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 [11], [12], [13], 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 [9] 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], [10] 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 several 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 IV.

Example 2. A simplified version of the second order ODE of a multirotor using body-centric representation is

$$\dot{q} = qv^{\wedge}$$
$$\mathcal{I}\dot{v} = \mathrm{ad}_v^{*}\mathcal{I}v + F\left(v, q\right)$$

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

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

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. Method Overview

We are interested in solving first order ordinary differential equations of the form

$$\dot{\gamma}(t) = X(t, \gamma(t)) \tag{1}$$

where M is a smooth manifold, $I \subseteq \mathbb{R}$, $\gamma(t) : I \to M$ is an integral curve of the vector field $X \in \mathfrak{X}(M)$ at the point $M \ni m = \gamma(0)$. In many cases, analytic solutions to the ODE do not exist, or at least we cannot solve them, and we must rely on numerical approximations to solve the ODE.

One of the most popular methods for numerical integration is the Runge-Kutta method (see [2], [7] for detailed information on this method, and appendix **??** for a simple overview). The Runge-Kutta (RK) method is limited to vector spaces, i.e. it isn't defined on manifolds that do not support addition and scalar multiplication. Fortunately, the explicit classical RK method was extended to manifolds on which the vector fields can be locally defined by an infinitesimal generator corresponding to a Lie group action. This method is called the Runge-Kutta-Munthe-Kaas (RK-MK) and was developed in the series of papers [11], [12], [13], and later optimized by Munthe-Kass and Owren in [14].

Depending on the manifold M, RK-MK can depend on the idea of vector fields with frozen coefficients developed by Crouch and Grossman in [4] which is well explained in [3]. To understand the basic idea consider the integral curve $v(t) : I \to V$ that evolves on the vector space V according to the simple linear first order ODE with a constant vector field $X : V \to TV$, $v \mapsto Av$ at the point $v(0) = v_0$ (TV is the tangent bundle of V) as depicted below

$$\begin{aligned}\dot{v}(t) &= X(v(t)) \\ &= Av(t).\end{aligned}$$

It is assumed that $Av(t)$ is defined for some representation of $v(t)$, and we have chosen to omit explicitly showing representations for clearity. The well known solution to the above ODE is

$$v(t) = \exp(At)\, v_0.$$

Now consider the same scenario, but let $X : \mathbb{R} \times V \to V$, $(t, v(t)) \mapsto A(t)\, v(t)$ be a vector field dependent on t, but still linear in $v(t)$, then the well known solution is

$$v(t) = \exp\left(\int_0^t A(t)\, dt\right) v_0.$$

Lastly consider the same scenario, but let X be a vector field dependent on t and not linearly dependent in $v(t)$, then the solution is not necessarily well known and it is not

$$v(t) = \exp\left(\int_0^t A(t, v(t))\, dt\right) v_0.$$

For the purposes of using RK-MK, we need to construct a vector field linear in t and $v(t)$ from a nonlinear vector field so that the solution can be approximated as

$$v(t) = \exp(At)\, v_0,$$

this is done using vector fields with frozen coefficients. The main idea is to take a vector field $X(t, v(t))$ and separate it into two parts

$$X(t, v(t)) = A(t, v(t))\, v(t)$$

and freeze the arguments of A at $t = t_0$ and $v(t) = v(t_0) = v_0$ to form the approximated vector field

$$X_{v_0}(t, v(t)) = A(t_0, v_0)\, v(t),$$

so that the approximated solution is

$$v(t) \approx \exp(A(t_0, v_0)\, t)\, v_0,$$

which is a Lie-Euler Method using vector fields with frozen coefficients. This method is only a first order approximation.

To improve this method, let $A(v(t))$ be the vector fields such that $X(v(t)) = A(v(t))\, v(t)$, and $\gamma(t)$ be the integral curve of the vector field $A(v(t))$ at the point $\gamma(t_0) = 0$, then the Taylor series of $\gamma(t)$ is

$$\gamma(t_0 + h) = \gamma(t_0) + h\left(A(v(t_0)) + \frac{d}{dt}A(v(t_0))\frac{h}{2!} + \frac{d^2}{dt^2}A(v(t_0))\frac{h^2}{3!} + \frac{d^3}{dt^3}A(v(t_0))\frac{h^3}{4!}\right) + \mathcal{O}\left(h^5\right).$$

Recall that $X(t, v(t)) = A(t, v(t))\, v(t)$. An RK method of order 4 generates a vector field

$$\bar{A}(h, v(t_0)) \approx A(v(t_0)) + \frac{d}{dt}A(v(t_0))\frac{h}{2!} + \frac{d^2}{dt^2}A(v(t_0))\frac{h^2}{3!} + \frac{d^3}{dt^3}A(v(t_0))\frac{h^3}{4!},$$

using only $A(v(t))$ to approximate the Taylor series up to order 4 so that the Taylor series of $\gamma(t)$ becomes

$$\gamma(t_0 + h) = \gamma(t_0) + h\bar{A}(h, v(t_0)) + \mathcal{O}\left(h^5\right).$$

The constructed vector field $\bar{A}(v(t))$ is then used to approximate the integral

$$\int_{t_0}^{t_0+h} A(v(t)) = h\bar{A}(h, v(t_0)) + \mathcal{O}\left(h^5\right),$$

and solve the ODE

$$v(t) \approx \exp\left(\bar{A}(h, v(t))\, h\right) v_0,$$

which is a lot more accurate than the 1^{st} order Lie-Euler Method. The purpose of this note is to lightly explain how to construct the vector field $\bar{A}$ using the RK-MK method.

Remark 3. Crouch and Crossman in [4] give a more precise and abstract definition and description for vector fields with frozen coefficients. We have simplified the definition so that it's application for the purposes of this note will be easily understood.

IV. FUNDAMENTALS

In this section we briefly review some of the fundamental theory needed to understand the RK-MK method.

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

$$\gamma(t) = \left(\cos(t)\, x_0 - \sin(t)\, y_0, \sin(t)\, x_0 + \cos(t)\, y_0\right), \tag{2}$$

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.

$$x(t) = \cos(t)\, x_0 - \sin(t)\, y_0 \tag{3a}$$

$$y(t) = \sin(t)\, x_0 + \cos(t)\, y_0. \tag{3b}$$

where $x(t)$ and $y(t)$ 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

$$\frac{d}{dt} f(t) := \lim_{h \to 0} \frac{f(t+h) - f(t)}{h}. \tag{4}$$

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

$$\frac{d}{dt} \gamma(t) := \lim_{h \to 0} \frac{\gamma(t+h) - \gamma(t)}{h}.$$

This expression for the derivative of $\gamma(t)$ is subtly misleading. The terms $\gamma(t+h)$ and $\gamma(t)$ 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 4. (**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 5. (**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

$$\phi : S^1 \to \mathbb{R},\ (x,y) \mapsto \arctan(y/x) \tag{5}$$

whose domain is $S^1 \backslash \{(0,-1)\}$. 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 $\gamma(t)$ and the chart ϕ to form the composite map

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

where $x(t)$ and $y(t)$ are defined in (3). 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.

$$
\begin{aligned}
\frac{d}{dt}\phi \circ \gamma(t) &= \lim_{h \to 0} \frac{\phi \circ \gamma(t+h) - \phi \circ \gamma(t)}{h} && \text{(6a)}\\
&= \frac{d}{dt}\arctan\left(y(t)/x(t)\right) && \text{(6b)}\\
&= \frac{\partial \arctan\left(y(t)/x(t)\right)}{\partial x}\frac{dx}{dt} + \frac{\partial \arctan\left(y(t)/x(t)\right)}{\partial y}\frac{dy}{dt} && \text{(6c)}\\
&= \frac{y(t)}{x^2(t)+y^2(t)}\frac{dx}{dt} - \frac{x(t)}{x^2(t)+y^2(t)}\frac{dy}{dt} && \text{(6d)}\\
&= y(t)\frac{dx}{dt} - x(t)\frac{dy}{dt} && \text{(6e)}\\
&= y^2(t) + x^2(t) && \text{(6f)}\\
&= 1. && \text{(6g)}
\end{aligned}
$$

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 6. (**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 7. 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], [10].

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 8. (**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 9. (**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 (6).

$$T\phi : TS^1 \to T\mathbb{R},\ ((x, y), (\dot{x}, \dot{y})) \to y\dot{x} - x\dot{y}, \tag{7}$$

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 (3), 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

$$x\dot{x} + y\dot{y} = 0. \tag{8}$$

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

$$x = \frac{1}{\sqrt{\tan^2(\theta) + 1}} \tag{9a}$$

$$y = \frac{\tan^2(\theta)}{\sqrt{1 + \tan^2(\theta)}}, \tag{9b}$$

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 (8) 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

$$\left(\theta, \dot{\theta}\right) \mapsto \begin{bmatrix} \frac{\tan^2(\theta)}{\sqrt{1+\tan^2(\theta)}} \\ -\frac{1}{\sqrt{\tan^2(\theta)+1}} \end{bmatrix} \dot{\theta}. \tag{10}$$

We can now use the inverse tangent map to map the tangent vector $\dot{\theta} = 1$ at θ found in (6) 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 (3).

