A Qualitative Introduction to Lie Theory With Applications To Quantum Physics

Monte Mahlum

June 2023

1 Axioms And Definitions[1]

Def: A *Lie Algebra* is a pair $(\mathfrak{g}, [\cdot,\cdot])$ where $\mathfrak{g}$ is a vector space over some field $\mathbb{K}$ and the *Lie Bracket*, $[\cdot,\cdot]$, is a bilinear map from $\mathfrak{g} \times \mathfrak{g}$ to $\mathfrak{g}$ satisfying the following two axioms.

$$\text{I. } [x, x] = 0 \text{ for all } x \in \mathfrak{g} \quad (\text{ anti-commutativity})$$
$$\text{II. } [x, [y, z]] + [z, [x, y]] + [y, [z, x]] = 0 \quad (\text{Jacobi Equation})$$

A first remark is that $\mathfrak{g}$ is often used to denote the pair $(\mathfrak{g}, [\cdot,\cdot])$ and the specific choice of a lowercase $\mathfrak{g}$ hints at the connection between lie algebras and groups which are often denoted as G. Secondly, Axiom I is referred to as anti-commutativity because it immediately implies the notion of anti-commutativity that we are familiar with:

$$0 = [x + y, x + y] = [x, x] + [x, y] + [y, x] + [y, y] = [x, y] + [y, x].$$

Further, $[x, y] = -[y, x]$ implies $[x, x] = 0$, so our theory is independent of whether we choose Axiom I to be the standard notion of anti-commutativity or the one defined above.

While this axiomatization and the abstract idea of a lie algebra may, at first glance, appear quite contrived, a simple investigation shows that the axioms are satisfied if $[x, y]$ is interpreted as the commutation $x \circ y - y \circ x$. Of course, for this interpretation to work, one would need the theory of the $\circ$ operation. On this note, an important result is the following:

Prop: Let A be a vector space over $\mathbb{K}$ and $\circ : A \times A \to A$ be a bilinear, associative map. Then

$$[x, y] := x \circ y - y \circ x$$

defines a lie bracket over A. The proof is easy to check.

Philosophy: It is very natural to think of the lie bracket as a commutator of an associative algebra because this was the initial motivation for the theory[2]. The main mission in physics is to develop more general theories which contains the old theories embedded as special cases (e.g., Newtonian Gravitation $\hookrightarrow$ General Relativity, QM $\hookrightarrow$ QFT). The situation in mathematics is no different which is why we resist the urge to restrict the conversation to commutation. The embedding is made more clear by the following definition:

Def: A *representation* of $\mathfrak{g}$ is a pair (V, ρ) where V is a also a vector space over $\mathbb{K}$ and ρ is a $[\cdot,\cdot]-$homomorphism from $\mathfrak{g}$ to $\mathfrak{gl}(V)$ which is the general linear group over V equipped with the natural commutator. In other words,

$$\rho : \mathfrak{g} \to \mathfrak{gl}(V) \qquad \text{such that}$$
$$\rho([x, y]_{\mathfrak{g}}) = [\rho(x), \rho(y)]_{\mathfrak{gl}(V)} := \rho(x) \circ \rho(y) - \rho(y) \circ \rho(x) \quad \forall x, y \in \mathfrak{g}$$

where $\circ$ here refers to the standard composition of linear endomorphisms of V.

A *sub-representation* of (V, ρ) is a subspace $W \subset V$ which is closed under the action of $\rho(\mathfrak{g})$, an *irreducible* representation is one with no sub-representations (analogous to a transitive $G-$set), and a *faithful* representation is one in which ρ is injective.

With this definition in hand, the reader should note that it is important to detach their perception of a lie algebra from its representations. The actual lie algebra is typically defined abstractly only in terms of its basis $\{e_1, ...\}$ and the so-called *structure constants* f_{ijk} which determine the commutation relations in the following way:

$$[e_i, e_j] = \sum_k f_{ijk} e_k$$

While these constants define the algebra in the most abstract manner, they are intimately tied to a specific representation which we will now define.

