Highlights of elementary Lie theory*

Abhimanyu Pallavi Sudhir

October 2019

Contents

1 Lie groups and algebras; exponential map

The motivation for the theory of Lie groups comes from a fundamental topological insight: some vague notion of "wiggle room" often has a lot of the properties of the notion of cardinality – the classic example of this in topology is *compact spaces*, which have a finite "wiggle room", generalising how finite sets have a finite cardinality. Even though the closed interval $[0, 1]$ has the same cardinality as $\mathbb{R}$, it is often "visualised" in a similar way to a set like $\{0, 1\}$.

Well, the same insight should be available in group theory – for instance, the circle group is analogous to a cyclic group, $(\mathbb{R}, +)$ is analogous to $(\mathbb{Z}, +)$, etc. The key notion we want to make sense of this is a *real-index power* (that is at least sometimes defined) on Lie groups generalising the standard integer-index power that is available to us for free.

---

*This is a quick walkthrough of some important ideas in Lie theory. For a fuller treatment (in-progress), see: `https://thewindingnumber.blogspot.com/p/1203.html` (for Lie theory) and `https://thewindingnumber.blogspot.com/p/2204.html` (for Topology).

A definition of real powers requires some notion of a parameterisation, which exists as input to an exponential function. The interesting fact about Lie groups is that they come with a very natural source of this parameterisation: the tangent space at the identity.
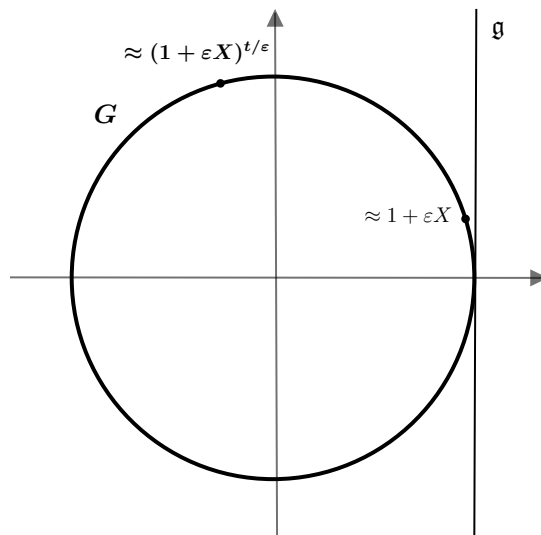

Figure 1: Lie algebras and the exponential map

Fig. 1 shows the intuition for why this makes sense – a group element infinitesimally close to the identity, $1 + \varepsilon X$ (where X is some basis element in the tangent space) exponentiates by a real power t/ε to the group element $(1+\varepsilon X)^{t/\varepsilon}$. With infinitesimal ε, this can be considered the definition of $\exp tX$ – and it's easy to check that this is equivalent to the "power series" definition.

(It's reasonable to ask how we can perform such "addition" between the group identity and vector space elements. Well, right now, we're considering the group and its Lie algebra to both be embedded in some $GL(\mathbb{R}^n)$ – we can then figure out what the properties are that an "abstract" Lie group or algebra ought to satisfy.)

2   Lie bracket; Lie homomorphisms

Much of the motivation behind Lie theory has to do with translating between ideas in the group and the tangent space – the ultimate hope is that we may be able to equip the tangent space with enough structure that it could characterise the Lie group (or at least its local behaviour) completely. Certainly for an Abelian Lie group just the linear space structure suffices, as exp is a local homeomorphism in this case. So it's sensible to expect that the remaining

structure in the general case has something to do with the non-commutativity of the group.

There are two equivalent ways to see the Lie bracket – the first makes the connection to non-commutativity more clear, but the second perhaps better illustrates the importance of the operation in connection to the conjugation map.

- $[X, Y]$ is the second-derivative at the identity of the commutator curve $e^{tX}e^{tY}e^{-tX}e^{-tY}$. Because the first-derivative is zero, replacing t with $\sqrt{t}$ shows that it is a member of the Lie Algebra.

