When does local imply global?

An Introduction to Lie Theory

Panos Tsimpos

December 2021

1 Motivation

In mathematics we study structures. A particularly good way to understand a structure is to simplify it. For example, it is simpler to present an n^{th} degree polynomial with complex coefficients as a product of its n roots or an Abelian Group as the direct product of cyclic groups of prime order. So, given a complex structure, we often want to break it down to a collection of simpler pieces that we understand better.

When one pairs abstract algebra with topology, something interesting happens: notions of algebraic simplicity and complexity meet a notion of nearness, a notion of shape. As such, there arise algebraic structures that we can understand *locally.* Near some element, we reduce to a familiar algebraic structure, but away from it, who knows what may happen? Such is the case in Matrix Lie Theory, the study of continuous symmetry using matrices. A Matrix Lie Group is a group with a topology. Near the identity of the group, we reduce to linear algebra: the simplest kind of an abstract algebra! This local object surrounding the identity we call the *Lie Algebra* of the Matrix Lie Group. It is therefore natural to ask what can we know for the group G if we understand its algebra? In other words, what does local behaviour say of global behaviour?

2 Matrix Lie Groups: An Introduction

First, what exactly is a Lie Group? Intuitively, one can think of Lie Group as a group G (the familiar algebraic object) that is *smooth*. Here, one can think of smoothness as a regularity condition in the way one "moves" in the group, that is, goes from one element of the group to another. The group operation is precisely that which allows "movement" in the group – analogously taking an inverse (the inverse operation) can be thought of as "movement" in the group – so we are motivated to demand our multiplication $m : G \times G \to G$ and inverse $i : G \to G$ to be smooth. For brevity, m and i are going to be referred to collectively as the "group operations" of G. To summarize:

Informal Definition 1. *A Lie Group is a group G with smooth group operations.*

Example 1. From abstract algebra, we note that the positive real numbers under multiplication form the group $(\mathbb{R}_{>0}, \cdot)$. By elementary calculus the group operations $(x, y) \mapsto x + y$ and $x \mapsto \frac{1}{x}$ are infinitely differentiable functions of $\mathbb{R}^2_{>0}$ and $\mathbb{R}_{>0}$, respectively. Thus, our group operations are smooth in the analytic sense, that is, they are differentiable n times, where n is any non-negative integer. Following our informal definition $(\mathbb{R}_{>0}, \cdot)$ is a Lie Group.

The avid reader will be skeptical, if not entirely enraged, at this point. We have "defined" a Lie Group as a group with smooth operations but we never defined smoothness in generality. A proper definition requires the introduction of **manifolds**, topological spaces that are locally similar to $\mathbb{R}^n$. Such a discussion, however, would be long and outside the scope of this paper. Thus, we will restrict our attention to Matrix Lie Groups, groups of matrices obeying "smoothness conditions". In the general treatment of Lie Groups, these matrix groups would (as they should) be Lie Groups. However, as opposed to general Lie Groups, for Matrix Lie Groups we can give smoothness conditions quickly and concisely.

First, recall that a Matrix Group is a set S of real matrices that forms a group. As such, $\forall A \in S, A^{-1}$ exists. First, to even speak of an inverse, A must be a square matrix, say an n by n matrix[1]. Moreover, by linear algebra, we have that A^{-1} exists if and only if $\det(A) \neq 0$. Thus, S is a subset of $GL_n(\mathbb{R})$ the set of n by n real matrices with non zero determinant. Recall that $GL_n(\mathbb{R})$ forms a group under matrix multiplication. Therefore, S is a subgroup of $GL_n(\mathbb{R})$.

Thus, we want to say that a Matrix Lie Group G is as subgroup of $GL_n(\mathbb{R})$ that satisfies some mystery smoothness condition. Here it is[2]:

Definition 1. *A Matrix Lie Group is a subgroup G of $GL_n(\mathbb{R})$ that satisfies: If $\{A\}_{n>0}$ a sequence in G and $\lim_{n\to\infty} A_n = B$ where B is invertible, then $B \in G$. This is called closure under non-singular limits. It is essential for the later discussion of the* log *and* exp *functions.*

The natural question arises: what exactly is a limit of matrices? To understand this, we have to swift our perspective:

Proposition 1. *$M_n(\mathbb{R})$ and $\mathbb{R}^{n^2}$ are isomorphic as vector spaces.*

Proof. Let $T : M_n(\mathbb{R}) \to \mathbb{R}^{n^2}$ be the map that takes $T(A)$ to the tuple $(x_1, x_2, \dots, x_{n^2})$ where $x_1 = A_{11}$, $x_2 = A_{12}$ and so on until we exhaust the first row of A, then $x_{n+1} = A_{21}$, $x_{n+2} = A_{22}$ and so on until we exhaust the second row of A and we carry on until we exhaust all entries of A. Note that this is well defined since the number of entries of A and the number of entries in a tuple of the codomain are the same. Clearly, this is a surjuctive map. Moreover, since addition and scalar multiplication of matrices are defined element wise, T is a linear map. Finally, if $T(A) = T(B)$ as elements of $\mathbb{R}^{n^2}$ then all of their entries are equal, one by one. Thus, $\forall (i, j)$ we have $A_{ij} = B_{ij}$ therefore $A = B$, proving injectivity. In other words, T is a bijective linear map between $M_n(\mathbb{R})$ and $\mathbb{R}^{n^2}$ thus an isomorphism. ☐

What this is telling us is that we are allowed to view n by n real matrices as points in $\mathbb{R}^{n^2}$ at least as far as linear algebra is concerned. Since, however, at our level the study of Matrix Lie Groups is almost based entirely on linear algebra, this isomorphism will suffice to pass seamlessly between the two objects. In fact, we will often do so without mentioning, and the above theorem will be at play in the background.

This fresh perspective will give us a notion of distance:

Definition 2. *A metric d in $M_n(\mathbb{R})$ is given by:*

$$d(A, B) = \sqrt{\sum_{i,j} (A_{ij} - B_{ij})^2}$$

Denote $d(A, B)$ as $|A - B|$.

This is exactly the Euclidean 2-distance of A,B as points in $\mathbb{R}^{n^2}$. To be precise, we can show this is well defined by observing that $|A - B| = d_2(T^{-1}(A), T^{-1}(B))$ where d_2 is the Euclidean 2-distance in $\mathbb{R}^{n^2}$ and T is the isomorphism discussed above.

So $(M_n(\mathbb{R}), |\cdot|)$ is a metric space! In particular, let the topology $\mathcal{T}$ of $M_n(\mathbb{R}$ be exactly that which is induced by $|\cdot|$. Now a limit of a sequence of n by n real matrices taken with respect to the $|\cdot|$ distance. Put differently, it is just the limit of a sequence of points in $\mathbb{R}^{n^2}$.

This allows us to give a more elegant definition of a Matrix Lie Group:

Definition 3. *A Matrix Lie Group G is a closed subset of $GL_n(\mathbb{R})$.*

Proposition 2. *Definitions 1 and 3 are equivalent.*

[1]To see this, appeal to the rank nullity theorem: if $T : V \to W$ is a linear map between vector spaces, $\dim(V) = \dim(\ker(T) + \dim(\text{rank}(T)$. For an inverse to exist, the map must be surjuctive, so $\dim(W) = \dim(\text{rank}(T))$. But if $\dim(V) < \dim(W)$ then $\dim(\ker(V)) < 0$, impossible. In the case that $\dim(V) > \dim(W)$ then $\dim(\ker(T)) > 1$ and so our map is not injective.

[2]As given by Stillwell.

Proof. Recall that a closed set, in a metric space, is a set that contains its limit points. Let $\{A_n\}$ be a sequence of matrices in $M_n(\mathbb{R})$ with limit A. Then $\{A_n\} \in GL_n(\mathbb{R})$ if and only if for all i, $\det A_i \neq 0$. If definition 1 holds then $\det A \neq 0$ so $A \in GL_n(\mathbb{R})$. But A was a limit point of a generic sequence in $GL_n(\mathbb{R})$ so definition 2 follows. Conversely, if definition 2 holds then, since A is a limit point of $\{A_n\}$, closure implies that $A \in GL_n(\mathbb{R})$, so the determinant of A must be nonzero. Therefore, A is non-singular. □

Let's now consider some examples:

Example 2. $SO(n) = \{A \in M_n(\mathbb{R}) | AA^T = \mathbf{1}, \det A > 0\}$ is a matrix lie group. Note that $\det(AA^T) = 1$ and $\det A > 0$ implies $\det A = 1$. Now $SO(n)$ is indeed a subgroup of $GL_n(\mathbb{R})$ by a direct calculation. Moreover, assuming that det is a continuous map[3] $\det : M_n(\mathbb{R}^n) \to \mathbb{R}^n$, for some sequence $\{A_n\}$ where $A_i \in SO(n)$, with limit A we have:

$$\det A = \det(\lim_{n\to\infty} A_n) = \lim_{n\to\infty} \det(A_n) = \lim_{n\to\infty} (+1) = 1$$

The second equality follows from the continuity of the determinant and the third equality follows from the definition of $SO(n)$. Thus, in particular, $\det A \neq 0$ so A is non-singular.