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.

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

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

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

B. Vector Fields and Flows

Definition 10. (**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

$$X(x, y) = \begin{bmatrix} y \\ -x \end{bmatrix}. \tag{11}$$

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

$$Y(\theta) = 1, \tag{12}$$

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 11. (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 $Y \in \mathfrak{X}\left(\mathbb{R}\right)$ are ϕ-related is rather trivial since we have already defined all of the vector fields, the map, and the tangent map. Using (5), (7), (11), and (12) we show that $X \sim_\phi Y$,

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

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 $\gamma\left(t\right)$ defined in (2). This curve is called integral curve of $X \in \mathfrak{X}\left(S^1\right)$ at (x_o, y_o).

Definition 12. (**Integral curve**). An integral curve of $X \in \mathfrak{X}\left(M\right)$ at $m \in M$ is a curve $\gamma\left(t\right)$ such that $\dot{\gamma}\left(t\right) = X\left(\gamma\left(t\right)\right)$ where t is an element of the interval $I \subseteq \mathbb{R}$ that includes the 0 element and $\gamma\left(0\right) = m$.

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 $X \in \mathfrak{X}\left(M\right)$ is a smooth vector field, then there is a unique integral curve $\gamma\left(t\right)$ 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 $M = \mathbb{R}^2 \backslash \left\{0\right\}$ with the vector field $\frac{\partial}{\partial x}$. The corresponding integral curve is $\gamma\left(t\right) = (x + t, y)$ 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 13. (**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 [10, 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 14. (**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

$$v_m = \left. \frac{d}{dt} \right|_{t=0} F(t, m). \tag{13}$$

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 (2), 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 13

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

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 15. *(**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 [10, 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 16. *(**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.

V. 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 17. (**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 $\mathbb{R}^2$.

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

Definition 18. (**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

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

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

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

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 $m_2 = \Phi\left(g, m_1\right)$ if the action is a left action or $m_2 = \Phi\left(m_1, g\right)$ 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 $SO\left(2\right)$, 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 $\Phi_g : M \to M,\ g \mapsto \Phi\left(g, \cdot\right)$ and $\Phi_m : G \to M,\ m \mapsto \Phi\left(\cdot, m\right)$ for left actions and $\Phi_g : M \to M,\ g \mapsto \Phi\left(\cdot, g\right)$ and $\Phi_m : G \to M,\ m \mapsto \Phi\left(m, \cdot\right)$ for right actions. A group action on itself will be denoted $L_g : G \to G,\ g \mapsto \Phi\left(g, \cdot\right)$ for left actions and $R_g : G \to G,\ g \mapsto \Phi\left(\cdot, g\right)$.

Definition 19. (**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

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

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

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

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 $\left(S^1, SO(2), \Phi\right)$ where S^1 is the manifold on which an ODE is evolving, $SO(2)$ 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

$$\Phi : SO(2) \times S^1 \to S^1,\ (R, s) \mapsto Rs$$
$$\Phi_R : S^1 \to S^1,\ s \mapsto Rs$$
$$\Phi_s : SO(2) \to S^1,\ R \mapsto Rs,$$

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

$$L_R : SO(2) \to SO(2),\ RH,$$

where $H \in SO(2)$. The Lie algebra of $SO(2)$ is $\mathfrak{so}(2) := \left\{u \in \mathbb{R}^{2\times 2} | u^\top + u = 0\right\}$, 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

$$B_R = \exp^{-1} \circ L_R \circ \exp$$
$$B_R(u) = \log(R \exp(u)).$$

We thus have the three equivalent composite maps

$$\Phi_s \circ \exp \circ B_R$$
$$\Phi_R \circ \Phi_s \circ \exp$$
$$\Phi_s \circ L_R \circ \exp$$

VI. 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 $\exp(\xi)$where $g = \exp(\xi)$ and $\xi \in \mathfrak{g}$, in the group actions. For example

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

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 20. (**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 18 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

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

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 (13)

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

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 21. (**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 (13) 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 (13) and is

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

where $ad_{z(t)}$ is the adjoint of $z\left(t\right)$. 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

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

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 (**??**) 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 $\xi \in \mathfrak{so}\left(2\right)$ we construct the one-parameter left group action ($\mathbb{R}-action$)

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

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

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

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

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

with the associated infinitesimal generator being

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

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

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

with the associated infinitesimal generator being

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

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

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

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

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

VII. Runge-Kutta-Munthe-Kaas

We now are now ready to present the RK-MK algorithm of any order. Let $\gamma(t) : I \to M$ be an integral curve of the vector field $X \in \mathfrak{X}(M)$ at $m_0 = \gamma(t)$, and assume that the vector field X can be represented as

$$X = \xi_M(t, m)\, m,$$

where ξ_M is the infinitesimal generator of ξ such that $\xi_M(t, m) \in \mathfrak{g}$. The s RK-MK algorithm from $\gamma(t_0)$ to $\gamma(t_0 + h)$ is

Algorithme 1 RK-MK

```
for i = 1, 2, ..., s
    u_i = h Σ_{j=i}^{s} a_ij k̃_j
    k_i = ξ_M (hc_i, λ(u_i, γ(t)))
    k̃_i = dexpinv(u_i, k_i)
end
v = h Σ_{j=1}^{s} b_j k̃_j
γ(t + h) = λ(v, γ(t))
```

The algorithm can be initially confusing, so to help build intuition we present the 4 stage RK-MK method which is also a 4^{th} order method in Algorithm 2.

While the RK-MK method has many advantages, one of the draw backs in the increase in the number of flops associated with the exponential mapping and dexpinv. Fortunately, Munth-Kaas and Owren optimized the RK-MK 4

Algorithme 2 RK-MK 4

$u_1 = 0$
$k_1 = \xi_M\left(0, \lambda\left(u_1, \gamma\left(t\right)\right)\right)$
$\tilde{k}_1 = k_1$
$u_2 = h\frac{1}{2}\tilde{k}_1$
$k_2 = \xi_M\left(\frac{1}{2}h, \lambda\left(u_2, \gamma\left(t\right)\right)\right)$
$\tilde{k}_3 = \text{dexpinv}\left(u_2, k_2\right)$
$u_3 = h\frac{1}{2}\tilde{k}_2$
$k_3 = \xi_M\left(\frac{1}{2}h, \lambda\left(u_3, \gamma\left(t\right)\right)\right)$
$\tilde{k}_3 = \text{dexpinv}\left(u_3, k_3\right)$
$u_4 = h\tilde{k}_3$
$k_4 = \xi_M\left(h, \lambda\left(u_4, \gamma\left(t\right)\right)\right)$
$\tilde{k}_4 = \text{dexpinv}\left(u_4, k_4\right)$
$v = h\left(\frac{1}{6}\tilde{k}_1 + \frac{1}{3}\tilde{k}_2 + \frac{1}{3}\tilde{k}_3 + \frac{1}{6}\tilde{k}_4\right)$
$\gamma\left(t+h\right) = \lambda\left(v, \gamma\left(t\right)\right)$

Algorithme 3 RK-MK 4 Optimized Right

$k_1 = h\xi_M\left(0, \lambda\left(0, \gamma\left(t\right)\right)\right)$
$k_2 = h\xi_M\left(\frac{1}{2}h, \lambda\left(\frac{1}{2}k_1, \gamma\left(t\right)\right)\right)$
$k_3 = h\xi_M\left(\frac{1}{2}h, \lambda\left(\frac{1}{2}k_2 - \frac{1}{8}\left[k_1, k_2\right], \gamma\left(t\right)\right)\right)$
$k_4 = h\xi_M\left(h, \lambda\left(k_3, \gamma\left(t\right)\right)\right)$
$v = \left(\frac{1}{6}k_1 + \frac{1}{3}k_2 + \frac{1}{3}k_3 + \frac{1}{6}k_4 - \frac{1}{12}\left[k_1, k_4\right]\right)$
$\gamma\left(t+h\right) = \lambda\left(v, \gamma\left(t\right)\right)$

method in [14]. We will not go over the details of the derivation, but only present it algorithms 3 for right-trivialized and 4 for left trivialized. Note that the dexpinv has been replaced by just two commutators!

In the case of example 1, the RK-MK 4 method reduces to the Lie-Euler method since the vector field is constant, and it's solution is

$$\gamma\left(t\right) = \exp\left(\omega t\right)\gamma\left(0\right).$$

To really demonstrate the power of RK-MK we will introduce another example in the following subsection.

Remark 22. It should be noted that there are some vector fields ξ_M such that $\left[\xi_M\left(m_1\right), \xi_M\left(m_2\right)\right] = 0$, $\forall m_1, m_2 \in M$. This implies that $\text{dexpinv}\left(\xi_M\left(m_1\right), \xi_M\left(m_2\right)\right) = \xi\left(m_2\right)$ which allows the RK-MK method to be simplified even more.