- The differential of $\mathrm{Ad} : G \to \mathrm{Aut}(G) := g \mapsto \lambda h. \; ghg^{-1}$ (the "Adjoint map" homomorphism) is $\mathrm{ad} : TG \to T\mathrm{Aut}(G) := X \mapsto [X, \cdot]$. One immediate consequence of this understanding is that ad is a Lie Algebra homomorphism, and thus preserves the Lie bracket: $\mathrm{ad}([X, Y]) = [\mathrm{ad}(X), \mathrm{ad}(Y)]$ (this is a form of the *Jacobi identity*).

Note that we haven't yet abstractly defined a Lie algebra or a Lie algebra homomorphism. When writing down *Lie's fundamental theorems*, it will be convenient to define a Lie algebra as simply a vector space equipped with a Lie bracket (and thus a homomorphism of Lie algebras preserves this Lie bracket), and the Lie bracket as a bilinear operator that is antisymmetric and satisfies the Jacobi identity – but Lie's theorems are what serve as the *motivation* for this definition.

Another example of a Lie algebra homomorphism induced by a Lie group homomorphism is the automorphism of conjugation on the Lie algebra, $\mathrm{Ad}(g)(X) = gXg^{-1}$, defined as the derivative of a conjugation automorphism $Ad(g)(h) = ghg^{-1}$. This map defines the "adjoint representation" of the Lie group, although only faithful iff the group has trivial center. If we want to understand $\mathrm{Ad}(g)$ as a "rotation" transformation, we would need to introduce the *Killing form* (in fact, for simple Lie algebras, this is the *unique* Ad-invariant bilinear form up to scaling) – we can then see, e.g. that $[X, Y]$ (the tangent to the orbit of Y under conjugation by a curve with derivative X at 0) is normal to X.

We can write down other examples of correspondences between Lie groups and their Lie algebras:

- A *subalgebra* of a Lie algebra is the tangent Lie algebra to a a Lie subgroup – one can confirm that it is a linear subspace closed under the Lie bracket.

- An *ideal* of a Lie algebra is the tangent Lie algebra to a normal subgroup – one can confirm that $\mathfrak{i}$ is an ideal of $\mathfrak{g}$ iff $[\mathfrak{g}, \mathfrak{i}] \subseteq \mathfrak{i}$, or equivalently iff it is the kernel of a Lie algebra homomorphism.

- The *centre* of a Lie algebra is the tangent space to the centre of the group – equivalently, the largest subalgebra (in fact ideal) $\mathfrak{z}$ such that $[\mathfrak{g}, \mathfrak{z}] = 0$.

It is then reasonable to ask if these implications go both ways – i.e. is every Lie algebra homomorphism the differential of a Lie group homomorphism; is every Lie subalgebra the tangent to a subgroup? This questions will be the subject of Lie's fundamental theorems.

3 Exceptional isomorphisms

Linear algebra is really simple on the whole isomorphism question – two linear spaces (over the same field) are isomorophic iff they have the same dimension. Things are much harder in topology and group theory.

Well, at least with the classical Lie groups, they are easy to categorise, so if we can define some invariants, we can perhaps decide what groups definitely aren't isomorphic and then worry about the remaining ones. Here are three invariants: dimension, center, rank (dimension of maximal torus).

The classical Lie groups we are interested in are the connected ones – so not $O(n)$ – we can start with $SO(n)$, $SU(n)$, $U(n)$, $Sp(n)$. The first thing we can do is look at how rank and dimension vary in each of these series: the rank of $U(n)$ is always n while its dimension is n^2. In $SU(n)$, the subgroup of diagonal matrices loses a degree of freedom, thus its rank is always $n-1$ while its dimension is $n^2 - 1$. The dimension of $SO(n)$ is $n(n-1)/2$ while its rank – well, you need two real dimensions to make a torus, so its rank is $\lfloor n/2 \rfloor$. The Symplectic group is a bit harder to visualise, but one can consider the diagonal matrices with $(1, i)$ entries to be a maximal torus so that the rank of $Sp(n)$ is n – and of course, its dimension can easily be seen to be $n(2n + 1)$.

