Solving ODE on Manifolds

I. Introduction

There are two main approaches to solving ordinary differential equations (ODE) using numerical integration on manifolds: extrinsic and intrinsic. In the extrinsic approach, the manifold in embedded in a higher dimensional Euclidean space which allows the use of the classical numerical integration techniques used in any vector space; however, the solution produced by these techniques are not guaranteed to stay on the manifold, and usually requires correction by projecting the solution back onto the manifold. In the intrinsic approach, the numerical integration technique uses properties of a differentiable manifold to evolve the solution of the ODE on the manifold. Because the intrinsic approach naturally evolves on the manifold, the solution incurs less error and never needs to be project back onto the manifold; however, it is typically more computationally cumbersome due to the need of the exponential map.

Fortunately, computers are becoming more sophisticated and better able to handle more computations, and as robotics advances there is a greater need for higher accuracy. Therefore intrinsic solutions of ODE is becoming not only computationally possible, but also necessary. The objective of this note is to present intrinsic numerical integration methods discovered by other researchers in an easy way to understand for the purposes of small unmanned air systems (sUAS) with examples. We will try to present them in a general case first, but then quickly apply them to sUAS. Our belief is that by seeing applications to sUAS along the way will improve understanding, versus exposing you to all the theory upfront and then presenting an application at the end.

The majority of the material that we will present will be from Crouch and Grossman [4], Munthe-Kaas [10], [11], [12], and Engø [5], [6]. Crouch and Grossman were one of the first to present the Runge-Kutta (RK) method to evolve on manifolds up to the third order. Their approach does the approximation on the Lie group which results in it being complicated and limited in the order. Munthe-Kaas took a similar approach, but instead of performing the RK approximations in the Lie group, the approximation is performed in the Lie algebra and then mapped to the Lie group. This approach allows the use of the classical RK method up to any order. Engø built upon Munthe-Kaas work by expanding it to other coordinates, and then developed the partitioned Runge-Kutta Muthe-Kaas (RK-MK) method used to solve partitioned differential equations which will be applicable to solving second order ODEs.

For additional references about implicit numerical integration on manifolds we refer the reader to [3] which provides an introduction to different methods, and to [8] which provides an extensive survey and overview of modern techniques. It is assumed that the reader has a basic understanding on theory of smooth manifolds, Lie groups and Lie algebras. Even though we will recall some of it in this note, we refer the reader to [1], [9] for a more in depth exposure on the subjects. Everything presented here is a compilation of the works of the authors already mentioned except for the multirotor example. All we hope to achieve is to add clarity and instruction on how to apply this theory to sUASs.

In this section we recall basic definitions that are paramount to understanding this material. Since we are simply recalling these definitions, we won't present them in their most abstract form, rather a simplified form that is sufficient for the needs of this note. For more precise abstract definitions we refer the reader to the authors mentioned in the introduction.

II. Examples

We will consider four examples throughout this note. The first exampled will be explored throughout the documentation, and the other examples will be presented fully in the appendix.

Example 1. Consider a pendulum that is unit length and constrained to move in a plane. This pendulum moves on the manifold $S^1 := \left\{(x,y) \in \mathbb{R}^2 | x^2+y^2=1\right\}$. The ordinary differential equation for this example will be developed in section III.

Example 2. Let $v \in S^2 := \left\{(x,y,z) \in \mathbb{R}^3 | x^2+y^2+z^2=1\right\}$ represent a point on the unit sphere in $\mathbb{R}^3$ and $\omega \in \mathfrak{X}^r\left(\mathbb{R}^3\right)$ denote the vector field on S^2, then a possible first order differential equation is

$$\dot{v} = \omega\left(v\right)$$

where ω is a skew symmetric vector field.

Example 3. Let $y \in SO(3)$ with the vector field $a(y) \in \mathfrak{X}^r(SO(3))$ defined as

$$a(y) = \frac{1}{2}\left(y - y^\top\right)$$

then the corresponding ODE is

$$\dot{y} = a(y)\, y.$$

This example is taken from [5], and similar variations have been used in both [12] and [8].

Example 4. A simplified version of the second order ODE of a multirotor using body-centric representation is

$$\begin{aligned} \dot{q} &= q v^\wedge \\ \mathcal{I}\dot{v} &= \mathrm{ad}_v^* \mathcal{I} v + F(v,q) \end{aligned}$$

where $q \in SE(3)$ is the pose of the multirotor, v is the generalized velocity of the rigid body w.r.t. the inertial frame expressed in the body frame, ad_v^* is the dual of the adjoint of v, $\mathcal{I}$ is the generalized inertial tensor, and $F(v,q) : TQ \to T^*Q$ is the force. In this example, the generalized velocity v has the form

$$v = \begin{bmatrix} \rho \\ \omega \end{bmatrix}$$

where $\rho \in \mathbb{R}^3$ denotes the translational velocity and $\omega \in \mathbb{R}^3$ denotes the angular velocity. The generalized inertial tensor in matrix form is a diagonal matrix with components

$$\mathcal{I} = \mathrm{diag}(m, m, m, J_x, J_y, J_z)$$

where m is the mass of the multirotor and $J_i\, i \in \{x, y, z\}$ are the moments of inertial about each axis assuming the multirotor is expressed in an NED frame. The dual adjoint of v has matrix form

$$\mathrm{ad}_v^* = \begin{bmatrix} \omega_\times & \rho_\times \\ & \omega_\times \end{bmatrix}^\top$$

where $\omega_\times$ and $\rho_\times$ are the skew symmetric matrices formed from ω and u corresponding to the body-centric representation, i.e.,

$$u_\times = \begin{bmatrix} 0 & -u_z & u_y \\ u_z & 0 & -u_x \\ -u_y & u_x & 0 \end{bmatrix}$$

where $u \in \{\omega, \rho\}$. The covector field $F(v,q)$ is

$$F(v,q) = \begin{bmatrix} mgR^\top e^i_{z/i} + T e^b_{z/b} \\ M \end{bmatrix}$$

where g is the gravitational constant, m is the mass, $R \in SO(3)$ is the rotation component of q, T is the thrust from the propellers, M is the moment produced by the propellers, $e^i_{z/i}$ is the unit vector along the z-axis of the inertial frame expressed in the inertial frame, and $e^b_{z/b}$ is the unit vector along the z-axis of the body frame expressed in the body frame. $e^i_{z/i}$ and $e^b_{z/b}$ have the matrix form $\begin{bmatrix} 0 & 0 & 1 \end{bmatrix}^\top$.

For our purposes, we will let the generalized inertial tensor have arbitrary values

$$\mathcal{I} = \mathrm{diag}(10, 10, 10, 0.5, 0.5, 0.1)$$

and the gravitational constant g will have value 9.81.

III. Fundamentals

A. Tangent Bundle

Consider the pendulum in example 1. Since the pendulum is constrained to rotate on a plane, the system evolves on the topological space $S^1 := \left\{(x,y) \in \mathbb{R}^2 | x^2 + y^2 = 1\right\}$, i.e., the unit circle. Let the pendulum's motion be defined by a curve $\gamma(t)$ that is a continuous map $\gamma : I \to S^1$ defined as

$$\gamma(t) = \left(\cos(t)\, x_0 - \sin(t)\, y_0, \sin(t)\, x_0 + \cos(t)\, y_0\right), \tag{1}$$

where I is an open interval on $\mathbb{R}$, and $(x_0, y_0) \in S^1$ is a fixed point in S^1. We can also express the curve in a more familiar form.

$$x\left(t\right) = \cos\left(t\right) x_0 - \sin\left(t\right) y_0 \tag{2a}$$

$$y\left(t\right) = \sin\left(t\right) x_0 + \cos\left(t\right) y_0. \tag{2b}$$

where $x\left(t\right)$ and $y\left(t\right)$ are diffeomorphism from the interval I to $\mathbb{R}$.

In order to form the corresponding ordinary differential equation, we need the derivative of the curve. Recall that the derivative of a function is defined as

$$\frac{d}{dt} f\left(t\right) \coloneqq \lim_{h \to 0} \frac{f\left(t+h\right) - f\left(t\right)}{h}. \tag{3}$$

Using the definition of a derivative we can find the derivative of the curve $\gamma\left(t\right)$ as

$$\frac{d}{dt} \gamma\left(t\right) \coloneqq \lim_{h \to 0} \frac{\gamma\left(t+h\right) - \gamma\left(t\right)}{h}.$$

This expression for the derivative of $\gamma\left(t\right)$ is subtly misleading. The terms $\gamma\left(t+h\right)$ and $\gamma\left(t\right)$ are elements of S^1, and the topological space has no addition operation defined. This means that the derivative cannot be computed in S^1. One possible solution is to embed S^1 into $\mathbb{R}^2$ which is a vector space that has addition operator defined and take the derivative in $\mathbb{R}^2$. This solution is an extrinsic method because it uses a larger dimensional space than the manifold. Extrinsic methods are good methods, but they generally produce larger errors than their counterpart intrinsic methods. Intrinsic methods uses local objects that support differentiation and then patches them together. In order to do this, we need to add additional structure to a topological space that maps subsets of the topological space to $\mathbb{R}^n$ in a way that supports differentiation.

Definition 5. (**Chart**). Let T be a topological space. A chart is a bijective map ϕ from an open subset U of T to an open subspace of $\mathbb{R}^n$. The chart is often identified as (U, ϕ) to indicate the domain on which the chart operates. A family of compatible charts whose union of domain is a cover of T forms an **atlas**.

Definition 6. (**Manifold**). A topological manifold (or manifold for short) is a topological space equipped with an atlas. A **differentiable manifold** is a manifold whose charts are C^k- diffeomorphisms. I refer the readers to [1] for a more precise definition.

