Any left invariant vector field is complete*

So if we call X the left invariant vector field associated to $v \in g$ we have that $\exists!\ \phi_X : \mathbb{R} \to G$:

$$\begin{cases} d\phi_X(\frac{d}{dt}) = X \\ \phi'_X(\tau) = X_{\phi_X(\tau)} \\ \phi'_X(0) = v = X_e \end{cases} \tag{1.7}$$

Definition 1.42. *The exponential function* $\exp$: $g \to G$ *is the function that to each $v \in g$ associates the $\phi_X(1)$ where ϕ_X satisfies the equation* (1.7)

ϕ_X is the integral curve of X passing through the identity so by the third property of theorem 1.18 we obtain that $\phi_X(t) = \exp(tX)$. In particular $\exp(0) = e$

By using the properties of the flows described in section 1.2 we immediately obtain that the exponential map is a group homomorphism.

Lemma 1.43. *Let G a Lie group and $\mathfrak{g}$ its Lie algebra, let $v \in \mathfrak{g}$. The map $t \to \exp(tv)$ is a group homomorphism between $\mathbb{R}$ and G. In particular:*

$$\exp[(s+t)v] = \exp(sv)\exp(tv)$$

$$(\exp(v))^{-1} = \exp(-v)$$

An important property of the exponential map is that near the origin is a local diffeomorphism.

Theorem 1.44. *Let G a Lie group and $\mathfrak{g}$ its Lie algebra. $d(\exp)_e = Id_G$ so the exponential map is a local diffeomorphism near the origin.*

Proof. Let $v \in \mathfrak{g}$ and $\phi_X(t) = \exp(tv)$ as in definition 1.42. By differentiating each member in $t = 0$ we obtain:

$$\phi'_X(0) = v = \frac{d}{dt}(\exp(tv))_{t=0} = d(\exp)_e(v)$$

□

Given $p \in G$, $X \in g$ if we define $\xi_p(t) = p\exp(tX)$ Using the fact that X is left invariant and $\exp(0) = e$ we obtain that ξ_p is the integral curve of X through the point p so given any $f \in C^\infty$ in a neighbourhood of p we obtain

$$X_p f(x) = \frac{d}{dt} f(\xi_p(t))_{t=0} = \frac{d}{dt} f(x\exp(tX))_{t=0} \tag{1.8}$$

We now turn our attention on matrix Lie group. In this case the exponential function can be expressed in terms of a convergent power series of elements of $M_n(\mathbb{R})$

Lemma 1.45. *Let* $\text{expm} : M_n(\mathbb{R}) \to M_n(\mathbb{R})$: $\text{expm}(A) = \sum_{k=0}^{\infty} \frac{A^k}{k!}$. *We have that:*

1. *The series is absolutely convergent*
2. $\text{expm}(M_n(\mathbb{R})) \subseteq GL(n, \mathbb{R})$
3. *for any $v \in M_n(\mathbb{R})$ the map $\gamma : \mathbb{R} \to GL(n, \mathbb{R})$: $\gamma(t) = \text{expm}(tv)$ is the integral curve of the left invariant vector field associated to v passing through the identity.*

Proof. the first statement is proven in chapter 3.1 of [31]. The second is lemma 3.2.2 of [31]. The last statement can be proven as follow: by theorem 3.2.6 of [31] γ solves the initial value problem $\gamma(0) = e$, $\gamma'(t) = \gamma(t)v$. Let X the left invariant vector field associated to v. By using that $X_p = dL_p v$ and lemma 1.39 we obtain the claim. □

With this lemma is straightforward to prove that **the matrix exponential is the exponential function of a matrix Lie group** (by using the uniqueness in lemma 1.40)

1.6 The adjoint representation of a Lie group and its Lie algebra

A representation of a group over a vector space V is a group G homomorphism $\pi : G \to Gl(V)$ from the group to the group of automorphisms over V.

A representation of a Lie algebra g in a finite dimensional vector space V is a lie algebra homomorphism between g and $\mathfrak{gl}(V)$

In section 1.3 we have defined the Lie algebra of a Lie group as the set of left invariant vector field. It is possible to define it as the set of right invariant vector field as well.

The Adjoint representation of the Lie group is the map $\mathrm{Ad} : G \to \mathfrak{gl}(\mathfrak{g})$ which allow us to change between the left invariant representation and the right invariant representation

Definition 1.46. *Let G a Lie group, $\mathfrak{g}$ its Lie algebra. For any $p \in G$ define the inner automorphism $\Psi_p : G \to G$: $\Psi_p(h) = php^{-1}$.*

Define $\Psi : G \to \mathrm{Aut}(G)$: $\Psi(p) = \Psi_p$. The map $\mathrm{Ad}_p = d\Psi_p : \mathfrak{g} \to \mathfrak{g}$ is an automorphism of $\mathfrak{g}$.

The map $\mathrm{Ad} : G \to GL(\mathfrak{g})$: $\mathrm{Ad}(p) = \mathrm{Ad}_p$ is a representation for the Lie group called the adjoint representation.

In terms of left and right multiplication we have that

$$\mathrm{Ad}(p)(v) = d(L_p \circ R_{p^{-1}})(v)$$

If $p \in G$ and $v \in \mathfrak{g}$ and X the left invariant vector field associated to v. Call γ the integral curve of X passing through the identity. The adjoint representation can be expressed as

$$\mathrm{Ad}_p(X_e) = \frac{d}{dt}(p\gamma(t)p^{-1})|_{t=0}$$

Example 1.47. *the adjoint representation of a matrix Lie group is given by*

$$\mathrm{Ad}_P(A) = PAP^{-1} \tag{1.9}$$

Indeed by lemma 1.39 and the chain rule we have that

$$\mathrm{Ad}_P(A) = d(L_P \circ R_{P^{-1}})(A) = dL_P(dR_P^{-1}(A)) = PAP^{-1}$$

Definition 1.48. *Let $\mathfrak{g}$ be a Lie algebra. The adjoint representation of $\mathfrak{g}$ is*

$$(\mathrm{ad}\, v)(w) := [v, w]$$

The set $\mathrm{ad}\,\mathfrak{g} = \{\mathrm{ad}\, v : v \in \mathfrak{g}\}$ is a subalgebra of $\mathfrak{gl}(\mathfrak{g})$.

By direct verification, it can be checked that for any $v \in \mathfrak{g}$, $\mathrm{ad}\, v$ is a derivation of $\mathfrak{g}$.

There is a close relation between this map and Ad. To see that we will use this result on the exponential function

Lemma 1.49 ([73] theorem 2.10.3)**.** *If G_1, G_2 are Lie group and $\mathfrak{g}_1, \mathfrak{g}_2$ their Lie algebras we have that, for any analytic homomorphism $f : G_1 \to G_2$, for any $v \in \mathfrak{g}_1$*

$$f(\exp(v)) = \exp(df(v))$$

Theorem 1.50. *Let G a Lie group and $\mathfrak{g}$ its Lie algebra.*

$$d(\mathrm{Ad})_e(v) = \mathrm{ad}(v)$$

for any $v \in \mathfrak{g}$.

Moreover

$$\mathrm{Ad}(\exp v) = \exp(\mathrm{ad}\, v) \tag{1.10}$$

Proof. We give the proof for the case of a matrix lie group. The proof of the general case can be found in [73, Theorem 2.13.2].

By using equation (1.9) with $P = \mathrm{expm}(tv)$ we obtain

$$\mathrm{Ad}_{\mathrm{expm}(tv)}(w) = \mathrm{expm}(tv)w\,\mathrm{expm}(-tv)$$

By taking the derivative for t=0 we obtain $\mathrm{ad}(v)(w) = [v, w]$.

Because ad is the differential of Ad and Ad is an analytic homomorphism (because left and right multiplication are analytic by the definition of Lie group) by lemma 1.49 we obtain equation (1.10). ∎

It is possible to express the differential of the exponential as a formal series in $\mathrm{ad}(v)$. Such series is actually invertible if some technical conditions are satisfied.

Theorem 1.51 ([73] theorem 2.14.3)**.** *Let G a Lie group and $\mathfrak{g}$ its Lie algebra. For any $v \in \mathfrak{g}$ we have that the differential of the exponential at v can be expressed as*

$$(d\exp)_v = \sum_{n=0}^{\infty} \frac{(-1)^n}{(n+1)!}(\mathrm{ad}(v))^n \tag{1.11}$$

In particular $(d\exp)_v$ is bijective if and only if $\mathrm{ad}(v)$ has no eigenvalues of the form $(-1)^{\frac{1}{2}}2k$ for some integer $k \neq 0$

Remark. *We use the convenction of [73] for the differential of the exponential, for which*

$$\frac{d}{dt}\exp(X(t)) = \exp(X(t))(d\exp)_X\left(\frac{d}{dt}X(t)\right)$$

Other sources as [30] define the differential of the exponential differently.

The equation (1.11) can be written as $\frac{\mathrm{expm}(\mathrm{ad}(v))-e}{\mathrm{ad}(v)}$.

By using this formula and the fact that $f(x) = \frac{x}{e^x-1}$ is the generating function of the Bernoulli numbers (see [1]) it is possible to compute a formula for the inverse of $d\exp$.

$$(d\exp)_v^{-1} = \frac{\mathrm{ad}(v)}{\mathrm{expm}(\mathrm{ad}(v)) - e} = \sum_{j=0}^{\infty} \frac{B_j}{j!}(\mathrm{ad}(v))^j \tag{1.12}$$

Where B_j are the Bernoulli numbers.

A truncated expression for equation (1.12) is given by

$$d\,\mathrm{expinv}(u, v, p) = \sum_{k=0}^{q-1} \frac{B_k}{k!}\,\mathrm{ad}_u^k(v) \tag{1.13}$$

2 Numerical Methods on Lie groups

2.1 Lie group and Lie algebra action

Definition 2.1. *An action of a Lie group G on a manifold M is a smooth map $\Lambda : G \times M \to M$ such that for any $x \in M$ $\Lambda(e, x) = x$ and*

$$\Lambda(p, \Lambda(r, x)) = \Lambda(pr, x)$$

If the relation holds for any $p, r \in G$ it is called a global action. It it holds only in a neighbourhood of the identity of G is called local action.

The action can be studied from an infinitesimal point of view, by studying the action on the manifold of the Lie algebra corresponding to the Lie group. This can be done by differentiating the Lie group action near the identity.

Definition 2.2. *Let, M be a manifold, G be a lie group and $\mathfrak{g}$ its lie algebra. Consider a smooth map $\lambda : \mathfrak{g} \times M \to M$. Consider the map $\lambda_* : \mathfrak{g} \to \mathcal{H}(M)$:*

$$\lambda_*(v)(p) = \frac{d}{dt}(\lambda(tv, p))_{t=0} \tag{2.1}$$

λ is a left Lie algebra action if this map is a Lie algebra antihomomorphism.

It is a right Lie algebra action if it is a Lie algebra homomorphism.

Theorem 2.3. *Let Λ be a left Lie group action. Define*

$$\lambda(v, x) = \Lambda(\exp(v), x)$$

This map is a left lie algebra action

If Λ^R is a right Lie group action it is possible to define λ^R, as above. It is a right Lie algebra action.

Proof. The proof that λ_*^R is a Lie algebra homomorphism can be found in [46, theorem 20.15].

Let Λ a left action and λ_* as in equation (2.1). You can define $\Lambda^R(x,g) = \Lambda(g^{-1}, x)$ This is a right action so the map $\lambda_*^R(g) = \lambda_*(g^{-1})$ is a lie algebra homomorphism. Because $\exp(tv)^{-1} = \exp(-tv)$ and using the bi-linearity of the Lie bracket you obtain that λ_* is an anti-homomorphism. $\square$

Example 2.4. *consider $\lambda : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ the translation on $\mathbb{R}^n$ ($\lambda(v,p) = v + p$).*

$\lambda_*(v)(p) = v$ so λ is a Lie algebra action of R^n onto itself.

Example 2.5. *Let $M = G$ a matrix Lie group. Let $\lambda : \mathfrak{g} \times G \to G : \lambda(v,p) = \mathrm{expm}(v)p$. As any Lie group act on itself by left multiplication, by theorem 2.3 we have that λ is a Lie algebra action*

$$\lambda_*(v)(p) = -\frac{d}{dt}(\mathrm{expm}(tv)p)_{t=0} = vp \tag{2.2}$$

2.2 ODEs on manifolds and the Runge-Kutta methods

We will now define a class of numerical methods, the Runge-Kutta methods. We can express such methods in terms of the Lie algebra action defined in equation (2.1)

Definition 2.6. *Let b_i, a_{ij} $(i,j = i..s)$ real coefficients. Let $c_i = \sum_{j=1}^{s} a_{ij}$ The s-stage implicit Runge-Kutta method (RK method) is defined as:*

$$k_i = f(t_0 + c_i h, y_0 + h\sum_{j=1}^{s} a_{ij}k_j) \tag{2.3}$$

$$y_1 = y_0 + h\sum_{j=1}^{s} b_i k_i$$

The RK method is said to have order p if

$$\lim_{h\to 0} y_1 - y(t_0 + h) = O(h^{p+1})$$

By using the B-series formalism described in [27, chapter 3] it is always possible to find order conditions on the coefficients in equation (2.3) so that the RK method has order p

Example 2.7. *Consider the ODE in $\mathbb{R}^n$*

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

Let $M = \mathbb{R}^n$, λ as in example 2.4. Let a_{ij}, b_j the coefficients of a s-stage, q-th order Runge-Kutta method. Let $c_j = \sum_i a_{ij}$.

Suppose $y(t)$ satisfies the ODE. The algorithm of definition 2.6 can be rewritten as

$$k_i = f(t_0 + c_i h, y_0 + h\sum_{j=1}^{s} a_{ij}k_j)$$

$$y_1 = \lambda(y_0, h\sum_{j=1}^{s} b_i k_i)$$

By assuming the existence of a Lie algebra action we can obtain a generalization for ODEs of differentiable manifolds.

Assumption. *Let M be a manifold and $y(t) \in M$ a curve $y(0) = y_0$ There exists a Lie algebra $\mathfrak{g}$, a left Lie algebra action λ and a function $f : \mathbb{R} \times M \to \mathfrak{g}$ such that the ODE for $y(t)$ is*

$$y' = (\lambda_* f(t,y))(y), \quad y(0) = y_0 \tag{2.4}$$

This assumption is automatically satisfied for any ODE if the Lie algebra action is transitive (i.e. give any two points x, y of M there is at least an element of $v \in \mathfrak{g}$ such that $x = \lambda(v,y)$).

Moreover, we have that the assumption is always satisfied at least in a chart if we choose $\mathfrak{g}$ to be the vector space generated by $(\frac{\partial}{\partial x^i})$ where x^i are the coordinates on the chart.

Example 2.8. *If $M = G$ is a matrix Lie group and λ is the action of example 2.5 by using lemma 1.39 we obtain that for any $v \in \mathfrak{g}$ and $p \in G$ $\lambda_* v(p) = vp$. So equation (2.4) is*

$$y' = A(t, y)y \tag{2.5}$$

Where $A : \mathbb{R} \times G \to \mathfrak{g} \subseteq \mathfrak{gl}(n, k)$

The key difference between $\mathbb{R}^n$ and any other example of Lie algebra action on a manifold is that the Lie bracket are not trivial. Consequently the exponential function is not the identity function and it appears in the definition of the Lie algebra action of theorem 2.3. If we consider only small t it is still possible to express the solution of the ODE (2.4) as the (left) action of an element of the Lie algebra on $y(0) = y_0$.

Lemma 2.9 ([55] lemma 8)**.** *Suppose $M, \lambda, f, y(t)$ as in equation (2.4). Let $\lambda_x(u) = \lambda(u, x)$. Call X the vector field such that $X(y_0) = \lambda_* f(y_0)(y_0)$ and $\tilde{f}(u) = d\exp_u^{-1}(f \circ \lambda_{y_0}(u))$.*
So $\tilde{f}$ and X are λ_{y_0}-related i.e $\lambda'_{y_0} \circ \tilde{f} = X \circ \lambda_{y_0}$

Theorem 2.10. *for sufficently small t the ODE (2.4) can be expressed as*

$$\begin{aligned} y(t) &= \lambda(u(t), y_0) \\ u' &= d\exp_u^{-1}(f(t, \lambda(u, y_0))) \\ u(0) &= 0 \end{aligned}$$

Proof. in the notation of lemma 2.9 $u' = \tilde{f}(t, u)$, while $y' = X(y)$, $y(0) = \lambda_{y_0}(0)$. So

$$X(t, \lambda_{y_0}(u(t))) = \lambda'_{y_0}(\tilde{f}(t, u(t))) = \lambda'_{y_0}(u'(t)) = y'(t)$$

Where we have used lemma 2.9 and the chain rule. So $y(t) = \lambda_{y_0}(u(t))$ is a solution of (2.4) □

2.3 The RKMK method

Let M a manifold, $\mathfrak{g}$ a Lie algebra which act on M with the left action λ. It is possible to generalize the Runge-Kutta method so that it works in this setting.

Definition 2.11 (**Runge-Kutta-Munthe-Kaas method**)**.** *Let $M, \mathfrak{g}, \lambda, f, y(t)$ as in equation 2.4, $y(0) = y_0$. Let a_{ij}, b_j coefficients of an s-stage q-th order Runge-Kutta method. Let $c_i = \sum_{j=1}^{s} aij$. for $i = 1, .., s$*

$$\begin{aligned} v_i &= \sum_{j=1}^{s} a_{ij}\tilde{k}_j \\ k_i &= hf(hc_i, \lambda(v_i, y_0)) \\ \tilde{k}_i &= d\,\mathrm{expinv}(v_i, k_i, q) \\ v &= \sum_{j=1}^{s} b_j\tilde{k}_j \\ y_1 &= \lambda(v, y_0) \end{aligned}$$

Where $d\,\mathrm{expinv}$ is the truncated expansion of $d\exp^{-1}$ given by equation (1.13)

As the motion on M given by the action λ evolves on M we have the following result.

Theorem 2.12. *The algorithm of definition 2.11 stays on the manifold*

$$M_{y_0} = \{x \in M : x = \lambda(v_k, ..., \lambda(v_1, y_0)); v_1, ..., v_k \in \mathfrak{g}\}$$

Next we will define an analogous of the Taylor series which work for ODEs on manifolds

Definition 2.13. *Let X a vector field and $\theta_{X,t}$ its flow. The Lie derivative of a smooth function f w.r.t. X is defined as*

$$X[f] = \frac{\partial}{\partial t}(\theta^*_{X,t} f)_{t=0} \tag{2.6}$$

Where $\theta^*_{X,t} f = f \circ \theta_{X,t}$ is the pullback of f along the flow.
This definition can be extended to any object for which the pullback along the flow is well defined (e.g tensor fields, see [46])

Definition 2.14. *The Lie series of a smooth function f w.r. to the vector field X is obtained by iterating equation (2.6).*

$$\theta^*_{X,t} f = \sum_{i=0}^{\infty} \frac{t^i}{i!} X^i[f]$$

Where $X^i[f] = X[X^{i-1}[f]]$

Similarly to the real case a numerical method $y_0 \to y_1(h)$ has order q if the first $q+1$ terms of the Lie series of $y_1(h)$ around $h = 0$ match the first $q+1$ terms of the Lie series of the analytical solution of the ODE (2.4).

Theorem 2.15. *The RKMK method of definition 2.11 has at least order q for any Lie group action Λ (and so for any Lie algebra action $\lambda(u, x) = \Lambda(\exp(u), x)$) on any manifold M*

Proof. By theorem 2.10 we know that the analytic solution of the ODE is

$$y(t) = \lambda(u(t), y_0)$$

Where $u(t)$ satisfied the ODE

$$u' = d\exp_u^{-1}(f(t, \lambda(u, y_0)))$$
$$u(0) = 0$$

$\mathfrak{g}$ is a vector space so we can apply a classical Runge-Kutta scheme to this ODE and obtain $u_1 \approx u(h)$. This is equivalent to the equations for v in definition 2.11.

By hypothesis the coefficients a_{ij}, b_j satisfies the order condition for a q-th order Runge-Kutta method in so u_1 has order q.

After that we obtain $y_1 = \lambda(u_1, y_0)$. Because λ is smooth The order of y_1 is at least equal to the order of u_1.

The approximation of the $d\exp_u^{-1}$ is of order q so it doesn't reduce the error □

2.4 The Magnus expansion

By theorem 2.10 we know that for small t the ODE (2.4) is the action on a point of the manifold of the element of the Lie algebra which satisfies the ODE

$$\begin{cases} u' = d\exp_u^{-1}(f(t, \lambda(u, y_0))) \\ u(0) = 0 \end{cases} \tag{2.7}$$

By using the Picard Iteration we obtain the following recursive relation

$$\begin{cases} u^{[0]}(t) = 0 \\ u^{[m+1]}(t) = \int_0^t d\exp^{-1}_{u^{[m]}(\xi)} f(\xi, \lambda(u^{[m]}(\xi), y_0) d\xi = \sum_{k=0}^{\infty} \frac{B_k}{k!} \int_0^t \mathrm{ad}_{u^{[m]}(\xi)} f(\xi, \lambda(u^{[m]}(\xi), y_0) d\xi \end{cases} \tag{2.8}$$

Theorem 2.16. *$u(t) - u^{[m]}(t)$ has at least order t^{m+1}*

Proof. λ is smooth so it at least preserves the order of $u^{[m]}$. This implies that the proof of the general case can be done in similar way of [16, theorem 2.1] (where it is proven the particular case of a matrix Lie group). □

The iterated integral of equation (2.8) can be calculated using a quadrature.

In [16] quadrature for order 2,3 and 4 for the case of matrix Lie group are calculated.

We now focus on the case of a linear ODE i.e the case in which $f(t, y) = f(t)$ in (2.4). In this case we have that $u(t)$ in (2.7) can be written in terms of binary rooted trees.

Theorem 2.17 ([34] chapter 4)**.** *Suppose $u(t)$ is a solution of the ODE (2.7) with $f(t, \lambda(u, y_0)) = f(t)$. For t small enough we have that*

$$u(t) = \sum_{k=0}^{\infty} \sum_{\tau \in \mathbb{T}_k} \alpha(\tau) \int_0^t C_\tau(\xi) d\xi$$

Where $\mathbb{T}_k$ is a subset of the set of binary rooted trees.

$\mathbb{T}_k$ and $C_\tau, \tau \in \mathbb{T}_k$ can be obtained by two composition rules

1. $\mathbb{T}_0 = \{\tau_0\}$ *and* $C_{\tau_0}(t) = u(t)$

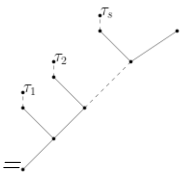
2. $\tau_1 \in \mathbb{T}_{m_1}$,$\tau_2 \in \mathbb{T}_{m_2}$ *imply that exists* $\tau \in \mathbb{T}_{m_1+m_2+1}$ *such that*

$$C_{tau}(t) = [\int_0^t C_{\tau_1}(\xi)d\xi, C_{\tau_2}(t)]$$

and $\alpha(\tau)$ *is a constant such that* $\alpha(\bullet) = 1$.

$$\alpha(\tau) = \frac{B_s}{s!}\prod_{i=1}^{s}\alpha(\tau_i)$$

Where τ *is decompose as*

$\tau =$ 

Such series is well defined and converges for small t

In either the RKMK method or the Magnus expansion we need to calculate the iterated commutators of elements of the Lie algebra considered. Such task is computational expensive, but using the framework of free Lie algebras it is possible to optimize this procedure reducing the number of commutators necessary to obtain the approximation.

2.5 The free Lie algebra

A magma is a couple formed by a set and a binary operation from the elements of the set. This operation is assumed to be closed.

Definition 2.18. *Suppose we have a set A. Give two element of this set we define a binary operation $[.,.]$ which send any couple of elements of $a_1, a_2 \in A$ to the formal bracket $[a_1, a_2]$.*

The free magma over A, $M(A)$ is defined as the smallest set which contains A and it is closed under this binary operation.

Equivalently we can define the free magma as the set of rooted planar trees with leaves in A with the operation of composition of trees

It is possible to associate a grading for the elements of the free magma. For each $a_i \in A$ let $\deg(a_i) = 1$ *the degree of a_i. For any $u, v \in M(A)$*

$$\deg([u, v]) := \deg(u) + \deg(v) \tag{2.9}$$

Given a set B and a field $\mathbb{K}$ it is possible to define the set of the formal linear combination of element of B with coefficients in $\mathbb{K}$. This is called the free $\mathbb{K} - module$ over B.