Classifying by rank, this means for rank at least 4,

$$\dim U < \dim SU < \dim SO(\text{even}) < \dim SO(\text{odd}) = \dim Sp$$

Well, for odd n, $SO(n)$ has trivial center while $Sp(n)$ has center $\{I, -I\}$. We can then look at the finite number of groups in ranks 1, 2 and 3 and confirm that the only cases where all three invariants are the same are:

- $SO(2)$ and $U(1)$

- $SU(2)$ and $Sp(1)$

In the first case, the isomorphism is simply the canonical real form of complex matrices – in the second, the isomorphism is simply the canonical complex form of quaternion matrices. So these cases indeed are isomorphisms. So these are the only isomorphisms we can write of the stated groups.

One could introduce some new Lie groups by quotienting over normal subgroups – well, in particular it happens that all the quotients over the centres (which are equivalently the inner automorphism groups) of the mentioned classical Lie groups: $Inn(SO(n))$, $Inn(SU(n))$, $Inn(U(n))$, $Inn(Sp(n))$ are simple.

4 Lie's theorems: aka "Why stop there?"

There are two decisions in the definition of a Lie algebra you'll find online that seem rather un-motivated: (1) why is the Lie bracket the only structure you endow the tangent space with to make a Lie algebra? (2) why are bilinearity, anticommutativity and the Jacobi identity sufficient to characterise a Lie algebra?

The answers, as you may expect, are: (2) all such algebras are Lie algebras to a Lie group, (1) the group (up to connectedness, coverings) is determined by just this information. These are Lie's third and second theorems (up to a special case) respectively:

1. **Lie's first theorem** – This refers to the relatively easy facts that (a) the Lie algebra of a Lie group is a Lie algebra, and two Lie groups have the same Lie algebra iff their connected components have the same universal cover, and (b) a homomorphism of Lie groups induces a homomorphism of their Lie algebras. The next two theorems are essentially the converse of (b) and (a) respectively.

2. **Lie's second theorem** – For a Lie algebra homomorphism $L : TG \to TH$, if G is simply-connected, there is a unique Lie group homomorphism $\ell : G \to H$ such that $L = d\ell$.

3. **Lie's third theorem** – Every Lie algebra is the Lie algebra of a unique simply-connected Lie group.

Together, these form the **Lie correspondence**. They have various consequences of a similar nature, e.g. the correspondence between subgroups and subalgebras, normal subgroups and ideals, centers and centers, abelian Lie groups and abelian Lie algebras, etc.

The first theorem is elementary. The second theorem can be proven quite trivially using the Baker-Cambell-Hausdorff theorem (below). The third theorem is quite hard.

Theorem 1. *Baker-Cambell-Hausdorff theorem (existence): The power series of $\log(e^X e^Y)$ (which is defined on some neighbourhood of the identity) is expressible purely in terms of the nested commutators of X, Y.*

Proof. We reproduce the proof by Eichler (1967) [1]. Group the terms by degree as F_n in the expansion of $\log(e^X e^Y)$, i.e.

$$\mu(X, Y) = \sum_{n=1}^{\infty} F_n(X, Y)$$

Noting that $F_1(X, Y) = X + Y$ and $F_2(X, Y) = \frac{1}{2}[X, Y]$, we will prove by strong induction that each F_n is a Lie polynomial. Suppose that all F_m are Lie polynomials for $m < n$. Then the main "stroke of genius" behind the proof is to nest:

$$(e^X e^Y)e^Z = e^X(e^Y e^Z)$$

$$\sum F_i\left(\sum F_j(X, Y), Z\right) = \sum F_i\left(X, \sum F_j(Y, Z)\right)$$

Looking at terms of degree n only, all terms except $F_n(X, Y) + F_n(X + Y, Z)$ on the left-hand-side and all terms except $F_n(X, Y + Z) + F_n(Y, Z)$ are nestings of F_m's of lower degree and are thus Lie polynomials themselves. Thus the

difference between the aforementioned exceptions is a Lie polynomial, which we write as:

$$F_n(X,Y) + F_n(X+Y,Z) \sim F_n(X,Y+Z) + F_n(Y,Z) \tag{1}$$

And $\sim$ can be checked to be an equivalence relation. Showing from here that $F_n(X,Y) \sim 0$ is like a 15 puzzle of sorts. The first couple of manipulations are at least somewhat predictable – considering the cases $Z = -Y$ and $X = -Y$ respectively yields:

$$F_n(X,Y) \sim -F_n(X+Y,-Y) \tag{2}$$

$$F_n(X,Y) \sim -F_n(-X,X+Y) \tag{3}$$

One can check that applying either condition to itself transitively just yields a trivial result. But we can apply Eq. (2) and Eq. (3) alternatingly to obtain:

$$F_n(X,Y) \sim -F_n(-Y,-X)$$

Since F_n is a iso-degree polynomial in X and Y of degree n, this is the same as:

$$F_n(X,Y) \sim (-1)^n F(Y,X) \tag{4}$$

We may also consider in Eq. (1) the cases $Z = -\frac{1}{2}Y$ and $X = -\frac{1}{2}Y$ (how to motivate this choice – I have no clue – the thing it simplifies to isn't very encouraging either) and similarly apply the results alternatingly to obtain:

$$F_n(X,Y) \sim F_n(\frac{1}{2}X,\frac{1}{2}Y) - F_n(-\frac{1}{2}X,\frac{1}{2}X+\frac{1}{2}Y)$$
$$- F_n(\frac{1}{2}X+Y,-\frac{1}{2}Y) + F_n(\frac{1}{2}X+Y,-\frac{1}{2}X-\frac{1}{2}Y)$$

Simplifying with Eq. (2) and Eq. (3) in third and fourth, and second, terms of the right-hand-side (once again, I cannot motivate this choice) respectively:

$$(1 - 2^{1-n})F_n(X,Y) \sim 2^{-n}F_n(X+Y,Y) - 2^{-n}F_n(Y,X+Y)$$

Simplifying with Eq. (4),

$$(1 - 2^{1-n})F_n(X,Y) \sim 2^{-n}(1 + (-1)^n)F_n(X+Y,Y) \tag{5}$$

Now successively apply Eq. (3) and Eq. (5) to $F_n(X,-Y)$ and iterate the result:

$$F_n(X,-Y) \sim -F_n(X-Y,Y)$$
$$\sim -2^{-n}(1 + (-1)^n)(1 - 2^{1-n})^{-1}F_n(X,Y)$$
$$F_n(X,Y) \sim 2^{-2n}(1 + (-1)^n)^2(1 - 2^{1-n})^{-2}F_n(X,Y)$$

Confirming that this coefficient on the right-hand-side is never equal to 1 for $n > 2$, this implies $F_n(X,Y)$ is Lie. $\qquad\square$

5 Examples

5.1 Special Orthochronous Indefinite Orthogonal Group

Preliminary definitions and facts:

- The indefinite orthogonal group $O(m, n)$ is the group of matrices that preserve the Minkowski form $\text{diag}(m, n)$(diagonal matrix of m 1s, n $-$1s).

- With $O(n)$, simply restricting to $SO(n)$ is enough to eliminate all reflections and produce a connected group, because an even number of reflections compose to a rotation. But with the indefinite orthogonal group, we may not want 180-degree rotations between time and space dimensions in our "restricted" Lorentz group.

- So while we can define $SO(m, n) := \{A \in O(m, n) \mid \det A = 1\}$, it is perhaps more useful to define

$$SO^+(m, n) := \{A \in O(m, n) \mid \det A = 1, \det A_t > 0\}$$

  The challenge is to prove that this is a group.

- There is a tedious "algebraic" proof of this in the case where either $m = 1$ or $n = 1$. But this does not generalise (see [2]).

The key idea comes from the physical insight that we're looking for precisely the transformations that an initially stationary observer can continuously boost or rotate into – i.e. we need to prove that $SO^+(m, n)$ forms the identity component of $O(m, n)$. It is clear (from considering suitable homomorphisms) that it is a union of connected components, and one may use the cos-cosh presentation of the elements to show path-connectedness.

5.2 Parallel parking

Lie theory happens to be useful in control theory, where one is required to study the motion of a system with some given degrees of freedom. An elementary example of such a system is driving – the set of possible motions of the car forms a Lie group. Fig. 2 shows the degrees of freedom of a car: (x, y) is the position of the steering, ϕ is the angle of the car from the horizontal (to determine the position of the back of the car) and θ is the steering angle – so the configuration space of the car is $\mathbb{R}^2 \times \mathbb{T}^2$.