A. *Example*

Example 23. Let $y \in SO\left(3\right)$ and $X \in \mathfrak{X}^r\left(SO\left(3\right)\right)$ defined as

$$X\left(y\right) = \left(3\sin\left(y - y^\top\right)y - y^\top 3\sin\left(y - y^\top\right)\right)y,$$

where the $\sin$ function acts on each component of y such that $\sin\left(y\right) \in \mathbb{R}^{3\times 3}$, then the corresponding ODE is

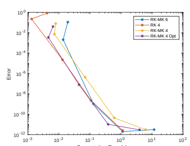
$$\dot{y} = X\left(y\right), \tag{14}$$

Algorithme 4 RK-MK 4 Optimized Left

$k_1 = h\xi_M\left(0, \lambda\left(0, \gamma\left(t\right)\right)\right)$
$k_2 = h\xi_M\left(\frac{1}{2}h, \lambda\left(\frac{1}{2}k_1, \gamma\left(t\right)\right)\right)$
$k_3 = h\xi_M\left(\frac{1}{2}h, \lambda\left(\frac{1}{2}k_2 + \frac{1}{8}\left[k_1, k_2\right], \gamma\left(t\right)\right)\right)$
$k_4 = h\xi_M\left(h, \lambda\left(k_3, \gamma\left(t\right)\right)\right)$
$v = \left(\frac{1}{6}k_1 + \frac{1}{3}k_2 + \frac{1}{3}k_3 + \frac{1}{6}k_4 + \frac{1}{12}\left[k_1, k_4\right]\right)$
$\gamma\left(t+h\right) = \lambda\left(v, \gamma\left(t\right)\right)$

Algorithme 5 Implementation Example of RK-MK 4

1: $u_1 = 0$
2: $k_1 = X_y\left(y_0\right)$
3: $\tilde{k}_1 = k_1$
4: $u_2 = h\frac{1}{2}\tilde{k}_1$
5: $k_2 = X_y\left(\exp\left(u_2\right)y_0\right)$
6: $\tilde{k}_3 = \text{dexpinv}\left(u_2, k_2\right)$
7: $u_3 = h\frac{1}{2}\tilde{k}_2$
8: $k_3 = X_y\left(\exp\left(u_3\right)y_0\right)$
9: $\tilde{k}_3 = \text{dexpinv}\left(u_3, k_3\right)$
10: $u_4 = h\tilde{k}_3$
11: $k_4 = X_y\left(\exp\left(u_4\right)y_0\right)$
12: $\tilde{k}_4 = \text{dexpinv}\left(u_4, k_4\right)$
13: $v = h\left(\frac{1}{6}\tilde{k}_1 + \frac{1}{3}\tilde{k}_2 + \frac{1}{3}\tilde{k}_3 + \frac{1}{6}\tilde{k}_4\right)$
14: $y\left(t_0 + h\right) = \exp\left(v\right)y_0$



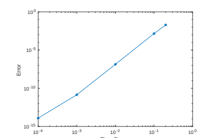


Figure 1. The image on the left shows the error as a function of computation time, and the image on the right shows the error in the determinant as a function of time step size.

which can be written as

$$\dot{y} = X_y\left(y\right)y,$$

where

$$X_y\left(y\right) = 3\sin\left(y - y^\top\right)y - y^\top 3\sin\left(y - y^\top\right).$$

The solution to the ODE is

$$y\left(t_0 + h\right) = \exp\left(\int_{t_0}^{t_0+h} X_y\left(y\right)dt\right)y_0,$$

where $y_0 = y\left(t_0\right)$. We note that the integral

$$\int_{t_0}^{t} X_y\left(y\right)dt$$

might not have an analytical solution; however, since $X_y\left(y\right) \in \mathfrak{so}\left(3\right)$, we can use RK-MK 4 to approximate the integral up to the 4th order. Since $X_y\left(y\right)$ is right trivialized, we need to use dexpinv corresponding to right-trivialization. Algorithm 5 shows the implementation for this example based on algorithm 2.

We finish this section with a results from a simulation. We numerically integrated the ODE described by (14) using RK4, RK-MK6, RK-MK4, RK-MK4 OPT over a time period of 10 seconds with interval steps of $h = \left\{5^{-1}, 10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}\right\}$. The result obtained from RK-MK6 using a time interval of $h = 10^{-4}$ was used as truth data. We calculated the error in the solution using the function

$$e = \left\|\log\left(R_T^\top R_E\right)\right\|,$$

where R_T is our true result and R_E is the estimated result obtained by one of the four algorithms. We also measured the total time required for each algorithm to perform the numerical integration. Lastly we measured in error in the determinant of the result obtained from the RK4 method. These results are shown in fig (1).

The results are surprising. Note that the computation time required for a certain numerical accuracy for the RK-MK4 OPT method is very similar to that of RK4; however, the RK4 method does leave the manifold slightly.

VIII. Partitioned RK-MK

The RK-MK method that we have discussed so far works on flows defined in Homogenous spaces that are a first order ordinary differential equation. The RK-MK method can be extended to flows that are second ordinary differential equations if these flows can be described as a group action on itself where the group is a tangent bundle of a Lie group. In other words we are working with the homogenous space (TG, TG, Φ) where G is a lie group, and TG is the tangent bundle of the Lie group G.

Since the tangent bundle of a Lie group is diffeomorphic to the direct product of the Lie group with its Lie algebra, i.e. $TG \approxeq G \times \mathfrak{g}$, we can work with the homogenous space $(G \times \mathfrak{g}, G \times \mathfrak{g}, \Phi)$. This allows us to work with a simpler structure. The group action Φ, in this case is defined as

$$\Phi : ((G \times \mathfrak{g}) \times (G \times \mathfrak{g})) \to (G \times \mathfrak{g})\,;\, (g_1, v_1) \times (g_2, v_2) \mapsto (g_1 g_2, v_1 + v_2)\,.$$

However, this doesn't properly represent the structure of the tangent bundle. To understand why, consider the two flows $F_1 = g_1 \exp(tv_1)$ and $F_2 = g_2 \exp(tv_2)$. We can construct a new flow that is the product of the two flows

$$F_3 = F_1 F_2 = g_1 \exp(tv_1)\, g_2 \exp(tv_2)\,,$$

taking the directional derivative yields

$$\begin{aligned} \left.\frac{d}{dt}\right|_{t=0} g_1 \exp(tv_1)\, g_2 \exp(tv_2) &= g_1 v_1 g_2 + g_1 g_2 v_2 \\ &= g_1 g_2 \left(\mathrm{Ad}_{g_2^{-1}} v_1 + v_2\right). \end{aligned}$$

This shows that the Lie algebra of the product of two flows is $\mathrm{Ad}_{g_2^{-1}} v_1 + v_2$, and not simply $v_1 + v_2$. Because of this reason, we will work with the semi-direct product $(G \ltimes_{\mathrm{Ad}} \mathfrak{g})$. For more information on the semi-direct product see [6] or appendix **??**.

A. Lie Algebra Action

Given the homogenous space $(G \ltimes_{\mathrm{Ad}} \mathfrak{g}, G \ltimes_{\mathrm{Ad}} \mathfrak{g}, \Phi)$, we know that we can represent flows on $G \ltimes_{\mathrm{Ad}} \mathfrak{g}$ as a group action of the group on itself. We do this by utilizing the Lie algebra of the Lie group. Let $(\mathfrak{g} \ltimes_{\mathrm{ad}} \mathfrak{h})$ be the Lie algebra of the semi-direct product of $G \ltimes_{\mathrm{Ad}} \mathfrak{g}$, then

$$\Phi_{\exp(t(\mathfrak{g} \ltimes_{\mathrm{ad}} \mathfrak{h}))} (G \ltimes_{\mathrm{Ad}} \mathfrak{g}) \tag{15}$$

represents a flow on $G \ltimes_{\mathrm{Ad}} \mathfrak{g}$. The automorphism $\Phi_{\exp(t(\mathfrak{g} \ltimes_{\mathrm{ad}} \mathfrak{h}))}$ can represent either a left or right group action. By taking the directional derivative of 15 with respect to time, we can derive the infinitesimal generator, which is the Lie Algebra Action;

$$\left.\frac{d}{dt}\right|_{t=0} \Phi_{\exp(t(\mathfrak{g} \ltimes_{\mathrm{ad}} \mathfrak{h}))} (G \ltimes_{\mathrm{Ad}} \mathfrak{g})\,.$$

What this is depends on the type of action and group. Let $(g, w) \in G \ltimes_{\mathrm{Ad}} \mathfrak{g}$ and $(v, u) \in \mathfrak{g} \ltimes_{\mathrm{ad}} \mathfrak{h}$. The infinitesimal generator corresponding to the left action is

$$\begin{aligned} &\left.\frac{d}{dt}\right|_{t=0} \left(g \exp(tv)\,, \mathrm{Ad}_{\exp(tv)^{-1}} w + \mathrm{dexp}_{(tv),l}(tu)\right) \\ &= (gv, \mathrm{ad}_{-v} w + u)\,, \end{aligned}$$

where $\frac{d}{dt} g = gv$ and $\frac{d}{dt} w = \mathrm{ad}_{-v} w + u$. The infinitesimal generator corresponding to the right action is

$$\begin{aligned} &\left.\frac{d}{dt}\right|_{t=0} \left(\exp(tv)\, g, w + \mathrm{Ad}_{g^{-1}} \left(\mathrm{dexp}_{(tv),r}(tu)\right)\right) \\ &= \left(vg, w + \mathrm{Ad}_{g^{-1}} u\right), \end{aligned}$$

where $\frac{d}{dt} g = vg$ and $\frac{d}{dt} w = \mathrm{Ad}_{g^{-1}} u + w$.