Def: Arguably the most important representation, and one that always exists, is the *adjoint* representation $(\mathfrak{g}, \text{ad})$:

$$\text{ad} : \begin{cases} \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g}) \\ x \mapsto \text{ad}_x \end{cases}$$

$$\text{ad}_x : \begin{cases} \mathfrak{g} \to \mathfrak{g} \\ y \mapsto [x, y]_{\mathfrak{g}} \end{cases}$$

Def: An *ideal* of $\mathfrak{g}$ is a sub-algebra $I \subset \mathfrak{g}$ such that $[I, \mathfrak{g}] \subset I$. Equivalently, $I \subset \mathfrak{g}$ is an ideal iff (I, ad) is a sub-representation of the adjoint representation, $(\mathfrak{g}, \text{ad})$.

The quotient vector space $\mathfrak{g}/I$ can be naturally interpreted as a lie algrebra with a lie bracket passed down from $\mathfrak{g}$ only when I is an ideal. More generally, given a sub-representation $W \subset V$, one can discuss the quotient representation where $\mathfrak{g}$ is made to act on V/W.

Philosophy: Given a representation, an interesting question is how to maximally reduce the information encoded in $\rho(\mathfrak{g})$. It turns out that in most relevant cases, there is a decomposition $V = W_1 \oplus ... \oplus W_n$ in which $(W_i)_i$ are irreducible sub-representations. One cannot help but be reminded of the orbit decomposition of a $G-$set. Further, the above remark on quotients should alert the reader to the following analogy:

$$\begin{array}{rcl} \text{closed under adjoint } \mathfrak{g} \text{ action} := \text{ideal} & \longleftrightarrow & \text{lie algebra} \\ & \wr & \\ \text{closed under } \mathfrak{g} \text{ action} := \text{sub-representation} & \longleftrightarrow & \text{representation} \\ & \wr & \\ \text{closed under conjugate } G \text{ action} := \text{normal subgroup} & \longleftrightarrow & \text{group} \\ & \wr & \\ \text{closed under multiplicative action} := \text{ideal} & \longleftrightarrow & \text{ring} \end{array}$$

Remark: Given a representation of $\mathfrak{g}$, (V, ρ),

$$\sigma_V : (x, y) \mapsto \text{tr}(\rho(x)\rho(y))$$

always defines a symmetric, $\mathfrak{g}-$invariant bilinear form. Here, $\mathfrak{g}-$invariance means $\sigma_V([x, y]_{\mathfrak{g}}, z) = \sigma_V(x, [y, z]_{\mathfrak{g}})$ for all $x, y, z \in \mathfrak{g}$. The importance of $\mathfrak{g}-$invariance is its equivalency to the following:

$$\sigma_V(\text{ad}_y(x), z) = -\sigma_V(x, \text{ad}_y(z)) \forall x, y, z \in \mathfrak{g}$$

which tells us that σ_V is invariant with respect to the following symmetry:

$$S : \begin{cases} \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g} \times \mathfrak{g} \\ (x, z) \mapsto (-\text{ad}^{-1}(x), \text{ad}(z)) \end{cases}$$

Def: The *Killing* form, $k : \mathfrak{g} \times \mathfrak{g} \to \mathbb{K}$, is the symmetric, $\mathfrak{g}-$invariant, bilinear form defined by

$$k(x, y) := \text{tr}(\text{ad}_x \circ \text{ad}_y)$$

Variations on the form σ_V, like k, are incredibly useful in both mathematics and physics. we will explore this in more detail in Section 7.

2 A Simple Example (sl_2)

This crucial example of a lie algebra will already be familiar to any student of quantum mechanics. It is the complex vector space of the 2×2 trace-free matrices. Consider the following objects:

$$h := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad e := \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad f := \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

We then define[2]

$$sl_2 := \text{span}_{\mathbb{C}}\{h, e, f\}$$

While the familiarity may not jump out at first glance, a further investigation reveals that sl_2 is simply the span of the Pauli Matrices. Here we have chosen a basis that is tailored to the mathematician, but the decision makes no difference. The commutation relations for sl_2 are presented as follows:

$$[h, e] = 2e, \ \ [h, f] = -2f, \ \ [e, f] = h$$

If we consider the sl_2 as the set of column vectors of the form $\begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{pmatrix} := \lambda_1 h + \lambda_2 e + \lambda_3 f$, then we can readily obtain the adjoint representation.

$$\text{ad}_h = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -2 \end{pmatrix}, \quad \text{ad}_e = \begin{pmatrix} 0 & 0 & 1 \\ -2 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \text{ad}_f = \begin{pmatrix} 0 & -1 & 0 \\ 0 & 0 & 0 \\ 2 & 0 & 0 \end{pmatrix}$$

This gives

$$k(h, e) = 0, \ \ k(h, f) = 0, \ \ k(e, f) = 2$$

While the adjoint representation is incredibly valuable, this example shows that it is certainly not the most compact way to represent the lie algebra.