$SO(n)$ is a special group. As discussed in linear algebra, $AA^T = 1$ is equivalent to $A\vec{x} \cdot A\vec{y} = \vec{x} \cdot \vec{y}$ for arbitrary $\vec{x}, \vec{y} \in \mathbb{R}^n$. In other words, A preserves the inner product. In $\mathbb{R}^n$ we can define a metric in terms of an inner product. Thus, if A preserves distances then A preserves distances in $\mathbb{R}^n$. That is, A is an *isometry*. Moreover, A clearly fixes the origin. Lastly, $\det A > 0$, that is, A preserves orientation[4]. In summary, $SO(n)$ is an orientation preserving isometry that fixes the origin. These are exactly the conditions obeyed by a rotation of the coordinate system about the origin in $\mathbb{R}^2$ and $\mathbb{R}^3$. A rotation *about the origin* fixes the origin by definition. Moreover, we visualize it as a "rigid motion" so any two points at distance d always remain at distance d. Finally, a rotation must "preserve orientation", that is, we do not allow reflections. Thus, it is natural to define a rotation in higher dimensions as:

Definition 4. *A rotation in $\mathbb{R}^n$ is any linear map $L : \mathbb{R}^n \to \mathbb{R}^n$ with a corresponding matrix M_L that obeys:*

$$M_L M_L^T = 1$$

and

$$det M_L > 0$$

Equivalently, a map L of which the matrix M_L is in SO(n)

SO(n) is the group of rotations[5] of n dimensional Euclidean space. We have shown that it is a Lie Group. In a sense, we should have worked backwards: start from SO(n), as the canonical example of what a Matrix Lie Group must be, and then generalize as we did above. Indeed, SO(n) joins the notions of smoothness and a group somewhat transparently. SO(n) is a group: if we rotate one way we can always rotate back to where we started, so every rotation has an inverse. Doing nothing can be viewed as a rotation by 0 degrees, so we have an identity. Finally, rotations are associative (although importantly not commutative). Thus, intuitively, rotations form a group. However, there is more to them. We can choose to rotate by an infinitesimally small angle $d\theta$ and then quotient the resulting matrix by $d\theta$, recovering some notion of a derivative. Note that this makes sense only because *an infinitesimal rotation* makes sense.

Recall the picture of a rotating vector $\vec{v}$ in the plane: it is valid to ask for the *generator* of a rotation. This ends up being a vector $\vec{w}$ orthogonal to $\vec{v}$: we visualize $\vec{w}$ with tail at the tip of $\vec{v}$. It, momentarily, pushes $\vec{v}$ in the orthogonal direction. When $\vec{v}$ moves a little bit, $\vec{w}$ tips over a little bit to remain orthogonal to $\vec{v}$ so it keeps pushing $\vec{v}$ in the orthogonal direction. This, we imagine, generates the rotation.

In summary, a Lie Group is an object that encodes *continuous symmetry* and $SO(n)$ exemplifies this. Indeed, the sphere in n dimensions, $\mathbb{S}^{n-1}$ is symmetric under the action of $SO(n)$.

We can generalize the above to matrices $\mathbb{C}$. Specifically:

---

[3]This is a fact that we will use again and again. One (cumbersome) yet direct way to see it is to recall that the determinant of a matrix M can be *expanded in co-factors of M*, an inductive process that involves sub-matrices of M resulting from deleting rows and columns of M. The result is that we can write the determinant as a polynomial in the entries of M. Its continuity, then, follows by analysis.

[4]Orientation here means the positive direction set by an arbitrary choice of basis

[5]It was recently pointed to me that SO(n) is best described of as the group of *simultaneous rotations* but this is a longer discussion.

Example 3. Let $SU(n) = \{A \in M_n(\mathbb{C}) | AA^T = \mathbf{1}, \det A = 1\}$. Observe the close relation to $SO(n)$: $SU(n)$ is the group of orientation preserving isometries that fix $\vec{0}$ in n dimensional complex space. Moreover, for some sequence $\{A_n\}$ where $A_i \in SU(n)$, with limit A, we have

$$|\det A| = |\det(\lim_{n\to\infty} A_n)| = |\lim_{n\to\infty} \det(A_n)| = \lim_{n\to\infty} |\det(A_n)| = \lim_{n\to\infty} (+1) = 1$$

where, as before, the second equality follows from the continuity of the determinant, the fourth equality follows from the definition of of $SO(n)$ and the third by the continuity of the complex norm[6]. Thus, in particular, $detA \neq 0$ so A is non-singular.

3 Tangent Spaces

In the linear algebra of 2 dimensions, one can think of the *tangent space* T at a point x of a (smooth) curve C as the line tangent to C at x. It encodes the "instantanious direction" of the curve C at x, given as a unit vector in T. Equivalently, the unit tangent vector of C at x encodes the direction "one may move starting at x whilst remaining on C". Similarly, one can think of the tangent space Q of a surface S at x as a plane generated by the tangent vectors of S at x. Again, it tells us the directions we can move towards, starting at x, and remain on S.

The tangent, space, then captures information about the local behaviour of the space. Motivated by this, we wish to define the tangent space of a Matrix Lie Group G. We wish that such a construction will capture the local behaviour of G. Not only do we succeed but, miraculously, we obtain much more: under conditions, the tangent space of G determines G. This is just baffling. Much work will be needed, however, to see this.

To know what a tangent space is, we first need to know how to "move" within a Matrix Lie Group. Thus, we have the following definition:

Definition 5. *Let S is a space of matrices. A smooth path in S is a map from $\phi : I \subset \mathbb{R} \to S$, $t \mapsto A$ such that for any i,j the map $t \mapsto A_{ij}(t)$ is differentiable. More formally, if $S \subset \mathbb{R}^{n^2}$ then map $\phi : I \subset \mathbb{R} \to S$ is a smooth path if $\forall 1 \leqslant i \leqslant n^2$, $\pi_i \circ \phi :\to \mathbb{R}$ is a differentiable function, where π_i is the standard projection of the i^{th} coordinate. The notations $A(t)$ in place of $\phi(t)$ will often be used.*

Moreover, multivariate calculus tells us how to take the derivative of such a ϕ, since the its codomain is nothing other than a subset of $\mathbb{R}^{n^2}$:

$$A'(t) = (a'_{11}(t), a'_{12}(t), \ldots, a'_{nn}(t)) = \begin{pmatrix} a'_{11}(t) & \cdots & a'_{1n}(t) \\ \vdots & \ddots & \vdots \\ a'_{n1}(t) & \cdots & a'_{nn}(t) \end{pmatrix}$$

Now, we are in a position to define the tangent space of a Matrix Lie Group.

Definition 6. *The tangent space of a Matrix Lie Group $G \subset M_n(\mathbb{R})$ at a point X is the set of $S \in M_n(\mathbb{R})$ that satisfie $S = A'(0)$ for some smooth path in G obeying $A(0) = X$. It is denoted by $T_X(G)$.*

Of special interest to us is the tangent space *at the origin* where we take our point $X \in G$ to be the identity matrix, denoted by $\mathbf{1}$.

Example 4. Recall the *orthogonal group* $O(n) = \{A \in M_n(\mathbb{R}) | AA^T = \mathbf{1}\}$. It is the natural generalization of $SO(n)$. Note that for $A \in O(n)$, $\det A = \pm 1$. O(n) is the group of isometries of $\mathbb{R}^n$ that fix the origin, reflections allowed. What is $T_{\mathbf{1}}(O(n))$? Well, this is a computation. If $X \in T_{\mathbf{1}}(O(n))$ then we must have a path $A(t)$ in O(n) such that $A(0) = \mathbf{1}$ and $A'(0) = X$. By definition, $A(t)A^T(t) = 1$ for all t. Note[7] that $[A(t)A^T(t)]' = A'(t)A^T(t) + A(t)A'(t)^T$. Note further that $[A^T(t)]' = A'(t)^T$, since the derivative acts on each entry separately. Therefore:

$$[A(t)A^T(t)]' = \mathbf{1}' \Rightarrow A'(t)A^T(t) + A(t)A'(t)^T = \mathbf{0}$$

[6]To see this, it suffices to recall the "inverse triangle inequality": for any two $x, y \in \mathbb{C}$ distinct we have $||x| - |y|| < |x - y|$

[7]This is the chain rule applied to the composition $m(A(t), A^T(t))$ where $m : G \times G \to G$ the multiplication map $(A, B) \mapsto AB$. Recall that $D(m)_{(A,B)}(X, Y) = AY + BX$

for all t. So, evaluating at $t = 0$ we have:

$$A'(0)A^T(0) + A(0)A'(0)^T = \mathbf{0} \Rightarrow X + X^T = \mathbf{0}$$

Thus, any matrix in $T_{\mathbf{1}}(O(n))$ is skew symmetric. So, in particular, $T_{\mathbf{1}}(O(n))$ is a subset of the set of skew symmetric matrices. But this is only half the story. What matrices do, in fact, live in the tangent space? For that, we need a new tool the *Matrix Exponential Function.*

