Håkon R. Fredheim

Magnus Expansion in Gauge Theories

July 2022

Master's thesis

2022

Håkon R. Fredheim

Master's thesis

NTNU
Norwegian University of
Science and Technology
Faculty of Natural Sciences
Department of Physics

NTNU
Norwegian University of
Science and Technology

NTNU
Norwegian University of
Science and Technology

NTNU
Norwegian University of Science and Technology

Magnus Expansion in Gauge Theories

Håkon R. Fredheim

Applied Physics and Mathematics
Submission date: July 2022
Supervisor: Kurusch Ebrahimi-Fard
Co-supervisor: Jon Andreas Støvneng

Norwegian University of Science and Technology
Department of Physics

Abstract

Some expressions for the solution of an ordinary differential equation (ODE) are derived. The method used involves Taylor expanding the vector field defining the ODE. A generalization of a known result is derived taking this approach. Consequences of this generalization are elaborated upon, and are shown to produce further insights into the solution of a related set of partial differential equations. Two physical applications of the results are presented. The first is a method for evaluating Feynman integrals, the second is a method for finding the path of swimming microbes. The results of this thesis have applications in many fields of Physics.

Sammendrag

Noen uttrykk for løsningen på en ordinær differensiallikning utledes ved hjelp av en metode som innebærer Taylor-rekkeutvikling av vektorfeltet som definerer differensiallikningen. En generalisering av et kjent resultat utledes ved hjelp av denne fremgangsmåten. Følger av denne generaliseringen drøftes, og det vises hvordan disse gir ytteligere innsikt i løsningen på en beslektet mengde partielle differnsiallikninger. To fysiske anvendelser av resultatene presenteres. Den første er en metode for å evaluere Feynman-integraler, og den andre er en metode for å finne svømmebanen til mikrober. Resultatene i denne oppgaven kan anvendes i mange fysiske fagfelt.

Contents

Preface

This is my thesis, concluding my master's degree in Applied Physics and Mathematics at NTNU. I have had to spend a lot of time getting to know new mathematical concepts and formalism, and I have been able to come up with new ideas for the problems I have been working on. It has been a tough, yet rewarding challenge.

I would like to thank my supervisor Kurusch Ebrahimi-Fard for the opportunity to do this project and very helpful discussions throughout the semester. I would also like to thank my friends and family for supporting me throughout this time.

Håkon Richard Fredheim
Kristiansand
3/7 - 2022

1 Introduction

Many problems in physics and mathematics boil down to solving a linear ordinary differential equation with initial value,

$$\dot{U}(t) = H(t)U(t),\ U(0) = 1, \tag{1}$$

where U takes values in some Lie group G, and H is a time dependent function that takes values in Lie G, the Lie algebra of G. 1 is the identity of G. A typical example of this situation is when U is the time-evolution operator in quantum mechanics. Eq. (1) is then a reformulation of the Schrödinger-equation: Given an initial state $|\psi_0\rangle$, the Schrödinger equation $\partial_t|\psi\rangle = H|\psi\rangle$ can be solved using the initial value problem (1), by posing $|\psi\rangle = U|\psi_0\rangle$.

In this thesis, we will develop and present a particular way of representing the solution of equations like (1). Namely, we wish to find some $\Omega(t) \in \mathrm{Lie}\, G$ such that $U(t) = \exp(\Omega(t))$. Our starting point is reformulating eq. (1) as an integral equation,

$$U(t) = 1 + \int_0^t d\tau H(\tau)U(\tau). \tag{2}$$

The integral equation (2) can be solved by Picard iteration, yielding the Dyson-Chen series

$$U = 1 + \overline{H} + \overline{H\overline{H}} + \overline{H\overline{H\overline{H}}} + \cdots, \tag{3}$$

where $\overline{H}(t) := \int_0^t d\tau H(\tau)$, and $\overline{H\overline{H}}(t) = \int_0^t d\tau \int_0^\tau d\tau' H(\tau)H(\tau')$, etc.[1] This is only a formal expression at this point, since we do not know whether the sum converges.[2] If H commutes with itself evaluated at different times, that is, if $H(t)H(t') = H(t')H(t)$ when $t \neq t'$, then $\overline{H}H = H\overline{H}$, and so

$$U = \exp(\overline{H}) := \sum_{n=0}^{\infty} \frac{\overline{H}^n}{n!}$$

solves eq. (1).

Now, if $H(t)H(t') \neq H(t')H(t)$, it is not obvious that U can be written as the exponential of some Lie element. If the Dyson-Chen series does not terminate, a closed expression for U may not be obtainable, and when approximating U by truncation of eq. (3), group properties may be lost, giving an expression not suitable for calculation.[3] Therefore, it is often convenient to express U as the exponential of some Lie element.

In quantum mechanics, eq. (1) appears as the Schrödinger equation, $\dot{U}(t) = H(t)U(t)$, where $H(t)$ is anti-Hermitian, meaning $H^\dagger(t) = -H(t)$, where $H^\dagger(t)$ is the conjugate-transpose (adjoint) of $H(t)$. The anti-Hermiticity of $H(t)$ implies that $U(t)$ is a unitary operator. A truncation of the Dyson-Chen series (3) does not necessarily yield a unitary operator, and so is not a desirable way of obtaining approximate solutions. It would be better to approximate solutions by the exponential of some anti-Herimitian matrix. It turns out that Ω defined above can be defined as an infinite sum of iterated brackets of H. This sum can be truncated to yield approximate solutions to the problem, which at all orders of truncation yield an operator with the desired anti-Hermitian property.

The quest for finding an efficient scheme of calculating Ω such that $U = \exp(\Omega)$, started a long time ago, the first paper providing a thorough investigation being that of Magnus in 1954 [37],

[1] The bar $(\bar{\cdot})$-notation meaning time integration is common in the literature. See e.g.[45]

[2] For instance, if H is a bounded operator on a Banach space, the sum converges: if $|H(t)| \leq M$ for some bound M, then $|U(t)| \leq 1 + |\overline{H}| + |\overline{H\overline{H}}| + \cdots \leq 1 + tM + \frac{1}{2}t^2M^2 + \cdots = e^{tM} < \infty$.

[3] If we are so lucky that the solution to our problem lies in the image of the exponential map, then we might be able to solve for Ω such that $\exp(\Omega) = U$. This is for instance not always the case for $G = \mathrm{SL}(2, \mathbb{C})$, whence the exponential map is not surjective although $\mathrm{SL}(2, \mathbb{C})$ is simply connected.

soon followed by important contributions by Chen and others. In this thesis, we shall be concerned with finding ways of expressing Ω as a means to solving eq. (1).

Magnus obtained in his seminal paper a scheme for calculating Ω, producing to the third order in H,

$$\Omega = \overline{H} - \frac{1}{2}\overline{[\overline{H}, H]} + \frac{1}{4}\overline{[[\overline{\overline{H}, H}], H]} + \frac{1}{12}\overline{[\overline{H}, [\overline{H}, H]]} + \cdots , \tag{4}$$

commonly called the Magnus expansion. Higher order terms will be of a similar form: They consist of different combinations of iterated Lie brackets of H which are iteratively time integrated, with some rational coefficients. Iserles & Nørsett [23] give a thorough review and exposition of methods to calculate the terms in eq. (4).

The fact that the problem of effectively approximating the solution of eq. (1) has remained a topic of active research for over eighty years reveals the complexity of the problem at hand. The fruits of this research is a wealth of iterative schemes and expressions for Ω, with properties suited for different versions of the problem. An extensive review of this is given by Blanes *et al.* [9].

We will in particular be concerned with calculating Ω in the context of gauge theories. In this case, H in eq. (1) appears as a connection on a principal bundle depending on time through a loop on the base space. This context is the main motivation for this thesis, and so we now state precisely what we mean by this. We suppose the reader to be familiar with some concepts from the theory of Lie groups and differential geometry. More precise definitions of the terms are given in appendix A. Relevant sources on this topic are classic texts on differential geometry, like [54] and [26], the well-known physics-directed work by Nakahara [43] and Frankel [16], and the very detailed work by Kolar *et al.* [27] containing proofs of most of the relevant results.

Let G be a Lie group with Lie algebra $\mathfrak{g}$. Let M be a manifold with tangent bundle TM, and suppose (W, x) is a chart[1] on M. We call a connection one-form ω a map $\omega : TM \to \mathfrak{g}$. Let $u : [0, 1] \to W$ be a path in $W \subset M$ such that $u(1) = u(0)$. We call this a closed path or a loop. We express ω in the local coordinates x as

$$\omega = A_i dx^i,$$

where $A_i : W \to \mathfrak{g}$. In the above equation, the Einstein summation convention is used, and it will be used throughout the thesis with the rule that repeated indices are summed over. If we evaluate ω on the path u, we get

$$\omega(\dot{u}(t)) = \dot{u}^i(t) A_i(u(t)).$$

We shall call the problem

$$\dot{g}(t) = -g(t)\dot{u}^i(t) A_i(u(t)),\ g(0) = 1 \tag{5}$$

the *holonomy problem* corresponding to the loop u. Evaluating g at $t = 1$ then gives a group element which we call the *holonomy element* corresponding to the loop u. The holonomy element corresponding to a loop is a purely geometric quantity depending on the loop and the connection.

The definition of the holonomy problem seems unmotivated at this point, but it has vast application in differential geometry and physics. We now mention a couple of examples. Berry showed in 1984 that a periodic variation of the parameters of a particular type of quantum mechanical problem yields an overall change in phase of the state of the system [5]. This phase can be expressed as the solution to an appropriately defined holonomy problem, in which the relevant gauge group is $G = U(1)$, and the connection one-form then comes from the Schrödinger equation. Guichardet

[1] $W \subset M$ is an open subset and $x : W \to \mathbb{R}^{\dim(M)}$ is a homeomorphism onto it's image, an open subset of $\mathbb{R}^{\dim(M)}$. The components of x are called local coordinates on M.

remarked in 1984 that vibrational motion in molecules generate an overall rotation [17], whence the gauge group is $G = SO(3)$, the rotation group, and the connection one-form is a reformulation of conservation of angular momentum. The same principle can be used to predict rotation of free floating objects, and is extensively used in control theory [42] [41]. We mention lastly the case of deformable bodies floating in a viscous fluid. This situation can also be reformulated as a holonomy problem, as was done by Shapere and Wilczek [50] [49]. In this case, the gauge group is SE(3), the 3D-Euclidean group, and the connection comes from some boundary conditions on the fluid.

We now introduce the notion of curvature of a connection, which is key to the main result of the thesis. Let g and u and ω be as defined before. Let

$$F_{ij} := \partial_i A_j - \partial_j A_i - [A_i, A_j], \tag{6}$$

where A_i, as before, are the components of ω with respect to the coordinates x and $\partial_i := \partial/\partial x^i$. We call $F := F_{ij} dx^i \wedge dx^j$ the curvature 2-form of the connection, and we can regard F as a 2-nd order covariant antisymmetric tensor with components F_{ij}, which we call the curvature. The name "curvature" is justified in the following sense that on a Riemannian manifold, the holonomy problem appears as the problem of parallel transport involving a Riemannian connection, and F can be identified with Riemannian curvature. As per intuition, we expect parallel transport along a contractible loop on a zero-(Riemannian)-curvature manifold to be trivial. Analogously, the holonomy element corresponding to a connection with $F = 0$ is trivial for contractible loops.[1]

It was shown by Radford and Burdick that the solution to the holonomy problem, eq. (5) can be expressed entirely in terms of the curvature F. The fact that $F = 0$ yields a trivial holonomy element then follows immediately from this expression. It is know that $F = 0$ implies path independence of the holonomy problem, that is, eq. (5) is independent of the path chosen between $u(0)$ and $u(1)$ whenever $F = 0$. This path independence only holds for homotopic paths; two non-homotopic paths might well produce different holonomy elements even though $F = 0$.

We now state the result of Radford and Burdick, which is the starting point of our discussion. Let $I = (i_1, ..., i_n)$ be a tuple of indices ranging from 1 to $\dim M$. The n-fold iterated integral of u is defined by

$$\int_0^t du^{i_1,\dots,i_n} := \int_0^t d\tau^1 \int_0^{\tau^1} d\tau^2 \cdots \int_0^{\tau^{n-1}} d\tau^n \dot{u}^{i_1}(\tau^1) \cdots \dot{u}^{i_n}(\tau^n), \tag{7}$$

where $\dot{u}^i$ is the i-th component of $\dot{u}$ with respect to the coordinates x.

An important fact is that these integrals are independent of reparametrization of u [13], and so the above iterated integral really only depends on the geometry of the path.

Radford and Burdick showed in [45] that the solution to the holonomy problem, eq. (5), for a closed path u, is $g(1) = g(0) \exp(\Omega)$, where[2]

$$\Omega = -\frac{1}{2} F_{ij}(u(0)) \int_0^1 du^{ij} - \frac{1}{3} \nabla_i F_{jk}(u(0)) \int_0^1 du^{ijk} + \cdots . \tag{8}$$

Here, F is as defined in eq. (6), and $\forall f : M \to \mathfrak{g}$: $\nabla_i f := \partial_i f - [A_i, f]$. We shall refer to the expansion on the right hand side of eq. (8) as the Radford expansion. For illustration, we only present the first two terms of this expansion here. The full expression will be presented in Theorem 4.1. The only path-dependent components of this expansion are the iterated integrals. The coefficients are all of the form $\nabla_{i_1} \cdots \nabla_{i_{n-2}} F_{i_{n-1} i_n}(u(0))$.

Although this expression did not appear before Radford & Burdick's paper [45], similar expressions have been alluded to earlier. Loos gives a similar expression in [34] for a special case of loops in his proof of a theorem by Chow (Theorem 4 in Loos's paper).

[1] This is theorem 9.1 in [26].

[2] A minus sign error in the original paper is often repeated when cited. Compare [38] and [36].

Curiously, this "covariant" version of the Magnus expansion has not gained a lot of attention outside the realm of control theory. The problem of finding the holonomy element corresponding to a path on the base space is ubiquitous in its physical application, so we believe that the Radford expansion (8) deserves to be put to use in physics proper. However, the derivation of eq. (8) as conveyed in [45] remains obscure, and leaves little room for deducing any relevant generalization. We will therefore find it useful, to provide a detailed explanation of how eq. (8) is derived, and then show that it can be derived from a more general theorem.

The traditional approach to the Magnus expansion is to work out and evaluate the terms in the expansion (4). In this thesis we shall take the approach of Radford, and work out ways of evaluating the Magnus expansion by Taylor expanding H. The reason this has not been a popular approach in practice is that H is not analytic in general, which is required for a Taylor expansion. However, if an analytic expression for H is obtainable, our approach seems the most suited for analysing the problem. By inspiration of how eq. (8) is derived in Radford & Burdick's paper [45], the program of this thesis will be to produce solutions to differential equations of the form of (1) in terms of exponentials of Taylor expansions of H.

This thesis is organised as follows: In Section 2 we present and clarify the notation that will be used and present some preliminary remarks. In Section 3, we review some basics on vector fields and flows. In Section 4, a new proof of Radford's theorem is presented. In Section 5 a couple of consequences of the new proof are elaborated upon. In Section 6, we review the theorems and expansions of Magnus, Dynkin, Fer and Wilcox. In Section 7, we rewrite these expansions in terms of products of Lie brackets and iterated integrals. In Section 8, the results of Section 7 are used to produce a new Radford-like expansion by means of the technique of the proof of Section 4. In Section 9, we present some applications of our results. In Appendix A we give a review of some notions from differential geometry, and in Appendix B, the original proof by Radford is presented in its differential geometric context.

2 Notation and preliminary remarks

2.1 Notation

Let f be a time dependent quantity, then we will often denote its time dependence by subscript f_t or superscript f^t.

$$\frac{d}{dt}f = \dot{f}$$

denotes the time derivative of f, and

$$\overline{f}_t := \int_0^t d\tau f_\tau.$$

Iterated overlines shall denote time-iterated integrals:

$$\overline{f\overline{f\overline{f}}}_t = \int_0^t d\tau^1 \int_0^{\tau^1} d\tau^2 \int_0^{\tau^2} d\tau^3 f_{\tau^1} f_{\tau^2} f_{\tau^3},$$

and

$$\overline{\overline{\overline{f}f}f}_t = \int_0^t d\tau^1 \int_0^{\tau^1} d\tau^2 \int_0^{\tau^2} d\tau^3 f_{\tau^3} f_{\tau^2} f_{\tau^1},$$

etc.

Let G be a Lie group. We denote the Lie algebra of G by Lie G or $\mathfrak{g}$. The Lie brackets on Lie G are denoted $[\cdot,\cdot]$, and we shall assume to be able to express the Lie bracket as $[A, B] := AB - BA$.

Let M be a manifold. Then TM denotes the tangent bundle of M. $\mathcal{X}(M)$ denotes the set of vector fields on M. $C^\infty(M)$ denotes the set of smooth functions on M.

We shall use the following abbreviations: "RHS" will mean "Right hand side", and "LHS" will mean "Left hand side". "IVP" will mean "Initial value problem". "ODE" will mean "Ordinary differential equation". "PDE" will mean "Partial differential equation". "iff" will mean "if and only if"

2.2 Iterated integrals of operator-valued functions

The inverse of U, as defined in eq. (3) is[1]

$$U^{-1} = 1 - \overline{H} + \overline{\overline{H}H} - \overline{\overline{\overline{H}H}H} + \cdots \tag{9}$$

and satisfies

$$\dot{U}^{-1} = -U^{-1}H. \tag{10}$$

From this follow several identities. From the requirement that $U^{-1}U = 1$, we can infer

$$\begin{aligned}
0 &= -\overline{H}\,\overline{H} + \overline{\overline{H}H} + \overline{H\overline{H}} \\
0 &= \overline{H}\,\overline{\overline{H}H} - \overline{H\overline{H}}\,\overline{H} + \overline{H\overline{H\overline{H}}} - \overline{\overline{\overline{H}H}H} \\
&\dots
\end{aligned} \tag{11}$$

These can equally be shown by integration by parts. It is custom to define the time ordering operator T by

$$TH(\tau)H(\tau') = \begin{cases} H(\tau)H(\tau') \text{ if } \tau > \tau' \\ H(\tau')H(\tau) \text{ if } \tau < \tau', \end{cases} \tag{12}$$

and analogous operations on higher order terms. Its adjoint T^* is defined by opposite ordering. Then,

$$\begin{aligned}
U &= T\exp(\overline{H}) \\
U^{-1} &= T^*\exp(-\overline{H}).
\end{aligned} \tag{13}$$

The time ordering operator T was first introduced by Feynman [15]. For reasons that will become clear later, we will adopt the notation $\mathrm{Ev}_H \equiv T\exp(\overline{H})$.

2.3 Iterated integrals and Taylor series

As mentioned in the introduction, we shall be concerned with the special case of eq. (3), where $H(t) = H(t, u(t))$, and that $u : [0, 1] \to M$ is a path on some manifold M. If further, $H(t) = \dot{u}^i(t)H_i(u(t))$, where $\dot{u}^i$ are the components of $\dot{u}$ with respect to some chart on M, then finding Ω in eq. (4) becomes the "holonomy problem", where H is interpreted as a connection on M. Ultimately, our goal is to express Ω in terms of derivatives of the $H_i's$ and iterated integrals, defined by eq. (7).

In anticipation of how eq. (8) is ultimately derived, note that we can Taylor expand H as

$$\begin{aligned}
H(t, u(t)) = \dot{u}^a(t)H_a(u(0)) &+ \dot{u}^a(t)\partial_i H_a(u(0))\int_0^t du^i + \dot{u}^a(t)\partial_i\partial_j H_a(u(0))\int_0^t du^{ij} \\
&+ \dot{u}^a(t)\partial_i\partial_j\partial_k H_a(u(0))\int_0^t du^{ijk} + \cdots
\end{aligned} \tag{14}$$

[1]This follows from $\frac{d}{dt}U^{-1}U = 0 \Rightarrow \dot{U}^{-1} = -U^{-1}\dot{U}U^{-1} = -U^{-1}H$.

Inserting this into the Magnus expansion, eq. (4) does indeed yield an expression of Ω in terms of derivatives of H and iterated integrals, but it seems *prima facie* impossible to regroup the terms in order to obtain eq. (8). We shall show that there is indeed a more transparent approach.

3 Vector fields, flows and evolution operators

In this chapter, we provide preliminaries on vector fields and flows, and address some concerns about the validity of our approach. The basic concepts presented here are largely well-established, and I have taken inspiration from some lecture notes by Schmeding [48].

3.1 Taylor expansions

We will represent the solution to $\dot{g} = Hg$ in terms of a Taylor expansion of H. This requires H to be analytic with respect to the parameter in which we Taylor expand. It also requires the manifold in question to be analytic. We shall always assume that we are allowed to Taylor expand, and so we guard this as an implicit assumption whenever Taylor expansion is performed. Let $\phi : \mathbb{R}^N \to \mathbb{R}$. Let $x_0 \in \mathbb{R}^N$, then the Taylor expansion $\tilde{\phi}$ of ϕ around x_0 is defined by

$$\tilde{\phi}(x) := \exp((x - x_0)^i \partial_i)\phi(x_0). \tag{15}$$

The RHS represents the infinite sum

$$\sum_{n=0}^{\infty} \frac{(x - x_0)^{i_1} \cdots (x - x_0)^{i_n}}{n!} \partial_{i_1} \cdots \partial_{i_n} \phi(x_0)$$

such that $\exp((x-x_0)^i\partial_i)$ represents the corresponding power series of partial differential operators. Notably, for $\tilde{\phi}$ to be well-defined, we require a certain regularity on the derivatives of ϕ at x_0. We shall throughout this thesis assume $\tilde{\phi}(x) = \phi(x)$, carrying the implicit assumption of analyticity.

Taylor expansions can be generalized to represent the solutions to ODEs in $\mathbb{R}^N$. Let $f : \mathbb{R}^N \to \mathbb{R}^N$, and define the ODE

$$\dot{x} = f(x),\ x(0) = x_0. \tag{16}$$

We call the solution to this ODE the *flow* of f and denote it by $\Phi^t_f(x_0) := x(t)$. The Picard-Lindelöf theorem guarantees that Φ_f is well defined and smooth in some neighborhood of $(0, x_0) \in \mathbb{R} \times \mathbb{R}^N$ as long as f is smooth.[1] Let $\phi : \mathbb{R}^N \to \mathbb{R}$. Under sufficient assumptions on analyticity for ϕ and f,

$$\exp(tf^i\partial_i)\phi(x_0) = \phi \circ \Phi^t_f(x_0) = \phi(x(t)).$$

We call the LHS the *evolution operator* Ev_f of f, and it is defined as the time dependent power series of differential operators acting on smooth functions $\phi : \mathbb{R}^N \to \mathbb{R}$:

$$\mathrm{Ev}^t_f \phi(x_0) = \exp(tf^i\partial_i)\phi(x_0) = \sum_{n=0}^{\infty} \frac{t^n}{n!} f^{i_1}\partial_{i_1} \cdots f^{i_n}\partial_{i_n}\phi(x_0).$$

We shall solve ordinary differential equations like (16) by manipulating the evolution operator of the particular vector field in question. We present a typical argument using this technique in the following.

[1] This is a classic result. See e.g. [20], pg. 100.

Let $f, g : \mathbb{R}^N \to \mathbb{R}^N$, and $\phi : \mathbb{R}^N \to \mathbb{R}$. Let $D(f) := f^i \partial_i$ and $D(g) := g^i \partial_i$. By *derivations*, we mean operators on $C^\infty(\mathbb{R})$ that satisfy the Leibniz rule. D identifies a map $f : \mathbb{R}^N \to \mathbb{R}^N$ with a derivation $D(f)$. It can be shown that this map is one-to-one and that the commutator of two derivations is again a derivation.

