MATHEMATICAL METHODS-2

TERM ESSAY

LIE GROUPS: GEOMETRY OF DYNAMICS AND GEOMETRY PRESERVING SIMULATORS

Shobhit Saheb Dey

Guided By - Prof. Jyotirmoy Bhattacharya

Department of Physics

Indian Institute of Technology Kharagpur

INTRODUCTION

The dynamics of various systems have been studied and simulated usually using Newtonian mechanics in physically the most trivial coordinates(mostly Cartesian) and this methodology demands additional diverging forces to accommodate the constraints and it completely misses to capture the geometrical features of the dynamics. On the other hand, Hamiltonian and Lagrangian mechanics allows to study the same systems on Non-Euclidean manifolds which are not only of reduced dimensionality(it is a submanifold of the Euclidean space) due to constraints but also inherently encodes the geometrical features of the dynamics. Manifolds induce a lot of geometric objects, some of which have their Euclidean analogues, which provide a higher and different rigour to physical quantities like system's state, system's velocity, conjugate momenta, state space trajectory etc. This essay is built on basic understanding, notations and conventions of differential geometry which can be found in these books([Fra11], [Nak03], [SS80]). Usually when the state space is constrained to a manifold, there is an *extrinsic* symmetry to it. And these symmetries can be characterized by particular continuous groups. For example, if some state space is constrained to a 2-sphere, the $SO(3)$ group characterizes it, a double pendulum stays on $S^1 \times S^1$ i.e 2-torus characterized by $SO(2) \times SO(2)$, a spin-1/2 particle by $SU(2)$ etc. The interesting feature of some continuous groups like $SO(3)$ is that group elements themselves can build a manifold where the chart will be the independent set of parameters that generate the group elements. These groups are called Lie groups. Now the piece that connects the symmetry of the state-space/configuration manifold and the associated Lie group is the action of any group element on the points of configuration manifold, i.e a specific diffeomorphism for every group element. Now here comes *a Non-Euclidean* way of describing the system's dynamics. Since we have an extrinsically symmetric configuration manifold, a group to characterize the symmetry and the idea of how any group element acts on the state, we can reach to any configuration by acting some unique group element on an initial configuration. When simulating the system at discrete time steps, our job reduces to finding a group element that transports the configuration between a specific interval. This method is structurally different from the generally used forward Euler scheme where given an equation $\dot{\mathbf{x}} = f(\mathbf{x}, t)$ it is discretized as $\mathbf{x}_{k+1} = \mathbf{x}_k + \Delta t f(\mathbf{x}_k, t_k)$. Now here if the configuration submanifold(which depends on $f(\mathbf{x}_k, t_k)$) is extrinsically curved in the space where $\mathbf{x}$ lies, then the forward Euler increment $\Delta t f(\mathbf{x}_k, t_k)$ is going to push the consecutive states away from the submanifold. Thus even for small Δt, the states after a large number of timesteps is going to be significantly away from the submanifold and the very essential geometrical feature is lost in this scheme. But if we update the states using finite valued actions of groups then the geometry is preserved inherently. This essay develops this technique starting from the basics of Lie Groups.

LIE GROUPS AND THEIR ACTION ON STATE SPACE

Mathematical structures that are designed keeping in mind a lot of caveats and implications must be stated in the most rigorous manner to keep things clearer. So the slightly rigorous definition of a Lie Group is as follows:

A manifold G which is also a group (G,μ) i.e there exists a map $\mu : G \times G \to G$, said to be the group operation here. The group operation μ and inversion $\lambda : x \to x^{-1}$, both of them should be smooth maps.