We can think of an element $\xi_{G \ltimes_{\mathrm{Ad}} \mathfrak{g}} \in \mathfrak{g} \ltimes_{\mathrm{ad}} \mathfrak{h}$ as a vector field of the tangent bundle that maps TG to TTG and is thus a function of TG and time. So the objective is to integrate the vector field $\xi_{G \ltimes_{\mathrm{Ad}} \mathfrak{g}}$ using the RK-MK method.

Example 24. Consider a small unmanned air vehicle (sUAV) with the following equations of motion

$$\begin{aligned}\dot{P}^i_{b/i} &= R^i_b v^b_{b/i} \\ m\dot{v}^b_{b/i} &= -\omega^b_{b/i} \times m v^b_{b/i} + F_1\left(R, P, \omega, v, t\right) \\ \dot{R}^i_b &= R^i_b \omega^b_{b/i} \\ J\dot{\omega}^b_{b/i} &= -\omega^b_{b/i} \times J\omega^b_{b/i} + F_2\left(R, P, \omega, v, t\right),\end{aligned}$$

where $P^i_{b/i}$ is the position of the sUAV w.r.t. the inertial frame expressed in the inertial frame, R^i_b is the rotation from the body frame to the inertial frame, $v^b_{b/i}$ is the translational velocity of the sUAV w.r.t. in the inertial frame and expressed in the body frame, $\omega^b_{b/i}$ is the angular velocity w.r.t. the inertial frame and expressed in the body frame, m is the mass of the sUAV, J is the moment of inertial about the center of mass, and F_1 and F_2 are forces on the system.

Since a sUAV moves on $SE\left(3\right)$, we can express the equations of motion as a Lie algebra action on the semi-direct product group $SE\left(3\right) \ltimes_{\text{Ad}} \mathfrak{se}\left(3\right)$. To show this, let $\left(x, z\right) \in SE\left(3\right) \ltimes_{\text{Ad}} \mathfrak{se}\left(3\right)$ where

$$x = \begin{bmatrix} R^i_b & P^i_{b/i} \\ 0 & 1 \end{bmatrix},$$

and

$$z = \begin{bmatrix} J\omega^b_{b/i} & m v^b_{b/i} \\ 0 & 0 \end{bmatrix},$$

and let $\left(a, b\right) \in \left(\mathfrak{se}\left(3\right) \ltimes_{\text{ad}} \mathfrak{se}\left(3\right)\right)$ where

$$a = \begin{bmatrix} \omega^b_{b/i} & v^b_{b/i} \\ 0 & 0 \end{bmatrix},$$

and

$$b = \begin{bmatrix} F_2 & F_1 \\ 0 & 0 \end{bmatrix}.$$

The infinitesimal generator corresponding to the left action is

$$\begin{aligned}&\left.\frac{d}{dt}\right|_{t=0} \left(x, z\right) \exp\left(t\left(a, b\right)\right) \\ &= \left(xa, \text{ad}_{-a} z + b\right).\end{aligned}$$

Thus we have this vector field $\xi_{SE(3)\ltimes_{\text{Ad}}\mathfrak{se}(3)} = \left(a, b\right)$ that maps elements from $TSE\left(3\right)$ to $TTSE\left(3\right)$.

B. Partitioned Runge-Kutta-Munthe-Kass Method

Let H denote a semi-direct product of a Lie group such that $H = G \ltimes_{\Phi} \mathfrak{g}$, and let ξ be the vector field on H that is the infinitesimal generator of the left multiplication action of H on itself. The vector field can be written as $\xi = \left(\zeta\left(H, t\right), \eta\left(H, t\right)\right)$ where $\zeta, \eta \in \mathfrak{g}$. Let $h \in H$ with the flow $h \exp\left(\int_{t_0}^t \xi\left(t, h\right)\right)$, our objective is to use the RK-MK method to numerically calculate the integral $\int_{t_0}^t \xi\left(t, h\right)$.

Appendix

In this appendix we quickly review the direct product of manifolds and spaces and the semi-direct product of homogenous spaces. We review this in preparation to discuss the partitioned RK-MK method that is discussed in section (VIII).

A. Direct Product of Spaces

Let M and N be two spaces (sets with additional structure), then the direct product of the spaces constructs a new space $M \times N := \{(m, n) \in M \times N | \, m \in M, \, n \in N\}$ that preserves the structure of M and N. For example let $\Delta : M \times M \to M$ be an operation defined on M and $\nabla : N \times N \to N$ be an operation defined on N, then we can define a new operation

$$* : (M \times N) \times (M \times N) \to M \times N, \, ((m_1, n_1), (m_2, n_2)) \mapsto (m_1 \Delta m_1, n_1 \nabla n_2).$$

For another example, let G be a group with operator $\Delta : G \times G \to G$ and H be another group with operator $\nabla : H \times H \to H$, then we can form the direct product of the spaces $G \times H$ and define a new operation

$$* : (G \times H) \times (G \times H) \to G \times H, \, ((g_1, h_1), (g_2, h_2)) \mapsto (g_1 \Delta g_2, h_1 \nabla h_2).$$

As a final example,let G be a group with operator $\Delta : G \times G \to G$ and M be a vector space, then we can form the direct product of the spaces $G \times M$ and define a new operation

$$* : (G \times M) \times (G \times M) \to G \times M, \, ((g_1, m_1), (g_2, m_2)) \mapsto (g_1 \Delta g_2, m_1 + m_2).$$

B. Semi-Direct Product of Groups

Given a Homogenous space (H, G, Φ) where H is a group equipped with the operator $\nabla : H \times H \to H$, G is a group with the operator $\Delta : G \times G \to G$ and $\Phi : G \to \text{Aut}(H)$ is a map from G to the automorphism of H such that $\Phi(g)(h) \in H$ where $g \in G$ and $h \in H$. For shorthand notation we denote $\Phi(g)$ as Φ_g. Note that Φ_g can be either a left, right or conjugate action on H, but we do not specify what type of action it is to keep it generic. We also note that Φ is a homomorphism which implies

$$\Phi_{g_2} \circ \Phi_{g_1}(h) = \Phi_{g_2 \Delta g_1}(h).$$

Using the automorphism, we can construct a new group which is the semi-direct product $G \ltimes_\Phi H := \{(g, h) : \, g \in G, \, h \in H\}$ with the operator $*$. The operator $*$ can be defined as either a left or right operator (sometimes referred as left and right group multiplication). In the case that $*$ is a right operator we define it as

$$* : (G \times H) \times (G \times H) \to G \times H, \, ((g_1, h_1), (g_2, h_2)) \mapsto (g_1 \Delta g_2, \Phi_{g_2}(h_1) \nabla h_2),$$

and in the case $*$ is a left operator we define it as

$$* : (G \times H) \times (G \times H) \to G \times H, \, ((g_1, h_1), (g_2, h_2)) \mapsto (g_1 \Delta g_2, h_1 \nabla \Phi_{g_1}(h_2)).$$

Since $G \ltimes_\Phi H$ is a group, it must also have an identity element and an inverse element. The identity element is simply (e_G, e_H) where e_G is the identity element in G and e_H is the identity element in H. To find the inverse, we use the identity

$$(g_1, h_1) * (g_1, h_1)^{-1} = (e_G, e_H),$$

where $*$ can be either the right or left operator. To fine the inverse element, let $(g_2, h_2) = (g_1, h_1)^{-1}$, then for the right group operation we have

$$(g_1 \Delta g_2, \Phi_{g_2}(h_1) \nabla h_2) = (e_G, e_H)$$

which holds when $g_2 = g_1^{-1}$ and $h_2 = \Phi_{g_2}(h_1)^{-1}$, thus

$$(g_1, h_1)^{-1} = \left(g_1^{-1}, \Phi_{g_1^{-1}}(h_1)^{-1}\right).$$

Similarly for the left group operation we get

$$(g_1, h_1)^{-1} = \left(g_1^{-1}, \Phi_{g_1^{-1}}(h_1)^{-1}\right).$$

Remark 25. Depending on the author, how the semi-direct product is defined can have subtle differences. We have tried to present it abstractly and simply as possible.