Definition 7. *The matrix exponential of $A \in M_n(\mathbb{R})$ is the power series:*

$$exp(A) = e^A = \mathbf{1} + A + \frac{A^2}{2!} + \frac{A^3}{3!} + \dots$$

where $\mathbf{1}$ *is the identity matrix in standard form.*

First things first: is this well defined? Observe that it converges absolutely. Consider:

$$\mathbf{1} + |A| + \frac{|A|^2}{2!} + \frac{|A|^3}{3!} + \dots$$

This series is just the power series expansion of $e^x : \mathbb{R} \to \mathbb{R}$ at 0, evaluated at $x = |A|$. We know from analysis that this power series converges back to e^x for any $x \in \mathbb{R}$. Thus, $\exp : \mathbb{R}^{n^2} \to \mathbb{R}^{n^2}$ converges absolutely for any $A \in \mathbb{R}^{n^2}$. An aside:

Remark 5. Since any power of A commutes with any other power of A, we can use the rules of algebra in a commutative ring to write:

$$\exp(A) = \lim_{n\to\infty} \left(\mathbf{1} + \frac{A}{n!}\right)^n$$

This tells us that exp(A), viewed as an operator, is 'built' by successive applications of A. This equation says: for some large $n \in \mathbb{Z}_{>\nVdash}$ applying the operator exp(A) is the same is taking A, scaling it down by n!, adding it to the identity and repeating all this n times. Since $n!$ is large, $\frac{A}{n!}$ is nominally small, so $\mathbf{1} + \frac{A}{n!}$ acts like a perturbation. We then compound this perturbation a large number of times (n of them) to get the desired result. For example, consider a rotation about the origin by an angle θ: we can view it as the result of many successive rotations, each of angle $d\theta$.

We now note a property of the exp:

Proposition 3. *If $A, B \in M_n(\mathbb{R})$ and AB=BA then $e^Ae^B = e^{A+B}$.*

Proof. We urge the reader to look at at Stillwell's *Naive Lie Theory*, section 5.2. The proof is purely an algebraic argument: we can expand both sides in power series and note that since A commutes with B, the calculation is identical to the case of real numbers. But in such a case, we know this is a true statement. ☐

We want to describe the tangent space of SO(n) at the origin. We know it is a subset of the set of skew symmetric matrices. Let X by any skew symmetric matrix. Consider the path e^{tX} for $0 \leqslant t \leqslant 1$. Observe that:

$$\frac{d(e^{tX})}{dt}|_{t=0} = Xe^{0X} = Xe^{\mathbf{0}} = X$$

and

$$e^{tX}|_{t=0} = e^{\mathbf{0}} = \mathbf{1}$$

Thus, this is a path with tangent X at the identity. If only we knew that e^{tX} is a path *in* SO(n), we would conclude that X is in the tangent space. This is just a calculation: First, observe that $(e^X)^T = e^{X^T}$ because for any $m \in \mathbb{Z}$ we have $(X^m)^T = (X^T)^m$ thus:

$$(e^X)^T = (\sum_n \frac{X^n}{n!})^T = \sum_n \frac{(X^n)^T}{n!} = \sum_n \frac{(X^T)^n}{n!} = e^{X^T}$$

where the second equality holds because taking a transpose is a linear operation. Now, since X is skew symmetric, we have $X^T = -X$. Then:

$$XX^T = X(-X) = (-X)X = X^T X$$

since multiplication by a scalar is commutative. So tX and tX^T commute! We can finally apply our previous proposition:

$$e^{tX}e^{tX^T} = e^{t(X+X^T)} = e^{\mathbf{0}} = \mathbf{1}$$

and we in fact have that e^{tX} is an orthogonal matrix, for any t. Now, we have to verify that $\det(e^{tX}) = 1$. This is another occasion where the continuity of the determinant is going to come to the rescue. Our path e^{tX} is a continuous map; its composition therefore with the determinant is also a continuous map. Note, however, that for all t, e^{tX} is an orthogonal matrix, thus $\det(e^{tX}) = \pm 1$. But any continuous function into a discrete space must be constant; evaluating at t=0 we have that

$$\det(e^{0X}) = \det(e^{\mathbf{0}}) = \det(\mathbf{1}) = +1$$

therefore $\det(e^{tX}) = +1$ for all t. We can finally conclude that our path lies in SO(n), as desired.

Some comments are in order: firstly, note that the same proof works for O(n); it's just the determinant argument that is redundant. So, in fact, $T_{\mathbf{1}}(SO(n)) = T_{\mathbf{1}}(O(n))$. Secondly, the exp map above seems to do a lot of the heavy lifting: it gives us a way to get from the tangent space to the Lie Group. Given an element X (tentatively) in the tangent space, we constructed an infinite family of elements e^{tX} in the Matrix Lie Group. This is not an accident. Consider another example:

Example 6. Consider $\mathbb{S}^1 \in \mathbb{C}$. The tangent space at the identity is $T_{\mathbf{1}} = \{i\theta | \theta \in \mathbb{R}\}$. This can be seen both from elementary calculus as well as intuitively: the tangent of a curve in $\mathbb{R}^2$ must be a line and the tangent at $(1, 0)$ of the circle is the vertical intersecting $(1, 0)$ namely the subspace generated by the vector $(0, 1) = i$. Now observe that $\exp(T_{\mathbf{1}}) = \mathbb{S}^1$ since any $x \in \mathbb{S}^1$ can be written as $x = e^{i\theta}$ for some real θ. So the exponential here is a *surjuctive* continuous map from the tangent space to the Lie Group. Any member of the Lie Group can be generated by an element of the tangent space. This is surprising. Remember, the tangent space was conceived as a local description of the Lie Group; here we are able to describe, at least as a set, *the entire group* from its tangent space.

Now, to extract information from the tangent space, we have to investigate its properties.

Theorem 4. *The tangent space at the origin of a Matrix Lie Group $T_{\mathbf{1}}(G)$ is a real vector space.*

Proof. $T_{\mathbf{1}}(G)$ being a vector space means that $\forall X, Y \in T_{\mathbf{1}}(G)$ we have $X + Y \in T_{\mathbf{1}}(G)$ and $\forall r \in \mathbb{R}, \forall X \in T_{\mathbf{1}}(G)$ we have $rX \in T_{\mathbf{1}}(G)$.

Given $X, Y \in T_{\mathbf{1}}(G)$ we have that $\exists A(t), B(t)$ paths in G such that $A(0) = B(0) = \mathbf{1}$ and $A'(0) = X, B'(0) = Y$. Let $C(t) = A(t)B(t)$. Then:

$$C(0) = A(0)B(0) = \mathbf{1}$$

and

$$C'(0) = A'(0)B(0) + A(0)B'(0) = X + Y$$

by differentiating the matrix product and evaluating at $t = 1$. For closure under scalar multiplication, given X, r as above, let $D(t) = A(rt)$. Then:

$$D(0) = A(r0) = \mathbf{1}$$

and

$$D'(0) = rA'(0) = rX$$

by the chain rule. □

In the proof, we observe that vector addition in the tangent space translates to matrix multiplication in the Lie Group. To find the path corresponding to the sum X, Y we took the paths of X and Y–nothing other than a parametrized family of matrices–and multiplied then at all corresponding values of the parameter. We would like to say that vector addition in the tangent space corresponds to matrix multiplication, the group operation, but that is just not true. For one, vector addition commutes whereas matrix multiplication does not. The need arises for a different kind of operation in the tangent space, one that captures the non-commutativity of the Lie Group.

Definition 8. *The commutator of real matrices is the operator* $[\cdot,\cdot]: M_n(\mathbb{R}) \times M_n(\mathbb{R}) \to M_n(\mathbb{R})$ *that takes* $X, Y \in M_n(\mathbb{R})$ *to:*

$$[X, Y] = XY - YX$$

We wanted a measure of non-commutativity. Well, here is the simplest possible one! Just measure how different a product is from its commuted counterpart. Now here is an interesting property:

Proposition 5. $T_{\mathbf{1}}(G)$ *is closed under the commutator.*

Proof. Here we use a novel idea: that of a path *of tangent vectors.* If X,Y are in the tangent space then as usual $\exists$ paths A(t) and B(t) in G such that $A(0) = B(0) = \mathbf{1}$ and $A'(0) = X$ and $B'(0) = Y$. We know fix some $s \in \mathbb{R}$ and define $C_s(t) = A(s)B(t)A^{-1}(s)$. Now at $t = 0$:

$$C_s(0) = A(s)B(0)A^{-1}(s) = A(s)A^{-1}(s) = \mathbf{1}$$

Moreover, differentiating and evaluating at t=0:

$$C_s'(0) = A(s)B'(0)A^{-1}(s) = A(s)YA^{-1}(s)$$

Thus, we conclude that for arbitrary s, $A(s)YA^{-1}(s) \in T_{\mathbf{1}}(G)$.