From a blurred view, a finite dimensional Lie Group is a manifold where each point is a group element endowed with an operation μ and we can assign charts to this manifold as sets of parameters that labels the continuous set of group elements. The most trivial Lie Group could be $\mathbb{R}^n$ with the operation of vector addition. A non-trivial group that contains many of the important continuous groups as its subgroups/submanifolds is the General Linear Group $\mathbb{GL}(n,\mathbb{R})$. This is the group on $n \times n$ invertible(non-singular) real matrices, hence it is n^2-dimensional manifold. $SO(n)$ is its subgroup and hence an n-dimensional submanifold. The same group can also be defined over the complex field. Now the important point here to note is that group elements of the manifold themselves are abstract mathematical objects while the system's dynamics resides and evolves in some separate, physically describable(computable) manifold. In ordinary group theory, the abstract groups act on the physical system usually described by state vectors through matrices(linear transformations) called representation of the group. Similarly the Lie Group act on system through diffeomorphisms on the configuration manifold to itself. A realization R of a Lie Group G on some manifold M is a map from $g \in G$ to some diffeomorphism $D(g) : M \to M$ such that (i) $D(e) = I$ i.e identity transformation, (ii) $D(g^{-1}) = (D(g))^{-1}$, (iii) $D(gh) = D(g) \circ D(h)$. Further if map $R : G \to D(M)$ is one-one, then the realization is said to be reliable. To distinguish between the 2-manifolds, consider a realization of $SO(3)$ on S^2 where every element $g \in SO(3)$ links to a map $D(g)$ that rotates S^2 about a some axis by some angle. If we consider $\mathbb{R}^3$ instead of S^2, then the realization is nothing but a representation in usual sense and also the group element itself.

Now coming to the 3 piece chain that actually provides dynamics to the system. The tangent space at e in the Lie Group, the pushforwarded tangent space via the realization in the configuration manifold and integral curves.

In the Lie Group, any neighbourhood of e is mapped to a neighbourhood of g when operated via μ through *left operation* of g. Let us call this map $\mu_g : N_e \to N_g$, so its pushforward $(\mu_g)_*$ will be a bijective map between the tangent spaces T_e and T_g. Consider a vector field $V(h)$ on G. If $V(g) = (\mu_g)_* V(e)$, then the field V is said to be left-invariant in G. If we choose any vector $V \in T_e$, then we can generate a whole vector field through pushforwards at all the other group elements. Hence the left invariant vector fields form an n-dimensional space and this space is called the Lie Algebra $\mathfrak{g}$ of the group Lie Group G. Reason it is called an algebra is that for any chosen set of n linearly independent vectors from T_e, which corresponds to n vector fields, the Lie-bracket of any 2 vector at any point is a linear combination of other vectors and the combination is universal throughout G, specific only to a chosen set/basis. In formal terms, $[V_i, V_j] = c^k_{ij} V_k$ and c^k_{ij} is a constant tensor(stating without proof).

The next piece shows how the Lie Algebra acts on the state space. For a realization R of Lie Group G to manifold M, e maps to identity transformation in M. But consider a slight deviation(slightness/infinitesimal defined via chart not any metric) from e towards an element h. This element realizes to a diffeomorphism $D(h)$ that maps the points in M to slightly close points(again by chart). So if the deviation $e \to h$ is along some vector V, then this induces a vector field $W(x)$ in M where $W(x)$ corresponds to direction and relative magnitude by which point x was infinitesimally deviated via a diffeomorphism $D(h)$. In the case of the realization of $SO(3)$ on S^2 the induced vector fields are directors for rotation of the sphere about some axis. The space of fields induced by the entire algebra $\mathfrak{g}$ will be generators(or Killing vectors) for S^2.

And now the final piece, the integral curves. Once we have chosen a vector $V \in T_e$, we have also induced a unique vector field on M via the realization. So if the system evolves in time along the direction of the vector induced at each point, then trajectory it takes is an integral curve of a set of vectors from the field. But as mentioned earlier that any point in the state space can be reached from any initial point via a transformation realized by a group element. Thus an integral curve in state space can be said to be generated by an unique integral curve in the Lie Group. Both the integral curves are parameterized by time(or some affine parameter in case of relativistic mechanics) as $g(t)$ and $x(t)$ such that $g(0) = e$ conventionally. If the system's evolution doesn't change the track, i.e if it follows the same integral curve, it would imply that pre-image vector V from the algebra $\mathfrak{g}$ is unchanged and hence the corresponding integral curve in the Lie-Group is the one generated by a left-invariant vector field. Such integral curves form one-parameter subgroups i.e $g(t)g(h) = g(t+h)$. In the specific case of $\mathbb{GL}(\mathbb{R}, n)$ groups that will be used to common symmetries in physics, the tangent spaceT_e is the space of *all* n matrices. And the integral curves generated by those matrices are computed using the following theorem:

The integral curve $g(t)$ with $g(0) = h$ on the group $\mathbb{GL}(\mathbb{R}, n)$, of the left-invariant vector field X_A generated by a vector $A \in T_e$ is given by:

$$g(t) = h \exp(At) \tag{1}$$

Where exp is the usual matrix exponential. This is a specific case of exponential map used to describe the one-parameter groups. And we now have all the analytic arsenal to build a geometry preserving algorithm for simulators.

DISCRETIZING THE ACTION OF LIE-GROUP

To understand how a Lie-Group can be used as a fundamental approach to simulate dynamical systems rather than the using the usual scheme of adding update vectors, we work through an example. Consider a system that evolves in $\mathbb{R}$ according to the following equation.

$$\begin{pmatrix} \dot{x_1} \\ \dot{x_2} \\ \dot{x_3} \end{pmatrix} = \begin{pmatrix} -x_2 + x_1 x_3^2 \\ x_1 + x_2 x_3^2 \\ -x_3(x_1^2 + x_2^2) \end{pmatrix} \tag{2}$$

A key feature we noting in the dynamics is that $\frac{\mathrm{d}(x_1^2+x_2^2+x_3^2)}{\mathrm{d}t} = 2(x_1\dot{x_1} + x_2\dot{x_2} + x_3\dot{x_3}) = 0$ by using Eq(2). Thus the states are constrained to stay on the S^2 submanifold. We know that this manifold is extrinsically curved, so any numerical scheme that uses update vectors as $\mathbf{x}_{k+1} = \mathbf{x}_k + \Delta\mathbf{x}$, after a long number of iterations, the state is going to be pushed away from the S^2 manifold(as shown in Figure 1). But from the previous discussion we know that any two states $\mathbf{x}(t)$ and $\mathbf{x}(0)$ can be connected by a realization of a group element $g \in SO(3)$. In this particular case, realizations are equivalent to *representations* of $SO(3)$ due to the choice of the atlas which is in turn equivalent to the group element itself due to the dimensionality of representation. So once we have set up the problem, we address the 3 pieces discussed in the previous section in the same order.

The group theory, the continuous groups are characterized by generators and parameters. The method to find the generators for linear groups like $SO(n)$ or $SU(n)$ is similar. The group elements whose representations are slight deviations about the identity are written as:

$$A = 1 + \Sigma_i (J_i \theta_i) \tag{3}$$

Here J_i are called generators w.r.t the parameters θ_i. But these slight deviations precisely