Example 26. (The Special Euclidean Group) Let $\mathbb{R}^n$ be an n-dimensional vector space and $SO(n)$ be the n-dimensional special orthogonal group which contains the set of all rotational isometries in n-dimensional space. We can construct the homogenous space $(\mathbb{R}^n, SO(n), \Phi)$ where Φ is a left group action

$$\Phi : SO(n) \times \mathbb{R}^n \to \mathbb{R}^n, \, (R, v) \mapsto Rv,$$

and

$$\Phi : SO(n) \to \mathrm{Aut}(\mathbb{R}^n).$$

A natural semi-direct product is $SO(n) \ltimes_\Phi \mathbb{R}^n$ where new right group operator $*$ is defined as

$$* : (SO(n) \times \mathbb{R}^n) \times (SO(n) \times \mathbb{R}^n) \to SO(n) \times \mathbb{R}^n,\ ((R_1, p_1), (R_2, p_2)) \mapsto (R_1 R_2, p_1 + \Phi(R_1, p_2)).$$

To make it more clear what we are doing, we can represent the new group $SO(n) \ltimes_\Phi \mathbb{R}^n$ in matrix notation as

$$E = \begin{bmatrix} R & p \\ 0_{3\times 3} & 1 \end{bmatrix},$$

where $R \in SO(n)$ and $v \in \mathbb{R}^n$. The right group multiplication in this representation is simply matrix multiplication

$$\begin{aligned} \begin{bmatrix} R_1 & p_1 \\ 0_{3\times 3} & 1 \end{bmatrix} \begin{bmatrix} R_2 & p_2 \\ 0_{3\times 3} & 1 \end{bmatrix} &= \begin{bmatrix} R_1 R_2 & p_1 + R_1 p_2 \\ 0 & 1 \end{bmatrix} \\ &= (R_1 R_2, p_1 + R_1 p_2). \end{aligned}$$

The identity element is

$$\begin{bmatrix} I & 0 \\ 0 & 1 \end{bmatrix},$$

and the inverse is

$$\begin{bmatrix} R^\top & -R^\top p \\ 0 & 1 \end{bmatrix}.$$

C. Semi-Direct Product of Linear Lie Groups Extended to Their Lie Algebras

Let $G \ltimes_\Phi H$ be a linear Lie group formed from the semi-direct product of the two Lie groups G and H, and be equipped with the operator $*$. Also, let $\mathfrak{g} \ltimes_\phi \mathfrak{h}$ be the semi-direct product of their Lie algebras. The objective of this subsection is to define the Lie bracket on $\mathfrak{g} \ltimes_\phi \mathfrak{h}$ using the adjoint action of the Lie Group $G \ltimes_\Phi H$ on $\mathfrak{g} \ltimes_\phi \mathfrak{h}$. To do this, we begin with a definition of the exponential and logarithm maps and use them with the conjugation action of $G \ltimes_\Phi H$ on itself to derive the Lie bracket on $\mathfrak{g} \ltimes_\phi \mathfrak{h}$.

In order to properly defined the exponential map, we will make use of the Ado theorem.

Theorem 27. ***(Ado).*** *Let $\mathfrak{gl}(n)$ denote the Lie algebra of $GL(n)$. Every finite-dimensional real Lie algebra is isomorphic to a subalgebra of $\mathfrak{gl}(n;\mathbb{R})$. Every finite-dimensional complex Lie algebra is isomorphic to a complex subalgebra of $\mathfrak{gl}(n;\mathbb{C})$.*

This theorem allows us to represent complex Lie algebras as matrices that are a subalgebra of $\mathfrak{gl}(n;\mathbb{C})$. In other words, let $\pi : \mathfrak{g} \ltimes_\phi \mathfrak{h} \to \mathfrak{gl}(n;\mathbb{C})$, then we can define matrix multiplication

$$\cdot : \pi(\mathfrak{g} \ltimes_\phi \mathfrak{h}) \times \pi(\mathfrak{g} \ltimes_\phi \mathfrak{h}) \to \mathfrak{gl}(2n;\mathbb{C}),$$

where the image of $\cdot$ is not necessarily $\pi(\mathfrak{g} \ltimes_\phi \mathfrak{h})$.

Theorem 28. *Every Linear Lie group is isomorphic to a matrix Lie group.*

This theorem allows us to map any Linear Lie group to a matrix Lie group. Essentially, since the Lie algebras and the Lie groups can be represented as matrices, we can easily define their exponential and logarithm maps using the matrix exponential and matrix logarithm as

$$\begin{aligned} \exp &: \mathfrak{g} \ltimes_\phi \mathfrak{h} \to G \ltimes_\Phi H,\ (v, u) \mapsto \sum_{k=0}^{\infty} \frac{1}{k!} (v, u)^k \\ \log &: G \ltimes_\Phi H \to \mathfrak{g} \ltimes_\phi \mathfrak{h},\ (g, h) \mapsto \sum_{k=1}^{\infty} (-1)^{k+1} \frac{((g, h) - (e_G, e_H))^k}{k}, \end{aligned}$$

where the mappings to their matrix representations are inferred, and will be inferred from here on out.

The conjugation action of a group on itself is defined as

$$\Psi : (G \ltimes_\Phi H) \times (G \ltimes_\Phi H) \to G \ltimes_\Phi H,\ ((g_1, h_1), (g_2, h_2)) \to (g_1, h_1) * (g_2, h_2) * (g_1, h_1)^{-1},$$

and the corresponding automorphism is

$$\Psi_{(g_1,h_1)} : (G \ltimes_\Phi H) \to (G \ltimes_\Phi H) .$$

The adjoint action of a Lie group on it's Lie algebra is define as

$$\left.\frac{d}{dt}\right|_{t=0} \Psi_{(g_1,h_1)} \left(\exp\left(t\left(v,u\right)\right)\right),$$

and is denoted

$$\mathrm{Ad}_{(g_1,h_1)} : \mathfrak{g} \ltimes_\phi \mathfrak{h} \to \mathfrak{g} \ltimes_\phi \mathfrak{h}.$$

The adjoint action of a Lie algebra on itself is defined as

$$\left.\frac{d}{dt}\right|_{t=0} \mathrm{Ad}_{\exp t(v_1,u_1)} \left(v_2, u_2\right),$$

and is denoted

$$\mathrm{ad}_{(v_1,u_1)} : \mathfrak{g} \ltimes_\phi \mathfrak{h} \to \mathfrak{g} \ltimes_\phi \mathfrak{h},$$

and is equivalent to the Lie bracket

$$\left[\left(v_1, u_1\right), \left(v_2, u_2\right)\right].$$

To help solidify what we discussed, we continue with example 26.

Example 29. (The Special Euclidean Group and Lie Algebra) The Lie algebra of the Lie group $SO\left(n\right) \ltimes_\Phi \mathbb{R}^n$ is $\mathfrak{so}\left(n\right) \ltimes_\phi \mathbb{R}^n$. The conjugate action of $SO\left(n\right) \ltimes_\Phi \mathbb{R}^n$ with $*$ being a left operator is

$$\begin{aligned}\Psi_{(R_1,v_1)}\left(R_2, v_2\right) &= \left(R_1, p_1\right) * \left(R_2, p_2\right) * \left(R_1, p_1\right)^{-1} \\ &= \left(R_1 R_2 R_1^{-1}, p_1 + R_1 p_2 - R_1 R_2 R_1^{-1} p_1\right).\end{aligned}$$

The adjoint action of the Lie group on it's Lie algebra is derived as

$$\begin{aligned}&\left.\frac{d}{dt}\right|_{t=0} \Psi_{(R_1,p_1)}\left(\exp\left(t\left(\omega,\rho\right)\right)\right) \\ &= \left.\frac{d}{dt}\right|_{t=0} \left(R_1, p_1\right) * \exp\left(t\left(\omega,\rho\right)\right) * \left(R_1, p_1\right)^{-1} \\ &= \left.\frac{d}{dt}\right|_{t=0} \left(R_1, p_1\right) * \sum_{k=0}^{\infty} \frac{t^k}{k!} \left(\omega,\rho\right)^k * \left(R_1, p_1\right)^{-1} \\ &= \left.\frac{d}{dt}\right|_{t=0} \left(R_1, p_1\right) * \left(I + t\left(\omega,\rho\right) + \mathcal{O}\left(t^2\right)\right) * \left(R_1, p_1\right)^{-1} \\ &= \left(R_1, p_1\right) \bullet \left(\omega,\rho\right) \bullet \left(R_1, p_1\right)^{-1} \\ &= \left(R_1 \omega R_1^{-1}, -R_1 \omega R_1^{-1} p_1 + R_1 \rho\right) \\ &= \mathrm{Ad}_{(R_1,p_1)}\left(\omega,\rho\right)\end{aligned}$$

where the operator $\bullet$ is dependent on the Lie group and Lie algebra, but in this case is matrix multiplication. In matrix notation we get

$$\begin{aligned}\left(R_1, p_1\right) \bullet \left(\omega,\rho\right) \bullet \left(R_1, p_1\right)^{-1} &\cong \begin{bmatrix} R_1 & p_1 \\ 0_{3\times3} & 1 \end{bmatrix} \begin{bmatrix} \omega & \rho \\ 0 & 0 \end{bmatrix} \begin{bmatrix} R_1^{-1} & -R_1^{-1} p_1 \\ 0_{3\times3} & 1 \end{bmatrix} \\ &\cong \begin{bmatrix} R_1 \omega R_1^{-1} & -R_1 \omega R_1^{-1} p_1 + R_1 \rho \\ 0 & 0 \end{bmatrix}.\end{aligned}$$