Definition 2.19. *Given a field $\mathbb{K}$ and a set A the free $\mathbb{K}$-algebra over A, denoted by $\mathcal{D}(A)$ is the free $\mathbb{K}$-module of the free magma of A. The algebra operation of $\mathcal{D}(A)$ is obtained by extending by linearity the magma operation of $M(A)$.*

$\mathcal{D}(A)$ has a natural structure of graded algebra given by the grading of $M(A)$

The free $\mathbb{K}$-algebra is the smallest algebra which contains A. By quotienting by the opportune ideal is it possible to give some properties to its operation. In particular we are interested in the Jacobi identity (definition 1.22).

Definition 2.20. *Let A a set and $\mathbb{K}$ a field. Let $\mathcal{D}(A)$ the free $\mathbb{K}$-algebra over A and $[.,.]$ its operation. Consider the ideal $I = \{[[x, y], z] + [[y, z], x] + [[z, x], y], [x, x] \ \forall \ \ x, y, z \in \mathcal{D}(A)\}$. The free Lie algebra over A is defined as*

$$\mathcal{L}(A) := \mathcal{D}(A)/I$$

with algebra multiplication inherit by $\mathcal{D}(A)$

Proposition 2.21 ([63] theorem 0.4)**.** *the multiplication of $\mathcal{L}(A)$ is well defined. $\mathcal{L}(A)$ is a Lie algebra and it is a graded algebra with the grading inherit by $\mathcal{D}(A)$.*

$\mathcal{L}(A)$ is generated as a Lie algebra by A and if $\deg(a_i) = 1$ for each $a_i \in A$ the component of degree 1 of $\mathcal{L}(A)$ are generated by $i(A)$, where $i : A \hookrightarrow \mathcal{L}(A)$ is the inclusion.

Moreover the free Lie algebra satisfies the following universal property: for any $\mathcal{L}$ Lie algebra over $\mathbb{K}$ and for any function $f : A \to \mathcal{L}$ $\exists!$ $\bar{f} : \mathcal{L}(A) \to \mathcal{L}$ such that the diagram below commutes

$$\begin{tikzcd} A \arrow[rr, "i"] \arrow[dr, "f"'] & & \mathcal{L}(A) \arrow[dl, "\bar{f}"] \\ & \mathcal{L} & \end{tikzcd}$$

Any Lie algebra which satisfies this universal property is isomorphism to the free Lie algebra.

This property is what makes the free Lie algebra a useful tool in the field of numerical computation of Lie algebras, because in some sense it is the most general possible object which satisfies the property of definition 1.22

If we need to do some calculation in a concrete Lie algebra $\mathfrak{g}$ we can do that in the abstract setting and by using the universal property of theorem 2.21 with $f(i \in B) = X_i \in \mathfrak{g}$ (where X_i are the elements of $\mathfrak{g}$ that we are interested to analyze) we can recover the concrete Lie algebra.

2.6 The universal enveloping algebra

We have seen that the elements of a Lie algebra associated to a Lie group G can be seen as the element of the tangent space T_eG i.e. as derivation of the Lie groups. The universal enveloping algebra allow to consider higher order differential operators.

Definition 2.22. *Let $\mathcal{L}$ a Lie algebra over a field $\mathbb{K}$. Consider the tensor algebra over $\mathcal{L}$. $T(\mathcal{L}) := \bigoplus_{n>0} \mathcal{L}^{\otimes n}$ where $\mathcal{L}^{\otimes n}$ is the tensor product of $\mathcal{L}$ with itself n times (see [3] for the definition).*

Consider the ideal $I = \{x \otimes y - y \otimes x - [x, y] \ \forall \quad x, y \in T(\mathcal{L})\}$.

$$A_0 = T(\mathcal{L})/I$$

with algebra multiplication inherit by the projection $T(\mathcal{L}) \xrightarrow{p} A_0$ is called the universal enveloping algebra of $\mathcal{L}$

Proposition 2.23 ([63] proposition 0.1)**.** *The multiplication of A_0 is well defined.*

Given A_0 exists an algebra homomorphism $\psi_0 : \mathcal{L} \to A_0$ such that given any associative algebra A and any Lie algebra homomorphism $\phi : \mathcal{L} \to A$ exists unique $\bar{\phi} : A_0 \to A$ such that the diagram below commutes

$$\begin{tikzcd} \mathcal{L} \arrow[rr, "\psi_0"] \arrow[dr, "\phi"'] & & A_0 \arrow[dl, "\bar{\phi}"] \\ & A & \end{tikzcd}$$

If there is another couple (A_1, ψ_1) which satisfies this universal property, A_1 is isomorphic to the universal enveloping algebra of $\mathcal{L}$.

ψ_0 can be defined as $\psi_0 : \mathcal{L} \overset{i}{\hookrightarrow} T(\mathcal{L}) \xrightarrow{p} A_0$ where $i : \mathcal{L} \hookrightarrow T(\mathcal{L})$ is the inclusion and $p : T(\mathcal{L}) \to A_0$ is the projection.

The free associative algebra $\tilde{A}$ over a set B is defined to be the associative algebra which satisfies the following universal property : for each associative algebra A and each map $B \xrightarrow{f} A$ exists unique an algebra homomorphism $\tilde{A} \xrightarrow{\tilde{f}} A$ such that the following diagram commutes:

$$\begin{tikzcd} B \arrow[rr, "i"] \arrow[dr, "f"'] & & \tilde{A} \arrow[dl, "\tilde{f}"] \\ & A & \end{tikzcd}$$

Definition 2.24. *Let A is a set. A word on the alphabet A is a finite sequence of elements of A. Denote the set of all the words with A^**

A non-commutative polynomial on A over the field $\mathbb{K}$ is a $\mathbb{K}$-linear combination of words on A. We can write any polynomial as $P = \sum_{w \in A^*} (P, w) w$ where all, but finitely many of the coefficients $(P, w) \in \mathbb{K}$ are 0.

If $P = \sum_{u \in A^*} (P,u)u$ and $Q = \sum_{v \in A^*} (Q,v)v$ define the concatenation product to be PQ with

$$(PQ, w) = \sum_{w=uv} (P,u)(Q,w)$$

Denote $K\langle A \rangle_.$ the set of all the non-commutative polynomial on A with the concatenation product.

$K\langle A \rangle_.$ is an associative algebra and satisfies the universal property of a free associative algebra generated by A [63, proposition 1.2]

We have the following characterization of the universal enveloping algebra of the free Lie algebra

Theorem 2.25. *Given a set B, the universal enveloping algebra of the free Lie algebra over B is the free associative algebra over B.*

Conversely if $K\langle A \rangle_.$ is the free associative algebra over a set B we can define a Lie bracket on $K\langle A \rangle_.$ as $[u,v] = uv - vu$ (where the juxtaposition represent the algebra operation of $K\langle A \rangle_.$). We have that $(K\langle A \rangle_., [.,.])$ is a Lie algebra. If we call $\mathcal{L}$ the smallest Lie subalgebra of $(K\langle A \rangle_., [.,.])$ which contains B we have that $\mathcal{L}$ is the free Lie algebra generated by B. $K\langle A \rangle_.$ is its universal enveloping algebra

Proof.

the proof of the first part can be found in [63] theorem 0.5.

The proof of the second statement can be found in [73] theorem 3.2.8 □

By the second statement of theorem 2.25 we can identify the free Lie algebra generated by a set B as the smallest set of non-commutative polynomial which contains B and is closed under the Lie bracket.

The elements of the free Lie algebra are so called **Lie polynomials**

2.7 Hall bases

We want to represent the free Lie algebra of definition 2.20 as linear combination of bases elements. There are various suitable bases for this vector space, one of the most common is the Hall bases.

Definition 2.26. *Let A a set and $M(A)$ its free magma. A set H is called an Hall set if:*

1. *$A \subseteq H$*
2. *H has a total ordering $<$ defined as follow: $\deg(u) < \deg(v) \to u < v$, where $\deg$ is the order of $M(A)$ defined in equation (2.9). If two elements have the same length they are ordered lexicographically.*
3. *Let $h \in M(A)$ such that $\deg(h) = 2$. $h \in H$ if and only if $h = [a_1, a_2]$ with $a_1, a_2 \in H$ and $a_1 < a_2$*
4. *Let $h \in M(A)$ such that $\deg(h) \geq 3$. $h \in H$ if and only if $h = [u,[v,w]]$ with $u, v, w, [v,w] \in H$ and $v \leq u < [v,w]$*

Such set always exists for any A ([8] chapter 2 proposition 11).

If $\mathcal{L}(A)$ is a free Lie algebra for A the immersion of the Hall set into $\mathcal{L}(A)$ is a (vector space) bases for $\mathcal{L}(A)$ ([8] chapter 2 theorem 1)

Remark. *Definition 2.26 is based on [8, page 132] and [58]. Some authors like [63] give a different definition of the Hall set (they reverse the order of the inequalities).*

The dimension of the component of length n can be calculated using the Witt formula

Theorem 2.27 ([8] chapter 2 theorem 2)**.** *Let A be a finite set and H is one of its Hall sets, the number of elements of A of length n i.e. the dimension of the module of homogeneous Lie polynomial of degree n is given by the* ***Witt formula***

$$\nu_n = \frac{1}{n} \sum_{d|n} \mu(d) s^{\frac{n}{d}}$$

where s is the number of generators and $\mu : \mathbb{Z}^+ \to \{-1, 0, 1\}$: if $d = p_1^{n_1} \ldots p_k^{n_k}$ is the prime factorization of d

$$\mu(d) = \begin{cases} 1 \text{ for } d = 1 \\ (-1)^k \text{ if all } n_i = 1 \\ 0 \text{ otherwise} \end{cases} \tag{2.10}$$

The Witt formula can be extended so that it consider other grading other that the length defined in equation (2.9). In particular given a grading function for the set A $\omega : A \to \mathbb{Z}^+$ such that $\omega(a_i) = \omega_i$ we can extend such function to the Hall bases by additivity

$$\omega([u, v]) = \omega(u) + \omega(v)$$

This induces a grading on the free Lie algebra of A

Example 2.28. *Suppose $P, Q \in M_n(\mathbb{R})$ are matrices which depend by a small parameter h as $P = O(h^{n_P})$, $Q = O(h^{n_Q})$. We have that $[P, Q] = O(h^{n_P + n_Q})$. So the function $\omega(P) = n_P$ is a grading function.*

Theorem 2.29 ([58] theorem 3.1)**.** *Lat A be a set of s elements and $\mathcal{L}(A)$ its free Lie algebra, let ω be a grading function for A such that $\omega(a_i) = \omega_i \ \forall \ a_i \in A$. Consider the polynomial*

$$P(T) = 1 - \sum_{i=1}^{s} T^{\omega_i}$$

and call $\{\lambda_i\}_{i=1}^{\max(\omega_i)}$ its roots. Call $\mathcal{L}(A)_n$ the subspace of the free Lie algebra composed by the elements of degree n with respect to the grading induced by ω. We have that the dimension of these subspaces is

$$\dim(\mathcal{L}(A)_n) = \frac{1}{n} \sum_{d|n} \Big(\sum_{i=1}^{\max(\omega_i)} \lambda_i^{\frac{-n}{d}} \Big) \mu(d) \tag{2.11}$$

where $\mu(d)$ is the Möbius function of equation (2.10)

2.8 Optimization of the RKMK algorithm

Theorem 2.29 allow us to reduce the number of commutators necessary for implementing the algorithm of definition 2.11 in the case of a linear ODE. Indeed with the same notations of (2.4) let

$$y' = \lambda_*(f(t))(y)$$

a linear ODE on a manifold M. The algorithm for a s-stage RKMK method is

$$\begin{cases} v_i = \sum_{j=1}^{s} a_{ij} \tilde{k}_j \\ k_i = hf(hc_i) \\ \tilde{k}_i = d\,\mathrm{expinv}(v_i, k_i, q) \\ v = \sum_{j=1}^{s} b_j \tilde{k}_j \\ y_1 = \lambda(v, y_0) \end{cases} \tag{2.12}$$

where a_{ij}, b_j, c_i and $d\,\mathrm{expinv}$ are the same of definition 2.11

Consider the grading function of example 2.28. The goal is to find a change of bases $\{k_i\} \to \{Q_i\}$ such that, with respect to that grading $\deg(Q_i) = i$. For doing this we define the Vandermonde matrix $(V_{ij})_{j=1}^{s} = (c_i^{j-1})$ and define the new basis as

$$(Q_1, .., Q_s)^T = V(c)^{-1}(k_1, ..., k_s)^T$$

We have that

$$Q_i = \frac{h^i}{(i-1)!} f^{(i-1)}(\xi_i) \text{ for some } \xi_i \in (0, h) \tag{2.13}$$

Another possible optimization is given by the time-reversal symmetry. Indeed if we consider the equation $y' = \lambda_*(-f(t-h))(y)$ we obtain a flow which in the time $t \in (0, h)$ go from y_1 to y_0. So if we use a Taylor series around $t = \frac{1}{2}h$ we obtain that under the symmetry $f(t) \to -f(h-t)$ $Q_i \to (-1)^i Q_i$.

Remember that $\lambda(v, x) = \Lambda(\exp(v), x)$. So under the time reverse symmetry we have that $v = v(Q_1, ..., Q_s) \to -v(Q_1, .., Q_s) = v((-1)Q_1, Q_2, ..., (-1)^s Q_s)$. So v depends only through terms on the Hall set of odd degree.

In the end the algorithm (2.12) becomes

$$\begin{cases} V = ((c_i + \frac{1}{2})^{j-1})_{i,j=1}^{s} \\ k_i = hf(hc_i) \\ Q_i = \sum_{j=1}^{s} (V^{-1})_{ij} k_j \\ v = v(Q_1, ..., Q_s) \\ y_1 = \lambda(v, y_0) \end{cases}$$

3 Riemannian geometry

3.1 Riemannian manifolds definition and examples

Definition 3.1. *let M a smooth manifold, TM its tangent bundle and T^*M its dual (the cotangent bundle). Define*

$$\mathcal{T}_s^r(M) := \bigotimes_r TM \otimes \bigotimes_s T^*M$$

An (r, s) tensor field is a smooth section of $\mathcal{T}_s^r$. In particular, $(0, 0)$ tensors are smooth functions and $(1, 0)$ tensors are vector fields

We are interested in a particular type of tensor:

Definition 3.2. *A pseudo-Riemannian structure on a smooth manifold M is a $(0, 2)$-tensor g such that*

$$g(X, Y) = g(Y, X)$$

and for each $p \in M$ g_p is not degenerate.

M has a Riemannian structure if g_p is positive definite for each $p \in M$. In such case g_p is a metric on T_pM and we will denote it as $g_p(.,.)$ $\langle .,. \rangle_p$

Proposition 3.3. *Any smooth manifold admits a Riemannian structure*

Proof. Let $\mathcal{A} := (U_\alpha, \psi_\alpha)_{\alpha \in I}$ a locally finite atlas on M. By lemma 1.5 exists a partition of unity $(\tau_\alpha)_{\alpha \in I}$ subordinated to $\mathcal{A}$. Because U_α is diffeomorphic to an open subset of $\mathbb{R}^n$ it is possible to define a metric on it as $g_\alpha := \psi_\alpha^* g^{\mathrm{eucl}} = g^{\mathrm{eucl}} \circ \psi^{-1}$, where g^{eucl} is the standard euclidean metric. Define

$$g = \sum_{\alpha \in I} \tau_\alpha g_\alpha$$

This sum is finite for any p because of the locally finiteness of the partitition of unity and g is a Riemannian metric for M □

Definition 3.4. *A map between Riemannian manifolds $f : (M, \langle .,. \rangle_M) \to (N, \langle .,. \rangle_N)$ is an isometry if for any $p \in M$ and $u, v \in T_pM$*

$$\langle u, v \rangle_M |_p = \langle df_p(u), df_p(v) \rangle_N |_{f(p)}$$

f is a local isometry at $p \in M$ if it is an isometry in a neighborhood of p

Example 3.5 (warp products). *Let (B, g_B), (F, g_F) two Riemannian manifolds with not zero dimension. Let $\pi_B : B \times F \to B$ and $\pi_F : B \times F \to F$ the natural projections of the product manifold. Let $f : B \to [0\infty)$ a positive, smooth function. The warped product $B \times_f F$ is defined as the Riemannian manifold $(B \times F, \langle .,. \rangle)$ where*

$$\langle X|X \rangle := g_B (\pi_B^*(X), \pi_B^*(X)) + f^2 (\pi_B(X)) \, g_F (\pi_F^*(X), \pi_F^*(X))$$

Example 3.6. *given a Lie group G a metric is said to be left invariant if L_x is an isometry for any $x \in G$. G is right invariant if R_x is an isometry for any $x \in G$ and is biinvariant if it is left and right invariant.*

Any Lie group admits a left (or right) invariant metric. Given any metric $\langle .,. \rangle_e$ on the Lie algebra $\mathfrak{g}$ we define

$$\langle u, v \rangle_x = \langle dL_{x^{-1}}(u), dL_{x^{-1}}(v) \rangle_e$$

Any compact Lie group admits a biinvariant metric (see [23] exercise 7) If $\langle .,. \rangle$ is a biinvariant metric on G we have that

$$\langle [u, w], v \rangle = - \langle u, [v, w] \rangle$$

(see [23] example 2.6)

There is a more general result which characterizes the Lie groups that admits a bi-invariant Riemannian structure

Proposition 3.7 ([52] lemma 7.5)**.** *A connected Lie group admits a bi-invariant metric if and only if it is isomorphic to the Cartesian product of a compact Lie group with an abelian Lie group*

3.2 Connnections and curvature

in $\mathbb{R}^n$ a vector field along a curve $c : I \to \mathbb{R}^n$ is said parallel if its directional derivative is 0. To define the concept of parallelism on a general smooth manifold we have to specify an affine connection

Definition 3.8. *Let M a smooth manifold and $\mathcal{H}(M)$ the set of all the vector field on M (see example 1.12). An affine connection ∇ on M is a map*

$$\nabla : \mathcal{H}(M) \times \mathcal{H}(M) \to \mathcal{H}(M)$$

that for any $X, Y, Z \in \mathcal{H}(M)$ and f, g smooth functions has the following properties:

1. *C^∞-**linearity** in the first component.*

$$\nabla_{fX+gY} Z = f\nabla_X Z + g\nabla_Y Z$$

2. *$\mathcal{H}(M)$-linearity and Leibnitz rule in the second component*

$$\begin{aligned}\nabla_X(Y+Z) &= \nabla_X Y + \nabla_X Z \\ \nabla_X(fY) &= f\nabla_X Y + XfY\end{aligned}$$

If (U, x^i) is a chart and $(\frac{\partial}{\partial x^i})$ is the corrresponding local section of TM we define the Christoffel symbols of second kind Γ^k_{ij} as the smooth function such that:

$$\nabla_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j}|_p = \Gamma^k_{ij}(p) \frac{\partial}{\partial x^j}|_p \tag{3.1}$$

for any $p \in U$. From now on we will omit the dependency by p when there is no risk of confusion.

By using equation (3.1) and the properties of the connection it is straightforward to see that if $X_p = X^i(p)\frac{\partial}{\partial x^i}|_p$ and $Y = Y^j(p)\frac{\partial}{\partial x^j}|_p$ for any p in a chart (U, x^i) that, for any $p \in U$

$$\nabla_X Y|_p = \left(X^i(p) Y^j(p) \Gamma^k_{ij}(p) + X^i(p) \frac{\partial}{\partial x^i}|_p Y^k(p) \right) \frac{\partial}{\partial x^k}|_p \tag{3.2}$$

Theorem 3.9 ([23] proposition 2.2)**.** *Given a smooth manifold with a connection (M, ∇) there exists a unique correspondence that associates to any vector field V along a curve $\gamma : I \to M$ another vector field along γ $\frac{D}{dt}V$ such that, for any V, W vector fields along γ and any f smooth in I*

1. *$\frac{D}{dt}(V + W) = \frac{D}{dt}V + \frac{D}{dt}W$*
2. *$\frac{D}{dt}(fV) = \frac{d}{dt}fV + f\frac{D}{dt}V$*
3. *if there is $Y \in \mathcal{H}(M) : V_t = Y_{\gamma(t)}$ then $\frac{D}{dt}V = \nabla_{\frac{d\gamma}{dt}} Y$*

$\frac{D}{dt}$ is called the covariant derivative along γ

It is possible to consider the covariant derivative of tensors:

Definition 3.10. *Let $T \in \mathcal{T}^0_s$ a $(0, s)$-tensor. The covariant differential ∇T is a $(0, s+1)$ tensor given by:*

$$\nabla T(Y_1, \cdots, Y_s, Z) = ZT(Y_1, \cdots, Y_s) - T(\nabla_Z, Y_1, \cdots Y_s) - \cdots - T(Y_1, \cdots, \nabla_Z Y_s)$$

for any $Y_1, \cdots, Y_s, Z \in \mathcal{H}(M)$.

The covariant derivative of T is the $(0, s)$-tensor

$$\nabla_Z T(Y_1, \cdots, Y_s) = \nabla T(Y_1, \cdots, Y_s, Z)$$

Definition 3.11. *Given an (r, s)-tensor on a Riemannian manifold (M, g, ∇) it is possible to define the contraction as the map $i : \mathcal{T}^r_s \to \mathcal{T}^{r-1}_{s+1}$ such that*

$$i\left(T^{a_1,\cdots a_r}_{b_1,\cdots b_s}\right) := g_{a_1,b_{s+1}} T^{a_1,\cdots a_r}_{b_1,\cdots b_s}$$

In this way is it possible to extend the definition of covariant differential also to controvariant tensors

We can now define the concept of parallelism for vector field on a manifold.

Definition 3.12. *Let (M, ∇) a smooth manifold with a connection and V a vector field along a curve $\gamma : I \to M$. V is parallel if $\frac{D}{dt}V = 0$ for any $t \in I$.*

Let (U, x^i) a chart and suppose $J \subseteq I$ is an interval such that $\gamma(J) \subseteq U$. Let $V_t = V^j(t)\frac{\partial}{\partial x^i}|_{\gamma(t)}$ for any $t \in [t_0, t_1]$. By using equation (3.2) we obtain that if V is a parallel vector field for each $t \in [t_0, t_1]$ and for any k

$$\frac{d}{dt}V^k + \Gamma^k_{ij}V^j\frac{d\gamma^i}{dt} = 0 \tag{3.3}$$

Given a tangent vector $v_0 \in T_{\gamma(t_0)}M$, where $t_0 \in I$ the coefficients of a vector field V parallel to γ such that $V_0 = v_0$ has to satisfies the ODE (3.3) in any chart that contains $\gamma(t_0)$.

By standard calculus and the compactness of $\gamma(I)$ we know that such vector exists and it is unique. It is called the **parallel transport** of v_0 along γ For Riemannian manifolds there are some particular choices of affine connection, namely those for which the covariant differential of the metric tensor is 0.

Definition 3.13. *Let $(M, \langle .,. \rangle, \nabla)$ a Riemannian manifold with an affine connection. ∇ is said compatible with the metric if, or any $X, Y, Z \in \mathcal{H}(M)$*

$$X\langle Y, Z\rangle = \langle \nabla_X Y, Z\rangle + \langle Y, \nabla_X Z\rangle$$

This condition is equivalent to:

$$\nabla_Z \langle X, Y\rangle = 0$$

Definition 3.14. *An affine connection ∇ is said torsionless (or symmetric) if, for any $X, Y \in \mathcal{H}(M)$*

$$\nabla_X Y - \nabla_Y X = [X, Y]$$

As $\left[\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\right] = 0$, we obtain

$$\Gamma^k_{ij}\frac{\partial}{\partial x^k} = \nabla_{\frac{\partial}{\partial x^i}}\frac{\partial}{\partial x^j} = \nabla_{\frac{\partial}{\partial x^j}}\frac{\partial}{\partial x^i} = \Gamma^k_{ji}$$

so for torsionless connections $\Gamma^k_{ij} = \Gamma^k_{jk}$ for any k

Theorem 3.15 (chapter 2 [23] theorem 3.6). *Let $(M, \langle .,. \rangle)$ a Riemannian manifold. There exists a unique connection that is torsionless and compatible with the metric. Such connection is called the Levi-Civita connection associated to the metric.*

If we call $g_{ij} := \left\langle \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\right\rangle$ and $g^{ij} = (g_{ij})^{-1}$ the Christoffel symbols of such connection are

$$\Gamma^m_{ij} = \frac{1}{2}g^{km}\left(\frac{\partial}{\partial x^i}g_{jk} + \frac{\partial}{\partial x^j}g_{ki} - \frac{\partial}{\partial x^k}g_{ij}\right)$$

We can now define the curvature tensor.