Let's return to example 1 where the topological space is S^1. We can equip it with the chart

$$\phi : S^1 \to \mathbb{R},\ (x, y) \mapsto \arctan\left(y/x\right) \tag{4}$$

whose domain is $S^1 \setminus \{(0, -1)\}$. Note that we haven't created an atlas since we have not constructed a family of compatible charts whose union of domain is a cover of S^1. But this is not necessary for our discussion. We can form the composition of the curve $\gamma\left(t\right)$ and the chart ϕ to form the composite map

$$\phi \circ \gamma\left(t\right) : I \to \mathbb{R},\ t \mapsto \arctan\left(y\left(t\right)/x\left(t\right)\right),$$

where $x\left(t\right)$ and $y\left(t\right)$ are defined in (2). Since the composite map $\phi \circ \gamma$ is a homeomorphism whose co-domain is a vector space, we can take the derivative with respect to (w.r.t.) t.

$$\frac{d}{dt} \phi \circ \gamma\left(t\right) = \lim_{h \to 0} \frac{\phi \circ \gamma\left(t+h\right) - \phi \circ \gamma\left(t\right)}{h} \tag{5a}$$

$$= \frac{d}{dt} \arctan\left(y\left(t\right)/x\left(t\right)\right) \tag{5b}$$

$$= \frac{\partial \arctan\left(y\left(t\right)/x\left(t\right)\right)}{\partial x} \frac{dx}{dt} + \frac{\partial \arctan\left(y\left(t\right)/x\left(t\right)\right)}{\partial y} \frac{dy}{dt} \tag{5c}$$

$$= \frac{y\left(t\right)}{x^2\left(t\right) + y^2\left(t\right)} \frac{dx}{dt} - \frac{x\left(t\right)}{x^2\left(t\right) + y^2\left(t\right)} \frac{dy}{dt} \tag{5d}$$

$$= y\left(t\right) \frac{dx}{dt} - x\left(t\right) \frac{dy}{dt} \tag{5e}$$

$$= y^2\left(t\right) + x^2\left(t\right) \tag{5f}$$

$$= 1. \tag{5g}$$

The derivative should make sense for our example since the map $\arctan$ maps the components of the ordered pair (x, y) to the angle between them, and since the curve $\gamma(t)$ is rotating the point (x_0, y_0) at a constant angular velocity of 1, we would expect the derivative of the angle to be 1. The derivative that we calculated is a tangent vector at the point $\phi \circ \gamma(t)$ which can be written as $T_{\phi\circ\gamma(t)}\mathbb{R} = 1$ where $T_{\phi\circ\gamma(t)}\mathbb{R}$ indicates the tangent space of $\mathbb{R}$ at the point $\phi \circ \gamma(t)$.

Definition 7. (**Tangent Space**). Let M be a manifold and $m \in M$, the tangent space of M at the point m is denoted $T_m M$ is the set of all tangent vectors at the point m.

Remark 8. There are many different ways to define the tangent space at a point on a manifold. We merely provide a simplified version of the coordinate approach. For other approaches see [1], [9].

So far for example 1 we have only defined one tangent vector at the point $\phi \circ \gamma(t)$. There are infinitely more tangent vectors that we could define at this point which spans all of $\mathbb{R}$. This means that $T_{\phi\circ\gamma(t)}\mathbb{R} = \mathbb{R}$, i.e., a vector space of $\mathbb{R}$ over the field of real numbers. In fact at each point $p \in \mathbb{R}$ there is a tangent space $T_p\mathbb{R}$. The union of all of these disjoint tangent spaces

$$T\mathbb{R} = \cup_{p\in\mathbb{R}} T_p\mathbb{R}(\text{disjoint}),$$

is called the tangent bundle of $\mathbb{R}$ and is the ordered pair $(p, v) \in T\mathbb{R}$ where p is a point in $\mathbb{R}$ and v is a tangent vector at p. This means that $T\mathbb{R} = \mathbb{R} \times \mathbb{R}$.

Definition 9. (**Tangent Bundle**). Let M be a manifold, $m \in M$ and $T_m M$ be the tangent space at the point m. The tangent bundle $TM = \cup_{m\in M} T_m M$ is the disjoint union of all of the tangent spaces. The projection map π_M is the map from the tangent bundle to the manifold. For example let $(m, v) \in TM$ where m is a tangent vector at the point m, then $\pi_M(m, v) = m$.

Returning to example 1. We have found the tangent vector of the curve γ in local coordinates to be the constant value 1. This vector resides in $T_{\phi\circ\gamma(t)}\mathbb{R}$, but the question is how to map it to $T_{\gamma(t)}S^1$. In order to do this, we need to construct the tangent bundle of our manifold and local coordinates, and establish a diffeomorphic map between them.

A quick aside on maps. A bijective map is a surjective and injective map which allows us to map elements from one space to another and back again. A homeomorphic map is a continuous bijective map. This is necessary when taking limits to ensure that map isn't "jumping around" in the co-domain. A diffeomorphic map is a differentiable homeomorphic map. The differential of a homeomorphic map, sometimes called a tangent or differential map, maps tangent vectors between spaces.

Definition 10. (**Tangent of f**) Let M and N be two manifolds with tangent bundles TM and TN. If $f : M \to N$ is of class C^1, then the tangent map or differential map is the map $Tf : TM \to TN$.

The tangent map $T\phi$ in example 1 is easy to calculate since we already calculated it in (5).

$$T\phi : TS^1 \to T\mathbb{R},\ ((x, y), (\dot{x}, \dot{y})) \to y\dot{x} - x\dot{y}, \tag{6}$$

or in matrix notation

$$\begin{bmatrix} y & -x \end{bmatrix} \begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix}.$$

The inverse tangent map is little tricky to calculate. We need to make use of the constraint on S^1 that $x^2 + y^2 = 1$. Recall that if we have a curve $\gamma(t) : I \to S^1$, then we can write an element of S^1 as $(x(t), y(t))$ as in (2), and write the constraint as $x^2(t) + y^2(t) = 1$. Taking the derivative of this constraint yields a holonomic constraint on the velocities

$$x\dot{x} + y\dot{y} = 0. \tag{7}$$

We also need to utilize the inverse map $\phi^{-1} : \mathbb{R} \to S^1$ which is defined as

$$x = \frac{1}{\sqrt{\tan^2(\theta) + 1}} \tag{8a}$$

$$y = \frac{\tan^2(\theta)}{\sqrt{1 + \tan^2(\theta)}}, \tag{8b}$$

where $\theta = \phi(x, y)$. Starting with the tangent map $T\phi$ we have

$$\dot{\theta} = y\dot{x} + x\dot{y},$$

where $\dot{\theta} = \frac{d}{dt}\phi \circ \gamma(t)$, and the dependence on t is implicit. Using the constraint (7) with $T\phi$ we get

$$\begin{aligned} \dot{x} &= \dot{\theta} y \\ \dot{y} &= -\dot{\theta} x, \end{aligned}$$

and finally using the definition of ϕ^{-1} we get

$$\begin{aligned} \dot{x} &= \dot{\theta}\frac{\tan^2(\theta)}{\sqrt{1+\tan^2(\theta)}} \\ \dot{y} &= -\dot{\theta}\frac{1}{\sqrt{\tan^2(\theta)+1}}. \end{aligned}$$

Using matrix notation we define the inverse tangent map $T\phi^{-1} : T\mathbb{R} \to TS^1$ as

$$\left(\theta, \dot{\theta}\right) \mapsto \begin{bmatrix} \frac{\tan^2(\theta)}{\sqrt{1+\tan^2(\theta)}} \\ -\frac{1}{\sqrt{\tan^2(\theta)+1}} \end{bmatrix} \dot{\theta}. \tag{9}$$

We can now use the inverse tangent map to map the tangent vector $\dot{\theta} = 1$ at θ found in (5) to $T_{\gamma(t)}S^1$. This is merely

$$\begin{bmatrix} -\frac{\tan^2(\theta)}{\sqrt{1+\tan^2(\theta)}} \\ \frac{1}{\sqrt{\tan^2(\theta)+1}} \end{bmatrix}.$$

Using ϕ^{-1} we can express it in terms of x and y,

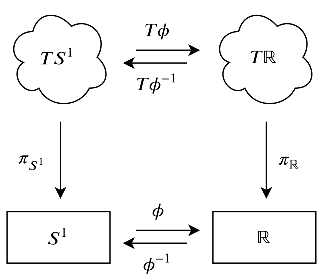
$$\begin{bmatrix} y \\ -x \end{bmatrix},$$

which implies that

$$\frac{d}{dt}\gamma(t) = \begin{bmatrix} y(t) \\ -x(t) \end{bmatrix},$$

where $x(t)$ and $y(t)$ are defined in (2).

Putting everything together that we have so far for example (1) we have the following diagram which shows how to map between the manifolds S^1 and $\mathbb{R}$, and their corresponding tangent bundles.



The last thing we need to talk about are vector fields and flows in order to complete the diagram.

$$\begin{tikzcd} TM \arrow[r, "T\phi"] & TN \\ M \arrow[u, "X"] \arrow[r, "\phi"'] & N \arrow[u, "Y"'] \end{tikzcd}$$

B. Vector Fields and Flows

Definition 11. (**Vector field**). A vector field $X \in \mathfrak{X}(M)$ is a map $X : M \to TM$ from the manifold M to the tangent space TM.

In example 1 we have already calculated a tangent vector in S^1 to be

$$\begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \begin{bmatrix} y \\ -x \end{bmatrix},$$

thus the vector field associated with this tangent vector is

$$X(x, y) = \begin{bmatrix} y \\ -x \end{bmatrix}. \tag{10}$$

It turns out that this specific vector field is a basis of all possible vector fields on S^1. Thus all vector fields on S^1 can be written in as

$$f(t, (x, y)) X(x, y)$$

where $f : \mathbb{R} \times S^1 \to \mathbb{R}$. The corresponding vector field in local coordinates is

$$Y(\theta) = 1, \tag{11}$$

which is also a basis for all possible vector fields on $\mathbb{R}^1$.

The two vector fields X and Y are ϕ-related.