The adjoint action of the Lie algebra on itself is derived as

$$\begin{aligned}
&\left.\frac{d}{dt}\right|_{t=0} \mathrm{Ad}_{\exp t(\omega_1,\rho_1)}(\omega_2,\rho_2) \\
&= \left.\frac{d}{dt}\right|_{t=0} \exp\left(t\left(\omega_1,\rho_1\right)\right) \bullet \left(\omega_2,\rho_2\right) \bullet \exp\left(t\left(\omega_1,\rho_1\right)\right)^{-1} \\
&= \left.\frac{d}{dt}\right|_{t=0} \left(I + t\left(\omega_1,\rho_1\right) + \mathcal{O}\left(t^2\right)\right) \bullet \left(\omega_2,\rho_2\right) \bullet \left(I - t\left(\omega_1,\rho_1\right) + \mathcal{O}\left(t^2\right)\right) \\
&= \left(\omega_1,\rho_1\right)\left(\omega_2,\rho_2\right) - \left(\omega_2,\rho_2\right)\left(\omega_1,\rho_1\right) \\
&= \mathrm{ad}_{(\omega_1,\rho_1)}\left(\omega_2,\rho_2\right) \\
&= \left[\left(\omega_1,\rho_1\right),\left(\omega_2,\rho_2\right)\right].
\end{aligned}$$

In matrix notation we get

$$\begin{aligned}
\mathrm{ad}_{(\omega_1,\rho_1)}\left(\omega_2,\rho_2\right) &= \begin{bmatrix}\omega_1 & \rho_1 \\ 0 & 0\end{bmatrix}\begin{bmatrix}\omega_2 & \rho_2 \\ 0 & 0\end{bmatrix} - \begin{bmatrix}\omega_2 & \rho_2 \\ 0 & 0\end{bmatrix}\begin{bmatrix}\omega_1 & \rho_1 \\ 0 & 0\end{bmatrix} \\
&= \begin{bmatrix}\omega_1\omega_2 - \omega_1\omega_2 & \omega_1\rho_2 - \omega_2\rho_2 \\ 0 & 0\end{bmatrix}.
\end{aligned}$$

The adjoint is a group operation of the lie group $\mathfrak{so}(n) \ltimes_\phi \mathbb{R}^n$ and can be represented as

$$\mathrm{ad}_{(\omega_1,\rho_1)}\left(\omega_2,\rho_2\right) = \left(\omega_1\omega_2 - \omega_2\omega_1, \phi\left(\left(\omega_1,\rho_1\right),\left(\omega_1,\rho_2\right)\right)\right),$$

where

$$\phi : \left(\mathfrak{so}(n) \ltimes_\phi \mathbb{R}^n\right) \times \left(\mathfrak{so}(n) \ltimes_\phi \mathbb{R}^n\right) \to \mathbb{R}^n,\ \left(\left(\omega_1,\rho_1\right),\left(\omega_1,\rho_2\right)\right) \mapsto \left(\omega_1\rho_2 - \omega_2\rho_2\right).$$

D. Semi-Direct Product of Linear Lie Groups with Their Lie Algebra

In this subsection we will assume we are working with matrix Lie groups and their corresponding Lie algebras since every linear Lie group is isomorphic to a matrix Lie group whose Lie Algebra is a subset of $\mathfrak{gl}(n)$. We can form a new group by taking the semi-direct product of a lie group G with it's Lie algebra $\mathfrak{g}$ where the action of G on $\mathfrak{g}$ is the adjoint. We denote this new group as $G \ltimes_{\mathrm{Ad}} \mathfrak{g}$.

The reason behind the construction of this new group is to create a group that is isomorphic to the tangent bundle of G. To understand how this is done, create the map

$$\Phi : G \to \mathrm{Aut}(G),$$

which can be a left or right group action of G on G. We can then extend it to the Lie algebra

$$\Phi \circ \exp : \mathfrak{g} \to \mathrm{Aut}(G),$$

and construct the infinitesimal generator

$$\begin{aligned}
&\left.\frac{d}{dt}\right|_{t=0} \Phi \circ \exp(t\xi)(g) \\
&= \xi_G(g),
\end{aligned}$$

where $\xi \in \mathfrak{g}$. We now have a vector field ξ_G on G. In the case the action is a right group action, it would look like

$$\dot{g} = g\xi \in T_g G.$$

By multiplying the above tangent vector on the left by g^{-1} we get

$$\begin{aligned}
g^{-1}\dot{g} &= g^{-1}g\xi \\
&= \xi \in \mathfrak{g}.
\end{aligned}$$

By now you should see that $T_g G \cong g \times \mathfrak{g}$, and thus $TG \cong G \times \mathfrak{g}$. However, this is not a semi-direct product yet. The derivation of the semi-direct product comes from considering two integral curves γ_1 and γ_2 over the vector fields ξ_1 and ξ_2 in G. These curves can be written in the form

$$\begin{aligned}
\gamma_1 &= g_1 \exp\left(t\xi_1\right) \\
\gamma_2 &= g_2 \exp\left(t\xi_2\right).
\end{aligned}$$

If we multiply the curves together we get

$$\begin{aligned}\gamma_1\gamma_2 &= g_1 \exp(t\xi_1)\, g_2 \exp(t\xi_2) \\ &= g_1 g_2 \exp\left(\mathrm{Ad}_{g_2^{-1}} t\xi_1\right) \exp(t\xi_2),\end{aligned}$$

and their corresponding infinitesimal generator is

$$\begin{aligned}&\left.\frac{d}{dt}\right|_{t=0} g_1 g_2 \exp\left(\mathrm{Ad}_{g_2^{-1}} t\xi_1\right) \exp(t\xi_2) \\ &= g_1 g_2 \left(\mathrm{Ad}_{g_2^{-1}} \xi_1 + \xi_2\right) \\ &= \xi_G(g_1 g_2),\end{aligned}$$

where $\xi_G = \mathrm{Ad}_{g_2^{-1}} \xi_1 + \xi_2$ is the infinitesimal generator. From this relation, we form the semi-direct product $G \ltimes_{\mathrm{Ad}} \mathfrak{g}$ and define the group operator $*$ to be

$$* : (G \ltimes_{\mathrm{Ad}} \mathfrak{g}) \times (G \ltimes_{\mathrm{Ad}} \mathfrak{g}) \to G \ltimes_{\mathrm{Ad}} \mathfrak{g},\ ((g_1, v_1), (g_2, v_2)) \mapsto \left(g_1 g_2, \mathrm{Ad}_{g_2^{-1}} v_1 + v_2\right).$$

Note that this definition of the group operator is for left trivialization of the vector field $\mathrm{Ad}_{g_2^{-1}} v_1 + v_2$. We could have defined the curves as

$$\begin{aligned}\gamma_1 &= \exp(t\xi_1)\, g_1 \\ \gamma_2 &= \exp(t\xi_2)\, g_2\end{aligned}$$

to get

$$\begin{aligned}\gamma_1\gamma_2 &= \exp(t\xi_1)\, g_1 \exp(t\xi_2)\, g_2 \\ &= \exp(t\xi_1) \exp(\mathrm{Ad}_{g_1} t\xi_2)\, g_1 g_2\end{aligned}$$

which has the corresponding infinitesimal generator $\xi_G = \mathrm{Ad}_{g_1} \xi_2 + \xi_1$ which is right trivialized. From this relation, we would form the group operator $*$ to be

$$* : (G \ltimes_{\mathrm{Ad}} \mathfrak{g}) \times (G \ltimes_{\mathrm{Ad}} \mathfrak{g}) \to G \ltimes_{\mathrm{Ad}} \mathfrak{g},\ ((g_1, v_1), (g_2, v_2)) \mapsto (g_1 g_2, \mathrm{Ad}_{g_1} v_2 + v_1).$$

Thus the definition of the group operator $*$ is dependent on if the vector field is left or right trivialized.

Using this operator, we can define a left and right group action on the group, the identity element, and the inverse element. These results are summarized in table I.

$G \ltimes_{\mathrm{Ad}} \mathfrak{g}$	Left Trivialized	Right Trivialized
$(g_1, v_1) * (g_2, v_2)$	$\left(g_1 g_2, \mathrm{Ad}_{g_2^{-1}} v_1 + v_2\right)$	$(g_1 g_2, \mathrm{Ad}_{g_1} v_2 + v_1)$
Identity	$(I, 0)$	$(I, 0)$
$(g_1, v_1)^{-1}$	$\left(g_1^{-1}, -\mathrm{Ad}_{g_1} v_1\right)$	$\left(g_1^{-1}, -\mathrm{Ad}_{g_1^{-1}} v_1\right)$
$R_{(g_1, v_1)}(g_2, v_2)$	$\left(g_2 g_1, \mathrm{Ad}_{g_1^{-1}} v_2 + v_1\right)$	$(g_2 g_1, \mathrm{Ad}_{g_2} v_1 + v_2)$
$L_{(g_1, v_1)}(g_2, v_2)$	$\left(g_1 g_2, \mathrm{Ad}_{g_2^{-1}} v_1 + v_2\right)$	$(g_1 g_2, \mathrm{Ad}_{g_1} v_2 + v_1)$

Table I
GROUP ACTIONS OF $G \ltimes_{\mathrm{AD}} \mathfrak{g}$

Since $G \ltimes_{\mathrm{Ad}} \mathfrak{g}$ is a Lie group, it has a Lie algebra $\mathfrak{g} \ltimes_{\mathrm{ad}} \mathfrak{h}$. The matrix exponential and logarithm maps between the Lie group $G \ltimes_{\mathrm{Ad}} \mathfrak{g}$ and its Lie algebra $\mathfrak{g} \ltimes_{\mathrm{ad}} \mathfrak{h}$. They are defined as

$$\begin{aligned}\exp &: \mathfrak{g} \ltimes_{\mathrm{ad}} \mathfrak{h} \to G \ltimes_{\mathrm{Ad}} \mathfrak{g},\ (v, u) \mapsto \sum_{k=0}^{\infty} \frac{1}{k!} (v, u)^k \\ \log &: G \ltimes_{\mathrm{Ad}} \mathfrak{g} \to \mathfrak{g} \ltimes_{\mathrm{ad}} \mathfrak{h},\ (g, w) \mapsto \sum_{k=1}^{\infty} (-1)^{k+1} \frac{((g, w) - (I, 0))^k}{k}.\end{aligned}$$

These mappings will be helpful when defining the Lie bracket in the Lie algebra. To define the Lie bracket, we follow the steps outlined in subsection C. We repeat them here to show that the derivation is very similar with only minor on the definition of the Lie group.