Definition 3.16. *Let $(M, \langle .,. \rangle, \nabla)$ a Riemannian manifold endowed with the Levi-Civita connection.*

The Riemann curvature tensor $R : \mathcal{H}(M) \times \mathcal{H}(M) \times \mathcal{H}(M) \to \mathcal{H}(M)$ is defined as

$$R(X, Y, Z) = \nabla_X\nabla_Y Z - \nabla_Y\nabla_X Z - \nabla_{[X,Y]}Z$$

Lemma 3.17. *Let R the Riemannian curvature tensor of $(M, \langle .,. \rangle, \nabla)$.*

For any $p \in M$ and for any $k \in \mathbb{N}$ we have that $R|_p$ and $\nabla^k R|_p$ are smooth functions on the tensor bundle.

Proof. Follow directly by the definition of $\mathcal{H}(M)$ and of the covariant derivative of a tensor $\square$

Definition 3.18. *Let $(M, \langle .,. \rangle, \nabla)$ a Riemannian manifold endowed with the Levi Civita connection, let R the Riemannian curvature tensor. Let u, v two linearly independent tangent vectors at the same point. The sectional curvature (or Gauss curvature) is*

$$K(u, v) := \frac{\langle R(u, v)v, u\rangle}{\langle u, u\rangle\langle v, v\rangle - (\langle u, v\rangle)^2}$$

We now give an expression for the curvature tensor of the warp product, defined in Example 3.5

Lemma 3.19 ([4] proposition 2.1)**.** *Let R the Riemannian curvature of the warp product $B \times_f F$ (see Example 3.5) and R^F the curvature tensor of F.It is possible to show that for any U, V, W vector fields of F, if we call $\tilde{U}$,$\tilde{V}$,$\tilde{W}$ their lifting over the warp product.*

$$R(\tilde{U}, \tilde{V})\tilde{W} = R^F(U, V)W + \frac{||\nabla f||^2}{f^2} |g_F(W, U)V - g(V, U)W|$$

In the case of a Lie group endowed with a bi-invariant metric it is possible to express the covariant derivative of a vector field and the curvature in terms of elements of the Lie algebra, more precisely,

Proposition 3.20 ([28] proposition 21.19)**.** *For any Lie group G equipped with a bi-invariant metric, the following properties hold:*

- *The Levi Civita connection is given by: $\nabla_X Y = [X, Y]$ for any X, Y left invariant vector fields*
- *The Riemannian curvature tensor is given by $R(u, v) = \frac{1}{4} \operatorname{ad}_{[u,v]}$ for each $u, v \in \mathfrak{g}$*
- *The sectional curvature is given by $K(u, v) = \frac{1}{4} \langle [u, v], [u, v] \rangle$ for any $u, v \in \mathfrak{g}$*

The connection also allow us to generalize the gradient and divergence to the case of a Riemannian manifold:

Definition 3.21. *Let (M, g, ∇) a Riemannian manifold endowed with the Levi Civita connection.*

*The **gradient** of a smooth function is defined as the algebraic dual of the differential i.e for any $X \in \mathcal{H}(M)$*

$$\langle \operatorname{grad} f, X \rangle = df(X) := Xf$$

where $\langle ., . \rangle = g(., .)$ as usual. In local coordinates the gradient can be expressed as

$$\operatorname{grad} f = g^{ij} \frac{\partial}{\partial x^i} f \frac{\partial}{\partial x^j}$$

*The **divergence** of a vector field is defined as the contraction i (see definition3.11) of the covariant differential of the vector field i.e*

$$\operatorname{div} X := i(\nabla X)$$

in local coordinates, if $X = a^i \frac{\partial}{\partial x^i}$ we have that

$$\operatorname{div} X = \frac{1}{\sqrt{\det(g)}} \frac{\partial}{\partial x^i} \left(\sqrt{\det(g)} a^i \right)$$

Definition 3.22. *Let (M, g, ∇) a Riemannian manifold endowed with the Levi Civita connection. The Laplace Beltrami operator is defined as $\Delta_M : C^\infty(M) \to C^\infty(M)$ such that*

$$\Delta_M f := \operatorname{div} \operatorname{grad} f$$

In local coordinates the Laplace Beltrami operator is expressed as

$$\Delta_M f = \frac{1}{\sqrt{\det(g)}} \frac{\partial}{\partial x^i} \left(\sqrt{\det(g)} g^{ij} \frac{\partial}{\partial x^i} f \right)$$

For any orthonormal bases $\{X_i\}$ of T_pM we have ([32] proposition 3.1.1)

$$\Delta_M f = \operatorname{trace} \nabla^2 f = \sum_{i=1}^{d} \nabla^2 f(X_i, X_i)$$

If $X_i = \frac{\partial}{\partial x^i}$ for any i we have

$$\Delta_M f(x) = g^{ij}(x) \left(\frac{\partial^2 f(x)}{\partial x^1 \partial x^j} - \Gamma^k_{ij}(x) \frac{\partial}{\partial x^k} f(x) \right) \tag{3.4}$$

3.3 Geodesics and the Riemannian exponential

Definition 3.23. *Let $(M, \langle .,. \rangle, \nabla)$ a Riemannian manifold endowed with the Levi-Civita connection. A parametrized curve $\gamma : I \to M$ is a geodesic if $\frac{D}{dt}\frac{d\gamma}{dt} = 0$, i.e. if the vector field $\frac{d\gamma}{dt}$ is parallel w.r.t. γ. If $\gamma(t) = (x^1(t), \cdots, x^n(t))$ in a chart $(U, \mathbf{x})$ around $\gamma(t_0)$, by equation (3.3) we obtain that γ is a geodesics if and only if*

$$\frac{d^2}{dt}x^i + \Gamma^k_{ij}\frac{dx^i}{dt}\frac{dx^j}{dt} = 0 \tag{3.5}$$

By using equation (3.5) and the existence and uniqueness theorem for second order linear ODE it is immediate to see that, given $(q, w) \in TM$ there is an unique geodesics defined in some interval I such that $\gamma(0) = q$, $\frac{d\gamma}{dt}(0) = w$

Lemma 3.24 ([23] chapter 3 proposition 2.7)**.** *Given $p \in M$ there exists a neighbourhood V of p and $\epsilon > 0$ and a smooth map $\gamma : (-2, 2) \times \mathcal{U} \to M$, where $\mathcal{U} := \{(q, w) \in TM : q \in V|w| \leq \epsilon\}$ such that $\gamma(t, (q, w))$ is the unique geodesic such that $\gamma(0) = q$, $\frac{d\gamma}{dt}(0) = w$ for every $w \in T_qM$ such that $|w| \leq \epsilon$*

with lemma 3.24 it is possible to define the geodesics exponential map

Definition 3.25. *Let $p \in M$ and $\mathcal{U} \subseteq TM$ and $\gamma : (-2, 2) \times \mathcal{U} \to M$ as in lemma 3.24. The geodesics exponential $\mathrm{Exp} : \mathcal{U} \to M$ is defined as*

$$\mathrm{Exp}(q, w) := \gamma(1, (q, w))$$

We have an analogous result to the case of the Lie group exponential

Proposition 3.26. *there exists $\epsilon > 0$ such that in the open ball $B_\epsilon(0) \subseteq T_qM$*

$$\mathrm{Exp}\,|_{T_qM} := \mathrm{Exp}_q : B_\epsilon(0) \to M$$

is a diffeomorphism onto its image

Proof.

$$d(\mathrm{Exp})_0(v) = \frac{d}{dt}|_{t=0}\gamma(1, q, tv) = \frac{d}{dt}|_{t=0}\gamma(t, q, v) = v$$

where we have used the homogeneity of a geodesic: $\gamma(\tau, q, tv) = \gamma(t\tau, q, v)$ ([23] lemma 2.6) □

Definition 3.27. *Let $p \in M$ and V a neighbourhood of the origin of T_pM such that the exponential map is a diffeormorphism onto its image in V. $\mathrm{Exp}_q(V)$ is called a normal neighborhood of M.*

In particular, if $B_\epsilon(0)$ is a ball whose closure is contained in V $\mathrm{Exp}_q(B_\epsilon(0))$ is called a geodesic ball

In case the Riemannian manifold is a Lie group equipped with a bi-invariant metric the definition of Riemannian and Lie group exponential coincide

Proposition 3.28 ([28] proposition 21.20)**.** *Let G a Lie group equipped with a bi-invariant metric. The geodesics passing through the origin are the curves $\gamma_u(t) = \exp(tu)$, where $\exp$ is the Lie group exponential. So we have that $\mathrm{Exp}_e = \exp$.*

Moreover, given any $p \in G$ the map $I_p : G \to G$ such that $I_p(q) = pqp^{-1}$ is an isometry which fixes p and such that, for any geodesics γ, passing through p, $I_p(\gamma(t)) = \gamma(-t)$. This implies that any Lie group that admits a bi-invariant metric is a Riemannian locally symmetric space (see [30])

The geodesics satisfy a local length minimizing property

Definition 3.29. *The arc length of a curve $gamma : [t_0, t_1] \to M$ is defined as*

$$l(\gamma) := \int_{t_0}^{t} \left|\frac{d\gamma}{d\tau}\right| d\tau$$

If γ is a geodesics it follows that

$$\frac{d}{dt}\left|\frac{d\gamma}{dt}\right| = 2\left\langle \frac{D}{dt}\frac{d\gamma}{dt}, \frac{d\gamma}{dt}\right\rangle = 0$$

so for any geodesics $\frac{d\gamma}{dt}$ has constant norm. If such norm is 1 the geodesics is said normalized.

Proposition 3.30 ([23] chapter 3 proposition 3.6 and corollary 3.9)**.** *Let $p \in M$, U a normal neighbourhood of p and $B_\epsilon(p) \subseteq U$ a geodesic ball. Let $\gamma : [0,1] \to B$ the segment of geodesic with $\gamma(0) = p$. For any piecewise differentiable curve joining $\gamma(0)$ and $\gamma(1)$ we have $l(\gamma) \leq l(c)$. The equality holds if and only if $\gamma([0,1]) = c([0,1])$.*

Conversely, if a piecewise differentiable differentiable curve $\gamma : [a,b] \to M$ with parameter proportional to the arc length has length

$$l(\gamma) = \min\{l(c) \textit{ with } c : [a,b] \to M \textit{piecewise differentiable curve} : c(a) = \gamma(a) \textit{ and } c(b) = \gamma(b)\}$$

then γ is a geodesic.

Given a Riemannian manifold it is always possible to define a distance as

$$d(p,q) := \inf\{l(c) \text{ with } c : [a,b] \to M \text{piecewise differentiable curve} : c(a) = p \text{ and } c(b) = q\} \quad (3.6)$$

Proposition 3.31. *$d(.,.)$ defined in equation (3.6) is indeed a distance and the topology induced by such distance coincide with the original topology of M.*

Moreover, for any fixed $p_0 \in M$ the function $f(p) := d(p,p_0)$ is continuous

Proof. The only non straightforward result in the definition of a distance is that if $d(p,q) = 0$ so $p = q$.

If this is not true there exists a geodesic ball $B_\epsilon(p)$ that doesn't contains q. Consider a curve $c : I \to M$ that join p and q. The length of the segment $c(I) \cap B_\epsilon(p)$ is greater or equal than ϵ for proposition 3.30, but this is impossible for the definition of infinimum.

By the definition of d is immediate to see that metric balls contains geodesics balls. the converse follow by proposition 3.30 if we take ϵ so that there exists a minimizing geodesic γ joining p and q. indeed in such case $d(p,q) = l(\gamma)$.

The continuity of f follow by the equivalence of the two topologies (the distance is continuous in the metric topology). $\square$

Definition 3.32. *A Riemannian manifold is said geodesically complete if for any $p \in M$ Exp_p is defined for all $v \in T_pM$ i.e if any geodesic starting at p is defined for any $t \in \mathbb{R}$*

Theorem 3.33 (Hopf and Rinow [23] chapter 7 theorem 2.8)**.** *Let M a Riemannian manifold. the following assertions are equivalent:*

1. *there exists $p \in M$ such that Exp_p is defined on all T_pM*
2. *M has the Heine-Borel property, i.e. any closed bounded set is compact*
3. *M is complete as a metric space*
4. *M is geodesically complete*

If any of such statement is true also holds that for any $p,q \in M$ there is a geodesics γ joining p and q such that $l(\gamma) = d(p,q)$.

Because any compact set is a complete metric space we have the following result:

Corollary 3.34. *Any compact Riemannian manifold is geodesically complete.*

Because closed subset of a complete space are complete we have also that

Corollary 3.35. *Any closed submanifold of a complete Riemannian manifold is geodesically complete in the induced metric. In particular, any closed submanifold of $\mathbb{R}^n$ is geodesically complete with respect to the induced metric.*

We end the section with this result on universal covering of complete Riemannian manifolds

Lemma 3.36. *Let M A smooth manifold and $\pi : \tilde{M} \to M$ the universal covering. Given any Riemannian metric g on M there exists a unique metric $\tilde{g}$ on $\tilde{M}$ such that π is a local isometry.*

If (M,g) is complete so it is $(\tilde{M},\tilde{g})$

Proof. Let $\tilde{g} := \pi^* g$. It is straightforward to check that w.r.t. such metric, π is a local isometry. If $\tilde{g}$ is a local isometry $\tilde{g}(X,Y)_q = g(\pi_* X, \pi_* Y)_{\pi(q)}$ so such metric is unique.

Suppose M complete and $\gamma : (a,b) \to \tilde{M}$ is an incomplete maximal geodesics. $\pi(\gamma)$ is a geodesic (because π is a local isometry) and can be extended to a geodesics $\gamma' : (a, b+\epsilon) \to M$. By lifting such geodesic we obtain an extension of γ, but this is a contradiction. $\square$

3.4 Horizontal vector spaces

We are interested in studying solution of stochastic differential equation on manifold. As we will see the definition of a solution of an SDE on a manifold will rely on the Whytney embedding theorem 1.11. To characterize a such solution it is possible to lift the processes to the fiber bundle.

In this section we will define such bundle and we will give a decomposition of its tangent space in its vertical and horizontal components. During all the section (M, g, ∇) denotes a Riemannian manifold endowed with the Levi-Civita connection

Definition 3.37 ([32] chapter 2.1)**.** *A frame at $p \in M$ is an $\mathbb{R}$-linear isomorphism $u : \mathbb{R}^d \to T_pM$*

The space of all the frames at $p \in M$ is denoted as $\mathfrak{F}(M)_p$. The frame bundle is defined as $\mathfrak{F}(M) := \bigsqcup_{p\in M} \mathfrak{F}(M)_p$. It is possible to show that is a smooth manifold.

Moreover, each fiber of $\mathfrak{F}(M)$ is diffeomorphic to $GL(d, \mathbb{R})$ and $M = \mathfrak{F}(M)/GL(d, \mathbb{R})$. i.e $(\mathfrak{F}(M), M, GL(n, d))$ is a principal bundle and $TM = \mathfrak{F}(M) \times_{GL(d,\mathbb{R})} \mathbb{R}^d$ with respect to the left action $(u, e) \to ue$.

The canonical projection $\pi : \mathfrak{F}(M) \to M$ is a smooth map.

The tangent space at u of the frame bundle is denoted as $T_u\mathfrak{F}(M)$. $X \in T_u\mathfrak{F}(M)$ is called vertical if it is tangent to $\mathfrak{F}(M)_{\pi u}$. The space of vertical tangent vectors at u is denoted with $V_u\mathfrak{F}(M)$.

A curve u_t in $\mathfrak{F}(M)$ is a smooth choice of frames at each point of the curve πu_t. u_t is an horizontal curve (w.r. to ∇) if for each $e \in \mathbb{R}^d$ u_te is parallel along the curve πu_t. A tangent vector is horizontal if it is the tangent vector of an horizontal curve. The space of all the tangent vectors is denoted as $H_u\mathfrak{F}(M)$. It is possible to show [32, chapter 2.1] that

$$T_u\mathfrak{F}(M) = V_u\mathfrak{F}(M) \oplus H_u\mathfrak{F}(M)$$

So given a frame u_p at p, the map $\pi_* : H_u\mathfrak{F}(M) \to T_{\pi u}M$ induced by the canonical projection is an isomorphism.

Definition 3.38 ([32] chapter 2.1)**.** *Given a curve γ_t in M and a frame u_t there is an unique horizontal curve $\{u_t\}$ such that $\pi u_t = \gamma_t$. It is called the horizontal lift of γ_t from u_0.*

For each $e \in \mathbb{R}^d$ we define H_e vector field on $\mathfrak{F}(M)$ by the relation

$$H_e(u) = \{\textit{the horizontal lift of } ue \in T_{\pi u}M \textit{ to } u\}$$

Given a set of coordinates $(e_i)_{i=1,\cdots d}$ on $\mathbb{R}^d$ we define the fundamental horizontal vector field of $\mathfrak{F}(M)$ as $H_i = H_{e_i}$

We have the following coordinate description of such vector fields

Theorem 3.39 ([32] proposition 2.1.3)**.** *Let (U, x^i) a local chart of M and consider the chart on $\mathfrak{F}(M)$ $(\pi^{-1}(U) = \tilde{U}, (x^i, e^i_j)$ where e^i_j : for a frame $u \in \tilde{U}$ we have $ue_j = e^i_j \frac{\partial}{\partial x^i}$.*

In this chart $V_u\mathfrak{F}(M)$ is spanned by $\frac{\partial}{\partial e^i_j}$ so $(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial e^i_j})$ span $T_u\mathfrak{F}(M)$ for every $u \in \tilde{U}$.

In this chart the fundamental vector field have the following local expression:

$$H_i(u) = e^j_i \frac{\partial}{\partial x^j} - e^j_l e^l_m \Gamma^k_{jl} \frac{\partial}{\partial e^k_m} \tag{3.7}$$

Definition 3.40. *Let $\{u_t\}$ the horizontal lift of a curve γ_t. The anti-development of γ_t is a curve in $\mathbb{R}^d$ defined as*

$$w_t := \int_0^t u_s^{-1} \frac{d}{dt}\gamma_s ds$$

It satisfies the following ODE on $\mathfrak{F}(M)$.

$$\frac{d}{dt}u_t = H_i(u_t)\frac{d}{dt}w^i_t \tag{3.8}$$

The development in $\mathfrak{F}(M)$ of a real curve w_t at the starting frame at x_0 u_0 is the unique solution of the ODE (3.8). The development of w_t in M is the projection of the development in $\mathfrak{F}(M)$

3.5 Manifolds with bounded geometry

An homogeneous space for a Lie group G is a not-empty smooth manifold M such that G act transitively on M.

In section 4.4 we will describe how to obtain order condition for a weak stochastic RK method in $\mathbb{R}^n$. Such estimate will depend from a formal series expansion that can be derived from the Kolmogorov backward formula i.e. the Talay-Tubaro expansion (I.2) (see also theorem 5.29).

We wish to expand such method to a general diffusion process over a Lie group or, more generally, over an homogeneous space.

We will now define a class of manifold for which the diffusion process will be a Feller process and it will satisfy the hypothesis of theorem 5.29, under some technical conditions.

Definition 3.41. *Given an (r,s)-tensor T in a Riemannian manifold (M,g) the pointwise norm $||.||_g$ is defined, for a point $p \in M$*

$$||T_x||^2 := g_{p_1,l_1} \cdots g_{p_r,l_r} g^{q_1,k_1} \cdots g^{q_s,k_s} T^{p_1,\cdots,p_r}_{q_1,\cdots,q_s} T^{l_1,\cdots,l_r}_{k_1,\cdots,k_s}$$

where $g_{ij} = g(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j})$ and $g^{ij} = g_{ij}^{-1}$ in a chart (U, x^i) containing p

Definition 3.42. *Let (M, g, ∇) a Riemannian manifold endowed with a connection. The injectivity radius at $p \in M$ is defined as*

$$r(p) = \sup\{r > 0 : \mathrm{Exp}_p : B_r(0) \subseteq T_pM \to M \text{ is a diffeomorphism}\}$$

The global injectivity radius is defined as

$$r(M) = \inf_{p \in M} r(p)$$

We have the following result:

Theorem 3.43 ([49] lemma 9)**.** *If a Riemannian manifold has a strictly positive global injective radius it is geodesically complete*

Definition 3.44. *A Riemannian manifold (M, g, ∇) endowed with the levi-Civita connection is said of k-th order bounded geometry if it has strictly positive injectivity radius and*

$$||||\nabla^k R||_g||_\infty < \infty$$

We will give two important examples of manifold of bounded geometry of any order.

We start with the following lemma:

Lemma 3.45. *Let (M, g) a compact Riemannian manifold, so it has positive injectivity radius.*

Proof. For any $p \in M$ there is a positive number $r(p) > 0$ such that p is contained in a $\delta(p)$-normal neighbourhood. Such neghbourhoods form a open cover of M, so they admits a finite subcover. So $r(M) = \min_{1 \le i \le N} r(p_i)$ that is stricly positive. □

Example 3.46. *Any compact Riemannian manifold M is of bounded geometry of any order*

Proof. By lemma 3.17 any covariant derivative of the curvature tensor is continuous, so its norm is upper semi-continuous.

It is a well known result of general topology that if $f : K \to \mathbb{R}$ is a upper semi-continuous function from a compact set over $\mathbb{R}$, it is bounded above, so the derivatives of the curvature tensor are bounded at any order. By lemma 3.45 we obtain that M is of bounded geometry. □

Example 3.47. *Smooth Riemannian manifold which possess a transitive group of isometries have bounded geometry. In particular, any homogeneous space with an invariant metric and any Lie group with a left (or right) invariant metric has bounded geometry*

Proof. Because there is a transitive group of isomorphism on the manifold the finite estimate for the injectivity radius and the covariant derivative translate to an uniform estimate for all the points of the manifold □

Other examples of manifold of bounded geometry and a description of their basic properties are shown in [24].

4 Stochastic differential equations

4.1 Theory of stochastic integration

We give a rapid survey of the necessary background to understand the theory of stochastic differential equation.

Given a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, where Ω is a sample space, $\mathcal{F}$ a σ-algebra and $\mathbb{P}$ is a probability measure and a measurable space(S, Σ) where S is measurable with respect to the σ-algebra Σ, a stochastic process is a collection of S-valued random variables

$$\{X_t : t \geq 0\}$$

The finite dimensional distribution of the process X are the measures $\mu_{t_1,..,t_n}(F_1 \times ... \times F_k) = \mathbb{P}[X_{t_1} \in F_1, ..., X_{t_k} \in F_k]$ where $F_1, ..., F_k \in \Sigma$.

Two process are said to be equal in law if they have the same finite dimensional distributions.

Definition 4.1. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ a probability space. Let $(\mathcal{F}_t)$ a filtration i.e an increasing family of σ-algebras of subset of Ω.*

A function $g : \mathbb{R} \times \Omega$ is said $\mathcal{F}_t$ adapted if $\omega \to g(t, \omega)$ is $\mathcal{F}_t$-measurable for any $t \geq 0$.

Given a probability space the smallest filtration for which a process X_t is adapted is called the filtration generated by the process.

Definition 4.2. *Let $(\Omega, \mathcal{F}, P)$ a probability space, $\mathcal{F}_t$ a filtration. A random variable τ is called a stopping time with respect to the filtration $\mathcal{F}_t$ if $\{\tau \leq t\} \in \mathcal{F}_t$ for each $t \geq 0$*

We will now define the most general process that can be the integrator of a stochastic imtegral: the semi-martingales.

Definition 4.3. *Let Ω an open set of $\mathbb{R}^n$. Let $u \in L^1(\Omega)$. The total variation of u is defined as*

$$V(u, \Omega) := \sup\left\{\int_\Omega u\nabla \cdot \phi : \phi \in C^1_c(\Omega)||\phi||_{L^\infty} \leq 1\right\}$$

The set of function of bounded variation over Ω is defined as

$$\mathrm{BV}(\Omega) := \{u \in L^1(\Omega) : V(u, \Omega) < \infty\}$$

Definition 4.4. *A stochastic process X_t is called a martingale with respect to a filtration $\mathcal{F}_t$ if X_t is $\mathcal{F}_t$-adapted and*

$$\begin{cases} \mathbb{E}[|X_t|] < \infty \\ \mathbb{E}[X_t|\mathcal{F}_s] = X_s \end{cases}$$

where $\mathbb{E}[.|\mathcal{F}_s]$ is the conditionated probability with respect to $\mathcal{F}_t$.