Let $Y(D(f), D(g)) = \mathrm{BCH}(D(f)+D(g), -D(g))$, where BCH denotes the Baker-Campbell-Hausdorff formula, defined by

$$\mathrm{BCH}(A, B) := \log(\exp(A)\exp(B)).$$

BCH can be expressed as an infinite series of iterated commutators. For later reference, the first few terms in BCH are [10]

$$\mathrm{BCH}(A, B) = A + B + \frac{1}{2}[A, B] + \frac{1}{12}([A, [A, B]] + [B, [B, A]]) - \frac{1}{24}[B, [A, [A, B]]] + \cdots \tag{17}$$

Since commutators of derivations are derivations, $Y(D(f), D(g))$ is a derivation. Define by $X(f, g) : \mathbb{R}^N \to \mathbb{R}^N$ the map with components defined by $D(X(f, g)) = X(f, g)^i \partial_i := Y(D(f), D(g))$. Since $Y(D(f), D(g))$ is a derivation, $X(f, g)$ is well defined. Performing a Taylor expansion,

$$\phi \circ \Phi^1_{f+g}(x_0) = \exp(D(f) + D(g))\phi(x_0).$$

In order for this equality to hold, we assume that Φ^1_{f+g} exists, and that f, g and ϕ are sufficiently analytic at x_0 for the power series $\exp(D(f)+D(g))\phi(x_0)$ to converge. We can formally manipulate this power series to obtain

$$\exp(D(f) + D(g))\phi(x_0) = \exp(Y(D(f), D(g)))\exp(D(g))\phi(x_0).$$

However, it is not a given that the RHS of this equality is well-defined as an operator on $C^\infty(\mathbb{R}^N)$. For instance, $f + g$ might be analytic while neither f nor g alone is. Furthermore, it is not obvious whether the analyticity of $X(f, g)$ can be inferred from the analyticity of g and f. Next, by assumption of analyticity we use the Taylor expansion

$$\exp(Y(D(f), D(g)))\exp(D(g))\phi(x_0) = \exp(D(X(f, g)))\phi \circ \Phi^1_g(x_0) = \phi \circ \Phi^1_g \circ \Phi^1_{Y(f,g)}(x_0).$$

Thus, we have obtained the equality

$$\phi \circ \Phi^1_{f+g} = \phi \circ \Phi^1_g \circ \Phi^1_{X(f,g)}.$$

Since ϕ is an arbitrary analytic function, this implies

$$\Phi^1_{f+g} = \Phi^1_g \circ \Phi^1_{X(f,g)}. \tag{18}$$

The critical reader may think that we can do away with the notions of flows and simply deal with evolution operators as formal objects. This would simplify arguments and would relieve us of carrying assumptions of analyticity. A paper by Kostonov is a good example of how this can be done [28]. Essentially, it relies on the observation that the evolution operator formally satisfies

$$\frac{d}{dt}\mathrm{Ev}_f = \mathrm{Ev}_f D(f).$$

The above equation holds in the sense that for analytic $\phi : \mathbb{R}^N \to \mathbb{R}$,

$$\frac{d}{dt}\mathrm{Ev}_f\phi = \mathrm{Ev}_f D(f)\phi, \tag{19}$$

where $D(f)\phi = f^i\partial_i\phi$. This follows from the chain rule:

$$\frac{d}{dt}\mathrm{Ev}_f^t\phi(x) = \frac{d}{dt}\phi\circ\Phi_f^t(x) = D(f)\phi(\Phi_f^t(x)) = \mathrm{Ev}_f^t D(f)\phi(x),$$

Formal manipulations like

$$\exp(D(f)+D(g)) = \exp(X(D(f),D(g)))\exp(D(g))$$

make perfect sense, and so we shall assume such arguments to hold. However, when inferring the equality of flows, as in eq. (18), we are making several assumptions on the analyticity and convergence of the objects involved. Hence, the convergence of solutions obtained this way has to be checked.

3.2 Generalizing to manifolds

The diffeomorphism group $\mathrm{Diff}(M)$ of a manifold M is the group of all diffeomorphisms $M \xrightarrow{\sim} M$ with product defined by composition of diffeomorphisms. Denote the set of vector fields on M by $\mathcal{X}(M)$, sections of the tangent bundle $TM \to M$.[1] The flow $\Phi_X \in \mathrm{Diff}(M)$ of a vector field $X \in \mathcal{X}(M)$ is defined by $\Phi_X(x_0)$ being the solution of

$$\dot{x} = X(x);\ x(0) = x_0,$$

i.e., $\Phi_X^t(x_0) = x(t)$. In this sense, the flow is a map $\Phi^t : \mathcal{X}(M) \to \mathrm{Diff}(M)$. $\mathcal{X}(M)$ with the Lie-Jacobi bracket $[\cdot,\cdot]$ is the Lie algebra of $\mathrm{Diff}(M)$ in the sense that Φ takes the role of the exponential map.[2] Let $\phi : M \to \mathbb{R}$. Then, we have by the chain rule

$$\frac{d}{dt}\phi\circ\Phi_X^t = X\phi\circ\Phi_X^t. \tag{20}$$

Thus, using the fundamental theorem of calculus,

$$\begin{aligned}\phi\circ\Phi_X^t &= \phi + \int_0^t d\tau X\phi\circ\Phi_X^\tau \\ &= \phi + \int_0^t d\tau\big(X\phi + \int_0^\tau d\tau' XX\phi\circ\Phi_X^{\tau'}\big).\end{aligned}$$

Repeating the iteration gives

$$\phi\circ\Phi_X^t = \big(1 + tX + \frac{t^2}{2!}X^2 + \cdots\big)\phi = \exp(tX)\phi.$$

In this sense, we relate Φ_X^t with the power series in X. From (20), we have

$$\frac{d}{dt}\exp(tX) = \exp(tX)X = X\exp(tX). \tag{21}$$

The above equation holds in the sense that $\forall\phi \in C^\infty(M)$:

$$\frac{d}{dt}\exp(tX)\phi = \exp(tX)X\phi = X\exp(tX)\phi.$$

[1] $\mathcal{X}(M)$ is a module over the ring $C^\infty(M)$ of smooth functions on M, and a Lie algebra with respect to the Lie-Jacobi commutator of vector fields.

[2] In fact, for general Lie groups, the exponential map is usually defined in terms of the flow [27]. Choosing $\mathrm{GL}(N,\mathbb{R})$, which is a subgroup of $\mathrm{Diff}(\mathbb{R}^N)$, and co-restricting the flow map to $\mathfrak{gl}(N,\mathbb{R})$, we get the usual definition of the Lie algebra exponential as matrix power series.

3.3 Time dependent vector fields

For time dependent vector fields, the exponential has to be replaced by the time oredered exponential. A time dependent vector field is a map $X : \mathbb{R} \to \mathcal{X}(M)$. Denote the set of time dependent vector fields on M by $\mathcal{TX}(M)$. We define the flow map $\Phi^t : \mathcal{TX}(M) \to \mathrm{Diff}(M)$ by writing the solution to

$$\dot{x}(t) = X_t(x(t));\ x(0) = x_0 \tag{22}$$

as $\Phi^t_X(x_0) := x(t)$. We note the t-dependence of vector fields by subscript, X_t, and we note the t-dependence of flows by superscript, Φ^t_X. Let $\phi \in C^\infty(M)$. Then, by definition of the flow map,

$$\frac{d}{dt}\phi \circ \Phi^t_X = X_t \phi \circ \Phi^t_X. \tag{23}$$

For clarity, the RHS of this equation can be more unambiguously written $(X_t\phi)\circ\Phi^t_X$, As mentioned in the introduction, we can integrate eq. (23) by Picard iteration:

$$\begin{aligned}
\phi \circ \Phi^t_X &= \phi + \overline{X}_t\phi + \overline{\overline{X}X}_t\phi + \overline{\overline{\overline{X}X}X}_t\phi + \cdots \\
&= (1 + \overline{X}_t + \overline{\overline{X}X}_t + \overline{\overline{\overline{X}X}X}_t + \cdots)\phi \\
&= T^* \exp(\overline{X}_t)\phi := \mathrm{Ev}^t_X \phi.
\end{aligned} \tag{24}$$

The evolution operator satisfies

$$\frac{d}{dt}\mathrm{Ev}^t_X = \mathrm{Ev}^t_X X_t, \tag{25}$$

and is defined as the formal solution to this equation with initial condition $\mathrm{Ev}^0_X = 1$, the identity operator. Magnus's theorem states that we can find a time dependent vector field Ω such that $\mathrm{Ev}^t_X = \exp(\Omega_t)$. We present Magnus's theorem in a form which will be demonstrated in section 6.3.

Theorem 3.1. *Magnus/Dynkin. The solution to* $\dot{Y} = YX;\ Y(0) = 1$ *is* $Y(t) = \exp(\Omega_t(X))$, *where*

$$\begin{aligned}
\Omega_t(X) = \sum_{j=1}^{\infty} \frac{1}{j} \sum_{l=1}^{j} \frac{(-1)^{l+1}}{l} \sum_{n_1+\cdots+n_l=j} \int_0^t d\tau^{1,\dots,n_1} \int_0^t d\tau^{n_1+1,\dots,n_1+n_2} \cdots \int_0^t d\tau^{n_1+\cdots+n_{l-1}+1,\dots,j} \\
[\cdots[X_{\tau^j}, X_{\tau^{j-1}}], \dots, X_{\tau^1}],
\end{aligned} \tag{26}$$

where

$$\int_0^t d\tau^{i_1,\dots,i_n} := \int_0^t d\tau^{i_1} \int_0^{\tau^{i_1}} d\tau^{i_2} \cdots \int_0^{\tau^{i_{n-1}}} d\tau^{i_n}.$$

The first few terms of Ω are

$$\Omega(X) = \overline{X} + \frac{1}{2}[\overline{\overline{X}, X}] + \frac{1}{3}[\overline{[\overline{\overline{X}, X}], X}] - \frac{1}{6}\big([[\overline{X}, \overline{\overline{X}], X}] + [[\overline{\overline{X}, X}], \overline{X}]\big) + \cdots . \tag{27}$$

We will review proofs of this and other formulas in section 6, all rendering the same idea: We are trying to find a way to express the time ordered exponential as a true exponential. Due to the symmetry of the problem, we have the following

Corollary 3.2.

$$\begin{aligned}\mathrm{Ev}_X &= \exp(\Omega(X))\\ \mathrm{Ev}_f^{-X} &= \exp(-\Omega(X))\\ \mathrm{Ev}_{-X} &= \exp(\Omega(-X))\\ \mathrm{Ev}_{-X}^{-1} &= \exp(-\Omega(-X))\end{aligned}$$

Corollary 3.3. *The solution to $\dot{Y} = XY$ is $Y = e^{-\Omega(-X)}$. This simply reverses the order of the brackets, such that*

$$-\Omega_t(-X) = \sum_{j=1}^{\infty} \frac{1}{j} \sum_{l=1}^{j} \frac{(-1)^{l+1}}{l} \sum_{n_1+\cdots+n_l=j} \int_0^t d\tau^{1,\ldots,n_1} \int_0^t d\tau^{n_1+1,\ldots,n_1+n_2} \cdots \int_0^t d\tau^{n_1+\cdots n_{l-1}+1,\ldots,j}$$
$$[X_{\tau^1}, \ldots, [X_{\tau^{j-1}}, X_{\tau^j}] \cdots],$$

and the first few terms are

$$-\Omega(-X) = \overline{X} + \frac{1}{2}\overline{[X, \overline{X}]} + \frac{1}{3}\overline{[X, \overline{[X, \overline{X}]}]} - \frac{1}{6}\big(\overline{[X, \overline{[\overline{X}, X]}} + [\overline{X}, \overline{[X, \overline{X}]} \cdots]\big) + \cdots .$$

Proof. This follows from the fact that

$$-[\cdots[-X_{\tau^j}, -X_{\tau^{j-1}}], \ldots, -X_{\tau^1}] = (-1)^{j+1}[\cdots[X_{\tau^j}, X_{\tau^{j-1}}], \ldots, X_{\tau^1}] = [X_{\tau^1}, \ldots, [X_{\tau^j}, X_{\tau^{j-1}}]]$$

□

Lemma 3.4. *Let X and Y be time dependent vector fields. Then,*

$$\mathrm{Ev}_{X+Y} = \mathrm{Ev}_{\mathrm{Ad}_{\mathrm{Ev}_X} Y}\mathrm{Ev}_X,$$

where $\mathrm{Ad}_{\mathrm{Ev}_X} Y := \mathrm{Ev}_X Y \mathrm{Ev}_X^{-1}$

Proof. We will prove the lemma by showing that $\mathrm{Ev}_{\mathrm{Ad}_{\mathrm{Ev}_X} Y}\mathrm{Ev}_X$ satisfies the same IVP as Ev_{X+Y}. Applying the Leibnitz rule,

$$\frac{d}{dt}\mathrm{Ev}_{\mathrm{Ad}_{\mathrm{Ev}_X} Y}\mathrm{Ev}_X = \frac{d}{dt}(\mathrm{Ev}_{\mathrm{Ad}_{\mathrm{Ev}_X} Y})\mathrm{Ev}_X + \mathrm{Ev}_{\mathrm{Ad}_{\mathrm{Ev}_X} Y}\frac{d}{dt}\mathrm{Ev}_X.$$

By definition,

$$\frac{d}{dt}(\mathrm{Ev}_{\mathrm{Ad}_{\mathrm{Ev}_X} Y}) = \mathrm{Ev}_{\mathrm{Ad}_{\mathrm{Ev}_X} Y}\mathrm{Ad}_{\mathrm{Ev}_X} Y,$$

and

$$\frac{d}{dt}\mathrm{Ev}_X = \mathrm{Ev}_X X.$$

Inserting these into the equation above, we get

$$\frac{d}{dt}\mathrm{Ev}_{\mathrm{Ad}_{\mathrm{Ev}_X} Y}\mathrm{Ev}_X = \mathrm{Ev}_{\mathrm{Ad}_{\mathrm{Ev}_X} Y}\mathrm{Ev}_X Y + \mathrm{Ev}_{\mathrm{Ad}_{\mathrm{Ev}_X} Y}\mathrm{Ev}_X X = \mathrm{Ev}_{\mathrm{Ad}_{\mathrm{Ev}_X} Y}\mathrm{Ev}_X(X+Y).$$

We thus have that $\mathrm{Ev}_{\mathrm{Ad}_{\mathrm{Ev}_X} Y}\mathrm{Ev}_X$ satisfies the same IVP as Ev_{X+Y}, which is

$$\frac{d}{dt}\mathrm{Ev}_{X+Y} = \mathrm{Ev}_{X+Y}(X+Y).\ \mathrm{Ev}^0_{X+Y} = 1,$$

□

As a consequence of Lemma 3.4,

$$\mathrm{Ev}_{\mathrm{Ad}_{\mathrm{Ev}_X} Y} = \mathrm{Ev}_{X+Y}\mathrm{Ev}_X^{-1},$$

and so, using Magnus's theorem,

$$\exp(\Omega(\mathrm{Ad}_{\mathrm{Ev}_X} Y)) = \exp(\Omega(X+Y))\exp(-\Omega(X)),$$

and taking the logarithm gives

Proposition 3.5.

$$\Omega(\mathrm{Ad}_{\mathrm{Ev}_X} Y) = \mathrm{BCH}(\Omega(X+Y), -\Omega(X)),$$

where

$$\mathrm{BCH}(A, B) = \log(\exp(A)\exp(B)).$$

Lastly, we derive an expression that will be useful in the next section. We present it as a lemma:

Lemma 3.6.

$$\mathrm{Ad}_{(\mathrm{Ev}_{-X})^{-1}} Y = Y + [\overline{X}, Y] + [\overline{X, [\overline{X}}, Y]] + [\overline{X, [\overline{X, [\overline{X}}}, Y]]] + \cdots,$$

Proof. Recall that by definition, the operator Ev_X satisfies

$$\frac{d}{dt}\mathrm{Ev}_X = \mathrm{Ev}_X X.$$

Composing Ev_X with its inverse must yield the identity:

$$(\mathrm{Ev}_X)^{-1}\mathrm{Ev}_X = 1.$$

Differentiating this expression with respect to time must then yield zero

$$\frac{d}{dt}((\mathrm{Ev}_X)^{-1})\mathrm{Ev}_X + (\mathrm{Ev}_X)^{-1}\frac{d}{dt}\mathrm{Ev}_X = 0.$$

Rearranging this equation, we find that $(\mathrm{Ev}_X)^{-1}$ satisfies

$$\frac{d}{dt}((\mathrm{Ev}_X)^{-1}) = -X(\mathrm{Ev}_X)^{-1}.$$

Ad is defined by the composition of operators

$$\mathrm{Ad}_A B := ABA^{-1}.$$

Using this, we have

$$\begin{aligned}
\frac{d}{d\tau}\mathrm{Ad}_{(\mathrm{Ev}^\tau_{-X})^{-1}} Y &= \frac{d}{d\tau}\big((\mathrm{Ev}^\tau_{-X})^{-1} Y \mathrm{Ev}^\tau_{-X}\big) \\
&= X_\tau(\mathrm{Ev}^\tau_{-X})^{-1} Y \mathrm{Ev}^\tau_{-X} + (\mathrm{Ev}^\tau_{-X})^{-1} Y \mathrm{Ev}^\tau_{-X}(-X) \\
&= [X_\tau, \mathrm{Ad}_{(\mathrm{Ev}^\tau_{-X})^{-1}} Y],
\end{aligned}$$

Using the fundamental theorem of Calculus and the above result, we have

$$\mathrm{Ad}_{(\mathrm{Ev}^t_{-X})^{-1}} Y = Y + \int_0^t d\tau \frac{d}{d\tau}\mathrm{Ad}_{(\mathrm{Ev}^\tau_{-X})^{-1}} Y = Y + \int_0^t d\tau [X_\tau, \mathrm{Ad}_{(\mathrm{Ev}^\tau_{-X})^{-1}} Y].$$

This is a recursive equation for $\mathrm{Ad}_{(\mathrm{Ev}^t_{-X})^{-1}} Y$. Iterating it yields the infinite series in the lemma. $\square$

The idea of Lemma 3.6 can be used to derive an identity that is useful in quantum mechanics. Recall that the evolution operator in quantum mechanics is defined as the solution to

$$\dot{U}(t) = H(t)U(t),\ U(0) = 1,$$

where $H(t)$ is anti-Hermitian. This makes U a unitary operator. A quantum state $|\psi(t)\rangle$ whose dynamics is governed by the Schrödinger equation,

$$\frac{d}{dt}\psi(t)\rangle = H(t)|\psi(t)\rangle,$$

can be given in terms of the evolution operator as $|\psi(t)\rangle = U(t)|\psi_0\rangle$, where $|\psi_0\rangle$ is the initial state of the system. We suppose the reader familiar with the terms "Schrödinger picture" and "Heisenberg picture" in this context. Let A_S be an operator in the Scrödinger picture. It's Heisenberg picture equivalent is

$$A_H := U A_S U^{-1}.$$

By doing the same trick as in the lemma,

$$A_H(t) = A_S(t) + \int_0^t d\tau \frac{d}{d\tau} U(\tau) A_S(t) U^{-1}(\tau) = A_S(t) + \int_0^t d\tau [H(\tau), U(\tau) A_S(t) U^{-1}(\tau)].$$

Iterating this equation gives

$$A_H = A_S + [\overline{H}, A_S] + [\overline{H, [\overline{H}}, A_S]] + [\overline{H, [\overline{H, [\overline{H}}}, A_S]]] + \cdots .$$

A truncation of this expansion gives an approximation for A_H.

4 Radford's theorem

In this section, We present the result from Radford & Burdick's paper [45] as a theorem, which we will call *Radford's Teorem*. We will show that it can be derived as a corollary of Proposition 3.5.

Theorem 4.1 (Radford)**.** *Let G be a Lie group and M a manifold, and (W, x) a chart on M. Let $u : [0, 1] \to W$ be a path in $W \subset M$ such that $u(1) = u(0)$. Let $A : TM \to \mathrm{Lie}\, G$ be an analytic connection one-form, that is, $A(\dot{u}(t)) = \dot{u}^i(t) A_i(u(t))$, where $A_i : W \to \mathrm{Lie}\, G$ is analytic.*[1] *Then, the solution to*

$$\dot{g} = -g A_i(u) \dot{u}^i$$

is given by

$$g(t) = g(0) \exp\Bigg(\sum_{j=2}^{\infty} \frac{1}{j} \sum_{l=1}^{\lfloor j/2 \rfloor} \frac{(-1)^l}{l} \sum_{|I_1| + \cdots + |I_l| = j} \int_0^t du^{I_1} \cdots \int_0^t du^{I_l} \nabla F_{I_1 \cdots I_l}(u(0)) \Bigg), \tag{28}$$

where $F_{ij} = \partial_i A_j - \partial_j A_i - [A_i, A_j]$, $\nabla_i = [A_i, \cdot] - \partial_i$, $\partial_i := \partial / \partial x^i$, $\nabla F_{i_1 \cdots i_n} := \nabla_{i_1} \cdots \nabla_{i_{n-2}} F_{i_{n-1} i_n}$ and I_i is a multi-index of size $|I_i|$. $\lfloor j/2 \rfloor$ denotes the largest integer less than or equal to $j/2$

The first few terms in the expansion above are

$$g(t) = g(0) \exp\Bigg(-\frac{1}{2} F_{ij}(u(0)) \int_0^t du^{ij} - \frac{1}{3} \nabla_i F_{jk}(u(0)) \int_0^t du^{ijk} + \cdots \Bigg),$$

This is the main theorem we wish to prove. It turns out that it can be derived from Proposition 3.5, and that it is possible to generalize the theorem to non-closed loops. These new results are what we will show in the following.

We will do this in steps. Proposition 3.5 is a general statement, and the equality holds for all vector fields, as long as the quantities involved are well defined. The vector fields we will put into the formula given in the proposition will be elements of $\mathcal{TX}(M \times G)$, that is, time dependent vector fields on $M \times G$. The functions involved in Theorem 4.1 are $u : [0, 1] \to W \subset M$, and $A : TM \to \mathrm{Lie}\, G$. We will construct some time dependent vector fields $U, LA \in \mathcal{TX}(M \times G)$ out of u and A respectively, put these into the formula in Propostion 3.5, and show that this gives a generalized identity that holds for non-closed loops u. We next show that Radford's theorem follows from this, and finally, we consider the special case of zero curvature, that is, if $F = 0$.

I. Constructing U and LA

Suppose $\dim M = d$. We shall assume for simplicity of the argument and without loss of generality that $W = M$, that is, that the chart (W, x) covers M. We can construct a time dependent vector

[1] We require A_i to be analytic in a neighborhood containing the path u.

field U out of u in the following way. Let $\{\partial_i\}_{i=1}^d$ be the coordinate frame induced by the chart (W, x) on M. These can be extended to define vector fields on $M \times G$ in a canonical way. Recall from differential geometry, that

$$T(M \times G) \cong TM \oplus TG.$$

This means that there is a canonical injection $i_M : TM \hookrightarrow TM \oplus TW$, such that $i_M \circ \partial_i : M \to TM \oplus TG$. Extend this to define a function on $M \times G$ by composing with the projection $\pi_M : M \times G \to M,\ (x, g) \mapsto x$. Then,

$$\tilde{\partial}_i := i_M \circ \partial_i \circ \pi_M : M \times G \to TM \oplus TG \cong T(M \times G)$$

defines a vector field on $M \times G$. We do this to all the basis vectors ∂_i in the frame $\{\partial_i\}$, and define the linear combination $U_t := \dot{u}^i(t)\tilde{\partial}_i$. U is then a time dependent vector field on $M \times G$. An important fact about this vector field, that we can note straight away, is that

$$\forall \phi \in C^\infty(M \times G) : U_t\phi(u(t), \cdot) = \dot{u}^i(t)\tilde{\partial}_i\phi(u(t), \cdot) = \frac{d}{dt}\phi(u(t), \cdot),$$

where $\phi(u(t), \cdot)$ denotes the map $G \to \mathbb{R},\ g \mapsto \phi(u(t), g)$.

In Theorem 4.1, $A : TM \to \text{Lie } G$ was defined as a Lie algebra valued one-form. This means that it can be written as the linear combination

$$A = A_i dx^i : TM \to \text{Lie } G,$$

where dx^i are the unit one-forms associated with the frame $\{\partial_i\}$. These are defined precisely by their action on the ∂_i's: $dx^i(\partial_j) := \delta^i_j$, where δ^i_j is the Kronecker delta symbol.

The linear combination $A_i dx^i$ is dependent on the chart x and so, is really only defined on the open subset W where x is defined. As mentioned before, we assume $M = W$. This is without loss of generality, since the theorem supposes the path u to be contained in W. A coordinate domain is by definition contractible, and it is a fact that a tangent bundle TM where M is contractible is trivializable [26], meaning $TM \cong M \times \mathbb{R}^d$, where $\mathbb{R}^d$ parametrizes the tangent vectors. Precisely,

$$\begin{aligned} A_i dx^i : M \times \mathbb{R}^d &\to \text{Lie } G, \\ (p, V) &\mapsto A_i(p)V^i. \end{aligned}$$

Now, we can fix the second coordinate of this map by evaluating it at $u(t)$:

$$\begin{aligned} A_i \dot{u}^i(t) : M &\to \text{Lie } G, \\ p &\mapsto A_i(p)\dot{u}^i(t). \end{aligned}$$

By composing this map with π_M, as we did before, we get a map

$$(A_i \dot{u}^i(t)) \circ \pi_M : M \times G \to \text{Lie } G. \tag{29}$$

We wish to extend this to define a vector field on all of $M \times G$.

Let $\xi \in \text{Lie } G$. We will define a map $L : \text{Lie } G \to \mathcal{X}(G)$, such that ξ can be identified with a vector field on G. Let

$$L(\xi)(g) := dg_1\xi,$$

where dg_1 is the differential of the map $G \to G : h \mapsto gh$ evaluated at 1. The map L is a Lie isomorphism with respect to the Lie-Jacobi bracket. We elaborate on the map L in Appendix A.2. If we are working with matrix Lie groups, the application $\xi \mapsto dg_1\xi$ is simply the left matrix product $g\xi$, and so we shall simply denote $dg_1\xi \equiv g\xi$. If a vector field $X \in \mathcal{X}$ lies in the image of L, we call it *left-invariant*. The inverse of L, taking a left-invariant vector field and giving the corresponding Lie element, simply amounts to evaluating the vector field at $1 \in G$: $L^{-1}X = X(1) \in \text{Lie } G$.

We can compose the map in eq. (29) with L and then the injection $i_G : TG \hookrightarrow TM \oplus TG \cong T(M \times G)$ to get a vector field defined on $M \times G$, which we will call LA:

$$\begin{aligned} LA : & M \times G \to T(M \times G), \\ & (x, g) \mapsto i_G(L(A_i(x)\dot{u}^i(t))(g)). \end{aligned}$$

We have defined time dependent vector fields U and LA in $M \times G$. We will put these into the formula in Proposition 3.5, and show that it produces the result of the theorem. However, first we need to show that this indeed gives the solution to the ODE defined in the theorem.