The first step is to define the conjugate action of the Lie group on itself as

$$\psi : (G \ltimes_{\mathrm{Ad}} \mathfrak{g}) \times (G \ltimes_{\mathrm{Ad}} \mathfrak{g}) \to G \ltimes_{\mathrm{Ad}} \mathfrak{g},\ ((g_1, w_1), (g_2, w_2)) \mapsto (g_1, w_1) * (g_2, w_2) * (g_1, w_1)^{-1},$$

and the corresponding automorphism is

$$\Psi_{(g_1, w_1)} : G \ltimes_{\mathrm{Ad}} \mathfrak{g} \to G \ltimes_{\mathrm{Ad}} \mathfrak{g}.$$

The adjoint action of the Lie group on it's Lie algebra is defined as

$$\left.\frac{d}{dt}\right|_{t=0} \Psi_{(g_1, w_1)} (\exp (t (v, u))),$$

and is denoted

$$\mathrm{Ad}_{(g_1, w_1)} : \mathfrak{g} \ltimes_{\mathrm{ad}} \mathfrak{h} \to \mathfrak{g} \ltimes_{\mathrm{ad}} \mathfrak{h}.$$

The adjoint action of a Lie algebra on itself is defined as

$$\left.\frac{d}{dt}\right|_{t=0} \mathrm{Ad}_{\exp t(v_1, u_1)} (v_2, u_2),$$

and is denoted

$$\mathrm{ad}_{(v_1, u_1)} : \mathfrak{g} \ltimes_{\mathrm{ad}} \mathfrak{h} \to \mathfrak{g} \ltimes_{\mathrm{ad}} \mathfrak{h},$$

and is equivalent to the Lie bracket.

The matrix exponential, matrix logarithm, adjoint actions, and Lie bracket have simplified versions for each left and right trivialized case. These are shown in table II. In the table we have used the notation $\exp_G$ and $\log_G$ to indicate that these mappings are the ones defined on the group. Also, where there is Ad_G or $\mathrm{ad}_{\mathfrak{g}}$, these are the adjoint actions defined in the group and Lie algebra. For more information about this table, see [6].

$\mathfrak{g} \ltimes_{\mathrm{ad}} \mathfrak{h}$	Left	Right
$\exp (v, u)$	$\left(\exp_G (v), \mathrm{dexp}_{v,l} (u)\right)$	$\left(\exp_G (v), \mathrm{dexp}_{v,r} (u)\right)$
$\log (g, w)$	$\left(\log_G (g), \mathrm{dexp}^{-1}_{\log(g),l} (w)\right)$	$\left(\log_G (g), \mathrm{dexp}^{-1}_{\log(g),r} (w)\right)$
$\mathrm{Ad}_{(g,w)} (v, u)$	$(\mathrm{Ad}_g v, \mathrm{Ad}_g (-\mathrm{ad}_v w + u))$	$(\mathrm{Ad}_g v, -\mathrm{ad}_{\mathrm{Ad}_g v} w + \mathrm{Ad}_g u)$
$[(v_1, u_1), (v_2, u_2)]$	$(\mathrm{ad}_{v_1} v_2, -\mathrm{ad}_{v_2} u_1 + \mathrm{ad}_{v_1} u_2)$	$(\mathrm{ad}_{v_1} v_2, -\mathrm{ad}_{v_2} u_1 + \mathrm{ad}_{v_1} u_2)$
$\mathrm{dexp}^{-1}_{(v_1, u_1)} (v_2, u_2)$	$\sum_{k=0}^{\infty} \frac{B_k}{k!} \left(-\mathrm{ad}_{(v_1, u_1)}\right)^k (u_2, v_2)$	$\sum_{k=0}^{\infty} \frac{B_k}{k!} \left(\mathrm{ad}_{(v_1, u_1)}\right)^k (u_2, v_2)$

Table II
SEMI-DIRECT PRODUCT LIE ALGEBRA OPERATIONS

Example 30. (Lie algebra of $SE(n) \ltimes_{\mathbf{Ad}} \mathfrak{se}(n)$). This example is a continuation of example 29. Recall that $SO(n) \ltimes_{\Phi} \mathbb{R}^n$ is isomorphic to $SE(n)$, in fact $SE(n)$ is just a matrix representation of $SO(n) \ltimes_{\Phi} \mathbb{R}^n$. $SE(n)$ has the familiar matrix form

$$\begin{bmatrix} R & p \\ 0 & 1 \end{bmatrix},$$

where $R \in SO(n)$ and $p \in \mathbb{R}^n$. It's Lie algebra has the form

$$\begin{bmatrix} \omega & \rho \\ 0 & 0 \end{bmatrix},$$

where $\omega \in \mathfrak{so}(n)$ and $\rho \in \mathbb{R}^n$. We define the semi-direct product $SE(n) \ltimes_{\mathbf{Ad}} \mathfrak{se}(n)$ with the group action being defined by the left trivialized case. An example of group multiplication is

$$\left(\begin{bmatrix} R_1 & p_1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} \omega_1 & \rho_1 \\ 0 & 0 \end{bmatrix}\right) * \left(\begin{bmatrix} R_2 & p_2 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} \omega_2 & \rho_2 \\ 0 & 0 \end{bmatrix}\right) = \left(\begin{bmatrix} R_1 R_2 & R_1 p_2 + p_1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} R_2 \omega_1 R_2^{-1} & -R_2 \omega_1 R_2^{-1} p_2 + R_2 \rho_1 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} \omega_2 & \rho_2 \\ 0 & 0 \end{bmatrix}\right)$$

The identity element is

$$\left(\begin{bmatrix} I & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}\right),$$

and the inverse element is

$$\left(\begin{bmatrix} R^{-1} & -R^{-1} p \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} -R \omega R^{-1} & R \omega R^{-1} p - R \rho \\ 0 & 0 \end{bmatrix}\right).$$

The matrix exponential is

$$\exp\left(\underbrace{\begin{bmatrix}\nu & \zeta\\ 0 & 0\end{bmatrix}}_{v},\underbrace{\begin{bmatrix}\upsilon & \eta\\ 0 & 0\end{bmatrix}}_{u}\right)=\left(\exp\left(v\right),\text{dexp}_{v,l}u\right),$$

where dexp is defined in appendix **??**.

The matrix logarithm is

$$\log\left(\underbrace{\begin{bmatrix}R & p\\ 0 & 1\end{bmatrix}}_{g},\underbrace{\begin{bmatrix}\omega & \rho\\ 0 & 0\end{bmatrix}}_{w}\right)=\Big(\log\left(g\right),\text{dexp}^{-1}_{\log(g),l}\left(w\right)\Big),$$

where dexp^{-1} is defined in appendix **??**.

The adjoint action of the Lie group on the Lie algebra is

$$\text{Ad}_{(g,w)}\left(v,u\right)=\left(\begin{bmatrix}R\nu R^{-1} & -R\nu R^{-1}p+R\zeta\\ 0 & 0\end{bmatrix},\text{Ad}_{g}\left(wv-vw+u\right)\right).$$

Lastly the Lie bracket is

$$\left[\left(\underbrace{\begin{bmatrix}\nu_1 & \zeta_1\\ 0 & 0\end{bmatrix}}_{v_1},\underbrace{\begin{bmatrix}\upsilon_1 & \eta_1\\ 0 & 0\end{bmatrix}}_{u_1}\right),\left(\underbrace{\begin{bmatrix}\nu_2 & \zeta_2\\ 0 & 1\end{bmatrix}}_{v_2},\underbrace{\begin{bmatrix}\upsilon_2 & \eta_2\\ 0 & 0\end{bmatrix}}_{u_2}\right)\right]=\left(\text{ad}_{v_1}v_2,-\text{ad}_{v_2}u_1+\text{ad}_{v_1}u_2\right),$$

where

$$\text{ad}_x y=xy-yx.$$

The derivative of the exponential mapping and it's inverse is commonly used in Lie theory such as in proving the Baker-Campbell-Hausdorff formula [8], and in many numerical integration techniques for solving ODE [9]. We present the general form of the derivative of the exponential mapping with it's inverse, and it's computationally efficient form for $SO\left(3\right)$ and $SE\left(3\right)$.