3 An Immediate Instantiation

Recall from quantum mechanics that for an electron around a nucleus we can diagonalize the Hilbert space $\mathcal{H}$ with respect to the operators H, J^2, J_z, S_z because they all commute. We often present this as[3]

$$\mathcal{H} = \text{span}_{\mathbb{C}}\{|n, l, m_l, m_s\rangle : n \in \mathbb{N}, 0 \le l \le n, -l \le m_l \le l, m_s = \pm 1/2\}$$

Because we already mentioned the connection between the Pauli Matricies and sl_2, the reader can probably recognize that for a fixed n, l, m_l, the vector space

$$V := \{|n, l, m_l, m_s\rangle : m_s = \pm 1/2\}$$

admits a two-dimensional representation of the sl_2.

Let us instead focus on the operators

$$J^2, J_z, J_+, J_- \in \mathfrak{gl}(\mathcal{H}).$$

J^2 measures total orbital angular momentum, J_z measures the projection of this orbital angular momentum onto the z-axis, and $J_\pm$ raise and lower the quantum number m_l.

$$J^2|n, l, m_l, m_s\rangle = \hbar^2 l(l+1)|n, l, m_l, m_s\rangle$$
$$J_z|n, l, m_l, m_s\rangle = \hbar m_l|n, l, m_l, m_s\rangle$$
$$J_\pm|n, l, m_l, m_s\rangle = \lambda_\pm|n, l, m_l \pm 1, m_s\rangle$$

Let us also recall the relevant commutation relations:

$$[J_z, J_\pm] = \pm\hbar J_\pm, \quad [J_+, J_-] = 2\hbar J_z$$

It is clear to see the analogy with the lie algebra sl_2 and if one sets $\hbar := 1$ and fixes the numbers n, l, m_s, it becomes even more apparent. What we have is a representation of sl_2, (V_{n,l,m_s}, ρ), defined as follows

$$V_{n,l,m_s} := \{|n, l, m_l, m_s\rangle : m_l \in \mathbb{Z} \cap [-j, +j]\}$$

$$\rho : \begin{cases} h \mapsto J_z \\ e \mapsto J_+ \\ f \mapsto J_- \end{cases}$$

Further, every finite and *even* dimensional irreducible representation of sl_2 is essentially realized by V_{n,l,m_s} for some l. Let us step back to revel in this idea. The quantum theory described here was postulated to explain the behavior of the electron around a nucleus and it is the beginning of a long story of representation theory in quantum physics. In this light, we see that encoded into the essence of every electron is the space of all even-dimensional irreducible representations of this primary example of a lie algebra, sl_2. The measurements corresponding to the operators H, J^2, S_z can tell you what representation your electron is operating in while a measurement defined by J_z projects this representation onto the z-axis of Spacetime.

Philosophy: If these operators are interpreted as measurement, shouldn't they really live in $\mathcal{H}$ itself (because measurements don't come from anywhere outside the universe)? Well, the measurement corresponding to an operator A is actually the map $\alpha : |\psi\rangle \mapsto \langle\psi|A|\psi\rangle$, so $\alpha \in \mathcal{H}^*$. Thankfully, we have the Reisz Representation Theorem, which tells us that $\alpha(\cdot) \equiv \langle\alpha|\cdot$ for some $\langle\alpha| \in \mathcal{H}$ [4]. Therefore, this idea of measurements living in the space itself is spared. The question is then of the space of operators. Is A a physical object or is it simply a mathematical tool to get to our measuring state $|\alpha\rangle$?

4 Root Space Decomposition

To gain more insight into the structure of a lie algebra, or any mathematical structure for that matter, one can separate it into smaller structures and investigate how the different structures interact. The main way to do this in lie theory is to use root space decomposition, a method largely pioneered by Élie Cartan. For this section, we will assume that $\mathfrak{g}$ is a *semisimple* lie algebra over $\mathbb{C}$.