A stochastic process is a local martigale (w. r. to a filtration) if it is an adapted stochastic process such that exists a sequence of $\mathcal{F}_t$-stopping times $\{\tau_n\}$ such that

$$\mathbb{P}\left[\tau_k < \tau_{k+1}\right] = 1$$

$$\mathbb{P}\left[\lim_{k\to\infty} \tau_k = \infty\right] = 1$$

$$X_{\min(t,\tau_k)} \text{ is a martingale for any } k \geq 0$$

A process is called a semi-martigale if can be decomposed as a sum of a local martingale and a cadlag process with bounded variation

Definition 4.5. *a d-dimensional stochastic process W_t is called a Wiener process (or Brownian motion) if*

1. *$W_0 = 0$*
2. *W_t has independent increments*
3. *W_t is adapted*
4. *$W_{t+u} - W_t$ is a Gaussian process with mean 0 and variance u*
5. *W has continuous paths i.e W_t is continuous for a.e t*

It is possible to show that such process exists ([40] definition 5.1)

Given a σ-algebra $\mathcal{F}$ it is always possible to consider its completion i.e. the σ-algebra generated by $\mathcal{F} \cup \mathcal{N}$ where $\mathcal{N}$ is the set of elements with measure 0 w.r.t. the probability measure. In similar way it is possible to define the completion of a filtration so that all the elements of $\mathcal{N}$ are measurable.

It is possible to show that if W_t is a Wiener process respect a particular filtration is a Wiener process, remains a Wiener process also respect its completion.([40] theorem 7.9)

Moreover it is possible to show that the completion of the filtration generated by the Wiener process satisfies

$$\mathcal{F}_t = \bigcap_{\epsilon > 0} \mathcal{F}_{t+\epsilon}$$

([40] proposition 7.7)

Given a d-dimensional semimartigale Z_t and an L^2 function $g : \mathbb{R} \times \mathbb{R}^{m \times d} \to \mathbb{R}^m$ it is possible to construct a new process, called stochastic integral , we refer to [40] or [59] for the details.

Consider a partition $0 = t_0 < t_1, ..., < t_N = 1$. It is possible to show that the sum

$$\sum_{i=1}^{N} g(\tau_i, Z_{\tau_i})(Z(t_i) - Z(t_{i-1})$$

converges in mean square to different values depending on $\tau_i := \theta t_i + (1-\theta) t_{i-1}$.

If $\theta = 0$ this sum converges to the Itô integral and it is denoted with $\int_0^T g(s, Z_s) dZ_s$

If $\theta = \frac{1}{2}$ this sum converges to the Statonovich integral and it is denoted with $\int_0^T g(s, Z_s) \circ dZ_s$ ([59])

Definition 4.6. *Given a filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, \mathcal{F}_t)$ we define the space $L^2_{ad}([0,t] \times \Omega)$ of square integrable, adapted processes. It is a normed space with norm*

$$\langle X_t, Y_t \rangle_{L^2_{ad}([0,t]\times\Omega)} := \mathbb{E}\left[\int_0^t X_s Y_s ds\right]$$

It is possible to show that if $g \in L^2$ adn W_t is a Brownian motion the Itô stochastic integral $\int_0^t g(s) dW_s$ is a martingale ([59] theorem 3.2.1). Moreover, we have the following important formula for the Itô integral:

Theorem 4.7 (**Itô Isometry** ([59] Corollary 3.1.7))**.** *Let W_t a Brownian motion and X_t, Y_t stochastic processes adapted to the natural filtration of the Brownian*

$$\mathbb{E}\left[\int_0^t X_s dW_s \int_0^t Y_s dW_s\right] = \mathbb{E}\left[\int_0^t X_t Y_t dt\right] \tag{4.1}$$

In other words the Itô integral thought as a function from the space $L^2_{ad}([0,t] \times \Omega)$ to $L^2(\Omega)$ is an isometry

Definition 4.8. *Let W_s is a d-dimensional Wiener process, $\mathcal{F}_t$ a filtration such that W_t is a martingale w.r.to the filtration. An Itô process is a stochastic process X_t on a probability space of the form*

$$X_t(\omega) = X_0 + \int_0^t u(s,\omega) ds + \int_0^t v(s,\omega) dW_s$$

denoted as $dX_t = u_t dt + v_t dW_t$ where u_t is $\mathcal{F}_t$-adapted and v_t is progressively measurable i.e

1. *$(t,\omega) \to f(t,\omega)$ is $B \times F$ measurable (where B is the Borel σ-algebra on $[0,\infty]$)*
2. *$f(t,\omega)$ is $\mathcal{F}_t$ adapted*
3. *$f \in L^2_{loc}([0,\infty])$*

Theorem 4.9 (**Itô's lemma** [59] theorem 4.2.1)**.** *Let $dX_t = u_t dt + v_d W_t$ an m-dimensional Itô process. Let $g : \mathbb{R}^+ \times \mathbb{R}^m$ a C^2 map. Than $g(t, X_t)$ is an Itô process whose components are given by*

$$d(g_k(t,X)) = L^0 g_k(t,X) dt + L^1 g_k(t,X) dW_t \tag{4.2}$$

where

$$L^0 = \frac{\partial}{\partial t} + u_t \frac{\partial}{\partial x} + \frac{1}{2} v_t^2 \sum_{i,j} \frac{\partial^2}{\partial x^i \partial x^j}$$

$$L^1 = v_t \sum_i \frac{\partial}{\partial x^i}$$

If the process is defines by using the Stratonovich integral, the Itô's formula becomes

$$d(g_k(t,X)) = (\frac{\partial}{\partial t} + u\frac{\partial}{\partial x})g_k(t,X) + L^1 g_k(t,X)dW_t$$

4.2 Solution of stochastic differential equation

We define now the strong solution for stochastic differential equations

Definition 4.10. *Let $(W_t)_{t\geq 0}$ be a Brownian motion with respect to a filtration $\mathcal{F}_t$. A progressively measurable process X_t is a (strong) solution with initial condition ξ if*

$$\begin{cases} X_t - X_0 = \int_0^t a(s,X_s)ds + \sum_r \int_0^t b_r(s,X_s)dW_r(s) \\ X_0 = \xi \end{cases} \tag{4.3}$$

holds for a.e. $t \geq 0$

Definition 4.11. *Let $a : \mathbb{R} \times \Omega$ and $b : \mathbb{R} \times \Omega$ two progressively measurable function. Let X_t a strong solution of equation (4.3). That is said to be pathwise unique if, given any other solution Y_t of (4.3) we have that*

$$\mathbb{P}(\sup_{t=0}(|X_t - Y_t|) = 0) = 1$$

We have the following result of (strong) existence and unicity for an SDE

Theorem 4.12 ([59], theorem 5.2.1)**.** *Let $T > 0$ and $a(.;.) : [0;T] \times \mathbb{R}^n \to \mathbb{R}^n$; $b(.;.) : [0;T] \times \mathbb{R}^n \to \mathbb{R}^{n\times m}$ be measurable functions such that for any $t \in [0,T]$, $x, y \in \mathbb{R}^n$ satisfy*

$$|a(t,x)| + |b(t,x)| \leq C(1 + |x|) \tag{4.4}$$

for some constant C.

$$|a(t,x) - a(t,y)| + |b(t,x) - b(t,y)| \leq D|x - y| \tag{4.5}$$

for some constant D.

Let Z be a random variable which is independent of the σ-algebra generated by the Wiener process W_t and such that

$$\mathbb{E}[|Z^2|] < \infty$$

Then the stochastic differential equation of equation (4.3) has a unique t-continuous solution $X_t(\omega)$ with the property that X_t is adapted to the filtration generated by Z and W_t and

$$\mathbb{E}\left[\int_0^T |X_t|^2 dt\right] < \infty$$

Given a semi-martingale Z_t and a matrix valued vector field V it is possible to define the Stratonovich SDE

$$X_t = V_i(X_t) \circ dZ_t^i$$

By converting the Stratonovich integral to the Itô integral (see [6] page 137) we obtain the equivalent formulation for the SDE:

$$X_t = V_i(X_t)dZ_t^i + \frac{1}{2}\nabla_{V_j}V_i(X_t)d\left\langle Z^i, Z^j\right\rangle_t$$

where $d\left\langle Z^i, Z^j\right\rangle_t$is the covariation process i.e

$$d\left\langle Z^i, Z^j\right\rangle_t = \lim_{|\pi|\longrightarrow 0}(Z^i_{t_k} - Z^i_{t_{k-1}})(Z^j_{t_k} - Z^j_{t_{k-1}})$$

where π with range over the partitions of the interval $[0,t]$.

In particular is known that if Z_t is a Brownian motion $d\left\langle Z^i, Z^j\right\rangle_t = \delta^{i,j}dt$ ([40] chapter 3.3) so in this case the Itô formulation became

$$X_t = V_i(X_t)dZ_t^i + \frac{1}{2}\sum_i \nabla_{V_i}V_i(X_t)dt$$

By the Itô's formula we obtain that for any $f \in C^2([0,T])$

$$f(X_t) = f(X_0) + \int_0^t V_i(f(X_t)) \circ dZ_t^i$$

$$f(X_t) = f(X_0) + \int_0^t V_i f(X_t) dZ_t^i + \frac{1}{2}\sum_i \nabla_{V_i} V_i(f(X_t))dt$$

In the next sections we will describe various numerical time-discretization methods for numerical solution of SDEs. In order to evaluate the accuracy of such methods there are two ways: strong convergence and weak convergence Suppose $\bar{X}_N$ be a numerical approximation to X_{t_N} after N steps with constant step size $h = \frac{t_N}{N}$. $\bar{X}_N$ is said to converges strongly to X with order p if $\exists C > 0$ and $\delta > 0$ such that for each $h \in (0, \delta)$

$$\mathbb{E}[|\bar{X}_N - X_{t_N}|] \leq Ch^p$$

$\bar{X}_N$ is said to converges weakly with order p to X if for each test function ϕ in a suitable space there is a constant $C > 0$ and $\delta > 0$ such that, for all $h \in (0, \delta)$ such that

$$|\mathbb{E}[\phi(\bar{X}_N)] - \mathbb{E}[(\phi(X_{t_N})]| \leq Ch^p$$

4.3 Itô-Taylor expansion

Let $a : \mathbb{R} \times \Omega \to \mathbb{R}^n, \{\sigma_r : \mathbb{R} \times \Omega \to \mathbb{R}^{n\times d}\}_{r=1,\cdots,q}$ progressively measurable functions. Let $\{W_r\}_{r=1,\cdots,n}$ n independent d-dimensional Brownian motions. Consider the system of (Itô) SDEs

$$dX = a(t, X)dt + \sum_{r=1}^{n} \sigma_r(t, X)dW_r(t) \tag{4.6}$$

It is possible to define an analogous to the Taylor expansion by recursively applying the Itô's formula (4.2).

Definition 4.13. *Let $\alpha = (j_1, \cdots, j_l)$ a multi-index. Define $\alpha- = (j_1, \cdots, j_{l-1})$. Let $f : \mathbb{R}\times\mathbb{R}^n \to \mathbb{R}^n$ a smooth function.*

Let ρ, τ two stopping times $0 < \rho(\omega) < \tau(\omega) < T$

The multiple Wiener (Itô) integral is defined recursively as follow:

$$I_\alpha[f(., X_.)]_{\rho,\tau} = \begin{cases} f(\tau) \ \text{if } \alpha = \emptyset \\ \int_\rho^\tau I_{\alpha-}[f(., X_.)]_{\rho,s}ds \ \text{if } j_l = 0 \\ \int_\rho^\tau I_{\alpha-}[f(.)]_{\rho,s}dW_s \ \text{if } j_l > 0 \end{cases}$$

We have the following result on the first momentum of the Itô multiple integral:

Theorem 4.14 ([41] lemma 5.7.1)**.** *Let α a multi-index such that at least one of its terms it is different by 0, f a smooth function and $0 < \rho < \tau < T$ two stopping times from the interval $[0, T]$. Then almost surely*

$$\mathbb{E}[I_\alpha[f]_{\rho,\tau}|\mathcal{F}_\rho] = 0$$

Theorem 4.15 ([41] theorem 5.5.1)**.** *A set of multi-indexes $\mathcal{A}$ is called hierarchical if it is not empty, all the multi-indexes of $\mathcal{A}$ are of finite length and for each $\alpha = (j_1, \cdots, j_l) \in \mathcal{A}/\emptyset$ we have that $-\alpha := (j_2, \cdots, j_l) \in \mathcal{A}$.*

The remainder set of a hierarchical set is defined as $\mathcal{B}(\mathcal{A}) = \{\alpha \text{ multi-index } : \alpha \notin \mathcal{A} : -\alpha \in \mathcal{A}\}$.

Given X_t solution of the SDE (4.6) define the diffusion operators of the process as

$$L^0 = \frac{\partial}{\partial t} + a^k \frac{\partial}{\partial x^k} \sum_i \sigma_i^k \sigma_i^l \frac{\partial^2}{\partial x^k \partial x^l}$$

for $j \in \{1, \cdots, n\}$ define

$$L^j = \sigma_j^k \frac{\partial}{\partial x^k}$$

Let ρ, τ 2 stopping times $0 \leq \rho(\omega) \leq \tau(\omega) < T$, $\mathcal{A}$ a hierarchical set and $f : \mathbb{R}^+ \times \mathbb{R}^d \to \mathbb{R}$ a smooth function. The Itô Taylor expansion

$$f(\tau, X_\tau) = f(\rho, X_\rho) + \sum_{\alpha=(j_1,\cdots,j_l)\in\mathcal{A}/\emptyset} I_\alpha[L^{j_1}, \cdots, L^{j_l} f(\rho, X_\rho)]_{\rho,\tau} + \sum_{\alpha=(j_1,\cdots,j_l,j_{l+1})\in\mathcal{B}(\mathcal{A})} I_\alpha[L^{j_1}, \cdots, L^{j_{l+1}} f(., X_.)]_{\rho,\tau}$$

holds if all the derivatives and the multiple integrals of the definition are well defined

Example 4.16. *If $n = 1$ and $\mathcal{A} = \{\emptyset\}$ we have that $\mathcal{B}(\mathcal{A}) = \{(0), (1)\}$ so the Taylor Itô expansion became:*

$$f(\tau, X_\tau) = f(\rho, X_\rho) + \int_\rho^\tau L^0 f(s, X_s) ds + \int_\rho^\tau L^1 f(s, X_s) dW_s$$

That is the Itô formula of equation (4.2)

We have the following result about the weak convergence of numerical schemes for SDEs.

Theorem 4.17 ([41] Corollary 5.12.1)**.** *Let $a : \mathbb{R} \times \Omega \to \mathbb{R}^n, \{\sigma_r : \mathbb{R} \times \Omega \to \mathbb{R}^{n\times d}\}_{r=1,\cdots,q}$ progressively measurable functions which satisfy the linear growth condition (4.4) and the Lipschitz condition (4.5). Suppose X_t is the solution of equation (4.6) with respect to a filtration $\mathcal{F}_t$.*

Let β a positive integer. $\Lambda_\beta = \{\alpha$ multi-indexes $: |\alpha| \leq \beta\}$.

for each $\alpha = (j_1, \cdots, j_n)$ let $f_\alpha = L^{j_1}, \cdots, L^{j_l} f$ where f is a smooth function. Let $\tilde{T}_{\alpha,t_0,.}$ some random variables such that there exists some constant K such that for any $\alpha_k \in \Lambda_\beta / \emptyset$

$$|\mathbb{E}[\prod_{k=1}^{l} I_{\alpha_k,t_0,t} - \prod_{k=1}^{l} \tilde{I}_{\alpha_k,to,t} | \mathcal{F}_{t_0}]| \leq K(t - t_0)^{\beta+1}$$

for $l = 1, \cdots, 2\beta + 1$. For $t \in [t_0, T]$ define

$$U_\beta(t) = \sum_{\alpha \in \Lambda_\beta} f_\alpha(t_0, X_{t_0}) \tilde{I}_{\alpha,t_0,t}$$

Than for any test function $\phi \in C^{2(\beta+1)}$ there exists a constant C_ϕ and a positive integer r such that

$$|\mathbb{E}\left[\phi(X_t) - \phi(U_\beta(t)) | \mathcal{F}_{t_0}\right]| \leq C_\phi (1 + |X_{t_0}|^{2r})(t - t_0)^{\beta+1}$$

with $X_t = X_t^{t_0, X_{t_0}}$

Example 4.18. *Consider the SDE of equation (4.3) with one Brownian motion. Let $\phi \in C^2$ a test function. Consider the truncated Taylor expansion given by*

$$\phi(X_{t+h}) = \phi(X_{t_0}) + \int_{t_0}^{t_0+h} L^0 \phi(X_s) ds \tag{4.7}$$

Let $X_{t_0} = x$. By using the Itô formula we obtain that up to order one

$$\mathbb{E}[\phi(X_s) | X_0 = x] = \phi(x) + hL^0\phi(x) \tag{4.8}$$

Suppose now X_1 is a numerical scheme for X_t. Suppose this schemes has a weak Taylor expansion

$$\mathbb{E}\left[\phi(X_1) | X_0 = x\right] = \phi(x) + hA_0\phi(x) + h^2 A_1 \phi(x) + \cdots$$

To have (local) weak order 1 it is necessary and sufficient that $A_0 = L^0$. (see theorem 5.29)

4.4 The exotic aromatic trees formalism

We will now describe how by extending the classical Buthcer tree formalism described in [27, chapter III] is it possible to simplifying the calculation that occurs in the SRK method. Following [45] we will consider the particular case of additive noise

$$dX(t) = f(X_t) + \sigma dW_t$$

where $\sigma > 0$ is constant, W_t is a d-dimensional Brownian motion and $f : \mathbf{R^d} \to \mathbf{R^d}$ is a smooth function that admits a C^∞ potential V such that $f(x) = \nabla V(x)$ is globally Lipschitz.

The stochastic RK scheme that we consider is

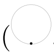
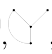
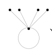
$$\begin{cases} X_0 = x \\ Y_i = X_n + h\sum_j a_{ij} f(Y_j) + \sum_k d_i^{(k)} \sigma \sqrt{h} \xi_n^{(k)} \\ X_{n+1} = X_n + h\sum_i b_i f(Y_i) + \sigma\sqrt{h}\xi_n^{(1)} \end{cases} \tag{4.9}$$

where a_{ij}, b_i, $d^{(k)}$ are coefficients and $\xi_n^{(k)}$ are independent standard Gaussian random variables (see [67] for an analysis on the convergence of this method).

In the study of numerical solution of SDE the divergence and the Laplacian operators arise naturally and the classical Butcher theory as no way to represent it.

Definition 4.19. *The set of aromatic forest is a directed graphs$\gamma = (V, E)$ in which any node has at most one outgoing edge. Its connected components are called aromatic trees.*

There are two kind of aromatic trees:

- Aromas that are aromatic trees connected with one single edge (, , )
- rooted trees

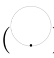
Definition 4.20. *The elementary differential of an aromatic forest γ is defined as follow: denote $\pi(v) = \{w \in V, (w, v) \in E\}$ the set of all predecessors of the node $v \in V$ and r the root of γ. We also call $V_0 = V/r = \{v_1, ..., v_m\}$ the other nodes of γ. Finally we introduce the notation $I_{\pi(v)} = (i_{q_1}, ..., i_{q_s})$, where the q_k are the predecessors of v and where*

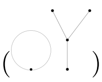
$$\partial_{I_{\pi(v)}} f = \frac{\partial f}{\partial_{x_{i_{q_1}}} ... \partial_{x_{i_{q_s}}}}$$

$F_f(\gamma)$ is defined as:

$$F_f(\gamma) = \sum_{i_{q_1},\dots,i_{q_s}}^{d} \Big(\prod_{v \in V_0} \partial_{I_{\pi(v)}} f_{i_v}\Big) \partial_{I_{\pi(r)}} f$$

Example 4.21.

$F_f($  $) = \nabla \cdot f$

$F_f($  $) = (\nabla \cdot f) f' f''(f, f)$

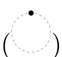
We further extend the set of aromatic trees by adding the concept of "liana", a new kind of edge that is represented as a dashed arc linking two given nodes. Such liana correspond to non oriented edges between two nodes of the forest.

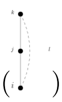
So an exotic aromatic tree is a triple $\gamma = (V, E, L)$ where V are the nodes, E are the edges and L are the lianas. We can define as well a new elemental differential.

Definition 4.22. *We name r the root of $\gamma = (V, E, L)$ and $V_0 = V/r = v_1, ..., v_m$ the other nodes of γ. We denote $l_1, ..., l_s$ the elements of L and for $v \in V$, $J_{\Gamma(v)}$ the multiindex $(j_{l_{x_1}}, ..., j_{l_{x_t}})$ where $\Gamma(v) = \{l_{x_1}, ..., l_{x_t}\}$. Then $F_f(\gamma)$ is defined as*

$$F_f(\gamma) = \sum_{i_{q_1},\dots,i_{q_s}}^{d} \sum_{j_{l_{x_1}},\dots,j_{l_{x_t}}}^{d} \Big(\prod_{v \in V_0} \partial_{I_{\pi(v)}} \partial_{J_{\Gamma(v)}} f_{i_v}\Big) \partial_{I_{\pi(r)}} \partial_{J_{\Gamma(r)}} f \tag{4.10}$$

Example 4.23.

$F_f($  $) = \Delta f$

$F_f($  $) = \sum_{i,j,k=1}^{d} \sum_{l=1}^{d} \partial_l f \partial_i f_j \partial_j \partial_l f_k = \sum_l \partial_l f'(f'(\partial_l f))$

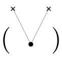
We also need a new kind of node to describe the Gaussian random variable which appears in equation (4.9).

Definition 4.24. *A grafted node is a new kind of node that is represented by a cross.*

Let V be a set of nodes whose subset of grafted nodes is V_g let E be a set of edges such that each node in V_g has exactly one outgoing edge and no ingoing edge, and let L be a set of lianas that link nodes in V/Vg; then $\gamma = (V, E, L)$ is a grafted exotic aromatic forest. If $\gamma = (V, E, L)$ is a grafted exotic aromatic rooted forest, $\phi : \mathbf{R}^d \to R$ a smooth function, and ξ a random vector of $\mathbf{R}^d$ whose components are independent and follow a standard normal law, the associated elementary differential of γ is, with the same notation of equation (4.10) and $V_0 = V/\{V_g \cup \{r\}\}$

$$F_f(\gamma) = \sum_{i_{q_1},\dots,i_{q_s}}^{d} \sum_{j_{l_{x_1}},\dots,j_{l_{x_t}}}^{d} \Big(\prod_{v \in V_0} \partial_{I_{\pi(v)}} \partial_{J_{\Gamma(v)}} f_{i_v}\Big)\Big(\prod_{v \in V_g} \xi_{i_v}\Big) \partial_{I_{\pi(r)}} \partial_{J_{\Gamma(r)}} \phi \tag{4.11}$$

Example 4.25.

$F_f($  $) = \phi^{''}(\xi, \xi)$

Definition 4.26. *Let $\mathcal{EAT}_g$ the set of grafted exotic aromatic forest.*

The order of $\tau \in \mathcal{EAT}_g$ is $|\tau| = N(\tau) - \frac{N_c(\tau)}{2} + N_l - 1$ where N is the number of nodes, N_c the number of grafted nodes and N_l the number of lianas.

A grafted exotic aromatic B-series is

$$B(a, \phi) = \sum_{\tau \in \mathcal{EAT}_g} h^{|\tau|} \alpha(\tau) F(\tau)(\phi) \tag{4.12}$$

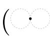
We can now use this new tree formalism to compute the expectation of the first stage of the stochastic RK method (4.9).

Theorem 4.27 ([45])**.** *Let $\gamma \in \mathcal{EAT}_g$ be a grafted exotic aromatic rooted forest with an even number of grafted nodes $2n$, let $\phi : \mathbb{R}^d \to \mathbb{R}$ be a smooth function, and let $V^\times = \{c_1, ..., c_{2n}\}$ be the set of grafted nodes of γ. We call $P_2(2n)$ the set of partitions by pair of $\{1, ..., 2n\}$.*

Finally we define $\psi_\gamma : P_2(2n) \to \mathcal{EAT}_g$ the application that maps the partition p of γ to the aromatic forest where the grafted nodes are linked by lianas according to p. Then, the expectation of $F(\gamma)(\phi)$ is given by

$$\mathbb{E}[(\gamma)(\phi)] = \sum_{p \in P_2(2n)} F(\psi_\gamma(p))(\phi) \tag{4.13}$$

Example 4.28.