E. General Form

Theorem 31. *Let $X\left(t\right)$ be a smooth matrix-valued function, then*

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

This theorem is proved in [8].

When we refer to the derivative of the matrix exponential we refer to

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

where the subscript l denotes that it is left trivialized. The right derivative of the matrix exponential is found using the adjoint

$$\begin{aligned}\text{dexp}_{X(t),r}&=Ad_{X(t)}\left(\text{dexp}_{X(t),l}\right)\\&=\frac{\exp\left(X\left(t\right)\right)-I}{ad_{X(t)}}.\end{aligned}$$

The derivative of the matrix exponential also have their geometric series representation

$$\begin{aligned}\text{dexp}_{X(t),l}&=\sum_{k=0}^{\infty}\left(-1\right)^k\frac{ad^k_{X(t)}}{(k+1)!}\\ \text{dexp}_{X(t),r}&=\sum_{k=0}^{\infty}\frac{ad^k_{X(t)}}{(k+1)!}.\end{aligned}$$

Their inverses are

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

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

where B_k are Bernoulli numbers.

Remark 32. Not every adjoint of a Lie algebra can be inverted. In these cases you must use the geometric series. In some special cases, the derivative of the matrix exponential and it's inverse have closed forms. We present the closed form solutions for $SO(3)$ and $SE(3)$ next.

F. $SO(3)$

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

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

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

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

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

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

Their corresponding inverses are

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

G. $SE(3)$

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 it's 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 $\text{dexp}_{x,l}$ is

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

We can solve for the closed form solution of this geometric series by solving for the closed form solution of the sub geometric series. We already know

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

since they were solved in subsection F. This leaves us to solve the other 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 $\text{dexp}_{x,r}$ is derived in a similar manner, which we will not do here; however, its closed form is

$$\text{dexp}_x = \begin{bmatrix} \text{dexp}_w & (-a_\theta + b_\theta (wv_\times + v_\times w) + q_x(w)) \\ & \text{dexp}_w \end{bmatrix}$$

where

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

The inverses of $\text{dexp}_{x,r}$ and $\text{dexp}_{x,l}$ 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

$$\begin{aligned} \text{dexp}_{x,l}^{-1} &= \begin{bmatrix} \text{dexp}_{-w}^{-1} & \text{dexp}_{-w}^{-1} B_{-x} \text{dexp}_{-w}^{-1} \\ 0 & \text{dexp}_{-w}^{-1} \end{bmatrix} \\ \text{dexp}_{x,r}^{-1} &= \begin{bmatrix} \text{dexp}_{w}^{-1} & \text{dexp}_{w}^{-1} B_{x} \text{dexp}_{w}^{-1} \\ 0 & \text{dexp}_{w}^{-1} \end{bmatrix} \end{aligned}$$

where

$$\begin{aligned} B_{x,l} &= (a_\theta v_\times + b_\theta (wv_\times + v_\times w) + q_{-x}(w)) \\ B_{x,r} &= (-a_\theta v_\times + b_\theta (wv_\times + v_\times w) + q_x(w)). \end{aligned}$$

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

$$\begin{aligned} y(t+h) &= y(t) + h\dot{y}(t) + \frac{h^2}{2!}\ddot{y}(t) + \frac{h^3}{3!}\dddot{y}(t) + \frac{h^4}{4!}\ddddot{y}(t) + \mathcal{O}\left(h^5\right) \\ &= y(t) + hf(t,y) + \frac{h^2}{2}\frac{d}{dt}f(t,y) + \frac{h^3}{3!}\frac{d^2}{dt^2}f(t,y) + \frac{h^4}{4!}\frac{d^3}{dt^3}f(t,y) + \mathcal{O}\left(h^5\right). \end{aligned}$$

The idea of the Runge-Kutta method is to construct a new vector field $\bar{f}(t,y)$ from $f(t,y)$ that is equivalent to Taylor series up to some order k such that

$$y(t+h) = y(t) + h\bar{f}(t,y) + \mathcal{O}\left(h^{k+1}\right).$$

In the case the RK method is of order 2, the new vector field $\bar{f}(t,y)$ would be equivalent to

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

and the Taylor series would be

$$y(t+h) = y(t) + h\bar{f}(t,y) + \mathcal{O}\left(h^3\right).$$

The derivation of RK methods of order greater than 2 is complex, and we refer the interested reader to [2]. However, we will derive the order 2 method and give an example of it.

H. Explicit RK Order 2

We return to the Taylor series expansion of $y(t)$

$$y(t+h) = y(t) + hf(t,y) + \frac{h^2}{2}\frac{d}{dt}f(t,y) + \frac{h^3}{3!}\frac{d^2}{dt^2}f(t,y) + \frac{h^4}{4!}\frac{d^3}{dt^3}f(t,y) + \mathcal{O}\left(h^5\right),$$

and take the derivative of the vector field $f(t,y)$ with respect to time

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

The derivative can be substituted into the Taylor series and then simplified to get

$$\begin{aligned}y(t+h) &= y(t) + hf(t,y) + \frac{h^2}{2}\left(f_t(t,y) + f_y(t,y)\,f(t,y)\right) + O\left(h^3\right)\\ &= y(t) + \frac{h}{2}f(t,y) + \frac{h}{2}\left(f(t,y) + hf_t(t,y) + hf_y(t,y)\,f(t,y)\right) + O\left(h^3\right).\end{aligned} \tag{17}$$

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

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

and can be written as

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

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 (17) 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

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

This particular second order method is known as the Heun's method.

To solidify what we have just shown, consider the simple ODE

$$\dot{y} = \frac{1}{2}y^2$$

whose Taylor series is

$$y(t+h) = y(t) + h\left(\frac{1}{2}y^2 + y\frac{h}{2!}\right) + \mathcal{O}\left(h^3\right).$$

Applying Heun's method we calculate the first two stages

$$
\begin{aligned}
k_1 &= \frac{1}{2}y^2 \\
k_2 &= \frac{1}{2}\left(y + \frac{h}{2}y^2\right)^2 \\
&= \frac{1}{2}\left(y^2 + hy + \frac{h^2}{4}y^4\right) \\
&= \frac{1}{2}\left(y^2 + hy\right) + \mathcal{O}\left(h^4\right)
\end{aligned}
$$

and get the final result

$$
\begin{aligned}
y\left(t+h\right) &= y\left(t\right) + \frac{1}{2}\left(k_1 + k_2\right) \\
&= y\left(t\right) + h\left(\frac{1}{2}y^2 + y\frac{h}{2!}\right) + \mathcal{O}\left(h^4\right),
\end{aligned}
$$

which is an accurate numerical integration up to the second order.

I. *Higher Stages*

An s-stage classical Runge-Kutta method is expressed as

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

where the coefficients c_i, b_i, a_{ij} are read off a Butcher Tableau. It should be noted that the number of stages involved does not determine the order of the method. For RK methods of order 4 or less, they have as many stages as their order. Afterward, you start to need more and more stages for a given order. Since this note focuses on explicit RK methods that are 4 stages, we include the corresponding Butcher Tableau

0				
1/2	1/2			
1/2	0	1/2		
1	0	0	1	
	1/6	1/3	1/3	1/6

References

[1] Ralph Abraham, Jerrold Mardsen, and Tudor Ratiu. *Manifolds, Tensor Analysis, and Applications*. Springer-Verlag, New York, first edition, 1998.

[2] J.C. Butcher. *Numerical Methods for Ordinary Differential Equations*. John Wiley & Sons, Ltd, second edi edition, 2008.

[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. Partitioned Runge-Kutta methods in Lie-group setting. *Reports in Informatics*, (202), 2000.

[6] Kenth Engø. Partitioned Runge-Kutta Methods in Lie-Group Setting. *BIT*, 43:21–39, 2003.

[7] E. Hairer, S. P. Norsett, and G. Wanner. *Solving ordinary differential equations. I. Nonstiff problems*. Springer, New York, 2nd rev edition, 1993.

[8] Brian C. Hall. *Lie Groups, Lie Algebras, and Representations: An Elementary Introduction*. Springer-Verlag New York, Inc, 2003.

[9] Arieh Iserles, Hans Z. Munthe-Kaas, Syvert P. Nørsett, and Antonella Zanna. Lie-group methods. *Acta Numerica*, 9(March 2015):215–365, 2000.

[10] John M. Lee. *An Introduction to Smooth Manifolds*. Springer Science+Business Media, LCC, 2013.

[11] Hans Munthe-Kaas. Lie-Butcher theory for Runge-Kutta methods. *BIT Numerical Mathematics*, 35(4):572–587, 1995.

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

[13] Hans Munthe-Kaas. High order Runge-Kutta methods on manifolds. *Applied Numerical Mathematics*, 29(1):115–127, 1999.

[14] Hans Munthe-Kaas and Brynjulf Owren. Computations in a free Lie algebra. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 357(1754):957–981, 1999.