Now if we set $D(s) = A(s)YA^{-1}A(s)$ and let s vary, we have a path of tangent vectors (as promised). Since $A(s), A^{-1}(s)$ are smooth functions of s, $D(s)$ is one too. Now observe that $D(0) = Y \neq \mathbf{1}$, so $D'(0)$ does not seem to fit the definition of a tangent at the identity. However, we can use the following fact: *the tangent space of a Matrix Lie Groups* $T_{\mathbf{1}}(G)$ *is closed under limits.* Since, then $D'(0)$ is a limit of elements of G, it will lie in G. We calculate[8]:

$$D'(0) = A'(0)YA^{-1}(0) + A(0)Y[-A^{-1}(0)A'(0)A^{-1}(0]) = XY - YX = [X, Y]$$

since $A(0) = \mathbf{1}$ and $A(0) = X$. So the image of the commutator restricted to the tangent space lies in the tangent space. ☐

Thus, we can define the following object:

Definition 9. *The Matrix Lie Algebra* $\mathfrak{g}$ *of a Matrix Lie Group G is* $T_{\mathbf{1}}(G)$ *equipped with the commutator* $[\cdot,\cdot]: T_{\mathbf{1}}(G) \times T_{\mathbf{1}}(G) \to T_{\mathbf{1}}(G)$, $[X, Y] = XY - YX$.

This is well defined precisely because the commutator maps $T_{\mathbf{1}}(G) \times T_{\mathbf{1}}(G)$ to $T_{\mathbf{1}}(G)$. Now, in Lie Theory there is a more general object called *the Lie Bracket*:

Definition 10. *Given a vector space V over F, a Lie Bracket on V is an operator* $[\cdot,\cdot]: V \times V \to V$ *that, for X, Y, Z* $\in V$ *and* $\lambda \in F$*, satisfies:*

1. $[X, Y] + [Y, Z] = 0$
2. $[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0$
3. $[\lambda X, Y] = [X, \lambda Y] = \lambda[X, Y]$
4. $[X, Y] + [Z, Y] = [X + Z, Y]$ *and* $[X, Y] + [X, Z] = [X, Y + Z]$

Remark 7. The last two properties are collectively often referred to as *multi-linearity.*

Example 8. A surprisingly familiar example is the cross product on $\mathbb{R}^3$. The first property is the statement that the cross product is skew-symmetric, a familiar fact from calculus. Explicitly:

$$\vec{x} \times \vec{y} = -\vec{y} \times \vec{x} \iff \vec{x} \times \vec{y} + \vec{y} \times \vec{x} = 0$$

The second statement is a computation involving vector identities. For the third statement, recall that the cross product respects scalar multiplication:

$$(c\vec{x}) \times \vec{y} = c(\vec{x} \times \vec{y})$$

The fourth statement is the fact that the cross product distributes over vector addition:

$$\vec{x} \times (\vec{y} + \vec{z}) = \vec{x} \times \vec{y} + \vec{x} \times \vec{z}$$

[8] Note that if $A: \mathbb{R} \to M_n(\mathbb{R})$ smooth then $[A^{-1}(t)]' = -A^{-1}(t)A'(t)A^{-1}(t)$. To see this, write $\mathbf{1} = A(t)A^{-1}$ and differentiate both sides, using the matrix product rule for the RHS.

Proposition 6. *Given a Matrix Lie Group G, the commutator on $T_{\mathbf{1}}(G)$ is a Lie Bracket. Recall that $T_{\mathbf{1}}(G)$ is, in fact, a vector space by the above proposition.*

Proof. This is a formal computation. In the interest of space, we omit it. □

The general notion of a Lie Bracket will not be of any more use to us. We mentioned it for the sake of completeness and in case the interested reader wishes to expand his study of Lie Theory elsewhere. We will stick with the commutator, the "canonical" example of a Lie Bracket for a Matrix Lie Group.

A natural question to ask is how dependent is a Lie Algebra on its corresponding Lie Group. Not as much as we would like:

Proposition 7. *There exist distinct Lie Groups with the same Lie Algebra.*

Proof. It suffices to present a counterexample.

Consider SO(n) and O(n). First, observe:

$$SO(n) \subset O(n) \implies T_{\mathbf{1}}(SO(n)) \subset T_{\mathbf{1}}(O(n))$$

That is so as any path $X(t) \in SO(n)$ is a path $X(t) \in O(n)$. On the other hand, $SO(n)$ is the path component of $O(n)$ that contains $\mathbf{1}$. Accepting for, a moment, that $SO(n)$ is path connected (proved in section 5) we can see this as follows: If $X \notin SO(n)$ but $X \in O(n)$ then necessarily $\det X = -1$. Let $X(t)$ be any path in $O(n)$ such that $X(0) = \mathbf{1}$ and $X(1) = X$. The map $\det X(t)$ is a continuous map into the discrete space $\{-1, +1\}$ so it must be constant. But

$$\det X(0) = \det \mathbf{1} = +1$$

and

$$\det X(1) = \det X = -1$$

a contradiction. Hence, $SO(n)$ is the maximally path connected subset of $O(n)$ that contains $\mathbf{1}$. Therefore, any path in $O(n)$ that intersects the identity necessarily lies in $SO(n)$. Hence:

$$T_{\mathbf{1}}(O(n)) \subset T_{\mathbf{1}}(SO(n)$$

. Thus:

$$T_{\mathbf{1}}(O(n)) = T_{\mathbf{1}}(SO(n)$$

□

So there are things the Lie Algebra does not capture, namely *topological information.* If we want the group to be determined by the algebra, tameness assumptions must be made on its topology. To be exact, we must assume that the Matrix Lie Groups under consideration are simply connected. But this is a discussion for later.

4 From Group to Algebra and Back

We have briefly seen that the exp map maps elements of the Lie Algebra[9] to the Lie Group. It is time we made this precise.

To do so, we are going to need another tool, namely the log function:

Proposition 8. *If $A \in M_n(\mathbb{R})$ is such that $|A| < 1$ then the power series:*

$$log(1 + A) := A - \frac{A^2}{2} + \frac{A^3}{3} - \frac{A^4}{4} \dots$$

converges absolutely. This defines the map $\log : B_1(\mathbf{1}) \to M_n(\mathbb{R})$.

[9]For the rest of the discussion *Matrix* is going to be dropped from *Matrix Lie Algebra* and *Matrix Lie Group*. Bare in mind, we always talk about the latter two objects. Moreover, *algebra* and *group* might be used when it is clear we are referring to a *Matrix Lie Algebra* and a *Matrix Lie Group*, respectively.

Proof. This follows just as in the matrix exponential case: we have absolute convergence because under the substitution $A \to |A|$ we get a series of real numbers that, by analysis, equals $log(1+|A|)$. But this converges whenever $|A| < 1$, which we have assumed. $\square$

The main use of the logarithm is that it is the inverse map of the matrix exponential.

Proposition 9. *If* $|e^X - \mathbf{1}| < 1$ *then* $\log(e^X) = X$.

Proof. Since $|e^X - \mathbf{1}| < 1$ then $log(e^X) = log(\mathbf{1} + (e^X - \mathbf{1}))$ converges. Now, we can use the definition of exp to write:

$$\log(e^X) = \log(\mathbf{1} + \sum_{n=1} \frac{X^n}{n!})$$

Then use the definition of the logarithm to write:

$$\log(e^X) = \sum_n \frac{X^n}{n!} - \frac{1}{2}\left(\sum_n \frac{X^n}{n!}\right)^2 + \frac{1}{3}\left(\sum_n \frac{X^n}{n!}\right)^3 + \dots$$

Now different powers of X commute with each other. Furthermore, since the series converge absolutely, we can justify any reordering of their terms. So we can treat the series as algebraic sums in a commutative field.

Now, if X was a real variable, the calculation would be exactly the same–precisely because X^m commutes with X^k in both cases, for arbitrary $m, k \in \mathbb{Z}_{>0}$, and we can reorder terms. So, purely formally, the calculation is the same. But in the case of the real variable, the statement is true! So, we have the desired result. $\square$

Of course, we also want:

Proposition 10. *If* $|\mathbf{1} - X| < 1$ *then* $e^{\log(X)} = X$.