Definition 12. (**ϕ-related**). Let $\phi : M \to N$ be a C^r mapping of manifolds. The vector fields $X \in \mathfrak{X}^{r-1}(M)$ and $Y \in \mathfrak{X}^{r-1}(N)$ are called ϕ-related, denoted $X \sim_\phi Y$, if $T\phi \circ X = Y \circ \phi$, i.e., the following diagram commutes

Showing that $X \in \mathfrak{X}\left(S^1\right)$ and $Y \in \mathfrak{X}(\mathbb{R})$ are ϕ-related is rather trivial since we have already defined all of the vector fields, the map, and the tangent map. Using (4), (6), (10), and (11) we show that $X \sim_\phi Y$,

$$\begin{aligned} T\phi \circ X &= Y \circ \phi \\ y\dot{x} - x\dot{y} &= Y \circ \arctan(y/x) \\ y^2 + x^2 &= 1 \\ 1 &= 1. \end{aligned}$$

In other words, the following diagram commutes

$$\begin{tikzcd} TS^1 \arrow[r, "T\phi"] & T\mathbb{R} \\ S^1 \arrow[u, "X"] \arrow[r, "\phi"] & \mathbb{R} \arrow[u, "Y"'] \end{tikzcd}$$

Since ϕ in example 1 is a diffeomorphism, the vector fields X and Y are not only ϕ-related, but also ϕ^{-1}-related, i.e., the following diagram commutes

$$\begin{tikzcd} TS^1 \arrow[r, shift left, "T\phi"] \arrow[d, shift right, "\pi_{S^1}"'] & T\mathbb{R} \arrow[l, shift left, "T\phi^{-1}"] \arrow[d, shift left, "\pi_{\mathbb{R}}"] \\ S^1 \arrow[u, shift right, "X"'] \arrow[r, shift left, "\phi"] & \mathbb{R} \arrow[l, shift left, "\phi^{-1}"] \arrow[u, shift left, "Y"] \end{tikzcd}$$

The vector field X from example 1 was derived from the curve $\gamma(t)$ defined in (1). This curve is called integral curve of $X \in \mathfrak{X}\left(S^1\right)$ at (x_o, y_o).

Definition 13. (**Integral curve**). An integral curve of $X \in \mathfrak{X}(M)$ at $m \in M$ is a curve $\gamma(t)$ such that $\dot{\gamma}(t) = X(\gamma(t))$ where t is an element of the interval $I \subseteq \mathbb{R}$ that includes the 0 element and $\gamma(0) = m$.

Finding integral curves is the same as finding the solution to an ODE. According to the **Local Existence, Uniqueness, and Smoothness Theorem** [1, THM 4.1.5], if a vector field $X \in \mathfrak{X}(M)$ is a smooth vector field, then there is a unique integral curve $\gamma(t)$ of X at m, and this integral curve is also unique in local coordinates. This theorem says that any smooth ODE has a unique integral curve (i.e., solution) in any local coordinates. This theorem is very applicable because we often do not start out with an integral curve and derive the vector fields as we did for example 1, rather we start out with an ODE and solve for the integral curve. This theorem tells us when a unique solution exists. This is very helpful even in circumstances where we cannot solve the ODE, because we can still move along the integral curve by moving along the vector fields in infinitesimal increments knowing that it is unique.

An integral curve might not exist at every point $m \in M$ or for all $t \in \mathbb{R}$. For example consider the manifold $M = \mathbb{R}^2 \backslash \{0\}$ with the vector field $\frac{\partial}{\partial x}$. The corresponding integral curve is $\gamma(t) = (x + t, y)$ for any $(x, y) \in M$. The integral curve exists until $(x + t) = 0$ since the curve is no longer on the manifold at this time. This means that the domain of the integral curve is $t \in (x, \infty)$ for $x < 0$ or $t \in (-\infty, -x)$ for $x > 0$.

Definition 14. (**Complete Vector Field**). Given a manifold M and a vector field X, let $\mathcal{D}_x \subseteq \mathbb{R} \times M$ be the set of the ordered pair $(m, t) \in \mathbb{R} \times M$ such that there is an integral curve $\gamma(t) : I \to M$ of X at m with $t \in I \subseteq \mathbb{R}$. If the set $\mathcal{D}_x = \mathbb{R} \times M$, then the vector field is called complete.

What this definition is saying is, if you start at any point in M and move along the vector field for any $t \in \mathbb{R}$ you will always stay in M. Showing that a vector field is complete can be very difficult. Fortunately there are theorems [9, Theorem 9.16 and Theorem 9.18] that show that on a compact smooth manifold, every smooth vector field is complete, and every left-invariant vector field on a Lie group is complete. Similarly, it can be proven that every right-invariant vector field on a Lie group is also complete.

In example 1, the manifold S^1 is a compact smooth manifold which means that every vector field on this manifold is complete and every associated integral curve exist at every point $s \in S^1$ for all $t \in \mathbb{R}$.

An object that is closely related to integral curves are flows.

Definition 15. (**Flow**). A flow on M is a continuous map $F : \mathbb{R} \times M \to M$ whose domain is called a **flow domain**. The flow domain is the open subset $\mathcal{D} \subseteq \mathbb{R} \times M$ with the properties that for each $(t, m) \in \mathcal{D}$, $F(t, m) \in M$, and for each $m \in M$, the set $\mathcal{D}^m := (t \in \mathbb{R} | (t, m) \in \mathcal{D})$ is an open interval containing 0. If the domain of a flow $\mathcal{D} = \mathbb{R} \times M$, the flow is called a **global flow**, otherwise the flow is called a **local flow**. A flow satisfies the group laws: for all $m \in M$,

$$F(0, m) = m,$$

and for all $s \in \mathcal{D}^m$ and $t \in \mathcal{D}^{F(t,m)}$ such that $s + t \in \mathcal{D}^m$,

$$F(t, F(s, m)) = F(t + s, m).$$

We can evaluate a flow at a specific point $m \in M$ or at a specific point $t \in \mathbb{R}$ to create other maps. These maps are defined as $F^t : M \to M, \mapsto F(t, \cdot)$, and $F^m : \mathbb{R} \to M$, $m \mapsto F(\cdot, m)$. By definition, F^m is an integral curve of a vector field at the point m.

If F is a smooth flow, the tangent vector $v_m \in T_m M$ corresponding to the flow F, is defined as

$$v_m = \left.\frac{d}{dt}\right|_{t=0} F(t, m). \tag{12}$$

The assignment $m \mapsto v_m$ is a smooth vector filed X on M and is called an **infinitesimal generator of F**. It is called an infinitesimal generator of **F** since the flow can be approximated by piecing together the tangent vectors generated by the vector field as you move along M. Also, F^m is the integral curve of X at m.

Returning to example 1 we defined an integral curve $\gamma(t)$ by (1), the corresponding flow is

$$\begin{aligned} &F : \mathbb{R} \times M \to M, \\ &(t, (x_o, y_o)) \mapsto (\cos(t)\, x_0 - \sin(t)\, y_0, \sin(t)\, x_0 + \cos(t)\, y_0). \end{aligned}$$

The infinitesimal generator is constructed using 12

$$\begin{aligned} X(s) &= \left.\frac{d}{dt}\right|_{t=0} F(t, \cdot) \\ &= (-\sin(t)\, x_0 - \cos(t)\, y_0, \cos(t)\, x_0 - \sin(t)\, y_0), \end{aligned}$$

or in matrix notation

$$X(s) = \begin{bmatrix} -\sin(t) & -\cos(t) \\ \cos(t) & -\sin(t) \end{bmatrix} \begin{bmatrix} x_0 \\ y_0 \end{bmatrix},$$

where $s = (x_0, y_0)$.

In a similar way in which vector fields on different manifolds can be related via a map, flows can be related via a map. To discuss this, we need to define additional notation. For each $t \in \mathbb{R}$, we define

$$M_t := \{m \in M | (t, m) \in \mathcal{D}\},$$

where $\mathcal{D}$ is the flow domain of the flow F.

Proposition 16. *(**Naturality of Flows**). Suppose M and N are smooth manifolds, $\phi : M \to N$, a smooth map, $X \in \mathfrak{X}(M)$, and $Y \in \mathfrak{X}(N)$. Let F_X be the flow of X and F_Y be the flow of N. If $X \sim_\phi Y$, then for each $t \in \mathbb{R}$, $\phi(M_t) \subseteq N_t$ and $F_Y \circ \phi = \phi \circ F_X$ on M_t:*

$$\begin{tikzcd} M_{-t} \arrow[r, "\phi"] & N_{-t} \\ M_t \arrow[u, "F_X"] \arrow[r, "\phi"'] & N \arrow[u, "F_Y"'] \end{tikzcd}$$

See [9, Proposition 9.13] for proof.

The Naturality of flows allows a flow defined on one manifold to be pulled back or pushed forward to another.

Corollary 17. *(**Diffeomorphism Invariance of Flows**). Let $\phi : M \to N$ be a diffeomorphism. If $X \in \mathfrak{X}(M)$ is a vector field defined on M, then we can define a vector field $Y \in \mathfrak{X}(N)$ using the push forward. The push forward of X is defined as $Y = \phi_* X = T\phi \circ X \circ \phi^{-1}$. Since X is defined, we have the flow F_X defined on M which we can push forward to N to define the flow F_Y on N whose corresponding vector field is $Y = \phi_* X$. The push forward of F_X is defined as $F_Y = \phi_* F_X = \phi \circ F_X \circ \phi^{-1}$. Likewise if Y and F_Y are defined on N, then we can pull the vector field and flow back to M. The pull back of Y is defined as $X = \phi^* Y = T\phi^{-1} \circ Y \circ \phi$, and the pull back of F_Y is defined as $F_X = \phi^* F_Y = \phi^{-1} \circ F_Y \circ \phi$.*

Returning to example 1 we have the vector fields $X \in \mathfrak{X}(S^1)$ and $Y \in \mathfrak{X}(\mathbb{R})$, the maps ϕ, ϕ^{-1}, the tangent maps $T\phi$, $T\phi^{-1}$, and the flows F_X and F_Y previously defined. We defined the vector field X by pulling back the vector filed Y, $X = \phi^* Y$, and we defined the flow F_Y by pushing forward the flow F_X, $F_Y = \phi_* F_X$.