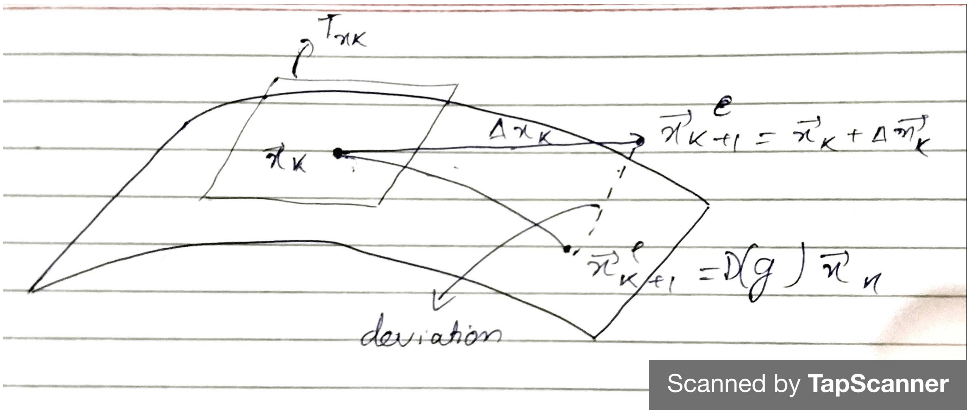


Figure 1: Difference between group action update and vector addition update

describe the tangent space of the Lie Group $SO(3)$ at e when the local coordinate chart is θ_i. Thus the set J_i actually contains the basis vectors for the tangent space T_e and hence for the Lie-Algebra denoted by $\mathfrak{so}(3)$. The standard choice(not exclusive though) for the generators is:

$$J_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}, \; J_2 = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \; J_3 = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{4}$$

Now what about the vector field induced by these so called generators. As mentioned earlier, in this specific case any matrix $A \in SO(3)$ is its own representation, so its action on $\mathbb{R}^3$ is just $\mathbf{x} \to A\mathbf{x}$, so the induced vector by J_i simply has the components of $J_i\mathbf{x}$. So the induced vectors are:

$$V_1 = x_3 \frac{\partial}{\partial x_2} - x_2 \frac{\partial}{\partial x_3}, \; V_2 = -x_3 \frac{\partial}{\partial x_1} + x_1 \frac{\partial}{\partial x_3}, \; V_3 = x_2 \frac{\partial}{\partial x_1} - x_1 \frac{\partial}{\partial x_2} \tag{5}$$

Now here is the discretization part. If the system evolved about a specific left-invariant vector field in $SO(3)$ then it would have implied a constant revolution about some orbit with a constant speed. But in general this isn't the case. If the axis or speed of rotation changes, it corresponds to change in a vector from $\mathfrak{g}$. So if we discretize the time in intervals of Δt, then between any 2 timesteps t_k and $t_k + \Delta t$ we can assume a the system rotates about a fixed axis by a constant angular velocity. Mathematically it implies that within this interval the trajectory in $SO(3)$ is the integral curve of a fixed left-invariant vector field through e. Using Eq(1), it would imply that we can transform between states $\mathbf{x}_k$ and $\mathbf{x}_{k+1}$ as:

$$\mathbf{x}_{k+1} = \exp(\Delta t A_k)\mathbf{x}_k \tag{6}$$

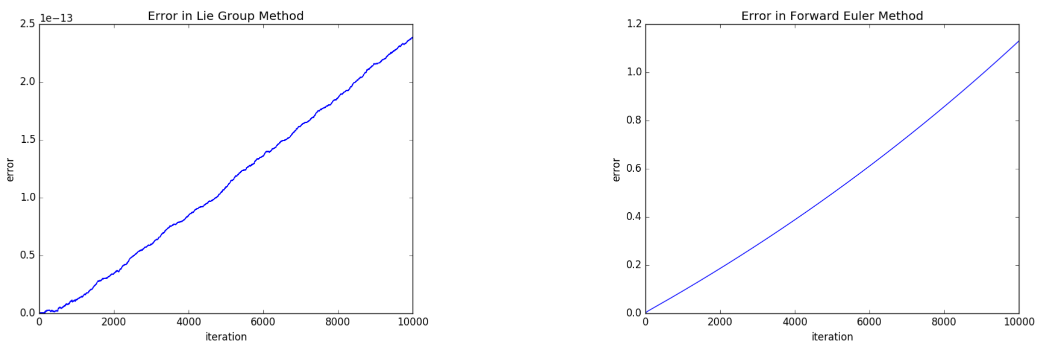
Now the job reduces to finding A_k[CMO14]. But how do we proceed? The key is that $\dot{\mathbf{x}}$ is created by the pushforward of A_k via the realization or representation in this case. So the linear combination (c_1, c_2, c_3) of tangent vectors in Eq(5) that results is $\dot{\mathbf{x}}(t_k)$, it is the same linear combination that produces the instantaneous tangent vector in $\mathfrak{g}$ i.e $A_k = \sum_i c_i J_i$. So to solve for (c_1, c_2, c_3) we write the problem in formal term:

$$\begin{pmatrix} -x_2 + x_1 x_3^2 \\ x_1 + x_2 x_3^2 \\ -x_3 x_1^2 - x_3 x_2^2 \end{pmatrix}_k = c_1 \begin{pmatrix} 0 \\ x_3 \\ -x_2 \end{pmatrix}_k + c_2 \begin{pmatrix} -x_3 \\ 0 \\ x_1 \end{pmatrix}_k + c_3 \begin{pmatrix} x_2 \\ -x_1 \\ 0 \end{pmatrix}_k \tag{7}$$

This equation is *exact* since we are working on the same tangent space. Solving this equation we get $c_1 = x_2 x_3,\ c_2 = -x_1 x_3,\ c_3 = -1$. So using $A_k = \sum_i c_i J_i$ we get:

$$A_k = \begin{pmatrix} 0 & -1 & x_1 x_3 \\ 1 & 0 & x_2 x_3 \\ x_1 x_3 & -x_2 x_3 & 0 \end{pmatrix} \tag{8}$$

In order to demonstrate the efficiency of Lie-group method, I simulated the system given by Eq(7) from the same initial state $(1,1,1)^T$ and using the same value for same time discretization of $\Delta t = 0.01$ and for same number of iterations: 10000. The geometric error $\left|\sqrt{\mathbf{x}^T \mathbf{x}} - \sqrt{3}\right|$. The vast difference can be clearly seen in figure 2. The Lie-Group method is about a 10^{13} times more efficient, i.e almost in machine precision, when it comes to preserving the geometric features of dynamics. Many systems studied in physics or dynamics can be described by



(a) Geometric deviation in Lie-Group method **(b)** Geometric deviation in Euler method

Figure 2: Note the difference in order of error between Lie Group method and Forward Euler method. Lie Group has an error less by an order of 10^{13}

$SO(3)$ symmetry. Another example in this class could be the symmetric top, described by the Euler equations. Systems in different topologies like SU(2) could be simulated by the same procedure. The only differences are the dimensionality(2 in this case) and generators(Pauli matrices in this case). Similarly, a double pendulum can be described in $S^2 \times S^2$ characterized by the Lie-Group $SO(2) \times SO(2)$.

Conclusion

If we can analytically derive the submanifold where the dynamics evolve we can deploy the Lie Group methods. The time evolution of system is generated by the Lie Group in the following way. We realize the group elements by diffeomorphisms on the configuration manifold. In the Lie group we compute a basis set that spans the associate Lie Algebra and hence generates the left-invariant vector fields. Each vector of the algebra induces a vector field in the configuration manifold and hence a family of trajectories the system can take. But due to the dynamics, the trajectories aren't geometrically trivial and hence there is an instantaneous vector from the Lie-algebra that predicts the direction of evolution, the instantaneous vector along the trajectory. Thus we have transformed the dynamics from configuration space to the coefficients of vectors in Lie Algebra. Now via the exponential map for integral curve in Lie Group, we characterize the system's trajectory. And while discretization we assume that system follows a fixed integral curve in Lie Group for a short time interval, analogous to assumption in Euler scheme of constant gradient for a short interval. And through group operations of finite effect we update the system's state but since it is a group operation, it is ensured that the geometric features are going to be preserved.

Bibliography

[SS80] Bernard F Schutz and Director Bernard F Schutz. *Geometrical methods of mathematical physics*. Cambridge university press, 1980.

[Nak03] Mikio Nakahara. *Geometry, topology and physics*. CRC press, 2003.

[Fra11] Theodore Frankel. *The geometry of physics: an introduction*. Cambridge university press, 2011.

[CMO14] Elena Celledoni, Håkon Marthinsen, and Brynjulf Owren. “An introduction to Lie group integrators–basics, new developments and applications”. In: *Journal of Computational Physics* 257 (2014), pp. 1040–1061.