Proof. As before, the proof is a purely formal manipulation of power series of commuting elements. We expand first the logarithm in terms of its defining power series and then the exponential. All terms are powers of X with real coefficients, thus commuting. A comparison to the case $e^{log(x)}, x \in \mathbb{R}$, where the result is known, concludes the proof. $\square$

Finally, one more useful result:

Proposition 11. *For* $A, B \in M_n(\mathbb{R})$, *if* $AB = BA$ *then:*

$$\log(A + B) = \log(A) + \log(B)$$

Proof. Let $X = \log(A)$ and $Y = \log(B)$. Now by definition of the logarithm:

$$X = \sum_{n=1} (-1)^{n+1} \frac{A^n}{n}$$

and

$$Y = \sum_{n=1} (-1)^{n+1} \frac{B^n}{n}$$

Since A commuted with B, all powers of A commute with all powers of B. So, in fact, $XY = YX$ and we can therefore use the corresponding property of the exponential:

$$\exp(X + Y) = \exp(X)\exp(Y)$$

Taking the logarithm of both sides and recalling that it is an inverse, we have:

$$\log(\exp(X + Y)) = X + Y = \log[\exp(X)\exp(Y)]$$

Plugging in the definition of X, Y:

$$\log(A) + \log(B) = \log[\exp(\log(A))\exp(\log(B))] = \log(AB)$$

$\square$

Now we are ready to prove the following:

Theorem 12. *If G is a Lie Group and X is in $T_{\mathbf{1}}(G)$ then e^X is in G. In other words, exp($T_{\mathbf{1}}(G)$) $\subset G$*

Proof. Let $X \in T_{\mathbf{1}}(G)$. Then, there exist a path A in G such that $A'(0) = X$ and $A(0) = \mathbf{1}$. The idea is to write X as a sum of $\log(A(h))$, where h is small and exponentiate. Then, we can use the sum to product property of the exponential and the fact that the logarithm is the inverse map to the exponential to express e^X in terms of $A(h)$ which by assumption lies in G. By definition:

$$A'(0) = \lim_{h \to 0} \frac{A(h) - \mathbf{1}}{h}$$

which can also be written as:

$$A'(0) = \lim_{n \to \infty} \frac{A(1/n) - \mathbf{1}}{1/n}$$

Expanding $\log(A(1/n))$ in a power series it can be shown that:

$$A'(0) = \lim_{n \to \infty} n \log A(1/n)$$

So:

$$\begin{aligned} e^{A'(0)} &= e^{\lim_{n \to \infty} n \log A(1/n)} \\ &= \lim_{n \to \infty} e^{n \log A(1/n)} \\ &= \lim_{n \to \infty} \left(e^{\log A(1/n)}\right)^n \\ &= \lim_{n \to \infty} A(1/n)^n \end{aligned}$$

The second equality follows from the continuity of the exponential, the third equality follows from proposition 3 and the fourth uses the fact that log is the inverse map of exp.

Now $A(1/n) \in G$ by assumption, thus by closure of G under multiplication, $A(1/n)^n \in G$ as well. Finally, this is a sequence of points in G with a non-singular limit – $e^{A'(0)}$ has the inverse $e^{-A'(0)}$ – so by the definition of a Matrix Lie Group:

$$e^{A'(0)} = e^X \in G$$

Since X was arbitrary, we conclude that:

$$\exp(T_{\mathbf{1}}(G)) \subset G$$

□

Finally, we have a way to produce elements of the Lie Group, given its Lie Algebra. One might say we are a third of the way there. Where? To giving a full description of the group from its algebra. We know how to get from the algebra to the group, namely via the exponential. We would like to know how to get from the algebra to the group as well as how to relate the operation of the algebra, namely the Lie Bracket, to the operation of the group, namely matrix multiplication. Miraculously, we can answer both questions. These are the two theorems we now explore.

Theorem 13. *If G is a Matrix Lie Group then there exists some neighborhood of the identity such that for any $X \in B_\delta(\mathbf{1}) = \{A \in M_n(\mathbb{R}) | A - \mathbf{1}| < \delta\}$ we have that $\log(X) \in T_{\mathbf{1}}(G)$. In other words, log($B_\delta(\mathbf{1})$) $\subset G$.*

To prove this theorem, we will need the following fact:

Proposition 14. *Let G be a matrix Lie Group. If $\{A_n\}$ is a sequence in G such that $\lim_{n \to \infty} A_n = \mathbf{1}$ and $\{\alpha_n\}$ is a sequence of real numbers such that $\lim_{n \to \infty}((A_n - \mathbf{1})/\alpha_n) = X$, for some $X \in M_n(\mathbb{R})$, then X is in $T_{\mathbf{1}}(G)$.*

Remark 9. This proposition reunites the geometric picture of tangent vectors in the low dimensional calculus of vectors with our more abstract picture of tangent vectors to a matrix group. In fact, the limit of X of the sequence of fractions is often called the *sequential tangent vector* in recognition of that fact. The sequence $A_n - \mathbf{1}$ can be interpreted as a sequence of lines that intersect the points $\mathbf{1}$ and A_n. The limit $A_n \to \mathbf{1}$ resembles the usual notion of a tangent to a curve. The ratio $\frac{A_n - \mathbf{1}}{\alpha_n}$ then computes the "slope" of the line, hence its limit $n \to \infty$ computes the slope of the tangent. Recall that an element of $T_{\mathbf{1}}(G)$ was defined as the derivative at $\mathbf{1}$ of a path in G.

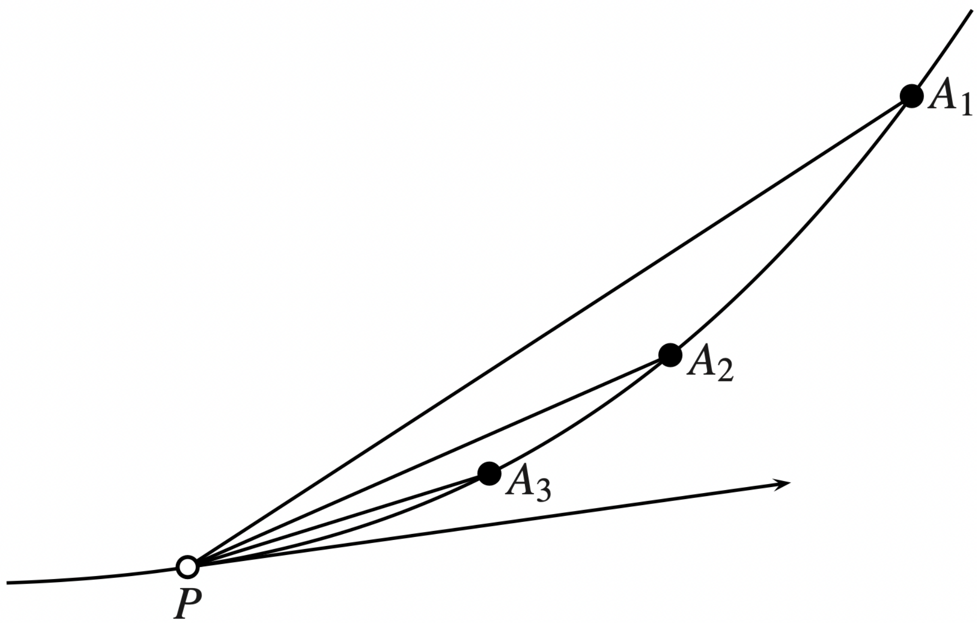

Figure 1: The tangent at P as a limit of lines that intersect P and points $A_m \to P$. Credit: J. Stillwell, *Naive Lie Theory*.

Proof. **(Of Proposition 13)** This is a proof by contradiction. If $\nexists$ such a neighborhood then $\exists$ a sequence $\{A_n\}$ in G such that $A_n \to \mathbf{1}$ that satisfies: $\forall n$, $\log A_n \notin G$. We can then decompose any

$$A_n = X_n + Y_n$$

where $X_n \in T_{\mathbf{1}}(G)$ and $Y_n \in T_{\mathbf{1}}(G)^{\perp}$. The idea of the proof is to construct a sequence with limit that both isn't and is in $T_{\mathbf{1}}(G)$. For the first part, picking a sequence in a closed subset of the complement of $T_{\mathbf{1}}(G)$ seems fitting – in fact, we will use a compact subset in $T_{\mathbf{1}}(G)^{\perp}$. For the second part, we will use the notion of sequential tangent vectors – notice that the definition of a sequential tangent vector works for *any* sequence in G with limit $\mathbf{1}$.

Observe that $X_m, Y_m \to \mathbf{0}$ as $A_m \to \mathbf{1}$ and log is a continuous map. Consider, now, the sequence:

$$\frac{Y_m}{|Y_m|} \in T_{\mathbf{1}}(G)^{\perp}$$

These elements have all norm 1, so they live in the unit sphere $\mathbb{S}^{n^2}$. Thus:

$$\frac{Y_m}{|Y_m|} \in \mathbb{S}^{n^2} \cap T_{\mathbf{1}}(G)^{\perp}$$

The latter is a compact set, since $\mathbb{S}^{n^2}$ is compact and $T_{\mathbf{1}}(G)^{\perp}$ is closed[10]. Therefore, there exists a subsequence of $Y_m/|Y_m|$ that converges to some $Y \in \mathbb{S}^{n^2} \cap T_{\mathbf{1}}(G)^{\perp}$ so, in particular, $Y \in T_{\mathbf{1}}(G)^{\perp}$ which implies

[10] Fact: Given a vector space V, the orthogonal orthogonal complement of any subspace of V is closed with respect to the metric topology.