C. Summary

The main idea from this section is that not every manifold is a vector space which means that we cannot perform calculus directly on the space. Instead, these manifolds can be mapped to Euclidean space using charts via local coordinates. In these local coordinates, we can perform calculus to define tangent vectors and vector fields. These tangent vectors can be mapped to the tangent space of the manifold, and the vector fields defined in Euclidean space can be pulled back to the manifold. In addition, we can perform integration or move along flows in local coordinates, and map them back to the manifold.

IV. Lie Group Actions and Equivariant Maps

In this section we explore how Lie group actions and their corresponding Lie algebra actions are flows on manifolds that generate infinitesimal generators. We will mostly follow the work of Engø in [5], [6].

Definition 18. (**Matrix Lie group**). A matrix Lie group is a subgroup of $GL(n; \mathbb{C})$ that is a smooth manifold G with the property that the multiplication map $m : G \times G \to G$ and inversion map $i : G \to G$ defined as

$$m(g, h) = gh, \quad i(g) = g^{-1}$$

are smooth maps.

One example of a matrix Lie group is $SO(2) := \left\{R \in \mathbb{R}^{2\times 2} | R^\top R = 1, \det(R) = 1\right\}$ which is the set of all rotation isometries in $\mathbb{R}^2$.

Definition 19. (**Group Action**). Let M be a manifold and let $(G, \cdot)$ be a group. A left action of G on M is a map $\Phi : G \times M \to M$ satisfying

$$\Phi(e, m) = m, \quad \Phi(g_2, \Phi(g_1, m)) = \Phi(g_2 \cdot g_1, m)$$

$$\begin{tikzcd} M \arrow[r, "\phi"] & N \\ M \arrow[u, "\Phi_g"] \arrow[r, "\phi"'] & N \arrow[u, "\Psi_g"'] \end{tikzcd}$$

for all $m \in M$. Similarly, a right action of G on M is a map $\Phi : M \times G \to M$ satisfying

$$\Phi(m, e) = m, \quad \Phi(\Phi(m, g_1), g_2) = \Phi(m, g_1 \cdot g_2)$$

for all $m \in M$. An action Φ is class C^r if Φ is a C^r-map, and is smooth if Φ is a C^∞-map.

An important property of groups actions to consider is the transitive property. A group action is called *transitive* if, for all $m_1, m_2 \in M$ there exists an element $g \in G$ such that $m_2 = \Phi(g, m_1)$ if the action is a left action or $m_2 = \Phi(m_1, g)$ if the action is a right action. Basically what this means is that from any element $m \in M$ the group action can map it to any other element in M.

In example 1, let the group be $SO(2)$, then the left group action in matrix notation is

$$s_2 = R \begin{bmatrix} x \\ y \end{bmatrix},$$

and the right group action is

$$s_2 = \begin{bmatrix} x & y \end{bmatrix} R,$$

where $(x, y) \in S^1$. It can also be shown that the group action is transitive.

Before we continue, we need to define additional maps. If Φ is a group action of G on M, then we define the maps $\Phi_g : M \to M,\ g \mapsto \Phi(g, \cdot)$ and $\Phi_m : G \to M,\ m \mapsto \Phi(\cdot, m)$ for left actions and $\Phi_g : M \to M,\ g \mapsto \Phi(\cdot, g)$ and $\Phi_m : G \to M,\ m \mapsto \Phi(m, \cdot)$ for right actions. A group action on itself will be denoted $L_g : G \to G,\ g \mapsto \Phi(g, \cdot)$ for left actions and $R_g : G \to G,\ g \mapsto \Phi(\cdot, g)$.

Definition 20. (**Equivariant map**). Let (M, G, Φ) and (N, G, Ψ) be two homogenous spaces. The map $\phi : M \to N$ is an equivariant map (or ϕ-equivariant) if $\phi \circ \Phi_g = \psi_g \circ \phi$, i.e. the following diagram commutes

Every Lie group has local coordinates referred to as its Lie algebra which we will denote as $\mathfrak{g}$. We will restrict the rest of our discussion to restricting the groups G to matrix Lie groups; however, this theory has been generalized to any group. There are different local coordinates that we can use for matrix Lie groups, but we will stick with using the canonical coordinates of the first kind which is the matrix exponential mapping defined as $\exp : \mathfrak{g} \to G$ and $\exp^{-1} : G \to \mathfrak{g}$ where $\exp^{-1}$ is the matrix logarithm. We quickly note that this is a chart of G whose domain is not all of G. The Lie algebra of the group G is a space where addition and scalar multiplication is defined; it is a space where we can perform numeric integration.

The goal is to relate Lie group actions to flows, and in order to do this, we need to construct an equivariant map between the Lie algebra, the Lie group, and the manifold M. We already have an equivariant map between the Lie group and Lie algebra, thus we need to pull back the group action of G on itself to $\mathfrak{g}$.

We can pull back a group action of G on itself to $\mathfrak{g}$ using the exponential mapping in a way similar to the pull back of a flow. Let $B_g = \exp_* R_g$ be the pullback of the right group action, then B_g is defined as

$$B_g \coloneqq \exp^{-1} \circ R_g \circ \exp$$
$$B_g(u) = \log(\exp(u)\, g).$$

B_g could have been using the left group action. With the construction of B_g, we have the following diagram that commutes

$$
\begin{tikzcd}
\mathfrak{g} \arrow[r, "\exp"] & G \arrow[r, "\Phi_m"] & M \\
\mathfrak{g} \arrow[u, "B_g"] \arrow[r, "\exp"'] & G \arrow[u, "R_g"] \arrow[r, "\Phi_m"'] & M \arrow[u, "\Phi_g"]
\end{tikzcd}
$$

Since the Lie algebra is a local coordinate of G, by definition of a chart, the exponential map must be bijective, and in fact it is a diffeomorphism. In addition, since G is a matrix Lie group, the actions of G are smooth. Since B_g is a composition of smooth maps, it is also smooth. Depending on the map ϕ_M, it may or may not be bijective, but we only need to require it to be of class C^1. The purpose of these properties will be made clear in the following sections.

Returning to example 1, we have the homogenous space $\left(S^1, SO\left(2\right), \Phi\right)$ where S^1 is the manifold on which an ODE is evolving, $SO\left(2\right)$ is a matrix Lie group, and Φ is the group action. In this example, we will define the group action as a left group action such that

$$
\begin{gathered}
\Phi : SO\left(2\right) \times S^1 \to S^1,\; \left(R, s\right) \mapsto Rs \\
\Phi_R : S^1 \to S^1,\; s \mapsto Rs \\
\Phi_s : SO\left(2\right) \to S^1,\; R \mapsto Rs,
\end{gathered}
$$

where s is a column matrix representation of an element of S^1. We also define the group action on itself as the left group action

$$
L_R : SO\left(2\right) \to SO\left(2\right),\; RH,
$$

where $H \in SO\left(2\right)$. The Lie algebra of $SO\left(2\right)$ is $\mathfrak{so}\left(2\right) := \left\{u \in \mathbb{R}^{2\times 2} | u^\top + u = 0\right\}$, i.e., the set of 2 dimensional skew symmetric matrices. The mapping between the Lie algebra and the Lie group is the exponential map, and the pull back of the Lie group action onto the Lie algebra is

$$
\begin{gathered}
B_R = \exp^{-1} \circ L_R \circ \exp \\
B_R\left(u\right) = \log\left(R \exp\left(u\right)\right).
\end{gathered}
$$

We thus have the three equivalent composite maps

$$
\begin{gathered}
\Phi_s \circ \exp \circ B_R \\
\Phi_R \circ \Phi_s \circ \exp \\
\Phi_s \circ L_R \circ \exp
\end{gathered}
$$

V. Lie algebra action

Using the exponential map, we can define an action of the Lie algebra on the manifold. This is done by replacing the group element g by $\exp\left(\xi\right)$where $g = \exp\left(\xi\right)$ and $\xi \in \mathfrak{g}$, in the group actions. For example

$$
\begin{gathered}
\Phi_g = \Phi_{\exp(\xi)} \\
R_g = R_{\exp(\xi)} \\
B_g = B_{\exp(\xi)},
\end{gathered}
$$

which is expressed in the following diagram

$$
\begin{tikzcd}
\mathfrak{g} \arrow[r, "\exp"] & G \arrow[r, "\Phi_m"] & M \\
\mathfrak{g} \arrow[u, "B_{\exp(\xi)}"] \arrow[r, "\exp"'] & G \arrow[u, "R_{\exp(\xi)}"] \arrow[r, "\Phi_m"'] & M \arrow[u, "\Phi_{\exp(\xi)}"]
\end{tikzcd}
$$

Since the Lie algebra is an algebra, scalar multiplication is defined such that $t\xi \in \mathfrak{g}$ where $\xi \in \mathfrak{g}$ and $t \in \mathbb{R}$, and $\exp(t\xi) \in G$. Using this idea, we can construct a one parameter group action from every element in $\mathfrak{g}$.

Definition 21. (**One Parameter Group Action**). A one-parameter group action is a continuous map $\Phi : \mathbb{R} \times M \to M$ (for left actions) or $\Phi : M \times \mathbb{R} \to M$ (for right actions) that adhere to the group action properties for left and right actions defined in definition 19 where the group operator is addition.