II. Showing that $L\Omega_t(-A_i(u)\dot{u}^i) = \Omega(-\mathrm{Ad}_{\mathrm{Ev}_U} LA(u(0), \cdot))$

By Theorem 3.1, we wish to find $\Omega_t(-A_i(u)\dot{u}^i)$. This gives the solution to the ODE defined in Theorem 4.1 by

$$g(t) = g(0)\exp(\Omega_t(-A_i(u)\dot{u}^i)).$$

Here,

$$\begin{aligned} -A_i(u)\dot{u}^i : & [0, 1] \to \mathrm{Lie}\, G \\ & t \mapsto -A_i(u(t))\dot{u}^i(t). \end{aligned}$$

It is worth noting for later reference that since $L : \mathrm{Lie}\, G \to \mathcal{X}(G)$ is a Lie isomorphism when restricted to its image, and since Ω is a Lie element,

$$L\Omega_t(-A_i(u)\dot{u}^i) = \Omega_t(L(-A_i(u)\dot{u}^i)),$$

or more intuitively stated, Ω commutes with L: $L \circ \Omega = \Omega \circ L$. Note that the Ω on the LHS of the equality above consists of iterated brackets of Lie elements, and that the Ω on the RHS consists of iterated brackets of left-invariant vector fields. It is safe to suppress this distinction since L is a Lie isomorphism onto its image, and so the identification is unambiguous.

Recall from eq. (14) that we can express the Taylor expansion of $A_i(\cdot)\dot{u}^i(t) : M \to \mathrm{Lie}\, G$ in terms of iterated integrals as

$$A_k(u(t))\dot{u}^k(t) = A_k(u(0))\dot{u}^k(t) + \int_0^t du^i \partial_i A_k(u(0))\dot{u}^k(t) + \int_0^t du^{ij} \partial_i \partial_j A_k(u(0))\dot{u}^k(t) + \cdots. \quad (30)$$

We will show that this corresponds to the action of $\mathrm{Ad}_{\mathrm{Ev}_U^t}$ on LA.

Firstly, since U_t is constant with respect to coordinates, we have $LA_t U_{t'} = 0$, and so,

$$[U_{t'}, LA_t] = U_{t'} LA_t - LA_t U_{t'} = U_{t'} LA_t = \dot{u}^i(t')\partial_i(L\dot{u}^k(t)A_k(\cdot)) = L\dot{u}^i(t')\partial_i(\dot{u}^k(t)A_k(\cdot)).$$

We used the fact that L commutes with the action of the ∂_i's:

$$\forall f : M \to \mathrm{Lie}\, G : \forall p \in M : L\partial_i f(p) = \partial_i Lf(p),$$

where Lf is understood to mean the map $Lf : M \to \mathcal{X}(G) : p \mapsto L(f(p))$.

Now if we integrate $[U_{t'}, LA_t]$ with respect to t', we get

$$[\overline{U}_{t''}, LA_t] = \int_0^{t''} dt'[U_{t'}, LA_t] = L\int_0^{t''} dt' \dot{u}^i(t')\partial_i(\dot{u}^k(t)A_k(\cdot)) = L\int_0^{t''} du^i \partial_i A_k(\cdot)\dot{u}^k(t).$$

Taking the bracket again and integrating gives

$$[\overline{U[\overline{U}_{t'''}}, LA_t]] = \int_0^{t'''} dt'' \int_0^{t''} dt'[U_{t''}, [U_{t'}, LA_t]] = L\int_0^{t'''} du^{ij} \partial_i \partial_j A_k(\cdot)\dot{u}^k(t).$$

Repeating the same argument then gives the general formula

$$\overbrace{[U, \cdots [\overline{U}, [\overline{\overline{U}}_{t'}, LA_t] \cdots]}^{n \text{ times}} = L \int_0^{t'} du^{i_1 \cdots i_n} \partial_{i_1 \cdots i_n} A_k(\cdot) \dot{u}^k(t),$$

where $\partial_{i_1 \cdots i_n} := \partial_{i_1} \cdots \partial_{i_n}$. Since L commutes with the ∂'s, this is equal to

$$\overbrace{[U, \cdots [\overline{U}, [\overline{\overline{U}}_{t'}, LA_t] \cdots]}^{n \text{ times}} = \int_0^{t'} du^{i_1 \cdots i_n} \partial_{i_1 \cdots i_n} LA.$$

Thus, applying L to the Taylor expansion (30), and evaluating at $u(0)$, we have

$$LA_t(u(t), \cdot) = LA_t(u(0), \cdot) + [\overline{U}_t, LA_t](u(0), \cdot) + [\overline{U}, [\overline{\overline{U}}_t, LA_t]](u(0), \cdot) + \cdots . \tag{31}$$

By Lemma 3.6, this sum of iterated brackets is equal to

$$LA_t(u(t), \cdot) = \mathrm{Ad}_{(\mathrm{Ev}^t_{-U})^{-1}} LA_t(u(0), \cdot).$$

Since the components of U are constant with respect to coordinates, we have $[U_t, U_{t'}] = 0$, and as remarked in the introduction, this implies $(\mathrm{Ev}^t_{-U})^{-1} = \mathrm{Ev}^t_U = \exp(\overline{U}_t)$. Using this, we get

$$LA_t(u(t), \cdot) = \mathrm{Ad}_{\exp(\overline{U}_t)} LA_t(u(0), \cdot).$$

Recall that in this subsection, we wish to show that $L\Omega_t(-A_i(u)\dot{u}^i) = \Omega_t(-\mathrm{Ad}_{\mathrm{Ev}_U} LA(u(0), \cdot))$. As remarked before,

$$L\Omega_t(-A_i(u)\dot{u}^i) = \Omega_t(-L(A_i(u)\dot{u}^i)).$$

By definition of L and LA, $L(A_i(u)\dot{u}^i) = LA(u, \cdot)$, and so, we have

$$L\Omega_t(-A_i(u)\dot{u}^i) = \Omega(-\mathrm{Ad}_{\mathrm{Ev}_U} LA(u(0), \cdot)),$$

which is what we set out to prove.

III. Showing that $\Omega(-\mathrm{Ad}_{\mathrm{Ev}_U} LA(u(0), \cdot)) = \Omega(-\mathrm{Ad}_{\mathrm{Ev}_U} LA)(u(0), \cdot)$

Recall that LA is defined as the vector field

$$M \times G \ni (p, g) \mapsto gA_i(p)\dot{u}^i(t) \in T_gG \subset T_pM \oplus T_gG \cong T_{(p,g)}(M \times G).$$

This vector field is by definition tangent to G. We will now use the following fact:

Let P, Q be manifolds, and suppose $X, Y \in \mathcal{X}(P \times Q)$ are tangent to Q. Let $p \in P$ and denote $X(p, \cdot) : Q \to TQ$ the map $q \mapsto X(p, q)$. Then,

$$[X(p, \cdot), Y(p, \cdot)] = [X, Y](p, \cdot),$$

that is, it does not matter whether we fix the first coordinate before or after taking the bracket.

Now, since LA is tangent to G, we indeed have

$$\Omega(-\mathrm{Ad}_{\mathrm{Ev}_U} LA(u(0), \cdot)) = \Omega(-\mathrm{Ad}_{\mathrm{Ev}_U} LA)(u(0), \cdot),$$

as we wanted to show. The fact that LA is tangent to G implies that $\mathrm{Ad}_{\mathrm{Ev}^t_U} LA_t$ is also tangent to G.

IV. The new result

We have shown that

$$L\Omega_t(-A_i(u)\dot{u}^i) = \Omega_t(-\mathrm{Ad}_{\mathrm{Ev}_U} LA)(u(0), \cdot).$$

The RHS of this equation is now of a form such that we can invoke Proposition 3.5:

$$\Omega_t(-\mathrm{Ad}_{\mathrm{Ev}_U} LA)(u(0), \cdot) = \mathrm{BCH}(\Omega_t(-LA + U), -\Omega_t(U))(u(0), \cdot).$$

This expression then solves $\dot{g} = -gA_i(u)\dot{u}^i$:

$$\begin{aligned}\Omega_t(-A_i(u)\dot{u}^i) &= L^{-1}L\Omega_t(-A_i(u)\dot{u}^i) = L^{-1}\Omega_t(-\mathrm{Ad}_{\mathrm{Ev}_U} LA)(u(0), \cdot)\\ &= L^{-1}\mathrm{BCH}(\Omega_t(-LA + U), \Omega_t(U))(u(0), \cdot).\end{aligned}$$

We thus have

$$g(t) = g(0)\exp(L^{-1}\mathrm{BCH}(\Omega_t(-LA + U), -\Omega_t(U))(u(0), \cdot)).$$

The map L^{-1} amounts to the same as evaluating the left-invariant vector field at $1 \in G$, so

$$g(t) = g(0)\exp(\mathrm{BCH}(\Omega_t(-LA + U), -\Omega_t(U))(u(0), 1)). \tag{32}$$

This is the main result of this section. It is a generalized version Radford's Theorem that holds for all t, whereas the original theorem only held for $t = 1$ with the requirement $u(1) = u(0)$. We next show how to produce the terms in the expansion in the exponential above.

As remaked before, $\mathrm{Ev}_U = \exp(\overline{U})$, and so $\Omega(U) = \overline{U}$. The terms in $\Omega_t(-LA + U)$ can be found using the Magnus expansion, given in Theorem 3.1. We write out the first couple of terms:

$$\Omega_t(-LA + U) = \overline{-LA + U} + \frac{1}{2}\overline{\overline{[\overline{-LA + U}, -LA + U]}} + \frac{1}{3}\overline{[\overline{\overline{[\overline{-LA + U}, -LA + U]}}, -LA + U]} + \cdots .$$

Now, recall that $LA_t = L\dot{u}^i(t)A_i$. Since L is a Lie algebra isomorphism,

$$[LA_t, LA_{t'}] = L[A_t, A_{t'}] = L\dot{u}^i(t)\dot{u}^j(t')[A_i, A_j].$$

From before, we know that $[U_t, LA_{t'}] = L\dot{u}^i(t)\dot{u}^j(t')\partial_i A_j$, and $[U_t, U_{t'}] = 0$. Combining these facts, we have that

$$[-LA_{t'} + U_{t'}, -LA_t + U_t] = -L\dot{u}^i(t)\dot{u}^j(t')F_{ij},$$

where, as defined before, $F_{ij} = \partial_i A_j - \partial_j A_i - [A_i, A_j]$. Using the same argument, we get

$$[[-LA_{t''} + U_{t''}, -LA_{t'} + U_{t'}], -LA_t + U_t] = -L\dot{u}^i(t)\dot{u}^j(t')\dot{u}^k(t'')\nabla_i F_{jk},$$

where $\nabla_i = [A_i, \cdot] - \partial_i$, and it is straightforward to deduce that

$$[\cdots[-LA_{t^n} + U_{t^n}, -LA_{t^{n-1}} + U_{t^{n-1}}], ..., -LA_{t^1} + U_{t^1}] = -\dot{u}^{i_1}(t^1)\cdots\dot{u}^{i_n}(t^n)L\nabla_{i_1}\cdots\nabla_{i_{n-2}}F_{i_{n-1}i_n}.$$

Performing an iterated integral. we get

$$\overline{[\cdots\overline{\overline{[\overline{-LA + U}, -LA + U]}}, ..., -LA + U]}_t = -\int_0^t du^{i_1\cdots i_n}\nabla_{i_1}\cdots\nabla_{i_{n-2}}F_{i_{n-1}i_n}.$$

Inserting these results into the expansion in Theorem 3.1, we have

$$\Omega_t(-LA+U) = \sum_{j=1}^{\infty}\frac{1}{j}\sum_{l=1}^{j}\frac{(-1)^l}{l}\sum_{|I_1|+\cdots+|I_l|=j}\int_0^t du^{I_1}\cdots\int_0^t du^{I_l}\nabla F_{I_1\cdots I_l}, \tag{33}$$

where $\nabla F_{i_1\cdots i_n} := L\nabla_{i_1}\cdots\nabla_{i_{n-2}}F_{i_{n-1}i_n}$, $\nabla F_i := -LA_i + \partial_i$, I_i is a multi-index of size $|I_i|$.

The first few terms of $\Omega_t(-LA+U)$ are

$$\begin{aligned}\Omega_t(-LA+U) = (-LA_i+\partial_i)\int_0^t du^i - \frac{1}{2}LF_{ij}\int_0^t du^{ij} \\ -\frac{1}{3}L\nabla_i F_{jk}\Big(\int_0^t du^{ijk} - \frac{1}{2}\Big(\int_0^t du^i\int_0^t du^{jk} + \int_0^t dx^{ij}\int_0^t du^k\Big)\Big) + \cdots .\end{aligned}$$

Recall that $\Omega_t(-LA+U)$ and $\overline{U}$ are defined as vector fields on $M\times G$. Thus, $\mathrm{BCH}(\Omega_t(-LA+U), -\overline{U})$ is also a vector field on $M\times G$. We can evaluate the first few terms in this expansion by using the expression (17) for BCH. The first few terms are

$$\begin{aligned}\mathrm{BCH}(\Omega_t(-LA+U), -\overline{U}_t) = LA_i\int_0^t du^i - \frac{1}{2}LF_{ij}\int_0^t du^{ij} + \frac{1}{2}L\partial_i A_j\int_0^t du^{i\sqcup j} \\ -\frac{1}{3}L\nabla_i F_{jk}\int_0^t du^{ijk} - \frac{1}{2}L\partial_i F_{jk}\int_0^t du^{i\sqcup jk} + \cdots .\end{aligned} \tag{34}$$

We denote the product of two iterated integrals by

$$\int_0^t du^{I\sqcup J} := \int_0^t du^I\int_0^t du^J,$$

where I and J are multi-indices. Notice that all the terms in eq. (34) are in the image of L, such that

$$L^{-1}\mathrm{BCH}(\Omega_t(-LA+U), -\overline{U}_t)(u(0),\cdot) = \mathrm{BCH}(\Omega_t(-LA+U), -\overline{U}_t)(u(0),1)$$

is an element of the Lie algebra, as we would expect for eq. (32) to make sense:

$$\begin{aligned}\mathrm{BCH}(\Omega_t(-LA+U), -\overline{U}_t)(u(0),1) = A_i(u(0))\int_0^t du^i - \frac{1}{2}F_{ij}(u(0))\int_0^t du^{ij} + \frac{1}{2}\partial_i A_j(u(0))\int_0^t du^{i\sqcup j} \\ -\frac{1}{3}\nabla_i F_{jk}(u(0))\int_0^t du^{ijk} - \frac{1}{2}\partial_i F_{jk}(u(0))\int_0^t du^{i\sqcup jk} + \cdots .\end{aligned} \tag{35}$$

V. The special case $\overline{U}=0$ implies Radford's theorem

If we assume that $u(1)=u(0)$, we have $\overline{U}_1 = 0$, and so only the first term in the BCH expansion survives:

$$\mathrm{BCH}(\Omega_1(-LA+U), -\overline{U}_1) = \Omega_1(-LA+U),$$

such that

$$g(t) = g(0)\exp(\Omega_1(-LA+U)(u(0),1)). \tag{36}$$

We can use eq. (33) to evaluate $\Omega_1(-LA+U)(u(0),1)$:

$$\Omega_1(-LA+U)(u(0),1) = \sum_{j=2}^{\infty}\frac{1}{j}\sum_{l=1}^{\lfloor j/2\rfloor}\frac{(-1)^l}{l}\sum_{|I_1|+\cdots+|I_l|=j}\int_0^1 du^{I_1}\cdots\int_0^1 du^{I_l}\nabla F_{I_1\cdots I_l}(u(0)), \tag{37}$$

which is indeed Radford's expansion as given in Theorem 4.1. We have used the fact that $\int_0^1 du^i = 0$ to reduce the number of terms in the sum over l. The first few terms of eq. (37) are

$$\begin{aligned}\Omega_1(-LA+U)(u(0),1) = &-\frac{1}{2}F_{ij}(u(0))\int_0^1 du^{ij} - \frac{1}{3}\nabla_i F_{jk}(u(0))\int_0^1 du^{ijk} \\ &-\frac{1}{4}\nabla_i\nabla_j F_{kl}(u(0))\Big(\int_0^1 du^{ijkl} - \frac{1}{2}\int_0^1 du^{ij}\int_0^1 du^{kl}\Big) + \cdots .\end{aligned} \tag{38}$$

VI. What if $F = 0$?

The expression for $\Omega_t(-LA+U)$, given in eq. (33) depends linearly on F to all orders. If $F = 0$, then all higher order terms vanish, and we are left with

$$\Omega_t(-LA+U) = -\overline{LA}_t + \overline{U}_t.$$

This provides a corollary to our generalized Radford theorem: If $F = 0$, then the solution to $\dot{g} = -gA_i(u)\dot{u}^i$ is given by

$$g(t) = g(0)\exp(\mathrm{BCH}(-\overline{LA}_t + \overline{U}_t), -\overline{U}_t)(u(0),1)). \tag{39}$$

As will be elaborated upon in the next section, this is a path-independent expression: It only depends on the endpoints $u(t)$ and $u(0)$ of the path u. This means that if curvature vanishes, then the solution to $\dot{g} = -gA_i(u)\dot{u}^i$ is path independent, consistent with the fact that for a manifold with vanishing curvature, the holonomy elements corresponding to two homotopic paths are the same.

5 Magnus expansion for vanishing curvature

In this section we provide some notes on the special case of $F = 0$ for the main result of the previous section.

Let G be a Lie group, and suppose we have the set of n PDEs

$$\frac{\partial}{\partial x^i} g = gA_i,\ i = 1, ..., n, \tag{40}$$

where $i = 1, \cdots, n$, $g : \mathbb{R}^n \to G$ and $A_i : \mathbb{R}^n \to \mathrm{Lie}\, G$. This set of PDEs comes up in several applications, notably in evaluating master integrals [21] and in the teory of moving frames in variational calculus [58]. Let u be a path in $\mathbb{R}^n$. Then, multiplying by $\dot{u}^i$, summing over i and evaluating at $x = u(t)$ gives

$$\frac{d}{dt}g(u(t)) = \dot{u}^i(t)\frac{\partial}{\partial x^i}g(u(t)) = g(u(t))\dot{u}^i(t)A_i(u(t)) =: g(u(t))A(u(t)), \tag{41}$$

an ODE which is typically solved in practice using the Magnus expansion. Using eq. (40), we have

$$0 = (\partial_i\partial_j - \partial_j\partial_i)g = \partial_i A_j - \partial_j A_i - [A_i, A_j] = F_{ij}. \tag{42}$$

Since F vanishes, Radford's expansion simplifies to eq. (39), giving a path independent expression.

Let $\Delta u(t) = u(t) - u(0)$ and

$$Y(A,B) := \mathrm{BCH}(A+B, -B) = A - \frac{1}{2}[A,B] - \frac{1}{12}[A,[A,B]] - \frac{1}{6}[B,[A,B]] + \cdots . \tag{43}$$

We can use eq. (39) to obtain an explicit expression for $g(t)$. Using $\overline{U} = \Delta u^i \partial_i$ and $\overline{A} = \Delta u^i A_i$, we have (39).

$$g(u(t)) = g(0) \exp(\mathrm{BCH}(\overline{A}_t + \overline{U}_t, -\overline{U}_t)(x_0) = g(0) \exp(Y(\overline{A}_t, \overline{U}_t)(x_0)). \tag{44}$$

Here, we have taken the liberty of dropping the L in eq. (39). It is understood that combinations like $[\partial_i, A_j]$ mean $\partial_i A_j$, while $[A_i, A_j]$ is the Lie algebra bracket of the two lie elements A_i and A_j. The expansion obtained this way is then evaluated at x_0:

$$\begin{aligned} g(u(t)) =& g(0) \exp\big(Y(\Delta u^i(t) A_i, \Delta u^j(t) \partial_j)(x_0)\big) \\ =& g(0) \exp\big(\Delta u^i(t) A_i(x_0) + \frac{1}{2} \Delta u^i(t) \Delta u^j(t) \partial_j A_i(x_0) \\ &+ \frac{1}{6} \Delta u^i(t) \Delta u^j(t) \Delta u^k(t) \big(\frac{1}{2}[A_i, \partial_j A_k](x_0) + \partial_i \partial_j A_k(x_0)\big) + \cdots \big). \end{aligned} \tag{45}$$

The expansion in the exponential above depends only on the endpoints of the path u, and so makes g path independent. If we choose a path $u : [0,1] \to \mathbb{R}^n$ such that $u(1) = x$ and $u(0) = x_0$, then

$$g(x) = g(u(1)) = g(x_0) \exp\big(Y(A_i \Delta x^i, \Delta x^j \partial_j)(x_0)\big), \tag{46}$$

where $\Delta x^i = x^i - x_0^i$. The exponent in eq. (46) is an infinite sum of iterated brackets. For instance, the second term is

$$-\frac{1}{2}[A_i \Delta x^i, \Delta x^i \partial_i](x_0) = \frac{1}{2} \Delta x^i \Delta x^j \partial_i A_j(x_0).$$

At this point, a sanity check is in order. Suppose we are faced with the ODE

$$\dot{g}(t) = g(t) A(t),$$

where $A : \mathbb{R} \to \mathrm{Lie}\, G$, then we can regard this as a special case of eq. (40), using t as the only coordinate A and g depend on (i.e. setting $n = 1$ in eq. (40)). This amounts to substituting t for Δu^i and A' for $\partial_i A_j$ in (45). We get

$$g(t) = g(0) \exp(tA(0) + \frac{t^2}{2} A'(0) + \frac{t^3}{6} A''(0) + \frac{t^3}{12}[A(0), A'(0)] + \cdots) \tag{47}$$

The expression in the exponential should match the usual Magnus expansion when A is Taylor expanded around zero. We can check this by Taylor expanding A around zero, inserting it into the Magnus expansion and performing the integrals. Let

$$A(t) = A(0) + tA'(0) + \frac{t^2}{2} A''(0) + \cdots .$$

Then, inserting into the Magnus expansion and performing the integrals, we get

$$\begin{aligned} \Omega_t &= \overline{A}_t + \frac{1}{2}[\overline{\overline{A}, A}]_t + \frac{1}{3}[[\overline{\overline{\overline{A}, A}], A}]_t + \cdots \\ &= \int_0^t d\tau \bigg(A(0) + \tau A'(0) + \frac{\tau^2}{2} A''(0) + \cdots \bigg) \\ &+ \frac{1}{2} \int_0^t d\tau \int_0^\tau d\tau' \bigg[A(0) + \tau' A'(0) + \cdots, A(0) + \tau A'(0) + \cdots \bigg] + \cdots \\ &= tA(0) + \frac{t^2}{2} A'(0) + \frac{t^3}{6} A''(0) + \frac{1}{2}\big(\frac{t^3}{6}[A'(0), A(0)] + \frac{t^3}{3}[A(0), A'(0)]\big) \\ &= tA(0) + \frac{t^2}{2} A'(0) + \frac{t^3}{6} A''(0) + \frac{t^3}{12}[A(0), A'(0)] + \cdots , \end{aligned} \tag{48}$$

giving the same terms as eq. (47). This gives a neat expression for the Taylor expanded Magnus expansion:

$$\Omega_t = Y\big(tA(\tau), t\frac{d}{d\tau}\big|_{\tau=0}\big) \tag{49}$$

This can be derived independently from the result in the previous section, and so we present it as an independent theorem:

Theorem 5.1. *Let $A : \mathbb{R} \to \operatorname{Lie} G$ be analytic, and suppose $g : \mathbb{R} \to G$ is the solution to*

$$\dot{g}(t) = g(t)A(t). \tag{50}$$

Then, $g(t) = g(0)\exp(\Omega(t))$, where

$$\Omega(t) = Y\left(tA(\tau), t\frac{d}{d\tau}\Big|_{\tau=0}\right), \tag{51}$$

where $Y(A, B) = \mathrm{BCH}(A + B, -B)$.

The above formula means the following: Calculate $Y(tA(\tau), td/d\tau)$ using the BCH formula and $[d/d\tau, A] = A'$, and evaluate at $\tau = 0$. This produces up to fourth order in t:

$$g(t) = g_0 \exp\big(tA(0) + \frac{t^2}{2}A'(0) + \frac{t^3}{12}[A, A'](0) + \frac{t^3}{6}A''(0) + \frac{t^4}{24}[A, A''](0) + \frac{t^4}{24}A'''(0) + \cdots\big)$$

Theorem 5.1 is a consequence of eq. (39), however it is possible to prove it directly:

Proof. Let $X := A + d/d\tau$ be the vector field on $\mathbb{R} \times G$ defined by

$$X(\tau, g) = gA(\tau) + \frac{d}{d\tau} = LA(\tau, g) + \frac{d}{d\tau},$$

where $d/d\tau$ is the unit vector in the $\mathbb{R}$-direction in $\mathbb{R}\times G$, and $LA(\tau, g) := gA(\tau)$. Then X generates the solution to (50) in the sense that

$$\Phi^t_X(0, g_0) = (t, g(t)).$$

This assertion comes from the following lemma:

Lemma 5.2. *Let $u : [0, 1] \to \mathbb{R}^n$ be a path in $\mathbb{R}^n$ and suppose $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$. Then let $y : [0, 1] \to \mathbb{R}^m$ be the solution of the IVP in $\mathbb{R}^m$*

$$\dot{y}(t) = f(u(t), y(t)),\ y(0) = y_0. \tag{52}$$

Now, define the time dependent vector field $U_t := \dot{u}^i(t)\partial_i$, where $\partial_i,\ i = 1, ..., n$ are the first n unit vectors on $\mathbb{R}^n \times \mathbb{R}^m$. Then, let $(\tilde{x}, \tilde{y})$ be the solution of the IVP on $\mathbb{R}^n \times \mathbb{R}^m$

$$\begin{aligned} \dot{\tilde{y}}(t) &= f(\tilde{x}(t), \tilde{y}(t)), & \tilde{y}(0) &= y_0, \\ \dot{\tilde{x}}(t) &= U_t, & \tilde{x}(0) &= u(0). \end{aligned} \tag{53}$$

Then, $(\tilde{x}(t), \tilde{y}(t)) = (u(t), y(t))$.