$\mathbb{E}[F_f($  $)(\phi)] = 3F_f($  $(\phi)) = 3\Delta^2 \phi$

The number 3 is the number of ways we can pair the grafted nodes.

More in general if you have γ as a tree with only the root and $2n$ grafted nodes the mean is

$$\mathbb{E}[F(\gamma)(\phi)] = \frac{(2n)!}{2^n n!} \Delta^n \phi$$

Let's compute the first stage of equation (4.9)

$$\begin{cases} Y_1 = x + h \sum_j a_{ij} f(Y_j) + \sum_k d_i^{(k)} \sigma \sqrt{h} \xi_n^{(k)} \\ X_1 = x + h \sum_i b_i f(Y_i) + \sigma \sqrt{h} \xi_n^{(1)} \end{cases} \tag{4.14}$$

By inserting the first equation into the second and computing the Taylor series around x we obtain (up to second order in h:

$$\begin{aligned} X_1 = x + \sigma\sqrt{h}\xi_n + h \sum_{i=1}^{s} b_i f + h^{\frac{3}{2}} \sigma \sum_{i=1}^{s} b_i \sum_{k=1}^{l} d_i^(k) f^{'}(\xi_n^{(k)}) + \\ + h^2 \left[\sum_{i=1}^{s} b_i \sum_{j=1}^{s} a_{ij} f'(f) + \frac{\sigma^2}{2} \sum_{i=1}^{s} b_i \sum_{k=1}^{l} (d_i^{(k)})^2 f^{''}(\xi_n^{(k)}, \xi_n^{(k)}) \right] + \cdots \end{aligned}$$

Where we have omitted the dependence on x of f.

Now let ϕ a test function with enough regularity (for example a polynomial). We are interest to calculate $\phi(X_1)$ up to second order. By expanding it in Taylor series around x we obtain:

$$\begin{aligned} \phi(X_1) = \phi + \phi' \Bigg[\sigma\sqrt{h}\xi_n + h \sum_{i=1}^{s} b_i f + h^{\frac{3}{2}} \sigma \sum_{i=1}^{s} b_i \sum_{k=1}^{l} d_i^{(k)} f^{'}(\xi_n^{(k)}) + \\ + h^2 \left[\sum_{i=1}^{s} b_i \sum_{j=1}^{s} a_{ij} f'(f) + \frac{\sigma^2}{2} \sum_{i=1}^{s} b_i \sum_{k=1}^{l} (d_i^{(k)})^2 f^{''}(\xi_n^{(k)}, \xi_n^{(k)}) \right] \Bigg] \\ + \phi^{''} \left[\sigma\sqrt{h}\xi_n + h \sum_{i=1}^{s} b_i f + h^{\frac{3}{2}} \sigma \sum_{i=1}^{s} b_i \sum_{k=1}^{l} d_i^{(k)} f^{'}(\xi_n^{(k)}) + o(h^{\frac{3}{2}}), - \right] \\ + \phi^{(3)} \left[\sigma\sqrt{h}\xi_n + h \sum_{i=1}^{s} b_i f + o(h), -, - \right] + \phi^{(4)} \left[\sigma\sqrt{h}\xi_n, -, -, - \right] \end{aligned}$$

By expanding the product, taking the middle value and remembering that all the not integer powers of h (corresponding to an odd number of normal distributions) disappear we have

$$\mathbb{E}[\phi(X_1)|X_0 = x] = \phi(x) + h\mathbf{L} + h^2 A_1$$

where

$$\mathbf{L} = \mathbb{E}[\sum_{i=1}^{s} b_i \phi^{'}(f) + \frac{\sigma^2}{2}\phi^{''}(\xi_n^{(1)}, \xi_n^{(1)})]$$

$$A_1 = \mathbb{E}[\sum_i b_i \sum_j a_{ij}\phi^{'}(f'(f)) + \frac{\sigma^2}{2}\sum_i b_i \sum_k \phi^{'}(f^{''}(\xi_n^{(k)}, \xi_n^{(k)})) + \sigma^2 \sum_i b_i d_i^{(1)}\phi^{''}(\xi_n^{(1)}, \xi_n^{(1)}) +$$
$$+ \frac{1}{2}\phi^{''}(f, f) + \frac{1}{2}\sigma^2\phi^{(3)}(\xi_n^{(1)}, \xi_n^{(1)}, f) + \frac{\sigma^4}{24}\phi^{(4)}(\xi_n^{(1)}, \xi_n^{(1)}, \xi_n^{(1)}, \xi_n^{(1)})]$$

In the formalism of exotic aromatic trees this two quantities can be rewritten as

$$A_1 = \mathbb{E}[\sum_i b_i \sum_j a_{ij} + \frac{\sigma^2}{2}\sum_i b_i \sum_k d_i^{(k)} + \sigma^2 \sum_i b_i d_i^{(1)} + \frac{1}{2} + \frac{\sigma}{2} + \frac{\sigma^4}{24}]$$

By using theorem 4.27 we obtain that

$$A_1 = \sum_i b_i \sum_j a_{ij} + \frac{1}{2} + \frac{\sigma^2}{2} + \sigma^2 \sum_i b_i d_i^{(1)} + \frac{\sigma}{2} + \frac{\sigma^4}{24}$$

An analysis on the order condition to obtain an approximation of weak order 2 is done in [45].

An important property of a class of exotic aromatic B-series is that they are unchanged under some affine transformations, namely.

Definition 4.29. *Given a subgroup $G \subseteq GL(d, \mathbb{R}) \rtimes \mathbb{R}^d$, a differential operator G is said G-equivariant if for all $(A, b) \in G$ and for all $f \in C^\infty(\mathbb{R}^d, \mathbb{R}^d)$*

$$\Phi(Af(A^{-1}(x - b)) = (A, b) \circ \Phi(f) \circ (A, b)^{-1}$$

The exotic aromatic B-series $B(a)$ with $a() = 1$, where is the empty tree are called exotic aromatic B-series methods. It is possible to prove [45, theorem 3.6] that such methods are isometric equivariant i.e are $O(d, \mathbb{R}) \rtimes \mathbb{R}^d$-equivariant.

The exotic aromatic trees formalism describes weak approximation of SDEs on Euclidean spaces. Finding an extension of such formalism for SDEs on Riemannian manifold is an actual subject of research and it can lead to a better understanding of the properties of the numerical schemes that approximate such processes.

5 Feller semi-groups

5.1 Feller processes

We will now describe a useful tool to calculate the solution of stochastic differential equation. We start with the following definition

Definition 5.1 (Markov processes). *Given a filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t\geq 0})$ and a measurable space S a stochastic adapted process $X_t : \Omega \to S$ is said to have the Markov property if for each measurable $A \in S$*

$$\mathbb{P}\left(X_{s+t} \in A|\mathcal{F}_s\right) = \mathbb{P}\left(X_{s+t} \in A|X_s\right)$$

where the $\mathbb{P}(.|\mathcal{F}_s)$ is the conditionated probability w.r.to the filtration $\mathcal{F}_t$ and $\mathbb{P}(.|X_s)$ are the conditionated probability w.r. to the filtration generated by the process.

If for every stopping time τ we have

$$\mathbb{P}\left(X_{t+\tau} \in A|\mathcal{F}_\tau\right) = \mathbb{P}\left(X_{t+\tau} \in A|X_\tau\right)$$

where $\mathcal{F}_\tau := \{A \in \mathcal{F} : \forall t \geq 0 \ \ (t \leq \tau) \in \mathcal{F}_t\}$,we say that the process has the strong Markov property.

Definition 5.2. *A d-dimensional Markov family is an adapted process X on some measure space $(\Omega, \mathcal{F})$ together with a family of probability measures $\{\mathbb{P}^x\}_{x\in\mathbb{R}^d}$ such that for all $x \in \mathbb{R}^d$ $\mathbb{P}^x(\mathcal{F})$ is universally measurable (i.e it is measurable w. r. to any complete probability measure which measure all the Borel sets)*

$$\mathbb{P}^x(X_0 = x) = 1$$

$$\mathbb{P}^x\left(X_{s+t} \in A|\mathcal{F}_s\right) = \mathbb{P}\left(X_{s+t} \in A|X_s\right)$$

$$\mathbb{P}^x\left(X_t \in A|X_s = y\right) = \mathbb{P}^y\left(X_t \in A\right) \quad P^x X_s^{-1} a.e.$$

Proposition 5.3. *If a Stochastic process has independent increments it is a Markov process. In particular, the Brownian motion is a Markov process*

Proof.

$$\mathbb{E}\left[X_{t+s}|X_s\right] = \mathbb{E}\left[X_{t+s} - X_s + X_S|X_s\right] = \mathbb{E}\left[X_{t+s} - X_s + x\right]_{x=X_s} = \mathbb{E}\left[X_{t+s}|\mathcal{F}_s\right]$$

where we have used lemma A.3 of [71] (the freezing lemma) □

A particular class of Markov families is given by the Feller processes, such processes are characterized by the property of their transition law.

Definition 5.4. *The transition kernel of a stochastic process $p(s, X, s+t, H)$ is given by*

$$p(s, X_s, s+t, H) = \mathbb{P}(X_{s+t} \in H|X_t)$$

a process is called time homogeneous if $p(s, X_s, s+t, H) = p(0, X_s, t, H) := \mu_t(X_s, H)$.

Given a time homogeneous stochastic process, if we call C_0 the set of continuous functions that vanishes at infinity the transition operator associated with the kernel of the process is given by $T_t : C_0 \to C_0$:

$$(T_t f)(x) := \int \mu_t(x, dy) f(y)$$

Because μ_t is a probability measure the transition operator is a positive contraction operator i.e if $0 \leq f \leq 1$ we have $0 \leq T_t f \leq 1$

Proposition 5.5. *The transition operator $\{T_t\}_{t\geq 0}$ associated to a time homogeneous Markov family is a semi-group*

Proof. $\mu_0(x, .) = \delta_x$ the Dirac delta centered in x. The operator associated to such kernel is the Identity operator.

By the Chapman-Kolmogorov equation (see e.g. [60] chapter 2.2) we obtain that $T_t \circ T_s f = T_{t+s} f$ □

Definition 5.6. *A Feller semi-group is a collection of positive contraction linear maps on C_0 such that they form a semi-group w.r. to the composition and $\lim_{t\to 0} ||T_t f - f|| = 0$ i.e it is strongly continuous.*

A Feller process is a time-homogeneous Markov process such that their transition operators form a Feller semi-group.

A Feller semi-group is characterized by its infinitesimal generator.

Definition 5.7. *The infinitesimal generator of a strongly continuous semi-group is given by*

$$\mathcal{L}f = \lim_{t\longrightarrow 0} t^{-1}(T_t f - f)$$

for any $f \in \mathcal{D}(A) := \{f \in C_0 : \exists \lim_{t\longrightarrow 0} t^{-1}(T_t f - f)\}$

Theorem 5.8. *A strongly continuous semi-group is uniquely determined by its infinitesimal generator*

Proof. Let T_t, S_t two strongly continuous semi-groups and let $(A, D(A))$ an operator such that

$$A = \lim_{t\longrightarrow 0} \frac{T_t f - f}{t} = \lim_{t\longrightarrow 0} \frac{S_t f - f}{t} \tag{5.1}$$

because $t \to ||T_t||$ and $t \to ||S_t||$ are continuous there is $C > 0 :$ $||T_t||||S_t|| < C$.

For a fixed $T > 0$ and $\epsilon > 0$, there exists a $\delta > 0$ such that, for any $f \in D(A)$, for any $0 \leq h \leq \delta$

$$h^{-1}||T_h f - S_h f|| < \frac{\epsilon}{TC}$$

Now by using the semi-group property and (5.1) as in [61] we obtain that

$$||T_t f - S_t f|| < \epsilon$$

The result follows by the arbitrariness of ϵ, T and f □

Example 5.9. *The Brownian motion transition operator is given by*

$$T_t f(x) = \frac{1}{2\pi\sqrt{t}} \int_0^t e^{-\frac{(x-y)^2}{t}} dy$$

It is easy to see that has the Feller property and using the Itô's formula we obtain that the infinitesimal generator is the Laplacian

We have the following properties of the Feller semi-group and its infinitesimal generator, of great importance for our analysis

Theorem 5.10 (**Kolmogorov backward and forward equations** [37] theorem 19.6)**.** *Let $\{T_t\}_{t\geq 0}$ a Feller semi-group and $(\mathcal{L}, \mathcal{D}(\mathcal{L}))$ its infinitesimal generator. T_t is strongly continuous and $T_t f$ is differentiable at 0 iif $f \in \mathcal{D}(\mathcal{L})$. Moreover, it satisfies the Kolmogorov backward and forward equations*

$$\frac{d}{dt}(T_t f) = \mathcal{L} T_t f \tag{5.2}$$

$$\frac{d}{dt}(T_t f) = T_t \mathcal{L} f \tag{5.3}$$

In particular, $\mathcal{L}$ commutes with T_t for any t

Example 5.11. *Let $T_t : C_0(\mathbb{R}) \to C_0(\mathbb{R})$ given by*

$$T_t f(x) = f(x+t)$$

It is strongly continuous and $||T_t f|| = ||f||$ so it is a Feller semi-group. Its infinitesimal generator is

$$\mathcal{L} f = \lim_{t \to 0} t^{-1}(f(x+t) - f(x)) = \frac{d}{dx} f(x)$$

The Kolmogorov backward (and forward) equations say that

$$\frac{d}{dt} f(x+t) = \frac{d}{dx} f(x+t)$$

A linear operator $A : D(A) \to H$ from a linear subspace $D(A)$ of a Banach space to an Hilbert space H is said closable if there is an extension $\tilde{A} : D(\tilde{A}) \to H$ with $D(A) \subseteq D(\tilde{A})$ that is closed (i.e $x_n \to x$ and $\tilde{A} x_n \to v$ imply $x \in D(\tilde{A})$ and $Ax = v$). We want to know when a linear operator is closable and its closure is the infinitesimal generator of a Feller semi-group

Theorem 5.12 (**Hille-Yoshida**([37] theorem 19.11))**.** *Let A a linear operator in C_0 with domain $\mathcal{D}(A)$. Then A is closable and its closure is the generator of a Feller semi-group on C_0 iif*

1. *$\mathcal{D}(A)$ is dense in C_0*
2. *there exists $\lambda_0 > 0$ such that $\mathrm{Im}(\lambda_0 I - A)$ is dense in C_0*
3. *(**positive maximum principle**) if $\max(f, 0) \leq f(x)$ for some $f \in \mathcal{D}(A)$ and some x then $Af(x) \leq 0$*

Theorem 5.13 (**Dynkin's formula** ([37] lemma 19.21))**.** *Let X_t a Feller process, $\mathcal{L}$ its infinitesimal generator. Than for any initial distribution ν of X the process and for any $f \in \mathcal{D}(\mathcal{L})$*

$$M_t^f = f(X_t) - f(X_0) - \int_0^t \mathcal{L} f(X_s) ds \tag{5.4}$$

is a martingale. in particular, for any bounded stopping time τ

$$\mathbb{E}\left[f(X_\tau) | X_0 = x\right] = f(x) + \mathbb{E}\left[\int_0^\tau \mathcal{L} f(X_s) ds | X_0 = x\right]$$

We continue by giving an abstract formulation of the Feynman-Kac formula.

Definition 5.14. *The resolvent of a C_0 semi-group T_t is its Laplace transform*

$$R_\lambda = \int_0^\infty e^{-\lambda t}(T_t f)dt$$

If T_t is a Feller semi-group and $\mathcal{L}$ is its generator,then $R_\lambda = (\lambda I - \mathcal{L})^{-1}$ (where I is the identity operator) ([37] theorem 19.4)

Theorem 5.15 (Feynman-Kac formula). *Let $(E, \mathcal{E})$ a measure space and v a not negative $\mathcal{E}$-measurable function such that $R_\lambda v(x) < \infty$, for each x(where R_λ is the resolvent of the transition kernel of X_t). Let $A_t := \int_0^t v(X_s)$, where X_s is a progressively measurable E-valued Feller process in some filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t\geq 0})$. Define*

$$T_t^v f(x) := \mathbb{E}\left[e^{-A(t)} f(X_t) | X_0 = x\right]$$

Let R_λ^v the resolvent of this semi-group and We have that

$$R_\lambda^v = R_\lambda - R_\lambda v R_\lambda^v$$

Moreover if v is a continuous non-negative function and we call $\mathcal{L}$ the generator of X, $\{P_t^v\}_{t\geq 0}$ is a Feller semi-group with generator

$$\mathcal{L}^v f(x) = \mathcal{L} f(x) - v(x) f(x) \tag{5.5}$$

(the theorem still holds in the more general case in which X_t is a Markov process , see ([64] chapter III.19)

Proof. We will give an idea of the proof.

A_t is a perfect, continuous, homogeneous, additive functional (PCHAF) of X (see [64] chapter III.18). In particular we need that it is an $\mathcal{F}_t$ adapted process such that $\mathcal{F}$-a.s. in Ω:

1. $t \to A_t$ is continuous and not decreasing and $A_0 = 0$
2. calling $\theta_t : \Omega \to \Omega$ the time shift $\theta_t(\omega)(s) = \omega(s+t)$ we have

$$A_{s+t}(\omega) = A_s(\omega) + A_t(\theta_t(\omega))\mathcal{F}\text{-a.s} \in \Omega$$

By the continuity of A_t and the additive property 2 and the Feller property of X_t we obtain that T_t^v is a Feller semi-group. Because of the measurability hypothesis we apply the Fubini theorem, so the resolvent of T_t^v is

$$R_\lambda^v f(x) = \mathbb{E}\left[\int_0^\infty dt e^{-\lambda t - A(t)} f(X_t) | X_0 = x\right]$$

for any bounded measurable function that vanish at infinite.

$$\begin{aligned}
&(R_\lambda - R_\lambda^v) f(x) = \\
=&\mathbb{E}\left[\int_0^\infty dt e^{-\lambda t - A(t)} f(X_t)(e^{A(t)} - 1) | X_0 = x\right] = \mathbb{E}\left[\int_0^\infty dt e^{-\lambda t - A(t)} f(X_t) \int_0^t ds\ v(X_s) e^{A(s)} | X_0 = x\right] = \\
=&\mathbb{E}\left[\int_0^\infty ds\ e^{-\lambda s} f(X_s) \int_0^\infty v(X_u \circ \theta_s) e^{-\lambda u - A_u \circ \theta_s} du | X_0 = x\right] = \mathbb{E}\left[\int_0^\infty e^{-\lambda s} v(X_s) R_\lambda^v f(X_s) ds | X_0 = x\right] = \\
=&R_\lambda v R_\lambda^v f(x)
\end{aligned}$$

□

There is a way to associate to any Feller semi-group an associated semigroup of Markov kernels. Indeed by the Markov-Riesz theorem given any linear functional T_t on C_0 there is an unique Radon measure μ_t such that

$$T_t f = \int \mu_t(x) f(x)$$

If such measures are probability measures the system is called conservative.

In that case by the Kolmogorov existence theorem ([40] chapter 2.2 theorem 2.2), given any probability distribution μ_0 there is a process X_t with finite dimensional distributions

$$\mathbb{P}(X_{t_1} \in A_1 \cdots X_{t_n} \in A_n) = \int \chi_{A_1}(x_1) \cdots \chi_{A_n}(x_n) \mu_{t_n - t_{n-1}}(x_{n_1}, dx_n) \cdots \mu_{t_1}(x_0, dx_1) \mu_0(dx_0)$$

Where χ_A is the indicator function of A ($\chi_A(x) = 1$ if $x \in A$, and it is 0 otherwise).

By this formula it is possible to characterize the action of a conservative Feller semi-group as

$$T_t f(x) = \mathbb{E}\left[f(X_t)|X_0 = x\right] \tag{5.6}$$

If the Feller semi-group on a space S is not conservative it is always possible to extend it to a conservative process in the space $S \cup \{\partial\}$, where ∂ is the point at infinity of the one point compactification if S is not compact and an isolated point if S is compact. Any function on $C_0(S)$ can be extended to $C_0(S \cup \{\partial\})$ by imposing $f(\partial) = 0$ and the corresponding Feller semi-group is

$$\hat{T}_t f = f(\partial) - T_t(f - f(\partial))$$

This is a Feller semi-group and generate Markov kernels μ_t such that $\mu_t(x, \{x\}) = 1$ ([37] lemma 19.13 and proposition 19.14) .

The random time $e(X) := \inf_{t \geq 0}\{X_t = \partial\}$ is called the explosion time of the process.

5.2 An example of Feller process: the Itô diffusion

A process is called Itô diffusion (on $\mathbb{R}^n$) if is a solution of the SDE

$$X_t^x = x + \int_0^t b(X_s^x) ds + \int_0^t \sigma(X_s^x) dW_s \tag{5.7}$$

with b, σ globally Lipschitz coefficients.

Proposition 5.16 ([71] theorem 19.9)**.** *The Itô diffusion is a Feller process with Feller semi-group given by*

$$T_t f(x) = \mathbb{E}\left[f(X_t^x)\right]$$

and infinitesimal generator given by

$$\mathcal{L} f(x) = b^j(x) \frac{\partial}{\partial x^j} f(x) + a^{ij}(x) \frac{\partial^2 f(x)}{\partial x^i \partial x^j}$$

where $a(x) = \sigma(x)\sigma^T(x)$

It is easy to see, by using the Itô's formula that M_t^f of equation (5.4) is a martingale.

If we call $u(t,x) = T_t f(x)$ equation (5.2) became

$$\frac{\partial}{\partial t} u(t,x) = \mathcal{L} u(t,x) \text{ and } u(0,x) = f(x)$$

that is the classical Kolmogorov backward formula (see [59] theorem 8.1.1)

Moreover by applying equation (5.5) to $u(t,x) := T_t^v f(x) = \mathbb{E}\left[e^{-A(t)} f(X_t)|X_0 = x\right] \in C_0$ and using equation (5.3) we obtain

$$\frac{\partial}{\partial t} u(t,x) = (\mathcal{L} - v) u(t,x)$$

that is the classical Feynman-Kac formula described in [59] theorem 8.2.1

5.3 Exponential map of semi-groups

Definition 5.17. *A strongly-continuous semi-group T_t is said uniformly continuous if*

$$\lim_{t \longrightarrow 0} T_t = I$$

where I is the identity operator

Example 5.18. *The operator* $e^{tA} := \sum_{n=1}^{\infty} \frac{t^n}{n!} A^n$ *is a uniformly continuous operator for any bounded operator A.*

Example 5.19. *Consider the Hilbert space l^2 and let $\{e_n\}$ its standard orthonormal bases. Define*

$$T_t f = \sum_{n=1}^{\infty} e^{-nt} \langle e_n, f \rangle e_n$$

Because $||T_t e_n - e_n|| = (1 - e^{-nt})$ it is a strongly continuous semi-group, but for the same reason it is impossible to make the norm arbitrarily close to 1 for all e_n so it is not uniformly continuous

The uniformly continuous semi-groups admits a nice representation in terms of their infinitesimal generator.

Theorem 5.20 ([61] theorem 1.2)**.** *Let $(\mathcal{L}, \mathcal{D})$ a linear operator. It is the infinitesimal generator of an uniformly continuous semi-group if and only if it is bounded.*

The semi-group can be represented as

$$T_t f(x) := e^{t\mathcal{L}} f(x) := \sum_{n=1}^{\infty} \frac{t^n}{n!} \mathcal{L}^n f(x) \tag{5.8}$$

We want to generalize equation (5.8) also in the case of unbounded generators, as the Laplacian (generator of the Brownian in $\mathbb{R}^n$) and more generally the Laplace-Beltrami operator associated to a Riemannian metric (see definition 3.22) are unbounded operator in C_0. In such case the series doesn't need to converges, but the set of functions for which this happens is dense in the domain of the operator.