We can study the Lie algebra of the system based on the two permitted motions of the car. The first is infinitesimal steering, given by

$$\text{Steer} = \partial/\partial\theta$$

The second is infinitesimal driving, which moves the steering forward – one can check that this pulls (x, y) forward by $\varepsilon(\sin(\theta + \phi), -\cos(\theta + \phi))$ and increases ϕ by $\varepsilon(-\cos\theta)$, so:

$$\text{Drive} = \sin(\theta + \phi)\frac{\partial}{\partial x} - \cos(\theta + \phi)\frac{\partial}{\partial y} - \cos(\theta)\frac{\partial}{\partial\phi}$$
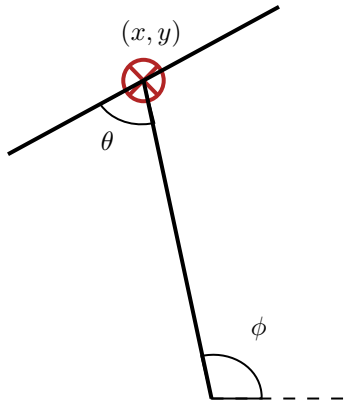
Figure 2: Degrees of freedom of a car

What's really interesting is this:

$$[[\text{Steer}, \text{Drive}], \text{Drive}] = \sin\phi\frac{\partial}{\partial x} - \cos\phi\frac{\partial}{\partial y}$$

Which is precisely the derivation corresponding to a *parallel-slide*. What this means is that the following sequence of tiny motions, each with order ε:

$$\text{Steer, Drive, } \overline{\text{Steer}}, \overline{\text{Drive}}, \text{Drive}, \overline{\text{Steer}}, \overline{\text{Drive}}, \text{Steer, Drive, } \overline{\text{Drive}}$$

(Where an overline represents reverse-direction) Or equivalently:

$$\text{Steer, Drive, } \overline{\text{Steer}}, \overline{\text{Steer}}, \overline{\text{Drive}}, \text{Steer}$$

Produces a parallel-slide of order ε^3. You can confirm with a pencil or something that this is correct.

6   Summary

Here is a summary of some important insights into various Lie-theoretic ideas:

- The notion of a **Lie group** itself – the idea comes from wanting to generalise what we know about discrete groups to more complicated contexts where the "manifold" structure of the group allows us to do so. Examples: **compactness** generalises finiteness, **one-parameter groups** generalise cyclic groups, etc.

- The **exponential map** – For one-parameter groups to generalise cyclic groups, we need a "generalisation" of the group power to allow "real-index powers". The general way to define a **real power** is through the exponential map. Well, this real power stuff isn't *always* defined as it turns

out (you need the exponential map to be surjective), but our motivation does explain why it "makes sense" that the **exponential map is surjective in the connected abelian case** (because then, the Lie algebra is basically a co-ordinate system on the Lie group – I'm aware exponential co-ordinates are defined in more generality, but it's certainly more well-behaved here).

- The **Lie algebra**, i.e. "why is the logarithm/parameter space the tangent space?" We'd like to generalise the notion of a generator to a Lie group – consider e.g. the circle group on the complex plane. An element near the identity generates a cyclic group, and as the element goes nearer to the identity – as it becomes an **infinitesimal generator**, the cyclic group it approaches the entire group. Well, an element close to the identity is of the form $1 + \varepsilon t X$, and generates a group element as $(1 + \varepsilon t X)^{1/\varepsilon} = e^{tX}$. This is also intuition for the compound-interest limit, and for Euler's identity.

- The **Lie bracket** is the second-derivative of the commutator curve $\gamma(t) = e^{tX} e^{tY} e^{-tX} e^{-tY}$. Well, it's also the derivative of $\gamma(\sqrt{t})$, which proves **closure under the Lie bracket**.