Generalizing this to products of manifolds, gives the lemma used to obtain the above assertion. In the particular case of the above assertion, u is simply the identity map $\mathbb{R} \to \mathbb{R}$.

Let $\phi : \mathbb{R} \times G \to \mathbb{R}$. Then, by Taylor expanding,

$$\phi(t, g(t)) = \phi(\Phi^t_X(0, g_0)) = \exp(tX)\phi(0, g_0).$$

Using the BCH-formula on the operator $\exp(tX)$,

$$\exp(tX)\phi(0, g_0) = \exp(Y(tLA, t\frac{d}{d\tau}))\exp(t\frac{d}{d\tau})\phi(0, g_0) = \exp(LY(tA, t\frac{d}{d\tau}))\phi \circ (t \times \mathrm{id})(0, g_0),$$

where $t \times \mathrm{id} : \mathbb{R} \times G \to \mathbb{R} \times G$ denotes the map $(u, g) \mapsto (u + t, g)$. This is the flow map of the vector field $d/d\tau$. LA is a left invariant vector field, and so is $Y(tLA, td/d\tau) = LY(tA, td/d\tau)$. We will use a lemma from differential geometry:

Lemma 5.3. *Let $X \in \mathcal{X}(G)$ be a left-invariant vector field. Then, $\Phi^t_X(g) = g\exp(t\xi)$, where ξ is the Lie algebra element associated with X by the isomorphism L defined in the previous section and $\exp : \mathrm{Lie}\, G \to G$ is the Lie exponential.*

Using this fact, we have

$$\begin{aligned}\exp(LY(tA, t\frac{d}{d\tau}))\phi \circ (t \times \mathrm{id})(0, g_0) &= \phi \circ (t \times \mathrm{id}) \circ (\mathrm{id} \times \Phi^1_{LY(tA, t\frac{d}{d\tau})})(0, g_0) \\ &= \phi\big(t, g_0 \exp\big(Y\big(tA, t\frac{d}{d\tau}|_{\tau=0}\big)\big)\big).\end{aligned}$$

We thus have

$$g(t) = g_0 \exp\big(Y\big(tA, t\frac{d}{d\tau}|_{\tau=0}\big)\big), \tag{54}$$

where $g_0 = g(0)$. □

The above theorem is of great utility if an analytic expression for A is known. We immediately obtain a corollary of the above theorem.

Corollary 5.4. *Let $A : \mathbb{R} \to \mathrm{Lie}\, G$ be analytic, and suppose $\tilde{g} : \mathbb{R} \to G$ is the solution to*

$$\dot{\tilde{g}}(t) = A(t)\tilde{g}(t). \tag{55}$$

Then, $\tilde{g}(t) = \exp(\Omega(t))\tilde{g}(0)$, where

$$\Omega(t) = -Y\bigg(-tA(\tau), t\frac{d}{d\tau}\bigg|_{\tau=0}\bigg), \tag{56}$$

where $Y(A, B) = \mathrm{BCH}(A + B, -B)$.

The corollary produces up to fourth order in t:

$$g(t) = g_0 \exp\big(tA(0) + \frac{t^2}{2}A'(0) - \frac{t^3}{12}[A, A'](0) + \frac{t^3}{6}A''(0) - \frac{t^4}{24}[A, A''](0) + \frac{t^4}{24}A'''(0) + \cdots\big)$$

Proof. The proof mirrors that of the previous theorem, with minor change in details. Let $X := A + d/d\tau$ be the vector field on $\mathbb{R} \times G$ defined by

$$X(\tau, g) = A(\tau)g + \frac{d}{d\tau} = RA(\tau, g) + \frac{d}{d\tau},$$

where $d/d\tau$ is the unit vector in the $\mathbb{R}$ direction in $\mathbb{R} \times G$, and $RA(\tau, g) := A(\tau)g$. Then X generates the solution to (55) in the sense that

$$\Phi^t_X(0, g_0) = (t, g(t)).$$

This assertion is clear by the same reasoning as in the proof of the previous theorem. The main difference to proof of the previous theorem is that $R : \mathrm{Lie}\, G \to \mathcal{X}(G)$ is not a Lie isomorphism, rather, as known from differential geometry, and shown in appendix A, R satisfies

$$R[A, B]_{\mathrm{Lie}\, G} = -[RA, RB],$$

where $[\cdot, \cdot]_{\mathrm{Lie}\, G}$ is the Lie bracket of the Lie algebra of G and $[\cdot, \cdot]$ is the bracket of vector fields on G. It can be shown that

$$Y(RA, \frac{d}{d\tau}) = -RY(-A, \frac{d}{d\tau}).$$

Basically, R "flips" all the Lie brackets around in the same manner as was remarked in the proof of Corollary 3.2. We state this as a lemma:

Lemma 5.5. *Let $X_1, ..., X_n \in \text{Lie } G$, and suppose $P(X_1, ..., X_n)$ is a Lie polynomial in the Lie elements $X_1, ..., X_n$. By this, we mean that P consists of Lie brackets of X's. Then, $P(RX_1, ..., RX_n) = -RP(-X_1, ..., -X_n)$.*

It's straightforward to prove the above lemma by considering the action of R on monomials:

$$\begin{aligned}[RX_1, ..., [RX_{n-1}, RX_n]...] &= -[RX_1, ..., [RX_{n-2}, R[X_{n-1}, X_n]]...] \\ &= (-1)^{n-1} R[X_1, ..., [X_{n-1}, X_n]...] = -R[-X_1, ..., [-X_{n-1}, -X_n]...].\end{aligned}$$

As mentioned in Corollary 3.2, this is the same as flipping all brackets, which is precisely what happens when we use R.

Let $\phi : \mathbb{R} \times G \to \mathbb{R}$. Then, by Taylor expanding,

$$\phi(t, g(t)) = \phi(\Phi^t_X(0, g_0)) = \exp(tX)\phi(0, g_0).$$

Using the BCH-formula on the operator $\exp(tX)$,

$$\exp(tX)\phi(0, g_0) = \exp(Y(tRA, t\frac{d}{d\tau})) \exp(t\frac{d}{d\tau})\phi(0, g_0) = \exp(-RY(-tA, t\frac{d}{d\tau}))\phi \circ (t \times \text{id})(0, g_0),$$

Next we have a lemma on the flow of a right-invariant vector field:

Lemma 5.6. *Let $X \in \mathcal{X}(G)$ be a right-invariant vector field. Then, $\Phi^t_X(g) = \exp(t\xi)g$, where ξ is the Lie algebra element associated with X by the anti-isomorphism R and* $\exp : \text{Lie } G \to G$ *is the Lie exponential.*

Using this fact, we have

$$\begin{aligned}\exp(-RY(-tA, t\frac{d}{d\tau}))\phi \circ (t \times \text{id})(0, g_0) &= \phi \circ (t \times \text{id}) \circ (\text{id} \times \Phi^1_{-RY(-tA, t\frac{d}{d\tau})})(0, g_0) \\ &= \phi\big(t, \exp\big(-Y\big(-tA, t\frac{d}{d\tau}|_{\tau=0}\big)\big)g_0\big).\end{aligned}$$

We thus have

$$g(t) = \exp\big(-Y\big(-tA, t\frac{d}{d\tau}|_{\tau=0}\big)\big)g_0, \tag{57}$$

where $g_0 = g(0)$. □

For completeness, we restate and make explicit the main result of this section as a theorem:

Theorem 5.7. *The solution to the set of n PDEs*

$$\frac{\partial}{\partial x^i} g(x) = g(x)A_i(x),\ g(x_0) = g_0,\ i = 1, ..., n \tag{58}$$

is given by

$$g(x) = g(0)\exp\big(Y(\Delta x^i A_i(\xi), \Delta x^i \partial_i|_{\xi=x_0})\big), \tag{59}$$

where $Y(A, B) = \text{BCH}(A + B, -B)$ and $\Delta x = x - x_0$.

The solution to the set of n PDEs

$$\frac{\partial}{\partial x^i}\tilde{g} = A_i(x)\tilde{g},\ \tilde{g}(x_0) = \tilde{g}_0,\ i = 1, ..., n \tag{60}$$

is given by

$$\tilde{g}(x) = \exp\big(-Y(-\Delta x^i A_i(\xi), \Delta x^i \partial_i|_{\xi=x_0})\big)\tilde{g}(0) \tag{61}$$

We evaluate the expansion in (59) up to fourth order in Δx:

$$
\begin{aligned}
g(x) = g(x_0)\exp\big(&\Delta x^i A_i(x_0) + \frac{1}{2}\Delta x^i\Delta x^j\partial_i A_j(x_0) + \frac{1}{12}\Delta x^i\Delta x^j\Delta x^k[A_i,\partial_j A_k](x_0)\\
&+\frac{1}{6}\Delta x^i\Delta x^j\Delta x^k\partial_i\partial_j A_k(x_0) + \frac{1}{24}\Delta x^i\Delta x^j\Delta x^k\Delta x^l[A_i,\partial_j\partial_j A_k](x_0)\\
&+\frac{1}{24}\Delta x^i\Delta x^j\Delta x^k\Delta x^l\partial_i\partial_j\partial_k A_l(x_0)+\cdots\big)
\end{aligned}
$$

Evaluating the second expansion, (61) just flips the sign of some of the terms:

$$
\begin{aligned}
\tilde{g}(x) = \exp\big(&\Delta x^i A_i(x_0) + \frac{1}{2}\Delta x^i\Delta x^j\partial_i A_j(x_0) - \frac{1}{12}\Delta x^i\Delta x^j\Delta x^k[A_i,\partial_j A_k](x_0)\\
&+\frac{1}{6}\Delta x^i\Delta x^j\Delta x^k\partial_i\partial_j A_k(x_0) - \frac{1}{24}\Delta x^i\Delta x^j\Delta x^k\Delta x^l[A_i,\partial_j\partial_j A_k](x_0)\\
&+\frac{1}{24}\Delta x^i\Delta x^j\Delta x^k\Delta x^l\partial_i\partial_j\partial_k A_l(x_0)+\cdots\big)\tilde{g}(x_0)
\end{aligned}
$$

Proof. Suppose $x, x_0 \in \mathbb{R}^n$ and let $u : [0,1] \to \mathbb{R}^n$ be any path connecting $u(0) = x_0$ with $u(1) = x$. Let $\gamma : \mathbb{R} \to G$ be the solution to

$$\dot{\gamma} = \gamma\dot{u}^i A_i(u)$$

Then the solution to (58) is $g(x) = \gamma(1)$. This follows from the chain rule: $g \circ u$ satisfies the same IVP as γ, and so

$$g(x) = g(u(1)) = \gamma(1).$$

Since $F_{ij} = 0$, we can invoke the result from the previous section, (39). In this case,

$$\overline{LA}_t = L\Delta u^i(t)A_i,$$

and

$$\overline{U}_t = \Delta u^i(t)\partial_i.$$

Inserting this into eq. (39) gives a time dependent left-invariant vector field on $M\times G$. Evaluating at $(x_0, 1)$ gives the corresponding Lie algebra element, given in the exponential in eq. (45). Evaluating at $t = 1$, then gives $\gamma(1) = g(x)$, which by eq. (46) gives the result of the theorem.

The analogous formula for $\tilde{g}$ can be infered from that of g by using Corollary 3.2. □

We can deduce a nice corollary from the theorem above:

Corollary 5.8. *Let* $\mathrm{M}_{m\times m}(\mathbb{R})$ *denote the set of* $m \times m$ *real matrices. Suppose* $I : \mathbb{R}^n \to \mathbb{R}^m$ *satisfy*

$$\frac{\partial I^a}{\partial x^j} = [A_j(x)]^a_b I^b,\ I(x_0) = I_0,\ j = 1, ..., n, \tag{62}$$

where $A : \mathbb{R}^n \to \mathrm{M}_{m\times m}(\mathbb{R})$ *is analytic. Then, the solution is given by*

$$I(x) = \exp(-Y(-\Delta x^i A_i(\xi), \Delta x^j\partial/\partial\xi^j|_{\xi=x_0}))I_0, \tag{63}$$

where $\Delta x = x - x_0$.

Proof. Suppose $\tilde{g} : \mathbb{R}^n \to \mathrm{GL}(m, \mathbb{R})$ is the solution to

$$\frac{\partial\tilde{g}}{\partial x^j} = A_j(x)\tilde{g}(x),\ \tilde{g}(0) = 1.$$

Then, $I(x) = \tilde{g}(x)I_0$. From Theorem 5.7, we know that

$$\tilde{g} = \exp(-Y(-\Delta x^i A_i(\xi), \Delta x^j\partial/\partial\xi^j|_{\xi=x_0})),$$

which by $I(x) = \tilde{g}(x)I_0$ implies the result. □

Something particular to note about the solutions in Theorem 5.7 is that they provide analytic expressions for the exponent to arbitrary accuracy, as long as the derivatives of A are sufficiently well behaved. This creates new possibilities for analysis in problems where (59) has to be solved. To the best of my knowledge, the expressions provided here are new.

I do not claim to have proven the results of this section to absolute mathematical rigor, that would be beyond the scope of my abilities. I am however confident that these results are true, and that the details needed to complete the proofs amount to technicalities that can be overcome. The "sanity check", showing the validity of Theorem 5.1 up to third order can be done to arbitrary order on a computer. I have not yet had time to do this, however.

5.1 Zassenhaus expansion for vanishing curvature

By a similar idea to the previous section, we can produce solutions to (40) based on the Zassenhaus formula, which we introduce next. The Zassenhaus formula presents the exponential of the sum of two operators as an infinite product of exponentials.

The left Zassenhaus formula is given by [11]

$$\exp(t(X+Y)) = \cdots \exp(t^4 Z_4(X,Y)) \exp(t^3 Z_3(X,Y)) \exp(t^2 Z_2(X,Y)) \exp(tY) \exp(tX),$$

where

$$Z_n(X,Y) = \frac{1}{n!}\frac{d^n}{dt^n}\bigg|_{t=0} \exp(t(X+Y)) \exp(-tX) \exp(-tY) \exp(-t^2 Z_2(X,Y)) \cdots \exp(-t^{n-1} Z_{n-1}(X,Y)).$$

The terms constituting the analogous right Zassenhaus formula,

$$\exp(t(X+Y)) = \exp(tX) \exp(tY) \exp(t^2 \tilde{Z}_2(X,Y)) \exp(t^3 \tilde{Z}_3(X,Y)) \exp(t^4 \tilde{Z}_4(X,Y)) \cdots$$

are related to the Z_n's simply by $\tilde{Z}_n = (-1)^{n+1} Z_n$.

Let g be the solution of

$$\frac{\partial}{\partial x^i} g = g A_i(x), \; g(x_0) = 1. \tag{64}$$

We know from before that the solution to (64) is

$$\begin{aligned} g(x) &= \exp\left(\mathrm{BCH}(\Delta x^i A_i(\xi) + \Delta x^i \partial/\partial \xi^i|_{\xi=x_0}, -\Delta x^i \partial/\partial \xi^i|_{\xi=x_0})\right) \\ &= \exp(\Delta x^i A_i(\xi) + \Delta x^i \partial/\partial \xi^i|_{\xi=x_0}) \exp(-\Delta x^i \partial/\partial \xi^i|_{\xi=x_0}). \end{aligned} \tag{65}$$

Using the left Zassenhaus formula by substituting $\Delta x^i \partial/\partial \xi^i|_{\xi=x_0}$ for X and $\Delta x^i A_i(\xi)$ for Y, eq. (65) can be rewritten as

$$g(x) = \cdots \exp\left(Z_3(\Delta x^j \partial/\partial \xi^j\big|_{\xi=x_0}, \Delta x^i A_i(\xi))\right) \exp\left(Z_2(\Delta x^j \partial/\partial \xi^j\big|_{\xi=x_0}, \Delta x^i A_i(\xi))\right) \exp(\Delta x^i A_i(x_0)), \tag{66}$$

an infinite left product of exponentials of homogeneous polynomials in $\Delta x = x - x_0$.

The Zassenhaus elements Z_n can be worked out in a systematic fashion, as shown in a paper by Casas et al. [11]. The first few are

$$\begin{aligned} Z_2(A,B) &= \frac{1}{2}[A,B] \\ Z_3(A,B) &= \frac{1}{3}[B,[A,B]] + \frac{1}{6}[A,[A,B]] \\ Z_4(A,B) &= \frac{1}{8}([B,[B,[A,B]]] + [B,[A,[A,B]]]) + \frac{1}{24}[A,[A,[A,B]]] \end{aligned}$$

such that, we have explicitly to fourth order in Δx:

$$
\begin{aligned}
g(x) = \cdots \exp\big(\Delta x^i \Delta x^j \Delta x^k \Delta x^l \big(\frac{1}{8}([A_i,[A_j,\partial_k A_l]](x_0) + [A_i,\partial_j\partial_k A_l](x_0)) + \frac{1}{24}\partial_i\partial_j\partial_k A_l(x_0)\big)\big) \\
\exp\big(\Delta x^i \Delta x^j \Delta x^k \big(\frac{1}{3}[A_i,\partial_j A_k](x_0) + \frac{1}{6}\partial_i\partial_j A_k(x_0)\big)\big) \exp\big(\Delta x^i \Delta x^j \frac{1}{2}\partial_i A_j(x_0)\big) \\
\exp\big(\Delta x^i A_i(x_0)\big)
\end{aligned}
\tag{67}
$$

We can make analogous use of the right Zassenhaus formula to obtain a solution to

$$
\frac{\partial}{\partial x^i}\tilde{g} = A_i(x)\tilde{g};\ \tilde{g}(x_0) = 1. \tag{68}
$$

We know the solution to (68) to be

$$
\tilde{g}(x) = \exp(-\mathrm{BCH}(-\Delta x^i A_i(\xi) + \Delta x^i \partial/\partial\xi^i|_{\xi=x_0}, -\Delta x^i \partial/\partial \xi^i|_{\xi=x_0}).
$$

Making use of then fact that $\mathrm{BCH}(A,B) = -\mathrm{BCH}(-B,-A)$, we have

$$
\begin{aligned}
\tilde{g}(x) &= \exp(-\mathrm{BCH}(-\Delta x^i A_i(\xi) + \Delta x^i \partial/\partial\xi^i|_{\xi=x_0}, -\Delta x^i \partial/\partial \xi^i|_{\xi=x_0}) \\
&= \exp(\Delta x^i \partial/\partial\xi^i|_{\xi=x_0}) \exp(\Delta x^i A_i(\xi) - \Delta x^i \partial/\partial\xi^i|_{\xi=x_0}).
\end{aligned}
\tag{69}
$$

By using the right Zassenhaus expansion, we thus get

$$
\tilde{g}(x) = \exp(\Delta x^i A_i(x_0)) \exp(\tilde{Z}_2(-\Delta x^j \partial/\partial\xi^j\big|_{\xi=x_0}, \Delta x^i A_i(\xi))) \exp\big(\tilde{Z}_3(-\Delta x^j \partial/\partial\xi^j\big|_{\xi=x_0}, \Delta x^i A_i(\xi))\big) \cdots . \tag{70}
$$

It is possible to produce a Zassenhaus-like expansion for the solution of an ODE. Let g be the solution of

$$
\dot{g}(t) = g(t)A(t);\ g(0) = 1. \tag{71}
$$

Then, by the same idea as above,

$$
g(t) = \cdots \exp\big(Z_3(t\frac{d}{d\tau}|_{\tau=0}, tA(\tau))\big) \exp\big(Z_2(t\frac{d}{d\tau}|_{\tau=0}, tA(\tau))\big) \exp\big(tA(0)\big). \tag{72}
$$

Explicitly, to fourth order in t, this gives

$$
\begin{aligned}
g(t) = \cdots \exp\big(\frac{t^4}{8}([A,[A,A']](0) + [A,A''](0)) + \frac{t^4}{24}A'''\big) \\
\exp\big(\frac{t^3}{3}[A,A''](0) + \frac{t^3}{6}A''(0)\big) \exp(\frac{t^2}{2}A'(0)\big) \exp\big(tA\big) g_0,
\end{aligned}
\tag{73}
$$

For the problem $\dot{\tilde{g}} = A\tilde{g}$ there is a corresponding solution in terms of the right Zassenhaus formula.

The expressions (66), (70) and (72) are to my knowledge new. Again, I do not claim to have proven these expressions to absolute mathematical rigor.

6 Theorems of Magnus, Dynkin and Fer

In this section, we revert to the general problem defined in the introduction in eq, (1). This is a well studied problem, and most of the results presented in this section are known. I provide an exposition of different approaches to the problem, and supply some of my own insights along the way. As well as containing the proof of Theorem 3.1, the results presented here are important for the discussion in Sections 7 and 8.

Let G be a Lie group with Lie algebra Lie G. Let $A : \mathbb{R} \to \text{Lie } G$, then the IVP

$$\dot{Y}(t) = Y(t)A(t), \ Y(0) = 1 \tag{74}$$

defines a path $Y : \mathbb{R} \to G$. The equation (74) can be Picard iterated to yield a power series expression for Y. Magnus's theorem states that it is in general possible to find formal expressions for Ω such that $Y = \exp(\Omega)$ [37]. First, we present Magnus's theorem.

6.1 Magnus expansion

This exposition and proof sketch of Magnus's theorem is my own. My conceptual contribution here is that the Magnus expansion itself can be understood as the time ordered exponential of a vector field, as illustrated in eq (85), although it is unclear whether this method has any computational benefit.

By hypothesis, we are seeking Ω such that $\exp(\Omega) = Y$. In consistence with Theorem 3.1, we note $\Omega = \Omega(A)$, but omit explicitizing the dependence on A here. Inserting $\exp(\Omega) = Y$ into eq. 74,

$$\exp(-\Omega)\frac{d}{dt}\exp(\Omega) = A. \tag{75}$$

We will show that

$$\exp(-\Omega)\frac{d}{dt}\exp(\Omega) = \sum_{j=0}^{\infty}\frac{(-1)^j}{(j+1)!}\mathrm{ad}_\Omega^j(\dot{\Omega}) = \frac{\exp(-\mathrm{ad}_\Omega)-1}{\mathrm{ad}_\Omega}(\dot{\Omega}), \tag{76}$$

where $\mathrm{ad}_\Omega(\dot{\Omega}) = [\Omega, \dot{\Omega}]$, $\mathrm{ad}_\Omega^n = \mathrm{ad}_\Omega \circ ... \circ \mathrm{ad}_\Omega$ is the n-fold composition of ad_Ω. We sketch a proof of eq. (76). Following Snider [53], we have

$$\frac{d}{dt}\exp(\Omega) = \int_0^1 d\alpha \exp((1-\alpha)\Omega)\dot{\Omega}\exp(\alpha\Omega),$$

such that

$$\exp(-\Omega)\frac{d}{dt}\exp(\Omega) = \int_0^1 d\alpha \exp(-\alpha\Omega)\dot{\Omega}\exp(\alpha\Omega).$$

It can be shown that[1]

$$\exp(-\alpha\Omega)\dot{\Omega}\exp(\alpha\Omega) = \sum_{n=0}^{\infty}\frac{(-\alpha)^n}{n!}\mathrm{ad}_\Omega^n\dot{\Omega}, \tag{77}$$

such that

$$\exp(-\Omega)\frac{d}{dt}\exp(\Omega) = \sum_{j=0}^{\infty}\frac{(-1)^{j+1}}{(j+1)!}\mathrm{ad}_\Omega^j(\dot{\Omega}) = \frac{\exp(-\mathrm{ad}_\Omega)-1}{\mathrm{ad}_\Omega}(\dot{\Omega}), \tag{78}$$

Then, inverting eq. (76) gives

$$\dot{\Omega} = \frac{\mathrm{ad}_\Omega}{\exp(-\mathrm{ad}_\Omega)-1}(A) = \sum_{n=0}^{\infty}\frac{(-1)^n B_n}{n!}\mathrm{ad}_\Omega^n(A), \ \Omega(0) = 0, \tag{79}$$

[1] See B. Hall [18], pg. 64.

where B_n is the n-th Bernoulli number.[1] We thus have an IVP for Ω. Define

$$f_\tau(\Omega) := \frac{\mathrm{ad}_\Omega}{\exp(-\mathrm{ad}_\Omega) - 1}(A(\tau)) \tag{80}$$

such that

$$\dot{\Omega}(t) = f_t(\Omega(t)). \tag{81}$$

Now, $f_t : \mathrm{Lie}\, G \to \mathrm{Lie}\, G$, and if we interpret f as a time dependent vector field on $\mathrm{Lie}\, G$, we can define its differential action on Lie polynomials by

$$f_\tau p(\Omega) := \frac{d}{dt}\bigg|_{\dot{\Omega}=f_\tau(\Omega)} p(\Omega(t)). \tag{82}$$

Let p_A^n be the Lie monomial

$$p_A^n(\Omega) := \mathrm{ad}_\Omega^n(A) = \overbrace{[\Omega, ..., [\Omega}^{n \text{ times}}, A(\tau)]...].$$

Then the action of f_τ on p^n is by the Leibniz rule

$$f_\tau p_A^n(\Omega) = [f_\tau(\Omega), [\Omega, ..., [\Omega, A]...] + ... + [\Omega, ..., [f_\tau(\Omega), A]...]. \tag{83}$$

such that

$$f_{\tau'} f_\tau(\Omega) = \sum_{n=0}^{\infty} \frac{(-1)^n B_n}{n!} f_{\tau'} p_A^n(\Omega). \tag{84}$$

By the same idea as in eq. (24),

$$\begin{aligned}
\Omega(t) &= \int_0^t d\tau \dot{\Omega}(\tau) = \int_0^t d\tau f_\tau(\Omega(\tau)) \\
&\overset{(1)}{=} \int_0^t d\tau \big[f_\tau(0) + \int_0^\tau d\tau' \frac{d}{d\tau'} f_\tau(\Omega(\tau'))\big] \\
&\overset{(2)}{=} \int_0^t d\tau \big[f_\tau(0) + \int_0^\tau d\tau' f_{\tau'} f_\tau(\Omega(\tau'))\big],
\end{aligned}$$

where in equality (1), we used the fundamental theorem of calculus and $\Omega(0) = 0$. In equality (2), we used eq. (81). Repeating the iteration yields a power series in f which can be used to evaluate the Magnus expansion:

$$\begin{aligned}
\Omega(t) &= \big[1 + \overline{f} + \overline{\overline{f} f} + \overline{\overline{\overline{f} f} f} + ...\big](0) \\
&= \overline{f}(0) + \overline{\overline{f} f}(0) + \overline{\overline{\overline{f} f} f}(0) + ... \\
&= \overline{A} + \frac{1}{2}\overline{[\overline{A}, A]} + \frac{1}{4}\overline{[\overline{[\overline{A}, A]}, A]} + \frac{1}{12}\overline{[\overline{A}, \overline{[\overline{A}, A]}]} + \cdots .
\end{aligned} \tag{85}$$

[1] Here, B_n are the Bernoulli numbers, defined by

$$\frac{z}{\exp(z) - 1} = \sum_{n=0}^{\infty} \frac{B_n}{n!} z^n.$$

The first few Bernoulli numbers are

$$B_0 = 1,\ B_1 = -\frac{1}{2},\ B_2 = \frac{1}{6},\ B_3 = 0,\ B_4 = -\frac{1}{30},\ \text{etc.}$$

In particular, $B_{2n+1} = 0$ for $n = 1, 2, 3, ...$

Several schemes exist for effectively calculating the terms in the Magnus expansion. Notably, the work of Iserles & Nørsett provide a good example [23].