Definition 5.21. *Let $\mathcal{L}, \mathcal{D}(\mathcal{L})$ a linear operator. Name $\mathcal{D}(\mathcal{L}^\infty) := \bigcap_n \mathcal{D}(\mathcal{L}^n)$.*

$f \in \mathcal{D}^\infty$ is called an analytic vector on a set $U \in \mathbb{R}$ if equation (5.8) converges for any $t \in U$

Example 5.22. *Let $T_t : C_0(\mathbb{R}) \to C_0(\mathbb{R})$ the time translation semi-group of example 5.11. We have that $\mathcal{D}(\mathcal{L}^\infty) = C^\infty(\mathbb{R})$ and*

$$e^{t\mathcal{L}} f(x) = \sum_{n=0}^{\infty} \frac{t^n}{n!} \frac{d^n}{dx^n} f(x)$$

i.e. $e^{t\mathcal{L}} f(x)$ is the Taylor series of $f(x)$ around 0. The set of analytical vectors are the functions that are analytic in 0. This set is properly contained in C^∞ and so in C^0. Indeed $\frac{d}{dx}$ is not a bounded operator in C_0

If the vector is analytic on $\mathbb{R}$ is called entire. If T_t is a strongly continuous group on a Banach space X the set of entire vectors is dense in X (see [26] exercise II 3.12(2)). A similar result holds in for Feller semi-groups.

We start by giving the following lemmas

Lemma 5.23 (**Yoshida approximation** ([37] lemma 19,7))**.** *Let T_t a Feller semi-group, $(\mathcal{L}, \mathcal{D})$ its infinitesimal generator and R_λ its resolvent. Let $\mathcal{L}^\lambda := \lambda \mathcal{L} R_\lambda$. It is a bounded operator and its associated semi-group is $T_t^\lambda := e^{t\mathcal{L}^\lambda}$. We have*

$$|T_t f - T_t^\lambda f| \leq t|\mathcal{L} f - \mathcal{L}^\lambda f|$$

Moreover, $\mathcal{L}^\lambda \xrightarrow{\lambda \to \infty} \mathcal{L}$ and $T_t^\lambda f \xrightarrow{\lambda \to \infty} T_t f$ for any $f \in C_0$ uniformly for bounded $t \geq 0$

Lemma 5.24 ([61] theorem 2.7)**.** *Let $(\mathcal{L}, \mathcal{D})$ the generator of a strongly continuous semi-group on a Banach space X and let $\mathcal{D}(\mathcal{L}^\infty)$ as in definition 5.21. $\mathcal{D}(\mathcal{L}^\infty)$ is dense in X*

Lemma 5.25. *Let $(\mathcal{L}, \mathcal{D})$ a closed and densely defined linear operator on a Banach space X. $g : [0, \infty) \to X$ a continuous function. If $g([0, \infty)) \in \mathcal{D}$ and $t \to \mathcal{L} g(t)$ is C_0*

$$\mathcal{L} \int_0^\infty g(t) dt = \int_0^\infty \mathcal{L} g(t) dt$$

Proof. The adjoint $\mathcal{L}^*$ of a closed, densely defines operator is densely defined so for any $x^* \in \mathcal{D}(\mathcal{L}^*)$

$$\left\langle \mathcal{L} \int_0^\infty g(t) dt, x^* \right\rangle = \int_0^\infty \langle g(t), \mathcal{L}^* x^* \rangle \, dt = \int_0^\infty \langle \mathcal{L} g(t), x^* \rangle \, dt = \left\langle \int_0^\infty \mathcal{L} g(t) dt, x^* \right\rangle$$

the result follow by the density of $\mathcal{D}(\mathcal{L}^*)$ □

Lemma 5.26. *The function $g_n : \mathbb{R} \to \mathbb{R}$ defined as*

$$g_n(x) := \begin{cases} \frac{\sqrt{n}}{2\sqrt{2\pi}x^2} e^{-\frac{n}{x^2}} & \text{if } x \neq 0 \\ 0 & \text{otherwise} \end{cases}$$

is a $C^\infty(\mathbb{R})$ function such that

$$\lim_{x \to 0} \frac{d^k}{dx^k} g_n(x) = 0 = \lim_{x \to \infty} \frac{d^k}{dx^k} g_n(x)$$

for any $k \geq 0$

Proof. Continuity in zero follow by $\lim_{x \to 0} x^2 e^{\frac{n}{x^2}} = \infty$ and the fact that $g_n(x)$ is positive.

In [46] lemma 2.20 it is proven that

$$f(t) = \begin{cases} e^{-\frac{1}{t}} & \text{if } x \neq 0 \\ 0 & \text{otherwise} \end{cases}$$

is C^∞ and all the derivatives vanishes, the same proof applies to $g_n(x)$ □

Theorem 5.27. *Let T_t a Feller semi-group on a Banach space X and $(\mathcal{L}, \mathcal{D})$ its infinitesimal generator. The set of analytic vectors on $[0, \infty)$ is dense in X and if f is an analytic vector on $U \subseteq [0, \infty)$ equation (5.8) holds for $T_t f$ for $t \in U$*

Proof. Suppose f is an analytic vector on U and let $t \in U$. By the Yoshida approximation (lemma 5.23) we know that

$$T_t f = \lim_{\lambda \to \infty} e^{t\mathcal{L}^\lambda} f$$

because $\mathcal{L}^\lambda f \to \mathcal{L} f$ and $e^{t\mathcal{L}} f$ is a well-defined function, the dominated convergence theorem implies that equation (5.8) holds.

Let now $f \in \mathcal{D}(\mathcal{L}^\infty)$ and g_n as in lemma 5.26 we define

$$S_n f := \int_0^\infty g_n(s) T_s f ds$$

As the semi-group is a contraction we have

$$||S_n f|| \leq \int_0^\infty g_n(s) ||f|| ds = ||f||$$

so the integral is well defined. By lemma 5.25, the Kolmogorov equation (5.2) and the integration by parts formula we obtain, for any $k > 0$

$$A^k S_n f = \int_0^\infty g_n(s) \frac{d}{ds} T_s A^{k-1} f ds = (-1)^k \int_0^\infty \frac{d^k}{ds^k} g_n(s) T_s f ds$$

This together with lemma 5.26 imply that the series (5.8) converges for $S_n f$ and any $t \in [0, \infty)$, i.e. $S_n f$ is an analytic vector on $[0, \infty)$ for any n.

By applying a change of variable and using the strong continuity and the contraction property of the semi-group

$$\lim_{n \to \infty} S_n f = \lim_{n \to \infty} \frac{1}{2\sqrt{2\pi}} \int_0^\infty e^{-u^2} T_{\frac{u}{\sqrt{n}}} f du = f$$

The boundedness of S_n and the Banach-Steinhaus theorem assure that the convergence is uniform. So the set of analytic vectors on $[0, \infty)$ is dense in $\mathcal{D}(\mathcal{L}^\infty)$ and $\mathcal{D}(\mathcal{L}^\infty)$ is dense in X □

A sufficient condition for a function to be an analytic vector on an interval $[0, \rho)$ is given by the following theorem, that is a consequence of the Cauchy-Hadamard theorem.

Proposition 5.28. *Let $(\mathcal{L}, \mathcal{D})$ a closed, densely defined operator and suppose that $f \in \mathcal{D}(\mathcal{L}^\infty)$ is such that*

$$\limsup_{n \to \infty} \left(\frac{1}{n!} ||\mathcal{L}^n f|| \right)^{\frac{1}{n}} =: \rho < \infty$$

then, f is an analytic vector for $t \in \left[0, \frac{1}{\rho}\right)$.

If $\rho = 0$, f is an analytic vector on $[0, \infty)$

We will now state a generalization of the Talay-Tubaro expansion (I.2). This will be the majot tool to specify the order conditions of a weak stochastic Runge-Kutta method on Lie groups.

Theorem 5.29. *Let T_t a Feller semi-group with generator $(\mathcal{L}, \mathcal{D})$. Suppose $f \in \mathcal{D}(\mathcal{L}^{n+1})$ and $||\mathcal{L}^{n+1} f|| \leq M < \infty$. Define the n-th remainder as*

$$R_{\mathcal{L}^n} f = T_t f - \sum_{i=0}^{n} \frac{t^i}{i!} \mathcal{L}^i f$$

We have

$$||R_{\mathcal{L}^n} f|| \leq \frac{M}{(n+1)!} t^{n+1}$$

Proof. By the contraction property of T_t

$$||\mathcal{L}^{n+1} T_t f|| = ||T_t \mathcal{L}^{n+1} f|| \leq ||\mathcal{L}^{n+1} f|| < M$$

By equation (5.2), the integration by part formula and $T_0 f = f$

$$\int_0^t \frac{(t-\tau)^n}{n!} T_\tau \mathcal{L}^{n+1} f d\tau = \int_0^t \frac{(t-\tau)^n}{n!} \frac{d^{n+1}}{d\tau^{n+1}} T_\tau f d\tau = \cdots = T_t f - \sum_{i=0}^{n} \frac{t^i}{i!} \mathcal{L}^i f = R_{\mathcal{L}^n} f$$

and

$$\left|\left| \int_0^t (t-\tau)^n T_\tau \mathcal{L}^{n+1} f d\tau \right|\right| \leq M \int_0^t \frac{|t-\tau|^n}{n!} = \frac{M}{(n+1)!} t^{n+1}$$

□

6 Stochastic Lie groups methods

We will now descibe stochastic generalization of the two methods that we have seen so far: the Magnus expansion and Runge-Kutta methods. In the first section, folowing [38] we will present a stochastic Magnus expansion for an Itô SDE on $GL(n, \mathbb{R})$. We will verify how, in the simplest case in which the vector fields are constant the formula will agree with the estimate given by theorem 5.29.

In the second section we will give a definition for a semi-martingale on a differentiable manifold. Such definition will rely on the Whitney embedding theorem 1.11. Using the manifold semi-martingales it is possible to give a definition for the solution of an SDE on a manifold. After that, by following [25] we will give a characterization of martingale on a manifold in terms of the connection.

In the third section we will further characterize SDEs on a manifold in terms of their development on the frame bundle and the corresponding anti-development in $\mathbb{R}^n$ (see section 3.4)

In the fourth section we will define the main subject of the chapter: diffusion on manifold. Any Feller process on a manifold will be shown to be a diffusion process. The converse is in general not true.

Following [49] we will show how in a manifold of bounded geometry (see section 3.5) it is sufficient to impose some condition on the coefficients of the diffusion operator to assure that the process associated to it is Feller.

We will use this result and theorem 5.29 to find a weak second order Lie Runge-Kutta method in a general matrix Lie group endowed with a left invariant metric.

As usual, through the whole chapter we will use the Einstein summation convention for repeated indexes.

6.1 The stochastic Magnus expansion for Itô integrals in the general linear group

The Magnus expansion described in section 2.4 can be generalized to the case of SDE in the Stratonovich (as outlined in [74]) or Itô formalism.

Here we will present the latter in the particular case of a process with a single Brownian motion. For the general case and the analysis of the convergence of such expansion can be found in [38].

We start in the case in which $GL(n, \mathbb{R})$. As the space is diffeomorphic to $\mathbb{R}^{n^2}$ it is possible to endowed it with the Euclidean metric. So the Itô formula of equation (4.2) holds.

The major theorem that we are going to reference is theorem 1 of [38]:

Theorem 6.1. *Let, ${G^i}_t$, $i = 0, \cdots q$ bounded, progressively measurable matrix valued processes on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t\geq 0})$ equipped with a standard q-dimensional Brownian motion $W = (W^1, \cdots, W^q)$.*

Suppose that the Itô SDE

$$\begin{cases} dX_t = {G^0}_t X_t dt + {G^j}_t X_t d{W_t}^j \\ X_0 = I_n \end{cases}$$

has a unique strong solution in $[0, T]$, then there exists a positive stopping time $\tau \leq T$ such that

1. *X_t has a real logarithm (inverse of the exponential matrix) $Y_t \in \mathfrak{gl}(n)$ up to time τ i.e $X_t = \operatorname{expm}(Y_t)$*
2. *such logarithm can be represented as an infinite series $Y_t = \sum_{n=0}^{\infty} Y_t^{(n)}$ which converges $\mathbb{P}$-almost surely up the time τ.*

By Lemma 1 of [38] we know that the second order differential of the exponential map at $v \in \mathfrak{gl}(n)$ is given by the map

$$(M, N) \to \mathfrak{D}_v(M, N) \operatorname{expm}(v)$$

where

$$\mathfrak{D}_v(M, N) = d\exp_v(M) d\exp_v(N) + \int_0^1 \tau[d\exp_v(N), \exp(\mathrm{ad}_{\tau v}(M)] d\tau =$$
$$= \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \frac{\mathrm{ad}_v^n(M)\, \mathrm{ad}_v^m(N)}{(n+1)!(m+1)!} + \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \frac{[\mathrm{ad}_v^n(N), \mathrm{ad}_v^m(M)]}{(n+m+2)(n+1)!m!}$$

Suppose now Y is a matrix-valued Itô process of the form

$$dY_t = \mu_t dt + \sigma_t^j dW_t^j$$

By using the Itô formula we obtain that

$$d\operatorname{expm}(Y_t) = \left(d\exp_{Y_t}(\mu_t) + \frac{1}{2} \sum_{j=1}^{q} \mathfrak{D}_{Y_t}(\sigma_t^i, \sigma_t^j) \right) \operatorname{expm}(Y_t) dt + d\exp_{Y_t}(\sigma_t^j) \operatorname{expm}(Y_t) dW_t^j$$

Consider now the process $X_t^{\epsilon,\delta}$ which solves the Itô SDE

$$\begin{cases} dX_t^{\epsilon,\delta} = \delta G_t^0 X_t^{\epsilon,\delta} dt + \epsilon G_t^j X_t^{\epsilon,\delta} dW_t^j \\ X_0^{\epsilon,\delta} = I_n \end{cases} \tag{6.1}$$

It is possible to show [38] that $X_t^{\epsilon,\delta}$ in (6.1) admits the exponential representation $X_t^{\epsilon,\delta} = \operatorname{expm}(Y_t^{\epsilon,\delta})$ and that $Y_t^{\epsilon,\delta}$ solves the Itô SDE

$$dY_t^{\epsilon,\delta} = \mu^{\epsilon,\delta}(t, Y_t^{\epsilon,\delta}) dt + \sigma_j^{\epsilon}(t, Y_t^{\epsilon,\delta}) dW_t^j \qquad Y_0^{\epsilon,\delta} = 0$$

With

$$\sigma_j^{\epsilon}(t, .) = \epsilon \sum_{n=0}^{\infty} \frac{B_n}{n!} \mathrm{ad}_.^n(G_t^j)$$

$$\mu^{\epsilon,\delta}(t, .) = \sum_{n=0}^{\infty} \frac{B_n}{n!} \mathrm{ad}_.^n \left(\delta G_t^0 - \frac{1}{2} \sum_{j=1}^{q} \mathfrak{D}_.(\sigma_j^{\epsilon}(t, .), \sigma_j^{\epsilon}(t, .) \right)$$

(B_n is the n-th Bernoulli number).

Moreover it is possible to express $Y_t^{\epsilon,\delta}$ as an infinite series

$$Y_t^{\epsilon,\delta} = \sum_{n=0}^{\infty} \sum_{r=0}^{n} Y_s^{(r,n-r)} \epsilon^r \delta^{n-r}$$

.

where we have weighted δ as ϵ^2 to respect the "probabilistic relation" $\sqrt{t} \approx W_t$, i.e

$$Y_s^{(r,n-r)} = o(h^{\frac{r}{2}(n-r)})$$

The general recursive formula for $Y_s^{(r,n-r)}$ can be found in [38]. Here we present the terms up to order 2 in the particular case in which $q = 1$. Let $0 < s < t < T$.

Define the anti-commutator $\{A,\ B\} = AB + BA$

$$\begin{aligned}
Y_t^{(0,0)} &= 0 \\
Y_t^{(1,0)} &= \int_0^t G_s^1 dW_s \\
Y_t^{(0,1)} &= \int_0^t G_s^0 ds \\
Y_t^{(2,0)} &= -\frac{1}{2}\int_0^t \left[Y_s^{(1,0)}, G_s^1\right] dW_s - \frac{1}{2}\int_0^t (G_s^1)^2 ds \\
Y_t^{(3,0)} &= \int_0^t -\frac{1}{2}\left[Y_s^{(2,0)}, G_s^1\right] + \frac{1}{12}\left[Y_s^{(1,0)}, \left[Y_s^{(1,0)}, G_s^1\right]\right] dW_s + \\
&+ \frac{1}{4}\int_0^t \frac{1}{3}\left[\left[Y_s^{(1,0)}, G_s^1\right], G_s^1\right] - \left[Y_s^{(1,0)}, (G_s^1)^2\right] ds \\
Y_t^{(1,1)} &= -\frac{1}{2}\int_0^t \left[Y_s^{(0,1)}, G_s^1\right] dW_s - \frac{1}{2}\int_0^t \left[Y_s^{(1,0)}, G_s^0\right] ds \\
Y_t^{(0,2)} &= -\frac{1}{2}\int_0^t \left[Y_s^{(0,1)}, G_s^o\right] ds \\
Y_t^{(2,1)} &= \int_0^t -\frac{1}{2}\left[Y_s^{(1,1)}, G_s^1\right] + \frac{1}{12}\left(\left[Y_s^{(1,0)}\left[Y_s^{(0,1)}, G_s^1\right]\right] + \left[Y_s^{(0,1)}\left[Y_s^{(1,0)}, G_s^1\right]\right]\right) dW_s + \\
&+ \frac{1}{4}\int_0^t \frac{1}{3}\left[\left[Y_s^{(0,1)}, G_s^1\right], G_s^1\right] + \left[Y_s^{(0,1)}, (G_s^1)^2\right] + \left[Y_s^{(2,0)}, G_s^0\right] ds \\
Y_t^{(4,0)} &= \int_0^t -\frac{1}{2}\left[Y_s^{(3,0)}, G_s^1\right] + \frac{1}{12}\left(\left[Y_s^{(2,0)}\left[Y_s^{(1,0)}, G_s^1\right]\right] + \left[Y_s^{(1,0)}\left[Y_s^{(2,0)}, G_s^1\right]\right]\right) dW_s + \\
&+ \frac{1}{2}\int_0^t -\frac{3}{4}\left(\left[Y_s^{(1,0)}, G_s^1\right]\right)^2 + \frac{1}{3}\left\{\left[Y_s^{(1,0)}, \left[Y_s^{(1,0)}, G_s^1\right]\right],\ G_s^1\right\} + \frac{1}{6}\left[\left[Y_s^{(2,0)}, G_s^1\right], G_s^1\right] ds + \\
&+ \frac{1}{4}\int_0^t \left[Y_s^{(2,0)}, (G_s^1)^2\right] - \frac{1}{3}\left[Y_s^{(1,0)}, \left[Y_s^{(1,0)}, (G_s^1)^2\right]\right] ds
\end{aligned}$$

We are interested in calculating the quantities $\mathbb{E}\left[\prod_{i,j} Y_t^{(m_i,n_j)}\right]$ with $\frac{1}{2}\sum_i m_i + \sum_j n_j \leq 2$. We start with a generalization of the Itô isometry

Lemma 6.2. *Let W_t a Brownian motion, $\{G_s^j\}_{i=0,\cdots p}$ bounded, progressively measurable, L^2, matrix-valued stochastic processes on the filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, \mathcal{F}_t)$, where $\mathcal{F}_t$ is the natural filtration of the Brownian so*

$$\mathbb{E}\left[\left(\int_0^t G_s^0 dW_s\right)\prod_{j=1}^{p-1} G_t^j \left(\int_0^t G_s^p dW_s\right)\right] = \mathbb{E}\left[\int_0^t G_s^0 \left(\prod_{j=1}^{p-1} G_t^j\right) G_s^p ds\right] \tag{6.2}$$

Proof. We will prove the result for elementary processes on $\mathfrak{gl}(n,\mathbb{R})$. The result for L^2 functions will follow by a density argument as in chapter 3.1 of [59]. So Let $G_t^i(\omega) := \sum_j e_j^i(\omega)\chi_{[t_j,t_{j+1}]}$, for $i = 0$ and $i = p$. The Itô integral of such processes is:

$$\int_0^t G_s^i(\omega) dW_s(\omega) := \sum_j e_j^i(\omega)\left(W_{t_{j+1}} - W_{t_j}\right)$$

$$\mathbb{E}\left[\left(\int_0^t G_s^0 dW_s\right)\prod_{j=1}^{p-1} G_s^j \left(\int_0^t G_s^p dW_s\right)\right] = \sum_{ij}\mathbb{E}\left[e_i^0\left(\prod_{k=1}^{p-1} G_s^k\right) e_j^p \Delta W_i \Delta W_j\right] =$$

$$= \sum_j \mathbb{E}\left[e_j^0\left(\prod_{k=1}^{p-1} G_s^k\right) e_j^p\right](t_{j+1} - t_j) = \int_0^t \mathbb{E}\left[G_s^0\left(\prod_{j=1}^{p-1} G_s^j\right) G_s^p\right] ds$$

Using Fubini theorem we obtain the equality (6.2) $\square$

To calculate the mean value of the product of more than two stochastic processes we will need the following result.