Def: A *semisimple* lie algebra $\mathfrak{g}$ is one which does not contain a *solvable* ideal. This means that there is no ideal $I \subset \mathfrak{g}$ in which the sequence

$$(I, [I, I], [[I, I], [I, I]], ...)$$

ultimately gives the zero subspace $\{0\} \subset \mathfrak{g}$. Another way to look at this is that there is no nontrivial degenerate ideal $I' \subset \mathfrak{g}$ such that $[I', I'] = 0$.

Def: An element $x \in \mathfrak{g}$ is said to be *semisimple* if ad_x is diagonalizable.

The following is a step-by-step guide on the process of root space decomposition[2].

1. Obtain the *cartan subalgebra* $\mathfrak{h}$. This is the maximal commuting subalgebra consisting of semisimple elements.

2. Find a basis, β, of $\mathfrak{g}$ which is an eigenbasis with respect to the operators $\{\mathrm{ad}_h : h \in \mathfrak{h}\}$. This is possible by the abelian nature of $\mathfrak{h}$.

3. For each basis element $b \in \beta$ find the elements of the dual space $\alpha_b \in \mathfrak{h}^*$ such that

$$\mathrm{ad}_h(b) = \alpha_b(h)b$$

   for all $h \in \mathfrak{h}$.

4. The set $\alpha := \{\alpha_b \neq 0 : b \in \beta\}$ is then your *root system*. Note that $|\alpha| = \dim(\mathfrak{g} \setminus \mathfrak{h})$ because $\alpha_h \equiv 0$ for all $h \in \beta \cap \mathfrak{h}$.

5. Find the *simple roots* $s_\alpha \subset \alpha$ which is the minimal set $s \subset \alpha$ such that

$$\alpha \subset \mathrm{span}_{\mathbb{Z}_{\geq 0}}(s) \cup \mathrm{span}_{\mathbb{Z}_{\leq 0}}(s)$$

   An important result is that s_α is precisely the set of roots that cannot be written as an integer sum of two other roots.

To summarize, s_α forms a sort of basis (we actually call it a *base*) of your root system. The root system consists of linear functionals $\alpha_b \in \mathfrak{h}$ which encodes the action of the cartan subalgebra on the eigenbasis element b. With a little more knowledge on the theory of root systems, one can almost entirely reconstruct the lie algebra $\mathfrak{g}$ from its base s_α. We will not get into these details here, but this is the beginning of the process of classifying all semisimple lie algebras. Further, one can define objects in the root system called *fundamental weights* by their relation to the simple roots. These weights allow us to classify the finite representations of our lie algebra. In this world of weights, we find analogs of the raising and lowering operators in sl_2 which we refer to as *ladder operations*. We have already seen in our angular momentum example that information about the possible representations of a lie algebra can be vital to our understanding of fundamental physics.

5 The Exponential Map[5]

Def: A *lie group* G is a group which is also a differentiable manifold. If $G \subset GL_n(\mathbb{K})$, then we can define its generating lie algebra, $\mathfrak{g}$, as

$$\mathfrak{g} := \{X \in \mathfrak{gl}(\mathbb{K}) : \exp(x) \in G\}$$

where the exponential map is defined by its Taylor series:

$$\exp : X \mapsto \mathbb{1} \sum_{n \in \mathbb{N}} \frac{X^n}{n!}$$

Using basic properties of the exponential, one can check that $\mathfrak{g}$ is indeed a vector space that is closed under commutation.

From a more geometric viewpoint, $\mathfrak{g}$ can be viewed as the tangent space at the identity element of G, $T_{\mathbb{1}}(G)$. To see how this works, begin with the knowledge that $\mathfrak{g}$ is a vector space so $e^{tX} \in G$ for all real t (assuming $\mathbb{K} = \mathbb{R}, \mathbb{C}$). Further, $e^0 = \mathbb{1}$ so e^{tX} is a differentiable path in G starting from identity. Thus, $\frac{d}{dt|_0} e^{tX} = X e^{tX}|_0 = X \in T_{\mathbb{1}}(G)$. Now to ensure a one-to-one correspondence, we must ensure that $\exp(A) \in G$ for all $A \in T_{\mathbb{1}}(G)$. In differential geometry, we define the geodesic from $\mathbb{1}$ in the direction of A to be the unique solution to

$$\frac{d}{dt}\gamma(t) = A\gamma(t), \quad \gamma(0) = \mathbb{1}$$

We know that in differentiable matrix groups, the usual exponential function is the solution to this ODE which gives the equivalency.