6.2 Fer expansion

Fer [14] showed that Y can be expressed as an infinite product of exponentials. Here, we expose a scheme for deriving this expansion, by inspiration of [8].

Suppose $\dot{Y} = YA$. Let $F_1 := \overline{A}$ and define $Y_1 \exp(F_1) := Y$. Then,

$$\begin{aligned} \dot{Y}_1 &= \dot{Y} \exp(-F_1) + Y \frac{d}{dt} \exp(-F_1) \\ &= Y_1 \big[\exp(F_1) A \exp(-F_1) + \exp(F_1) \frac{d}{dt} \exp(-F_1) \big] := Y_1 A_1 \end{aligned}$$

Using eqs. (77) and (78),

$$\begin{aligned} A_1 &= \sum_{n=0}^{\infty} \left[\frac{1}{n!} \mathrm{ad}_{F_1}^n A - \frac{1}{(n+1)!} \mathrm{ad}_{F_1}^n \dot{F}_1 \right] = \sum_{n=0}^{\infty} \left[\frac{1}{n!} - \frac{1}{(n+1)!} \right] \mathrm{ad}_{F_1}^n A \\ &= \sum_{n=1}^{\infty} \frac{n}{(n+1)!} \mathrm{ad}_{F_1}^n A \end{aligned} \tag{86}$$

Let $F_{n+1} := \overline{A}_n$ and $Y_n =: \exp(F_{n+1}) Y_{n+1}$. Then, define by the same procedure as above

$$A_{n+1} = \sum_{n=1}^{\infty} \frac{n}{(n+1)!} \mathrm{ad}_{F_{n+1}}^n A_n. \tag{87}$$

Fer showed that Y is expressible as the infinite product of exponentials

$$Y = \cdots \exp(F_3) \exp(F_2) \exp(F_1). \tag{88}$$

Suppose $\dot{\tilde{Y}} = A\tilde{Y}$, and let $\tilde{F}_{n+1} = \overline{\tilde{A}}_n$, where

$$\tilde{A}_{n+1} = \sum_{n=1}^{\infty} \frac{n(-1)^n}{(n+1)!} \mathrm{ad}_{\tilde{F}_{n+1}}^n \tilde{A}_n, \tag{89}$$

and $A_0 := A$. Then, as shown in [8], $\tilde{Y}$ can be expressed as the right infinite product of exponentials

$$\tilde{Y} = \exp(\tilde{F}_1) \exp(\tilde{F}_2) \exp(\tilde{F}_3) \cdots . \tag{90}$$

6.3 Dynkin expansion

The approach taken here is known in the literature, and is elaborated upon in [45] amongst other places. My contribution is the rewriting of the Dynkin expansion to the form of equation (94), which to my knowledge has not appeared explicitly before.

$\Omega(A)$ can be calculated directly by taking the logarithm of the power series produced by $\dot{Y} = YA$. Let

$$Y = 1 + \overline{A} + \overline{\overline{A}A} + \overline{\overline{\overline{A}A}A} + \cdots = \exp(\Omega(A)). \tag{91}$$

Y is seen to satisfy $\dot{Y} = YA$ by term-wise differentiation of (91). For ease of notation, $\top^n A := \overline{\overline{A} \cdots A}$ the n-fold left iterated integral. Taking the logarithm of Y,

$$\Omega = \log\left(1 + \sum_{n=1}^{\infty} \top^n A\right) = \sum_{l=1}^{\infty} \frac{(-1)^{l+1}}{l} \sum_{n_1,\ldots,n_l=1}^{\infty} (\top^{n_1} A) \cdots (\top^{n_l} A). \tag{92}$$

By Magnus's theorem, $\Omega(A)$ is a Lie element.[1]. Next, we make use of Dynkin's theorem [46]: Let P be a homogenous polynomial of Lie elements order n. Then, $(1/n)[P] = P$, where $[\,\cdot\,]$ is the left-iterated Lie bracket defined on monomials by $[X_1 \cdots X_n] = [\cdots[X_1, X_2], \cdots X_n]$, and extended linearly to Lie power series.

Making use of Dynkin's theorem, and the fact that Ω is a Lie element,

$$\Omega(A) = \sum_{l=1}^{\infty} \frac{(-1)^{l+1}}{l} \sum_{n_1,\ldots,n_l=1}^{\infty} \frac{[(\top^{n_1} A)...(\top^{n_l} A)]}{n_1 + ... + n_l}. \tag{93}$$

The sum can be rewritten as[2]

$$\Omega(A) = \sum_{j=1}^{\infty} \frac{1}{j} \sum_{l=1}^{j} \frac{(-1)^{l+1}}{l} \sum_{n_1+\ldots+n_l=j} [(\top^{n_1} A)...(\top^{n_l} A)] \tag{94}$$

The first few terms are

$$\Omega(A) = \overline{A} + \frac{1}{2}[\overline{\overline{A}, A}] + \frac{1}{3}[[\overline{\overline{\overline{A}, A}], A}] - \frac{1}{6}\big([[\overline{A}, \overline{\overline{A}], A}] + [[\overline{\overline{A}, A}], \overline{A}]\big) + \cdots . \tag{95}$$

This demonstrates Theorem 3.1. By Corollary 3.2, the solution to

$$\dot{\tilde{Y}} = A\tilde{Y}, \ \tilde{Y}(0) = 1,$$

is given by $\tilde{Y} = \exp(-\Omega(-A))$, which by Corollary 3.3 is given by

$$-\Omega(-A) = \sum_{j=1}^{\infty} \frac{1}{j} \sum_{l=1}^{j} \frac{(-1)^{l+1}}{l} \sum_{n_1+\ldots+n_l=j} [(\top_R^{n_1} A)...(\top_R^{n_l} A)]_R, \tag{96}$$

where $\top_R^n A$ is the right-iterated integral $\top_R^n A = \overline{A \cdots \overline{A}}$, and $[\,\cdot\,]_R$ is the right-iterated Lie bracket, defined by $[X_1 \cdots X_n]_R = [X_1, ..., [X_{n-1}, X_n]...]$. The first few terms in the expansion (96) are

$$-\Omega(-A) = \overline{A} + \frac{1}{2}[\overline{A, \overline{A}}] + \frac{1}{3}[\overline{A, [\overline{A, \overline{A}}]}] - \frac{1}{6}\big([\overline{A, [\overline{A}}, \overline{A}] + [\overline{A}, [\overline{A, \overline{A}}]...]\big) + \cdots .$$

6.4 Mielnik-Plebanski-Strichartz formula

It was shown by Mielnik and Plebanski [39], and later elaborated by Strichartz [56] that the Magnus expanison can be rewritten as,

$$\Omega = \sum_{l=1}^{\infty} \sum_{\pi \in S^l} \frac{\lambda(\pi)}{l!} \int d\tau^{1,\ldots,n} [A(\tau^{\pi(1)}) \cdots A(\tau^{\pi(n)})], \tag{97}$$

[1]The fact that $\Omega(A)$ is a Lie element can also be shown by using Friedrich's criterion [56]

[2]This is shown in section 3.1.4.1 in [10].

where S^l is the set of permutations of n symbols, and

$$\int^t d\tau^{1,...,n} := \int_0^t d\tau^1 \int_0^{\tau^1} d\tau^2 ... \int_0^{\tau^{n-1}} d\tau^n, \tag{98}$$

and

$$\frac{\lambda(\pi)}{l!} = \frac{(-1)^{d(\pi)}}{l^2 \binom{l-1}{d(\pi)}}, \tag{99}$$

where $d(\pi)$ is the cardinality of the descent set $D(\pi)$ defined by $D(\pi) := \{i \in (1, ..., l-1) |\ \pi(i+1) < \pi(i)\}$.

6.5 Wilcox expansion

Wilcox introduced in a 1967 paper [57] a method of expressing Y as an infinite product of exponentials, much like the method of Fer. Wilcox actually attributed his infinite product of exponentials to Fer, although it was in fact a novelty [25]. Here, we present Wilcox's product expansion.

If we scale A by a variable λ, such that $\dot{Y} = Y\lambda A$, then $\Omega(\lambda A)$ becomes a power series in λ:

$$\Omega(\lambda A) = \lambda\Omega_1(A) + \lambda^2\Omega_2(A) + \lambda^3\Omega_3(A) + ..., \tag{100}$$

where

$$\Omega_n(A) = \frac{1}{n!} \frac{d^n}{d\lambda^n}\bigg|_{\lambda=0} \log(Y). \tag{101}$$

Eq. (101) produces the terms in the Magnus expansion. In a similar idea to this, one can construct $S_1, S_2, S_3, ...$ such that

$$Y = \cdots \exp(\lambda^3 S_3) \exp(\lambda^2 S_2) \exp(\lambda S_1) \tag{102}$$

Then,

$$S_1 = \frac{d}{d\lambda}\bigg|_{\lambda=0} Y, \tag{103}$$

and

$$S_n = \frac{1}{n!} \frac{d^n}{d\lambda^n}\bigg|_{\lambda=0} Y \exp(-\lambda S_1) \cdots \exp(-\lambda^{n-1} S_{n-1}). \tag{104}$$

Using

$$Y = 1 + \lambda\overline{A} + \lambda^2\overline{\overline{A}A} + \lambda^3\overline{\overline{\overline{A}A}A} + \cdots \tag{105}$$

The prescription produces

$$
\begin{aligned}
S_1 &= \overline{A} \\
S_2 &= \overline{\overline{A}A} - \frac{1}{2!}S_1^2 \\
S_3 &= \overline{\overline{\overline{A}A}A} - S_2 S_1 - \frac{1}{3!}S_1^3 \\
S_4 &= \overline{\overline{\overline{\overline{A}A}A}A} - S_3 S_1 - \frac{1}{2!}S_2^2 - \frac{1}{2!}S_2 S_1^2 - \frac{1}{4!}S_1^4 \\
&\cdots \\
S_n &= \overbrace{\overline{\overline{\overline{A}A}\cdots A}}^{n \text{ times}} - \sum_{\substack{l\geq 2;\ i_1 n_1+\cdots+i_l n_l=n \\ n-1\geq i_1\geq\cdots\geq i_l\geq 1}} \frac{S_{i_1}^{n_1}\cdots S_{i_l}^{n_l}}{n_1!\cdots n_l!}
\end{aligned}
\tag{106}
$$

It is possible to construct an analogous left-infinite product for the problem

$$
\dot{\tilde{Y}} = \lambda A \tilde{Y},\ \tilde{Y}(0) = 1,
$$

such that

$$
\tilde{Y} = \exp(\lambda \tilde{S}_1)\exp(\lambda^2 \tilde{S}_2)\exp(\lambda^3 \tilde{S}_3)\cdots, \tag{107}
$$

where $\tilde{S}_1 = A$, and

$$
\tilde{S}_n = \overbrace{\overline{A\cdots\overline{A\overline{A}}}}^{n \text{ times}} - \sum_{\substack{l\geq 2;\ i_1 n_1+\cdots+i_l n_l=n \\ 1\leq i_1\leq\cdots\leq i_l\leq n-1}} \frac{\tilde{S}_{i_1}^{n_1}\cdots \tilde{S}_{i_l}^{n_l}}{n_1!\cdots n_l!} \tag{108}
$$

Wilcox gives another procedure in his paper [57]. Observe that

$$
\begin{aligned}
\dot{\tilde{Y}} = \lambda A \tilde{Y} &= (e^{\lambda \tilde{S}_1})' e^{\lambda^2 \tilde{S}_2} e^{\lambda^3 \tilde{S}_3}\cdots + e^{\lambda \tilde{S}_1}(e^{\lambda^2 \tilde{S}_2})' e^{\lambda^3 \tilde{S}_3}\cdots \\
&= \left[(e^{\lambda \tilde{S}_1})' e^{-\lambda \tilde{S}_1} + e^{\lambda \tilde{S}_1}(e^{\lambda^2 \tilde{S}_2})' e^{-\lambda^2 \tilde{S}_2} e^{-\lambda \tilde{S}_1} + \cdots\right]\tilde{Y}.
\end{aligned}
\tag{109}
$$

Using identities (76) and (77),

$$
\lambda A = -\sum_{n=0}^{\infty} \frac{\lambda^{n+1}}{(n+1)!}\mathrm{ad}_{S_1}^n \dot{S}_1 - \sum_{m,n=0}^{\infty} \frac{\lambda^{n+2(m+1)}}{n!(m+1)!}\mathrm{ad}_{S_1}^n \mathrm{ad}_{S_2}^m \dot{S}_2 + \cdots \tag{110}
$$

By equating powers of λ, one verifies by induction that the $\tilde{S}$'s are indeed Lie elements. By an analogous approach, the S's defined above are also Lie elements. The expression (110) is a bit too implicit to give a transparent idea of how to produce the S's, so we will prefer eqs. (106).

The following is an idea for evaluating the Magnus expansion that can be useful. We can evaluate Ω_n by comparing powers of λ:

$$
\exp(\Omega) = \sum_{n=0}^{\infty} \lambda^n \sum_{l\geq 1;\ i_1 n_1+\cdots+i_l n_l=n} \frac{\Omega_{i_1}^{n_1}\cdots\Omega_{i_l}^{n_l}}{l!} = \sum_{n=0}^{\infty} \lambda^n \left[\Omega_n + \sum_{l\geq 2;\ i_1 n_1+\cdots+i_l n_l=n} \frac{\Omega_{i_1}^{n_1}\cdots\Omega_{i_l}^{n_l}}{l!}\right] \tag{111}
$$

Equating powers in λ,

$$\Omega_n = \overbrace{\overline{\overline{\overline{A}A}\cdots A}}^{n \text{ times}} - \sum_{l\geq 2;\; i_1 n_1+\cdots+i_l n_l = n} \frac{\Omega_{i_1}^{n_1}\cdots\Omega_{i_l}^{n_l}}{l!}. \tag{112}$$

The first few Ω_n's are

$$\begin{aligned}
\Omega_1 &= \overline{A} \\
\Omega_2 &= \overline{\overline{A}A} - \frac{1}{2!}\Omega_1^2 \\
\Omega_3 &= \overline{\overline{\overline{A}A}A} - \frac{1}{2!}(\Omega_2\Omega_1 + \Omega_1\Omega_2) - \frac{1}{3!}\Omega_1^3 \\
\Omega_4 &= \overline{\overline{\overline{\overline{A}A}A}A} - \frac{1}{2!}(\Omega_3\Omega_1 + \Omega_1\Omega_3 + \Omega_2^2) - \frac{1}{3!}(\Omega_2\Omega_1^2 + \Omega_1^2\Omega_2) - \frac{1}{4!}\Omega_1^4.
\end{aligned} \tag{113}$$

By Magnus's theorem, these Ω_n's are guaranteed to be Lie elements.

7 Iterated integrals

In many applications, it will suffice to consider the case where we are working with a Lie algebra with finite basis,$\{X_i\}_{i=1}^d$ so that A can be written as

$$A(t) = \alpha^i(t)X_i. \tag{114}$$

By the summation convention, this is a sum over i from 1 to d. Then, expanding the iterated integrals, the solution of $\dot{Y} = AY$ can be written as

$$Y(t) = 1 + \sum_{n=1}^{\infty}\sum_{i_1,\ldots,i_n=1}^{d}\int_0^t d\alpha^{i_1\cdots i_n} X_{i_1}\cdots X_{i_n}, \tag{115}$$

where

$$\int_0^t d\alpha^{i_1\cdots i_n} := \int_0^t d\tau^1\alpha^{i_1}(\tau^1)\int_0^{\tau^1} d\tau^2\alpha^{i_2}(\tau^2)\cdots\int_0^{\tau^{n-1}} d\tau^n d\alpha^{i_n}(\tau^n). \tag{116}$$

The iterated integrals are interesting objects in their own right and were extensively investigated by Chen [13, 12]. For instance, as a consequence of eq. (9),

$$Y^{-1} = 1 + \sum_{n=1}^{\infty}(-1)^n\sum_{i_1,\ldots,i_n}\int^t d\alpha^{i_n\cdots i_1} X_{i_1}\cdots X_{i_n}, \tag{117}$$

and thus, from the requirement $Y^{-1}Y = 1$, we have

$$\int d\alpha^{i_1\cdots i_n} - \int d\alpha^{i_1\cdots i_{n-1}}\int d\alpha^{i_n} + \cdots + \int d\alpha^{i_1}\int d\alpha^{i_n\cdots i_2} - \int d\alpha^{i_n\cdots i_1} = 0, \tag{118}$$

so $(-1)^n\int d\alpha^{i_n\cdots i_1}$ is in some sense the inverse of $\int d\alpha^{i_1\cdots i_n}$ for an appropriately defined product. The iterated integrals do indeed turn out to have a group structure as well as many other interesting features. One important feature worth mentioning is the property

$$\int d\alpha^I \int d\alpha^J = \int d\alpha^{I \sqcup\!\sqcup J}, \tag{119}$$

where $I = (i_1 \cdots i_n)$ and $J = (j_1 \cdots j_m)$ are multi-indices and $\int d\alpha^{I \sqcup\!\sqcup J}$ is a sum over iterated integrals where each integral is indexed by a "shuffle" of I and J. We refer to [24] for details on this.

For now, consider

$$\Omega = \log(Y) = \sum_{l=1}^{\infty} \frac{(-1)^{l+1}}{l} \sum_{I_1,\dots,I_l} \int d\alpha^{I_1} \cdots \int d\alpha^{I_l} X_{I_1 \cdots I_l}, \tag{120}$$

where $X_I = X_{i_1,\dots,i_n}$ Using Dynkin's theorem,

$$\Omega = \log(Y) = \sum_{l=1}^{\infty} \frac{(-1)^{l+1}}{l} \sum_{I_1,\dots,I_l} \int d\alpha^{I_1} \cdots \int d\alpha^{I_l} \frac{[X_{I_1 \cdots I_l}]}{|I_1| + \cdots + |I_l|}. \tag{121}$$

A priori, there is nothing to suggest that the series in eq. (120) should be a Lie element. However, Ree showed [46] that the condition (119) on the coefficients in eq. (120) precisely means that Ω is a Lie element, so the use of Dynkin's theorem is justified.

We can rephrase other expressions defined in the previous section in terms of iterated integrals. Let

$$\Omega = \Omega_1 + \Omega_2 + \Omega_3 + \cdots \tag{122}$$

be the usual Magnus expansion. Then, from eq. (94), we get

$$\Omega = X_i I_1^i + [X_i, X_j] I_2^{ij} + [X_i, [X_j, X_k]] I_3^{ijk} + \cdots \tag{123}$$

where

$$I_n^{i_1 \cdots i_n} = \frac{1}{n} \sum_{l=1}^{n} \frac{(-1)^{l+1}}{l} \sum_{1 \le n_1 < \cdots < n_l \le n} \int d\alpha^{i_1 \cdots i_{n_1}} \cdots \int d\alpha^{i_{n_l} \cdots i_n} \tag{124}$$

or equally, from eq. (112)

$$I_n = \frac{1}{n} \left[\int d\alpha^n - \sum_{l \ge 2;\; i_1 n_1 + \cdots + i_l n_l = n} \frac{1}{l!} I_{i_1}^{n_1} \cdots I_{i_l}^{n_l} \right] \tag{125}$$

or, following the Mielnik-Plebanski-Strichartz formula (97),

$$I_n = \sum_{\pi \in S^n} \frac{\lambda(\pi)}{l!} \int d\alpha \circ \pi. \tag{126}$$

Similarly, for the Wilcox formula (102)

$$S_n = \frac{1}{n} [X_{i_1}, \dots, [X_{I_{n-1}}, X_{i_n}] \dots] J_n^{i_1 \cdots i_n}, \tag{127}$$

where

$$J_n = \int d\alpha^n - \sum_{\substack{l\geq 2;\ i_1 n_1+\cdots+i_l n_l = n \\ i_1\geq\cdots\geq i_l}} \frac{J_{i_1}^{n_1}\cdots J_{i_l}^{n_l}}{n_1!\cdots n_l!}. \tag{128}$$

8 Covariant Magnus and Wilcox expansions

The idea presented in the last chapter provides a way to separate expansions of iterated Lie brackets into series of right iterated brackets multiplied by some iterated integral. In the proof of Radford's theorem, we used Dynkin's expansion to produce Radford's expansion, essentially by the substitutions

$$\overline{[\cdots[\overline{\overline{X},X}],...,X]} \to -\nabla F_{i_1\cdots i_n}\int du^{i_n\cdots i_1}$$

and

$$\overline{[X,...,[\overline{X,\overline{X}}]\cdots]} \to -\nabla F_{i_1\cdots i_n}\int du^{i_1\cdots i_n}$$

In this section, we remark that this works in general: We can transform expressions derived in the previous section into covariant Radford-like expressions essentially by this substitution. I have not had time to write out the proof of this in detail; this chapter only contains the sketch of a proof.

As in Section 5, let

$$\dot{g} = -g\dot{u}^i(t)A_i(u(t)),$$

where $g(t) \in G$ and $u(t) \in M$, and let U be the time dependent vector field $U_t = \dot{u}^i(t)\partial_i$ on $M \times G$. Let $X = -LA + U$, the same time dependent vector field as defined in Sect. 5. Suppose $u(1) = u(0)$. Then, $\overline{U} = 0$, and we have equation (36),

$$g(t) = g(0)\exp(\Omega_1(-LA+U)(u(0),1)). \tag{129}$$

The map $G \to G : g \mapsto g\exp(\Omega_1(-LA+U)(u(0),1))$ is the flow map of the left-invariant vector field

$$L\big(\Omega_1(-LA+U)(u(0),1)\big) = \Omega_1(-LA+U)(u(0),\cdot).$$

evaluated at $t = 1$.

$X = -LA + U$ is a time dependent vector field on $M \times G$, and we have

$$\Phi_X^t(u(0),g(0)) = (u(t),g(t)).$$

This is essentially another application of Lemma 5.2. Let $\phi \in C^\infty(M \times G)$. Then,

$$\phi(u(t),g(t)) = \mathrm{Ev}_X^t\phi(u(0),g(0)).$$

Using Wilcox's expansion on the operator Ev_X^t, we have from eq. (102),

$$\mathrm{Ev}_X^t = \cdots\exp(S_3^t(X))\exp(S_2^t(X))\exp(S_1^t(X)).$$

Applying this to ϕ, we have

$$\mathrm{Ev}_X^t\phi = \phi\circ\Phi^1_{S_1^t(X)}\circ\Phi^1_{S_2^t(X)}\circ\Phi^1_{S_3^t(X)}\circ\cdots.$$

Evaluating at $t = 1$, we have $\overline{U} = 0$, and $S_1^1(X) = 0$. The higher order S's will become left-invariant vector fields, such that

$$\phi\circ\Phi^1_{S_1^t(X)}\circ\Phi^1_{S_2^t(X)}\circ\Phi^1_{S_3^t(X)}\circ\cdots(u(0),g(0)) = \phi(u(0),g(0)\cdots\exp(S_3^1(X))\exp(S_2^1(X)).$$

As was noted in Ch. 5,

$$\overline{[X,...,[\overline{X,\overline{X}}]\cdots]} = -L\nabla F_{i_1\cdots i_n}\int du^{i_1\cdots i_n}. \tag{130}$$

We can use the Wilcox expansion, eq. (102) and eq. (128) to obtain

$$g(1) = g(0) \cdots \exp(-\nabla_i \nabla_j F_{kl}(u(0)) J_4^{ijkl}) \exp(-\nabla_i F_{jk}(u(0)) J_3^{ijk}) \exp(-F_{ij}(u(0)) J_2^{ij}) \tag{131}$$

as the solution to

$$\dot{g} = -g\dot{u}^i(t) A_i(u(t)). \tag{132}$$

This "covariant" version of the Wilcox expansion is new. Regrettably I have not had the time to write out the details of the proof in a proper way.