Lemma 6.3 (Isserlis' theorem [36]). *If $(X_1, \cdots, X_n)$ is a zero mean multivariate normal random variable then*

$$\mathbb{E}[\prod_{i=1}^{n} X_i] = \sum_{p\in P_n^2} \prod_{(i,j)\in p} \mathbb{E}[X_i X_j]$$

where P_n^2 is the set of all the possible partitions in pairs of $\{1 \cdots n\}$

By using lemma 6.2 , The Isserlis' theorem 6.3, Fubini theorem and theorem 4.14 we obtain

$$\mathbb{E}[Y_t^{(0,1)}] = Y_t^{(0,1)}$$
$$\mathbb{E}[Y_t^{(0,2)}] = Y_t^{(0,2)}$$
$$\mathbb{E}[Y_t^{(2,0)}] = -\frac{1}{2}\int_0^t (G_r^1)^2 dr$$
$$\mathbb{E}[Y_t^{(2,1)}] = \frac{1}{4}\int_0^t \frac{1}{3}\left[\left[\int_0^s G_r^0 dr, G_s^1\right], G_s^1\right] - \frac{1}{2}\left[(G_S^1)^2, G_s^0\right] ds$$
$$\mathbb{E}[Y_t^{(4,0)}] = \frac{1}{2}\int_0^t -\frac{3}{4}\left(\left[\int_0^s G_r^1 dr, G_s^1\right]\right)^2 + \frac{1}{3}\left\{\left[\int_0^s G_r^1 dr, \left[\int_0^s G_r^1 dr, G_s^1\right]\right],\ G_s^1\right\} ds +$$
$$-\frac{1}{4}\int_0^t \frac{1}{3}\left[\left[\int_0^s (G_r^1)^2 dr, G_s^1\right], G_s^1\right] + \frac{1}{2}\left[\int_0^s (G_r^1)^2 dr, (G_s^1)^2\right] + \frac{1}{3}\left[\int_0^s G_r^1 dr, \left[\int_0^s G_r^1 dr, (G_s^1)^2\right]\right] ds$$

$$\mathbb{E}[Y_t^{(1,0)} \cdot Y_t^{(1,0)}] = \int_0^t (G_s^1)^2 dr$$
$$\mathbb{E}[Y_t^{(0,1)} \cdot Y_t^{(0,1)}] = (\int_0^t (G_s^0) dr)^2$$
$$\mathbb{E}[Y_t^{(2,0)} \cdot Y_t^{(2,0)}] = \frac{1}{4}\left(\left(\int_0^s (G_r^1)^2 dr\right)^2 + \int_0^t \left[\int_0^s G_r^1 dr, G_s^1\right]^2 ds\right)$$
$$\mathbb{E}[Y_t^{(0,1)} \cdot Y_t^{(2,0)}] = -\frac{1}{2}\int_0^t G_s^0 dr \int_0^t (G_s^1)^2 dr$$
$$\mathbb{E}[Y_t^{(1,0)} \cdot Y_t^{(1,1)}] = -\frac{1}{2}\cdot\int_0^t G_s^1 \left[\int_0^s G_r^0 dr, G_s^1\right] ds$$
$$\mathbb{E}[Y_t^{(1,0)} \cdot Y_t^{(3,0)}] = -\frac{1}{2}\cdot\int_0^t G_s^1 \left(\left[-\frac{1}{2}\int_0^s (G_r^1)^2 dr, G_s^1\right] + \frac{1}{12}\left[\int_0^s G_r^1 dr, \left[\int_0^s G_r^1 dr, G_s^1\right]\right]\right) ds$$
$$\mathbb{E}[(Y_t^{(1,0)})^2 Y_t^{(0,1)}] = \int_0^t (G_s^1)^2 dr \int_0^t G_s^0 dr$$
$$\mathbb{E}[Y_t^{(1,0)} Y_t^{(0,1)} Y_t^{(1,0)}] = \int_0^t G_s^1 \left(\int_0^t G_s^0 ds\right) G_s^1 ds$$
$$\mathbb{E}[(Y_t^{(1,0)})^2 Y_t^{(2,0)}] = -\frac{1}{2}\left(\int_0^t (G_s^1)^2 dr\right)^2$$
$$\mathbb{E}[(Y_t^{(1,0)})^4] = 3\left(\int_0^t G_s^1 ds\right)^4$$

while all the other means are either 0 or can be obtained by permutating the factors of the multiplication of the integrals above

Using this formulas it is possible to define a numerical scheme of weak order 2 for the case of constant coefficients G_0 and G_1

Example 6.4. *Let X_t the unique solution of the SDE*

$$dX_t = G^0 X_t dt + G^1 X_t dW_t$$
$$X_0 = I_n$$

where G^0 and G^1 are constant.

$$
\begin{aligned}
Y_t^{\left(\frac{1}{2}\right)} &= Y_t^{(1,0)} \\
Y_t^{(1)} &= tG^0 + Y_t^{(2,0)} \\
Y_t^{\left(\frac{3}{2}\right)} &= \sum_{2\epsilon+\delta=\frac{3}{2}} Y_t^{(\epsilon,\delta)} \\
Y_t^{(2)} &= \sum_{2\epsilon+\delta=2} Y_t^{(\epsilon,\delta)}
\end{aligned}
$$

Let ξ a standard Gaussian random variable, $h \in [0,T]$ small enough and define the random variables

$$
\begin{aligned}
\hat{Y}_h^{\left(\frac{1}{2}\right)} &= \sqrt{h}G^1\xi \\
\hat{Y}_h^{(1)} &= h\left(G^0 - \frac{1}{2}(G^1)^2\right) \\
\hat{Y}_h^{\left(\frac{3}{2}\right)} &= -\frac{1}{4}h^{\frac{3}{2}}\left[G^0, G^1\right]\xi \\
\hat{Y}_h^{(2)} &= h^2\left(\frac{1}{12}\left[\left[G^0, G^1\right], G^1\right] - \frac{1}{8}\left[(G^1)^2, G^0\right]\right)
\end{aligned}
$$

By direct verification we obtain that given any set of indexed $\alpha_i \in \left\{\frac{1}{2}, 1, \frac{3}{2}, 2\right\}$ such that $\sum_i \alpha_i \leq 2$

$$\mathbb{E}\left[\prod_i Y_h^{(\alpha_i)} - \prod_i \hat{Y}_h^{(\alpha_i)}\right] = O(h^3)$$

or, more explicitly

$$\mathbb{E}\left[X_t | X_0 = I_n\right] = I_n + hG^0 + h^2\frac{1}{2}(G^0)^2 + o(h^2) \tag{6.3}$$

This is consistent with the result of theorem 5.29, Indeed with the indetification $M_n(\mathbb{R}) \cong \mathbb{R}^{n^2}$ the coefficients are globally Lipschitz, so by theorem 5.7 the process is Feller. Its infinitesimal generator can be written in a base as

$$\mathcal{L}f(x) = \left((G_0x)^i\frac{\partial}{\partial x^i} + \frac{1}{2}(G_1x)^i(G_1x)^j\frac{\partial^2}{\partial x^i\partial x^j}\right)f(x)$$

Let now $f : M_n(\mathbb{R}) \to \mathbb{R}$ be a coordinate function. Let K a compact set of $M_n(\mathbb{R})$ and U a neighbourhood of K. Let f_K a smooth function whose support is contained in U and such that it agrees with f on K. Let τ the exit time from K. By theorem 5.29, for any $t < \tau$

$$\mathbb{E}[f_K(X_t)|X_0 = I_n] = I_n + \mathcal{L}f_K(I_n) + \mathcal{L}^2 f_K(I_n) + O(h^3)$$

Because the second derivative of a coordinate function is zero this formula agrees with equation (6.3)

6.2 Semi-martingales and SDEs on Riemannian manifolds

Definition 6.5. *Let M a differentiable manifold and $(\Omega, \mathcal{F}, \mathbb{P}, \mathcal{F}_t)$ a filtered probability space and let τ a stopping time.*

A continuous process X defined on $[0,\tau)$ is an M-valued semi-martingale if $f(X)$ is a real semi-martingale on $[0,\tau)$ for any $f \in C^\infty(M)$ (see definition 4.4)

Let $\hat{M} = M \cup \{\partial\}$ the one point compactification of a non-compact manifold (if M is compact $\hat{M} = M$) and define the path space

$$\hat{W}(M) := \left\{w : [0,\infty) \to \hat{M} : \text{ w is continuous, } w(0) \in M \text{ and if } w(t) = \partial \text{ then } \forall\ t' > tw(t') = \partial\right\}$$

and let $\mathcal{B}(\hat{W}(M))$ the σ-algebra generated by the Borel cylinders sets and define the explosion time as

$$e(w) = \inf\{t : w(t) = \partial\} \tag{6.4}$$

Definition 6.6. *Let $\{G_i\}_{i=0}^k$ a family of vector fields over M and let X_t a $\mathcal{F}_t$-adapted $\hat{W}(M)$-valued random variable over a filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, \mathcal{F}_t)$ and define $\mathcal{F}_t$-Brownian motions W_t^i.*

X_t is the solution of the SDE

$$dX_t = G_0(X_t)dt + G_i(X_t) \circ dW_t^i \tag{6.5}$$

if, for any $f \in C_c^\infty(M)$ (the set of compactly supported smooth functions over M) $f(X_t)$ if

$$f(X_t) - f(X_0) = \int_0^t G_i f(X_s) \circ dW_s^i + \int_0^t G_0 f(X_s) ds \tag{6.6}$$

where the Stochastic integral is intended in the Stratonovich sense.

By the Whitney embedding theorem 1.11 any manifold can be thought as an embedded submanifold of $\mathbb{R}^N$ for N big enough. We define the coordinates function $f : M \to \mathbb{R}^N$ such that $f^i(x) = x^i$, where x^i is the i-th coordinate of x in $\mathbb{R}^N$. We have the following theorem

Theorem 6.7 ([32] proposition 1.2.7)**.** *Suppose M is a closed submanifold of $\mathbb{R}^N$ and let $\{f^i\}_{i=1\cdots N}$ be the coordinate functions and let X_t an M-valued continuous process*

1. *X_t is a semi-martingale on M if and only if $f^i(X_t)$ is a real semi-martingale for each $i = 1 \cdots N$*

2. *X_t is a solution of the SDE of equation (6.5) up to a stopping time σ if and only if equation (6.6) holds for any f^i, $i = 1 \cdots N$ and for each $0 \le t < \sigma$*

Given any SDE on a manifold as in equation (6.5) it is possible to extend the vector fields (considering that as smooth function from M to $\mathbb{R}^N$ to a vector field on $\mathbb{R}^N$. Because vector fields are tangent to M if $X_0 \in M$ the SDE stays on M.

Moreover, If we define a solution in $\mathbb{R}^N$ up to its explosion time (defined as in equation (6.4)) it is possible to weaken the sufficient condition for existence and uniqueness of theorem 5.8. In particular, we don't need the coefficients to have linear growth anymore (see [32] proposition 1.1.9). So, because the Locally Lipschitz condition is already satisfies we have an uniqueness result for an SDE on a manifold. In summary

Theorem 6.8. *Let $\tilde{G}_i$ the extension of the vector fields in equation (6.5) and suppose X_t is the solution of the extended equation*

$$dX_t = \tilde{G}_0(X_t)dt + \tilde{G}_i(X_t) \circ dW_t^i$$

up to the explosion time $e(X)$ and suppose $X_0 \in M$, then $X_t \in M$ for any $0 \le t < e(X)$ ([32] proposition 1.2.8).

Moreover the solution of the SDE (6.5) exists up to the explosion time and is unique ([32] theorem 1.2.9)

Remark. *This formalism for SDE on manifolds relies on a non-canonical embedding in an Euclidean space. It is possible to define the stochastic integral and the solution of SDEs on a Riemannian manifold endowed with a traceless connection intrinsically. This requires to define the space of second order tangent vectors [25, chapter 6].*

If the manifold is a Lie group G with Lie algebra $\mathfrak{g}$ this space is a submodule of the universal enveloping algebra of $\mathfrak{g}$.

We continue with the definition of a M-valued martingale

Definition 6.9. *Let (M, g, ∇) a Riemannian manifold endowed with a torsion-free connection and X an M-valued semi-martingale.*

X is a martingale w.r.to the connection ∇ if, for any $f \in C^\infty(M)$ we have

$$f(X_t) = f(X_0) + M_t + \int_0^t \nabla^2 f(dX, dX) \tag{6.7}$$

where M_t is a real local martingale. It is possible to characterize a martingale in local coordinates (see [25] 4.20). By partitioning the interval $[0, e(X))$ in an increasing sequence of stopping times $\{\tau_n\}$ so that the process stays in a local chart for any interval $[\tau_n, \tau_{n+1}]$.

A semi-martingale X_t is a martingale if and only if there exists a real local martingale such that in $[\tau_n, \tau_{n+1}]$

$$X_t^i = X_0^i + M_t^i - \frac{1}{2}\int_0^t \Gamma_{jk}^i(X_s) d\left\langle M^j, M^k \right\rangle_s$$

where $d\left\langle M^j, M^k \right\rangle_s$ is the quadratic variation process and Γ_{jk}^i are the Christoffel symbols

6.3 Stochastic developments

In section 3.4 we have defined the development of a curve in M and the anti-development of a curve in $\mathfrak{F}(M)(M)$. There is a stochastic analogous to that.

In all the section (M, g, ∇) is a Riemannian manifold endowed with the Levi-Civita connection and all the processes are defined on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, \mathcal{F}_t)$ and are $\mathcal{F}_t$-adapted. We denote $\pi : \mathfrak{F}(M) \to M$ the projection from the frame bundle to the manifold.

Definition 6.10. *Consider the following SDE on $\mathfrak{F}(M)$:*

$$dU_t = \sum_i H_i(U_t) \circ dW_t^i \tag{6.8}$$

where W is an $\mathbb{R}^d$ semi-martingale and U_i are the fundamental horizontal vector fields on $\mathfrak{F}(M)$ (see equation (3.7)).

An $\mathfrak{F}(M)$ valued semi-martingale U is said to be horizontal if there exists an $\mathbb{R}^d$-valued semi-martingale W such that equation (6.8) holds. W is called the anti-development of U (or of $X = \pi U$).

A stochastic development in $\mathfrak{F}(M)$ of an $\mathbb{R}^d$-valued semi-martingale W at the starting frame U_0 is a solution of equation (6.8). Its projection is a stochastic development of W in M.

U $\mathfrak{F}(M)$-valued semi-martingale is a stochastic horizontal lift of an M-valued semi-martingale X if $\pi U = X$

There is a 1-1 correspondence between stochastic processes on a manifold, their stochastic development and the corresponding anti-development as shown in the next theorems

Theorem 6.11 ([32] theorem 2.3.5 and lemma 2.3.7)**.** *Let X an M-valued semimartingale up to a stopping time τ and U_0 a $\mathfrak{F}(M)$-valued $\mathcal{F}_0$-random variable such that $\pi U_0 = X_0$. Then, there is an unique horizontal lift $\{U_t t \in [0, \tau)\}$ of X starting at U_0.*

The horizontal lift of X is defined up to τ i.e. there is no explosion in the vertical direction.

Lemma 6.12 ([32] lemma 2.3.3)**.** *Suppose M is a closed submanifold of R^N and let $P(x) : \mathbb{R}^N \to T_x M$ the orthogonal projection from R^N to $T_x M$ (thought as a subspace of $\mathbb{R}^N$). Let X an M-valued stochastic process. We have that*

$$dX_t = P_\alpha(X_t) \circ dX_t^\alpha$$

Theorem 6.13 ([32] theorem 2.3.4)**.** *An horizontal semi-martingale U on $\mathfrak{F}(M)$ has a unique anti-development W_t, namely*

$$W_t = \int_0^t U_s^{-1} P_\alpha(X_s) \circ dX_s^\alpha$$

Where $P(x) : R^N \to T_x M$ is the orthogonal projection defined in lemma 6.12

We have the following result that allows us to find an SDE for the horizontal lift of a process

Theorem 6.14 ([32] proposition 2.3.8)**.** *Let X be a semi-martingale on M and suppose it is the solution of the SDE*

$$dX_t = V_i \circ dZ_t^i$$

Let V_i^* the horizontal lift of V_i and U_0 such that $\pi U_0 = X_0$, then the horizontal lift U of X, starting at U_0 is the solution of the SDE

$$dU_t = V_i^* dZ_t^i$$

the anti-development of X is given by

$$W_t = \int_0^t U_s^{-1} V_i(X_t) \circ dZ_s^i$$

6.4 Diffusion processes on a manifold

As in the case of $\mathbb{R}^n$ on a manifold is it possible to define a Markov process whose infinitesimal generator satisfies the Dynkin formula of equation (5.13).

For all the chapter $(M, \langle ., . \rangle, \nabla)$ will denote a Riemannian manifold endowed with the Levi Civita connection.

Definition 6.15. *Given a linear differential operator $(P, \mathcal{D}(\mathcal{P}) \subseteq C_0(M))$. Let $(U, \mathbf{x})$ a chart. The total symbol of P is defined as*

$$P(\mathbf{x}, \xi) = e^{-\langle \mathbf{x}, \xi \rangle} P(e^{\langle \mathbf{x}, \xi \rangle})$$

In local coordinates this is equivalent to the map $P := a^\alpha D_\alpha \to a^\alpha \xi_{\alpha_1} \cdots \xi_{alpha_k}$, where $\alpha = (\alpha_1, \cdots, \alpha_k)$ is a multi-index and $D_\alpha := \frac{\partial}{\partial x^{\alpha_1}} \cdots \frac{\partial}{\partial x^{\alpha_k}}$. The principal symbol of P is defined as the homogeneous component of maximum degree.

An operator is called elliptic in $p \in M$ if its principal symbol is positive definite in p.

A second order operator is said uniformly elliptic w.r.t. to the metric g if exists a constant $C > 0$ such that

$$P^{ij}(x)\xi_i\xi_j \geq C g^{ij}(x)\xi_i\xi_j \tag{6.9}$$

Proposition 6.16. *Let (M, g, ∇) a compact Riemannian manifold, so any elliptic operator is uniformly elliptic.*

Proof. In any local chart g^{ij} is bounded and so equation (6.9) holds. □

Given an elliptic operator it is possible to define a diffusion process generated by L

Definition 6.17 ([32] definition 1.3.1)**.** *Let L a differentiable elliptic operator of second order and let $(\Omega, \mathcal{B}(W(M)), \mathbb{P}, (\mathcal{F})_t)$ a filtered probability space.*

A L-diffusion is a M-valued, $\mathcal{F}_t$-adapted stochastic process, defined up its explosion time $X : \Omega \to W(M)$ defined up to its explosion time $e(X)$ such that for any $f \in C^\infty(M)$ $M^f(X)_t$ defined in equation (5.4) for any $t < e(X)$

An L diffusion measure is a probability measure over the path space $(W(M), \mathcal{B}(W(\hat{M}))$ such that, for any $f \in C^\infty(M)$, for any $\omega \in W(M)$ and for any $t < e(\omega)$

$$M^f(\omega)_t := f(\omega_t) - f(\omega_0) - \int_0^t Lf(\omega_s)ds$$

is a $\mathcal{B}(W(\hat{M}))$-martingale .

It is immediate to see that the law of a L-diffusion is an L diffusion measure and that if μ is an L-diffusion measure the coordinate process $X_t(\omega) := \omega_t$ on $(W(M), \mathcal{B}(W(\hat{M})), \mu)$ is an L-diffusion.

There is a result of existence and uniqueness for the L-diffusion measures

Theorem 6.18 ([32] theorem 1.3.4 and 1.3.6)**.** *given a differentiable elliptic operator of the second order $\mathcal{L}$ and a probability distribution μ_0 on M there exists an L diffusion measure with initial distribution μ_0.*

An $\mathcal{L}$-diffusion measure with a given initial distribution is unique.

While the $\mathcal{L}$-diffusion satisfies the strong Markov property ([32] Theorem 1.3.7) nothing assures that such process will be Feller, indeed there are some examples of Riemannian manifolds in which the Brownian motion is not a Feller process (see, e.g. [62] Example 8.2).

We will show that under some reasonable conditions the diffusion has the Feller property.

Definition 6.19. *Define a second order differential operator $\mathcal{L}, \mathcal{D}(\mathcal{L})$ on a smooth manifold M as*

$$\mathcal{L}_0 f(x) := \frac{1}{2}\sum_{k=1}^{r} G_k(x)G_k(x)f(x) + G_0(x)f(x) \tag{6.10}$$

for any $x \in M$ and $f \in \mathcal{D}(\mathcal{L})$.

In a chart (U, x^i), if $G_k = G_k^i \frac{\partial}{\partial x^i}$. for any $x \in U$ it is possible to express the operator as

$$\mathcal{L}_0 f(x) = \frac{1}{2} a^{ij}(x) \frac{\partial^2}{\partial x^i \partial x^j} f(x) + b^i(x) \frac{\partial}{\partial x^i} f(x) \tag{6.11}$$

where $b^i = G_0^i + \sum_k G_k^j \frac{\partial}{\partial x^j} G_k^i$ and $a^{ij} = \sum_k G_k^i G_k^j$. Because a^{ij} is already semi-positive definite it is elliptic at $p \in M$ if and only if its symbol is not degenerate

Definition 6.20. *Let (M, g, ∇) a manifold of bounded geometry (see definition 3.44) and let $r(M)$ its injectivity radius.*

A function is said C^k-bounded if $f \in C^k(M)$ and for every $r_0 \in (0, r(M))$ and every multi-index α there exists a constant $C_\alpha < \infty$ such that $|D_\alpha|_x f(x)| \leq C_\alpha$.

A function is C^k-bounded if and only if $||\nabla^k f||_\infty$ is bounded ([49] Remark 19). The set of C^k-bounded function over M is denoted by $C^k_b(M)$

A differential operator P of order n is said C^∞-bounded if, for every $r_0 \in (0, r(M))$ and for every pair of multi-indexes there is a constant $C_{\alpha,\beta} \geq 0$ such that $D_\beta|_x P_\alpha(x)| \leq C_{\alpha,\beta}$ in every geodesic chart $(B_{r_0}(p), \mathrm{Exp}_p^{-1})$, for any $p \in M$.

Proposition 6.21 ([49] Remark 22). *Every C^∞-bounded vector field G satisfies the following conditions:*

1. *Consider the pointwise metric $||.||_g$ defined in equation (3.41)*

$$||||\nabla^k G||_g||_\infty \leq a_k < \infty$$

2. *$\mathcal{L}_0$ defined in equation (6.10) is C^∞ bounded if and only if the vector fields G_k are C^∞-bounded*

3. *In a compact Riemannian manifold any smooth vector field is C^∞-bounded*

Theorem 6.22 ([49] theorem 28). *Let $(M, g\nabla)$ a manifold with bounded geometry and let $\mathcal{L}_0$ an uniformly elliptic second order differential operator defined as in (6.10) with any $(G_h)_{h=0,\cdots,r}$ are real smooth and C^∞-bounded vector fields. Let*

$$D_k := \{f \in C_0(M) \cap C^\infty(M) \cap C^k_b(M) : \mathcal{L}_0 f \in C_0(M)\} \tag{6.12}$$

If $\mathcal{L}$ the closure of $\mathcal{L}_0|_{D_k}$ it is the generator of a Feller semi-group in $C_0(M)$. $\mathcal{L}$ doesn't depend from k

To apply theorem 5.29 to the operator $\mathcal{L}_0$ we will consider a restriction of the domain D_k. The space of compactly supported smooth functions $C^\infty_c(M)$ is a natural candidate.

Proposition 6.23. *Let $\mathcal{L}_0$ as in equation (6.10) and suppose that the coefficients G_h are smooth and C^∞-bounded vector fields. Given any $f \in C^\infty_c(M)$ and any $n > 0$ there exists $M = M(f, n)$ such that*

$$||\mathcal{L}^n_0 f||_{C_0} \leq M$$

Proof. Let $\{U_\alpha\}_{\alpha\in I}$ an atlas of M. It is an open cover for $\mathrm{supp}(f)$. Let $\{U_i\}_{i=1,\cdots N}$ a finite subcover and let $C_h, h = 0, \cdots r$, C_f real numbers such that

$$C_h \geq \max_{i=1,\cdots N} \left(||G_h||_{C^{2n}(U_i)}\right)$$

$$C_f \geq \max_{i=1,\cdots N} \left(||f||_{C^{2n}(U_i)}\right)$$

Such number exists because the coefficients are C^∞-bounded and f is a smooth compactly supported function. It is straightforward to check that

$$||\mathcal{L}^n_0 f||_{C_0} \leq \left(\frac{1}{2}\sum_{h=1}^{r} C_h^2 + C_0\right)^n C_f =: M < \infty$$

□

Consider now the Laplace-Beltrami operator of definition 3.22. Because its principal symbols is the metric tensor g^{ij} (see equation (3.4)) the operator is uniformly elliptic. By using theorem 6.22 we obtain that in any manifold of bounded geometry such operator is the generator of a Feller semi-group.

Theorem 6.24 ([49] theorem 39). *Let M of bounded geometry. The C_0-closure of the Laplace Beltrami operator is the generator of a Feller semi-group. Its domain is given by D_k defined by equation (6.12) and doesn't depend from k.*

While theorem 6.22 gives us a sufficient condition for a diffusion process to be Feller, such condition is not necessary, as shown in the next example.

Proposition 6.25 ([62] corollary 7.2). *Let M a Cartan-Hadamard manifold i.e a Riemannian manifold simply connected, complete and with not positive sectional curvature. The Laplace Beltrami operator of M is the generator of a Feller semi-group*

Example 6.26. *Consider now the warp product $(M = \mathbb{R} \times_f S^1, g)$, where $f(t) = e^{t^3}$. The curvature tensor is not bounded (see lemma 3.19), so it isn't a manifold with bounded geometry.*

Let $\pi : \tilde{M} \to M$ the universal covering of M. The universal covering map is a local isometry w.r.t. the pullback metric $\pi^* g$ so the Riemannian manifold $(\tilde{M}, \pi^* g)$ isn't of bounded geometry (local isometries preserve the connection, so they preserve the curvature tensor).

It is possible to show [62, example 8.2] that the sectional curvature of the manifold is

$$k(t, \theta) = -\frac{f''(t)}{f(t)} \leq 0$$

Because π is a local isometry that implies that the sectional curvature of $\tilde{M}$ is not positive as well.

Let now $d : \mathbb{R} \times S^1 \to \mathbb{R}^+$ the Riemannian distance (definition 3.6) of $\mathbb{R} \times S^1$ w.r.t. the usual metric. By the Hopf-Rinow theorem 3.33 such metric is complete. Let now $d_f : \mathbb{R} \times_f S^1 \to \mathbb{R}^+$ the geodesic distance of the warp product. Because $0 < e^{t^3} \leq L < \infty$ on compact sets, the two metric are equivalent, so M is complete. By lemma 3.36 we have that $\tilde{M}$ is complete. So it is a Cartan-Hadamard manifold and by proposition 6.25 the Laplace Beltrami operator is the generator of a Feller semi-group.

6.5 Brownian motion on a manifold

Let (M, g, ∇) a Riemannian manifold endowed with the Levi Civita connection and let Δ_M the Laplace Beltrami operator (definition 3.22). It is an uniformly elliptic operator because its principal symbol is the metric, so by theorem 6.18 there is $\frac{1}{2}\Delta_M$-diffusion measure on $W(M)$.