If $\xi \in \mathfrak{g}$, $\exp : \mathfrak{g} \to G$, $s, t \in \mathbb{R}$, and $m \in M$, then the right one-parameter group action is

$$\Phi(t, m) = m \exp(t\xi)$$

with the properties

$$\begin{aligned} \Phi(0, m) &= m \\ \Phi(t, \Phi(s, m)) &= \Phi(t + s, m) . \end{aligned}$$

Other continuous maps can be constructed from the one parameter group action by evaluating it at an element $t \in \mathbb{R}$ and $m \in M$ defined as

$$\begin{aligned} \Phi_t(m) &:= \Phi(t, \cdot) = \exp(t\xi)(m) \\ \Phi_m(t) &:= \Phi(\cdot, m) = m(G) . \end{aligned}$$

From the one-parameter group action on the manifold M, we can construct the corresponding group actions on G and $\mathfrak{g}$

$$\begin{aligned} \Phi_g &= \Phi_{\exp(t\xi)} \\ R_g &= R_{\exp(t\xi)} \\ B_g &= B_{\exp(t\xi)}, \end{aligned}$$

which is expressed in the following diagram

$$\begin{tikzcd} \mathfrak{g} \arrow[r, "\exp"] & G \arrow[r, "\Phi_m"] & M \\ \mathfrak{g} \arrow[u, "B_{\exp(t\xi)}"] \arrow[r, "\exp"'] & G \arrow[u, "R_{\exp(t\xi)}"] \arrow[r, "\Phi_m"'] & M \arrow[u, "\Phi_{\exp(t\xi)}"] \end{tikzcd}$$

Note that the one parameter group action describes a flow on the manifold M, and Φ_m is the corresponding integral curve, and we constructed the flow from a Lie group action on the manifold. Not only that, but it can be shown that the corresponding vector field (or infinitesimal generator) is an element of the Lie algebra. The infinitesimal generator of the right one-parameter group action $\Phi(t, m) = m \exp(t\xi)$ is calculated using (12)

$$\begin{aligned} \left.\frac{d}{dt}\right|_{t=0} m \exp(t\xi) &= m\xi \\ &= \xi_M(m) . \end{aligned}$$

and $\xi_M = \xi \in \mathfrak{g}$. More importantly, any vector field $X \in \mathfrak{X}(M)$ that is liner in M can be represented as a Lie algebra if the G-action is transitive. In the cases where the vector field is not linear in M, we can approximate it by using the method of frozen frames discussed in [4] which we will briefly discuss in a later section.

Definition 22. (**Lie-Algebra Action**). If $\mathfrak{g}$ is an arbitrary finite-dimensional Lie algebra, any Lie algebra homomorphism $\lambda : \mathfrak{g} \to X \in \mathfrak{X}(M)$ is a Lie algebra action on M. The right Lie algebra action on M for matrix Lie groups is

$$\lambda := \left.\frac{d}{dt}\right|_{t=0} m \exp(t\xi) ,$$

and the left Lie algebra action on M for matrix Lie groups is

$$\lambda := \left.\frac{d}{dt}\right|_{t=0} \exp(t\xi)\, m.$$

Note that right Lie algebra actions on M produce left invariant vector fields on M.

Since the map Φ_m isn't necessarily a diffeomorphism, we cannot pull back a vector field defined on M to G or $\mathfrak{g}$; however, since we have the equivariant maps $\exp$ and Φ_m w.r.t. the flows

$$\begin{aligned} \Phi_g &= \Phi_{\exp(t\xi)} \\ R_g &= R_{\exp(t\xi)} \\ B_g &= B_{\exp(t\xi)} \end{aligned}$$

the infinitesimal generators of these flows are going to be related by the maps $\exp$ and Φ_m. They are calculated using (12) as

$$\begin{aligned} \xi_M &= \left.\frac{d}{dt}\right|_{t=0} \Phi_{\exp(t\xi)} \\ \xi_G &= \left.\frac{d}{dt}\right|_{t=0} R_{\exp(t\xi)} \\ \xi_{\mathfrak{g}} &= \left.\frac{d}{dt}\right|_{t=0} B_{\exp(t\xi)} \end{aligned}$$

and their relations are denoted as

$$\begin{aligned} \xi_{\mathfrak{g}} &\sim_{\exp \circ \Phi_m} \xi_M \\ \xi_{\mathfrak{g}} &\sim_{\exp} \xi_G. \end{aligned}$$

The following diagram depicts the relation between vector fields

$$\begin{tikzcd} T\mathfrak{g} \arrow[r, "T\exp"] & TG \arrow[r, "T\Phi_m"] & TM \\ \mathfrak{g} \arrow[u, "\xi_{\mathfrak{g}}"] \arrow[r, "\exp"'] & G \arrow[u, "\xi_G"] \arrow[r, "\Phi_m"'] & M \arrow[u, "\xi_M"] \end{tikzcd}$$

where $T\Phi_m$ is the tangent map of Φ_m and $T\exp$ is the tangent map of $\exp$. The question remains, what is the vector field $\xi_{\mathfrak{g}}$. Let $u \in \mathfrak{g}$, and using the definition of the flow $B_{\exp(t\xi)}$ we get

$$\begin{aligned} B_{\exp(t\xi)}(u) &= \log(\exp(u)\exp(t\xi)). \\ &= z(t) \end{aligned}$$

The infinitesimal generator is constructed using (12) and is

$$\begin{aligned} \xi_{\mathfrak{g}} &= \left.\frac{d}{dt}\right|_{t=0} B_{\exp(t\xi)}(u) \\ &= \left.\frac{d}{dt}\right|_{t=0} \log(\exp(u)\exp(t\xi)) \\ &= \left(\frac{I - \exp\left(ad_{z(t)}\right)}{ad_{z(t)}}\right)^{-1} \xi, \end{aligned}$$

where $ad_{z(t)}$ is the adjoint of $z(t)$. Recall that ξ is an element of $\mathfrak{g}$ which we can define as the vector field on M by ξ_M, thus

$$\xi_{\mathfrak{g}} = \left(\frac{I - \exp\left(-ad_{z(t)}\right)}{ad_{z(t)}}\right)^{-1} \xi_M.$$

This formula is based on the *left trivializations*, i.e. tangents at a point $g \in G$ being written as $G\xi$, $\xi \in \mathfrak{g}$ and is denoted as

$$\operatorname{dexp}_{z,l}^{-1} = \left(\frac{I - \exp\left(-ad_{z(t)}\right)}{ad_{z(t)}}\right)^{-1}$$

where the subscript l denotes that it is based on the *left trivializations* or is a left invariant vector field constructed from a left group action. See appendix (A) for more information about the derivative of the matrix exponential and it's inverse.

Now that we have a vector field defined on the Lie algebra, we can use any classical numerical integration technique to solve the differential equation in the Lie algebra and map the solution to the manifold. We will cover a few techniques in the next section, but before moving on we will apply what we presented in this section to example (1).

In example (1) we constructed the left group action. By using elements in the Lie algebra $\xi \in \mathfrak{so}(2)$ we construct the one-parameter left group action ($\mathbb{R} - action$)

$$\Phi_{\exp(t\xi)} : \mathbb{R} \times S^1 \to S^1,\ (t,s) \mapsto \exp(t\xi)\, s.$$

A unique $\mathbb{R} - action$ can be constructed from the every element of the Lie algebra. The corresponding Lie algebra action is the infinitesimal generator

$$\begin{aligned} \xi_{S^1}(s) &= \left.\frac{d}{dt}\right|_{t=0} \exp(t\xi)\, s \\ &= \xi s, \end{aligned}$$

thus $\xi_{S^1} = \xi$. The corresponding $\mathbb{R} - action$ on the Lie group is

$$L_{\exp(t\xi)} : \mathbb{R} \times SO(2) \to SO(2),\ (t,R) \mapsto \exp(t\xi)\, R$$

with the associated infinitesimal generator being

$$\begin{aligned} \xi_{SO(2)}(R) &= \left.\frac{d}{dt}\right|_{t=0} \exp(t\xi)\, R \\ &= \xi R. \end{aligned}$$

Lastly the corresponding $\mathbb{R} - action$ on the Lie group is

$$B_{\exp(t\xi)} : \mathbb{R} \times \mathfrak{g} \to \mathfrak{g},\ (t,u) \mapsto \log(\exp(t\xi)\exp(u))$$

with the associated infinitesimal generator being

$$\begin{aligned} \xi_{\mathfrak{g}}(u) &= \left.\frac{d}{dt}\right|_{t=0} \log(\exp(t\xi)\exp(u)) \\ &= \left(\frac{\exp\left(ad_{z(t)}\right) - I}{ad_{z(t)}}\right)^{-1} \xi \\ &= \operatorname{dexp}_{z,r}^{-1}(\xi), \end{aligned}$$

where $z(t) = B_{\exp(t\xi)}(u)$.

In example (1) we defined the integral curve $\gamma(t)$ at the point $s = (x_0, y_0)$ as, which can be written in matrix notation as

$$\gamma(t) = \underbrace{\begin{bmatrix} \cos(t) & -\sin(t) \\ \sin(t) & \cos(t) \end{bmatrix}}_{R} \begin{bmatrix} x_0 \\ y_0 \end{bmatrix},$$

where

$$R(t) = \begin{bmatrix} \cos(t) & -\sin(t) \\ \sin(t) & \cos(t) \end{bmatrix} \in SO(2),$$

and is a left $\mathbb{R} - action$. The corresponding infinitesimal generator is

$$\xi_{S^1}(s) = \left.\frac{d}{dt}\right|_{t=0} \begin{bmatrix} \cos(t) & -\sin(t) \\ \sin(t) & \cos(t) \end{bmatrix} s$$
$$= \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} s$$

where

$$\xi = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix},$$

is an element of the Lie algebra. Thus we have the ODE on S^1

$$\frac{d}{dt}\gamma(t) = \xi R(t) s$$

and the corresponding ODE on $SO(2)$ is

$$\frac{d}{dt}R(t) = \xi R(t)$$

such that

$$T\Phi_s \circ \xi_{SO(3)} = \xi_{S^1} \circ \Phi_s,$$

where $T\Phi_s : TSO(2) \to TS^1$, $v \mapsto vs$. The corresponding ODE on $\mathfrak{so}(2)$ is

$$\frac{d}{dt}z(t) = \mathrm{dexp}^{-1}_{z(t),r}(\xi).$$

VI. Lie-Euler Method

A. Right Action

Let R_g denote the right action of a Lie group G on M, and $\rho : G \to GL(M)$; $g \to \rho(g)$ be a representation of G, then the right action is defined as $R_g : G \times M \to M$; $(g, m) \mapsto m\rho(g)$. Using the exponential map $\exp : \mathfrak{g} \to G$, we can construct a right action R_v of the Lie algebra $\mathfrak{g}$ on M by $R_v : \mathfrak{g} \times M \to M$; $(v, m) \mapsto m\rho(\exp(v))$ where $v \in \mathfrak{g}$. We can use this right action to construct infinitesimal generators of R_v on M using the directional derivative

$$\xi_v(m) = \left.\frac{d}{dt}\right|_{t=0} m\rho(\exp(tv)) \tag{13}$$

where $\xi_v \in \mathfrak{X}^\infty(M)$ is the vector field induced by R_v. The set of infinitesimal generators of R_v is isomorphic to the Lie Algebra of M. Because of this isomorphism, some authors call it the Lie Algebra of M [9], [2]. By this convention, the map $\theta_r : \mathrm{Lie}(G) \to \mathrm{Lie}(M)$ defined by (13) is a Lie algebra homomorphism [9, Theorem 20.15]. It is also instructive to note the set of infinitesimal generators of R_v are left invariant vector fields.