$Y \notin T_{\mathbf{1}}(G)$. Now define another sequence:

$$T_m = e^{-X_m} A_m$$

By our previous theorem $e^{-X_m} \in G$ since $X_m \in T_{\mathbf{1}}(G)$ so $T_m \in G$. Using that log is inverse to exp write:

$$A_m = e^{X_m + Y_m}$$

A computation involving the power series of log and exp shows (cf. Stillwell p. 148):

$$T_m = \mathbf{1} + Y_m$$

So finally:

$$\lim_{m \to \infty} \frac{T_m - \mathbf{1}}{|Y_m|} = \lim_{m \to \infty} \frac{Y_m}{|Y_m|} = Y$$

with:

$$Y \notin T_{\mathbf{1}}(G)$$

But $Y_m \to \mathbf{0}$ implies $T_m \to \mathbf{1}$. Moreover, we showed $T_m \in G$. Lastly, if $Y_m \to 0$ then surely $|Y_m| \to \mathbf{0}$. So the limit:

$$\lim_{m \to \infty} \frac{T_m - \mathbf{1}}{|Y_m|} = Y$$

is a sequential tangent vector of G at $\mathbf{1}$. It therefore must lie in $T_{\mathbf{1}}(G)$ by Proposition 14. Hence, we have a contradiction. □

Immediately, we have:

Corollary 15. *If G is a Matrix Lie Group then there exists some neighborhood of the identity that is homeomorphic to $T_{\mathbf{1}}(G)$.*

Proof. By the above two theorems, $\exp : T_{\mathbf{1}}(G) \to G$ and $\log : G \to T_{\mathbf{1}}(G)$ are well defined continuous map. Moreover, by earlier propositions, they are inverse to one one another. Thus, log is such a homeomorphism. □

This is either nice or really bad, depending on your perspective[11]. We are now at a position to go back and forth between the Lie Algebra and part of the Lie Group, namely some neighborhood of the identity. However, this precursors, in a sense, our failure of describing the group fully from its algebra, without imposing tameness conditions on its topology. Indeed, the theorem is telling us that it is only *local* topological information of the group that are captured by the algebra.

Despite the above issue, our work is not done. To begin with, we have not related the operation of the algebra to that of the group. Moreover, we can improve the above theorem by asking for a sufficiently nice topology. The first consideration we treat here. The second, in the next section.

As discussed, the (important) operation of the Lie Algebra is the Lie Bracket. The operation of the Lie Group is matrix multiplication. Here is how we relate them:

Theorem 16. *(Campbell-Baker-Hausdorff) If G is a Matrix Lie Group and $\mathfrak{g}$ is its algebra, then for $X, Y \in \mathfrak{g}$:*

$$e^X e^Y = e^{X + Y + \frac{1}{2}[X,Y] + F(X,Y)}$$

where F(X,Y) is a linear combination (with real coefficients) of X, Y and possibly nested commutators [X,Y].

Proof. Stillwell gives us *Eichler's Proof* which is, remarkably, pure algebra. He writes $F(A, B) = \sum_i F_i(A, B)$ where each F_i is a homogeneous polynomial of degree n in the variables A and B, after computing all commutators of A and B. He calls such an F_i a *Lie polynomial of degree i*. He then works inductively. The case $n = 0$ is a calculation with power series, arrived at by writing $\log(e^X e^Y)$ and collecting terms to first order. After assuming the result for $m < n$, that is, assuming that F_m is a Lie polynomial for $m < n$, the hard work is to prove that F_n is Lie.

[11] The half-empty half-full glass question, yet again.

The first step is noting associativity: matrix products $ABC\ldots$ are associative so we expect $e^Ae^Be^C$ to be associative, by an expansion in power series. Thus, one has:

$$(e^Ae^B)e^C = e^A(e^Be^C)$$

Then, let W be such that $e^Ae^Be^C = W$ and using the expression above, write W in terms of the F_i polynomials in two different ways:

$$\sum_{i=1}^{\infty}(F_i(\sum_{j=1}^{\infty}F_j(A,B),C) = \sum_{i=1}^{\infty}(F_i(A,\sum_{j=1}^{\infty}F_j(B,C))$$

Then, the inductive hypothesis says that F_m is a Lie polynomial for $m < n$. We wish to prove that F_n is Lie. Recall that this is a homogeneous polynomial in A,B so its terms are going to come from products like F_lF_k where $k + l = n$. But by the inductive hypothesis, if $l < n$ and $k < n$ we are done, as we know that both F_l and F_k are Lie hence their product is going to be Lie. The only mystery terms come from products of the form F_1F_n. To get such products, we need to necessarily take $i = 1, j = n$ or $i = n, j = 1$ in the above equations. The LHS yields:

$$F_n(A,B) + C + F_n(A+B,C)$$

whereas the RHS yields:

$$F_n(B,C) + A + F_n(A,B+C)$$

So we have the reduced the problem to 4 polynomials. That is progress! The rest of the argument involves a sequence of convoluted algebraic manipulations that eventually establish the result. Due to length constraints, we omit them. The are given in Stillwell, p.154-157. □

We have shown that the Lie Bracket "determines" the operation in the group. Given $\mathfrak{g}$ we can produce elements of G and compute their product without knowing *anything whatsoever* about the product operation in G. Just take any two $X, Y \in \mathfrak{g}$, calculate their commutator and use the prescription of Campbell-Baker-Hausdorff (commonly abbreviated CBH) to find what the product of their corresponding G elements is, namely e^Xe^Y. This is a milestone!

What about arbitrary elements of G though? Given $A, B \in G$, we would like to find their counterparts, say $a, b \in \mathfrak{g}$, then compute $[a, b]$ and hope this tells us something about AB. But here is the issue: for given G, there is only *a neighborhood* of the identity from which we know how to get to $\mathfrak{g}$, via the logarithm. The solution is to tame the topology of G.

5 Taming the Topology

Definition 11. *If G is a Matrix Lie Group and $X, Y \in G$ then A is path connected to B if there exists a continuous path $\gamma : [0,1] \to G$ in G that satisfies: $\gamma(0) = A$ and $\gamma(1) = B$. Recall that continuity here is continuity of a map $\mathbb{R} \to \mathbb{R}^{n^2}$.*

And so:

Definition 12. *A Matrix Lie Group G is path connected if for every pair $A, B \in G$, A is path connected to B.*

The rest of the our well-known topological notions of connectedness apply as usual, notably the definition of a *path component.* As promised, here is why SO(n) is path connected:

Example 10. SO(n) is path connected. To show this, one may induct on the dimension n.

The base case deserves some attention. There two approaches, a geometric and a formal one. The latter is less work, so it is presented first. For $n = 1$, any $M \in M_1(\mathbb{R})$ is $M = [x]$ for some $x \in \mathbb{R}$ so in fact we can identify $\mathbb{R}$ with $M_1(\mathbb{R})$. Now, $M^T = M$ always, i.e. all one dimensional real matrices are symmetric. If $M \in O(n)$ then $M^{-1} = M^T$ so $M^{-1} = M$. Thus $M = \pm 1$. Finally, an appeal to the co-factor expansions of the determinant (cf. footnote 3), shows that $\det M = \det[x] = x$, so if we want $M \in SO(1)$ then

$$\det M = +1 \iff x = 1$$

Thus, $SO(1)$ has a single element, $+1$. Therefore, it is path connected.

For the geometric approach (cf. Stillwell) we use $n = 2$ as a base case. We identify $SO(2)$ with the unit circle $\mathbb{S}^1$ in the complex plane.[12] Then, given any two u, v in $\mathbb{S}^1$ we can always find a path in the unit circle that travels from u to v: if

$$u = e^{i\theta_1}$$

and

$$v = e^{i\theta_2}$$

then let $\gamma : [0,1] \to \mathbb{S}^1$ be:

$$\gamma(t) = e^{i[\theta_1 + t(\theta_2 - \theta_1)]}$$

We just go from u to v along the arc that connects them!

Now for the inductive step. Assume $SO(n-1)$ is path connected. We will show that given any $A \in SO(n)$ there exists a path from A to $\mathbf{1}$. Then, we can connect any pair $X, Y \in SO(n)$ by connecting both X and Y to the identity (i.e. taking the concatenation of the paths, with one traversed in reverse). *Disclaimer*: a full proof requires a fair amount of writing and linear algebra. The idea is as follows: Fix $A \in SO(n)$ and some basis $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$. To connect $\mathbf{1}$ to A it suffices to find a path of matrices $R(t)$, where $t \in [0,1]$ in $S)(n)$ such that $R_1\mathbf{e}_i = A\mathbf{e}_i$ for $1 \leqslant i \leqslant n$. Since $R9(t)$ is inner product preserving its action can be thought of as "rigid motion". So, intuitively, we want to move the n vectors $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ by a continuous rigid motion to the n vectors $A\mathbf{e}_1, A\mathbf{e}_2, \ldots, A\mathbf{e}_n$.

If $\mathbf{e}_1$ and $A\mathbf{e}_1$ are distinct, they define a plane $\mathbb{P}$. Up to a change of basis, we can find a matrix path $R(t)$ in $SO(2)$ that takes $\mathbf{e}_1$ to $A\mathbf{e}_1$ when $t = 1$. This is possible because $SO(2)$ is path connected, by the base case. Then, $R(1)$ takes the rest of our basis to vectors $R(1)\mathbf{e}_1, R(1)\mathbf{e}_2, \ldots, R(1)\mathbf{e}_n$. Now we a need a key fact and a key observation.