Definition 6.27. *A Brownian motion on a manifold M with respect to a filtration $\mathcal{F}_t$ is an M-stochastic process $X : \Omega \to W(M)$ that is strong Markov w.r.to $\mathcal{F}_t$ and it is a $\frac{1}{2}\Delta_M$ diffusion process*

We have the following characterization of the Brownian motion

Theorem 6.28 ([32] proposition 3.2.1)**.** *$X : \Omega \to W(M)$, defined in a filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, \mathcal{F}_t)$ such that it is strong Markov w.r. to the filtration is a Brownian motion in the sense of definition 6.27 if and only if X is a $\mathcal{F}_t$-semi-martingale on M whose anti-development is a standard Euclidean Brownian motion.*

A process defined on $[0, \tau)$ where τ is a $\mathcal{F}_t$ stopping time is a Brownian motion if its anti-development is a local martingale up to τ with quadratic variation $\langle W, Z \rangle = \mathrm{Id}\, t$

If $\{D_n\}$ is an exhaustion by compact of M we define the heat kernel $p_{D_n}(t, x, y)$ as the fundamental solution of the heat operator on M, namely if

$$T_t f(x) := \int_{D_n} p_{D_n}(t, x, dy) f(y)$$

we have $(\frac{\partial}{\partial t} - \frac{1}{2}\Delta_M) T_t f(t, x) = 0$ on the closure of D_n and $\lim_{t \to 0} T_t f(x) = f(x)$. (these conditions are exactly the Kolmorov backward equation 5.2 and the strong continuity property of the semi-group).

The minimal heat kernel is defined as $p_M(t, x, y) = \lim_n p_{D_n}(t, x, y)$. It is possible to prove ([32] proposition 4.1.6) that for any Riemannian manifold endowed with the Levi-Civita connection $p_M(t, x, y)$ the minimal heat kernel is the transition density function of the Brownian motion. If M is of bounded geometry this is an immediate condequence of theorem 6.24 and the results of section 5.1.

6.6 Weak second order Lie Runge-Kutta method for matrix Lie groups

Let (M, g, ∇) a manifold with bounded geometry, D_k as in equation (6.12) and let $G_i(x)$, $i = 0, 1$ vector field over M such that $G_1(x) = G_1$ and the operator $\mathcal{L}$ whose expression in any local chart (U, x^i) of M can be written

$$\mathcal{L}f(x) := \left((G_0(x)x)^i + (G_1 x)^i G_1^i\right) \partial_i + \frac{1}{2}(G_1 x)^j (G_1 x)^k \partial_{jk} \tag{6.13}$$

is uniformly elliptic and C^∞-bounded (in a compact manifold this is satisfies for every smooth vector field G_0 by lemma 6.21).

By theorem 6.22 the closure of the operator $L|_{D_k}$ is the generator of a Feller semi-group and by the discussion in the end of section 5.1, up to a random time $e(X)$ such semi-group can be expressed as

$$T_t f(x) = \mathbb{E}\left[f(X_t)|X_0 = x\right] \tag{6.14}$$

where X_t is the solution of the Stratonovich SDE

$$\begin{aligned} dX_t &= G_0(X_t)X_t dt + G_1 X_t dW_t \\ X_0 &= x \end{aligned}$$

where W_t is an Euclidean Brownian motion.

We can now use the result of theorem 5.29 to give an approximation of equation (6.14) in terms of the infinitesimal generator

Proposition 6.29. *Let $\mathcal{L}$ the operator defined in equation (6.13) and let $f \in D(\mathcal{L}^3)$ such that $||\mathcal{L}^3 f||_{C_0} < \infty$ (e.g $f \in C_c^\infty(M)$) and let $h > 0$. On any chart (U, x^i) we have that*

$$\begin{aligned}
&\mathbb{E}\left[f(X_h)|X_0 = I_n\right] = I_n + h\left(\left(G_0^i + \frac{1}{2}G_1^i G_1^i\right)\partial_i f + G_1^i G_1^j \partial_{ij} f\right) + \\
&\frac{h^2}{2}\left(G_0^j \partial_j G_0^i + (G_0^i)^2 + G_0^i (G_1^j)^2 + \frac{1}{2}(G_1^j)^2 \partial_j G_0^i + \frac{1}{4}(G_1^i)^4 + \frac{1}{2}G_1^j G_1^k \partial_{jk} G_0^i + \frac{1}{2}G_1^j G_1^k \partial_j G_0^i\right)\partial_i f + \\
&\frac{h^2}{2}\left(G_0^i G_0^j + (G_1^j)^2 G_0^i + \frac{3}{4}(G_1^j)^2 (G_1^i)^2 + 2G_0^j G_1^j G_1^i + (G_1^i)^3 G_1^j + \frac{1}{2}G_1^j G_1^k \partial_k G_0^i\right)\partial_{ij} f + \\
&\frac{h^2}{2}\left(G_0^i G_1^j G_1^k + \frac{3}{2}(G_1^i)^2 G_1^j G_1^k\right)\partial_{ijk} f + \frac{h^2}{8}G_1^i G_1^j G_1^k G_1^l \partial_{ijkl} f + O(h^3)
\end{aligned}$$

Definition 6.30. *Let G a matrix Lie group. Define*

$$\begin{aligned}
Y^i &= \operatorname{expm}\left(\sqrt{h} d G_1 h + c_k^i G_0(Y^k)\right) \\
X_{n+1} &= \operatorname{expm}\left(\sqrt{h}\beta G_1 + h\alpha_j G_0(Y^j) + h^{\frac{3}{2}}\sum_j \gamma\left[G_1, G_0(Y^j)\right]\right) X_n \\
X_0 &= I_n
\end{aligned}$$

where α_j, c_k^i are constants and d, β and γ are random variables.

Remark. *If $G_1 = 0$ i.e. the SDE is an ODE, it is known that this class of methods can have at most order 2 (see [55])*

We will now expand such scheme and impose the right conditions on the coefficients so that its mean value will approximate equation (6.14) with $x = I_n$ up to second order. Choice now a chart on G such that $G_0 = \{G_0^i\}$ and $G_1 = \{G_1^i\}$. By Expanding G_0 as a Taylor series around I_n

$$Y^i = I_n + \sqrt{h} d G_1 + h\left(\sum_k c_k^i G_0 + \frac{1}{2}d^2 (G_1)^2\right) + O(h^{\frac{3}{2}}) \tag{6.15}$$

We now plug equation (6.15) in the expression of X_1 and we do a taylor expansion of G_0 around the identity. If we denote $b := \sum_m \alpha_m$ the expansion became

$$\begin{aligned}
X_1 = I_n &+ \sqrt{h}\beta G_1 + h\left(b G_0 + \frac{\beta^2}{2}(G_1)^2\right) + \\
&+ h^{\frac{3}{2}}\left(b d G_0' G_1 + \beta b G_0 G_1 + \frac{\beta^3}{6}(G_1)^3 + \gamma\left[G_1, G_0\right]\right) + \\
&+ h^2\left(\frac{1}{2}b^2 (G_0)^2 + \frac{1}{2}b\beta^2 G_0 (G_1)^2 + \frac{\beta^4}{24}(G_1)^4 + \sum_k \alpha_j c_k^j G_0' G_0\right) + \\
&+ h^2\left(\frac{1}{2}b d^2 G_0'' (G_1)^2 + \frac{1}{2}b d G_0' (G_1)^2 + \gamma\beta\left[G_1, G_0\right] G_1 + \gamma d\left[G_1, G_0' G_1\right]\right) + O(h^{\frac{5}{2}})
\end{aligned}$$

In local coordinates the first commutator is equal to

$$\left[G_1^j\frac{\partial}{\partial x^j},G_0^i(0)\frac{\partial}{\partial x^i}\right]==G_1^j\partial_jG_0^i(0)\partial_i$$

Let now f a $C^\infty(M)$ function. Suppose $\mathbb{E}[\beta]=\mathbb{E}[\gamma]=\mathbb{E}[d]=\mathbb{E}[\gamma d]=0$. Consider the Taylor expansion of $\mathbb{E}[f(X_1)|X_0=I_n]$ around the identity.

$$\begin{aligned}
&\mathbb{E}\left[f(X_1)|X_0=I_n\right]=I_n+h\left(\partial_if\left(bG_0^i+\frac{\mathbb{E}[\beta^2]}{2}G_1^iG_1^i\right)+\frac{\mathbb{E}[\beta^2]}{2}\partial_{ij}fG_1^iG_1^j\right)+\\
&+h^2\partial_if\left(\frac{1}{2}\sum_n\alpha_mc_n^mG_0^j\partial_jG_0^i+\frac{1}{2}b^2(G_0^i)^2+\frac{1}{2}b\mathbb{E}[\beta^2]G_0^i(G_1^j)^2+\frac{\mathbb{E}[\beta^4]}{24}(G_1^i)^4\right)+\\
&+h^2\partial_if\left(\frac{1}{2}b\mathbb{E}\left[d^2\right](G_1^j)^2\partial_jG_0^i+\frac{1}{2}b^2\mathbb{E}\left[d^2\right]G_1^jG_1^k\partial_{jk}G_0^i+\mathbb{E}\left[\gamma\beta\right]G_1^k\partial_kG_0^iG_1^i\right)+\\
&+\frac{h^2}{2}\partial_{ij}f\left(b\mathbb{E}[\beta^2]G_0^i(G_1^j)^2+\frac{\mathbb{E}[\beta^4]}{4}(G_1^i)^2(G_1^j)^2+\frac{\mathbb{E}[\beta^4]}{6}(G_1^i)^3G_1^j\right)+\\
&+\frac{h^2}{2}\partial_{ij}f\left(b^2G_0^iG_0^j+2b\mathbb{E}[\beta^2]G_0^iG_1^iG_1^j+2\left(b\mathbb{E}\left[d\beta\right]+\mathbb{E}\left[\gamma\beta\right]\right)G_1^iG_1^k\partial_kG_0^j\right)+\\
&+\frac{h^2}{2}\partial_{ijk}f\left(b\mathbb{E}[\beta^2]G_0^iG_1^jG_1^k+\frac{\mathbb{E}[\beta^4]}{2}(G_1^i)^2G_1^jG_1^k\right)+\frac{h^2}{24}\mathbb{E}[\beta^4]\partial_{ijkl}f\left(G_1^iG_1^jG_1^kG_1^l\right)+O(h^{\frac{5}{2}})
\end{aligned}$$

Any Lie group which admits with a left (or right) metric is a manifold of bounded geometry w.r.t. that metric as shown in example 3.47. We can now state the main theorem

Theorem 6.31. *Let G a matrix Lie group with a left (or right) invariant metric. Let $\mathcal{L}$, f as in proposition 6.29, $T_tf(x)$ as in equation (6.14) and $0<h<e(X_t)$.*

Let X_{n+1} the numerical scheme of definition 6.30. If the coefficients $\alpha_j, c_k^i, d, \beta$ and γ satisfies the order condition

$$\begin{aligned}
&\mathbb{E}[\beta]=\mathbb{E}[\gamma]=\mathbb{E}[d]=0 && \\
&\sum_m\alpha_m=:b=1 && \sum_n\alpha_mc_n^m=1\\
&\mathbb{E}\left[\beta^2\right]=1 && \mathbb{E}\left[\beta^4\right]=3\\
&\mathbb{E}\left[d^2\right]=\frac{1}{2} && \mathbb{E}\left[\gamma d\right]=0\\
&\mathbb{E}\left[\gamma\beta\right]=\frac{1}{4} && \mathbb{E}\left[d\beta\right]=0
\end{aligned}$$

we have that:

$$\mathbb{E}\left[f(X_t)-f(X_1)|X_0=I_n\right]=O(h^{\frac{5}{2}})$$

In particular, if we choice β to be a standard Gaussian random variable , η to be a standard Gaussian random variable indipendent by β i.e. $\mathbb{E}\left[\beta\eta\right]=0$ and we define $\gamma=\frac{1}{4}\beta$, $d=\frac{1}{2}\eta$ the random variables satisfy all the order conditions

7 Conclusions

The method of definition 6.30 provides a weak second order integrator for diffusion SDE on Lie groups whose infinitesimal generator has the form (6.13).

The results on the deterministic method (see [55]) suggest that to obtain methods of order higher than two it will be necessary to use more refined schemes similar to the RKMK methods described in section 2.3.

While the weak Lie RK scheme require less commutators than the schemes used for strong approximation (see e.g. [14]) we will expect that the number of commutator will grow sensibly for higher orders. As far as we know it isn't known if the tools of the free Lie algebras described in chapter 2 can be extended to the stochastic case.

Moreover, the order conditions of theorem 6.31 heavily rely on the hypothesis that the coefficient G_1 in (6.13) is constant. The presence of a second order term and the form of the Itô Magnus expansion of section 6.1 suggest that approximating elements of the universal enveloping algebra can help to handle the non-constant case, but further research will be needed to verify this statement.

As far as we know It is yet to be explored if the exotic aromatic tree formalism can be extended to the case of SDE on Lie groups. If such an algebraic structure exists it is possible that will simplify the calculation and reduce the complexity of the numerical scheme.

We hope that generality of the Talay Tubaro expansion of theorem 5.29 could be used to find Runge Kutta methods for more general classes of manifold widely used on practical applications, like Grassmannian manifolds or even general homogeneous manifolds.

References

[1] M. Abramowitz and I. Stegun. *Handbook of Mathematical Functions: With Formulas, Graphs, and Mathematical Tables*. Applied mathematics series. Dover Publications, 1965.

[2] S. Anmarkrud and A. Kværnø. Order conditions for stochastic Runge-Kutta methods preserving quadratic invariants of Stratonovich SDEs. *J. Comput. Appl. Math.*, 316:40–46, 2017.

[3] M. Atiyah. *Introduction To Commutative Algebra*. CRC Press, mar 2018.

[4] N. Bhunia, S. Pahan, and B. Arindam. Pseudo-projective curvature tensor on warped product manifolds and its applications in space-times. *SUT Journal of Mathematics*, 57:93–107, 02 2022.

[5] T. Björk. *On the Geometry of Interest Rate Models*, pages 133–215. Springer Berlin Heidelberg, Berlin, Heidelberg, 2004.

[6] J. F. Blowey, J. P. Coleman, and A. W. Craig, editors. *Theory and Numerics of Differential Equations*. Springer Berlin Heidelberg, 2001.

[7] G. Bogfjellmo. Algebraic structure of aromatic b-series, 2019.

[8] N. Bourbaki. *Lie Groups and Lie Algebras: Chapters 1-3*. Bourbaki, Nicolas: Elements of mathematics. Springer, 1989.

[9] E. Bronasco. Exotic B-series and S-series: algebraic structures and order conditions for invariant measure sampling. *Submitted*, 2022.

[10] E. Bronasco and A. Laurent. Clumped exotic forests and substitution law for exotic S-series. *In preparation*, 2023.

[11] C. Brouder. Runge–Kutta methods and renormalization. *The European Physical Journal C*, 12(3):521–534, feb 2000.

[12] K. Burrage and P. M. Burrage. High strong order explicit Runge-Kutta methods for stochastic ordinary differential equations. *Appl. Numer. Math.*, 22(1-3):81–101, 1996. Special issue celebrating the centenary of Runge-Kutta methods.

[13] 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*, 133(1-4):34–48, Jan. 1999.

[14] K. Burrage and P. M. Burrage. Order conditions of stochastic Runge-Kutta methods by B-series. *SIAM J. Numer. Anal.*, 38(5):1626–1646, 2000.

[15] J. Butcher. An algebraic theory of integration methods. *Mathematics of Computation - Math. Comput.*, 26:79–79, 01 1972.

[16] F. Casas and A. Iserles. Explicit Magnus expansions for nonlinear equations. *Journal of Physics A: Mathematical and General*, 39(19):5445–5461, apr 2006.

[17] G. Ciccotti, T. Lelievre, and E. Vanden-Eijnden. Projection of diffusions on submanifolds: application to mean force computation. *Comm. Pure Appl. Math.*, 61(3):371–408, 2008.

[18] A. Connes and D. Kreimer. Hopf algebras, renormalization and noncommutative geometry. *Communications in Mathematical Physics*, 199(1):203–242, dec 1998.

[19] K. Debrabant. Runge-Kutta methods for third order weak approximation of SDEs with multidimensional additive noise. *BIT Numer. Math.*, 50(3):541–558, 2010.

[20] K. Debrabant and A. Kværnø. B-series analysis of stochastic Runge-Kutta methods that use an iterative scheme to compute their internal stage values. *SIAM J. Numer. Anal.*, 47(1):181–203, 2008/09.

[21] K. Debrabant and A. Kværnø. Composition of stochastic B-series with applications to implicit Taylor methods. *Appl. Numer. Math.*, 61(4):501–511, 2011.

[22] P. Diaconis, S. Holmes, and M. Shahshahani. Sampling from a manifold. In *Advances in modern statistical theory and applications: a Festschrift in honor of Morris L. Eaton*, volume 10 of *Inst. Math. Stat. (IMS) Collect.*, pages 102–125. Inst. Math. Statist., Beachwood, OH, 2013.

[23] M. P. do Carmo. *Riemannian Geometry.* Birkhäuser Boston, 1992.

[24] J. Eldering. *Manifolds of Bounded Geometry*, pages 35–74. Atlantis Press, Paris, 2013.

[25] M. Emery. *Stochastic Calculus in Manifolds.* Springer-Verlag, 1989.

[26] K.-J. Engel and R. Nagel. *One-Parameter Semigroups for Linear Evolution Equations.* Springer-Verlag, 2000.

[27] C. L. Ernst Hairer, Gerhard Wanner. *Geometric Numerical Integration.* Springer-Verlag, 2006.

[28] J. Gallier and J. Quaintance. Metrics, connections, and curvature on Lie groups. In *Differential Geometry and Lie Groups*, pages 617–663. Springer International Publishing, 2020.

[29] E. Hairer and G. Wanner. On the Butcher group and general multi-value methods. *Computing*, 13(1):1–15, mar 1974.

[30] S. Helgason. Symmetric spaces. In *Differential Geometry, Lie Groups, and Symmetric Spaces*, pages 197–228. American Mathematical Society, jun 1979.

[31] J. Hilgert and K.-H. Neeb. *Structure and Geometry of Lie Groups.* Springer New York, 2012.

[32] E. Hsu. *Stochastic Analysis on Manifolds.* American Mathematical Society, feb 2002.

[33] N. Ikeda and S. Watanabe. *Stochastic Differential Equations and Diffusion Processes.* Kodansha scientific books. North-Holland, 1989.

[34] A. Iserles, H. Munthe-Kaas, S. Norsett, and A. Zanna. Lie group methods. *Acta Numerica*, 9:215–, 01 2000.

[35] A. Iserles and S. Nørsett. On the solution of linear differential equations in Lie groups. *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 357(1754):983–1019, apr 1999.

[36] L. Isserlis. On a formula for the product-moment coefficient of any order of a normal frequency distribution in any number of variables. *Biometrika*, 12(1/2):134–139, 1918.

[37] O. Kallenberg. *Foundations of Modern Probability.* Springer New York, 2002.

[38] K. Kamm, S. Pagliarani, and A. Pascucci. On the stochastic Magnus expansion and its application to SPDEs. *Journal of Scientific Computing*, 89(3), oct 2021.

[39] K. Kamm, S. Pagliarani, and A. Pascucci. Numerical solution of kinetic spdes via stochastic Magnus expansion. *Mathematics and Computers in Simulation*, 207:189–208, 2023.

[40] I. Karatzas and S. E. Shreve. *Brownian Motion and Stochastic Calculus.* Springer New York, 1998.

[41] P. E. Kloeden and E. Platen. *Numerical Solution of Stochastic Differential Equations.* Springer Berlin Heidelberg, 1992.

[42] Y. Komori, T. Mitsui, and H. Sugiura. Rooted tree analysis of the order conditions of ROW-type scheme for stochastic differential equations. *BIT Numer. Math.*, 37(1):43–66, 1997.

[43] D. Küpper, A. Kværnø, and A. Rößler. A Runge-Kutta method for index 1 stochastic differential-algebraic equations with scalar noise. *BIT Numer. Math.*, 52(2):437–455, 2012.

[44] A. Laurent and H. Z. Munthe-Kaas. The universal equivariance properties of exotic aromatic Butcher-series. *In preparation*, 2023.

[45] A. Laurent and G. Vilmart. Exotic aromatic b-series for the study of long time integrators for a class of ergodic SDEs. *Mathematics of Computation*, 89(321):169–202, jun 2019.

[46] J. M. Lee. *Introduction to Smooth Manifolds.* Springer New York, 2012.

[47] T. Lelièvre, M. Rousset, and G. Stoltz. Langevin dynamics with constraints and computation of free energy differences. *Math. Comp.*, 81(280):2071–2125, 2012.

[48] W. Magnus. On the exponential solution of differential equations for a linear operator. *Communications on Pure and Applied Mathematics*, 7(4):649–673, 1954.

[49] S. Mazzucchi, V. Moretti, I. Remizov, and O. Smolyanov. Chernoff approximations of feller semigroups in riemannian manifolds. *Mathematische Nachrichten*, 296(3):1244–1284, 2020.

[50] R. I. McLachlan, K. Modin, H. Munthe-Kaas, and O. Verdier. B-series methods are exactly the affine equivariant methods. *Numerische Mathematik*, 133(3):599–622, aug 2015.

[51] R. I. McLachlan, K. Modin, H. Munthe-Kaas, and O. Verdier. Butcher series: A story of rooted trees and numerical methods for evolution equations, 2017.

[52] J. Milnor. Curvatures of left invariant metrics on Lie groups. *Advances in Mathematics*, 21(3):293–329, 1976.

[53] G. Milstein and M. Tretyakov. Computing ergodic limits for langevin equations. *Physica D: Nonlinear Phenomena*, 229(1):81–95, 2007.

[54] M. Muniz, M. Ehrhardt, M. Günther, and R. Winkler. Higher strong order methods for linear itô SDEs on matrix Lie groups. *BIT Numerical Mathematics*, 62(4):1095–1119, jan 2022.

[55] H. Munthe-Kaas. High order Runge-Kutta methods on manifolds. *Applied Numerical Mathematics*, 29(1):115–127, jan 1999.

[56] H. Munthe-Kaas and O. Verdier. Aromatic Butcher series. *Foundations of Computational Mathematics*, 16(1):183–215, feb 2015.

[57] H. Z. Munthe-Kaas and K. K. Føllesdal. Lie-Butcher series, geometry, algebra and computation, 2017.

[58] H. Munthe–Kaas and B. Owren. Computations in a free Lie algebra. *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 357(1754):957–981, apr 1999.

[59] B. Øksendal. *Stochastic Differential Equations.* Springer Berlin Heidelberg, 2003.

[60] G. A. Pavliotis. *Stochastic Processes and Applications.* Springer New York, 2014.

[61] A. Pazy. *Semigroups of Linear Operators and Applications to Partial Differential Equations.* Springer New York, 1983.

[62] S. Pigola and A. G. Setti. The Feller property on riemannian manifolds. *Journal of Functional Analysis*, 262(5):2481–2515, 2012.

[63] C. Reutenauer. *Free Lie Algebras*. LMS monographs. Clarendon Press, 1993.

[64] L. C. G. Rogers and D. Williams. *Diffusions, Markov Processes, and Martingales*, volume 1 of *Cambridge Mathematical Library*. Cambridge University Press, 2 edition, 2000.

[65] A. Rößler. Stochastic Taylor expansions for the expectation of functionals of diffusion processes. *Stochastic Anal. Appl.*, 22(6):1553–1576, 2004.

[66] A. Rößler. Rooted tree analysis for order conditions of stochastic Runge-Kutta methods for the weak approximation of stochastic differential equations. *Stoch. Anal. Appl.*, 24(1):97–134, 2006.

[67] A. Rößler. Runge-Kutta methods for Itô stochastic differential equations with scalar noise. *BIT Numer. Math.*, 46(1):97–110, 2006.

[68] A. Rößler. Strong and weak approximation methods for stochastic differential equations—some recent developments. In *Recent developments in applied probability and statistics*, pages 127–153. Physica, Heidelberg, 2010.

[69] S. C. R.W. Sharpe. *Differential Geometry, Cartan's Generalization of Klein's Erlangen Program*. Springer New York, 1997.

[70] J. M. Sanz-Serna and L. Abia. Order conditions for canonical runge–kutta schemes. *SIAM Journal on Numerical Analysis*, 28(4):1081–1096, 1991.

[71] R. Schilling, L. Partzsch, and B. Böttcher. *Brownian Motion: An Introduction to Stochastic Processes*. De Gruyter graduate. De Gruyter, 2012.

[72] D. Talay and L. Tubaro. Expansion of the global error for numerical schemes solving stochastic differential equations. *Stochastic Anal. Appl.*, 8(4):483–509 (1991), 1990.

[73] V. S. Varadarajan. *Lie Groups, Lie Algebras, and Their Representations*. Springer New York, 1984.

[74] Z. Wang, Q. Ma, Z. Yao, and X. Ding. The Magnus expansion for stochastic differential equations. *Journal of Nonlinear Science*, 30(1):419–447, sep 2019.