Since vector fields are maps $\mathfrak{X}^\infty(M) : M \to TM$, they provide us with ODEs of the form

$$\dot{m} = \xi_v(m)$$

or using a representation of $\mathfrak{g}$

$$\dot{m} = m\rho(v)$$

where $\rho : \mathfrak{g} \to GL(M)$. One way to solve this is in terms of flow. Assume that we start at the point m_0 and move along the flow generated by $\rho(v)$, then the solution to the ODE is

$$m(t) = m_0 \exp(t\rho(v)).$$

B. Left Action

Let L_g denote the left action of a Lie group G on M, and $\rho : G \to GL(M)\,;\; g \to \rho(g)$ be a representation of G, then the left action is defined as $L_g : G \times M \to M;\; (g, m) \mapsto \rho(g)\, m$. Using the exponential map $\exp : \mathfrak{g} \to G$, we can construct a left action L_v of the Lie algebra $\mathfrak{g}$ on M by $L_v : \mathfrak{g} \times M \to M;\; (v, m) \mapsto \rho(\exp(v))\, m$ where $v \in \mathfrak{g}$. We can use this left action to construct infinitesimal generators of L_v on M using the directional derivative

$$\xi_v(m) = \left.\frac{d}{dt}\right|_{t=0} m\rho(\exp(tv)) \tag{14}$$

where $\xi_v \in \mathfrak{X}^\infty(M)$ is the vector field induced by L_v. The set of infinitesimal generators of L_v is not isomorphic to the Lie Algebra of M. This is because the map $\theta_l : \mathrm{Lie}(G) \to \mathfrak{X}^\infty(M)$ defined by (14) is an antihomomorphism (a linear map satisfying $\phi_l([v_1, v_2]) = -[\phi_l(v_1), \phi_l(v_2)]$ for all $v_1, v_2 \in \mathfrak{g}$) [9, Theorem 20.18]. It is also instructive to note the set of infinitesimal generators of L_v are right invariant vector fields.

Since vector fields are maps $\mathfrak{X}^\infty(M) : M \to TM$, they provide us with ODEs of the form

$$\dot{m} = \xi_v(m)$$

or using a representation of $\mathfrak{g}$

$$\dot{m} = \rho(v)\, m$$

where $\rho : \mathfrak{g} \to GL(M)$. As before, one solution is

$$m(t) = \exp(t\rho(v))\, m_0.$$

C. Example

In this example we will work with the unit sphere in 3-dimensions $M = S^2 := \left\{(x, y, z) \in \mathbb{R}^3 | x^2 + y^2 + z^2 = 1\right\}$, the group $G = SO(3) := \left\{R \in GL(3, \mathbb{R})\, | R^T R = 1,\ \det(R) = 1\right\}$, and the associated Lie algebra $\mathfrak{g} = \mathfrak{so}(3) := \left\{v \in \mathbb{R}^{3\times 3} | v + v^\top = 0\right\}$. The left action of the Lie group $SO(3)$ on S^2 is

$$L_R : SO(3) \times S^2 \to S^2;\; (R, s) \mapsto \rho(R)\, \rho(s)\,, \tag{15}$$

where $\rho : SO(3) \to SO(3)$ is the standard representation and $\rho : S^2 \to \mathbb{R}^3;\; s \to [s]$ is the matrix representation of an element of S^2 where $s = (x, y, z)$. For clarity I will write out (15) using the representations of each group

$$L_R : SO(3) \times S^2 \to S^2;\; (R, s) \mapsto R \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

Note that can just as easily define the right action

$$R_R : SO(3) \times S^2 \to S^2;\; (R, s) \mapsto \begin{bmatrix} x & y & z \end{bmatrix} R$$

but for the purposes of this example we will work with the left action.

The left action of the Lie algebra $\mathfrak{so}(3)$ on S^2 is then

$$L_v : \mathfrak{so}(3) \times S^2 \to S^2;\; (v, s) \mapsto \exp(v)\, \rho(s)\,.$$

This left action can be used to generate vector fields on S^2 from $\mathfrak{so}(3)$ using the directional derivative

$$\begin{aligned} \xi_v(s) &= \left.\frac{d}{dt}\right|_{t=0} \exp(tv)\, \rho(s) \\ &= v\rho(s)\,. \\ &= v \begin{bmatrix} x \\ y \\ z \end{bmatrix} \end{aligned}$$

Using the vector fields on S^2, we can construct the first order ODE

$$\dot{s} = v\rho(s)$$

whose Lie-Euler solution is

$$s(t) = \exp(tv)\, \rho(s_0)\,,$$

where s_0 is the initial value.

Algorithm 1

D. Summary

The Lie-Euler method provides an exact solution as long as the vector field doesn't change over the period of integration. If the vector field does change, then we can either integrate over smaller time period, or use another method of integration such as the Runge-Kutta-Munthe-Kaas method.

VII. Runge-Kutta-Munthe-Kaas

In the section VI, we found an exact solution to a first order differential equation provided that the vector field was constant during the time interval of integration. When the vector field on a manifold is non-constant during the interval of integration, we need another method of numerical integration. A popular method is the Runge-Kutta (RK) method that performs numerical integration on vector spaces. Unfortunately not every manifold is a vector space. Crouch and Grossman in [4] were amongst the first to modify the RK method to evolve on manifolds without the use of local coordinates.

VIII. Numerical Integration of ODE on Manifolds (Crouch)

A. Introduction

The flow of many mechanical systems preserve certain constraints: energy, rotation, confined to a manifold. Some numerical methods do not keep the constraints.

This paper is focused on numerical algorithms that naturally evolve on a constraint manifold.

Integration in local coordinates can become difficult when the trajectory of the system leave the coordinate chart where the local representation is valid.

Integration by embedding the manifold in some higher dimensional space can cause the result to leave the submanifold. You can project it back onto the submanifold though.

This paper develops a numerical integration procedure in which iterates of the integration algorithm always evolve on the submanifold

1) Vector Fields with Frozen Coefficients: Let $X_n \in \Gamma^\infty(M)$ be a vector field basis such that any vector field X can be represented as $X = \sum_{i=1}^n f_i(t,m) X_n$ where $f \in C^\infty(M)$. Then a vector in TM can be expressed as

$$\dot{m} = \sum_{i=1}^n f_i(t,m) X_i.$$

If we freeze the coefficients f_i for a period δt at time τ and configuration $p \in M$ then the differential equation becomes

$$\dot{m} = \sum_{i=1}^n a_i X_i$$

over the time period δt where $a_i = f_i(\tau, p)$. Let the solution to the system be denoted as $m(t) = F(t, m_0)$ where $F(t, m_0)$ is the flow of the system. If $m_0 \in M$ then $F(t, m_0) \in M$.

B. Classical Numerical Integration Algorithms

Used to solve differential equations written in the form

$$\dot{z} = F(z), z \in \mathbb{R}^N z(0) \in z_0.$$

1) Classical (Explicit) Runge-Kutta Algorithms: asf

2) Classical (Explicit) Multistep Algorithms:

IX. High Order Runge-Kutta methods on Manifolds (Hans) [12]

A. RKMK methods of general order

Appendix A
The Derivative of the Exponential Mapping

The derivative of the exponential mapping is commonly used in Lie theory such as in proving the Baker-Campbell-Hausdorff formula [7], and in many numerical integration techniques for solving ODE [8]. We present the general form of the derivative of the exponential mapping, and it's computationally efficient form for $SO(3)$ and $SE(3)$.

A. General Form

Theorem 23. *Let $X(t)$ be a smooth matrix-valued function, then*

$$\frac{d}{dt}\exp(X(t)) = \exp(X(t))\frac{I-\exp\left(-ad_{X(t)}\right)}{ad_{X(t)}}\left(\frac{dX}{dt}\right). \tag{16}$$

This theorem is proved in [7]. Note that the term

$$\text{dexp}_{-X(t)} = \frac{I-\exp\left(-ad_{X(t)}\right)}{ad_{X(t)}}$$

is a left invariant vector field. This vector field can be mapped to a right invariant vector field using the Adjoint

$$Ad_{X(t)}\left(\text{dexp}\left(X(t)\right)\right)\exp(X(t)) = \exp(X(t))\,\text{dexp}_{X(t)},$$

which gives us the other form for the derivative of the exponential mapping

$$\frac{d}{dt}\exp(X(t)) = \text{dexp}_{-X(t)}\left(\frac{dX}{dt}\right)\exp(X(t)), \tag{17}$$

where

$$\text{dexp}_{X(t)} = \frac{\exp(X(t))-I}{ad_{X(t)}}.$$

The terms $\text{dexp}_{X(t)}$ and $\text{dexp}_{-X(t)}$ are derived from the geometric series

$$\text{dexp}_{-X(t)} = \sum_{k=0}^{\infty}(-1)^k\frac{ad_{X(t)}^k}{(k+1)!}$$

$$\text{dexp}_{X(t)} = \sum_{k=0}^{\infty}\frac{ad_{X(t)}^k}{(k+1)!}$$

and have the inverses

$$\begin{aligned}\text{dexp}_{-X(t)}^{-1} &= \frac{I-\exp\left(-ad_{X(t)}\right)}{ad_{X(t)}}\\ &= \sum_{k=0}^{\infty}(-1)^k\frac{B_k}{k!}ad_{X(t)}^k\end{aligned}$$

$$\begin{aligned}\text{dexp}_{X(t)}^{-1} &= \frac{ad_{X(t)}}{\exp(X(t))-I}\\ &= \sum_{k=0}^{\infty}\frac{B_k}{k!}ad_{X(t)}^k,\end{aligned}$$

where B_k are Bernoulli numbers.