- The real justification for the Lie bracket, however, comes from the fundamental fact that $\mathrm{ad} : \mathfrak{g} \to \mathrm{Der}(\mathfrak{g}) := X \to [X, \cdot]$ is the differential of the adjoint map $\mathrm{Ad} : G \to \mathrm{Aut}(G) := g \mapsto \lambda x, gxg^{-1}$, which is a group homomorphism. In particular, the preservation of the Lie Bracket by the differential of a group homomorphism is precisely the **Jacobi identity**: $\mathrm{ad}([x,y]) = [\mathrm{ad}(x), \mathrm{ad}(y)]$. The basic point is that we are trying to reduce Lie group problems to Lie algebra ones as much as possible, and conjugation is an important idea that we'd like to see the map induced by on the Lie algebra – we are seeing the result of the obvious fact that $T\mathrm{Aut}(G) \subseteq \mathrm{Der}(TG)$ (and also $T\mathrm{Aut}(M) = \mathrm{Der}(M)$ – the fact that the automorphisms of an object form a group is equivalent to the derivations on an object forming a Lie algebra). Some more examples of the "study the Lie algebra approach":

  - The uniqueness of the determinant as a map from $G \to \mathbb{R} - \{0\}$.
  - An **ideal** is a subalgebra "induced" on the Lie algebra by a normal subgroup of the Lie group. This immediately provides the interpretation as "kernels of Lie algebra homomorphisms" as well as the condition $[\mathfrak{g}, \mathfrak{i}] \subseteq \mathfrak{i}$.

- The idea behind the manifold-structure of a Lie group is that the flows are produced by left-multiplication by group elements, so those must be homeomorphisms. This motivation can be confirmed through various topological consequences, e.g.

  - **A neighbourhood of the identity generates the connected component.** The idea behind the proof is this: if an entire open neighbourhood of the identity is contained in the subgroup, it means

you can "flow in any direction" from the subgroup – but to bring these flows to an arbitrary point of the manifold, you need left-multiplication to be a homeomorphism.

– **The identity component is a (normal) subgroup.** Because left-multiplication and inversion are continuous, they cannot tear the connected component apart (generalised "intermediate value theorem"), so it is closed under multiplication.

– **Compact Lie groups** – How can a Lie group possibly "close in on itself"? Surely we keep "extending" an open neighbourhood W of the identity by observing that xW must be in the subgroup? The idea is that these translations of W form an **open cover of the group, if it has a finite subcover**, then it makes sense for the group to close in on itself. By playing around with different open neighbourhoods W and taking some suitable unions, one can see that this is equivalent to the condition that every open cover has a finite subcover, i.e. the group is compact.

– **Characterisation of Abelian Lie groups** – "Compact Connected Abelian Lie Group is a torus" is a generalisation of "finite Abelian group is a product of cyclic groups" – the idea is that the exponential map "wraps" the Lie algebra around into the Lie group – this just gives the quotient of the Lie algebra by the kernel of the exponential map, which is topologically $\mathbb{R}^n/\mathbb{Z}^n$. The characterisation of a connected Abelian Lie group as a cylinder $\mathbb{R}^{n+k}/\mathbb{Z}^k$ follows similarly.

• **Lie's fundamental theorems**, consequences of the **BCH theorem**, convey the notion that nestings of the Lie bracket alone are sufficient to determine the local structure of a Lie group (I do not have an intuitive explanation of *why* this is true – of why the BCH theorem makes sense). The "up to **simply connectedness**" condition can be readily understood – it's obvious why the Lie algebra cannot see disconnectedness; the reason it cannot see covering spaces is that the is the quotient of its universal cover by a discrete subgroup, and an application of the fundamental theorem of homomorphisms and the fact that the Lie algebra of a discrete subgroup is trivial implies our result.

• The **Killing form** is the natural way to define an Ad-invariant bilinear form on a Lie algebra, in fact it is unique for **simple Lie algebras**. It allows the interpretation of Ad as a "rotation" of the Lie algebra, as the tangents to its contours are perpendicular to the radial vectors.

Stuff not covered in this text: abstract Lie theory (representations, Ado's theorem, universal enveloping algebras); infinite-dimensional Lie theory; Killing stuff (classification of Lie groups, Cartan's criterion); abstract algebraic things (solvability, nilpotence, Levi decomposition); differential geometry on a Lie group.

References

[1] M. Eichler, *A new proof of the Baker-Cambell-Hausdorff formula*. J. Math. Soc. Japan: 40 (1-2), 1968.

[2] Physics Stack Exchange (Accessed 27-09-2019): *Orthochronous indefinite orthogonal group* $\mathrm{O}^{+}(m,n)$ *forms a group.* `https://physics.stackexchange.com/q/494260`