9 Applications

In this section, we present two applications of our results.

9.1 Feynman diagrams and master integrals

The first application presented here is a method of evaluating Feynman integrals, called the Method of Differential Equations. This method, first proposed by Kotikov [29] has recently received substantial attention as a general algorithm for evaluating Feynman integrals. A nice review and exposition of this method is given in a Review paper by Argeri [2]. My contribution here is the application of the results of Section 5 to this method.

We first provide the general idea of evaluating Feynman integrals by the master integral approach. The main source of the exposition is [52].

Suppose we have a Feynman integral on the form

$$\begin{aligned} F(a_1, ..., a_N) &= \int_{-\infty}^{\infty} d^n k_1 \cdots \int_{-\infty}^{\infty} d^n k_L \frac{1}{D_1^{a_1} \cdots D_N^{a_N}}, \\ D_\alpha &= A_\alpha^{ij} q_i \cdot q_j - m_\alpha^2, \\ q &= (k_1, ..., k_L, p_1, ..., p_E), \end{aligned} \tag{133}$$

where E is the number of external momenta, L the number of independent internal momenta and $N = E + L$. F depends on the space-time dimension n, external momenta p_i, the masses m_i and exponents a_i of the propagators. We shall be concerned the dependence of F on the exponents $a_1, ..., a_N$. The goal of the differential equations approach, is to express $F(a_1, ..., a_n) \equiv F(\mathbf{a})$ in terms of simpler integrals. The starting point are the so-called integration by parts (IBP) identities,

$$\int_{-\infty}^{\infty} d^n k_1 \cdots \int_{-\infty}^{\infty} d^n k_L \frac{\partial}{\partial k_i} \cdot q_j \frac{1}{D_1^{a_1} \cdots D_N^{a_N}} = 0, \tag{134}$$

$i = 1, ..., L;\ j = 1, ..., E + L$. These follow from Stokes' Theorem, provided that the integrand vanishes quickly enough at infinity. Carrying out the differentiation inside the integral gives

$$nF(\mathbf{a})\delta_j^i - \sum_{u=1}^{N} a_u \int_{-\infty}^{\infty} d^n k_1 \cdots \int_{-\infty}^{\infty} d^n k_L \frac{(A_u^{ia} + A_u^{ai}) q_j \cdot q_a}{D_1^{a_1} \cdots D_u^{a_u+1} \cdots D_N^{a_N}} = 0. \tag{135}$$

Next, we need to assume that $(A_u^{ia} + A_u^{ai}) q_j \cdot q_a$ can be written as a linear combination of the D's, that is, $(A_u^{ia} + A_u^{ai}) q_j \cdot q_a = B_u^{ij,b} D_b$. This can be assumed in general (See [52], sect. 6 for details). The B's will in general be rational functions of the masses m_i^2. We have

$$
\begin{aligned}
nF(\mathbf{a})\delta^i_j - \sum_{u,b=1}^{N} a_u B^{ij,b}_u F(a_1, ..., a_b - 1, ..., a_u + 1, ..., a_N) = 0 \\
= \left[n\delta^i_j - \sum_{u,b=1}^{N} B^{ij,b}_u (\mathbf{n}_u - 1 + \delta_{ub})\mathbf{u}^+\mathbf{b}^-\right] F(\mathbf{a}) = 0 \\
=: \mathbf{O}_{ij} F(\mathbf{a}) = 0,
\end{aligned}
\tag{136}
$$

where we have defined the operators $\mathbf{n}_u, \mathbf{u}^+, \mathbf{u}^-$ by $\mathbf{n}_u F(a_1, ..., a_N) = a_u F(a_1, ..., a_N)$ and $\mathbf{u}^\pm F(a_1, ..., a_N) = F(a_1, ..., a_u \pm 1, ..., a_N)$. This relates the value of F at $a \in \mathbb{Z}^N$ to values at neighboring points in $\mathbb{Z}^N$.

Let $\mathcal{A}$ be the algebra generated by the operators $\{\mathbf{n}_u, \mathbf{u}^+, \mathbf{u}^-\}_{u=1}^N$ acting on F. Then, the operators $\mathbf{O}_{ij}$ generate a left ideal $\mathcal{J} \subset \mathcal{A}$.[1] Let

$$
F(a_1, ..., a_N) = (\mathbf{1}^+)^{a_1-1}(\mathbf{2}^+)^{a_2-1} \cdots (\mathbf{N}^+)^{a_N-1} F(1, ..., 1) =: \mathbf{Y}^{a_1-1,...,a_N-1} F(1, ..., 1). \tag{137}
$$

It is a fact that $\mathcal{A}/\mathcal{J}$ is finite dimensional.[2] This means that there is a finite (non-unique) basis $\mathbf{Y}^{I_1}, ..., \mathbf{Y}^{I_S}$ such that

$$
F(a_1, ..., a_N) = \sum_{i=1}^{S} c_i(a_1, ..., a_N)\mathbf{Y}^{I_i} F(1, ..., 1) =: \sum_{i=1}^{S} c_i(a_1, ..., a_N) I^i \tag{138}
$$

for all $(a_1, ..., a_N) \in \mathbb{N}^N$. The I^i's defined here are called Master Integrals. They are all of the form $I^i = F(\mathbf{a}^j)$ for some $\mathbf{a}^j \in \mathbb{N}^N$. Hopefully, one is able to find a basis such that the master integrals are simpler to evaluate that the original problem. The choice of basis is largely arbitrary. There exist several algorithms for reduction to Master Integrals, and several computer codes for this purpose have been produced, notably by Laporta [31].

Master integrals form a basis for $\{F(a) | a \in \mathbb{N}^N\}$, and there is a particularly neat way of evaluating them exploiting this fact: Let $\mathbf{a}^j \in \mathbb{N}^N$ such that $I^j = F(\mathbf{a}^j)$. Then, since $\{I^i\}_{i=1}^S$ is a basis,

$$
\begin{aligned}
\frac{\partial}{\partial m_i^2} I^j &= -a^j_i F(a^j_1, ..., a^j_i + 1, ..., a^j_N) \\
&= -\sum_{u=1}^{S} a^j_i c_u(a^j, ..., a^j_i + 1, ..., a^j_N) I^u =: \sum_{u=1}^{S} (A_i)^j_u I^u.
\end{aligned}
\tag{139}
$$

We have obtained a set of PDE's satisfied by the master integrals exactly on the same form as eq. (62). We can thus use Corollary 5.8 to evaluate them. Let $x = (m_1^2, ..., m_N^2)$, and let $x_0 \in \mathbb{R}^N$ be arbitrary. Then, by Corollary 5.8,

$$
\begin{aligned}
I(x) = \exp\big(&\Delta x^i A_i(x_0) + \frac{1}{2}\Delta x^i \Delta x^j \partial_i A_j(x_0) - \frac{1}{12}\Delta x^i \Delta x^j \Delta x^k [A_i(x_0), \partial_j A_k(x_0)] \\
&+ \frac{1}{6}\Delta x^i \Delta x^j \Delta x^k \partial_i \partial_j A_k(x_0) - \frac{1}{24}\Delta x^i \Delta x^j \Delta x^k \Delta x^l [A_i(x_0), \partial_j \partial_j A_k(x_0)] \\
&+ \frac{1}{24}\Delta x^i \Delta x^j \Delta x^k \Delta x^l \partial_i \partial_j \partial_k A_l(x_0) + \cdots \big) I(x_0).
\end{aligned}
$$

[1] In fact, the $\mathbf{O}$ also span a Lie algebra in $\mathcal{A}$, but this is not relevant for the discussion here.

[2] The proof of this is nontrivial: It means that it is in general possible to find a finite basis of master integrals for F. Smirnov proves this in [51], but I have not been able to understand the proof, and so the exact conditions on F for this to hold remain unclear to me. Smirnov remarks that it is an empirical fact that a finite master integral basis can be found by use of algorithms based on the ideas presented here.

Henn and Smirnov show in detail in [22] how $I(x)$ can be evaluated by parametrizing a path between x_0 and x, and evaluating the time ordered exponential. I conjecture that my approach here may give the same results. One way to check is to perform the same calculations as Henn and Smirnov in [22] and compare. This would be a very algebra-heavy exercise, which has yet to be done.

9.2 Deformable body in a low Reynolds number fluid

In this section, another application is presented. We will discuss how Radford's expansion can be used to analyze the dynamics of a deformable body immersed in a fluid. Much of this section is my own contribution, but should be taken with a grain of salt, as I'm not an expert. My main conceptual contribution is that the type of constraint given in eq. (143) gives an expression for a connection sought by Shapere and Wilczek. This has probably been remarked before in the literature, though I have not found a reference for it.

It has long been understood that fish propel themselves in water by constantly transferring momentum to the water by beating with tails and fins. This is the mechanism of any propulsion in a high-Reynolds number medium; propellers on boats and aircraft work by the same principle. However, at low Reynolds number, transfer of momentum to the fluid is not possible, as there is no self-sustained movement of the fluid. Serious investigations into the mechanisms of propulsion at low Reynolds number was not conducted until the seminal work of Lighthill in 1952 [32] and Saffman in 1967 [47] demonstrating that it is possible for a body immersed in a perfect low Reynolds number fluid to self-propel, i.e. bacteria can indeed swim. The problem received sporadic attention over the subsequent two decades. Blake showed in a 1971 paper [7] that a spherical microswimmer can self-propel by surface distortions that do not change the overall shape of the body. Purcell's 1977 lecture [44] introduced the famous scallop theorem, stating that a body with one trivial degree of freedom cannot propel itself at low Reynolds number. In 1987, Shapere and Wilczek, [50, 49] reframed the problem in terms of the holonomy of a gauge potential. Miloh and Galper [40, 30] derived in 1993 a set of relations connecting instantaneous body deformations with the euclidean movement of the body (i.e. the Lie algebra of the euclidean group) which turn out to constitute the connection sought by Shapere and Wilczek. In this chapter, we will attempt to make this connection explicit, by means of a set of boundary conditions introduced by Stone [55]. The main observation connecting this problem to the earlier discussion of the thesis is that the movement through the fluid of a body going through a sequence of deformations is the holonomy element of the Shapere-Wilczek connection, with respect to the path in the some space parameterizing the deformations.

The starting point of our discussion here will be Stone's paper [55]. We will consider a near-spherical body immersed in a low-Reynolds number fluid and show how the Shapere-Wilczek connection arises from the boundary conditions on the flow of the fluid.

A low-Reynolds number flow, or Stokes flow $\mathbf{u}$ models the velocity field of a fluid moving at low Reynolds number. We suppose the fluid to be incompressible. Then, the fluid flow is governed by Stokes's equations

$$\begin{aligned} -\nabla p + \mu \nabla^2 \mathbf{u} &= 0, \\ \nabla \cdot \mathbf{u} &= 0, \end{aligned} \tag{140}$$

where p is the pressure and μ is the viscosity of fhe fluid. These equations are well established and have been studied for a long time.[1] A deformation of an immersed body will induce a Stokes flow around it by means of the boundary condition on the interface between the body and the fluid.

By a deformation of a spherical body, we will mean the following. Let $\mathbf{r} : \mathbb{S}^2 \hookrightarrow \mathbb{R}^3$ be a orientation preserving immersion that preserves the centroid of the sphere,

[1]Many textbooks cover Stokes flows. A classic reference is [19] and a modern one is [4].

$$\int_{\mathrm{im}\,\mathbf{r}} dS\mathbf{r} = 0, \tag{141}$$

where dS is an area element. We call $\mathbf{r}$ a deformation of the sphere. By a sequence of deformations, we will mean a map $\mathbf{r} : [0,1] \times \mathbb{S}^2 \to \mathbb{R}^3$ such that $\mathbf{r}_t$ is a deformation. Denote the image of $\mathbb{S}^2$ under $\mathbf{r}_t$ by ∂B_t. The immersed body, whose movement we seek to study, will be identified with the volume enclosed by the boundary ∂B_t.

As defined, we call any orientation preserving immersion that preserves the centroid of the sphere a deformation. A sequence of deformations induces a flow around the body. This flow obeys eqs. (140), and is defined by a boundary condition on ∂B. Following [55], we impose a non-slip boundary condition, meaning

$$\mathbf{u}_t\big|_{\partial B_t} = \frac{d}{dt}\mathbf{r}_t\big|_{\partial B_t}. \tag{142}$$

By this boundary condition, the sequence of deformations $\mathbf{r}_t$ affect a net force and torque on the fluid. As our body floats freely in the fluid, this implies a net translation and rotation of the body, counteracting the force and torque exactly. We can model this translation and rotation as a path $(R, V) : [0,1] \to \mathrm{SE}(3)$ in the Euclidean group. Let $x \in \mathbb{S}^2$. Then, under the sequence of deformations $\mathbf{r}_t$, the point x follows the path $(R(t), V(t)) \cdot \mathbf{r}_t(x) = R(t)\mathbf{r}_t(x) + V(t)$ in $\mathbb{R}^3$. Let

$$\mathbf{v} := \frac{d}{dt}(R, V) \cdot \mathbf{r} = \dot{R}\mathbf{r} + R\dot{\mathbf{r}} + \dot{V}.$$

Stone showed that using the Reciprocal theorem for Stokes flow, conservation of linear and angular momentum implies that (R, V) obeys

$$\begin{aligned} \int_{\partial B_t} \mathbf{v} dS = 0, \\ \int_{\partial B_t} R\mathbf{r} \times \mathbf{v} dS = 0, \end{aligned} \tag{143}$$

where $\mathbf{n}$ is the outward unit normal on ∂B_t, dS is a surface element and $\times$ is the cross-product.[1] The equations (143) constitute the Shapere-Wilczek connection, defining the path (R, V) for a given sequence of deformations $\mathbf{r}$.

Suppose we restrict to surface deformations that do not alter the overall area of the body B, which we suppose to be 4π for convenience. Then, eq. (141) together with the first equation in (143) imply

$$\dot{V} = -\frac{1}{4\pi} R \int_{\partial B_t} \dot{\mathbf{r}}_t dS. \tag{144}$$

By eq. (141), $\int \mathbf{r} dS = 0$, and so, the second equality in (143) implies

$$\int_{\partial B_t} R\mathbf{r} \times \dot{R}\mathbf{r} dS = -\int_{\partial B_t} R\mathbf{r} \times R\dot{\mathbf{r}} dS. \tag{145}$$

We can define the angular velocity matrix Ω by $\Omega := R^{-1}\dot{R}$. Then, using the property of the cross product $R\mathbf{a} \times R\mathbf{b} = R(\mathbf{a} \times \mathbf{b})$, eq. (145) can be rewritten as

$$\int_{\partial B_t} \mathbf{r} \times \Omega\mathbf{r} dS = -\int_{\partial B_t} \mathbf{r} \times \dot{\mathbf{r}} dS. \tag{146}$$

[1] These relations have been used in other in-depth analyses of the problem of swimming at low Reynolds number, see [6].

Let M and N be the matrices defined by

$$\begin{aligned} M^{ij} &= \int_{\partial B_t} r^i r^j dS \\ N^{ij} &= -\int_{\partial B_t} (r^i \dot{r}^j - r^j \dot{r}^i) dS \end{aligned}$$

M is then symmetric and invertible, and N is antisymmetric. Let $i : \mathbb{R}^3 \xrightarrow{\sim} \mathfrak{so}(3)$ be defined by

$$i(x) = \begin{pmatrix} 0 & -x^3 & x^2 \\ x^3 & 0 & x^1 \\ -x^2 & x^1 & 0 \end{pmatrix}$$

i is a Lie algebra isomorphism with respect to the cross product and Lie bracket respectively. In particular, i satisfies $i(a \times b)^{ij} = a^j b^i - a^i b^j$, $i(a)b = a \times b$ and $i(a \times b) = [i(a), i(b)]$. By a direct calculation, it can be shown that

$$i \int_{\partial B_t} \mathbf{r} \times \Omega \mathbf{r} dS = \Omega M + M\Omega \tag{147}$$

We assume the application $\mathcal{M} : \mathfrak{so}(3) \to \mathfrak{so}(3) : \Omega \mapsto \Omega M + M\Omega$ to be invertible,[1] such that

$$\Omega = -\mathcal{M}^{-1} N, \tag{148}$$

which implies

$$\dot{R} = -R\mathcal{M}^{-1} N \tag{149}$$

If we restrict to spherical deformations, then $M = \frac{4\pi}{3}\mathbf{1}$, such that $\mathcal{M}^{-1} = \frac{3}{8\pi}\mathbf{1}$. This is consistent with the results of Stone [55]: If we restrict to spherical deformations, i.e. deformations that stretch and contract the surface without changing the overall shape, then eq. (145) can be rewritten as

$$\dot{R} = RA(\dot{\mathbf{r}}), \tag{150}$$

where $A(\dot{\mathbf{r}})$ is the antisymmetric matrix

$$A^{ij}(\dot{\mathbf{r}}) = \frac{3}{8\pi} \int_{\partial B_t} (r^i \dot{r}^j - r^j \dot{r}^i) dS. \tag{151}$$

In order to get an intuition about how Radford's theorem is relevant for solving eq. (150), we will invoke a specific model for a swimming bacterium, considered in a paper by Blake [7]. Suppose, for simplicity, that we have deformation that does not alter the radius of the body. Such a deformation can be modeled by a change in spherical coordinates, $(\theta, \phi) \mapsto (\Theta, \Phi)$. Suppose ϵ a small parameter, and let

$$\begin{aligned} \Theta_t(\theta, \phi) &= \theta + \epsilon \sum_{l=0}^{\infty} \sum_{m=-l}^{l} c_{l,m}(t) Y^{l,m}(\theta, \phi), \\ \Phi_t(\theta, \phi) &= \phi + \epsilon \sum_{l=0}^{\infty} \sum_{m=-l}^{l} d_{l,m}(t) Y^{l,m}(\theta, \phi), \end{aligned} \tag{152}$$

[1] I do not know whether this holds generally, as I have not been able to find an inverse for this map. For spherical deformations, M is indeed invertible, and so, deformations that deviate only little from a spherical deformation will be invertible.

where $0 < \theta < \pi$; $0 < \phi < 2\pi$ are the usual spherical coordinates parameterizing the sphere. $Y_{l,m}$ are real spherical harmonics defined by

$$Y_{l,m}(\theta,\phi) = \begin{cases} P_{l,m}(\theta)\cos(m\phi), & \text{if } m \geq 0 \\ P_{l,|m|}(\theta)\sin(m\phi), & \text{if } m < 0, \end{cases}$$

where P_l^m is the associated Legendre polynomial, defined as the solution to

$$\frac{d}{dx}\left[(1-x^2)\frac{d}{dx}P_l^m(x)\right] + \left[l(l+1) - \frac{m^2}{1-x^2}\right]P_l^m(x) = 0.$$

Any small $O(\epsilon)$ deformation that leaves the overall shape of the body spherical can be approximated to arbitrary accuracy by a truncation of the series (152). Say, we use n terms in the sums, such that

$$\begin{aligned} \Theta_t(\theta,\phi) &= \theta + \epsilon \sum_{l=0}^{n}\sum_{m=-l}^{l} c_{l,m}(t) Y^{l,m}(\theta,\phi), \\ \Phi_t(\theta,\phi) &= \phi + \epsilon \sum_{l=0}^{n}\sum_{m=-l}^{l} d_{l,m}(t) Y^{l,m}(\theta,\phi), \end{aligned}$$

Then, the configuration depends on the $N := (n+2)(n+1)$ coordinates $c := \{c_{l,m}, d_{l,m}\} \in \mathbb{R}^N$. Using Θ and Φ, we can write $\mathbf{r}$ at time t as

$$\begin{aligned} r^x &= \cos(\Phi_t)\sin(\Theta_t) \\ r^y &= \sin(\Phi_t)\sin(\Theta_t) \\ r^z &= \cos(\Theta_t). \end{aligned}$$

The connection in eq (150) can then be expanded by order in ϵ:

$$\dot{R} = R(\epsilon A_1(c) + \epsilon^2 A_2(c) + \epsilon^3 A_3(c) + \cdots) \tag{153}$$

A sequence of such deformations can be described by a path $c(t) \in \mathbb{R}^N$. Radford's theorem can then be invoked to evaluate $R(t)$. Obtaining the terms in (153) is a lengthy calculation, and has yet to be done.

We can demonstrate that a bacterium can make itself rotate about any axis by means of the type of surface deformation described above. Suppose

$$\begin{aligned} \Theta(\theta,\phi) &= \theta, \\ \Phi(\theta,\phi) &= \phi + \epsilon\cos(n\phi - \omega t). \end{aligned}$$

This is a periodic surface deformation around the z-axis. We can expand the terms in the integral (151) by order in ϵ, up to second order:

$$\begin{aligned} dS &= dS_0 + \epsilon dS_1 + O(\epsilon^2), \\ \mathbf{r} &= \mathbf{r}_0 + \epsilon\mathbf{r}_1 + O(\epsilon^2), \\ \dot{\mathbf{r}} &= \epsilon\dot{\mathbf{r}}_1 + \epsilon^2\dot{\mathbf{r}}_2 + O(\epsilon^3). \end{aligned}$$

Using this, to expand A by order in ϵ,

$$\begin{aligned} A_1^{ij} &= \frac{3}{8\pi}\int dS_0(n_0^i\dot{r}_1^j - n_0^j\dot{r}_1^i), \\ A_2^{ij} &= \frac{3}{8\pi}\int dS_1(n_0^i\dot{r}_1^j - n_0^j\dot{r}_1^i) + dS_0(n_1^i\dot{r}_1^j - n_1^j\dot{r}_1^i) + dS_0(n_0^i\dot{r}_2^j - n_0^j\dot{r}_2^i). \end{aligned} \tag{154}$$

Explicitly,

$$\begin{aligned} dS &= d\phi\sin(\theta)d\theta(1-\epsilon n\cos(n\phi-\omega t)), \\ \mathbf{r} &= \begin{pmatrix}\cos\phi\sin\theta \\ \sin\phi\sin\theta \\ \cos\theta\end{pmatrix} + \epsilon\cos(n\phi-\omega t)\begin{pmatrix}-\sin\phi\sin\theta \\ \cos\phi\sin\theta \\ 0\end{pmatrix} + O(\epsilon^2), \\ \dot{\mathbf{r}} &= \epsilon\omega\sin(n\phi-\omega t)\begin{pmatrix}-\sin\phi\sin\theta \\ \cos\phi\sin\theta \\ 0\end{pmatrix} + \epsilon^2\cos(n\phi-\omega t)\sin(n\phi-\omega t)\begin{pmatrix}-\cos\phi\sin\theta \\ -\sin\phi\sin\theta \\ 0\end{pmatrix} + O(\epsilon^3). \end{aligned}$$

Using these to preform the integrals in eq. (154), we get $A_1^{ij} = 0$, and

$$A_2 = -\frac{1}{2}n\omega\begin{pmatrix}0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0\end{pmatrix},$$

such that

$$\Omega = -\frac{1}{2}\epsilon^2 n\omega\begin{pmatrix}0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0\end{pmatrix}.$$

This means that the deformation induces a rotation around the z axis. Stone showed that a similar periodic surface deformation in the θ-direction allows the organism to propel itself forward with speed $\propto \epsilon^2 n\omega$. Our result here shows that the organism can rotate by the same mechanism around any axis with angular velocity $\propto \epsilon^2 n\omega$. The combination of these two results imply that the path of the organism in the fluid is completely controllable; it can follow any path in SE(3) by a combination of rotations and translations. An investigation into the controllability of microswimmers has already been done in a paper by Lohéac and Munnier [33].

10 Conclusion

The main result of this thesis is the generalization and (arguably) simplified proof of Radford's Theorem given in Section 4. This was derived by use of a formalism of evolution operators vector fields. This formalism is known, yet as the approach is unpopular, the literature is sparse. Several of the ideas surrounding the proof are therefore developed here. Some of the consequences in Section 5 of the generalization of Radford's Theorem are new, yet are expected: The fact that zero curvature implies path independence for iterated integrals is known, but I have not been able to find explicit formulas exhibiting this elsewhere. Sections 6 and 7 contain many things already known, but I provide my own insights and ideas along the way. In Section 8, I presented an idea for generalizing Radford's theorem to produce other expansions, such as the one based on Wilcox's expansion. I did not have time to complete the proof. The physical applications of the results are relevant, but Subsection 9.2 failed to provide a direct application. Rather, Subsection 9.2 presents an idea of how to apply the results of Section 4, supplied with a result about how spherical microswimmers can self-rotate.

Further work

Much of the derivations and proofs provided here are not mathematically rigorous enough to proclaim that the main results of this thesis should be accepted without scrutiny. A proper proof of the result would need more care and detail than provided here. This remains to be done. Convergence analyses of the results remain to be done. One could devise algorithms to produce the terms in the generalized Radford expansion (35), and the expansions provided in Theorem 5.7 and eq. (66). Numerical convergence analyses of these results remain to be done. As mentioned at the end of Subsection 9.1, the results and numerical efficacy of the method could be compared to the work done by Henn & Smirnov by means of the application to Master Integrals. The validity of the analysis of Subsection 9.2 needs to be checked by an expert on low Reynolds number dynamics, as I am not an expert myself. The terms in eq. (153) remain to be calculated, and it would be nice to supply a proper numerical simulation of a moving deformable body using this method. Otherwise, I have not had time to double-check all the calculations in this thesis, so there may be some errors.