The geometric series form of $\text{dexp}_{X(t)}$ and $\text{dexp}_{-X(t)}$ and their inverses are more useful than the analytical forms since the term $ad_{X(t)}$ is not always invertible. This is the case when $X(t)$ is an element of $\mathfrak{so}(3)$ or $\mathfrak{se}(3)$ which are two of the Lie algebras most used in mechanics. Fortunately the derivative of the matrix exponential have closed forms in $SO(3)$ and $SE(3)$.

B. $SO(3)$

Let $x \in \mathfrak{so}(3)$ with norm $\theta = \|x\|$ then

$$\begin{aligned}\text{dexp}_{-x} &= I + \frac{\cos(\theta)-1}{\theta^2}x + \frac{\theta-\sin(\theta)}{\theta^3}x^2\\ \text{dexp}_{x} &= I + \frac{1-\cos(\theta)}{\theta^2}x + \frac{\theta-\sin(\theta)}{\theta^3}x^2\end{aligned}$$

It is instructive to see the derivation for at least one of the above terms. In the derivation we will use the property of a skew symmetric matrix that

$$\begin{aligned} ad_x &= x \\ ad_x^2 &= x^2 \\ ad_x^{2n+1} &= (-1)^n \left\| x \right\|^{2n} x \\ ad_x^{2n+2} &= (-1)^n \left\| x \right\|^{2n} x^2, \end{aligned}$$

for $n \geq 0$. The derivation proceeds as follows

$$\begin{aligned} \text{dexp}_{-x} &= \sum_{k=0}^{\infty} (-1)^k \frac{ad_{X(t)}^k}{(k+1)!} \\ &= \sum_{k=0}^{\infty} (-1)^k \frac{x^k}{(k+1)!} \\ &= I - \frac{1}{2!}x + \frac{1}{3!}x^2 + \frac{\left\| x \right\|^2}{4!}x - \frac{\left\| x \right\|^2}{5!}x^2 \cdots \\ &= I + \left[\left(1 - \frac{\left\| x \right\|^2}{2!} + \frac{\left\| x \right\|^4}{4!} \cdots \right) - 1 \right] \frac{x}{\left\| x \right\|^2} + \left[1 - \left(\left\| x \right\| - \frac{\left\| x \right\|^3}{3!} + \frac{\left\| x \right\|^5}{5!} \cdots \right) / \left\| x \right\| \right] \frac{x^2}{\left\| x \right\|^2} \\ &= I + \frac{\cos(\theta) - 1}{\theta^2} x + \frac{\theta - \sin(\theta)}{\theta^3} x^2. \end{aligned}$$

Their corresponding inverses are

$$\begin{aligned} \text{dexp}_{-x}^{-1} &= I + \frac{1}{2}x - \frac{\left\| x \right\| \cot\left(\frac{\left\| x \right\|}{2}\right) - 2}{2\left\| x \right\|^2} x^2 \\ \text{dexp}_{x}^{-1} &= I - \frac{1}{2}x - \frac{\left\| x \right\| \cot\left(\frac{\left\| x \right\|}{2}\right) - 2}{2\left\| x \right\|^2} x^2 \end{aligned}$$

C. $SE(3)$

An element $x \in \mathfrak{se}(3)$ has the form

$$x = \begin{bmatrix} w & v \\ 0_{3\times 1} & 0 \end{bmatrix}$$

where $w \in \mathfrak{so}(3)$ and $v \in \mathbb{R}^3$. The adjoint of x and powers are

$$\begin{aligned} ad_x &= \begin{bmatrix} w & v_\times \\ 0_{3\times 3} & w \end{bmatrix} \\ ad_x^2 &= \begin{bmatrix} w^2 & (wv_\times + v_\times w) \\ & w^2 \end{bmatrix} \\ ad_x^3 &= -\theta^2 ad_x - 2\left(\left(w^\vee\right)^\top v \right) \begin{bmatrix} 0 & w \\ 0 & 0 \end{bmatrix} \\ ad_x^{2n+1} &= (-1)^n \left(\theta^{2n} ad_x + \begin{bmatrix} 0 & \alpha(n) w \\ 0 & 0 \end{bmatrix} \right) \\ ad_x^{2n+2} &= (-1)^n \left(\theta^{2n} ad_x^2 + \begin{bmatrix} 0 & \alpha(n) w^2 \\ 0 & 0 \end{bmatrix} \right), \end{aligned}$$

for $n \geq 0$ and where

$$\alpha(n) = (2n)\left(\left(w^\vee\right)^\top v \right) \theta^{2(n-1)}.$$

The geometric series of dexp_{-x} is

$$\begin{aligned}
\operatorname{dexp}_{-x} &= \sum_{k=0}^{\infty} (-1)^k \, \frac{ad_{X(t)}^k}{(k+1)!} \\
&= 1 + \sum_{n=0}^{\infty} \frac{(-1)^{2n+1}}{(2n+2)!} ad_x^{2n+1} + \sum_{n=0}^{\infty} \frac{(-1)^{2n+2}}{(2n+3)!} ad_x^{2n+2} \\
&= 1 + \sum_{n=0}^{\infty} \frac{(-1)^{2n+1}}{(2n+2)!} (-1)^n \left(\theta^{2n} ad_x + \begin{bmatrix} 0 & \alpha(n)\, w \\ 0 & 0 \end{bmatrix} \right) + \sum_{n=0}^{\infty} \frac{(-1)^{2n+2}}{(2n+3)!} (-1)^n \left(\theta^{2n} ad_x^2 + \begin{bmatrix} 0 & \alpha(n)\, w^2 \\ 0 & 0 \end{bmatrix} \right) \\
&= 1 + \sum_{n=0}^{\infty} \frac{(-1)^{2n+1}}{(2n+2)!} (-1)^n \left(\theta^{2n} ad_x + \begin{bmatrix} 0 & \alpha(n)\, w \\ 0 & 0 \end{bmatrix} \right) + \sum_{n=0}^{\infty} \frac{(-1)^{2n+2}}{(2n+3)!} (-1)^n \left(\theta^{2n} ad_x^2 + \right) \\
&= 1 + \sum_{n=0}^{\infty} \frac{(-1)^{n+1}}{(2n+2)!} \theta^{2n} ad_x + \sum_{n=0}^{\infty} \frac{(-1)^{n+2}}{(2n+3)!} \theta^{2n} ad_x^2 + \sum_{n=0}^{\infty} \frac{(-1)^{n+1}}{(2n+2)!} \begin{bmatrix} 0 & \alpha(n)\, w \\ 0 & 0 \end{bmatrix} + \sum_{n=0}^{\infty} \frac{(-1)^{n+2}}{(2n+3)!} \begin{bmatrix} 0 & \alpha(n)\, w^2 \\ 0 & 0 \end{bmatrix}.
\end{aligned}$$

The geometric series

$$\begin{aligned}
\sum_{n=0}^{\infty} \frac{(-1)^{n+1}}{(2n+2)!} \theta^{2n} ad_x &= \frac{\cos(\theta) - 1}{\theta^2} ad_x \\
\sum_{n=0}^{\infty} \frac{(-1)^{n+2}}{(2n+3)!} \theta^{2n} ad_x^2 &= \frac{\theta - \sin(\theta)}{\theta^3} ad_x^2
\end{aligned}$$

were solved in subsection A-B. This leaves us to solve the last two series. We being with

$$\sum_{n=0}^{\infty} \frac{(-1)^{n+1}}{(2n+2)!} \begin{bmatrix} 0 & \alpha(n)\, w \\ 0 & 0 \end{bmatrix}$$

which can be written as

$$\left(w^{\vee}\right)^{\top} v \left(\sum_{n=0}^{\infty} \frac{(-1)^{n+1} (2n)}{(2n+2)!} \theta^{2(n-1)} \right) \begin{bmatrix} 0 & w \\ 0 & 0 \end{bmatrix},$$

thus were left with solving for

$$\begin{aligned}
\sum_{n=0}^{\infty} \frac{(-1)^{n+1} (2n)}{(2n+2)!} \theta^{2(n-1)} &= 0 + \frac{2}{4!} - \frac{4}{6!}\theta^2 + \frac{6}{8!}\theta^4 - \frac{8}{10!}\theta^6 + \frac{10}{12!}\theta^8 \cdots \\
&= \left(\frac{1}{3!} - \frac{1}{5!}\theta^2 + \frac{1}{7!}\theta^4 - \frac{1}{9!}\theta^6 \cdots \right) + \left(\frac{-2}{4!} + \frac{2}{6!}\theta^2 - \frac{2}{8!}\theta^4 + \frac{2}{10!}\theta^6 \cdots \right) \\
&= \frac{1}{\theta^2} - \frac{1}{\theta^3} \left(\theta - \frac{1}{3!}\theta^3 + \frac{1}{5!}\theta^5 - \frac{1}{7!}\theta^7 + \frac{1}{9!}\theta^9 \cdots \right) + -\frac{1}{\theta^2} + \frac{2}{\theta^4} - \frac{2}{\theta^4} \left(1 - \frac{1}{2}\theta^2 + \frac{1}{4!}\theta^4 - \frac{1}{6!}\theta^6 + \frac{1}{8!}\theta^8 \right. \\
&= -\frac{1}{\theta^3} \sin(\theta) + 2 \left(\frac{1 - \cos(\theta)}{\theta^4} \right),
\end{aligned}$$

where we used the identity

$$\frac{n-2}{n!} = \frac{1}{(n-1)!} - \frac{2}{n!}.$$

Note that you can calculate the geometric series with the first few terms (about 2-3) and have very little error (about 10^{-7}). We calculate the second series in a similar manner. The series

$$\sum_{n=0}^{\infty} \frac{(-1)^{n+2}}{(2n+3)!} \begin{bmatrix} 0 & \alpha(n)\, w^2 \\ 0 & 0 \end{bmatrix}$$