The key fact, requiring some linear algebra, is that $R(1)$ is an orientation preserving isometry for all $\mathbf{e}_i$, not just for elements in $\mathbb{P}$. This says that $R(1)\mathbf{e}_2, \ldots, R(1)\mathbf{e}_n$ are orthogonal to $R(1)\mathbf{e}_1$ because $\mathbf{e}_2, \ldots, \mathbf{e}_n$ are orthogonal to $\mathbf{e}_1$ and R(t) preserves inner products. The key observation is that $\mathbf{e}_2, \ldots, \mathbf{e}_n$ are orthogonal to $A\mathbf{e}_2, \ldots, A\mathbf{e}_n$ because $\mathbf{e}_2, \ldots, \mathbf{e}_n$ are orthogonal to $\mathbf{e}_1$ and A preserves inner products. So

$$\{R(1)\mathbf{e}_2, \ldots, R(1)\mathbf{e}_n\}, \{A\mathbf{e}_2, \ldots, A\mathbf{e}_n\} \in \operatorname{span}(\mathbf{e}_1)^\perp$$

Note that $\operatorname{span}(\mathbf{e}_1)^\perp$ is an $n-1$ dimensional subspace of $\mathbb{R}^n$. By the inductive hypothesis $SO(n-1)$ is path connected so there exists a path $T(t) \in SO(n-1)$ such that $T(1)R(1)\mathbf{e}_i = A\mathbf{e}_i$[13] Finally, extend $T(t)$ to a map $\tilde{T(t)} : [0,1] \times \mathbb{R}^n \to \mathbb{R}^n$ by defining it to be the identity on $\operatorname{span}(\mathbf{e}_1)$ – concretely, take the $n-1$ by $n-1$ matrix $T(t)$ and embed it into an n by n matrix where the n^{th} column and n^{th} row have zeros everywhere but in their last entry.

Finally, concatenate the two paths: do $R(t)$ first and when you are done, do $\tilde{T(t)}$. $R(t)$ is going to take care of the $\mathbf{e}_1$ and $\tilde{R(t)}$ is going to take care of $\mathbf{e}_2, \ldots, \mathbf{e}_n$ without messing up $A\mathbf{e}_1$. Each map is an orientation preserving isometry, for all t, in the whole of $\mathbb{R}^n$ so their composition is going to be an orientation preserving isometry in $\mathbb{R}^n$, for all t, that is to say, an path in $SO(n)$.

Example 11. $GL_n(\mathbb{C})$ is path connected. Pick $A, B \in GL_n(\mathbb{C})$. Let's look at matrices of the form $zA + (1-z)B$ where $z \in \mathbb{C}$. This defines a plane $\mathbb{P}$ parametrized by z, that includes A and B, obtained at $z = 0$ and $z = 1$ respectively. We will look for a path $A(t)$ in this plane. Our condition is $\det M(t) \neq 0$, so we wish to avoid matrices for which:

$$\det(zA + (1-z)B) = 0$$

Now expanding the determinant (once again) in co-factors, we see that $\det(zA + (1-z)B)$ is a polynomial of degree n in z with complex coefficients. By the fundamental theorem of algebra, it has exactly n roots,

[12] This identification is valid because there exists a homeomorphism $\phi : \mathbb{S}^1 \to SO(2)$ given by

$$e^\theta \mapsto \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix}$$

So given a path $\psi : \mathbb{S}^1 \to \mathbb{S}^1$ we can always construct a path $\tilde{\psi} : SO(2) \to SO(2)$ by taking $\tilde{\psi} = \phi \circ \psi \circ \phi^{-1}$.

[13] Again, some linear algebra is needed. Since both R(1) and A are orientation preserving isometries, $\{A\mathbf{e}_2, \ldots, A\mathbf{e}_n\}$ and $\{R(1)\mathbf{e}_2, \ldots, R(1)\mathbf{e}_n\}$ are both orthonormal bases with the same orientation. It is a linear algebra fact that such bases can be transformed to each other by an element of SO(n).

so there are exactly n points in $\mathbb{P}$ that we have to avoid. But then we are done: it is a topological fact that $\mathbb{R}^n - \{x_1, x_2, ;x_n\}$ is path connected if $n > 1$. No singleton is a cut point of $\mathbb{R}^n$ if $n > 1$. Here:

$$\mathbb{P} \cong \mathbb{R}^2$$

Thus, if $z_1, \ldots, z_n$ are the roots of $\det(zA + (1-z)B) = 0$ then

$$\mathbb{P} - \{z_1, \ldots, z_n\}$$

is a path connected set that contains A and B.

Remark 12. Observe that $GL_n(\mathbb{R})$ is not path connected. This is so as there exist non-path connected subgroups of it, such as O(n). Note that in the proof of the connectedness of $GL_n(\mathbb{C})$ we used the fact that a polynomial in $\mathbb{C}$ has exactly n solutions. Such a statement cannot be made for $\mathbb{R}$.

As promised, let's see the tameness of the topology in action:

Proposition 17. *If G is path connected Matrix Lie Group and $B_\epsilon(\mathbf{1})$ is a neighborhood of the identity, any element of G is a product of elements of $B_\epsilon(\mathbf{1})$.*

Proof. Fix some ball around the identity $B_\epsilon(\mathbf{1})$. Since G is path connected, for any $A \in G$ there exist a path $A(t)$ in G such that $A(0) = \mathbf{1}$ and $A(t) = A$. If S is a subset of $M_n(\mathbb{R})$ we introduce the notation $A(t)S$ for the set:

$$A(t)S = \{A(t)B : B \in S\}$$

Let V be an open set that contains $\mathbf{1}$. Note that for all t, $A(t) \in A(t)V$. Now, matrix multiplication is a continuous map $M_n(\mathbb{R}) \times M_n(\mathbb{R}) \to M_n(\mathbb{R})$ so, in particular, for fixed A, the map $m_A : M_n(\mathbb{R}) \to M_n(\mathbb{R})$ given by $X \mapsto AX$ is a continuous map. Since $m_{\mathbf{1}} = \mathbf{1}$ we further have that, if A non-singular, m_A is a homeomorphism with inverse $m_{A^{-1}}$. That is because $m_{A^{-1}}$ is again continuous and we can compute[14]

$$m_{A^{-1}} \circ m_A = m_A \circ m_{A^{-1}} = m_{AA^{-1}} = m_{\mathbf{1}} = \mathbf{1}$$

Therefore, m_A is an open map, for A non-singular. Above, $A(t) \in G$ for all t, so $A(t)$ is non-singular for all t as $G \subset GL_n(\mathbb{R})$. Thus, $m_{A(t)}$ is open. Now observe:

$$A(t)V = m_{A(t)}(V)$$

Thus $A(t)V$ is open, for all t, as the image of the open set V under the open map $m_{A(t)}$.

With the above, we have that the family:

$$\mathcal{V} = \{A(t)V\}_{t \in [0,1]}$$

is an (ambient) open cover of the image of the path $A(t)$, denoted by ImA. Now since $[0, 1]$ is compact and $A(t)$ is by assumption continuous, we have that ImA is compact. Thus, there exist $t_1, t_2 \ldots t_n$ such that

$$\text{Im}A \subset \bigcup_{i=1}^{n} A(t_i)V$$

We want to find a finite sequence of points $\mathbf{1} = A_1, A_2 \ldots A_m = A$ such that for all $1 \leqslant i \leqslant m$ there exists a $1 \leqslant j \leqslant n$ so that

$$A_i A_{i+1} \in A(t_j)V$$

To do this, appeal to the *Lebesgue Number Lemma*: for any cover $\mathcal{U}$ of a compact metric space (X, d) there exist some $\delta > 0$ such that any ball of radius less than δ is contained in some element of $\mathcal{U}$.[15] ImA is a compact metric space (with induced metric from the ambient space) so with cover $\tilde{\mathcal{V}} = \mathcal{V} \cap \text{Im}A$, let δ as in

[14] A map that is both a homomorphism and a homeomorphism

[15] This is slightly incorrect: the true statement is that for any cover $\mathcal{U}$ of a compact metric space (X, d) there exist some $\delta > 0$ such that any *set of diameter* less than δ is contained in some element of $\mathcal{U}$, where diam$S = \sup\{|x - y| : x, y \in S\}$. But the two statements are equivalent.