Bibliography

[1] R. Abraham and J.E. Marsden. *Foundations of Mechanics*. AMS Chelsea publishing. AMS Chelsea Pub./American Mathematical Society, 2008. ISBN: 978-0-821-84438-0. URL: https://books.google.no/books?id=4Y-ownk6ilsC.

[2] M. Argeri and P. Mastrolia. 'Feynman diagrams and differential equations'. In: *International Journal of Modern Physics A* 22.24 (2007), pp. 4374–4436. ISSN: 0217751X. DOI: 10.1142/S0217751X07037147.

[3] V. I. Arnold and B. A. Khesin. *Topological methods in hydrodynamics*. Applied Mathematical Sciences. Springer Cham, 1992. ISBN: 978-3-030-74280-5. DOI: 10.1007/978-3-030-74278-2.

[4] D. Barthès-Biesel. *Microhydrodynamics and Complex Fluids*. CRC Press, 2012. ISBN: 978-0-429-09939-7. DOI: 10.1201/b12147.

[5] B.M. Berry. 'Quantal phase factors accompanying adiabatic changes'. In: *Proc. R. Soc. Lond. A* 392 (1984), pp. 45–57.

[6] L. Berti, L. Giraldi and C. Prud'homme. 'Swimming at low Reynolds number'. In: *ESAIM: Proceedings and Surveys* 67 (2020), pp. 46–60. DOI: 10.1051/proc/202067004.

[7] J. R. Blake. 'A spherical envelope approach to ciliary propulsion'. In: *Journal of Fluid Mechanics* 46.1 (1971), pp. 199–208. DOI: 10.1017/S002211207100048X.

[8] S. Blanes et al. 'Magnus and Fer expansions for matrix differential equations: The convergence problem'. In: *Journal of Physics A: Mathematical and General* 31.1 (1998). ISSN: 03054470. DOI: 10.1088/0305-4470/31/1/023.

[9] S. Blanes et al. 'The Magnus expansion and some of its applications'. In: *Physics Reports* 470.5-6 (2009). ISSN: 03701573. DOI: 10.1016/j.physrep.2008.11.001.

[10] A. Bonfiglioli and R. Fulci. *Topics in noncommutative Algebra: The Theorem of Campbell, Baker, Hausdorff and Dynkin*. Springer Berlin, Heidelberg, 2012. ISBN: 978-3-642-22596-3. DOI: 10.1007/978-3-642-22597-0.

[11] F. Casas, A. Murua and M. Nadinic. 'Efficient computation of the Zassenhaus formula'. In: *Computer Physics Communications* 183.11 (2012). ISSN: 00104655. DOI: 10.1016/j.cpc.2012.06.006.

[12] K. Chen. 'Algebraic paths'. In: *Journal of Algebra* 10.1 (1968), pp. 8–36. ISSN: 0021-8693. DOI: 10.1016/0021-8693(68)90102-6.

[13] K. Chen. 'Integration of Paths, Geometric Invariants and a Generalized Baker- Hausdorff Formula'. In: *The Annals of Mathematics* 65.1 (1957). ISSN: 0003486X. DOI: 10.2307/1969671.

[14] F. Fer. 'Résolution de l'équation matricielle dU/dt=pU par produit infini d'exponentielles matricielles'. In: *Acad. Roy. Belg. Bull. Cl. Sci.* 44 (1958), pp. 818–829.

[15] R. P. Feynman. 'An operator calculus having applications in quantum electrodynamics'. In: *Physical Review* 84.1 (1951). ISSN: 0031899X. DOI: 10.1103/PhysRev.84.104.

[16] T. Frankel. *The Geometry of Physics*. Cambridge University Press, 2011. ISBN: 978-1-139-06137-7. DOI: 10.1017/cbo9781139061377.

[17] A Guichardet. 'On rotation and vibration motions of molecules'. In: *Annales de l'I.H.P. Physique théorique* 40.3 (1984). ISSN: 0246-0211.

[18] B. C. Hall. *Lie Groups, Lie Algebras, and Representations: An Elementary Introduction*. Springer Cham, 2013. ISBN: 978-3-319-13466-6. DOI: 10.1007/978-3-319-13467-3.

[19] J. Happel and H. Brenner. 'Low Reynolds number hydrodynamics'. In: *Fluid Mechanics: Foundations and Applications of Mechanics*. Cambridge University Press, 1973. ISBN: 978-1-316-13403-0. DOI: 10.1017/cbo9781316134030.008.

[20] P. Hartman. *Ordinary differential equations*. 2nd ed. Classics in applied mathematics 38. Society for Industrial and Applied Mathematics, 1982. ISBN: 3764330686.

[21] J. M. Henn. 'Multiloop integrals in dimensional regularization made simple'. In: *Physical Review Letters* 110.25 (2013). ISSN: 00319007. DOI: 10.1103/PhysRevLett.110.251601.

[22] J. M. Henn and V. A. Smirnov. 'Analytic results for two-loop master integrals for Bhabha scattering I'. In: *Journal of High Energy Physics* 2013 (11 2013). ISSN: 10298479. DOI: 10.1007/JHEP11(2013)041.

[23] A. Iserles and S. P. Nørsett. 'On the solution of linear differential equations in Lie groups'. In: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 357.1754 (1999). ISSN: 1364503X. DOI: 10.1098/rsta.1999.0362.

[24] M. Kawski. 'The combinatorics of nonlinear controllability and noncommuting flows'. In: *Lectures given at the Summer School on Mathematical Control Theory*. 2001.

[25] S. Klarsfeld and J. A. Oteo. 'Exponential infinite-product representations of the time-displacement operator'. In: *Journal of Physics A: Mathematical and General* 22.14 (1989). ISSN: 03054470. DOI: 10.1088/0305-4470/22/14/019.

[26] S. Kobayashi and K. Nomizu. *Foundations of Differential Geometry*. Foundations of Differential Geometry vol. 1. Interscience Publishers, 1963. ISBN: 978-0-470-49648-0. URL: https://books.google.no/books?id=wn4pAQAAMAAJ.

[27] I. Kolář, J. Slovák and P. W. Michor. *Natural Operations in Differential Geometry*. Springer Berlin, Heidelberg, 1993. DOI: 10.1007/978-3-662-02950-3.

[28] A. V. Kotikov. 'Differential equations method. New technique for massive Feynman diagram calculation'. In: *Physics Letters B* 254.1 (1991), pp. 158–164. ISSN: 0370-2693. DOI: 10.1016/0370-2693(91)90413-K. URL: https://www.sciencedirect.com/science/article/pii/037026939190413K.

[29] A. V. Kotikov. 'Differential equations method. New technique for massive Feynman diagram calculation'. In: *Physics Letters B* 254.1-2 (1991). ISSN: 03702693. DOI: 10.1016/0370-2693(91)90413-K.

[30] L. Landweber and T. Miloh. 'Unsteady Lagally theorem for multipoles and deformable bodies'. In: *Journal of Fluid Mechanics* 96.1 (1980). ISSN: 14697645. DOI: 10.1017/S0022112080002005.

[31] S. Laporta. 'High-precision calculation of multiloop Feynman integrals by difference equations'. In: *International Journal of Modern Physics A* 15.32 (2000). ISSN: 0217751X. DOI: 10.1142/S0217751X00002159.

[32] M. J. Lighthill. 'On the squirming motion of nearly spherical deformable bodies through liquids at very small reynolds numbers'. In: *Communications on Pure and Applied Mathematics* 5.2 (1952). ISSN: 10970312. DOI: 10.1002/cpa.3160050201.

[33] J. Lohéac and A. Munnier. 'Controllability of 3D low reynolds number swimmers'. In: *ESAIM - Control, Optimisation and Calculus of Variations* 20.1 (2014). ISSN: 12928119. DOI: 10.1051/cocv/2013063.

[34] H. G. Loos. 'Internal holonomy groups of Yang-Mills fields'. In: *Journal of Mathematical Physics* 8.10 (1967). ISSN: 00222488. DOI: 10.1063/1.1705128.

[35] M. Lothaire. *Combinatorics on Words*. 2nd ed. Cambridge Mathematical Library. Cambridge University Press, 1997. DOI: 10.1017/CBO9780511566097.

[36] D. Lucarelli. 'Control aspects of holonomic quantum computation'. In: *Journal of Mathematical Physics* 46.5 (2005). ISSN: 00222488. DOI: 10.1063/1.1888028.

[37] W. Magnus. 'On the exponential solution of differential equations for a linear operator'. In: *Communications on Pure and Applied Mathematics* 7.4 (1954). ISSN: 10970312. DOI: 10.1002/cpa.3160070404.

[38] J. B. Melli, C. W. Rowley and D. S. Rufat. 'Motion planning for an articulated body in a perfect planar fluid'. In: *SIAM Journal on Applied Dynamical Systems* 5.4 (2006). ISSN: 15360040. DOI: 10.1137/060649884.

[39] B. Mielnik and J. F. Plebański. 'Combinatorial approach to Baker-Campbell-Hausdorff exponents'. In: *Annales De L Institut Henri Poincare-physique Theorique* 12 (1970), pp. 215–254.

[40] T. Miloh and A. Galper. 'Self-propulsion of general deformable shapes in a perfect fluid'. In: *Proceedings of the Royal Society of London. Series A: Mathematical and Physical Sciences* 442.1915 (1993). ISSN: 0962-8444. DOI: 10.1098/rspa.1993.0104.

[41] R. Montgomery. ‘Gauge theory of the falling cat’. In: *Dynamics and Control of Mechanical Systems: The Falling Cat and Related Problems*. Fields Institute Communications, July 1993. ISBN: 978-0-8218-9200-8. DOI: 10.1090/fic/001/09.

[42] R. Montgomery. ‘Nonholonomic Control and Gauge Theory’. In: *Nonholonomic Motion Planning*. Ed. by Zexiang Li and J. F. Canny. Boston, MA: Springer US, 1993, pp. 343–377. ISBN: 978-1-4615-3176-0. DOI: 10.1007/978-1-4615-3176-0_9. URL: https://doi.org/10.1007/978-1-4615-3176-0_9.

[43] M. Nakahara. *Geometry, Topology and Physics*. 2nd. Edition. CRC Press, 2003. ISBN: 978-1-315-27582-6. DOI: 10.1201/9781315275826.

[44] E. M. Purcell. ‘Life at low Reynolds number’. In: *American Journal of Physics* 45.1 (1977). ISSN: 0002-9505. DOI: 10.1119/1.10903.

[45] J. E. Radford and J. W. Burdick. ‘Local Motion Planning for Nonholonomic Control Systems Evolving on Principal Bundles’. In: *In Proc. Mathematical Theory of Networks and Systems*. 1998.

[46] R. Ree. ‘Lie Elements and an Algebra Associated With Shuffles’. In: *The Annals of Mathematics* 68.2 (1958). ISSN: 0003486X. DOI: 10.2307/1970243.

[47] P. G. Saffman. ‘The self-propulsion of a deformable body in a perfect fluid’. In: *Journal of Fluid Mechanics* 28.2 (1967). ISSN: 14697645. DOI: 10.1017/S0022112067002149.

[48] A. Schmeding. *An introduction to infinite-dimensional differential geometry*. 2021. DOI: 10.48550/ARXIV.2112.08114. URL: https://arxiv.org/abs/2112.08114.

[49] A. Shapere and F. Wilczek. ‘Gauge kinematics of deformable bodies’. In: *American Journal of Physics* 57.6 (1989). ISSN: 0002-9505. DOI: 10.1119/1.15986.

[50] A. Shapere and F. Wilczek. ‘Self-propulsion at low reynolds number’. In: *Physical Review Letters* 58.20 (1987). ISSN: 00319007. DOI: 10.1103/PhysRevLett.58.2051.

[51] A.r V. Smirnov and A. V. Petukhov. ‘The Number of Master Integrals is Finite’. In: *Letters in Mathematical Physics* 97.1 (2011). ISSN: 03779017. DOI: 10.1007/s11005-010-0450-0.

[52] V. A. Smirnov. *Analytic Tools for Feynman Integrals*. Springer Tracts in Modern Physics. Springer Berlin, Heidelberg, 2012. DOI: 10.1007/978-3-642-34886-0.

[53] R. F. Snider. ‘Perturbation variation methods for a quantum Boltzmann equation’. In: *Journal of Mathematical Physics* 5.11 (1964). ISSN: 00222488. DOI: 10.1063/1.1931191.

[54] N. Steenrod. *The Topology of Fibre Bundles*. Princeton University Press, 1951. DOI: 10.1515/9781400883875.

[55] H. A. Stone and A. D.T. Samuel. ‘Propulsion of Microorganisms by Surface Distortions’. In: *Physical Review Letters* 77.19 (1996). ISSN: 10797114. DOI: 10.1103/PhysRevLett.77.4102.

[56] R. S. Strichartz. ‘The Campbell-Baker-Hausdorff-Dynkin formula and solutions of differential equations’. In: *Journal of Functional Analysis* 72.2 (1987). ISSN: 10960783. DOI: 10.1016/0022-1236(87)90091-7.

[57] R. M. Wilcox. ‘Exponential operators and parameter differentiation in quantum physics’. In: *Journal of Mathematical Physics* 8.4 (1967). ISSN: 00222488. DOI: 10.1063/1.1705306.

[58] M. Zadra and E. L. Mansfield. ‘Using lie group integrators to solve two and higher dimensional variational problems with symmetry’. In: *Journal of Computational Dynamics* 6.2 (2019). ISSN: 21582505. DOI: 10.3934/jcd.2019025.

A Notes on differential geometry and Lie groups

I decided to include this chapter on differential geometry and Lie groups in order to neatly summarize some of the results and facts used in the thesis. The material presented here is well-known, and can be found in many classic textbooks. The textbooks most familiar to me, which serve as sources for this material, are classic texts on differential geometry, like [54] and [26], the well-known physics-directed work by Nakahara [43] and Frankel [16], and the very detailed work by Kolar *et al.* [27]. Arnold & Khesin's book [3] contains a very nice review of Lie groups. Abraham & Marsden's book [1] is also a useful source taking the mechanics-perspective.

This section is not meant as a pedagogical introduction to differential geometry, but serves to establish notation and present some important results used in the thesis. Firstly, we introduce basic notions from differential geometry. This is mainly in order to establish notation and definitions, such that this section is as self-contained as possible. Secondly, we review some notions from the theory of Lie groups. Lastly, we review notions on principal bundles and how they relate to notions in gauge theory, as a physicist might know it.

A.1 Differential Geometry

A.1.1 Definition of Manifolds and Atlases

We define a d-dimensional real manifold M to be a second-countable Hausdorff topological space that is locally homeomorphic to $\mathbb{R}^d$. Such a local homeomorphism is called a chart, noted as a pair (U, x), where $U \subset M$ is the domain of the homeomorphism $x : U \xrightarrow{\sim} x(U) \subset \mathbb{R}^d$. A collection of charts such that their domains cover M is called an atlas. Let $\mathcal{A}$ be an atlas of M. Then, if $\forall x, y \in \mathcal{A} : y \circ x^{-1} \in C^\infty(\mathbb{R}^d, \mathbb{R}^d)$, we call $\mathcal{A}$ a smooth atlas. We define two smooth atlases to be compatible if their union is again a smooth atlas. Compatibility of smooth atlases defines an equivalence relation on atlases. We define a smooth manifold to be a manifold M equipped with a smooth atlas $\mathcal{A}$ and define charts on M to be any chart compatible with a chart in $\mathcal{A}$. Equivalently, we can define a smooth manifold to be a manifold with a prescribed equivalence relation of smooth atlases. For clarity of notation, $C^\infty(M, N)$ denotes smooth functions between manifolds M and N. A map $f : M \to N$ is called smooth if it is smooth in charts, that is, if $y \circ f \circ x^{-1}$ is smooth for all charts x in M and y in N.

A.1.2 The Tangent Space of a Manifold

We will now introduce the notion of tangent space of a manifold. The details here are not important, but a noteworthy point is that tangent vectors are defined in terms of equivalence classes of derivatives of some path. Tangent vectors are thus readily identified with directional derivatives, which can act on smooth functions. This is an important point.

Let (U, x) be a chart, and let $p \in U \subset M$. We say x is a chart at p if $x(p) = 0$. In the following, we shall for simplicity denote functions $(t \mapsto \gamma(t))$ by $\gamma(t)$. Let

$$C^\infty_m(\mathbb{R}, M) := \{\gamma \in C^\infty(\mathbb{R}, M) | \gamma(0) = m, \},$$

and let $\gamma \sim \tilde{\gamma}$ iff there is a chart ϕ at m such that

$$(\phi \circ \gamma)'(0) = (\phi \circ \tilde{\gamma})'(0).$$

$C^\infty_x(\mathbb{R}, M)/\!\sim \,=: T_m M$ comes with a natural vector space structure. This comes from the fact that

$$\phi^{-1}((\phi \circ \tilde{\gamma})'(0)t + (\phi \circ \gamma)(0)) \sim \gamma,$$

such that addition can be defined by

$$[\gamma] + [\tilde{\gamma}] := [\phi^{-1}(((\phi \circ \gamma)'(0)t + (\phi \circ \tilde{\gamma})'(0))t + (\phi \circ \gamma)(0))],$$

and scalar multiplication by

$$\alpha[\gamma] := [\phi^{-1}(\alpha(\phi \circ \gamma)'(0)t + (\phi \circ \gamma)(0)].$$

It can be shown that these definitions of addition and scalar multiplication on T_mM are well defined, by showing that they are independent of the choice of chart ϕ.

We denote the projection of a path into it's equivalence class by $d/dt|_{t=0}$:

$$\left.\frac{d}{dt}\right|_{t=0} : C_m^\infty(\mathbb{R}, M) \to T_mM : \gamma \mapsto [\gamma] \tag{155}$$

For $X \in T_mM$, we shall denote by γ_X an arbitrary element of X. It is a classic result that T_mM with it's vector space structure is equivalent to the derivations on $C_m^\infty(M, \mathbb{R})$, such that any element of T_mM can be regarded as a first order linear differential operator. This motivates the following notation.

Let (U, x) be a chart at $p \in M$. Let $e_i := (0, ..., 1, 0, ..., 0)$ be the i-th unit vector in $\mathbb{R}^d$. We denote $\forall q \in U$:

$$\frac{\partial}{\partial x^i}(q) := \left.\frac{d}{dt}\right|_{t=0} x^{-1}(te_i + x(q)). \tag{156}$$

As before, "$te_i + x(q)$" really means $(t \mapsto te_i + x(q))$. Then, the tangent vectors $\{\partial/\partial x^i(q)|i = 1, 2, ..., d\}$ span T_qM, and we call $U \to TM^d : q \mapsto \{\partial/\partial x^i(q)\}_{i=1}^d$ a coordinate frame on U. Let $f \in C^\infty(\mathbb{R}^M, \mathbb{R}^N)$, then we denote by Df the $M \times N$ matrix, whose i, j'th entry is $Df_j^i := \partial f^i/\partial x^j$. Let $\phi \in C^\infty(M, N)$, and $p \in M$. The *differential* of ϕ at p is a map $d\phi_p : T_mM \to T_{\phi(p)}N$ defined by $\forall X \in T_pM$:

$$d\phi_p X := \left.\frac{d}{dt}\right|_{t=0} \phi \circ \gamma_X.$$

Given two charts x, y with nonempty intersection of domains, we have

$$\frac{\partial}{\partial y^i}(q) = \frac{\partial x^j}{\partial y^i}(q)\frac{\partial}{\partial x^j}(q), \tag{157}$$

where $\partial x^j/\partial y^i(q) := D(x^j \circ y^{-1})_i(y(q))$. This generalises naturally to higher order tensors.

A.1.3 The Tangent Bundle of a Manifold

Next, we are ready to define the important notion of tangent bundle of a manifold. The tangent bundle of a manifold M if dimension d will be denoted TM and is itself a $2d$-dimensional manifold equipped with a projection map onto M.

Let

$$TM := \bigsqcup_{q \in M} T_qM.$$

TM has an obvious vector space structure inherited from T_qM, and we give TM a manifold structure by declaring that any chart (U, x) of M induces a chart $\mathbb{R}^{2d} \to TM$: $(x^1, ..., x^d, \alpha^1, ..., \alpha^d) \mapsto \alpha^i\partial/\partial x^i(x^{-1}(x^1, ..., x^d))$. The map $\pi : TM \to M$: $\alpha^i\partial/\partial x^i(p) \mapsto p$ is well defined and smooth. We call the combination of manifolds and maps $TM \xrightarrow{\pi} M$ the tangent bundle of M.

A *vector field* X is a section of the tangent bundle, that is, $X : M \to TM$ such that $\pi \circ X = \mathrm{id}_M$. Let $\mathcal{X}(M)$ denote the set of vector fields on M.

Next we remark a lemma that identifies derivations with vector fields. This is from section 3.3 in Kolar *et al.* [27].

A *derivation* D on M is a map $D : C^\infty(M) \to C^\infty(M)$ satisfying the Leibiniz rule: $\forall \phi, \psi \in C^\infty(M) : D(\phi\psi) = \phi D(\psi) + D(\phi)\psi$. There is a one-to-one correspondence of vector fields and derivations on M.

A.2 Lie groups

A Lie group G is a manifold with group structure such that $G \to G : g \mapsto g^{-1}$ is smooth and $G \times G \to G : (g, h) \mapsto gh$ is smooth. Let $1 \in G$ denote the unit element. Let $A \in T_1G$. Then, A generates a vector field on G by the action of G on itself. Define the left-invariant vector field LA on G generated by A by

$$LA(g) := dg_1 A,$$

where dg is the differential of the map $G \to G : h \mapsto gh$. We call $A \in \mathcal{X}(G)$ *left-invariant* if

$$\forall g \in G : g^* LA = LA,$$

meaning

$$\forall g, h \in G : dg_{gh}^{-1} LA(gh) = LA(h).$$

This allows us to define a Lie bracket on T_1G, turning it into a Lie algebra. Let $A, B \in T_1G$, and

$$[A, B]_{\text{Lie}} := dg_g^{-1} [LA, LB](g).$$

This is independent of g, since, in fact,

$$[LA, LB] = L[A, B]_{\text{Lie}} . \tag{158}$$

Denote by $\mathcal{L}(G)$ the Lie algebra of left-invariant vector fields on G. By eq. (158), $\mathcal{L}(G)$ is closed under the vector field bracket. Denote $\text{Lie}\, G := (T_1G, [\cdot, \cdot]_{\text{Lie}})$, the algebra formed by the tangent space at identity with product defined by the Lie bracket. From any $X \in \mathcal{L}(G)$, we obtain an element of T_1G by evaluating at $1 \in G$. This provides an inverse for the map $L : \text{Lie}\, G \to \mathcal{L}(G)$, which becomes a Lie algebra isomorphism. This is a classic result, and can be found in many textbooks on Lie groups.

Proposition A.1. *$L : \text{Lie}\, G \to \mathcal{L}(G)$ is a Lie algebra isomorphism.*

Proof. Let $X \in \text{Lie}\, G$. Then, $LX(1) = X$. Let $\Xi \in \mathcal{L}(G)$, then, $\Xi(g) = dg_1\Xi(1)$, and $L(\Xi(1))(g) = dg_1\Xi(1) = \Xi(g)$. This shows that L is bijective. It remains to shown that L respects the brackets. By definition, $\forall A, B \in T_1G$, $[LA, LB] = L[A, B]_{\text{Lie}}$. Let $\Xi, \mathrm{H} \in \mathcal{L}(G)$. Then, $[\Xi, \mathrm{H}](1) = L[\Xi(1), \mathrm{H}(1)]_{\text{Lie}}(1) = [\Xi(1), \mathrm{H}(1)]_{\text{Lie}}$. $\square$

What is so special about the Left action? We could equally define the Lie bracket in terms of the Right action, defined analogously. Suppose $A \in T_1G$, and let $RA(g) := dR(g)_1 A$, where $R(g) : G \to G,\ h \mapsto hg$ is the right action of G on itself. Note $dR(g)_1 A \equiv Adg_1$ for simplicity. Let

$$[A, B]_{\text{Lie, Right}} := [RA, RB](g) dg_g^{-1}$$

It is a fact that $[A, B]_{\text{Lie, Right}} = -[A, B]_{\text{Lie}}$ (see Kolar *et al.* [27], 4.11).