can be written as

$$\left(w^{\vee}\right)^{\top} v \left(\sum_{n=0}^{\infty} \frac{(-1)^{n+2} (2n)}{(2n+3)!} \theta^{2(n-1)} \right) \begin{bmatrix} 0 & w^2 \\ 0 & 0 \end{bmatrix},$$

thus we are left with solving for

$$
\begin{aligned}
\sum_{n=0}^{\infty} \frac{(-1)^{n+2}(2n)}{(2n+3)!}\theta^{2(n-1)} &= 0 - \frac{2}{5!} + \frac{4}{7!}\theta^2 - \frac{6}{9!}\theta^4 + \frac{8}{11!}\theta^6 \cdots \\
&= \left(\frac{3}{5!} - \frac{3}{7!}\theta^2 + \frac{3}{9!}\theta^4 - \frac{3}{11!}\theta^6 \cdots\right) + \left(-\frac{1}{4!} + \frac{1}{6!}\theta^2 - \frac{1}{8!}\theta^4 + \frac{1}{10!}\theta^6 \cdots\right) \\
&= -\frac{3}{\theta^4} + \frac{3}{3!\theta^2} + \frac{3}{\theta^5}\left(\theta - \frac{1}{3!}\theta^3 + \frac{1}{5!}\theta^5 - \frac{1}{7!}\theta^7 + \frac{1}{9!}\theta^9 - \frac{1}{11!}\theta^{11} \cdots\right) + \frac{1}{\theta^4} - \frac{1}{2!\theta^2} - \frac{1}{\theta^4}\left(1 - \frac{1}{2!}\theta^2 + \right. \\
&= -\frac{2}{\theta^4} + \frac{3}{\theta^5}\sin(\theta) - \frac{1}{\theta^4}\cos(\theta),
\end{aligned}
$$

where we used the identity

$$
\frac{n-3}{n!} = \frac{1}{(n-1)!} - \frac{3}{n!}.
$$

For clarity, let

$$
\begin{aligned}
a_\theta &= \frac{\cos(\theta) - 1}{\theta^2} \\
b_\theta &= \frac{\theta - \sin(\theta)}{\theta^3} \\
c_\theta &= -\frac{1}{\theta^3}\sin(\theta) + 2\left(\frac{1 - \cos(\theta)}{\theta^4}\right) \\
d_\theta &= -\frac{2}{\theta^4} + \frac{3}{\theta^5}\sin(\theta) - \frac{1}{\theta^4}\cos(\theta) \\
q_{-x}(w) &= \left(\left(w^\vee\right)^\top v\right)\left(c_\theta w + d_\theta w^2\right).
\end{aligned}
$$

Putting all of the pieces together we get

$$
\mathrm{dexp}_{-x} = \begin{bmatrix} \mathrm{dexp}_{-w} & \left(a_\theta v_\times + b_\theta\left(w v_\times + v_\times w\right) + q_{-x}(w)\right) \\ 0 & \mathrm{dexp}_{-w} \end{bmatrix}.
$$

The closed form solution for dexp_x is derived in a similar manner, which we will not do here; however, its closed form is

$$
\mathrm{dexp}_{x} = \begin{bmatrix} \mathrm{dexp}_{w} & \left(-a_\theta + b_\theta\left(w v_\times + v_\times w\right) + q_{x}(w)\right) \\ & \mathrm{dexp}_{w} \end{bmatrix}
$$

where

$$
q_x(w) = \left(\left(w^\vee\right)^\top v\right)\left(-c_\theta w + d_\theta w^2\right).
$$

The inverses of $\mathrm{dexp}_{X(t)}$ and $\mathrm{dexp}_{-X(t)}$ are easily computed since a square block matrix M with the from

$$
M = \begin{bmatrix} A & B \\ 0 & A \end{bmatrix}
$$

has an inverse

$$
M^{-1} = \begin{bmatrix} A^{-1} & -A^{-1}BA^{-1} \\ 0 & A^{-1} \end{bmatrix}.
$$

Using this identity we compute

$$
\begin{aligned}
\mathrm{dexp}_{-x}^{-1} &= \begin{bmatrix} \mathrm{dexp}_{-w}^{-1} & \mathrm{dexp}_{-w}^{-1} B_{-x} \mathrm{dexp}_{-w}^{-1} \\ 0 & \mathrm{dexp}_{-w}^{-1} \end{bmatrix} \\
\mathrm{dexp}_{x}^{-1} &= \begin{bmatrix} \mathrm{dexp}_{w}^{-1} & \mathrm{dexp}_{w}^{-1} B_{x} \mathrm{dexp}_{w}^{-1} \\ 0 & \mathrm{dexp}_{w}^{-1} \end{bmatrix}
\end{aligned}
$$

where

$$
\begin{aligned}
B_{-x} &= \left(a_\theta v_\times + b_\theta\left(w v_\times + v_\times w\right) + q_{-x}(w)\right) \\
B_{x} &= \left(-a_\theta v_\times + b_\theta\left(w v_\times + v_\times w\right) + q_{x}(w)\right).
\end{aligned}
$$

Appendix B
Classical Runge-Kutta Methods

Consider the 1st order ordinary differential equation with initial conditions

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

The Taylor series expansion of $y(t)$ is

$$\begin{aligned} y(t+h) &= y(t) + h\dot{y}(t) + \frac{h^2}{2}\ddot{y}(t) + O\left(h^3\right) \\ &= y(t) + hf(t, y) + \frac{h^2}{2}\frac{d}{dt}f(t, y) + O\left(h^3\right) \end{aligned}$$

note that the term

$$\begin{aligned} \frac{d}{dt}f(t, y) &= \frac{\partial f}{\partial t}\frac{dt}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} \\ &= \frac{\partial f}{\partial t}1 + \frac{\partial f}{\partial y}f(t, y) \\ &= f_t(t, y) + f_y(t, y) f(t, y) \end{aligned}$$

can be substituted into the original equation and then simplified to get

$$\begin{aligned} y(t+h) &= y(t) + hf(t, y) + \frac{h^2}{2}\left(f_t(t, y) + f_y(t, y) f(t, y)\right) + O\left(h^3\right) \\ &= y(t) + \frac{h}{2}f(t, y) + \frac{h}{2}\left(f(t, y) + hf_t(t, y) + hf_y(t, y) f(t, y)\right) + O\left(h^3\right). \end{aligned} \tag{18}$$

Before we can simplify this expression more, we need to look at the bivariate Taylor series expansion.

The Bivariate Taylor series expansion is defined as

$$x(t+h, z+k) = \sum_{n=0}^{\infty}\frac{1}{n!}\left(h\frac{\partial}{\partial t} + k\frac{\partial}{\partial z}\right)^n f(t, y)$$

\theta^{2}, i.e.

$$x(t+h, z+k) = x(t, z) + hx_t(t, z) + kf_z(t, z) + O\left(h^2, k^2\right).$$

Using the Bivariate Taylor series expansion we can approximate

$$f(t+h, y+k) \approx f(t, y) + hf_t(t, y) + kf_y(t, y).$$

If we let $k = hf(t, y)$ then

$$f(t+h, y+hf(t, y)) \approx f(t, y) + hf_t(t, y) + hf(t, y) f_y(t, y)$$

We can then substitute this into (18) to get

$$y(t+h) = \frac{h}{2}f(t, y) + \frac{h}{2}f(t+h, y+hf(t, y)) + O\left(h^3\right)$$

which gives us the 2-stage Runge-Kutta Method

$$\begin{aligned} k_1 &= hf(t, y) \\ k_2 &= hf(t+h, y+hk_1) \\ y(t+h) &= \frac{1}{2}k_1 + \frac{1}{2}k_2 \end{aligned}$$

Higher stages of the classical Runge-Kutta method are derived the same way by expanding out the Taylor series even more. An s-stage classical Runge-Kutta method is expressed as

$$\begin{aligned} k_i &= f\left(t + c_i h, y_n + h\sum_{j<i}^{i} a_{ij}k_j\right) \\ y(t+h) &= y(t) + \sum_{i=1}^{s} b_i k_i \end{aligned}$$

where the coefficients c_i, b_i, c_{ij} are read off a Butcher Tableau.

REFERENCES

[1] Ralph Abraham, Jerrold Mardsen, and Tudor Ratiu. *Manifolds, Tensor Analysis, and Applications*. Springer-Verlag, New York, first edition, 1998.
[2] Francesco Bullo and Andrew D. Lewis. *Geometric Control of Mechanical Systems*. Springer Science+Business Media, LCC, New York, 2005.
[3] Elena Celledoni, Håkon Marthinsen, and Brynjulf Owren. An introduction to Lie group integrators - basics, new developments and applications. *Journal of Computational Physics*, 257(PB):1040–1061, 2014.
[4] P. E. Crouch and R. Grossman. Numerical integration of ordinary differential equations on manifolds. *Journal of Nonlinear Science*, 3(1):1–33, 1993.
[5] Kenth Engø. On the construction of geometric integrators in the RKMK class. *BIT Numerical Mathematics*, 40(1):41–61, 2000.
[6] Kenth Engø. Partitioned Runge-Kutta Methods in Lie-Group Setting. *BIT*, 43:21–39, 2003.
[7] Brian C. Hall. *Lie Groups, Lie Algebras, and Representations: An Elementary Introduction*. Springer-Verlag New York, Inc, 2003.
[8] Arieh Iserles, Hans Z. Munthe-Kaas, Syvert P. Nørsett, and Antonella Zanna. Lie-group methods. *Acta Numerica*, 9(March 2015):215–365, 2000.
[9] John M. Lee. *An Introduction to Smooth Manifolds*. Springer Science+Business Media, LCC, 2013.
[10] Hans Munthe-Kaas. Lie-Butcher theory for Runge-Kutta methods. *BIT Numerical Mathematics*, 35(4):572–587, 1995.
[11] Hans Munthe-Kaas. Runge-Kutta methods on Lie groups. *BIT Numerical Mathematics*, 38(1):92–111, 1998.
[12] Hans Munthe-Kaas. High order Runge-Kutta methods on manifolds. *Applied Numerical Mathematics*, 29(1):115–127, 1999.