the statement of the Lebesgue Number Lemma. Then if we choose the points A_i such that $|A_i - A_{i+1}| < \delta$, the ball $B_\delta(A_i)$ contains both A_i and A_{i+1}. By the Lebesgue Number Lemma there exist j such that

$$B_\delta(A_i) \subset \tilde{V}_j \in \tilde{\mathcal{V}}$$

that is

$$B_\delta(A_i) \subset A(t_j)V$$

by recalling the definition of $\tilde{\mathbb{V}}$. So finally:

$$A_i,\, A_{i+1} \in A(t_j)V$$

With this sequence in hand, write:

$$A = A_1 A_1^{-1} A_2 A_2^{-1} \dots A_{m-1}^{-1} A_m$$

recalling that $A_m = A$. By associativity of matrix multiplication:

$$A = (A_1) \cdot (A_1^{-1} A_2) \cdot (A_2^{-1} A_3) \cdot \dots \cdot A_m$$

If we can show that each term in parenthesis in is $B_\epsilon(\mathbf{1})$ we are done. To do this we use the following fact, that will be proven as a lemma below: if $B_r(\mathbf{1})$ some ball around the identity then there exists an open set U with $\mathbf{1} \in U$ such that for any $A, B \in V$, $AB^{-1} \in B_r(\mathbf{1})$. In a nutshell, this is due to the continuity of both matrix multiplication and matrix inverse taking.We could have chosen our V to satisfy this property, with $r = \epsilon$. Finally, for any A_i there exists a $B_i \in V$ and a t_j so that

$$A_i = A(t_j) B_i$$

Therefore:

$$\begin{aligned} A_i^{-1} A_{i+1} &= (A(t_j)B_i)^{-1} A(t_j) B_{i+1} \\ &= B_{i+1}^{-1} A(t_j)^{-1} A(t_j) B_i \\ &= B_{i+1}^{-1} B_i \end{aligned}$$

But B_{i+1}, $B_i \in V$ so by the choice of V we have $B_{i+1}^{-1} B_i \in B_\epsilon(\mathbf{1})$ and we are done. □

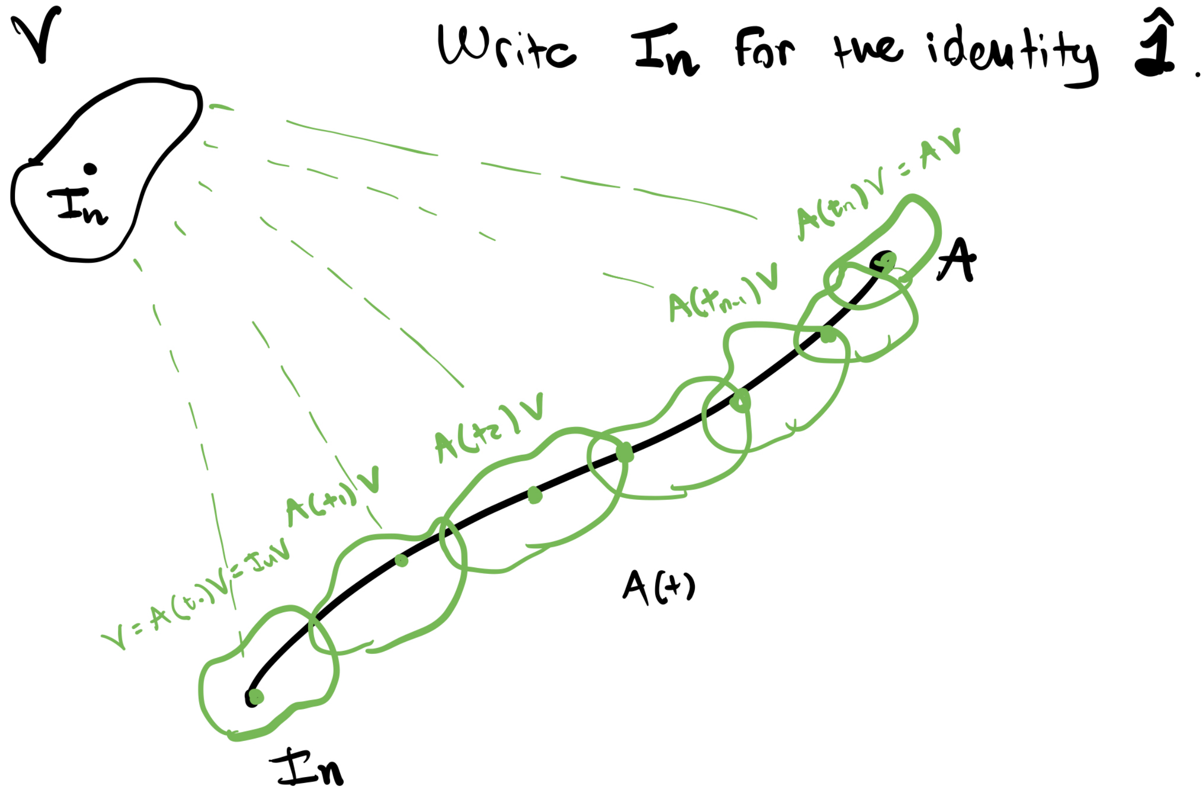

Figure 2: A finite covering of the path $A(t)$ by n "translations" of the set V by $A(t_i)$, $1 \leqslant n \leqslant 1$.

Remark 13. Note the importance of being able to write A_i and A_{i+1} in terms of the *same* $A(t_j)$, which we got from the Lebesgue Number Lemma: this is what allowed cancellation of the $A(t_j)$ in the product.

Lemma 18. *If $B_r(\mathbf{1})$ some ball around the identity then there exists an open set U with $\mathbf{1} \in U$ such that for any $A, B \in V$, $AB^{-1} \in B_r(\mathbf{1})$*

Proof. Let $m : GL_n(\mathbb{R}) \times GL_n(\mathbb{R}) \to GL_n(\mathbb{R})$ be the multiplication map $(A, B) \mapsto AB$ and $i : GL_n(\mathbb{R}) \to GL_n(\mathbb{R})$ be the inverse taking map $A \mapsto A^{-1}$. We seek open V with $\mathbf{1} \in V$ such that

$$m(V, V) \subset B_r(\mathbf{1})$$

and

$$i(V) \subset B_r(\mathbf{1})$$

Since m is continuous $W_1 \times W_2 = m^{-1}(B_r(\mathbf{1}))$ is open. Moreover, $\mathbf{1} \in m^{-1}(B_r(\mathbf{1}))$ clearly. So by definition of the product topology, there exist open sets U_1, U_2 such that $(\mathbf{1}, \mathbf{1}) \in U_1 \times U_2 \subset m^{-1}(B_r(\mathbf{1}))$. Now since i is continuous, and $\mathbf{1} \in i^{-1}(B_r(\mathbf{1})$ there exists open set $\mathbf{1} \in U_3$ such that $U_3 \subset i^{-1}(B_r(\mathbf{1}))$. So let

$$V = U_1 \cap U_2 \cap U_3$$

V is open as it is an intersection of finitely open sets, non empty because all U_i contain $\mathbf{1}$. So finally we have:

$$m(V, V) \subset m(U_1, U_2) \subset B_r(\mathbf{1})$$

and

$$m(V) \subset m(U_3) \subset B_r(\mathbf{1})$$

proving the claim. ∎

As such:

Theorem 19. *If G is a path connected Matrix Lie Group then for $A \in G$ and $X_1, X_2, \dots X_n \in \mathfrak{g}$ we have:*

$$A = e^{X_1} e^{X_2} \dots e^{X_n}$$

Proof. Pick $A \in G$. Let $\delta > 0$ by the number that ensures $\log(B_\epsilon(\mathbf{1})) \in T_{\mathbf{1}}(G)$. By the above theorem, there exist $A_1 \dots A_n \in B_\epsilon(\mathbf{1})$ so that

$$A = A_1 A_2 \dots A_n$$

But now for all i

$$log(A_i) = X_i \in T_{\mathbf{1}}(G)$$

So

$$A_i = e^{X_i}$$

which allows us to write

$$A = e^{X_1} e^{X_2} \dots e^{X_n}$$

□

Finally, we have what we wanted. The third step! The first was relating every element of the algebra to an element of the group, achieved with the exp map. The second was to express the group operation in terms of the (relevant) operation in the algebra, namely the commutator. That was the Campbell-Baker-Hausdorff Theorem. The third step is that we can relate any element of the group to a collection of elements in the algebra. Crucially, the last step was achieved by assuming the the topology is tame enough, namely that the group is connected. Now if we go *just* a bit further, there is much more we can say. Let G and H be *simply connected* Matrix Lie Groups. Then their algebras $\mathfrak{g}$ and $\mathfrak{h}$ are *isomorphic* if and only if G and H are *isomorphic*. This discussion can be done with almost exclusively what is on in this paper. However, due too length constraints we simply state the two relevant theorems:

Theorem 20. *For any Lie group homomoprhism $\Phi : G \to H$ of matrix Lie groups with corresponding Lie algebras $\mathfrak{g}$ and $\mathfrak{h}$, there is an induced Lie homomoprhism $\phi : \mathfrak{g} \to \mathfrak{h}$ such that:*

$$\phi(A'(0)) = (\Phi \circ A)'(0)$$

Proof. Page 191. □

Theorem 21. *If $\mathfrak{g}$ and $\mathfrak{h}$ are Lie algebras of the simply connected Lie groups G and H then any Lie algebra homomorphism $\phi : \mathfrak{g} \to \mathfrak{h}$ is induced by a Lie group homomorphism $\Phi : G \to H$.*

Proof. Page 198. □