How does the Lie bracket defined above give rise to the matrix Lie bracket used when we have a finite-dimensional representation of our Lie group? Matrix Lie groups in finite dimension appear as subgroups of $\text{GL}(n, \mathbb{R})$. The Lie bracket defined above then exactly coincides with the usual matrix bracket. We now show how to make this identification. $\text{GL}(n, \mathbb{R})$ is an open subset of $\mathbb{R}^{n^2}$, since $\text{GL}(n, \mathbb{R}) := \det^{-1} \mathbb{R}^*$, the preimage of the open set $\mathbb{R}^* := \mathbb{R} - \{0\}$ by a continuous function det : $\mathbb{R}^{n^2} \to \mathbb{R}$. The embedding $\mathrm{GL}(n,\mathbb{R}) \hookrightarrow \mathbb{R}^{n^2}$ thus defines a global chart, whose components we use as coordinates. In the usual matrix notation, we note these coordinates as $g^i_j : \mathrm{GL}(n,\mathbb{R}) \to \mathbb{R}$, $i, j = 1, ..., n$. Now, let $A \in \mathfrak{gl}(n,\mathbb{R}) := T_1\mathrm{GL}(n,\mathbb{R})$. Then, $LA(g) = g^i_j A^j_k \partial/\partial g^i_k$. It is straightforward to calculate the brackets using standard calculus rules:

$$\begin{aligned}[LA, LB](g) &= \big[g^i_j A^j_k \frac{\partial}{\partial g^i_k}, g^a_b B^b_c \frac{\partial}{\partial g^a_c}\big] = g^i_j A^j_k \frac{\partial}{\partial g^i_k} g^a_b B^b_c \frac{\partial}{\partial g^a_c} - g^a_b B^b_c \frac{\partial}{\partial g^a_c} g^i_j A^j_k \frac{\partial}{\partial g^i_k} \\ &= g^i_j \left(A^j_a B^a_k - B^j_a A^a_k \right) \frac{\partial}{\partial g^i_k} = g^i_j [A, B]^j_k \frac{\partial}{\partial g^i_k},\end{aligned}$$

where $[A, B]$ is the matrix bracket of A and B. An analogous calculation results in

$$[RA, RB](g) = \big[A^i_j g^j_k \frac{\partial}{\partial g^i_k}, B^a_b g^b_c \frac{\partial}{\partial g^a_c}\big] = \cdots = -[A, B]^i_j g^j_k \frac{\partial}{\partial g^i_k}.$$

A.3 Principal bundles

Since our exposition and proof of Radford's theorem in appendix C relies on the formalism of principal bundles, we will here define some concepts and state some result that are used later.

A *left principal G-bundle* $P \xrightarrow{\pi} X$ consists of a manifold P called the *total space*, a manifold X called the *base space*, a submersion π, a Lie group G and a smooth *right action* on P, which is a smooth map $\mu : G \times P \to P$, such that $\mu(g, \cdot) \circ \mu(h, \cdot) = \mu(\cdot, gh)$. We will denote $\mu(g, p) \equiv gp$. We require that $\forall g \in G : \pi(gp) = \pi(p)$

Let $p \in P$ and $\mu(\cdot, p) : G \to P,\ g \mapsto gp$. The *vertical subspace* of P at p is defined as

$$V_p := \mathrm{im}\ d(\mu(\cdot, p))_1 \subset T_pP.$$

This identifies the Lie algebra of G with a subspace of T_pP. A vector field on P is called vertical if is contained in the vertical subspace at every point of P.

A connection $\omega : TP \to \mathrm{Lie}\ G$ is defined pointwise as a right-inverse to $d(\mu(\cdot, p)_1$:

$$\omega_p \circ d(\mu(\cdot, p))_1 = \mathrm{id}_{\mathrm{Lie}\ G}$$

B Radford's proof

In this section, we quote and present the proof of Radford's theorem in the context of principal bundles. This appendix presents the original statement by Radford [45], and it's original proof. I have provided the details to the proof ginven in Radford's paper.

B.1 Statement of the Theorem

We quote the original statement by Radford [45]:

Let G be a Lie group, M a manifold, TM its tangent bundle, and $A : TM \to \mathrm{Lie}\ G$ a Lie algebra valued one-form. Suppose $u : [0, 1] \to M$ is a loop ($u(1) = u(0)$). Then, the solution to

$$\dot{g} = -gA(u(t));\ g(0) = 1 \tag{159}$$

is given by $g = \exp(\Omega_1(A))$, where

$$\Omega_1(A) := \sum_{l=1}^{\infty} \frac{(-1)^{l+1}}{l} \sum_{I_1,\ldots,I_l} \frac{\int du^{I_1} \cdots \int du^{I_l}}{|I_1| + \cdots + |I_l|} \nabla F_{I_1 \cdots I_l}, \tag{160}$$

where $F_{ij} = \partial_i A_j - \partial_j A_i - [A_i, A_j]$, $\nabla_i = [A_i, \cdot] - \partial_i$, $\partial_i := \partial/\partial x^i$, $\nabla F_{i_1 \cdots i_n} := \nabla_{i_1} \cdots \nabla_{i_{n-2}} F_{i_{n-1} i_n}$ and I_i is a multi-index of size $|I_i|$.

A priori, the only condition on A is that it be a Lie G-valued one-form on M, $A \in T^*M \otimes \text{Lie } G$, and no principal bundle has been mentioned. The proof in terms of the language of principal bundles is convenient, since we can canonically construct one that fits the situation. This however obscures the proof somewhat for the untrained eye. A proof not involving this superfluous formalism might be more enlightening, and was given in section 4.

In this appendix, we follow Radford, and present a detailed proof of the above in the principal bundle language.

The product $Q := M \times G$ can be given a principal bundle structure by defining the left action $h(x, g) := (x, hg)$ and projection $\pi(x, g) = x$. A connection $\omega : TQ \to \text{Lie } G$ can be constructed in terms of A: Let $v \oplus \gamma \in T_{(x,g)}Q$. Then,

$$\omega(v \oplus \gamma) = A(v) + g^{-1}\gamma.$$

Here, $g^{-1}\gamma := \pi_G^* \Gamma(\gamma)$, where Γ is the Maurer-Cartan form on G and $\pi_G : Q \to G$ is the projeciton $(x, g) \mapsto g$. The horizontal lift $h : \mathcal{X}(M) \to \mathcal{X}(Q)$ can be defined by

$$h(X)(x, y) \equiv X^h(x, g) := X(x) - gA(X(x)). \tag{161}$$

One verifies immediately that $\omega \circ h = 0$:

$$\omega \circ h(X) = \omega(X - gA(X)) = A(X) - A(X) = 0.$$

Let U be the time dependent vector field on M defined by $U(x) = \dot{u}^i \partial_i$. Now, the problem of finding g as defined in eq. (159) can be restated as solving

$$\dot{q} = U^h(q), \; q(0) = u(0). \tag{162}$$

Then, $g(1)$ is the unique element in G such that $q(1) = q(0)g(1)$.

We proceed with a general statement for finding the holonomy element corresponding to a loop in the base space of a principal bundle.

B.2 Derivation of Radford's expression

Let $Q \to M$ be a principal bundle with connection one-form ω. Let $u : [0, 1] \to M$ be a path in M. By the horizontal lift induced by ω, we define a time-dependent vector field on Q by $U_t^h := \dot{u}^h = \dot{u}^i \partial_i^h$, where $\partial_i = \partial/\partial x^i$, where x is a chosen chart on M.[1] Our goal is to integrate $\dot{q} = X_t^h(q)$, giving a path $q : [0, 1] \to Q$ such that $\pi(q(t)) = u(t)$. Let σ be a local section of the principal bundle, $\sigma : M \to Q$. Define the holonmy path $g : [0, 1] \to G$ by $q(t) = g(t)\sigma(u(t))$. Since $\sigma(u(t))$ is known, finding $q(t)$ is then equivalent to finding $g(t)$, which obeys the ODE $\dot{g} = -gA$. In the particular case that u is a loop in M, then $g(1)$ gives us the corresponding holonomy element. Then, we are able to identify Lie brackets in the Magnus expansion with covariant derivatives of the curvature, yielding Radford's expansion in terms of geometric invariants.

Let $q : [0, 1] \to Q$ be the path defined by $\dot{q} = U_t^h$, and let $\phi \in C^\infty(Q)$. Since

$$\frac{d}{dt}\phi(q(t)) = U_t^h \phi(q(t)),$$

we have

[1] The horizontal lift $(\cdot)^h : TM \to TP$ is a family of linear maps uniquely defined by ω, such that im $(\cdot)^h = \ker \omega$ and $\pi \circ (\cdot)^h = id_M$.

$$\phi(q(t)) = \phi(q(0)) + \int_0^t d\tau (U^h_\tau \phi)(q(\tau)).$$

Repeating this argument for $U^h_\tau \phi$, we have

$$\begin{aligned}(X^h_\tau \phi)(q(\tau)) &= (U^h_\tau \phi)(q(0)) + \int_0^\tau d\tau' \frac{d}{d\tau'}(U^h_\tau \phi)(q(\tau')) \\ &= (U^h_\tau \phi)(q(0)) + \int_0^\tau d\tau' (U^h_{\tau'} U^h_\tau \phi)(q(\tau')),\end{aligned}$$

such that

$$\phi(q(t)) = \phi(q(0)) + \int_0^t d\tau (U^h_\tau \phi)(q(0)) + \int_0^t d\tau \int_0^\tau d\tau' (U^h_{\tau'} U^h_\tau \phi)(q(\tau')).$$

Repeating this iteration yields the time ordered exponential:

$$\phi(q(t)) = \phi(q(0)) + \sum_{n=1}^\infty \int_0^t d\tau^1 ... \int_0^{\tau^{n-1}} d\tau^n (U^h_{\tau^n} ... U^h_{\tau^1} \phi)(q(0)) =: \mathrm{Ev}^t_{U^h} \phi(q(0)). \tag{163}$$

Equivalently, $\mathrm{Ev}^t_{U^h} : C^\infty(Q) \to C^\infty(Q)$ can be defined by $\mathrm{Ev}^t_{U^h}\phi(q_0) = \phi(\Phi^t_{U^h}(q_0))$, where $\Phi^t : \mathcal{X}(Q) \to \mathrm{Diff}(Q)$ is the flow of map.

We wish to express $\mathrm{Ev}^t_{U^h}$ as the formal exponential of some vector field Ω_t. We suppose Q to be regular, meaning that we suppose $\exp : \mathcal{X}(Q) \to \mathrm{Diff}(Q)$ to be well defined. Chen [13] showed that Ω_t is given by

$$\dot{\Omega}_t = f(ad_{\Omega_t}) U^h_t, \tag{164}$$

where $f(x) := x/(1 - e^{-x})$. The first few terms are

$$\Omega_t = \overline{U^h}_t - \frac{1}{2}\overline{[\overline{U^h}, U^h]}_t + \frac{1}{4}\overline{[[\overline{U^h}, U^h], U^h]}_t + \frac{1}{12}\overline{[\overline{U^h}, [\overline{U^h}, U^h]]}_t + ... \tag{165}$$

At this point, we have not required anything of U^h or made any assumption other than that the expressions involved are well defined.

Using the fact that $U^h_t = \dot{u}^i(t)\partial^h_i$, we have

$$\begin{aligned}\mathrm{Ev}^t_{U^h} &= 1 + \sum_{n=1}^\infty \int_0^t d\tau^1 \dot{u}^{i_1}(\tau^1) \cdots \int_0^{\tau^{n-1}} d\tau^n \dot{u}^{i_n}(\tau^n) \partial^h_{i_1} \cdots \partial^h_{i_n} \\ &= 1 + \sum_{n=1}^\infty \sum_{|I|=n} \int du^I \partial^h_I\end{aligned} \tag{166}$$

The coefficients $\int du^I$ satisfy Ree's theorem [46], which we render here:

Theorem B.1. *Let $\{X_i\}$ be a finite number of indeterminates. The series*

$$S := 1 + \sum_{n=1}^\infty a^{i_1 \cdots i_n} X_{i_1} \cdots X_{i_n} \tag{167}$$

is the exponential of a Lie series iff the coefficients a satisfy the shuffle algebra, e.g., $a^\alpha a^{i_1 \cdots i_n} = a^{\alpha i_1 \cdots i_n} + a^{i_1 \alpha i_2 \cdots i_n} + \cdots + a^{i_1 \cdots i_n \alpha}$.

For the first few a's this means $a^i a^j = a^{ij} + a^{ji}$, $a^i a^{jk} = a^{ijk} + a^{jik} + a^{jki}$. In our case, the coefficients in (166) indeed satisfy Ree's theorem, such that $\log \mathrm{Ev}^t_{U^h}$ is indeed a Lie series.

$$\begin{aligned}\log \mathrm{Ev}^t_{U^h} &= \sum_{l=1} \frac{(-1)^{l+1}}{l} \left[\sum_{n=1}^{\infty} \int_{u_t} du^I \partial^h_I \right]^l \\ &= \sum_{l=1} \frac{(-1)^{l+1}}{l} \sum_{|I_i|\geq 1} \int du^{I_1} ... \int du^{I_l} \partial^h_{I_1 ... I_l}.\end{aligned} \tag{168}$$

Since we know that this to be a Lie series, we can invoke Dynkin's theorem [35]:

Theorem B.2. *Let $\{X_i\}$ be a set of indeterminates. The series*

$$S := \sum_{n=1}^{\infty} a^{i_1 \cdots i_n} X_{i_1} \cdots X_{i_n} \tag{169}$$

is a Lie series iff $[\![S]\!] = S$, where $[\![\cdot]\!]$ is defined termwise by $[\![X_{i_1} \cdots X_{i_n}]\!] = (1/n)[X_{i_1} \cdots X_{i_n}] := (1/n)[X_{i_1}, ..., [X_{i_{n-1}}, X_{i_n}]...]$.

Applying this to 168, we finally get,

$$\Omega_t := \log \mathrm{Ev}^t_{U^h} = \sum_{l=1} \frac{(-1)^{l+1}}{l} \sum_{|I_i|\geq 1} \int du^{I_1} \cdots \int du^{I_l} \frac{[\partial^h_{I_1 ... I_l}]}{|I_1| + \cdots + |I_l|}. \tag{170}$$

Writing out the first few terms, we get unsurprisingly the same as in (165). Combining terms again,

$$\Omega_t = \log \mathrm{Ev}^t_{X^h} = \sum_{l=1} \frac{(-1)^{l+1}}{l} \sum_{|k_i|\geq 1} \frac{[\overline{U^h_{k_1}} \cdots \overline{U^h_{k_l}}]}{k_1 + \cdots + k_l}. \tag{171}$$

where

$$\overline{U^h_n} := \overbrace{\overline{U^h \cdots \overline{U^h \overline{U^h}}}}^{n \text{ times}}$$

The first couple of terms are

$$\Omega_t = \overline{U}_t + \frac{1}{2}[\overline{U, \overline{U}}]_t + \frac{1}{3}[\overline{U, [\overline{U, \overline{U}}]}]_t - \frac{1}{6}\left([\overline{U}, [\overline{U, \overline{U}}]]_t + [\overline{U}, [\overline{\overline{U}}, \overline{U}]]_t \right) \tag{172}$$

We have found a vector field Ω_t such that $\mathrm{Ev}^t_{U^h}\phi = \exp(\Omega_t)\phi$. We have the following theorem

Theorem B.3. *Let Q be a manifold, Z a time independent vector field on Q and $\phi \in C^\infty(Q)$. Then,*

$$(\exp(tZ)\phi)(q) = \phi(\Phi^t_Z(q)), \tag{173}$$

where we denote the flow of Z by Φ^t_Z. If Z is time dependent,

$$T\exp(\overline{Z})\phi(q) = \phi(\Phi^t_Z(q)) \tag{174}$$

where $T\exp(\overline{Z})$ is the time ordered exponential.

If Z is time dependent, we take Z_t to mean the time independent vector field obtained by evaluating Z at time t.

Proof. The LHS of (173) is $\phi(q) + (Z\phi)(q) + (Z(Z\phi))(q)/2! + ...$ The RHS is $\phi(\Phi^1_Z(q)) = \phi(q) + \int_0^1 dt(Z\phi)(\Phi^t_Z(q))$, since $d/dt\phi(\Phi^t_Z(q)) = (Z\phi)(\Phi^t_Z(q))$. Reiterating gives the desired result. □

We have shown that $\phi(q(t)) = (\mathrm{Ev}^t_{U^h}\phi)(q(0)) = \exp(\Omega_t)\phi(q(0)) = \phi(\Phi^1_{\Omega_t}(q(0))$, implying $q(t) = \Phi^1_{\Omega_t}(q(0))$. Let $e : \mathrm{Lie}\, G \to G$ denote the Lie exponential, mapping the Lie algebra of G into G. It will turn out that when $u(1) = u(0)$ then $\Phi^1_{\Omega_t}(q(0)) = e^{\xi}q(0)$ for some lie algebra element ξ whose expression we seek to find. On our way there, some crucial facts need to be established.

Proposition B.4. *Ω_t is G-equivariant, meaning that $\forall g \in G, q \in Q : \Omega_t(gq) = g\Omega_t(q)$.*

In order to establish the above, we need to return to the language of connections on principal bundles.

Let $A \in \mathrm{Lie}\, G$. Then, A_Q denotes a vector field on Q defined by $A_Q(q) := d/dt|_{t=0}e^{tA}q$.

We next quote the following theorem from Radford:

Theorem B.5. *Let X and Y be vector fields on M, W the curvature 2-form on Q, defined by $\forall X, Y \in \mathcal{X}(Q) : W(X, Y) := d\omega(X, Y) - [\omega(X), \omega(Y)]$. Let $f : Q \to \mathrm{Lie}\, G$ an Ad-equivariant function, meaning $\forall g \in G, q \in Q : f(gq) = Ad_g f(q)$. Then,*

$$\begin{aligned} [X^h, Y^h] &= [X, Y]^h - W_Q(X^h, Y^h) \\ [X^h, f_Q] &= (X^h f)_Q \end{aligned} \tag{175}$$

This means in particular that $[\partial^h_i, \partial^h_j] = -W_Q(\partial^h_i, \partial^h_j)$. Let $f : Q \to \mathrm{Lie}\, G$. We call f Ad-equivariant if $f(gq) = \mathrm{Ad}_g f(q)$. Let $X \in \mathcal{X}(Q)$. We call X G-equivariant if $X(gq) = dg_q X(q)$. One important property of W is that it is Ad-equivariant. By the following lemma,

Lemma B.6. *$f : Q \to \mathrm{Lie}\, G$ is Ad-equivariant iff f_Q is G-equivariant.*

Theorem B.5 together with the above lemma and the fact that W is Ad-equicariant, imply that $[\partial^h_I] = -(\partial^h_{i_1}...\partial^h_{i_{n-2}}\Omega(\partial^h_{n-1}, \partial^h_n))_Q$, and that all the brackets $[\partial^h_I]$ are in fact G-equivariant vector fields. In particular ∂^h_i is G-equivariant, and since Ω_t then is a sum of equivariant vector fields, this establishes that Ω_t is equivariant.

Let $X \in \mathcal{X}(Q)$. We call X vertical if $d\pi \circ X = 0$. We call X horizontal if $d\pi(X - \omega_Q(X)) = 0$. It can be shown that, for a vertical vector field X, $\omega_Q X = X$.

In the case that $u(1) = u(0)$, Ω_1 gains the key property, that it is vertical. This will allow us to identify Ω_1 with a unique $\xi \in \mathrm{Lie}\, G$.

Proposition B.7. *If $u(1) = u(0)$, then Ω_1 is a vertical vector field.*

Proof. In the expansion of Ω_t in equation (165), the only horizontal part is the terms $\overline{U^h_t}$. This is because the bracket of any two horizontal vector fields is vertical, and the bracket of a horizontal and a vertical vector field is again vertical. When $u(1) = u(0)$, the horizontal terms become zero, and Ω_1 is purely vertical. □

If Ω is equivariant, then $\omega(\Omega) : Q \to \mathrm{Lie}\, G$ is Ad-equivariant: $\omega(\Omega(gq)) = \omega(gZ(q)) = Ad_g\omega(Z(q))$. Now, if $f : Q \to \mathrm{Lie}\, G$ is Ad-equivariant, then f_Q is a G-equivariant vertical vector field: $f_Q(gq) = d/dt\; e^{tf(gq)}gq = d/dt\; e^{Ad_g tf(q)}gq = d/dt\; ge^{tf(q)}q = gf_Q(q)$.

Theorem B.8. *Let $f : Q \to \mathrm{Lie}\, G$ be Ad-equivariant. Then, $\Phi^1_{f_Q}(q) = e^{f(q)}q$.*

Proof. Let $\phi \in C^\infty(Q)$. Taylor expanding the LHS gives

$$\phi(\Phi_{f_Q}(q, 1)) = \phi(q) + (f_Q\phi)(q) + (f_Q(f_Q\phi))(q)/2! + \cdots,$$

and Taylor expanding the RHS gives

$$\phi(e^{f(q)t}q) = \phi(q) + \cdots + d^n/dt^n|_0\ \phi(e^{f(q)t}q)/n! + \cdots .$$

By definition of the vector field f_Q, we have $d/dt|_0\ \phi(e^{f(q)t}q) = (f_Q\phi)(q)$, and the higher derivatives are $d^n/dt^n|_0\ \phi(e^{f(q)t}q) = (f_Q \cdots f_Q\phi)(q)$, which equates the two sides. $\square$

This then implies the following fact: Let Ω_1 be a G-equivariant vertical vector field, then there is a unique Ad-equivariant function $\xi : Q \to \mathrm{Lie}\, G$ s.t. $\Omega_1 = \xi_Q$. In that case, $\Phi^1_{\Omega_1}(q) = e^{\xi(q)}q$. The desired Lie element ξ is then given by $\omega(\Omega_1(q(0)))$. By theorem B.5, $\omega([\partial^h_I]) = -\partial^h_{i_1}...\partial^h_{i_{n-2}}\Omega(\partial^h_{n-1}, \partial^h_n)$.

Suppose $h : M \to \mathrm{Lie}\, G$. We can define the covariant derivative of h as $\nabla h := dh - [A, h]$. If X is a vector field on M, then, denote $\nabla_X h \equiv \nabla h(X) = dh(X) - [A(X), h]$.

Lemma B.9. *Let $f : Q \to \mathrm{Lie}\, G$ be Ad-equivariant and X a vector field on M. Then, $\sigma^*(X^h f) = \nabla_X \sigma^* f$*

Proof. $\nabla_X\sigma^* f = \sigma^* df(X) - [A(X), \sigma^* f] = \sigma^*(df(\sigma X) - [\omega(\sigma X), f]) = \sigma^*(\sigma X \cdot f - \omega_Q(\sigma X) \cdot f) = \sigma^*((\sigma X - \omega_Q(\sigma X)) \cdot f) = \sigma^* X^h f$, were we denote $\sigma X := (q \mapsto g(q)d\sigma_{\pi(q)}X(\pi(q))$. We used the fact that for f Ad-equivariant, $[\omega(X), f] = \omega_Q(X) \cdot f$. This follows from $[\omega(X), f](q) = d/dt|_0\ Ad\{e^{\omega(X)t}\}f(q) = d/dt_0 f(e^{\omega(X)t}q) = \omega_Q(X)f(q)$. $\square$

Let $F := \sigma^* W$. The previous lemma then immediately implies

$$\sigma^*\partial^h_{i_1}...\partial^h_{i_{n-2}}W(\partial^h_{n-1}, \partial^h_n) = \nabla_{i_1}...\nabla_{i_{n-2}}F_{n-1,n},$$

such that $\sigma^*\omega[\partial^h_I](q(0)) = -\nabla_{i_1}...\nabla_{i_{n-2}}F_{n-1,n}(u(0))$[1] The Lie element we seek to find is

$$\begin{aligned}
\xi = \omega(Z_T(q(0))) &= \omega \sum_{l=1} \frac{(-1)^{l+1}}{l} \sum_{|I_i|\geq 1} \int du^{I_1} \cdots \int du^{I_l} \frac{[\partial^h_{I_1\cdots I_l}](q(0))}{|I_1| + \cdots + |I_l|} \\
&= \sum_{l=1} \frac{(-1)^{l+1}}{l} \sum_{|I_i|\geq 1} \int du^{I_1} \cdots \int du^{I_l} \frac{\omega[\partial^h_{I_1\cdots I_l}](q(0))}{|I_1| + \cdots + |I_l|} \\
&= \sum_{l=1} \frac{(-1)^{l+1}}{l} \sum_{|I_i|\geq 1} \int du^{I_1} \cdots \int du^{I_l} \frac{\nabla F_{I_1\cdots I_l}(x(0))}{|I_1| + \cdots + |I_l|},
\end{aligned} \tag{176}$$

where $\nabla F_I = -\nabla_{i_1} \cdots \nabla_{i_{n-2}}F_{n-1,n}$. Using the fact that we are integrating over a closed path, we have

$$\int du^i = 0,$$

so all first order integrals vanish. The sum can then be rewritten as

$$\xi = \sum_{j=2}^{\infty} \frac{1}{j} \sum_{l=1}^{\lfloor j/2 \rfloor} \frac{(-1)^{l+1}}{l} \sum_{|I|=j;\ |I_i|\geq 2} \int du^{I_1} \cdots \int du^{I_l} \nabla\partial_{I_1\cdots I_l}(x(0)). \tag{177}$$

Note $U^h \equiv X$. Then,

$$\Omega_1 = \frac{1}{2}\overline{[X,\overline{X}]} + \frac{1}{3}\overline{[X,\overline{[X,\overline{X}]}]} + \frac{1}{4}\left(\overline{[X,\overline{[X,\overline{[X,\overline{X}]}]}]} - \frac{1}{2}\overline{[X,\overline{[\overline{X},\overline{[X,\overline{X}]}]}]}\right) + \cdots , \tag{178}$$

where $\overline{[X,\overline{X}]} = \int_0^T d\tau[X_\tau, \int_0^\tau d\tau' X_{\tau'}]$ etc. The corresponding Lie algebra series $\xi = \omega(\Omega_1)$ is

$$\xi = -\frac{1}{2}F_{ij}\int du^{ij} - \frac{1}{3}\nabla_i F_{jk}\int du^{ijk} - \frac{1}{4}\nabla_i\nabla_j F_{kl}\left(\int du^{ijkl} - \frac{1}{2}\int du^{ij}\int du^{kl}\right) + \cdots . \tag{179}$$

[1]This is lemma 2, Ch. III.9, in [26].