6 The Schrodinger Equation[3]

In searching for a guiding differential equation for our quantum state $|\psi_0\rangle$, we begin with the idea that time evolution should be a unitary operation. In other words, $|\psi(t)\rangle = U(t)|\psi_0\rangle$ for some unitary operator U. We set this requirement so that our state remains normalized throughout time: unitary operators are those in which $U^\dagger U = \mathbb{1}$ so $\langle\psi(t)|\psi(t)\rangle = \langle\psi_0|U^\dagger(t)U(t)|\psi_0\rangle = 1$. With the knowledge that unitary matrices make up a differentiable manifold, we can safely assume that $U(t) = \exp(tH')$ for some $H \in \mathfrak{u}$. If we now differentiate with time we get

$$\frac{d}{dt}|\psi(t)\rangle = \frac{d}{dt}\exp(tH')|\psi_0\rangle = H'\exp(tH')|\psi_0\rangle = H'|\psi(t)\rangle$$

From here, one imposes that H' should be proportional to the Hamiltonian operator H (a consequence of Noether's Theorem in classical mechanics), that it should have real eigenvalues, and that these eigenvalues should be discretized by $\hbar$. These requirements then produce the ansatz that $H' = \frac{i}{\hbar}H$.

7 A Brief Introduction To Yang-Mills Theory[6]

It is important to first remark that the author is only beginning his journey in understanding gauge theory and the following should be taken as a very general and potentially incorrect overview.

With that out of the way, consider the following steps which are normally taken to obtain the lagrangian density of a Yang-Mills Theory with gauge group G.

1. Obtain the lie algebra $\mathfrak{g}$ of the gauge group G.
2. Identify the generators T^a of the *fundamental* representation, (V, ρ), of $\mathfrak{g}$ and ensure that they are properly orthonormalized:
$$\sigma_V(T^a, T^b) = \frac{1}{2}\delta^{ab}$$
The reader can take the *fundemental* representation to be defined as the minimal dimensional, faithful representation. In reality, the definition uses the concept "weights" which is a generalization of the root system to arbitrary representations of the algebra.

   From here on out, we will identify $\mathfrak{g}$ with its fundamental representation.
3. Define the gauge potential over Spacetime
$$A_\mu(x) = A^a_\mu(x)T^a$$
This object is very ambiguous and hovers somewhere between the lie algebra and the cotangent space of Spacetime. To gain more insight, observe how it acts on a vector field V^μ.
$$A_\mu V^\mu = A^a_0 T^a V^0 + ... + A^a_3 T^a V^3 \in \mathfrak{g}$$
so essentially, $A_\mu : TM \to \mathfrak{g}$ and this map is defined by the gauge field A^a_μ. The gauge field is really just a field of linear transformations over Spacetime. In this light, the μ^{th} column of the matrix A^a_μ is the vector $A_\mu \in \mathfrak{g}$.
4. Define the field strength
$$F_{\mu\nu} := \partial_\mu A\nu - \partial_\nu A_\mu - i[A_\mu, A_\nu]_{\mathfrak{g}}$$
This is essentially the curvature of the gauge potential.
5. The Yang-Mills Lagrangian density $\mathcal{L}_{YM}$ is then defined as
$$\mathcal{L}_{YM} := \mathrm{tr}(F_{\mu\nu}F^{\mu\nu})$$
The object $F_{\mu\nu}F^{\mu\nu}$ lies inside $\mathfrak{g}$. Denote $T := F_{\mu\nu}F^{\mu\nu}$. If T admits a square root within $\mathfrak{g}$, then we have the following equality:
$$\mathcal{L}_{YM} := \sigma_V(\sqrt{T}, \sqrt{T})$$

References

[1] R. Ó Buachalla. *An Introduction To Lie Algebras and Lie Groups*. Lecture Notes, Charles University, 2023.

[2] K. Erdmann and M.J. Wildon. *An Introduction to Lie Algebras*. Springer Undergraduate Mathematics Series ISSN 1615-2085, 2006.

[3] K. Agarwal. *Honours Quantum Physics I*. Lecture Notes, McGill University, 2022.

[4] W. Rudin. *Functional Analysis*. International Series in Pure and Applied Mathematics. Vol. 8 (Second ed.). 1991.

[5] P. Woit. *Quantum Theory, Groups, and Representations*. Springer, 2017.

[6] D. Tong. *Yang-Mills Theory*. Lecture Notes, University of Cambridge.